1. Summary

Natural nucleic acids are the molecules of life. They contain the genetic information of every organism and are involved in many biological processes. Additionally, they were involved for the development of life on earth. All nucleic acids contain natural modified nucleotides, which are incorporated in complex processes and are of tremendous importance for regulation and accuracy of transcription and translation in every organism. Over 120 different modifications were detected in all natural nucleic acids and the number is still increasing. tRNA is the most heavily modified nucleic acid and it additionally contains the broadest structural variety of modifications. So far, the function and role of single tRNA modifications has been investigated. Natural modified nucleosides were only rarely investigated in a context based manner. To overcome this limitation, we developed a method which allows quantification of in principle all tRNA modifications in parallel. This is similar to *proteomics* and *metabolomics* research in which the complete proteome or metabolome of a cell is quantified in parallel, respectively.

The method is based on HPLC-ESI-MS analysis. It enables precise quantification of tRNA modifications by using isotope-labeled stable internal standard molecules for each modification (*Figure 1*). In addition, the method was extended to allow quantification of DNA modifications.

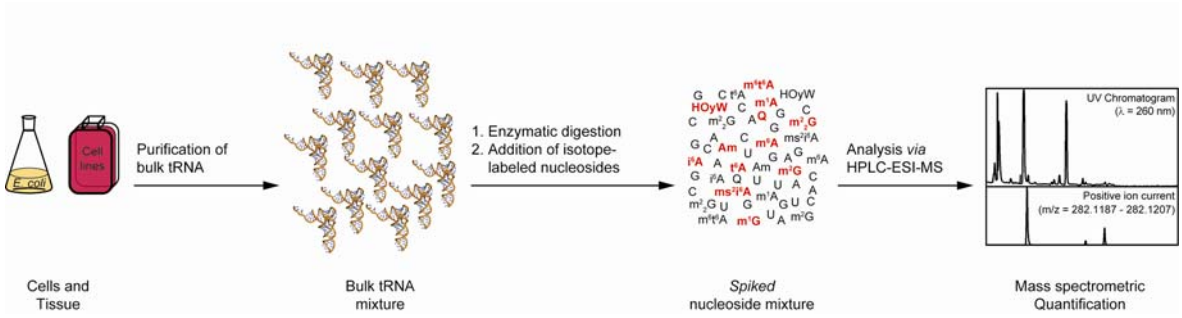

Figure 1: Representative workflow of the LC-MS based quantification method for tRNA nucleosides which is applicable for all kind of cells. Extension of this method allows quantification of DNA, rRNA, and mRNA modifications.

The method starts with isolation and purification of tRNA or DNA, followed by complete enzymatic digestion of tRNA or DNA to the nucleosides. Afterwards the isotope-labeled nucleosides are spiked to the natural nucleoside mixture, which is analyzed *via* HPLC-ESI-MS experiments (*Figure 1*). In these measurements, the specific mass area of the natural

modification is compared with the area of the added corresponding isotope-labeled nucleoside.

Modified tRNA nucleosides

Availability of isotope-labeled reference compounds is essential for accurate quantification of the corresponding natural occurring modified nucleoside. Therefore, in total 13 nucleosides including six isotope-labeled derivatives were synthesized and used in this thesis work (*Figure 2*).

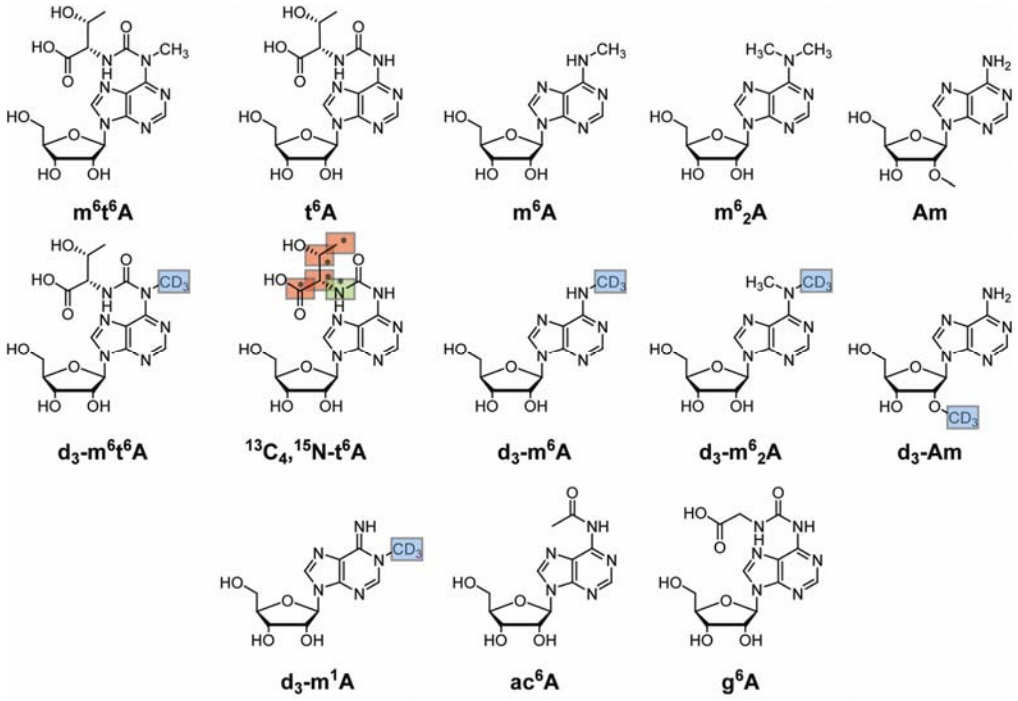

Figure 2: The 13 modified nucleosides synthesized in this Ph.D. thesis work.

These six isotope-labeled nucleosides as well as eleven other modifications synthesized in the group of Prof. T. Carell were used as internal standards for quantification of tRNA modifications in different bacterial species, mammalian tissues, and human cell lines.

In a first project we identified quantitative differences between healthy tissues and cancer cell lines especially for the modified tRNA nucleosides t^6A and ms^2i^6A (*Figure 3A*). The modification t^6A is upregulated in all cell lines compared to the liver tRNA values. In addition, the mitochondria specific modification ms^2i^6A could not be detected in any cancer cell line, but is present in significant amounts in liver tRNA. The absence in cancer cells is attributed to the Warburg effect, which describes an impaired mitochondrial activity in tumors. Furthermore, we analyzed the ms^2i^6A content in different tissues and found high

levels in tissues with high mitochondrial activities (*Figure 3B*). The data correlate well with the mitochondria specific Cytochrome C oxidase activity (*Figure 3C*). In summary, the ms^2i^6A content represents mitochondrial activity and can be used as marker to differentiate between healthy and tumor tissues.

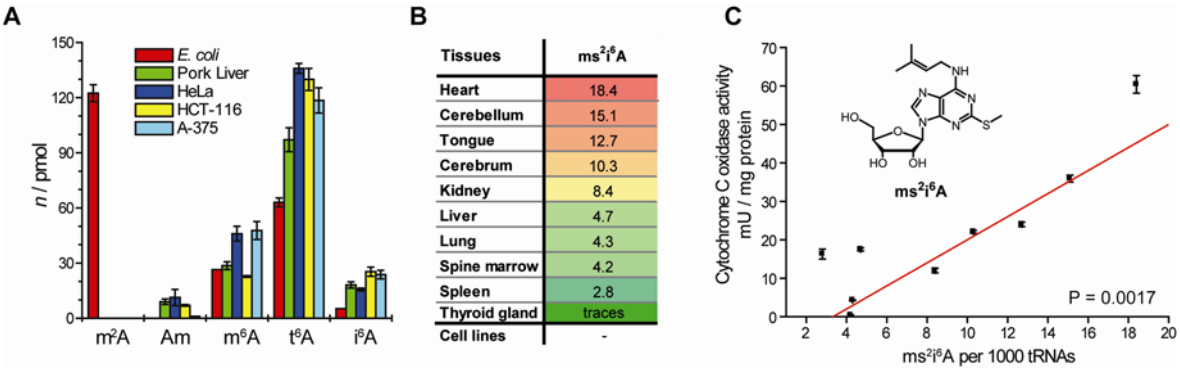

Figure 3: A) Differences of tRNA modification levels between *E. coli*, liver tissue and three cancer cell lines. B) Tissue dependency of ms^2i^6A. C) Correlation of ms^2i^6A with Cytochrome C oxidase activity.

The tRNA modification levels of 11 further nucleosides were determined in 10 different cell lines and 10 different tissues (*Figure 4A*). These results revealed that tissues contain significantly different modification levels. tRNA from liver and cerebellum is most heavily modified, whereas heart and cerebrum exhibit lowest modification levels. These results are in line with *in vivo* protein synthesis rates from literature. Therefore, an *in vitro* translation system was established and used in this thesis to further support this hypothesis. High *in vitro* translation activity correlated with the modification level (*Figure 4B*). All cancer cell lines showed similar or higher levels than liver as the most heavily modified tissue. This can be explained by the high proliferation rates of cancer cells, which necessitates high protein synthesis rates (*Figure 4A*).

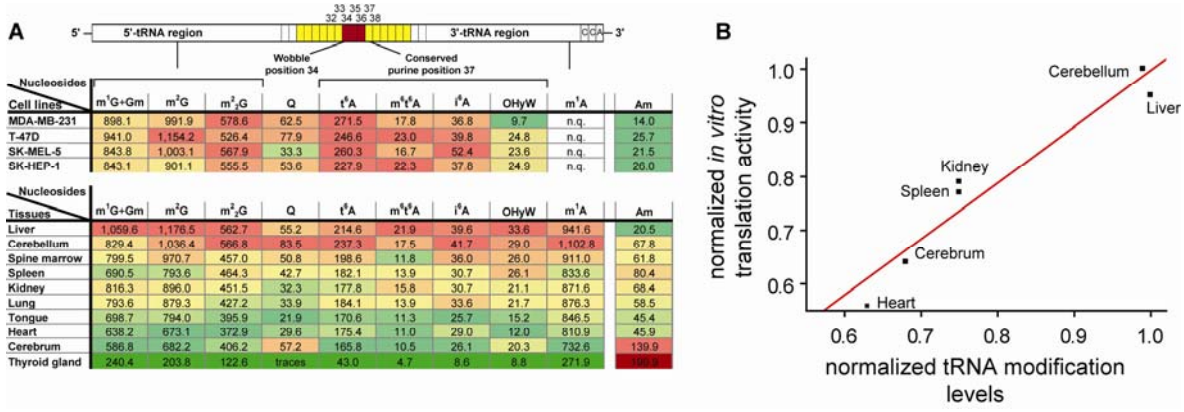

Figure 4: A) Quantitative data of cancer cell lines and tissues colored according to the amount. B) Correlation of total tRNA *in vitro* translation activity with normalized nucleoside levels.

In a further project the tRNA modification levels of 11 different prokaryotic species were investigated. Analysis of bacteria from different parts of the phylogenetic tree revealed large tRNA modification level differences. With these data, combined the mammalian tissue values and data from two yeast strains, we performed a Cluster correlation analysis. This analysis yielded clustering according to phylogenetic correlations (*Figure 5A*). Eukaryotic and prokaryotic organisms cluster separately from each other and the two yeast strains are differentiated from mammalian tissues. Furthermore, Gram-positive and Gram-negative bacteria are clearly separated from each other and even bacteria from the same genus can clearly be differentiated. These results show the high accuracy of our quantitative data and hint at an evolutionary controlled development of tRNA modifications.

In addition, the response of the tRNA modification pattern to external stimulation was analyzed in *E. coli*. Indeed, variations in the tRNA modification values were detected after applying different pH stress conditions or antibiotic treatment (*Figure 5B*).

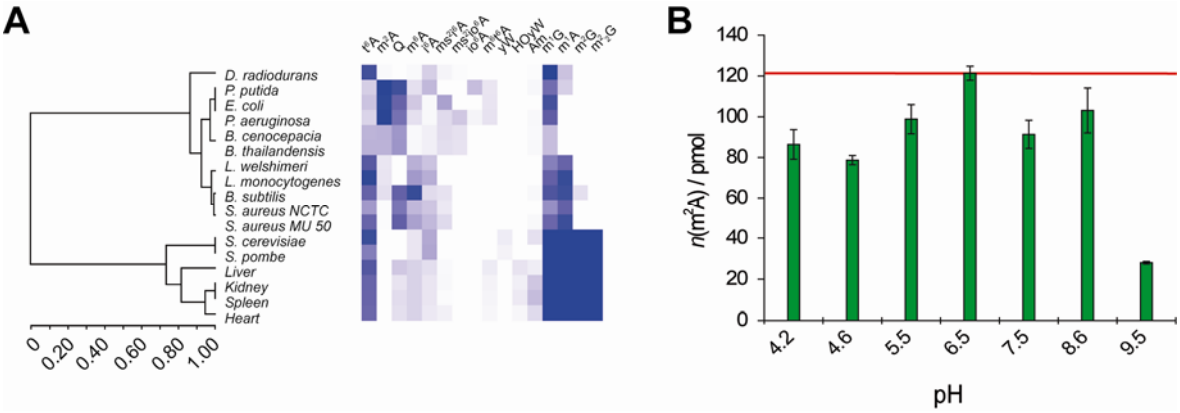

Figure 5: A) Cluster analysis of determined quantitative tRNA modification data for prokaryotes and eukaryotes. B) tRNA modifications variations depending on pH stress.

The sixth DNA base 5-hydroxymethylcytosine

Methylation of cytosine at the 5-position (mC) is an epigenetic marker, which is known for many decades. It is of high importance to block expression of specific genes. In 2009, the modification 5-hydroxymethylcytosine (hmC) was detected as a novel base in purkinje neurons of the mammalian cerebellum. Enzymes of the *Tet* family were identified to convert mC to hmC. These observations indicate that hmC has an epigenetic role, which is not clarified yet.

Using our quantification method, we analyzed the hmC and mC content in the mammalian body. Analysis of different tissues revealed that hmC is distributed over the whole

mammalian body with significantly varying amounts from 0.03% to 0.7% hmC/dG depending on the tissue type (*Figure 6A*). However, mC values are constant around 4.2%. Interestingly, tissues from the central nervous system (CNS) contain the highest amount of hmC. Medium values were found for tissues like kidney, heart, and lung. Lowest values could be detected in liver, spleen and pituitary gland, which is located in the brain. All these results indicate an important role of hmC in the nervous system. Therefore, we analyzed the mammalian brain in more detail and again found strongly varying values of hmC in different brain regions (*Figure 6B*). An interesting fact is that regions with a high cognitive role (cerebral cortex, hippocampus) contain high hmC values. Furthermore, an age dependency for the hmC values in hippocampus could be shown (*Figure 6C*).

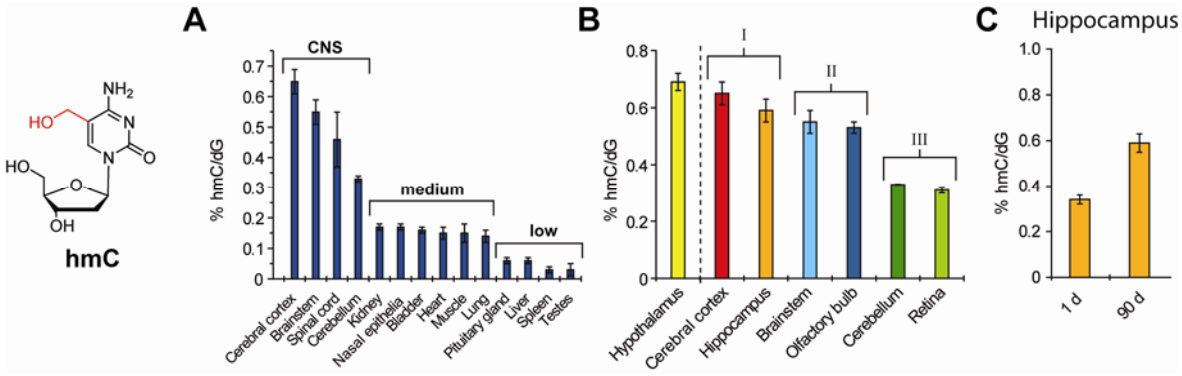

Figure 6: A) hmC distribution in the mammalian body. B) hmC distribution in different brain regions. C) Age dependency of hmC in hippocampus.