# Chromatin remodelers couple inchworm motion with twist-defect formation to slide nucleosomal DNA

Giovanni B. Brandani, Shoji Takada†

Department of Biophysics, Graduate School of Science, Kyoto University, Kyoto 606-8502, Japan

† takada@biophys.kyoto-u.ac.jp

## Abstract

ATP-dependent chromatin remodelers control genome organization by repositioning, ejecting, or editing nucleosomes, activities that confer them essential regulatory roles on gene expression and DNA replication. Here, we investigate the mechanism of active DNA sliding by remodelers, which underlies most remodeling functions, via molecular dynamics simulations of the Snf2 remodeler in complex with a nucleosome. During its inchworm motion driven by ATP consumption, the remodeler overwrites the original nucleosome energy landscape via steric and electrostatic interactions to induce sliding of nucleosomal DNA. Repositioning is initiated at the remodeler binding location via the generation of twist defects, which then spontaneously propagate to complete sliding throughout the entire nucleosome. We also reveal how remodeler mutations and DNA sequence control active nucleosome repositioning. These results offer a mechanistic understanding of chromatin remodeling important for the interpretation of experimental studies.

## Introduction

Eukaryotic genomes are compacted into the cell nucleus via the formation of nucleosomes, each of them consisting of ~147 base pairs (bp) of DNA wrapping around a protein histone octamer[1]. After having been initially considered as passive building blocks of chromatin organization, nucleosomes became to be recognized as active regulators of DNA transcription and replication[2]. An origin of this regulation is the steric effect that inhibits other DNA-binding proteins, such as transcription factors, from accessing nucleosomal DNA[3], suggesting the requirement of fine control of nucleosome positioning along the genomic sequence[2,4]. For instance, repositioning of nucleosomes located next to transcription start sites enables the dynamic regulation of gene expression in response to stresses such as heat shock[5,6]. *In vitro*, nucleosome positions along DNA are determined by DNA mechanics[7]; in particular, sequence locations where the intrinsic bending of DNA lowers the free energy of nucleosome assembly are favored[8]. While the sequence still plays an important role *in vivo*[9], in this case the precise nucleosome positioning is controlled by the action of chromatin remodelers[2,4,10]. Remodelers are molecular motors that consume ATP to perform a wide variety of functions related to genome organization[11,12]: facilitating nucleosome assembly[13] and precise spacing[10,14], controlling DNA accessibility via nucleosome sliding[15] or histone ejection[16], and nucleosome editing via exchange between different histone variants[17]. These activities enable remodelers to maintain chromatin organization after disruptive events such as replication[2,18], and to regulate gene expression via the dynamic control of nucleosome positions[2,5,6,16,19].

Although the changes in chromatin organization induced by remodelers have been widely documented[2,10,14], the precise molecular mechanisms are still far from being clear[11,12,20]. The complexity comes in part from the existence of a wide variety of remodelers with different structures and functions[11], which has led to several possible classifications into remodeler sub-families[21]. Each remodeler consists of many distinct domains, which act in concert to confer specificity to the remodeling activity (e.g. nucleosome sliding vs ejection)[11,12] and to fine-tune it via substrate recognition (e.g. of histone tail modifications)[21,22]. Despite this complexity, all remodelers share a conserved translocase domain[11]: an ATPase motor capable of unidirectional sliding along DNA via binding and hydrolysis of ATP between its two RecA-like lobes, structurally similar to those found in helicases[11,12,21]. The translocase domain of most remodelers binds nucleosomes at the superhelical location (SHL) 2 [11,23,24], i.e. two DNA turns away from the dyad symmetry axis (SHL 0)[25] (Fig. 1a). Remodelers induce sliding of nucleosomal DNA towards the dyad from the translocase binding location[11,26]. This may represent a shared fundamental mechanism at the basis of all remodeling activities; the interactions with the additional domains would then confer specificity to the remodeler, allowing for substrate recognition and determining whether the final outcome is nucleosome repositioning, histone ejection or histone exchange[11].

There is much experimental evidence suggesting that the translocase domain in remodelers, as well as some helicases, slides unidirectionally along DNA via an inchworm mechanism[11,27] (sometimes referred as ratchet[24,28]), processing by 1 bp every ATP cycle. A minimal inchworm model requires 3 distinct chemical states, apo, ATP-bound, and ADP-bound, which are coupled to conformational changes of the translocase (Fig. 1b). The two lobes are either distant in the open form or in contact in the closed form. On top, conformational changes modulate interaction strengths with DNA[28]. To translocate along DNA in the direction from lobe 1 to lobe 2 (which would correspond to sliding nucleosomal DNA in the direction from SHL 2 towards the dyad, as in remodelers[23]), ATP binding to the translocase first induces the transition from an open to a closed conformation (see the second and third cartoons in Fig. 1b). During closure, lobe 2-DNA interactions are stronger than those of lobe 1, so that lobe 1 will detach from the DNA and move towards lobe 2 by 1 bp, which maintains its position. Then, ATP hydrolysis is accompanied by weakening of the lobe 2 interactions with DNA relative to those of lobe 1, so that during the conformational change from the closed to the open state lobe 2 now moves away from lobe 1 by 1 bp (fourth and fifth cartoons in Fig. 1b)[11]. The cycle is then completed by the release of ADP and the change of the interaction strengths to their initial apo-state values. This mechanism was firstly suggested for helicases from their crystal structures that show ATPase closure and changes in lobes-DNA interactions as a function of the chemical state[27,28]. For remodelers, the same mechanism was suggested based on the structural similarity to helicases[12,21,23]. Recent cryo-EM structures of nucleosome-bound remodelers in the open and closed conformations brought further evidence to this mechanism[23,24]. However, while the inchworm model can explain the motion of remodelers along naked DNA[29], its application to nucleosome repositioning is far from trivial, since such motion would eventually result into steric clashes between the remodeler and the histone octamer. Furthermore, complete nucleosome repositioning necessarily involves breakage of the many histone-DNA contacts that stabilize the nucleosome structure[25], and it is not clear how the remodeler may perturb the contacts far away from the binding location at SHL 2.

Insights from research on spontaneous nucleosome repositioning may shed light into the more complex active case[30]. Indeed, DNA sliding on nucleosomes can also simply proceed via transient structural changes

driven by thermal fluctuations[30,31]. Modes of nucleosome repositioning can be classified in two types depending on whether sliding is accompanied by the rotation of DNA around its axis[32]. In the rotation-uncoupled mode, sliding proceeds via large steps of about a DNA turn (~10 bp)[32–34], possibly facilitated by the formation of loops[33,34]. On the other hand, in the rotation-coupled mode, DNA sliding proceeds at small steps of 1 bp via a screw-like motion[32], facilitated by the formation of twist defects[1,35–39]. Twist defects are the structural deformations of DNA allowing to accommodate different numbers of base pairs between the strong histone-DNA contact points, which correspond to half-integer SHL locations[25,37]. Between two adjacent contact points, while canonical DNA turns contains ~10-bp, 9 bp (a missing bp) or 11 bp (an extra bp) may also be accommodated, corresponding to deformations referred to as under-twist and over-twist defects respectively (see an example illustrated in Fig. 1a, where the anti-clockwise motion of DNA at the contact point at SHL 1.5 (the red arrow) generates DNA under-twisting at SHL 2 (green) and DNA over-twisting at SHL 1 (brown)). The spontaneous formation and propagation of twist defects around the nucleosome causes repositioning by 1 bp at the time[36,38]. While some remodelers have been shown to induce the formation of large DNA loops in nucleosomes[40], this may be an effect due to extra domains rather than the origin of nucleosome sliding, since more recent experimental evidence, such as the small step size (~1 bp) of active repositioning[15,41], points towards the importance of DNA twisting in the molecular mechanism[20]. Furthermore, having the same characteristic step size, sliding via twist defects is more compatible with the inchworm mechanism of the translocase domain, with one possibility being that the conformational changes occurring during the ATP cycle initiate the formation and propagation of twist defects from the binding site at SHL 2 (Fig. 1a)[42].

Due to the ubiquitous importance of remodelers for chromatin organization, gene expression and replication[2], the detailed understanding of their molecular mechanism of action would be extremely valuable. Coarse-grained molecular dynamics (MD) simulations[43] represent an ideal tool for approaching this problem, since their resolution can be high enough to accurately represent the potential key steps occurring during active nucleosome repositioning[32,38,44], while achieving a speed-up of several orders of magnitude relative to all-atom simulations[43], for which the system size and the relevant time scales would result in an exceedingly large computational cost. Notably, coarse-graining approaches have been successfully applied to the study of nucleosome dynamics[8,45–47], including spontaneous repositioning[32,34,38], and ATP-dependent molecular motors[44,48,49].

In this work, we investigate the fundamental mechanism of active repositioning in a minimal system consisting of the ATPase-translocase domain of the Snf2 remodeler from yeast in complex with the nucleosome[23]. Firstly, we test that our model can reproduce the expected inchworm mechanism and unidirectional sliding during ATP consumption. By comparing to the spontaneous case, we show how the remodeler induces directed repositioning by modifying the nucleosome free energy landscape via steric effects and long-range electrostatic interactions, explaining past experimental data on Snf2 mutants[23]. Nucleosome repositioning occurs by coupling the ATPase inchworm motion to the formation and propagation of twist defects starting from the remodeler binding location at SHL 2. Finally, we reveal how DNA sequence can be exploited to control the kinetics of the system, consistently with its role in determining the repositioning outcome of many remodelers[10,50,51].

## Results

### MD simulations of remodeler translocase sliding on naked and nucleosomal DNA via an inchworm mechanism

We performed coarse-grained MD simulations of the ATP-dependent translocase domain of the Snf2 remodeler both on naked DNA and when bound to nucleosomes (Fig. 1c). The nucleosome model is the same as that previously employed to study spontaneous nucleosome repositioning[32,38], whereas the remodeler model and its interactions with the DNA are based on the cryo-EM structure of the Snf2-nucleosome complex[23]. Our computational model coarse-grains proteins at the level of individual residues[52] and DNA at the level of sugar, phosphate, and base groups, capturing the sequence-dependent flexibilities of base steps[53,54] (see Methods and Refs.[32,52,53] for more details). In most simulations, we used a 2-bp periodic sequence formed by repeating ApG base steps (polyApG). This sequence was chosen because it was shown to display an intermediate flexibility on the nucleosome[38]. We also investigated the effect on active repositioning due to strong nucleosome positioning sequences[9,55], and the introduction of sequence motifs with different flexibilities at target regions.

Based on the inchworm model (Fig. 1b), each remodeler chemical state (apo, ATP or ADP) corresponds to slightly different force-field parameters of the coarse-grained potential, and we simulate an ATP cycle (apo→ATP→ADP→apo) via switching the potential during the MD simulation, a common strategy in coarse-grained studies of molecular motors[48,49]. Initially, in the open apo state, the remodeler configuration and the strengths of the interactions between ATPase lobes and DNA are as found in the structure with PDB id 5X0Y[23], then switching to the ATP-bound potential enhances the attraction between the two lobes, favoring the closed conformation of the remodeler. ATP hydrolysis is emulated by switching to the ADP-bound potential, which reduces the lobe 2-DNA interactions by a factor of 0.8 and weakens the attractive interaction between the two lobes to favor the open conformation. In all our MD trajectories (40 on naked DNA, 100 on nucleosome), switching from apo to ATP states occurs at time 0 after $2x10^7$ MD equilibration steps in the open conformation, ATP hydrolysis occurs after $10^7$ MD steps, which are sufficient for the full relaxation of the system in the closed conformation, and finally switching back to the apo state occurs after $10^7$ steps (which are sufficient to observe translocase opening). More details on the simulation protocol are provided in the Methods section.

To analyze the repositioning dynamics in our MD trajectories, we track the motion of the DNA base pairs relative to the two individual ATPase lobes and relative to the histone octamer at the 14 histone-DNA contact points, located at the half-integer SHLs where the DNA minor groove faces the octamer. We refer to these collective variables as the contact indexes: $\Delta bp_{L1}$ and $\Delta bp_{L2}$ for the remodeler lobes and $\Delta bp_i$ for the histone contacts, where i is the half-integer valued SHL of the contact (these contacts will be indicated by their SHL value, e.g. contact point 1.5). As in our previous work[38], the nucleosome contact indexes are evaluated relative to the 147-bp conformation found in the crystal structure with PDB id 1KX5[56], which does not display twist defects. Using these contact indexes, we can fully characterize the remodeler's inchworm dynamics, the sliding of DNA in the nucleosome, and the potential role played by twist defects. These DNA deformations are distributed around integer-valued SHLs lacking direct histone-DNA contacts, and can either involve DNA over-twisting accommodating an extra base pair relative to the 1KX5 reference (over-twist defect), or DNA under-twisting accommodating a missing base pair (under-twist defect)[1,35,36,38,57]. A twist defect at SHL i can be simply

evaluated by the difference between the neighboring contact indexes (i-1/2 and i+1/2): a defect value close to zero corresponds to the standard non-defect case, a value close to 1 to an over-twist defect and a value close to -1 to an under-twist defect (see an example in Fig. 1a).

In this section, we present the simulation results of both Snf2-naked DNA and Snf2-nucleosome systems, but focusing on the motion of the remodeler relative to the DNA. In Fig. 2a, we show that by switching between the remodeler chemical states during the MD simulation, this can slide along both naked and nucleosomal DNA by 1-bp (as evidenced by the change in the average lobe contact index during the ATP cycle). While sliding on naked DNA is not a key function of remodelers, this process has been documented in experiments[29]. Interestingly, we note that under the current computational settings sliding by 1 bp by the end of an ATP cycle occurs with higher probability when in complex with nucleosomes (98%) than when on naked DNA (45%), whereas in the remaining cases the remodeler simply goes back to its original position.

Figure 2b displays two representative trajectories (one on the naked DNA and one on the nucleosome) projected onto the contact indexes of the two separate ATPase lobes. In both cases, this projection clearly highlights the inchworm motions. Specifically, starting from the apo state in the open conformation (bottom left in the figure), switching the potential to the ATP state induces the closure of the remodeler, with lobe 1 moving by 1 bp towards lobe 2, which maintains its position due to its stronger grip to the DNA (bottom right). Then, simulating ATP hydrolysis via switching to the ADP-state potential induces the domain opening, but since the lobe 2-DNA interactions are also decreased, now it is this lobe that usually moves by 1 bp away from lobe 1 (top right). Switching again to the apo-state potential simply restores the original lobes-DNA interaction strengths, maintaining the same open configuration and completing a full ATP cycle with the remodeler shifted by 1 bp relative to where it started. On naked DNA, this mechanism is sufficient to explain the translocase's unidirectional motion (see Supplementary Movie 1 for a visualization of the trajectory in Fig. 2b). However, what is not clear from this analysis is how the translocase motion may induce sliding of nucleosomal DNA; the next sections are devoted to the characterization of the complete active repositioning process.

**Remodelers couple inchworm motion to nucleosome sliding via steric and electrostatic interactions**

Our MD simulations show that the ATP-driven translocase closure is followed by sliding of nucleosomal DNA. Specifically, the DNA at the remodeler binding location slides unidirectionally towards the dyad, as indicated by the 1-bp increase in the average nucleosome contact index around SHL 2, $(\Delta bp_{1.5}+\Delta bp_{2.5})/2$ (see two representative trajectories and the cumulative distribution at the end of the ATP-state in Fig. 3a, upper panel). This direction of repositioning is consistent with the experimental evidence[23]. On the other hand, in the absence Snf2, sliding of the same polyApG nucleosomal DNA occurs in a random direction (Fig. 3a, lower panel).

To better characterize the origin of the unidirectional motion, in Fig. 3b, we compare the free-energy profiles of nucleosome sliding along the average contact index around SHL 2 for different scenarios. In the absence of Snf2, as expected for the uniform polyApG sequence and the random motion reported in Fig. 3a, sliding by 1 base pair in either direction does not change the free energy of the system, but it involves climbing significant free-energy barriers (~6 $k_BT$). The presence of the remodeler modifies the original nucleosome landscape in a chemical-state-dependent fashion. In the initial open conformation before ATP binding, there is a single free-energy minimum at $(\Delta bp_{1.5}+\Delta bp_{2.5})/2=0$, so that nucleosomal DNA sliding is strongly inhibited. Instead, after ATP binding and translocase closure, a second deeper free-energy minimum appears around

($\Delta bp_{1.5}+\Delta bp_{2.5}$)/2=1, favoring DNA sliding towards the dyad. After the last opening conformational change following ATP hydrolysis, the nucleosome landscape returns to have a single free-energy minimum now at ($\Delta bp_{1.5}+\Delta bp_{2.5}$)/2=1, so that further sliding is inhibited. The switching among these different free-energy landscapes reveals a clear ratchet mechanism, as often employed for the theoretical modeling of molecular motors[58]. The changes in the free energy profiles can be in part explained by the inchworm motion of the translocase domain and steric effects. In the open conformation, DNA sliding at SHL 2 by 1 bp in either direction would cause steric overlap between lobe 2 and histone octamer on one side or overlap between lobe 1 and the opposite DNA gyre around SHL -6 on the other side (see inset in Fig. 3b), blocking nucleosome repositioning and explaining the single free energy minimum when Snf2 is in the apo state. Since the translocase closure upon ATP binding involves the motion of lobe 1 towards lobe 2, DNA sliding is now allowed to occur towards the dyad, causing the translocase to swing on the opposite side of its binding location (see cartoons in Fig. 3a). However, this argument does not explain the large extent to which the closed ATPase favors unidirectional repositioning, i.e. the decrease in free energy by ~4 $k_BT$ from ($\Delta bp_{1.5}+\Delta bp_{2.5}$)/2=0 to 1.

From the crystal structure of the nucleosome-bound Snf2 remodeler[23], it was shown that apart from the main interactions at SHL 2, the translocase domain also interacts with the opposite DNA gyre around SHL -6 via long-range electrostatics mediated by residues K855, R880 and K885, located within lobe 1 (see bottom view in Fig. 1c and inset in Fig. 3b). It was also experimentally shown that changing these residues from positively- to negatively-charged markedly reduced the remodeling activity of Snf2[23]. To investigate this effect, we performed MD simulations where the three key residues have all been mutated to glutamic acid (K855E-R880E-K885E mutant). While still possible, DNA sliding around SHL 2 is no longer accompanied by a large decrease in free energy (Fig. 3b). This change can be understood in terms of the movement of the ATPase lobe 1 during repositioning. In the open state, lobe 1 is close to the contact point 1.5 and also interacts with the opposite gyre at SHL -6 via the basic patch in wild-type (WT) Snf2 (K855, R880 and K885). After ATP binding, lobe 1 moves by 1 bp towards lobe 2, becoming further apart from both contact point 1.5 and the DNA at SHL -6, weakening the electrostatic interaction (specifically, the average distance between the center of mass of the lobe 1 patch and the DNA phosphate backbone increases from ~6.1 Å to ~8.3 Å upon translocase closure). The sliding of nucleosomal DNA causes lobe 1 to swing back towards the dyad, restoring also the original interactions between the basic patch and SHL -6. Comparing to the initial open apo structure, the translocase closure and subsequent sliding of DNA at SHL 2 makes it appear that lobe 2 moved by 1 bp towards lobe 1, and not the opposite. Notably, this observation is consistent with the recent cryo-EM structure of the nucleosome-Chd1 complex in the presence of an ATP analog[24], where lobe 1 overlaps with the corresponding lobe in the open conformation of the Snf2 remodeler[23], whereas lobe 2 appears to have moved by 1 bp[24].

While so far we only considered a simple uniform polyApG sequence, genomes are rich in positioning motifs that contribute to specify the optimal location of nucleosomes along DNA[4]. These motifs, such as T/A base steps periodically spaced every 10 bp, cause the intrinsic bending of DNA, which lowers the free energy cost of nucleosome assembly, and favor a specific rotational setting, as they preferentially locate where the DNA minor groove faces the histone octamer[8]. These signals strongly inhibit nucleosome sliding relative to random DNA sequences, since repositioning would proceed either by DNA screw-like motion via a high-energy intermediate with a non-optimal rotational setting[36,38], or via alternative repositioning mechanisms uncoupled with DNA

rotations, which involve the energetically-costly breakage of many histone-DNA contacts[32,34]. Nevertheless, chromatin remodelers are still able to actively reposition nucleosomes made with strong positioning sequences such as 601[15,23,55].

To investigate the robustness of the active repositioning mechanism against changes in DNA sequence, we next run MD simulations of Snf2 in complex with nucleosomes made with the 601 sequence[55]. In the starting configuration, we shifted the DNA by 3 bp relative to the optimal configuration found in the 5X0Y structure, in the direction from the remodeler site towards the dyad. We refer to this sequence as $601_{\Delta 3}$. Because of the non optimal location of the T/A steps relative to the histone octamer (see cartoon in Fig. 3c), starting from here in the absence of the remodeler will be most likely followed by sliding backward away from the dyad, i.e. towards the optimal configuration (in about half of the cases within $10^7$ MD steps, Fig. 3c, lower panel). Instead, not only the remodeler prevents backward sliding, but upon ATP binding, in about half of the cases, it can also induce sliding of nucleosomal DNA forward towards the dyad (Fig. 3C, upper panel), in the same way as observed with the uniform polyApG sequence. A comparison of the free-energy landscapes along DNA sliding at SHL 2 with and without remodeler (Fig. 3d) shows indeed that in the case without remodeler the free energy strongly increases with sliding forward towards the dyad and decreases away from the dyad, whereas the closed translocase is able to lower the free energy cost of forward sliding to ~0 $k_BT$, while preventing sliding backward in the opposite direction via steric effects. The free energy profile obtained with the K855E-R880E-K885E Snf2 bound to $601_{\Delta 3}$ nucleosomes, shows that this mutant cannot slide these strong positioning sequences, due to an extra free energy penalty of ~3 $k_BT$ upon sliding by 1 bp. This is consistent with the results from experiments on similar Snf2 charge mutants sliding 601 nucleosomes[23]. While the limitations of our computational model (e.g. the assumptions on the precise ATP hydrolysis kinetics) prevent us from making quantitative predictions of remodeling activity, our simulations provide a mechanistic understanding of the important role of electrostatic interactions in directing repositioning[23].

**Nucleosome repositioning proceeds via the formation and propagation of twist defects from the remodeler binding location**

So far, we focused on the inchworm motion of the translocase domain and on nucleosomal DNA sliding at the SHL 2 binding site, establishing how these two are tightly coupled. However, a full characterization of the repositioning mechanism requires the analysis of DNA sliding at the individual histone-DNA contact points on the entire nucleosome. In Fig. 4a, we plot the timelines of the contact index coordinates of both remodeler and nucleosome for two representative trajectories during which repositioning by 1 bp occurs. These plots show how nucleosomal DNA sliding is initiated near the translocase binding location at the contact point at SHL 1.5, with the creation of opposite-type twist defects at the neighboring SHLs. The diffusion of these defects then completes repositioning of the entire nucleosome. To aid the understanding of the dynamics, we label the key metastable conformations of the system according to the following rules: the first letter, o or c, corresponds respectively to open or closed translocase conformation; when it is closed (c), the domain can adopt distinct configurations with DNA and histone octamer within its binding site at SHL 2, which will be indicated by a capital letter as A, B, C or D (see below for definition); finally, a last integer number, 0, 1 or 2, indicates the number of over-twist defects which may form near the dyad at the three central SHLs (SHL -1, 0, and +1, these defects are most favorably found at SHLs +/-1).

In the first trajectory (the left panel in Fig. 4a), starting from an open translocase bound to a nucleosome in a standard 1KX5-like configuration lacking twist defects (state o0), switching to the ATP-bound potential at time 0 quickly induces the closure of the remodeler via the motion of lobe 1 towards lobe 2 ($0.03 \times 10^6$ MD steps, state cA0). In this first closed configuration, the lobe 1-DNA interface is destabilized relative to the one observed in the reference 5X0Y structure and the motion towards lobe 2 is only partial ($\Delta bp_{L1}$~0.6). Only after some time ($0.09 \times 10^6$ MD steps) the motion of lobe 1 is complete ($\Delta bp_{L1}$~1, state cB0). From this closed configuration, we observe motion of nucleosomal DNA towards the dyad relative to the histone octamer starting from SHL 1.5 ($0.13 \times 10^6$ MD steps, state cC1). In this configuration, the DNA motion causes DNA over-twisting at SHL 1, which now accommodates an extra base pair, and DNA under-twisting at SHL 2 (where the remodeler is bound), which now accommodates a missing base pair. Soon afterwards ($0.2 \times 10^6$ MD steps), the nucleosomal DNA further slides by 1 bp from the remodeler site up to the closest nucleosome entry/exit, releasing the under-twist defect (state cD1). As highlighted in the previous section, while these two steps do not involve remodeler's motion relative to the DNA, the DNA motion relative to the histone octamer causes the remodeler to swing by 1 bp towards the dyad and enables to re-establish the electrostatic contacts between lobe 1 and SHL -6, which were lost during the initial ATPase closure. State cD1, for the polyApG sequence considered here, is the most stable configuration among the closed ones. Repositioning is usually completed only after ATP hydrolysis ($10 \times 10^6$ MD steps), which causes translocase opening via lobe 2 motion by 1 bp (state o1) (the full pathway is o0→ cA0→ cB0→ cC1→ cD1→ o1→ o0). The very last step consists of the sliding of nucleosomal DNA from the translocase up to the far nucleosome entry/exit, releasing the over-twist defect near the dyad (state o0, see Supplementary Movies 2 and 3 for visualizations of this trajectory). The second trajectory (the right panel) is qualitatively similar to the first, except that all states have an additional defect near the dyad at the starting time (the full pathway is then o1→ cA1→ cB1→ cC2→ cD2→ cD1→ o1 for trajectory 2). In particular, motion at the remodeler and nucleosome contact points proceeds in the same order. These two trajectories are representative of the most common repositioning pathways observed in our 100 MD trajectories. In all cases, repositioning involves twist-defect formation and propagation starting from the remodeler binding location.

Trajectories can be projected onto a low dimensional space defined by the sum of the contact indexes around the remodelers binding location ($\Delta bp_{L1}+\Delta bp_{L2}+\Delta bp_{1.5}+\Delta bp_{2.5}$, the horizontal axis in Fig. 3b) and by the size of the twist defect around the dyad ($\Delta bp_{1.5}-\Delta bp_{-1.5}$, the vertical axis in Fig. 4b). On this space, all the key metastable states involved in repositioning can be clearly separated (in Fig. 4b we show trajectories 1 and 2 from panel a). From this figure we notice that most key conformational changes occur in the closed conformation (between two dotted lines). To test the importance of the system relaxation in this portion of the phase space, we induced ATP hydrolysis after only $10^6$ MD steps from ATP binding (instead of $10^7$). Even if limiting our analysis to the trajectories where the remodeler successfully closed before hydrolysis (64%), only in 36% of the cases the remodeler can complete sliding by 1 bp, compared to the original 98% success rate (see Supplementary Fig. 1 for timeline of the contact indexes during one of these unsuccessful repositioning events). Failing to reposition occurs when the DNA does not have enough time to slide at the remodeler binding location (reaching state cD1), which is essential to avoid the steric overlap between lobe 2 and histones upon opening.

In Fig. 4c, we plot the changes in the average size of the twist defects at the SHLs near the remodeler during the ATP cycle. ATP binding and remodeler closure cause an enhancement of DNA over-twisting at SHL 1 and DNA under-twisting at SHL 2. Conversely, ATP hydrolysis and remodeler opening restore the former level of DNA twisting. Therefore, the closed remodeler lowers the free energy of opposite-type twist defects near its binding location at SHLs 1 and 2, which in turn will favor the initiation of DNA sliding towards the dyad starting from contact point 1.5, as observed in the trajectories in Fig. 4a.

**Defect-mediated repositioning is controlled by DNA sequence**

The twist-defect-mediated mechanism highlighted by our MD simulations may offer a further route to control the remodeling activity via DNA sequence (apart from the effects of positioning motifs described above), potentially explaining the significant sequence-dependence observed in experiments[10,32,34,38,50,51]. DNA sequence was already found to have a strong effect on the time-scales of spontaneous repositioning due to variations in DNA flexibility and energetics of twist defects[38]. When nucleosome diffusion is mediated by the formation of twist defects, repositioning proceeds much faster on sequences such as polyTpA, which are very flexible and easily accommodate twist defects, than on sequences such as poly(dA:dT) tracts (ApA repeats), which are stiffer and unlikely to display twist defects[38]. Perhaps surprisingly, it was shown that TpA repeats inhibit nucleosome repositioning by the Chd1 remodeler when the element is located at SHL 2[51]. The origin of the influence of sequence on active repositioning is not clear, and there is likely also a strong dependence on the considered remodeler itself, due to the many different domains that can interact with the nucleosome apart from the translocase[11]. We begin exploring the defect-mediated sequence effects on the behavior of our minimal remodeler system by considering targeted changes from reference polyApG sequence at specific regions: the insertion of a 10-bp poly(dA:dT) tract at SHL 1 (polyApG-$ApA_{SHL1}$ sequence), and a 10-bp TpA repeat at SHL 2 (polyApG-$TpA_{SHL2}$ sequence) (see Fig. 5a for the locations of these sequence elements within the Snf2-nucleosome complex). To characterize these effects, we reconstructed the free energy landscapes and kinetics of the systems via Markov state modeling (see Methods). In particular, we study the region of phase space where the translocase is in its closed (ATP-bound) configuration and lobe 1 fully completed the motion by 1 bp towards lobe 2, so that we can solely focus on the key formation and propagation of twist defects mediating repositioning.

In Fig. 5b, we display the free energy landscapes along the average contact index around SHL 2 (horizontal axis, $(\Delta bp_{1.5}+\Delta bp_{2.5})/2$, the same coordinate used to investigate sliding in Fig. 3) and the number of twist defects at the three SHLs around the dyad (the vertical axis, $\Delta bp_{1.5}-\Delta bp_{-1.5}$, which enables to track how repositioning proceeds further away from SHL 2) for the three DNA sequences. These landscapes are qualitatively similar and have several local minima corresponding to the twist-defect intermediates observed in our trajectories. Indeed, the representative trajectories reported in Supplementary Fig. 2 shows how the role of twist defects in mediating active repositioning is robust against changes in DNA sequence and remodeler mutations. However, the changes considered here do have significant effects on the kinetics of repositioning. For instance, if we focus on the pathway of twist-defect propagation observed in trajectory 2 in Fig. 4 (cB1→cC2→cD2→cD1, which is the most common for polyApG, highlighted in black in Fig. 5b), then the addition of the $ApA_{SHL1}$ and $TpA_{SHL2}$ elements greatly increase the free energies of the intermediate states along the minimum energy pathways (Fig. 5c). This change is also evidenced by an increase in the mean first

passage time to reach the final cD1 state (by a factor of ~2.4 and ~1.3 respectively). These changes are explained by the different DNA flexibilities and by the changes in defect pattern along the pathway of repositioning. ApA repeats where previously found to inhibit DNA over-twisting around the dyad, whereas TpA repeats where found to inhibit DNA under-twisting at SHLs +/-2[38]. Since states cB1 and cD1 have only one defect around the dyad (at SHL +1 or -1), whereas the intermediate states cC2 and cD2 have two defects (at both SHLs +1 and -1), the $ApA_{SHL1}$ element increases the free energy of these two intermediate states. Similarly, despite its flexibility, the $TpA_{SHL2}$ element significantly increases the free energy of the cC2 state, where DNA under-twisting at SHL 2 is observed. Therefore, the sequence-dependent formation and propagation of twist defects during remodeling provides an explanation of past experimental observations[51].

Interestingly, we note that the repositioning activity may also depend on the DNA sequence at the translocase SHL 2 binding site via changes in the relative strength of lobe 1 and lobe 2 interactions with DNA. In particular, while the initial translocase closure occurring upon ATP binding always involves the motion of lobe 1 towards lobe 2 for polyApG, polyApG-$ApA_{SHL1}$ and $601_{\Delta 3}$, in ~28% of the polyApG-$TpA_{SHL2}$ simulations lobe 2 moves towards lobe 1, due to the weaker lobe 2-DNA interactions. When this occurs, there is no sliding of nucleosomal DNA, since the original lobe 1 electrostatic interactions with SHL -6 remain unaffected, and the translocase opening following ATP hydrolysis simply brings the system back to its starting configuration. This sequence effect resulting in lower remodeling activity on TpA repeats at SHL 2 is also consistent with past experiments[51].

## Discussion

The recent structural information of the Snf2-nucleosome complex[23] and insights from related molecular motors[27,28] allowed us to design an efficient computational model that reproduces the expected inchworm motion of the translocase domain along DNA[11]. More importantly, our MD simulations revealed the detailed molecular mechanism by which the inchworm motion of the translocase is converted into nucleosome repositioning. Specifically, we establish the fundamental role of electrostatic interactions and nucleosome twist defects. The interaction between a basic patch located on lobe 1 and the DNA gyre around SHL -6 acts as an electrostatic spring to direct the sliding of nucleosomal DNA from the translocase binding site at SHL 2 towards the dyad. Firstly, ATP binding drives the closure of the translocase via the motion of lobe 1 towards lobe 2 by 1 bp. This initial motion increases the distance between lobe 1 and SHL -6, weakening the electrostatic interactions. The sliding of nucleosomal DNA towards the dyad then causes the translocase to swing towards dyad together with DNA, re-establishing the original electrostatic contacts. The steric repulsion between the remodeler and the histone octamer allows nucleosomal DNA sliding to occur in the closed conformation only, and prevents sliding in the opposite direction away from the dyad. This mechanism is robust enough to enable active repositioning of nucleosomes formed with strong positioning sequences such as 601[55], where sliding has to proceed against a high free-energy uphill due to the preferred rotational register of T/A positioning motifs. Consistent with previous experiments, our MD simulations of Snf2 mutants show that the basic patch on lobe 1 is essential for the ability to slide 601 nucleosomes[23].

The electrostatics-driven sliding of nucleosomal DNA starts from the histone-DNA contact point at SHL 1.5, via the formation of a pair of twist defects of opposite type at the neighboring SHLs: DNA under-twisting at SHL

2 accommodates the missing base pair, DNA over-twisting at SHL 1 accommodates the extra base pair. The importance of twist defects lies in facilitating the initiation of nucleosomal DNA sliding locally from the translocase binding site, without simultaneously breaking distant histone-DNA contacts. The initial paired twist-defect conformation represents an intermediate state with tension accumulated around the remodeler location. Release of this tension via the propagation of the two twist defects in opposite directions completes the sliding of DNA by 1 bp throughout the entire nucleosome. The ATP cycle is completed after hydrolysis and ADP release, during which the motion of the weakened lobe 2 away from lobe 1 brings the translocase in the initial open conformation, only with the nucleosomal DNA slid by 1 bp relative to both histone octamer and remodeler. Interestingly, the twist defects that favor active nucleosome repositioning are the same that are commonly found in nucleosome crystal structures (at SHL 2)[1,25], and those that have been shown to play an important role in spontaneous repositioning (SHL 1 and 2)[38], suggesting a reason for targeting the specific translocase binding location.

Our results are consistent with the wave-ratchet-wave model of nucleosome repositioning[11,26], where sliding is initiated at the translocase site via the generation of tension (DNA twisting). This has been suggested as a unifying fundamental mechanism of chromatin remodeling, where additional domains control tension release and interactions with other nucleosome regions to perform substrate recognition and to define the remodeling outcome (e.g. sliding vs histone ejection vs histone exchange)[11]. While we considered a minimal remodeler system consisting of nucleosome and translocase domain only, the molecular mechanism highlighted here may offer an understanding of more complex scenarios. For instance, the ISWI remodeler was found to slide nucleosomes via coordinated 3-bp entry and 1-bp exit steps[41]. These findings may be explained by our model if the additional HAND-SANT-SLIDE domain of ISWI blocks DNA sliding on the entry side until enough tension from under-twist defects is accumulated (3 bp), while on the exit side DNA sliding is unrestrained and it proceeds at 1-bp steps via the release of over-twist defects[11]. Although our model is currently not well tested for the study of histone conformational changes, the strong accumulation of tension from DNA twisting on the entry side may be the origin of the histone octamer distortions observed during remodeling by ISWI[59]. Future studies should also address how remodelers may couple sliding of nucleosomal DNA to histone exchange[60].

Finally, our simulations reveal how the twist-defect mechanism allows a kinetic control of nucleosome repositioning via DNA sequence. This is due to the significant sequence-dependence in the formation of the twist-defect intermediates that facilitate repositioning. In our previous work[38], we showed how the free-energy cost of DNA over-twisting and under-twisting changes as a function of superhelical location (SHL) and sequence. With this knowledge, it is possible to target precise remodeling intermediates to control nucleosome sliding. Specifically, we tested the effect of introducing a 10-bp TpA repeat at SHL 2, which inhibits under-twist defects at this location, and a 10-bp poly(dA:dT) tract at SHL 1 ($ApA_{SHL1}$), which inhibits over-twist defects. Consistent with the important role of twist defects at these regions, both DNA sequence modifications slow-down sliding relative to the pure polyApG case, due to the increase in the free-energy barrier along the repositioning pathway. Notably, our TpA simulations explain recent experiments showing that the addition of this sequence element blocks repositioning by the Chd1 remodeler at SHL 2[51]. These results may be considered surprising, because TpA base steps are known for their high flexibility[54] and expected to favor nucleosome sliding. Indeed, polyTpA sequences were found to be among those with the highest spontaneous

nucleosome diffusivity[38]. However, even in that case, DNA sliding at SHL 2 represented the main kinetic bottleneck to sliding due to the high twist-defect cost[38], further supporting the tight relationship between spontaneous and active scenarios. Many experiments have also highlighted the role of poly(dA:dT) tracts in controlling remodeling outcomes[10,50,51]. Of note, while these sequences are well known for inhibiting their wrapping into nucleosomes[61], recent experiments suggested that the increase in the free energy of nucleosome assembly is not sufficient to explain the formation of nucleosome free regions[50], and that these are instead created via the direct action of chromatin remodelers[10,50]. Our molecular dynamics simulations reveal the detailed molecular mechanism by which DNA sequence changes affecting twist-defect formation may be exploited to control remodeling activities.

## Methods

Our coarse-grained model employs the AICG2+ structure-based force field for proteins[52] and the 3SPN.2C force field for DNA[53]. This combination has been successfully applied in many past studies of nucleosome dynamics[8,32,34,38,45,46], and it enables a good compromise between computational speed-up and accuracy, necessary to reach the time-scales relevant to nucleosome repositioning[32,34,38]. According to these models, each amino acid is coarse-grained to a single bead located at the corresponding Cα atom[52], whereas each nucleotide is represented by 3 beads corresponding to sugar, phosphate and base groups[53]. The reference native conformations of the histone octamer and the Snf2 ATPase are respectively taken from the nucleosome crystal structure with PDB id 1KX5[56] and the nucleosome-Snf2 cryo-EM structure with PDB id 5X0Y[23]. Histone tails and remodeler disordered regions not visible in the reference structures are modeled according to a sequence-dependent local statistical potential[62]. The DNA model was parametrized against several experimental thermodynamic quantities, such as melting temperature,[53] and the sequence-dependent elasticity of DNA[54]. Notably, this model has been successfully employed to predict the role of DNA sequence on the free energy of nucleosome assembly[8] and the formation of twist defects[38], suggesting its suitability for investigating the sequence-dependent effects present in chromatin remodeling[10,50,51].

Histone octamer, remodeler and DNA interact via excluded volume, long-range electrostatics and hydrogen-bonds. For excluded volume, we employ bead-type dependent radii derived from a database of protein-protein and protein-DNA complexes[32,38,63]. Following our previous protocol[32,38], electrostatics is modeled according to Debye-Hückel theory, with standard unit charges placed on DNA phosphate groups and protein residues in flexible regions, and fractional charges on protein residue in folded regions derived using the RESPAC method, which optimizes the coarse-grained electrostatic potential against the all-atom one in the folded conformation[64]. For intra-DNA electrostatics, phosphate charges are rescaled by 0.6 to implicitly account for counter ion condensation[53]. In all simulations the salt concentration was set to 250 mM of monovalent ions. Histone-DNA hydrogen bonds are modeled using a recently-developed distance- and orientation-dependent potential between protein residues and phosphate groups[32,38]; the potential, unlike a Go-like contacts[45], is invariant under a rotation-coupled motion of DNA, allowing the study of nucleosome repositioning. These hydrogen bonds are defined from those observed in the 1KX5[56] and 3LZ0[65] crystal structures[32,38], and we employed the same hydrogen bond strength $\varepsilon_{HB}$ = 1.8 $k_BT$ used in Ref.[38], which is an intermediate value among those giving a nucleosome disassembly profile consistent with experiments[32]. Properly accounting for hydrogen-bond

interactions is necessary for reproducing the experimental twist-defect metastability playing an important role in spontaneous nucleosome repositioning[36,38].

Translocase-DNA contacts are identified from the protein hydrogen donors within 5 Å from phosphate-group oxygen acceptors in the 5X0Y Snf2-nucleosome structure, and are represented with the same potential used for histone-DNA hydrogen bonds (reference distances and angles are also taken from 5X0Y). With this choice, lobe 1 (residue id 743-940) and lobe 2 (residue id 1046-1212) have each 16 contacts with DNA. These interactions are not strictly speaking only hydrogen bonds, but contacts introduced in the coarse-grained model to specify the remodeler-DNA binding mode. The relatively high cutoff is necessary to ensure that the motion of the remodeler lobes relative to the DNA occurs only during the translocase opening or closure steps induced during the ATP cycle, since spontaneous remodeler motion in the absence of ATP is inconsistent with the inchworm mechanism suggested from experiments[11,27,28]. To model the remodeler in the closed conformation adopted upon ATP binding, we created additional lobe 1-lobe 2 Go-like contacts identified from the Cα atoms within 7.5 Å in a homology model generated after aligning the two Snf2 lobes to the corresponding lobes found in the closed conformation of the NS3 helicase with PDB id 3KQU[28]. This reference structure was chosen because it was shown that the apo-state open conformations of Snf2[23] (PDB id 5X0Y) and NS3[28] (PDB id 3KQH) display a high degree of structural similarity[23].

As described at the beginning of the Results section, we simulate the remodeling activity of Snf2 during a full ATP cycle by switching the underlying potential to model changes in the translocase chemical state, as previously done for modeling other protein motors[48,49]. Each MD simulation cycle consists of $2x10^7$ MD equilibration steps in the apo-state, $10^7$ steps in the ATP-bound state, and $10^7$ steps in the ADP-bound state. In the apo state, translocase-DNA contacts have the same strength as histone-DNA hydrogen bonds ($\varepsilon_{1\text{-DNA}} = \varepsilon_{2\text{-DNA}} = 1.8\ k_BT$) and lobe 1-lobe 2 contacts are very weak ($\varepsilon_{1\text{-}2} = 0.17\ k_BT$), so that the other native interactions based on the 5X0Y reference dominate and the translocase is preferably in its open conformation. In the ATP-bound state, translocase-DNA interactions remain unchanged, but the interactions between lobe 1 and lobe 2 are strengthened ($\varepsilon_{1\text{-}2} = 1.0\ k_BT$) to stabilize the closed conformation. After ATP hydrolysis, lobe 2-DNA interactions are weakened by a factor of 0.8 ($\varepsilon_{2\text{-DNA}} = 1.44\ k_BT$). Completing the cycle requires a last switch to the apo-state potential, but this only changes the translocase-DNA interactions, without inducing any conformational change, therefore this step is not shown in the results. With these settings, we generate with very high probability successful nucleosome repositioning by 1 bp in the correct direction[23] (DNA sliding from the remodeler site towards the dyad) via the inchworm mechanism suggested from past experimental studies[11,27,28]. Variations to this simulation protocol, such as using the recent nucleosome-Chd1 cryo-EM structure to model the ATP-bound closed conformation[24] or changes to the Snf2-DNA interactions, have been also tested, but they do not affect the key features of the active repositioning mechanism mediated by twist defects presented in the Results section.

MD simulations have been performed using the software CafeMol 3.0[66] (available at http://www.cafemol.org/), integrating the equations of motion using the default settings via Langevin dynamics at 300 K. For the cases where the remodeler is in complex with the nucleosome, we ran 100 MD simulation cycles for each system, whereas for the remodeler sliding on naked DNA we ran 40 cycles. In the former cases, nucleosomal DNA is made by 223 bp (central 147 bp plus 38 bp for each linker), whereas in the latter by 40 bp. Nucleosome-bound

simulations start from the conformation observed in the 5X0Y structure[23] after a short energy minimization with the steepest descent method.

As in our previous work[38], we follow nucleosome repositioning in our MD trajectories by tracking the DNA base pair indexes relative to a reference conformation at the contact points with the proteins (14 histone contacts plus 1 contact for each ATPase lobe). We refer to these as the contact indexes Δbp and they are computed from the system coordinates[38] using path collective variables[67] (for the 14 nucleosome contacts we use the PDB structure 1KX5 lacking twist defects as a reference; for the two ATPase lobes we use the PDB structure 5X0Y). To characterize nucleosome repositioning induced by the remodeler in the ATP state (Figs. 3 and 5), we calculated free energy landscapes and kinetics via Markov state modeling (MSM)[68]. To this aim, clustering of the conformations was performed via the Density Peak algorithm[69], computing the local density on the space defined by the contact indexes of the ATPase lobes and the nucleosome at SHLs -1.5, 1.5 and 2.5 ($\Delta bp_{L1}$, $\Delta bp_{L2}$, $\Delta bp_{-1.5}$, $\Delta bp_{1.5}$ and $\Delta bp_{2.5}$), using a Gaussian kernel of 0.1 bp. These coordinates are sufficient to describe the slow formation of nucleosome twist defects and translocase motion that controls the repositioning, as DNA motion at other regions is faster[38]. The clustering correctly identifies the metastable conformations corresponding to the local free energy minima in the landscapes of Fig. 4b and Fig. 5b. The Markov state models were generated using the software PyEMMA[70] with a lag time of $2.5 \times 10^5$ MD steps, after which the estimates of the relaxation time scales of the models remain nearly constant (Supplementary Fig. 3). For the MSMs of the remodeler (wild type or mutant) on $601_{\Delta 3}$ nucleosomes, the 100 unbiased trajectories were sufficient to obtain a connected set of transitions covering the phase space. For all other systems, due to the higher free energy differences, it was necessary to run 10 additional $10^8$-MD-steps simulations with an extra harmonic potential along the Snf2-DNA rmsd to explore the full phase space, and we obtained the free energy landscapes and kinetics of the unbiased systems using the transition-based reweighting analysis method (TRAM)[71] for multiensemble Markov models. The free energy profiles reported in Figs. 3 and 5 were generated by reweighting the populations of the MSM clusters along the corresponding collective variables. Errors on the free energy profiles were evaluated by computing the standard deviation of the profiles obtained from distinct sets of trajectories, giving in all cases errors within ~0.3 $k_BT$. The free energy profiles of spontaneous repositioning in the absence of Snf2 (in Fig. 3) were generated by reweighting the conformations observed in $8 \times 10^8$-MD-steps simulations with a linear bias along the average contact index at SHL 2 and two harmonic walls to prevent sliding by further than 1.5 bp in either direction away from $(\Delta bp_{1.5}+\Delta bp_{2.5})/2=0$. The minimum-energy paths in Fig. 5 were determined via the nudged elastic band optimization method[72].

## Funding

This work was supported by JSPS KAKENHI grants [25251019 to S.T., 16KT0054 to S.T., and 16H01303 to S.T.], by the MEXT as “Priority Issue on Post-K computer” (S.T.), and by the RIKEN Pioneering Project “Dynamical Structural Biology” (S.T.). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

## References


1. Richmond, T. J. & Davey, C. A. The structure of DNA in the nucleosome core. *Nature* **423,** 145–150 (2003).

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

3. Field, Y. *et al.* Distinct Modes of Regulation by Chromatin Encoded through Nucleosome Positioning Signals. *PLoS Comput. Biol.* **4,** e1000216 (2008).

4. Struhl, K. & Segal, E. Determinants of nucleosome positioning. *Nat. Struct. Mol. Biol.* **20,** 267–273 (2013).

5. Shivaswamy, S. *et al.* Dynamic Remodeling of Individual Nucleosomes Across a Eukaryotic Genome in Response to Transcriptional Perturbation. *PLoS Biol.* **6,** e65 (2008).

6. Reja, R., Vinayachandran, V., Ghosh, S. & Pugh, B. F. Molecular mechanisms of ribosomal protein gene coregulation. *Genes Dev.* **29,** 1942–54 (2015).

7. Morozov, A. V *et al.* Using DNA mechanics to predict in vitro nucleosome positions and formation energies. *Nucleic Acids Res.* **37,** 4707–4722 (2009).

8. Freeman, G. S., Lequieu, J. P., Hinckley, D. M., Whitmer, J. K. & de Pablo, J. J. DNA Shape Dominates Sequence Affinity in Nucleosome Formation. *Phys. Rev. Lett.* **113,** 168101 (2014).

9. Segal, E. *et al.* A genomic code for nucleosome positioning. *Nature* **442,** 772–778 (2006).

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

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

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

13. Torigoe, S. E., Urwin, D. L., Ishii, H., Smith, D. E. & Kadonaga, J. T. Identification of a Rapidly Formed Nonnucleosomal Histone-DNA Intermediate that Is Converted into Chromatin by ACF. *Mol. Cell* **43,** 638–648 (2011).

14. Gkikopoulos, T. *et al.* A Role for Snf2-Related Nucleosome-Spacing Enzymes in Genome-Wide Nucleosome Organization. *Science* **333,** 1758–1760 (2011).

15. Harada, B. T. *et al.* Stepwise nucleosome translocation by RSC remodeling complexes. *Elife* **5,** 339–346 (2016).

16. Boeger, H., Griesenbeck, J., Strattan, J. S. & Kornberg, R. D. Removal of Promoter Nucleosomes by Disassembly Rather Than Sliding In Vivo. *Mol. Cell* **14,** 667–673 (2004).

17. Bruno, M. *et al.* Histone H2A/H2B Dimer Exchange by ATP-Dependent Chromatin Remodeling Activities. *Mol. Cell* **12,** 1599–1606 (2003).

18. Vasseur, P. *et al.* Dynamics of Nucleosome Positioning Maturation following Genomic Replication. *Cell Rep.* **16,** 2651–2665 (2016).

19. Lorch, Y. & Kornberg, R. D. Chromatin-remodeling and the initiation of transcription. *Q. Rev. Biophys.* **48,** 465–470 (2015).

20. Mueller-Planitz, F., Klinker, H. & Becker, P. B. Nucleosome sliding mechanisms: new twists in a looped history. *Nat. Struct. Mol. Biol.* **20,** 1026–32 (2013).

21. Flaus, A. & Owen-Hughes, T. Mechanisms for ATP-dependent chromatin remodelling: the means to the end. *FEBS J.* **278,** 3579–3595 (2011).

22. Blossey, R. & Schiessel, H. Kinetic proofreading of gene activation by chromatin remodeling. *HFSP J.* **2,** 167–170 (2008).

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

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

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

26. Saha, A., Wittmeyer, J. & Cairns, B. R. Chromatin remodeling through directional DNA translocation from an internal nucleosomal site. *Nat. Struct. Mol. Biol.* **12,** 747–755 (2005).

27. Velankar, S. S., Soultanas, P., Dillingham, M. S., Subramanya, H. S. & Wigley, D. B. Crystal Structures of Complexes of PcrA DNA Helicase with a DNA Substrate Indicate an Inchworm Mechanism. *Cell* **97,** 75–84 (1999).

28. Gu, M. & Rice, C. M. Three conformational snapshots of the hepatitis C virus NS3 helicase reveal a ratchet translocation mechanism. *Proc. Natl. Acad. Sci. U. S. A.* **107,** 521–8 (2010).

29. Sirinakis, G. *et al.* The RSC chromatin remodelling ATPase translocates DNA with high force and small step size. *EMBO J.* **30,** 2364–72 (2011).

30. Blossey, R. & Schiessel, H. The dynamics of the nucleosome: thermal effects, external forces and ATP. *FEBS J.* **278,** 3619–3632 (2011).

31. Meersseman, G., Pennings, S. & Bradbury, E. M. Mobile nucleosomes--a general behavior. *EMBO J.* **11,** 2951–2959 (1992).

32. Niina, T., Brandani, G. B., Tan, C. & Takada, S. Sequence-dependent nucleosome sliding in rotation-coupled and uncoupled modes revealed by molecular simulations. *PLOS Comput. Biol.* **13,** e1005880 (2017).

33. Schiessel, H., Widom, J., Bruinsma, R. F. & Gelbart, W. M. Polymer Reptation and Nucleosome Repositioning. *Phys. Rev. Lett.* **86,** 4414–4417 (2001).

34. Lequieu, J., Schwartz, D. C. & de Pablo, J. J. In silico evidence for sequence-dependent nucleosome sliding. *Proc. Natl. Acad. Sci. U. S. A.* **44,** E9197–E9205 (2017).

35. van Holde, K. E. & Yager, T. D. in *Structure and Function of the Genetic Apparatus* 35–53 (Springer US, 1985). doi:10.1007/978-1-4684-5024-8_4

36. Kulić, I. M. & Schiessel, H. Chromatin Dynamics: Nucleosomes go Mobile through Twist Defects. *Phys. Rev. Lett.* **91,** 148103 (2003).

37. Edayathumangalam, R. S., Weyermann, P., Dervan, P. B., Gottesfeld, J. M. & Luger, K. Nucleosomes in Solution Exist as a Mixture of Twist-defect States. *J. Mol. Biol.* **345,** 103–114 (2005).

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

39. Shaytan, A. K. *et al.* Coupling between Histone Conformations and DNA Geometry in Nucleosomes on a Microsecond Timescale: Atomistic Insights into Nucleosome Functions. *J. Mol. Biol.* **428,** 221–237 (2016).

40. Zhang, Y. *et al.* DNA Translocation and Loop Formation Mechanism of Chromatin Remodeling by SWI/SNF and RSC. *Mol. Cell* **24,** 559–568 (2006).

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

42. Holde, K. van & Yager, T. Models for chromatin remodeling: a critical comparison. *Biochem. Cell Biol.* **81,** 169–172 (2003).

43. Takada, S. *et al.* Modeling Structural Dynamics of Biomolecular Complexes by Coarse-Grained Molecular Simulations. *Acc. Chem. Res.* **48,** 3026–3035 (2015).

44. Flechsig, H. & Mikhailov, A. S. Tracing entire operation cycles of molecular motor hepatitis C virus helicase in structurally resolved dynamical simulations. *Proc. Natl. Acad. Sci. U. S. A.* **107,** 20875–80 (2010).

45. Kenzaki, H. & Takada, S. Partial Unwrapping and Histone Tail Dynamics in Nucleosome Revealed by Coarse-Grained Molecular Simulations. *PLoS Comput. Biol.* **11,** e1004443 (2015).

46. Chang, L. & Takada, S. Histone acetylation dependent energy landscapes in tri-nucleosome revealed by residue-resolved molecular simulations. *Sci. Rep.* **6,** 34441 (2016).

47. Zhang, B., Zheng, W., Papoian, G. A. & Wolynes, P. G. Exploring the Free Energy Landscape of Nucleosomes. *J. Am. Chem. Soc.* **138,** 8126–8133 (2016).

48. Koga, N. & Takada, S. Folding-based molecular simulations reveal mechanisms of the rotary motor F 1 –ATPase. *Proc. Natl. Acad. Sci.* **103,** 5367–5372 (2006).

49. Yao, X.-Q., Kenzaki, H., Murakami, S. & Takada, S. Drug export and allosteric coupling in a multidrug transporter revealed by molecular simulations. *Nat. Commun.* **1,** 117 (2010).

50. Lorch, Y., Maier-Davis, B. & Kornberg, R. D. Role of DNA sequence in chromatin remodeling and the formation of nucleosome-free regions. *Genes Dev.* **28,** 2492–7 (2014).

51. Winger, J. & Bowman, G. D. The Sequence of Nucleosomal DNA Modulates Sliding by the Chd1 Chromatin Remodeler. *J. Mol. Biol.* **429,** 808–822 (2017).

52. Li, W., Wang, W. & Takada, S. Energy landscape views for interplays among folding, binding, and allostery of calmodulin domains. *Proc. Natl. Acad. Sci. U. S. A.* **111,** 10550–10555 (2014).

53. Freeman, G. S., Hinckley, D. M., Lequieu, J. P., Whitmer, J. K. & de Pablo, J. J. Coarse-grained modeling of DNA curvature. *J. Chem. Phys.* **141,** 165103 (2014).

54. Olson, W. K., Gorin, A. A., Lu, X. J., Hock, L. M. & Zhurkin, V. B. DNA sequence-dependent deformability deduced from protein-DNA crystal complexes. *Proc. Natl. Acad. Sci. U. S. A.* **95,** 11163–11168 (1998).

55. Lowary, P. . & Widom, J. New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning. *J. Mol. Biol.* **276,** 19–42 (1998).

56. Davey, C. A., Sargent, D. F., Luger, K., Maeder, A. W. & Richmond, T. J. Solvent Mediated Interactions in the Structure of the Nucleosome Core Particle at 1.9Å Resolution. *J. Mol. Biol.* **319,** 1097–1113 (2002).

57. Richmond, T. J. & Widom, J. in *Chromatin structure and gene expression* **35,** 1–19 (Oxford University Press, UK, 2000).

58. Jülicher, F., Ajdari, A. & Prost, J. Modeling molecular motors. *Rev. Mod. Phys.* **69,** 1269–1282 (1997).

59. Sinha, K. K., Gross, J. D. & Narlikar, G. J. Distortion of histone octamer core promotes nucleosome mobilization by a chromatin remodeler. *Science.* **355,** eaaa3761 (2017).

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

61. Segal, E. & Widom, J. Poly(dA:dT) tracts: major determinants of nucleosome organization. *Curr. Opin. Struct. Biol.* **19,** 65–71 (2009).

62. Terakawa, T. & Takada, S. Multiscale ensemble modeling of intrinsically disordered proteins: p53 N-terminal domain. *Biophys. J.* **101,** 1450–1458 (2011).

63. Tan, C., Terakawa, T. & Takada, S. Dynamic Coupling among Protein Binding, Sliding, and DNA Bending Revealed by Molecular Dynamics. *J. Am. Chem. Soc.* **138,** 8512–8522 (2016).

64. Terakawa, T. & Takada, S. RESPAC: Method to Determine Partial Charges in Coarse-Grained Protein Model and Its Application to DNA-Binding Proteins. *J. Chem. Theory Comput.* **10,** 711–721 (2014).

65. Vasudevan, D., Chua, E. Y. D. & Davey, C. A. Crystal Structures of Nucleosome Core Particles Containing the '601' Strong Positioning Sequence. *J. Mol. Biol.* **403,** 1–10 (2010).

66. Kenzaki, H. *et al.* CafeMol: A Coarse-Grained Biomolecular Simulator for Simulating Proteins at Work. *J. Chem. Theory Comput.* **7,** 1979–1989 (2011).

67. Branduardi, D., Gervasio, F. L. & Parrinello, M. From A to B in free energy space. *J. Chem. Phys.* **126,** 54103 (2007).

68. Prinz, J.-H. *et al.* Markov models of molecular kinetics: Generation and validation. *J. Chem. Phys.* **134,** 174105 (2011).

69. Rodriguez, A. & Laio, A. Clustering by fast search and find of density peaks. *Science* **344,** 1492–6 (2014).

70. Scherer, M. K. *et al.* PyEMMA 2: A Software Package for Estimation, Validation, and Analysis of Markov Models. *J. Chem. Theory Comput.* **11,** 5525–5542 (2015).

71. Wu, H., Paul, F., Wehmeyer, C. & Noé, F. Multiensemble Markov models of molecular thermodynamics and kinetics. *Proc. Natl. Acad. Sci. U. S. A.* **113,** E3221-30 (2016).

72. Sheppard, D., Terrell, R. & Henkelman, G. Optimization methods for finding minimum energy paths. *J. Chem. Phys.* **128,** 134106 (2008).

## Figures

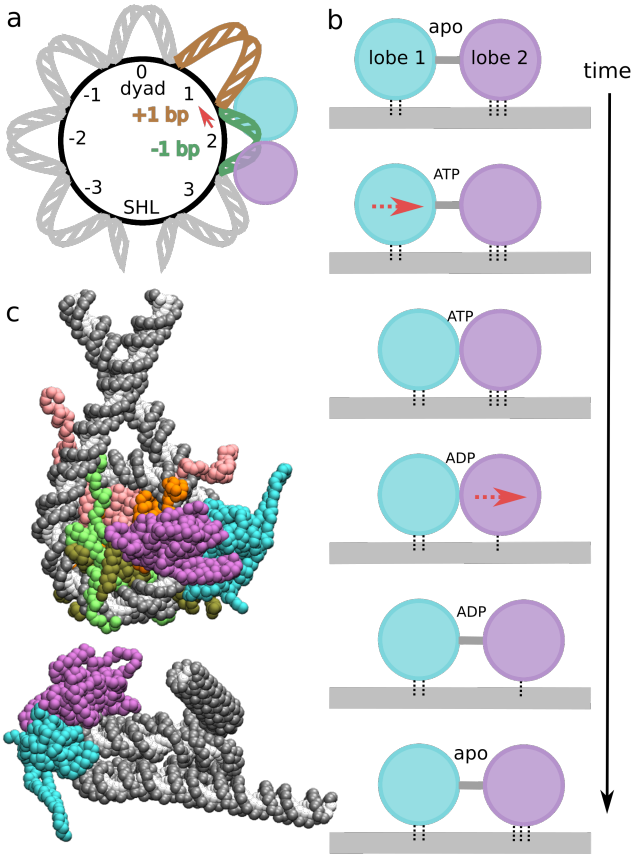


Fig. 1: Structure of the translocase-nucleosome complex. (a) Cartoon where nucleosome regions are indicated by the number of DNA turns from the dyad symmetry axis: superhelical location (SHL) 0 corresponds to the dyad, the ATPase domain (lobe 1 in cyan, lobe 2 in purple) binds DNA at SHL 2, the strong histone-DNA contact points are located at the half-integer SHLs where the DNA minor groove faces the octamer (for clarity, we only depict the region from SHL -4 to +4). To analyze DNA sliding, we track the base pair indexes $\Delta bp_i$ at these contact points relative to the initial conformation. If, for example, nucleosome sliding starts with the motion of DNA at contact point 1.5 by 1 bp towards the dyad (red arrow), the contact point $\Delta bp_{1.5}$ will increase from ~0 to ~1, and this will also indicate the formation of an over-twist defect at SHL 1 (brown) and an under-twist defect at SHL 2 (green), which now accommodate respectively an extra and a missing base pair relative to the reference nucleosome conformation in the crystal structure with PDB id 1KX5. (b) Inchworm mechanism of translocase motion along DNA. ATP binding induces a conformational change from open to closed, with the motion of lobe 1 (cyan) towards lobe 2 (purple) by 1 bp, due to the weaker DNA contacts of the former. ATP hydrolysis weakens the lobe 2-DNA contacts and induces opening via the motion of lobe 2 away from lobe 1 by 1 bp. ADP release completes the cycle. (c) Two views of the initial Snf2-nucleosome structure for our coarse-grained MD simulations: DNA backbone (phosphate and sugar groups) in gray, bases in white, translocase lobes in cyan and purple, histones H3 in pink, H4 in orange, H2A and H2B in light and dark green respectively. In the lower picture, histones are not drawn for clarity.

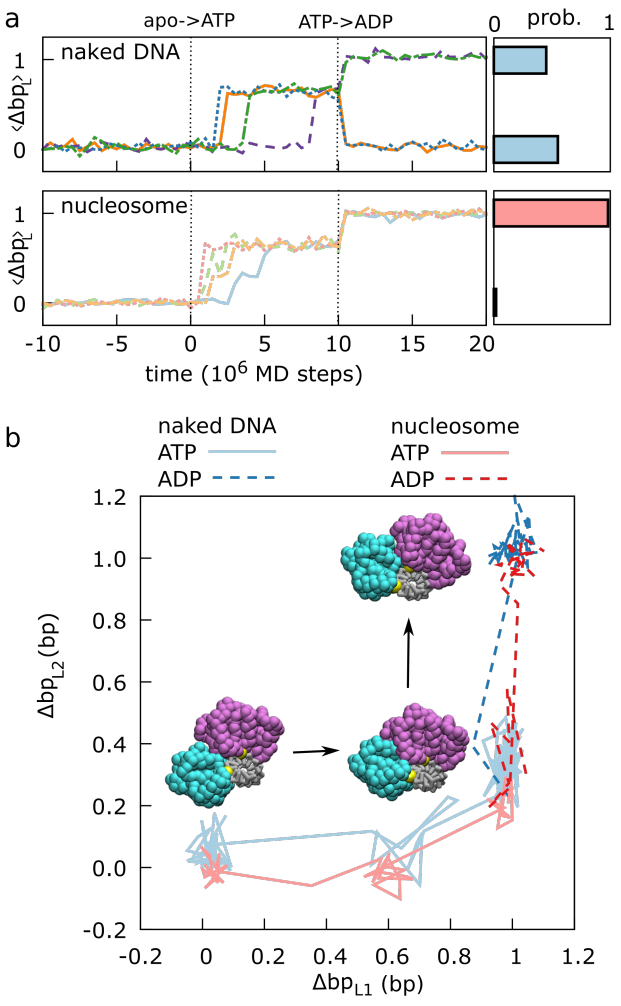


Fig. 2: Snf2 translocase motion relative to DNA via an inchworm mechanism. (a) Representative trajectories of the average lobe contact index $<\Delta bp_L>=(\Delta bp_{L1}+\Delta bp_{L2})/2$ during the ATP cycle in our MD simulations (ATP binding occurs after equilibration at time 0, hydrolysis occurs after $10^7$ MD steps) on naked (upper left) and nucleosomal (lower left) DNAs with polyApG sequence. In both cases we find unidirectional motions; a 1 bp step of the translocase in direction from lobe 1 to lobe 2 occurs every ATP cycle with probabilities of ~0.45 and ~0.98 on naked DNA (upper right) and on nucleosomal DNA (lower right), respectively. (b) Projections of two representative trajectories (blue for naked, red for nucleosomal DNAs) on the lobe contact indexes $\Delta bp_{L1}$ and $\Delta bp_{L2}$ (lighter solid lines for the ATP state, darker dashed lines for the ADP state after hydrolysis), highlighting the inchworm mechanism. Snapshots of translocase moving on naked DNA are also shown (lobe 1 in cyan, lobe 2 in purple, DNA in gray, two reference phosphates in yellow).

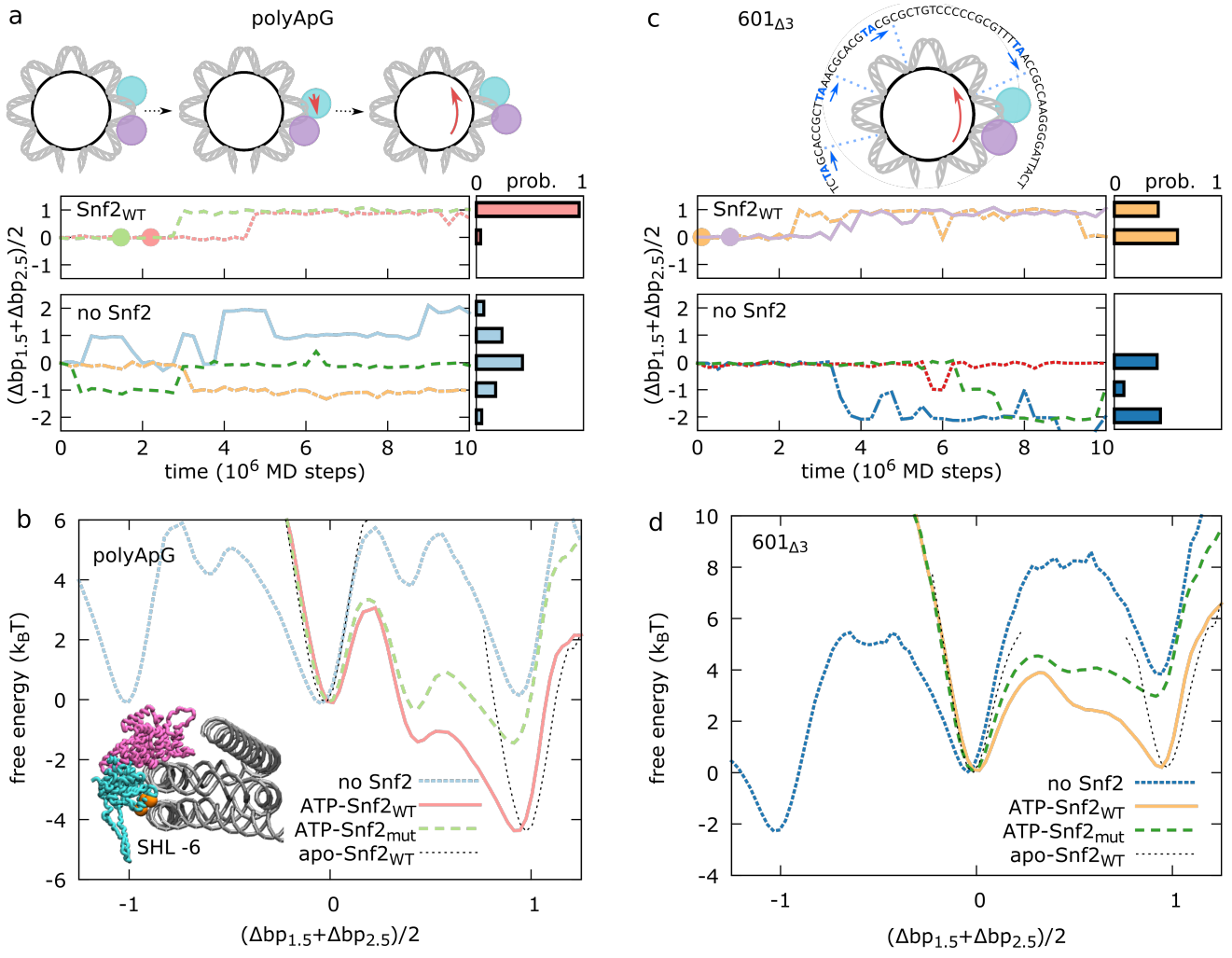
a
polyApG
Snf2WT
no Snf2
prob.
(Δbp1.5+Δbp2.5)/2
time (10^6 MD steps)
b
polyApG
SHL -6
no Snf2
ATP-Snf2WT
ATP-Snf2mut
apo-Snf2WT
free energy (kBT)
(Δbp1.5+Δbp2.5)/2
c
601Δ3
Snf2WT
no Snf2
prob.
(Δbp1.5+Δbp2.5)/2
time (10^6 MD steps)
d
601Δ3
no Snf2
ATP-Snf2WT
ATP-Snf2mut
apo-Snf2WT
free energy (kBT)
(Δbp1.5+Δbp2.5)/2

Fig. 3: Mechanism of directed nucleosomal DNA sliding by Snf2. (a,c) For polyApG (a) and $601_{\Delta 3}$ (c) sequences, we show representative trajectories of the average contact index at SHL 2, $(\Delta bp_{1.5}+\Delta bp_{2.5})/2$, as a function of time (left panel) and the repositioning histogram at the end of 100 trajectories of $10^7$ MD steps (right) for the remodeler-bound case (upper), and in the absence of remodeler (spontanous sliding) (lower). For the remodeler case, we report only the portion of trajectory in the ATP state, where nucleosome sliding occurs after the translocase closure (indicated by the circles). (b,d) For polyApG (b) and $601_{\Delta 3}$ (d) sequences we report the free energy profiles along the average contact index at SHL 2 for different systems: spontaneous sliding in the absence of the remodeler (no Snf2), in the complexes with the bound open apo-state Snf2 (apo-$Snf2_{WT}$), the closed ATP-bound Snf2 (ATP-$Snf2_{WT}$), and the closed ATP-bound K855E-R880E-K885E charge mutant Snf2 (ATP-$Snf2_{mut}$). Errors on the free-energy profiles, which are not explicitly displayed for clarity, are on the order of ~0.3 $k_BT$. The cartoons in panel a indicate how the inchworm motion of the translocase (lobe 1 in cyan, lobe 2 in purple) is coupled to nucleosome sliding (key translocase and DNA motions in red). The structure of the Snf2-nucleosome complex in the inset in panel b highlights the location of the K855, R880 and K885 residues (orange) targeted by the mutation and their interaction with the DNA gyre at SHL -6. In panel c we also displayed the central 63 base pairs of the $601_{\Delta 3}$ sequence highlighting the location of the TpA positioning steps within the Snf2-nucleosome complex (in blue). In the initial configuration for the $601_{\Delta 3}$ simulations, the DNA is shifted by 3 bp relative to the optimal structure with PDB id 5X0Y (with the TpA steps located where the DNA minor groove faces the histone octamer, indicated with a blue dashed line).

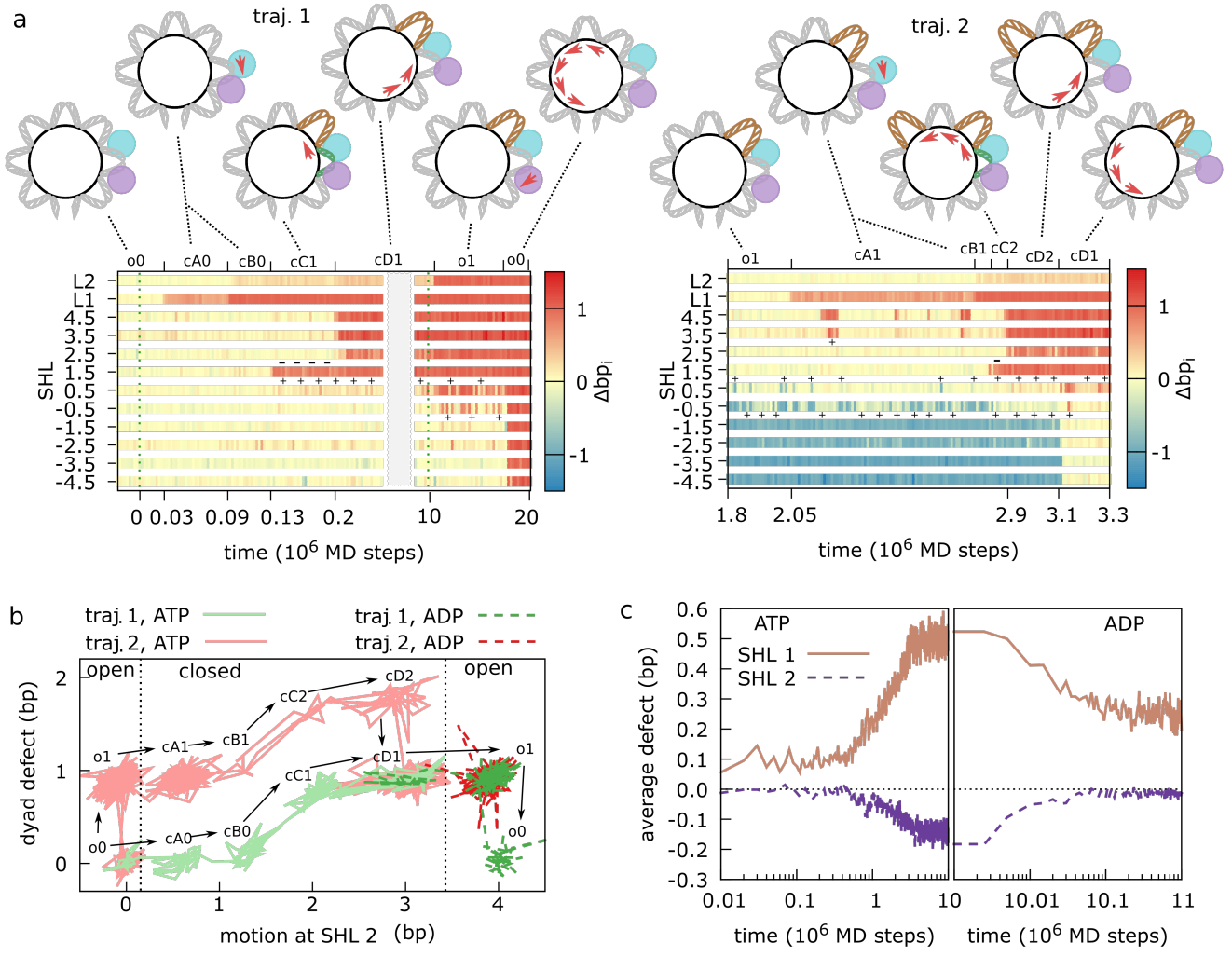
a
traj. 1
traj. 2
o0
cA0
cB0
cC1
cD1
o1
o0
cA1
cB1
cC2
cD2
cD1
SHL
L2
L1
4.5
3.5
2.5
1.5
0.5
-0.5
-1.5
-2.5
-3.5
-4.5
Δbp_i
1
0
-1
0 0.03 0.09 0.13 0.2 10 20
1.8 2.05 2.9 3.1 3.3
time (10^6 MD steps)
b
traj. 1, ATP
traj. 2, ATP
traj. 1, ADP
traj. 2, ADP
open
closed
open
dyad defect (bp)
motion at SHL 2 (bp)
c
ATP
ADP
SHL 1
SHL 2
average defect (bp)
0.01 0.1 1 10 10.01 10.1 11
time (10^6 MD steps)

Fig. 4: Active nucleosome repositioning via twist defect propagation for polyApG sequence. (a) Timelines of the translocase (L1 and L2) and nucleosome contact indexes (SHL -4.5 to 4.5) for two representative trajectories where nucleosomal DNA slides by 1 bp relative to the initial configuration. These plots highlight how the remodeler conformational changes (closure and opening) occurring during the ATP cycle (ATP binding at time 0, hydrolysis after $10^7$ MD steps) are coupled to nucleosomal DNA sliding via the formation and propagation of twist defects. The intermediate configurations along the repositioning pathway are indicated by the corresponding labels (described in the main text) and cartoons. The key twist defects facilitating repositioning are highlighted by a plus sign for DNA over-twisting (in brown in the cartoons) and by a minus sign for DNA under-twisting (in green in the cartoons). DNA and translocase lobes (1 in cyan, 2 in purple) motions are indicated by red arrows. (b) 2-dimensional projections of the trajectories in panel a (traj. 1 in green, traj. 2 in red; lighter solid lines for the ATP state, darker dashed lines after hydrolysis; time increases in the direction indicated by the arrows). The x-axis represents the sum of the translocase and nucleosome contact indexes around the remodeler binding location at SHL 2: $\Delta bp_{L1}+\Delta bp_{L2}+\Delta bp_{1.5}+\Delta bp_{2.5}$. The y-axis represents the size of the twist defects at the three central SHLs: $\Delta bp_{1.5}-\Delta bp_{-1.5}$. The intermediate configurations along the pathway are labeled as in panel a. (c) Twist defects at SHLs 1 (brown, solid line) and 2 (purple, dashed) averaged over 100 MD trajectories as a function of time, showing how twist defects are formed after ATP binding and translocase closure (left panel), and how they are released after hydrolysis and translocase opening (right panel).

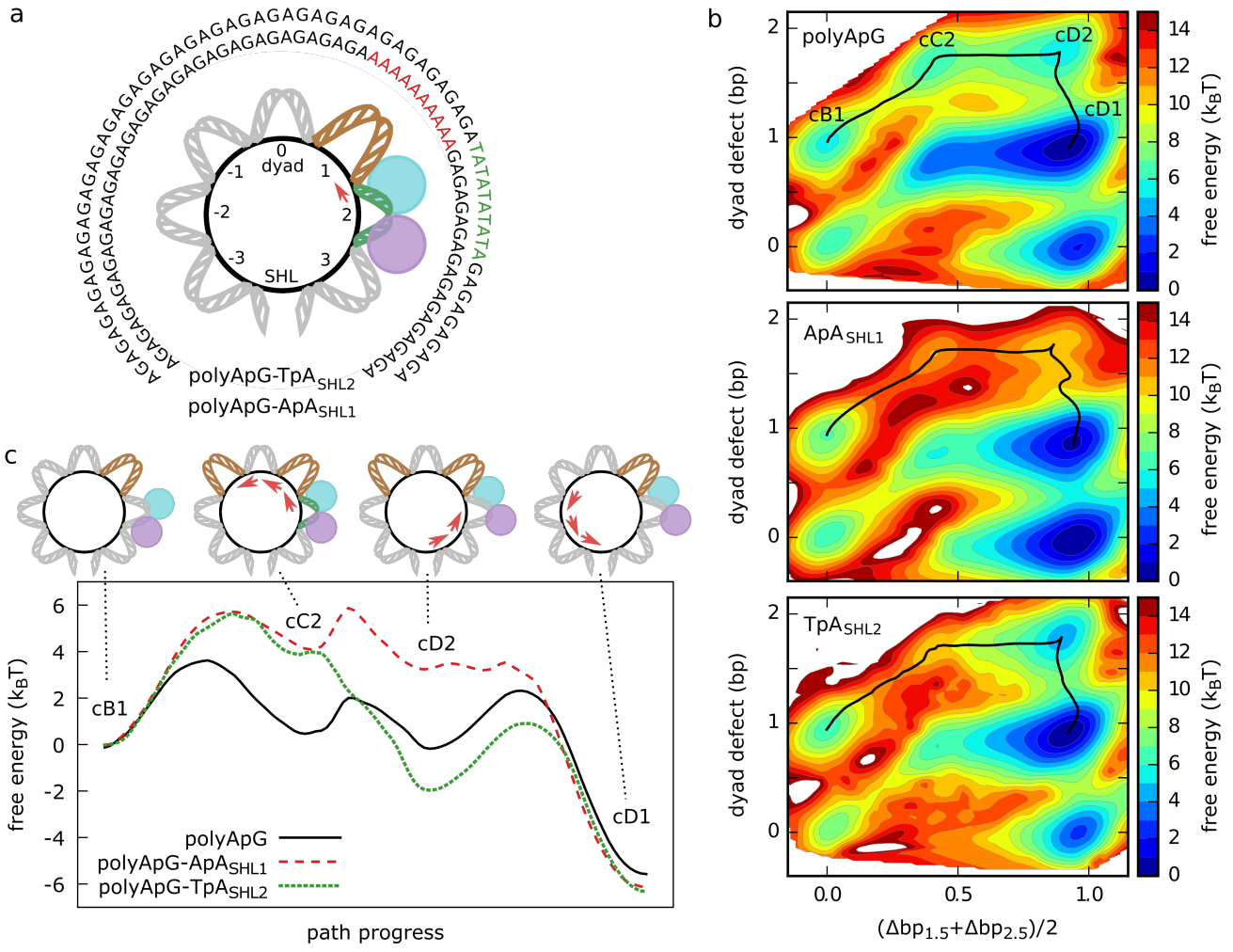


Fig. 5: Twist-defect-mediated effects due to DNA sequence. (a) Considered DNA sequences apart from the reference polyApG: polyApG-ApA$_{SHL1}$ (inner) and polyApG-TpA$_{SHL2}$ (outer). We displayed the central 71 base pairs highlighting their location relative to the translocase-nucleosome complex; note the poly(dA:dT) tract at SHL 1 (ApA, in red) and the TpA repeat at SHL 2 (in green). (b) Free energy landscapes of the Snf2-nucleosome complex in the closed conformation along the average contact index at SHL 2, ($\Delta bp_{1.5}+\Delta bp_{2.5}$)/2, and the size of the twist defects around the dyad ($\Delta bp_{1.5}-\Delta bp_{-1.5}$), for the three DNA sequences polyApG (top), polyApG-ApA$_{SHL1}$ (center) and polyApG-TpA$_{SHL2}$ (bottom). (c) Comparison of the free energy profiles along the minimum paths of repositioning from state cB1 to state cD1 (corresponding to the representative trajectory 2 in Fig. 3a) among the 3 sequences; polyApG (black, solid line), polyApG-ApA$_{SHL1}$ (red, dashed), and polyApG-TpA$_{SHL2}$ (green, dotted). The pathways are indicated on the landscapes in panel b by black solid lines. Errors on the profiles are within ~0.3 $k_BT$.