The society of proteins

Having realized that proteins usually do their jobs by combining to form transient complexes, biologists are queuing up to study these structures using a powerful electron-microscopy technique. Alison Abbott reports.

Giulio Superti-Furga likes to call himself a protein sociologist. His team has provided powerful evidence for an emerging biological concept: that proteins generally don't work alone, but instead assemble into complexes until their job is done. This can mean a long-lasting association, but often it is just a fleeting alliance. "There are two types of protein — those that prefer long-term relationships, and those that prefer one-night stands," quips Superti-Furga, who is vice-president of Cellzome, a biotech spin-off from the European Molecular Biology Laboratory (EMBL) in Heidelberg, Germany.

Behind Superti-Furga's humour lies an important shift in biological thinking. It has long been known that some proteins form stable complexes to perform a specific task. The ribosome, for instance, which manufactures new proteins from amino acids, consists of some 50 proteins plus three large RNA molecules. But the classic view of many cellular processes involves proteins interacting with one another in linear pathways, coming together as they shift around in the cell's cytoplasm.

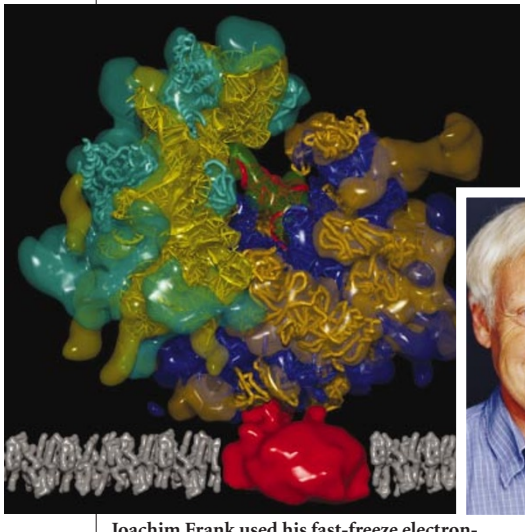
In recent years, however, biologists have realized that many important cellular functions are actually carried out by protein complexes that act as molecular 'machines'[1].

CELLZOME

CELLZOME

Hunted down: Giulio Superti-Furga's team has identified more than 200 protein complexes in yeast, here represented by coloured bubbles.

BECKMANN *ET AL. COLD SPRING HARBOR SYMP. QUANT. BIOL.* **66**, 543–554; 2002



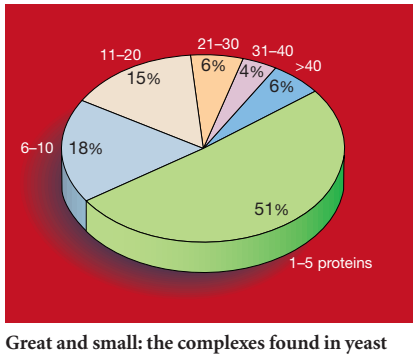
Joachim Frank used his fast-freeze electron-microscopy technique to analyse the ribosome.

About a dozen such complexes have already been subjected to intense investigation. And structural biologists recently started to get a handle on what some of these machines look like — and how the individual proteins that form their working parts move — using new methods in electron microscopy.

It was a pair of papers[2,3] published in *Nature* in January, one from Superti-Furga's group at Cellzome and the EMBL, the other from a Danish–Canadian collaboration headed by Mike Tyers of the University of Toronto, that really opened biologists' eyes to the ubiquity of protein complexes. The two teams tagged individual proteins from the yeast *Saccharomyces cerevisiae* with peptides that could be recognized by specific antibodies. This allowed the proteins to be used as 'bait' to pull out other proteins that were associated with them from yeast cell extracts. The researchers then purified the captured complexes and identified their constituent proteins using mass spectrometry.

Both teams found surprising numbers of protein complexes. The analysis by Superti-Furga and his colleagues used some 1,700 bait proteins, representing almost 30% of the yeast's total complement of expressed genes, and defined 232 distinct complexes. Superti-Furga is convinced that most are genuine functional entities, rather than experimental artefacts. "We checked that each component is found in the same subcellular compartment as the others in the same complex," he says. "We also pulled out the same complexes using different component proteins as bait."

The complexes vary considerably in size (see chart, below). Some are familiar —



SOURCE: REF. 2

Great and small: the complexes found in yeast vary in the numbers of proteins they contain.

Arp2/3, for instance, is an assembly of seven proteins that causes the polymerization of actin, a protein that interacts with myosin to generate cell movement and make muscles contract. The structure of this complex had been published only a few months before[4].

But most of the complexes were previously unknown. Nevertheless, by looking at the proteins within each, the researchers were able to infer possible functions for some of them. As a result, the Cellzome/EMBL team suggested new cellular functions for 344 proteins, including 231 for which there was previously no information.

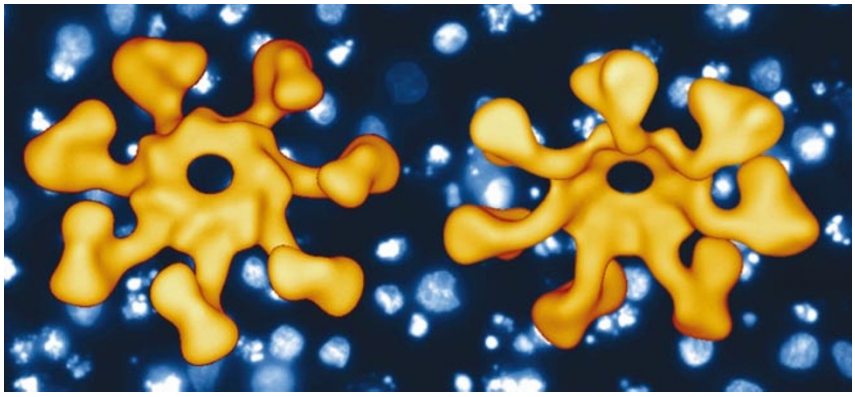
Promiscuous partners

Superti-Furga believes that this rich harvest of new complexes will include many transient structures whose component proteins are one-night-standers, "and also very promiscuous". Indeed, almost 40% of the proteins identified in Superti-Furga's project take part in more than one complex. "Everything seems to be interconnected," concurs Tyers, "and the most exciting thing is the number of unanticipated connections we have found."

To leading cell biologists, this makes sense. Tony Pawson of the Samuel Lunenfeld Research Institute at Mount Sinai Hospital in Toronto, a member of Tyers's collaboration, points out that many cell-signalling pathways share proteins. "We also know that the same protein can do different things in different cellular states," he adds.

One example is cytochrome *c*, a key player in respiration. This protein is usually anchored to the inner membranes of organelles called mitochondria, the main function of which is to generate adenosine triphosphate, the energy currency of the cell. But when the cell receives a signal to undergo apoptosis — or 'cell suicide' — cytochrome *c* leaks into the cytoplasm. Here, it associates with another protein called Apaf-1 to create the apoptosome, a complex that activates the enzyme procaspase-9, which breaks up the cell's components for recycling[5].

The apoptosome is one of a series of protein complexes whose three-dimensional structures have been sketched out using a technique called cryo-electron microscopy, or cryo-EM. In particular, a method called single-particle analysis, which was developed by Joachim Frank of the Wadsworth Center and the State University of New York, both in Albany, allows protein complexes to be analysed in three-dimensions from a set of electron-microscope images[6]. It does not deliver the high resolution of X-ray crystallography, the traditional method for analysing whole protein structures, but its main advantage is that it doesn't require the proteins to be in crystalline form. This is key, because protein complexes, which are temperamental at the best of times, are notoriously difficult to crystallize.



X. WANG & C. AKEY LABS

Wheel of death: the apoptosome's characteristic seven-spoked shape has now been visualized.

Put simply, Frank's technique involves applying a relatively pure solution containing the complex as a thin film onto a standard grid used to mount samples for EM, and then freezing it very rapidly at the temperature of liquid nitrogen. This freezing is so fast that the molecules have no time to move, and are captured in every possible orientation.

Electron microscopes work by shooting electron beams through a sample to generate a two-dimensional image. The complexes in the frozen film are so sensitive that they are destroyed if more than a single shot of electrons is used. But fortunately, one shot is all that is required, provided that there are sufficient complexes present, and you have enough computing power to calculate the typical three-dimensional structure of the complex from the resulting series of two-dimensional images.

All mapped out

Data on 5,000–10,000 copies of a complex are usually enough to produce a three-dimensional map of its structure with a resolution of 15–30 ångströms. Getting down to 10 Å requires up to 100,000 complexes. By comparison, X-ray crystallography can produce structures with a resolution of 2–3 Å. If these are available for any of the complex's component proteins, however, they can be superimposed onto the lower-resolution cryo-EM map.

Cryo-EM has already helped to reveal the workings of the apoptosome. "The ability to see what it looks like in three dimensions is enormously helpful in learning about function," says Christopher Akey, a structural biologist at Boston University who led a group that in February published the apoptosome's structure at a resolution of 27 Å (ref. 7). Akey's team showed that seven molecules each of Apaf-1 and cytochrome *c* organize themselves into a shape that looks like a seven-spoked wheel. This "wheel of death", as Akey calls it, does not spin. Instead, it appears to serve as a scaffold for seven molecules of procaspase-9 to bind to its hub. The molecules would then extend from the hub into solution, allowing each to bind to another procaspase-9 molecule to form the pairs that constitute the enzyme's active form.

Cryo-electron microscopy provides a rapid means of unlocking protein-complex structures.

Perhaps the best demonstration of the power of cryo-EM in understanding biological processes comes from Frank's own detailed analysis of the ribosome. A ribosome consists of two individual subunits that join together when the smaller subunit encounters and binds to a messenger RNA (mRNA) carrying the instructions to make a particular protein. Amino acids are joined together on the ribosome according to the mRNA code, each being delivered in sequence by a specific transfer RNA (tRNA) molecule. In 1995, Frank's team first determined the shape of the empty space between the joined subunits in their cryo-EM maps[8]. "We were struck by just how similar it was to the shape of a tRNA molecule," says Frank. "The perfection of the fit convinced us that the tRNAs must pass through this space to do their job."

Equally revealing was the fact that the

At CIMBio (main picture), Ron Milligan hopes that cryo-electron microscopy can be automated.

shape of this space is the same in ribosomes from all sorts of organisms — bacteria, wheat, yeast, rats — despite the fact that the ribosomes themselves in these organisms are very different in composition[9]. This indicates that ribosomes must have evolved in partnership with tRNAs over the past billion years or so, says Frank. And it makes a bigger point that is beginning to dawn on biologists — that the overall shapes of protein complexes doing a particular job are more likely to be conserved across species than those of the individual protein components.

Shopping spree

Biologists worldwide are now clamouring to purchase cyro-EM equipment, and to hire experts in the technique — still a rare breed — so that they can begin to analyse a multitude of complexes.

With this in mind, Superti-Furga and his EMBL colleagues are planning to launch a project called EMBA-C (European Molecular Biology Alliance for the analysis of Protein Complexes) to coordinate efforts to generate structural information on protein complexes, and to collate and archive the data. EMBL research groups will have first pick of the protein complexes identified by the Cellzome/EMBL team. "But we hope that the alliance will eventually be open to everyone," says Luis Serrano, who coordinates the structures and biocomputing programme at the EMBL. The effort will also extend to protein complexes identified by other groups.

Across the Atlantic, Ron Milligan, director of the newly opened Center for Integrative Molecular Biosciences (CIMBio) at the Scripps Research Institute in La Jolla, California, hopes to speed the rate of progress by automating cryo-EM. "I was extremely excited when I read the papers in *Nature*," he says. "Structural analysis of previously unknown complexes is exactly what we want to do with the new technologies being developed at CIMBio."

"EM structure determination is a bit of a craft at the moment," Milligan explains. "It involves sitting for hours at the microscope, and manually processing data." Together with his colleagues Bridget Carragher and Clint Potter, Milligan has proved in principle that the picture taking, data processing and generation of the three-dimensional map can all be automated, using the structure of the tobacco mosaic virus as an example. "But we are a long way from generalizing this automation to all samples," Milligan warns.

Other groups are aiming to focus on particular classes of protein complex. The Structural Cell Biology of DNA Repair Machines collaboration, led by John Tainer of the Scripps Research Institute and Priscilla Cooper of the Lawrence Berkeley National Laboratory in California, kicked off this year with a five-year grant from the National Cancer Institute. It aims to study the transient protein complexes that form when proteins that monitor DNA for damage find a problem.

Quick fix

According to the type of damage involved, particular combinations of DNA-repair enzymes join forces to achieve the necessary fix. "These transient machines are only created under particular conditions, and they are made up of parts which can recombine with other complexes to form different machines," says Eva Nogales, a cryo-EM expert at the University of California, Berkeley, and the Lawrence Berkeley National Laboratory, who is a member of the collaboration.

Milligan and Steve Almo, a protein crystallographer at the Albert Einstein College of Medicine in New York, this month launched discussions with about a dozen colleagues from different labs to work out a strategy for generating new protein complexes on a large scale and analysing their structures — and for seeking funding. They are also talking with Cellzome and the EMBL about creating an international 'complexomics' consortium. "Cellzome's resources could serve as a foundation for getting such a programme on its feet," says Almo. "A synergistic international effort will maximize everyone's productivity — and fun."

One top priority for many researchers is extending work on complexes to human proteins. Cellzome's scientists, for instance, are already looking at pathways in different human cells, such as liver cells and those of the immune system, that are likely to be involved in human diseases. "We have shown that many complexes in yeast and humans are nearly identical — they seem to share the same sociology, as if they are conserved at the machine rather than the protein level," says Superti-Furga. So Cellzome has been tagging all genes that have been shown to be associated with a particular disease, such as Alzheimer's or specific cancers. In this way, Cellzome researchers hope to identify machines, rather than proteins, that can serve as a targets for drug development.

As more complexes are identified and analysed, biologists are anticipating fundamental insights into the detailed workings of the cell that could also advance our understanding of development and evolution at the molecular level. Pawson, for instance, suspects that structural information on transient complexes may explain how different growth-factor receptors manage to transmit their own particular message to the cell's nucleus despite using the same signalling pathway.

Whatever the new discipline of analysing protein complexes ends up being called — complexomics, protein sociology, or some yet-to-be-coined term — it seems certain to bloom over the next few years. "Half of my lab wants to stop their present projects and work on complexes," says Almo. ■

Alison Abbott is *Nature*'s senior European correspondent.

1. Alberts, B. *Cell* **92,** 291–294 (1998).
2. Gavin, A.-C. *et al. Nature* **415,** 141–147 (2002).
3. Ho, Y. *et al. Nature* **415,** 180–183 (2002).
4. Robinson, R. C. *et al. Science* **294,** 1679–1684 (2001).
5. Zou, H., Henzel, W. J., Liu, X., Lutschg, A. & Wang, X. *Cell* **90,** 405–413 (1997).
6. Frank, J. *BioEssays* **23,** 725–732 (2001).
7. Acehan, D. A. *et al. Mol. Cell* **9,** 423–432 (2002).
8. Frank, J. *et al. Nature* **376,** 441–444 (1995).
9. Spahn, C. M. T. *et al. Cell* **107,** 373–386 (2001).