

PROTEOME ANALYSIS IN BIOLOGY

LENKA HERNYCHOVA, EVA MICHALOVA, BORIVOJ VOJTESEK

Regional Centre for Applied Molecular Oncology, Masaryk Memorial Cancer Institute,
Zlutý kopec 7, 656 53, Brno, Czech Republic

ABSTRACT - The term proteome was created from two words protein and genome in 1994 by Marc Wilkins and depicts the complete set of proteins expressed by a genome. Proteomes are studied by proteomics which involves different experimental procedures estimating structures and functions of proteins. Mass spectrometry represents the main tool used in modern proteomics. This method enables us to identify, quantify, and characterize large sets of proteins present in biological samples. Advanced procedures connected with mass spectrometry are successfully used in cancer research.

PROTEOME ANALYSIS NEEDS GENOME DATA

Gregor Johann Mendel published the first genetic study on pea plants in 1866. This Mendelian inheritance of biological characteristics was later applied to a variety of organisms and enabled a partial explanation of biological evolution. Since the discovery of DNA molecule structure (WATSON *et al.*, 1953) big progress was made through the investigation of chromosomes, processes concerning DNA during cell life, genes, genetic code and sequencing of genomes. This time is known as *genomic era*. In 2001, the project for human genome sequencing (Human Genome Project) started to better understand the human body and its disorders. The project was completed in 2004 offering information into 20,000–25,000 genes in human beings (International Human Genome Sequencing Consortium, 2004). Nevertheless, the knowledge was not enough for the characterization of human as a biological system. It was already known that one gene expresses more than one protein and therefore, the study of sets of proteins in a given organism under defined conditions has begun. In 1994 Marc Wilkins (WILKINS *et al.*, 2009) used for the first time the term *proteome* and defined it as all the proteins expressed by the genome of a cell, a tissue, or an organism. This time, known as *post genomics era*, opens the way to study gene expression and protein function and to build the relationship between genes and their proteins.

WHY IS THE PROTEOME MORE COMPLEX THAN THE GENOME?

Proteins are synthesized on the basis of the information saved in the DNA sequence. Transcription initiation at alternative promoters, differential transcription termination, and alternative splicing of the transcript are mechanisms that generate different mRNA transcripts from a single gene. The human transcriptome is estimate to count over 100,000 transcripts (Fig. 1). Moreover, proteins are usually modified after the translation step (posttranslational modification, PTM) and these modifications increase the functional diversity of the proteins by the addition of functional groups or proteolytic cleavage of