The blossoming field of proteomics promises better ways of managing and treating serious ailments in livestock and humans. **Steve Davidson** reports.

pioneering proteins

A remarkable renaissance is taking place in protein science. Commentators are talking about 'biology's biggest boom industry'. This stems largely from the explosion of data generated by the renowned Human Genome Project and associated advances in biotechnology tools and equipment.

Genes, after all, essentially consist of recipes for making proteins in cells, but while we now know that humans have about 30 000 genes in all, scientists reckon we have somewhere between 200 000 and two million proteins. A few years ago the estimate was just 100 000. The old textbook axiom that 'one gene codes for one protein' has been discarded.

So while cataloguing the full complement of genes in humans was an impressive achievement, doing the same for our proteome, the sum of all human proteins, will be a Herculean effort. And the same can be said for plants, animals and microbes.

The sheer weight of numbers is not the only challenge. Genes basically remain unchanged from year to year, but proteins are moving targets. Proteins in cells vary between tissue types and with an individual's age, gender, health and so on. Just drinking a cup of coffee or a glass of wine alters our proteome!

What's more, whereas DNA has just four chemical bases (abbreviated to A, C, G and T), proteins are assembled from 20 different amino acids. Genes tell the cell which amino acids to string together to form a given protein.

To further complicate matters, proteins fold into all sorts of shapes, and scientists are still working on new ways to classify proteins according to these contortions.

What then is fuelling the tremendous interest in protein structure and function, especially in private enterprise? The short answer is pharmaceuticals.

Defective or missing proteins can cause serious diseases in humans, including cancer, Parkinson's and Alzheimer's disease, diabetes, psoriasis and arthritis. This is also true for many diseases affecting livestock and other organisms.

The goal of much protein research is to develop new or better drugs that act directly on abnormal proteins that cause disease, or to use newly identified proteins as drugs to counter disease.

For pharmaceutical companies, the ultimate prize is a lucrative patent on a protein, while pure research may simply reveal how proteins form various tissues, or how they interact and network together as enzymes and hormones to run the body.

Proteomics

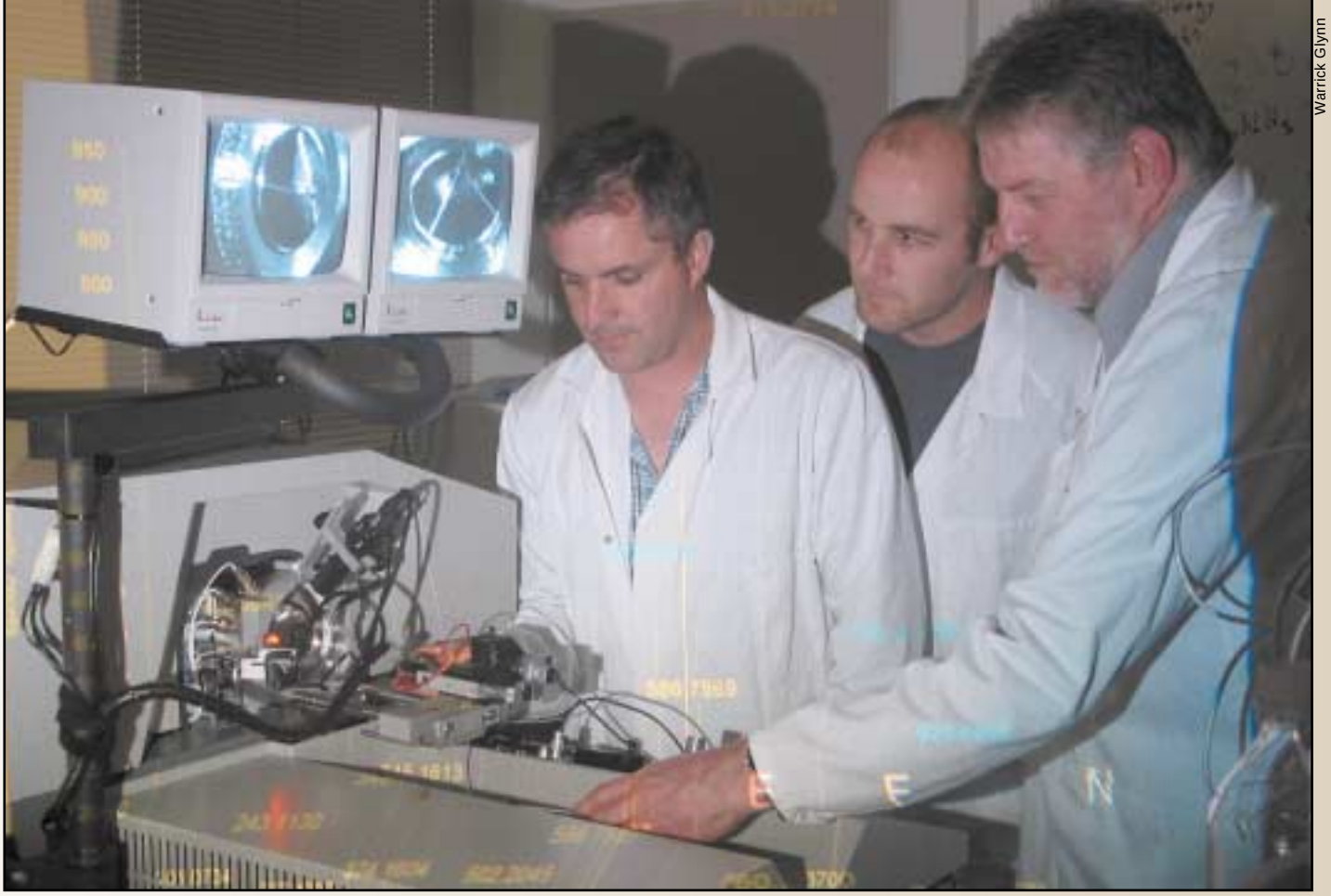
If the proteome is 'the protein complement encoded by a genome', the science of 'proteomics' must involve the study of those proteins. But pinning down a definition isn't easy.

In its strictest sense, proteomics, a term coined in 1994, is the protein equivalent of genomics: an approach that involves deciphering the full complement of proteins in an organism, organ, tissue or cell, using rapid, high-throughput analysis. At the other extreme, however, it has come to mean just protein science: the general study of protein structure and function.

Proteomics is really best thought of as the formation of enormous databases of information about a proteome that can later be interrogated.

For example, such a database will allow comparisons between healthy and cancer-affected tissues, or perhaps normal cells and those affected by a toxin.

The former could identify a protein that is overproduced in cancer-affected tissues and close study of the protein structure should allow the design of drugs to block the activity of the harmful protein in the cancer. This is just one of many possible applications for proteomics.



Warrick Glynn

Tristan Wallis, Dean Whelan and Jeff Gorman of CSIRO Health Sciences load a sample into the Q-star hybrid quadrupole/TOF mass spectrometer. This instrument is a powerful tool in proteomics for analysing peptide sequences and modifications. Researchers at CSIRO are working in several areas of proteomics and in related aspects of protein science.

Scientific American journal says most scientists agree that proteomics can be broken down into three main activities: identifying all the proteins made in a given cell, tissue or organism; determining how those proteins form cooperative networks; and investigation of the precise three-dimensional structures of proteins to identify targets for drugs, that is, ways to switch their activities on or off.

Researchers at CSIRO are working in several of these areas of proteomics and in related aspects of protein science.

Dr Jeff Gorman heads the proteomics group in CSIRO Health Sciences and Nutrition, based in Parkville, Victoria, but he will soon relocate to Brisbane to coordinate the proteomics activities of CSIRO Livestock Industries as leader of the Proteomics Group of the Institute for Molecular Bioscience.

His group is particularly interested in what happens to proteins after they are translated, or assembled. For example, proteins can split into shorter chains, or can be modified when other molecules latch onto them, changing their function.

Gorman says CSIRO and other Australian research institutions can make an important contribution to proteomics, but we risk falling behind in the global race. He speaks with some envy of a Japanese proposal to spend US$88 million on a single, five-year proteomics initiative.

In the face of such massive competition, it will be essential to form effective partnerships between Australian groups and to integrate into well-funded international initiatives.

It was a partnership between Gorman's group and Dr Murray Whitelaw's group at the University of Adelaide's Department of Molecular Biosciences that led to a breakthrough on the regulation of the activity of a protein known as Hypoxia-Inducible Factor.

Publication of this work has led to an international race to develop ways of inhibiting development of blood vessels, which are necessary for the development of tumours.

The breakthrough could not have been made using a genomics approach alone, and highlights the necessity of proteomics for the comprehensive study of biological systems.

Structural science

Another area of protein science making great strides is structural biology.

Dr Jose Varghese, leader of Structural Biology at CSIRO Health Sciences, says his group, with Professor Richard Simpson's team at the Ludwig Institute for Cancer Research, has resolved the structure of a protein receptor involved in several human cellular responses.

Interleukin-6 is a vital hormone of the immune system, playing a pivotal role in triggering our response to injury or disease, such as the inflammation of tissues.

But abnormally high levels of the hormone are associated with diseases ranging from multiple myeloma and various types of cancer, to rheumatoid arthritis and cardiovascular disease.

The scientists are endeavouring to control this undesirable response by interfering with the receptor system for the hormone, that is, its method of recognising and attaching to cells that are sensitive to the hormone.

If they can work out the precise atomic structure of the proteins in this signalling system, the scientists should be able to identify targets for therapeutic drugs against the diseases.

Picturing proteins of potential

CHARACTERISING all the proteins, or genes, in an organism generates vast amounts of data in quick time. The total worldwide pool of biological information doubles every 14 months!

The job of dealing with these huge databases is known as bioinformatics: the application of information technology, statistics and mathematics to huge volumes of such complex data to address biological problems.

Dr Mervyn Thomas of CSIRO Mathematical and Information Sciences says the search for new pharmaceuticals is a two-step process.

First one identifies potential protein targets for drug activity based on genetic and proteomic data. Candidate drug compounds are then screened for biologically significant activity on those targets in a number of stages.

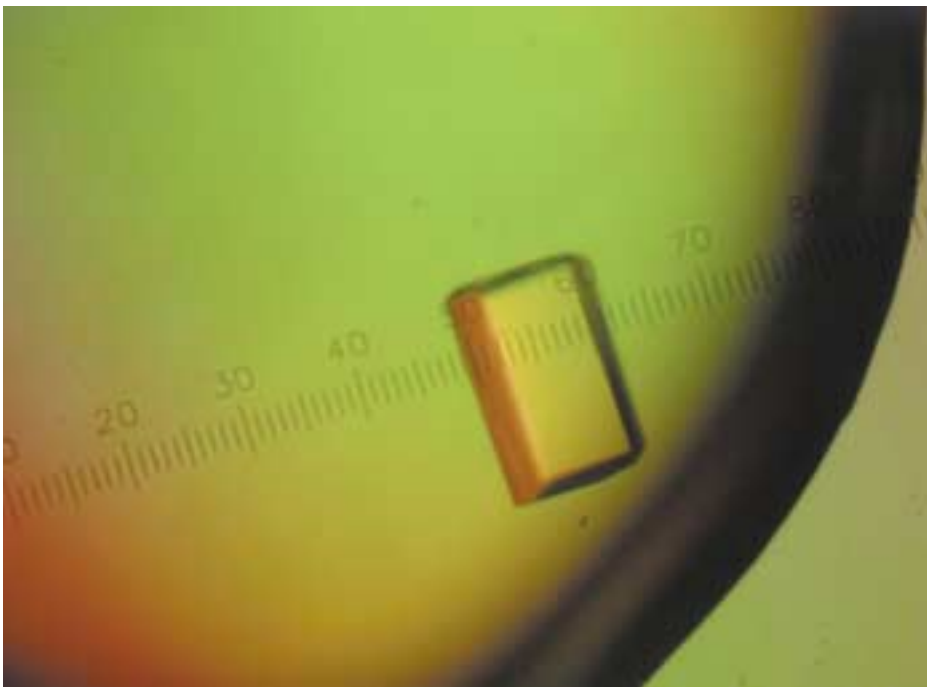
CSIRO is working with Proteome Systems Ltd (PSL) on several projects that involve applying CSIRO-developed image analysis and bioinformatics technologies to identifying promising protein targets for drug discovery

For example, two-dimensional gel image analysis is being used to identify and characterise the many forms of a particular protein encoded by a gene.

CSIRO has developed advanced image-analysis techniques to ensure that identical proteins in different gels are recognised as identical. The techniques have helped PSL to develop software for an integrated instrument, the Xcise, for high-speed processing of protein gels.

The instrument can acquire and automatically archive high-resolution gel images; detect, excise, process and analyse protein spots; and perform a 'trend analysis' of protein spots in a gel.

Thomas says that this system, which combines sophisticated hardware and software, is an important advance in high-throughput screening of proteins to find novel biomarkers, including drug targets.

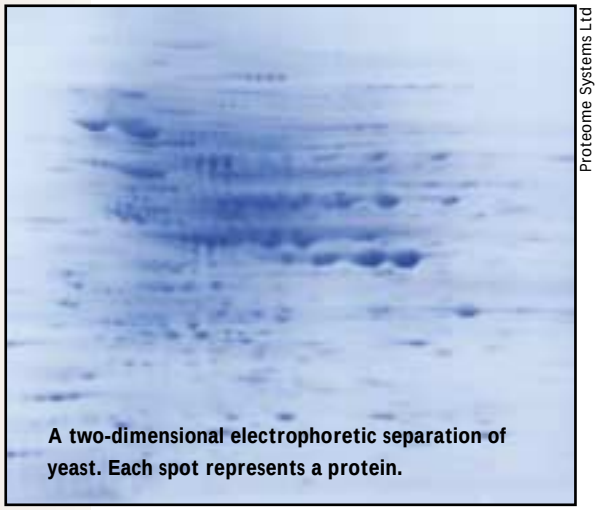


The team has constructed an atomic model of the hormone's signalling complex using the crystal structure of the Interleukin-6 receptor protein.

This is an important step towards the goal of more effective treatments for rheumatoid arthritis by administering small molecule inhibitors that prevent the complex forming.

Confounding cancer cells

Fellow protein chemist at CSIRO Health Sciences, Dr Colin Ward, and his colleagues collaborated with Dr Tom Garrett, a crystallographer now at The Walter and Eliza Hall Institute, to win a 20-year-long race to determine the exact three-dimensional structure of a protein that could lead to a whole new class of anti-cancer drugs.



Proteome Systems Ltd

A two-dimensional electrophoretic separation of yeast. Each spot represents a protein.

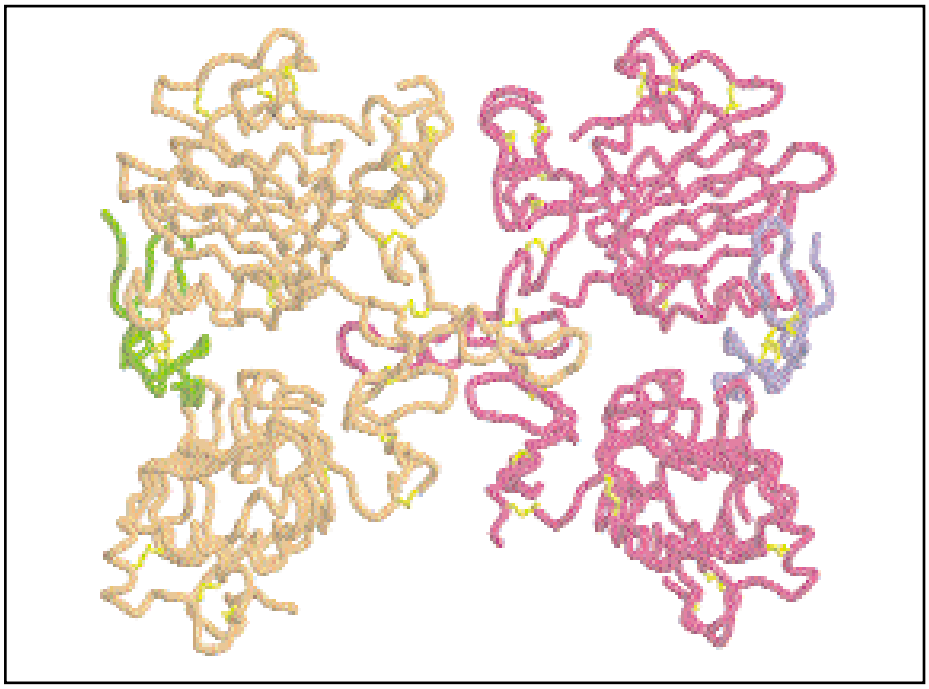
The protein is an EGF, or epidermal growth factor receptor. It sits on the surface of cancer cells, allowing the growth-factor molecule to latch on and stimulate growth of cancer tumours.

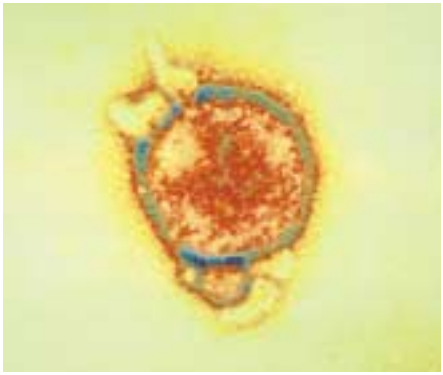
The idea is to eventually develop drugs that might specifically block the receptor site and so inhibit cancer growth. Antibodies to the EGF receptor have inhibited growth of certain cell lines in the laboratory and are being used in clinical trials.

The EGF receptor is in the same superfamily of receptors as those for insulin and IGF (insulin-like growth factor). All these receptors sit on the surface of cells and detect vital chemical messengers such as hormones and growth factors. When one of these attaches to its particular receptor, it in turn commands the cell to perform tasks such as growth or the processing of sugar.

Ward, Garrett, and their colleagues, try to understand the atomic structure of this family of receptors by coaxing the purified protein of interest to crystallise, and then bombarding it with a beam of x-rays to produce a diffraction pattern. This reveals the exact three-dimensional structure of the protein and is known as x-ray crystallography.

Their success with the EGF receptor followed the team's earlier hallmark determination of the structure of one half of the IGF receptor, which is similar to the insulin receptor. This has increased understanding of the chemical dynamics





Above: Proteomics is being applied to better understand the Hendra virus which appeared in Queensland in 1994–95, killing two humans and fifteen horses in two separate outbreaks. It normally infects fruit bats.

Far left: Determining the structure of a crucial human protein implicated in cancer. A crystal of the epidermal growth factor receptor and its attached hormone (the growth factor itself, TGFalpha) is analysed by x-ray crystallography.

Left: The elegant three-dimensional structure of the receptor protein bound to its growth factor hormone. Two identical receptors (coloured brown and magenta) hold arms while the two growth hormones bound to them (coloured green and purple) are held out to the sides.

of diabetes and represents a prominent milestone on the long road to a cure for the disease.

Improving life for livestock

At CSIRO Livestock Industries' Australian Animal Health Laboratory (AAHL), scientists are taking a proteomic approach to understanding Johne's disease in sheep and cattle; deadly viruses called Hendra and Nipah, which have been discovered in fruit bats; and the effects of plant toxins on grazing animals.

Dr Wojtek Michalski, who leads the AAHL Protein Biochemistry and Proteomics Group, says his team is arguably a world-leader in the application of proteomics to Johne's disease, a chronic intestinal infection that debilitates ruminant animals worldwide.

The scientists are identifying and characterising a variety of antigens in proteins derived from the bacterium that causes the disease. These may be suitable for better early diagnosis and disease prevention.

With the entire genome of the pathogen now sequenced, and with improved proteomic methods, it is possible to accelerate the identification of promising antigens, identify the genes encoding for them and produce the proteins using recombinant DNA methods.

Similarly, the researchers are applying proteomics to learning more about a family of plant-derived toxins known as corynetoxins. These are responsible for annual ryegrass toxicity, floodplain staggers and Stewart's Range syndrome in livestock and cause high stock losses.

Corynetoxins have a devastating effect when ingested because they inhibit an enzyme that kicks off the synthesis of a group of glycoproteins, compounds essential for normal cell, tissue and organ function.

The researchers are using an approach known as toxicoproteomics to study how the plant toxins affect livestock health. They are comparing protein maps for blood and tissues from healthy and poisoned animals to compile baseline data to apply in later studies of long-term, low-level dietary exposure to these toxins. This will help to minimise the risk to our export markets.

In 1994, the life of horse trainer Vic Rail and 13 of his thoroughbreds were lost to a deadly paramyxovirus. The Hendra virus was new to science (see *Ecos* 82). In 1999, another member of this virus family was discovered. This Nipah virus emerged in Malaysia, killing more than 100 people and resulting in the culling of over one million pigs.

Dr Michalski and his team study deactivated samples of the viruses, using proteomic approaches to identify receptor molecules. He says the work will lead to better diagnostic tools for the viruses and possibly virus specific treatments.

More about proteomics

Ezzell C (2002) Proteins rule. *Scientific American* 286:26–33.

Service RF (2001) High-speed biologists search for gold in proteins. *Science* 294: 207–2083.

Abstract: An explosion of data generated by the Human Genome Project and associated advances in biotechnology tools and equipment are fuelling interest in protein science, especially proteomics. However, studying the proteome or full protein complement of an organism is a much bigger challenge than its genome. While commercial enterprises generally concentrate on high-throughput screening of compounds in search of new pharmaceuticals, CSIRO scientists are using proteomics, structural biology and related approaches to study a range of health problems in humans and livestock. Others in the organisation are using bioinformatics to deal with the huge databases generated by proteomics and genomics.

Keywords: proteomics, proteins, growth factor receptors, bioinformatics, plant toxins, corynetoxins, pharmacology.