Mechanism of Polycomb recruitment to CpG islands revealed by inherited disease-associated mutation

Valentina S. Caputo¹, Joana R. Costa¹,², Kalliopi Makarona¹, Elisabeth Georgiou¹,³, D. Mark Layton¹, Irene Roberts¹ and Anastasios Karadimitris¹,∗

¹Centre for Haematology, Department of Medicine, Imperial College London, Du Cane Road, London W12 0NN, UK, ²CEDOC, Faculdade de Ciências Médicas, FCM, Universidade Nova de Lisboa, Lisboa 1169-056, Portugal and ³Department of Biological Chemistry, School of Medicine, Aristotle University of Thessaloniki, Thessaloniki, Greece

Received January 21, 2013; Revised March 29, 2013; Accepted April 5, 2013

INTRODUCTION

Polycomb is a highly conserved transcriptional repressive complex critical for transcriptional regulation of genes involved in development (1). Polycomb repressed genes are poised, i.e. can be either activated or permanently silenced during developmental and cell differentiation processes. The functional hierarchy of in cis and trans regulators of transcription that determines Polycomb recruitment to gene promoters of mammalian cells and in particular to the promoters of housekeeping genes in somatic cells remains to be established. To generate an integrated network of Polycomb and of in cis and trans regulators and its influence on transcription of active genes in somatic cells, we took advantage of the pathogenic C>G mutation in the promoter of the housekeeping gene PIGM (2). We previously showed that the pathogenic C>G mutation falls in a −270GC-rich box in the core promoter of PIGM thereby abolishing binding of TF Sp1 in vitro (2, 4). This GC-rich box is part of a 697 bp CpG island (CGI) spanning the promoter of PIGM (Fig. 1A). Using chromatin immunoprecipitation in conjunction with real-time PCR (ChIP–RQ–PCR), we show that although binding of Sp1 to the Mut −270GC-rich box in vivo in B cells is severely decreased, it is not completely abrogated (Fig. 1B and Supplementary Material, Fig. S1A) whereas PIGM mRNA levels in Mut B cells were 1% of normal indicating a strong repressing effect of the pathogenic mutation (Fig. 1C). Although not tested directly, it is possible that other Sp factors, such as Sp3, also bind to the −270GC-rich motif and thus may be implicated in the pathogenesis of IGD.

RESULTS AND DISCUSSION

How the transcription repressing complex Polycomb interacts with transcriptional regulators at housekeeping genes in somatic cells is not well understood. By exploiting a CpG island (CGI) point mutation causing a Mendelian disease, we show that DNA binding of activating transcription factor (TF) determines histone acetylation and nucleosomal depletion commensurate with Polycomb exclusion from the target promoter. Lack of TF binding leads to reversible transcriptional repression imposed by nucleosomal compaction and consolidated by Polycomb recruitment and establishment of bivalent chromatin status. Thus, within a functional hierarchy of transcriptional regulators, TF binding is the main determinant of Polycomb recruitment to the CGI of a housekeeping gene in somatic cells.

We found histone (H) 3 and 4 hypoacetylation in the mutant PIGM promoter (Fig. 1D and Supplementary Material, Fig. S1B) concomitant with the lack of Sp1 binding both of which are restored to normal levels upon exposure to histone deacetylases (HDAC) inhibitors including butyrate (5) (Fig. 1E and F, Supplementary Material, Fig. S1C and D), thus implicating HDAC as important in gene repression in IGD. This is supported by the ability of butyrate to control treatment-intractable epilepsy in patients with IGD (4). H4 re-acetylation,
Figure 1. Sp1-dependent reversible transcriptional repression associated with the PIGM core promoter −270GC-rich box C > G mutation. (A) Schematic representation of PIGM, a single exon gene. Note that −270 refers to distance from the start codon ATG. (B) ChIP analysis of Sp1 relative enrichment at the PIGM promoter of WT and Mut B cells, with a DHFR promoter region used as a positive control (n = 3). (C) PIGM mRNA levels relative to GAPDH in WT, Het and Mut B cells (n = 3). (D) ChIP analysis for histone 3 (H3Ac) and histone 4 (H4Ac) acetylation enrichment at the PIGM promoter of WT and Mut B cells (n = 3). (E–G) ChIP analysis of Sp1(E) and H4Ac enrichment (F) at the PIGM promoter and PIGM mRNA levels (G) in response to butyrate (Bu; 3 mM) treatment over 36 h. Representative of three-independent experiments. Data are normalized to WT 0 h values. (H) FACS analysis of cell surface GPI expression after staining with FLAER of Mut, WT and Het B cells (left panel) and of Mut B cells upon Bu treatment (right panel) for the time points shown. All data are shown as mean ± SEM.
increased Sp1 binding and an increase in PIGM mRNA levels in Mut B cells in vitro were observed at 30 min, 1 and 3 h, respectively, post-exposure to butyrate reaching levels comparable with those of WT cells by 24 h (Fig. 1G and Supplementary Material, Fig. S1C–E), thereby resulting in restoration of the GPI biosynthetic pathway as assessed by the expression of GPI at the cell surface (Fig. 1H and Supplementary Material, Fig. S1F). These effects of butyrate on the mutant PIGM promoter acetylation and transcription are specific as they are not observed with GAPDH, another housekeeping gene (Supplementary Material, Fig. S1G).

The increase in PIGM mRNA was not affected by the protein synthesis inhibitor cyclohexamide (CHX), suggesting a direct effect of butyrate on PIGM transcriptional activation (Supplementary Material, Fig. S1H). Consistent with the idea that baseline transcriptional activation of PIGM in WT cells requires Sp1 and that de-repression of PIGM transcription in Mut cells in response to butyrate requires restoration of Sp1 binding, transcriptional activation of PIGM in WT, Het and in Mut B cells treated with butyrate was abrogated by mithramycin A (MitA) (Supplementary Material, Fig. S1I), an agent that inhibits binding of GC-rich motif-binding TF, including Sp1 (6–8). Together, these data show that Sp1 binding to the −270GC-rich motif is required for histone acetylation and transcriptional activation of PIGM and conversely, re-acetylation of histones is associated with restoration of Sp1 binding and PIGM expression.

On the basis of these observations, we first explored the role of CGI in nucleosome dynamics. CGI are promoter GC-rich DNA modules critical for transcriptional regulation of active genes (9, 10) and are reported to inherently de-stabilize nucleosome modules critical for transcriptional regulation of active genes (CGI) in nucleosome dynamics. CGI are promoter GC-rich DNA dependent of activating TF binding is unclear. By employing a cells (Fig. 2A) consistent with nucleosomal compaction in the C

270GC-rich motif and the consequent histone
acetylation appears to be a critical requirement for nucleosomal configuration that is compatible with active gene transcription.

In view of the dynamic nature of nucleosome compaction, and taking into account the temporal relationship of the in cis and trans events shown in Fig. 1E–G, we propose that the C > G mutation does not completely abolish binding of Sp1 to the mutated −270GC-rich motif but, instead, that it reduces it sufficiently so that it leads to reduced recruitment of histone acetyltransferases (HAT) by Sp1, to decreased nucleosome acetylation and compaction and, thus, to reduced accessibility of the −270GC-rich motif. These events are likely to operate in a self-enforcing, co-operative manner until the mutant motif is inaccessible to Sp1 most of the time. This model of self-enforcement may also unfold in reverse in response to HDAC inhibitors: increased histone acetylation, leading to nucleosome relaxation and increased accessibility of the −270GC-rich motif, even in its mutant form, thereby allowing some Sp1 binding. This would lead to the Sp1-dependent recruitment of HAT, promoting further histone acetylation, nucleosome relaxation and motif accessibility and, eventually, full restoration of Sp1 binding and transcriptional activation.

Since transcriptional repression by nucleosome compaction can be further consolidated by changes in DNA and histone methylation (13), we analyzed the mutant promoter for the repressive H3K9me2 and H3K27me3 histone modifications, characteristic marks of heterochromatin and developmentally regulated/poised genes, respectively (14). While there was no difference in H3K9me2 levels (Fig. 3A), compared with WT, we observed a significant enrichment for the H3K27me3 mark in the mutant promoter (Fig. 3B and Supplementary Material, Fig. S3A and B), consistent with transcriptional repression of PIGM by Polycomb (1, 15, 16). Supporting this, Suz12, EzH2 and Ring1B, components of the PcG PRC2 and PRC1 complexes, respectively (1), were significantly enriched in the mutant promoter as was Jarid2, a de-methylase that contributes to the recruitment of PcG to GC-rich domains (17, 18) (Fig. 3C and D). Treatment of Mut cells with 3-deazaneplanocin A (DZNep), a pharmacological agent that depletes PRC2 components (19–21), including EzH2, increased PIGM mRNA levels by up to 70-fold (Fig. 3E), supporting the importance of PcG-mediated transcriptional repression of the mutant promoter. Further, restoration of Sp1 binding in response to butyrate caused a time-dependent decrease in the enrichment of EzH2 and H3K27me3 but not of H3K4me3 marks in the mutant PIGM promoter (Fig. 3F; see below). Thus, PcG recruitment and occlusion of NDR (Fig. 2C). Upon treatment of Mut cells with butyrate, the ‘fixed’ pattern reverted to a pattern resembling that of the ‘moving’ nucleosomes.
Figure 2. Nucleosomal positioning and occupancy associated with lack of Sp1 binding to CGI. (A) Restriction enzyme accessibility assay in WT and Mut B cells. AatII digestion was performed in isolated nuclei to monitor accessibility within the core promoter of PIGM. Digestion levels were quantified by RQ–PCR relative to a control region of the PIGM promoter lacking AatII restriction site (amplicons 2 and 5 in Fig. 1A; n = 3). (B) Nucleosome mapping by the MNase protection assay on WT, Mut and Bu-treated (3 mM, 15 h) Mut B cells. The abundance of twelve 100 bp amplicons that overlap by 50 bp covering the length of the PIGM promoter region was quantified within equal amounts of MNase-digested or undigested DNA by RQ–PCR. GAPDH promoter amplicon is shown as an accessibility control for WT, Mut and butyrate-treated B cells. Plots show mean levels of relative protection ± SEM (n = 3). Amplicon b is shown in bold. (C) Ligation-mediated PCR analysis in WT and Mut B cells with and without Bu treatment using an amplicon b forward primer and an adaptor-specific reverse primer. (D) PIGM promoter DNA sequence. AatII and BstUI sites are underlined. C > G mutation is shown with a star. 270 Sp1 binding site is shown in italics. The nucleotide sequence of the promoter DNA corresponding to the nucleosome spanning the TSS (arrow) and the −270 GC-rich box in the Mut promoter is highlighted in grey (see also Supplementary Material, Fig. S2B).
deposition of PcG-mediated H3K27me3 marks to the mutant CGI consolidates the nucleosome compaction caused by the C > G mutation and the lack of Sp1 binding to the core promoter –270GC-rich motif. These data also conclusively show that the lack of activating TF binding and gene transcription at GC-rich promoters leads to PcG recruitment.

To analyze the impact of the C > G mutation on the activating histone mark status, we performed ChIP–RQ–PCR for the H3K4me3 modification which, as expected for a housekeeping gene (14), was found highly enriched in the WT PIGM gene (Fig. 4A) and in particular in the core promoter (Supplementary Material, Fig. S4A). Surprisingly, we also found comparable or higher than in WT H3K4me3 enrichment in the mutant promoter (Fig. 4A) suggesting the presence of a bivalent chromatin state of the antagonistic H3K4me3/H3K27me3 marks. Co-existence of these marks in the promoters of developmental genes in embryonic stem cells is required for their dynamic regulation during development and differentiation (22–25). To establish that the bivalent status represents true bivalency on the same promoter, we studied histone methylation in parental Het B cells (Supplementary Material, Fig. S4B). Taking advantage of the novel BstU I RE site created by the C > G mutation (Fig. 1A and Supplementary Material, Fig. S4B) we found that upon BstUI digestion in Het cells, levels of enrichment of H3K4me3 and CFP1, a CGI-binding protein required for H3K4 trimethylation (26), decreased by ~50%, (Fig. 4B and C) while enrichment of H3K27me3, EzH2, Ring1B and Suz12 decreased by ~70% (Fig. 4D and E; Supplementary Material, Fig. S4C). These results are consistent with simultaneous presence at the Mut PIGM promoter of H3K4me3, H3K27me3, and EzH2 at the PIGM promoter of Mut B cells after 5 and 24 h treatment with Bu (3 mM; n = 3). All data are shown as mean ± SEM.

**Figure 3.** Lack of Sp1 binding to CGI results in Polycomb recruitment and repressive H3K27me3 mark deposition. (A–D) ChIP analysis of H3K9me2 (A), H3K27me3 (B), Jarid2, Suz12 (C), EzH2 and Ring1B (D) enrichment at the PIGM promoters of WT and Mut B cells (n = 3). HOXA2, a member of the Hox genes that are known to be repressed by PcG and H3K27 trimethylation is shown as a positive control. (E) PIGM mRNA levels in Mut, WT and Het B cells after treatment for 72 h with the EzH2 inhibitor DzNep (3 μM). Data normalized to DMSO-treated cells (n = 3). (F) Enrichment of H3K27me3, H3K4me3, and EzH2 at the PIGM promoter of Mut B cells after 5 and 24 h treatment with Bu (3 mM; n = 3). All data are shown as mean ± SEM.
Figure 4. Absence of Sp1 binding to CGI leads to bivalent chromatin domain-mediated transcriptional repression but not to DNA methylation. (A) ChIP analysis of H3K4me3 enrichment at the PIGM promoter of WT and Mut B cells (n = 3). (B–E) Allele-specific ChIP analysis (see Supplementary Material, Fig. S4B) in Het1 cells. ChIP analysis of H3K4m3, CFP1, H3K27m3, EzH2, Suz12 and Ring1B enrichment levels at the PIGM promoter of Het B cells with and without BstUI digestion (see Fig. 1A). Enrichment for H3K27me3, EzH2, Suz12 and Ring1B at an area of HOXA2 gene that lacks BstUI restriction site is shown as a positive control (n = 3). (F and G) Allele specific, reciprocal re-ChIP analysis of H3K27me3-H3K4me3 co-occupancy in Het1 B cells. Large graphs show enrichment in the second ChIP, with and without BstUI digestion, normalized to IgG control. Insets show enrichment as % of Input in the first ChIP of the re-ChIP assay. All data are shown as mean ± SEM (n = 3). (H) PIGM CGI bisulphite sequencing of two different Mut and one WT B cell line. Filled circles represent methylated CpG, novel CpG created by the C > G mutation is indicated by an arrowhead. TSS is shown by an arrow, % methylation is shown on the right. Methylation at the Oct4 gene is shown as a positive control (right).
Finally, bisulphite sequencing showed that the *PIGM* CGI is unmethylated in WT cells (Fig. 4G), consistent with active gene status (30). The DNA methylation status of the *PIGM* CGI in Mut cells did not change and there was no change in *PIGM* mRNA levels and surface GPI expression in these cells when they were treated with the DNA de-methylating agent 5′-azacytidine (5′-Aza) (Supplementary Material, Fig. S4D). These findings are contrary to earlier as well as more recent data that directly or indirectly link lack of TF, including of Sp1, binding to cognate motifs to methylation of regulatory DNA (31–33). Instead, they indicate that lack of activating TF binding or PcG-mediated transcriptional repression is not by themselves sufficient for CGI DNA methylation and gene silencing. Instead, they are in line with genome-wide and structural data (34, 35) showing that the presence of H3K4me3 marks protects CGI promoters from DNA methylation and thus permanent transcriptional silencing. They are also consistent with previous work showing that in Sp1−/− murine embryonic stem cells, CGI of genes regulated by Sp1 are not methylated (36).

In conclusion, our data strongly support a model of a functional hierarchy, which, with activating TF as its central hub, determines recruitment or exclusion of PcG from CGI (Supplementary Material, Fig. S5). Specifically, we propose that histone acetylation and nucleosomal depletion, PcG exclusion and thus active transcription are all dependent on activating TF binding to CGI. In contrast, lack of TF-binding histone is associated with conservation of activating H4K3me3 histone marks, PcG recruitment and imposition of repressing H3K27me3 methylation and thus conversion of an active promoter to a repressed, bivalent chromatin state. Preservation of H4K3me3 histone marks at CGI, even at the repressed state, prevents DNA methylation by DNA methyltransferases, possibly through steric hindrance (35), thus averting permanent gene silencing. In addition, our observations constitute the first example of PcG-mediated bivalent chromatin domain transcriptional repression as a pathogenetic mechanism in Mendelian disease.

**MATERIALS AND METHODS**

**Cell lines and treatment with inhibitors**

Normal (WT), patient (Mut) and parental (Het) EBV-immortalized lymphoblastoid B cells were cultured in RPMI1640 with 10% FBS, 2 mM L-glutamine and 1 unit/ml penicillin—1 μg/ml streptomycin. Mut1, Mut2, Het1, Het2 and Het3 are from members of the same family, whereas Mut3 are cells from a patient from a different family.

Sodium butyrate (Bu), MitA, 5′-Aza, CHX and trichostatin A (TSA) were from Sigma. Vorinostat—suberoylanilide hydroxamic acid (SAHA) (Selleck Chemicals), DzNep (Cambridge Bioscience). Bu and 5′-Aza were prepared fresh in RPMI complete medium at 600 and 5 mM stocks, respectively. MitA, CHX, TSA, SAHA and DzNep were dissolved in DMSO at 1 mM, 100 mg/ml, 50, 200, 20 mM stock solutions and diluted prior to use at the indicated final concentrations.

**Gene expression analysis**

Total RNA was extracted using the GeneJet RNA purification kit (Fermentas-THERMO). After DNA traces were removed by DNase treatment (Fermentas-THERMO), cDNA was synthesized with random primers using the RevertAid First strand cDNA Synthesis Kit (Fermentas-THERMO). RQ–PCR experiments were carried out on AB 7500 Real-time PCR system using Maxima SYBR Green (Fermentas-THERMO). The primers used are listed in Supplementary Material, Table S1.

**Chromatin immunoprecipitation**

This was performed using a ChIPAb + Sp1 Millipore kit with some modifications. Specifically, cells were collected by centrifugation at 800 g. Chromatin was prepared using the Shearing ChIP kit (Diagenode). Chromatin was sonicated at 4°C for 20 min in a 0.5/0.5 on/off cycle using high intensity in a Bioruptor UCD-200 (Diagenode).

Post-sonication fragments of average 500 bp length were confirmed on a 1.5% agarose gel. The chromatin was diluted at least 10 times in ChIP dilution buffer (0.01% SDS, 1.1% Triton X-100, 1.2 mM EDTA, 16,7 mM Tris–HCL pH 8, 167 mM NaCl) containing protein inhibitors. The diluted chromatin was pre-cleared by incubation with magnetic beads (Dynabeads Protein G form Invitrogen) for 1 h at 4°C on a rotating wheel. Pre-cleared chromatin was incubated with 3 μg of antibodies overnight at 4°C on a rotating wheel.

Protein G magnetic beads were added and incubated for 2–4 h at 4°C on a rotating wheel. The immunoprecipitates were washed for 5 min at 4°C in a rotating wheel with low salt buffer (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris–HCl pH 8, 150 mM NaCl), high salt buffer (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris–HCl pH 8, 500 mM NaCl), LiCl buffer (0.25 M LiCl, 1% IGEPAL, 1% sodium deoxicolate, 1 mM EDTA, 10 mM Tris–HCl pH 8) and two times with TE.

Immuno complexes were eluted by adding 150 ml elution buffer (50 mM Tris–HCl pH 8, 50 mM NaCl, 1 mM EDTA and freshly added 1% SDS and 20 mg/ml of RNaseA) at 65°C for 4 h to overnight and a second time for 30 min. Both elutions were pooled and treated with proteinase K (Fermentas-THERMO).

DNA was purified by phenol extraction and DNA precipitation. RQ–PCR was performed using Maxima SYBR Green (Fermentas-THERMO) on an AB 7500 Real-time PCR system.

Enrichment of the target sequence was assessed against an isotypic IgG antibody—immunoprecipitated aliquot. The ΔΔCt method was used to calculate the enrichment of the target sequence, compared with the negative IgG control: \( \Delta \Delta C_t = \Delta C_t (C_{\text{sample}} - C_{\text{input}}) - \Delta C_t (C_{\text{IgG}} - C_{\text{input}}) \). The fold difference of the experimental sample relative to the IgG control (both normalized to the input sample) was expressed by the power \( 2^{-\Delta \Delta C_t} \).

**Allele-specific ChIP assay**

ChIP from Het B cell lines was performed as described above. After purification, equal amounts of DNA were diluted in RE buffer and treated or not with BstU (NEB) under the same conditions. Under these conditions, while the *PIGM* mutant promoter allele is sensitive to BstU digestion (Fig. 1A), the undigested DNA corresponds to both, mutated and unmutated alleles. Digestion efficiency was assessed by RQ–PCR using a
SUPPLEMENTARY MATERIAL

Supplementary Material is available at HMG online.

ACKNOWLEDGEMENTS

This was work was supported by the Medical Research Council (V.S.C., M.L., A.K.), the Fundacao para a Ciencia e Tecnologia, Lisboa, Portugal (J.R.C.), the Foundation of the Hellenic Society (V.S.C., M.L., A.K.), the Fundacao para a Ciencia e Tecnologia, Portugal, J. (2007) Entropically-driven binding of mithramycin in the minor groove of C/G-rich DNA sequences. Nucleic Acids Res., 35, 2215–2226.


