A Conserved Insulator That Recruits CTCF and Cohesin Exists between the Closely Related but Divergently Regulated Interleukin-3 and Granulocyte-Macrophage Colony-Stimulating Factor Genes[∇]

Sarion R. Bowers,¹ Fabio Mirabella,¹§ Fernando J. Calero-Nieto,¹†§ Stephanie Valeaux,¹ Suzana Hadjur,² Euan W. Baxter,¹‡ Matthias Merkenschlager,² and Peter N. Cockerill^{1*}

Experimental Haematology, Leeds Institute of Molecular Medicine, University of Leeds, St. James's University Hospital, Leeds LS9 7TF, United Kingdom,¹ and Lymphocyte Development Group, MRC Clinical Sciences Centre, Imperial College, Du Cane Road, London W12 0NN, United Kingdom²

Received 8 September 2008/Returned for modification 5 November 2008/Accepted 9 January 2009

The human interleukin-3 (IL-3) and granulocyte-macrophage colony-stimulating-factor (GM-CSF, or CSF2) gene cluster arose by duplication of an ancestral gene. Although just 10 kb apart and responsive to the same signals, the IL-3 and GM-CSF genes are nevertheless regulated independently by separate, tissue-specific enhancers. To understand the differential regulation of the IL-3 and GM-CSF genes we have investigated a cluster of three ubiquitous DNase I-hypersensitive sites (DHSs) located between the two genes. We found that each site contains a conserved CTCF consensus sequence, binds CTCF, and recruits the cohesin subunit Rad21 in vivo. The positioning of these sites relative to the IL-3 and GM-CSF genes and their respective enhancers is conserved between human and mouse, suggesting a functional role in the organization of the locus. We found that these sites effectively block functional interactions between the GM-CSF enhancer and either the IL-3 or the GM-CSF promoter in reporter gene assays. These data argue that the regulation of the IL-3 and the GM-CSF promoters depends on the positions of their enhancers relative to the conserved CTCF/cohesin-binding sites. We suggest that one important role of these sites is to enable the independent regulation of the IL-3 and GM-CSF genes.

The closely related interleukin-3 (IL-3) and granulocyte-macrophage colony-stimulating factor (GM-CSF, or CSF2) genes arose from a gene duplication event and are activated via the same signaling pathways but are differentially regulated (9, 28, 29). The expression of these cytokines needs to be very tightly controlled, as inappropriate expression can contribute to leukemia and inflammation. Both genes are highly inducible, being activated by T-cell receptor signaling pathways in T cells and similar pathways in mast cells (3, 4, 9, 10, 16, 18, 28, 29). However, the GM-CSF gene is also expressed by many additional cell types, such as macrophages, endothelial cells, and fibroblasts, that do not express IL-3 (3, 9, 10, 28, 29). In the human IL-3/GM-CSF locus, each gene is associated with clusters of conserved upstream sequences which include inducible DNase I-hypersensitive sites (DHSs) that function as inducible enhancers specifically in the appropriate cell types where each gene is expressed (3, 9, 10, 16). Furthermore, the cluster of DHSs upstream of the IL-3 gene is only present in IL-3-ex-

pressing cells or their precursors, whereas the DHSs just upstream of the GM-CSF gene exist within a much broader range of GM-CSF-expressing cells (9, 10, 16). It is likely, therefore, that these two genes are independently regulated. This concept is further supported by the fact that a 10-kb segment of DNA encompassing the human GM-CSF gene plus the GM-CSF enhancer is correctly regulated as a transgene in the absence of any additional elements from the IL-3 locus (10). What remains unclear, however, is whether any mechanisms exist to prevent activation of the IL-3 promoter by the GM-CSF enhancer.

In this study we investigated the functions of a poorly understood cluster of three ubiquitous DHSs known to exist immediately downstream of the human IL-3 gene and 5 kb upstream of the GM-CSF enhancer (10) (see Fig. 1A, below). Unlike the inducible DHSs in the locus, our previous studies of these ubiquitous DHSs failed to find a role for these elements as transcriptional enhancers (P. N. Cockerill and A. Hawwari, unpublished data). These ubiquitous DHSs also effectively represent a dividing line that separates the IL-3/GM-CSF locus into two independent functional units controlled by independent enhancers. Significantly, in this study we found that both the transcription factor CTCF and a member of the cohesin complex, Rad21, were associated with each one of the three ubiquitous DHSs. CTCF is often associated with transcriptional insulator activities (2, 40, 42) and was also recently found to recruit Rad21 at known insulators and other CTCF-binding sites in the genome (13, 33, 34, 36, 38, 41). We found that this cluster of CTCF-binding elements downstream of the IL-3

* Corresponding author. Mailing address: Experimental Haematology, Leeds Institute of Molecular Medicine, Wellcome Trust Brenner Building, St. James's University Hospital, Leeds LS9 7TF, United Kingdom. Phone: 44 113 3438639. Fax: 44 113 3438502. E-mail: p.n.cockerill@leeds.ac.uk.

† Present address: Department of Haematology, Cambridge Institute for Medical Research, Wellcome Trust/MRC Building, Cambridge, United Kingdom.

‡ Present address: YCR p53 Research Unit, Department of Biology, University of York, York YO10 5DD, United Kingdom.

§ F.M. and F.J.C.-N. contributed equally to this study.

[∇] Published ahead of print on 21 January 2009.

gene functions as an efficient enhancer/blocker in transfection assays of GM-CSF enhancer function. This suggests that the insulator may help to coordinate correct communication between enhancers and promoters within the IL-3/GM-CSF locus.

MATERIALS AND METHODS

Cell culture. The Jurkat human T-cell line was cultured in RPMI (Invitrogen) containing 10% fetal calf serum, 4 g/liter D-glucose, 1 mM sodium pyruvate, 1× minimal essential medium with essential amino acids (Invitrogen), 1× minimal essential medium with nonessential amino acids (Invitrogen), 100 U/ml penicillin, and 100 µg/ml streptomycin. T lymphoblasts and mast cells were cultured from line C42 human IL-3/GM-CSF transgenic mice as described previously (4). Other transgenic cell types were prepared as follows: fibroblasts were prepared by culturing embryonic fibroblasts, macrophages were prepared by culturing adherent bone marrow cells in IL-3, stem cell factor, and M-CSF, and myeloid progenitor cells were prepared by culturing nonadherent fetal liver cells in IL-3 plus stem cell factor.

Nuclease-hypersensitive site analysis. Nuclease digestions of nuclei isolated from Jurkat human T cells and mapping of hypersensitive sites downstream of the human IL-3 locus were performed essentially as described previously (10), using 2 µg/ml (4 U/ml) DNase I (Worthington) or 50 U/ml micrococcal nuclease (Worthington). DNase I digestions of permeabilized T cells cultured from human IL-3/GM-CSF transgenic mice were performed essentially as described previously (4). DHSs downstream of the mouse IL-3 gene were mapped using PCR-amplified mouse DNA probes as follows. The kb +2.5 and +3.7 sites were mapped within BamHI-digested DNA using a probe generated from the primers GGACAGATACGTTATCAAGTCC and TGACCATGGGCCATGAGGAAC. The kb +10.1 and +12.8 sites were mapped within EcoRI-digested DNA using a probe generated from the primers CTCCATGGTTAGGAATGTGC and CTGAAGACTGGTCTGAGTCTG. Briefly, all nuclease digestions of permeabilized cells were performed in nuclei digestion buffer containing 60 mM KCl, 15 mM NaCl, 5 mM MgCl₂, 10 mM Tris, 0.3 M sucrose, pH 7.4. Cells were initially suspended at a concentration of 3×10^7 to 4×10^7 cells/ml in digestion buffer, to which was added an equal volume of nuclease in digestion buffer plus 0.4% Nonidet P-40 and 2 mM CaCl₂. Cells were digested with a final concentration of 5 to 15 µg/ml (30 to 100 U/ml) DNase I for 3 min at 22°C before stopping the reactions, purifying the DNA, and mapping hypersensitive sites as described previously (8). Digested DNA samples were run on a 0.8% agarose gel and the DNA transferred to a Hybond N membrane (Amersham). Mapping of nuclease cleavage sites was performed by indirect end labeling, as described previously (8).

In vivo footprinting analyses. Footprinting analyses were performed essentially as described previously (37). C42 transgenic mouse T lymphoblasts were either stimulated for 4 h with 20 ng/µl phorbol 12-myristate 13-acetate (PMA) and 1 µM calcium ionophore A23187, or left unstimulated, before treatment with 0.2% dimethyl sulfate (DMS) in phosphate-buffered saline (PBS) for 5 min at room temperature. The reaction was stopped by addition of 40 volumes of ice-cold PBS, and cells were washed twice with 25 volumes of ice-cold PBS. Cells were resuspended and lysed in lysis buffer containing 10 mM Tris · HCl, pH 8.0, 85 mM NaCl, 20 mM EDTA, 0.5% sodium dodecyl sulfate. Proteinase K was added to give a concentration of 0.5 µg/ml, before incubation at room temperature overnight, followed by purification of the DNA. Methylated guanine residues were cleaved by treatment with 0.1 M piperidine for 10 min at 90°C, and the DNA was butanol extracted and ethanol precipitated. A cleaved methylated guanine DNA control was prepared by *in vitro* DMS treatment using a modification (37) of the method of Maxim and Gilbert (27). DNase I digestions were performed as above using permeabilized unstimulated C42 T lymphoblasts digested with 12.5 µg/ml DNase I.

Specific DNA cleavage sites were detected using ligation-mediated PCR (LM-PCR) as described previously (37), essentially by the method of Herskovitz and Riggs (17) with modifications described by Kontaraki et al. (20), and using a linker duplex consisting of the 25-base LP25 nucleotide, GCGGTGACCCGGGAGATCTGAATTC, annealed to the 21-base LP21 nucleotide, GAATTCAGATCTCCCGGGTCA. LM-PCR was performed with 0.5 to 1.0 µg of template DNA, using a biotinylated primer to extend one round of DNA synthesis to the cleavage sites before ligating the LP25-LP21 linker duplex to the extension products, binding the biotinylated products to magnetic beads, performing a nested PCR for 22 cycles with a region-specific primer and LP25 primer, performing seven cycles of another nested PCR with a ³²P-end-labeled primer, and analyzing the products on a 6% polyacrylamide gel under DNA-denaturing conditions. Betaine (1.25 M) was included in the PCR mixtures to increase

oligonucleotide primer stringency and specificity. The DNA was visualized using a phosphorimager.

LM-PCR analyses were performed with three sets of three nested primers specific for the three DHSs, plus the linker primer. The three primers within each set are designated B for the biotinylated primer used at the first stage, P for the primer used at the second PCR stage, and L for those used in the final labeling stage, with F indicating a 5' primer and R indicating a 3' primer. The oligonucleotide primers in each set of primers have the following sequences (the annealing temperatures used for the primers are shown in parentheses): for the kb +2.9 region, D2.9BF, AAATCCACAGAAGTCTAC (58°C), D2.9PF, AGCCGTTCTGAGCATTTTACCTAT (60°C), and D2.9LF, CTATCTAATCTATTCCTACCCCAACCAA (64°C); for the kb +4.2 region, D4.2BR, GATGAAAA CAGGACTAGCT (58°C), D4.2PR, GAGCTGGGTGTGCAGGAGACTT (60°C), and D4.2LR, TGGGTGTGCAGGAGACTTGGGAAAC (64°C); for the kb +4.9 region, D4.9BR, GCTCAACAATAGTTT (58°C), D4.9PR, ATAAAATTCCTGGCTGAGAGCTGT (60°C), and D4.9LR, GAGAGCTGT AATCCACCTTCCCCCTGA (64°C).

Electrophoretic mobility shift assays (EMSAs). Nuclear extracts were prepared from unstimulated Jurkat cells, as described previously (11). Oligonucleotides were labeled with [α -³²P]dCTP. Four µg of nuclear extract was incubated with 0.2 ng of labeled probe, and 4 µg of poly(dI-dC) in a 16-µl reaction mixture containing 10% glycerol, 20 mM HEPES, 30 mM KCl, 30 mM NaCl, 0.1 mM ZnCl₂, 3 mM MgCl₂, 1% dithiothreitol, 0.1 mM phenylmethylsulfonyl fluoride, 5 µg/ml aprotinin, 5 µg/ml leupeptin for 20 min at room temperature (~22°C) before analysis by polyacrylamide gel electrophoresis. Polyacrylamide gels (4%) containing 25 mM Tris borate, 0.5 mM EDTA were pre-electrophoresed for 1 h at room temperature at 10 V/cm and run at 13 V/cm for 1.5 h at room temperature. Gels were dried and the DNA visualized on a phosphorimager screen. For supershift assays 0.25 µl of CTCF antibody (07-729; Upstate Biotechnology) was incubated with the nuclear extract for 5 minutes on ice before addition to the above reaction mixture. Competition assays were carried out by incubation of 4 µg of nuclear extract with 100 ng of unlabeled oligonucleotide duplex for 5 min, at room temperature, before the addition of 0.2 ng of labeled probe and further incubation for 15 min at room temperature.

ChIP assays. Chromatin immunoprecipitation (ChIP) assays, presented in Fig. 4 below, were performed essentially as previously described (1, 31), with the following minor modifications. Cells from transgenic T lymphoblasts were fixed in 0.33 M (1%) formaldehyde for 10 min, before the addition of 4 volumes of ice-cold PBS containing 0.125 M glycine, to give an approximately twofold molar excess of glycine over formaldehyde. DNA was purified after proteinase K digestion using Qiagen PCR purification columns. Chromatin was immunoprecipitated with 1 µg of antibody raised against human Rad21 (Abcam) or 6 µl of rabbit antiserum raised against human CTCF (07-729; Upstate Biotechnology) per 5×10^6 cells, with normal rabbit immunoglobulin G (IgG; Upstate Biotechnology) used as a control. The amount of precipitated DNA was quantified using real-time quantitative PCR, with an ABI 7700 or 7900HT sequence detection system (Perkin-Elmer Life Sciences). All values were calculated by using a standard curve of serially diluted sonicated transgenic mouse DNA. Data were normalized by expressing the amount of specific DNA as a ratio, with DNA from an inactive nonconserved region of mouse chromosome 1 (UCSC genome browser, July 2007 build 37 chromosome 1 coordinates 159,762,928 to 159,763,258). The sequences of the primers were as follows: kb -10.4 IL-3, CACACAGCACTCCCGTGATC and CCTGGAGCATAAGTGCCCAAGAG; kb -4.1 IL-3, CCACAAAGGAGGTTTCCCCTAA and AGACCTGGTCCCC TGTAAGATG; kb -0.2 IL-3, TATGGAGGTTCCATGCATAGATAAAGAT and AACAGCCTCCCGCTTATATG; kb +2.9 IL-3, TGAAGTTCGGCCTG GTTGAG and GGGAGACATGCCACAGGATTA; kb +4.2 IL-3, GGGATG TCTTGTGTTTGGACTTAGC and CCTGAGACCACTGGCACACA; kb +4.9 IL-3, TCCCACTGTGATGGCATTAG and GGGCCTTCTCTGGTGG TTCT; kb +6.7 IL-3, AGACCCCAAGCTCAACTCTACCT and CCTCACAC CAGAAGGATTCCAT; m chromosome 1 control sequence, TGCTCCACAG TGTCATGTACA and AGCAATTTTCATGGGTGAGAGAAG.

ChIP assays on small interfering RNA (siRNA)-treated 293T cells expressing FLAG-tagged Rad21 were performed as previously described (37). Samples were amplified by real-time PCR with the above primers. Cells were transfected with CTCF siRNA, Rad21 siRNA, or control siRNA. We confirmed that there was efficient knockdown of Rad21 or CTCF expression by performing a Western blot assay of transfected cells (37).

Mapping of CG methylation. Five µg of DNA extracted from each of a number of cell types isolated from C42 human IL-3/GM-CSF transgenic mice was digested with BamHI and HhaI, a methylation-sensitive enzyme. DNA was resolved on a 0.8% agarose gel, transferred to a Hybond N nitrocellulose membrane (Amersham), and hybridized with a ³²P-labeled 4.6-kb BamHI-digested

fragment spanning the region 3' of the IL-3 gene and encompassing the three DHSs.

Plasmid construction. The previously described constructs pGM and pGM-GME, containing the -627 to +28 region of the human GM-CSF promoter (GM627) in the absence or presence of a 717-bp BglII fragment of the human GM-CSF enhancer, respectively (10), were inserted into the firefly luciferase reporter gene plasmid pXPG (3).

pGMC was created by inserting the sequence TCGATCCTGTGGCATGTC TCCCCTAGTGGTCTTCCAGAACTGCAGCCGCCGCC, encompassing the IL-3 kb +2.9 CTCF site into the XhoI site of pGM. pGMCEC was created from pGMC by inserting the sequence GATCTCCTGTGGCATGTC TCCCCTAGTGGTCTTCCAGAACTGCAGCCGCCGCC, encompassing the kb +2.9 CTCF site into the BamHI site of pGMC and then inserting the GM-CSF enhancer into the BglII site as in pGM-GME.

The plasmid pGM628 was created from a PCR-generated fragment and includes the native bp -628 to +33 region fused directly to the HindIII and NcoI sites of pXPG, so as to superimpose the ATG translation initiation sites of the GM-CSF promoter and luciferase gene. This was performed by replacing the GGATGT sequence at the GM-CSF ATG element with CCATGG to create an NcoI site. pGM4.3 contains the natural kb -4.3 to +33 region of the GM-CSF promoter and was created by inserting the 3.7-kb HindIII fragment from kb -4.3 to -0.6 into the HindIII site of pGM628. pKS was constructed by inserting a 2.6-kb KpnI/StuI fragment (KS) encompassing the IL-3 kb +2.9, +4.2, and +4.9 CTCF sites (plus some linker sequences) upstream of the promoter in pGM628. pEKS was constructed by inserting the 717-bp BglII fragment of the GM-CSF enhancer upstream of the KS fragment in pKS. pGM4.3C and pGM4.3C3 were created by inserting either one copy or three direct repeats of the sequence TCATTCTGTGGCATGTCTCCCCTAGTGGTCTTCCAGAACTGCAGCCGCCGCCG encompassing the kb +2.9 CTCF site into the kb -1.6 BpI site of the GM-CSF promoter in pGM4.3.

The plasmid pIL3, previously known as pXPG-IL3H, contains the -559 to +50 region of the human IL-3 promoter (16). pIL3-GME contains the 717-bp BglII fragment of the GM-CSF enhancer inserted upstream of the IL-3 promoter. pIL3-C3-GME contains the array of three CTCF sites from pGM4.3C3 inserted between the GM-CSF enhancer and the IL-3 promoter.

All DNA fragments and sequences were inserted into the plasmids in the same orientation as they exist in the GM-CSF/IL-3 locus.

Transient transfections and luciferase assays. Cells were transfected by electroporation of 4.5×10^6 Jurkat cells with plasmid DNA purified by two rounds of CsCl gradient centrifugation. In some experiments (see Fig. 7A, 8, and 9C, below), the firefly luciferase plasmids were linearized with AatII prior to transfection. Cells were transfected with either (i) 5 μ g of pIL3-based plasmid plus 1 μ g of the thymidine kinase promoter control *Renilla* luciferase plasmid, pRL-TK (Promega), or (ii) 5 μ g of pGM or pGM628 reporter plasmids, or an equimolar amount of larger derivatives of these plasmids, plus 1 μ g of a control plasmid (pXRL-GME) containing the GM-CSF enhancer and promoter that was created by replacing the Luc⁺ gene of pGM-GME with the *Renilla* luciferase reporter gene derived from pRL-TK. Subsequent to transfection, cells were cultured for 40 h before stimulation with 20 ng/ml PMA and 1 μ M calcium ionophore A23187 for 8 h. Cells were washed in PBS and assayed for luciferase reporter gene activities using the Promega dual luciferase assay kit and a Berthold Mithras LB-940 microplate luminometer.

RESULTS

The IL-3 and GM-CSF genes are separated by a conserved cluster of ubiquitous DHSs. Located just downstream of the human IL-3 gene is a cluster of three ubiquitous constitutive DHSs (Fig. 1A) (10). In the human locus these DHSs are located 2.9, 4.2, and 4.9 kb downstream of the IL-3 transcription start site and are hypersensitive to both DNase I and micrococcal nuclease in Jurkat human T cells (Fig. 1B). To determine whether equivalent DHSs exist in the mouse locus, we mapped DHSs between the IL-3 and GM-CSF genes in mouse thymus and cultured splenic T cells. Two such DHSs were found 2.5 kb and 3.7 kb downstream of the mouse IL-3 transcription start site (Fig. 1C) in positions analogous to the human IL-3 kb +2.9 and +4.3 DHSs, respectively. However, a large block of repetitive DNA is present immediately down-

stream of these two DHSs in the mouse but not the human locus. Consequently, we probed further downstream of this repetitive region and located two additional DHSs (Fig. 1D) in mouse T cells. One DHS located at kb +10.1 appeared to be the counterpart of the human IL-3 kb +4.9 DHS, and the other DHS at kb +12.8 represented a site that has been previously observed upstream of the GM-CSF gene in mouse EL-4 T cells and defined at that time as the kb -4.9 GM-CSF DHS (32).

The three ubiquitous DHSs each encompass conserved binding sites for the insulator factor CTCF. A pairwise comparison of the human and mouse genome sequences spanning the three DHSs using ecbrowser at the NCBI dcode website failed to find any significant homology or conservation of any transcription factor-binding sites (<http://www.dcode.org>). However, a broader parallel 28-species analysis of conserved regions, using the UCSC Genome Browser (<http://genome.ucsc.edu/>), revealed some very limited conservation of discrete segments of DNA within the three human DHSs (Fig. 1E) and the equivalent three mouse DHSs. The degree of conservation within these DHSs was much less extensive than that seen at the GM-CSF and IL-3 promoters and enhancers (Fig. 1E) and was restricted to very short stretches of sequence (Fig. 2A). In comparison to other more complex regulatory elements, these segments would not normally be ranked as significant conserved noncoding sequences. However, closer inspection revealed that each conserved region closely resembled the CTCF consensus binding sequence (Fig. 2A) (19, 44) and in each case comprised two short stretches of adjacent conserved sequence that corresponded very closely to the predicted binding sites for the two separate clusters of zinc finger DNA-binding domains of CTCF (5, 35). The short DNA sequences shown in Fig. 2A are highly conserved among a wide range of mammalian species. Curiously, the CTCF consensus sequences located at the kb +2.9 and +4.2 DHSs share complete identity across the predicted binding sites both for Zn finger domains 8 to 11 at the 5' end and for Zn finger domains 4 to 7 at the 3' end (5). In Fig. 2A we present an expanded CTCF consensus sequence based on our own and previously published analyses, so as to include two distinct sites within DNA that are contacted by CTCF. This is our attempt to build upon the recently published consensus sequences that represent just the predicted binding region for Zn fingers 3 to 7 (19, 44).

The three ubiquitous DHSs each recruit CTCF. As a model system for studying the IL-3/GM-CSF locus, we employed the C42 line of transgenic mice carrying a 130-kb AgeI fragment of DNA that encompasses the entire human locus, including its regulatory elements (4). This AgeI segment extends far upstream and downstream of the IL-3 and GM-CSF genes and is sufficient to support correct levels of expression of both IL-3 and GM-CSF in transgenic mice (F. Mirabella and P. N. Cockcroft, unpublished data). In vivo footprinting was performed on cultured splenic T cells incubated with dimethyl sulfate and on permeabilized T cells digested briefly with DNase I. LM-PCR was performed across the conserved regions within each of the kb +2.9, +4.2, and +4.9 DHSs to reveal sites bound by nuclear factors in vivo. For each DHS, the CTCF elements were the only regions showing any evidence of factor binding. Each of the three regions revealed a biphasic pattern of two closely spaced protected segments flanked by sites of enhanced cleav-

FIG. 1. Mapping of three constitutive DNase I-hypersensitive sites downstream of the IL-3 gene. (A) Schematic maps of the IL-3/GM-CSF locus in humans and mice. Vertical arrows indicate the positions of the DNase I-hypersensitive sites. Dotted lines show sites equivalent between humans and mice. Fragments generated by restriction enzyme digestion are indicated by horizontal lines, and the probes subsequently used to detect these fragments are depicted by black boxes with arrows. (B) Mapping of DNase I- and micrococcal nuclease (MNase)-hypersensitive sites at kb +2.9, +4.2, and +4.9 within a BamHI fragment downstream of the IL-3 gene in human Jurkat T cells, as described previously (10). (C) Mapping of the kb +2.5 and +3.7 DHSs within a BamHI fragment downstream of the IL-3 gene in mouse thymus and T cells. (D) Mapping of the kb +10.1 and +12.8 DHSs within an EcoRI fragment downstream of the IL-3 gene in mouse T cells. (E) Sequence conservation between vertebrate species of the key regulatory elements of the IL-3/GM-CSF locus and the three DHSs downstream of the IL-3 gene. Regulatory elements and DHSs are marked with vertical arrows.

age in the sequences spanning the conserved CTCF sites (Fig. 2B to D). As summarized in Fig. 2E, the predicted Zn finger-binding DNA sequences exhibited protection from DNase I and DMS, and the DNA sequences predicted to bind Zn fingers 3 to 7 were flanked by regions hypersensitive to DNase I and DMS. Based on this information we predicted that CTCF was making contacts with each of the three consensus sequences via both of the clusters of Zn fingers that are believed to mediate DNA binding by CTCF.

EMSA of Jurkat cell nuclear extracts were used to demonstrate that each of the three CTCF consensus sequences had the capacity to specifically associate with CTCF-like complexes that comigrated with CTCF complexes bound to the chicken β -globin gene insulator CTCF site (HS4) (2) (Fig. 3A). Efficient binding was dependent upon inclusion of ZnCl_2 in the binding reaction mixture (data not shown). Formation of each CTCF-like complex was specifically inhibited by inclusion of the β -globin CTCF site as a competitor, but not by an oligonucleotide in which critical CTCF-binding residues were mu-

tated (HS4 Δ) (Fig. 3A). Inclusion of an antibody raised against CTCF reduced the intensity of the specific complex, and there was some evidence for the appearance of supershifted complexes in the presence of antibody (Fig. 3B). The complex was unaffected by the inclusion of normal IgG as a control (Fig. 3B). The kb +2.9 and +4.2 probes, which share considerable sequence identity, were equally efficient at forming CTCF complexes, to about the same extent as the globin CTCF site. In contrast, the kb +4.9 CTCF site was considerably less efficient at binding CTCF, and this reflected the fact that this site is a poorer match to the CTCF consensus sequence (Fig. 2A), produces weaker *in vivo* footprints (Fig. 2D), and forms a weaker DNase I-hypersensitive site than the kb +2.9 and +4.2 CTCF sites (Fig. 1B).

ChIP assays of transgenic T cells were employed to further confirm that CTCF was binding in a constitutive manner to each of the three DHSs downstream of the IL-3 gene but not to any of the other regions assayed within the IL-3 locus (Fig. 4A), in either the presence or absence of stimulation with

FIG. 2. DMS and DNase I in vivo footprinting of CTCF sites downstream of the IL-3 gene. (A) Sequence homology and conservation of the CTCF-binding consensus sequences located within the three DHSs 3' of the IL-3 gene and comparison with previously defined CTCF-binding sites and consensus sequences. Shown above the list of binding sites are two consensus sequences that we derived from known sites. On the right is a compilation of published consensus sequences (19, 44) that represent the predicted binding sites for Zn finger domains 3 to 7. On the left is a consensus sequence predicted to bind Zn finger domains 8 to 11 that we and others (5) derived by comparing known binding sites and the three IL-3 gene CTCF sites. The two lines of the consensus sequence represent alternate bases allowed at each position, with highly conserved bases shown in uppercase letters and less-well-conserved bases shown in lowercase letters. Predicted binding sites for CTCF Zn finger domains are indicated by asterisks below, and individual Zn finger-binding domains are numbered as described in reference 5. Within the IL-3 gene CTCF sites, underlined regions indicate strong conservation between vertebrate species, bold lettering indicates regions protected from DNase I digestion, and hyphens indicate variable spacing between the two separate CTCF-binding segments. (B to D) LM-PCR of DMS and DNase I footprints within the human IL-3 kb +2.9 DHS (B), kb +4.2 DHS (C), and kb +4.9 DHS (D) in transgenic mouse T cells that were either unstimulated (-) or stimulated for 4 h with 20 ng/ml PMA and 2 μ M calcium ionophore A23187 (+). DMS treatment was performed on two independent cultures of T cells. Black circles indicate hypermethylated guanines and open circles show hypomethylated guanines, compared to two separate DMS-treated naked DNA controls (G1 and G2). Numbers on the left indicate nucleotide positions relative to the IL-3 gene transcription start site. A DNase I-digested DNA control is shown in lane C. (E) Summary of footprinting data obtained in panels B to D. Nucleotide bases involved in CTCF binding, as determined by DMS treatment, are indicated by open and closed circles, indicating hypo- or hypermethylation, respectively. Protected nucleotide bases, as determined by DNase I treatment, are shown in bold and bases hypersensitive to DNase I are underlined. Predicted CTCF contacts (5) are indicated by asterisks.

FIG. 3. EMSA results for CTCF binding to each of the three 3' IL-3 gene DHSs. (A and B) EMSAs of nuclear extracts from Jurkat T cells in either (A) the presence and absence (nil) of an excess of unlabeled competitor oligonucleotide duplex containing a CTCF site from the chicken β -globin insulator (HS4) or the same site with key CTCF-binding residues mutated (Δ HS4), or (B) the presence and absence of either an antibody against CTCF or control IgG. DNA probes used in each EMSA are indicated above the images. (C) Sequences of CTCF EMSA probes used in panels A and B. Mutated residues in Δ HS4 are shown in bold lowercase letters, and the approximate positions of predicted binding sites are indicated by asterisks.

PMA and calcium ionophore (data not shown). This is in agreement with a global microarray CTCF ChIP analysis that also revealed CTCF binding across this region in human fibroblasts (19).

CTCF sites in the IL-3 locus are not DNA methylated. In some instances, such as at the H19 imprinting control region, the recruitment of CTCF can be suppressed by DNA methylation at the binding site (15). We observed that the kb +4.9 CTCF-binding site encompasses a CG sequence within the region that is protected from DNase I cleavage in T cells (Fig. 2E). This CG also exists within a recognition site for the methylation-sensitive restriction enzyme HhaI, so we were able to determine whether this site was methylated *in vivo* in a wide range of tissues from transgenic mice. We performed Southern blot hybridization analysis on a 4.6-kb BamHI fragment of genomic DNA that was also digested with HhaI (Fig. 5). The presence of strong 1.5-kb and 3.1-kb bands, and the near-complete absence of any predicted band of a size where the CTCF HhaI site is uncut, indicated that the kb +4.9 CTCF site exists predominantly in an unmethylated state in all tissues. The presence of a weak 2.9-kb band indicated that the HhaI site located 0.2 kb upstream of the CTCF site is also unmethylated in a small proportion of cells, whereas the HhaI site at the far 5' end appeared to be fully methylated. These data suggest that the constitutive binding of CTCF to the kb +4.9 site may protect this site from DNA methylation. The kb +2.9 and +3.9 CTCF sites are unlikely to be influenced by DNA methylation because they lack CG elements within the consensus binding sequences.

CTCF mediates recruitment of cohesin protein Rad21. During mitosis, chromosomal cohesion is mediated by the cohesin complex, of which Rad21 is an essential component (25, 38).

Cohesin has additional functions in interphase (25, 38), and it was recently observed that CTCF can recruit cohesin (33, 36, 41). By performing ChIP on the human IL-3 gene in transgenic T cells, we also found that each of the CTCF-binding sites was associated with Rad21 (Fig. 4B), and the levels of Rad21 remained unchanged under both unstimulated and stimulated conditions (data not shown). In an additional ChIP analysis we employed human 293T cells carrying a FLAG-tagged Rad21 expression vector (33) and confirmed that the three CTCF sites downstream of the IL-3 gene were recruiting both CTCF and Rad21, while the IL-3 promoter and a flanking sequence did not (Fig. 6). Furthermore, by using siRNA to diminish levels of Rad21 or CTCF in these cells, we demonstrated that CTCF was required for the efficient recruitment of Rad21 to each of the three IL-3 gene CTCF sites and to a control CTCF-binding region within the UPK2 promoter (h11:118333) (33) (Fig. 6A). In contrast, Rad21 was not required for recruitment of CTCF to any site (Fig. 6B). In these studies we also confirmed that the siRNA oligonucleotides employed efficiently knocked down expression of the target genes at the level of protein expression (33).

The cluster of CTCF sites functions as an insulator in enhancer-blocking assays. CTCF has come to prominence largely because many CTCF-binding sites are associated with transcriptional insulators that have the ability to block the action of distal enhancers (2). For this reason we investigated whether the correct pattern of specific communication between the IL-3 and GM-CSF promoters and enhancers might also be controlled in part by a CTCF-dependent insulator located between the two genes. To this end we performed transfection assays with luciferase reporter gene plasmids containing the human GM-CSF promoter linked to the GM-CSF

FIG. 4. ChIP assays of CTCF and Rad21. (A and B) ChIP analyses of C42 transgenic mouse T cells performed with either control IgG or antibodies against CTCF (A) or Rad21 (B). Results are plotted as relative enrichment, calculated as the ratio of immunoprecipitated DNA over input DNA, and were normalized by dividing by the values obtained using an amplicon from a nontranscribed nonconserved region of mouse chromosome 1. Results shown are the averages of three independent experiments with the standard deviations. (C) Map of the IL-3 locus, with DHSs indicated by vertical arrows above and the locations of the primers used for each amplicon, and their positions relative to the transcription start site, indicated below.

enhancer and/or the intact cluster of CTCF-binding DHSs. We initially chose the GM-CSF gene upstream sequences to test insulator function because the GM-CSF enhancer is normally located 3 kb upstream of the GM-CSF promoter. This made it possible to simply replace an ~2-kb segment of natural intervening sequence with the intact cluster of CTCF-binding DHSs, thus maintaining the enhancer in a position about 3 kb upstream, so as to avoid the complications of distance effects.

In the plasmids summarized in Fig. 7A, the potential insu-

FIG. 5. DNA methylation analysis of the kb +4.9 DHS. DNA purified from a range of cell types from C42 transgenic mice was digested with BamHI and the methylation-sensitive enzyme HhaI and analyzed by Southern blot hybridization analysis using the 4.6-kb BamHI fragment at the 3' end of the IL-3 locus as a probe. The upper 4.6-kb band is the intact BamHI fragment, undigested by HhaI. The 3.1-, 2.9-, and 1.5-kb bands are subfragments generated by HhaI digestion at the kb +4.7 or +4.9 HhaI sites. A schematic of the region is shown below, with restriction sites for BamHI (B) and HhaI (H) indicated. The sequence of the kb +4.9 CTCF-binding site is also shown, with the HhaI site underlined. No cleavage of the 5' HhaI site was detected.

lator was assayed as a 2.6-kb KpnI-StuI fragment of the human locus, referred to here as the KS element, which is located 221 bp downstream of the IL-3 gene polyadenylation signal. These constructs contained either (i) the kb -628 to +33 segment of the human GM-CSF promoter (pGM628), (ii) the kb -4.3 to +33 segment of the GM-CSF locus in which the GM-CSF enhancer exists 3 kb upstream of the promoter (pGM4.3), (iii) the KS element positioned between the GM-CSF enhancer and promoter, in place of 2 kb of the natural intervening sequence normally present between the -2.6 BglII site and the -0.6 HindIII site (pEKS), or (iv) just the KS element inserted into the kb -0.6 HindIII site upstream of the promoter (pKS). Because these circular constructs contain just one potential insulator, they were first linearized with AatII, which cuts once within the vector backbone. Plasmids were transfected into Jurkat cells, left for 40 h to allow the plasmid DNA to recruit nuclear proteins, and then stimulated for 8 h with PMA and calcium ionophore to activate the promoter and enhancer (Fig. 7). In its natural context the enhancer increased promoter activity by 850% (Fig. 7A, pGM4.3). In the presence of the intervening KS element containing the three CTCF-binding sites, the enhancer only increased promoter activity by 130%, which amounts to a sixfold decrease in activity (Fig. 7A, pEKS). This confirmed that the cluster of CTCF-binding

FIG. 6. Requirements for colocalization of CTCF and Rad21. CTCF and Rad21 ChIP assays were performed in 293T or in FLAG-Rad21-transfected 293T cells, as previously described (33). (A) FLAG-Rad21 ChIP assay in the presence of CTCF siRNA or control siRNA. (B) CTCF ChIP in the presence of Rad21 siRNA or control siRNA. Primers are as indicated in Fig. 4. The control region is a previously identified CTCF/Rad21 site (Chr. 11:118,332,119) (33) located 130 bp upstream of the human uroplakin 2 (UPK2) gene, as shown in Fig. 2A. Values are calculated as the ratio of immunoprecipitated DNA over input DNA and were normalized by dividing by the values obtained using the UPK2 CTCF site as a control. Error bars represent the standard deviations of two independent experiments.

DHSs can indeed function as an insulator to block the actions of the GM-CSF enhancer on the promoter. In the absence of the enhancer, the KS element did not significantly influence GM-CSF promoter activity (Fig. 7A, pKS), demonstrating that it is not merely a repressor of transcription. The KS element also substantially reduced the activity of the GM-CSF enhancer when these constructs were assayed as circular plasmids (data not shown).

In order to obtain evidence that the enhancer-blocking activity of this element was due to CTCF, we generated an additional series of plasmids to test the effect of just the kb +2.9 CTCF site alone on the GM-CSF enhancer and promoter (Fig. 7B and 8). In the plasmids pGM4.3C and pGM4.3C3, either a single copy or a triplet of a 58-bp sequence encompassing the kb +2.9 CTCF site was inserted into pGM4.3 at a BspI site located 1.6 kb upstream of the transcription start site (Fig. 7B). The DNA was transfected into Jurkat T cells and assayed as described above, except that we employed uncut

FIG. 7. Enhancer-blocking assay of DHSs downstream of the IL-3 gene. Jurkat T cells were transfected with pGM628 firefly luciferase reporter gene plasmids containing different combinations of the GM-CSF promoter (P) and enhancer (GM-E) and either a 2.6-kb KS fragment encompassing the three DHSs 3' of the IL-3 gene (A) or one or three copies of a 58-bp oligonucleotide encoding the kb +2.9 CTCF-binding site (indicated by C and CCC in the map below) (B). After transfection, cells were incubated for 40 h before stimulation for 8 h with 20 ng/ml PMA and 1 μ M calcium ionophore A23187. Activity is expressed relative to the activity of a control cotransfected plasmid containing the GM-CSF enhancer and promoter attached to the *Renilla* luciferase reporter gene. Each point shown is the average of three assays of each of two independent plasmid preparations. Plasmids used in panel A were linearized with AatII prior to transfection, whereas circular plasmids were used in panel B.

circular plasmids. Upon insertion of a single CTCF site, the increase in promoter activity observed in the presence of the enhancer was reduced from 345% to 130%, a 2.5-fold reduction in activity. Insertion of three CTCF sites resulted in a further reduction, with the enhancer increasing promoter ac-

FIG. 8. Enhancer-blocking assay of single CTCF-binding sites. Jurkat T cells were transfected with linearized pGM-based plasmids as in Fig. 7. Plasmids contained the GM-CSF promoter in the presence and absence of the GM-CSF enhancer (GME) and in the presence and absence of one or two copies of the kb +2.9 DHS CTCF site (indicated by C).

tivity by just 80%, a fourfold reduction in enhancer activity. This suggests that CTCF can indeed account for the activity of the KS element.

To further assay the function of the kb +2.9 CTCF site in another context, we tested the properties of single CTCF sites directly flanking the GM-CSF enhancer. We generated an additional collection of linearized plasmids based on pGM (10), which contains the bp -627 to +28 region of the GM-CSF promoter (Fig. 8). When just the GM-CSF enhancer was inserted directly upstream of the promoter (pGM-GME) it supported a 530% increase in promoter activity. When the enhancer was flanked on either side by single copies of the kb +2.9 CTCF site (pGMCEC), its ability to enhance the activity of the promoter was reduced to just 150%. When a single CTCF site was inserted either upstream of the promoter (pGMC) or upstream of the enhancer (data not shown), there was essentially no change in activity. Overall, this body of data indicates that the function of the CTCF sites is position dependent and conforms to the definition of an enhancer-blocking insulator.

In its natural context, one function of the IL-3/GM-CSF insulator may be to block inappropriate activation of the IL-3 promoter by the GM-CSF enhancer. As a preliminary test of this hypothesis, we determined whether the insulator had any capacity to block communication between the GM-CSF enhancer and IL-3 promoter. For this purpose we placed the 717-bp GM-CSF enhancer sequence directly upstream of a -559 to +50 segment of the human IL-3 promoter in the plasmid pIL3 (Fig. 9A). We found that the enhancer mediated a four- to sixfold increase in promoter activity in either circular or linear plasmids (Fig. 9B and C). To test the CTCF-dependent activities of the IL-3/GM-CSF insulator, we employed the same approach used in Fig. 7B. We inserted three copies of the oligonucleotide encompassing the kb +2.9 IL-3 CTCF site in between the GM-CSF enhancer and IL-3 promoter. Signifi-

FIG. 9. Insulator assay of the IL-3 promoter. Jurkat T cells were transfected with circular (B) or linearized (C) pIL3-based plasmids as in Fig. 7. (A) Plasmids contained the -559 to +50 region of the human IL-3 promoter in the presence and absence of the GM-CSF enhancer (GME) in the presence and absence of three copies of the kb +2.9 DHS CTCF site (indicated by CCC) inserted downstream of the enhancer. (B and C) Luciferase activities, expressed relative to the activity of the cotransfected thymidine kinase promoter plasmid pRL-TK. Values represent the means of six independent transfections, employing three of each of two independent clones of each construct (which gave equivalent values) with the standard deviations.

cantly, we found that this short segment of DNA was capable of completely suppressing GM-CSF enhancer activity in linearized plasmids (Fig. 9C) and reduced enhancer activity by fivefold in circular plasmids (Fig. 9B). This residual enhancer activity observed in circular plasmids may represent activation of the IL-3 promoter activation from a position several kilobases downstream of the IL-3 promoter, a situation potentially analogous to what may occur in the endogenous IL-3/GM-CSF locus in the absence of the insulator. However, formal testing of the *in vivo* functions of the insulator will require further studies involving genetic manipulation of the mouse genome.

DISCUSSION

The IL-3/GM-CSF locus exists as two independently regulated units. There are many examples of genes in the genome that have undergone duplication to create clusters of genes with related activities and expression patterns. In some cases, such as in the β -globin locus (14) and the IL-4/IL-13 locus (23), genes have evolved with locus control regions (LCRs) that

regulate clusters of genes. At the β -globin locus the genes are expressed in a temporal sequence during development and the LCR appears to only activate a single gene at a time (43). However, in the case of the IL-3/GM-CSF locus, the genes can be either coexpressed or differentially regulated. This may have created a need for a different pattern of regulation, and it is apparent that the IL-3/GM-CSF locus does not have just a single dominant LCR. This problem appears to have been solved via the parallel evolution of separate enhancers with distinct activities that can in principle control each gene independently. The GM-CSF gene has an inducible enhancer that is active in a very broad range of cells that express NFAT and AP-1 (3, 10), which represents the majority of GM-CSF-expressing cell types. The IL-3 gene has a much narrower expression pattern, as it only expressed in a restricted subset of hematopoietic cell types that includes T cells and mast cells. Furthermore, the IL-3 gene is controlled by its kb -4.5 enhancer, which is active in T cells and mast cells (16) but not other myeloid lineages or in endothelial cells, epithelial cells, or fibroblasts, all of which express GM-CSF.

Our finding that an insulator exists between the IL-3 and GM-CSF genes provides an additional mechanism that may aid in the differential regulation of these two genes. Our observations may also help to account for the fact that some myeloid cells fail to express IL-3, even though they appear to express the necessary complement of transcription factors for promoter function. For example, the primitive myeloid cell line KG1a does not express IL-3 (3) and yet it expresses factors such as NFAT, AP-1, GATA, and Runx family proteins, which are required for IL-3 promoter activity (3, 4), and it has a DHS at kb -4.1 (16) that is associated with early activation events in the IL-3 locus (9). Although the kb -4.5 IL-3 enhancer is completely inactive in KG1a cells, the GM-CSF enhancer is highly active (3), and if allowed to act in an unobstructed manner, one might expect the GM-CSF enhancer to be able to act in cooperation with the above-mentioned transcription factors to activate the IL-3 gene. The GM-CSF enhancer is certainly capable of activating the IL-3 promoter in transfection assays (Fig. 9) (11). However, the presence of an insulator between the two genes may ensure that activation of the IL-3 gene does not occur via the GM-CSF enhancer in the absence of an active IL-3 enhancer. Thus, it makes good biological sense to have an insulator between the IL-3 and GM-CSF genes to ensure that these genes are expressed in the appropriate cell types.

CTCF can direct IL-3/GM-CSF insulator function. A striking feature of the IL-3 gene insulator is that it is comprised of three DHSs with very limited overall sequence conservation. Sequence homology is restricted entirely to the bases thought to be required for CTCF binding. Furthermore, the CTCF sites are the only elements that display any evidence of factor binding in *in vivo* footprinting analyses. Therefore, we propose that CTCF is likely to be the only DNA-binding factor required for the enhancer-blocking function of the IL-3 gene insulator. The high degree of sequence conservation observed for each of the three CTCF consensus sequences makes it likely that this insulator played an important role in mammalian evolution.

CTCF is a protein that can utilize different combinations of up to 11 Zn fingers to contact DNA, and at times this has made

it difficult to predict CTCF-binding sites from DNA sequences (5). Recent studies have defined better consensus sequences for CTCF (19, 33, 44), but these are invariably restricted to the predicted contacts for Zn fingers 3 to 7 (5). Because CTCF can also contact DNA and create distinct footprints via Zn fingers 8 to 11, we attempted to create an expanded consensus sequence. We found that each of the IL-3 gene CTCF sites was composed of two distinct highly conserved regions which each directed *in vivo* footprints. These regions were easily aligned with similar regions in other well-defined CTCF sites (5), and so we endeavored to construct a second loose consensus sequence that roughly matches the predicted binding sites for Zn fingers 9 to 11 (Fig. 2A). This is centered on a well-conserved TG sequence, and further computational analyses by others will be required to validate the accuracy of this second consensus.

The IL-3 gene insulator is now just one of many enhancer-blocking insulators known to be dependent on CTCF. One of the best-described examples is the imprinted H19/Igf2 locus, where CTCF binding and insulator function is controlled by DNA methylation, and changes in DNA methylation patterns result in altered patterns of enhancer-promoter communication and gene expression (15). In contrast, the IL-3 gene insulator has three CTCF sites that either lack DNA methylation sites or are essentially fully demethylated in all cells. This implies that the IL-3 and GM-CSF genes are likely to be insulated from each other at all times and in all cells.

Mechanisms of insulator action. Insulators can function in various ways, and there appear to be quite distinct properties that allow insulators to function either to block enhancer activity or as barrier elements preventing the spread of heterochromatin (21, 22, 33, 40, 42). Mammalian enhancer-blocking insulators are almost invariably associated with CTCF. Although the precise mechanism of action of enhancer-blockers is not fully understood, it is known that CTCF can potentially establish distinct topological domains within chromatin by tethering insulators to specific sites associated with nucleoli (45). Hence, insulators may function by either segregating genes within defined chromatin domains or by acting as decoys to divert DNA elements away from sites of transcription (21, 40, 42).

The significant finding that CTCF can recruit cohesin adds an exciting new dimension to molecular mechanisms of insulator function (13, 33, 34, 36, 38, 41). Cohesin exists as a component of a protein ring that is large enough to encircle two strands of chromatin and thereby physically delineate the boundaries of looped DNA domains in the genome (25). There is also evidence that insulators function in pairs (6, 21, 22, 26), whereby (i) genes flanked by insulators are shielded from outside enhancers and (ii) enhancer elements can bypass genes flanked by a pair of insulators to activate a more distant gene. Cohesin may represent the glue that mediates interactions between pairs of insulators.

It was also recently observed that the binding of CTCF and cohesin can in some cases be tightly regulated as part of an inducible gene regulatory circuit. CTCF and cohesin both bind to an insulator in the chicken lysozyme locus, where they function to block the actions of an upstream enhancer. The activation of this gene involves the induction of a non-

coding RNA transcript that displaces CTCF and cohesin, by repositioning a nucleosome over the CTCF site, thereby allowing the upstream enhancer to engage the lysozyme promoter (24).

In contrast to their role in enhancer interference, CTCF sites do not appear to function as boundaries of active chromatin domains. For example, CTCF can remain bound within the introns of actively transcribed genes, meaning that insulators do not simply create rigid independently regulated chromatin domains (33).

CTCF appears to have many diverse functions, and there is evidence that CTCF-binding elements do not universally function simply as insulators. CTCF was first identified as a factor associated with the *myc* promoter, where it is thought to function as a classical repressor by recruiting histone deacetylases (30). Conversely, the amyloid precursor protein promoter requires CTCF for efficient activation (39). Even the CTCF sites in the archetypal H19 ICR are now known to be required for H19 expression (12), and efficient H19 expression also requires cohesin (41). Furthermore, CTCF sites are also themselves able to either recruit or accumulate RNA polymerase (7), and our own preliminary studies indicate that the IL-3/GM-CSF insulator elements also recruit active RNA polymerase (Mirabella and Cockerill, unpublished). Hence, insulators may best be viewed as regulatory elements within the genome that function to maintain correct communication between promoters and enhancers while preventing inappropriate activation by disconnected regulatory elements.

ACKNOWLEDGMENTS

We thank P. Lefevre for advice on ChIP assays and providing antibodies, M. Hoogenkamp and C. Bonifer for assistance setting up in vivo footprinting assays and cell culture systems, F. Ponchel for advice on real-time PCRs, E. Klenova for discussions regarding insulators, and D. Brooke, A. Horner, and L. Williams for creating and managing the transgenic mice.

This work was supported by the Biotechnology and Biological Sciences Research Council, the Association for International Cancer Research, and Yorkshire Cancer Research.

REFERENCES

- Baxter, E. W., W. J. Cummings, and R. E. K. Fournier. 2005. Formation of a large, complex domain of histone hyperacetylation at human 14q32.1 requires the serpin locus control region. *Nucleic Acids Res.* **33**:3313–3322.
- Bell, A. C., A. G. West, and G. Felsenfeld. 1999. The protein CTCF is required for the enhancer blocking activity of vertebrate insulators. *Cell* **98**:387–396.
- Bert, A. G., J. Burrows, A. Hawwari, M. A. Vadas, and P. N. Cockerill. 2000. Reconstitution of T cell-specific transcription directed by composite NFAT/Oct elements. *J. Immunol.* **165**:5646–5655.
- Bert, A. G., B. V. Johnson, E. W. Baxter, and P. N. Cockerill. 2007. A modular enhancer is differentially regulated by GATA and NFAT elements that direct different tissue-specific patterns of nucleosome positioning and inducible chromatin remodeling. *Mol. Cell. Biol.* **27**:2870–2885.
- Burcin, M., R. Arnold, M. Lutz, B. Kaiser, D. Runge, F. Lottspeich, G. N. Filippova, V. V. Lobanenko, and R. Renkawitz. 1997. Negative protein 1, which is required for function of the chicken lysozyme gene silencer in conjunction with hormone receptors, is identical to the multivalent zinc finger repressor CTCF. *Mol. Cell. Biol.* **17**:1281–1288.
- Cai, H. N., and P. Shen. 2001. Effects of cis arrangement of chromatin insulators on enhancer-blocking activity. *Science* **291**:493–495.
- Chernukhin, I., S. Shamsuddin, S. Y. Kang, R. Bergstrom, Y. W. Kwon, W. Yu, J. Whitehead, R. Mukhopadhyay, F. Docquier, D. Farrar, I. Morrison, M. Vigneron, S. Y. Wu, C. M. Chiang, D. Loukinov, V. Lobanenko, R. Ohlsson, and E. Klenova. 2007. CTCF interacts with and recruits the largest subunit of RNA polymerase II to CTCF target sites genome-wide. *Mol. Cell. Biol.* **27**:1631–1648.
- Cockerill, P. N. 2000. Identification of DNase I hypersensitive sites within nuclei. *Methods Mol. Biol.* **130**:29–46.
- Cockerill, P. N. 2004. Mechanisms of transcriptional regulation of the human IL-3/GM-CSF locus by inducible tissue-specific promoters and enhancers. *Crit. Rev. Immunol.* **24**:385–408.
- Cockerill, P. N., A. G. Bert, D. Roberts, and M. A. Vadas. 1999. The human granulocyte-macrophage colony-stimulating factor gene is autonomously regulated in vivo by an inducible tissue-specific enhancer. *Proc. Natl. Acad. Sci. USA* **96**:15097–15102.
- Cockerill, P. N., M. F. Shannon, A. G. Bert, G. R. Ryan, and M. A. Vadas. 1993. The granulocyte-macrophage colony-stimulating factor/interleukin 3 locus is regulated by an inducible cyclosporin A-sensitive enhancer. *Proc. Natl. Acad. Sci. USA* **90**:2466–2470.
- Engel, N., J. L. Thorvaldsen, and M. S. Bartolomei. 2006. CTCF binding sites promote transcription initiation and prevent DNA methylation on the maternal allele at the imprinted H19/Igf2 locus. *Hum. Mol. Genet.* **15**:2945–2954.
- Gause, M., C. A. Schaaf, and D. Dorsett. 2008. Cohesin and CTCF: cooperating to control chromosome conformation? *Bioessays* **30**:715–718.
- Grosfeld, F., G. B. van Assendelft, D. R. Greaves, and G. Kollias. 1987. Position-independent, high-level expression of the human beta-globin gene in transgenic mice. *Cell* **51**:975–985.
- Hark, A. T., C. J. Schoenherr, D. J. Katz, R. S. Ingram, J. M. Levorso, and S. M. Tilghman. 2000. CTCF mediates methylation-sensitive enhancer-blocking activity at the H19/Igf2 locus. *Nature* **405**:486–489.
- Hawwari, A., J. Burrows, M. A. Vadas, and P. N. Cockerill. 2002. The human IL-3 locus is regulated cooperatively by two NFAT-dependent enhancers that have distinct tissue-specific activities. *J. Immunol.* **169**:1876–1886.
- HersHKovitz, M., and A. D. Riggs. 1997. Ligation-mediated PCR for chromatin-structure analysis of interphase and metaphase chromatin. *Methods* **11**:253–263.
- Johnson, B. V., A. G. Bert, G. R. Ryan, A. Condina, and P. N. Cockerill. 2004. Granulocyte-macrophage colony-stimulating factor enhancer activation requires cooperation between NFAT and AP-1 elements and is associated with extensive nucleosome reorganization. *Mol. Cell. Biol.* **24**:7914–7930.
- Kim, T. H., Z. K. Abdullaev, A. D. Smith, K. A. Ching, D. I. Loukinov, R. D. Green, M. Q. Zhang, V. V. Lobanenko, and B. Ren. 2007. Analysis of the vertebrate insulator protein CTCF-binding sites in the human genome. *Cell* **128**:1231–1245.
- Kontaraki, J., H. H. Chen, A. Riggs, and C. Bonifer. 2000. Chromatin fine structure profiles for a developmentally regulated gene: reorganization of the lysozyme locus before trans-activator binding and gene expression. *Genes Dev.* **14**:2106–2122.
- Kuhn, E. J., and P. K. Geyer. 2003. Genomic insulators: connecting properties to mechanism. *Curr. Opin. Cell Biol.* **15**:259–265.
- Kuhn, E. J., M. M. Viering, K. M. Rhodes, and P. K. Geyer. 2003. A test of insulator interactions in *Drosophila*. *EMBO J.* **22**:2463–2471.
- Lee, G. R., P. E. Fields, T. J. Griffin, and R. A. Flavell. 2003. Regulation of the Th2 cytokine locus by a locus control region. *Immunity* **19**:145–153.
- Lefevre, P., J. Witham, C. E. Lacroix, P. N. Cockerill, and C. Bonifer. 2008. The LPS-induced transcriptional upregulation of the chicken lysozyme locus involves CTCF eviction and noncoding RNA transcription. *Mol. Cell* **32**:129–139.
- Losada, A. 2007. Cohesin regulation: fashionable ways to wear a ring. *Chromosoma* **116**:321–329.
- Maksimenko, O., A. Golovnin, and P. Georgiev. 2008. Enhancer-promoter communication is regulated by insulator pairing in a *Drosophila* model bigenic locus. *Mol. Cell. Biol.* **28**:5469–5477.
- Maxam, A. M., and W. Gilbert. 1977. A new method for sequencing DNA. *Proc. Natl. Acad. Sci. USA* **74**:560–564.
- Nicola, N. A. 1989. Hemopoietic cell growth factors and their receptors. *Annu. Rev. Biochem.* **58**:45–77.
- Nimer, S. D., and H. Uchida. 1995. Regulation of granulocyte-macrophage colony-stimulating factor and interleukin 3 expression. *Stem Cells* **13**:324–335.
- Ohlsson, R., R. Renkawitz, and V. Lobanenko. 2001. CTCF is a uniquely versatile transcription regulator linked to epigenetics and disease. *Trends Genet.* **17**:520–527.
- Orlando, V., H. Strutt, and R. Paro. 1997. Analysis of chromatin structure by in vivo formaldehyde cross-linking. *Methods* **11**:205–214.
- Osborne, C. S., M. A. Vadas, and P. N. Cockerill. 1995. Transcriptional regulation of mouse granulocyte-macrophage colony-stimulating factor/IL-3 locus. *J. Immunol.* **155**:226–235.
- Parelho, V., S. Hadjur, M. Spivakov, M. Leleu, S. Sauer, H. C. Gregson, A. Jarmuz, C. Canzonetta, Z. Webster, T. Nesterova, B. S. Cobb, K. Yokomori, N. Dillon, L. Aragon, A. G. Fisher, and M. Merkenschlager. 2008. Cohesins functionally associate with CTCF on mammalian chromosome arms. *Cell* **132**:422–433.
- Peric-Hupkes, D., and B. van Steensel. 2008. Linking cohesin to gene regulation. *Cell* **132**:925–928.
- Renda, M., I. Baglivo, B. Burgess-Beusse, S. Esposito, R. Fattorusso, G. Felsenfeld, and P. V. Pedone. 2007. Critical DNA binding interactions of the insulator protein CTCF: a small number of zinc fingers mediate strong

- binding, and a single finger-DNA interaction controls binding at imprinted loci. *J. Biol. Chem.* **282**:33336–33345.
36. **Stedman, W., H. Kang, S. Lin, J. L. Kissil, M. S. Bartolomei, and P. M. Lieberman.** 2008. Cohesins localize with CTCF at the KSHV latency control region and at cellular c-myc and H19/Igf2 insulators. *EMBO J.* **27**:654–666.
 37. **Tagoh, H., P. N. Cockerill, and C. Bonifer.** 2006. In vivo genomic footprinting using LM-PCR methods. *Methods Mol. Biol.* **325**:285–314.
 38. **Uhlmann, F.** 2008. Molecular biology: cohesin branches out. *Nature* **451**:777–778.
 39. **Vostrov, A. A., M. J. Taheny, and W. W. Quitschke.** 2002. A region to the N-terminal side of the CTCF zinc finger domain is essential for activating transcription from the amyloid precursor protein promoter. *J. Biol. Chem.* **277**:1619–1627.
 40. **Wallace, J. A., and G. Felsenfeld.** 2007. We gather together: insulators and genome organization. *Curr. Opin. Genet. Dev.* **17**:400–407.
 41. **Wendt, K. S., K. Yoshida, T. Itoh, M. Bando, B. Koch, E. Schirghuber, S. Tsutsumi, G. Nagae, K. Ishihara, T. Mishiro, K. Yahata, F. Imamoto, H. Aburatani, M. Nakao, N. Imamoto, K. Maeshima, K. Shirahige, and J. M. Peters.** 2008. Cohesin mediates transcriptional insulation by CCTC-binding factor. *Nature* **451**:796–801.
 42. **West, A. G., M. Gaszner, and G. Felsenfeld.** 2002. Insulators: many functions, many mechanisms. *Genes Dev.* **16**:271–288.
 43. **Wijgerde, M., F. Grosveld, and P. Fraser.** 1995. Transcription complex stability and chromatin dynamics in vivo. *Nature* **377**:209–213.
 44. **Xie, X., T. S. Mikkelsen, A. Gnirke, K. Lindblad-Toh, M. Kellis, and E. S. Lander.** 2007. Systematic discovery of regulatory motifs in conserved regions of the human genome, including thousands of CTCF insulator sites. *Proc. Natl. Acad. Sci. USA* **104**:7145–7150.
 45. **Yusufzai, T. M., H. Tagami, Y. Nakatani, and G. Felsenfeld.** 2004. CTCF tethers an insulator to subnuclear sites, suggesting shared insulator mechanisms across species. *Mol. Cell* **13**:291–298.