

## Summary

RNA polymerase II mainly transcribes protein-encoding genes, and the transcriptional activity of many of those genes is tightly regulated. Two cooperating components are key in regulating RNAP II transcription. One component acts *in cis* and requires an element that is located on the same DNA molecule as the transcribed gene, the so-called *cis*-regulatory DNA elements. The other component acts *in trans*, i.e. it requires molecules separate from the gene containing DNA molecule, here referred to as *trans*-acting factors.

Despite the fact that *cis*-regulatory DNA elements have different functional definitions, there are two features that they have in common: they can be detected as DNase I hypersensitive sites (HSs) *in vivo* and they are relatively small DNA fragments (200-300bp) that contain multiple binding sites for certain *trans*-acting factors and their interacting partners. *Trans*-acting factors are a diverse group of proteins (and RNA), which includes sequence-specific transcription factors, components of the general transcription machinery, and co-factors. Many *trans*-acting factors can interact and as such influence each other activities.

How do *cis*-regulatory DNA elements and *trans*-acting factors collaborate efficiently to regulate RNAP II transcription? Furthermore, *cis*-regulatory DNA elements somehow need to communicate with each other. However, many gene loci have widely separated regulatory elements that all appear to contribute to proper transcriptional regulation. One important aspect of (bio)chemistry is that the reaction rate is proportional to the concentration of molecules involved. An economical way to bring the different components together might be through compartmentalization.

Indeed, the nucleus of eukaryotes, which holds the genetic material (DNA), is organized into many distinct compartments each dedicated to drive particular biochemical processes. The DNA of the nucleus is packaged into chromatin, which has constrained mobility and consequently a particular gene and its regulatory sequences are limited to a tiny portion of the nucleus. In contrast, *trans*-acting factors diffuse 'freely' and can roam the entire nuclear space. A rapid flux of proteins on/off macromolecules (e.g. chromatin) and in/out of compartments generates a dynamic but stable nuclear organization.

In this thesis the  $\beta$ -globin loci of man and mouse were used as model systems to study their transcriptional regulation in the context of the living nucleus. These loci consist of several genes and *cis*-regulatory DNA elements. Transcription of the genes is regulated in a tissue- and developmental stage-specific manner. One of the *cis*-regulatory DNA elements is the Locus Control Region (LCR), which is thought to create an open chromatin conformation and activate transcription of the genes. A hypothesis that explains this activity is the looping model. This model presumes that the LCR directly and stably interacts with individual genes, thereby activating it. The intervening chromatin between LCR and the gene is thought to loop out.

In chapter 3, this model was tested on the mouse  $\beta$ -globin locus *in vivo*, using a novel technique called Chromosome Conformation Capture (3C). The results showed that in erythroid cells the LCR and actively transcribed genes were in close spatial proximity, although they are separated by 40-60 kb of intervening DNA. Inactive genes located between the LCR and active genes did not interact, but looped out. In nonerythroid cells, which do not transcribe globin genes, no spatial interactions were found between the LCR and any of the genes. Interestingly, there were additional *cis*-regulatory DNA elements that joined the LCR-active gene interactions. The upstream 5' HS-62.5/60.7 participate in this interaction, again with the intervening DNA looping out. At the other end of the locus the 3' HS1 is also involved in the contacts, but we have no evidence for DNA looping out between the genes and 3' HS1. We propose that clustering of *cis*-regulatory DNA elements is essential for creating and maintaining active chromatin domains and regulating transcription. We refer to this clustering as the 'active chromatin hub' (ACH).

In chapter 4, we examined whether the ACH is formed at other developmental stages and if its formation is evolutionary conserved between the mouse and human loci. The 3C results showed that during different stages of development an erythroid-specific spatial cluster of *cis*-regulatory elements and active genes is formed in the  $\beta$ -globin cluster. The core of this spatial cluster consists of the locus control region (LCR) and distal regulatory elements, a so-called core ACH. This core ACH is structurally conserved during development, while a switch occurs in globin genes interacting with the core ACH, correlating with the switch in their transcriptional activity. This spatial organisation of the  $\beta$ -globin gene cluster is conserved between mouse and man. Furthermore, we found in erythroid progenitor cells that the locus adopts a conformation that differs from brain and erythroid cells. The distal hypersensitive regions and the 5' side of the LCR interact, while the inactive genes and the 5' HS1-3 of the LCR seem to loop out. These data strongly suggest that the *cis*-regulatory DNA elements surrounding the  $\beta$ -globin genes create an erythroid-specific developmentally stable nuclear compartment dedicated to RNA polymerase II transcription. This spatial clustering of regulatory elements results in a high local concentration of DNA binding sites for cognate transcription factors and their interacting partners, which consequently accumulate at the site. In chapter 5, we discuss several speculative properties of the ACH that may explain its role in efficient transcription of the globin genes.

Globin genes are transcribed alternately rather than at the same time and transcription can switch back and forth between genes in a flip-flop mechanism. This implies that interactions between the different components of the ACH are a dynamic rather than static process. However, the 3C measurements represent steady-state average levels. Consequently, the dynamic interactions cannot be detected. In chapter 2, we describe an approach to monitor these dynamic interactions in living cells. We 'marked' the human  $\beta$ -globin locus with a dual tagging system, using Lac<sup>O</sup>/ LacR-cyan fluorescent protein (CFP) in conjunction with a Tet<sup>O</sup>/TetR-yellow fluorescent protein (YFP). This work is still in progress.