In order to understand the legal treatment of 3-D protein structure-related claims, a thorough understanding of basic proteomic concepts is necessary. This chapter therefore provides a brief introduction into the scientific background of the subject. Since proteomics is a rapidly growing and dynamically changing field, it is of course unrealistic to provide a complete and exhaustive treatment. Instead, the focus will be on issues that are indispensable as a background and relevant from the point of view of intellectual property rights.

After defining the term "proteomics", the role of proteins in biological organisms will be reviewed, with special emphasis on theories of amino acid structure and protein folding. Since many pharmaceutical applications of proteomics deal with specific folding details, the concepts of "structurally similar, sequence dissimilar proteins" as well as the basic idea of "posttranslational modifications" have to be introduced. In section C, the role played by genetic information in the shaping of proteomic structures will be assessed. This subject is of prime importance for two reasons. First, from a biological point of view, recent proteomic research has significantly changed the conceptual treatment of protein encoding. In particular, the close association of genetic code and protein functionality has become increasingly blurred. Second (and closely related), this may have important implications for the legal treatment of proteomics, since questions of patent dependency have to be evaluated in light of the relative importance of genetic information. The chapter closes with a description of the most important proteomic research techniques. The diversity of active research areas shows how dynamic the field of proteomics is, and provides a sense of the types of issues confronting the patent system.

A. Definition of the Term

"Proteomics" is derived from the term "proteome", which was first used in 1994 during a scientific conference in Siena, Italy. At that time, following rapid advances in analytical techniques, it had become possible for biochemists to identify and to examine many new proteins. Consequently, the possibility for large-scale protein studies seemed attainable.[15] The proteome was defined as the total set of proteins expressed in a given cell at a given time, the study of which is termed 'proteomics'[16].

15 See Patterson, Scott D./Aebersold, Ruedi H., Proteomics: The First Decade and Beyond, 33 Nature Genetics Supplement 2003, 311, 314.

16 See Dove, Alan, Proteomics: Translating Genomics into Products? 17 Nature Biotechnology 1999, 233. A comprehensive glossary of biotechnological terms and definitions is provided

'Proteomics' conjures up two distinct but interdependent associations. First, it refers to the general analysis of proteins. Here, the objective is to gain insights into the composition, function and further development of protein structures. This analysis is carried out against the backdrop of the effects that protein structural changes can have on biological organisms.[17] Second, the term was coined to make an analogy with genomics, to indicate proteomics' potential to become the major "next step" of biotechnological analysis.[18] 'Proteomics' particularly focuses on the complex relations between proteins and gene sequences, taking into account that the specific function of the genome can only be determined with knowledge of the genome's product, the protein. Starting from the encoding of proteins by the genome, proteomics can therefore also be defined as the systematic study of proteins, with the aim of understanding the whole and detailed function of gene sequences.[19] Proteomics aims to provide information about (a) the conditions under which predicted gene products are translated, (b) the timing of the translation, and (c) the extent of 'post-translational' modifications, i.e. changes to the structure of proteins not *directly* related to the genetic code and the process of translation. It is worth noting that none of these elements is necessarily predicted by the nucleotide acid sequence alone.[20] Consequently, one of the major aims of proteomics is to identify the forces that determine the exact structure of gene products apart from the genetic code.[21]

It is worth mentioning that the term 'proteomics' is sometimes used differently, depending on the context. In the scientific community, it is used very broadly, encompassing everything from protein characterization techniques (such as mass spectrometry and two-dimensional gel electrophoresis) to anything remotely related to the quantitative determination of proteins. At the same time, biotechnology firms engaged in any kind of protein analyses often describe themselves as "proteomic firms", using the term as a cachet to signal attractiveness for potential investors.[22]

on the website of the "Human Genome Project Information", at http://www.ornl.gov/. See also Patterson, Scott D./Aebersold, Ruedi H., Proteomics: The First Decade and Beyond, 33 Nature Genetics Supplement, 311, 314. Hall provides a slightly modified definition of "proteomics", as "the science and technology of cataloguing and describing the behavior of all the proteins encoded in a particular organism's genome", see: Hall, Stephan S., Revitalizing Drug Discovery, Technology Review October 2003, 39, 44.

17 Patterson, Scott D./Aebersold, Ruedi H., Proteomics: The First Decade and Beyond, 33 Nature Genetics Supplement, 311, 314.

18 http://www.wikipedia.org/wiki/Protemics, last checked on January 22, 2008.

19 Patterson, Scott D./Aebersold, Ruedi H., Proteomics: The First Decade and Beyond, 33 Nature Genetics Supplement, 311, 314; Mullner, S/Neumann, T./Lottspeich, F., Proteomics–A new Way for Drug Target Discovery, 48 Arzneimittelforschung 1998, 93.

20 Humphery-Smith, I., Blackstock, W., Proteome Analysis: Genomics via the Output rather than the Input Code, 16 Journal of Protein Chemistry 1997, 537.

21 The proteomic analysis of complete complements of proteins encompasses not only the identification and quantification of proteins, but also the determination of their localization, modifications, interactions, activities, and function, see Fields, Stanley, Proteomics in Genomeland, 291 Science 2001, 122f, 122.

22 Dove, Alan, Proteomics: Translating Genomics into Products?, 17 Nature Biotechnology 1999, 233.

Finally, the term 'functional proteomics' should be introduced, as it refers to the 3-D structure determination of all proteins encoded by the genome of key organisms, a major focus of this study. The major goal of functional proteomics is the analysis of protein structures by an integrated approach combining computer-based technologies of bioinformatics and the in-depth analysis of 3-D protein structures through physical methods, such as nuclear magnetic resonance (NMR) spectroscopy or x-ray crystallography (see below).

B. Proteins and the biological organism

Proteins[23] support every aspect of biological activity.[24] Through their structural stability, diversity, and chemical reactivity, proteins influence and enable most of the key processes associated with life. They operate as catalysts[25], provide mechanical support and immune protection, transport and store other molecules such as oxygen, cause movement[26], transmit nerve impulses, and direct growth and differentiation.[27]

I. Amino acid sequences

In order to understand the functioning of proteins one must be aware that the term "protein structure" refers to three distinct levels of organization: primary, secondary, and tertiary. The primary structure refers to the amino acid sequence as such. The secondary structure describes the conformation or spatial relationship adopted by local regions of the polypeptide chain. Finally, "tertiary structure" expresses the entire folding of the polypeptide chain.[28]

23 The origin of the word "protein" is usually attributed to Jöns Jakob Berzelius (1779-1848) and has been ascribed to derivation from the Latin word primarius, or from the Greek word for "first thing" (in Greek πρωτεϊνη = first element), see Whitford, David, Proteins – Structure and Function, Chichester, West Sussex, England, 2005, 1.

24 Whitford, David, Proteins: Structure and Function, Chichester, West Sussex, U.K., 2005, 9.

25 The term catalyst refers to substances that accelerate chemical reactions.

26 Schwaiger, Ingo/Sattler, Clara/Hostetter, Daniel R./Rief, Matthias, The Myosin coiled-coil is a truly elastic Protein Structure, 1 Nature Materials 2002, 232.

27 Berg, Jeremy M./Tymoczko, John L./Stryer Lubert, Biochemistry, New York, NY, 2005, 41.

28 Whitford, David, Proteins: Structure and Function, Chichester, West Sussex, U.K., 2005, 81. The term "quaternary structure" further refers to a certain association of multiple 3-D folded proteins to form multi-subunit complexes.