CTCF is a highly conserved transcription factor which recognizes and binds to various target DNA sequences (28, 31, 54, 72). Different regulatory roles performed by CTCF include promoter activation (33, 86) or repression (32), hormone-responsive gene silencing (16), regulation of cell growth and proliferation (77, 84), differentiation (56, 61, 83), and apoptosis (26). CTCF is also involved in the regulation of methylation-dependent chromatin insulation (11, 42) and chromatin barrier (91) functions and also affects the transcription of rRNA (84). It is well established that PARylation modulates the activity of various nuclear proteins (88) and posttranslational modifications, such as phosphorylation (29, 53), SUMOylation (64), and poly(ADP-ribosyl)ation (PARylation) (18, 52). In particular, PARylation was found previously to regulate CTCF insulator (92) and chromatin barrier (91) functions and also affects the transcription of rRNA (84).

PARylation is a covalent modification of proteins catalyzed by the poly(ADP-ribose) polymerases (PARPs), a large family of 18 proteins encoded by different genes (44, 80), of which the best-studied isoform is PARP-1 (6). Recent evidence, however, suggests that there may be only six true PARPs and that the remaining family members are mono(ADP-ribosyl) transferases (44, 51). The PARylation reaction involves a processive sequential transfer of ADP-ribose moieties from coenzyme NAD$^+$ to an acceptor protein (44, 60). Although it is generally perceived that there is no specific consensus site for the PARylation reaction (43), it has been reported previously that glutamic and aspartic acid (3, 43) and lysine (4) residues of putative acceptor proteins can be utilized as ADP-ribose acceptor sites. The catalyzed reaction results in an ADP-ribose polymer (PAR) chain of variable length, from a few to 200 ADP-ribose units, attached to the protein. This polymer chain may also be branched in structure, with a frequency of branching of 1 per 20 to 30 ADP-ribose residues (69). The modification is transient, as the PARs are rapidly degraded by poly-(ADP-ribose) glycohydrolase (PARG) or other proteins with phosphodiesterase activity (14, 37, 79).

It is well established that PARylation modulates the activities of PARP-1 and various nuclear proteins (21, 23, 57–59, 85).
and is implicated in DNA repair, recombination, cell proliferation, cell death, and the regulation of nuclear and chromatin functions (21, 44, 68, 80). Intriguingly, CTCF appears to act as a link between PARylation and DNA whereby CTCF activates PARP-1, leading to DNA hypomethylation (40).

Although we have gained knowledge from previous studies, the role of PARylation in the regulation of different CTCF functions and also mechanistic aspects of this regulation are still not well understood. In this investigation, we identified PARylation sites in CTCF and generated a mutant deficient in PARylation, which was employed to investigate the importance of CTCF PARylation in transcriptional regulation and the control of cell proliferation. Regulation of insulator function was also examined with a newly generated isogenic insulator reporter cell line. In all functional tests, we observed a loss of function of the non-PARylated CTCF. Our observation of PARylation-independent association of CTCF and PARP-1 at the mouse H19 ICR, in the new isogenic insulator system, led to a model whereby PARP-1 (and possibly PARG) molecules form complex functional complexes with CTCF at sites on DNA. We suggest that PARylation and de-PARylation of CTCF are dynamic processes in response to cellular signals and/or the environment and are an essential mechanism in the regulation of CTCF function.

MATERIALS AND METHODS

Cell lines and transfections. Human cell lines used in this study were HeLa (cervical carcinoma) and 293T (embryonic kidney) cells and B4 cells (transgenic mouse cervical carcinoma) and 293T (embryonic kidney) cells and B4 cells (transgenic mouse cervical carcinoma) and MDA435 (breast carcinoma) and MCF7 (breast carcinoma) cells, and NeoNeo (Lonza), and MDA435 (breast carcinoma) and MCF7 (breast carcinoma) cells, maintained in Dulbecco’s modified Eagle’s medium supplemented with 1-glutamine, glucose pyridoxine-HCl, and NaHCO3, (Lonza), and MDA435 (breast carcinoma) and MCF7 (breast carcinoma) cells, maintained in RPMI 1640 with 1-glutamine (Lonza). All media were supplemented with 10% fetal bovine serum (Biowest) and 50 μg/ml gentamicin (Invitrogen). Transient transfections of cell lines were carried out using a calcium phosphate transfection protocol (26). The β-galactosidase expression construct pCM110 (Invitrogen) was used for normalization of the transfection efficiency. The nucleoplasmic extracts from MCF7 cells (106) enriched with BrdU by cells was visualized by incubation with anti-BrdU antibody (Sigma) and then a tetramethyl rhodamine isocyanate (TRITC)-conjugated secondary antibody (Dako), staining of cells with DAPI (4′,6-diamidino-2-phenylindole), and fluorescence microscopy analysis. A minimum of 1,000 cells were assessed for proliferation based upon the incorporation of BrdU (red fluorescence) and positivity for EGFP (green fluorescence). An average percentage of proliferating cells from each experiment, carried out in triplicate and repeated three times, was calculated.

PARylation. The nucleoplasmic extracts from MCF7 cells (106) and from the cell line expressing CTCF-180 isoform (84), were obtained as described previously (92). The extracts were incubated with 1, 2, 3, or 5 mU of PARG enzyme (Axxora) and then subjected to affinity purification on coupled polyclonal antibodies as described previously (15). The immunopurified extracts were analyzed by Western blotting using an anti-CTCF antibody (Abcam).

Indirect immunostaining and analysis of colocalization of CTCF and PARP-1. Immunostaining and analysis of colocalization were performed for both wild-type and mutant forms, were prepared by excision of the relevant CTCF constructs using XbaI and EcoRI and insertion of the fragments into pEGFP-C1 (Clontech). A pEGFP-C1-luciferase gene reporter construct was used in the reporter assays, as described previously (33). The luciferase reporter assay kit was used to measure luciferase activity according to the instructions of the manufacturer (Promega). The luciferase reporter assay kit was used to measure luciferase activity according to the instructions of the manufacturer (Promega).

Generation of the CTCF-expressing vectors and the CTCF mutants deficient in PARylation. A CTCF expression construct was prepared by excision of the sequence encoding His-tagged CTCF from the vector for bacterial expression of CTCF, pET7.1 (7). The fragment containing CTCF cDNA was excited from pET7.1 by using Xhol and Nhel and inserted into the eukaryotic expression vector pCI (Promega) digested with Xhol and Xbal. The resulting cloning expression vector produced a full-length His-tagged wild-type CTCF protein (CTCF WT).

Mutant CTCF expression vectors were produced by site-directed mutagenesis using CTCF WT as a template, by the QuikChange method according to the instructions of the QuikChange manufacturer (Stratagene). The primers used in the generation of CTCF mutants were as follows: for CTCF mutant 1 (Mut1), forward primer 5′ GATGGTGCTCGCTGATATTTTGCAGCAAGCAGCA GGAGG 3′ and reverse primer 5′ CCTCCTGCGGTGCTGGCGCAAAAAATCG TACGCAGACACATC 3′; for Mut2, forward primer 5′ GGCGTTATACAGGCG GGGCGCAGAAAATG 3′ and reverse primer 5′ CATCTTTGCGCGCTGT TATAACGC 3′; and for Mut3 and Mut4, forward primer 5′ ACAGCAGCGG GTCGCTGTACGCGGGTTAAATGCGCGCGAAGTTGTTG 3′ and reverse primer 5′ CAACATTTCCGCCCCATTAACGGGTACGACCCCCCT GCTGT 3′.

The CTCF-enhanced green fluorescent protein (EGFP) fusion constructs, both wild-type and mutant forms, were prepared by excision of the relevant CTCF constructs using Xhol and EcoRI and insertion of the fragments into pEGFP-C1 (Clontech). A p9SARF-luciferase gene reporter construct was used in the reporter assays, as described previously (33).

Luciferase and β-galactosidase assays. The human cell lines were subjected to a luciferase reporter assay, following cotransfections and normalization of transfection efficiency by using a β-galactosidase assay, as described previously (66). The luciferase reporter assay kit was used to measure luciferase activity according to the instructions of the manufacturer (Promega).

BrdU assay. Forty-eight hours posttransfection of 293T, HeLa, and MDA435 cells with either EGFP-CTCF WT or EGFP-CTCF mutants, cells were incubated with bromodeoxyuridine (BrdU) as described previously (35). Incorporation of BrdU by cells was visualized by incubation with anti-BrdU antibody (Sigma) and then a tetramethyl rhodamine isocyanate (TRITC)-conjugated secondary antibody (Dako), staining of cells with DAPI (4′,6-diamidino-2-phenylindole), and fluorescence microscopy analysis. A minimum of 1,000 cells were assessed for proliferation based upon the incorporation of BrdU (red fluorescence) and positivity for EGFP (green fluorescence). An average percentage of proliferating cells from each experiment, carried out in triplicate and repeated three times, was calculated.

PARG assay. The nucleoplasmic extracts from MCF7 cells (106) and from the cell line expressing CTCF-180 isoform (84), were obtained as described previously (92). The extracts were incubated with 1, 2, 3, or 5 mU of PARG enzyme (Axxora) as reported earlier (15). The incubated extracts were analyzed by Western blotting using an anti-CTCF antibody (Abcam).

ISPLA. The in situ proximity ligation assay (ISPLA) was performed as described previously (34), with modifications. Briefly, cells were seeded for 48 h onto eight-well glass slides and fixed with 70% ethanol overnight at −20°C. Blocking was performed for 2 h at room temperature with blocking solution containing 3% bovine serum albumin (BSA) in phosphate-buffered saline (PBS). Slides were then incubated simultaneously with anti-rabbit PARP-1 antibody (diluted 1:400; a gift from L. Kraus) and anti-mouse CTCF antibody (diluted 1:50; from BD Transduction Laboratories) at 4°C overnight. After three washes in PBS with gentle shaking (each at least 15 min), secondary antibodies labeled with DyLight 488 (green; anti-mouse antibody) and DyLight 594 (red; anti-rabbit antibody) (both from Thermo Scientific) were mixed, and the slides were incubated at room temperature for 2 h and then washed with PBS. Images for PARP-1 and CTCF were taken using the same exposure times for both red and green channels before using the colocalization module (Volocity; Improvision), which provides the facility to visualize proteins and quantify colocalization in a pair of images. The colocalization was calculated at the same time (autocorrelation metric) location in the two images; by using the colocalization module, an overlay with a scatter plot including statistics with a Pearson coefficient was automatically generated (65).

ISPLA. The in situ proximity ligation assay (ISPLA) was performed as described previously (34), with modifications. Briefly, cells were seeded for 48 h onto eight-well glass slides and fixed with 70% ethanol overnight at −20°C. Cells were then incubated for 2 h at room temperature with blocking solution containing 3% BSA and PBS. Both primary antibodies, anti-CTCF (BD Transduction Laboratories) and anti-PARP-1 (a gift from L. Kraus), were individually incubated in blocking solution for 15 min and then mixed together and applied to the slides, and the slides were incubated overnight at 4°C. The slides were washed four times with PBS for 15 min with gentle shaking. Proximity probes were covalently linked via their 5′ ends to affinity-purified polyclonal antibodies by using a hydrazine linker (SoluLink, San Diego, CA). The donkey anti-mouse antibody (Jackson ImmunoResearch Laboratories) was linked to the nonpriming amine-modified oligonucleotide NH2-AAA AAA AAA AGA CGC TAA TAG TTA AGA AGC CT (T/UUU). Modified secondary antibodies were separately incubated in blocking solution for 15 min and then added to the slides, and the slides were incubated for 2 h at 37°C in a humid chamber and washed with PBS four times, each for 15 min. Hybridization, ligation, and rolling-circle amplification were performed using the Duolink detection kit (Olink Bioscience). At the final step, when the detection bond is applied, the biotin-labeled secondary antibodies (R&D Systems) were also added to visualize the individual primary antibodies. Finally, six washes in PBS at room temperature for 15 min each were performed. Slides were mounted with Vectorshield mounting medium containing DAPI (Vector Laboratories). Images were taken using a grid micro-
scope and analyzed by the Velocity software (Improvision). Controls for ISPLA are shown in Fig. S7A in the supplementary material.

IP. Immunoprecipitation (IP) was performed as described previously (20). In brief, MCF7 cells (2 × 10^6) were lysed in 500 µl of IP buffer, comprised of 25 mM Tris–HEPES (pH 8.0), 2 mM EDTA, 0.5% Tween 20, 0.25 M NaCl, and protease inhibitors. The lysates were incubated for 4 h with either anti-PAR-10H mouse monoclonal antibody (Alexis) or anti-PAR rabbit polyclonal antibody (Calbiochem) and then incubated for 2 h with either protein G or protein A Sepharose Fast Flow beads. Immunoprecipitated proteins were analyzed by Western blotting.

ChIP and re-ChIP (serial ChIP) assays. A chromatin IP (ChIP) assay was performed as described previously (19), using B4 cells either treated or not treated with PJ34 and either untransfected or transfected with the pC empty vector (pC EV) or the vector expressing CTCF WT or CTCF Mut4. In this assay, 10^7 cells were cross-linked with formaldehyde. Instead of being subjected to the usual sonication method to fragment the DNA, the DNA-protein complexes were digested overnight at 37°C with the Apol restriction endonuclease (New England Biolabs). In all ChIP experiments, the complete digestion of the cross-linked DNA by Apol was assessed by PCR analysis of the input material using the site 1 forward primer and the site 2 reverse primer (see below) for amplification: absence of the product confirmed the efficiency of digestion (data not shown). The DNA-protein complexes were immunoprecipitated using anti-His tag and anti-CTCF (for endogenous CTCF) or anti-PARP antibody (Alexis) and protein A4 Sepharose Fast Flow beads (Sigma).

PCR amplification of the purified DNA was performed using primers corresponding to the selected DNA sites within the integrated H19 1CR, as follows: site 1, forward primer 5′-AAAGACGGAGCATCTCTAT-3′ and reverse primer 5′-GGGGGCTCTTAGGTTT-3′; and site 2, forward primer 5′-GCGAGGACATGCTATT-3′ and reverse primer 5′-GACTAGCTGAACCTCG-3′. PCR conditions were 94°C for 3 min; 33 cycles of 94°C for 1 min, 57°C for 1 min and 72°C for 2 min; and 72°C for 2 min.

The re-ChIP assay was performed following the standard single ChIP assay (see above). For re-ChIP, the immunoprecipitates from the single ChIP were eluted with the buffer of 5 M urea–50 mM 2-mercaptoethanol, after which the precipitates were dialyzed against a solution of 10 mM Tris–HEPES (pH 8.0), 1 mM EDTA, 150 mM NaCl, and 0.5% Triton X-100. Re-ChIP was performed with anti-His tag, anti-PARP, or anti-CTCF antibodies and protein A or protein G4 Sepharose Fast Flow beads (Sigma). PCR amplification was performed using the purified DNA from either the single ChIP or the re-ChIP, primers corresponding to the selected DNA sites within the integrated H19 1CR, and the conditions described above for the single ChIP.

Real-time PCR analyses of the single-ChIP or re-ChIP samples were carried out in triplicate using 2 µl of the immunoprecipitated DNA sample or 2 µl of input DNA together with 200 nM primers diluted in a final volume of 25 µl of SensiFAST SYBR No-ROX (Quanta). PCR conditions were per the instructions of the solution manufacturer. Primers used were as follows: site 1, forward primer 5′-AGACCATGCTATTCTTTG-3′ and reverse primer 5′-GCTTCTTAGGTTGGCC-3′; site 2, forward primer 5′-GCGAGGACATGCTATT-3′ and reverse primer 5′-CTGAGAAGGTCAATGTC-3′; chromosome 7, forward primer 5′-AGAGCAACATTAAACAGGAGG-3′ and reverse primer 5′-GCCATACATTCAATCCTG-3′; chromosome 11, forward primer 5′-GTTCTAGGAACTTCGGACATC-3′ and reverse primer 5′-TCTGGCTCACTTGTTGACC-3′; and chromosome 21, forward primer 5′-GATCCTCTGAAGTGOAACATACAG-3′ and reverse primer 5′-CTGTCCTTCTTGTGTTACC-3′.

Amplification, data acquisition, and analysis were carried out using a Chromos 4 real-time PCR system (Bio-Rad). The percentage of DNA brought down by ChIP (percent input) was calculated as follows: percent input = AE(input) × Fd × 100, where AE is the amplification efficiency, CTChIP and CTInput are the threshold cycle values obtained from the exponential phase of quantitative PCR, and Fd is a dilution factor for the input DNA to balance the difference in amounts of ChIP samples.

Western blot analysis. Lysates from different cell lines were prepared and Western blot analysis was performed as described previously (26). In brief, cell lysates were resolved by SDS-PAGE, immunoblotted, and probed with anti-CTCF (Abcam, Cambridge, Cambs, United Kingdom), anti-His tag (Cell Signalling), anti-PAR-10H mouse monoclonal (Alexis), anti-PAR rabbit polyclonal (Calbiochem), antibiotin (Sigma), anti-PARP (Alexis), anti-PAR (Abcam, Cambridge, Cambs, United Kingdom), and anti-α-tubulin (Sigma) antibodies. The secondary antibodies used were horseradish peroxidase (HRP)-conjugated anti-rabbit and anti-mouse antibodies (both from Abcam) and anti-protein A–HRP (Sigma). Immunocomplexes were detected by using an ECL kit according to the instructions of the manufacturer (Amersham Biosciences, Little Chalfont, Bucks, United Kingdom).

Biotinylated-NAD labeling and selection of labeled proteins. To confirm the modification of CTCF by PARylation, nonradioactive labeling of cells with 6-biotin-17-NAD (biotinylated NAD+), which results in the incorporation of biotinyl-ADP-ribose into acceptor proteins, was used. MCF7 cells (7.5 × 10^4) were incubated with biotinylated NAD (Trevigen) as described previously (8). The cells were lysed in 500 µl IP buffer (see “IP” above), and labeled proteins were selected using Dynal M-280 streptavidin beads per the instructions of the manufacturer (Invitrogen).

ChIP-on-ChIP data comparison to investigate the colocalization of CTCF and PARP-1. A comparison of the microarray data obtained from two independent ChIP-ChIP experiments was made. The two chosen data sets both contained data from genome binding/occupancy profiling using tiling arrays featuring the ENCODE regions and therefore enabled a direct comparison of the data. The data were obtained from the Gene Expression Omnibus (GEO) database (http://www.ncbi.nlm.nih.gov/geo/). As input for CTCF ChIP localization analysis, we used the statistically processed data deposited under access number GSM52941 (ChIP-ChIP data for CTCF in HL60 cells) (13). For PARP-1 input, we used raw data supplied under accession numbers GSM239063 and GSM239061 (ChIP-ChIP data for PAR-P1 in MCF7 cells) (59) that have been statistically processed in a way similar to data for CTCF (13) to make them comparable. In brief, a two-tailed ranked Wilcoxon test was used to compare probes in each IP versus input conditions in a 1,000-bp sliding window with 250-bp steps. A >−10 logP value and a Hodges-Lehmann estimator were determined for each window. Chromosome coordinates corresponding to statistically significant P values for both CTCF and PARP-1 ChIP data were compared. The colocalization region was considered when CTCF and PARP-1 positions, expressed in chromosome coordinates, were within 500 bp of each other. The colocalization sites were mapped to the relevant human chromosomes to obtain a hot spot map.

Mass spectrometry. Samples were loaded onto a thin SDSE-PAGE gel, and the affinity-purified proteins were visualized by Coomassie blue staining. Mass spectrometry analysis was performed as described earlier (19). In brief, the band of interest was excised and subjected to in-gel tryptic digestion, and the tryptic peptides were subsequently extracted and analyzed using a liquid chromatography-tandem mass spectrometry ion trap (Esquire; Bruker Daltonics). Ion searching was performed using both in-house-built Mascot (http://www.matrixscience.com) and the X!Tandem tandem mass spectrometry search tool (http://thegpm.org) with SwissProt database release 57.8.

RESULTS

CTCF-130 and CTCF-180 are differentially poly(ADP-ribose)ylated. It has been generally assumed that CTCF is a protein with the apparent molecular mass of 130 kDa (CTCF-130), found predominately in the nuclear extracts of cells (26, 55). A further form of CTCF that migrated at 180 kDa (CTCF-180) and carried a PARylation mark was later reported (84, 92). We verified this modification in a PARG enzyme assay using MCF7 cell nuclear extracts enriched with CTCF-180 following treatment with sodium butyrate (15). In this experiment, a shift from CTCF-180 to CTCF-130 was observed with increasing doses of PARG, suggestive of the removal of the PARylation mark from CTCF-180 by PARG (Fig. 1A). We anticipated the presence of intermediate forms of CTCF following treatment with PARG, due to the nature of PARylation. However, only discrete forms of CTCF (CTCF-180 and CTCF-130) were detected.

Findings from studies using in vitro PARylation and IP with the mouse monoclonal 10H anti-PAR antibody had suggested that CTCF-180 was the PARylated form (92). This antibody recognizes only linear PAR chains of more than 20 PAR residues (47). We reasoned that if CTCF contains less than 20 residues of PAR, it is unlikely to be detected with this antibody. In the present investigation, we refined this analysis to further assess the two CTCF forms with regard to the level of
FIG. 1. CTCF-180 and CTCF-130 are differentially poly(ADP-ribosyl)ated. (A) Removal of PARylation from CTCF-180 using PARG. MCF7 cell nucleoplasmic extracts containing only CTCF-180 were incubated with PARG enzyme in a dose-dependent manner. The incubated lysates were analyzed by Western blotting using an anti-CTCF polyclonal antibody (Abcam). A complete shift from CTCF-180 to CTCF-130 is observed following increases in units of PARG. (B) The anti-PAR rabbit polyclonal antibody and mouse 10H antibody immunoprecipitate specific forms of CTCF. MCF7 cell extracts were immunoprecipitated with either the anti-PAR rabbit polyclonal antibody or anti-PAR mouse (10H) antibody. The immunoprecipitated proteins were analyzed by Western blotting with the anti-CTCF polyclonal antibody. (Left) The anti-PAR mouse 10H antibody, which recognizes more than 20 PAR residues, immunoprecipitates only CTCF-180. (Right) The anti-PAR rabbit polyclonal antibody is able to recognize as few as two PAR residues and is able to immunoprecipitate both CTCF-130 and CTCF-180. No AB, control with no antibody in the IP reaction mixture. (C) Labeling with biotinylated NAD\(^+\) reveals two forms, CTCF-180 and CTCF-130. MCF7 cells \((7.5 \times 10^5)\) were incubated with biotinylated NAD\(^+\), leading to the incorporation of biotinylated ADP-ribose residues attached to protein acceptors. The labeled proteins were selected using M-280 streptavidin magnetic beads. (Top) The input was analyzed for the presence of the CTCF forms by a Western blot assay. The membrane was probed with the anti-CTCF rabbit polyclonal antibody (Abcam), stripped, and probed with antibiotin antibody. Both CTCF-130 and CTCF-180 forms were present in the extracts (left), and bands of similar sizes (indicated by arrows) were observed when the membrane was probed with the antibiotin antibody (right). (Bottom) For selected biotinylated proteins, the elution (El) and flowthrough (F/T) fractions were analyzed by a Western blot assay for the presence of CTCF. (D) Analysis of the 130- and 180-kDa bands of the PARylated proteins selected using M-280 streptavidin magnetic beads and mass spectrometry. Proteins eluted from streptavidin magnetic beads were subjected to SDS-PAGE and stained with Coomassie blue. Bands corresponding to 130 and 180 kDa were excised and analyzed by mass spectrometry. The identified CTCF peptides are shown. (E) Schematic diagram of the CTCF N-terminal domain amino acid sequence showing the mutations produced by site-directed mutagenesis. The wild-type sequence is depicted at the top. Amino acid substitutions are highlighted. The amino acids are numbered according to Filippova et al. (32). (F) Western blot analysis of the proteins produced by the mutant variants of CTCF. 293T cells were transfected with 5 \(\mu\)g of EV or a vector expressing the CTCF WT or a CTCF mutant. Twenty-four hours posttransfection, cells were lysed and the cell extracts were divided in two and either subjected to IP with the anti-PAR rabbit polyclonal antibody or used as an input control. (Top) The immunoprecipitated proteins were resolved by SDS-PAGE, blotted, and probed with the anti-His tag antibody. The membrane was reprobed with an anti-CTCF antibody, confirming the presence of endogenous CTCF-130, retained by the anti-PAR antibody, in all cell lysates. (Bottom) The input materials were resolved by SDS-PAGE and probed with both the anti-His tag antibody and anti-CTCF, showing equal amounts of the exogenous and endogenous CTCF proteins.
PARylation by using the rabbit polyclonal anti-PAR antibody, which recognizes as few as two PAR residues and has higher affinities for longer chain lengths (1). We compared the two anti-PAR antibodies, monoclonal 10H and rabbit polyclonal, in IP experiments using MCF7 cell lysates and then performed Western analysis with the anti-CTCF polyclonal antibody. This experiment revealed that both CTCF-180 and CTCF-130 were present in the sample immunoprecipitated with the rabbit anti-PAR polyclonal antibody but that only CTCF-180 was present in the sample immunoprecipitated with the monoclonal 10H anti-PAR antibody (Fig. 1B).

To confirm the modification of CTCF by PARylation, MCF7 cells were labeled with biotinylated NAD\(^+\) and the lysates were subjected to Western blotting using either the anti-CTCF or the antibiotin antibody. With the use of the antibiotin antibody, two proteins were observed whose sizes coincided with CTCF-180 and CTCF-130 identified in the same lysates with the polyclonal anti-CTCF antibody (Fig. 1C, top). The labeled samples were then purified using M-280 streptavidin magnetic beads, which selected proteins containing biotinyl-ADP-ribose residues. As shown in Fig. 1C (bottom), two bands coincided with CTCF-180 and CTCF-130 identified in the same lysates with the polyclonal anti-CTCF antibody.

To confirm the presence of CTCF-130 and CTCF-180 among the proteins migrating as 130 and 180 kDa, the proteins labeled with biotinylated NAD\(^+\) and selected using streptavidin beads were resolved by SDS-PAGE and stained with Coomassie blue; bands corresponding to 130- and 180-kDa proteins were excised from the gel (Fig. 1D). The gel slices were subjected to in-gel digestion, and peptides were extracted and analyzed by mass spectrometry. Database searching with the recovered peptide sequences showed that sequences from both the 130- and 180-kDa bands matched human CTCF (SwissProt protein database accession number P49711); these peptide sequences are presented in Fig. 1D (right).

From the results of these experiments, we conclude that both forms of CTCF were PARylated, with CTCF-130 containing from 1 to several ADP-ribose residues and CTCF-180 containing more than 20 ADP-ribose residues.

Production of a series of CTCF mutants deficient in poly-(ADP-ribosyl)ation. Next, we aimed to generate a CTCF variant deficient in PARylation. It is generally accepted that there is no consensus site for PARylation (25); therefore, the following experimental observations were considered to determine the possible PARylated sequences in CTCF. First, we discovered that the N-terminal domain of CTCF was modified by PARylation in the in vitro PARylation assay (92). Second, by using a combination of alkaline hydrolysis of CTCF produced in baculovirus and mass spectrometry, a PAR-modified peptide was identified in the N-terminal domain (see Fig. S1A in the supplemental material). This peptide is highly conserved among species (see Fig. S1B in the supplemental material) and is also rich in glutamic acids, which we propose comprises possible PARylation sites (Fig. 1E).

This information was used to generate a series of CTCF mutants. The glutamic acid residues of the His-tagged CTCF expressed by the pCI vector were replaced by alanine residues by sequential site-directed substitutions until a mutant deficient in PARylation (CTCF Mut4) was obtained. The mutants were produced in the following order: CTCF Mut1, CTCF Mut2, CTCF Mut3 (an unexpected mutant identified following sequencing of Mut4), and CTCF Mut4 (Fig. 1E). Each mutant was characterized for the presence of the PARylation mark in transfection experiments with either the CTCF WT or mutant CTCF, followed by IP experiments using the polyclonal anti-PAR antibody (Fig. 1F). Western analysis using an anti-His tag antibody revealed that CTCF Mut1 and CTCF Mut2 could be immunoprecipitated by the anti-PAR antibody, whereas only trace amounts of CTCF Mut3 and no CTCF Mut4 variants were detected (Fig. 1F, top). Thus, CTCF Mut4 was the only exogenous CTCF protein that did not carry a PARylation mark. The membrane was then reprobed with an anti-CTCF antibody; this experiment confirmed that endogenous CTCF-130 proteins in all samples were equally and efficiently immunoprecipitated (Fig. 1F, top). The results also implied that a population of the endogenous CTCF-130 carried a PARylation mark, although this was likely to be short. Western analysis of the input material confirmed the presence of the exogenous CTCF in the samples (Fig. 1F, bottom). The generated mutants did not show any difference from the CTCF WT in their nuclear distribution/localization, stability, and levels of expression in the cell (Fig. 1F and data not shown).

We therefore concluded that CTCF Mut4 is a non-PARylated CTCF protein, which can be used in functional tests as a dominant negative mutant.

Inhibition of PARylation by PJ34 and PARylation deficiency of a CTCF mutant compromise CTCF functions as a transcriptional regulator and suppressor of cell proliferation. CTCF is a transcriptional factor that negatively regulates cell proliferation (33, 77, 83). To examine how inhibition of PARylation affects transcription, we employed transient reporter assays with three cell lines of different tissue origins, 293T, MDA435, and HeLa, using a luciferase gene driven by the p9arf promoter activated by the CTCF WT (33).

In the first series of experiments, cotransfections were carried out using either the CTCF WT or an EV and cells were treated with PJ34, a potent PARP inhibitor. The concentration of 10 \(\mu\)M PJ34 was found to be optimal for all three cell lines, with no effect on proliferation observed (see Fig. S2 in the supplemental material). However, at this concentration, PJ34 was found to effectively reduce general PARylation of proteins in these cell lines following oxidative stress caused by hydrogen peroxide (data not shown). As expected, activation of the p9arf promoter by the CTCF WT was observed in all cell lines; however, this activation was considerably reduced in cells treated with PJ34 (Fig. 2A). The luciferase reporter system was also used to examine the effects of the mutated non-PARylated version of CTCF (CTCF Mut4). CTCF Mut4 showed reduced ability to activate the p9arf promoter by the CTCF WT compared to the CTCF WT, although equivalent amounts of the protein were generated from the expression vectors in all three cell lines (Fig. 2B). Other CTCF mutants (Mut1, Mut2, and Mut3) were also used in this reporter assay, but no marked reduction in activation of the p9arf promoter was observed (data not shown).

We then asked whether the changes in CTCF Mut4 would compromise CTCF function as an inhibitor of cell proliferation. To investigate this issue, the EGFP-CTCF fusion proteins, both wild type and Mut4, were generated and utilized in a BrdU cell proliferation assay with three cell lines, 293T,
MDA435, and HeLa. Following transfection with an EV expressing EGFP (the control) or a vector expressing EGFP-CTCF WT or EGFP-CTCF Mut4, the percentage of EGFP-positive proliferating cells was assessed by fluorescence microscopy (Fig. 3A). In all cell lines, inhibition of proliferation by the CTCF WT, but not CTCF-Mut4, was observed. The loss of activation by the CTCF WT following PJ34 treatment is observed in all cell lines. Bars represent luciferase activity. Each bar shows an average of results from three experiments performed in triplicate. Error bars indicate standard deviations. The CTCF WT significantly activates the p19ARF promoter in the untreated cells, but following treatment with PJ34, the luciferase expression is greatly reduced. The loss of activation by the CTCF WT following PJ34 treatment is observed in all cell lines. Bars represent luciferase activity. Each bar shows an average of results from three experiments performed in triplicate. Error bars indicate standard deviations. The p19ARF promoter is activated by the CTCF WT, but reduced activation by CTCF Mut4 is observed. Bars represent luciferase activity. Each bar shows an average of results from three experiments performed in triplicate. Error bars indicate standard deviations. The levels of exogenously produced CTCF in the transfected cell lysates were assessed by Western blot analysis using the anti-His tag antibody. The membrane was stripped and reprobed with an antitubulin antibody used as a loading control. The results of the Western blot analysis are shown at the bottom.

FIG. 2. Activation of the p19ARF promoter by CTCF is poly(ADP-ribosyl)ation dependent. (A) Activation of the p19ARF promoter is abrogated by PJ34, a PARP inhibitor. 293T, MDA435, and HeLa cells were cotransfected with 2 μg of the luciferase reporter construct driven by the p19ARF promoter and either 0.5 μg of EV or 0.5 μg of a CTCF WT-expressing plasmid; 0.5 μg of the β-galactosidase vector was added for normalization. Transfected cells were either treated or not treated with PJ34 (10 μM), and luciferase expression was quantified by a luciferase assay. Luciferase values were normalized using a β-galactosidase assay. The CTCF WT significantly activates the p19ARF promoter in the untreated cells, but following treatment with PJ34, the luciferase expression is greatly reduced. The loss of activation by the CTCF WT following PJ34 treatment is observed in all cell lines. Bars represent luciferase activity. Each bar shows an average of results from three experiments performed in triplicate. Error bars indicate standard deviations. (B) Activation of the p19ARF promoter is abrogated by the CTCF mutant deficient in PARylation (CTCF Mut4). The 293T, MDA435, and HeLa cell lines were transfected with 0.5 μg of either an EV or a vector expressing the CTCF WT or CTCF Mut4, together with 2.0 μg of the p19ARF-luciferase reporter construct and 0.5 μg of a β-galactosidase expression vector. A luciferase assay of the transfected cells was performed, and values were normalized for transfection efficiency by using a β-galactosidase assay. The p19ARF promoter is activated by the CTCF WT, but reduced activation by CTCF Mut4 is observed. Bars represent luciferase activity. Each bar shows an average of results from three experiments performed in triplicate. Error bars indicate standard deviations. The levels of exogenously produced CTCF in the transfected cell lysates were assessed by Western blot analysis using the anti-His tag antibody. The membrane was stripped and reprobed with an antitubulin antibody used as a loading control. The results of the Western blot analysis are shown at the bottom.

Of note, the EGFP fusion constructs were functionally and biochemically equivalent to their nonfusion counterparts, as determined by several criteria: (i) similar patterns of regulation of the reporter construct (p19ARF-luciferase reporter gene) in a reporter assay, (ii) identical patterns of distribution in the nucleus, and (iii) similar amounts of the fusion proteins, as determined by Western blotting (Fig. 3B and data not shown).

From the results of these experiments, we concluded that deficiency in PARylation compromises CTCF functions as a transcriptional regulator and also as a suppressor of cell proliferation. Inhibition of PARylation by PJ34 and PARylation deficiency of a CTCF mutant compromise the CTCF insulator function in a cell-based epigenetic system and in an endogenous system. In our study, the insulator function of CTCF at the Igf2-H19 locus was examined using a novel isogenic insulator reporter.
cell model (a stable exogenous insulator system, or epigenetic system). The insulator reporter cell line was generated by stable transfection of 293T cells with a construct containing a luciferase gene regulated by the mouse *Igf2-H19* insulator sequence, an enhancer, and *Igf2* promoter 3 (yielding B4 cells) (Fig. 4A). The rationale for the generation of such a system was to create a quantifiable assay for insulator function that did not require colony counting. Investigation of the endogenous *Igf2-H19* locus is impeded by a number of factors: (i) *Igf2* is expressed at low levels in most normal cell lines; (ii) cancer cell
FIG. 4. CTCF insulator function at the H19 ICR in a transgenic model is dependent on PARylation. (A) In vitro epigenetic reporter system: a schematic diagram of the stably integrated luciferase reporter construct in the 293T transgenic cell line B4. The integrated construct contains a luciferase gene under the control of the mouse Igf2 promoter (P3), an H19 enhancer, and the upstream H19 ICR. The ICR contains four identified binding sites for the CTCF protein (indicated). (B) The activity of the luciferase gene controlled by the Igf2 promoter is increased in the presence of the PARP inhibitor PJ34. B4 cells (6 × 10^5) were plated and either treated or not treated with 10 μM PJ34. Twenty-four hours
lines often have rearrangements at the locus, and therefore, results obtained by assaying Igf2 expression after transfection with CTCF mutants are difficult to interpret; (iii) primary cell lines are not always available, and transfection of primary cell lines is difficult; (iv) many cell lines do not have the appropriate polymorphisms to detect the parent from which the expressed Igf2 allele originates; and (v) the endogenous Igf2 mRNA levels are influenced by RNA stability, multiple promoters, and posttranscriptional processing. We developed the stable exogenous insulator system which was devoid of these deficiencies and therefore used in our tests.

Prior to our experiments, this system was thoroughly assessed. Single-copy integration of the luciferase-neo construct was confirmed by PCR and Southern blotting (see Fig. S3A, B, and C in the supplemental material). Since CTCF binds only the maternal unmethylated Igf2-H19 ICR in order to exhibit its insulator function (46), the DNA methylation status of the integrated H19 ICR was inspected. Bisulfite sequence analysis revealed that a length of the DNA of the H19 ICR containing all four CTCF binding sites was unmethylated (see Fig. S4 in the supplemental material). ChIP analysis confirmed that CTCF is bound to the exogenous insulator sequence (Fig. 4). Taken together, these data demonstrate that the epigenetic system is an adequate model for the maternal allele of the Igf2-H19 locus.

 Luciferase assays were carried out following treatment of B4 cells with PJ34, and a comparison between treated and untreated cells was made. We assumed that the endogenous CTCF bound to the transgenic insulator sequence would function with both the CTCF WT and CTCF Mut4 at both binding sites and additionally in both treated and untreated cells in comparison with the controls (Fig. 4D; see also Fig. S5A in the supplemental material).

ChIP experiments were also carried out using B4 cells transfected with the vector expressing His-tagged CTCF WT or CTCF Mut4, and the anti-His tag antibody was used to select for only the bound exogenous CTCF. Results showed enrichment with both the CTCF WT and CTCF Mut4 at both binding sites (Fig. 4E; see also Fig. S5B in the supplemental material), thereby confirming that exogenous CTCF was bound to site 1 and site 2 at the H19 ICR, having replaced the endogenous protein.

The effects of PARYlation inhibition and the expression of CTCF Mut4 on the endogenous Igf2-H19 locus were then tested using A911M and A911P mouse-human hybrid cells, later, the cells were counted and lysed and the luciferase assay was performed. The luciferase values were normalized based upon cell number. Bars represent luciferase activity in relative luciferase units (RLU). Each bar shows an average of results from three experiments performed in triplicate. Error bars indicate standard deviations. The levels of CTCF in the cell lysates were assessed by Western blotting using the anti-CTCF antibody. The membrane was reprobed with an antitubulin antibody as an internal control for cell number (results are shown at the bottom). (C) The activity of the luciferase gene controlled by the Igf2 promoter is increased in cells transfected with the vector expressing CTCF Mut4, which is deficient in PARYlation. B4 cells (4 × 10^5) were transfected with 5 μg of either the EV or a vector expressing the CTCF WT or CTCF Mut 4 and 0.5 μg of the β-galactosidase-expressing construct. Forty-eight hours posttransfection, the cells were counted, lysed, and subjected to the luciferase assay. The luciferase values were initially normalized by cell number, and the transfection efficiency was then normalized using the β-galactosidase assay. Bars represent luciferase activity in relative luciferase units. Each bar shows an average of results from three experiments performed in triplicate. Error bars indicate standard deviations. The levels of exogenous CTCF were assessed by Western blotting using the anti-His tag antibody. The membrane was reprobed with an antitubulin antibody as an internal control for cell number (results are shown at the bottom). (D) Endogenous CTCF is associated with the CTCF binding sites 1 and 2 at the H19 ICR in B4 cells not treated and treated with PJ34. B4 cells (5 × 10^5) were either treated or not treated with 10 μM PJ34 for 24 h, cells were cross-linked with formaldehyde, and the standard ChIP assay was performed to assess the in vivo CTCF occupancies at the DNA. To separate two CTCF binding sites, site 1 and site 2 at the H19 ICR, having replaced the endogenous protein.

The effects of PARYlation inhibition and the expression of CTCF Mut4 on the endogenous Igf2-H19 locus were then tested using A911M and A911P mouse-human hybrid cells,
which contain a single copy of human chromosome 11 of known parental origin (maternal or paternal) (67). The imprinted status is maintained in both cell lines, with monoallelic expression of IGF2 observed in A911P cells (those with the paternal chromosome) but not in A911M cells (those with the maternal chromosome) (67). Both cell lines were either treated with PJ34 or transfected with the EV or a vector expressing the CTCF WT or CTCF Mut4, and the expression level of IGF2 was assessed by real-time PCR. Expression of IGF2 was detected in A911P cells under all conditions, whereas in A911M cells (those with the maternal chromosome), IGF2 expression was observed only following treatment with PJ34 or transfection with the CTCF Mut4 plasmid (see Fig. S5C in the supplemental material).

**CTCF and PARP-1 colocalize at the CTCF binding sites.** Our results support a model described in previous reports (52, 92) which proposes a mechanism of epigenetic control based on reversible PARylation of CTCF bound to DNA. In this model, the PARylated CTCF at the H19 ICR represents the default state and functions as a constitutive insulator protein; the de-PARylated CTCF remains bound to DNA, but its function as an insulator is compromised. However, although CTCF was reported previously to interact with PARP-1 (93), it had not been determined whether PARP-1 and CTCF were associated with the same DNA sequences. To experimentally test this proposition, we first asked whether PARP-1 (endogenous), CTCF (exogenous), and PARG (endogenous) are all associated at the exogenous insulator in B4 cells. We performed coimmunoprecipitation/Western analyses using B4 cells transfected with a plasmid expressing either the CTCF WT or CTCF Mut4 (non-PARylated). Anti-PARP-1 or anti-His tag antibodies were used for IP. As shown in Fig. 5A, endogenous PARP-1 is associated with both the CTCF WT and CTCF Mut4 (non-PARylated), but endogenous PARG was not detected in the immunoprecipitated complexes. No signal was detected in the control samples.

The colocalization of the endogenous CTCF and PARP-1 was observed in nuclei of different cell types by using ISPLA (Fig. 5B; see also Fig. S7B in the supplemental material) and immunofluorescence staining (Fig. 5C and data not shown). In the latter analysis, the Pearson coefficient was 0.697, indicating good correlation between the staining patterns of the two proteins. However, these results also indicate that there are populations of CTCF and PARP-1 which are not colocalized, suggesting that only a subset of CTCF binding sites have the potential to recruit CTCF–PARP-1 complexes.

We then asked if both PARP-1 and CTCF are present at site 1 and site 2 of the H19 ICR. We first performed single ChIP assays using either the anti-PARP-1, anti-His tag, or antitubulin (control) antibody with transfected B4 cells. Following the single ChIP with the PARP-1 antibody, both sites could be specifically amplified from all samples, confirming the presence of PARP-1 at both sites (Fig. 5D; see also Fig. S6, upper panels, in the supplemental material). Single ChIP was also performed using the anti-His tag antibody, thereby selecting only the DNA-protein complexes containing exogenously expressed CTCF (WT and Mut4). Signals for both CTCF binding sites were detected in cells transfected with a plasmid expressing the CTCF WT or CTCF Mut4 but not in cells transfected with the EV. The results of the single ChIP experiments confirmed that PARP-1 and the PARylated and non-PARylated forms of CTCF are associated with both CTCF binding sites.

In order to confirm the simultaneous presence of PARP and CTCF (WT and Mut4) at the two CTCF binding sites, a re-ChIP assay was performed by using combinations of anti-PARP-1, anti-His tag, and antitubulin (control) antibodies. As shown in Fig. 5D and Fig. S6 (lower panels) in the supplemental material, site 1 and site 2 could be specifically amplified from the samples subjected to re-ChIP. No signal was detected in the controls when the antitubulin antibody was used as the first antibody in the re-ChIP with either the anti-His tag or anti-PARP-1 antibody. These results confirmed the simultaneous association of PARP-1 with both the CTCF WT and CTCF Mut4 at the H19 ICR in B4 cells. Importantly, this association was dependent upon the PARylation status of CTCF.

We therefore suggest a model in which CTCF forms a functional complex with PARP-1 at CTCF binding sites, independently of the PARylation status of CTCF. Such association...
between CTCF and PARP-1 will then allow dynamic reversible modification of CTCF by PARP-1, thus maintaining the PARylated state of CTCF and modulating CTCF function (Fig. 5E). It is possible that a PAR degradation enzyme, such as PARG, may also be present in this complex, although we were unable to prove it experimentally.

We next asked if CTCF and PARP-1 colocalization may be a frequent occurrence across the genome by taking advantage of existing databases recording CTCF and PARP-1 distribution patterns in the human genome (13, 59). The sites of overlap (hot spots) for CTCF and PARP-1 are summarized in relation to specific chromosomes in Fig. 6A, and the corresponding chromosome positions are shown in Fig. 6B. In order to confirm that CTCF and PARP-1 occupied the identified sites of overlap, we randomly chose three hot spots on chromosomes 7, 11, and 21. Using the ChIP samples obtained from the B4 cell experiments (Fig. 5D), we observed enrichment for both CTCF and PARP-1 at all three sites tested, although to different extents (Fig. 6C).

Searches were then performed and positions of hot spot clusters were mapped with the Ensembl genome browser; these positions are shown as the relevant “chromosome region in detail” in Fig. 6D and Fig. 8A to N in the supplemental material. The sequence of hot spot 4 on chromosome 11, obtained from the CTCF motif reported previously (50), is presented in Fig. 6D (bottom) with a putative CTCF binding site highlighted. The genome locations for all five hot spots in a cluster on chromosome 11 are shown in Fig. 6E.

This analysis revealed that there is overlap in the localization patterns of CTCF and PARP-1 on many of the chromosomes, although no such overlap was found on chromosomes 1, 9, 13, 14, 15, 20, and X. The locations of the overlap sites were both intergenic and intragenic, as exemplified on chromosome 11 (Fig. 6D) and several other chromosomes (see Fig. S8 in the supplemental material). No information was available for CTCF localization on chromosomes 3, 17, and Y or for PARP-1 localization on chromosome Y.

The distribution of regions of CTCF/PARP-1 overlap varied among chromosomes, and interestingly, the sites of overlap appeared to be clustered on the chromosomes, suggesting that these features may be linked to specific functions of CTCF and PARP-1 at these sites.

**DISCUSSION**

In this study, we further characterized two forms of CTCF, CTCF-130 and CTCF-180, with regard to their PARylation status. CTCF-130 was previously suggested to be the non-PARylated form of CTCF (92) whose predominance in all cellular contexts has been well established (26, 33, 54, 55, 72, 77, 83). Further detailed analysis of CTCF-130 and CTCF-180 revealed the complex nature of CTCF PARylation. Both forms appear to be PARylated, although to different extents (Fig. 1B and C): CTCF-180 is likely to be highly PARylated, whereas CTCF-130 may contain only a few ADP-ribose residues. Two lines of evidence support this proposition: (i) CTCF-130 can be detected only by the rabbit polyclonal antibody (Fig. 1B), and (ii) PARylated CTCF-130 comigrates with the non-PARylated form of CTCF (Fig. 1F). These data demonstrate that the CTCF-130 band seen upon SDS-PAGE represents different populations of CTCF, which may contain mono(ADP-ribose), oligo(ADP-ribose), and unmodified CTCF; similar observations have been made in a recent report (91). Our findings suggest that the two PARylated forms of CTCF, CTCF-130 and CTCF-180, may be involved in functional regulation in different cell contexts (see below).

Intriguingly, PARylation of CTCF-130 with PARP-1 (92) and enzymatic degradation of CTCF-180 with PARG (Fig. 1A) result in the appearance of discrete bands of CTCF-180 and CTCF-130, respectively. The absence of a discernible CTCF protein smear, indicating varying modification levels, suggested that a distinct mechanism may contribute to the dramatic increase/decrease in the molecular mass of CTCF following PARylation and de-PARylation. This phenomenon requires further investigation.

The coexistence of differentially ADP-ribosylated forms of CTCF supports the current view of the PARylation reaction, which requires preformed mono(ADP-ribosyl)ated substrates. Several enzymes, such as mono(ADP-ribosyl)transferase, PARP-1 itself, other PARP family members, and heteromers of these proteins, may mono(ADP-ribosyl)ate CTCF, which can then be used for the PARylation reaction (43, 73).

Although CTCF-180 usually represents only a minor fraction of the total CTCF population, our recent study shows that CTCF-180 is the predominant isoform in many normal human tissues and the only isoform in normal breast tissues, with CTCF-130 associated with cell proliferation, whereas CTCF-180 is associated with nondividing cells (27). In this context, the low levels of endogenous CTCF-180 in the cell lines, as well as the absence/very low levels of CTCF-180 produced from the exogenous CTCF-expressing constructs (reference 40 and this study), may be explained by rapid degradation of PAR, typical for proliferating cells in culture, where the half-life of PAR is less than a minute (5, 90).

To specifically examine the role of PARylation in CTCF functions, we generated a CTCF mutant deficient in PARylation by mutating the cluster of glutamic acid residues, the preferred acceptors of PARylation, in the N-terminal domain of CTCF (Fig. 1E). To our knowledge, this is the first example of the generation of a mutant protein completely deficient in PARylation in vivo; a recently described PARylation-deficient p53 mutant still demonstrated residual in vivo PARylation (45). The non-PARylated CTCF Mut4 was identical to the CTCF WT by the following criteria: (i) protein size of 130 kDa upon SDS-PAGE, protein stability, and similar protein levels produced from the plasmid vectors (Fig. 1F, 2B, and 4C and data not shown), (ii) localization and distribution patterns in the nucleus (Fig. 3B), and (iii) ability to bind to DNA (Fig. 4E and 5D). These findings are in contrast with the reports that PARylation alters the stability and DNA binding activity of proteins (81, 89).

In the present study, in all experimental systems tested, the differences in CTCF PARylation resulted in the loss of CTCF function. In the presence of PJ34, a potent PARP inhibitor, or transfection with the plasmid expressing CTCF Mut4, the activity of the p9ARF promoter-driven reporter was considerably reduced, suggesting that CTCF function as a transcriptional regulator is PARylation dependent. Notably, the effects of PARylation in cell lines of different origins seem to vary,
A

<table>
<thead>
<tr>
<th>Chromosome number</th>
<th>No of Hotspots</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>7</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>9</td>
</tr>
<tr>
<td>6</td>
<td>16</td>
</tr>
<tr>
<td>7</td>
<td>67</td>
</tr>
<tr>
<td>8</td>
<td>6</td>
</tr>
<tr>
<td>10</td>
<td>4</td>
</tr>
<tr>
<td>11</td>
<td>13</td>
</tr>
<tr>
<td>12</td>
<td>17</td>
</tr>
<tr>
<td>16</td>
<td>2</td>
</tr>
<tr>
<td>18</td>
<td>16</td>
</tr>
<tr>
<td>19</td>
<td>5</td>
</tr>
<tr>
<td>21</td>
<td>17</td>
</tr>
<tr>
<td>22</td>
<td>2</td>
</tr>
</tbody>
</table>

B

FIG. 6—Continued on following page
thus indicating that cell context may play a “tuning” role in CTCF regulation.

The inhibitory role of CTCF in cell proliferation was also compromised by CTCF Mut4 (Fig. 3). It is conceivable that this regulation occurs through transcripitional regulation of CTCF target genes controlling proliferation, such as p19ARF (32, 33, 62), c-myc (62), pRb (24), BRCA1 (17), and possibly p21 and p27 (76). Recent studies also implicated CTCF in the regulation of ribosomal biogenesis (84) and replication timing (12). Thus, since CTCF effects on cell proliferation are likely to involve various regulatory pathways, alterations in CTCF PARylation may lead to global changes in cell biology.

In our previous investigations, we used similar assays to assess the effects of CTCF phosphorylation by protein kinase CK2 on CTCF functions, which generally resemble those of CTCF PARylation (29, 53). Cross talk between the two pathways may occur and may explain the similar functional outcomes; these observations require further examination. Functional regulation by both modifications, PARylation and phosphorylation, is not unusual and has been reported previously for p53 and histones (30, 41, 70).

Finally, the insulator function of CTCF, tested using a stable isogenic insulator reporter cell system and a mouse-human hybrid cell line, was also perturbed by PJ34 and CTCF Mut4 (Fig. 4). Using this system, we explored a hypothesis that CTCF functional activities regulated by PARylation are facilitated mechanistically by the close proximity of CTCF and PARP (Fig. 5E). In this model, PARylation of CTCF is established by PARPs present at the same DNA site and can be reversed by PARG. We were not able to establish the presence
FIG. 6. Colocalization of CTCF and PARP-1 across the genome. (A) Summary of the microarray data obtained from two independent ChIP-ChIP experiments studying CTCF and PARP-1, deposited in the GEO database. Numerous colocalization sites within human DNA were identified. The table indicates the number of identified CTCF and PARP-1 colocalization hot spots with regard to the specific human chromosomes. (B) Correlated sites (hot spot clusters) of CTCF and PARP-1 colocalization mapped to human chromosomes. The numbers of hot spots on the chromosomes are shown. (C) Validation of the identified CTCF and PARP-1 colocalization sites using three randomly chosen sites positioned on chromosomes 7, 11, and 21. Real-time PCR was performed using a selection of the ChIP samples obtained in the B4 cell experiments (Fig. 5D). The single ChIP samples were analyzed with anti-His tag, anti-PARP-1, and antitubulin antibodies, and re-ChIP samples were analyzed with anti-His tag/anti-PARP-1, anti-PARP-1/antitubulin, and antitubulin/anti-PARP-1 antibodies. The efficiency of ChIP was calculated as a percentage of the starting material (percent input). At all three chromosomal sites, enrichment is observed, indicating colocalization of CTCF and PARP-1 at these chromosomal positions. (D) Chromosome 11 is shown as an example of a “chromosome region in detail.” The Ensembl genome browser was used to map the sites of CTCF/PARP-1 colocalization. (Top) The chromosomal position of the hot spot cluster comprising five hot spots is shown. (Middle) The region in detail shows genes within the area of CTCF/PARP-1 colocalization. (Bottom) One of the identified colocalization hot spots, hot spot 4, is expanded to reveal the DNA sequence, and a putative CTCF binding site is highlighted. ncRNA, noncoding RNA. (E) Chromosomal positions of the five identified hot spots in a cluster on chromosome 11.
of PARG at the insulator or association with PARP-1 and CTCF due to the absence of ChIP-grade anti-PARG antibodies; however, interaction between PARP-1 and PARG has been reported in the literature (36, 48).

In the proposed model, a feedback axis can be envisaged whereby CTCF can directly regulate PARP-1, which in turn may lead to DNA hypomethylation (40). Cooperation between CTCF and PARP-1 may provide a link between PARylation and DNA methylation, thereby adding another layer of complexity to the model. These aspects deserve further investigation.

The notions that CTCF is stably bound to its DNA targets and that regulatory signals are responsible for the modulation of CTCF activity are supported by previously published observations (63, 92). However, such close association between CTCF and PARP-1 may not be required for CTCF regulation by PARylation at all CTCF targets (91).

The genome-scale association between CTCF and the poly(ADP-ribose)ylation mark was shown previously with the mouse microarray library of CTCF binding sites (71, 92). In the present investigation, considerable overlap between CTCF and PARP-1 in the nuclei in different cell lines was confirmed (Fig. 5B and C; see also Fig. 5B in the supplemental material). Furthermore, bioinformatics analysis of the distributions of CTCF (13) and PARP-1 (59) in the ENCODE regions of the genome in human cell lines revealed sizable areas of overlap between CTCF and PARP-1 on different chromosomes (Fig. 6D and E; see also Fig. S8 in the supplemental material), with three of the identified CTCF/PARP-1 colocalization sites validated (Fig. 6C). Although the ChIP-ChIP data for CTCF and PARP-1 were obtained from different cell lines, previous reports indicate that CTCF occupancy at its binding sites is conserved among different cell lines (9, 13, 50).

In this context, it is important to acknowledge the substantial overlap between CTCF and cohesin binding sites throughout the genome, which is believed to have important functional implication in the regulation of gene expression and is the subject of much current research (references 38, 39, and 74 and references therein). The results of our study, indicating the frequency of colocalization of CTCF and PARP-1, lead us to hypothesize that CTCF, PARP-1, and cohesin may all appear in one complex. This hypothesis deserves further investigation.

A common theme of this study is that CTCF and PARP-1 are important elements of the same regulatory pathway(s) and likely to be linked by a feedback mechanism. Therefore, PARP-1 dysfunction may lead to loss of CTCF PARylation and, as a result, dramatic changes in gene regulation and development of disease, in particular cancer. Alternatively, deregulation of PARP-1 by aberrant CTCF protein may also contribute to tumorigenesis. In this context, it is significant that both CTCF and PARP-1 display features of tumor suppressors (2, 31, 54, 68, 78). However, the interrelationship between CTCF and the PARylation enzymatic machinery appears to be more complex, as CTCF and PARylation enzymes are likely to be shared by different functional complexes. Such sharing may be important for functional integration of different cellular processes in response to various stimuli.

ACKNOWLEDGMENTS

We thank F. Doquier, A. Harrison, R. Meyer, M. Meyer-Ficca, and N. Curtin for helpful discussions. We are grateful to V. Lobanenkov for the gift of the anti-CTCF monoclonal antibody. L. Kraus for the gift of the anti-PARP-1 antibody, and M. Oshima and A. Feinberg for A911M and A911P mouse-human hybrid cells. Funding for this work was from the Medical Research Council (to D.F., S.R., and E.K.), the Breast Cancer Campaign (to S.R. and E.K.), the University of Essex (to E.K. and I.C.), the Swedish Science Research Council (to R.O.), the Swedish Cancer Research Foundation (to R.O.), the Swedish Pediatric Cancer Foundation (to R.O.), the Lundberg Foundation (to R.O.), HEROIC and ChILL (European Union integrated projects; to R.O.), a Cancer Research United Kingdom senior cancer research fellowship (to A.M.), the Wenner-Gren Foundation (to M.J.), and an Association of International Cancer Research grant (to A.M. and Y.I.). We acknowledge the support of the University of Cambridge, Cancer Research UK, and Hutchison Whampoa Limited (to A.M., M.J., and Y.I.).

REFERENCES


Klenova, E. M., H. C. Morse III, R. Ohlsson, and V. V. Lobanenkov. 2002. The novel BORIS + CTCF gene family is uniquely involved in the epige-


