these can provide the basis for a start-up company.

Upstart start-up

Take CellZome, which has spun off from the European Molecular Biology Laboratory (EMBL) in Heidelberg. This company is so new that it hasn't even got a website. Gitte Neubauer is a mass spectrometrist and one of six founding scientists from EMBL. Three are biologists, two are mass spectrometrists and one is a bioinformatician.

Using novel technology developed by Bertrand Seraphin (one of the co-founders of CellZome), the company plans to purify complexes of interacting proteins and identify them using mass spectroscopy and bioinformatics. It will complete the work by carrying out the biochemistry to validate its conclusions. The company is currently looking for collaborators interested in applying its techniques and scientific know-how to drug discovery.

Taken together, Oxford GlycoSciences, Glaxo Wellcome and CellZome give a good sense of the vibrancy of the newly emerging field of proteomics. Although they vary in both size and age, all of them are acutely aware of the need for skills in mass spectroscopy and bioinformatics. People armed with those skills should find themselves in demand.

Helen Gavaghan

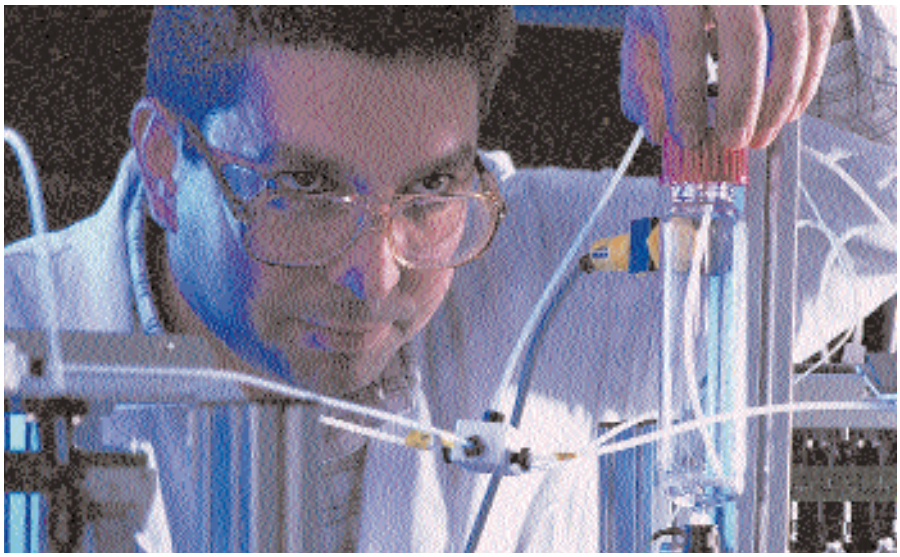
Further information

For a good sense of the techniques used in proteomics see:

1. Kell, D. B. & King, R. D. *Trends Biotechnol.* **18,** 93–98 (2000).
2. Brent, R. *Cell* **100,** 169–183 (2000).
3. Abbott, A. *Nature* **402,** 715–720 (1999).
4. Blackstock, W. P. & Weir, M. P. *Trends Biotechnol.* **17,** 121–127 (1999).

For proteomics opportunities and information see:

- Oxford GlycoSciences — jobs and technologies
  http://www.ogs.com
- Glaxo Wellcome
  http://science.glaxowellcome.com/start.htm
- SmithKline Beecham
  http://www.sb.com
- Incyte Genomics offers surfers a proteomics tour
  http://www.incyte.com/science/
- Human Proteome Initiative
  http://ebi.ac.uk
- Proteomics, protein databases and bioinformatics
  http://www.expas.ch and http://www.isb-sib.ch
- Links to academic proteomics sites
  http://www.proteome.med.umich.edu/links/proteo_sites.html



GLAXO WELLCOME

Growth industry: understanding protein function is key to future drug discovery.

Training

United States gives priority to skills shortage

Bioinformatics marries together a wide range of scientific disciplines, but with a global shortage of skilled researchers, training is high on the agenda.

Washington

Industry is draining bioinformatics talent from universities faster than it can be replenished. This is good news for the people getting such training, but bad news for the institutions that are scrambling to provide it, says Francis Ouellette, at the University of British Columbia's Center for Molecular Medicine and Therapeutics. Ouellette and Christoph Sensen at Canadian Bioinformatics Resource, in Halifax, Nova Scotia, run a four-part survey series (one week each for bioinformatics, genomics, proteomics and tools development), which introduces people to the field. Ouellette worries that the series is only a temporary patch on the problem.

Sensen stresses the difficulties academic groups have in finding and retaining talent. "In two years of looking I haven't found a person willing to be an understudy in this environment. PhDs either go to a company or to a nice warm place in the United States where they also get more money. But there is an urgent need for more PhDs in academia because that's where much of the real science is done."

Chris Lee, of the Bioinformatics Institute at the University of California, Los Angeles, concurs. Industry has the data, he says. But it lacks the "intellectual diversity" of a full-service university, as well as the freedom to "sit around talking about problems with people from different backgrounds".

The gap between supply and demand in bioinformatics is receiving official recognition in the United States. The US National Institutes of Health (NIH) supports bioinformatics mainly through two institutes, the National Human Genome Research Institute and the National Library of Medicine. However, for the field to grow, centres outside the NIH must also arise. The NIH approves the concept of developing such "centres of excellence", but has been slow to commit funding and develop infrastructure.

The National Institute of General Medical Sciences has also committed itself to

> The gap between supply and demand in bioinformatics is receiving official US recognition.

funding training slots, and a fourth branch of the NIH, the National Center for Research Resources (which is not an institute), has put itself behind shared bio-computational resources at more than a dozen centres nationwide. The Department of Energy's Argonne and Oak Ridge laboratories are also huge funders of bioinformatics work, as is, to a somewhat smaller extent, the Department of Defense.

On the private side, the Howard Hughes Medical Institute (HHMI) has declared that it will appoint investigators in computational biology, a departure for an organization that until now has avoided funding research in what it viewed as engineering disciplines. Now, however, it is becoming clear that biocomputation is not only part of HHMI's biomedical mission, but is one of its most critical elements.

Other support is also issuing from the Alfred P. Sloan Foundation, which has recently called for proposals to fund academic units that create professional master's degrees in biology. Traditionally, these degrees have not carried the same weight in biology as in engineering or business, where they are terminal degrees with full professional credentials. A two-year training cycle might help to relieve industry's acute need, where competent support staff, rather than technically savvy principal investigators, is what is needed.

Recognition of need

The NIH and the HHMI, respectively the largest public and private funders of biomedical science in the United States, have both previously acknowledged the need for more training in computational biology. But the money to pay for large-scale programmes has not materialized. Now the two agencies have signalled that this may change.

Last June, an NIH working group recommended a sweeping national programme to promote computational biology, its cornerstone being funding for up to 20 national centres (see http://www.nih.gov/about/director/060399.htm). Although the goals of this Biomedical Information Science and Technology Initiative (BISTI) have not been met, Carol Dahl, NIH assistant to the director of special technologies, remains optimistic.

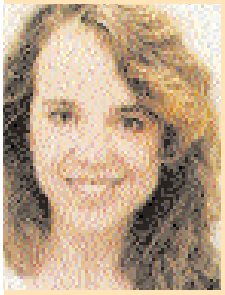
BISTI will be considered by all NIH institute directors this spring, and money for the programme could be in the budget for 2001. President Bill Clinton's 2001 budget proposal contains more money for large centres rather than new money for additional investigator-initiated grants.

The HHMI has sent similar signals. When Tom Cech was appointed director of the institute, he was quick to single out computational biology as an area in which it might increase its support (see *Nature* **398,** 256; 1999). But, because of the necessary shift towards collaboration and data collection, academia may need to rethink the way such work is rewarded, Cech warns. Tenure might not come easily to a researcher whose credit to a paper is lost in a sea of names. Other systems of acknowledgement, where each person's individual contributions are noted, may become necessary as authorship lists to collaborative papers become longer.

Last winter, the HHMI hosted a symposium to examine how computation and biology could be better merged and funded, from both public and private sources. However, according to an institute spokesman, no new funds have been committed for computational biology.

Potter Wickware

Borrowing methods and models



Kimmen Sjölander's early work in using Hidden Markov Models (HMMs) in genomics illustrates how bioinformatics reaches backwards to move forwards. HMMs originated in machine learning and were developed for speech recognition, but they turned out to be "fantastically successful" for modelling proteins, says the scientist from Celera Genomics, Foster City, California. "This typifies what bioinformatics is all about, to recognize that a method in one context can be transferred to another."

People with these skills are valuable to industry. "Pharma companies are desperately trying to organize and prioritize their drug targets," Sjölander says. "To do this you begin by making inferences about functions that are embodied in sequence."

However, that skill does not come easily. To develop fully the ability to make inferences, one needs to feel comfortable moving back and forth between computational and experimental data. "It's important for people who want to work in this field to have a true basis of understanding in more than one field. In addition to molecular biology, math, statistics and some elements of machine learning would be helpful," she says. "You can't do it with knowledge of just one specialty."

Sjölander thinks that bioinformatics should be a science in its own right. The demand certainly exists. The sheer quantity of data being generated from such a wide variety of sources requires the constant creation of new methods using mathematical or statistical modelling. "There's so much data being generated and we need to have methods that will respond to its particularities."

To increase the supply of computational biologists to deal with the demand being created by the data deluge, Sjölander believes that a cultural change is in order. "Historically, biologists haven't felt comfortable talking to mathematicians and computer scientists, who in turn don't really understand a lot of the issues in biology," she says. "That has to change."

P.W.

Training

Europe seeks solution to bioinformatics shortfall

London

A shortage of skills in bioinformatics is not limited to the United States. Europe is also feeling the pinch. The need for bioinformaticians to make sense of the data emerging from the Human Genome Project is well known among the industrialists and academics working in the field. It is something that was emphasized repeatedly to Lesley Thompson, who runs the life sciences interface programme of the Engineering and Physical Science Research Council, as she toured the United Kingdom's universities to identify ways in which the physical sciences could help life sciences.

For proteomics, bioinformaticians are, of course, essential, but they are not very easy to find. Both Walter Blackstock, of Glaxo Wellcome, and Gitte Neubauer, of the start-up company CellZome in Heidelberg, say that when recruiting for their proteomics effort they received only ten applicants for posts in bioinformatics, compared with 100 for biology.

One way to begin to tackle this imbalance, says Graham Cameron, joint head of the European Bioinformatics Institute (EBI) in Cambridge, an outpost of the European Molecular Biology Laboratory, would be to make sure that training in bioinformatics is included in all undergraduate biology courses. "Bioinformatics needs to be so pervasive in biology that it transcends its geeky image," he says.

Another way of addressing the imbalance is by establishing more academic programmes at all levels. Five of the United Kingdom's research councils have banded together to strengthen research in bioinformatics. The Biotechnology and Biological