

Proteomics: understanding the recipes for life

The genome

Imagine owning a recipe book that has nothing but an index. Imagine further that this index has all the page numbers, but only a few of these page numbers are actually accompanied by the name of a recipe.

This is the state of the art in the world of genomics, where researchers are determining the genetic makeup of living things. The “index” is the map of a genome, or full set of genes. Genes are contained in chromosomes, which are inside the nucleus of nearly all living cells. They are made of deoxyribonucleic acid, or DNA, and carry the information needed to build and maintain an organism.

So far, researchers have mapped the genome of several different organisms, including a yeast, a worm, a fruit fly and a few dozen bacteria. In December 2000, a concerted effort by researchers around the world yielded the full genome map of *Arabidopsis thaliana*, a weedy relative of the canola plant that is a valuable model for plant science research. Another international effort, the Human Genome Project, has produced the first “rough draft” of the more than 100,000 genes that contain the instruction set for *Homo sapiens* – human beings.

Working molecules

It's a long way from having the index of a book to knowing the recipe on every page. This is where the study of proteins – proteomics – comes in.

Genes don't do anything by themselves except hold the blueprints for proteins. In addition, while there are in the order of 100 thousand genes, there are probably millions of different proteins. They come in a multitude of different forms, from muscle tissue to enzymes and hormones that regulate body functions.

Chemically, a protein is a complex molecule, composed of carbon, hydrogen, oxygen and usually, sulphur. Some proteins also incorporate phosphorous, iron, zinc and copper. These basic elements combine into molecules called amino acids, which are the building blocks of proteins. There are 20 amino acids found in most animals (including people), and plants. In plants, however, in addition to their 20 basic amino acids, over 250 non-protein amino acids can be present. These non-protein amino acids, also known as secondary metabolites are intermediates in the synthesis or breakdown of protein amino acids. Many of these secondary metabolites are thought to play a role as defensive agents.

To make a protein, amino acids are linked together in a chain, and each link has a different function. Some amino acids act as scaffolding to hold the molecule together, others act to keep the protein dissolved in water, and some do the chemical reaction work of the molecule. This work may be sending a hormonal message, breaking down other molecules as an enzyme, or joining together with other molecules to build structures like plant seeds or animal muscles.

Getting to know the proteome

While a genome map tells where the genes are, it doesn't fully say what proteins they make, or what the proteins are supposed to do. Scientists do know, however, that proteins must have a specific shape to work properly. An antibody protein, for example, is folded in just the right shape to bind onto the surface of a pathogen, like a virus or bacteria. If the antibody protein is the wrong shape, it won't fit, and the invader will be free to go about its business of making the host organism sick.

Sometimes, a mistake in the genetic code itself will cause faulty proteins to be made. These misfolded proteins can also be introduced from the outside. Alzheimer's disease, cystic fibrosis, mad cow disease and many types of cancer are believed to be caused by misfolded proteins.

Proteins also constantly change, folding themselves into new shapes depending on what the cell is doing. This folding can be almost instantaneous, happening in a mere millionth of a second.

Understanding the proteome – the full set of proteins for an organism – is a daunting task, but the potential rewards are great. Most modern medicines, for example, are either proteins or work by binding to proteins. In agriculture, a better understanding of how proteins work could lead to the development of crops with increased frost and drought tolerance, or a host of other valuable traits.

Researchers use a range of tools to study and catalogue proteins. One of the most valuable is the computer, which can be used to store vast databases of protein information, as well as to analyze this data. Modelling programs can also help predict how a protein will fold under various conditions. Using this information, researchers can learn the shape of each protein and get an idea of what it does.

Protein research on your home computer

One of the great challenges of proteomics is the vast quantity of information that scientists have to handle. Even powerful and expensive supercomputers can be overwhelmed by the sheer volume of information that must be processed.

Protein researchers have taken an idea from the Search for Extraterrestrial Intelligence (SETI) to help solve the problem. The SETI@home project takes advantage of a huge under-used resource: the 300 million or so computers connected to the Internet, most of which sit idle 90 per cent of the time.

SETI asked Internet users to load their screen saver program, which would process radio telescope data while people were away from their computers. Over two million users have downloaded the SETI software. Together, they have processed information that would have taken a single computer over 400,000 years.

Scientists at Stanford University have launched a similar "distributed computing" program to tackle proteomics. As of December 2000, the Folding@home project had over 10,000 participants, with the numbers growing daily.

Further reading

For more information about proteomics, see the Folding@home Web site at <http://foldingathome.stanford.edu>. The site offers a good primer on proteins and how they work, as well as the Folding@home software you can use to help researchers understand these remarkable and complex molecules. Software for Linux and Windows based machines is already available, and versions for Mac OS and various flavours of UNIX are in the works.

To find out more about agricultural biotechnology, contact:

Ag-West Biotech Inc.

101-111 Research Drive

Saskatoon, Saskatchewan

Canada S7N 3R2

Telephone: (306) 975-1939

Fax: (306) 975-1966

E-mail: agwest@agwest.sk.ca