These stations are quite expensive but, just as core facilities for genome sequencing sprang up once the equipment came of age, the same is likely to happen with protein characterization. This should ensure that smaller academic and commercial labs will share in the advance of knowledge. And smaller labs might still be able to automate individual steps, such as spot picking or digestion, finding new ways to integrate steps that might be overlooked in larger, more streamlined organizations.

Alternatives for eliminating, rather than integrating, such steps are also emerging. One fairly new strategy involves transferring the gel to a membrane made of polyvinylidene difluoride (PVDF), then probing the membrane directly with mass spectrometry. This bypasses the spot-cutting step between electrophoresis and mass spectrometry.

Improvements also extend to mundane but essential items such as stains. Coomassie blue, a staple in most labs, can interfere with the digestion of gel spots by trypsin, so new stains such as zinc imidazole and noncovalent fluorescent SYPRO dyes, which do not have this limitation, are being introduced.

Mass-spectrometry output

It was not until the early 1990s that the mass spectrometers, now virtually essential components in the proteomics pipeline, could be used to analyse proteins.

Mass spectrometry relies on the fact that a substance carrying a net electric charge — an ion — can be made to move in a predictable way in an electromagnetic field. Ions are sorted by their charge-to-mass ratio, and from these a 'mass fingerprint' of the sample can be derived. Software, such as the University of California's Prospector package, can then be used to match the fingerprint to a protein database such as Amos Bairoch's Swiss-Prot (see 'Setting standards', overleaf).

In earlier models, excessive ionization energies would blast delicate molecules such as DNA and proteins into indecipherable particles. But innovations using a matrix such as MALDI, which protects the sample by modulating the ionizing laser beam, have helped to overcome this limitation.

Nevertheless, the technique still has its limits. A mass fingerprint will not be enough for identification if the protein is not registered in a database, or if post-translational modifications have changed its observed mass from the predicted value. In these instances, more information can be obtained from secondary protein fragments by re-routing the ions from the first analysis down a second channel and then analysing these fragments with the spectrometer. Of course, more complete databases will also help. And pairing mass spectrometry with other techniques, such as some kinds of protein-detector chip (see 'Alternative approaches', below) may make the method even more useful.

Future challenges

Automating and integrating the protein-characterization process is a good start, but there is no simple way forward. Although effective with adequate sample sizes, automated processes in general are not effective with very small amounts (less than 10 femtomoles of material).

It is hard enough to describe a single protein in a particular state. But things get even more difficult when trying to

Leroy Hood (right) and Ruedi Aebersold.

CHIPS

Alternative approaches

Proteins lack DNA's copying ability and do not readily undergo amplification, making separation and fractionation more important — especially for small amounts of proteins. And the inherent complexity and diversity of proteins makes a viable protein array an even more difficult goal. But the need to process proteins en masse is so urgent that heroic efforts are under way to develop a workable protein chip.

Leading the field at present are designs based on antibodies tethered to a solid surface. Large Scale Biology in Germantown, Maryland, and Biosite Diagnostics in San Diego, California, are developing an array of antibodies against 2,000–5,000 protein targets from the former's human protein index database. Biosite will use its omniclonal phage display technology to generate high-affinity antibodies against the targets. The companies hope the system will be available in the second half of 2002.

But an inherent drawback of antibody chips — or any protein chip, for that matter — is the destructive effect of proteases that may be lurking in the analyte mixture. "You have to use protease inhibitors if you're sampling microdissected tissue," says Lance Liotta of the US National Cancer Institute's Center for Cancer Research, who invents tools for proteomics and has surveyed the existing technology. "Process the tissue, lyse it, stain it and pray that these manipulations don't affect the 3D state of the protein."

Perhaps the biggest challenge is the accurate quantification of low-abundance protein. The faint signal of a protein of interest may easily be swamped by the much higher concentrations of other surrounding proteins.

Ciphergen in Fremont, California, is selling a device that helps scientists to detect low-abundance proteins. The company's chip uses specific surface chemistries to affinity-capture minute quantities of proteins. "A peak in one sample but not the other says a variation exists, but you still have to figure out what it is," says Mike Baldwin, a chemist at the University of California, San Francisco. "It's an interesting approach, but not mainstream proteomics — at least, not yet."

Another recent quantitative protein-expression and -identification technique using mass spectrometry is isotope-coded affinity tagging (ICAT), a kind of labelling invented by Ruedi Aebersold at the Institute for Systems Biology in Seattle. The start-up company Sense Proteomic, based in Cambridge, UK, is trying to use smaller numbers of mounted proteins to assay for suspected protein–protein interactions such as those known to play a role in toxicity.

Other chip approaches towards proteomics include atomic-force microscopy, aptamer libraries and biosensors.