Epigenome-wide effects of vitamin D and their impact on the transcriptome of human monocytes involve CTCF

Sabine Seuter†, Antonio Neme‡ and Carsten Carlberg*

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

Received July 09, 2015; Revised December 15, 2015; Accepted December 19, 2015

ABSTRACT

The physiological functions of vitamin D are mediated by its metabolite 1α,25-dihydroxyvitamin D₃ (1,25(OH)₂D₃) activating the transcription factor vitamin D receptor (VDR). In THP-1 human monocytes we demonstrated epigenome-wide effects of 1,25(OH)₂D₃ at 8979 loci with significantly modulated chromatin accessibility. Maximal chromatin opening was observed after 24 h, while after 48 h most sites closed again. The chromatin-organizing protein CTCF bound to 14% of the 1,25(OH)₂D₃-sensitive chromatin regions. Interestingly, 1,25(OH)₂D₃ affected the chromatin association of CTCF providing an additional mechanism for the epigenome-wide effects of the VDR ligand. The 1,25(OH)₂D₃-modulated transcriptome of THP-1 cells comprised 1284 genes, 77.5% of which responded only 24 h after stimulation. During the 1,25(OH)₂D₃ stimulation time course the proportion of down-regulated genes increased from 0% to 44.9% and the top-ranking physiological function of the respective genes shifted from anti-microbial response to connective tissue disorders. The integration of epigenomic and transcriptomic data identified 165 physiologically important 1,25(OH)₂D₃ target genes, including HTT and NOD2, whose expression can be predicted primarily from epigenomic data of their genomic loci. Taken together, a large number of 1,25(OH)₂D₃-triggered epigenome-wide events precede and accompany the transcriptional activation of target genes of the nuclear hormone.

INTRODUCTION

Vitamin D has a number of physiological functions, such as maintaining the balance of calcium and phosphorus homeostasis, controlling innate and adaptive immunity and modulating cellular growth, via its biologically most active metabolite 1,25(OH)₂D₃ (1–3). The nuclear hormone is a high affinity ligand of vitamin D receptor (VDR) (4), which is a transcription factor being expressed in most human tissues (http://biogps.org/#goto=generereport&id=7421). Taking all vitamin D responsive organs and cell types together, a few hundred genes are primary targets of VDR and 1,25(OH)₂D₃ (5). However, for a most comprehensive insight on the physiological actions of vitamin D not only short-term effects of 1,25(OH)₂D₃ on primary target genes need to be understood, but also long-term, secondary effects in the time frame of 1–2 days have to be taken into account.

VDR is a member of the nuclear receptor superfamily (6) that in humans comprises 48 members, many of which are modulated in their activity by small lipophilic molecules in the size of cholesterol (7). The principles of nuclear receptor signaling are well understood and seem to be very comparable for most members of the superfamily (8). For example, for VDR the canonical nuclear receptor signaling model is based on DNA binding sites formed by a direct repeat of two hexameric binding motifs spaced by three nucleotides (DR3), on which the receptor binds as a heterodimer with retinoid X receptor (RXR) (9,10). However, the genome-wide binding profile of VDR, as determined by the method chromatin immunoprecipitation sequencing (ChIP-seq) in six human cell culture models (11), indicated that only a minority of 11.5% of the 23 409 individual VDR binding sites carry a DR3-type sequence below their summit. This suggests that VDR has to use additional mechanisms to recognize its genomic targets (12).

After specific binding of their cognate ligand(s) nuclear receptors undergo a conformational change in their ligand-binding domain, which affects the detailed structure of the domain’s outer surface and its potential for protein-protein interaction with other nuclear proteins, such as co-activators and co-repressors (13). These co-factors then form larger complexes with chromatin modifying enzymes, such as histone deacetylases (HDACs), histone acetyltransferases and others (14). In most cases the net result of ligand

*To whom correspondence should be addressed. Tel: +358 40 355 3062; Email: carsten.carlberg@uef.fi
†These authors contributed equally to the paper as first authors.

© The Author(s) 2015. Published by Oxford University Press on behalf of Nucleic Acids Research.
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com
activation of nuclear receptors is the transient opening of chromatin at specific enhancer and transcription start site (TSS) regions resulting in the stimulation of gene transcription (15).

The main components of chromatin are nucleosomes formed of eight histone proteins around whom genomic DNA is wrapped (16). The default state of chromatin is densely packed heterochromatin that protects genes from un-controlled activation (17). This gene silencing is achieved by histone modifications, such as methylation of lysine residues, and the methylation of genomic DNA at cytosines. The positions of these epigenomic modifications are cell-specific and can change in response to environmental changes, such as the activation of signal transduction cascades (18). Accessible chromatin loci can be detected by methods as formaldehyde-assisted isolation of regulatory elements sequencing (FAIRE-seq) (19) and represent only a few percent of the whole genome, such as TSS and enhancer regions. In addition, the 3-dimensional organization of chromatin into larger and smaller scale loops has an important impact on the coordination of gene expression (20).

This chromatin looping does not only bring enhancer regions binding transcription factors, such as VDR, into close vicinity of TSS regions, but also sub-divides the genome into functional units, referred to as chromosome domains (21). A key protein in the latter process is the transcription factor CCCTC-binding factor (CTCF) that together with the protein cohesin defines insulator regions separating chromosomal domains from each other (22,23). Experimentally this can be monitored by the method chromatin interaction analysis by paired-end tag sequencing (ChIA-PET) (24).

In this study, we applied FAIRE-seq to 1,25(OH)2D3-stimulated THP-1 human monocytes and demonstrated that the chromatin accessibility of nearly 9000 loci distributed along the genome was significantly modulated by the VDR ligand. We characterized these sites for their (i) dynamics in a time frame of 48 h, (ii) overlap with TSS regions, (iii) VDR occupancy and (iv) occurrence of DR3-type binding sites. A screening for non-DR3-type binding motifs below the summits of FAIRE peaks indicated CTCF motifs as top ranking. CTCF ChIP-seq in THP-1 cells did not only confirm this prediction but also demonstrated that genomic CTCF binding can be modulated by 1,25(OH)2D3.

RNA isolation, cDNA synthesis and PCR
Total RNA was extracted using the High Pure RNA Isolation Kit (Roche) according to the manufacturer’s instructions. RNA quality was assessed by native agarose gel electrophoresis. cDNA synthesis and qPCR were performed as described previously (30). qPCR reactions were performed with the LightCycler® 480 System (Roche) using 250 nM of reverse and forward primers, 2 μl cDNA and the LightCycler 480 SYBR Green I Master mix (Roche). Primer-specific temperatures and sequences are listed in Supplementary Table S7. Relative mRNA expression levels were determined using the formula 2^(ΔΔCt), where ΔCt is Ct(target gene) – Ct(reference gene). The genes B2M, GAPDH and HPRT1 were used as references as described previously (28).

**MATERIALS AND METHODS**

**Cell culture**

The human acute monocytic leukemia cell line THP-1 (25) 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 (26–29). 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 mRNA or chromatin extraction, cells were first 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).

**Dicer substrate siRNA (DsiRNA) silencing**

THP-1 cells were transfected with either non-specific control (DsiRNA) oligomers (NC1) or specific DsiRNAs targeting CTCF mRNA (Supplementary Table S6). In addition, mock-transfected controls (no DsiRNA) were prepared. The cells were grown for 2 days in RPMI 1640 medium supplemented with 10% FCS, 2 mM L-glutamine, 0.1 mg/ml streptomycin and 100 U/ml penicillin. The transfection was performed in an Amaxa nucleofector I (Lonza) using the Cell Line Nucleofector Kit V according to the manufacturer’s instructions with the following modifications. Per sample, 2 x 10⁶ cells were harvested by centrifugation and re-suspended in 100 μl Nucleofector Solution V. A total of 200 pmol of a mixture of three different DsiRNA oligonucleotides was added and the cell suspension was transferred to an electroporation cuvette. The Nucleofector T-12 program was applied. Immediately after the pulse, 500 μl pre-warmed RPMI 1640 medium supplemented with 5% charcoal-stripped FCS, 2 mM L-glutamine, 0.1 mg/ml streptomycin and 100 U/ml penicillin was added and the cell suspension was transferred to pre-equilibrated 6-well plates containing 2.5 ml of medium per well. The extent of the knock-down was determined on mRNA level after an incubation for 48 h including ligand or solvent treatment for 24 h. The transfection efficiency has been determined after nuleofection of a fluorescently labeled DsiRNA oligo (IDT, Leuven, Belgium) by confocal microscopy to be ≥ 88%.
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 0.125 M and incubating at room temperature for 10 min on a rocking platform. The cells were collected by centrifugation and washed twice with ice cold phosphate-buffered saline (PBS). The cell pellets were subsequently resuspended twice in 10 ml cell lysis buffer (0.1% 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 diluted 1:5 in IP dilution buffer (1% Triton X-100, 2 mM EDTA, 150 mM NaCl, protease inhibitors, 20 mM Tris-HCl, pH 7.5). 4 μl anti-CTCF antibody (Millipore, 07-729) was coated to 60 μl Dynabeads Protein G (Invitrogen) in an overnight incubation at 4°C. The pre-formed bead-antibody complexes were then washed twice with beads wash buffer (0.1% Triton X-100, PBS, protease inhibitors) and added to the chromatin aliquots. The samples were incubated for overnight at 4°C on a rotating wheel to form and collect the immuno-complexes. The beads were washed sequentially for 5 min on a rotating wheel with 1 ml of the following buffers, each: twice cell lysis buffer, once high salt buffer (0.1% SDS, 1% Triton X-100, 10 mM EDTA, 350 mM NaCl, 0.1% sodium deoxycholate, 50 mM HEPES-KOH, pH 7.5), once ChIP wash buffer (250 mM LiCl, 1% Nonidet P-40, 0.5% sodium deoxycholate, 1 mM EDTA, 10 mM Tris-HCl, pH 8.0) and twice TE buffer (1 mM EDTA, 10 mM Tris-HCl, pH 8.0). Then, the immune complexes were eluted using 250 μl ChIP elution buffer (1% SDS, 10 mM EDTA, 50 mM Tris-HCl, pH 7.5) at 37°C for 30 min with rotation. The elution was repeated with a 10 min rotation and the supernatants were combined. The immune complexes were reverse cross-linked at 50°C for 2 h in the presence of proteinase K (Fermentas) in a final concentration of 160 μg/ml. The genomic DNA was isolated with the ChIP DNA Clean&Concentrator Kit (Zymo Research).

FAIRE
FAIRE analysis was performed according to the protocol published by Giresi et al. (32) with some modifications. In short, 18 x 10^6 THP-1 cells were cross-linked identically as for ChIP. After 10 min cross-linking and stopping with glycine the washed cell pellets were resuspended and incubated sequentially in 2 ml of buffer L1, 2 ml of buffer L2 and 700 μl of buffer L3. The lysates were sonicated in a Bioruptor to result in DNA fragments of 200 to 500 bp and cellular debris was removed by centrifugation. Input samples were reverse cross-linked overnight at 65°C. The FAIRE samples and reverse cross-linked reference samples were subjected to two sequential phenol/chloroform/isoamyl alcohol (25/24/1) extractions, resuspended in 10 mM Tris-HCl (pH 7.4) and treated with 1 μl of RNase A (10 mg/ml) for 1 h at 37°C. The genomic DNA was purified using the ChIP DNA Clean&Concentrator Kit.

ChIP-seq, FAIRE-seq and RNA-seq analyses
ChIP and FAIRE DNA templates and total RNA samples were sequenced at 50 bp read length using standard manufacturer protocols at the Gene Core at the EMBL (Heidelberg, Germany). In parallel, public ENCODE ChIP-seq data sets (33) from K562 cells for CTCF (GSM749690) and the transcription factors ETS1 (GSM803442), GABPA (GSM803524), NR2F2 (GSM1010782) and NR4A1 (GSM777637) were downloaded. All ChIP-seq data were reanalyzed at harmonized settings: alignment with the human reference genome version hg19 using Bowtie software version 1.1.1 (34) with the following essential command line arguments: bowtie -n 1 -m 1 -k 1 -c 70 –best. The aligned input and CTCF reads were converted to sorted BAM format using samtools (35) and, after merging the read sets per sample, converted to TDF format using igvtools, in order to allow efficient visualization in the Integrative Genomics Viewer (IGV) genome browser (36). Statistically significant ChIP-seq peaks were identified using Model-based Analysis of ChIP-Seq data (MACS) version 2 (37) with the following essential command line arguments: macs2 callpeak –bw 150 –keep-dup 1 –g hg19 –qvalue 0.01 –m 50 –bdg. Otherwise, default parameters were used.

For FAIRE-seq data analysis statistically significant peaks were identified using the Zinba program package version 2.02 by setting the mean fragment length at 200 bp and using other settings as recommended for FAIRE-seq in the Zinba website (http://code.google.com/p/zinba/wiki/UsingZINBA) including peak refinement (38). To enable comparisons across the full time series for all positions with a significant FAIRE peak in at least one of the time points, common peak borders were identified by detecting the local maxima of original peak border density and filtering out all resulting intervals that were > 2 kb. The peak statistics (except the false discovery rate (FDR)) were recalculated as done in MACS version 1.3.7.1 and all intervals where the maximal fold enrichment over the entire data set was < 5 were filtered out.

RNA-seq reads were mapped to the reference genome (hg19) using Bowtie (http://bowtie-bio.sourceforge.net/index.shtml). Then, raw counts and fragments per virtual kb million were obtained for profiling gene expression. Finally, differential expression was calculated with Cuffdiff (http://cole-trapnell-lab.github.io/cufflinks/cuffdiff) based on a t-test over the ratio of normalized counts for the treated and untreated conditions using a cutoff FDR of 0.5%. CTCF ChIP-seq, FAIRE-seq and RNA-seq raw data are available at Gene Expression Omnibus (GEO; www.ncbi.nlm.nih.gov/geo) at GSE69303.

Functional characterization of FAIRE peak loci
More than 80 000 FAIRE-seq peaks were identified at one or more time points. In order to obtain higher accuracy in data integration (see below), we were interested in persistent peaks and filtered out some 18 000 peaks (22% of
all) that were not present at all four time points (0, 2, 24 and 48 h). The summit regions (±100 bp) of these 62 231 peaks were specifically searched for DR3-type sequences and other transcription factor binding motifs by using HOMER (39) with score 7. For a better characterization of the dynamics of the 8979 sites of 1,25(OH)2D3-modulated chromatin accessibility along the four time points, we applied self-organizing maps (SOM). SOM (40) is an unsupervised method that allows the identification of representative profiles of the inspected data. In the context of this contribution, SOM received as input vectors the peak intensity along the four time points, or the fold change (FC), of the aforementioned 8979 peaks. Each component of the vector represented a specific time point, and thus the input space is four-dimensional. SOM classifies these peaks in one of nine possible profiles in a 3 × 3 lattice, which allows a visual inspection of the similarities of the chromatin accessibility as a function of time.

Gene ontology analysis

Gene ontology analysis was conducted through the Ingenuity Pathway Analysis (IPA) software (www.qiagen.com/ingenuity). The genes with a significantly differential expression as calculated by Cuffdiff as well as their logFC were provided as input to the software. Three different data sets were constructed, one for each time point (2.5, 4 and 48 h).

Statistics

From 62 231 FAIRE-seq peaks we computed a significant modulation (ligand-effect) comparing the intensity of the peaks at 2, 24 and 48 h to the co-located peak intensity at time point 0 h. Three replicate experiments were conducted for each of the four time points allowing to compute statistical significance. At each time point, a peak was declared to have a ligand-effect, if the P-value of a paired Student’s t-test applied to the peak intensity of the three replicates at that time point and the three replicates at 0 h was lower than 0.05. We focused on the group of 8979 peaks that presented a significant ligand-effect at one or more time points. Since the distribution of the FC does not fit to a Gaussian distribution (Supplementary Figure S2), a non-parametric analysis of variance by the Kruskal–Wallis test was calculated (H(3) = 780.61, 0.0063). This test offers the basis to reject the null hypothesis that all three data are coming from the same distribution, and at least one sample dominates at least one of the other two. Next, we applied a pairwise Mann–Whitney–Wilcoxon test and found also that the FC was significant in all cases (U(2 h, 24 h) = 87.3, 0.00038; U(2 h, 48 h) = 94.2, 0.00092; U(24 h, 48 h) = 102.8, 0.00024).

RESULTS

Genome-wide changes in chromatin accessibility in response to 1,25(OH)2D3

In three independent experiments THP-1 human monocytes were stimulated for 0, 2, 24 and 48 h with 1,25(OH)2D3 and chromatin was isolated for FAIRE-seq analysis. This approach identified 62 231 genomic regions that showed at all four time points significantly (FDR < 0.1%) accessible chromatin. Importantly, 8979 (14.4%) of these FAIRE peaks were at least at one time point significantly (P < 0.05) modulated by VDR ligand treatment (Supplementary Figure S1A and Table S1). In reference to untreated cells (0 h), 7716 (85.9%) regions were ligand-responsive only at one, 1197 (13.3%) at two and 66 (0.7%) at three time points. Already after 2 h stimulation with 1,25(OH)2D3 3323 genomic regions were affected, of which 2754 (82.8%) showed increased chromatin accessibility, and after 24 h even 4262 (92.9%) out of 4586 ligand-responsive chromatin regions opened up (Supplementary Figure S1B). In contrast, after 48 h only 689 (28.6%) out of 2399 modulated FAIRE peaks were induced by VDR ligand, i.e. at this late time point the majority of the 1,25(OH)2D3-responsive chromatin loci were less accessible than at their basal state. Four representative example regions illustrate the increase of chromatin accessibility (Supplementary Figure S1C) or chromatin closing (Supplementary Figure S1D) in response to VDR ligand treatment.

The global effect of 1,25(OH)2D3 treatment on the number of open chromatin regions and their grade of ligand inducibility at different time points was illustrated by histogram plots (Supplementary Figure S2). The average increase in chromatin accessibility of all significantly ligand-responsive FAIRE peaks after 2 h ligand treatment was already 1.57-fold and increased even to 1.76-fold after 24 h (Figure 1A). In contrast, after 48 h the average induction of the ligand-responsive regions was only 0.89-fold, i.e. they were 1.11-fold repressed compared to the basal state. This suggested that the average chromatin accessibility of the genome in human monocytes (i) increased significantly (P < 0.001) after 2 h stimulation with 1,25(OH)2D3, (ii) was maximal at 24 h and (iii) after 48 h the chromatin was even slightly less accessible than before onset of stimulation. A very comparable result was obtained when all 22 autosomal chromosomes and chromosome X were analyzed individually (Figure 1B), i.e. on each chromosome there was always more 1,25(OH)2D3-mediated chromatin opening after 24 h than after 2 h, while after 48 h chromatin was closed again. A more detailed analysis using 5000 equally sized (621 kb) windows over the whole human genome (3.1 Gb, Figure 1C), indicated 390 regions that (i) contain one to six ligand-responsive FAIRE peaks and (ii) in average were more than 2-fold up- or down-regulated (Supplementary Table S1). Only two (0.5%) of these regions were observed at all three time points and 19 (4.9%) at two time points, i.e. the vast majority (369 (94.6%)) of them were observed only at one time point. In addition, we identified 95 regions that contain six or more ligand-modulated FAIRE peaks (Supplementary Table S1).

In summary, in THP-1 cells 8979 chromatin regions, i.e. approximately every seventh accessible site, were significantly modulated by 1,25(OH)2D3. At earlier time points (2 and 24 h) most regions opened up, but later (48 h) 1,25(OH)2D3 reduced the accessibility of the majority of them.
Figure 1. Genome-wide affects of 1,25(OH)\(_2\)D\(_3\) on chromatin accessibility. In triplicate independent experiments THP-1 cells were stimulated for 0, 2, 24 and 48 h with 100 nM 1,25(OH)\(_2\)D\(_3\) and chromatin was isolated for FAIRE-seq analysis. Box plots display the range of the statistically significant ($P < 0.001$) FC (in log scale) of the FAIRE peaks after 2, 24 and 48 h (A). The indentation of the boxes indicate the average FC at the 3323, 4586 and 2399 regions at time points 2, 24 and 48 h, respectively. A radar plot visualizes for each chromosome the average FC (in linear scale) at time points 2, 24 and 48 h (B). For each of 5000 fragments that cover with an even window size of 621 kb the whole human genome the average FC (in log scale) of the FAIRE peaks within the respective region is displayed for the 1,25(OH)\(_2\)D\(_3\) stimulation times 2, 24 and 48 h (C). For in total 390 regions (marked in Supplementary Table S1) the average chromatin accessibility is more than 2-fold up- or down-regulated in response to 1,25(OH)\(_2\)D\(_3\) treatment (horizontal black lines at logFC 1 and -1).
Properties of 1,25(OH)\(_2\)D\(_3\)-modulated chromatin regions

From 62 231 regions of the human genome that in THP-1 cells were located within accessible chromatin, 12 515 (20.1%) overlapped with TSS regions (Supplementary Figure S3A). Also in the ligand-responsive subsets of 3323 (2 h), 4586 (24 h) and 2399 (48 h) FAIRE peaks the TSS overlap rate was 18.9%, 19.8% and 20.0%, respectively, being very similar in all time points. Next, we performed SOM analysis of the time course response of the absolute chromatin accessibility of all 1,25(OH)\(_2\)D\(_3\)-responsive genomic regions (Supplementary Figure S3B). This approach subdivided the 8979 time course profiles into nine groups (I to IX) containing 69 to 1779 genomic sites, each. The groups III, VI and VII showed a more than 2.5-times higher overlap with TSS regions (50.7–55.8%) than the average (20%). In total, these three groups contained 722 genomic regions that co-locate with 400 TSS regions displaying the overall highest levels of open chromatin. In addition, these regions were characterized by a rapid increase of chromatin accessibility within the first 2 h after ligand treatment. In contrast, groups V and VIII showed with 7.3% and 7.6% a more than 2.5-times lower TSS overlap rate than average, i.e. the 3062 members of these groups may primarily represent enhancer regions. Furthermore, rather low absolute levels of open chromatin and minor ligand-dependent increases of chromatin accessibility characterized these regions. The remaining four groups (I, II, IV and IX) represented 5220 genomic regions that showed average TSS rates (16.4–29.6%) and maximal chromatin accessibility 24 h after onset of stimulation with 1,25(OH)\(_2\)D\(_3\).

Traditional characteristics of 1,25(OH)\(_2\)D\(_3\) signaling are the presence of VDR and the involvement of DR3-type binding sites at the concerned genomic loci (12). Based on VDR ChIP-seq data from THP-1 cells (11,29) only 824 (1.3%) of the 62 231 sites of open chromatin were occupied by VDR (Figure 2A). In parallel, HOMER screening identified DR3-type motifs only within 3626 (5.8%) accessible chromatin regions. Of these 3626 DR3 sites only 170 overlapped with VDR binding loci. This means that at only 20.6% of all 824 VDR-occupied sites the receptor had access to a DR3-type sequence for the formation of heterodimers with RXR. In the subsets of 1,25(OH)\(_2\)D\(_3\)-responsive chromatin regions at 2, 24 and 48 h the VDR occupancy rate increased to 1.7%, 2.6% and 1.9%, respectively, while the DR3 rate raised slightly to 6.7%, 7.0% and 7.2% (Figure 2B). In addition, the DR3 overlap rate of the VDR binding sites at 2, 24 and 48 h ligand treatment increased to 25.5%, 33.9% and 34.8% (in average 29.4%), respectively. However, this means that the majority of all accessible DR3-type sites (87.1–93.7%) were not occupied by VDR and that the receptor used in some 70% of its binding events an alternative mode of complex formation than VDR-RXR heterodimers on DR3-type sequences (Figure 2C).

In order to identify the most interesting subsets of 1,25(OH)\(_2\)D\(_3\)-responsive chromatin sites, we repeated the SOM analysis with the 8979 FC-derived profiles of the time course. This again resulted in nine groups (A to I) with 60 to 2565 members (Figure 2D). Groups A and B showed with 18.3% and 10.2% not only a clearly increased VDR rate but with 30.0% and 11.1%, respectively, their DR3 rate was far higher than average. The 168 members of both groups were characterized by an in average more than 4-fold increase in chromatin accessibility displaying a maximum at 24 h after ligand treatment. The 2204 members of groups E and H showed a similar behavior, but displayed with VDR rates of 3.3% and 3.4% and DR3 rates of 9.4% and 8.6%, respectively, less pronounced enrichments of these key characteristics of 1,25(OH)\(_2\)D\(_3\) signaling. The 119 members of group D showed a prominent induction of chromatin accessibility within the first 24 h, but they did not overlap with any VDR binding site and displayed with 6.7% a rather mean DR3 rate. The 1292 members of group G also showed with 0.3% a very low VDR rate and with 4.6% a slightly reduced DR3 rate, but they represented the only subset of the time course series that displayed continuous chromatin closing at all time points after ligand treatment. Finally, groups C, F and I showed with VDR rates of 1.5%, 1.5% and 1.1% and DR3 rates of 6.6%, 5.3% and 7.0%, respectively, rather average characteristics. The 5196 members of these groups demonstrated less prominent opening of chromatin at time points 2 and 24 h and in contrast to all other time courses their chromatin accessibility returned after 48 h back to the basal state. However, in accordance with the results presented in Figure 1, all time course subsets showed lower chromatin accessibility after 48 h ligand treatment than after 24 h. Chromatin regions representing groups A and B, respectively, are illustrated in Figure 2E (left example: group A, right example: group B).

Taking together, SOM analysis allowed the distinction of different profile subsets within the group of 8979 sites displaying 1,25(OH)\(_2\)D\(_3\)-modulated chromatin accessibility. Regions overlapping with TSS regions showed high levels of chromatin accessibility and were quickly inducible, while enhancer-type regions were less prominent and responded delayed. Furthermore, on regions with a high VDR occupancy rate chromatin was opening prominently, whereas low VDR rates mostly indicated ligand-dependent chromatin closing.

1,25(OH)\(_2\)D\(_3\)-inducible CTCF binding

Only 608 (6.8%) of the 8979 ligand-modulated chromatin sites contained a DR3-type motif and even only 170 of all sites (1.9%) were occupied by VDR (see Figure 2B). In fact, only 50 (29.4%) of these 170 VDR binding sites contained a DR3-type sequence. This suggested that (i) the very most of the effects of 1,25(OH)\(_2\)D\(_3\) on chromatin opening or closing could not be explained by VDR binding to the same site and (ii) even in cases where VDR was occupying such a ligand-dependent chromatin locus there must be further mechanistic explanations for the effects of 1,25(OH)\(_2\)D\(_3\) in addition to VDR-RXR heterodimers binding to DR3-type sequences. Therefore, we used HOMER to screen the sequences below the summits (±100 bp) of all 8979 ligand-modulated FAIRE peaks for transcription factor binding motifs. The 10 most significantly enriched motifs were binding sites for the insulator proteins CTCF and BORIS, the pioneer factor PU.1 and several members of the ETS transcription factor family (Supplementary Figure S4A). This suggested that these transcription factors could occupy 1,25(OH)\(_2\)D\(_3\)-modulated, accessible chromatin sites
Figure 2. Overlap of FAIRE peaks with VDR binding sites and DR3-type motifs. Stacked column charts display the overlap of all 62,231 regions of open chromatin (A) and of the 8,979 1,25(OH)2D3-responsive FAIRE peaks (B) with DR3-type motifs, sites of VDR binding or both. Schematic representation of the understanding of three mechanistic classes of VDR interacting with genomic DNA that are differentiated in A and B (C). For all 8,979 significantly 1,25(OH)2D3-responsive genomic regions the time course profiles of the FC of their chromatin accessibility was analyzed in a 3 × 3 SOM lattice. Each of the nine resulting groups (A to I, marked in Supplementary Table S1) is represented by a graph displaying the average FC (in log scale) of the chromatin accessibility over time. Groups A and B significantly overlap with more VDR and DR3 sites than the reference (see A), while two other groups have no (group D) or a very low number (group G) of overlapping VDR sites. The IGV browser was used to visualize representative examples of genomic loci (±2 kb of the peak summits) monitoring 1,25(OH)2D3-dependent chromatin opening at VDR binding sites in the absence and presence of a DR3-type binding site (E). The peak tracks display data from the triplicate FAIRE-seq time course experiment (gray for input lanes and turquoise for 1,25(OH)2D3 (1,25D) treatments for the indicated time periods) and a VDR ChIP-seq experiment (red (29)). Gene structures are shown in blue.
and may contribute to the mechanisms of action of the VDR ligand on the epigenome.

We focused on the top ranking transcription factor CTCF and performed ChIP-seq analysis with chromatin from THP-1 cells that were stimulated for 24 h with 1,25(OH)_{2}D_{3} in comparison to vehicle-treated cells. In total, we found 71,993 CTCF peaks, but from the 46,691 and 46,922 CTCF binding events in presence and absence of 1,25(OH)_{2}D_{3}, respectively, only 21,690 (≈46%) were common (Figure 3A and Supplementary Table S2). 52.2% of all 71,993 CTCF peaks in THP-1 cells were found at the same position as in the ENCODE tier 1 cell line K562 (33), for which we harmonized re-analysis of the CTCF ChIP-seq data indicated 98,515 peaks (Supplementary Figure S4B). In the absence of ligand 65.9% of the 46,920 CTCF sites overlapped with K562 data and for the 21,690 common CTCF sites the rate was even 91.7%.

From the 8979 ligand-dependent FAIRE peaks 1249 (13.9%) overlapped with these common CTCF sites (Supplementary Table S1). These 1249 CTCF binding sites colocalized with 234 (18.7%) TSS regions and 21 (1.7%) VDR sites, i.e. these were average percentages (compare Supplementary Figure S3 and Figure 2) not indicating any enrichment for these types of sites. Interestingly, 5915 common CTCF sites were 2-fold or more induced or reduced after 24 h treatment with 1,25(OH)_{2}D_{3} (Figure 3B and Supplementary Table S2). Moreover, 1752 common CTCF sites overlapped with ligand-dependent FAIRE peaks, 597 of which were also sensitive to 1,25(OH)_{2}D_{3} (Figure 3B and Supplementary Table S2). Four representative examples illustrate the co-locating of ligand-modulated FAIRE peaks with sites of either up- or down-regulated CTCF association after treatment of the cells with 1,25(OH)_{2}D_{3} (Figure 3C).

From the 5915 ligand-dependent CTCF sites, 609 (10.3%) overlapped with a TSS region, which was a 2.3-fold increased percentage compared to all common CTCF sites (4.4%, Supplementary Table S2). Similarly, the percentage of overlap with 1,25(OH)_{2}D_{3}-modulated FAIRE peaks was increased 2.5-fold from 4.0% for all 71,993 CTCF sites to 8.8% for the 21,690 common CTCF sites. In contrast, comparing both groups, there was no significant increase for the rate of DR3-type sequence (7.4 versus 8.6%). Moreover, a HOMER analysis for enriched motifs below the CTCF peaks did not provide any evidence for an enrichment of VDR binding events. In parallel, there were only 37 (0.6%) ligand-dependent CTCF sites overlapping with a VDR binding site (Supplementary Table S2). Compared to the transcription factors ETS1, GABPA and NR2F2 (also called COUP-TF2), which in the ENCODE tier 1 cell line K562 (33) showed overlap rates with CTCF of 10.1%, 7.7% and 15.1%, respectively (Supplementary Figure S4C), VDR showed a very low rate. Also from the perspective of the rather low total number of 1142 VDR binding sites in THP-1 cells 37 VDR peaks overlapping with CTCF represent only 3.2% of all, which was even 4-fold lower than respective rates for the nuclear receptor NR4A1 (also called NUR77). This indicates that VDR is not directly involved in the 1,25(OH)_{2}D_{3}-dependent association or dissociation of CTCF with its binding sites, i.e. most likely VDR and CTCF are not part of the same nuclear protein complex.

The functional impact of CTCF for the regulation of 1,25(OH)_{2}D_{3} target genes by the ligand was tested by silencing CTCF mRNA expression (Figure 3D and Supplementary Figure S5). After the transfection of THP-1 cells with DsiRNA oligonucleotides against CTCF, the cells were treated with 1,25(OH)_{2}D_{3} for 24 h. The knockdown of CTCF mRNA expression was 70% and 67% for solvent- and ligand-treated cells, respectively, compared to cells transfected with a negative control DsiRNA (NC1). The additionally performed mock-transfected controls differed in average by 9.6% from the NC1 control (standard deviation 9.6%, data not shown). From in total 20 tested 1,25(OH)_{2}D_{3} target genes, six genes showed significant modulations of either their relative mRNA expression or their ligand-inducibility due to the CTCF silencing. We observed both increased (CAMP and ALOX5) and decreased (NOD2, ZFP36, THBD and MYC) mRNA levels and the changes ranged from –34% (NOD2) to +37% (CAMP). Likewise, the induction of the 1,25(OH)_{2}D_{3} target genes by the ligand was both positively and negatively modulated by the CTCF silencing. The CAMP gene showed with 33% the strongest attenuation of its ligand response, while the induction of the ALOX5 gene increased by even 69%. In contrast, the ZFP36 mRNA was reduced to a similar extent for both solvent- and ligand-treated cells (22% and 24%, respectively), thus its inducibility by 1,25(OH)_{2}D_{3} did not change. Interestingly, at the ligand-dependent CTCF sites (Figure 3B) closest to the TSS regions of the six genes, which showed sensitivity to CTCF knock-down, CTCF association was more prominently induced or reduced by 1,25(OH)_{2}D_{3} than at those of the remaining 14 genes that were not affected by CTCF silencing. An exception to this observation is the CTCF site 86 kb downstream of the TMEM37 gene TSS, which is nearly 5-fold induced (Supplementary Table S2). However, it is possible that it lies outside of the chromatin domain of the TMEM37 gene and therefore does not influence the gene’s expression.

In summary, ligand-modulated chromatin sites were not only highly significantly enriched for CTCF binding motifs, but 13.9% of these 8979 loci also bound the transcription factor. Surprisingly, the association of CTCF with some of these genomic sites could be up- or down-regulated by 1,25(OH)_{2}D_{3}, which provided an additional mechanistic explanation for the epigenomic effects of the VDR ligand. VDR is not directly involved in the ligand-modulated binding of CTCF, though. The actions of CTCF affected 30% of all tested 1,25(OH)_{2}D_{3} target genes.

Transcriptome-wide effects of 1,25(OH)_{2}D_{3} stimulation

Next, we performed triplicate RNA-seq experiments with THP-1 cells that were for 2.5, 4 and 24 h stimulated with 1,25(OH)_{2}D_{3} (Supplementary Table S3). From 23,622 human genes 9220 (39.0%) were not expressed, but from the 14,402 expressed genes 1246 (8.7%) were at least at one time point significantly (P ≤ 0.0001) modulated by the VDR ligand (Supplementary Figure S6A). Of note, 38 genes, such as the well-known VDR target gene CYP24A1 (41), were at basal level below the expression threshold (0.3) but passed it after stimulation with 1,25(OH)_{2}D_{3} (gray in Supplementary Table S3). Therefore, in the following we refer to in to-
Figure 3. Overlap of 1,25(OH)2D3-modulated CTCF binding with open chromatin. THP-1 cells were treated for 24 h with 1,25(OH)2D3 (1,25D) or vehicle (EtOH) and ChIP-seq was performed for CTCF binding. A Venn diagram illustrates the number of genomic regions on which CTCF binding sites in 1,25(OH)2D3-treated cells (left) overlap with that found in vehicle-treated cells (right) (A). The center indicates 21,690 genomic regions, on which both dataset overlaps. From these 21,690 common CTCF sites 5,915 are more than 2-fold induced or reduced by 1,25(OH)2D3; 597 of the 1,752 ligand-dependent FAIRE peaks overlap with this subset of ligand-dependent CTCF loci (B). The IGV browser was used to visualize representative examples of genomic loci (±2 kb of the peak summits) monitoring 1,25(OH)2D3-dependent chromatin opening at sites where the ligand also modulates CTCF binding (C). The peak tracks display data from the triplicate FAIRE-seq time course experiment (gray for input lanes and turquoise for 1,25(OH)2D3 treatments for the indicated time periods), from the CTCF ChIP-seq experiment in absence and presence of ligand (blue) and from a VDR ChIP-seq experiment (red (29)). Gene structures are shown in blue. THP-1 cells were transfected with a mixture of three DsiRNA oligonucleotides directed against CTCF and after 24 h they were stimulated with solvent (EtOH) or 100 nM 1,25(OH)2D3 for 24 h (D). qPCR was performed in order to determine changes in the expression of CTCF and four 1,25(OH)2D3 target genes. Columns show the average of three independent experiments, each performed in triplicate individual transfections, and bars indicate standard deviations. Two-tailed Student’s t-tests were performed to determine the significance of (i) the mRNA changes by the CTCF knockdown (black) in reference to control DsiRNA-transfected cells, (ii) the modulation of the ligand effect (blue) in reference to control DsiRNA-transfected cells and (iii) the induction by 1,25(OH)2D3 (red) in reference to solvent-treated cells (*P < 0.05; **P < 0.01; ***P < 0.001).
tal 1284 1,25(OH)_{2}D_{3} target genes. Forty-six of these genes were already responding 2.5 h after onset of 1,25(OH)_{2}D_{3} stimulation, 288 after 4 h and 1204 after 24 h (Figure 4A). The expression of only 45 genes was affected by the VDR ligand at all three time points (Supplementary Table S3). Interestingly, after 2.5 h stimulation all 46 differentially expressed genes were up-regulated, after 4 h 170 (59.0%) out of 288 and after 24 h 663 (55.1%) out of 1204 genes (Figure 4B, Supplementary Figure S6B and S6C). This suggests that, like in the epigenome analysis (Figure 1), at later time points the 1,25(OH)_{2}D_{3}-dependently down-regulated part of the transcriptome had a far larger impact. A gene ontology analysis using IPA software of the 1,25(OH)_{2}D_{3}-regulated genes at each of the three time points indicated that the top-ranking biofunction represented by the respective genes shifts from ‘anti-microbial response’ at 2.5 h to ‘endocrine system disorder’ at 4 h and ‘connective tissue disorders’ at 24 h (Figure 4C and Supplementary Table S4). The second ranking term changed from ‘gene expression’ to ‘renal and urological disease’ and ‘hepatic system disease’.

Taken together, the 1,25(OH)_{2}D_{3}-modulated transcriptome of THP-1 cells comprised 1284 genes, of which 681 (53.0%) were up-regulated and 603 (47.0%) down-regulated. The vast majority of the genes (995 (77.5%)) responded with a delay of 24 h to stimulation with 1,25(OH)_{2}D_{3}. Interestingly, during the time course of 1,25(OH)_{2}D_{3} treatment the proportion of down-regulated genes increased from 0 to 44.9% and in parallel the main biofunction of the genes shifted from anti-microbial response to connective tissue disorder.

Integration of epigenome- and transcriptome-wide responses to 1,25(OH)_{2}D_{3}

After testing several models for epigenome and transcriptome data integration, such as Pearson product-moment between pairs of variables presented in Supplementary Tables S1 and S3, that delivered only low statistical correlations, we followed a machine learning perspective (42). This approach provided the clearest, although not univariate, prediction of the expression state of the TSS regions of the 165 selected genes (Supplementary Figure S7D and S7E). Importantly, for 138 out of 165 genes their inducibility after 24 h 1,25(OH)_{2}D_{3} treatment could be accurately (FDR < 5%) predicted by the six most relevant parameters listed in Supplementary Figure S7C (Supplementary Figure S7F).

Two representative examples of the 165 selected 1,25(OH)_{2}D_{3} target genes were presented in Figure 5. CTCF Chi-PET data from K562 cells indicated that the HTT gene is co-located in the same chromatin domain with the genes GRK4 and HTT-AS (Figure 5A). However, the HTT gene was more prominently expressed than the two other genes and the only gene that after 4 and 24 h was significantly induced by 1,25(OH)_{2}D_{3} (Figure 5B). The HTT gene is widely expressed and is required for normal development, but in cases when it acquires additional unstable trinucleotide repeats it can cause the neurodegenerative disorder Huntington’s disease, which is characterized by loss of striatal neurons (43). At two loci, 450 and 4985 bp downstream of the HTT TSS, VDR binding sites overlapped with ligand-inducible FAIRE sites (Figure 5A). In contrast, the chromatin domain of the NOD2 gene did not contain any VDR binding site, but the chromatin accessibility at its TSS region was induced after treatment with 1,25(OH)_{2}D_{3} (Figure 5C). The NOD2 protein acts as an intracellular pattern recognition receptor binding bacterial peptidoglycans and stimulates the innate immune response (44). The NOD2 gene is co-located with the genes NDK1, SNX20 and CYLD, but only NOD2 responded significantly to 1,25(OH)_{2}D_{3} after 24 h stimulation (Figure 5D).

In summary, the integration of epigenomic and transcriptomic data identified 165 1,25(OH)_{2}D_{3} target genes, the expression of which could be predicted primarily from epigenomic data of their genomic loci. For example, the 1,25(OH)_{2}D_{3} target gene HTT is regulated directly via VDR binding close to its TSS region, while 1,25(OH)_{2}D_{3}-dependently opening chromatin at the NOD2 TSS contributes to the regulation of the gene.

DISCUSSION

This study describes the epigenome-wide effects of 1,25(OH)_{2}D_{3} on chromatin accessibility of human monocytes. During a time course of 48 h the natural VDR ligand significantly affected the opening and closing of chromatin at nearly 9000 sites, which was one in seven of all accessible genomic regions in this cellular model. Interestingly, both in number of affected loci as well as in average strength of the chromatin opening, the 24 h stimulation with 1,25(OH)_{2}D_{3} was more prominent than the treatment for 2 h or 48 h. Epigenomic effects are assumed to be mediated primarily by chromatin modifying enzymes that act quickly (45). In fact, there was a subset of 1,25(OH)_{2}D_{3}-modulated chromatin sites, in particular at TSS regions, that already showed maximal accessibility 2 h after onset of stimulation. However, the majority of the sites reached their maximal levels of chromatin open-
Figure 4. Transcriptome-wide and physiological effects of 1,25(OH)2D3 treatment. In triplicate independent experiments THP-1 cells were treated for 2.5, 4 and 24 h with vehicle (EtOH) or 100 nM 1,25(OH)2D3 and RNA-seq analysis was performed. A Venn diagram illustrates the overlap of genes being significantly (P < 0.001) regulated by 1,25(OH)2D3 at the respective time points (A). Stacked bar charts indicate the number of significantly up- and down-regulated genes 2.5, 4 and 24 after stimulation with 1,25(OH)2D3 (B). Of note, after 2.5 h there is no down-regulated gene. Gene ontology analysis of the significantly regulated genes for their involvement in biofunctions (Supplementary Table S4) provided for each time point a different ranking of the three most significant functions (C).
Figure 5. Large-scale genomic view of the HTT and NOD2 gene loci. The IGV browser was used to show the whole chromosomal domain (defined by K562 CTCF ChIA-PET data (33)) of the 1,25(OH)2D3 (1,25D) target genes HTT (A) and NOD2 (C). The peak tracks display FAIRE-seq data (turquoise), VDR ChIP-seq data (red (29)) and CTCF ChIP-seq data (blue) before and after 1,25(OH)2D3 treatment (all from THP-1 cells). Gene structures are shown in blue. The absolute expression of all genes at the two loci 2.5, 4 and 24 h after stimulation with 1,25(OH)2D3 was determined by RNA-seq (B and D). Columns represent the means of three independent experiments and the bars indicate standard deviations. Two-tailed Student’s t-tests were performed to determine the significance of the mRNA induction by 1,25(OH)2D3 in reference to solvent-treated cells (*P < 0.05; **P < 0.01).
ing delayed after 24 h. This suggests that the process of 1,25(OH)_{2}D_{3}-mediated chromatin opening involves multiple steps and that the effects observed 24 h after onset of stimulation are not primary responses but rather secondary effects on chromatin accessibility. The observation that most 1,25(OH)_{2}D_{3}-sensitive chromatin sites return after 48 h back to basal levels indicates that the epigenome-wide effects of the stimulation with VDR ligand are transient. This suggests that for a persistent effect a constant vitamin D supplementation is required.

Since VDR is the only nuclear protein that binds 1,25(OH)_{2}D_{3} with high affinity (46), the nuclear receptor must be central to the mechanisms initiating the epigenome-wide effects of its natural ligand. However, in accordance with previous genome-wide studies on vitamin D signaling (11,29) only a very low number of 1,25(OH)_{2}D_{3}-sensitive chromatin sites (1.9%) contain a DR3-type binding site that is essential for the well-understood binding of VDR-RXR heterodimers. Even the small subset of 168 rapidly responding genomic regions with high VDR occupancy have a DR3 rate of only 17.9%, which is in the same order as obtained in the VDR ChIP-seq data set meta-analysis (11). This suggests that the canonical model of VDR-RXR heterodimers binding to DR3-type sequences applies only for a subset of the primary responding chromatin sites and already in the majority of the fast responding sites VDR seems to be involved in different types of complexes. In some of these complexes VDR may be partnered with other nuclear proteins than RXR, in order to contact DNA. As suggested already a longer time ago (47,48), these could be other members of the nuclear receptor superfamily. However, motif screenings indicated that primarily rather ubiquitously binding transcription factors of the ETS family co-locate with genomic VDR binding sites (11). It has not yet been shown that VDR forms any heterodimeric complexes with these proteins, but they may act as pioneer factors supporting VDR in opening chromatin regions. In parallel, it is possible that at a number of sites VDR does not associate directly to genomic DNA but may bind ‘piggyback’ on DNA-associated transcription factors (12).

In previous studies we already described that several members of the HDAC family are primary 1,25(OH)_{2}D_{3} target genes (49) and that HDAC proteins and their small molecule inhibitors are involved in the control of VDR target genes (28,50). In this study, we demonstrated that CTCF co-locates with some 14% of all 1,25(OH)_{2}D_{3}-sensitive chromatin sites. This provides an additional mechanistic explanation for the epigenome-wide actions of the VDR ligand. However, we could not find any evidence that VDR is directly involved in the association of CTCF with DNA, i.e. VDR and CTCF appear not to be part of the same nuclear protein complex. Interestingly, silencing of CTCF affects also the action of other transcription factors, including other nuclear receptors. In fact, CTCF seems to co-locate with a number of these regulatory proteins in the same nuclear complex, i.e. CTCF may use additional mechanisms for the interference with the respective signaling pathways.

Within the first 24 h of stimulation with 1,25(OH)_{2}D_{3} at some 90% of the 7193 genomic loci, which were accessible after 2 h ligand treatment, chromatin accessibility increased while in parallel the RNA level of 681 genes was up-regulated and of 603 genes down-regulated. The more than 5-times higher number of 1,25(OH)_{2}D_{3}-sensitive chromatin loci that bind CTCF showed only low percentages of VDR occupancy and TSS overlap. CTCF tends to associate preferentially with loci where chromatin accessibility was modulated early after ligand treatment and in particular then the association of CTCF with genomic DNA increased. This suggests that CTCF may be involved in the early epigenomic effects of 1,25(OH)_{2}D_{3}. We assume that CTCF affects also the action of other transcription factors, including other nuclear receptors. In fact, CTCF seems to co-localize with a number of these regulatory proteins in the same nuclear complex, i.e. CTCF may use additional mechanisms for the interference with the respective signaling pathways.

The novel finding that the associations of CTCF with its genomic targets were modulated by 1,25(OH)_{2}D_{3} suggests that the CTCF-dependent organization of chromatin can be modulated by vitamin D. This could affect the size and function of chromosomal domains, i.e. such a mechanism has wider consequences than a 1,25(OH)_{2}D_{3}-dependent action at any specific chromatin site. The 1,249 1,25(OH)_{2}D_{3}-sensitive chromatin loci that bind CTCF showed only low percentages of VDR occupancy and TSS overlap. CTCF tends to associate preferentially with loci where chromatin accessibility was modulated early after ligand treatment and in particular then the association of CTCF with genomic DNA increased. This suggests that CTCF may be involved in the early epigenomic effects of 1,25(OH)_{2}D_{3}. We assume that CTCF affects also the action of other transcription factors, including other nuclear receptors. In fact, CTCF seems to co-localize with a number of these regulatory proteins in the same nuclear complex, i.e. CTCF may use additional mechanisms for the interference with the respective signaling pathways.
to be found within their regulatory region but instead effects of 1,25(OH)2D3 on the accessibility of the respective chromatin regions are detected. This emphasizes the impact of epigenome-wide data collection, such as chromatin accessibility as well as binding of VDR, CTCF and other nuclear proteins, over pure RNA quantifications, in order to understand the mechanisms of regulation of these 1,25(OH)2D3 target genes. Moreover, even in cases, where the regulatory mechanism is rather obvious, such as for the two VDR binding sites close to the TSS of the HTT gene, the 1,25(OH)2D3-dependent modulation of chromatin accessibility at these sites provides a confirmation of their impact for the control of the respective genomic region and the genes located at these sites. Nevertheless, the fact that the integration of epigenomic and transcriptomic data can so far only correctly predict the action of 165 of 1284 1,25(OH)2D3 target genes, indicates that we are still missing insight into the complexity of vitamin D signaling, i.e. additional data need to be assessed and integrated into the model. The Roadmap Epigenomics Consortium (55) has recently presented the integration of 111 human epigenomes. They based their findings on an impressive set of epigenome-wide data, including different histone modifications and DNA methylation, from a large variety of tissues, but could only provide a model for the basal state of the human epigenome. This means that for the different challenges of the human body, such as a supplementation with vitamin D, additional data need to be collected.

The shift of the main physiological function ‘anti-microbial response’ represented by the primary vitamin D targets to ‘connective tissue disorders’ by secondary targets characterizes and reflects the functional profile of VDR and its ligand 1,25(OH)2D3. The primary function of ancestral, probably orphan nuclear receptors was the control of metabolism (56) and nowadays still many members of the nuclear receptor superfamily have this property. During evolution the superfamily diversified and the different members not only acquired high affinity ligand binding but also additional roles in physiology. As a side effect of evolutionary and developmental pressures, VDR became a key controller of innate immunity but also to affect cellular growth in many of its target tissues (57).

In conclusion, in this study we demonstrated the dynamic nature of the epigenetic landscape of human monocytes in response to stimulation with 1,25(OH)2D3. The characterization of the epigenomic changes within the chromosomal domain of 1,25(OH)2D3 target genes contributes to the understanding of the involved mechanisms. The variety of observed effects of 1,25(OH)2D3 is an indication for a larger number of different mechanisms comprising many nuclear proteins, such as the ones demonstrated here for CTCF. The results of this study present mechanisms and factors that significantly extend the canonical model of 1,25(OH)2D3 signaling.

**ACKNOWLEDGEMENTS**

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

**FUNDING**

Funding for open access charge: Academy of Finland and Juelius Foundation.

**Conflict of interest statement.** None declared.

**REFERENCES**

1. Verstuyf,A., Carmeliet,G., Bouillon,R. and Mathieu,C. (2010) Vitamin D: a pleiotropic hormone. *Kidney Int.*, 78, 140–145.

2. Feldman,D., Krishnan,A.V., Swami,S., Giovannucci,E. and Feldman,B.J. (2014) The role of vitamin D in reducing cancer risk and progression. *Nat. Rev. Cancer*, 14, 342–357.

3. Carlberg,C. (2014) The physiology of vitamin D far more than calcium and bone. *Front. Physiol.*, 5, 335.

4. Carlberg,C. (2014) Genome-wide (over)view on the actions of vitamin D. *Front. Physiol.*, 5, 167.

5. Campbell,M.J. (2014) Vitamin D and the RNA transcriptome: more than mRNA regulation. *Front. Physiol.*, 5, 181.

6. Evans,R.M. and Mangelsdorf,D.J. (2014) Nuclear Receptors, RXR, and the Big Cell. *Cell*, 157, 255–266.

7. Chawla,A., Repa,J.J., Evans,R.M. and Mangelsdorf,D.J. (2001) Nuclear receptors and lipid physiology: opening the X-files. *Science*, 294, 1866–1870.

8. Mangelsdorf,D.J., Thummel,C., Beato,M., Herrlich,P., Schütz,G., Umesono,K., Blumberg,B., Kastner,P., Mark,M., Chambon,P. et al. (1995) The nuclear receptor superfamily: the second decade. *Cell*, 83, 835–839.

9. Pike,J.W., Meyer,M.B., Wang,W., Kim,S., Zella,L.A., Fretz,J.A., Yamazaki,M. and Shevde,N.K. (2007) Perspectives on mechanisms of gene regulation by 1,25-dihydroxyvitamin D3 and its receptor. *J. Steroid Biochem. Mol. Biol.*, 103, 389–395.

10. Haussler,M.R., Haussler,C.A., Bartik,L., Whitfield,G.K., Hsieh,J.C., Slater,S. and Jurutka,P.W. (2008) Vitamin D receptor: molecular signaling and actions of nutritional ligands in disease prevention. *Nutr. Rev.*, 66, S98–S112.

11. Tuoresmäki,P., Väisänen,S., Neme,A., Heikkinen,S. and Carlberg,C. (2014) Patterns of genome-wide VDR locations. *PLoS One*, 9, e96105.

12. Carlberg,C. and Campbell,M.J. (2013) Vitamin D receptor signaling mechanisms: Integrated actions of a well-defined transcription factor. *Steroids*, 78, 127–136.

13. Rastinejad,F., Ollendorff,V. and Polikarpov,I. (2015) Nuclear receptor full-length architectures: confronting myth and illusion with high resolution. *Trends Biochem. Sci.*, 40, 16–24.

14. Perissi,V. and Rosenfeld,M.G. (2005) Controlling nuclear receptors: mechanisms and actions of cofactor cycles. *Nat. Rev. Mol. Cell Biol.*, 6, 542–554.

15. Carlberg,C. and Molnár,F. (2015) Vitamin D receptor signaling and its therapeutic implications: Genome-wide and structural view. *Can. J. Physiol. Pharmacol.*, 93, 311–318.

16. Helin,K. and Dhanak,D. (2013) Chromatin proteins and modifications as drug targets. *Nature*, 502, 480–488.

17. Smith,Z.D. and Meissner,A. (2013) DNA methylation: roles in mammalian development. *Nat. Rev. Genet.*, 14, 204–220.

18. Badeaux,A.I. and Shi,Y. (2013) Emerging roles for chromatin as a signal integration and storage platform. *Nat. Rev. Mol. Cell Biol.*, 14, 211–224.

19. Giresi,P.G., Kim,J., McDaniel,R.M., Iyer,V.R. and Lieb,J.D. (2007) FAIRE (Formaldehyde-Assisted Isolation of Regulatory Elements) isolates active regulatory elements from human chromatin. *Genome Res.*, 17, 877–885.

20. Van Bortle,K. and Corces,V.G. (2013) The role of chromatin insulators in nuclear architecture and genome function. *Curr. Opin. Genet. Dev.*, 23, 212–218.

**SUPPLEMENTARY DATA**

Supplementary Data are available at NAR Online.
21. Sexton, T. and Cavalli, G. (2015) The role of chromosome domains in shaping the functional genome. *Cell*, 160, 1049–1059.
22. Li, Y., Huang, W., Niu, L., Umbach, D.M., Covo, S. and Li, L. (2013) Characterization of constitutive CTCF/cohesin loci: a possible role in establishing topological domains in mammalian genomes. *BMC Genomics*, 14, 553.
23. Phillips, J.E. and Corces, V.G. (2009) CTCF: master weaver of the genome. *Cell*, 137, 1194–1211.
24. Handoko, L., Xu, H., Li, G., Nga, C.Y., Chew, E., Schnapp, M., Lee, C.W., Ye, C., Ping, J.L., Mulawadi, F. et al. (2011) CTCF-mediated functional chromatin interaction in pluripotent cells. *Nat. Genet.*, 43, 630–638.
25. Tschiya, S., Yamabe, M., Yamaguchi, Y., Kobayashi, Y., Konno, T. and Tada, K. (1980) Establishment and characterization of a human acute monocytic leukemia cell line (THP-1). *Int. J. Cancer*, 26, 171–176.
26. Matilainen, J.M., Husso, T., Toropainen, S., Seuter, S., Turunen, M.P., Gynter, P., Ylä-Herttuala, S., Carlberg, C. and Väisänen, S. (2010) Primary effect of 1α,25(OH)2D3 on IL-10 expression in monocytes is short-term down-regulation. *Biochem. Biophys. Acta*, 1803, 1276–1286.
27. Gynter, P., Toropainen, S., Matilainen, J.M., Seuter, S., Carlberg, C. and Väisänen, S. (2011) Mechanism of 1α,25-dihydroxyvitamin D3-dependent repression of interleukin-12B. *Biochem. Biophys. Acta*, 1813, 810–818.
28. Seuter, S., Heikkinen, S. and Carlberg, C. (2013) Chromatin acetylation at transcription start sites and vitamin D receptor binding regions relates to effects of 1α,25-dihydroxyvitamin D3 and histone deacetylase inhibitors on gene expression. *Nucleic Acids Res.*, 41, 110–124.
29. Heikkinen, S., Väisänen, S., Pehkonen, P., Seuter, S., Benes, V. and Carlberg, C. (2011) Nuclear hormone 1α,25-dihydroxyvitamin D3 elicits a genome-wide shift in the locations of VDR chromatin occupancy. *Nucleic Acids Res.*, 39, 9181–9193.
30. Seuter, S., Rynnänen, J. and Carlberg, C. (2014) The ASAP2 gene is a primary target of 1,25-dihydroxyvitamin D in human monocytes and macrophages. *J. Steroid Biochem. Mol. Biol.*, 144, 12–18.
31. Zhang, J., Poh, H.M., Peh, S.Q., Sia, Y.Y., Li, G., Mulawadi, F.H., Goh, Y., Fulwood, M.J., Sung, W.K., Ruan, X. et al. (2012) ChIA-PET analysis of transcriptional chromatin interactions. *Methods*, 58, 289–299.
32. Giresi, P.G. and Lieb, J.D. (2009) Isolation of active regulatory elements from eukaryotic chromatin using FAIRE (Formaldehyde Assisted Isolation of Regulatory Elements). *Methods*, 48, 233–239.
33. ENCODE-Project-Consortium. Bernstein, B.E., Birney, E., Dunham, I., Green, E.D., Gunter, C. and Snyder, M. (2012) An integrated encyclopedia of DNA elements in the human genome. *Nature*, 489, 57–74.
34. Langmead, B., Trapnell, C., Pop, M. and Salzberg, S.L. (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.*, 10, R25.
35. Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G. and Durbin, R. (2009) The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, 25, 2078–2079.
36. Robinson, J.T., Thorvaldsdottir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G. and Mesirov, J.P. (2011) Integrative genomics viewer. *Nat. Biotechnol.*, 29, 24–26.
37. Zhang, Y., Liu, T., Meyer, C.A., Eckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W. et al. (2008) Model-based analysis of ChIP-Seq (MACS). *Genome Biol.*, 9, R137.
38. Rashid, N.U., Giresi, P., Ibrahim, J.G., Sun, W. and Lieb, J.D. (2011) ZINBA integrates local covariates with DNA-seq data to identify broad and narrow regions of enrichment, even within amplified genomic regions. *Genome Biol.*, 12, R67.
39. Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H. and Glass, C.K. (2010) Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol. Cell.*, 38, 576–589.
40. Kohonen, T. (1995) Self-organizing maps. *Series Inform. Sci.*, 30.
41. Väisänen, S., Dunlop, T.W., Sinkkonen, L., Frank, C. and Carlberg, C. (2005) Spatio-temporal activation of chromatin on the human CYP24 gene promoter in the presence of 1α,25-Dihydroxyvitamin D3. *J. Mol. Biol.*, 350, 65–77.
42. Cheng, C., Yan, K.K., Yip, K.Y., Rozowsky, J., Alexander, R., Shou, C. and Gerstein, M. (2011) A statistical framework for modeling gene expression using chromatin features and application to modENCODE datasets. *Genome Biol.*, 12, R15.
43. Walker, F.O. (2007) Huntington's disease. *Lancet*, 369, 218–228.
44. Kufer, T.A., Banks, D.J. and Philipott, D.J. (2006) Innate immune sensing of microbes by Nod proteins. *Ann. N. Y. Acad. Sci.*, 1072, 19–27.
45. Zentner, G.E. and Henikoff, S. (2013) Regulation of nucleosome dynamics by histone modifications. *Nat. Struct. Mol. Biol.*, 20, 259–266.
46. Carlberg, C. and Molnár, F. (2012) Current status of vitamin D signaling and its therapeutic applications. *Curr. Top. Med. Chem.*, 12, 528–547.
47. Schröder, M., Müller, K.M., Nayeri, S., Kahlen, J.P. and Carlberg, C. (1994) VDR-T3 R receptor heterodimer polarity directs ligand sensitivity of transactivation. *Nature*, 370, 382–386.
48. Schröder, M., Bendik, J., Becker-Andre, M. and Carlberg, C. (1993) Interaction between retinoic acid and vitamin D signaling pathways. *J. Biol. Chem.*, 268, 17830–17836.
49. Malinen, M., Saramäki, Å., Ropponen, A., Degenhardt, T., Väisänen, S. and Carlberg, C. (2008) Distinct HDACs regulate the transcriptional response of human cyclin-dependent kinase inhibitor genes to trichostatin A and 1α,25-dihydroxyvitamin D3. *Nucleic Acids Res.*, 36, 121–132.
50. Malinen, M., Rynnänen, J., Heinäniemi, M., Väisänen, S. and Carlberg, C. (2011) Cyclical regulation of the insulin-like growth factor binding protein 3 gene in response to 1α,25-dihydroxyvitamin D3. *Nucleic Acids Res.*, 39, 502–512.
51. Ren, L., Wang, Y., Shi, M., Wang, X., Yang, Z. and Zhao, Z. (2012) CTCF mediates the cell-type-specific spatial organization of the Kcnq5 locus and the local gene regulation. *PLoS One*, 7, e31416.
52. Yang, F., Deng, X., Ma, W., Berleth, J.B., Rabia, N., Wei, G., Moore, J.M., Filippova, G.N., Xu, J., Liu, Y. et al. (2015) The IncRNA Firre anchors the inactive X chromosome to the nucleolus by binding CTCF and maintains H3K27me3 methylation. *Genome Biol.*, 16, 52.
53. Rybakova, K.N., Bruggeman, F.J., Tomaszewska, A., Mone, M.J., Carlberg, C. and Westerhoff, H.V. (2015) Multiplex eukaryotic transcription (in)activation: timing, bursting and cycling of a ratchet clock mechanism. *PLoS Comput. Biol.*, 11, e1004236.
54. Turunen, M.M., Dunlop, T.W., Carlberg, C. and Väisänen, S. (2007) Selective use of multiple vitamin D response elements underlies the 1α,25-dihydroxyvitamin D3-mediated negative regulation of the human CYP2B1 gene. *Nucleic Acids Res.*, 35, 2734–2747.
55. Roadmap Epigenomics, C., Kundaje, A., Meuleman, W., Ernst, J., Bilenky, M., Yen, A., Heravi-Moussavi, A., Kheradpour, P., Zhang, Z., Wang, J. et al. (2015) Integrative analysis of 111 reference human epigenomes. *Nature*, 518, 317–330.
56. Escrivá, H., Bertrand, S. and Laudet, V. (2004) The evolution of the nuclear receptor superfamily. *Essays Biochem.*, 40, 11–26.
57. Boullon, R. and Suda, T. (2014) Vitamin D: calcium and bone homeostasis during evolution. *Bone Rep.*, 3, 480.