

**SHORT THESIS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY (PhD)**

**Investigation of the function of Retinoid X Receptor in mouse  
bone marrow-derived macrophages using novel genomic  
approaches**

**Bence Dániel**

Supervisor: Prof. Dr. László Nagy



UNIVERSITY OF DEBRECEN

DOCTORAL SCHOOL OF MOLECULAR CELL AND IMMUNE BIOLOGY

DEBRECEN, 2014

# Investigation of the function of Retinoid X Receptor in mouse bone marrow-derived macrophages using novel genomic approaches

By **Bence Dániel**

Supervisor: Prof. Dr. László Nagy, PhD, DSc, member of HAS

Doctoral School of Molecular Cell and Immune Biology, University of Debrecen

Head of the **Examination Committee:** Prof. Dr. János Szöllősi, PhD, DSc

Members of the Examination Committee:

Dr. Péter Nagy, PhD, DSc

Dr. Ágota Apáti, PhD

The Examination takes place at Department of Biophysics and Cell Biology, Faculty of Medicine, University of Debrecen; at 13 PM; 10<sup>th</sup> of November 2014.

Head of the **Defense Committee:** Prof. Dr. Gábor Szabó, PhD, DSc

Reviewers: Dr. Péter Bay, PhD

Dr. Eric Soler, PhD

Members of the Defense Committee: Prof. Dr. Iannis Talianidis, PhD, ext. member of HAS

Prof. Dr. László Tora, PhD, ext. member of HAS

The PhD Defense takes place at the lecture hall of In Vitro Diagnostic Building, Faculty of Medicine, University of Debrecen; at 10 AM, 18<sup>th</sup> of November, 2014.

## **Introduction**

### **Gene expression regulation**

The basic mechanisms of gene expression have been described more than 60 years ago by the famous scientist duo, Francois Jacob and Jacques Monod. Their research group used prokaryotic systems and they were the very first to show that trans-acting elements are able to regulate the expression level of the neighboring gene. The relevance of these discoveries is quite huge because later on it has been recognized that the regulatory mechanisms described by them are supported the basics of transcription regulation in higher order organisms also. The observation that trans-acting factors are regulated at the level of transcription by other trans-acting factors by using the cis-elements embedded into the DNA. At the beginning investigators characterized these gene/cis-element networks in the close proximity of the gene of interest. The aims of the field are totally transformed with the appearance of genome-wide methods which offers a global picture about gene regulation.

Based on the experiments carried out by Jacob and Monod, the prokaryotic gene regulatory systems largely differ from the eukaryotic one. The difference is due to the location and the number of the above mentioned cis-regulatory elements. Cis-elements/sequences can be found in the DNA and holds the possibility to recruit transcription factors (trans-acting factor). In bacterial systems the investigation of these elements are not the most challenging if one can use the old molecular biology technics like promoter bashing linked to reporter systems, because the regulation of the operons are achieved mainly by proximal regulatory elements. Eukaryotic systems are much more complex than the prokaryotic, due to the fact that regulatory elements are dispersed in the genome. The limitations of the technology at that time and before the genomic era did not allow the identification of the long-range cis-elements. Altogether, a technological breakthrough was needed in order to find the needle in the haystack.

## **The technological breakthrough**

The technological advent came with the availability of next-generation sequencing. The first high throughput method which is able to recognize the changing transcripts was first used in 1982 aimed at detecting the level of 382 transcripts between normal and tumor tissues. Actually this was a filter paper spotted technology which is known as the ancestor of the well-known microarray methodology. Although the miniaturized microarray technology was first used to measure the expression of 45 Arabidopsis genes. These studies were the very first attempts to enlarge our view on gene expression and its regulation. As another example, the profiling of transcription factor binding sites (TFBSs) and histone modification patterns along with nucleosome positions have been also determined by linking Chromatin Immunoprecipitation (ChIP) to microarray technology. During the last several years, sequencing technologies have greatly evolved and almost fully replaced the microarray-based methods. With the ability to sequence tens of millions of reads (DNA sequences) in a parallel fashion, more applications, which could be only imagined before, became a reality. The linking of NGS (Next-Generation Sequencing) with simple molecular biology approaches proved to be very useful and effective and led to the immediate generation of genome-wide data in a number of areas including whole genome sequencing, mRNA sequencing and revealing MNase and DNase I hypersensitive sites. Our capabilities to detect intra- and interchromosomal interactions have also expanded quite a bit. The availability of technologies, such as 3C (Chromosome Conformation Capture), greatly increased our opportunities to construct the interaction map of cis-elements. However the biggest technological advent in the field was the development of Global Run-On sequencing (GRO-seq). The technology is based on the classical nuclear run-on by which we can determine the amount of freshly transcribed nascent RNA and by linking it to NGS it holds the potential to do it in a genome-wide manner and also the direction and position. The features of GRO-seq make it very useful to identify active enhancers, due to the fact that enhancers are transcribed into so called enhancer RNAs (eRNA), while regulating their target genes.

The above mentioned techniques alone were able to revolutionize the field, but soon it has been recognized that their combination is needed to understand those molecular mechanisms that fine tune the expression of genes thus contributes to the diversity of higher order organisms.

### **Diversity at a magisterial level**

Higher order organisms contain many different cell types which possess exactly the same genetic material, however their functionality is widespread. In the human body there are approximately 220 cell types with different cellular functions. What kind of driving forces are responsible for this enormous diversity? By using the above mentioned technologies we might have the chance to answer this question, but it is clear that their combination is needed to pinpoint the exact mechanisms behind this phenomenon which are probably linked to different cis-element usage.

### **ENCODING the cis-element landscape in human**

The explanation for the aforementioned diversity in cellular functions must be in the non-coding part of our genome which contains an enormous amount of cis-regulatory elements. Humans are estimated to have about 20000 protein-coding genes, but this covers only 1.5% of the entire genome. The other 98% accounts for intronic and intergenic regions, non-coding RNAs and short or long interspersed elements. In the ENCODE project several research groups aimed at revealing all the cis-acting elements in the human genome as a consortium. This initiative is the continuation of the Human Genome Project, but in this particular case the final aim was the delineation of functional DNA sequences that act at the protein and RNA levels in a given cell type. Using 1600 data sets from 147 different cell types they could show that 80% of the human genome has cis-element function. Based on chromatin modification maps they predicted that approximately 400000 enhancer elements and 70000 promoter elements are part of the human regulatory element toolkit.

The only question remained elusive is the most important one: How can a cell decide which differentiation program to choose and which cis-element to use?

## **Fine tuners of gene expression regulation, pioneering, bookmarking and the role of higher order chromatin structure**

In the ENCODE project the main aim was to identify all the possible cis-acting elements in the human genome. Each of the genes encoded in the human or mammalian genetic material is likely to have many individual cis-acting elements distributed across tens to hundreds of kilobases. These regulatory sequences act in collaboration to fine-tune gene expression in a highly tissue and signal specific manner. What kind of factors dictate the functionality of such elements in the genome leading to differential gene expression patterns in various cell types?

From a mechanistic point of view enhancer elements must be prepared before activation. It has been shown that the forkhead transcription factor A (FoxA) binds to inactive genomic regions and by remodeling the nucleosome and histone modification pattern, it is able to recruit other transcription factors leading to enhancer activation. This observation leads to the establishment of the pioneering factor concept. It is well-known that there are two main states of the chromatin: heterochromatin and euchromatin. The heterochromatin represents the transcriptionally inactive regions of the genome, while the euchromatin is an opened chromatin structure favors to transcriptional events. More publications in the field showed that these can be found in well-separated regions of the nucleus. The heterochromatin is located closer to the nuclear lamina, while the euchromatic regions are more likely to be in the middle of the nucleus. It seems that positioning in the nucleus has clear effect on the expression of the genes and the discovery of insulator elements helped us a lot to understand the basics of this system. Insulators are representing a new type of cis-regulatory elements and these sequences were identified based on their ability to block promoter-enhancer interactions. Later on the insulator binding transcription factor has been also identified and named as CTCF based on its recognition sequence CCCTC. After the discovery of CTCF it has been shown that its main function is the separation of transcriptionally active, passive gene domains as a border. Once its interaction with the Cohesin complex has been described which has indispensable functions in sister chromatid cohesion during the metaphase of the cell cycle, CTCF has been linked more clearly to the regulation of higher order chromatin structure. In our days more and more genome-wide data show that these

factors collaboratively regulate the conformation of the chromatin structure, thus implicated in gene regulatory processes.

Taken together, the cell type specific enhancer landscape is evolved through differentiation due to the appearance of cell type specific pioneer factors and then the transcriptional activity of these elements and their target genes are further regulated by the previously mentioned CTCF/Cohesin complex which is responsible for the shaping of the higher order chromatin structure. These are the most common components that are responsible for cell type specific gene expression profile.

### **Enhancers as the main components of the cis-element toolkit**

Enhancers are cis-elements harboring the ability to recruit various transcription factors and positively regulate their target genes. Their appropriate action is indispensable to maintain normal cellular functions. It has been shown also that these sequences are able to function from long distances and their localization can be up- or downstream relative to the target gene. Due to their widespread localization in the genome their identification and characterization remained impossible before the technological development. After the appearance of Next-generation sequencing and its combinations with other molecular biology approaches greatly improved our possibilities to identify them. Chromatin immunoprecipitation followed by sequencing was the first very effective method to visualize them genome-wide. These studies were unequivocally confirmed that these elements are mainly located to the non-coding part of the genome, very often covering long distanced from the putative regulated gene. The vast amount of ChIP-seq data more and more described the main features of enhancers, thus creating a consensus on the field that the chromatin modifications like H3K4me1, H3K4me2 and H3K27ac and P300 which is a coactivator are very good markers of enhancers.

The RNA- and GRO-seq have also revealed another important feature of enhancers, namely that in the immediate vicinity of the transcription factor binding sites, transcription occurs. We call these RNA molecules as enhancer RNAs (eRNA) due to their origin. Soon it has been shown also that their production upon activating stimuli which acts on the enhancer and the dynamics of the up regulation shows positive correlation with the

putative target gene. These observations helped a lot in those studies where the final aim was to identify the enhancer network for a given transcription factor and assign these to the regulated genes. The most interesting question regarding these molecules is whether they have regulatory capacity or just merely byproducts. More groups aimed at identifying their functional relevance. The most relevant studies showed that specifically degrading the eRNA negatively affected the target gene's expression. Another striking result was when these eRNA sequences were inserted into reporter vectors. Surprisingly, the reporter containing the core sequence responsible for transcription factor binding and the full length eRNA-coding sequence had the highest transcriptional readout. Inverting the eRNA-coding region supported the most striking result because changing the sequence of the eRNA diminished its enhancing effect, suggesting that the sequence of these short RNAs can be very important in the context of gene regulation. According to these results, it seems that eRNAs possess regulatory function, but the question is how.

It has been demonstrated that enhancers actively participating in looping with their targeted promoters transcribe higher level of eRNAs. Recently, a very important finding came to light showing that knocking down the eRNAs immediately next to ERα (Estrogen Receptor Alpha) bound enhancers reduced the enhancer/promoter interactions and resulted in a reduced expression level of the corresponding gene. Potentially, these ERα-mediated eRNAs are taking part in the modulation of looping, which was further supported by the fact that eRNAs could pull-down the subunits of the cohesin complex.

Overall these studies imply that at least in some cases eRNAs possess regulatory function and contributes to gene expression regulation. Clearly, further studies are needed to clarify their precise role in gene expression regulation.

### **Nuclear receptors as signal specific transcription factors acting on enhancer elements**

Nuclear receptors are ligand activated transcription factors which are able to activate or repress gene expression. Their natural ligands are small lipophilic molecules, like steroids, retinoids and various metabolites, which are able to freely diffuse through the



cell membrane. Functionally these proteins are playing very important roles in the body because they transmit the effects of hormones and metabolites to the genome.

In our study we focused on the Retinoid X Receptor (RXR). RXR is a unique nuclear receptor because it functions as an obligate heterodimerization partner of other receptors. It is expressed in every cell type and is required for post-natal life in mice. Its presumed main molecular function is to regulate the activity of a dozen or so nuclear receptors. There is also evidence that it can form homodimers and/or have heterodimer-independent signaling capacity. A key concept regarding RXR signaling is the permissiveness, non-permissiveness mutually exclusive dual paradigm. According to this, in certain heterodimers such as RXR:PPAR, RXR:LXR ligand activation of RXR results in transcriptional activation, hence these are permissive heterodimers, whilst in other heterodimers such as RAR:RXR, TR:RXR and VDR:RXR RXR is suppressed or “subordinated” and therefore these so-called non-permissive heterodimers cannot be activated from the RXR side. Therefore the activation of all permissive heterodimers present in a particular cell type might lead to pleiotropic gene activation and engagement of potentially conflicting pathways.

Ligand activation of RXR is believed to regulate distinct and coherent gene expression in various cell types and it has been shown to have beneficial pharmacological effects in various animal models such as increasing insulin sensitivity in diabetic animals and more recently selective activation of RXR in microglia and astrocytes has been shown to clear  $\beta$ -amyloid and reverse deficits in an AD mouse model although the extent of the improvement was recently questioned. These data support that ligand activation of RXR is biologically and therapeutically relevant. Therefore understanding the liganded receptor's genomic and cistromic activity is essential to further map its biological role and therapeutic potential.

In macrophages there are several heterodimeric receptors with key cellular roles such as PPAR $\gamma$  regulating oxLDL uptake and processing, LXR regulating cholesterol efflux and immune function, NR4A1 (NUR77) regulating inflammatory response. These heterodimeric receptors have been linked to the development of atherosclerosis and also immune function and provide means to reprogram macrophages. Therefore it is

biologically important to understand how activation of RXR contributes to these pathways and potentially to novel ones, and regulates gene expression in macrophages.

## **Aims**

1. Validation of the binding of STAT6 by CHIP-RT-QPCR on the enhancer elements of lysosomal genes in macrophages.
2. Determination of the effect of RXR activation at the level of the macrophage transcriptome.
3. Identification of the RXR bound genomic regions in macrophages by CHIP-seq.
4. Determination of the primarily RXR activated genes and the corresponding enhancer elements by GRO-seq.
5. Validation of enhancer/promoter interactions at the chromatin level by 3C methods.
6. Investigating the possible unique effect of RXR on gene expression regulation.

## **Methods**

### **Materials**

Ligands: LG268, LG1208 gifts from M. Leibowitz (Ligand Pharmaceuticals), RSG (Sigma), GW3965 was a gift from T. M Wilson (GlaxoSmithKline), AM580 (Sigma).

### **Differentiation of bone marrow-derived macrophages**

Bone-marrow was flushed from the femur of wild-type and RXR $\alpha$ / $\beta$  DKO C57Bl6/J male animals. Cells were purified through a Ficoll-Paque gradient (Amersham Biosciences, Arlington Heights, IL) and cultured in DMEM containing 20% endotoxin-reduced fetal bovine serum and 30% L929 conditioned medium for 5 days.

### **RNA-Seq**

RNA-Seq library was prepared from two biological replicates by using TruSeq RNA Sample Preparation Kit (Illumina) according to manufacturer protocol. Briefly, 2.5 $\mu$ g total RNA was used for the library preparation.

### **Real-Time Quantitative PCR**

RNA was isolated with Trizol Reagent (Molecular Research Center). For mRNA measurements, Tetro Reverse Transcriptase (Bioline) Kit was used and transcript quantification was performed by Quantitative Real-Time PCR reaction using SYBR green dye and Universal Mastermix (Diagenode). For eTranscript measurements RNA was reverse transcribed using High-Capacity cDNA Reverse Transcription Kit (Life Technologies) and quantification was performed with LightCycler<sup>®</sup> 480 SYBR Green I Master (Roche). Transcript levels were normalized to Ppia or Rplp0 in the case of eTranscript measurements.

### **ChIP-seq**

Approximately 40 million of cells were crosslinked with DSG (Sigma) for 30 minutes and then with formaldehyde (Sigma) for 10 minutes. After fixation chromatin were sonicated with Diagenode Bioruptor to generate 200-1000 bp fragments. Chromatin was immunoprecipitated with antibodies against pre-immune IgG (Millipore, 12-370), RXR (sc-774), P300 (sc-585), PU.1 (sc-352), CTCF (Millipore, 07-729), RAD21 (ab992), H4ac (Millipore, 06-866), H3K27ac (ab4729), H3K4me2 (Upstate, 07-030) and H3K4me3

(ab8580). Chromatin antibody complexes were precipitated with Protein A coated paramagnetic beads (Life Technologies). After 6 washing steps complexes were eluted and reverse crosslinked. DNA fragments were column purified (Qiagen, MinElute). The amount of immunoprecipitated DNA was quantified with Qubit fluorometer (Invitrogen). DNA was applied for QPCR analysis or library preparation. ChIP-Seq library was prepared with Ovation Ultralow Library Systems (NuGen) at least from two biological replicates according to manufacturer instructions.

### **GRO-seq**

Approximately 10 million of nuclei were isolated and Run-On reactions were performed for 5 minutes. RNA was then isolated by acid:phenol-chloroform extraction and base hydrolysis was performed for 20 minutes in the presence of 1N NaOH followed by DNase I treatment. Nuclear run-on RNA was immunoprecipitated with Br-dUTP antibody coated agarose beads and then washed to remove nonspecifically bound substances. Specifically bound RNA was eluted from the beads and end repair was performed in the presence of Tobacco acid pyrophosphatase and then in the presence of Polynucleotide Kinase followed by RNA extraction with acid:phenol-chloroform. For adapter ligation the RNA was incubated for 4 hours on the lab bench either in the presence of the 5'- or the 3'- adapter oligo. RNA enrichment was performed with the above mentioned anti-deoxy Br-UTP agarose beads. The eluted RNAs were then reverse transcribed and the remaining RNAs were degraded by the addition of a RNase mixture. Sequencing libraries were amplified with Phusion high fidelity DNA polymerase using the primers according to Illumina. Libraries were run on a non-denaturing 1xTBE, 8% acrylamid gel and fragments larger than 90 nucleotides were extracted from the gel and then libraries were then extracted, quantified and sequenced.

### **Chromosome conformation capture**

Nuclei were isolated in buffer containing 10mM Tris-HCl pH7.5, 10mM NaCl, 0.2% NP40 (Sigma), and protease inhibitor tablets (Roche). Chromatin was digested with 400U of HindIII (Fermentas) restriction enzyme at 37 °C for 16 hours and for an additional 1 hour with 100U. Chromatin fragments were ligated with 100U of T4 DNA ligase (Fermentas) at 16 °C for 4 hours. After ligation chromatin was de-crosslinked overnight at 65 °C.

Ligation products were column purified (Roche, High Pure PCR Template Preparation Kit) and DNA concentration was determined by Nanodrop. DNA fragments were submitted to QPCR reactions using TaqMan probes designed to the assayed enhancer region. Tandem primers were designed in the close proximity of the restriction enzyme cutting sites. BAC control DNA pools were used to determine primer efficiency in each analyzed genomic region and GAPDH was used as a loading control.

### **3C-sequencing**

After the first digestion and ligation the 3C DNA pool was purified with phenol/chloroform/ isoamyl alcohol (25:24:1) (Sigma). Second restriction digestion was performed by using DpnII (NEB) for 16 hours per manufacturer's instruction. Second ligation was performed at 16C for 6 hours with 200U of T4 DNA ligase. DNA was then purified again with phenol/chloroform/isoamyl alcohol (25:24:1) followed by QIAquick gel purification column (Qiagen) purification. Bait specific inverse PCRs were performed using primers coupled to Universal Illumina adapters and Barcode sequences. Reactions were purified by QIAquick gel purification columns. Amplicon libraries were quantified and qualified by Agilent using DNA 7500 chip cartridge. Amplicon libraries were sequenced on Illumina MiSeq and HiSeq2000 sequencer.

### **Reporter construct preparation**

Enhancer sequences were PCR amplified from BACs or genomic DNA with overhangs on the forward (atataagctt-) and reverse (tataggatcc-) primers, digested with BamHI and HindIII and cloned into pUC18 HSV TK-LUC. BACs were ordered from BACPAC Resources Center.

### **Transient transfection**

Transient transfections were carried out as previously described [92]. Briefly, COS1 cells were transfected with reporter constructs along with plasmids encoding  $\beta$ -galactosidase and full-length receptors in triplicates using polyethylenimine. Six hours after transfection, cells were exposed to ligands. After 48 hours of incubation luciferase activity was determined by Luciferase Assay System (Promega) and normalized to  $\beta$ -galactosidase activity.

## **ELISA**

ELISA experiments were carried out based on manufacturer instructions (R&D Systems). Cells from three biological replicates were counted and plated. After ligand exposure, supernatants were collected and centrifuged. Clear supernatants were applied to the ELISA plates. Unbound substances were removed by washing and the specific polyclonal antibody was added to the wells. Unbound complexes were eliminated by washing, then substrates were added to the wells and enzyme activity was determined.

## **Chorioallantoic membrane (CAM) assay**

Fertilized white leghorn chicken eggs (Wyverkens, Halle, Belgium) were incubated for 3 days at 37°C before removing 3ml of albumen to detach the shell from the developing CAM. After making a window in the eggshell to expose the CAM, it was covered with cellophane. Eggs were then further incubated until day 9, when the test conditions were applied. Sterile absorbable gelatin sponges (1-2 mm<sup>3</sup>; Hospithera, Brussels, Belgium) were impregnated with 3x10<sup>4</sup> BMDM treated with 100nM LG268 for 6 hours and then placed on the CAM. LG268 alone was used as a negative control and recombinant hVEGF-A<sub>165</sub> was used as a positive control. The windows were again covered and the incubation continued until day 13, when angiogenesis was assessed. Membranes were fixed with 4% paraformaldehyde at room temperature for 2h. A large area around the contact or insertion sites was removed, placed on a glass slide, and photographed under using a Zeiss Lumar V.12 stereomicroscope. Digital images were captured using an AxioCam MRc5 and processed with Axiovision 4.5 Software (Zeiss). To determine the number of blood vessels, a grid containing three concentric circles with diameters of 4, 5, and 6 mm was positioned on the surface of the CAM and all vessels radiating from the sample spot and intersecting the circles were blindly counted under a stereomicroscope.

## **Statistical tests**

QPCR, ELISA and CAM assay were presented as means +/-SD. We made at least three biological replicates and we performed paired (two-tailed) t-tests and results were considered significant with  $p < 0.05$ . Interaction frequencies of the chosen genomic

regions were considered significantly different with  $p < 0.0001$  according to the unpaired (two-tailed) t-test.

## **Results**

### **STAT6 regulates the expression of many lysosomal genes during alternative macrophage activation**

In this study we used gene expression correlation analyses to identify transcription factors which might regulate lysosomal function. We found STAT6 as the most significant hit, which is signal dependent transcription factor involved in the regulation of immune system. The background of such analysis is that if there is a transcriptional regulator which coordinated action is responsible for the expression of a group of lysosomal genes, it might be possible to identify such a factor by using correlation analyses across a great number of microarray data.

Clustering of all known lysosomal genes led to the determination of several subgroups whose expression appears to be coordinated. Importantly STAT6 was correlated with the biggest lysosomal cluster consisting of known acidic hydrolases and vacuolar H<sup>+</sup> ATPase subunits. Importantly, 14 out of the 15 subunits building up the vacuolar H<sup>+</sup> ATPase were shown to be regulated by STAT6 and three subunits (*Atp6v0a1*, *Atp6v0d2* and *Atp6v1b2*) were among the most strongly induced lysosomal genes by IL-4/STAT6. Taken together, in this study we assigned a new transcription factor to the lysosomal gene panel in macrophages which can fine tune their expression. The data presented here identified a novel pathway that is implicated in the lysosomal function during alternative macrophage activation.

### **The transcriptional readout of the activated RXR in murine macrophages**

In order to study the effect of the RXR activation in macrophages we set out to systematically dissect the genomic events following RXR liganding in bone marrow-derived macrophages. We used a selective synthetic RXR agonist throughout these studies. First we identified the variable transcripts by RNA-Seq and found that selective activation of RXR affects hundreds of genes at the steady state mRNA levels. In order to identify the immediate early responder targets to RXR activation we wanted to analyze

and determine in detail both the genomic binding events of RXR (RXR cistrome) and the enhancers where the receptor acts and responsible for the altered transcription.

### **The liganding of RXR results in enhanced RXR binding on preformed PU.1 marked cis-regulatory elements and leads to P300 recruitment and modification of surrounding chromatin structure**

We performed ChIP-seq experiments to reveal the RXR, PU.1 which is the master regulator of myeloid cell development and P300 cistrome in the absence and presence of the activator ligand (LG268). On the other hand, we also investigated the chromatin modification pattern of these cells, thus we carried out ChIP-seq experiments against the well-known active chromatin markers: H3K27ac, H3K4me2 and H4ac.

We identified approximately 5200 RXR binding sites. Interestingly the we could not observe major redistribution in the RXR cistrome after one hour ligand activation, however the receptor's occupancy increased at its predetermined binding regions. Next we performed a *de novo* motif analysis under the detected peaks which showed us PU.1 as the most enriched motif followed by various nuclear receptor binding motifs. Investigation regarding the receptor's genomic localization revealed that RXR almost exclusively (almost 90% of the receptor binding sites) bind to intergenic regions. The combination of the ChIP-seq results showed us that more than 65% of the RXR binding sites are overlapped with the lineage determining transcription factor PU.1, while P300 dynamically recruited to RXR upon ligand activation. On these P300 bound sites the local chromatin acetylation increased on the H4 histones, while did not show any change regarding H3K27ac or H3K4me2. With the help of these experiments we identified the components of the RXR cistrome and the changes on the surrounding chromatin environment in response to ligand activation. Our initial aim was to link these binding regions to the regulated genes, but we were not able to do it in a reliable way due to the localization of the RXR sites.

### **Annotation of the RXR cis-element toolkit to the regulated genes**

As a next step, we wanted to assign the detected RXR bound cis-elements to the regulated genes. Unfortunately the above mentioned methods did not allow the easy



identification of enhancer/gene pairs because the RXR binding sites are often located more than 100kb far from the putative regulated genes. In order to solve this issue, we used a novel technology called GRO-seq which has the ability to map the transcriptionally engaged polymerase products or nascent RNAs. The fact that transcription can be also detected on the enhancer regions and the observation that their immediate induction profile upon an activating stimuli positively correlates with the expression profile of the regulated genes offered a plausible solution for our annotation problem. The combination of ChIP-seq, GRO-seq and RNA-seq helped us to identify more than 400 cis-elements to which we could assign more than 250 RXR regulated genes.

Out of the 400 cis-elements, we characterized 45 in reporter systems, but we could also measure the freshly synthesized eRNAs on several of them by QPCR method which suggested the existence of RXR selective enhancers. Next we wanted to validate our gene/enhancer pairs with the so called 3C assay. We have chosen 3 genes and their enhancer regions and interestingly we could validate all of them. Surprisingly the addition of the RXR ligand further strengthens the interaction frequency between the enhancers and the genes. Coupling 3C to sequencing we could show also that probably the RXR enhancers and their targeted genes are able to co-localize in the so called transcription factories.

### **Activation of RXR regulates and angiogenesis network in macrophages**

Using the gene set to which we could assign the enhancer elements, we performed an interaction network analysis which revealed the existence of an angiogenic gene network. We identified more crucial factors with angiogenic potential for instance: Vegfa (Vascular endothelial growth factor), Hbegf (Heparin-binding EGF-like growth factor) and Litaf (Lipopolysaccharide-induced TNF factor). Next we wanted to test these macrophages in an in vivo relevant angiogenic assay.

We performed our experiments to measure the Vegfa gene and protein product and first we could show that RXR activation induced the level of the gene. Later on we could also confirm these finding at the protein level using ELISA experiments. By using RXR $\alpha\beta$

knock out macrophages we unequivocally showed that the angiogenic phenotype of macrophages is depending on the presence of the receptor. Finally, we tested the angiogenic capacity of macrophages in an *in vivo* relevant angiogenesis assay, namely the Chorioallantois membrane assay. We were able to show that RXR programmed macrophages are more angiogenic than the non-treated and the receptor knock out macrophages. Altogether, these experiments supported evidence that RXR activates an angiogenesis related gene program which shifts macrophages toward a pro-angiogenic phenotype.

## **Discussion**

The most important and desired aim regarding signal specific transcription factors is to determine its action on the genome in order to reveal its main biological roles in a given cell type. Here we introduced a combined approach which can serve as a proof of concept to dissect the roles of transcription factors. We showed that transcriptomics data integrated with cistromics data is not enough to reliably link the RXR bound cis-elements to the regulated genes, due to the fact that they are widespread in the non-coding part of the genome. Therefore we used the GRO-seq technology which helped us to generate a much clearer picture about the enhancer/gene pairs and lead to the identification of a novel RXR function in bone marrow-derived macrophages. We successfully confirmed this finding in an *in vivo* relevant angiogenesis assay. Based on our results, RXR seems to be a relevant drug target in tumor biology and regenerative medicine.

## Summary

In our study we focused on the action of the activated RXR nuclear receptor in bone marrow-derived macrophages. Taking advantage of the technological developments we determined the action of the receptor in a genome-wide manner.

First, we identified the regulated gene network by using RNA-sequencing, then we performed chromatin immunoprecipitation followed by deep sequencing in order to find the receptor bound genomic loci. Using these methods we found approximately 800 regulated genes and 5200 binding sites for the receptor. Interestingly, almost 90% of the RXR sites are located to intergenic regions, often more than 100kb far from the closest genes thus their annotation to the regulated genes remained elusive. By using global run-on sequencing we analyzed the amount of the freshly synthesized nascent RNA on gene bodies and as reported previously also the transcripts originated from cis-elements known as enhancer RNAs. Based on the fact that enhancer RNAs are induced in a very similar manner as their targeted genes we assigned them to their potential targets, yielding the transcriptionally active RXR cis-element/gene pair network. The combination of the aforementioned methods let us to identify 252 target genes to which we could assign 414 cis-acting RXR elements. To our surprise, these genes showed a convergent angiogenic network consisting well-known angiogenesis inducers, for example: *Vegfa*, *Hbegf*, *Litaf*. In order to characterize the RXR programmed macrophages in the context of angiogenesis we carried out the well-accepted chorioallantoic membrane assay, which unequivocally shed light on the improved angiogenic capacity of macrophages harboring the activated RXR program. Taken together, these results show that the RXR regulated gene/cis-element network promotes the establishment of a pro-angiogenic macrophage phenotype.

Our highly integrated method is capable of identifying the regulated genes and their cis-element network for any kind of signal specific transcription factor thus able to reveal new biological functions in a transcription factor and cell type dependent manner.



Register number: DEENKÉTK/261/2014.  
Item number:  
Subject: Ph.D. List of Publications

Candidate: Bence Dániel  
Neptun ID: FVHSAG  
Doctoral School: Doctoral School of Molecular Cell and Immune Biology

### List of publications related to the dissertation

1. **Dániel, B.**, Nagy, G., Nagy, L.: The intriguing complexities of mammalian gene regulation: How to link enhancers to regulated genes. Are we there yet?  
*FEBS Lett.* 588 (15), 2379-2391, 2014.  
DOI: <http://dx.doi.org/10.1016/j.febslet.2014.05.041>  
IF:3.341 (2013)
2. **Dániel, B.**, Nagy, G., Hah, N., Horváth, A., Czimmerer, Z., Póliska, S., Gyuris, T., Keirsse, J., Gysemans, C., Van Ginderachter, J.A., Bálint, B.L., Evans, R.M., Barta, E., Nagy, L.: The active enhancer network operated by liganded RXR supports angiogenic activity in macrophages.  
*Genes Dev.* 28 (14), 1562-1577, 2014.  
DOI: <http://dx.doi.org/10.1101/gad.242685.114>  
IF:12.639 (2013)
3. Brignull, L.M., Czimmerer, Z., Saidi, H., **Dániel, B.**, Villela, I., Bartlett, N.W., Johnston, S.L., Meira, L.B., Nagy, L., Nohturfft, A.: Reprogramming of lysosomal gene expression by interleukin-4 and Stat6.  
*BMC Genomics.* 14 (1), 1-20, 2013.  
DOI: <http://dx.doi.org/10.1186/1471-2164-14-853>  
IF:4.041





### List of other publications

4. **Dániel, B.**, Bálint, B.L., S. Nagy, Z., Nagy, L.: 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, 2014.  
DOI: [http://dx.doi.org/10.1007/978-1-4939-1346-6\\_2](http://dx.doi.org/10.1007/978-1-4939-1346-6_2)
5. Brázda, P., Krieger, J., **Dániel, B.**, Jonás, D., Szekeres, T., Langowski, J., Tóth, K., Nagy, L., Vámosi, G.: Ligand binding shifts highly mobile RXR to chromatin-bound state in a coactivator-dependent manner as revealed by single cell imaging.  
*Mol. Cell. Biol.* "accepted by publisher" (2014)  
DOI: <http://dx.doi.org/10.1128/MCB.01097-13>  
IF:5.036 (2013)
6. Nagy, G., **Dániel, B.**, Jonás, D., Nagy, L., Barta, E.: A novel method to predict regulatory regions based on histone mark landscapes in macrophages.  
*Immunobiology.* 218 (11), 1416-1427, 2013.  
DOI: <http://dx.doi.org/10.1016/j.imbio.2013.07.006>  
IF:3.18

**Total IF of journals (all publications): 28,237**

**Total IF of journals (publications related to the dissertation): 20,021**

The Candidate's publication data submitted to the iDEa Tudóstér have been validated by DEENK on the basis of Web of Science, Scopus and Journal Citation Report (Impact Factor) databases.

11 September, 2014



