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

Mechanistic insights into subgroup III SLFN proteins and the INO80 chromatin remodeling complex

Felix Johannes Metzner
aus
Marburg, Deutschland

2023

Erklärung

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

Versicherung an Eides statt

Diese Dissertation wurde eigenständig und ohne unerlaubte Hilfe erarbeitet.

München, am 09. Mai 2023

____

Felix Johannes Metzner

Dissertation eingereicht am:	11. Mai 2023
1. Gutachter:	Prof. Dr. Karl-Peter Hopfner
2. Gutachterin:	Prof. Dr. Elena Conti
Mündliche Prüfung am:	07. Juli 2023

This thesis has been prepared from May 2018 to May 2023 in the laboratory of Prof. Dr. Karl-Peter Hopfner at the Gene Center of the Ludwig-Maximilians-University in Munich, Germany.

This is a cumulative thesis based on the following publications:

Felix J. Metzner*, Elisabeth Huber*, Karl-Peter Hopfner, Katja Lammens (2022) Structural and biochemical characterization of human Schlafen 5. *Nucleic Acids Research*, Volume 50, Issue 2, pp. 1147-1161.

Felix J. Metzner*, Simon J. Wenzl*, Michael Kugler*, Stefan Krebs, Karl-Peter Hopfner, Katja Lammens (2022) Mechanistic understanding of human SLFN11. *Nature Communications*, Volume 13, Article 5464.

Franziska Kunert*, Felix J. Metzner*, James Jung*, Markus Höpfler, Stephan Woike, Kevin Schall, Dirk Kostrewa, Manuela Moldt, Jia-Xuan Chen, Susanne Bantele, Boris Pfander, Sebastian Eustermann*, Karl-Peter Hopfner (2022) Structural mechanism of extranucleosomal DNA readout by the INO80 complex. *Science Advances*, Volume 8, Issue 49.

* These authors contributed equally.

Table of contents

Summary

The human body has evolved numerous mechanisms to counteract the threat of viruses and other pathogens. The innate immune response to viral infections includes the upregulation of interferons that induce the expression of interferon-stimulated genes (ISGs). Many ISGs serve as restriction factors that collectively establish an antiviral environment and target virtually all steps of viral replication cycles. Several members of the human Schlafen (Slfn) protein family are upregulated in response to interferon signaling and contribute to the inhibition of viral infections. The human subgroup III Slfn proteins SLFN5 and SLFN11 target viral transcription and translation, respectively. Additionally, SLFN5 is associated with cancer cell proliferation and migration and SLFN11 sensitizes cancer cells to DNA-damaging agents. Although the biological functions of SLFN5 and SLFN11 have been extensively studied in a cellular context, little was known about the underlying molecular mechanisms. Structures of both of these Slfn proteins were not available and a comprehensive biochemical characterization of the putative nuclease and helicase domains was largely missing, raising several questions.

What is the overall structural organization of SLFN5 and SLFN11? What is the mode of interaction with nucleic acid substrates and how do the individual domains contribute to it? Do SLFN5 and SLFN11 exhibit tRNase activity and if yes, how is tRNA recognized and cleaved? The aim of the first part of this thesis was the structural and biochemical characterization of SLFN5 and SLFN11 to answer these questions.

Cryo-electron microscopy (cryo-EM) analysis of full-length human SLFN5 and SLFN11 apoenzymes revealed a monomeric SLFN5 state and a dimeric SLFN11 conformation. Biochemical experiments showed that SLFN5 binds ATP, while the SLFN11 ATPase domain adopts an autoinhibited conformation in which the ATP binding site is sterically blocked. However, a phosphomimetic mutation within the SLFN11 helicase domain induces a conformational change, resembling the SLFN5 state, and enabling the SLFN11 mutant to bind ATP. SLFN5 binds various nucleic acid substrates, including double-stranded DNA and tRNA. SLFN11 also binds tRNA and has a high affinity to single-stranded DNA. A structure of SLFN11 bound to single-stranded DNA revealed the DNA binding mode of the helicase domain and suggests a mechanism for SLFN11 recruitment to stalled replication forks. The phosphomimetic mutation within the SLFN11 helicase domain was identified as a conformational switch, regulating dimerization and DNA binding. SLFN11 cleaves tRNA close to the 3' end within the acceptor stem, while SLFN5 shows no tRNase activity. A cryo-EM reconstruction revealed that one tRNA molecule is bound between the Slfn core domains of a SLFN11 dimer and, together with structure-guided mutagenesis studies, suggests a mechanism for the endoribonuclease reaction. Taken together, the data presented here give structural insights into human subgroup III Slfn proteins and, in combination with the biochemical data, offer an important basis for future in cellulo studies.

The second part of this thesis focuses on the characterization of the INO80 chromatin remodeling complex. ATP-dependent chromatin remodelers shape the chromatin landscape by assembling, sliding, editing, or disassembling of nucleosomes and, thereby, control DNA accessibility. This has important implications for the regulation of DNA-dependent processes such as transcription, DNA replication, or DNA repair. The activity of chromatin remodelers relies on a Snf2-type ATPase that uses the energy of ATP hydrolysis to translocate DNA. INO80 is a large multi-subunit remodeler with a modular architecture. While its interaction with the nucleosome core particle has been structurally characterized, a comprehensive structural understanding of the DNA-interaction by the regulatory A-module was missing. This is of particular interest, as the A-module serves as an important allosteric regulator of the nucleosome sliding reaction that acts as molecular ruler and, at least in yeast, reads out shape features of extranucleosomal DNA.

The aim of the second part of the thesis was the structural and biochemical characterization of the INO80 A-module with a focus on its interaction with extranucleosomal DNA.

Cryo-EM analysis of the INO80-nucleosome complex revealed that the A-module binds to approximately 40 bp of extranucleosomal DNA. The helicase/SANT-associated (HSA)/post-HSA domains form a continuous helix that chemomechanically couples the DNA-bound A-module to the motor domain. Structural and biochemical data confirmed that the Ino80 HSA domain and the N-terminus of actin-related protein 8 are critical for DNA binding. The effects of DNA-binding mutants of the A-module on nucleosome sliding support the hypothesis that A-module-DNA interactions are essential for coupling of ATP hydrolysis to productive DNA translocation. Sites of DNA bending and minor groove widening suggest a mechanism of DNA shape readout by the A-module and the motor domain. Finally, cryo-EM analysis of the INO80 A-modules from three different species revealed an evolutionarily conserved A-module anchor that is prevalent in species-specific subunits of INO80 and SWI/SNF family remodelers. Together, this work gives new mechanistic insight into the regulation of INO80-mediated nucleosome remodeling by extranucleosomal DNA.

1. Introduction

1.1. Discovery of the *Schlafen* gene family

The *Schlafen* (*Slfn*) gene family was first described in a study investigating murine thymocyte development [1]. The identified *Slfn* genes showed differential regulation during thymocyte maturation and ectopic expression of Slfn1 was shown to induce cell cycle arrest in thymocytes. This gave rise to the gene family name '*Schlafen'*, which means *'sleep'* in German [1]. Later, a new Slfn subgroup was identified, which harbors a C-terminal domain with similarity to superfamily I DNA/RNA helicases [2]. A subsequent study identified *Slfn* genes in a wide range of mammals, including humans, but also in one species of amphibian (*Xenopus laevis*), fish (*Callorhinchus milii*) and in orthopox viruses [3]. In the human genome, the six identified *SLFN* paralogs (*SLFN5, SLFN11, SLFN12, SLFN12L, SLFN13,* and *SLFN14*) cluster on chromosome 17 [3]. The high degree of conservation within the *Slfn* gene family suggests that they evolved from a common ancestor by duplication or recombination events [3, 4]. Proteins encoded by clustered genes are generally related with respect to function but may be differentially expressed and play divergent roles depending on the biological context [5]. In line with that, Slfn proteins have been linked to a variety of biological processes, such as inhibition of viral replication [6-10], cancer cell proliferation and migration [11-15], and sensitizing of cancer cells to DNA damaging agents [16-21].

1.2. The three Slfn subgroups

The Slfn family members can be divided into three subgroups according to their size and domain composition (Figure 1) [2]. Subgroup I Slfn members consist of a Slfn core domain which is conserved among all Slfn members and has been linked to nucleic acid binding and processing [22-24]. Subgroup II Slfn members contain an additional linker domain, which may serve as interaction site for binding partners [25, 26] and subgroup III Slfn members contain an additional superfamily I DNA/RNA helicase domain [2]. The mouse genome comprises nine *Slfn* genes, two of which are subgroup I Slfn members (*Slfn1, Slfn2*), two are subgroup II Slfn members (*Slfn3, Slfn4*), and five are subgroup III Slfn members (*Slfn5, Slfn8, Slfn9, Slfn10,* and *Slfn14*) [2, 3]. The human genome lacks subgroup I Slfn members and comprises two subgroup II Slfn members (*SLFN12* and *SLFN12L*) and four subgroup III Slfn members (*SLFN5, SLFN11, SLFN13,* and *SLFN14*) [3]. Thus, SLFN5 and SLFN14 are the only human Slfn members with direct orthologs in mice.

1.2.1. Slfn subgroup I and the Slfn core domain

Slfn subgroup I comprises the shortest Slfn proteins, consisting only of a Slfn core domain [1, 2]. The Slfn core domain is highly conserved and can be found in all three subgroups, making it the defining feature of the Slfn family [1, 2, 4]. A particularly conserved region within the Slfn core domain possesses sequence similarity to a domain that is associated with transcriptional regulation in prokaryotes and eukaryotes [2, 27]. This region has been termed Slfn box. Furthermore, the Slfn core domain has been predicted to possess a divergent AAA domain [28, 29]. Typically, AAA domain containing proteins bind and hydrolyze ATP and can be involved in various processes such as membrane trafficking [30] and protein quality control [31].

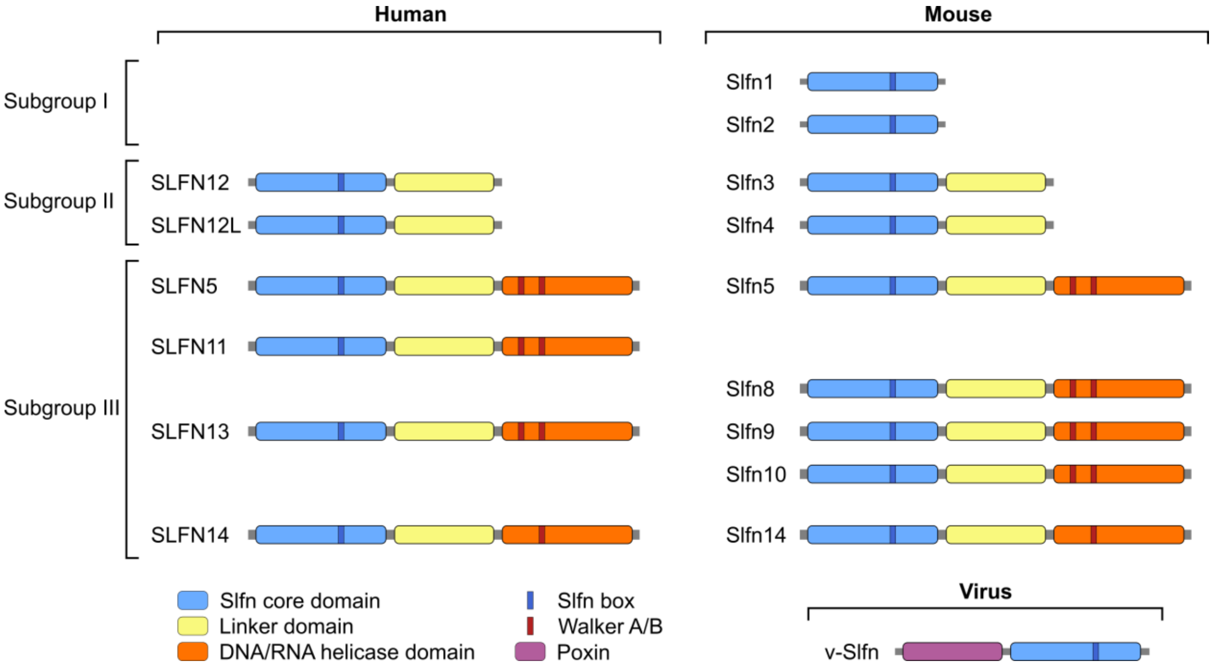


Figure 1: Overview of human, mouse and viral Slfn family members. The Slfn family members can be divided into three subgroups based on protein size and domain composition. Subgroup I Slfn members consist of a Slfn core domain (blue), Subgroup II Slfn members consist of a Slfn core and linker domain (yellow) and Subgroup III Slfn members consist of a Slfn core, linker and DNA/RNA helicase domain (orange). The Slfn box with the putative nuclease active site and the Walker A/B motifs are indicated in dark blue and red, respectively. Viral v-Slfn consists of a poxin domain (purple) and a Slfn core domain. Figure based on [10].

Slfn1 and Slfn2 are preferentially expressed in murine lymphoid tissues and are differentially regulated during thymocyte maturation [1]. Both are expressed in resting T cells and are downregulated upon T cell activation and infection with *Listeria monocytogenes* [1, 2]. Overexpression of Slfn1 in T cells or fibroblasts causes cell cycle arrest [1, 28]. Mechanistically, Slfn1 causes the arrest during G1 phase by inhibiting the induction of cycline D1 [28]. Slfn2 blocks T cell maturation and contributes to maintaining T cells and hematopoietic stem cells in a quiescent state [32-35]. Furthermore, Slfn2 acts as a negative regulator of metastasis and growth of malignant cells [36, 37]. The quiescent state is a resting state, which is characterized by a small cell size, low metabolic activity and resistance to apoptosis [32, 38]. Thus, preventing uncontrolled immune activation. However, upon stimulation with foreign antigens, T cells can leave the quiescent state and start proliferating again [39]. A chemically induced mutation in the Slfn2 core domain was shown to cause inherited immunodeficiency with susceptibility to bacterial and viral infections and decreased numbers of T cells [32, 33]. This so-called *elektra* phenotype is suggested to be a result of a loss of immune cell quiescence [32, 33]. T cell activation is accompanied by changes in the metabolism and an increase in reactive oxygen species (ROS) [40]. The oxidative stress induces certain ribonucleases, such as angiogenin that inhibits translation by tRNA cleavage [41-43]. Slfn2 directly binds tRNAs and protects them from angiogenin-mediated cleavage, thus facilitating protein synthesis in activated T cells [24]. Taken together, Slfn1 and Slfn2 contribute to maintaining the quiescent state of T cells in mice [32, 44].

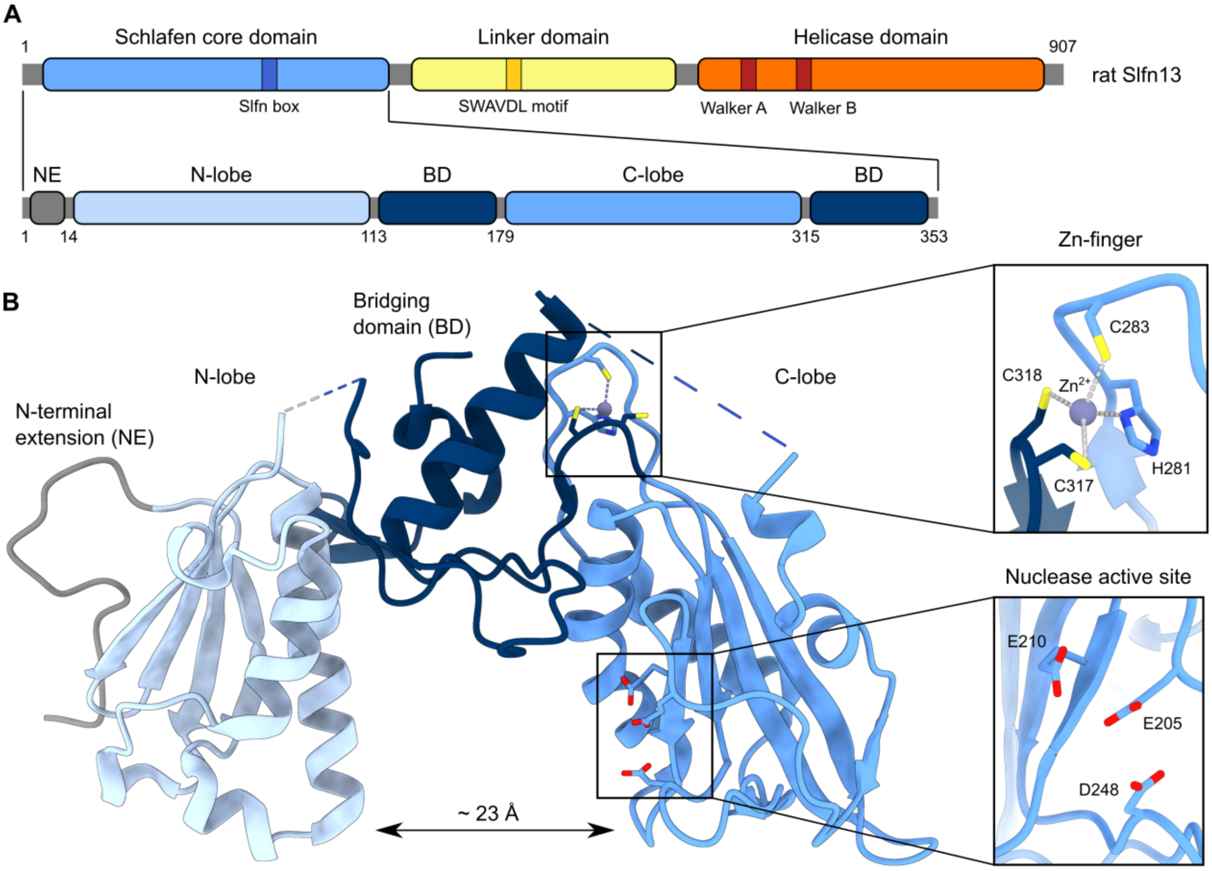


Figure 2: Structure of the rat Slfn13 core domain. (A) Domain overview of full-length rat Slfn13 (top). The Slfn core domain (blue), linker domain (yellow) and helicase domain (orange) are color-coded. The Slfn box, SWAVDL motif and Walker A/B motifs are highlighted. Detailed domain overview of the rat Slfn13 core domain (bottom). The N-terminal extension (NE), Slfn core N- and C-lobes and bridging domains (BD) are color-coded. (B) Structure of rat Slfn13$^{14-353}$ (PDB: 5YD0) [22]. Domains are indicated and colored as in (A). Residues of the Zn-finger motif and the putative nuclease active site are shown. Detailed views of the Zn-finger motif and the putative nuclease active site are shown as inserts.

Structural information on subgroup I Slfn proteins is lacking, though, structures of the Slfn core domains of subgroup II and III proteins are available. Yang et al. [22] reported the crystal structure of the Slfn core domain of rat Slfn13 (rSlfn13$^{14-353}$) (Figure 2). Furthermore, structures of the SLFN12-PDE3A complex provide insights into the SLFN12 Slfn core domain [25, 26]. These studies illustrate that the Slfn core domain is structurally conserved and adopts a horseshoe-like shape (Figure 2B). The domain consists of an N- and C-lobe that are connected by two bridging domains. In the case of rSlfn13$^{14-353}$, the arrangement can be described as pseudo symmetric [22]. The C-lobe harbors a highly conserved zinc finger motif that consists of three Cys and one His residue coordinating a zinc ion (Figure 2B). The N- and C-lobes of the Slfn core domain form a channel, with positively charged patches on each lobe (Figure 3). The inner channel of rSlfn13$^{14-353}$ measures approximately 23 Å, which could fit base-paired DNA or RNA and has been suggested to be a binding platform for nucleic acids [22]. In line with this, rSlfn13$^{14-353}$ cleaves tRNA and rRNA in an Mg^{2+}/Mn^{2+} dependent manner [22]. A conserved three-carboxylate triad (E205, E210 and D248) is essential for the endonuclease activity of rSlfn13 (Figure 2B) [22]. Furthermore, RNase activity has been shown for mouse Slfn8, SLFN11, SLFN12, and SLFN14 [22, 23, 26, 45, 46]. While Slfn2 counteracts translational inhibition by the protection of tRNAs from nucleolytic cleavage [24], SLFN11 and SLFN12 cleave certain tRNAs and inhibit translation in a codon-usage-dependent manner [6, 45, 46].

Taken together, several Slfn members play a role in translational regulation via their Slfn core domain that is involved in RNA binding and processing.

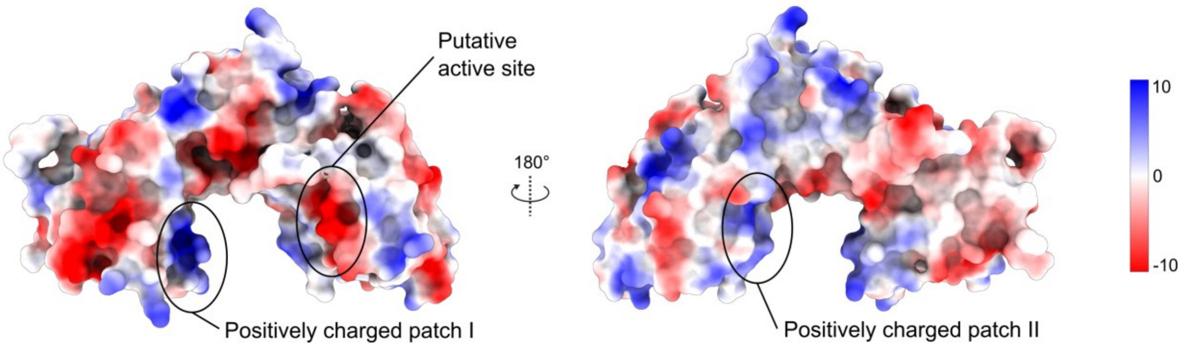


Figure 3: Electrostatic surface potential of rat Slfn13 core domain. Front and back view of the surface representation of the rat Slfn13 core domain (PDB: 5YD0) [22]. The surface is colored according to its electrostatic potential. The values of the color-bar are in units of kcal mol^{-1} e^{-1} at 298 K. The positively charged patches as well as the negatively charged putative active site are indicated.

1.2.2. Slfn subgroup II and the linker domain

Two subgroup II Slfn members are encoded in the mouse (*Slfn3, Slfn4*) and human genome (*SLFN12*, *SLFN12L*), respectively (Figure 1) [1, 2]. In addition to the Slfn core domain, subgroup II Slfn members harbor a so-called linker domain (Figure 4A). This domain has also been referred to as SWADL or SWAVDL domain due to a conserved S-W-(A/S)-(V/G/L)-D-(L/I/V) motif [27, 47].

Similar to mouse subgroup II Slfn members, SLFN12 is predominantly localized in the cytoplasm [27, 48]. While the native biological functions of subgroup II Slfn members are largely uncharacterized, SLFN12 came into focus as a factor that sensitizes cancer cells to a subset of phosphodiesterase 3A (PDE3A) modulators [49]. In a phenotypic screen for new cancer drugs, the cytotoxicity of several PDE3A inhibitors correlated with the expression of PDE3A, a phosphodiesterase that hydrolyses the second messenger 3',5'-cyclic adenosine monophosphate (cAMP) [49, 50]. The identified inhibitors, such as the small molecule DNMDP, promote a physical interaction between PDE3A and SLFN12 [49]. The cytotoxic effects of DNMDP do not solely rely on PDE3A expression but correlate with both PDE3A and SLFN12 expression, while knockdown of PDE3A or SLFN12 reduces sensitivity to DNMDP [49]. In addition to synthetic PDE3A inhibitors, the endogenous human steroid hormone 17-β-estradiol can induce apoptosis by binding to PDE3A and recruitment of SLFN12 [51]. The cytotoxic effect of the identified small molecules does not rely on the inhibition of PDE3A phosphodiesterase activity but rather induces a novel activity in complex with SLFN12, as SLFN12 knockdown is sufficient to abrogate cytotoxicity of certain PDE3A inhibitors [52, 53]. Furthermore, the natural product nauclefine induces apoptosis via the PDE3A-SLFN12 pathway without inhibiting PDE3A phosphodiesterase activity [54].

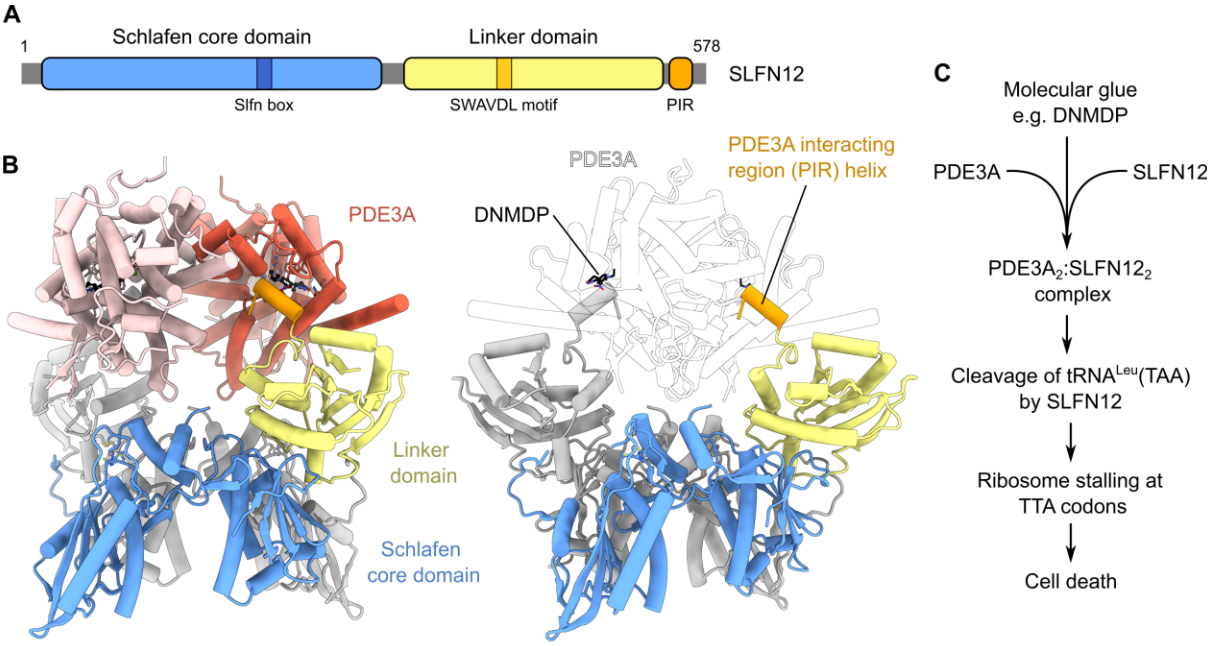


Figure 4: Structure of the PDE3A-SLFN12 complex. (A) Domain overview of full-length SLFN12. The Slfn core domain (blue), linker domain (yellow) and PDE3A interacting region (PIR; orange) are color-coded. The Slfn box and SWAVDL motif are highlighted. (B) Structure of the PDE3A^{CAT}-SLFN12-DNMDP heterotetramer (PDB: 7LRD) [26]. One SLFN12 protomer is colored as in (A) and the second protomer is depicted in grey. The PDE3A protomers are depicted in red and pink (left) or transparent (right). The PIR helix is indicated and DNMDP is shown in black. (C) Overview of the PDE3A-SLFN12 pathway in response to DNMDP treatment. Molecular glues, such as DNMDP, induce complex formation between PDE3A and SLFN12. Complex formation stimulates the SLFN12 nuclease activity that cleaves tRNA-Leu-TAA. This causes ribosome stalling at TTA codons and eventually leads to cell death. Figure based on [46].

Insights into the mechanism of action of DNMDP as well as other PDE3A modulators could be gained by structural analysis of the DNMDP bound PDE3A-SLFN12 complex [25, 26]. The catalytic domain of PDE3A and SLFN12 form a C2-symmetric heterotetrameric complex consisting of two subunits each (Figure 4B). DNMDP binds a pocket near the PDE3A active site and stabilizes an interaction between PDE3A and the C-terminal α-helix of SLFN12. Thus, DNMDP acts as a molecular glue, which are small molecules that mediate the interaction of otherwise non-interacting proteins [25]. The C-terminal α-helix of SLFN12 that interacts with PDE3A has been termed PDE3A interacting region (PIR) and is essential for complex formation and the cytotoxic activity of DNMDP [26]. The two PDE3A protomers interact via a large interface while the interaction between PDE3A and SLFN12 is largely restricted to the PIR helix (Figure 4B) [25, 26]. The two SLFN12 protomers interact via a large interface between the Slfn core domains in a head-to-tail orientation. The linker domain is a globular domain that interacts with the N-lobe of the Slfn core domain as well as with the PDE3A catalytic domain. The structure of the SLFN12 core domain is largely similar to the rat Slfn13 core domain and harbors the conserved three-carboxylate triad and the zinc finger [22, 25, 26]. The two Slfn core domains form a positively charged channel with a diameter of approximately 20 Å [25, 26]. SLFN12 has RNase activity, which is stimulated by DNMDP-induced SLFN12-PDE3A complex formation [26]. Mutation of the conserved three-carboxylate triad strongly reduces RNase activity and renders cells chemoresistant to DNMDP treatment, showing that RNase activity is required for DNMDP-mediated cancer cell killing [26]. SLFN12 RNase activity is regulated by multiple phosphorylation sites within the linker domain [55]. Upon complex formation with PDE3A, SLFN12 is dephosphorylated, which promotes its RNase activity and cytotoxic function [55]. A recent report could identify tRNA-Leu-TAA as the physiological substrate of SLFN12 [46]. Molecular glue induced cleavage of tRNA-Leu-TAA by SLFN12 causes ribosome pausing at Leu-TTA codons, resulting in translational inhibition of proteins encoded by UUA-rich mRNAs. As many factors that are involved in mRNA translation and protein homeostasis fall into this category, it has been speculated that SLFN12-mediated cancer cell killing is a result of global translational inhibition (Figure 4C) [46].

1.2.3. Slfn subgroup III and the helicase domain

Slfn subgroup III represents the largest Slfn subgroup in mouse (*Slfn5, Slfn8*, *Slfn9*, *Slfn10*, and *Slfn14*) and human (*SLFN5, SLFN11*, *SLFN13*, and *SLFN14*) (Figure 1) [2, 3]. Subgroup III Slfn proteins were first discovered in mice and shown to be differentially regulated in hematopoietic cells, e.g. during T cell activation, suggesting a cell type specific role in cell development and growth [2]. In addition to the Slfn core domain and the linker domain, subgroup III members harbor a C-terminal domain with similarity to superfamily I DNA/RNA helicases [2]. Based on sequence alignment, the Slfn helicase domain shows the signature motifs of P-loop ATPases that consist of two RecA-like domains and harbors the conserved Walker A and B motifs that are required for ATP binding and hydrolysis [2]. Due to their functional importance, Walker A and B motifs are highly conserved and consist of the consensus sequences GxxxxGK(T/S) (x: any amino acid) and hhhhD(D/E) (h: hydrophobic amino acid), respectively [56]. In contrast to subgroup I and II Slfn proteins, most subgroup III members have a nuclear localization signal (NLS) within the helicase domain [27].

SLFN5 and SLFN11 are the best characterized subgroup III members and will be discussed in the later sections in more detail (Sections 1.4 to 1.6). In brief, SLFN5 functions as a restriction factor for certain viruses by targeting viral transcription (Section 1.4.2) [9, 57]. Furthermore, it plays a role in transcriptional regulation in different cancers and can have an inhibitory [11, 58-63] or stimulatory [14, 64-68] effect on tumorigeneses (Section 1.5). SLFN11 also possesses antiviral activity but targets translation of viral proteins in a codon-usage-dependent manner (Section 1.4.1) [6, 69]. This function depends on the SLFN11 core domain, but at least in the case of some viruses, also on the helicase domain [70]. SLFN11 was further shown to irreversibly block DNA replication in response to treatment with certain DNA damaging agents (Section 1.6) [71, 72]. Mutagenesis studies of the Walker motifs could show that the ATPase activity of the helicase domain is essential for the replication block [71, 72].

In contrast to SLFN5 and SLFN11, SLFN13 lacks an NLS and is located in the cytoplasm [22]. It can cleave tRNA and rRNA in vitro which requires a conserved three-carboxylate triad within the core domain [22]. The expression of SLFN13 is stimulated by human immunodeficiency virus 1 (HIV-1) infection and SLFN13 restricts HIV-1 in a nuclease-dependent manner [22]. SLFN13 inhibits the expression of viral proteins but also disrupts translation on a global scale [22]. While the SLFN13 core domain alone shows antiviral activity, the effect is stronger in the case of the full-length protein, indicating that the linker and/or helicase domain might contribute to its antiviral activity [22].

Similar to other subgroup III Slfn members, SLFN14 is mostly localized in the nucleus and shows antiviral activity against certain RNA and DNA viruses [8]. In a study on rabbit reticulocytes, a type of immature red blood cells that do not have a nucleus, a C-terminally truncated version of SLFN14 was found to be strongly overexpressed and bound to ribosomes [23]. SLFN14 interacts with ribosomes via its linker domain and cleaves rRNA and ribosome-associated mRNA, resulting in degradation of ribosomal subunits [23]. The endoribonuclease reaction is Mg^{2+}- and Mn^{2+}-dependent and independent of ATP [23]. However, in nucleated cell types, only full-length SLFN14 was detected and located in the nucleus [23]. Several point mutations in the C-lobe of the SLFN14 core domain (K218E, K219N, and V220D) have been associated with inherited thrombocytopenia, a disease characterized by low platelet counts and an increased risk of bleeding [73]. In vitro studies suggest that these SLFN14 mutants still bind to ribosomes and degrade rRNA, but show significantly reduced protein levels due to post-translational degradation of the partially misfolded SLFN14 versions [74]. In vivo experiments further support the finding that SLFN14 acts as a regulator in mammalian hematopoiesis and is involved in platelet and erythroid lineage commitment in a species-specific manner [75].

Taken together, subgroup III Slfn members share a conserved three-partite domain composition and many act as antiviral restriction factors. However, the current understanding of their biological functions is only fragmentary and seems to be rather diverse.

1.3. *Slfn* family belongs to interferon-stimulated genes

Cytokines are a versatile class of small proteins that are involved in cell signaling and act through cell surface receptors [76]. Interferons (IFNs) are a subset of cytokines that are released by host cells in response to infections [77]. Accordingly, the name interferon is based on their ability to “interfere” with viral infections [78]. As part of the innate immune response, IFNs play a role in the first line of defense against viral but also bacterial or fungal infections [77]. In humans, IFNs can be classified into three distinct families [79]. Type I IFNs represent the largest family with IFNα and IFNβ being the most broadly expressed members [79]. All type I IFN family members bind the type I IFN receptor (IFNAR) [79]. IFNγ is the only type II IFN family member and is recognized by type II IFN receptor (IFNGR) [80]. The type III IFN family consist of IFNλ1-4 and binds to a cell surface receptor which is mostly expressed on epithelial cells [81].

IFNα/β can be produced by most cell types of the body in response to the stimulation of pattern recognition receptors (PRRs) (Figure 5) [82, 83]. Examples for PRRs are retinoic acid-inducible gene I (RIG-I) that senses viral RNA [84] or cyclic GMP-AMP synthase (cGAS) that senses cytosolic dsDNA [85]. Upon stimulation of PRRs, diverse signaling pathways lead to the activation of transcription factors such as the IFN-regulatory factors (IRFs) that activate the transcription of type I IFNs (Figure 5) [77, 86]. The type I INFs are released from the infected host cells and stimulate the Janus kinase/signal transducer and activator of transcription (JAK-STAT) pathway by binding to the cell surface receptor IFNAR [77, 86]. IFNAR is a heterodimeric transmembrane receptor consisting of the subunits IFNAR1 and IFNAR2 [87]. IFN binding induces dimerization of the receptor subunits which activates the receptor-associated tyrosine kinases Janus kinase 1 (JAK1) and tyrosine kinase 2 (TYK2) [88]. Activation of the kinases results in autophosphorylation of the kinases as well as phosphorylation of IFNAR1/2 [88]. This permits the recruitment of signal transducer and activator of transcription 1 (STAT1) and STAT2 that are phosphorylated, leading to dimerization [89, 90]. The phosphorylated STAT1/2 dimer binds the IFN-regulatory factor 9 (IRF9), forming the IFN-stimulated-gene factor 3 (ISGF3) complex [86]. ISGF3 translocates into the nucleus and binds to IFN-stimulated response elements (ISREs) in promoters of IFN-stimulated genes (ISG), which leads to the transcription of ISGs (Figure 5) [86]. This way, IFNα/β signaling induces the transcription of a variety of ISGs, which collectively create an antiviral state within the cell [91].

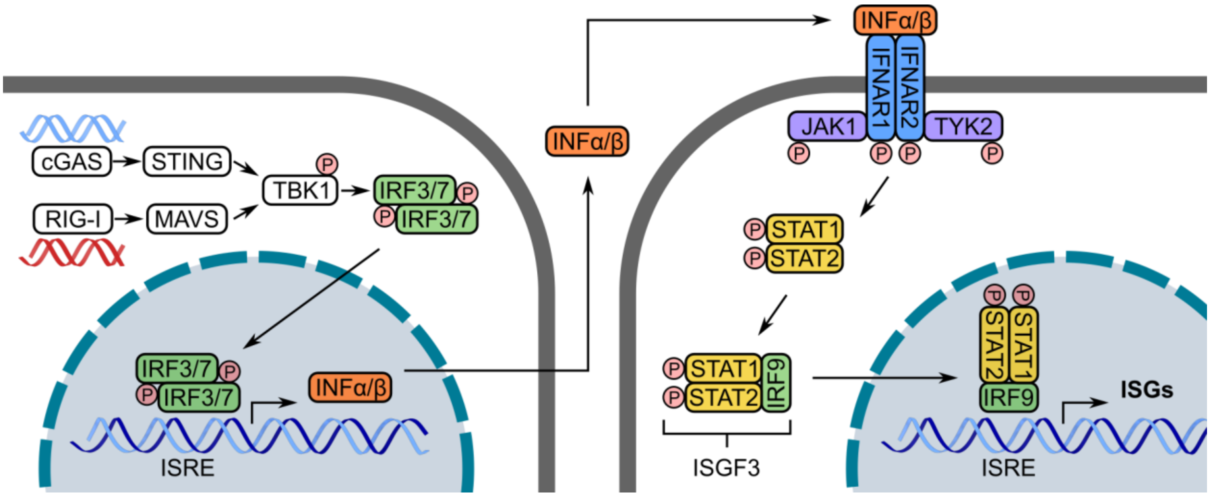


Figure 5: Schematic overview of type I interferon induction and signaling. Pattern recognition receptors such as cGAS or RIG-I detect cytosolic nucleic acids. This triggers signaling pathways via the stimulator of interferon genes protein (STING) or mitochondrial antiviral-signaling protein (MAVS) and the TANK-binding kinase 1 (TBK1), which phosphorylates IFN regulatory factor (IRF)3 and/or IRF7. Phosphorylation of the IRF transcription factors leads to homodimerization, nuclear translocation and induction of type I IFN expression (IFNα/β). The type I IFNs are released from the cell and act on adjacent cells via binding to the heterodimeric type I IFN receptor (IFNAR)1/2. This activates the receptor-associated JAK1 and TYK2 kinases, resulting in autophosphorylation of the kinases and phosphorylation of INFAR1/2. This permits the recruitment of STAT1 and STAT2, that heterodimerize upon phosphorylation and bind IRF9 to form the ISGF3 complex. The transcription factor complex enters the nucleus and induces the transcription of ISGs by binding to ISRE. Figure based on [92].

The class of ISGs comprises PRRs, IRFs, and STAT1/2 that are involved in pathogen detection and IFN signaling but also factors that negatively regulate IFN signaling, for example by inhibiting the JAK-STAT pathway [93]. Negative regulators are essential for IFN-desensitization following IFN signaling, as dysregulated IFN production or signaling may cause autoimmune disorders [93]. However, a large number of ISGs are restriction factors with direct antiviral activity, which evolved to target almost any step of the viral life cycle [93].

Several mouse and human Slfn family members are inducible by type I IFNα, indicating that the Slfn family belongs to the class of ISGs [94]. In mice, treatment with IFNα induces transcription of subgroup I (*Slfn1* and *Slfn2*), subgroup II (*Slfn3*) and subgroup III (*Slfn5* and *Slfn8*) Slfn members in melanoma and renal cell carcinoma cells [37]. In human melanocytes, IFNα treatment induces the gene expression of *SLFN5*, *SLFN11*, *SLFN12* and *SLFN13* [58]. Furthermore, stimulation of monocyte-derived dendritic cells with human rhinovirus, lipopolysaccharide or IFNα induces the gene expression of *SLFN5*, *SLFN11*, *SLFN12*, *SLFN12L* and *SLFN13*, showing that the upregulation is type I interferon dependent [95]. The degree of upregulation of Slfn family members showed similar kinetics but was less pronounced compared to the upregulation of a classical ISG (*MxA*), possibly due to different numbers of ISRE sites [10, 95]. However, basal *SLFN* expression levels differ between different cell types and the strength of induction is also cell type dependent [58, 95].

SLFN5 gained special attention as it not only functions as an ISG but also acts as a negative regulator of IFN signaling [64]. While stimulation with IFNα induces the expression of *SLFN5*, it also promotes the interaction of SLFN5 with STAT1. The resulting complex was suggested to bind to ISREs, leading to the repression of ISG transcription. In line with that, SLFN5 knockout cells showed increased ISG expression compared to wild type cells [64]. Together, this points towards a role of SLFN5 in a negative-feedback loop [64].

1.4. Antiviral activity of Slfn proteins

Many ISGs function as antiviral restriction factors that have evolved to target almost any step of the viral life cycle [93]. HIV-1 for example, is a single-stranded enveloped RNA retrovirus that primarily infects CD4$^+$ T cells and macrophages [96]. When a HIV virion encounters a CD4 expressing cell, it binds the CD4 receptor via its spike protein (Figure 6). Engagement of the co-receptor CC-chemokine receptor 5 (CCR5) leads to fusion of the viral and host cell membranes. The viral capsid that contains the single-stranded RNA genome as well as viral enzymes such as reverse transcriptase, proteases and integrase is released into the cell. The RNA is reverse transcribed into double-stranded DNA that is imported into the nucleus. Integration of the DNA into the host genome by integrases forms the so-called provirus. Transcription of the provirus yields viral mRNA that is exported into the cytosol. Translation of the mRNA yields viral proteins that, together with viral RNA, assemble into viral particles. After budding and release of the viral particles, protease-mediated maturation leads to the formation of infectious viral particles [96].

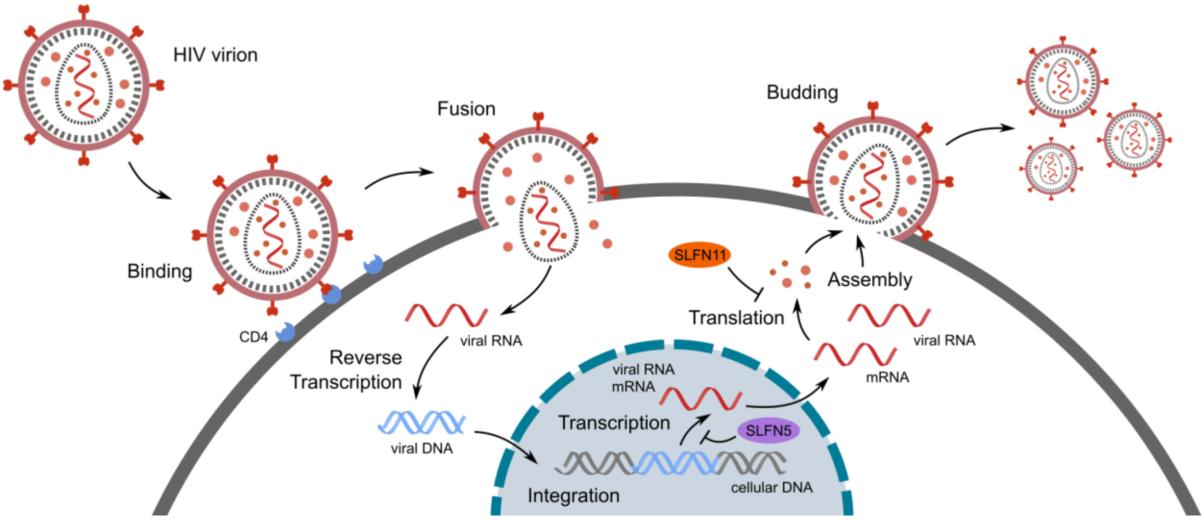


Figure 6: Schematic overview of the HIV-1 replication cycle. The virus binds the CD4 receptor and co-receptors of target cells via its spike protein, leading to the fusion of the viral and host cell membranes. Viral RNA and enzymes (e.g. reverse transcriptase, integrase) are released into the target cell and the viral RNA is reverse transcribed into viral dsDNA. After nuclear import, the viral DNA is integrated into the host genome. Transcription of viral DNA yields viral mRNA which is exported into the cytoplasm. Translation of viral mRNA yields viral proteins that, together with viral RNA, assemble into viral particles. After budding and release of the viral particles, protease-mediated maturation leads to the formation of infectious viral particles. Figure based on [96]. SLFN5 (purple) inhibits HIV-1 replication by suppressing the transcription while SLFN11 (orange) inhibits the translation of viral proteins in a codon-usage-dependent manner [6, 57].

HIV-1 restriction factors target for example the fusion of the virus with the host cell, reverse transcription, nuclear import, viral assembly or budding of the viral particles [97]. The ISG and restriction factor sterile alpha motif and histidine/aspartic acid (HD) domain containing protein 1 (SAMHD1) for example, is a triphosphohydrolase that converts deoxynucleoside triphosphates (dNTPs) to deoxynucleosides [98]. This leads to a depletion of the intracellular dNTP pool, which ultimately inhibits reverse transcription of viral RNA. Another restriction factor is tetherin, that restricts HIV-1 replication at a much later stage of the viral life cycle by tethering the budding virion to the cell surface and inhibiting the release of the viral particle [99].

Due to stimulated expression of Slfn proteins in response to IFNs, viral infections or pathogen-associated molecular patterns (PAMPs), Slfns have been proposed to play a role in antiviral defense

[95, 100, 101]. Indeed, several Slfn members, especially those belonging to subgroup III, were found to restrict viral infections by multiple different mechanisms [10].

1.4.1. SLFN11 is a restriction factor that targets translation

SLFN11 was the first human SLFN family member, discovered to have antiviral activity [6]. In 2012, Li et al. [6] reported that SLFN11 inhibits the replication of HIV-1 in a codon-usage-dependent manner. This is in line with the observation that enhanced SLFN11 expression in CD4$^+$ T cells correlates with reduced HIV-1 RNA levels during antiretroviral therapy [102]. Individuals that naturally suppress HIV-1 to low levels, so called HIV-1 "elite controllers", show elevated SLFN11 expression levels in CD4$^+$ T cells [7].

A systematic analysis of the influence of SLFN11 on the HIV-1 replication cycle revealed that SLFN11 does not inhibit the early steps such as reverse transcription of viral RNA, integration of viral DNA or transcription of the provirus [6]. It also has no inhibitory effect on the budding of viral particles. Instead, SLFN11 interferes with retroviral replication at a much later stage, by inhibiting the expression of viral proteins (Figure 6 and 7) [6]. The genomes of many viruses, such as HIV-1, have a different codon usage compared to the human genome [103, 104]. This viral codon bias is characterized by a lower G/C content compared to the host cell codon usage [6, 105, 106]. The HIV-1 genome shows a particularly high frequency of A/U in the third position of the codon which relates to rare codons [6, 105, 106]. Usage of rare codons can cause ribosomal pausing and is associated with inefficient translation [107]. While SLFN11 inhibits the translation of the codon biased viral group-specific antigen (gag), it has no inhibitory effect on the expression of an artificially codon-optimized gag protein [6]. The antiviral activity of SLFN11 is dependent on the proteins N-terminal domain, which binds tRNAs [6]. The SLFN11 activity manifests in the counteraction of virus-induced changes of the tRNA pool in HIV-1 infected cells [6, 108].

However, codon-usage-dependent inhibition of protein expression is not limited to viral proteins, as expression of human proteins with a disproportionately high frequency of certain non-optimal codons is also inhibited [45]. The SLFN11-dependent translational inhibition of these host proteins is the result of cleavage of distinct type II tRNAs upon treatment of tumor cells with DNA damaging agents (DDAs) [45]. Type II tRNAs include Ser and Leu tRNAs and differ from type I tRNAs by the presence of an extended variable loop [109]. While multiple type II tRNAs are cleaved by SLFN11, cleavage of tRNA-Leu-TAA seems to be responsible for the translational inhibition [45]. The depletion of this already rare tRNA by SLFN11 is speculated to become a rate-limiting factor during translation, causing translational inhibition of TTA codon rich mRNAs [45]. In fact, SLFN11 is active in the absence of viral infections and might thereby establish a general antiviral state within the cell [6, 45, 110].

A study on human and non-human primate SLFN11 could show that it evolved under recurrent positive selection in primates [110]. Its antiviral activity towards retroviruses and the ability to inhibit translation of non-codon optimized transcripts differs between species and is more potent in non-human primates compared to humans [110]. Besides HIV-1, prototype foamy virus (PFV) and equine infectious anemia virus (EIAV), which both belong to the family of retroviruses, are inhibited by SLFN11 in a codon-usage-dependent manner [69, 70]. In the case of PFV, residues of the proposed nuclease active site within the SLFN11 core domain as well as a functional Walker B motif within the helicase domain are essential for the antiviral activity [70]. Furthermore, dephosphorylation of three residues

(S219, T230, S753), that were shown to regulate SLFN11 activity in the DNA damage response, is essential for its antiviral activity [111].

It was speculated, that positive-sense ssRNA viruses are more susceptible to translational inhibition by SLFN11 compared to negative-sense ssRNA viruses, as they rely on immediate translation of their RNA genome in order to replicate [108]. Indeed, positive-sense flaviviruses such as West Nile virus (WNV), dengue virus (DENV) and Zika virus (ZIKV) are sensitive to SLFN11, while replication of several negative-sense ssRNA viruses is unaffected [108]. The extend by which the replication of flaviviruses is affected by SLFN11 differs between viruses, suggesting that some viruses evolved a mechanism to counteract SLFN11 [10, 108].

A viral antagonistic mechanism was recently described in a human cytomegalovirus (HCMV) infection setting [112]. SLFN11 potently restricts HCMV infections, showing that it not only targets RNA viruses but also DNA viruses. However, HCMV evolved a mechanism to counteract the restriction factor SLFN11. The viral protein RL1 targets SLFN11 for proteasomal degradation via recruitment of the Cullin4-RING E3 Ubiquitin Ligase (CRL4) complex (Figure 7) [112]. Thus, targeting of SLFN11 antagonizing mechanisms might offer therapeutic potential.

1.4.2. SLFN5 is a restriction factor that targets transcription

SLFN5 expression is upregulated by IFNα signaling and rhinovirus infections [36, 95]. Similar to SLFN11, SLFN5 has been described as a restriction factor, targeting certain RNA and DNA viruses [9, 57]. However, SLFN5 and SLFN11 seem to employ different modes of action.

Overexpression of SLFN5 suppresses HIV-1 replication, while depletion of endogenous SLFN5 leads to an increase in viral replication [57]. The antiviral activity of SLFN5 depends on its nuclear localization, as disruption of its nuclear localization signal leads to a loss of function [57]. The antiviral activity resides in the N-terminal domain, as a SLFN5 construct that is lacking the helicase domain is still able to restrict HIV-1 replication [57]. SLFN5 binds to specific sites within the promoter region of the proviral DNA, which represses recruitment of RNA polymerase II to transcription initiation sites [57]. SLFN5 recruits the histone methyltransferase polycomb repressive complex 2 (PRC2) to the viral promoter [57]. This promotes deposition of the repressive epigenetic marks H3K27me2 and H3K27me3, leading to chromatin compaction and inhibition of proviral transcription (Figure 6 and 7). However, the inhibitory effect of SLFN5 on HIV replication differs between different HIV strains and is not observed in HIV-2 [57].

SLFN5 also functions as a restriction factor for the DNA virus herpes simplex virus (HSV-1) [9]. By binding to viral DNA, SLFN5 has been proposed to limit accessibility of RNA polymerase II to viral promoters [9]. This leads to a direct inhibition of transcription of the viral genome. In line with its function in the HIV-1 setting, the antiviral activity of SLFN5 does not require a functional ATPase [9].

HSV-1 has evolved a potent antagonistic mechanism to overcome SLFN5-mediated restriction [9]. The viral E3 ubiquitin-protein ligase ICP0 mediates the ubiquitination of SLFN5, targeting it for proteasomal degradation (Figure 7). Thus, the antiviral effect of SLFN5 is only prominent in an ICP0 knockout HSV-1 strain, as the antagonistic mechanism compensates the restriction factor activity in wild type HSV-1 [9].

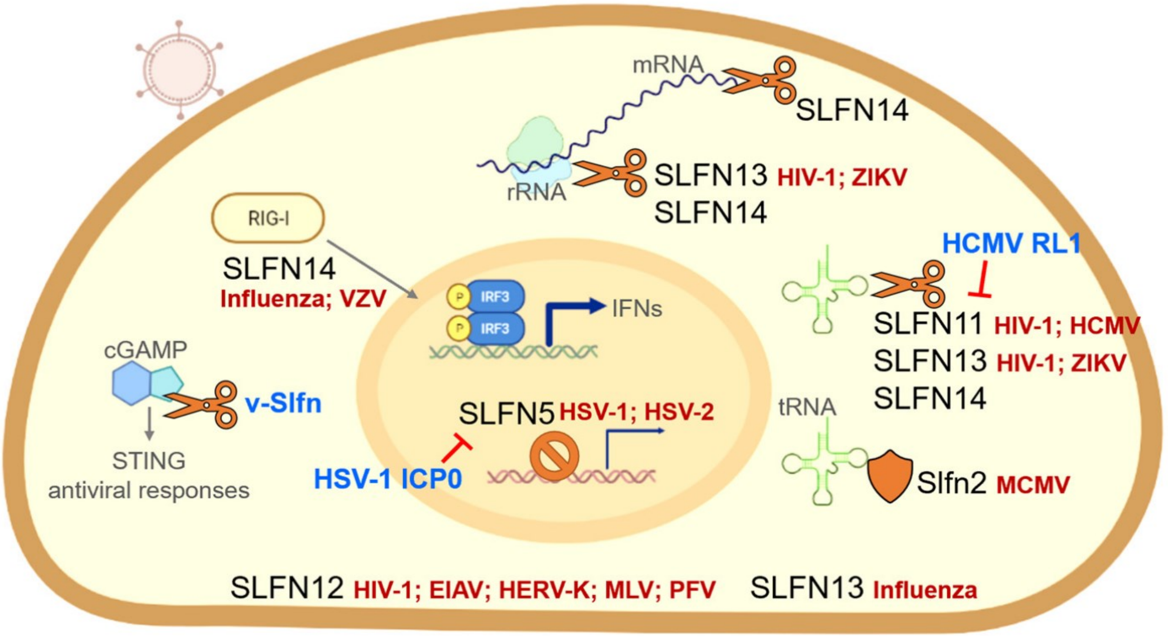


Figure 7: Schematic overview of the restriction factor functions of Slfn proteins. Slfn members employ different mechanisms to exert their antiviral activity. The nuclear Slfn member SLFN5 inhibits the transcription of viral genes e.g. in HSV-1. SLFN11, SLFN12, SLFN13 and SLFN14 interfere with translation by cleaving of tRNA, rRNA or mRNA. Slfn2 protects tRNA from oxidative stress-induced cleavage. Viral antagonistic mechanisms include ubiquitin-mediated degradation of SLFN5 in HSV-1 or SLFN11 in HCMV. Furthermore, v-Slfns that consist of a Slfn core domain and a poxin domain interfere with the cGAS-STING-mediated antiviral response by degrading cGAMP. Figure taken from [10].

1.4.3. Antiviral activity of other SLFN members

SLFN13 exhibits antiviral activity against HIV-1, shows a moderate restriction of ZIKV, but does not restrict the DNA virus HSV-1 [22]. Inhibition of HIV-1 replication by SLFN13 is caused by inhibition of protein synthesis of viral proteins [22]. This is dependent on the nucleolytic activity of SLFN13, which cleaves tRNA and rRNA in vitro (Figure 7). The secondary structure of tRNAs is recognized by the SLFN13 core domain and three acidic residues are involved in the Mg^{2+} and Mn^{2+} dependent endoribonuclease reaction [22].

SLFN13 and *SLFN14* gene expression levels are upregulated in response to influenza virus infections [8]. Depletion of SLFN13 leads to increased replication of influenza B virus, while SLFN14 overexpression inhibits replication of influenza A. This establishes SLFN13 and SLFN14 as restriction factors for negative-sense ssRNA viruses [8].

Similar to SLFN13, SLFN14 exhibits Mg^{2+} and Mn^{2+} dependent endoribonuclease activity on different RNA targets in vitro [23]. A C-terminally truncated rabbit SLFN14 isoform cleaves tRNA and binds to ribosomes, where it cleaves rRNA and ribosome-associated mRNA (Figure 7). However, full-length SLFN14 was found to neither cleave RNA nor does it associate with ribosomes [23]. Whether the endoribonuclease activity of SLFN14 is responsible for its previously observed antiviral activity is yet to be determined.

While the restriction factors described above belong to SLFN subgroup III, little is known about the role of subgroup I and II SLFN members in antiviral immunity [10]. A point mutation in the subgroup I protein mouse Slfn2 has been described to lead to a reduced number of T cells and enhanced susceptibility to viral infections [32]. In screens for ISGs with antiviral activity, subgroup II member

SLFN12 was identified as a restriction factor targeting different viruses, e.g. the retrovirus HIV-1 (Figure 7) [113, 114].

Genes with sequence similarity to subgroup I Slfns have been identified in the genomes of different orthopoxviruses (OPV) [1]. Due to this similarity, it has been proposed, that these so-called viral *Slfn* genes (*v-Slfn*), were horizontally transferred from rodents to OPVs [3]. In addition to the SLFN core domain, v-Slfns comprise an N-terminal poxvirus immune nuclease (poxin) domain [115, 116]. The poxin domain degrades the second messenger cyclic GMP-AMP (cGAMP), which is produced by cGAS in response to cytosolic DNA, thereby inhibiting the cGAS-STING pathway [116, 117]. In mousepox disease, v-Slfn significantly contributes to virulence in an in vivo setting [117]. However, the effect of v-Slfn is dependent on the poxin domain, as deletion of the poxin domain shows a similar phenotype as a v-Slfn knockout (Figure 7) [117]. In fact, several OPVs encode poxins that are not fused to a Slfn core domain [117]. Thus, the exact function of the Slfn core domain of v-Slfns is yet to be determined.

1.5. SLFN5 in cancer

SLFN5 is a subgroup III SLFN member that is localized in the nucleus [2, 27]. It is expressed in many tissues but shows differential up- or downregulation in various tumors [68, 118]. Its role has been investigated in different tumor types, showing that SLFN5 can either inhibit [11, 58-63] or stimulate [14, 64-68] tumorigenesis. These results imply that SLFN5 acts through different mechanisms in a cancer cell type specific fashion. However, the current understanding of the underlying mechanisms is only fragmentary.

The first report that directly connected a SLFN member to the regulation of cell invasion was a study on IFNα-sensitive malignant melanoma [58]. This type of cancer is inhibited by treatment with IFNα; however, the exact mechanism is not fully understood. While the investigated SLFN members (SLFN5, SLFN11, SLFN12, SLFN13, and SLFN14) show basal expression levels in malignant melanoma cells, only SLFN5 is significantly upregulated upon treatment with IFNα. Furthermore, SLFN5 expression is suppressed in malignant melanoma cells and knockdown of SLFN5 results in increased anchorage-independent growth and invasiveness in vitro [58]. Together, this suggest that SLFN5 contributes to the antitumor effects of IFNα in malignant melanoma.

In renal cell carcinoma (RCC) patients, high SLFN5 expression correlates with better overall survival [11]. Similar to malignant melanoma cells, SLFN5 is the only SLFN member that is significantly upregulated in RCC cells upon treatment with IFNα [11, 58]. In RCC cells, SLFN5 negatively controls motility and invasiveness by downregulating the expression of genes involved in cell motility such as matrix metalloproteinases 1 and 13 (MMP-1 and MMP-13) [11]. Overexpression of SLFN5 reduces motility and invasiveness, while knockdown of SLFN5 leads to increased motility. Due to the changes in the transcriptome induced by the knockdown of SLFN5, it is speculated that SLFN5 may serve as a transcriptional repressor.

Another study showed that SLFN5 expression is significantly higher in non/low invasive cancer cell lines compared to highly invasive cancer cell lines [59]. Knockdown of SLFN5 in non/low invasive cancer cell lines stimulates MMP-14 expression, leading to increased cancer cell migration and invasion. This effect is due to activation of the β-catenin signaling pathway that regulates MMP-14 expression. On the other hand, SLFN5 overexpression in highly invasive cancer cell lines inhibits MMP-14 expression, leading to decreased migration and invasion in vitro and in vivo. Knockdown of MMP-14 reverses the upregulation of migration and invasion that is observed in SLFN5 knockdown cells, indicating that MMP-14 contributes to SLFN5-dependent regulation of migration and invasion in cancer cells.

A positive correlation between SLFN5 expression levels and overall survival has also been reported in patients suffering from non-small cell lung cancer [60]. Knockdown of SLFN5 stimulates the expression of MMP-9 in vitro, resulting in increased cell migration.

Together, this shows that SLFN5 contributes to the regulation of motility and invasiveness in different cancer cell types by regulating the expression of migration-associated matrix metalloproteinases.

In breast cancer (BRCA) cell lines, SLFN5 inhibits metastasis while its knockdown induces epithelial-mesenchymal transition (EMT) and stimulates invasiveness of BRCA cells [61]. Mechanistically, SLFN5 inhibits the expression of the transcriptional repressor zinc finger E-box-binding homeobox 1 (ZEB1), a transcription factor that contributes to EMT [61, 119]. The inhibitory effect is due to SLFN5 binding to a sequence motif within the *ZEB1* promoter and requires the C-terminal RNA/DNA helicase domain of SLFN5 [61]. Downregulation of ZEB1 leads to upregulation of ZEB1-regulated phosphatase and tensin homologue on chromosome 10 (PTEN), which is a known cancer suppressor [62]. A role of PTEN signaling was further reported in lung adenocarcinoma where SLFN5 was shown to inhibit proliferation and promote apoptosis [63].

In contrast to these reports that show inhibitory effects on tumorigenesis, SLFN5 has also been shown to stimulate tumorigenesis in certain types of cancer [14, 64-68].

Intestinal metaplasia is an abnormal transformation of epithelium cells of the stomach that is typically caused by infections with *Helicobacter pylori* and delineates a precursor for gastric cancer [120]. Analysis of SLFN5 expression revealed that high expression levels correlate with progression from intestinal metaplasia to gastric cancer [14]. Furthermore, high SLFN5 expression is associated with a shorter overall survival in gastric cancer patients [68].

SLFN5 promotes the malignant phenotype in glioblastoma multiform (GBM), a type of tumor of the central nervous system, and its expression correlates with a shorter overall survival in GBM patients [64]. In vitro, SLFN5 promotes motility and invasiveness of GBM cells. Injection of SLFN5 wild type or knockout GBM cells into mice showed delayed tumor formation and decreased growth of SLFN5 knockout cells in vivo. Mechanistically, SLFN5 interferes with IFN-signaling by acting as a transcriptional co-repressor of STAT1 [64]. Type I IFN-treatment stimulates complex formation of SLFN5 with STAT1 that represses STAT1-mediated transcription. The dual role of SLFN5 as an ISG and as a repressor of IFN-signaling suggests the existence of a negative-feedback loop that may play a role in suppression of antitumor immunity in GBM.

In patients suffering from prostate cancer, high SLFN5 expression correlates with shorter relapse-free survival and increased risk of metastasis [67]. On the other hand, depletion of SLFN5 leads to impaired tumor growth in vivo. Mechanistically, SLFN5 interacts with the activating transcription factor 4 (ATF4) and controls the expression of ATF4 target genes. One of those targets is the large neutral amino acids transporter 1 (LAT1) which, if upregulated, alters amino acid homeostasis within the cells and is associated with an increased risk of metastasis [121].

SLFN5 is overexpressed in pancreatic ductal adenocarcinoma (PDAC) and high expression levels correlate with shorter overall survival [66]. Knockout of SLFN5 results in decreased proliferation of PDAC cells and reduced tumorigenesis in vivo. The effect of SLFN5 on tumorigenesis depends on binding to the transcriptional repressor E2F7 [66]. Binding of SLFN5 to E2F7 blocks its function and leads to the transcription of E2F7-controlled target genes, many of which are regulators of cell cycle progression and stimulate S phase progression.

Taken together, the precise role of SLFN5 in tumorigenesis seems to be highly tumor specific and can be either inhibitory or stimulatory [118]. However, in cases where SLFN5 promotes the malignant phenotype, targeting of SLFN5 may offer therapeutic potential [64, 66, 67].

1.6. SLFN11

The tremendous advances in next generation sequencing technologies enabled the analysis of the genomic and transcriptomic landscape of a large number of cancer cell lines, facilitating the identification of new biomarkers [122]. In the search for predictive markers for cancer therapy, libraries of gene expression data from many different cancer cell lines were analyzed regarding their correlation with the cytotoxicity profiles of different anticancer drugs [17, 123]. SLFN11 expression was found to be a predictive biomarker for the sensitivity of many cancer cell lines to various DNA damaging agents (DDAs), such as topoisomerase I (TOP1) inhibitors and cisplatin [17, 123]. Subsequent work established SLFN11 as a regulator in DNA damage response with replication checkpoint activity, as it can irreversibly block stalled replication forks [45, 71, 124, 125].

1.6.1. DNA replication and replication stress

DNA replication is a highly regulated process that ensures that the DNA is replicated exactly once per cell cycle [126, 127]. The semiconservative replication of the double-stranded DNA occurs mainly during S phase of the cell cycle and involves various multi-subunit protein complexes.

Replication is initiated during G1 phase by the licensing of the origin recognition complex (ORC) and cell division cycle 6 (CDC6) to replication origins [126, 128, 129]. The recruitment of two heterohexameric DNA helicase minichromosome maintenance complexes 2-7 (MCM2-7) and chromatin licensing and DNA replication factor 1 (CDT1) leads to the formation of the pre-replicative complex (pre-RC) (Figure 8) [130]. Subsequently, the licensing factors CDC6 and CDT1 are released from the pre-RC and cycline-dependent kinases (CDKs) and Dbf4-dependent kinase (DDK) phosphorylate several replication factors, facilitating their loading onto the origin and assembly of the pre-initiation complex (pre-IC) (Figure 8) [128, 131, 132]. Two of those factors are cell division cycle 45 (CDC45) and the GINS complex that, together with MCM2-7, form the replicative CDC45-MCM2-7-GINS (CMG) helicase [133]. Phosphorylation of MCM2-7 by CDKs and DDK triggers firing of the CMG helicase, which marks the beginning of the S phase [134]. The active CMG helicase unwinds the DNA duplex into single-stranded DNA (ssDNA), forming the characteristic replication fork [135]. This triggers the recruitment of a variety of additional factors to the replication fork to form the replisome (Figure 8) [126, 128]. These factors include replication protein A (RPA) that binds to ssDNA [136], DNA polymerase ε (Pol ε) for leading strand synthesis [137], the DNA polymerase α (Pol α)–DNA primase complex that synthesizes and extends RNA primers on the lagging strand [138], DNA polymerase δ (Pol δ) for lagging strand synthesis [139], and proliferating cell nuclear antigen (PCNA) that acts as a DNA clamp and processivity factor for Pol ε and Pol δ [140]. Topoisomerases act ahead of the CMG helicase to resolve the torsional stress that is associated with DNA duplex unwinding [141]. The replisome moves along and replicates the DNA until it converges with another replication fork, which marks the end of the DNA replication and S phase [142]. Of note, only a subset of origins fire during

DNA replication with most origins remaining silent, even though they may be licensed [143, 144]. Thus, the exact set of active origins appears flexible and differs between cells.

DNA replication can be disturbed by different factors, resulting in replication stress [145-147]. This can have endogenous or exogenous sources and is linked to genome instability and disease. Endogenous sources include DNA secondary structures such as G-quadruplexes that can act as natural barriers for the replisome, DNA damage caused by e.g. oxidation or reactive metabolites, limited chromatin accessibility or collisions of replication forks with sites of active transcription [145-147]. Exogenous sources include DNA damage induced by ionizing radiation and DNA damaging chemotherapeutic agents [145-147]. DNA damaging agents (DDAs) act via different mechanisms, often by introducing cytotoxic DNA lesions such as DNA base mismatches, DNA breaks, DNA-protein crosslinks or interstrand crosslinks [145, 148].

Many factors that cause replication stress can lead to stalling of the replication fork [146]. Interstrand crosslinks for example prevent the unwinding of the DNA duplex by the CMG helicase, causing it to stop [149]. Bulky DNA adducts on the other hand, can lead to an uncoupling of the DNA polymerase from the CMG helicase [150]. In that case, the helicase continues to unwind the DNA duplex, while the polymerase stalls at the DNA lesion, leading to the formation of extended stretches of single-stranded DNA (Figure 9) [150]. RPA coated single-stranded DNA is a signal for replication stress and serves as a platform for the recruitment of replication stress response proteins, including the central replication stress regulator ATM- and Rad3-related (ATR) [146]. The signaling kinase ATR is recruited via the ATR-interacting protein (ATRIP) that binds RPA coated single-stranded DNA [151]. ATR phosphorylates several targets, including the downstream checkpoint kinase 1 (CHK1), which, upon activation, causes cell cycle arrest, inhibits firing of late origins and stabilizes the replication fork [152, 153]. The transient nature of the ATR/CHK1-mediated cell cycle arrest allows for DNA repair and subsequent replication restart [154].

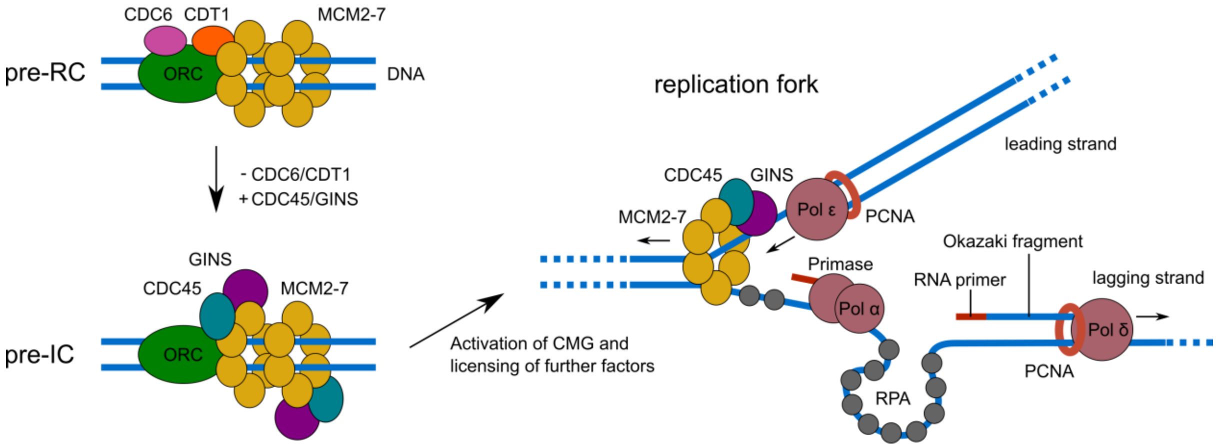


Figure 8: Overview of the assembly of the human replisome. In G1 phase, the pre-RC is assembled at an origin of replication and consists of the ORC, CDC6, CDT1 and the helicase MCM2-7. Release of CDC6 and CDT1 and kinase dependent recruitment of several replication factors such as CDC45 and GINS to MCM2-7 leads to the formation of the pre-IC. Firing of the activated CMG helicase at the onset of S phase leads to unwinding of the DNA duplex and the formation of the replication fork. Additional factors, such as DNA polymerases and RPA are recruited to the replication fork, forming the replisome. Pol ε facilitates leading strand synthesis, while primase synthesizes RNA primers on the lagging strand, that are extended by Pol α and Pol δ. Stretches of single-stranded DNA are protected by RPA. Figure based on [126, 128].

Different types of DNA damage are repaired via different DNA repair mechanisms [155]. DNA interstrand crosslinks that block progression of the CMG helicase are repaired via the Fanconi anemia

(FA) pathway [156]. Furthermore, FA proteins as well as DNA repair protein RAD51 homolog 1 (RAD51) are involved in the stabilization and protection of 4-way DNA junctions, so called “chicken foots” [157]. These DNA structures are generated upon reverse remodeling of stalled replication forks and require protection from nucleolytic degradation, to ensure genome stability and replication restart [157]. Certain DNA damages can bypass the CMG helicase and lead to uncoupling of the polymerase from the helicase activity and subsequent stalling of the polymerase [150]. In a process called translesion synthesis (TLS), specialized TLS polymerases with low proofreading activity temporarily replace the replicative polymerase and incorporate nucleotides at sites that harbor a damaged DNA template [158]. This mechanism allows postponing of DNA repair until after the DNA replication has finished. Further mechanisms that cope with replication stress include repriming of DNA synthesis downstream of DNA lesions [159, 160] or activation of dormant origins proximal to stalled replication forks [161].

While low levels of replication stress are associated with genome instability and carcinogenesis, high levels of replication stress can lead to cell death [145]. DNA damaging chemotherapeutic agents take advantage of this, as fast dividing cancer cells are especially sensitive to high levels of DDA-induced replication stress [162]. SLFN11 has emerged as an important factor in this process as it possesses replication checkpoint activity and contributes to the cytotoxicity of DDAs by the irreversible blocking of stalled replication forks [45, 71, 124, 125].

1.6.2. Mechanisms underlying SLFN11 replication checkpoint activity

DDAs such as cisplatin or topoisomerase inhibitors represent exogenous sources of DNA damage that cause replication stress and stalling of replication forks [162-164]. SLFN11 expression levels correlate with the cytotoxicity profiles of many DDAs, making SLFN11 a predictive biomarker for the sensitivity of cancer cells to these chemotherapeutic agents [17, 123]. This correlation is based on the ability of SLFN11 to irreversibly block stressed replication forks [71, 165, 166]. The underlying molecular mechanisms are only partially understood and it became apparent that SLFN11 acts in diverse ways (Figure 9).

In response to DNA damage and replication stress, SLFN11 is recruited to RPA-coated single-stranded DNA [71, 124]. Single-stranded DNA can accumulate upon DNA damage due to uncoupling of helicase and polymerase activity at stressed replication forks or by nucleolytic processing of DNA double-strand breaks [150, 167]. RPA is a heterotrimeric complex composed of RPA1, RPA2 and RPA3 that specifically binds to single-stranded DNA to protect and stabilize it [136]. RPA-coated single-stranded DNA serves as a signal and platform for the recruitment of a variety of factors with diverse biochemical functions. In addition to SLFN11, this includes the aforementioned ATR-ATRIP kinase complex (Figure 9) [151]. SLFN11 directly interacts with RPA1 via the C-lobe of the SLFN11 helicase domain [124]. Upon DNA damage induced by the topoisomerase inhibitor camptothecin (CPT), SLFN11 co-localizes with RPA at sites of DNA end resection as well as at replication foci [71, 124]. However, knockdown of RPA disrupts the recruitment of SLFN11 to these sites [71, 124]. At stalled replication forks, SLFN11 interacts with MCM3, a subunit of the MCM helicase [71, 130], and DExH-box helicase 9 (DHX9), a RNA/DNA helicase involved in DNA damage repair [18, 71, 168]. SLFN11 irreversibly blocks stalled replication forks, which eventually leads to cell death (Figure 9) [71]. ATR on the other hand, which is also recruited to stalled replication forks, blocks replication only transiently, allowing DNA repair and cell survival [153, 154]. The replication block by SLFN11 is independently of ATR, as ATR inhibition has little influence on CPT mediated cell killing in SLFN11 positive cells [71]. However, SLFN11 negative cells show chemoresistance to CPT, which can be overcome by ATR inhibition [71, 165]. The mechanism of action seems to be different from the ATR-CHK1 pathway, as SLFN11 inhibits neither CDC45 nor PCNA loading [71]. The ATPase activity of SLFN11 is required for replication inhibition but not for the recruitment of SLFN11 to chromatin [71].

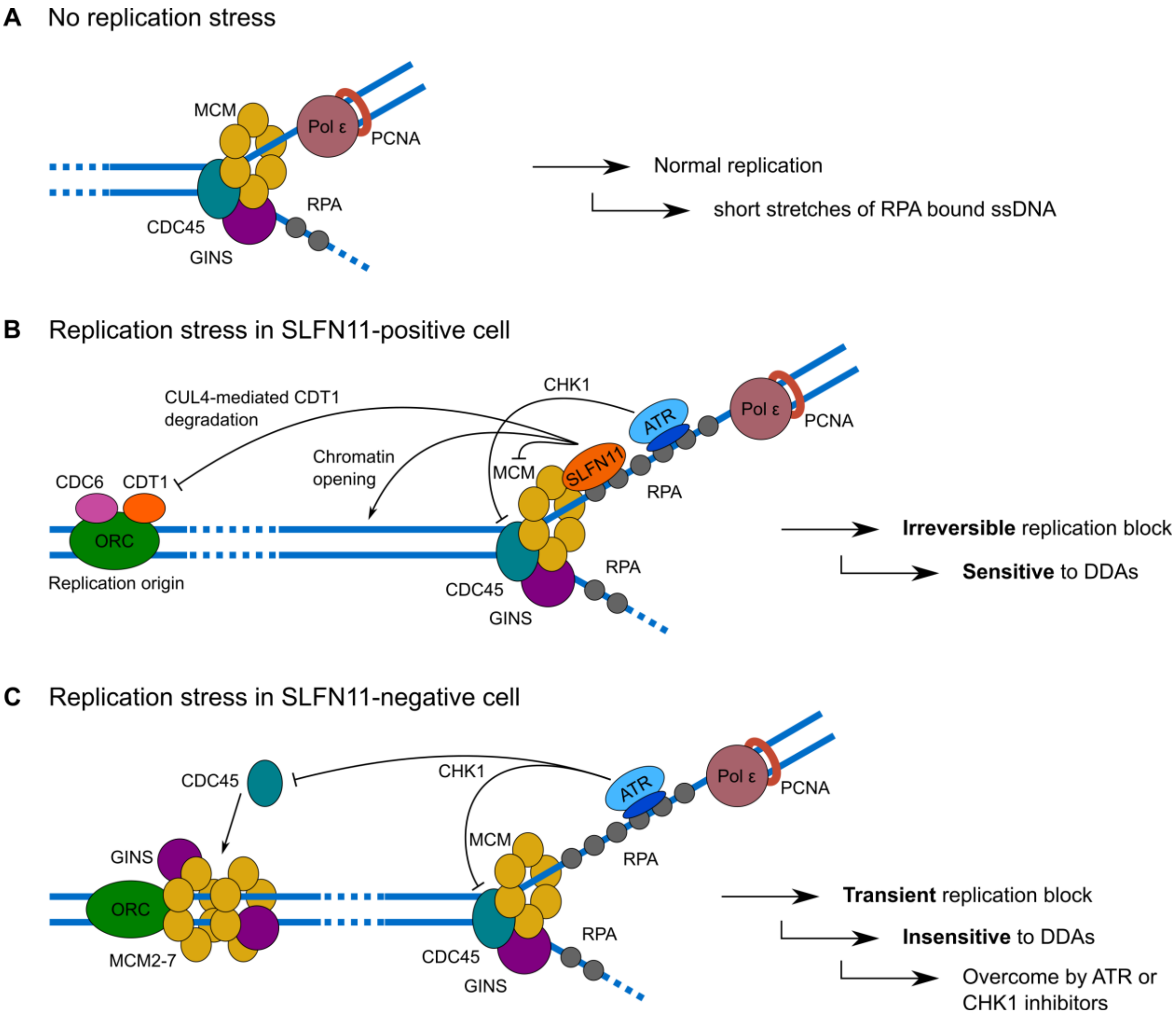


Figure 9: Overview of stalled DNA replication forks in SLFN11 positive and negative cells. (A) DNA replication fork in stress-free state. The dsDNA is opened by the CDC45-MCM-GINS (CMG) helicase and DNA is replicated by the polymerase (only leading strand polymerase visualized). Only short RPA-coated ssDNA stretches are present. (B) Upon treatment of SLFN11 positive cells with DDAs, ATR-ATRIP and SLFN11 are recruited to stalled replication forks by long RPA-coated ssDNA stretches, which can result from helicase-polymerase uncoupling. SLFN11 directly interacts with MCM3, causes chromatin opening in front of the replication fork and inhibits replication origin firing by the degradation of the replication licensing factor CDT1. This leads to an irreversible replication block upon DDA treatment that eventually leads to cell death. (C) Upon treatment of SLFN11 negative cells with DDAs, ATR-ATRIP is recruited to stalled replication forks by long RPA-coated ssDNA stretches. ATR signals via the ATR-CHK1 kinase pathway that inhibits the recruitment of CDC45, slows down and stabilized replication forks, leading to a transient replication block. The transient nature of the block can cause chemoresistance to DDAs, as it allows the cell to repair the damage. However, this can be overcome by inhibition of ATR or CHK1. Figure based on [169].

Upon CPT-induced replication stress, SLFN11 increases chromatin accessibility in the vicinity of replication initiation sites in an ATPase dependent manner [71]. Chromatin opening ahead of the replication fork has been speculated to block replication fork progression (Figure 9B) [71]. In addition, SLFN11 increases genome-wide chromatin accessibility at pre-existing chromatin accessible regions such as active promotors and induces the transcription of immediate early genes (IEGs), some of which are known to contribute to cell cycle arrest [72].

Furthermore, SLFN11 promotes the proteasomal degradation of the replication licensing factor CTD1 in response to CPT treatment (Figure 9B) [165]. This is believed to prevent firing of late origins and irreversibly block cell cycle progression. Mechanistically, this is due to a direct interaction of SLFN11 with DNA damage-binding protein 1 (DDB1), a subunit of the DDB1-CUL4^{CDT2} E3 ubiquitin ligase complex which is part of the replisome [165, 170]. SLFN11 interacts with DDB1 via its helicase domain and is suggested to act as a cofactor that activates CUL4 E3 ligase activity [165]. The ATPase activity of SLFN11 is critical for SLFN11-induced CDT1 degradation, as it is for replication arrest in response to CPT.

It has been found that SLFN11 boosts the Fanconi anemia (FA) phenotype by promoting the degradation of stalled replication forks [166]. As mentioned before, FA is a disease that is characterized by hypersensitivity to interstrand crosslinks (ICLs) [157]. SLFN11 knockdown in FA cells or wild type cells improves cell survival upon treatment with ICL-inducing agents [166]. Mechanistically, SLFN11 negatively affects the recruitment of RAD51 to reversed replication forks, in an SLFN11 ATPase dependent manner. Without the protection by RAD51, the reversed fork is subject to nucleolytic degradation by double-strand break repair protein MRE11 and DNA replication ATP-dependent helicase/nuclease DNA2, resulting in genome instability. Thus, SLFN11 is discussed as a therapeutic target for the treatment of FA.

RPA-coated single-stranded DNA recruits SLFN11 not only to stalled replication forks but also to end resected DNA double strand breaks [124]. SLFN11 directly interacts with RPA and reduces RPA foci in CPT treated cells, indicating that SLFN11 might destabilize the RPA-ssDNA complex. Cells expressing high levels of SLFN11 show defects in homologous recombination (HR) repair and an inhibition of CHK1 DNA damage checkpoint maintenance [124].

SLFN11 has also been shown to inhibit the translation of the Ser/Thr kinases ATR and ataxia telangiectasia mutated (ATM) in response to DDA treatment [45]. The codon-usage-dependent inhibition is due to cleavage of type II tRNAs, which includes serine and leucine tRNAs [45, 109]. The N-terminal Slfn core domain is essential for tRNA cleavage and a mutant lacking the C-terminal helicase domain is still able to inhibit translation [45]. mRNAs with high usage of the rarely used TTA (Leu) codon show high susceptibility to translational suppression by SLFN11. These mRNAs include the ones coding for ATR, ATM but also for many viral proteins as mentioned before (Section 1.4) [6, 45].

The endonuclease activity of SLFN11 is regulated by phosphorylation [111]. Three phosphorylation sites were identified, two of which are located in the Slfn core domain and one in the helicase domain. SLFN11 phosphomimetics show decreased tRNA endonuclease activity and cells expressing these mutants are less susceptible to CPT treatment than SLFN11^{WT} cells. Phosphatase 1 catalytic subunit γ (PPP1CC) acts as the upstream enzyme that dephosphorylates SLFN11 so that it becomes fully active. Upon CPT treatment, SLFN11 phosphorylation is reduced over time which is in line with the induction of SLFN11 activity upon DDA treatment.

Overall, SLFN11 acts via diverse mechanisms to sensitize cells to DNA damage, making it a potential predictive biomarker for the cancer treatment with DDAs.

1.6.3. SLFN11 as a predictive biomarker in cancer

As SLFN11 can sensitize cancer cells to various replication stress-inducing chemotherapeutic agents, its expression level may serve as a predictive biomarker of response to cancer therapeutics (Figure 10)

[17, 123]. The predictive value of SLFN11 expression has been established in preclinical studies for different DDAs including alkylating and crosslinking agents (cyclophosphamide, temozolomide, cisplatin, carboplatin and oxaliplatin) [17-20, 171-176], TOP1 inhibitors (camptothecin, topotecan, indotecan, and irinotecan) [16, 17, 20, 71, 123, 177-182], TOP2 inhibitors (etoposide, mitoxantrone, epirubicin, and doxorubicin) [17, 177], and DNA synthesis inhibitors (gemcitabine, cytarabine, and hydroxyurea) [17, 183]. Furthermore, SLFN11 expression correlates with the sensitivity of cancer cells to poly(ADP-ribose)-polymerase (PARP) inhibitors (olaparib, veliparib, talazoparib, and niraparib) [19, 125, 173, 184, 185]. While the induction of replication stress is a hallmark of these chemotherapeutic agents, they do so via different mechanisms. Platinum drugs for example can cause ICLs that block the CMG helicase, leading to replication fork stalling [186]. TOP1 inhibitors covalently trap topoisomerase-DNA cleavage complexes, which can lead to replication-associated DSBs [187]. DNA synthesis inhibitors may act by direct inhibition of the DNA polymerases or by depletion of the dNTP pool [188]. PARP inhibitors block the ADP-ribose-polymerase activity and can trap PARP-DNA complexes, which can lead to replication blocks or replication-associated DSBs [189]. However, SLFN11 expression does not correlate with sensitivity to chemotherapeutic agents that do not induce replication stress, such as kinase inhibitors or tubulin poisons [17].

A correlation between SLFN11 expression levels and sensitivity of cancer cells to DDAs has been shown for a variety of cancer cell lines or patient derived xenograft models, including small-cell lung cancer (SCLC) [19, 184, 190, 191], breast cancer [20, 21, 180], ovarian cancer [17, 172, 173], prostate cancer [175], gastric cancer [174, 192], colorectal cancer [16, 193], bladder cancer [171], sarcoma [181, 194-196], and mesothelioma [185].

However, SLFN11 is only expressed in about 50% of cancer cell lines and a lack of SLFN11 expression correlates with chemoresistance to DDAs [4, 17, 123, 127, 197]. The lack of SLFN11 expression is largely due to epigenetic regulation. Epigenetic silencing by CpG island hypermethylation of the *SLFN11* gene promoter suppresses SLFN11 expression and correlates with chemoresistance to DDAs and poor prognosis [18, 192, 193, 198]. In addition, epigenetic silencing of SLFN11 may be acquired during treatment with DDAs and contribute to acquired chemoresistance [174, 199]. In SCLC, acquired chemoresistance is associated with the histone modification H3K27me3 within the *SLFN11* gene, which is placed by the histone methyltransferase Enhancer of Zeste Homolog 2 (EZH2) and induces local chromatin condensation and gene silencing [199].

Several strategies have been tested to re-express SLFN11 in order to sensitize chemoresistant SLFN11 low/negative cells to DDAs (Figure 10) [197]. Treatment of *SLFN11* hypermethylated breast or gastric cancer cell lines with a DNA methylation inhibitor resensitized the cancer cells to platinum drugs [18, 174]. Another study restored SLFN11 expression by the inhibition of EZH2 in a SCLC model [199]. Finally, histone deactelylase (HDAC) inhibitors were shown to induce SLFN11 expression in cells without promoter hypermethylation, sensitizing them to DDA treatment [183, 191]. These results underscore the synergistic potential for the combination of epigenetic drugs with DDAs in cancer therapy.

An alternative approach to sensitize chemoresistant SLFN11 low/negative cells to DDAs is the inhibition of the ATR/CHK1 pathway [71, 125, 165]. In SLFN11 positive cells, SLFN11 "overwrites" the ATR/CHK1-mediated transient replication block by irreversibly blocking the replication fork, eventually leading to cell death. However, in SLFN11 low/negative cells, DDA treatment activates the ATR/CHK1 pathway, which leads to a transient replication block that allows for DNA damage repair and a subsequent DNA replication restart. This is an important factor contributing to chemoresistance of SLFN11 low/negative cells towards certain DDAs. Thus, inhibition of ATR or CHK1 can resensitize chemoresistant SLFN11 low/negative cells to DDA treatment (Figure 10) [21, 190, 200].

As preclinical data indicate that SLFN11 expression levels are a promising predictive biomarker of response to certain cancer therapeutics, efforts have been made to translate this knowledge to the clinics. It has been shown that immunohistiochemical assays can be used to reliably monitor SLFN11 expression status in different cancer types [21, 184, 201]. In fact, the predictive value of SLFN11 expression is actively investigated in several clinical trials (NCT04334941, NCT03880019) [202-204]. First results show that SLFN11 expression correlates with improved progression-free survival (PFS) and overall survival (OS) in SCLC patients, treated with a combination of DDA and PAPR inhibitor [202].

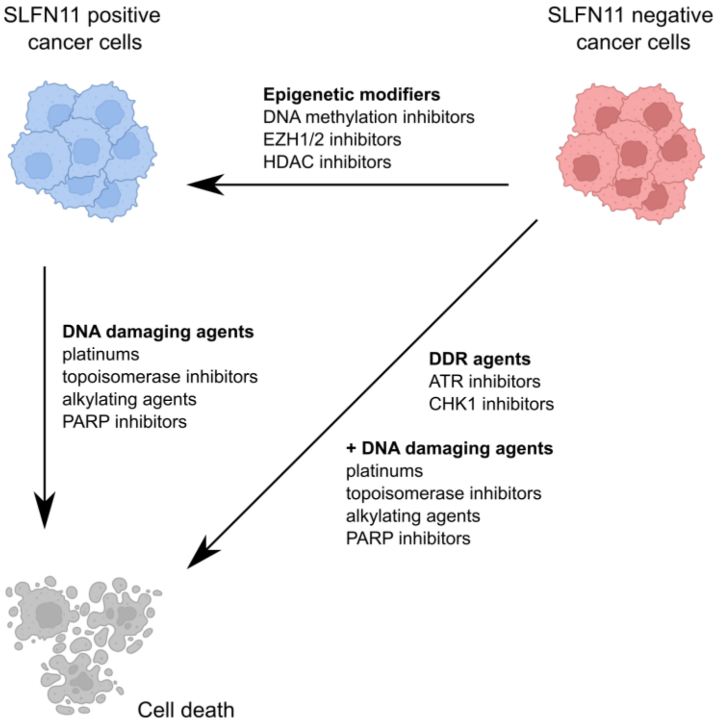


Figure 10: Therapeutic strategies for inducing SLFN11-mediated cell death. Treatment of SLFN11 positive cancer cells with DDAs cause a SLFN11-mediated irreversible block of replication forks, leading to cell death. If SLFN11 is epigenetically silenced in cancer cells, a treatment with epigenetic modifiers such as DNA methylation inhibitors may restore SLFN11 expression and make them vulnerable for DDA treatment. Alternatively, SLFN11 negative cells may be co-treated with ATR or CHK1 inhibitors and DDAs to overcome chemoresistance. Created with BioRender.com. Figure based on [197].

Besides its role as a predictive biomarker, the prognostic value of SLFN11 expression is also actively investigated. In hepatocellular carcinoma (HCC), overexpression of SLFN11 inhibits cancer cell proliferation, migration, and invasion in vitro, and impedes HCC growth and metastasis in vivo [205]. Mechanistically, this is due to an interaction of SLFN11 with the oncogenic ribosomal protein S4 X-Linked (RPS4X), which facilitates the inactivation of the mammalian target of rapamycin (mTOR) signaling pathway. Interestingly, SLFN11 was mostly located in the cytoplasm of HCC cells as shown by immunohistiochemical staining, which is contrary to the nuclear localization of SLFN11 in other cancer cell lines [21, 205]. Furthermore, high SLFN11 expression was correlated with better OS compared to patients with low SLFN11 expression in CRC and BC patients [176, 180]. In contrast, high SLFN11 expression was an unfavorable prognostic marker in patients suffering from bladder cancer as shown by reduced OS [171]. In a patient-derived glioblastoma xenograft model, SLFN11 knockout showed delayed tumor growth and improved survival [206].

Taken together, SLFN11 expression may offer prognostic value; however, its role in tumorigenesis is not well understood and seems to be tumor specific as it can be inhibitory or stimulatory [205, 206].

1.7. The INO80 chromatin remodeling complex

1.7.1. Chromatin architecture

In eukaryotes, DNA is organized in the form of chromatin, a protein-DNA complex that compacts the DNA in the nucleus and plays a role in the spatial and temporal regulation of vital processes such as transcription, DNA replication and cell division [207, 208]. The basic unit of chromatin is the nucleosome core particle (NCP), which consist of approximately 147 bp DNA wrapped around histone proteins (Figure 11A) [209-212]. The histone octamer that constitutes the core of the NCP consists of a H3-H4 tetramer and two H2A-H2B dimers, resulting in two copies of each of the four histone proteins H2A, H2B, H3 and H4 (Figure 11B) [209, 213]. The NCP exhibits twofold symmetry and the center of the approximately 1.7 superhelical turns of nucleosomal DNA that wrap around the core is termed the dyad [209, 212]. At the dyad, the major groove faces the histone core and is defined as super helical location 0 (SHL 0) [209]. The SHL changes by +/- 1 at each successive turn of the DNA helix, resulting in a numbering from SHL -7 to SHL +7 (Figure 11C) [209]. Together with flanking extranucleosomal linker DNA, the NCP forms the nucleosome [214].

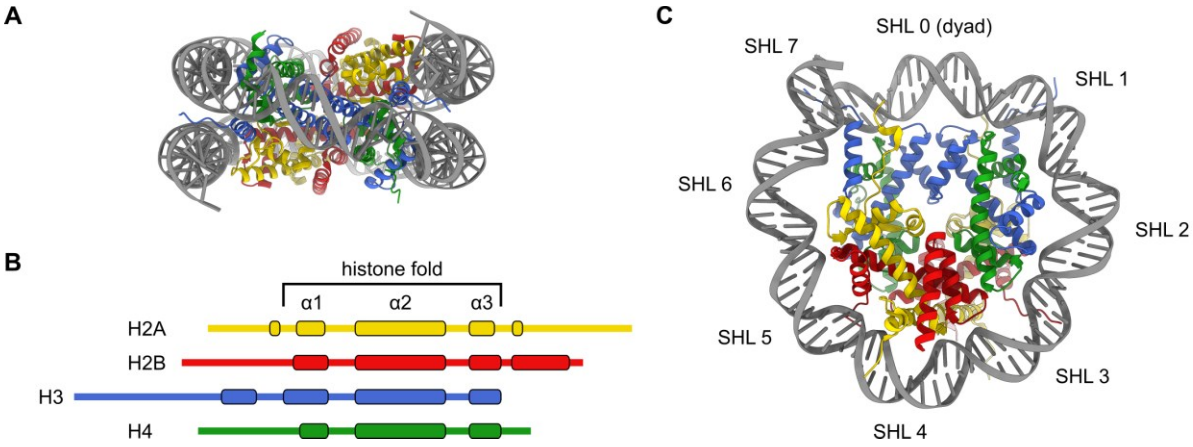


Figure 11: Structure of the nucleosome core particle. (A) Structure of the nucleosome core particle (PDB: 7OHC) [215]. DNA is colored in grey and the histones in yellow (H2A), red (H2B), blue (H3), and green (H4). (B) Domain architecture of the four canonical histones H2A (yellow), H2B (red), H3 (blue), and H4 (green). Squares represent α-helical regions. The histone fold is indicated and conserved α-helices are labeled. (C) Top view of the nucleosome core particle (PDB: 7OHC) [215]. Only one DNA gyre is shown for clarity. The superhelical locations (SHLs) of one DNA gyre are labeled, starting at the dyad (SHL 0). Figure based on [214].

The proteinaceous surface of the nucleosome harbors an acidic residue-rich cluster formed by H2A and H2B, referred to as the acidic patch [209, 216]. This conserved region serves as a binding site for various chromatin-associated factors that interact with the acidic patch via a so-called arginine anchor [217-219]. Furthermore, the H4 tail can interact with the acidic patch of neighboring nucleosomes, contributing to the formation of higher order chromatin structures [209, 220]. In addition to their structurally conserved histone folds, all histones exhibit unstructured N-terminal, and in case of H2A also C-terminal, extensions that are referred to as histone tails (Figure 11B) [209, 214, 221]. Histone tails are a hotspot for posttranslational modifications (PTMs) such as phosphorylation, acetylation or methylation [222, 223]. These histone marks are deposited, interpreted or removed by factors referred to as writers, readers and erasers and impact chromatin structure and chromatin-dependent processes such as transcription, DNA replication or DNA repair [224]. Canonical histones can be replaced by histone variants such as H2AZ that is enriched at promoter regions or H2AX that is phosphorylated in response to DNA double-strand breaks [225-227]. Furthermore, linker histone H1 can bind to the extranucleosomal linker DNA next to the NCP to form the chromatosome [228], resulting in a more compact and transcriptionally inactive chromatin conformation [229, 230].

The average nucleosome repeat length is proposed to range around 190 bp, resulting in an estimated 30 million nucleosomes per diploid human cell [231, 232]. In a relaxed state, nucleosomal arrays resemble "beads on a string", where nucleosomes are the beads connected by linker DNA [211]. In the nucleus, chromatin is tightly packed and its architecture depends on the cell cycle [233]. Chromatin can organize into topologically associated domains (TADs), which are chromatin regions that show strong internal interactions while showing less interactions with neighboring TADs [234-236]. Expression patterns within a TAD are often regulated by the same enhancer elements and TADs are conserved between different cell types [235, 236]. Chromatin can be further classified into euchromatin and heterochromatin [237]. Euchromatin forms a rather loose and dynamic structure, is transcriptionally active and enriched for RNA polymerase [238]. Heterochromatin on the other hand is more condensed and transcriptionally inactive.

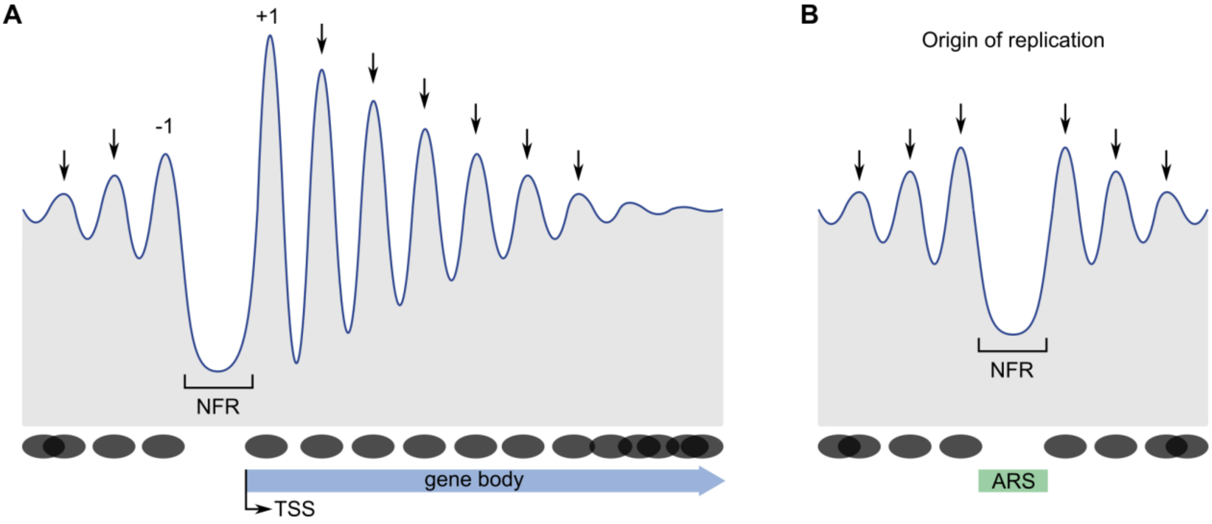


Figure 12: Nucleosome organization at transcription start sites and origins of replication. (A) Schematic view of nucleosome occupancy and positioning (blue line) relative to a transcription start site (TSS) at a gene promoter. Peaks indicate positions of high nucleosome occupancy and arrows mark the predicted dyad positions. The +1 and -1 nucleosomes that form the border of the nucleosome free region (NFR) are indicated. Nucleosomes are represented as grey ellipses to indicate the increasing fuzziness of the nucleosome positioning when moving into the gene body. Figure based on [239]. (B) Schematic view of nucleosome occupancy and positioning (blue line) relative to an autonomous replicating sequence (ARS) at an origin of replication. Peaks indicate positions of high nucleosome occupancy and arrows mark the predicted dyad positions. Nucleosomes are represented as grey ellipses to indicate the formation of an NFR around the ARS. Figure based on [239].

While nucleosomes enable the tight packing of DNA into the nucleus, they impair access to the genome, which is essential for DNA-dependent processes such as transcription or DNA replication [240]. Therefore, factors such as chromatin remodelers and histone modifying enzymes actively orchestrate nucleosome positioning and composition [241]. Transcriptionally active promoters exhibit a nucleosome-depleted or nucleosome free region (NDR/NFR), flanked by well positioned +1 and -1 nucleosomes and phased nucleosomal arrays, so-called genic nucleosomal arrays (Figure 12A) [242, 243]. Similarly, active origins of replication also harbor an NFR around the autonomously replicating sequence with flanking nucleosomal arrays (Figure 12B) [239, 244]. Genomic nucleosome positioning is best described for *S. cerevisiae*, where the +1 nucleosome is located at the transcription start site (TSS) [245]. The +1 nucleosome often incorporates the H2AZ histone variant and marks the beginning of phased nucleosomal arrays with a nucleosome repeat length of approximately 165 bp [242, 246]. Genic nucleosome arrays extend into the gene bodies but diffuse with increasing distance from the +1 nucleosome (Figure 12A) [239].

Nucleosome positioning is influenced by the DNA sequence, as the underlying DNA shape properties promote or inhibit nucleosome assembly as shown for +1 nucleosomes or NFRs [247-250]. However, ATP-dependent chromatin remodelers are essential for proper nucleosome positioning by generating NDRs, positioning of +1 and -1 nucleosomes and spacing/phasing of nucleosomal arrays [250, 251].

1.7.2. ATP-dependent chromatin remodelers

ATP-dependent chromatin remodelers harness the energy of ATP hydrolysis to shape the nucleosomal landscape by e.g. removal, deposition, sliding or spacing of nucleosomes or by incorporation or removal of histone variants [241, 252]. They are characterized by an ATPase domain that belongs to the superfamily 2 (SF2) RNA/DNA helicases and more specifically to the subfamily of sucrose non-fermenting 2 (Snf2)-type ATPases [253]. The Snf2-type ATPase domain serves as the motor domain in chromatin remodelers and functions by stepwise translocation of double-stranded DNA [254]. The ATPase domain consist of two RecA-like folds that are also referred to as N- and C-lobes and harbors the conserved Walker A and B motifs that are essential for ATP binding and hydrolysis [253]. Chromatin remodelers can be classified into four subfamilies based on their domain architecture; imitation switch (ISWI), chromodomain helicase DNA-binding (CHD), switch/sucrose non-fermenting (SWI/SNF), and inositol auxotroph mutant 80 (INO80) [252].

ISWI and CHD family members form small complexes or function as single-subunit remodelers while SWI/SNF and INO80 family members form large multi-subunit complexes [241]. All chromatin remodelers possess auxiliary domains besides the Snf2-type ATPase that fulfill different functions such as interaction with histone tails or extranucleosomal DNA [241]. The large remodeler complexes harbor additional subunits, many of which are involved in substrate recognition or regulation of the motor domain [241, 255]. ISWI remodelers have nucleosome sliding and spacing activity and are associated with the generation of nucleosomal arrays [250, 256, 257]. Furthermore, they harbor a HAND-SANT-SLIDE domain that interacts with the unmodified H3 tail and extranucleosomal DNA [258, 259]. CHD family members are associated with nucleosome spacing and the incorporation of histone variant H3.3 [260-262]. They possess a tandem chromodomain that in case of human CHD1 recognizes tri-methylated H3K4 [263] and a DNA-binding domain (DBD) that interacts with entry DNA [264]. The megadalton chromatin remodeling complexes of the SWI/SNF and INO80 families share a helicase/SANT-associated (HSA) domain followed by a post-HSA domain [241, 255]. The HSA domain serves as binding platform for additional subunits, including nuclear actin or actin-related proteins (ARPs) [265]. Members of the SWI/SNF family, that comprises the SWI/SNF and RSC (remodeling the structure of chromatin) complexes in *S. cerevisiae* and the BAF (BRG1/BRM-associated factors) and PBAF (polybromo-associated BAF) complexes in *H. sapiens*, are involved in nucleosome sliding and eviction [250, 266-268]. They possess a C-terminal bromo domain that interacts with acetylated H3 tails to facilitate binding to modified nucleosomes [269]. The subunit composition of human BAF complexes can vary between different cell types and mutations of BAF subunits are often observed in cancer [270, 271]. Chromatin remodelers of the INO80 family show nucleosome sliding or histone exchange activities [272-274]. Next to their HSA/post-HSA domains, they are characterized by a large insertion domain between the N- and C-lobes of the motor domain, that serves as an assembly platform for a RuvB-like protein 1 and 2 (Rvb1/2) heterohexamer [275, 276]. In *S. cerevisiae*, the INO80 family comprises the SWR1 (SWI2/SNF2-related 1) and INO80 complexes [277]. SWR1 edits nucleosome composition by exchanging the canonical H2A-H2B histone dimer with the variant dimer H2AZ-H2B [274]. While the reverse reaction has been suggested to be catalyzed by INO80 [278-280], this is still under debate [281]. In contrast to SWR1, INO80 shows nucleosome sliding and spacing activity and is involved in the positioning of +1 nucleosomes [248, 250, 273].

Over the past years, structures of nucleosome-bound chromatin remodelers from all four subfamilies elucidated how they engage NCPs and suggested possible mechanisms for DNA translocation [264, 275, 276, 282-290]. The motor domains of members of the ISWI and CHD subfamilies interact with SHL -2 [264, 285], while the motor domain of SWI/SNF member Swi2/Snf2 can bind to SHL -2 or SHL -6 [284]. The motor domain of the INO80 complex binds to SHL -6, while Arp5 forms a counter grip at SHL -2 [275]. Interestingly, SWR1 that also belongs to the INO80 subfamily of remodelers engages the NCP with its motor domain at SHL -2 and an Arp6 counter grip at SHL -6 [276]. For DNA translocation, different models have been postulated [291, 292]. The twist diffusion model describes a mechanism in which the remodeler motor domain pumps DNA in 1-2 bp steps into the nucleosome, creating a local over- or underwinding defect of the nucleosomal DNA that subsequently propagates around the nucleosome [283, 291]. The stepwise diffusion of the twist defect may allow to overcome the high energy barrier, that is posed by the combined histone-DNA interactions, by breaking it down into small increments. The bulge propagation model assumes the formation of a DNA loop between the motor domain and the counter grip by successive DNA translocation steps [278, 291]. The bulging of the DNA would distort the DNA-histone contacts and once a critical size is reached, the bulge may propagate around the nucleosome in a wave like motion. In support of that, DNA translocation step sizes of 10-15 bp were observed for the INO80 complex [278].

1.7.3. Structure and function of the INO80 complex

The INO80 complex plays a regulatory role in various DNA-dependent processes, such as transcription, DNA replication and DNA damage repair [293, 294]. INO80 expression correlates with the up- or downregulation of certain transcripts, suggesting a general role in transcriptional regulation [294-297]. In *S. cerevisiae,* INO80 was enriched at the NFRs and +1 nucleosomes of more than 90% of all genes [298], which is in line with its ability to position +1 nucleosomes in vitro [248, 250]. In fact, INO80 is the only remodeler known to establish a proper nucleosome architecture at promoters on its own, including positioning of +1 and -1 nucleosomes and formation of NFRs and genic nucleosomal arrays [250]. Positioning of +1 nucleosomes at *S. cerevisiae* promoter sequences is facilitated by a DNA shape/mechanics readout of nucleosomal and linker DNA by the INO80 complex [248]. Furthermore, INO80 regulates transcription of non-coding transcripts, as disruption of INO80 causes increased transcription of long non-coding RNAs [299]. The ability of INO80 to exchange histone variant dimer H2AZ-H2B for the canonical H2A-H2B dimer might contribute to transcriptional regulation [278-280]; however, whether INO80 really possesses histone dimer exchange function is still under debate [281]. In addition, INO80 plays a role in the restart of stalled replication forks, and might facilitate progression of the replication fork through chromatin [300-302]. When the replisome collides with the transcription machinery, INO80 promotes the release of RNA polymerase II, which is subsequently targeted for proteasomal degradation [303, 304]. INO80 is also enriched at sites of DNA double-strand breaks, where it co-localizes with phosphorylated H2AX [295, 305] and is associated with the exchange of histone variant dimer H2AZ-H2B for the canonical H2A-H2B dimer [306, 307].

The INO80 chromatin remodeling complex can be divided into three modules: the N-module, A-module, and C-module [308, 309]. The Ino80 protein that harbors the Snf2-type motor domain forms a scaffold for the assembly of the three modules. The A- and C-modules represent the conserved part of the complex and are essential for its nucleosome sliding activity, whereas the species-specific N-module is less conserved and is not required for nucleosome sliding [308-310].

In *S. cerevisiae*, INO80 forms a megadalton complex consisting of 15 different subunits, seven of which are able to bind ATP (Figure 13A) [294, 309]. The C-module comprises the Ino80 ATPase, Ino eighty subunit 2 (Ies2), Ies6, actin-related protein 5 (Arp5), and RuvB-like protein 1 and 2 (Rvb1/2) [308, 309]. Cryo-EM structures of the *Chaetomium thermophilum* and *Homo sapiens* nucleosome-bound INO80 C-modules gave insight into the architecture of the conserved core complex (Figure 13B) [275, 282]. The Rvb1/2 heterohexamer that forms around the Ino80 insertion domain forms an assembly scaffold for the C-module and may act as “stator” in the context of DNA translocation [275]. The Ino80 ATPase engages the NCP at SHL -6, while Arp5 and Ies6 form a counter grip at SHL -2/-3 [275]. Nucleosomal DNA binding by the Ino80 motor domain leads to unwrapping of approximately 15 bp of DNA and causes minor groove widening [275]. Ies2 interacts with the Ino80 motor domain, binds the NCP with its “throttle helix” at SHL +2 and interacts with the distal acidic patch [275]. On the opposite side, Arp5 engages the proximal acidic patch via its “grappler” element [275].

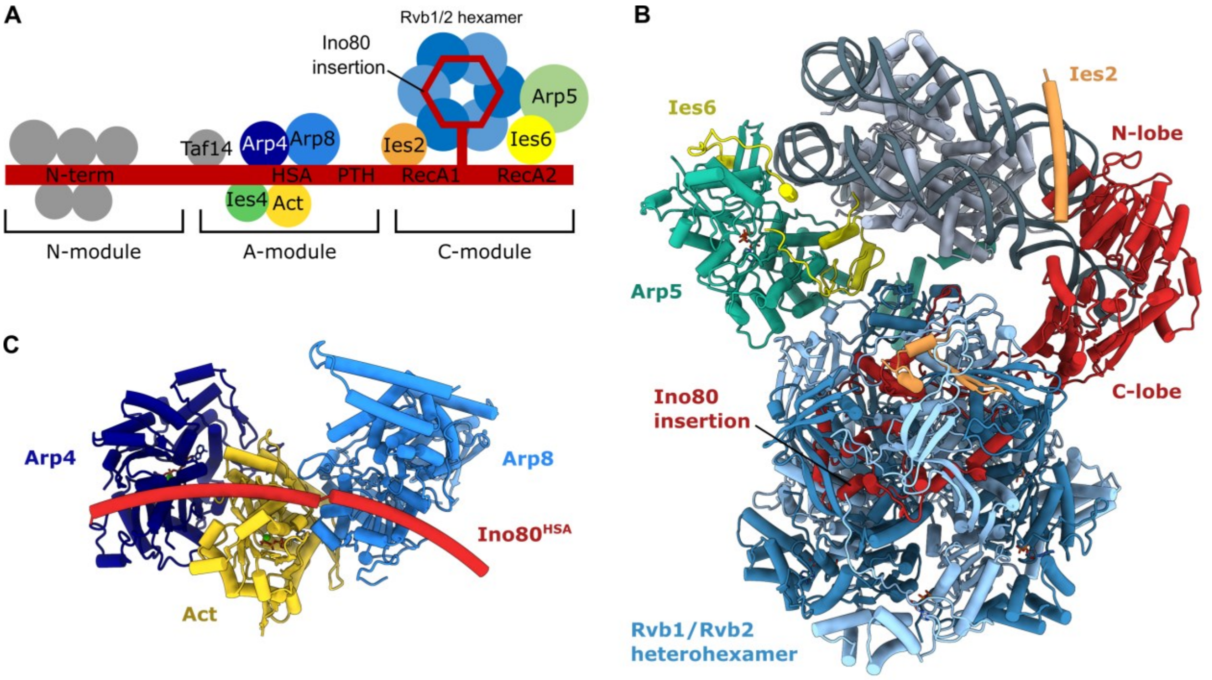


Figure 13: Structure of the INO80 complex. (A) Schematic view of the modular architecture of the *S. cerevisiae* INO80 complex. The N-, A-, and C-modules are indicated and the subunits of the conserved A- and C-modules are labeled and color-coded (Ino80: red, Arp5: green; Ies6: light yellow, Ies2: orange, Rvb1/Rvb2: blue, Arp4: dark blue, Act: yellow, Arp8: bright blue, Ies4: lime green, Taf14: grey). Several domains of the Ino80 protein (red) are indicated including the N-terminus (N-term), the HSA and post-HSA (PTH) domains, the insertion domain and the ATPase RecA1 (N-lobe) and RecA2 (C-lobe). Figure based on [311]. (B) Structure of the *C. thermophilum* INO80 C-module-nucleosome complex (PDB: 6FML) [275]. The INO80 subunits are labeled and colored as in (A). The nucleosome is colored in grey (DNA: dark grey, histones: light grey). (C) Structure of the *S. cerevisiae* INO80 A-module (PDB: 5NBN) [311]. The INO80 subunits are labeled and colored as in (A).

The INO80 A-module comprises the Ino80 HSA domain, actin, the actin-related proteins Arp4 and Arp8, and species-specific subunits like Ies4 and TBP associated factor 14 (Taf14) in *S. cerevisiae* or yin yang 1 (YY1) in *H. sapiens* [308, 309]. A crystal structure of the conserved part of the *S. cerevisiae* A-module revealed that Arp8, actin and Arp4 assemble along the helical Ino80 HSA domain (Figure 13C) [311]. A low-resolution cryo-EM reconstruction as well as biochemical data indicate that the A-module binds to approximately 40 bp of extranucleosomal DNA near the Ino80 motor domain [275, 311, 312]. Mutations of the Ino80 HSA domain that reduce the affinity of the A-module to DNA result in strongly reduced nucleosome sliding without losing nucleosome binding and ATPase activity, indicating that the A-module is important for productive nucleosome remodeling [311]. It has been suggested that the A-module functions as a molecular ruler that governs nucleosome positioning with regard to barrier factors such as the general regulatory factor Reb1 and nucleosome density-dependent spacing (Figure 14) [313]. Furthermore, the A-module is important for +1 nucleosome positioning, possibly by reading out DNA shape/mechanics features of extranucleosomal DNA [248, 250]. Together, this suggests that the A-module functions as an allosteric regulator of the INO80 nucleosome remodeling activity [311, 312]. Actin or actin-related proteins are not only part of the INO80 complex, but are integral components of INO80 and SWI/SNF family remodelers in general, highlighting their evolutionary conservation [265, 314]. *S. cerevisiae* SWR1 for example also harbors actin and Arp4 [274, 315], while *S. cerevisiae* SWI/SNF and RSC harbor an Arp7-Arp9 pair [286, 316, 317]. All A-modules assemble on a HSA domain and connect to the respective Snf2-type motor domain via a post-HSA domain [265]. In the case of RSC, it has been shown that its nucleosome sliding and ejection activity is regulated by an interaction of the post-HSA domain with protrusion I, a conserved part of the motor domain [318, 319].

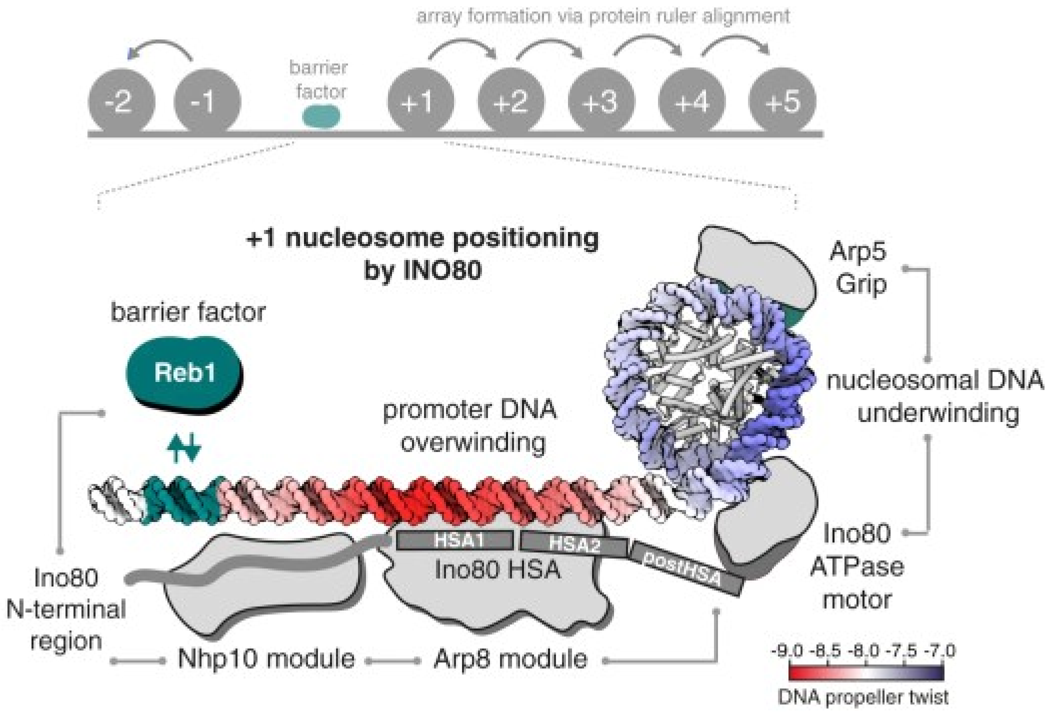


Figure 14: Model of +1 nucleosome positioning by the *S. cerevisiae* INO80 complex. INO80 reads-out and utilizes genomic information such as DNA shape features (DNA propeller twist) and the position of barrier factors such as Reb1 to position +1 nucleosomes in *S. cerevisiae*. The N-module (Nhp10 module) and A-module (Arp8 module) bind to extranucleosomal DNA, while the Ino80 motor and Arp5 grip of the C-module bind to nucleosomal DNA. INO80 positioned nucleosomes show on average underwinding (blue) of nucleosomal DNA and overwinding (red) of extranucleosomal DNA, which may correlate with A-module binding. Figure taken and modified from [248].

The species-specific N-module of the *S. cerevisiae* INO80 complex comprises the N-terminal portion of Ino80, non-histone protein 10 (Nhp10), Ies1, Ies3 and Ies5 [294, 309]. While the N-module is dispensable for nucleosome sliding [272, 309, 310], it may have a regulatory role by sensing the length of extranucleosomal DNA, as the N-module inhibits nucleosome sliding in a DNA length-dependent and switch-like manner [320]. The N-module has a high affinity for dsDNA, and may be involved in the recruitment of INO80 to promoter sites and sites of DNA damage [272, 298, 305, 321]. For the recruitment of the N-module to DNA, the Nhp10 subunit might be essential as it harbors a high mobility group (HMG) box domain, which is a known DNA binding domain [321].

1.8. Objectives

Members of the Slfn protein family are induced by interferons and play roles in inhibition of viral replication [6], cell cycle control [28, 94], T cell quiescence [32], and DNA damage response [71]. Studies on human and mouse Slfn members paint a picture of a diverse protein family with little conservation between species. This complicates the study of Slfn proteins as suitable model organisms are missing. The human subgroup III Slfn members SLFN5 and SLFN11 comprise a three-domain architecture including a helicase and a putative nuclease domain [3]. SLFN5 is involved in tumorigenesis and acts as a transcriptional regulator [11, 62]. Additionally, it shows antiviral activity by targeting transcription of viral genes [9]. SLFN11 targets translation of viral proteins in a codon-usage-dependent manner [6] and sensitizes cancer cells to DNA damaging agents by irreversibly blocking stalled replication forks [71]. While the importance of SLFN5 and SLFN11 in the mentioned processes is well established, little is known about the underlying molecular mechanisms.

Characterization of the nucleic acid binding specificity of SLFN5 and SLFN11 can shed light on their role in transcriptional regulation and DNA replication quality control, respectively. As structural data is not available on SLFN5 or SLFN11, their structural characterization can offer insights into their molecular mechanisms and facilitate the design of structure-guided mutations. It has been shown that the SLFN11 tRNase activity contributes to sensitizing cells to DDA treatment and presumably underlies its antiviral activity; however, the mechanism is unclear [6, 45]. Mechanistic insights into the SLFN11 nuclease activity can deepen the understanding of this critical enzymatic activity with implications for other RNase proficient Slfn proteins. This work addresses these questions by investigating SLFN5 and SLFN11 by a combination of structural, biochemical and biophysical methods.

The second part of this work addresses the characterization of the INO80 chromatin remodeling complex with a focus on its interaction with extranucleosomal DNA. INO80 has nucleosome sliding activity and is involved in the establishment of a proper nucleosome architecture at active promoters, including positioning of +1 and -1 nucleosomes and formation of NFRs [250]. Within the group of ATP-dependent chromatin remodelers, INO80 has the specific characteristic of engaging the NCP at SHL -6 with its Snf2-type motor domain and to interact with extranucleosomal DNA via the A-module [275, 311, 312]. The A-module that comprises actin and the actin-related proteins Arp4 and Arp8 serves as an allosteric regulator of the Ino80 motor domain and processes genomic information such as DNA shape features [248, 311]. Furthermore, the A-module acts as a molecular ruler that helps to position nucleosomes with respect to barrier factors or DNA ends [313]. While structural and biochemical studies have characterized the interaction of INO80 with the nucleosome core particle [275, 282], a comprehensive structural model for the regulation of INO80 by extranucleosomal DNA is missing.

Structural analysis of the A-module with and without DNA can help to identify conserved structural features and elucidate its DNA binding mode. The effect of structure-guided A-module mutants on DNA binding and the enzymatic activities of INO80 can shed light on the regulatory function of the A-module. Extending the scope to study the INO80 A/C-complex bound to a nucleosome can give insights into the mode of extranucleosomal DNA interaction by the A-module and additional regulatory features. Answering these questions will help to understand the complex regulation of the INO80 chromatin remodeler.

2. Publications

2.1. Structural and biochemical characterization of human Schlafen 5

Felix J. Metzner*, Elisabeth Huber*, Karl-Peter Hopfner, Katja Lammens (2022) Structural and biochemical characterization of human Schlafen 5. *Nucleic Acids Research*, Volume 50, Issue 2, pp. 1147-1161.

* These authors contributed equally.

DOI: 10.1093/nar/gkab1278
URL: https://academic.oup.com/nar/article/50/2/1147/6509097

Summary

This publication describes the first structure of a full-length subgroup III Slfn protein. The cryo-EM reconstruction of human SLFN5 reveals the structural arrangement of the Slfn core, linker and helicase domains. The horseshoe-shaped Slfn core domain exhibits a stable hydrophobic interface with the linker domain, which interacts with the helicase N-lobe. A high-resolution crystal structure of the SLFN5 core domain reveals the presence of a conserved zinc finger motif and together with biochemical data, shows that the Slfn core domain does not resemble an ATPase-like fold, contrary to previous sequence-based predictions. The helicase domain, of which the N-lobe is resolved in the cryo-EM reconstruction, harbors Walker A and B motifs and binds ATP. However, DNA or RNA do not stimulate ATP hydrolysis by SLFN5. Investigation of nucleic acid binding shows that the SLFN5 core domain, as well as the full-length protein, bind to dsDNA and tRNA. Single point mutants of positively charged residues within the Slfn core domain decrease its affinity to dsDNA and tRNA. In line with a previous study that demonstrated endoribonuclease activity towards tRNA for the rSlfn13 core domain, we observe endoribonuclease activity for the mSlfn8 core domain and full-length SLFN11 towards type II tRNA^{Ser}. In contrast, SLFN5 shows no endonucleolytic activity towards tRNA or DNA. Overall, we present a comprehensive structural and biochemical analysis of SLFN5 and discuss the divergent enzymatic functions within the subgroup III Slfn family.

Author contribution

I established the expression of full-length SLFN5 and SLFN11 in HEK293 cells and developed purification strategies for full-length SLFN5, SLFN11 and mSlfn8^{N}. I expressed and purified SLFN5, SLFN11, mSlfn2 and mSlfn8^{N}. I performed the biochemical characterization of full-length SLFN5, including DNA and tRNA binding studies, ATP binding and hydrolysis assays and nuclease assays. Together with Katja Lammens, I prepared grids for cryo-EM. Furthermore, Katja Lammens and I performed cryo-EM data collection, processing and model building of SLFN5. I prepared the manuscript together with Elisabeth Huber and Katja Lammens.

Nucleic Acids Research, 2022, Vol. 50, No. 2 **1147–1161**
https://doi.org/10.1093/nar/gkab1278

Structural and biochemical characterization of human Schlafen 5

Felix J. Metzner[†], Elisabeth Huber[†], Karl-Peter Hopfner and Katja Lammens[*]

Department of Biochemistry, Gene Center, Feodor-Lynen-Straße 25, 81377 München, Germany

Received September 21, 2021; Revised December 08, 2021; Editorial Decision December 09, 2021; Accepted December 14, 2021

ABSTRACT

The Schlafen family belongs to the interferon-stimulated genes and its members are involved in cell cycle regulation, T cell quiescence, inhibition of viral replication, DNA-repair and tRNA processing. Here, we present the cryo-EM structure of full-length human Schlafen 5 (SLFN5) and the high-resolution crystal structure of the highly conserved N-terminal core domain. We show that the core domain does not resemble an ATPase-like fold and neither binds nor hydrolyzes ATP. SLFN5 binds tRNA as well as single- and double-stranded DNA, suggesting a potential role in transcriptional regulation. Unlike rat Slfn13 or human SLFN11, human SLFN5 did not cleave tRNA. Based on the structure, we identified two residues in proximity to the zinc finger motif that decreased DNA binding when mutated. These results indicate that Schlafen proteins have divergent enzymatic functions and provide a structural platform for future biochemical and genetic studies.

INTRODUCTION

The Schlafen (Slfn) family (1,2) belongs to the class of interferon-stimulated genes (ISGs) (3,4). Ten murine and six human Schlafen proteins have been identified, which can be categorized into three subgroups according to their size and domain architecture (1,2,4). The Schlafen proteins play roles in various cellular processes such as regulation of cell cycle (5,6), T cell quiescence (7–11), differentiation and proliferation (3,12–14), tumorigenesis (3,13,15), response to DNA damaging agents (16–23) and inhibition of viral replication (24–26).

All Schlafen family members share a highly conserved N-terminal core region of approximately 340 amino acids (1,2). Based on sequence similarity, this region has been predicted to contain a divergent AAA ATPase associated domain (27). In this work, the N-terminal region will be referred to as Schlafen core domain. While subgroup I Schlafen proteins consist of the Schlafen core domain only, subgroups II and III harbor a C-terminal linker domain of unknown function. In addition, subgroup III Schlafen proteins, such as human SLFN5 and SLFN11, possess an additional C-terminal domain with sequence homology to the family of SF1 DNA/RNA helicases (2). Subgroup I and II members are predominantly located in the cytoplasm, while subgroup III members were mainly detected in the nucleus (28).

The Schlafen protein family was initially described as a regulator in thymocyte maturation (1,2) and T cell quiescence in mice (7). The *elektra* mouse, which carries a point mutation in the *mSlfn2* gene, is characterized by immune deficiency and susceptibility to bacterial and viral infections. Several studies indicated that mSlfn2 might maintain the quiescent state of T-cells by promoting the expression of 'quiescence' genes and inactivation of genes required for proliferation or differentiation (7,10,29). Additionally, Fischietti *et al.* suggested that mSlfn2 plays a role in the transcriptional regulation of ISGs via the balancing of type I IFN-mediated activation of STAT1 and NF-κB (30).

The human Schlafen family member SLFN11 came into focus due to its ability to promote cancer cell death in response to DNA-damaging agents (16–17,20). The N-terminal domain of SLFN11 was shown to specifically cleave type II tRNAs (31). This leads to translational downregulation of a range of proteins such as Ataxia telangiectasia and Rad3-related (ATR), resulting in inhibition of the DNA damage-repair pathway (31). In addition, it sensitizes cancer cells to DNA targeting therapies by blocking replication in response to DNA damage sites (32). Chromatin opening and replication stalling seem to be independent of the downregulation of ATR (22), but depend on the C-terminal helicase domain (32). SLFN11 has been described as a promising biomarker in different types of cancer (16–17,33–39). Expression levels of SLFN11 can help to predict the response to a wide range of DNA-damaging anti-cancer agents across multiple cancer types. Hence, it was proposed that SLFN11 could have clinical applications for matching patients to DNA-damaging chemotherapies. More recently, SLFN11 was also reported to destabilize stalled replication forks (40) and was found to play a role as a regulator of protein quality control (41).

---

[*] To whom correspondence should be addressed. Tel: +49 89 2180 76987; Email: klammens@genzentrum.lmu.de

[†] The authors wish it to be known that, in their opinion, the first two authors should be regarded as Joint First Authors.

Furthermore, SLFN11 specifically abrogates the replication of HIV by selectively inhibiting viral protein synthesis in HIV infected cells in a codon-usage dependent manner (24). Inhibition of viral protein translation is achieved by tRNA binding, which counteracts the virus-induced shift of the tRNA pool towards A/U (24). Furthermore, endoribonuclease activity has been observed for several Schlafen family members (42–44). C-terminally truncated rabbit and human SLFN14 (residues 1–400) were shown to bind ribosomes and cleave rRNA (43,44). Other studies showed cleavage of tRNA and rRNA by rat Slfn13 (rSlfn13), human SLFN13, mouse mSlfn8 and human SLFN12, suggesting a role in translational regulation (42,45). Recently, Garvie *et al.* presented that a tetrameric complex of two phosphodiesterases PDE3A and two SLFN12 molecules lead to a cytotoxic response in cancer cells (45). The small molecule DNMDP stabilized binding to PDE3A and increased the RNase activity of SLFN12, which was important for its cytotoxic function.

Although human SLFN5 is linked to tumorigenesis and is being investigated as a biomarker, its molecular functions are poorly characterized. Depending on the tumor type, SLFN5 can have inhibitory (4,46–47) or stimulatory (48,49) effects on tumorigenesis. A recent study showed an inhibitory effect on the transcription of the transcription factor ZEB1, which leads to the inhibition of the AKT signaling pathway in BRCA cells (47). The suppression is mediated by direct binding of SLFN5 to the promoter DNA of ZEB1 and requires the proteins C-terminal domain (50). Furthermore, SLFN5 suppresses cancer cell migration by inhibiting expression of the membrane-type metalloprotease MT1-MMP (51). Since SLFN5 interacts with STAT1 and functions not only as an ISG but also as a repressor of ISG transcription, the existence of a negative-feedback regulatory loop is speculated (48).

SLFN5 was shown to have antiviral activity upon infection with herpes simplex virus 1 (HSV-1). It represses HSV-1 transcription by binding to viral DNA and in turn, preventing RNA polymerase II from accessing viral promoters. However, in the presence of the viral E3 ubiquitin ligase ICP0, SLFN5 is ubiquitinated and subject to proteasomal degradation (26).

To understand the molecular mechanism of human SLFN5, we determined the cryo-EM structure of full-length human SLFN5 and the crystal structure of the SLFN5 core domain. The crystal structure, at a resolution of 1.8 Å, revealed a highly conserved zinc finger motif. We confirmed binding to various nucleic acid substrates and identified residues involved in nucleic acid binding in proximity to this motif. In contrast to sequence-based predictions, the SLFN5 core domain does not resemble an ATPase like fold and neither binds nor hydrolysis ATP. Despite the partial sequence conservation of the active site residues identified in rSlfn13, we did not observe ribonuclease activity for human SLFN5 towards type II tRNAs or DNA. Further, we compared the ribonuclease activity of mouse mSlfn2, mSlfn8 and human SLFN11 and structurally discussed the mSlfn2 *elektra* mutation. The cryo-EM map of full-length SLFN5 gives insights into the domain architecture and domain interfaces of group III Schlafen proteins. The full-length protein shows a high affinity to double-stranded DNA and binds ATP. Overall, we present a comprehensive structural and biochemical analysis of SLFN5 and group III Schlafen proteins and discuss similarities and differences throughout this diverse family.

MATERIALS AND METHODS

Protein expression and purification

The gene constructs encoding the N-terminal domain M1-D336 of human SLFN5, the N-terminal domain M1-D351 of murine mSlfn8 or full-length murine mSlfn2 were cloned into pET21a vector (Novagen) using NdeI/XhoI (Thermo) restriction enzymes. Site-directed mutagenesis for the generation of the SLFN5$^{1-336}$ R271E and R326E and the mSlfn2 I135N mutations were performed using the Quickchange (Stratagene) protocol. *E. coli* Rosetta (DE3) cells containing the plasmids were grown at 37°C to an OD_{600} of 0.8. Protein expression was induced with 0.2 mM IPTG and cells were incubated over night at 18°C and harvested by centrifugation.

The cells expressing C-terminally His$_6$-tagged SLFN5$^{1-336}$ were resuspended in lysis buffer (50 mM HEPES, pH 8.0, 200 mM NaCl, 7 mM imidazole, 2 mM $MgCl_2$, 4 mM β-mercaptoethanol) supplemented with a protease inhibitor cocktail (0.176 g/l phenylmethylsulfonyl fluoride, 0.316 g/l benzamidine hydrochloride, 1.372 mg/l pepstatin, 0.256 mg/l leupeptin, 0.2 mg/l chymostatin). The cells were disrupted by sonication and the insoluble cell debris was separated from the supernatant by centrifugation at 30 000 × g at 4°C for 30 min. The supernatant was applied onto a Ni-NTA column (Qiagen) and extensively washed with lysis and washing buffer (lysis buffer supplemented with 31 mM imidazole). The protein was eluted by applying elution buffer (50 mM HEPES, pH 8.0, 200 mM NaCl, 250 mM imidazole, 2 mM $MgCl_2$, 4 mM β-mercaptoethanol). Further purification was performed by negative anion exchange chromatography (Q-HP, GE Healthcare), where SLFN5$^{1-336}$ remained in the flow-through, and a subsequent Superdex 200 size-exclusion chromatography (GE Healthcare) by using the following buffer condition: 50 mM HEPES, pH 8.0, 200 mM NaCl, 2 mM $MgCl_2$, 4 mM β-mercaptoethanol. The expression and purification of the SLFN5$^{1-336}$ mutants were performed accordingly.

C-terminally His$_6$-tagged mSlfn2 and mSlfn2 I135N were expressed and purified accordingly but the pH of the buffers was set to pH 7.5.

Cells expressing C-terminally His$_6$-tagged murine Slfn8$^{1-351}$ were harvested by centrifugation, resuspended in lysis buffer (50 mM Tris pH 8.2, 300 mM NaCl, 2 mM $MgCl_2$, 10 mM imidazole) supplemented with protease inhibitor cocktail and disrupted by sonication. The lysate was cleared by centrifugation at 30 000 × g at 4°C for 30 minutes and the supernatant was incubated with pre-equilibrated Ni-NTA resin (Qiagen). The Ni-NTA resin was washed (50 mM Tris pH 8.2, 300 mM NaCl, 2 mM $MgCl_2$, 15 mM imidazole) and the His$_6$-tagged protein was eluted by applying 50 mM Tris pH 8.2, 300 mM NaCl, 2 mM $MgCl_2$, 250 mM imidazole. The elution fractions were dialyzed against 20 mM Tris pH 8.2, 50 mM NaCl, 2 mM $MgCl_2$, 0.5 mM DTT overnight. The dialyzed protein was further purified using anion exchange chromatography (HiTrap Q HP, GE Healthcare) and heparin chromatography (HiTrap Heparin HP, GE Healthcare). The protein was eluted by applying a linear gradient from 50 mM NaCl to 1 M NaCl in 20 mM Tris pH 8.2, 2 mM $MgCl_2$, 0.5 mM DTT, respectively. For further purification, mSlfn8$^{1-351}$ was applied onto a Superdex 200 16/60 column (GE Healthcare), pre-equilibrated with 20 mM Tris pH 8.2, 200 mM NaCl, 2 mM $MgCl_2$, 0.5 mM DTT.

Prior to SEC and freezing, the proteins were concentrated with centrifugal concentrators (Amicon® Ultra Centrifugal Filters, Merck). The purified proteins were flash frozen in liquid nitrogen and stored at -80°C until further use.

Crystallization and structure determination of SLFN5$^{1-336}$

1 μl of SLFN5$^{1-336}$ purified from Rosetta cells and 1 μl of the reservoir solution were mixed and crystals were obtained by the hanging drop vapor diffusion method. Cubic crystals formed at a protein concentration of 7 mg/ml and 200 mM NaCl, 100 mM MES pH 5.8, 20% (v/v) PEG 2000 MME as reservoir conditions. Reservoir solution supplemented with 25% (v/v) ethylene glycol was used as cryo-protectant prior to flash freezing in liquid nitrogen. The concentration of SLFN5$^{1-336}$ was set to 3.2 mg/ml for the needle shaped crystals and 0.1 M sodium acetate pH 5.0 and 1.5 M ammonium sulfate were used as reservoir. Reservoir solution supplemented with 25% (v/v) glycerol was used as cryo-protectant prior to flash freezing in liquid nitrogen.

Diffraction data were collected at the beamlines X06SA (PXI) and X06DA (PXIII) (Swiss Light Source, Paul-Scherrer Institute, Villigen, Switzerland) at 100 K. Data were integrated and scaled with XDS (52,53). Experimental phases were determined using a 3.2 Å Zn-SAD dataset from the needle shaped crystals measured at 1.28 Å wavelength. For the generation of the Zn^{2+}-substructure and a poly-alanine model HySS (54), AutoSol (55) and Autobuild (56) within the Phenix (57) software and Chainsaw (58) within the CCP4 package (59) were used. The poly-alanine model was used as model for molecular replacement phasing of the 1.8 Å native dataset from the cubic crystals. This was done using Phaser (60,61) and the initial model was automatically rebuild with Autobuild (56). The final structure of SLFN5$^{1-336}$ was solved at 1.8 Å by iterative refinement cycles in PHENIX (62) or Refmac (63,64). The structure was manually completed with COOT (65). Prior to model building and refinement, we randomly omitted 5% of the reflections for monitoring the free R value. Data collection and model statistics are stated in Supplementary Table S2. All figures were prepared using PyMOL Molecular Graphic Systems (version 2.0, Schrödinger, LLC) or UCSF ChimeraX (66).

Right-angle light scattering measurement

The molecular weight of SLFN5$^{1-336}$ was determined by size-exclusion chromatography (SEC)-coupled right-angle light scattering. The experiment was performed using a Superdex 200 10/300 Increase column (GE Healthcare), coupled to a right-angle laser static light scattering device and refractive index detector (Malvern/Viscotek). BSA was used to calibrate the system and the evaluation was performed using the OmniSEC software (Malvern/Viscotek).

Nucleic acid substrates

DNA and RNA oligonucleotides were purchased from Metabion (Planegg, Germany) and Biomers (Ulm, Germany), respectively.

For the generation of double-stranded nucleic acid substrates, the single strands were mixed in an equimolar ratio, heated to 95°C for 10 min and slowly cooled down to room temperature. The nucleic acid substrates used in this work are summarized in Supplementary Table S3.

Affinity measurement by fluorescence anisotropy

Initial protein dilutions (0, 1, 2, 3, 5, 10, 15, 20, 30, 40 and 60 μM if not stated otherwise) of SLFN5$^{1-336}$ wild type and mutants were prepared in 2× assay buffer (100 mM HEPES pH 8.0, 100 mM NaCl) and then mixed with 6-FAM labeled DNA or RNA (at a final concentration of 100 nM) in a 1:1 (v/v) ratio. The reaction was incubated on ice for 30 min and the fluorescence anisotropy was subsequently measured at an excitation wavelength of 470 nm and an emission wavelength of 520 nm. The data sets were analyzed with Prism (GraphPad Software) and fit to a Hill model.

[γ-^{32}P] ATP hydrolysis assay

For the ATPase assay 5 μM SLFN5$^{1-336}$ was incubated in presence or absence of 0.2 μM single-stranded 60-mer poly (dT) DNA in 50 mM Tris pH 7.5, 150 mM KCl, 5 mM $MgCl_2$, 1.5 mM ATP and 10 nM [γ-^{32}P] ATP (Hartmann Analytik, Germany) at 37°C for 0 or 60 min. For analysis, 1 μl of the reaction mixture was applied onto polyethyleneimine cellulose TLC plates (Sigma-Aldrich, Germany) and free phosphate was separated from ATP by thin layer chromatography in TLC running buffer (1 M formic acid, 0.5 M LiCl). [γ-^{32}P] ATP was detected using a Typhoon FLA 9000 imaging system (GE healthcare).

SLFN5^{FL} expression and purification

A construct encoding for full-length SLFN5 with an N-terminal double FLAG-tag and a HRV 3C cleavage site was purchased from GenScript. The construct was inserted into a pcDNA3.1 vector using a codon-optimized sequence for human expression systems. Expi293F cells (Thermo Fisher Scientific) were transfected with the SLFN5 expression vector using polyethylenimine (PEI, MW 40 000, Polysciences). Cells were cultured in Expi293 Expression Medium (Thermo Fisher Scientific) at 37°C and 5% CO_2. After 72 h, the cells were harvested by centrifugation, resuspended in lysis buffer (50 mM Tris pH 7.1, 400 mM NaCl, 2 mM $MgCl_2$) supplemented with protease inhibitor (0.18 g/l PMSF, 0.32 g/l benzamidine, 1.37 mg/l pepstatin A, 0.26 mg/l leupeptin, 0.2 mg/l chymostatin) and disrupted by sonication. The lysate was cleared by centrifugation at 30 000 × g at 4°C for 45 min and the supernatant was incubated with pre-equilibrated ANTI-FLAG

M2 Affinity Gel (Sigma-Aldirch) for 60 min. The resin was washed with wash buffer (25 mM Tris pH 7.1, 250 mM NaCl, 2 mM $MgCl_2$) and wash buffer supplemented with 1 mM ATP. After washing with buffer A (25 mM Tris pH 7.1, 120 mM NaCl, 2 mM $MgCl_2$, 1 mM DTT), the protein was eluted in 4 × 1.1 ml elution buffer (buffer A supplemented with 0.2 mg/ml Flag-peptide) over 60 min. The eluate was loaded onto a HiTrap Heparin HP column (GE Healthcare) and the protein was eluted by a linear salt gradient (100 % buffer A to 100 % buffer B (25 mM Tris pH 7.1, 1 M NaCl, 2 mM $MgCl_2$, 1 mM DTT) over 12 CV). The peak fractions were combined and applied onto a Superdex 200 5/150 column (GE Healthcare), pre-equilibrated with buffer A. The peak fractions were combined and flash frozen in liquid nitrogen. FLAG-tagged SLFN11 was purified following a similar protocol, with the difference, that the pH of the buffers was adjusted to pH 7.5.

Cryo-EM grid preparation

Freshly purified SLFN5 was diluted to a final concentration of 3 μM using cryo-EM buffer (50 mM glycine pH 9, 50 mM NaCl, 2.5 mM $MgCl_2$, 1 mM DTT). *n*-octyl-β-D-glucoside was added (0.045 %) and 4.5 μl of the sample was applied onto a glow discharged UltrAuFoil® R2/2 holy gold grid. The sample was flash-frozen in liquid ethane, using an EM GP plunge freezer (Leica, 10°C and 90 % humidity).

Cryo-EM data collection

The datasets were collected using a FEI Titan Krios G3 transmission electron microscope (300 kV), equipped with a GIF quantum energy filter (slit width 20 eV) and a Gatan K2 Summit direct electron detector. The data was automatically acquired using EPU (FEI). Three datasets were collected with 1765 (dataset I), 550 (dataset II) and 798 (dataset III) movies. Datasets II and III were collected at a tilt angle of 25°. All datasets were collected with a pixel size of 1.046 Å and 40 frames over 8 s. Dataset I and II were collected with a total electron dose of 41.2 $e^-/Å^2$ and dataset III with 40.9 $e^-/Å^2$. Defocus values ranging from –1.1 to –2.9 μm were applied.

Cryo-EM data processing and 3D reconstruction

Motion correction of the movie frames was done using MotionCor2 (67). Unless stated otherwise, all subsequent processing steps were performed in cryoSPARC v3.2.0 (68) and the resolutions reported here are calculated based on the gold-standard Fourier shell correlation criterion (FSC = 0.143). The CTF parameters of the three datasets were determined using patch CTF estimation (multi) in cryoSPARC (v3.2.0). The exact processing scheme is depicted in Supplementary Figure S2. The data collection and refinement statistics are summarized in Supplementary Table S1. Initial particle picking was done using Blob picker on dataset III, yielding 607 362 particles, which were extracted with a box size of 256 px and a pixel size of 1.046 Å/px. The particles were subject to 2D classification and classes with clearly defined features were selected (43 089 particles). The selected particles were used as input for a Topaz train job, followed by particle extraction and 2D classification. The classes with clearly defined features yielded 102 102 particles, which were used as input for another round of Topaz train. The resulting Topaz model was used to pick on all three datasets. 592 223 particles were extracted from dataset I, 302 247 particles from dataset II and 344 426 particles from dataset III. Each particle set was subject to one round of 2D classification and ab-initio reconstruction. The resulting 192 301 particles from dataset I, 69 376 particles from dataset II and 140 963 particles from dataset III were combined (402 640 particles) and subject to another round of 2D classification and ab-initio reconstruction with five different classes. Three classes with 293 165 particles were selected and subject to heterogeneous refinement with four classes. The *ab-initio* reconstructions were used as input volumes for the heterogeneous refinement job. The class that showed the most defined features was selected (140 715 particles) and used for further refinement. The final resolution of the reconstruction after non-uniform refinement (69) was 3.44 Å.

Model building

The $SLFN5^{1-336}$ crystal structure as well as the AlphaFold v2.0 model (70) of the linker domain and the ATPase N-lobe were rigid body docked into the cryo-EM density using UCSF ChimeraX (66). Model building in COOT (65) and real space refinement in PHENIX (62) were performed iteratively using the 3.44 Å map.

Purification of mononucleosomes

Canonical human histones were purchased from The Histone Source. For octamer assembly, the histones were resuspended in 7 M guanidinium chloride and mixed at a 1.2-fold excess of H2A and H2B. The mixture was dialyzed against 2 M NaCl for 16 h. The histone octamer was purified by size-exclusion chromatography using a Superdex 200 16/60 column (GE Healthcare). The Widom 601 DNA (71) with 80 bp of extranucleosomal DNA was amplified by PCR and purified by anion-exchange chromatography (Supplementary Table S3). DNA and histone octamer were mixed at a 1.1-fold excess of DNA in 2 M NaCl and diluted to 50 mM NaCl over 16 h at 4°C. Finally, the nucleosomes were purified by anion-exchange chromatography, dialyzed against 50 mM NaCl, concentrated and stored at 4°C.

Electrophoretic mobility shift assay (EMSA)

EMSAs were conducted to analyze the interaction between SLFN5 and various nucleic acid substrates. Increasing amounts of SLFN5 were titrated to fluorescently labeled substrates (6-FAM or Cy5 labelled, 40 nM) in EMSA buffer (25 mM HEPES pH 7.0, 60 mM NaCl, 5 % glycerol, 2 mM $MgCl_2$, 0.5 mM DTT) and incubated on ice for 30 minutes. The samples were analysed by native PAGE on 3–12 % acrylamide Bis-Tris gels (Invitrogen). The electrophoresis was performed in 1× NativePAGE Running Buffer (Invitrogen) at 4°C and 100 V for 120 min. The gels were visualized using a Typhoon FLA 9000 imaging system (GE healthcare). The images were analysed and integrated using GIMP v2.10.2 and ImageJ (72).

Nuclease assay

The 6-FAM labeled nucleic acid substrate (tRNA$_{Ser}$ or DNA) (50 nM) was incubated with the indicated Schlafen protein (250 nM) in nuclease buffer (25 mM Tris pH 7.3, 120 mM NaCl, 4 mM MgCl$_2$, 1 mM DTT) at 37°C for 45 min. Where indicated, MnCl$_2$ (2 mM) or EDTA (10 mM) was added. For the nuclease assay with DNA as a substrate, DNase I (Thermo Fisher Scientific) was used as positive control. The samples were analyzed on 15% denaturing polyacrylamide gels (Rotiphorese® DNA sequencing system) in 1× TBE buffer. Gels were run at 270 V for 50 min and visualized using a Typhoon FLA 9000 imaging system (GE healthcare). The images were analysed and integrated using GIMP v2.10.2.

Nano differential scanning fluorimetry (nanoDSF)

Interaction of SLFN5 with nucleotides was analyzed using nanoDSF (Tycho NT.6, NanoTemper Technologies). Full-length SLFN5 (500 nM) was incubated with or without nucleotides (1 mM) in buffer A (25 mM Tris pH 7.1, 120 mM NaCl, 2 mM MgCl$_2$, 1 mM DTT) for 15 min on ice. The samples were loaded into glass capillaries and the internal fluorescence at 330 nm and 350 nm was measured while a thermal ramp was applied. The internal Tycho NT.6 software was used for data analysis, smoothing and calculation of derivatives. For SLFN5$^{1-336}$, SLFN5$^{1-336}$ R271E and SLFN5$^{1-336}$ R326E, a protein concentration of 2 μM was used and the pH was adjusted to pH 7.5.

ATP hydrolysis assay

A fluorescence-based ATPase assay was conducted to determine the ATPase rate of SLFN5. SLFN5 (50 nM) was incubated with DNA or RNA substrates (150 nM) in assay buffer (25 mM Tris pH 7.5, 50 mM NaCl, 2 mM MgCl2, 0.1 mg/ml BSA, 1 mM DTT) at 25°C. In the assay, ATP (1 mM) hydrolysis is enzymatically coupled to the oxidation of NADH (0.1 mM) via phosphenolpyruvate (0.5 mM) by pyruvate kinase and lactate dehydrogenase (25 U/ml each, Sigma). Hexokinase from *Saccharomyces cerevisiae* was used as positive control (Sigma-Aldrich). The reaction volumes, of 50 μl each, were transferred to black non-binding 384-well plates (Greiner) and the fluorescence of NADH was measured using an Infinite M1000 PRO microplate photometer (TECAN). The reaction was monitored for 45 min (20 sec intervals) using an excitation wavelength of 340 nm and an emission wavelength of 460 nm.

RESULTS AND DISCUSSION

Cryo-EM structure of full-length SLFN5

As a subgroup III Schlafen family member (2), SLFN5 possess a tripartite domain architecture (101 kDa). The N-terminal Slfn core domain (residues 1–336) is followed by a linker domain (residues 337–552) and a C-terminal helicase/ ATPase domain (residues 553–891). To allow for the biochemical and structural characterization of the full-length protein, a purification strategy was established. A human expression system was chosen, as bacterial expression did not yield soluble protein. Purification via the N-terminal FLAG-tag followed by heparin-affinity chromatography yielded protein of high purity (Supplementary Figure S1A, C). Analytical size exclusion chromatography showed a single peak with the elution volume approximately corresponding to a monomer (Supplementary Figure S1B). To gain insights into the structural organization of subgroup III Schlafen proteins, we employed single particle cryo electron microscopy (cryo-EM). The full-length SLFN5 protein was vitrified in the absence of nucleotides. Due to an orientation bias of the particles, the data was partially acquired at a tilt angle of 25°. A 3.5 Å cryo-EM density was calculated, giving insights into the overall domain arrangement (Figure 1A, B, Supplementary Figure S2A–F, Supplementary Table S1, Movie S1). A majority of the protein could be resolved (residues 3–684) with exception of the C-terminal ATPase lobe (residues 685–891), which is likely due to its flexibility (Figure 1C). Residues 143–165 are not resolved, indicating an unordered loop. The most N-terminal loop (residues 7–13) is shifted towards a hydrophobic patch of the N-lobe (Ile57, Leu96, Phe98), resulting in hydrophobic interaction via a semi-conserved aromatic residue (Phe11) (Figure 1D). Furthermore, this loop contributes to interaction with the SLFN5 linker domain via a salt bridge between Glu13 and Lys475. The N-lobe of the Slfn core forms a second large interface with the linker domain, including several hydrophobic (Ile57, Met89, Phe98, Val479, Tyr514, Pro515, Tyr518) and ionic interactions (Asp87 to Arg487) (Figure 1D). In addition, the loop connecting the Slfn core to the linker domain (residues 335–366) holds the two domains together. The globular linker domain exhibits a mixed α/β topology and connects the Slfn core and the ATPase domain. The highly conserved SWAVDL motif (residues 424–429) seals the hydrophobic core of the linker domain and interacts with the N-terminal lobe of the helicase (Figure 1E). Density for the C-terminal ATPase lobe was not observed. The data was acquired in the absence of nucleotides, suggesting that the missing density of the second ATPase lobe could be due to relative flexibility between the lobes. A model of the C-lobe of the helicase, as calculated by AlphaFold (70), does not cause any steric clashes (Figure 1C). The N-terminal helicase lobe is anchored to the linker domain by an α-helix (residues 561–568) (Figure 1E). Further residues that are involved in the interaction are Gln432 and Arg590 of the linker domain. Apart from these charged residues, the helicase–linker domain interface is mostly hydrophobic (Figure 1E). Additionally, the loop preceding the ATPase domain, interacts with the ATPase N-lobe in close proximity to the putative ATP binding site via Phe540 and Phe543 (Figures 1E, 2B). The N-terminal ATPase lobe is highly conserved and harbors the characteristic residues of Walker A and B motifs that are essential for ATP binding and hydrolysis (Supplementary Figure S3). We superimposed the N-terminal ATPase lobe of SLFN5 with the nucleotide and ssDNA bound structure of the related SF1 helicase DNA2 (PDB: 5EAX) (Figure 2A, B). This suggests an

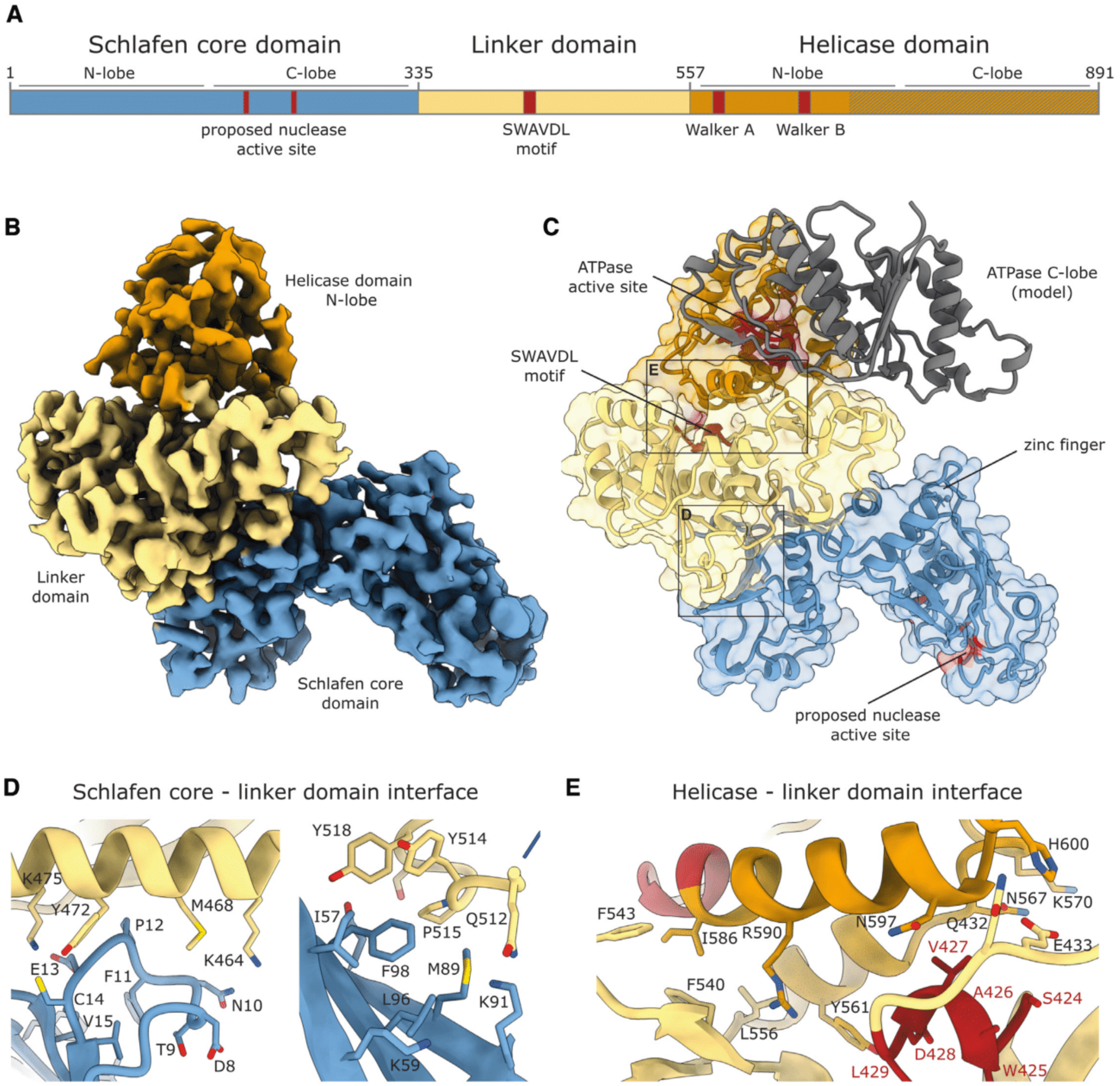


Figure 1. Structure of full-length SLFN5. (**A**) Domain architecture of Slfn5. (**B**) Cryo-EM reconstruction of full-length Slfn5. The Slfn core domain is depicted in blue, the linker domain in yellow and the helicase N-lobe in orange. (**C**) Ribbon and surface representation of Slfn5. Colored according to (A) with indicated motifs depicted in red. A model of the helicase C-lobe, as calculated by AlphaFold (70), is depicted in gray. (**D**) Detailed views of the interface between the linker and Schlafen core domain. (**E**) Detailed view of the interface between the linker and helicase domain N-lobe. The SWAVDL and Walker A motifs are colored in red.

accessible ATP binding site and sufficient space for the second ATPase lobe to bind. Assuming a similar DNA binding mode as in DNA2, the DNA binding site would be located on top of the ATPase domain (Figure 2A). The electrostatic surface potential of SLFN5 (Supplementary Figure S4) illustrates that the side of the molecule with the zinc finger and proposed DNA binding site in the helicase domain is positively charged, whereas the proposed nuclease active site would be located at the opposite side of SLFN5 with no clear surface charge potential. The distance between the positively charged patch next to the zinc finger motif and the predicted DNA binding site of the helicase domain is ~40–45 Å (Supplementary Figure S4).

In summary, we present the first structure of a full-length subgroup III Schlafen family member, giving new insights into the structural organization of this class of proteins.

Characterization of full-length SLFN5

As a subgroup III Schlafen family member, SLFN5 possess a C-terminal domain with homology to SF1 DNA/RNA helicases (2). SLFN5 harbours all characteristic motifs of SF1 DNA/RNA helicases, including the Walker A and Walker B motifs, which are involved in ATP binding and hydrolysis (Figure 2D). While the Walker A motif is highly conserved throughout subgroup III Schlafen family mem-

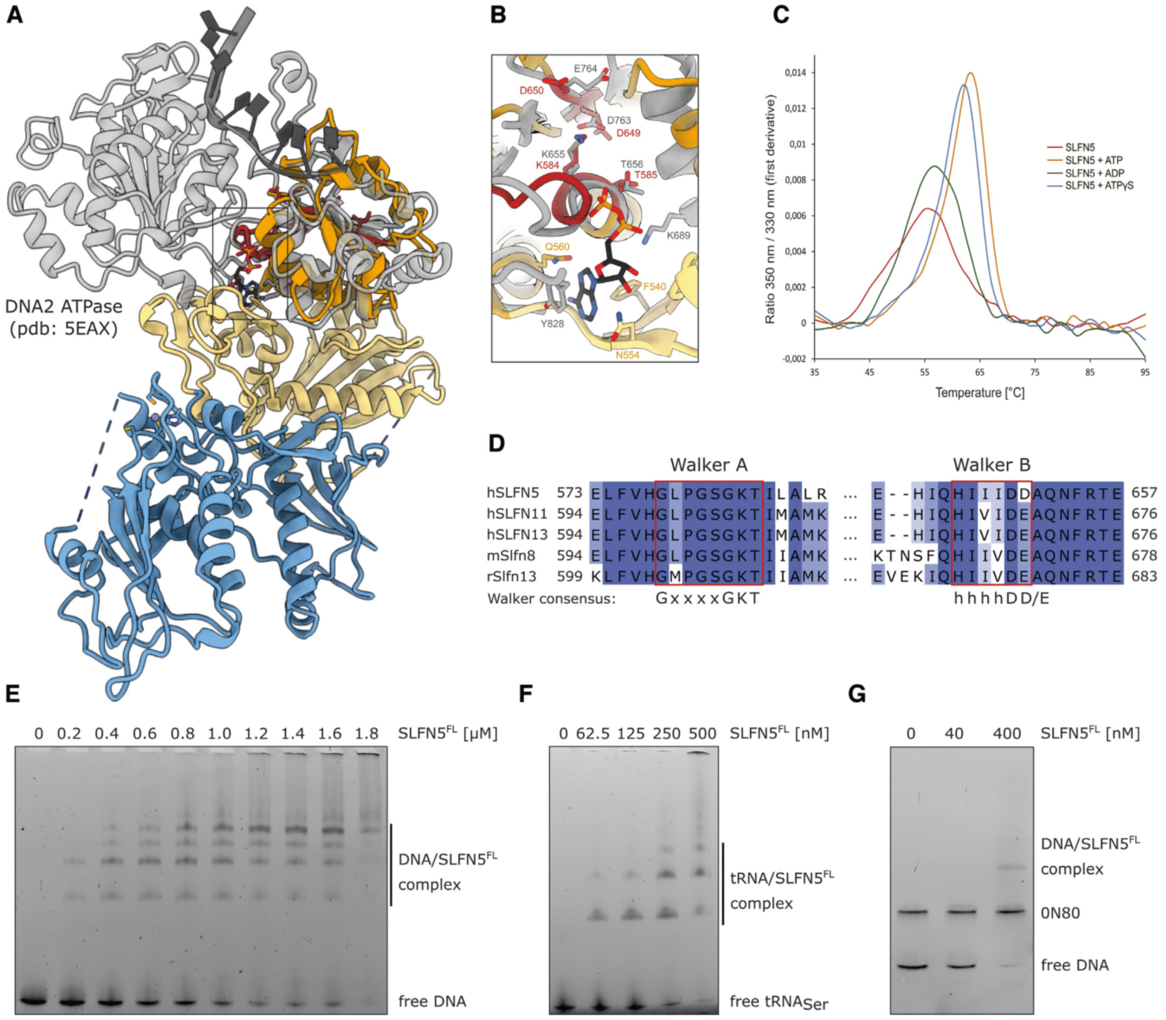


Figure 2. Biochemical characterization of full-length SLFN5. (**A**) Comparison of SLFN5 and the ATPase domain of DNA2 (gray) bound to ADP and ssDNA (only ATPase domain shown, PDB: 5EAX). (**B**) Superposition of ADP bound ATPase active site of DNA2 (gray) and the nucleotide free active site of SLFN5 (orange). The nucleotide and active site residues are displayed as sticks. (**C**) NanoDSF measurements of SLFN5 in presence of different nucleotides or without nucleotide. (**D**) Multiple sequence alignment of Walker A and B motifs of selected human, mouse and rat subgroup III Schlafen family members. Walker A and B motifs are highlighted in red. **(E)** Interaction of SLFN5 with 60 bp DNA monitored by electrophoretic mobility shift assay. (**F**) Interaction of SLFN5 with $tRNA_{Ser}$ monitored by electrophoretic mobility shift assay. (**G**) Competition interaction analysis between SLFN5 and 0N80 nucleosome or 227 bp DNA monitored by electrophoretic mobility shift assay.

bers, the last residue of the Walker B motif of SLFN5 is an aspartate compared to glutamate (Figure 2D). We tested nucleotide binding in a thermal unfolding assay using nano differential scanning fluorimetry (nanoDSF) (Figure 2C). Addition of ATP or ATPγS led to a shift of the inflection point towards higher temperature, indicating nucleotide binding. ADP binding resulted in a weaker shift of the inflection temperature. This could indicate structural differences between the diverse nucleotide states. Despite ATP binding by SLFN5, we could not detect ATP hydrolysis in a fluorescence-based ATPase assay (Supplementary Figure S5A-D). Neither single- or double-stranded DNA nor RNA or tRNA led to a stimulation of the ATPase. This could indicate that none of the tested substrates can stimulate the ATPase or that an essential additional factor, e.g. interaction partner, is still missing.

Compared to $SLFN5^{1-336}$, full-length SLFN5 showed an increase in affinity to double-stranded DNA as well as tRNA (Figure 2E, F). In an electrophoretic mobility shift assay with a 60 bp substrate, four distinct band shifts were visible with the second and fourth shift showing higher intensity (Figure 2E). This could hint towards a cooperative binding model. At high SLFN5 concentrations, the protein-DNA complex did not enter the gel, indicating either aggregation or formation of large complexes. In competitive shift assays with $tRNA_{Ser}$ and dsDNA, SLFN5 showed comparable affinities to $tRNA_{Ser}$ and 50 bp DNA, but a preference for 196 bp DNA over tRNA_{Ser} (Supplementary Figure S6A, B).

SLFN11 was previously shown to be a regulator of chromatin structure, leading to increased accessibility of promoter sites. This activity is ATPase dependent (73). Based on its similarity to SLFN11, we investigated whether SLFN5 could interact with nucleosomes, as a majority of the DNA in the nucleus is organized in nucleosomes. In a competition assay with 227 bp DNA and a 0N80 nucleosome that was assembled on the same sequence (147 bp nucleosomal DNA and 80 bp of extranucleosomal DNA) as substrates, SLFN5 showed a clear preference for free DNA over nucleosomes (Figure 2G).

Structural framework of $\text{SLFN5}^{1–336}$

To gain high resolution structural insights into human SLFN5, we crystallized the SLFN5 core domain ($\text{SLFN5}^{1–336}$) (Figure 3A, B). The core domain (38.9 kDa) behaves as a monomer in solution, as determined by right angle light scattering (RALS) (Supplementary Figure S7A, B). $\text{SLFN5}^{1–336}$ crystallized in the space groups $P3_221$ and $P2_1$, diffracting to 3.2 and 1.8 Å, respectively (Supplementary Figure S8A, Figure 3B). The structure was determined by SAD using the intrinsically bound zinc ion and was refined to a resolution of 1.8 Å (Figure 3B, C and Supplementary Figure S8B). The detailed refinement statistics are summarized in Supplementary Table S2.

The $\text{SLFN5}^{1- 336}$ structure depicts a horseshoe-like shape with a mixed α/β topology consisting of 10 α-helices and 14 β-sheets (Figure 3B, Supplementary Figure S8D). The approximate dimensions are 63 × 40 × 25 Å with the inner tunnel measuring 18 Å. The domain consists of an N-terminal and C-terminal lobe with respective bridging domain (Figure 3B). Each lobe consists of four α-helices and five β-sheets and each bridging domain of one helix and two sheets (Supplementary Figure S8D). The $\text{SLFN5}^{1–336}$ structure comprises a zinc finger motif with the calculated anomalous map confirming the presence of a zinc ion (Figure 3C and Supplementary Figure S8B). The zinc ion is coordinated by a histidine and three cysteine residues (H266, C268, C302, C303) (Figure 3C) that are highly conserved throughout the entire Schlafen protein family (Supplementary Figure S3). In most parts, the crystal structure of $\text{SLFN5}^{1–336}$ corresponds to the cryo-EM density.

The electron density for residues 145–168 could not be traced, indicating that this region consists of a flexible loop. We identified additional electron density close to the predicted active site, which could originate from a sulfate ion, as the protein was crystallized at an ammonium sulfate concentration of 1.5 M (Supplementary Figure S8C). The sulfate is coordinated by arginine 211 and via backbone interactions. As sulfates have been described to mimic phosphates of nucleic acids (74–76), this region could be a putative DNA binding or nuclease active site (Figure 3D).

The structure of $\text{SLFN5}^{1–336}$ shows similarity to the published structure of the N-terminal domain of rat Slfn13 ($\text{rSlfn13}^{14–353}$) (Supplementary Figure S8E) (42). However, the N-termini of the two structures adopt different conformations. The pseudo symmetry between the N- and C-lobes as reported for $\text{rSlfn13}^{14–353}$ is broken in $\text{SLFN5}^{1–336}$, as the sequential order of the secondary structure elements of each lobe differ slightly. $\text{SLFN5}^{1–336}$ lacks a helix between α3 and α4 that is present in $\text{rSlfn13}^{14–353}$. Moreover, the tunnel between the two lobes measures 18 Å in $\text{SLFN5}^{1–336}$ but 23 Å in $\text{rSlfn13}^{14–353}$. Taken together, the overall structure and the zinc finger motif of $\text{SLFN5}^{1–336}$ are similar to $\text{rSlfn13}^{14–353}$. However, differences between the secondary structural elements are evident.

SLFN5^{1-336} nucleic acid binding properties

Several members of the Schlafen family have been shown to interact with nucleic acids. While rSlfn13, human SLFN11 and rabbit Slfn14 are involved in tRNA or rRNA binding and processing (31,42,44), human SLFN5 has been shown to bind DNA, acting as a transcriptional regulator of ISGs (48). Therefore, we investigated the nucleic acid binding properties of $\text{SLFN5}^{1–336}$ to various nucleic acid substrates (Supplementary Table S3) by fluorescence anisotropy experiments (FA) (Figure 4A, B).

The SLFN5 core domain revealed comparable affinities to all RNA types tested and binds tRNA_{Ser} with an equilibrium dissociation constant (K_d-value) of 4.3 μM. Binding of $\text{SLFN5}^{1–336}$ to 30- and 60-mer poly (rU) ssRNA resulted in K_d values of 5.1 and 3.6 μM, respectively (Figure 4A).

Furthermore, we analyzed $\text{SLFN5}^{1–336}$ binding to various DNA substrates. The anisotropy measurements in the presence of single-stranded 5-,10-, 20- or 30-mer poly (dT) DNA (ssDNA) indicate low affinity to short ssDNA substrates (Figure 4B). With increasing DNA length, the affinity of $\text{SLFN5}^{1–336}$ to single-stranded DNA was enhanced with K_d values of 11 μM and 7.8 μM for 30- and 60-mer poly (dT) ssDNA, respectively. The highest affinity was measured for 60-mer dsDNA with a K_d of 2.5 μM (Figure 4B).

In order to identify new residues involved in nucleic acid binding, besides the residues that were previously described (42), structure guided mutants were designed. Two arginine residues in close proximity to the zinc finger region emerge from the surface, suggesting an essential role of these residues in nucleic acid binding (Figure 4C). Based on this, two $\text{SLFN5}^{1–336}$ mutants with substitutions of arginine to glutamate (R271E or R326E) were generated. Both mutants showed reduced binding to ssDNA, dsDNA and tRNA_{Ser} in anisotropy measurements (Figure 4D, E, Supplementary Figure S9A). The K_d values towards ssDNA, dsDNA and tRNA_{Ser} were increased compared to wild type $\text{SLFN5}^{1–336}$, indicating that the identified residues R271 and R326 are involved in substrate binding. The inflection temperatures of the mutants are not decreased compared to wild type SLFN5^{1-336}, indicating that the mutants are folded correctly (Supplementary Figure S9B). In summary, we could demonstrate a moderate binding preference of $\text{SLFN5}^{1–336}$ for dsDNA and the importance of the zinc finger region for DNA binding.

The Schlafen core domain has no ATPase activity

Based on sequence similarity, the N-terminal Schlafen domain was predicted to comprise a putative ATPase domain (27). Generally, the Walker A and B consensus sequences

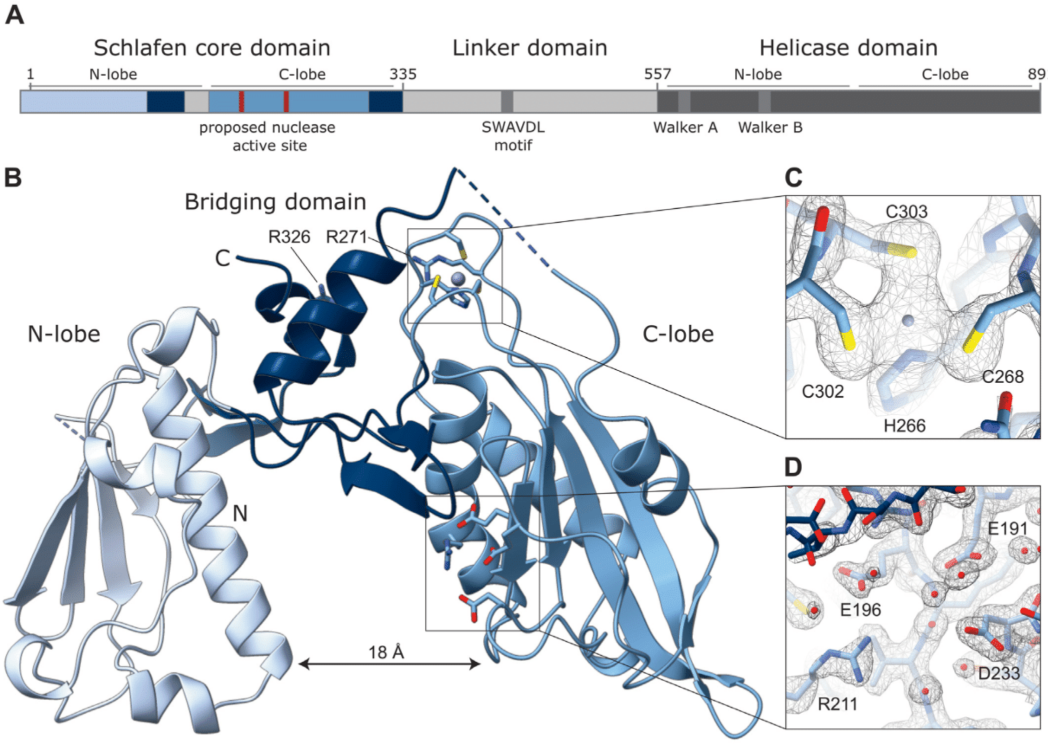


Figure 3. Structure of human SLFN5$^{1-336}$ N-terminal domain. (**A**) Scheme of the domain architecture of human SLFN5. (**B**) Cartoon representation of SLFN5$^{1-336}$ crystal structure solved in space group P2$_1$. The residues of the zinc finger, the predicted nuclease active site and R271 and R326 are visualized as sticks. (**C**) Close-up view of the zinc coordinating residues (H266, C268, C302 and C303). The $2F_o - F_c$ electron density map is colored in black and contoured at $\sigma = 1$. (**D**) Electron density around the predicted active site residues (E191, E196, D233) indicate the 1.8 Å resolution. The $2F_o - F_c$ electron density map is colored in black and contoured at $\sigma = 1$.

consist of GxxxxGK(T/S) (where x is any amino acid) and hhhhD(D/E) (where h is a hydrophobic amino acid) (77). In human SLFN14, the conserved aspartate residues D248 and D249 were predicted to be part of a Walker B motif (44). However, in this model, the Walker B motif would be inserted into a disrupted Walker A motif (Supplementary Figure S10A).

When comparing these residues in the crystal structures of SLFN5$^{1-336}$ and rSlfn13 (Supplementary Figure S10B, C) to the active sites of other ATPases, no explicit structural similarity is evident. In SLFN5 and rSlfn13, the putative Walker A and B motifs are overlapping, whereas in functional ATPases, the motifs generally face each other to fulfill their function. Furthermore, the secondary structure elements in SLFN5$^{1-336}$ and rSlfn13 are in a reverse order compared to known ATPases. The glycine-rich loop is preceded by a helix and followed by a β-sheet (Supplementary Figure S10B, C), while it is typically the other way around in known ATPases (77).

We tested nucleotide binding of SLFN5$^{1-336}$ in a thermal unfolding assay using nano differential scanning fluorimetry (nanoDSF). None of the nucleotides tested (ADP, ATP or ATPγS) resulted in a change in the inflection temperature, suggesting that SLFN5$^{1-336}$ does not bind to ATP (Supplementary Figure S10D). To test for ATP hydrolysis, SLFN5$^{1-336}$ was incubated with radioactively labeled [γ-^{32}P] ATP and the release of $^{32}P_i$ was monitored by thin layer chromatography (TLC). However, regardless of the presence of 60-mer poly (dT) ssDNA, no significantly increased release of $^{32}P_i$ was detected, indicating that no ATP hydrolysis took place (Supplementary Figure S10E).

The structural and biochemical data of SLFN5$^{1-336}$ show that it does not resemble an ATPase-like fold, nor does it bind or hydrolyze ATP, confirming that the Schlafen core domain is not an ATPase.

The Schlafen proteins share a similar fold, but differ in the predicted active site

Several Schlafen family members have been described to cleave and process RNAs. The N-terminal domain of rSlfn13 endonucleolytically cleaves tRNA$_{Ser}$ in a Mg^{2+}/Mn^{2+} dependent manner (42) and human SLFN11 was shown to cleave type II tRNAs (31). Therefore, we tested the nuclease activity of different Schlafen protein members on tRNA$_{Ser}$.

SLFN5$^{1-336}$ and full-length SLFN5 showed no nuclease activity on tRNA$_{Ser}$, regardless of the presence of Mn^{2+} (Figure 5A, B). Murine Slfn8, a homologue of rat rSlfn13, was used as positive control. We confirm nuclease activity of mSlfn8$^{1-351}$ on tRNA$_{Ser}$ in a Mg^{2+}/Mn^{2+} dependent manner, as the addition of EDTA prevented cleavage (Figure

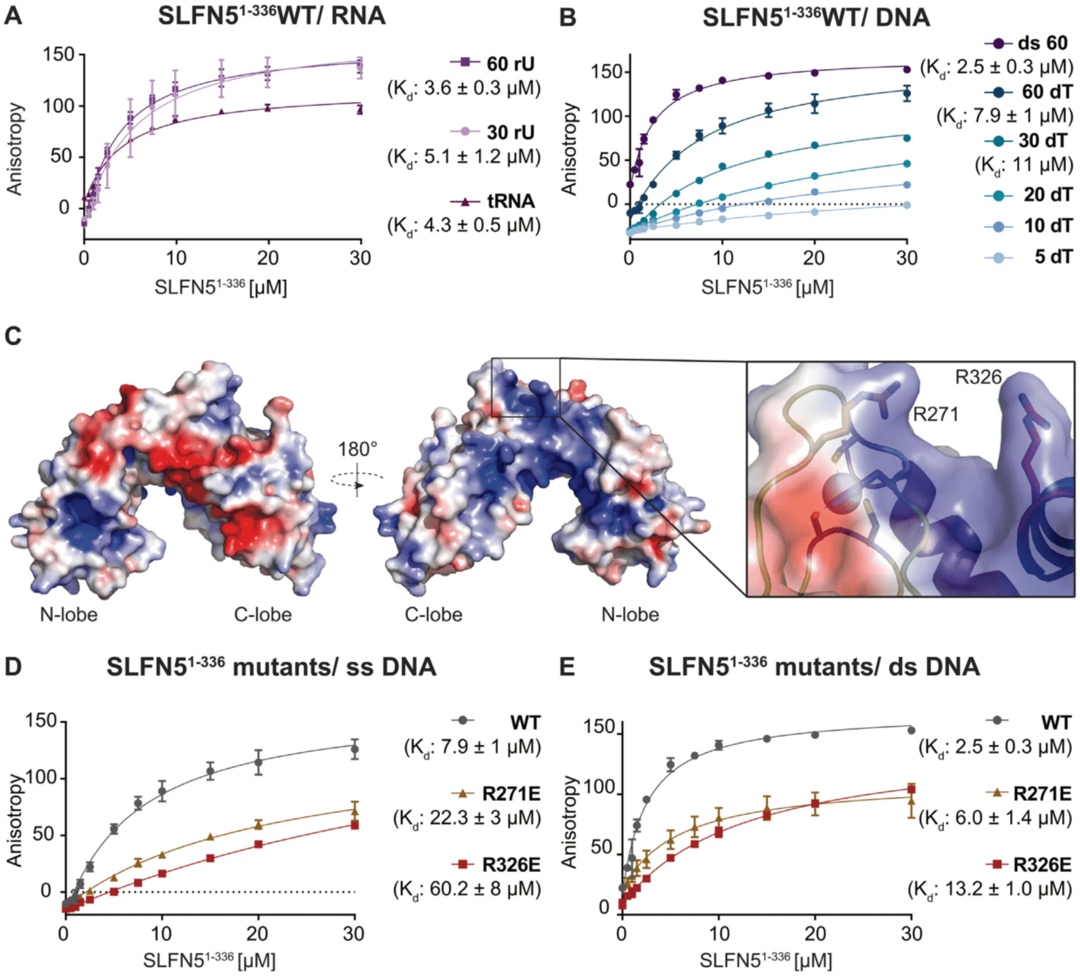


Figure 4. Analysis of SLFN5$^{1-336}$ nucleic acid binding properties. (A, B) Fluorescence anisotropy assay to monitor the binding of SLFN5$^{1-336}$ to different nucleic acid substrates. Different nucleic acid ligands are depicted in different colors. The data were fit to a 1 to 1 binding equation. Error bars represent the standard deviation from three independent experiments. (**A**) Binding of SLFN5$^{1-336}$ to single-stranded 30-, 60-mer RNA and tRNA$_{Ser}$. (**B**) Binding of SLFN5$^{1-336}$ to single-stranded 5-, 10-, 20-, 30- and 60-mer poly d(T) DNA and double-stranded 60-mer DNA. (**C**) Electrostatic surface potential of SLFN5$^{1-336}$ colored from red (–4kT/e) to blue (4kT/e) and close-up view of SLFN5$^{1-336}$ R271 and R326, which are located in close proximity to the zinc finger. (D, E) Interaction of SLFN5$^{1-336}$ and the charge reverse mutants SLFN5$^{1-336}$ R271E and R326E to single-stranded 60-mer (**D**) and double-stranded 60-mer DNA (**E**) by the change in fluorescence anisotropy. The data were fit to a 1 to 1 binding equation. Error bars represent the standard deviation from three independent experiments.

5A). The cleavage pattern was similar to the one previously observed for rSlfn13 (42). Murine mSlfn2 showed no nuclease activity towards tRNA$_{Ser}$ (Figure 5A), regardless if Mn^{2+} was present or not.

However, full-length SLFN11 exhibited Mn^{2+}-dependent cleavage activity on tRNA$_{Ser}$ (Figure 5B). The migration behavior of the cleavage product suggests the cleavage site to be located ∼10 nt from the 3' end of the tRNA. This places the cut site between the acceptor arm and the T-arm of the tRNA. As SLFN5 binds, but does not cleave tRNA, we tested whether it could inhibit the tRNA cleavage activity of SLFN11. However, at equimolar concentrations, SLFN5 showed no inhibitory effect on tRNA cleavage by SLFN11 (Figure 5B). Furthermore, SLFN5$^{1-336}$ showed no nuclease activity on ssDNA or dsDNA (Supplementary Figure S11).

Yang *et al.* identified a conserved three-carboxylate triad (E205, E210 and D248) that is responsible for the endonuclease activity of rSlfn13 (Figure 5C–E) (42). In human SLFN5$^{1-336}$ the carboxylate motif, consisting of residues E191, E196 and D233, is conserved (Figure 5D). The additional density, which we interpreted as a sulfate ion in our crystal structure (Supplementary Figure S8C), is located in close proximity to the carboxylate motif, suggesting that this region could be involved in nucleic acid interaction, as sulfates are known to mimic the phosphate groups of nucleic acids (74–76). When comparing Schlafen 5 homologues of various species, the carboxylate triad exhibits little conser-

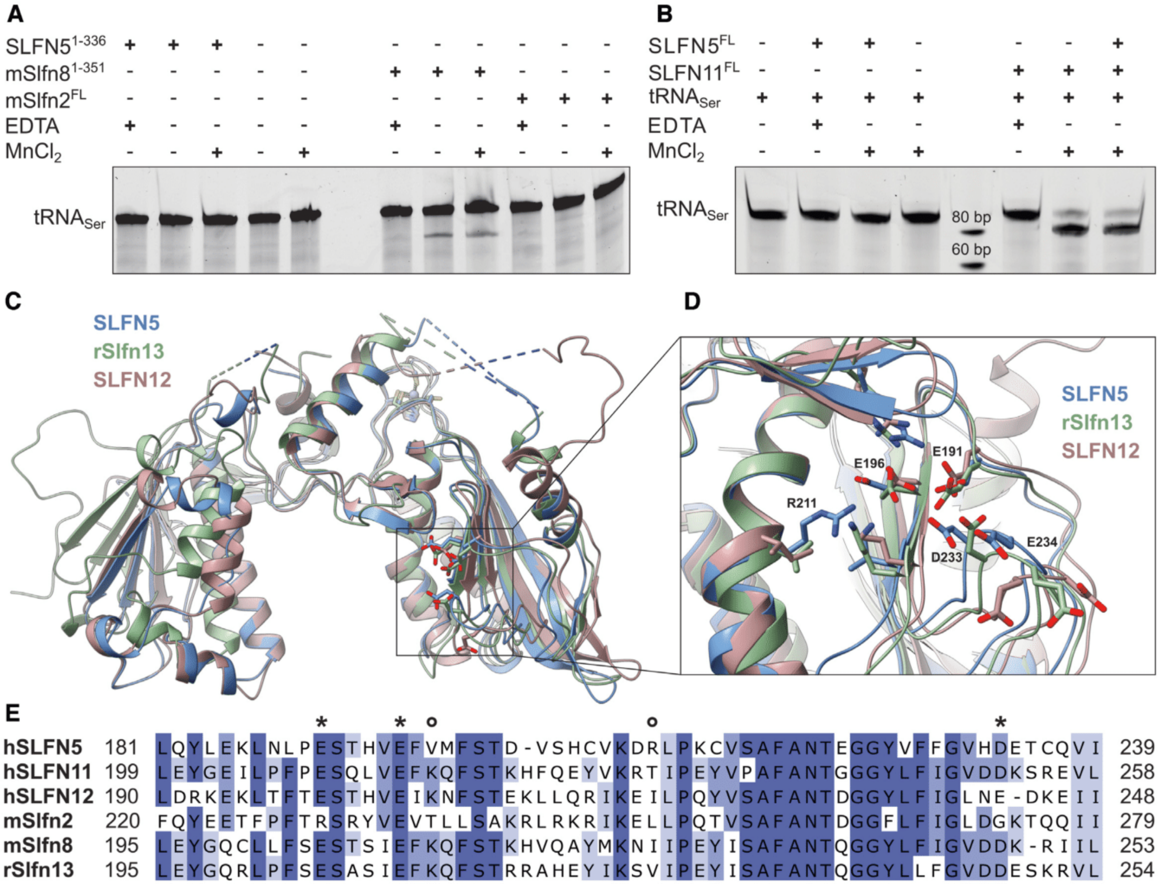


Figure 5. Comparison of the proposed nuclease active site between different Schlafen protein family members. (**A**) Endonuclease assay of SLFN5$^{1-336}$, mSlfn8$^{1-351}$ and mSlfn2^{FL} using tRNA$_{Ser}$ as substrate. (**B**) Endonuclease assay of SLFN5 and SLFN11 using tRNA$_{Ser}$ as substrate. (**C**) Overlay of the Slfn core domains of hSLFN5 (blue), rSlfn13 (green, PDB: 5YD0) and hSLFN12 (red, PDB: 7LRE). Residues of the proposed nuclease sites and the zinc finger are displayed as sticks. (**D**) Detailed view of the predicted nuclease site. Naming of the residues corresponds to SLFN5. (**E**) Multiple sequence alignment of the proposed active site regions of selected Schlafen proteins. The proposed catalytic residues are indicated (E/D = *, K/R = o)

vation. E196 (E210 in rSlfn13) is the only residue of the predicted carboxylate triad that shows a high degree of conservation, while E191 (E205 in rSlfn13) and D233 (D248 in rSlfn13) are only partially conserved (Supplementary Figure S3).

Since human SLFN12 and rSlfn13 have been shown to cleave rRNA and tRNAs, respectively (42,45), we superimposed the available structures to further investigate the active sites (Figure 5C, D). The superposition illustrates the similarities within their active site residues (Figure 5D). Additional to the carboxylate triad, both proteins harbor a lysine residue (K212 in rSlfn13 and K207 in SLFN12) in close proximity to the active site, which might be involved in the ribonuclease reaction. This lysine residue cannot be found in SLFN5 nor in mSlfn2. Instead, arginine R211 is involved in the coordination of the sulfate ion in the SLFN5$^{1-336}$ structure (Supplementary Figure S8C). However, R211 in SLFN5 is also only partially conserved between different species.

The subgroup I Schlafen protein member mSlfn2 shows no cleavage activity on tRNA$_{Ser}$ (Figure 5A). This is in line with the observation that murine mSlfn1, another Schlafen I protein subgroup member, lacks endonuclease activity as well (42). Out of the three proposed essential active site residues, only E235 (E196 in SLFN5) is present in mSlfn2 (Supplementary Figure S3). Instead, it harbors several positively charged amino acids (R247, R230, K274).

The lack of the lysine residue in SLFN5 and mSlfn2, which is conserved in cleavage proficient Schlafen family members mSlfn8, rSlfn13, hSLFN11 and hSLFN12, as well as the low conservation score for the residues of the three-carboxylate triad amongst Schlafen 5 proteins between different species, agree with the absence of nuclease activity in our assays.

In summary, our results indicate that even though the overall fold of the Slfn core domain is conserved, the enzymatic activity differs between the different Schlafen proteins.

The *elektra* mutation could lead to misfolding and aggregation of the mSlfn2 protein

Elektra mice are highly immunodeficient due to a point mutation in the *mSlfn2* gene, which leads to a substitution of isoleucine 135 to asparagine (I135N) (7). In order to understand the influence of the *elektra* mutation on the protein structure, we generated an AlphaFold model of full-length mSlfn2 (70). I135 is located in the N-lobe of mSlfn2, surrounded by several hydrophobic residues forming a hydrophobic patch (Supplementary Figure S12A). These residues are highly conserved throughout all murine and human Schlafen members (Supplementary Figure 12B), indicating that the hydrophobic patch region is present in all Schlafen proteins. The introduction of a polar side chain into the hydrophobic patch by the I135N mutation might interfere with proper folding of the N-lobe, presumably resulting in aggregation. To support this hypothesis we expressed both wild type mSlfn2 and the I135N mutant in a bacterial expression system. We were able to solubly express and purify the wild type protein (Supplementary Figure S12C). On the other hand, the protein harboring the *elektra* mutation was insoluble, which might be a sign of aggregation.

CONCLUSION

The Schlafen (Slfn) protein family belongs to the interferon-stimulated genes, which play key roles in immune defense and pathogen control (4). Although the Schlafen family members share many highly conserved sequence regions, their biological roles and enzymatic functions differ. While some family members such as rat Slfn13 and human SLFN11 influence the translation machinery (31,42), others, such as SLFN5, are involved in transcriptional regulation (26,47,50). In particular, the molecular mechanism of human SLFN5 in tumor control is not well understood.

Here, we present the cryo-EM structure of human full-length SLFN5 together with the high-resolution X-ray structure of the N-terminal core domain and biochemical data. The N-terminal domain is the common core domain of all Schlafen family members and supposedly involved in nucleic acid substrate recognition. The overall structural organization of the Schlafen core domain, including the zinc finger, is similar to rat Slfn13 (42). Nevertheless, they differ in some secondary structure elements and active site residues. Rat Slfn13 has been identified as an endoribonuclease with an active site consisting of a carboxylate triad in the C-lobe. However, the tRNA endoribonuclease activity could not be confirmed for human SLFN5, mSlfn1 or mSlfn2 by us and by others (42,78). In addition, the active site residues proposed by Yang *et al.* are not entirely conserved through all Schlafen family members (42). This confirms a divergent enzymatic function within the protein family. Despite performing nuclease activity assays with a variety of nucleic acids, we could not identify a substrate which is cleaved by SLFN5 or mSlfn2. However, we cannot exclude the necessity of additional factors for enzymatic activity. In order to understand the biochemical and enzymatic functions of the Schlafen core domain of SLFN5, we analyzed the nucleic acid binding properties. Our data show that the affinity of SLFN5$^{1-336}$ to single-stranded DNA and RNA steadily increases in a length dependent manner from 20 bases up to 60 bases. The highest affinity was detected for a 60-mer double-stranded DNA, followed by 60-mer single-stranded RNA and tRNA$_{Ser}$. Full-length SLFN5 binds double-stranded DNA with high affinity, which is in line with recent reports, showing SLFN5 to play a role in transcriptional regulation (26,50). The analysis of the electrostatic surface potential of the SLFN5 core domain revealed a positively charged patch. This region is in close proximity to the highly conserved zinc finger region and on the opposite site of the molecule compared to the previously proposed nuclease active site. To prove the involvement of this region in nucleic acid recognition, we mutated the residues R271 or R326 to glutamate, which decreased the binding affinity to DNA two to ten-fold and to tRNA$_{Ser}$ two to five-fold, respectively. Thus, we identified an additional region involved in substrate binding. Furthermore, structural analysis as well as ATPase assays of SLFN5$^{1-336}$ corroborated that the Schlafen core domain neither has the necessary ATPase motifs nor possess ATP hydrolysis activity, disproving earlier sequence-based predictions. We confirm a computational analysis, reporting that the Schlafen core domain is not an ATPase (79).

In contrast to the N-terminal core domain, our data demonstrate that full-length SLFN5 is able to bind ATP. Despite the fact that all necessary residues for an active ATPase exist in the C-terminal helicase domain of SLFN5, we could not observe ATP hydrolysis. The protein remained inactive in presence of different DNA and RNA substrates. We cannot exclude that additional cofactors, conformational rearrangements, interaction partners or substrates are needed to activate the enzymatic activity of the helicase domain.

The presented cryo-EM structure identified the linker domain of SLFN5 as a connector between the N-terminal SLFN5 core domain and the C-terminal helicase domain. The YPXSY motif (reidue 514–518) in the linker region as well as residue F98 involved in the interaction of the Schlafen core with the linker domain are highly conserved between subgroup II and III Schlafen family members including SLFN11 (Supplementary Figure S3). Moreover, the interface between the helicase and linker domain is conserved between the long Schlafen proteins, with the exception of Slfn14. This high degree of sequence conservation suggests that the identified domain organization and interfaces are preserved in both subgroups. However, Slfn14 seems to differ in the helicase - linker domain interface, which could indicate a divergent structural organization. This might be necessary for its function in ribosome degradation (43). In SLFN11, the clinically most relevant Schlafen protein, the domain interface motifs are conserved, suggesting a homologue structural arrangement. Recently, SLFN12 also came into focus of cancer research. It has been described that small molecule compounds, which stabilize complex formation between the phosphodiesterase PDE3A and SLFN12, cause selective cancer cell killing (45). The PDE3A–SLFN12 complex is formed by the interaction of the C-terminal extending PIR (phosphodiesterase interacting region) helix of SLFN12 with PDE3A. The structural comparison of

SLFN5 with SLFN12 indicates that the PIR helix is only found in SLFN12 (Supplementary Figure S13A–D), excluding a similar interaction between SLFN5 and PDE3A. Based on sequence alignment and AlphaFold models (data not shown), the PIR helix cannot be found in subgroup III Schlafen proteins.

In summary, our study provides valuable insights into Schlafen subgroup III family members and will support the investigation of their diverse enzymatic functions.

DATA AVAILABILITY

The coordinates and structure factors of the 1.8 Å and 3.2 Å SLFN5$^{1\text{-}336}$ structures have been deposited in the Protein Data Bank under the accession codes 7Q3Z and 6RR9, respectively. The coordinates of the full-length SLFN5 structure have been deposited in the Protein Data Bank under the accession code 7PPJ. The SLFN5 reconstruction is available at the Electron Microscopy Data Bank under the EMDB accession code EMD-13581.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We acknowledge the Swiss Light Source at the Paul-Scherrer Institute (Villigen, Switzerland) for the provision of experimental facilities. We thank S. Wenzl for help with SLFN11 protein purification. We thank F. Kunert for help with nucleosome preparation. We thank Dr. F. Schlauderer for providing the pET21-SLFN5$^{1\text{-}336}$, pET21-mSlfn2^{FL} and pET21-mSlfn8$^{1\text{–}351}$ plasmids. Furthermore, we thank all members of the Hopfner lab for helpful discussions.

Author contributions: E.H., F.J.M. and K.L. designed and conducted all structural and biochemical experiments. F.J.M. and K.L. carried out the cryo-EM data collection and analysis of full-length SLFN5. E.H. crystallized SLFN5$^{1\text{–}336}$, E.H. and K.L. collected and processed the crystallographic data and carried out the structural determination and analysis. K.-P.H. helped with the analysis and interpretation of the results. E.H., F.J.M. and K.L. wrote the manuscript.

FUNDING

Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) [Project-ID 210592381 (SFB1054)]; Gottfried Wilhelm Leibniz-Prize (to K.-P.H.); European Research Council [ERC Advanced Grant INO3D to K.-P.H.]; International Max-Planck Research School of Molecular Life Sciences (to E.H.). Funding for open access charge: CRC[1054].

Conflict of interest statement. None declared.

REFERENCES

1. Schwarz,D.A., Katayama,C.D. and Hedrick,S.M. (1998) Schlafen, a new family of growth regulatory genes that affect thymocyte development. *Immunity*, **9**, 657–668.
2. Geserick,P., Kaiser,F., Klemm,U., Kaufmann,S.H. and Zerrahn,J. (2004) Modulation of t cell development and activation by novel members of the schlafen (slfn) gene family harbouring an RNA helicase-like motif. *Int. Immunol.*, **16**, 1535–1548.
3. Katsoulidis,E., Carayol,N., Woodard,J., Konieczna,I., Majchrzak-Kita,B., Jordan,A., Sassano,A., Eklund,E.A., Fish,E.N. and Platanias,L.C. (2009) Role of schlafen 2 (SLFN2) in the generation of interferon alpha-induced growth inhibitory responses. *J. Biol. Chem.*, **284**, 25051–25064.
4. Katsoulidis,E., Mavrommatis,E., Woodard,J., Shields,M.A., Sassano,A., Carayol,N., Sawicki,K.T., Munshi,H.G. and Platanias,L.C. (2010) Role of interferon {alpha} (IFN{alpha})-inducible schlafen-5 in regulation of anchorage-independent growth and invasion of malignant melanoma cells. *J. Biol. Chem.*, **285**, 40333–40341.
5. Brady,G., Boggan,L., Bowie,A. and O'Neill,L.A. (2005) Schlafen-1 causes a cell cycle arrest by inhibiting induction of cyclin d1. *J. Biol. Chem.*, **280**, 30723–30734.
6. Kuang,C.Y., Yang,T.H., Zhang,Y., Zhang,L. and Wu,Q. (2014) Schlafen 1 inhibits the proliferation and tube formation of endothelial progenitor cells. *PLoS One*, **9**, e109711.
7. Berger,M., Krebs,P., Crozat,K., Li,X., Croker,B.A., Siggs,O.M., Popkin,D., Du,X., Lawson,B.R., Theofilopoulos,A.N. *et al.* (2010) An slfn2 mutation causes lymphoid and myeloid immunodeficiency due to loss of immune cell quiescence. *Nat. Immunol.*, **11**, 335–343.
8. Puck,A., Aigner,R., Modak,M., Cejka,P., Blaas,D. and Stockl,J. (2015) Expression and regulation of schlafen (SLFN) family members in primary human monocytes, monocyte-derived dendritic cells and t cells. *Results Immunol*, **5**, 23–32.
9. Puck,A., Hopf,S., Modak,M., Majdic,O., Cejka,P., Bluml,S., Schmetterer,K., Arnold-Schrauf,C., Gerwien,J.G., Frederiksen,K.S. *et al.* (2017) The soluble cytoplasmic tail of CD45 (ct-CD45) in human plasma contributes to keep t cells in a quiescent state. *Eur. J. Immunol.*, **47**, 193–205.
10. Goldshtein,A., Zerbib,S.M., Omar,I., Cohen-Daniel,L., Popkin,D. and Berger,M. (2016) Loss of T-cell quiescence by targeting slfn2 prevents the development and progression of T-ALL. *Oncotarget*, **7**, 46835–46847.
11. Omar,I., Lapenna,A., Cohen-Daniel,L., Tirosh,B. and Berger,M. (2016) Schlafen2 mutation unravels a role for chronic ER stress in the loss of t cell quiescence. *Oncotarget*, **7**, 39396–39407.
12. Basson,M.D., Wang,Q., Chaturvedi,L.S., More,S., Vomhof-DeKrey,E.E., Al-Marsoummi,S., Sun,K., Kuhn,L.A., Kovalenko,P. and Kiupel,M. (2018) Schlafen 12 interaction with serpinb12 and deubiquitylases drives human enterocyte differentiation. *Cell. Physiol. Biochem.*, **48**, 1274–1290.
13. Kovalenko,P.L. and Basson,M.D. (2014) Schlafen 12 expression modulates prostate cancer cell differentiation. *J. Surg. Res.*, **190**, 177–184.
14. Lee,N.K., Choi,H.K., Yoo,H.J., Shin,J. and Lee,S.Y. (2008) RANKL-induced schlafen2 is a positive regulator of osteoclastogenesis. *Cell. Signal.*, **20**, 2302–2308.
15. Mavrommatis,E., Arslan,A.D., Sassano,A., Hua,Y., Kroczynska,B. and Platanias,L.C. (2013) Expression and regulatory effects of murine schlafen (Slfn) genes in malignant melanoma and renal cell carcinoma. *J. Biol. Chem.*, **288**, 33006–33015.
16. Tian,L., Song,S., Liu,X., Wang,Y., Xu,X., Hu,Y. and Xu,J. (2014) Schlafen-11 sensitizes colorectal carcinoma cells to irinotecan. *Anticancer Drugs*, **25**, 1175–1181.
17. Zoppoli,G., Regairaz,M., Leo,E., Reinhold,W.C., Varma,S., Ballestrero,A., Doroshow,J.H. and Pommier,Y. (2012) Putative DNA/RNA helicase schlafen-11 (SLFN11) sensitizes cancer cells to DNA-damaging agents. *Proc. Natl. Acad. Sci. U.S.A.*, **109**, 15030–15035.
18. de Waal,L., Lewis,T.A., Rees,M.G., Tsherniak,A., Wu,X., Choi,P.S., Gechijian,L., Hartigan,C., Faloon,P.W., Hickey,M.J. *et al.* (2016) Identification of cancer-cytotoxic modulators of PDE3A by predictive chemogenomics. *Nat. Chem. Biol.*, **12**, 102–108.
19. Barretina,J., Caponigro,G., Stransky,N., Venkatesan,K., Margolin,A.A., Kim,S., Wilson,C.J., Lehar,J., Kryukov,G.V., Sonkin,D. *et al.* (2012) The cancer cell line encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature*, **483**, 603–607.

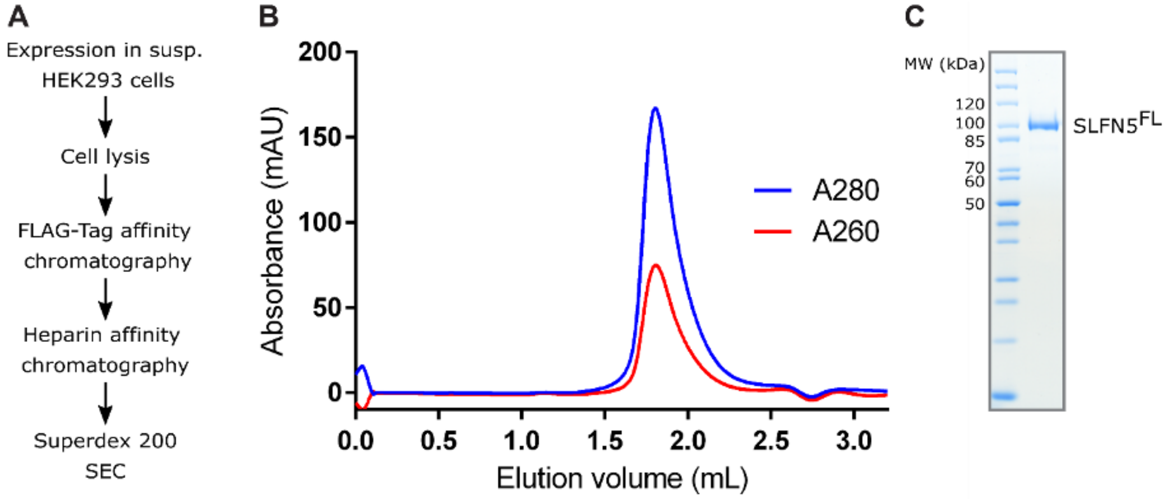
20. Nogales,V., Reinhold,W.C., Varma,S., Martinez-Cardus,A., Moutinho,C., Moran,S., Heyn,H., Sebio,A., Barnadas,A., Pommier,Y. *et al.* (2016) Epigenetic inactivation of the putative DNA/RNA helicase SLFN11 in human cancer confers resistance to platinum drugs. *Oncotarget*, **7**, 3084–3097.
21. Lok,B.H., Gardner,E.E., Schneeberger,V.E., Ni,A., Desmeules,P., Rekhtman,N., de Stanchina,E., Teicher,B.A., Riaz,N., Powell,S.N. *et al.* (2017) PARP inhibitor activity correlates with SLFN11 expression and demonstrates synergy with temozolomide in small cell lung cancer. *Clin. Cancer Res.*, **23**, 523–535.
22. Murai,J., Feng,Y., Yu,G.K., Ru,Y., Tang,S.W., Shen,Y. and Pommier,Y. (2016) Resistance to PARP inhibitors by SLFN11 inactivation can be overcome by ATR inhibition. *Oncotarget*, **7**, 76534–76550.
23. Gardner,E.E., Lok,B.H., Schneeberger,V.E., Desmeules,P., Miles,L.A., Arnold,P.K., Ni,A., Khodos,I., de Stanchina,E., Nguyen,T. *et al.* (2017) Chemosensitive relapse in small cell lung cancer proceeds through an EZH2-SLFN11 axis. *Cancer Cell*, **31**, 286–299.
24. Li,M., Kao,E., Gao,X., Sandig,H., Limmer,K., Pavon-Eternod,M., Jones,T.E., Landry,S., Pan,T., Weitzman,M.D. *et al.* (2012) Codon-usage-based inhibition of HIV protein synthesis by human schlafen 11. *Nature*, **491**, 125–128.
25. Seong,R.K., Seo,S.W., Kim,J.A., Fletcher,S.J., Morgan,N.V., Kumar,M., Choi,Y.K. and Shin,O.S. (2017) Schlafen 14 (SLFN14) is a novel antiviral factor involved in the control of viral replication. *Immunobiology*, **222**, 979–988.
26. Kim,E.T., Dybas,J.M., Kulej,K., Reyes,E.D., Price,A.M., Akhtar,L.N., Orr,A., Garcia,B.A., Boutell,C. and Weitzman,M.D. (2021) Comparative proteomics identifies schlafen 5 (SLFN5) as a herpes simplex virus restriction factor that suppresses viral transcription. *Nat Microbiol*, **6**, 234–245.
27. Chaturvedi,L., Sun,K., Walsh,M.F., Kuhn,L.A. and Basson,M.D. (2014) The P-loop region of schlafen 3 acts within the cytosol to induce differentiation of human caco-2 intestinal epithelial cells. *Biochim. Biophys. Acta*, **1843**, 3029–3037.
28. Neumann,B., Zhao,L., Murphy,K. and Gonda,T.J. (2008) Subcellular localization of the schlafen protein family. *Biochem. Biophys. Res. Commun.*, **370**, 62–66.
29. Omar,I., Rom,O., Aviram,M., Cohen-Daniel,L., Gebre,A.K., Parks,J.S. and Berger,M. (2017) Slfn2 mutation-induced loss of T-cell quiescence leads to elevated de novo sterol synthesis. *Immunology*, **152**, 484–493.
30. Fischietti,M., Arslan,A.D., Sassano,A., Saleiro,D., Majchrzak-Kita,B., Ebine,K., Munshi,H.G., Fish,E.N. and Platanias,L.C. (2018) Slfn2 regulates type i interferon responses by modulating the NFkB pathway. *Mol. Cell. Biol.*, **38**, e00053-18.
31. Li,M., Kao,E., Malone,D., Gao,X., Wang,J.Y.J. and David,M. (2018) DNA damage-induced cell death relies on SLFN11-dependent cleavage of distinct type II tRNAs. *Nat. Struct. Mol. Biol.*, **25**, 1047–1058.
32. Murai,J., Tang,S.W., Leo,E., Baechler,S.A., Redon,C.E., Zhang,H., Al Abo,M., Rajapakse,V.N., Nakamura,E., Jenkins,L.M.M. *et al.* (2018) SLFN11 blocks stressed replication forks independently of ATR. *Mol. Cell*, **69**, 371–384.
33. Winkler,C., Armenia,J., Jones,G.N., Tobalina,L., Sale,M.J., Petreus,T., Baird,T., Serra,V., Wang,A.T., Lau,A. *et al.* (2021) SLFN11 informs on standard of care and novel treatments in a wide range of cancer models. *Br. J. Cancer*, **124**, 951–962.
34. Rathkey,D., Khanal,M., Murai,J., Zhang,J., Sengupta,M., Jiang,Q., Morrow,B., Evans,C.N., Chari,R., Fetsch,P. *et al.* (2020) Sensitivity of mesothelioma cells to PARP inhibitors is not dependent on BAP1 but is enhanced by temozolomide in cells with high-schlafen 11 and Low-O6-methylguanine-DNA methyltransferase expression. *J. Thorac. Oncol.*, **15**, 843–859.
35. Takashima,T., Taniyama,D., Sakamoto,N., Yasumoto,M., Asai,R., Hattori,T., Honma,R., Thang,P.Q., Ukai,S., Maruyama,R. *et al.* (2021) Schlafen 11 predicts response to platinum-based chemotherapy in gastric cancers. *Br. J. Cancer*, **125**, 65–77.
36. Shee,K., Wells,J.D., Jiang,A. and Miller,T.W. (2019) Integrated pan-cancer gene expression and drug sensitivity analysis reveals SLFN11 mRNA as a solid tumor biomarker predictive of sensitivity to DNA-damaging chemotherapy. *PLoS One*, **14**, e0224267.
37. Coussy,F., El-Botty,R., Château-Joubert,S., Dahmani,A., Montaudon,E., Leboucher,S., Morisset,L., Painsec,P., Sourd,L., Huguet,L. *et al.* (2020) BRCAness, SLFN11, and RB1 loss predict response to topoisomerase i inhibitors in triple-negative breast cancers. *Sci. Transl. Med.*, **12**, eaax2625.
38. Kang,M.H., Wang,J., Makena,M.R., Lee,J.S., Paz,N., Hall,C.P., Song,M.M., Calderon,R.I., Cruz,R.E., Hindle,A. *et al.* (2015) Activity of MM-398, nanoliposomal irinotecan (nal-IRI), in ewing's family tumor xenografts is associated with high exposure of tumor to drug and high SLFN11 expression. *Clin. Cancer Res.*, **21**, 1139–1150.
39. Allison Stewart,C., Tong,P., Cardnell,R.J., Sen,T., Li,L., Gay,C.M., Masrorpour,F., Fan,Y., Bara,R.O., Feng,Y. *et al.* (2017) Dynamic variations in epithelial-to-mesenchymal transition (EMT), ATM, and SLFN11 govern response to PARP inhibitors and cisplatin in small cell lung cancer. *Oncotarget*, **8**, 28575–28587.
40. Okamoto,Y., Abe,M., Mu,A., Tempaku,Y., Rogers,C.B., Mochizuki,A.L., Katsuki,Y., Kanemaki,M.T., Takaori-Kondo,A., Sobeck,A. *et al.* (2021) SLFN11 promotes stalled fork degradation that underlies the phenotype in fanconi anemia cells. *Blood*, **137**, 336–348.
41. Murai,Y., Jo,U., Murai,J., Jenkins,L.M., Huang,S.N., Chakka,S., Chen,L., Cheng,K., Fukuda,S., Takebe,N. *et al.* (2021) SLFN11 inactivation induces proteotoxic stress and sensitizes cancer cells to ubiquitin activating enzyme inhibitor TAK-243. *Cancer Res.*, **81**, 3067–3078.
42. Yang,J.Y., Deng,X.Y., Li,Y.S., Ma,X.C., Feng,J.X., Yu,B., Chen,Y., Luo,Y.L., Wang,X., Chen,M.L. *et al.* (2018) Structure of schlafen13 reveals a new class of tRNA/rRNA- targeting Rnase engaged in translational control. *Nat. Commun.*, **9**, 1165.
43. Fletcher,S.J., Pisareva,V.P., Khan,A.O., Tcherepanov,A., Morgan,N.V. and Pisarev,A.V. (2018) Role of the novel endoribonuclease SLFN14 and its disease-causing mutations in ribosomal degradation. *RNA*, **24**, 939–949.
44. Pisareva,V.P., Muslimov,I.A., Tcherepanov,A. and Pisarev,A.V. (2015) Characterization of novel ribosome-associated endoribonuclease SLFN14 from rabbit reticulocytes. *Biochemistry*, **54**, 3286–3301.
45. Garvie,C.W., Wu,X., Papanastasiou,M., Lee,S., Fuller,J., Schnitzler,G.R., Horner,S.W., Baker,A., Zhang,T., Mullahoo,J.P. *et al.* (2021) Structure of PDE3A-SLFN12 complex reveals requirements for activation of SLFN12 RNase. *Nat. Commun.*, **12**, 4375.
46. Sassano,A., Mavrommatis,E., Arslan,A.D., Kroczynska,B., Beauchamp,E.M., Khuon,S., Chew,T.L., Green,K.J., Munshi,H.G., Verma,A.K. *et al.* (2015) Human schlafen 5 (SLFN5) is a regulator of motility and invasiveness of renal cell carcinoma cells. *Mol. Cell. Biol.*, **35**, 2684–2698.
47. Gu,X., Wan,G., Yang,Y., Liu,Y., Yang,X., Zheng,Y., Jiang,L., Zhang,P., Liu,D., Zhao,W. *et al.* (2020) SLFN5 influences proliferation and apoptosis by upregulating PTEN transcription via ZEB1 and inhibits the purine metabolic pathway in breast cancer. *Am J Cancer Res*, **10**, 2832–2850.
48. Arslan,A.D., Sassano,A., Saleiro,D., Lisowski,P., Kosciuczuk,E.M., Fischietti,M., Eckerdt,F., Fish,E.N. and Platanias,L.C. (2017) Human SLFN5 is a transcriptional co-repressor of STAT1-mediated interferon responses and promotes the malignant phenotype in glioblastoma. *Oncogene*, **36**, 6006–6019.
49. Companioni Napoles,O., Tsao,A.C., Sanz-Anquela,J.M., Sala,N., Bonet,C., Pardo,M.L., Ding,L., Simo,O., Saqui-Salces,M., Blanco,V.P. *et al.* (2017) SCHLAFEN 5 expression correlates with intestinal metaplasia that progresses to gastric cancer. *J. Gastroenterol.*, **52**, 39–49.
50. Wan,G., Zhu,J., Gu,X., Yang,Y., Liu,Y., Wang,Z., Zhao,Y., Wu,H., Huang,G. and Lu,C. (2020) Human schlafen 5 regulates reversible epithelial and mesenchymal transitions in breast cancer by suppression of ZEB1 transcription. *Br. J. Cancer*, **123**, 633–643.
51. Wan,G., Liu,Y., Zhu,J., Guo,L., Li,C., Yang,Y., Gu,X., Deng,L.L. and Lu,C. (2019) SLFN5 suppresses cancer cell migration and invasion by inhibiting MT1-MMP expression via AKT/GSK-3beta/beta-catenin pathway. *Cell. Signal.*, **59**, 1–12.
52. Kabsch,W. (2010) Integration, scaling, space-group assignment and post-refinement. *Acta Crystallogr. D. Biol. Crystallogr.*, **66**, 133–144.
53. Kabsch,W. (2010) Xds. *Acta Crystallogr. D. Biol. Crystallogr.*, **66**, 125–132.

54. Grosse-Kunstleve,R.W. and Adams,P.D. (2003) Substructure search procedures for macromolecular structures. *Acta Crystallogr. Sect. D*, **59**, 1966–1973.
55. Terwilliger,T.C., Adams,P.D., Read,R.J., McCoy,A.J., Moriarty,N.W., Grosse-Kunstleve,R.W., Afonine,P.V., Zwart,P.H. and Hung,L.-W. (2009) Decision-making in structure solution using bayesian estimates of map quality: the PHENIX autosol wizard. *Acta Crystallogr. Sect. D*, **65**, 582–601.
56. Terwilliger,T.C., Grosse-Kunstleve,R.W., Afonine,P.V., Moriarty,N.W., Zwart,P.H., Hung,L.W., Read,R.J. and Adams,P.D. (2008) Iterative model building, structure refinement and density modification with the PHENIX autobuild wizard. *Acta Crystallogr. D. Biol. Crystallogr.*, **64**, 61–69.
57. Adams,P.D., Afonine,P.V., Bunkoczi,G., Chen,V.B., Davis,I.W., Echols,N., Headd,J.J., Hung,L.W., Kapral,G.J., Grosse-Kunstleve,R.W. *et al.* (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D. Biol. Crystallogr.*, **66**, 213–221.
58. Stein,N. (2008) CHAINSAW: a program for mutating pdb files used as templates in molecular replacement. *J. Appl. Crystallogr.*, **41**, 641–643.
59. Winn,M.D., Ballard,C.C., Cowtan,K.D., Dodson,E.J., Emsley,P., Evans,P.R., Keegan,R.M., Krissinel,E.B., Leslie,A.G., McCoy,A. *et al.* (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr. D. Biol. Crystallogr.*, **67**, 235–242.
60. McCoy,A.J., Grosse-Kunstleve,R.W., Adams,P.D., Winn,M.D., Storoni,L.C. and Read,R.J. (2007) Phaser crystallographic software. *J. Appl. Crystallogr.*, **40**, 658–674.
61. McCoy,A.J., Nicholls,R.A. and Schneider,T.R. (2013) SCEDS: protein fragments for molecular replacement in phaser. *Acta Crystallogr. Sect. D*, **69**, 2216–2225.
62. Afonine,P.V., Grosse-Kunstleve,R.W., Echols,N., Headd,J.J., Moriarty,N.W., Mustyakimov,M., Terwilliger,T.C., Urzhumtsev,A., Zwart,P.H. and Adams,P.D. (2012) Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. D. Biol. Crystallogr.*, **68**, 352–367.
63. Murshudov,G.N., Vagin,A.A. and Dodson,E.J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D*, **53**, 240–255.
64. Vagin,A.A., Steiner,R.A., Lebedev,A.A., Potterton,L., McNicholas,S., Long,F. and Murshudov,G.N. (2004) REFMAC5 dictionary: organization of prior chemical knowledge and guidelines for its use. *Acta Crystallogr. Sect. D*, **60**, 2184–2195.
65. Emsley,P., Lohkamp,B., Scott,W.G. and Cowtan,K. (2010) Features and development of coot. *Acta Crystallogr. Sect. D*, **66**, 486–501.
66. Pettersen,E.F., Goddard,T.D., Huang,C.C., Meng,E.C., Couch,G.S., Croll,T.I., Morris,J.H. and Ferrin,T.E. (2021) UCSF chimerax: structure visualization for researchers, educators, and developers. *Protein Sci.*, **30**, 70–82.
67. Zheng,S.Q., Palovcak,E., Armache,J.P., Verba,K.A., Cheng,Y. and Agard,D.A. (2017) MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods*, **14**, 331–332.
68. Punjani,A., Rubinstein,J.L., Fleet,D.J. and Brubaker,M.A. (2017) cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods*, **14**, 290–296.
69. Punjani,A., Zhang,H. and Fleet,D.J. (2020) Non-uniform refinement: adaptive regularization improves single-particle cryo-EM reconstruction. *Nat. Methods*, **17**, 1214–1221.
70. Jumper,J., Evans,R., Pritzel,A., Green,T., Figurnov,M., Ronneberger,O., Tunyasuvunakool,K., Bates,R., Žídek,A., Potapenko,A. *et al.* (2021) Highly accurate protein structure prediction with alphafold. *Nature*, **596**, 583–589.
71. Lowary,P.T. and Widom,J. (1998) New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning. *J. Mol. Biol.*, **276**, 19–42.
72. Schneider,C.A., Rasband,W.S. and Eliceiri,K.W. (2012) NIH image to imagej: 25 years of image analysis. *Nat. Methods*, **9**, 671–675.
73. Murai,J., Zhang,H., Pongor,L., Tang,S.W., Jo,U., Moribe,F., Ma,Y., Tomita,M. and Pommier,Y. (2020) Chromatin remodeling and immediate early gene activation by SLFN11 in response to replication stress. *Cell Rep.*, **30**, 4137–4151.
74. Jeong,S.Y., Kumagai,A., Lee,J. and Dunphy,W.G. (2003) Phosphorylated claspin interacts with a phosphate-binding site in the kinase domain of chk1 during ATR-mediated activation. *J. Biol. Chem.*, **278**, 46782–46788.
75. Zhang,X., Wu,J., Du,F., Xu,H., Sun,L., Chen,Z., Brautigam,C.A., Zhang,X. and Chen,Z.J. (2014) The cytosolic DNA sensor cGAS forms an oligomeric complex with DNA and undergoes switch-like conformational changes in the activation loop. *Cell Rep.*, **6**, 421–430.
76. Tanaka,Y. and Chen,Z.J. (2012) STING specifies IRF3 phosphorylation by TBK1 in the cytosolic DNA signaling pathway. *Sci. Signal*, **5**, ra20.
77. Walker,J.E., Saraste,M., Runswick,M.J. and Gay,N.J. (1982) Distantly related sequences in the alpha-and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. *EMBO J.*, **1**, 945–951.
78. Yue,T., Zhan,X., Zhang,D., Jain,R., Wang,K.W., Choi,J.H., Misawa,T., Su,L., Quan,J., Hildebrand,S. *et al.* (2021) SLFN2 protection of tRNAs from stress-induced cleavage is essential for t cell-mediated immunity. *Science*, **372**, eaba4220 .
79. Chen,J. and Kuhn,L.A. (2019) Deciphering the three-domain architecture in schlafens and the structures and roles of human schlafen12 and serpinB12 in transcriptional regulation. *J. Mol. Graph. Model.*, **90**, 59–76.

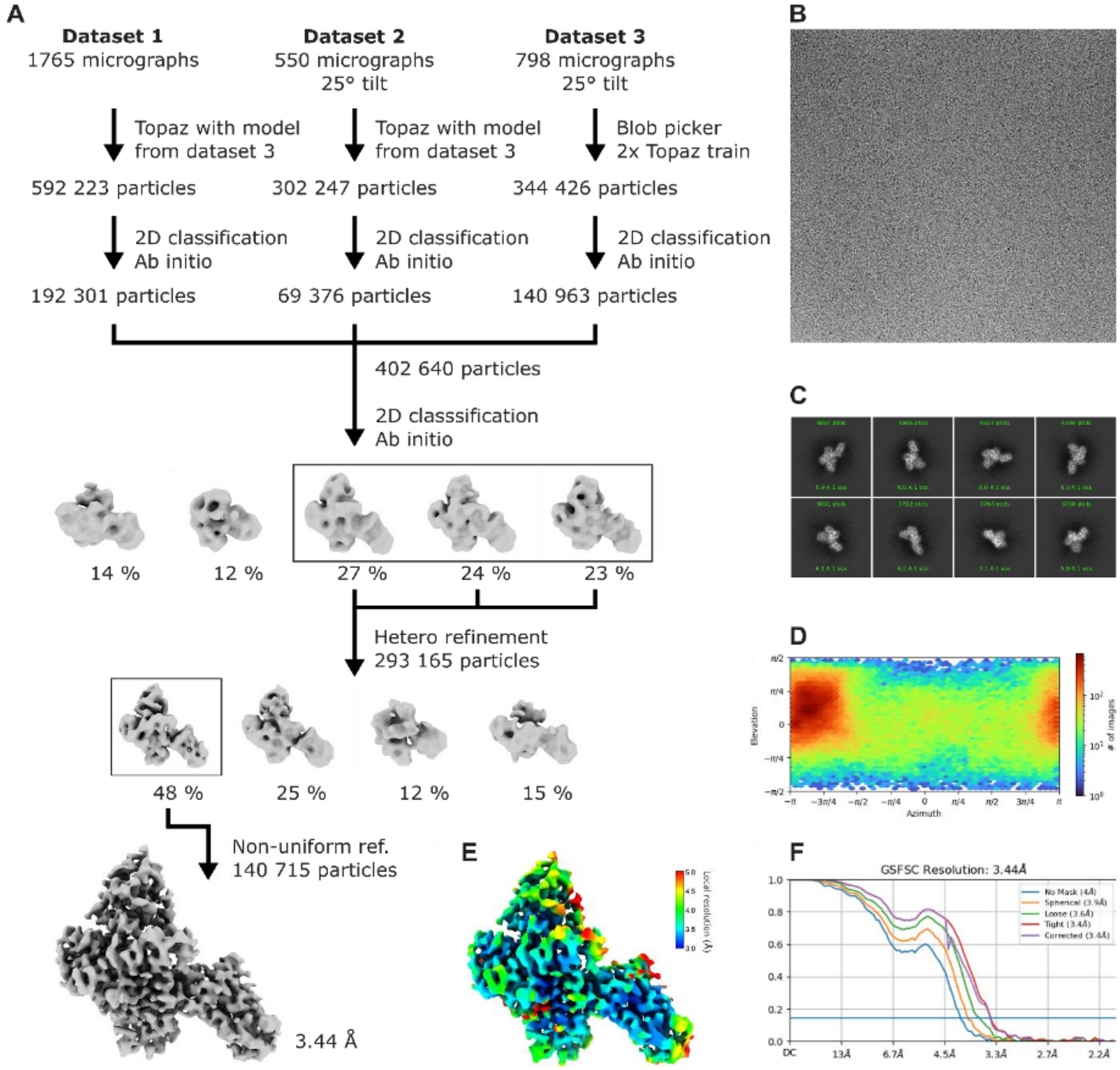
SUPPLEMENTARY DATA

Structural and biochemical characterization of human Schlafen 5

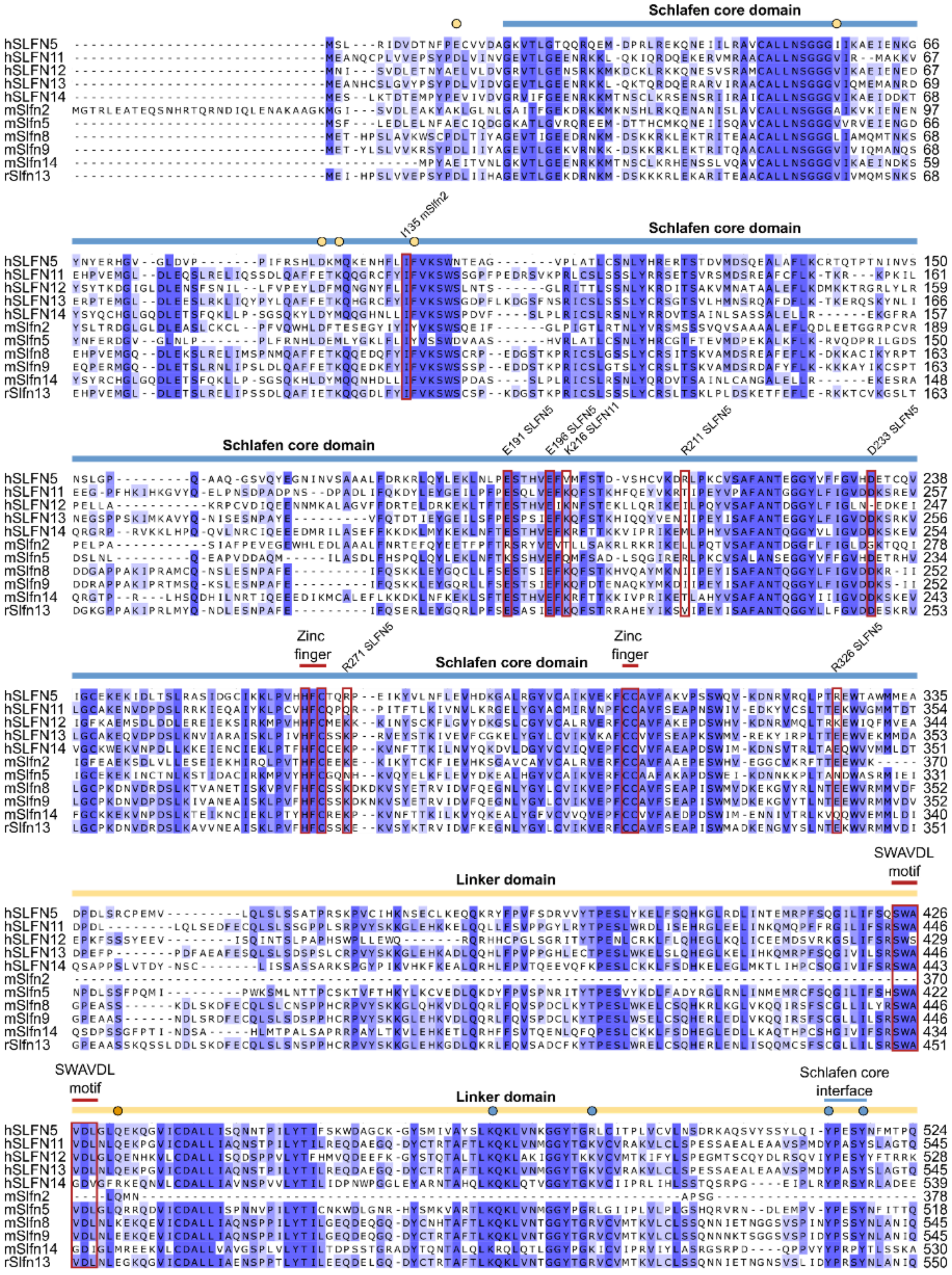
Felix J. Metzner, Elisabeth Huber, Karl-Peter Hopfner and Katja Lammens

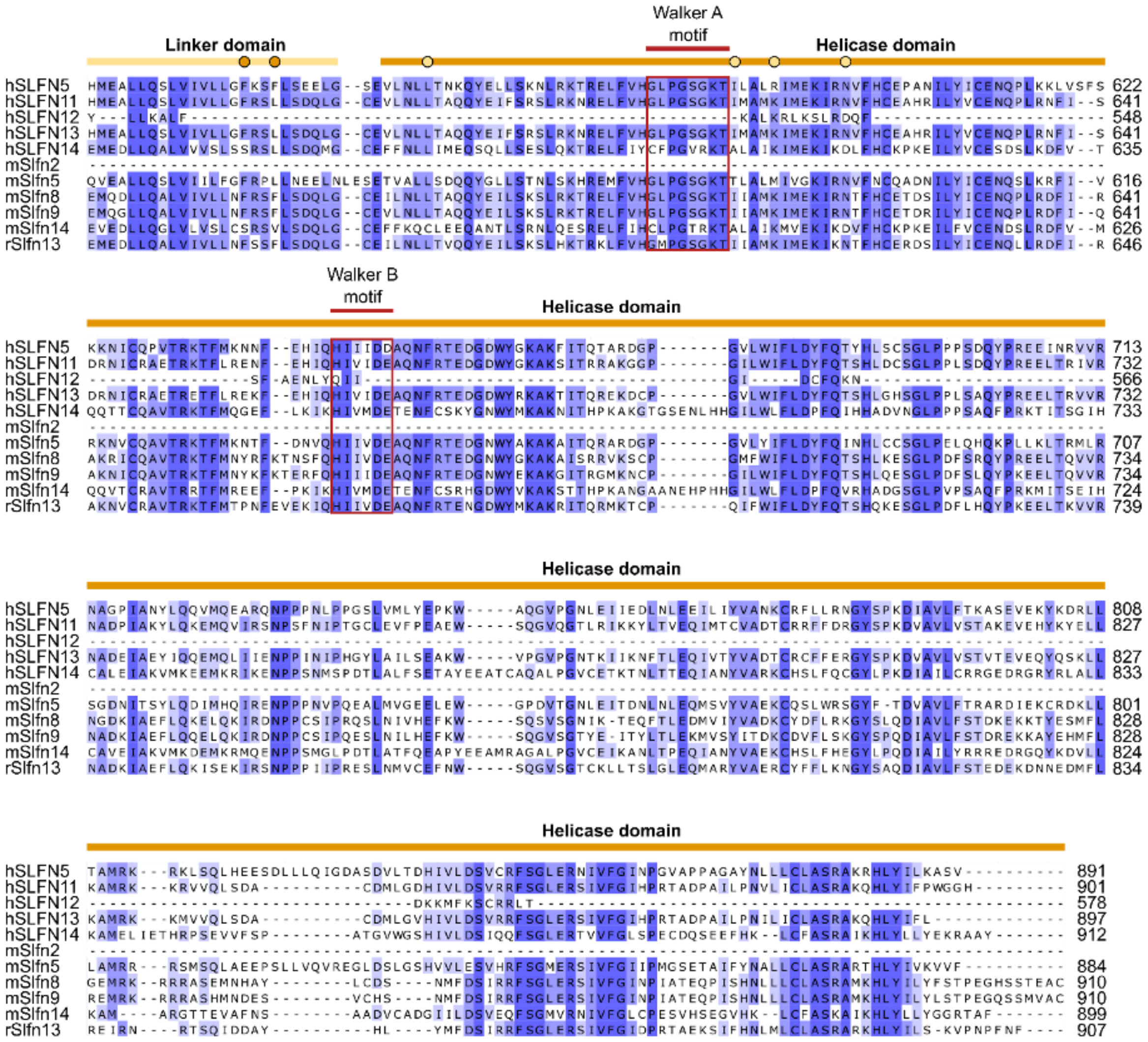


Supplementary Figure S1. Purification of full-length SLFN5. **(A)** Schematic of expression and purification of full-length SLFN5. **(B)** SEC elusion profile from a Superdex 200 5/150 column. Absorbance at 260 nm (red) and 280 nm (blue) is shown. **(C)** SDS-PAGE of purified full-length SLFN5 (stained with InstantBlue).

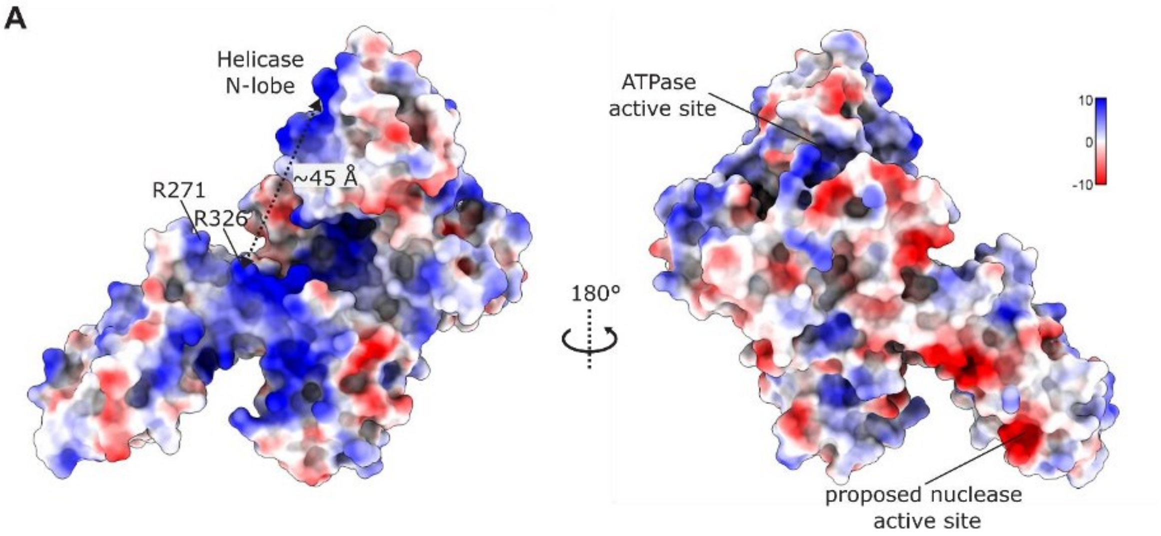


Supplementary Figure S2. Cryo-EM data analysis. **(A)** Cryo-EM data processing workflow of full-length SLFN5 using cryoSPARC v3.2.0 (1). **(B)** Representative micrograph of full-length SLFN5. **(C)** Representative classes of a 2D classification of the particles used for the final SLFN5 reconstruction. **(D)** Angular distribution of the particles used for the final SLFN5 reconstruction. **(E)** Visualization of local resolution calculated in cryoSPARC. Blue indicates higher resolution and red indicates lower resolution. **(F)** Gold-standard Fourier shell correlation (FSC) curves of the final SLFN5 reconstruction. The blue line indicates the 0.143 cutoff criterion, indicating a nominal resolution of 3.44 Å.

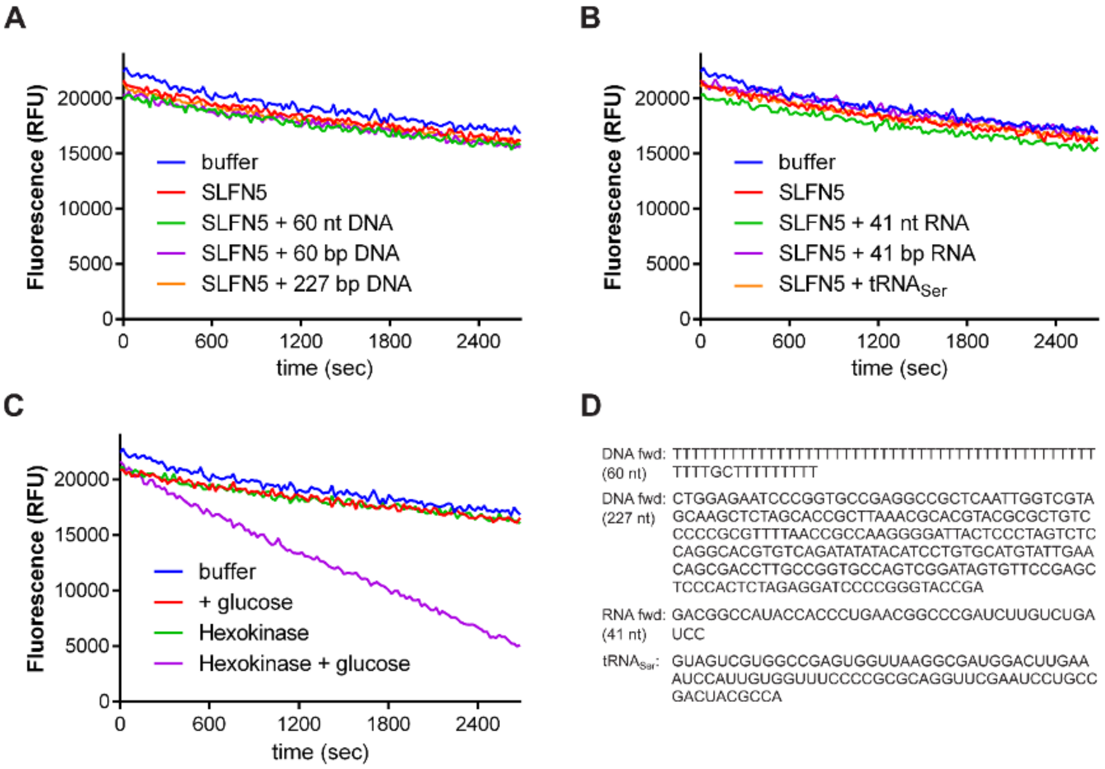

Schlafen core domain
hSLFN5
hSLFN11
hSLFN12
hSLFN13
hSLFN14
mSlfn2
mSlfn5
mSlfn8
mSlfn9
mSlfn14
rSlfn13
I135 mSlfn2
E191 SLFN5
E196 SLFN5
K216 SLFN11
R211 SLFN5
D233 SLFN5
Zinc finger
R271 SLFN5
R326 SLFN5
Linker domain
SWAVDL motif
Schlafen core interface



Supplementary Figure S3. Multiple sequence alignment of SLFN5 and selected Schlafen proteins from human (SLFN5, SLFN11, SLFN12, SLFN13, SLFN14), mouse (Slfn2, Slfn5, Slfn8, Slfn9, Slfn14) and rat (Slfn13). The alignment was calculated using MAFFT (2). Residues are colored according to percentage identity (dark blue = more conserved, white = less conserved). The Schlafen core domain is indicated in blue, the linker domain in yellow and the helicase domain in orange. The residues of the zinc finger, the SWAVDL motif, the Walker A and B motif and the position of the *elektra* mutation are highlighted in red. Residues involved in the Schlafen core domain – linker domain and the linker domain – helicase domain interaction are marked with circles.



Supplementary Figure S4. Electrostatic surface potential of SLFN5. **(A)** Coulombic electrostatic surface potential of SLFN5 colored from red (-10 kcal/mol**e*) to blue (10 kcal/mol**e*). Values were calculated in UCSF ChimeraX (3). R271 and R326, the distance between R326 and the helicase domain and the putative nuclease and ATPase active sites are indicated.



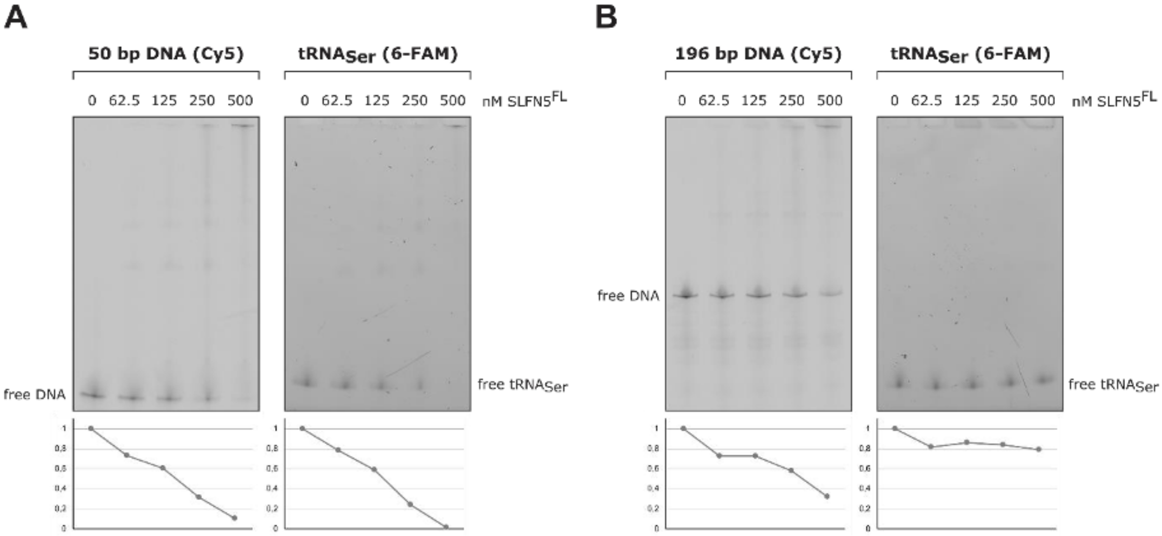
DNA fwd: TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT
(60 nt) TTTTGCTTTTTTTTT

DNA fwd: CTGGAGAATCCCGGTGCCGAGGCCGCTCAATTGGTCGTA
(227 nt) GCAAGCTCTAGCACCGCTTAAACGCACGTACGCGCTGTC
CCCCGCGTTTTAACCGCCAAGGGGATTACTCCCTAGTCTC
CAGGCACGTGTCAGATATATACATCCTGTGCATGTATTGAA
CAGCGACCTTGCCGGTGCCAGTCGGATAGTGTTCCGAGC
TCCCACTCTAGAGGATCCCCGGGTACCGA

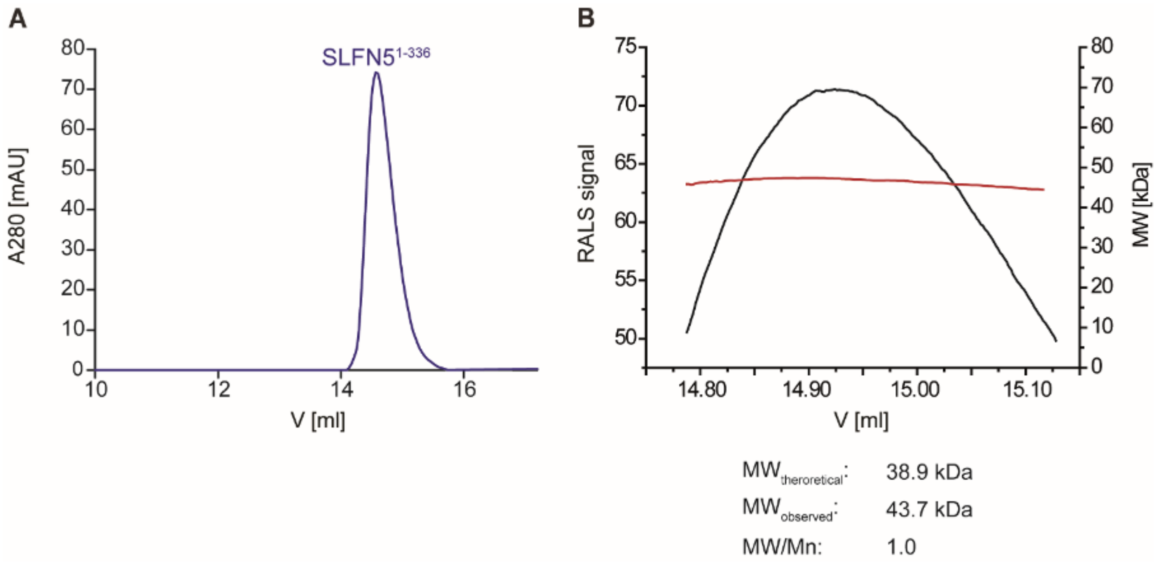
RNA fwd: GACGGCCAUACCACCCUGAACGGCCCGAUCUUGUCUGA
(41 nt) UCC

tRNA$_{Ser}$: GUAGUCGUGGCCGAGUGGUUAAGGCGAUGGACUUGAA
AUCCAUUGUGGUUUCCCCGCGCAGGUUCGAAUCCUGCC
GACUACGCCA

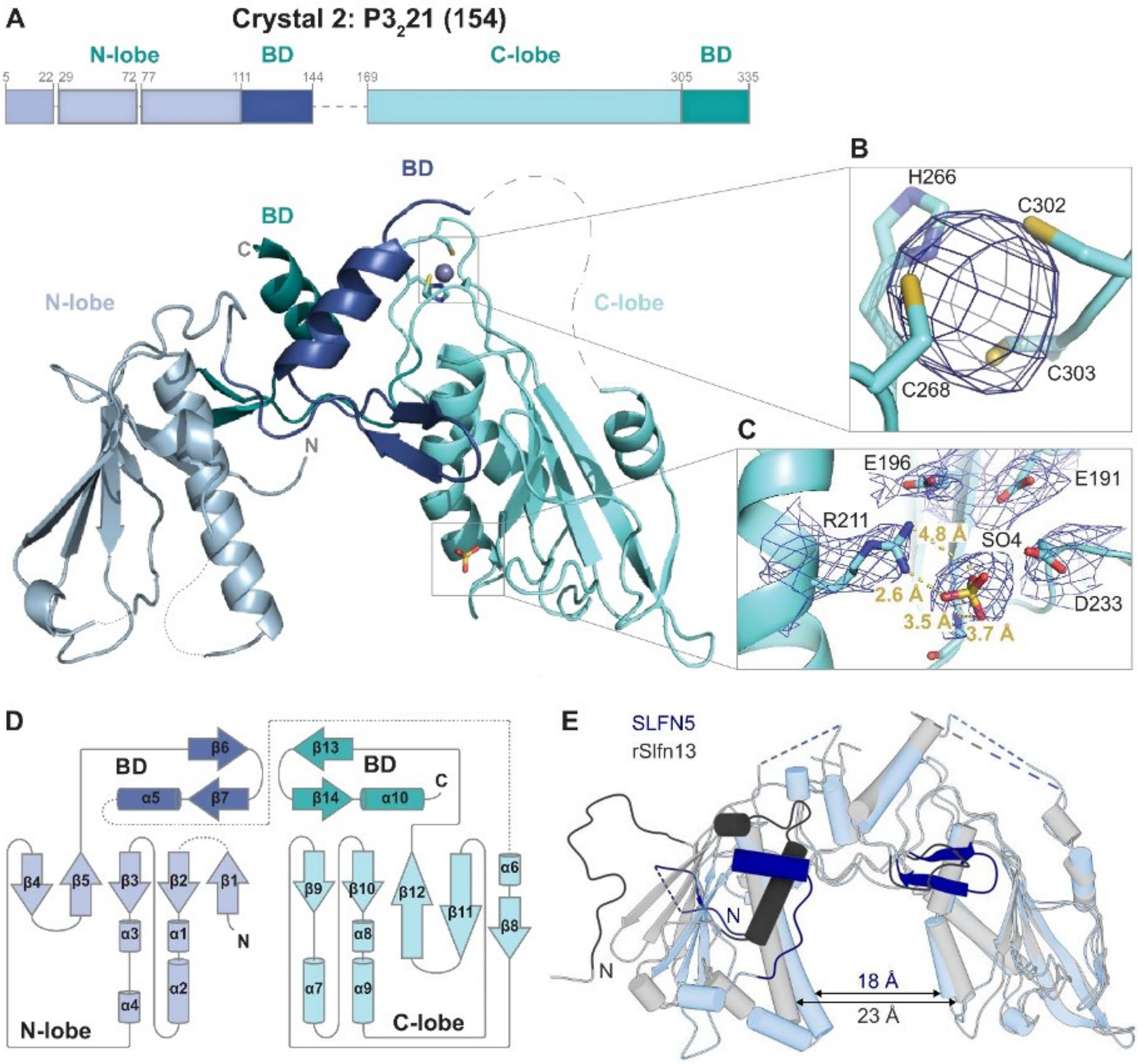
Supplementary Figure S5. Fluorescence-based ATPase assay of SLFN5. Assay was performed in triplicates but a single replicate is shown due to visualization purposes. Data of buffer and SLFN5 is shown multiple times. **(A)** SLFN5 (50 nM) ATPase activity basal or in presence of ssDNA (60 nt, 150 nM) or dsDNA (60 bp and 227 bp, 150 nM). **(B)** SLFN5 (50 nM) ATPase activity basal or in presence of ssRNA (41 nt, 150 nM), dsRNA (41 bp, 150 nM) or tRNA$_{Ser}$ (150 nM). **(C)** Positive control: ATPase activity of hexokinase (1 nM) basal or in presence of glucose (200 µM). **(D)** Nucleotide sequences of substrates used in the ATPase assay (only forward sequences shown).



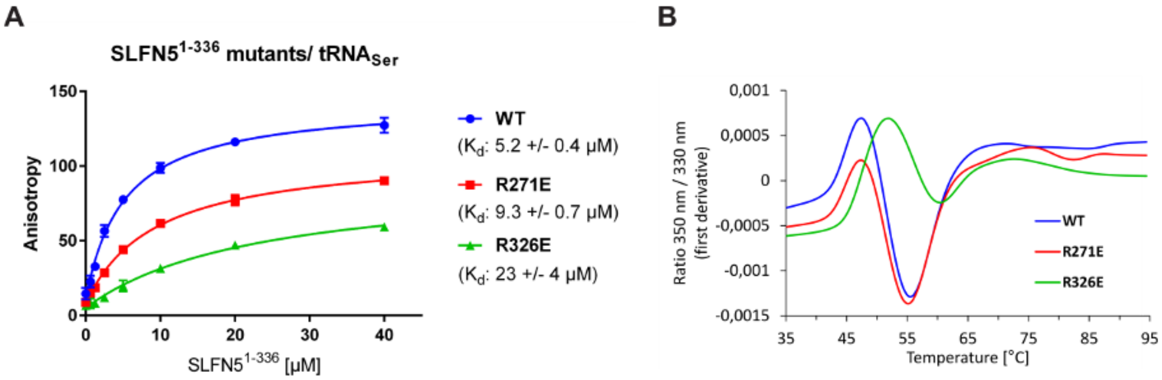
Supplementary Figure S6. Analysis of dsDNA and tRNA binding properties by electrophoretic mobility shift assay (EMSA) of SLFN5. **(A)** Competition EMSA of SLFN5 with Cy5 labeled 50 bp DNA and 6-FAM labeled tRNA$_{\text{Ser}}$ (20 nM each). The same gel was scanned twice, using the Cy5 (left) or FAM (right) channel. Bottom: Quantification of free DNA and tRNA bands using ImageJ (4). First lane was set to 1. **(B)** Competition EMSA of SLFN5 with Cy5 labeled 196 bp DNA and 6-FAM labeled tRNA$_{\text{Ser}}$ (20 nM each). The same gel was scanned twice, using the Cy5 (left) or FAM (right) channel. Bottom: Quantification of free DNA and tRNA bands using ImageJ. First lane was set to 1.



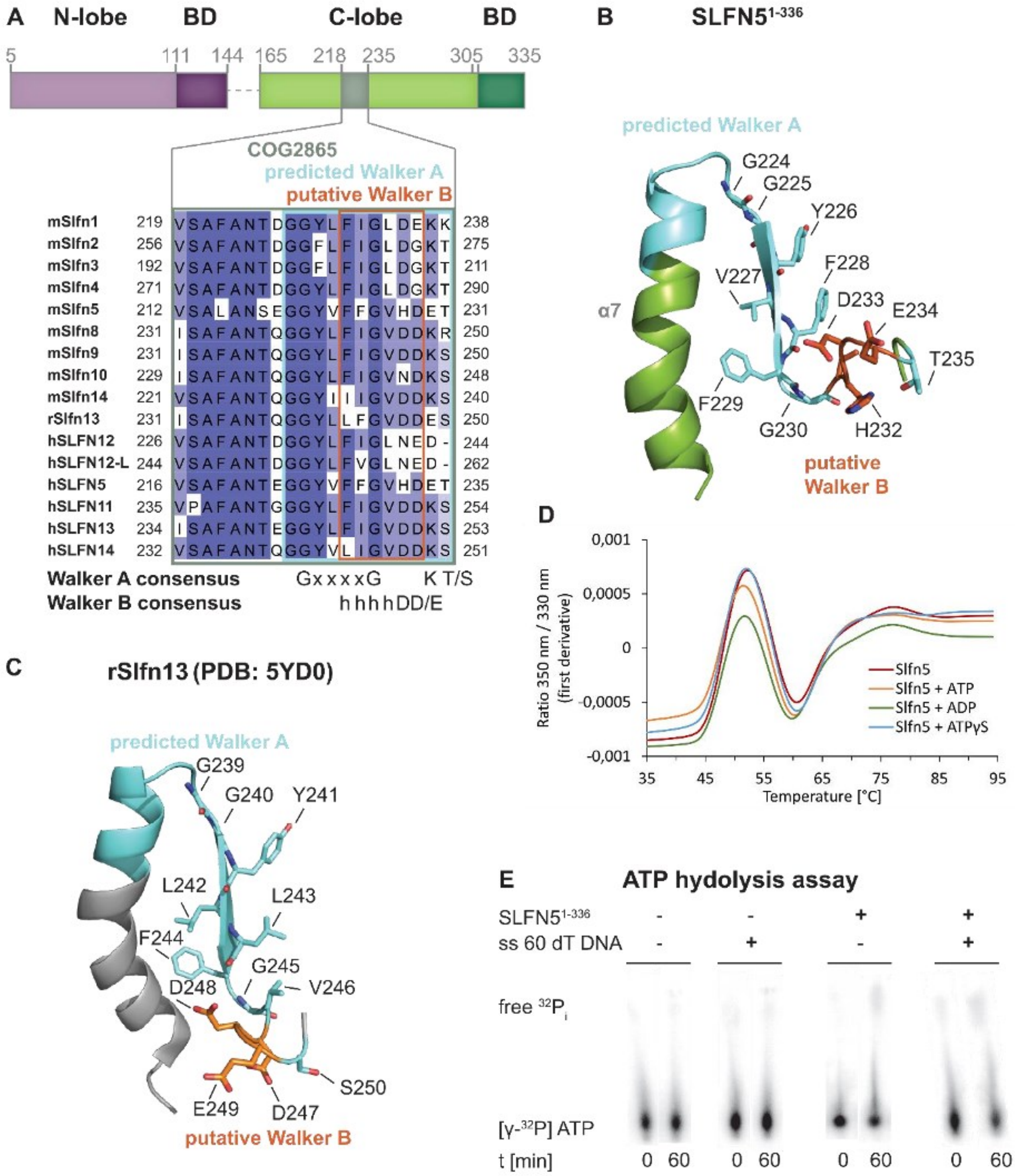
Supplementary Figure S7. Molecular weight determination of SLFN5$^{1\text{-}336}$ by SEC-RALS. **(A)** Analytical size exclusion chromatography using a Superdex 200 10/300 column. SLFN5$^{1\text{-}336}$ elutes in a single peak at approximately 15 ml. The blue line indicates the absorption at 280 nm. (**B)** Right-angle light scattering (RALS) coupled to SEC. The SLFN5$^{1\text{-}336}$ sample is monodisperse and the observed molecular weight was calculated to be 43.7 kDa, which corresponds to a monomer in solution. MW: molecular weight, Mn: number average molecular weight.



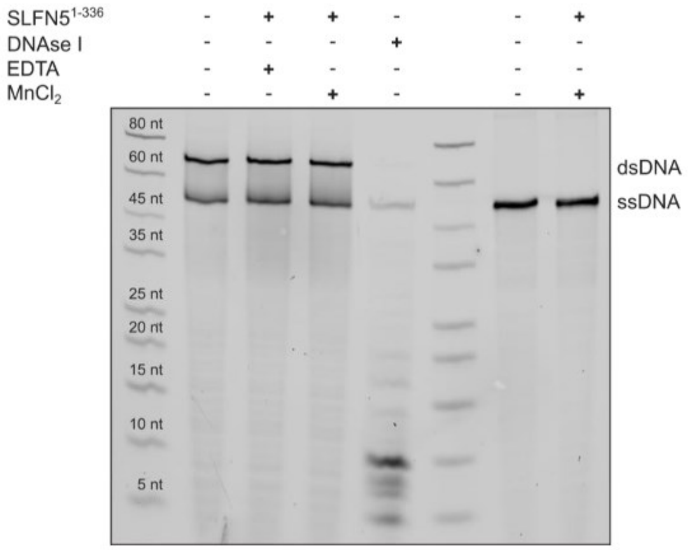
Supplementary Figure S8. Crystal structure of SLFN5$^{1-336}$ crystallized in space group P3$_2$21. **(A)** Schematic overview and structure of SLFN5$^{1-336}$ (P3$_2$21). N-lobe: light blue, N-bridging domain (BD): blue, C-lobe: cyan, C-BD: dark cyan. **(B)** Close-up view of the zinc finger motif. The anomalous electron density map is colored in dark blue and contoured at σ=5. **(C)** Close-up view of the bound sulfate and neighboring residues. The 2F$_0$-F$_c$ electron density map is colored in dark blue and contoured at σ=1. The interaction distances are indicated in yellow. **(D)** Secondary structure topology diagram of SLFN5$^{1-336}$. **(E)** Overlay of SLFN5$^{1-336}$ (blue) and rSlfn13$^{14-353}$ (gray, PDB: 5YD0). Selected regions showing structural differences are highlighted.



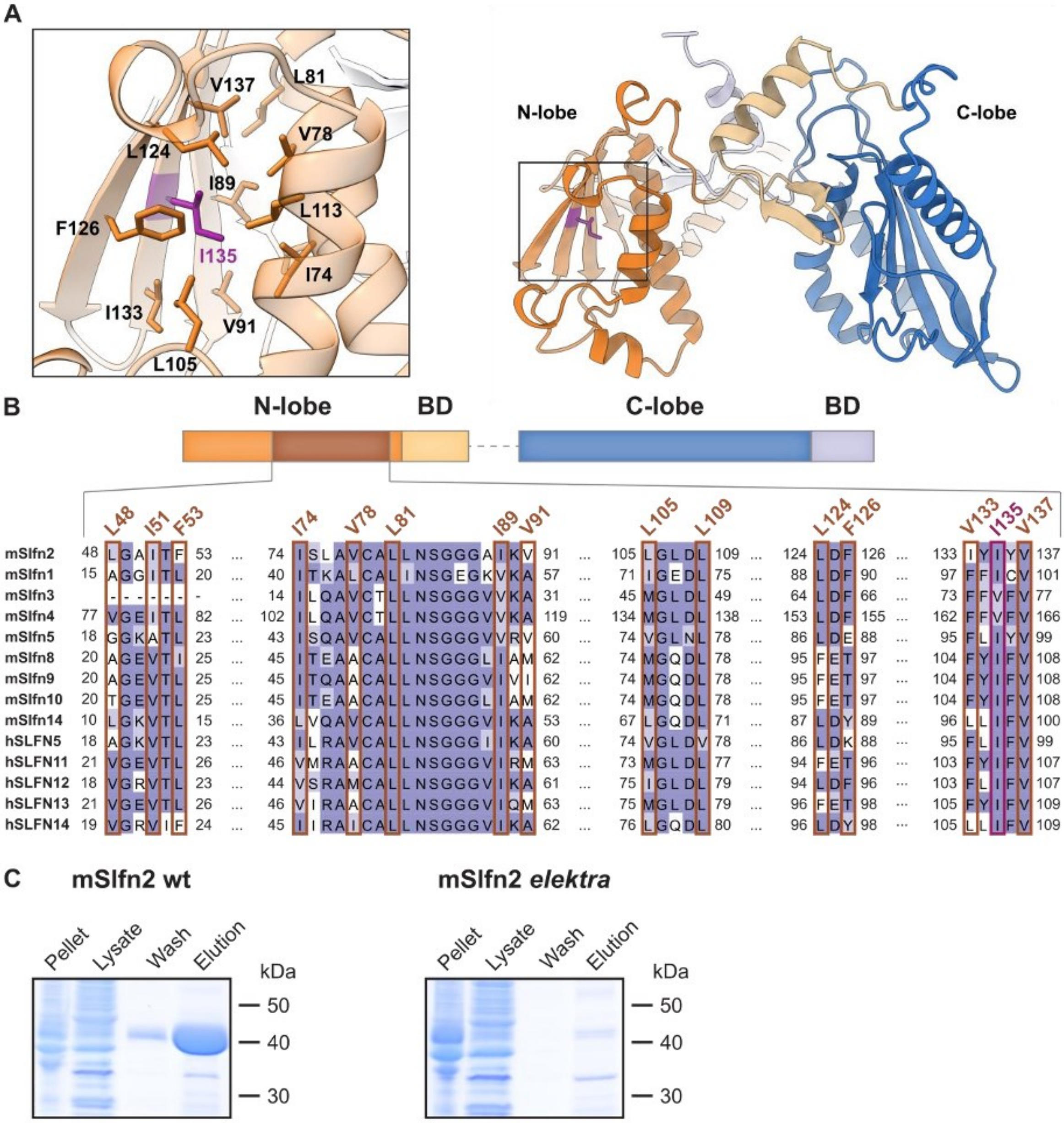
Supplementary Figure S9. Analysis of SLFN5$^{1-336}$ tRNA$_{Ser}$ binding properties. **(A)** Fluorescence anisotropy assay to monitor the binding of SLFN5$^{1-336}$ (blue), SLFN5$^{1-336}$ R271E (red) and SLFN5$^{1-336}$ R326E (green) to tRNA$_{Ser}$. The final protein concentrations were 0, 0.625, 1.25, 2.5, 5, 10, 20 and 40 µM. The data were fit to a 1 to 1 binding equation. Error bars represent the standard deviation from three experiments. **(B)** NanoDSF measurements of SLFN5$^{1-336}$ (blue), SLFN5$^{1-336}$ R271E (red) and SLFN5$^{1-336}$ R326E (green).



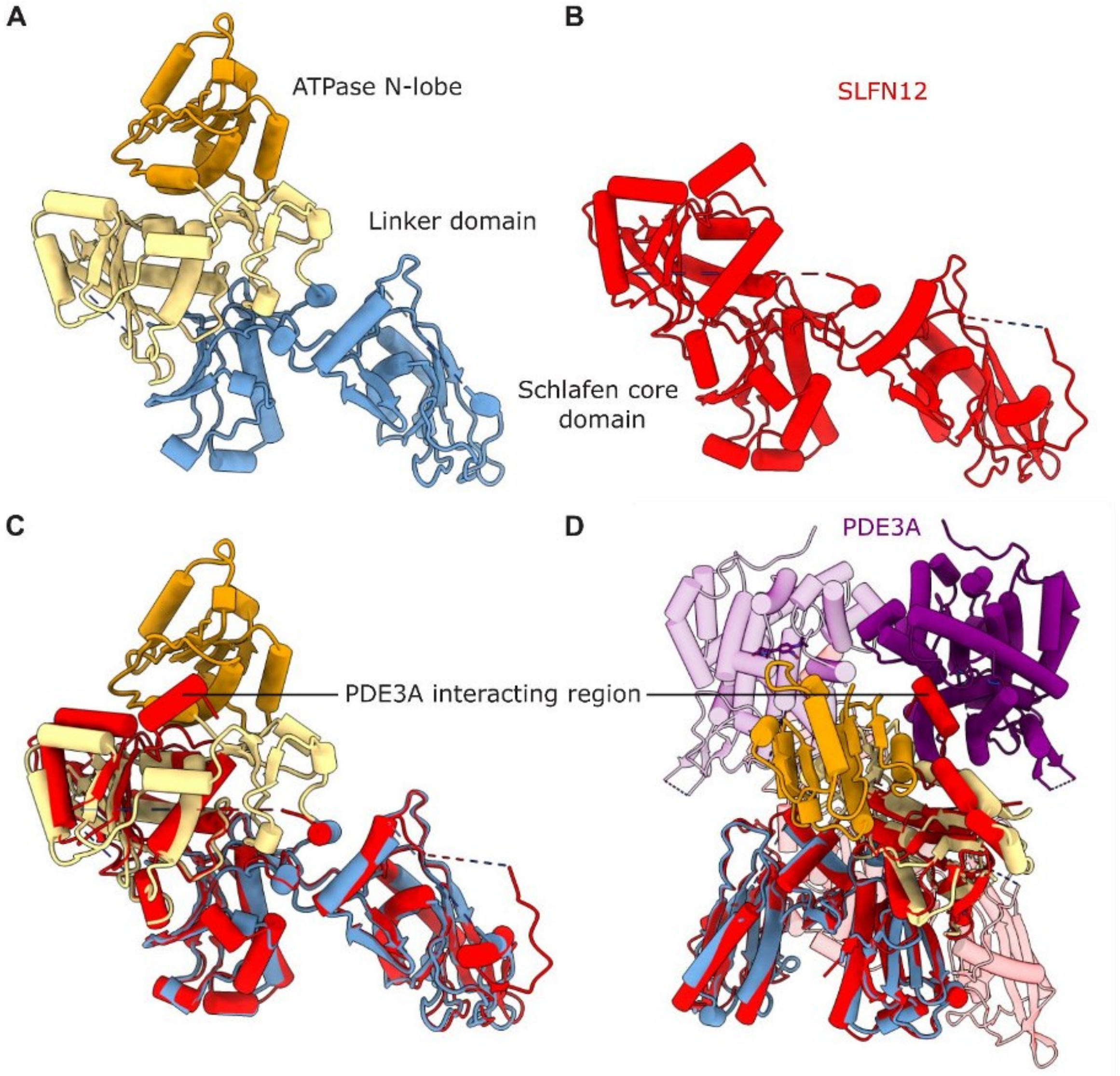
Supplementary Figure S10. Structural and functional analysis of the predicted ATPase site within the Slfn core domain. **(A)** Schematic overview of the domain organization of SLFN5$^{1-336}$ and multiple sequence alignment of the predicted Walker A/ Walker B motifs of murine and human Schlafen members. **(B-C)** Structural framework of the predicted Walker A/ Walker B motifs of SLFN5 (B) and rSlfn13 (PDB: 5YD0) (C). The predicted Walker A and B motifs are colored in aquamarine and orange, respectively. **(D)** NanoDSF measurements of SLFN5$^{1-336}$ in presence of different nucleotides or without nucleotide. **(E)** ATP hydrolysis assay of SLFN5$^{1-336}$ in presence or absence of a single-stranded 60-mer DNA substrate.



Supplementary Figure S11. Analysis of nuclease activity of SLFN5$^{1-336}$ on DNA. SLFN5$^{1-336}$ shows no nuclease activity on 50 bp dsDNA (lanes 2,3) and 50 nt ssDNA (lane 6). DNase I was used as positive control (lane 4).



Supplementary Figure S12. Structural consequences of the *elektra* mutation (I135N) in mSlfn2. **(A)** AlphaFold model of full-length mSlfn2 (5) and close-up view of the hydrophobic patch region harboring the *elektra* mutation (I135, colored in purple). Amino acids are colored by heteroatom. **(B)** Schematic view of domain architecture of mSlfn2 and multiple sequence alignment of human and murine Schlafen family members. The conserved residues that form the hydrophobic core are highlighted in orange. **(C)** Ni-NTA purification of mSlfn2 wt and mSlfn2 *elektra* (I135N). Pellet, lysate, wash and elution fractions were subject to SDS-PAGE (stained with Coomassie blue). Similar volumes were loaded for both purifications. mSlfn2 has a molecular weight of approximately 42.5 kDa.



Supplementary Figure S13. Structural comparison of SLFN5 and SLFN12 (PDB: 7LRD). **(A)** Cartoon representation of SLFN5. The Schlafen core domain is depicted in blue, the linker domain in yellow and the helicase N-lobe in orange. **(B)** Cartoon representation of SLFN12 depicted in red (PDB: 7LRD). **(C)** Overlay of SLFN5 with SLFN12 (PDB: 7LRD). **(D)** Overlay of SLFN5 with the SLFN12-PDE3A complex (PDB: 7LRD). SLFN12 is depicted in red and PDE3A in purple.

Supplementary Table S1. Data collection, 3D reconstruction and refinement statistics

	H. sapiens **SLFN5 dataset I**	***H. sapiens*** **SLFN5 dataset II**	***H. sapiens*** **SLFN5 dataset III**
Data collection and processing			
Magnification	130 000	130 000	130 000
Voltage (kV)	300	300	300
Electron exposure (e^-/Å^2)	41.2	41.2	40.9
Defocus range (µm)	-1.1 to -2.9	-1.1 to -2.9	-1.1 to -2.9
Pixel size (Å)	1.046	1.046	1.046
Tilt angle (°)	0	25	25
Symmetry imposed	C1	C1	C1
Initial particle images (no.)	592 223	302 247	344 426
Data processing (combined data)			
Final particle images (no.)	140 715		
Map resolution (Å)/ FSC threshold	3.44/ 0.143		
Refinement			
Initial model used (PDB code)	6RI1		
Model resolution (Å)	3.7		
FSC threshold	0.5		
Model resolution range (Å)	3.2 -3.7		
Map-sharpening *B* factor (Å^2)	175.9		
Model composition			
Nonhydrogen	1 Zn		
Protein residues	643		
R.m.s. deviations			
Bond lengths (Å)	0.005		
Bond angles (°)	0.709		
Validation			
MolProbity score	2.19		
Clashscore	12.5		
Poor rotamers (%)	0.51		
Ramachandran plot			
Favored (%)	92.62		
Allowed (%)	7.38		
Disallowed (%)	0.0		
PDB ID	7PPJ		

Supplementary Table S2. Data collection and refinement statistics

Data collection and processing	**SLFN5$^{1\text{-}336}$**	**SLFN5$^{1\text{-}336}$**
Wavelength [Å]	1.0	1.28
Space group	P 1 2_1 1	P 3_2 2 1
Unit cell dimensions		
a, b, c [Å]	49.88, 64.04, 104.58	101.1, 101.1, 114.5
α, β, γ [°]	90, 93, 90	90, 90, 120
Resolution range [Å]	45.9 – 1.85 (1.92 – 1.85)	19.8 – 3.4
R_{meas}	9.0 (130.6)	15.21 (125.1)
I/σI	12.86 (1.43)	17.1 (2.2)
Completeness [%]	98.8 (97.6)	97.5 (83.7)
Redundancy	6.6 (6.3)	18.2 (14.9)
Solvent content [%]	44	72
Matthews coefficient [A^3/Da]	2.2	4.44
Refinement		
Number reflections	55695 (5454)	9144 (779)
R_{work}/R_{free}	0.18/0.21	0.21/0.26
Number of atoms		
Protein	4992	2483
Ligands	10	32
Water	289	
Average B-factor [Å]	42.361	64.22
RMSD		
Bond lengths [Å]	0.007	0.013
Bond angles [°]	0.85	1.41
Ramachandran plot		
Favored [%]	97.66	84.04
Allowed [%]	2.34	13.36
Outliers [%]	0.0	2.61
PDB ID	7Q3Z	6RR9

Supplementary Table S3. Oligonucleotides (5' to 3')

Purpose	Name	Sequence	Modification
Cloning	SLFN5_R271E_fwd	CATCATTTCTGCACACAGGAGCCTGAGATAAAATATGTC	
	SLFN5_R271E_rev	GACATATTTTATCTCAGGCTCCTGTGTGCAGAAATGATG	
	SLFN5_R326E_fwd	GACAATTGCCCACAGAAGAATGGACTGCTTG	
	SLFN5_R326E_rev	CAAGCAGTCCATTCTTCTGTGGGCAATTGTC	
	mSlfn2_I135N_fwd	GAGCGAAGGTTACATCTATAACTATGTTAAATCCTGGAGC	
	mSlfn2_I135N_rev	GCTCCAGGATTTAACATAGTTATAGATGTAACCTTCGCTC	
Biochemistry	5 dT	TTTTT	5'-FAM
	10 dT	TTTTTTTTTT	5'-FAM
	20 dT	TTTTTTTTTTTTTTTTTTTT	5'-FAM
	30 dT	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTT	5'-FAM
	60 dT	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT TTTT	5'-FAM
	60bp_A/T_GC_fwd	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTGCTTTTT TTTT	5'-FAM
	60bp_A/T_GC_rev	AAAAAAAAAGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAA	
	30 rU	UUUUUUUUUUUUUUUUUUUUUUUUUUUUUU	5'-FAM
	60 rU	UUUUUUUUUGCUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUU UUUUUUUUUUUU	5'-FAM
	tRNA(Ser)	GUAGUCGUGGCCGAGUGGUUAAGGCGAUGGACUUGAAAUCCAUUG UGGUUUCCCCGCGCAGGUUCGAAUCCUGCCGACUACGCCA	5'-FAM
	41_nt_RNA_fwd	GACGGCC AUACCA CCC UGAACGGCCC GAUCUUGUCUGAUCC	5'-FAM
	41_nt_RNA_rev	GGAUCAGACAAGAUCGGGCCGUUCAGGGUGGUAUGGCCGUC	
	50_bp_DNA_fwd	AATTGGTCGTAGCAAGCTCTAGCACCGCTTAAACGCACGTACGCGCTGTC	5'-Cy5
	50_bp_DNA_rev	GACAGCGCGTACGTGCGTTTAAGCGGTGCTAGAGCTTGCTACGACCAATT	
	0N80 _fwd	CTGGAGAATCCCGGTGCCGAGG	5'-FAM
	0N80 _rev	TCGGTACCCGGGGATCCTCTAGAGTG	
	0N49 _fwd	CTGGAGAATCCCGGTGCCGAGG	
	0N49 _rev	TCGGAACACTATCCGACTGGCACC	5'-Cy5
	0N49/80 template	CTGGAGAATCCCGGTGCCGAGGCCGCTCAATTGGTCGTAGCAAGCTCTAG CACCGCTTAAACGCACGTACGCGCTGTCCCCCGCGTTTTAACCGCCAAGG GGATTACTCCCTAGTCTCCAGGCACGTGTCAGATATATACATCCTGTGCAT GTATTGAACAGCGACCTTGCCGGTGCCAGTCGGATAGTGTTCCGAGCTCC CACTCTAGAGGATCCCCGGGTACCGA	

REFERENCES

1. Punjani, A., Rubinstein, J.L., Fleet, D.J. and Brubaker, M.A. (2017) cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat Methods*, **14**, 290-296.
2. Katoh, K., Misawa, K., Kuma, K. and Miyata, T. (2002) MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res*, **30**, 3059-3066.
3. Pettersen, E.F., Goddard, T.D., Huang, C.C., Meng, E.C., Couch, G.S., Croll, T.I., Morris, J.H. and Ferrin, T.E. (2021) UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Sci*, **30**, 70-82.
4. Schneider, C.A., Rasband, W.S. and Eliceiri, K.W. (2012) NIH Image to ImageJ: 25 years of image analysis. *Nat Methods*, **9**, 671-675.
5. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A. *et al.* (2021) Highly accurate protein structure prediction with AlphaFold. *Nature*, **596**, 583-589.

2.2. Mechanistic understanding of human SLFN11

Felix J. Metzner*, Simon J. Wenzl*, Michael Kugler*, Stefan Krebs, Karl-Peter Hopfner, Katja Lammens (2022) Mechanistic understanding of human SLFN11. *Nature Communications*, Volume 13, Article 5464.

* These authors contributed equally.

DOI: 10.1038/s41467-022-33123-0
URL: https://www.nature.com/articles/s41467-022-33123-0

Summary

In this publication, we present cryo-EM structures of the human SLFN11 apoenzyme, SLFN11 bound to tRNA and in complex with ssDNA. SLFN11 forms a ring-shaped homodimer with interfaces between the Slfn core domains and the helicase domains. Compared to the structure of SLFN5, the SLFN11 helicase domain is rotated by almost 180°. The SLFN11 dimer binds tRNA at the positively charged central channel formed by the Slfn core domains. The endonuclease active sites are located in this central groove in close proximity to the tRNA molecule and SLFN11 cleaves type II tRNA^{Ser} more efficiently than type I tRNA^{Met}. The endonuclease reaction is Mn^{2+}-dependent and the cleavage site is located 10 nt from the tRNA 3′ end. Residues that are essential for endonuclease activity are identified by structure-guided mutagenesis. Dimer interface mutants as well as a trans-complementation assay demonstrate that the SLFN11 dimer is the nuclease active species. SLFN11 binds and is stabilized by ssDNA but not by dsDNA. The structure reveals that each helicase domain of the SLFN11 dimer binds a stretch of five nucleotides of ssDNA that are pointing in opposing 5′ to 3′ directions. A phosphomimetic mutant of a previously described phosphorylation site within the helicase domain abolishes the affinity to ssDNA. This suggests a functional connection between ssDNA binding to the helicase domain and the nuclease activity of the Slfn core domain. In contrast to SLFN5, SLFN11 shows no ATP binding in vitro. The dimeric structure of SLFN11 reveals an arrangement of the inter-domain region that sterically blocks the ATP binding site. This suggests an autoinhibited conformation of the helicase domain. Taken together, our data reveal the structure of full-length human SLFN11, give detailed insights into tRNA recognition and processing, ssDNA binding, as well as its regulation by dimerization and phosphorylation.

Author contributions

I developed purification strategies for full-length human SLFN5 and SLFN11. Simon Wenzl, Michael Kugler and I conducted protein expression and purification. I performed biochemical and biophysical assays, including nuclease assays, tRNA and DNA binding studies, mass photometry measurements as well as ATP binding and hydrolysis assays together with Simon Wenzl and Michael Kugler. Together with Simon Wenzl and Michael Kugler, I prepared grids for cryo-EM analysis. Cryo-EM data collection and processing was done by me, Katja Lammens, Michael Kugler and Simon Wenzl. Katja Lammens, Michael Kugler and I build the SLFN11 structures. I prepared the manuscript together with Katja Lammens and Michael Kugler and input from all other authors.

nature communications

Article

Mechanistic understanding of human SLFN11

Received: 1 April 2022

Accepted: 1 September 2022

Felix J. Metzner[1,2], **Simon J. Wenzl**[1,2], **Michael Kugler**[1,2], **Stefan Krebs**[1], **Karl-Peter Hopfner**[1] & **Katja Lammens**[1]

Schlafen 11 (SLFN11) is an interferon-inducible antiviral restriction factor with tRNA endoribonuclease and DNA binding functions. It is recruited to stalled replication forks in response to replication stress and inhibits replication of certain viruses such as the human immunodeficiency virus 1 (HIV-1) by modulating the tRNA pool. SLFN11 has been identified as a predictive biomarker in cancer, as its expression correlates with a beneficial response to DNA damage inducing anticancer drugs. However, the mechanism and interdependence of these two functions are largely unknown. Here, we present cryo-electron microscopy (cryo-EM) structures of human SLFN11 in its dimeric apoenzyme state, bound to tRNA and in complex with single-strand DNA. Full-length SLFN11 neither hydrolyses nor binds ATP and the helicase domain appears in an autoinhibited state. Together with biochemical and structure guided mutagenesis studies, our data give detailed insights into the mechanism of endoribonuclease activity as well as suggestions on how SLFN11 may block stressed replication forks.

Human Schlafen 11 (SLFN11) acts as a potent restriction factor of certain retroviruses, such as the human immunodeficiency virus 1 (HIV-1)[1]. SLFN11 binds tRNAs and counteracts changes in the tRNA repertoire induced by HIV-1 infections[1]. This results in the inhibition of viral protein expression in a codon-usage-dependent manner[2]. In addition to retroviruses, SLFN11 impairs the replication of DNA viruses like human cytomegalovirus (HCMV)[3] as well as positive-strand RNA Flaviviruses like West Nile virus (WNV), Dengue virus (DENV), and Zika virus (ZIKV)[4]. Hence, SLFN11 is an important antiviral factor that targets different types of viruses, offering therapeutic potential[5].

SLFN11 expression levels have been reported to show a strong positive correlation with the sensitivity of tumour cells to DNA-damaging agents (DDAs). The downregulation of certain tRNAs by SLFN11 inhibits the translation of the central DNA damage response (DDR) proteins ataxia-telangiectasia mutated (ATM) and Rad3-related protein (ATR)[6]. However, SLFN11 is also directly recruited to sites of DNA damage and stalled replication forks in response to replication stress induced by DDAs[7]. It interacts with Replication Protein A (RPA1)[7,8] and minichromosome-maintenance 3 (MCM3) at replication foci and selectively blocks fork progression by chromatin opening in the vicinity of replication initiation sites[7,9].

SLFN11 can serve as a biomarker to predict the response to platinum-based DDAs[7,10–12], topoisomerase inhibitors[9,13–17], poly(-ADP-ribose) polymerase (PARP) inhibitors[18–20] and DNA synthesis inhibitors[14,21]. As these drugs lead to replication fork stalling and cell cycle checkpoint activation, replication stress appears to be the common mechanism by which SLFN11 sensitizes cancer cells to DDAs[22].

While much has been learned about the biological and medical relevance of SLFN11, the mechanism of how the protein achieves its endoribonucleolytic and replication fork binding functions is still not well understood. Although crystal structures of the N-terminal domains of SLFN5 and rat rSLFN13 as well as cryo-EM structures of full-length SLFN5 and SLFN12 have been solved by us and others, information about substrate recognition and processing as well as the C-terminal helicase domain is still missing[23–26].

In this work, we provide the cryo-EM structures of the full-length SLFN11 apoenzyme and in complex with single-strand DNA (ssDNA) as well as tRNA. SLFN11 forms a homodimer with the helicase domains adopting an autoinhibited conformation. Guided by the structures, we predict amino acid exchanges that selectively diminish ssDNA binding and endonucleolytic activity, which we

---

[1]Gene Center and Department of Biochemistry, Ludwig-Maximilians-Universität München, Feodor-Lynen Straße 25, 81377 Munich, Germany. [2]These authors contributed equally: Felix J. Metzner, Simon J. Wenzl, Michael Kugler. e-mail: klammens@genzentrum.lmu.de

verify by biochemical analysis. Together, the data give detailed insights into substrate recognition and processing by SLFN11 as well as its regulation by phosphorylation. The mode of ssDNA binding by dimeric SLFN11 suggests a mechanism how the protein might block stalled replication forks. Taken together, our manuscript describes a new avenue to a mechanistic understanding of Schlafen proteins and specifically how SLFN11 acts as a double-edged sword with two functions.

Results and discussion

Overall structure of dimeric SLFN11 apoenzyme

SLFN11 consists of an N-terminal endonuclease domain (residues 1–353), termed Slfn core domain, followed by a linker domain (residues 354–576) and a C-terminal domain with homology to superfamily I (SF I) DNA/RNA helicases (residues 577–901) (Fig. 1a). To provide a structural basis for the nuclease and ATPase functions, we used cryo-EM to solve the structure of full-length human SLFN11 (Fig. 1b). We recorded

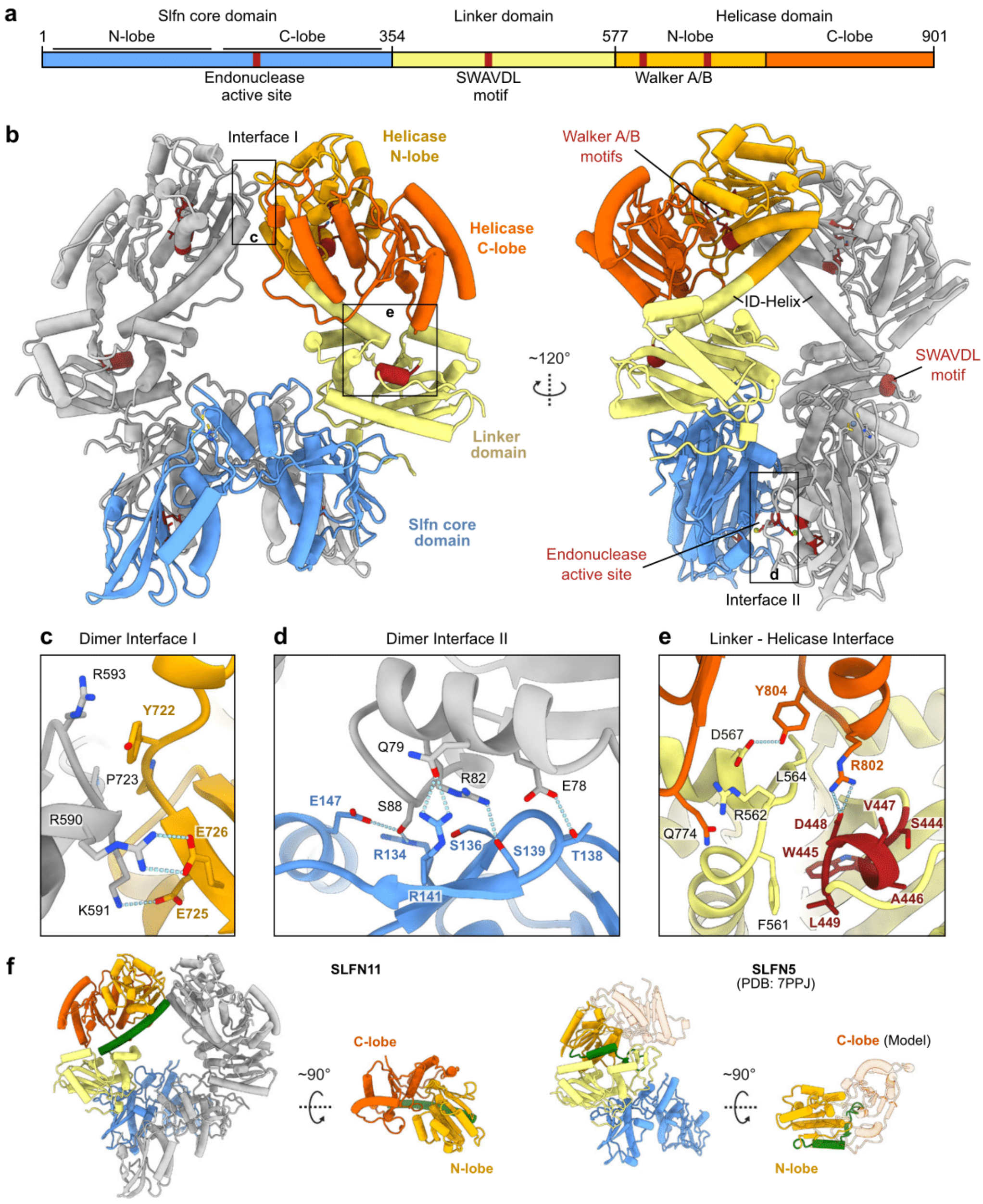


Fig. 1 | Structure of full-length human SLFN11. a Domain architecture of SLFN11 with indicated key functional features. **b** Ribbon representation of the full-length human SLFN11 dimer with highlighted structural features in dark red. **c** Detailed view of the dimer interface I between helicase domains. **d** Detailed view of the dimer interface II between Slfn core domains. **e** Detailed view of the linker-helicase interface. SWAVDL motif and corresponding residues are coloured in dark red. **f** Structural differences of SLFN11 and SLFN5 (PDB code 7PPJ) in the conformation of the helicase domains. The SLFN11 ID-helix is depicted in green. The position of the missing C-lobe of the helicase domain of SLFN5 is depicted as a transparent cartoon model based on the AlphaFold model[42].

datasets of SLFN11^{wt} and SLFN11^{E209A} and 2D classifications yielded classes of monomeric as well as C2 symmetric dimeric SLFN11. 3D reconstruction of the SLFN11^{wt} dimer resulted in a map with a global resolution of 2.86 Å, allowing model building of residues 7–899 (Fig. 1b). Furthermore, monomeric and dimeric SLFN11^{E209A} were reconstructed from two different datasets, resulting in maps with global resolutions of 4.0 Å and 3.25 Å, respectively (Supplementary Fig. 1a). Regarding a single protomer, the Slfn core domain forms a horseshoe-like shape as previously described for other Schlafen family members[23–26]. A zinc ion is coordinated by residues H285, C287, C321 and C322 forming a zinc finger motif (Supplementary Fig. 1b).

The SLFN11 dimer exhibits a ring-like structure with the individual protomers interacting via two interfaces (Fig. 1b–d). Interface I is located between the N-lobes of the helicase domains and is stabilized by several salt bridges (R590 to E726 and K591 to E725) (Fig. 1c). Interface II is located between the Slfn core domains that interact in a head-to-tail orientation and is formed by mostly polar and charged residues (E78, Q79, R82, S88, R134, T138, S139, R141, E147) (Fig. 1d). A related dimer interface was reported for SLFN12 which is stabilized in its dimeric form by small molecule-induced PDE3A binding[25,26] (Supplementary Fig. 1c).

In solution, two peaks corresponding to the molecular weights of monomeric (104 kDa) and dimeric (208 kDa) SLFN11 were detected by mass photometry analysis (Supplementary Fig. 1d). In line with the polar nature of the dimer interfaces, the equilibrium between monomer and dimer is highly salt-sensitive (Supplementary Fig. 1d).

The linker domain, harbouring the conserved SWAVDL motif, interacts with the helicase C-lobe (Fig. 1e) and connects to the helicase domain via a long α-helix that we will refer to as the inter-domain (ID)-helix (Fig. 1f). Notably, the structure of human SLFN5 revealed a strikingly different conformation of this region, despite a high sequence conservation[23] (Fig. 1f and Supplementary Fig. 2). Due to the different folds, the helicase N-lobes of SLFN5 and SLFN11 adopt relative conformations which are rotated by approximately 180° to each other with regard to the respective linker domains. While monomeric SLFN11 exhibits the same overall conformation as the protomers in the dimer, the Slfn core C-lobe is not visible in the map of monomeric SLFN11 (Supplementary Fig. 1a). This suggests conformational flexibility and indicates that dimerization is needed to stabilize the C-lobes of the Slfn core domains in a defined conformation.

Mechanism of the SLFN11 endoribonuclease activity

SLFN11 has been reported to cleave type II tRNAs, leading to translational inhibition in a codon usage-dependent manner[6]. By modulation of the tRNA pool, SLFN11 can inhibit viral protein translation during e.g. HIV infections or translation of human ATM and ATR in response to stalled replication forks[1,6]. In the dimer structure, the lined-up Slfn core domains form a central channel. The endonuclease active sites are located in this central groove in close proximity to interface II (Fig. 2). To identify the binding mode to tRNA, we structurally analysed SLFN11 in the presence of a mixture of different yeast tRNAs. 3D

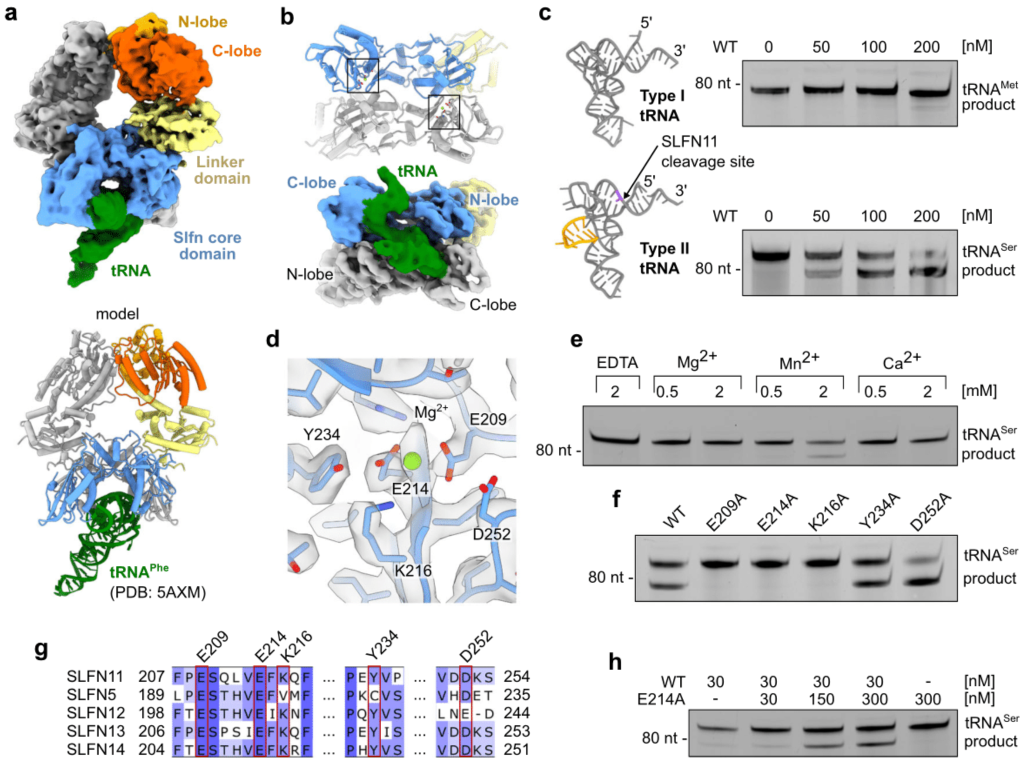


Fig. 2 | SLFN11 tRNA recognition and nuclease activity. **a** Cryo-EM density map for SLFN11 bound to tRNA with proposed structural model of docked yeast tRNA^{Phe} (PDB code 5AXM). **b** Bottom view of Slfn core domains with highlighted nuclease active sites. Cryo-EM density map for tRNA bound between N- and C-lobes of Slfn core domains. **c** Structural models of type I (tRNA^{Met}, PDB code 2FMT) and type II (tRNA^{Ser}, model created by RNAComposer[53, 54]). Variable loop of type II tRNA is coloured in orange and SLFN11 cleavage site in pink (nt 75). SLFN11 endonuclease activity towards respective tRNAs demonstrated in nuclease assays. **d** Close-up view of the nuclease active site with active site residues (E209 and E214) coordinating a magnesium ion. **e** Metal ion dependency of SLFN11 endonuclease activity examined by nuclease assay. **f** Effect of mutations of nuclease active site residues on cleavage of tRNA^{Ser} monitored by nuclease assay. **g** Multiple sequence alignment of nuclease active site residues of human Schlafen family members with mutated residues highlighted in red. **h** Characterization of dimerization-induced nuclease activity of SLFN11 monitored by nuclease assay. Experiments in **c**, **e**, **f**, **h** were performed in duplicates. One representative replicate is shown. Source data for **c**, **e**, **f**, **h** are provided as a Source Data file.

reconstruction showed additional density in the groove between the Slfn core domain N- and C-lobes, fitting one tRNA molecule (Fig. 2a). This shows that a single tRNA molecule is bound and cleaved by the SLFN11 dimer. Two conserved positively charged patches, previously described to contribute to efficient tRNA cleavage by rSLFN13, are in close proximity to the tRNA[24] (Supplementary Fig. 3a). Although the resolution of the map did not allow for detailed tRNA model building, the density indicates that substrate recognition occurs in the Slfn core N- and C-lobes of both protomers of the dimer (Fig. 2b).

We analysed binding of SLFN11 to type I ($tRNA^{Met}$) and type II ($tRNA^{Ser}$) tRNAs in vitro, showing that it binds both tRNAs (Supplementary Fig. 3b). However, $tRNA^{Ser}$ is more efficiently cleaved compared to $tRNA^{Met}$ (Fig. 2c). Type I and type II tRNAs differ in the presence of a variable loop that is present in type II but not in type I tRNAs (Fig. 2c). The observed cleavage pattern places the cutting site approximately 10 nucleotides from the 3′ end, between the acceptor stem and the T-loop (Fig. 2c and Supplementary Fig. 3c, d). To precisely map the SLFN11 endonucleolytic cutting site, we sequenced $tRNA^{Ser}$ cleavage products (Supplementary Fig. 3e). SLFN11 cleaves $tRNA^{Ser}$ mainly at a single cutting site, positioned 10 nucleotides from the 3′ end, between nucleotides 75 and 76. This is in line with the cleavage pattern observed for rSLFN13[24]. In the cryo-EM map, SLFN11 interacts with the acceptor stem and T-loop of the tRNA. It is positioned in close proximity to both endonuclease active sites, approximately 10 and 20 nucleotides from the 3′ end of the tRNA, respectively. However, based on the endonuclease activity assays and the sequencing results, only the active site that is closer to the 3′ end is cleavage proficient. The variable loop which is not clearly visible in our density, could be involved in binding, since it would be positioned in proximity to the Slfn core domain. Hence, the specificity of the enzyme is presumably determined at several different recognition sites.

In the apoenzyme structure additional density for a metal ion appears at the proposed nuclease active site which is coordinated by residues E209 and E214 (Fig. 2d). In cleavage reactions, addition of Mn^{2+} resulted in endonuclease activity, while Mg^{2+} and Ca^{2+} did not stimulate tRNA cleavage (Fig. 2e). Density for a metal ion could not be observed in the cryo-EM map of mutant $SLFN11^{E209A}$ (Supplementary Fig. 3f). To clarify, which residues contribute to the endonuclease activity, we introduced point mutations around the proposed nuclease active site. Mutation of the ion coordinating residues (E209A, E214A) abolished the nuclease activity completely (Fig. 2f). This is in accordance with the active site residues identified for rSLFN13[24]. In contrast, mutation D252A, the third residue of the proposed three carboxylate triad, resulted in a slight increase in nuclease activity. Mutation Y234A showed wild type-like activity, while K216A rendered the nuclease inactive (Fig. 2f). The essential residues E209, E214 and K216 are conserved among RNA cleavage proficient human Schlafen proteins (SLFN11, SLFN12, SLFN13, and SLFN14) (Fig. 2g). Schlafen family members lacking one of these three active site residues, like SLFN5[23] or mouse mSLFN2[27], have been shown to be endonuclease deficient.

Malone et al. reported[28] that ribonuclease activity of SLFN11 is inhibited by phosphorylation of residues S219, T230 and S753. Thus, we investigated the phosphorylation status of the protein used in our studies, which was expressed in insect cells. Mass spectrometry analysis unambiguously identified tryptic peptides with unphosphorylated S219, T230 and S753 (Supplementary Fig. 4). Corresponding phosphorylated peptides remained undetected. This confirms that the unphosphorylated protein is the enzymatically active form. Residues S219 and T230 are located within the Slfn core C-lobe in close proximity to the nuclease active site and might directly influence substrate binding or cleavage[28] (Supplementary Fig. 4a). The impact of a phosphorylation at S753 which is located within the helicase domain is not obvious from a structural point of view and will be discussed subsequently (Supplementary Fig. 4b).

To biochemically verify whether SLFN11 dimerization is required for endonuclease activity, we performed an in trans complementation assay where the nuclease inactive $SLFN11^{E214A}$ was titrated to a constant and limiting concentration of $SLFN11^{wt}$. The nuclease activity was increased with the addition of $SLFN11^{E214A}$, while $SLFN11^{E214A}$ alone was inactive (Fig. 2h). This confirms that dimeric SLFN11 is the nuclease active species and that a single active site is sufficient for tRNA cleavage. A double mutant of dimer interface I located in the helicase domains (K591D, Y722A) did not influence tRNA cleavage while the introduction of a third mutation in dimer interface II (R82D) rendered the nuclease inactive (Supplementary Fig. 3g). In summary, dimerization of SLFN11 is necessary for substrate specificity and cleavage site orientation in the active site. In line with our results, it has been hypothesized that the ribonuclease activity of subgroup II Schlafen family member SLFN12 might be stimulated by dimerization[25].

SLFN11 binding to single-strand DNA

SLFN11 is recruited to stalled replication forks and blocks them in an irreversible manner, eventually leading to cell death[7,18,29]. We found that SLFN11 binds single-strand DNA (ssDNA) with high affinity but does not bind to double-strand DNA (dsDNA) (Fig. 3a and Supplementary Fig. 5). SLFN11 binding to 50 nucleotide (nt) ssDNA resulted in a defined species in an electrophoretic mobility shift assay (EMSA) (Fig. 3a). Fluorescence anisotropy was employed to verify the binding preference for ssDNA over dsDNA and yielded an apparent K_d of ~30 nM for ssDNA (Supplementary Fig. 5a). Binding of ssDNA to SLFN11 increased its inflection temperature by 7.6 °C in a thermal unfolding assay (nanoDSF) while dsDNA had no stabilizing effect (Fig. 3b).

To structurally analyse the interaction between SLFN11 and ssDNA, we solved the cryo-EM structure of SLFN11 bound to 60 nt ssDNA at 3.16 Å resolution (Fig. 3c). The overall fold is similar to the structure of the apoenzyme dimer. Both protomers bind a stretch of five nucleotides via a positively charged patch between the helicase N- and C-lobes (Fig. 3d). Since the helicase domains of the two protomers are rotated by 180° relative to each other, the ssDNA strands are pointing in opposing 5′ to 3′ directions (Fig. 3c). The majority of the interactions are located along the phosphate backbone (helicase N-lobe: N633, Q634, T650, K652, T653, R656, R674; helicase C-lobe: S853, R855, R856) (Fig. 3d). A single charge reversal mutation (K652D) abolishes the ssDNA binding completely (Fig. 3e and Supplementary Fig. 5b, c). The mutation has no effect on the ribonuclease activity confirming the correct folding of the protein (Supplementary Fig. 5d). Most DNA binding residues are conserved amongst human subgroup III Schlafen family members except for SLFN14 and with minor differences in SLFN5 and SLFN13 (Supplementary Fig. 2). The shape of the ssDNA binding groove in SLFN11 seems to sterically prevent dsDNA binding (Supplementary Fig. 5e). NanoDSF measurements of SLFN11 in the presence of dsDNA and at different buffer conditions showed no effect on the inflection temperature (Supplementary Fig. 5f–h). This indicates that SLFN11 does not bind to dsDNA.

The sigmoid shape of the fluorescence anisotropy data suggests cooperative binding of ssDNA (Fig. 3e and Supplementary Fig. 5a), consistent with the observation that both DNA binding sites are occupied in the cryo-EM structure. Mass photometry data show that the addition of ssDNA shifts the equilibrium towards the dimer state, indicating ssDNA-induced stabilization of the SLFN11 dimer (Supplementary Fig. 5i). The opposite 5′ to 3′ direction of the two bound ssDNAs suggests that the SLFN11 dimer could simultaneously bind to both single strands at a stalled replication fork and may thereby block fork progression[7,9].

It has been reported that SLFN11 can be phosphorylated at S753 and that the phospho-mimetic mutant S753D is incapable of reducing type II tRNA levels[28]. The phosphorylation site at S753 is located near the ssDNA binding groove within the helicase domain (Supplementary Fig. 6a). Thus, we mutated S753 to a phospho-mimetic aspartate.

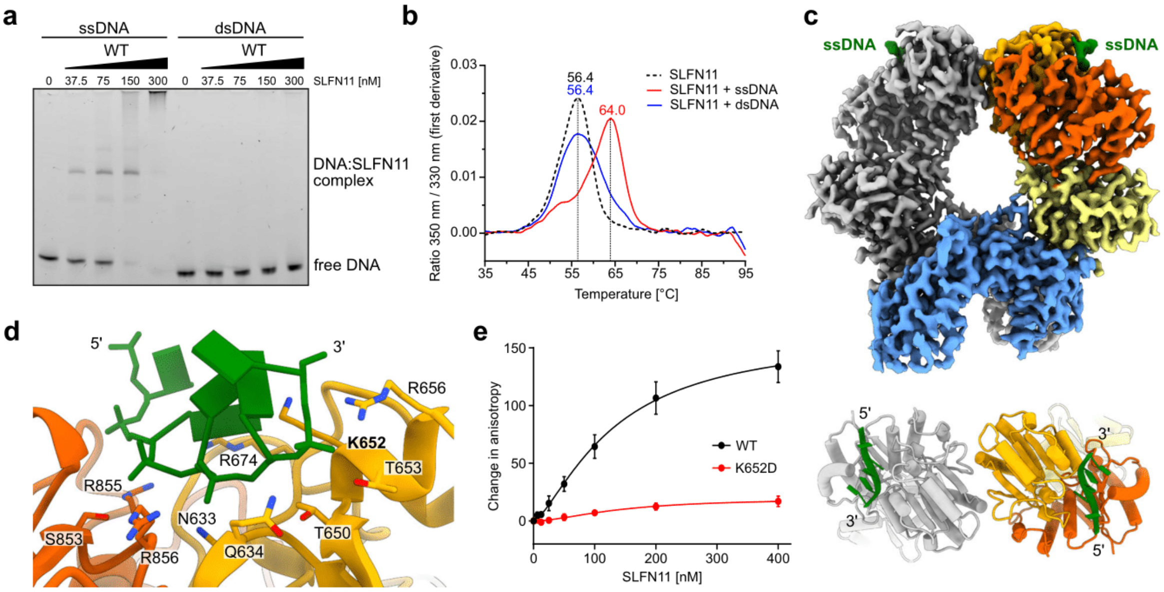


Fig. 3 | SLFN11 binding to DNA. **a** ssDNA and dsDNA binding ability of SLFN11 monitored by electrophoretic mobility shift assay. The experiment was performed in duplicates. One representative replicate is shown. **b** NanoDSF measurements of SLFN11 in the presence of ssDNA and dsDNA. **c** Cryo-EM reconstruction of the SLFN11^{wt} dimer bound to ssDNA. Top view of ssDNA bound between the helicase N- and C-lobes of SLFN11 with indicated 5´ to 3´ direction. **d** Detailed view of the DNA binding region with labelled DNA-interacting residues. Mutated residue K652 is highlighted in bold. **e** Fluorescence anisotropy assay documenting the effect of K652D mutation on the binding of SLFN11 towards 50 nt ssDNA. The data were fit to a cooperative binding equation. Data are represented as mean values + /− SEM from four independent experiments. Source data for **a**, **b**, **e** are provided as a Source Data file.

Although there is no direct interaction between S753 and ssDNA in the cryo-EM structure, the S753D mutant is deficient in DNA binding as monitored by EMSA (Supplementary Fig. 6b, c). This hints towards a functional connection between ssDNA binding to the helicase domain and the nuclease activity of the Slfn core domain.

SLFN11 exhibits an autoinhibited ATPase conformation

The helicase domain of SLFN11 harbours the essential Walker A and B motifs of SF I DNA/RNA helicases which usually couple ATP hydrolysis to translocation along or unwinding of DNA or RNA (Supplementary Fig. 2). Mutation of the Walker B motif has been shown to cause defects in chromatin opening activity of SLFN11 in response to replication stress[9]. Thus, we analysed nucleotide binding and hydrolysis by SLFN11. Addition of different nucleotides (ATP, ADP and ATPγS) causes only minor changes in the inflection temperature of SLFN11, indicating no interaction with the tested nucleotides (Fig. 4a). In contrast, SLFN5 shows a significant increase in unfolding temperature in presence of ATP and ATPγS (Fig. 4a).

We structurally overlaid the helicase domains of SLFN11 and DNA2, another SF I DNA/RNA helicase in its ADP-bound state (Fig. 4b). This illustrates that in SLFN11, ATP binding is sterically blocked by the ID-helix. The conserved glutamine residue of the predicted Q-motif points away from the nucleotide-binding site (Fig. 4b). In addition, we could not detect ATPase activity of SLFN11 alone or in the presence of different DNA/RNA substrates or RPA (Supplementary Fig. 7a). Together, this indicates that the ATPase domains of dimeric SLFN11 are locked in an inactive state, unable to bind ATP.

Structural comparison of SLFN11 and SLFN5 reveals large differences in the conformation of the helicase domains and linker domains (Fig. 4c). In SLFN11, the conserved residue F561 and R802 of the helicase C-lobe interact with the SWAVDL motif (amino acid 444–449) (Fig. 1e). These interactions might suggest a regulatory role of this amongst Schlafen proteins highly conserved motif. In SLFN5, the helicase N-lobe forms an interface with the linker domain and the different conformation of the ID-region potentially opens space for a nucleotide to bind (Fig. 4b). The SLFN11 helicase domain appears to be locked in an autoinhibited conformation via the dimer interface I and the helicase-linker domain interface (Figs. 1c, e and 4c).

The monomeric cryo-EM structure adopts a similar helicase domain arrangement and ID-helix fold as the SLFN11 dimer (Supplementary Fig. 1a), indicating a second necessary step to trigger an active ATPase conformation. However, the characteristic strand separating Pin motif that is found in most SF I helicases with double strand unwinding activity is missing in SLFN11[30,31] (Supplementary Fig. 7b). Together with the absence of dsDNA binding elements (SF1A helicase domains 1B and 2B) in SLFN11, this demonstrates that the protein alone is not a strand opening helicase.

In summary, human SLFN11 is a manganese-dependent type II tRNA endoribonuclease. One tRNA molecule is bound in the positively charged groove of the SLFN11 core domain dimer and both protomers are involved in substrate recognition. Although the tRNA is in close proximity to both nuclease active sites, only one cleavage reaction occurs 10 nucleotides from the 3′ end. The active site residues E209, E214 and K216 are essential for catalysis. These residues are conserved in human SLFN11, 12, 13 and 14 that were shown to be nuclease proficient[1,24,25,32,33]. Previous studies implied that helicase or ATPase activity is needed for replication fork blockage and/or chromatin opening[7]. SLFN11 on its own is not proficient in ATP hydrolysis nor ATP binding in vitro. Hence, an additional factor e.g. binding partner, modification or signal is needed to activate ATPase activity. Despite the high sequence conservation of the Schlafen family members, the subtle differences in the endonuclease active site, DNA binding groove and helicase domain lead to divergent functions within this protein family. The presented data shed light on the structural and functional attributes of human SLFN11 and how it might bind to replication forks. The structural and mutational data explain the regulatory and inhibitory effects resulting from the phosphorylation of SLFN11. The inability of the phospho-mimetic S753D mutant to cleave tRNA and to bind

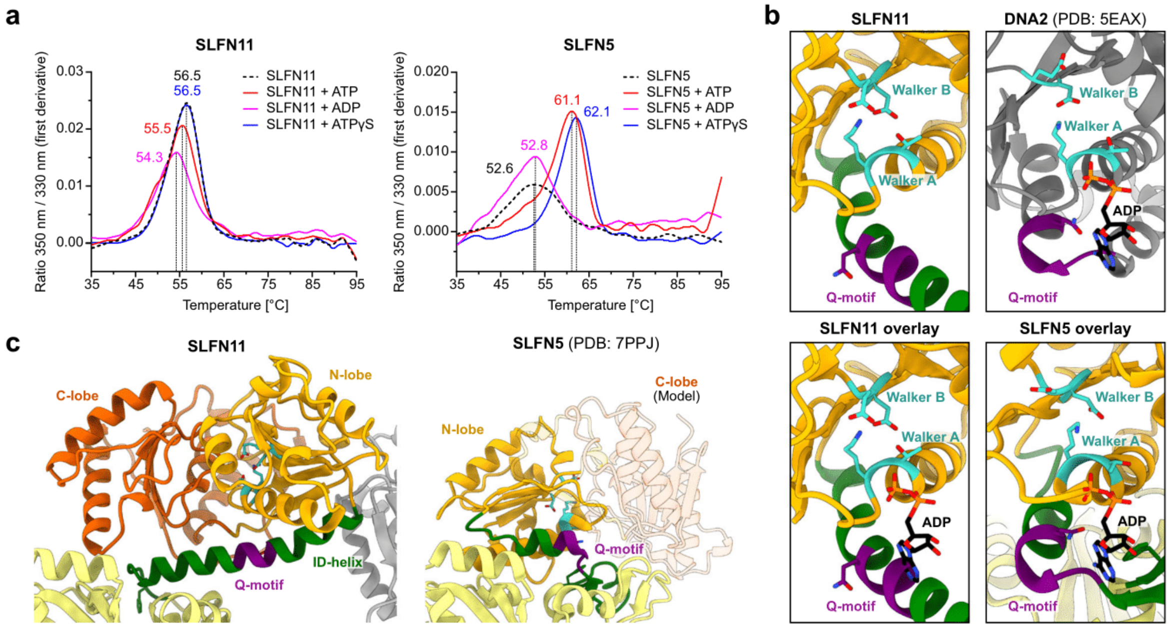


Fig. 4 | Characterization of the SLFN11 helicase domain. a NanoDSF measurements of SLFN11 and SLFN5 in presence of different nucleotides or without nucleotide. **b** Structural comparison of helicase domains of SLFN11, SLFN5 (PDB code 7PPJ) and DNA2 in ADP bound state (PDB code 5EAX). Walker motifs and Q-motif are highlighted in teal and purple, respectively. ADP is coloured in black. **c** Detailed view of the interdomain helix (ID-helix) conformational differences between SLFN11 and SLFN5 (PDB code 7PPJ). Walker motifs are highlighted in teal. Source data for **a** are provided as a Source Data file.

ssDNA might be a hint at a functional connection between the nuclease and helicase domains. We conclude that SLFN11 may act as a double-edged sword with two functions within one protein which might be regulatory connected. Further *in cellulo* studies are needed to uncover the whole regulatory picture of SLFN11 activation and to clarify the possible coupling or separation of these two functions. Our biochemical and structural investigation of full-length SLFN11 presents an important basis to facilitate these studies.

Methods

Protein expression and purification

A construct encoding for full-length SLFN11 with an N-terminal double FLAG-tag and a HRV 3 C cleavage site was purchased from GenScript. SLFN11 was cloned into the pFASTBac1 expression vector using Gibson assembly for expression in insect cells[34]. *Spodoptera frugiperda* Sf21 insect cells (Thermo Fisher) were used for virus generation. Expression was carried out in *Trichoplusia ni* High Five cells (Invitrogen) at 27 °C and 95 rpm for 72 h. After 72 h, the cells were harvested by centrifugation, resuspended in lysis buffer (50 mM Tris pH 7.5, 400 mM NaCl, 2 mM $MgCl_2$) supplemented with protease inhibitors (0.18 g l^{-1} PMSF, 0.32 g l^{-1} benzamidine, 1.37 mg l^{-1} pepstatin A, 0.26 mg l^{-1} leupeptin, 0.2 mg l^{-1} chymostatin) and disrupted by sonication. The lysate was clarified by centrifugation at 30,000 g at 4 °C for 45 min and the supernatant was incubated with pre-equilibrated ANTI-FLAG M2 Affinity Gel (Sigma-Aldrich) for 90 min. The resin was washed with wash buffer (25 mM Tris pH 7.5, 250 mM NaCl, 2 mM $MgCl_2$). After washing with buffer A (25 mM Tris pH 7.5, 120 mM NaCl, 2 mM $MgCl_2$, 1 mM DTT), the protein was eluted iteratively for five times in elution buffer (buffer A supplemented with 0.2 mg ml^{-1} Flag-peptide) over 60 min. The eluate was loaded onto a HiTrap Heparin HP column (GE Healthcare) and the protein was eluted by a linear salt gradient (100 % buffer A to 100 % buffer B (25 mM Tris pH 7.5, 1 M NaCl, 2 mM $MgCl_2$, 1 mM DTT) over 12 CV). The peak fractions were combined and flash frozen in liquid nitrogen. For preparation of cryo-EM samples, SLFN11 was directly concentrated using a centrifugal filter unit (Amicon, MWCO 30 kDa). Concentrated SLFN11 was applied onto either Superdex 200 5/150 or Superose 6 increase 5/150 column (GE Healthcare) equilibrated in buffer A. Peak fractions were used for cryo-EM grid preparation.

SLFN11 mutants were prepared by site-directed mutagenesise PCR and expressed and purified as the wild-type protein. SLNF5 was expressed and purified following a similar protocol, with the difference, that the pH of the buffers was adjusted to pH 7.1[23].

RPA was cloned into the pBIG1a expression vector using the biGBac system[35]. *Spodoptera frugiperda* Sf21 insect cells (Thermo Fisher) were used for virus generation. RPA was expressed in *Trichoplusia ni* High Five cells (Invitrogen) similarly to SLFN11. Cell pellets were resuspended in lysis buffer (20 mM HEPES pH 7.8, 300 mM NaCl, 2 mM $MgCl_2$, 5 mM KCl, 0.1 mM EDTA, 4 mM imidazole) supplemented with protease inhibitors. Cells were lysed by sonication and the lysate was centrifuged at 30,000 g at 4 °C for 60 min. The supernatant was applied onto pre-equilibrated Ni^{2+}-NTA agarose beads (Qiagen), and incubated for 45 min. Beads were applied onto a 5 ml column (Bio-Rad) and washed with lysis buffer. RPA was eluted by adding 5 CV imidazole buffer (20 mM HEPES pH 7.8, 300 mM NaCl, 2 mM $MgCl_2$, 5 mM KCl, 0.1 mM EDTA, 350 mM imidazole). The eluate was pooled and dialyzed overnight into buffer A (20 mM HEPES pH 7.8, 100 mM NaCl, 2 mM $MgCl_2$, 5 mM KCl, 0.1 mM EDTA). The clarified protein solution was applied onto a pre-equilibrated HiTrap Heparin HP column (GE Healthcare). RPA was eluted by applying a linear NaCl gradient (10 CV) using buffer B (20 mM HEPES pH 7.8, 1 M NaCl, 2 mM $MgCl_2$, 5 mM KCl, 0.1 mM EDTA). Peak fractions were pooled and flash frozen in liquid nitrogen.

Sample preparation and cryo-EM data acquisition

Purified $SLFN11^{wt}$ was dialyzed into dialysis buffer (20 mM HEPES pH 7.5, 60 mM NaCl, 2 mM $MgCl_2$, 1 mM DTT) and pre-incubated for 30 min with brewer's yeast tRNAs mixture (Sigma-Aldrich). The sample was diluted in cryo-EM buffer (20 mM HEPES pH 7.5, 2 mM $MgCl_2$, 1 mM DTT) to a final concentration of 3.5 µM SLFN11 and 10 µM tRNA mixture. 4.5 µl was applied onto a glow discharged QUANTIFOIL® R2/1 + 2 nm carbon Cu200 grid. The sample was vitrified in liquid ethane using an EM GP plunge freezer (Leica, 10 °C and 90% humidity).

Freshly purified SLFN11^{E209A} was diluted in cryo-EM buffer (50 mM Glycine pH 9, 200 mM NaCl, 2 mM $MgCl_2$, 1 mM DTT) to a final concentration of 5 µM. 4.5 µl was applied onto a glow discharged QUANTIFOIL® R2/1 Cu200 grid. The sample was vitrified in liquid ethane using an EM GP plunge freezer (Leica, 10 °C and 90% humidity).

Freshly purified SLFN11^{wt} was pre-incubated for 30 min with 60 nt ssDNA and 1 mM ADP. The sample was diluted in cryo-EM buffer (100 mM Glycine pH 9, 2 mM $MgCl_2$, 1 mM DTT) to a final concentration of 5 µM SLFN11, 2 µM ssDNA, and 1 mM ADP. 4.5 µl was applied onto a glow discharged QUANTIFOIL® R2/1 Cu200 grid. The sample was vitrified in liquid ethane using an EM GP plunge freezer (Leica, 10 °C and 90% humidity).

Cryo-EM data collection

Cryo-EM data were collected using an FEI Titan Krios G3 transmission electron microscope (300 kV) equipped with a GIF quantum energy filter (slit width 20 eV) and a Gatan K2 Summit direct electron detector (software used: EPU 2.12.1.278REL, TEM User interface Titan 2.15.4, Digital Micrograph 3.22.1461.0). For the structure of SLFN11^{wt} and SLFN11^{wt} bound to tRNA 8,569 movies were collected with a total electron dose of 49.65 $e^- Å^{-2}$, fractionated into 40 movie frames over 8 s. For the structure determination of SLFN11^{E209A} dimer 7,078 movies were collected (of which 4,420 movies at a tilt angle of 25°) with a total electron dose of 44.03 $e^- Å^{-2}$, fractionated into 40 movie frames over 8 s. For the structure of SLFN11^{E209A} monomer 3,212 movies were collected with a total electron dose of 43.58 $e^- Å^{-2}$, fractionated into 40 movie frames over 8 s. For the structure determination of SLFN11^{wt} bound to ssDNA 6,419 movies were collected (of which 4,088 movies at a tilt angle of 25°) with a total electron dose of 43.33 $e^- Å^{-2}$, fractionated into 40 movie frames over 8 s. All datasets were collected with defocus values ranging from -1.1 to -2.9 µm and a pixel size of 1.046 Å.

Cryo-EM image processing

Movie frames were motion corrected using MotionCor2 1.4.5[36]. All subsequent cryo-EM data processing steps were carried out using cryoSPARC 3.3.1[37] and the resolutions reported here are calculated based on the gold-standard Fourier shell correlation criterion (FSC = 0.143). In total, five datasets were processed. The CTF parameters of the datasets were determined using patch CTF estimation (multi). The exact processing schemes are depicted in Supplementary Fig. 8–10. The data collection and refinement statistics are summarized in Supplementary Table 1.

For the SLFN11^{wt} with tRNA mixture dataset (Supplementary Fig. 8), particles were initially picked on 320 micrographs using Blob picker. Reasonable 2D classes were selected and used as input for Topaz train[38,39]. The resulting Topaz model was used as template for particle picking on 8,569 micrographs yielding 2,756,604 particles extracted with a box size of 320 px and a pixel size of 1.046 Å. The particles were subject to 2D classification, ab-initio reconstruction, and heterogenous refinement and the class with clearly defined features was selected (1,603,620 particles). The obtained particles were further sorted by 2D classification and heterogenous refinement resulting in five classes. The class that showed the most defined features of SLFN11 alone was selected (522,494 particles) and used for further refinement. The final resolution of the SLFN11^{wt} reconstruction after non-uniform refinement[40] was 2.86 Å. From these above-mentioned five classes, two classes (249,140 particles and 126,239 particles) were analysed by 3D variability[41] for presence of an additional density belonging to bound tRNA. Resulting volumes of 3D variability served as inputs for two rounds of heterogenous refinement of the previously obtained set of particles (1,603,620 particles). From four classes, one class (96,514 particles) showed clear features of bound tRNA to SLFN11. The final resolution of the tRNA bound SLFN11^{wt} reconstruction without masking was 3.98 Å.

Cryo-EM data processing of SLFN11^{E209A} dimer (Supplementary Fig. 9) was carried out in a similar fashion compared to the SLFN11^{wt} with tRNA mixture dataset. For the SLFN11^{E209A} data -set particles were initially picked on 2,658 micrographs using blob picker. Reasonable 2D classes were selected and used as input for Topaz train. The resulting topaz model was used as template for particle picking on 2,658 micrographs (untilted) and 4,420 micrographs (25° tilt angle), respectively. Particles corresponding to dimer classes were extracted with a box size of 320 px and a pixel size of 1.046 Å. Particles were further sorted by 2D classification, ab-initio reconstruction, and heterogenous refinement. A total of 280,205 particles were combined and subject to Ctf refinement and non-uniform refinement. The final resolution of the SLFN11^{E209A} dimer reconstruction in C2 symmetry was 3.25 Å.

An analogous processing strategy was used for the SLFN11^{E209A} monomer (Supplementary Fig. 9). From 3,312 micrographs, a total of 789,190 particles corresponding to monomer classes were extracted with a box size of 256 px and a pixel size of 1.046 Å. After further processing, one class with 262,713 particles was used for final 2D classification and non-uniform refinement. Processing resulted in a 4.0 Å map of the SLFN11^{E209A} monomer containing 223,491 particles.

Cryo-EM data processing of SLFN11^{wt} bound to ssDNA (Supplementary Fig. 10) was carried out in a similar fashion as in the SLFN11^{E209A} dimer dataset. 2,331 micrographs were used for initial particle picking. The resulting Topaz model was used as template for particle picking on 2,331 micrographs and 4,088 micrographs (25° tilt angle), respectively. A total of 152,738 particles (untilted) and 310,225 particles (tilted) was extracted and combined. Further processing implying heterogenous and non-uniformed refinement with C2 symmetry applied, resulted in a 3.16 Å map of the SLFN11^{wt} dimer bound to ssDNA (247,654 particles).

Model building and refinement

Atomic models were built by rigid body docking of models predicted by AlphaFold2[42] into the cryo-EM density. For building of ssDNA, ssDNA bound to DNA2 (PDB code 5EAX (DNA2 in complex with ssDNA)) was used as a starting model. As the sequence of the bound ssDNA cannot be determined from the map, the five nucleotides closest to the 5′ end of the 60 nt sequence were modelled into the C2 symmetric ssDNA bound SLFN11 map. The models were partially rebuilt in Coot 0.9-pre[43]. Missing parts were built de-novo. Atomic models were improved by ISOLDE 1.2.2[44] and real space refinement in PHENIX 1.17[45,46] using the maps with highest resolution, respectively. The model of SLFN11^{wt} bound tRNA was generated by docking of tRNA^{Phe} (PDB code 5AXM] (Thg1 like protein (TLP) with tRNA^{Phe})) into the corresponding density map using UCSF ChimeraX 1.2[47]. All structure figures were prepared with UCSF ChimeraX[47].

Mass photometry

The molecular mass of SLFN11 in solution was determined by mass photometry. All mass photometry measurements were carried out using a OneMP mass photometer (Refeyn). Prior to each measurement the focus was adjusted by applying 19 µl mass photometry buffer (25 mM Tris pH 7.5, 2 mM $MgCl_2$, 1 mM DTT with variable concentrations of NaCl) to a new flow chamber. SLFN11 was diluted in sterile filtered mass photometry buffer to a final concentration of 50 nM, immediately prior to mass photometry measurements. For ssDNA stabilization experiment, 60 nt ssDNA was added with a final concentration of 100 nM or 300 nM. Movies were recorded for 60 s and data were analysed using AcquireMP (Refeyn) 2.3.

Nuclease assay

Nuclease activity of SLFN11 was examined by a gel-based nuclease assay. The nuclease reaction was performed in nuclease buffer (25 mM Tris pH 7.5, 120 mM NaCl, 2 mM $MgCl_2$, 1 mM DTT) with 50 nM SLFN11 and 50 nM 6-FAM labelled nucleic acid substrate, unless otherwise indicated. 2 mM $MnCl_2$ was added if not stated otherwise. Reactions were started by adding the substrate and incubated at 37 °C for 45 min. Samples were mixed with loading dye (15% Ficoll, 20 mM Tris pH 7.6, 40 mM NaCl) and applied to a self-cast 15% denaturing polyacrylamide gel (Rotiphorese® DNA sequencing system). Gels were run in 0.5× TBE at 270 V (Bio-Rad) for 50 min. Gels were imaged by a Typhoon™ FLA 7000 (GE Healthcare) and analysed using GIMP 2.10.2 and ImageJ 1.8.0[48].

The effect of dimerization on the nuclease activity of SLFN11 was examined using a gel-based nuclease assay with settings described above. 30, 150, and 300 nM $SLFN11^{E214A}$ was titrated to 30 nM $SLFN11^{wt}$. Reactions were started by adding 50 nM 6-FAM labelled $tRNA^{Ser}$.

The effect of ssDNA on the nuclease activity of SLFN11 was examined using a gel-based nuclease assay as described before. 50 nM $SLFN11^{wt}$ or $SLFN11^{K652D}$ were incubated with or without 100 nM 50 nt ssDNA. Reactions were started by adding 50 nM 6-FAM labelled $tRNA^{Ser}$.

Uncropped gels are provided in the Source Data file.

Electrophoretic mobility shift assay (EMSA)

Binding of SLFN11 to nucleic acid substrates was monitored by electrophoretic mobility shift assay (EMSA). Either 37.5–300 nM SLFN11 (for DNA substrates) or 62.5–500 nM SLFN11 (for tRNA substrates) was incubated with 40 nM 6-FAM labelled substrates at 4 °C for 30 min in EMSA buffer (25 mM HEPES pH 7.5, 60 mM KCl, 8% glycerol, 2 mM $MgCl_2$, 1 mM DTT). Samples were mixed with loading dye (15% Ficoll, 20 mM Tris pH 7.6, 40 mM NaCl) and applied to a NativePAGE 3–12% Bis-Tris gel (Thermo Fisher). The electrophoresis was performed in 1× NativePAGE running buffer (Thermo Fisher) at 100 V for 120 min at 4 °C. The gels were imaged using a Typhoon™ FLA 7000 (GE Healthcare) and analysed using GIMP. Uncropped gels are provided in the Source Data file.

Affinity measurement by fluorescence anisotropy

Initial protein dilutions (0, 6.25, 12.5, 25, 50, 100, 200, 400, and 800 nM) of SLFN11 were prepared in assay buffer (25 mM Tris pH 7.5, 120 mM NaCl, 2 mM $MgCl_2$, 1 mM DTT). Protein dilutions were then mixed with 6-FAM labelled DNA (Supplementary Table 2) in assay buffer without NaCl (final DNA concentration of 10 nM) in a 1:1 (v/v) ratio (final volume: 20 µl, Greiner Flat Bottom Black 384 well plate). The reaction was incubated at 25 °C for 30 min and the fluorescence anisotropy was subsequently measured at an excitation wavelength of 470 nm and an emission wavelength of 520 nm using an Infinite M1000 microplate photometer (Tecan). Experiments were performed at least three times. The background signal (no protein sample) was subtracted from each value of a dilution series and the datasets were analysed with Prism 6.07 (GraphPad Software). The datasets were fit to a Hill model to determine the apparent dissociation constants.

Nano differential scanning fluorimetry (nanoDSF)

Binding of SLFN11 to various substrates was examined by nanoDSF (Tycho NT.6, NanoTemper Technologies). 300 nM SLFN11 was incubated with 300 nM of 50 nt ssDNA or 50 bp dsDNA in nanoDSF buffer (25 mM Tris pH 7.5, 60 mM NaCl, 2 mM $MgCl_2$, 1 mM DTT) for 30 min on ice, respectively. For testing of dependence on salt (NaCl) and ions (Mg^{2+}, Mn^{2+}, Ca^{2+}, and Zn^{2+}), nanoDSF buffer was adjusted accordingly. Interaction with nucleotides was performed similarly, where SLFN11 or SLFN5 (300 nM) were incubated with or without corresponding nucleotides (1 mM) in nanoDSF buffer. The samples were loaded into glass capillaries and the internal fluorescence at 330 nm and 350 nm was measured while a thermal gradient was applied. Data were analysed using the internal Tycho NT.6 software 1.3.2.878 and plotted with Prism (GraphPad Software).

ATP hydrolysis assay

A fluorescence-based ATPase assay was conducted to determine the ATPase rate of SLFN11[49]. SLFN11 (250 nM) was incubated with 150 nM of different DNA or RNA substrates in ATPase buffer (25 mM Tris pH 7.5, 50 mM NaCl, 1 mM DTT, 2 mM $MgCl_2$, 0.1 mg ml^{-1} BSA) at 4 °C for 30 min. RPA was used at a concentration of 250 nM. SLFN11:substrate complexes were combined with 0.1 mM NADH in reaction buffer (25 mM Tris pH 7.5, 50 mM NaCl, 1 mM DTT, 2 mM $MgCl_2$, 0.1 mg ml^{-1} BSA, 0.5 mM PEP (phosphoenolpyruvate) , 1 mM ATP, 25 U ml^{-1} lactate dehydrogenase/ pyruvate kinase (Sigma-Aldrich)) in a 384 well plate (Greiner). Hexokinase from *Saccharomyces cerevisiae* (1.5 nM, Sigma-Aldrich) supplemented with 300 µM glucose served as the positive control. The fluorescence of NADH was measured at 37 °C using an Infinite M1000 microplate photometer (Tecan). The reaction was monitored for 45 min (20 s intervals) using an excitation wavelength of 340 nm and an emission wavelength of 460 nm. Data were analysed using Prism (GraphPad Software).

tRNA sequencing

For the identification of the SLFN11 cleavage position in $tRNA^{Ser}$, a nuclease reaction was performed as described above. Briefly, 1 µM tRNA substrate in nuclease buffer (25 mM Tris pH 7.5, 120 mM NaCl, 2 mM $MgCl_2$, 1 mM DTT, 2 mM $MnCl_2$) was incubated with 1.1 µM SLFN11 at 37 °C for 45 min. 15 µl of the reaction mixture were purified on a Sephadex G-25 column (Roche Quick Spin columns for radiolabeled RNA). The eluate (20 ng $µl^{-1}$) was converted to an Illumina sequencing library with the SMARTer smRNA-Seq kit (Takara) following the manufacturer's instruction. The RNA was enzymatically polyadenylated, then reversely transcribed with an oligo dT primer and a template-switching primer, obtaining a cDNA that was extended with primer-templated sequences on both ends of the original RNA. The cDNA was amplified with barcoded primers, converting it into a sequencing-ready Illumina-compatible library. The cDNA library was sequenced on a NextSeq1000 (Thermo Fischer) in 60 bp paired-end mode. The obtained sequencing reads were demultiplexed and poly-A tails at the end as well as three nucleotides at the beginning, that were introduced by the template-switching mechanism, were removed. The reads were mapped to the sequence of $tRNA^{Ser}$ with Burrows-Wheeler Aligner 0.17.7[50]. The start positions of the mapped reads were visualized in a histogram using Prism (GraphPad Software).

Mass spectrometry

2 µl of purified SLFN11 (3.2 µM, purified from *Trichoplusia ni* High Five cells) were diluted with 5 µl 100 mM NH_4HCO_3 and 3 µl water. Proteins were reduced by addition of 1 µl DTE (50 mM in 50 mM NH_4HCO_3) and incubated for 30 min at 37 °C. Carbamidomethylation of cysteines was performed by adding 2 µl of iodoacetamide (100 mM in 50 mM NH_4HCO_3) and 30 min incubation at RT. For protein digestion, 20 ng of Trypsin (Promega) was added to the sample and incubated at 37 °C overnight. After digestion, the sample was acidified with 2 µl 15% formic acid. Liquid chromatography-mass spectrometry analysis was performed on an Ultimate 3000 nano-LC system (Thermo Fisher) coupled with a Q Exactive HF-X mass spectrometer (Thermo Fisher). Peptides were separated with an EasySpray reversed-phase column (PepMap RSLC C18, 50 cm length, 75 µm ID, Thermo Fisher) at a flow rate of 250 nl min^{-1}. Solvent A consisted of 0.1% formic acid in water and solvent B of 0.1% formic acid in acetonitrile. The chromatography method included gradients from 3% to 25% solvent B in 30 min and from 25% to 40% B in 5 min. For data acquisition, a top 12 data-dependent acquisition method was used. Spectra were searched using MASCOT 2.4[51] (Matrix Science Ltd) and the human subset of the Swiss-Prot database[52].

Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The data that support this study are available from the corresponding author upon request. The coordinates of the SLFN11^{wt}, SLFN11^{E209A} dimer, and SLFN11^{wt} ssDNA-bound structure have been deposited in the Protein Data Bank (PDB) under the accession code 7ZEL, 7ZEP, and 7ZES, respectively. The SLFN11^{wt} dimer cryo-EM reconstruction is available at the Electron Microscopy Data Bank (EMDB) under the EMBD accession code EMD-14690. The SLFN11 dimer reconstruction of SLFN11^{wt} bound to tRNA is available at the EMDB under the EMDB accession code EMD-14695. The SLFN11 monomer and dimer reconstruction of SLFN11^{E209A} is available at the EMDB under the EMDB accession code EMD-14693 and EMD-14691, respectively. The SLFN11 dimer reconstruction of SLFN11^{wt} bound to ssDNA is available at the EMDB under the accession code EMD-14692. MS spectra were searched using the human subset of the Swiss-Prot database [https://www.uniprot.org/]. Source data are provided with this paper.

References

1. Li, M. et al. Codon-usage-based inhibition of HIV protein synthesis by human schlafen 11. *Nature* **491**, 125–128 (2012).
2. Stabell, A. C. et al. Non-human Primate Schlafen11 Inhibits Production of Both Host and Viral Proteins. *PLoS Pathog.* **12**, e1006066 (2016).
3. Nightingale, K. et al. Human cytomegalovirus protein RL1 degrades the antiviral factor SLFN11 via recruitment of the CRL4 E3 ubiquitin ligase complex. *Proc. Natl Acad. Sci. USA* **119**, e2108173119 (2022).
4. Valdez, F. et al. Schlafen 11 Restricts Flavivirus Replication. *J. Virol.* **93**, e00104–e00119 (2019).
5. Kim, E. T. & Weitzman, M. D. Schlafens Can Put Viruses to Sleep. *Viruses* **14**, 442 (2022).
6. Li, M. et al. DNA damage-induced cell death relies on SLFN11-dependent cleavage of distinct type II tRNAs. *Nat. Struct. Mol. Biol.* **25**, 1047–1058 (2018).
7. Murai, J. et al. SLFN11 Blocks Stressed Replication Forks Independently of ATR. *Mol. Cell* **69**, 371–384.e376 (2018).
8. Mu, Y. et al. SLFN11 inhibits checkpoint maintenance and homologous recombination repair. *EMBO Rep.* **17**, 94–109 (2016).
9. Murai, J. et al. Chromatin Remodeling and Immediate Early Gene Activation by SLFN11 in Response to Replication Stress. *Cell Rep.* **30**, 4137–4151.e4136 (2020).
10. Nogales, V. et al. Epigenetic inactivation of the putative DNA/RNA helicase SLFN11 in human cancer confers resistance to platinum drugs. *Oncotarget* **7**, 3084–3097 (2016).
11. Rathkey, D. et al. Sensitivity of Mesothelioma Cells to PARP Inhibitors Is Not Dependent on BAP1 but Is Enhanced by Temozolomide in Cells With High-Schlafen 11 and Low-O^6-methylguanine-DNA Methyltransferase Expression. *J. Thorac. Oncol.* **15**, 843–859 (2020).
12. Iwasaki, J. et al. Schlafen11 Expression Is Associated With the Antitumor Activity of Trabectedin in Human Sarcoma Cell Lines. *Anticancer Res.* **39**, 3553–3563 (2019).
13. Barretina, J. et al. The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature* **483**, 603–607 (2012).
14. Zoppoli, G. et al. Putative DNA/RNA helicase Schlafen-11 (SLFN11) sensitizes cancer cells to DNA-damaging agents. *Proc. Natl Acad. Sci. USA* **109**, 15030–15035 (2012).
15. Coussy, F. et al. BRCAness, SLFN11, and RB1 loss predict response to topoisomerase I inhibitors in triple-negative breast cancers. *Sci. Transl. Med.* **12**, eaax2625 (2020).
16. Marzi, L. et al. The Indenoisoquinoline TOP1 Inhibitors Selectively Target Homologous Recombination-Deficient and Schlafen 11-Positive Cancer Cells and Synergize with Olaparib. *Clin. Cancer Res.* **25**, 6206–6216 (2019).
17. Kaur, S. et al. Identification of Schlafen-11 as a Target of CD47 Signaling That Regulates Sensitivity to Ionizing Radiation and Topoisomerase Inhibitors. *Front. Oncol.* **9**, 994 (2019).
18. Murai, J. et al. Resistance to PARP inhibitors by SLFN11 inactivation can be overcome by ATR inhibition. *Oncotarget* **7**, 76534–76550 (2016).
19. Lok, B. H. et al. PARP Inhibitor Activity Correlates with SLFN11 Expression and Demonstrates Synergy with Temozolomide in Small Cell Lung Cancer. *Clin. Cancer Res.* **23**, 523–535 (2017).
20. Allison Stewart, C. et al. Dynamic variations in epithelial-to-mesenchymal transition (EMT), ATM, and SLFN11 govern response to PARP inhibitors and cisplatin in small cell lung cancer. *Oncotarget* **8**, 28575–28587 (2017).
21. Winkler, C. et al. SLFN11 informs on standard of care and novel treatments in a wide range of cancer models. *Br. J. Cancer* **124**, 951–962 (2021).
22. Murai, J., Thomas, A., Miettinen, M. & Pommier, Y. Schlafen 11 (SLFN11), a restriction factor for replicative stress induced by DNA-targeting anti-cancer therapies. *Pharmacol. Ther.* **201**, 94–102 (2019).
23. Metzner, F. J., Huber, E., Hopfner, K. P. & Lammens, K. Structural and biochemical characterization of human Schlafen 5. *Nucleic Acids Res.* **50**, 1147–1161 (2022).
24. Yang, J. Y. et al. Structure of Schlafen13 reveals a new class of tRNA/rRNA-targeting RNase engaged in translational control. *Nat. Commun.* **9**, 1165 (2018).
25. Garvie, C. W. et al. Structure of PDE3A-SLFN12 complex reveals requirements for activation of SLFN12 RNase. *Nat. Commun.* **12**, 4375 (2021).
26. Chen, J. et al. Structure of PDE3A-SLFN12 complex and structure-based design for a potent apoptosis inducer of tumor cells. *Nat. Commun.* **12**, 6204 (2021).
27. Yue, T. et al. SLFN2 protection of tRNAs from stress-induced cleavage is essential for T cell-mediated immunity. *Science* **372**, eaba4220 (2021).
28. Malone, D., Lardelli, R. M., Li, M. & David, M. Dephosphorylation activates the interferon-stimulated Schlafen family member 11 in the DNA damage response. *J. Biol. Chem.* **294**, 14674–14685 (2019).
29. Jo, U. et al. SLFN11 promotes CDT1 degradation by CUL4 in response to replicative DNA damage, while its absence leads to synthetic lethality with ATR/CHK1 inhibitors. *Proc. Natl Acad. Sci. USA* **118**, e2015654118 (2021).
30. Raney, K. D., Byrd, A. K. & Aarattuthodiyil, S. Structure and Mechanisms of SF1 DNA Helicases. *Adv. Exp. Med. Biol.* **767**, 17–46 (2013).
31. Fairman-Williams, M. E., Guenther, U. P. & Jankowsky, E. SF1 and SF2 helicases: family matters. *Curr. Opin. Struct. Biol.* **20**, 313–324 (2010).
32. Pisareva, V. P., Muslimov, I. A., Tcherepanov, A. & Pisarev, A. V. Characterization of Novel Ribosome-Associated Endoribonuclease SLFN14 from Rabbit Reticulocytes. *Biochemistry* **54**, 3286–3301 (2015).
33. Fletcher, S. J. et al. Role of the novel endoribonuclease SLFN14 and its disease-causing mutations in ribosomal degradation. *Rna* **24**, 939–949 (2018).
34. Gibson, D. G. et al. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* **6**, 343–345 (2009).
35. Weissmann, F. et al. biGBac enables rapid gene assembly for the expression of large multisubunit protein complexes. *Proc. Natl Acad. Sci. USA* **113**, E2564–E2569 (2016).
36. Zheng, S. Q. et al. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods* **14**, 331–332 (2017).

37. Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods* **14**, 290–296 (2017).
38. Bepler, T. et al. Positive-unlabeled convolutional neural networks for particle picking in cryo-electron micrographs. *Nat. Methods* **16**, 1153–1160 (2019).
39. Bepler, T., Kelley, K., Noble, A. J. & Berger, B. Topaz-Denoise: general deep denoising models for cryoEM and cryoET. *Nat. Commun.* **11**, 5208 (2020).
40. Punjani, A., Zhang, H. & Fleet, D. J. Non-uniform refinement: adaptive regularization improves single-particle cryo-EM reconstruction. *Nat. Methods* **17**, 1214–1221 (2020).
41. Punjani, A. & Fleet, D. J. 3D variability analysis: Resolving continuous flexibility and discrete heterogeneity from single particle cryo-EM. *J. Struct. Biol.* **213**, 107702 (2021).
42. Jumper, J. et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).
43. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D.* **66**, 486–501 (2010).
44. Croll, T. I. ISOLDE: a physically realistic environment for model building into low-resolution electron-density maps. *Acta Crystallogr. D.* **74**, 519–530 (2018).
45. Liebschner, D. et al. Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr. D.* **75**, 861–877 (2019).
46. Afonine, P. V. et al. Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. D.* **68**, 352–367 (2012).
47. Pettersen, E. F. et al. UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Sci.* **30**, 70–82 (2021).
48. Schneider, C. A., Rasband, W. S. & Eliceiri, K. W. NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* **9**, 671–675 (2012).
49. Kiianitsa, K., Solinger, J. A. & Heyer, W. D. NADH-coupled microplate photometric assay for kinetic studies of ATP-hydrolyzing enzymes with low and high specific activities. *Anal. Biochem.* **321**, 266–271 (2003).
50. Li, H. & Durbin, R. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* **26**, 589–595 (2010).
51. Perkins, D. N., Pappin, D. J., Creasy, D. M. & Cottrell, J. S. Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* **20**, 3551–3567 (1999).
52. Bairoch, A. & Boeckmann, B. The SWISS-PROT protein sequence data bank: current status. *Nucleic Acids Res.* **22**, 3578–3580 (1994).
53. Antczak, M. et al. New functionality of RNAComposer: an application to shape the axis of miR160 precursor structure. *Acta Biochim. Pol.* **63**, 737–744 (2016).
54. Popenda, M. et al. Automated 3D structure composition for large RNAs. *Nucleic Acids Res.* **40**, e112 (2012).

Acknowledgements

We acknowledge Dr. Thomas Fröhlich from the Gene Center at the LMU for mass spectrometry analysis and data evaluation. We acknowledge Dr. Daniel Bollschweiler from the Max-Planck Institute of Biochemistry for the mass photometry experiments and data evaluation. We thank all members of the Hopfner lab for helpful discussions. We acknowledge support by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project-ID 210592381 – SFB 1054 (to K.L.), the Gottfried Wilhelm Leibniz-Prize (to K.-P.H.) and the European Research Council (ERC Advanced Grant INO3D, to K.-P.H.).

Author contributions

F.J.M., S.J.W. and K.L. conceived the project. F.J.M., S.J.W., M.K. and K.L. designed all structural and biochemical experiments. F.J.M., S.J.W. and M.K., conducted all structural and biochemical experiments. F.J.M. S.J.W., M.K. and K.L. carried out the cryo-EM data collection and analysis. S.K. performed tRNA sequencing analysis. K.-P.H. helped with the analysis and interpretation of the results. F.J.M., S.J.W., M.K. and K.L. wrote the manuscript. K.L. and K.-P.H. provided funding. All authors discussed and commented on the results and the manuscript.

Funding

Open Access funding enabled and organized by Projekt DEAL.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-022-33123-0.

Correspondence and requests for materials should be addressed to Katja Lammens.

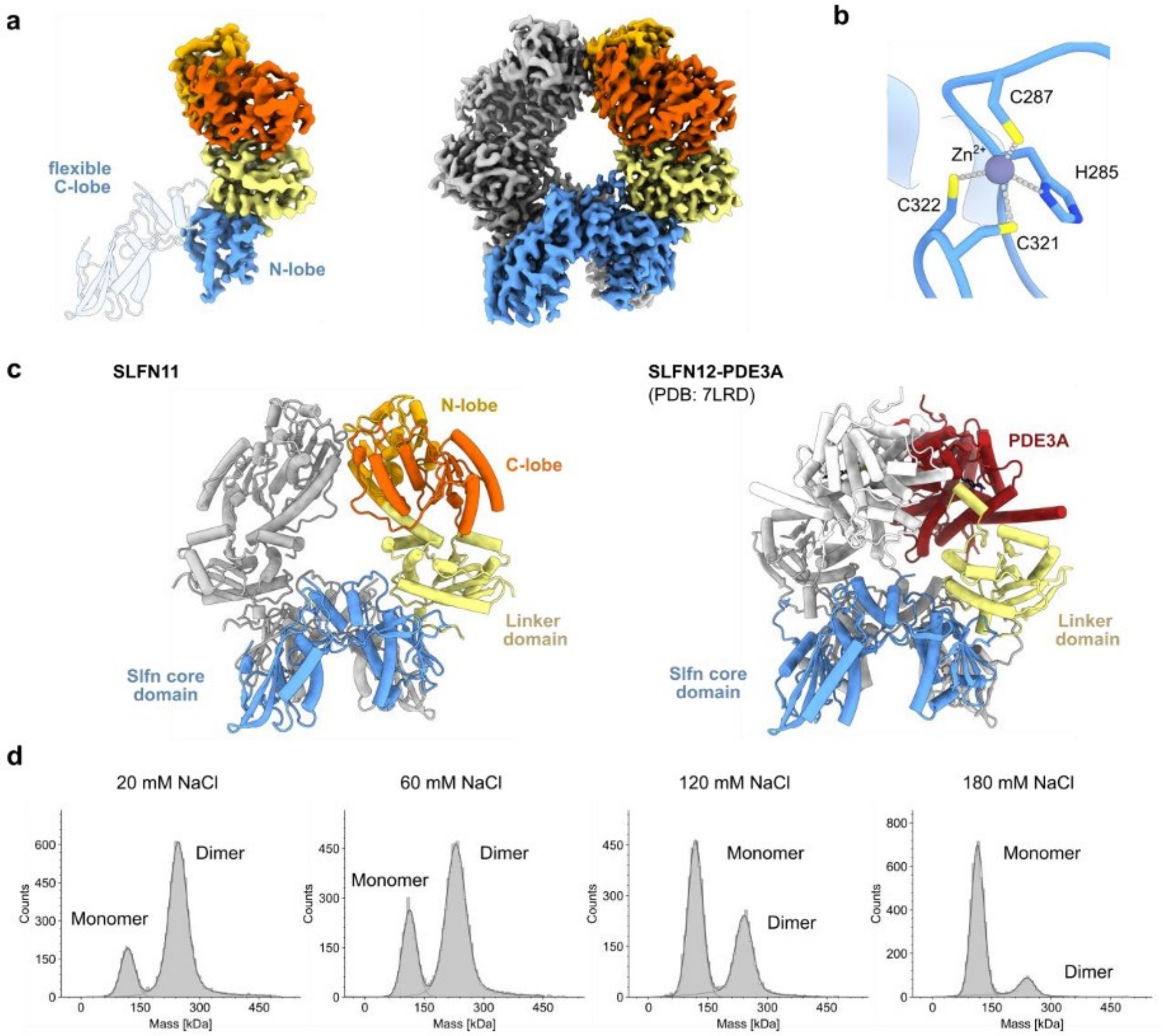
Peer review information *Nature Communications* thanks Christian Dienemann and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Supplementary Information

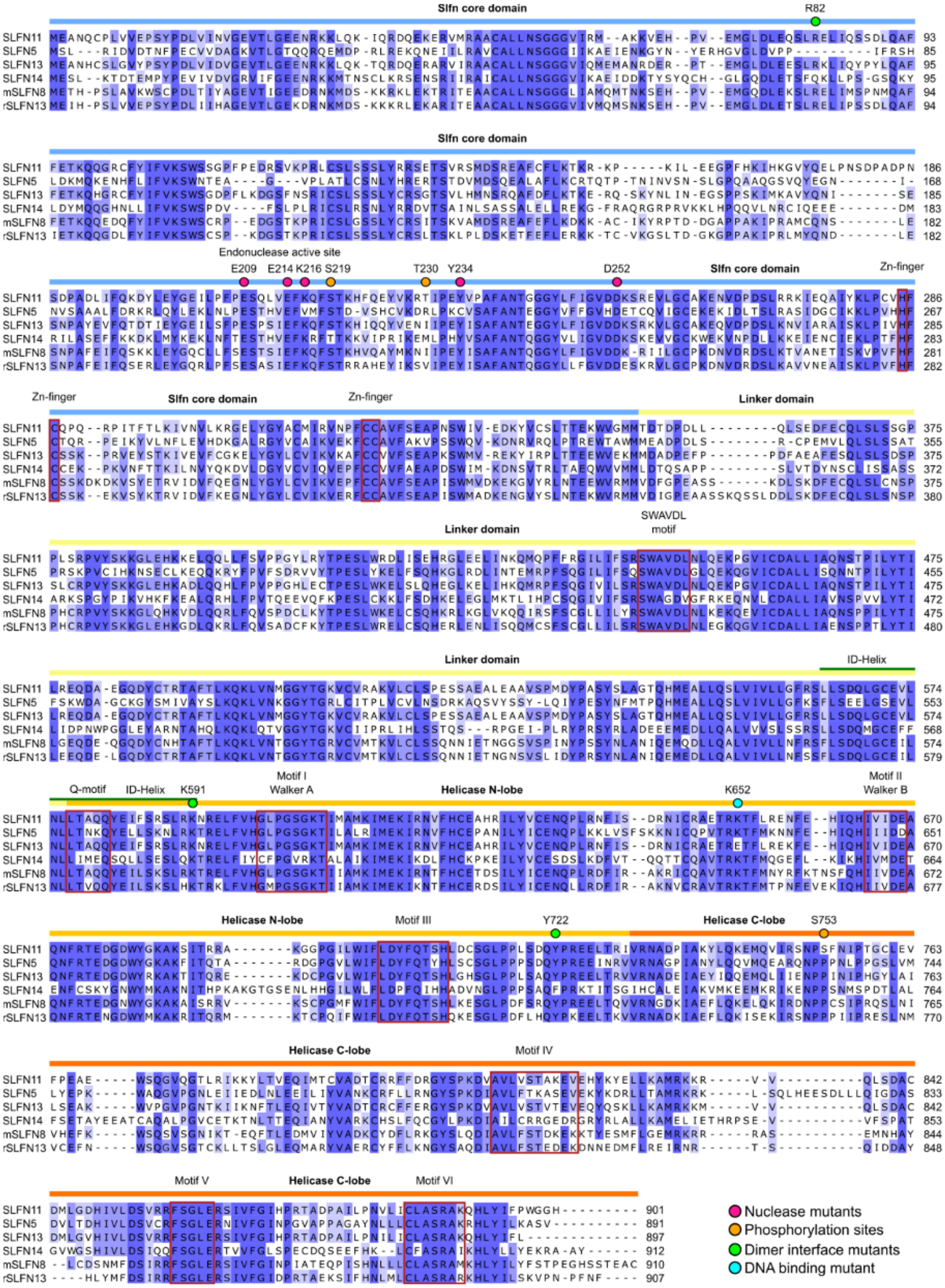
Mechanistic understanding of human SLFN11

Felix J. Metzner, Simon J. Wenzl, Michael Kugler, Stefan Krebs, Karl-Peter Hopfner and Katja Lammens

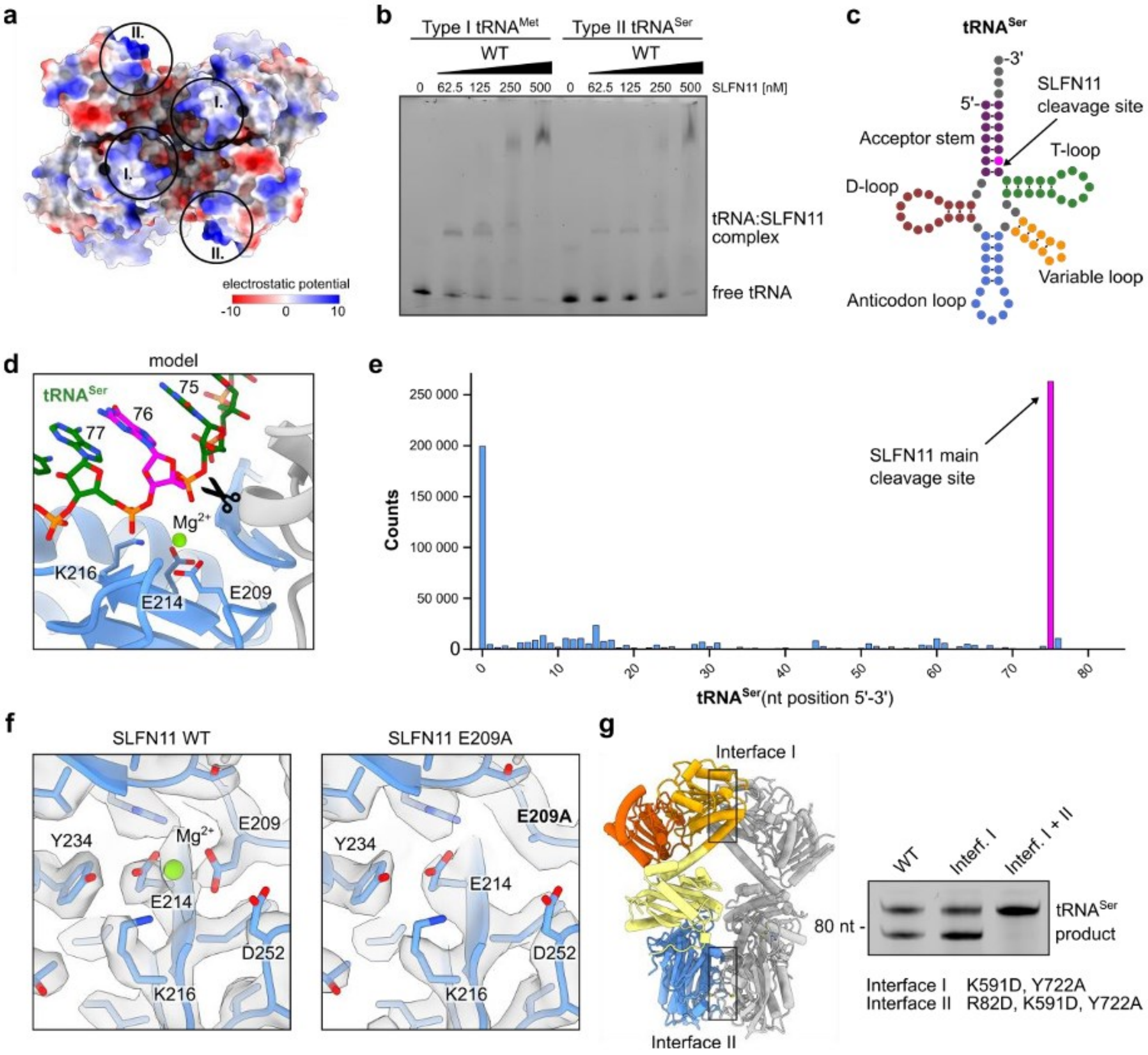
Supplementary Figures



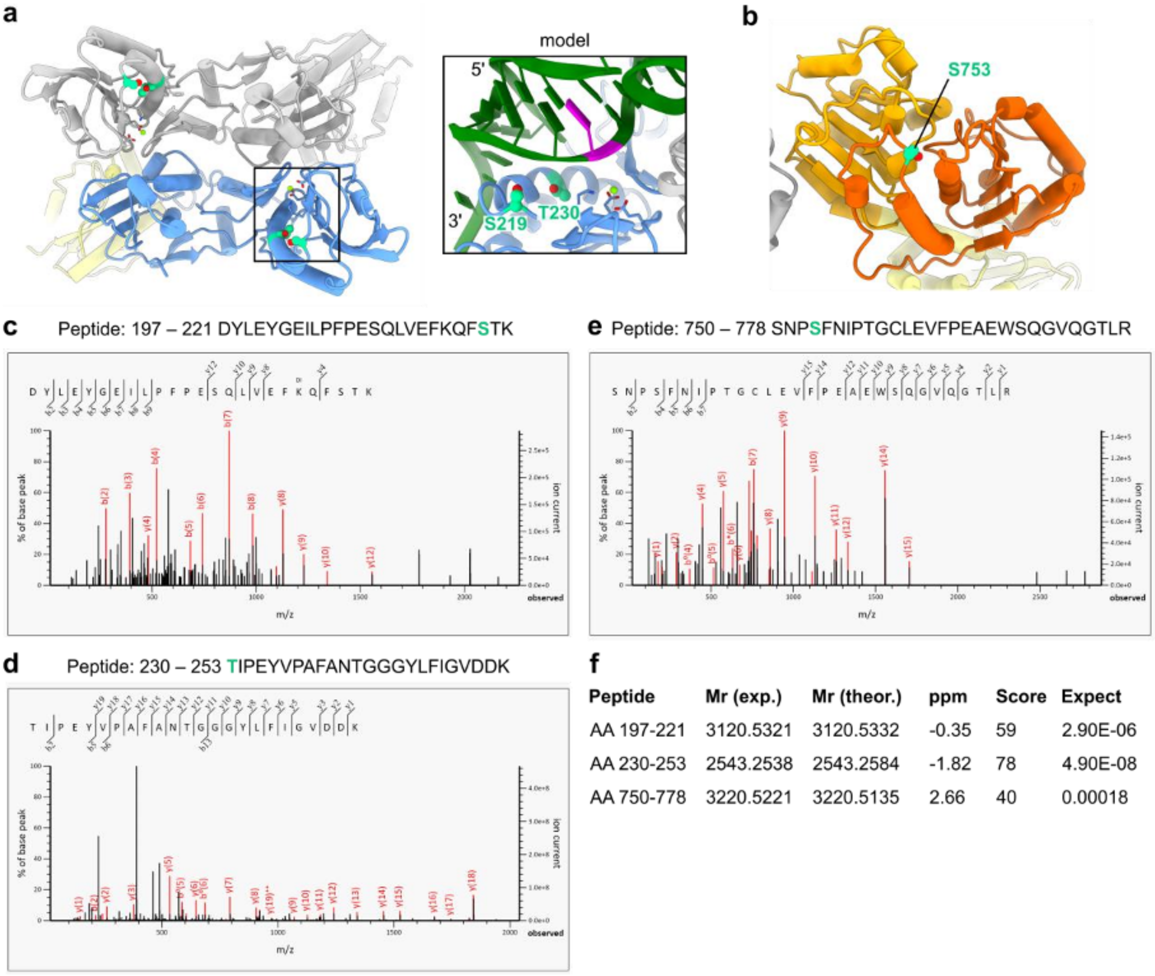
Supplementary Fig. 1: Analysis of SLFN11 monomeric and dimeric form. a, Cryo-EM reconstruction of the SLFN11^{E209A} monomer at 4.0 Å. The position of the missing C-lobe of Slfn core domain is depicted as a transparent cartoon model based on the SFLN11 dimeric structure. Cryo-EM reconstruction of the SLFN11^{E209A} dimer at 3.25 Å. **b**, Close-up view on SLFN11 zinc finger with depicted residues coordinating a zinc ion. **c**, Structural comparison of SLFN11 and SLFN12-PDE3A (PDB code 7LRD). The opposing second protomers of SLFN12 and PDE3A are coloured in grey and white, respectively. **d**, Mass distribution of SLFN11 monomers and dimers in the presence of variable concentrations of NaCl observed by mass photometry.



Supplementary Fig. 2: Multiple sequence alignment of SLFN11 and selected Schlafen proteins from human (SLFN5, SLFN13, SLFN14), mouse (mSLFN8) and rat (rSLFN13). The alignment was calculated using T-Coffee[1]. Residues are coloured according to percentage identity (dark blue = more conserved, white = less conserved). The Slfn core domain is indicated in blue, the linker domain in yellow and the helicase domain in orange. The individual motifs are boxed in red squares. Mutated residues and phosphorylation sites are highlighted in circles in corresponding colour based on the legend.

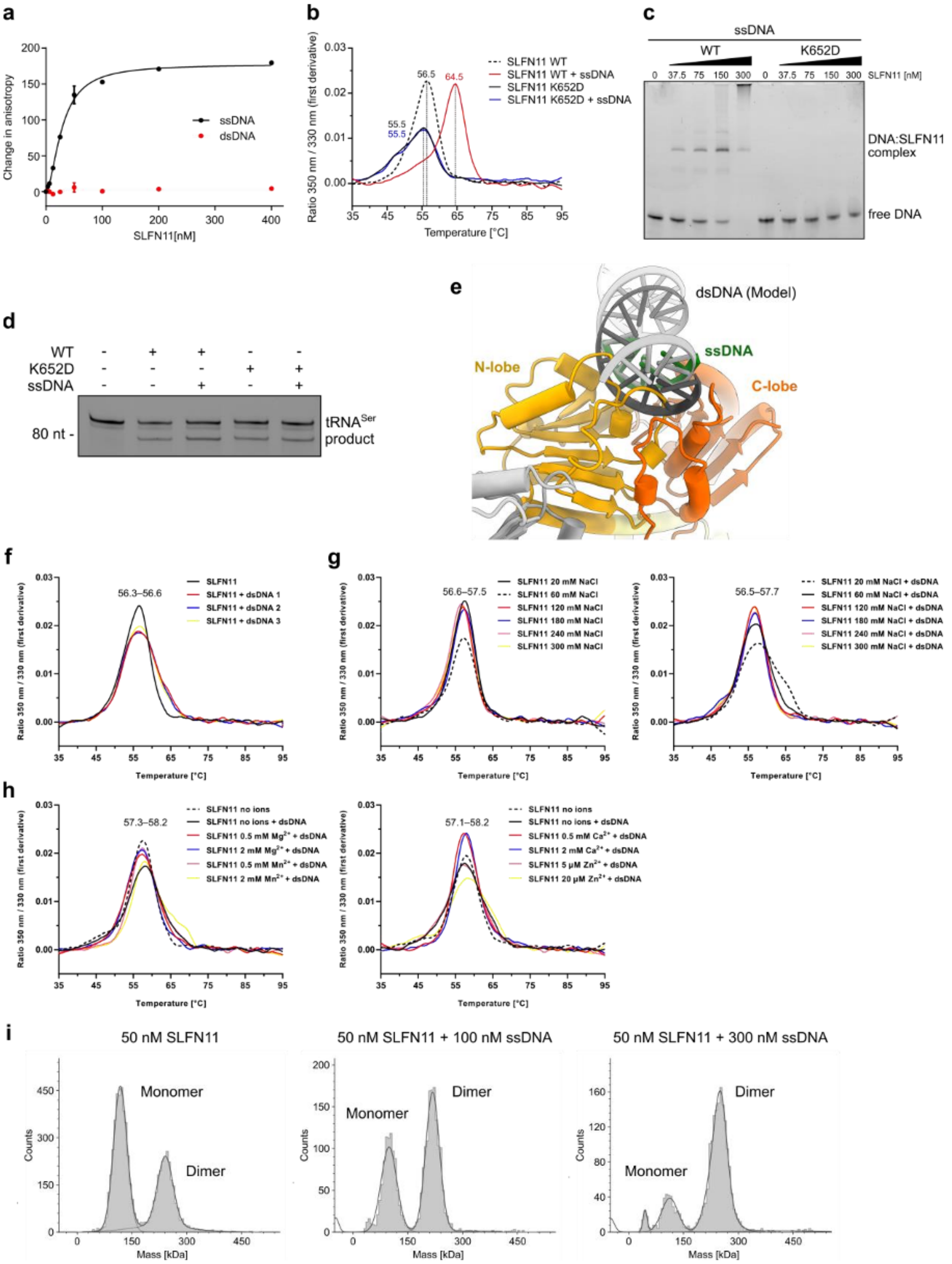


Supplementary Fig. 3: Binding of tRNA to SLFN11 and its regulation. a, Bottom view of the Slfn core domain shown as an electrostatic surface with positively charged patches (I. and II.) responsible for tRNA binding. The electrostatic potential values are in units of kcal mol^{-1} e^{-1} at 298 K. **b**, SLFN11 binding towards tRNA^{Met} and tRNA^{Ser} monitored by electrophoretic mobility shift assay. **c**, Schematic representation of tRNA^{Ser}. SLFN11 cleavage site is coloured in pink. **d**, Structural model of tRNA^{Ser} bound to SLFN11. Residues of the nuclease active site and the metal ion are shown. tRNA^{Ser} bases are numerically labelled from 5' to 3'. The cleavage site between position 75 and 76 (pink) is indicated. **e**, Analysis of tRNA^{Ser} cleavage products by tRNA sequencing. Read counts are plotted against the start positions of the mapped sequencing reads. The main cleavage site between position 75 and 76 is indicated (pink). **f**, Comparison of SLFN11 WT and E209A mutant nuclease active sites with corresponding cryo-EM density map. **g**, Localization of interface I (K591D, Y722A) and interface II (R82D, K591D, Y722A) mutants of SLFN11 and their ability to cleave tRNA^{Ser} as observed in nuclease assay. Experiments in **b**, **g** were performed in duplicates. One representative replicate is shown. Source data for **b**, **e**, **g** are provided as a Source Data file.



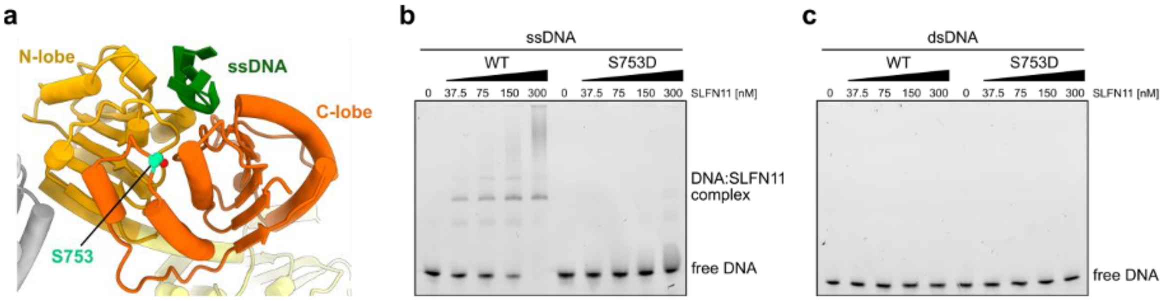
Peptide	Mr (exp.)	Mr (theor.)	ppm	Score	Expect
AA 197-221	3120.5321	3120.5332	-0.35	59	2.90E-06
AA 230-253	2543.2538	2543.2584	-1.82	78	4.90E-08
AA 750-778	3220.5221	3220.5135	2.66	40	0.00018

Supplementary Fig. 4: Phosphorylation sites of SLFN11. a, Left: Bottom view of Slfn core domains. SLFN11 phosphorylation sites S219 and T230 are coloured in light green. Nuclease active site residues are shown. Right: Structural model of tRNA^{Ser} (green) bound to SLFN11. SLFN11 cleavage site is coloured in pink. Phosphorylation sites S219 and T230 are coloured in light green. Residues of the nuclease active site are shown. **b**, Position of phosphorylation site in SLFN11 helicase domain (S753). Residue S753 is shown and coloured in light green. **c**, MS/MS spectra of SLFN11 peptides AA 197–221 covering S219. **d**, MS/MS spectra of SLFN11 peptides AA 230-253 covering T230. **e**, MS/MS spectra of SLFN11 peptides AA 750-778 covering S753. **f**, Overview of Mascot[2] results including experimental (Mr exp.) and theoretical (Mr theor.) relative molecular mass, mass deviation in parts per million (ppm), Mascot ions scores (Score) and expectation values (Expect).

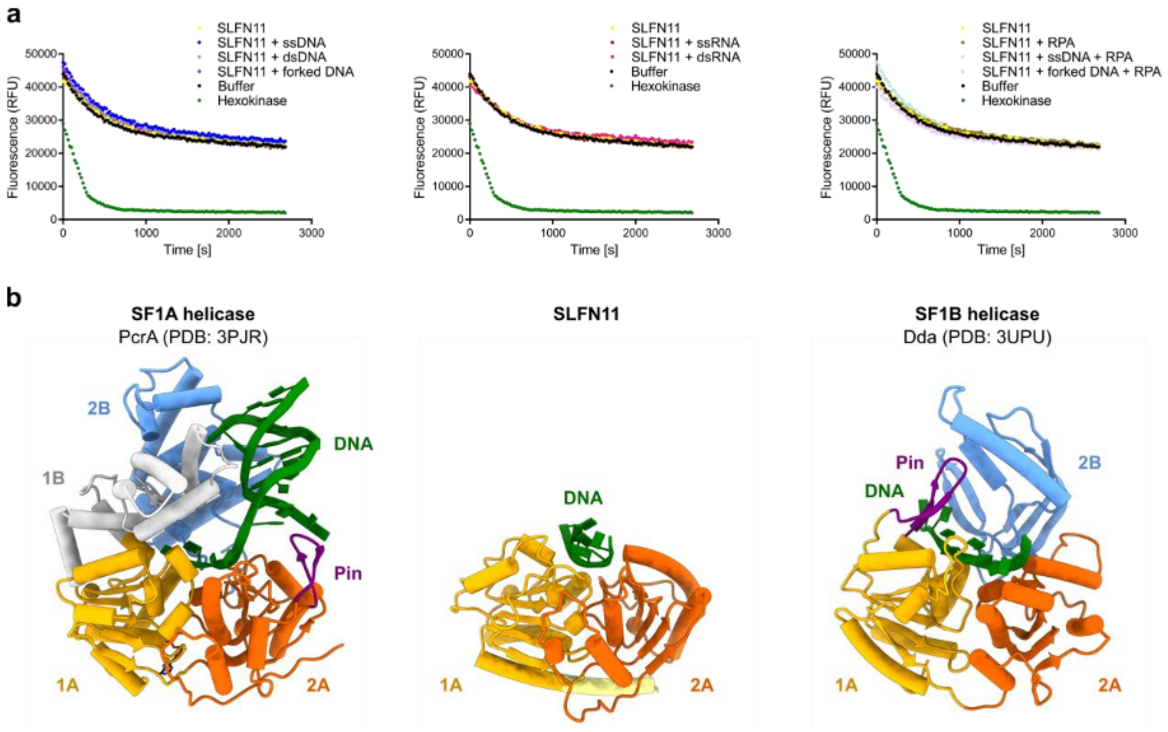


Supplementary Fig. 5: Characterization of DNA binding by SLFN11. a, Fluorescence anisotropy assay illustrating difference in binding ability of SLFN11 towards 50 nt ssDNA and 50 bp dsDNA. The data were fit to a cooperative binding equation. Data are represented as mean values +/- SEM from three independent experiments. **b**, NanoDSF measurements of

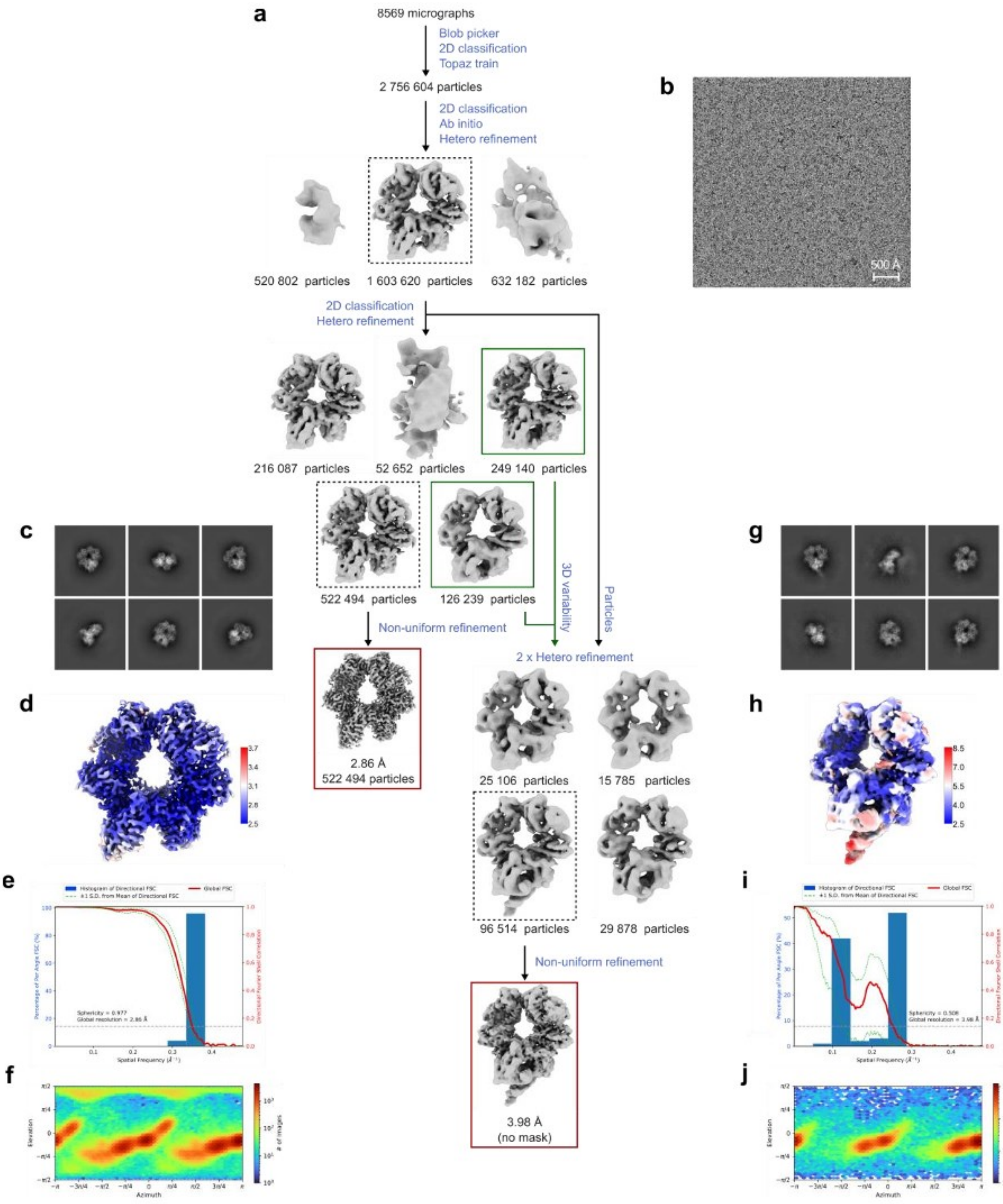
$SLFN11^{wt}$ and $SLFN11^{K652D}$ in the presence of 50 nt ssDNA. **c**, Abolished ssDNA binding ability of SLFN11 by single-point mutation K652D monitored by electrophoretic mobility shift assay. **d**, The effect of ssDNA-induced dimerization of SLFN11 on the nuclease activity of SLFN11 examined by nuclease assay. **e**, Structural model of dsDNA not fitting into the SLFN11 DNA-binding groove between the helicase N- and C-lobes. **f**, NanoDSF measurements of $SLFN11^{wt}$ in the presence of 50 bp dsDNAs with different sequences. **g**, NanoDSF measurements of $SLFN11^{wt}$ at different NaCl concentration without (left) or with (right) dsDNA present. **h**, NanoDSF measurements of $SLFN11^{wt}$ in the presence of different bivalent ions (left: Mg^{2+}, Mn^{2+}; right; Ca^{2+}, Zn^{2+}) with or without dsDNA present. **i**, Stabilization of SLFN11 dimer by ssDNA observed in mass photometry. Experiments in **c**, **d** were performed in duplicates. One representative replicate is shown. Source data for **a**, **b**, **c**, **d**, **f**, **g**, **h** are provided as a Source Data file.



Supplementary Fig. 6: Characterization of SLFN11 phosphomimetic mutant S753D. a, Position of SLFN11 phosphorylation site in ssDNA bound helicase domain (S753). Residue S753 is shown and coloured in light green. **b**, Abolished ssDNA binding ability of SLFN11 by single-point mutation S753D monitored by electrophoretic mobility shift assay. **c**, dsDNA binding ability of SLFN11 by single-point mutation S753D monitored by electrophoretic mobility shift assay. Experiments in **b**, **c** were performed in duplicates. One representative replicate is shown. Source data for **b**, **c** are provided as a Source Data file.

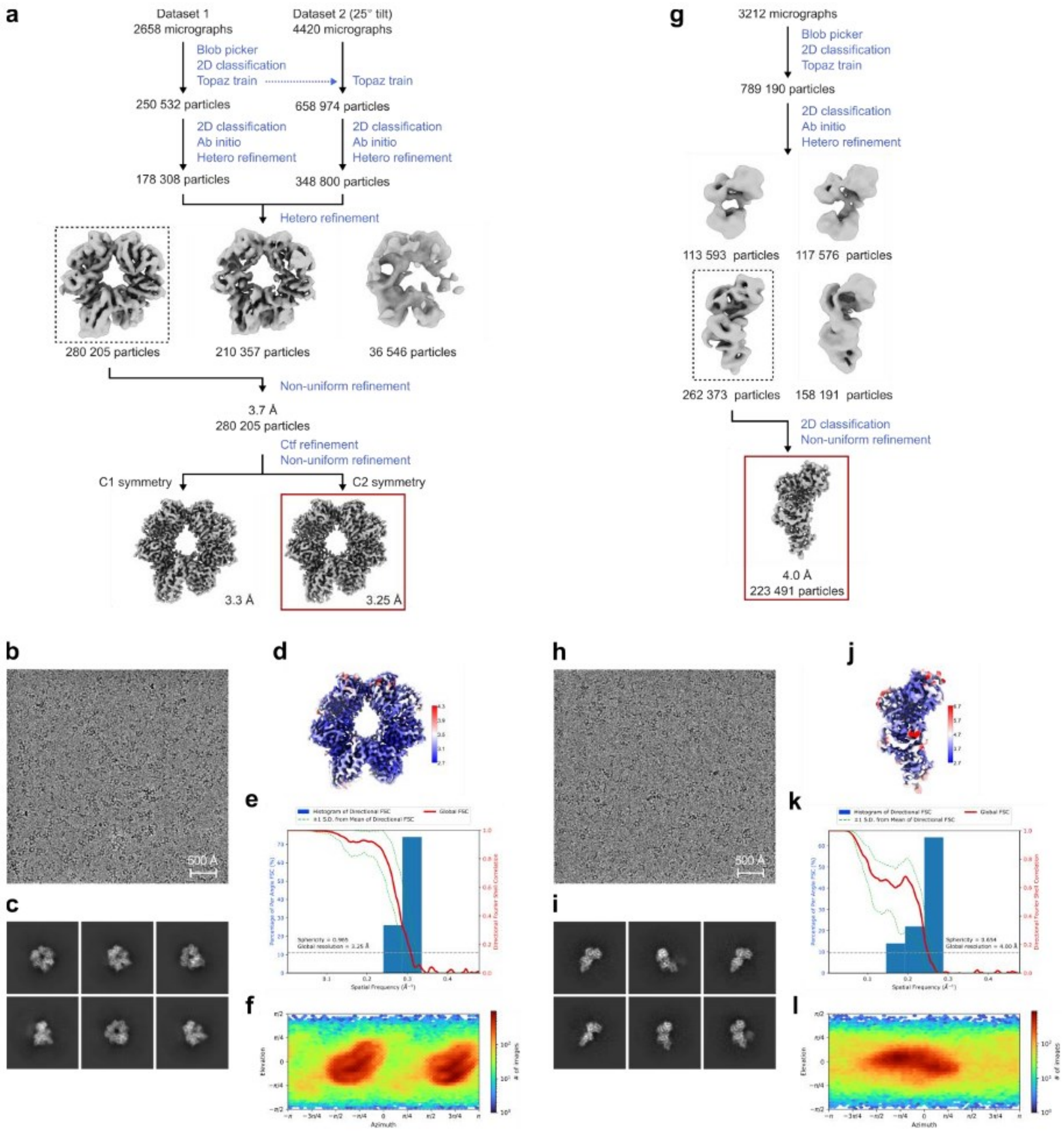


Supplementary Fig. 7: SLFN11 ATPase is locked in autoinhibited state. a, Fluorescence-based ATPase assay of SLFN11 (negative control: buffer; positive control: hexokinase). The fluorescence of NADH is measured over time. ATP consumption is enzymatically coupled to the oxidation of NADH, causing a decrease in fluorescence upon ATP hydrolysis. The details of individual substrates are shown in Supplementary Table 2. **b**, Structural comparison of SLFN11 helicase domain with PcrA SF1A helicase (PDB code 3PJR) and Dda helicase SF1B (PDB code 3UPU). The characteristic Pin in purple required for splitting of the incoming DNA duplex is missing in SLFN11. Source data for **a** are provided as a Source Data file.



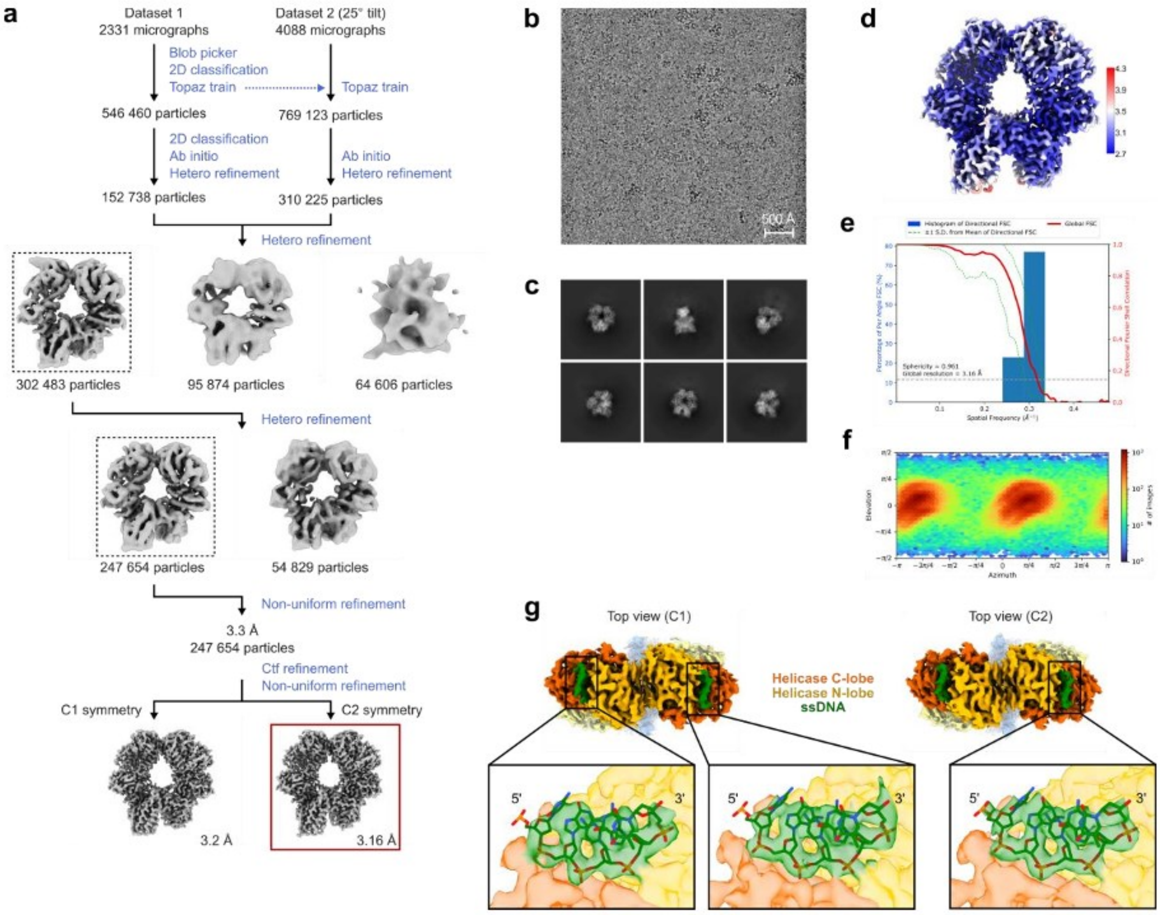
Supplementary Fig. 8: Cryo-EM data analysis of SLFN11^{wt} and SLFN11^{wt} with tRNA. **a**, Cryo-EM data processing workflow of SLFN11^{wt} and SLFN11^{wt} with tRNA using cryoSPARC[3]. **b**, Representative micrograph of SLFN11^{wt} and SLFN11^{wt} with tRNA dataset. The displayed micrograph is representative of 8,569 movies collected. **c**, Representative classes of a 2D classification of the particles used for the final SLFN11^{wt} reconstruction. **d**, Visualization of local resolution of SLFN11^{wt} calculated in cryoSPARC. Blue indicates higher resolution and red indicates lower resolution. **e**, Histogram of directional Gold-standard Fourier shell correlation (FSC)[4] (blue) and global FSC curve (red) of the final SLFN11^{wt} reconstruction.

The spread of directional resolution values is defined as ±1σ (dashed green lines). The grey dashed line indicates the 0.143 cutoff criterion, indicating a nominal resolution of 2.86 Å. **f**, Angular distribution of the particles used for the final SLFN11^{wt} reconstruction. **g**, Representative classes of a 2D classification of the particles used for the SLFN11^{wt} with tRNA reconstruction. **h**, Visualization of local resolution of SLFN11^{wt} with tRNA calculated in cryoSPARC. Blue indicates higher resolution and red indicates lower resolution. **i**, Histogram of directional FSC[4] (blue) and global FSC curve (red) of the final SLFN11^{wt} with tRNA reconstruction. The spread of directional resolution values is defined as ±1σ (dashed green lines). The grey dashed line indicates the 0.143 cutoff criterion, indicating a nominal resolution of 3.98 Å (no mask applied). **j**, Angular distribution of the particles used for the final SLFN11^{wt} with tRNA reconstruction.



Supplementary Fig. 9: Cryo-EM data analysis of SLFN11^{E209A} dimer and monomer. **a**, Cryo-EM data processing workflow of SLFN11^{E209A} dimer using cryoSPARC[3]. **b**, Representative micrograph of SLFN11^{E209A} dimer. The displayed micrograph is representative of 7,078 movies collected. **c**, Representative classes of a 2D classification of the particles used for the final SLFN11^{E209A} dimer reconstruction. **d**, Visualization of local resolution of SLFN11^{E209A} calculated in cryoSPARC. Blue indicates higher resolution and red indicates lower resolution. **e**, Histogram of directional FSC[4] (blue) and global FSC curve (red) of the final SLFN11^{E209A} dimer reconstruction. The spread of directional resolution values is defined as ±1σ (dashed green lines). The grey dashed line indicates the 0.143 cutoff criterion, indicating a nominal resolution of 3.25 Å. **f**, Angular distribution of the particles used for the final

SLFN11^{E209A} dimer reconstruction. **g**, Cryo-EM data processing workflow of SLFN11^{E209A} monomer using cryoSPARC[3]. **h**, Representative micrograph of SLFN11^{E209A} monomer. The displayed micrograph is representative of 3,212 movies collected. **i**, Representative classes of a 2D classification of the particles used for the final SLFN11^{E209A} monomer reconstruction. **j**, Visualization of local resolution of SLFN11^{E209A} monomer calculated in cryoSPARC. Blue indicates higher resolution and red indicates lower resolution **k**, Histogram of directional FSC[4] (blue) and global FSC curve (red) of the final SLFN11^{E209A} monomer reconstruction. The spread of directional resolution values is defined as $\pm 1\sigma$ (dashed green lines). The grey dashed line indicates the 0.143 cutoff criterion, indicating a nominal resolution of 4.0 Å. **l**, Angular distribution of the particles used for the final SLFN11^{E209A} monomer reconstruction.



Supplementary Fig. 10: Cryo-EM data analysis of SLFN11^{wt} with ssDNA. a, Cryo-EM data processing workflow of SLFN11^{wt} with ssDNA using cryoSPARC[3]. **b**, Representative micrograph of SLFN11^{wt} with ssDNA. The displayed micrograph is representative of 6,419 movies collected. **c**, Representative classes of a 2D classification of the particles used for the final SLFN11^{wt} with ssDNA reconstruction. **d**, Visualization of local resolution of SLFN11^{wt} with ssDNA calculated in cryoSPARC. Blue indicates higher resolution and red indicates lower resolution. **e**, Histogram of directional FSC[4] (blue) and global FSC curve (red) of the final SLFN11^{wt} with ssDNA reconstruction. The spread of directional resolution values is defined as $\pm 1\sigma$ (dashed green lines). The grey dashed line indicates the 0.143 cutoff criterion, indicating a nominal resolution of 3.16 Å. **f**, Angular distribution of the particles used for the final SLFN11^{wt} with ssDNA reconstruction. **g**, Top views of cryo-EM reconstructions of SLFN11^{wt} with ssDNA with C1 (left) or C2 (right) symmetry imposed. Helicase lobes and DNA are colour coded. Bottom: Detailed views of the cryo-EM densities around the ssDNA for the C1 (left) and C2 (right) reconstructions.

Supplementary Tables

Supplementary Table 1: Cryo-EM data collection, refinement and validation statistics.

	SLFN11 WT (EMD-14690) (PDB 7ZEL)	SLFN11 E209A dimer (EMD-14691) (PDB 7ZEP)	SLFN11 E209A monomer (EMD-14693)	SLFN11 WT + ssDNA (EMD-14692) (PDB 7ZES)	SLFN11 WT + tRNA (EMD-14695)
Data collection and processing					
Magnification	130,000	130,000	130,000	130,000	130,000
Voltage (kV)	300	300	300	300	300
Electron exposure (e$^-$/Å^2)	49.7	44.0	43.6	43.3	49.7
Defocus range (μm)	-1.1 to -2.9	-1.1 to -2.9	-1.1 to -2.9	-1.1 to -2.9	-1.1 to -2.9
Pixel size (Å)	1.046	1.046	1.046	1.046	1.046
Symmetry imposed	C1	C2	C1	C2	C1
Initial particle images (no.)	2,756,604	909,506	789,190	1,315,583	2,756,604
Final particle images (no.)	522,494	280,205	223,491	247,654	96,514
Map resolution (Å)	2.86	3.25	4.0	3.16	3.98
FSC threshold	0.143	0.143	0.143	0.143	0.143
Refinement					
Initial model used	AlphaFold	AlphaFold		AlphaFold	
Map sharpening *B* factor (Å^2)	132.7	157.3		146.8	
Model composition					
Non-hydrogen atoms	13,334	13,324		13,538	
Protein residues	1,654	1,654		1,654	
Ligands	2 Mg, 2 Zn	2 Zn		2 Mg, 2 Zn	
Nucleotide				10	
B factors (Å^2)					
Protein	42.40	52.29		42.02	
Ligand	57.56	91.95		57.92	
Nucleotide				84.10	
R.m.s. deviations					
Bond lengths (Å)	0.003	0.002		0.005	
Bond angles (°)	0.569	0.475		0.609	
Validation					
MolProbity score	1.25	1.27		1.39	
Clashscore	4.85	5.15		5.90	
Poor rotamers (%)	0.00	0.00		0.00	
Ramachandran plot					
Favored (%)	98.41	98.29		97.68	
Allowed (%)	1.59	1.71		2.32	
Disallowed (%)	0.00	0.00		0.00	

Supplementary Table 2: List of oligonucleotides (5’ to 3’).

Purpose	Name	Sequence	Modification
Cloning	SLFN11_E209A_fwd	CCATTTCCCGCCAGCCAGCTCGTG	
	SLFN11_E209A_rev	CACGAGCTGGCTGGCGGGAAATGG	
	SLFN11_E214A_fwd	CCAGCTCGTGGCCTTCAAGCAGTTTTCC	
	SLFN11_E214A_rev	GGAAAACTGCTTGAAGGCCACGAGCTGG	
	SLFN11_K216A_fwd	CTCGTGGAGTTCGCGCAGTTTTCCACC	
	SLFN11_K216A_rev	GGTGGAAAACTGCGCGAACTCCACGAG	
	SLFN11_Y234A_fwd	CAATCCCTGAGGCGGTGCCAGCCTTC	
	SLFN11_Y234A_rev	GAAGGCTGGCACCGCCTCAGGGATTG	
	SLFN11_D252A_fwd	GTTTATCGGCGTGGACGCGAAGAGCAGAGAGGTG	
	SLFN11_D252A_rev	CACCTCTCTGCTCTTCGCGTCCACGCCGATAAAC	
	SLFN11_K652D_fwd	GCCGAGACAAGAGACACATTCCTGCGG	
	SLFN11_K652D_rev	CCGCAGGAATGTGTCTCTTGTCTCGGC	
	SLFN11_R82D_fwd	GAGCAGAGCCTGGATGAGCTGATCCAG	
	SLFN11_R82D_rev	CTGGATCAGCTCATCCAGGCTCTGCTC	
	SLFN11_K591D_fwd	CGGTCCCTGAGAGATAACAGGGAGCTG	
	SLFN11_K591D_rev	CAGCTCCCTGTTATCTCTCAGGGACCG	
	SLFN11_Y722A_fwd	CTGTCTGACCAGGCTCCCAGGGAGGAG	
	SLFN11_Y722A_rev	CTCCTCCCTGGGAGCCTGGTCAGACAG	
	SLFN11_S753D_fwd	CGCTCTAACCCTGACTTCAATATCCC	
	SLFN11_S753D_rev	GGGATATTGAAGTCAGGGTTAGAGCG	
Biochemistry	tRNA^{Met}	AGUAAGGUCAGCUAAAUAAGCUAUCGGGCCCAUACCCCGAAAAUGU UGGUUAUACCCUUCCCGUACUACCA	5´-FAM
	tRNA^{Ser}	GUAGUCGUGGCCGAGUGGUUAAGGCGAUGGACUUGAAAUCCAUUGU GGUUUCCCCGCGCAGGUUCGAAUCCUGCCGACUACGCCA	5´-FAM
	ssDNA 50 nt	TCTTTTTTTTTTTTGTTCTTTTTTTTGATTCCGGTTTCTTTGAAATTTTT	5´-FAM
	dsDNA 50 bp	TCTTTTTTTTTTTTGTTCTTTTTTTTGATTCCGGTTTCTTTGAAATTTTT	5´-FAM
	dsDNA 2 50 bp	TGTCTACTTGAAATTCTAATTCATATTTTTTTTGTTTGGATAGAATATCA	
	dsDNA 3 50 bp	CAAAAGTGACATCCACAGCAAGCTGGACAGGTAAATTGCCTCATACAA TC	
	ssDNA 60 nt EM	CGCGTACGTGCGTTTAGAGCTTGCTACGTGCGTTTAAGCGGTGCTAGA GCTTGCTACGAC	
	ssDNA 60 nt MF	CGCGTACGTGCGTTTAGAGCTTGCTACGTGCGTTTAAGCGGTGCTAGA GCTTGCTACGAC	
	ssDNA 60 nt ATPase	CGCGTACGTGCGTTTAGAGCTTGCTACGTGCGTTTAAGCGGTGCTAGA GCTTGCTACGAC	
	dsDNA 60 bp ATPase	CGCGTACGTGCGTTTAGAGCTTGCTACGTGCGTTTAAGCGGTGCTAGA GCTTGCTACGAC	
	ssDNA 50 nt (Fig. 3e)	AATTGGTCGTAGCAAGCTCTAGCACCGCTTAAACGCACGTACGCGCTG TC	5´-FAM
	forked DNA	$(dT)_{40}$CTACGACCAATTGAGCGGCCTCGGCACCGGGATTCTCCAG	5´-Cy5
	ssRNA 41 nt	GACGGCCAUACCACCCUGAACGGCCCGAUCUUGUCUGAUCC	5´-FAM
	dsRNA 41 bp	GACGGCCAUACCACCCUGAACGGCCCGAUCUUGUCUGAUCC	5´-FAM

References

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

2 Perkins, D. N., Pappin, D. J., Creasy, D. M. & Cottrell, J. S. Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* **20**, 3551-3567 (1999).

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

4 Tan, Y. Z. *et al.* Addressing preferred specimen orientation in single-particle cryo-EM through tilting. *Nat. Methods* **14**, 793-796 (2017).

2.3. Structural mechanism of extranucleosomal DNA readout by the INO80 complex

Franziska Kunert*, Felix J. Metzner*, James Jung*, Markus Höpfler, Stephan Woike, Kevin Schall, Dirk Kostrewa, Manuela Moldt, Jia-Xuan Chen, Susanne Bantele, Boris Pfander, Sebastian Eustermann*, Karl-Peter Hopfner (2022) Structural mechanism of extranucleosomal DNA readout by the INO80 complex. *Science Advances*, Volume 8, Issue 49.

* These authors contributed equally.

DOI: 10.1126/sciadv.add3189
URL: https://www.science.org/doi/10.1126/sciadv.add3189

Summary

In this publication, we employ a combination of structural and biochemical methods to analyze how the INO80 chromatin remodeling complex interacts with nucleosomal and extranucleosomal DNA. Cryo-EM structures of regulatory INO80 A-modules from three different species (*C. thermophilum*, *S. cerevisiae*, and *H. sapiens*) reveal a conserved architecture, including a “2W-hairpin” motif that allows the binding of species-specific subunits. Structural alignments suggest the 2W-hairpin to be an evolutionarily conserved Actin-Arp4 anchor that is found throughout A-modules of INO80 and SWI/SNF family remodelers. Structures of the INO80 A-module bound to DNA as well as biochemical and yeast in vivo studies clarify the mode of DNA binding by the A-module. Different interactions between the A-module and extranucleosomal DNA, which are largely mediated by the Ino80 HSA domain and the Arp8 N-terminus, can have either inhibitory or stimulatory effects on nucleosome sliding. A reconstruction of the INO80 A- and C-module bound to a nucleosome shows how the A-module senses extranucleosomal DNA and is coupled to the motor domain by a continuous Ino80 HSA/post-HSA helix. This mechanical connection may underlie the allosteric regulatory function of the A-module. The ATPase rate of the motor domain is further regulated by an Arp5 insertion domain that contacts entry DNA and appears essential for nucleosome sliding. DNA bending at the A-module and at the INO80 motor domain suggest a mechanism for DNA shape readout and may explain the unusual position of the motor domain at SHL -6. Together, our data provide insight into the regulation of the INO80 complex by nucleosomal and extranucleosomal DNA.

Author contributions

I cloned and purified the *S. cerevisiae* A-module. Together with James Jung, I prepared grids of the *S. cerevisiae* A-module in different nucleotide states for cryo-EM analysis. Cryo-EM data collection and processing of the *S. cerevisiae* A-module datasets was done by James Jung and me. Furthermore, I performed fluorescence anisotropy assays to investigate the impact of mutations within the *C. thermophilum* A-module on DNA binding. Lastly, I contributed to manuscript preparation.

STRUCTURAL BIOLOGY

Structural mechanism of extranucleosomal DNA readout by the INO80 complex

Franziska Kunert[1]†, Felix J. Metzner[1]†, James Jung[1]†‡, Markus Höpfler[2]§||, Stephan Woike[1]§, Kevin Schall[1]¶, Dirk Kostrewa[1], Manuela Moldt[1], Jia-Xuan Chen[3], Susanne Bantele[2]#, Boris Pfander[2]††, Sebastian Eustermann[1,4]†, Karl-Peter Hopfner[1]***

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

INTRODUCTION

Chromosomal DNA is predominantly organized in the form of nucleosome core particles (NCPs)—~147 base pairs (bp) of DNA wrapped around the histone octamer (two copies of histones 2A, 2B, 3, and 4)—along with interspersed extranucleosomal linker DNA as well as larger nucleosome-free regions (NFRs) or nucleosome-depleted regions (NDRs) (*1*). NFRs and NDRs are important regulatory regions and are found at promoters, enhancers, and origins of replication in *Saccharomyces cerevisiae* (*2*). Nucleosomal packaging not only condenses and protects DNA but also generates epigenetic information in the form of nucleosome occupation, histone modifications, and histone variant composition (*2*).

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

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

A comprehensive mechanistic framework for the different biochemical activities of INO80 and how they are regulated or work together is still largely elusive. For instance, the detailed structural mechanism by which INO80 determines +1 and −1 nucleosome positions remains unclear. Not only NFR located barrier factors such as *S. cerevisiae* Reb1 and DNA ends but also NFR features such as promoter DNA mechanics and shape recently emerged as regulators of INO80-mediated nucleosome positioning in whole-genome chromatin reconstitutions (*15*–*17*). In mammals, INO80 might be regulated, in part, by the DNA sequence because the DNA binding transcription factor YY1 (Yin Yang 1), an early developmental regulator and structuring factor of promoter-enhancer elements, is a component of the human INO80 complex. Together, current evidence suggests that INO80 acts as an information processing hub that integrates diverse sources of information to properly shape chromatin around gene promoter regions (*15*, *16*).

Structural studies on INO80 and other remodelers revealed basic principles of how these molecular machines (or subcomplexes) bind

---

[1] Gene Center, Department of Biochemistry, Ludwig-Maximilians-Universität München, Munich, Germany. [2] DNA Replication and Genome Integrity, Max Planck Institute of Biochemistry, Martinsried, Germany. [3] Institute of Molecular Biology (IMB), Mainz, Germany. [4] European Molecular Biology Laboratory (EMBL), Heidelberg, Germany.
†These authors contributed equally to this work.
‡Present address: Australian Infectious Diseases Research Centre, School of Chemistry and Molecular Biosciences, The University of Queensland, Brisbane, QLD, Australia.
§These authors contributed equally to this work.
||Present address: MRC Laboratory of Molecular Biology, Francis Crick Avenue, Cambridge CB2 0QH, UK.
¶Present address: Crelux GmbH, Martinsried, Germany.
#Present address: Novo Nordisk Foundation Center for Protein Research, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark.
**Present address: Genome Maintenance Mechanisms in Health and Disease, Institute of Aerospace Medicine, German Aerospace Center (DLR), Cologne, Germany.
††Present address: Genome Maintenance Mechanisms in Health and Disease, Institute of Genome Stability in Ageing and Disease, CECAD Research Center, University of Cologne, Cologne, Germany.
*Corresponding author. Email: hopfner@genzentrum.lmu.de

nucleosomes and mobilize nucleosomal DNA using cycles of ATP binding and hydrolysis (*18*–*25*). Even in light of this process, we are far from understanding how complex remodeling reactions are carried out in a highly regulated manner, owed in part to their complex, dynamic, and modular architecture. INO80 contains more than 15 subunits, organized in three structural modules that we denote "N," "A," and "C." Up to now, structural information is available for the C-module bound to the nucleosome, as well as parts of the A-module in the absence of DNA. The Ino80 polypeptide itself carries the core ATPase motor activity and acts as a scaffold for the three modules. The C-module is the core nucleosome sliding unit: It contains the Swi2/Snf2 ATPase motor domain of Ino80p ($\text{Ino80}^{\text{motor}}$), the scaffolding AAA^{+} ATPases Rvb1 and Rvb2, and nucleosome binding subunits Ies2 (Ino eighty subunit 2), Ies6, and Arp5 (actin-related protein 5) (*20*, *21*). The NCP is bound by $\text{Ino80}^{\text{motor}}$-Ies2 at DNA superhelical location SHL-6 and by Arp5-Ies6 at DNA SHL-2. Furthermore, the Arp5 "grappler" insertion domain interacts with the nucleosome "acidic patch," a motif at the H2A/H2B interface that is a binding site for numerous chromatin proteins (*20*). In this configuration, $\text{Ino80}^{\text{motor}}$ pumps extranucleosomal entry DNA into the NCP, a model that can explain its sliding activity (*10*, *20*, *22*). The function of N- and A-modules is less clear. The N-module is evolutionarily rather divergent, binds DNA, and has autoregulatory functions to ensure switch-like activation of INO80 by extranucleosomal DNA (*26*). The A-module is highly conserved in evolution and contains an HSA (helicase-SANT–associated) domain ($\text{Ino80}^{\text{HSA}}$) in the middle of the Ino80p polypeptide chain, along with actin (Act1), Arp4, Arp8, Ies4, and Taf14. The complex of $\text{Ino80}^{\text{HSA}}$ with Arp4, actin, and Arp8 has been crystallized, and low-resolution structural along with functional analysis suggests that the $\text{Ino80}^{\text{HSA}}$ domain acts as an extranucleosomal DNA sensor, which is required for robust nucleosome sliding (*27*, *28*) and positioning in whole-genome chromatin reconstitution (*15*, *16*).

It is yet unclear how the A-module binds DNA and how it regulates the C-module. A-modules are found in all multisubunit

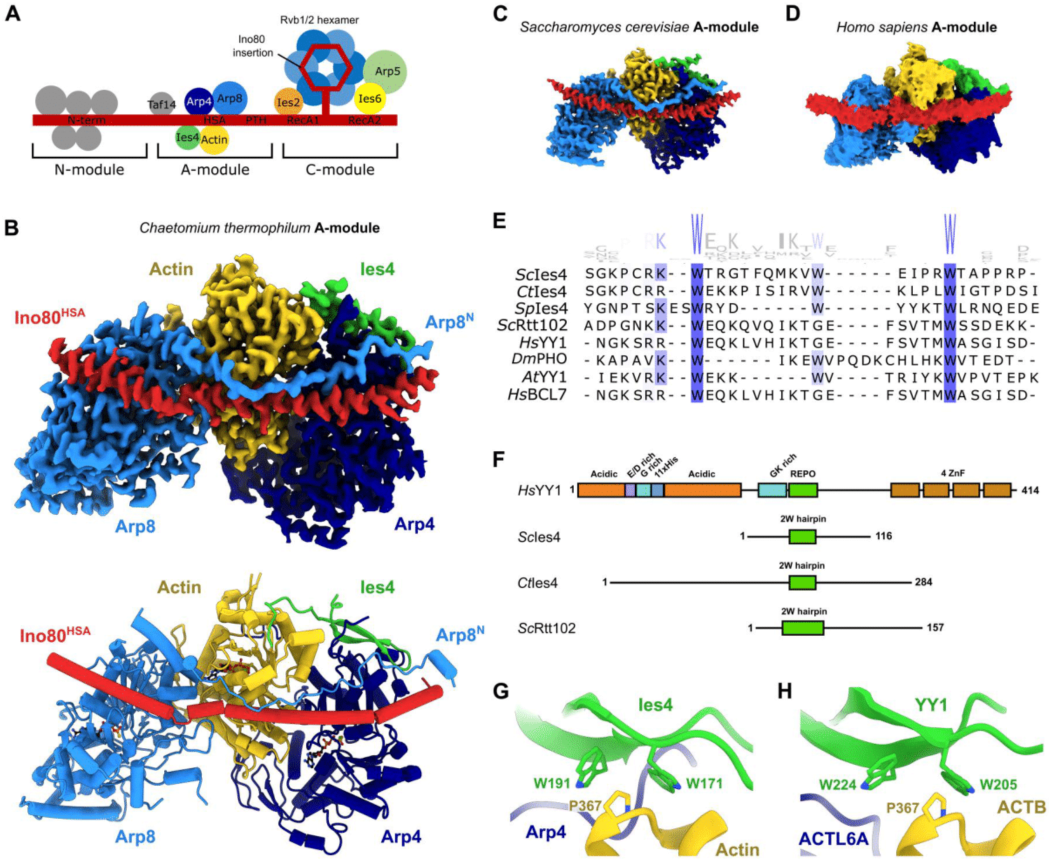


Fig. 1. Structure of the INO80 A-module. (**A**) Schematic of INO80 complex submodule and subunit organization. (**B**) Cryo-EM reconstruction (top) and structural model (bottom) of *C. thermophilum* (*Ct*) A-module. The protein subunits are color-coded and annotated. (**C**) Cryo-EM reconstructions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

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

RESULTS

Architecture of the INO80 regulatory A-module

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

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

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

The C-terminal part of Arp8 N-terminal extension forms a helix that binds into the interface of actin and Arp8 and thus might be affected by the nucleotide state of the actin folds. To test this, we imaged A-modules in the presence of different nucleotides (fig. S1, A to C). In the order ADP>ATPγS>ATP, we observe a very small conformational change in the *S. cerevisiae* Arp8-actin pair and an ordering of the N-terminal segment of Arp8 along actin and Arp4 in the ATP state (fig. S1, D and E). This may indicate a potential differential role of ATP/ADP at *S. cerevisiae* Arp8. Typical for actin fold proteins, the underlying conformational changes are very subtle, making it difficult to distinguish them from experimental variability in the cryo-EM analyses at this stage. In the case of *C. thermophilum* A-module, imaging without nucleotides (DNA-bound classes) or adding ATPγS (without DNA) resulted in the presence of ATP/ATPγS at the nucleotide binding sites of all three actin fold proteins (fig. S2, A and B). In any case, in both *S. cerevisiae* and *C. thermophilum* A-modules, we observed constitutive ATP binding at Arp4 and actin, while nucleotide binding to *S. cerevisiae* Arp8 is at least variable (figs. S1, A to C, and S2, A and B).

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

The two tryptophans bind a Gly^{366}-Pro^{367} linker between the last two helices of actin (subdomain I; Fig. 1G and figs. S2C and S3A). Here, Pro^{367} is situated in an aromatic "corner" formed by the nearly right-angled tryptophan side chains. A similar type of interaction to human ACTB (β-actin) Pro^{367} through two tryptophans organized in a β-stranded structure is seen in the extracellular actin sensor C-type lectin DNGR-1, suggesting a more widely evolved actin interaction principle (fig. S3B) (*30*). Furthermore, the β-stranded fold and the presence of two tryptophans are broadly similar to the WW domains that bind proline-rich peptides (fig. S3C) (*31*, *32*).

Ies4 also interacts with the N-terminal tail of Arp8 as well as with Ino80^{HSA} (Fig. 1B and fig. S2, D and E). These contacts are mediated by the tip of the β-hairpin element and are probably important to assemble a "defined" INO80 A-module because the Arp4-actin pair is also present in BAF/PBAF, SWR1, and NuA4 complexes as part of different molecular assemblies. The 2W-hairpin motifs of Ies4 are furthermore interesting, as they resemble the structure of Rtt102 bound to Arp7-Arp9. Arp7-Arp9 are the orthologs of Arp4-actin in *S. cerevisiae* SWI/SNF family remodelers SWI/SNF and RSC

(Fig. 1, E and F). Rtt102 displays a similar 2W attachment to Arp9 as Ies4 (Fig. 1E and fig. S3D), revealing an architectural conservation of A-modules across INO80 and SWI/SNF remodelers that goes well beyond the Arp4-actin pair and the HSA domain.

Mammalian YY1 is the structural homolog of yeast and fungal Ies4

Mammalian INO80 does not have a clearly recognizable Ies4 homolog based on sequence conservation. However, YY1, a GLI/Krüppel-like transcription factor associated with chromosome loop formation, stem cell biology, and early development, has been shown to interact with a module of human INO80 containing $\text{INO80}^{\text{HSA}}$, ACTL6A (Arp4 homolog), and ACTR8 (Arp8 homolog) (*33*). To see whether YY1 could be the evolutionary ortholog of Ies4, we produced recombinant *H. sapiens* A-module $\text{INO80}^{\text{HSA}}$, ACTL6A , ACTB, ACTR8 , and YY1. These proteins assemble in a stoichiometric and stable complex that we used for cryo-EM analysis (fig. S4A). From 25,652 particles, we obtained a map with a resolution of 7.5 Å (Fig. 1D and fig. S4B), but a high degree of particle orientation bias impeded a higher-resolution reconstruction. Still, it allowed unambiguous interpretation with models derived from the crystal structure of ACTR8and AlphaFold2 models of ACTL6Aand ACTB. In general, the arrangement of actin-related proteins and ACTB along $\text{INO80}^{\text{HSA}}$ is very similar to that found in fungal and yeast complexes (Fig. 1, B to D). After docking of the actin fold proteins, residual density at the hydrophobic rim of ACTL6A–ACTB matches very well the density corresponding to the hairpin region of Ies4 on the surface of yeast and fungal Arp4-actin (Fig. 1D). Sequence analysis (Fig. 1E) and AlphaFold2 prediction of YY1 indicated that residues 201 to 226 have the appropriate β-hairpin structure with two conserved, flanking tryptophans. This part has also been crystallized in a complex with the polycomb group protein MBTD1 and shows a 2W-hairpin motif (fig. S3E) (*34*). AlphaFold2 modeling of a complex of ACTL6A and the 2W-hairpin of YY1 (Fig. 1H) resulted in a complex that matches the corresponding surface density of the *Hs*A-module. Chemical cross-linking and mass spectrometry (CX-MS) also identifies a cross-link, consistent with this location of YY1 (fig. S4, C and D). Notably, binding of the 2W-hairpin motif (denoted also REPO domain) (*35*) to ACTL6A–ACTB is distinct from its interaction with MBTD1. Superposition of both complexes via the YY1 element indicates partially overlapping binding sites to the hairpin region (fig. S3F), which may explain partitioning of the *Drosophila* YY1 ortholog Pho into INO80 and *Drosophila melanogaster* polycomb group protein Sfmbt (*35*).

Besides the INO80 complex, we identified the 2W-hairpin motif in the AlphaFold2 predictions of complex subunits of INO80 family (*S. cerevisiae*: Ies4 in INO80, Swc4 in SWR1, and NuA4; *H. sapiens*: YY1 in INO80, DMAP1 in SRCAP, and TIP60) and SWI/SNF family remodelers (*S. cerevisiae*: Rtt102 in SWI/SNF and RSC; *H. sapiens*: BCL7 in BAF and PBAF) (*36*), hinting at a pervasive binding motif between remodeler families (fig. S3, G to I).

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

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

Previous biochemical work established that the INO80 A-module is important for extranucleosomal DNA recognition and nucleosome sliding in vitro (*27*, *28*). To this end, we previously identified a series of positively charged residues on HSA$^{\alpha1}$ and HSA$^{\alpha2}$ that, upon mutation to glutamines, severely affected the nucleosome sliding in vitro (denoted HSA$^{\text{Q1}}$ and HSA$^{\text{Q2}}$) (*27*). We introduced these mutants, along with *arp8*Δ, *arp8ΔN* (*28*), and a Walker B mutation in Ino80 that affects ATP hydrolysis (*ino80*^{E842A}) into *S. cerevisiae* (W303 background; tables S2 and S3). Because these mutants were designed before the experimental DNA complex was obtained (see below), we generated an additional set of more structure-informed (*C. thermophilum*) K/R→A mutants in *S. cerevisiae* HSA$^{\alpha2}$ (denoted HSA$^{\text{A2}}$), which led to similar effects as the HSA$^{\text{Q2}}$. We tested for viability under unchallenged conditions as well as in the presence of different stresses that had previously been linked to the INO80 function (*37*–*39*). While a wild-type (WT) *INO80* construct was able to complement the *INO80* deletion, *ino80-HSA*^{Q1} gave poor growth already at unchallenged conditions and was unable to support growth upon heat stress, in the absence of inositol, under anaerobic conditions or upon induction of a DSB (DNA double-strand break; Fig. 2A; see fig. S5, A to C, for expression levels of mutant proteins). *ino80-HSA*^{Q2} and *HSA*^{A2} cells showed similar but slightly less severe phenotypes. The *ino80-HSA*$^{Q1+Q2}$ double mutant was unable to support viability in W303 background, similar to strains lacking *INO80* or the *ino80*^{E842A} mutant (Fig. 2B), suggesting an additive contribution of DNA binding by HSA$^{\alpha1}$ and HSA$^{\alpha2}$. Furthermore, deletion of *ARP8* showed a growth phenotype under all stresses, but was only mildly affecting growth under nonperturbed conditions (Fig. 2C). Expression of *arp8ΔN* partially rescued the *arp8Δ* heat stress phenotype, but not the homologous recombination–dependent DSB repair function as tested in growth and ectopic recombination assays (Fig. 2, C and D). Together, these data validate the importance of putative DNA interacting residues of the HSA domain in rendering INO80 functional and indicate that the INO80 DNA binding surfaces might affect the diverse functional roles to different degrees.

Structural basis of DNA binding by the INO80 A-module

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

In both states, most DNA interactions are formed by $\text{Ino80}^{\text{HSA}}$, consistent with the robust effects of $\text{Ino80}^{\text{HSA}}$ mutations in the in vivo analysis. Additional interactions are contributed by the N-terminal extension of Arp8 and by Ies4 (Fig. 3, A and C). While most $\text{Ino80}^{\text{HSA}}$-mediated DNA interactions appear to be peripheral electrostatic interactions between Lys and Arg side chains and DNA, a central contact side is at HSA$^{\text{L1}}$, a loop that disrupts the $\text{Ino80}^{\text{HSA}}$

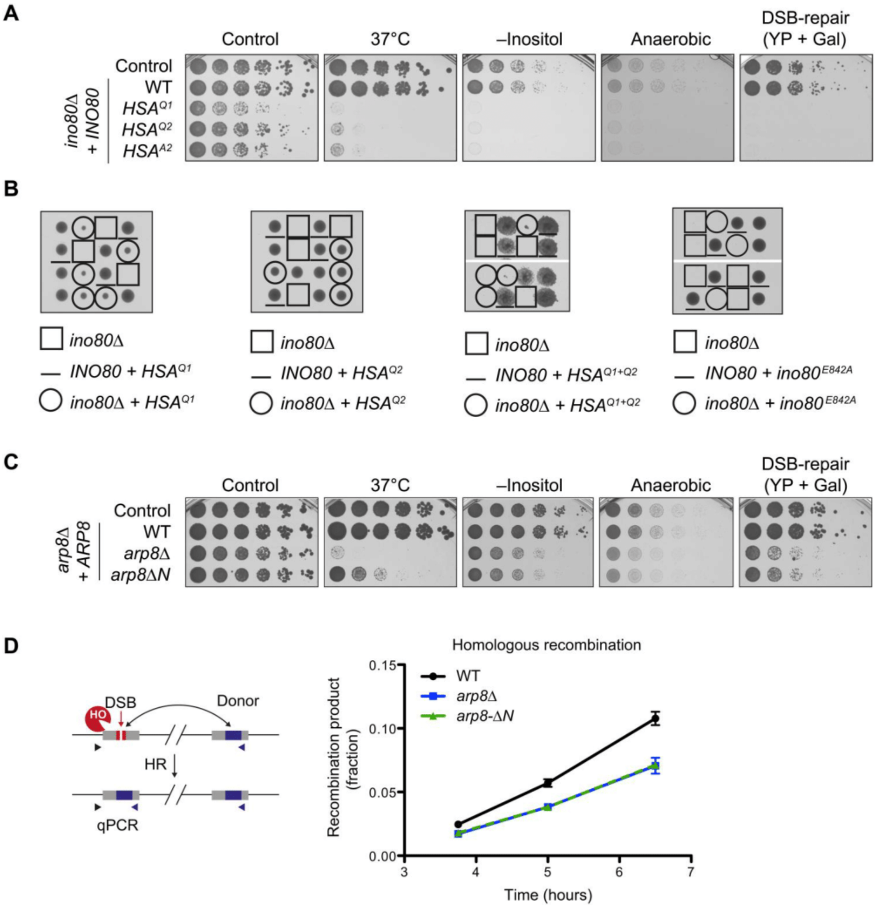


Fig. 2. HSA surface residues are critical for INO80 function in budding yeast. (**A**) Fivefold serial dilutions of budding yeast expressing indicated Ino80 variants were grown for 2 to 7 days. (**B**) Tetrad analysis of yeast cells bearing the indicated *INO80* alleles, where each row represents four colonies of a tetrad from a single diploid progenitor cell. *ino80-HSA*^{Q1} and *ino80-HSA*^{Q2} alleles (circled colonies in left two panels) partially rescue the *ino80Δ* lethality. The *ino80-HSA*$^{Q1+Q2}$ double mutant allele showed severely impaired or no growth (circled, third panel). The Walker B mutation (*ino80*^{E842A}) is lethal (fourth panel). (**C**) The Arp8 N-terminal region is critical for tolerance to elevated temperatures (37°C), growth on medium lacking inositol, under anaerobic conditions, and for DSB repair via homologous recombination (HR). *arp8Δ* cells were complemented with a full-length *ARP8* allele (WT) or an allele lacking the N-terminal 197 amino acids (*arp8ΔN*) and subjected to spot dilution growth assays as in (A). (**D**) The N-terminal region is required for Arp8 function in DSB repair by HR. Left: Schematic of the quantitative real-time PCR (qPCR)–based analysis of HR (*39*). Cells express a galactose-inducible HO endonuclease that cuts a single defined HO-cleavage site (red, ChrIV 491 kb). The DSB can be repaired by HR using a noncleavable donor site as repair template (blue, ChrIV 795 kb), and HR can be quantified by amplifying a recombination-dependent PCR product (triangles indicate primer positions). Right: Emergence of the recombination product after HO endonuclease induction (t = 0) was normalized to completed recombination (value = 1) for the strains indicated. n = 3, with error bars denoting SDs.

element into two helices HSA$^{\alpha 1}$ and HSA$^{\alpha 2}$ (Fig. 3B). Here, the N-terminal turn of HSA$^{\alpha 2}$ binds a DNA backbone phosphate through main-chain amide nitrogens. This interaction shows a remarkable similarity to the DNA interactions of the innate immune sensor cGAS and could provide a phosphate "registry lock" (*40*). The central contact side is reinforced by a "hook" element of the Arp8 N-terminus that binds to the DNA backbone as well as to two minor groove base pairs (Fig. 3C). The hook element is stabilized by Ies4 (residues 148 to 156), which is also in direct binding distance to the DNA backbone and may contribute further interactions. Similar folds of the hook region in the *apo* states of the *C. thermophilum* and *S. cerevisiae* A-modules suggest evolutionary conservation of this DNA binding element.

In the case of curved DNA, we also observe DNA contacts around SHL-11, mediated predominantly via the HSA$^{\alpha 1}$ region and a helix near the very N-terminus of Arp8^{N} (Fig. 3D). Binding of curved DNA is noteworthy, as it might be influenced by DNA mechanical properties. Geometrically, it is a result of the curved shape of Ino80^{HSA} at the Arp4-actin pair, which is incompatible with binding of linear DNA along the entire length of the A-module (Fig. 3, E and F).

A 7.5 Å resolution structure of *S. cerevisiae* A-module bound to DNA could also be reconstructed from 69,226 particles (fig. S1F). Here, we see predominantly contacts at HSA$^{\alpha 1}$ and the Arp8 N-terminal extension. Again, DNA appears to be curved at this side, but the rather low resolution prevents a more detailed analysis.

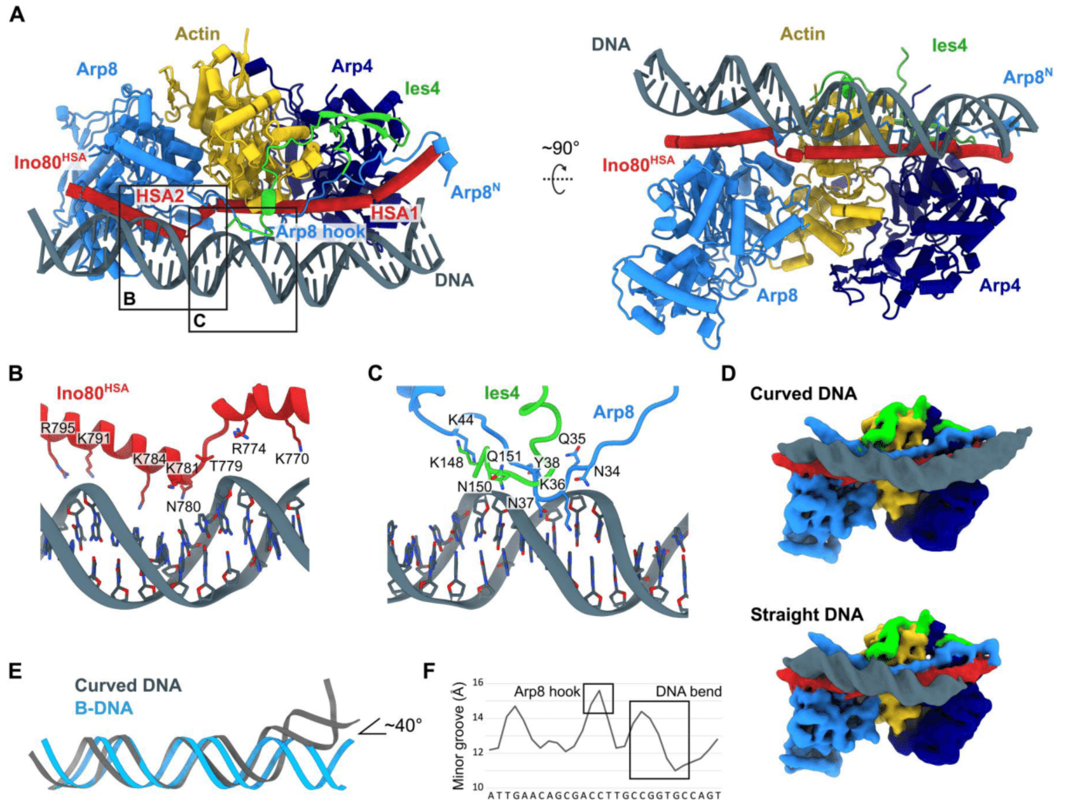


Fig. 3. Structure of the *C. thermophilum* A-module bound to DNA. (**A**) Structural model of *C. thermophilum* A-module bound to DNA. (**B**) Detailed view of the Ino80^{HSA}-DNA interaction. (**C**) Detailed view of the Arp8 hook and Ies4-DNA interaction. (**D**) Cryo-EM reconstructions of A-module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

Predominant binding of DNA at HSA$^{\alpha1}$ is consistent with the somewhat stronger growth defects of HSA^{Q1} mutations in *S. cerevisiae* in vivo (Fig. 2A).

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

Biochemical analysis

The observation that *Ct*INO80 A-module can bind both linear and curved DNA prompted us to perform more detailed biochemical studies to analyze the role of different DNA binding sites (Fig. 4, A to F) on *Ct*INO80$^{\Delta N}$ remodeling (Fig. 4G and fig. S6, A and B), *Ct*INO80$^{\Delta N}$ ATP hydrolysis (Fig. 4H and fig. S6C), and the A-module DNA binding activities (Fig. 4I and fig. S6D) in vitro. To this end, we evaluated various structure-derived mutations in the Ino80^{HSA} and Arp8 subunit (fig. S6, B, E, and F).

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

Arp8.1 (N34A, Q35A, K36A, N37A, Y38A, and K44A), carrying mutations in the hook as well as the N-terminal helix, leads to a reduction in sliding in the same range as Arp8$^{\Delta N}$, but this effect appears to be caused by defects other than a simple reduction of DNA affinity (Fig. 4, G and I). Again, this argues for a defective geometry of the active complex or a particular conformational state. The most remarkable effect showed the Ino80.1 mutant (K721A, K725A, R736A, and R740A), which carries mutations in the very distal extranucleosomal DNA binding region of Ino80^{HSA}. Ino80.1 strongly reduces binding to the DNA but increases sliding and ATPase rate of *Ct*INO80$^{\Delta N}$ (Fig. 4, G to I). These effects could be explained if the A-module can also negatively regulate INO80 and that such a role is affected by the Ino80.1 mutation.

In summary, the mutations all affect various functions and the biochemical properties of INO80, validating our structural results.

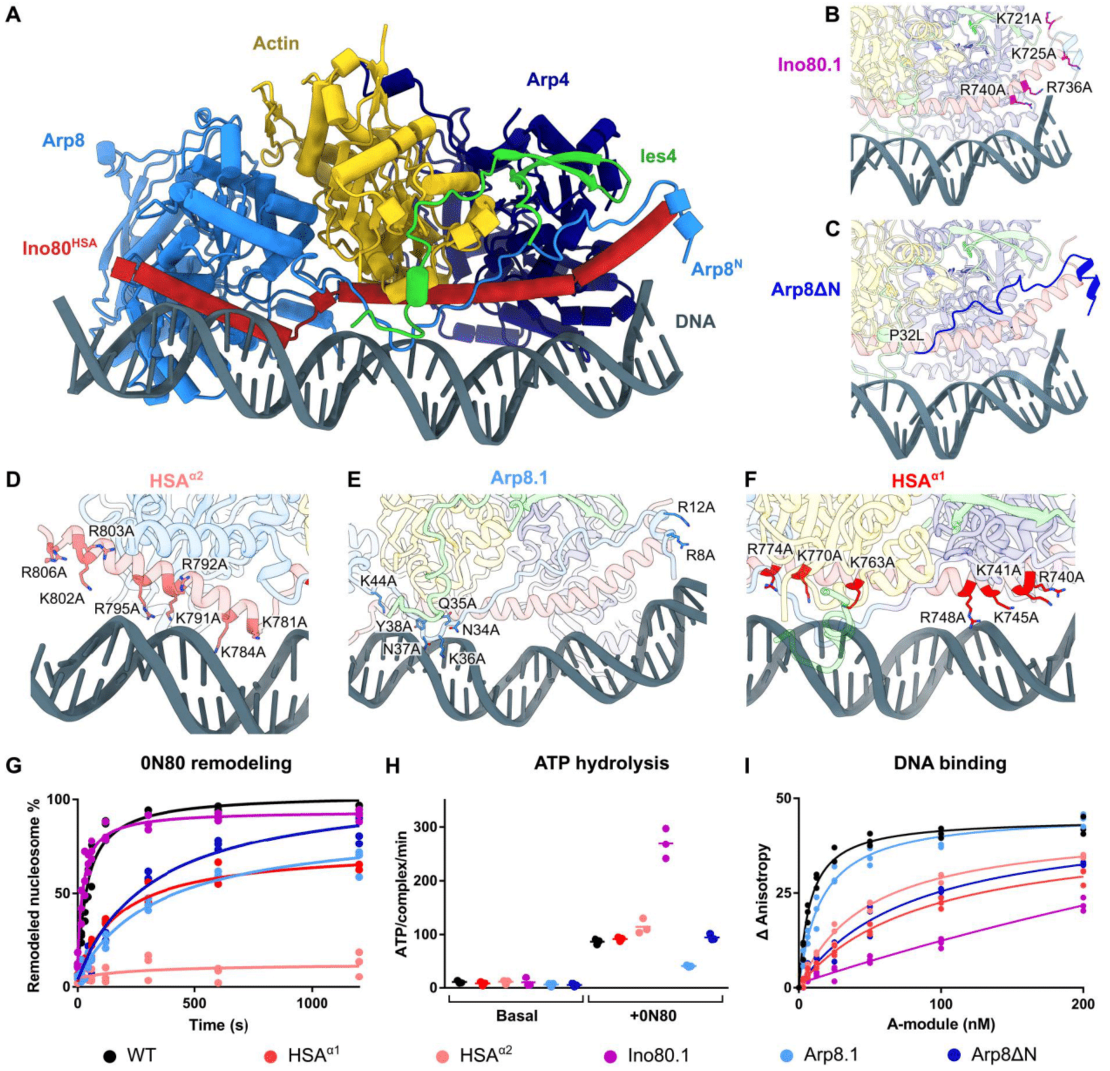


Fig. 4. Structural basis of DNA binding by the INO80 A-module. (**A**) Structural model of *C. thermophilum* A-module bound to DNA. (**B**) Ino80.1 mutation probes the distal region of HSA$^{\alpha 1}$. (**C**) Illustration of the truncated portion of the Arp8 N-terminus; P32L truncation site. (**D**) HSA$^{\alpha 2}$ mutations. (**E**) Arp8 mutations in hook and N-terminal helix. (**F**) HSA$^{\alpha 1}$ mutations probe the central region of HSA$^{\alpha 1}$. (**G**) Evaluation of the remodeling activity of *Ct*INO80$^{\Delta N}$ mutants. Band intensities of remodeled and unremodeled nucleosome species were quantified, and the fraction of remodeled nucleosomes was plotted against time. Data points were fitted using an exponential equation. Mean and individual data points ($n = 3$, technical replicates). (**H**) ATPase rate of *Ct*INO80$^{\Delta N}$ mutants with and without stimulation by nucleosomes. Rates were calculated from the linear area of the raw data and were corrected by a buffer blank. Mean and individual data points ($n = 3$, technical replicates). (**I**) Fluorescence anisotropy assay to monitor the binding of *C. thermophilum* A-module and mutants to a 50-bp DNA. The data were fitted to a nonlinear noncooperative 1:1 binding model. Individual data points of three independent experiments are plotted.

However, they indicate that the A-module plays a more complex, pleiotropic regulatory role with activating and inhibitory roles on remodeling.

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

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

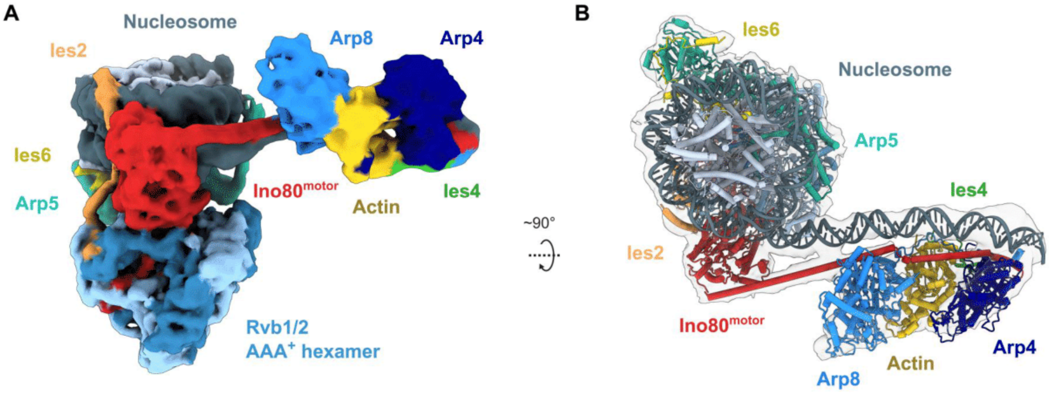


Fig. 5. The INO80 A-module interacts with extranucleosomal DNA. (**A**) Cryo-EM reconstitution (multibody refined) of the *C. thermophilum* A- and C-modules binding to nucleosome and extranucleosomal DNA. (**B**) Structural model of INO80$^{\Delta N}$ based on structures of C-module bound to the nucleosome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

We first focused on the motor domain to see what effects ATP binding have on the way Ino80$^{\text{motor}}$ interacts with DNA. In the *apo* state (no nucleotide), Ino80$^{\text{motor}}$ is well resolved and substantially bends DNA as previously described (Fig. 6A). In the *apo* state, we now see clear density for the post-HSA domain, which was missing in previous analyses. It interacts as a continuous, long helix with the N-lobe of the motor domain. It occupies the same region on the motor as the regulatory elements auto-N of ISWI and the post-

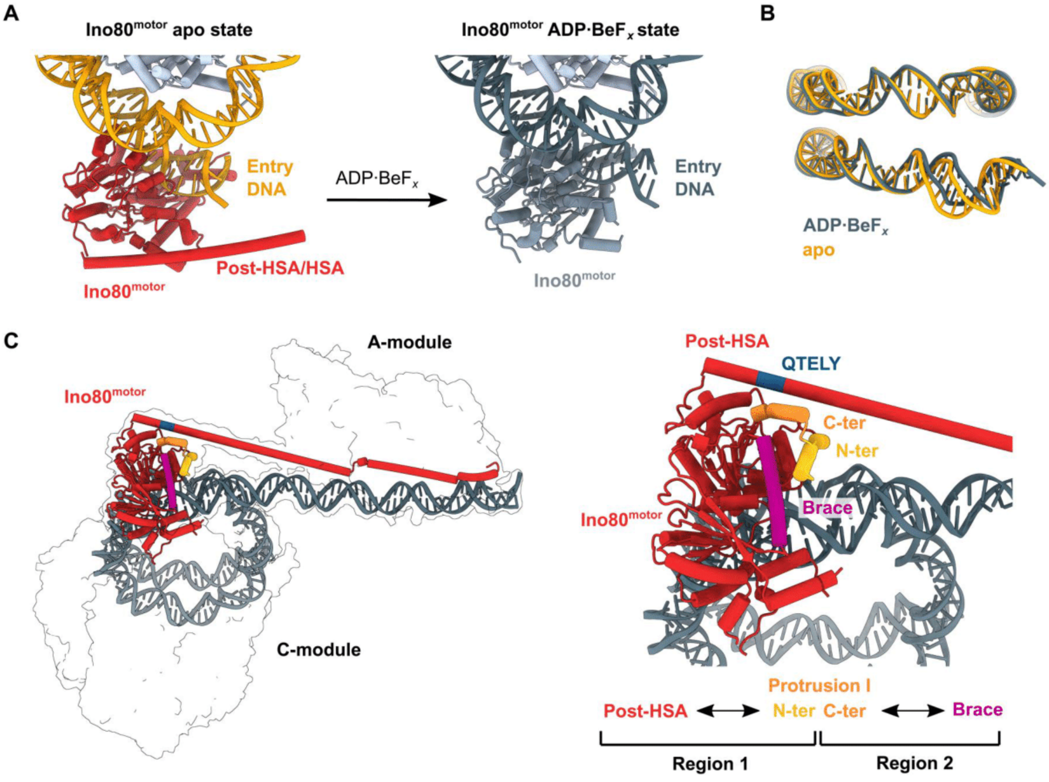


Fig. 6. Ino80$^{\text{motor}}$ conformations in *apo* and ADP•BeF$_x$ states. (**A**) Structural model of *C. thermophilum* Ino80$^{\text{motor}}$ interacting with the nucleosome at SHL-6 in *apo* state (left) and ADP•BeF$_x$ state (right). Structured post-HSA domain is visible in *apo* state. (**B**)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

HSA domain of Snf2 (*41*, *42*), showing a high degree of conservation of motor regulatory elements among different remodelers (fig. S7A). However, we note that the interactions of post-HSA domains of Ino80 and Swi2 are somewhat shifted, although other parts of the motor superimpose and match very well. It was previously suggested that movements of the post-HSA could be coupled to motor activation (*42*, *43*).

ADP•BeF$_x$ binding leads to a straightening of the DNA at the motor compared to the bent conformation with widened minor groove in the apo state (Fig. 6, A and B; fig. S7, B and C; and movie S2). Furthermore, in the ADP•BeF$_x$ bound state, the post-HSA domain is not visible anymore, suggesting that it is a rather dynamic feature that could be coupled to the nucleotide state of the motor and/or the relative location of the A-module with respect to the motor (see below; Fig. 5A).

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

In any case, the dynamics of the post-HSA motor contacts point toward a more profound allosteric communication between A- and C-modules on both ends of the HSA/post-HSA domain, and we set out to obtain an overall structure to see how A- and C-modules could communicate. While A- and C-modules appear to be generally mobile with respect to each other in most picked particles, we could identify a subset of particles in the dataset without nucleotide addition that showed a more defined orientation between the A- and C-modules. This set of particles resulted in a 7.7 Å map that allowed us to place high-resolution structures of A- and C-modules and model the entire HSA/post-HSA helix that links A- and C-modules (Fig. 5, A and B). In this structure, the A-module is situated at SHL-9 to SHL-11, orientated such that Arp8 faces the Ino80^{motor} domain, while Arp4 points away. The HSA/post-HSA region forms a continuous helix all the way from Arp8 to the N-lobe of the motor domain. In this state, the Ino80^{HSA} domain could even use further DNA contacts between Ino80^{motor} and Arp8, supported by several Lys/Arg side chains in the vicinity of DNA. Because of the structural flexibility and moderate resolution of this state, the DNA grooves are not well defined. However, modeling canonical B-DNA into the map indicates that the A-module is not exactly bound to DNA in the same way as we observe in the high-resolution individual reconstructions. It appears to be shifted along DNA by approximal $^1/_2$ helical turn, suggesting that the A-module might not be fully engaged and aligned with the DNA grooves in this state as observed on most of the DNA-bound classes of the A-module alone (Fig. 3, A to C). The observed configuration could be a nucleosome "sliding" state, where loose DNA binding of the A-module does not slow down nucleosome sliding, yet promotes post-HSA motor contacts. Such an interpretation would be consistent with the observation that some mutants in the HSA/A-module actually lead to a speeding up of nucleosome sliding, while others slow down or abolish sliding. In the ADP•BeF$_x$ dataset, we were not able to identify an equivalent subset of particles with well-defined arrangements of A- and C-modules. The absence of the post-HSA domain density may suggest a more dynamic mutual orientation of A- and C-modules. From sterically considerations, the different angle of entry DNA at the motor in the ADP•BeF$_x$ state may not allow mutual binding of the A-module to DNA along with a linear HSA/post-HSA domain between A-module and motor. It is therefore possible that post-HSA–motor interactions are quite dynamic in the ATPase cycle or additional conformations of HSA/post-HSA and their attachment to the motor exist. Recent functional and structural studies on the RSC complex (*25*) identified an intriguing structural arrangement of the post-HSA domain at protrusion I of the motor. Considering the conserved arrangement of these regulatory domains in Swi2/Snf2 ATPases, it is likely that an equivalent regulatory hub exists in INO80. Notably, the conserved QTELY motif, a homolog of the conserved SWI/SNF QTXX[F/Y] motif, forms the post-HSA contact site toward protrusion I, hinting at a critical interface for modular allostery by the A-module (Fig. 6C). Together, the mode of interaction between A- and C-modules through HSA/post-HSA, and its modulation by nucleotide binding at Ino80^{motor}, provides an obvious direct chemomechanical link between Ino80^{motor} and binding of the A-module to extranucleosomal DNA.

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

The improved maps and AlphaFold2 structure predictions allowed us to model the complete Arp5 protein, in particular its unique grappler insertion element (Fig. 7A). This led to clarification of the way the grappler "foot" binds the acidic patch of the nucleosome and allowed us to identify two additional critical DNA contacts (fig. S8A). As described previously (*20*), we observe two remarkably distinct grappler configurations (fig. S8, B and C). In the "parallel" state, its two main helical arms are arranged in a near-parallel fashion and bind DNA around the nucleosome dyad. In the "cross" configuration, one helical arm binds along the DNA gyre, placing its tip at the entry DNA opposite the motor domain. Using 3D variability analysis (movie S3), formation of contacts between the tip of the cross arm appears to coincide with a better-defined HSA/post-HSA and a properly curved entry DNA, suggesting a functional link. We noticed two patches of Arg/Lys residues in loop regions that are properly placed to interact with the entry DNA and may account for this effect. Although the density map is not good enough to directly visualize these loops, the supporting helical elements are nevertheless defined well enough to confidently provide a location for the positively charged loops using AlphaFold2 models (Fig. 7, B and C).

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

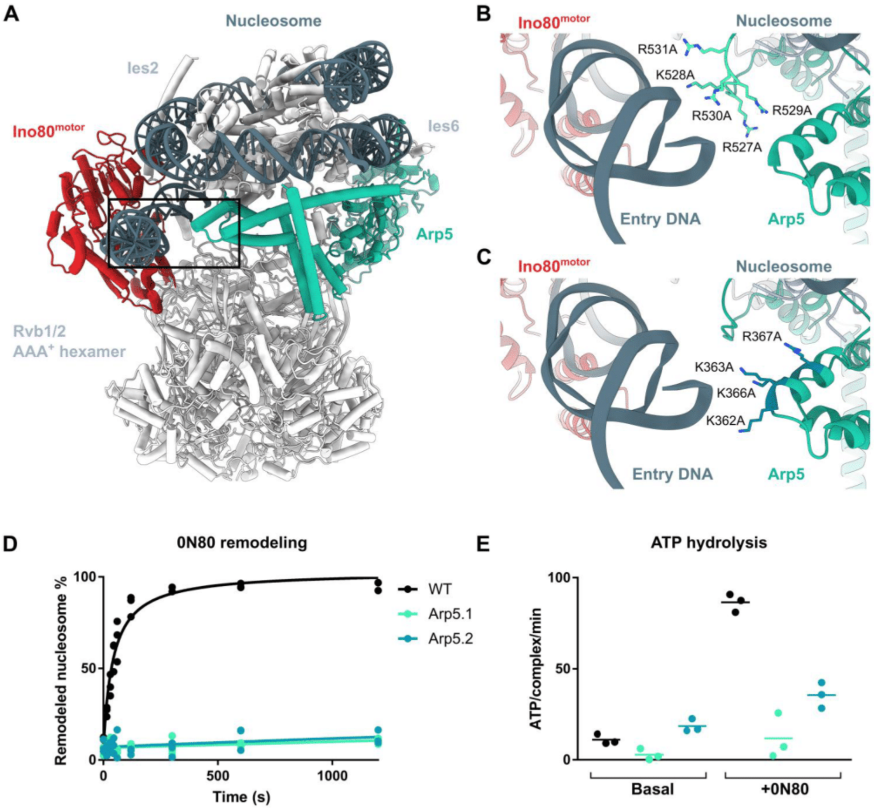


Fig. 7. The Arp5 grappler interacts with entry DNA and regulates the motor domain. (**A**) Structural model of the *C. thermophilum* C-module, highlighting Arp5 (green, in cross configuration), Ino80^{motor} (red), and nucleosomal DNA (dark gray). (**B** and **C**) Docking model of two loop regions at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

Regulation of remodeling by DNA features

Both the path of DNA around Ino80^{motor}/Arp5 and the A-module show curved DNA regions, which are geometrically linked with the relative placement of A- and C-modules and a linear HSA/post-HSA helix, or the binding of extranucleosomal DNA along the entire Ino80^{HSA} domain. Previous experimental and statistical analysis indicated that the *S. cerevisiae* INO80 remodeling activity is influenced by DNA shape/mechanical features in extranucleosomal DNA. To test the generality of these observations for the *C. thermophilum* complex and also clarify the contribution of different modules of INO80 to DNA feature readout, we replaced the sequence of our model substrate with an A/T-rich, rigid sequence cassette derived from the *URA3* promoter at four different locations (Fig. 8, A and B, and fig. S9A), probing contributions of distal (SHL-10/11) and proximal (SHL-8/9) extranucleosomal DNA binding sites of the A-module and the motor domain (SHL-6/7) and inside the nucleosome behind Ino80^{motor} (SHL-4/5).

Rigid DNA at SHL-8/9 and SHL-10/11 and inside the nucleosome (SHL-4/5) had a nearly equal, moderately reducing effect on nucleosome sliding by the *Ct*INO80$^{\Delta N}$ complex, whereas changing the DNA at the Ino80^{motor} binding site more markedly reduced sliding efficiency (Fig. 8C and fig. S9B). The sliding efficiencies did not correlate well with alterations in the ATPase rates of *Ct*INO80$^{\Delta N}$ because only the SHL-4/5 insertion had a reduced ATPase rate, while all nucleosome variants showed similar binding efficiency (Fig. 8D and fig. S9, C and D). ATP hydrolysis by Ino80^{motor} and sliding efficiency were also not correlated in the analysis of A-module mutations, hinting toward futile ATP cycles when stiff DNA is located at Ino80^{motor} or in extranucleosomal DNA. When DNA is inserted into the nucleosome, it is plausible that the underlying nucleosome is structurally weakened, leading to proficient sliding despite reduced ATPase rates. It should be noted that in this analysis, DNA elements are obviously pushed along different DNA binding sites during the remodeling reaction, and thus, the effects might be, to some extent, integrated. Nevertheless, the strongest effect is observed at the motor domain, which is also the site where DNA shows the most profound bend. In summary, these data show that inserting DNA cassettes with a stiff DNA sequence leads to a general reduction of nucleosome sliding, not only corroborating the influence of extranucleosomal DNA

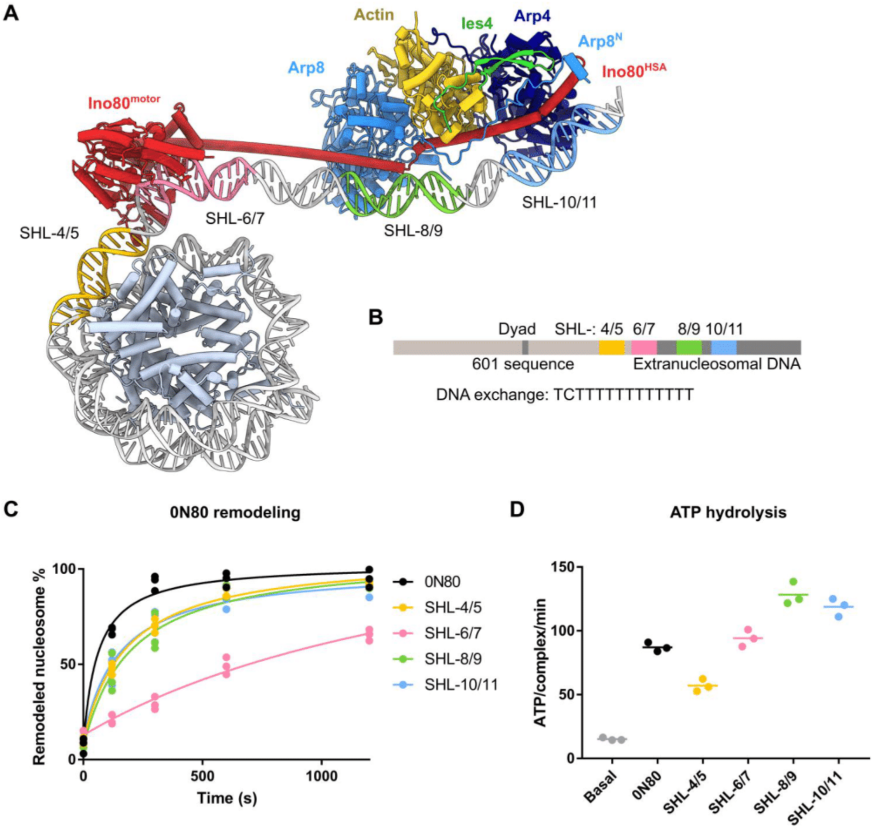


Fig. 8. Influence of A/T-rich DNA on *Ct*INO80$^{\Delta N}$ nucleosome remodeling. (**A**) Location of exchanged DNA cassettes on the structural docking model. (**B**) Schematic visualization of exchanged DNA sequence cassettes in distance to the dyad of the nucleosome. (**C**) Evaluation of the sliding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

sequence but also revealing that, in particular, the motor domain is sensitive to DNA features.

DISCUSSION

In the past years, groundbreaking structures of different remodelers bound to the nucleosome shed light on the basic principles of nucleosome recognition (*18*–*25*), while structural and functional analyses of selected single-subunit remodelers (*4*, *19*) suggested paths of allosteric activation by core nucleosome binding. While we begin to understand from these studies how remodelers grip and move DNA at nucleosomes, revealing the large-scale nucleosome reconfiguration steps and their regulation at atomic detail is the next frontier. For instance, the INO80 complex shows pleiotropic biochemical activities such as nucleosome spacing and editing, as well as the positioning of nucleosomes at NFR flanking regions (*12*). These diverse reactions depend on a basic nucleosome or hexasome sliding/mobilization activity (*11*), where the ATPase motor pumps extranucleosomal entry DNA into the nucleosome. To place a nucleosome at the +1 position, rather than sliding it further into the NFR, however, requires a regulation of the ATPase activity itself or the coupling between Ino80^{motor} and nucleosome sliding. Regulatory signals could arise when the remodeler encounters a neighboring nucleosome, a barrier factor at the NFR/NDR, and, at least in the case of *S. cerevisiae*, DNA with particular mechanical or shape features such as those found in NFRs/NDRs (*15*–*17*). Functional and previous structural work suggested that a key regulative principle is the sensing of extranucleosomal DNA by the INO80 A-module (*15*, *16*, *27*, *28*).

Here, we provide a structural basis for this regulation and reveal how INO80 interacts simultaneously with nucleosomal and extranucleosomal DNA (Fig. 9). This work extends the analysis of multisubunit remodelers from NCP binding to recognition of linker DNA and reveals how binding of extranucleosomal DNA by the A-module is chemomechanically coupled to the remodeling motor. We provide details of the linker DNA binding and identify multiple sites where DNA shape features might tune the biochemical activity. Hereby, in particular, not only do Arp8 and Arp5 subunits emerge as critical regulators but also the Ino80^{motor} domain itself might play a central role in DNA feature readout. These

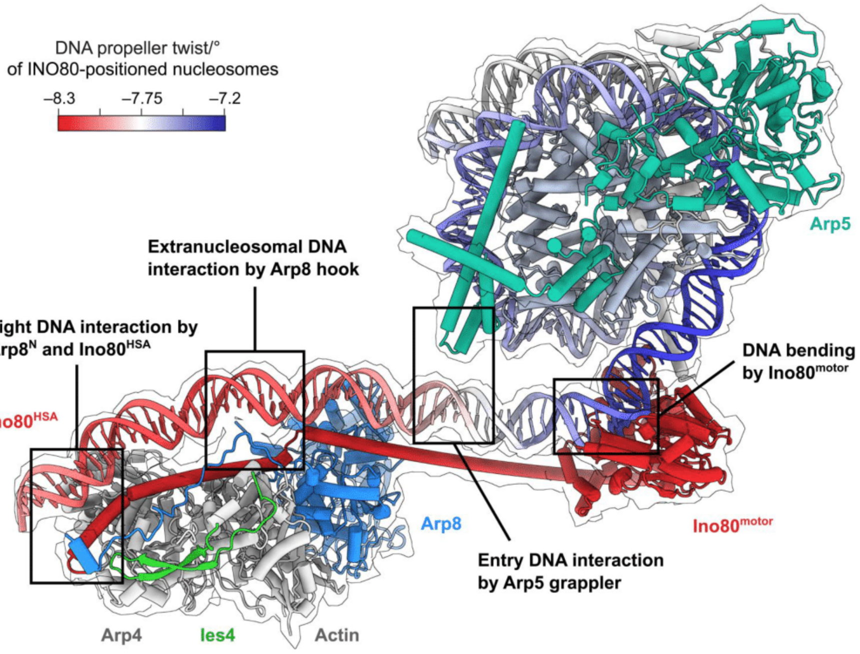


Fig. 9. Model of multivalent INO80$^{\Delta N}$-DNA interactions. The unified model integrates our structural and biochemical analysis. Ino80^{motor} engages the nucleosome at SHL-6. The Arp5 grappler contacts entry DNA opposite of Ino80^{motor}. The A-module binds extranucleosomal DNA and is linked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

properties of the motor domain could provide an explanation for the unique way INO80 has evolved to interact with the nucleosome, compared to other remodelers (*5*).

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

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

In YY1, the 2W-hairpin was characterized as the REPO (recruitment of polycomb) domain because it facilitates recruitment of polycomb group complexes (PcGs) (*34*). This mirrors the bivalent nature of YY1′s context-dependent transcriptional activation and repression: YY1 recruits either activating (INO80) or repressive (PcG) complexes to their respective genomic loci dependent on the cellular context. As an integral subunit of the *Hs*INO80 complex, YY1 has been implicated in the recruitment of INO80 to promoter sites. A coactivation between YY1 and INO80 was proposed because, as a transcription factor acting on accessible DNA, YY1 also relies on the INO80 nucleosome sliding activity to gain access to its cognate promoter sites (*33*). Epigenetic regulation of YY1 binding to DNA could also influence INO80 recruitment. YY1 binding is inhibited by methylation of certain CpG sites (*50*, *51*), which might conceivably control INO80 engagement, or activity, at promoter sites. Recently, a YY1-dependent recruitment of not only INO80 but also the BAF complex was shown in embryonic stem cells (*52*). Because BAF and INO80 share the BAF53^{Arp4}/

ACTB pair, our structural results offer a possible mechanistic explanation, although whether YY1 binds BAF's A-module like it binds the INO80 A-module needs further investigation.

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

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

In addition to the A-module, the Arp5 subunit emerges as a critical regulator of the remodeling reaction, thus identifying all actin-related proteins of the INO80 as core regulators. We consistently observe two major configurations in the helical insertion domain of Arp5 denoted grappler (*20*), which might point to rather complex functional roles in sliding or exchange reactions. While understanding the function of the parallel state and the precise role of the conformational switch needs to be addressed in future work, the cross state of the Arp5 grappler binds to the entry DNA opposing the $\text{Ino80}^{\text{motor}}$ domain. This interaction appears to stabilize the path of entry DNA, allowing a continuous HSA helix to chemomechanically couple the extranucleosomal DNA-bound A-module to the N-lobe of the motor domain. To enable this configuration, DNA needs to be bent at or ahead of $\text{Ino80}^{\text{motor}}$, which provides a possible additional DNA feature sensing mechanism. Placing rigid DNA at this region severely affects sliding efficiency; thus, $\text{Ino80}^{\text{motor}}$ directly, or indirectly through the $\text{Ino80}^{\text{HSA}}$–A-module geometry, could be responsive to DNA mechanical properties. For instance, the extended $\text{Ino80}^{\text{HSA}}$ domain and the A-module could act like a lever arm in this regard. Such a scenario might also help rationalize the peculiar NCP recognition mode of INO80 complexes as opposed to other remodelers. In INO80, the motor is placed at SHL-6 on the entry DNA, while in others, the motor is placed within the nucleosome at SHL-2. When the motor is positioned at SHL-6, it is able to not only pump DNA into the nucleosome but also monitor DNA features at the same time. In contrast, a motor at SHL-2 might be more blind to shape features because the histones prebend DNA anyway.

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

In summary, we provide a detailed mechanism for extranucleosomal DNA binding by $\text{Ino80}^{\text{HSA}}$ and A-module and reveal how it is chemomechanically coupled to the motor of the C-module. The overall architecture reveals multiple instances of extranucleosomal curved DNA, indicating an integrative monitoring of DNA features [propeller twist (*15*)] as one way to tune INO80 sliding (Fig. 9). Future studies need to address how the INO80 complex interacts with other substrates such as hexasomes and nucleosomal arrays. This will allow us to gain further insights into the conformational spectrum of the complex, the way INO80 has "ruler" functions in the generation of nucleosomal arrays (*16*), and possibly understand the suggested histone exchange activities as well (*10*). To this end, it will be important to visualize the evolutionary highly variable N-terminal modules, which may add not only additional targeting but also negative regulatory activities (*26*). Nevertheless, our analysis provides an important step forward in the mechanistic understanding of these complex and fascinating chromatin-shaping molecular machines.

MATERIALS AND METHODS

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

Subunits of the *Ct*INO80 complex and mutants were cloned and expressed by using the MultiBac technology. The gene coding for $\text{Ino80}^{718-1848}$ ($Ct\text{Ino80}^{\Delta\text{N}}$)–2xFLAG® was cloned in pACEBac1; genes for Rvb1 and Rvb2 were cloned in pIDC; and genes coding for Arp5, Ies6, and Ies2 were cloned in the pIDK vector. Together, they were combined in one bacmid. $\textit{Ino80}^{545-850}$ (INO80 A-module) was also cloned in pACEBac1. $\textit{Ino80}^{1-850}$ (INO80ΔC) and *Ies1* were also cloned in pACEBac1. Genes coding for HMG and Iec3 were cloned in pIDC. Genes coding for ZnF and FHA were cloned in the pIDS vector. Genes coding for Ies4 and Taf14 were cloned in pACEBac1, and genes coding for Arp8, actin, and Arp4 were cloned in a pIDK vector and combined on a separate bacmid. *PirHC* (Geneva Biotech) and *Escherichia coli* XL1-Blue (Stratagene) cells were used for all recombination steps by the addition of the Cre recombinase (NEB). From each bacmid (generated in *E. coli* DH10 MultiBac cells), baculoviruses were generated in *Spodoptera frugiperda* (Sf21) insect cells (Thermo Fisher Scientific, #11497013). Each baculovirus (1:100) was transferred to 1 liter of *Trichoplusia ni* High Five cells (Invitrogen, #B85502), thereby coinfecting the cells. Cells were cultured for 60 hours at 27°C and harvested by centrifugation at 4°C.

Cells were disrupted in lysis buffer [30 mM Hepes (pH 7.8), 400 mM NaCl, 10% glycerol, 0.25 mM dithiothreitol (DTT), pepstatin A (0.28 μg/ml), phenylmethylsulfonyl fluoride (PMSF; 0.17 mg/ml), benzamidine (0.33 mg/ml), and 2 mM MgCl_2] for complex purification and gently sonified for 2 min (duty cycle, 50% and output control, 5). Raw lysate was cleared by centrifugation at 20,500*g* and 4°C for 30 min. Supernatant was incubated with 2 ml of ANTI-FLAG® M2 Affinity Gel (Sigma-Aldrich) for 1 hour and washed with 50 ml of lysis buffer and 75 ml of wash buffer [30 mM Hepes (pH 7.8), 150 mM NaCl, 5% glycerol, and 0.25 mM DTT]. The protein was eluted from the matrix by incubation with 4.5 ml of wash buffer [supplemented with FLAG® peptide (0.2 mg/ml)] in three incubation steps of 20 min each.

The elution fractions were loaded onto a Mono Q 5/50 GL column (GE Healthcare) and eluted by an increasing salt gradient (200 mM NaCl to 1 M NaCl), resulting in highly pure INO80. The Ino80$^{\Delta N}$ and A-module mutants were generated by site-directed mutagenesis and expressed and purified as described above (table S4).

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

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

For purification of the INO80 complexes, cells were resuspended in lysis buffer [50 mM tris (pH 7.9), 500 mM NaCl, 10% glycerol, 0.25 mM DTT, pepstatin A (0.28 μg/ml), PMSF (0.17 mg/ml), and benzamidine (0.33 mg/ml)] and disrupted by sonication (4 × 1 min; duty cycle, 50%; and output control, 5). The lysate was cleared by centrifugation at 20,500*g* and 4°C for 40 min. Supernatant was incubated with 3 ml of ANTI-FLAG® M2 Affinity Gel (Sigma-Aldrich) for 1 hour and washed with 50 ml of wash 1 buffer [25 mM Hepes (pH 8), 500 mM KCl, 10% glycerol, 0.05% IGEPAL CA630, 4 mM $MgCl_2$, and 0.25 mM DTT], 50 ml of wash 2 buffer [25 mM Hepes (pH 8), 200 mM KCl, 10% glycerol, 0.05% IGEPAL CA630, 4 mM $MgCl_2$, and 0.25 mM DTT], and 10 ml of buffer A [25 mM Hepes (pH 8), 150 mM KCl, 2 mM $MgCl_2$, and 0.25 mM DTT]. The protein was eluted from the matrix by incubation with 4.5 ml of buffer A [supplemented with FLAG® peptide (0.2 mg/ml)] in four incubation steps of 15 min each.

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

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

Human A-module (*ACTR8*, *ACTB*, *ACTL6A*, and *YY1*) open reading frames (ORFs) were ordered and optimized for insect cell expression at GeneArt (Thermo Fisher Scientific) and assembled on a single pBIG1ab vector using the biGBac cloning system. The 2xFLAG®-tagged *INO80*^{HSA} (*INO80*$^{267-487}$-2xFLAG®) was cloned separately on a pBIG1a vector. After virus generation in Sf21 cells (*S. frugiperda,* Thermo Fisher Scientific, #11497013), the complexes were recombinantly expressed in High Five insect cells (*T. ni*; Invitrogen, #B85502) by adding the two viruses at 1:150 (volume virus: medium) to 3 liters of insect cell culture. The cells were incubated for 60 hours at 27°C and harvested by centrifugation at 4°C. For lysis, the pellet was resuspended in lysis buffer [20 mM Hepes (pH 8.0), 500 mM NaCl, 0.25 mM DTT, and 1× protease inhibitor (cOmplete, Roche)] and gently sonicated three times for 1.5 min. The lysate was incubated with ANTI-FLAG® M2 affinity gel (Sigma-Aldrich) for 1.5 hours and submitted to a gravity flow column. First, the agarose beads were washed with 10 column volumes (CV) lysis buffer [20 mM Hepes (pH 8.0), 500 mM NaCl, and 0.25 mM DTT] followed by 20 CV wash buffer [20 mM Hepes (pH 8.0), 150 mM NaCl, and 0.25 mM DTT]. Next, the protein complex was eluted three times by incubation with 1 CV wash buffer supplemented with 1xFLAG® peptide (0.4 mg/ml) for 15 min each. The elution fractions were applied onto a Capto HiRes Q 5/50 column (Cytiva), and the protein complex was separated via a salt gradient (100 mM NaCl to 1000 mM NaCl) using buffer A [20 mM Hepes (pH 8.0), 100 mM NaCl, 20 mM $ZnCl_2$, 4 mM $MgCl_2$, and 0.25 mM DTT] and buffer B [20 mM Hepes (pH 8.0), 100 mM NaCl, 20 mM $ZnCl_2$, 4 mM $MgCl_2$, and 0.25 mM DTT]. Protein target peak fractions were concentrated to 2 mg/ml in centrifugal filters (Centricon; 70-kDa cutoff, Millipore) and flash-frozen in liquid nitrogen.

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

Purification of nucleosomes

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

Widom 601 DNA with 80-bp extranucleosomal DNA in the 0N80 orientation for reconstituting nucleosomes was used as a DNA template (table S5). DNA was amplified by polymerase chain reaction (PCR), followed by purification using anion exchange chromatography, and the DNA was concentrated in vacuum after the DNA was dialyzed to H_2O overnight. DNA was mixed at a 1.1-fold excess with the histone octamer at 2 M NaCl. The NaCl concentration was decreased to 50 mM over 16 hours at 4°C. After this, nucleosomes were purified by anion exchange chromatography using a SourceQ 1-ml column, and fractions containing nucleosomes were pooled and dialyzed to 50 mM NaCl, concentrated to 1 mg/ml (Centricon; 10-kDa cutoff, Millipore), and stored at 4°C.

Nucleosome sliding assays

0N80 nucleosomes with 5′-fluorescein–labeled extranucleosomal DNA were used for monitoring the sliding activity of *Ct*INO80$^{\Delta N}$ and mutants. Nucleosome (150 nM) was incubated with 50 nM *Ct*INO80$^{\Delta N}$ in sliding buffer [25 mM Hepes (pH 8), 60 mM KCl, 7% glycerol, bovine serum albumin (BSA; 0.10 mg/ml), 0.25 mM DTT, and 2 mM MgCl$_2$] at 25°C. By the addition of 1 mM ATP, the sliding reaction was started and stopped at several time points (15, 30, 45, 60, 120, 300, 600, and 1200 s) by addition of Lambda DNA (0.2 mg/ml; NEB). Nucleosome species were separated by native polyacrylamide gel electrophoresis (PAGE) on a 3 to 12% acrylamide bis-tris gel (Invitrogen) and visualized using the Typhoon imaging system (GE Healthcare). Experiments were performed in triplicates. For gel band quantification, ImageJ was used and the fraction of remodeled band was plotted against the reaction time in percent. Data describe a saturation curve and was fitted in Prism (GraphPad) using an exponential equation.

NADH-coupled ATPase assay

NADH (reduced form of nicotinamide adenine dinucleotide)–coupled ATPase assays were used to determine the ATPase rate of *Ct*INO80$^{\Delta N}$ and mutants. *Ct*INO80$^{\Delta N}$ (30 nM) was incubated in assay buffer [25 mM Hepes (pH 8), 50 mM KCl, 1 mM DTT, 2 mM MgCl$_2$, and BSA (0.1 mg/ml)] with 0.5 mM phosphoenolpyruvate, 1 mM ATP, 0.1 mM NADH, and lactate dehydrogenase (25 U/ml) and pyruvate kinase (Sigma-Aldrich) at 25°C in a final volume of 50 μl. Decreasing NADH concentration was monitored fluorescently over 1 hour in nonbinding, black, 384-well plates (Greiner Bio-One) using 340 nm for excitation and an emission of 460 nm with Tecan Infinite M100 (Tecan). Where indicated, ATPase rate was determined in the presence of 200 nM nucleosome. Experiments were performed in triplicates. ATP turnover was calculated using maximal initial linear rates, corrected for a buffer blank.

Affinity measurement by fluorescence anisotropy

Increasing protein concentrations of the *Ct*INO80 A-module and mutants (final concentrations for *C. thermophilum* A-module: 0, 3.125, 6.25, 12.5 25, 50, 100, and 200 nM) were prepared in assay buffer [25 mM Hepes (pH 8), 100 mM KCl, 2 mM MgCl$_2$, 2% glycerol, 0.01% Triton X-100, and 1 mM DTT] and mixed with 50-bp 6-FAM–labeled DNA (table S5) in assay buffer (final concentration, 5 nM) in a 1:1 (v/v) ratio (final volume: 20 μl; Greiner Flat Bottom Black 384-well plate). The reaction was incubated for 30 min at room temperature, and the fluorescence anisotropy was subsequently measured at an excitation wavelength of 470 nm and an emission wavelength of 520 nm using a TECAN Infinite M1000 plate reader. Experiments were performed in triplicates. The background signal (no protein sample) was subtracted from each value of a dilution series, and the datasets were analyzed with Prism (GraphPad Software). Data were analyzed and fitted to a nonlinear, noncooperative 1:1 binding model [$y = A_f - (A_f - A_b) \times (x/(K_d + x))$, where y is the anisotropy, A_f is the anisotropy of free ligand, A_b is the anisotropy of bound ligand, K_d is the dissociation constant, and x is the receptor concentration] to calculate the apparent dissociation constants.

Electrophoretic mobility shift assay

Electrophoretic mobility shift assays were used to monitor the interaction between INO80 and 0N80 nucleosomes. Nucleosomes were labeled at the 5′ end of their extranucleosomal DNA with fluorescein. Nucleosome (40 nM) was incubated with 80 nM INO80 in electrophoretic mobility shift assay buffer [25 mM Hepes (pH 8), 60 mM KCl, 7% glycerol, 0.25 mM DTT, and 2 mM CaCl$_2$] for 30 min on ice. Samples were analyzed at 4°C by native PAGE on a 3 to 12% acrylamide bis-tris gel (Invitrogen) and visualized using the Typhoon imaging system (GE Healthcare).

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

*Ct*INO80$^{\Delta N}$ and 0N80 nucleosomes were mixed in a ratio of 2:1 and dialyzed to binding buffer [20 mM Hepes (pH 8), 60 mM KCl, 0.25 mM CaCl$_2$, 20 μM ZnCl$_2$, and 0.25 mM DTT] for 1 hour in Slide-A-Lyzer dialysis tubes (Thermo Fisher Scientific). The complex was vitrified at a concentration of 1 mg/ml on Quantifoil R2/1 grids in the presence of 0.05% octyl-β-glucoside using a Leica EM GP (Leica). The same was done for INO80ΔC and INO80 A-module with or without DNA and nucleotides. *Ct*INO80A-module with ATPγS bound was purified further and mildly cross-linked by GraFix using an SW40-Ti rotor (Beckman Coulter). The glycerol (10 to 30%) and glutaraldehyde (0 to 0.025%) co-gradient was generated using Gradient Station *ip* 153 (BioComp Instruments). The samples were fractionated and monitored for 280/260-nm absorbance using Triax UV Flow Cell (BioComp Instruments). The fractions were visually inspected and selected by uranyl acetate (2%) negative staining.

S. cerevisiae N–A-module and DNA (58 bp; table S5) were mixed in an equimolar ratio (1.5 μM each) in cryo-EM buffer [20 mM Hepes (pH 8), 60 mM KCl, 2 mM MgCl$_2$, and 1 mM DTT]. The respective nucleotide was added (final concentration: 1 mM), and the sample was incubated on ice for 10 min. Octyl-β-glucoside was added (0.045%), and 4.5 μl was applied onto a glow discharged Quantifoil R2/1 Cu200 grid. The sample was vitrified in liquid ethane using an EM GP plunge freezer (Leica; 10°C and 90% humidity).

Data collection

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

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

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

For the *S. cerevisiae* A-module bound to ATP (fig. S10, A and C), the contrast transfer function (CTF) parameters of the dataset (4543 micrographs) were determined using patch CTF estimation (multi). The exact processing scheme and data collection and refinement statistics are summarized in table S1. Initial particle picking was done on 2048 micrographs using Blob picker. The particles were subjected to 2D classification and ab initio reconstruction, and classes with clearly defined features were selected. The selected particles were used as input for a Topaz train job on 4543 micrographs. After three rounds of Topaz, 1,028,485 particles were extracted with a box size of 256 pixels and a pixel size of 1.059 Å. The particles were subjected to multiple rounds of 2D classification and heterogeneous refinement. The class that showed the most defined features was selected (327,293 particles) and used for further refinement. The final resolution of the ATP-bound A-module reconstruction after nonuniform refinement was 3.3 Å. To identify a subset of DNA-bound particles, particles were reextracted in Relion (fig. S10B) and subjected to three rounds of 3D classification. The A-module bound to ATP and DNA was reconstructed from 69,226 particles, and the final resolution after 3D refinement was 7.5 Å.

For the *S. cerevisiae* A-module bound to ADP (fig. S11, A and B), the CTF parameters of the dataset (5550 micrographs) were determined using CTFFIND4.1. All subsequent cryo-EM data processing steps were carried out using Relion-3.0 (*58*). Data collection and refinement statistics are summarized in table S1. A total of 2,264,013 particles were picked using Autopicking, and particles were extracted with a box size of 256 pixels and a pixel size of 1.059 Å. 3D classification with five classes was performed using the A-module bound to ATPγS (filtered to 40 Å) as reference. After another round of 3D classification, 970,407 particles were selected and used for further refinement. The final resolution of the ADP-bound A-module reconstruction after postprocessing was 3.2 Å.

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

Beam-induced motions of particles were corrected using MotionCor2 (Relion-3.0) in 5 × 5 patches per frame (*56*, *58*). CTF parameters were estimated from sums of three movie frames using CTFFIND4.1 (*59*). The particles were automatically picked ab initio and qualitatively filtered using WARP (fig. S12, A and B) (*60*). The particles were boxed and extracted from the micrographs in Relion with the particle coordinates exported from WARP using the PyEM scripts developed by D. Asarnow (https://github.com/asarnow/pyem). The initial 3D reconstructions were carried out ab initio using *cis*TEM (fig. S12C). Iterative rounds of 3D classifications were carried out using Relion (*58*). The initial 3D refinements were carried out in Relion-3 using the ab initio 3D reference generated in *cis*TEM (*61*). The final resolution of the ATPγS-bound A-module reconstruction after postprocessing was 3.2 Å. The exact processing scheme is depicted in fig. S12 (A to D). Data collection and refinement statistics are summarized in table S1.

Cryo-EM data processing of *C. thermophilum* A-modules and INO80^{ΔN}

Beam-induced motions of particles were corrected using MotionCor2 (Relion-3.0) in 5 × 5 patches per frame (*56*, *58*). CTF parameters were estimated from sums of three movie frames using CTFFIND4.1. The particles were automatically picked ab initio and qualitatively filtered using WARP (*60*). The particles were boxed and extracted from the micrographs in Relion with the particle coordinates exported from WARP using the PyEM scripts developed by D. Asarnow (https://github.com/asarnow/pyem). The initial 3D reconstructions were carried out ab initio using *cis*TEM. Iterative rounds of 3D classifications were carried out using Relion-3 to remove unbound nucleosomes and separate subtly different C- and A-module conformations. The initial 3D refinements were carried out in Relion-3 using the ab initio 3D reference generated in *cis*TEM (*61*). The exact processing schemes are depicted in figs. S13 to S15. Data collection and refinement statistics are summarized in table S1.

Cryo-EM data processing of *C. thermophilum* INO80 C-module and nucleosome (ADP•BeF$_x$)

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

Initial particle picking was done using Blob picker. Particles were subjected to 2D classification and ab initio reconstruction. Classes with clearly defined features were selected and used as input for a Topaz train job on all micrographs, followed by particle extraction and 2D classification. After three rounds of Topaz, 304,000 particles were extracted with a box size of 360 pixels and a pixel size of 1.059 Å. After selecting 2D classes with clearly defined features, one round of ab initio reconstruction with three classes was performed. Classes with the most defined features were selected and subjected to heterogeneous refinement with two classes. The ab initio reconstructions were used as input volumes for the heterogeneous refinement job. Both classes were selected for further refinement. The final resolution of the reconstruction after nonuniform refinement were 3.5 Å for parallel grappler and 3.8 Å for the cross grappler (fig. S16C).

For a detailed analysis of Ino80^{motor}, beam-induced motions of particles were corrected using MotionCor2 (Relion-3.0) in 5 × 5 patches per frame (*56*, *58*). CTF parameters were estimated from sums of three movie frames using CTFFIND4.1. The exact processing scheme is depicted in fig. S16B. A total of 13,704,000 particles were automatically picked ab initio. In total, 1,242,248 manually picked particles were extracted with a box size of 360 pixels and a pixel size of 1.059 Å. Iterative rounds of 3D classifications and 3D refinement were carried out. After the last round of 3D classification, 137,900 particles were selected and used for further refinement. The final resolution of Ino80^{motor} after postprocessing was 3.6 Å (fig. S16D). Data collection and refinement statistics are summarized in table S1.

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

The movie frames were motion-corrected using MotionCor2 (*56*). All subsequent processing steps were performed in cryoSPARC v3.2.0 (*57*), and the resolutions reported here are calculated on the basis of the gold-standard Fourier shell correlation criterion (FSC = 0.143). The exact processing scheme is depicted in fig.

S17A. Data collection and refinement statistics are summarized in table S1.

Initial particle picking was done using Blob picker. Particles were subjected to 2D classification. Classes with clearly defined features were selected and used as input for a Topaz train job on all micrographs, followed by particle extraction and 2D classification. After three rounds of Topaz, 15,000 particles were extracted with a box size of 256 pixels and a pixel size of 1.059 Å. After selecting 2D classes with clearly defined features, one round of ab initio reconstruction with one class was performed. The final resolution of the reconstruction after nonuniform refinement was 7.5 Å (fig. S17B).

Model building and refinement

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

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

Yeast manipulation and methods

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

For growth assays, cells were grown overnight and adjusted to 0.5 OD_{600} (optical density at 600 nm) units and fivefold serial dilutions were spotted on YPD plates (1% yeast extract, 2% peptone, and 2% glucose), YP + Gal plates (2% galactose), or SD-inositol plates [yeast nitrogen base (6.9 g/liter) without inositol, Formedium CYN37CFG, supplemented with adenine (40 mg/liter), uracil (40 mg/liter), tryptophan (40 mg/liter), histidine (40 mg/liter), leucine (80 mg/liter), and 2% glucose]. Cells were then grown at 30°C for 2 to 5 days unless indicated otherwise. For anaerobic growth conditions, plates were incubated in an anaerobic chamber.

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

Recombination assay

To measure the efficiency of homologous recombination, a quantitative PCR (qPCR)–based gene conversion assay was used as described previously (*39*). Briefly, yeast strains were deficient of endogenous HO endonuclease cleavage sites and engineered with galactose-inducible HO endonuclease, a single HO cut site at ChrIV 491 kb, and a recombination donor sequence at ChrIV 795 kb with a mutated HO cut site and an additional unique 23-bp sequence to allow qPCR analysis. Yeast cells of the indicated genotypes were grown to logarithmic phase in YP + 2% raffinose medium, and HO endonuclease expression was induced by addition of 2% galactose. Aliquots equivalent to one OD_{600} unit were harvested at the indicated time points, and genomic DNA was isolated using the Epicentre MasterPure Yeast DNA Purification Kit (MPY80200). qPCR was performed on a LightCycler 480 instrument (Roche) using LightCycler 480 SYBR Green I Master (Roche 04707516001) with primers designed to detect the completed recombination product (5′-CATACTGTCTCACTCGCTTGGA-3′ and 5′-TTGTTTGCCATTTCGTCAGCTAG-3′). Data were normalized to an unrelated control locus (*MDV1* locus, primers 5′-GCGTGCCTGGTCACAGGTTCATACGAC-3′ and 5′-TCATACGGCCCAAATATTTACGTCCC-3′) and plotted using the GraphPad Prism software as the relative amount of recombination product over time (where 100% recombination = 1). Notably, yeast growth on YP + Gal plates in spot dilution provided a qualitative readout for homologous recombination efficiency as well.

Protein cross-linking

Snap-frozen stock solutions of *H. sapiens* INO80 complex [20 mM Hepes/NaOH (pH 8.0), 200 mM NaCl, 0.5 mM $CaCl_2$, 20 μM $ZnCl_2$, and 0.5 mM DTT] and 0N80 nucleosome [20 mM Hepes/NaOH (pH 7.5), 50 mM NaCl, 0.5 mM DTT, and 10% glycerol] were thawed on ice and mixed in equimolar amounts in reconstitution buffer [20 mM Hepes/NaOH (pH 7.9), 60 mM KCl, 0.5 mM $CaCl_2$, 20 μM $ZnCl_2$, and freshly added 0.5 mM DTT]. The INO80-nucleosome complex mixture was incubated on ice for 30 min to allow for reconstitution. Afterward, 4 μl of freshly prepared BS3 cross-linker stock solution (2 μg/μl in reconstitution buffer; Thermo Fisher Scientific) was added to the reconstituted complex. The complex was cross-linked at 4°C for 2 hours. After that, the reaction was quenched by adding 4 μl of 2 M ammonium bicarbonate, followed by incubation at 4°C for 30 min. Thereafter, half of the cross-linked product was processed by in-gel digestion, and the other half was processed by ethanol precipitation and in-solution digestion.

In-gel digestion

One half of the cross-linked product was mixed with LDS sample buffer, separated in a 4 to 12% NuPAGE bis-tris gel, and stained with Coomassie blue (Colloidal Blue Staining Kit; Thermo Fisher

Scientific). The highlighted area of the gel (fig. S4C) was excised and cut into small gel cubes, followed by destaining in 50% ethanol/50 mM ammonium bicarbonate. The proteins were then reduced in 10 mM DTT at 56°C and alkylated by 50 mM iodoacetamide in the dark at room temperature. Afterward, proteins were digested by trypsin (1 μg per sample) in 50 mM ammonium bicarbonate at 37°C overnight. Following peptide extraction sequentially using extraction buffer (0.1% formic acid in 30% acetonitrile) and 100% acetonitrile, the sample volume was reduced in a centrifugal evaporator to remove residual acetonitrile. The peptides were then acidified with 0.1% formic acid and purified by solid-phase extraction in C18 StageTip (*70*).

Ethanol precipitation and in-solution digestion

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

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

Liquid chromatography tandem mass spectrometry

Cross-linked peptides were analyzed using an Orbitrap Exploris 480 mass spectrometer (Thermo Fisher Scientific) coupled to EASY-nLC 1200 UHPLC system (Thermo Fisher Scientific). Peptides were separated in an in-house packed 55-cm analytical column (inner diameter: 75 μm; ReproSil-Pur 120 C18-AQ 1.9-μm silica particles, Dr. Maisch GmbH) by online reversed-phase chromatography through a 90-min gradient of 2.4 to 33.6% acetonitrile with 0.1% formic acid at a nanoflow rate of 250 nl/min. The eluted peptides were sprayed directly by electrospray ionization into the mass spectrometer. Each sample was injected twice and measured using two different combinations of collision energies in stepped mode (*71*). Mass spectrometry measurement was conducted in data-dependent acquisition mode using a top15 method with one full scan [resolution, 60,000; scan range, 300 to 1650 mass/charge ratio (*m*/*z*); target value, 3×10^6; maximum injection time, 40 ms] followed by 15 fragmentation scans via higher-energy collision dissociation (HCD; normalized collision energy in stepped mode, 25, 30, and 35% or 27, 30, and 33%; resolution, 15,000; target value, 1×10^5; maximum injection time, 40 ms; isolation window, 1.4 *m*/*z*). Only precursor ions of +3 to +8 charge state were selected for fragmentation scans. In addition, precursor ions already isolated for fragmentation were dynamically excluded for 25 s.

Mass spectrometry data analysis

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

Supplementary Materials

This PDF file includes:
Figs. S1 to S17
Tables S1 to S5

Other Supplementary Material for this manuscript includes the following:
Movies S1 to S3
Raw data file

View/request a protocol for this paper from *Bio-protocol*.

REFERENCES AND NOTES

1. G.-C. Yuan, Y.-J. Liu, M. F. Dion, M. D. Slack, L. F. Wu, S. J. Altschuler, O. J. Rando, Genome-scale identification of nucleosome positions in S. cerevisiae. *Science* **309**, 626–630 (2005).
2. W. K. M. Lai, B. F. Pugh, Understanding nucleosome dynamics and their links to gene expression and DNA replication. *Nat. Rev. Mol. Cell Biol.* **18**, 548–562 (2017).
3. C. R. Clapier, J. Iwasa, B. R. Cairns, C. L. Peterson, Mechanisms of action and regulation of ATP-dependent chromatin-remodelling complexes. *Nat. Rev. Mol. Cell Biol.* **18**, 407–422 (2017).
4. I. M. Nodelman, S. das, A. M. Faustino, S. D. Fried, G. D. Bowman, J. P. Armache, Nucleosome recognition and DNA distortion by the Chd1 remodeler in a nucleotide-free state. *Nat. Struct. Mol. Biol.* **29**, 121–129 (2022).
5. A. Jungblut, K. P. Hopfner, S. Eustermann, Megadalton chromatin remodelers: Common principles for versatile functions. *Curr. Opin. Struct. Biol.* **64**, 134–144 (2020).
6. J. Jin, Y. Cai, T. Yao, A. J. Gottschalk, L. Florens, S. K. Swanson, J. L. Gutiérrez, M. K. Coleman, J. L. Workman, A. Mushegian, M. P. Washburn, R. C. Conaway, J. W. Conaway, A mammalian chromatin remodeling complex with similarities to the yeast INO80 complex. *J. Biol. Chem.* **280**, 41207–41212 (2005).
7. X. Shen, G. Mizuguchi, A. Hamiche, C. Wu, A chromatin remodelling complex involved in transcription and DNA processing. *Nature* **406**, 541–544 (2000).
8. K. Yen, V. Vinayachandran, B. F. Pugh, SWR-C and INO80 chromatin remodelers recognize nucleosome-free regions near +1 nucleosomes. *Cell* **154**, 1246–1256 (2013).
9. M. Udugama, A. Sabri, B. Bartholomew, The INO80 ATP-dependent chromatin remodeling complex is a nucleosome spacing factor. *Mol. Cell. Biol.* **31**, 662–673 (2011).
10. S. Brahma, M. I. Udugama, J. Kim, A. Hada, S. K. Bhardwaj, S. G. Hailu, T.-H. Lee, B. Bartholomew, INO80 exchanges H2A.Z for H2A by translocating on DNA proximal to histone dimers. *Nat. Commun.* **8**, 15616 (2017).
11. L. J. Hsieh, M. A. Gourdet, C. M. Moore, E. N. Muñoz, N. Gamarra, V. Ramani, G. J. Narlikar, A hexasome is the preferred substrate for the INO80 chromatin remodeling complex, allowing versatility of function. *Mol. Cell* **82**, 2098–2112.e4 (2022).
12. N. Krietenstein, M. Wal, S. Watanabe, B. Park, C. L. Peterson, B. F. Pugh, P. Korber, Genomic nucleosome organization reconstituted with pure proteins. *Cell* **167**, 709–721.e12 (2016).
13. A. A. Reyes, R. D. Marcum, Y. He, Structure and function of chromatin remodelers. *J. Mol. Biol.* **433**, 166929 (2021).
14. Y. Yin, W. T. C. Lee, D. Gupta, H. Xue, P. Tonzi, J. A. Borowiec, T. T. Huang, M. Modesti, E. Rothenberg, A basal-level activity of ATR links replication fork surveillance and stress response. *Mol. Cell* **81**, 4243–4257.e6 (2021).

15. E. Oberbeckmann, N. Krietenstein, V. Niebauer, Y. Wang, K. Schall, M. Moldt, T. Straub, R. Rohs, K. P. Hopfner, P. Korber, S. Eustermann, Genome information processing by the INO80 chromatin remodeler positions nucleosomes. *Nat. Commun.* **12**, 3231 (2021).
16. E. Oberbeckmann, V. Niebauer, S. Watanabe, L. Farnung, M. Moldt, A. Schmid, P. Cramer, C. L. Peterson, S. Eustermann, K. P. Hopfner, P. Korber, Ruler elements in chromatin remodelers set nucleosome array spacing and phasing. *Nat. Commun.* **12**, 3232 (2021).
17. A. Basu, D. G. Bobrovnikov, Z. Qureshi, T. Kayikcioglu, T. T. M. Ngo, A. Ranjan, S. Eustermann, B. Cieza, M. T. Morgan, M. Hejna, H. T. Rube, K. P. Hopfner, C. Wolberger, J. S. Song, T. Ha, Measuring DNA mechanics on the genome scale. *Nature* **589**, 462–467 (2021).
18. X. Liu, M. Li, X. Xia, X. Li, Z. Chen, Mechanism of chromatin remodelling revealed by the Snf2-nucleosome structure. *Nature* **544**, 440–445 (2017).
19. L. Farnung, S. M. Vos, C. Wigge, P. Cramer, Nucleosome-Chd1 structure and implications for chromatin remodelling. *Nature* **550**, 539–542 (2017).
20. S. Eustermann, K. Schall, D. Kostrewa, K. Lakomek, M. Strauss, M. Moldt, K. P. Hopfner, Structural basis for ATP-dependent chromatin remodelling by the INO80 complex. *Nature* **556**, 386–390 (2018).
21. R. J. Aramayo, O. Willhoft, R. Ayala, R. Bythell-Douglas, D. B. Wigley, X. Zhang, Cryo-EM structures of the human INO80 chromatin-remodeling complex. *Nat. Struct. Mol. Biol.* **25**, 37–44 (2018).
22. R. Ayala, O. Willhoft, R. J. Aramayo, M. Wilkinson, E. A. McCormack, L. Ocloo, D. B. Wigley, X. Zhang, Structure and regulation of the human INO80–Nucleosome complex. *Nature* **556**, 391–395 (2018).
23. A. B. Patel, C. M. Moore, B. J. Greber, J. Luo, S. A. Zukin, J. Ranish, E. Nogales, Architecture of the chromatin remodeler RSC and insights into its nucleosome engagement. *eLife* **8**, e54449 (2019).
24. S. He, Z. Wu, Y. Tian, Z. Yu, J. Yu, X. Wang, J. Li, B. Liu, Y. Xu, Structure of nucleosome-bound human BAF complex. *Science* **367**, 875–881 (2020).
25. R. W. Baker, J. M. Reimer, P. J. Carman, B. Turegun, T. Arakawa, R. Dominguez, A. E. Leschziner, Structural insights into assembly and function of the RSC chromatin remodeling complex. *Nat. Struct. Mol. Biol.* **28**, 71–80 (2021).
26. C. Y. Zhou, S. L. Johnson, L. J. Lee, A. D. Longhurst, S. L. Beckwith, M. J. Johnson, A. J. Morrison, G. J. Narlikar, The yeast INO80 complex operates as a tunable DNA length-sensitive switch to regulate nucleosome sliding. *Mol. Cell* **69**, 677–688.e9 (2018).
27. K. R. Knoll, S. Eustermann, V. Niebauer, E. Oberbeckmann, G. Stoehr, K. Schall, A. Tosi, M. Schwarz, A. Buchfellner, P. Korber, K. P. Hopfner, The nuclear actin-containing Arp8 module is a linker DNA sensor driving INO80 chromatin remodeling. *Nat. Struct. Mol. Biol.* **25**, 823–832 (2018).
28. S. Brahma, M. Ngubo, S. Paul, M. Udugama, B. Bartholomew, The Arp8 and Arp4 module acts as a DNA sensor controlling INO80 chromatin remodeling. *Nat. Commun.* **9**, 3309 (2018).
29. J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Žídek, A. Potapenko, A. Bridgland, C. Meyer, S. A. A. Kohl, A. J. Ballard, A. Cowie, B. Romera-Paredes, S. Nikolov, R. Jain, J. Adler, T. Back, S. Petersen, D. Reiman, E. Clancy, M. Zielinski, M. Steinegger, M. Pacholska, T. Berghammer, D. Silver, O. Vinyals, A. W. Senior, K. Kavukcuoglu, P. Kohli, D. Hassabis, Applying and improving AlphaFold at CASP14. *Proteins* **89**, 1711–1721 (2021).
30. P. Hanč, T. Fujii, S. Iborra, Y. Yamada, J. Huotari, O. Schulz, S. Ahrens, S. Kjær, M. Way, D. Sancho, K. Namba, C. R. e Sousa, Structure of the complex of F-actin and DNGR-1, a C-type lectin receptor involved in dendritic cell cross-presentation of dead cell-associated antigens. *Immunity* **42**, 839–849 (2015).
31. X. Huang, F. Poy, R. Zhang, A. Joachimiak, M. Sudol, M. J. Eck, Structure of a WW domain containing fragment of dystrophin in complex with beta-dystroglycan. *Nat. Struct. Biol.* **7**, 634–638 (2000).
32. M. A. Verdecia, M. E. Bowman, K. P. Lu, T. Hunter, J. P. Noel, Structural basis for phosphoserine-proline recognition by group IV WW domains. *Nat. Struct. Biol.* **7**, 639–643 (2000).
33. Y. Cai, J. Jin, T. Yao, A. J. Gottschalk, S. K. Swanson, S. Wu, Y. Shi, M. P. Washburn, L. Florens, R. C. Conaway, J. W. Conaway, YY1 functions with INO80 to activate transcription. *Nat. Struct. Mol. Biol.* **14**, 872–874 (2007).
34. C. Alfieri, M. C. Gambetta, R. Matos, S. Glatt, P. Sehr, S. Fraterman, M. Wilm, J. Müller, C. W. Müller, Structural basis for targeting the chromatin repressor Sfmbt to Polycomb response elements. *Genes Dev.* **27**, 2367–2379 (2013).
35. F. H. Wilkinson, K. Park, M. L. Atchison, Polycomb recruitment to DNA in vivo by the YY1 REPO domain. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 19296–19301 (2006).
36. C. Kadoch, D. C. Hargreaves, C. Hodges, L. Elias, L. Ho, J. Ranish, G. R. Crabtree, Proteomic and bioinformatic analysis of mammalian SWI/SNF complexes identifies extensive roles in human malignancy. *Nat. Genet.* **45**, 592–601 (2013).
37. J. Poli, S. M. Gasser, M. Papamichos-Chronakis, The INO80 remodeller in transcription, replication and repair. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **372**, 20160290 (2017).
38. W. Yao, D. A. King, S. L. Beckwith, G. J. Gowans, K. Yen, C. Zhou, A. J. Morrison, The INO80 complex requires the Arp5-Ies6 subcomplex for chromatin remodeling and metabolic regulation. *Mol. Cell. Biol.* **36**, 979–991 (2016).
39. C. A. Lademann, J. Renkawitz, B. Pfander, S. Jentsch, The INO80 complex removes H2A.Z to promote presynaptic filament formation during homologous recombination. *Cell Rep.* **19**, 1294–1303 (2017).
40. F. Civril, T. Deimling, C. C. de Oliveira Mann, A. Ablasser, M. Moldt, G. Witte, V. Hornung, K. P. Hopfner, Structural mechanism of cytosolic DNA sensing by cGAS. *Nature* **498**, 332–337 (2013).
41. L. Yan, H. Wu, X. Li, N. Gao, Z. Chen, Publisher correction: Structures of the ISWI-nucleosome complex reveal a conserved mechanism of chromatin remodeling. *Nat. Struct. Mol. Biol.* **26**, 389 (2019).
42. M. Li, X. Xia, Y. Tian, Q. Jia, X. Liu, Y. Lu, M. Li, X. Li, Z. Chen, Mechanism of DNA translocation underlying chromatin remodelling by Snf2. *Nature* **567**, 409–413 (2019).
43. C. R. Clapier, N. Verma, T. J. Parnell, B. R. Cairns, Cancer-associated gain-of-function mutations activate a SWI/SNF-family regulatory hub. *Mol. Cell* **80**, 712–725.e5 (2020).
44. H. L. Schubert, J. Wittmeyer, M. M. Kasten, K. Hinata, D. C. Rawling, A. Héroux, B. R. Cairns, C. P. Hill, Structure of an actin-related subcomplex of the SWI/SNF chromatin remodeler. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 3345–3350 (2013).
45. M. Sudol, Structure and function of the WW domain. *Prog. Biophys. Mol. Biol.* **65**, 113–132 (1996).
46. Y. A. Chen, C. Y. Lu, T. Y. Cheng, S. H. Pan, H. F. Chen, N. S. Chang, WW domain-containing proteins YAP and TAZ in the hippo pathway as key regulators in stemness maintenance, tissue homeostasis, and tumorigenesis. *Front. Oncol.* **9**, 60 (2019).
47. A. J. Morrison, J. A. Kim, M. D. Person, J. Highland, J. Xiao, T. S. Wehr, S. Hensley, Y. Bao, J. Shen, S. R. Collins, J. S. Weissman, J. Delrow, N. J. Krogan, J. E. Haber, X. Shen, Mec1/Tel1 phosphorylation of the INO80 chromatin remodeling complex influences DNA damage checkpoint responses. *Cell* **130**, 499–511 (2007).
48. P. Kapoor, Y. Bao, J. Xiao, A. Espejo, L. Yang, M. T. Bedford, G. Peng, X. Shen, Phosphorylation-dependent enhancement of Rad53 kinase Activity through the INO80 chromatin remodeling complex. *Mol. Cell* **58**, 863–869 (2015).
49. P. P. Singh, M. Shukla, S. A. White, M. Lafos, P. Tong, T. Auchynnikava, C. Spanos, J. Rappsilber, A. L. Pidoux, R. C. Allshire, Hap2-Ino80-facilitated transcription promotes de novo establishment of CENP-A chromatin. *Genes Dev.* **34**, 226–238 (2020).
50. A. G. Chapman, A. M. Cotton, A. D. Kelsey, C. J. Brown, Differentially methylated CpG island within human XIST mediates alternative P2 transcription and YY1 binding. *BMC Genet.* **15**, 89 (2014).
51. J. Kim, A. Kollhoff, A. Bergmann, L. Stubbs, Methylation-sensitive binding of transcription factor YY1 to an insulator sequence within the paternally expressed imprinted gene, Peg3. *Hum. Mol. Genet.* **12**, 233–245 (2003).
52. J. Wang, X. Wu, C. Wei, X. Huang, Q. Ma, X. Huang, F. Faiola, D. Guallar, M. Fidalgo, T. Huang, D. Peng, L. Chen, H. Yu, X. Li, J. Sun, X. Liu, X. Cai, X. Chen, L. Wang, J. Ren, J. Wang, J. Ding, YY1 positively regulates transcription by targeting promoters and super-enhancers through the BAF complex in embryonic stem cells. *Stem Cell Rep.* **10**, 1324–1339 (2018).
53. P. Elizarev, K. Finkl, J. Müller, Distinct requirements for Pho, Sfmbt, and Ino80 for cell survival in Drosophila. *Genetics* **219**, iyab096 (2021).
54. B. Turegun, R. W. Baker, A. E. Leschziner, R. Dominguez, Actin-related proteins regulate the RSC chromatin remodeler by weakening intramolecular interactions of the Sth1 ATPase. *Commun. Biol.* **1**, 1 (2018).
55. C. R. Clapier, M. M. Kasten, T. J. Parnell, R. Viswanathan, H. Szerlong, G. Sirinakis, Y. Zhang, B. R. Cairns, Regulation of DNA translocation efficiency within the chromatin remodeler RSC/Sth1 potentiates nucleosome sliding and ejection. *Mol. Cell* **62**, 453–461 (2016).
56. S. Q. Zheng, E. Palovcak, J. P. Armache, K. A. Verba, Y. Cheng, D. A. Agard, MotionCor2: Anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods* **14**, 331–332 (2017).
57. A. Punjani, J. L. Rubinstein, D. J. Fleet, M. A. Brubaker, cryoSPARC: Algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods* **14**, 290–296 (2017).
58. J. Zivanov, T. Nakane, B. O. Forsberg, D. Kimanius, W. J. H. Hagen, E. Lindahl, S. H. W. Scheres, New tools for automated high-resolution cryo-EM structure determination in RELION-3. *eLife* **7**, e42166 (2018).
59. A. Rohou, N. Grigorieff, CTFFIND4: Fast and accurate defocus estimation from electron micrographs. *J. Struct. Biol.* **192**, 216–221 (2015).
60. D. Tegunov, P. Cramer, Real-time cryo-electron microscopy data preprocessing with Warp. *Nat. Methods* **16**, 1146–1152 (2019).
61. T. Grant, A. Rohou, N. Grigorieff, cisTEM, user-friendly software for single-particle image processing. *eLife* **7**, e35383 (2018).
62. E. F. Pettersen, T. D. Goddard, C. C. Huang, E. C. Meng, G. S. Couch, T. I. Croll, J. H. Morris, T. E. Ferrin, UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Sci.* **30**, 70–82 (2021).

63. P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501 (2010).
64. K. Yamashita, C. M. Palmer, T. Burnley, G. N. Murshudov, Cryo-EM single-particle structure refinement and map calculation using Servalcat. *Acta Crystallogr. D Struct. Biol.* **77**, 1282–1291 (2021).
65. T. I. Croll, ISOLDE: A physically realistic environment for model building into low-resolution electron-density maps. *Acta Crystallogr. D Struct. Biol.* **74**, 519–530 (2018).
66. T. Burnley, C. M. Palmer, M. Winn, Recent developments in the CCP-EM software suite. *Acta Crystallogr. D Struct. Biol.* **73**, 469–477 (2017).
67. P. D. Adams, P. V. Afonine, G. Bunkóczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L. W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, A. J. McCoy, N. W. Moriarty, R. Oeffner, R. J. Read, D. C. Richardson, J. S. Richardson, T. C. Terwilliger, P. H. Zwart, PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 213–221 (2010).
68. M. Knop, K. Siegers, G. Pereira, W. Zachariae, B. Winsor, K. Nasmyth, E. Schiebel, Epitope tagging of yeast genes using a PCR-based strategy: More tags and improved practical routines. *Yeast* **15**, 963–972 (1999).
69. C. Janke, M. M. Magiera, N. Rathfelder, C. Taxis, S. Reber, H. Maekawa, A. Moreno-Borchart, G. Doenges, E. Schwob, E. Schiebel, M. Knop, A versatile toolbox for PCR-based tagging of yeast genes: New fluorescent proteins, more markers and promoter substitution cassettes. *Yeast* **21**, 947–962 (2004).
70. J. Rappsilber, Y. Ishihama, M. Mann, Stop and go extraction tips for matrix-assisted laser desorption/ionization, nanoelectrospray, and LC/MS sample pretreatment in proteomics. *Anal. Chem.* **75**, 663–670 (2003).
71. C. E. Stieger, P. Doppler, K. Mechtler, Optimized fragmentation improves the identification of peptides cross-linked by MS-cleavable reagents. *J. Proteome Res.* **18**, 1363–1370 (2019).
72. J. Cox, M. Mann, MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* **26**, 1367–1372 (2008).
73. S. Lenz, S. H. Giese, L. Fischer, J. Rappsilber, In-search assignment of monoisotopic peaks improves the identification of cross-linked peptides. *J. Proteome Res.* **17**, 3923–3931 (2018).
74. M. L. Mendes, L. Fischer, Z. A. Chen, M. Barbon, F. J. O'Reilly, S. H. Giese, M. Bohlke-Schneider, A. Belsom, T. Dau, C. W. Combe, M. Graham, M. R. Eisele, W. Baumeister, C. Speck, J. Rappsilber, An integrated workflow for crosslinking mass spectrometry. *Mol. Syst. Biol.* **15**, e8994 (2019).
75. C. Notredame, D. G. Higgins, J. Heringa, T-coffee: A novel method for fast and accurate multiple sequence alignment. *J. Mol. Biol.* **302**, 205–217 (2000).
76. E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, T. E. Ferrin, UCSF Chimera—A visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
77. S. Li, W. K. Olson, X. J. Lu, Web 3DNA 2.0 for the analysis, visualization, and modeling of 3D nucleic acid structures. *Nucleic Acids Res.* **47**, W26–W34 (2019).

Acknowledgments: We thank K. Lammens and J. Bartho for help with data collection and V. Doebler for discussion and biochemical support in nucleosome purification. We also thank T. Fröhlich from the LMU functional genome analysis. **Funding:** This work is supported by the Deutsche Forschungsgemeinschaft DFG (CRC1064 to K.-P.H. and B.P., CRC1361 to K.-P.H. and J.-X.C., PFA794-5/1 to B.P., and RTG1721 and the Gottfried Wilhelm Leibniz Prize to K.-P.H.), the European Research Council (ERC Advanced Grant 833613 INO3D to K.-P.H.), and the Max Planck Society (to B.P.). S.E. acknowledges an EMBO long-term fellowship, and K.S. acknowledges funding by Quantitative Biosciences Munich (QBM). **Author contributions:** Conceptualization: K.-P.H. and S.E. Construct design: S.E., K.S., F.K., F.J.M., and J.J. Protein purification: F.K., F.J.M., M.M., S.W., K.S., and J.J. Cryo-EM sample preparation: F.K., F.J.M., S.W., J.J., S.E., and K.S. Processing: J.J., K.-P.H., F.K., F.J.M., S.W., S.E., and K.S. Model building: D.K., J.J., K.-P.H., S.E., and K.S. In vitro biochemistry: F.K., F.J.M., and J.-X.C. In vivo experiments: M.H., S.B., and B.P. Visualization: F.K., F.J.M., S.W., and M.H. Supervision: K.-P.H. and B.P. Writing—original draft: K.-P.H., B.P., F.K., F.J.M., S.W., M.H., J.J., and D.K. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** The coordinates of the *S. cerevisiae* A-modules (ATP and ATPγS state) have been deposited in the PDB under the accession codes 8A5A and 8A5O, respectively. The *S. cerevisiae* A-module (ATP, ATPγS state, and ADP state) and the DNA-bound A-module cryo-EM reconstructions are available at the EMDB under the EMDB accession codes EMD-15163, EMD-15179, EMD-15177, and EMD-15186. *The C. thermophilum* A-module (ATPγS state), the A-module bound to curved DNA (*apo*) and straight DNA (ADP•AlF$_x$), the INO80 C-module bound to nucleosome (cross grappler, *apo*), and the nucleosome-bound Ino80^{motor} (ADP•BeF$_x$) coordinates have been deposited in the PDB under the accession codes 8A5D, 8A5P, 8A5Q, 8AV6, and 8ATF, respectively. *The C. thermophilum* A-module (ATPγS state) and the A-module bound to curved DNA (*apo*) and straight DNA (ADP•AlF$_x$), the INO80 C-module bound to nucleosome (cross grappler, ADP•BeF$_x$), and the nucleosome-bound Ino80^{motor} (ADP•BeF$_x$) cryo-EM reconstructions are available at the EMDB under the EMDB accession codes EMD-15165, EMD-15180, EMD-15184, EMD-15688, and EMD-15647, respectively. The *C. thermophilum* A-module (bound to curved DNA in the ADP•AlF$_x$ state) cryo-EM reconstruction and the INO80 C-module bound to nucleosome (parallel grappler, ADP•BeF$_x$) cryo-EM reconstruction have been deposited at the EMDB under the EMDB accession codes EMD-15187 and EMD-15188, respectively. *C. thermophilum* INO80$^{\Delta N}$ bound to a nucleosome cryo-EM reconstruction has been deposited at the EMDB under the accession code EMD-15211. The *H. sapiens* A-module cryo-EM reconstruction has been deposited at the EMDB under the accession code EMD-15164. All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials.

Submitted 8 June 2022
Accepted 27 October 2022
Published 9 December 2022
10.1126/sciadv.add3189

Science Advances

Structural mechanism of extranucleosomal DNA readout by the INO80 complex

Franziska Kunert, Felix J. Metzner, James Jung, Markus Hpfler, Stephan Woike, Kevin Schall, Dirk Kostrewa, Manuela Moldt, Jia-Xuan Chen, Susanne Bantele, Boris Pfander, Sebastian Eustermann, and Karl-Peter Hopfner

Sci. Adv., **8** (49), eadd3189.
DOI: 10.1126/sciadv.add3189

Science Advances (ISSN) is published by the American Association for the Advancement of Science. 1200 New York Avenue NW, Washington, DC 20005. The title *Science Advances* is a registered trademark of AAAS.

Science Advances
AAAS

Supplementary Materials for

Structural mechanism of extranucleosomal DNA readout by the INO80 complex

Franziska Kunert *et al.*

Corresponding author: Karl-Peter Hopfner, hopfner@genzentrum.lmu.de

Sci. Adv. **8**, eadd3189 (2022)
DOI: 10.1126/sciadv.add3189

The PDF file includes:

Figs. S1 to S17
Tables S1 to S5
Legends for movies S1 to S3
Legend for raw data file

Other Supplementary Material for this manuscript includes the following:

Movies S1 to S3
Raw data file

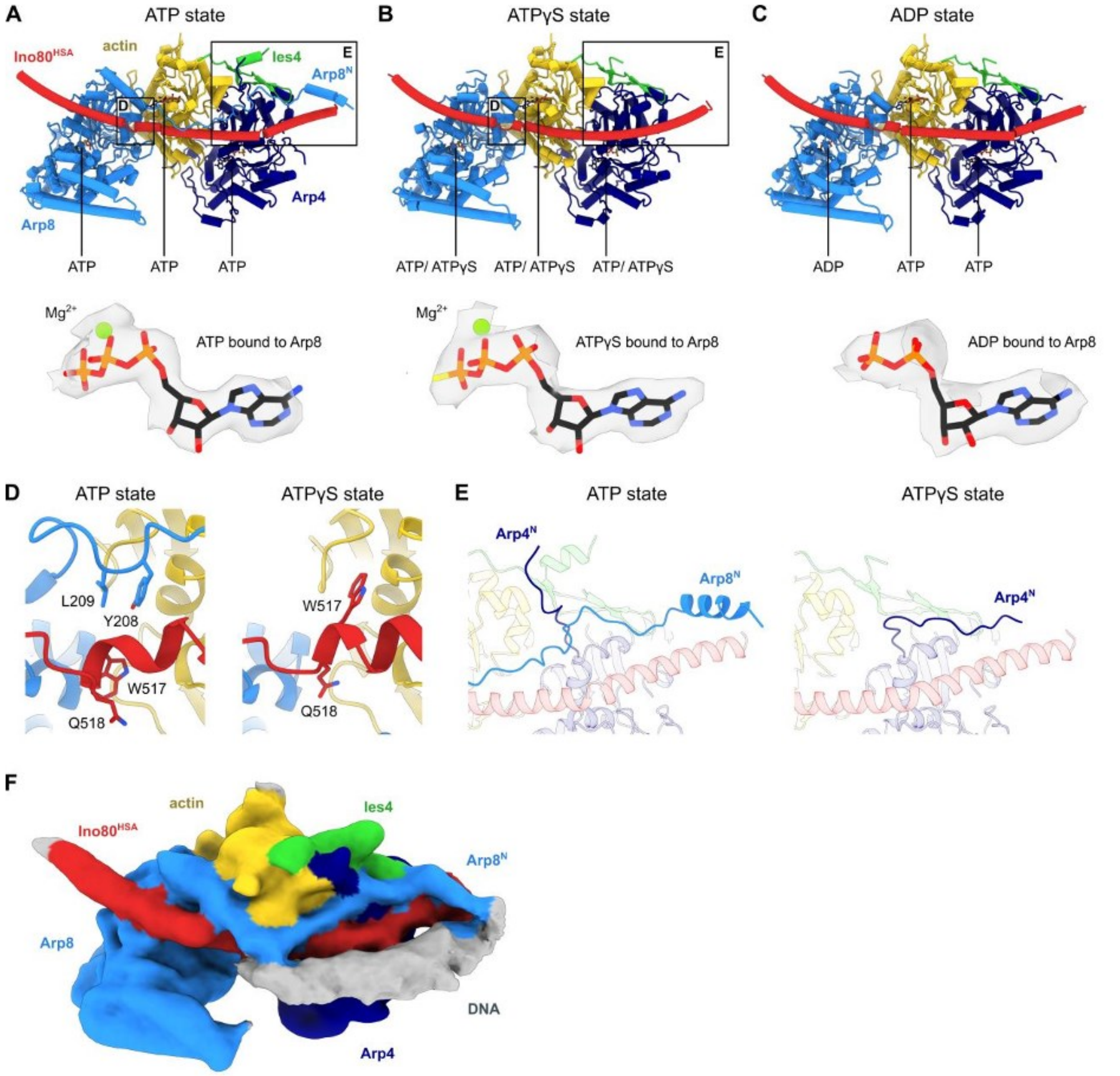


Fig. S1. The Arp8 N-terminus of the *S. cerevisiae* A-module binds DNA. (A-C) Structural model of the *S. cerevisiae (Sc)* A-module in (**A**) ATP state, (**B**) ATPγS state and (**C**) ADP state (ADP state is not a fully refined model, ATPγS state docked into ADP state reconstruction and ADP modelled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

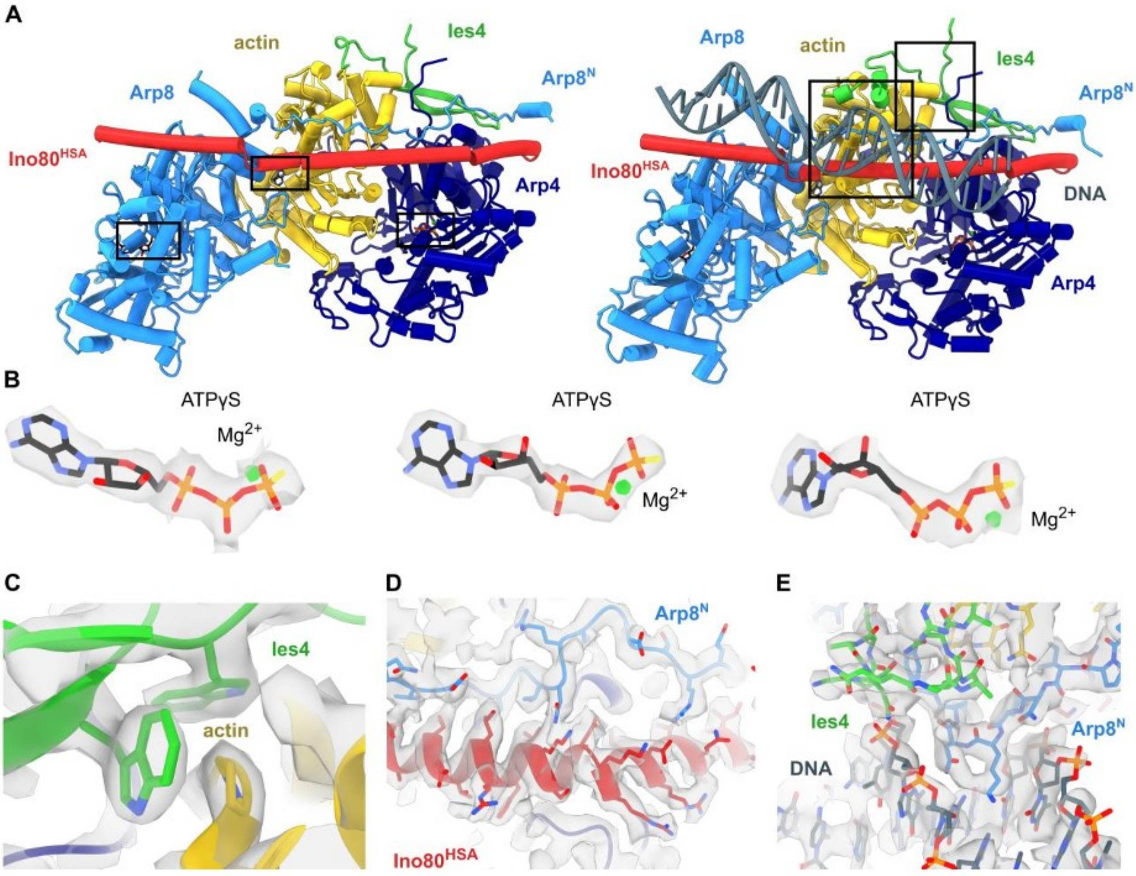


Fig S2: Examples of cryo-EM density map of the *C. thermophilum* A-module. (**A**) Structural model of *C. thermophilum (Ct)* A-module (left) without DNA or (right) with DNA bound. The protein subunits are color coded and annotated. (**B**) Cryo-EM density maps of nucleotides (ATPγS) bound to (left to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

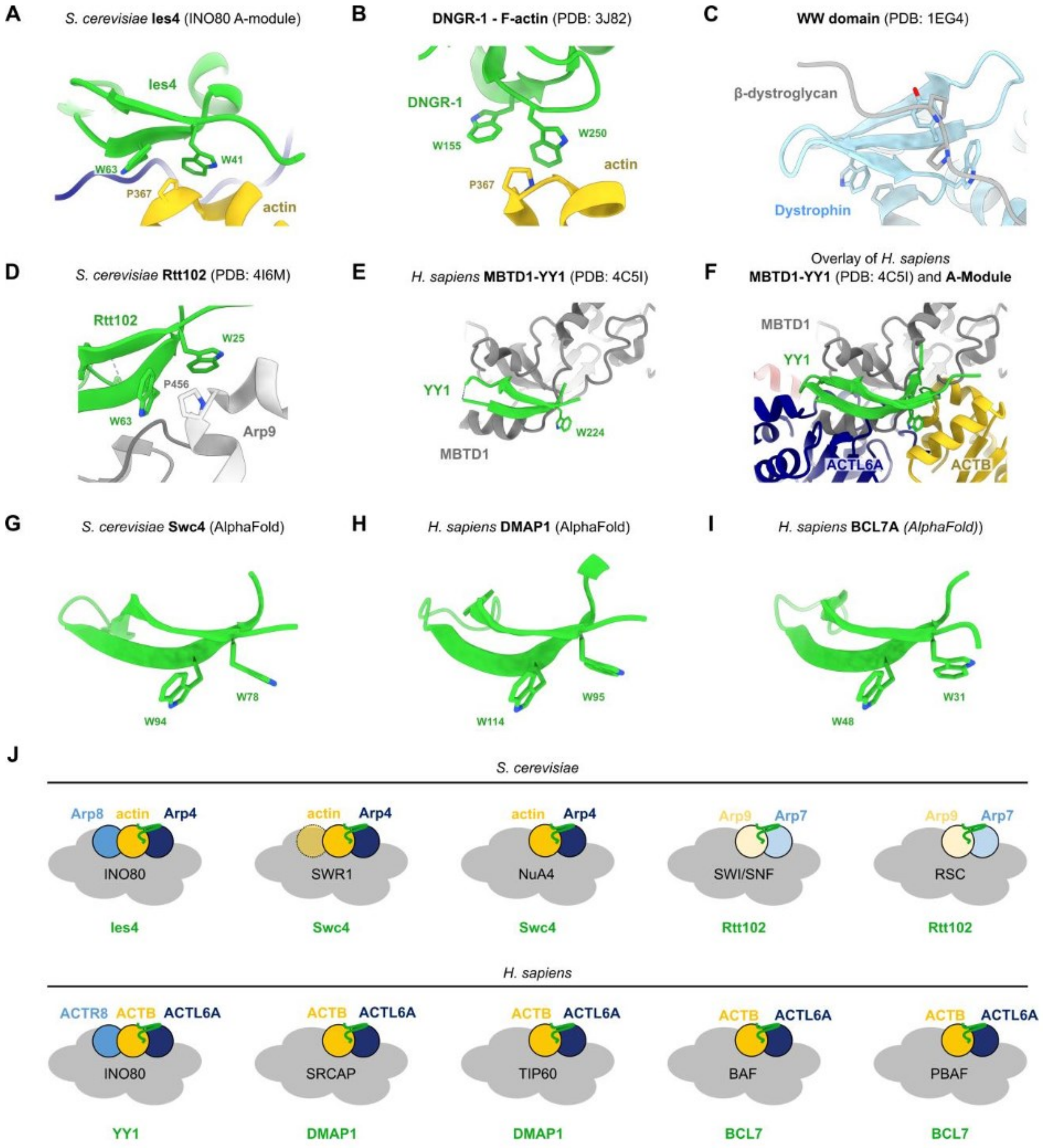


Fig. S3. The 2W-hairpin shows a conserved interaction mode with actin/ARPs. (**A**) Detailed view of the Ies4-actin interface in the *S. cerevisiae* A-module. The conserved tryptophan and proline residues are shown. (**B**) Detailed view of the DNGR-1-actin interface (PDB: 3J82). The conserved tryptophan and proline residues are shown. (**C**) Structure of the dysotrophin WW-domain in complex with a β-dystroglycan peptide (PDB: 1EG4). Conserved tryptophan and proline residues are shown. (**D**) Detailed view of the Rtt102-Arp9 interface in *S. cerevisiae* SWI/SNF A-module (PDB: 4I6M). The conserved tryptophan and proline residues are shown.

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

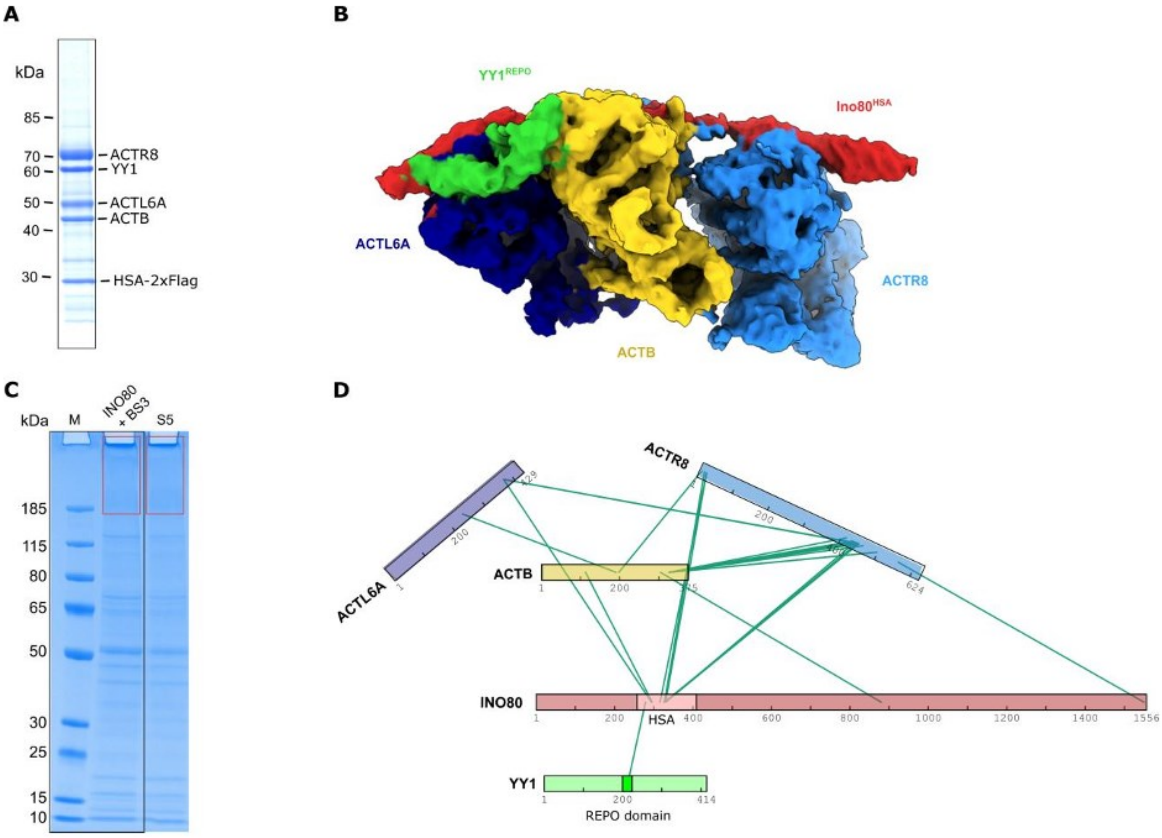


Fig. S4. YY1 interacts with the *H. sapiens* Ino80$^{\mathbf{HSA}}$. (**A**) Coomassie-stained SDS-PAGE gel of the purified *H. sapiens (Hs)* A-module (**B**) Cryo-EM reconstruction of *Hs*A-module. The density is colored according to the underlying protein subunits. (**C**) Coomassie-stained SDS-PAGE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

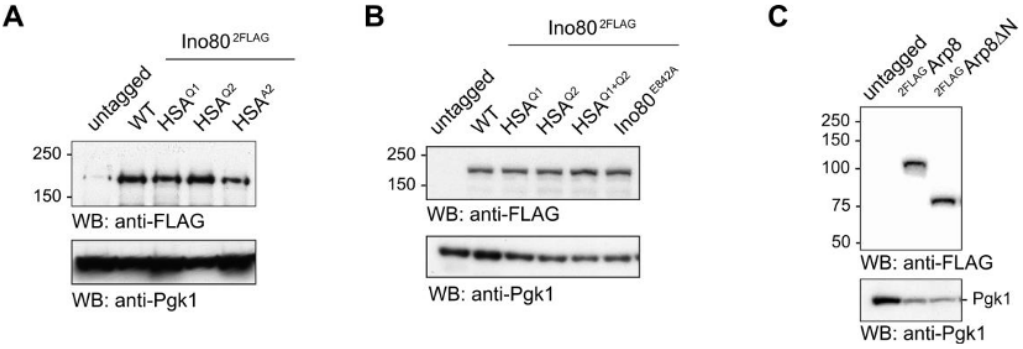
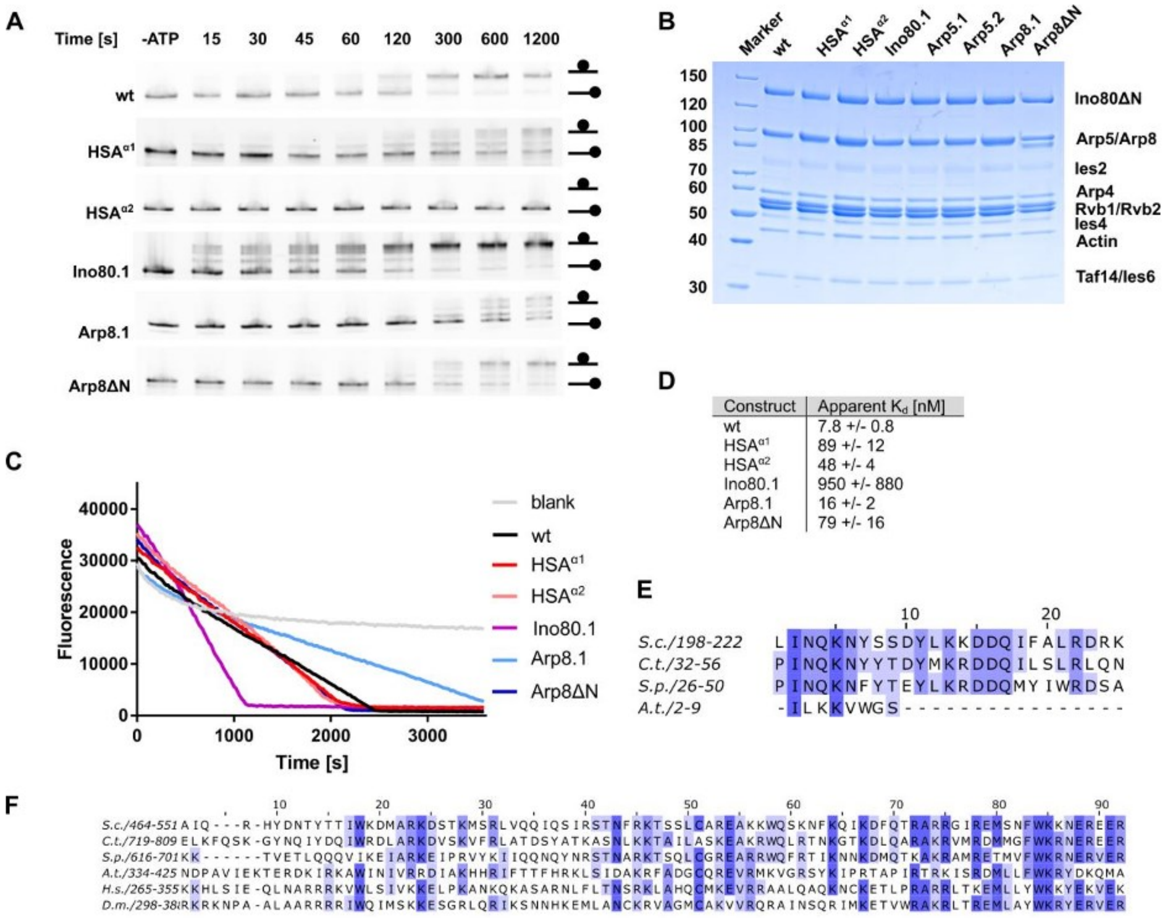


Fig. S5. HSA surface residues are critical for INO80 function in budding yeast. (**A**) Ino80 variants are expressed at levels similar to the Ino80 WT protein. Yeast cells as in Fig. 2A expressing the indicated 2FLAG-tagged Ino80 variants under the control of the endogenous *INO80* promoter were subjected to total protein analysis by western blotting using an anti-FLAG antibody. Pgk1 levels served as control. (**B**) Ino80 variants are expressed at levels similar to the Ino80 WT protein. Diploid yeast cells used for tetrad analysis in Fig. 2B expressing the indicated 2FLAG-tagged Ino80 variants under the control of the endogenous *INO80* promoter were subjected to total protein analysis by western blotting using an anti-FLAG antibody. Pgk1 levels served as control. (**C**) The Arp8ΔN variant is expressed similar to Arp8 WT levels. Yeast cells as in Fig. 2C and D expressing the indicated 2FLAG-tagged Arp8 variants under the control of the endogenous *ARP8* promoter were subjected to total protein analysis by western blotting using an anti-FLAG antibody. Pgk1 levels served as control.



Construct	Apparent K_d [nM]
wt	7.8 +/- 0.8
HSA$^{\alpha1}$	89 +/- 12
HSA$^{\alpha2}$	48 +/- 4
Ino80.1	950 +/- 880
Arp8.1	16 +/- 2
Arp8ΔN	79 +/- 16

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

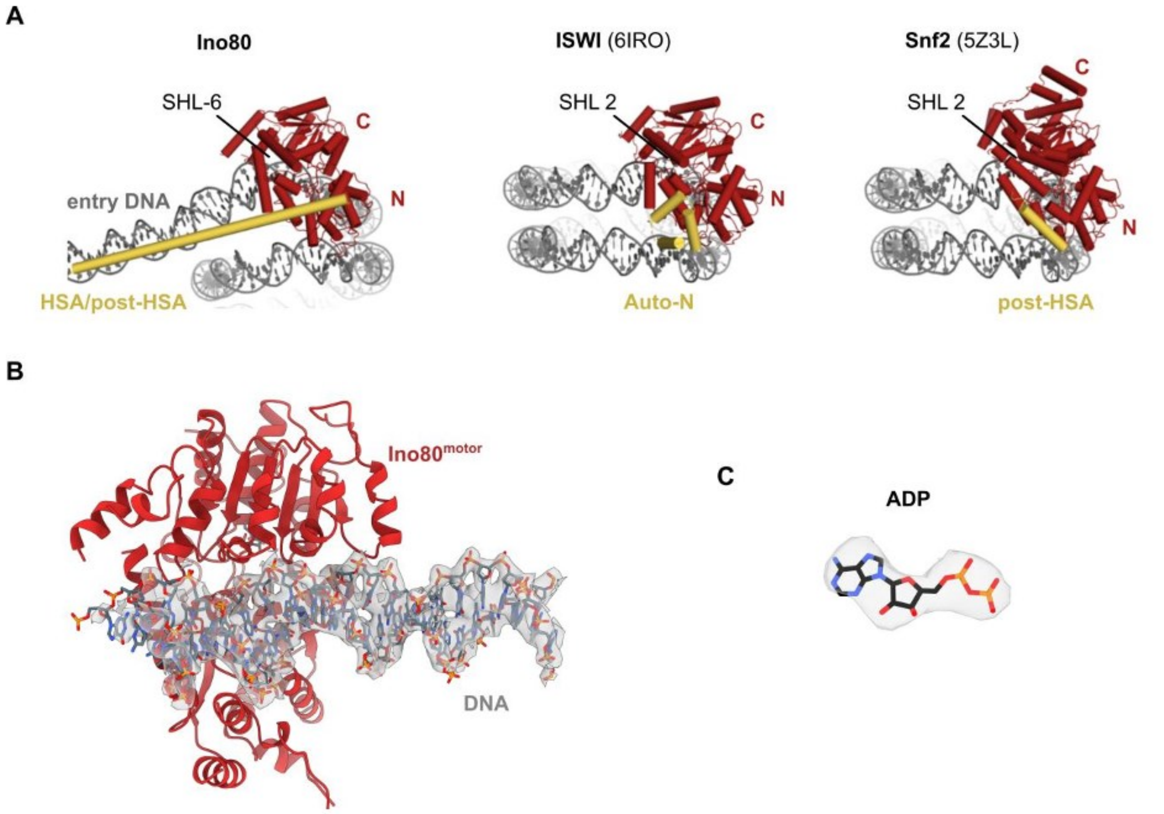


Fig. S7. Example of cryo-EM density map of the Ino80^{motor} (ADP·BeF$_x$ state). (**A**) Comparison of Ino80, ISWI and Snf2 interacting with nucleosomes and similarity of post-HSA and Auto-N. Ino80 binds the nucleosome at SHL-6 while Isw1 and Snf2 bind at SHL-2. (**B**) Detailed view of the cryo-EM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

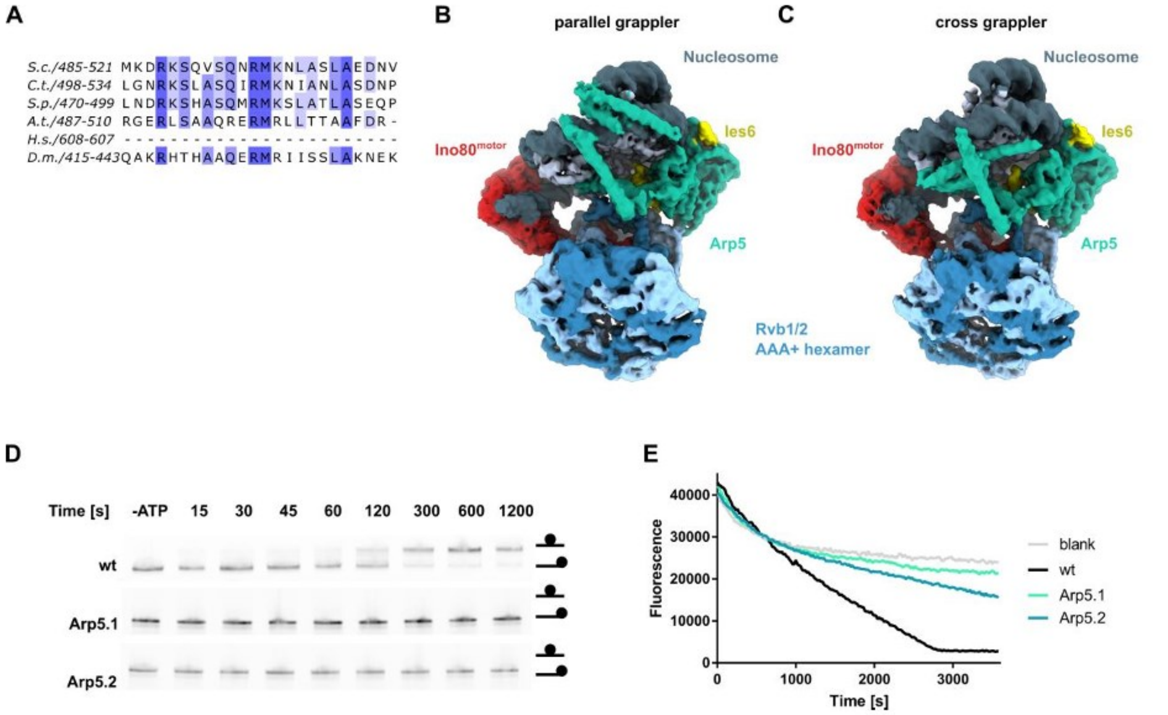


Fig. S8. Arp5 entry DNA interaction is important for *Ct*INO80$^{\Delta N}$ sliding and ATPase activity. (**A**) Multiple sequence alignment *(75)* of the grappler "foot". *S.c., S. cerevisiae; C.t., C. thermophilum; S.p., S. pombe; A.t., A. thaliana; H.s., Homo sapiens; D.m., D. melanogaster*. (**B**-**C**)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

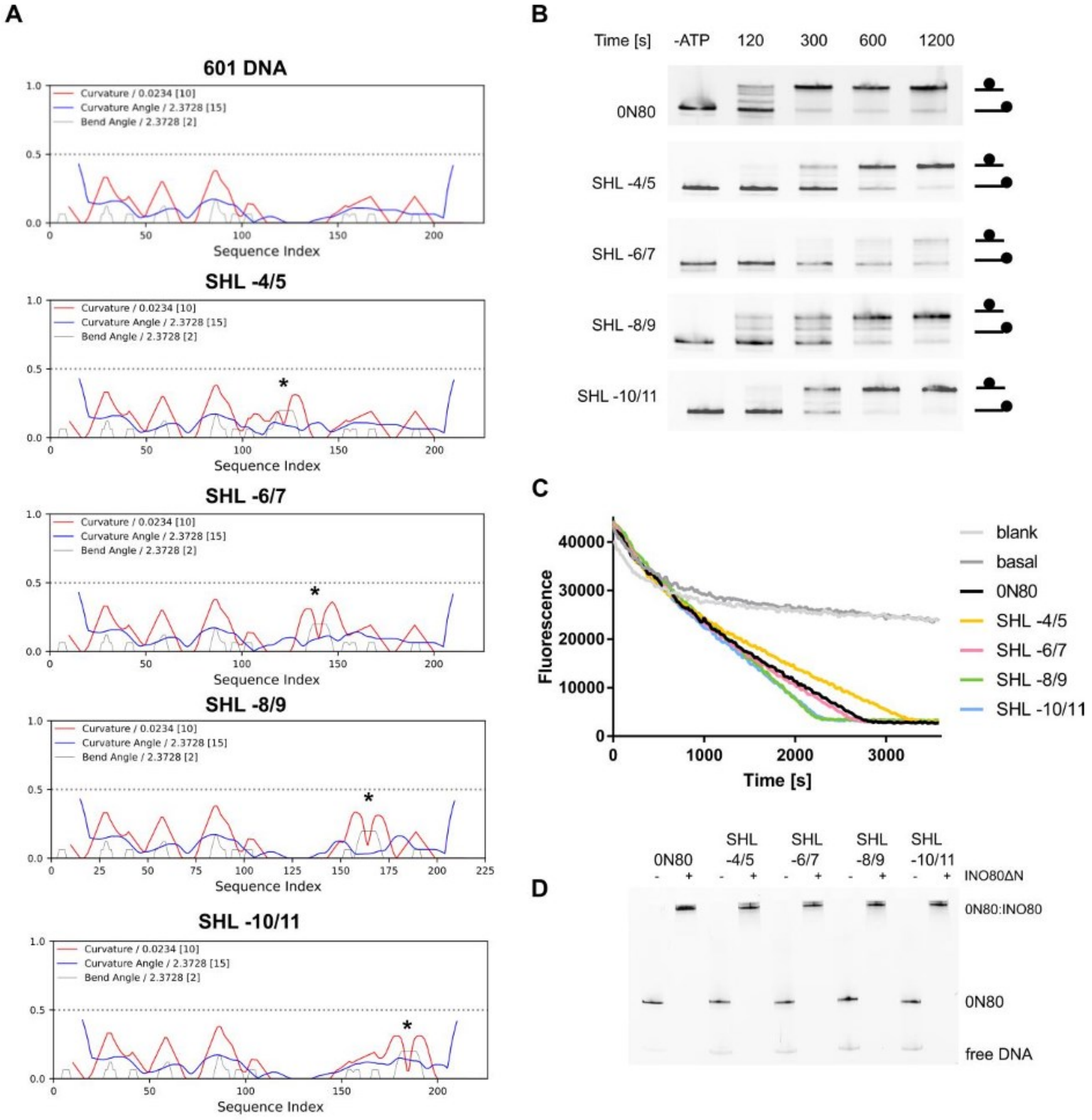


Fig. S9. Influence of A/T-rich DNA on *Ct*INO80$^{\Delta N}$ nucleosome remodeling. (**A**) DNA curvature analysis of the 601-based 0N80 DNA template sequence and sequences with replaced A/T-rich DNA cassettes (https://github.com/cgohlke/dnacurve). (**B**) Sliding of different 0N80 nucleosomes by *Ct*INO80$^{\Delta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

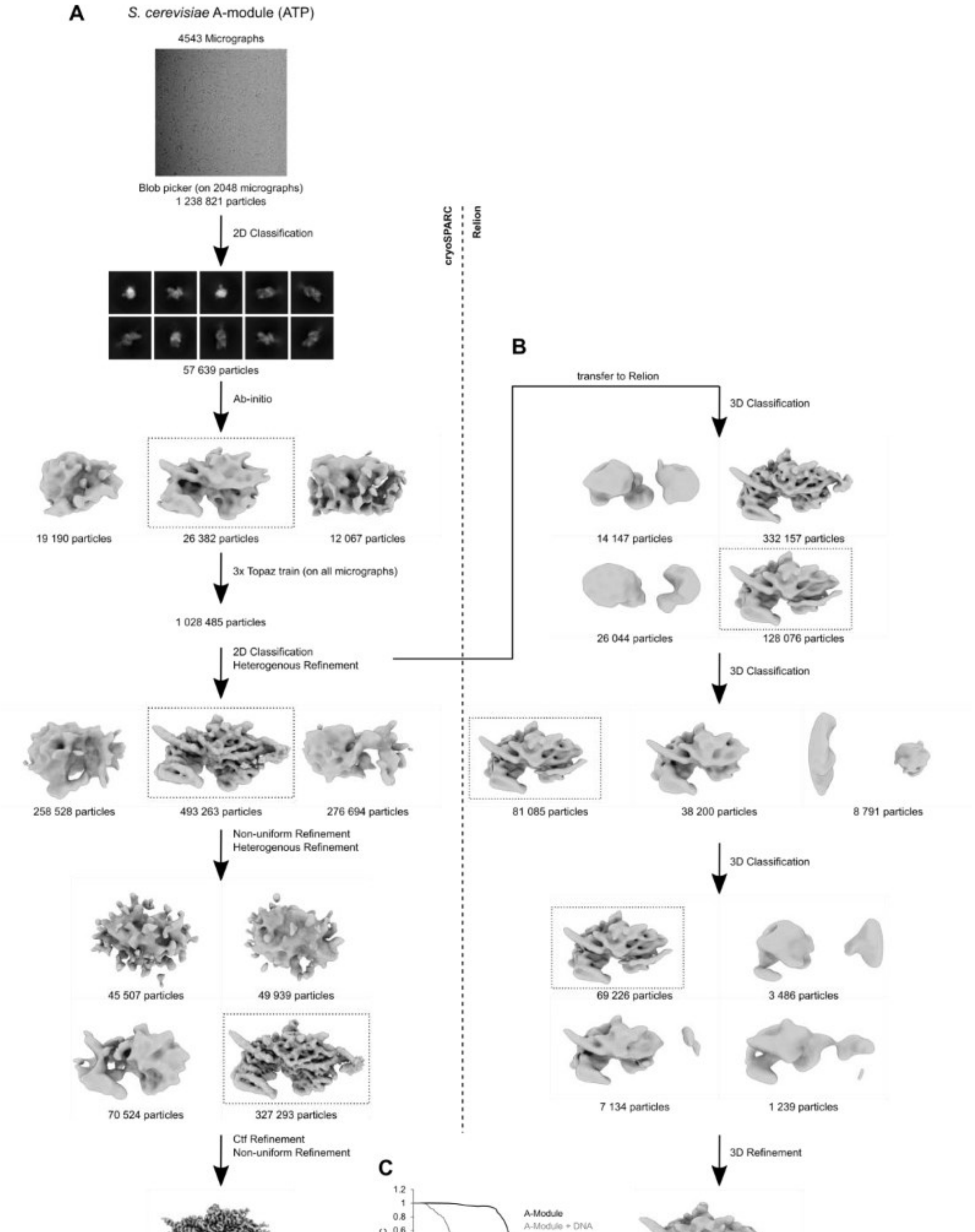


Fig. S10. Cryo-EM data analysis of *S. cerevisiae* A-module (ATP-state). (**A**) Cryo-EM data processing workflow of *S. cerevisiae* A-module using cryoSPARC v3.2.0 *(57)* and (**B**) A-module bound to DNA using cryoSPARC v3.2.0 *(57)* and Relion-3.0 *(58)*. A representative micrograph, representative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

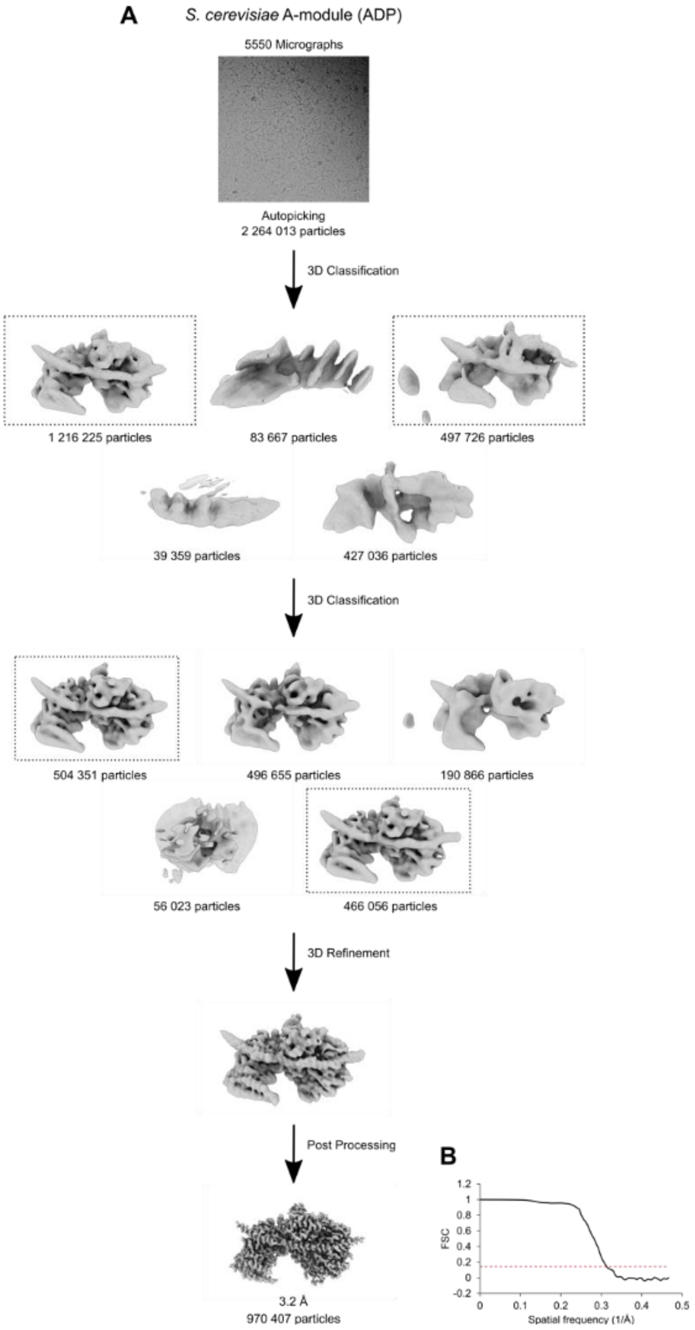


Fig. S11. Cryo-EM data analysis of *S. cerevisiae* A-module (ADP-state). (**A**) Cryo-EM data processing workflow of *S. cerevisiae* A-module using Relion-3.0 *(58)*. A representative micrograph and the cryo-EM data processing workflow are shown. (**B**) Gold-standard Fourier shell correlation (FSC)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

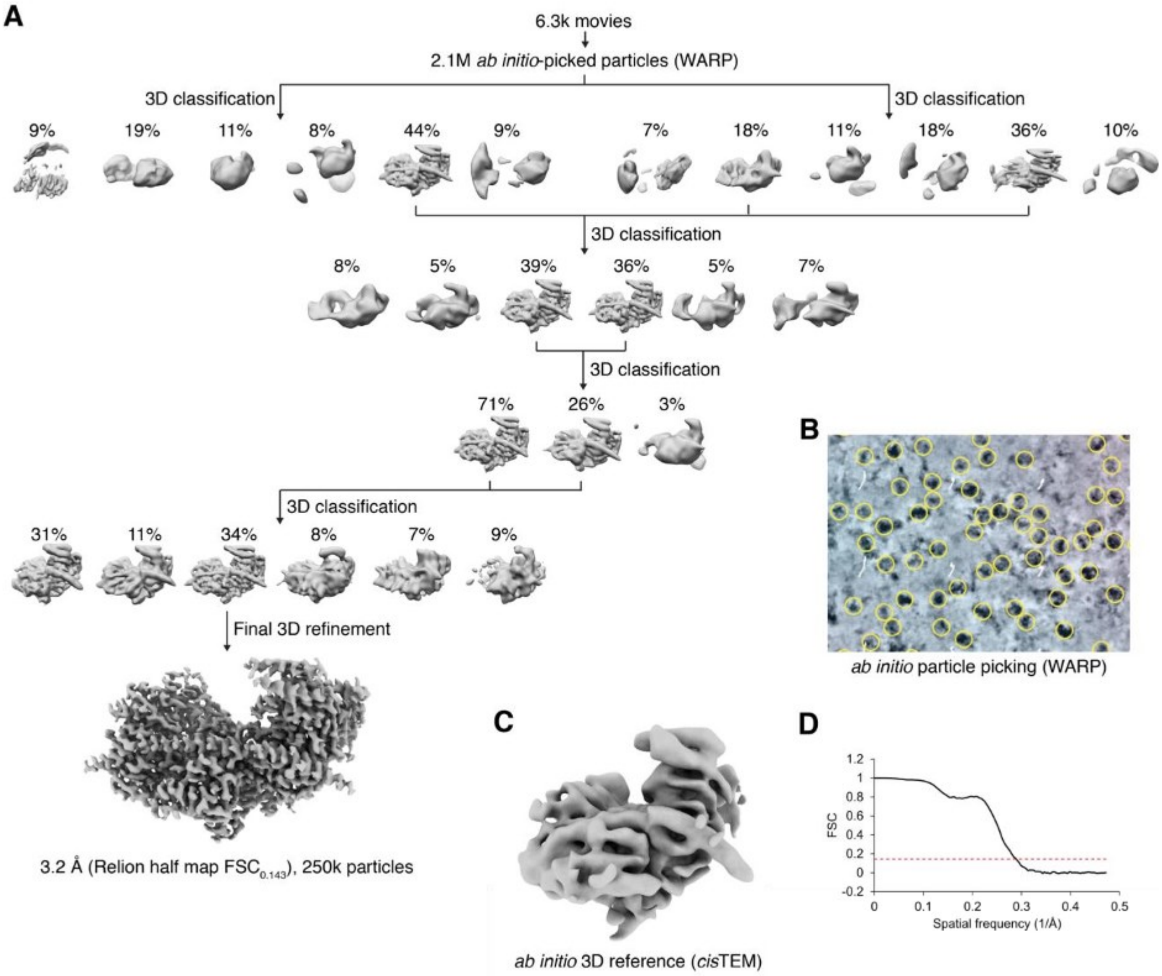
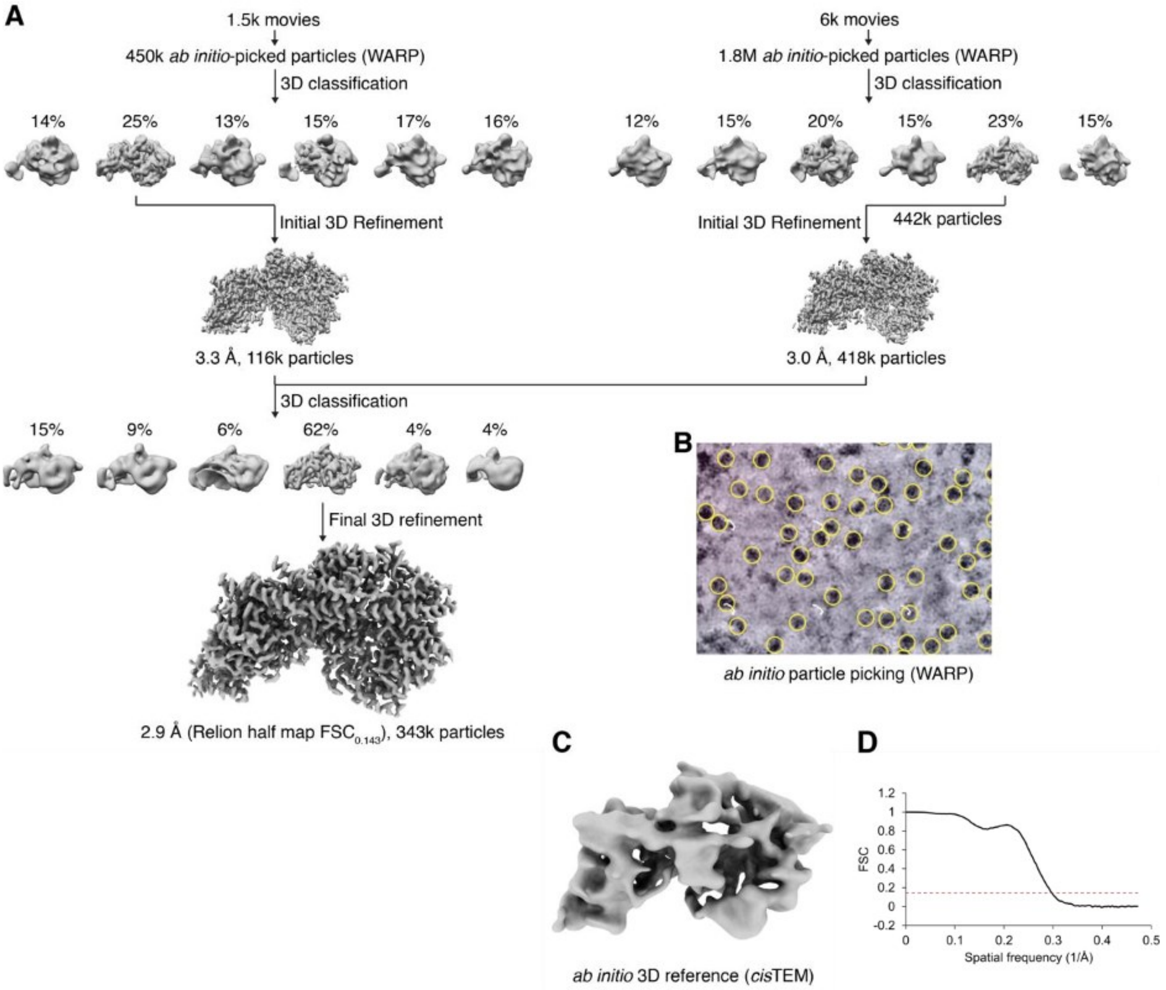


Fig. S12. Cryo-EM data analysis of *S. cerevisiae* A-module (ATPγS-state). (**A**) Good particles were selected through a series of focused 3D classifications and subsequently polished in Relion *(58)*. The final 3D reconstructions were generated and the resolution values calculated by Relion independent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



ab initio particle picking (WARP)

ab initio 3D reference (*cis*TEM)

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

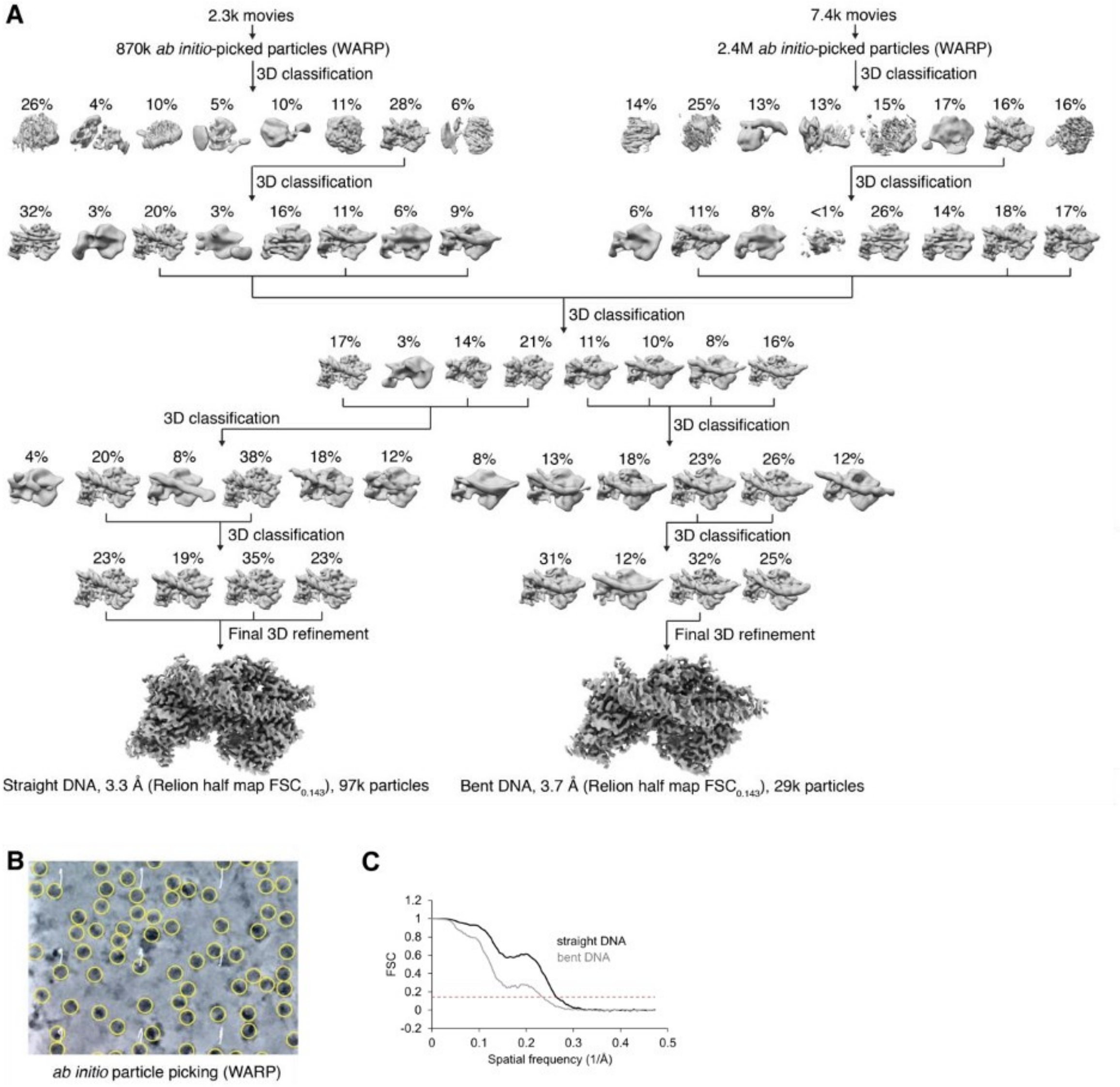


Fig. S14. Cryo-EM data analysis of *C. thermophilum* A-module and DNA (ADP·AlF$_x$). The *ab initio* 3D model without DNA bound was generated in *cis*TEM *(61)* and used to avoid bias in DNA presence and conformation. (**A**) DNA-bound A-module species were isolated from DNA-free species, followed by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

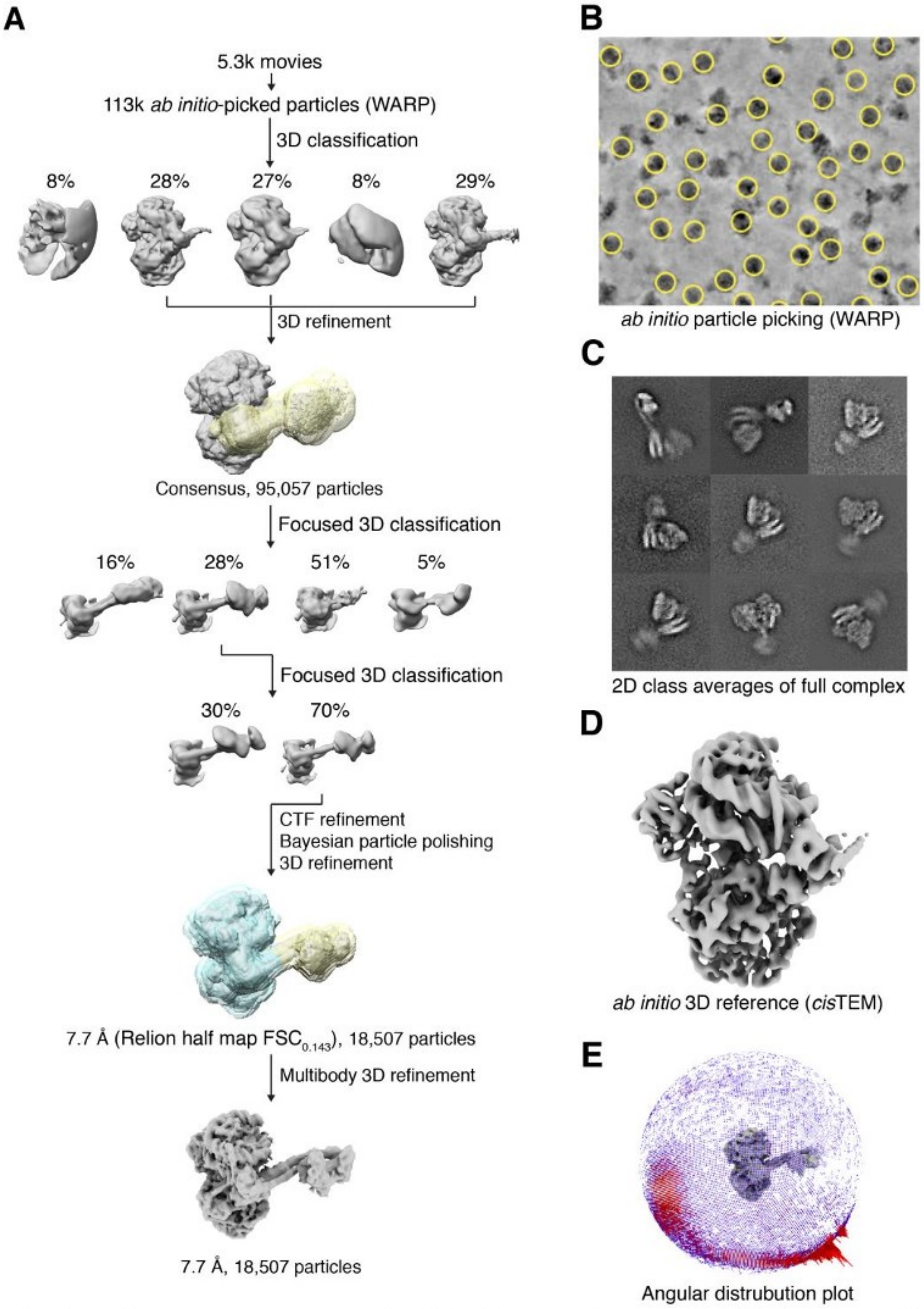


Fig. S15. Cryo-EM data analysis of *C. thermophilum* INO80$^{\Delta N}$. (**A**) A small subpopulation with the C- and A-modules more coherently placed were isolated through a series of focused 3D classifications in Relion and the final 3D reconstruction was generated by Relion multibody 3D refinement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

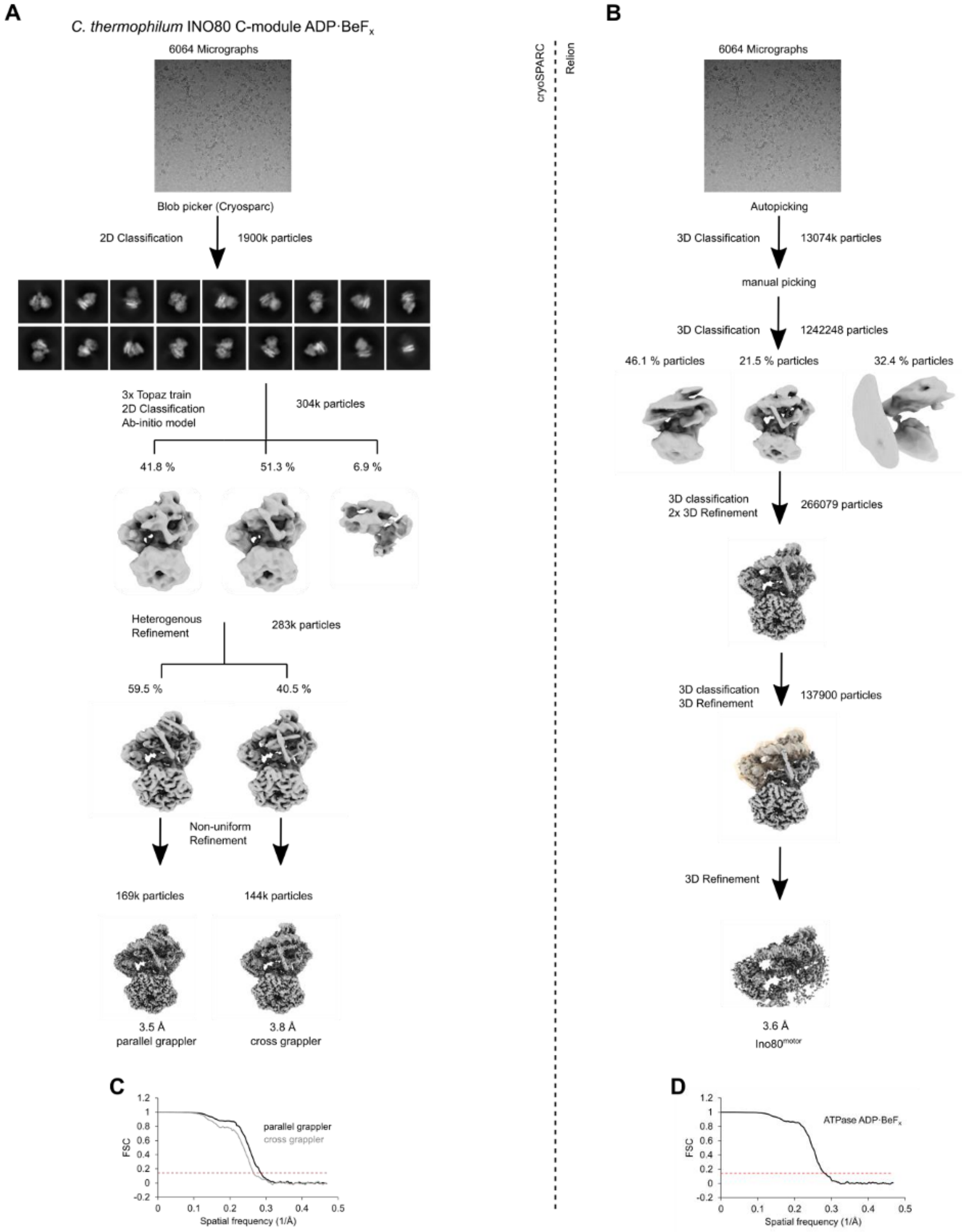


Fig. S16. Cryo-EM data analysis of *C. thermophilum* INO80 C-module with 0N80 nucleosomes. (**A**) Cryo-EM data processing workflow o*f C. thermophilum* INO80 C-module with

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

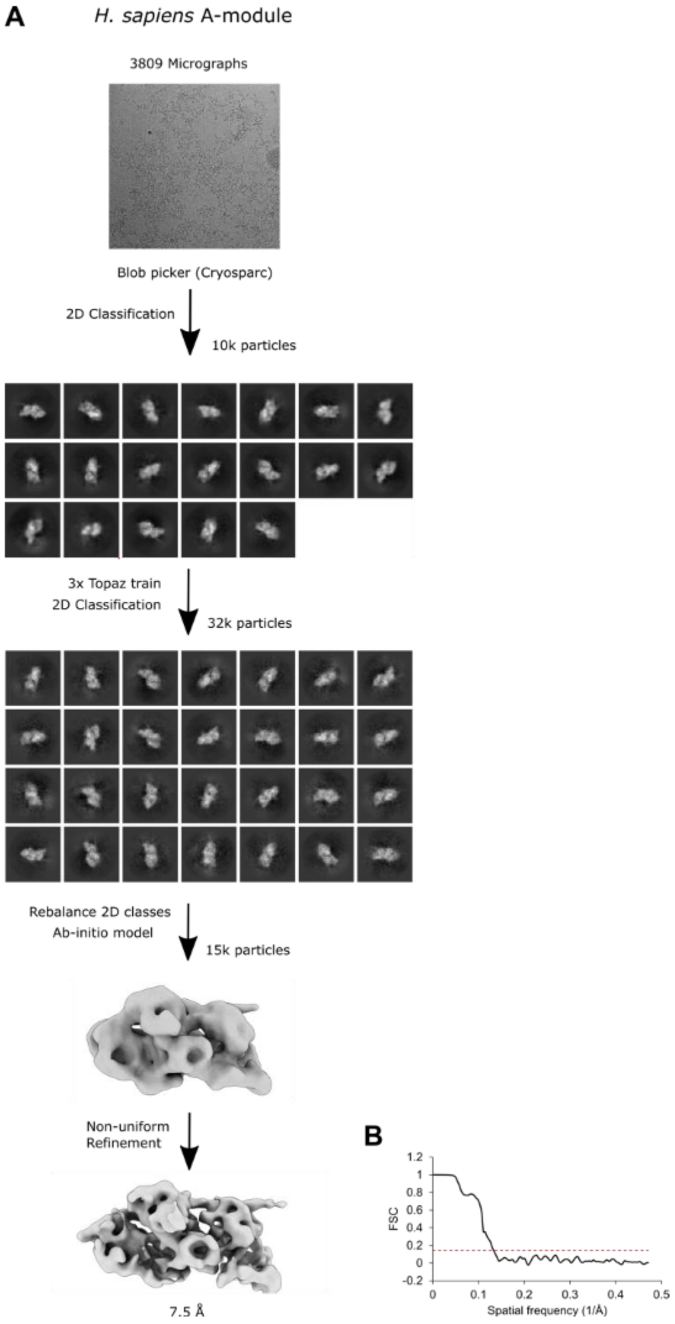


Fig. S17. Cryo-EM data analysis of *H. sapiens* A-module. (**A**) Cryo-EM data processing workflow of *Hs*A-module using cryoSPARC v3.2.0 *(57)*. Representative micrograph of *Hs*A-module and representative classes of a 2D classification of the particles used for the final A-module reconstruction after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

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

Complex	S.c A-module + ATPγS	S.c. A-module + ATP	S.c. A-module + ATP+ DNA	S.c. A-module + ADP	C.t. A-module + ATPγS	C.t. A-module + curved DNA (*apo*)	C.t. A-module + curved DNA (ADP-AlF$_x$)	C.t. A-module + straight DNA (ADP-AlF$_x$)	C.t. Ino80 C-module: cross grappler (ADP-BeF$_x$)	C.t. Ino80 C-module: parallel grappler (ADP-BeF$_x$)	C.t. Ino80motor + 0N80 nucleosome (ADP-BeF$_x$)	C.t. Ino80ΔN+0N80 nucleosome (*apo*)	h.s. A-module (*apo*)	C.t. Ino80 C-module (PDB: 6FML and EMDB 4277 and 4278)
Data collection and Processing														
Magnification	130	130	130	130	130	130	130	130	130	130	130	130	130	
Voltage (kV)	300	300	300	300	300	300	300	300	300	300	300	300	300	
Electron exposure (e-/Å^2)	47.37	44.68	44.68	45	44	50.5	42.4	42.4	48.4	48.4	48.4	50.5	52	
Defocus range (μm)	-1.1 to -2.9	-1.1 to -2.9	-1.1 to -2.9	-1.1 to -2.9	-1.1 to -2.9	-1.1 to -2.9	-1.1 to -2.9	-1.1 to -2.9	-1.1 to -2.9	-1.1 to -2.9	-1.1 to -2.9	-1.1 to -2.9	-1.1 to -2.9	
Pixel size (Å)	1059	1059	1059	1059	1059	1059	1059	1059	1059	1059	1059	1059	1059	
Symmetry imposed	C1	C1	C1	C1	C1	C1	C1	C1	C1	C1	C1	C1	C1	
Initial particle images (no.)	2100k	1230k	1230k	2300k	2250k	1180k	3240k	3240k	1900k	1900k	13074k	113k	2787k	
Final particle images (no.)	250k	327k	69k	970k	343k	325k	29k	97k	169k	144k	138k	18k	15k	
Map Resolution [Å]	3.2	3.3	7.5	3.2	2.9	3.4	3.7	3.3	3.8	3.5	3.6	7.7	7.5	
R-factor	0.29	0.27		0.27	0.27	0.28	0.29	0.28			0.24			
FSC	0.84	0.88		0.88	0.86	0.86	0.84	0.83			0.89			
Refinement														
Initial model used														
Model composition														
Non-hydrogen atoms	12571	13193			12225	13937		13537			15719			45030
Protein residues	1555	1629			1519	1550		1552			1219			4996
DNA nucleotides	–	–			–	72		52			286			290
Ligands (ATPγS, ATP, ADP)	3	3			3	3		3			1			8
Ions (Mg^{2+})	3	3			3	3		3			–			–
B-factors [Å^2]														
Protein	122.6	197.4			109.8	146.3		117.6			142.1			197.3
DNA	–	–			–	478.5		348.6			177.1			291.0
Ligands & Ions	73.7	138.6			77.5	98.7		74.2			145.1			129.8
RMS deviations (REFMAC5)														
Bond lengths [Å]	0.010	0.011			0.012	0.009		0.008			0.010			0.009
Bond angles [°]	1.71	1.88			1.75	1.82		1.72			2.31			1.78
Molprobity statistics/validation														
Molprobity Score	1.02	1.49			0.90	1.23		0.97			1.65			1.68
All-Atom Clashscore	1	2.44			0.86	2.33		1.13			4.15			4.15
Rotamer Outliers [%]	0.43	1.58			0.23	0.9		0.38			1.81			1.81
Ramachandran plot														
Favored [%]	96.5	95.47			97.27	96.61		97.2			96.23			96.23
Allowed [%]	3.31	4.22			2.60	3.26		2.74			3.1			3.1
Outliers [%]	0.19	0.31			0.13	0.13		0.07			0.67			0.67

Table S2. ***In vivo*** **yeast strains.**

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

Table S3. ***In vivo*** **yeast mutants.**

BP5148	*MATa, ade3::PGAL-HO, hmlΔ::pRS-1 hmrΔ::pRS-2, ChrIV491kb::GFPHOcs-hphNT1, ChrIV795kb::GFPHOinc-kanMX4, matHOcsΔ::pBR-1, URA3::YIplac211-empty*
BP5150	*MATa, ade3::PGAL-HO, hmlΔ::pRS-1 hmrΔ::pRS-2, ChrIV491kb::GFPHOcs-hphNT1, ChrIV795kb::GFPHOinc-kanMX4, matHOcsΔ::pBR-1, ino80::natNT2, URA3::INO80-2FLAG*
BP5155	*MATa, ade3::PGAL-HO, hmlΔ::pRS-1 hmrΔ::pRS-2, ChrIV491kb::GFPHOcs-hphNT1, ChrIV795kb::GFPHOinc-kanMX4, matHOcsΔ::pBR-1, ino80::natNT2, URA3::ino80-HSA-Q1-2FLAG*
BP5159	*MATa, ade3::PGAL-HO, hmlΔ::pRS-1 hmrΔ::pRS-2, ChrIV491kb::GFPHOcs-hphNT1, ChrIV795kb::GFPHOinc-kanMX4, matHOcsΔ::pBR-1, ino80::natNT2, URA3::ino80-HSA-Q2-2FLAG*
BP5415	*MATa, ade3::PGAL-HO, hmlΔ::pRS-1 hmrΔ::pRS-2, ChrIV491kb::GFPHOcs-hphNT1, ChrIV795kb::GFPHOinc-kanMX4, matHOcsΔ::pBR-1, ino80::natNT2, URA3::ino80-HSA-A2-2FLAG*
BP5137	*MATa/MATα, 2n[ade3::PGAL-HO, hmlΔ::pRS-1 hmrΔ::pRS-2, ChrIV491kb::GFPHOcs-hphNT1, ChrIV795kb::GFPHOinc-kanMX4], matHOcsΔ::pBR-1, ino80::natNT2, Δura3/URA3::ino80-HSA-Q1-2FLAG*
BP5138	*MATa/MATα, 2n[ade3::PGAL-HO, hmlΔ::pRS-1 hmrΔ::pRS-2, ChrIV491kb::GFPHOcs-hphNT1, ChrIV795kb::GFPHOinc-kanMX4], matHOcsΔ::pBR-1, ino80::natNT2, Δura3/URA3::ino80-HSA-Q2-2FLAG*
BP5139	*MATa/MATα, 2n[ade3::PGAL-HO, hmlΔ::pRS-1 hmrΔ::pRS-2, ChrIV491kb::GFPHOcs-hphNT1, ChrIV795kb::GFPHOinc-kanMX4], matHOcsΔ::pBR-1, ino80::natNT2, Δura3/URA3::ino80-HSA-Q1+Q2-2FLAG*
BP5140	*MATa/MATα, 2n[ade3::PGAL-HO, hmlΔ::pRS-1 hmrΔ::pRS-2, ChrIV491kb::GFPHOcs-hphNT1, ChrIV795kb::GFPHOinc-kanMX4], matHOcsΔ::pBR-1, ino80::natNT2, Δura3/URA3::ino80-E842A-2FLAG*
BP5128	*MATa, ade3::PGAL-HO, hmlΔ::pRS-1 hmrΔ::pRS-2, ChrIV491kb::GFPHOcs-hphNT1, ChrIV795kb::GFPHOinc-kanMX4, matHOcsΔ::pBR-1, arp8::natNT2, URA3::2FLAG-ARP8*
BP5125	*MATa, ade3::PGAL-HO, hmlΔ::pRS-1 hmrΔ::pRS-2, ChrIV491kb::GFPHOcs-hphNT1, ChrIV795kb::GFPHOinc-kanMX4, matHOcsΔ::pBR-1, ura3Δ, arp8::natNT2*
BP5133	*MATa, ade3::PGAL-HO, hmlΔ::pRS-1 hmrΔ::pRS-2, ChrIV491kb::GFPHOcs-hphNT1, ChrIV795kb::GFPHOinc-kanMX4, matHOcsΔ::pBR-1, arp8::natNT2, URA3::2FLAG-arp8-aa1-197Δ*

Table S4. Oligonucleotides (5’ to 3’) for *C. thermophilum* and *S. cerevisiae* cloning.

C. thermophilum	
Ino80-720-1855: open-vector-fwd	ATATTTATAGGTTTTTTTATTACAAAACTG
Ino80-720-1855: open-vector-rev	TCTAGAGCCTGCAGTCTCG
Ino80dN-rev	ACTGCAGGCTCTAGATTAATTAATCACTTGTCGTCATCA
Ino80dN- fwd	AGCTCAAGTTTCAGTCCAAAGGCTACAACC
Ino80-1-850:NA- fwd	CGCAAAGAAGAAGCCGACTACAAGGACGACGATGACAAG
Ino80-1-850NA-rev	CGTCGTCCTTGTAGTCGGCTTCTTCTTTGCGAGCATTC
Ino80-545-850A- fwd	ATGCAGCGTAAGCGGCGTCGCGAGAAGTCG
Ino80-545-850A-rev	ATTTATAGGTTTTTTTATTACAAAACTG
HSA1-insert- fwd	ACCAGATTTACGATCAGATCTGGCGTGATCTGGCTGCCG
HSA1-insert-rev	CGTTTCGCCCTAGCCTGGAGATCCTTTGTGCCCTTGTTGGTACGC
HSA1-vector- fwd	CTCCAGGCTAGGGCGAAACGTGTCATGCG
HSA1-vector-rev	GATCTGATCGTAAATCTGGTTGTAGCCTTTGGACTGAAACTTGAGCTCC
Insert HSA1	TGGCGTGATCTGGCTGCCGCGGATGTTAGCGCAGTTTTCGCTCTGGCGACGGACTCGTACGCTACCAAAGCATCCAACCTCAAGGCGACGGCCATCCTGGCTTCCGCGGAAGCAAAAGCCTGGCAACTGCGTACCAACAAGGGCACAAAGGAT
HSA2-insert- fwd	CTTCCAACCTCAAGAAGACGGCCATCCTGGCTTCCAAGGAAGC
HSA2-insert-rev	CGAGCATTCTCGAGTTCGAGCCTTTCAGCAGCCTTGCGTAAATCG
HSA2-vector- fwd	CTCGAACTCGAGAATGCTCGCAAAGAAGAAGC
HSA2-vector-rev	CGTCTTCTTGAGGTTGGAAGCTTTGG
Insert HSA2	GCCATCCTGGCTTCCAAGGAAGCCAAACGCTGGCAACTGCGTACCAACGCTGGCACAGCCGATCTCCAGGCTAGGGCGGCAGCGGTCATGGCTGACATGATGGGCTTCTGGGCCGCAAACGAGGCGGAAGAGCGCGATTTACGCAAGGCTGCTGAAAGG
Ino80.1- fwd	CCCCTGGAGCTCGCGTTTCAGTCCGCAGGCTACAACCAG
Ino80.1-rev	CTGGTTGTAGCCTGCGGACTGAAACGCGAGCTCCAGGGG
Ino80.2- fwd	GATCAGATCTGGGCTGATCTGGCTGCCAAAGATGTTAGC
Ino80.2-rev	GCTAACATCTTTGGCAGCCAGATCAGCCCAGATCTGATC
Arp5.1- fwd	TGCGGCAGGCGGCGCCGGCGCAGAC
Arp5.1-rev	GCTGCGGCGCTTCCCGAGCCGGTCGGGTTG
Arp5.2- fwd	GAAAGCGCAGCGGAAATCGCGGCTCTCCTCGACG
Arp5.2-rev	CGTCGAGGAGAGCCGCGATTTCCGCTGCGCTTTC
Arp8.1-vector-rev	CATATTTATAGGTTTTTTTATTACAAAACTGTTACGAAAACAGTAAAATACTTATTTATTTGCGAGATGG
Arp8.1-insert-fwd	TAAAAAAACCTATAAATATGGTGGGAAAAGTGAGCGAGGCCG
Arp8.1-vector- fwd	GATCAGATCCTGTCGCTGCGCCTCCAGAACGAGG
Arp8.1-insert-rev	CGCAGCGACAGGATCTGATCATCGCGGGCCATATAATCAGTATAGGCG
Insert Arp8.1	GTGGGAAAAGTGAGCGAGGCCGTGCTTGCCGCCGAAGGGCTGGAACGCACCGACAATGGCATGAGGCAAACAAGCTGGCCTGAGGTGCAGCCCATTGCCGCAGCTGCCGCATATACTGATTATATGGCCCGCGAT
Arp8dN-rev	CATATTTATAGGTTTTTTTATTACAAAACTGTTACGAAAACAG
Arp8dN- fwd	CTGATTAACCAAAAGAACTATTATACTGATTATATGAAGCGCG
S. cerevisiae	
Ino80-598-fwd	GTATTCGCATTTCATTGGAAGGAAAGACTACAAGGACGACGATGACAAG
Ino80-598-rev	CGTCCTTGTAGTCTTTCCTTCCAATGAAATGCGAATACAATTC

Table S5. Oligonucleotides (5’ to 3’) for nucleosome preparation, anisotropy and cryo-EM.

Nucleosome-0N80-fwd	CTGGAGAATCCCGGTGCCGAGG
Nucleosome-0N80-rev	CGGTACCCGGGGATCCTCTAG
DNA templates (5´FAM)	
601 sequence	CTGGAGAATCCCGGTGCCGAGGCCGCTCAATTGGTCGT AGCAAGCTCTAGCACCGCTTAAACGCACGTACGCGCTG TCCCCCGCGTTTTAACCGCCAAGGGGATTACTCCCTAGT CTCCAGGCACGTGTCAGATATATACATCCTGTGCATGT ATTGAACAGCGACCTTGCCGGTGCCAGTCGGATAGTGT TCCGAGCTCCCACTCTAGAGGATCCCCGGGTACCGA
601 sequence SHL -4/5	CTGGAGAATCCCGGTGCCGAGGCCGCTCAATTGGTCGT AGCAAGCTCTAGCACCGCTTAAACGCACGTACGCGCTG TCCCCCGCGTTTTAACCGCCAAGGGGATTACTCCCTAGT TCTTTTTTTTTTTTCAGATATATACATCCTGTGCATGTAT TGAACAGCGACCTTGCCGGTGCCAGTCGGATAGTGTTC CGAGCTCCCACTCTAGAGGATCCCCGGGTACCGA
601 sequence SHL -6/7	CTGGAGAATCCCGGTGCCGAGGCCGCTCAATTGGTCGT AGCAAGCTCTAGCACCGCTTAAACGCACGTACGCGCTG TCCCCCGCGTTTTAACCGCCAAGGGGATTACTCCCTAGT CTCCAGGCACGTGTCAGATCTTTTTTTTTTTTGCATGTA TTGAACAGCGACCTTGCCGGTGCCAGTCGGATAGTGTT CCGAGCTCCCACTCTAGAGGATCCCCGGGTACCGA
601 sequence SHL -8/9	CTGGAGAATCCCGGTGCCGAGGCCGCTCAATTGGTCGT AGCAAGCTCTAGCACCGCTTAAACGCACGTACGCGCTG TCCCCCGCGTTTTAACCGCCAAGGGGATTACTCCCTAGT CTCCAGGCACGTGTCAGATATATACATCCTGTGCATGT ATTGATCTTTTTTTTTTTTGGTGCCAGTCGGATAGTGTTC CGAGCTCCCACTCTAGAGGATCCCCGGGTACCGA
601 sequence SHL -10/11	CTGGAGAATCCCGGTGCCGAGGCCGCTCAATTGGTCGT AGCAAGCTCTAGCACCGCTTAAACGCACGTACGCGCTG TCCCCCGCGTTTTAACCGCCAAGGGGATTACTCCCTAGT CTCCAGGCACGTGTCAGATATATACATCCTGTGCATGT ATTGAACAGCGACCTTGCCGGTGCTCTTTTTTTTTTTTTC CGAGCTCCCACTCTAGAGGATCCCCGGGTACCGA
Anisotropy DNA	
Aniso_fwd (5´FAM)	AATTGGTCGTAGCAAGCTCTAGCACCGCTTAAACGCAC GTACGCGCTGTC
Aniso_rev	GACAGCGCGTACGTGCGTTTAAGCGGTGCTAGAGCTTG CTACGACCAATT
cryo-EM DNA	
EM_fwd	CTTACCCTGCGTGCCCGCGCACCCTGGCGACTTCGCCTC GTTTTGGCGATTTTCTTAG
EM_rev	CTAAGAAAATCGCCAAAACGAGGCGAAGTCGCCAGGG TGCGCGGGCACGCAGGGTAAG

Supplementary Movie 1 legend

Movie 1 illustrates the conformational change of the DNA bound to the *Ct*A-module with different interaction sites (3D variability analysis in cryoSPARC).

Supplementary Movie 2 legend

Structural morph illustrating the conformational changes of the Ino80^{motor} bound to the nucleosome in the apo-state or bound to ADP·BeF$_x$.

Supplementary Movie 3 legend

Movie 3 illustrates the conformational change of the Arp5 grappler and its contacts to the entry DNA of the nucleosome (3D variability analysis in cryoSPARC).

Raw data file legend

Excel file including raw data for Fig. 2D, Fig. 4G to I, Fig. 7D and E, Fig. 8C and D, Fig. S4D, Fig. S5, Fig. S6B, Fig. S9D.

3. Discussion

The human subgroup III Slfn members SLFN5 and SLFN11 are two of the functionally best described Slfn proteins. SLFN5 is involved in tumorigenesis and was suggested to act as a transcriptional regulator [11, 62]. It shows antiviral activity by targeting transcription of viral genes [9], while SLFN11 targets translation of viral proteins in a codon-usage-dependent manner [6]. Furthermore, SLFN11 sensitizes cancer cells to DNA damaging agents by irreversibly blocking stalled replication forks [71]. However, little is known about the underlying molecular mechanisms of SLFN5 and SLFN11 and structural as well as biochemical in vitro characterization of these proteins is largely missing.

Thus, open questions regarding these Slfn proteins include:

- What is the overall structural organization of SLFN5 and SLFN11?
- What are the functions of the individual domains?
- Does SLFN5/SLFN11 recognize and process tRNA?
- What is the substrate specificity of the SLFN11 helicase domain?
- How is SLFN11 regulated by post-translational modifications?

These questions will be discussed in the following sections with the goal to gain mechanistic insights into SLFN5 and SLFN11 and subgroup III Slfn proteins in general.

3.1. Structural insights into human subgroup III Slfn members SLFN5 and SLFN11

Subgroup III Slfn members share a conserved domain composition, consisting of an N-terminal Slfn core domain, a central linker domain and a C-terminal helicase domain (Figure 15A) [2]. So far, a crystal structure of the Slfn core domain of rat Slfn13 was the only available structure of any subgroup III Slfn protein [22]. Therefore, structural insights into a full-length Slfn protein could significantly advance the understanding of this protein family.

The cryo-EM structure of the human SLFN5 apoenzyme is the first structure of a full-length subgroup III Slfn protein, giving novel insights into the overall domain arrangement (Figure 15B) [322]. The Slfn core domain, that is generally associated with nucleic acid binding and RNA cleavage, adopts a horseshoe-like shape as previously seen in the crystal structure of the rat Slfn13 core domain and the recently reported structure of the SLFN12-PDE3A complex [22, 25, 26, 322]. This includes a highly conserved Cys3His zinc finger as part of the Slfn core C-lobe. The channel formed by the SLFN5 core domain measures approximately 18 Å and is lined by several positively charged residues [322]. Thus, it is wide enough to serve as a potential binding platform for base paired nucleic acids. Even though the basic residues are only partially conserved, positively charged patches along the channel are present in the other Slfn structures as well [22]. Indeed, alanine mutations of the positively charged residues in rat Slfn13 resulted in reduced endoribonuclease activity [22]. The linker domain (also called SWA(V)DL-domain) has a globular fold and sterically separates the Slfn core domain from the C-terminal helicase domain (Figure 15B) [322]. The highly conserved SWAVDL-motif is part of the hydrophobic core, but also interacts directly with the helicase N-lobe. The C-lobe of the helicase is not resolved in the cryo-EM map, presumably due to conformational flexibility in the nucleotide-free state (Figure 15C). However, superimposing of an AlphaFold model [323] of the entire helicase domain with the SLFN5 structure indicates that the helicase lobes could adopt a closed (ATP-bound) state without causing steric clashes [322].

We further solved the structure of the full-length SLFN11 apoenzyme, revealing striking conformational differences between SLFN5 and SLFN11 despite a high degree of sequence conservation [324]. SLFN11 forms a ring-shaped symmetric dimer with interfaces between the Slfn core domains and the helicase domains (Figure 15B). The two Slfn core domains align to form a continuous positively charged channel with a width of approximately 20 Å. A similar dimerization mode has previously been reported for the subgroup II Slfn protein SLFN12, which forms a molecular glue-mediated complex with the phosphodiesterase PDE3A [25, 26]. This complex shows a similar interface between the Slfn core domains; however, the second interface is formed by the PDE3A protomers (Figure 4B). The similarity between the dimer structures suggests that the SLFN11 helicase domain and PDE3A in case of the SLFN12-PDE3A complex serve a structural function in the context of dimerization. As dimerization can be observed for subgroup II and III Slfn members, it seems plausible that other Slfn family members form dimers as well. In line with this, oligomerization was observed for a truncated SLFN14 construct in vitro [74]. In contrast to the SLFN5 structure, both lobes of the helicase domain are resolved in the SLFN11 dimer structure (Figure 15B) [322, 324]. Interestingly, the helicase domain is rotated by approximately 180° compared to its orientation in the SLFN5 structure and it interacts with the linker domain via the helicase C-lobe, while the N-lobes form the dimer interface [324]. This dramatic change is possible due to a rearrangement of the inter-domain (ID)-region that connects the linker with the helicase domain (Figure 15B). In SLFN5, the ID-region forms a hairpin-like structure with mixed α/β topology, which allows the helicase N-lobe to interact with the linker domain [322]. In SLFN11, the ID-region forms a continuous helix, the ID-helix, that extends outwards from the linker domain and interacts with both helicase lobes (Figure 15B) [324]. This places the SLFN11 helicase N-lobe away from the linker domain and allows the helicase C-lobe to interact with the linker domain. The ID-helix appears to lock the helicase lobes in a closed conformation and could therefore act as a brace. Interestingly, monomeric SLFN11 apoenzyme adopts the same helicase conformation as the protomers within the dimer, suggesting that this conformation is not a result of dimerization but rather a prerequisite for dimerization [324]. However, the cryo-EM map of monomeric SLFN11 lacks density for the Slfn core C-lobe and the bridging domains, suggesting conformational flexibility (Figure 15C) [324]. Thus, dimerization might be necessary to lock the lobes of the Slfn core domains in a stable conformation. In contrast to SLFN11, the C-lobe of the SLFN5 core domain adopts a stable conformation [322]. This might be connected to the conformation of the region connecting the Slfn core domain with the linker domain, as it is visible in the SLFN5 structure but flexible in the SLFN11 structure (Figure 15B) [322, 324]. In the SLFN5 structure, this region interacts with the Slfn core domain, linker domain and the ID-region, possibly stabilizing the overall fold.

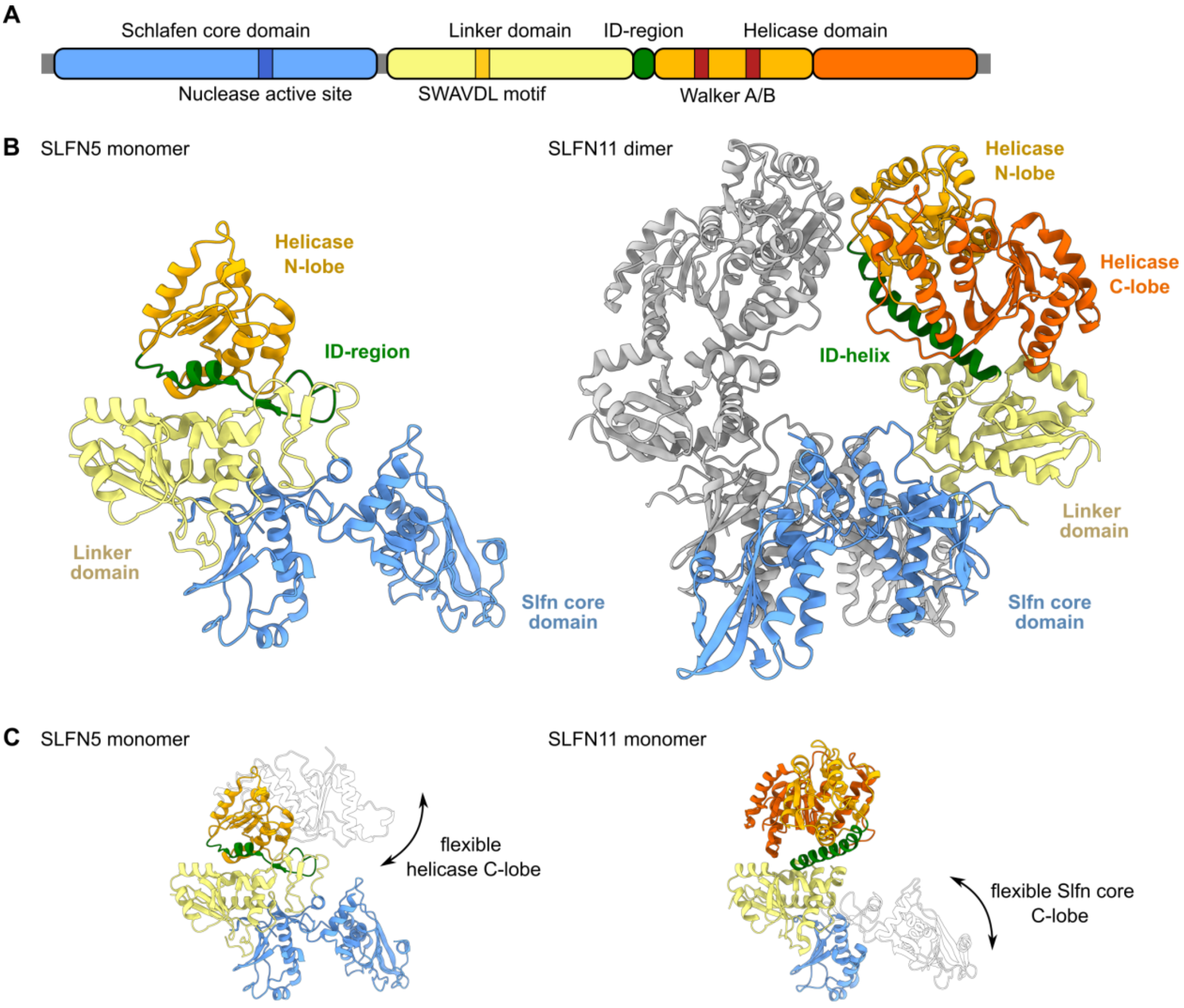


Figure 15: Structures of human SLFN5 and SLFN11 apoenzymes. (A) Schematic domain overview of subgroup III Slfn proteins. The Slfn core domain (blue), linker domain (yellow), ID-region (green) and helicase domain (orange) are color-coded. The putative nuclease active site, SWAVDL motif and Walker A/B motifs are highlighted. (B) Ribbon representation of full-length human SLFN5 (PDB: 7PPJ) [322] and dimeric full-length human SLFN11 (PDB: 7ZEL) [324]. Domains are labeled and colored as in (A). The second SLFN11 protomer is depicted in grey. (C) Structural comparison of monomeric SLFN5 (PDB: 7PPJ) [322] and monomeric SLFN11. Domains are colored according to (A). The SLFN5 helicase C-lobe that is not resolved in the SLFN5 reconstruction is indicated as a transparent cartoon model based on an AlphaFold prediction [323, 325]. The SLFN11 Slfn core C-lobe that is not resolved in the monomeric SLFN11 apoenzyme reconstruction (EMDB: 14693) is indicated as a transparent cartoon model based on the structure of the SLFN11 dimer (PDB: 7ZEL) [324]. The ID-regions are depicted in green to visualize the conformational differences between the presented SLFN5 and SLFN11 structures.

3.2. Nuclease activity of SLFN5 and SLFN11

Several human Slfn proteins have been shown to exhibit ribonuclease activity. This includes tRNA cleavage by SLFN11, SLFN12, and SLFN13 [22, 45, 46], rRNA cleavage by SLFN12, SLFN13, and SLFN14 [22, 23, 26], and mRNA cleavage by SLFN14 [23]. As RNase activity seems to be a common feature of Slfn proteins, we investigated the tRNA cleavage activity of SLFN5 and SLFN11 [322, 324].

While SLFN11 cleaves tRNA^{Ser} in an endonucleolytic fashion, SLFN5 shows no tRNase activity [322, 324]. However, this does not exclude the possibility that SLFN5 has a different substrate specificity. Mutational analysis of potential nuclease active site residues in SLFN11 revealed three essential residues, namely E209, E214 and K216 that are located in the C-lobe of the Slfn core domain [324]. This is in partial agreement with a study on rat Slfn13 that identified a three-carboxylate-triad that is essential for Slfn13 endonuclease activity [22]. While two of the residues correspond to E209 and E214, the third corresponding residue of the three-carboxylate-triad is not essential for SLFN11 endonuclease activity. The three essential residues identified in SLFN11 are conserved in all human Slfn proteins except for SLFN5, which lacks the conserved lysine residue [324]. This suggests that SLFN5 has no RNase activity while the other human Slfn proteins are RNase active. In support of that, mouse Slfn2, which lacks two of the three active site residues, also lacks RNase activity [24, 322]. In the cryo-EM map of SLFN11, we identified unassigned density near the nuclease active site, which we interpreted as Mg^{2+}-ion, as $MgCl_2$ was present in the buffer [324]. However, the resolution is not high enough to unambiguously tell whether one or two ions are bound. It appears that E209 and E214 coordinate the Mg^{2+}-ion and the E209A mutation disrupts the coordination of the ion as shown by a loss of the additional density in the $SLFN11^{E209A}$ cryo-EM map [324]. This nicely explains why E209 and E214 in SLFN11 or the corresponding residues in SLFN12 or SLFN13 are essential for the nuclease activity [22, 46]. Screening of different bivalent cations in tRNA cleavage reactions showed that SLFN11-mediated cleavage is Mn^{2+}-dependent and Mg^{2+}-independent [324]. In contrast, RNase activity of SLFN13 and SLFN14 seem to be less ion-specific as both show nuclease activity in the presence of Mn^{2+} or Mg^{2+}, with a higher activity in case of Mn^{2+} [22, 23]. This seems plausible, as Mn^{2+} and Mg^{2+} have the same charge and very similar ionic radii [326].

A recent study showed that DDA-induced cell death relies on SLFN11-mediated cleavage of type II tRNAs, which differ from type I tRNAs in the presence of an extended variable loop and include all leucine and serine tRNAs [45]. Particularly cleavage of tRNA-Leu-TAA renders genes with high TTA codon usage susceptible to translational suppression by SLFN11 [45]. In line with that, SLFN11 rapidly cleaves type II $tRNA^{Ser}$ in vitro, while cleavage of type I $tRNA^{Met}$ is much less pronounced [324]. This indicates that the endonuclease activity is highly substrate specific, which might result from recognition of specific tertiary structures. To gain insights into the mode of substrate recognition by SLFN11, we analyzed the SLFN11-tRNA complex by cryo-EM. The cryo-EM reconstruction explains how a single tRNA molecule is recognized by a SLFN11 dimer (Figure 16) [324]. The positively charged channel formed by the Slfn core domains of the two protomers forms a substrate specific binding site for tRNA. The overall conformation of SLFN11 is similar to the apoenzyme structure, indicating that the SLFN11 dimer is poised for tRNA binding. SLFN11 mostly interacts with the acceptor stem and the T-loop of the tRNA, suggesting that these regions contribute to tRNA recognition and substrate specificity. As the tertiary structure of this region is highly conserved between different tRNAs, it is not surprising that binding of both type I and type II tRNAs can be observed in vitro [324]. The variable loop seems to be important for the regulation of the RNase activity [45]. The resolution of the cryo-EM reconstruction did not allow for detailed model building; however, it is clear that the variable loop is in close proximity to the Slfn core domain and might therefore contribute to tRNA binding and substrate recognition (Figure 16C) [324]. Interestingly, a recent study revealed that type II tRNA-Leu-TAA is the physiological substrate of the SLFN12 RNase, which indicates striking similarities in substrate specificity between SLFN11 and SLFN12 [46]. The study showed that changes of the sequence of the acceptor stem or variable loop have a significant impact on SLFN12 RNase activity, indicating that SLFN12 recognizes distinct RNA sequences. Considering the high degree of conservation between Slfn proteins, it is likely that the tRNA binding mode observed for SLFN11 is conserved for multiple Slfn proteins including SLFN12 [324]. Thus, the protein-tRNA interactions observed in the model are in good agreement with the hypothesis of tRNA sequence recognition by SLFN12, indicating that this model might be applicable to other Slfn proteins as well [46]. Despite the similarities, some differences appear in the substrate specificity as SLFN11 downregulates several type II tRNAs (Leu and Ser) while SLFN12 appears to be more specific for tRNA-Leu-TAA [45, 46]. Interestingly, SLFN12 cleaves different tRNAs in vitro, something that has also been shown for SLFN11 and SLFN13 [22, 45, 46]. However, in a cellular context, only tRNA-Leu-TAA was significantly downregulated by SLFN12 [46]. Thus, it is possible that relatively small differences in in vitro cleavage assays translate to major differences in more complex in vivo systems. Therefore, studies that systematically identify the physiological substrates of different Slfn proteins are needed.

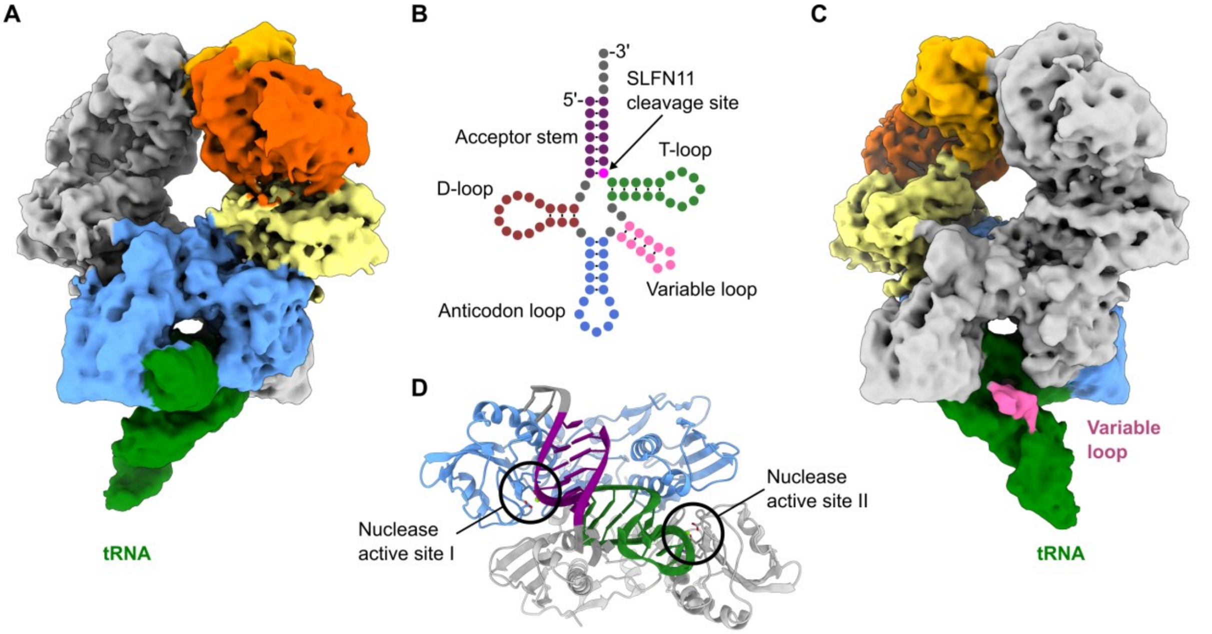


Figure 16: Structural basis for tRNA recognition and cleavage by SLFN11. (A) Cryo-EM reconstruction of tRNA bound SLFN11 dimer (front view; EMDB: 14695) [324]. The Slfn core domain (blue), linker domain (yellow), helicase domain (orange) and tRNA (green) are color-coded. The second SLFN11 protomer is colored in grey. (B) Schematic representation of type II tRNA^{Ser}. The structural features of the tRNA are labeled and color-coded. The SLFN11 cleavage site is indicated. (C) Cryo-EM reconstruction of tRNA bound SLFN11 dimer (back view; colored as in (A); EMDB: 14695) [324]. The variable loop is highlighted in pink. (D) Structural docking model of tRNA (PDB: 5AXM) [327] bound by SLFN11 dimer (bottom view; PDB: 7ZEL) [324]. SLFN11 is colored as in (A) and the tRNA is colored according to (B). The tRNA acceptor stem and T-loop are depicted. The nuclease active sites of both SLFN11 protomers are highlighted.

The cryo-EM map of tRNA bound SLFN11 reveals that the tRNA is clamped between the lobes of the Slfn core domains and is in close proximity to the nuclease active sites of both SLFN11 protomers, approximately 10 and 20 nucleotides from the 3’ end of the tRNA, respectively (Figure 16) [324]. However, RNA sequencing analysis shows that SLFN11 cleaves tRNA^{Ser} mainly 10 nucleotides from the 3’ end within the acceptor stem while the potential second cleavage site, which would be located within the T-loop, is not cleaved by SLFN11 [324]. The position of the endonucleolytic cleavage site identified for SLFN11 is in good agreement with the major cleavage sites identified for SLFN12 and SLFN13, which are 10 and approximately 11 nucleotides from the 3’ end, respectively (Figure 16B) [22, 46]. Thus, these Slfn proteins may share a common tRNA binding and cleavage mode. Interestingly, cleavage of tRNA^{Ser} by an N-terminal rat Slfn13 construct yields a secondary cleavage product that corresponds to a cleavage event within the tRNA T-loop [22]. This could indicate that some Slfn proteins have the ability to cleave tRNA at both sites or that the full-length protein is required for regulation of the cleavage reaction.

The trans-complementation assay indicates that dimerization of SLFN11 stimulates its endonuclease activity [324]. Together with the point mutants that disrupt the dimer interface, this suggests that dimeric SLFN11 is the nuclease active species and that a single active site is sufficient for tRNA cleavage. In line with that, molecular glue and PDE3A-mediated dimerization of SLFN12 has been shown to stimulate its RNase activity [26]. Of note, mutation of dimer interface I between the SLFN11 helicase domains does not abolish nuclease activity and SLFN12 still shows weak nuclease activity without molecular glue and PDE3A present [26, 324]. While this allows the interpretation that these Slfn proteins could be nuclease active in the monomeric state, it may as well mean that under those conditions the monomer-dimer equilibrium is shifted towards the monomeric state, but that dimerization is still required for tRNA cleavage. Therefore, we suggest that dimerization of Slfn proteins is essential for substrate recognition, binding, positioning and cleavage.

It has been shown that SLFN11 restricts various viruses such as HIV-1 in a codon usage dependent manner and that the double mutant E209A/E214A abrogates its antiviral activity [6, 70]. While SLFN11 cleaves multiple type II tRNAs, downregulation of tRNA-Leu-TAA appears to be responsible for the translational inhibition of Leu-UUA rich transcripts [45]. On average, the Leu-UUA codon usage frequency (Leu-UUA/all Leu codons) is about 8% in the human genome, but only about 2% in the 24 most highly expressed human proteins [45]. In contrast, the HIV-1 genome displays an average Leu-UUA codon usage of approximately 30%, making HIV-1 transcripts more susceptible to translational inhibition by tRNA-Leu-TAA depletion [45]. Interestingly, many factors that are involved in DNA damage response signaling also display an unusually high Leu-UUA codon usage (ATR: 21%, ATM: 23%) [45]. In line with that, it was shown that DNA-damage induced cell death relies on cleavage of tRNA-Leu-TAA by SLFN11, which inhibits the translation of ATM and ATR [45]. Similarly, SLFN12 also inhibits HIV-1 replication and translation of Leu-UUA codon rich transcripts [46, 328]. In the context of molecular glue-mediated PDE3A complex formation, SLFN12 was shown to cleave tRNA-Leu-TAA, which leads to ribosome pausing at Leu-UUA codons, global inhibition of protein synthesis and cell death [46]. Thus, we hypothesize related molecular mechanisms for the antiviral activity of SLFN11 and SLFN12. Interferon stimulated upregulation of these Slfns in response to viral infection may lead to cleavage of tRNA-Leu-TAA, which inhibits the translation of Leu-UUA rich viral and host transcripts. This raises the question of why some viruses employ a different codon usage compared to their hosts since it opens a point of attack for antiviral restriction factors and seems to impair viral replication. While this question is not entirely answered yet, there is evidence that viral codon usage is more similar to that of symptomatic hosts than to that of natural asymptomatic hosts [329]. As the survival of the virus relies on long-term coexistence with its natural host, a good fitness of the host might be essential for the long-term survival of the virus [329]. Thus, the viral codon usage bias has to be large enough to ensure sufficient fitness of the natural host. Surprisingly, SLFN11 and SLFN12 RNase activity seems not only stimulated by viral infections but also by DDA-induced replication stress in the case of SLFN11 and molecular glue-mediated PDE3A complex formation in the case of SLFN12 [26, 45, 111].

One interesting open question is where in the cell SLFN11 exerts its endonuclease activity. SLFN11 harbors a nuclear localization signal and while most studies have mapped it to the nucleus [21, 71], there is evidence that SLFN11 can also be found in the cytoplasm [205]. tRNA is generated in the nucleus as pre-tRNA but also mature tRNAs can enter the nucleus via retrograde nuclear import, e.g. in response to oxidative stress [330]. However, the majority of tRNA is found in the cytoplasm, where it is needed for translation of mRNA [331]. Thus, it is possible that SLFN11 nuclease activity is limited to the nucleus, even though cleavage of cytoplasmic tRNAs may be a more efficient way to deplete the tRNA-Leu-TAA pool. As SLFN12 is located in the cytoplasm and also cleaves tRNA-Leu-TAA, it seems possible that SLFN11 and SLFN12 work simultaneously upon viral infections to quickly deplete tRNA-Leu-TAA in the nucleus as well as in the cytoplasm [328]. Whether Slfn-generated tRNA cleavage products themselves serve further biological functions, as shown for other tRNA-derived small RNAs, is yet to be investigated [332].

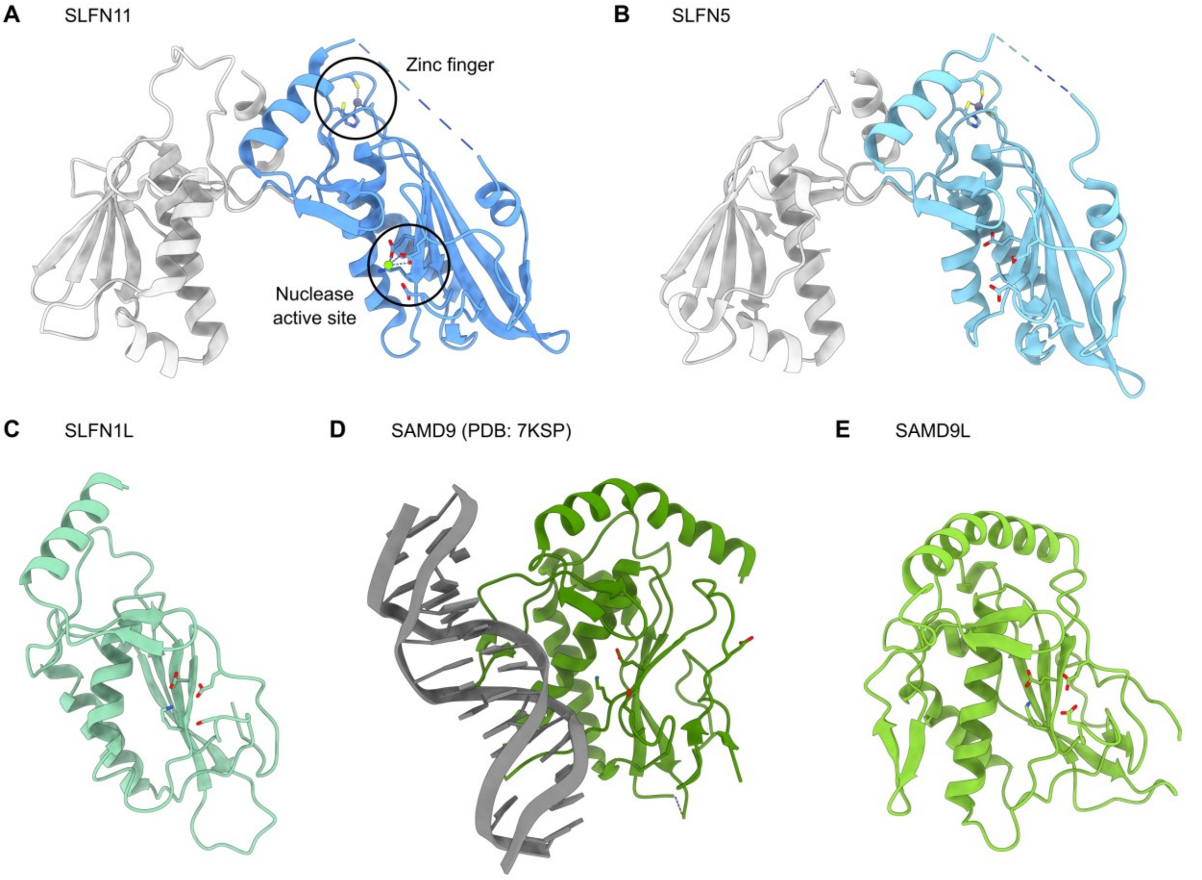


Figure 17: Human proteins with structural homology to the Slfn core domain. (A) Ribbon representation of the SLFN11 core domain (PDB: 7ZEL) [324]. The first bridging domain and the C-lobe are depicted in blue. Residues of the zinc finger and the nuclease active site are depicted as sticks. (B) Ribbon representation of the SLFN5 core domain (PDB: 7Q3Z) [322]. The first bridging domain and the C-lobe are depicted in blue. Residues of the zinc finger and the putative nuclease active site are depicted as sticks. (C) Ribbon representation of the conserved part of SLFNL1 as predicted by AlphaFold [323, 333]. Residues of the putative nuclease active site are depicted as sticks. (D) Structure of the SAMD9 DNA-binding domain (green) in complex with DNA (grey) (PDB: 7KSP) [334]. Residues of the putative nuclease active site are depicted as sticks. (E) Ribbon representation of the conserved part of SAMD9L as predicted by AlphaFold [323, 333]. Residues of the putative nuclease active site are depicted as sticks.

While the full Slfn core domain is only conserved among Slfn family members, several proteins share a domain with similarity to the Slfn core C-lobe, the so-called Slfn box (Figure 17). Structural homology search of the C-lobe of the SLFN11 core domain against the human subset of the AlphaFold Protein Structure Database [333, 335] yields several hits with high similarity. The uncharacterized Schlafen-like protein 1 (SLFNL1) harbors a domain with similarity to the Slfn core C-lobe and bridging domain (Figure 17C). A recent study showed that *C. elegans* Schlafen-like proteins form a trimeric complex with RNA endonuclease activity [336]. Human Sterile alpha motif domain-containing protein 9 (SAMD9) and SAMD9-like (SAMD9L) that act as restriction factors of poxviruses, also harbor a domain with similarity to the Slfn core C-lobe and bridging domain (Figure 17D and E) [337, 338]. In addition, SAMD9L restricts HIV-1 by targeting translation of viral proteins [337]. Mutation of the potential nuclease active site residues abrogates its ability to restrict HIV-1, suggesting that SAMD9L may also have nuclease activity [337]. This indicates a degree of functional similarity between Slfn box containing proteins.

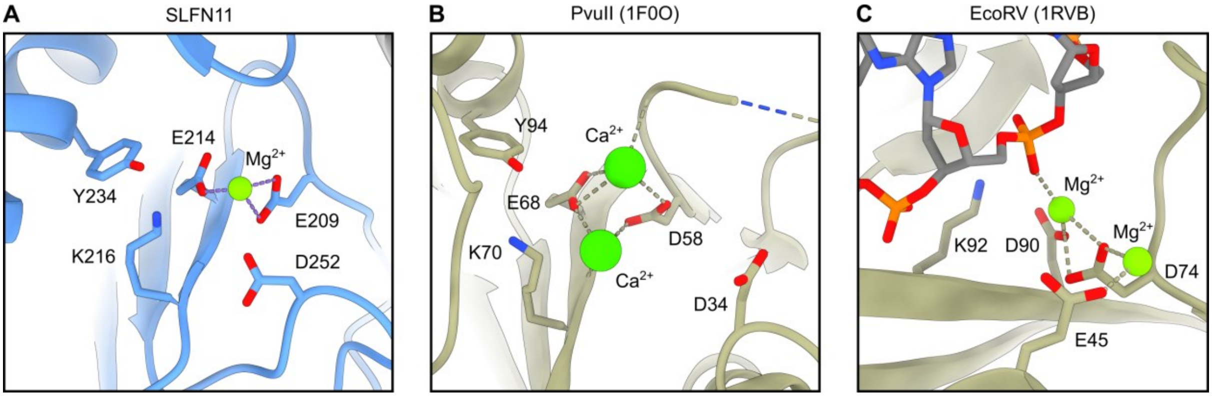


Figure 18: Structural comparison of the nuclease active sites of SLFN11 and PD…D/ExK restriction enzymes. (A) Detailed view of the SLFN11 nuclease active site (PDB: 7ZEL) [324]. Key residues and the coordinated Mg^{2+}-ion are depicted and labeled. (B) Detailed view of the PvuII restriction endonuclease active site (PDB: 1F0O) [339]. Key residues and the coordinated Ca^{2+}-ions are depicted and labeled. For clarity, the bound DNA is not visualized. (C) Detailed view of the endonuclease active site of EcoRV in complex with DNA (PDB: 1RVB) [340]. Key residues and the coordinated Mg^{2+}-ions are depicted and labeled.

While the number of proteins with Slfn box-related domains is rather small, we noticed striking similarities between the nuclease active sites of SLFN11 and type II restriction endonucleases of the PD…D/ExK family (Figure 18). Type II restriction endonucleases generally form homodimers that cleave dsDNA at palindromic sequences and require Mg^{2+} or Mn^{2+} as cofactor [341]. Due to these characteristics, enzymes of this class are an important tool in molecular biology [342]. PD…D/ExK family members such as PvuII or EcoRV have two conserved acidic residues that are involved in metal ion coordination and an essential lysine residue that might position and activate the nucleophilic water (Figure 18B and C) [341]. The active site of SLFN11 harbors a PE…ExK motif with similar geometry to the active site of PD…D/ExK enzymes (Figure 18A and B). Thus, the catalytic mechanism of Slfn proteins and PD…D/ExK enzymes may be related. However, the catalytic mechanism of type II restriction endonucleases is not well understood but several models were hypothesized [341]. High-resolution crystal structures of type II restriction endonucleases show that the two acidic active site residues coordinate the metal ions (Figure 18B and C) [339]. While type II restriction endonucleases generate 5′-phosphoryl and 3′-hydroxyl terminal ends, different structures found either one or two metal ions bound at the active site, suggesting one- or two-metal ion catalytic mechanisms [339, 341, 343]. The conserved lysine residue might function as a general base; however, due to its high pKa-value and its predicted protonated state, it is suggested to rather position and activate a water molecule for the nucleophilic attack on the phosphate backbone [341]. Interestingly, the similarities do not end at the three active site residues, but several PD…D/ExK family members such as EcoRV harbor a third acidic residue in a similar position to D252 in SLFN11 (Figure 18). Similarly to D252 in SLFN11, E45 in EcoRV is not essential for its nuclease activity [324, 344]. Furthermore, PvuII harbors a tyrosine residue at a similar position to Y234 in SLFN11 (Figure 18A and B). In PvuII, this tyrosine is suggested to contribute to Mg^{2+} binding and a mutant of it shows reduced nuclease activity [345]. While SLFN11 Y234A still shows RNase activity, it is worth considering that we did not measure cleavage kinetics but a single time point [324].

Taken together, we suggest that the endoribonuclease activity of Slfn proteins such as SLFN11 depend on the conserved E…ExK motif as well as on divalent metal cofactors, while loss of one or more residues of the essential E…ExK motif results in RNase deficiency as seen for SLFN5 [322].

3.3. Insights into the C-terminal helicase domain of subgroup III Slfn members

All subgroup III Slfn proteins harbor a C-terminal domain with similarity to SF1 DNA/RNA helicases [2, 4]. While this classification is based on sequence alignment and conservation of essential helicase motifs, no structural or comprehensive biochemical characterization of the Slfn helicase domain has been available.

The structures of SLFN5 and SLFN11 highlight similarities but also surprising differences between their Slfn helicase domains. The helicase C-lobe is not visible in the cryo-EM reconstruction of monomeric SLFN5, indicating conformational flexibility [322]. In contrast, both lobes of the SLFN11 helicase domain are resolved in the reconstruction of the SLFN11 dimer [324]. The ATPase consist of two RecA-like lobes and comprises the characteristic helicase motifs, including Walker A and B motifs as part of the N-lobe [56, 346]. The conservation of the essential motifs suggests that the SLFN5 and SLFN11 helicase domains possess the capability to bind and hydrolyze ATP. The finding that a SLFN11 Walker B mutant fails to sensitize cells to DNA damaging agents further supports the assumption of a functional integrity of this domain [71]. Interestingly, SLFN11 does not bind ATP in vitro while SLFN5 shows nucleotide binding [322, 324]. A comparison of the structures of SLFN5 and SLFN11 with a nucleotide bound structure of the SF1 DNA/RNA helicase DNA2 indicates that the conformation of the ID-region might regulate ATP binding [324]. The ID-helix in SLFN11 sterically blocks the ATP binding site, while the different conformation of the ID-region in SLFN5 leaves space for a nucleotide to bind. Furthermore, the Q-motif with its conserved glutamine residue that is generally involved in ATP binding [347], is part of the ID-helix in SLFN11 and the glutamine residue points away from the nucleotide-binding pocket [324]. In contrast, SLFN5 shows the typical conformation of a functional Q-motif with the glutamine residue pointing towards the nucleotide-binding site [322]. This indicates that the SLFN11 ATPase is locked in an autoinhibited conformation that is unable to bind ATP [324]. In line with that, we could not observe ATP hydrolysis by SLFN11 in vitro even with different DNA/RNA substrates or RPA present [324]. However, we also could not detect ATP hydrolysis by SLFN5 despite its ability to bind ATP [322]. Activation of the SLFN5 and SLFN11 ATPases may therefore require different steps. While SLFN5 can bind ATP, the absence of ATPase activity indicates that an activating factor, such as a specific substrate, binding partner or modification, is missing. SLFN11, on the other hand, requires an initial conformational rearrangement to overcome the autoinhibition and enable ATP binding. As a second step, it might also require further factors to stimulate its ATPase activity. The autoinhibited conformation of SLFN11 might be stabilized by dimerization but does not require dimerization as the ATPase domain of monomeric SLFN11 adopts the same autoinhibited conformation as seen in the dimer structure [324]. Instead, recent data suggests that ATP binding by SLFN11 is regulated through a phosphorylation-dependent conformational change in the SLFN11 helicase domain, which will be discussed later on [324]. While the ATPase domains of most subgroup III Slfn proteins are highly conserved, SLFN14 seems to be an exception as it shows multiple changes in essential helicase motifs, including the Walker A motif [322]. Thus, it is possible that SLFN14 lost its ability to bind and hydrolyze ATP and that the residual helicase-like domain serves alternative functions, such as stabilization of a SLFN14 dimer or as a binding site for other factors.

SF1 helicases are involved in different aspects of DNA and RNA metabolism and couple ATP hydrolysis to various processes such as translocation along single- or double-stranded DNA/RNA or unwinding of double-stranded DNA/RNA [346]. As subgroup III Slfn proteins lack insertion domains within the helicase domain, such as a pin motif or additional dsDNA binding elements, that are characteristic for strand-opening SF1 helicases [346], it appears unlikely that Slfn proteins are involved in the unwinding of double-stranded DNA or RNA [324]. To gain insights into the potential function of the SLFN5 and SLFN11 helicase domains, we investigated their substrate specificity. SLFN5 binds double-stranded DNA in vitro [322], which is in line with its roles as a transcriptional repressor [61, 62] and antiviral factor that binds viral or proviral DNA [9, 57]. Interestingly, the transcriptional repression of ZEB1 by SLFN5 depends on its C-terminal helicase domain [61], while the restriction of HIV-1 or HSV-1 depends on the N-terminal SLFN5 core domain [9, 57]. This indicates that SLFN5 may have two different DNA binding sites, one in the helicase and one in the Slfn core domain. Indeed, analysis of the SLFN5 core domain (SLFN5$^{1-336}$) revealed that it preferentially binds dsDNA over ssDNA or tRNA [322]. Full-length SLFN5 seems to have a higher affinity to dsDNA than the SLFN5 core domain alone [322]. Thus, SLFN5 may either bind dsDNA at two independent binding sites or via an extended binding mode that comprises both sites. To solve this conundrum, structural and mutational studies are needed.

As SLFN11 is recruited to stalled replication forks upon treatment with DDAs, we investigated its interaction with DNA [71, 124]. SLFN11 binds single-stranded DNA with high affinity ($K_{d,app}$ ~ 30 nM) but does not bind to double-stranded DNA [324]. This finding adds to previous reports that showed that SLFN11 colocalizes with RPA at sites of DNA damage and at stalled replication forks [71, 124]. SLFN11 was suggested to be recruited to ssDNA via a direct interaction between its C-terminal helicase domain with RPA1 [124]. However, as we do not see stable complex formation between SLFN11 and RPA in vitro (data not shown), the described colocalization may also be a result of independent binding of RPA and SLFN11 to ssDNA. It has been reported that knockdown of RPA1, MRE11 or deletion of the SLFN11 helicase C-lobe result in the absence of SLFN11 foci in CPT treated cells [124]. While the deletion of the SLFN11 helicase C-lobe likely disrupts ssDNA binding, knockdown of MRE11 likely results in reduced ssDNA formation. However, as a knockdown of RPA1 also impairs SLFN11 foci formation, an interaction between RPA and SLFN11 appears plausible [124]. This interaction might be rather weak or dependent on post-translational modifications, as we do not see stable complex formation in vitro. It has further been suggested that SLFN11 destabilizes RPA-ssDNA complexes and inhibits the accumulation of RAD51 at stalled replication forks, promoting stalled fork degradation [124, 166]. SLFN11-mediated stalled fork degradation in FA cells depends on the SLFN11 ATPase activity that affects RAD51 recruitment to reversed forks [166]. However, whether SLFN11 actually destabilizes the RPA-ssDNA complex is still an open question, as the observed reduction in RPA binding could also be a result of the replication fork block by SLFN11 [71, 124]. The cryo-EM reconstruction of ssDNA bound SLFN11 shows that each protomer of the SLFN11 dimer binds a stretch of five nucleotides between the helicase N- and C-lobes [324]. The overall conformation of the ssDNA-bound SLFN11 dimer is very similar to that of the SLFN11 apoenzyme, indicating that the apoenzyme is already poised for ssDNA binding. This comes as a surprise, as the ATPase adopts an autoinhibited state that is unable to bind ATP [324]. Thus, the question arises whether and how DNA binding and ATP hydrolysis may be coupled in the active enzyme. Interestingly, a SLFN11 Walker B mutant colocalizes with replication foci but fails to block replication fork progression [71]. As the ssDNA binding function of SLFN11 was previously unknown, it would be of great interest to see whether the DNA binding deficient SLFN11 mutant still colocalizes with replication foci and whether it retained the ability to block stressed replication forks or inhibit viral infections. SLFN11 interacts mostly with the phosphate backbone of ssDNA, which would suggest low sequence specificity [324]. However, we observed small differences in sequence specificity for different ssDNA substrates (data not shown), suggesting that the DNA sequence might contribute to governing SLFN11 binding to a certain extend. Overall, low sequence specificity seems plausible as it may allow SLFN11 to act at stalled replication forks, independently of their position within the genome. The anisotropy data indicate cooperative binding of ssDNA, which is in line with the observation that both SLFN11 protomers are bound to ssDNA in the cryo-EM reconstruction [324]. Furthermore, mass photometry data indicate that ssDNA binding stabilizes the SLFN11 dimer, as shown by a shift of the monomer-dimer equilibrium towards the dimer state [324]. As the SLFN11 dimer is poised for tRNA binding and represents the nuclease active state, ssDNA binding may stimulate the SLFN11 RNase activity by stabilizing the dimeric state. This could imply that in the absence of a stimulus, SLFN11 may adopt a monomeric conformation in vivo that would be nuclease deficient. This hypothesis is in line with a report showing that DDA treatment induces SLFN11-mediated cleavage of type II tRNAs, eventually leading to cell death [45]. The SLFN11 dimer interface mutants would be an excellent tool to investigate whether dimerization is essential for replication fork blockage by SLFN11.

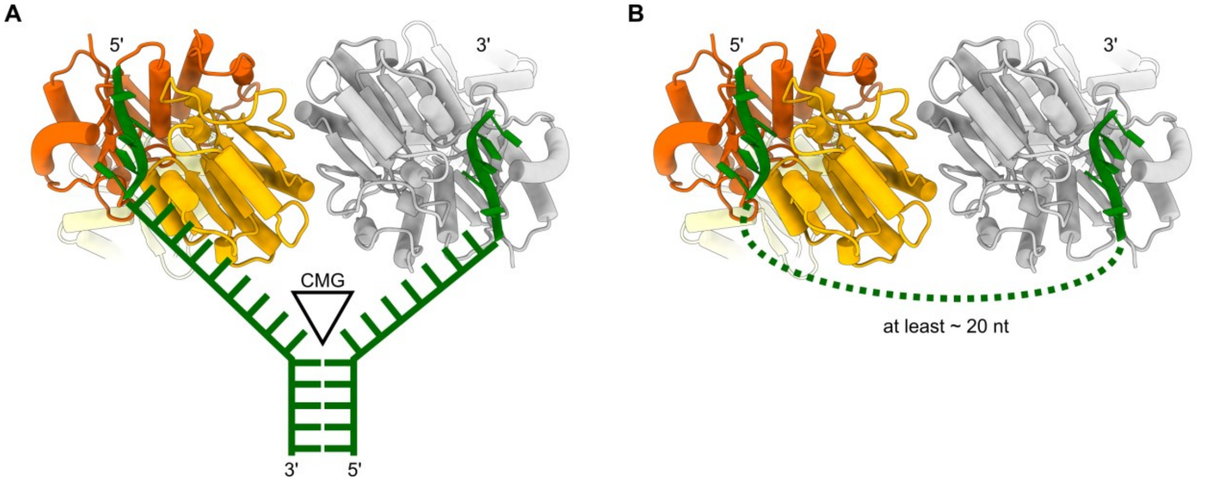


Figure 19: Potential models for ssDNA binding by SLFN11. (A) Schematic representation of a SLFN11 dimer (orange and grey; PDB: 7ZES) [324] binding to both DNA strands (green) at a replication fork. The CMG helicase that unwinds the DNA double-strand is represented as triangle. (B) Schematic representation of a SLFN11 dimer (orange and grey; PDB: 7ZES) [324] binding to one strand of ssDNA (green). At least 20 nt DNA are required to span the distance between the SLFN11 DNA binding sites, resulting in a minimal theoretical ssDNA footprint of 30 nt.

The two DNA fragments bound to the SLFN11 dimer have opposing 5’ to 3’ directions [324]. Assuming that both protomers bind ssDNA, different binding modes are possible (Figure 19). Dimeric SLFN11 could bind a continuous stretch of ssDNA, which, based on the structure, would translate to a ssDNA footprint of at least 30 nucleotides (nt) (2 x 5 nt bound + 20 nt linker). Alternatively, the SLFN11 dimer could simultaneously bind to both single-strands of a stalled replication fork and may thereby block fork progression [324]. While such a static model of a physical replication block seems appealing, data shows that the ATPase activity of SLFN11 is essential for replication fork blockage [71]. Thus, the process likely involves conformational changes that would allow SLFN11 to exit the autoinhibited state. Future studies may focus on the identification of such activating factors, which could comprise certain DNA or RNA structures, post-translational modifications or novel binding partners.

3.4. Phosphorylation induces a conformational change that regulates SLFN11 activity

Recently, three phosphorylation sites were identified in SLFN11 that regulate its RNase activity and its ability to sensitize cells to DDA treatment [111]. Two phosphorylation sites are located in the Slfn core domain (S219, T230) in close proximity to the nuclease active site, whereas the third site is located in the helicase domain (S753). Mutagenesis studies revealed that phosphomimetic mutants (Asp) of these sites render SLFN11 inactive with respect to type II tRNA cleavage and sensitizing cells to DDA treatment, while Ala mutants showed wild type-like activity, suggesting that dephosphorylation of SLFN11 is required to become functionally active [111]. Protein phosphatase 1 catalytic subunit γ (PPP1CC) was identified to be responsible for SLFN11 dephosphorylation and consequently for its activation [111]. Interestingly, CPT treatment induces a reduction in total SLFN11 phosphorylation, while showing no significant reduction in case of the SLFN11 S753D mutant, suggesting that the phosphorylation of S753 is important for the regulation of SLFN11 activity [111]. Tight regulation of the SLFN11 enzymatic activity might be important to avoid undesired SLFN11-mediated effects under stress-free conditions. Uncontrolled cleavage of type II tRNAs, for example, could negatively affect translation of endogenous transcripts, a fitness penalty that might be the lesser of two evils in case of a viral infection, but may be avoided in uninfected cells by phosphorylation-dependent inhibition of SLFN11 RNase activity. Similarly, uncontrolled blockage of stalled replication forks by SLFN11 might negatively affect cell survival by preventing replication restart after ATR/CHK1-mediated transient replication stop and DNA damage repair. The phosphorylation sites that are located in close proximity to the SLFN11 nuclease active site (S219, T230) are likely to interfere with tRNA binding, possibly due to repulsive forces of the negatively charged phosphate groups of the tRNA backbone and the phosphorylated S219 and/or T230 [324]. Due to the tRNA binding mode by dimeric SLFN11, these phosphorylations might affect tRNA interactions in both protomers, which could amplify their effectiveness. How a phosphorylation within the SLFN11 helicase domain (S753) would affect tRNA cleavage appears more enigmatic. Located in the region that connects the two ATPase lobes, S753 is far away from the tRNA binding site [324]. On the other hand, it is located in close proximity to the ssDNA binding groove but does not interact with DNA in the ssDNA-bound SLFN11 structure [324].

Interestingly, the phosphomimetic mutant S753D is ssDNA binding deficient, binds ATP and appears as a monomer in solution (Figure 20) [324]. This is in contrast to wild-type SLFN11 that binds ssDNA and adopts an autoinhibited dimeric conformation that is ATP binding deficient (Figure 20) [324]. Initial structural investigation of ATP-bound SLFN11^{S753D} revealed that it adopts a monomeric conformation, which is highly similar to that of SLFN5 (Figure 20D). The SLFN11^{S753D} reconstruction shows a conformational change of the ID-region that results in a rotation of the helicase domain by approximately 180°, freeing space for a nucleotide to bind. Both ATPase lobes are visible and adopt a closed, nucleotide-bound conformation. In contrast to the wild-type SLFN11 monomer reconstruction [324], both lobes of the Slfn core domain are resolved. However, whether the SLFN11 S753D mutation accurately reflects the S753 phosphorylated state (SLFN11^{S753p}) requires further investigation. Assuming this is the case, the SLFN11^{S753D} conformation that allows for ATP binding might represent an ATPase proficient state. However, whether SLFN11^{S753D} itself might be ATPase active or whether

SLFN11^{WT} can adopt a similar conformation, e.g. through engagement with a binding partner, is yet to be elucidated. As SLFN11^{S753D} does not bind to ssDNA, it would be a good tool to test whether the recruitment of SLFN11 to stalled replication forks depends on ssDNA or on other factors, such as RPA [324]. The conformation of SLFN11^{S753D} does not exclude the possibility for dimerization, as docking of the SLFN11^{S753D} conformation onto the dimer structure does not cause significant clashes. However, the loss of interface I as a result of the rotated helicase domains destabilizes the dimer, explaining why SLFN11^{S753D} adopts a monomeric state in solution (Figure 20E). This would also explain the impaired tRNA cleavage activity by SLFN11^{S753D} [111], as dimerization promotes efficient tRNA cleavage [324].

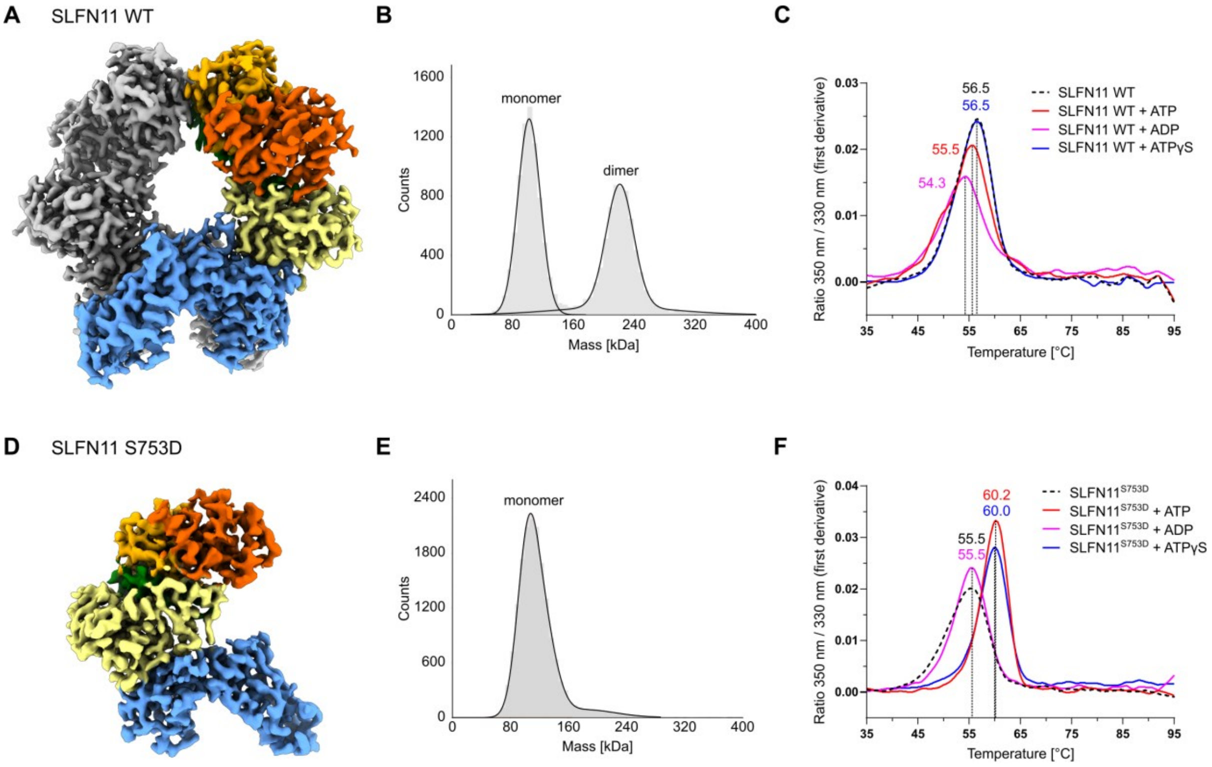


Figure 20: Comparison of SLFN11^{WT} with the phosphomimetic mutant SLFN11^{S753D}. (A) Cryo-EM reconstruction of the wild-type SLFN11 dimer (EMDB: 14690) [324]. The Slfn core domain (blue), linker domain (yellow), ID-region (green), and helicase domain (orange) are color-coded. The second SLFN11 protomer is colored in grey. (B) Mass photometry analysis of the monomer-dimer distribution of wild-type SLFN11 at 80 mM NaCl. Figure modified from [324]. (C) NanoDSF analysis of wild-type SLFN11 in presence of different nucleotides (ADP: pink, ATP: red, ATPγS: blue) or without nucleotides present (black). The inflection temperatures are indicated. Figure modified from [324]. (D) Cryo-EM reconstruction of the SLFN11^{S753D} phosphomimetic mutant. The Slfn core domain (blue), linker domain (yellow), ID-region (green), and helicase domain (orange) are labeled and color-coded. (E) Mass photometry analysis of the monomer-dimer distribution of SLFN11^{S753D} at 80 mM NaCl. (C) NanoDSF analysis of SLFN11^{S753D} in presence of different nucleotides (ADP: pink, ATP: red, ATPγS: blue) or without nucleotides present (black). The inflection temperatures are indicated.

Thus, a possible regulatory model could look as follows (Figure 21): SLFN11 might be phosphorylated at S753 in a stress-free state [111], preventing dimerization and efficient tRNA cleavage, as well as ssDNA binding and a subsequent block of stalled replication forks. The normal level of occasionally occurring stalled replication forks and DNA damage might not be sufficient to activate SLFN11 as it would otherwise impair DNA replication. However, upon replication stress induced by DDAs, a critical stress-level might be reached that results in the dephosphorylation of SLFN11^{S753p} by PPP1CC [111], releasing the phosphorylation-mediated conformational inhibition. This may recruit SLFN11 to stalled replication forks via its high affinity for ssDNA. Additional factors could activate its ATPase that is essential for a SLFN11-mediated irreversible block of stalled replication forks [71]. Furthermore, the high local concentration of SLFN11 may promote SLFN11 dimerization, which activates its tRNase, resulting in translational inhibition of certain DNA damage response and repair factors such as ATM and ATR [45].

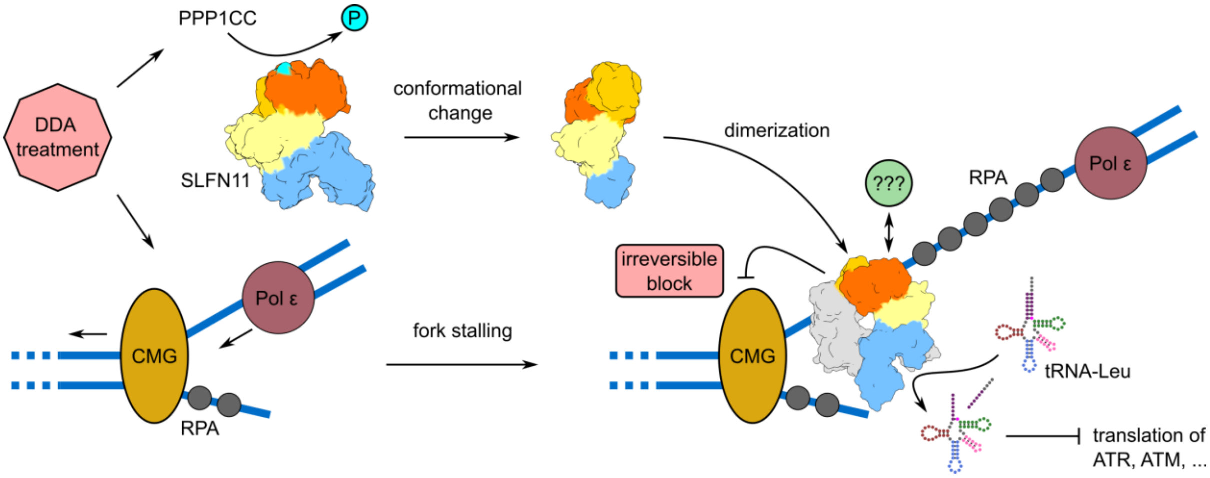


Figure 21: Model for the activation of SLFN11 by DDAs. Under stress-free conditions, ssDNA binding, dimerization, and nuclease activity of SLFN11 (Slfn core domain: blue, linker domain: yellow, helicase domain: orange) is inhibited by phosphorylation of S753. Treatment of cells with certain DDAs, such as CPT, induces replication stress and fork stalling [71]. It also leads to the dephosphorylation of SLFN11^{S753p} by protein phosphatase 1 (PPP1CC) [111]. The dephosphorylation induces a conformational rearrangement of the helicase domain and ID-region, allowing SLFN11 to bind ssDNA and dimerize. SLFN11 is recruited to extended stretches of RPA-coated ssDNA at stalled replication forks [71, 124] where its nuclease activity is stimulated by dimerization. Cleavage of type II tRNAs, such as tRNA-Leu-TAA, by SLFN11 inhibits the translation of several proteins involved in DNA damage response, such as ATR and ATM [45]. Furthermore, SLFN11 interacts with the CMG helicase and blocks the replication fork in an ATPase-dependent manner [71]. The ATPase activity might be induced by the interaction with additional factors (green) and leads to an irreversible replication block and eventually to cell death.

ATPase-dependent replication block and tRNA cleavage both contribute to sensitizing cells to DDA treatment; however, the contribution of the RNase activity is not well understood [45, 71]. As suggested by Malone et al. [111], dephosphorylation of SLFN11 may occur in a sequential manner, starting with the dephosphorylation of SLFN11^{S753p}, allowing for a scenario where phosphorylated S219 and T230 inhibits the RNase activity of ATPase active SLFN11. Differential phosphorylation and dephosphorylation of SLFN11 could therefore govern a separation of function, as phosphorylation of S753 might regulate its replication checkpoint activity, while phosphorylation of S219 and T230 may regulate its antiviral activity. However, the functional connection of the SLFN11 nuclease and helicase domains is not well understood and requires further research.

A search for the S753 phosphorylation motif against a substrate specificity atlas of the human Ser/Thr kinome [348] ranks CHK1 as one of the kinases most likely to be involved in S753 phosphorylation. This would indicate a regulatory role of the ATR/CHK1 pathway on SLFN11 activity. For example, SLFN11 activity could be inhibited through the ATR/CHK1 pathway in response to low levels of replication stress in order to prevent irreversible fork stalling, whereas high levels of replication stress as induced by DDA treatment might activate dephosphorylation of SLFN11 to an extend that cannot be counteracted by CHK1.

Interestingly, S753 is not conserved between human subgroup III Slfn proteins and is only found in SLFN11 [324]. Thus, this specific mode of regulation by modulating the DNA binding affinity and the oligomeric state might be unique to SLFN11. However, it has also been reported that the RNase activity of SLFN12 is regulated by phosphorylation, potentially by regulating its interaction with PDE3A [55]. Furthermore, it has been suggested that SLFN5 plays a role in cell cycle progression and cellular proliferation and that this is regulated by CDK-mediated phosphorylation [349]. Taken together, phosphorylation appears to play an important role in the functional regulation of Slfn proteins but further research is required to get a better picture of the underlying regulatory network.

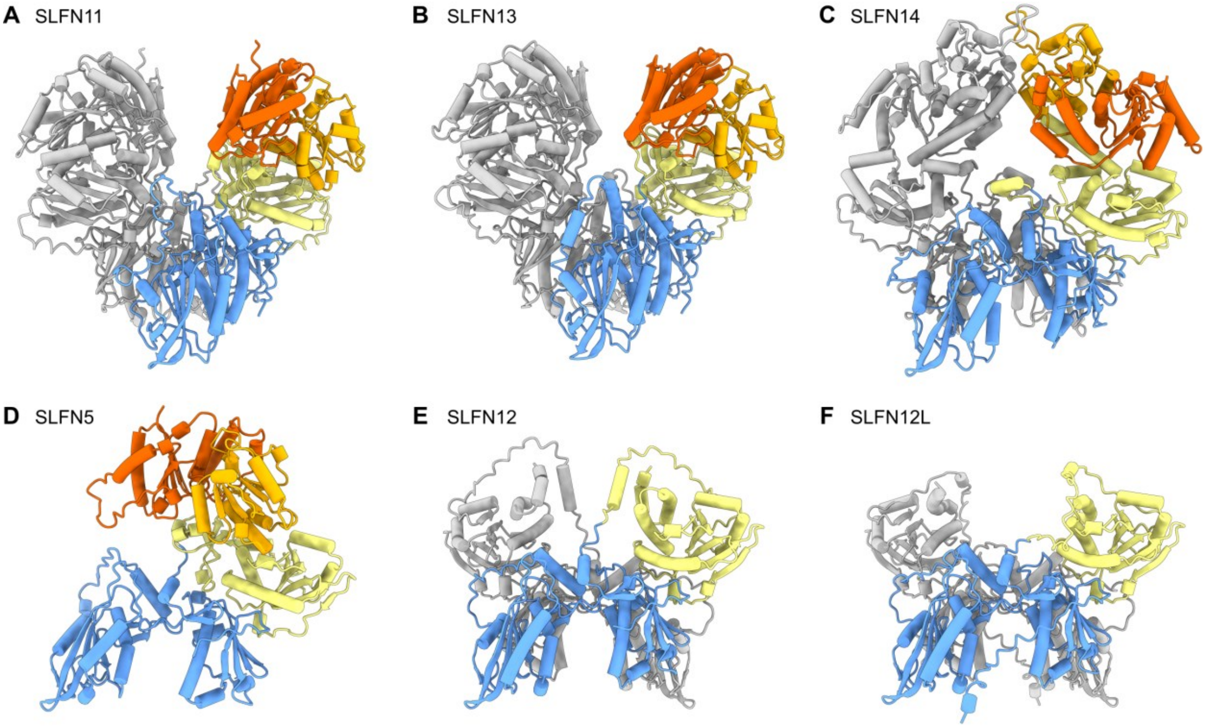


Figure 22: AlphaFold predictions of human Slfn protein dimers. Dimer models for all human full-length Slfn proteins were predicted using AlphaFold2-multimer [323, 325, 350]. The Slfn core domains (blue), linker domains (yellow), and helicase domains (orange) are color-coded. The second protomers are colored in grey. (A) SLFN11 prediction with dimer interface between Slfn core domains. (B) SLFN13 prediction with dimer interface between Slfn core domains. (C) SLFN14 prediction with dimer interfaces between Slfn core domains and helicase domains. (D) Monomeric SLFN5 as it was not predicted to form a dimer. (E) SLFN12 prediction with dimer interface between Slfn core domains. (F) SLFN12L prediction with dimer interface between Slfn core domains.

The relatively high degree of sequence conservation between human Slfn proteins may result in structural and functional similarities. AlphaFold predictions [323, 325] of human subgroup III proteins result in dimeric models for SLFN11, SLFN13 and SLFN14 but a monomeric model for SLFN5 (Figure 22). This is in line with the structures of monomeric SLFN5 and dimeric SLFN11 [322, 324]. Furthermore, this is in agreement with the biochemical data [322], as a loss of tRNase activity in SLFN5 might eliminate the need for dimerization, assuming that the reason for dimerization is to form a platform for tRNA binding. AlphaFold predicts the protomers of the SLFN11 and SLFN13 dimers in the ATP binding proficient conformation (Figure 22A and B), while it models the protomers of the SLFN14 dimer in the autoinhibited conformation (Figure 22C), similar to the cryo-EM structure of the SLFN11 dimer [323-325]. This is particularly striking as SLFN14 harbors an altered Walker A motif [322], which likely results in ATPase deficiency and the inability to bind nucleotides. As we have established that SLFN11 can adopt different conformations with variable biochemical properties [324], it seems plausible that other subgroup III Slfn members exhibit a similar conformational plasticity. Hence, the

SLFN11 structures hint at a general regulatory mechanism that might be conserved throughout subgroup III SLFN family members.

3.5. A conserved A-module anchor in INO80 and SWI/SNF family remodelers

DNA-dependent nuclear processes, such as transcription, DNA replication or DNA repair, are regulated by nucleosome positioning and composition, which depend on the activity of ATP-dependent chromatin remodelers [241, 252]. These enzymes share a conserved Snf2-type ATPase for DNA translocation and range from small single-subunit remodelers to large multi-subunit assemblies like the INO80 complex [241, 253]. As single-subunit remodelers are capable of catalyzing nucleosome sliding, the question arises why INO80 and other multi-subunit remodelers assemble into much larger and more complex molecular machines [241, 351]. Biochemical data suggests that the modular architecture of the INO80 complex contributes to the regulation of its chromatin remodeling activities by reading of chromatin-associated information, such as nucleosome composition, nucleosome density, presence of barrier factors, or DNA shape features [248, 250, 298, 309, 312, 313, 320]. The INO80 A-module processes such information, at least in parts through interaction with extranucleosomal DNA, and serves as an allosteric regulator of nucleosome remodeling [248, 311-313]. In this work, we employed a bottom-up approach to structurally investigate extranucleosomal DNA readout by the A-module alone or in context of the conserved A/C-module complex to gain insights into regulatory mechanisms underlying nucleosome remodeling by INO80.

The structural analysis of INO80 A-modules from three different species (*C. thermophilum*, *S. cerevisiae*, and *H. sapiens*) [352] adds to a previously published crystal structure of the conserved portion of the *S. cerevisiae* A-module (Arp8-actin-Arp4-Ino80^{HSA}) [311] by including full-length proteins and species-specific subunits. Yeast/fungal Ies4 and human YY1 display a conserved mode of interaction with the Arp4-actin pair via a β-hairpin (Figure 23A) [352]. As the hairpin harbors two conserved tryptophan residues that form a hydrophobic pocket for the interaction with actin, it has been termed "2W-hairpin" [352]. In line with the conservation of the Arp4-actin pair in INO80 family and SWI/SNF family members (or Arp7-Arp9 in *S. cerevisiae* SWI/SNF family remodelers) [265], the 2W-hairpin motif can be found in structures or structure predictions of A-module associated subunits of several multi-subunit chromatin remodelers, indicating that the 2W-hairpin is an evolutionarily conserved A-module anchor [352]. Structures of the *S. cerevisiae* SWI/SNF and RSC A-modules reveal that regulator of Ty1 transposition protein 102 (Rtt102) interacts with Arp7-Arp9 via a 2W-hairpin motif, as does SWR1-complex protein 4 (Swc4) with Arp4-actin in the *S. cerevisiae* NuA4 complex (Figure 23B and C) [317, 353, 354]. Furthermore, B-cell CLL/lymphoma 7 (BCL7), a subunit of the human SWI/SNF complexes BAF and PBAF, may bind the A-module via a 2W-hairpin, as structure predictions of BCL7A/B/C predict the presence of a 2W-hairpin [333, 352, 355]. Interestingly, different remodelers recruit different 2W-hairpin containing subunits, despite the conservation of the Arp4-actin pair, indicating that the HSA domain or additional subunits contribute to the correct assembly of the A-module. In addition, the expression pattern of the client proteins may determine the complex compositions in cases like the human SWI/SNF complexes that can incorporate either BCL7A, BCL7B or BCL7C [355].

As the human INO80 subunit YY1 is a transcription factor that binds specific DNA sequences, it has been suggested to recruit INO80 to YY1 binding sites, resulting in transcriptional activation [356]. However, YY1 also recruits repressive polycomb group complexes, indicating that it has opposing functions in a context dependent manner [357]. Interestingly, YY1 interacts with the polycomb group protein MBT domain-containing protein 1 (MBTD1) via the 2W-hairpin, suggesting that it cannot bind INO80 and MBTD1 at the same time [358]. Further 2W-hairpin containing proteins also possess DNA binding activity, suggesting that a function of these client proteins may be the recruitment of chromatin remodelers to specific locations within the genome. The SWR1 subunit Swc4, for example, preferentially binds AT-rich DNA and may therefore recruit SWR1 to its target genes [359]. Ies4 on the other hand lacks any known DNA binding domain but plays a role in the DNA damage checkpoint response [360]. Thus, 2W-hairpin containing client proteins may add specific functions or features to the conserved part of the A-module (Figure 23D).

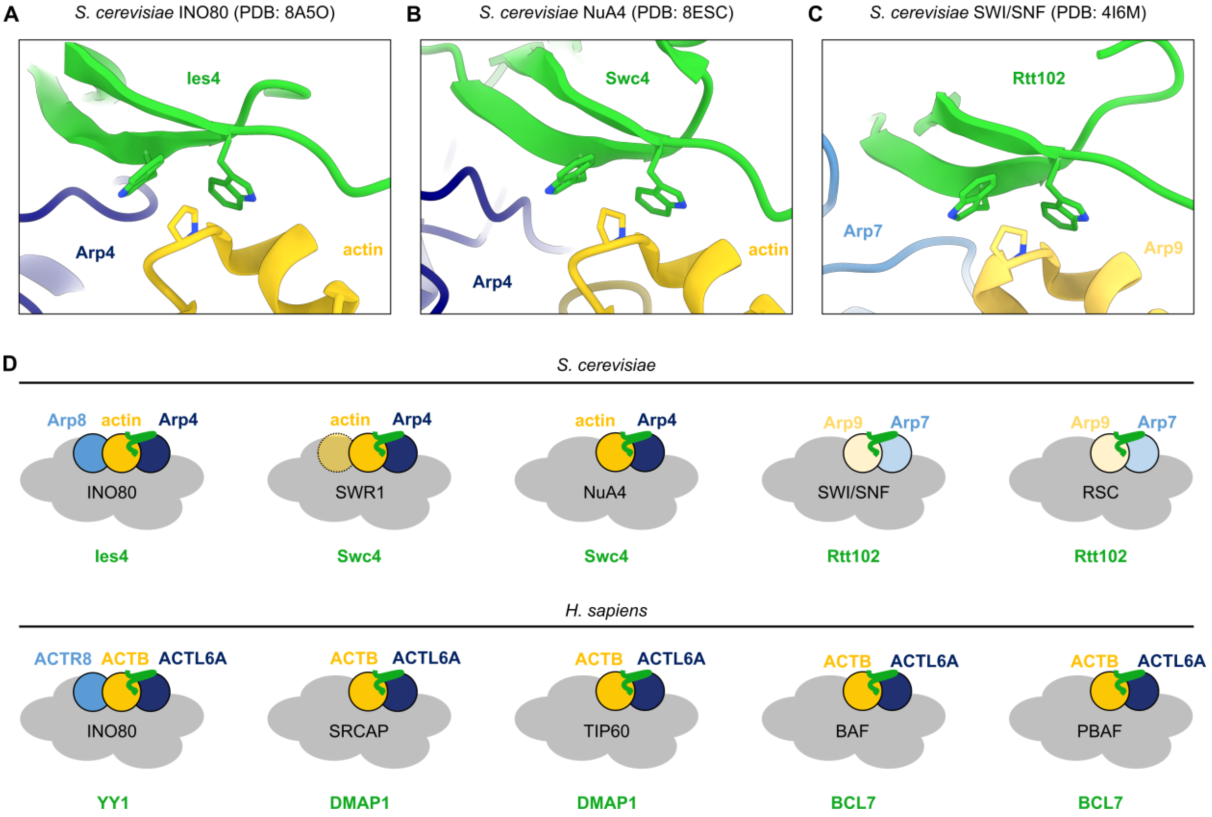


Figure 23: The 2W-hairpin is a conserved A-module anchor. (A) Detailed view of the Ies4-actin interface within the *S. cerevisiae* INO80 A-module (PDB: 8A5O) [352]. The conserved tryptophan residues of the 2W-hairpin and actin residue P367 are shown. Ies4 (lime green), Arp4 (dark blue) and actin (yellow) are labeled. (B) Detailed view of the Swc4-actin interface within the *S. cerevisiae* NuA4 A-module (PDB: 8ESC) [353]. The conserved tryptophan residues of the 2W-hairpin and actin residue P367 are shown. Swc4 (lime green), Arp4 (dark blue) and actin (yellow) are labeled. (C) Detailed view of the Rtt102-Arp9 interface within the *S. cerevisiae* SWI/SNF A-module (PDB: 4I6M) [317]. The conserved tryptophan residues of the 2W-hairpin and Arp9 residue P459 are shown. Rtt102 (lime green), Arp7 (light blue) and Arp9 (light yellow) are labeled. (D) Schematic overview of the conservation of the Arp4-actin heterodimer and 2W-hairpin subunits in *S. cerevisiae* and *H. sapiens* INO80 and SWI/SNF family chromatin remodelers. The respective 2W-hairpin subunits are indicated in green. Figure modified from [352].

Taf14, which was suggested to be an additional subunit of the *S. cerevisiae* A-module, is not resolved in the cryo-EM reconstruction [309, 352]. Thus, it is not a structural component of the A-module. Structure prediction of the Ino80-Taf14 interaction places Taf14 upstream of the A-module, flexibly connected by a 40 amino acid loop region [323, 350]. This could allow a conformation in which the A-module engages extranucleosomal DNA, while Taf14 binds to acetylated/crotonylated H3K9 (Figure 24) [361].

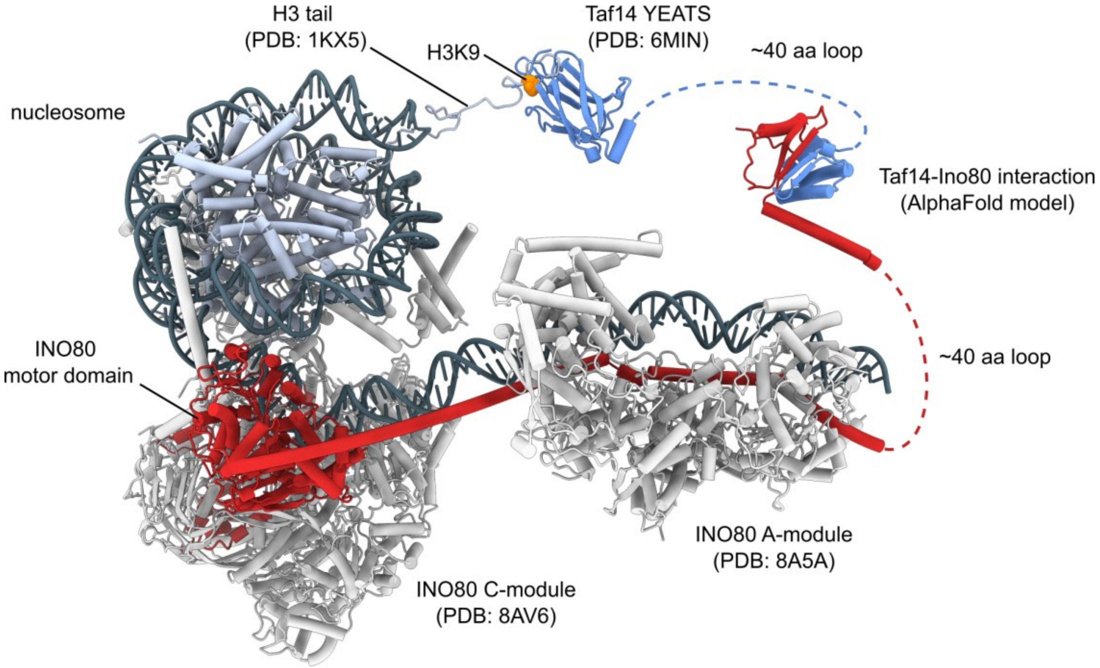


Figure 24: Structural model of the nucleosome-bound INO80 A/C-complex including Taf14. Composite model of the INO80-nucleosome complex with the A-module bound to extranucleosomal DNA. *S. cerevisiae* Taf14 (blue) is predicted to bind Ino80 (red) approximately 40 amino acids (aa) ahead of the A-module. The *S. cerevisiae* Taf14 YEATS domain (blue) recognizes acetylated/crotonylated H3K9 (orange). The model is composed of the *C. thermophilum* C-module (PDB: 8AV6) and the *S. cerevisiae* A-module (PDB: 8A5A) that are modeled into (EMDB: 15211) [352], an AlphaFold prediction of the *S. cerevisiae* Taf14-Ino80 interaction (Taf14 aa 176-244, Ino80 aa 332-406) [323], a structure of the *S. cerevisiae* Taf14 YEATS domain bound to crotonylated H3K9 (PDB: 6MIN) [362] and the H3 histone tail (PDB: 1KX5) [363]. The dotted lines indicate flexible loop region in Ino80 (red) and Taf14 (blue).

3.6. INO80-DNA interactions regulate nucleosome remodeling

The structure of the *C. thermophilum* A-module bound to DNA shows that the Ino80 HSA domain and the Arp8 N-terminus are involved in DNA binding [352]. This explains previous findings, which showed that the Ino80 HSA domain as well as the Arp8 N-terminus are crucial for extranucleosomal DNA binding and nucleosome sliding [311, 312]. The A-module binds a DNA segment of approximately 40 bp, which is in line with the observation that more than 40 bp of extranucleosomal DNA are required for efficient nucleosome sliding by INO80 [273, 311, 320, 352]. It also supports the idea that the INO80 A-module functions as a ruler element in nucleosome spacing and phasing, possibly by sensing neighboring nucleosomes or barrier factors [313]. The A-module bound DNA is markedly bend and the Arp8 hook element widens the minor groove upon binding [352]. This could enable DNA shape readout, which contributes to INO80-mediated +1 nucleosome positioning in *S. cerevisiae* [248]. Interestingly, mutations of the HSA domain have differential effects on the allosteric regulation of the nucleosome sliding reaction in vitro [352]. While nucleosome remodeling is inhibited by DNA-binding mutations of the C-terminal part of the HSA domain, it is stimulated by DNA-binding mutations in the N-terminal part of the HSA domain. Thus, the A-module can function as an allosteric activator or repressor of the sliding reaction, possibly by adopting different conformations or alternating DNA binding modes.

The structure of the *C. thermophilum* A/C-module bound to a nucleosome visualizes how the A-module binds to linker DNA in the context of the nucleosome-bound complex [352]. The HSA/post-HSA domain forms a continuous helix that connects the A-module to the N-lobe of the Ino80 motor domain (Figure 25A). This chemomechanical link might couple extranucleosomal DNA sensing by the A-module to productive nucleosome sliding. The interaction of the Ino80 post-HSA domain with protrusion I, a motif of the ATPase N-lobe, resembles the structural arrangement of the post-HSA domain and protrusion I in the RSC complex, which form a regulatory hub that controls ATPase activity and nucleosome sliding [318, 319, 352]. Of note, the extranucleosomal DNA-bound conformation is only visible in the apo state but not in the ADP·BeF_x state [352]. However, DNA-binding mutations of the A-module that impair nucleosome remodeling have only a moderate effect on the Ino80 ATPase rate, suggesting that the regulatory role of the A-module relies on the regulation of directional DNA-translocation rather than on the regulation of the ATP hydrolysis rate [352]. This would agree with the bulge propagation model, where a function of the INO80 A-module may be to prevent back slippage of the DNA loop, formed between the motor domain and the Arp5 counter grip [275, 291].

Arp5 is also involved in the regulation of the remodeling reaction by binding the NCP at SHL -2/-3 and by contacting the acidic patch with an Arp5 insertion domain, termed “grappler” [275]. The Arp5 grappler can adopt at least two defined conformations, a parallel and a cross grappler state. In the cross conformation, the Arp5 grappler contacts entry DNA opposite to the Ino80 motor domain. Mutations of two positively charged loops of the Arp5 grappler that might contribute to entry DNA interaction almost abolish nucleosome sliding and strongly reduce the INO80 ATPase rate, indicating a regulatory role of Arp5 on the ATPase activity [352].

Binding of the A-module to linker DNA bends the DNA along the Ino80 HSA domain and at the Ino80 motor domain [352]. Mapping of DNA shape features of INO80-positioned nucleosomes onto the structural model shows that extreme DNA propeller twist values correlate with the A-module binding site and the region between the motor domain and Arp5 [248]. Introduction of more rigid A/T-rich DNA cassettes, which resemble poly-A/T tracts that are found in yeast promoters, results in decreased nucleosome remodeling by INO80 [352, 364]. A/T-rich DNA most strikingly affects nucleosome sliding, when introduced at the Ino80 motor domain [352]. Thus, the motor domain itself might be involved in DNA shape readout and nucleosome positioning. Therefore, it could contribute to the establishment of NFRs at A/T-rich DNA regions and +1 nucleosome positioning, as nucleosome sliding away from A/T-rich sequences might be kinetically favored over nucleosome sliding into A/T-rich DNA [365]. The unusual position of the Ino80 motor domain at SHL -6, together with the extranucleosomal DNA bound A-module, might thus allow INO80 to monitor DNA shape features of incoming DNA [352].

Besides INO80, A-modules are also found in INO80 family and SWI/SNF family chromatin remodelers [265]. The conserved A-module core comprises an Arp4-actin heterodimer (or Arp7-Arp9 in *S. cerevisiae* SWI/SNF family remodelers) that assembles on the HSA domain [265]. As the Arp8 subunit is exclusive to INO80 and is critical for DNA interaction, a conserved DNA binding mode between the INO80 A-module and A-modules of other remodelers seems unlikely [312, 352]. Indeed, structurers of nucleosome-bound SWI/SNF family remodelers from yeast and human show a different conformation of the A-module (Figure 25B) [287-290]. While the motor domain of these remodelers engage the NCP at SHL -2, the A-module is folded back and contacts the NCP around SHL +6, which is close to the exit DNA. The interaction of the SWI/SNF HSA domain with the nucleosomal DNA rather resembles the binding mode of the Ies2 throttle helix in INO80 than that of the Ino80 HSA-DNA interaction (Figure 25) [275, 352]. In the structures of the SWI/SNF complexes, the post-HSA domain also interacts with protrusion I but does not form a continuous helix with the HSA domain (Figure 25B)

[287-290]. This indicates that the function of the A-module as an allosteric regulator is conserved, whereas its interaction with DNA is variable. Whether the INO80 A-module can adopt alternative conformations, more similar to the situation in SWI/SNF remodelers, or vice versa, requires further investigation.

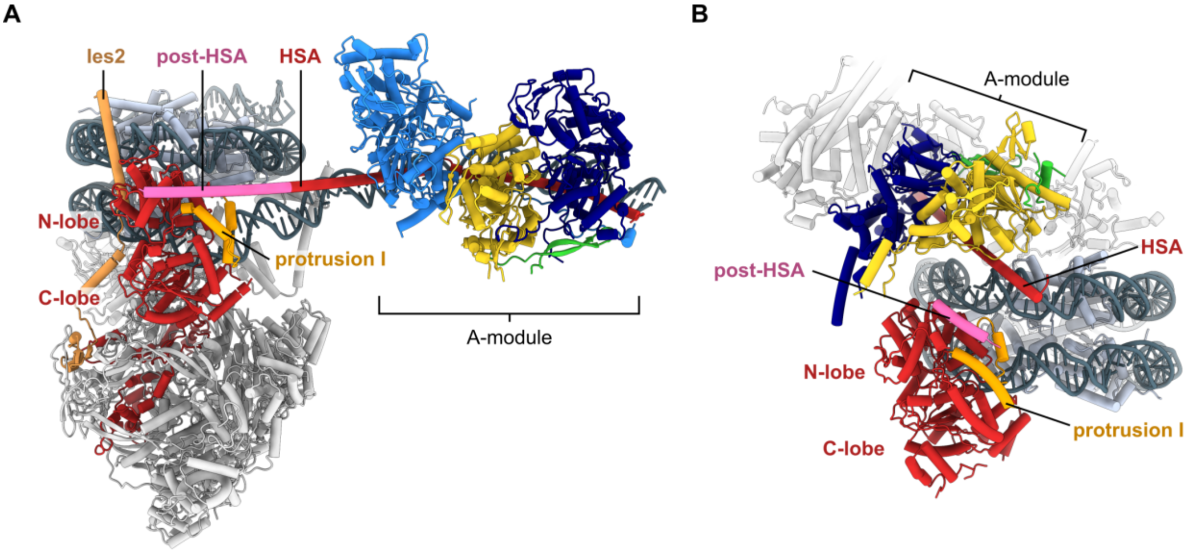


Figure 25: Structural comparison of INO80 and SWI/SNF chromatin remodelers. (A) Structural model of the *C. thermophilum* INO80 A/C-module complex bound to a nucleosome, with the A-module binding to extranucleosomal DNA and the motor domain at SHL -6 (based on EMDB: 15211) [352]. The Ino80 HSA domain (red), post-HSA domain (hot pink) and protrusion I (orange) are labeled and color-coded. Subunits of the A-module and Ies2 are colored (Ies2: orange, Arp8: bright blue, actin: yellow, Arp4: dark blue, Ies4: lime green). The A-module and the lobes of the motor domain are labeled. (B) Structure of the *S. cerevisiae* SWI/SNF complex bound to a nucleosome, with the A-module binding to nucleosomal DNA and the motor domain at SHL -2 (PDB: 6UXW) [288]. The Snf2 HSA domain (red), post-HSA domain (hot pink) and protrusion I (orange) are labeled and color-coded. Subunits of the A-module are colored (Arp9: yellow, Arp7: dark blue, Rtt102: lime green). The A-module and the lobes of the motor domain are labeled.

Taken together, INO80-mediated chromatin remodeling is a highly regulated process that relies on its modular architecture and is mediated by multivalent (extra)nucleosomal DNA interactions, especially by its actin-related subunits [352]. Future studies may include the species-specific N-module, which adds another regulatory layer as a sensor of extranucleosomal DNA, or investigate the interaction of INO80 with more complex substrates such as nucleosomal arrays.

References

1. Schwarz, D.A., C.D. Katayama, and S.M. Hedrick, *Schlafen, a new family of growth regulatory genes that affect thymocyte development.* Immunity, 1998. **9**(5): p. 657-68.
2. Geserick, P., et al., *Modulation of T cell development and activation by novel members of the Schlafen (slfn) gene family harbouring an RNA helicase-like motif.* Int Immunol, 2004. **16**(10): p. 1535-48.
3. Bustos, O., et al., *Evolution of the Schlafen genes, a gene family associated with embryonic lethality, meiotic drive, immune processes and orthopoxvirus virulence.* Gene, 2009. **447**(1): p. 1-11.
4. Jo, U. and Y. Pommier, *Structural, molecular, and functional insights into Schlafen proteins.* Exp Mol Med, 2022. **54**(6): p. 730-738.
5. Graham, G.J., *Tandem genes and clustered genes.* J Theor Biol, 1995. **175**(1): p. 71-87.
6. Li, M., et al., *Codon-usage-based inhibition of HIV protein synthesis by human schlafen 11.* Nature, 2012. **491**(7422): p. 125-8.
7. Abdel-Mohsen, M., et al., *Expression profile of host restriction factors in HIV-1 elite controllers.* Retrovirology, 2013. **10**: p. 106.
8. Seong, R.K., et al., *Schlafen 14 (SLFN14) is a novel antiviral factor involved in the control of viral replication.* Immunobiology, 2017. **222**(11): p. 979-988.
9. Kim, E.T., et al., *Comparative proteomics identifies Schlafen 5 (SLFN5) as a herpes simplex virus restriction factor that suppresses viral transcription.* Nat Microbiol, 2021. **6**(2): p. 234-245.
10. Kim, E.T. and M.D. Weitzman, *Schlafens Can Put Viruses to Sleep.* Viruses, 2022. **14**(2).
11. Sassano, A., et al., *Human Schlafen 5 (SLFN5) Is a Regulator of Motility and Invasiveness of Renal Cell Carcinoma Cells.* Mol Cell Biol, 2015. **35**(15): p. 2684-98.
12. Al-Marsoummi, S., E. Vomhof-DeKrey, and M.D. Basson, *Schlafen12 Reduces the Aggressiveness of Triple Negative Breast Cancer through Post-Transcriptional Regulation of ZEB1 That Drives Stem Cell Differentiation.* Cell Physiol Biochem, 2019. **53**(6): p. 999-1014.
13. Al-Marsoummi, S., et al., *Schlafen 12 Is Prognostically Favorable and Reduces C-Myc and Proliferation in Lung Adenocarcinoma but Not in Lung Squamous Cell Carcinoma.* Cancers (Basel), 2020. **12**(10).
14. Companioni Nápoles, O., et al., *SCHLAFEN 5 expression correlates with intestinal metaplasia that progresses to gastric cancer.* J Gastroenterol, 2017. **52**(1): p. 39-49.
15. Isnaldi, E., et al., *Schlafen-11 expression is associated with immune signatures and basal-like phenotype in breast cancer.* Breast Cancer Res Treat, 2019. **177**(2): p. 335-343.
16. Tian, L., et al., *Schlafen-11 sensitizes colorectal carcinoma cells to irinotecan.* Anticancer Drugs, 2014. **25**(10): p. 1175-81.
17. Zoppoli, G., et al., *Putative DNA/RNA helicase Schlafen-11 (SLFN11) sensitizes cancer cells to DNA-damaging agents.* Proc Natl Acad Sci U S A, 2012. **109**(37): p. 15030-5.
18. Nogales, V., et al., *Epigenetic inactivation of the putative DNA/RNA helicase SLFN11 in human cancer confers resistance to platinum drugs.* Oncotarget, 2016. **7**(3): p. 3084-97.

19. Allison Stewart, C., et al., *Dynamic variations in epithelial-to-mesenchymal transition (EMT), ATM, and SLFN11 govern response to PARP inhibitors and cisplatin in small cell lung cancer.* Oncotarget, 2017. **8**(17): p. 28575-28587.

20. Shee, K., et al., *Integrated pan-cancer gene expression and drug sensitivity analysis reveals SLFN11 mRNA as a solid tumor biomarker predictive of sensitivity to DNA-damaging chemotherapy.* PLoS One, 2019. **14**(11): p. e0224267.

21. Winkler, C., et al., *SLFN11 informs on standard of care and novel treatments in a wide range of cancer models.* Br J Cancer, 2021. **124**(5): p. 951-962.

22. Yang, J.Y., et al., *Structure of Schlafen13 reveals a new class of tRNA/rRNA- targeting RNase engaged in translational control.* Nat Commun, 2018. **9**(1): p. 1165.

23. Pisareva, V.P., et al., *Characterization of Novel Ribosome-Associated Endoribonuclease SLFN14 from Rabbit Reticulocytes.* Biochemistry, 2015. **54**(21): p. 3286-301.

24. Yue, T., et al., *SLFN2 protection of tRNAs from stress-induced cleavage is essential for T cell-mediated immunity.* Science, 2021. **372**(6543).

25. Chen, J., et al., *Structure of PDE3A-SLFN12 complex and structure-based design for a potent apoptosis inducer of tumor cells.* Nat Commun, 2021. **12**(1): p. 6204.

26. Garvie, C.W., et al., *Structure of PDE3A-SLFN12 complex reveals requirements for activation of SLFN12 RNase.* Nat Commun, 2021. **12**(1): p. 4375.

27. Neumann, B., et al., *Subcellular localization of the Schlafen protein family.* Biochem Biophys Res Commun, 2008. **370**(1): p. 62-6.

28. Brady, G., et al., *Schlafen-1 causes a cell cycle arrest by inhibiting induction of cyclin D1.* J Biol Chem, 2005. **280**(35): p. 30723-34.

29. Chaturvedi, L., et al., *The P-loop region of Schlafen 3 acts within the cytosol to induce differentiation of human Caco-2 intestinal epithelial cells.* Biochim Biophys Acta, 2014. **1843**(12): p. 3029-37.

30. Ye, Y., H.H. Meyer, and T.A. Rapoport, *The AAA ATPase Cdc48/p97 and its partners transport proteins from the ER into the cytosol.* Nature, 2001. **414**(6864): p. 652-6.

31. Glickman, M.H., et al., *The regulatory particle of the Saccharomyces cerevisiae proteasome.* Mol Cell Biol, 1998. **18**(6): p. 3149-62.

32. Berger, M., et al., *An Slfn2 mutation causes lymphoid and myeloid immunodeficiency due to loss of immune cell quiescence.* Nat Immunol, 2010. **11**(4): p. 335-43.

33. Omar, I., et al., *Schlafen2 mutation unravels a role for chronic ER stress in the loss of T cell quiescence.* Oncotarget, 2016. **7**(26): p. 39396-39407.

34. Omar, I., et al., *Slfn2 mutation-induced loss of T-cell quiescence leads to elevated de novo sterol synthesis.* Immunology, 2017. **152**(3): p. 484-493.

35. Warsi, S., et al., *Schlafen2 is a regulator of quiescence in adult murine hematopoietic stem cells.* Haematologica, 2022.

36. Mavrommatis, E., et al., *Expression and regulatory effects of murine Schlafen (Slfn) genes in malignant melanoma and renal cell carcinoma.* J Biol Chem, 2013. **288**(46): p. 33006-15.

37. Katsoulidis, E., et al., *Role of Schlafen 2 (SLFN2) in the generation of interferon alpha-induced growth inhibitory responses.* J Biol Chem, 2009. **284**(37): p. 25051-64.

38. Yusuf, I. and D.A. Fruman, *Regulation of quiescence in lymphocytes.* Trends Immunol, 2003. **24**(7): p. 380-6.

39. Chapman, N.M., M.R. Boothby, and H. Chi, *Metabolic coordination of T cell quiescence and activation.* Nat Rev Immunol, 2020. **20**(1): p. 55-70.

40. Sena, L.A., et al., *Mitochondria are required for antigen-specific T cell activation through reactive oxygen species signaling.* Immunity, 2013. **38**(2): p. 225-36.

41. Ivanov, P., et al., *Angiogenin-induced tRNA fragments inhibit translation initiation.* Mol Cell, 2011. **43**(4): p. 613-23.

42. Yamasaki, S., et al., *Angiogenin cleaves tRNA and promotes stress-induced translational repression.* J Cell Biol, 2009. **185**(1): p. 35-42.

43. Fu, H., et al., *Stress induces tRNA cleavage by angiogenin in mammalian cells.* FEBS Lett, 2009. **583**(2): p. 437-42.

44. Zhang, Y., et al., *The Hsp40 family chaperone protein DnaJB6 enhances Schlafen1 nuclear localization which is critical for promotion of cell-cycle arrest in T-cells.* Biochem J, 2008. **413**(2): p. 239-50.

45. Li, M., et al., *DNA damage-induced cell death relies on SLFN11-dependent cleavage of distinct type II tRNAs.* Nat Struct Mol Biol, 2018. **25**(11): p. 1047-1058.

46. Lee, S., et al., *Velcrin-induced selective cleavage of tRNA(Leu)(TAA) by SLFN12 causes cancer cell death.* Nat Chem Biol, 2022.

47. Chen, J. and L.A. Kuhn, *Deciphering the three-domain architecture in schlafens and the structures and roles of human schlafen12 and serpinB12 in transcriptional regulation.* J Mol Graph Model, 2019. **90**: p. 59-76.

48. Basson, M.D., et al., *Schlafen 12 Interaction with SerpinB12 and Deubiquitylases Drives Human Enterocyte Differentiation.* Cell Physiol Biochem, 2018. **48**(3): p. 1274-1290.

49. de Waal, L., et al., *Identification of cancer-cytotoxic modulators of PDE3A by predictive chemogenomics.* Nat Chem Biol, 2016. **12**(2): p. 102-8.

50. Hunter, R.W., C. Mackintosh, and I. Hers, *Protein kinase C-mediated phosphorylation and activation of PDE3A regulate cAMP levels in human platelets.* J Biol Chem, 2009. **284**(18): p. 12339-48.

51. Li, D., et al., *Estrogen-Related Hormones Induce Apoptosis by Stabilizing Schlafen-12 Protein Turnover.* Mol Cell, 2019. **75**(6): p. 1103-1116.e9.

52. An, R., et al., *PDE3A inhibitor anagrelide activates death signaling pathway genes and synergizes with cell death-inducing cytokines to selectively inhibit cancer cell growth.* Am J Cancer Res, 2019. **9**(9): p. 1905-1921.

53. Wu, X., et al., *Mechanistic insights into cancer cell killing through interaction of phosphodiesterase 3A and schlafen family member 12.* J Biol Chem, 2020. **295**(11): p. 3431-3446.

54. Ai, Y., et al., *An alkaloid initiates phosphodiesterase 3A-schlafen 12 dependent apoptosis without affecting the phosphodiesterase activity.* Nat Commun, 2020. **11**(1): p. 3236.

55. Yan, B., et al., *Multiple PDE3A modulators act as molecular glues promoting PDE3A-SLFN12 interaction and induce SLFN12 dephosphorylation and cell death.* Cell Chem Biol, 2022. **29**(6): p. 958-969.e5.

56. Walker, J.E., et al., *Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold.* Embo j, 1982. **1**(8): p. 945-51.

57. Ding, J., et al., *Schlafen 5 suppresses human immunodeficiency virus type 1 transcription by commandeering cellular epigenetic machinery.* Nucleic Acids Res, 2022. **50**(11): p. 6137-53.

58. Katsoulidis, E., et al., *Role of interferon {alpha} (IFN{alpha})-inducible Schlafen-5 in regulation of anchorage-independent growth and invasion of malignant melanoma cells.* J Biol Chem, 2010. **285**(51): p. 40333-41.

59. Wan, G., et al., *SLFN5 suppresses cancer cell migration and invasion by inhibiting MT1-MMP expression via AKT/GSK-3β/β-catenin pathway.* Cell Signal, 2019. **59**: p. 1-12.

60. Yang, X., et al., *The association between SLFN5 expression and the prognosis of non-small cell lung cancer.* Int J Clin Exp Med, 2019. **12**: p. 682-689.

61. Wan, G., et al., *Human Schlafen 5 regulates reversible epithelial and mesenchymal transitions in breast cancer by suppression of ZEB1 transcription.* Br J Cancer, 2020. **123**(4): p. 633-643.

62. Gu, X., et al., *SLFN5 influences proliferation and apoptosis by upregulating PTEN transcription via ZEB1 and inhibits the purine metabolic pathway in breast cancer.* Am J Cancer Res, 2020. **10**(9): p. 2832-2850.

63. Gu, X., et al., *Human Schlafen 5 Inhibits Proliferation and Promotes Apoptosis in Lung Adenocarcinoma via the PTEN/PI3K/AKT/mTOR Pathway.* Biomed Res Int, 2021. **2021**: p. 6628682.

64. Arslan, A.D., et al., *Human SLFN5 is a transcriptional co-repressor of STAT1-mediated interferon responses and promotes the malignant phenotype in glioblastoma.* Oncogene, 2017. **36**(43): p. 6006-6019.

65. Guo, L., Z. Liu, and X. Tang, *Overexpression of SLFN5 induced the epithelial-mesenchymal transition in human lung cancer cell line A549 through β-catenin/Snail/E-cadherin pathway.* Eur J Pharmacol, 2019. **862**: p. 172630.

66. Fischietti, M., et al., *Schlafen 5 as a novel therapeutic target in pancreatic ductal adenocarcinoma.* Oncogene, 2021. **40**(18): p. 3273-3286.

67. Martinez, R.S., et al., *SLFN5 Regulates LAT1-Mediated mTOR Activation in Castration-Resistant Prostate Cancer.* Cancer Res, 2021. **81**(13): p. 3664-3678.

68. Xu, J., et al., *Schlafen family is a prognostic biomarker and corresponds with immune infiltration in gastric cancer.* Front Immunol, 2022. **13**: p. 922138.

69. Lin, Y.Z., et al., *Equine schlafen 11 restricts the production of equine infectious anemia virus via a codon usage-dependent mechanism.* Virology, 2016. **495**: p. 112-21.

70. Guo, G., et al., *Human Schlafen 11 exploits codon preference discrimination to attenuate viral protein synthesis of prototype foamy virus (PFV).* Virology, 2021. **555**: p. 78-88.

71. Murai, J., et al., *SLFN11 Blocks Stressed Replication Forks Independently of ATR.* Mol Cell, 2018. **69**(3): p. 371-384.e6.

72. Murai, J., et al., *Chromatin Remodeling and Immediate Early Gene Activation by SLFN11 in Response to Replication Stress.* Cell Rep, 2020. **30**(12): p. 4137-4151.e6.

73. Fletcher, S.J., et al., *SLFN14 mutations underlie thrombocytopenia with excessive bleeding and platelet secretion defects.* J Clin Invest, 2015. **125**(9): p. 3600-5.

74. Fletcher, S.J., et al., *Role of the novel endoribonuclease SLFN14 and its disease-causing mutations in ribosomal degradation.* Rna, 2018. **24**(7): p. 939-949.

75. Stapley, R.J., et al., *Heterozygous mutation SLFN14 K208N in mice mediates species-specific differences in platelet and erythroid lineage commitment.* Blood Adv, 2021. **5**(2): p. 377-390.

76. Dinarello, C.A., *Historical insights into cytokines.* Eur J Immunol, 2007. **37 Suppl 1**(Suppl 1): p. S34-45.

77. McNab, F., et al., *Type I interferons in infectious disease.* Nat Rev Immunol, 2015. **15**(2): p. 87-103.

78. Isaacs, A. and J. Lindenmann, *Virus interference. I. The interferon.* Proc R Soc Lond B Biol Sci, 1957. **147**(927): p. 258-67.

79. Pestka, S., C.D. Krause, and M.R. Walter, *Interferons, interferon-like cytokines, and their receptors.* Immunol Rev, 2004. **202**: p. 8-32.

80. Schoenborn, J.R. and C.B. Wilson, *Regulation of interferon-gamma during innate and adaptive immune responses.* Adv Immunol, 2007. **96**: p. 41-101.

81. Witte, K., et al., *IL-28A, IL-28B, and IL-29: promising cytokines with type I interferon-like properties.* Cytokine Growth Factor Rev, 2010. **21**(4): p. 237-51.

82. Li, D. and M. Wu, *Pattern recognition receptors in health and diseases.* Signal Transduct Target Ther, 2021. **6**(1): p. 291.

83. Goubau, D., S. Deddouche, and C. Reis e Sousa, *Cytosolic sensing of viruses.* Immunity, 2013. **38**(5): p. 855-69.

84. Yoneyama, M., et al., *The RNA helicase RIG-I has an essential function in double-stranded RNA-induced innate antiviral responses.* Nat Immunol, 2004. **5**(7): p. 730-7.

85. Sun, L., et al., *Cyclic GMP-AMP synthase is a cytosolic DNA sensor that activates the type I interferon pathway.* Science, 2013. **339**(6121): p. 786-91.

86. Platanias, L.C., *Mechanisms of type-I- and type-II-interferon-mediated signalling.* Nat Rev Immunol, 2005. **5**(5): p. 375-86.

87. Novick, D., B. Cohen, and M. Rubinstein, *The human interferon alpha/beta receptor: characterization and molecular cloning.* Cell, 1994. **77**(3): p. 391-400.

88. Silvennoinen, O., et al., *Interferon-induced nuclear signalling by Jak protein tyrosine kinases.* Nature, 1993. **366**(6455): p. 583-5.

89. Darnell, J.E., Jr., I.M. Kerr, and G.R. Stark, *Jak-STAT pathways and transcriptional activation in response to IFNs and other extracellular signaling proteins.* Science, 1994. **264**(5164): p. 1415-21.

90. Darnell, J.E., Jr., *STATs and gene regulation.* Science, 1997. **277**(5332): p. 1630-5.

91. Der, S.D., et al., *Identification of genes differentially regulated by interferon alpha, beta, or gamma using oligonucleotide arrays.* Proc Natl Acad Sci U S A, 1998. **95**(26): p. 15623-8.

92. Mogensen, T.H., *IRF and STAT Transcription Factors - From Basic Biology to Roles in Infection, Protective Immunity, and Primary Immunodeficiencies.* Front Immunol, 2018. **9**: p. 3047.

93. Schneider, W.M., M.D. Chevillotte, and C.M. Rice, *Interferon-stimulated genes: a complex web of host defenses.* Annu Rev Immunol, 2014. **32**: p. 513-45.

94. Mavrommatis, E., E.N. Fish, and L.C. Platanias, *The schlafen family of proteins and their regulation by interferons.* J Interferon Cytokine Res, 2013. **33**(4): p. 206-10.

95. Puck, A., et al., *Expression and regulation of Schlafen (SLFN) family members in primary human monocytes, monocyte-derived dendritic cells and T cells.* Results Immunol, 2015. **5**: p. 23-32.

96. Engelman, A. and P. Cherepanov, *The structural biology of HIV-1: mechanistic and therapeutic insights.* Nat Rev Microbiol, 2012. **10**(4): p. 279-90.

97. Boso, G. and C.A. Kozak, *Retroviral Restriction Factors and Their Viral Targets: Restriction Strategies and Evolutionary Adaptations.* Microorganisms, 2020. **8**(12).

98. Goldstone, D.C., et al., *HIV-1 restriction factor SAMHD1 is a deoxynucleoside triphosphate triphosphohydrolase.* Nature, 2011. **480**(7377): p. 379-82.

99. Neil, S.J., T. Zang, and P.D. Bieniasz, *Tetherin inhibits retrovirus release and is antagonized by HIV-1 Vpu.* Nature, 2008. **451**(7177): p. 425-30.

100. Sohn, W.J., et al., *Novel transcriptional regulation of the schlafen-2 gene in macrophages in response to TLR-triggered stimulation.* Mol Immunol, 2007. **44**(13): p. 3273-82.

101. van Zuylen, W.J., et al., *Macrophage activation and differentiation signals regulate schlafen-4 gene expression: evidence for Schlafen-4 as a modulator of myelopoiesis.* PLoS One, 2011. **6**(1): p. e15723.

102. Abdel-Mohsen, M., et al., *Select host restriction factors are associated with HIV persistence during antiretroviral therapy.* Aids, 2015. **29**(4): p. 411-20.

103. Coleman, J.R., et al., *Virus attenuation by genome-scale changes in codon pair bias.* Science, 2008. **320**(5884): p. 1784-7.

104. Meintjes, P.L. and A.G. Rodrigo, *Evolution of relative synonymous codon usage in Human Immunodeficiency Virus type-1.* J Bioinform Comput Biol, 2005. **3**(1): p. 157-68.

105. Berkhout, B. and F.J. van Hemert, *The unusual nucleotide content of the HIV RNA genome results in a biased amino acid composition of HIV proteins.* Nucleic Acids Res, 1994. **22**(9): p. 1705-11.

106. Haas, J., E.C. Park, and B. Seed, *Codon usage limitation in the expression of HIV-1 envelope glycoprotein.* Curr Biol, 1996. **6**(3): p. 315-24.

107. Liu, Y., *A code within the genetic code: codon usage regulates co-translational protein folding.* Cell Commun Signal, 2020. **18**(1): p. 145.

108. Valdez, F., et al., *Schlafen 11 Restricts Flavivirus Replication.* J Virol, 2019. **93**(15).

109. Kim, Y., et al., *Type-II tRNAs and Evolution of Translation Systems and the Genetic Code.* Int J Mol Sci, 2018. **19**(10).

110. Stabell, A.C., et al., *Non-human Primate Schlafen11 Inhibits Production of Both Host and Viral Proteins.* PLoS Pathog, 2016. **12**(12): p. e1006066.

111. Malone, D., et al., *Dephosphorylation activates the interferon-stimulated Schlafen family member 11 in the DNA damage response.* J Biol Chem, 2019. **294**(40): p. 14674-14685.

112. Nightingale, K., et al., *Human cytomegalovirus protein RL1 degrades the antiviral factor SLFN11 via recruitment of the CRL4 E3 ubiquitin ligase complex.* Proc Natl Acad Sci U S A, 2022. **119**(6).

113. Kane, M., et al., *Identification of Interferon-Stimulated Genes with Antiretroviral Activity.* Cell Host Microbe, 2016. **20**(3): p. 392-405.

114. Liu, S.Y., et al., *Systematic identification of type I and type II interferon-induced antiviral factors.* Proc Natl Acad Sci U S A, 2012. **109**(11): p. 4239-44.

115. Gubser, C., et al., *Camelpox virus encodes a schlafen-like protein that affects orthopoxvirus virulence.* J Gen Virol, 2007. **88**(Pt 6): p. 1667-1676.

116. Eaglesham, J.B., et al., *Viral and metazoan poxins are cGAMP-specific nucleases that restrict cGAS-STING signalling.* Nature, 2019. **566**(7743): p. 259-263.

117. Hernáez, B., et al., *Viral cGAMP nuclease reveals the essential role of DNA sensing in protection against acute lethal virus infection.* Sci Adv, 2020. **6**(38).

118. Al-Marsoummi, S., E.E. Vomhof-DeKrey, and M.D. Basson, *Schlafens: Emerging Proteins in Cancer Cell Biology.* Cells, 2021. **10**(9).

119. Caramel, J., M. Ligier, and A. Puisieux, *Pleiotropic Roles for ZEB1 in Cancer.* Cancer Res, 2018. **78**(1): p. 30-35.

120. Correa, P., M.B. Piazuelo, and K.T. Wilson, *Pathology of gastric intestinal metaplasia: clinical implications.* Am J Gastroenterol, 2010. **105**(3): p. 493-8.

121. Wang, Q., et al., *Androgen receptor and nutrient signaling pathways coordinate the demand for increased amino acid transport during prostate cancer progression.* Cancer Res, 2011. **71**(24): p. 7525-36.

122. Roychowdhury, S. and A.M. Chinnaiyan, *Translating cancer genomes and transcriptomes for precision oncology.* CA Cancer J Clin, 2016. **66**(1): p. 75-88.

123. Barretina, J., et al., *The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity.* Nature, 2012. **483**(7391): p. 603-7.

124. Mu, Y., et al., *SLFN11 inhibits checkpoint maintenance and homologous recombination repair.* EMBO Rep, 2016. **17**(1): p. 94-109.

125. Murai, J., et al., *Resistance to PARP inhibitors by SLFN11 inactivation can be overcome by ATR inhibition.* Oncotarget, 2016. **7**(47): p. 76534-76550.

126. Burgers, P.M.J. and T.A. Kunkel, *Eukaryotic DNA Replication Fork.* Annu Rev Biochem, 2017. **86**: p. 417-438.

127. Jo, U., et al., *Precision Oncology with Drugs Targeting the Replication Stress, ATR, and Schlafen 11.* Cancers (Basel), 2021. **13**(18).

128. Fragkos, M., et al., *DNA replication origin activation in space and time.* Nat Rev Mol Cell Biol, 2015. **16**(6): p. 360-74.

129. Speck, C., et al., *ATPase-dependent cooperative binding of ORC and Cdc6 to origin DNA.* Nat Struct Mol Biol, 2005. **12**(11): p. 965-71.

130. Remus, D., et al., *Concerted loading of Mcm2-7 double hexamers around DNA during DNA replication origin licensing.* Cell, 2009. **139**(4): p. 719-30.

131. Heller, R.C., et al., *Eukaryotic origin-dependent DNA replication in vitro reveals sequential action of DDK and S-CDK kinases.* Cell, 2011. **146**(1): p. 80-91.

132. Tanaka, S., et al., *CDK-dependent phosphorylation of Sld2 and Sld3 initiates DNA replication in budding yeast.* Nature, 2007. **445**(7125): p. 328-32.

133. Ilves, I., et al., *Activation of the MCM2-7 helicase by association with Cdc45 and GINS proteins.* Mol Cell, 2010. **37**(2): p. 247-58.

134. Douglas, M.E., et al., *The mechanism of eukaryotic CMG helicase activation.* Nature, 2018. **555**(7695): p. 265-268.

135. Burnham, D.R., et al., *The mechanism of DNA unwinding by the eukaryotic replicative helicase.* Nat Commun, 2019. **10**(1): p. 2159.

136. Zou, Y., et al., *Functions of human replication protein A (RPA): from DNA replication to DNA damage and stress responses.* J Cell Physiol, 2006. **208**(2): p. 267-73.

137. Pursell, Z.F., et al., *Yeast DNA polymerase epsilon participates in leading-strand DNA replication.* Science, 2007. **317**(5834): p. 127-30.

138. Kilkenny, M.L., et al., *Structures of human primase reveal design of nucleotide elongation site and mode of Pol α tethering.* Proc Natl Acad Sci U S A, 2013. **110**(40): p. 15961-6.

139. Johnson, R.E., et al., *A Major Role of DNA Polymerase δ in Replication of Both the Leading and Lagging DNA Strands.* Mol Cell, 2015. **59**(2): p. 163-175.

140. Moldovan, G.L., B. Pfander, and S. Jentsch, *PCNA, the maestro of the replication fork.* Cell, 2007. **129**(4): p. 665-79.

141. Pommier, Y., et al., *Human topoisomerases and their roles in genome stability and organization.* Nat Rev Mol Cell Biol, 2022. **23**(6): p. 407-427.

142. Dewar, J.M., M. Budzowska, and J.C. Walter, *The mechanism of DNA replication termination in vertebrates.* Nature, 2015. **525**(7569): p. 345-50.

143. Cayrou, C., et al., *Genome-scale analysis of metazoan replication origins reveals their organization in specific but flexible sites defined by conserved features.* Genome Res, 2011. **21**(9): p. 1438-49.

144. Akerman, I., et al., *A predictable conserved DNA base composition signature defines human core DNA replication origins.* Nat Commun, 2020. **11**(1): p. 4826.

145. Muñoz, S. and J. Méndez, *DNA replication stress: from molecular mechanisms to human disease.* Chromosoma, 2017. **126**(1): p. 1-15.

146. Zeman, M.K. and K.A. Cimprich, *Causes and consequences of replication stress.* Nat Cell Biol, 2014. **16**(1): p. 2-9.

147. Gaillard, H., T. García-Muse, and A. Aguilera, *Replication stress and cancer.* Nat Rev Cancer, 2015. **15**(5): p. 276-89.

148. Cheung-Ong, K., G. Giaever, and C. Nislow, *DNA-damaging agents in cancer chemotherapy: serendipity and chemical biology.* Chem Biol, 2013. **20**(5): p. 648-59.

149. Noll, D.M., T.M. Mason, and P.S. Miller, *Formation and repair of interstrand cross-links in DNA.* Chem Rev, 2006. **106**(2): p. 277-301.

150. Byun, T.S., et al., *Functional uncoupling of MCM helicase and DNA polymerase activities activates the ATR-dependent checkpoint.* Genes Dev, 2005. **19**(9): p. 1040-52.

151. Zou, L. and S.J. Elledge, *Sensing DNA damage through ATRIP recognition of RPA-ssDNA complexes.* Science, 2003. **300**(5625): p. 1542-8.

152. Liu, Q., et al., *Chk1 is an essential kinase that is regulated by Atr and required for the G(2)/M DNA damage checkpoint.* Genes Dev, 2000. **14**(12): p. 1448-59.

153. Cimprich, K.A. and D. Cortez, *ATR: an essential regulator of genome integrity.* Nat Rev Mol Cell Biol, 2008. **9**(8): p. 616-27.

154. Saldivar, J.C., D. Cortez, and K.A. Cimprich, *The essential kinase ATR: ensuring faithful duplication of a challenging genome.* Nat Rev Mol Cell Biol, 2017. **18**(10): p. 622-636.

155. Cortez, D., *Replication-Coupled DNA Repair.* Mol Cell, 2019. **74**(5): p. 866-876.

156. Ceccaldi, R., P. Sarangi, and A.D. D'Andrea, *The Fanconi anaemia pathway: new players and new functions.* Nat Rev Mol Cell Biol, 2016. **17**(6): p. 337-49.

157. Schlacher, K., H. Wu, and M. Jasin, *A distinct replication fork protection pathway connects Fanconi anemia tumor suppressors to RAD51-BRCA1/2.* Cancer Cell, 2012. **22**(1): p. 106-16.

158. Sale, J.E., A.R. Lehmann, and R. Woodgate, *Y-family DNA polymerases and their role in tolerance of cellular DNA damage.* Nat Rev Mol Cell Biol, 2012. **13**(3): p. 141-52.

159. García-Gómez, S., et al., *PrimPol, an archaic primase/polymerase operating in human cells.* Mol Cell, 2013. **52**(4): p. 541-53.

160. Mourón, S., et al., *Repriming of DNA synthesis at stalled replication forks by human PrimPol.* Nat Struct Mol Biol, 2013. **20**(12): p. 1383-9.

161. Kawabata, T., et al., *Stalled fork rescue via dormant replication origins in unchallenged S phase promotes proper chromosome segregation and tumor suppression.* Mol Cell, 2011. **41**(5): p. 543-53.

162. Kitao, H., et al., *DNA replication stress and cancer chemotherapy.* Cancer Sci, 2018. **109**(2): p. 264-271.

163. Donaldson, K.L., G.L. Goolsby, and A.F. Wahl, *Cytotoxicity of the anticancer agents cisplatin and taxol during cell proliferation and the cell cycle.* Int J Cancer, 1994. **57**(6): p. 847-55.

164. Avemann, K., et al., *Camptothecin, a specific inhibitor of type I DNA topoisomerase, induces DNA breakage at replication forks.* Mol Cell Biol, 1988. **8**(8): p. 3026-34.

165. Jo, U., et al., *SLFN11 promotes CDT1 degradation by CUL4 in response to replicative DNA damage, while its absence leads to synthetic lethality with ATR/CHK1 inhibitors.* Proc Natl Acad Sci U S A, 2021. **118**(6).

166. Okamoto, Y., et al., *SLFN11 promotes stalled fork degradation that underlies the phenotype in Fanconi anemia cells.* Blood, 2021. **137**(3): p. 336-348.

167. Raderschall, E., E.I. Golub, and T. Haaf, *Nuclear foci of mammalian recombination proteins are located at single-stranded DNA regions formed after DNA damage.* Proc Natl Acad Sci U S A, 1999. **96**(5): p. 1921-6.

168. Lee, T., et al., *Suppression of the DHX9 helicase induces premature senescence in human diploid fibroblasts in a p53-dependent manner.* J Biol Chem, 2014. **289**(33): p. 22798-22814.

169. Murai, J., et al., *Schlafen 11 (SLFN11), a restriction factor for replicative stress induced by DNA-targeting anti-cancer therapies.* Pharmacol Ther, 2019. **201**: p. 94-102.

170. Jin, J., et al., *A family of diverse Cul4-Ddb1-interacting proteins includes Cdt2, which is required for S phase destruction of the replication factor Cdt1.* Mol Cell, 2006. **23**(5): p. 709-21.

171. Taniyama, D., et al., *Prognostic impact of Schlafen 11 in bladder cancer patients treated with platinum-based chemotherapy.* Cancer Sci, 2022. **113**(2): p. 784-795.

172. Winkler, C., et al., *SLFN11 captures cancer-immunity interactions associated with platinum sensitivity in high-grade serous ovarian cancer.* JCI Insight, 2021. **6**(18).

173. Willis, S.E., et al., *Retrospective analysis of Schlafen11 (SLFN11) to predict the outcomes to therapies affecting the DNA damage response.* Br J Cancer, 2021. **125**(12): p. 1666-1676.

174. Takashima, T., et al., *Schlafen 11 predicts response to platinum-based chemotherapy in gastric cancers.* Br J Cancer, 2021. **125**(1): p. 65-77.

175. Conteduca, V., et al., *SLFN11 Expression in Advanced Prostate Cancer and Response to Platinum-based Chemotherapy.* Mol Cancer Ther, 2020. **19**(5): p. 1157-1164.

176. Deng, Y., et al., *High SLFN11 expression predicts better survival for patients with KRAS exon 2 wild type colorectal cancer after treated with adjuvant oxaliplatin-based treatment.* BMC Cancer, 2015. **15**: p. 833.

177. Sousa, F.G., et al., *Alterations of DNA repair genes in the NCI-60 cell lines and their predictive value for anticancer drug activity.* DNA Repair (Amst), 2015. **28**: p. 107-15.

178. Marzi, L., et al., *Novel Fluoroindenoisoquinoline Non-Camptothecin Topoisomerase I Inhibitors.* Mol Cancer Ther, 2018. **17**(8): p. 1694-1704.

179. Marzi, L., et al., *The Indenoisoquinoline TOP1 Inhibitors Selectively Target Homologous Recombination-Deficient and Schlafen 11-Positive Cancer Cells and Synergize with Olaparib.* Clin Cancer Res, 2019. **25**(20): p. 6206-6216.

180. Coussy, F., et al., *BRCAness, SLFN11, and RB1 loss predict response to topoisomerase I inhibitors in triple-negative breast cancers.* Sci Transl Med, 2020. **12**(531).

181. Gartrell, J., et al., *SLFN11 is Widely Expressed in Pediatric Sarcoma and Induces Variable Sensitization to Replicative Stress Caused By DNA-Damaging Agents.* Mol Cancer Ther, 2021. **20**(11): p. 2151-2165.

182. Jo, U., et al., *TOP1-DNA Trapping by Exatecan and Combination Therapy with ATR Inhibitor.* Mol Cancer Ther, 2022. **21**(7): p. 1090-1102.

183. Tang, S.W., et al., *Overcoming Resistance to DNA-Targeted Agents by Epigenetic Activation of Schlafen 11 (SLFN11) Expression with Class I Histone Deacetylase Inhibitors.* Clin Cancer Res, 2018. **24**(8): p. 1944-1953.

184. Lok, B.H., et al., *PARP Inhibitor Activity Correlates with SLFN11 Expression and Demonstrates Synergy with Temozolomide in Small Cell Lung Cancer.* Clin Cancer Res, 2017. **23**(2): p. 523-535.

185. Rathkey, D., et al., *Sensitivity of Mesothelioma Cells to PARP Inhibitors Is Not Dependent on BAP1 but Is Enhanced by Temozolomide in Cells With High-Schlafen 11 and Low-O6-methylguanine-DNA Methyltransferase Expression.* J Thorac Oncol, 2020. **15**(5): p. 843-859.

186. Wagner, J.M. and L.M. Karnitz, *Cisplatin-induced DNA damage activates replication checkpoint signaling components that differentially affect tumor cell survival.* Mol Pharmacol, 2009. **76**(1): p. 208-14.

187. Pommier, Y., *DNA topoisomerase I inhibitors: chemistry, biology, and interfacial inhibition.* Chem Rev, 2009. **109**(7): p. 2894-902.

188. Berdis, A.J., *DNA polymerases as therapeutic targets.* Biochemistry, 2008. **47**(32): p. 8253-60.

189. Rose, M., et al., *PARP Inhibitors: Clinical Relevance, Mechanisms of Action and Tumor Resistance.* Front Cell Dev Biol, 2020. **8**: p. 564601.

190. Kundu, K., et al., *SLFN11 biomarker status predicts response to lurbinectedin as a single agent and in combination with ATR inhibition in small cell lung cancer.* Transl Lung Cancer Res, 2021. **10**(11): p. 4095-4105.

191. Yin, Y.P., et al., *FK228 potentiates topotecan activity against small cell lung cancer cells via induction of SLFN11.* Acta Pharmacol Sin, 2021.

192. Peng, Y., et al., *Methylation of SLFN11 promotes gastric cancer growth and increases gastric cancer cell resistance to cisplatin.* J Cancer, 2019. **10**(24): p. 6124-6134.

193. He, T., et al., *Methylation of SLFN11 is a marker of poor prognosis and cisplatin resistance in colorectal cancer.* Epigenomics, 2017. **9**(6): p. 849-862.

194. Kang, M.H., et al., *Activity of MM-398, nanoliposomal irinotecan (nal-IRI), in Ewing's family tumor xenografts is associated with high exposure of tumor to drug and high SLFN11 expression.* Clin Cancer Res, 2015. **21**(5): p. 1139-50.

195. Tang, S.W., et al., *SLFN11 Is a Transcriptional Target of EWS-FLI1 and a Determinant of Drug Response in Ewing Sarcoma.* Clin Cancer Res, 2015. **21**(18): p. 4184-93.

196. Iwasaki, J., et al., *Schlafen11 Expression Is Associated With the Antitumor Activity of Trabectedin in Human Sarcoma Cell Lines.* Anticancer Res, 2019. **39**(7): p. 3553-3563.

197. Zhang, B., et al., *A wake-up call for cancer DNA damage: the role of Schlafen 11 (SLFN11) across multiple cancers.* Br J Cancer, 2021. **125**(10): p. 1333-1340.

198. Krushkal, J., et al., *Epigenome-wide DNA methylation analysis of small cell lung cancer cell lines suggests potential chemotherapy targets.* Clin Epigenetics, 2020. **12**(1): p. 93.

199. Gardner, E.E., et al., *Chemosensitive Relapse in Small Cell Lung Cancer Proceeds through an EZH2-SLFN11 Axis.* Cancer Cell, 2017. **31**(2): p. 286-299.

200. Mao, S., et al., *Resistance to Pyrrolobenzodiazepine Dimers Is Associated with SLFN11 Downregulation and Can Be Reversed through Inhibition of ATR.* Mol Cancer Ther, 2021. **20**(3): p. 541-552.

201. Takashima, T., et al., *Immunohistochemical analysis of SLFN11 expression uncovers potential non-responders to DNA-damaging agents overlooked by tissue RNA-seq.* Virchows Arch, 2021. **478**(3): p. 569-579.

202. Pietanza, M.C., et al., *Randomized, Double-Blind, Phase II Study of Temozolomide in Combination With Either Veliparib or Placebo in Patients With Relapsed-Sensitive or Refractory Small-Cell Lung Cancer.* J Clin Oncol, 2018. **36**(23): p. 2386-2394.

203. Federico, S.M., et al., *A phase I trial of talazoparib and irinotecan with and without temozolomide in children and young adults with recurrent or refractory solid malignancies.* Eur J Cancer, 2020. **137**: p. 204-213.

204. Byers, L.A., et al., *Veliparib in Combination with Carboplatin and Etoposide in Patients with Treatment-Naïve Extensive-Stage Small Cell Lung Cancer: A Phase 2 Randomized Study.* Clin Cancer Res, 2021. **27**(14): p. 3884-3895.

205. Zhou, C., et al., *SLFN11 inhibits hepatocellular carcinoma tumorigenesis and metastasis by targeting RPS4X via mTOR pathway.* Theranostics, 2020. **10**(10): p. 4627-4643.

206. Fischietti, M., et al., *SLFN11 negatively regulates non-canonical NFkB signaling to promote glioblastoma progression.* Cancer Res Commun, 2022. **2**(9): p. 966-978.

207. Klemm, S.L., Z. Shipony, and W.J. Greenleaf, *Chromatin accessibility and the regulatory epigenome.* Nat Rev Genet, 2019. **20**(4): p. 207-220.

208. Groth, A., et al., *Chromatin challenges during DNA replication and repair.* Cell, 2007. **128**(4): p. 721-33.

209. Luger, K., et al., *Crystal structure of the nucleosome core particle at 2.8 A resolution.* Nature, 1997. **389**(6648): p. 251-60.

210. Kornberg, R.D., *Chromatin structure: a repeating unit of histones and DNA.* Science, 1974. **184**(4139): p. 868-71.

211. Olins, A.L. and D.E. Olins, *Spheroid chromatin units (v bodies).* Science, 1974. **183**(4122): p. 330-2.

212. Finch, J.T., et al., *Structure of nucleosome core particles of chromatin.* Nature, 1977. **269**(5623): p. 29-36.

213. Kornberg, R.D. and J.O. Thomas, *Chromatin structure; oligomers of the histones.* Science, 1974. **184**(4139): p. 865-8.

214. Cutter, A.R. and J.J. Hayes, *A brief review of nucleosome structure.* FEBS Lett, 2015. **589**(20 Pt A): p. 2914-22.

215. Wang, H., L. Xiong, and P. Cramer, *Structures and implications of TBP-nucleosome complexes.* Proc Natl Acad Sci U S A, 2021. **118**(30).

216. Fan, J.Y., et al., *H2A.Z alters the nucleosome surface to promote HP1alpha-mediated chromatin fiber folding.* Mol Cell, 2004. **16**(4): p. 655-61.

217. Barbera, A.J., et al., *The nucleosomal surface as a docking station for Kaposi's sarcoma herpesvirus LANA.* Science, 2006. **311**(5762): p. 856-61.

218. McGinty, R.K. and S. Tan, *Recognition of the nucleosome by chromatin factors and enzymes.* Curr Opin Struct Biol, 2016. **37**: p. 54-61.

219. McGinty, R.K., R.C. Henrici, and S. Tan, *Crystal structure of the PRC1 ubiquitylation module bound to the nucleosome.* Nature, 2014. **514**(7524): p. 591-6.

220. Dorigo, B., et al., *Chromatin fiber folding: requirement for the histone H4 N-terminal tail.* J Mol Biol, 2003. **327**(1): p. 85-96.

221. Arents, G. and E.N. Moudrianakis, *The histone fold: a ubiquitous architectural motif utilized in DNA compaction and protein dimerization.* Proc Natl Acad Sci U S A, 1995. **92**(24): p. 11170-4.

222. Kouzarides, T., *Chromatin modifications and their function.* Cell, 2007. **128**(4): p. 693-705.

223. Strahl, B.D. and C.D. Allis, *The language of covalent histone modifications.* Nature, 2000. **403**(6765): p. 41-5.

224. Bannister, A.J. and T. Kouzarides, *Regulation of chromatin by histone modifications.* Cell Res, 2011. **21**(3): p. 381-95.

225. Redon, C., et al., *Histone H2A variants H2AX and H2AZ.* Curr Opin Genet Dev, 2002. **12**(2): p. 162-9.

226. Adam, M., et al., *H2A.Z is required for global chromatin integrity and for recruitment of RNA polymerase II under specific conditions.* Mol Cell Biol, 2001. **21**(18): p. 6270-9.

227. Rogakou, E.P., et al., *DNA double-stranded breaks induce histone H2AX phosphorylation on serine 139.* J Biol Chem, 1998. **273**(10): p. 5858-68.

228. Simpson, R.T., *Structure of the chromatosome, a chromatin particle containing 160 base pairs of DNA and all the histones.* Biochemistry, 1978. **17**(25): p. 5524-31.

229. Fan, Y., et al., *Histone H1 depletion in mammals alters global chromatin structure but causes specific changes in gene regulation.* Cell, 2005. **123**(7): p. 1199-212.

230. Shen, X. and M.A. Gorovsky, *Linker histone H1 regulates specific gene expression but not global transcription in vivo.* Cell, 1996. **86**(3): p. 475-83.

231. Schwartz, U., et al., *Characterizing the nuclease accessibility of DNA in human cells to map higher order structures of chromatin.* Nucleic Acids Res, 2019. **47**(3): p. 1239-1254.

232. Valouev, A., et al., *Determinants of nucleosome organization in primary human cells.* Nature, 2011. **474**(7352): p. 516-20.

233. Ma, Y., K. Kanakousaki, and L. Buttitta, *How the cell cycle impacts chromatin architecture and influences cell fate.* Front Genet, 2015. **6**: p. 19.

234. Dixon, J.R., D.U. Gorkin, and B. Ren, *Chromatin Domains: The Unit of Chromosome Organization.* Mol Cell, 2016. **62**(5): p. 668-80.

235. Dixon, J.R., et al., *Topological domains in mammalian genomes identified by analysis of chromatin interactions.* Nature, 2012. **485**(7398): p. 376-80.

236. Nora, E.P., et al., *Spatial partitioning of the regulatory landscape of the X-inactivation centre.* Nature, 2012. **485**(7398): p. 381-5.

237. Allshire, R.C. and H.D. Madhani, *Ten principles of heterochromatin formation and function.* Nat Rev Mol Cell Biol, 2018. **19**(4): p. 229-244.

238. Ahringer, J. and S.M. Gasser, *Repressive Chromatin in Caenorhabditis elegans: Establishment, Composition, and Function.* Genetics, 2018. **208**(2): p. 491-511.

239. Lai, W.K.M. and B.F. Pugh, *Understanding nucleosome dynamics and their links to gene expression and DNA replication.* Nat Rev Mol Cell Biol, 2017. **18**(9): p. 548-562.

240. Jiang, C. and B.F. Pugh, *Nucleosome positioning and gene regulation: advances through genomics.* Nat Rev Genet, 2009. **10**(3): p. 161-72.

241. Clapier, C.R., et al., *Mechanisms of action and regulation of ATP-dependent chromatin-remodelling complexes.* Nat Rev Mol Cell Biol, 2017. **18**(7): p. 407-422.

242. Yuan, G.C., et al., *Genome-scale identification of nucleosome positions in S. cerevisiae.* Science, 2005. **309**(5734): p. 626-30.

243. Lee, C.K., et al., *Evidence for nucleosome depletion at active regulatory regions genome-wide.* Nat Genet, 2004. **36**(8): p. 900-5.

244. Eaton, M.L., et al., *Conserved nucleosome positioning defines replication origins.* Genes Dev, 2010. **24**(8): p. 748-53.

245. Schones, D.E., et al., *Dynamic regulation of nucleosome positioning in the human genome.* Cell, 2008. **132**(5): p. 887-98.

246. Weber, C.M., S. Ramachandran, and S. Henikoff, *Nucleosomes are context-specific, H2A.Z-modulated barriers to RNA polymerase.* Mol Cell, 2014. **53**(5): p. 819-30.

247. Zhang, Z., et al., *A packing mechanism for nucleosome organization reconstituted across a eukaryotic genome.* Science, 2011. **332**(6032): p. 977-80.

248. Oberbeckmann, E., et al., *Genome information processing by the INO80 chromatin remodeler positions nucleosomes.* Nat Commun, 2021. **12**(1): p. 3231.

249. Segal, E., et al., *A genomic code for nucleosome positioning.* Nature, 2006. **442**(7104): p. 772-8.

250. Krietenstein, N., et al., *Genomic Nucleosome Organization Reconstituted with Pure Proteins.* Cell, 2016. **167**(3): p. 709-721.e12.

251. Yen, K., et al., *Genome-wide nucleosome specificity and directionality of chromatin remodelers.* Cell, 2012. **149**(7): p. 1461-73.

252. Clapier, C.R. and B.R. Cairns, *The biology of chromatin remodeling complexes.* Annu Rev Biochem, 2009. **78**: p. 273-304.

253. Flaus, A., et al., *Identification of multiple distinct Snf2 subfamilies with conserved structural motifs.* Nucleic Acids Res, 2006. **34**(10): p. 2887-905.

254. Yan, L. and Z. Chen, *A Unifying Mechanism of DNA Translocation Underlying Chromatin Remodeling.* Trends Biochem Sci, 2020. **45**(3): p. 217-227.

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

256. Ito, T., et al., *ACF, an ISWI-containing and ATP-utilizing chromatin assembly and remodeling factor.* Cell, 1997. **90**(1): p. 145-55.

257. Längst, G. and P.B. Becker, *Nucleosome mobilization and positioning by ISWI-containing chromatin-remodeling factors.* J Cell Sci, 2001. **114**(Pt 14): p. 2561-8.

258. Clapier, C.R. and B.R. Cairns, *Regulation of ISWI involves inhibitory modules antagonized by nucleosomal epitopes.* Nature, 2012. **492**(7428): p. 280-4.

259. Grüne, T., et al., *Crystal structure and functional analysis of a nucleosome recognition module of the remodeling factor ISWI.* Mol Cell, 2003. **12**(2): p. 449-60.

260. Lusser, A., D.L. Urwin, and J.T. Kadonaga, *Distinct activities of CHD1 and ACF in ATP-dependent chromatin assembly.* Nat Struct Mol Biol, 2005. **12**(2): p. 160-6.

261. Pointner, J., et al., *CHD1 remodelers regulate nucleosome spacing in vitro and align nucleosomal arrays over gene coding regions in S. pombe.* Embo j, 2012. **31**(23): p. 4388-403.

262. Konev, A.Y., et al., *CHD1 motor protein is required for deposition of histone variant H3.3 into chromatin in vivo.* Science, 2007. **317**(5841): p. 1087-90.

263. Flanagan, J.F., et al., *Double chromodomains cooperate to recognize the methylated histone H3 tail.* Nature, 2005. **438**(7071): p. 1181-5.

264. Farnung, L., et al., *Nucleosome-Chd1 structure and implications for chromatin remodelling.* Nature, 2017. **550**(7677): p. 539-542.

265. Szerlong, H., et al., *The HSA domain binds nuclear actin-related proteins to regulate chromatin-remodeling ATPases.* Nat Struct Mol Biol, 2008. **15**(5): p. 469-76.

266. Mohrmann, L. and C.P. Verrijzer, *Composition and functional specificity of SWI2/SNF2 class chromatin remodeling complexes.* Biochim Biophys Acta, 2005. **1681**(2-3): p. 59-73.

267. Whitehouse, I., et al., *Nucleosome mobilization catalysed by the yeast SWI/SNF complex.* Nature, 1999. **400**(6746): p. 784-7.

268. Cairns, B.R., et al., *RSC, an essential, abundant chromatin-remodeling complex.* Cell, 1996. **87**(7): p. 1249-60.

269. Hassan, A.H., et al., *Function and selectivity of bromodomains in anchoring chromatin-modifying complexes to promoter nucleosomes.* Cell, 2002. **111**(3): p. 369-79.

270. Ho, L. and G.R. Crabtree, *Chromatin remodelling during development.* Nature, 2010. **463**(7280): p. 474-84.

271. Kadoch, C., et al., *Proteomic and bioinformatic analysis of mammalian SWI/SNF complexes identifies extensive roles in human malignancy.* Nat Genet, 2013. **45**(6): p. 592-601.

272. Shen, X., et al., *Involvement of actin-related proteins in ATP-dependent chromatin remodeling.* Mol Cell, 2003. **12**(1): p. 147-55.

273. Udugama, M., A. Sabri, and B. Bartholomew, *The INO80 ATP-dependent chromatin remodeling complex is a nucleosome spacing factor.* Mol Cell Biol, 2011. **31**(4): p. 662-73.

274. Mizuguchi, G., et al., *ATP-driven exchange of histone H2AZ variant catalyzed by SWR1 chromatin remodeling complex.* Science, 2004. **303**(5656): p. 343-8.

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

276. Willhoft, O., et al., *Structure and dynamics of the yeast SWR1-nucleosome complex.* Science, 2018. **362**(6411).

277. Willhoft, O. and D.B. Wigley, *INO80 and SWR1 complexes: the non-identical twins of chromatin remodelling.* Curr Opin Struct Biol, 2020. **61**: p. 50-58.

278. Brahma, S., et al., *INO80 exchanges H2A.Z for H2A by translocating on DNA proximal to histone dimers.* Nat Commun, 2017. **8**: p. 15616.

279. Papamichos-Chronakis, M., et al., *Global regulation of H2A.Z localization by the INO80 chromatin-remodeling enzyme is essential for genome integrity.* Cell, 2011. **144**(2): p. 200-13.

280. Watanabe, S., et al., *A histone acetylation switch regulates H2A.Z deposition by the SWR-C remodeling enzyme.* Science, 2013. **340**(6129): p. 195-9.

281. Wang, F., et al., *Comment on "A histone acetylation switch regulates H2A.Z deposition by the SWR-C remodeling enzyme".* Science, 2016. **353**(6297): p. 358.

282. Ayala, R., et al., *Structure and regulation of the human INO80-nucleosome complex.* Nature, 2018. **556**(7701): p. 391-395.

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

284. Liu, X., et al., *Mechanism of chromatin remodelling revealed by the Snf2-nucleosome structure.* Nature, 2017. **544**(7651): p. 440-445.

285. Yan, L., et al., *Structures of the ISWI-nucleosome complex reveal a conserved mechanism of chromatin remodeling.* Nat Struct Mol Biol, 2019. **26**(4): p. 258-266.

286. Ye, Y., et al., *Structure of the RSC complex bound to the nucleosome.* Science, 2019. **366**(6467): p. 838-843.

287. Wagner, F.R., et al., *Structure of SWI/SNF chromatin remodeller RSC bound to a nucleosome.* Nature, 2020. **579**(7799): p. 448-451.

288. Han, Y., et al., *Cryo-EM structure of SWI/SNF complex bound to a nucleosome.* Nature, 2020. **579**(7799): p. 452-455.

289. He, S., et al., *Structure of nucleosome-bound human BAF complex.* Science, 2020. **367**(6480): p. 875-881.

290. Yuan, J., et al., *Structure of human chromatin-remodelling PBAF complex bound to a nucleosome.* Nature, 2022. **605**(7908): p. 166-171.

291. Bowman, G.D., *Mechanisms of ATP-dependent nucleosome sliding.* Curr Opin Struct Biol, 2010. **20**(1): p. 73-81.

292. Mueller-Planitz, F., H. Klinker, and P.B. Becker, *Nucleosome sliding mechanisms: new twists in a looped history.* Nat Struct Mol Biol, 2013. **20**(9): p. 1026-32.

293. Poli, J., S.M. Gasser, and M. Papamichos-Chronakis, *The INO80 remodeller in transcription, replication and repair.* Philos Trans R Soc Lond B Biol Sci, 2017. **372**(1731).

294. Shen, X., et al., *A chromatin remodelling complex involved in transcription and DNA processing.* Nature, 2000. **406**(6795): p. 541-4.

295. van Attikum, H., et al., *Recruitment of the INO80 complex by H2A phosphorylation links ATP-dependent chromatin remodeling with DNA double-strand break repair.* Cell, 2004. **119**(6): p. 777-88.

296. Yao, W., et al., *The INO80 Complex Requires the Arp5-Ies6 Subcomplex for Chromatin Remodeling and Metabolic Regulation.* Mol Cell Biol, 2016. **36**(6): p. 979-91.

297. Ebbert, R., A. Birkmann, and H.J. Schüller, *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.* Mol Microbiol, 1999. **32**(4): p. 741-51.

298. Yen, K., V. Vinayachandran, and B.F. Pugh, *SWR-C and INO80 chromatin remodelers recognize nucleosome-free regions near +1 nucleosomes.* Cell, 2013. **154**(6): p. 1246-56.

299. Alcid, E.A. and T. Tsukiyama, *ATP-dependent chromatin remodeling shapes the long noncoding RNA landscape.* Genes Dev, 2014. **28**(21): p. 2348-60.

300. Shimada, K., et al., *Ino80 chromatin remodeling complex promotes recovery of stalled replication forks.* Curr Biol, 2008. **18**(8): p. 566-75.

301. Vincent, J.A., T.J. Kwong, and T. Tsukiyama, *ATP-dependent chromatin remodeling shapes the DNA replication landscape.* Nat Struct Mol Biol, 2008. **15**(5): p. 477-84.

302. Papamichos-Chronakis, M. and C.L. Peterson, *The Ino80 chromatin-remodeling enzyme regulates replisome function and stability.* Nat Struct Mol Biol, 2008. **15**(4): p. 338-45.

303. Poli, J., et al., *Mec1, INO80, and the PAF1 complex cooperate to limit transcription replication conflicts through RNAPII removal during replication stress.* Genes Dev, 2016. **30**(3): p. 337-54.

304. Lafon, A., et al., *INO80 Chromatin Remodeler Facilitates Release of RNA Polymerase II from Chromatin for Ubiquitin-Mediated Proteasomal Degradation.* Mol Cell, 2015. **60**(5): p. 784-796.

305. Morrison, A.J., et al., *INO80 and gamma-H2AX interaction links ATP-dependent chromatin remodeling to DNA damage repair.* Cell, 2004. **119**(6): p. 767-75.

306. Lademann, C.A., et al., *The INO80 Complex Removes H2A.Z to Promote Presynaptic Filament Formation during Homologous Recombination.* Cell Rep, 2017. **19**(7): p. 1294-1303.

307. Alatwi, H.E. and J.A. Downs, *Removal of H2A.Z by INO80 promotes homologous recombination.* EMBO Rep, 2015. **16**(8): p. 986-94.

308. Chen, L., et al., *Subunit organization of the human INO80 chromatin remodeling complex: an evolutionarily conserved core complex catalyzes ATP-dependent nucleosome remodeling.* J Biol Chem, 2011. **286**(13): p. 11283-9.

309. Tosi, A., et al., *Structure and subunit topology of the INO80 chromatin remodeler and its nucleosome complex.* Cell, 2013. **154**(6): p. 1207-19.

310. Chen, L., R.C. Conaway, and J.W. Conaway, *Multiple modes of regulation of the human Ino80 SNF2 ATPase by subunits of the INO80 chromatin-remodeling complex.* Proc Natl Acad Sci U S A, 2013. **110**(51): p. 20497-502.

311. Knoll, K.R., et al., *The nuclear actin-containing Arp8 module is a linker DNA sensor driving INO80 chromatin remodeling.* Nat Struct Mol Biol, 2018. **25**(9): p. 823-832.

312. Brahma, S., et al., *The Arp8 and Arp4 module acts as a DNA sensor controlling INO80 chromatin remodeling.* Nat Commun, 2018. **9**(1): p. 3309.

313. Oberbeckmann, E., et al., *Ruler elements in chromatin remodelers set nucleosome array spacing and phasing.* Nat Commun, 2021. **12**(1): p. 3232.

314. Olave, I.A., S.L. Reck-Peterson, and G.R. Crabtree, *Nuclear actin and actin-related proteins in chromatin remodeling.* Annu Rev Biochem, 2002. **71**: p. 755-81.

315. Cao, T., et al., *Crystal structure of a nuclear actin ternary complex.* Proc Natl Acad Sci U S A, 2016. **113**(32): p. 8985-90.

316. Szerlong, H., A. Saha, and B.R. Cairns, *The nuclear actin-related proteins Arp7 and Arp9: a dimeric module that cooperates with architectural proteins for chromatin remodeling.* Embo j, 2003. **22**(12): p. 3175-87.

317. Schubert, H.L., et al., *Structure of an actin-related subcomplex of the SWI/SNF chromatin remodeler.* Proc Natl Acad Sci U S A, 2013. **110**(9): p. 3345-50.

318. Turegun, B., et al., *Actin-related proteins regulate the RSC chromatin remodeler by weakening intramolecular interactions of the Sth1 ATPase.* Commun Biol, 2018. **1**: p. 1.

319. Clapier, C.R., et al., *Regulation of DNA Translocation Efficiency within the Chromatin Remodeler RSC/Sth1 Potentiates Nucleosome Sliding and Ejection.* Mol Cell, 2016. **62**(3): p. 453-461.

320. Zhou, C.Y., et al., *The Yeast INO80 Complex Operates as a Tunable DNA Length-Sensitive Switch to Regulate Nucleosome Sliding.* Mol Cell, 2018. **69**(4): p. 677-688.e9.

321. Ray, S. and A. Grove, *The yeast high mobility group protein HMO2, a subunit of the chromatin-remodeling complex INO80, binds DNA ends.* Nucleic Acids Res, 2009. **37**(19): p. 6389-99.

322. Metzner, F.J., et al., *Structural and biochemical characterization of human Schlafen 5.* Nucleic Acids Res, 2022. **50**(2): p. 1147-1161.

323. Jumper, J., et al., *Highly accurate protein structure prediction with AlphaFold.* Nature, 2021. **596**(7873): p. 583-589.

324. Metzner, F.J., et al., *Mechanistic understanding of human SLFN11.* Nat Commun, 2022. **13**(1): p. 5464.

325. Mirdita, M., et al., *ColabFold: making protein folding accessible to all.* Nat Methods, 2022. **19**(6): p. 679-682.

326. Yang, W., *Nucleases: diversity of structure, function and mechanism.* Q Rev Biophys, 2011. **44**(1): p. 1-93.

327. Kimura, S., et al., *Template-dependent nucleotide addition in the reverse (3'-5') direction by Thg1-like protein.* Sci Adv, 2016. **2**(3): p. e1501397.

328. Kobayashi-Ishihara, M., et al., *Schlafen 12 restricts HIV-1 latency reversal by a codon-usage dependent post-transcriptional block in CD4+ T cells.* bioRxiv, 2022: p. 2022.11.09.512793.

329. Chen, F., et al., *Dissimilation of synonymous codon usage bias in virus-host coevolution due to translational selection.* Nat Ecol Evol, 2020. **4**(4): p. 589-600.

330. Schwenzer, H., et al., *Oxidative Stress Triggers Selective tRNA Retrograde Transport in Human Cells during the Integrated Stress Response.* Cell Rep, 2019. **26**(12): p. 3416-3428.e5.

331. Chatterjee, K., et al., *tRNA dynamics between the nucleus, cytoplasm and mitochondrial surface: Location, location, location.* Biochim Biophys Acta Gene Regul Mech, 2018. **1861**(4): p. 373-386.

332. Liu, B., et al., *Deciphering the tRNA-derived small RNAs: origin, development, and future.* Cell Death Dis, 2021. **13**(1): p. 24.

333. Varadi, M., et al., *AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models.* Nucleic Acids Res, 2022. **50**(D1): p. D439-d444.

334. Peng, S., et al., *Structure and function of an effector domain in antiviral factors and tumor suppressors SAMD9 and SAMD9L.* Proc Natl Acad Sci U S A, 2022. **119**(4).

335. Holm, L., *Dali server: structural unification of protein families.* Nucleic Acids Res, 2022. **50**(W1): p. W210-5.

336. Podvalnaya, N., et al., *piRNA processing by a trimeric Schlafen-domain nuclease.* bioRxiv, 2023: p. 2023.01.19.524756.

337. Legrand, A., et al., *The inflammatory and tumor suppressor SAMD9L acts through a Schlafen-like box to restrict HIV and inhibit cell translation in SAAD/ATXPC.* bioRxiv, 2023: p. 2023.01.19.524725.

338. Meng, X., et al., *A paralogous pair of mammalian host restriction factors form a critical host barrier against poxvirus infection.* PLoS Pathog, 2018. **14**(2): p. e1006884.

339. Horton, J.R. and X. Cheng, *PvuII endonuclease contains two calcium ions in active sites.* J Mol Biol, 2000. **300**(5): p. 1049-56.

340. Kostrewa, D. and F.K. Winkler, *Mg2+ binding to the active site of EcoRV endonuclease: a crystallographic study of complexes with substrate and product DNA at 2 A resolution.* Biochemistry, 1995. **34**(2): p. 683-96.

341. Pingoud, A., et al., *Type II restriction endonucleases: structure and mechanism.* Cell Mol Life Sci, 2005. **62**(6): p. 685-707.

342. Pingoud, A., G.G. Wilson, and W. Wende, *Type II restriction endonucleases--a historical perspective and more.* Nucleic Acids Res, 2014. **42**(12): p. 7489-527.

343. Jeltsch, A., et al., *Substrate-assisted catalysis in the cleavage of DNA by the EcoRI and EcoRV restriction enzymes.* Proc Natl Acad Sci U S A, 1993. **90**(18): p. 8499-503.

344. Groll, D.H., et al., *Does the restriction endonuclease EcoRV employ a two-metal-Ion mechanism for DNA cleavage?* Biochemistry, 1997. **36**(38): p. 11389-401.

345. Spyridaki, A., et al., *Structural and biochemical characterization of a new Mg(2+) binding site near Tyr94 in the restriction endonuclease PvuII.* J Mol Biol, 2003. **331**(2): p. 395-406.

346. Fairman-Williams, M.E., U.P. Guenther, and E. Jankowsky, *SF1 and SF2 helicases: family matters.* Curr Opin Struct Biol, 2010. **20**(3): p. 313-24.

347. Tanner, N.K., et al., *The Q motif: a newly identified motif in DEAD box helicases may regulate ATP binding and hydrolysis.* Mol Cell, 2003. **11**(1): p. 127-38.

348. Johnson, J.L., et al., *An atlas of substrate specificities for the human serine/threonine kinome.* Nature, 2023. **613**(7945): p. 759-766.

349. Vit, G., et al., *Human SLFN5 and its Xenopus Laevis ortholog regulate entry into mitosis and oocyte meiotic resumption.* Cell Death Discov, 2022. **8**(1): p. 484.

350. Evans, R., et al., *Protein complex prediction with AlphaFold-Multimer.* bioRxiv, 2021: p. 2021.10.04.463034.

351. Qiu, Y., et al., *The Chd1 Chromatin Remodeler Shifts Nucleosomal DNA Bidirectionally as a Monomer.* Mol Cell, 2017. **68**(1): p. 76-88.e6.

352. Kunert, F., et al., *Structural mechanism of extranucleosomal DNA readout by the INO80 complex.* Sci Adv, 2022. **8**(49): p. eadd3189.

353. Zukin, S.A., et al., *Structure and flexibility of the yeast NuA4 histone acetyltransferase complex.* Elife, 2022. **11**.

354. Baker, R.W., et al., *Structural insights into assembly and function of the RSC chromatin remodeling complex.* Nat Struct Mol Biol, 2021. **28**(1): p. 71-80.

355. Mashtalir, N., et al., *Modular Organization and Assembly of SWI/SNF Family Chromatin Remodeling Complexes.* Cell, 2018. **175**(5): p. 1272-1288.e20.

356. Cai, Y., et al., *YY1 functions with INO80 to activate transcription.* Nat Struct Mol Biol, 2007. **14**(9): p. 872-4.

357. Wilkinson, F.H., K. Park, and M.L. Atchison, *Polycomb recruitment to DNA in vivo by the YY1 REPO domain.* Proc Natl Acad Sci U S A, 2006. **103**(51): p. 19296-301.

358. Alfieri, C., et al., *Structural basis for targeting the chromatin repressor Sfmbt to Polycomb response elements.* Genes Dev, 2013. **27**(21): p. 2367-79.

359. Gómez-Zambrano, Á., et al., *Arabidopsis SWC4 Binds DNA and Recruits the SWR1 Complex to Modulate Histone H2A.Z Deposition at Key Regulatory Genes.* Mol Plant, 2018. **11**(6): p. 815-832.

360. Morrison, A.J., et al., *Mec1/Tel1 phosphorylation of the INO80 chromatin remodeling complex influences DNA damage checkpoint responses.* Cell, 2007. **130**(3): p. 499-511.

361. Chen, G., et al., *Taf14 recognizes a common motif in transcriptional machineries and facilitates their clustering by phase separation.* Nat Commun, 2020. **11**(1): p. 4206.

362. Klein, B.J., et al., *Structural insights into the π-π-π stacking mechanism and DNA-binding activity of the YEATS domain.* Nat Commun, 2018. **9**(1): p. 4574.

363. Davey, C.A., et al., *Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 a resolution.* J Mol Biol, 2002. **319**(5): p. 1097-113.

364. Suter, B., G. Schnappauf, and F. Thoma, *Poly(dA.dT) sequences exist as rigid DNA structures in nucleosome-free yeast promoters in vivo.* Nucleic Acids Res, 2000. **28**(21): p. 4083-9.

365. Barnes, T. and P. Korber, *The Active Mechanism of Nucleosome Depletion by Poly(dA:dT) Tracts In Vivo.* Int J Mol Sci, 2021. **22**(15).

Acknowledgements

First, I would like to thank my research supervisors, Prof. Dr. Karl-Peter Hopfner and Dr. Katja Lammens, for giving me the opportunity to work on these exciting projects, for providing valuable advice and supervision and for giving me the freedom to develop and follow my own ideas.

I want to thank all former and present members of the SLFN subgroup (Lisi, Simon, and Michael) for the great teamwork and the productive work environment. In particular, I would like to thank Dr. Michael Kugler for inspiring scientific discussions and Simon Wenzl for his great work during the preparation of his master's thesis. I want to extend this thank-you to all former and present members of the INO80 subgroup (Franzi, Stephan, Manu, Vanessa, Mariia, Sebastian, Dirk, James, Priyanka, Manmohan, and Alberto). In particular, I want to thank Franziska Kunert and Stephan Woike for the excellent work environment and the great time in our INO80 bay and Manuela Moldt for support with cloning and protein expression.

I would like to thank Dr. Katja Lammens and Dr. Joseph Bartho for support with the microscopes and PD Dr. Gregor Witte for organizing the biophysics facility. I also want to thank Brigitte Keßler, Manuela Moldt, Alexandra Schele and Olga Fettscher for the excellent organization of the lab.

I am very grateful to all former and present members of the Hopfner lab for the great and supportive working atmosphere and for making the lab such a special place.

I also want to thank our collaboration partners, especially Prof. Dr. Boris Pfander together with Dr. Markus Höpfler and Leonhard Karl, Dr. Sebastian Eustermann, Dr. Jia-Xuan Chen, Dr. Stefan Krebs, Dr. Thomas Fröhlich, and Dr. Daniel Bollschweiler.

I want to thank Dr. Katja Lammens, Dr. Michael Kugler and Franziska Kunert for the proofreading of this thesis.

I would like to thank my second reviewer Prof. Dr. Elena Conti and the co-reviewers Prof. Dr. Julian Stingele, Prof. Dr. Klaus Förstemann, PD Dr. Gregor Witte, and Prof. Dr. Johannes Stigler for reviewing this thesis.

Furthermore, I want to thank the SFB1054 and the IRTG for the financial and scientific support.

Most importantly, I would like to thank my parents, my sister, Elaine and my family and friends for their unlimited support and love, which reminds me of what truly matters.