The EMBO Journal (2011) 30, 2596–2609 | © 2011 European Molecular Biology Organization | Some Rights Reserved 0261-4189/11
www.embojournal.org

THE EMBO JOURNAL

EMBO open

The DNA-binding domain of the Chd1 chromatin-remodelling enzyme contains SANT and SLIDE domains

Daniel P Ryan[1], Ramasubramanian Sundaramoorthy[1], David Martin[2], Vijender Singh[1] and Tom Owen-Hughes[1,*]

[1]Wellcome Trust Centre for Gene Regulation and Expression, College of Life Sciences, University of Dundee, Dundee, UK and [2]Biological Chemistry and Drug Discovery, College of Life Sciences, University of Dundee, Dundee, UK

The ATP-dependent chromatin-remodelling enzyme Chd1 is a 168-kDa protein consisting of a double chromodomain, Snf2-related ATPase domain, and a C-terminal DNA-binding domain. Here, we show the DNA-binding domain is required for *Saccharomyces cerevisiae* Chd1 to bind and remodel nucleosomes. The crystal structure of this domain reveals the presence of structural homology to SANT and SLIDE domains previously identified in ISWI remodelling enzymes. The presence of these domains in ISWI and Chd1 chromatin-remodelling enzymes may provide a means of efficiently harnessing the action of the Snf2-related ATPase domain for the purpose of nucleosome spacing and provide an explanation for partial redundancy between these proteins. Site directed mutagenesis was used to identify residues important for DNA binding and generate a model describing the interaction of this domain with DNA. Through inclusion of Chd1 sequences in homology searches SLIDE domains were identified in CHD6–9 proteins. Point mutations to conserved amino acids within the human CHD7 SLIDE domain have been identified in patients with CHARGE syndrome.

The EMBO Journal (2011) **30,** 2596–2609. doi:10.1038/emboj.2011.166; Published online 27 May 2011
Subject Categories: chromatin & transcription; structural biology
Keywords: Chd1; DNA binding; nucleosomes; SANT; SLIDE

Introduction

Packaging of eukaryotic DNA into nucleosomes and higher order chromatin structures is fundamental for the efficient housing of genomes within nuclei. However, this packaging poses a barrier to genetic processes such as transcription, replication, DNA repair, and recombination. The controlled manipulation of chromatin structure offers an effective means of regulating these processes. This regulation is achieved through an assortment of mechanisms including the modification of histone proteins (including methylation, acetylation, phosphorylation, ubiquitination/sumoylation, and ADP ribosylation) (Kouzarides, 2007), methylation of DNA (Bird, 2002), changes in the histone content of chromatin involving histone variants and their chaperones (Park and Luger, 2008), and the ATP-dependent remodelling of nucleosomes via Snf2 proteins (Clapier and Cairns, 2009).

Snf2 proteins contain a helicase-like ATPase domain (Snf2 domain) that consists of two RecA folds and is related to that found in all members of superfamily 2 helicase-related proteins (Flaus *et al*, 2006). The Snf2 domain is thought to facilitate ATP-dependent translocation of Snf2 proteins along DNA, which generates force that can be channelled into the movement or destabilisation of nucleosomes (Cairns, 2007).

Despite the strong conservation of the Snf2 domain among Snf2 proteins there is a broad spectrum of output among the different family members. Some Snf2 proteins disrupt nucleosomes by repositioning and/or dissociation (RSC and SWI/SNF), others generate ordered arrays of nucleosomes (Chd1 and ISWI), whereas others facilitate the exchange of histones/histone variants within a nucleosome (INO80 and SWR1) or act on non-nucleosomal substrates (Mot1) (Flaus *et al*, 2006; Cairns, 2007; Clapier and Cairns, 2009). These differences may in part be explained by sequence differences within the Snf2 domain that are conserved within distinct subfamilies of Snf2 proteins and which appear to correlate with known biochemical activities of the subfamilies. However, it is more likely that accessory domains and/or interaction partners have a much larger role in directing the mechanical output of the Snf2 proteins. Many Snf2 proteins reside in large multi-protein complexes, for example, the yeast RSC and SWI/SNF complexes consist of ~15 and ~12 subunits, respectively, and as such there is much scope for tuning the Snf2 domains that lie at the core of these complexes. In contrast, the *Saccharomyces cerevisiae* protein Chd1 appears to act as a monomer (Tran *et al*, 2000) and dissecting how the accessory domains within this protein direct the activity of its Snf2 domain is a more tractable problem.

Chd1, or chromodomain helicase DNA-binding protein 1, is conserved from yeast to humans. Chd1 proteins have been associated with an assortment of activities, including the efficient assembly and spacing of nucleosomes (Robinson and Schultz, 2003; Lusser *et al*, 2005; Stockdale *et al*, 2006).

*Corresponding author. Division of Gene Regulation and Expression, Wellcome Trust Centre for Gene Regulation and Expression, College of Life Sciences, University of Dundee, Dow Street, Dundee DD1 5EH, UK. Tel.: +44 138 238 5796; Fax: +44 138 234 8072;
E-mail: t.a.owenhughes@dundee.ac.uk

Received: 24 February 2011; accepted: 21 April 2011; published online: 27 May 2011

In *S. cerevisiae*, Chd1 is partially redundant with the action of other spacing enzymes (Tsukiyama *et al*, 1999). In *Drosophila melanogaster*, Chd1 is important for the appropriate deposition of the histone variant H3.3 (Konev *et al*, 2007); but in contrast in *Schizosaccharomyces pombe*, Chd1 is proposed to remove nucleosomes from DNA (Walfridsson *et al*, 2005). Assembly and removal of nucleosomes are likely to occur during transcription and DNA replication and numerous studies link Chd1 with these processes (Alen *et al*, 2002; Simic *et al*, 2003; Biswas *et al*, 2007, 2008; McDaniel *et al*, 2008). In mice, Chd1 is necessary for maintaining the pluripotency of embryonic stem cells (Gaspar-Maia *et al*, 2009).

Chd1 is comprised of a double chromodomain motif N-terminal to the Snf2 domain and a putative DNA-binding domain in the C-terminal third of the protein (Delmas *et al*, 1993; Stokes and Perry, 1995). Significant focus has been placed on understanding the role of the chromodomains within Chd1. In humans, the chromodomains of Chd1 bind to methylated lysine 4 in the histone H3 tail (H3K4), and it is believed this may target the activity of Chd1 to specific areas of chromatin—trimethylated H3K4 marks regions of active transcription (Flanagan *et al*, 2007). However, the tandem chromodomains of yeast Chd1 cannot bind to methylated H3K4 peptides (Sims *et al*, 2005; Flanagan *et al*, 2007; Okuda *et al*, 2007), and in *Drosophila* the chromodomains are not required for the chromatin localisation of CHD1 (Morettini *et al*, 2011). The crystal structure of a fragment of yeast Chd1, encompassing the chromodomains and the Snf2 domain, has shed light on how the chromodomains negatively regulate the ATPase activity of the Snf2 domain (Hauk *et al*, 2010).

In contrast, little is known about the C-terminal DNA-binding region. A region of ~250–300 amino acids in the C-terminus of mouse Chd1 was originally identified as being required for proper association with DNA and chromatin (Stokes and Perry, 1995). However, there is still insufficient annotation of this region and little further investigation has been undertaken to delineate the function of this domain. Here, we have probed this region in *S. cerevisiae* Chd1, showing that, like the mouse homologue, a domain of ~300 amino acids is necessary for DNA/nucleosome binding and is essential for the nucleosome-remodelling activity of Chd1. We have refined the boundaries of this domain and subsequently solved the structure by X-ray crystallography. The structure shows two main modules similar to a SANT (*S*WI3, *A*DA2, *N*-CoR, and *T*FIIIB) domain (Aasland *et al*, 1996) and a SLIDE (*S*ANT-*l*ike *I*SWI *d*omain) domain (Grune *et al*, 2003). The arrangement of these modules shows a striking resemblance to that of the nucleosome-recognition module of the *D. melanogaster* ISWI (DmISWI) chromatin remodeller (Grune *et al*, 2003). This finding suggests this domain may be an important conserved motif aimed at tuning Snf2 domains towards generating ordered nucleosome arrays—a characteristic property of Chd1 and ISWI remodelling enzymes.

Results

A C-terminal region of Chd1 is required for nucleosome binding and remodelling

We previously observed that Chd1 binds preferentially to nucleosomes bearing linker DNA (Stockdale *et al*, 2006), and sought to investigate whether the C-terminal DNA-binding region of the protein contributed to this. As a first step, a series of C-terminal truncations (Figure 1A; Supplementary Figure S1A) were screened for their ability to bind nucleosomes using electrophoretic mobility shift assays (EMSAs). Figure 1B shows that full-length Chd1 was able to bind nucleosomes, forming two complexes—presumably corresponding to the binding of one and two Chd1 proteins (Figure 1B, lanes 1–7). While truncations to residue 1305 did not prevent nucleosome binding, extending the truncation to residue 1010 or 860 resulted in no detectable nucleosome binding (Figure 1B , lanes 8–25), similar results were also seen when binding to DNA was tested (Supplementary Figure S1B). These observations indicate that a region N-terminal to residue 1305 and C-terminal to residue 1010 are required for Chd1 to bind nucleosomes and DNA.

To determine whether deletions that affect nucleosome binding also influence Chd1 remodelling activity, the effect of these truncations on the ability of Chd1 to reposition nucleosomes was investigated. Figure 1C shows Chd1 is unable to slide end-positioned nucleosomes when truncations beyond residue 1305 were used. The ability of these truncations to hydrolyse ATP was also tested. Truncation to residue 1305 resulted in no obvious changes in ATPase activity; however, extending beyond this resulted in a loss of nucleosome and DNA-stimulated ATPase activity (Figure 1D). In concert, these observations suggest that the existence of a DNA-binding feature C-terminal to the Snf2 domain of Chd1 is required for both engagement with nucleosomes and the catalytic action.

A minimal domain comprising ~270 residues binds 20 bp of DNA

With the aim of defining a minimum region required for DNA-binding activity, a series of truncated Chd1 fragments within the vicinity of residues 1010–1305 were prepared (Figure 2A). When incubated with DNA, several of these fragments formed complexes stable during native gel electrophoresis (Figure 2B). Of these, the shortest fragment consisted of residues 1009–1274 (referred to as Chd1-DBD; Figure 2B , lanes 17–19); further deletion into this region from either the N- or C-terminus resulted in a loss of DNA binding (Figure 2B , lanes 8–10 and 14–16).

To investigate the approximate footprint of DNA recognised by Chd1-DBD, binding to a series of progressively shorter DNA fragments was tested. Binding was observed with fragments ranging from 45 to 20 bp but not to a 15-bp fragment (Figure 2C, lanes 21–25). Finally, the ability of Chd1-DBD to bind to nucleosomal DNA was investigated. Chd1-DBD was able to bind to nucleosomes bearing linker DNA (Figure 2D, lanes 1–5), but not efficiently to nucleosome core particles (Figure 2D, lanes 6–10), which mirrors the behaviour of the full-length protein (Stockdale *et al*, 2006).

Crystal structure of the Chd1 DNA-binding domain

Inspection of common domain databases (PFAM, SMART, PROSITE and InterPro) shows little annotation of the region encompassing Chd1-DBD. InterPro identifies a short region of around 30–40 amino acids as a homeodomain-related motif; however, the remainder of Chd1-DBD has no predicted similarity with known domain sequences. To gain insight into the structural elements of this domain, we crystallised Chd1-DBD and solved the structure using a three-wavelength

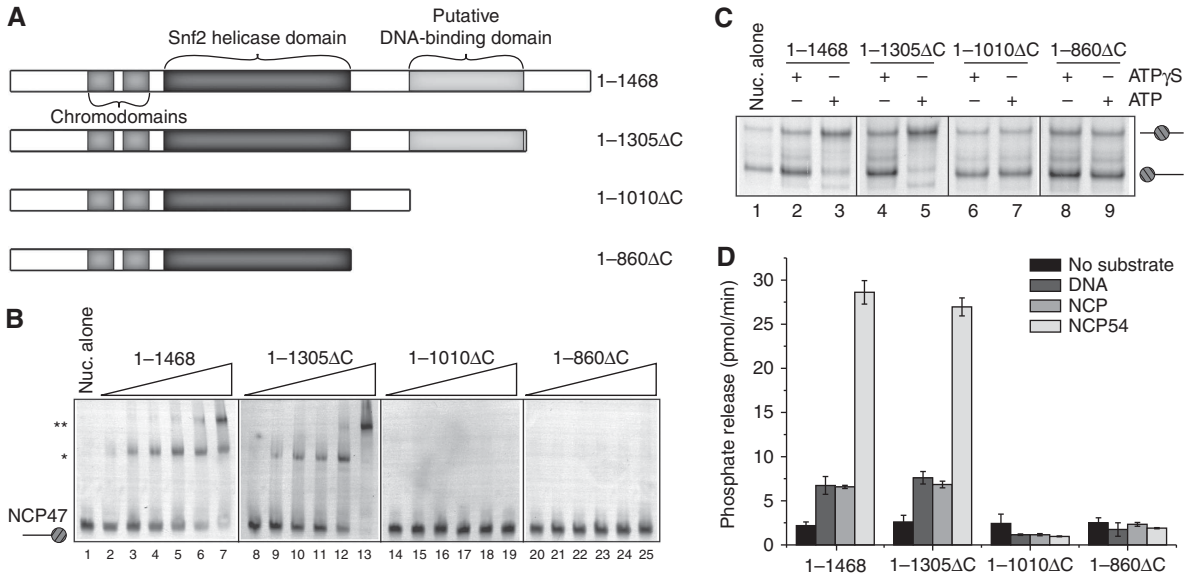

Figure 1 The C-terminus of Chd1 is required for nucleosome-binding and remodelling activity. (**A**) Schematic of the domain structure of full-length Chd1 (1–1468) and the C-terminal truncation constructs used in (**B**–**D**). (**B**) Nucleosome-binding activity of full-length Chd1 and C-terminal truncations (1–1305ΔC, 1–1010ΔC, and 1–860ΔC). Binding reactions contained 15 nM Cy3-labelled mononucleosomes assembled on 194 bp 601 sequence (Thastrom *et al*, 1999) DNA (NCP47) and 3.125, 6.25, 12.5, 25, 50, and 100 nM Chd1, 1–1305ΔC, 1–1010ΔC, or 1–860ΔC. Chd1-bound nucleosome species (star) were resolved on native polyacrylamide gels. (**C**) Nucleosome-sliding activity of recombinant Chd1 and C-terminal truncations. End-positioned mononucleosomes (0.5 pmol) assembled on 201 bp MMTV NucA DNA were incubated with 0.5 nM Chd1, 1–1305ΔC, 1–1010ΔC, or 1–860ΔC for 45 min in the presence of ATP or the non-hydrolysable ATP-analogue ATPγS. Only full-length Chd1 and 1–1305ΔC repositioned nucleosomes to the central location (slower migrating species) in an ATP-dependent manner. (**D**) DNA- and nucleosome-stimulated ATPase activity is abolished in 1–1010ΔC or 1–860ΔC. ATP hydrolysis catalysed by 5 nM Chd1, 1–1305ΔC, 1–1010ΔC, or 1–860ΔC was monitored in real time using a fluorescent-based phosphate sensor assay in the absence of substrate or in the presence of 100 nM 147 bp DNA, mononucleosomes assembled on 147 bp DNA (NCP), or mononucleosomes assembled on 201 bp DNA (NCP54). Errors bars indicate ± s.d.

MAD strategy. The final structure was refined against the native diffraction data to a resolution of 2.0 Å and had an R_{work} and R_{free} of 20.4 and 24.8%, respectively (Table I). The structure is of excellent quality with MolProbity ranking it in the 100th percentile of structures with a similar resolution (Chen *et al*, 2010), and 97.8% of residues are in the favoured regions of the Ramachandran plot and no residues in the unfavoured regions. The 2Fo-Fc density was generally well defined (Supplementary Figure S2A–C); however, for 37 residues electron density was severely broken and of poor quality (Supplementary Figure S2D). Five of these residues were at the C-terminus (residues 1270–1274), with the remainder belonging to a single loop spanning residues 1213–1244 inclusive. This loop is poorly conserved within Chd1 proteins, even among yeast species, and in higher eukaryotes it is trimmed to just two or three residues (Figure 3A), suggesting it is not likely to fulfil a general functional role, although we cannot exclude the possibility of a specific role within the context of the *S. cerevisiae* protein.

The fold (Figure 3B) is made up of 1 β-hairpin and 10 α helices, the longest being α3, which has 13 turns, and the shortest being α6 and α9, which are just a single turn each. Overall, the molecule measures $\sim 85 \times 55 \times 45$ Å at the widest point in each dimension, and resembles an 'axe' or 'hammer' in shape, with the 'head' of the axe made up of α1–2 and α5–10, the 'handle' α3 and α4 (*HL1*), and the β-linker (βL) securing the base of the 'head' to the 'handle'. The *HL1* region comprising an insertion of some 40 residues specific to budding yeast is unlikely to form such a prominent feature or perform a conserved function in other species.

Chd1-DBD contains SANT and SLIDE domains

The structure was subject to a DALI search to identify structural homologues (Holm and Rosenström, 2010). No single hit matched the entire structure. The top-scoring hit (Z-score = 7.3, RMSD = 2.5 Å) was the SLIDE domain of the DmISWI chromatin remodeller, which matches a region of Chd1-DBD encompassing helices α5–10 (Figure 4A). SLIDE domains have only been identified in ISWI proteins and the DmISWI C-terminal HAND-SANT-SLIDE domain structure (DmISWI-HSS; PDB ID 1OFC) is the only entry for a SLIDE domain in the PDB. As such, this structure represents the first evidence for the presence of SLIDE-like domains outside the ISWI family. The SLIDE domain was initially reported (Grune *et al*, 2003) as a SANT-related domain with several conserved differences, including an extended loop preceding the first helix, an extended loop between first and second helices and the insertion of an extra helix between the second and third helix. Chd1-DBD contains a longer insertion before the first helix (α7), but lacks the extension of the loop between helices one (α7) and two (α8) (Figure 4B). However, the main feature conserved between the two domains, and that sets them apart from classical SANT domains, is the insertion of a short helix (α9) between helices two (α8) and three (α10; Figure 4A and D).

The remaining DALI hits were Myb/homeodomain-like structures that matched to α1, α2, and part of α3 of Chd1-DBD. DmISWI contains a SANT domain N-terminal to the SLIDE domain (Figure 4A). SANT domains are structurally related to Myb-like domains and are common motifs found in chromatin interacting proteins (Aasland *et al*, 1996; Boyer *et al*, 2004). When the SANT region of DmISWI is

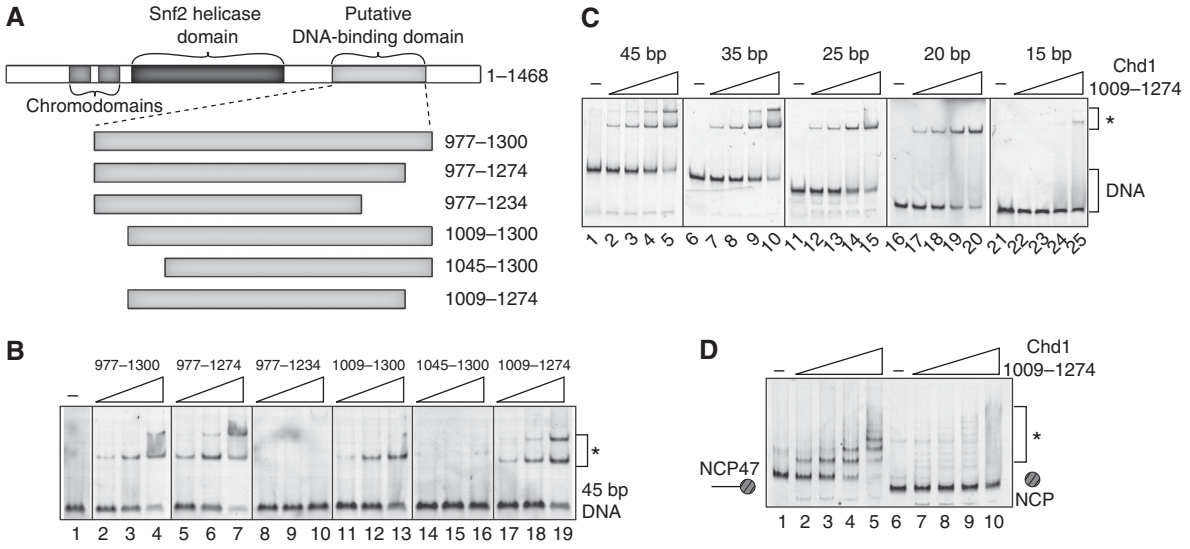

Figure 2 Defining the minimal DNA-binding domain of Chd1. (**A**) Schematic representation of the constructs used to determine the domain boundaries of the C-terminal DNA-binding domain of Chd1. Numbers indicate the boundary residues for each construct. (**B**) Binding of the constructs described in (**A**) at 50, 100, and 200 nM to 45 bp Cy3-labelled duplex DNA (50 nM). (**C**) Binding of Chd1-DBD (55, 110, 220, and 440 nM) to different lengths of Cy3-labelled duplex DNA (50 nM). (**D**) Binding of Chd1-DBD (55, 110, 220, and 440 nM) to Cy3-labelled nucleosomes (50 nM) assembled on the 601 sequence with (NCP47) or without (NCP) extranucleosomal DNA. Asterisks (*) indicate Chd1-DBD containing complexes.

Table I Data collection, phasing, and refinement statistics for Chd1-DBD

	Native	SeMet		
Data collection				
Space group	$P2_12_12_1$	$P2_12_12_1$		
Cell dimensions				
a, *b*, *c* (Å)	37.45, 91.12, 98.510	35.97, 86.98, 94.39		
α, β, γ (deg)	90, 90, 90	90, 90, 90		
		Peak	Inflection	Remote
Wavelength	0.9775	0.9792	0.9794	0.9074
Resolution (Å)	49.26–2.00 (2.11)	33.61–3.00 (3.16)	33.58–3.00 (3.16)	33.65–3.00 (3.16)
R_{sym}	0.079 (0.819)	0.077 (0.172)	0.086 (0.212)	0.077 (0.174)
$I/\sigma I$	18.1 (2.5)	27.7 (15.2)	25.5 (12.8)	28.2 (14.8)
Completeness (%)	97.5 (95.9)	99.9 (100)	99.9 (100)	99.9 (100)
Redundancy	9.7 (9.4)	13.9 (14.4)	13.8 (14.4)	13.8 (14.3)
Refinement				
Resolution (Å)	49.26–2.00			
No. of reflections				
Total	221 470			
Unique	22 943			
R_{work}/R_{free}	0.2050/0.2479			
No. of atoms (non-H atoms)				
Protein	1858			
Ligand/ion	14			
Water	109			
B-factors (non-H atoms; Å^2)				
Protein	49.4			
Ligand/ion	62.4			
Water	49.9			
RMSD				
Bond lengths (Å)	0.018			
Bond angles (deg)	1.349			

Values in parentheses are for highest resolution shell.

superimposed with the Myb-like region (α1–α3) from Chd1-DBD they overlay well with an RMSD of 1.6 Å (Figure 4C). Chd1-DBD and DmISWI SANT domains also contain similarities at the amino-acid level. For example, both have several surface-exposed basic residues on the first helix and both have a number of acidic residues protruding from the third helix.

The arrangement of the SANT and SLIDE domains with respect to one another in the two proteins is similar; however, the spacing between the two domains is different—when the

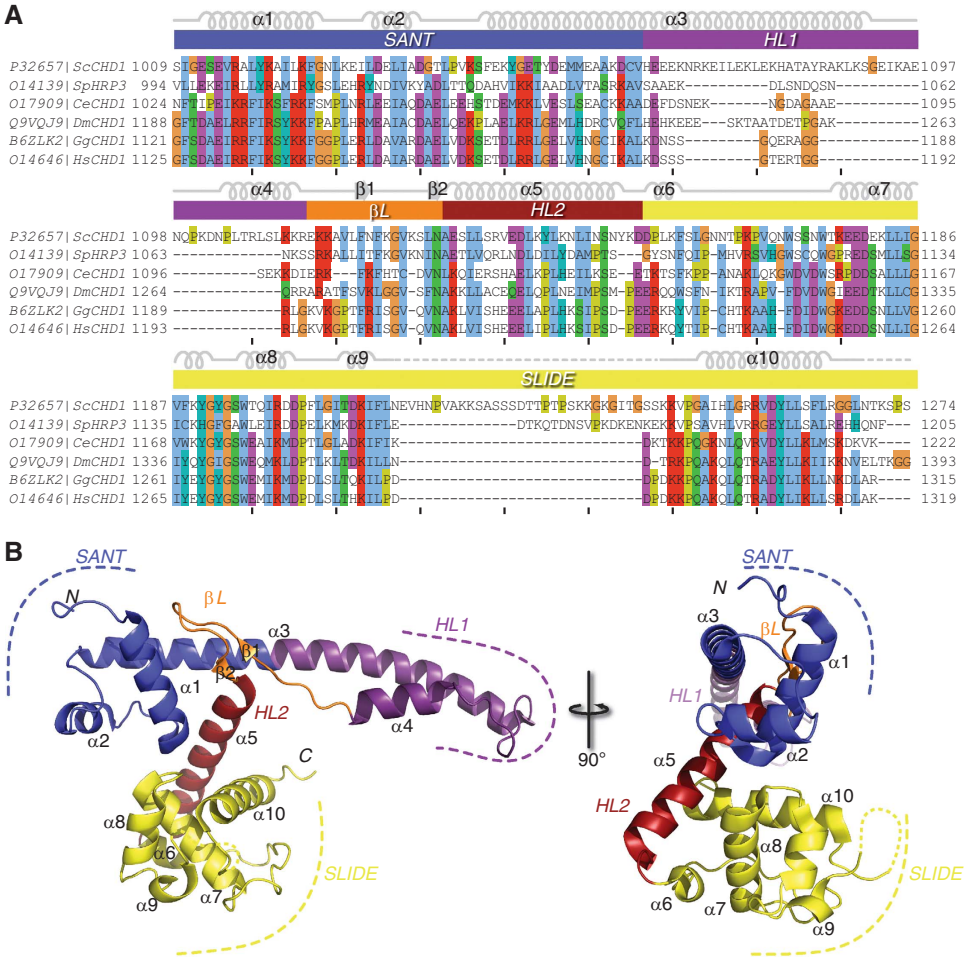

Figure 3 Structure of *S. cerevisiae* Chd-DBD. (**A**) Chd1-DBD sequence aligned to the corresponding regions of Chd1 homologues from *S. pombe*, *C. elegans*, *D. melanogaster*, *G. gallus*, and *H. sapiens*. Secondary structure definitions are from the Chd1-DBD crystal structure. The SANT (blue), SLIDE (yellow), helical linker-1 (purple; HL1) and -2 (red; HL2) and β-linker (orange; βL) regions are marked. The sequence alignment is coloured using the ClustalX colouring scheme (Thompson *et al*, 1997). (**B**) Cartoon representation of the crystal structure of Chd1-DBD is shown in two orientations. The domains are coloured and marked as denoted in (**A**).

SLIDE domains from both proteins are superimposed the SANT domain from Chd1 is shifted ~15 Å towards the SLIDE domain (Figure 4D). The SANT and SLIDE regions are well conserved in both Chd1 and ISWI proteins, leading us to hypothesise that this domain configuration may be an important component for the nucleosome-spacing activity exhibited by these enzymes.

Identification of a DNA-binding interface

The electrostatic surface potential of Chd1-DBD shows two distinct faces to the molecule. One face has a mixture of acidic, basic, and neutral surface residues (Figure 5A). The most noticeable feature on this surface is a contiguous patch of acidic residues formed primarily by the tip of the α2 (E1030, D1033, and E1034) and outer face of the α3 (E1047, E1051, D1054, E1055, E1058, and D1062) helices of the SANT domain (Figure 5A). This same region in DmISWI is also negatively charged and likely incompatible with DNA binding (Grune *et al*, 2003).

In contrast, the opposite face of the molecule is dominated by positive charge (Figure 5B). A large positive groove is formed by residues on the α1 helix in the SANT domain (R1016, K1020, and K1024), the βL (R1113, K1115, and K1126), and the α6–α7 loop (K1166), and α10 (R1255) helix in the SLIDE domain (Figure 5B). The basic nature of this face makes it a likely candidate for mediating interactions with DNA and it is a conserved feature among Chd1 proteins—basic residues either at, or in very similar positions, are found across Chd1 proteins. Furthermore, ISWI proteins also have conserved basic residues in structurally similar positions, such as K810 and R954 that lie on the first helix of the SANT domain and last helix of the SLIDE domain in the DmISWI structure, respectively.

To test whether the basic face of Chd1-DBD directly interacts with DNA, key basic residues were mutated to alanine (R1016A/K1020A, R1113A/K1115A, and R1255A). Mutant Chd1-DBD proteins were expressed, purified, and tested for binding to a 20-bp DNA fragment (Figure 5C). Each of the mutants has diminished DNA binding, albeit to differing degrees. R1255A had the least effect, but this may be due to only a single point mutation being introduced, rather than two as in the other mutants. When we combine these mutations to make triple and quadruple mutants (R1016A/K1020A/R1255A, R1016A/R1020A/R1113A/K1115A, or R1113A/K1115A/R1255A) DNA binding is abolished in our EMSAs over the concentrations tested (Figure 5C). These data

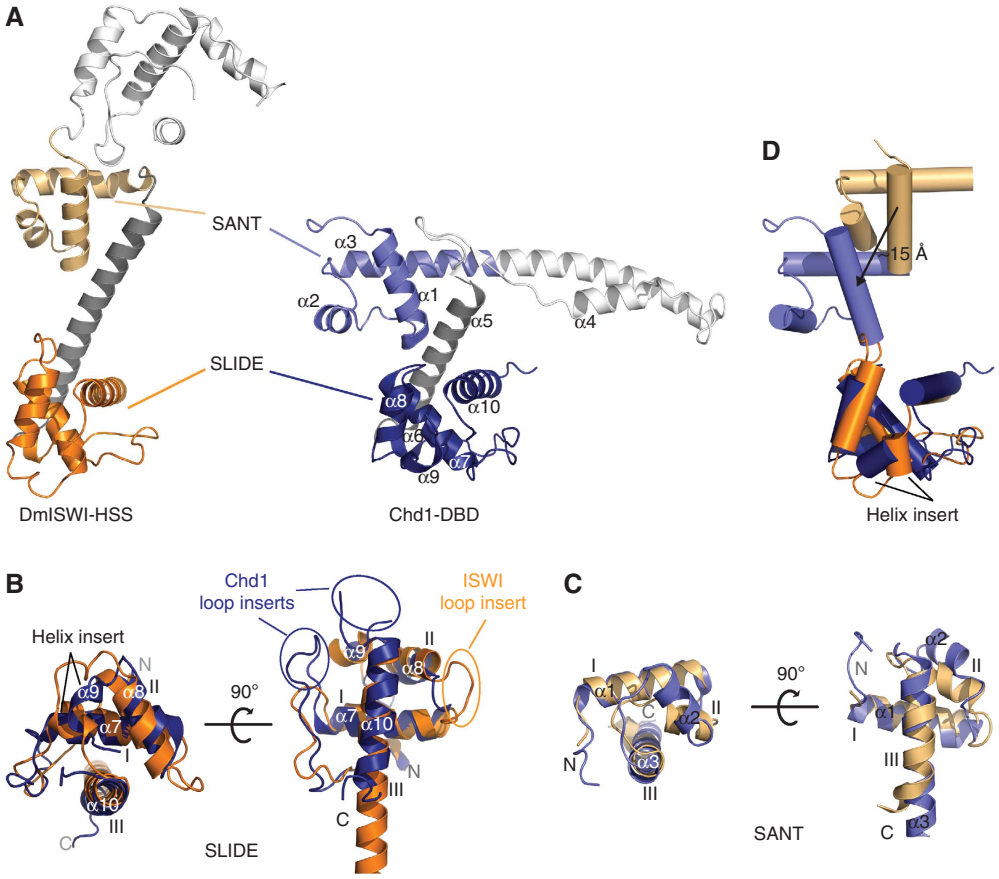

Figure 4 Comparison of Chd1-DBD with the DmISWI-HAND-SANT-SLIDE (DmISWI-HSS) structure (PDB ID 1OFC). (**A**) Crystal structures of DmISWI-HSS and Chd1-DBD are shown side by side. The SANT and SLIDE domains in each structure are in the light and dark shade (orange or blue), respectively. (**B**) Structural superposition of the SLIDE domains of Chd1-DBD and DmISWI-HSS calculated using DALI (RMSD = 2.6 Å). The differences in the inserted loop regions between the two structures are marked along with the common inserted helix (α9; also shown in (**D**)). I, II, and III denote the canonical three helices for SANT/Myb-like domains. (**C**) Similarly, superpositions of the SANT domains are shown (RMSD = 1.6 Å). (**D**) The shift (∼15 Å) in spacing between the SANT domains of ISWI (orange) and Chd1 (blue) when SLIDE domains are superposed, the other structural elements have been removed for clarity.

indicate that residues in the SANT domain (R1016/K1020), βL (R1113/K1115), and SLIDE domain (R1255) are important for interactions with DNA. Using this mutational data to define ambiguous interaction restraints for the HADDOCK web server (de Vries *et al*, 2010), we generated an ensemble of computationally derived models describing how the SANT and SLIDE domains are likely to interact with DNA—similar models are then clustered and the clusters ranked on the overall HADDOCK score (additional details of the modelling can be found in Supplementary Figure S3 and Supplementary Table SI). The lowest energy model from the top-scoring cluster (RMSD = 2.8 ± 1.8 Å for the cluster, $n = 8$) is shown in Figure 5D. The models predict the SANT α1 helix binds across the DNA-phosphate backbone and the SLIDE α10 helix docks into the minor groove with the βL region forming additional contacts with the DNA backbone. Models from the top clusters were very consistent, with all showing similar modes of binding but just differing in the register or orientation of the DNA (Supplementary Figure S3).

The DNA-binding interface of Chd1-DBD is required for the remodelling activity of Chd1 in vitro and in vivo

These DNA-binding surface mutations were introduced into full-length Chd1 (Supplementary Figure S4A) to assess the role of DNA binding by the SANT and SLIDE domains in nucleosome binding and remodelling. First, the ability of full-length Chd1 mutants to bind nucleosomes was examined. The introduction of the R1016A/K1020A, R1113A/K1115A, or R1255A mutations had little effect on nucleosome-binding affinity over the concentrations tested. However, combined mutation of R1016A/K1020A/R1255A, R1016A/R1020A/R1113A/K1115A, or R1113A/K1115A/R1255A resulted in significantly reduced nucleosome binding (Figure 6A). For wild-type and R1016A/K1020A, R1113A/K1115A, and R1255A mutant proteins, >95% of the nucleosomes are bound at the highest protein concentrations (Figure 6A), whereas ∼80% of the nucleosome signal remains unbound for R1016A/K1020A/R1255A and R1016A/K1020A/R1113A/K1115A and ∼30% for R1113A/K1115A/R1255A, at the equivalent protein concentration. This is consistent with previous observations made on DmISWI (Grune *et al*, 2003), where deletion of individual SANT or SLIDE domains did not significantly affect nucleosome binding, but combined deletion of both domains severely compromised binding. These data suggest the SANT-SLIDE motif recognises DNA/nucleosomes as a single cooperative unit. The attenuated effect of our point mutations on the full-length protein compared with isolated Chd1-DBD (Figure 5C) indicates other regions within Chd1 contribute

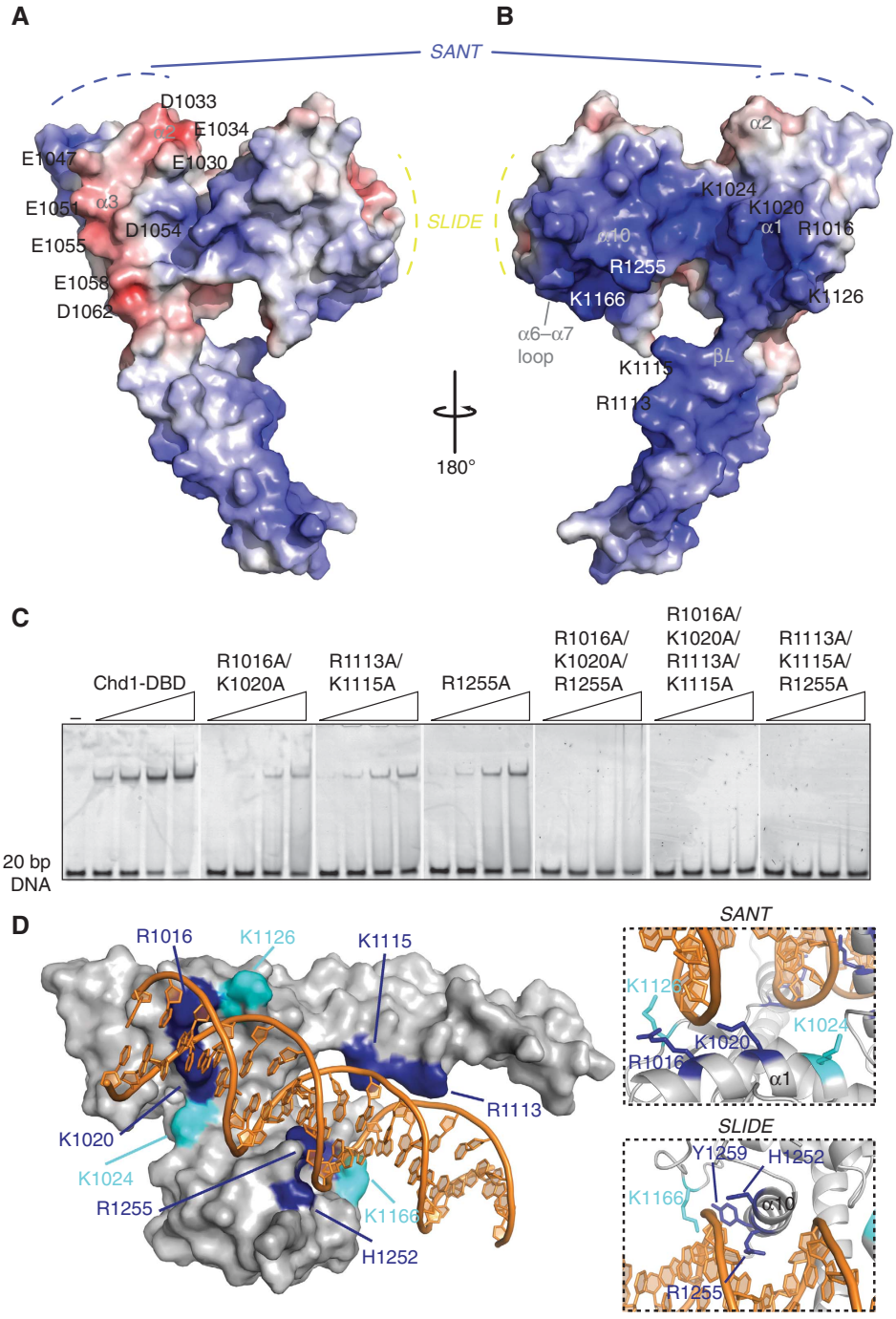

Figure 5 The basic surface of Chd1-DBD is necessary for DNA binding. (**A**) One face of the molecule exhibits a mixture of acidic, basic, and neutral surfaces. The labelled amino acids within the SANT domain comprise an acidic patch on this face. (**B**) The opposite face of the molecule is almost exclusively positively charged. Key residues involved in DNA binding and underlying structural elements are marked. Electrostatic surface features of Chd1-DBD were calculated using APBS (Baker *et al*, 2001) (±7 kT/e); blue and red represent positive and negative electrostatic potential, respectively. (**C**) Wild-type and mutant Chd1-DBD (55, 110, 220, and 440 nM) were tested for binding to Cy3-labelled 20 bp duplex DNA (50 nM). (**D**) Top-scoring HADDOCK-generated model of Chd1-DBD bound to 20 bp DNA. Upper inset, the interaction between the SANT α1 helix and DNA. Lower inset, the interaction between the SLIDE α10 helix and the minor groove of DNA. Residues in dark blue denote 'active' residues in the HADDOCK modelling, residues in cyan are other basic residues found on the binding surface.

to nucleosome binding. Indeed, recently the chromodomains of Chd1 have been shown to occlude a DNA-binding motif within the Snf2 domain (Hauk *et al*, 2010), which may be responsible for stabilising interactions with nucleosomes.

Next, the ability of the mutants to reposition nucleosomes was assessed. The R1016A/K1020A, R1113A/K1115A, and R1255A mutants showed a visible decrease in the rate of nucleosome sliding, and these data are summarised in Figure 6B, which shows the overall rate of accumulation of a final centrally positioned nucleosome starting from an end-positioned nucleosome assembled on the MMTV NucA positioning sequence (Stockdale *et al*, 2006). The triple and quadruple mutants (R1016A/K1020A/R1255A, R1016A/R1020A/R1113A/K1115A, or R1113A/K1115A/R1255A) exhibited

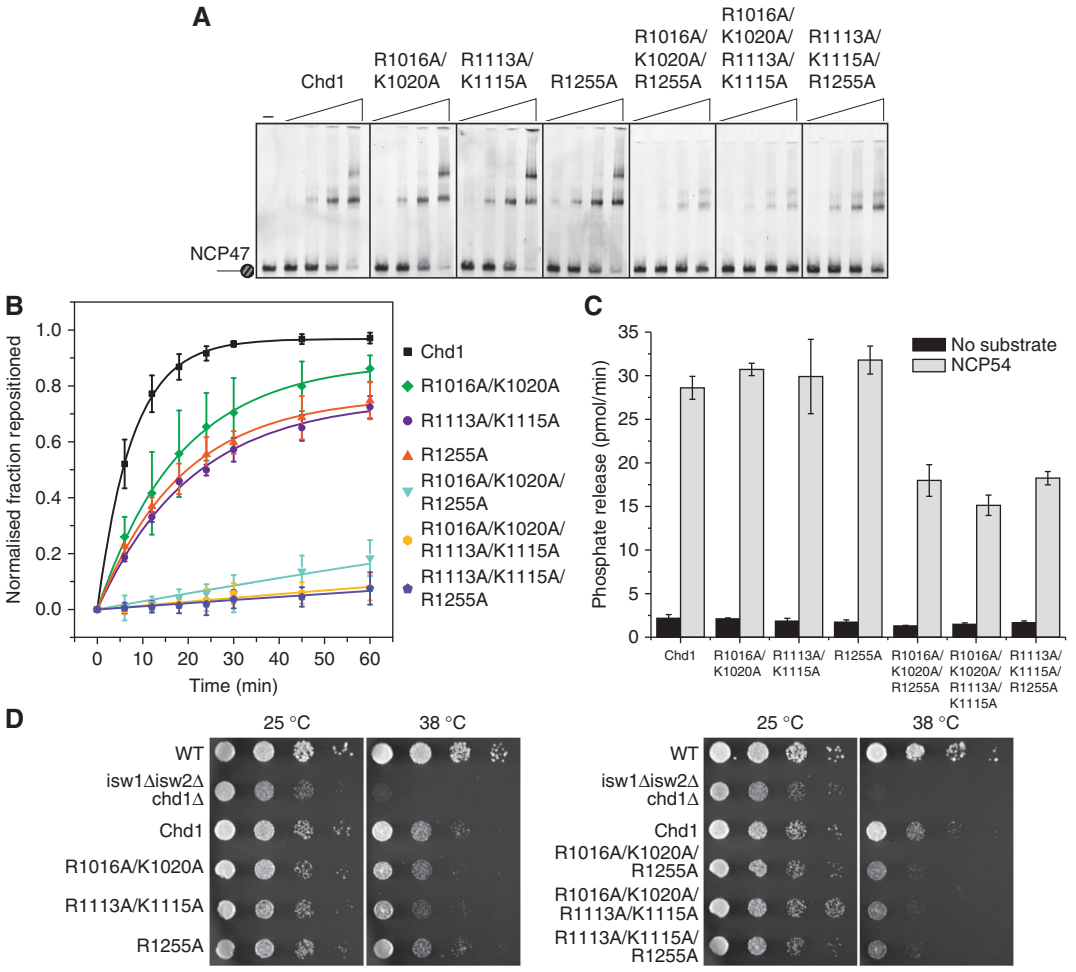

Figure 6 The basic surface of Chd1-DBD is important for nucleosome remodelling by Chd1. (**A**) Full-length wild-type and mutant Chd1 (7.5, 15, 30, and 60 nM) were tested for binding to Cy3-labelled NCP47 nucleosomes (25 nM) (**B**) Quantification of nucleosome sliding assays using full-length wild-type and mutant Chd1. End-positioned nucleosomes (0.5 pmol) were incubated with Chd1 (0.5 nM) for the times indicated, the products resolved on native polyacrylamide gels and the bands quantitated. Symbols represent the average of at least three experiments and error bars indicate ± s.d. Solid lines are the fits to the data of a single exponential model. (**C**) ATPase activity of Chd1 mutants (5 nM) in the absence (no substrate) or presence (NCP54) of 100 nM nucleosomes. Data for wild-type Chd1 are taken from Figure 1D for comparison. (**D**) Yeast complementation assays assessing the ability of full-length wild-type and mutant Chd1 proteins to rescue the temperature sensitive growth defect of an *isw1*Δ, *isw2*Δ, *chd1*Δ triple knockout strain. The triple knockout strain was transformed with single-copy plasmids encoding wild-type or mutant Chd1. Serial dilutions of the transformants and wild-type strain (BMA64) were spotted at equal densities onto selection plates and incubated at 25 or 38°C for 2 days.

significantly decreased sliding activity under the same conditions. The data in Figure 6B were fit with a single exponential function to determine the relative changes in the rate of movement (Figure 6B, solid lines). This analysis revealed the overall movement of nucleosomes into the more centrally located position is decreased, 2.4-, 2.7-, and 2.4-fold for R1016A/K1020A, R1113A/K1115A, and R1255A, respectively, relative to wild type, whereas for R1016A/K1020A/R1255A, R1016A/R1020A/R1113A/K1115A, or R1113A/K1115A/R1255A these decreases were ~43-, 92-, and 113-fold.

Next, we assessed the effects of these mutations on the nucleosome-stimulated ATPase activity of Chd1, using the same conditions as those shown in Figure 1D. R1016A/K1020A, R1113A/K1115A, and R1255A each had ATPase activity comparable to wild-type levels when stimulated with nucleosomes bearing 54 bp linker DNA (Figure 6C). The combined mutations, R1016A/K1020A/R1255A, R1016A/R1020A/R1113A/K1115A, or R1113A/K1115A/R1255A, have stimulated ATPase levels 50–60% of wild type (Figure 6C). This is surprising given the much reduced levels of nucleosome sliding displayed by these mutants. The ATPase experiments were performed with higher amounts of remodeller and nucleosomes compared with the sliding experiments and the nucleosomes used were assembled on the stronger 601 positioning sequence. Thus, to confirm the observed levels of ATPase activity were concomitant with defects in sliding we performed further sliding reactions using nucleosomes assembled on the 601 sequence and with the same concentrations of remodeller and nucleosomes as in the ATPase experiments (Supplementary Figure S4B). Again, defects in all of the mutants were apparent with R1016A/K1020A/R1255A, R1016A/R1020A/R1113A/K1115A, and R1113A/K1115A/R1255A generating very little nucleosome sliding over the course of the experiment. The observation of increased ATP hydrolysis relative to nucleosome sliding suggests that the DNA-binding domain is important for efficient conversion of the energy derived from ATP hydrolysis into that required for nucleosome remodelling.

In order to determine whether the residues that contribute to DNA binding are also important for the *in vivo* function of Chd1, we assessed the ability of mutant Chd1 proteins to rescue a temperature sensitive growth defect phenotypic of an *isw1*Δ, *isw2*Δ, *chd1*Δ triple knockout yeast strain (Tsukiyama *et al*, 1999). In our assay, the R1016A/K1020A, R1113A/K1115A, and R1255A mutants were comparable to wild type in their ability to restore growth at the non-permissive temperature of 38°C (Figure 6D). However, the R1016A/K1020A/R1255A, R1016A/R1020A/R1113A/K1115A, or R1113A/K1115A/R1255A mutants did not support growth to the same extent (Figure 6D). These results are in agreement with the *in vitro* data, where the greatest effects on remodelling activity were seen in the R1016A/K1020A/R1255A, R1016A/R1020A/R1113A/K1115A, and R1113A/K1115A/R1255A mutants. This indicates that the DNA-binding domain contributes to Chd1 function *in vivo*.

Chd1 and ISWI SANT domains are closely related

The presence of a SANT/Myb-like domain at this location in Chd1 was not expected from the sequence. This prompted us to examine the sequence relationships between Chd1 SANT domains and other SANT domain-containing proteins. Initially, SANT domain sequences were collected from a range of well-documented SANT domain-containing proteins, including Nuclear Corepressor (N-Cor), REST Corepressor (R-Cor), Metastasis-associated (MTA) proteins, Ada2, Swi3, Rsc8, TFIIIb, ISWI, and DnaJC. These sequences were aligned with the SANT domain sequences of a selection of Chd1 proteins from yeast to humans. Myb-repeat sequences were also included in the alignment for comparison. In total, 128 sequences were used in the alignment (Supplementary Figure S5), which was used to construct the neighbour-joining tree in Figure 7A. The tree shows that SANT sequences of Chd1 proteins are the most distantly related group of sequences relative to other SANT/Myb sequences, and are more diverse than other SANT proteins. This provides some explanation for why Chd1 sequences have not been previously identified as containing a SANT domain at this position. The tree also shows that SANT sequences from ISWI proteins are most closely related to those from Chd1 proteins. Two features appear to be characteristics of ISWI and Chd1 SANT domains; first, a higher proportion of basic residues in the first helix; second, a higher proportion of acidic residues in the second helix and adjacent loops (Supplementary Figure S5).

The relatedness of Chd1 and ISWI SANT sequences was independently confirmed using iterative Hidden–Markov model (HMM) sequence profile searches of the UniRef90 database. For this, 25 SANT domain sequences from Chd1 proteins were aligned and a HMM constructed using HMMER (Eddy, 1998). Three rounds of searching and HMM construction using the results were performed. Outside of Chd1 sequences, ISWI SANT sequences were the next family of sequences returned in this process—the results are summarised in Supplementary Figure S6. The identification of ISWI sequences in this way indicates they are the most closely related sequences to Chd1 SANT domains outside of the Chd1 family. Chd1 SANT sequences cluster into three distinct subfamilies that correspond to fungi, metazoan, and plant sequences, with sequences from protists (apicomplexa and slime moulds) forming minor subfamilies closely related to fungi (Supplementary Figure S6). The plant subfamily is most closely related to the ISWI sequences. Chd1 SANT sequences are diverse when compared with those of ISWI, with each of the Chd1 subfamilies occupying similar sequence space to that of all the ISWI sequences combined.

SLIDE domains are found in other Chd proteins

Iterative searching with HMM profiles was also used to identify other proteins that may contain putative SLIDE domains. The sequences of SLIDE domains from the same set of 25 Chd1 sequences used for the SANT HMM searches were used to seed the process. The loop region between α9 and α10 (Figures 3A and 7B) in Chd1 SLIDE sequences is highly variable, in both the length and amino-acid composition. This variation meant more remotely related sequences (even within the Chd1 family) often scored poorly in the search process and automated alignments were often incorrect in this region. Thus, careful inspection of the results was necessary after each round and alignments were adjusted manually to correct obvious mistakes. Similar to the SANT domain search results the majority of hits belonged to Chd1 sequences. Outside of the Chd1 family the most significant hits belonged to the CHD7 subfamily of Snf2 proteins, which consists of CHD6–9 (Flaus *et al*, 2006). The highest scoring of these was identified in the second round of searches as a Kismet-related sequence from *Apis mellifera* (UPI0000DB76A1; *E*-value = 0.053, bit score = 0.7)—Kismet is the *D. melanogaster* homologue of CHD9 (Daubresse *et al*, 1999). Figure 7B shows a representative alignment of CHD6–9 sequences against the SLIDE domain sequences from several Chd1 and ISWI proteins. The most well-conserved region aligns to the sequence spanning α7–α9 from Chd1-DBD (Figure 7B). Like some Chd1 sequences, the CHD7 sequences contain a long insert after the putative α9 helix, which is less well conserved than the surrounding sequence. Secondary structure predictions indicate that the CHD7 sequences are likely to have secondary structural features similar to that of Chd1 and ISWI (Figure 7B). The similarity of the CHD7 family sequences with the SLIDE domain was further confirmed by PSI-BLAST searches, which use a different algorithm to the HMM method. Using a Chd1 SLIDE sequence as the seed sequence resulted in CHD7 family sequences being returned, or *vice versa* in the reciprocal search (data not shown). The SLIDE region of ISWI proteins also turned up in our HMM searches; however, these were typically lower scoring, and so were excluded from HMM construction for subsequent rounds of searches. The low scores of ISWI sequences are likely due to the insertions surrounding the second helix and the lack of insertion preceding the final helix (Figure 7B), each of these features are well conserved among ISWI proteins but absent from Chd1 and CHD7 sequences. Sequences from CHD3/4 proteins also appeared in our search results but were poor scoring and ranked alongside a number of general Myb/SANT-containing sequences (Supplementary Dataset 1). Thus, CHD3/4 sequences may also contain a SLIDE-related motif, but this is significantly divergent from that of Chd1 and further functional and structural analysis will be required to determine the significance of this relationship.

Discussion

We have solved the structure of the C-terminal DNA-binding domain from Chd1 and shown that full DNA-binding activity of this domain is required for the function of Chd1. Our

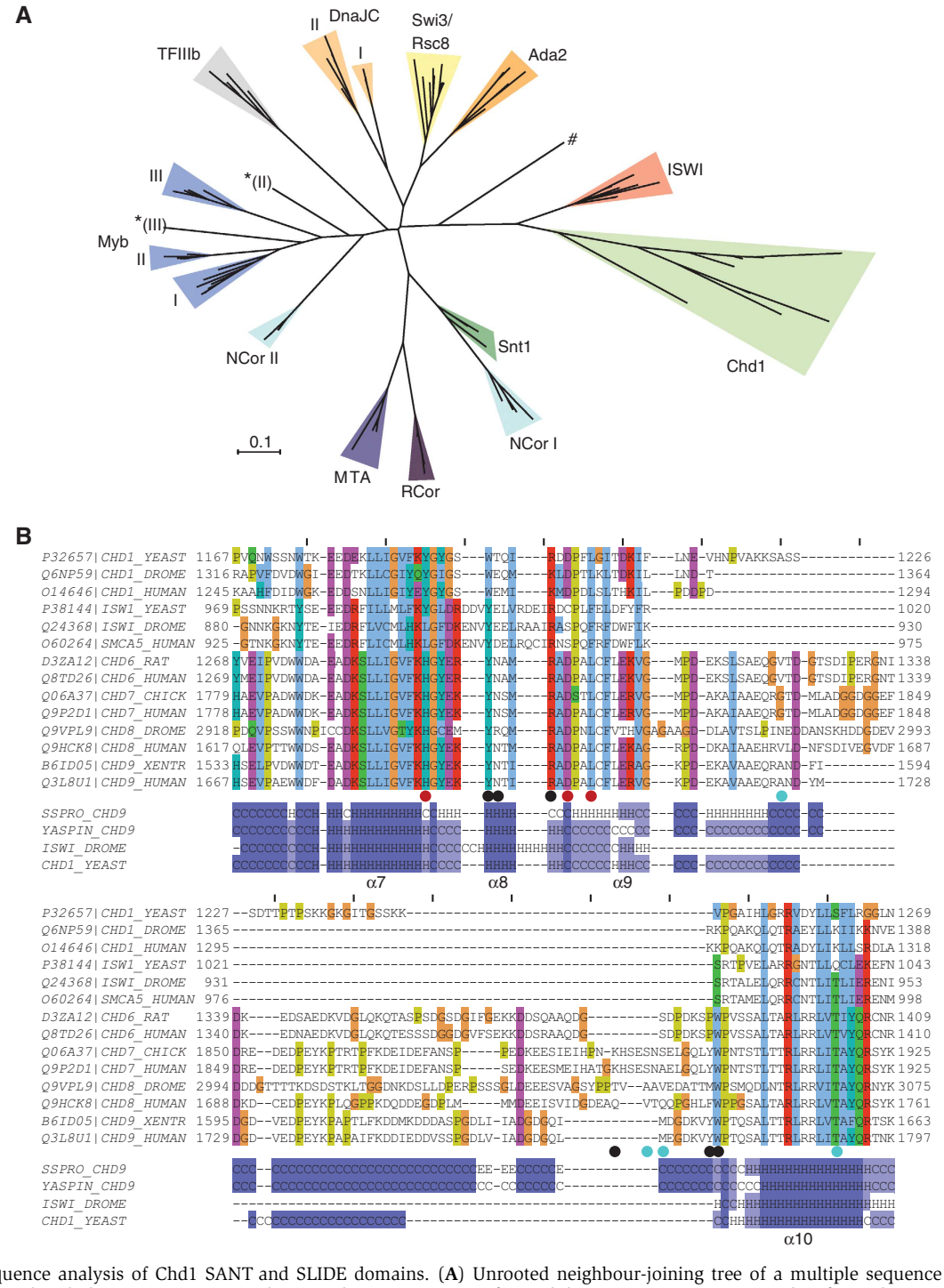

Figure 7 Sequence analysis of Chd1 SANT and SLIDE domains. (**A**) Unrooted neighbour-joining tree of a multiple sequence alignment of known SANT and Myb domain sequences with SANT domain sequences from Chd1 proteins. Roman numerals refer to the repeat number for proteins containing multiple SANT/Myb domains. *Indicates Myb (II and III) repeats from *S. cerevisiae Bas*I (P22035) and # marks the SANT domain of *D. melanogaster* dADA2a (Q7KSD8), both outliers for their respective subfamilies. (**B**) Alignment of SLIDE sequences from Chd1 and ISWI proteins against SLIDE-related sequences from CHD6–9 proteins. Coloured dots indicate amino acids affected by missense (red), nonsense (black), and frameshift (cyan) mutations in the human CHD7 gene of CHARGE syndrome patients. Secondary structure predictions from YASPIN (Lin *et al*, 2005) and SSPRO (Cheng *et al*, 2005) structure prediction servers for human CHD9 aligned with the secondary structures of Chd1-DBD and DmISWI SLIDE domains, H = helix, C = coil, E = sheet. Similar predictions were made for the other CHD6–9 sequences. Corresponding helices from the Chd1-DBD structure are indicated.

biochemical analysis of the DNA-binding activity of this domain within the context of the activity of the full-length protein indicates that this domain, although required for binding to nucleosomes, also has a significant role during the nucleosome-remodelling cycle post engagement with nucleosomes. This is evident even with the modest mutation of one or two residues in the SANT (R1016A/K1020A), SLIDE (R1255A), or βL region (R1113A/K1115A), where the mutant proteins demonstrated wild-type levels of nucleosome binding and ATPase activity but had defects in their ability to slide

nucleosomes. These effects were then exaggerated when the mutations were combined, indicating the whole SANT-SLIDE motif acts as a cooperative unit to bind DNA. There are two obvious explanations that could explain why the DNA-binding domain influences the efficiency of nucleosome sliding. First, the domain may act as a processivity factor, such that full DNA-binding affinity is required to maintain an association with the nucleosome so that sufficient cycles of remodelling can occur to complete the sliding reaction. The second possibility, and perhaps the more compatible with our data, is that the DNA-binding domain acts as a 'grip' that allows the enzyme to generate enough tension/force to move the nucleosome. When the strength of the 'grip' is reduced the enzyme, although still associated with the nucleosome, 'slips' and continues to hydrolyse ATP such that the net movement of the nucleosome is greatly reduced with respect to the amount of ATP used.

Our structure has revealed previously unrecognised structural homology with the ISWI subfamily of Snf2 proteins (Grune *et al*, 2003). There is significant evidence, both *in vitro* and *in vivo*, linking the activities of Chd1 and ISWI proteins. Biochemically, both Chd1 and ISWI demonstrate a dependence on linker DNA for nucleosome binding and remodelling (Stockdale *et al*, 2006; Yang *et al*, 2006; Gangaraju and Bartholomew, 2007), as well as exhibiting a dependence on the N-terminal tail of histone H4 (Clapier *et al*, 2001; Hamiche *et al*, 2001; Fazzio *et al*, 2005; Ferreira *et al*, 2007). In yeast, deletion of combinations of Isw1, Isw2, and Chd1, but not the individual deletions, results in synthetic phenotypes, including sensitivity to temperature and formamide (Tsukiyama *et al*, 1999). These genetic interactions are consistent with shared or redundant functions for these proteins. Furthermore, deletion of SANT and SLIDE domains has been shown to prevent DmISWI from repositioning nucleosomes *in vitro* (Grune *et al*, 2003) and *S. cerevisiae* Isw1 from repositioning nucleosomes at the MET16 promoter *in vivo* (Pinskaya *et al*, 2009). Coupled with our observations that mutations to the DNA-binding surfaces of Chd1 SANT and SLIDE domains negatively affect the activity of Chd1, these data raise the possibility that the redundant functions of these proteins relate to the role of the SANT and SLIDE domains in directing nucleosome organising activities.

Chd1 and ISWI enzymes are well documented as being able to generate regularly spaced arrays of nucleosomes (Corona *et al*, 1999; Tsukiyama *et al*, 1999; Grune *et al*, 2003; Lusser *et al*, 2005). The human ACF complex, which contains the ISWI homologue SNF2h, achieves this by acting as a dimer that continuously samples linker DNA on either side of the nucleosome, moving the nucleosome until both linker lengths are equal (Blosser *et al*, 2009; Racki *et al*, 2009). This coupled with the observation the SANT-SLIDE region of yeast Isw2 contacts extranucleosomal DNA (Dang and Bartholomew, 2007) strongly implicate the SANT-SLIDE motif as having a direct role in the nucleosome-spacing mechanism through interactions with linker DNA. Our identification of a SANT-SLIDE motif in Chd1 suggests it may use an analogous mechanism, and it will be of significant interest to determine whether Chd1 also acts as a dimer; as a point of interest, our EMSAs indicate more than one molecule of Chd1 can interact with the nucleosome. The spacing between the SANT and SLIDE domains could be envisaged to affect the spacing between nucleosomes by defining the size of the footprint of the SANT-SLIDE motif on the linker DNA. The reduced spacing between these domains in Chd1 compared with DmISWI (Figure 4D) would be anticipated to generate nucleosome arrays with decreased linker DNA lengths. This has been observed with *D. melanogaster* ACF and CHD1, which generate nucleosome repeat lengths of ~175 and ~162 bp, respectively (Lusser *et al*, 2005). However, this interpretation is complicated by the fact that ISWI proteins are often part of multi-protein complexes where the additional subunits may influence remodelling activity. For example, yeast Isw1 and Isw2, which have highly similar SANT-SLIDE regions, are part of distinct complexes that generate nucleosomes spaced ~175 and ~200 bp apart (Tsukiyama *et al*, 1999), respectively. Furthermore, the ACF1 subunit of the ACF and CHRAC complexes alters the linker DNA length dependence of the ISWI subunit and its remodelling activity (Eberharter *et al*, 2001; Yang *et al*, 2006). Ionic conditions can also influence the outcome of nucleosome-spacing reactions (Blank and Becker, 1995). Thus, while SANT-SLIDE motifs appear to be the characteristic of nucleosome-spacing Snf2 proteins it is still unclear whether this motif is a determinant for the distance by which adjacent nucleosomes are separated.

SANT and SLIDE domains are related to Myb DNA-binding domains; however, they may interact with DNA in a non-canonical manner. In particular, basic residues important for Myb–DNA interactions are not conserved in Chd1 SANT (or DmISWI SANT; Grune *et al*, 2003), and have been substituted for a number of acidic residues. Instead, we have shown that basic residues on the first helix (α1) of Chd1 SANT are important for DNA interactions. Our modelling of the Chd1-DBD:DNA complex suggests that a number of interactions with the DNA-phosphate backbone are made, along with several non-specific base contacts within the minor groove (see Figure 5D and Supplementary Figure S3). This mode of binding appears largely sequence independent, and is presumably a requirement for Chd1 to be able to interact with and remodel nucleosomes located throughout the genome. Interactions with the minor groove were a feature of all of our top-scoring models, and this is consistent with data on the C-terminal domain of mouse Chd1 that show it has a preference for the minor groove of AT-rich sequences *in vitro* (Stokes and Perry, 1995), and our own observations that indicate double-stranded poly-dAdT and poly-dIdC, which have similar structural properties in the minor groove but not the major groove, compete for Chd1-DBD DNA binding more effectively than poly-dGdC (Supplementary Figure S7). Although these data are suggestive of some sequence specificity, *in vivo*, mouse Chd1 was not seen to colocalise with AT-rich regions of the genome (Stokes and Perry, 1995), and yeast Chd1-DBD clearly binds to sequences that are not necessarily AT rich (the GC content of DNA used in our EMSAs was 50–60%). Thus, AT-rich sequences may display structural features favourable for Chd1-DBD binding; however, they are unlikely to be a genuine sequence-specific target *in vivo*. The basic residues we identified as being important for Chd1 DNA binding are also generally conserved in ISWI proteins, and so our Chd1-DBD:DNA model may be a good approximation of ISWI:DNA interactions. In addition, cross-linking experiments performed on yeast Isw2 (Dang and Bartholomew, 2007)

predict DNA binds to the equivalent surface of the SANT-SLIDE motif in this protein.

The identification of new examples of SANT and SLIDE domains provides an opportunity to refine the sequence definitions of these domains and search for their presence in other proteins. Our identification of SLIDE-related domains in the CHD7 subfamily of Snf2 proteins suggests the SLIDE domain may be an important motif for other chromatin-remodelling activities. The SLIDE region in CHD7 proteins is flanked by further blocks of conserved sequence, which interestingly, includes a predicted SANT domain directly C-terminal to the SLIDE sequence. This suggests, like Chd1 and ISWI, it forms part of a larger domain structure that may interact with DNA. Mutations to human CHD7 are found in ~60% of cases of CHARGE syndrome, which results in congenital abnormalities including choanal atresia and malformations of the heart, inner ear, and retina (Zentner *et al*, 2010). Interestingly, several of the CHD7 mutations found in CHARGE syndrome patients map to the SLIDE domain (Figure 7B). Although the number of these mutations that have been sequenced to date is small it may be significant that missense point mutations are enriched at highly conserved residues, such as H1801P or L1815P, in comparison with nonsense and frameshift mutations (Zentner *et al*, 2010) (Figure 7B), suggesting the structural integrity of the SLIDE domain could be important for the activity of CHD7 in development.

With the addition of the structure reported here Chd1 now represents the most structurally well-characterised Snf2 protein, with structures available for all three of the major domains found in the enzyme (Flanagan *et al*, 2007; Hauk *et al*, 2010). Although work is still required to establish exactly how the SANT and SLIDE domains direct engagement of Chd1 with nucleosomal DNA, the increased availability of structural information provides essential details for future studies aimed at dissecting the molecular mechanism of nucleosome remodelling by spacing enzymes.

Materials and methods

Cloning, protein expression, and purification

Full-length wild-type Chd1 was cloned into the pGEX-6P expression vector (GE Healthcare) from yeast genomic DNA and engineered to have a C-terminal 6 × His tag (SSHHHHHH). C-terminal truncations (1–1305ΔC, 1–1010ΔC, or 1–860ΔC) were made from the full-length clone using an inverse PCR strategy. DNA-binding domain constructs were subcloned into pGEX-6P from the full-length clone using standard molecular biological methods. Point mutations were introduced using QuikChange® mutagenesis (Agilent). All proteins were expressed in Rosetta2 (DE3) pLysS *Escherichia coli* cells at 20°C overnight in LB medium following induction with 0.4 mM IPTG, except for selenomethionine (SeMet) substituted protein, which was expressed in B834 cells harbouring the pRARE2 plasmid (Novagen) growing in SeMet expression media (Molecular Dimensions).

Full-length Chd1, 1–1305ΔC, 1–1010ΔC, or 1–860ΔC were purified as described for the Fun30 chromatin-remodelling enzyme (Awad *et al*, 2010) via tandem metal- and glutathione-affinity purification.

DNA-binding domain GST fusion proteins were purified by glutathione affinity on SuperGlu agarose (Generon) in 20 mM Tris pH 7.5, 350 mM NaCl, 0.05% β-mercaptoethanol, containing protease inhibitors (0.2 mM 4-(2-aminoethyl)benzenesulphonyl fluoride, 2 μM E64, 2.6 mM aprotinin, and 1 μM pepstatin). Bound protein was released from GST by cleavage with Prescission protease overnight at 4°C in 20 mM Tris pH 7.5, 350 mM NaCl, and 1 mM DTT. The protein was either concentrated at this stage and exchanged into 20 mM Tris pH 7.5, 150 mM NaCl, 1 mM DTT, and 10% glycerol, snap-frozen and stored at −80°C, or purified by ion exchange at pH 7.2 on either HiTrap SP High-Performance columns (GE Healthcare) or Pierce Strong Ion-exchange Spin Columns (Thermo Scientific), according to manufacturer's instructions. For crystallisation experiments Chd1-DBD was exchanged into 20 mM Na Cacodylate pH 7.1, 100 mM NaCl, and 1 mM DTT.

Crystallisation, data collection, and structure solution of Chd1-DBD

Native and SeMet crystals were grown at 20°C using hanging drop vapour diffusion by mixing a 20 mg/ml protein solution with reservoir solution (0.1 M Tris pH 8.5, 28–30% PEG 4000, 0.35–0.45 M $MgCl_2$) at ratios of 1:2–0.5:2.5. Crystals grew within 2–3 days as rectangular needles measuring up to 1000 × 100 × 100 μM. Crystals were harvested in reservoir solution containing 20% (v/v) glycerol and flash frozen in liquid nitrogen. Native diffraction data were collected at Diamond light source beamline I04 and MAD data at ESRF beamline BM-16. Data were processed in iMOSFLM and scaled in SCALA from the CCP4 suite (Bailey, 1994). PHENIX (Adams *et al*, 2010) was used for structure determination and refinement. Phenix.autosol was used to determine heavy atom positions and experimental phases using the three-wavelength MAD data (figure-of-merit after SOLVE and RESOLVE stages were 0.59 and 0.82, respectively) followed by initial model building using the SeMet-phased maps, and then a single round of automated model extension using phenix.autobuild. Further refinement was carried out in phenix.refine against the native data. TLS parameters generated using the TLSMD web server (Painter and Merritt, 2006), and riding hydrogens were included in the refinement strategy. Manual model building was performed in COOT (Emsley *et al*, 2010). Structural figures were prepared using PyMol (http://www.pymol.org).

Nucleosome assembly and remodelling assays

Nucleosomes were assembled using standard salt-gradient dialysis methods using recombinant *Xenopus laevis* histones and PCR-generated DNA fragments containing either the 601 (Thastrom *et al*, 1999) or MMTV NucA positioning sequences, as described previously (Ferreira *et al*, 2007). Nucleosome sliding assays were performed in TSM buffer (50 mM NaCl, 50 mM Tris (pH 7.5), and 3 mM $MgCl_2$), and 1 mM ATP, incubated at 30°C with the amount of Chd1 specified in the figure legends. Reactions were stopped at the times indicated by addition of 500 ng of *Hin*dIII-digested λ DNA and sucrose to 5% (w/v) before electrophoresis on 0.2 × Tris-Borate EDTA (TBE), 5% native polyacrylamide gels. For time course experiments, bands were quantitated in AIDA software and data fit in MicroCal Origin 7 software using the following equation:

$$y = a(1 - e^{-kx})$$

where a and k are constants representing the asymptote and the relative reaction rate, respectively.

ATP-hydrolysis assays

ATP hydrolysis was monitored in real time using the phosphate-sensitive fluorescent sensor MDCC-PBP as described previously (Ferreira *et al*, 2007) in TSM buffer using 5 nM Chd1 and 100 nM DNA or nucleosomes (as specified in the figures). Averages were taken from at least three experiments.

Electrophoretic mobility shift assays

EMSAs were prepared in TSM buffer supplemented with 0.1 mg/ml BSA and either 3% (w/v) Ficoll-400 for reactions containing full-length and C-terminally truncated Chd1, or 5% (w/v) sucrose and 1 mM DTT for reactions containing DNA-binding domain constructs. Amounts of protein and DNA/nucleosomes are indicated in the figure legends. Reactions were assembled on ice and incubated for 30 min before electrophoresis on 0.5 × TBE acrylamide gels (5% for nucleosomes and 10–12% for short DNA fragments). DNA sequences used in the EMSAs are derived from the linker DNA flanking the 601 positioning sequence used in our nucleosome-binding studies; 45 bp—TGTAGGGGATTCTCTAGAGTCGACCTGCAG GCATGCAAGCTTGAG, 35 bp—GGGATTCTCTAGAGTCGACCTGCAG GCATGCAAGC, 25 bp—TCTCTAGAGTCGACCTGCAGGCATG, 20 bp—CTAGAGTCGACCTGCAGGCA, and 15 bp—AGAGTCGACCTGCAG.

Yeast complementation assays

The *S. cerevisiae* strains used in this study were BMA64: *MATα ura3-1 ade2-1 his3-11,5 trp1Δ leu2-3,112 can1-100* (Baudin *et al*, 1993) and MP28: *MATα ura3-1 ade2-1 his3-11,5 trp1Δ leu2-3,112 can1-100*

isw1::*URA3*, *isw2*::*TRP1*, *chd1*::*KanMX* (a kind gift from Dr J Mellor). Full-length and mutant yeast Chd1 were subcloned into pRS313 plasmid (Sikorski and Hieter, 1989), along with the promoter region of the *CHD1* gene and *ADH1* terminator. Plasmids were transformed into the MP28 strain using standard protocols and the transformants selected on synthetic dropout (SD) plates lacking histidine (−His). For temperature sensitivity assays, strains were resuspended in sterile water to $OD_{600\,nm}$ 1.0 and serially diluted (10-fold), spotted onto SD-His plates and incubated at 25 or 38°C for 2–3 days.

Sequence searches and modelling

Sequence alignments were performed, visualised, and manipulated using Jalview version 2 (Waterhouse *et al*, 2009). HMM construction, searches, and alignments were conducted using HMMER (2.3.2) (Eddy, 1998). Models of Chd1-DBD bound to a 20-bp double-stranded DNA were generated using the HADDOCK web server (de Vries *et al*, 2010). Further information is provided in the Supplementary data.

Structure deposition

Coordinates and structure factors for Chd1-DBD have been deposited in the PDB under ID 2XB0.

Supplementary data

Supplementary data are available at *The EMBO Journal* Online (http://www.embojournal.org).

Acknowledgements

We thank Bill Hunter for access to crystal screening facilities. This work was funded by a Wellcome Trust Senior Fellowship 064414 (DPR, RS, VS, and TOH) and initially by a NHMRC Australia CJ Martin Fellowship (457137; DR) and BBSRC project grant BB/E022286/1 (RS). We also thank Katarzyna Kozyrska who assisted with the establishment of the *in vivo* complementation assay.

Author contributions: DPR designed and carried out experiments and analysed data; RS and VS assisted with experiments and data analysis; DM assisted with bioinformatic analysis; TOH supervised the project, designed experiments, and analysed data; DPR and TOH wrote the manuscript.

Conflict of interest

The authors declare that they have no conflict of interest.

References

Aasland R, Stewart AF, Gibson T (1996) The SANT domain: a putative DNA-binding domain in the SWI-SNF and ADA complexes, the transcriptional co-repressor N-CoR and TFIIIB. *Trends Biochem Sci* **21:** 87–88

Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung LW, Kapral GJ, Grosse-Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner R, Read RJ, Richardson DC, Richardson JS, Terwilliger TC, Zwart PH (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **66:** 213–221

Alen C, Kent NA, Jones HS, O'Sullivan J, Aranda A, Proudfoot NJ (2002) A role for chromatin remodeling in transcriptional termination by RNA polymerase II. *Mol Cell* **10:** 1441–1452

Awad S, Ryan D, Prochasson P, Owen-Hughes T, Hassan AH (2010) The Snf2 homolog Fun30 acts as a homodimeric ATP-dependent chromatin-remodeling enzyme. *J Biol Chem* **285:** 9477–9484

Bailey S (1994) The Ccp4 suite—programs for protein crystallography. *Acta Crystallogr D* **50:** 760–763

Baker NA, Sept D, Joseph S, Holst MJ, McCammon JA (2001) Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc Natl Acad Sci USA* **98:** 10037–10041

Baudin A, Ozierkalogeropoulos O, Denouel A, Lacroute F, Cullin C (1993) A simple and efficient method for direct gene deletion in Saccharomyces cerevisiae. *Nucleic Acids Res* **21:** 3329–3330

Bird A (2002) DNA methylation patterns and epigenetic memory. *Genes Dev* **16:** 6–21

Biswas D, Dutta-Biswas R, Stillman DJ (2007) Chd1 and yFACT act in opposition in regulating transcription. *Mol Cell Biol* **27:** 6279–6287

Biswas D, Takahata S, Xin H, Dutta-Biswas R, Yu YX, Formosat T, Stillman DJ (2008) A role for Chd1 and Set2 in negatively regulating DNA replication in Saccharomyces cerevisiae. *Genetics* **178:** 649–659

Blank TA, Becker PB (1995) Electrostatic mechanism of nucleosome spacing. *J Mol Biol* **252:** 305–313

Blosser TR, Yang JG, Stone MD, Narlikar GJ, Zhuang X (2009) Dynamics of nucleosome remodelling by individual ACF complexes. *Nature* **462:** 1022–1027

Boyer LA, Latek RR, Peterson CL (2004) The SANT domain: a unique histone-tail-binding module? *Nat Rev Mol Cell Bio* **5:** 158–163

Cairns BR (2007) Chromatin remodeling: insights and intrigue from single-molecule studies. *Nat Struct Mol Biol* **14:** 989–996

Chen VB, Arendall III WB, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, Murray LW, Richardson JS, Richardson DC (2010) MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* **66:** 12–21

Cheng J, Randall AZ, Sweredoski MJ, Baldi P (2005) SCRATCH: a protein structure and structural feature prediction server. *Nucleic Acids Res* **33:** W72–W76

Clapier CR, Cairns BR (2009) The biology of chromatin remodeling complexes. *Annu Rev Biochem* **78:** 273–304

Clapier CR, Langst G, Corona DF, Becker PB, Nightingale KP (2001) Critical role for the histone H4 N terminus in nucleosome remodeling by ISWI. *Mol Cell Biol* **21:** 875–883

Corona DF, Langst G, Clapier CR, Bonte EJ, Ferrari S, Tamkun JW, Becker PB (1999) ISWI is an ATP-dependent nucleosome remodeling factor. *Mol Cell* **3:** 239–245

Dang WW, Bartholomew B (2007) Domain architecture of the catalytic subunit in the ISW2-nucleosome complex. *Mol Cell Biol* **27:** 8306–8317

Daubresse G, Deuring R, Moore L, Papoulas O, Zakrajsek I, Waldrip WR, Scott MP, Kennison JA, Tamkun JW (1999) The Drosophila kismet gene is related to chromatin-remodeling factors and is required for both segmentation and segment identity. *Development* **126:** 1175–1187

de Vries SJ, van Dijk M, Bonvin AM (2010) The HADDOCK web server for data-driven biomolecular docking. *Nat Protoc* **5:** 883–897

Delmas V, Stokes DG, Perry RP (1993) A mammalian DNA-binding protein that contains a chromodomain and an SNF2/SWI2-like helicase domain. *Proc Natl Acad Sci USA* **90:** 2414–2418

Eberharter A, Ferrari S, Langst G, Straub T, Imhof A, Varga-Weisz P, Wilm M, Becker PB (2001) Acf1, the largest subunit of CHRAC, regulates ISWI-induced nucleosome remodelling. *EMBO J* **20:** 3781–3788

Eddy SR (1998) Profile hidden Markov models. *Bioinformatics* **14:** 755–763

Emsley P, Lohkamp B, Scott WG, Cowtan K (2010) Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **66:** 486–501

Fazzio TG, Gelbart ME, Tsukiyama T (2005) Two distinct mechanisms of chromatin interaction by the Isw2 chromatin remodeling complex *in vivo*. *Mol Cell Biol* **25:** 9165–9174

Ferreira H, Flaus A, Owen-Hughes T (2007) Histone modifications influence the action of Snf2 family remodelling enzymes by different mechanisms. *J Mol Biol* **374:** 563–579

Flanagan JF, Blus BJ, Kim D, Clines KL, Rastinejad F, Khorasanizadeh S (2007) Molecular implications of evolutionary differences in CHD double chromodomains. *J Mol Biol* **369:** 334–342

Flaus A, Martin DM, Barton GJ, Owen-Hughes T (2006) Identification of multiple distinct Snf2 subfamilies with conserved structural motifs. *Nucleic Acids Res* **34:** 2887–2905

Gangaraju VK, Bartholomew B (2007) Mechanisms of ATP dependent chromatin remodeling. *Mutat Res* **618:** 3–17

Gaspar-Maia A, Alajem A, Polesso F, Sridharan R, Mason MJ, Heidersbach A, Ramalho-Santos J, McManus MT, Plath K, Meshorer E, Ramalho-Santos M (2009) Chd1 regulates open chromatin and pluripotency of embryonic stem cells. *Nature* **460:** 863–868

Grune T, Brzeski J, Eberharter A, Clapier CR, Corona DF, Becker PB, Muller CW (2003) Crystal structure and functional analysis of a nucleosome recognition module of the remodeling factor ISWI. *Mol Cell* **12:** 449–460

Hamiche A, Kang JG, Dennis C, Xiao H, Wu C (2001) Histone tails modulate nucleosome mobility and regulate ATP-dependent nucleosome sliding by NURF. *Proc Natl Acad Sci USA* **98:** 14316–14321

Hauk G, McKnight JN, Nodelman IM, Bowman GD (2010) The chromodomains of the Chd1 chromatin remodeler regulate DNA access to the ATPase motor. *Mol Cell* **39:** 711–723

Holm L, Rosenström P (2010) Dali server: conservation mapping in 3D. *Nucleic Acids Res* **38:** W545–W549

Konev AY, Tribus M, Park SY, Podhraski V, Lim CY, Emelyanov AV, Vershilova E, Pirrotta V, Kadonaga JT, Lusser A, Fyodorov DV (2007) CHD1 motor protein is required for deposition of histone variant h3.3 into chromatin *in vivo*. *Science* **317:** 1087–1090

Kouzarides T (2007) Chromatin modifications and their function. *Cell* **128:** 693–705

Lin K, Simossis VA, Taylor WR, Heringa J (2005) A simple and fast secondary structure prediction method using hidden neural networks. *Bioinformatics* **21:** 152–159

Lusser A, Urwin DL, Kadonaga JT (2005) Distinct activities of CHD1 and ACF in ATP-dependent chromatin assembly. *Nat Struct Mol Biol* **12:** 160–166

McDaniel IE, Lee JM, Berger MS, Hanagami CK, Armstrong JA (2008) Investigations of CHD1 function in transcription and development of Drosophila melanogaster. *Genetics* **178:** 583–587

Morettini S, Tribus M, Zeilner A, Sebald J, Campo-Fernandez B, Scheran G, Worle H, Podhraski V, Fyodorov DV, Lusser A (2011) The chromodomains of CHD1 are critical for enzymatic activity but less important for chromatin localization. *Nucleic Acids Res* **39:** 3103–3115

Okuda M, Horikoshi M, Nishimura Y (2007) Structural polymorphism of chromodomains in Chd1. *J Mol Biol* **365:** 1047–1062

Painter J, Merritt EA (2006) Optimal description of a protein structure in terms of multiple groups undergoing TLS motion. *Acta Crystallogr D Biol Crystallogr* **62:** 439–450

Park YJ, Luger K (2008) Histone chaperones in nucleosome eviction and histone exchange. *Curr Opin Struct Biol* **18:** 282–289

Pinskaya M, Nair A, Clynes D, Morillon A, Mellor J (2009) Nucleosome remodeling and transcriptional repression are distinct functions of Isw1 in Saccharomyces cerevisiae. *Mol Cell Biol* **29:** 2419–2430

Racki LR, Yang JG, Naber N, Partensky PD, Acevedo A, Purcell TJ, Cooke R, Cheng Y, Narlikar GJ (2009) The chromatin remodeller ACF acts as a dimeric motor to space nucleosomes. *Nature* **462:** 1016–1021

Robinson KM, Schultz MC (2003) Replication-independent assembly of nucleosome arrays in a novel yeast chromatin reconstitution system involves antisilencing factor Asf1p and chromodomain protein Chd1p. *Mol Cell Biol* **23:** 7937–7946

Sikorski R, Hieter P (1989) A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in Saccharomyces cerevisiae. *Genetics* **122:** 19–27

Simic R, Lindstrom DL, Tran HG, Roinick KL, Costa PJ, Johnson AD, Hartzog GA, Arndt KM (2003) Chromatin remodeling protein Chd1 interacts with transcription elongation factors and localizes to transcribed genes. *EMBO J* **22:** 1846–1856

Sims RJ, Chen CF, Santos-Rosa H, Kouzarides T, Patel SS, Reinberg D (2005) Human but not yeast CHD1 binds directly and selectively to histone H3 methylated at lysine 4 via its tandem chromodomains. *J Biol Chem* **280:** 41789–41792

Stockdale C, Flaus A, Ferreira H, Owen-Hughes T (2006) Analysis of nucleosome repositioning by yeast ISWI and Chd1 chromatin remodeling complexes. *J Biol Chem* **281:** 16279–16288

Stokes DG, Perry RP (1995) DNA-binding and chromatin localization properties of Chd1. *Mol Cell Biol* **15:** 2745–2753

Thastrom A, Lowary PT, Widlund HR, Cao H, Kubista M, Widom J (1999) Sequence motifs and free energies of selected natural and non-natural nucleosome positioning DNA sequences. *J Mol Biol* **288:** 213–229

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res* **25:** 4876–4882

Tran HG, Steger DJ, Iyer VR, Johnson AD (2000) The chromo domain protein chd1p from budding yeast is an ATP-dependent chromatin-modifying factor. *EMBO J* **19:** 2323–2331

Tsukiyama T, Palmer J, Landel CC, Shiloach J, Wu C (1999) Characterization of the imitation switch subfamily of ATP-dependent chromatin-remodeling factors in Saccharomyces cerevisiae. *Genes Dev* **13:** 686–697

Walfridsson J, Bjerling P, Thalen M, Yoo EJ, Park SD, Ekwall K (2005) The CHD remodeling factor Hrp1 stimulates CENP-A loading to centromeres. *Nucleic Acids Res* **33:** 2868–2879

Waterhouse AM, Procter JB, Martin DMA, Clamp M, Barton GJ (2009) Jalview Version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics* **25:** 1189–1191

Yang JG, Madrid TS, Sevastopoulos E, Narlikar GJ (2006) The chromatin-remodeling enzyme ACF is an ATP-dependent DNA length sensor that regulates nucleosome spacing. *Nat Struct Mol Biol* **13:** 1078–1083

Zentner GE, Layman WS, Martin DM, Scacheri PC (2010) Molecular and phenotypic aspects of CHD7 mutation in CHARGE syndrome. *Am J Med Genet A* **152A:** 674–686