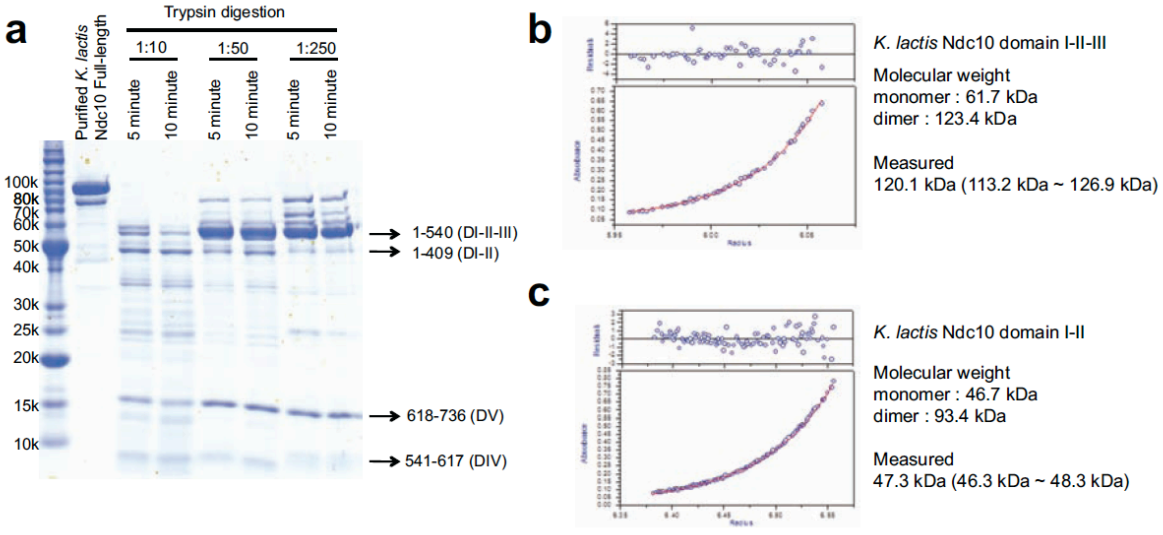
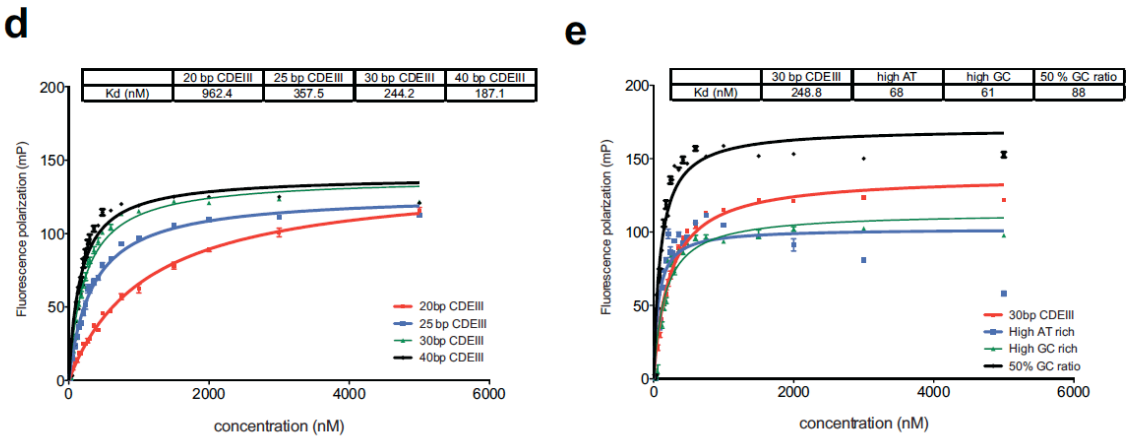
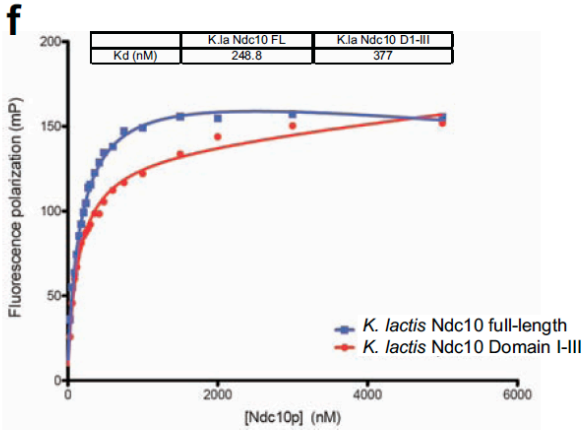
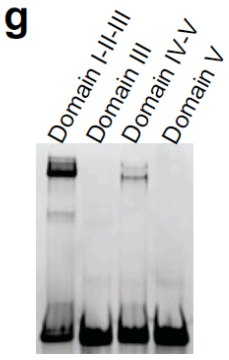
Supplementary Information

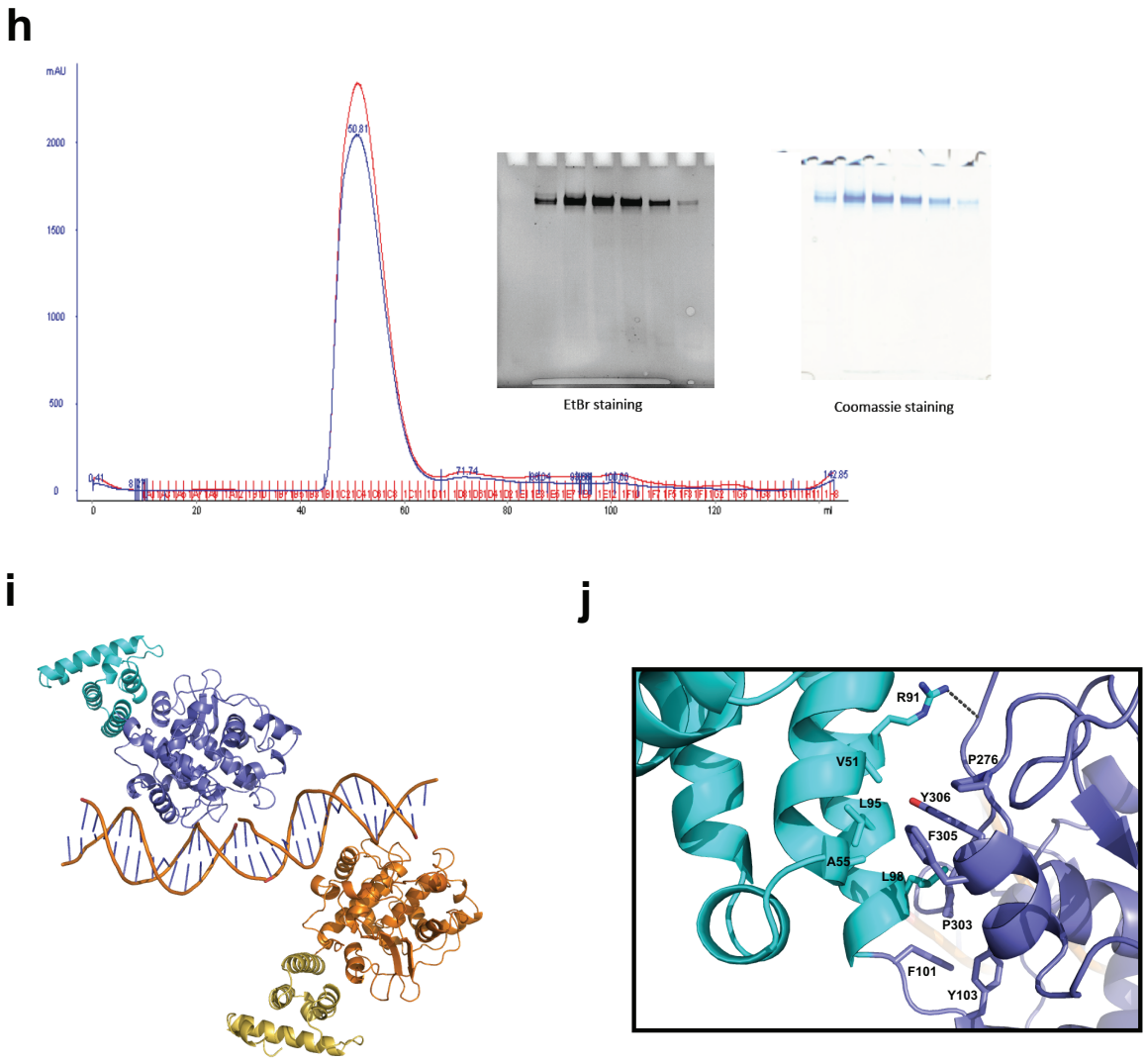
Supplementary Data





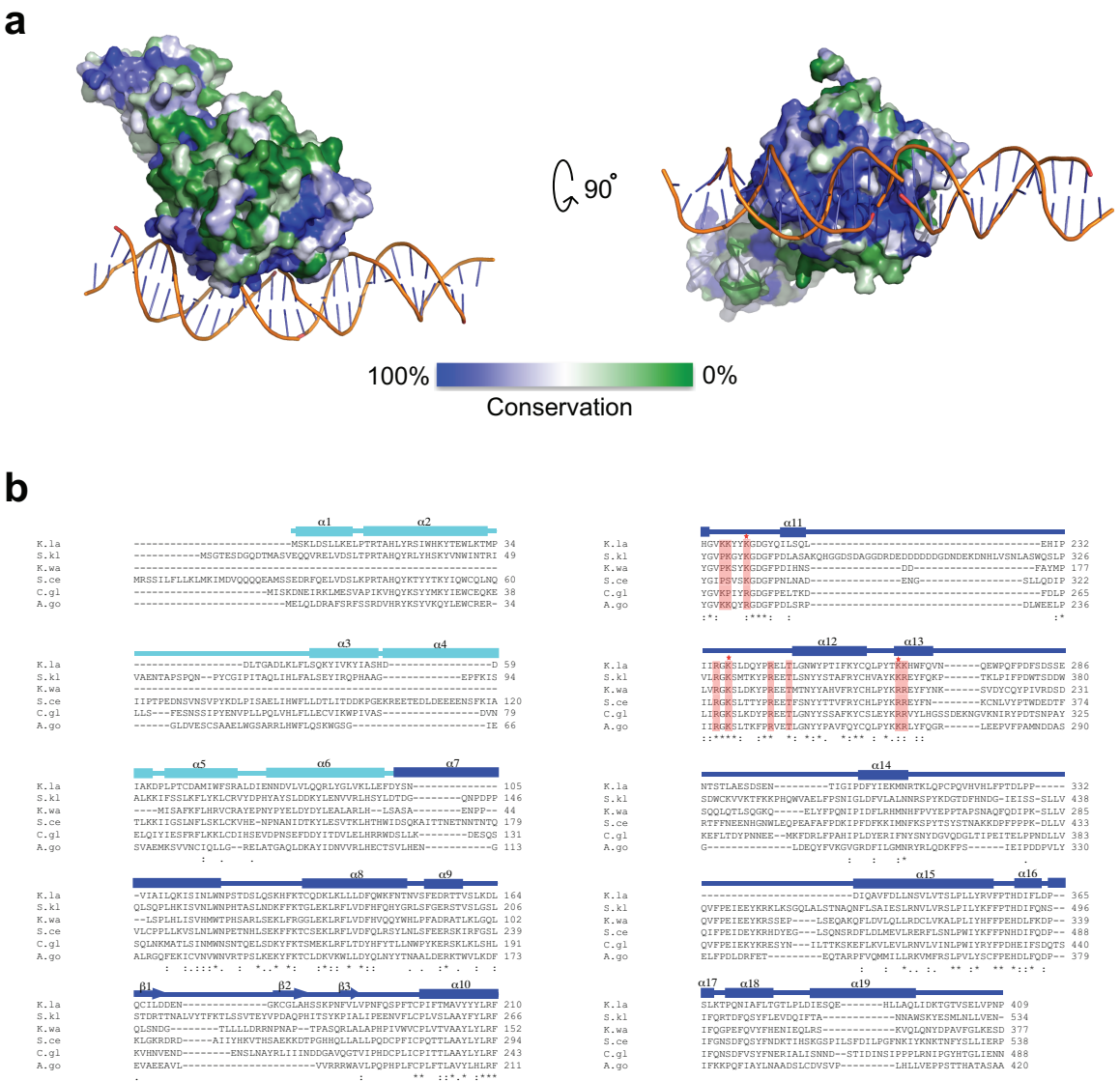






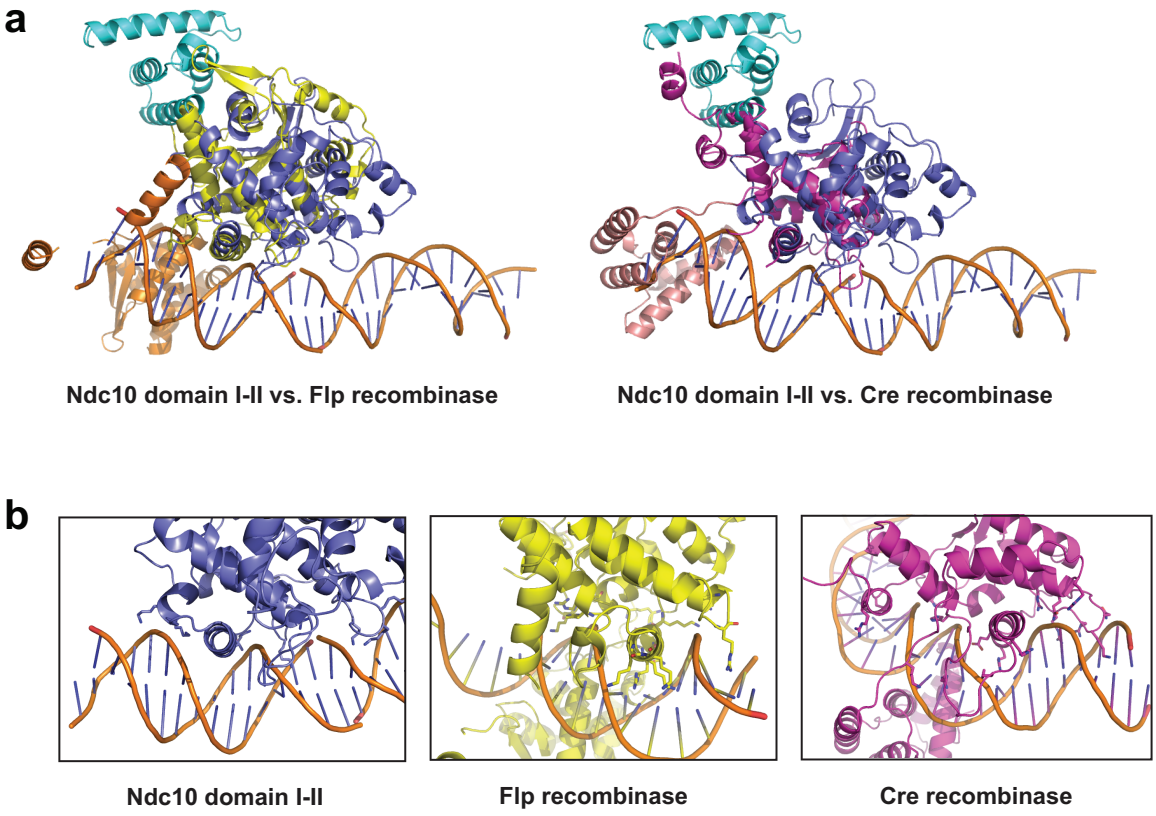
Supplementary Figure 1. Biochemical properties and structure of *K. lactis* Ndc10. **(a)** Limited trypsin proteolysis of *K. lactis* full-length Ndc10. Trypsin was mixed in ratios of 1:10, 1:50, 1:250 (w/w) with purified *K. lactis* Ndc10 and incubated for 5 or 10 minutes at room temperature. The 1:50 (w/w) treated sample (10 minutes) was analyzed by mass spectrometry, and four major fragments were identified: residues 1–540 (DI–III), 1–409 (DI–II), 618–736 (DV), and 541–617 (DIV). **(b–c)** Sedimentation equilibrium analysis of *K. lactis* Ndc10 DI–III and DI–II. The measured molecular mass of Ndc10 DI–III corresponds to a dimer; that of Ndc10 DI–II, to a monomer. **(d–f)** Fluorescence

anisotrophy measurements of *K. lactis* Ndc10 with 5'-fluorescein-tagged DNA. **(d)** Binding of full-length *K. lactis* Ndc10 with DNA of various lengths. **(e)** Binding of full-length *K. lactis* Ndc10 with DNAs of different GC-content. **(f)** Binding of full-length Ndc10 and Ndc10 DI–III with 30 bp CDEIII DNA. **(g)** Electrophoretic mobility shift assay of Ndc10 domains with 30 bp CDEIII DNA. The TBE-buffered, 10% polyacrylamide gel was stained with ethidium bromide (EtBr). **(h)** Purification of *K. lactis* Ndc10 with CDEIII DNA. Full-length Ndc10 was added to 30 bp of CDEIII DNA (including one A/T overhang) and applied to a Superdex 200 size-exclusion column (GE Healthcare). In the elution profile, the blue line represents abosrbance$_{280nm}$; the red line, absorbance$_{260nm}$. Fractions were loaded (without denaturation) onto a TBE-buffered, 10% polyacrylamide gel, which was stained after electrophoresis with both EtBr and Coomassie blue. **(i)** Symmetry-related Two Ndc10 DI–II monomers related by a crystallographic twofold screw axis parallel to c (and along the end-to-end stacked DNA) in space group $C222_1$. Ndc10 DI–II contacts more than 20 bp of CDEIII DNA but shares a 30 bp length of DNA with another, symmetry-related molecule. **(j)** Interface of Ndc10 DI and DII, showing well-packed hydrophobic residues and a hydrogen bond (Arg91).

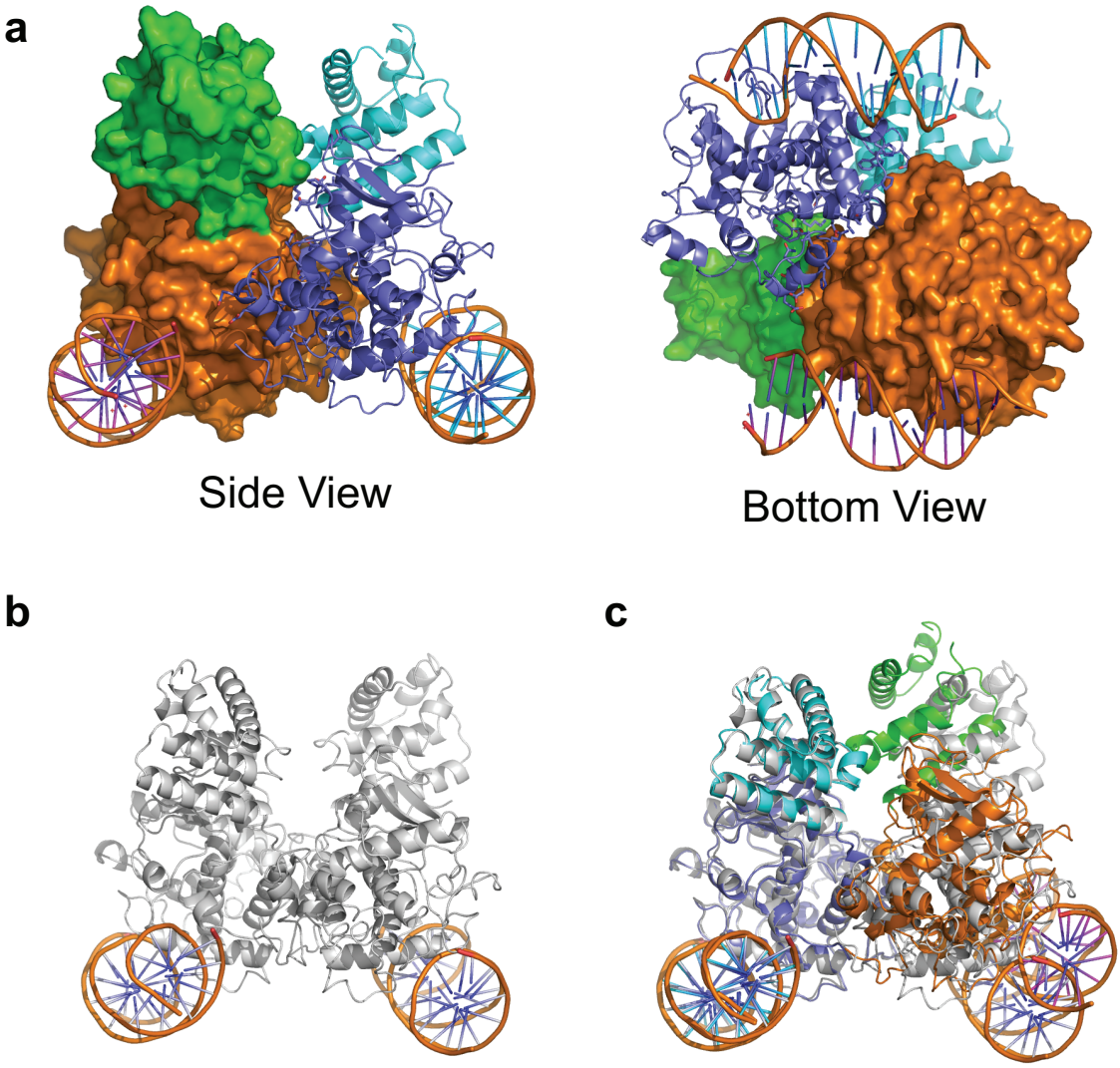


Supplementary Figure 2. Sequence conservation and multiple sequence alignment of Ndc10 DI–II from point-centromere yeasts. (a) Surface views of *K. lactis* Ndc10 DI–II, colored according to amino-acid conservation among point centromere yeasts, based on the multiple sequence alignment (blue=conserved, green=variable). **(b)** Alignment of Ndc10 DI–II sequences from *Kluyveromyces lactis (K.la), Saccharomyces kluyveri (S.kl), Kluyveromyces waltii (K.wa), Saccharomyces cerevisiae (S.ce), Candida glabrata (C.gl),* and *Ashbya gossypii (A.go)*. Residues in contact with DNA contact are highlighted in

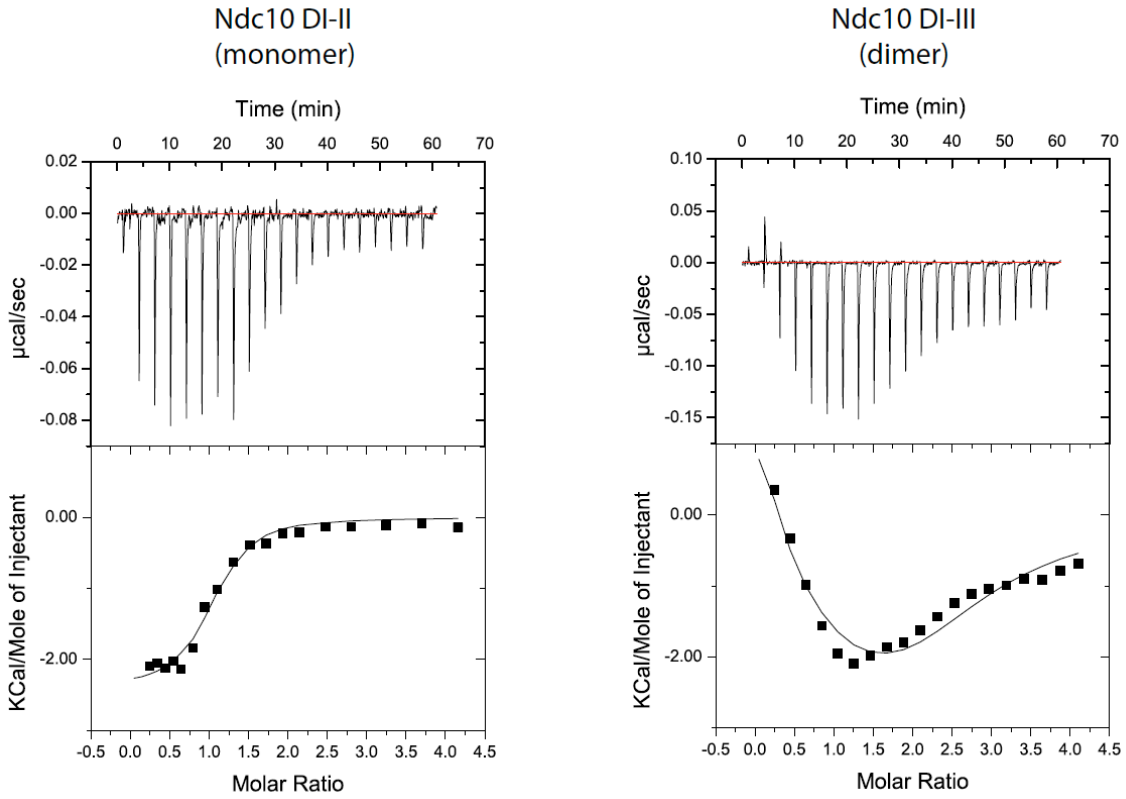
orange; those that donate hydrogen bonds to phosphates from main-chain amide groups are also marked with a red asterisk.



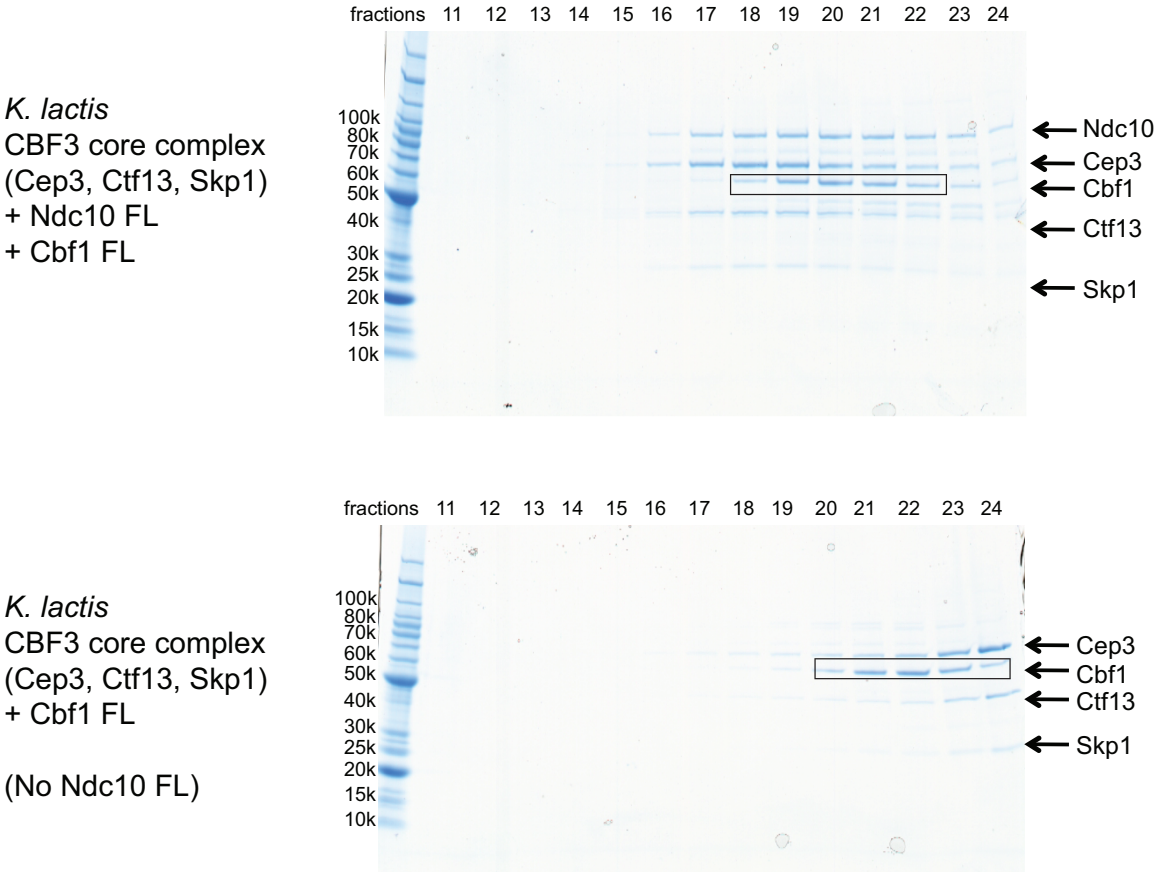
Supplementary Figure 3. Structural overlay of Ndc10 DI–II with Flp and Cre. (a) Structures of Flp and Cre DNA-binding domain are aligned with the structure of *K. lactis* Ndc10 DII. The figure shows only the 30 bp CDEIII DNA from the Ndc10 structure: the Holliday junction DNA from the Cre and Flp complexes would overlap the CDEIII DNA in the region of protein contact. The color scheme is the same as in Fig. 3. **(b)** DNA contacts of Ndc10 DI–II, Flp, and Cre. An α-helix in Flp (α12) and Cre (α10) contacts DNA base-pairs in the major groove, but the corresponding α-helix in Ndc10 DI–II (α12) does not.



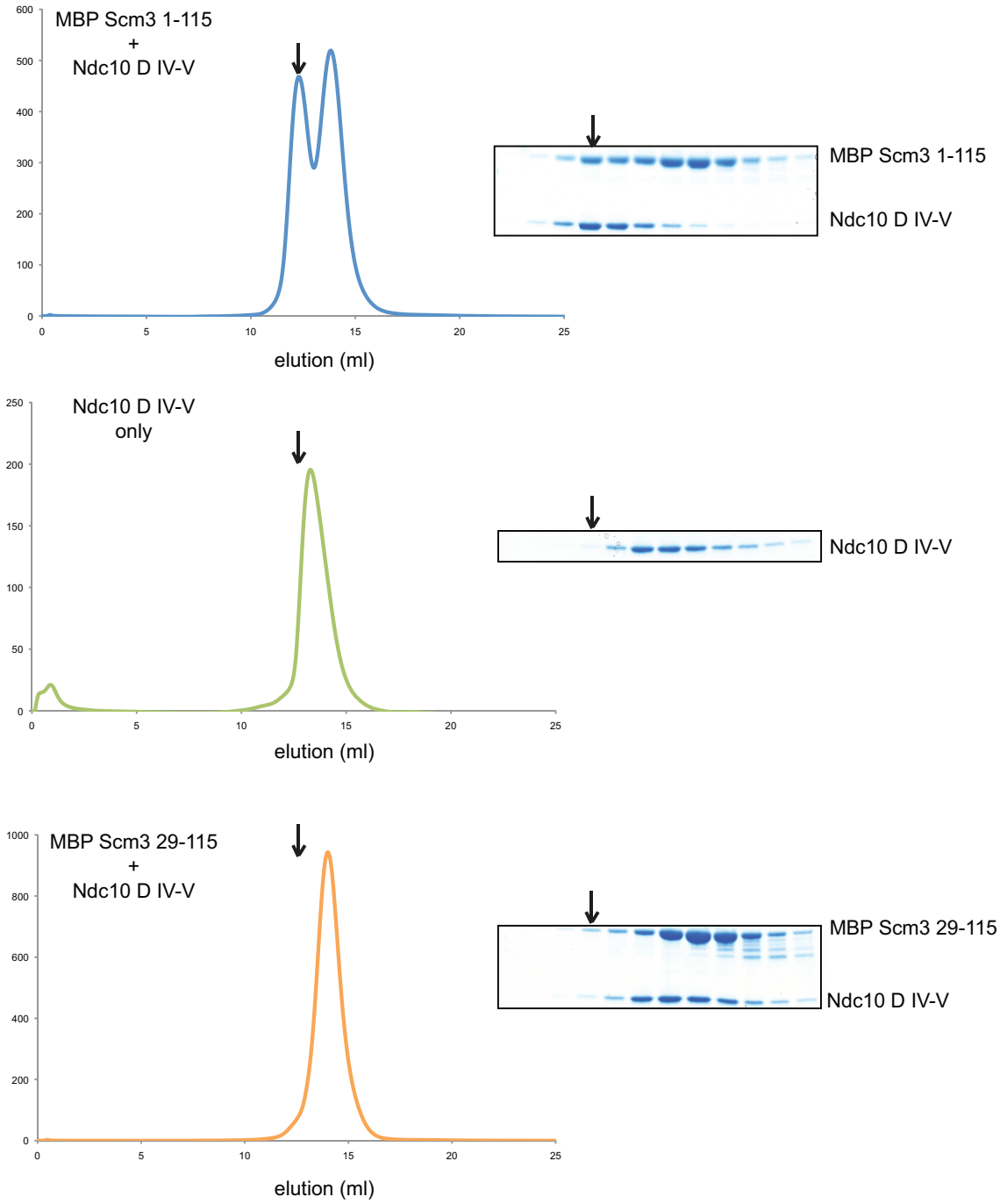
Supplementary Figure 4. Twofold interface in the crystal structure of *K. lactis* Ndc10 DI–II. (a) Potential Ndc10 dimer, related by a crystallographic twofold axis parallel to b axis in space group $C222_1$. Buried surface area is ~800 Å^2. **(b)** The dimer of Ndc10 DI–II in space group $P2_12_12_1$. The packing environment in the two crystal forms ($C222_1$ and $P2_12_12_1$) is quite different, except for the interface shown. **(c)** Superposition of Ndc10 DI–II dimer complexed with CDEIII DNA in space groups $C222_1$ and $P2_12_12_1$. The $P2_12_12_1$ structure from is in gray; the $C222_1$ structure colored as in panel **a**.



Supplementary Figure 5. Isothermal titration calorimetry of Ndc10 DI–II and Ndc10 DI–III with 30 bp CDEIII DNA. Either Ndc10 DI–II (monomer) or Ndc10 DI–III (dimer), at a concentration of 25 μM, was in the reaction cell and 0.5 mM 30 bp CDEIII DNA, in the syringe. Raw data and integrated heats are shown on the top and bottom charts, respectively. Ndc10 DI–II monomer binds to DNA in a 1:1 molar ratio; Ndc10 DI–III dimer binds DNA in a 1:2 molar ratio. Each of the initial steps of DNA addition to Ndc10 dimer, up to about a 1:1 molar ratio, shows an overlapped mixture of a rapid exothermal phase and slightly slower endothermal phase. We interpret this observation as either an allosteric effect of binding Ndc10 dimer to DNA[1] or as a contribution from non-specific interaction of the other subunit in a dimer (with high effective local concentration) at low DNA: protein ratios[2].

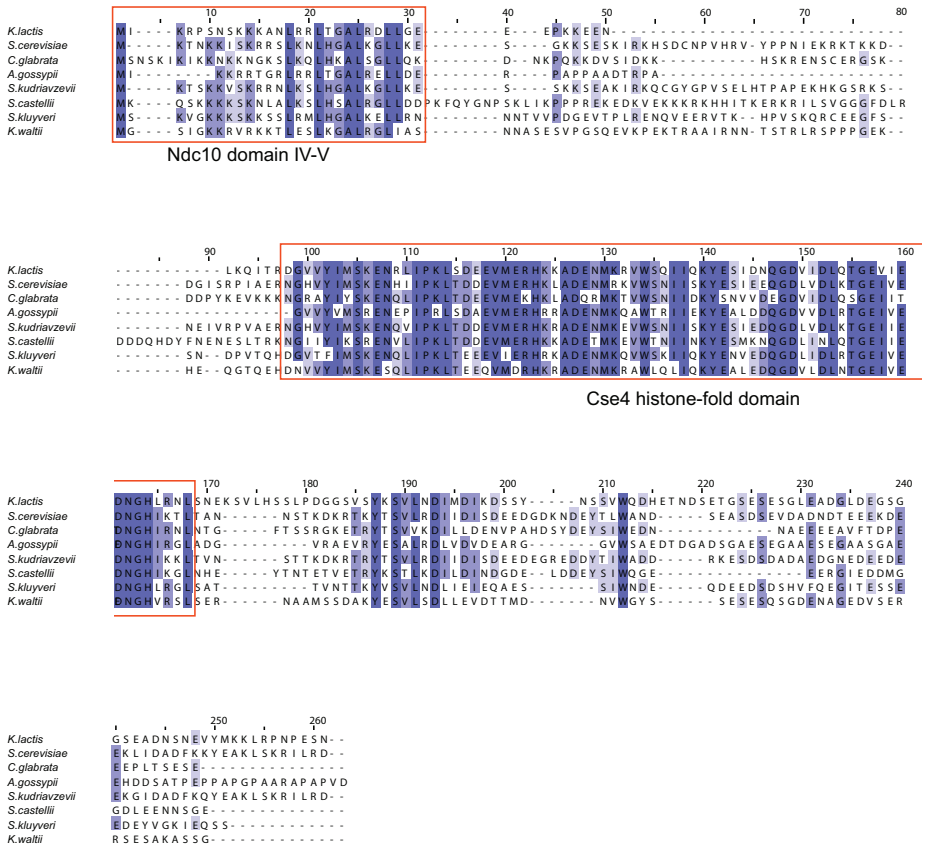


Supplementary Figure 6. Analytical size-exclusion chromatography of *K. lactis* CBF3 core-Ndc10-Cbf1 and CBF3 core-Cbf1. Equimolar quantities of *K. lactis* CBF3 core [Ctf13:Skp1:$(Cep3)_2$] and Cbf1 were incubated in either the presence or absence of Ndc10 and the mixture applied to a size-exclusion column (Superdex 200 10/300: GE Healthcare). The figure shows SDS-PAGE analysis of fractions. The upper panel shows that Cbf1 co-elutes with holo-CBF3 (CBF3 core plus one Ndc10 dimer); the lower panel shows that Cbf1 and CBF3 core do not elute together.



Supplementary Figure 7. Analytical size-exclusion chromatography of Ndc10 DIV–V alone, Ndc10 DIV–V with MBP tagged residues 1–115 of Scm3, and Ndc10 DIV–V with MBP tagged residues 29–115 of Scm3. Purified Ndc10 DIV–V was incubated with MBP tagged Scm3 constructs and ran a size-exclusion chromatography (Superdex

200 10/300: GE Healthcare). Black arrows represent the position of the peak fraction for the Ndc10 DIV–V: MBP Scm3 1–115 complex. As shown, Scm3 1–115, but not Scm3 29–115, forms a complex with Ndc10 DIV–V.



Supplementary Figure 8. Multiple sequence alignment of Scm3 among point centromere yeasts. The N-terminal region (Ndc10 interaction) and the central region (Cse4:H4 interaction) of Scm3 show the conserved feature among point centromeres (red box).

Supplementary Methods

DNA sequences (fluorescence polarization experiment) :

20 bp *K. lactis* CEN1 CDEIII DNA: 5'-TGCTTTATGTTTCCGAAAAT-3'

25 bp *K. lactis* CEN1 CDEIII DNA: 5'-TGCTTTATGTTTCCGAAAATTATTT-3'

30 bp *K. lactis* CEN1 CDEIII DNA:

5'-TGCTTTATGTTTCCGAAAATTATTTTATTA-3'

40 bp *K. lactis* CEN1 CDEIII DNA:

5'-TGCTTTATGTTTCCGAAAATTATTTTATTAGTGAA-3'

High GC (86.7 % of GC ratio):

5'-GTGCGCCTCCGGGCGGAGCCCCACCCCCGG-3'

High AT (13.3 % of GC ratio):

5'-AGTATAAGAATTTATTCTAAAACAAAAATT-3'

50% GC (50% of GC ratio):

5'-GCGGTAATACGGTTATCCACAGAATCAGGG-3'

Cloning, expression, and purification of *K. lactis* Ndc10

A clone of full-length (FL) *K. lactis* Ndc10 was obtained by PCR amplification from a genomic library of *Kluyveromyces lactis* and sub-cloned into a pET vector modified for LIC (ligation independent cloning) and encoding an N-terminal His-tag and a TEV protease cleavage site. This construct was then transformed into Rosetta (DE3) pLysS *E. coli* competent cells (Novagen) and used for protein expression. Cells were grown at 37°C in ZYM-5052 auto-inducible medium; the temperature was switched to

25°C when the O.D. reached 1.0, and the incubation continued overnight at the lower temperature[3]. Cells were harvested and resuspended in binding buffer (30 mM Tris-HCl, pH8.0, 500 mM NaCl, 3 mM β-mercaptoethanol) with protease inhibitor cocktail tablets. Cells were sonicated and the cell debris removed by centrifugation at 20,000 xg for 1 hour. Soluble fractions were then passed through a 0.45mm filter and applied to a Ni-NTA column (Qiagen) pre-equilibrated with binding buffer. The column was washed with binding buffer and washing buffer (30 mM TrisHCl, pH8.0, 500 mM NaCl, 30 mM imidazole, 3 mM β-mercaptoethanol) and eluted with elution buffer (30 mM Tris-HCl. pH8.0, 500 mM NaCl, 300 mM imidazole, 3 mM β-mercaptoethanol). TEV protease was added to the eluted fraction in a 1:100 (w/w) ratio, and the sample dialyzed into 30 mM Tris-HCl, pH8.0, 100 mM NaCl, 1 mM DTT overnight at 4°C. The dialyzed proteins were re-applied to the Ni-NTA column to eliminate His-tag fragments, and flow-through fractions were then applied into HiTrap Q HP column 5 ml (GE healthcare) pre-equilibrated with 30 mM Tris-HCl, pH8.0, 100 mM NaCl, 1 mM DTT and eluted by gradually increasing NaCl concentration up to 0.5 M. Fractions containing target proteins were pooled, concentrated, and applied to a Superdex 200 size exclusion column (Prep grade 16/60: GE healthcare) pre-equilibrated with 30 mM Tris-HCl, pH8.0, 100 mM NaCl, 1 mM TCEP (tris(2-carboxyethyl)phosphine)(buffer A). Fractions containing target proteins were pooled and concentrated and used for further biochemical studies. All other *K. lactis* Ndc10 constructs including mutants were generated, expressed, and purified with the same procedure. *K. lactis* CBF3 core was sub-cloned as separate pairs of Ctf13–Skp1 and Cep3–Sgt1, into pFastBac-Duet plasmids (Invitrogen), modified for ligation independent cloning, with an N-terminal His-tag on one member of each pair.

The two baculoviruses were co-infected in Hi-5 cells and harvested after 72 hours. *K. lactis* Cbf1 was sub-cloned into a pET vector, modified for ligation independent cloning, containing a N-terminal His-tag. Proteins were further purified by affinity, ion-exchange, and size-exclusion column chromatography.

Sedimentation equilibrium analytical ultracentrifugation

Sedimentation equilibrium experiments were carried out with a Beckman-Coulter ProteomeLab XL-I analytical ultracentrifuge equipped with 12-mm Epon double-sector cells in an An-60 Ti rotor (Beckman). Solutions were prepared in buffer A at three different protein concentrations, corresponding to optical densities at 280nm between 0.25 and 0.75 (Ndc10 DI–III : 2–6 μM, Ndc10 DI–II : 5–15 μM), and measurements were made at three different speeds. At each speed, equilibrium was achieved within 30 hours. A multiple fit alignment was performed using the XL-I software package. The fitting model used was:

$$C_r=C_{r0}\{\exp[(\omega^2/2RT)M(1-\nu\rho)(r^2-r_0^2)]\} + \Delta E$$

where C_r is the concentration at radius r, C_{r0} is the concentration of the protein at the reference radius r_0, ω is the angular velocity, R is the gas constant, T is temperature in Kelvin, M is the protein molecular weight, ν is the partial specific volume of the solute, and ρ is the density of the solvent. ΔE is the baseline offset parameter used in fitting.

Isothermal titration calorimetry

Isothermal titration calorimetry was performed with an iTC200 microcalorimeter (Microcal) interfaced with a computer running Origin 7.0 software. Protein, at 25 μM in

30 mM Tris-HCl [pH8.0], 100 mM NaCl, and 1 mM TCEP, was titrated with 0.5 mM 30 bp CDEIII DNA in the same buffer. The reaction cell volume was 200 μl and the injection syringe volume, 60 μl. The titration used a first injection of 0.5 μl that was not used in data analysis, followed by 19 injections of 1~3 μl each.

Purification and crystallization of a *K. lactis* CEN1 CDEIII - *K. lactis* Ndc10 DI–II Complex

Purified *K. lactis* Ndc10 DI–II was mixed in 1:1 molar ratio with a DNA fragment corresponding to 29 bp of CDEIII plus a single A/T overhang, and the sample was applied to a Superdex200 (10/300) size exclusion column (GE healthcare) pre-equilibrated with buffer A (Supplementary Fig. 1h). The protein-DNA complex formation was confirmed by measuring absorbance at 280 nm and 260 nm, and by native gel electrophoresis followed by EtBr and Coomassie staining (Supplementary Fig. 1h). The protein-DNA complex was concentrated to 10 mg/ml and used for crystallization. The complex crystallized in 20% PEG400, 0.1 M MES (2-(*N*-morpholino) ethanesulfonic acid) pH 6.0, 0.2 M NaCl. The needle-shaped crystals, in space group $C222_1$, a=117.29, b=147.51, c=95.63, were improved by micro-seeding. For x-ray diffraction data collection, crystals were transferred to 35 % PEG4000 and flash frozen in liquid nitrogen. Diffraction data were recorded to a minimum Bragg spacing of 2.8 Å on the NE-CAT 24C/E beamlines at the Advanced Photon Source (Argonne National Laboratory).

Seleno(Se)-methionine labeled protein was expressed with PASM-5052 autoinducible medium[3], and the protein-DNA complex was purified with the procedure described above. Single wavelength anomalous dispersion (SAD) data (3.0 Å minimum

Bragg spacing) were collected at NE-CAT beamline 24C at the peak wavelength of selenium. ShelxD[4] located 4 of the 5 methionine sites, but the anomalous signal was too weak to visualize protein secondary structure in the resulting density map. The map did, however, reveal DNA density clearly enough that we could build a 15 bp B-DNA duplex (one complete asymmetric unit). Experimental phases were then combined, using Phaser[5], with phases derived from the 15 bp DNA model and a new map calculated. This map revealed sufficiently clear protein density to commence iterative rounds of model building and phase combination. Models were built using the program COOT[6], and refinement calculations carried out using PHENIX.REFINE[7]. The refined model, with all residues from 1 to 402, except for disordered loops from 36 to 39 and 283 to 292, has $R_{work} = 0.193$ and $R_{free} = 0.251$.

As in many crystals of protein–DNA complexes, the DNA fragments stack end-to-end in a pseudo-continuous DNA duplex. The A/T overhang allows successive fragments to overlap by forming a base pair. In the $C222_1$ crystals, the DNA stacks along the c-axis twofold screw, which relates one half of the 30 bp fragment to the other. The map thus shows averaged densities for base pairs, and we chose to build randomly selected A:T or T:A base pairs, rather than a defined sequence.

A second crystal form (also needles) of the same complex was obtained from 25% PEG 10,000, 0.1 M Tris-HCl, pH8.5. These crystals were flash frozen with 20% glycerol as cryo-protectant. Diffraction data to a minimum Bragg spacing of 3.6 Å were recorded at NE-CAT beamline 24C. The space group was $P2_12_12_1$, a=93.1, b=99.7, c=125.9, with a protein dimer and 30 bp DNA in the asymmetric unit. Molecular replacement using the refined Ndc10 DI–II ($C222_1$) structure was used to determine the structure, which was

refined with Phenix[7]. The refined model (residues 1–402, except for disordered loops from 36 to 39 and 283 to 292) has R_{work} = 28.6 and R_{free} = 33.1.

Supplementary References

1. Jin, L. et al. Asymmetric allosteric activation of the symmetric ArgR hexamer. *J. Mol. Biol.* **346**, 43-56 (2005).
2. Holbrook, J.A., Tsodikov, O.V., Saecker, R.M. & Record, M.T., Jr. Specific and non-specific interactions of integration host factor with DNA: thermodynamic evidence for disruption of multiple IHF surface salt-bridges coupled to DNA binding. *J. Mol. Biol.* **310**, 379-401 (2001).
3. Studier, F.W. Protein production by auto-induction in high density shaking cultures. *Protein Expr. Purif.* **41**, 207-234 (2005).
4. Sheldrick, G.M. A short history of SHELX. *Acta Crystallogr. A* **64**, 112-122 (2008).
5. McCoy, A.J. et al. Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658-674 (2007).
6. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126-2132 (2004).
7. Adams, P.D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 213-221 (2010).