QSAR methods applied to biomolecules.

Gabriele Cruciani, DCBB, University of Perugia

Research today is focusing on biomolecules. In fact, the genome, the genetic information heritage of our organism, contains only the mold for making biomolecules, but the real actors who do most of the work in cells are precisely proteins.

And are always the proteins that make it possible to distinguish the different types of cells. Although all cells essentially have the same genome, they differ according to which genes are active and which proteins are produced. Similarly, diseased cells often produce absent proteins in healthy ones or vice versa. It is therefore no coincidence that researchers are trying to catalog all human proteins and find out how they interact with each other. The aim is to soon be able to conceive better drugs with fewer side effects. However, achieving this goal will not be so simple as the study of proteins is even more difficult than that of genes.

The science that studies the cataloguing, expression, quantification and de-orphanization of proteins is called proteomics. Proteomics is also a great business and the global market for proteomic tools, reagents, software and services reached about $25 billion in 2020. And this estimate does not include the revenues generated by the drugs and diagnostic kits that will be developed thanks to proteomics studies. Investment in research and development in this area reached $100 billion in 2010, and about 60 bio-drugs are approved today, though many are in the process of being approved.

In this scenario, artificial intelligence methods (AI) are extremely useful both in speeding up search results and in reducing their costs. However, AI methods must be created, nurtured, managed and applied wisely. Too often artificial intelligence is used by staff who are poorly prepared to understand its potential and limitations. This paper aims to provide the approaches developed in more than 30 years of research in the field, providing examples and showing applications in the field of industrial research.