Transcriptional control of transglutaminase 2 expression in mouse apoptotic thymocytes

Katalin Sándor[a,1], Bence Dániel[b1,2], Bea Kiss[a], Fruzsina Kovács[a] and Zsuzsa Szondy[a]

[a]Division of Dental Biochemistry, Signaling and Apoptosis Research Group, [b]Nuclear Receptor Research Group, Research Center of Molecular Medicine, Department of Biochemistry and Molecular Biology, University of Debrecen, Nagyerdei krt.98., Debrecen, H-4012 Hungary

Running title: Transglutaminase 2 expression in apoptotic thymocytes

Page number: 38

Address correspondence and reprint requests to:
Zsuzsa Szondy M.D., Ph.D., D.Sc.
Department of Biochemistry and Molecular Biology
University of Debrecen
H-4012 Debrecen
Nagyerdei krt.98.
Hungary
Tel: +36 52 416432
Fax: +36 52 314 989 e-mail: szondy@med.unideb.hu

Footnotes

[1]Katalin Sándor and Bence Dániel contributed equally to the experiments presented in this work and share first authorship.

[2]Present address: Sanford-Burnham-Prebys Medical Discovery Institute at Lake Nona, 6400 Sanger Road, Orlando, FL 32827 USA

Abbreviations

Atf1; activating transcription factor 1 ; CBP, CREB binding protein; CCTF, CCCTC-binding factor; CEBP, CCAAT/enhancer-binding protein; ChIP-seq, chromatin immunoprecipitation followed by deep sequencing; 9cRA, 9-cis retinoic acid; Crem, cAMP response element modulator; CREB, cAMP response element binding protein; CRTC, CREB-regulated transcription coactivator; DNase-seq, DNase I hypersensitivity analysis followed by deep sequencing; eRNA, enhancer RNA; IGV, Integrative Genomics Viewer; RAR, retinoic acid receptor; RNAPII, RNA polymerase II; RNA-seq, RNA sequencing; RXR, retinoid X receptor; TGF-β, transforming growth factor β; TGM2, transglutaminase 2; TSS, transcription start site

Highlights

▪ A new approach was used to find the functional enhancers of the *Tgm2* gene in apoptotic thymocytes

▪ DNase-seq and ChIP-seq results from the ENCODE data base were integrated to map putative enhancers

▪ Then their eRNA levels were determined in thymocytes exposed to signals known to regulate TGM2 expression

▪ Their transcription factor recruitment was simultaneously determined by ChIP-qPCR

▪ Retinoids, TGF-β and dbcAMP selectively activate enhancers at the *Tgm2* locus

Abstract

Transglutaminase 2 (TGM2) is a ubiquitously expressed multifunctional protein, which participates in various biological processes including thymocyte apoptosis. As a result, the transcriptional regulation of the gene is complex and must depend on the cell type. Previous studies from our laboratory have shown that in dying thymocytes the expression of TGM2 is induced by external signals derived from engulfing macrophages, such as retinoids, transforming growth factor (TGF)-β and adenosine, the latter triggering the adenylate cyclase signaling pathway. The existence of TGF-β and retinoid responsive elements in the promoter region of *Tgm2* has already been reported, but the intergenic regulatory elements participating in the regulation of *Tgm2* have not yet been identified. Here we used publicly available results from DNase I hypersensitivity analysis followed by deep sequencing and chromatin immunoprecipitation followed by deep sequencing against CCCTC-binding factor, H3K4me3, H3K4me1 and H3K27ac to map a putative regulatory element set for *Tgm2* in thymocytes. By measuring eRNA expressions of these putative enhancers in retinoid, rTGF-β or dibutiryl cAMP-exposed thymocytes we determined which of them are functional. By applying ChIP-qPCR against SMAD4, retinoic acid receptor, retinoid X receptor, cAMP response element binding protein, P300 and H3K27ac under the same conditions, we identified two enhancers of *Tgm2*, which seem to act as integrators of the TGF-β, retinoid and adenylate cyclase signaling pathways in dying thymocytes. Our study describes a novel strategy to identify and characterize the signal specific functional enhancer set of a gene by integrating genome-wide datasets and measuring the production of enhancer specific RNA molecules.

Key words: T cells, Apoptosis, Gene Regulation, Enhancer RNA, Transglutaminase 2, ChIP-seq

1. Introduction

Transglutaminases [1] are a family of thiol- and Ca^{2+}-dependent acyl transferases that catalyze the formation of a covalent bond between the γ–carboxamide groups of peptide-bound glutamine residues and various primary amines, including the ε–amino group of lysine in certain proteins. The reaction results in post-translational modification of proteins by establishing ε–(γ–glutamyl)lysine cross-linkages and/or covalent incorporation of mono- or polyamines and into proteins. Transglutaminase 2 (TGM2) is very unique in the transglutaminase family, because besides being a transglutaminase it also possesses GTPase, protein disulphide isomerase and protein kinase enzymatic activities. In addition, TGM2 can also function in various biological settings as a protein/protein interaction partner [2]. TGM2 has been known for a long time to be associated with the *in vivo* apoptosis program of various cell types including T cells [3]. TGM2 expression is induced in thymocytes dying *in vivo* following exposure to various apoptotic signals [4], and TGM2 also appears in the dying T lymphocytes of HIV-infected individuals [5]. While, however, TGM2 is strongly induced in dying thymocytes *in vivo,* no induction of TGM2 was observed, when thymocytes were induced to die by the same stimuli *in vitro* [6] indicating that signals arriving from the tissue environment contribute to the *in vivo* induction of the enzyme in apoptotic thymocytes. Later studies from our laboratory have demonstrated that in dying thymocytes the expression of TGM2 is induced by signals derived from engulfing macrophages, such as retinoids [7], transforming growth factor (TGF)-β [8] and adenosine, the latter triggering the adenylate cyclase signaling pathway [9].

Due to its ubiquitous appearance and multiple functions, the expression of the *Tgm2* gene must be tightly controlled. The *in vivo* appearance of TGM2 protein in dying thymocytes seems to be regulated at the level of transcription, since in mice, which carry the beta-galactosidase reporter gene under the control of a 3.8 kilobase fragment of the *Tgm2*

promoter, the beta-galactosidase expression showed strong correlation with the endogenous TGM2 expression [6]. Other studies also indicate that the complex regulation of TGM2 expression is mediated at the level of transcription. Thus research has led to the identification of functional retinoid [10] and TGF-β [11] response elements in the promoter region of the gene. In addition, response elements for nuclear factorκB [12] and the hypoxia inducible factor [13] have been also identified in the core promoter. The limitation of these studies was that they could focus only on the promoter element at that time.

Recent advances in sequencing technologies allow genome-wide mapping of all the intergenic regulatory elements including enhancers. The combination of the DNase I hypersensitivity analysis followed by deep sequencing (DNase-seq) and chromatin immunoprecipitation followed by deep sequencing (ChIP-seq) datasets proved to be useful in identifying and characterizing the chromatin signatures of enhancer elements [14]**.** These include enrichment of histone 3 lysine 4 monomethylation (H3K4me1) as compared to H3K4me3, which peaks at promoter elements [15]**.** However, enhancers exhibiting these modifications are not necessary functional, because most of the open regulatory elements are enriched for at least one of these markers. Functional enhancers are rather associated with histone 3 acetylated at lysine 27 (H3K27ac) [16], with the transcription co-regulators P300 or CREB binding protein (CBP) [17], and with the actively transcribing polymerase II (RNAPII) [18]. The appearance of RNA sequencing and Global Run-On sequencing (GRO-seq) technologies revealed the existence of new types of RNA species including enhancer RNAs (eRNAs) [19,20]. These short RNAs are transcribed from active enhancers only and follow the induction profile of their target genes. Recently, it has been shown that the modified version of GRO-seq by enriching 5'-capped RNAs is a very powerful tool to identify active transcriptional regulatory elements based on its improved eRNA detection ability. The analysis of such datasets demonstrated the enrichment of expression quantitative trait loci,

disease-associated polymorphisms, H3K27ac and transcription factor binding sites at transcriptionally active regulatory elements [21]. According to these findings, eRNAs represent an accepted functional enhancer feature [22]. Hence, it has already been exploited to assign the active enhancer element repertoire of certain transcription factors, such as ESR1 [23] or RXR [24] to their target genes genome-wide.

Increasing evidence suggests that genes and their enhancers are located in functional domains separated from each other by proteins bound to insulator sequences. Insulators provide a barrier function to prevent repressive heterochromatin from spreading into a neighboring domain, provide an enhancer-blocking function when positioned between the enhancer and promoter, and allow three-dimensional looping of genomic regions [25,26]. CCCTC-binding factor (CTCF) is known to act as an insulator, but also implicated in the regulation of higher order chromatin structure. Recent studies investigating the spatial organization of the genome show that the direction of the CTCF motif determines the 3D folding of the genome [27]. In addition, it has been also shown that CTCF anchors play a vital role in connecting enhancers to their target gene promoter [28].

Based on advanced sequencing technologies, in 2014 the ENCODE (Encyclopedia of DNA elements) Consortium has released genome-wide datasets of the mouse genome [29]. These studies aimed to identify the cis-acting element repertoire of several human and mouse cell lines. In addition, they provided genome-wide data also from various mouse tissues.

In the present work we decided to identify those enhancers which regulate the expression of *Tgm2* in dying thymocytes exposed to retinoids, TGF-β and/ or following activation of the adenylate cyclase signaling pathway. Having access to the ENCODE generated datasets from mouse thymus, first we integrated DNase-seq and ChIP-seq results against CTCF, H3K4me3, H3K4me1 and H3K27ac in order to map the putative regulatory elements participating in the regulation of *Tgm2*. Then by measuring their basal and induced

eRNA levels, we determined which of them are functional. Finally, we analyzed the basal and induced binding of SMAD4, retinoic acid (RAR), retinoid X (RXR) receptor and cAMP response element binding protein (CREB), transcription factors mediating the effects of the retinoid, TGF-β and adenylate cyclase signaling pathways, to these enhancers. Our study describes the thymocyte specific enhancer elements of *Tgm2* and presents a novel strategy to identify and characterize the functional contribution of individual enhancers to gene regulation.

2. Materials and Methods

2.1. Materials

All reagents were obtained from Sigma-Aldrich (Budapest, Hungary) except indicated otherwise.

2.2. Experimental animals

Most of the experiments were carried out with 4-wk-old C57Bl/6 mice. Mice were maintained in specific pathogen-free condition in the Central Animal Facility of our university, and all animal experiments were approved by the Animal Care and Use Committee of University of Debrecen (DEMÁB).

2.3. Thymocyte cultures

Isolated thymocytes ($1x10^7$/ml) were cultured in RPMI medium 1640 supplemented with 10% charcoal stripped FBS, 2 mM glutamine, 1 mM Na pyruvate, 100 U/ml penicillin and 100

mg/ml streptomycin at 37°C in 5% CO_2. Cells were exposed to 0.3 μM 9-cis retinoic acid (9cRA) (Tocris Bioscience, Eching, Germany), 1 μM AM580, an RAR agonist, 100 nM LG268, an RXR agonist analogue alone, or in combination with 100 μM db-cAMP, a cell permeable cAMP or with 5 ng/ml recombinant human TGF-β1 (AbD Serotec, Kidlington, UK) and for the indicated time periods.

2.4 Real-Time Quantitative PCR (qRT-PCR) for detecting eRNA levels

For this purpose we designed eRNA specific primer pairs targeting the enhancer sequences upstream to their DNase hypersensitive sites, but on the enriched histone modifications (Supplementary Fig. 1). We used primer3 software to design the eRNA specific primers with standard conditions except the followings: amplicon length: 90-150bp; Tm: 59-60°C [30]. RNA was isolated with Trizol Reagent according to the manufacturer's guidelines. For mRNA measurements, High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems, Budapest, Hungary) was used and transcript quantification was performed by Quantitative Real-Time PCR reaction using SYBR green dye (gb SG PCR Master Mix, Generi Biotech, Czech Republic). For enhancer RNA measurements, RNA was isolated with the same Trizol based method, but samples were digested with DNaseI enzyme for 1 hour at 37C. RNA was then reverse transcribed using High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems, Budapest, Hungary) and quantification was performed with LightCycler® 480 SYBR Green I Master (Roche, USA). The following PCR program was used for enhancer RNA and mRNA detection: 95°C-3 min; 95°C-30 sec; 60°C-1 min. The last two steps were repeated 40 times. Transcript levels were normalized to *Ppia*. Primer sequences:

Tgm2 +30kb eRNA

Fw: 5’-CCACTGCAGCCTTCCTAGTT-3’

Rev: 5’-TCAGCCAGCACCTCTAGACA-3’

Tgm2 -7.9kb eRNA

Fw: 5’-CTCTAGGGCCCCAGTTGAG-3’

Rev: 5’-CAGGAGAGATGCCCTTCCAG-3’

Tgm2 -13kb eRNA

Fw: 5’-GTGAACTGTGCTGCTGAACA-3’

Rev: 5’-CAATCTCCAAAGCACCCCAC-3’

Tgm2 -20kb eRNA

Fw: 5’-AGCCTGAACCCACAAGTGTA-3’

Rev: 5’-TGTGTGTGCGTGTGTGTTTT-3’

Tgm2 -28kb eRNA

Fw: 5’-AGCTTTCCTGACCTGGTGTT-3’

Rev: 5’-TGCCTCACCACCATCATCA-3’

Prmt8 -1.7kb

Fw: 5’-CGTGAGCAGAGGTGAGGAGT-3’

Rev: 5’-GGTTAACCCAAGCTTCTTGCT-3’

2.5. Real-Time Quantitative PCR (qRT-PCR) for detecting *Tgm2* and *Rprd1b* mRNA levels

RNA was isolated with Trizol Reagent according to the manufacturer's guidelines. For mRNA measurements, High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems, Budapest, Hungary) was used and transcript quantification was performed by qRT-PCR reaction using SYBR green dye (gb SG PCR Master Mix, Generi Biotech, Czech Republic). RNA was then reverse transcribed using High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems, Budapest, Hungary) and quantification was performed with LightCycler® 480 SYBR Green I Master (Roche, USA).

Tgm2

Fw: 5'-AGGACATCAACCTGACCCTG-3'

Rev: 5'-CTTGATTTCGGGATTCTCCA-3'

Rprd1b

Fw: 5'-GACGTGTCGCTGCTAGAGAA-3'

Rev: 5'-CCGCTAGTAACAGACATGCTTC-3'

2.6. Chromatin immunoprecipitation followed by qPCR

ChIP was performed as previously described [31] with minor modifications. Briefly, cells were crosslinked with Di(N-succinimidyl) glutarate (ProteoChem, USA) for 30 minutes and then with formaldehyde for 10 min. Crosslinking reaction was terminated by the addition of glycine. Chromatin was sonicated with Diagenode Bioruptor to generate 200-1000 bp fragments. Immunoprecipitation was performed with antibodies against pre-immune IgG (Millipore, 12-370), RXR (sc-774), P300 (sc-585), CBP (sc-369X), CREB1 (ab31387), SMAD4 (sc-7154X), H3K27ac (ab4729) from Merck (Budapest, Hungary), Santa Cruz Biotechnology (Dallas, USA) or Abcam (Cambridge, UK). Chromatin antibody complexes

were precipitated with Protein A-coated paramagnetic beads (Life technologies, USA). After 6 washing steps complexes were eluted and reverse crosslinked. DNA fragments were column purified (Qiagen, MinElute). DNA was diluted and used for qPCR analysis. The following primer sequences were used:

Tgm2 +30kb

Fw: 5'-TCTGCCTAAGATGTGCAGAGG-3'

Rev: 5'-CCCTCTCTGAGGCGTTCC-3'

Tgm2 -7.9kb

Fw: 5'-CCAGAATAGAGGAGTGCTGGTAA-3'

Rev:5'-CAGTATTTTCTTGGTTGACTGTGG-3'

Tgm2 -13kb

Fw: 5'-GCAGAGTTGGCTCTCGTCA-3'

Rev: 5'-GTGTCTGCCTGTCCTGTCTG-3'

Tgm2 -20kb

Fw: 5'-TGGCATACACACAAGGAAGTG-3'

Rev: 5'-TTCCTGATCTGGTTTGTCAGTG-3'

Tgm2 -28kb

Fw: 5'-GAAGTCACCCCCAATCTGAG-3'

Rev: 5'-ATGCAGGTTTGCTGACTGCT-3'

Prmt8 -1.7kb

Fw: 5'-CGTGAGCAGAGGTGAGGAGT-3'

Rev: 5'-GGTTAACCCAAGCTTCTTGCT-3'

2.7. Accession numbers for ChIP-seq experiments

ChIP-seq experiments performed by the ENCODE Consortium have been downloaded and visualized by the Integrative Genomics Viewer (IGV) genome browser [32]. More specifically, ChIP-seq datasets were used against CTCF (GSM918734), H3K4me3 (GSM1000101), H3K4me1 (GSM1000102) and H3K27ac (GSM1000103). In addition DNaseI-seq results were also used (GSM1003827).

2.8. Chromosome conformation capture (3C)

Chromosome conformation capture experiments were carried out essentially as described previously [33,34]. Briefly, thymocytes were isolated and cross-linked using 2% formaldehyde for 10 min. Cross-linking reaction was quenched with glycine and nuclei were isolated. Chromatin was digested with HindIII restriction enzyme overnight at 37°C. On the following day digested chromatin was ligated in the presence of T4 DNA ligase (50U) at 16°C for 16 h. Ligated DNA fragments were isolated by phenol:chloroform extraction. PCR reactions were performed by using approximately 100 ng 3°C template.

Tgm2 -20kb (BAIT)

Fw: 5'- CACACACACATACGGTTCTGTT -3'

Tgm2 +11kb

Fw: 5'- CAGTCCTAGGGAAGTGGAGC -3'

Tgm2 +13kb

Fw: 5'- TTATGGGTTTGGGGACTGCA -3'

Tgm2 +30kb

Fw: 5'- GGCAGACCTTGTTTGAATGTGA -3'

Loading control primers (chr10:74788584-74788803):

Fw: 5' - CTC CAC CAT GAT GTA CAC CG – 3'

Rev: 5'- CATGGTTTCGGGAGATGCAG - 3'

2.9. Statistical analysis

All data are presented as means +/-SD. We made three biological replicates in all the experiments and we performed paired (two tailed) t tests. Results were considered significant with $p < 0.05$.

3. Results

3.1. Combination of DNase-seq and ChIP-seq allows identification of the putative enhancer set of *Tgm2*

To identify the putative enhancer elements, which might play a role in the regulation of *Tgm2* in mouse thymocytes, we exploited several mouse thymus derived datasets generated by the ENCODE Consortium. As a very first step, we tried to exploit the recently described feature about the motif orientation of CTCF and its effect on long-range chromatin looping and demarcating functionally different genomic regions. Tandem orientation of the CTCF motifs has been proposed to play role in gene regulation by allowing the enhancers to find their targets, while convergent motifs are more likely to be implicated in the formation of higher-

order chromatin structure [27,28]. We found three high affinity CTCF peaks flanking a 157kb wide genomic region around *Tgm2* (Fig. 1A). Analysis of the orientation of the motifs showed one convergent motif pair located at the borders of the 157kb genomic area, while the third, transcription start site (TSS) specific motif constitutes a tandem motif pair with the upstream CTCF border (Fig. 1A). These data suggest that the regulation of the gene might take place in a genomic region spanning 157 kb pairs. While convergent motifs are likely to function in establishing topologically associated domain borders, there are still many exceptions in which case potential enhancers could be overlooked. To avoid this possibility, we overlapped the interaction profile of a genomic loci of ~800kbp containing the *Tgm2* locus with the available CTCF ChIP-seq and DNase-seq from the CH12.LX mouse B lymphoblast cells [27] and the same datasets from thymus along with the histone modification ChIP-seq profiles. We found that interactions between the two proposed boundaries are clearly enriched suggesting that the regulation of *Tgm2* most probably takes place in this genomic unit and the enhancers do not communicate with the neighboring genes (Supplementary Figure 2). In addition, we also measured the mRNA expression of the neighboring gene (*Rprd1b*, chr2:157854530-157903056), which seems to be active based on the H3K4me3 histone modification present at the transcription start site of the gene. However, we could not detect any difference in its expression level in the presence of those compounds [7-9], which affected the expression *Tgm2* (data not shown).

Next we took advantage of the available DNase-seq dataset and identified several hypersensitive sites around the *Tgm2* gene located between the tandem CTCF sites, and an additional one located between the TSS specific and the downstream CTCF site, which show convergent motif orientation (Fig.1A). Since integration of DNase-seq with ChIP-seq against H3K4me3, H3K4me1 and H3K27ac proved to be useful in the identification of putative functional enhancers [15-17], [35], we applied this approach to find six (+30kb, TSS, -7.9kb, -

13kb, -20kb, -28kb) hypersensitive regions which showed enrichment for H3K4me3 and for H3K4me1. H3K4me3 is a hallmark of actively transcribed protein coding gene promoters [14], but can be detected occasionally together with enhancers due to a strong enhancer/promoter interaction. H3K4me1 is associated with most of the accessible regulatory elements in the genome and represents active, poised or primed enhancers depending on the presence of other histone modifications [15-16]. Out of the six elements, only two (-13kb and -20kb) showed enrichment for H3K27ac, which is a widely accepted functionally active enhancer marker [16]. Thus using the ENCODE generated genome-wide datasets from mouse thymus we have identified five intergenic enhancers for detailed analysis excluding the TSS/promoter region, which has been studied thoroughly.

3.2.The acetylation levels of the putative enhancers of *Tgm2* from the whole thymus tissue and from *in vitro* cultured thymocytes show similar pattern

The thymus is composed of various types of cells. Although thymocytes constitute the majority of cells in the thymus [36], one should be careful of using whole thymus for enhancer mapping, because other cell types, such as macrophages, dendritic cells and epithelial cells are also present. Therefore, signals are not reliably originated from the cell type of interest. That is why we isolated thymocytes from the mouse thymus to check whether the selected putative enhancers of *Tgm2* in these cells show a similar pattern of H3K27ac. As shown in Figure 1B, our ChIP-qPCR experiments demonstrated that similar to the ENCODE data originated from the whole thymus (Fig.1A), in *in vitro* cultured thymocytes the two (-13kb and -20kb) enhancer elements are much stronger acetylated as compared to the other putative enhancers of *Tgm2* (Fig. 1B). In accordance we found that these two highly acetylated enhancers have a concomitant enrichment also for the two well-known histone

acetyltransferases P300 and CBP. These results proved the adaptability of the two systems showing positive correlation between the whole thymus derived genome-wide datasets and our ChIP experiments on individual enhancers from thymocytes cultured *in vitro*.

3.3.**Enhancer RNA measurements identify functional enhancers for the *Tgm2* gene**

Previous studies from our laboratory have shown that in thymocytes TGM2 expression is regulated by external signals derived from engulfing macrophages, such as retinoids [7], TGF-β [8] and adenosine, the latter activating adenosine A2A receptors and the consequent adenylate cyclase pathway [9]. Among these signals 9cRA was the most effective in inducing TGM2 mRNA levels, resulting in a 5 fold induction. Activation of both the adenylate cyclase and the TGF-β signaling pathway synergized with the retinoid signaling pathway in inducing TGM2 mRNA levels.

To determine whether the putative enhancers respond to these signals, we decided to measure the amount of eRNAs transcribed from them. To detect the response of the individual enhancers, primary thymocytes cultured *in vitro* were exposed for 1 h to 0.3 μM 9cRA, a ligand of the nuclear retinoic acid receptors (RAR/RXR) [37], to 5 ng/ml rTGF-β leading to trimerization of SMAD4 with phosphorylated SMAD2 or 3 and its nuclear transport [36], or to 100 μM dbcAMP, a membrane permeable analogue of cAMP, resulting in the activation of protein kinase A and in subsequent phosphorylation of CREB [39]. However, phosphorylation and activation of CREB can occur also via activation of both the TGF-β signaling pathway, which activates protein kinase A in a cAMP-independent manner leading to the release of its regulatory subunit [40] and the retinoid signaling pathway, which can lead to

CREB phosphorylation in a protein kinase A-dependent [41,42] and independent [43,44] manner.

Interestingly, all of the enhancers showed elevated eRNA expression in response to 9-cRA, but only two of them, located -13kb and -20kb, appeared to be regulated also by TGF-β or dbcAMP (Fig. 2A, B). In addition, these two enhancers showed clear differences in their retinoid response showing a much robust eRNA upregulation as compared to the +30kb, -7.9kb and -28kb enhancers.

Our knowledge about these signaling pathways allowed the investigation of the binding of the signal specific transcription factors on these enhancers. We applied ChIP-qPCR to measure the enrichment of CREB, SMAD4, RAR and RXR in the presence of their activators on the identified enhancer elements (Fig. 2C). Exposure to dbcAMP efficiently recruited CREB to each enhancer elements, but the occupancy of CREB appeared to be the highest when we tested its binding on enhancers located at -13kb and -20kb. In accordance with reports, which demonstrated CREB activation in the retinoid and TGF-β signaling pathways, enhanced CREB binding could be detected also following exposure to 9cRA or rTGF-β, but again the occupancy of CREB appeared to be the highest when we tested its binding on the enhancers located at -13kb and -20kb (Fig. 2C). Similarly, when we assessed the binding of SMAD4 in the presence of rTGF-β, we found a similar pattern of dynamic recruitment on the investigated enhancers with the exception of the -7.9kb and -28kb enhancers which did not show any enrichment. ChIP-qPCR experiments performed against RAR or RXR in the presence of 9cRA did not show robust recruitment on the investigated enhancers. Still, an enhanced binding of RAR and RXR, as a 9cRA response, was found on all the investigated enhancers. In line with the eRNA measurements, these results indicate that the majority of RARs and RXRs are already bound to all of the investigated sites even in the absence of the externally added agonist, as it was previously shown by others [45], [24]. In

our negative control experiments, in which we performed the same experiments on the neuronal specific *Prmt8* gene *-1.7kb* region, neither eRNA formation, nor transcription factor bindings were found (Supplementary figure 3).

Our results demonstrate that the strategy developed by us proved to be efficient in the identification of signal selective enhancers. By measuring the synthesized eRNA species from individual enhancers and combining the results with transcription factor binding data, we identified five signal selective enhancers of *Tgm2* in mouse thymocytes.

3.4. The signaling pathway induced by TGF-β selectively potentiates the retinoid signaling on *Tgm2* associated enhancer elements

Recently, it has been shown by our group that the adenylate cyclase and TGF-β mediated signaling synergize with the retinoid signaling on TGM2 leading to a more pronounced induction of TGM2 at protein level [7]. In line with the protein data, we found by using qRT-PCR analysis that administration of TGF-β could significantly increase the 9cRA-induced mRNA expression of *Tgm2* (Fig.3A). Based on these results we decided to test, whether the activation of TGF-β signaling could potentiate the retinoid-mediated signaling pathways in the context of *Tgm2* regulation. To identify those enhancers on which the collaboration between the two signaling pathways might take place, we determined enhancer specific eRNA levels in the presence of various signal combinations. As shown in Figure 3B, enhancer elements located at +30kb and at -28kb did not show any sign of collaboration in agreement with our previous results, which indicated that on these elements no significant TGF-β-induced SMAD4 binding was found. On the other hand, we could detect significantly increased eRNA expression on the -7.9kb, -13kb and -20kb enhancers if both signals were present, indicating that the TGF-β and the retinoid pathways crosstalk on these enhancers. To

analyze this crosstalk in more detail, we performed ChIP-qPCR experiments using RXR specific antibodies. While 9-cRA alone resulted only in a slight increase in the RXR recruitment, a significant increase in the recruitment of RXR to the enhancers located -7.9kb, -13kb and -20kb could be detected, when TGF-β was also administered (Fig. 3C). Interestingly, in the presence of both signals the SMAD4 recruitment increased also. Moreover, the recruited P300 and CBP showed also elevated levels on these enhancer elements in response to TGF-β/9cRA co-treatment as compared to 9cis-RA alone. These results suggest that in the presence of both signals the accessibility of these sites increases for both transcription factors, which ultimately leads to an increased coactivator binding and more efficient transcription. Interestingly, in the case of the -7.9 kb enhancer the interaction existed between the two pathways despite the lack of significant binding of SMAD4 to this enhancer in the presence of TGF-β alone.

Knowing the fact that 9cRA is able to activate various RXR heterodimers including RAR/RXR, we applied a selective synthetic agonist of RARα (AM580) and that of RXR (LG268) to determine the target of interaction between the retinoid and the TGF-β signaling pathways on the signal integrator enhancers (-13kb and -20kb). We found that both the RARα and the RXR activator alone could significantly induce the level of enhancer transcripts at the signal integrator enhancers. Interestingly, when we measured the activation of these enhancers in the presence of the signal combinations of TGF-β/AM580 or TGF-β/LG268, we found a more pronounced eRNA expression in both cases and on both enhancers (Fig. 3D). These data indicate that TGF-β signaling is able to potentiate the transcriptional activity of the RXR/RARα heterodimer, in line with the fact that we detected both RAR and RXR binding on these enhancers. In addition, administration of the pan RAR antagonist AGN194310 reduced LG268/TGF-β-induced eRNA levels to the TGF-β/AGN194310-induced eRNA levels, and the AM580/TGF-β- and the LG268/TGF-β-induced eRNA levels to that of the

AM580/AGN194310- and the LG268/AGN194310-induced eRNA levels, respectively, indicating that only RAR/RXR heterodimers interact with the TGF-β signaling on these enhancers.

3.5. The adenylate cyclase pathway also selectively potentiates retinoid signaling on *Tgm2* associated enhancer elements

Since previous studies indicated that the adenylate cyclase pathway also enhances TGM2 protein levels in the thymocytes *in vivo* [9], we determined *Tgm2* mRNA levels in the presence of dbcAMP as well. As shown in Figure 4A exposure to dbcAMP alone enhanced the steady state levels of *Tgm2* mRNA, and when dbcAMP was applied together with 9cRA, it synergized with retinoid signaling in regulating *Tgm2* mRNA levels, though less efficiently as TGF-β did (Fig. 3A). Next by measuring eRNA levels on the various enhancers, we examined whether the two signaling pathways synergize also in the context of *Tgm2* regulation. As shown in Figure 4B, enhancer elements located at +30kb, -7.9kb and -28kb did not show any sign of collaboration in agreement with our previous results, which indicated that on these elements no significant dbcAMP-induced CREB binding was found. On the other hand, we could detect significantly increased eRNA expression on the -13kb and -20kb enhancers if both signals were present, indicating that the adenylate cyclase and the retinoid signaling pathways crosstalk on these enhancers.

To analyze this crosstalk in more detail, we performed similar ChIP-qPCR experiments for transcription factor binding, as we did for studying the TGF-β/retinoid signaling crosstalk. However, as seen in Figure 4C, exposure to dbcAMP did not enhance the 9cRA-induced RXR recruitment to any of the enhancers. The recruitment of CREB, on the other hand, on most enhancers was higher in the presence of the two signals, than in the

presence of 9cRA alone, but not higher than in the presence of dbcAMP alone. Thus, while in the case of the TGF-β/retinoid crosstalk we detected increased enhancer accessibility for RXR, when the two signals were present together, no such an enhancement for RXR binding was detected in the adenylate cyclase/retinoid signaling crosstalk. In addition, despite the fact that more CREB was recruited to most of the enhancers in the presence of the two signals, neither more CBP, nor more P300 coactivator binding were detected in the presence of dbcAMP, as compared to 9cRA alone.

3.6. Chromosome conformation capture experiments reveal interactions between the -20kb enhancer and the *Tgm2* gene

As a control for our approach, we performed chromosome conformation capture experiments [33,34] to prove that the enhancers, we identified, interact with *Tgm2* gene. For this purpose we selected the +11kb and the +30 kb (-3' UTR) regions of the gene, which are DNase hypersensitive and as a control the +13kb region, which is a lowly accessible region. The limitation of the method related to the HindIII restriction sites allowed us to investigate the interaction of these regions with the potential -20kb enhancer. As shown in Figure 5, we were able to demonstrate the interaction between the enhancer located -20kb and the intronic region of the gene. Interestingly, we also found an interaction with the 3' UTR of the gene, where we also identified a DNase hypersensitive site, but could not show the interaction with the intervening sequence (+13kb region). Importantly, none of these contacts appeared to show induced interactions in the presence of the applied agonists, indicating that these are stable, pre-formed loops between the enhancer and the gene.

4. Discussion

Previous studies from our laboratory have shown that the expression of the TGM2 protein is associated with the *in vivo* apoptosis of thymocytes and is regulated transcriptionally by at least three extracellular signals arriving from the neighboring engulfing macrophages [6-9]. However, the regulation of TGM2 expression must be very complex, since the protein is ubiquitously expressed and participates in many biological processes [2]. During the years several publications appeared in the literature about the regulation of the *Tgm2* gene [10-13]. However, these studies were limited to the investigation of one single promoter element, and did not address the cell type specific transcriptional regulation of the gene. With the help of the next-generation sequencing coupled approaches we had the opportunity to specifically map and identify the cell type- and gene-specific regulatory elements genome-wide as shown by others [14-18]. In the present study, we took advantage of the ENCODE generated datasets from mouse thymus. Using the integration of DNase-seq and ChIP-seq approaches we revealed the putative regulatory element repertoire of *Tgm2*. We have exploited the recently described phenomenon about CTCF motif orientation [27,28] in order to find the most probable, functional enhancer set for *Tgm2*. These approaches lead to the identification of five intergenic elements (+30kb, -7.9kb, -13kb, -20kb, -28kb) as potential regulators of the gene in dying thymocytes. The functional contribution of these elements to transcription regulation could be studied with the well-known, though sometimes artificial luciferase based reporter assays. However, if the cell type of interest is not a good subject for modification, such as the fast dying thymocytes, it is almost impossible to get reliable information about the function of a given enhancer element under the cell type specific setting. Several studies have recently described that activation of a signal specific transcription factor leads to a similar induction profile on the target genes as on the contributing enhancer elements at the nascent RNA level

[20,23,24]. As a result, eRNA production from an enhancer has been accepted as one if not the best marker for enhancer activation [22]. Although, their function in transcription regulation is debated, more and more studies report eRNA specific functions in gene regulation [46-48]. In addition, the presence of eRNAs is indicative of loop formation between the enhancer and gene promoter. Thus, as a novel possible tool, so far not applied for clarification of the functional contribution of the individual enhancer elements to the regulation of a gene, we decided to determine the production of eRNAs [19] synthesized from active enhancers of the *Tgm2* gene in the presence of TGF-β, dbcAMP or retinoids. These measurements revealed three retinoid selective enhancers located +30kb, -7.9kb and -28kb respectively. Surprisingly, we found two distant enhancer elements located at -13kb and -20kb away from the gene, at which we detected eRNA transcription in response to all of the above mentioned activators. Actually, these two regions show the highest DNase hypersensitive profile in agreement with their histone acetylation level, histone acetyltransferase enrichment determined by the presence of P300 and CBP, and the occupancy of the signal specific transcription factors in response to stimulation. In addition, combinations of 9cRA with dbcAMP or TGF-β showed a potentiated retinoid response on these enhancers. Furthermore, we were also able to show that the more distant -20kb enhancer interacts with the intronic- and 3'UTR-region of the gene using 3C approach. These data indicate that these enhancers might act as integrators, and might function also as shadow enhancers [49] in the regulation of the *Tgm2* gene expression.

In the case of TGF-β and retinoid synergism the enhanced eRNA transcription at these sites was attributable to more RXR, SMAD4, P300 and CBP recruitment in the presence of the two signals. Interestingly, though we detected a crosstalk between the adenylate cyclase and retinoid signaling pathways as well, this latter interaction did not result in enhanced CBP or P300 binding, despite the fact that an enhanced CREB binding was also detected. However,

CREB works in concert not only with the coactivator CBP, but depending on the cellular context, also with a family of coactivators called CREB-regulated transcription coactivators (CRTCs) [50]. Of the CRTC family members, CRTC2 is thought to be the major mediator of regulated CREB target gene expression [51] and is known to be expressed by T cells as well [52]. Alternatively, the adenylate cyclase signaling pathway regulates these enhancers via other transcription factors as well and not only via CREB. Since in the ENCODE database we have not found ChIP-seq datasets for transcription factor bindings related to mouse thymus, we decided to check the motifs under the identified DNase hypersensitive regions using the JASPAR database. We probed all the regulatory sites for the motifs of the following transcription factors known to be regulated by the adenylate cyclase signaling pathway: CREB1, CCAAT/enhancer-binding protein alpha (CEBPA), activating transcription factor 1 (Atf1), Atf3, Creb3l2, cAMP response element modulator (Crem), CREB3, CEBPB, CEBPD, CEBPE, CEBPG, CREB3L1, Creb5. Interestingly, many of these transcription factor binding motifs are enriched under the identified regulatory regions (Supplementary Table 1). The two cAMP-responsive elements showed enrichment for these motifs also, but we found only CREB1 motif at the enhancer located -20kb from the gene. This prediction validated our approach because we found the strongest CREB binding at this sites using ChIP-qPCR. At the -13kb enhancer we found several CREB1 motifs and also other CREB motifs, but importantly ATF motifs were also enriched. Thus the crosstalk between the adenylate cyclase and the retinoid signaling pathways might be mediated via proteins that we did not investigate in our study.

In summary, in the present study, we dissected the regulation of an apoptosis related gene in mouse thymocytes in a so far unprecedented manner. Our strategy to map and study the contribution of individual enhancers in the regulation of *Tgm2* serves as a proof of concept, thus can be applied to solve similar problems. However, we would like to emphasize

that the usage of genome engineering is inevitable in order to support direct evidence about the contribution of a single enhancer element in gene regulation. Our results shed light on a very intriguing regulatory process at the level of individual enhancers and show that such enhancer clusters are able to integrate various signaling pathways, as it has been proposed recently by the super-enhancer concept [51]. This theory suggests that members of the super-enhancers can physically and functionally interact with each other in order to fine tune the activity of the target gene. We think that the enhancers of *Tgm2* in thymocytes serve a similar function by enabling the dying cells to respond very precisely to the environmental signals in the thymus. Among these signals macrophage-derived TGF-β has been shown to contribute also to the differentiation of regulatory T cells which, as a result, depends on the constantly ongoing apoptosis and engulfment of the improperly produced thymocytes in the thymus after birth [52].

Acknowledgements

The help of Zsolt Sarang in putting together the figures and the technical assistance of Edit Komóczi and Zsolt Hartmann are gratefully acknowledged. This study was supported by Hungarian grants from the National Research Fund (OTKA K104228 and NK 105046) and the TÁMOP 4.2.2.A-11/1/KONV-2012-0023 "VÉD-ELEM" project co-financed by the European Social Fund and the European Regional Development Fund.

References

1. Folk JE, Chung SI (1985) Transglutaminases. Methods Enzymol 113:358-375
2. Fésüs L, Piacentini M (2002) Transglutaminase 2: an enigmatic enzyme with diverse functions. Trends Biochem Sci 27:534-539
3. Fésüs L, Szondy Z (2005) Transglutaminase 2 in the balance of cell death and survival. FEBS Lett 579:3297-3302
4. Szondy Z, Molnar P, Nemes Z, Boyiadzis M, Kedei N, Tóth R, Fésüs L (1997a) Differential expression of tissue transglutaminase during in vivo apoptosis of thymocytes induced via distinct signalling pathways. FEBS Lett 404:307-313
5. Amendola A, Gougeon ML, Poccia F, Bondurand A, Fesus L, Piacentini M (1996) Induction of "tissue" transglutaminase in HIV pathogenesis: evidence for high rate of apoptosis of CD4+ T lymphocytes and accessory cells in lymphoid tissues. Proc Natl Acad Sci USA 93:11057-11062
6. Szegezdi E, Szondy Z, Nagy L, Nemes Z, Friis RR, Davies PJ, Fésüs L (2000) Apoptosis-linked in vivo regulation of the tissue transglutaminase gene promoter. Cell Death Differ 7:1225-1233
7. Garabuczi É, Kiss B, Felszeghy S, Tsay GJ, Fésüs L, Szondy Z (2013) Retinoids produced by macrophages engulfing apoptotic cells contribute to the appearance of transglutaminase 2 in apoptotic thymocytes. Amino Acids 44:235-244
8. Szondy Z, Sarang Z, Molnar P, Nemeth T, Piacentini M, Mastroberardino PG, Falasca L, Aeschlimann D, Kovacs J, Kiss I, Szegezdi E, Lakos G, Rajnavolgyi E, Birckbichler PJ, Melino G, Fesus L (2003) Transglutaminase 2-/- mice reveal a phagocytosis-associated crosstalk between macrophages and apoptotic cells. Proc Natl Acad Sci USA 100:7812-7817

9. Sándor K, Pallai A, Duró E, Legendre P, Couillin I, Sághy T, Szondy Z (2016) Adenosine produced from adenine nucleotides through an interaction between apoptotic cells and engulfing macrophages contributes to the appearance of transglutaminase 2 in dying thymocytes. Amino Acids (in press)
10. Nagy L, Saydak M, Shipley N, Lu S, Basilion JP, Yan ZH, Syka P, Chandraratna RA, Stein JP, Heyman RA, Davies PJ (1996) Identification and characterization of a versatile retinoid response element (retinoic acid receptor response element-retinoid X receptor response element) in the mouse tissue transglutaminase gene promoter. J Biol Chem 271:4355-4365
11. Ritter SJ, Davies PJ (1998) Identification of a transforming growth factor-beta1/bone morphogenetic protein 4 (TGF-beta1/BMP4) response element within the mouse tissue transglutaminase gene promoter. J Biol Chem 273:12798-12806
12. Brown KD (2013) Transglutaminase 2 and NF-κB: an odd couple that shapes breast cancer phenotype. Breast Cancer Res Treat 137:329-336
13. Penumatsa KC, Toksoz D, Warburton RR, Hilmer AJ, Liu T, Khosla C, Comhair SA, Fanburg BL (2014) Role of hypoxia-induced transglutaminase 2 in pulmonary artery smooth muscle cell proliferation. Am J Physiol Lung Cell Mol Physiol 307:L576-585
14. Visel A, Rubin EM, Pennacchio LA (2009) Genomic views of distant-acting enhancers. Nature 461:199-205
15. Heintzman ND, Stuart RK, Hon G, Fu Y, Ching CW, Hawkins RD, Barrera LO, Van Calcar S, Qu C, Ching KA, Wang W, Weng Z, Green RD, Crawford GE, Ren B (2007) Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. Nat Genet 39:311-318
16. Creyghton MP, Cheng AW, Welstead GG, Kooistra T, Carey BW, Steine EJ, Hanna J, Lodato MA, Frampton GM, Sharp PA, Boyer LA, Young RA, Jaenisch R. (2010)

Histone H3K27ac separates active from poised enhancers and predicts developmental state. Proc Natl Acad Sci U S A 107:21931-21936

17. Visel A, Blow MJ, Li Z, Zhang T, Akiyama JA, Holt A, Plajzer-Frick I, Shoukry M, Wright C, Chen F, Afzal V, Ren B, Rubin EM, Pennacchio LA (2009) ChIP-seq accurately predicts tissue-specific activity of enhancers. Nature 457:854-858
18. De Santa F, Barozzi I, Mietton F, Ghisletti S, Polletti S, Tusi BK, Muller H, Ragoussis J, Wei CL, Natoli G (2010) A large fraction of extragenic RNA pol II transcription sites overlap enhancers. PLoS Biol 8:e1000384
19. Kim TK, Hemberg M, Gray JM, Costa AM, Bear DM, Wu J, Harmin DA, Laptewicz M, Barbara-Haley K, Kuersten S, Markenscoff-Papadimitriou E, Kuhl D, Bito H, Worley PF, Kreiman G, Greenberg ME (2010) Widespread transcription at neuronal activity-regulated enhancers. Nature 465:182-187
20. Wang D, Garcia-Bassets I, Benner C, Li W, Su X, Zhou Y, Qiu J, Liu W, Kaikkonen MU, Ohgi KA, Glass CK, Rosenfeld MG, Fu XD (2011) Reprogramming transcription by distinct classes of enhancers functionally defined by eRNA. Nature 474:390-394
21. Danko CG, Hyland SL, Core LJ, Martins AL, Waters CT, Lee HW, Cheung VG, Kraus WL, Lis JT, Siepel A (2015) Identification of active transcriptional regulatory elements from GRO-seq data. Nature methods 12:433-438
22. Lam MT, Li W, Rosenfeld MG, Glass CK (2014) Enhancer RNAs and regulated transcriptional programs. Trends Biochem Sci 39:170-182
23. Hah N, Murakami S, Nagari A, Danko CG, Kraus WL (2013) Enhancer transcripts mark active estrogen receptor binding sites. Genome Res 23:1210-1223
24. Daniel B, Nagy G, Hah N, Horvath A, Czimmerer Z, Poliska S, Gyuris T, Keirsse J, Gysemans C, Van Ginderachter JA, Balint BL, Evans RM, Barta E, Nagy L (2014)

The active enhancer network operated by liganded RXR supports angiogenic activity in macrophages. Genes Dev 28:1562-1577

25. Chung JH, Whiteley M, Felsenfeld G (1993) A 5' element of the chicken beta-globin domain serves as an insulator in human erythroid cells and protects against position effect in Drosophila. Cell 74:505-514
26. Ong CT, Corces VG (2014) CTCF: an architectural protein bridging genome topology and function. Nat Rev Genet 15: 234-246
27. Rao SS, Huntley MH, Durand NC, Stamenova EK, Bochkov ID, Robinson JT, Sanborn AL, Machol I, Omer AD, Lander ES, Aiden EL (2014) A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. Cell 159: 1665-1680
28. Tang Z, Luo OJ, Li X, Zheng M, Zhu JJ, Szalaj P, Trzaskoma P, Magalska A, Wlodarczyk J, Ruszczycki B, Michalski P, Piecuch E, Wang P, Wang D, Tian SZ, Penrad-Mobayed M, Sachs LM, Ruan X, Wei CL, Liu ET, Wilczynski GM, Plewczynski D, Li G, Ruan Y (2015) CTCF-Mediated Human 3D Genome Architecture Reveals Chromatin Topology for Transcription. Cell 163:1611-1627
29. Yue F, Cheng Y, Breschi A, Vierstra J, Wu W, Ryba T, Sandstrom R, Ma Z, Davis C, Pope BD, Shen Y, Pervouchine DD, Djebali S, Thurman RE, Kaul R, Rynes E, Kirilusha A, Marinov GK, Williams BA, Trout D, Amrhein H, Fisher-Aylor K, Antoshechkin I, DeSalvo G, See LH, Fastuca M, Drenkow J, Zaleski C, Dobin A, Prieto P, Lagarde J, Bussotti G, Tanzer A, Denas O, Li K, Bender MA, Zhang M, Byron R, Groudine MT, McCleary D, Pham L, Ye Z, Kuan S, Edsall L, Wu YC, Rasmussen MD, Bansal MS, Kellis M, Keller CA, Morrissey CS, Mishra T, Jain D, Dogan N, Harris RS, Cayting P, Kawli T, Boyle AP, Euskirchen G, Kundaje A, Lin S, Lin Y, Jansen C, Malladi VS, Cline MS, Erickson DT, Kirkup VM, Learned K, Sloan

CA, Rosenbloom KR, Lacerda de Sousa B, Beal K, Pignatelli M, Flicek P, Lian J, Kahveci T, Lee D, Kent WJ, Ramalho Santos M, Herrero J, Notredame C, Johnson A, Vong S, Lee K, Bates D, Neri F, Diegel M, Canfield T, Sabo PJ, Wilken MS, Reh TA, Giste E, Shafer A, Kutyavin T, Haugen E, Dunn D, Reynolds AP, Neph S, Humbert R, Hansen RS, De Bruijn M, Selleri L, Rudensky A, Josefowicz S, Samstein R, Eichler EE, Orkin SH, Levasseur D, Papayannopoulou T, Chang KH, Skoultchi A, Gosh S, Disteche C, Treuting P, Wang Y, Weiss MJ, Blobel GA, Cao X, Zhong S, Wang T, Good PJ, Lowdon RF, Adams LB, Zhou XQ, Pazin MJ, Feingold EA, Wold B, Taylor J, Mortazavi A, Weissman SM, Stamatoyannopoulos JA, Snyder MP, Guigo R, Gingeras TR, Gilbert DM, Hardison RC, Beer MA, Ren B; Mouse ENCODE Consortium (2014) A comparative encyclopedia of DNA elements in the mouse genome. Nature 515:355-364

30. Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, Rozen SG (2012) Primer3--new capabilities and interfaces. Nucleic Acids Res 40:e115
31. Daniel B, Balint BL, Nagy ZS, Nagy L (2014) Mapping the genomic binding sites of the activated retinoid X receptor in murine bone marrow-derived macrophages using chromatin immunoprecipitation sequencing. Methods Mol Biol 1204:15-24
32. Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, Mesirov JP (2011) Integrative genomics viewer. Nat Biotechnol 29:24-26
33. Dekker J, Rippe K, Dekker M, Kleckner N (2002) Capturing chromosome conformation. Science 295:1306-1311
34. van de Werken HJ, de Vree PJ, Splinter E, Holwerda SJ, Klous P, de Wit E, de Laat W (2012) 4C technology: protocols and data analysis. Methods in enzymology 513:89-112

35. Zhu Y, Sun L, Chen Z, Whitaker JW, Wang T, Wang W (2013) Predicting enhancer transcription and activity from chromatin modifications. Nucleic Acids Res 41:10032-10043
36. Bellamy D, Alkufaishi H. (1972) The cellular composition of thymus: a comparison between cortisol-treated and aged C57-BL mice. Age Ageing 1:88-98
37. Heyman RA, Mangelsdorf DJ, Dyck JA, Stein RB, Eichele G, Evans RM, Thaller C (1992) 9-cis retinoic acid is a high affinity ligand for the retinoid X receptor. Cell 68:397-406
38. Massagué J (2008) TGFβ signalling in context. Nat Rev Mol Cell Biol 13:616-630
39. Sands WA, Palmer TM (2008) Regulating gene transcription in response to cyclic AMP elevation. Cell Signal 20:460-466
40. Zhang L, Duan CJ, Binkley C, Li G, Uhler MD, Logsdon CD, Simeone DM (2004) A Transforming Growth Factor β-Induced Smad3/Smad4 Complex Directly Activates Protein Kinase A. Mol Cell Biol 24:2169-2180
41. Manna PR, Slominski AT, King SR, Stetson CL, Stocco DM (2014) Synergistic activation of steroidogenic acute regulatory protein expression and steroid biosynthesis by retinoids: involvement of cAMP/PKA signaling. Endocrinology 155:576-591
42. Al-Wadei HA, Schuller HM (2006) Cyclic adenosine monophosphate-dependent cell type-specific modulation of mitogenic signaling by retinoids in normal and neoplastic lung cells. Cancer Detect Prev 30:403-411
43. Cañón E, Cosgaya JM, Scsucova S, Aranda A. (2004) Rapid effects of retinoic acid on CREB and ERK phosphorylation in neuronal cells. Mol Biol Cell 15:5583-5592

44. Aggarwal S, Kim SW, Cheon K, Tabassam FH, Yoon JH, Koo JS (2006) Nonclassical action of retinoic acid on the activation of the cAMP response element-binding protein in normal human bronchial epithelial cells. Mol Biol Cell 17:566-575

45. Brázda P, Krieger J, Daniel B, Jonas D, Szekeres T, Langowski J, Tóth K, Nagy L, Vámosi G (2014) Ligand binding shifts highly mobile retinoid X receptor to the chromatin-bound state in a coactivator-dependent manner, as revealed by single-cell imaging. Mol Cell Biol 34:1234-1245

46. Lam MT, Cho H, Lesch HP, Gosselin D, Heinz S, Tanaka-Oishi Y, Benner C, Kaikkonen MU, Kim AS, Kosaka M, Lee CY, Watt A, Grossman TR, Rosenfeld MG, Evans RM, Glass CK (2013) Rev-Erbs repress macrophage gene expression by inhibiting enhancer-directed transcription. Nature 498:511-515

47. Li W, Notani D, Ma Q, Tanasa B, Nunez E, Chen AY, Merkurjev D, Zhang J, Ohgi K, Song X, Oh S, Kim HS, Glass CK, Rosenfeld MG (2013) Functional roles of enhancer RNAs for oestrogen-dependent transcriptional activation. Nature 498:516-520

48. Melo CA, Drost J, Wijchers PJ, van de Werken H, de Wit E, Oude Vrielink JA, Elkon R, Melo SA, Léveillé N, Kalluri R, de Laat W, Agami R (2013) eRNAs are required for p53-dependent enhancer activity and gene transcription. Mol Cell 49: 524-535

49. Hong JW, Hendrix DA, Levine MS (2008) Shadow enhancers as a source of evolutionary novelty. Science. 2008 321:1314

50. Altarejos JY, Montminy M (2011) CREB and the CRTC co-activators: sensors for hormonal and metabolic signals. Nat Rev Mol Cell Biol 12:141-151

51. Conkright MD, Canettieri G, Screaton R, Guzman E, Miraglia L, Hogenesch JB, Montminy M (2003) TORCs: transducers of regulated CREB activity. Mol Cell 12:413-423

52. Hernandez JB, Chang C, LeBlanc M, Grimm D, Le Lay J, Kaestner KH, Zheng Y, Montminy M (2015) The CREB/CRTC2 pathway modulates autoimmune disease by promoting Th17 differentiation. Nat Commun 6:7216
53. Hnisz D, Schuijers J, Lin CY, Weintraub AS, Abraham BJ, Lee TI, Bradner JE, Young RA (2015) Convergence of developmental and oncogenic signaling pathways at transcriptional super-enhancers. Mol Cell 58:362-370
54. Konkel JE, Jin W, Abbatiello B, Grainger JR, Chen W. (2014) Thymocyte apoptosis drives the intrathymic generation of regulatory T cells. Proc Natl Acad Sci U S A 111:E465-473

Figure Legends

Figure 1 Finding putative enhancers for the *Tgm2* gene in mouse thymocytes. (A) IGV genome browser view about the *Tgm2 l*ocus. ChIP-seq results against CTCF, H3K4me1, H3Kme3 and H3K27ac along with the DNase hypersensitive sites are shown from mouse thymus. CTCF motif orientations are indicated with black horizontal arrows and the presumable interactions are depicted as connecting lines between them (top). Intergenic DNase hypersensitive sites and the TSS of *Tgm2* are shown with vertical arrows, and their relative distance to the TSS is also highlighted (under the CTCF track). TSSs of *Tgm2* and *Rprd1b* genes are also shown with the direction of the gene highlighted by horizontal arrows. (B) ChIP-qPCR measurements against CBP, P300 and H3K27ac from *in vitro* cultured mouse thymocytes are shown on the indicated intergenic regulatory elements. As a control, we also performed the same ChIP-qPCRs on the neuron specific *Prmt8* gene -17kb region, which in the thymic genome appears as a non-DNase hypersensitive region (non-DHS). Data represent mean±SD of triplicate measurements.

Figure 2 Combined eRNA measurements and determination of transcription factor binding identify functional enhancers for the *Tgm2* gene. (A) *Tgm2* mRNA levels determined by RT-qPCR analysis in the presence of the indicated compounds (0.3 μM 9cRA, 5 ng/ml rTGF-β or 100 μM dbcAMP). **(**B) Intergenic DNase hypersensitive sites and their relative distance to the TSS of *Tgm2* are shown (top). (C) Enhancer RNA levels determined by RT-qPCR analysis on the indicated *Tgm2* enhancers in the presence of the indicated molecules (0.3 μM 9cRA, 5 ng/ml rTGF-β or to 100 μM dbcAMP) in *in vitro* cultured thymocytes after 1 h treatment. (D) CREB, SMAD4, RXR and RAR binding on the indicated enhancer elements determined by ChIP-qPCR analysis in the presence of the indicated molecules after 1 h exposure.

Background was determined by using rabbit IgG and subtracted from the values presented on the graphs. CTR: vehicle treated control. As a negative control, we also performed the same eRNA and ChIP-qPCR measurements on the neuron specific *Prmt8* gene -17kb region, which in the thymic genome appears as a non-DNase hypersensitive region (Supplementary Figure 3). Data represent mean±SD of triplicate determinations. *Significantly different ($P<0.05$).

Figure 3 The signaling pathway induced by TGF-β selectively potentiates the retinoid signaling on *Tgm2* associated enhancer elements. (A) *Tgm2* mRNA levels determined by RT-qPCR analysis in the presence of the indicated compounds (0.3 μM 9cRA, 5 ng/ml rTGF-β or 0.3 μM 9cRA/5 ng/ml rTGF-β) after 6 h treatment. mRNA levels are expressed as fold induction as compared to the DMSO (0.5%) exposed samples. (B) Enhancer RNA levels determined by RT-qPCR analysis on the indicated *Tgm2* enhancers in the presence of the indicated molecules in *in vitro* cultured thymocytes after 1 h treatment. (C) RXR, SMAD4, P300 and CBP binding on the indicated enhancer elements determined by ChIP-qPCR analysis in the presence of the indicated compounds after 1 h treatment. Background was determined by using rabbit IgG and subtracted from the values presented on the graphs. (D) Enhancer RNA levels determined by RT-qPCR analysis on the indicated *Tgm2* enhancers in the presence of 1 μM AM580, an RARα agonist, 100 nM LG268, an RXR agonist, 5 ng/ml rTGF-β, 100 nM AGN194310, a pan RAR agonist in *in vitro* cultured thymocytes after 1 h treatment. CTR: vehicle treated control. As a negative control, we also performed the same eRNA and ChIP-qPCR measurements on the neuron specific *Prmt8* gene -17kb region, which in the thymic genome appears as a non-DNase hypersensitive region (Supplementary Figure 3). Data represent mean±SD of triplicate determinations. *Significantly different ($P<0.05$).

Figure 4 The adenylate cyclase pathway also selectively potentiates retinoid signaling on *Tgm2* associated enhancer elements. (A) *Tgm2* mRNA levels determined by RT-qPCR analysis in the presence of the indicated compounds (0.3 μM 9cRA, 100 μM dbcAMP or 0.3 μM 9cRA/100 μM dbcAMP) after 6 h treatment. mRNA levels are expressed as fold induction as compared to the DMSO (0.5%) exposed samples. (B) Enhancer RNA levels determined by RT-qPCR analysis on the indicated *Tgm2* enhancers in the presence of the indicated molecules in *in vitro* cultured thymocytes after 1 h treatment.(C) RXR, CREB, P300 and CBP binding on the indicated enhancer elements determined by ChIP-qPCR analysis in the presence of the indicated compounds after 1 h treatment. Background was determined by using rabbit IgG and subtracted from the values presented on the graphs. CTR: vehicle treated control. As a negative control, we also performed the same eRNA and ChIP-qPCR measurements on the neuron specific *Prmt8* gene, which in the thymic genome appears as a non-DNase hypersensitive region (Supplementary Figure 3). Data are represented as mean± SD of triplicate determinations. *Significantly different ($P<0.05$).

Figure 5 Chromosome Conformation Capture (3C) experiments reveal interactions between the -20kb enhancer and the intronic-, 3'UTR region of the *Tgm2* gene. (A) IGV genome browser representation of DNase hypersensitive regions on the *Tgm2* locus in mouse thymocytes. HindIII cutting sites are indicated all over the locus, along with the used restriction fragments for 3C analysis. BAIT represents the starting point of the analysis overlapping with the -20kb enhancer. (B) Gel (1.5% agarose) image of 3C-PCR products combining BAIT-specific with the +11kb-, +13kb-, +30kb-specific primers. Thymocytes were treated with the indicated compounds for 1 h before fixation. No ligase controls are also carried out and indicated (middle). Gel image (1.5% agarose) of PCR products using primers specific to chr10:74788584-74788803 genomic region as a loading control (bottom).

Supplementary figure 1 Finding the position in the DNA to which primers are designed to detect eRNA transcription from a putative enhancer.

Supplementary figure 2 Integration of Hi-C interaction data, CTCF ChIP-seq and DNase-seq from CH12.LX cells along with mouse thymocyte data for DNase-seq and ChIP-seq against CTCF, H3K4me1, H3K4me3 and H3K27ac. Contact matrix heatmap was generated using the Interactive Hi-C Data Browser available at http://www.3dgenome.org. Hi-C data was generated by Rao et al. from the CH12.LX mouse B lymphoblast cell line. Data is shown with 5kb resolution on the indicated genomic region (top). IGV browser representation of ChIP-seq and DNase-seq experiments from the indicated cell types. CH12.LX-derived datasets are available under accessions: GEO:GSM912909 (CTCF ChIP-seq), GEO:GSM1014153 (DNase-seq) The *Tgm2* locus is marked with dashed lines and the motif direction of the two CTCF anchors (black arrows) are also shown used by us to predict the functional domain of the gene in which the regulation might take place (bottom).

Supplementary figure 3 Enhancer RNA levels and RXR, SMAD4, P300, CBP and CREB binding to an enhancer of the neuron specific *Prmt8* gene, -1.7 kb far from the TSS region. (A) Enhancer RNA levels determined by RT-qPCR analysis on the *Prmt8* -1.7 kb region in the presence of the indicated molecules in *in vitro* cultured thymocytes after 1 h treatment. (B) RXR, SMAD4, P300, CBP and CREB binding on the *Prmt8* -1.7 kb region determined by ChIP-qPCR analysis in the presence of the indicated compounds after 1 h treatment. Background was determined by using rabbit IgG and subtracted from the values presented on the graphs. Compounds used were: 0.3 μM 9cRA, 5 ng/ml rTGF-β, 100 μM dbcAMP, 1 μM AM580, 100 nM LG268, 100 nM AGN194310. CTR: vehicle treated control. Data represent mean±SD of triplicate determinations.