CTCF binding at the *H19* imprinting control region mediates maternally inherited higher-order chromatin conformation to restrict enhancer access to *Igf2*

Sreenivasulu Kurukuti [†‡], Vijay Kumar Tiwari [†‡], Gholamreza Tavoosidana [†‡], Elena Pugacheva[§], Adele Murrell [¶], Zhihu Zhao [†], Victor Lobanenkov [§], Wolf Reik

Laboratory of Developmental Genetes TJ /T1_4 1 Tf 24.45 0 Td (H19

Fig. 2. Analysis of parent-of-origin-speci c patterns of physical proximity between ICR and the Igf2 H19 domain in neonatal liver. The 3C analysis was performed by comparing relative crosslinking frequencies between the xed ICR of the maternal allele and the rest of the locus after normalizing the maternal allele frequencies to 100% (blue line). Also see Fig. 11 for direct comparison of frequencies of interactions. Because of a lack of signal for the maternal allele in some instances, the blue line makes a dip, as indicated. 3C analysis and allelic bias were corrected for as described in Fig. 1 (see Fig. 12). The bottommost image exempli es a hot-stop PCR analysis of 3C samples, which were digested with FauI to identify the SD7 allele. See Material and Methods for additional information.

the same locations as described above (Fig. 2). As for the enhancer–chromatin fiber interaction, albeit less marked, the ICR interactions reflect the proximity of the sequences involved with reduced interactions with more distal elements (Fig. 11, which is published as supporting information on the PNAS web site). A FauI polymorphic restriction site 61 bp downstream of the EcoRI restriction site within the ICR of specifically the SD7 allele was used to discriminate between parental alleles. After corrections for allelic bias in the PCR amplification steps (Fig. 12, which is published as supporting information on the PNAS web site), Fig. 2 displays the signal of the paternal allele (red bars) in relation to that of the maternal allele (blue lines) normalized to 100%.

This experiment shows that the ICR is in close physical proximity to all regions examined in the entire Igf2 H19 domain on both parental alleles, with three significant exceptions. First, the ICR is excluded from contact with P1 DMR2 specifically on the maternal chromosome and in close contact with P1 DMR2 only on the paternal chromosome. Second, the ICR is in close contact with DMR1 exclusively on the maternal chromosome. These results corroborate the findings by Murrell et al. (29). Third, and most interestingly, on the 3 side of Igf2, the ICR is in contact with MAR3 exclusively on the maternal chromosome. Therefore, on the maternal allele, the Igf2 gene is located in a tight pocket made of contacts between the ICR, DMR1, and MAR3, which we suggest excludes the gene (P1 DMR2) from interactions with the enhancers.

Significantly, maternal transmission of the 142* allele (with the CTCF binding site mutations) led to a loss of the maternal interactions among the ICR, DMR1, and MAR3 and a gain of interaction between the ICR and P1 DMR2 (Fig. 2), showing that CTCF binding is the key regulator of the tight pocket between the ICR, DMR1, and MAR3.

The *H19* ICR–CTCF Complex Controls the Epigenetic Status at Igf2 DMR1 2. To understand the DMR1–ICR interaction in more detail, we performed 3C analysis using HindIII-digested chromatin (Fig. 3A). Fig. 3 B and C confirms that only the maternal DMR1 allele is engaged in the interaction with the H19 ICR and that this maternal interaction was lost when the mutated H19 ICR allele was inherited maternally. To demonstrate that CTCF is actually present in this complex, we combined the 3C analysis with chromatin immunopurification (ChIP), a method termed ChIP-loop assay (35). Formaldehyde crosslinked chromatin from livers of wild-type mice was digested with HindIII in the initial 3C step as outlined above and were immunopurified by a CTCF antibody (13–15). The purified DNA–protein complex was ligated and amplified for 3C analysis as above. Fig. 3 shows that the expected PCR product diagnostic of the H19 ICR–Igf2 DMR1 complex was specifically present in the ChIP material but absent in controls, including ChIP samples obtained with control serum. This result documents that CTCF is part of a complex that includes the Igf2 DMR1 and H19 ICR in close physical proximity.

Our earlier demonstration that different CTCF–DNA complexes can interact physically with each other (5) prompted us to examine the possibility that the ICR–DMR1 complex involved CTCF target sites on both sides of the loop. EMSA revealed that only two overlapping DNA fragments in DMR1, 2 and 3 in Fig. 13 Upper, which is published as supporting information on the PNAS web site, displayed an ability to significantly interact in vitro with CTCF. Interestingly, the CTCF-positive fragments cover five of the seven CpGs analyzed in the bisulfite analysis for differential methylation of Igf2 DMR1 (see below). We therefore determined whether the CTCF binding to DMR1 was methylation-sensitive. Fig. 4A shows that only one DNA fragment, which encompasses CpG nr 5 (see below), interacted with CTCF in a methylation-sensitive manner.

Fig. 3. CTCF target sites control the interaction between the H19 ICR and Igf2 DMR1 regions. (A) Schematic map of the Igf2 and H19 loci. The Igf2 DMR1 and H19 ICR domains are expanded to show the locations of 3C primers (marked with roman numerals and thick arrows to indicate their directions). The numbers indicate their distance from the HindIII sites. Primers IV V span a polymorphic restriction site for DraI speci c to the SD7 allele. (B) Three independent samples from each cross were subjected to the 3C assay. The PCR products were digested with DraI and subjected to Southern blot hybridization analysis to verify speci city of the ampli ed DNA fragments. The amplication of HindIII-digested and ligated yeast arti cial chromosome DNA covering the entire Igf2 H19

Next, we examined the possibility that DMR1 interacted with CTCF in vivo by using ChIP assays. Fig. 4B shows that CTCF indeed interacts specifically with the maternal DMR1 allele. Strikingly, CTCF binding to the maternal DMR1 allele was lost when the mutated H19 ICR allele was inherited maternally. This result suggests that in vivo CTCF is recruited to DMR1 through the physical interaction between the H19 ICR and DMR1.

Given that CTCF binds to the maternal DMR1 allele in vivo dependent on CTCF binding sites within the H19 ICR, it was possible that this CTCF binding conferred protection against DNA methylation. We therefore carried out methylation analysis of DMR1, comparing paternal (control) to maternal inheritance of the CTCF binding mutant. With paternal inheritance of the 142* allele (control), the maternal DMR1 allele was comparatively less methylated than the paternal one (Fig. 4E), as described for the wild-type situation (27). However, when the mutant H19 ICR allele was maternally inherited, there was a significant overall increase of methylation of the maternal DMR1 allele. Particularly striking was the change at CpG nr 5, for which we showed methylation-dependent CTCF binding in vitro. In the control cross, the paternal allele was highly methylated, whereas the maternal allele was hypomethylated, but both alleles were hypermethylated when the

Fig. 4. ICR–DMR1, CTCF, and long-range epigenetic coordination. (A) (Upper) Effects of in vitro CpG methylation on CTCF binding to positive DNA fragments 2, 3, and 7 (the last fragment covers only CpG site nr 5; see Fig. 13). Digestion of unmethylated methylated probe with the methylation-sensitive enzyme HpaII is depicted in lane d for each panel. Three pair of panels with both unmethylated and methylated [^{32}P]DNA probes and in vitro-translated proteins are depicted (lanes are marked as in Fig. 13 Upper). (Lower) Shown is a map of core DMR1 fragment with CpG numbers and the positions of the overlapping DNA fragments. (B and C) CTCF interacts with the Igf2 DMR1 in vivo. ChIP analysis of neonatal liver derived from a cross between C57BL 6 and SD7 (B) or from 142* SD7 and SD7 142* crosses (C). The DraI polymorphism speci c for the M. spretus allele of the DMR1 in the recombinant SD7 mouse strain was exploited (see Fig. 3 A). (D) Bisul te sequencing data of DMR1 in neonatal liver. Filled and open circles represent methylated and unmethylated CpGs, respectively. Maternal and paternal alleles were distinguished using the DraI polymorphism. (E) Summary of the overall bisul te data or, speci cally, the fth CpG site, given in percentage of CpG methylation.

mutant ICR was maternally derived. Clearly, removal of CTCF binding to the unmethylated CpG Nr 5 in DMR1 had removed the protection against de novomethylation. In addition to DMR1, we found that the CTCF binding sites in the ICR prevent de novo methylation of the maternally inherited DMR2 allele (Fig. 14, which is published as supporting information on the PNAS web site).

Discussion

We have carried out a systematic 3C analysis in an imprinting cluster. Although we have taken care to perform this study as quantitatively as possible, it is important to point out that all major conclusions are based on observing qualitative interactions. The significant conclusions from this work first include the observation that the distal enhancers, which are required both for Igf2

organization of higher-order chromatin structure in cis is a general property of these elements.

Materials and Methods

3C Assay. Neonatal mouse liver cells were dispersed by immediate mashing through a 70-mm nylon cell strainer into Dulbecco's modified Eagle's medium. The 3C assays were done essentially as described in refs. 29 and 40. The linear range of amplification was determined for the liver samples by serial dilution. The PCR products were subjected to Southern blot hybridization or hot-stop PCR ([-^{32}P]-ATP-labeled en4-R and ICR-R primer, respectively) (41) analysis, and the results were quantified by using MULTIGAUGE version 2.2 PhosphorImager system (Fuji). All information about primers and PCR conditions are summarized in Table 1, which is published as supporting information on the PNAS web site. Consult the Supporting Materials and Methods, which is published as supporting information on the PNAS web site, for a complete description of how 3C digestion, ligation, and PCR were controlled, and how quantitative values were normalized. Briefly, we normalized for 3C efficiency between experiments by using the unrelated gene locus Ercc3 for ligation and PCR bias of each primer pair with a Igf2-H19 yeast artificial chromosome and for parental PCR bias between M. m. domesticus and M. spretus alleles by mixing the 3C PCR products of the two strains in defined ratios. Each experiment was done on three different liver specimens from each mouse cross that is described in the study. Each liver sample was processed for 3C analysis three times.

ChIP-Loop Assay. Formaldehyde-crosslinked chromatin was subjected to a ChIP-loop assay (35). Briefly, DNA–protein complexes were digested with HindIII, precleared for 4 h with protein G4 Fast Flow Sepharose beads (Amersham Pharmacia Biosciences), and then incubated with mouse monoclonal CTCF antibody (BD Biosciences, Franklin Lakes, NJ) overnight. After incubation with protein G4 Fast Flow Sepharose beads (and washing the complex four times), the beads were suspended in ligation buffer and subjected to the 3C analysis as described above.

Bisulfite Sequencing Analyses. Genomic DNA was isolated from 1-day postpartum liver tissues (Promega). Approximately 1 g of EcoRI-digested DNA was subjected to bisulfite treatment (42) and PCR amplification (27). The resulting PCR products were gel-purified (Qiagen, Hilden, Germany) and ligated into pCR2.1 (Invitrogen). The sequencing analyses were performed with the BigDye Terminator cycle sequencing kit (Applied Biosystems).

Nuclear Extracts and *in Vitro* Transcription–Translation. Full-length human CTCF and the 11ZF CTCF-binding domain were in vitro translated from pET-7.1 and pET-11ZF, respectively (43), by using the TnT reticulocyte lysate-coupled in vitro transcription–translation system (Promega).

EMSA and *in Vitro* CpG Methylation. The in vitro analysis of CTCF binding sites and CpG methylation effects was performed as described in ref. 15. A detailed account of the primer sequences is presented in Supporting Materials and Methods and Table 1.

We thank Drs. G. Smits and W. de Laat for helpful discussions. This work