Vitamin D-dependent chromatin association of CTCF in human monocytes

Antonio Neme 1, Sabine Seuter 1, Carsten Carlberg *

School of Medicine, Institute of Biomedicine, University of Eastern Finland, FI 70211, Kuopio, Finland

Abstract

CCCTC-binding factor (CTCF) is a transcription factor being involved in 3D chromatin organization and displays a highly conserved genome-wide binding pattern. In this study, we report the cistrome of CTCF in THP-1 human monocytes and confirm that from the 40,078 CTCF binding sites nearly 85% are identical with those found in K562 monocytes. Quadruplicate chromatin immunoprecipitation sequencing (ChIP-seq) demonstrated that at 2130 loci the association strength of CTCF with genomic DNA was significantly (p < 0.05) modulated by stimulation with the natural vitamin D receptor (VDR) ligand 1α,25-dihydroxyvitamin D3 (1,25(OH)2D3). Some 55% of these CTCF sites contribute to DNA looping and mark the anchors of 587 putative topologically associating domains (TADs) containing at least one VDR binding site and one 1,25(OH)2D3 target gene. These TADs can explain the regulatory scenarios of up to 70% of all 1,25(OH)2D3 target genes. A self-organizing map approach subdivided the vitamin D-sensitive CTCF sites into seven classes that can be distinguished by participation in DNA loop formation, binding to open chromatin, carrying binding motifs for CTCF or its relative BORIS, overlap with transcription start site (TSS) regions and binding of VDR. These variant molecular profiles suggest different mechanisms of the 1,25(OH)2D3-dependent action of CTCF. The co-location of VDR and 1,25(OH)2D3-dependent CTCF sites increases in the context of accessible chromatin and TSS regions but does not show any significant correlation with classical DNA binding mechanisms of CTCF. In conclusion, vitamin D-sensitive CTCF sites provide further mechanistic details to the epigenome-wide understanding of 1,25(OH)2D3-mediated gene regulation.

© 2016 Elsevier B.V. All rights reserved.

Keywords:
Vitamin D
CTCF
ChIP-seq
Gene regulation
Open chromatin
Vitamin D target genes

1. Introduction

CTCF is an evolutionarily conserved multi-zinc finger transcription factor [1,2] that is involved in diverse functions ranging from organizing chromatin into active and inactive regions, such as the gene clusters of high-expression genes [3] and HOXA [4], control of gene imprinting, like at the IGF2 locus [5], or gene repression, such as of the genes MYC [6] and LYZ [7]. In vertebrates CTCF is the major protein binding to insulators [8]. The latter are stretches of genomic DNA that have either a barrier function between heterochromatin and euchromatin or act as enhancer-blockers [1]. On both types of insulators the binding of CTCF mediates the formation of genomic DNA loops leading to a 3D chromosomal network. In a scale of hundreds of kilobases to a few megabases these loops are referred to as TADs [9]. These genomic regions contain either active or repressed compartments of chromosomes [10,11]. Based on high-throughput 3C (Hi-C) analyses Dixon et al. [10] suggested that the human genome is subdivided into some 2000 TADs. More recently, Rao et al. [12] provided higher resolution Hi-C maps indicating that a number of smaller loops are formed, when CTCF binds to genomic loci within TADs. Then, the protein mediates its enhancer blocking activity by shielding TSS regions of certain genes against activation by transcription factors binding to enhancers in their vicinity [13]. Thus, CTCF-CTCF loops function as insulated neighborhoods for gene regulation and their modulation or even loss can lead to changes in gene activity at the respective genomic region [14]. CTCF is ubiquitously expressed and chromatin immunoprecipitation sequencing (ChIP-seq) approaches in a variety of human cell types [15–21] indicated that the genome-wide binding pattern of CTCF
shows a conservation rate that is unusually high for a transcription fac-
tor. Furthermore, the method chromatin interaction analysis by paired-
end tag sequencing (ChIA-PET) has already been used for CTCF in a
number of tissues, such as the ENCODE tier 1 cell lines K562 (mono-
cytes) and MCF-7 (breast cancer), and demonstrated that also the
CTCF-CTCF loops are very conserved [22,23].

In a recent study [24], we identified CTCF as a transcription factor
involved in the epigenome-wide process of VDR signaling. The nuclear
receptor VDR is a widely expressed transcription factor that serves as
the unique high affinity target of the biologically most active vitamin
D metabolite 1,25(OH)2D3 [25,26]. VDR ChIP-seq in six human cell
culture models indicated that the nuclear receptor is able to bind to
>23,000 loci genome-wide [27]. Ligand-activated VDR interacts with
a number of nuclear proteins, such as co-activators and chromatin modi-
fiers [28], in order to transiently open chromatin at specific enhancer
and TSS regions. This results in the modulation of the expression of
>1000 VDR/1,25(OH)2D3 target genes [24]. Furthermore, by applying
the method formaldehyde-assisted isolation of regulatory elements
sequencing (FAIRE-seq) to 1,25(OH)2D3-stimulated THP-1 cells we
responded and physiologically meaningful model system for the
provides further insight into the mechanisms of 1,25(OH)2D3-mediated
gene regulation.

These data support the action of VDR. In this way, ligand-responsive CTCF sites
create a number of different gene regulatory scenarios, in which CTCF
binds to target genes and ligand-sensitive chromatin sites that
generate the pairs of different gene regulatory scenarios, in which CTCF
supports the action of VDR. In this way, ligand-responsive CTCF sites
provide further insight into the mechanisms of 1,25(OH)2D3-mediated
gene regulation.

2. Material and methods

2.1. Cell culture

The human acute monocytic leukemia cell line THP-1 [29] is a well
responding and physiologically meaningful model system for the
investigation of 1,25(OH)2D3-triggered physiological processes, such as
innate immunity and cellular growth [30–33]. The cells were grown
in RPMI 1640 medium supplemented with 10% fetal calf serum,
2 mM l-glutamine, 0.1 mg/ml streptomycin and 100 U/ml penicillin
and were kept at 37 °C in a humidified 95% air/5% CO2 incubator. Prior
to treatment, the cells were grown overnight in phenol
red-free medium supplemented with charcoal-stripped fetal calf
serum and then treated with vehicle (0.1% ethanol (EtOH)) or 100 nM
1,25(OH)2D3 (Sigma-Aldrich).

2.2. CTCF ChIP

ChIP assays were performed as described by Zhang et al. [34] with
some modifications. After treatment of 2 × 107 THP-1 cells, nuclear
proteins were cross-linked to genomic DNA by adding formaldehyde
directly to the medium to a final concentration of 1% and incubating at
room temperature for 10 min on a rocking platform. Cross-linking was
stopped by adding glycine to a
final concentration of 160
μg/ml. The genomic DNA was isolated
from K562 cells (human myelogenous leukemia, GSM749690) were
downloaded. All ChIP-seq data were (re)analyzed at harmonized set-
tings: alignement with the human reference genome version hg19
using Bowtie software version 1.1.1 [35] with the following essential command line arguments: bowtie -n 1 -m 1
the data integration we used only those 44,078 peaks that were present
in at least three of the four repeats at either vehicle- or 1,25(OH)2D3-
treated samples (Table S1). Since we were interested in determining
the chromatin elution buffer (1% SDS, 1 mM EDTA, 150 mM NaCl, 1% Triton X-100, 0.1% sodium deoxycholate, protease
inhibitors, 50 mM HEPES-KOH, pH 7.5) and once in 10 ml nuclear lysis
buffer (1% SDS, 1 mM EDTA, 150 mM NaCl, 1% Triton X-100, 0.1% sodium
deoxycholate, protease inhibitors, 50 mM HEPES-KOH, pH 7.5). After
two washes with cell lysis buffer, the chromatin pellet was resuspended
in 700 μl of SDS lysis buffer (1% SDS, 10 mM EDTA, protease inhibitors,
50 mM Tris-HCl, pH 8.1) and the lysates were sonicated in a Bioruptor
Plus (Diagenode) to result in DNA fragments of 200 to 500 bp. Cellular
debris was removed by centrifugation. 340 μl aliquots of the lysate were
In this context, we identified CTCF binding sequences as the top-ranking transcription factor binding
motif below the summits of these 1,25(OH)2D3-modulated FAIRE
peaks.

In this study, we performed in four biological repeats CTCF ChIP-seq
in the absence and presence of 1,25(OH)2D3 in order to obtain a highly
accurate cistrome of the transcription factor in THP-1 cells. We found
44,078 common CTCF binding sites, 2130 of which were significantly
(p < 0.05) modulated by stimulation with 1,25(OH)2D3. These ligand-
dependent CTCF sites display co-location patterns with VDR, TSS regions
of 1,25(OH)2D3 target genes and ligand-sensitive chromatin sites that
create a number of different gene regulatory scenarios, in which CTCF
supports the action of VDR. In this way, ligand-responsive CTCF sites
provide further insight into the mechanisms of 1,25(OH)2D3-mediated
gene regulation.
replicates, i.e., it starts at the position of the peak with the most up-stream starting locus and ends at the most downstream ending locus. Two peak regions are considered to overlap if a portion of one of them is contained in the second, or if the distance between the closest extremes of the two regions is lower than 50 bp. Loci active in both untreated and 1,25(OH)₂D₃ treated conditions were tested for significant (p < 0.05) differences in peak strength, as reported by MACS2, through a paired Student's t-test. The summit regions (± 100 bp) of the 44,078 peak loci were searched de novo for transcription factor binding motifs by using the webtool HOMER [39]. To characterize the dynamics of the 2130 sites of 1,25(OH)₂D₃-modulated CTCF binding, we applied self-organizing maps (SOMs). SOM [40] is an unsupervised method that allows the identification of representative profiles of the inspected data via dimensionality reduction.

2.5. Statistics

Statistical significance was calculated by applying a paired Student's t-test. The Chi-square test was applied, in order to provide evidence, if there is a significant difference between the expected and observed frequencies in a given group of categories and instances. The k-means algorithm was used, in order to find a pre-specified number of centroids in data while taking the density into account. This allows a partition based on the actual distribution of data instead of a biased partition in

![Image](image-url)
cases where data are equally distributed. In addition, it creates categories with individual ranges.

3. Results

3.1. CTCF cistrome in THP-1 cells

In order to obtain a highly reliable cistrome of the transcription factor CTCF, we performed four independent ChIP-seq experiments with THP-1 cells that had been treated for 24 h with either vehicle (EtOH) or VDR ligand (1,25(OH)2D3). This approach identified a set of 44,078 genomic loci that bound significantly (false discovery rate (FDR) < 0.1%) CTCF at any of the treatment conditions and in at least three of the four biological replicates (Table S1). There are 36,376 CTCF sites present at both vehicle- and 1,25(OH)2D3-treated conditions, while 2161 and 5541 loci exclusively appeared only without or with ligand treatment, respectively (Fig. 1A). The average intensity of the regions lost and gained after 1,25(OH)2D3 stimulation is 3.55 and 4.09, respectively. In contrast, the peak intensity of those regions present in both conditions is 11.91 after 1,25(OH)2D3 treatment, and 10.01 for the EtOH treated. The CTCF loci are highly conserved, since 36,874 (83.7%) of them overlap with CTCF peaks that we obtained from harmonized re-analysis of ChIP-seq data from the related cell line K562 [22] (Fig. 1B and Table S2). Interestingly, 2130 (4.8%) of the 44,078 common CTCF peaks were significantly (p < 0.05) modulated by 1,25(OH)2D3 treatment (Fig. 1C and Table S2). At the very most of these sites (97.7%) CTCF binding was increased. At 1062 (49.3%) of all the sites CTCF binding was only between 1.20- and 1.49-fold increased, further 858 sites were between 1.50- and 1.99-fold enhanced and only 89 sites reached an induction of >2-fold (Fig. 1D and Table S2). We used HOMER for de novo transcription factor motif searches below the summits (± 100 bp) of the 2130 ligand-dependent CTCF loci or all 44,078 CTCF peaks and found binding sites for CTCF, the CTCF relative BORIS [41] and the E-box helix-loop-helix transcription factor NEUROD1 [42] as top ranking (sorted by p-value, Fig. 1D). The ligand-dependent CTCF loci showed a significantly higher percentage (based on Chi-square proportion test, \( \chi^2 = 4.27, p < 0.05 \)) for all three motifs than the group of all CTCF binding sites. Finally, we evaluated all 44,078 CTCF sites for an overlap with DNA loop anchor sites that had been identified by CTCF CHIA-PET assays in K562 cells (Fig. 1E and Table S2). Based on the high conservation of CTCF binding and function [23], this approach suggests that in THP-1 cells 20,396 CTCF sites (46.2%) were involved in DNA looping, 1106 of which were sensitive to stimulation with 1,25(OH)2D3.

In summary, in THP-1 cells the CTCF cistrome is formed by 44,078 genomic sites, 2130 of which are responsive to the VDR ligand 1,25(OH)2D3. In almost a third of these CTCF loci binding motifs for genomic sites, 2130 of which are responsive to the VDR ligand 1,25(OH)2D3. This TAD size distribution monitored a change in the difference of TADs between consecutive groups after the first three groups suggesting a threshold at 4.56 Mb. We considered for further analysis only those TADs with a maximum of this size. This approach linked 759 target genes (59.1% of all) with 861 CTCF loci and suggested the formation of 427 TADs, in which both anchors associate with vitamin D-responsive CTCF sites. Despite the TAD size limitation, 556 of all 587 TADs (94.7%) were identical to those obtained from the K562 ChiA-PET data without such limitation (Table S3).

Two examples of TADs flanked by a pair of vitamin D-sensitive CTCF sites are displayed. A TAD in the size of 331 kb contains the TSS of only one gene (CYP26B1) and three VDR binding sites 5, 21 and 54 kb downstream of this TSS (Fig. 2B). The transcription of the CYP26B1 gene was induced 23.1-fold by 1,25(OH)2D3 and the VDR sites are located within ligand-sensitive chromatin regions as determined by FAIRE-seq [24]. A second more complex example is a 430 kb TAD containing seven genes, of which only the erythrocyte membrane protein band 4.1 like 1 (EPB41L1) gene was 2.94-fold up-regulated by 1,25(OH)2D3 and controlled by four VDR sites 26 kb upstream and 7, 14 and 112 kb downstream of the gene’s TSS (Fig. 2C).

Taken together, vitamin D-sensitive CTCF sites mark the anchors of 587 TADs that contain at least one VDR binding site and one 1,25(OH)2D3 target gene. The 427 TADs with a size below 4.5 Mb can classify the regulatory scenarios of nearly 60% of all 1,25(OH)2D3-modulated targets in THP-1 cells.

3.3. Functional profiles of vitamin D-sensitive CTCF sites

We characterized the 1,25(OH)2D3-sensitive CTCF sites by a SOM approach that was based on the parameters overlap with i) a TSS region, ii) a VDR binding site (ChIP-seq from THP-1 cells), iii) 1,25(OH)2D3-sensitive or iv) all sites of accessible chromatin (FAIRE-seq from THP-1 cells [24]) or the presence of motifs for v) CTCF, vi) BORIS and vii) NEUROD1, and viii) overlap with a chromatin loop anchor (ChiA-PET data from K562 cells [22]). This sorted the 2130 CTCF sites into seven classes (Fig. 3 and Table S2).

The majority of the 833 THP-1 CTCF peaks involved in chromatin looping belong to classes 9 (296), 7 (242) and 3 (151). All the 399 peaks in class 9 are confirmed by ChiA-PET as anchor sites in K562, although not all of them are identified as anchor sites for THP-1. All the chromatin looping-related CTCF sites in class 9 overlap with accessible chromatin, while all sites of class 3 that co-locate with open chromatin contain a CTCF motif. From the 296 members in class 1 only four overlap with a TSS region, but none overlaps with VDR or co-locates with open chromatin. However, 290 of them contain a CTCF motif and all 296 contain a BORIS motif. In contrast, the 1297 ligand-sensitive CTCF sites that are not involved in DNA loop formation are primarily found in SOM classes 2 and 8. From the 333 CTCF sites in class 8, only one contains a CTCF motif but all co-locate with accessible chromatin.

In summary, a SOM approach subdivided the 2130 vitamin D-sensitive CTCF sites into seven classes that can be distinguished by i) participating in DNA loop formation, ii) overlapping with open chromatin, iii) carrying CTCF and/or BORIS motifs and iv) overlapping with a TSS region and VDR binding site. These variant molecular profiles suggest different mechanisms of action of CTCF.

3.4. Impact of VDR for ligand-dependent CTCF association

The SOM analysis directed towards more detailed inspections of the relation of CTCF binding to sites of open chromatin. At basal conditions triplicate FAIRE-seq assays in THP-1 cells indicated 62,231 genomic regions of accessible chromatin [24]. At 20,118 of these loci (32.3%) also CTCF binding was detected (Fig. 4A and Table S2), i.e. nearly half (45.6%) of the 44,078 common CTCF sites were associated with accessible chromatin. Similarly, at 1020 of the 2130 ligand-modulated sites (47.8%) CTCF was found within open chromatin. Interestingly, at 2139 of the 20,118 accessible CTCF loci (10.6%) our THP-1 ChIP-seq data confirmed the presence of VDR. For comparison, at non-accessible chromatin only 1.3% (319 out of 23,960) of the CTCF sites also bound
VDR, i.e. at accessible chromatin the VDR association rate was 8.1-times higher. At ligand-sensitive CTCF sites the VDR overlap percentage increased by even 10.7-fold from 0.9% (11 of 1110) to 9.6% (98 of 1020). An illustrative example is the 1,25(OH)2D3 target gene sirtuin 4 (SIRT4) that carries a ligand-dependent prominent VDR binding site 6.5 kb downstream of its TSS region (Fig. S1A). The ligand-

**Fig. 2.** TADs including VDR sites and 1,25(OH)2D3 target genes. Size distribution of 587 TADs (Table S3) defined by 1,25(OH)2D3-responsive CTCF sites overlapping with K562 cell ChIA-PET sites (A). Three size categories including 426 TADs with a size below 4.56 Mb clearly discriminate from the seven classes of 161 larger TADs. Please note that the k-means ranking created 10 categories with individual ranges of TAD size. The IGV browser was used to visualize TAD examples including the 1,25(OH)2D3 target genes CYP26B1 (B) and EPB41L1 (C). The peak tracks display merged data tracks from the quadruplicate CTCF ChIP-seq (purple), triplicate VDR ChIP-seq (unpublished, red) and triplicate FAIRE-seq ([24], turquoise) obtained from THP-1 cells treated for 24 h with vehicle (EtOH) or 1,25(OH)2D3 (1,25D). Gene structures are shown in blue and 1,25(OH)2D3 target genes are highlighted in red; their induction after 24 h ligand treatment is based on triplicate RNA-seq assays [24]. Regions of ligand-dependent CTCF and VDR binding are shaded in grey. The interpretation of CTCF and VDR binding in relation to TSS regions and DNA looping are schematically indicated (right).
dependence of both the CTCF and the VDR binding site was supported by 1,25(OH)2D3-dependent chromatin opening. A simple interpretation of this gene regulatory scenario is a DNA looping event, in which VDR was attracted to CTCF and the basal transcriptional machinery on the SIRT4 TSS, so that the transcription of the gene increased 2.62-fold within 24 h.

In total, 15,942 (36.1%) common CTCF sites carry either a CTCF, BORIS or NEUROD1 motif (Fig. 4B and Table S2). A slightly higher percentage was observed for ligand-dependent CTCF sites (817 of 2130, i.e. 38.3%). Similarly, for the individual transcription factors the overlap percentage with their binding motifs did not significantly change when the group of all CTCF sites was compared with its subset of vitamin D-sensitive sites (Fig. S2). Moreover, the VDR overlap rate of the CTCF sites was not majorly influenced by the presence of any of the three transcription factor motifs, i.e. the presence of none of these motifs seems to affect the vitamin D-sensitivity of CTCF sites.

At basal conditions 14,786 of 23,487 genes (63.0%) were expressed in THP-1 cells [24] and 4282 CTCF sites overlapped with TSS regions of expressed genes. From the latter genes 775 (18.1%) also bound VDR and 110 are 1,25(OH)2D3 target genes (Fig. 4C and Table S2). In contrast, at the 39,797 non-TSS CTCF loci the CTCF-VDR co-occupancy rate was only 4.3%, i.e. it was 4.2-fold higher at TSS regions. The average VDR-TSS co-location rate of ligand-dependent CTCF sites was 1.9% (41 of 2130) and most of them are found in SOM classes 8 (21) and 9 (18). However, only 6 of these TSS regions mark 1,25(OH)2D3 target genes (two up-regulated (SLC25A4 and CD274) and four down-regulated (SLC2A1, TRIM13, MIR3648 and CNR1)). For example, the CD274 gene showed on its TSS region 1.25(OH)2D3-stimulated CTCF and VDR binding as well as chromatin opening (Fig. S1B). In addition, a strong ligand-dependent VDR binding site is found 1 kb upstream of the CD274 TSS and 12 kb further upstream locates the TSS region of the 1,25(OH)2D3 target gene plasminogen receptor with a C-terminal lysine (PLGKRT). An obvious interpretation of the binding pattern is CTCF-mediated loop formation bringing the prominent VDR site into close vicinity of both TSS regions. This then leads to the accumulation of the transcripts of the PLGKRT gene by 1.93-fold and of the CD274 gene even by 23.71-fold. In this genomic region the VDR binding site is so prominent that, in contrast to the SIRT4 gene (Fig. S1A), even on both TSS regions ligand-induced VDR ChIP-seq peaks were detectable. These VDR peaks may result exclusively from looping events, i.e. VDR does not bind directly to the TSS regions of these two 1,25(OH)2D3 target genes.

Taken together, the co-location of VDR and 1,25(OH)2D3-dependent CTCF sites increases in the context of accessible chromatin and TSS regions but does not show any significant correlation with CTCF, BORIS or NEUROD1 binding motifs. Many vitamin D target genes are regulated by the combination of CTCF and VDR binding at genomic regions of open chromatin.

4. Discussion

This study presents the characterization of the highly accurate cistrome of the transcription factor CTCF in THP-1 cells being composed of 44,078 binding sites. This number of CTCF sites is more than double of those being reported in IMR90 human fibroblasts.
and 126,200 cell type-specific CTCF-binding sites, of which approximately 19,200 were ubiquitous quadruplicate CTCF dataset of 1,25(OH)2D3-stimulated cells in comparison based on 38 different cell lines [45] listed in total 326,840 CTCF-binding sites, of which approximately 19,200 were ubiquitous and 126,200 cell type-specific. The 44,078 CTCF sites in THP-1 cells primarily belong to the class of common binding sites and therefore their overlap with cellular models of comparable origin, such as K562 cells, is with nearly 85% very high. Accordingly, our dataset from THP-1 cells does not add many unique sites to those described in K562 cells (19,572) [15], HeLa human cervix carcinoma (19,308) [15], Jurkat human T cells (19,572) [15] and CD4+ T cells (20,262) [44]. A comparison based on 38 different cell lines [45] listed in total 326,840 CTCF-binding sites, of which approximately 19,200 were ubiquitous and 126,200 cell type-specific. The 44,078 CTCF sites in THP-1 cells primarily belong to the class of common binding sites and therefore their overlap with cellular models of comparable origin, such as K562 cells, is with nearly 85% very high. Accordingly, our dataset from THP-1 cells does not add many unique sites to those described in the database CTCFBSDB 2.0 (http://insulatordb.uthsc.edu) [46], but confirms that previous approaches of using CTCF binding data extrapolations from K562 cells were appropriate [47,48].

The most important finding of this study is that through the use of a quadruplicate CTCF dataset of 1,25(OH)2D3-stimulated cells in comparison to vehicle-treated cells 2130 CTCF binding loci were shown to be significantly sensitive to vitamin D. In parallel, we detected at 2161 and 5541 genomic sites CTCF binding exclusively in the absence or presence of VDR ligand, respectively. However, for the latter 7702 ligand-modulated CTCF sites no statistics and FC could be calculated. A similar dynamic component of the CTCF cistrome was already reported in a model of adipocyte differentiation [49]. Even if part of the CTCF binding dynamics of that study was caused by variations in the ChIP-seq assays and insufficient number of biological repeats, it clearly indicates that the association of CTCF with genomic DNA is not as constant as previously assumed.

A functional interaction of CTCF with other members of the nuclear receptor superfamily had been indicated before. The thyroid hormone receptor (THR) was found to co-locate with CTCF [50] and on a few selected sites thyroid hormone-dependent association of CTCF was monitored [51]. Moreover, also the estrogen receptor (ER) had been described to associate with CTCF [52], but genome-wide no estrogen-dependent effects on CTCF binding could be detected [53]. Thus, this study is the first providing genome-wide evidence that a nuclear hormone, such as 1,25(OH)2D3, can modulate the association of CTCF with genomic DNA.

The 2130 significantly ligand-sensitive CTCF sites represent only a rather small subset (4.8%) of the 44,078 common CTCF sites. In this study, we took advantage of the high conservation of CTCF-mediated loop formation, as monitored by ChIA-PET and Hi-C assays in K562 cells and other model systems, and extrapolated the location of the loop anchors to our CTCF data in THP-1 cells. Oti et al. [54] recently demonstrated that such an approach is valid. In this way, we observed that >800 of the 1,25(OH)2D3-dependent CTCF sites are located at positions that are important to insulate prominent VDR binding sites from inappropriate activation in their genomic environment. Thus, our study suggests that 587 TADs may be affected depending on the threshold settings, the expression levels of an as high number as 3700 genes are statistically significantly (p < 0.05) modulated by a 24 h stimulation with 1,25(OH)2D3 ([24] and unpublished results). However, we assume that the vast majority of these
apparent but mostly small ligand responses represent rather transcriptional noise resulting from a still insufficient insolation within the vitamin D signaling process than a regulation with a physiological purpose. Thus, 1,25(OH)2D3-induced tighter CTCF binding to TAD borders increases the specificity of vitamin D signaling. Future research may even uncover additional regulatory scenarios, where vitamin D may lead to the reorganization of TADs, such as demonstrated in classical cases of genomic imprinting.

Our SOM analysis indicated that there are a number of different constellations in which CTCF binds in a vitamin D-dependent fashion to these genomic loci. Since VDR is the only high affinity target of 1,25(OH)2D3 in the nucleus [25], the most simple mechanism to understand vitamin D-sensitive CTCF sites is their co-location with VDR. However, this applies only to 131 (9.9%) of all 1,25(OH)2D3-sensitive CTCF sites. Interestingly, the VDR-CTCF co-location rate was 8-times higher at accessible chromatin than at closed chromatin. In parallel, open chromatin was confirmed as one of the preferential pre-requisites for TAD anchor loci. This pointed towards another mechanism for the understanding of 1,25(OH)2D3-sensitive CTCF sites, which is the overlap with vitamin D-sensitive chromatin sites [24]. However, this scenario is found only for 132 of the 2130 sites (6.2%), 15 of which are overlapping with the first mechanism. A more indirect mechanisms may take place via the location of 1,25(OH)2D3-sensitive CTCF sites with TSS regions, in particular with that of vitamin D target genes to which VDR is looping from enhancer sites. Although TSS regions show a 4-fold increase in VDR-CTCF co-occupancy, this scenario also applies only to 197 vitamin D-sensitive CTCF sites (9.2% of all). Taken together, mechanisms that link directly or indirectly to VDR so far explain only some 25% of the vitamin D-sensitive CTCF sites.

Based on traditional understanding of gene regulation, transcription factors, such as CTCF and VDR, associate with their consensus binding motifs. However, in the case of VDR we already demonstrated previously that <20% of the known 23,000 genomic binding loci are traditional DR3-type sites [27]. This indicated that VDR must use a number of alternative binding modes [28]. Similarly in this study, we found in <50% of the CTCF peaks a CTCF consensus binding motif. In analogy to VDR, also CTCF seems to use a number of alternative scenarios, in order to contact genomic DNA. Interestingly, the presence of a CTCF, BORIS of NEUROD1 motif does not indicate any enrichment for VDR binding at the same site. This suggests that CTCF rather uses alternative, presently uncharacterized binding modes for its vitamin D-sensitive subset.

In conclusion, in this study we monitored the dynamic nature of the CTCF-organized epigenetic landscape of human monocytes in response to stimulation with 1,25(OH)2D3. The observation that genome-wide 2130 CTCF binding sites are significantly vitamin D-sensitive contributes to a further understanding of the mechanisms of 1,25(OH)2D3 signaling.

Supplementary data to this article can be found on line at http://dx.doi.org/10.1016/j.bbagrm.2016.08.008.

Conflict of interest

There is no conflict of interests between the authors and all authors disclose any financial interest that might be constructed to influence the results or interpretation.

Transparency document

The Transparency document associated with this article can be found in online version.

Acknowledgements

This work was supported by the Academy of Finland (#267067) and the Jusselius Foundation. Kind thanks to the Gene Core Facility at the EMBL in Heidelberg, Germany, for massively parallel sequencing services.

References
