ABSTRACT

The *O*-GlcNAc modification is, akin to phosphorylation, an abundant modification which plays an important role in cellular processes. The addition of *O*-GlcNAc modification to proteins is regulated by *O*-GlcNAc transferase (OGT). This enzyme is ubiquitously expressed in mammals. The human OGT spans approximately 43 kb of genomic DNA, and to date 3 OGT isoforms have been sequenced (two nuclear and one mitochondrial isoforms). In mammals OGT exist in different isoforms, such as in rat and mouse which contain two different isoforms differing by 30 nucleotides in their *N*-terminal (exon 2) . In this study the mouse OGT isoforms were investigated. In addition, the expression of the full OGT transcript and various transcripts was studied by real-time PCR in mouse tissues (liver, heart, kidney, testis, lung, spleen, ear, stomach, intestine, bone, tail and brain). The highest level of full length mouse OGT was found in spleen followed by testis and lung, though human OGT already have been documented to have highest expression in pancreas. Further investigations showed that mouse expresses both isoforms (one isoform which uses both exon 2A and 2B and one isoform which skips exon 2A and only uses exon 2B) to produce OGT. The first isoform showed highest expression in heart and aorta and the second isoform in mouse liver and lung, suggesting that mouse OGT exist in different isoforms with varying abundance in tissues.

Post-translational modifications (PTMs) are the major regulators of protein biological functions. PTMs are catalyzed by their respective enzymes which are sequence specific. Depending on the environment and signaling contexts, proteins are modified and instantly perform their function. In this study 4 different modifications consensus sequences

were investigated by utilizing the bio-informatic tool MAPRes. This tool mines association rules of a modified residue in peptides. Phosphorylated Ser/Thr/Tyr showed a polar sequence environment with Pro at various positions; *O*-GlcNAc modified Ser/Thr occurs in a polar environment with vicinal Val (-1 position) and Pro (-2 and -3 positions); acetylated Lys occurs in a basic environment with a preference for His or Tyr at +1 and Ser at +7 positions. Methylated Lys also shows a preference for basic amino acids, but compared with acetylated Lys, which have a high preference for vicinal Lys, methylated Lys shows a higher preference for Arg. Methylated Arg shows a high preference for Gly both up- and dowmstream in the peptide chain. In addition Yin Yang sites sequence environment was also investigated and showed that Ser/Thr Yin Yang sites are located in a polar environment with Pro at various positions. These results suggest that Pro is a very important amino acid, which is located nearby modified amino acids. Furthermore the position of Pro may be a determining factor in deciding whether a residue is phosphorylated, *O*-GlcNAc modified or both (Yin Yang site).