

# The origin recognition complex requires chromatin tethering by a hypervariable intrinsically disordered region that is functionally conserved from sponge to man

Olubu A. Adiji, Brendan S. McConnell and Matthew W. Parker \*

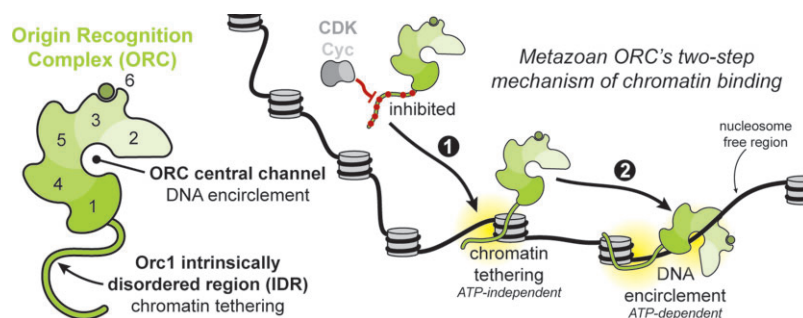
Department of Biophysics, University of Texas Southwestern Medical Center, Dallas, TX 75235, USA

\*To whom correspondence should be addressed. Tel: 1 214 648 1374; Email: matthew.parker@utsouthwestern.edu

## Abstract

The first step toward eukaryotic genome duplication is loading of the replicative helicase onto chromatin. This ‘licensing’ step initiates with the recruitment of the origin recognition complex (ORC) to chromatin, which is thought to occur via ORC’s ATP-dependent DNA binding and encirclement activity. However, we have previously shown that ATP binding is dispensable for the chromatin recruitment of fly ORC, raising the question of how metazoan ORC binds chromosomes. We show here that the intrinsically disordered region (IDR) of fly Orc1 is both necessary and sufficient for recruitment of ORC to chromosomes *in vivo* and demonstrate that this is regulated by IDR phosphorylation. Consistently, we find that the IDR confers the ORC holocomplex with ATP-independent DNA binding activity *in vitro*. Using phylogenetic analysis, we make the surprising observation that metazoan Orc1 IDRs have diverged so markedly that they are unrecognizable as orthologs and yet we find that these compositionally homologous sequences are functionally conserved. Altogether, these data suggest that chromatin is recalcitrant to ORC’s ATP-dependent DNA binding activity, necessitating IDR-dependent chromatin tethering, which we propose poises ORC to opportunistically encircle nucleosome-free regions as they become available.

## Graphical abstract



## Introduction

Eukaryotic DNA replication initiation occurs in two temporally separate steps known as licensing and firing. Replication licensing constitutes loading of the replicative helicase, the heterohexameric Mcm2–7 complex, onto replication start sites, or ‘origins’, during late mitosis and early G1 phase of the cell cycle [reviewed in (1)]. During the synthesis (S) phase of the cell cycle, Mcm2–7 scaffolds replisome assembly and replication initiates, or ‘fires’, leading to DNA strand separation and duplication of cellular genetic material. *In vitro* reconstitution studies have led to a precise mechanistic understanding of the DNA replication licensing reaction. In the first step, the origin recognition complex (ORC, composed of Orc1–6) binds and encircles DNA. This nucleates assembly of the pre-replication complex (pre-RC), a macromolecular machine composed of ORC, Cdc6, Cdt1 and Mcm2–7. The pre-RC then loads Mcm2–7 around duplex DNA as a

double hexamer (2,3), where it is poised to be activated in S phase.

The first step of origin licensing is recruitment of ORC to chromatin. This process has been studied intensely in budding yeast (*Saccharomyces cerevisiae*) and with recombinant fly and human ORCs. A common theme across species is the formation of an ATP-dependent ORC–DNA ternary complex (4–8), which structural studies show manifests as DNA encirclement within the ORC central channel (8–10). In addition to being an essential intermediate in pre-RC assembly, this DNA binding mode is also thought to mediate the initial recruitment of ORC to chromatin. This mechanism of DNA binding is consistent with yeast ORC’s constitutive association with chromosomes (11), but it is less obvious how it relates to the regulated chromatin association observed for metazoan ORC (12,13). This mechanism is contrasted with that of *Schizosaccharomyces pombe* ORC, which utilizes a two-step mecha-

Received: July 3, 2023. Revised: January 26, 2024. Editorial Decision: January 29, 2024. Accepted: February 8, 2024

© The Author(s) 2024. Published by Oxford University Press on behalf of Nucleic Acids Research.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

nism of chromatin binding. In the first step, *Sp*ORC is tethered to chromatin via *ATP-independent* DNA binding facilitated by AT-hook motifs embedded within the *Sp*Orc4 N-terminus (14–17). In the second step, the assembly transitions to a salt-stable complex that, based on conservation of the ORC core complex, likely represents DNA encirclement (18). Why *Sp*ORC requires a two-step mechanism, while *S. cerevisiae* and metazoan ORCs appear to proceed by direct ATP-dependent DNA binding and encirclement is currently unclear.

We previously observed that *Drosophila melanogaster* ORC's ATP-binding (Walker A) and hydrolysis motifs (Walker B)—regions necessary for ATP-dependent DNA binding *in vitro* (5)—are dispensable for the recruitment of ORC to chromosomes *in vivo* (19). Consistently, a large fraction of recombinant fly and human ORCs' *in vitro* DNA binding activity is ATP independent (5,7,20). These data suggest that metazoan ORC may first be tethered to chromatin via an *ATP-independent* mechanism, reminiscent of the two-step binding path of *S. pombe* ORC. Indeed, metazoan ORC contains multiple DNA and chromatin binding elements that could, in theory, facilitate this. For example, we previously identified an intrinsically disordered region (IDR) in metazoan Orc1 that binds and phase separates with DNA *in vitro* and contributes to chromatin recruitment in tissue culture cells (19). N-terminal to the Orc1 IDR is a bromo adjacent homology (BAH) domain that in chordates binds histone H4 dimethylated at lysine 20 (21) and is important for chromatin localization (22). Finally, the TFIIB-like domains of *D. melanogaster* Orc6 bind DNA *in vitro* (23). It is currently unknown which of these chromatin binding elements, if any, facilitate chromatin tethering of metazoan ORC.

Here, we demonstrate that the fly Orc1 IDR is necessary and sufficient for recruitment of ORC to chromosomes *in vivo*. We show that the Orc1 IDR can directly bind both DNA and chromatin and that cyclin-dependent kinases (CDKs) regulate these interactions through IDR phosphorylation. Combined with our previous work, these data strongly suggest that chromatin is recalcitrant to ATP-dependent DNA encirclement by ORC but not IDR-dependent chromatin recruitment. Interestingly, we find that all metazoan Orc1 orthologs possess an IDR and yet share little to no linear sequence similarity in this region. Despite their divergence, we demonstrate functional conservation of metazoan Orc1 IDRs from sponge to man. Collectively, these findings suggest that metazoan ORC implements a two-step chromatin binding mechanism, with IDR-dependent chromatin tethering necessarily preceding and poisoning ORC for ATP-dependent DNA binding, which we propose occurs opportunistically as nucleosome-free regions become available. More broadly, we provide evidence that IDRs with no sequence similarity can nonetheless be functionally conserved, and we develop the concept of compositional homology to explain this.

## Materials and methods

### Generation of transgenic fly lines

The full-length *D. melanogaster* Orc1 (*DmOrc1*) gene was cloned from genomic DNA (OregonR fly line) by polymerase chain reaction (PCR) amplification of the protein coding region in addition to 1 kb of regulatory sequence upstream and downstream of the start and stop codons. This sequence was inserted into vector pattB and Gibson assembly was used to

replace the start codon with the coding sequence for the fluorescent protein mNeonGreen (mNG). This construct was used as a template for around-the-horn mutagenesis to generate an IDR deletion construct ( $\Delta 249$ –541, *DmOrc1* <sup>$\Delta$ IDR</sup>). We generated two transgenes that express the IDR alone, one that was PCR amplified from genomic DNA and spanned two exons (exon 1 and exon 2) and one that was PCR amplified from complementary DNA (cDNA). These constructs encode residues 187–549 of the Orc1 gene and were inserted into pattB with an N-terminal mNeonGreen tag. These constructs (pattB-mNG-g*DmOrc1*, pattB-mNG-g*DmOrc1* <sup>$\Delta$ IDR</sup>, pattB-mNG-g*DmOrc1*<sup>IDR</sup> and pattB-mNG-c*DmOrc1*<sup>IDR</sup>, respectively) were injected into embryos for site-specific PhiC31 integration into chromosome 3 at position attP2:68A4 (Genetivision Corporation). The resulting transgene lines were balanced and crossed to produce homozygous lines. Genomic DNA from homozygous lines was used as a template for PCR amplification of the inserted transgenes, which were then confirmed by sequencing.

### Fluorescent imaging of *D. melanogaster* embryos

Homozygous transgenic fly lines were amplified and combined in bottles (200–500 flies per bottle) that were inverted onto agar plates smeared with yeast paste to induce egg laying. The collected embryos were dechorionated in 100% bleach for 2 min followed by extended washing with H<sub>2</sub>O. The embryos were then transferred into  $\mu$ -slide four-well glass bottom imaging dishes (Ibidi), covered with halocarbon oil and then imaged. Images were acquired on a Nikon Ti2E microscope equipped with a Yokogawa CSU X1 spinning disk using a 60 $\times$  oil immersion objective with the appropriate filter set and images collected with 488 nm laser power (12.7%) and a 200 ms exposure. Z-stacks were collected at 1  $\mu$ m intervals through the full embryo volume. Z-stacks were acquired every 30 s until nuclear cycle 14. Image processing was done with FIJI. For each image set, a maximum intensity projection was generated and the measured mNeonGreen-Orc1 intensity was used to calculate the ratio between Orc1 signal on chromosomes versus the cytosol during mitosis. The doubling of nuclei within the embryo was used to stage cells in mitosis.

### Generation of transgenic *D. melanogaster* S2 cells

To assess Orc1 cellular dynamics in *Drosophila* S2 cells, plasmids were constructed by cloning Orc1 and Orc1 IDRs into our custom pCopiaFP(mNG) vector that appends inserts with an N-terminal mNeonGreen tag. Coding regions were amplified from cDNA and included full-length *Drosophila* Orc1, *Drosophila* Orc1<sup>IDR</sup> (residues 187–549), *Drosophila* Orc1<sup>IDR-P-dead</sup> (residues 187–549 with every '[S/T]P' mutated to 'AP') and metazoan Orc1<sup>IDR</sup> orthologs [human Orc1 residues 177–484, *Fasciola hepatica* (Platyhelminthes) Orc1 residues 1–267, *Hypsibius dujardini* (Tardigrada) Orc1 residues 1–278 and *Amphimedon queenslandica* (Porifera) Orc1 residues 177–593]. All cloning was carried out using ligation-independent cloning (LIC). To visualize chromatin, *DmHistone2A* was fluorescently labeled at its N-terminus with mTurquoise2 by cloning into our custom pCopiaFP(mTurquoise2) vector. Subsequently, individual Orc1 constructs were co-transfected with vectors p8HCO (providing a methotrexate-resistant gene) and pCopiaFP(mTurquoise2)-*DmHis2A* into *Drosophila* S2 cells (Expression Systems) maintained at 27°C in ESF 921 medium (Expression Sys-

tems). For transfection, S2 cells were seeded in six-well plates at a density of  $2 \times 10^6$  cells/well. After 24 h, the medium was removed and supplemented with fresh medium. Transfections were carried out using Effectene reagent following the manufacturer's protocol (Qiagen). Forty-eight hours post-transfection, the medium was replaced with fresh medium supplemented with 0.1  $\mu\text{g}/\text{ml}$  methotrexate. Subsequently, the stably transfected cells were selected by replacing medium with methotrexate-supplemented fresh medium every 3 days for 5 weeks.

### Fluorescent imaging of *D. melanogaster* S2 cells

Stably transfected S2 cells were prepared for imaging by gently transferring 1 ml of culture at  $\approx 1 \times 10^6$  cells/ml to a  $\mu$ -Dish 35 mm imaging dish (Ibidi). The cells were allowed to adhere for 20–30 min prior to imaging by spinning disk confocal fluorescence microscopy (Nikon Ti2E with Yokogawa CSU X1 spinning disk). Images were taken with a  $60\times$  oil immersion objective and 405 and 488 nm lasers were used to excite mTurquoise2 (His2A) and mNeonGreen (Orc1), respectively. Samples were scanned to identify mitotic cells with chromosomes aligned at the metaphase plate. The  $x, y$  coordinates of multiple mitotic cells were marked and the cellular dynamics of Orc1 assessed by time lapse imaging with a  $z$ -stack (8–14  $\mu\text{m}$  thick section at 0.3  $\mu\text{m}$  intervals) taken every 3 min with 300 ms exposure throughout mitosis.

FIJI was used for image processing and quantitation. For each image set, a maximum intensity projection was generated and a median filter was applied (pixel radius = 2) before the image was split into blue and green channels and background intensity subtracted. Regions of interest (ROIs) were generated for chromatin by auto-thresholding on mTurquoise2 signal. The His2A and Orc1 chromosome intensity was measured within these regions for each time point. Similarly, cytosolic mNeonGreen-Orc1 intensity was measured for each time point. The fold change in chromosome intensity was calculated and normalized to the signal at metaphase. To determine the chromosome partitioning of Orc1, mNeonGreen-Orc1 signal on chromosomes was multiplied by the chromosome area and divided by the sum of the same plus mNeonGreen-Orc1 cytosolic intensity times cytosolic area (as shown in Figure 3A). The average and standard deviation of Orc1 chromosome partitioning in metaphase and telophase were calculated for each construct in at least six different mitotic cells.

### Transfection, fractionation and imaging of HeLa cells

The human Orc1<sup>IDR</sup> coding sequence (residues 177–484) was inserted by LIC into vector 6D (QB3, Macrolab) that appends a GFP tag at the coding region's C-terminus. HeLa cells stably expressing mCherry-tagged human histone H2B (24) (provided by Dr Bryan Gibson) were transfected with 6D-HsOrc1<sup>IDR</sup> using the jetPRIME transfection reagent (Polyplus) following the manufacturer's protocol. Briefly, HeLa cells were cultured in Dulbecco's modified Eagle medium supplemented with 5% fetal bovine serum and maintained at 37°C and 5% CO<sub>2</sub>. A six-well plate was seeded with  $2 \times 10^6$  cells/well and incubated for 24 h to achieve  $\leq 60$ –80% confluence. For the transfection, DNA was diluted in jetPRIME buffer, and then mixed with the transfection reagent. The mixture was incubated at room temperature for 10 min before

adding dropwise to cells and then gently mixing. The plate was incubated for 48–72 h before imaging. To image, samples were transferred to a  $\mu$ -Dish 35 mm imaging dish (Ibidi) and imaged by spinning disk confocal fluorescence microscopy. The imaging setup was as described for S2 cells, with the exception that imaging was done with a  $40\times$  air objective and a 568 nm laser was used to excite mCherry. Images were processed as described for S2 cells.

Transfected cells were also subject to cellular fractionation and western blotting to confirm the interaction between HsOrc1<sup>IDR</sup> and chromatin. For this experiment, HeLa cells were transfected with either 6D-HsOrc1<sup>IDR</sup> or the 6D empty vector (as a control). After 48–72 h, the cells were harvested, washed in phosphate-buffered saline and fractionated into cytosolic, nuclear and chromatin-bound fractions according to the manufacturer's protocol (Abcam, cat # ab219177). Fractions were run on 10% sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–PAGE) gels and transferred to polyvinylidene fluoride membranes (Bio-Rad, cat # 1704156) using the Trans-Blot Turbo Transfer system (Bio-Rad). The membranes were blocked with 5% bovine serum albumin for 30 min at room temperature, washed and then incubated with primary antibodies at room temperature for 1 h (1/5000 dilution of anti-GFP, Proteintech, cat # 66002-1-Ig and 1/5000 dilution of anti-histone H3, Proteintech, cat # 68345-1-Ig). Each membrane was then washed and incubated with secondary antibody for 30 min (1/5000 dilution of StarBright Blue 700 Goat Anti-Mouse IgG, cat # 12004159). After washing, blots were visualized on a ChemiDoc MP (Bio-Rad) imaging system.

### Cloning, expression and purification of the ORC holocomplex and CDK/Cyc complexes

The ORC holocomplex (*Dm*ORC), the holocomplex lacking the Orc1 IDR (*Dm*ORC <sup>$\Delta$ IDR</sup>), *Dm*CDK1/CycA and *Dm*CDK2/CycE were expressed and purified from Sf9 cells. The coding region of *Dm*Orc1 was cloned into vector 438B (QB3 Macrolab) for expression with an N-terminal hexahistidine tag. *Dm*Orc2, *Dm*Orc3 and *Dm*Orc5 were cloned into vector 438A (no tag, QB3 Macrolab). Gibson cloning was used to clone *Dm*Orc4 into vector 438B replacing the hexahistidine coding sequence with a TEV-cleavable maltose-binding protein (MBP) tag. Orc1–5 were combined into a single multibac expression vector. *Dm*Orc6 was cloned into pFastBac. Plasmids for production of *Dm*ORC1–6 <sup>$\Delta$ IDR</sup> were the same except that IDR residues 248–549 were deleted from 438B-*Dm*Orc1 and this was not combined in a multibac with Orc2–5. *Dm*CDK1 and CDK2 coding sequences were cloned into our custom 438-MBP vector for expression as an N-terminal TEV-cleavable MBP fusion. Cyclin subunits (CycA or CycE) were cloned into vector 438B. Bacmid DNA was generated from these vectors according to previously established methods (19). Subsequently, bacmid DNAs were transfected into Sf9 cells (Expression Systems) maintained in ESF921 using Cellfectin II following the manufacturer's protocol (Fisher Scientific). P0 virus was harvested from the transfected cells and amplified twice prior to infection of Sf9 cells for protein expression. Cells were co-infected with multiple viruses for production of holocomplexes (*Dm*ORC, *Dm*ORC <sup>$\Delta$ IDR</sup>, *Dm*CDK1/CycA and *Dm*CDK2/CycE) and harvested after 2 days. Cell pellets were isolated by centrifugation and frozen at  $-80^\circ\text{C}$  until protein purification.



*Dm*ORC was purified from 2 l of cultured cells and the cell pellet resuspended in 80 ml of lysis buffer [50 mM Tris, pH 7.5, 300 mM KCl, 50 mM imidazole, 10% glycerol, 200  $\mu$ M phenylmethylsulfonyl fluoride (PMSF), 1 mM  $\beta$ -mercaptoethanol (BME), 1  $\mu$ M benzonase and 1 $\times$  cOmplete EDTA-free Protease Inhibitor Cocktail (Sigma–Aldrich)]. Cells were lysed by sonicating on ice (five cycles of 15 s 100% power followed by 1 min rest) and the lysate centrifuged at 18 000 rpm for 1 h at 4°C. The supernatant was collected and filtered using aPES 0.45  $\mu$ m bottle-top filter unit (Nalgene Rapid-Flow, Thermo Fisher) and then subject to an ammonium sulfate precipitation (975 mM ammonium sulfate) and incubated under gentle rotation for 30 min at 4°C. The lysate was further clarified by centrifugation at 18 000 rpm for 1 h at 4°C. The sample was filtered again prior to nickel purification. The filtrate was loaded onto a 5 ml HisTrap HP column (GE Healthcare), washed with 12 column volumes (CV) of wash buffer (50 mM Tris, pH 7.5, 300 mM KCl, 50 mM imidazole, 10% glycerol, 1 mM BME) and the protein eluted with a 6 CV linear gradient from 0% to 100% elution buffer (50 mM Tris, pH 7.5, 300 mM KCl, 250 mM imidazole, 10% glycerol, 1 mM BME). Amylose purification was used as a second affinity purification step. The sample was loaded onto a column packed with amylose resin and subsequently washed with 3 CV of amylose wash buffer (50 mM Tris, pH 7.5, 300 mM KCl, 10% glycerol, 1 mM BME) and eluted with 2 CV of amylose elution buffer (50 mM Tris, pH 7.5, 300 mM KCl, 10% glycerol, 1 mM BME, 20 mM maltose). The His6 and MBP tags were cleaved by addition of TEV (1/10, w/w) and a 12-h incubation at 4°C. Finally, TEV digested protein was concentrated to <2 ml using an Amicon Ultra-15 concentrator (Millipore) and then purified by size exclusion chromatography. Samples were loaded onto a HiPrep 16/60 S300 HR and run with 1.2 CV of sizing buffer (50 mM HEPES, pH 7.5, 300 mM potassium glutamate, 10% glycerol, 1 mM BME). Peak fractions were assessed by SDS–PAGE and subsequently pooled, concentrated, aliquoted and flash frozen in liquid nitrogen and stored in  $-80^{\circ}\text{C}$ . The same purification process was applied to CDK/Cyc complexes except that the TEV digest and sizing exclusion steps were omitted from the workflow.

### Cloning, expression and purification of Orc1 IDR constructs

*Escherichia coli* BL21(DE3) cells were used to express and purify Orc1 IDRs, including *Dm*Orc1<sup>IDR</sup> (*D. melanogaster* Orc1 residues 187–549), synthesized (Twist Biosciences) metazoan Orc1<sup>IDR</sup>s (*Acropora millepora* Orc1 residues 170–485, *Brachionus plicatilis* Orc1 residues 1–158, *Mizuhopecten yessoensis* Orc1 residues 175–621, *Dimorphilus gyrociatus* Orc1 residues 169–412, *H. dujardini* residues 1–278 and *A. queenslandica* Orc1 residues 177–593) and *Dm*Orc1 IDR variants ( $\Delta 1$  = deletion of residues 187–246,  $\Delta 2$  = deletion of residues 247–306,  $\Delta 3$  = deletion of residues 307–366,  $\Delta 4$  = deletion of residues 367–426,  $\Delta 5$  = deletion of residues 427–486,  $\Delta 6$  = deletion of residues 487–549 and  $\Delta\text{BP}$  = deletion of residues 520–533). Each Orc1 IDR coding sequence was cloned into QB3 Macrolab vector 1C for expression as a TEV-cleavable N-terminal His6–MBP fusion. *Dm*Orc1<sup>IDR</sup> deletion constructs were generated by around-the-horn PCR-based mutagenesis using 1C-*Dm*Orc1<sup>IDR</sup> as a template. Each construct was trans-

formed into BL21(DE3) cells (NEB). Subsequently, cells were grown in liquid culture to an OD<sub>600</sub> = 0.8 and expressed in overnight cultures at 20°C upon 1 mM isopropyl  $\beta$ -D-1-thiogalactopyranoside induction. Cells were harvested by centrifugation and the collected cell pellets were stored at  $-80^{\circ}\text{C}$  until protein purification.

For protein purification, cells from 2 l of culture were resuspended in 80 ml of lysis buffer [20 mM Tris, pH 7.5, 500 mM NaCl, 30 mM imidazole, 10% glycerol, 200  $\mu$ M PMSF, 1 $\times$  cOmplete EDTA-free Protease Inhibitor Cocktail (Sigma–Aldrich), 1 mM BME and 0.1 mg/ml lysozyme] and sonicated on ice before centrifugation at 18 000 rpm for 1 h at 4°C. The sample was applied to an aPES 0.45  $\mu$ m bottle-top filter unit (Nalgene Rapid-Flow, Thermo Fisher) and the filtrate loaded onto a 5 ml HisTrap HP column (GE Healthcare). The column was washed with 10 CV of nickel wash buffer (20 mM Tris, pH 7.5, 500 mM NaCl, 30 mM imidazole, 10% glycerol, 200  $\mu$ M PMSF, 1 mM BME) and the protein eluted with 5 CV of nickel elution buffer (20 mM Tris, pH 7.5, 150 mM NaCl, 500 mM imidazole, 10% glycerol, 1 mM BME). The protein was then applied to a HiTrap Heparin HP column (GE Healthcare), which was then washed with 10 CV of heparin binding buffer (20 mM Tris, pH 7.5, 150 mM NaCl, 10% glycerol, 1 mM BME, 400  $\mu$ M PMSF) and eluted with a 10 CV linear gradient of increasing salt from 150 mM to 1 M NaCl. Elution fractions were assessed by SDS–PAGE and the fractions containing nonproteolyzed protein were pooled and TEV digested overnight at 4°C. The free tag, uncleaved protein and TEV were subsequently removed by an ortho nickel affinity purification step and the sample concentrated to <2 ml using a 10K Amicon Ultra-15 concentrator (Millipore). Finally, the sample was loaded on a HiPrep 16/60 Sephacryl S-300 HR column pre-equilibrated and run in sizing buffer (50 mM HEPES, pH 7.5, 300 mM potassium glutamate, 10% glycerol, 1 mM BME). Peak fractions were assessed by SDS–PAGE before being pooled, concentrated, flash frozen in liquid nitrogen and stored at  $-80^{\circ}\text{C}$ .

### In vitro phosphorylation reactions

In general, the molar ratio of substrate to kinase (CDK/Cyc complex) in each phosphorylation reaction was 5:1. Specifically, 4  $\mu$ M ORC was mixed with 0.8  $\mu$ M of CDK/Cyc and 12  $\mu$ M Orc1<sup>IDR</sup> was mixed with 2.4  $\mu$ M CDK/Cyc in the presence or absence of 1 mM ATP in phosphorylation reaction buffer (50 mM HEPES, pH 7.5, 300 mM potassium glutamate, 10% glycerol, 5 mM MgOAc, 1 mM BME). The reaction mixtures were incubated for 30 min at 25°C. Phosphorylation was assessed by SDS–PAGE with Coomassie staining, where a shift in the band of the phosphorylated protein compared to the unphosphorylated control confirmed the success of the reaction.

We also performed a large-scale preparation of phosphorylated *Dm*Orc1<sup>IDR</sup>. In this case, the size exclusion chromatography peak fractions containing purified *Dm*Orc1<sup>IDR</sup> were pooled, concentrated and then subject to CDK2/CycE-dependent phosphorylation. The phosphorylation reaction mixture contained a 5:1 molar ratio of *Dm*Orc1<sup>IDR</sup> to CDK2/CycE in addition to 4 mM ATP in reaction buffer (50 mM HEPES, pH 7.5, 300 mM potassium glutamate, 10% glycerol, 5 mM MgOAc, 1 mM BME). The mixture was incubated for 30 min at room temperature (25°C). Afterward, the reaction mixture was further supplemented with 2 mM

ATP and CDK2/CycE at a molar ratio of 20:1 and then incubated for an additional 30 min to drive the phosphorylation reaction to completion. The reaction mixture was then subjected to size exclusion chromatography in order to remove CDK2/CycE and ATP. Peak fractions containing phosphorylated *DmOrc1*<sup>IDR</sup> (p*DmOrc1*<sup>IDR</sup>) were assessed by SDS-PAGE and then pooled, concentrated, flash frozen in liquid nitrogen and stored at  $-80^{\circ}\text{C}$ .

### Analysis of DNA binding by electrophoretic mobility shift assays

Serial dilutions of Orc1 IDR variants were prepared in assay buffer (50 mM HEPES, pH 7.5, 150 mM potassium glutamate, 10% glycerol, 1 mM BME) containing 2 nM Cy5-dsDNA (5'-GAAGCTAGACTTAGGTGTCATATTGAACCTACTATGCCGAAGTACGAGCTATAACC-3'). The mixtures were incubated at room temperature for 20 min and then run on a 1% agarose gel at 100 V for 30 min. The gel was imaged on a Bio-Rad ChemiDoc MP using imaging settings appropriate for the excitation and emission spectra of the fluorescent label [cyanine 5 (Cy5) or fluorescein isothiocyanate (FITC)]. The acquired images were analyzed by quantitating the loss of free DNA and calculating the fraction bound for each protein concentration. These data were plotted in GraphPad Prism and fit with a Hill equation to determine the dissociation constant. The data from three independent experiments were used to calculate mean and standard deviation, which are reported.

### Analysis of DNA binding by fluorescence polarization

Fluorescence anisotropy was performed using a FITC-labeled double-stranded DNA (dsDNA) (5'-GAAGCTAGACTTAGGTGTCATATTGAACCTACTATGCCGAAGTACGAGCTATAACC-3'). Reactions invariably contained 2 nM FITC-dsDNA but varying levels of potassium glutamate (150 or 300 mM) and ATP (0 or 1 mM). The remaining buffer components were always the same: 50 mM HEPES (pH 7.5), 10% glycerol, 5 mM MgOAc and 1 mM BME. Concentrations of ORC and Orc1<sup>IDR</sup> are as indicated in figure legends. DNA/protein mixtures were incubated for 30 min at room temperature before being added to a 384-well black bottom multi-well plate and transferred to a CLARIOstar BMG LABTECH plate reader for fluorescence polarization measurement and calculation of fluorescence polarization. All data were background corrected for a buffer blank before calculating polarization. At least two biological replicates, each with technical duplicates, were performed, and the mean and standard deviation of three independent experiments were plotted with respect to each protein concentration. The data were analyzed by GraphPad Prism by fitting a Hill equation to the data to calculate dissociation constants.

Time-resolved fluorescence polarization DNA binding assays were carried out for *DmOrc1*<sup>IDR</sup> at 150 mM potassium glutamate in the presence of CDK2/CycE. The reaction mixture contained 5  $\mu\text{M}$  *DmOrc1*<sup>IDR</sup>, 50 nM Cdk2/CycE, 5 mM MgOAc and 2 nM of FITC-dsDNA. The phosphorylation reaction was initiated by adding 1 mM ATP, gentle mixing, and then immediately and repeatedly taking fluorescence polarization readings. The control reaction was set up the same way but with the exclusion of ATP. Fluorescence polarization mea-

surements were taken every 2.5 min for  $\sim 2$  h and measurements were plotted as a function of time.

### Mass spectrometry analysis of CDK/Cyc-dependent phosphorylation

Intact protein samples (50  $\mu\text{L}$ , 0.3 mg/ml) were analyzed by liquid chromatography/mass spectrometry, using a Sciex X500B QTOF mass spectrometer coupled to an Agilent 1290 Infinity II HPLC. Samples were injected onto a POROS R1 reverse-phase column (2.1 mm  $\times$  30 mm, 20  $\mu\text{m}$  particle size, 4000 Å pore size) and desalted. The mobile phase flow rate was 300  $\mu\text{L}/\text{min}$  and the gradient was as follows: 0–3 min: 0% B; 3–4 min: 0–15% B; 4–16 min: 15–55% B; 16–16.1 min: 55–80% B; 16.1–18 min: 80% B. The column was then re-equilibrated at initial conditions prior to the subsequent injection. Buffer A contained 0.1% formic acid in water and buffer B contained 0.1% formic acid in acetonitrile (ACN). The mass spectrometer was controlled by Sciex OS v.3.0 using the following settings: ion source gas 1, 30 psi; ion source gas 2, 30 psi; curtain gas, 35; CAD gas, 7; temperature,  $300^{\circ}\text{C}$ ; spray voltage, 5500 V; declustering potential, 135 V; collision energy, 10 V. Data were acquired from 400 to 2000 Da with a 0.5 s accumulation time and four time bins summed. The acquired mass spectra for the proteins of interest were deconvoluted using BioPharmaView v. 3.0.1 software (Sciex) in order to obtain the molecular weights. The peak threshold was set to  $\geq 5\%$  and reconstruction processing was set to 20 iterations with a signal-to-noise threshold of  $\geq 20$  and a resolution of 2500.

For identification of phospho sites, samples were digested overnight with trypsin (Pierce) following reduction and alkylation with DTT and iodoacetamide (Sigma-Aldrich). The samples then underwent solid-phase extraction cleanup with an Oasis HLB plate (Waters) and the resulting samples were injected onto a QExactive HF mass spectrometer coupled to an Ultimate 3000 RSLC-Nano liquid chromatography system. Samples were injected onto a 75  $\mu\text{m}$  i.d., 15 cm long EasySpray column (Thermo) and eluted with a gradient from 0% to 28% buffer B over 90 min with a flow rate of 250 nL/min. Buffer A contained 2% (v/v) ACN and 0.1% formic acid in water, and buffer B contained 80% (v/v) ACN, 10% (v/v) trifluoroethanol and 0.1% formic acid in water. The mass spectrometer operated in positive ion mode with a source voltage of 2.5 kV and an ion transfer tube temperature of  $275^{\circ}\text{C}$ . MS scans were acquired at 120 000 resolution in the Orbitrap and up to 20 MS/MS (tandem mass spectrometry) spectra were obtained for each full spectrum acquired using higher-energy collisional dissociation for ions with charges 2–8. Dynamic exclusion was set for 20 s after an ion was selected for fragmentation.

Raw MS data files were analyzed using Proteome Discoverer v2.4 SP1 (Thermo), with peptide identification performed using Sequest HT searching against the *D. melanogaster* protein database from UniProt along with the sequence of Orc1<sup>IDR</sup>. Fragment and precursor tolerances of 10 ppm and 0.02 Da were specified, and three missed cleavages were allowed. Carbamidomethylation of Cys was set as a fixed modification, with oxidation of Met and phosphorylation of Ser, Thr and Tyr set as variable modifications. Phosphorylated sites were localized using the IMP-ptmRS node within Proteome Discoverer. The false discovery rate cutoff was 1% for all peptides.

## Bioinformatic analysis of Orc1 IDR orthologs

Sequence alignments were done with Geneious software (Dot-matics). Pairwise comparisons of the compositional differences between metazoan Orc1 orthologs were performed using a modified version of the chi-square test of homogeneity that quantifies differences in the fractional content of amino acids (25). Instead of using the test statistic to confirm or reject a null hypothesis, the score is normalized so that compositionally identical sequences receive a score of 0, while sequences with no residues in common receive a score of 1. Applied intra-sequentially, the chi-score analysis was used to identify regions of distinct amino acid composition within the *D. melanogaster* Orc1 IDR. This was done by using the pairwise chi-scores between subsequences to predict the positions of module boundaries, which were then optimized and validated statistically. Boundaries were iteratively removed until only those with  $z$ -scores corresponding to a confidence level of 95% or higher remained. The application of the chi-square method to compare sequence composition between proteins and local bias within a contiguous region of disorder is described in detail in a recent publication (25).

## Statistical analyses

The precise number of replicates varied between experiments and we have therefore included these details in the figure legends. Where applicable, a  $t$ -test (two-tailed) was used to compare mean values and significance assigned for  $P < 0.05$ .

## Results

### The Orc1 IDR is necessary and sufficient for regulated chromosome binding

The ATP-binding and hydrolysis motifs of *Drosophila* Orc1—motifs that are indispensable for ATP-dependent DNA binding *in vitro* (5,26)—are dispensable for chromatin recruitment *in vivo* (19). We therefore sought to identify the essential chromatin recruitment element of *Drosophila* ORC. Our previous work found that fly Orc1 contains an IDR that mediates DNA-dependent phase separation *in vitro* and chromatin binding in tissue culture cells (19). This led us to ask whether the Orc1 IDR may be the essential chromatin tethering element *in vivo*. To test this, we generated transgenic fly lines expressing mNeonGreen-tagged full-length *Drosophila* Orc1 (mNG-Orc1), a construct lacking the IDR (mNG-Orc1<sup>ΔIDR</sup>) or the IDR alone (mNG-Orc1<sup>IDR</sup>). These transgenes were cloned from genomic DNA with endogenous promoter and stop sequences (Supplementary Figure S1A–C) and transgene chromatin binding dynamics were assessed by confocal fluorescence microscopy in live embryos 1.5–2 h after fertilization (nuclear cycles 10–13) (Figure 1A–C).

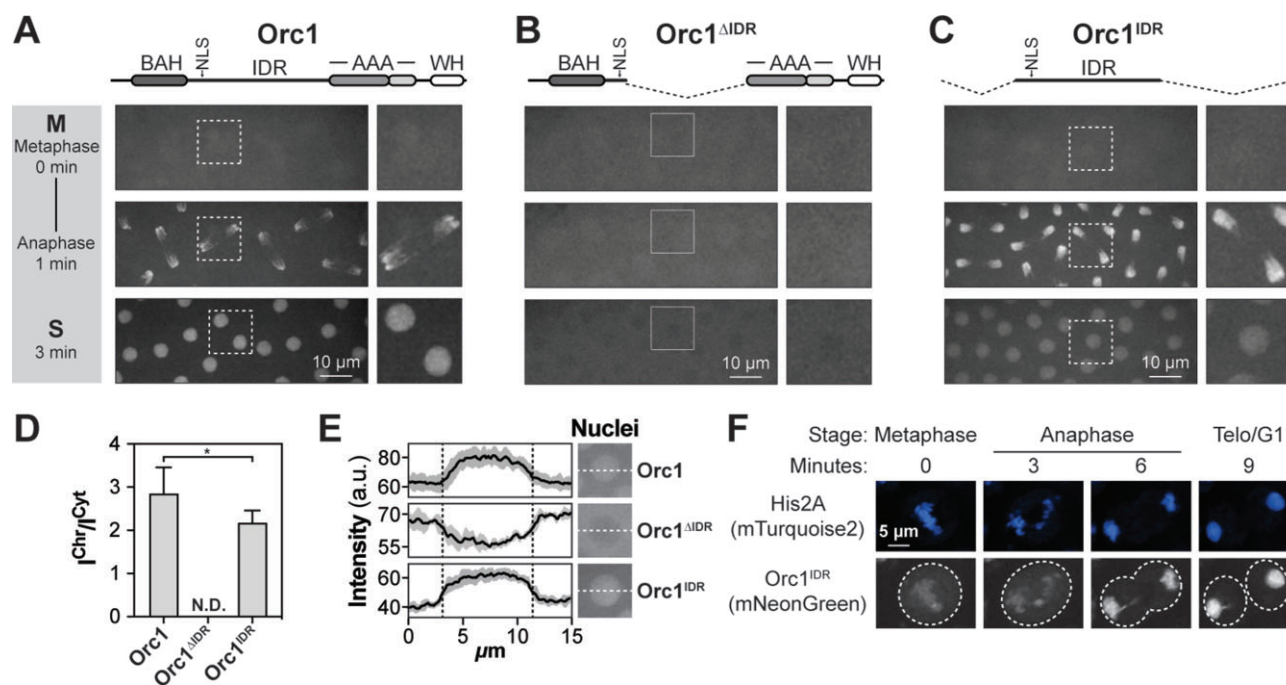
We first assessed the dynamics of the full-length Orc1 transgene and found that mNG-Orc1 is recruited to chromatin in anaphase where it appears to uniformly coat chromosomes (Figure 1A). These dynamics are consistent with the timing of Mcm2–7 loading in the early embryo (27). During anaphase, the chromosome intensity of mNG-Orc1 is 2.9-fold ( $\pm 0.60$ ) above its cytosolic levels (Figure 1D). At this stage of *Drosophila* development, S phase begins immediately after mitosis (i.e. there is no gap phase) and our results show that Orc1 remains enriched in the nucleus throughout this cell cycle phase (Figure 1A, bottom panel, and Figure 1E, top panel). Notably, the transition from S to M phase is evident from the

loss of mNG-Orc1 signal in the nucleus that is the result of nuclear envelope breakdown upon mitotic entry (Figure 1A, compare bottom and top panels) (19). Indeed, the open mitosis of flies enables us to distinguish between mitotic chromatin binding and nuclear enrichment since the nuclear envelope is absent when we see Orc1 binding in anaphase (28). Conversely, we cannot distinguish between nuclear localization and chromatin binding outside of mitosis, although the rapid loss of Orc1 signal concomitant with nuclear envelope breakdown suggests it is not chromatin associated in late S phase. These results establish a baseline understanding of Orc1 cellular dynamics and are consistent with previous reports of ORC dynamics in the early embryo (19,29).

We next assessed the cellular dynamics of mNG-Orc1<sup>ΔIDR</sup>, a construct that lacks most of the predicted disordered region but retains the predicted N-terminal nuclear localization signal. mNG-Orc1<sup>ΔIDR</sup> was fully defective in recruitment to chromatin in anaphase (Figure 1B, middle panel) and, consequently, was depleted from the nucleus upon entry into S phase (Figure 1B, bottom panel, and Figure 1E, middle panel). Interestingly, in late S phase, mNG-Orc1<sup>ΔIDR</sup> briefly transitions to a nuclear enriched state before it again disperses after nuclear envelope breakdown (Supplementary Figure S1D). We are unsure how to interpret this change in nuclear enrichment except to say that mNG-Orc1<sup>ΔIDR</sup> retains nuclear localization capabilities. Together, these data demonstrate that the Orc1 IDR is required for anaphase chromosome recruitment and provides further mechanistic rationale for our previous report that deletion of the Orc1 IDR is embryonic lethal (19).

We next assessed the *in vivo* chromatin binding dynamics of the Orc1 IDR alone (residues 187–549) to determine whether it is sufficient for chromatin recruitment. To test this, we prepared two transgenes: a genomic construct (mNG-Orc1<sup>IDR</sup>) that spans part of two exons and the intervening intron, and a cDNA construct (mNG-Orc1<sup>IDR</sup>) (Supplementary Figure S1C). No difference was observed in the chromatin binding dynamics of these two transgenes (compare Figure 1C and Supplementary Figure S1E) and the data that follow derive from the cDNA construct. Despite lacking the globular domains that we generally associate with Orc1 function (e.g. BAH and AAA+ domains), the bulk cellular dynamics of mNG-Orc1<sup>IDR</sup> was visually indistinguishable from that of the full-length protein (Figure 1C). Indeed, we find that mNG-Orc1<sup>IDR</sup> is homogeneously distributed as the embryo enters mitosis (Figure 1C, top panel); it then rapidly and uniformly binds chromosomes in anaphase (Figure 1C, middle panel) and remains nuclear enriched throughout S phase (Figure 1C, bottom panel, and Figure 1E, bottom panel). The anaphase chromosome intensity of mNG-Orc1<sup>IDR</sup> is 2.2-fold ( $\pm 0.28$ ) above its cytosolic level (Figure 1D), slightly lower than what we observed for the wild-type form of the protein. We also assessed Orc1<sup>IDR</sup> dynamics in *D. melanogaster* tissue culture cells (S2 cells) to determine whether the function of this region is conserved in other developmental stages. We observed chromatin binding dynamics equivalent to what we observed in embryos, with Orc1<sup>IDR</sup> showing loading onto anaphase chromosomes (Figure 1F). The only obvious distinction from embryos was a low level of Orc1<sup>IDR</sup> binding to metaphase chromosomes. Together, these data demonstrate that the Orc1 IDR is both necessary and sufficient for regulated association with mitotic chromosomes and provides a molecular explanation for ORC's ATP-independent chromatin binding capabilities (19).





**Figure 1.** The fly Orc1 IDR is necessary and sufficient for chromatin recruitment *in vivo*. **(A)** Fly embryos expressing an mNeonGreen-tagged full-length Orc1 transgene were imaged live and Orc1 chromatin binding dynamics assessed by confocal fluorescence microscopy during nuclear cycles 10–13. **(B)** Chromatin binding dynamics of an mNeonGreen-tagged Orc1 deletion construct that lacks most of the IDR. **(C)** Chromatin binding dynamics of an mNeonGreen-tagged Orc1 deletion construct that lacks all globular domains but retains the IDR. **(D)** Ratio of mNeonGreen-Orc1 intensity on chromosomes ( $I^{Chr}$ ) versus the cytosol ( $I^{Cyt}$ ). No chromosome enrichment was observed for Orc1<sup>ΔIDR</sup> (N.D., not determined). A *t*-test was used to compare the mean level of Orc1 chromosome enrichment across multiple chromosomes from at least three embryos ( $*P < 0.05$ ). **(E)** Line intensity profile of mNeonGreen-Orc1 signal within and outside the nucleus during S phase. **(F)** An mNeonGreen-tagged Orc1<sup>IDR</sup> was expressed in *D. melanogaster* S2 tissue culture cells where we observed chromatin binding dynamics similar to embryos. Chromatin was visualized by co-expression of mTurquoise2-tagged His2A.

### The Orc1 IDR mediates ATP-independent DNA binding *in vitro*

Recent studies show that ORC's *in vitro* DNA binding activity requires ATP but not the Orc1 IDR (8,30). Paradoxically, we observe that ORC's chromatin recruitment *in vivo* requires the Orc1 IDR but not ATP binding (Figure 1) (19). To understand this discrepancy, we used fluorescence polarization to measure the *in vitro* DNA binding affinity of recombinant ORCs representative of the transgenes we produced for *in vivo* imaging (Figure 1), including the full-length ORC holocomplex (ORC), an ORC holocomplex lacking the Orc1 IDR (ORC<sup>ΔIDR</sup>) and the isolated Orc1 IDR (Orc1<sup>IDR</sup>) (Supplementary Figure S2A). ORC binds DNA in a sequence nonspecific fashion (31) and we therefore used a random 60-bp duplex DNA labeled with fluorescein (FITC-dsDNA) as a substrate in fluorescence polarization DNA binding assays.

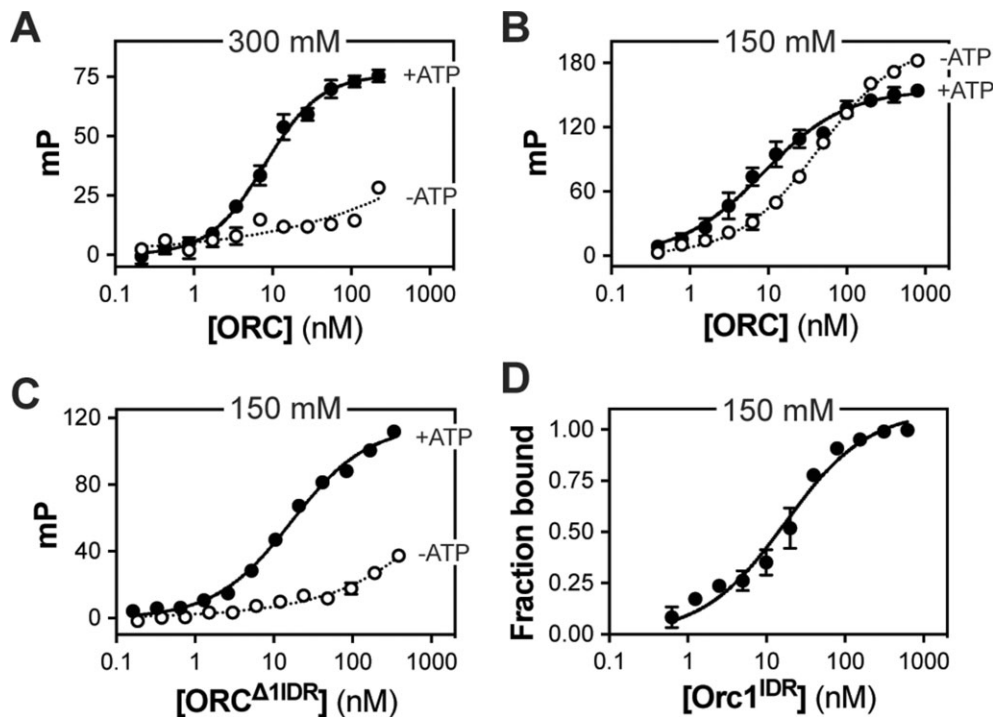
We first assayed DNA binding under conditions similar to those used previously (8) and confirmed that ORC possesses high-affinity DNA binding activity that is strictly dependent on ATP (Figure 2A,  $K_d = 7 \pm 0.3$  nM). We reasoned that these assay conditions, which contain relatively high, nonphysiological concentrations of potassium glutamate ([KGlut] = 300 mM), may selectively impede an IDR-dependent interaction with DNA to impose a dependence on ATP. We therefore repeated the DNA binding assay under the same conditions, except using more physiological levels of salt ([KGlut] = 150 mM) (32,33) (Figure 2B). These conditions unveiled an ATP-independent mechanism of high-affinity DNA binding (Figure 2B, dotted line,  $K_d = 34 \pm 10$  nM). The addition of ATP still

modestly stimulated ORC's affinity for naked DNA (Figure 2B, solid line,  $K_d = 6 \pm 3$  nM).

To test whether the ATP-independent DNA binding we observe *in vitro* (Figure 2B) is mechanistically equivalent to the IDR-dependent chromatin recruitment we observe *in vivo* (Figure 1), we produced a mutant ORC holocomplex that lacks the Orc1 IDR (ORC<sup>ΔIDR</sup>; Supplementary Figure S2A) and assayed DNA binding. ORC<sup>ΔIDR</sup> retained high-affinity ATP-dependent DNA binding activity, although with an ~3-fold reduction in affinity compared to the full-length complex (Figure 2C, solid line,  $K_d = 19 \pm 9$  nM). Conversely, ATP-independent DNA binding by ORC<sup>ΔIDR</sup> was severely impaired (Figure 2C, dotted line,  $K_d > 1$  μM). Finally, we produced the isolated Orc1 IDR (Orc1<sup>IDR</sup>; Supplementary Figure S2A) and assessed its affinity for FITC-dsDNA by EMSA (raw data shown in Supplementary Figure S2B). We quantified the loss of free DNA with increasing concentrations of Orc1<sup>IDR</sup> and found that it binds dsDNA with low nanomolar affinity (Figure 2D, solid line,  $K_d = 15 \pm 4$  nM). These data demonstrate that ORC possesses two mechanistically separable DNA binding modes, one that requires the IDR and the other that requires ATP, and both are required *in vivo* (Figure 1) (19).

### Phosphorylation of the Orc1 disordered region underlies regulated chromatin binding

Anaphase onset is coordinated with a cessation in CDK activity that, through an unknown mechanism, promotes ORC binding to chromosomes (29). This observation, together with



**Figure 2.** The Orc1 IDR facilitates ATP-independent DNA binding *in vitro*. **(A)** A fluorescence polarization assay was used to measure the affinity of the ORC holocomplex for FITC-dsDNA in the presence (black line,  $K_d = 7 \pm 0.3$  nM) and absence of ATP (dotted line,  $K_d =$  N.D.). Reaction conditions were 50 mM HEPES (pH 7.5), 300 mM potassium glutamate, 10% glycerol, 1 mM BME and 5 mM MgOAc ( $\pm 1$  mM ATP). **(B)** Same as panel (A), except salt concentration was reduced to physiological levels (150 mM potassium glutamate). ORC's DNA binding affinity was assessed in the presence (black line,  $K_d = 6 \pm 3$  nM) and absence of ATP (dotted line,  $K_d = 34 \pm 10$  nM). **(C)** The DNA binding affinity of an ORC holocomplex lacking the disordered region of Orc1 (ORC $\Delta$ 1IDR) was assessed in the presence (black line,  $K_d = 19 \pm 9$  nM) and absence of ATP (dotted line,  $K_d =$  N.D.) at 150 mM potassium glutamate. **(D)** Quantitation of an electrophoretic mobility shift assay (EMSA) used to measure binding of the isolated Orc1 IDR to Cy5-dsDNA in the absence of ATP (Orc1<sup>IDR</sup>,  $K_d = 15 \pm 4$  nM). The reported mean and standard deviation are calculated from the results of three independent experiments.

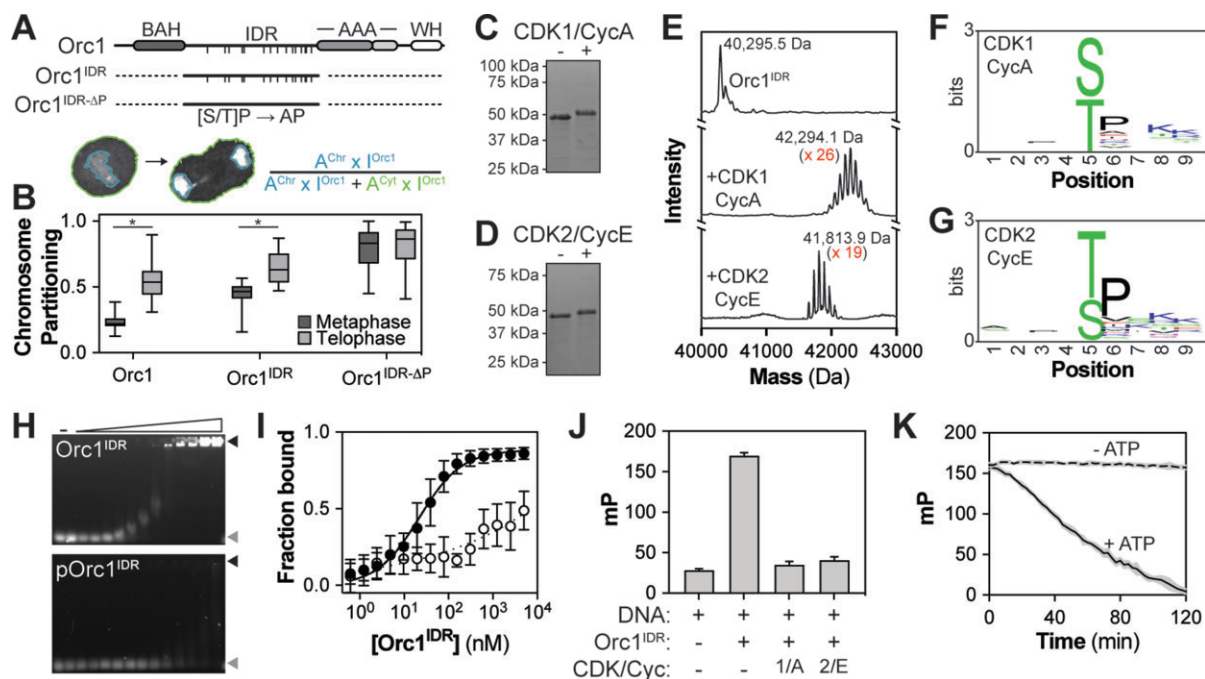
our discovery of the DNA and chromatin binding activity of the Orc1 IDR, led us to hypothesize that phosphorylation of the Orc1 IDR may govern ORC's chromatin binding dynamics. Consistently, the *D. melanogaster* Orc1 IDR possesses an abundance of CDK consensus motifs ('[S/T]P') (19,34). To test this idea, we generated stable *D. melanogaster* S2 tissue culture cell lines that express mTurquoise2-*Dm*Histone2A together with mNeonGreen-tagged wild-type Orc1 (mNG-Orc1), the Orc1 IDR alone (mNG-Orc1<sup>IDR</sup>) or an Orc1 IDR variant where every CDK/Cyc consensus motif has been mutated ([S/T]P $\rightarrow$ AP, Orc1<sup>IDR- $\Delta$ P</sup>) (Figure 3A). We then used confocal fluorescence microscopy to assess the chromatin binding dynamics of each Orc1 transgene throughout mitosis. Each image set was thresholded for either mTurquoise2-*Dm*Histone2A or mNeonGreen-Orc1 to create ROIs encompassing the chromosomes or cytosol, respectively (Figure 3A), and the measured intensity of mNeonGreen within each ROI was used to calculate chromosome partitioning of Orc1 in metaphase versus telophase, the interval over which ORC loads onto chromosomes (Figure 1).

Analysis of mNG-Orc1 chromosome partitioning revealed that, on average, 24% ( $\pm 6\%$ ) of the total Orc1 signal is associated with chromosomes in metaphase and that this more than doubles in telophase ( $56\% \pm 14\%$ , Figure 3B). This is consistent with our *in vivo* results (Figure 1), except that Orc1 showed no binding to metaphase chromosomes in the early embryo. Similarly, mNG-Orc1<sup>IDR</sup> showed a cell cycle-dependent increase in chromosome partitioning, increasing from 44% ( $\pm 10\%$ ) in metaphase to 65% ( $\pm 12\%$ ) in telophase

(Figure 3B). While mNG-Orc1 and mNG-Orc1<sup>IDR</sup> had overall similar dynamics, we note that the IDR alone showed significantly higher levels of chromosome partitioning in both metaphase and telophase, but we do not currently understand why this occurs. Finally, we assessed chromosome partitioning of mNG-Orc1<sup>IDR- $\Delta$ P</sup>, the Orc1 IDR variant that can no longer be phosphorylated. In striking contrast to mNG-Orc1 and mNG-Orc1<sup>IDR</sup>, mNG-Orc1<sup>IDR- $\Delta$ P</sup> was highly enriched even on metaphase chromosomes ( $79\% \pm 15\%$ ) and showed no significant increase as the cells progressed into telophase ( $81\% \pm 15\%$ , Figure 3B). These data demonstrate that in the absence of IDR phosphorylation Orc1 becomes constitutively associated with chromosomes.

We next sought to determine whether phosphorylation directly inhibits the IDR's ability to interact with DNA. We therefore reconstituted IDR phosphorylation *in vitro* with the purified Orc1 IDR (Orc1<sup>IDR</sup>) in combination with CDK1/CycA or CDK2/CycE complexes purified from insect cells. In the presence of ATP, the addition of CDK1/CycA (Figure 3C) and CDK2/CycE (Figure 3D) resulted in a reduction of Orc1<sup>IDR</sup> mobility on SDS-PAGE, strongly suggestive of phosphorylation. Similarly, treatment of the ORC holocomplex with kinase results in a noticeable shift of the Orc1 band as assessed by SDS-PAGE (Supplementary Figure S3A). We confirmed IDR phosphorylation using intact mass spectrometry, which also enabled quantitation of the number of sites phosphorylated by each kinase (Figure 3E). Although the Orc1 IDR has only 15 putative CDK/Cyc consensus motifs, CDK1/CycA treatment added an average of 26 phosphate





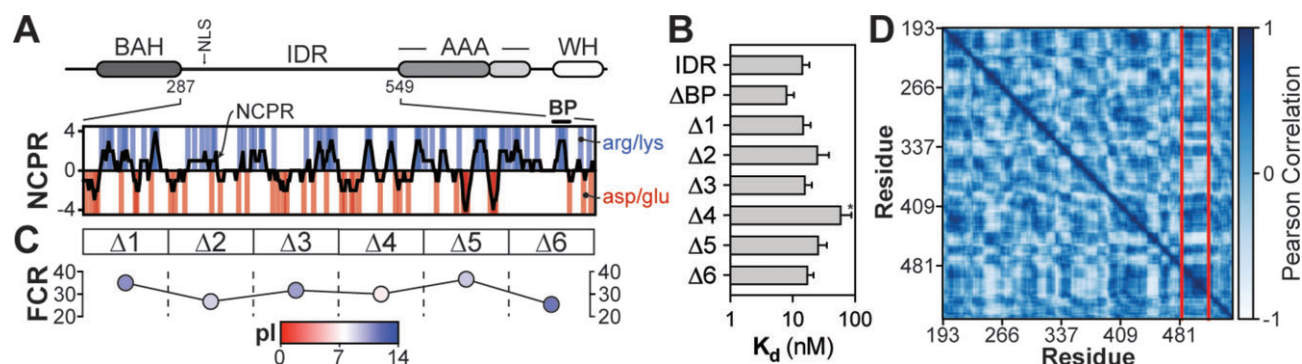
**Figure 3.** Phosphorylation of the Orc1 IDR regulates DNA and chromatin binding. **(A)** Orc1 contains multiple putative CDK/Cyc phosphorylation sites within its IDR (indicated by hash marks). *Drosophila melanogaster* S2 cell lines were generated that stably express full-length Orc1, the IDR alone or an IDR variant that can no longer be phosphorylated. Each construct was tagged at the N-terminus with mNeonGreen and co-expressed with mTurquoise2-His2A, and mitotic chromosome partitioning was measured in live cells. **(B)** The chromosome partitioning of each Orc1 construct was assessed in metaphase and telophase. More than 20 cells were imaged for each construct and a *t*-test was used to calculate significance (\**P* < 0.05). **(C–D)** *In vitro* reconstitution of Orc1<sup>IDR</sup> phosphorylation (12 μM) by **(C)** CDK1/CycA (2.4 μM) and **(D)** CDK2/CycE (2.4 μM). **(E)** Intact mass spectrometry was used to quantitate the level of Orc1<sup>IDR</sup> (top, MW = 40295.8 Da) phosphorylation induced by CDK1/CycA (middle) or CDK2/CycE (bottom). Indicated MW corresponds to highest intensity peak. **(F)** Sequence logo of the CDK1/CycA-dependent phosphorylation sites identified by trypsin digestion of the phosphorylated samples and subsequent liquid chromatography–mass spectrometry. **(G)** Same as panel (F), except for CDK2/CycE phosphorylation of the Orc1<sup>IDR</sup>. **(H)** Unphosphorylated (Orc1<sup>IDR</sup>, top) and CDK2/CycE phosphorylated Orc1<sup>IDR</sup> (pOrc1<sup>IDR</sup>, bottom) was titrated (1 nM to 5 μM) with dsDNA (2 nM) and binding was measured by EMSA. **(I)** Quantitation of panel (H) (Orc1<sup>IDR</sup> *K*<sub>d</sub> = 21 nM and pOrc1<sup>IDR</sup> *K*<sub>d</sub> = N.D.). **(J)** The purified Orc1<sup>IDR</sup> (5 μM) was treated with either CDK1/CycA (1 μM) or CDK2/CycE (1 μM) in the presence of ATP (1 mM) and subsequently added to FITC-dsDNA (2 nM) and DNA binding measured by fluorescence polarization. The data represent the mean and standard deviation of two independent experiments. **(K)** Time-resolved fluorescence polarization was used to assess how progressive phosphorylation of the Orc1<sup>IDR</sup> impacts DNA binding. Orc1<sup>IDR</sup> (5 μM) was premixed with CDK2/CycE (53 nM) and ATP (0 or 1 mM) added at time point = 0, at which point fluorescence polarization readings were begun.

ions (Figure 3E, middle) and CDK2/CycE treatment added an average of 19 phosphate ions (Figure 3E, bottom). The precise phosphorylation sites were mapped by trypsin proteolysis followed by liquid chromatography and mass spectrometry, which revealed significant flexibility in CDK/Cyc motifs (Figure 3F and G). These data are visualized as a sequence logo derived from an alignment of the eight residues immediately surrounding each modified amino acid and, as expected, reveal a preference for a proline residue immediately C-terminal to the phospho site (either Ser or Thr).

We next assessed the impact of phosphorylation on the Orc1 IDR's *in vitro* DNA binding activity. We therefore repeated our purification of Orc1<sup>IDR</sup>, except integrated CDK2/CycE-dependent phosphorylation into our workflow and removed the kinase during the final size exclusion chromatography step. This resulted in a highly pure phosphorylated Orc1<sup>IDR</sup> (pOrc1<sup>IDR</sup>) but with a slightly reduced level of phosphorylation (15 phosphates added; Supplementary Figure S3B and C) compared to our previous assay executed at more concentrated reaction conditions (19 phosphates added; Figure 3D and E). We then used EMSAs to directly compare the DNA binding affinity of Orc1<sup>IDR</sup> and pOrc1<sup>IDR</sup> (Figure 3H and I). Compared to the unphosphorylated sequence, pOrc1<sup>IDR</sup> showed a >100-fold reduction in DNA binding

affinity. This effect was confirmed for CDK1/CycA using fluorescence polarization DNA binding assays (Figure 3J). Here, 5 μM Orc1<sup>IDR</sup> was incubated with CDK1/CycA (1 μM) or CDK2/CycE (1 μM) in the presence of ATP before adding FITC-dsDNA and measuring fluorescence polarization. Notably, pretreatment with either kinase markedly reduced the IDR's DNA binding activity, suggesting that many CDK/Cyc pairs are likely sufficient to regulate Orc1.

Our experiments show that the Orc1 IDR is heavily phosphorylated by CDK/Cyc and raises the question of whether progressive phosphorylation events have a cooperative or linear impact on the IDR's DNA binding activity. To address this, we premixed Orc1<sup>IDR</sup> (5 μM), CDK2/CycE (50 nM) and FITC-dsDNA (2 nM) in the absence of ATP and placed the sample on ice. Subsequently, ATP was added, the sample was mixed, and fluorescence polarization readings were immediately and repeatedly taken for the phosphorylated sample (Figure 3K, solid line) and an unphosphorylated control (no ATP) (Figure 3K, dashed line). Over the course of the experiment (2 h), we observed a linear reduction in fluorescence polarization in the presence of ATP but not in its absence. Assuming the rate of phosphorylation is linear, these data suggest that instead of acting as a binary switch, phosphorylation functions as a rheostat to tune the IDR's affin-



**Figure 4.** The Orc1 disordered region possesses multiple redundant DNA binding motifs. **(A)** The domain architecture of *DmOrc1* and a plot of the IDR's net charge per residue (NCPR) and location of charged amino acids. 'BP' denotes the previously identified basic patch in fly Orc1 (8). **(B)** EMSAs were used to calculate the DNA binding affinity of Orc1 IDR variants. All variants bound DNA with a  $K_d < 100$  nM. The data represent the mean and standard deviation of three independent experiments. A *t*-test was used to compare the mean  $K_d$  of the variants versus wild type ('IDR') ( $*P < 0.05$ ). **(C)** For each deleted region ( $\Delta 1$ – $\Delta 6$ ), the fraction of charged residues (FCR) was calculated and plotted. Each point is colored by the *pI* of the deleted region. **(D)** Chi-score analysis (25) of Orc1 IDR modularity identifies a single compositionally biased region.

ity for DNA. These data may help explain why Orc1 partitions onto metaphase chromosomes in tissue culture cells but not in embryos (Figure 1A and F), two systems that may have a differential balance of kinase and phosphatase activity and therefore differing basal levels of IDR phosphorylation in metaphase. Altogether, these data suggest that regulatory CDK/Cyc-dependent phosphorylation blocks ORC activity by directly inhibiting the ability of the IDR to engage chromatin.

### The Orc1 IDR possesses multiple redundant DNA binding motifs

Our data show that the Orc1 IDR is an essential chromatin tethering element and we next sought to understand its mechanism of DNA binding. To identify functionally relevant regions, we assessed basic chemical features of the 363 amino acid long sequence, including the location of charged residues and the net charge per residue along the sequence (Figure 4A). This approach was motivated by our observation that ORC's ATP-independent DNA binding activity, which is dependent on the IDR, is salt sensitive (Figure 2A–C). In general, the Orc1 IDR contains more basic residues than it does acidic (isoelectric point,  $pI = 10.1$ ) and the charged residues do not appear to be organized in any obvious pattern. The C-terminal residues represent the most basic portion of the sequence (residues 487–549,  $pI = 11.5$ ) and contain a cluster of basic residues (known as the basic patch, BP) that were previously shown to be important for ORC's ATP-dependent DNA binding (8). We reasoned that the basic patch may likewise impart DNA binding activity to the IDR alone and we therefore purified an Orc1 IDR construct with a basic patch deletion ( $\Delta 520$ –533, Orc1<sup>IDR- $\Delta$ BP</sup>) and assayed its affinity for FITC-dsDNA by EMSA (Figure 4B and Supplementary Figure S4A and B). Orc1<sup>IDR- $\Delta$ BP</sup> bound dsDNA tightly ( $K_d = 8 \pm 2$  nM) with no significant difference in affinity compared to the full-length IDR (Figure 2D,  $K_d = 15 \pm 4$  nM).

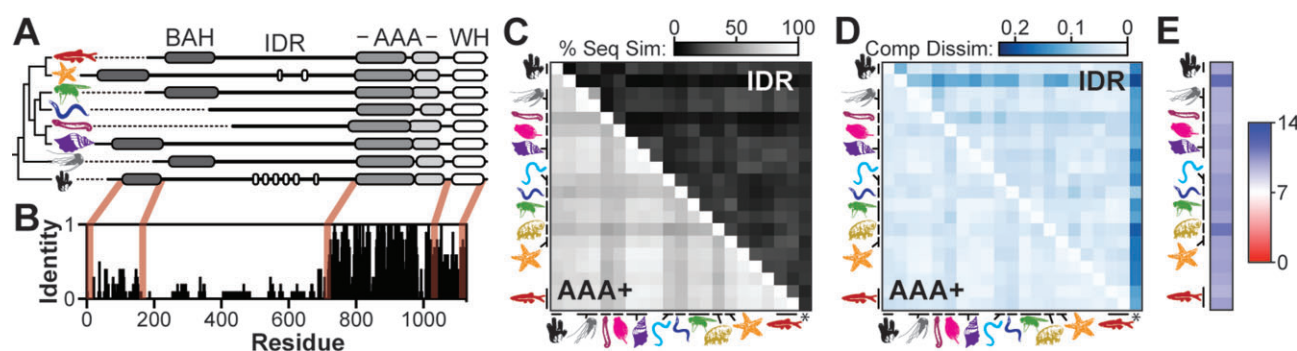
Given the absence of other obvious sequence features targetable by mutagenesis, we took an unbiased, deletion mapping approach to identify DNA binding motifs within the IDR. We purified six IDR variants containing consecutive 60 amino acid deletions spanning the length of the IDR (Supplementary Figure S4A, Orc1<sup>IDR- $\Delta$ 1</sup>–Orc1<sup>IDR- $\Delta$ 6</sup>). Each deleted region has a relatively high fraction of charged

residues (ranging from 0.25 to 0.37) and, except for the fourth region that is slightly acidic, all regions have a basic  $pI$  (ranging from 8.9 to 11.5) (Figure 4C). These analyses provided no obvious candidate DNA binding region(s), and when we measured the DNA binding affinity of each deletion construct, we found that all retained high-affinity DNA binding with a  $K_d < 100$  nM (Figure 4B and Supplementary Figure S4C–H). The only variant with a  $K_d$  significantly different from the full-length IDR was Orc1<sup>IDR- $\Delta$ 4</sup>, which nonetheless still bound DNA with high affinity ( $K_d = 62 \pm 26$  nM).

In a final effort to identify DNA binding regions, we analyzed the Orc1 IDR with a bioinformatic algorithm we have recently developed that uses the chi-square test statistic to quantify local variance in amino acid composition across a sequence. This approach led us to the unexpected discovery that many IDRs contain nonrandom, sequence-spanning compositional patterns that organize the sequence into juxtaposed modules of distinct compositional bias (25). However, unlike many other disordered regions (25), the fly Orc1 IDR is relatively unremarkable with weak modularity and no repetitive module types (Figure 4D). There are only two nonrandom boundaries (at a 95% confidence interval) that delimit a threonine-rich region from a large N-terminal and shorter C-terminal module, which both have high sequence complexity and no obvious sequence patterns. The threonine-rich module is contained entirely in the region deleted in Orc1<sup>IDR- $\Delta$ 6</sup>, which had no effect on DNA binding (Figure 4B). Altogether, these results afforded no further insight into the mechanism of DNA binding and suggest that the fly Orc1 IDR has a relatively uniform sequence landscape with highly redundant DNA binding sequence features scattered throughout its length.

### Metazoan Orc1 IDR orthologs have compositional homology but weak linear sequence similarity

We next assessed natural sequence variation among Orc1 IDR orthologs with the prediction that redundant DNA binding motifs would be conserved and that sequence alignments would divulge their location. We first compiled a list of 20 Orc1 orthologs selected from species representing many of the major metazoan phyla, including Porifera (sponges), Cnidaria (e.g. jellyfish), Arthropoda (insects), Echinodermata (e.g. starfish) and Chordata, among others. The domain organization of each ortholog was predicted using Pfam (35)



**Figure 5.** Metazoan Orc1 IDR orthologs have compositional homology but lack linear sequence similarity. **(A)** The domain architecture of Orc1 is largely conserved across metazoans. Represented phyla include (from top to bottom) Chordata, Echinodermata, Arthropoda, Nematoda, Platyhelminthes, Mollusca, Cnidaria and Porifera. Echinodermata and Porifera Orc1 orthologs possess AT-hook motifs (white regions) embedded within their IDR. **(B)** Sequence identity across 20 Orc1 orthologs. **(C)** Pairwise comparison of linear sequence similarity between 20 Orc1 AAA+ and IDR orthologs. The '\*' column indicates a comparison between each Orc1 IDR and a randomly scrambled version of itself. **(D)** Pairwise comparison of compositional dissimilarity between 20 Orc1 AAA+ and IDR orthologs. Lighter colors indicate sequences with similar sequence composition. The '\*\*' column indicates a comparison between each Orc1 IDR and a 'standardized' sequence (5% of each amino acid). **(E)** All metazoan Orc1 IDRs assessed have a basic pI.

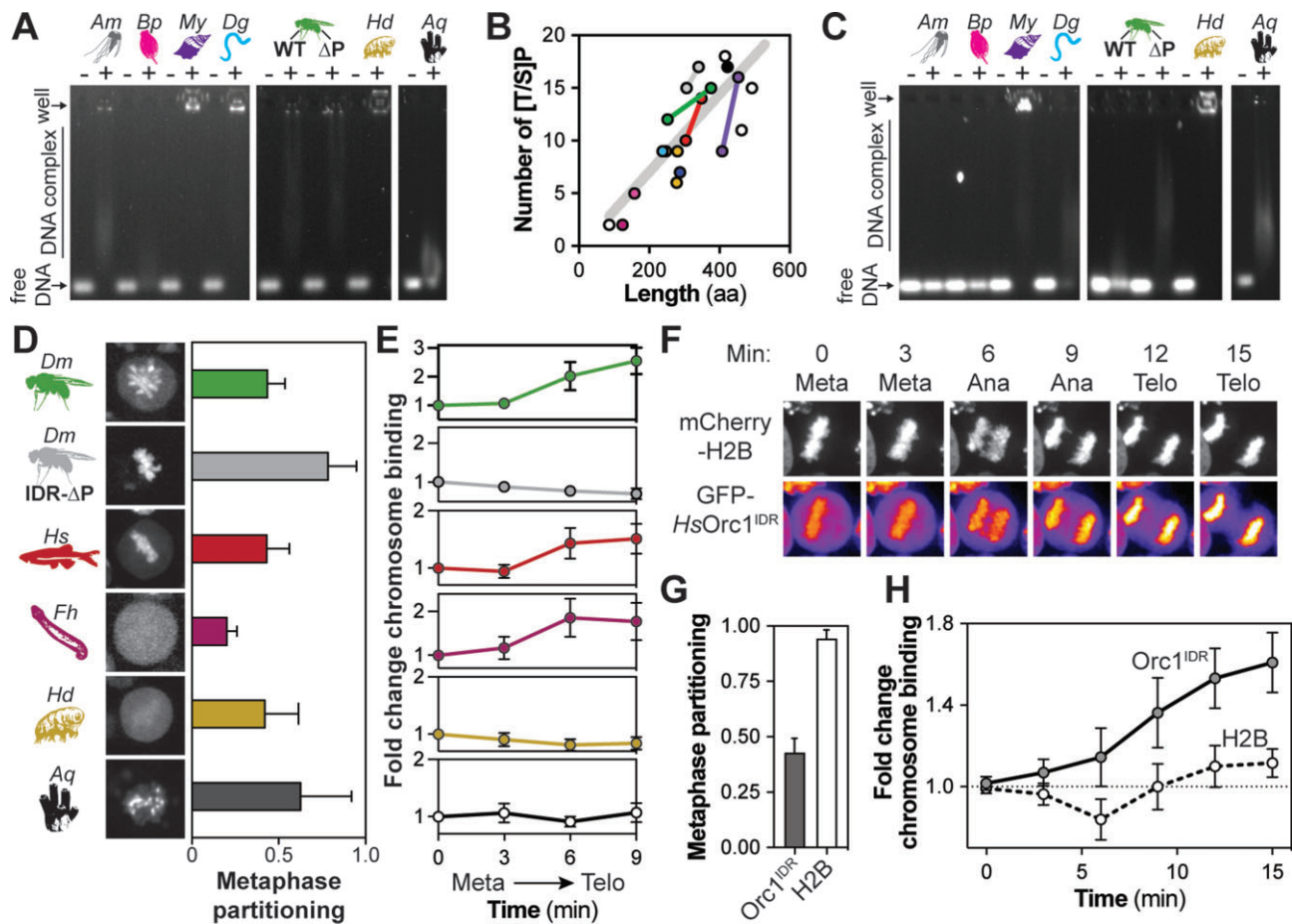
and Metapredict (36) to identify globular domains and regions of intrinsic disorder, respectively. Every Orc1 ortholog we assessed possessed both a AAA+ and a winged helix (WH) domain (Figure 5A), regions that are known to be integral components of the ring-shaped ORC core complex conserved across eukaryotes (37). Each ortholog also possessed an IDR N-terminal to the AAA+ domain. In the sequences sampled, Orc1 IDRs were highly variable in length, with the shortest being only 90 amino acids (Placozoa) and the longest nearly 500 amino acids (Echinodermata). Interestingly, both Echinodermata and Porifera Orc1 orthologs were found to possess AT-hook motifs embedded within their IDR. This DNA binding motif interacts with AT-rich sequences and is the same motif that underlies chromatin binding of *S. pombe* ORC (14,15). Our alignments also revealed that Orc1's BAH domain has been lost in certain phyla, including Tardigrada and Platyhelminthes (flatworms) (Figure 5A).

Using multiple sequence alignment, we measured the sequence identity across the Orc1 gene (Figure 5B). This analysis showed that the AAA+ domain is the most highly conserved region within Orc1, followed by the WH domain and then the BAH domain. Interestingly, our alignments failed to identify regions of sequence similarity within the disordered region. While it is known that IDRs evolve more rapidly than globular sequences (38), we were expecting some level of conservation based on the essential role this region plays in flies. We reasoned that the vast evolutionary timescale represented in our analysis [~800 million years (39)] may mask conserved regions and we thus assessed pairwise sequence similarity for both the AAA+ domain and IDR (Figure 5C). The AAA+ domain (bottom left of diagonal) was highly conserved for each pair of sequences, possessing 54–94% sequence similarity. Conversely, the IDR (top right of diagonal) showed such weak conservation (4–55% sequence similarity) that in many cases it was unclear whether the alignments revealed significant levels of homology or simply spurious registration of short sequences. To answer this, we generated alignments of each Orc1 IDR ortholog with a randomly scrambled version of itself (column marked with an asterisk) and found that sequence similarity ranged from 10% to 19%. Of the 190 pairs of Orc1 IDR orthologs, 82 have a score similar to or below that observed for the randomly scrambled sequences suggest-

ing that in many cases the Orc1 IDR is so highly diverged that no sequence similarity remains.

The sequence divergence among Orc1 IDRs failed to advance our search for DNA binding motifs and further suggested that the function of the Orc1 IDR may not be conserved. Another intriguing possibility, however, is that the function of the Orc1 IDR does not depend on the linear ordering of amino acids, and instead relies only on sequence composition. We therefore assessed whether Orc1 IDRs are *compositionally homologous* using a newly developed metric that quantitatively compares the fractional composition of amino acids between sequences (25). The results were scaled such that sequences with the same composition have a score of 0 and sequences that have no shared amino acids have a score of 1. We calculated compositional homology between orthologous Orc1 AAA+ domains (Figure 5D, bottom left) and orthologous Orc1 IDRs (Figure 5D, top right). As a comparison, we calculated compositional homology between each IDR and a standardized sequence composed of 5% of each of the 20 amino acids (column marked with an asterisk). From this analysis, it was clear that Orc1 IDR orthologs have a sequence composition much more similar to one another than to the standardized sequence, and that their level of compositional homology is, in many cases, on par with what is observed between AAA+ domains. Consistent with their shared amino acid composition, all metazoan Orc1 IDRs have a basic pI (Figure 5E). We also assessed compositional homology between Orc1 IDR orthologs and other functionally related (i.e. IDRs from other licensing factors, including human Cdt1, Cdc6 and Orc2) and functionally unrelated classes of IDRs. This latter class includes disordered regions from proteins involved in transcription (human Rpb1, Med1 and Spt6), the nuclear pore complex (human Nup98) and RNA processing (human FUS) (Supplementary Figure S5). These data emphasize the shared amino acid composition of Orc1 IDR orthologs and reveal a higher degree of compositional homology among licensing factor IDRs (i.e. Orc1, Cdt1 and Cdc6) than between licensing factor IDRs and the disordered regions from other functional classes. Altogether, these data suggest that the function of Orc1 IDRs could be conserved despite an absence of traditional forms of sequence homology.





**Figure 6.** The Orc1 IDR is functionally conserved across metazoans. **(A)** EMSAs were used to assess binding of Orc1 IDR orthologs (600 nM) to Cy5-dsDNA (2 nM). Orthologs from the following phyla were used: Cnidaria (Am), Rotifera (Bp), Mollusca (My), Annelida (Dg), Arthropoda (Dm), Tardigrada (Hd) and Porifera (Aq). 'ΔP' is the *Dm*Orc1 IDR with mutated phosphorylation sites (all '[S/T]P' → 'AP'). **(B)** Graph showing the number of phosphorylation sites ('[T/S]P') versus the length of 20 Orc1 IDR orthologs. Each point represents an IDR and is colored by phyla. IDRs from the same phyla are connected by a line. The gray line indicates the line of best fit. **(C)** Each Orc1 IDR ortholog (600 nM) was phosphorylated by CDK2/CycE and DNA binding assessed by EMSA. **(D)** Metaphase images of *Drosophila* S2 cells expressing Orc1 IDR orthologs tagged with mNeonGreen. Orthologs from the following phyla were used: Arthropoda (*Dm*), Chordata (*Hs*, human), Platyhelminthes (*Fh*), Tardigrada (*Hd*) and Porifera (*Aq*). **(E)** Live imaging experiments in *Drosophila* S2 cells were used to calculate the fold change in mitotic chromosome recruitment for each Orc1 IDR ortholog. **(F)** HeLa cells (stably expressing mCherry-H2B) were transiently transfected with GFP-tagged human Orc1<sup>IDR</sup> (*HsOrc1*<sup>IDR</sup>) and live imaging used to assess *HsOrc1*<sup>IDR</sup> recruitment to mitotic chromosomes. **(G)** Metaphase partitioning of *HsOrc1*<sup>IDR</sup> and H2B. **(H)** Live imaging experiments were used to calculate the fold change in chromosome recruitment of *HsOrc1*<sup>IDR</sup> and H2B throughout HeLa cell mitosis. The mean and standard deviation are reported for seven independent mitotic events.

### The function of the Orc1 IDR is conserved across metazoans

The key functional features of the *D. melanogaster* Orc1 IDR are its ability to bind DNA and chromatin (Figures 1–3) and we therefore set out to test whether Orc1 IDR orthologs share these functionalities. We synthesized the Orc1 IDR coding region from multiple metazoan organisms representing a diverse set of phyla and we were able to successfully express and purify the Orc1 IDRs from *A. millepora* (Cnidaria), *B. plicatilis* (Rotifera), *M. yessoensis* (Mollusca), *D. gyrocoliatilis* (Annelida), *H. dujardini* (Tardigrada) and *A. queenslandica* (Porifera). These IDRs range in size from 158 to 454 amino acids and all purified as a single band on SDS-PAGE, except for *A. queenslandica* that was a mixture of full-length and proteolyzed products (Supplementary Figure S6A). To assay DNA binding, we combined each IDR (600 nM) with FITC-dsDNA (2 nM) and assessed binding by EMSA (Figure 6A). For comparison, we also assayed DNA binding with

the fly Orc1 IDR (*DmOrc1*<sup>IDR</sup>) and the mutated variant that lacks phosphorylation sites (*DmOrc1*<sup>IDR-ΔP</sup>). At this concentration, each of the Orc1 IDRs caused a near-complete loss of free FITC-dsDNA and we observed a heterogeneous mixture of either well-shifted or smeared IDR-DNA complexes, indicative of variable stoichiometry or possibly a dynamic complex. We also assessed DNA binding by fluorescence polarization and found that with the exception of *BpOrc1*<sup>IDR</sup> and *HdOrc1*<sup>IDR</sup> all orthologs possess high affinity for dsDNA ( $K_d \approx 100$  nM or less; Supplementary Figure S6B–I). Despite possessing AT-hook motifs, which are *bona fide* DNA binding elements, *AqOrc1*<sup>IDR</sup> had relatively weak DNA binding affinity by EMSA. We suspect this is due to incompatibility of our probe DNA (which has 50% AT content) and the specificity of AT-hook motifs for AT-rich regions of DNA.

We next asked whether the DNA binding activity of metazoan Orc1 IDRs is regulated by phosphorylation. First, we assessed whether CDK/Cyc sites ('[S/T]P') are present in meta-

zoan Orc1 orthologs and found that the number of sites is linearly correlated with the length of the IDR (Figure 6B). Next, we combined purified *D. melanogaster* CDK2/CycE (0.5  $\mu$ M) with each IDR (2.5  $\mu$ M) and, after a 30-min incubation, we incubated each IDR (600 nM) with FITC-dsDNA (2 nM) and assessed binding by EMSA (Figure 6C). As expected, treatment with CDK2/CycE inhibited the DNA binding activity of the *DmOrc1*<sup>IDR</sup>—as evidenced by the presence of a free DNA band—but not the variant with mutated phosphorylation sites (*DmOrc1*<sup>IDR- $\Delta$ P</sup>). Interestingly, we found that about half of the purified Orc1 IDR orthologs are regulated by CDK2/CycE-dependent phosphorylation (including *Am*, *Bp* and *Dm*), and the other half are clearly not regulated (including *My*, *Hd* and *Aq*) or are only partially (*Dg*). One possible explanation for the observed differences is that the *Drosophila* CDK2/CycE complex may be effective against only a subset of our sequences. Alternatively, some sequences may legitimately lack regulation, such as *HdOrc1*<sup>IDR</sup>, which has multiple minimal CDK/Cyc recognition motifs ('[S/T]P') but only a single optimal site ('[S/T]PX[R/K]'). Likewise, we suspect that the more sequence-specific DNA binding facilitated by the AT-hook motifs of *AqOrc1*<sup>IDR</sup> may be insensitive to phosphorylation. Collectively, these data indicate that despite an absence of sequence similarity, metazoan Orc1 IDRs universally possess the ability to bind DNA and suggest that phosphoregulation is broadly but not universally conserved.

The conservation of DNA binding activity prompted us to assess whether Orc1 IDR orthologs can also facilitate chromatin tethering within the cell. Testing this in each organism individually would be impractical and we therefore assessed functionality in a heterologous system. Specifically, we generated stable *D. melanogaster* cell lines (S2 cells) that co-express mNeonGreen-tagged Orc1 IDR orthologs alongside mTurquoise2-tagged Histone2A (for chromosome visualization). We were successful in generating stable cell lines for some Orc1 IDRs but not others, resulting in datasets on six orthologs [*HsOrc1*<sup>IDR</sup> (Chordata), *FhOrc1*<sup>IDR</sup> (Platyhelminthes), *HdDmOrc1*<sup>IDR</sup> (Tardigrada), *AqOrc1*<sup>IDR</sup> (Porifera), *DmOrc1*<sup>IDR</sup> (Arthropoda) and *DmOrc1*<sup>IDR- $\Delta$ P</sup>], four of which are represented in our *in vitro* DNA binding assays (Figure 6A–C). Each cell line was individually imaged by time-course spinning disk confocal fluorescence microscopy. We began imaging in metaphase (Figure 6D) and continued through early G1 to assess how Orc1<sup>IDR</sup> chromatin binding changes throughout mitosis (Figure 6E). The basal level of chromosome association observed in metaphase cells served as the reference for calculating the fold change in chromosome binding through anaphase and telophase. The *DmOrc1*<sup>IDR</sup> (regulated chromatin binding) and *DmOrc1*<sup>IDR- $\Delta$ P</sup> (constitutive chromatin binding) data are derived from Figure 3A and B and are presented for comparison with other orthologs. As observed for the fly Orc1 IDR (metaphase partitioning = 44%), *HsOrc1*<sup>IDR</sup> (Chordata) was found to be moderately enriched on metaphase chromosomes (Figure 6D, red, metaphase partitioning = 44%) and showed enhanced chromatin binding as cells progressed through mitosis (Figure 6E, red). *FhOrc1*<sup>IDR</sup> (Platyhelminthes) did not appear visibly enriched on nor excluded from metaphase chromosomes (Figure 6D, purple, partitioning = 21%), but progression into the later stages of mitosis resulted in significant chromosome recruitment (Figure 6E, purple). There was relatively high cell-to-cell variability in the chromatin binding dynamics of *HdOrc1*<sup>IDR</sup> (Tardigrada, Fig-

ure 6D, gold, partitioning = 43%) and *AqOrc1*<sup>IDR</sup> (Porifera, Figure 6D, black, partitioning = 63%), but on average they were moderately enriched on chromatin in metaphase, and in both cases we observed no change in chromosome partitioning as the cells progressed through mitosis (Figure 6E, gold and black). Notably, *AqOrc1*<sup>IDR</sup> possesses a punctate distribution on chromosomes that we speculate is pericentromeric heterochromatin on the basis that the chromosomal protein D1, the factor that normally occupies these chromosomal loci, also has multiple AT-hook motifs. The lack of regulation we observed for *HdOrc1*<sup>IDR</sup> and *AqOrc1*<sup>IDR</sup> is consistent with our *in vitro* data where we observed DNA binding regardless of CDK2/CycE phosphorylation (Figure 6A–C). Consistent with our *in vitro* DNA binding studies, these data suggest that chromatin binding is a universally conserved function of metazoan Orc1 IDRs, and that some orthologous sequences likely mediate a constitutive association with chromatin (e.g. *Hd* and *Aq* Orc1 IDRs).

As a final test of the functional conservation of Orc1 IDRs, we assayed the chromatin binding dynamics of the human Orc1 IDR (*HsOrc1*<sup>IDR</sup>) in human tissue culture cells. We constructed a transgene that expresses the human Orc1 IDR coding region fused at its C-terminus to GFP and transiently transfected the construct into HeLa cells stably expressing mCherry-tagged histone H2B (24). Cells were plated in glass bottom dishes that were scanned by confocal fluorescence microscopy for mitotic cells. The position of each mitotic cell was marked and subsequently imaged with an automated imaging routine that collected a z-stack every 3 min through mitosis (Figure 6F, representative image set). As seen in *D. melanogaster* S2 cells, *HsOrc1*<sup>IDR</sup> was moderately enriched on metaphase chromosomes (Figure 6F and G, metaphase partitioning = 44%) and, as cells progressed through mitosis, chromosome partitioning of the human Orc1 IDR increased to 1.6 times the level seen in metaphase (Figure 6H). The chromosome binding activity of *HsOrc1*<sup>IDR</sup> was also validated by cell fractionation and western blotting, where it was found to be strongly enriched in the chromatin-bound fraction (Supplementary Figure S7). We also quantitated mCherry-H2B dynamics through mitosis and, as expected, protein intensity remained essentially unchanged, being fully partitioned onto chromatin regardless of mitotic stage (Figure 6F–H). These data demonstrate that the Orc1 IDR is a DNA and chromatin binding element and that it is functionally conserved across the metazoan lineage.

## Discussion

We report here that the fly Orc1 IDR is necessary and sufficient for *in vivo* chromosome recruitment of ORC and that CDK/Cyc-dependent multisite phosphorylation inhibits chromatin binding. These findings provide the molecular logic behind metazoan ORC's regulated association with chromatin. Together with our previous work (19), these data suggest that chromatin is recalcitrant to ORC's ATP-dependent DNA binding and encirclement activity, and we suggest that this necessitates chromatin tethering by the Orc1 IDR. Interestingly, we find that metazoan Orc1 IDR orthologs share little to no sequence similarity, but nonetheless have a conserved functionality. We propose that a similar amino acid composition is sufficient to maintain the function of Orc1 IDR orthologs. These findings revise our understanding of the mechanism of metazoan DNA replication licensing and, more broadly, expand the

relationship between a protein's primary structure and function with the concept of compositional homology.

Over 30 years ago, *S. cerevisiae* ORC was identified on the basis of its ATP-dependent association with yeast origins of replication (4). Subsequent studies found that human and fly ORC also possess ATP-dependent DNA binding activity, with the caveat that nonspecific, nucleotide-independent interactions with DNA do occur (5,7,20). These data led to the notion that metazoan ORC, like ScORC, is recruited to chromatin via ATP-dependent DNA binding. Recent studies support this, showing that ATP is strictly required for fly ORC's *in vitro* DNA binding activity (8,30). We have now shown that ATP binding is dispensable for chromatin recruitment of ORC *in vivo* (19) and that the metazoan Orc1 IDR is the essential chromatin tethering element (Figure 1). These findings present an apparent paradox with the reported ATP dependence of ORC's *in vitro* DNA binding activity (8). Our study rationalizes this discrepancy with *in vitro* experiments showing that ATP is only required for ORC's DNA binding activity when salt concentrations are relatively high ([KGlut] = 300 mM). Indeed, at physiological salt concentrations (150 mM), we observe high-affinity, ATP-independent DNA binding by ORC, and this requires the Orc1 IDR (Figure 2).

Our results suggest that the mechanism of metazoan ORC chromatin binding is more akin to *S. pombe* than it is to *S. cerevisiae*. The metazoan Orc1 IDR seems to be the functional analog of the AT-hook-containing SpOrc4 N-terminus in the limited sense that both are required for chromatin tethering (14,15). This raises the question of why chromatin tethering is needed at all, since ORC's ATP-dependent DNA binding and encirclement activity is clearly conserved across eukaryotes and is sufficient for high-affinity DNA binding *in vitro* (Figure 2) (30). We propose that in metazoans chromatin is inherently restrictive to ATP-dependent DNA binding, and that this necessitates a chromatin tethering mechanism. This idea is most clearly supported by the observation that ORC<sup>ΔIDR</sup> still possesses high-affinity ATP-dependent binding to naked DNA *in vitro* (Figure 2) and yet deletion of the Orc1 IDR *in vivo* is lethal (19) and the protein can no longer bind chromatin (Figure 1). In hindsight, this conclusion appears self-evident, for we know that ORC, which requires at least 40 bp of naked DNA for ATP-dependent binding (8), is ineffective at displacing nucleosomes (40,41). Further, internucleosomal linker DNA is relatively short [25–50 bp on average (42,43)] and occluded by histone H1 (44). Thus, the lack of naked DNA *in vivo* precludes ATP-dependent DNA binding and encirclement by ORC. This problem is further emphasized when one considers the nearly 100 bp needed for full pre-RC assembly (45). It seems that this problem is evaded altogether by ScORC as it is constitutively associated with chromatin and binds specific DNA sequences that are maintained in a nucleosome-free state (46). For these reasons, we think an accessory chromatin tethering appendage is required for *S. pombe* and metazoan ORCs (and likely most other eukaryotes) but not the *S. cerevisiae* complex.

Of course, ATP-dependent DNA binding and encirclement must and does occur within the context of the chromosome, and we think that chromatin tethering promotes this. We propose that ORC is first nonspecifically and dynamically tethered to chromatin via the Orc1 IDR and that this state, though not a structurally resolvable pre-RC intermediate, positions ORC to opportunistically bind and encircle nucleosome-free regions in an ATP-dependent fashion. Thus, metazoan ORC

does not select origins *per se*, but rather exploits the chromatin remodeling capabilities of other chromatin contextualized processes with which it has no direct connection. In line with this, genomic studies demonstrate that ORC is enriched at transcription start sites (TSSs) and other regions of high nucleosome turnover (47–49). This mechanism would endow the licensing machinery with an inherent flexibility to adapt to the cell type-specific transcriptional programs of multicellular organisms. Alternatively, ORC may be stabilized at select TSSs through specific interactions *in trans* (50). For example, it has recently become clear that disordered regions are essential for targeting (51) and assembling (52,53) the transcriptional machinery at specific chromosomal loci and the Orc1 IDR may integrate into these inter-IDR interaction networks. We suspect that the DNA binding IDRs of metazoan Cdt1 and Cdc6 (19) play a similar role by maintaining relatively high concentrations of these factors on chromatin to enable the efficient assembly and recycling of pre-RC components, and this may be further supported by the ability of these proteins to phase separate with DNA (19,54).

While this work clearly demonstrates that the Orc1 IDR tethers ORC to chromatin, precisely how it does this remains to be determined. Our results show that the Orc1 IDR's DNA binding motifs are highly redundant (Figure 4) and that Orc1 IDR orthologs retain DNA and chromatin binding capabilities (Figure 6) but, oddly, have little to no sequence similarity (Figure 5). Orc1 IDR orthologs do, however, possess a similar amino acid composition and all have a basic pI (Figure 5). We propose the term compositional homology to describe such sequence sets and our newly developed chi-score metric provides a quantitative means to classify these (25). Simply put, these results suggest that the precise ordering of amino acids in the Orc1 IDR is not important for chromatin binding and that these sequences likely interact nonspecifically with the DNA backbone (19). This is consistent with the promiscuous *in vitro* DNA binding activity of metazoan ORC (7,31) and the absence of sequence-specific origins of replication in metazoans (48,47,55). Similarly, the Cdt1 IDR, which is compositionally homologous to the Orc1 IDR (19), retains the ability to bind DNA even when randomly scrambled and binds DNA nonspecifically *in vitro* (56). The sponge Orc1 IDR is somewhat unique in that it contains seven AT-hook motifs (e.g. AqOrc1<sup>473–482</sup>: 'RKRGRPRKEE'), the very same DNA binding element found in the SpOrc4 N-terminus (15,16). Interestingly, degenerate AT-hook motifs are also present in echinoderm (501-KQGRPKK-507 and 564-RKRGRPRSVKK) and fission yeast (226-RGRGRPRK-233) Orc1 IDRs, suggesting the intriguing possibility that an AT-hook containing IDR was the ancient ancestral sequence from which Orc1 IDRs have diverged and have maintained DNA binding capabilities but with reduced specificity.

The Orc1 IDR's nonspecific, electrostatic-based interactions with DNA raise the question of how specific binding to DNA/chromatin is achieved. We have shown in a previous study that metazoan licensing factors undergo DNA-dependent phase separation, that their IDRs are required for this (19) and, at least for Cdt1, that other polyanions (RNA and polyglutamate) function equally well as duplex DNA in stimulating phase separation (56). We therefore cannot rule out the possibility that, in addition to DNA, the Orc1 IDR is targeted to chromatin through interactions with chromatin-associated RNAs (57), proteins that present an acidic surface (e.g. the H2A/H2B acidic patch), or through direct in-



teraction with cognate IDRs on the transcriptional machinery (as was discussed above). In fact, we favor the idea that the IDR can interact with many different chromatin features that may provide context-independent chromatin association to support genome-wide licensing. In support of this, it has been shown that the Orc1 IDR interacts with chromatin-localized RNA and that this is important for both normal cellular replication and licensing of viral genomes (58–60). How then is the IDR targeted specifically to chromatin and not to other RNA-enriched structures? This question remains to be answered, but one interesting possibility is that IDR dephosphorylation is spatially restricted to the chromosome surface through chromatin-associated phosphatases, such as the RepoMan–protein phosphatase 1 complex that, like ORC, binds chromatin at anaphase onset (61). However, CDK/Cyc activity will rise in G1 phase and licensing nonetheless continues (62). We thus speculate that either local dephosphorylation of the Orc1 IDR retains ORC at specific chromosomal loci [e.g. PP2A (63)] or the phosphorylated IDR acquires new interactions with chromatin in trans. Additionally, other chromatin binding domains of ORC, such as the Orc1 BAH domain (22,21), may play a more prominent role in origin selection while the IDR remains phosphorylated.

This study, in addition to our previous work with Cdt1 (56), adds the metazoan licensing factor IDRs to a short list of intrinsically disordered sequences whose function is known to rely solely on sequence composition and not on the linear ordering of amino acids. Other proteins in this category include linker histone H1, which retains the ability to compact DNA even when the sequence is randomly scrambled (64), and the prion domains of yeast Sup35p and Ure2p, which can be scrambled without losing the ability to induce amyloid formation (65,66). Collectively, this limited set of experiments suggests that the function of certain IDR classes depends only on the combined attributes of length and fractional content of amino acids, and that functionally homologous IDRs could be identified by compositional homology alone. An interesting future direction is to extend these observations to other sets of IDR orthologs to compare the extent of compositional homology versus linear sequence similarity, which would likely provide important mechanistic insight into how function is encoded in disordered sequences.

In conclusion, this work demonstrates that metazoan ORC engages chromatin in a two-step process, with IDR-dependent chromatin tethering preceding ATP-dependent DNA binding and encirclement. We propose that chromatin tethering is likely a general solution to overcome the restrictive nature of chromatin, and evolution appears to have implemented this in various ways, including with the metazoan Orc1 IDR, the AT-hook motifs of *S. pombe* Orc4 and possibly hitherto unidentified mechanisms, such as the predicted Zn finger in *Ara-bidopsis* Orc1. Our data suggest that the function of metazoan Orc1 IDRs does not depend on the linear ordering of amino acids, and we provide evidence that amino acid composition alone is their defining feature. We anticipate that compositional homology will be an important concept for understanding the functional conservation of other classes of disordered domains that show limited or no linear sequence similarity.

## Data availability

All data underlying this article are available in the article, in its online supplementary material and in

the article's associated Data Dryad online repository (<https://datadryad.org/stash/share/zx-lUj8gB3lrmZ-kLBbYbvlHuNIG2G9LmNwO2NCgH5c>).

## Supplementary data

Supplementary Data are available at NAR Online.

## Acknowledgements

We thank members of the Parker lab for helpful discussion and advice. We also thank Andrew Lemoff of the UTSW Proteomics Core for guidance in assessing ORC phosphorylation by mass spectrometry. M.W.P. is the Cecil H. and Ida Green Endowed Scholar in Biomedical Computational Science. The funders had no role in study design, data collection and analysis, decision to publish or preparation of manuscript.

## Funding

Welch Foundation [I-2074-20210327 to M.W.P.]; Cancer Prevention and Research Institute of Texas (CPRIT) [RR200070 to M.W.P.]; National Science Foundation (NSF) [2308642 to M.W.P.]. Funding for open access charge: University of Texas Southwestern Medical Center.

## Conflict of interest statement

None declared.

## References

1. Parker, M.W., Botchan, M.R. and Berger, J.M. (2017) Mechanisms and regulation of DNA replication initiation in eukaryotes. *Crit. Rev. Biochem. Mol. Biol.*, **52**, 107–144.
2. Remus, D., Beuron, F., Tolun, G., Griffith, J.D., Morris, E.P. and Diffley, J.F. (2009) Concerted loading of Mcm2–7 double hexamers around DNA during DNA replication origin licensing. *Cell*, **139**, 719–730.
3. Evrin, C., Clarke, P., Zech, J., Lurz, R., Sun, J., Uhle, S., Li, H., Stillman, B. and Speck, C. (2009) A double-hexameric MCM2–7 complex is loaded onto origin DNA during licensing of eukaryotic DNA replication. *Proc. Natl Acad. Sci. U.S.A.*, **106**, 20240–20245.
4. Bell, S.P. and Stillman, B. (1992) ATP-dependent recognition of eukaryotic origins of DNA replication by a multiprotein complex. *Nature*, **357**, 128–134.
5. Chesnokov, I., Remus, D. and Botchan, M. (2001) Functional analysis of mutant and wild-type *Drosophila* origin recognition complex. *Proc. Natl Acad. Sci. U.S.A.*, **98**, 11997–12002.
6. Bielinsky, A.K., Blitzblau, H., Beall, E.L., Ezrokhi, M., Smith, H.S., Botchan, M.R. and Gerbi, S.A. (2001) Origin recognition complex binding to a metazoan replication origin. *Curr. Biol.*, **11**, 1427–1431.
7. Vashee, S., Cvetič, C., Lu, W., Simancek, P., Kelly, T.J. and Walter, J.C. (2003) Sequence-independent DNA binding and replication initiation by the human origin recognition complex. *Genes Dev.*, **17**, 1894–1908.
8. Bleichert, F., Leitner, A., Aebbersold, R., Botchan, M.R. and Berger, J.M. (2018) Conformational control and DNA-binding mechanism of the metazoan origin recognition complex. *Proc. Natl Acad. Sci. U.S.A.*, **115**, E5906–E5915.
9. Li, N., Lam, W.H., Zhai, Y., Cheng, J., Cheng, E., Zhao, Y., Gao, N. and Tye, B.K. (2018) Structure of the origin recognition complex bound to DNA replication origin. *Nature*, **559**, 217–222.
10. Jaremko, M.J., On, K.F., Thomas, D.R., Stillman, B. and Joshua-Tor, L. (2020) The dynamic nature of the human origin

- recognition complex revealed through five cryoEM structures. *eLife*, **9**, e58622.
11. Liang, C. and Stillman, B. (1997) Persistent initiation of DNA replication and chromatin-bound MCM proteins during the cell cycle in *cdc6* mutants. *Genes Dev.*, **11**, 3375–3386.
  12. Findeisen, M., El-Denari, M., Kapitza, T., Graf, R. and Strausfeld, U. (1999) Cyclin A-dependent kinase activity affects chromatin binding of ORC, Cdc6, and MCM in egg extracts of *Xenopus laevis*. *Eur. J. Biochem.*, **264**, 415–426.
  13. Li, C.J., Vassilev, A. and DePamphilis, M.L. (2004) Role for Cdk1 (Cdc2)/cyclin A in preventing the mammalian origin recognition complex's largest subunit (Orc1) from binding to chromatin during mitosis. *Mol. Cell. Biol.*, **24**, 5875–5886.
  14. Lee, J.K., Moon, K.Y., Jiang, Y. and Hurwitz, J. (2001) The *Schizosaccharomyces pombe* origin recognition complex interacts with multiple AT-rich regions of the replication origin DNA by means of the AT-hook domains of the spOrc4 protein. *Proc. Natl Acad. Sci. U.S.A.*, **98**, 13589–13594.
  15. Chuang, R.Y. and Kelly, T.J. (1999) The fission yeast homologue of Orc4p binds to replication origin DNA via multiple AT-hooks. *Proc. Natl Acad. Sci. U.S.A.*, **96**, 2656–2661.
  16. Moon, K.Y., Kong, D., Lee, J.K., Raychaudhuri, S. and Hurwitz, J. (1999) Identification and reconstitution of the origin recognition complex from *Schizosaccharomyces pombe*. *Proc. Natl Acad. Sci. U.S.A.*, **96**, 12367–12372.
  17. Kong, D. and DePamphilis, M.L. (2001) Site-specific DNA binding of the *Schizosaccharomyces pombe* origin recognition complex is determined by the Orc4 subunit. *Mol. Cell. Biol.*, **21**, 8095–8103.
  18. Houchens, C.R., Lu, W., Chuang, R.Y., Frattini, M.G., Fuller, A., Simanek, P. and Kelly, T.J. (2008) Multiple mechanisms contribute to *Schizosaccharomyces pombe* origin recognition complex–DNA interactions. *J. Biol. Chem.*, **283**, 30216–30224.
  19. Parker, M.W., Bell, M., Mir, M., Kao, J.A., Darzacq, X., Botchan, M.R. and Berger, J.M. (2019) A new class of disordered elements controls DNA replication through initiator self-assembly. *eLife*, **8**, e48562.
  20. Austin, R.J., Orr-Weaver, T.L. and Bell, S.P. (1999) *Drosophila* ORC specifically binds to ACE3, an origin of DNA replication control element. *Genes Dev.*, **13**, 2639–2649.
  21. Kuo, A.J., Song, J., Cheung, P., Ishibe-Murakami, S., Yamazoe, S., Chen, J.K., Patel, D.J. and Gozani, O. (2012) The BAH domain of ORC1 links H4K20me2 to DNA replication licensing and Meier–Gorlin syndrome. *Nature*, **484**, 115–119.
  22. Noguchi, K., Vassilev, A., Ghosh, S., Yates, J.L. and DePamphilis, M.L. (2006) The BAH domain facilitates the ability of human Orc1 protein to activate replication origins *in vivo*. *EMBO J.*, **25**, 5372–5382.
  23. Balasov, M., Huijbregts, R.P. and Chesnokov, I. (2007) Role of the Orc6 protein in origin recognition complex-dependent DNA binding and replication in *Drosophila melanogaster*. *Mol. Cell. Biol.*, **27**, 3143–3153.
  24. Samwer, M., Schneider, M.W.G., Hoefler, R., Schmalhorst, P.S., Jude, J.G., Zuber, J. and Gerlich, D.W. (2017) DNA cross-bridging shapes a single nucleus from a set of mitotic chromosomes. *Cell*, **170**, 956–972.
  25. McConnell, B.S. and Parker, M.W. (2023) Protein intrinsically disordered regions have a non-random, modular architecture. *Bioinformatics*, **39**, btad732.
  26. Klemm, R.D., Austin, R.J. and Bell, S.P. (1997) Coordinate binding of ATP and origin DNA regulates the ATPase activity of the origin recognition complex. *Cell*, **88**, 493–502.
  27. Su, T.T. and O'Farrell, P.H. (1997) Chromosome association of minichromosome maintenance proteins in *Drosophila* mitotic cycles. *J. Cell Biol.*, **139**, 13–21.
  28. Katsani, K.R., Karess, R.E., Dostatni, N. and Doye, V. (2008) *In vivo* dynamics of *Drosophila* nuclear envelope components. *Mol. Biol. Cell*, **19**, 3652–3666.
  29. Baldinger, T. and Gossen, M. (2009) Binding of *Drosophila* ORC proteins to anaphase chromosomes requires cessation of mitotic cyclin-dependent kinase activity. *Mol. Cell. Biol.*, **29**, 140–149.
  30. Schmidt, J.M. and Bleichert, F. (2020) Structural mechanism for replication origin binding and remodeling by a metazoan origin recognition complex and its co-loader Cdc6. *Nat. Commun.*, **11**, 4263.
  31. Remus, D., Beall, E.L. and Botchan, M.R. (2004) DNA topology, not DNA sequence, is a critical determinant for *Drosophila* ORC–DNA binding. *EMBO J.*, **23**, 897–907.
  32. Park, J.O., Rubin, S.A., Xu, Y.F., Amador-Noguez, D., Fan, J., Shlomi, T. and Rabinowitz, J.D. (2016) Metabolite concentrations, fluxes and free energies imply efficient enzyme usage. *Nat. Chem. Biol.*, **12**, 482–489.
  33. Andersen, O.S. (2013) Cellular electrolyte metabolism. In: *Encyclopedia of Metalloproteins*. Springer, NY, pp. 580–587.
  34. Remus, D., Blanchette, M., Rio, D.C. and Botchan, M.R. (2005) CDK phosphorylation inhibits the DNA-binding and ATP-hydrolysis activities of the *Drosophila* origin recognition complex. *J. Biol. Chem.*, **280**, 39740–39751.
  35. Patil, A., Teraguchi, S., Dinh, H., Nakai, K. and Standley, D.M. (2012) Functional annotation of intrinsically disordered domains by their amino acid content using IDD navigator. In: *Pacific Symposium on Biocomputing*. Kohala Coast, Hawaii, USA, pp. 164–175.
  36. Emenecker, R.J., Griffith, D. and Holehouse, A.S. (2021) Metapredict: a fast, accurate, and easy-to-use predictor of consensus disorder and structure. *Biophys. J.*, **120**, 4312–4319.
  37. Attali, I., Botchan, M.R. and Berger, J.M. (2021) Structural mechanisms for replicating DNA in eukaryotes. *Annu. Rev. Biochem.*, **90**, 77–106.
  38. Brown, C.J., Takayama, S., Campen, A.M., Vise, P., Marshall, T.W., Oldfield, C.J., Williams, C.J. and Keith Dunker, A. (2002) Evolutionary rate heterogeneity in proteins with long disordered regions. *J. Mol. Evol.*, **55**, 104–110.
  39. Kumar, S., Suleski, M., Craig, J.M., Kasprowicz, A.E., Sanderford, M., Li, M., Stecher, G. and Hedges, S.B. (2022) TimeTree 5: an expanded resource for species divergence times. *Mol. Biol. Evol.*, **39**, msac174.
  40. Azmi, I.F., Watanabe, S., Maloney, M.F., Kang, S., Belsky, J.A., MacAlpine, D.M., Peterson, C.L. and Bell, S.P. (2017) Nucleosomes influence multiple steps during replication initiation. *eLife*, **6**, e22512.
  41. Simpson, R.T. (1990) Nucleosome positioning can affect the function of a *cis*-acting DNA element *in vivo*. *Nature*, **343**, 387–389.
  42. Perišić, O., Collepardo-Guevara, R. and Schlick, T. (2010) Modeling studies of chromatin fiber structure as a function of DNA linker length. *J. Mol. Biol.*, **403**, 777–802.
  43. Baldi, S., Krebs, S., Blum, H. and Becker, P.B. (2018) Genome-wide measurement of local nucleosome array regularity and spacing by nanopore sequencing. *Nat. Struct. Mol. Biol.*, **25**, 894–901.
  44. Woodcock, C.L., Skoultschi, A.I. and Fan, Y. (2006) Role of linker histone in chromatin structure and function: H1 stoichiometry and nucleosome repeat length. *Chromosome Res.*, **14**, 17–25.
  45. Miller, T.C.R., Locke, J., Greiwe, J.F., Diffley, J.F.X. and Costa, A. (2019) Mechanism of head-to-head MCM double-hexamer formation revealed by cryo-EM. *Nature*, **575**, 704–710.
  46. Eaton, M.L., Galani, K., Kang, S., Bell, S.P. and MacAlpine, D.M. (2010) Conserved nucleosome positioning defines replication origins. *Genes Dev.*, **24**, 748–753.
  47. Eaton, M.L., Prinz, J.A., MacAlpine, H.K., Tretyakov, G., Kharchenko, P.V. and MacAlpine, D.M. (2011) Chromatin signatures of the *Drosophila* replication program. *Genome Res.*, **21**, 164–174.
  48. MacAlpine, H.K., Gordan, R., Powell, S.K., Hartemink, A.J. and MacAlpine, D.M. (2010) *Drosophila* ORC localizes to open chromatin and marks sites of cohesin complex loading. *Genome Res.*, **20**, 201–211.

49. Deal,R.B., Henikoff,J.G. and Henikoff,S. (2010) Genome-wide kinetics of nucleosome turnover determined by metabolic labeling of histones. *Science*, **328**, 1161–1164.
50. Knott,S.R.V., Peace,J.M., Ostrow,A.Z., Gan,Y., Rex,A.E., Viggiani,C.J., Tavaré,S. and Aparicio,O.M. (2012) Forkhead transcription factors establish origin timing and long-range clustering in *S. cerevisiae*. *Cell*, **148**, 99–111.
51. Brodsky,S., Jana,T., Mittelman,K., Chapal,M., Kumar,D.K., Carmi,M. and Barkai,N. (2020) Intrinsically disordered regions direct transcription factor *in vivo* binding specificity. *Mol. Cell*, **79**, 459–471.
52. Chong,S.S., Dugast-Darzacq,C., Liu,Z., Dong,P., Dailey,G.M., Cattoglio,C., Heckert,A., Banala,S., Lavis,L., Darzacq,X., *et al.* (2018) Imaging dynamic and selective low-complexity domain interactions that control gene transcription. *Science*, **361**, 378.
53. Sabari,B.R., Dall'Agnese,A., Boija,A., Klein,I.A., Coffey,E.L., Shrinivas,K., Abraham,B.J., Hannett,N.M., Zamudio,A.V., Manteiga,J.C., *et al.* (2018) Coactivator condensation at super-enhancers links phase separation and gene control. *Science*, **361**, eaar3958.
54. Hossain,M., Bhalla,K. and Stillman,B. (2021) Multiple, short protein binding motifs in ORC1 and CDC6 control the initiation of DNA replication. *Mol. Cell*, **81**, 1951–1969.
55. Hyrien,O. and Mechali,M. (1993) Chromosomal replication initiates and terminates at random sequences but at regular intervals in the ribosomal DNA of *Xenopus* early embryos. *EMBO J.*, **12**, 4511–4520.
56. Parker,M.W., Kao,J.A., Huang,A., Berger,J.M. and Botchan,M.R. (2021) Molecular determinants of phase separation for *Drosophila* DNA replication licensing factors. *eLife*, **10**, e70535.
57. Li,X. and Fu,X.D. (2019) Chromatin-associated RNAs as facilitators of functional genomic interactions. *Nat. Rev. Genet.*, **20**, 503–519.
58. Norseen,J., Thomae,A., Sridharan,V., Aiyar,A., Schepers,A. and Lieberman,P.M. (2008) RNA-dependent recruitment of the origin recognition complex. *EMBO J.*, **27**, 3024–3035.
59. Zhang,A.T., Langley,A.R., Christov,C.P., Kheir,E., Shafee,T., Gardiner,T.J. and Krude,T. (2011) Dynamic interaction of Y RNAs with chromatin and initiation proteins during human DNA replication. *J. Cell Sci.*, **124**, 2058–2069.
60. Mas,A.M., Goñi,E., de los Mozos,I.R., Arcas,A., Statello,L., Gonzalez,J., Blázquez,L., Lee,W.T.C., Gupta,D., Sejas,Á., *et al.* (2023) ORC1 binds to *cis*-transcribed RNAs for efficient activation of replication origins. *Nat. Commun.*, **14**, 4447.
61. Trinkle-Mulcahy,L., Andersen,J., Yun,W.L., Moorhead,G., Mann,M. and Lamond,A.I. (2006) Repo-Man recruits PP1 $\gamma$  to chromatin and is essential for cell viability. *J. Cell Biol.*, **172**, 679–692.
62. Mei,L. and Cook,J.G. (2021) Efficiency and equity in origin licensing to ensure complete DNA replication. *Biochem. Soc. Trans.*, **49**, 2133–2141.
63. Zheng,H., Qi,Y., Hu,S., Cao,X., Xu,C., Yin,Z., Chen,X., Li,Y., Liu,W., Li,J., *et al.* (2020) Identification of Integrator–PP2A complex (INTAC), an RNA polymerase II phosphatase. *Science*, **370**, eabb5872.
64. Lu,X., Hamkalo,B., Parseghian,M.H. and Hansen,J.C. (2009) Chromatin condensing functions of the linker histone C-terminal domain are mediated by specific amino acid composition and intrinsic protein disorder. *Biochemistry*, **48**, 164–172.
65. Ross,E.D., Edskes,H.K., Terry,M.J. and Wickner,R.B. (2005) Primary sequence independence for prion formation. *Proc. Natl Acad. Sci. U.S.A.*, **102**, 12825–12830.
66. Ross,E.D., Baxa,U. and Wickner,R.B. (2004) Scrambled prion domains form prions and amyloid. *Mol. Cell Biol.*, **24**, 7206–7213.