Published in final edited form as:

J Proteome Res. 2018 October 05; 17(10): 3475–3484. doi:10.1021/acs.jproteome.8b00395.

SILAC-MS profiling of reconstituted human chromatin platforms for the study of transcription and RNA regulation

Maggie M. Balas^{1,2,3}, Allison M. Porman^{1,2,4}, Kirk C. Hansen², and Aaron M. Johnson^{1,2,3,4,5,*}

¹Molecular Biology Program, University of Colorado Denver Anschutz Medical Campus 12801 East 17th Ave., Aurora, CO, United States

²Department of Biochemistry and Molecular Genetics, University of Colorado Denver Anschutz Medical Campus 12801 East 17th Ave., Aurora, CO, United States

³RNA Bioscience Initiative, University of Colorado Denver Anschutz Medical Campus 12801 East 17th Ave., Aurora, CO, United States

⁴Program in Cancer Biology, University of Colorado Denver Anschutz Medical Campus 12801 East 17th Ave., Aurora, CO, United States

⁵Linda Crnic Institute for Down Syndrome, University of Colorado Denver Anschutz Medical Campus 12801 East 17th Ave., Aurora, CO, United States

Abstract

DNA packaged into chromatin is the core structure of the human genome. Nearly all eukaryotic genome regulation must interface with this genomic structure and modification to the chromatin can influence molecular mechanisms that regulate the underlying DNA. Many processes are governed by regulated step-wise assembly mechanisms that build complex machinery on chromatin to license a specific activity such as transcription. Transcriptional activators drive the initial steps of gene expression, regulated in part by chromatin. Here we describe tools to study the step-wise assembly of protein complexes on chromatin in a highly-controlled manner using reconstituted human chromatin platforms and quantitative proteomic profiling. We profile the early steps in transcriptional activation and highlight the potential for understanding the multiple ways chromatin can influence transcriptional regulation. We also describe modifications of this approach to study the activity of a long noncoding RNA to act as a dynamic scaffold for proteins to be recruited to chromatin. This approach has the potential for a more comprehensive understanding of important macromolecular complex assembly that occurs on the human genome. The reconstituted nature of the chromatin substrate offers a tuneable system, able to be trapped at specific sub-steps, to understand how chromatin interfaces with genome regulation machinery.

Keywords

chromatin; SILAC; transcription activation; activator; Mediator; lncRNA; Xist; RepA; TDP-43; RNA tethering

*Corresponding Author: Aaron M. Johnson PhD, Aaron.m.johnson@ucdenver.edu, Tel: +1(303)724-3224.

Introduction

The assembly of eukaryotic genomes into chromatin presents a barrier to many DNA-dependent processes and an opportunity for regulation through post-translational modification of the histone proteins that package the DNA¹. Many processes of genome regulation have been biochemically characterized and are known to be driven by regulated recruitment to chromatin (Figure 1A). For example, transcription by RNA Polymerase II is regulated first at the step of initiation of transcription through a multi-step process involving association of transcription machinery with a promoter. Many distinct components of the transcriptional machinery are multi-subunit complexes, adding to the complexity of studying these processes. Traditional methods to analyze this mechanism have relied on using Western analysis with antibodies for specific subunits of the transcription complexes as a proxy for assembly of each complex. We and others have used label-based quantitative proteomics as a way to interrogate to a greater extent the full mechanism of eukaryotic transcription initiation on chromatin using the budding yeast system²⁻³. We also were able to profile which events in the transcription mechanism were blocked by the assembly of yeast heterochromatin². This system has allowed for detailed investigation of the molecular assemblies that occur on chromatin and the approach is broadly applicable to other organisms. Stable isotope labeling of amino acids in cell culture coupled to mass spectrometry (SILAC-MS) has been previously used to study recombinant nucleosomes from metazoan histones with specific post-translational modifications⁴⁻⁵. Here we describe SILAC approaches to study molecular assembly on human chromatin, using transcriptional activation and RNA-scaffolded assembly as proof-of-principle contexts.

Using a model yeast system, the ability of transcriptional activators to recruit gene expression and chromatin modification machinery to reconstituted nucleosomal substrates was previously profiled using label-based, comparative quantitative proteomics methods²⁻³. Additionally, label-free proteomic methods have been used to study activation of the human transcriptional machinery under various conditions⁶⁻⁸. Transcriptional activators are typically composed of at least two domains, with separable functions: a DNA binding domain and a transactivation domain⁹. For example, the Herpes virus VP16 transcriptional activator is a well-studied viral protein that activates transcription. VP16 has been used as a model for transcription activation, often with the VP16 activation domain fused to a separate DNA Binding domain from the yeast Gal4 protein. The typical transactivation domain requires an overall acidic amino acid sequence and is relatively unstructured, yet differences in amino acid sequence likely still underlie observed differences in how activators function mechanistically¹⁰.

Long noncoding RNAs (lncRNAs) are another class of molecules that can associate with chromatin to regulate transcription, in part through direct recruitment of proteins to the genome¹¹. Long noncoding RNAs have been shown to regulate multiple genomic processes including gene repression, gene activation, and genome stability, often with implications for disease when these mechanisms are mis-regulated¹². The Xist lncRNA is the most well-studied of these molecules. Xist regulates mammalian dosage compensation, the process of limiting the expression of one X chromosome (ChrX) copy in females to maintain similar levels to males. Xist interacts with ChrX chromatin and coordinates the near-complete

silencing of one copy of ChrX through conversion to a heterochromatic state, called X-inactivation¹³. The human Xist lncRNA is ~19 kb, composed of multiple regions with specific functions¹⁴. One key region, known as Repeat A (RepA), is located near the 5' end of Xist. RepA can be expressed on its own or as part of the full-length transcript. RepA has been identified as one of the important Xist regions for carrying out the gene silencing that occurs across the X chromosome. Additionally, RepA is a key scaffold for proteins that interact with Xist to contribute to X-inactivation¹⁵.

Here we describe the use of reconstituted human chromatin substrates to study aspects of transcription initiation by SILAC profiling (Figure 1B). In addition, we demonstrate how similar approaches can be used to study proteins that associate with an RNA molecule that interacts with the genome. These systems provide a tightly-controlled robustly comparable framework for investigating regulated recruitment of proteins to chromatin and how the specific patterns of chromatin modifications can influence the mechanisms of transcription and large-scale chromosome repression.

EXPERIMENTAL PROCEDURES

Cell culture and nuclear extract preparation:

For SILAC experiments with HeLa cells, we used SILAC DMEM medium (Gibco) supplemented with 10% dialyzed FBS (Gibco; Life Technologies), 4 mM glutamine, 0.339 mM of either heavy L-Arginine HCL (¹³C₆ only or ¹³C₆/¹⁵N₄; Cambridge Isotopes) or light (Sigma), and 0.798 mM of either heavy L-Lysine-2HCL (¹³C₆/¹⁵N₂, Cambridge Isotopes) or light lysine (L-Lysine 2HCL, Sigma). Cells were grown in SILAC medium for at least one week prior to nuclear extract preparation. Nuclear extracts were prepared as previously described¹⁶.

DNAs, RNAs, and Proteins

To generate a DNA substrate for chromatin reconstitution, a PCR product consisting of a 601 positioning sequence¹⁷, six LexA binding sites, five Gal4 DNA binding sites, and a viral E4 promoter was cloned into pGEM3z/601 vector backbone via isothermal assembly. PCR with a biotinylated reverse primer was used to add a biotin tag to the 3' end of the 1.1-kb DNA substrate. RepA was PCR amplified from HeLa genomic DNA with Phusion polymerase, followed by PCR cleanup with the E.Z.N.A. Cycle Pure Kit (Omega Biotek). The RepA DNA template for in vitro transcription was generated by adding a 5' T7 promoter sequence and 3' RAT tag sequence using tailed PCR primers. 5' RAT tagged RNA derived from the antisense sequence of the firefly luciferase mRNA (anti-Luc) and the PP7-LexA fusion protein were prepared as previously described¹⁸. The Gal4-VP16 fusion protein was prepared as previously described¹⁹. In vitro transcription of all RNA was performed using the T7 enzyme (NEB) for 4 h at 37°C. RNA was treated with DNase1 (Ambion) for 15 min, and RNA was purified with the RNeasy kit (QIAGEN)

Chromatin Assembly

Chromatin was assembled as previously described¹⁸. Briefly, chromatin was reconstituted through a previously developed enzymatic assembly method using a histone chaperone and

nucleosome spacing factor^{20–21}. Biotinylated DNA, human histone octamers (H2A, H2B, H3.1, and H4), human histone chaperone Nap1, nucleosome positioning factor Isw1a, and an ATP regeneration system (final concentrations: 30 mM creatine phosphate, 3 mM ATP, 4.1 mM MgCl₂, and 6.4 µg/mL creatine kinase) were incubated at 30°C for 5 h in R+ Buffer (10 mM Hepes pH 7.5, 10 mM KCl, 1.5 mM MgCl₂, 500 µM EGTA, 10% Glycerol, 2.5 mM β-glycerophosphate, 200 µM PMSF, 1 mM DTT). Chromatin was subsequently conjugated to streptavidin magnetic beads (Life Technologies) for 2 h at RT as previously described²².

RNA Tethering

Equal molar quantities of 3' RAT-tagged RepA (16.4 µg) or 5' RAT tagged anti-luciferase (16.8 µg) IVT RNA were incubated with 750 ng LexA-PP7 fusion protein for 1 h at RT to prebind. Prebound RNA LexA-PP7 fusion protein was incubated with 1.8 µg biotinylated chromatin conjugated to magnetic beads, for 1 h at RT to bind chromatin. All samples were incubated in RB buffer (20mMTris–HCl pH 7.5, 100mM NaCl, 10mM MgCl₂, 500 µM EGTA, 0.01% NP-40). Untethered RNA and fusion protein were removed by 2X magnet separation washes in RB buffer, followed by resuspension in PB100 buffer (50 mM Hepes pH 7.5, 500 mM NaCl, 10 mM Magnesium acetate, 100 µM EDTA, 0.02% NP-40, 5% Glycerol, 5 mM EGTA, 1 mM PMSF, and 1 mM DTT) and RNase inhibitor (Roche).

Chromatin Pulldown for Mass spectrometry

RNA–chromatin substrates were incubated with 450 µg of heavy or light SILAC HeLa cell nuclear extracts in PB100 buffer for 1 h at RT, followed by 1 h at 4°C, both with end-over-end rotation. Samples were then washed with 1.5 mL cold PB100 at 4°C. Proteins were eluted from the streptavidin beads by resuspending beads in cold 100 µL 50 mM Hepes pH 7.5, 2 M NaCl, incubating at 4 °C 15 minutes, and magnet separating. Elutions from heavy and light experiments were mixed in equal volume, diluted to 1 mL with 50 mM Hepes pH 7.5, and precipitated in trichloroacetic acid.

Mass spectrometry

Samples were prepared for mass spectrometry using a modified FASP protocol²³. Samples were processed in 8 M urea, alkylated with 25 mM iodoacetamide for 30 min, and placed in a YM-30 Microcon column (Sigma). Samples were washed twice with 8 M urea, twice with 2 M urea, and then digested with 0.5 µg of trypsin (Promega MS grade) with 0.2 µL of 0.5 M CaCl₂ at 37 °C for 6 hrs, and eluted by centrifugation. A second elution was performed with 1% formic acid, and desalted using C-18 columns (Pierce). The samples were eluted from the C-18 resin in 70% acetonitrile, 0.1% tri-fluoroacetic acid and placed in a Speedvac. After bringing the samples back up in 1% formic acid they were analyzed on an Orbitrap Q Exactive mass spectrometer (Thermo Fisher Scientific) coupled to an nanoEASY LC system through a nanoelectrospray LC – MS interface as previously described²⁴. A volume of 8 µL of sample was injected into a 10 µL loop using the autosampler. The analytical column was loaded at 600 nL/min (in house-made 100 µm i.d. × 200 mm fused silica capillary packed with 4 µm 80 Å Synergi Hydro C18 resin (Phenomex, Torrance, CA)). After 10 min of sample loading, the flow rate was adjusted to 350 nL/min, and a 180-min linear gradient of 2–40% acetonitrile with 0.1% formic acid was used to separate the peptides. LC mobile

phase solvents and sample dilutions used 0.1% formic acid in water (Buffer A) and 0.1% formic acid in acetonitrile (Buffer B) (Chromasolv LC–MS grade, Sigma-Aldrich). Data acquisition was performed using the instrument supplied Xcalibur™ (version 2.1) software. The mass spectrometer was operated in the positive ion mode. Each survey scan of m/z 400–2,000 was followed by Higher-energy C-trap dissociation (HCD) MS/MS of the twenty most intense precursor ions. Singly charged ions were excluded from HCD selection. Normalized collision energies were employed using helium as the collision gas. The 40 most-abundant contaminants and proteins from the extract that were unchanged in the control and variable experiments were used as an exclusion list during the Q Exactive runs.

The resulting raw files were processed, quantified and searched using MaxQuant version 1.3.0.51²⁵ and the Andromeda peptide search engine searching against the Uniprot *human* database. Parameters defined for the search were: trypsin as digesting enzyme, allowing two missed cleavages; a minimum length of 6 amino acids; fixed modifications of +8 on lysine residues ($^{13}\text{C}_6^{15}\text{N}_2$ L-lysine) and +6 or +10 on arginines ($^{13}\text{C}_6$ or $^{13}\text{C}_6^{15}\text{N}_4$ L-arginine), carbamidomethylation at cysteine residues as fixed modification, oxidation at methionine and protein N-terminal acetylation as variable modifications. The maximum allowed mass deviation was 10 ppm for the MS and 20 ppm for the MS/MS scans.

MaxQuant results were processed in the following manner prior to enrichment analysis. All recombinant proteins and sample handling contaminants were removed (porcine trypsin and human keratins). For proteins with an intensity value for only heavy or light, an estimate was used for the other value, based on an average of the lowest order of magnitude values in the run where only a single peptide was identified. For both Gal4-VP16 experiments, this value was $\sim 8.1\text{e}5$. For both RepA experiments the values were $\sim 2.5\text{--}2.8\text{e}6$. These values were multiplied by the number of peptides identified in the condition with an intensity value and this “minimum intensity*peptide number” was used as the estimate for the condition without an intensity value. Certain mediator complex subunits were in this category and served as a useful validation of this approach, as the enrichment values estimated were within two-fold of those subunits with MaxQuant determined enrichments. MaxQuant-determined normalized ratios were otherwise used.

The mass spectrometry proteomics data have been deposited in the MassIVE database run by the Center for Computational Mass Spectrometry at the University of California-San Diego (<https://massive.ucsd.edu>) with ProteomeXchange dataset identifier PXD009921.

RNA Immunoprecipitation

TDP43-bound RNA was quantified in the immunoprecipitates obtained from formaldehyde-crosslinked MCF-7 cells using a previously described method²⁶ with some modifications. Cells were trypsinized, washed twice with ice cold PBS and fixed with 1% v/v methanol-free formaldehyde in PBS for 10 min at 25°C on a rotor. Formaldehyde was quenched by adding glycine to a final concentration of 0.25 M and then incubating at 25°C for 5 min. Fixed cells were washed 3 times with ice-cold PBS and resuspended in RIPA buffer (50 mM Tris-HCl pH 7.4, 100 mM NaCl, 1% NP-40, 0.1% SDS, 0.5% sodium deoxycholate) with protease inhibitors (Roche) and 1 mM DTT per 3 million cells. Cells were incubated in the Bioruptor® Pico sonication device (Diagenode) for 10 cycles at 30 s ON, 30 s OFF. Lysates

were incubated on ice 10 min and subjected to DNase digestion for 5 minutes at 37°C with mixing at 1000 rpm (2 µl TURBO DNase per 0.5 ml lysate). Tubes were immediately transferred to ice and incubated for 5 min. Lysates were clarified by centrifugation at 21,000g at 4°C for 10 min. TDP-43 antibody was bound to RIPA-washed Protein A/G magnetic beads (Thermo Fisher). Antibody-bound beads were then washed with RIPA buffer, added to the lysate for immunoprecipitation, and incubated at 4°C overnight. Rabbit IgG antibody (Santa Cruz) was used as a control. Beads were washed five times with RIPA buffer containing 1 M NaCl and 1 M Urea and resuspended in 100 µl RNA elution buffer (50 mM Tris-HCl pH 7.4, 5 mM EDTA, 10 mM DTT, 1% SDS). Formaldehyde-induced crosslinks were reversed by incubation at 70°C for 30 min with mixing at 1000 rpm. Supernatant was removed from beads and mixed with Trizol LS (Thermo Fisher), and coimmunoprecipitated RNA was purified according to the manufacturer's instructions. Glycoblue (Thermo Fisher) was used to visualize the RNA pellet. Purified RNA was reverse transcribed with random hexamers using Super Script IV reverse transcriptase. RNA molecules were detected using qPCR primers (Table S3) and Takyon qPCR Mix (Eurogentek). Fraction recovered was calculated by normalizing quantification to the input sample and total amount of RNA recovered in each sample.

Results

Chromatin substrate assembly

We assembled short arrays of nucleosomes using a previously-developed protocol^{18, 22}. Briefly, recombinant, *E. coli*-expressed human histones were reconstituted into the histone octamer and isolated by size-exclusion chromatography. Histone octamers were assembled into regularly-spaced nucleosomal arrays through incubation with the Nap1 histone chaperone and a chromatin remodeling factor in the presence of a linear DNA substrate (Figure 1C,D). Regular-spacing of nucleosomes was assessed via limited micrococcal nuclease digestion of the sample and gel analysis (Figure 1E). The DNA substrate used was a synthetic sequence PCR product composed of a viral transcriptional promoter (adenovirus E4 protein promoter) downstream of five Gal4 binding sites²⁷. Upstream of this transcription cassette was an array of six LexA binding sites for use in later experiments (see below). Flanking these sequences on either side were 601 nucleosome positioning sequences, which are thermodynamically-favorable for nucleosome assembly²⁸. Finally, this PCR product was generated with one biotinylated primer to allow for bead conjugation.

Profiling transcriptional activator-dependent recruitment to chromatin by SILAC-MS

Most RNA Polymerase II transcription that occurs in eukaryotic nuclei is dependent on one or more transcriptional activators. Transcriptional activators stimulate transcription by direct recruitment of core transcriptional machinery such as TFIID and Mediator. Activators also help to reduce the barrier to transcription that is posed by packaging into chromatin with recruitment of chromatin modifying and remodeling factors. Different activators work by varying mechanisms to achieve the goal of stimulated Pol II transcription¹⁰.

Pre-assembled human chromatin, prepared as above and conjugated to magnetic beads, was used to profile recruitment events mediated by a transcriptional activator. We used the Gal4-

VP16 fusion construct, a strong transcriptional activator frequently used for biochemical analyses of transcription²⁹. Chromatin substrates were incubated either in the presence or absence of Gal4-VP16, along with an ATP-regeneration system and a nuclear extract prepared from HeLa cells. The HeLa cells were grown either in heavy or light arginine/lysine medium for multiple passages (Fig. 1B), allowing full incorporation of the isotopically-enriched amino acids (validated by mass spectrometry, data not shown). Nuclear extracts were prepared based on previous protocols to generate a transcription- and splicing-competent extract¹⁶.

Non-covalent, stably-bound factors that were bound to the chromatin substrates were captured by magnetic isolation and washing of the bead-bound chromatin. Proteins were stripped from the DNA with a high-salt elution from beads, equal volumes of heavy and light samples were mixed, and all proteins were TCA precipitated. TCA pellets were resuspended, trypsinized and analyzed by LC-MS/MS for peptide identification and quantification. Each sample, in the presence or absence of Gal4-VP16 activator (Figure 2A), was incubated in both heavy and light medium, mixed accordingly, and the results were analyzed.

The presence of Gal4-VP16 bound to the chromatin substrate produced an enriched profile of transcription-associated factors (Fig. 2B-D, Table S4). Most-notably and with consistent reproducibility was the enrichment of subunits of the Mediator complex. The VP16 activation domain interacts directly with the Med17 and Med25 subunits of Mediator^{30–31}, facilitating recruitment of the entire ~1 MDa mediator complex of up to 30 subunits. Med17 and Med25 proteins were identified as enriched in both replicates of the Gal4-VP16 SILAC experiment, as were 10 other Mediator subunits (Figure 2B). Mediator is composed of four “modules”, based on biochemical and structural studies (Figure 2C)^{32–33}. Three modules, the head, middle, and tail, compose a stable, core complex generally purified in stoichiometric amounts of each subunit. The 12 subunits of Mediator that we reproducibly identified as enriched by Gal4-VP16 were all part of the core Mediator, spanning the three core modules. Since Mediator core operates as a homogeneous complex, we expect that the 12 subunits that were reproducibly-identified were representative of full core Mediator being recruited to chromatin by Gal4-VP16, based on the sampling efficiency of the proteomic experiment. This expectation is supported by the fact that each replicate experiment identified additional core Mediator subunits, bringing the total identified to 20/26 (Figure 2C). In addition, this analysis identified subunits of the Mediator kinase module, which is known to be dissociable.

Based on the identification of additional Mediator subunits in single experiments, including the dynamically-associated kinase module, we analyzed the lists of Gal4-VP16 enriched proteins in each individual SILAC experiment to identify further trends related to transcriptional activation on chromatin. These results identified many additional proteins involved in chromatin-specific transcriptional regulation (Fig. 2D), including subunits of the human SWI/SNF, MLL, INO80, and CAF-I complexes. In addition, multiple proteins that promote Pol II catalytic activity itself were identified, including subunits of the polymerase associated factor-1 (PAF1) and TFIIH complexes. We also noted that multiple proteins involved in inhibiting transcription were identified, including the HEXIM1 protein that

mediates Pol II polymerase pausing. Gene Ontology analysis of all proteins enriched >1.5-fold in each of the experiments demonstrated that the majority of the most-enriched molecular function pathways are those involving transcription mechanisms (Supplementary Table S1), further confirming that the experimental design fosters transcription-associated mechanisms on the chromatin substrate.

The conditions for this experiment clearly support certain key events of transcriptional activation, but there is not a clear signature for a fully-active promoter, such as robust Gal4-VP16 dependent recruitment of core transcriptional machinery such as Pol II itself. Human chromatin requires histone modifications, acetylation and methylation, in particular, to be fully-activated for transcription. Lack of histone modifications and perhaps certain transcription inhibitors may together keep the substrates in the above experiment positioned at an early stage of transcriptional activation. In addition, the p300 histone acetyltransferase competes with the general transcription factor TFIID and prevents downstream Pol II recruitment³⁴.

Recruitment of the Mediator by Gal4-VP16 to the chromatin substrate is a proof of principle that this system is amenable to profiling transcription-activation events in the context of chromatin using the SILAC-MS approach. Additional identified activator-enriched proteins suggest that this is a powerful experimental design for further study of transcription mechanisms in a chromatin context.

Using Tethered RNAs on chromatin to understand lncRNA-mediated genome regulation

As a further test of the utility of profiling different chromatin substrates through SILAC-MS, we focused on a long noncoding RNA that directs chromatin repression. We set out to profile the RepA domain of the Xist lncRNA and its interactions with proteins on chromatin. A number of studies have recently profiled the proteins that are cross-linked and isolated with mouse Xist during oligo-capture of the RNA from cells with an inactive X chromosome^{35–37}. With these previous studies to compare to, we designed a method to profile the interactions that can be fostered by Xist RepA on chromatin using our above-described SILAC-MS profiling method for reconstituted chromatin.

Though there is much evidence for Xist associating with chromatin from RNA capture approaches^{38–39}, the exact nature of Xist-chromatin interaction is not well-defined. To place RepA on chromatin for subsequent analysis of the interactions that are assembled on the RNA-chromatin substrate, we used an RNA tethering strategy, which we had previously designed to place the lncRNA HOTAIR on chromatin¹⁸. Briefly, this involves using a fusion protein containing a LexA DNA-binding domain fused to an RNA aptamer-binding protein, in this case the bacteriophage PP7 coat protein/aptamer pair. The DNA template used in the SILAC profiling contained six tandem LexA sites upstream of the Gal4 binding sites, which allow for the LexA-PP7 protein to bind. Pre-binding of LexA-PP7 to tagged RNA is performed, followed by a second step of tethering the PP7-aptamer containing RNA to the chromatin. RNAs are transcribed in vitro using T7 polymerase from PCR products that add the PP7 aptamer to the RNA. This tethering scheme typically results in 1–2 molecules of RNA tethered to a chromatin substrate (data not shown). To rule out non-specific RNA binding protein interactions that do not associate preferentially to RepA, we used a control

RNA, the antisense sequence of the Luciferase mRNA (anti-Luc) which we have used previously¹⁸.

Using SILAC nuclear extracts we profiled the proteins that stably assembled on a RepA-tethered chromatin substrate versus the anti-Luc-tethered control chromatin (Figure 3B). The proteins reproducibly-enriched (label-swap replicates, Table S5) identified a select set of primarily RNA-binding proteins and transcriptional regulators (Figure 3C,D). As a proof of principle that the *in vitro*-transcribed, tethered RepA behaves similarly to RepA in the cell, we identified proteins that have previously been seen in the proteomic investigations of Xist interactions *in vivo*. Most-notably, TDP-43, ELAV1, FUBP1, and FUBP3 have previously been found by proteomic analysis of proteins cross-linked to mouse Xist in living cells with an inactive ChrX^{35–37}. These studies did not perform any follow-up experiments on these interactions, instead focusing on other proteins identified. TDP-43 enrichment on RepA-chromatin was validated by Western analysis, as was enrichment of the RNA binding protein PDCD7 (also known as U11 59K) (Figure 3E). To further validate by an additional orthogonal approach, we cross-linked the breast cancer cell line MCF-7, which expresses Xist, and performed an immunoprecipitation for TDP-43, followed by reverse transcription-qPCR (Figure 3F). We found that TDP-43 IP specifically isolated Xist RepA RNA, compared to controls of GAPDH and the noncoding 7SL RNA, validating that this interaction occurs *in vivo*. Previous reports of TDP-43 interaction with mouse Xist has suggested a role in Xist stability⁴⁰, though TDP-43 mediates many processes in RNA metabolism.

Analysis of the RepA-enriched proteins from each replicate (Table 1) provides additional potential mechanisms for Xist-mediated repression. Previous work has highlighted the essential role of the protein Spen/SHARP/MINT in X inactivation^{35–37, 41}. Our approach identified Spen but did not demonstrate an enrichment with RepA compared to the control RNA. Spen binds to many RNAs and our results may be reflective of this low-specificity. However, Spen interacts directly with the N-CoR repressor proteins NCOR1 and NCOR2/SMRT⁴², both of which were identified as enriched in the RepA-chromatin samples (Table 1).

Gene Ontology Analysis of proteins in Table 1 showed that the most-enriched molecular function pathways were for RNA-binding and transcriptional regulation (Supplementary Table S2), correlating with the molecular nature of Xist as a noncoding RNA transcriptional regulator. Interestingly, both positive and negative regulators of transcription were identified in the RepA-enriched experiments. Negative transcriptional regulators may work directly with Xist to silence the X Chromosome. It is possible that interactions of Xist RepA with positive transcriptional regulators, such as Cyclin H and TFIIF (Table 1) lead to interference with the normal positive role these factors play in transcription, similar to the repression mechanism of FIR described above.

Our RNA tethering approach demonstrates that we can place RepA on reconstituted chromatin to foster known RNA-protein interactions. This approach also enabled the identification of new interactions that may be facilitated by the chromatin context. This system is highly-manipulatable, allowing for the study of how the Xist RNA can interact

with different types of chromatin, including diverse histone modification patterns and histone variants that can be introduced with the reconstituted system. This powerful flexibility is relevant in the context of dosage compensation, because Xist must interact with many types of chromatin, bearing specific modification patterns, as it works across an entire chromosome to convert active chromatin to a repressed state. In addition, our RepA-chromatin profiling results are consistent with growing evidence that Xist works by multiple mechanisms to achieve dosage compensation, likely harnessing multiple regions of the long RNA to scaffold interactions that facilitate these mechanisms.

DISCUSSION

Chromatin, as a complex of histones and DNA, provides a multi-faceted platform for assembly of genome regulators. The comprehensive study of the assembly of proteins that form in a specific chromatin context requires both methods for isolation of the native complexes from the cell and controlled biochemical approaches that allow for manipulation and homogeneity that is difficult to achieve *in vivo*. Efforts to profile the proteins in mammalian cells that interact with specific types of chromatin and specific genomic loci have improved in recent years^{43–45} and have great potential. We present a framework for complementary biochemical and proteomic approaches that can, together with *in vivo* analysis, lead toward a comprehensive picture of how chromatin can serve as a platform for transcription or RNA-dependent genome regulation.

Our system to establish the transcriptional activation mechanism on human chromatin using the Gal4-VP16 system highlights the potential to use SILAC-MS as a robust readout for studying the biochemistry of this molecular process. We demonstrate the efficient and specific recruitment of the Mediator complex to establish the early steps in transcription activation, as has been observed for similarly-designed biochemistry experiments using label-free MudPIT mass spectrometry⁶. This leaves the chromatin poised for further regulatory steps that require specific histone modifications to remodel chromatin and provide binding sites for specific transcriptional co-activators. There are many variations of transcriptional activation that can be explored with this system. For example, the mechanisms of different transcriptional activators, other than VP16, can be studied in a more-comprehensive way on chromatin by our SILAC-MS approach, similar to the recent label-free approach used to study estrogen receptor mechanism⁸. Other potential analyses that can be performed using this approach include the study of the molecular mechanism behind pharmacological agents that target the transcription initiation machinery such as Mediator⁴⁶.

LncRNAs that work on chromatin, whether at their site of transcription or elsewhere, are involved in many biological processes from development to disease¹¹. In general, the molecular mechanisms underlying lncRNAs are poorly understood, therefore proteomic methods have the potential to uncover binding partners that are candidates to function with the lncRNA. The tethering approach we describe is a useful way to bypass an unknown mechanism of chromatin interaction, such as the way that Xist interacts across the X Chromosome, which is still unclear. Our tethering approach led to enrichment of some proteins that had been previously found to associate with RepA, such as the FUBP family,

TDP-43, and ELAVL1^{35–37} while identifying additional novel protein interactors. The challenge of spreading across and silencing an entire chromosome is thought to be met through harnessing multiple pathways by Xist-protein interactions, which may be reflected in the many binding events we and others have identified. The diversity of these binding profiles also potentially hints at the fact that Xist likely acts as a “dynamic scaffold” where protein interactions occur transiently, rather than generating a singular homogeneous RNP. Employing cross-linking approaches may be one way to capture at an even deeper level weak and transient, yet physiologically important, interactions.

From the reproducibly-enriched proteins associated with RepA-chromatin, one potential mechanism for how Xist may mediate gene repression is suggested. The FUBP proteins are known to interact with a protein called PUF60, or alternatively named FUBP-interacting repressor (FIR)⁴⁷. FIR has been shown to directly interact with the TFIID general transcription factor to interfere with transcription initiation⁴⁷. PUF60/FIR was also reproducibly-enriched on RepA-chromatin, likely through interaction with FUBP1/3. These interactions may serve as one “transcription repressor” module to mediate shutdown of gene expression on the inactivating X chromosome. The FIR/FUBP interaction with Xist was also identified in a recent study using pulldown of a biotinylated synthetic RNA with a RepA sequence included⁴⁸, suggesting that the RNA itself may mediate the interaction directly, rather than through DNA/chromatin. FUBPs are able to bind both single-strand DNA and RNA⁴⁹ and may in this case be directly binding the Xist sequence.

Tethering of RNA to chromatin is also a way to mimic the generation of an RNA that occurs as the polymerase transcribes a nascent transcript. In the future, this approach could be used to study lncRNAs that work in *cis*, at the site of transcription. In addition, other processes that occur on chromatin and require assembly of proteins on nascent RNAs, such as splicing, can be profiled from the approaches we have described. A growing amount of evidence suggests that splicing activity feeds back to chromatin⁵⁰, potentially marking specific regions as splice sites for further co-transcriptional splicing events. The full extent of histone modification and reader interactions that interface splicing and chromatin has yet to be determined and presents a significant analytical challenge.

Conclusions

Biochemical methods such as the ones presented here are important tools to uncover and dissect the complex regulatory mechanisms of genome regulation. The approaches presented here allow for a more-comprehensive profiling of interactions that occur on chromatin in a manipulatable manner that is amenable to introduction of the many types of chromatin modification that tune the genome. Using these methods can shed light on general mechanisms of transcriptional control, the complex process of dosage compensation, and many other molecular mechanisms at work in the nucleus.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We thank M. Dzieciatkowska of the University of Colorado School of Medicine Biological Mass Spectrometry Core Facility for assistance with mass spectrometry. This work was supported by NIH grant R35GM119575 (A.M.J.), a research grant from the Linda Crnic Institute for Down Syndrome (A.M.J.), the University of Colorado-Denver Cancer Biology Training grant T32CA190216 (A.M.P.), and an RNA Scholar Award from the University of Colorado RNA Bioscience Initiative (M.M.B).

References

1. Rothbart SB; Strahl BD, Interpreting the language of histone and DNA modifications. *Biochim Biophys Acta* 2014, 1839 (8), 627–43. [PubMed: 24631868]
2. Johnson A; Wu R; Peetz M; Gygi SP; Moazed D, Heterochromatic Gene Silencing by Activator Interference and a Transcription Elongation Barrier. *J Biol Chem* 2013, 288 (40), 28771–28782. [PubMed: 23940036]
3. Sikorski TW; Joo YJ; Ficarro SB; Askenazi M; Buratowski S; Marto JA, Proteomic Analysis Demonstrates Activator- and Chromatin-specific Recruitment to Promoters. *J Biol Chem* 2012, 287 (42), 35397–408. [PubMed: 22902623]
4. Nikolov M; Stutzer A; Mosch K; Krasauskas A; Soeroes S; Stark H; Urlaub H; Fischle W, Chromatin affinity purification and quantitative mass spectrometry defining the interactome of histone modification patterns. *Mol Cell Proteomics* 2011, 10 (11), M110 005371.
5. Bartke T; Vermeulen M; Xhemalce B; Robson SC; Mann M; Kouzarides T, Nucleosome-interacting proteins regulated by DNA and histone methylation. *Cell* 2010, 143 (3), 470–84. [PubMed: 21029866]
6. Lin JJ; Lehmann LW; Bonora G; Sridharan R; Vashisht AA; Tran N; Plath K; Wohlschlegel JA; Carey M, Mediator coordinates PIC assembly with recruitment of CHD1. *Genes Dev* 2011, 25 (20), 2198–209. [PubMed: 21979373]
7. Chen XF; Lehmann L; Lin JJ; Vashisht A; Schmidt R; Ferrari R; Huang C; McKee R; Mosley A; Plath K; Kurdiani SK; Wohlschlegel J; Carey M, Mediator and SAGA have distinct roles in Pol II preinitiation complex assembly and function. *Cell Rep* 2012, 2 (5), 1061–7. [PubMed: 23177621]
8. Foulds CE; Feng Q; Ding C; Bailey S; Hunsaker TL; Malovannaya A; Hamilton RA; Gates LA; Zhang Z; Li C; Chan D; Bajaj A; Callaway CG; Edwards DP; Lonard DM; Tsai SY; Tsai MJ; Qin J; O'Malley BW, Proteomic analysis of coregulators bound to ERalpha on DNA and nucleosomes reveals coregulator dynamics. *Mol Cell* 2013, 51 (2), 185–99. [PubMed: 23850489]
9. Naar AM; Lemon BD; Tjian R, Transcriptional coactivator complexes. *Annu Rev Biochem* 2001, 70, 475–501. [PubMed: 11395415]
10. Ebmeier CC; Taatjes DJ, Activator-Mediator binding regulates Mediator-cofactor interactions. *Proc Natl Acad Sci U S A* 2010, 107 (25), 11283–8. [PubMed: 20534441]
11. Bonasio R; Shiekhattar R, Regulation of transcription by long noncoding RNAs. *Annu Rev Genet* 2014, 48, 433–55. [PubMed: 25251851]
12. Balas MM; Johnson AM, Exploring the mechanisms behind long noncoding RNAs and cancer. *Noncoding RNA Res* 2018, 3 (3), 108–117. [PubMed: 30175284]
13. Gendrel AV; Heard E, Noncoding RNAs and epigenetic mechanisms during X-chromosome inactivation. *Annu Rev Cell Dev Biol* 2014, 30, 561–80. [PubMed: 25000994]
14. Monfort A; Wutz A, Progress in understanding the molecular mechanism of Xist RNA function through genetics. *Philos Trans R Soc Lond B Biol Sci* 2017, 372 (1733).
15. Pintacuda G; Young AN; Cerase A, Function by Structure: Spotlights on Xist Long Non-coding RNA. *Front Mol Biosci* 2017, 4, 90. [PubMed: 29302591]
16. Zhou Z; Reed R, Purification of functional RNA-protein complexes using MS2-MBP. *Curr Protoc Mol Biol* 2003, Chapter 27, Unit 27 3.
17. Li G; Widom J, Nucleosomes facilitate their own invasion. *Nat Struct Mol Biol* 2004, 11 (8), 763–9. [PubMed: 15258568]

18. Meredith EK; Balas MM; Sindy K; Haislop K; Johnson AM, An RNA matchmaker protein regulates the activity of the long noncoding RNA HOTAIR. *RNA* 2016, 22 (7), 995–1010. [PubMed: 27146324]
19. Cho EJ; Buratowski S, Evidence that transcription factor IIB is required for a post-assembly step in transcription initiation. *J Biol Chem* 1999, 274 (36), 25807–13. [PubMed: 10464320]
20. An W; Roeder RG, Reconstitution and transcriptional analysis of chromatin in vitro. *Methods Enzymol* 2004, 377, 460–74. [PubMed: 14979045]
21. Vary JC Jr.; Fazzio TG; Tsukiyama T, Assembly of yeast chromatin using ISWI complexes. *Methods Enzymol* 2004, 375, 88–102. [PubMed: 14870661]
22. Johnson A; Li G; Sikorski TW; Buratowski S; Woodcock CL; Moazed D, Reconstitution of heterochromatin-dependent transcriptional gene silencing. *Mol Cell* 2009, 35 (6), 769–81. [PubMed: 19782027]
23. Nielsen PA; Olsen JV; Podtelejnikov AV; Andersen JR; Mann M; Wisniewski JR, Proteomic mapping of brain plasma membrane proteins. *Mol Cell Proteomics* 2005, 4 (4), 402–8. [PubMed: 15684408]
24. Hill RC; Calle EA; Dzieciatkowska M; Niklason LE; Hansen KC, Quantification of extracellular matrix proteins from a rat lung scaffold to provide a molecular readout for tissue engineering. *Mol Cell Proteomics* 2015, 14 (4), 961–73. [PubMed: 25660013]
25. Cox J; Matic I; Hilger M; Nagaraj N; Selbach M; Olsen JV; Mann M, A practical guide to the MaxQuant computational platform for SILAC-based quantitative proteomics. *Nat Protoc* 2009, 4 (5), 698–705. [PubMed: 19373234]
26. Niranjankumari S; Lasda E; Brazas R; Garcia-Blanco MA, Reversible cross-linking combined with immunoprecipitation to study RNA-protein interactions in vivo. *Methods* 2002, 26 (2), 182–90. [PubMed: 12054895]
27. Carey M; Leatherwood J; Ptashne M, A potent GAL4 derivative activates transcription at a distance in vitro. *Science* 1990, 247 (4943), 710–2. [PubMed: 2405489]
28. Thastrom A; Bingham LM; Widom J, Nucleosomal locations of dominant DNA sequence motifs for histone-DNA interactions and nucleosome positioning. *J Mol Biol* 2004, 338 (4), 695–709. [PubMed: 15099738]
29. Lin YS; Carey MF; Ptashne M; Green MR, GAL4 derivatives function alone and synergistically with mammalian activators in vitro. *Cell* 1988, 54 (5), 659–64. [PubMed: 3044607]
30. Ito M; Yuan CX; Malik S; Gu W; Fondell JD; Yamamura S; Fu ZY; Zhang X; Qin J; Roeder RG, Identity between TRAP and SMCC complexes indicates novel pathways for the function of nuclear receptors and diverse mammalian activators. *Mol Cell* 1999, 3 (3), 361–70. [PubMed: 10198638]
31. Yang F; DeBeaumont R; Zhou S; Naar AM, The activator-recruited cofactor/Mediator coactivator subunit ARC92 is a functionally important target of the VP16 transcriptional activator. *Proc Natl Acad Sci U S A* 2004, 101 (8), 2339–44. [PubMed: 14983011]
32. Conaway RC; Conaway JW, The Mediator complex and transcription elongation. *Biochim Biophys Acta* 2013, 1829 (1), 69–75. [PubMed: 22983086]
33. Poss ZC; Ebmeier CC; Taatjes DJ, The Mediator complex and transcription regulation. *Crit Rev Biochem Mol Biol* 2013, 48 (6), 575–608. [PubMed: 24088064]
34. Black JC; Choi JE; Lombardo SR; Carey M, A mechanism for coordinating chromatin modification and preinitiation complex assembly. *Mol Cell* 2006, 23 (6), 809–18. [PubMed: 16973433]
35. McHugh CA; Chen CK; Chow A; Surka CF; Tran C; McDonel P; Pandya-Jones A; Blanco M; Burghard C; Moradian A; Sweredoski MJ; Shishkin AA; Su J; Lander ES; Hess S; Plath K; Guttman M, The Xist lncRNA interacts directly with SHARP to silence transcription through HDAC3. *Nature* 2015, 521 (7551), 232–6. [PubMed: 25915022]
36. Minajigi A; Froberg JE; Wei C; Sunwoo H; Kesner B; Colognori D; Lessing D; Payer B; Boukhali M; Haas W; Lee JT, Chromosomes. A comprehensive Xist interactome reveals cohesin repulsion and an RNA-directed chromosome conformation. *Science* 2015, 349 (6245).

37. Chu C; Zhang QC; da Rocha ST; Flynn RA; Bharadwaj M; Calabrese JM; Magnuson T; Heard E; Chang HY, Systematic discovery of Xist RNA binding proteins. *Cell* 2015, 161 (2), 404–16. [PubMed: 25843628]
38. Simon MD; Pinter SF; Fang R; Sarma K; Rutenberg-Schoenberg M; Bowman SK; Kesner BA; Maier VK; Kingston RE; Lee JT, High-resolution Xist binding maps reveal two-step spreading during X-chromosome inactivation. *Nature* 2013, 504 (7480), 465–9. [PubMed: 24162848]
39. Engreitz JM; Pandya-Jones A; McDonel P; Shishkin A; Sirokman K; Surka C; Kadri S; Xing J; Goren A; Lander ES; Plath K; Guttman M, The Xist lncRNA exploits three-dimensional genome architecture to spread across the X chromosome. *Science* 2013, 341 (6147), 1237973. [PubMed: 23828888]
40. Polymenidou M; Lagier-Tourenne C; Hutt KR; Huelga SC; Moran J; Liang TY; Ling SC; Sun E; Wanciewicz E; Mazur C; Kordasiewicz H; Sedaghat Y; Donohue JP; Shiue L; Bennett CF; Yeo GW; Cleveland DW, Long pre-mRNA depletion and RNA missplicing contribute to neuronal vulnerability from loss of TDP-43. *Nat Neurosci* 2011, 14 (4), 459–68. [PubMed: 21358643]
41. Moindrot B; Cerase A; Coker H; Masui O; Grijzenhout A; Pintacuda G; Schermelleh L; Nesterova TB; Brockdorff N, A Pooled shRNA Screen Identifies Rbm15, Spen, and Wtap as Factors Required for Xist RNA-Mediated Silencing. *Cell Rep* 2015, 12 (4), 562–72. [PubMed: 26190105]
42. Ariyoshi M; Schwabe JW, A conserved structural motif reveals the essential transcriptional repression function of Spen proteins and their role in developmental signaling. *Genes Dev* 2003, 17 (15), 1909–20. [PubMed: 12897056]
43. Becker JS; McCarthy RL; Sidoli S; Donahue G; Kaeding KE; He Z; Lin S; Garcia BA; Zaret KS, Genomic and Proteomic Resolution of Heterochromatin and Its Restriction of Alternate Fate Genes. *Mol Cell* 2017, 68 (6), 1023–1037 e15. [PubMed: 29272703]
44. Fujita T; Fujii H, Isolation of specific genomic regions and identification of associated molecules by engineered DNA-binding molecule-mediated chromatin immunoprecipitation (enChIP) using CRISPR. *Methods Mol Biol* 2015, 1288, 43–52. [PubMed: 25827874]
45. Dejardin J; Kingston RE, Purification of proteins associated with specific genomic Loci. *Cell* 2009, 136 (1), 175–86. [PubMed: 19135898]
46. Phillips AJ; Taatjes DJ, Small molecule probes to target the human Mediator complex. *Isr J Chem* 2013, 53 (8), 588–595. [PubMed: 25705053]
47. Liu J; He L; Collins I; Ge H; Libutti D; Li J; Egly JM; Levens D, The FBP interacting repressor targets TFIID to inhibit activated transcription. *Mol Cell* 2000, 5 (2), 331–41. [PubMed: 10882074]
48. Pintacuda G; Wei G; Roustan C; Kirmizitas BA; Solcan N; Cerase A; Castello A; Mohammed S; Moindrot B; Nesterova TB; Brockdorff N, hnRNPK Recruits PCGF3/5-PRC1 to the Xist RNA Repeat to Establish Polycomb-Mediated Chromosomal Silencing. *Mol Cell* 2017, 68 (5), 955–969 e10. [PubMed: 29220657]
49. Zhang J; Chen QM, Far upstream element binding protein 1: a commander of transcription, translation and beyond. *Oncogene* 2013, 32 (24), 2907–16. [PubMed: 22926519]
50. Naftelberg S; Schor IE; Ast G; Kornbliht AR, Regulation of alternative splicing through coupling with transcription and chromatin structure. *Annu Rev Biochem* 2015, 84, 165–98. [PubMed: 26034889]
51. Tsai KL; Tomomori-Sato C; Sato S; Conaway RC; Conaway JW; Asturias FJ, Subunit architecture and functional modular rearrangements of the transcriptional mediator complex. *Cell* 2014, 157 (6), 1430–44. [PubMed: 24882805]
52. Mertins P; Qiao JW; Patel J; Udeshi ND; Clauser KR; Mani DR; Burgess MW; Gillette MA; Jaffe JD; Carr SA, Integrated proteomic analysis of post-translational modifications by serial enrichment. *Nat Methods* 2013, 10 (7), 634–7. [PubMed: 23749302]
53. Udeshi ND; Svinkina T; Mertins P; Kuhn E; Mani DR; Qiao JW; Carr SA, Refined preparation and use of anti-diglycine remnant (K-epsilon-GG) antibody enables routine quantification of 10,000s of ubiquitination sites in single proteomics experiments. *Mol Cell Proteomics* 2013, 12 (3), 825–31. [PubMed: 23266961]

Figure 1. Profiling regulated protein complex assembly on reconstituted chromatin platforms. A) Assembly of protein complexes on chromatin regulates many genomic processes. B) Scheme for SILAC labeling and profiling of the interactions of nuclear proteins with specific reconstituted chromatin substrates. C) DNA template used in all chromatin profiling experiments in the current study: a biotinylated PCR product containing an adenovirus E4 promoter with upstream Gal4 binding sites and LexA sequences flanked on each side by 601 nucleosome positioning sequences. D) General chromatin assembly scheme. *E. coli*-purified human histones, assembled into octamers were enzymatically assembled into chromatin using the DNA template described above using histone chaperone and nucleosome spacing factors. E) Limited micrococcal nuclease digestion of assembled chromatin demonstrating nucleosome-protection of DNA in a regular array.

Figure 2. SILAC-based profiling of early steps in transcriptional activation by Gal4-VP16.
A) Scheme of chromatin substrates compared by SILAC-mass spectrometry. B) SILAC profile of proteins enriched by Gal4-VP16 binding to chromatin. Replicate experiments with swapped, labeled nuclear extracts were compared. Mediator subunits are highlighted in red.
C) Analysis of Mediator subunits identified in the SILAC profiling of GV enrichment on chromatin. (left) Subunits of Mediator subcomplexes listed with indicated identification in each SILAC experiment. (right) Cryo-EM structure of core Mediator from ⁵¹ (permission will be obtained upon manuscript acceptance) with appropriate color-coding. D) Additional selected chromatin and transcription-associated proteins identified as enriched with Gal4-VP16 in either heavy (pink) or light (blue) sample only.

Figure 3. Profiling of a long noncoding RNA involved in mammalian X chromosome inactivation. A) Dosage compensation in mammals is dependent on the lncRNA Xist for silencing of one X chromosome (ChrX) to an inactive state (Xi). (bottom) domain structure for Xist lncRNA, depicting the repeat regions A-E. Expanded view of the RepA transcript, which can be transcribed independently of full-length Xist. RepA harbors critical RNA regions important for chromosome silencing. B) Schematic of the two types of chromatin platforms profiled by SILAC-MS. RNAs are tethered to reconstituted chromatin via a LexA-PP7 fusion protein that binds LexA DNA sites on the chromatin and also binds PP7 aptamer-tagged in vitro transcribed RNAs. The control long noncoding RNA in this experiment is derived from the antisense sequence of the luciferase mRNA, which is similar in length and GC content to RepA. C) Profile of replicated SILAC-MS experiments relative to enrichment in the RepA-containing chromatin samples. D) Expanded view from C (red box) highlighted the specifically RepA-enriched proteins. E) Western validation for RepA enrichment of TDP-43

and PDCD7. Note, the most enriched band in the PDCD7 blot likely corresponds to a low-abundance multiply-ubiquitinated form identified in previous surveys of ubiquitination sites^{52–53}. F) RNA immunoprecipitation of TDP-43 validating its interaction with RepA in MCF7 female breast cancer cells expressing Xist.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 1.

Table of RepA-enriched proteins.

Gene IDs for all proteins identified as 1.5-fold or greater enriched in the RepA-tethered chromatin sample. Gene ID column highlights indicate a protein ID'd in only the Light (red) or Heavy (blue) RepA sample. Abridged Description column highlights in green indicate annotated RNA binding proteins. Additional Notes column color-coding indicates proteins that are related either by homology or physical association. ID'd in other studies column uses citations to indicate a protein that was identified in another study on Xist-protein interactions.

Gene ID	Abridged Description	Additional Notes (shading for related proteins)	ID'd in other studies
PDCD7	Programmed cell death 7	59K of U11 snRNP. Binds RBM40.	
FUBP1	Far upstream element-binding protein 1	Interacts with PUF60.	Refs. 37,48
RANGAP1	Ran GTPase-activating protein 1		
TIA1	Nucleolysin TIA-1	Binds U1 SnRNP. Related to TIAR	
CCNH	HUMAN Cyclin-H	TFIIH subunit.	
TARDBP(TDP-43)	TAR DNA binding protein		Refs. 37,48
HNRPF	Heterogeneous nuclear ribonucleoprotein F		
ELAVL1	ELAV-like protein 1	Related to CELF1.	Refs. 36,37
TIAL1	Nucleolysin TIAR	Related to TIA-1.	
ZGRF1	GRF-type zinc finger domain-containing protein 1		
HNRNPM	Heterogeneous nuclear ribonucleoprotein M		Refs. 35,37,48
CLOCK	Circadian locomotor output cycles protein kaput		
MYH10	Myosin-10		
PCF11	Pre-mRNA cleavage complex 2 protein		
PUF60 (FIR)	FUBP-interacting repressor. Poly(U)-binding-splicing factor.	Interacts with FUBPs.	Ref. 48
TFEB	Transcription factor EB		
FUBP3	Far upstream element-binding protein 3	Interacts with PUF60.	Refs. 37,48
BAP18	Chromatin complexes subunit	MLL and NURF component.	
MBNL1	Muscleblind-like protein 1		
KIF4A	Chromosome-associated kinesin		Ref. 36
SPTY2D1	SPT2 homolog	Histone chaperone coupled to transcription.	
SETD8	H4K20-specific histone methyltransferase	Marks heterochromatin, including Xi.	
NUP160	Nuclear pore complex protein		
RNF220	E3 ubiquitin-protein ligase		
SMARCA1	global transcription activator SNF2L	NURF component.	

Gene ID	Abridged Description	Additional Notes (shading for related proteins)	ID'd in other studies
GTF2F1	General transcription factor IIF subunit 1	Important for transcription elongation.	
CELF1	CUGBP Elav-like family member 1	Related to ELAVL1.	Refs. 35,37,48
SYF2	Pre-mRNA-splicing factor		
KAT5	Histone acetyltransferase	NuA4 catalytic subunit.	
NCOR1	Nuclear receptor corepressor 1	Interacts with Spen.	
NCOR2 (SMRT)	Nuclear receptor corepressor 2	Interacts with Spen.	
<u>Key</u>			
Only ID'd in Light	RNA Binding Protein		
Only ID'd in Heavy			