Supplemental Information

TFIIIC Binding to Alu Elements

Controls Gene Expression via Chromatin Looping and Histone Acetylation

Roberto Ferrari, Lara Isabel de Llobet Cucalon, Chiara Di Vona, François Le Dilly, Enrique Vidal, Antonios Lioutas, Javier Quilez Oliete, Laura Jochem, Erin Cutts, Giorgio Dieci, Alessandro Vannini, Martin Teichmann, Susana de la Luna, and Miguel Beato
Figure S1. Tumorigenic and normal cell lines show TFIIIC (but not TFIIIB nor Pol III) binding to AEs close to Pol II promoters of cell cycle-related genes upon serum starvation (SS), related to Figure 1.

(A) Schematic view of the experimental design.

(B) Cell cycle profile of T47D cells grown in normal conditions (+S, orange) or upon 16 h of SS (–S grey). Note that T47D cells show a cell cycle profile highly enriched in the G1 population in the normal growth conditions; although a slight increase in the proportion of cells in G1 phase with a concomitant reduction in the S and G2/M phases was detected following serum depletion, it was not statistically significant (Chi-square test, p-value = 0.09).

(C) Proportional Venn diagram of total AEs bound by TFIIIC detected vs those containing ETC sites (with only B-box) in the serum-starved cells.

(D) Bar plots of gene ontology (GO) enrichment for Pol II genes with TFIIIC-bound within 5 kb of their TSSs) in the serum-starved cells. GO terms were calculated by DAVID (Molecular Function and Biological Processes combined) and ranked from the lowest to the highest p-value of the first nine terms found.

(E) Heatmap of AEs density across all human TSSs spanning a 6 kb-region and sorted by high to low AEs density. TFIIIC-associated promoters are shown at the top. Color bar scale with increasing shades of color stands for higher AEs density.

(F) CEAS plots of the average binding to AEs bound by TFIIIC (top panels) or to tDNAs bound by TFIIIC (bottom panels) for BDP1 (left) and Pol III/RPC39 (right) in T47D grown in normal condition (+S, red) or upon SS (-S, grey) (plotted is the –log10 of the Poisson p-value).

(G) CEAS plots of TFIIIC average binding for MCF10A, T98G and IMR90 cells in normal growth conditions (+S, red) and following SS (–S, grey, see Methods for specific conditions for each cell line). The graphs are plotted over the summit of TFIIIC peaks in the –S condition (plotted is the –log10 of the Poisson p-value). The enrichment in peaks corresponds to AEs (not shown).

(H) Genome browser view of CCNE1, MDM4 and UBE2V2 loci with ChIP-seq data for Pol III, BDP1 and TFIIIC in MCF10A and T47D breast cell lines. The graph includes the tracks for AEs and tDNAs. Highlighted in grey is the AE bound by TFIIIC close to the TSS of the indicated genes and arrow shows direction of transcription.

(I) CEAS plot of TFIIIC average enrichment for T98G cells grown in normal conditions (+S, red), SS (–S, grey), or SS followed by serum addition for 30 min (–S/+S-30 min, purple). The graphs are plotted over TFIIIC-peaks summit in the –S condition (plotted is the –log10 of the Poisson p-value).
Fig. S2

A

B

C

D

E

F

G
Figure S2. ADNP depletion causes TFIIIC loss from TFIIIC-associated promoters upon serum starvation, Related to Figure 2.

(A) Western Blot analysis of tagged-immunopurification of TFIIIC and ADNP complexes as described in Figure 2B. FLAG- and HA-purified complexes were probed for the presence of TFIIIC (GTF3C2 subunit) and ADNP.

(B) Pie chart showing the percentage of mouse Adnp peaks found in mESCs belonging to repetitive elements and non-repetitive elements (analysis from GSE97945). Notice that almost all the binding of this factor lays on repetitive elements.

(C) Pie chart showing the percentage of Adnp peaks mapping to repetitive elements and belonging to B-elements or to other type of repetitive elements. Notice that more than 50% of Adnp peaks mapping to repetitive elements lay on B-elements.

(D) Average plot for ADNP-eGFP (GSE105573) enrichment across TFIIIC-associated Pol II promoters (red) spanning a 6 kb-region. The profile of a random set of genes of the same size (purple), as well as the average for all human TSSs (black) is also shown.

(E) qRT-PCR expression analysis of ADNP in serum-starved T47D cells (siCTRL and siADNP) corresponding to the experiment in Figure 2D. The value in siCTRL cells was arbitrarily set as 1. Note that the knockdown of ADNP reaches values almost 80% of its control.

(F) Immunoblot probing the levels of ADNP protein across three cellular fractions: cytoplasm (Cyt), nucleoplasm (Nuclear) and chromatin (Chrom) in serum-starved T47D cells transfected with siADNP or control siCTRL. Note that the largest pool of ADNP is found in the chromatin fraction and is strongly reduced upon knock down. α-Tubulin and histone H1.2 are used as controls for the subcellular fractionation.

(G) CEAS profile of TFIIIC enrichment over TFIIIC-associated Pol II promoters upon depletion of ADNP (siADNP, purple) compared to control cells (siCTRL, red) in serum-starved T47D. Note the strong reduction in TFIIIC occupancy (from -logP ~ 5 to -logP ~ 2.5) in ADNP knocked down cells.
Fig. S3

A. Average H3K27ac (−logP)

B. Average H3K18ac (−logP)

C. Average p300 binding (−logP)

D. Nucleosome density (H2P enrichment score)

E. Relative distance from TFIIIC-AEs [−S] summit (kb)

F. Relative distance from TFIIIC-AEs bound AEs (kb)

G. Relative distance to T47D TFIIIC center (kb)

H. Frequency

I. Pearson Correlation

J. Gene profiles

K. Enrichment of gene expression

Legend:
- H3K18ac
- H3K18a
- TFIIIC
- p300
- siCTRL
- siGTF3C1
- siGTF3C5
-HELLS
-FEM1A

Supplementary Figure S3

A-C. Metagene analysis showing the enrichment of histone acetylation and p300 binding at different distances from TFIIIC-AEs.

D. Nucleosome density profile showing the enrichment score at different distances to TFIIIC.

E. Relative distance from TFIIIC-AEs [−S] summit (kb).

F. Relative distance from TFIIIC-AEs bound AEs (kb).

G. Relative distance to T47D TFIIIC center (kb).

H. Frequency of gene expression enrichment.

I. Pearson Correlation analysis.

J. Gene profiles showing the enrichment at different conditions.

K. Enrichment of gene expression with different conditions.
**Figure S3.** Both H3K18ac and H3K27ac mark AEs occupied by TFIIIC, but only H3K18ac is changed upon serum starvation, Related to Figure 3.

(A) Sitepro profile of H3K27ac enrichment in T47D grown in the presence (+S, red) or absence of serum (−S, grey) over TFIIIC-bound AEs (plotted is the −log10 of the Poisson p-value). Note that the H3K27ac levels at these sites are independent of the growth conditions.

(B) Sitepro profile of H3K9ac enrichment in T47D grown in the presence (+S, red) or absence of serum (−S, grey) over TFIIIC-bound AEs (plotted is the −log10 of the Poisson p-value). Note the low levels of H3K9ac at these sites and their independence of the growth conditions.

(C) Sitepro profile of H3K18ac enrichment in T47D grown in the presence (+S, red) or absence (−S, grey) of serum at tDNAs (plotted is the −log10 of the Poisson p-value).

(D) Nucleosome density as mapped by MNase digestion across all TFIIIC-bound AEs in T47D cells using the iNPS score. The plot has been produced using NucMap (Zhao et al., 2019). Note the presence of at least one nucleosome over the region of the AE (represented as a grey rectangle).

(E) Genome browser view of representative TFIIIC-associated genes FEM1A and HELLS with ChIP-seq data for p300 and H3K18ac in T47D in the presence (+S) or absence (−S) of serum. The AE bound by TFIIIC in each locus is shown by a grey rectangle. The gene structure and the direction of transcription (arrow) are shown at the bottom. Note that p300 is not recruited at the AEs bound by TFIIIC (grey boxes) as it is for other adjacent intergenic regions (yellow box).

(F) Average plot for p300 occupancy across all TFIIIC-bound AEs spanning a 6 kb-region in T98G grown in the presence (+S, red) or absence (−S, grey) of serum. Data was from GSE21026.

(G) Average plot for GTF3C1 occupancy across all TFIIIC-bound AEs spanning a 6 kb-region in T47D grown in the absence of serum.

(H) Immunoblot probing GTF3C1 protein levels in serum-starved T47D cells transfected with siGTF3C1 or siCTRL as in Figure 3H. Histone H1.2 is shown as loading control.

(I) Histogram plot of Pearson’s correlation frequencies of H3K18ac colocalization with DAPI staining of Figure 3H. The large majority of cells in T47D siCTRL cells had H3K18ac colocalizing with DAPI, whereas cellular ablation of GTF3C1 caused the loss of H3K18ac and consequently its colocalization with DAPI.

(J) Immunoblot probing GTF3C5 protein levels in T47D cells transfected with increasing concentration of siGTF3C5 or control siCTRL.

(K) ChIP-qPCR showing loss of H3K18ac enrichment at two AEs bound by TFIIIC (UHRF1 and HELLS loci) in serum-starved T47D upon knock down of GTF3C5 by siRNA. The graph shows mean and SD of 2 independent experiments. Depletion levels are shown in Figure S3J.
Fig. S4

A) T47D: serum starvation vs serum

B) Eigen Vector profile

C) Gene expression profile T47D

D) Average Pol II (-logP)

E) Replicate 1

F) Replicate 2

G) Replicate 3

H) Replicate 4

I) KS Test < 0.001

J) Normalized interaction score for AEs

K) Gene expression profile T47D

L) GO TAGs
Figure S4. Global gene expression analysis in T47D grown in normal conditions and after serum deprivation and dependence on TFIIIC, Related to Figure 4.

(A) Eigen Vector profile (-log10 adjusted p-value) for changes in A and B compartments in Hi-C experiments (4 replicates for each condition) across all chromosomes for T47D cells grown in the absence or presence of serum. Yellow represents switches to A compartment and dark purple represents switches to B compartment upon serum deprivation. Black horizontal bars represent threshold of p-val>0.001. Genomic positions are indicated at the bottom.

(B) Barplot representing the total number of TADs for Hi-C experiments in T47D grown in the presence or absence of serum (4 independent replicates for each condition). No significant differences were observed comparing replicates and conditions (Fisher exact test = 0.326).

(C) Scatter plot of Pol II gene expression in T47D cells grown in the presence (+S) or absence (–S) of serum (red, up-regulated in +S; blue, down-regulated in +S; grey, no changes). The number of genes up- or down-regulated in the +S condition (-1.5 < FC > 1.5; adjusted p-value < 0.05) is indicated.

(D) Boxplot of mRNA expression for T47D cells grown in the presence (+S, red) or absence (–S, grey) of serum for TFIIIC-associated Pol II promoters. A random dataset of the same size has been used as a control (**p-value < 0.001; two-tailed paired t-test). Notice that TFIIIC-associated promoters show higher level of expression compared to a random control. No significant changes were detected when the two growth conditions were compared (two-tailed paired t-test).

(E) CEAS profile of average total Pol II enrichment at TFIIIC-associated promoters for T47D cells grown in the presence (+S, red) or absence (-S, grey) of serum. The graphs are plotted over the TSS of the TFIIIC-associated promoters in the –S condition (plotted is the –log10 of the Poisson p-value). No significant differences were detected (two-tailed paired t-test).

(F) Box plot of oneD-normalized interaction scores calculated for all the AEs bound by TFIIIC for T47D grown in the presence (+S) or absence (-S) of serum. Note the significant increase in the interaction score (Kolmogorov-Smirnov Test).

(G) Sitepro profile of strand-specific tRNA expression in T47D in conditions of siCTRL or siGTF3C5 in the absence of serum across all tDNAs spanning a 400 bp region (±200 bp relative to the tRNA TSS; plotted is the average expression in RPM). Note no differences in the two profiles.

(H) Volcano plot comparing mRNA-seq data of siGTF3C5 vs siCTRL in T47D grown in the absence of serum (–S) (plotted the –log10 of the adjusted p-value vs the -log2 ratio of siGTF3C5 vs siCTRL). The genes that scored significant (adjusted p-value < 0.05) are indicated in red (FC > 1.5) and blue (FC < -1.5). GTF3C5 is found among the most downregulated genes. See Table S2 for more information.

(I-J) Scatter plot of gene expression comparing siGTF3C5 (I) and siCTRL (J) treated cells in the absence of serum vs mock-transfected cells in the same growth condition. The number of genes up- or down-regulated (-1.5 < FC > 1.5; adjusted p-value < 0.05) is indicated in red or blue, respectively.

(K) Bar plots of GO enrichment (according to DAVID) of TFIIIC-activated genes (genes downregulated in siGTF3C5 cells).

(L) FACS cell cycle profile of T47D cells grown in the absence (–S) of serum or transfected with siCTRL or siGTF3C5 in the same growth conditions (n = 2 independent experiments). Note that no effect was observed by the siCTRL treatment, but further cell cycle arrest detected in siGTF3C5 cells (*, p-value for squared χ² test = 0.00045).
Fig. S5

A

B

C

D

E

F

G

H

I

J
Figure S5. TFIIIC and CTCF increased interaction upon SS does not depend on changes in total levels of TFIIIC, Related to Figure 5.

(A) Venn diagram of overlapping peaks for CTCF and TFIIIC in T47D grown in the absence of serum.

(B) Motif enrichment analysis of CTCF-bound regions in TFIIIC-looped genes. The p-value of the motif’s enrichment is shown.

(C) Sitepro analysis of RAD21 enrichment at CTCF-bound regions in the TFIIIC-looped genes (red) compared to a random set of regions of the same size (blue). Note that there is no difference in the RAD21 enrichment between the two conditions.

(D) Venn diagram of overlapping peaks for the Cohesin component RAD21 and the CTCF peaks in condition of serum starvation.

(E) Western blot for different TFIIIC subunits (GTF3C1, GTF3C2, GTF3C4 and GTF3C5) in T47D cells grown in the presence (+S) or absence (–S) of serum. For each panel, a loading control with α-tubulin is also shown.

(F-G) Virtual 4C (from Hi-C data) representation of the CENPE (F) and CCND1 (G) loci, two of the genes repressed upon siGTF3C5 (Figure 4C). The data is represented for the conditions of presence (+S red) and absence (–S, blue) of serum. The viewpoint of the virtual 4C is the AE bound by TFIIIC. Arrows point to the location of the gene, the AE bound by TFIIIC (AE-TFIIIC) and the peak of H3K27ac. The y-axis corresponds to the corrected contact score. Regions with changes in the frequency of interaction of the AE-TFIIIC and the target genes have been zoomed out for a clearer view. Grey boxes represent the location of the gene compared to the 4C signal.

(H) Sitepro analysis of ADNP-eGFP binding in K562 cells (GSE105573) at AEs of TFIIIC-associated promoters (red), all AEs-bound by TFIIIC (TFIIIC-AEs, blue) and TFIIIC-looped genes (yellow). Significant higher levels of ADNP-eGFP are found at TFIIIC-looped genes (p-values for squared Chi test comparing TFIIIC-looped genes vs AEs at TFIIIC-associated promoters and TFIIIC-AEs is reported). X-axis is either the center of the AE (for the AEs at TFIIIC-associated promoters or TFIIIC-AEs, or the TSS of the TFIIIC looped genes).

(I-J) CEAS plot of H3K18ac (I) and H3K27ac (J) average at the TSS of all human genes in T47D grown in the presence (+S, red) or absence (–S, grey) of serum (plotted is the −log10 of the Poisson p-value). Note how H3K18ac drastically changed upon serum starvation, whereas H3K27ac remained unaffected.
**Fig. S6**

A

B

C

D

E

F

G
Figure S6. The deletion of a TFIIIC-bound AE affects DNA looping and expression of the distal *UHRF1* locus, Related to Figure 6.

(A) Boxplot of the normalized interaction score of Hi-C data for siCTRL and siGTF3C5 between promoters and TFIIIC-bound AEs for TFIIIC-looped genes. Two Hi-C biological replicates were used (p-value from Friedman X² test is indicated).

(B) qRT-PCR expression analysis of *GTF3C5* in T47D cells treated with siCTRL and siGT3C5 in the absence of serum and released from serum starvation by serum addition for the indicated times. Samples correspond to Figure 6D. The graph represents the mean ± SEM from two biological experiments, in which the value in siCTRL cells was arbitrarily set as 1 at each time point. Note that the knockdown of TFIIIC always reaches values of more than 70% at each time point analyzed.

(C) PCR result for the screen of CRISPR-Cas9 T47D clones with primers Up3 and Down (see Supplementary Materials and Methods for details and schematic representation in Figure 6E): the upper band corresponds to the WT allele, whereas the lower band correspond to the deleted allele. Representative clones are shown, but almost all clones analyzed were heterozygous for the deletion. For further analysis, clone 11 was selected. The DNA marker size is shown. * indicates a non-specific band. NC corresponds to no DNA sample. Dotted line represents a cut of the gel. However, all the remaining lanes are from the same gel.

(D) Cartoon model for the proposed TFIIIC mechanism of action.

(E-F) Kaplan–Meier plots of breast tumor samples for TFIIIC-associated promoters or TFIIIC-activated genes expression, respectively. TFIIIC-associated promoters were divided in three main groups according to their expression levels within brackets (with blue being the highest, red the intermediate and grey the lowest). P-values from a Mantel-Cox test are indicated. Higher expression of TFIIIC-associated promoters is associated with poor prognosis for overall survival and distance metastasis free survival (DMFS), respectively. Plots are generated using (GOBO) (Ringner et al., 2011).

(G) Boxplots of expression of TFIIIC-associated promoters from all tumor samples across the three breast cancer grades. Box plots are generated by using (GOBO). TFIIIC-associated promoters show higher expression in most aggressive tumors (3rd grade), p-value is also indicated.