Article

Phase separation directs ubiquitination of gene-body nucleosomes

https://doi.org/10.1038/s41586-020-2097-z

Received: 10 June 2019

Accepted: 11 February 2020

Published online: 11 March 2020

Check for updates

Laura D. Gallego[1,3], Maren Schneider[1,3], Chitvan Mittal[2,3], Anete Romanauska[1], Ricardo M. Gudino Carrillo[1], Tobias Schubert[1], B. Franklin Pugh[2] & Alwin Köhler[1]✉

The conserved yeast E3 ubiquitin ligase Bre1 and its partner, the E2 ubiquitin-conjugating enzyme Rad6, monoubiquitinate histone H2B across gene bodies during the transcription cycle[1]. Although processive ubiquitination might—in principle—arise from Bre1 and Rad6 travelling with RNA polymerase II[2], the mechanism of H2B ubiquitination across genic nucleosomes remains unclear. Here we implicate liquid–liquid phase separation[3] as the underlying mechanism. Biochemical reconstitution shows that Bre1 binds the scaffold protein Lge1, which possesses an intrinsically disordered region that phase-separates via multivalent interactions. The resulting condensates comprise a core of Lge1 encapsulated by an outer catalytic shell of Bre1. This layered liquid recruits Rad6 and the nucleosomal substrate, which accelerates the ubiquitination of H2B. In vivo, the condensate-forming region of Lge1 is required to ubiquitinate H2B in gene bodies beyond the +1 nucleosome. Our data suggest that layered condensates of histone-modifying enzymes generate chromatin-associated 'reaction chambers', with augmented catalytic activity along gene bodies. Equivalent processes may occur in human cells, and cause neurological disease when impaired.

Eukaryotic cells require the precisely timed activation of genes, which coincides with chromatin alterations. The fundamental unit of chromatin is the nucleosome core particle (NCP), which is assembled from histones and DNA, and iterated to resemble 'beads on a string'. Post-translational modifications of histones affect chromatin structure and thus gene activity. How histone-modifying enzymes are concentrated in discrete parts of the genome to install histone marks locally is a key question. Histone H2B is monoubiquitinated at lysine 123 (H2BK123ub) in yeast and at lysine 120 in mammals[1,2]. This mark is found across promoter, genic and termination regions of most genes in *Saccharomyces cerevisiae*[4]. H2BK123ub influences DNA transcription, repair and replication, and RNA processing[1]. H2BK123ub exerts its transcriptional effects by regulating NCP assembly, stability[4] and specific histone methylations[1]. Aberrant levels of H2B ubiquitination are found in various diseases, including cancer[1]. It is therefore important to understand how H2BK123ub is established in specific regions of the genome.

Ubiquitination is a three-step enzymatic cascade, which requires a ubiquitin-activating (E1), a ubiquitin-conjugating (E2) and a ubiquitin-ligating (E3) enzyme. In yeast, H2BK123ub is generated by the E3 Bre1 (an orthologue of the human RNF20 and RNF40 proteins (hereafter, RNF20/RNF40)), together with the E2 Rad6 and the E1 Uba1[5–9]. A common view is that H2BK123ub is deposited in a cotranscriptional manner[1,2]. However, H2B ubiquitination does not strictly depend on Bre1–Rad6 binding to RNA polymerase (Pol) II[10]. Thus, the mechanism of H2B ubiquitination across genic NCPs is unclear.

To understand how the ubiquitination machinery is enriched at the correct genomic positions, we considered the Large 1 (Lge1) protein, which has an as-yet unknown biochemical role, copurifies with Bre1 and is essential for H2B monoubiquitination in yeast[7]. Except for a short C-terminal coiled coil, Lge1 is predicted to consist entirely of an intrinsically disordered region (IDR) (Fig. 1a, Extended Data Fig. 2a, b). The human Bre1 orthologues RNF20/RNF40 interact with WAC, a protein that is similar to Lge1 in its IDR and coiled-coil domain architecture[11] (Extended Data Fig. 2b). Some proteins that contain an IDR participate in multivalent interactions, which can lead to liquid–liquid phase separation (LLPS) and the formation of biomolecular condensates[3,12,13]. LLPS has previously been implicated in chromatin organization[14], heterochromatin compaction[15] as well as promoter and enhancer function[16–19].

Here we define a direct role of LLPS in stimulating histone ubiquitination specifically across gene bodies. Lge1, Bre1, Rad6 and nucleosomes together form spatially organized reaction chambers that stimulate H2B ubiquitination.

Lge1 promotes Bre1 oligomerization

We first explored interactions between Lge1 and Bre1, and their capacity to phase-separate. Bre1 consists of a series of coiled-coils, flanked by an N-terminal Rad6-binding domain and a C-terminal RING domain[20] (Fig. 1a). Bre1 associates with Lge1 when purified from yeast (Extended Data Fig. 1a). Lge1 tagged with glutathione *S*-transferase (GST) purified from *Escherichia coli* bound the 'middle' part of recombinant Bre1, which we therefore call the Lge1-binding domain (LBD; comprising residues 308–632) (Extended Data Fig. 1b, c, Supplementary Fig. 2, Supplementary Table 1). The Lge1 coiled coil alone (Lge1(CC), comprising amino acids 243–332) was necessary and sufficient for binding to Bre1 (Extended Data Fig. 1d). Thus, coiled-coil contacts mediate a direct interaction between Lge1 and Bre1.

[1]Max Perutz Labs, Medical University of Vienna, Vienna Biocenter Campus (VBC), Vienna, Austria. [2]Department of Biochemistry and Molecular Biology, Center for Eukaryotic Gene Regulation, Pennsylvania State University, University Park, PA, USA. [3]These authors contributed equally: Laura D. Gallego, Maren Schneider, Chitvan Mittal. ✉e-mail: alwin.koehler@mfpl.ac.at

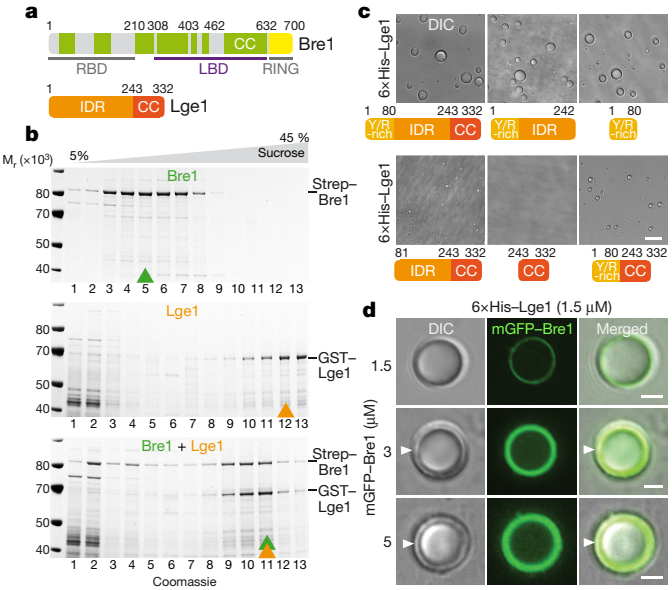


Fig. 1 | Lge1 forms core–shell condensates with Bre1. a, Organization of Bre1 and Lge1. CC, coiled-coil domain (green; non-coiled-coil regions, grey); LBD, Lge1-binding domain; RBD, Rad6-binding domain; RING, really interesting new gene domain. **b**, Sucrose gradient analysis of Strep-tagged Bre1 (Strep–Bre1) and GST–Lge1. Arrowheads label peak fractions. See Supplementary Fig. 1 for uncropped gels. **c**, LLPS assay with 6×His–Lge1 constructs (1.5 μM) (cartoons). Scale bar, 10 μm. See Supplementary Video 1 for condensate fusion events. **d**, Reconstitution of core–shell condensates with increasing amounts of mGFP–Bre1 added to preformed 6×His–Lge1 condensates. Arrowheads label the shell in differential interference contrast (DIC) images. Scale bar, 2 μm.

We next analysed the oligomerization state of Bre1 and how this state might be influenced by the IDR of Lge1. Bre1 showed a broad peak in the low-molecular-weight region of a sucrose gradient, whereas Lge1 sedimented in a distinct high-molecular-weight region. When mixed with Bre1, Lge1 shifted a substantial portion of Bre1 into the high-molecular-weight region (Fig. 1b, Supplementary Fig. 1). The Lge1(CC) alone sedimented in the low-molecular-weight region, whereas the Lge1 IDR alone (Lge1(IDR), comprising amino acids 1–242) sedimented in the high-molecular-weight region (Extended Data Fig. 1e). Lge1(IDR) (which does not bind Bre1) left Bre1 in the low-molecular-weight region. Thus, Lge1 is a scaffold protein that—via its IDR—promotes Bre1 oligomerization.

Lge1 phase-separates in vitro

Lge1 has the hallmarks of proteins that undergo LLPS[3], including intrinsic disorder, low sequence complexity and blocks of alternating charges (Extended Data Fig. 2a, b). Recombinant full-length 6×His-tagged Lge1 indeed formed condensates, which grew in size to micrometre-sized spheres by fusion (Fig. 1c, Extended Data Fig. 1f–h, Supplementary Video 1). Phase separation occurred at 0.1–0.5 μM under physiological salt concentrations without crowding agents, and was inhibited by 1,6-hexanediol (Extended Data Fig. 4a, b). The sedimentation of Lge1 in sucrose gradients could thus be explained by LLPS (Extended Data Fig. 1k). Deletion of the IDR of Lge1 abolished LLPS, whereas the Lge1(IDR) alone was sufficient for LLPS (Fig. 1c, Extended Data Fig. 1g). The N terminus (amino acids 1–80) of Lge1 is enriched in tyrosine (Y) and arginine (R) residues (Extended Data Fig. 2a). Truncating this Y/R-rich region strongly decreased LLPS (Fig. 1c, Extended Data Fig. 1g). Conversely, the Y/R-rich region was sufficient for LLPS and showed LLPS when fused to Lge1(CC), albeit with a lower efficiency. In sum, a short Y/R-rich 'sticker' region at the Lge1 N terminus may act as a seed for LLPS.

A Bre1 catalytic shell encapsulates Lge1

Bre1 does not phase-separate on its own (Extended Data Fig. 4c, d); Lge1 therefore behaves as a scaffold and Bre1 as a client. When we used recombinant Bre1 tagged with monomeric green fluorescent protein (mGFP–Bre1) at equimolar ratio to Lge1, Bre1 penetrated the Lge1 condensate only to a certain depth and formed a shell around an Lge1 core. This shell grew in thickness with increasing concentrations of Bre1 (Fig. 1d, Extended Data Fig. 1j). To understand how the core–shell architecture is established, we added mGFP–Bre1 to condensates formed by Lge1(IDR) that lacked the coiled coil. In these experiments, Bre1 no longer assembled a shell and instead diffused into the condensate (Extended Data Fig. 1i, j). The Bre1(LBD) was sufficient for shell formation (albeit with lower efficiency), and this specifically depended on the coiled coil of Lge1 (Extended Data Fig. 3a, b). To further test the role of the Lge1–Bre1 interaction in shell formation, we created hybrid condensates composed of varying ratios of full-length Lge1 mixed with Lge1(IDR). Bre1 could penetrate deeper into the core when fewer Lge1 coiled coils were present (Extended Data Fig. 3c, d). This suggests that the coiled-coil domain of Lge1 captures and organizes Bre1 as it infiltrates the structure from the periphery. Notably, the Bre1 shell inhibited condensate fusion (Extended Data Fig. 3e, Supplementary Video 2) and thereby restricted condensate growth (Extended Data Fig. 4d). Thus, Lge1 and Bre1 form a distinct membraneless compartment with a catalytic E3 shell that encapsulates a liquid-like Lge1 core.

The material properties of a condensate determine which molecules are incorporated into or excluded from it. We characterized the effective mesh size using fluorescently labelled dextrans that ranged from 35 to 2,000 kDa in size[21]. None of the probes was excluded, and the partitioning ratios were similar with or without a Bre1 shell (Extended Data Fig. 4e, f). The 2,000-kDa probe has a hydrodynamic radius (R_h) of about 27 nm in dilute aqueous buffer, similar to a 16-unit nucleosomal array (Extended Data Fig. 4e, g). Thus, Lge1–Bre1 condensates permit the penetration of large molecules.

A key question about LLPS is how particular interactions shape the mesh and function of a condensate[3,22]. Both Y>A and R>K mutations in the Lge1 sticker region strongly decreased LLPS (Extended Data Fig. 5a–c). In the context of full-length Lge1, Y>A mutations in the sticker region decreased LLPS, and three additional Y>A mutations within amino acids 1–102 of Lge1 eliminated LLPS. The R>K mutations in full-length Lge1 showed only a modest decrease in LLPS. In sum, a multivalent network underlies the dynamic behaviour of Lge1 condensates, and tyrosines have a key role.

LLPS accelerates H2B ubiquitination

To understand the function of the core–shell condensates, we explored their influence on H2B ubiquitination. Fluorescently labelled Rad6 was rapidly recruited to the Bre1 shell of preformed Lge1–Bre1 condensates (Fig. 2a). Subsequently, Rad6 diffused further into the core, and co-enrichment with Bre1 in the shell still persisted. When Bre1 was omitted, Rad6 did not transiently accumulate in the outer shell, but instead became rapidly distributed throughout the condensate (Extended Data Fig. 6a). The recruitment of Rad6 to the Bre1 shell depended on a direct interaction, as shown by the fact that the Bre1(LBD), which lacks the RING domain and the Rad6-binding domain, did not concentrate Rad6 in the shell (Extended Data Fig. 6b). This demonstrates that Lge1 condensates can co-enrich both the E3 Bre1 and the E2 Rad6 in a catalytic shell.

For the condensates to qualify as a reaction chamber for ubiquitination, nucleosomes should also become concentrated. A 16-unit NCP array, reconstituted from yeast histones on a 16× Widom sequence, was recruited to the catalytic shell of the Lge1–Bre1 condensate (Fig. 2b, Extended Data Fig. 6c, d). Similar to Rad6, the chromatin substrate diffused into the condensate over time. A similar partitioning behaviour

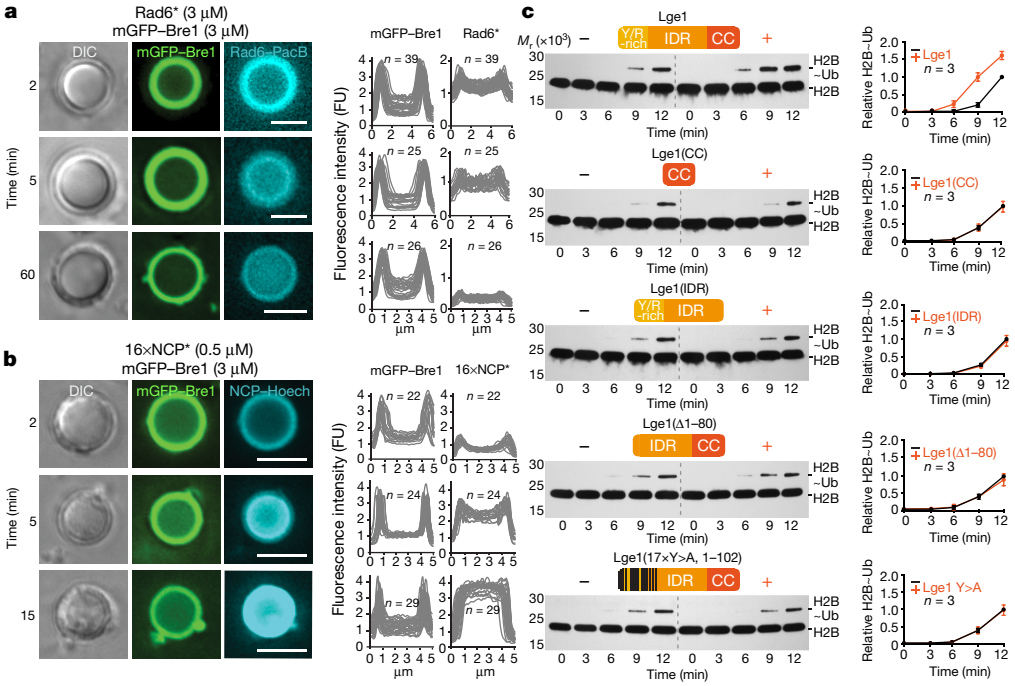


Fig. 2 | Lge1–Bre1 condensates promote H2BK123ub. a, Recruitment of Pacific-Blue-labelled 6×His–Rad6 (here denoted 'Rad6*') to Lge1 condensates with a Bre1 shell. Fluorescence intensities were quantified across single condensates. FU, arbitrary fluorescence units. *n*, number of condensates. Scale bar, 5 μm. **b**, Partitioning of Hoechst (Hoech)-labelled, reconstituted yeast oligonucleosome arrays (here denoted '16×NCP*') into Lge1–Bre1 condensates. Scale bar, 5 μm. **c**, Time-resolved in vitro H2B ubiquitination assay performed under LLPS conditions. ±, the indicated Lge1 constructs used at equimolar concentration. Intensity of the H2B-Ub band was quantified and normalized to 12-min time point without Lge1 for each assay. Lge1(17×Y>A, 1–102), Lge1 with mutations within amino acids 1–102 (see 'Protein purification' in Methods for details of the substitutions). *n*, independent experiments. Mean and s.d. are indicated.

was observed for mononucleosomes (Extended Data Fig. 6e). When the Bre1 shell was missing, chromatin immediately diffused into the interior of the Lge1 condensate (Extended Data Fig. 6f). DNA alone showed no strong accumulation (Extended Data Fig. 6g). In sum, LLPS spatially organizes the ubiquitination machinery and its substrate into a distinct core–shell structure.

To test the functional importance of these reaction chambers, we reconstituted NCP ubiquitination from defined components under phase-separating conditions (Fig. 2c). The addition of full-length Lge1 accelerated H2B ubiquitination. Importantly, the effect was abolished by removing the IDR of Lge1, leaving only the coiled-coil domain that interacts with Bre1. The Lge1(IDR) alone, which undergoes LLPS but does not assemble a Bre1 shell (Extended Data Fig. 3b), could not enhance the reaction. The stimulatory effect of full-length Lge1 required the sticker region—specifically, the tyrosine residues that promote LLPS. Thus, H2B ubiquitination was accelerated by conditions that correspond to the formation of Lge1–Bre1 core–shell condensates.

Endogenous Lge1 and Bre1 form large complexes

To determine whether Lge1–Bre1 condensates exist in yeast cells, we analysed the sedimentation of Lge1, Bre1 and Pol II at endogenous levels in whole-cell extracts using sucrose density gradients. Bre1 tagged with haemagglutinin was found throughout the gradient, and overlapped with the broad peak of Lge1 tagged with a tandem affinity purification tag in the high-molecular-weight region (Fig. 3a, Extended Data Fig. 6h, i). Pol II sedimented in the low-molecular-weight region, which suggests that Lge1–Bre1 multimerization is not due to an association with Pol II. When *LGE1* was deleted, Bre1 relocalized into the low-molecular-weight region of the gradient, consistent with our reconstitution experiments (Fig. 1b). Deletion of the IDR of Lge1 was sufficient to shift the remaining Lge1(CC) and Bre1 into the low-molecular-weight region. Removing only the sticker region had an effect that was similar to—but weaker than—removing the entire IDR of Lge1. Consistent with Lge1(IDR) condensates forming in the absence of Bre1 in vitro (Fig. 1c, Extended Data Fig. 1e), Lge1(IDR) alone sedimented in the high-molecular-weight region but did not shift Bre1 into this region. Overall, the multimerization state of Lge1 and Bre1 in cell extracts is consistent with the formation of core–shell structures that we observed in vitro (Fig. 1d).

To visualize Lge1 and Bre1 in cells, we co-overexpressed them under a strong promoter. Lge1 and Bre1 colocalized into prominent nuclear foci (Fig. 3b, Extended Data Fig. 7a, b). When expressed from their endogenous promoters, Lge1 and Bre1 were homogenously dispersed in the yeast nucleus (Extended Data Fig. 7c–f). To selectively visualize Bre1 that is bound to Lge1, we used bimolecular fluorescence complementation (BiFC) (Fig. 3c). Bre1 fused to the C-terminal fragment of Venus exhibited BiFC with full-length Lge1 fused to the N-terminal fragment of Venus (Fig. 3d, Extended Data Fig. 7h, i), confirming their interaction. The BiFC signal showed multiple puncta (Fig. 3d, Extended Data Fig. 7g), which may reflect condensates. We quantified puncta formation by measuring the heterogeneity (coefficient of variation) of fluorescent intensities within the nucleus (Fig. 3e). Deleting the Lge1 IDR or the Lge1 sticker region (resulting in Lge1(Δ1–80)) reduced the formation of BiFC puncta (Fig. 3d, e, Extended Data Fig. 7i). However, a homogenous BiFC signal was maintained (Fig. 3d), consistent with the presence of the coiled-coil domain of Lge1, which interacts with Bre1. As in the other assays, the effect of deleting the entire IDR was stronger than deleting the sticker region only (Figs. 1c, 3a). Together, these data suggest that the nuclear BiFC puncta correspond to the pool of Bre1 that undergoes LLPS when bound to Lge1.

Lge1 IDR affects gene-body ubiquitination

To address how LLPS contributes to H2B ubiquitination in cells, we measured global H2BK123ub levels. Deleting the IDR of Lge1 strongly reduced H2B ubiquitination (Extended Data Fig. 8a), consistent with

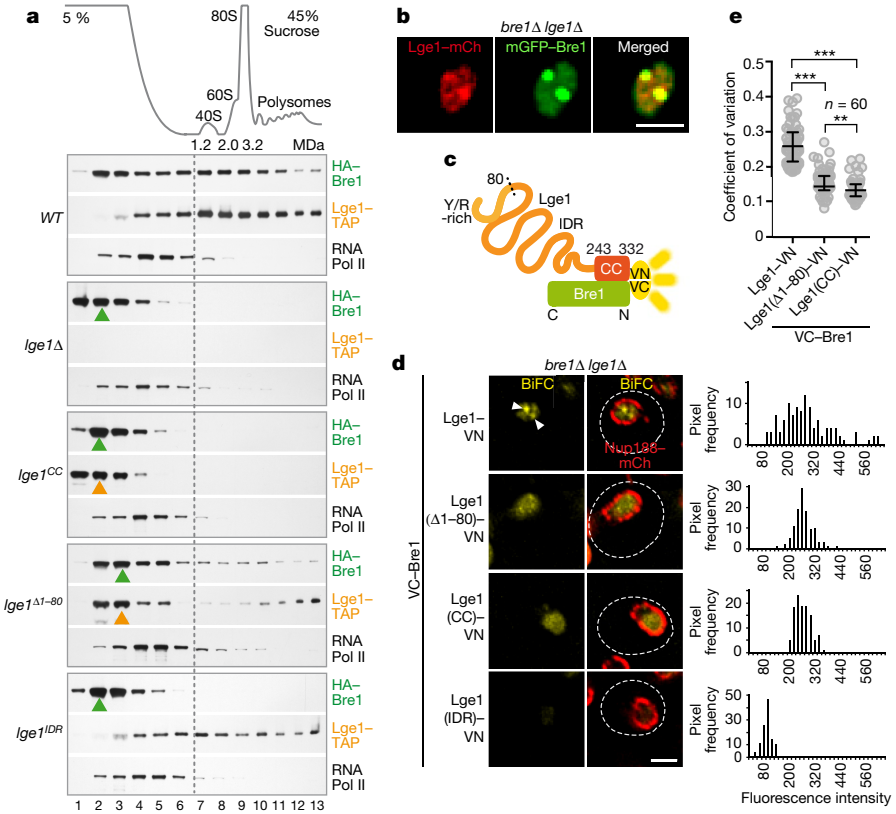


Fig. 3 | Endogenous Lge1 and Bre1 form large complexes and nuclear puncta. **a**, Sucrose gradient analysis of cell extracts from *lge1*Δ cells transformed with the indicated plasmids. Ribosomal species serve as size markers. Gradient fractions were analysed by immunoblotting. Peak fractions are highlighted (arrowheads). HA, haemagglutinin; TAP, tandem affinity purification tag; WT, wild-type *LGE1*. **b**, Live imaging of bre1Δ lge1Δ cells that coexpress the indicated constructs from a strong GPD promoter. Scale bar, 2 µm. mCh, mCherry. **c**, BiFC design. VN and VC denote the complementary Venus fragments (N-terminal and C-terminal fragments, respectively). **d**, Live imaging of bre1Δ lge1Δ cells expressing VC–Bre1 and the indicated Lge1–VN constructs from their endogenous promoters. Arrowheads label nuclear BiFC puncta, Nup188–mCherry labels the nuclear envelope and dashed line represents the cell contour. Histograms represent pixel frequencies of fluorescence intensity values. Scale bar, 2 µm. **e**, Coefficient of variation of nuclear BiFC fluorescence-intensity profiles (Extended Data Fig. 7i). The higher the coefficient of variation, the greater the heterogeneity of the BiFC signal. Median and interquartile range are indicated. *n*, number of analysed cells. $^{**}P = 0.0037$, $^{***}P < 0.001$, determined by two-sided Mann–Whitney test.

LLPS accelerating the enzymatic reaction in vitro (Fig. 2c). Deletion of the sticker region had a weaker effect than removal of the entire IDR of Lge1. Lge1(IDR) alone did not promote H2B ubiquitination, as it is not connected to Bre1 (see also Fig. 3a, d). To determine the genome-wide effect of Lge1–Bre1 LLPS, we analysed the levels of H2B and H2BK123ub by sequencing exonuclease-treated chromatin immunoprecipitates (ChIP-exo)[23], which resolved both subunits of H2B within each averaged nucleosome position (Fig. 4a, left). A *lge1Δ* strain showed increased nucleosome occupancy (approximated by H2B) downstream of the first (+1) gene-body nucleosome. Lge1(CC) was sufficient to sustain proper nucleosome occupancy and positioning, which suggests that Lge1–Bre1, but not Lge1-directed LLPS, is involved in these aspects of chromatin organization.

In examining H2BK123ub, we separated all genes into three well-studied groups—ribosomal-protein, SAGA-dominated and TFIID-dominated genes—that have distinct types of regulation[24,25] (Fig. 4a, Extended Data Figs. 8b, c, 9d, e). As expected, a *lge1Δ bre1Δ* strain showed substantial loss of H2BK123ub. An *lge1Δ* strain also showed reduced H2BK123ub across all gene-body nucleosomes. The greater loss in the*lge1Δ bre1Δ* strain indicates that Bre1 retains some activity in the absence of Lge1 (see also Extended Data Fig. 8a). When the Bre1-interacting coiled-coil domain of Lge1 was removed (*NLS-Lge1^{IDR}* strain), the H2BK123ub pattern resembled that in the *lge1Δ* strain, confirming the general importance of the Bre1–Lge1 interaction in regulating H2B ubiquitination in vivo (Extended Data Fig. 8a, b). By contrast, when the IDR of Lge1 was deleted, H2BK123ub levels were mainly reduced downstream of +1 nucleosomes (Fig. 4a). This suggests that LLPS is especially important for ubiquitination within gene bodies.

To examine whether any IDR could execute the LLPS function of Lge1, we tested two IDRs from distant species. The IDR of the *Caenorhabditis elegans* P-granule component LAF1 (amino acids 1–169) undergoes LLPS[26], but has no relation to H2B ubiquitination. *Homo sapiens* WAC (amino acids 1–318) was picked as the probable functional counterpart of Lge1[11] (Extended Data Fig. 2b). Using ChIP-exo, we found that the IDR of human WAC, but not the IDR of the unrelated LAF1, partially substituted for the IDR of yeast Lge1 when fused to Lge1(CC) (Fig. 4a, Extended Data Figs. 8b, 10a, b, d–g). This was further confirmed by measuring global H2BK123ub levels with immunoblotting (Extended Data Fig. 8a), which suggests that the IDRs of human WAC and yeast Lge1 have analogous functions.

The genome-wide binding profile of Lge1 showed enrichment at all gene classes, but was strongest at ribosomal-protein and SAGA-dominated genes (Extended Data Fig. 9a, b). Enrichment was maximal downstream of +1 nucleosomes (Extended Data Fig. 9c), consistent with the effect of deleting the IDR of Lge1 on H2BK123ub. The genome-wide distribution of Lge1 was largely unaffected by loss of Bre1 (Extended Data Fig. 9b, c). Thus, intrinsic features of Lge1 (or as-yet unknown factors) may guide Lge1 to its target locations. These results support a nucleosome-position-specific function of Lge1 condensates, downstream of the +1 nucleosome. Ubiquitination at the +1 position depends on Bre1, but not through Lge1-directed LLPS.

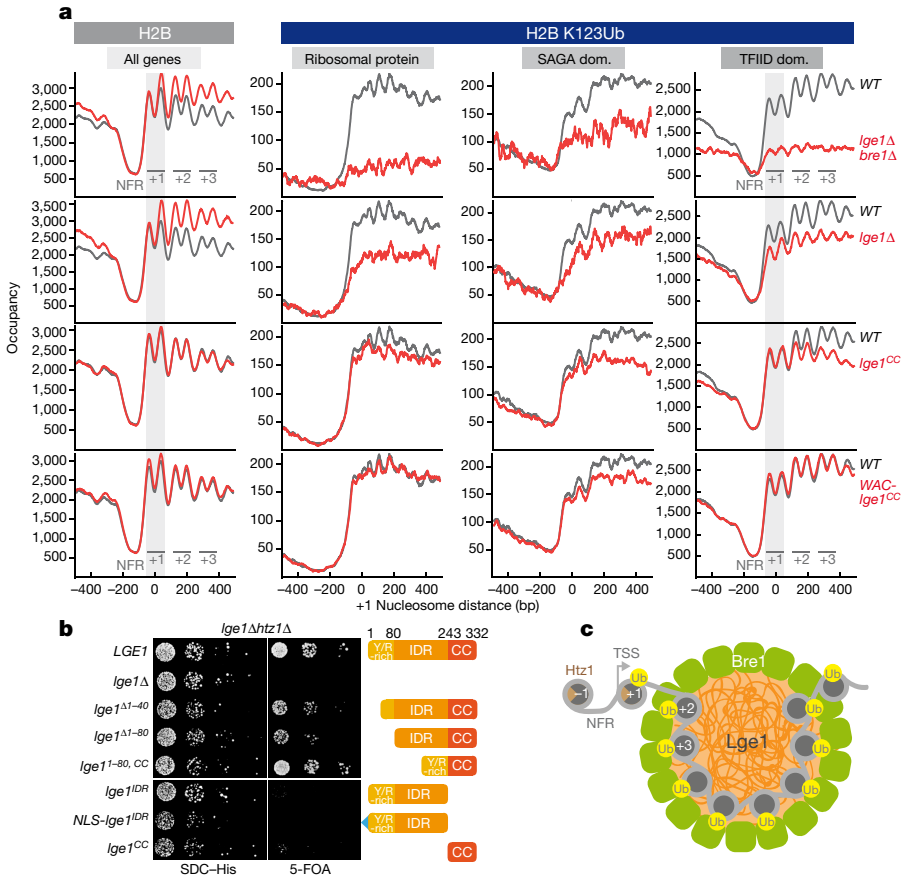


Fig. 4 | The IDR of Lge1 enhances H2BK123ub in gene bodies and is required for viability in *htz1Δ* cells. a, H2B (left) or H2BK123ub (right) ChIP-exo tag 5′ ends were plotted relative to the +1 nucleosome of all genes for wild type (grey trace) and the indicated *lge1* mutants (red traces). The first three genic nucleosomes are labelled +1, +2 and +3. Two H2B peaks are observed per nucleosome position. H2BK123ub patterns are shown separately for each gene class[29]. dom., dominated; NFR, nucleosome free region. **b**, Genetic interaction analysis. Double deletion strains containing a wild-type *LGE1* cover plasmid (Ura marker) were cotransformed with the indicated plasmids (His marker). Growth (tenfold serial dilutions) was followed on SDC–His (control) and on SDC + 5-fluoroorotic acid (5-FOA), which shuffles out the Ura cover plasmid. **c**, LLPS-based ubiquitination model. The Htz1-containing +1 NCP is depicted outside of the 'reaction chamber' as it is ubiquitinated independently of Lge1 LLPS. For simplicity, Rad6 and the E1 are omitted and only a single layer of Bre1 molecules in the shell is shown. TSS, transcription start site.

Lge1 IDR affects viability in *htz1Δ* cells

We next asked whether Lge1–Bre1 condensates affect cell physiology. Both *LGE1* and *BRE1* are essential for viability in cells that lack the conserved histone variant H2A.Z (yeast Htz1)[7] (Fig. 4b, Extended Data Fig. 7f). Htz1 assembles into +1 nucleosomes[27]—the only gene-body nucleosomes that do not depend on Lge1 LLPS for ubiquitination. A truncation of the sticker region of Lge1 (either Lge1(Δ1–80) or Lge1(Δ1–40)) showed a synthetically enhanced growth phenotype when combined with *htz1*Δ (a genetic interaction resulting in a synergistic growth defect), as did Y>A mutations in the N terminus of Lge1 (amino acids 1–102) (Fig. 4b, Extended Data Fig. 10h, i). The deletion of the middle region of Lge1 (amino acids 81–242) displayed no obvious growth defect with *htz1*Δ. Notably, however, deletion of the entire IDR was synthetically lethal with *htz1*Δ—despite the presence of the Bre1-interacting coiled-coil domain of Lge1. Expression of Lge1(IDR) alone, which cannot bind Bre1, was similarly synthetic lethal (Fig. 4b). The IDR of human WAC—but not that of LAF1—partially compensated for the deletion of the IDR of Lge1 (Extended Data Fig. 10c, e), consistent with their effect on H2B ubiquitination of gene-body nucleosomes. Overall, these genetic assays suggest that Lge1 LLPS becomes essential for viability when Htz1 is absent from +1 nucleosomes, reinforcing the physiological importance of Lge1–Bre1 condensates.

We propose that the Bre1 shell has a direct catalytic role, whereas the core concentrates the E2 Rad6 and the chromatin substrate. By confining the reactants in a small space at high concentration, LLPS increases the opportunity for productive interactions many fold. Moreover, by avidity, the strength and specificity of multiple simultaneous Bre1–NCP interactions will be greater than the sum of individual binding events. The human orthologues of Bre1 (RNF20/RNF40) copurify with WAC, which may perform a function analogous to that of Lge1 in human cells[11]. Despite no obvious sequence similarity, WAC and Lge1 share a similar domain organization (Extended Data Fig. 2b). Our results provide an entry point to study whether LLPS contributes to DeSanto–Shinawi syndrome, a neurodevelopmental disorder that is caused by mutations that truncate WAC[28] (Extended Data Fig. 2b). In conclusion, we have discovered that LLPS augments the catalytic activity of the chromatin modifier Bre1. We propose that core–shell condensates form dynamic histone ubiquitination hubs, which target gene-body nucleosomes (Fig. 4c) and thereby shape gene architecture and expression.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-020-2097-z.

1. Fuchs, G. & Oren, M. Writing and reading H2B monoubiquitylation. *Biochim. Biophys. Acta* **1839**, 694–701 (2014).
2. Weake, V. M. & Workman, J. L. Histone ubiquitination: triggering gene activity. *Mol. Cell* **29**, 653–663 (2008).
3. Banani, S. F., Lee, H. O., Hyman, A. A. & Rosen, M. K. Biomolecular condensates: organizers of cellular biochemistry. *Nat. Rev. Mol. Cell Biol.* **18**, 285–298 (2017).
4. Batta, K., Zhang, Z., Yen, K., Goffman, D. B. & Pugh, B. F. Genome-wide function of H2B ubiquitylation in promoter and genic regions. *Genes Dev.* **25**, 2254–2265 (2011).
5. Cucinotta, C. E., Young, A. N., Klucevsek, K. M. & Arndt, K. M. The nucleosome acidic patch regulates the H2B K123 monoubiquitylation cascade and transcription elongation in *Saccharomyces cerevisiae*. *PLoS Genet.* **11**, e1005420 (2015).
6. Gallego, L. D. et al. Structural mechanism for the recognition and ubiquitination of a single nucleosome residue by Rad6–Bre1. *Proc. Natl Acad. Sci. USA* **113**, 10553–10558 (2016).
7. Hwang, W. W. et al. A conserved RING finger protein required for histone H2B monoubiquitination and cell size control. *Mol. Cell* **11**, 261–266 (2003).
8. Kim, J. et al. RAD6-mediated transcription-coupled H2B ubiquitylation directly stimulates H3K4 methylation in human cells. *Cell* **137**, 459–471 (2009).
9. Wood, A. et al. Bre1, an E3 ubiquitin ligase required for recruitment and substrate selection of Rad6 at a promoter. *Mol. Cell* **11**, 267–274 (2003).
10. Van Oss, S. B. et al. The histone modification domain of Paf1 complex subunit Rtf1 directly stimulates H2B ubiquitylation through an interaction with Rad6. *Mol. Cell* **64**, 815–825 (2016).
11. Zhang, F. & Yu, X. WAC, a functional partner of RNF20/40, regulates histone H2B ubiquitination and gene transcription. *Mol. Cell* **41**, 384–397 (2011).
12. Alberti, S. Phase separation in biology. *Curr. Biol.* **27**, R1097–R1102 (2017).
13. Shin, Y. & Brangwynne, C. P. Liquid phase condensation in cell physiology and disease. *Science* **357**, eaaf4382 (2017).
14. Gibson, B. A. et al. Organization of chromatin by intrinsic and regulated phase separation. *Cell* **179**, 470–484 (2019).
15. Larson, A. G. et al. Liquid droplet formation by HP1α suggests a role for phase separation in heterochromatin. *Nature* **547**, 236–240 (2017).
16. Boija, A. et al. Transcription factors activate genes through the phase-separation capacity of their activation domains. *Cell* **175**, 1842–1855 (2018).
17. Cho, W. K. et al. Mediator and RNA polymerase II clusters associate in transcription-dependent condensates. *Science* **361**, 412–415 (2018).
18. Chong, S. et al. Imaging dynamic and selective low-complexity domain interactions that control gene transcription. *Science* **361**, eaar2555 (2018).
19. Sabari, B. R. et al. Coactivator condensation at super-enhancers links phase separation and gene control. *Science* **361**, eaar3958 (2018).
20. Turco, E., Gallego, L. D., Schneider, M. & Köhler, A. Monoubiquitination of histone H2B is intrinsic to the Bre1 RING domain–Rad6 interaction and augmented by a second Rad6-binding site on Bre1. *J. Biol. Chem.* **290**, 5298–5310 (2015).
21. Wei, M. T. et al. Phase behaviour of disordered proteins underlying low density and high permeability of liquid organelles. *Nat. Chem.* **9**, 1118–1125 (2017).
22. Wang, J. et al. A molecular grammar governing the driving forces for phase separation of prion-like RNA binding proteins. *Cell* **174**, 688–699 (2018).
23. Rossi, M. J., Lai, W. K. M. & Pugh, B. F. Simplified ChIP-exo assays. *Nat. Commun.* **9**, 2842 (2018).
24. Huisinga, K. L. & Pugh, B. F. A genome-wide housekeeping role for TFIID and a highly regulated stress-related role for SAGA in *Saccharomyces cerevisiae*. *Mol. Cell* **13**, 573–585 (2004).
25. Reja, R., Vinayachandran, V., Ghosh, S. & Pugh, B. F. Molecular mechanisms of ribosomal protein gene coregulation. *Genes Dev.* **29**, 1942–1954 (2015).
26. Elbaum-Garfinkle, S. et al. The disordered P granule protein LAF-1 drives phase separation into droplets with tunable viscosity and dynamics. *Proc. Natl Acad. Sci. USA* **112**, 7189–7194 (2015).
27. Lai, W. K. M. & Pugh, B. F. Understanding nucleosome dynamics and their links to gene expression and DNA replication. *Nat. Rev. Mol. Cell Biol.* **18**, 548–562 (2017).
28. DeSanto, C. et al. WAC loss-of-function mutations cause a recognisable syndrome characterised by dysmorphic features, developmental delay and hypotonia and recapitulate 10p11.23 microdeletion syndrome. *J. Med. Genet.* **52**, 754–761 (2015).
29. Vinayachandran, V. et al. Widespread and precise reprogramming of yeast protein-genome interactions in response to heat shock. *Genome Res.* **28**, 357–366 (2018).

Methods

No statistical methods were used to predetermine sample size. The experiments were not randomized and investigators were not blinded to allocation during experiments and outcome assessment.

Protein purification

All recombinant proteins used in this study were expressed as published[20]. Proteins were cloned by PCR-based methods from genomic DNA. Mutations and deletions were generated by PCR-based methods and confirmed by sequencing. StrepII–Bre1 constructs were purified on a 5-ml Strep-Tactin Superflow column (GE Healthcare) as previously described[20]. 6×His–Lge1–StrepII constructs were purified on Ni-NTA Sepharose 6 FastFlow beads (GE Healthcare) coated with Co^{2+} ($Co(NO_3)_2$) at room temperature. Lysis was performed in a buffer containing 10 mM Tris, pH 8.0, 100 mM NaH_2PO_4, 1 mM TCEP and 8 M urea, followed by a washing step in washing buffer (10 mM Tris, pH 7.4, 100 mM NaH_2PO_4, 1 mM TCEP and 8 M urea). A linear gradient to the final buffer (containing 20 mM Tris, pH 7.4, 500 mM NaCl, 1 mM TCEP, 3 M urea and 10% v/v glycerol) was applied. Additional washing steps in 10 mM Tris, pH 5.9, 100 mM NaH_2PO_4, 1 mM TCEP, 3 M urea and 10% v/v glycerol, and the same buffer with pH 4.5 were performed. Proteins were eluted in 10 mM Tris, pH 7.2, 300 mM or 1M NaCl, 1 mM TCEP and 1 M imidazole, and directly used for phase separation assays in buffer containing 100 mM NaCl and 100 mM imidazol.

The Lge1 mutant with R>K substitutions (Lge1(11×R>K)) (R11K, R18K, R20K, R27K, R28K, R30K, R39K, R48K, R63K, R70K and R77K) was ordered from Invitrogen. Mutant Lge1(17×(Y>A), 1–102) (Y4A, Y9A, Y12A, Y23A, Y38A, Y39A, Y45A, Y49A, Y53A, Y68A, Y69A, Y73A, Y79A, Y80A, Y87A, Y95A and Y102A) was ordered from GenScript.

StrepII–mGFP and StrepII–mGFP–Bre1 constructs were affinity-purified on a 5-ml StrepII-Tactin Superflow column (GE Healthcare) in buffer containing 50 mM Tris, pH 7.5 and 100 mM NaCl. Proteins were eluted in 5 mM desthiobiotin and stored in 10% v/v glycerol at −80 °C. 6×His–TEV–Rad6 was cloned, expressed and purified as previously described[20]. Where stated, purified 6×His–TEV–Rad6 was labelled with Pacific Blue C5-maleimide (ThermoFisher) as follows. Rad6 was incubated with Pacific Blue dissolved in DMSO for 20 min at 4 °C, before dialysis (18 h) in 50 mM Tris, pH 8.0, 50 mM NaCl, 50 mM KCl, 10 mM $MgCl_2$ and 1 mM TCEP. Unbound fluorophore was removed by size-exclusion chromatography on a Superdex 200 16/60 column (GE Healthcare) equilibrated with the same buffer. Purified protein was stored in final 10% v/v glycerol at −80 °C.

LAF1 (residues 1–169) was amplified by PCR from a plasmid containing full-length LAF1 (obtained from the laboratory of C. Brangwynne) and fused with Lge1(CC) in an expression vector. The coding region for WAC (amino acids 1–318) was amplified by PCR from human cDNA and was fused to Lge1(CC) in an expression vector. 6×His–LAF1(1–169)–Lge1(CC)–StrepII and 6×His–WAC(1–318)–Lge1(CC)–StrepII were purified in urea as described for 6×His–Lge1–StrepII, followed by buffer exchange to 10 mM Tris, pH 8.0, 0.2 mM EDTA and 100 mM NaCl.

In vitro binding assays

Purified GST–Lge1–StrepII constructs were mixed with StrepII–Bre1 constructs in a 1:3 molar ratio in binding buffer (50 mM Tris, pH 7.5, 100 mM NaCl, 0.5 mM DTT and 50 μg of BSA) and incubated with glutathione-sepharose 4B beads (GE Healthcare) for 1 h at 4 °C. Beads were washed with binding buffer (without BSA) and bound proteins were eluted in 0.006 mg/ml glutathione. After trichloroacetic acid (TCA) precipitation, samples were analysed by SDS–PAGE (4–12% gel, MOPS buffer) and Coomassie staining, or by immunoblotting with anti-StrepII antibody (Qiagen, no. 34850).

Sucrose gradient sedimentation assay

For recombinant proteins, 30 μg of purified proteins were applied to a 12-ml 5–45% sucrose gradient in 20 mM Tris, pH 7.5, 10 mM KCl and 5 mM $MgCl_2$, and sedimented at 27,000 rpm for 15 h at 4 °C in a SW40-Ti rotor (Beckman Coulter). Gradients were fractionated in 13× 1-ml fractions by pumping Fluoriniert FC-40 (3M) into the bottom of the gradient while collecting the fractions from the top using a Brandel gradient fractionation system. Fractions were precipitated with TCA, washed twice with acetone and analysed by SDS–PAGE (4–12% gel, MOPS buffer) and Coomassie staining. For the Bre1–Lge1 complex, 30 μg of each protein was incubated for 30 min at 4 °C before loading the gradient. For the fractionation of cell lysates, cells were grown exponentially, collected and resuspended in buffer 20 mM Tris, pH 7.5, 10 mM KCl, 5 mM $MgCl_2$ and 1mM DTT. Cells were lysed by vortexing with glass beads. Lysates were clarified by centrifugation at 14,000 rpm for 5 min at 4 °C. The total protein amount was quantified by Bradford assay (BioRad) and about 350 μg were layered onto 12-ml 5–45% sucrose gradients, sedimented and fractionated as described. Ribosomal profiles were recorded as a reference by monitoring absorbance at an optical density of 254 nm. After TCA precipitation, samples were analysed by immunoblotting with anti-HA (Sigma, no. 11666606001), anti-protein A (Sigma, P3775) and anti-RNA polymerase II (Rpb1; Covance, no. MMS-126R). Input samples were also analysed with anti-Pgk1 (Abcam, no. Ab113687) as loading control.

Lge1 phase-separation assay

Purified Lge1 proteins in a total volume of 20 μl were placed in a 16-well glass-bottom ChamberSLIP slide (Grace, BioLabs). Slides were pretreated with 1% (w/v) pluronic F-127 (PF127, Sigma Aldrich) and washed twice with buffer 25 mM Tris, pH 7.5, 50 mM NaCl, 1 mM DTT and 10% v/v PEG_{6000}. A final buffer composition of 20 mM Tris, pH 7.5, 100 mM NaCl and 1 mM DTT was used for the dilution of the proteins. 1,6-Hexanediol was used at a final concentration of 5 or 10% (w/v) (Sigma Aldrich, no. 240117). Final protein concentration and incubation times for the experiments are detailed in the corresponding figure legends. DIC imaging was performed on a temperature-controlled DeltaVision Elite microscope (GE Healthcare) at 20 °C. Images were acquired with a 60× oil immersion objective coupled to a CoolSNAP HQ2 charge-coupled device camera (Photometrics). Image deconvolution was carried out using the SoftWoRx software (GE Healthcare). Images were processed with ImageJ. Quantification of the total area (in μm^2) of randomly selected condensates was analysed in ImageJ. One hundred condensates were measured for each condition. Statistical significance was evaluated by two-sided Mann–Whitney test using the GraphPad Prism software.

Phase-separation assays with fluorescent proteins

6×His–Lge1 constructs (1.5 μM) were mixed with fluorescent proteins (StrepII–mGFP, StrepII–mGFP–Bre1 constructs, 6×His–Rad6*, 1×NCP*, 16×NCP* or 16× 601 Widom DNA*) in pretreated and washed 16-well glass-bottom ChamberSLIP slides at 20 °C in a final volume of 20 μl. A buffer composition of 20 mM Tris, pH 7.5, 100 mM NaCl (50 mM NaCl for NCP assays, 1 mM DTT) was used. The concentration of fluorescent constructs and incubation times for each experiment are described in the figure legends. Samples were either directly imaged or, where stated, the experiment was followed over time. Image acquisitions parameters for images of the same experiment are always identical. Images were processed with ImageJ. The peak fluorescence intensity of the condensates with fluorescent-tagged proteins was measured by drawing a line across at least 25 randomly selected condensates. Line-scan graphs were generated in ImageJ. For Bre1 shell formation, 6×His–Lge1–StrepII (1.5 μM) and StrepII–mGFP–Bre1 constructs (3 μM) were preincubated in pretreated and washed 16-well glass-bottom ChamberSLIP slides for 15 min at 20 °C. Fluorescent shell thickness was quantified as follows: on the line-scan graphs, background signal inside the condensate was subtracted from the maximum fluorescent intensity of the peaks. Final thickness values correspond to the length (in μm) of the resulting baseline of the peaks. Statistical significance was evaluated by two-sided *t*-test analysis with the GraphPad Prism

software. Image acquisitions parameters for images of the same experiment are always identical. The number of analysed condensates (*n*) is indicated in each figure.

Turbidity assay

Two hundred microlitres of each purified protein (buffer: 20 mM Tris pH 7.5, 100 mM NaCl, 100 mM imizadol and 1mM DTT) were measured at 450 nm on a bottom-clear 96-well plate (Greiner Bio-One) in a Victor Nivo plate reader (Perkin Elmer).

Dextran experiments

Tetramethylrhodamine isothiocyanate (TRITC)-labelled dextran of an average molecular weight of 35–45 (TdB Consultancy AB, no. TD40), 65–85 (Sigma Aldrich, no. T1162), 155 (Sigma Aldrich, no. T1287) or 2,000 kDa (ThermoFisher, no. D7139) was dissolved in 20 mM Tris, pH 7.5 and 100 mM NaCl. The dextrans with a final concentration of 0.05 mg/ml were added to samples containing 6×His–Lge1–StrepII (1.5 μM), 6×His–Lge1(IDR)–StrepII (1.5 μM) or 6×His–Lge1–StrepII (1.5 μM) with a preformed Bre1 shell, each in 20 μl. The final buffer composition was 20 mM Tris, pH 7.5, 100 mM NaCl and 1 mM DTT. Samples were incubated for 15 min at 20 °C in pretreated and washed 16-well glass-bottom ChamberSLIP slides. Samples were imaged simultaneously in DIC, FITC and TRITC channels. Partition ratios were calculated on the basis of the background-corrected TRITC fluorescent intensity inside and outside of a condensate quantified with ImageJ. A total of 60 randomly selected condensates were considered for each condition.

Dynamic light scattering

The hydrodynamic radius and molecular weight of recombinant proteins were measured at 20 °C in a ProteinSolution DynaPro-99-E50 DLS module (Wyatt).

Nucleosome array reconstitution and labelling

Mononucleosomes (1×NCP) were reconstituted as previously described[20]. For chromatin reconstitution (16×NCP array), DNA consisting of 16 repeats of the 167-bp 601 Widom sequence was generated from a pUC19-16x601Widom plasmid (a gift from T. Richmond) by triple digestion with PstI, BamHI and DpnI for 20 h at 37 °C. Forty per cent (w/v) PEG_{6000}–NaCl precipitation was used to separate the digested plasmid from array DNA, which was precipitated in EtOH and finally resuspended in 10 mM Tris, pH 8.0, 0.1 mM EDTA. Chromatin was reconstituted with the purified histone octamer and the 16× 601Widom DNA, in a histone octamer:DNA ratio (mg:mg) of 10, followed by KCl dialysis and Mono Q ion-exchange chromatography, as previously described[20]. Samples were then dialysed in 20 mM Tris, pH 7.5, 1 mM EDTA and 1 mM DTT for 18 h. Histone stoichiometry was assessed by SDS–PAGE and Coomassie staining. Nucleosome occupancy was tested by ScaI digestion and electrophoresis in a 1% agarose gel. Both 1× and 16×NCPs were labelled with Hoechst 33258 (Sigma Aldrich) (1 μg/ml) by incubation for 30 min at 4 °C, followed by dialysis in 20 mM Tris, pH 7.5 and 1 mM DTT for 20 h at 4 °C. Labelled 1×NCP* and 16×NCP* were stored at 4 °C. The 16× DNA was labelled with Hoechst (final 1 μg/ml) by incubation for 18h at 4 °C and stored at −20 °C.

Negative-stain electron microscopy

The quality of the 16×NCP array was also assessed by electron microscopy. Samples in 20 mM Tris, pH 7.5, were adsorbed on glow-discharged electron microscopy grids coated with carbon film and stained with 2% uranyl acetate (Merck), pH 4.0. Electron microscopy samples were examined on a FEI Morgagni 268D TEM operated at 80 kV. Digital images were acquired using an 11-megapixel Morada charge-coupled device camera from Olympus-SIS.

NCP ubiquitination assay

For the in vitro H2B ubiquitination assay with Lge1–Bre1 condensates, 6×His–Lge1 (0.5 μM), Bre1 (3 μM), 155-kDa dextran (0.65 μM), 1×NCP (1 μM), E1 (100 nM), Rad6 (3 μM) and ubiquitin (36 μM) were incubated in 20 mM Tris HCl, pH 7.5, 75 mM NaCl and 1 mM DTT. Reactions were started by addition of ATP (3 mM) and carried out for 12 min with shaking at 300 rpm at 30 °C. Reactions were stopped at the indicated time points by adding 4×SDS loading buffer and boiling (5 min at 95 °C), and analysed by SDS–PAGE and immunoblotting with anti-Flag antibody (Sigma Aldrich, no. F1804). Each experiment was reproduced at least three times. Condensate formation was verified by microscopy.

Yeast strains and tandem affinity purification

All yeast strains used in this study are listed in the Supplementary Table. Genes in yeast were tagged or deleted by a standard one-step PCR-based technique. Microbiological techniques followed standard procedures. Cells were grown in yeast-extract peptone dextrose (YPD) or, when transformed with plasmids, in selective synthetic dextrose complete (SDC) drop-out medium. Tandem affinity purifications (TAPs) from yeast were performed according to standard protocols.

Live-cell imaging of yeast

Exponentially growing cells were immobilized on microscope slides with agarose pads and imaged on a DeltaVision Elite microscope (GE Healthcare). Images were acquired with a 60× oil immersion objective and recorded with a CoolSNAP HQ2 charge-coupled device camera (Photometrics). Deconvolution was carried out using the SoftWoRx software (GE Healthcare). Images were processed with ImageJ. Cell contours were marked with a dashed white line based on bright-field imaging. For the 1,6-hexanediol experiment, cells were incubated with 10% (w/v) 1,6-hexanediol and 10 μg/ml digitonin (or only 10 μg/ml digitonin as control) for 10 min at 30 °C before imaging.

To record Lge1–Bre1 BiFC intensity in a histogram, a contour was drawn around the nucleus and fluorescence intensities were analysed using the ImageJ 'Histogram' function. For quantification of the mean nuclear BiFC intensity, mean fluorescence-intensity values from the histograms were used. The coefficient of variation values represent the ratio of the standard deviation to the mean (both values determined from histogram data).

To quantify the corrected total-cell fluorescence, a contour was drawn around a cell and the integrated density was measured using ImageJ. To quantify the fluorescence intensity of Lge1 constructs in cells, a line was drawn across the nucleus and line-scan graphs were obtained by ImageJ. The number of analysed cells (*n*) is indicated in each figure. Statistical significance was evaluated by two-sided Mann–Whitney test using the GraphPad Prism software.

Yeast genetic interaction analysis

Double mutant strains containing a wild-type *LGE1* cover plasmid (Ura marker) were cotransformed with plasmids carrying *LGE1*, *WAC* or *LAF1* (His marker). Growth was followed on SDC−His (loading control) and on SDC + 5-FOA plates to shuffle out the Ura cover plasmid. Cells were spotted in tenfold serial dilutions and incubated for 2 (SDC−His) or 3 days (5-FOA) at 30 °C.

Protein expression levels in whole-cell extracts

Yeast lysates were prepared, normalized for protein concentration and analysed by western blotting according to standard procedures. Antibodies were used according to the manufacturer's instructions: anti-mCherry (Abcam, no. Ab125096), anti-protein A (Sigma, no. P3775) and anti-Pgk1 (Abcam, no. Ab113687).

Global levels of H2B ubiquitination

LGE1 was deleted in a strain carrying Flag-tagged H2B. This strain was then transformed with plasmid-based versions of wild-type *LGE1-mCherry*, empty vector and the indicated mutants, all of which were expressed from the endogenous *LGE1* promoter. Cell lysates were subjected to anti-Flag immunoprecipitation (M2 beads, Sigma, no. A2220).

Recovered proteins were analysed by SDS–PAGE and immunoblotting with an anti-Flag antibody (Sigma, no. A8592).

ChIP-exo version 5.0

The antibodies rabbit monoclonal antibody against ubiquityl-H2B(Lys120) (Cell Signaling, no. 5546), and rabbit polyclonal antibody against H2B (Abcam, no. ab1790)[10] were used to probe for the specific ubiquitin mark and the underlying nucleosome, respectively. For TAP-tagged strains, rabbit IgG (Sigma) conjugated to Dynabeads was used. The protein A module of the TAP tag was the target.

Indicated mutant strains were grown in selective CSM − His medium, and the wild-type strains were grown in CSM + all medium to an optical density at 600 nm of 0.8 at 25 °C. Cells were crosslinked with formaldehyde for 15 min at room temperature at a final concentration of 1%, and quenched with glycine for 5 min at a final concentration of 125 mM. Cells were then centrifuged and washed with ST buffer (10 mM Tris-HCl, pH 7.5 and 100 mM NaCl) at 4 °C. Supernatant was removed and cell pellets were flash-frozen and stored at −80 °C until further use. Fifty-millilitre culture aliquots were resuspended and lysed in 750 µl FA lysis buffer (50 mM Hepes-KOH, pH 7.5, 150 mM NaCl, 2 mM EDTA, 1% Triton X-100, 0.1% sodium deoxycholate and complete protease inhibitor (Roche)) along with 1 ml of 0.5-mm zirconia beads by bead- beating for 4 cycles of 3-min on/7-min off in a Mini-Beadbeater-96 machine (Biospec). Whole-cell lysates were transferred to microcentrifuge tubes and spun at 14,000 rpm for 3 min at 4 °C to obtain chromatin pellets. The supernatant (representing the cytoplasmic fraction) was discarded and the chromatin pellet was resuspended in 200 µl FA lysis buffer supplemented with 0.1% SDS and 75 µl 0.1-mm zirconia beads. The resuspended chromatin sample was sonicated in a Bioruptor (Diagenode) for 4 cycles of 30-s on/30-s off. The tubes were centrifuged at 14,000 rpm for 10 min to collect the sonicated chromatin fraction. Chromatin obtained from 50 ml culture was used for a single ChIP experiment. A Reb1–TAP-tagged strain (Open Biosystems) was used a positive control to determine the success of the ChIP experiments. Appropriate antibodies (5 µg per ChIP) were conjugated to protein A magnetic sepharose resin (GE Healthcare Life Sciences) at 4 °C for 6–8 h. Unconjugated antibody was removed and chromatin sample was added and incubated at 4 °C overnight to allow for immunoprecipitation. All steps after immunoprecipitation were essentially performed as previously described[23]. Prepared libraries were gel-purified and were sequenced in paired-end mode with a NextSeq 500. Sequence reads were aligned to the *S. cerevisiae* genome (sacCer3) using bwa-mem (v.0.7.9a)[30]. Aligned reads obtained were filtered to remove PCR duplicates and any non-unique alignments. ChIP-exo read-1 5′ ends had their coordinates shifted by 6 bp in the 3′ direction, to reflect the offset of exonuclease stops and the site of crosslinking. At least two biological replicates were performed for ChIP-exo experiments, unless otherwise indicated. These methods and information on data analysis and scripts can be accessed online (https://github.com/CEGRcode/Gallego_2019).

Structural bioinformatics

Charge profiles for Lge1 and WAC protein sequences were generated using scripts written for this purpose, with a window of 10 amino acid residues used for smoothing. Hydrophobicity profiles for Lge1 and WAC protein sequences were generated and superimposed using VOLPES server[31] (http://coil.msp.univie.ac.at/app.html).

Statistics and reproducibility

The experiments in Fig. 1b, c, d were performed 4, 2 and 3 times, respectively; the quantification of condensates is included in Extended Data Fig. 1j. Experiments in Fig. 2a, b were performed 3 times. Experiments in Fig. 3a, b, d were performed 2, 3 and 3 times, respectively; quantification is included in Extended Data Fig. 7i. The experiment shown in Fig. 4a is representative of 2 replicates. The experiment in Fig. 4b was performed 6 times.

Experiments were performed the following numbers of times: 2 (Extended Data Figs. 1i, 3c, 4a, g, 6a, d, g, 10a), 3 (Extended Data Figs. 1k, 3a, b, 4b, e, 6b, 7a–d, f–h, 8a, 10d, f), 4 (Extended Data Figs. 1d, e, 5a, c, 6e, f, h, 10c, e, h, i), 5 (Extended Data Fig. 1b, c) or 6 (Extended Data Fig. 1a)

In Extended Data Fig. 1f, Lge1(1–80) was purified 2 times, Lge1(1–80, CC) was purified 3 times and the other constructs were purified >10 times. The time lapse for droplet fusion in Extended Data Fig. 1h was recorded in 2 independent experiments. In Extended Data Fig. 2b, sequences were analysed as described in the figure legend and in Supplementary Methods. The time lapse in Extended Data Fig. 3e was recorded in 2 independent experiments. The experiment in Extended Data Fig. 4c was performed 10 times as control. In Extended Data Fig. 6c, the purification of mononucleosomes (1×NCP) was performed 29 times and the purification of 16× NCP was performed 11 times. The experiment shown in Extended Data Fig. 6i corresponds to Fig. 3a (performed 2 times). The numbers of randomly selected cells in Extended Data Fig. 7e were: Lge1 = 22; Lge1 *bre1Δ* = 33; Lge1(Δ1–40) = 24; Lge1(Δ1–80) = 25; Lge1(1–80, CC) = 25; Lge1(IDR) = 28; NLS–Lge1(IDR) = 21; and Lge1(CC) = 27. The quantification for Extended Data Fig. 7h is shown in Extended Data Fig. 7i. Data shown in Extended Data Fig. 8b, c are representative of 2 replicates. In Extended Data Fig. 9a, the dot spots were performed 4 times, and whole cell extracts and western blotting were performed 2 times. Data shown in Extended Data Fig. 9b–e are based on 2 replicates.

Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

Data availability

ChIP-exo sequencing files can be accessed at NCBI Gene Expression Omnibus (https://www.ncbi.nlm.nih.gov/geo/), with the accession number GSE131639. All source data (that is, uncropped gels and western blots) associated with the paper are provided in Supplementary Figs. 1, 2.

Code availability

Data analysis software, scripts and parameters used to analyse ChIP-exo datasets can be accessed via https://github.com/CEGRcode/Gallego_2019.

30. Li, H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. Preprint at https://arxiv.org/abs/1303.3997 (2013).
31. Bartonek, L. & Zagrovic, B. VOLPES: an interactive web-based tool for visualizing and comparing physicochemical properties of biological sequences. *Nucleic Acids Res.* **47**, W632–W635 (2019).
32. Armstrong, J. K., Wenby, R. B., Meiselman, H. J. & Fisher, T. C. The hydrodynamic radii of macromolecules and their effect on red blood cell aggregation. *Biophys. J.* **87**, 4259–4270 (2004).
33. Yassour, M. et al. *Ab initio* construction of a eukaryotic transcriptome by massively parallel mRNA sequencing. *Proc. Natl Acad. Sci. USA* **106**, 3264–3269 (2009).

Acknowledgements We thank B. Zagrovic and A. Polyansky for critical insights into the multivalency of Lge1 LLPS, and Lge1 and WAC sequence analyses; O. R. Abdi for technical support; and G. Warren and D. Gerlich for discussions. A.K. was funded in part by a NOMIS Pioneering Research Grant, L.D.G. by a L'Oréal-UNESCO-OeAW Austria Fellowship and A.R. by an OeAW DOC Fellowship. B.F.P. and C.M. were funded by the National Institutes of Health grant HG004160.

Author contributions L.D.G. and M.S. performed all biochemical and genetic experiments with support from R.M.G.C. C.M. conducted ChIP-exo and analysed the data with B.F.P. T.S. contributed to sucrose gradient analyses. A.R. performed all cell biology experiments. A.K. designed the study and wrote the paper with input from all authors.

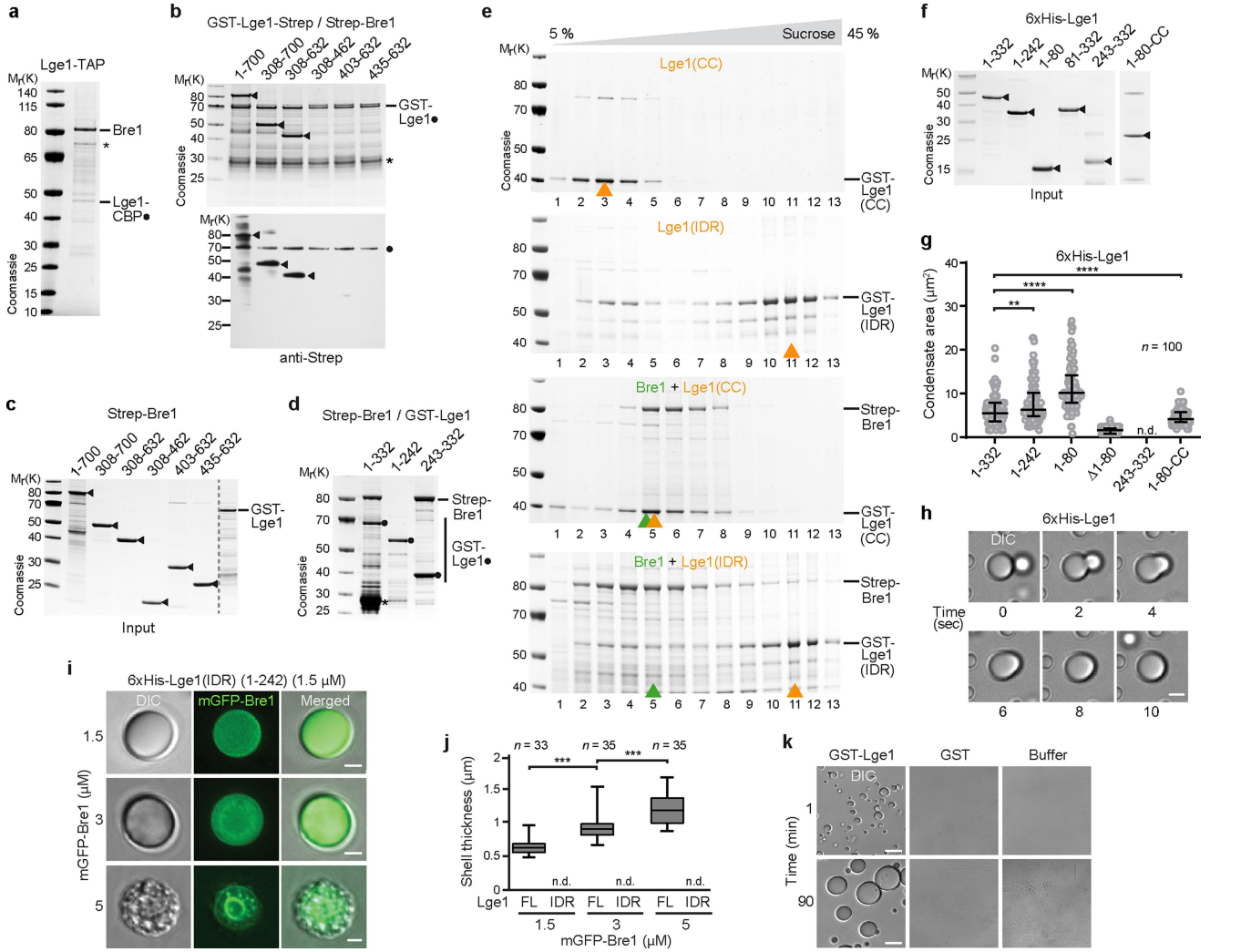
Competing interests B.F.P. has a financial interest in Peconic LLC, which uses the ChIP-exo technology implemented in this study and could potentially benefit from the outcomes of this research. The remaining authors declare no competing interests.

Additional information

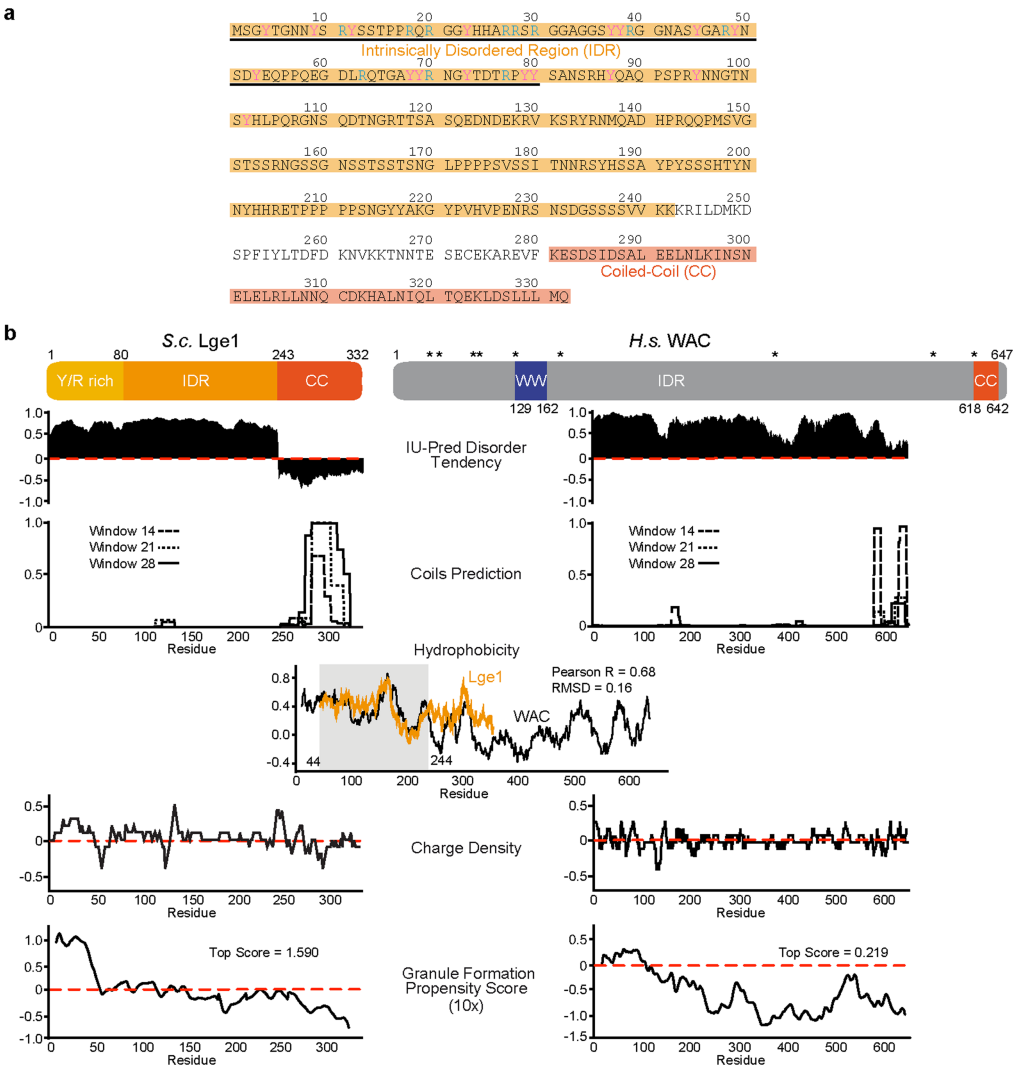
Supplementary information is available for this paper at https://doi.org/10.1038/s41586-020-2097-z.

Correspondence and requests for materials should be addressed to A.K.

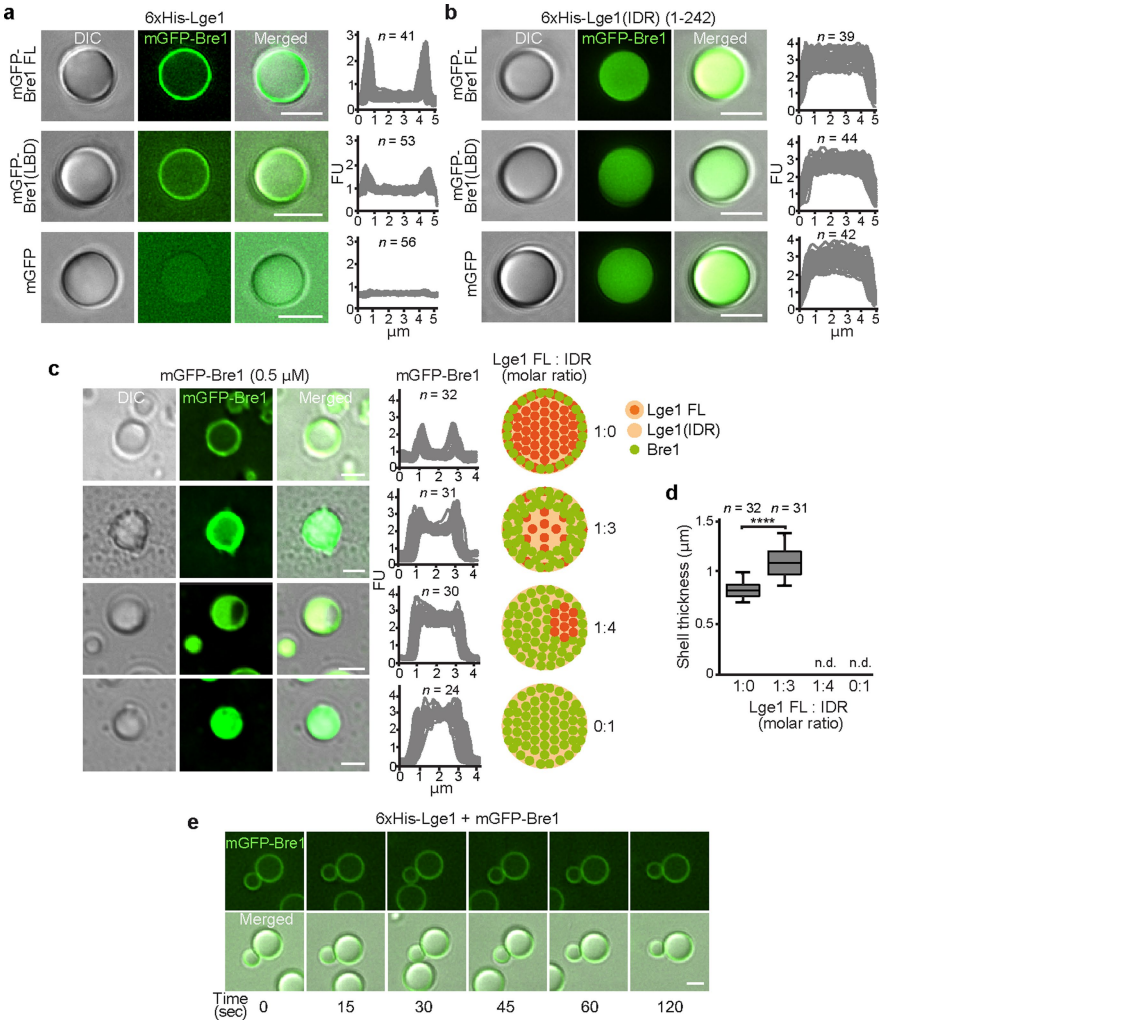
Peer review information *Nature* thanks Fred van Leeuwen, Tanja Mittag and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.



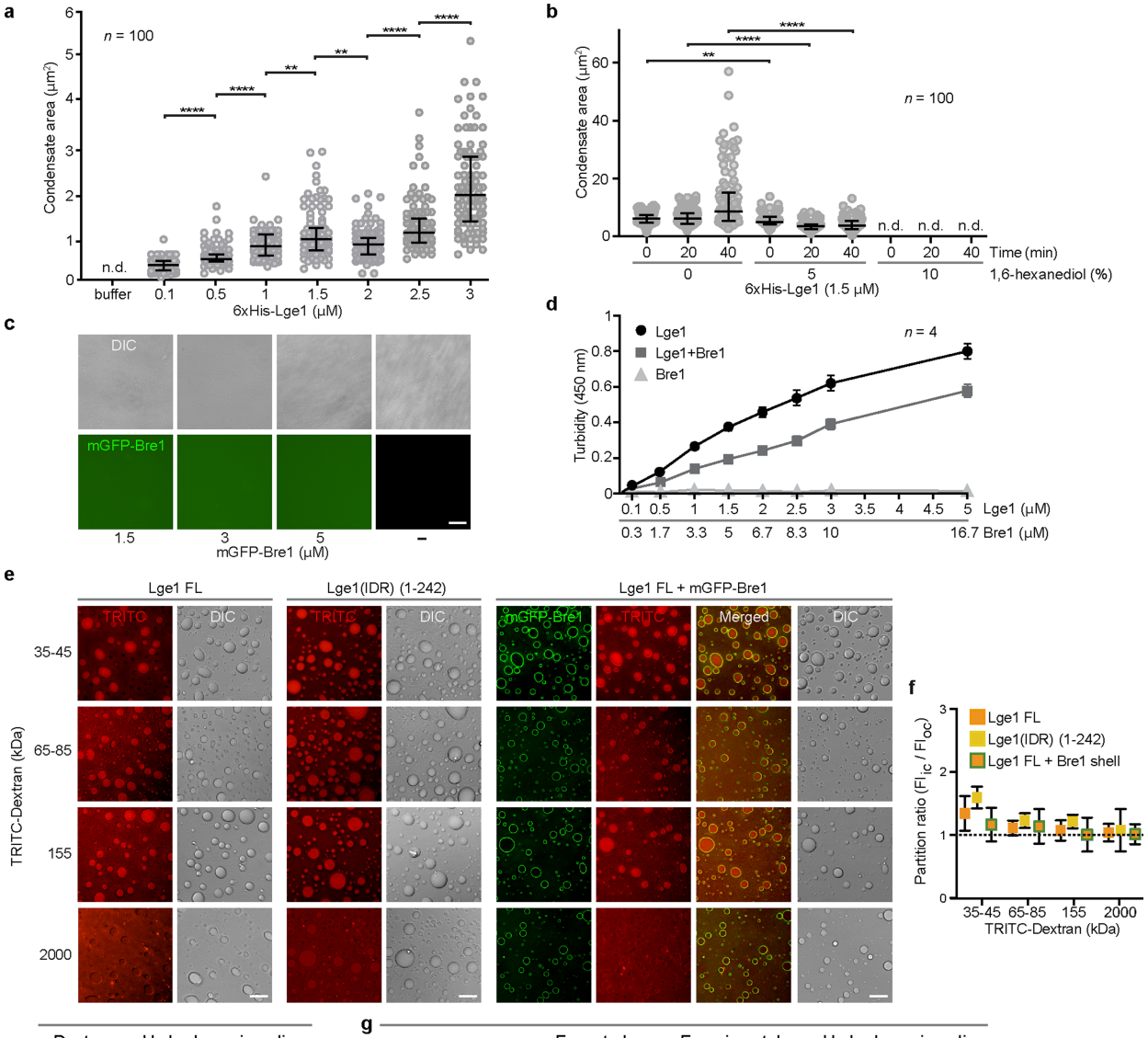
Extended Data Fig. 1 | Lge1 interaction with Bre1 and LLPS. a, Bre1 copurifies with Lge1. Lge1–TAP was tandem-affinity-purified from yeast. Bre1 and CBP (calmodulin binding peptide)-tagged Lge1 (Lge1–CBP; a remnant after protease cleavage) were confirmed by mass spectrometry. Asterisk indicates a Bre1 degradation product. **b**, Bre1 interacts with Lge1 in vitro. Binding assay using recombinant GST–Lge1–Strep as bait (black dot) with Strep–Bre1 constructs (arrowheads) (1:3 molar ratio). Proteins were analysed by Coomassie staining and anti-Strep immunoblotting. **c**, Input proteins for **b**. **d**, Same setup as in **b**. Black dots indicate immobilized Lge1 constructs; asterisk indicates a degradation product. **e**, Sucrose gradient sedimentation assays (5–45%) of recombinant Strep–Bre1 and GST–Lge1 proteins. Peak fractions are highlighted (arrowheads). **f**, Input protein for 6×His–Lge1 constructs used in phase-separation assays. **g**, Quantification of condensate sizes in Fig. 1c. *n*, number of condensates. Dot plots with median and interquartile range. $^{**}P = 0.0189$, $^{****}P < 0.0001$, determined by two-sided Mann–Whitney test. n.d., not determinable. **h**, Analysis of condensate growth using DIC imaging (6×His–Lge1, 5 μM) (Supplementary Video 1). Scale bar, 2 μm. **i**, Increasing amounts of mGFP–Bre1 were added to preformed 6×His–Lge1(IDR) condensates (which lack the coiled-coil domain of Lge1). The unrestricted diffusion of Bre1 (5 μM) into Lge1(IDR) condensates caused their interior to collapse into coarse aggregates. Scale bar, 2 μm. **j**, Quantification of mGFP–Bre1 shell thickness in **i** and Fig. 1d. Box-and-whisker plot shows median, interquartile range, and minimum and maximum values. $^{***}P$ value < 0.001, determined with two-sided *t*-test (1.5 μm–3 μm, $t = 7.6$, degrees of freedom = 67; 3 μm–5 μm, $t = 14$, degrees of freedom = 69). *n*, number of condensates; n.d. not determinable. **k**, GST–Lge1 used in Fig. 1b phase-separates in vitro. Recombinant GST–Lge1 (7 μM), GST (7 μM) and buffer-only (20 mM Tris pH 7.5, 10 mM KCl and 5 mM $MgCl_2$) were visualized by DIC microscopy at 20 °C for the indicated times (min). Scale bar, 10 μm. See Supplementary Fig. 2 for uncropped gels and western blots.



Extended Data Fig. 2 | Lge1 structural properties, and comparison with WAC. a, Amino acid sequence of *S. cerevisiae* Lge1. The IDR (amino acids 1–242) is highlighted in light orange; the predicted coiled-coil domain (amino acids 281–332) is highlighted in in dark orange; and the Y/R-rich sticker region (amino acids 1–80) is underlined. The mutated Y and R residues are labelled in magenta and blue, respectively. **b**, Comparative sequence analyses of yeast Lge1 (left) and human WAC (WW-domain-containing adaptor protein with coiled-coil) (right). Cartoon shows predicted Lge1 and WAC domain organization; boundaries are drawn to scale. WW, protein–protein interaction domain with two tryptophans (W). Asterisks indicate residues the mutation of which has previously been implicated in the pathogenesis of DeSanto–Shinawi syndrome[28]. This neurodevelopmental disorder causes a developmental delay and dysmorphic facial features. It is caused by mutations that are predicted to truncate the IDR of WAC, and therefore disrupt the WAC interaction with RNF20/RNF40 (that is, nonsense and frameshift mutations leading to nonsense-mediated decay or protein truncation)[28]. For disorder prediction, disorder scores were calculated with the IUPred algorithm. Both Lge1 and WAC show extensive intrinsic disorder. For coiled-coil domain predictions, the COILS software was used, showing putative coiled-coil elements at the C termini of both proteins. For hydrophobicity analysis, the VOLPES web server was used to plot hydrophobicity profiles of protein sequences[31] (further details are provided in the Methods). Lge1 and WAC display a similar hydrophobicity pattern in their N-terminal regions. The region with the best match is highlighted with a grey rectangle. Similarity in this region is indicated by Pearson correlation coefficient ($R = 0.68$) and root-mean-squared deviation (r.m.s.d. = 0.16). In their charge distribution profile, both Lge1 and WAC exhibit alternating blocks of negative and positive charge. The sequence charge density was calculated using a custom-made script. A window of ten residues was used for smoothing. For condensate formation, the catGRANULE algorithm was used to predict the propensity for condensate formation. Top scores are indicated. The high-scoring sequence in Lge1 corresponds approximately to the Y/R-rich region.



Extended Data Fig. 3 | Mechanism of Bre1 shell formation.
a, **b**, Reconstitution of condensates with core–shell architecture. Recombinant mGFP-tagged proteins (1.5 μM) were added to preformed 6×His–Lge1 condensates or 6×His–Lge1(IDR). Samples were incubated for 15 min before imaging by DIC and fluorescence microscopy. Scale bar, 2 μm. **c**, Reconstitution of hybrid condensates, with varying ratios of 6×His–Lge1:6×His–Lge1(IDR) show a differential partitioning of mGFP–Bre1 into the core. Proteins were mixed at the indicated molar ratios and incubated for 15 min at 20 °C. mGFP–Bre1 (0.5 μM) was added to the preformed condensates and incubated for 15 min before imaging by DIC and fluorescent microscopy. Fluorescent intensities were quantified across single condensates. Cartoons indicate putative assembly state of hybrid condensates, with a deterioration of the core–shell structure upon reduction of available Lge1 coiled-coil domains. *n*, number of condensates. Scale bar, 2 μm. **d**, Quantification of mGFP–Bre1 shell thickness in **c**. Box-and-whisker plot shows median, interquartile range, and minimum and maximum values. ****$P < 0.0001$, determined with two-sided *t*-test ($t = 9.4$, degrees of freedom = 62). *n*, number of condensates; n.d. not determinable. **e**, Analysis of condensate fusion. 6×His–Lge1 condensates (1.5 μM) with an mGFP–Bre1 shell (1.5 μM) were followed over time by microscopy. Condensates collide but do not fuse (Supplementary Video 2). Compare to Extended Data Fig. 1h for fusion dynamics. Scale bar, 2 μm.

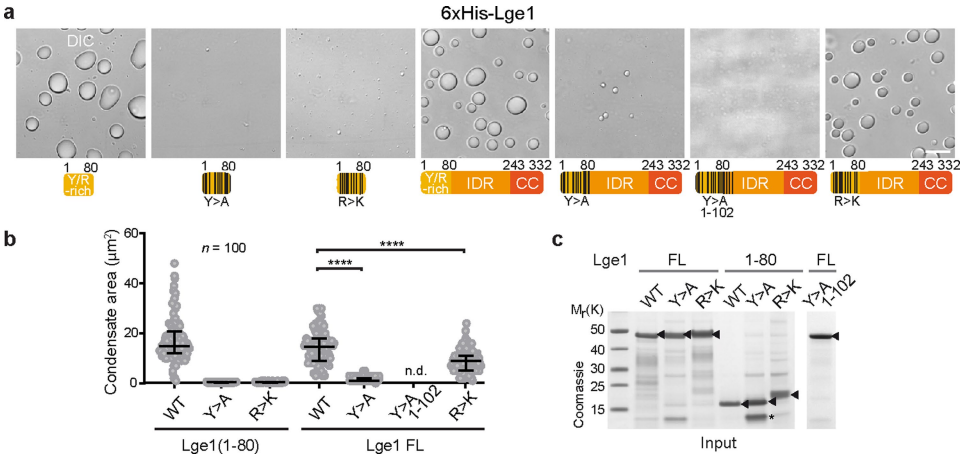


Dextran (kDa)	Hydrodynamic radius (R_h, nm)
35-45	4.78
65-85	6.49
110	7.82
155	n.d.
2000	26.89

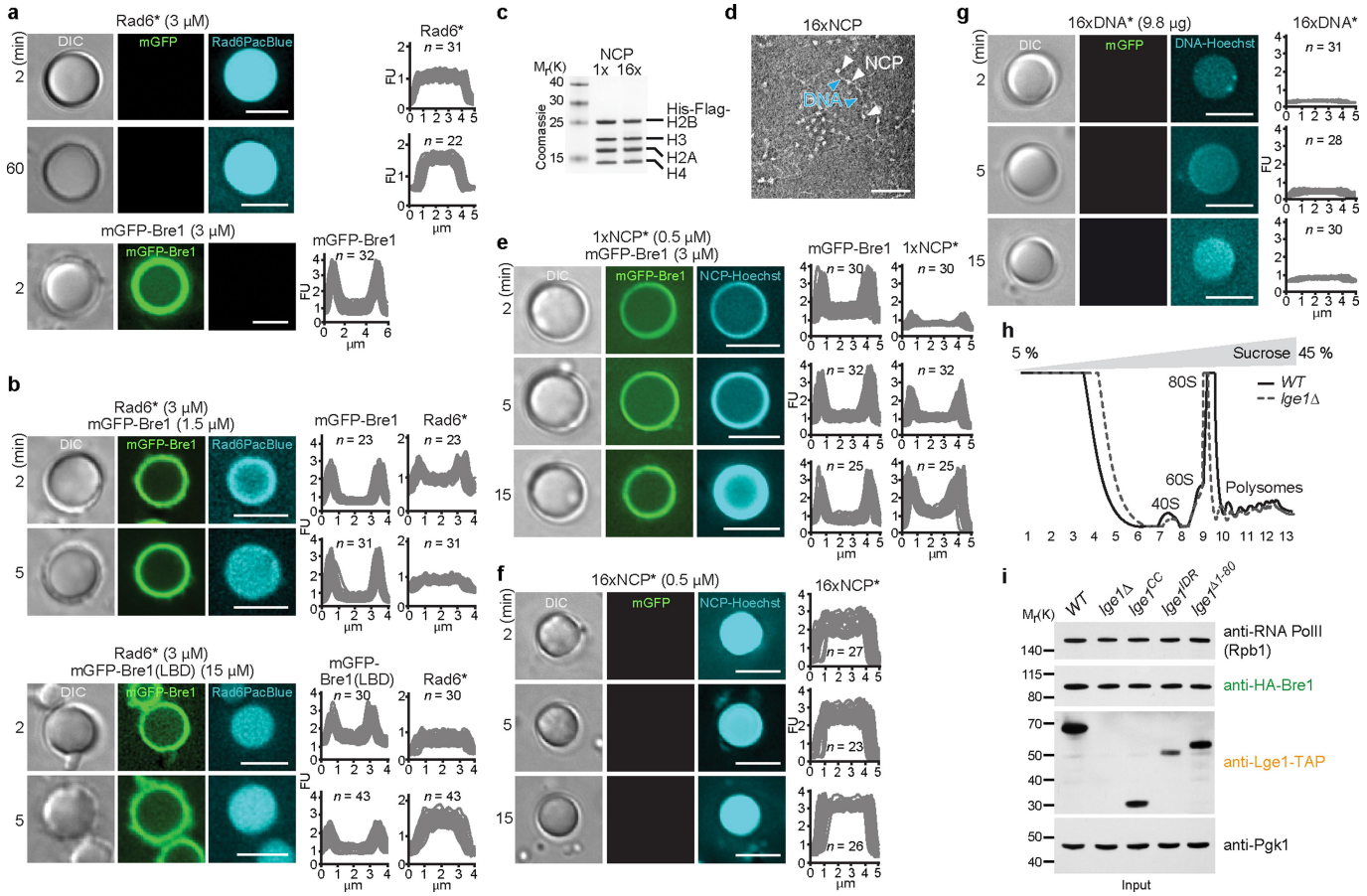
g

Protein	Expected (kDa)	Experimental (kDa)	Hydrodynamic radius (R_h, nm)
mGFP	26.80	32.50	2.60
mGFP-Bre1	108.90	n.d.	n.d.
mGFP-Bre1(LBD)	65.20	106.00	4.30
1xNCP*	220.00	241.50	6.20
16xNCP*	3520.00	7446.70	26.90

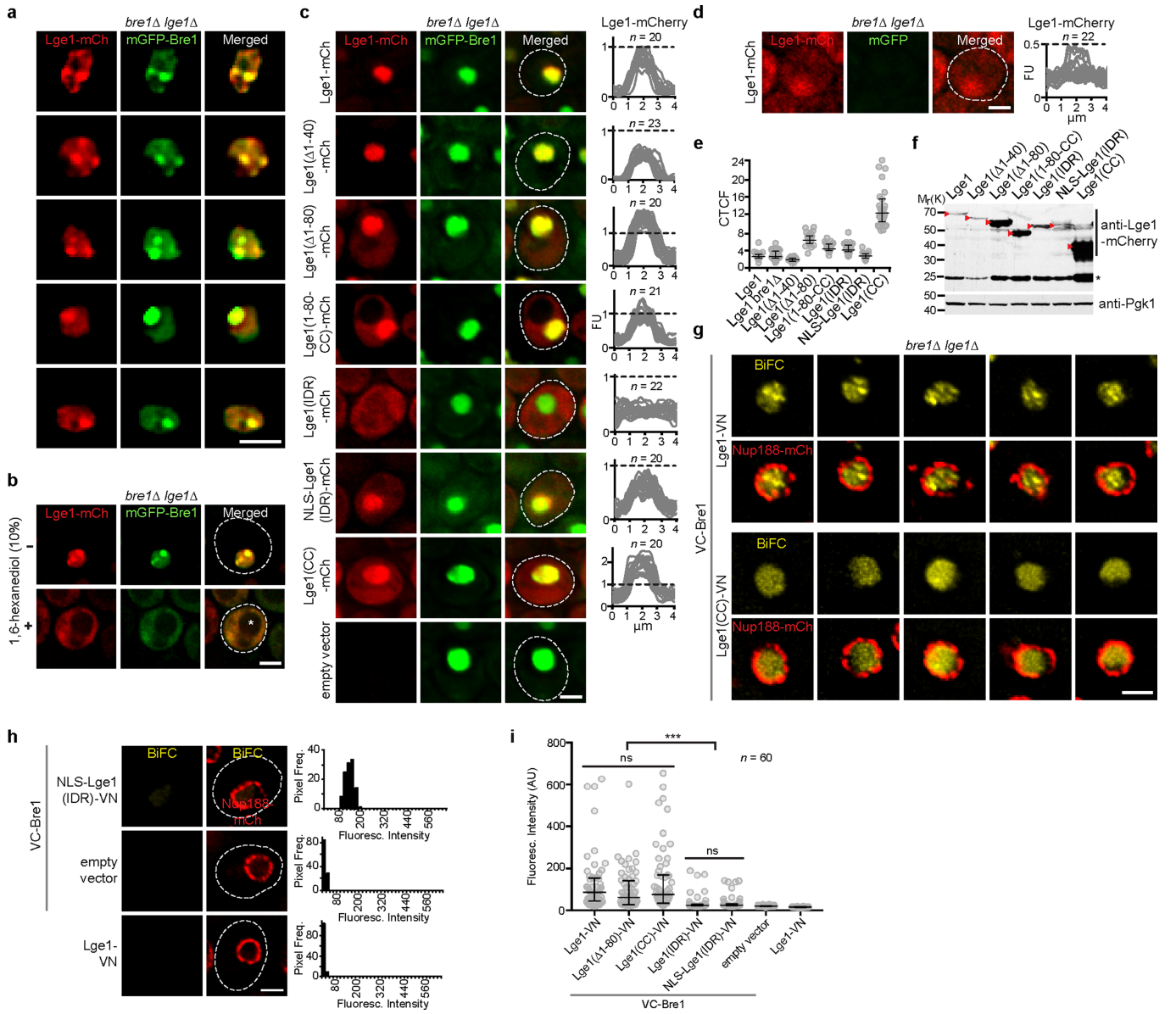
Extended Data Fig. 4 | Material properties of Lge1 condensates. **a**, Quantification of 6×His–Lge1 condensate sizes at different protein concentrations after 5 min of incubation at 20 °C. Quantification was done using ImageJ. *n*, number of condensates. Dot plots show median and interquartile range. $^{**}P$ (for 1 versus 1.5) = 0.0046, $^{**}P$ (for 1.5 versus 2) = 0.0029, $^{****}P < 0.0001$, determined by two-sided Mann–Whitney test. n.d. not determinable. **b**, Quantification of 6×His–Lge1 condensate size in the presence of 1,6-hexanediol indicates an inhibition of LLPS. Concentrated Lge1 protein was diluted to 1.5 µM in buffer with 1,6-hexanediol (%, w/v) and incubated for 15 min before imaging. *n*, number of condensates. Dot plots show median and interquartile range. $^{**}P = 0.0032$, $^{****}P < 0.0001$, determined by two-sided Mann–Whitney test. n.d. not determinable. **c**, Strep–mGFP–Bre1 does not phase-separate under the conditions we tested. Concentrated proteins were diluted and incubated at 20 °C for 5 min before DIC microscopy. Scale bar, 10 µm. **d**, Turbidity measurements of 6×His–Lge1 at 450 nm, with or without Strep–Bre1. Proteins were mixed at the indicated molar ratios. LLPS of Lge1 occurred already at 0.1 µM, Strep–Bre1 shows no LLPS under the conditions we tested. Mean and s.d. are indicated. **e**, Condensates of 6×His–Lge1, 6×His–Lge1(IDR) or 6×His–Lge1 with an mGFP–Bre1 shell were incubated with TRITC-labelled dextran of different sizes (final dextran concentration 0.05 mg ml^{-1}) for 15 min at 20 °C. Samples were imaged by DIC and fluorescence microscopy. Scale bar, 10 µm. The table below shows the hydrodynamic radius (R_h) for different dextrans in aqueous buffer, and is adapted from previously published work[32]. **f**, Lge1–Bre1 condensates are permeable to dextran of different sizes. Dextran is never excluded (partition ratios ≥ 1). Mean and s.d. are indicated; *n* = 60 condensates. **g**, Average R_h of recombinant proteins used in this work as measured by dynamic light scattering at 20 °C. The expected molecular mass was calculated according to the amino acid composition of the protein, and compared to the experimental molecular mass obtained by dynamic light scattering. Final data correspond to the average of at least two independent measures. n.d. not determinable.



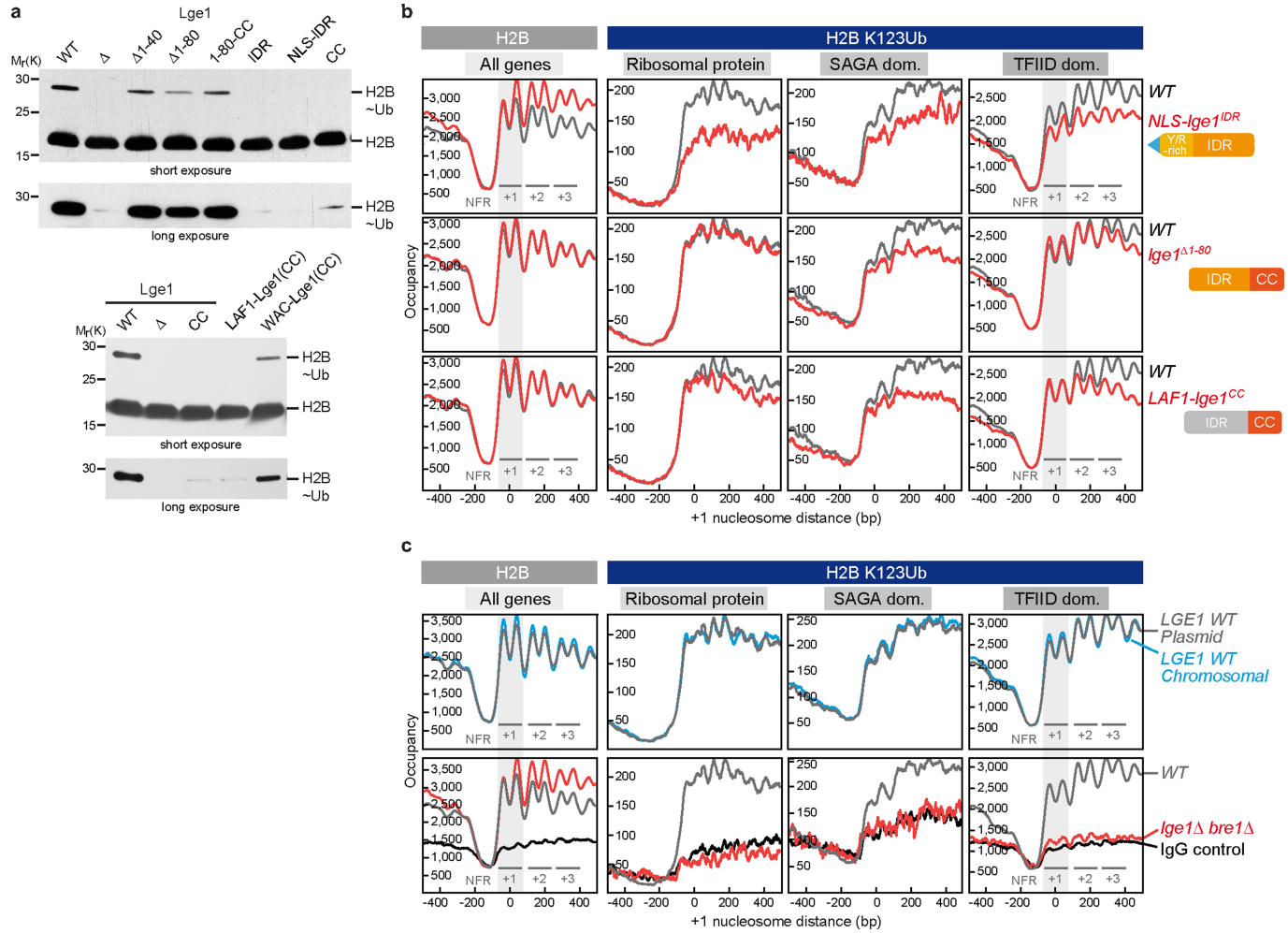
Extended Data Fig. 5 | Lge1 tyrosine residues are critical for LLPS. a, Phase-separation assay with the indicated 6×His–Lge1 constructs (10 µM). Scale bar, 10 µm. Cartoons are drawn to scale. Lge1(Y>A, 1–102) contains three additional mutations within amino acids 1–102 besides the Y>A mutations in the sticker region (amino acids 1–80) (Extended Data Fig. 2a), which increases the disruption of Lge1 LLPS in vitro. **b**, Quantification of condensate sizes (6×His–Lge1 constructs, 10 µM). *n*, quantified condensates. Dot plot shows median and interquartile range. ****P value < 0.0001, determined by two-sided Mann–Whitney test. n.d. not determinable. **c**, Input gel for **a**. Asterisk indicates degradation product, arrowheads indicate Lge1 constructs. See Supplementary Fig. 2 for uncropped gels.



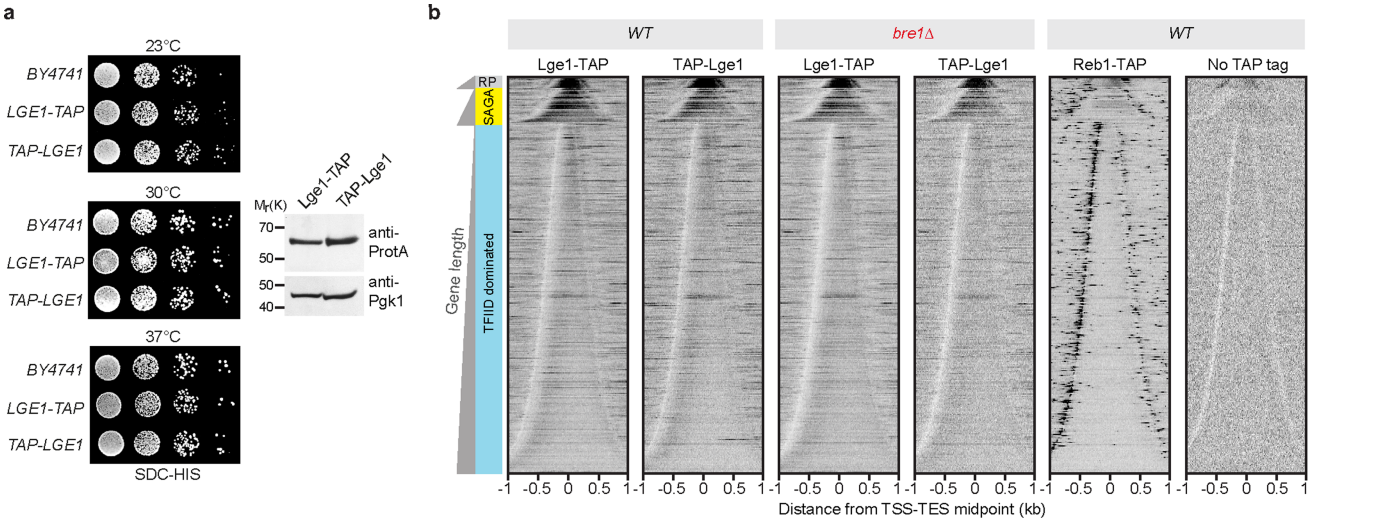
Extended Data Fig. 6 | Lge1–Bre1 condensates recruit the E2 Rad6 and chromatin. a, Recruitment of Pacific-Blue-labelled 6×His–Rad6 (here denoted 'Rad6*') (3 μM) or mGFP–Bre1 (3 μM) to Lge1 condensates. Experimental conditions as in Fig. 2a. Fluorescent intensities were quantified across single condensates. *n*, number of condensates. Scale bar, 5 μm. **b**, The recruitment of Rad6 to condensates was examined by adding Rad6* to 6×His–Lge1 condensates in the presence and absence of an mGFP–Bre1 (1.5 μM) or mGFP–Bre1(LBD) (15 μM) shell. The Bre1(LBD) construct has a weaker affinity to Lge1 than full-length Bre1, and therefore requires a higher concentration in the assay. Microscopy was performed immediately after adding Rad6* (3 μM) and followed over time. Fluorescent intensities were quantified across single condensates. *n*, number of condensates. Scale bar, 5 μm. **c**, Reconstituted mononucleosomes (1×NCP) and oligonucleosomes (16×NCP) were analysed by SDS–PAGE and Coomassie staining to assess purity and stoichiometry. **d**, Negative-stain electron microscopy was performed to assess the structure of oligonucleosomes. White arrowheads label individual nucleosomes, blue arrowheads indicate the linker DNA. Scale bar, 100 nm. **e**, Recruitment of mononucleosomes to Lge1 condensates with an mGFP–Bre1 shell. Hoechst-labelled, reconstituted mononucleosomes (1×NCP*, 0.5 μM) were added to 6×His–Lge1 condensates with an mGFP–Bre1 shell (3 μM), and imaged over time (min). *n*, number of condensates. Scale bar, 5 μm. **f**, Recruitment of oligonucleosomes to Lge1 condensates. The 16×NCPs* (0.5 μM) were added to 6×His–Lge1 condensates and imaged over time (min). *n*, number of condensates. Scale bar, 5 μm. **g**, Diffusion and retention of 601 Widom DNA into Lge1 condensates. Same setup as in **f** but with Hoechst-labelled 16× 601 Widom DNA (16×DNA*) added to 6×His–Lge1 condensates. *n*, number of condensates. Scale bar, 5 μm. **h**, Traces (optical density at 254 nm) of the sucrose gradient sedimentation assays (5–45%) show a reproducible fractionation pattern for cell extracts prepared from the indicated strains. The different ribosomal species and fractions are indicated and correspond to the fractions in Fig. 3a. **i**, Protein levels in cell extracts used for sucrose gradient assays. Anti-Pgk1 serves as a loading control. See Supplementary Fig. 2 for uncropped gels and western blots.

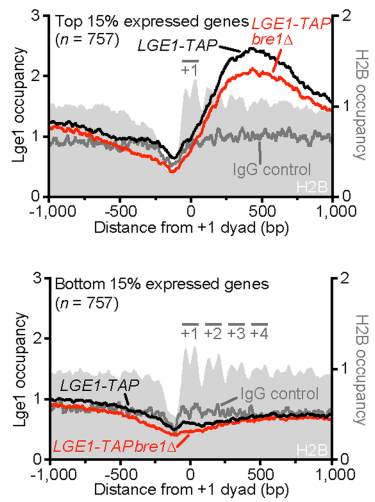


Extended Data Fig. 7 | Analysis of Lge1 and Bre1 in vivo. a, Gallery of representative images of bre1Δ lge1Δ cells co-expressing the indicated constructs from the strong GPD promoter. Scale bar, 2 µm. **b**, 1,6-Hexanediol treatment (10%, w/v) reduces the formation of Lge1–Bre1 puncta, but also affects nuclear import and thus complicates the interpretation of hexanediol effects in cells. White asterisk labels the vacuole; dashed white line shows the cell contour. Scale bar, 2 µm. **c**, **d**, Live imaging of bre1Δ lge1Δ cells expressing mGFP–Bre1 (**c**) or an empty vector (**d**), and the indicated Lge1–mCherry constructs. Dashed white line shows the cell contour. The fluorescence intensity of Lge1–mCherry constructs was quantified across a line spanning the nucleus. For comparison, the arbitrary fluorescence unit value = 1 is marked with a horizontal dashed line (except for **d**, in which the dashed line indicates a value of 0.5). *n*, number of randomly selected cells. Scale bar, 2 µm. **e**, Quantification of background-corrected total cell fluorescence (CTCF) of mCherry in **c**, **d**. Dot plots show median and interquartile range. **f**, Comparison of protein-expression levels of different Lge1–mCherry constructs in **c**. Cell lysates were analysed by SDS–PAGE and immunoblotting with anti-mCherry antibody. Anti-Pgk1 serves as a loading control. Asterisk indicates a degradation product. Red arrowheads indicate Lge1 constructs according to their predicted sizes. See Supplementary Fig. 2 for uncropped western blot. **g**, Gallery of representative images of bre1Δ lge1Δ cells expressing VC–Bre1 and Lge1–VN or Lge1(CC)–VN. Nup188–mCherry marks the nuclear envelope. Scale bar, 2 µm. **h**, Live imaging of bre1Δ lge1Δ cells expressing VC–Bre1 and the indicated Lge1–VN constructs from their endogenous promoters. Nup188–mCherry marks the nuclear envelope, and the dashed line marks the cell contour. Histograms represent pixel frequencies of fluorescence intensity values. Scale bar, 2 µm. **i**, Quantification of mean nuclear BiFC intensity in **h** and Fig. 3d. Median and interquartile range are indicated. *n*, number of cells. ***$P < 0.001$, determined by two-sided Mann–Whitney test. ns, not significant.



Extended Data Fig. 8 | Effects of the IDR of Lge1 on global H2B ubiquitination, and additional ChIP-exo analyses. a, Global levels of H2BK123ub. A lge1Δ Flag-H2B strain was transformed with plasmid-based variants of *LGE1-mCherry* or an empty plasmid. Cell lysates were subjected to anti-Flag immunoprecipitation, and analysed by SDS–PAGE and immunoblotting with anti-Flag antibody. Different exposures of the monoubiquitinated H2B band are shown. See Supplementary Fig. 2 for uncropped western blots. Expression levels of the Lge1–mCherry constructs in cell extracts are shown in Extended Data Figs. 7f, 10e. **b**, H2B (left) or H2BK123ub (right) ChIP-exo tag 5′ ends were plotted relative to the +1 nucleosome of all genes for wild type (grey trace) and the indicated *lge1* mutants (red traces). Sequencing tags were normalized across datasets to a 30-bp window centred in the nucleosome-free region (NFR) of all genes, representing unbound background regions. The first three genic nucleosomes are labelled +1, +2 and +3. Two H2B peaks are observed per nucleosome position. H2BK123ub patterns are shown separately for ribosomal-protein genes, and SAGA- and TFIID-dominated genes. Analyses were performed as in Fig. 4a. **c**, ChIP-exo analysis shows that Lge1 function is not compromised by the plasmid-based approach used in this study (top). H2B (left) or H2BK123ub (right) occupancy in strains with plasmid-based expression of *LGE1* under its endogenous promoter (grey trace) was compared to strains with *LGE1* in its chromosomal context (blue trace). Analysis of ChIP-exo background using a non-specific IgG antibody (bottom). H2BK123ub enrichment in an lge1Δ bre1Δ strain (red trace) is reduced to background levels (IgG negative control, black trace). Analyses were performed as described in **b**.





Gene class	Percentage represented Top 15%	Bottom 15%
RP	84	1
SAGA	23	11
TFIID	11	14

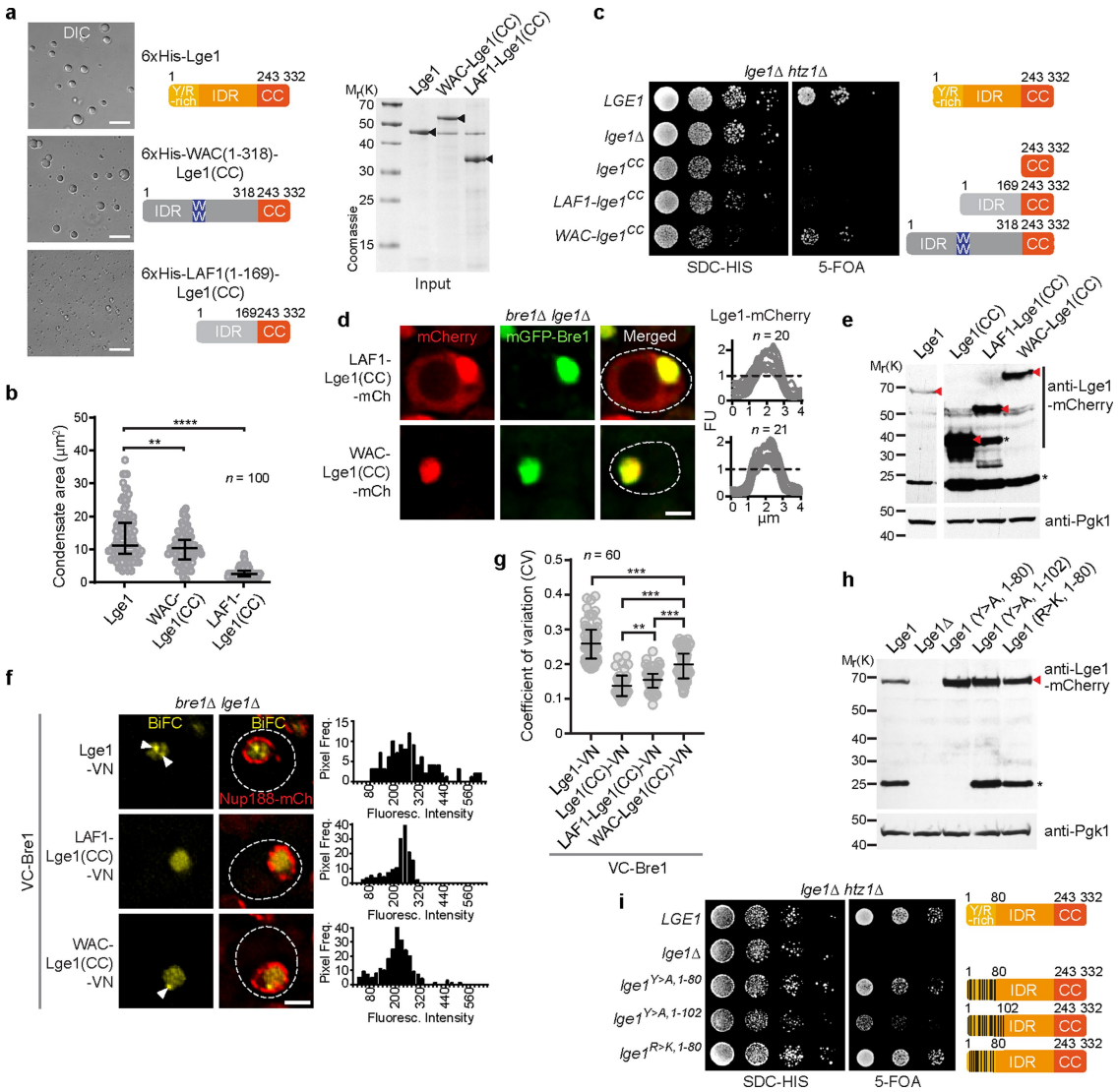
d

Strain		$[Ub/H2B]_{mut}$ / $[Ub/H2B]_{WT}$ [Rep1] / [Rep2] +1 nuc	+2 nuc	+3 nuc
BY4741	RP	1.02/N/A	1.01	1.08
	$SAGA_{-dom}$	1.09	1.04	1.04
	$TFIID_{-dom}$	1.01	0.99	0.98
lge1Δ bre1Δ	RP	0.25/0.24	0.22/0.21	0.29/0.20
	$SAGA_{-dom}$	0.65/0.69	0.48/0.59	0.52/0.54
	$TFIID_{-dom}$	0.48/0.43	0.38/0.34	0.34/0.31
lge1Δ	RP	0.51/0.66	0.49/0.61	0.52/0.49
	$SAGA_{-dom}$	0.67/0.88	0.58/0.79	0.54/0.65
	$TFIID_{-dom}$	0.74/0.83	0.61/0.72	0.57/0.68
NLS-lge1^{IDR}	RP	0.55/0.54	0.52/0.59	0.55/0.55
	$SAGA_{-dom}$	0.73/0.82	0.62/0.81	0.62/0.74
	$TFIID_{-dom}$	0.78/0.82	0.66/0.73	0.61/0.69
lge1$^{Δ1-80}$	RP	1.00/0.99	0.98/0.94	0.99/0.85
	$SAGA_{-dom}$	0.87/1.05	0.85/0.98	0.80/0.88
	$TFIID_{-dom}$	1.01/1.09	0.99/1.02	0.92/0.93
lge1^{CC}	RP	0.91/0.91	0.83/0.83	0.90/0.78
	$SAGA_{-dom}$	0.85/0.98	0.84/0.99	0.79/0.96
	$TFIID_{-dom}$	0.99/0.95	0.90/0.83	0.84/0.75
LAF1-lge1^{CC}	RP	0.92/0.86	0.88/0.79	0.86/0.73
	$SAGA_{-dom}$	0.85/0.87	0.82/0.84	0.78/0.76
	$TFIID_{-dom}$	0.95/0.94	0.89/0.85	0.83/0.77
WAC-lge1^{CC}	RP	0.93/0.87	0.95/0.85	1.01/0.85
	$SAGA_{-dom}$	0.88/0.95	0.88/0.93	0.87/0.89
	$TFIID_{-dom}$	0.98/0.98	0.98/0.96	0.97/0.93

e

Strain	Replicate correlation (R) H2B	H2BK123ub
WT	0.98	0.95
lge1Δ bre1Δ	0.96	0.98
lge1Δ	0.98	0.88
NLS-lge1^{IDR}	0.95	0.93
lge1$^{Δ1-80}$	0.95	0.92
lge1^{CC}	0.97	0.90
LAF1-lge1^{CC}	0.91	0.86
WAC-lge1^{CC}	0.98	0.95

Extended Data Fig. 9 | Enrichment of Lge1 along gene bodies coincides with establishment of H2B ubiquitination pattern. a, TAP-tagging does not impair Lge1 function under the conditions we tested (related to ChIP-exo in **b**). C- or N-terminally TAP-tagged versions of *LGE1* were transformed into *lge1*Δ cells, and their growth was compared to an untagged BY4741 control strain. Lge1 proteins were expressed at similar levels. Cell lysates were analysed by SDS–PAGE and immunoblotting with anti-protein A antibody. Anti-Pgk1 serves as a loading control. See Supplementary Fig. 2 for uncropped western blots. **b**, Genome-wide binding profiles of Lge1. Frequency distribution of 5′ tags of Lge1 ChIP-exo were mapped to the midpoint between gene transcript start (TSS) and end (TES) for ribosomal-protein, SAGA-dominated and TFIID-dominated gene classes. Each class was sorted by gene length, thus generating bell plots. Lge1 binding profiles in wild-type and *bre1*Δ backgrounds are depicted. Insertion of the TAP tag at the N or C terminus gave similar enrichment patterns. TAP-tagged Reb1 and 'No TAP tag' serve as positive (site-specific binding) and negative controls, respectively. **c**, Lge1 binding is tied to the expression of the target gene, and is largely independent of Bre1 (black versus red traces). Composite plots of Lge1 enrichment at top- and bottom-15% expressed genes[33] are shown in the top and bottom panels, respectively. The IgG negative control is depicted as a grey trace. ChIP-exo tag 5′ ends were mapped to the +1 nucleosome dyad (as defined by micrococcal nuclease (MNase) H3 ChIP sequencing). H2B occupancy is depicted as a filled light grey trace. The +1, +2, +3 and +4 nucleosome positions are highlighted. The table represents the percentage of the members of each gene class included in the top and bottom 15%. **d**, Quantification of H2BK123ub density for the first three genic nucleosome positions in various Lge1 mutants. Table depicts the fold enrichment of H2BK123ub density (ChIP-exo H2BK123ub/H2B) at canonical nucleosome positions +1, +2, and +3 for the indicated mutants relative to wild type. Data for two biological replicates are shown. BY4741 is a positive control and correlates with the strain carrying wild-type *LGE1* on a plasmid. Ratios for each gene class are indicated, to directly compare with graphs shown in Fig. 4a, Extended Data Fig. 8b. **e**, Correlation between two biological replicates for H2B and H2BK123ub datasets are shown. ChIP-exo tag 5′ ends were binned in 500-bp intervals, and the coefficient of correlation between the two datasets was calculated.



Extended Data Fig. 10 | Specific features of yeast Lge1 and human WAC contribute to H2BK123ub and cell viability. a, The IDRs of WAC and LAF1 promote LLPS. Phase-separation assay of recombinant 6×His–Lge1 and the fusion constructs 6×His–WAC(1–318)–Lge1(CC) and 6×His–LAF1(1–169)–Lge1(CC) (both proteins, 5 µM; buffer, 20 mM Tris pH 7.5, 100 mM NaCl and 1 mM DTT, 20 °C). Scale bar, 10 µm. Protein inputs are shown on the right (black arrows). **b**, Quantification of condensate sizes (in µm^2) in **a**. *n*, number of condensates. Dot plot showing median and interquartile range. $^{**}P = 0.004$, $^{****}P < 0.0001$, determined by two-sided Mann–Whitney test. **c**, A synthetic genetic approach was used to investigate the functionality of Lge1 LLPS in vivo. Cells were inviable when *LGE1* and *HTZ1* were deleted together, indicating a functional relationship. Double-deletion strains containing a wild-type *LGE1* cover plasmid (Ura marker) were cotransformed with the indicated plasmids (His marker). Growth was followed on SDC–His (loading control) and on SDC + 5-FOA, which shuffles out the Ura cover plasmid. Cells were spotted in tenfold serial dilutions and incubated for two (SDC–His) or three days (5-FOA) at 30 °C. **d**, Live imaging of bre1Δ lge1Δ cells expressing mGFP–Bre1 and WAC(1–318)–Lge1(CC)–mCherry or LAF1(1–169)–Lge1(CC)–mCherry shows protein import into the nucleus. Dashed white line indicates the cell contour. Fluorescence intensity of the mCherry construct was quantified across a line spanning the nucleus. For comparison, the arbitrary fluorescence unit value = 1 is marked with a horizontal dashed line. *n*, number of cells. Scale bar, 2 µm. **e**, Cell lysates of strains in **c** and **d** were analysed by SDS–PAGE and immunoblotting with anti-mCherry antibody. Anti-Pgk1 serves as a loading control; asterisks indicate degradation products. Red arrowheads indicate Lge1 constructs according to their predicted sizes. **f**, Live imaging of bre1Δ lge1Δ cells expressing VC–Bre1 and WAC (1–318)–Lge1(CC)–VN or LAF1(1–169)–Lge1(CC)–VN constructs from *LGE1* endogenous promoter. Arrowheads label nuclear BiFC puncta, Nup188–mCherry labels the nuclear envelope and dashed lines indicate the cell contours. Histograms represent pixel frequency of fluorescent intensity values. Scale bar, 2 µm. **g**, Coefficient of variation of the fluorescence-intensity distribution of BiFC signals in **f** and Fig. 3e. The higher the coefficient of variation, the greater the heterogeneity of the BiFC signal. A propensity for LLPS is suggested by an increased coefficient of variation of the WAC(1–318)–Lge1(CC)–VN construct. Dot plot showing median and interquartile range. *n*, number of cells. $^{**}P = 0.0024$, $^{***}P < 0.001$, determined by two-sided Mann–Whitney test. **h**, Expression levels of Lge1–mCherry constructs used in **i**. Cells lysates were analysed by SDS–PAGE and immunoblotting with anti-mCherry antibody. Anti-Pgk1 serves as a loading control. Asterisk indicates degradation products. Red arrowhead indicates Lge1 constructs. **i**, Genetic interaction analysis, set up as in **c** with the indicated plasmids. See Supplementary Fig. 2 for uncropped gels and western blots.

nature research

Corresponding author(s): Alwin Köhler

Last updated by author(s): Feb 29, 2020

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒ A description of all covariates tested
☒	☐ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒ For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒ Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	High-throughput DNA sequencing for ChIp-exo was performed using Illumina NextSeq 500 in a paired-end mode to produce 2x40 bp reads. Obtained sequence reads were aligned to the yeast genome (sacCer3) using bwa-mem (v0.7.9a). PCR duplicates and non unique alignments were removed from aligned reads to obtain uniquely mapped reads. Sequence hydrophobicity for Lge1 and WAC were obtained by GROMACS 5.1.4. VOLPES.
Data analysis	For all graphs included and statistic analyses, the GraphPad Prism 7.0e was used. For analysis of microscopy images, the SoftWoRx software 6.5.2 (GE Healthcare) for deconvolution and ImageJ v1.46r or v1.51s were used. Figure 4a, Extended Data Figures 8b and 8c - Aligned BAM files (deduplicated) were converted to TAB format using Scriptmanager v0.12. TAB files were shifted by 6 bp using script shift_tags.py. Shifted files were normalized to a 30 bp NFR region using script normalization_script.py and mapped to coordinate files using script map_shited_tags_to_ref.py. Composite plots were generated using script composite_plots_Vinesh_v2.py. Extended Data Figures 9b and 9c - Scriptmanager v0.12 was used to map deduplicated BAM files to various coordinate files and Prism 7 was used to generate composite plots. Data analysis parameters are described in detail in the Methods section and on GITHUB (https://github.com/CEGRcode/Gallego_2019).

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

ChIP-exo sequencing files for Figure 4 and Extended Data Figure 8 and 9 can be publicly accessed at NCBI Gene Expression Omnibus (https://www.ncbi.nlm.nih.gov/geo/) under accession number GSE131639. Additional replicates of Reb1-TAP can be found at GSE110681.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Two independent experiment replicates were obtained for all ChIP-exo experiments, unless otherwise mentioned. This is a standard practice in genomics.
Data exclusions	Binding data at t-RNA and non coding genes were excluded. These exclusion criteria were pre-established, since the focus of the study is on RNA Pol2 transcribed coding genes only.
Replication	Datasets shown in Figure 4a were successfully replicated. Correlation coefficients between the two replicates were calculated and reported in Extended Data Figures 9d and 9e. In addition, individual replicates were analyzed and compared to ensure obtained results were consistent. In Extended Data Figure 9b, indicated factors are probed, with the TAP tags on either N or C-terminus and thus serving as replicates.
Randomization	Randomization was not performed using data analysis, since negative control experiments were performed in parallel.
Blinding	Studies were not blinded because the experiments were performed and analyzed by the same group of scientists. In vitro biochemistry, microscopy and in vivo functional experiments were performed independently by different researchers and their findings support one another, thereby providing independent verification.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology
☒	☐ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data

Methods

n/a	Involved in the study
☐	☒ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used

- Mouse monoclonal antibody against HA; clone 12CA5 (Sigma Aldrich, 11666606001), lot #43852500, lot #21900900 (Western Blotting 1:1000)
- Rabbit polyclonal antibody anti-Protein A (Sigma Aldrich, P3775), lot #129K4786 (Western Blotting 1:10 000)
- Mouse monoclonal antibody against CTD repeats of subunit Rpb1 of RNA Polymerase II; clone 8WG16 (Covance, MMS-126R), lot #B216274 (Western Blot 1:1000)
- Mouse monoclonal anti-Pgk1; clone 22C5D8 (Abcam, Ab113687), lot #GR3286258-2 (Western Blotting 1:10 000)
- Mouse monoclonal antibody anti-Flag; clone M2 (Sigma Aldrich, F1804) for recombinant protein detection, lot #SLBF6631V, lot #SLBX2256, lot #SLCC6485 (Western Blotting 1:2000)

- Mouse monoclonal anti-Flag M2-Peroxidase (Sigma Aldrich, A8592) for global levels of H2B ubiquitination, lot #SLBN8495V, lot #SLBV3799 (Western Blotting 1:1000)
- Mouse monoclonal anti-mCherry; clone 1C51 (Abcam, Ab125096), lot #GR271170-1, lot #GR3264408-1 (Western Blotting 1:2000)
- Rabbit monoclonal antibody against ubiquityl-Histone H2B (Lys120)(Cell Signaling, 5546, lot #5) and rabbit polyclonal to histone H2B (Abcam, ab1790, lot #GR310932-1) (Van Oss et al., 2017) were used to probe for the specific ubiquitin mark and the underlying nucleosome, respectively.
- Mouse monoclonal antibody against the Strep-tag II (Qiagen, 34850), lot #157031640, (Western Blotting 1:5000)
- For TAP tagged strains, rabbit IgG (Sigma, i5006, lot #SLBM2617V) conjugated to Dynabeads was used.

Validation

- Mouse monoclonal anti-HA (Sigma Aldrich, 11666606001) https://www.sigmaaldrich.com/catalog/product/roche/roaha?lang=de&region=AT
- Rabbit polyclonal anti-Protein A (Sigma Aldrich, P3775) https://www.sigmaaldrich.com/catalog/product/sigma/p3775?lang=de&region=AT
- Mouse monoclonal anti-Flag (Sigma Aldrich, F1804) https://www.sigmaaldrich.com/catalog/product/sigma/f1804?lang=de&region=AT
- Mouse monoclonal anti-Flag M2-Peroxidase (Sigma Aldrich, A8592) https://www.sigmaaldrich.com/catalog/product/sigma/a8592?lang=de&region=AT
- Mouse monoclonal anti-Pgk1 (Abcam, Ab113687) https://www.abcam.com/pgk1-antibody-22c5d8-ab113687.html
- Mouse monoclonal anti-mCherry (Abcam, Ab125096) https://www.abcam.com/mcherry-antibody-1c51-ab125096.html
- Mouse monoclonal antibody against CTD repeats of subunit Rpb1 of RNA Polymerase II (Covance, MMS-126R). Used original stock from Covance. The product is discontinued. https://www.antibodypedia.com/gene/3208/POLR2A/antibody/1457935/MMS-126R
- Mouse monoclonal anti-Strep (Qiagen, 34850) https://www.qiagen.com/at/products/discovery-and-translational-research/protein-purification/tagged-protein-expression-purification-detection/strep-tag-antibody/?clear=true#orderinginformation
- Rabbit polyclonal to histone H2B (Abcam, ab1790) https://www.abcam.com/histone-h2b-antibody-chip-grade-ab1790.html
- Rabbit monoclonal antibody against ubiquityl-Histone H2B (Lys120)(Cell Signaling, 5546) https://www.cellsignal.com/products/primary-antibodies/ubiquityl-histone-h2b-lys120-d11-xp-rabbit-mab/5546
Ref: Van Oss et al., 2017
- For TAP tagged strains, rabbit IgG (Sigma, i5006) conjugated to Dynabeads was used. https://www.sigmaaldrich.com/content/dam/sigma-aldrich/docs/Sigma/Product_Information_Sheet/2/i5006pis.pdf

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)

Saccharomyces cerevisiae (budding yeast) was used in this study. Strains were purchased or generated by the Koehler lab:
- S. cerevisiae strain BY4741 WT, purchased from Euroscarf Y00000
- S. cerevisiae strain LGE1-TAP, purchased from Dharmacon (www.horizondiscovery.com) YSC1178-202233656.
- Derivatives of these strains were generated during the study and are described in the Supplementary Table.
- TAP tagged strains shown in Extended Data Figure 9 were obtained from the Pugh lab (through Open Biosystems) .

Authentication

All yeast strains (commercial or self-made) were verified by PCRs, or where possible by additional immunoblotting and fluorescence microscopy.

Mycoplasma contamination

Not applicable, since no mammalian cell lines were used.

Commonly misidentified lines
(See ICLAC register)

Not applicable, since no mammalian cell lines were used.

ChIP-seq

Data deposition

☒ Confirm that both raw and final processed data have been deposited in a public database such as GEO.

☒ Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

Data access links
May remain private before publication.

All Datasets generated by ChIP-exo can be publicly accessed at NCBI Gene Expression Omnibus (https://www.ncbi.nlm.nih.gov/geo/) using accession number GSE131639.

Files in database submission

All files can be accessed on GEO using accession number GSE131639.

Genome browser session
(e.g. UCSC)

No longer applicable.

Methodology

Replicates

Two biological replicates for each experiment were performed, unless otherwise specified.

Sequencing depth

High-throughput DNA sequencing for all ChIP-exo samples was performed in a paired-end mode to produce 2x40 bp reads. For Figure 4a, average total reads for H2B data ranged from 3.67 to 6.04 million, with 2.90 to 4.32 million unique deduplicated reads, respectively. For H2BK123ub datasets, sequencing reads ranged from 813,000 to 7.48 million total

	reads resulting in 600,000 to 4.77 million unique deduplicated mapped reads. For Extended Data Figure 9, total reads ranged from 501,559 to 6.03 million resulting in 209,419 to 3.33 deduplicated uniquely mapped reads.
Antibodies	- Rabbit monoclonal antibody against ubiquityl-Histone H2B (Lys120)(Cell Signaling, 5546, lot #5) and rabbit polyclonal to histone H2B (Abcam, ab1790, lot #GR310932-1) (Van Oss et al., 2017) were used to probe for the specific ubiquitin mark and the underlying nucleosome, respectively. - For TAP tagged strains, rabbit IgG (Sigma, i5006, lot #SLBM2617V) conjugated to Dynabeads was used. https://www.sigmaaldrich.com/content/dam/sigma-aldrich/docs/Sigma/Product_Information_Sheet/2/i5006pis.pdf
Peak calling parameters	Parameters for data analysis can be accessed at https://github.com/CEGRcode/Gallego_2019.
Data quality	Parameters for data analysis can be accessed at https://github.com/CEGRcode/Gallego_2019.
Software	Scriptmanager v0.11 was used for majority of the data analyses, as described in GITHUB (https://github.com/CEGRcode/Gallego_2019).