Supplemental Data Temporal Profiling of the Chromatin Proteome Reveals System-wide Responses to Replication Inhibition

Guennadi A. Khoudoli, Peter J. Gillespie, Graeme Stewart, Jens S. Andersen, Jason R. Swedlow, and J. Julian Blow

Supplemental Experimental Procedures

Materials and Equipment

Trypsin Gold, Mass Spectrometry Grade, was purchased from Promega. Expression and purification of non-degradable geminin^{DEL} were described in [S1]. Mcm2, Mcm3, Smc2 ,and Nup153 antibodies were as described previously [S2]. Rabbit polyclonal antiserum was raised against a synthetic peptide corresponding to the alpha-peptide of *Xenopus* pol δ (GTKQA SIMGFFQKK) as described [S3].

Xenopus Chromatin Manipulations

Mitotically arrested *Xenopus* egg extract preparation, chromatin reconstitution, and isolation was carried out as previously described [S4, S5], with small modifications. Extracts were characterized before being used for proteomic data collection and were rejected if they did not support efficient nuclear assembly and DNA replication, with a sharp entry and exit into S phase. Defrosted extracts were precleared by centrifugation at 20,000 g for 15 min at 4°C. The assembly reactions including 6000 sperm nuclei/ μ I were combined on ice and released into interphase by addition of 0.3 mM CaCl₂ and transfer to 23°C.

Samples of reconstituted chromatin were isolated at 5 min and then at 10 min intervals between 10 and 90 min. One hundred and fifty microliter aliquots of extract were diluted with 300 µl nuclear isolation buffer (NIB: 50 mM KCl, 50 mM Hepes [pH 7.6], 5 mM MgCl₂, 2 mM dithiothreitol [DTT], 0.5 mM spermidine, 0.15 mM spermine, and 1 µg/ml each leupeptin, pepstatin, and aprotinin), supplemented with 2.5 mM Mg-ATP and 0.1% Triton X-100, and chromatin was pelleted through a 30% sucrose cushion at 6000 g for 15 min at 4°C. Chromatin proteins were eluted in 90 µl XBE2 (10 mM K-Hepes [pH 7.7], 0.1 M KCl, 2 mM MgCl₂, 0.1 mM CaCl₂, 5 mM ethylene glycol tetraacetic acid [EGTA], and 50 mM sucrose), supplemented with 0.1% Triton X-100 and 0.5 M NaCl, and were then chloroform/methanol precipitated. Protein pellets were dissolved and propagated for trypsin digestion as previously described [S6]. So that technical variations could be minimized and enough protein could be accumulated for analysis, experiments were repeated with the same extract at least twice for all conditions, and samples from corresponding time points were pooled.

Mcm2- and Mcm3-associated chromatin proteins were isolated by immunoprecipitation as described [S2]. In brief, replicating chromatin was isolated and treated with micrococcal nuclease to digest DNA to < 160 bp fragments. After centrifugation, soluble material was immunoprecipitated with antibodies against Mcm2 or Mcm3 or with nonimmune antibodies coupled to protein G Sepharose beads. After the washing, bound protein was eluted with 500 mM NaCl and was analyzed by LC-MS/MS.

Mass Spectrometry and Quantification

LC-MS analysis was performed with a quadruple time-of-flight instrument (QSTAR-XL, ABI-MDS-Sciex, Toronto, Canada) or a linear ion trap Fouriertransform ion-cyclotron resonance mass spectrometer (LTQ-FT-ICR, Thermo-Finnigan, Bremen). Samples were loaded onto a fused silica capillary with a 75 μ m ID and an 8 μ m tip opening (New Objective, Woburn, Massachusetts) filled with Reprosil 3 μ m reverse phase material (Dr. Maisch, Ammerbuch, Germany). Peptides were eluted with a 140 min linear gradient of 92% buffer A (0.5% acetic acid in H₂O) to 50% buffer B (80% acetonitrile, 0.5% acetic acid in H₂O).

Mass spectra were collected automatically with information depended acquisition. The QSTAR-XL was operated in the pulsing mode to enhance the signal in both MS and MS/MS. Precursor ion spectra (from m/z 350–1500) were collected for 1 s. The four most intense ions were isolated and product ion spectra (m/z 70–1500) were collected for 1.5 s. The LTQ-FT-ICR instrument was also operated in the data-dependent mode to acquire high-resolution precursor ion spectra (from m/z 300–1500, R = 25,000, and ion accumulation to a target value of 5,000,000) in the ICR cell. The five most intense ions were sequentially isolated for accurate mass measurements by SIM scans (10 Da mass window, R = 50,000, and a target accumulation value of 50,000). The ions were simultaneously fragmented in the

linear ion trap with a normalized collision energy setting of 27% and a target value of 2,000.

Combined peak lists were searched in the NCBI database with the Mascot program (Matrix Science, London, United Kingdom) and specifying Xenopus laevis species. QSTAR-XL data were searched with initial mass tolerance of 0.2 Da for precursors and product ions. Iterative calibration algorithms on the basis of identified peptides were used to achieve an average absolute mass accuracy of better than 20 ppm in both precursors and product ions. Proteins identified with a combined peptide score of higher than 60 were considered significant, and lower-scoring proteins were manually verified or rejected. LTQ-FT-ICR data were searched with a peptide mass tolerance of 5 ppm and a fragment mass tolerance of 0.8 Da. Iterative calibration algorithms on the basis of identified peptides resulted in an average absolute peptide mass accuracy of better than 1 ppm. Stringent criteria were used for protein identification based on the LTQ-FT-ICR data: at least two matching peptides per protein, a mass accuracy within 3 ppm, a Mascot score for individual ions of > 20, and a delta score of > 5. In those cases where protein identification was the only target, the public version of Phenyx software (http://www.phenyx-ms.com) was also employed.

MSQuant software [S7] was used to extract information from the Mascot HTML database search files and to manually validate the certainty in peptide identification, and in peptide abundance, on the basis of extracted ion chromatograms. A few proteins with a single high-scored peptide identified in at least seven time points were included in the analysis. To avoid rejection of polypeptides transiently associating with chromatin, we also analyzed proteins identified by only two or more highly scoring peptides at a single time point.

Mass-spectrometric analysis was repeated three times for the replicating chromatin and once for chromatin reconstituted in the presence of replication inhibitors. Temporal profiles from technical repeats were averaged at the protein level.

Clustering of Temporal Profiles

For each protein identified, the extracted ion current of up to the 12 most intense peptides were linearly transformed to the same mean value and averaged at each time point. The resulting temporal profiles were subjected to two rounds of smoothing (Figure 1B) and normalized so that over the time series the maximum abundance of each protein was set to 1. Proteins were combined into a single entry if the homology between related isoforms was > 93%. The resulting temporal profiles from replicating chromatin were clustered with fuzzy c mean (FCM) soft clustering procedure available as Mfuzz toolbox for R [S8]. The software package Cluster 3.0 [S9] was used for hierarchical cluster analysis.

DAVID [S10] was used to carry out annotation term enrichment analysis and functional clustering of FCM groups. The annotation categories are considered to be strongly enriched if p values (Figure 2) associated with them are equal or smaller than 0.05. The overall group enrichment score (Figure 2: GES column) is the negative log of the geometric mean of the statistical significance (p value) for all members of the cluster. For statistical calculation, all genes that were identified in our experiments were used as a background.

To estimate the reproducibility of temporal profiles, we used 276 proteins identified in three LCMS replicates of untreated chromatin and separated them into 12 FCM clusters. We defined the level of reproducibility (Rp) for cluster as a percentage of cases where at least two individual profiles we present in a cluster together with corresponding averaged one.

References

- S1. Ferenbach, A., Li, A., Brito-Martins, M., and Blow, J.J. (2005). Functional domains of the Xenopus replication licensing factor Cdt1. Nucleic Acids Res. 33, 316–324.
- S2. Gillespie, P.J., Khoudoli, G.A., Stewart, G., Swedlow, J.R., and Blow, J.J. (2007). ELYS/MEL-28 chromatin association coordinates nuclear



D

| Correlation | P interval | Percentage |
|-------------|------------|------------|
| Large | 0.5 - 1.0 | 72.7 |
| Medium | 0.3 - 0.5 | 6.0 |
| Small | 0.1 - 0.3 | 4.0 |

С



Figure S1. Data Presentation and Reproducibility Analysis

(A) Comparison of raw ion current values (blue) and smoothed values for Mcm2. Below, a heat map of the smoothed values is shown (red, highest; green, lowest; black, intermediate). Error bars indicate the SD.

(B) Reproducibility of protein behavior observed in three LCMS runs of untreated chromatin estimated by Pearson coefficient of correlation and graded according to Cohen [S11]. Coefficients of correlation were calculated pair wise for 276 proteins identified in all three LCMS runs of untreated chromatin.

(C) Regression analysis of correlation between reproducibility of temporal profiles assignment to cluster and amplitude of profiles variation presented by averaged cluster standard deviation. Only proteins identified in all three repeated LCMS analysis of untreated chromatin (276) were used for comparison.

pore complex assembly and replication Licensing. Curr. Biol. 17, 1657-1662.

S3. Hirano, T., Kobayashi, R., and Hirano, M. (1997). Condensins, chromosome condensation protein complexes containing XCAP-C, XCAP-E and a Xenopus homolog of the Drosophila Barren protein. Cell 89, 511–521.

- S4. Chong, J.P., Thommes, P., Rowles, A., Mahbubani, H.M., and Blow, J.J. (1997). Characterization of the Xenopus replication licensing system. Methods Enzymol. 283, 549–564.
- Murnion, M.E., Adams, R.R., Callister, D.M., Allis, C.D., Earnshaw, W.C., and Swedlow, J.R. (2001). Chromatin-associated protein phosphatase 1 regulates aurora-B and histone H3 phosphorylation. J. Biol. Chem. 276, 26656–26665.
- Andersen, J.S., Wilkinson, C.J., Mayor, T., Mortensen, P., Nigg, E.A., and Mann, M. (2003). Proteomic characterization of the human centrosome by protein correlation profiling. Nature 426, 570–574.
- Schulze, W.X., and Mann, M. (2004). A novel proteomic screen for peptide-protein interactions. J. Biol. Chem. 279, 10756–10764.
- Futschik, M.E., and Carlisle, B. (2005). Noise-robust soft clustering of gene expression time-course data. J. Bioinform. Comput. Biol. 3, 965–988.
- S9. Eisen, M.B., Spellman, P.T., Brown, P.O., and Botstein, D. (1998). Cluster analysis and display of genome-wide expression patterns. Proc. Natl. Acad. Sci. USA 95, 14863–14868.
- S10. Dennis, G., Jr., Sherman, B.T., Hosack, D.A., Yang, J., Gao, W., Lane, H.C., and Lempicki, R.A. (2003). DAVID: Database for Annotation, Visualization, and Integrated Discovery. Genome Biol. 4, 3.
- Cohen, J. (1988). Statistical Power Analysis for the Behavioral Sciences, Second Edition (Hillsdale, New Jersey: Lawrence Erlbaum Associates).

| Untreated | Geminin | Rocovitine | Peptide number |
|--------------------|-------------------------|--|---|
| | | | 69c 34g 21r gil562793 KDF4A. kinesin-like protein 1 107c 107g 41r gil1722855 SmC4. Structural maintenance of chromosome 4 ; 68c 106g 28r gil3764087 38 condensin XCAP-D2 subunit 42c 59g 28r gil3764087 38 condensation protein XCAP-G 106g 47r gil1722856 SmC2. Structural maintenance of chromosome 2 ; 23c 37g 14r gil33417152 Brra1-prov (91%ident to13S condensin XCAP-H)] |
| | | | 90c 66g 44r ri 51703659 ATP-binding cassette. sub-family F 280c 252g 216r ri 54035281 B4 protein, Linker histone B4 |
| | | A CONTRACTOR OF A CONTRACTOR OF | 30c 40g 34r gi 2499719 Vdac2. Voltage-dependent 192c 174g 163r gi 51258124 PKM2 protein |
| THE OWNER WATER OF | Distant in | | 52c 35a 24r ail214157 DNA-bindina protein _32c _25g _22r gi 1175534 Y-box binding protein 2-B |
| | | Contraction of the local division of the loc | 124c 109g 105r ri 46015180 Chain F. Crystallographic Studies 54c 46g 46r ri 50603936 replication-dependent histone H2A |
| | | | 53c 41g 37r gi 50603982 Ribosomal protein S23 . NG(64c 39g 41r gi 32450055 ribosomal protein large P2 |
| | | | 62c 0g 25r gi 559623 sperm-specific basic nuclear protein 3. SP3 70c 0g 32r gi 1384124 sperm-specific nuclear basic protein (SP4) |
| | | Land Street of | 30c 0g 26r ri 32450165 Vrk1-prov protein 16c _0g _9r gi 1729803 beta 1 subunit of heterotrimeric GTP-binding |
| | | | 59c 63a 44r cri 64771 Histone H1C (127c 144g 109r fi 28277305 Vim4 protein |
| | | | 1 76c 94 α 61r ril27370869 Canzb-prov protein 64c 74 α 54r ril33416674 Arpc3-prov protein |
| | and the second second | Contract of the | 142c 168g 107r r. 27696928 Similar to ARP2 actin-related protein 2 homol(110c 136g 82r r. 29126978 Actr3-prov protein |
| | | | 77c 73g 44r gil49522851 Actin related protein 2/3 complex. subunit 4, _47c _52g _35g _gil64909 nuclear actin-binding protein chain a |
| | A DECK | | 117c 179g 158r gi 65066 replication protein & (RPA) _43c _42g _32r _fi 48734630 RPA2, replication protein &2, 32kDa, |
| | and the second second | | 141c 130g 142r gi 623602 importin alpha 1a 141c 101g 106r gi 1079301 importin 2 - Africa |
| | | | 113c 122a 72r ai 4586285 DUF87 208c 172g 106r gi 4586287 DUF140 |
| | | The second second | 45c 68g 101r n 28422180 chromodomain helicase DNA binding protein 4, _14c _21g _21r n 51704120 AB21 protein |
| | | | 85c 92g 61r ri 47940311 X12PC protein 12c 32g 19r ri 62944550 egg envelope glycoprotein |
| | | | 81.C 58cr 86cr ril27371199 Ywhab-r 107c 78cr 88r ril27924430 Ywhae-r |
| | | | 109c _69g _73r gi 1360640 14-3-3 |
| | | and interest of the | 878c 820g 619r fi 54035188 cytoplasmic beta actin, MGC52661 |
| | No. of Concession, Name | | 1 - 420 + 210 + 521 + 311310320 + 3001 + 30010 + 205 - 300 + 31030 + 300 + 111310320 + 3001 + 300 + |
| | | A COLUMN TWO IS NOT | 21c 25a 21r ri 27694605 Psmbl-brow protein 13c 57a 34r ri 217694605 psmbl-brow protein 13c 57a 34r ri 51703928 proteasome (prosome, macropain) 26S subunit. |
| | | | 159c 158g 166g gil65167 Tubulin. aluha 7 359c 317g 366g gil54311209 Tubulin. beta 3 LOC495319 protein |
| | | | 121c 131g 137r ri 50414753 Tubulin alpha-4. Mec-12-prov protein 140c 157g 166r ri 27735456 Alphatub84b-prov protein |
| | | | 15c 26g 0r ri 51262162 PTK9. Protein tvrosine kinase 9c _12g 0r gi 963087 phosphorylase phosphatase |
| | | | 25c 13g 26r ri 49118559 similar to Nucleonorin SEH1 . NGC82845 protein 114c 83g 84r ri 12230782 WD-repeat protein 1 (Actin-interacting protein 1) |
| | | | 58c 40g 33r ri 28461382 Tbi-brov brotein 158c 112g 96r ri 32450571 Eno1-prov protein |
| | | | 134c 139u 228r ril47682287 cvtoskeletal keratin 19, NGC83069 |
| | | | 29c 29c 34r fil27124671 RanB77. importin 7, MGC79934 prot 113c 82g 154r fil32766495 Krt18-prov protein |
| | in the second second | | 284c 249g 212r ri 49115532 Clathrin. heavy polypeptide (Hc) 127c 74g 76r ri 50603997 FSCN1(fascin-1, actin-bundling t |
| | Contraction of the | | 1 49c 50cr 0r ri 28422594 Nucleolar protein 5 , 1 10c 8cr 0r ri 3414965 Lamin B2 |
| | | | 244c 2030 222r cil64880 lamin LIII |
| A DESCRIPTION OF | | | 1 23C 12 (0 14 (7 71604 (733) Lamin = 1111) (Lamin = 5) MG(13122 () 24C 15 (17 1154261464 Hoin, NHP2 non-histone chromosome protein 2-lik 18C 33g 33 gi 64882 Lamin B1 |
| | | | 80c 112α 164r ril51895818 NUP205. Nucleonorin NUP205, MGC83295 protein 46c 73α 96r ril2850225 NUP155-prov protein |
| | | | 90c 88cr 118r ri 27924242 Num93. Nuclear pore complex protein 93kDa. Dve- 19c 25g 29r ri 77748171 NUP62, nuclear pore complex glycoprotein p62, I |
| | | | 16c 36g 29g gi 50414519 Rae1/Gle2 17c 26g 41g gi 45126627 importin alpha 5.1 protein |
| | | | 19c 36g 18r xi 83318359 simiar to Dkc1, dyskerin, _33c 21g 18r gi 119856 Fibrillarin |
| | 1000 | | 39c 35g 52r xi 51950243 NUP153 protein _15c 14g _20r yi 50415052 NUP98 protein |
| | | | 5c 0g 8r gi 6093543 Nuclear transport factor 2 _9c _0g _11r j 40286642 NUP88B, nucleoporin Nup88B |
| | | | 42c 66g 82r ci 51258556 nucleonorin 85 . MGC86335 24c 43g 75r ci 55250537 transcription factor ELVS |
| | | | · _ · · · _ · · · · · · · · · · · · · · |
| | Time | | |

Figure S2. Protein Groups with Highly Correlated Temporal Behavior on Interphase Chromatin across All Experimental Treatments The combined dataset was generated by integration of profiles from inhibitor-free, geminin, and roscovitine experiments for each protein. Hierarchical clustering was performed on complete combined dataset. Branches of hierarchical trees with the correlation level above the threshold defined by the MCM licensing complex are presented. The color coding is as follows: 0.0, green; 0.5, black; and 1.0, red. Number of peptides identified for proteins in each treatment is also indicated.



в



Time

(Figure S3. Continued)



Figure S3. Hierarchical Clustering of Temporal Profiles for Proteins Identified on Chromatin Only in the Presence of Inhibitors

(A) Proteins whose binding to chromatin was induced by both inhibitors.

(B) Proteins identified on chromatin only in the presence of geminin.

(C) Proteins identified on chromatin only in the presence of roscovitine.

С

| Table S Replica | 1. Funct | tional Ann bitors | otation Clustering of Proteins Ind | nduced by | |
|--------------------|----------|----------------------|---|-----------|--|
| Drug | FG | GES | Overrepresented Annotation Terms with Highest p Value | p Value | |
| Gem (7 | 2) | | | | |
| | 1 | 1.84 0.87 | WD-40repeat (IN) | 1.00E-02 | |

| | 6 | 0.61 | DNA repair (GB) | 1.20E-02 |
|--------|-----------|------------|--|-------------------|
| | 7 | 0.42 | Phosphorylation (GB) | 3.60E-01 |
| Ros (9 | 4) | | | |
| | 1 | 3.41 | Phosphorylation (GB) | 4.10E-08 |
| | 2 | 1.05 | transporter activity (GM) | 1.90E-02 |
| | 5 | 0.66 | cellular physiological process (GB) | 5.10E-02 |
| | 8 | 0.38 | Armadillo-like helical (IN) | 2.60E-02 |
| Numbe | er of ger | nes unique | for each treatment is indicated | l in parentheses. |

Number of genes unique for each treatment is indicated in parentheses. "FG" indicates functional groups identified by DAVID; "GES" indicates group enrichment score. The annotation source is indicated next to a term: GB, GOTERM_BP_ALL; GM, GOTERM_MF_ALL; GC, GOTERM_CC_ALL; IN, INTERPRO_NAME; SN, SMART_NAME; SP, SP_PIR_KEYWORDS; KP, KEGG_PATHWAY; and HC, Hierarchical Clustering.

| | Number of Identified Peptide | |
|---------------------------|------------------------------|---------|
| Gene Name | Mcm2 IP | Mcm3 IP |
| MCM2 | 45 | 44 |
| MCM3 | 32 | 9 |
| MCM4 | 57 | 51 |
| MCM5 | 35 | 3 |
| MCM6 | 30 | 31 |
| MCM7 | 37 | 37 |
| ORC1 | _ | 7 |
| ORC2 | - | 4 |
| ORC3 | - | 2 |
| ORC4 | - | 2 |
| ORC5 | - | 1 |
| UBF1 | 2 | - |
| HP1 | 2 | 3 |
| Cdc6 | _ | 1 |
| DNA polymerase epsilon | 1 | _ |
| RuvBL2 | 2 | 2 |
| Duf87 | 1 | _ |
| AND-1 | _ | 1 |
| RECQL4 | - | 1 |
| MCM10 | - | 1 |
| Actr3 | 13 | 7 |
| ARP2 | 7 | 4 |
| Arp2/3 complex, subunit 2 | 4 | 4 |
| Arp2/3 complex, subunit 4 | 2 | 3 |
| Arpc1a | 1 | 4 |
| Capzb | 6 | 4 |
| Cortactin | 4 | 11 |
| Aurora B | - | 13 |
| XL-INCENP | - | 5 |
| Survivin | - | 5 |
| DasraA protein | - | 1 |
| Plk1 | 1 | 2 |
| ELYS/Mel-28 | 2 | 1 |
| Nup160 | - | 1 |
| Nup107 | 3 | 2 |
| Nup 85 | 3 | 1 |
| Nup53 | 1 | - |
| Nup43 | 1 | - |
| SEH1 | 1 | _ |