Chromosome Research (2005) 13:809–818
DOI: 10.1007/s10577-005-1015-4

Allele-specific methylation of a functional CTCF binding site upstream of *MEG3* in the human imprinted domain of 14q32

Alberto L. Rosa[1], Yuan-Qing Wu[2†], Bernard Kwabi-Addo[2†], Karen J. Coveler[2], V. Reid Sutton[2] & Lisa G. Shaffer[1*]
[1]*Health Research and Education Center, Washington State University, PO Box 1495, Spokane, WA 99210-1495, USA; Tel: +1-509-368-6710; Fax: +1-509-358-7627; E-mail: lshaffer@wsu.edu;* [2]*Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX, USA*
[†]*These authors contributed equally to this work*
[*]*Correspondence*

Received 7 July 2005. Received in revised form and accepted for publication by Mike Goldman 28 October 2005

Key words: epigenetics, imprinting, methylation, protein binding sites, UPD

Abstract

The gene *MEG3* is located in the imprinted human chromosomal region on 14q32. Imprinting of a structurally homologous region *IGF2*/*H19* on 11p15 is mediated through cytosine methylation-controlled binding of the protein CTCF to target sites upstream of *H19*. We identified five new CTCF binding sites around the promoter of *MEG3*. Using an electrophoretic mobility shift assay, we showed that these sites bind CTCF *in vitro*. Using one of these sites, chromatin immunoprecipitation (ChIP) analysis confirmed CTCF binding *in-vivo*, and differential allele-specific methylation was demonstrated in seven individuals with either maternal or paternal uniparental disomy 14 (UPD14). The site was unmethylated on the maternally inherited chromosomes 14 and methylated on the paternally inherited chromosomes 14, suggesting parent-specific methylation of sequences upstream of *MEG3*. We speculate that this CTCF-binding region may provide a mechanism for the transcriptional regulation of *MEG3* and *DLK1*.

Introduction

Genomic imprinting is an epigenetic phenomenon in which the activity of a gene is reversibly modified depending upon the parent of origin. This leads to unequal expression from the maternal and paternal alleles of a diploid locus. Thus, imprinted genes are functionally haploid in normal individuals. A number of human disorders result from alterations in genomic imprinting (Morison & Reeve 1998). In humans and mice, at least 50 imprinted loci have been identified in which maternal or paternal alleles may either be expressed or repressed during ontogeny (http://www.mgu.har.mrc.ac.uk/imprinting/imprin-generef.html). Chromosome abnormalities, such as uniparental disomy (UPD), can result in abnormal phenotypes that unmask genomic imprinting. In the simplest sense, UPD is the situation in which both chromosomes of a pair are inherited from one parent, with no contribution from the other parent (Engel 1980). Uniparental disomy of chromosome 14 (UPD14) results in one of two distinct abnormal phenotypes depending upon the parent of origin. The phenotype associated with maternal disomy 14 is mild relative to that of paternal disomy 14 (Sutton & Shaffer 2000). Maternal UPD [upd(14)mat] of chromosome 14 is associated with low birth weight, short stature, small hands and feet, motor delay, and precocious puberty, whereas paternal UPD14 [upd(14)pat] is associated with a more severe musculoskeletal phe-

notype including bell-shaped thorax and rib anomalies, resulting in respiratory insufficiency and mental retardation (Sutton & Shaffer 2000). The abnormal phenotypes associated with UPD14 presumably result from the over-expression or under-expression of imprinted genes on this chromosome. Many studies have provided evidence that imprinted genes are located within the 14q24-q32 region (Robin *et al*. 1997, Georgiades *et al*. 1998, Sutton *et al*. 2002). Subsequently, a cluster of imprinted genes including *MEG3* and *DLK1* was identified and localized to 14q32 (Kobayashi *et al*. 2000, Miyoshi *et al*. 2000, Schmidt *et al*. 2000, Takada *et al*. 2000, Wylie *et al*. 2000, Charlier *et al*. 2001a,b).

MEG3 is located within 14q32 and lies approximately 90 kb distal to the paternally expressed gene *DLK1*. *MEG3* expression studies using single nucleotide polymorphisms (SNPs) demonstrated monoallelic expression in fetal heart, kidney, liver, lung and brain, with the transcripts originating exclusively from the maternal allele (Wylie *et al*. 2000). Furthermore, bisulfite sequencing of normal tissues showed chromosome-specific methylation of three CpG islands upstream of *MEG3*; however, no cases of UPD14 were examined to investigate correlation between regions of methylation, parental origin, and differential expression (Wylie *et al*. 2000). Although it has been demonstrated that *MEG3* and *DLK1* are expressed exclusively from the maternal and paternal allele, respectively, the regulatory elements that modulate their transcription are poorly understood.

Insulator elements containing binding sites for the zinc-finger protein CTCF can regulate the expression of genes. The binding of CTCF to insulator elements mediates promoter–enhancer interactions (Bell *et al*. 1999, Hark *et al*. 2000, Lewis & Murrell 2004). Bell & Felsenfeld (2000) first demonstrated that the imprinted genes *Igf2* and *H19* are regulated by differential methylation of a segment of DNA that contains binding sites for the protein CTCF. Binding of CTCF to its unmethylated cognate binding sites blocks access of the enhancers, located 3′ of *H19*, to the *Igf2* promoter, thereby silencing the *Igf2* gene. However, the enhancers can still interact with the H19 promoter to allow for transcription of *H19*. Methylation of CpG sequences at CTCF target sites impairs binding of this protein to chromatin, inhibits interaction of enhancers with the *H19* promoter and permits transcription of the *Igf2* gene (for a review see Lewis & Murrell 2004). It has recently been demonstrated that patients affected with Beckwith–Wiedemann syndrome can carry inherited microdeletions abolishing CTCF binding sites at the *Igf2*/*H19* region (Sparago *et al*. 2004, Prawitt *et al*. 2005).

Previous work has identified two consensus CTCF binding sites located in the second intron of *MEG3* (Wylie *et al*. 2000). Methylation within and surrounding the two putative human CTCF consensus sequences was also reported (Wylie *et al*. 2000). Although these observations would be consistent with allele-specific differential methylation at this locus, this was not confirmed in these studies, nor was parent-specific methylation demonstrated (Wylie *et al*. 2000). Herein, we identify five additional CTCF binding sites around the promoter of *MEG3*: one additional site is located within the second intron, whereas four sites are located upstream of the gene. We demonstrate that CTCF binds to these sites *in vitro*. In addition, *in-vivo* binding of CTCF to one of the upstream *MEG3* CTCF sites as well as parent-specific methylation of this site is shown.

Materials and methods

Patients and cell lines

Patient information and samples were obtained using a Baylor College of Medicine Institutional Review Board-approved consent form. Genomic DNA was isolated from lymphoblast, amniocyte, or fibroblast cell lines established on seven UPD 14 patients reported previously (Table 1) and their available parents. The patients' clinical features and diagnostic methods are described elsewhere (Pentao *et al*. 1992, Antonarakis *et al*. 1993, Walter *et al*. 1996, Sutton & Shaffer 2000, Towner *et al*. 2001a,b, Berend *et al*. 2002, Coveler *et al*. 2002, McGowan *et al*. 2002). We confirmed the diagnosis of UPD14 in all cases using microsatellite markers on 14q, as previously described (Berend *et al*. 2002).

Oligonucleotide synthesis and probe preparations

Oligonucleotide sequences were synthesized by Integrated DNA Technologies, Inc., and modified oligonucleotide sequences were synthesized by New England Biolabs. Table 2 shows the sense strand

Table 1. Summary of seven UPD cases studied for parent-specific differential methylation.

Case	*Parental origin of UPD*	*Karyotype*	*Reference*
1	Maternal	45,XX, rob(14q15q)	Berend *et al*. 2002
2	Maternal	45,XX,i(14)(q10)	Pentao *et al*. 1992, Sutton & Shaffer 2000
3	Paternal	45,XY,i(14)(q10)	Walter *et al*. 1996
4	Paternal	45,XX,i(14)(q10)	McGowan *et al*. 2002
5	Maternal	46,XX,rob(13q14q),+14[5]/45,XX, rob(13q14q)[95]	Antonarakis *et al*. 1993
6	Maternal	46,XX	Towner *et al*. 2001a,b
7	Paternal	46,XY	Towner *et al*. 2001a,b, Coveler *et al*. 2002

sequence for each oligonucleotide. Complementary oligonucleotide sequences were annealed by boiling for 5 min equal amounts (10 μg) of the sense and anti-sense sequences in 10 mM Tris-HCl pH 8.5 and 50 mM NaCl and then slowly cooling the solution to room temperature. To make the probes, the 5′ termini of the annealed oligonucleotides (0.1 μg) were radiolabelled with T4 polynucleotide kinase (Invitrogen, Life Technologies) at 2 units/μg in a 50 μl reaction volume containing 10 μCi [γ-^{32}P] dCTP (Amersham; 3000 Ci/mmol), and 1 X polynucleotide kinase reaction buffer (Invitrogen, Life Technologies). The reactions were incubated at 37°C for 30 min followed by inactivation of the enzyme by incubation at 70°C for 10 min. Unincorporated [γ-^{32}P] dCTP were removed by purifying the probes using a G-25 (Fine) Sephadex Quick spin columns (Roche).

Preparation of CTCF protein

For the *in-vitro* translation of CTCF protein, 2 μg of the chicken CTCF plasmid (p4bfull; a gift from Gary Felsenfeld) was used in a 50 μl reaction volume according to the manufacturers' protocol (TNT R T7 Coupled Reticulocyte Lysate System, Promega). Chicken CTCF has been used successfully in these types of experiments due to the high similarity with the human CTCF protein (Ohlsson *et al*. 2001). Nuclear extracts from confluent human fibroblast cells were prepared as published previously (Schiff *et al*. 2000). The supernatant containing total nuclear proteins was used in the electrophoretic mobility shift assays.

Electrophoretic mobility shift assay (EMSA)

The EMSAs were carried out as described previously (Zimarino & Wu 1987). Briefly, the ^{32}P-labelled oligonucleotide probes were incubated with or without protein extract in a total reaction volume of 25 μl containing the binding assay buffer (50 mM Tris-HCl, pH 7.4; 10 mM $MgCl_2$; 30 mM NaCl; 1 mg/ml BSA; 1 μg/ml poly dI-dC; 20% Ficoll). The reactions were started by the addition of either the *in-vitro* translated protein (5 μl per reaction) or nuclear extract (5 μl per reaction) and incubated at room temperature for 20 min. In the control reactions the interaction of the crude nuclear extract and the radiolabelled oligonucleotides were competed with

Table 2. Summary of oligonucleotide sequences used as probes in EMSA studies.

Probe	*Site*	*Sense strand oligonucleotide sequence*[a]
F11	β-globin	AGGCGCGCCCCCAGGGATGTAATTACGTCCCTCCC**CCGC**TA**GGGGGC**AGCAGGCGCGCCT
MutF11	β-globin	AGGCGCGCCCCCAGGGATGTAATTACGTCCCTCCC**AATA**TA**TTGTTA**AGCAGGCGCGCCT
59910	A	GTGACATTGTGTTTATGGGGTGGTCAGCCATC**CCGC**CA**GGTGCA**CAACGTGTGGCTCGGT
64130	B	GGCTTCTCACTGGGAGGCTCAACTTC**CCGC**AG**GGGCGC**CAGGCCCTGACAGGAGAGACTG
65310	C	ATGTGCCAGCTCCGGAGCCGAGG**CCGC**GG**CAGGGC**TCGGCGCAACATGTGTCGCTGCCTG
65840	D	CGATGGATGTTCCGAAAA**CCGC**CA**GGTGTG**GGATCTGCGCCCCGACAGCCCCACCTTGGCC
68030	E	GGCGCTCCAGAGGGGCCTGGCGCAGGT**CCGC**TG**GGT**G**GC**CCCACCTTCCTCTGGGGCTGGG
68261	F	GCACCCCCTGTCCATCCAGGC**CCGC**CG**GGCGCG**GGCCGGGGCGCCGCAGGTGGCTGCTGC

[a]Bold letters indicate conserved bases with other CTCF-binding sites.

100-fold molar excess of the corresponding unlabelled oligonucleotides.

Chromatin immunoprecipitation (ChIP)

Chromatin immunoprecipitation (ChIP) was performed as described previously (Filippova *et al.* 2001) with the following modifications. Briefly, formaldehyde fixation buffer (10X buffer: 50 mM HEPES, pH 8.0, 1 mM EDTA, 0.5 mM EGTA, 11 mM NaCl and 11% formaldehyde) was added to the media of cultured human fibroblasts for 10 min with gentle shaking and allowing the reaction to quench with 125 mM glycine. Cells were rinsed twice with cold PBS and lysed on ice for 5 min with 600 μl of buffer A (1% SDS, 10 mM EDTA, pH 8.0, 50 mM Tris-HCl, pH 8.1, and protease inhibitor cocktail) and then sonicated for 2 min. After the centrifugation, the supernatant was pre-cleared with 100 μl of protein A/G beads solution (Santa Cruz) that had been pre-absorbed with 1.5 μg of sonicated single-strand salmon sperm DNA. An aliquot of the supernatant was saved as total input chromatin. This was processed with the eluted immunoprecipitation beginning at the proteinase K digestion step as described below. Immunoprecipitation was carried out by mixing 5 μg CTCF pre-immune serum with or without 5 μg of anti-CTCF antibody (Upstate Biotech) and incubating at 4°C overnight with rotation. Beads were pelleted by centrifugation, washed three times in buffer B (0.1% SDS, 1% Triton X-100, 2 mM EDTA, pH 8.0, 150 mM NaCl, 20 mM Tris-HCl, pH 8.1) and washed once in buffer C (0.1% SDS, 1% Triton X-100, 2 mM EDTA, pH 8.0, 500 mM NaCl, 20 mM Tris-HCl, pH 8.1). Chromatin was eluted from beads by the addition of 450 μl of buffer D (1% SDS, 0.1 M NaHCO3) with rotation at room temperature for 15 min. Proteinase K (100 μg/ml) was added to the elution and then heated at 65°C overnight. The DNA solution was then extracted twice with phenol/chloroform followed by ethanol precipitation and resuspended in 50 μl of sterile distilled water. The DNA solution (5–10 μl) was used as a template for PCR with primers to site C and H19. The following primer pairs were used:

Site C: C-Forward 5′ –CTGTCACCCTAAGATAG ATCCTCGG– 3′;
C-Reverse 5′ –GGAAAGTACTTTTAGAG GGCGGACT– 3′;

H19: H19 Forward 5′ –GTGCACTATTGAGGGTTG GGAAG– 3′;
H19 Reverse 5′ –CCATGAGTGTCCTATACC TCACGAC– 3′.

Methylation-Sensitive Restriction Digestion Assay

For this analysis, we digested the genomic DNA samples with *Eco*RI (Amersham Pharmacia), a methylation-insensitive enzyme, to yield a 3.9-kb fragment that can be separated and visualized on a 1.2% agarose gel. Each sample was then digested with *Sac*II (Amersham Pharmacia), which cannot cut methylated CpG dinucleotides (methylation-sensitive). After transfer to nylon membrane, the blot was probed with a ~2 kb amplification product of genomic sequence located approximately 3.8 kb upstream of *MEG3*. The ~2 kb probe (see Figure 1) was amplified from the BAC (GenBank accession no. AL117190) using the following primers (Forward: 5′ –CTT TCT CTC TTT GCC TCG CAT TC–3′ and Reverse: 5′–CGC TTA CAA CTC ATG TTC GCG–3′). Hybridization and washing were performed according to standard protocols.

Results

Identification of CTCF binding sites of MEG3

To identify additional consensus sequence motifs for binding of the protein CTCF around the *MEG3* transcriptional start site, we carried out a BLAST search of the human chromosome 14 genomic sequence of BAC RP11-123M6, which contains *MEG3* (GenBank accession no. AL117190). Results revealed seven sites that share similar consensus sequence motifs (CCGCNNGGNGNC) as potential CTCF-binding sites (Figure 1 and Table 3). Four sites are located upstream of *MEG3* (sites A to D), whereas three sites (sites E to G) are located within intron 2 (Figure 1 and Table 3). Site E was identified previously by Wylie *et al.* (2000), who also identified an additional downstream site (Wylie *et al.* 2000). Based on their published sequence and reported nucleotide position (Wylie *et al.* 2000) we assumed that this additional site is 'G' (Figure 1 and Table 3), which is located on the complementary strand. We recognized a third site within intron 2 (site F),

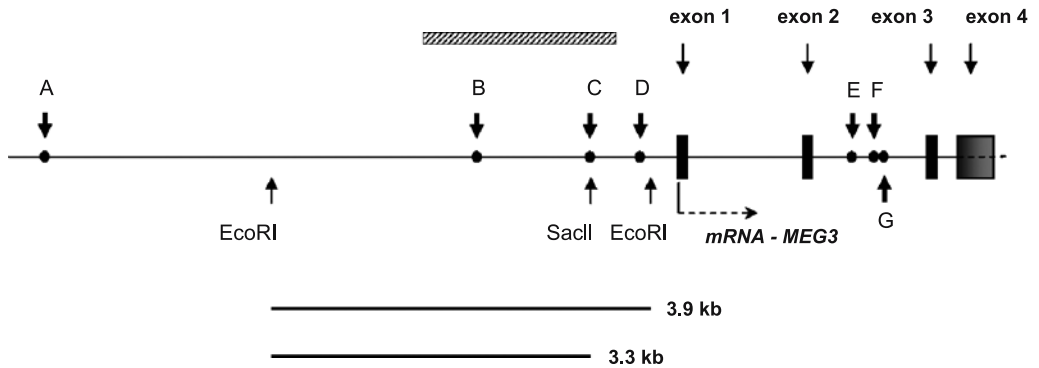

Figure 1. Schematic representation of the genomic region around the promoter of the gene *MEG3*. Putative transcription start site is indicated (dotted arrow, mRNA-*MEG3*). Exons 1, 2, 3 and 4 (incomplete) are represented with filled rectangles. Upstream (**A** to **D**) and intronic (**E** to **F**) CTCF consensus-target sites are indicated with filled circles. The position of *Eco*RI and *Sca*II restriction sites mentioned in the text, as well as the expected restriction fragments (horizontal lines below, 3.9 kb and 3.3 kb), is indicated. The probe used in Figure 2 is represented by the hatched rectangle.

located between sites E and G, which may correspond to an 'intronic site' reported by Paulsen *et al.* (2001).

We performed an EMSA using 60 bp of DNA sequence corresponding to the CTCF binding consensus motifs and flanking sequences unique to sites A to F (Table 2). We used the chicken β-globin CTCF binding sequence, F11 (Chung *et al.*1993, 1997), as a positive control, and a mutated CTCF binding consensus motif of F11, mutF11, as a negative control. We observed a single DNA–protein complex with the same mobility in all the sites investigated when compared to the positive F11 control (Figure 2A). Competition with 100-fold molar excess of unlabelled probe diminished the binding of the CTCF protein to each fragment (Figure 2A). Although the mutant F11 showed reduced binding (Figure 2A), the excess unlabelled mutF11 was not able to compete effectively with the radiolabelled F11 for CTCF binding (Figure 2A). Binding of mutF11 by CTCF is not unexpected because CTCF is known to be quite promiscuous in its binding to DNA (Hark *et al.* 2000). However, the specificity of binding is demonstrated by the inability of the mutF11 to compete with wild-type F11 sequence.

Table 3. Summary of CTCF-binding consensus sequences 5′ of *MEG3*

Genomic position	*Site*	*Sequence*[a]
59938	A	CCATC **CCGC** CA **GG** T **G**CA CAAC
64152	B	ACTTC **CCGC** AG **GG** G C**GC** CAGG
65329	C	CGAGG **CCGC** GG CA G **GGC** TCGG
65853	D	CGAAA **CCGC** CA **GG** T **G**TG GGAT
68053	E[b]	CAGGT **CCGC** TG **GG** T **GGC** CCAC
68277	F	CAGGC **CCGC** CG **GG** C **G**CG GGCC
68345	G[b,c]	ACACA **CCGC** TA **GG** G **GGC** GCGC

[a]Bold letters indicate conserved bases with other CTCF-binding sites.
[b]Reported previously by Wylie *et al.* (2000).
[c]Sequence on the complementary strand.

A chromatin immunoprecipitation (ChIP) strategy was used to demonstrate binding of CTCF to the DNA targets identified in this study. As proof of principle, we chose site C for further investigation. Figure 2B shows that CTCF *in-vivo* is bound to site C of *MEG3* in chromatin of primary cultured fibroblasts from a normal control individual. Binding was also demonstrated for an H19 control DNA target.

A CTCF binding site upstream of MEG3 is differentially methylated

Analyses of the sequences of the putative CTCF binding sites identified a site for the restriction enzyme *Sac*II, which overlaps site C (Figure 1). The site was amenable for molecular studies designed to confirm the functionality of site C.

Previous methylation analyses of *MEG3* intronic CTCF sites were consistent with allele-specific differential methylation (Wylie *et al.* 2000). However, this was not confirmed in those studies, nor was parent-specific methylation demonstrated (Wylie *et*

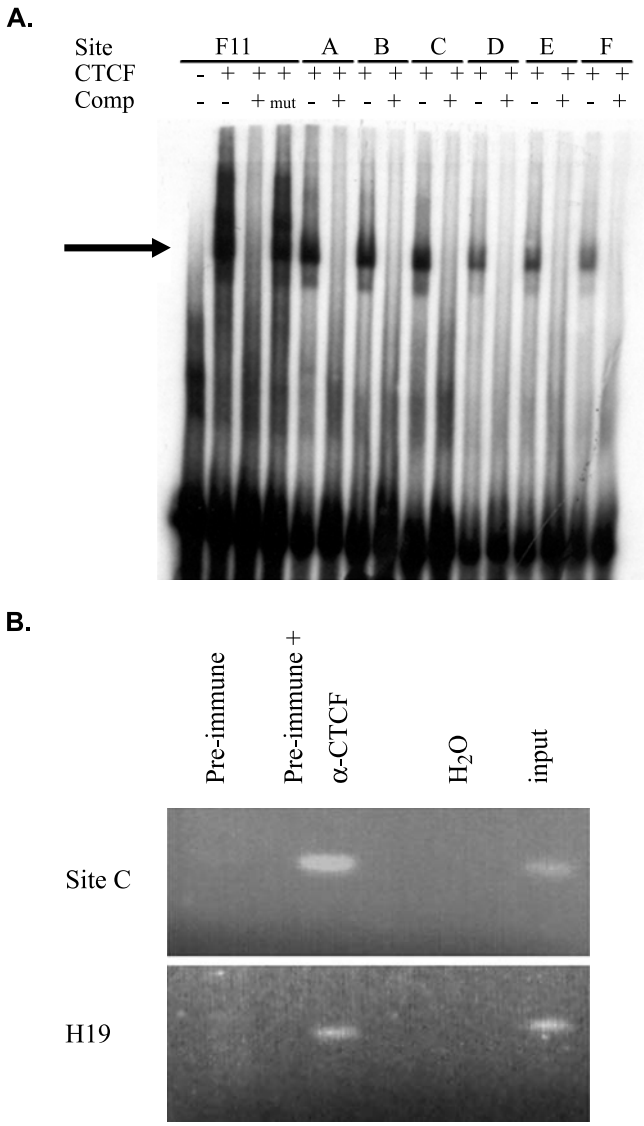

Figure 2. CTCF target sites A to F from *MEG3* bind the CTCF protein. A to F correspond to the six CTCF consensus binding motifs on 14q32 analysed in this study (Tables 2 and 3). Lanes labelled F11 and mutF11 correspond to wildtype and mutated CTCF binding site at the chicken β-globin locus, respectively (Table 2). Arrows indicate the position of DNA–protein complexes. **A:** CTCF protein was obtained from *in-vitro* translation and added to all reactions (+) except for a control F11 lane (–). The chicken β-globin F11 oligonucleotide (Table 2) is competed with non-radiolabelled F11, but not mutated F11 (mut). The various oligonucleotides containing sites A to F form a complex with CTCF. Specificity of these complexes is confirmed in competition assays using the corresponding non-radiolabelled oligonucleotides (+ in the 'Comp' row). **B:** Chromatin immunoprecipitation assay shows *in-vivo* binding of CTCF to the *MEG3* site C. *Total input* chromatin was used as a positive control for the PCR reaction. A pre-immune serum and a mixture of pre-immune and α-CTCF antibodies were used to immunoprecipitate CTCF–DNA complexes. CTCF bound to *H19*, on human chromosome 11, was used as a control in these studies. Lane 3 shows no amplification in the water control and lane 4 shows amplification of all three sites in total input chromatin. For experimental details see Materials and Methods section.

al. 2000). We investigated differential DNA methylation of a CpG island upstream of *MEG3* in several cell lines derived from patients with either maternal or paternal UPD14. Our search for CpG islands was based on the known genomic sequence and potential methylation-sensitive restriction enzyme sites located in and around *MEG3* (GenBank accession no. AL117190). Differential methylation was examined

using a methylation-sensitive restriction enzyme digestion, followed by Southern hybridization of genomic DNA isolated from maternal or paternal UPD14 patients. The probe is a ~2 kb amplification product of genomic sequence (see Materials and Methods section) overlapping the putative CTCF target sites B and C upstream of *MEG3* (Figure 1).

According to the anticipated restriction digestion map of the genomic sequence (GenBank accession no. AL117190), the probe was expected to hybridize to a single 3.9-kb fragment of genomic DNA following digestion with *Eco*RI (Figure 1). If site C were unmethylated, it is expected that digestion with *Sac*II would generate 3.3-kb and 0.6-kb fragments. Alternatively, if the region were methylated, *Sac*II would be unable to cut the methylated CpG dinucleotide at –3849 bp, and a single 3.9-kb band (the same as the *Eco*RI single-digestion fragment) would be seen (Figure 1).

Patients 1 and 2 (Figure 3A) and the child in family 1 (Figure 3B), all of whom have maternal UPD 14, show only the smaller 3.3-kb band after double digestion with *Eco*RI and *Sac*II. This pattern is consistent with two unmethylated alleles on chromosome 14 at the *Sac*II site resulting from maternal disomy 14. Patient 5 shows a difference in hybridization intensity between the 3.9-kb and 3.3-kb fragment (Figure 3A), with the 3.3-kb fragment (unmethylated *Sac*II site) more intense than the 3.9-kb fragment. This result is consistent with the previously reported low-level mosaicism for trisomy 14 and a predominant maternal UPD14 cell line (Antonarakis *et al*. 1993). Patients 3 and 4 and the child in family 2, all of whom have paternal UPD14, show a hybridization pattern consistent with methylation on both chromosomes 14 at the *Sac*II site (Figure 3). Thus, the region containing the putative CTCF binding site C is differentially methylated, with methylation of the paternal allele.

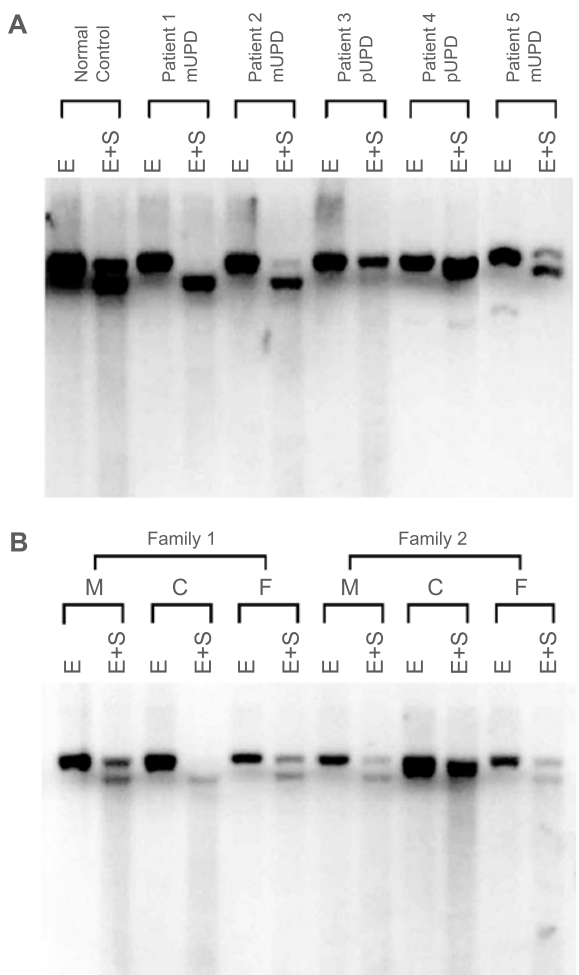

Figure 3. Methylation-sensitive restriction enzyme digestion of *MEG3* site C in cell lines derived from patients with UPD14. **A:** Genomic DNA extracted from a normal control (*Normal control*) and cell lines derived from patients with UPD14 (*Patients 1 to 5*) was digested with *Eco*RI (E) or with *Eco*RI followed by *Sac*II (E+S). In the *Normal control* (*E+S*) both the 3.9-kb and 3.3-kb fragments are visible, corresponding to the methylated paternal allele and unmethylated maternal allele, respectively. E+S digestion of genomic DNA derived from maternal UPD14 patients shows only the unmethylated (3.3-kb) allele (*Patients 1 and 2*). DNA derived from a mosaic trisomy 14/maternal UPD 14 patient shows a minor band at 3.9 kb (*Patient 5*), indicating the presence of a methylated paternal allele in the trisomy 14 cell line and a more intense band at 3.3 kb corresponding to the two copies of an unmethylated allele present in both the trisomic cell line and the predominant maternal UPD 14 cell line. DNA derived from paternal UPD14 (*Patients 3 and 4*) shows no digestion of the *Eco*RI fragment with *Sac*II, demonstrating the presence of two methylated alleles at the *Sac*II site. **B:** Analysis of methylation status at the *Sac*II site in two families. DNA digestions from the mother (M), child with UPD14 (C) and father (F) are shown for families 1 and 2. Parental E+S lanes (lanes 2, 6, 8, 12) show the normal pattern of both the 3.9-kb and 3.3-kb fragments. The child in family 1 shows only the 3.3-kb fragment following E+S digestion (lane 4), consistent with two unmethylated alleles and maternal UPD14. The child in family 2 shows only the 3.9-kb fragment following E+S digestion (lane 10), consistent with two methylated alleles and paternal UPD14.

Discussion

CTCF is an evolutionarily conserved, zinc finger phosphoprotein that binds to ~50 bp target sites through combinatorial use of its 11 zinc fingers. These binding sites have remarkable sequence variation (Chung *et al*. 1993, 1997). The CTCF protein is quite promiscuous in its DNA interactions, as it has interacted with a great number of insulators examined (Hark *et al*. 2000). Recently, it has been demonstrated that CTCF-dependent chromatin insulation is regulated by poly(ADP-ribosyl)ation of the protein (Yu *et al*. 2004). CTCF has a key role in the control of imprinting at the *Igf2*/*H19* region on 11p15 (Szabo *et al*. 2004).

Wylie *et al*. (2000) reported the identification of two putative CTCF binding sites at the second intron of *MEG3*. One of the sites reported by Wylie *et al*. (2000) is present on the sense strand (site E; Figure 1), whereas a downstream site is located on the complementary strand (site G; Figure 1). We identified an additional potential CTCF binding site (site F; Figure 1) located between the intronic sites recognized by Wylie *et al*. (2000). This region was originally supposed to be upstream of the transcriptional start site of *MEG3* (Wylie *et al*. 2000). Paulsen *et al*. (2001), however, demonstrated that the potential CTCF target sites reported by Wylie *et al*. are located in intron 2 (Paulsen *et al*. 2001). Their conclusion was based on the sequence of human ETS AW163035, which contains transcribed sequences from the previously considered intron 1 (Paulsen *et al*. 2001). A recent assembly of the *MEG3* coding sequences, which includes an additional alternatively spliced exon of the gene (exon 5b; Zhang *et al*. 2003), did not consider the ETS reported by Paulsen *et al*. (2001) and, for this reason, lacks exon 2. These conflicting reports concerning the *MEG3* coding sequence (Schmidt *et al*. 2000, Wylie *et al*. 2000, Zhang *et al*. 2003) may be explained if *MEG3* has complex alternative splicing patterns as has been described for the homologous mouse gene *Gtl2* (Schuster-Gossler *et al*. 1998, Croteau *et al*. 2003).

In addition to the intron 2 CTCF consensus binding motif sites mentioned above, we found four more sites located between 0.4 and 6.0 kb upstream of *MEG3* (sites A–D; Figure 1). Using conventional analyses of protein–DNA complexes, we demonstrated CTCF binding to all of these sites. In addition, using chromatin immunoprecipitation analysis we demonstrated *in-vivo* binding of CTCF to site C. Also, we showed that site C, which is amenable to restriction enzyme analyses, has parental-specific methylation: a *Sac*II restriction site, overlapping site C, was unmethylated on the maternally inherited chromosomes 14 and methylated on the paternally inherited chromosomes 14. These results are consistent with the previous reports of monoallelic (maternal) expression of *MEG3* and allele-specific methylation near this gene in humans, mice, and sheep (Miyoshi *et al*. 2000, Schmidt *et al*. 2000, Takada *et al*. 2000, Wylie *et al*. 2000, Charlier *et al*. 2001a,b). Our findings support a correlation between the methylation status of the upstream CpG island and the monoallelic expression of *MEG3*.

Like the *IGF2*/*H19* domain, *DLK1* and *MEG3* are separated by approximately 90 kb of genomic sequence (Wylie *et al*. 2000). Both *IGF2* and *DLK1* are paternally expressed and maternally methylated, whereas *H19* and *MEG3* are maternally expressed and paternally methylated and are thought to function as noncoding RNAs (Wylie *et al*. 2000). The *IGF2*/*H19* imprinted domain has been extensively studied and is well characterized (Bird 1995, Kanduri *et al*. 2000a,b, Sasaki *et al*. 2000). A differentially methylated region, located approximately 4 kb from the start site of *H19* and containing seven CTCF binding sites, functions as an insulator regulating expression of *IGF2* and *H19*. On the unmethylated allele, CTCF binds to DNA and facilitates enhancer interactions allowing for transcription of *H19* but not *IGF2*. On the methylated allele, CTCF cannot bind to the methylated binding sites, which permits transcription of *IGF2* but fails to enhance *H19* expression (Wolffe 2000).

Given the substantial similarity between the genomic arrangement and the expression patterns of the *IGF2*/*H19* domain and the *DLK1*/*MEG3* domain, we speculate that there is a common regulatory mechanism governing the differential expression of these imprinted genes. The CTCF binding sites identified in this study may function as an insulator, regulating the expression of *DLK1* and *MEG3* in a manner similar to the regulation of the *IGF2*/*H19* domain. Binding of the insulator by CTCF would allow for interaction with a putative enhancer and allow for expression of *MEG3* but not *DLK1*, whereas methylation of the CTCF DNA binding sites would prevent binding of the protein and block interactions of an enhancer with *MEG3,* while allowing for interaction

of an enhancer with *DLK1* and expression of this paternally expressed gene. Further investigation is required to establish the role of CTCF binding on the imprinting of *DLK1/MEG3*.

Electronic Database Information

Accession numbers and URLs referenced in this article are as follows:

Imprinted gene loci and references, http://www.mgu.har.mrc.ac.uk/imprinting/imprin-generef.html (for information on imprinted loci) GenBank accession number AL117190, http://www.ncbi.nlm.nih.gov/entrez/ (for human genomic sequence of this region).

Acknowledgements

The authors thank the families who provided samples for this investigation and the clinicians involved in the ascertainment of families and specimen collection. Additionally, the authors thank Dr. J. Hoover (Johns Hopkins University, Baltimore, MD) for the cell line on patient 5, Drs G. Felsenfeld and A. West (NIH, Bethesda, MD) for the chicken CTCF plasmid (p4bfull), the CTCF preimmune IgG serum and for helpful discussions, and Drs A.L. Beaudet, Y.H. Jiang, and T. Sahoo (Baylor College of Medicine) for their critical reviews of the manuscript. This research was supported in part by grants from the National Institutes of Health R03 HD38433 (L.G.S.) and K23 HD40843 (V.R.S.).

References

Antonarakis SE, Blouin JL, Maher J, Avramopoulos D, Thomas G, Talbot CC, Jr (1993) Maternal uniparental disomy for human chromosome 14, due to loss of a chromosome 14 from somatic cells with t(13;14) trisomy 14. *Am J Hum Genet* **52**: 1145–1152.

Bell AC, Felsenfeld G (2000) Methylation of a CTCF-dependent boundary controls imprinted expression of the Igf2 gene. *Nature* **405**: 482–485.

Bell AC, West AG, Felsenfeld G (1999) The protein CTCF is required for the enhancer blocking activity of vertebrate insulators. *Cell* **98**: 387–396.

Berend SA, Bejjani BA, McCaskill C, Shaffer LG (2002) Identification of uniparental disomy in phenotypically abnormal carriers of isochromosomes or Robertsonian translocations. *Am J Med Genet* **111**: 362–365.

Bird AP (1995) Gene number, noise reduction and biological complexity. *Trends Genet* **11**: 94–100.

Charlier C, Segers K, Karim L *et al*. (2001a) The callipyge mutation enhances the expression of coregulated imprinted genes in cis without affecting their imprinting status. *Nat Genet* **27**: 367–369.

Charlier C, Segers K, Wagenaar D *et al*. (2001b) Human–ovine comparative sequencing of a 250-kb imprinted domain encompassing the callipyge (clpg) locus and identification of six imprinted transcripts: DLK1, DAT, GTL2, PEG11, antiPEG11, and MEG8. *Genome Res* **11**: 850–862.

Chung JH, Whiteley M, Felsenfeld G (1993) A 5′ element of the chicken beta-globin domain serves as an insulator in human erythroid cells and protects against position effect in *Drosophila*. *Cell* **74**: 505–514.

Chung JH, Bell AC, Felsenfeld G (1997) Characterization of the chicken beta-globin insulator. *Proc Natl Acad Sci USA* **94**: 575–580.

Coveler KJ, Yang SP, Sutton R *et al*. (2002) A case of segmental paternal isodisomy of chromosome 14. *Hum Genet* **110**: 251–256.

Croteau S, Charron MC, Latham KE, Naumova AK (2003) Alternative splicing and imprinting control of the Meg3/Gtl2-Dlk1 locus in mouse embryos. *Mamm Genome* **14**: 231–241.

Engel E (1980) A new genetic concept: uniparental disomy and its potential effect, isodisomy. *Am J Med Genet* **6**: 137–143.

Filippova GN, Thienes CP, Penn BH *et al*. (2001) CTCF-binding sites flank CTG/CAG repeats and form a methylation-sensitive insulator at the DM1 locus. *Nat Genet* **28**: 335–343.

Georgiades P, Chierakul C, Ferguson-Smith AC (1998) Parental origin effects in human trisomy for chromosome 14q: implications for genomic imprinting. *J Med Genet* **35**: 821–824.

Hark AT, Schoenherr CJ, Katz DJ, Ingram RS, Levorse JM, Tilghman SM (2000) CTCF mediates methylation-sensitive enhancer-blocking activity at the H19/Igf2 locus. *Nature* **405**: 486–489.

Kanduri C, Holmgren C, Pilartz M *et al*. (2000a) The 5′ flank of mouse H19 in an unusual chromatin conformation unidirectionally blocks enhancer–promoter communication. *Curr Biol* **10**: 449–457.

Kanduri C, Pant V, Loukinov D *et al*. (2000b) Functional association of CTCF with the insulator upstream of the H19 gene is parent of origin-specific and methylation-sensitive. *Curr Biol* **10**: 853–856.

Kobayashi S, Wagatsuma H, Ono R *et al*. (2000) Mouse Peg9/Dlk1 and human PEG9/DLK1 are paternally expressed imprinted genes closely located to the maternally expressed imprinted genes: mouse Meg3/Gtl2 and human MEG3. *Genes Cells* **5**: 1029–1037.

Lewis A, Murrell A (2004) Genomic imprinting: CTCF protects the boundaries. *Curr Biol* **14**: R284–R286.

McGowan KD, Weiser JJ, Horwitz J *et al*. (2002) The importance of investigating for uniparental disomy in prenatally identified balanced acrocentric rearrangements. *Prenat Diagn* **22**: 141–143.

Miyoshi N, Wagatsuma H, Wakana S *et al*. (2000) Identification of an imprinted gene, Meg3/Gtl2 and its human homologue MEG3, first mapped on mouse distal chromosome 12 and human chromosome 14q. *Genes Cells* **5**: 211–220.

Morison IM, Reeve AE (1998) A catalogue of imprinted genes and

parent-of-origin effects in humans and animals. *Hum Mol Genet* **7**: 1599–1609.

Ohlsson R, Renkawitz R, Lobanenkov V (2001) CTCF is a uniquely versatile transcription regulator linked to epigenetics and disease. *Trends Genet* **17**: 520–527.

Paulsen M, Takada S, Youngson NA *et al*. (2001) Comparative sequence analysis of the imprinted Dlk1-Gtl2 locus in three mammalian species reveals highly conserved genomic elements and refines comparison with the Igf2-H19 region. *Genome Res* **11**: 2085–2094.

Pentao L, Lewis RA, Ledbetter DH, Patel PI, Lupski JR (1992) Maternal uniparental isodisomy of chromosome 14: association with autosomal recessive rod monochromacy. *Am J Hum Genet* **50**: 690–699.

Prawitt D, Enklaar T, Gartner-Rupprecht B *et al*. (2005) Microdeletion of target sites for insulator protein CTCF in a chromosome 11p15 imprinting center in Beckwith–Wiedemann syndrome and Wilms' tumor. *Proc Natl Acad Sci USA* **102**: 4085–4090.

Robin NH, Harari-Shacham A, Schwartz S, Wolff DJ (1997) Duplication 14(q24.3q31) in a father and daughter: delineation of a possible imprinted region. *Am J Med Genet* **71**: 361–365.

Sasaki H, Ishihara K, Kato R (2000) Mechanisms of Igf2/H19 imprinting: DNA methylation, chromatin and long-distance gene regulation. *J Biochem (Tokyo)* **127**: 711–715.

Schiff R, Reddy P, Ahotupa M *et al*. (2000) Oxidative stress and AP-1 activity in tamoxifen-resistant breast tumors *in-vivo*. *J Natl Cancer Inst* **92**: 1926–1934.

Schmidt JV, Matteson PG, Jones BK, Guan XJ, Tilghman SM (2000) The Dlk1 and Gtl2 genes are linked and reciprocally imprinted. *Genes Dev* **14**: 1997–2002.

Schuster-Gossler K, Bilinski P, Sado T, Ferguson-Smith A, Gossler A (1998) The mouse Gtl2 gene is differentially expressed during embryonic development, encodes multiple alternatively spliced transcripts, and may act as an RNA. *Dev Dyn* **212**: 214–228.

Sparago A, Cerrato F, Vernucci M, Ferrero GB, Silengo MC, Riccio A (2004) Microdeletions in the human H19 DMR result in loss of IGF2 imprinting and Beckwith–Wiedemann syndrome. *Nat Genet* **36**: 958–960.

Sutton VR, Shaffer LG (2000) Search for imprinted regions on chromosome 14: comparison of maternal and paternal UPD cases with cases of chromosome 14 deletion. *Am J Med Genet* **93**: 381–387.

Sutton VR, Coveler KJ, Lalani SR, Kashork CD, Shaffer LG (2002) Subtelomeric FISH uncovers trisomy 14q32: lessons for imprinted regions, cryptic rearrangements and variant acrocentric short arms. *Am J Med Genet* **112**: 23–27.

Szabo PE, Tang SH, Silva FJ, Tsark WM, Mann JR (2004) Role of CTCF binding sites in the Igf2/H19 imprinting control region. *Mol Cell Biol* **24**: 4791–4800.

Takada S, Tevendale M, Baker J *et al*. (2000) Delta-like and gtl2 are reciprocally expressed, differentially methylated linked imprinted genes on mouse chromosome 12. *Curr Biol* **10**: 1135–1138.

Towner D, Yang SP, Shaffer LG (2001a) Prenatal ultrasound findings in a fetus with paternal uniparental disomy 14q12-qter. *Ultrasound Obstet Gynecol* **18**: 268–271.

Towner DR, Shaffer LG, Yang SP, Walgenbach DD (2001b) Confined placental mosaicism for trisomy 14 and maternal uniparental disomy in association with elevated second trimester maternal serum human chorionic gonadotrophin and third trimester fetal growth restriction. *Prenat Diagn* **21**: 395–398.

Walter CA, Shaffer LG, Kaye CI *et al*. (1996) Short-limb dwarfism and hypertrophic cardiomyopathy in a patient with paternal isodisomy 14: 45,XY,idic(14)(p11). *Am J Med Genet* **65**: 259–265.

Wolffe AP (2000) Transcriptional control: imprinting insulation. *Curr Biol* **10**: R463–R465.

Wylie AA, Murphy SK, Orton TC, Jirtle RL (2000) Novel imprinted DLK1/GTL2 domain on human chromosome 14 contains motifs that mimic those implicated in IGF2/H19 regulation. *Genome Res* **10**: 1711–1718.

Yu W, Ginjala V, Pant V *et al*. (2004) Poly(ADP-ribosyl)ation regulates CTCF-dependent chromatin insulation. *Nat Genet* **36**: 1105–1110.

Zhang X, Zhou Y, Mehta KR *et al*. (2003) A pituitary-derived MEG3 isoform functions as a growth suppressor in tumor cells. *J Clin Endocrinol Metab* **88**: 5119–5126.

Zimarino V, Wu C (1987) Induction of sequence-specific binding of Drosophila heat shock activator protein without protein synthesis. *Nature* **327**: 727–730.