2 PURPOSE & AIMS

Sequencing the human genome and subsequently a number of other organisms has been a tremendous step towards understanding the mechanisms of life. However, the sequences in themselves shed little light on questions of differentiation and specification as the DNA is the same in nearly all cells and remains so throughout life. It has been suggested to view the sequence itself more as an encyclopedia, where different chapters (transcriptional units) are read depending on the biological requirements of a given time point. In this comparison the histone modifications have been suggested to play the role of the indexing system which makes the required chapters readily accessible. Depending on the type of modification an open euchromatic or compact heterochromatic state can be induced. Histone-modifying enzymes place and remove these bookmarks thereby facilitating differentiation and making adaptations to changing environmental influences possible. The opening of the chromatin structure exposes certain stretches of the DNA sequence which are recognized by transcription factors which then either repel or recruit the polymerase that finally reads the transcriptional unit.

In this allegory the mechanism of transcriptional regulation appears simple and straightforward. However, it is becoming ever clearer, that within and between the different levels of regulation a complex interplay is taking place and many fundamental questions remain unanswered. Understanding of transcriptional regulation is not only of academic interest, but malfunctioning of any of the involved factors can result in serious disease.

The observation that histone modifications are key determinants in transcriptional regulation led to the formulation of the *histone code hypothesis* by Strahl and Allis in 2000[78]: 'distinct histone modifications on one or more tails act sequentially or in combination to form a *histone code* that is, read by other proteins to bring about distinct downstream events'. It has since been debated[36,63,223-225] to which degree modification combinations may encode distinct read-outs and serve as signaling marks. In this context it is of particular interest, whether they are placed prior to transcription, indicating a signaling function, or during the process. Transcription factors (TFs) generally have short binding motifs which frequently occur in the genome and only few of the binding sites are actually functional. Moreover, the choice of binding sites is frequently cell-type-dependent. A possibility to direct cell specificity are steric changes in the TF induced by cofactor binding. The recent discovery of TFs which specifically recognize individual histone modifications may be a further mechanism.

The purpose of this study was to gain insight into these fundamental questions regarding histone modifications and transcription factors and understanding their regulatory

networks. Due to the vast number of factors involved the investigation was focused onto a small subset of four histone modifications and four transcription factors. Moreover, the study was concentrated on heart and skeletal muscle. The heart was chosen as cardiac malformations are the most frequent defect in live births occurring at an incidence of 1% and are associated with high morbidity and mortality. Therefore, understanding of transcriptional processes in this organ is particularly relevant. To understand cell-type specific differences in histone modification pattern a closely related cell type was to be included in the study. For this purpose skeletal muscle cells were chosen, as these have the additional advantage of being able to differentiate in cell culture. By means of ChIP-chip and expression arrays in these three cell types this set-up allowed the characterization of the combinatorial relationship between the four modifications and transcription and to follow changes during differentiation.

The histone modifications to be investigated were selected by a number of different criteria. Firstly, modifications should all be described to be equally associated with higher transcript levels in yeast[25,226,227] and higher eukaryotes[228-230]. This allowed investigation of combination effects without the possible confounding factor of a dominant repressive mark. Additionally, the confinement to activating modifications allowed quantification of the association with transcript levels and monitoring of gradual changes due to combinations. As the investigation of combination effects is feasible only for modifications often occurring at the same chromosomal location the investigation was further limited to modifications reported to frequently colocalize[25,226-230]: acetylation of histones 3 and 4 (H3ac and H4ac) as well as di- and trimethylation of lysine 4 on histone 3 (H3K4me2, H3K4me3).

An ever growing number of transcription factors has been linked to cardiovascular disease. To decide on a comprehensive set such transcription factors were chosen which were considered to be clinically relevant: The TFs should be evolutionarily highly conserved and mutations were shown to result in severe cardiac malformations, as known both from animal models and mutation analyses in patients. Formation of the heart is undoubtedly regulated by a number of interlocked regulatory networks. To obtain one comprehensive subnetwork it was further required that between the TFs at least one regulatory interaction is known. This resulted in the choice of Gata4, Mef2a, Nkx2.5 and Srf. The function of these TFs was to be investigated in cardiomyocytes by ChIP and siRNA knockdown. As the four chosen histone modifications were to be analyzed in the same cell line the TF data was to be integrated with the information on histone modifications. Thereby conclusions regarding the interplay between histone modifications and transcription factors might be gained.