UCSF
UC San Francisco Electronic Theses and Dissertations

Title
The mechanisms of assembly and DNA length sensing by the chromatin remodeling complex INO80.

Permalink
https://escholarship.org/uc/item/44c6q1gm

Author
Zhou, Coral Yishan

Publication Date
2017

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

The mechanisms of assembly and DNA length sensing by the chromatin remodeling complex INO80

by

Coral Yishan Zhou

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Biochemistry and Molecular Biology

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA, SAN FRANCISCO

Acknowledgements:

I dedicate this thesis to a few key individuals, without whom this would not have been possible.

To Geeta, my advisor and confidant, thank you for allowing me to be myself, both scientifically and personally. Also, thank you and Alok for bringing your amazing daughter Kalpana into this world, who reminds me every time I see her that having kids can actually be cool.

To all of my labmates over the years, who were my mentors, my closest friends, my shoulders to cry on, my voices of reason, my buds, my eyes and ears. A special thank you to Lisa and Daniele, who will always be my scientific heroes.

To my wonderful collaborators, Mike Trnka, Ignacia Echeverria, Caitlin Stoddard, and Eugene Palovcak, for helping me making science more fun than I could have ever imagined.

To Adam, who taught me how to paint with all the colors of the rainbow.

To Lindsey, my beautiful rock from Canton, for being the yin to my yang.

To my parents, who made me who I am today.

Acknowledgement of previously published work in this thesis.

Chapter 2 of this thesis was published in the following scientific journal:

Zhou, C.Y., Stoddard, C.I., Johnston, J.B., Trnka, M.J., Echeverria, I., Palovcak, E., Sali, A., Burlingame, A.L., Cheng, Y., and Narlikar, G.J. (2017). Regulation of Rvb1/Rvb2 by a Domain within the INO80 Chromatin Remodeling Complex Implicates the Yeast Rvbs as Protein Assembly Chaperones. *Cell Rep* **19**, 2033–2044.

Abstract:

The mechanisms of assembly and DNA length sensing by the chromatin remodeling complex INO80.

By

Coral Y. Zhou

The fundamental basis of gene regulation is the physical accessibility of genetic material to downstream machinery that convert DNA into proteins. In eukaryotes, accessibility to the genome is ultimately regulated by the structure of the nucleosome—the fundamental unit of chromatin that is comprised of 147 base pairs of DNA tightly wrapped around a core of eight histone proteins. Of the many factors that affect nucleosome structure, ATP-dependent chromatin remodeling enzymes are unique in their ability to convert the chemical energy of ATP into physical changes in the nucleosome that would otherwise be energetically disfavorable. Of these remodelers, the INO80 remodeling enzyme is a 15-subunit complex from yeast that is unique both in its modular structure and its critical roles at sites of genomic instability such as double strand breaks and replication forks. However, very little is known about how the INO80 machinery is assembled or how it achieves specificity in nucleosome movement. In the first half of this thesis, we explore the potential roles of Rvb1 and Rvb2 (Rvbs), two poorly understood AAA+ ATPases that are integral components of the INO80 complex, as well as other multi-

subunit complexes in the cell. While the Rvbs were traditionally thought of as DNA helicases, we have new evidence for their role as protein chaperones. We discover the first protein client of Rvbs, a small insertion region of the Ino80 ATPase (Ino80INS), which robustly stimulates the ATPase activity of the Rvbs while promoting dodecamerization of the normally hexameric Rvbs. Using a combination of structural methods including crosslinking mass spectrometry, native mass spectrometry, cryo-EM and integrative modeling, we find that two Ino80INS molecules bind asymmetrically and adjacent to one another along the dodecamerization interface, resulting in a conformationally flexible dodecamer that collapses into hexamers upon ATP addition. Our results demonstrate the chaperone-like potential of Rvb1/Rvb2 and suggest a model where binding of multiple clients like Ino80 stimulates ATP-driven cycling between hexamers and dodecamers, providing iterative opportunities for correct subunit assembly. In the second half of the thesis, we explore the molecular mechanism of nucleosome movement by INO80. While it is known that INO80 can center nucleosomes on a short piece of DNA, it is unknown how the motor accomplishes this task. Using a combination of enzymology and single molecule FRET, we find that INO80 exhibits a switch-like response to flanking DNA length, largely regulated by subcomplex of INO80 called the Nhp10 module. We also find that remodeling by INO80 is extremely processive: once sliding is initiated, INO80 moves the nucleosome rapidly at least 20 bp without re-assessing flanking DNA length, and can change the direction of nucleosome sliding without dissociation. Taken together, these findings provide unprecedented insight into INO80's lifetime in the nucleus: from its assembly by Rvb1/Rvb2 to its action on chromatin.

Table of Contents:

Chapter 1: Introduction 1

References 22

Chapter 2: Regulation of Rvb1/Rvb2 by a domain within the INO80 chromatin remodeling complex implicates the yeast Rvbs as protein assembly chaperones 38

Abstract 39

Introduction 39

Methods 41

Results 45

Discussion 56

References 61

Supplemental references 112

Appendix to Chapter 2: unpublished data 123

Methods 124

Results 126

Discussion 133

References 136

Chapter 3: The yeast INO80 complex operates as a tunable DNA-length sensitive Switch for regulating nucleosome sliding 150

Abstract 151

Introduction 151

Methods 153

Results 157

Discussion 168

References 174

Appendix to Chapter 3 204

Methods 205

Results 205

Discussion 207

List of Tables:

Chapter 2

Table S1 (related to Figure 1 and Figure S1): Summary of XL-MS data for Rvb1/Rvb2 and Rvb1/Rvb2/Ino80INS 94

Table S2 (related to Figure 2, Figure S2): Analysis of the Rvb1/Rvb2/Ino80INS complex by native mass spectrometry 95

Table S3 (related to Figure 3, Figure S3): Predicted molecular weights for different stoichiometries of Rvb1/Rvb2 hexamers 96

Table S4A (related to Figure 4 and Table S3): Molecular mass distribution of Rvb1/Rvb2 hexamers by native mass spectrometry 97

Table S4B (related to Figure 4 and Table S3): Molecular mass distribution of Rvb1/Rvb2 dodecamers by native mass spectrometry 98

Table S5 (related to Figure 4 and Figure S4): Molecular weight predictions from MALS Experiments 99

Table S6 (related to Figure 6, S6, S7): Summary of integrative structure determination of the Rvb1/Rvb2/Ino80INS complex 100

List of Figures:

Chapter 1

Figure 1: Case studies of nucleic acid motors 33

Figure 2: General properties of nucleosomes.....35

Figure 3: INO80 family of remodeling motors 36

Figure 4: Remodeling motors act in conjunction *in vivo* 37

Chapter 2

Figure 1: The insertion domain of the Ino80 ATPase (Ino80INS) is a direct interactor and activator of Rvb1/Rvb2 72

Figure S1 (related to Figure 1): XL-MS of Rvb1/Rvb2 and Rvb1/Rvb2/Ino80INS 74

Figure 2: The Ino80 insertion promotes dodecamerization of Rvb1/Rvb2.....75

Figure S2 (related to Figure 2, Table S2): Native-MS and negative stain EM of the Rvb1/Rvb2/Ino80INS complex 77

Figure 3: Nucleotide state drives the switch between hexameric and dodecameric forms of the Rvb1/Rvb2 complex.....79

Figure S3 (related to Figure 3, Table S3, and Table S4): Native-MS of Rvb1/Rvb2 complexes.....81

Figure 4: The ATPase activity of the Rvb1/Rvb2 complex is enhanced by Dodecamerization.....83

Figure S4 (related to Figure 4 and Table S5): SEC-MALS of Rvb1/Rvb2 complexes.....85

Figure 5: Cryo-EM of the Rvb1/Rvb2/Ino80INS complex reveals conformational heterogeneity 86

Figure S5 (related to Figure 5): Cryo-EM recontruction of the Rvb1/Rvb2/Ino80INS complex.....88

Figure 6: Ino80INS binds to Rvb1/Rvb2 dodecamer asymmetrically 89

Figure S6 (related to Figure 6, Figure S7, Table S6): The 4-stage scheme for integrative structure modeling of the Rvb1/Rvb2/Ino80INS complex 90

Figure S7 (related to Figure 6, Figure S6, Table S6): Analysis of integrative structure modeling of Rvb1/Rvb2/Ino80INS complex.....91

Figure 7: Model for activation of Rvb1/Rvb2 by client proteins.....93

Appendix to Chapter 2

Appendix Figure 1: Sequence and structure of hexameric yeast Rvb1 and Rvb2 138

Appendix Figure 2: Nucleotide binding by WT Rvb1/Rvb2 detected by native-MS.....139

Appendix Figure 3: Nucleotide binding profiles of different Walker B mutants of Rvb1/Rvb2 by native-MS.....140

Appendix Figure 4: Comparative spectral counting analysis of Rvb1/Rvb2/Ino80INS reveals major changes in DII domains when Rvbs are bound to the Ino80INS.....141

Appendix Figure 5: The Ino80INS pushes the DII domains of Rvb1/Rvb2 toward a conformation that is more compatible with dodecamerization.....143

Appendix Figure 6: Fractionation of Rvb1/Rvb2/Ino80INS reveals a gradient of different specific activities.....144

Appendix Figure 7: The Swr1 insertion (Swr1INS) also interacts directly with Rvb1/Rvb2, enhancing its ATPase activity while promoting dodecamerization of the Rvbs.....145

Appendix Figure 8: MBP-Ies2 and MBP-Ies6 do not co-purify with Rvb1/Rvb2, but do stimulate ATPase activity of Rvbs.....147

Appendix Figure 9: Results of labeling and exchanging bacterially expressed Rvb1 and Rvb2 into the INO80 complex purified from yeast.....148

Chapter 3: The yeast INO80 complex operates as a tunable DNA-length sensitive switch for regulating nucleosome sliding

Figure 1: Flanking DNA length modulates nucleosome sliding by INO80, but not ATPase activity.....180

Figure S1: DNA sequence affects final product distribution, but not remodeling rates, by INO80.....182

Figure 2: Remodeling by INO80 at the single nucleosome level is preceded by a long pause, followed by rapid nucleosome translocation 183

Figure S2: Relationship between FRET and bp flanking DNA on the Cy5-labeled exit side of the nucleosome.....185

Figure S3: Comparison of ensemble FRET and smFRET remodeling rates.....187

Figure S4: Additional characteristics of the pauses and the translocation phase.....189

Figure S5: Effect of DNA sequence on sliding of individual nucleosomes by INO80 192

Figure 3: INO80 processively and continuously moves a nucleosome bi-directionally on long DNA 193

Figure S6: Summary of mass spectrometry data of the Ino80 protein isolated from either WT (top) or *Δnhp10* (bottom) yeast 195

Figure S7: DNA sequences used in this work 196

Figure S8: Summary of the Traces smFRET analysis software 197

Figure S9: Schematic summary of the Slopey algorithm for quantifying translocation events.....198

Appendix to Chapter 3

Appendix Figure 1: The presence of 3' or 5' overhangs inhibits the ATPase activity of INO80 211

Appendix Figure 2: Discrepancy between remodeling rates by gel and by FRET.....212

Appendix Figure 3: INO80 has a 5-fold preference for nucleosomes over DNA 213

Chapter 1: Introduction

Chapter 1: Introduction

Packaging of DNA into chromatin provides the cell a means of regulating all processes that require access to the genetic material. The basic building block of chromatin consists of a nucleosome, which contains ~150 base pairs (bp) of DNA wrapped around an octamer composed of two copies each of histones H2A, H2B, H3, and H4. Multiple nucleosomes further interact with each other and other cellular factors to achieve higher order compaction (1). Since the primary consequences of nucleosome formation are the compaction and occlusion of the wrapped DNA, in early years chromatin was thought of mainly as a packaging material that presented a barrier to DNA-based processes (2). However, the discovery over the last two decades of a variety of enzymatic complexes, which catalyze both covalent and conformational changes in chromatin, has helped shape current views that chromatin also plays sophisticated regulatory roles (3).

Non-covalent conformational changes of chromatin are catalyzed by chromatin remodeling motors that use the energy of ATP hydrolysis to alter nucleosome structure (4). These motors carry out diverse transformations on chromatin, and play central roles in all nuclear processes including transcription, DNA replication and repair, and recombination. The recent identification of mutations in these motors as drivers of various cancers underscores the centrality of their biological roles (5, 6). In this review, we first summarize case studies of other well-studied nucleic acid motors to highlight general mechanistic principles, and then use these principles to discuss the current state of understanding of chromatin remodeling motors. We address the unique challenges faced by chromatin remodeling motors and the regulatory principles these motors may use to overcome these challenges.

Nucleic acid remodeling ATPases

Molecular motors play an integral role in nucleic acid biology. These motors include polymerases, which catalyze the formation of covalent bonds during DNA replication and transcription, as well as motors that catalyze non-covalent rearrangements of their nucleic acid substrates. Here we focus on the second class of motors, which are classified under the SF1 and SF2 superfamilies of ATPases (7). Members of the SF1 superfamily are *bona fide* helicases that processively translocate along single stranded DNA, while members of the SF2 family show a greater diversity of activities and include ATP-dependent chromatin remodeling motors (8). The core feature of both families is an ATPase domain containing two RecA-like lobes (see Sidebar and Figure 1A). A closer look at the RecA-fold architecture provides structural clues to how these motors couple ATP hydrolysis to conformational changes.

RecA is the ancestral nucleic acid remodeling motor

Bacterial RecA plays a key role in catalyzing many of the DNA rearrangements that occur during homologous recombination, a major pathway cells use to repair damaged DNA. The main activity of RecA is ATP-dependent oligomerization along single stranded DNA (reviewed in (9)). The core domain of RecA contains the prototypical RecA fold, which binds ATP and DNA, while the N- and C-terminal domains mediate oligomerization (10). RecA oligomerizes along DNA such that ATP binds at the interface between two RecA monomers (11). The structure of the active site in the presence of an ATP analog ADP-AlF_4 provides an elegant physical explanation for how the ATPase cycle is coupled to RecA oligomerization. ATP binding (Walker A) and hydrolysis (Walker B) motifs are provided by one monomer, while two lysine residues from the neighboring monomer stabilize the transition state mimicked by ADP-AlF_4. ATP hydrolysis causes disassembly of the RecA oligomer by weakening the monomer-monomer interface and DNA binding. Duplication of the RecA fold in SF1 and SF2 family proteins

converts the ATP-driven inter-molecular assembly cycle seen in the RecA oligomer into intra-molecular conformational changes between the two RecA-containing lobes to catalyze diverse structural rearrangements in nucleic acid substrates.

Structural modifications of RecA lead to diversity of function in chromatin remodeling motors

The two copies of the RecA fold in the SF1 and SF2 superfamily are asymmetric. The first copy retains the well-conserved ATP binding (Walker A) and ATP hydrolysis (Walker B) motifs found in the ancestral RecA protein, while the second copy has lost these motifs and instead supplies an arginine finger motif. These arginine fingers play a key role in ATP hydrolysis by stabilizing the transition state. A consequence of this split architecture is that closure of the two lobes is necessary for efficient ATP hydrolysis, providing a mechanism for these motors to couple ATP hydrolysis to movement (12). In current models of SF1 and SF2 motors, this movement of the RecA lobes relative to each other lies at the core of these motors' basic activities of translocation and rearrangement of nucleic acid structures. This general architecture of the ATPase domain is then further elaborated in different classes of motors depending on the type of reaction catalyzed (Figure 1B) (7). The elaborations include (*i*) a diverse set of N and C terminal domains that extend beyond the RecA-like lobes and (*ii*) sequence insertions and variations within the RecA-like lobes. These extra domains, insertions and variations regulate the ATPase domain of the motor and also allow for interactions with accessory subunits, thus providing an additional layer of specificity. The structural diversity of these ATPases and their respective accessory subunits is the basis for the functional diversity in the rearrangements they can catalyze.

Efforts to achieve a mechanistic understanding of nucleic acid motors raise three major questions:

1. How is the ATPase cycle coordinated with distortion of the substrate?
2. How is the activity of the core ATPase domain modulated either by accessory domains on the same protein or accessory subunits as part of a larger complex?
3. How does the motor achieve specificity for its biological substrate and directionality of output?

Next, we will illustrate each of these concepts through three case studies of well-studied nucleic acid motors. In the next section, we will then discuss how some of the basic strategies used in these case studies help us understand the more complex molecular transformations catalyzed by chromatin remodeling motors.

Mss116**:** Mss116 provides an excellent example of how a motor couples ATP binding and hydrolysis to distortion of its substrate. Mss116 (Fig. 1C) is an RNA chaperone that catalyzes folding of group I and II introns *in vivo* (13, 14). It is a member of the DEAD-box family of SF2 motors, whose main activity is to unwind short (<15 bases) duplex RNAs without translocating (reviewed in (15)). ATP and double-stranded RNA each bind to a single RecA lobe, which stabilizes an activated closed conformation of the enzyme (16). The structure of this closed conformation indicates that the ATPase physically bends the RNA duplex, destabilizing the interactions between the two single strands of RNA (17). After one strand of RNA is released, ATP is hydrolyzed, allowing the second strand of RNA to be released along with ADP and free phosphate. Thus, physical unwinding of the substrate is coupled to ATP binding, while ATP hydrolysis promotes substrate release and recycling of the enzyme (18).

RecBCD**:** The RecBCD system exemplifies how regulation of a motor can be achieved in a multi-subunit complex. RecBCD (Fig. 1D) is a 3-subunit machine from bacteria that facilitates double-strand break repair by catalyzing the first steps in homologous recombination (19). The RecBCD complex binds to the end of a double-strand break, then unwinds and degrades the DNA as it translocates away from the break. Once the complex reaches a sequence specific crossover hotspot instigator ("Chi") site, the complex alters its activity to form a 3' overhang. This overhang serves as a platform for subsequent loading of RecA and the start of homologous recombination. RecB and RecD are both SF1-type helicases but have opposite polarities (3'-5' and 5'-3' respectively), allowing them to translocate in the same direction along double stranded DNA. Two types of subunit coordination are responsible for the biological activity of the RecBCD machine. The first is an enhancement of processivity of RecB by RecC (20). RecB itself is a weak helicase but the crystal structure of RecBCD (Fig. 1D) reveals that RecC contributes a critical pin motif, which prevents reannealing of the duplex as RecB separates the DNA strands (21). The second type of coordination relates to the regulation of the nuclease activity of RecB by the Chi recognition motif of RecC. The nuclease is connected to the rest of the motor via a long linker, allowing it the flexibility to cut either the 3' or the 5' end of the DNA. However, once the complex reaches a Chi site, RecC binds tightly to this site on the 3' end of the DNA, protecting it from nuclease and allowing RecA to load onto the exposed single-stranded DNA.

Mot1**:** Mot1 exemplifies how a motor achieves specificity and drives directionality of the reaction output. Mot1 (Modifier of transcription 1; Fig. 1E) is an SF2 motor that facilitates eukaryotic transcriptional regulation by removing a transcriptional activator, TATA-binding protein (TBP), from DNA upstream of promoters, enabling its redistribution across the genome. TBP binds DNA in a high-affinity, long-lived complex, hence the necessity of a specialized ATPase to remove it (22-24).

Mot1 consists of an ATPase domain, containing the two conserved RecA-like lobes, and a large N-terminal domain that binds TBP through a series of HEAT repeats (25). Unlike the other nucleic acid motors discussed above, Mot1 has negligible DNA binding affinity and its ATPase activity is not stimulated by DNA, but instead by TBP (25). Interestingly, an N-terminal deletion that isolates the ATPase domain *can* weakly bind DNA, suggestive of an autoinhibitory role for the N-terminal region (24). This autoinhibitory feature may contribute to Mot1's specificity by, for example, sequestering the ATPase domain until a TBP-DNA complex has been bound by the N-terminus. Although details of the mechanism by which Mot1 couples ATP hydrolysis to TBP removal remain unclear, the Mot1:TBP crystal structure sheds light on how Mot1 achieves directionality, namely through an alpha helix connected to the N-terminal TBP binding domain by a flexible linker called the "latch". This latch binds the DNA-binding surface of TBP exposed by the action of Mot1, preventing TBP-DNA re-binding, much as the "pin" in RecC prevents the re-association of the nucleic acid strands separated by RecB (25).

A universal theme that emerges from these case studies is the ability of nucleic acid motors to act directionally. This directionality can be defined in terms of the direction of motor movement, such as the 3'-5' versus 5'-3' directionality of RecB versus RecD translocation on single stranded DNA. Directionality of *output,* however, refers to the ability of nucleic acid motors to promote the forward reaction while inhibiting the reverse reaction. This type of directionality requires both specificity for the substrate over the product of the reaction, as well as the coupling of this specificity to the irreversibility of ATP hydrolysis. In the next section, we will explore this theme both mechanistically and structurally in the context of chromatin remodeling motors. Unlike the motors just described, these chromatin remodeling motors must act on a vastly more complicated substrate, the nucleosome (Figure 2A).

Activities of remodeling complexes

Chromatin remodeling motors are classified into four families based on homology within the ATPase subunits: the ISWI family, CHD1 family, INO80 family, and SWI/SNF family (3). These complexes have overlapping roles in key nuclear processes such as transcriptional regulation and DNA repair, yet catalyze distinct transformations of nucleosomes, underscoring the complexity of regulating these processes in the context of chromatin. The activities catalyzed by these complexes include sliding an intact histone octamer, disassembling the histone octamer, exchanging histone variants and altering nucleosome conformation (Figure 2). Chromatin remodelers and their diverse activities have been the topic of a number of excellent reviews (3, 4, 26). Here we build on these reviews to address how remodeling motors elaborate on the basic RecA architecture introduced in the previous section to achieve a diversity of outputs. We begin by introducing the biophysical properties of a nucleosome, focusing on uncatalyzed transitions between the canonical structure and other states that are accessible at physiological temperatures (Figure 2B). We then discuss how remodeling motors might alter the relative populations of accessible chromatin states or induce new chromatin states to generate a precise product.

Intrinsic biophysical properties of a nucleosome

Under physiological conditions, nucleosome structure is stabilized by electrostatic interactions between the phosphate backbone of DNA and positively charged residues on histones (27). The histones make few specific interactions with the DNA bases. However, wrapping of the DNA around the octamer requires bending of the DNA and results in a different type of sequence specificity. For example, GC base pairs favor major groove compression while AT base pairs favor minor groove compression, and so DNA sequences that contain GC and AT base pairs

spaced 5 bp apart show increased affinity for the histone octamer compared to sequences without such periodicity(28, 29).

This structure allows for a diverse set of alternative states that are lowly populated under physiological conditions. For example, the DNA wrapped around the histone octamer can transiently unwrap on the millisecond time-scale, although the equilibrium lies in favor of the wrapped state (30-32). DNA sites closer to the dyad unwrap with a million-fold lower probability than sites closer to the entry/exit site of a nucleosome. In contrast to DNA unwrapping, sliding of the octamer with respect to the DNA is very slow under physiological conditions but can be detected by raising temperature and salt. Finally, it has been shown that increasing salt can promote sequential loss of the H2A/H2B dimer followed by the H3/H4 tetramer, or promote a state in which the H2A/H2B dimer remains attached to the DNA but is detached from the H3/H4 tetramer (33). These intrinsic properties can be further regulated by post-translational modifications of the canonical histone proteins and by replacement with histone variants (1, 34, 35). Such elaborations of the basic nucleosome structure play a major role in conferring biological specificity on the action of chromatin remodeling motors through recruitment and allosteric effects.

These uncatalyzed nucleosomal transformations are qualitatively similar to those achieved by chromatin remodeling motors and provide a conceptual framework to ask if and how remodeling motors exploit these lowly populated states to drive specific outcomes. This concept is best illustrated by a free energy diagram (Figure 2D), where a remodeling motor may either stabilize physiologically accessible transition states or create new states by lowering the barriers for alternate paths. Below, we focus on the INO80 family of remodeling motors.

INO80 family

SWR and INO80 are large, multi-subunit complexes belonging to a family of remodeling enzymes called INO80 (3). *In vivo*, SWR and INO80 act at similar regions in the genome (double-strand breaks, replication forks and transcription start sites, but their contributions at these regions can be distinct (36-38). INO80 and SWR also differ significantly in their activities *in vitro*. While SWR's only known activity is to exchange H2A/H2B dimers for H2A.Z/H2B, INO80 centers mononucleosomes and can evenly space a nucleosomal array (39, 40). Interestingly, in some cases, INO80 can perform the opposite reaction of SWR, taking off H2A.Z/H2B dimers and replacing them with H2A/H2B (41). In this section, we will focus on the mechanism of SWR, as more is known about how SWR performs histone exchange. We will then compare and contrast SWR and INO80 in terms of structure and mechanism.

Demands on the motor: In addition to the four canonical histones that comprise the nucleosome, the incorporation of histone variants that can replace their canonical counterparts creates an additional layer of complexity in the regulation of chromatin structure. H2A.Z, a common variant of H2A, plays broad and essential roles in regulating transcription and maintaining genomic stability across all eukaryotes. H2A.Z is enriched at promoters of genes, where it is thought to interact with transcriptional machinery to regulate gene expression (42) (also reviewed in (43, 44)). While H2A.Z and H2A share only 60% identity in amino acid sequence, crystal structures of nucleosomes containing either H2A.Z or H2A reveal surprisingly similar architectures (45). SWR's ability to perform directional histone variant exchange requires that the complex distinguish between substrate (H2A-containing nucleosomes) and product (H2A.Z-containing nucleosomes), which is a major challenge considering these two structures are highly similar. Specificity for H2A.Z- versus H2A-containing nucleosomes can in principle be achieved by binding modules that recognize small differences between the two histones. However, to

achieve directional histone exchange, binding specificity needs to be coupled to the irreversibility of ATP hydrolysis. Below we explore the known modules of the SWR complex that contribute to its specificity for H2A.Z- versus H2A-containing nucleosomes and speculate as to how these modules coordinate with the ATPase subunit Swr1 to generate directionality.

Specificity for H2A.Z is encoded in separate SWR modules: Careful work from the last decade has revealed that the specificity of SWR's histone exchange activity relies on two separate modules of the complex: (*i*) Swc2, a subunit of the SWR complex; and (*ii*) Swr1-Z, a small, 30-amino acid region immediately N-terminal of the RecA-like ATPase domain of the Swr1 ATPase. Both Swc2 and Swr1-Z specifically interact with H2A.Z/H2B but not H2A/H2B (46, 47). Deletion of either module impairs the ability of SWR to perform directional histone exchange *in vitro* (47, 48). In fact, deletion of the Swc2 subunit changes the directionality of exchange such that it can replace H2A.Z-containing nucleosomes with canonical H2A/H2B dimers (48). The Swr1-Z region acts as a *bona fide* chaperone for H2A.Z/H2B dimers, as it can incorporate H2A.Z/H2B dimers onto pre-formed DNA-$(H3/H4)_2$ tetrasomes in the absence of the rest of the Swr1 ATPase domain.

Together, these data have led to a chaperone-like model of histone exchange by SWR (Fig. 3A). In this model, Swc2 and Swr1-Z bind to H2A.Z/H2B, delivering the H2A.Z/H2B dimer close to the canonical nucleosome. The exchange reaction must require a hand-off mechanism in which the interactions between SWR and H2A.Z/H2B are replaced by those between H2A.Z/H2B and H3/H4, but it is not clear if eviction of H2A/H2B and deposition of H2A.Z/H2B are coupled processes. Eviction of H2A/H2B may also require a separate module of SWR that has specificity for H2A/H2B over H2A.Z/H2B. This module could act analogously to the "latch" domain of Mot1 to help disrupt interactions between the H3/H4 and the H2A/H2B dimer, preventing re-association of H2A/H2B (47). While this module has not yet been identified, a

subunit of SWR called Yaf9 has been shown to bind specifically to H3 and H4, and not to H2A or H2B (49). In fact, Yaf9 is required for histone exchange, raising the intriguing possibility that Yaf9 prevents re-association of an evicted H2A-H2B dimer by blocking interactions with the tetramer (46).

Kinetic measurements of the exchange reaction show that histone exchange occurs in a stepwise manner, with a heterotypic nucleosome (containing one H2A/H2B dimer and one H2A.Z/H2B dimer) as an intermediate of the full exchange reaction (50). Consistent with stepwise exchange, *in vivo*, H2A.Z-containing nucleosomes either have a single copy of H2A.Z/H2B (heterotypic) or two copies of H2A.Z/H2B (homotypic), raising the possibility that the number of copies of H2A.Z provides a source of biological regulation. Further, since SWR has two separate modules that specifically recognize H2A.Z/H2B (Swc2 and Swr1-Z), it is also possible that a single SWR can processively catalyze two exchange reactions. However, the processivity would need to be regulated to generate the heterotypic nucleosomes seen *in vivo*.

Coupling of substrate specificity to ATP hydrolysis: It is important to note that when both H2A.Z/H2B dimers have been incorporated in the nucleosome, SWR cannot catalyze the reverse reaction, even when provided a large excess of H2A/H2B dimers (50). Similar to the motors discussed in Section 2, this enforcement of directionality during histone exchange appear to be tightly coupled to its ATPase activity. Unlike remodelers that slide nucleosomes, SWR is not stimulated at all by naked DNA. In fact, SWR ATPase activity is only stimulated by canonical nucleosomes, and not by H2A.Z-containing nucleosomes. There is a further 2.5-fold increase in ATPase rate when excess H2A.Z/H2B dimers, but not canonical H2A/H2B, are provided in addition to canonical nucleosomes. Moreover, both types of dimers, in the absence of a nucleosome, fail to stimulate SWR's ATPase activity, suggesting that a unique intermediate in which SWR binds both the nucleosome and a free H2A.Z/H2B dimer represents an activated

state of the enzyme. Interestingly, in comparison with the structure of H2A.Z in a nucleosome, H2A.Z bound to Swr1-Z has an extended alpha helical structure, which could reflect the structural changes induced by the activated SWR-nucleosome-H2A.Z/H2B complex (47). This altered structure of H2A.Z/H2B could reflect a high-energy intermediate on the histone exchange pathway that is stabilized by SWR (Fig. 2D).

Despite the tremendous progress in understanding how SWR is able to achieve specificity during histone exchange, we still do not know how these H2A.Z-specific modules regulate the ATPase at different steps of the reaction. One hypothesis is that the Swc2 subunit and the Swr1-Z motif regulate the ATPase domain directly and potentially in concert (46, 51). Swr1-Z is located directly adjacent to the ATPase domain, consistent with having a direct effect on ATPase activity. Deletion of the Swc2 subunit reduces all histone-specific ATPase stimulation to basal levels, suggestive of a direct regulatory function (48). However, whether or not Swc2 makes direct contacts with the ATPase domain is not clear (46). It would be informative to investigate how the Swc2 and Swr1-Z regulatory modules work together in regulating the activity of the complex, and whether the two modules act in the same step or different steps along the histone exchange pathway.

Comparison with INO80: similarities and differences in structure and function: The INO80 family of remodeling enzymes are defined by a unique 300 amino acid insertion between the two RecA-like lobes of the ATPase subunit (Figure 3B) (52). In human cells, replacing the sequence of the insertion in the Ino80p ATPase with that of SRCAP, the human homolog of SWR, is sufficient for the replacement of several INO80-specific subunits with SRCAP-specific subunits (53). This result suggests that subunit composition is regulated by the identity of the insert and leads to the attractive hypothesis that the subunits recruited by the insert dictate the respective activities of the SWR and INO80 complexes. Indeed, unlike ISWI enzymes, Ino80

and Swr1 ATPases alone are not sufficient for remodeling activity (46, 53). This observation raises the question of how other subunits regulate the activity of the complex and how this regulation contributes to substrate and product specificity.

Biochemical and structural approaches have shown that both SWR and INO80 exhibit a modular architecture of subunit organization in which different regions of Swr1 and Ino80 ATPases act as a scaffold for association of different sub-complexes, or "modules" (51, 54, 55). The term "modules" in this context refers to clusters of 3-6 subunits that form a functional unit of the larger complex. For example, the SWR complex is organized into two major modules, one module that associates with the N terminus of the Swr1 ATPase and another module, which contains Swc2, that associates with the C terminus of the ATPase (54). For INO80, further complementary functional studies have isolated specific roles for each module. For example, deleting the Arp8 module (subunits lost include: Arp8, Arp4, Act1, Taf14, and Ies4) does not substantially affect nucleosome binding, but instead greatly diminishes remodeling and ATPase activity (56). The differences in functional modules in SWR and INO80 may contribute to how these complexes use substrate cues such as flanking DNA. SWR preferentially binds to nucleosomes with longer flanking DNA (57). This property is thought to recruit the enzyme to the promoters of genes, which are immediately downstream of a nucleosome free region (NFR). On the other hand, INO80 preferentially remodels nucleosomes with longer flanking DNA length while binding is unaffected (40). These and other structure/function studies are just beginning to provide an image of the moving parts involved in nucleosome remodeling by these large machines.

Open questions: How do two highly related enzymes in the same family have such different activities *in vitro*? Given what we know about the dynamics of nucleosome structure, it is tempting to speculate that both motors take advantage of one type of uncatalyzed

conformational change in the nucleosome to enable different outputs (Figure 4C): DNA unwrapping from one end of the nucleosome, which breaks DNA contacts with H2A/H2B. In the case of INO80, this unwrapping may provide a precursor for propagating the unwrapped DNA across the histone octamer to enable sliding. In the case of SWR, unwrapping poises the H2A/H2B dimer for exchange while the DNA propagation pathway may be blocked by an accessory subunit or unidentified domain of the Swr1 ATPase. Thus, both sliding by INO80 and histone exchange by SWR may share a common, physiologically accessible intermediate. Another major open question for the INO80 family is how movement of the RecA lobes of the ATPase subunit enables remodeling. Unlike ISWI, the INO80 family has not been shown to translocate along naked DNA, suggesting that the RecA-like lobes may have a completely different function, such as directly disrupting histone-histone contacts or regulating large-scale conformational changes within the complex during remodeling.

Additional complexities imposed on remodeler action in vivo

Our current mechanistic understanding of chromatin remodelers provides a starting point for asking how these motors may function in the complex environment of the nucleus. Below, we briefly discuss how additional features of the chromatin template as well other chromatin regulators may influence the activities of remodeling complexes.

DNA sequence: DNA sequences that are more bendable in a manner favored for nucleosome formation, termed positioning sequences, have been shown to form more stable nucleosomes (29, 58). *In vivo*, such DNA sequences are correlated with well-positioned nucleosomes, raising the question of whether DNA sequence can influence the activity of remodelers by regulating the stability of the nucleosome (59-61). In this context it is instructive to assess the dynamic range of stability that is achievable by varying DNA sequence. One of the highest affinity

sequences is a synthetic DNA sequence obtained by *in vitro* selection assays, called 601. At 4°C, 601 has a ΔΔG of formation of ~3 kcal/mol compared to a natural nucleosome positioning sequence termed 5S (62). This corresponds to a roughly 100-fold greater overall nucleosomal stability, suggesting in principle that DNA sequence could regulate remodeling rates over a 100-fold dynamic range. However, when measured at physiological temperatures, the 100-fold difference in stability diminishes to a 4-fold range (ΔΔG ~0.8 kcal/mol) (63). Consistent with this smaller range, nucleosomes assembled on either the 601 sequence or 5S are repositioned by ISWI and SWI/SNF family members at the same rate (62). Does this mean that DNA sequence cannot regulate remodeling activities? Not necessarily. Unlike overall nucleosome stability, DNA unwrapping is dramatically affected by DNA sequence at physiological temperatures (up to 100-fold greater for 5S compared to 601) (31). Thus in principle, remodeling motors that use DNA unwrapping as part of their mechanism could be greatly influenced by DNA sequence. Sequence could also affect other aspects of nucleosome dynamics that remodelers could exploit. For example, it has been shown that certain DNA sequences promote octamer transfer by SWI/SNF, while others, specifically AT-rich sequences, promote octamer eviction by RSC in the presence of histone chaperones (64).

Higher-order chromatin structures and histone modifications: Most of the studies we have discussed so far were carried out with mononucleosomes or defined arrays of extended nucleosomes, and with recombinant histones that do not contain post-translational modifications (PTMs) (65). While these substrates are powerful tools for capturing core properties of remodeling motors, there remains a large gap in the types of chromatin structures that we can study *in vitro* and the types of structures that exist *in vivo*. *In vivo*, inter-nucleosomal interactions drive chromatin into higher-order states of folding (1). These states are thought to be driven by electrostatic interactions between the histone tails and DNA, and between the H4 tail and a

negatively charged patch between H2A and H2B called the acidic patch (27). Specific proteins also participate in stabilizing certain types of condensed chromatin, including the linker histone H1 as well as heterochromatin associated proteins such as HP1 and Polycomb (66-69). Furthermore, *in vivo,* histones are decorated with many different types of PTMs, which regulate remodelers either directly or through effects on chromatin structure (70). How different remodelers act on physiological nucleosomal substrates is unclear, though recent work has begun to shed light on this topic. For example, *Drosophila* ISWI remodeling activity is only minimally inhibited by nucleosome arrays containing H1, while SWI/SNF remodeling is more significantly inhibited on this more *in vivo*-like substrate. Similarly, underscoring the importance of histone modifications, the CHD and SWI/SNF families each contain domains that recognize specific PTMs, which affect their localization *in vivo* and their activity *in vitro* (70). Further progress towards addressing these questions, however, has been limited both by the difficulty of reconstituting these substrates homogeneously and by the limited range of assays which can directly report on remodeling of these complex substrates.

Mechanisms of collaboration with other nuclear factors: *In vivo*, a remodeler acts in the context of many other factors that are competing for the opportunity to interact with the chromatin template. How do remodelers coordinate their actions with these other factors to achieve a defined outcome? While much remains to be learned, some biochemical studies are beginning to shed light on how such coordination may be achieved.

(i) Sequence-specific DNA binding factors: In addition to passively taking advantage of DNA binding sites made accessible by chromatin remodeling motors, specific DNA binding factors can directly recruit chromatin remodeling complexes. For example, in *Drosophila*, the ISWI-family member NURF is recruited by the GAGA factor to the Hsp70 promoter to promote

transcriptional activation. This role is particularly important as to date no remodeling complex has been shown to possess high levels of DNA sequence specificity, yet many are known to be specifically recruited to particular DNA loci *in vivo* (71). DNA binding factors can also act as barriers to the movement of nucleosomes. In the case of ISWI, *in vitro* single-molecule experiments have shown that a bound transcription factor can act as a barrier, such that the ISWI motors move nucleosomes away from the bound factor, in contrast to SWI/SNF, which uses the nucleosome to displace the bound factor (72).

(ii) Other chromatin remodeling motors: Many sites of chromatin regulation involve action by multiple families of remodelers. For example, in yeast, promoters are defined by a nucleosome-free region at the transcriptional start site followed by a series of three to four well-positioned nucleosomes (73). Several remodelers have been shown to be necessary for generating this architecture *in vivo* (74). How and why are these combinations of remodelers required? One possibility is their action is temporally controlled, so that only one remodeler is recruited and active at a time, paving the way for the next remodeler. Alternatively, all the remodelers act at the same time such that the steady state reflects the outcome seen *in vivo*.

(iii) Other nuclear machinery: Chromatin remodeling is required for facilitating essentially all other nuclear processes. For example, the yeast SWI/SNF family member RSC is required for RNA polymerase elongation at coding regions of actively transcribed genes, possibly acting alongside histone chaperones to evict and re-assemble histones (75). The INO80 and SWR complexes play essential roles in transcription as well as DNA repair, where the action of these complexes are required for recruitment of repair factors to the site of DNA damage (76, 77). An exciting area of current research aims to understand the interplay between chromatin remodeling enzymes and the other nuclear machinery, including (*i*) which remodelers are

directly recruited by other nuclear machines, versus those that act as the primary recruiting factors (*e.g.* at sites of DNA damage), (*ii*) which types of nucleosome remodeling events are required for other processes, and *(iii)* whether remodelers directly regulate the activity of other nuclear machines.

Thesis outline

Chromatin remodeling motors use a diverse set of strategies to regulate gene expression, but the mechanisms of how these motors harness the energy of ATP to create structural changes in chromatin are not well understood. In this thesis, we focus on the INO80 complex from *S. cerevisiae*, and address several types of questions related to both the assembly and nucleosome remodeling mechanisms of this tightly regulated complex.

The first half of the thesis will focus on the role of Rvb1 and Rvb2 (Rvbs, Rvb1/Rvb2) in nucleosome remodeling by INO80. The Rvbs are members of the AAA+ ATPase family of molecular motors, and are integral components of the INO80 family of chromatin remodeling enzymes, as well as other multi-subunit complexes in the cell (78). The Rvbs were initially proposed to act as DNA helicases, though it was recently proposed that they may act as chaperones to help assemble active forms of these complexes. In chapter 2, we summarize our work directed at distinguishing between these two models. Briefly, we discovered the first potential client of the Rvbs, a small insertion within the Ino80 ATPase termed Ino80INS. We found that Ino80INS directly interacts with the Rvbs and robustly stimulates its ATPase activity, likely by promoting dodecamerization of the normally hexameric Rvb1/Rvb2. Based on these and other lines of evidence presented in chapter 2, we propose that the Rvbs' chaperone-like activity is related to its ability to couple changes in oligomeric state (hexamer and dodecamer) to its ATPase cycle. The more active, dodecameric state of the Rvbs is also energetically more favorable for binding to clients, facilitating the association of multiple client proteins along the

dodecameric interface. Once assembly has occurred, the Rvb dodecamer falls apart, resulting in the stable association of a Rvb1/Rvb2 hexamer with the rest of the complex. The collapse and reformation of the dodecamer is coupled to stages of the ATPase cycle, allowing the Rvbs to catalyze many rounds of assembly if necessary. An important feature of our model is that it addresses the controversy over the oligomeric state of Rvb1/Rvb2 by assigning different roles to each state. We believe our findings are only scratching the surface of how the Rvbs function *in vivo*, and the appendix of chapter 2 summarizes several thought-provoking unpublished results that may be the beginnings for future investigations.

The second half of the thesis focuses on a more traditional aspect of remodeling motor mechanism: how does INO80 integrate substrate cues on a nucleosome to create a specific and predictable result? It was shown previously that INO80 could slide nucleosomes toward the center of a piece of short DNA, and center a trinucleosomal array (40), but required at least 30 base pairs (bp) of flanking DNA to accomplish these tasks. Using a combination of bulk and single molecule approaches, we found that unlike other DNA length sensing chromatin remodelers, INO80 behaves more switch-like. When there is ≤ 40 bp of flanking DNA on one side of the nucleosome, INO80 will only move the nucleosome extremely slowly. However, when the flanking DNA is > 40 bp, the rate of nucleosome sliding is greatly enhanced. To achieve this switch-like behavior, we found that a set of subunits containing Nhp10 is likely responsible for autoinhibiting sliding on nucleosomal constructs with short (≤ 40 bp) flanking DNA. Once this autoinhibition is relieved by longer flanking DNA, INO80 can travel long distances (>25 bp) without re-assessing length, a significant departure from what is known for ISWI-like motors which also center and evenly sapce nucleosomes. These results, which are summarized in chapter 3, offer two major advances for the chromatin remodeling field. First, they challenge the long-standing “stop-and-reassess” model of ISWI motors, suggesting that INO80 may play a different type of role *in vivo* that does not require precise spacing of

nucleosomes. Secondly, our nhp10 data suggest that the association or dissociation of this module may be an additional way of regulating nucleosome spacing in vivo.

Taken together, the findings presented in this thesis paint a complex picture of how INO80 lives and breathes during its lifetime in the nucleus. First, INO80 must be assembled from its many components, with the help of the molecular chaperone Rvb1/Rvb2. Our results suggest that this assembly process may be iterative, in tune with the ebb and flow of the Rvbs' ATPase cycle as well as the dynamics of client-Rvb interactions. Once assembled into an active form, INO80 can assess a nucleosomal substrate for its flanking DNA length, with the help of its nhp10 module. If it determines the DNA is long enough, INO80 will quickly slide the nucleosome along the DNA. Depending on the biological context, this could mean away from a DNA end, such as during double strand break repair. Or, INO80 may enforce the creation of the nucleosome free region (NFR) at transcription start sites by quickly moving nucleosomes away from the NFR. At any point during its lifetime, INO80 may be disassembled and re-assembled by the Rvbs, depending on the needs of the cell. We hope this thesis helps foster many more generations of research toward better understanding the fascinating world of INO80.

References

1. Luger, K., Dechassa, M. L., and Tremethick, D. J. (2012) New insights into nucleosome and chromatin structure: an ordered state or a disordered affair? *Nat. Rev. Mol. Cell Biol.* 13, 436–447

2. Kornberg, R. D., and Lorch, Y. (1992) Chromatin structure and transcription. *Annu. Rev. Cell Biol.* 8, 563–587

3. Clapier, C. R., and Cairns, B. R. (2009) The biology of chromatin remodeling complexes. *Annu. Rev. Biochem.* 78, 273–304

4. Narlikar, G. J., Sundaramoorthy, R., and Owen-Hughes, T. (2013) Mechanisms and functions of ATP-dependent chromatin-remodeling enzymes. *Cell* 154, 490–503

5. Garraway, L. A., and Lander, E. S. (2013) Lessons from the cancer genome. *Cell* 153, 17–37

6. Kadoch, C., and Crabtree, G. R. (2015) Mammalian SWI/SNF chromatin remodeling complexes and cancer: Mechanistic insights gained from human genomics. *Sci Adv* 1, e1500447–e1500447

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

8. Hopfner, K.-P., Gerhold, C.-B., Lakomek, K., and Wollmann, P. (2012) Swi2/Snf2 remodelers: hybrid views on hybrid molecular machines. *Curr. Opin. Struct. Biol.* 22, 225–233

9. Cox, M. M. (2007) Motoring along with the bacterial RecA protein. *Nat. Rev. Mol. Cell Biol.* 8, 127–138

10. Story, R. M., and Steitz, T. A. (1992) Structure of the recA protein-ADP complex. *Nature* 355, 374–376

11. Chen, Z., Yang, H., and Pavletich, N. P. (2008) Mechanism of homologous recombination from the RecA-ssDNA/dsDNA structures. *Nature* 453, 489–484

12. Singleton, M. R., Dillingham, M. S., and Wigley, D. B. (2007) Structure and mechanism of helicases and nucleic acid translocases. *Annu. Rev. Biochem.* 76, 23–50

13. Séraphin, B., Simon, M., Boulet, A., and Faye, G. (1989) Mitochondrial splicing requires a protein from a novel helicase family. *Nature* 337, 84–87

14. Huang, H.-R., Rowe, C. E., Mohr, S., Jiang, Y., Lambowitz, A. M., and Perlman, P. S. (2005) The splicing of yeast mitochondrial group I and group II introns requires a DEAD-box protein with RNA chaperone function. *Proc. Natl. Acad. Sci. U.S.A.* 102, 163–168

15. Jarmoskaite, I., and Russell, R. (2014) RNA helicase proteins as chaperones and remodelers. *Annu. Rev. Biochem.* 83, 697–725

16. Mallam, A. L., Del Campo, M., Gilman, B., Sidote, D. J., and Lambowitz, A. M. (2012) Structural basis for RNA-duplex recognition and unwinding by the DEAD-box helicase Mss116p. *Nature* 490, 121–125

17. Del Campo, M., and Lambowitz, A. M. (2009) Structure of the Yeast DEAD box protein Mss116p reveals two wedges that crimp RNA. *Mol. Cell* 35, 598–609

18. Mallam, A. L., Sidote, D. J., and Lambowitz, A. M. (2014) Molecular insights into RNA and DNA helicase evolution from the determinants of specificity for a DEAD-box RNA helicase. *Elife* 3, e04630

19. Wigley, D. B. (2013) Bacterial DNA repair: recent insights into the mechanism of RecBCD, AddAB and AdnAB. *Nat. Rev. Microbiol.* 11, 9–13

20. Yu, M., Souaya, J., and Julin, D. A. (1998) The 30-kDa C-terminal domain of the RecB protein is critical for the nuclease activity, but not the helicase activity, of the RecBCD enzyme from Escherichia coli. *Proc. Natl. Acad. Sci. U.S.A.* 95, 981–986

21. Singleton, M. R., Dillingham, M. S., Gaudier, M., Kowalczykowski, S. C., and Wigley, D. B. (2004) Crystal structure of RecBCD enzyme reveals a machine for processing DNA breaks. *Nature* 432, 187–193

22. Parvin, J. D., McCormick, R. J., Sharp, P. A., and Fisher, D. E. (1995) Pre-bending of a promoter sequence enhances affinity for the TATA-binding factor. *Nature* 373, 724–727

23. Perez-Howard, G. M., Weil, P. A., and Beechem, J. M. (1995) Yeast TATA binding protein interaction with DNA: fluorescence determination of oligomeric state, equilibrium binding, on-rate, and dissociation kinetics. *Biochemistry* 34, 8005–8017

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

25. Wollmann, P., Cui, S., Viswanathan, R., Berninghausen, O., Wells, M. N., Moldt, M., Witte, G., Butryn, A., Wendler, P., Beckmann, R., Auble, D. T., and Hopfner, K.-P. (2011) Structure and mechanism of the Swi2/Snf2 remodeller Mot1 in complex with its substrate TBP. *Nature* 475, 403–407

26. Bartholomew, B. (2014) Regulating the chromatin landscape: structural and mechanistic perspectives. *Annu. Rev. Biochem.* 83, 671–696

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

28. Shrader, T. E., and Crothers, D. M. (1990) Effects of DNA sequence and histone-histone interactions on nucleosome placement. *J. Mol. Biol.* 216, 69–84

29. Widom, J. (2001) Role of DNA sequence in nucleosome stability and dynamics. *Q. Rev. Biophys.* 34, 269–324

30. Polach, K. J., and Widom, J. (1995) Mechanism of protein access to specific DNA sequences in chromatin: a dynamic equilibrium model for gene regulation. *J. Mol. Biol.* 254, 130–149

31. Anderson, J. D., and Widom, J. (2000) Sequence and position-dependence of the equilibrium accessibility of nucleosomal DNA target sites. *J. Mol. Biol.* 296, 979–987

32. Li, G., Levitus, M., Bustamante, C., and Widom, J. (2005) Rapid spontaneous accessibility of nucleosomal DNA. *Nat. Struct. Mol. Biol.* 12, 46–53

33. Böhm, V., Hieb, A. R., Andrews, A. J., Gansen, A., Rocker, A., Tóth, K., Luger, K., and Langowski, J. (2011) Nucleosome accessibility governed by the dimer/tetramer interface. *Nucleic Acids Res.* 39, 3093–3102

34. Andrews, A. J., and Luger, K. (2011) Nucleosome structure(s) and stability: variations on a theme. *Annu Rev Biophys* 40, 99–117

35. Killian, J. L., Li, M., Sheinin, M. Y., and Wang, M. D. (2012) Recent advances in single molecule studies of nucleosomes. *Curr. Opin. Struct. Biol.* 22, 80–87

36. Morrison, A. J., and Shen, X. (2009) Chromatin remodelling beyond transcription: the INO80 and SWR1 complexes. *Nat. Rev. Mol. Cell Biol.* 10, 373–384

37. Horigome, C., Oma, Y., Konishi, T., Schmid, R., Marcomini, I., Hauer, M. H., Dion, V., Harata, M., and Gasser, S. M. (2014) SWR1 and INO80 Chromatin Remodelers Contribute to DNA Double-Strand Break Perinuclear Anchorage Site Choice. *Mol. Cell* 55, 626–639

38. Gerhold, C.-B., Hauer, M. H., and Gasser, S. M. (2015) INO80-C and SWR-C: guardians of the genome. *J. Mol. Biol.* 427, 637–651

39. Shen, X., Ranallo, R., Choi, E., and Wu, C. (2003) Involvement of actin-related proteins in ATP-dependent chromatin remodeling. *Mol. Cell* 12, 147–155

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

41. Papamichos-Chronakis, M., Watanabe, S., Rando, O. J., and Peterson, C. L. (2011) Global regulation of H2A.Z localization by the INO80 chromatin-remodeling enzyme is essential for genome integrity. *Cell* 144, 200–213

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

43. Zlatanova, J., and Thakar, A. (2008) H2A.Z: view from the top. *Structure* 16, 166–179

44. Subramanian, V., Fields, P. A., and Boyer, L. A. (2015) H2A.Z: a molecular rheostat for transcriptional control. *F1000Prime Rep* 7, 01

45. Suto, R. K., Clarkson, M. J., Tremethick, D. J., and Luger, K. (2000) Crystal structure of a nucleosome core particle containing the variant histone H2A.Z. *Nat. Struct. Biol.* 7, 1121–1124

46. Wu, W.-H., Alami, S., Luk, E., Wu, C.-H., Sen, S., Mizuguchi, G., Wei, D., and Wu, C. (2005) Swc2 is a widely conserved H2AZ-binding module essential for ATP-dependent histone exchange. *Nat. Struct. Mol. Biol.* 12, 1064–1071

47. Hong, J., Feng, H., Wang, F., Ranjan, A., Chen, J., Jiang, J., Ghirlando, R., Xiao, T. S., Wu, C., and Bai, Y. (2014) The catalytic subunit of the SWR1 remodeler is a histone chaperone for the H2A.Z-H2B dimer. *Mol. Cell* 53, 498–505

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

49. Wang, A. Y., Schulze, J. M., Skordalakes, E., Gin, J. W., Berger, J. M., Rine, J., and Kobor, M. S. (2009) Asf1-like structure of the conserved Yaf9 YEATS domain and role in H2A.Z deposition and acetylation. *Proc. Natl. Acad. Sci. U.S.A.* 106, 21573–21578

50. Luk, E., Ranjan, A., FitzGerald, P. C., Mizuguchi, G., Huang, Y., Wei, D., and Wu, C. (2010) Stepwise histone replacement by SWR1 requires dual activation with histone H2A.Z and canonical nucleosome. *Cell* 143, 725–736

51. Wu, W.-H., Wu, C.-H., Ladurner, A., Mizuguchi, G., Wei, D., Xiao, H., Luk, E., Ranjan, A., and Wu, C. (2009) N terminus of Swr1 binds to histone H2AZ and provides a platform for subunit assembly in the chromatin remodeling complex. *J. Biol. Chem.* 284, 6200–6207

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

53. Chen, L., Conaway, R. C., and Conaway, J. W. (2013) 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.* 110, 20497–20502

54. Nguyen, V. Q., Ranjan, A., Stengel, F., Wei, D., Aebersold, R., Wu, C., and Leschziner, A. E. (2013) Molecular architecture of the ATP-dependent chromatin-remodeling complex SWR1. *Cell* 154, 1220–1231

55. Gerhold, C. B., and Gasser, S. M. (2014) INO80 and SWR complexes: relating structure to function in chromatin remodeling. *Trends Cell Biol.* 24, 619–631

56. Tosi, A., Haas, C., Herzog, F., Gilmozzi, A., Berninghausen, O., Ungewickell, C., Gerhold, C. B., Lakomek, K., Aebersold, R., Beckmann, R., and Hopfner, K.-P. (2013) Structure and subunit topology of the INO80 chromatin remodeler and its nucleosome complex. *Cell* 154, 1207–1219

57. Ranjan, A., Mizuguchi, G., FitzGerald, P. C., Wei, D., Wang, F., Huang, Y., Luk, E., Woodcock, C. L., and Wu, C. (2013) Nucleosome-free region dominates histone acetylation in targeting SWR1 to promoters for H2A.Z replacement. *Cell* 154, 1232–1245

58. Cloutier, T. E., and Widom, J. (2004) Spontaneous sharp bending of double-stranded DNA. *Mol. Cell* 14, 355–362

59. Hayes, J. J., Tullius, T. D., and Wolffe, A. P. (1990) The structure of DNA in a nucleosome. *Proc. Natl. Acad. Sci. U.S.A.* 87, 7405–7409

60. Widlund, H. R., Cao, H., Simonsson, S., Magnusson, E., Simonsson, T., Nielsen, P. E., Kahn, J. D., Crothers, D. M., and Kubista, M. (1997) Identification and characterization of genomic nucleosome-positioning sequences. *J. Mol. Biol.* 267, 807–817

61. Segal, E., Fondufe-Mittendorf, Y., Chen, L., Thåström, A., Field, Y., Moore, I. K., Wang, J.-P. Z., and Widom, J. (2006) A genomic code for nucleosome positioning. *Nature* 442, 772–778

62. Partensky, P. D., and Narlikar, G. J. (2009) Chromatin remodelers act globally, sequence positions nucleosomes locally. *J. Mol. Biol.* 391, 12–25

63. Andrews, A. J., Chen, X., Zevin, A., Stargell, L. A., and Luger, K. (2010) The histone chaperone Nap1 promotes nucleosome assembly by eliminating nonnucleosomal histone DNA interactions. *Mol. Cell* 37, 834–842

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

65. Luger, K., Rechsteiner, T. J., Flaus, A. J., Waye, M. M., and Richmond, T. J. (1997) Characterization of nucleosome core particles containing histone proteins made in bacteria. *J. Mol. Biol.* 272, 301–311

66. Canzio, D., Larson, A., and Narlikar, G. J. (2014) Mechanisms of functional promiscuity by HP1 proteins. *Trends Cell Biol.* 24, 377–386

67. Eissenberg, J. C., and Elgin, S. C. R. (2014) HP1a: a structural chromosomal protein regulating transcription. *Trends Genet.* 30, 103–110

68. Simon, J. A., and Kingston, R. E. (2013) Occupying chromatin: Polycomb mechanisms for getting to genomic targets, stopping transcriptional traffic, and staying put. *Mol. Cell* 49, 808–824

69. Harshman, S. W., Young, N. L., Parthun, M. R., and Freitas, M. A. (2013) H1 histones: current perspectives and challenges. *Nucleic Acids Res.* 41, 9593–9609

70. Swygert, S. G., and Peterson, C. L. (2014) Chromatin dynamics: interplay between remodeling enzymes and histone modifications. *Biochim. Biophys. Acta* 1839, 728–736

71. Yen, K., Vinayachandran, V., Batta, K., Koerber, R. T., and Pugh, B. F. (2012) Genome-wide nucleosome specificity and directionality of chromatin remodelers. *Cell* 149, 1461–1473

72. Li, M., Hada, A., Sen, P., Olufemi, L., Hall, M. A., Smith, B. Y., Forth, S., McKnight, J. N., Patel, A., Bowman, G. D., Bartholomew, B., and Wang, M. D. (2015) Dynamic regulation of transcription factors by nucleosome remodeling. *Elife* 4,

73. Zhang, Z., Wippo, C. J., Wal, M., Ward, E., Korber, P., and Pugh, B. F. (2011) A packing mechanism for nucleosome organization reconstituted across a eukaryotic genome. *Science* 332, 977–980

74. Hartley, P. D., and Madhani, H. D. (2009) Mechanisms that specify promoter nucleosome location and identity. *Cell* 137, 445–458

75. Spain, M. M., Ansari, S. A., Pathak, R., Palumbo, M. J., Morse, R. H., and Govind, C. K. (2014) The RSC complex localizes to coding sequences to regulate Pol II and histone occupancy. *Mol. Cell* 56, 653–666

76. Morrison, A. J., Highland, J., Krogan, N. J., Arbel-Eden, A., Greenblatt, J. F., Haber, J. E., and Shen, X. (2004) INO80 and gamma-H2AX interaction links ATP-dependent chromatin remodeling to DNA damage repair. *Cell* 119, 767–775

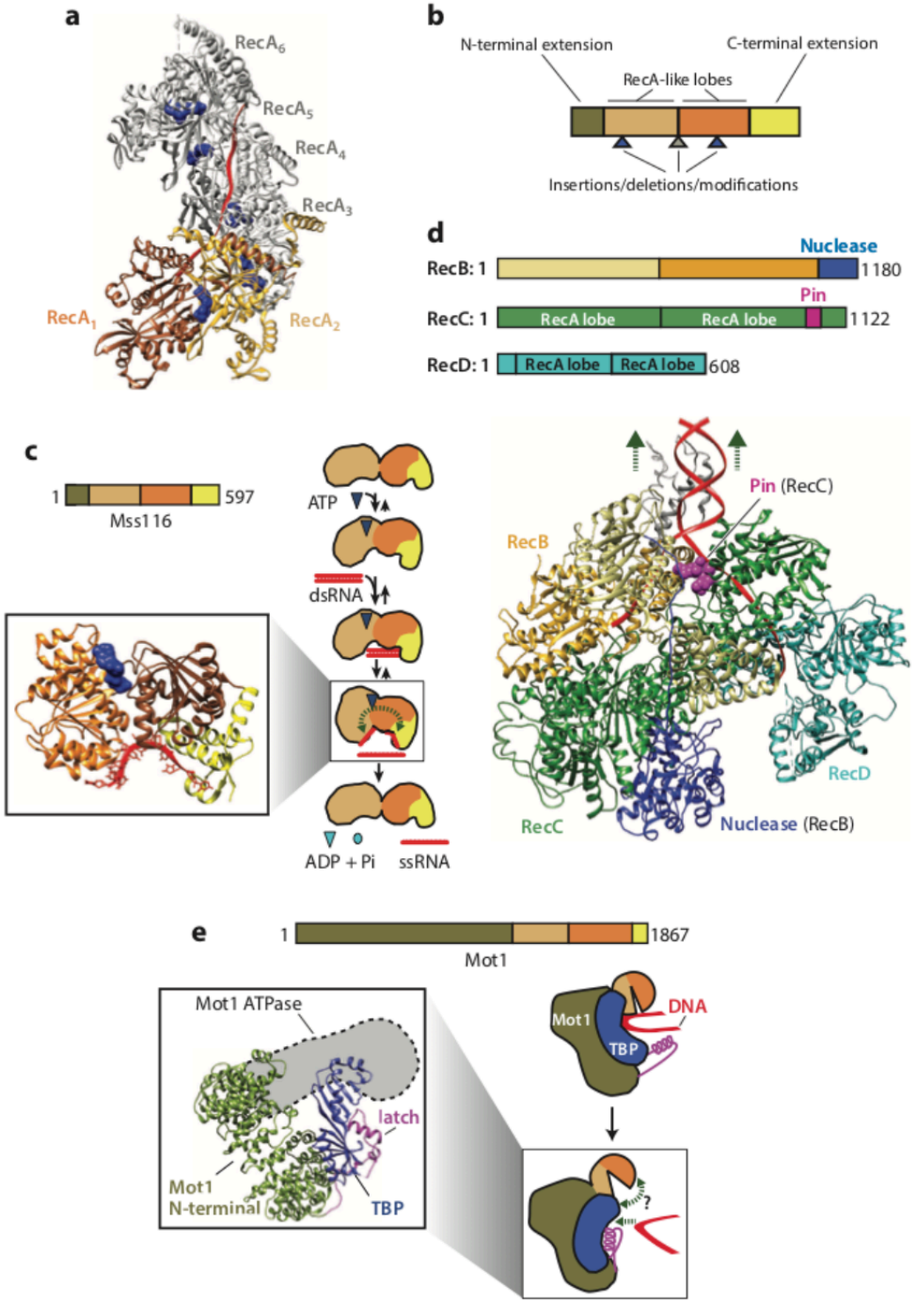
77. van Attikum, H., Fritsch, O., Hohn, B., and Gasser, S. M. (2004) Recruitment of the INO80 complex by H2A phosphorylation links ATP-dependent chromatin remodeling with DNA double-strand break repair. *Cell* 119, 777–788

78. Nano, N., and Houry, W. A. (2013) Chaperone-like activity of the AAA+ proteins Rvb1

and Rvb2 in the assembly of various complexes. *Philos. Trans. R. Soc. Lond., B, Biol. Sci.* 368, 20110399

79. Ranjan, A., Wang, F., Mizuguchi, G., Wei, D., Huang, Y., and Wu, C. (2015) H2A histone-fold and DNA elements in nucleosome activate SWR1-mediated H2A.Z replacement in budding yeast. *Elife* 4,

Figure 1: Case studies of nucleic acid motors



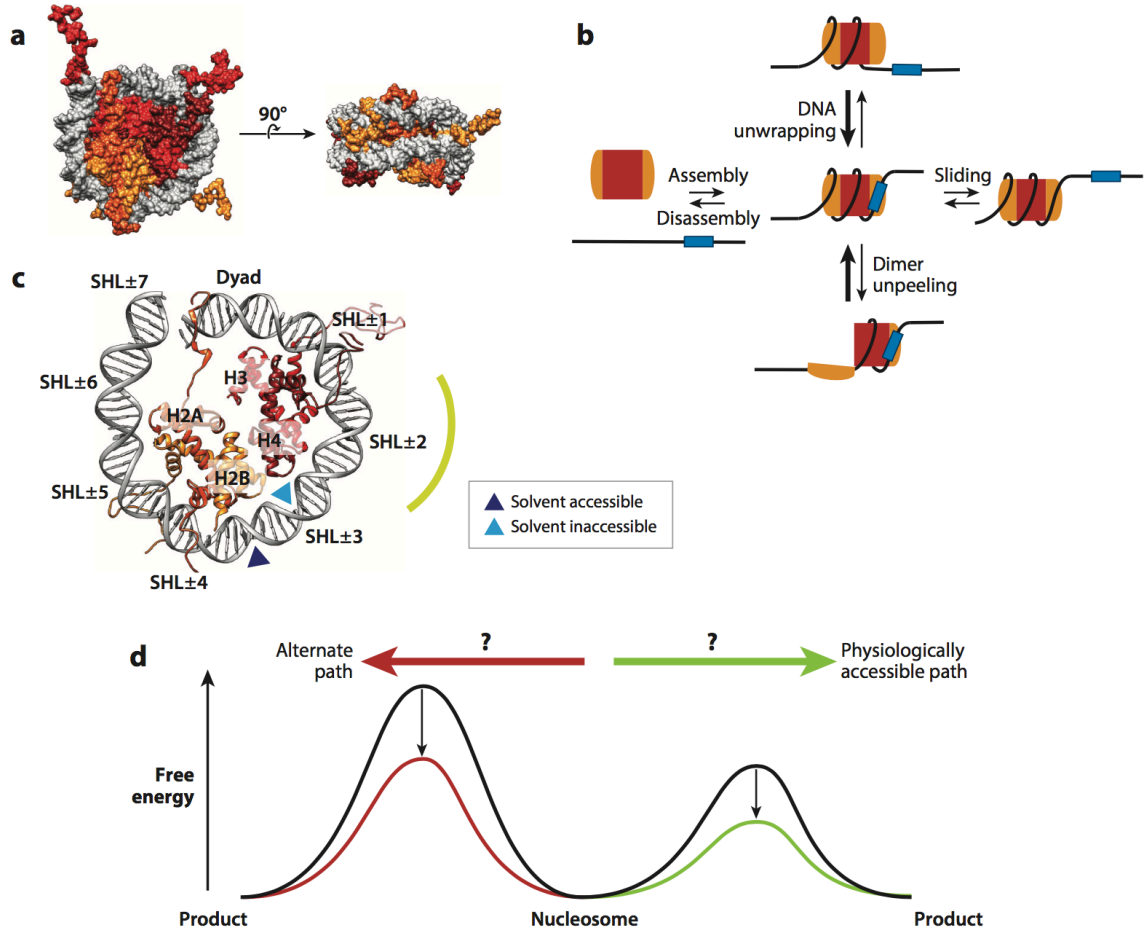
(A) Crystal structure of a RecA filament bound to single-stranded DNA (red ribbon) with ADP-AlF4 (blue space-fill) (PDBID: 3CMU) (11).

(B) Domain architecture of SF1 and SF2 ATPases. The conserved RecA-like folds (orange and tan) can be elaborated by insertions or modifications between the RecA folds (gray triangles) or within either RecA-like fold (blue triangles), as well as by N- and C-terminal extensions (green and yellow, respectively). In (*c*)-(*d*), for simplicity, schematics show only N- and C-terminal extensions without indication of insertions and modifications in the RecA folds.

(C) Mss116. Structure of Mss116 bound to duplex RNA and AMP-PNP (17) (PDBID: 3i5x). The N-terminal extension is disordered in the crystal structure and is not shown. (D) RecBCD. Structure of the RecBCD complex and partially unwound DNA (red ribbons) with the pin motif shown in magenta (PDBID: 3K70).

(E) Mot1. Crystal structure of the N-terminal domain of Mot1 bound to TBP, with the approximate location of the ATPase domain, based on electron microscopy and crosslinking data, shown in gray (PDBID: 3OC3) (25). Although Mot1's mechanism is not fully understood, it is likely that movement of the RecA-fold-containing lobes of the ATPase is coupled to TBP removal. Crystal structure representations in all figures were made with Chimera (http://www.cgl.ucsf.edu/chimera).

Figure 2: General properties of nucleosomes



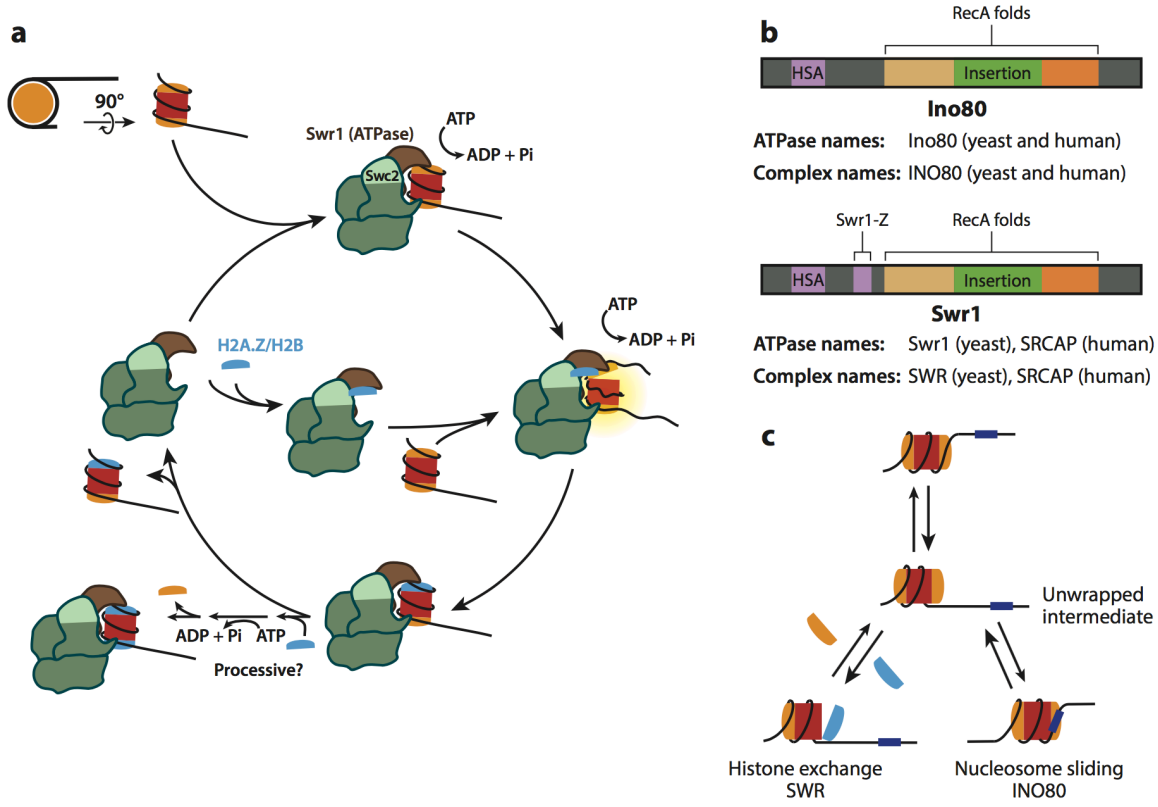
(A) Two space-filling views of a nucleosome core particle (PDB ID: 1KX5), with the DNA in gray and the histone proteins in red/orange {Davey:2002ir}.

(B) Schematic of all physiologically accessible states for a nucleosome observed to date. H2A/H2B dimers are orange and H3/H4 tetramer is red.

(C) Top view of a nucleosome with histones and Super Helical Locations (SHL) on the DNA labeled. SHL±2, the binding location for several remodeler families, is highlighted in yellow.

(D) Hypothetical free energy profiles for a nucleosome remodeling reaction. The same remodeled product can be made through a physiologically accessible (right) or alternate (left) path.

Figure 3: INO80 family of remodeling motors

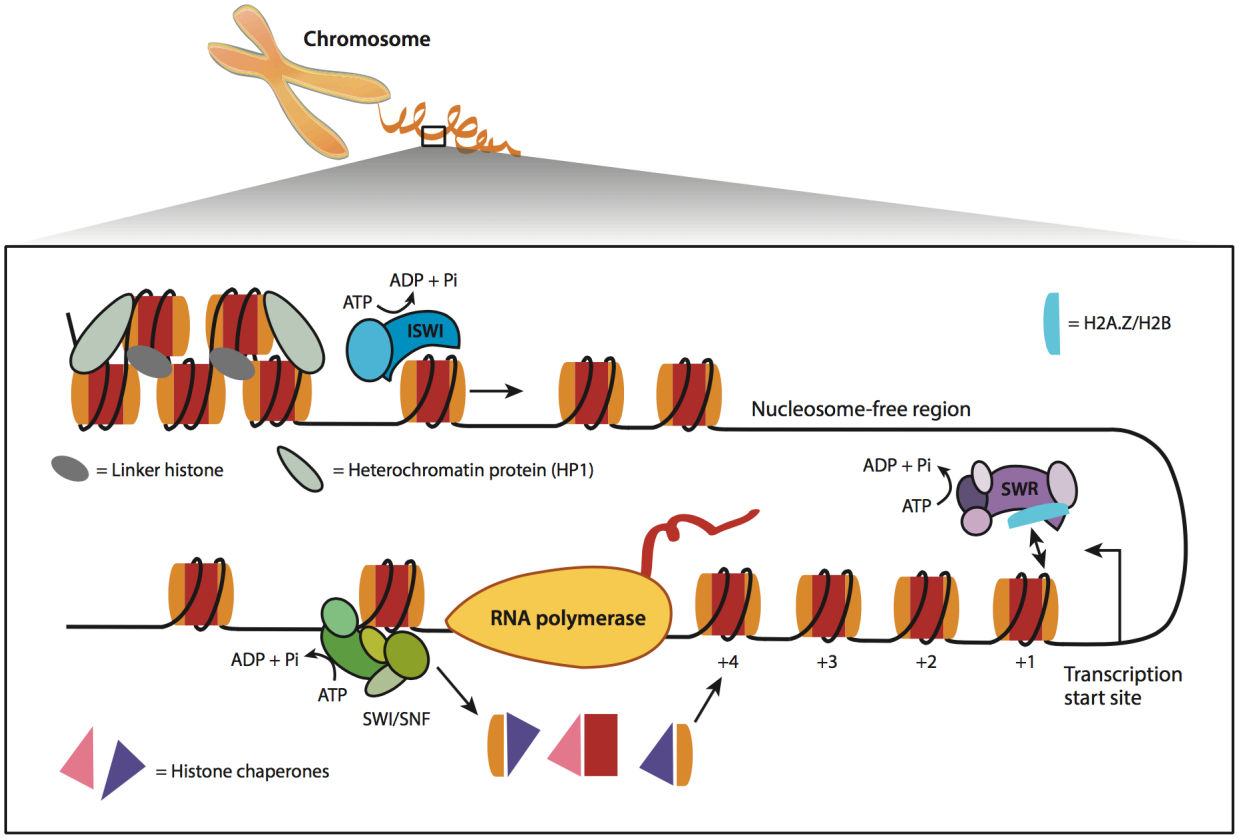


(A) Model for directional histone exchange by SWR (adapted from (50)). The Swr1 ATPase (brown) engages a nucleosome containing a canonical H2A/H2B dimer (blue) at SHL±2 on the distal side relative to flanking DNA (79). The Swc2 subunit and SWR1 both engage an H2A.Z/H2B dimer. When both the nucleosome and H2A.Z are bound, Swr1's ATPase activity is maximally stimulated and a single H2A.Z containing dimer is exchanged. Swr1 may also act processively until a fully H2A.Z nucleosome is formed.

(B) Domain architectures of the Ino80 and Swr1 ATPases.

(C) The INO80 and SWR complexes may share a common intermediate in which DNA is unwrapped from one side of the nucleosome. SWR may use this intermediate to facilitate histone exchange with H2A.Z/H2B while INO80 may use it to slide the histone octamer toward the site of unwrapping.

Figure 4: **Remodeling motors act in conjunction *in vivo***



In vivo, different families of remodeling enzymes collaborate with each other and with other nuclear factors to regulate gene expression. The mechanisms behind this collaboration are not well understood, and are an area of continuing interest in the field.

Chapter 2: Regulation of Rvb1/Rvb2 by a domain within the INO80 chromatin remodeling complex implicates the yeast Rvbs as protein assembly chaperones

Abstract

The hexameric AAA+ ATPases Rvb1 and Rvb2 (Rvbs) are essential for a wide range of processes including metabolic signaling, transcription, telomere maintenance and chromatin remodeling. Despite their expansive roles in biology, there are competing models for how the Rvbs function in these processes. Traditionally the Rvbs have been thought to be helicases. More recently it has been hypothesized that the Rvbs may function as chaperones for multi-subunit complexes, though mechanistic evidence for such an activity is lacking. Here, we identify a potential protein client of the Rvbs, a domain in the ATPase subunit (Ino80INS) of the INO80 chromatin remodeling complex, termed the Ino80INS. Ino80INS is sufficient for interacting with the Rvbs *in vitro*, and stimulates the Rvbs' ATPase activity by 16-fold. Interestingly, Ino80INS activates the Rvbs by promoting dodecamerization of Rvb1/Rvb2. Using a combination of crosslinking mass spectrometry, cryo-EM, native mass spectrometry, and integrative modeling, we find that the Ino80INS binds asymmetrically along the Rvb1/Rvb2 dodecamerization interface, creating a conformationally dynamic, metastable dodecamer that collapses into hexamers upon addition of ATP. We further find this nucleotide-coupled transition between dodecamer and hexamer to be a core activity of the Rvbs. Together, our results demonstrate the chaperone-like potential of Rvb1/Rvb2 and suggest a model where binding of multiple client proteins such as Ino80 stimulates ATP-driven cycling between hexamers and dodecamers, providing iterative opportunities for correct subunit assembly.

Introduction

AAA+ ATPases (ATPases associated with diverse cellular activities) are a highly conserved family of molecular motors that use ATP binding and hydrolysis to drive a diverse set of structural rearrangements in their substrates (1, 2). Well-studied examples include the microtubule translocase dynein, the protein unfoldase ClpX, and the replicative helicase MCM2-

7. Rvb1 (also known as RuvBL1, Tip49a, Pontin, Tih1p, ECP-54, p50, Tap54b) and Rvb2 (also known as RuvBL2, TIP49b, reptin, Tih2p, ECP-51, p47, Tap54a) are two AAA+ ATPases from *S. cerevisiae* that are essential for a wide range of cellular processes including transcription, cellular differentiation, cell signaling, mitosis, snoRNP assembly, and DNA damage repair (3-5). The importance of Rvb1/Rvb2 (Rvbs) in human biology is highlighted by their role in cancer, as both human Rvb1 (RuvBL1) and Rvb2 (RuvBL2) are over-expressed in colon and liver cancers. Inhibition of the ATPase activity of RuvBL2 results in decreased tumor growth, making the Rvbs a potential target for anti-cancer drugs (6-9).

Despite their extensive involvement in biology, the roles of Rvb1 and Rvb2 in cellular processes are still unknown. Even the basic enzymatic function of these motors remains poorly defined. Early studies showed that while Rvb1 and Rvb2 are individually weak ATPases, the activity of Rvb1/Rvb2 complexes is slightly higher, suggesting that the two proteins act cooperatively (10). Furthermore, there are competing models for how ATPase activity contributes to the function of the Rvbs. Based on structural similarity of the Rvb1 and Rvb2 to the bacterial helicase RuvAB, which also forms hexameric rings, it was initially hypothesized that the Rvbs function as DNA helicases (11). While a few groups have been able to detect DNA unwinding activity (10-12), other groups have reported a lack of such an activity (13-15). More recently, it has been proposed that the Rvbs act as protein assembly chaperones, based on the observation that the Rvbs co-purify with a diverse set of multi-subunit protein complexes including the R2TP complex, telomerase, snoRNP complexes and chromatin remodeling complexes (4). However, any chaperone like properties of the Rvbs have not yet been experimentall y demonstrated.

Here we investigate the possibility that the Rvbs act as assembly chaperones. We chose to use the yeast INO80 complex as a model system for this investigation due to the wealth of existing biochemical and structural information on the INO80 family of chromatin remodeling complexes. We found that a small insertion within the Ino80 remodeling ATPase (Ino80INS) is

sufficient for interacting with the Rvbs, creating a Rvb1/Rvb2/Ino80INS subcomplex with a 16-fold higher ATPase activity than Rvbs alone. We further characterized the structural basis of this activation using a combination of cutting-edge mass spectrometry, cryo-EM and integrative modeling. Using crosslinking mass spectrometry, we determined residue-level interactions within the Rvb1/Rvb2/Ino80INS complex and discovered that the Ino80INS binds to a surface of the Rvbs that is known to regulate both activity and oligomeric state of the Rvbs. We next used native mass spectrometry to analyze the distribution of oligomeric states of intact Rvb1/Rvb2 complexes under different nucleotide states. Finally, we determined the cryo-EM structure of the Rvb1/Rvb2/Ino80INS complex and used integrative modeling to comprehensively analyze the structural consequence of Ino80INS binding to the Rvbs. Taken together, these data suggest a model in which (i) the dodecameric Rvb1/Rvb2/Ino8INS state is an activated state that is poised to bring together several protein clients in close proximity and (ii) nucleotide binding and/or hydrolysis collapses the activated intermediate, helping assemble multi-subunit complexes containing a stably associated and less active Rvb1/Rvb2 hexamer.

Methods

Also see supplemental methods, at the end of this chapter, for additional information.

Protein expression and purification

Rvb1 and Rvb2 were co-expressed in Rosetta (DE3) *E. coli* on a pet28a plasmid (a kind gift from the Hopfner lab). Rvb2 had a TEV-cleavable N-terminal 8xhistidine tag, which was only left uncleaved in experiments with Rvb1/his-Rvb2 (Figures 1 and 2). Rvb1/Rvb2 was purified using cobalt affinity resin followed by size exclusion chromatography.

SEC-MALS

Molecular weight was determined using SEC (Wyatt 050S5 column) with an Ettan LC (GE Healthcare) and in-line DAWN HELEOS MALS and Optilab rEX differential refractive index

detectors (Wyatt Technology Corporation). SEC was performed in 80 mM PIPES (pH 6.9), 150 mM KCl, 5 mM $MgCl_2$. Data were analyzed by the ASTRA VI software package (Wyatt Technology Corporation).

XL-MS

Preparations of Rvb1/Rvb2, Rvb1/Rvb2/Ino80INS and Rvb1/Rvb2/Swr1IN were crosslinked with BS3 reagent, digested with trypsin, fractionated by size exclusion chromatography and analyzed by LC-MS using a QExactive Plus mass spectrometer. Spectra were searched for crosslinked peptides using Protein Prospector and further classified using a support vector machine (SVM) model (16, 17).

Native-MS

Native mass spectrometry was performed on the Rvb1/Rvb2 protein complexes using the Exactive Plus EMR instrument (Thermo Scientific, San Jose, CA). The samples were buffer-exchanged into 150 mM ammonium acetate, pH7.5 and then sprayed from Au/Pd-coated borosilicate emitters into mass spectrometer using static infusion.

Negative stain EM

Rvb1/Rvb2/Ino80INS complexes were prepared as described above and dialyzed overnight into sample buffer (25 mM Hepes pH 7.5, 150 mM KCl). Samples for negative stain were adsorbed onto 30-second glow discharged copper grids coated with carbon at concentrations between 5 and 10 μM Rvb monomer in a volume of 2.5 μL. Samples were blotted and stained with 3 μL 0.75% uranyl formate solution for five applications with blotting in between each application. Images were collected on a Technai T12 microscope (FEI Company) equipped with a LaB_6 filament and operated at 120 kV. Images were recorded with an Ultrascan 4096 x 4096 pixel CCD camera (Gatan) at a nominal magnification of 52,000x.

ATPases

ATP hydrolysis rates were measured using an NADH enzyme coupled reaction assay (Lindsley, 2001). The final reaction included 10 U/mL of pyruvate kinase (Sigma) and 10 U/mL of lactate dehydrogenase (LDH, Sigma), 2 mM phosphoenylpyruvate (PEP, Sigma), between 2-10 μM Rvb1/2 complex, 0.18 mM NADH, 4 mM ATP, 4.5 mM $MgCl_2$, 150 mM KCl, and 25 mM Hepes pH 7.5 in a final volume of 80 μL. ATP concentration was empirically determined to be saturating for all constructs tested. Every experiment also included a negative control with no enzyme, which had negligible detected activity. All components other than Rvb1/2 were assembled and incubated at 30°C in a 384 well plate (clear bottom, Corning) for 10 minutes. Enzyme was mixed in, and absorbance at 340 nm was monitored using a SpectraMax M5e plate reader. After background subtraction, the data were fit to a linear regression, with the slope being the rate of hydrolysis (μM/min).

Cryo-electron microscopy image acquisition

Rvb1/Rvb2/Ino80 insertion complexes were prepared as described above and dialyzed overnight into sample buffer. For cryo-EM grid preparation, 2.5 μL samples at an approximate concentration of 10 μM Rvb monomer were applied to Quantifoil holey carbon grids (1.2/1.3 spacing) that had been glow discharged for 30 seconds. Using a Mark I Vitrobot (FEI Company), grids were blotted 5.5 or 6 s. with -3 mm offset at 100% humidity and plunge frozen into liquid ethane cooled by liquid nitrogen.

Images were collected on a Titan Krios microscope (FEI Company) operated at 300 kV under low-dose conditions at Janelia Research Campus. All images were recorded with a K2 Summit direct electron detector camera (Gatan) using super resolution counting mode, with a pixel size of 1.31 Å/pixel after 2x2 binning for motion correction and subsequent processing. The camera dose rate was set to ~1.4 e-/pixel/frame. Dose-fractionated stacks were collected over an

exposure time of 8 s. with a sub-frame exposure of 200 ms. for a total of 40 subframes, resulting in a total dose of ~32.6 e-/Å^2 for the entire stack and 0.8 e-/Å for each subframe.

Cryo-electron microscopy image processing

All dose-fractionated image stacks were motion corrected using MotionCor2 and the sum of all subframes was used for subsequent image processing (18). Estimation of the Contrast Transfer Function and defocus was done using Gctf, a GPU-accelerated CTF estimation program (19). A set of ~1400 particles were manually picked and extracted from corrected micrographs using RELION 1.4. Manually picked particles were subjected to 2D classification in RELION 1.4 and good classes were selected as a template for reference-based automatic particle picking. Multiple rounds of 2D classification yielded a particle stack of ~366,000 particles that was used to generate a consensus model with the RELION 3D auto-refine function, using C1 symmetry. The structure of the *T. acidophilum* 20S proteasome (EMDB 5623) filtered to 50 Å was used an initial model because of its largely cylindrical structure. All subsequent 3D classification was performed using either the consensus model, a model generated by prior 3D classification, or the *C. thermophilum* Rvb1/Rvb2 crystal structure (PDB 4WVY), all filtered to at least 40 Å, as an initial model. Classes of interest were subjected to further 3D classification with C1 symmetry and refined using gold-standard procedures within RELION 1.4 (20).

Integrative modeling of the Rvb1/Rvb2/Ino80INS complex

We computed a structural model of the yeast Rvb1/Rvb2/Ino80INS complex using an integrative approach based on data from X-ray crystallography, cross-linking mass spectrometry and cryo-EM. The complex consists of 2 copies of the Ino80INS-MBP construct and 2 Rvb1/Rvb2 rings, each ring consisting of 3 copies Rvb1 and Rvb2. Our integrative structure determination proceeds through four stages (21, 22) (Figure S6A): (1) gathering of data, (2) representation of subunits and translation of the data into spatial restraints, (3) configurational sampling and

scoring to produce an ensemble of structures that satisfies the restraints, and (4) analysis and validation of the ensemble structures. The modeling protocol (i.e., stages 2, 3, and 4) was scripted using the Python Modeling Interface (PMI), a library for modeling macromolecular complexes based on our open-source Integrative Modeling Platform (IMP) package, version 2.5 (http://integrativemodeling.org) (22). Further details of the integrative modeling procedures are provided in the Supplemental. Files containing the input data, scripts, and output structures are available online (https://github.com/integrativemodeling/Rvbs).

Data Availability: For the cryo-EM structure, the data that support the findings of this study are available from the corresponding author upon reasonable request.

Code Availability: For integrative structure modeling, files containing the input data, scripts, and output structures are available online (https://github.com/integrativemodeling/Rvbs).

Results

The insertion domain of the Ino80 ATPase (Ino80INS) is a potent activator of Rvb1/Rvb2.

In our search for a minimal system to study the potential chaperone-like activity of the Rvbs, we turned to the chromatin remodeling complex INO80. An early study showed that the ATP hydrolysis activity of Rvb2 is required for formation of active INO80 complexes in yeast (23). More recently, it was shown that a small insertion within the Ino80 chromatin remodeling ATPase subunit of the INO80 complex (Ino80INS, Figure 1A, top) is required for association with Rvb1 and Rvb2 and other subunits of the complex that are critical for activity (24). These results suggested that the Ino80INS may be used as a scaffold by the Rvbs for assembling an active INO80 complex. To test this possibility, we asked if the Ino80INS on its own could directly interact with the Rvb1/Rvb2 complex and whether or not this interaction influences the ATPase activity of the Rvbs.

We co-expressed untagged Rvb1 and Rvb2 along with a his-tagged, MBP-Ino80INS fusion in *Escherichia coli* (Figure 1A, top). These three proteins form a stable complex that purifies as a single peak over a size exclusion column (Figure 1A, bottom). This interaction

required the Ino80INS, as MBP alone was unable to pull down the Rvb1/Rvb2 complex (Figure 1A, bottom). Efforts to purify the complex without an MBP tag on the Ino80INS or to purify MBP-Ino80INS alone were unsuccessful, suggesting that even MBP-tagged Ino80INS is inherently unstable, but can be stabilized via interactions with Rvb1 and Rvb2. Next, we tested whether the activity of the Rvbs is affected by the presence of the Ino80INS. We found that the Rvb1/Rvb2/Ino80INS complex has a 16-fold greater k_{cat} for ATP hydrolysis compared to the Rvbs alone. As expected, mutating the residues required for ATP hydrolysis (Walker B) in both Rvb1 and Rvb2 greatly reduced the ATPase activity of the Rvb1/Rvb2/Ino80INS complex (Figure 1B). A fold activation of this magnitude, to our knowledge, is unprecedented in the Rvb literature. In contrast, a large excess of double stranded DNA did not stimulate the Rvbs ATPase activity, consistent with previously published results with human Rvb1/Rvb2 (25). Interestingly, the partially unfolded region of staphylococcal nuclease △131△, which is routinely used as a model substrate for the chaperone Hsp90 (26), modestly stimulated the ATPase activity of the Rvbs (2-3 fold, Figure 1B). This result further suggests that protein, not DNA, is the biologically relevant substrate for yeast Rvb1/Rvb2.

We next used crosslinking mass spectrometry (XL-MS) to investigate the structural basis of the interaction between the Rvbs and the Ino80INS. This technique, which uses chemical crosslinking followed by identification of crosslinked residues by mass spectrometry, allows for the identification of residue-level interactions within a protein complex (27). From two separate XL-MS experiments with the Rvb1/Rvb2/Ino80INS (Figure 1C) and the Rvb1/Rvb2 complexes (Figure S1A), we identified 519 unique cross-linked residue pairs (“crosslinks”) at a false discovery rate of 1.7% (Figure S1B) where each individual experiment contained between 300 and 400 unique crosslinks (Table S1). Strikingly, a majority of the high confidence crosslinks containing the Ino80INS mapped to the DII regions of Rvb1 and Rvb2 (Figure 1C). The DII domains are of particular interest because of their role in regulating the activity and oligomeric state of the Rvbs (28, 29). Importantly, very few residues from Rvb1 and Rvb2 crosslinked to

MBP (Figure 1C), confirming that the complex is predominantly stabilized by direct interactions between the DII regions of Rvb1/Rvb2 and the insertion itself. Also, our data recapitulate over 80% of the interactions found in a cross-linking study of the complete yeast INO80 complex (30), suggesting that the Rvb1/Rvb2/Ino80INS subcomplex maintains the same interactions as seen within the native INO80 complex. Additionally, we identified 324 new crosslinking sites from the subcomplex, (Table S1). In summary, we discovered the Rvb1/Rvb2/Ino80INS as a biologically relevant subcomplex of the full INO80 complex where the Rvbs are in an activated state.

The Ino80 insertion promotes dodecamerization of Rvb1/Rvb2

The Rvbs are predominantly hexameric, but a small population of naturally occurring dodecamers has also been reported (31). Yet, the mechanistic role of either oligomeric state is unknown. Our observation that the Ino80INS binds predominantly to the DII regions of the Rvbs raised the possibility that the Ino80INS may activate the Rvbs' ATPase activity by altering the oligomeric state of Rvb1/Rvb2. To investigate this possibility, we first used size exclusion chromatography-multi angle light scattering (SEC-MALS) to determine the molecular weight of Rvb1/Rvb2/Ino80INS and found that the complex has a molecular weight that is most consistent with a dodecamer of Rvb1/Rvb2 bound to at least two copies of the Ino80 insertion (Figure 2A). In contrast, Rvb1/Rvb2 alone gave a molecular weight most consistent with hexamers. To more precisely analyze the distribution of species in the Rvb1/Rvb2/Ino80INS sample, we turned to native mass spectrometry (native-MS). Unlike traditional mass spectrometry, native-MS allows for analysis of intact protein complexes with very high mass resolution (32-34). Using this technique, we found two distinct charge envelopes in the Rvb1/Rvb2/Ino80INS sample with masses consistent with either two or three copies of the Ino80INS bound to a single Rvb1/Rvb2 dodecamer (Figure 2B, Table S2, Figure S2A-B). Along with the SEC-MALS results, these data suggest that more than one Ino80INS can bind to a single dodecamer of Rvb1/Rvb2. In addition

to being the first biologically activator of the Rvbs, this is the first observation, to our knowledge, of a biologically relevant molecule that promotes dodecamerization of the Rvbs. In addition to the dodecamer, we also observed a small population of Rvb1/Rvb2 hexamer (Species 1 in Figures 2B, S2B, Table S2) as well as Rvb1/Rvb2 hexamer bound to a single MBP-Ino80INS (Species 2 in Figures 2B, S2B, Table S2). These results suggest the existence of multiple states of the Rvb1/Rvb2/Ino80INS complex, which may play a role in the potential chaperone activities of these ATPases (see Discussion).

Intriguingly, addition of ATP-Mg^{2+} to the Rvb1/Rvb2/Ino80INS complex caused the dodecameric species to convert to hexameric species, as shown by native-MS (Figure 2C). The hexamers generated in this process have a molecular mass inconsistent with the presence of Ino80INS (Table S2), suggesting the Ino80 insertions dissociate from the Rvb1/Rvb2 complex after addition of ATP-Mg^{2+}. Dissociation into hexamers was also suggested by negative-stain EM, as barrel shaped Rvb1/Rvb2 dodecamers fall apart to ring-shaped hexamers upon addition of ATP-Mg^{2+} (Figure S2C). These results suggest that destabilization of the activated Rvb1/Rvb2/Ino80INS dodecamer is coupled to the nucleotide state of the Rvbs.

Nucleotide state drives the switch between hexameric and dodecameric forms of Rvb1/Rvb2

Our observation that the Rvb1/Rvb2/Ino80INS complex collapses in an ATP-dependent manner raised the possibility that the transition between dodecamer and hexamer is a core activity of the Rvbs, which may be related to the Rvbs' ability to function as a chaperone. To gain a more complete understanding of how the oligomeric state of the Rvbs themselves are regulated by ATP state, we used native-MS to analyze the relative populations of oligomers of Rvb1/Rvb2 as a function of different nucleotide states. The high mass resolution of this method also allowed us to more accurately determine subunit stoichiometry and the nucleotide binding state of each oligomer. Untagged yeast Rvb1/Rvb2 complexes ionize as hexamers with a mass that is most consistent with a 3:3 stoichiometry of Rvb1: Rvb2 (Figure 3A, Table S3), in contrast

to a 4:2 stoichiometry reported for human Rvbs (28). Species smaller than hexamers were not observed (Figure S3B). To further investigate the stoichiometry of the yeast Rvb1/Rvb2 system, we performed intact denatured mass spectrometry on the same complex. We indeed found an equimolar ratio of Rvb1 and Rvb2 in the complex, confirming our observation that the Rvb1/Rvb2 hexamer contains a 3:3 stoichiometry of Rvb1: Rvb2 (Figure S3C). Adding ATP-Mg^{2+} to Rvb1/Rvb2 did not change the equilibrium between hexamer and dodecamer, though the mass of the hexamer did increase slightly, likely due to nucleotide loading in the active sites (Figure S3D, Table S4).

We next addressed how the Rvb1/Rvb2 dodecamer might be affected by nucleotide state. This experiment could not be done with untagged Rvb1/Rvb2, as the complex is predominantly a hexamer at concentrations compatible with native-MS (1 μM). Previous work with yeast Rvb1/Rvb2 showed that an N-terminal his-tag promotes dodecamerization, providing a means to study the dodecameric state at low concentrations (31, 35). We thus used the his-tag on Rvb2 as a tool to bias the complex towards dodecamer (Figure S3A). As expected, his-tagged Rvbs ionized as a dodecamer with a 6:6 stoichiometry of Rvb1: Rvb2 (Figure 3B). When an excess of ATP-Mg^{2+} was added, a small proportion (14%) of the dodecamer dissociated into hexamers (Figure 3C, Table S4), consistent with our observation for the Rvb1/Rvb2/Ino80INS complex (Figure 2C).

To test which stage(s) of the ATPase cycle affects oligomerization, we used an excess of the non-hydrolyzable ATP analogs AMP-PNP and ATP-γ-S to mimic the ATP bound state, and ADP to mimic the post-hydrolysis state. Surprisingly, AMP-PNP, ATP-γ-S and ADP all substantially increased the proportion of the hexameric state (Figures 3D-E, Figure S3E). Importantly, this increase was greater than what was observed with ATP (Figures 3C vs. 3D-E, Table S4). The quantitative difference in the effects we observed with AMP-PNP vs. ATP does not appear to be due to differences in affinity of the nucleotide for Rvb1/Rvb2, as similar amounts of nucleotide loading was observed for both ATP-Mg^{2+} and AMP-PNP-Mg^{2+} (Table S4).

To further investigate the role of nucleotide binding vs. hydrolysis, we repeated the experiment with mutations in the Walker B motif of both Rvb1 and his-Rvb2, which diminish ATP hydrolysis by at least 20-fold (Figure S3F), but do not reduce ATP binding. Compared to wild-type Rvb1/his-Rvb2 as well as wild-type Rvb1/his-Rvb2 with added ATP, the population of hexamers in the Walker B Rvb1/his-Rvb2 mutants increased dramatically, from 0% to 84% (Figure 3F, Table S4). The molecular weight of the his-tagged Walker B mutant Rvbs is consistent with at least two and six ATP's bound to the hexamer and dodecamer respectively (Table S4). Further addition of excess ATP-Mg^{2+} did not change the relative proportions of the two oligomeric states (Figure S3G). These results suggest a model in which nucleotide binding promotes a conformational state of the Rvb proteins that stabilizes the hexameric state, while an intermediate in the ATP hydrolysis reaction stabilizes a conformation that promotes formation of a metastable dodecamer (Figure 3G).

ATPase activity of the Rvb1/Rvb2 complex is enhanced by dodecamerization

We previously observed that the stimulation of Rvbs by the Ino80INS by 16-fold is concurrent with dodecamerization of Rvb1/Rvb2. We next asked whether this correlation exists in a context outside of Ino80INS. Previous work had shown that either adding an N-terminal his-tag or replacing the DII domain with a flexible linker can increase dodecamerization of Rvbs, but there is conflicting evidence for how these mutants affect the Rvbs' specific activity in the yeast system (28, 29, 31). Using SEC-MALS we confirmed that wild type Rvb1/Rvb2 has a molecular weight most consistent with a hexamer, while the Rvb1/his-Rvb2 complex has a molecular weight consistent with a dodecamer (Figure 4A). Next, we examined the molecular weight of Rvb1ΔDII/Rvb2ΔDII. Importantly, the DII domain constitutes approximately 20% (11 kDa) of the monomer, thus deleting this domain significantly reduces the expected molecular weight of the protein (Table S5). The Rvb1ΔDII/Rvb2ΔDII complex has a molecular weight (341 kDa) that is in between what is expected for a dodecamer (474 kDa) and hexamer (237 kDa) (Figure 4A and

Table S5). We reasoned that the intermediate molecular weight could reflect either a rapidly exchanging mixture of hexamer and dodecamer that is not resolved by the sizing column, or a homogeneous population of an oligomeric state between a dodecamer and hexamer (Table S5). Increasing the concentration of Rvb1ΔDII/Rvb2ΔDII from 10 μM to 30 μM gave a molecular weight (437 kDa) that is approximately the expected molecular weight of a dodecamer (474 kDa, Figure S4A, Table S5). In contrast, wild-type Rvb1/Rvb2 remains hexameric even at concentrations as high as 80 μM (Figure S4B), suggesting that deleting the DII domains results in a dodecamer that is more unstable than the dodecamer formed by Rvb1/his-Rvb2. Consistent with the concentration-dependent oligomerization of Rvb1ΔDII/Rvb2ΔDII, 1 μM of this mutant flew as predominantly hexamer by native-MS, though a small (3%) population of dodecamer did exist (Figure S4C).

We next compared the maximal ATPase activity (k_{cat}) of the different Rvb1/Rvb2 mutants. We found that adding a his-tag to Rvb2 modestly increased the k_{cat} for ATP hydrolysis (2-fold), while deleting DII domains of Rvb1 and Rvb2 more dramatically increased the specific activity (7-fold) (Figure 4B). When considered in combination with the SEC-MALS data, these results suggest that while the ATPase activity is consistently stimulated by dodecamerization, the degree of stimulation is higher for a system in which the dodecamer is metastable and can more readily dissociate into hexamers. Taken together with the data in Figure 2, these data suggest that the ATPase cycle of Rvb1/Rvb2 drives a constant inter-conversion between hexameric and dodecameric states, resulting in a defined steady state distribution of the two oligomeric states with different specific activities (Figure 4C). The regulation of Rvbs' activity by Ino80INS serves as a natural extension of this model, as binding of Ino80INS to the DII domains of the Rvbs promotes an activated, metastable dodecamer that easily dissociates into hexamers upon addition of ATP (Figures 1-2).

Cryo-EM reconstruction of the Rvb1/Rvb2/Ino80INS complex suggests high degree of conformational flexibility

We were next curious about the structural basis of how the Ino80INS engages the DII domains of Rvb1/Rvb2 to create a metastable dodecamer. Previous structural studies of Rvb1/Rvb2 have reported a high degree of conformational flexibility within the DII regions (29, 31, 36-38), which appear to form the dodecameric interface, and to which the Ino80INS binds. This flexibility is most clearly exemplified in a recent crystal structure of the Rvb1/Rvb2 complex from the thermophillic fungus *Chaetomium thermophilum* (39), which shows a state in which the DII domains of Rvb1 and Rvb2 in two different conformations: the Rvb2 DII domains fold back toward the hexameric ring ("in" state), and one in which the Rvb1 DII domains extend toward a second hexamer ("out" state), appearing to form the dodecamerization interface (Figure 4D). We imagined two extreme models for the structural consequence of the Ino80INS binding to the Rvbs: (i) the Ino80INS "locks" the conformationally flexible dodecameric interface into one predominant conformation, or (ii) the Ino80INS binds in a way that maintains the natural conformational flexibility of the DII domains.

To test these hypotheses, we determined the cryo-EM density map of the Rvb1/Rvb2/Ino80INS complex at ~12.0 Å resolution (Figure 5B, S5B). Both raw images and 2D class averages show a distinct barrel shape created by the two hexameric rings flanking the DII domains of the Rvbs (Figure 5A, S5A). We also observed both bent and straight conformations of the Rvb1/Rvb2/Ino80INS dodecamer in our 2D classes (Figure 5A), likely due to flexibility within the DII domains that sit in between the two hexameric rings. These features are similar to those reported in a recently published cryo-EM structure of Rvb1/Rvb2 dodecamers without an activator bound (38, 40), in support of our hypothesis that Ino80INS binding does not diminish the natural dynamics of the Rvbs. Our observation that the binding of the Ino80INS does not

grossly change the overall shape or dynamics of the Rvb1/Rvb2 dodecamer is also consistent with the Ino80INS being a natural activator for the Rvbs.

Next, we further investigated the extent of the conformational dynamics in the Rvb1/Rvb2/Ino80INS complex. All particles used for 2D classification were also used for 3D refinement to generate a consensus cryo-EM map of the Rvb1/Rvb2/Ino80INS complex (Figure 5B), with no imposed symmetry. Interestingly, the bent feature observed in the 2D classes is retained in the consensus map, suggesting that it is a common feature of most of the particles. In an effort to find higher-resolution subpopulations of the dataset, we further subclassified particles to generate twelve different 3D classes (Figure 5C). Each of the twelve classes was populated with a similar number of particles and the resolutions measured between classes were similar to one another and to that for the consensus map (Figure 5C). Despite many rounds of further subclassification, we were unable to achieve higher resolution, implying a high degree of conformational heterogeneity in the system. This conformational heterogeneity is most prominent in the densities surrounding the DII domains, as well as in the degree of tilt observed between the two hexameric rings (Figure 5C). In addition to the conformational heterogeneity, a second general feature of the Rvb1/Rvb2/Ino80INS system is its asymmetry. This asymmetry is retained in both the consensus map and each of the twelve subclasses, as one face of the dodecamer (Face A, Figures 5B-C) consistently shows bulkier density than the opposite face of the dodecamer (Face B, Figures 5B-C). Since this type of asymmetry in the dodecameric interface has not been previously reported for the Rvb1/Rvb2 system, we reasoned it was at least in part due to binding of Ino80INS.

Ino80INS bind asymmetrically on the Rvb1/Rvb2 dodecamer

To obtain more detailed insight into the conformational flexibility and asymmetry of the Rvb1/Rvb2/Ino80INS complex, we mapped all unique Rvb-Rvb crosslinks (residue pairs) observed in the Rvb1/Rvb2/Ino80INS complex onto a comparative structural model of the yeast

Rvb1/Rvb2 dodecamer. We created this model based on sequence identities between the *S. cerevisiae* Rvb1 and Rvb2 subunits and those found in the *S. thermophilus* structure of the Rvb1/Rvb2 dodecamer (39) (see Methods). Strikingly, this mapping revealed that only 132 (58%) of the unique 226 Rvb-Rvb crosslinks are satisfied (Figure S7A). It is unlikely that the low percentage of satisfied crosslinks is due to inaccurate spectral assignments given the low false discovery rate (FDR) of our crosslinking experiment (less than 2%, Figure S3B). It is also unlikely that our comparative model of the Rvb1/Rvb2 dodecamer is insufficiently accurate to explain the XL violations, given the high sequence similarity between the *S. cerevisiae* and S. *thermophilus* Rvbs (73% for Rvb1 and 72% for Rvb2) and the high structural conservation in this family (41). Thus, the high proportion of crosslinks violated by the model is likely due to the additional conformations that the Rvb1/Rvb2/Ino80INS complex can adopt beyond the one exhibited in the comparative model of the Rvb1/Rvb2 dodecamer. Importantly, the violated unique crosslinks are defined by lower spectral counts (average spectral counts of 1.5 for violated crosslinks versus 7.1 for satisfied), suggesting that these additional conformations of Rvb1/Rvb2/Ino80INS may be in relatively low abundance in the system. Interestingly, of the violated crosslinks, 73% are between or to residues in the DII domains, consistent with the types of conformational heterogeneity observed in our cryo-EM data.

To more precisely characterize this conformational heterogeneity, we applied an integrative structure determination approach based on all information available for the Rvb1/Rvb2/Ino80INS complex including the XL-MS dataset (Figures 3C, Table S4), the consensus cryo-EM map (Figure 5B), the comparative model of the Rvb1/Rvb2 dodecamer, the crystal structure of MBP (42), and the predicted secondary structure of Ino80INS (Figure S6, Table S6, Supplementary Methods). The resulting integrative models are expected to be more accurate and precise than models based on a subset of data (16, 21, 22, 43). We imposed a stoichiometry of two copies of Ino80INS bound to one Rvb1/Rvb2 dodecamer based on the major species found by native-MS (Figure 2C). The ensemble of models, consistent with input

information, was computed by satisfying spatial restraints implied by the data, allowing us to quantify the structural heterogeneity and asymmetry observed in the cryo-EM data as well as to predict the binding site of the Ino80INS.

The ensemble contains three clusters of models, each of which represent possible conformations of the Rvb1/Rvb2/Ino80INS dodecamer. For Cluster 1, the average pairwise RMSD of the models in the cluster (i.e., precision) is 25 Å, and the precisions of the Rvb1/Rvb2 dodecamer and Ino80INS in particular are 15.8 and 39.3 Å, respectively. These values represent the average fluctuations of the individual residues or beads in 3-D space across the ensemble of solutions. The precisions of the other two clusters are comparable (Table S6). Each cluster of models satisfies the XLs approximately within the expected tolerance (Table S6): 97%, 94%, and 91% of all crosslinks are satisfied by at least one structure in Clusters 1, 2, and 3, respectively (Figure 6A), and 98% of all crosslinks are satisfied by at least one structure in the three clusters. The remaining 2% unsatisfied cross-links are likely explained by the 2% false positive rate of cross-linking, sample heterogeneity, insufficient conformational sampling, and coarse-grained representation of the modeled components. Each of the three clusters also matches the EM density map, as demonstrated by the overlaps of the cluster localization density maps with the consensus cryo-EM density map (cross-correlation of 0.94, 0.96, and 0.87, respectively), when all maps are contoured to match their volumes (Figure 6B). A localization density map for a set of models is defined as the probability of observing a model component at any point in space (16).

Because each cluster satisfies the cross-links and EM map comparably well, and each cluster contains a comparable number of solutions, it is likely that the localization density maps represent the underlying conformational heterogeneity in the sample rather than a lack of information used for modeling. Therefore, it is also likely that the actual sample contains the structures represented by all three clusters. In all clusters, Ino80INS binds to the central region of the barrel, contacting the DII domains (Figure 6A and S7C). Furthermore, the two Ino80INS

domains bind asymmetrically on the Rvb1/Rvb2 dodecamer, contacting the same side of the barrel. Examination of the MBP crosslinks shows that some crosslinks are necessarily inter-molecular, suggesting that the two MBP domains are close to each other, though they do not necessarily interact directly with one another. Due to the proximity of the Ino80 and MBP domains in the linear sequence, this observation further supports a model of asymmetric binding of the INO80 insertions to the Rvbs dodecamer (Figure S7B).

Discussion

Despite being essential for a wide array of cellular processes, the biochemical function of the Rvbs remains elusive. Here we identify a new activity of yeast Rvb1/Rvb2 that is likely related to the proposed chaperone-like potential of these complexes. We found a region of the Ino80 ATPase that is sufficient for interacting with the Rvbs and activates the Rvbs' ATPase activity by creating a metastable dodecamer of Rvb1/Rvb2 that dissociates into hexamers upon addition of ATP. We also present data suggesting that the ATP-driven cycling between hexamers and dodecamers is a core feature of the yeast Rvb1/Rvb2 system. Implications of these findings are discussed below.

In addition to the INO80 complex, Rvb1 and Rvb2 have been found to associate with a number of multi-subunit complexes, including snoRNP complexes (44), PIKK signaling complexes (45) and telomerase (46), leading to the hypothesis that Rvbs may act as chaperones for assembling and/or remodeling these complexes (4). Based on our data and prior work, we propose the following model for how the Rvbs may help assemble multi-subunit complexes such as INO80 (Figure 7). In this model, nucleotide binding (ATP or ADP alone) is coupled to a conformation of the Rvb monomer that is autoinhibitory for ATP hydrolysis and incompatible with dodecamerization (Figure 3G). In contrast, ATP hydrolysis (either the transition state or the ADP-P_i product state) is coupled to a conformation of the Rvb monomer that promotes dodecamerization (Figure 3G). Binding of clients such as the Ino80INS would

then drive the Rvbs into the more active dodecameric state, yet one that is metastable and is more likely to dissociate into hexamers (Figure 2). Thus, a window of opportunity is created for other client proteins to bind to adjacent DII domains of the Rvb1/Rvb2 complex. We speculate that the Rvb1/Rvb2 dodecamer is able to "sense" co-localization of multiple client proteins by coupling the association of these proteins to nucleotide-driven dissociation of the dodecamer into hexamers. This model is consistent with our observation that two copies of Ino80INS can bind asymmetrically to one side of the Rvb1/Rvb2 dodecamer (Figure 6). We speculate that this observation reflects the ability of the Rvb1/Rvb2 dodecamer to bind to multiple different client proteins *in vivo*.

A major feature of our model is that binding of a client protein such as the Ino80INS promotes formation of a metastable dodecamer that is prone to collapsing into hexamers upon nucleotide binding. Our cryo-EM and integrative modeling data provide structural insight into how such a metastable dodecamer is created, as the dodecameric interface is highly flexible and dynamic (Figures 5 and 6). The ability of ATP to further destabilize this interface is consistent with recently published cryo-EM structures of the Rvb1/Rvb2 dodecamers alone, which suggest that addition of nucleotide causes a significant reconfiguration of the DII domains, resulting in a shift in the relative orientations of the two hexameric rings (38). We therefore speculate that creation of a metastable dodecamer though reconfiguration of the DII domains is important for both the ATP-driven cycling between hexamers and dodecamers and the chaperone-like activity of the Rvbs.

Based this model we would most simply expect that a hexameric Rvb1/Rvb2 bound to Ino80INS would be created during the ATPase cycle of the Rvbs. Instead, we only observe Rvb1/Rvb2 hexamers by native-MS upon addition of ATP (Figure 2C). We speculate that the different steps in ATP hydrolysis cause the Rvbs to cycle between strong and weak binding for their client proteins analogous to other chaperones such as Hsp90 (26, 47). In the context of the full-length Ino80 there would be additional interactions to keep it bound to the Rvbs during the

ATPase cycle, whereas in the context of the Ino80 insert, these additional binding interactions could be missing, causing it to dissociate.

Our model also addresses the potentially separate functions of the hexameric and dodecameric forms of the Rvbs. While both of these oligomeric states been observed in native Rvb1/Rvb2 systems (31, 38), mechanistic roles for either oligomeric state have not yet been proposed. We speculate that the dodecameric Rvb1/Rvb2/Ino80INS complex is representative of an intermediate that is on pathway toward assembling a multi-subunit INO80 complex. Completion of assembly would then result in the stable binding of the less active, hexameric state of Rvb1/Rvb2 to the rest of the INO80 complex. In this model, the transition from dodecamer to hexamer is governed by both the presence of bound client proteins as well as the ATPase cycle, ultimately resulting in an equilibrium among dodecameric assembly intermediates and fully assembled complexes containing a hexamer of Rvb1/Rvb2. Importantly, the predominant (>85%) stoichiometry of the components within the Rvb1/Rvb2/Ino80INS complex (6:6:2) is consistent with recently published EM data of the full INO80 complex, where a large number of the INO80 complexes contain a single Ino80 ATPase associated with a hexamer of Rvb1/Rvb2 (48).

We favor the model described above to explain our results because previous work has suggested that the Rvbs are required for assembly of active INO80 complexes (Chen et al., 2013; Jónsson et al., 2004). However, we propose at least two alternative models that are also compatible with our results. In the first alternative model, the ATPase stimulation caused by Ino80INS serves to limit the number of Ino80 proteins bound, as increasing the number of bound Ino80 proteins would further destabilize the already metastable dodecamer. In this model, the Rvbs would primarily play a role in regulating subunit stoichiometry. In the second alternative model, the ATPase stimulation caused by Ino80INS may reflect the Rvbs' potential role in protein homeostasis, as Rvb1 and Rvb2 have been hypothesized to be involved in

disaggregation of amyloid fibrils (49) and chromosome decondensation (50). Importantly, our observation that protein, not nucleic acid, is a robust activator of Rvb1/Rvb2 helps exclude models involving a helicase- like role of the Rvbs *in vivo.*

While we believe we have identified the first client of Rvb1/Rvb2, further studies are required to fully demonstrate the chaperone-like potential of the yeast Rvbs. In the future, an ability to reconstitute all the steps in the reaction catalyzed by the Rvbs would be invaluable for obtaining insights into the mechanism of the process. In learning from more well-studied protein assembly chaperones, such as Hsp90, it is clear that such reconstitution is non-trivial due in part to the challenges of identifying an appropriate and well-behaved model system. In the context of the Rvbs we anticipate that the yeast INO80 complex can serve as a powerful model system for future mechanistic studies for the following reasons: (i) the organization and stoichiometry of the INO80 subunits within the native complex, down to residue-level interactions, has been fully characterized (30) and (ii) the proper assembly of INO80 is required for *in vitro* nucleosome remodeling activity (24, 51). Thus, the INO80 complex provides an ideal opportunity for a complete *in vitro* reconstitution of the steps required in assembling an active multi-subunit complex. These experiments, while technically challenging, would provide unprecedented mechanistic insight into this class of essential yet poorly understood AAA+ ATPases.

Acknowledgments

We thank the Hopfner lab (Ludwig-Maximilians University in Munich) for the expression plasmid for Rvb1 and Rvb2. We thank Lindsey Pack, Scott Coyle, Phillip Dumesic, Laura Hsieh, Stephanie Johnson and Daniele Canzio for critical reading of the manuscript, and members of the Narlikar lab for helpful discussions. We thank Daniel Elnatan for teaching C.Y.Z. how to perform the NADH enzyme coupled ATPase assay and for helpful discussions about ATPases, and Andrew Lyon for teaching C.Y.Z. how to perform SEC-MALS experiments. This work was

funded by a National Institutes of Health Ruth L. Kirschstein National Research Service Award (5F31CA180651-02) to C.Y.Z., a National Science Foundation fellowship (1144247) to C.I.S., a P41 center grant (8P41GM10348) to A.L.B., and a UCSF Enabling Technology Advisory Committee Instrument Grant (1S10D016229) to A.L.B., and the NIH R01 (GM083960) and P41 (GM083960) grants to A.S.

Author contributions

C.Y.Z. and G.J.N conceptualized the study and designed most of the experiments presented in the manuscript. C.Y.Z. cloned and purified the proteins, performed ATPase assays and SEC-MALS experiments. J.B.J. performed the native-MS experiments and M.J.T. performed the XL-MS experiments, created homology models of Rvb1/Rvb2, and analyzed the data. C.I.S. performed the EM experiments and analyzed the data with help from E.P. I.E. performed the integrative modeling. All authors contributed to preparation of the manuscript. Y.C., A.S., and A.L.B. provided expertise and support for the project.

References

1. Singleton, M. R., Dillingham, M. S., and Wigley, D. B. (2007) Structure and mechanism of helicases and nucleic acid translocases. *Annu. Rev. Biochem.* 76, 23–50

2. White, S. R., and Lauring, B. (2007) AAA+ ATPases: achieving diversity of function with conserved machinery. *Traffic* 8, 1657–1667

3. Matias, P. M., Baek, S. H., Bandeiras, T. M., Dutta, A., Houry, W. A., Llorca, O., and Rosenbaum, J. (2015) The AAA+ proteins Pontin and Reptin enter adult age: from understanding their basic biology to the identification of selective inhibitors. *Front Mol Biosci* 2, 17

4. Nano, N., and Houry, W. A. (2013) Chaperone-like activity of the AAA+ proteins Rvb1 and Rvb2 in the assembly of various complexes. *Philos. Trans. R. Soc. Lond., B, Biol. Sci.* 368, 20110399

5. Rosenbaum, J., Baek, S. H., Dutta, A., Houry, W. A., Huber, O., Hupp, T. R., and Matias, P. M. (2013), p mr1.

6. Grigoletto, A., Lestienne, P., and Rosenbaum, J. (2011) The multifaceted proteins Reptin and Pontin as major players in cancer. *Biochim. Biophys. Acta* 1815, 147–157

7. Grigoletto, A., Neaud, V., Allain-Courtois, N., Lestienne, P., and Rosenbaum, J. (2013) The ATPase activity of reptin is required for its effects on tumor cell growth and viability in hepatocellular carcinoma. *Mol. Cancer Res.* 11, 133–139

8. Huber, O., Ménard, L., Haurie, V., Nicou, A., Taras, D., and Rosenbaum, J. (2008) Pontin and reptin, two related ATPases with multiple roles in cancer. *Cancer Res.* 68, 6873–6876

9. Ménard, L., Taras, D., Grigoletto, A., Haurie, V., Nicou, A., Dugot-Senant, N., Costet, P., Rousseau, B., and Rosenbaum, J. (2010) In vivo silencing of Reptin blocks the progression of human hepatocellular carcinoma in xenografts and is associated with replicative senescence. *J. Hepatol.* 52, 681–689

10. Gribun, A., Cheung, K. L. Y., Huen, J., Ortega, J., and Houry, W. A. (2008) Yeast Rvb1 and Rvb2 are ATP-dependent DNA helicases that form a heterohexameric complex. *J. Mol. Biol.* 376, 1320–1333

11. Kanemaki, M., Kurokawa, Y., Matsu-ura, T., Makino, Y., Masani, A., Okazaki, K., Morishita, T., and Tamura, T. A. (1999) TIP49b, a new RuvB-like DNA helicase, is included in a complex together with another RuvB-like DNA helicase, TIP49a. *J. Biol. Chem.* 274, 22437–22444

12. Makino, Y., Kanemaki, M., Kurokawa, Y., Koji, T., and Tamura, T. A. (1999) A rat RuvB-like protein, TIP49a, is a germ cell-enriched novel DNA helicase. *J. Biol. Chem.* 274, 15329–15335

13. Ikura, T., Ogryzko, V. V., Grigoriev, M., Groisman, R., Wang, J., Horikoshi, M., Scully, R., Qin, J., and Nakatani, Y. (2000) Involvement of the TIP60 histone acetylase complex in DNA repair and apoptosis. *Cell* 102, 463–473

14. Matias, P. M., Gorynia, S., Donner, P., and Carrondo, M. A. (2006) Crystal structure of the human AAA+ protein RuvBL1. *J. Biol. Chem.* 281, 38918–38929

15. Qiu, X. B., Lin, Y. L., Thome, K. C., Pian, P., Schlegel, B. P., Weremowicz, S., Parvin, J. D., and Dutta, A. (1998) An eukaryotic RuvB-like protein (RUVBL1) essential for growth. *J. Biol. Chem.* 273, 27786–27793

16. Robinson, P. J., Trnka, M. J., Pellarin, R., Greenberg, C. H., Bushnell, D. A., Davis, R., Burlingame, A. L., Sali, A., and Kornberg, R. D. (2015) Molecular architecture of the yeast Mediator complex. *Elife* 4, 44

17. Trnka, M. J., Baker, P. R., Robinson, P. J. J., Burlingame, A. L., and Chalkley, R. J. (2014) Matching cross-linked peptide spectra: only as good as the worse identification. *Mol. Cell Proteomics* 13, 420–434

18. Zheng, S. Q., Palovcak, E., Armache, J.-P., Verba, K. A., Cheng, Y., and Agard, D. A. (2016) MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat Meth* 14, 331–332

19. Zhang, K. (2016) Gctf: Real-time CTF determination and correction. *J. Struct. Biol.* 193, 1–12

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

21. Alber, F., Dokudovskaya, S., Veenhoff, L. M., Zhang, W., Kipper, J., Devos, D., Suprapto, A., Karni-Schmidt, O., Williams, R., Chait, B. T., Rout, M. P., and Sali, A. (2007) Determining the architectures of macromolecular assemblies. *Nature* 450, 683–694

22. Russel, D., Lasker, K., Ben Webb, Velázquez-Muriel, J., Tjioe, E., Schneidman-Duhovny, D., Peterson, B., and Sali, A. (2012) Putting the Pieces Together: Integrative Modeling Platform Software for Structure Determination of Macromolecular Assemblies. *PLOS Biol* 10, e1001244

23. Jónsson, Z. O., Jha, S., Wohlschlegel, J. A., and Dutta, A. (2004) Rvb1p/Rvb2p recruit Arp5p and assemble a functional Ino80 chromatin remodeling complex. *Mol. Cell* 16, 465–477

24. Chen, L., Conaway, R. C., and Conaway, J. W. (2013) 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.* 110, 20497–20502

25. Puri, T., Wendler, P., Sigala, B., Saibil, H., and Tsaneva, I. R. (2007) Dodecameric structure and ATPase activity of the human TIP48/TIP49 complex. *J. Mol. Biol.* 366, 179–192

26. Street, T. O., Lavery, L. A., and Agard, D. A. (2011) Substrate binding drives large-scale conformational changes in the Hsp90 molecular chaperone. *Mol. Cell* 42, 96–105

27. Leitner, A., Faini, M., Stengel, F., and Aebersold, R. (2016) Crosslinking and Mass Spectrometry: An Integrated Technology to Understand the Structure and Function of Molecular Machines. *Trends Biochem. Sci.* 41, 20–32

28. Niewiarowski, A., Bradley, A. S., Gor, J., McKay, A. R., Perkins, S. J., and Tsaneva, I. R. (2010) Oligomeric assembly and interactions within the human RuvB-like RuvBL1 and RuvBL2 complexes. *Biochem. J.* 429, 113–125

29. Gorynia, S., Bandeiras, T. M., Pinho, F. G., McVey, C. E., Vonrhein, C., Round, A., Svergun, D. I., Donner, P., Matias, P. M., and Carrondo, M. A. (2011) Structural and functional insights into a dodecameric molecular machine - the RuvBL1/RuvBL2 complex. *J. Struct. Biol.* 176, 279–291

30. Tosi, A., Haas, C., Herzog, F., Gilmozzi, A., Berninghausen, O., Ungewickell, C., Gerhold, C. B., Lakomek, K., Aebersold, R., Beckmann, R., and Hopfner, K.-P. (2013) Structure and subunit topology of the INO80 chromatin remodeler and its nucleosome complex. *Cell* 154, 1207–1219

31. Jeganathan, A., Leong, V., Zhao, L., Huen, J., Nano, N., Houry, W. A., and Ortega, J. (2015) Yeast Rvb1 and Rvb2 Proteins Oligomerize As a Conformationally Variable Dodecamer with Low Frequency. *J. Mol. Biol*,

32. Heck, A. J. R. (2008) Native mass spectrometry: a bridge between interactomics and structural biology. *Nat. Methods* 5, 927–933

33. Rose, R. J., Damoc, E., Denisov, E., Makarov, A., and Heck, A. J. R. (2012) High-sensitivity Orbitrap mass analysis of intact macromolecular assemblies. *Nat. Methods* 9, 1084–1086

34. Konermann, L., Vahidi, S., and Sowole, M. A. (2014) Mass spectrometry methods for studying structure and dynamics of biological macromolecules. *Anal. Chem.* 86, 213–232

35. Cheung, K. L. Y., Huen, J., Kakihara, Y., Houry, W. A., and Ortega, J. (2010) Alternative oligomeric states of the yeast Rvb1/Rvb2 complex induced by histidine tags. *J. Mol. Biol.* 404, 478–492

36. Petukhov, M., Dagkessamanskaja, A., Bommer, M., Barrett, T., Tsaneva, I., Yakimov, A., Queval, R., Shvetsov, A., Khodorkovskiy, M., Käs, E., and Grigoriev, M. (2012) Large-scale conformational flexibility determines the properties of AAA+ TIP49 ATPases. *Structure* 20, 1321–1331

37. López-Perrote, A., Muñoz-Hernández, H., Gil, D., and Llorca, O. (2012) Conformational transitions regulate the exposure of a DNA-binding domain in the RuvBL1-RuvBL2 complex. *Nucleic Acids Res.* 40, 11086–11099

38. Ewens, C. A., Su, M., Zhao, L., Nano, N., Houry, W. A., and Southworth, D. R. (2016) Architecture and Nucleotide-Dependent Conformational Changes of the Rvb1-Rvb2 AAA+ Complex Revealed by Cryoelectron Microscopy. *Structure* 24, 657–666

39. Lakomek, K., Stoehr, G., Tosi, A., Schmailzl, M., and Hopfner, K.-P. (2015) Structural Basis for Dodecameric Assembly States and Conformational Plasticity of the Full-Length AAA+ ATPases Rvb1·Rvb2. *Structure*, 23, 483-495

40. Silva-Martin, N., Daudén, M. I., Glatt, S., Hoffmann, N. A., Kastritis, P., Bork, P., Beck, M., and Müller, C. W. (2016) The Combination of X-Ray Crystallography and Cryo-Electron Microscopy Provides Insight into the Overall Architecture of the Dodecameric Rvb1/Rvb2 Complex. *PLoS ONE* 11, e0146457

41. Iyer, L. M., Leipe, D. D., Koonin, E. V., and Aravind, L. (2004) Evolutionary history and higher order classification of AAA+ ATPases. *J. Struct. Biol.* 146, 11–31

42. Kim, M. (2013) Beta conformation of polyglutamine track revealed by a crystal structure of Huntingtin N-terminal region with insertion of three histidine residues. *Prion* 7, 221–228

43. Shi, Y., Fernandez-Martinez, J., Tjioe, E., Pellarin, R., Kim, S. J., Williams, R., Schneidman-Duhovny, D., Sali, A., Rout, M. P., and Chait, B. T. (2014) Structural characterization by cross-linking reveals the detailed architecture of a coatomer-related heptameric module from the nuclear pore complex. *Mol. Cell Proteomics* 13, 2927–2943

44. McKeegan, K. S., Debieux, C. M., and Watkins, N. J. (2009) Evidence that the AAA+ proteins TIP48 and TIP49 bridge interactions between 15.5K and the related NOP56 and NOP58 proteins during box C/D snoRNP biogenesis. *Mol. Cell. Biol.* 29, 4971–4981

45. Izumi, N., Yamashita, A., Iwamatsu, A., Kurata, R., Nakamura, H., Saari, B., Hirano, H., Anderson, P., and Ohno, S. (2010) AAA+ proteins RUVBL1 and RUVBL2 coordinate PIKK activity and function in nonsense-mediated mRNA decay. *Sci Signal* 3, ra27–ra27

46. Venteicher, A. S., Meng, Z., Mason, P. J., Veenstra, T. D., and Artandi, S. E. (2008) Identification of ATPases pontin and reptin as telomerase components essential for holoenzyme assembly. *Cell* 132, 945–957

47. Verba, K. A., Wang, R. Y.-R., Arakawa, A., Liu, Y., Shirouzu, M., Yokoyama, S., and Agard, D. A. (2016) Atomic structure of Hsp90-Cdc37-Cdk4 reveals that Hsp90 traps and stabilizes an unfolded kinase. *Science* 352, 1542–1547

48. Watanabe, S., Tan, D., Lakshminarasimhan, M., Washburn, M. P., Erica Hong, E.-J., Walz, T., and Peterson, C. L. (2015) Structural analyses of the chromatin remodelling enzymes INO80-C and SWR-C. *Nat Commun* 6, 7108

49. Zaarur, N., Xu, X., Lestienne, P., Meriin, A. B., McComb, M., Costello, C. E., Newnam, G. P., Ganti, R., Romanova, N. V., Shanmugasundaram, M., Silva, S. T., Bandeiras, T. M., Matias, P. M., Lobachev, K. S., Lednev, I. K., Chernoff, Y. O., and Sherman, M. Y. (2015) RuvbL1 and RuvbL2 enhance aggresome formation and disaggregate amyloid fibrils. *EMBO J.* 34, 2363–2382

50. Magalska, A., Schellhaus, A. K., Moreno-Andrés, D., Zanini, F., Schooley, A., Sachdev, R., Schwarz, H., Madlung, J., and Antonin, W. (2014) RuvB-like ATPases function in chromatin decondensation at the end of mitosis. *Dev. Cell* 31, 305–318

51. Jónsson, Z. O., Dhar, S. K., Narlikar, G. J., Auty, R., Wagle, N., Pellman, D., Pratt, R. E., Kingston, R., and Dutta, A. (2001) Rvb1p and Rvb2p are essential components of a chromatin remodeling complex that regulates transcription of over 5% of yeast genes. *J. Biol. Chem.* 276, 16279–16288

52. Webb, B., Lasker, K., Schneidman-Duhovny, D., Tjioe, E., Phillips, J., Kim, S. J., Velázquez-Muriel, J., Russel, D., and Sali, A. (2011) Modeling of proteins and their assemblies with the integrative modeling platform. *Methods Mol. Biol.* 781, 377–397

53. Kim, S. J., Fernandez-Martinez, J., Sampathkumar, P., Martel, A., Matsui, T., Tsuruta, H., Weiss, T. M., Shi, Y., Markina-Inarrairaegui, A., Bonanno, J. B., Sauder, J. M., Burley, S. K., Chait, B. T., Almo, S. C., Rout, M. P., and Sali, A. (2014) Integrative structure-function mapping of the nucleoporin Nup133 suggests a conserved mechanism for membrane anchoring of the nuclear pore complex. *Mol. Cell Proteomics* 13, 2911–2926

54. LoPiccolo, J., Kim, S. J., Shi, Y., Wu, B., Wu, H., Chait, B. T., Singer, R. H., Sali, A., Brenowitz, M., Bresnick, A. R., and Backer, J. M. (2015) Assembly and Molecular Architecture of the Phosphoinositide 3-Kinase p85α Homodimer. *J. Biol. Chem.* 290, 30390–30405

55. Luo, J., Cimermančič, P., Viswanath, S., Ebmeier, C. C., Kim, B., Dehecq, M., Raman, V., Greenberg, C. H., Pellarin, R., Sali, A., Taatjes, D. J., Hahn, S., and Ranish, J. (2015) Architecture of the Human and Yeast General Transcription and DNA Repair Factor TFIIH. *Mol. Cell* 59, 794–806

56. Sali, A., and Blundell, T. L. (1993) Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* 234, 779–815

57. Söding, J., Biegert, A., and Lupas, A. N. (2005) The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res.* 33, W244–8

58. Pieper, U., Webb, B. M., Dong, G. Q., Schneidman-Duhovny, D., Fan, H., Kim, S. J., Khuri, N., Spill, Y. G., Weinkam, P., Hammel, M., Tainer, J. A., Nilges, M., and Sali, A. (2014) ModBase, a database of annotated comparative protein structure models and associated resources. *Nucleic Acids Res.* 42, D336–46

59. McGuffin, L. J., Bryson, K., and Jones, D. T. (2000) The PSIPRED protein structure prediction server. *Bioinformatics* 16, 404–405

60. Drozdetskiy, A., Cole, C., Procter, J., and Barton, G. J. (2015) JPred4: a protein secondary structure prediction server. *Nucleic Acids Res.* 43, W389–94

61. Ward, J. J., McGuffin, L. J., Bryson, K., Buxton, B. F., and Jones, D. T. (2004) The DISOPRED server for the prediction of protein disorder. *Bioinformatics* 20, 2138–2139

62. Kawabata, T. (2008) Multiple subunit fitting into a low-resolution density map of a macromolecular complex using a gaussian mixture model. *Biophys. J.* 95, 4643–4658

63. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, É. (2011) Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research* 12, 2825–2830

64. Algret, R., Fernandez-Martinez, J., Shi, Y., Kim, S. J., Pellarin, R., Cimermančič, P., Cochet, E., Sali, A., Chait, B. T., Rout, M. P., and Dokudovskaya, S. (2014) Molecular architecture and function of the SEA complex, a modulator of the TORC1 pathway. *Mol. Cell Proteomics* 13, 2855–2870

65. Rieping, W., Habeck, M., and Nilges, M. (2005) Inferential structure determination. *Sci Signal* 309, 303–306

66. Shen, M.-Y., and Sali, A. (2006) Statistical potential for assessment and prediction of protein structures. *Protein Sci.* 15, 2507–2524

67. Shi, Y., Pellarin, R., Fridy, P. C., Fernandez-Martinez, J., Thompson, M. K., Li, Y., Wang, Q. J., Sali, A., Rout, M. P., and Chait, B. T. (2015) A strategy for dissecting the architectures of native macromolecular assemblies. *Nat. Methods* 12, 1135–1138

68. Erzberger, J. P., Stengel, F., Pellarin, R., Zhang, S., Schaefer, T., Aylett, C. H. S., Cimermančič, P., Boehringer, D., Sali, A., Aebersold, R., and Ban, N. (2014) Molecular architecture of the 40S·eIF1·eIF3 translation initiation complex. *Cell* 158, 1123–1135

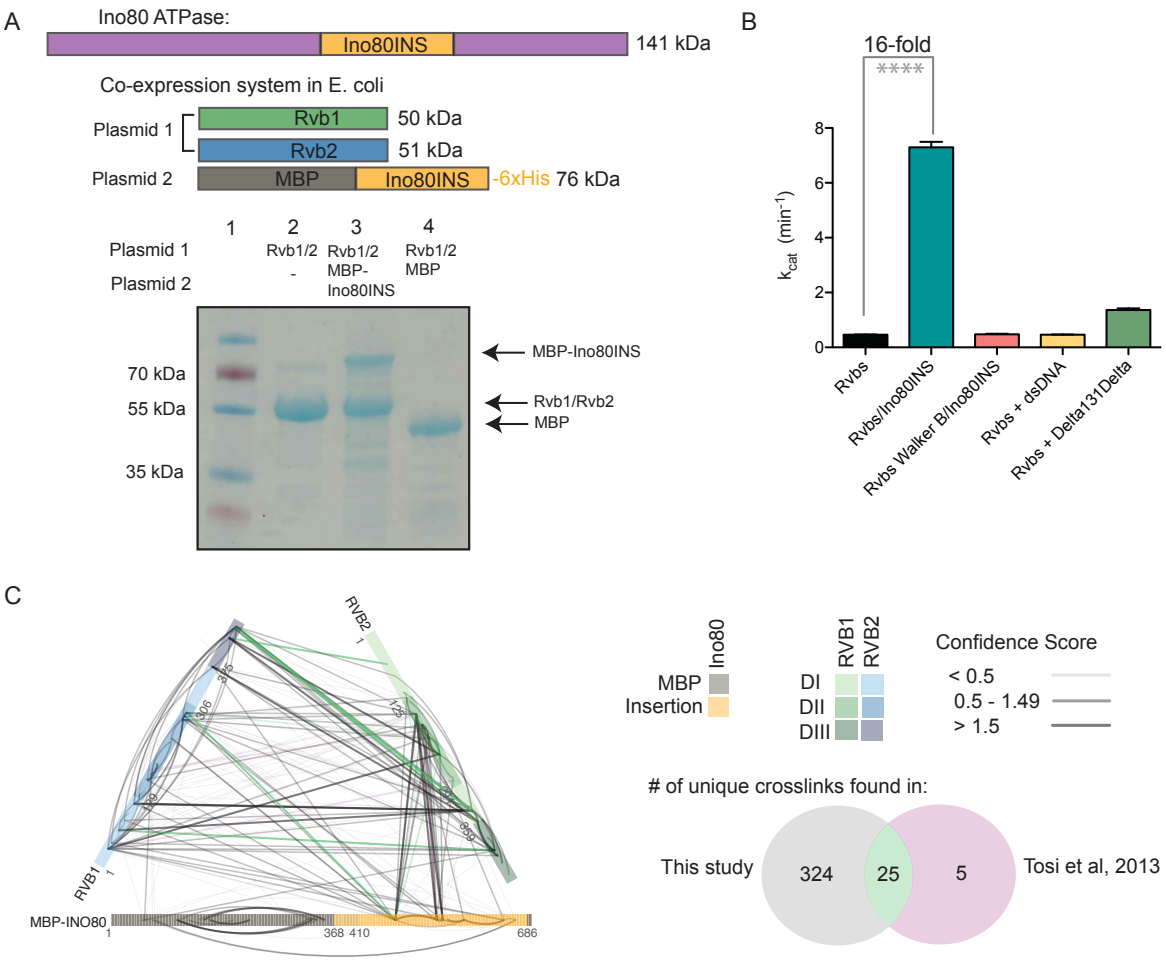
69. Sali, A., Berman, H. M., Schwede, T., Trewhella, J., Kleywegt, G., Burley, S. K., Markley, J., Nakamura, H., Adams, P., Bonvin, A. M. J. J., Chiu, W., Peraro, M. D., Di Maio, F., Ferrin, T. E., Grünewald, K., Gutmanas, A., Henderson, R., Hummer, G., Iwasaki, K., Johnson, G., Lawson, C. L., Meiler, J., Marti-Renom, M. A., Montelione, G. T., Nilges, M., Nussinov, R., Patwardhan, A., Rappsilber, J., Read, R. J., Saibil, H., Schröder, G. F., Schwieters, C. D., Seidel, C. A. M., Svergun, D., Topf, M., Ulrich, E. L., Velankar, S., and Westbrook, J. D. (2015) Outcome of the First wwPDB Hybrid/Integrative Methods Task Force Workshop. *Structure* 23, 1156–1167

70. Zeng-Elmore, X., Gao, X.-Z., Pellarin, R., Schneidman-Duhovny, D., Zhang, X.-J., Kozacka, K. A., Tang, Y., Sali, A., Chalkley, R. J., Cote, R. H., and Chu, F. (2014) Molecular architecture of photoreceptor phosphodiesterase elucidated by chemical cross-linking and integrative modeling. *J. Mol. Biol.* 426, 3713–3728

71. McDonald, J. H., University of Delaware (2009) *Handbook of Biological Statistics*

72. Alber, F., Dokudovskaya, S., Veenhoff, L. M., Zhang, W., Kipper, J., Devos, D., Suprapto, A., Karni-Schmidt, O., Williams, R., Chait, B. T., Sali, A., and Rout, M. P. (2007) The molecular architecture of the nuclear pore complex. *Nature* 450, 695–701

Figure 1: The insertion domain of the Ino80 ATPase (Ino80INS) is a direct interactor and activator of Rvb1/Rvb2



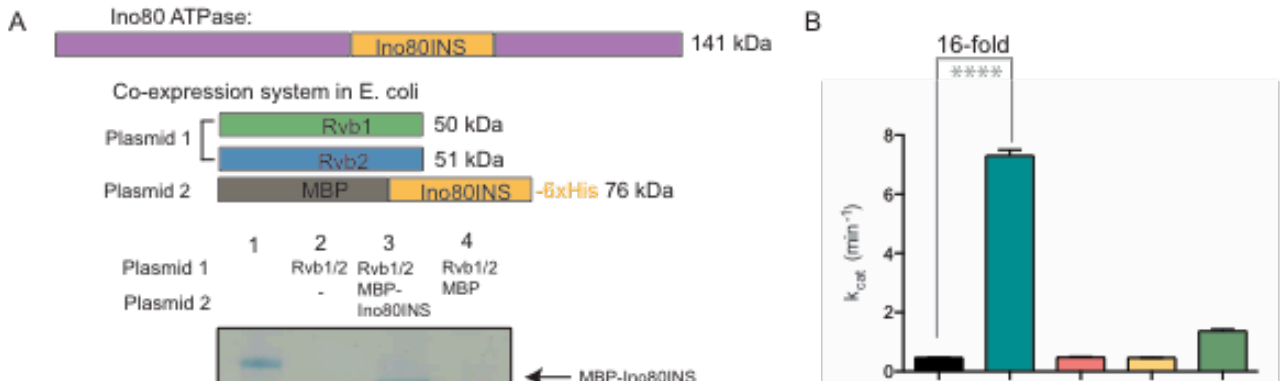
(A) Top, Sequence architecture of the Ino80 remodeling ATPase from yeast, with the Ino80INS in yellow. Below, co-expression system for Rvb1/Rvb2/Ino80INS in *E. coli* is shown. Plasmid 1 contains both Rvb1 and Rvb2. Rvb2 Plasmid 2 contains the Ino80 insertion (Ino80INS, residues 1022-1294) expressed as an MBP-fusion and with a C-terminal histidine tag. Bottom, SDS-PAGE gel showing results of pull down with co-expressed MBP-Ino80INS and Rvb1/Rvb2 (Lane 3). Lane 4 shows the result of the pull down with Rvb1/Rvb2 and MBP alone. Purification of co-expressed Rvb1 and Rvb2 is shown in Lane 2 for comparison.

(B) Comparison of maximal rates of ATP hydrolysis by different Rvb1/Rvb2 and Rvb1/Rvb2/Ino80INS complexes. Stimulation of 5 μM Rvbs with 25 μM double stranded DNA and MBP-Δ131Δ are shown in yellow and green, respectively. Bars represent the mean ± SEM of 3 replicates. **** = $p < 0.0001$.

(C) XL-MS data for the Rvb1/Rvb2/Ino80INS complex. Rvb1 (green) and Rvb2 (blue) are shown with different shades representing DI, DII, and DIII (see legend, right). The sequence positions of unique crosslinks in our dataset are shown as different colored lines according to the venn diagram on right. Crosslinks found only in our dataset are shown in gray lines, while crosslinks only found in the native INO80 complex (Tosi et al., 2013) are shown in pink, and crosslinks found in both datasets are shown in green. The confidence of the MS identification (SVM score) is represented by the thickness and transparency of the lines (Supplemental Experimental Procedures).

See also Figure S1 and Table S1.

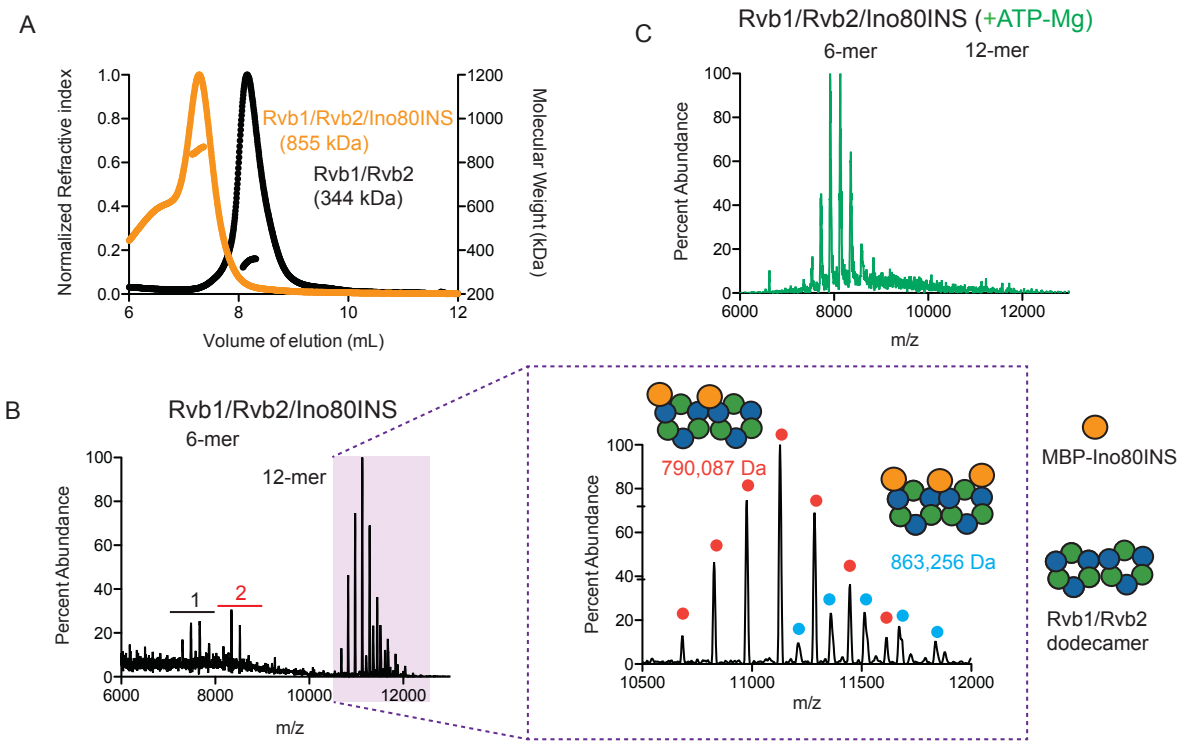
Figure S1 (related to Figure 1): XL-MS of Rvb1/Rvb2 and Rvb1/Rvb2/Ino80INS



(A) XL-MS data for Rvb1/Rvb2 alone. Rvb1 (green) and Rvb2 (blue) are shown with different shades representing DI, DII, and DIII (see legend, right). The confidence of the MS identification (SVM score) is represented by the thickness and transparency of the lines (Supplemental Experimental Procedures).

(B) Decoy analysis of cross-links. The distribution of cross-linked residue pairs in which both cross-linked amino acids matched the target protein database are plotted in blue with respect to the SVM confidence score used for classification. Crosslinked residue pairs in which either residue matched the decoy (scrambled) sequence database are in red. An SVM score of 0.0 (marked by green dotted line) was taken as the acceptance criteria resulting in a residue-pair level false discovery rate of 1.7%.

Figure 2: The Ino80 insertion promotes dodecamerization of Rvb1/Rvb2



(A) SEC-MALS of Rvb1/Rvb2/Ino80INS complexes. Normalized refractive index (left y-axis) and molecular weight of the major peak (right y-axis) are plotted against the volume of elution. For each experiment, either 10 μM of Rvb1/Rvb2/Ino80INS (orange) or 30 μM Rvb1/Rvb2 (black) was injected. The molecular weights shown in corresponding colors are an average of the calculated molecular weight values across the main peak.

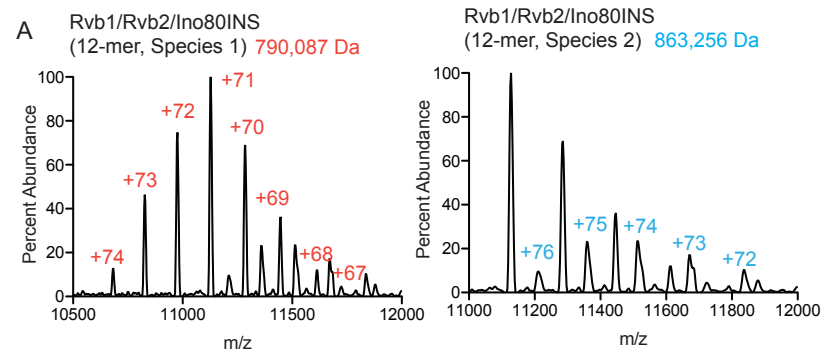
(B) Left, raw spectrum of Rvb1/Rvb2/Ino80 complexes using native-MS. Charge envelopes for the hexameric species are shown in black (1) and red (2). Right, an expanded view of the charge envelopes representing the dodecameric species. Blue and red spheres indicate the charge states of two separate charge envelopes, with the deconvoluted masses in corresponding colors. Cartoons representing the MBP-Ino80INS (orange spheres) and

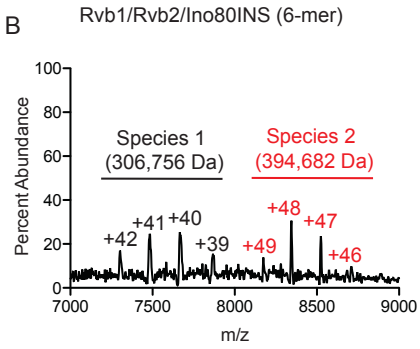
Rvb1/Rvb2 (blue and green spheres) illustrate the predicted stoichiometries represented by the two charge envelopes.

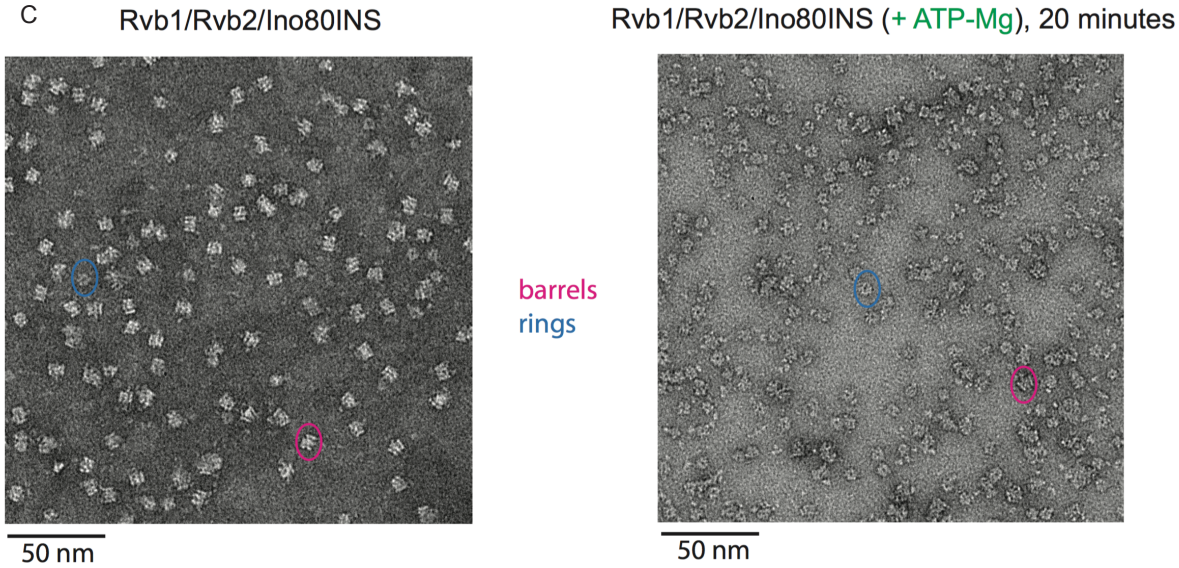
(C) Raw spectrum of 1 μM Rvb1/Rvb2/Ino80INS complexes with 20 μM ATP-Mg added, by native-MS.

See also Figure S2 and Table S2.

Figure S2 (related to Figure 2, Table S2): Native-MS and negative stain EM of the Rvb1/Rvb2/Ino80INS complex







(A) Full raw spectrum of Rvb1/Rvb2/Ino80INS by native-MS.

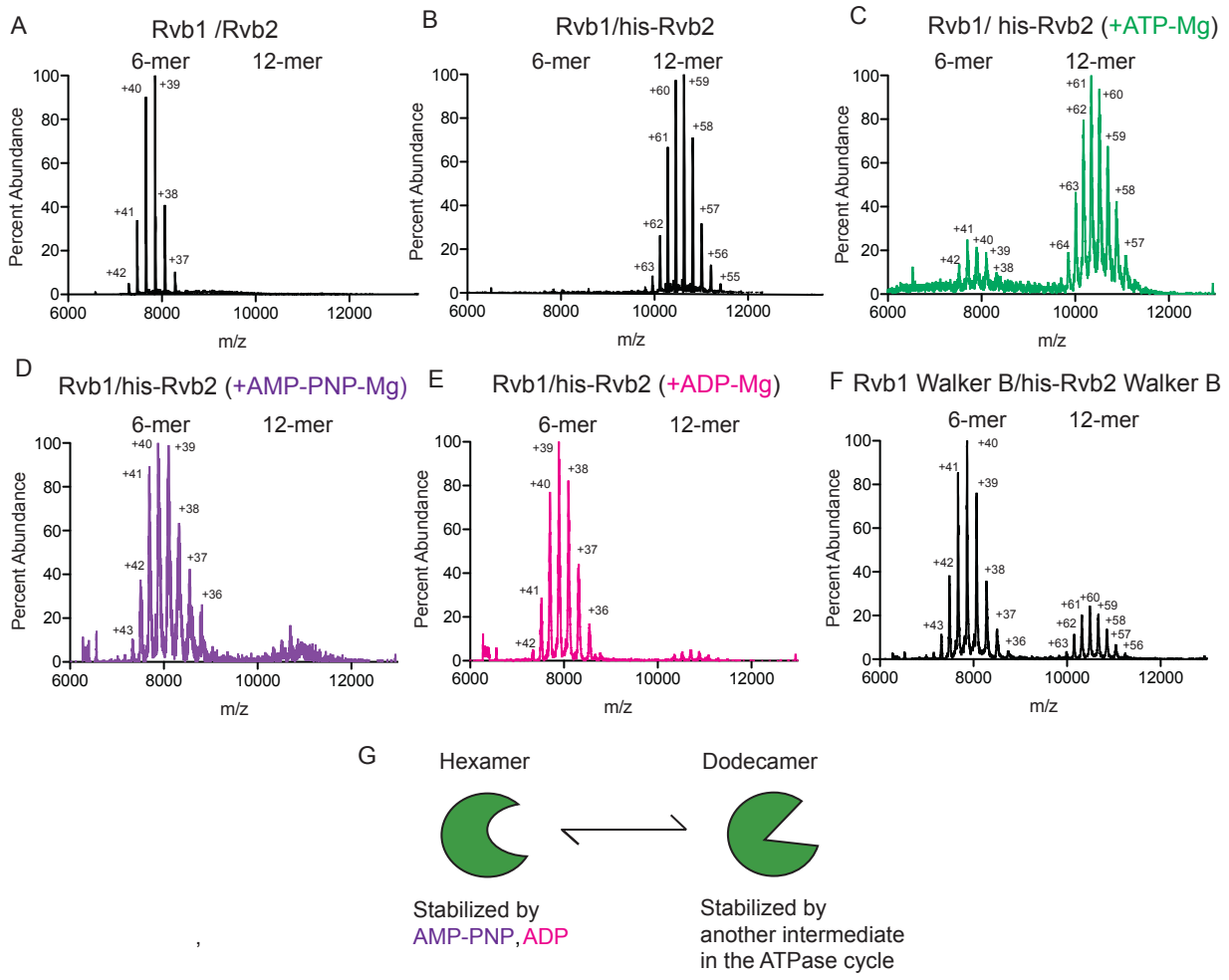
(B) Charge states for Species 1 (red dots in Figure 4B, right) of the Rvb1/Rvb2/Ino80INS sample that contains a Rvb1/Rvb2 dodecamer.

(C) Charge states for Species 2 (blue dots in Figure 4B, right) of the Rvb1/Rvb2/Ino80INS sample that contains a Rvb1/Rvb2 dodecamer. Deconvoluted masses are shown in the corresponding colors.

(D) Charge states for two species of the Rvb1/Rvb2/Ino80INS sample that contains a Rvb1/Rvb2 hexamer. Species 1 (black) has a mass that is most consistent with a Rvb1/Rvb2 heterohexamer while Species 2 (red) has a mass that is most consistent with a Rvb1/Rvb2 heterohexamer bound to an Ino80INS.

(E) Left, negative stain image of 2 mM Rvb1/Rvb2/Ino80INS complexes. Right, the same sample of Rvb1/Rvb2/Ino80INS with 20 mM ATP-Mg added, and after a 20-minute incubation at room temperature. For each image, barrel-like structures are highlighted by pink circles and ring-like structures are highlighted by blue circles.

Figure 3: Nucleotide state drives the switch between hexameric and dodecameric forms of the Rvb1/Rvb2 complex



(A) Raw spectrum of 1 µM of untagged Rvb1/Rvb2 using native-MS.

(B) Raw spectrum of 1 µM Rvb1/his-Rvb2 using native-MS.

(C) Raw spectrum of 1 µM Rvb1/Rvb2-his tag after addition of 20 µM ATP-Mg.

(D) Raw spectrum of 1 µM Rvb1/Rvb2-his tag after addition of 20 µM AMP-PNP-Mg, using native-MS.

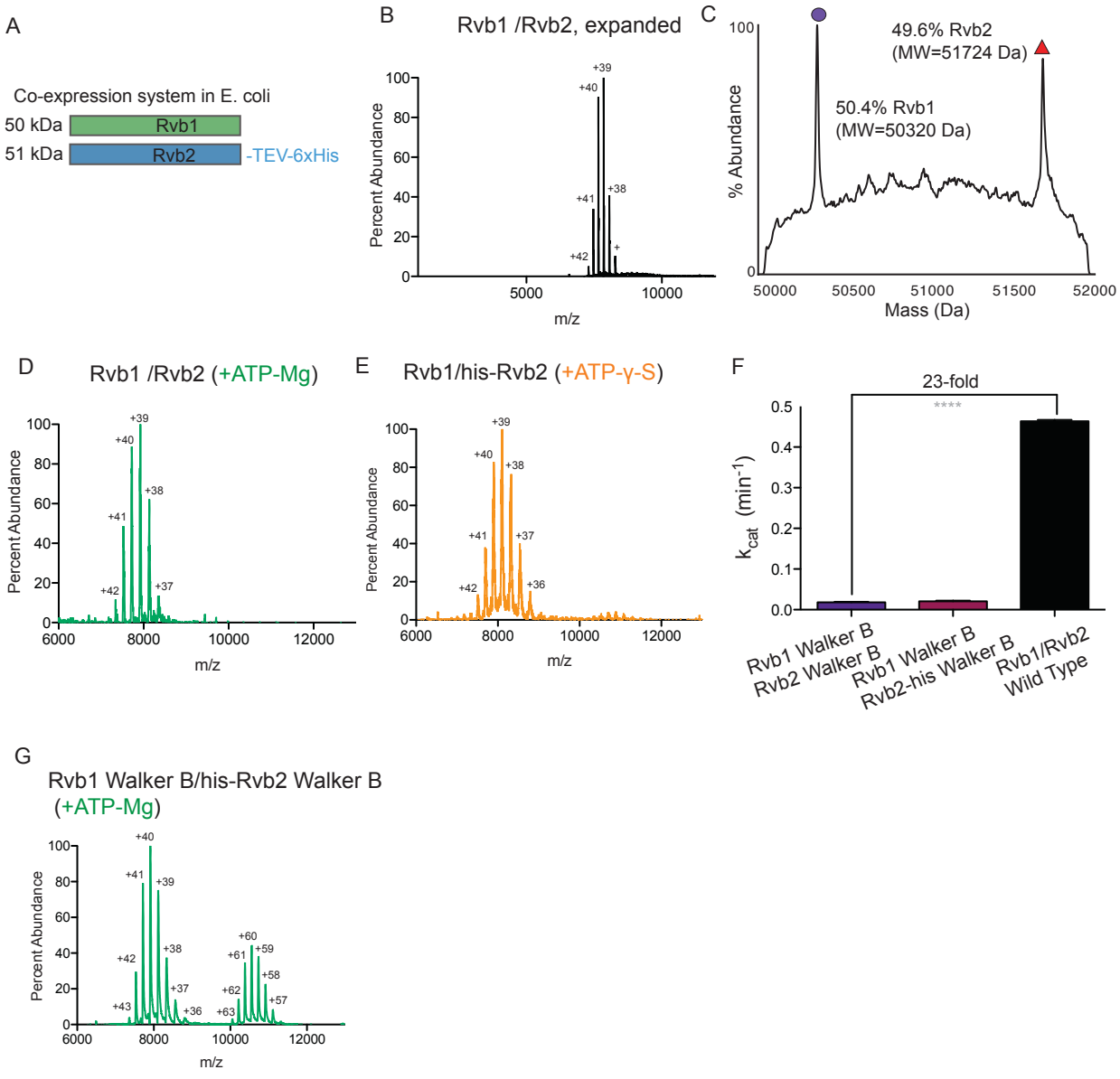
(E) Raw spectrum of 1 µM Rvb1/Rvb2-his tag after addition of 20 µM ADP-Mg, using native-MS.

(F) Raw spectrum of 1 µM his-tagged Walker B mutants of Rvb1 (D311N) and Rvb2 (D296N) using native-MS.

(G) Model for how hexameric and dodecameric states of Rvbs are regulated by nucleotide state.

See also Figure S3, Table S3, and Table S4.

Figure S3 (related to Figure 3, Table S3, and Table S4): Native-MS of Rvb1/Rvb2 complexes



(A) Co-expression system used for experiments with Rvb1/Rvb2 alone. The his-tag on Rvb2 is removed by tev cleavage unless otherwise noted. See also experimental procedures.

(B) Expanded view of Rvb1/Rvb2 spectrum.

(C) Deconvoluted ESI-mass spectrum of intact Rvb1/Rvb2 complexes. 0.25 μM of Rvb1/Rvb2 was denatured by addition of 50% acetonitrile and 10% acetic acid. The percent abundance of each peak was measured by taking the area under each peak.

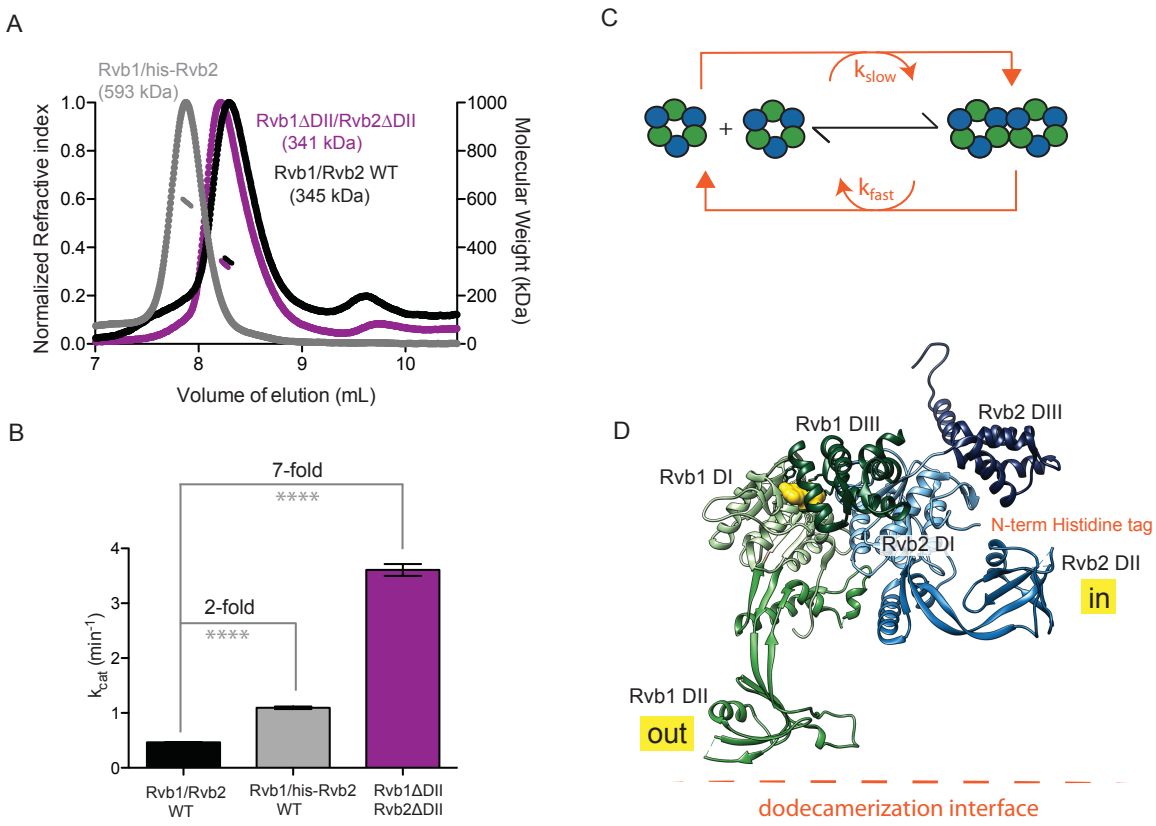
(D) Raw spectrum of 1 μM Rvb1/Rvb2 with 20 μM ATP-Mg added.

(E) Maximal rate contants (k_{cat}) for ATP hydrolysis activity by Walker B mutants of Rvb1/Rvb2. k_{cat} values for Rvb1 D311N/Rvb2 D296N (purple bar) and Rvb1 D311N/his-Rvb2 D296N (red bar) are plotted next to k_{cat} value of wild-type Rvb1/Rvb2. Experiments were performed using 10 mM Rvb monomer and saturating ATP. Bars represent the mean ± SEM of 3 replicates. **** = $p < 0.0001$.

(F) Raw spectrum of 1 μM Rvb1 Walker B/his-Rvb2 Walker B with 20 μM ATP-Mg added.

(G) Raw spectrum of 1 μM Rvb1/his-Rvb2 with 20 μM ATP-γ-S added.

Figure 4: The ATPase activity of the Rvb1/Rvb2 complex is enhanced by dodecamerization



(A) SEC-MALS of different Rvb1/Rvb2 complexes. Normalized refractive index (left y-axis) and molecular weight of the major peak (right y-axis) are plotted against the volume of elution. For each experiment, 10 μM of either Rvb1/Rvb2 (black), Rvb1/Rvb2-his (gray), or Rvb1ΔDII, Rvb2ΔDII (purple) was injected. The molecular weights shown in corresponding colors are an average of the calculated molecular weight values across the main peak.

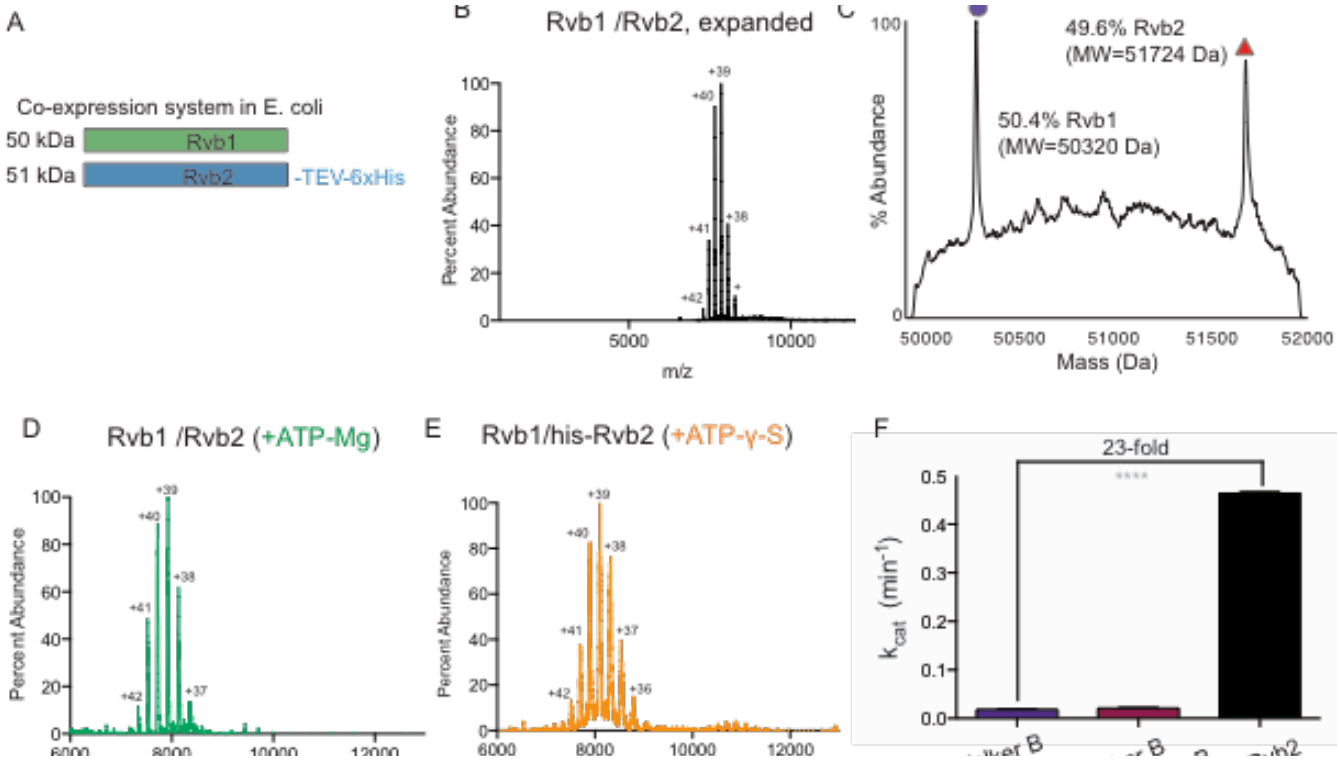
(B) Maximal rate constants (k_{cat}) for ATP hydrolysis activity by different Rvb1/Rvb2 complexes. Experiments were performed using 10 mM Rvbs and saturating ATP. For all experiments, a sample containing no enzyme was always included as a negative control, which has negligible activity. Bars represent the mean ± SEM of 3 replicates. **** = $p < 0.0001$.

(C) Model for how Rvb1/Rvb2 oligomeric state is coupled to ATPase cycle. Rvb1/Rvb2 complexes exist at equilibrium between hexamer and dodecamer (black arrows). In the presence of nucleotide (orange arrows), Rvb1/Rvb2 complexes are constantly interchanging between the two oligomeric states, creating a steady state distribution of the hexamer and dodecamer, where hexamers hydrolyze ATP slower (k_{slow}) than dodecamers (k_{fast}).

(D) Top, domain architecture of Rvb1 and Rvb2 from *S. cerevisiae*. Domains of Rvb1 are colored as different shades of green and domains of Rvb2 are colored as different shades of blue. Bottom, model of Rvb1 (bound to ATP) and Rvb2 (apo) adapted from the crystal structure of Rvb1/Rvb2 complex from *Chaetomium theramophilum* (Lakomek et al, 2015, PDB: 4WVY). Domains I (light green in Rvb1 and light blue in Rvb2) and III (dark green in Rvb1 and dark blue in Rvb2) contain the Walker A and Walker B residues for ATP binding and hydrolysis, respectively. Domain II (medium green in Rvb1 and medium blue in Rvb2) folds as a separate module. The location of the N-terminal histidine tag, which does not appear in the density, is shown in orange and a space-filled model of ATP is in yellow. The DII domains of Rvb1 and Rvb2 are in two distinct conformations, “in” and “out” (yellow boxes). The DII “out” state appears to form the dodecameric interface in this model (orange dashed line).

See also Figure S4 and Table S5.

Figure S4 (related to Figure 4 and Table S5): SEC-MALS of Rvb1/Rvb2 complexes

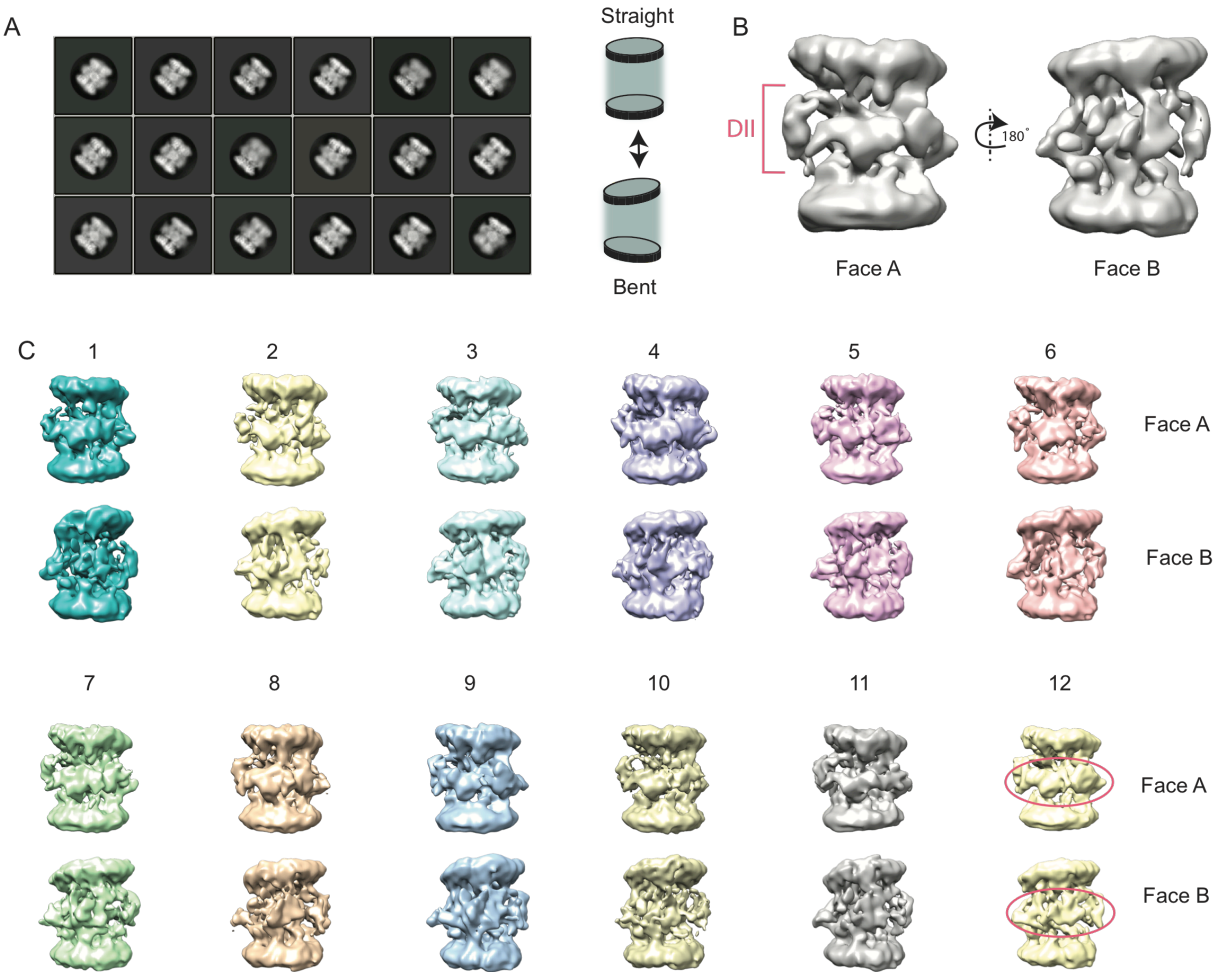


(A) SEC-MALS of Rvb1△DII, Rvb2△DII complexes at 10 μM (purple) and 30 μM (green). The molecular weights shown are an average of the calculated molecular weight values across the main peak.

(B) SEC-MALS of wild-type Rvb1/Rvb2 at 80 μM, along with calculated molecular weight.

(C) Native-MS raw spectrum of 1 μM Rvb1△DII/Rvb2△DII.

Figure 5: Cryo-EM of the Rvb1/Rvb2/Ino80INS complex reveals conformational heterogeneity



(A) Left: Side views of reference-free 2D class averages of Rvb1/Rvb2/Ino80INS particles extracted from micrographs represented in Figure S5. Right: Scheme displaying the types of conformational heterogeneity in the Rvb1/Rvb2/Ino80INS dodecamer, as observed in both 2D class averages and the 3D classes.

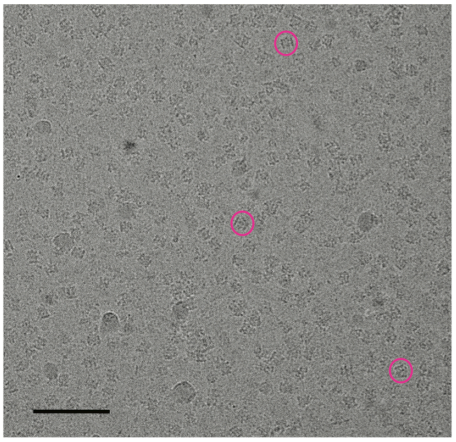
(B) Rvb1/Rvb2/Ino80INS consensus model generated by refinement of all 366,000 particles used for 2D classification in (A).

(C) Left: Twelve 3D classes representing all of the particles used for the consensus structure in (B). The top and bottom rows of each class represent two opposite faces of the dodecamer. The particles are approximately evenly distributed between all 12 classes, with approximately 30,000 per class.

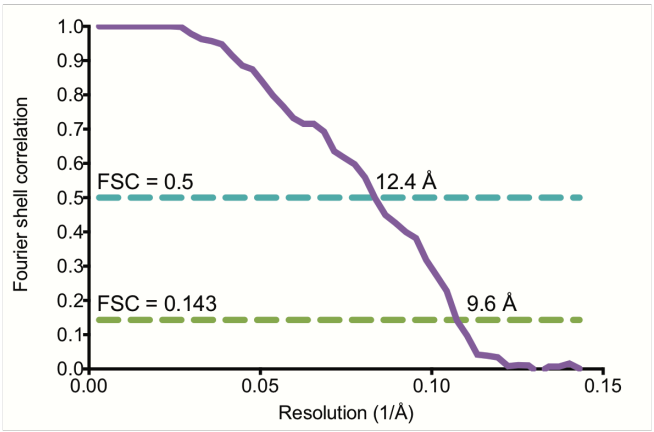
See also Figure S5.

Figure S5 (related to Figure 5): Cryo-EM recontruction of the Rvb1/Rvb2/Ino80INS complex

A



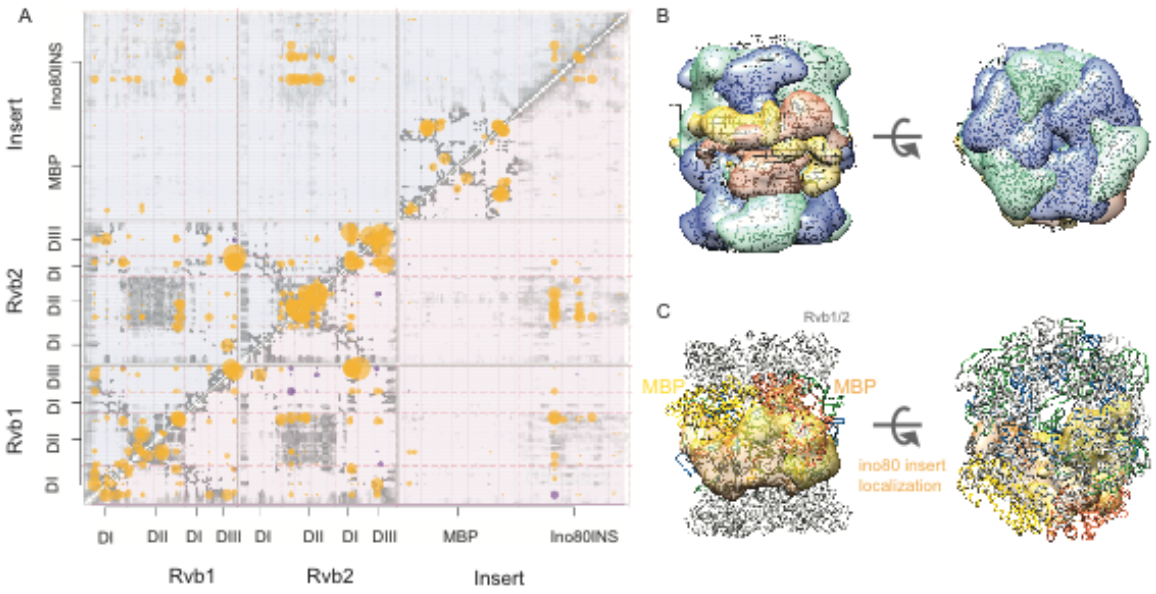
B



(A) Representative cryo-EM micrograph showing mono-dispersed and relatively homogeneous Rvb1/Rvb2Ino80INS complex particles. Scale bar is equal to 50 nm.

(B) Fourier shell correlation (FSC) curve generated after refining two separate half-maps using Relion 3D auto-refine.

Figure 6: Ino80INS binds to Rvb1/Rvb2 dodecamer asymmetrically



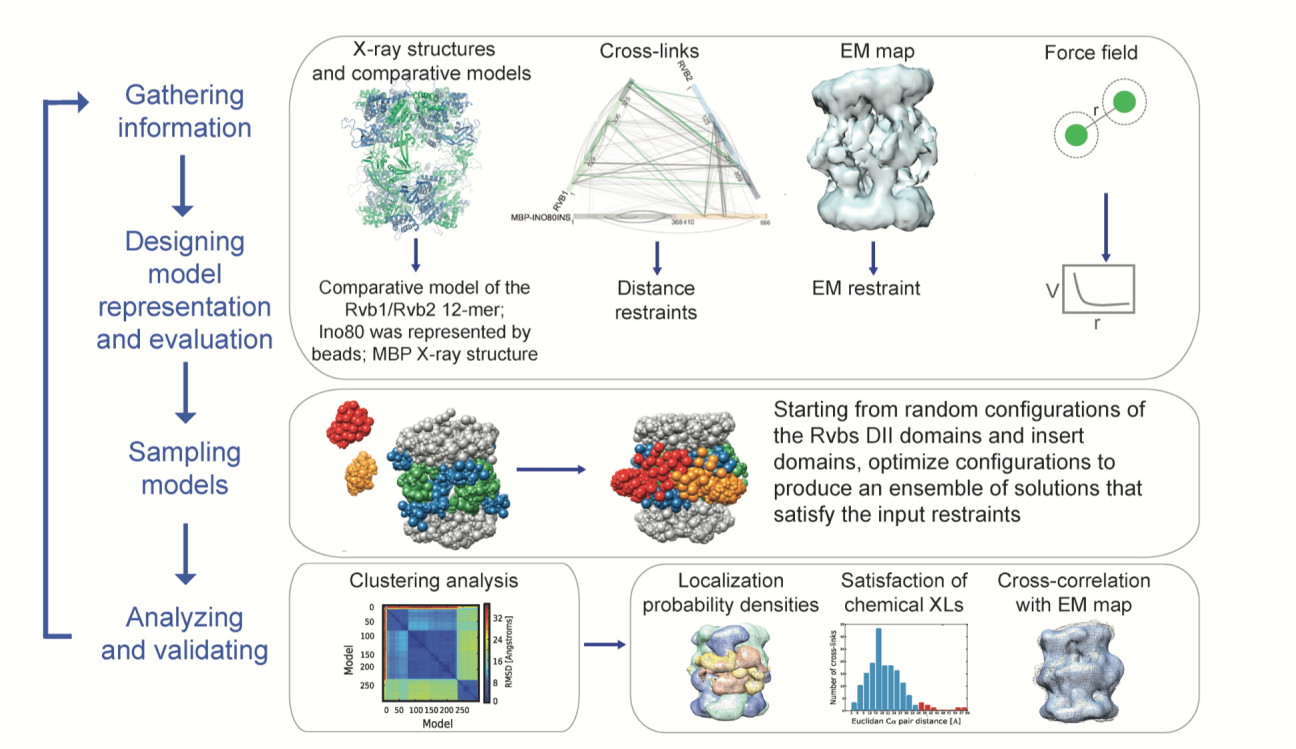
(A) Characterization of Clusters 1 (top half, shaded light blue) and 2 (bottom half, shaded light red). The gray bins indicate pairs of proximal beads representing the model, with the intensity of grey proportional to the fraction of models in the cluster whose distance is closer than the cutoff of 12 Å. The yellow circles correspond to the cross-links satisfied by at least one model in a cluster, with the size of the circle proportional to the spectral count. The purple circles indicate the violated cross-links, which are not satisfied by any model in the cluster.

(B) Side and top views of the localization density maps for Rvb1 (green), Rvb2 (blue), and MBP-Ino80INS (orange and yellow) calculated for Cluster 1 (thresholded at 0.15). The grey mesh shows the consensus EM density map.

(C) Top and side views of a representative model from Cluster 1. The distributions of the two copies of Ino80INS are shown by their localization densities (orange and yellow, respectively).

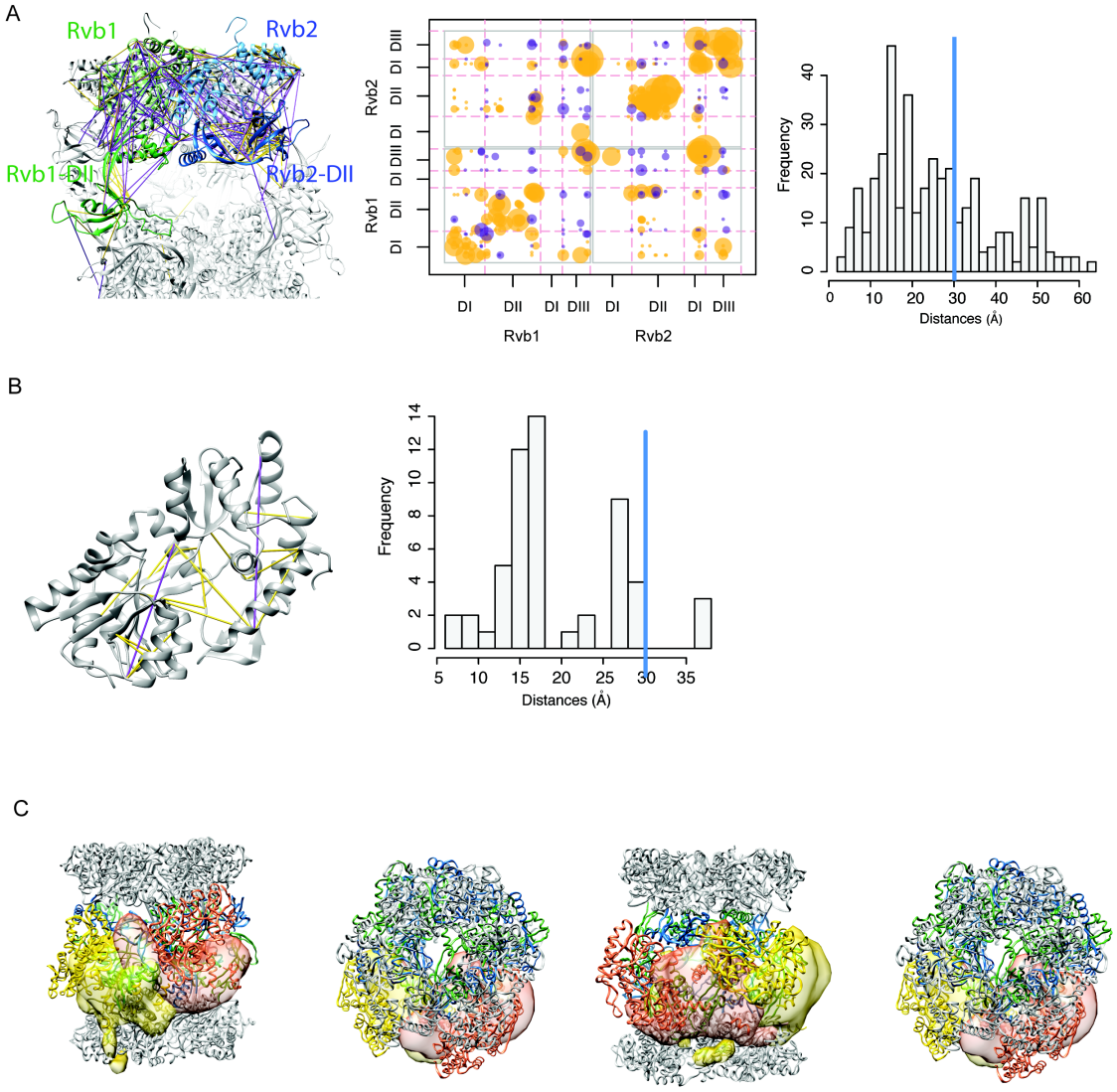
Also see Figure S6, Figure S7, Table S6.

Figure S6 (related to Figure 6, Figure S7, Table S6): The 4-stage scheme for integrative structure modeling of the Rvb1/Rvb2/Ino80INS complex



Our integrative approach proceeds through four stages: (1) gathering of data, (2) representation of subunits and translation of the data into spatial restraints, (3) configurational sampling to produce an ensemble of models that satisfies the restraints, and (4) analysis of the ensemble.

Figure S7 (related to Figure 6, Figure S6, Table S6): Analysis of integrative structure modeling of Rvb1/Rvb2/Ino80INS complex



(A) Analysis of the Rvb-Rvb crosslinked residue pairs in the Rvb1/Rvb2/Ino80INS dataset. Left: 132 satisfied XLs (58%; yellow) and 94 violated XL's (42%; purple) are mapped onto a model of the Rvb1/Rvb2 dodecamer, using a 30 Å cutoff. For clarity, only XLs for the first subunit of Rvb1 (green) and Rvb2 (blue) are shown. Middle: the yellow and purple circles correspond to the XLs that are satisfied or violated in the X-ray structure, respectively. The size of each circle is proportional to the spectral count of the cross-link. Right: histogram of the

distances between crosslinked residues in the X-ray structure. The vertical blue line represents the 30 Å cutoff used to define satisfied and violated XLs.

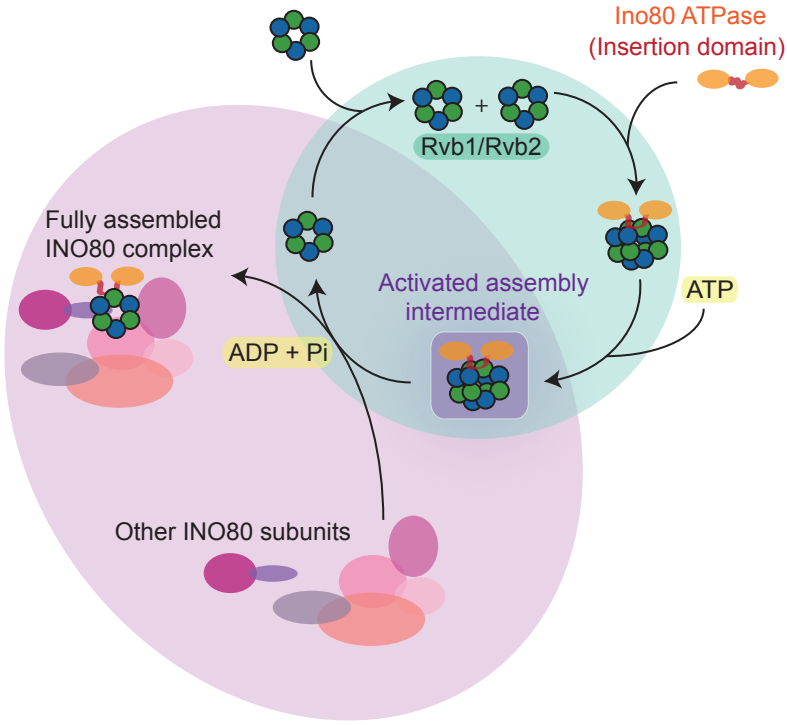
(B) Analysis of the cross-linked residues in MBP

Left: Mapping MBP-MBP crosslinks onto the X-ray structure of shows that two of the MBP XLs are violated (in purple). This observation indicates that these cross-links are inter-molecular.

Right: Histogram of the distances between cross-linked residues in the X-ray structure. The vertical blue line represents the 30 Å cutoff.

(C) Structure of the Rvb1/Rvb2/Ino80INS complex Cluster 2 (left) and Cluster 3 (right). The Rvb1/Rvb2 dodecamer is represented by the centroid structure of the cluster. The top and bottom rings are colored grey, while the Rvb1 and Rvb2 domains are colored green and blue, respectively. The Ino80INS domains are shown as localization densities (orange and yellow). The MBPs are shown in ribbon representation (orange and yellow).

Figure 7: Model for activation of Rvb1/Rvb2 by client proteins



Binding of client proteins to DII domains of Rvb1/Rvb2 promotes the dodecameric state, perhaps by stabilizing the flexible DII regions that comprise the dodecamerization interface. The two clients bind asymmetrically to the dodecamer, creating an activated intermediate that is highly competent for ATP hydrolysis. Addition of ATP causes the dodecamer to fall apart, potentially allowing the two client proteins to be assembled into the full INO80 complex. Meanwhile, Rvb1/Rvb2 hexamers are recycled and poised to repeat the reaction.

Table S1 (related to Figure 1 and Figure S1): Summary of XL-MS data for Rvb1/Rvb2 and Rvb1/Rvb2/Ino80INS

Rvb1/Rvb2 crosslinks observed				
	Rvb1	Rvb2	Ino80	Total [a]
Rvb1	111			
Rvb2	156	106		
Ino80	na	na	na	
Total	267	106	0	373 (2300)
Rvb1/Rvb2/Ino80INS crosslinks observed				
	Rvb1	Rvb2	Ino80	Total [a]
Rvb1	69			
Rvb2	85	73		
Ino80	42	38	42	
Total	196	111	42	349 (1420)

[a] Numbers of unique cross-linked residue pairs. Numbers in parentheses are spectral counts.

Table S2 (related to Figure 2, Figure S2): Analysis of the Rvb1/Rvb2/Ino80INS complex by native mass spectrometry

Rvb1/Rvb2/Ino80INS Hexameric Species							
Hexamer Species 1				Hexamer Species 2			
Mass observed (Da)	Standard devation	Predicted stoichiometry	% Error	Mass observed (Da)	Standard deviation	Predicted stoichiometry	% Error
306,756	73	3 Rvb1 3 Rvb2	0.7%	394,682	8232	3 Rvb1 3 Rvb2 1 MBP-Ino80INS	3.0%
Rvb1/Rvb2/Ino80INS Dodecameric Spcies							
Dodecameric Species 1				Dodecameric Species 2			
Mass observed (Da)	Standard devation	Predicted stoichiometry	% Error	Mass observed (Da)	Standard deviation	Predicted stoichiometry	% Error
790,087	387	6 Rvb1 6 Rvb2 2 MBP-Ino80INS	3.0%	863,256	917	6 Rvb1 6 Rvb2 3 MBP-Ino80INS	2.4%
Rvb1/Rvb2/Ino80INS + ATP-Mg							
Hexameric Species				Dodecameric Species			
Mass observed (Da)	Standard devation	Predicted stoichiometry	% Error	Mass observed (Da)	Standard deviation	Predicted stoichiometry	% Error
308813	47	3 Rvb1 3 Rvb2 5 nucleotides	0.9%	n.a.	n.a.	n.a.	n.a.

Table S3 (related to Figure 3, Figure S3): Predicted molecular weights for different stoichiometries of Rvb1/Rvb2 hexamers

Rvb1:Rvb2	Predicted MW (Da)	% Error (based on observed mass)
3:3	306930	0.19%
4:2	305526	0.46%
2:4	308334	0.46%
1:5	309738	0.91%
5:1	304122	0.91%

Table S4A (related to Figure 4 and Table S3): Molecular mass distribution of Rvb1/Rvb2 hexamers by native mass spectrometry

Sample	Hexamer				
	% of total area	Expected mass (Da)	Observed mass (Da)	Standard deviation (Da) [a]	# of NT bound [b]
Rvb1/Rvb2	100	306,930	306,352	40	0
Rvb1/Rvb2 +ATP-Mg	100	306,930	308,693	250	4-5
Rvb1/his-Rvb2	0	313,692	n.a.	n.a.	n.a.
Rvb1/his-Rvb2 + ATP-Mg	14	313,692	315,729	71	4-5
Rvb1/his-Rvb2 + ADP-Mg	100	313,692	315,745	210	5-6
Rvb1/his-Rvb2 + AMP-PNP-Mg	100	313,692	315,452	175	4-5
Rvb1/his-Rvb2 + ATP-γ-S-Mg	100	313,692	315,796	220	2-3
Rvb1 Walker B/ his-Rvb2 Walker B	83	313,686	314,564	83	1-2
Rvb1 Walker B/ his-Rvb2 Walker B+ ATP-Mg	80	313,686	316,884	494	6-7

[a] Standard deviation based on n=2
[b] Number of nucleotides bound is estimated by dividing the difference between observed and expected mass by the mass of ATP Mg (531 Da), ADP-Mg (451 Da), or AMP-PNP-Mg (528 Da).

Table S4B (related to Figure 4 and Table S3): Molecular mass distribution of Rvb1/Rvb2 dodecamers by native mass spectrometry

Sample	Dodecamer				
	% of total area	Expected mass (Da)	Observed mass (Da)	Standard deviation (Da) [a]	# of NT bound[b]
Rvb1/Rvb2	0	n.a.	n.a.	n.a.	n.a.
Rvb1/Rvb2 +ATP-Mg	0	n.a.	n.a.	n.a.	n.a.
Rvb1/his-Rvb2	100	627,384	627,407	102	1-2
Rvb1/his-Rvb2 + ATP-Mg	86	627,384	631,079	424	8-9
Rvb1/his-Rvb2 + ADP-Mg	0	627,384	n.a.	n.a.	n.a.
Rvb1/his-Rvb2 + AMP-PNP-Mg	0	627,384	n.a.	n.a.	n.a.
Rvb1/his-Rvb2 + ATP-γ-S-Mg	0	627,384	n.a.	n.a.	n.a.
Rvb1 Walker B/ his-Rvb2 Walker B	17	627,372	629,314	129	5-6
Rvb1 Walker B/ his-Rvb2 Walker B+ ATP-Mg	20	627,372	633,685	494	13-14

[a] Standard deviation based on n=2
[b] Number of nucleotides bound is estimated by dividing the difference between observed and expected mass by the mass of ATP Mg (531 Da), ADP-Mg (451 Da), or AMP-PNP-Mg (528 Da).

Table S5 (related to Figure 4 and Figure S4): Molecular weight predictions from MALS experiments

Sample	Observed molecular weight (kDa)	Expected molecular weight for hexamer (kDa)[a]	Expected molecular weight for dodecamer (kDa)[a]
Rvb1/Rvb2	345 kDa	307 kDa	614 kDa
Rvb1/his-Rvb2	593 kDa	314 kDa	627 kDa
Rvb1ΔDII, Rvb2ΔDII	341 kDa	237 kDa	474 kDa

[a] Expected molecular weights were calculated for a 3:3 stoichiometry of Rvb1:Rvb2 for the hexamer and 6:6 stoichiometry for Rvb1:Rvb2 for the dodecamer.

Table S6 (related to Figure 6, S6, S7): Summary of integrative structure determination of the Rvb1/Rvb2/Ino80INS complex

Modeling programs	• Python Modeling Interface (PMI), version b47d68e • Integrative Modeling Platform (IMP, version 2.5) • MODELLER 9.13
Homology detection and structure prediction	• HHPred, PSIPRED, DISOPRED
Structured components	• Rvb1/Rvb2 dodecamer comparative models based on PDB 4WVY • MBP X-ray structure (PDB 4EFB) • Ino80 helices residues 1137-1144, 1148-1156, 1193-1201, 1127-1237 and 1251-1261.
Unstructured components	• Rvb1 residues: 1-51, 143-165, 211-238 • Rvb2 residues: 142-155, 205-229, 249-266, 444-354 • Ino80INS, expect helices defined as structured parts.
Stoichiometry	6:6:2(=Rvb1:Rvb2:MBP-Ino80INS)
Spatial restraints	Sequence connectivity, protein excluded volume, chemical cross-links and 3D electron microscopy *See Supplemental Experimental Procedures.*
Sampling method	Replica Exchange Gibbs sampling, based on the Metropolis Monte Carlo algorithm. 32 replicas were used in 20 independent runs, at the temperature range of 1.0 - 2.5.
Monte Carlo moves	- Random translation and rotation of rigid bodies (up to 1.5 Å and 0.03 radians, respectively). - Random translation of individual beads in the flexible segments (up to 2 Å).
Clustering analysis	3 clusters of 196, 180 and 124 structures. See figures 6 and S7.
Sampling exhaustiveness	p-value = 0.15
Precision of the clusters	• Cluster 1: 25 Å (all components); 15.9 Å (Rvbs) 39.3 Å (Ino80ins) • Cluster 2: 26.7 Å (all components); 14.8 Å (Rvbs) 40.2 Å (Ino80ins) • Cluster 3: 12.1 Å (all components); 6.2 Å (Rvbs) 14.3 Å (Ino80ins)
Chemical cross-links satisfied in the clusters	• Cluster 1: 97% • Cluster 2: 94% • Cluster 3: 91% • All clusters: 98%
3D EM map cross-correlation coefficients (ccc)	• Cluster 1: ccc = 0.94 • Cluster 2: ccc = 0.96 • Cluster 3: ccc = 0.87
Visualization and plotting	UCSF Chimera 1.10, XlinkAnalyzer, R

Supplemental Experimental Procedures

Protein expression and purification

Rvb1 and Rvb2 were co-expressed on a pet28a-DUET plasmid, a kind gift from the Hopfner laboratory at the Ludwig Maximilian University in Munich. The Ino80 insertion (Ino80INS, residues 1022-1294 of the yeast Ino80 ATPase) or the Swr1 insertion (Swr1INS, residues 1003-1237 of the yeast Swr1 ATPase) were cloned into the pMAL vector (NEB) containing a C-terminal His_{x8} tag, a gift from the Lim lab at UCSF. For each prep, the appropriate constructs were freshly transformed into Rosetta (DE3) *E. coli* (Novagen). 1 L cultures of transformed cells were grown to ~0.3-0.4 at 37°C in 2X LB media with 1X NaCl, then switched to 18°C and induced at OD_{600} = 0.6-0.7 with 0.3 mM IPTG for approxminately 18 hours. Cells were then harvested, resuspended in lysis buffer (25 mM HEPES pH 7.5, 300 mM KCl, 7.5 mM imidazole, 10% glycerol v/v, 2 µg/mL aprotinin, 1 µg/mL peptstatin A, 3 µg/mL leupeptin,1 mM PMSF) and lysed by high pressure with an Emulsiflex-C3 homogenizer (Avestin). Lysates were cleared by spinning at 30,000 x g for 30 minutes, and cleared lysates were incubated with TALON cobalt resin (Clontech) (1.5 mL slurry per 1 L culture) for 1 hour. Resin was batch-washed in 50 mL conicals for one minute before spinning down (1,000 x g) and removing the supernatant. For Rvb1/Rvb2 alone preps, resin was incubated with lysis buffer containing 100 µM ATP-Mg for 30 minutes, and then batch washed with lysis buffer for 30 minutes. For Rvb1/Rvb2/ID preps, the ATP wash step was skipped. Resin was transferred to a 20-mL disposable column (BioRad) and washed with 50 mLs of lysis buffer. Protein of interest was eluted in elution buffer (25 mM HEPES pH 8, 300 mM KCl, 500 mM imidazole). For Rvb1/Rvb2 untagged preps, TEV protease (homemade) was added to a final concentration of 0.075-0.15 mg/mL, and the mixture was dialyzed overnight into size exclusion (SEC) buffer, which contains 25 mM Hepes pH 7.5, 10% glycerol v/v, 300 mM KCl, and 2 mM dithiothreitol (DTT). For Rvb1/his-Rvb2 or Rvb1/Rvb2/ID preps, TEV protease was not added and the dialysis step was skipped. Proteins were then purified by size exclusion chromatography on a HiLoad Superdex 200 26/60 column (320 mL

bed volume; GE Lifesciences) in SEC buffer. Fractions were analyzed on an SDS-PAGE gel, pooled, concentrated, aliquoted, flash frozen in liquid nitrogen, and stored at -80°C.

Native-MS

Native mass spectrometry was carried out using the Exactive Plus EMR instrument (33) (Thermo Scientific, San Jose, CA) that was externally calibrated using a 5 mg/mL CsI solution prepared in water. Prior to analysis, the protein samples were buffer-exchanged into 150 mM ammonium acetate, pH 7.5 using MicroBiospin-6 columns (Bio-Rad, Hercules, CA) that had been pre-equilibrated in the same buffer. Protein samples were introduced into the mass spectrometer using offline Au/Pd-coated borosilicate emitters (NanoES Spray Capillaries, Medium, ES380, Thermo Scientific) at a flow rate of 10 – 40 nL/min. Spectra were acquired over the range *m/z* 500 – 20,000 in positive ion mode, were averaged, and then exported for deconvolution and subsequent generation of the zero-charge mass values using PeakSeeker (Lu et al., 2015). Samples were analyzed with the following experimental parameters: spray voltage (0.8 – 1.5 kV), injection flatapole = 5; inter flatapole lens = 5; bent flatapole = 5; transfer multipole = 6.1; C-trap entrance lens = 8.9, source DC offset (25 V), fragmentation energies (CE = 20 – 50 and CID = 50 – 100), injection times (200 usec), trapping gas pressure (7.5), resolution (17,500 arbitrary units), capillary temperature (250 °C), S-len RF levels (200 V), microscans (10), and AGC ($1e^6$). Nucleotide – Mg complexes were prepared fresh by incubating equimolar ratio of nucleotide and $MgCl_2$ for 30 minutes at 4 °C and then freshly diluting to the final working concentration. Protein (~ 0.5 to 1 μM) was incubated with excess nucleotide (20 μM) immediately prior to analysis.

XL-MS

Cross-linking analysis was generally performed as in “Procedure B” of Robinson et al., 2015 (16), with minor modifications. 80 μM of Rvb1/Rvb2, 46 μM of Rvb1/Rvb2/MBP-ino80 insertion,

and 44 μM of Rvb1/Rvb2/Swr1IN insertion were buffer exchanged into 25 mM Hepes, pH 7.5, 150 mM KCl and crosslinked with 2 mM Bis[sulfosuccinimidyl]suberate (BS3,Thermoscientific Pierce) on ice for 30 minutes. Samples were quenched with 50 mM Tris, pH 8.5. Crosslinked proteins were pelleted by centrifuging at 21,000 x g. The pellet was washed once with cold acetone, and then resuspended in 50 mM ammonium bicarbonate containing 8M Urea and 10 mM Tris(2-carboxyethyl)phosphine **(**TCEP). After incubation at 56 degrees for 20 minutes, the sample was cooled, iodoacetamide was added to a final concentration of 20 mM to alkylate –SH groups. Sample was then trypsinized overnight at 37°C. Following digestion, peptides were desalted using a Macrotrap C_{18} cartridge (Michrom Bioresources). Cross-linked products were then enriched by size-exclusion chromatography (Superdex Peptide, GE Healthcare Life Sciences). Fractions eluting between 0.9 and 1.4 ml were dried, resuspended in 0.1% formic acid and then analyzed with a Q-Exactive Plus mass spectrometer (Thermo Scientific) coupled with a nanoelectrospray ion source (Easy-Spray, Thermo) and NanoAcquity UPLC system (Waters). Enriched fractions were separated on a 15 cm x 75 μm ID PepMap C_{18} column (Thermo) using a 90-minute gradient from 3-27% solvent B (A: 0.1% formic acid in water, B: 0.1% formic acid in acetonitrile). Precursor MS scans were measured in the Orbitrap scanning from 350-1500 m/z (mass resolution: 70,000). The ten most intense triply charged or higher precursors were isolated in the quadrupole (isolation window: 4 m/z), dissociated by HCD (normalized collision energy: 25), and the product ion spectra were measured in the Orbitrap (mass resolution: 17,500). A dynamic exclusion window of 15 sec was applied and the automatic gain control targets were set to 3e6 (precursor scan) and 5e4 (product scan).

XL-MS data analysis

Peaklists were generated using Proteome Discoverer 1.4 (Thermo) and searched for cross-linked peptides with an in-house version of Protein Prospector 5.14.0 (17). 85 of the most intense peaks from each product ion spectrum were searched against a database containing 11

target protein sequences (yeast Rvb1, Rvb2, Ino80 and Swr1 insertion constructs as well as several minor contaminating proteins identified in a prior, unbiased search of the data) concatenated with 110 randomized decoy protein sequences using mass tolerances of 9 ppm (precursor scan) and 20 ppm (product scan). Carbamidomethylation of cysteine was considered as a fixed modification. N-terminal methionine loss with and without acetylation, peptide N-terminal glutamine conversion to pyroglutamate, oxidation of methionine, and “dead-end” modification of lysine and the protein N-terminus by Tris quenched BS3 were considered as variable modifications in addition to cross-linking by BS3. Up to 3 variable modifications per peptide were considered. Trypsin specificity with 3 missed cleavages was used to generate theoretical peptides.

Cross-link spectral matches (CSM) were initially discarded if the following Protein Prospector parameters fell outside the threshold values: peptide score below 20, peptide/protein, and score difference below 0. A linear support vector machine (SVM) model was constructed based on three parameters (“score”, “score difference”, and “peptide 2 rank”) to further classify CSMs between decoy and target classes (16, 17). 5,208 spectra were classified with an SVM confidence score above the acceptance threshold of 0 at a false discovery rate (FDR) of 0.6%. The FDR at the residue pair level was 1.7%. CSMs where either peptide was less than 4-amino acids long were discarded. When spectra could be interpreted by multiple cross-linked peptide-pairs, CSMs were considered to be unambiguous if the second best match had a confidence score 0.3 lower than the top match. The remaining spectral redundancy came exclusively from ambiguous site localizations, where there was sufficient evidence to identify both peptides, but the position of the modified amino acid remained ambiguous. In these cases, the CSM was flagged as ambiguous and all possible sites were reported.

Homology models of yeast Rvb1Rvb2 dodecamers were constructed using modeller 9.12 using the crystal structure of the Rvb1 Rvb2 complex from *C. thermophilum* (pdb:4WVY) as the template. The sequences are over 70% identical.

Cross-linking results and annotated HCD spectra may be viewed online using Protein Prospector's MS-Viewer program:

http://prospector2.ucsf.edu/prospector/cgi-bin/msform.cgi?form=msviewer

Search key: ***uvtzc60yzg***

Integrative structure modeling of the Rvb1/Rvb2/Ino80INS complex

We computed a structural model of the yeast Rvb1/Rvb2/Ino80INS complex using an integrative approach based on data from X-ray crystallography, cross-linking mass spectrometry, and cryo-EM. The complex consists of 2 copies of the Ino80INS-MBP construct and 2 Rvb1/Rvb2 rings, each ring consisting of 3 copies of Rvb1 and Rvb2. Our integrative structure determination proceeds through four stages (Figure S6A) (21, 22): (1) gathering of data, (2) representation of subunits and translation of the data into spatial restraints, (3) configurational sampling and scoring to produce an ensemble of structures that satisfies the restraints, and (4) analysis and validation of the ensemble structures. The modeling protocol (i.e., stages 2, 3, and 4) was scripted using the Python Modeling Interface (PMI), a library for modeling macromolecular complexes based on our open-source Integrative Modeling Platform (IMP) package, version 2.5 (http://integrativemodeling.org) (22). Further details of the integrative modeling procedures are provided in previous publications (16, 52-55). Files containing the input data, scripts, and output structures are available online (http://salilab.org/Rvbs).

1. Gathering of data: A comparative model of the yeast Rvb1/Rvb2 double-ring hexamer was computed based on the atomic X-ray structure of the double-heterohexameric Rvb1/Rvb2 rings from *Chaetomium thermophilum* (PDB access code 4WVY) (39), using MODELLER 9.15 (56). The atomic X-ray structure of the MBP domain was extracted from the PDB (4FED (42)). The structure of the yeast Ino80INS is not known and no putative homologs of known structure were detected, using HHpred (57) and ModWeb (58). Secondary structure and disordered regions of the Ino80INS domain were predicted by PSIPRED (59), JPRED (60), and DISOPRED (61).
We used the 350 unique BS3 cross-links of the Rvb1/Rvb2/Ino80INS complex (Figures 1 and S1, and Table S1) and the consensus cryo-EM model of the Rvb1/Rvb2/Ino80INS determined in this study (Figure 5).

2. Representation of subunits and translation of the data into spatial restraints: To maximize computational efficiency while avoiding using too coarse a representation, we represented the Rvb1/Rvb2/Ino80INS complex in a multi-scale fashion. In particular, the domains of the Rvb1/Rvb2/Ino80INS complex subunits were coarse-grained using beads of varying sizes representing either a rigid body or a flexible string, based on the available crystallographic structures and comparative models, as follows. First, the comparative models and crystallographic structures of the Rvb1s, Rvb2s and MBPs components were coarse-grained into three representations at different resolutions. Two of the representations corresponded to beads obtained at different levels of coarse-graining of the atomic structure. At the first scale (i.e., fine scale), each bead corresponded to an individual residue, and was centered at the position of the Cα atom. At the second scale (i.e., coarse scale), each bead represented 10-residue segments and was positioned at the center of mass of all atoms of the corresponding segment. At the third scale, the system was represented by a 3D density map corresponding to the Gaussian mixture model (GMM) (62). The atomic structures were converted into a GMM by first sampling points from the density, and then fitting the sample

using the expectation-maximization algorithm implemented in scikit-learn (63). For each component, we set the number of Gaussians to be approximately the number of residues in the component divided by 50. We excluded loop regions Rvb1$^{1-51}$, Rvb1$^{143-165}$, Rvb1$^{211-238}$, Rvb2$^{444-454}$, Rvb2$^{142-155}$, Rvb2$^{205-229}$, and Rvb2$^{249-266}$, which were represented as flexible strings of beads, as described below. Second, the predicted helical regions of the Ino80INS components (Ino80$^{1137-1144}$, Ino80$^{1148-1156}$, Ino80$^{1193-1201}$, Ino80$^{1227-1237}$, and Ino80$^{1251-1261}$) were modeled as ideal helices and represented as described for crystallographic structures and comparative models; for each helix, we used two Gaussians for the GMM representation. Finally, the remaining regions without a crystallographic structure, comparative model, or predicted secondary structure were represented by a flexible string of beads corresponding to up to 10 residues each. In addition, each of these beads was represented by a spherical Gaussian (the centers of a bead and its Gaussian were identical). The radius of the bead and the variance of the Gaussians were set to describe the average molecular volume and the molecular electron density of the polypeptide segments, respectively.

Rigid body definition: We defined the Rvb1 domains as follows: Rvb1$^{1-128}$, Rvb1$^{129-305}$, and Rvb1$^{306-463}$ correspond to domains DI, DII, and DIII, respectively. Similarly, we defined the Rvb2 domains as follow: Rvb2$^{1-124}$, Rvb2$^{125-290}$, and Rvb2$^{291-471}$ correspond to domains DI, DII, and DIII, respectively. The DI/DIII domain of Rvb1/2 copies 1-3 were defined as a single rigid body (top ring; Rvb1.1$^{1-128}$-Rvb1.1$^{306-463}$- Rvb2.1$^{1-124}$-Rvb2.1$^{125-290}$- Rvb1.2$^{1-128}$-Rvb1.2$^{306-463}$- Rvb2.2$^{1-124}$-Rvb2.2$^{125-290}$ -Rvb1.3$^{1-128}$-Rvb1.3$^{306-463}$- Rvb2.3$^{1-124}$-Rvb2.3$^{125-290}$). Similarly, The DI/DIII domain of Rvb1/2 copies 4 through 6 were defined as a single rigid body (bottom ring; Rvb1.4$^{1-128}$-Rvb1.4$^{306-463}$- Rvb2.4$^{1-124}$-Rvb2.4$^{125-290}$- Rvb1.5$^{1-128}$-Rvb1.5$^{306-463}$- Rvb2.5$^{1-124}$-Rvb2.5$^{125-290}$ -Rvb1.6$^{1-128}$-Rvb1.6$^{306-463}$- Rvb2.6$^{1-124}$-Rvb2.6$^{125-290}$). Each of the 12 Rvb1/2 DII domains was defined as a single independent rigid body. Consequently, the Rvb1/Rvb2 complex was described by 14 rigid bodies. Each copy of MBP as well of each of the ideal helices of Ino80INS

domains were represented as a single rigid body. In each rigid body, the beads have their relative distances constrained during configurational sampling, whereas in a flexible string the beads are restrained by the sequence connectivity, as described in our previous studies (43, 54, 64).

Scoring function: With this representation in hand, we next encoded the spatial restraints into a Bayesian scoring function (43, 65) based on the information gathered in Stage 1, as follows.

First, the excluded volume restraints were applied to each bead in 10-residue (or the closest) bead representations, using the statistical relationship between the volume and the number of residues that it covered (21, 54, 66).

Second, we applied the sequence connectivity restraint, using a harmonic upper bound on the distance between consecutive beads in a subunit, with a threshold distance equal to four times the sum of the radii of the two connected beads. The bead radius was calculated from the excluded volume of the corresponding bead, assuming standard protein density (21, 43, 54).

Third, the collected BS3 cross-links were used to construct the Bayesian scoring function that restrained the distances spanned by the cross-linked residues (35 A) (54, 67, 68), taking into account the ambiguity due to multiple copies of identical subunits (54, 64); the ambiguous cross-link restraint considers all possible pairwise distances in multiple copies of identical subunits, scoring only the shortest distance. The cross-link restraint was applied to the fine scale representation for the X-ray structures and comparative models as well as to flexible beads.

Fourth, the EM 3D restraint corresponded to the cross-correlation coefficient between the GMM representation of each component and the GMM representation of the consensus EM map (16, 62). The EM map was approximated by a 680 Gaussian GMM computed using the expectation-maximization as implemented in scikit-learn (63); 680 Gaussians appeared to be sufficient to reproduce the significant features of the map.

3. Configurational sampling and scoring: Structural models of the Rvb1/Rvb2/Ino80INS complex were computed using Replica Exchange Gibbs sampling, based on the Metropolis Monte Carlo algorithm (65, 67). The Monte Carlo moves included random translation and rotation of rigid bodies (up to 1.5 Å and 0.03 radians, respectively) and random translation of individual beads in the flexible segments (up to 2 Å). 32 replicas were used for each run, with temperatures ranging between 1.0 and 2.5. A structure model was saved every 10 Gibbs sampling steps, each consisting of a cycle of Monte Carlo steps that moved every rigid body and flexible bead once. The sampling produced a total of ~500,000 conformations from the 20 independent runs. 500 top-scoring structures from Step 3 were subjected to the subsequent analysis in Stage 4.

4. Analysis and validation of the ensemble:

Input information and output structures were analyzed to estimate structure precision and accuracy, detect inconsistencies among input data and missing information, and to suggest more informative future experiments. Assessment begins with structural clustering of the modeled structures produced by sampling, followed by assessment of the thoroughness of structural sampling, estimating structure precision based on variability in the ensemble of good-scoring structures and quantification of the structure fit to the input information. These validations are based on the nascent wwPDB effort on archival, validation, and dissemination of integrative structure models (69). We now discuss each one of these points in turn.

Clustering: A prerequisite for analysis of the ensemble of structures generated by satisfying the input data is clustering of these structures (21, 43, 70). We used Cα root- mean-square deviation (RMSD) quality-threshold clustering (43). For the Rvb1/Rvb2/Ino80INS complex, the

500 best scoring models grouped into 3 clusters. The average RMSD between the major (196 structures) and minor clusters (180 and 124 structures) was 86.4 Å and 77.3 Å respectively. Despite the large RMSD between clusters, the localization of all components is effectively identical between the clusters, differing mainly in the orientation of the Ino80INS-MBP domain (Figure S7C). Most importantly, our functional interpretation of the structure is the same regardless of which cluster we analyze.

Sampling convergence: For stochastic methods, thoroughness of sampling can be assessed by showing that two independent runs or set of runs (e.g., using random starting configurations or different random number generator seeds) do not result in significantly different solutions (21, 43, 64). Given two or more sets of structures from independent runs, we first cluster structures from all sets together, followed by assessing whether or not the runs contribute evenly to the population of each cluster, using the p-value from the χ-square contingency test for homogeneity of proportions (71). For the Rvb1/Rvb2/Ino80INS complex, the significant p-value of 0.15 indicated that our Monte Carlo algorithm sampled all top-scoring solutions at the resolution better than the precision of the dominant cluster. The caveat is that passing this sampling test is not absolute evidence of thorough sampling; a positive outcome of the test may be misleading if, for example, the landscape contains only a narrow, and thus difficult to find, pathway to the pronounced minimum corresponding to the correct structure.

Estimating structure precision based in variability in the ensemble of good-scoring structures: The ensemble of the top-scoring structures was analyzed in terms of the precision of its structural features (21, 72). The spread around the maximum describes how precisely the feature is determined from the input information. The precision of component position is quantified as the average root-mean-square fluctuation (RMSF) across all pairs of in the cluster,

after least-squares superposition onto the centroid structure structures (43) (Figure S5C), likely provided the lower bound on accuracy.For Cluster 1, the precision of the complex is 25 Å. The precision of the Rvbs dodecamer and the Ino80INS-MBP domain is 15.9 and 39.9 Å, respectively. This precision is sufficiently high to pinpoint the locations of the Ino80INS and DII domains (Figures 1A and S5C, and Table 1), demonstrating the quality of the data including the cross-links and 3D-EM map. For Clusters 2 and 3, the precision of the complex is 25 Å (14.8 Å for the Rvbs and 40.2 Å for the Ino80INS-MBP domains) and 12.2 Å (6.2 Å for the Rvbs and 14.3 Å for the Ino80INS-MBP domains), respectively. The localization probability density maps of every Rvb1/Rvb2/Ino80INS subunit as well as the whole complex were computed from the dominant cluster of the 189 solutions (Figures 6B) and the other two clusters (Figure S7C).

Fit to input data: The ensemble of solutions was assessed in terms of how well they satisfied information from which they were computed, including the cross-links, the 3D-EM, the excluded volume and sequence connectivity restraints.

First, 97%, 95%, and 91% of all crosslinks are satisfied by at least one structure in Clusters 1, 2, and 3, respectively. 98% of all crosslinks are satisfied by at least one structure in the three clusters; a cross-link restraint was satisfied by a cluster if the Cα-Cα distance between the cross-linked residues (considering restraint ambiguity) was less than 30 Å in any of the structures in the cluster.

Second, the solutions fit the 3D-EM consensus map with an average cross-correlation of 0.94.

Third, 99% of the top 500 solutions satisfied the excluded volume and sequence connectivity restraints under a combined score threshold of 50.

Supp lemental References

1. Singleton, M. R., Dillingham, M. S., and Wigley, D. B. (2007) Structure and mechanism of helicases and nucleic acid translocases. *Annu. Rev. Biochem.* 76, 23–50

2. White, S. R., and Lauring, B. (2007) AAA+ ATPases: achieving diversity of function with conserved machinery. *Traffic* 8, 1657–1667

3. Matias, P. M., Baek, S. H., Bandeiras, T. M., Dutta, A., Houry, W. A., Llorca, O., and Rosenbaum, J. (2015) The AAA+ proteins Pontin and Reptin enter adult age: from understanding their basic biology to the identification of selective inhibitors. *Front Mol Biosci* 2, 17

4. Nano, N., and Houry, W. A. (2013) Chaperone-like activity of the AAA+ proteins Rvb1 and Rvb2 in the assembly of various complexes. *Philos. Trans. R. Soc. Lond., B, Biol. Sci.* 368, 20110399

5. Rosenbaum, J., Baek, S. H., Dutta, A., Houry, W. A., Huber, O., Hupp, T. R., and Matias, P. M. (2013), p mr1.

6. Grigoletto, A., Lestienne, P., and Rosenbaum, J. (2011) The multifaceted proteins Reptin and Pontin as major players in cancer. *Biochim. Biophys. Acta* 1815, 147–157

7. Grigoletto, A., Neaud, V., Allain-Courtois, N., Lestienne, P., and Rosenbaum, J. (2013) The ATPase activity of reptin is required for its effects on tumor cell growth and viability in hepatocellular carcinoma. *Mol. Cancer Res.* 11, 133–139

8. Huber, O., Ménard, L., Haurie, V., Nicou, A., Taras, D., and Rosenbaum, J. (2008) Pontin and reptin, two related ATPases with multiple roles in cancer. *Cancer Res.* 68, 6873–6876

9. Ménard, L., Taras, D., Grigoletto, A., Haurie, V., Nicou, A., Dugot-Senant, N., Costet, P., Rousseau, B., and Rosenbaum, J. (2010) In vivo silencing of Reptin blocks the progression of human hepatocellular carcinoma in xenografts and is associated with replicative senescence. *J. Hepatol.* 52, 681–689

10. Gribun, A., Cheung, K. L. Y., Huen, J., Ortega, J., and Houry, W. A. (2008) Yeast Rvb1 and Rvb2 are ATP-dependent DNA helicases that form a heterohexameric complex. *J. Mol. Biol.* 376, 1320–1333

11. Kanemaki, M., Kurokawa, Y., Matsu-ura, T., Makino, Y., Masani, A., Okazaki, K., Morishita, T., and Tamura, T. A. (1999) TIP49b, a new RuvB-like DNA helicase, is included in a complex together with another RuvB-like DNA helicase, TIP49a. *J. Biol. Chem.* 274, 22437–22444

12. Makino, Y., Kanemaki, M., Kurokawa, Y., Koji, T., and Tamura, T. A. (1999) A rat RuvB-like protein, TIP49a, is a germ cell-enriched novel DNA helicase. *J. Biol. Chem.* 274, 15329–15335

13. Ikura, T., Ogryzko, V. V., Grigoriev, M., Groisman, R., Wang, J., Horikoshi, M., Scully, R., Qin, J., and Nakatani, Y. (2000) Involvement of the TIP60 histone acetylase complex in DNA repair and apoptosis. *Cell* 102, 463–473

14. Matias, P. M., Gorynia, S., Donner, P., and Carrondo, M. A. (2006) Crystal structure of the human AAA+ protein RuvBL1. *J. Biol. Chem.* 281, 38918–38929

15. Qiu, X. B., Lin, Y. L., Thome, K. C., Pian, P., Schlegel, B. P., Weremowicz, S., Parvin, J. D., and Dutta, A. (1998) An eukaryotic RuvB-like protein (RUVBL1) essential for growth. *J. Biol. Chem.* 273, 27786–27793

16. Robinson, P. J., Trnka, M. J., Pellarin, R., Greenberg, C. H., Bushnell, D. A., Davis, R., Burlingame, A. L., Sali, A., and Kornberg, R. D. (2015) Molecular architecture of the yeast Mediator complex. *Elife* 4, 44

17. Trnka, M. J., Baker, P. R., Robinson, P. J. J., Burlingame, A. L., and Chalkley, R. J. (2014) Matching cross-linked peptide spectra: only as good as the worse identification. *Mol. Cell Proteomics* 13, 420–434

18. Zheng, S. Q., Palovcak, E., Armache, J.-P., Verba, K. A., Cheng, Y., and Agard, D. A. (2016) MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat Meth* 14, 331–332

19. Zhang, K. (2016) Gctf: Real-time CTF determination and correction. *J. Struct. Biol.* 193, 1–12

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

21. Alber, F., Dokudovskaya, S., Veenhoff, L. M., Zhang, W., Kipper, J., Devos, D., Suprapto, A., Karni-Schmidt, O., Williams, R., Chait, B. T., Rout, M. P., and Sali, A. (2007) Determining the architectures of macromolecular assemblies. *Nature* 450, 683–694

22. Russel, D., Lasker, K., Ben Webb, Velázquez-Muriel, J., Tjioe, E., Schneidman-Duhovny, D., Peterson, B., and Sali, A. (2012) Putting the Pieces Together: Integrative Modeling Platform Software for Structure Determination of Macromolecular Assemblies. *PLOS Biol* 10, e1001244

23. Jónsson, Z. O., Jha, S., Wohlschlegel, J. A., and Dutta, A. (2004) Rvb1p/Rvb2p recruit Arp5p and assemble a functional Ino80 chromatin remodeling complex. *Mol. Cell* 16, 465–477

24. Chen, L., Conaway, R. C., and Conaway, J. W. (2013) 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.* 110, 20497–20502

25. Puri, T., Wendler, P., Sigala, B., Saibil, H., and Tsaneva, I. R. (2007) Dodecameric structure and ATPase activity of the human TIP48/TIP49 complex. *J. Mol. Biol.* 366, 179–192

26. Street, T. O., Lavery, L. A., and Agard, D. A. (2011) Substrate binding drives large-scale conformational changes in the Hsp90 molecular chaperone. *Mol. Cell* 42, 96–105

27. Leitner, A., Faini, M., Stengel, F., and Aebersold, R. (2016) Crosslinking and Mass Spectrometry: An Integrated Technology to Understand the Structure and Function of Molecular Machines. *Trends Biochem. Sci.* 41, 20–32

28. Niewiarowski, A., Bradley, A. S., Gor, J., McKay, A. R., Perkins, S. J., and Tsaneva, I. R. (2010) Oligomeric assembly and interactions within the human RuvB-like RuvBL1 and RuvBL2 complexes. *Biochem. J.* 429, 113–125

29. Gorynia, S., Bandeiras, T. M., Pinho, F. G., McVey, C. E., Vonrhein, C., Round, A., Svergun, D. I., Donner, P., Matias, P. M., and Carrondo, M. A. (2011) Structural and functional insights into a dodecameric molecular machine - the RuvBL1/RuvBL2 complex. *J. Struct. Biol.* 176, 279–291

30. Tosi, A., Haas, C., Herzog, F., Gilmozzi, A., Berninghausen, O., Ungewickell, C., Gerhold, C. B., Lakomek, K., Aebersold, R., Beckmann, R., and Hopfner, K.-P. (2013) Structure and subunit topology of the INO80 chromatin remodeler and its nucleosome complex. *Cell* 154, 1207–1219

31. Jeganathan, A., Leong, V., Zhao, L., Huen, J., Nano, N., Houry, W. A., and Ortega, J. (2015) Yeast Rvb1 and Rvb2 Proteins Oligomerize As a Conformationally Variable Dodecamer with Low Frequency. *J. Mol. Biol*,

32. Heck, A. J. R. (2008) Native mass spectrometry: a bridge between interactomics and structural biology. *Nat. Methods* 5, 927–933

33. Rose, R. J., Damoc, E., Denisov, E., Makarov, A., and Heck, A. J. R. (2012) High-sensitivity Orbitrap mass analysis of intact macromolecular assemblies. *Nat. Methods* 9, 1084–1086

34. Konermann, L., Vahidi, S., and Sowole, M. A. (2014) Mass spectrometry methods for studying structure and dynamics of biological macromolecules. *Anal. Chem.* 86, 213–232

35. Cheung, K. L. Y., Huen, J., Kakihara, Y., Houry, W. A., and Ortega, J. (2010) Alternative oligomeric states of the yeast Rvb1/Rvb2 complex induced by histidine tags. *J. Mol. Biol.* 404, 478–492

36. Petukhov, M., Dagkessamanskaja, A., Bommer, M., Barrett, T., Tsaneva, I., Yakimov, A., Queval, R., Shvetsov, A., Khodorkovskiy, M., Käs, E., and Grigoriev, M. (2012) Large-scale conformational flexibility determines the properties of AAA+ TIP49 ATPases. *Structure* 20, 1321–1331

37. López-Perrote, A., Muñoz-Hernández, H., Gil, D., and Llorca, O. (2012) Conformational transitions regulate the exposure of a DNA-binding domain in the RuvBL1-RuvBL2 complex. *Nucleic Acids Res.* 40, 11086–11099

38. Ewens, C. A., Su, M., Zhao, L., Nano, N., Houry, W. A., and Southworth, D. R. (2016) Architecture and Nucleotide-Dependent Conformational Changes of the Rvb1-Rvb2 AAA+ Complex Revealed by Cryoelectron Microscopy. *Structure* 24, 657–666

39. Lakomek, K., Stoehr, G., Tosi, A., Schmailzl, M., and Hopfner, K.-P. (2015) Structural Basis for Dodecameric Assembly States and Conformational Plasticity of the Full-Length AAA+ ATPases Rvb1·Rvb2. *Structure*,

40. Silva-Martin, N., Daudén, M. I., Glatt, S., Hoffmann, N. A., Kastritis, P., Bork, P., Beck, M., and Müller, C. W. (2016) The Combination of X-Ray Crystallography and Cryo-Electron Microscopy Provides Insight into the Overall Architecture of the Dodecameric Rvb1/Rvb2 Complex. *PLoS ONE* 11, e0146457

41. Iyer, L. M., Leipe, D. D., Koonin, E. V., and Aravind, L. (2004) Evolutionary history and higher order classification of AAA+ ATPases. *J. Struct. Biol.* 146, 11–31

42. Kim, M. (2013) Beta conformation of polyglutamine track revealed by a crystal structure of Huntingtin N-terminal region with insertion of three histidine residues. *Prion* 7, 221–228

43. Shi, Y., Fernandez-Martinez, J., Tjioe, E., Pellarin, R., Kim, S. J., Williams, R., Schneidman-Duhovny, D., Sali, A., Rout, M. P., and Chait, B. T. (2014) Structural characterization by cross-linking reveals the detailed architecture of a coatomer-related heptameric module from the nuclear pore complex. *Mol. Cell Proteomics* 13, 2927–2943

44. McKeegan, K. S., Debieux, C. M., and Watkins, N. J. (2009) Evidence that the AAA+ proteins TIP48 and TIP49 bridge interactions between 15.5K and the related NOP56 and NOP58 proteins during box C/D snoRNP biogenesis. *Mol. Cell. Biol.* 29, 4971–4981

45. Izumi, N., Yamashita, A., Iwamatsu, A., Kurata, R., Nakamura, H., Saari, B., Hirano, H., Anderson, P., and Ohno, S. (2010) AAA+ proteins RUVBL1 and RUVBL2 coordinate PIKK activity and function in nonsense-mediated mRNA decay. *Sci Signal* 3, ra27–ra27

46. Venteicher, A. S., Meng, Z., Mason, P. J., Veenstra, T. D., and Artandi, S. E. (2008) Identification of ATPases pontin and reptin as telomerase components essential for holoenzyme assembly. *Cell* 132, 945–957

47. Verba, K. A., Wang, R. Y.-R., Arakawa, A., Liu, Y., Shirouzu, M., Yokoyama, S., and Agard, D. A. (2016) Atomic structure of Hsp90-Cdc37-Cdk4 reveals that Hsp90 traps and stabilizes an unfolded kinase. *Science* 352, 1542–1547

48. Watanabe, S., Tan, D., Lakshminarasimhan, M., Washburn, M. P., Erica Hong, E.-J., Walz, T., and Peterson, C. L. (2015) Structural analyses of the chromatin remodelling enzymes INO80-C and SWR-C. *Nat Commun* 6, 7108

49. Zaarur, N., Xu, X., Lestienne, P., Meriin, A. B., McComb, M., Costello, C. E., Newnam, G. P., Ganti, R., Romanova, N. V., Shanmugasundaram, M., Silva, S. T., Bandeiras, T. M., Matias, P. M., Lobachev, K. S., Lednev, I. K., Chernoff, Y. O., and Sherman, M. Y. (2015) RuvbL1 and RuvbL2 enhance aggresome formation and disaggregate amyloid fibrils. *EMBO J.* 34, 2363–2382

50. Magalska, A., Schellhaus, A. K., Moreno-Andrés, D., Zanini, F., Schooley, A., Sachdev, R., Schwarz, H., Madlung, J., and Antonin, W. (2014) RuvB-like ATPases function in chromatin decondensation at the end of mitosis. *Dev. Cell* 31, 305–318

51. Jónsson, Z. O., Dhar, S. K., Narlikar, G. J., Auty, R., Wagle, N., Pellman, D., Pratt, R. E., Kingston, R., and Dutta, A. (2001) Rvb1p and Rvb2p are essential components of a chromatin remodeling complex that regulates transcription of over 5% of yeast genes. *J. Biol. Chem.* 276, 16279–16288

52. Webb, B., Lasker, K., Schneidman-Duhovny, D., Tjioe, E., Phillips, J., Kim, S. J., Velázquez-Muriel, J., Russel, D., and Sali, A. (2011) Modeling of proteins and their assemblies with the integrative modeling platform. *Methods Mol. Biol.* 781, 377–397

53. Kim, S. J., Fernandez-Martinez, J., Sampathkumar, P., Martel, A., Matsui, T., Tsuruta, H., Weiss, T. M., Shi, Y., Markina-Inarrairaegui, A., Bonanno, J. B., Sauder, J. M., Burley, S. K., Chait, B. T., Almo, S. C., Rout, M. P., and Sali, A. (2014) Integrative structure-function mapping of the nucleoporin Nup133 suggests a conserved mechanism for membrane anchoring of the nuclear pore complex. *Mol. Cell Proteomics* 13, 2911–2926

54. LoPiccolo, J., Kim, S. J., Shi, Y., Wu, B., Wu, H., Chait, B. T., Singer, R. H., Sali, A., Brenowitz, M., Bresnick, A. R., and Backer, J. M. (2015) Assembly and Molecular Architecture of the Phosphoinositide 3-Kinase p85α Homodimer. *J. Biol. Chem.* 290, 30390–30405

55. Luo, J., Cimermančič, P., Viswanath, S., Ebmeier, C. C., Kim, B., Dehecq, M., Raman, V., Greenberg, C. H., Pellarin, R., Sali, A., Taatjes, D. J., Hahn, S., and Ranish, J. (2015) Architecture of the Human and Yeast General Transcription and DNA Repair Factor TFIIH. *Mol. Cell* 59, 794–806

56. Sali, A., and Blundell, T. L. (1993) Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* 234, 779–815

57. Söding, J., Biegert, A., and Lupas, A. N. (2005) The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res.* 33, W244–8

58. Pieper, U., Webb, B. M., Dong, G. Q., Schneidman-Duhovny, D., Fan, H., Kim, S. J., Khuri, N., Spill, Y. G., Weinkam, P., Hammel, M., Tainer, J. A., Nilges, M., and Sali, A. (2014) ModBase, a database of annotated comparative protein structure models and associated resources. *Nucleic Acids Res.* 42, D336–46

59. McGuffin, L. J., Bryson, K., and Jones, D. T. (2000) The PSIPRED protein structure prediction server. *Bioinformatics* 16, 404–405

60. Drozdetskiy, A., Cole, C., Procter, J., and Barton, G. J. (2015) JPred4: a protein secondary structure prediction server. *Nucleic Acids Res.* 43, W389–94

61. Ward, J. J., McGuffin, L. J., Bryson, K., Buxton, B. F., and Jones, D. T. (2004) The DISOPRED server for the prediction of protein disorder. *Bioinformatics* 20, 2138–2139

62. Kawabata, T. (2008) Multiple subunit fitting into a low-resolution density map of a macromolecular complex using a gaussian mixture model. *Biophys. J.* 95, 4643–4658

63. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, É. (2011) Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research* 12, 2825–2830

64. Algret, R., Fernandez-Martinez, J., Shi, Y., Kim, S. J., Pellarin, R., Cimermančič, P., Cochet, E., Sali, A., Chait, B. T., Rout, M. P., and Dokudovskaya, S. (2014) Molecular architecture and function of the SEA complex, a modulator of the TORC1 pathway. *Mol. Cell Proteomics* 13, 2855–2870

65. Rieping, W., Habeck, M., and Nilges, M. (2005) Inferential structure determination. *Sci Signal* 309, 303–306

66. Shen, M.-Y., and Sali, A. (2006) Statistical potential for assessment and prediction of protein structures. *Protein Sci.* 15, 2507–2524

67. Shi, Y., Pellarin, R., Fridy, P. C., Fernandez-Martinez, J., Thompson, M. K., Li, Y., Wang, Q. J., Sali, A., Rout, M. P., and Chait, B. T. (2015) A strategy for dissecting the architectures of native macromolecular assemblies. *Nat. Methods* 12, 1135–1138

68. Erzberger, J. P., Stengel, F., Pellarin, R., Zhang, S., Schaefer, T., Aylett, C. H. S., Cimermančič, P., Boehringer, D., Sali, A., Aebersold, R., and Ban, N. (2014) Molecular architecture of the 40S·eIF1·eIF3 translation initiation complex. *Cell* 158, 1123–1135

69. Sali, A., Berman, H. M., Schwede, T., Trewhella, J., Kleywegt, G., Burley, S. K., Markley, J., Nakamura, H., Adams, P., Bonvin, A. M. J. J., Chiu, W., Peraro, M. D., Di Maio, F., Ferrin, T. E., Grünewald, K., Gutmanas, A., Henderson, R., Hummer, G., Iwasaki, K., Johnson, G., Lawson, C. L., Meiler, J., Marti-Renom, M. A., Montelione, G. T., Nilges, M., Nussinov, R., Patwardhan, A., Rappsilber, J., Read, R. J., Saibil, H., Schröder, G. F., Schwieters, C. D., Seidel, C. A. M., Svergun, D., Topf, M., Ulrich, E. L., Velankar, S., and Westbrook, J. D. (2015) Outcome of the First wwPDB Hybrid/Integrative Methods Task Force Workshop. *Structure* 23, 1156–1167

70. Zeng-Elmore, X., Gao, X.-Z., Pellarin, R., Schneidman-Duhovny, D., Zhang, X.-J., Kozacka, K. A., Tang, Y., Sali, A., Chalkley, R. J., Cote, R. H., and Chu, F. (2014) Molecular architecture of photoreceptor phosphodiesterase elucidated by chemical cross-linking and integrative modeling. *J. Mol. Biol.* 426, 3713–3728

71. McDonald, J. H., University of Delaware (2009) *Handbook of Biological Statistics*

72. Alber, F., Dokudovskaya, S., Veenhoff, L. M., Zhang, W., Kipper, J., Devos, D., Suprapto, A., Karni-Schmidt, O., Williams, R., Chait, B. T., Sali, A., and Rout, M. P. (2007) The molecular architecture of the nuclear pore complex. *Nature* 450, 695–701

Appendix to Chapter 2: Unpublished data

Included in this section are unpublished experimental results that may help foster future investigations into the Rvbs' function, particularly in their potential role in assembling the INO80 complex.

Methods

Native-MS of Rvb1/Rvb2 with different nucleotides

Native-MS was performed as described in the Methods section of Chapter 2. Importantly, nucleotide – Mg complexes were prepared fresh by incubating equimolar ratio of nucleotide and $MgCl_2$ for 30 minutes at 4 °C and then freshly diluting to the final working concentration. Protein (~ 0.5 to 1 μM) was incubated with excess nucleotide (20 μM) immediately prior to analysis.

Spectral counting analysis to visualize structural consequences of Ino80INS binding to Rvbs

Cross-link spectral matches (CSM) were initially discarded if the following Protein Prospector parameters fell outside the threshold values: peptide score below 20, peptide/protein, and score difference below 0. A linear support vector machine (SVM) model was constructed based on three parameters ("score", "score difference", and "peptide 2 rank") to further classify CSMs between decoy and target classes (1, 2). 5,208 spectra were classified with an SVM confidence score above the acceptance threshold of 0 at a false discovery rate (FDR) of 0.6%. The FDR at the residue pair level was 1.7%. CSMs where either peptide was less than 4-amino acids long were discarded. When spectra could be interpreted by multiple cross-linked peptide-pairs, CSMs were considered to be unambiguous if the second best match had a confidence score 0.3 lower than the top match. The remaining spectral redundancy came exclusively from ambiguous site localizations, where there was sufficient evidence to identify both peptides, but the position of the modified amino acid remained ambiguous. In these cases, the CSM was flagged as ambiguous and all possible sites were reported. Spectral counts were determined from the number of CSMs mapping to each unique residue pair. Visualizations of cross-link data against

protein domains were created using Cytoscape 3.3. Two-dimensional cross-linking maps were created with in-house matplotlib scripts.

Fractionation of Rvb1/Rvb2/Ino80INS by size exclusion chromatography

Rvb1/Rvb2/Ino80INS was expressed as described in the methods section of Chapter 2. After size exclusion chromatography, fractions were pooled and collected separately (a total of 8 pools, two fractions each). Each pool was separately quantified by gel using BSA standards.

Co-purification of Rvb1/Rvb2 with MBP-Swr1INS

Residues 1003-1237 of the Swr1 ATPase were cloned into the Rvb1/Rvb2 was co-expressed with MBP-Swr1INS (residues 1003-1237 of the Swr1 ATPase) as described previously for the Rvb1/Rvb2/Ino80INS complex.

Purification of MBP-Ies2 and MBP-Ies6

MBP-Ies2 and MBP-Ies6 were cloned into the pACYC vector with C-terminal his-tags. We attempted to co-express and co-purify each fusion protein individually with Rvb1/Rvb2, as we had done successfully for MBP-Swr1INS and MBP-Ino80INS. Next, we expressed and purified MBP-Ies2 and MBP-Ies6 individually in a similar procedure as described previously for Rvb1/Rvb2. Briefly, the proteins were bound to cobalt resin, eluted in buffer containing 500 mM imidazole, and further purified by size exclusion chromatography on an S200 sizing column.

Labeling of Rvb1 and Rvb2 with fluorescent dyes

Rvb1 and Rvb2 were separately expressed in bacteria using the gb11 vector. After removal of his-tags by TEV cleavage, the proteins were purified further using size exclusion chromatography. Enzyme was concentrated and stored in storage buffer (25 mM Hepes pH 7.5, 300 mM KCl, 2 mM DTT, 10% glycerol). On the day of labeling, enzyme was thawed and diluted

to 20 uM in enzyme storage buffer for a final volume of 150 uL. An equal volume of labeling buffer was added (600 mM KCl, 50 mM Hepes pH 7.5, 2 mM fresh TCEP), before buffer exchanging using the Zeba 7000 desalting spin columns. Next, either cy3-maleimide or cy5-maleimide were added to Rvb1 or Rvb2 respectively and the labeling reaction was incubated for 2 hours at room temperature in the dark. Next, ATP-Mg was added for a final concentration of 1 mM, and allowed labeling to occur further for 8 hours on ice. Used zeba columns again to spin away free dye, and protein was dialyzed into buffer containing 25 mM pH 75, 10% glycerol, 300 mM KCl, 5 mM BME.

Exchange of labeled Rvb1 and Rvb2 into the INO80 complex

Freshly labeled Rvbs and Ino80-flag were combined at the following concentrations: 12 nM Ino80-flag, 680 nM Rvb1-cy3, 1.4 uM Rvb1-cy5. Salt concentration was adjusted for a final concentration of 1 M KCl. An additional dose of ATP-Mg was also added at 1 mM, and the exchange reaction was incubated at 4 degrees overnight. The following morning, the exchange reaction was added to 300 uL of pre-equilibrated anti-flag resin (Sigma) and allowed to incubate at 4 degrees while rotating. Beads were washed twice with H0.5 buffer (Ino80-flag prep buffer containing 500 mM KCl, then exchanged into H0.1 buffer (Ino80-flag prep buffer containing 100 mM KCl). The remaining Ino80-flag complex was eluted with 1 mg/mL flag peptide, and beads were spun away using 0.22 um spin filters. The resulting eluate was concentrated to about 40 uL and the extent of the exchange was analyzed by SDS-PAGE.

Results

Rvb1/Rvb2 hexamers exhibit a distribution of nucleotide-bound states by native-MS.

Rvb1 and Rvb2 form hexameric ring structures with a potential of six separate active sites that can bind and hydrolyze ATP. A crystal structure of full length Rvb1/Rvb2 from *Chaetomium thermophilum* shows that indeed, each protomer can potentially bind to a single nucleotide

(Appendix Figure 1). However, it is unclear if the individual protomers act cooperatively or individually to bind nucleotide, a major outstanding question for AAA+ ATPases in general (3). To address this question, we used native-MS to analyze the distribution of nucleotide-bound states for both wild-type and Walker B mutants of Rvb1/Rvb2. For this study, we focused solely on the hexameric species of Rvbs. While a majority of Rvb1/Rvb2 purified from bacteria has a molecular mass most consistent with the apo state, there is a small population of Rvbs that is likely bound to nucleotide, as suggested by a separately resolved peak with a slightly higher m/z value (Appendix Figure 2A-B). The mass difference between the major and minor peaks is approximately the value of a nucleotide, suggesting that this smaller peak corresponds to a population of hexamers bound to a single nucleotide. Addition of an excess of ATP-Mg to this sample shifted these peaks even further toward higher m/z values (Appendix Figure 2C), suggesting the formation of states with more nucleotides bound per hexamer. These new peaks deconvolve to masses consistent with the Rvb hexamer bound to three, four, and five nucleotides bound, with the most populated state being one where four nucleotides are bound. We observed a similar behavior with the addition of ADP-Mg (Appendix Figure 2D).

Next, we repeated these experiments with Walker B mutations in either one or both of the Rvbs. This mutation greatly diminishes ATP hydrolysis activity without affecting ATP binding activity, often resulting in a higher population of nucleotide-bound species, as the nucleotide is “stuck” in the binding pocket. Indeed, without adding extra nucleotide, we found that mutating the Walker B residue in either one or both Rvbs results in an increase in the peak that corresponds to a nucleotide-bound state (Appendix Figure 3). Addition of ATP-Mg or ADP-Mg shifted these peaks toward higher values of m/z (Appendix Figure 3), similar to the results seen for wild type Rvbs. Taken together, these results suggest that nucleotide binding by Rvbs is not extremely cooperative, as several stoichiometries of nucleotide binding can exist.

Comparative spectral counting analysis of Rvb1/Rvb2/Ino80INS reveals major changes in DII domains when Rvbs are bound to the Ino80INS

With complete XL-MS datasets for Rvb1/Rvb2 alone as well as Rvb1/Rvb2/Ino80INS, we were next able to search for major differences between the two datasets, which may be indicative of structural changes that occur as a result of Ino80INS binding. The metric we used for this analysis is called spectral counting, defined as the absolute number of MS spectra identifying a given crosslink and is well-established as a quantitative measure of protein abundance. We visualized the regions of Rvb1/Rvb2 that showed the greatest amount of change upon Ino80INS binding by plotting the difference in spectral counts between the Rvb1/Rvb2 dataset and the Rvb1/Rvb2/Ino80INS dataset as a contact map (Appendix Figure 4). Crosslinks that increase in spectral abundance upon INO80INS binding are plotted in pink, while residues that decrease in spectral abundance are plotted in blue. The relative area of each circle is scaled to the difference in the number of counts. As expected, only positive crosslinks (pink) appeared between the Rvbs' and the Ino80INS domain. For the crosslinks within the Rvb1/Rvb2 complex, we observed only crosslinks that decreased in spectral count (blue). We focused on the most significantly affected regions of Rvb1/Rvb2 by plotting only those crosslinks in the Rvb1/Rvb2 dataset that change by greater than 65% and by at least 5 absolute counts. We expected that decreased spectral counts of Rvb-Rvb crosslinks would correlate with increased Ino80ID-Rvb counts at those positions. Crosslinks that met this expectation were classified as Group A, whereas residue pairs that show decreased spectral counts in Rvb-Rvb interations but with no corresponding increase in crosslinks to Ino80INS were classified as group Group B.

Group A crosslinks were found in the DII domains of both Rvb1 and Rvb2, and to a lesser extent the DIII domain of Rvb2. These interactions are best explained by direct binding of Ino80INS to these positions or by a local change in conformation of the DII domain. A striking example is the crosslink between K31 on the DI domain of Rvb1 and K192 on the Rvb1 DII domain. This pair of crosslinked residues is among the most abundant in the Rvb1/Rvb2 dataset

(64 counts), yet completely vanishes in the data set for Rvb1/Rvb2/Ino80INS (Appendix Figure 4-5). K31 and K193 are in close proximity (14 Å) when the DII domain of Rvb1 is in the “in” state, but exceed the expected crosslinking range when Rvb1 DII is in the “out” state (44.6 Å) (Appendix Figure 5A). This analysis suggests a model where Ino80 binding to the DII domain causes a conformational change of the DII domain from the “in” state towards the “out” state. Group B residue pairs fall outside of the DII domain and map to regions with no observed Ino80INS binding, along the outer edge of the Rvb1/Rvb2 ring (Appendix Figure 5B). These results raise the possibility that binding by the Ino80INS can induce an allosteric change both locally at the DII domain, and on a completely different interface of the Rvb1/Rvb2 complex.

Fractionation of Rvb1/Rvb2/Ino80INS reveals a gradient of different specific activities

A key finding from our work is the ability of Ino80INS to act as a robust activator of Rvbs’ ATPase activity. Further, our integrative modeling and native-MS data with suggest that multiple copies of Ino80INS can bind to a single Rvb1/Rvb2 dodecamer. This observation raises the possibility that the number of copies of Ino80INS could further regulate the specific activity of the Rvbs. To test this possibility, the broad peak that corresponds to pure Rvb1/Rvb2/Ino80INS from the sizing column was collected as eight separate fractions, with earlier fractions corresponding to species of larger molecular weight (Appendix Figure 6A). We reasoned that these larger molecular weight species may correspond to populations of Rvb1/Rvb2/Ino80INS with a higher number of Ino80INS bound per Rvb1/Rvb2 dodecamer. Next we measured the specific activity for ATP hydrolysis of these eight fractions and found that the population of Rvb1/Rvb2/Ino80INS that eluted earlier had a significantly higher (2-3 fold) specific activity than the enzyme in the fractions that eluted later (Appendix Figure 6B). This data suggests that the number of Ino80INS bound per dodecamer could further regulate the specific activity of the Rvbs.

The insertion domain of the Swr1 ATPase stimulates the ATPase activity and promotes dodecamerization of Rvb1/Rvb2.

As described in Figure 1 of Chapter 2, the ATPase activity of Rvb1/Rvb2 is greatly enhanced (16-fold) by the presence of a small domain of the Ino80 ATPase, termed the Ino80 insertion (Ino80INS). The Swr1 ATPase, belonging to the SWR remodeling complex, also contains an insertion domain, which we hypothesized may bind to the Rvbs as well. To test this hypothesis, we replaced the DNA sequence of the Ino80INS expression plasmid with that of the insertion of Swr1 (residues 1003-1237). Using the exact same co-expression and co-purification protocol that was used to create the Rvb1/Rvb2/Ino80INS complex, we were also able to pull down the Rvb1/Rvb2/Swr1INS complex (Appendix Figure 7A). The Rvb1/Rvb2/Swr1INS complex also displayed enhanced ATPase activity compared to Rvbs alone (Appendix Figure 7B), and promotes dodecamerization of the Rvbs (Appendix Figure 7C). These data suggest that the insertion domain, which is only present in the INO80 family of remodeling enzymes, serves as a specific binder and activator of the Rvbs. Implications of this finding is discussed further in the following section.

MBP-Ies2 and MBP-Ies6 do not co-purify with Rvb1/Rvb2, but do stimulate ATPase activity of Rvbs.

Our observation that the Ino80INS is a direct activator of Rvb1/Rvb2 suggested that the Ino80INS may serve as a client for the Rvbs' chaperone activity. We reasoned that other subunits of the INO80 complex could also act similarly as clients by binding to the DII domains of Rvb1/Rvb2 and stimulating its ATPase activity. In this model, the Rvbs have the potential to act as a scaffold on which different proteins could bind and interact, eventually leading to the assembly of a multi-subunit complex. To test this hypothesis, we chose two candidate proteins, Ies2 and Ies6, both subunits of the INO80 complex that have been shown in a crosslinking-MS study to have direct interactions with the DII domains of the Rvbs in the context of the native

INO80 complex from yeast (4). We cloned Ies2 and Ies6 and individually co-expressed them along with Rvb1 and Rvb2, as we had done before for Ino80INS. When we performed a pull down assay by pulling on this his-tag of either MBP-Ies2 or MBP-Ies6, we were unable to detect a direct interaction between the potential client and the Rvbs (Appendix Figure 8A). This result did not necessarily rule out the potential interaction between the Rvbs and the Ies proteins, only that the interaction is not strong enough to overcome the limitations of the pull down assay. Next, we purified MBP-Ies2 and MBP-Ies6 individually, without co-expression with Rvb1/Rvb2. When we added either MBP-Ies2 and MBP-Ies6 to Rvb1/Rvb2 in trans, we observed a small but significant stimulation of ATPase activity (Appendix Figure 8B), consistent with Ies2 and Ies6 acting as clients for the Rvbs. In the case of Ies6, the extent of the stimulation was higher when the concentration of Ies6 was increased, consistent with Ies6 being a lower affinity binder of Rvb1/Rvb2 compared to Ino80INS. Importantly, efforts to purify Ies2 or Ies6 without the MBP tag resulted in heavy degradation, suggesting that these two proteins are structurally unstable on their own.

Next, we used negative stain EM to explore the structural consequences of MBP-Ies2 and MBP-Ies6 binding on Rvb1/Rvb2 structure. Rvb1/Rvb2 alone form symmetrical ring shapes on negative stain grids, while addition of Ies2 or Ies6 resulted in those ring shapes to be deformed, with many particles looking crescent shaped, as if one or more subunits in the hexamer are missing (Appendix Figure 8C). These data, while preliminary, further suggest that there is a direct and detectable interaction between these Ies proteins and the Rvbs that may be of functional significance for the Rvbs' chaperone-like activity.

Labeling and exchanging bacterially expressed Rvb1 and Rvb2 into the INO80 complex purified from yeast

The role of the Rvbs in the INO80 complex, as well as other essential complexes in the cell, is not known. There are two general hypotheses for how the Rvbs may function in the context of

chromatin remodeling. In one model, the Rvbs solely assemble an active complex, after which their activity is no longer required. In a second model, the Rvbs play an active role in nucleosome remodeling, perhaps by assembling and re-assembling different states of the INO80 complex during the remodeling reaction. To distinguish between these two models, it would be ideal to be able to make inactivating mutations in the Rvbs in the context of the native INO80 complex. This is a difficult task however, since both Rvb1 and Rvb2 are essential genes. To get around this challenge, one group placed the wild-type Rvb1 and Rvb2 genes on a degron plasmid, while making Walker B mutations to the endogenous locus of Rvb2 (5). Once degradation of the wild-type protein was initiated, only inactive Rvb2 should be produced and assembled into the Ino80 complex. The group found that mutating the Walker B residue of Rvb2 did not destroy remodeling, but did cause a key subunit Arp5 to be lost from the purified complexes. These results are contradictory, since it was recently shown that Arp5 is critical for activity of the INO80 (6).

As an alternative test of the models, we decided to use an in vitro exchange assay, where an excess of bacterially purified Rvbs would be used to replace the Rvbs in natively purified INO80 complexes from yeast. If it works, this assay would be a powerful way of assessing the effect of Rvb mutants on the assembly and activity of the INO80 complex. First we bulk-labeled Rvb1 and Rvb2 with either cy3 or cy5 dyes, relying on the low-cysteine content of these proteins to prevent any major deleterious effects on function. Next, we combined an excess of labeled Rvbs with native INO80 complexes purified from yeast, along with ATP-Mg to facilitate the exchange. After an overnight incubation, we re-purified the native INO80 complexes by binding to anti-flag beads and washing with high salt. Presumably, the only contents left bound to the beads upon elution would be proteins that specifically form the INO80 complex. In a preliminary test of the experiment, we found that Rvbs could be successfully labeled through bulk labeling approaches, and at least some amount of these labeled Rvbs were incorporated into the final INO80 complex (Appendix Figure 9). However, it is unclear how

much of the endogenous Rvbs are left in the sample, and the quantities recovered were insufficient for testing in a remodeling assay. Interestingly, the gel of the final eluted INO80 complex showed a loss in the band corresponding to Arp5, consistent with the previous study showing that Arp5 is lost upon mutation of Rvb2. While preliminary, these results suggest that there is potential in using an in vitro exchange assay to test the effects of Rvb mutants in the assembly and activity of INO80.

Discussion

While the Rvbs have been proposed to act as chaperones for assembling multi-subunit complexes, there is still little evidence for this hypothesis. Our published finding that the Ino80INS is a potent activator of the Rvbs' ATPase activity is the first experimental result that demonstrates the potential for a chaperone-like activity (7), yet raises further questions about how the activation of the Rvbs' ATPase activity leads to the assembly of a multi-subunit complex. The unpublished experimental results presented above help to frame these new questions, as discussed below.

How general of a chaperone are the Rvbs?

Our results with the Swr1INS being able to activate the Rvbs is exciting, as it demonstrates the ability of the Rvbs to act on more than a single potential client. Not only are SWR and INO80 the only chromatin remodeling complexes to have these long insertion regions within their ATPases, but they are also the only remodelers that contain the Rvbs (8). Furthermore, the insertion region is required for the recruitment of subunits that are specific to either SWR or INO80, both known to be structurally and functionally modular multi-subunit complexes. From these observations an exciting idea emerges: could the insertion regions of Ino80 and Swr1 have evolved to co-opt the chaperone function of the Rvbs? One might imagine that a driving force for this co-evolution is the modularity of these complexes, which provide a more sophisticated

form of regulation that is not found in complexes where the remodeling ATPase alone is sufficient for nucleosome movement.

In our search for additional clients of Rvb1/Rvb2, we found that Ies2 and Ies6, despite having distinct interactions with the Rvbs in the context of the native INO80 complex, do not pull down with the Rvbs the way the insertion regions of Ino80 and Swr1 did. This result emphasizes the uniqueness of the insertion regions, and further suggests that there may be an ordered process of assembly that we have not yet been able to reconstitute. Although we could not detect a direct interaction between the Rvbs and the Ies proteins, we did observe a modest stimulation in ATPase activity of the Rvbs when either Ies2 or Ies6 were added in trans. We observed a similar effect with the addition of Δ131Δ, a model substrate normally used for studies of Hsp90 (7). MBP alone, however, does not stimulate the ATPase of Rvbs (data not shown). Taken together, these results demonstrate the dynamic range in how Rvbs may interact with different proteins. To push our understanding of Rvb-client interactions forward, it would be important to know what aspects of protein structure determines the specificity of these interactions. A common feature among client proteins is likely the presence of unstructured regions, as all of the proteins that we have found to activate Rvb activity are too unstable to be purified on their own, or without an MBP tag. Future studies will be needed to reveal additional specificities, which will undoubtedly help expedite the identification of biologically relevant clients, and vice versa.

What is the ATPase activity actually being used for?

The ability of the Rvbs to bind to different client proteins and activate the ATPase is only the beginning of how it helps assemble multi-subunit complexes. What happens next? We know from our own work that the key to this puzzle is likely the DII domains, which not only bind to the client proteins, but also regulate activity and dodecamerization. The importance of these DII domains is highlighted by our spectral counting analysis, which show that the residues

undergoing the largest changes in spectral counts upon Ino80INS binding are in the DII domains (Appendix Figure 4 and 5). We also showed that larger molecular weight species of Rvb1/Rvb2/Ino80INS display the largest specific activity. If these larger species correspond to more copies of Ino80INS being bound to the Rvb dodecamer, it would suggest that the DII domains can act individually, its ATPase being activated by the binding of a client protein.

With limited information on how the Rvbs may act on a client, we can imagine at least two models. In one model, the ATPase activity of the Rvbs, which is likely regulated by the conformation of the DII domains, is simply used as a timer for assembly, where clients must be properly assembled on the dodecameric interface before ATP is hydrolyzed. Otherwise the dodecamer will fall apart and re-assembly must be initiated by the re-binding of client. If this dynamic does exist, we could not capture it, likely due to the unstable nature of our minimal system. In a second model, the Rvbs' ATPase activity is used to distort the client protein in some way, perhaps toward a conformation that is more favorable for interacting with an additional subunit. Future work toward reconstituting the steps of assembly will be necessary to distinguish between these models.

References

1. Robinson, P. J., Trnka, M. J., Pellarin, R., Greenberg, C. H., Bushnell, D. A., Davis, R., Burlingame, A. L., Sali, A., and Kornberg, R. D. (2015) Molecular architecture of the yeast Mediator complex. *Elife* 4, 44

2. Trnka, M. J., Baker, P. R., Robinson, P. J. J., Burlingame, A. L., and Chalkley, R. J. (2014) Matching cross-linked peptide spectra: only as good as the worse identification. *Mol. Cell Proteomics* 13, 420–434

3. Stinson, B. M., Nager, A. R., Glynn, S. E., Schmitz, K. R., Baker, T. A., and Sauer, R. T. (2013) Nucleotide binding and conformational switching in the hexameric ring of a AAA+ machine. *Cell* 153, 628–639

4. Tosi, A., Haas, C., Herzog, F., Gilmozzi, A., Berninghausen, O., Ungewickell, C., Gerhold, C. B., Lakomek, K., Aebersold, R., Beckmann, R., and Hopfner, K.-P. (2013) Structure and subunit topology of the INO80 chromatin remodeler and its nucleosome complex. *Cell* 154, 1207–1219

5. Jónsson, Z. O., Jha, S., Wohlschlegel, J. A., and Dutta, A. (2004) Rvb1p/Rvb2p recruit Arp5p and assemble a functional Ino80 chromatin remodeling complex. *Mol. Cell* 16, 465–477

6. Yao, W., King, D. A., Beckwith, S. L., Gowans, G. J., Yen, K., Zhou, C., and Morrison, A. J. (2016) The INO80 Complex Requires the Arp5-Ies6 Subcomplex for Chromatin Remodeling and Metabolic Regulation. *Mol. Cell. Biol.* 36, 979–991

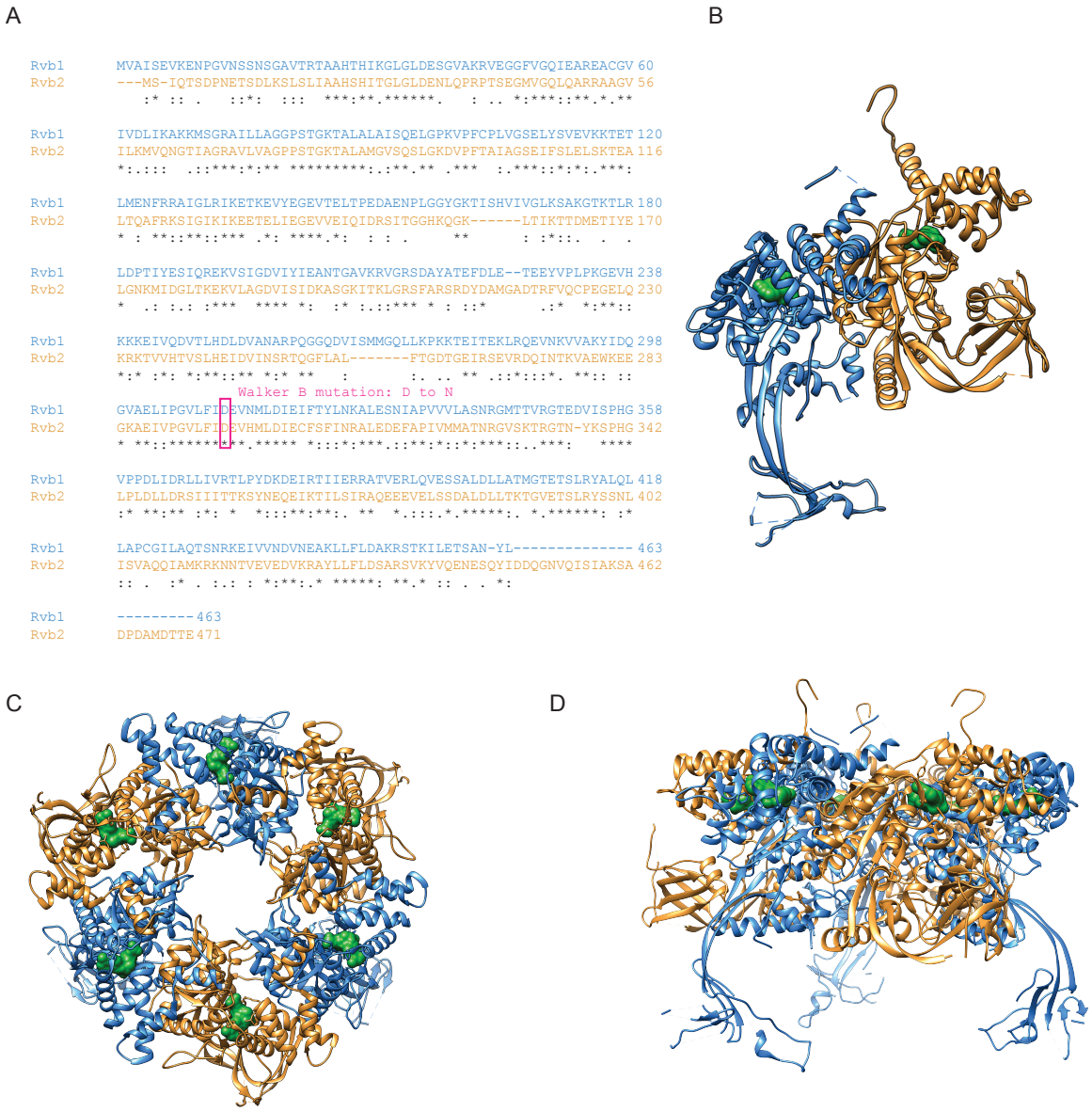
7. Zhou, C. Y., Stoddard, C. I., Johnston, J. B., Trnka, M. J., Echeverria, I., Palovcak, E., Sali, A., Burlingame, A. L., Cheng, Y., and Narlikar, G. J. (2017) Regulation of Rvb1/Rvb2 by a Domain within the INO80 Chromatin Remodeling Complex Implicates the Yeast Rvbs as

Protein Assembly Chaperones. *Cell Rep* 19, 2033–2044

8. Zhou, C. Y., Johnson, S. L., Gamarra, N. I., and Narlikar, G. J. (2016) Mechanisms of ATP-Dependent Chromatin Remodeling Motors. *Annu Rev Biophys* 45, 153–181

9. Lakomek, K., Stoehr, G., Tosi, A., Schmailzl, M., and Hopfner, K.-P. (2015) Structural Basis for Dodecameric Assembly States and Conformational Plasticity of the Full-Length AAA+ ATPases Rvb1·Rvb2. *Structure*, 23, 483-495

Appendix Figure 1: Sequence and structure of hexameric yeast Rvb1 and Rvb2

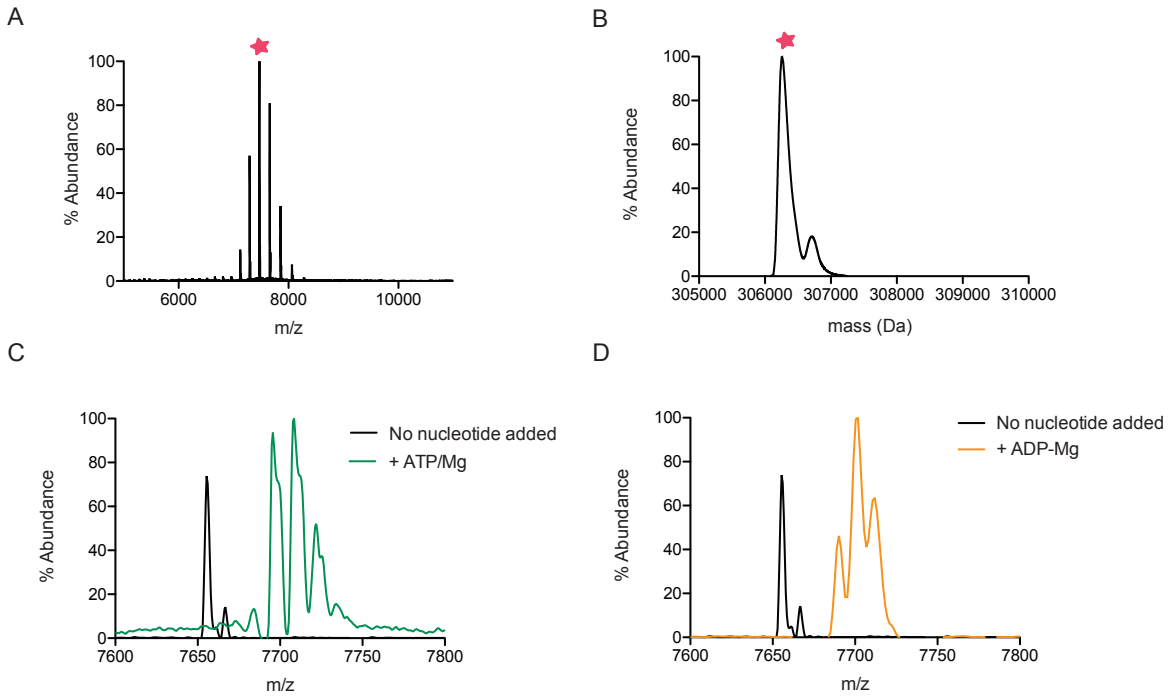


(A) Sequence alignments of Rvb1 (blue) and Rvb2 (orange). The pink box indicates the Walker B residue that is mutated in some experiments in this chapter.

(B) Homology model of the yeast Rvb1/Rvb2 heterodimer, built based on a published structure of Rvb1/Rvb2 from *Chaetomium thermophilum* ((9). Space filling model of ATP is shown in green.

(C) Top view and (D) side view of the Rvb1/Rvb2 hexamer.

Appendix Figure 2: Nucleotide binding by WT Rvb1/Rvb2 detected by native-MS



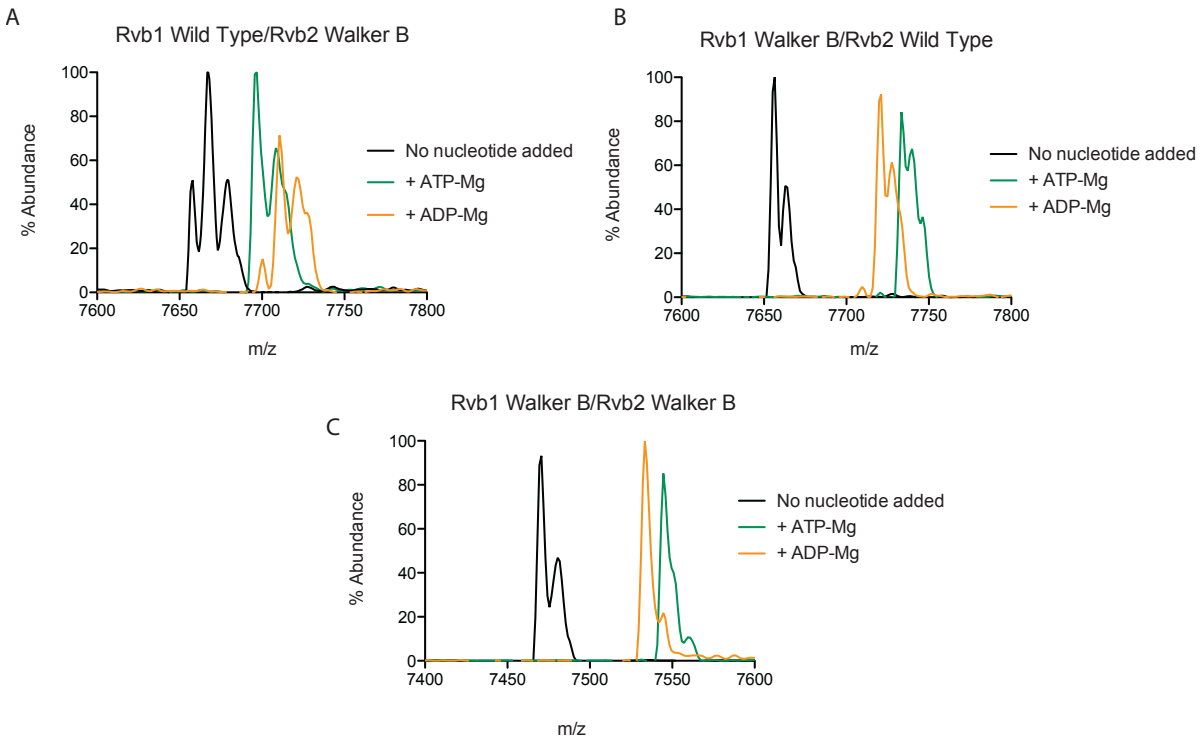
(A) Raw spectrum of 1 μM of WT Rvb1/Rvb2 by native-MS. Pink star indicates the peak that is shown in the deconvolved spectrum in (B)

(B) Deconvoluted mass spectrum of WT Rvb1/Rvb2 showing that the peak indicated in (A) deconvolves to two peaks: a major peak corresponding to the apo state of Rvb1/Rvb2 and a smaller peak that corresponds to Rvb1/Rvb2 bound to a single nucleotide.

(C) Raw mass spectrum of Rvb1/Rvb2 with and without additon of ATP-Mg.

(D) Raw mass spectrum of Rvb1/Rvb2 with and without additon of ADP-Mg.

Appendix Figure 3: Nucleotide binding profiles of different Walker B mutants of Rvb1/Rvb2 by native-MS

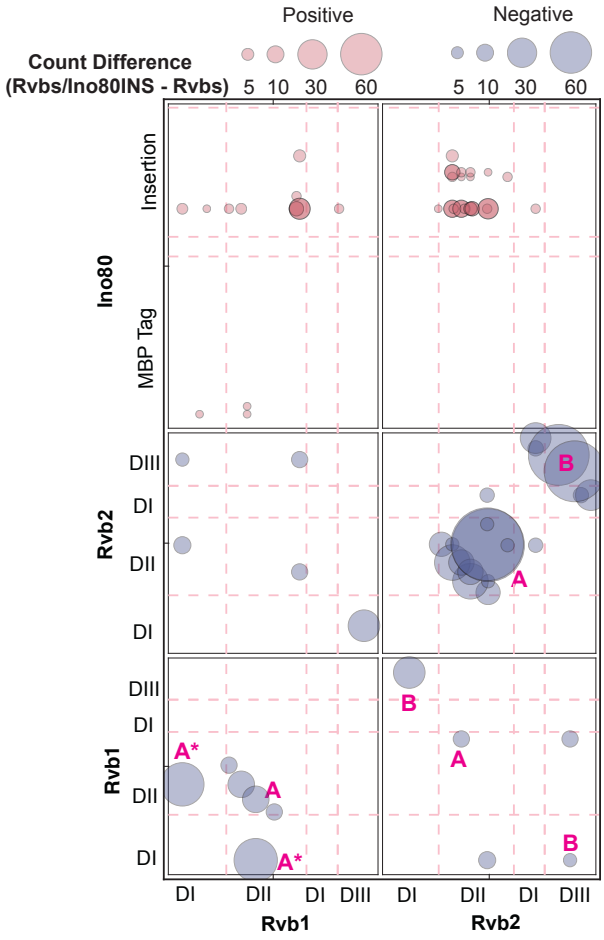


(A) Raw mass spectrum of Rvb1 WT/Rvb2 WB with and without addition of nucleotide.

(B) Raw mass spectrum of Rvb1 WB/Rvb2 WT with and without addition of nucleotide.

(C) Raw mass spectrum of Rvb1 WB/Rvb2 WB with and without addition of nucleotide.

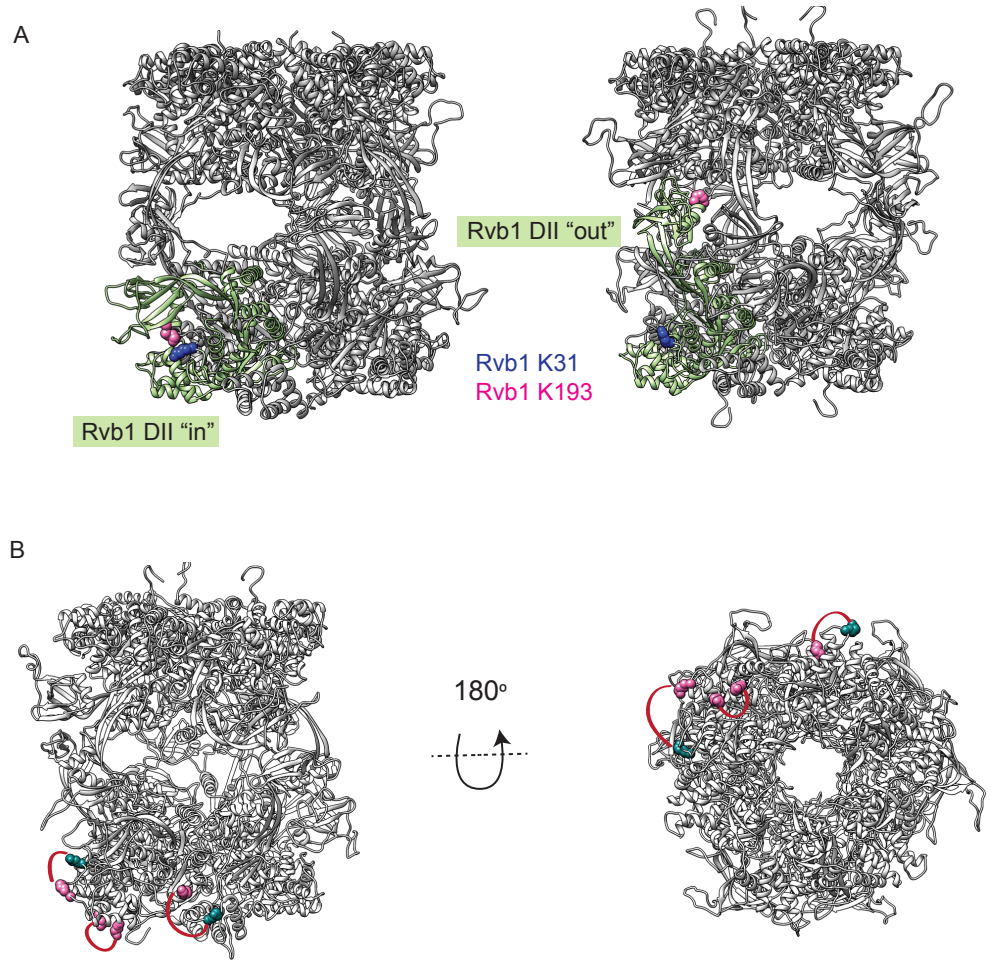
Appendix Figure 4: Comparative spectral counting analysis of Rvb1/Rvb2/Ino80INS reveals major changes in DII domains when Rvbs are bound to the Ino80INS



Difference map of spectral counts between Rvb1/Rvb2 crosslinks and Rvb1/Rvb2/Ino80INS crosslinks. The amino acid sequences of Rvb1 and Rvb2 are plotted on both x- and y-axes, while the amino acid sequence of MBP-Ino80INS is only plotted on the y-axis. Pink dotted lines delineate the boundaries between the three domains of Rvb1 and Rvb2 (DI, DII, DIII). Crosslinks observed are plotted as circles, with negative differences in spectral counts plotted in blue and positive differences in spectral counts plotted in pink. The size of each circle is scaled to the count difference between the two datasets, with the epicenter of the circle representing the location of the crosslink. All residues that crosslink to MBP-Ino80INS are plotted (pink circles), whereas only residues that have at least 5 counts in the Rvb1/Rvb2 dataset and with a

spectral count difference of more than 65% upon MBP-Ino80INS binding are plotted (blue circles). Clusters of crosslinks between Rvb1 and Rvb2 are grouped based on their proximity to residues that crosslink to INO80INS (pink letters, A and B). Group A residues are near other residues that crosslink to Ino80INS while Group B residues are not near any residues that crosslink to INO80INS.

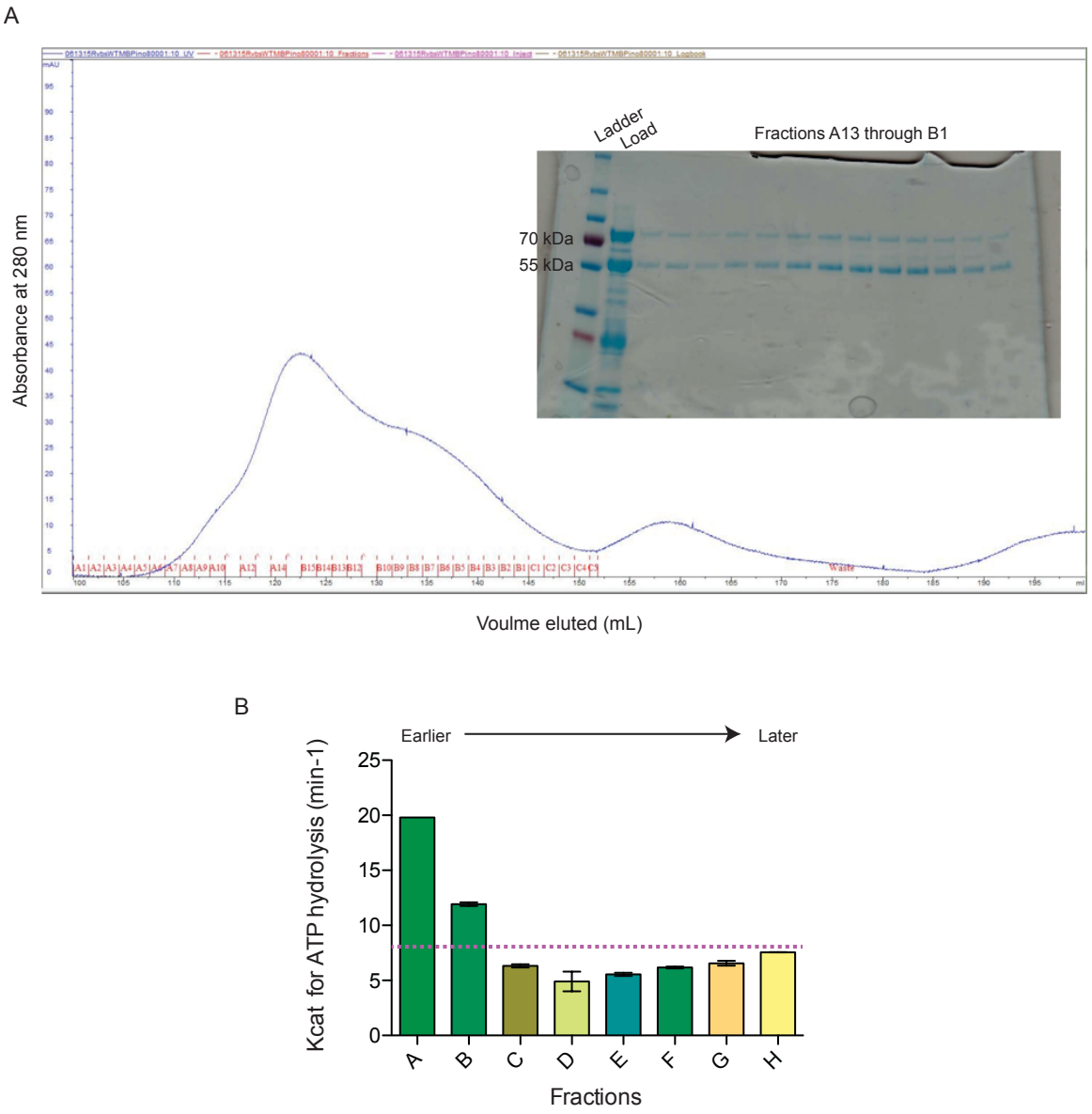
Appendix Figure 5: The Ino80INS pushes the DII domains of Rvb1/Rvb2 toward a conformation that is more compatible with dodecamerization



(A) Left, homology model of Rvb1/Rvb2 from *S. cerevisiae* structure based on published structure of Rvb1/Rvb2 from *C. thermophilum* where Rvb1 DII is "in" (9). A single Rvb1 protomer is show in green, with the rest of the structure in gray. Within the Rvb1 protomer, K31 is labeled as blue spheres and K193 is labeled is pink spheres. Right, homology model of Rvb1/Rvb2 where Rvb1 DII is "out". Same coloring is shown as in the left panel.

(B) Left, side view of Rvb1/Rvb2 showing crosslinked residues in Group B from the spectral counting difference map. Rvb1 residues are shown in pink and Rvb2 residues are show in teal, with red lines indicating the crosslinking that is present in the Rvb1/Rvb2 dataset but is largely diminished in the Rvb1/Rvb2/Ino80ID dataset. Right, top view of the same structure.

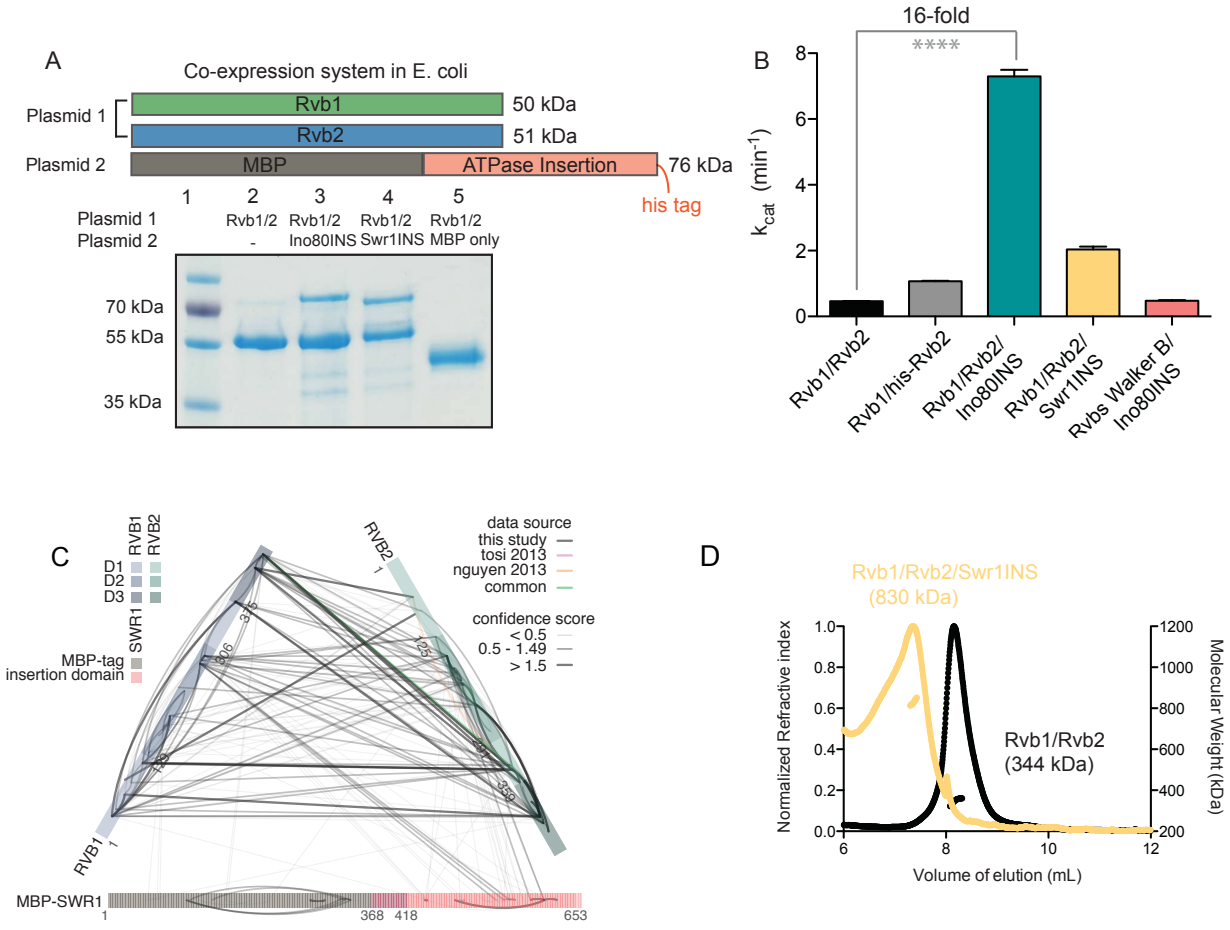
Appendix Figure 6: Fractionation of Rvb1/Rvb2/Ino80INS reveals a gradient of different specific activities



(A) Trace from a size exclusion run (S200) of Rvb1/Rvb2/Ino80INS. The large broad peak eluting between fractions A12-B2 corresponds to the elution of Rvb1/Rvb2/Ino80INS. Inset: SDS-PAGE gel of the fractions indicated on the chromatogram. These fractions were pooled separately (two fractions per pool), and labeled A → H.

(B) Activity assays of the fractions from (A), showing specific activity on the y-axis. Dotted line indicates the average specific activity measured for unfractionated Rvb1/Rvb2/Ino80INS.

Appendix Figure 7: The Swr1 insertion (Swr1INS) also interacts directly with Rvb1/Rvb2, enhancing its ATPase activity while promoting dodecamerization of the Rvbs



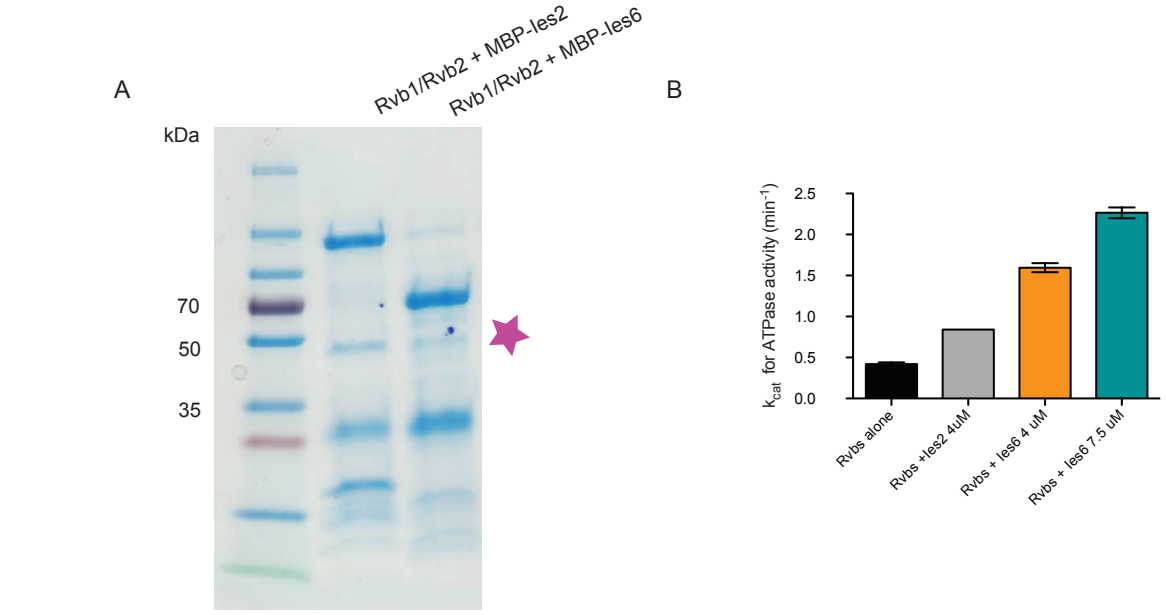
(A) Co-epxression of Rvb1/Rvb2 along with either MBP-Ino80INS (residues 1022-1294) or MBP-Swr1INS (residues 1003-1237) leads to a direct interaction among the three proteins as shown by a pull down assay.

(B) Comparison of maximal rates of ATP hydrolysis by different Rvb1/Rvb2 and Rvb1/Rvb2/ATPaseINS complexes. Bars represent the mean ± SEM of 3 replicates. **** = $p < 0.0001$.

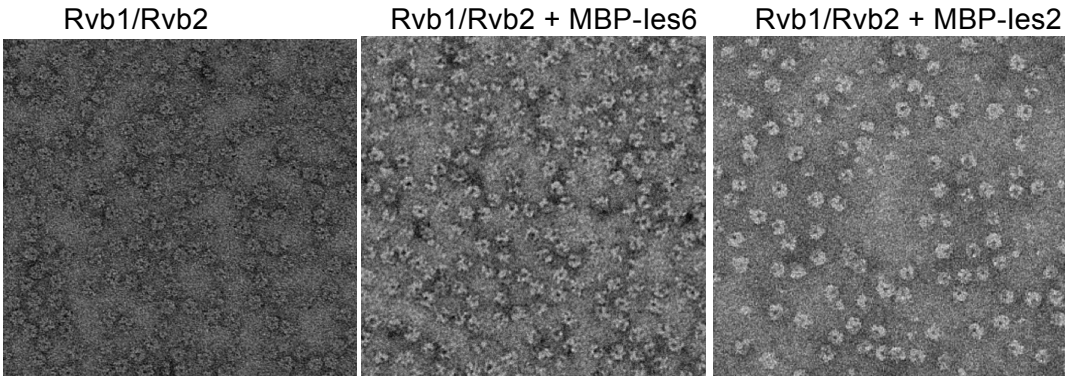
(C) XL-MS data with Rvb1/Rvb2/Swr1INS.

(D) SEC-MALS data comparing the molecular mass obtained from injecting 10 μM of Rvb1/Rvb2 or 12 μM of Rvb1/Rvb2/SwrINS. The molecular mass for Rvb1/Rvb2/SwrINS is consistent with the dodecamerization of Rvb1/Rvb2.

Appendix Figure 8: MBP-Ies2 and MBP-Ies6 do not co-purify with Rvb1/Rvb2, but do stimulate ATPase activity of Rvbs



C

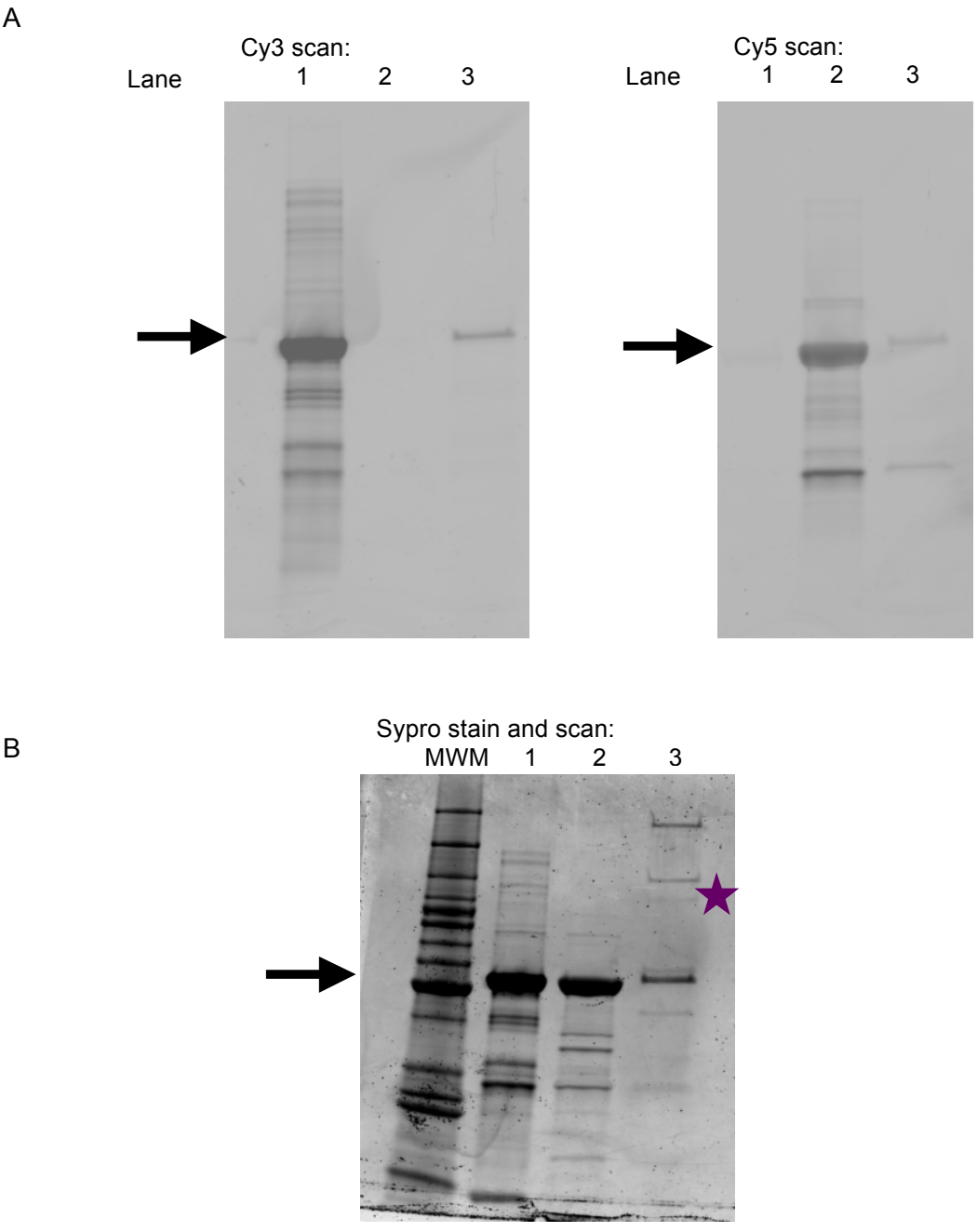


(A) SDS-PAGE gel showing the results of pull down with his-tagged MBP-Ies2 (120 kDa) or MBP-Ies6 (84 kDa) in the presence of co-expressed untagged Rvb1/Rvb2. Pink star indicates where Rvbs would be expected to run on the gel.

(B) Results of ATPase assays with either Rvbs alone or Rvbs in the presence of MBP-Ies2 or MBP-Ies6 added in trans. Concentration of Rvb1/Rvb2 was 1 μM.

(C) Negative stain EM images of Rvb1/Rvb2 alone (left), or with MBP-Ies2 (middle) or MBP-Ies6 (right) added.

Appendix Figure 9: Results of labeling and exchanging bacterially expressed Rvb1 and Rvb2 into the INO80 complex purified from yeast



(A) Cy3 scan of an SDS-PAGE gel assessing the incorporation of Rvb1-cy3 and Rvb2-cy5 into INO80 complexes purified from yeast. Lane 1: Rvb1-cy3, Lane 2: Rvb2-cy5, Lane 3-Ino80-flag

after exchange with the purified labeled Rvbs. Black arrows indicate where the Rvbs are expected to run on the gel.

(B) Same as (A), except a cy5 scan.

(C) Same as (A), except the SDS-PAGE gel was stained with Sypro Ruby and imaged acoordingly.

Chapter 3: The yeast INO80 complex operates as a tunable DNA length-sensitive switch for regulating nucleosome sliding

Abstract

The yeast INO80 chromatin remodeling complex plays essential roles in regulating DNA repair, replication and promoter architecture. INO80's role in these processes is likely related to its ability to slide nucleosomes, but the underlying mechanism is poorly understood. Here we use ensemble and single-molecule enzymology to study INO80-catalyzed nucleosome sliding. We find that INO80 exhibits a switch-like response to flanking DNA length, with nucleosome sliding increasing ~100-fold when DNA length is increased from 40 bp to 60 bp. Furthermore, once sliding is initiated, INO80 moves the nucleosome rapidly at least 20 bp without re-assessing flanking DNA length, and can change the direction of nucleosome sliding without dissociation. Finally, we identify a new auto-inhibitory role for the Nhp10 module of INO80, which controls INO80's switch-like response to flanking DNA. Our results indicate that INO80 is a highly processive remodeling motor that is tightly regulated by both substrate cues and non-catalytic subunits.

Introduction

The genetic material in cells is highly dynamic, with these dynamics gating and coordinating access to the genome to regulate all DNA-based processes. In eukaryotes, many of these dynamics originate at the basic organizational unit of chromatin, called the nucleosome, consisting of 147 base pairs (bp) of DNA wrapped around a core of eight histone proteins. ATP-dependent chromatin remodeling enzymes convert the chemical energy of ATP into mechanical forces that break and re-form histone-DNA contacts. Compounded by effects of other chromatin factors, these small-scale changes to nucleosome structure translate into large-scale changes to the chromatin landscape, which ultimately regulate every aspect of genome biology from transcription to DNA damage repair (1, 2). Of the four main families of remodelers (ISWI, SWI/SNF, CHD1, and INO80), the INO80 family is the most recently discovered and its biochemical mechanisms remain the most elusive (1, 3, 4).

The INO80 remodeler from *S. cerevisiae* is a multi-subunit complex consisting of a core remodeling ATPase (Ino80), related to the ATPase subunits of the other chromatin remodeling families, and 14 other subunits (3, 4). In vivo, INO80 has roles in regulating transcription, DNA damage repair, and replication, and in metabolic regulation (5). In vitro, INO80 has been shown to slide mononucleosomes to the center of short DNAs and to evenly space a tri-nucleosomal array (6). Two groups have also shown that INO80 can replace nucleosomes containing H2A.Z/H2B with canonical H2A/H2B in vitro (7, 8), but this finding is currently a subject of debate (9, 10).

A direct role for the nucleosome sliding activity of INO80 is suggested by studies on the nucleosome architecture of the yeast transcription start site (TSS). In yeast, nucleosome positioning at the TSS is well defined, with a nucleosome-free region (NFR) immediately preceding the TSS, followed by four evenly spaced nucleosomes (11, 12). Genome-wide studies have shown that ATP-dependent remodeling activity is required for establishing the nucleosome architecture at the TSS (5, 13). Using in vitro reconstituted yeast chromatin, it was recently shown that INO80 is the only remodeler that is sufficient for positioning the +1 nucleosome (14), consistent with earlier data showing that several INO80 subunits bind specifically to the +1 nucleosome, on the side proximal to the NFR (15, 16). In combination with the known biochemical activities of INO80, these data suggest that the nucleosome sliding activity of INO80 observed for single nucleosomes may play a critical role for positioning nucleosomes at the TSS, particularly the +1 nucleosome.

However, much is still unclear about the mechanism of nucleosome sliding by INO80, including how substrate cues are “read” by the enzyme and which motifs in the enzyme itself regulate this activity. Here we describe a combined approach using a variety of ensemble nucleosome remodeling assays as well as single-molecule FRET to probe the underlying mechanism of how INO80 uses flanking DNA length as a substrate cue for nucleosome movement. We find that INO80 exhibits a switch-like response to flanking DNA length, and that

the regulation of nucleosome sliding by flanking DNA length in INO80 is not directly coupled to ATPase activity as in other remodeler families. In addition, INO80 is capable of sliding a nucleosome long distances continuously and bi-directionally. We also describe a new auto-inhibitory activity of INO80 that is specific for preventing sliding of nucleosomes with short flanking DNA. Taken together, our results suggest that INO80 has a distinct mechanism from other remodeling families that likely has implications for INO80's distinct roles in vivo at the TSS and in double-strand break repair.

Methods

Expression and purification of INO80 complexes from yeast

INO80 was purified by FLAG immunoprecipitation based on previously published methods (17). Briefly, *S. cerevisae* with endogenously FLAG-tagged INO80 (17) was grown in YEPD at 30ºC to saturation. Cells were pelleted by centrifugation for 10 min at 5000 rpm, resuspended with buffer H0.3 (25 mM HEPES, pH 7.5, 1 mM EDTA, pH 8.0, 10% glycerol, 0.02% NP-40, 0.3 M KCl), and pelleted again. Pelleted cells were then extruded through a 60 mL syringe into liquid nitrogen to create "noodles". Cell "noodles" were then lysed using a Tissue Lyser II (Qiagen, Hilden Germany) or a freezer mill (SPEX 6970 EFM), cooled in liquid nitrogen. Frozen lysate powder was resuspended in equal volume of H0.3 and spun in an SW28 rotor for 2 hr at 25,000 rpm at 4ºC. Clarified lysate was mixed with equal volume buffer H0.3 and applied to FLAG M2-affinity resin (1 mL bead slurry per 40 mL of cleared lysate) equilibrated with H0.3 and incubated for 3 hours at 4ºC. An additional dose of protease inhibitors was added halfway through the incubation. Resin was washed with 3x50 mL buffer H0.5 (H0.3 buffer except with 0.5 M KCl) followed by 3x10 mL washes with buffer H0.1 (0.1 M KCl) and eluted with H0.1 supplemented with 1mg/mL FLAG peptide. Eluate was concentrated, aliquoted, flash frozen in liquid nitrogen, and stored at -80ºC. INO80 concentration was determined by SDS-PAGE with BSA standards, based on the intensity of the Ino80-FLAG band.

To generate the INO80(*Δ2-200-ino80*) construct, amino acids 2-200 of Ino80 were deleted by knock-in at the endogenous locus in the Ino80-FLAG strain, using a NAT marker -700 bp upstream of the ORF. The mutation was verified by colony PCR and by sequencing. INO80(*Δnhp10*) was made similarly except that a KanMX marker was knocked into the endogenous *nhp10* locus. Mutant complexes were purified as described above for WT INO80.

Nucleosome labeling and reconstitution

Recombinant *Xenopus laevis* histones were expressed and purified from *E. coli* as previously described (18). Histone octamer was reconstituted as previously described (19). FRET-labeled nucleosomes, with the donor Cy3 (for smFRET) or the acceptor Cy5 (for ensemble) on histone H3, were generated via a cysteine introduced at position 33, and were labeled prior to histone octamer assembly via cysteine-maleimide chemistry. Octamer for smFRET was assembled using a 2:1 unlabeled:labeled H3 mixture, to generate nucleosomes with mostly one H3 or neither H3 labeled; labeled octamer for ensemble FRET was assembled with all labeled histone H3. Nucleosomes for REA were not labeled on the histones. Cy3-labeled (for ensemble assays) and Cyanine 5 SE-labeled and biotinylated DNAs (for smFRET) were generated by PCR with HPLC-purified, labeled primers (Cy5 end-labeled primers: TriLink Biotechnologies, San Diego, CA; Cy5 internally labeled primers for centered constructs, IBA Life Sciences, Göttingen, Germany; Cy3 and biotinylated primers: IDT, Coralville, IA) and purified by PAGE. In most cases, nucleosomes were assembled using the 601 nucleosome positioning (20). Where indicated, the naturally occurring 5S sequence from *X. borealis*, with a different arbitrary sequence in the flanking DNA, was used instead (21). These DNAs were assembled with either labeled or unlabeled octamers by salt gradient dialysis, purified by glycerol gradient centrifugation, and quantified by native gel (19). DNA sequences used in this work are given in Figure S7.

Native gel remodeling assay

All gel remodeling reactions were performed under single turnover conditions (enzyme in excess of nucleosomes). Reactions were performed at 30ºC with 10 nM FRET-labeled nucleosomes, saturating enzyme (15 nM for WT INO80, 30 nM for INO80(*Δnhp10*) and INO80(*Δ2-200-ino80*)), 40 mM Tris pH 7.5, 60 mM KCl, 2 mM ATP•$MgCl_2$, 1.1 mM $MgCl_2$, 0.02% NP40, 1%(v/v) glycerol, and 0.5 mg/mL FLAG peptide. Reactions were assembled without ATP-Mg^{2+} and incubated at 30 ºC for 10 minutes prior to the addition of ATP. At the times indicated, a small portion of the reaction was removed and quenched with excess ADP and plasmid DNA. Time points were then resolved by native PAGE (6% acrylamide, 0.5XTBE) and scanned on a Typhoon variable mode imager (GE Life Sciences, Pittsburgh, PA) by scanning for Cy3. Gels were then quantified by densitometry using ImageJ.

ATPase assay

ATPase reactions were performed under multiple turnover conditions (ATP in excess of enzyme). Reactions were performed at 30ºC with 5 nM nucleosomes, 15 nM INO80, 40 mM Tris pH 7.5, 50 mM KCl, 100 mM ATP•$MgCl_2$, 0.5 mM $MgCl_2$, and trace amounts of γ-^{32}P-ATP. Reactions were assembled without ATP-Mg^{2+} and incubated at 30 ºC for 10 minutes prior to the addition of ATP. 2.5 µL time points were quenched with an equal volume of 50 mM Tris pH 7.5, 3% SDS, and 100 mM EDTA. Inorganic phosphate was resolved from ATP on a PEI-cellulose TLC plate (Select Scientific) with 0.5 M LiCl/1M formic acid mobile phase. Plates were dried, exposed to a phosphorscreen overnight, and scanned on a Typhoon variable mode imager. Rate constants were determined by fitting a line through the first 10% of inorganic phosphate generated using Prism.

Restriction enzyme accessibility (REA) assays

All REA reactions were performed under single turnover conditions (enzyme in excess of nucleosomes) and at 30ºC. Final conditions were: saturating enzyme (15 nM for WT INO80,

30 nM for INO80(*Δnhp10*)), 10 nM nucleosomes, 40 mM Tris pH 7.5, 60 mM KCl, 2 mM ATP•$MgCl_2$, 5 mM $MgCl_2$, 0.02% NP40, 0.5 mg/mL FLAG peptide, and 3 U/uL Pst1 (NEB). Reactions were assembled without ATP-Mg and incubated at 30ºC for 10 minutes before addition of ATP. Time points were quenched with an equal volume of stop solution (70 mM EDTA, 20 mM Tris, pH 7.5-7.7, 2% SDS, 20% glycerol, 0.2 mg/ml xylene cyanole and 0.2 bromophenol blue). After all time points were completed, Proteinase K was added to each sample. Time points were then resolved by native PAGE (6% acrylamide, 0.5XTBE) and scanned on a Typhoon variable mode imager (GE Life Sciences, Pittsburgh, PA) by scanning for Cy3. Fraction of DNA cut was quantified by densitometry using ImageJ. The data were fit to a single exponential decay using Prism 6 (GraphPad, La Jolla, CA) (Equation 1),

$$y=(y0-p)e-kobs\,t+p \quad (1)$$

where y_0 is the initial fraction un-cut, k_{obs} is the observed rate constant (min^{-1}), and p is the fraction DNA un-cut at plateau.

Ensemble FRET remodeling assay

Ensemble FRET remodeling assays were performed under the same conditions as gel remodeling assays, except that 5 nM nucleosomes were used. Reactions were initiated by addition of ATP and then measured in a K2 fluorometer (ISS) equipped with a 550 nm short pass excitation filter and a 535 nm long pass emission filter. Reactions were excited at 515 nm and emission was measured at 665 nm. The resulting curves were fit to a two-phase exponential decay (Equation 2),

$$y=p+(y0-p)(\,ffast\,e-kfast\,t+(1-\,ffast)\,e-kslow\,t)) \quad (2)$$

where f_{fast} is the fraction in the fast phase and k_{fast} and k_{slow} are the remodeling rates of the fast and slow phase respectively. All rate constants reported for ensemble FRET assays are k_{fast}, with the exception of the ensemble FRET rate constants reported in Figures S3 and S4. These data were obtained at 20ºC rather than 30ºC, and so these slower data were fit to a single exponential decay, with the reported rate constant being the decay rate of the single exponential.

Results

Flanking DNA modulates the overall remodeling rate of INO80, but not its ATPase activity

Previous work has shown that yeast INO80 can slide mononucleosomes toward the center of DNA, and evenly space a tri-nucleosomal array, but both activities require at least 50 base pairs (bp) of extranucleosomal DNA (6). We used the strong nucleosome positioning sequence 601 to construct initially end-positioned nucleosomes with varying amounts of extranucleosomal DNA, and found that by native gel remodeling, efficient nucleosome sliding requires more than at least 40 bp of flanking DNA (Figure 1A), in agreement with this previous work (6). We also noticed that while nucleosomes with 60 bp of flanking DNA ("0/60") are moved to a more centered position that runs as a single discrete band on a native gel, nucleosomes with 80 bp of flanking DNA ("0/80") are slid toward the center, but result in a final distribution of products that runs as a doublet on a native gel (Figure 1A). We hypothesized that this distribution of products could represent sequence-specific variations in the strength of the histone-DNA contacts near the center position of the 0/80 construct's DNA. To test this possibility, we repeated the experiment using the naturally occurring 5S sequence, which is 100-fold weaker at positioning nucleosomes (22), and which also contains a different 80 bp flanking sequence outside of the positioning sequence. We found that with 0/80 nucleosomes positioned by 5S, the final distribution of products by native gel becomes more like that of the 0/60 nucleosomes

positioned by 601, with a discrete centered product band (Figure S1). However, the overall remodeling rate constant remains unchanged, as measured by an ensemble FRET assay (Figure S1). These results suggest that DNA sequence can play a role in remodeling by INO80 by affecting the final product distribution but not the rate, similar to what was reported for ISWI remodelers (22) and to what has been reported for INO80 in a reconstituted yeast chromatin system (14).

We next used a FRET-based remodeling assay to quantify the maximal rate constant (k_{max}) as a function of flanking DNA length. We found that varying the flanking DNA from 40 bp to 80 bp increases the k_{max} for remodeling by 300-fold (Figure 1B-C). However, the largest increase (~100-fold) occurs with an increase in flanking DNA length from 40 bp to 60 bp. These results suggest a switch-like response between 40 bp and 60 bp of flanking DNA. As these experiments were performed under saturating concentrations of INO80, the differences in rate constant arise from differences in a step that is post-binding.

These observations of a modulation of nucleosome sliding with flanking DNA length are reminiscent of the well-studied ISWI remodeling enzymes, which move nucleosomes faster toward longer flanking DNA by coupling length sensing to the ATPase activity of the motor (23). To test whether this paradigm holds true for the INO80 family of motors, we next measured the maximal ATPase activity of INO80 when stimulated by nucleosomes with either 40 bp or 80 bp of flanking DNA, using the same INO80 and nucleosome concentrations as in the FRET-based assay. Surprisingly, both nucleosomal substrates stimulate the ATPase activity of INO80 equally (Figure 1D; data not shown). This result is in contrast to what is known for ISWI-family motors, and raises the possibility that the INO80 family of motors uses a different mechanism to sense flanking DNA length and couple this sensitivity to remodeler output. These results also indicate that the substantially slower rate of sliding with 0/40 nucleosomes is not due to an inability of INO80 to bind these nucleosomes, because the rate of ATP hydrolysis with 0/40 nucleosomes is the same as with 0/80 nucleosomes.

This observation also raises the possibility that the ATPase activity of INO80 is used to generate another type of remodeled intermediate that is not detected by native gel or FRET. To test this hypothesis, we used a restriction enzyme accessibility (REA) assay to measure rates of remodeling by INO80 (Figure 1E) (24). In principle, the REA assay can report on changes to the accessibility of the nucleosomal DNA arising from changes that do not require nucleosome sliding, such as DNA unpeeling or a conformational change in the nucleosome. By REA, we found that nucleosomes with 40 bp of flanking DNA show an increase in DNA accessibility at a rate that is 10-fold faster than the observed rate of FRET remodeling (0.05 ± 0.02 min^{-1} vs 0.005 ± 0.006 min^{-1}, respectively) (Figure 1F,G). No change in accessibility was observed in the absence of ATP (data not shown). This observation suggests that INO80 can alter histone-DNA contacts of a nucleosome with shorter flanking DNA, but cannot efficiently convert this altered nucleosome into a slid product unless the nucleosome has greater than 40 bp of flanking DNA. Interestingly, the 300-fold difference in k_{max} between nucleosomes with 40 bp vs. 80 bp of flanking DNA measured by FRET decreases to 12-fold when measured by REA. This result suggests that nucleosome sliding in INO80 is much more sensitive to flanking DNA length than the generation of the restriction enzyme accessible intermediate. However, the rate of REA is still sensitive to flanking DNA length, which does not explain our result that ATPase activity is insensitive to flanking DNA length. The possible origins of this discrepancy are addressed in the Discussion.

Remodeling by INO80 is dominated by a long initial pause, followed by rapid nucleosome mobilization

To explore how ATPase rates and remodeling rates have different relationships with flanking DNA length in INO80, we turned to a single-molecule assay to identify the steps in the INO80 remodeling reaction and to determine which step(s) are sensitive to DNA length and/or ATP concentration. Single-molecule FRET (smFRET) has proven to be a powerful method for

dissecting the remodeling reactions of several other families, notably the ISWI family (25). We therefore adapted the smFRET assay to the study of remodeling by INO80 (Figure 2A).

In order to put our smFRET results in the broader context of previously described chromatin remodelers, we first make some general comparisons between remodeler behaviors observed at the single-nucleosome level. The SWI/SNF family member RSC, which slides nucleosomes towards DNA ends but is not sensitive to flanking DNA length, has been shown to mobilize single nucleosomes quickly, with no reaction steps other than nucleosome translocation having been identified (25). In contrast, ISWI family remodelers exhibit a stereotyped behavior when observed at the single nucleosome level: remodeling proceeds through an alternating series of “pauses”, during which no change in nucleosome position is detected, and translocation events, during which the nucleosome is slid (26-28). As noted above, most ISWI remodelers are sensitive to flanking DNA length. It was shown by smFRET that both DNA length sensing and the detection of another ISWI substrate cue, the histone H4 tail, occur during the pauses (27, 28). Moreover, ATP is hydrolyzed during these pauses (26-28), suggesting that ATP hydrolysis is used during the pauses to couple flanking DNA length sensing to the regulation of sliding activity. An emerging model is therefore that the pause phases in the ISWI reaction represent a rate-limiting regulatory step, during which substrate cues are assessed and used to gate translocation activity (1). In this model, the interruption of translocation by pauses allows for a periodic re-assessment of the flanking DNA length.

ISWI and SWI/SNF thus represent two ends of a spectrum when viewed at the single-nucleosome level, from highly regulated, periodic translocation events gated by substrate-cue-sensitive pause phases, to rapid translocation that is apparently not interrupted by obvious regulatory events. Given that INO80 represents a different family of remodelers that shares length-sensing with ISWI, but large, multi-subunit complex size with SWI/SNF, we investigated if and how INO80 differs from ISWI and SWI/SNF at the single-nucleosome level.

Figure 2B shows several examples of smFRET measurements of INO80 remodeling a single end-positioned nucleosome with 78 base pairs of flanking DNA on one side. As expected, the overall FRET efficiency decreases over time, consistent with INO80 sliding the nucleosome away from the DNA end, towards the 78 bp flanking DNA (see also Figure 1B). Although we cannot rule out the formal possibility that part or all of the observed change in FRET is due to a process other than sliding–for example, unpeeling of the DNA away from the octamer surface, which would also result in a change in inter-dye distance–we note that the behavior shown in Figure 2B is ATP-dependent, with changes in FRET that are not instantaneous (Figure S3, Figure S4E).

At a simple level, the data in Figure 2B suggest that INO80 remodeling represents an intermediate between ISWI- and SWI/SNF-style remodeling. On the one hand, translocation is preceded by an extremely long pause, during which the FRET does not change (and the nucleosome is not being slid). Translocation is often (about 75% of the time; Figure S4B,C) interrupted by at least one additional pause that appears to occur at roughly half-maximal FRET (Figure 2C and Figure S4B,C). These pauses are candidates for regulatory steps as in the ISWI paradigm and are discussed further below. On the other hand, translocation is fast, even at 20°C (~7 bp/s at saturating ATP; Figure S4E), comparable to the translocation rate of ACF at 30°C (~7-9 bp/s (27)). Furthermore, INO80 moves the nucleosome at least 10 bp before being interrupted by a pause (Figure 2C and Figure S4B,C), in contrast to ISWI remodelers, which move the nucleosome at most 7 bp, and more often 3 bp, between pauses (26-28).

The initial pause is ATP dependent, but not DNA length sensitive

We next examined whether the initial and secondary pauses, and/or the translocation phases, are sensitive to substrate cues such as DNA length. Previous work with ISWI complexes showed that increasing flanking DNA length decreased the durations of the pauses, but did not affect translocation rates or the distance the nucleosome was moved during the translocation

phases, implying that flanking DNA length is sensed in the pause phases. Furthermore, a reduction in ATP concentration was shown to result in longer pauses and slower translocation for ISWI complexes, implying that ATP is required for both the pause and translocation phases (26-28).

We performed similar experiments with INO80 to ask which step(s) in its remodeling reaction are sensitive to DNA length or ATP concentration. As with ISWI remodelers, the translocation phases of INO80 are sensitive to ATP concentration and not to flanking DNA length (Figure S4E). The initial pause ($p_{initial}$) is sensitive to ATP concentration, but, surprisingly, is not sensitive to flanking DNA length (Figure 2D,E). The ATP dependence of $p_{initial}$ quantitatively recapitulates the ATP dependence of the overall remodeling rate constant measured by ensemble FRET, suggesting that $p_{initial}$ contains the main ATP-dependent step observed in ensemble assays (Figure S4D).

Even more surprisingly, the secondary pauses are not sensitive to either flanking DNA length or ATP concentration (Figure 2D,E). By ensemble FRET, 0/60 nucleosomes are remodeled 3-fold more slowly than 0/80 nucleosomes (Figure 1C). If either $p_{initial}$ or p_{second} were sensitive to flanking DNA length, we would have expected these pause durations to decrease with increasing flanking DNA. However, we do not observe a statistically significant trend in either $p_{initial}$ or p_{second}, much less a 3-fold decrease in either pause duration, as a function of increasing flanking DNA length from 60 to 80 bp. These results represent a departure from the ISWI paradigm, and suggest that the secondary pauses do not play a regulatory role.

The secondary pauses of INO80 also differ from ISWI in where they occur. The regulatory ISWI pauses exhibit a distinctly non-random pattern of an initial large, ~7 bp movement of the nucleosome during the first translocation event, followed by smaller (~3 bp) movements (26-28). This 7-3-3 pattern is independent of flanking DNA length as well as DNA sequence. In contrast, the secondary pauses in the INO80 reaction occur predominantly after 10 bp of movement, or less commonly after 5 or 15 bp of movement (Figure 2C, Figure S4B,C).

These values are intriguingly close to the rotational periodicity of nucleosomal DNA (29), suggesting that INO80 might pause during translocation not as a regulatory aspect of its mechanism, but at sites where the DNA forms favorable contacts with the histone core {Hall:2009er}. Such behavior would be similar to RNA Polymerase II pausing at points of strong DNA-histone interactions while translocating through a nucleosome (30, 31), as well as with the results described above, in which a sequence-specific difference in product distribution is observed by native gel (Figure S1). To further test this hypothesis, we quantified the durations of $p_{initial}$ and p_{second} and the locations at which p_{second} occurs with a 3/78 construct containing the 5S positioning sequence instead of the 601 sequence. Although the pause durations with 5S-positioned nucleosomes are within error of those with 601-positioned nucleosomes, the distribution of locations at which p_{second} occurs broadens (Figure S5). This result is consistent with the 5S positioning sequence, as a weaker positioning sequence than 601, having less of a preference for particular rotational conformations of the octamer on the DNA (22). These results are in contrast to what has been observed with ACF, in which sequence does not affect the locations of the pauses (27). Thus we favor an interpretation of our results in which the secondary pauses of the INO80 remodeling reaction are not regulatory, and the INO80 remodeling reaction is not sensitive to flanking DNA length at least within the first ~20 bp of movement.

These results have new implications for the mechanism of remodeling by INO80. First, it appears that once translocation has begun, it is not interrupted by regular pauses that re-interrogate flanking DNA length after the nucleosomes has been moved a short distance, at least not within the 20-25 bp that the nucleosome is in FRET range (Figure S2). We note that for the 3/60 nucleosomes in particular, this means that the nucleosome is moved almost all the way to the center of the DNA without a length-sensitive interruption of translocation. It is interesting that the ISWI-family remodeler ACF, which has a similar length-sensing regime to

INO80 in terms of the modulation of its overall remodeling rate, has about three length-sensitive pauses within the first 20 bp of nucleosome movement, whereas INO80 has none.

Second, our result that $p_{initial}$, which represents the slowest step, is sensitive to ATP but not to flanking DNA length helps explain the results of Figure 1, where we see that ensemble ATPase activity is not sensitive to flanking DNA length, while the overall remodeling rate by ensemble FRET is. A comparison between our ensemble and single molecule FRET results also implies that flanking DNA length sensitivity arises at a step that is not observed in these smFRET experiments. The possible origins of INO80's length sensitivity are addressed in the Discussion.

INO80 processively and continuously moves a centered nucleosome significant distances back and forth along the DNA

The results of Figure 2 are also suggestive of another important feature of INO80's mechanism: processivity, which we define as a measure of how many base-pairs the nucleosome can be translocated before dissociation of INO80. In Figure 2, INO80 was introduced into the sample chamber and allowed to bind end-positioned, surface-attached nucleosomes, before the addition of the ATP that started the remodeling reaction. The addition of ATP was performed in a buffer exchange step that chased off unbound INO80. Under these conditions, INO80 is capable of sliding end-positioned nucleosomes at least 20 bp without dissociation.

A better test for processivity, however, would be under conditions in which the nucleosome remains within FRET range during a longer portion of the remodeling reaction. Such an experiment would also allow us to interrogate the ability of the motor to switch directions during translocation. Such conditions have been obtained for the ISWI remodeler ACF by using a nucleosome that starts in the center of a long piece of DNA. ACF has been shown to continuously (and processively) move such a nucleosome back and forth around the center position, keeping the nucleosome within FRET range for a large part of the reaction (27).

It is unclear whether INO80 possesses similar bi-directional nucleosome sliding activity: no nucleosome sliding was observed for yeast INO80 with a 70/70 nucleosome (6), but recent work with human INO80 suggests it can mobilize a centered nucleosome with 70 bp flanking DNAs, moving it slightly off-center (32).

To test for processivity, we generated “78/78” nucleosomes, with the 601 positioning sequence flanked on both sides by 78 bp. These DNAs also contain an internal Cy5 dye 9 bp from the edge of the 601 sequence on the flanking DNA, opposite the side fixed to the surface of the microscope slide (Figure 3A). As in Figure 2, nucleosomes labeled on the H3 proximal to this internal Cy5 were selected for further analysis, resulting in a population of nucleosomes that start in mid-high FRET, consistent with our FRET-to-bp calibration curve (Figure S2). If INO80 slides the nucleosome towards the surface, the FRET will decrease according to our calibration curve; if the nucleosome is moved towards the dye, the FRET will increase, and then decrease as the dye-labeled base-pair enters the nucleosome. We note that once the dye has moved into the nucleosome, the relationship between FRET and bp will not follow our calibration curve.

We find that INO80 does indeed move an initially centered nucleosome, quickly sliding it out of FRET range as with the end-positioned constructs (Figure 3B). INO80 is also capable of sliding a nucleosome back and forth along the DNA without releasing it, as indicated by the alternating gains and losses of FRET in the bi-directional and multi-directional remodeling behaviors we observe (Figure 3B). In combination with the data presented in Figure 2, these results indicate that INO80 can catalyze many rounds of translocation without dissociating. We note that the percentage of trajectories showing direction reversals with these initially centered constructs is double the percentage with end-positioned constructs (40% for the 78/78 construct, versus 20% of all end-positioned constructs in Figure 2).

To confirm that the bi-directional, processive behavior we observe is due to the same INO80 complex that initially binds the nucleosome, we performed “double chase” experiments, in which a second buffer exchange into 1 mM ATP was performed roughly a minute after the

addition of ATP that started the remodeling reaction (Figure 3B, bottom panels). As a control, we also chased with buffer only, after which remodeling stops, indicating that we are fully exchanging the buffer in the sample chamber during these washes (Figure 3C). Thus INO80 is capable of moving a nucleosome long distances bi-directionally along a DNA strand without dissociating. Although it is currently not clear how INO80 switches its direction of translocation under chase conditions, one possibility is that it can bind a nucleosome as a dimer. Cooperative binding and sliding of nucleosomes has recently been shown for a human INO80 core complex, as has the potential of two human INO80 core complexes to bind a single nucleosome (32).

Deletion of nhp10 results in impaired length sensing by INO80

We next investigated what motif(s) in INO80 might be involved in regulating the switch-like increase in nucleosome sliding upon increasing flanking DNA length beyond 40 bp. Structural studies of both yeast INO80 (33) and human INO80 (34) have shown that the complex is organized into different modules, each composed of 3-4 subunits that make distinct contributions to nucleosome remodeling. For example, the Arp8 module, composed of Arp8, Arp4, Act1, Taf14 and Ies4, is marginally involved in nucleosome and DNA binding, but is critical for the catalysis of nucleosome remodeling. In contrast, the yeast-specific Nhp10 module, composed of Nhp10, Ies1, Ies3 and Ies5, is not critical for nucleosome movement or ATP hydrolysis but exhibits strong DNA and nucleosome binding properties on its own (33, 35). Based on this observation, we hypothesized that the Nhp10 module may contribute to INO80's ability to discriminate between nucleosome substrates based on flanking DNA length.

To test this possibility, we purified INO80 complexes from yeast containing a deletion of *nhp10*. As seen previously, when compared to wild-type (WT) INO80 on an SDS-PAGE gel, the INO80(*Δnhp10*) complex shows the absence of bands corresponding to components of the Nhp10 module, such as Ies3 (purple star, Figure 4A) (33, 36). We also noticed that in our preps of INO80(*Δnhp10*), the majority of the Ino80 band ran consistently lower than that observed for

WT INO80 (green star, Figure 4A). We reasoned that this lower band may correspond to a truncation of Ino80 that occurs as a result of deleting *nhp10*. Based on mass spectrometry data, we indeed found that the first ~200 amino acids of the N-terminus of Ino80 is missing in INO80(*Δnhp10*) but remains intact in WT INO80 (Figure S6). Previous work has shown that the Nhp10 module interacts with these first 200 amino acids of Ino80 (33), suggesting that deletion of the *nhp10* causes both the loss of the Nhp10 module from the INO80 complex as well as the destabilization and degradation of the N-terminal region of the Ino80 ATPase. To further test this possibility, we purified INO80 complexes from yeast containing an N-terminal deletion of Ino80. The INO80(*Δ2-200-ino80*) complexes also lacked the same subunits of the Nhp10 module (Figure 4A), further supporting the idea that the N-terminus of Ino80 and the Nhp10 module are structurally inter-dependent, and perhaps even co-fold.

We next compared the remodeling rates of these different mutant INO80 complexes on nucleosomes with either 40 or 80 bp flanking DNA, using ensemble FRET. We found that neither INO80(*Δ2-200-ino80*) nor INO80(*Δnhp10*) has a defect in remodeling nucleosomes with 80 bp flanking DNA, consistent with published data (33). However, to our surprise, we observed a striking 100-fold increase in remodeling activity on nucleosomes with 40 bp flanking DNA with the mutant complexes compared to WT (Figure 4B). Similarly, by native gel, we observe the efficient generation of a more centered product with 0/40 nucleosomes and INO80(*Δ2-200-ino80*) or INO80(*Δnhp10*), suggesting that the low-FRET product formed by INO80(*Δ2-200-ino80*) and INO80(*Δnhp10*) is indeed a slid nucleosome (Figure 4C,D). Taken together, these results suggest that either the Nhp10 module or the N-terminus of Ino80 is auto-inhibitory for remodeling nucleosomes with short flanking DNA.

As noted above, a comparison between remodeling rates as measured by FRET versus REA with WT INO80 suggests a 25-fold difference in flanking length discrimination, leading to the hypothesis that nucleosome sliding is more sensitive to flanking DNA length compared to the creation of a nucleosome intermediate that can be cut by a restriction enzyme. To further

test this hypothesis, we next measured remodeling rate constants by REA for INO80(*Δnhp10*) on 0/40 versus 0/80 nucleosomes. Surprisingly, we found no difference in remodeling rates by REA between WT and INO80(*Δnhp10*), for either nucleosome substrate (Figure 4E-G). This result stands in contrast to the 100-fold difference observed by FRET for nucleosome sliding with 40 bp of flanking DNA, and is consistent with our hypothesis that generation of the REA-sensitive intermediate and sliding happen in distinct kinetic steps. These data also raise the possibility that the inhibition of sliding of nucleosomes with ≤40 bp of flanking DNA occurs after the generation of the REA intermediate.

Discussion

Compared to chromatin remodeling complexes from other families such as ISWI and SWI/SNF, which were discovered earlier, the biophysical mechanism of INO80 is less well understood. Our study, which uses a complementary set of single molecule and ensemble assays, indicates that while INO80 shares some remodeling behaviors with the ISWI and SWI/SNF families, it displays substantial mechanistic differences in how it slides nucleosomes. Nucleosome sliding by INO80 has been suggested to play a part in its critical roles in positioning nucleosomes at the transcription start site (TSS) (14) and in facilitating double strand break (DSB) repair (3). While flanking DNA has been identified as a key substrate cue in the sliding reaction, the mechanism behind coupling of this cue to nucleosome movement has been unclear.

Our ensemble and single-molecule results provide a starting point for building a kinetic model for how INO80 preferentially slides nucleosomes with longer flanking DNA lengths. Toward assembling this model, we first summarize three key findings: (*i*) by ensemble FRET, increasing the flanking DNA length from 40 bp to 80 bp results in a switch-like increase in remodeling rates, with the greatest increase (~100 fold) occurring between 40 bp and 60 bp, and a smaller (~3-fold) but reproducible increase occurring between 60 bp and 80 bp (Figure

1C); (*ii*) by a restriction enzyme accessibility assay, nucleosomes with 40 bp flanking DNA are remodeled 10-fold faster than by ensemble FRET, and flanking DNA length regulates DNA accessibility by only 12-fold compared to the 300-fold dynamic range observed for sliding (Figure 1G), and (*iii*) flanking DNA length does not affect either the translocation rate measured by smFRET, or any pause duration within at least the first 20 bp that the nucleosome is moved (Figure 2D,E, Figure S4).

Based on these data, we propose the following model for how INO80 couples flanking DNA length sensing and nucleosome movement (Figure 5). After the addition of ATP, an intermediate that is detectable by REA is formed, a process to which we assign a rate constant k_{conf}. This intermediate can collapse ($k_{collapse}$) back to its original structure, or be converted to a sliding competent state (k_{slide}). We propose that the ratio of $k_{slide}/k_{collapse}$ is sensitive to flanking DNA length, such that it increases with increasing flanking DNA length. Thus nucleosomes with 40 bp flanking DNA will more often collapse back to the original structure than nucleosomes with 80 bp flanking DNA. An alternative interpretation of our data is that the REA-sensitive intermediate is not on-pathway for nucleosome sliding, but is an end product itself, or an intermediate for other products that could be observed under different assay conditions.

A prediction of our model is that the generation of an REA-sensitive nucleosome (k_{conf}) occurs independently of sliding (k_{slide}). Our data with INO80(*Δnhp10*) serve as a test of this prediction, since this mutant increases the sliding rate of nucleosomes with 40 bp of flanking DNA by 100-fold, but leaves the rate of cutting observed by REA on 0/40 nucleosomes the same as WT (Figure 4). This result most simply suggests that the Nhp10 module or the N-terminus of Ino80 regulates length-sensitive sliding by altering k_{slide}, but not k_{conf}. We note in particular that the 12-fold difference in REA as a function of flanking DNA is retained in the INO80(*Δnhp10*) mutant (Figure 4G), further supporting our model that Nhp10 does not regulate k_{conf}.

We propose that the flanking DNA length sensitivity contained in $k_{slide}/k_{collapse}$ represents the basal length sensitivity of INO80, and that the Nhp10 module adds a step-like response on top of this behavior between 40 bp and 60 bp. Interestingly, Nhp10 is not found in humans, but human INO80 does contain several metazoan-specific subunits that also bind to the N-terminus of the Ino80 ATPase (37). Indeed, a core complex of human INO80 lacking these subunits does not exhibit the switch-like behavior as a function of DNA length that we observe with yeast INO80 (32). Rather than a 100-fold increase in overall remodeling rate constant between 40 bp and 60 bp flanking DNA, followed by a smaller increase between 60 bp and 80 bp, the core human INO80 complex exhibits a gradual increase in remodeling rate between 20 bp and 60 bp flanking DNA. While the sensitivity to flanking DNA length has not yet been tested for the full human INO80 complex, this observation raises the possibility that there exist human INO80 subunits that play a similar role to that of the Nhp10 module in yeast.

In our model (Figure 5), k_{conf}, $k_{collapse}$, and k_{slide} are all contributors to the duration of the initial pause we observe by smFRET. However, $p_{initial}$ is ATP-dependent, but not length sensitive (Figure 2D,E). In contrast, the overall rate of remodeling as measured by ensemble FRET, which also encompasses the same set of rate constants, is length sensitive. Why then do we not see a dependence of $p_{initial}$ on flanking DNA length between 60 bp and 80 bp? We note that the overall rate of remodeling by ensemble FRET is slower (by about 2-3 fold) than what we observe by single molecule (Figure S3). We propose that, while $k_{slide}/k_{collapse}$ is length-sensitive, the products generated by recycling through the collapsed state are generated too slowly to be detected by smFRET before photobleaching occurs. Thus by smFRET we primarily observe the fraction of nucleosomes that are converted to a sliding competent state and go on to the translocation phase without recycling through the collapse pathway, decreasing the potential flanking DNA length sensitive contribution of $k_{collapse}$. However, the recycling process through $k_{collapse}$ does contribute to the overall remodeling rate observed by ensemble FRET, which therefore appears slower and also sensitive to flanking DNA length.

This model already represents a departure from the paradigm of the better-studied, length-sensitive ISWI-family remodelers. Our single molecule results also highlight a second significant departure from the ISWI length-sensing paradigm: the lack of a regulatory, substrate-cue-sensitive pause that interrupts translocation after the nucleosome has moved only a short distance (~3-7 bp). Such a re-interrogation of flanking DNA length after the nucleosome has moved only a short distance is thought to be essential to the ability of some ISWI enzymes to center a nucleosome, and to maintain a centered nucleosome near the center position (1). However, this behavior results in a remodeling reaction dominated by regulatory pauses that lead to only short-range nucleosome movements of 3-7 bp (26-28). In contrast, once translocation is initiated, INO80 moves a nucleosome quickly (~7 bp/s) by at least 20-25 bp without a re-assessment of flanking DNA length (Figure 2, Figure S4).

Despite our observation that INO80 does not pause within at least the first 20 bp to re-assess DNA length, native gel remodeling assays show that the population-averaged nucleosome position after the reaction has been stopped is close to the center of the DNA for 0/60 and 0/80 nucleosomes (Figure 1; (6)). Additionally, translocation occurs bi-directionally on long (78/78) constructs. What “stop cue” halts the translocation phase to enable the nucleosome to end, on average, away from DNA ends, and for INO80 to switch the direction of sliding? This stop cue could be in the substrate (e.g., a certain distance, ~40 bp, from a DNA end or another nucleosome), or in INO80 (e.g., a certain distance translocated, >20 bp). A distance-translocated stop cue, if it were to result in a pause during which DNA length were re-assessed, would make INO80 similar to ISWI, but with a larger step size (>20 bp rather than ~3-7 bp). However, the possibility of a substrate cue such as distance from another nucleosome providing the “stop”, combined with the ability of INO80 to move nucleosomes long distances quickly, provides an intriguing speculative mechanism for the recent observation that in a reconstituted system, INO80 was sufficient to position the +1 nucleosome at the NFR, but with the wrong spacing between the +1 and +2 nucleosomes (14). We speculate that in vivo, INO80 mobilizes

nucleosomes long distances, grouping them close to the TSS, where other true spacing enzymes (i.e. ISWI enzymes) and other factors can then make small adjustments to the spacing to obtain the native promoter organization, as was shown to occur in the reconstituted system (14). Similarly, upon recruitment to a double strand break, INO80 can very quickly move nucleosomes away from the site of damage, while other remodeling motors may fine-tune the nucleosome spacing to facilitate repair. A similar conceptual picture to our substrate-based stop cue has been proposed for human INO80, though with significant differences in the relationship between the "stop" cue and the length sensing that occurs prior to the initiation of translocation (32). We note that regardless of whether the putative "stop" cue is substrate-based or enzyme-based, Nhp10 does not appear to affect it, since INO80(*Δnhp10*) and INO80(*Δ2-200-ino80*) generate similar product distributions as WT INO80 on 0/80 nucleosomes (Figure 4D).

Auto-inhibitory motifs are being discovered in several chromatin remodeling enzymes. In one example, the chromodomains of CHD1 have been shown to inhibit ATPase activity, but have only small effects on remodeling rates (38). In another example, the AutoN motif of ISWI motors has been shown to generally auto-inhibit both ATPase and remodeling rates (39). Both of these mechanisms are distinct from that observed for the Nhp10 module, which only inhibits remodeling of nucleosomes with short flanking DNA (Figure 4), and has no effect on ATPase rates (33). More similar to Nhp10 is the NegC module of ISWI, which has been shown to regulate the remodeling activity of the human ISWI ATPase subunit SNF2h in a DNA-length dependent manner, with the strongest effect (10-fold) on nucleosomes without flanking DNA (40).

However, an important distinction between NegC and Nhp10 is that Nhp10 is a completely separate polypeptide from the remodeling ATPase, suggesting that the presence of this module is a potential source of regulation in vivo. Such differential regulation based on subunit composition has already been demonstrated for a positive regulator of INO80, the Arp5/Ies6 module, which has been shown for yeast INO80 to be required for both ATPase and

remodeling activity (5, 34). Interestingly, in vivo, Arp5/Ies6 differentially associates at promoters of INO80-regulated genes (5). The most highly transcribed genes also contain more Arp5, suggesting that different levels of Arp5/Ies6 may have different effects on nucleosome architecture at the TSS. Our data with *Δnhp10* suggest that the Nhp10 module may act analogously, perhaps providing a means to regulate nucleosome sliding at sites of DNA damage. Our study, in combination with previous work on INO80, demonstrates how remodeling activity can be tuned by subunit composition. We speculate that such tunability is important to meet the dynamic chromatin remodeling needs of different genomic contexts.

References

1. Zhou, C. Y., Johnson, S. L., Gamarra, N. I., and Narlikar, G. J. (2016) Mechanisms of ATP-Dependent Chromatin Remodeling Motors. *Annu Rev Biophys* 45, 153–181

2. Clapier, C. R., and Cairns, B. R. (2009) The biology of chromatin remodeling complexes. *Annu. Rev. Biochem.* 78, 273–304

3. Morrison, A. J., and Shen, X. (2009) Chromatin remodelling beyond transcription: the INO80 and SWR1 complexes. *Nat. Rev. Mol. Cell Biol.* 10, 373–384

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

5. Yao, W., King, D. A., Beckwith, S. L., Gowans, G. J., Yen, K., Zhou, C., and Morrison, A. J. (2016) The INO80 Complex Requires the Arp5-Ies6 Subcomplex for Chromatin Remodeling and Metabolic Regulation. *Mol. Cell. Biol.* 36, 979–991

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

7. Papamichos-Chronakis, M., Watanabe, S., Rando, O. J., and Peterson, C. L. (2011) Global regulation of H2A.Z localization by the INO80 chromatin-remodeling enzyme is essential for genome integrity. *Cell* 144, 200–213

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

9. Watanabe, S., and Peterson, C. L. (2016) Response to Comment on "A histone acetylation switch regulates H2A.Z deposition by the SWR-C remodeling enzyme". *Science* 353, 358–358

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

11. Yuan, G.-C., Liu, Y.-J., Dion, M. F., Slack, M. D., Wu, L. F., Altschuler, S. J., and Rando, O. J. (2005) Genome-scale identification of nucleosome positions in S. cerevisiae. *Science* 309, 626–630

12. Rando, O. J., and Chang, H. Y. (2009) Genome-wide views of chromatin structure. *Annu. Rev. Biochem.* 78, 245–271

13. Zhang, Z., Wippo, C. J., Wal, M., Ward, E., Korber, P., and Pugh, B. F. (2011) A packing mechanism for nucleosome organization reconstituted across a eukaryotic genome. *Science* 332, 977–980

14. Krietenstein, N., Wal, M., Watanabe, S., Park, B., Peterson, C. L., Pugh, B. F., and Korber, P. (2016) Genomic Nucleosome Organization Reconstituted with Pure Proteins. *Cell* 167, 709–721.e12

15. Yen, K., Vinayachandran, V., Batta, K., Koerber, R. T., and Pugh, B. F. (2012) Genome-wide nucleosome specificity and directionality of chromatin remodelers. *Cell* 149, 1461–1473

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

17. Shen, X. (2004) Preparation and analysis of the INO80 complex. *Meth. Enzymol.* 377, 401–412

18. Luger, K., Rechsteiner, T. J., and Richmond, T. J. (1999) Preparation of nucleosome core particle from recombinant histones. *Meth. Enzymol.* 304, 3–19

19. Zhou, C. Y., and Narlikar, G. J. (2016) Analysis of Nucleosome Sliding by ATP-Dependent Chromatin Remodeling Enzymes. *Meth. Enzymol.* 573, 119–135

20. 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

21. Wolffe, A. P., Jordan, E., and Brown, D. D. (1986) A bacteriophage RNA polymerase transcribes through a Xenopus 5S RNA gene transcription complex without disrupting it. *Cell* 44, 381–389

22. Partensky, P. D., and Narlikar, G. J. (2009) Chromatin remodelers act globally, sequence positions nucleosomes locally. *J. Mol. Biol.* 391, 12–25

23. Yang, J. G., Madrid, T. S., Sevastopoulos, E., and Narlikar, G. J. (2006) The chromatin-remodeling enzyme ACF is an ATP-dependent DNA length sensor that regulates nucleosome spacing. *Nat. Struct. Mol. Biol.* 13, 1078–1083

24. Polach, K. J., and Widom, J. (1995) Mechanism of protein access to specific DNA sequences in chromatin: a dynamic equilibrium model for gene regulation. *J. Mol. Biol.* 254, 130–149

25. Harada, B. T., Hwang, W. L., Deindl, S., Chatterjee, N., Bartholomew, B., and Zhuang, X. (2016) Stepwise nucleosome translocation by RSC remodeling complexes. *Elife* 5, 3653

26. Deindl, S., Hwang, W. L., Hota, S. K., Blosser, T. R., Prasad, P., Bartholomew, B., and Zhuang, X. (2013) ISWI remodelers slide nucleosomes with coordinated multi-base-pair entry steps and single-base-pair exit steps. *Cell* 152, 442–452

27. Blosser, T. R., Yang, J. G., Stone, M. D., Narlikar, G. J., and Zhuang, X. (2009) Dynamics of nucleosome remodelling by individual ACF complexes. *Nature* 462, 1022–1027

28. Hwang, W. L., Deindl, S., Harada, B. T., and Zhuang, X. (2014) Histone H4 tail mediates allosteric regulation of nucleosome remodelling by linker DNA. *Nature* 512, 213–217

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

30. Bondarenko, V. A., Steele, L. M., Ujvári, A., Gaykalova, D. A., Kulaeva, O. I., Polikanov, Y. S., Luse, D. S., and Studitsky, V. M. (2006) Nucleosomes can form a polar barrier to transcript elongation by RNA polymerase II. *Mol. Cell* 24, 469–479

31. Gaykalova, D. A., Kulaeva, O. I., Volokh, O., Shaytan, A. K., Hsieh, F.-K., Kirpichnikov, M. P., Sokolova, O. S., and Studitsky, V. M. (2015) Structural analysis of nucleosomal barrier to transcription. *Proc. Natl. Acad. Sci. U.S.A.* 112, E5787–95

32. Willhoft, O., McCormack, E. A., Aramayo, R. J., Bythell-Douglas, R., Ocloo, L., Zhang, X., and Wigley, D. B. (2017) Cross-talk within a functional INO80 complex dimer regulates nucleosome sliding. *Elife* 6, e25782

33. Tosi, A., Haas, C., Herzog, F., Gilmozzi, A., Berninghausen, O., Ungewickell, C., Gerhold, C. B., Lakomek, K., Aebersold, R., Beckmann, R., and Hopfner, K.-P. (2013) Structure and subunit topology of the INO80 chromatin remodeler and its nucleosome complex. *Cell* 154, 1207–1219

34. Chen, L., Conaway, R. C., and Conaway, J. W. (2013) 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.* 110, 20497–20502

35. Shen, X., Xiao, H., Ranallo, R., Wu, W.-H., and Wu, C. (2003) Modulation of ATP-dependent chromatin-remodeling complexes by inositol polyphosphates. *Science* 299, 112–114

36. Shen, X., Ranallo, R., Choi, E., and Wu, C. (2003) Involvement of actin-related proteins in ATP-dependent chromatin remodeling. *Mol. Cell* 12, 147–155

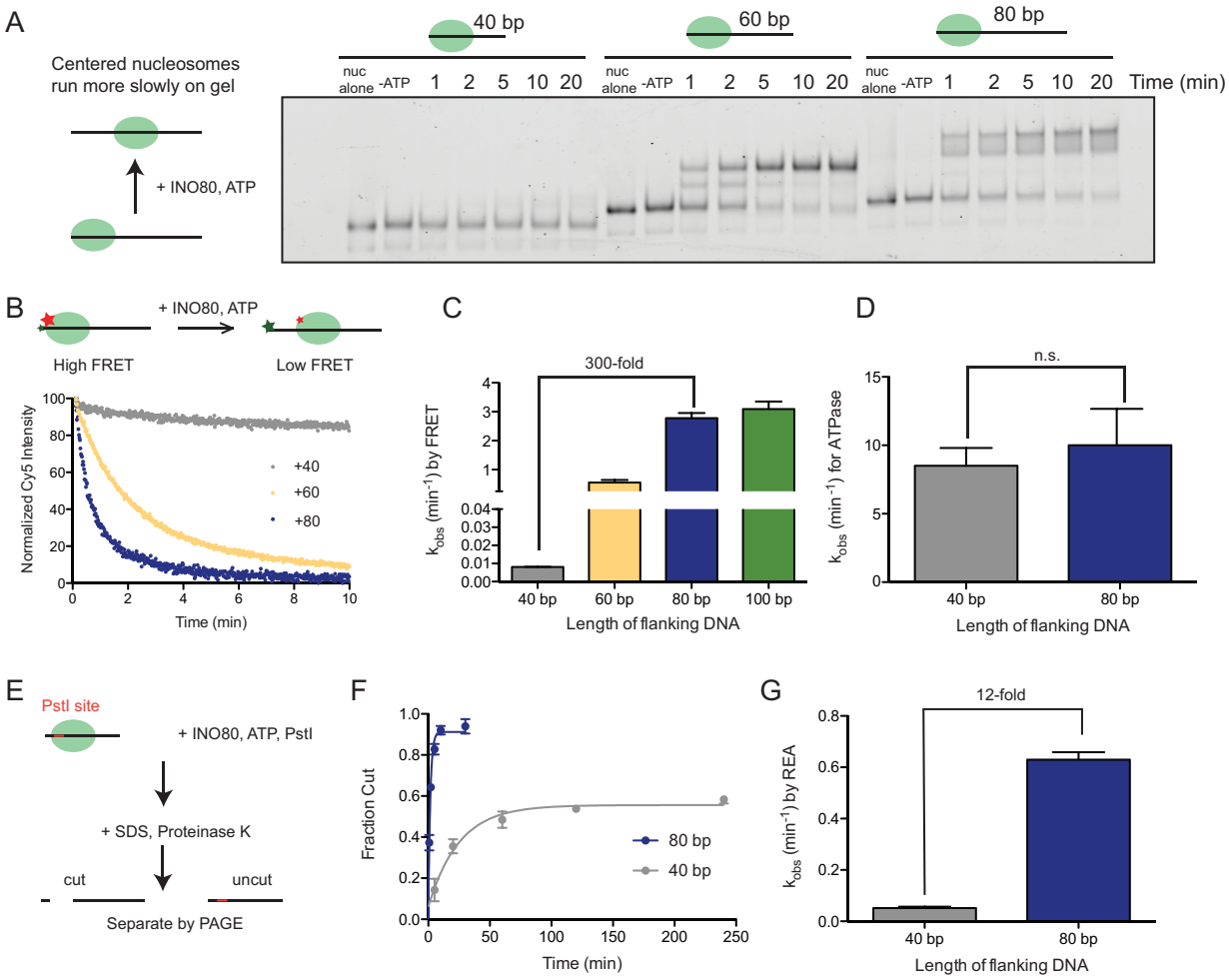
37. Chen, L., Cai, Y., Jin, J., Florens, L., Swanson, S. K., Washburn, M. P., Conaway, J. W., and Conaway, R. C. (2011) Subunit organization of the human INO80 chromatin remodeling complex: an evolutionarily conserved core complex catalyzes ATP-dependent nucleosome remodeling. *J. Biol. Chem.* 286, 11283–11289

38. Hauk, G., McKnight, J. N., Nodelman, I. M., and Bowman, G. D. (2010) The chromodomains of the Chd1 chromatin remodeler regulate DNA access to the ATPase motor. *Mol. Cell* 39, 711–723

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

40. Leonard, J. D., and Narlikar, G. J. (2015) A nucleotide-driven switch regulates flanking DNA length sensing by a dimeric chromatin remodeler. *Mol. Cell* 57, 850–859

Figure 1: Flanking DNA length modulates nucleosome sliding by INO80, but not ATPase activity



(A) Left, schematic of the native gel assay. Right, native gel showing the remodeling of nucleosomes over time with 40 bp, 60 bp, or 80 bp of flanking DNA ("0/40", "0/60", and "0/80" nucleosomes), under single turnover conditions and with saturating enzyme and ATP.

(B) Left, schematic of the ensemble FRET remodeling assay and example raw data for 0/40, 0/60, and 0/80 nucleosomes.

(C) Rate constants for remodeling with 0/40, 0/60, and 0/80 nucleosomes by ensemble FRET. These assays were performed under single turnover conditions and with saturating enzyme and ATP. Data represent the mean ± the S.E.M. for three replicates.

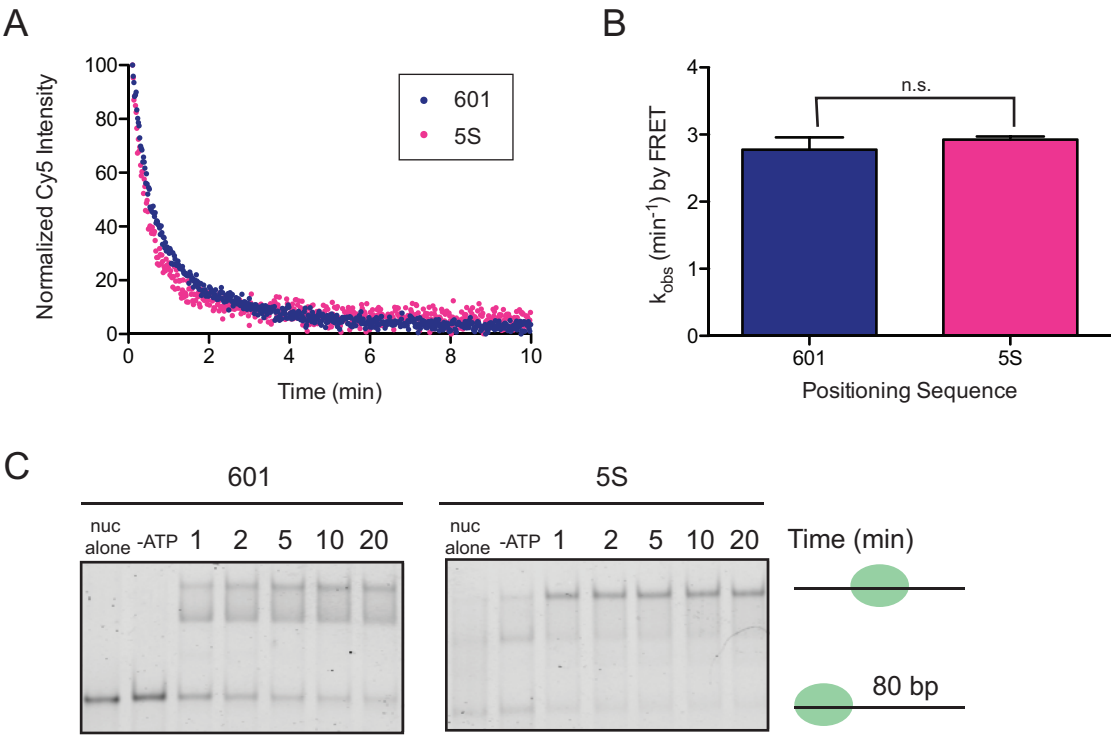
(D) Rate constants for ATP hydrolysis for 0/40 and 0/80 nucleosomes derived from initial rates. Data represent the mean ± the S.E.M. for three replicates. These assays were performed under multiple turnover conditions, with IN080 in excess and saturating over nucleosomes, and ATP in excess of INO80.

(E) Schematic of the REA assay. Remodeling by INO80 (DNA unpeeling, a conformational change, or nucleosome sliding) exposes a Pst1 site located 18 bp from the short end of the DNA. After digestion of the histones with Proteinase K, uncut and cut products are separated by native gel.

(F) Quantification of the fraction of the DNA cut by the Pst1 restriction enzyme as a function of time for 0/40 and 0/80 nucleosomes.

(G) Quantification of the rate constant for cutting by Pst1 for 0/40 and 0/80 nucleosomes. Data represent the mean ± the S.E.M. for three replicates. These assays were performed under single turnover conditions and with saturating enzyme and ATP.

Figure S1: DNA sequence affects final product distribution, but not remodeling rates, by INO80

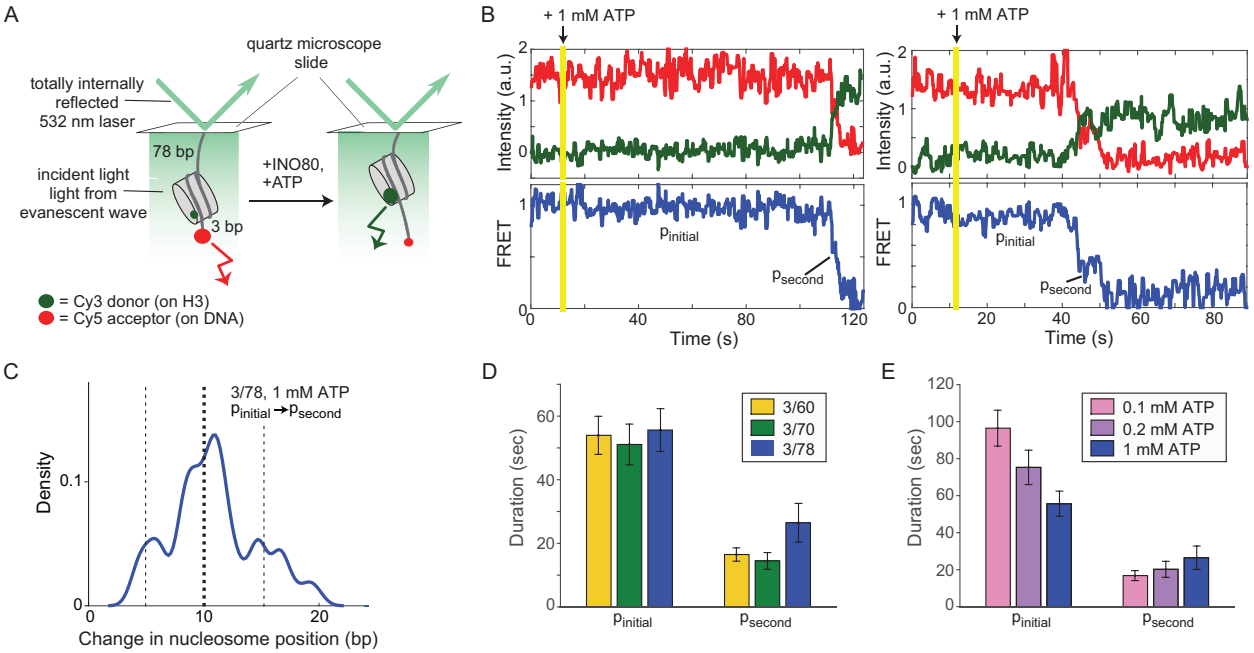


(A) Raw data showing FRET remodeling of 0/80 nucleosomes positioned with either the 601 (blue) or 5S (pink) sequence.

(B) Quantification of the rate constants of FRET remodeling for either 601 or 5S 0/80 nucleosomes. These experiments were performed under single turnover conditions and with saturating enzyme and ATP. Data represent the mean ± S.E.M for three replicates.

(C) Native gel of remodeling timecourse with 601- or 5S-positioned 0/80 nucleosomes.

Figure 2: Remodeling by INO80 at the single nucleosome level is preceded by a long pause, followed by rapid nucleosome translocation



(A) Schematic of the smFRET assay. Nucleosomes labeled on histone H3 with a Cy3 donor dye and on the DNA end with a Cy5 acceptor dye are immobilized on the surface of microscope slides and imaged with a prism-based TIRF microscope (see also Methods). In the end-positioned "3/78" construct shown here, the nucleosome starts 3 bp from the Cy5-labeled DNA end, with a 78 bp linker attaching the nucleosome to the surface, resulting in an initial high FRET efficiency (Figure S2). As remodeling proceeds, the nucleosome is moved towards the 78 bp flanking DNA and the FRET efficiency is reduced.

(B) Example timecourses of remodeling from individual, surface-attached, 3/78 nucleosomes in the presence of saturating INO80 and ATP (15 nM and 1 mM respectively), imaged at 7.4 Hz and smoothed (for visualization only) with a median filter with a 1-second window. Vertical yellow line indicates the time at which ATP is introduced into the sample chamber to start the remodeling reaction. Both examples here show the long initial pause ($p_{initial}$) exhibited by all

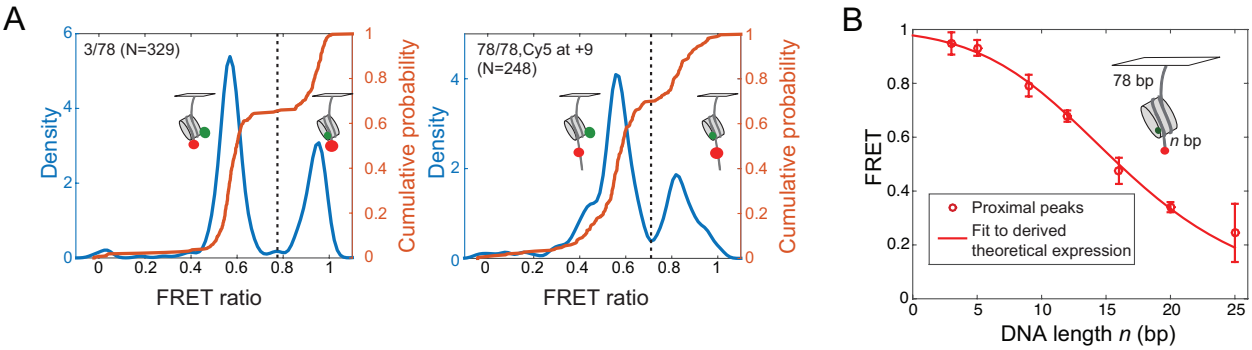
trajectories, as well as a secondary pause (p_{second}), exhibited in about three quarters of remodeling events (Figure S4B,C).

(C) Kernel density estimation (KDE) plot of the change in nucleosome position between $p_{initial}$ and p_{second} for the 3/78 construct at saturating ATP. KDEs are conceptually similar to histograms, but have a smoothing parameter rather than a bin size, which reduces the potential for generating artificial clusters or "peaks" in the data based on choice of bin width and location. The majority of the trajectories that show secondary pauses have translocated the nucleosome about 8-12 bp before the secondary pause is encountered. KDE has a Gaussian kernel with bandwidth 0.75.

(D) Quantification of the average durations of the initial and secondary pauses as a function of entry DNA length. Nucleosomes initially have 3 bp on one side, and 60, 70, or 78 bp on the opposite side. Some traces exhibit more than one secondary pause; however these traces are too few to gain accurate statistics beyond the p_{second} pause. Data represent means±S.E.M. obtained via a bootstrapping approach (see Methods).

(E) Quantification of the average durations of the initial and secondary pauses as a function of ATP concentration, for 3/78 nucleosomes. 1 mM ATP is saturating, while 0.1 mM and 0.2 mM are both sub-saturating (Figure S4D; (6)). Dark blue data are the same as in C. Data represent means±S.E.M. obtained via a bootstrapping approach (see Methods).

Figure S2: Relationship between FRET and bp flanking DNA on the Cy5-labeled exit side of the nucleosome

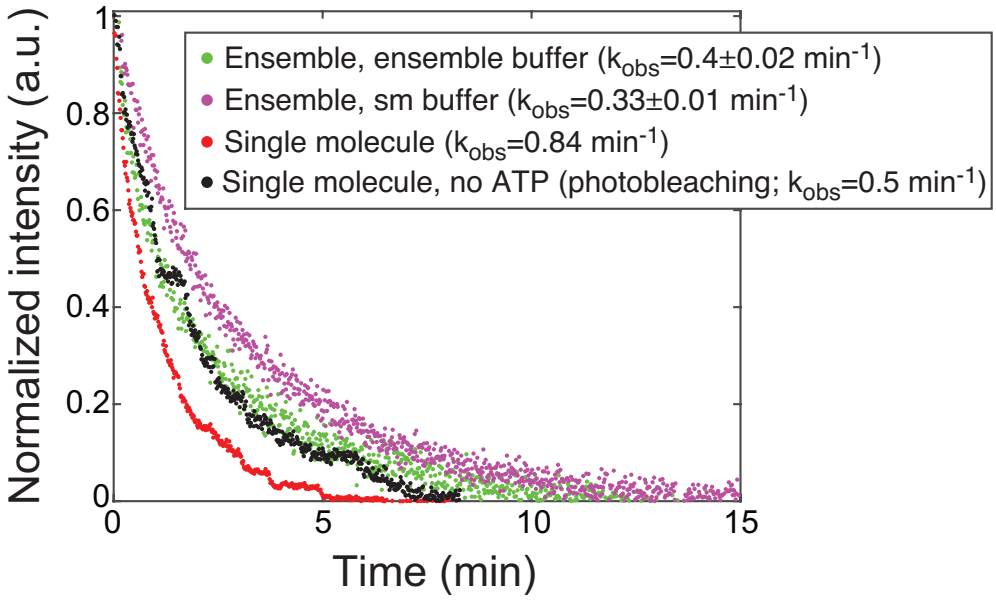


(A) Kernel density estimation plots (KDEs, blue) and empirical cumulative probabilities (CDFs, orange) of initial FRET values for 3/78 nucleosomes with the Cy5 on the DNA end (left), and 78/78 nucleosomes with the Cy5 at 9 bp from the edge of 601 (right). The two peaks in the KDE plots, which are similar to histograms but are smoothed rather than binned, correspond to two expected populations of nucleosomes: one population in which the H3 proximal to the Cy5-labeled DNA end is Cy3 labeled, resulting in higher initial FRET, and one in which the distal H3 has the Cy3 label. Nucleosomes with neither H3 labeled are invisible; those with both H3's labeled with Cy3 exhibit two-step photobleaching in the Cy3 channel and are excluded from these histograms as well as from remodeling datasets. The CDFs represent the same information as the KDEs but without the smoothing filter; peaks in the KDEs appear in the CDFs as steep slopes. Nucleosomes here are in the presence of 1 mM ATP but in the absence of remodeler. *N* indicates the number of nucleosomes included in the KDE and CDF; vertical black dashed lines indicate the lower threshold FRET value for nucleosomes to be classified as starting with a proximal H3 label and thus included for further analysis. KDE bandwidths are 0.01 (left) and 0.025 (right).

(B) Calibration curve showing the relationship between the average FRET value of the proximal Cy5 dye, and bp flanking DNA, *n*. This calibration curve allows us to convert FRET values measured during remodeling to bp DNA moved out of the nucleosome. Red curve is a fit to our

derived theoretical expression for FRET as a function of *n* (see Methods; (30)). Errors on the data are the standard deviations of the Gaussians obtained by fitting Gaussian mixture models to CDFs of FRET values for each construct (see panel A for the CDF for the *n* = 3 case).

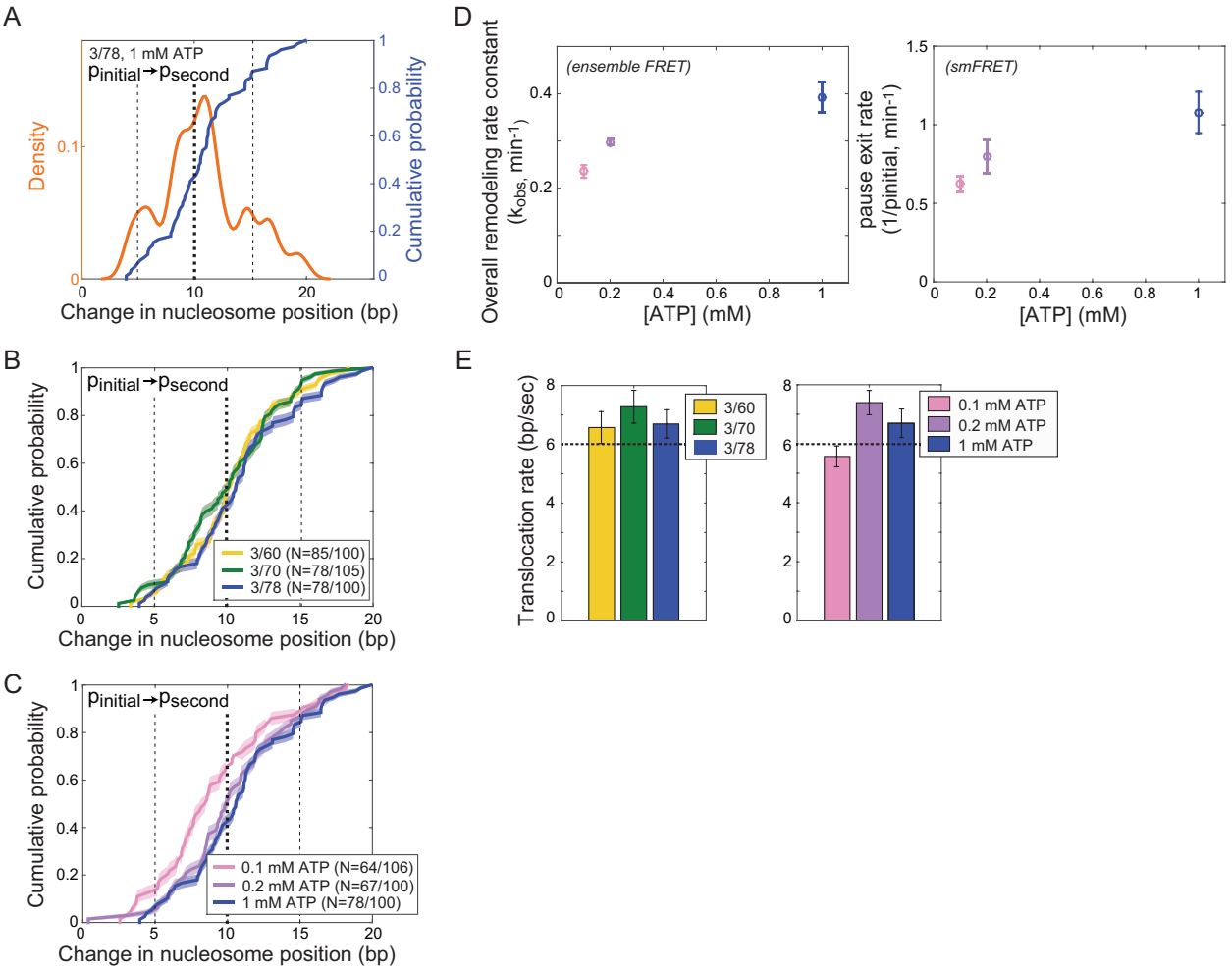
Figure S3: Comparison of ensemble FRET and smFRET remodeling rates



Comparison between ensemble FRET remodeling rates in ensemble buffer versus the imaging buffer used for smFRET, and between these rates and sliding rates observed by smFRET. Data are with 3/78 nucleosomes, saturating ATP (1 mM) and saturating enzyme (15 nM), at 20°C and under single turnover conditions. Ensemble FRET rate constants are means and S.E.M. from three replicates. smFRET data were converted to "pseudo-ensemble" data by summing the intensities of all Cy5 molecules included in the smFRET data, and then binning into 1-second intervals as in our ensemble FRET data. That is, for all nucleosomes that exhibit the sliding behavior shown in the examples in Figure 2B in the main text, the Cy5 intensity data were summed and then binned. Note that this results in a change in FRET as a function of time only for events that are clearly sliding. Events that look like photobleaching (an instantaneous transition from high FRET to zero FRET within one frame) were excluded. Note the difference between these data (red) and the pseudo-ensemble smFRET data in the absence of ATP (black), in which all quantified events were photobleaching events. We note that ensemble FRET remodeling (green and purple) is slower than the sliding rates observed by smFRET (red), suggesting that attachment of the nucleosome to the surface of the microscope slide does

not adversely impact remodeling (which is also the case for other remodelers, including large complexes like RSC (25)). We note also that because of the way these experiments are performed, the smFRET data are under chase conditions, while the ensemble FRET data are not.

Figure S4: Additional characteristics of the pauses and the translocation phase



(A) An example combined KDE and CDF plot (see also Figure S2) for the distance the nucleosome is moved in the first translocation phase (that is, the distance the nucleosome is moved between $p_{initial}$ and p_{second}). The major peak in the KDE corresponds to the steepest slope in the CDF, and indicates that most nucleosomes are moved about 10 bp between $p_{initial}$ and p_{second} (alternatively, note that the cumulative probability is roughly 0.5 at 10 bp, indicating that half the nucleosomes are moved 10 bp or fewer during the first translocation event). Although KDEs are more intuitive, being conceptually similar to histograms, they are less preferable for making quantitative comparisons, because peaks can “appear” or “disappear” in the data as the smoothing parameter is varied. Moreover, KDEs are not normalized, again making quantitative

comparisons between data sets difficult. Therefore in B and C, we show only CDFs in order to make comparisons between data sets. The KDE here has a Gaussian kernel with bandwidth 0.75.

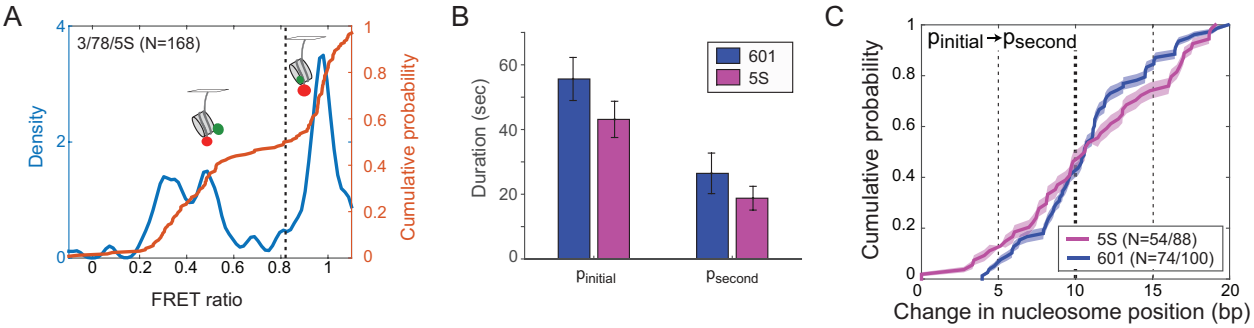
(B) CDFs showing the distance a nucleosome is moved between $p_{initial}$ and p_{second}, as a function of flanking DNA length. The steepest slope occurs for all three constructs primarily at 10 bp of movement, and secondarily at 5 or 15 bp. An alternative summary of the information in these CDFs is that roughly half of the nucleosomes are moved 10 bp or fewer during the first translocation event; but only about 5%-10% of them move fewer than 7-8 bp, and ~80% move 12 bp or fewer. This again indicates that 8-12 bp is the predominant distance the nucleosomes is moved between the initial and secondary pauses, for the traces that have a secondary pause within FRET range. *N* in the legend indicates how many traces were included in these plots; only traces with distinct secondary pauses within FRET range (about 75% of the data) were included. Errors are indicated as shaded areas around the solid curves and are obtained via a bootstrapping procedure (see Methods).

(C) Same as B but as a function of ATP concentration. It appears that at 0.1 mM ATP, the step size (that is, the number of bp the nucleosome is moved between $p_{iniital}$ and $p_{secondary}$) decreases. This could suggest that at low ATP concentrations, INO80 may get “stuck” at more locations as the DNA is moved over the histone octamer, rather than pausing only at the locations of the most favorable DNA-histone contacts.

(D) Left, overall ensemble FRET remodeling rate as a function of ATP concentration for 3/78 nucleosomes. These experiments were performed at 20°C, under single turnover conditions, with saturating enzyme and ATP. Data represent means±S.E.M. for three replicates. Right, the inverse of the $p_{initial}$ durations in Figure 2E, as a function of ATP concentration. Most, if not all, of the observed reduction in overall remodeling rate measured by ensemble FRET can be attributed to an increase in the duration of $p_{initial}$ (i.e., a decrease in $1/p_{initial}$).

(E) Rate of nucleosome sliding as a function of flanking DNA length (left) or ATP concentration (right). The small but statistically significant decrease in translocation rate below 0.2 mM ATP (indicated by dashed line) is comparable to the effect of ATP on the translocation rate of ACF in a similar concentration range (32). Data represent means±S.E.M. obtained via a bootstrapping procedure (see Methods).

Figure S5: Effect of DNA sequence on sliding of individual nucleosomes by INO80

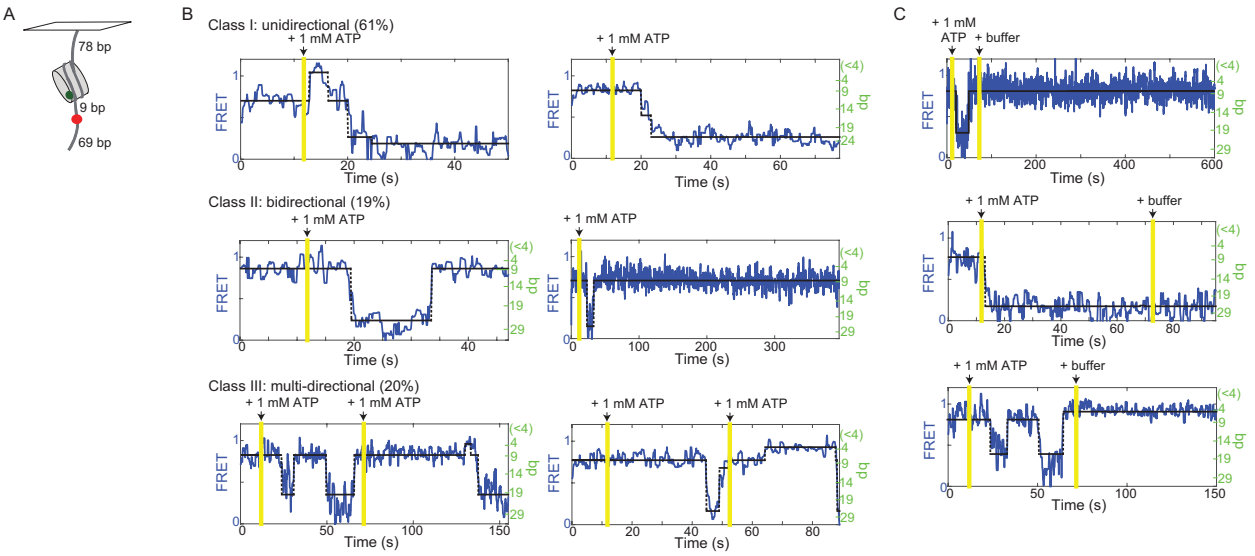


(A) KDE and CDF of initial FRET values of 5S nucleosomes (see also Figure S2A).

(B) Initial and secondary pause durations with 3/78 nucleosomes containing either the 601 positioning sequence or the 5S positioning sequence. We observe similar overall remodeling rates with 5S and 601 by ensemble FRET (Figure S1), which is consistent with $p_{initial}$ being insensitive to sequence. (Although there appears to be a trend towards shorter pauses with 5S nucleosomes, this trend is not statistically significant.) Data represent means±S.E.M. obtained via a bootstrapping procedure (see Methods).

(C) CDFs of the step size as a function of positioning sequence (see also Figure S4). The overall slope of the CDF is reduced for 5S compared to 601, indicating that with 5S nucleosomes, there is not as strong a preference for moving the nucleosome 10 bp. That is, roughly half of the nucleosomes are moved 10 bp or fewer in the first translocation event with both 601 and 5S nucleosomes; but more of them (~10%) are moved 5 bp or fewer with 5S, compared to 601, in which only about 5% are moved 5 bp or fewer. Similarly, ~80% of the nucleosomes are moved 15 bp or fewer with 601, but only about 75% with 5S. This increased diversity in step size is consistent with the 5S sequence having a reduced preference for its rotational positioning on the octamer compared to 601 (22).

Figure 3: INO80 processively and continuously moves a nucleosome bi-directionally on long DNA



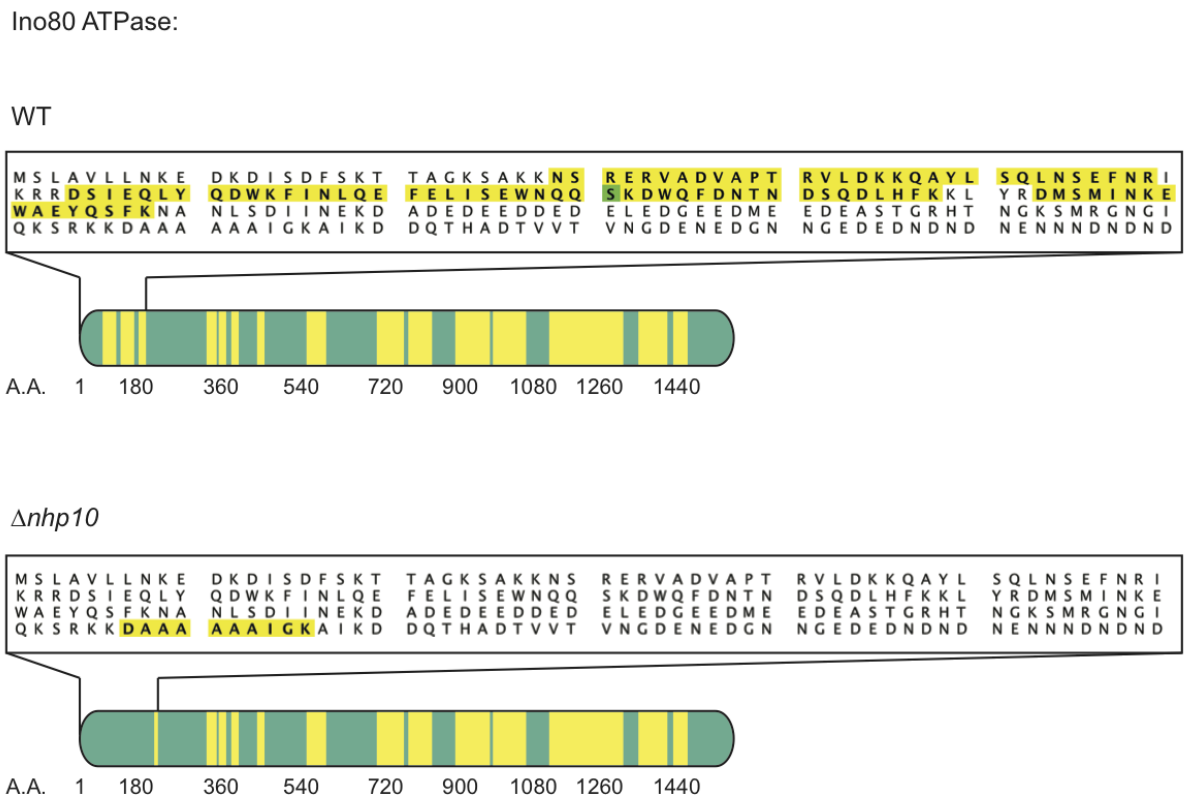
(A) Schematic of the centered 78/78 construct. These nucleosomes have an internal Cy5 9 bp from one edge of the 601 sequence that initially positions the nucleosome, such that they start in mid-high FRET.

(B) Example traces of remodeling of 78/78 nucleosomes with saturating INO80 and ATP (15 nM and 1 mM respectively). All traces have been cropped where either the donor or acceptor dye photobleached. We observe three kinds of behavior with these centered constructs: "unidirectional" remodeling, in which the nucleosome is moved out of FRET range and does not come back into FRET range before one of the dyes photobleaches (Class I); bidirectional movement, in which the nucleosome initially moves out of FRET range but then returns, and can persist in mid to high FRET even for very long durations, as in the right panel (Class II); and multi-directional movement, in which the nucleosome is continuously moved in and out of FRET range before photobleaching occurs (Class III). All of these traces were acquired under chase conditions, in that, as in Figure 2, unbound INO80 is washed out as ATP is washed in. As an additional demonstration of processivity, in some cases a “double chase” experiment was

performed, as in the two Class III examples. In these experiments, a second buffer exchange into 1 mM ATP was performed, as indicated by the second vertical yellow line. The percent of the ~100 total traces (from the single-chase and double-chase experiments combined) that we classified into each type of behavior are given in parenthesis. We note that our calibration curve (Figure S2), which allows us to convert the FRET on the left axes to bp on the right axes, applies only when the dye is in the flanking DNA, not when the nucleosome has been moved over the dye's position; thus one of the Class I traces, which shows only movement towards the dye (and then beyond the dye on the other side), has not been assigned a bp scale.

(C) As an additional control to demonstrate that the behavior in B is in fact due to processive motion under true chase conditions, no remodeling was observed after exchange into buffer only (no ATP).

Figure S6: Summary of mass spectrometry data of the Ino80 protein isolated from either WT (top) or *Δnhp10* (bottom) yeast



The secondary structure of the Ino80 protein is shown in teal, with peptides identified indicated in yellow. The sequence of the N-terminus is shown as a box above each structure, with the identified peptide sequences highlighted in yellow.

Figure S7: DNA sequences used in this work

```
"0/100":
5′-/5Cy3/CTGGAGAATCCCGGTGCCGAGGCCGCTCAATTGGTCGTAGACAGCTCTAGCACCGCTTAAACGCACGTACGCGCTGTCCCCCGCGTTTTAACCG...
3′-      GACCTCTTAGGGCCACGGCTCCGGCGAGTTAACCAGCATCTGTCGAGATCGTGGCGAATTTGCGTGCATGCGCGACAGGGGGCGCAAAATTGGC...

...CCAAGGGGATTACTCCCTAGTCTCCAGGCACGTGTCAGATATATACATCCTGTGCATCTATTGAACAGCGACCTTGCCGGTGCCAGTCGGATAGTGTTCCGAG...
...GGTTCCCCTAATGAGGGATCAGAGGTCCGTGCACAGTCTATATATGTAGGACACGTAGATAACTTGTCGCTGGAACGGCCACGGTCAGCCTATCAGAAGGCTC...

...CTCCCACTCTAGAGGATCCCCGGGTACCGAGCTCGAATTCGCCCTATA-3′
...GAGGGTGAGATCTCCTAGGGGCCCATGGCTCGAGCTTAAGCGGGATAT-5′

"3/78":
5′-/5Cy5/GCCCTGGAGAATCCCGGTCTGCAGGCCGCTCAATTGGTCGTAGACAGCTCTAGCACCGCTTAAACGCACGTACGCGCTGTCCCCCGCGTTTTAACCG...
3′-      CGGGACCTCTTAGGGCCAGACGTCCGGCGAGTTAACCAGCATCTGTCGAGATCGTGGCGAATTTGCGTGCATGCGCGACAGGGGGCGCAAAATTGGC...

...CCAAGGGGATTACTCCCTAGTCTCCAGGCACGTGTCAGATATATACATCCTGTGCATCTATTGAACAGCGACCTTGCCGGTGCCAGTCGGATAGTGTTCCGAG...
...GGTTCCCCTAATGAGGGATCAGAGGTCCGTGCACAGTCTATATATGTAGGACACGTAGATAACTTGTCGCTGGAACGGCCACGGTCAGCCTATCAGAAGGCTC...

...CTCCCACTCTAGAGGATCCCCGGGTACC-3′
...GAGGGTGAGATCTCCTAGGGGCCCATGG/5Bioteg/-5′

"3/78/5S" and "0/80/5S":
5′-/5Cy5/CCAGGCCCGACCCTGCTTGGCTGCAGAGATCAGACGATATCGGGCACTTTCAGGGTGGTATGGCCGTAGGCGAGCACAAGGCTGACTTTTCCTCCCC...
3′-      GGTCCGGGCTGGGACGAACCGACGTCTCTAGTCTGCTATAGCCCGTGAAAGTCCCACCATACCGGGATCCGCTCGTGTTCCGACTGAAAAGGAGGGG...

...TTGTGCTGCCTTCTGGGGGGGGCCCAGCCGGATCCCCGGGCGAGCTCGAATTCGTAATCATGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAAT...
...AACACGACGGAAGACCCCCCCCGGGTCGGCCTAGGGGCCCGCTCGAGCTTAAGCATTAGTACAGTATCGACAAAGGACACACTTTAACAATAGGCGAGTGTTA...

...TCCACACAACATACGAGCCGGAAGCATAA-3′
...AGGTGTGTTGTATGCTCGGCCTTCGTATT/5Bioteg/-5′

"78-78":
5′-GGATCCTAATGACCAAGGAAAGCATGATTCTTCACACCGAGTTCATCCCTTATGTGATGGACCCTATAC*T*CGGCCGCCCTGGAGAATCCCGGTCTGCAGGC...
3′-CCTAGGATTACTGGTTCCTTTCGTACTAAGAAGTGTGGCTCAAGTAGGGAATACACTACCTGGGATATG C GCCGGCGGGACCTCTTAGGGCCAGACGTCCG...

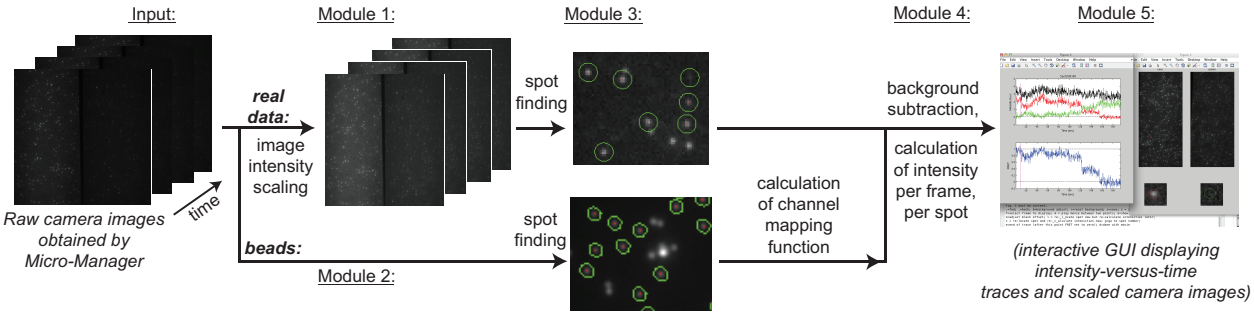
...CGCTCAATTGGTCGTAGACAGCTCTAGCACCGCTTAAACGCACGTACGCGCTGTCCCCCGCGTTTTAACCGCCAAGGGGATTACTCCCTAGTCTCCAGGCACG...
...GCGAGTTAACCAGCATCTGTCGAGATCGTGGCGAATTTGCGTGCATGCGCGACAGGGGGCGCAAAATTGGCGGTTCCCCTAATGAGGGATCAGAGGTCCGTGC...

...TGTCAGATATATACATCCTGTGCATCTATTGAACAGCGACCTTGCCGGTGCCAGTCGGATAGTGTTCCGAGCTCCCACTCTAGAGGATCCCCGGGT...
...ACAGTCTATATATGTAGGACACGTAGATAACTTGTCGCTGGAACGGCCACGGTCAGCCTATCAGAAGGCTCGAGGGTGAGATCTCCTAGGGGCCCA...

...ACC-3′
...TGG/5Bioteg/-5′
```

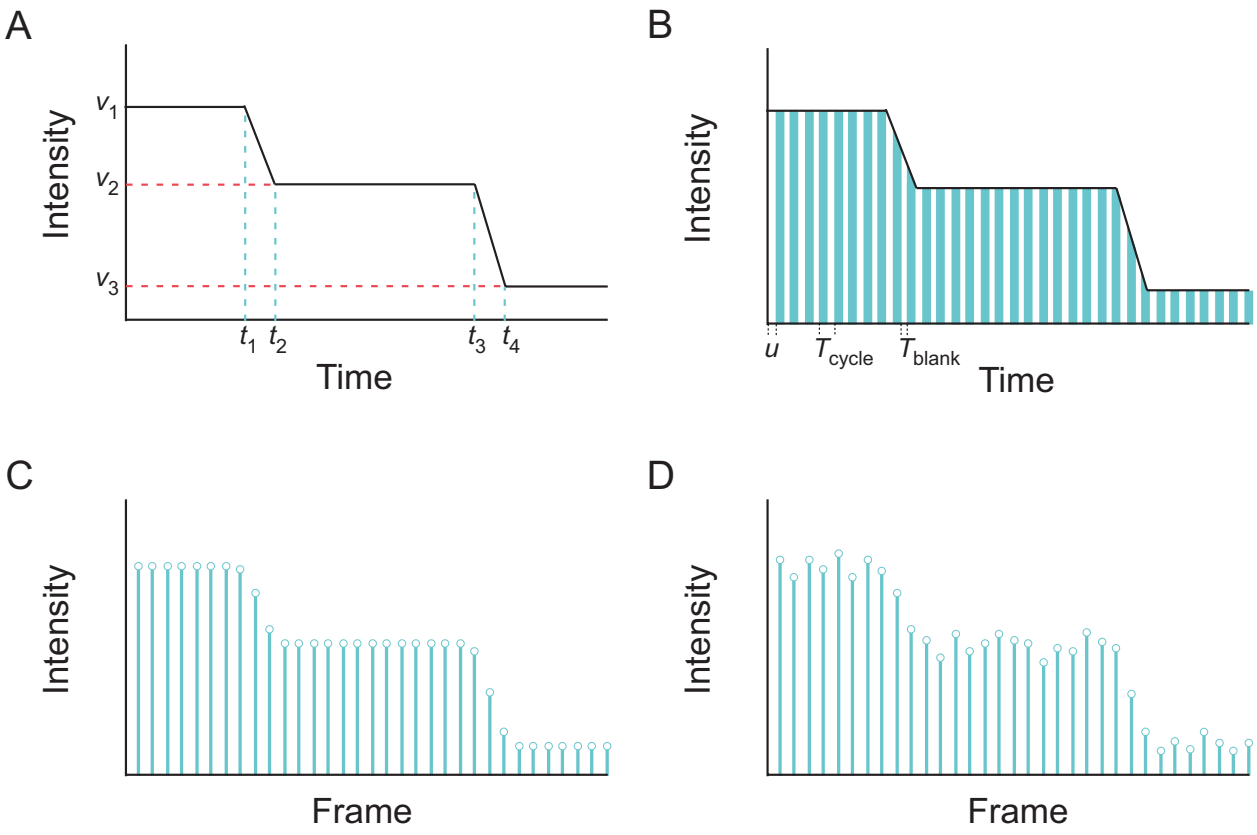
The nucleosome positioning sequence (601 for all constructs except "3/78/5S" and "0/80/5S") is underlined; a Pst1 restriction enzyme recognition site, if present, is in boldface blue. The nucleotide in 78/78 to which the internal Cy5 is conjugated is flanked by asterisks. The two additional bases on the long end of 0/80/5S, compared to 3/78/5S, are highlighted in yellow. The 0/80/5S construct also has a Cy3 label instead of the Cy5 shown here, and no biotin on the other end. All other DNAs used in this work can be derived from these sequences (e.g. 0/40, 0/60, and 0/80 DNAs used for ensemble FRET are truncations of the 0/100 construct; the 3/60 and 3/70 constructs used for smFRET are truncations of 3/78; etc.) Constructs shown here are for FRET-labeled nucleosomes; REA constructs have the Cy3 on the 5' end of the opposite DNA strand (that is, on the long flanking DNA).

Figure S8: Summary of the Traces smFRET analysis software



The workflow for analyzing smFRET data can be divided into 5 modules: (1) data pre-processing, which includes scaling each image to take advantage of the full dynamic range of the numeric type used; (2) finding a mapping between donor and acceptor channels, which will never be perfectly aligned in any setup, and which we accomplish by fitting a polynomial transformation to paired images of surface-attached, microscopic fluorescent beads; (3) spot finding, or identifying which pixels in each channel contain real fluorescence vs. background fluctuations; (4) calculation of the intensity of each identified spot in each frame of a movie; and (5) displaying these time trajectories ("traces") for the user to pick out results for further analysis. The GUI that accomplishes Module 5 in Traces also provides options for the user to refine the localization of each spot (for example, if the bead-based polynomial transformation was inaccurate for part of the field of view), adjust the background to be subtracted, crop parts of the trace for downstream analysis, and inspect the camera images that generated each trace.

Figure S9: Schematic summary of the Slopey algorithm for quantifying translocation events



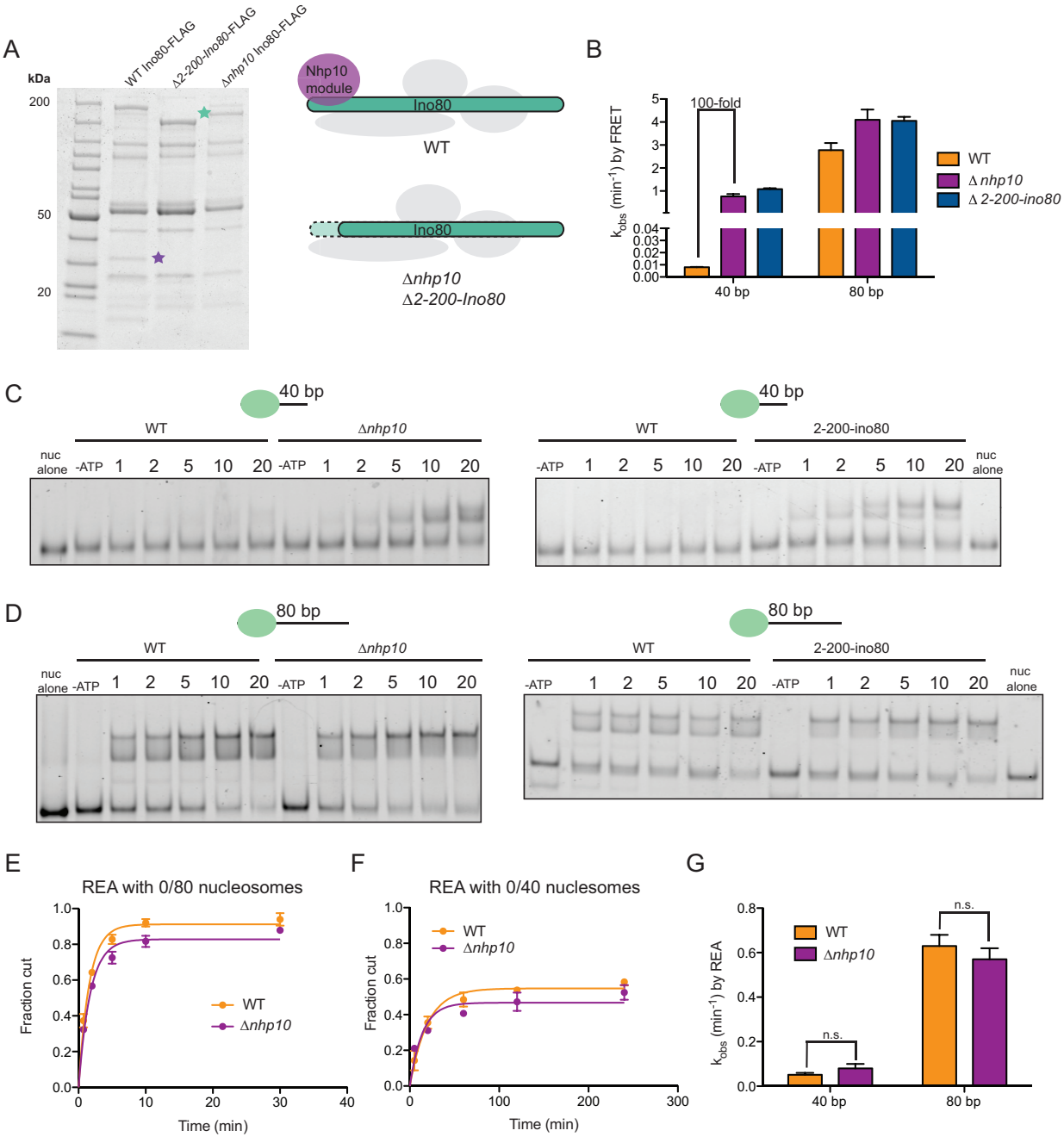
(A) Idealized trajectory of acceptor dye intensity over time. The pause phases have acceptor intensities of v_1, v_2, and v_3; transitions from pause phases to translocation phases occur at times t_1 and t_3, while transitions from translocation phases to pauses occur at times t_2 and t_4. (B) Schematic of the camera model. Blue bars represent the times during which the camera is recording fluorescence intensity. An image frame acquired by the camera, which has total duration T_{cycle}, consists of the blue bars (the exposure time), plus T_{blank}, during which the information is transferred off of the camera chip into memory. Light intensity is not recorded during T_{blank}. We collect data with frame transfer off, such that T_{cycle} = 135.6 seconds and T_{blank} = 35.6 seconds for 100 ms exposure. Note that translocation events can begin at any point during

T_{cycle}, including during T_{blank}. The offset parameter u represents the time between the beginning of the data trajectory and the first frame recorded by the camera.

(C) Idealized total intensities collected during each frame. The height of each bar represents the area in each blue box in B. Note that if a translocation event begins mid-frame, as at time t_3, the total intensity collected during that time will be less than during the preceding frame, but not as much less as if the frame were fully within the translocation phase.

(D) We add i.i.d. (independent and identically distributed) Gaussian noise with a variance of 0.075 to the idealized camera output in C to obtain our final model of acceptor intensity data. Given the data, the prior and this model of the data, Slopey infers u, $\{v_n\}$, and $\{t_n\}$ and the channel transform parameters a and b (see Methods).

Figure 4: Deletion of *nhp10* results in impaired DNA length sensing by INO80



(A) Left, denaturing gels showing purifications of Ino80-FLAG(WT), Ino80-FLAG(Δ*2-200-ino80*) and Ino80-FLAG(Δ*nhp10*). Green and purple stars represent where the Ino80 ATPase and a

major component of the Nhp10 module run on the gel, respectively. Right, schematic illustrating the deletion of the N-terminus of Ino80 observed in ∆*nhp10* and ∆*2-200*-ino80 strains.

(B) Remodeling rate constants measured by ensemble FRET for either 0/40 or 0/80 nucleosomes and the various INO80 mutants described in A. These assays were performed under single turnover conditions and with saturating enzyme and ATP. Data represent means±S.E.M. for three replicates.

(C) Native gel remodeling of 0/40 nucleosomes by INO80(*∆nhp10)* and INO80(*∆2-200-ino80*). These assays were performed under single turnover conditions and with saturating enzyme and ATP.

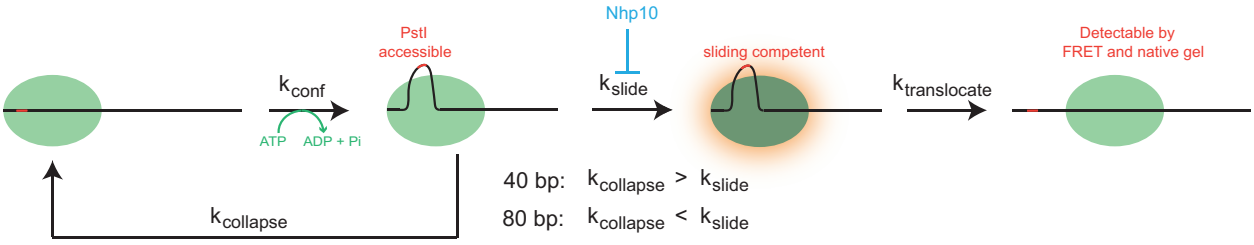
(D) Same as C but with 0/80 nucleosomes.

(E) Quantification of fraction cut by REA on 0/80 nucleosomes, with either INO80(WT) or INO80(*∆nhp10)*.

(F) Same as E but with 0/40 nucleosomes.

(G) Rate constants measured by REA for 0/40 and 0/80 nucleosomes with INO80(WT) or INO80(*∆nhp10)*. These assays were performed under single turnover conditions and with saturating enzyme and ATP. Data represent means±S.E.M. for three replicates.

Figure 5: Model for the length-dependent regulation of translocation initiation by INO80



The first catalytic step in the INO80 remodeling reaction is the ATP-dependent creation of an intermediate that is accessible to cleavage by Pst1 (k_{conf}). This intermediate either goes on to a sliding competent state (k_{slide}) which can then be slid ($k_{translocate}$), or collapses back to its original conformation ($k_{collapse}$). The ratio between $k_{collapse}$ and k_{slide}, and thus the probability that a nucleosome will be slid by INO80, is determined by the length of flanking DNA, and is also modulated by the Nhp10 module.

Acknowledgements:

We thank Julia Tretyakova for histone purifications, and all of the Narlikar lab for helpful discussions. We thank Toshio Tsukiyama for the yeast plasmids to make the *Δnhp10* strain. This work was supported by a grant from the NIH to G.J.N. (R01GM073767), from the NIH to A.M. (5R35GM119580), by a Ruth L. Kirschstein National Research Service Award (5F31CA180651-02) to C.Y.Z., and a Leukemia and Lymphoma Society Career Development Program Fellow award to S.L.J.

Author contributions:

Conceptualization: C.Y.Z., S.L.J., L.J.L., and G.J.N.; Methodology: C.Y.Z., S.L.J., L.J.L., and G.J.N.; Software: S.L.J. and M.J.J.; Investigation: C.Y.Z., S.L.J. L.J.L., and A.L.; Resources: S.B. and A.M.; Writing–Original Draft: C.Y.Z., S.L.J., and G.J.N.; Writing–Reviewing & Editing: C.Y.Z., S.L.J., L.J.L., A.L., S.B., A.M., and G.J.N.; Supervision: G.J.N.

Appendix to Chapter 3

In this section, we summarize some unpublished findings related to the story presented in Chapter 4. While preliminary, we hope that these results will help future investigators design new experiments to probe the mechanism of nucleosome movement by INO80.

Methods

Purification of DNA substrates with and without overhangs

Oligos 1 (5' TTGCACTGCAATGACTCTAG 3'), 2 (5' CTAGAGTCATTGCAGTGCAA 3'), 3 (5' AGGAAGCGGATAACAAGTTCCTAGAGTCATTGCAGTGCAA 3'), 4 (5' AAGCGGATAACAAG TTCTTGTTGCACTGCAATGACTCTAG 3'), 5 (5' CTAGAGTCATTGCAGTGCAAAGCGGATAA CAGTTTCACAC 3'), 6 (5' TTGCACTGCAATGACTCTAGAGCGGATAACAGTTTCACAC 3') were ordered from IDT. The oligos were resuspended in TE (10 mM Tris pH 8.0, 1 mM EDTA). To form the blunt substrate, Oligos 1 and 2 were annealed in the PCR machine using the following program: 95°C for 5 minutes, -1° every 40 seconds (ramp rate of 3°C/s) until temperature reaches 20°C. The same process was used to anneal Oligos 3 and 4 to make the substrate with 5' overhangs, or Oligos 5 and 6 to make the substrate with 3' overhangs. The annealed DNA's were gel-purified using a 10% acrylamide gel. DNA was extracted from the gel pieces by incubation with TE for 24 hours, followed by ethanol precipitation.

The other methods used in this section (FRET and gel remodeling, ATPase assays) are the same as those described in the methods section of Chapter 4. The exact experimental condition for each result is detailed in the figure legend below each figure.

Results

The presence of 3' or 5' overhangs inhibits the ATPase activity of INO80

INO80 plays a crucial and specific role in repairing double strand breaks and relieving stalled replication forks. Both of these genomic contexts share a unique feature: the presence of DNA

ends or overhangs. One model for how INO80 may achieve specificity at these sites is through recognition of these unique DNA structures. To test this possibility, we measured ATPase activity of INO80 in the presence of three different DNA substrates that have either blunted ends or single stranded overhangs (Appendix Figure 1A). We used saturating concentrations of these DNA substrates to measure to k_{cat} of ATP hydrolysis activity under multiple turnover conditions. To our surprise, we found that the presence of either a 3' or 5' overhang on the DNA substrate significantly inhibited the ATPase activity of INO80 by 10-15 fold. The extent of the inhibition was the same regardless of which type of overhang was present.

We reasoned that perhaps the single stranded overhang was inhibiting the ATPase activity of INO80 because single stranded DNA itself is inhibitory. To test this possibility, we performed similar experiments as above using single stranded DNA (either a 20mer or 40mer) as the sole substrate in the ATPase assay (Appendix Figure 1B). Although we were unable to achieve saturation with this substrate, it is clear that single stranded DNA on its own is still stimulatory, suggesting that there is a more complex reason for why the substrates containing overhangs were inhibitory to the INO80 ATPase.

Discrepancy between remodeling rates by gel and by FRET

A powerful strategy for dissecting the kinetic steps of nucleosome remodeling by INO80 is to obtain rate constants for remodeling using different remodeling assays. Each remodeling assay reports on a different set of kinetic steps leading to a remodeled nucleosome, allowing us to build these steps into an overall kinetic scheme for remodeling. Understanding the physical changes to the nucleosome that are required for an observable read-out for each remodeling assay allows us to further constrain the models based on our knowledge of nucleosome structure. To build upon the kinetic scheme we developed in the main chapter, we quantified rate constants of remodeling by gel and compared them to the rate constants we had obtained by FRET, for either +60 or +80 nucleosomes (Appendix Figure 2). For +60 nucleosomes, the

rate constants for remodeling were similar regardless of which assay was used. For +80 nucleosomes, however, remodeling by FRET was 20-fold faster than remodeling by gel. This result suggests that the two assays are capturing different steps in the remodeling reaction.

INO80 has a 5-fold preference for nucleosomes over DNA

Chromatin remodeling enzymes evolved from bacterial DNA motors to act on a more complex substrate—the nucleosome. As such, remodelers contain unique domains that recognize epitopes specific to the nucleosome. To begin test whether or not INO80 holds true to this principle, we compared ATP hydrolysis rate constants for nucleosomal DNA vs. the corresponding nucleosome. We found that under saturating conditions, nucleosomes consistently stimulated the ATPase activity of INO80 5-fold more than DNA. This result is at least consistent with INO80 having specific epitopes that recognize certain regions of the histones.

Discussion

Here we built upon the results presented in the main part of the chapter by further investigating the substrate specificity of INO80. Our findings and their implications are discussed below.

We found that nucleosomes are the preferred substrate of INO80 compared to naked DNA, as the ATPase activity of INO80 is consistently 5-fold higher for the nucleosome compared to the corresponding nucleosomal DNA (Appendix Figure 3). This trend is maintained even with nucleosomes that have longer flanking DNA. These results are somewhat consistent with published results for INO80 (1), and suggest that there are specialized motifs on the histones that are being specifically recognized the INO80 complex, and being communicated to the Ino80 ATPase. While the mechanisms behind this specificity remain to be determined, an exciting possibility is that INO80 may specifically recognize the H2A/H2B dimer. This hypothesis is supported by a few lines of evidence: (i) INO80 binds to a region of the nucleosome that is in

close proximity to the H2A/H2B dimer (Bartholomew lab, unpublished), (ii) INO80 has been shown to have dimer exchange activity (2), and (iii) the highly related remodeling complex belonging to the same remodeling family, SWR, contains several motifs that recognize the H2A.Z/H2B dimer specifically (3, 4).

In addition to measuring substrate specificity for nucleosomes over DNA, we also obtained some puzzling results for how INO80 treats different DNA end structures. First, we found that INO80's ATPase activity is inhibited by single stranded DNA overhangs. While it remains to be seen how this inhibition is exhibited in the context of a nucleosome, it does raise some interesting possibilities for how INO80 may be regulated *in vivo*. It has been shown that INO80 is involved in homologous recombination and in rescuing stalled replication forks (5), both places in the genome where single stranded DNA overhangs are found. Our results suggest that INO80 specifically recognizes these overhang structures and subsequently downregulates its ATPase activity as a result. It is tempting to speculate that the presence of the overhang actually regulates the way INO80 remodels chromatin at those sites, perhaps by generating a different set of products. Future work using nucleosomal substrates with DNA overhangs will be required to determine the true impact of the results shown here.

In another puzzling twist, we found that ratio between the rates of remodeling by gel vs. by FRET varies with DNA length. If we consider that a FRET change only requires some change in distance between the cy3 and cy5 probes while a change in mobility of the nucleosome by gel requires a complete change in positioning of the octamer on the DNA, this result suggests some uncoupling between two kinetic steps of remodeling when flanking DNA becomes longer than 60 base pairs. This interpretation becomes even more exciting in light of the possibility that INO80 may be able to exchange H2A.Z/H2B dimers for canonical H2A/H2B (2). INO80's published ability to exchange H2A.Z/H2B dimers for canonical H2A/H2B suggests that at the very least INO80 can loosen the contacts between dimer and tetramer, potentially forming a different type of product that may or may not be converted to a slid nucleosome. In

such a model, an additional level of flanking DNA length sensing that is coupled to product distribution. Further assay development would be required to test whether or not there is indeed a different intermediate being formed, whether or not the formation of this intermediate is sensitive to flanking DNA length, and whether the intermediate leads to a slid nucleosome.

References

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

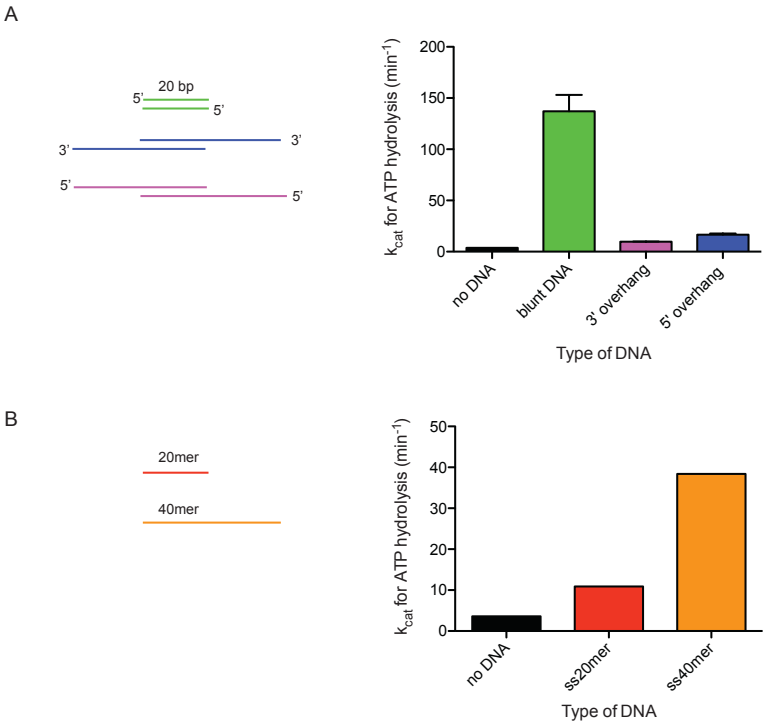
2. Papamichos-Chronakis, M., Watanabe, S., Rando, O. J., and Peterson, C. L. (2011) Global regulation of H2A.Z localization by the INO80 chromatin-remodeling enzyme is essential for genome integrity. *Cell* 144, 200–213

3. Wu, W.-H., Wu, C.-H., Ladurner, A., Mizuguchi, G., Wei, D., Xiao, H., Luk, E., Ranjan, A., and Wu, C. (2009) N terminus of Swr1 binds to histone H2AZ and provides a platform for subunit assembly in the chromatin remodeling complex. *J. Biol. Chem.* 284, 6200–6207

4. Wu, W.-H., Alami, S., Luk, E., Wu, C.-H., Sen, S., Mizuguchi, G., Wei, D., and Wu, C. (2005) Swc2 is a widely conserved H2AZ-binding module essential for ATP-dependent histone exchange. *Nat. Struct. Mol. Biol.* 12, 1064–1071

5. Morrison, A. J., and Shen, X. (2009) Chromatin remodelling beyond transcription: the INO80 and SWR1 complexes. *Nat. Rev. Mol. Cell Biol.* 10, 373–384

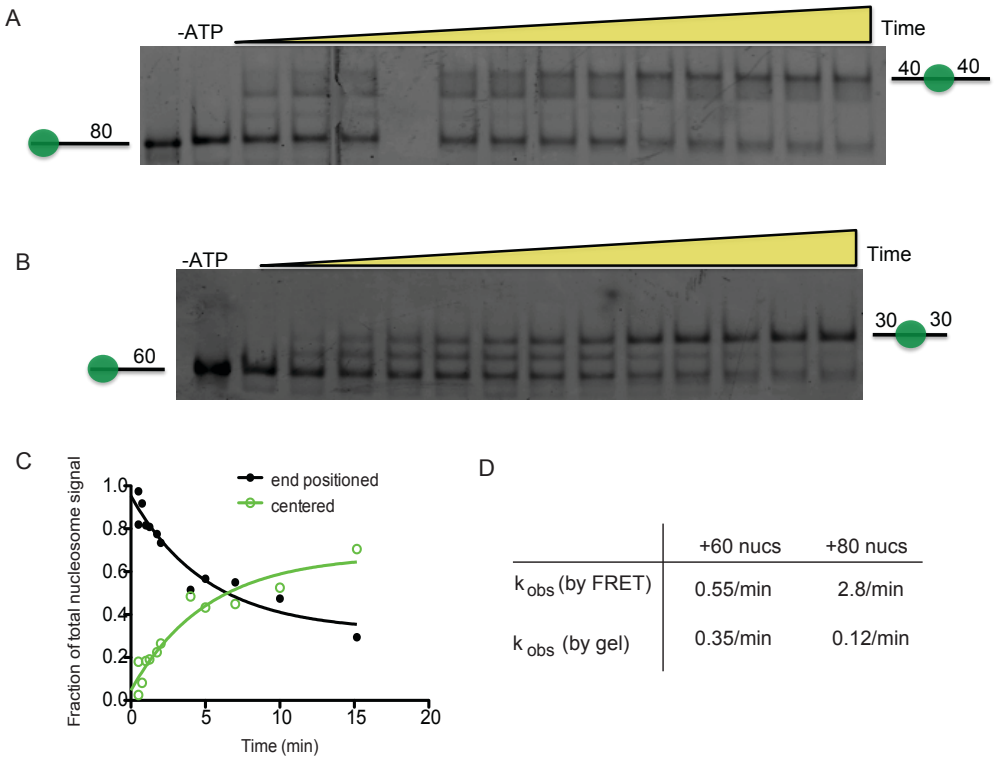
Appendix Figure 1: The presence of 3' or 5' overhangs inhibits the ATPase activity of INO80



(A) Left, different DNA substrates used in the ATPase assay results shown on the right. For each substrate, the duplex region corresponds to 20 base pairs of double stranded DNA. For the two substrates with overhangs, an extra 20 bases of DNA are added on either end. The conditions for the ATPases are multiple turnover, where [DNA] is in excess and saturating for [INO80].

(B) Left, different DNA substrates used in the ATPase assay results shown on the right. The conditions for the ATPases are multiple turnover, where [DNA] is in excess but not saturating for [INO80].

Appendix Figure 2: Discrepancy between remodeling rates by gel and by FRET



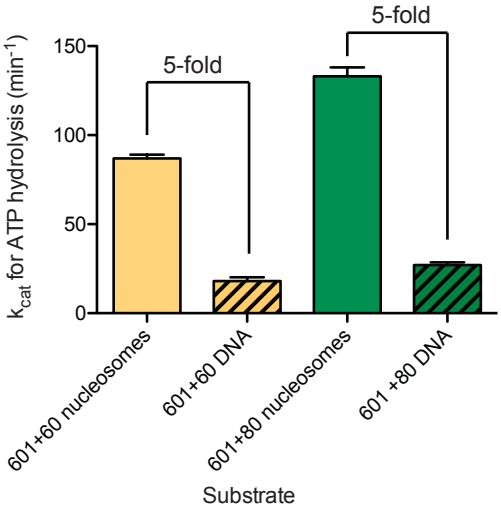
(A) Gel remodeling experiment with +80 nucleosomes, under single turnover conditions where [INO80] is in excess and saturating over nucleosomes.

(B) Gel remodeling experiment with +60 nucleosomes, under single turnover conditions where [INO80] is in excess and saturating over nucleosomes.

(C) An example fit of the raw data, where fraction of either the centered band or the end positioned band is plotted against time. The intensitites of the bands were quantified using ImageJ and the data were fit using a one-phase exponential function.

(D) Chart comparing the values for k_{obs} determined by FRET vs. gel for either +60 or +80 nucleosomes.

Appendix Figure 3: INO80 has a 5-fold preference for nucleosomes over DNA



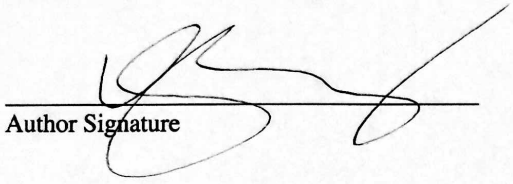
INO80 ATPase activity was stimulated by either nucleosomes or the corresponding nucleosomal DNA and k_{cat} for ATPase activity was measured. The nucleosome experiments were done under single turnover conditions, where [INO80] was in excess and saturating over [nucleosomes]. The DNA experiments were done under multiple turnover conditions, where [DNA] was in excess and saturating over [INO80].

Publishing Agreement

It is the policy of the University to encourage the distribution of all theses, dissertations, and manuscripts. Copies of all UCSF theses, dissertations, and manuscripts will be routed to the library via the Graduate Division. The library will make all theses, dissertations, and manuscripts accessible to the public and will preserve these to the best of their abilities, in perpetuity.

Please sign the following statement:

I hereby grant permission to the Graduate Division of the University of California, San Francisco to release copies of my thesis, dissertation, or manuscript to the Campus Library to provide access and preservation, in whole or in part, in perpetuity.



Author Signature

08/14/17

Date