The Mapping of a Molecule

by Richard E. Dickerson

Caltech chemists are beginning to work out the structure of cytochrome c, a molecule which is essential to virtually all life on earth

One of nature's smallest biological engines, a molecule which is essential to virtually all life on earth, is gradually emerging from the invisible world to the visible. With a technique known as x-ray diffraction, a Caltech team has succeeded in enlarging the image of the engine about one hundred million times.

The "engine" is a molecule of a protein called cytochrome *c* and is shaped like an egg one ten-millionth of an inch across. For comparison, if the molecule were the size of a hen's egg, then a man to the same scale standing on the earth could reach halfway to the moon. An analysis of approximately 4,500 diffraction data from crystals of cytochrome *c*, with and without labeling of molecules with platinum and mercury, has led to the calculation of a low resolution map in which the general plan of the molecule is visible.

This molecule is one small part of a complicated factory, the mitochondrion, which converts food into usable energy for all of the body's processes. Dozens of these tiny factories are to be found working like miniature power packs in nearly every living cell. Cytochrome *c* itself is near the end of a production line (the terminal oxidation chain) which breaks down the fuel in small steps and saves the energy by storing it in molecules of adenosine triphosphate (ATP), which are then used to energize all plant and animal activity. The mechanisms by which this process is carried out are unknown, but finding out the structures of the machines involved is a step in this direction.

Cytochrome *c* is interesting not only because of its key role in energy production but also for what it tells us about evolution. Cytochrome *c* evolved shortly after oxygen-using, one-celled life appeared on earth roughly two to three billion years ago. All living things that have descended from these simple organisms have cytochrome *c*; and cytochome *c* from wheat germ, yeast, screw-worm fly, horse, and man are essentially the same.

The chemical makeup of cytochrome *c* is well known. The molecule has a molecular weight of 12,400 and is a chain of 104 links, each an amino acid with particular chemical properties. Emanuel Margoliash at Abbott Laboratories in North Chicago and Emil Smith at UCLA have studied the sequences of these links in the chains of more than 30 species including man, monkey, and other mammals; fish, birds, reptiles, insects, yeasts; and one plant, wheat germ. Some important parts of the chain are absolutely identical in all species; in other parts it seems to make little difference what kinds of links are present. One of our goals at Caltech is to find out where these constant and variable parts of the chain are situated on the molecule and why their influence on the molecule is so different.

Margoliash and Smith have shown that the degree of similarity or difference between chains from two species checks exactly with their places in the evolutionary tree. For example, man and yeast differ by 44 links in the chain, man and horse by 12 out of the 104, and man and rhesus monkey by only two. Cytochrome *c* from man and from chimpanzee are identical. Darwin's theory of evolution is thus checked out using not only evidence which Darwin did not know about but also a substance whose existence was completely unknown in Darwin's time.

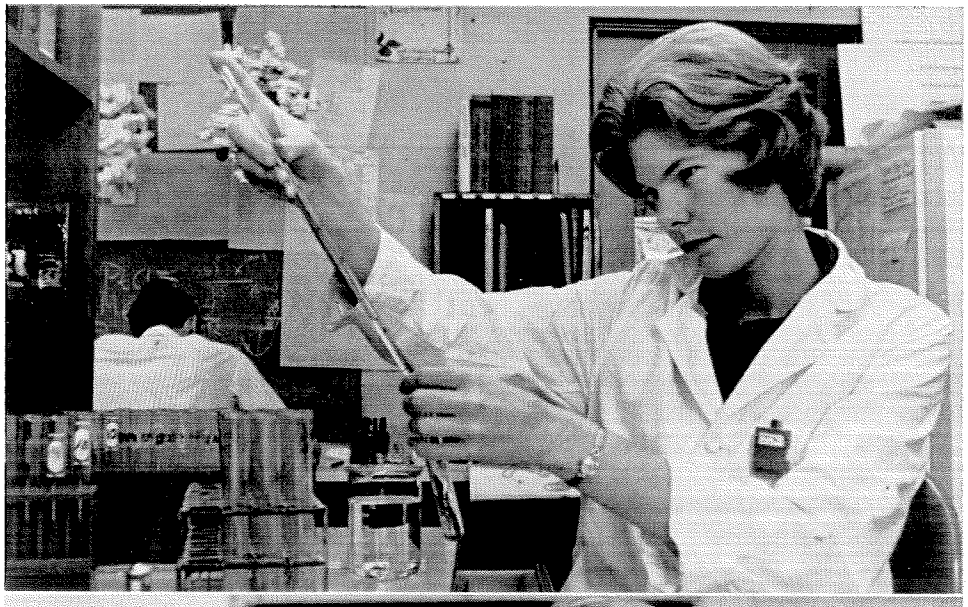
The heart of this tiny molecular engine is a heme group—an atom of iron surrounded by a cluster of atoms called a porphyrin molecule—very similar to that in hemoglobin. But instead of carrying oxygen where it is needed, as hemoglobin does, cytochrome *c* is designed to accept an electron from its predecessor in the mitochondrial energy chain and then to give it up again at the proper place. How it does this is not known. Some chemists think that the molecule shuttles back and forth from one site to

another. Others have proposed that it stays in one place but rotates or wobbles so as to bring its iron-containing heme group close to the electron donor and acceptor in turn. Chemical and physical evidence suggests that the molecule opens up to some extent when it loses an electron and closes down when it receives one. In order to check this idea out and find the mechanism of oxidation and reduction of the iron, two independent x-ray analyses of the molecule are required—one in the oxidized and one in the reduced state.

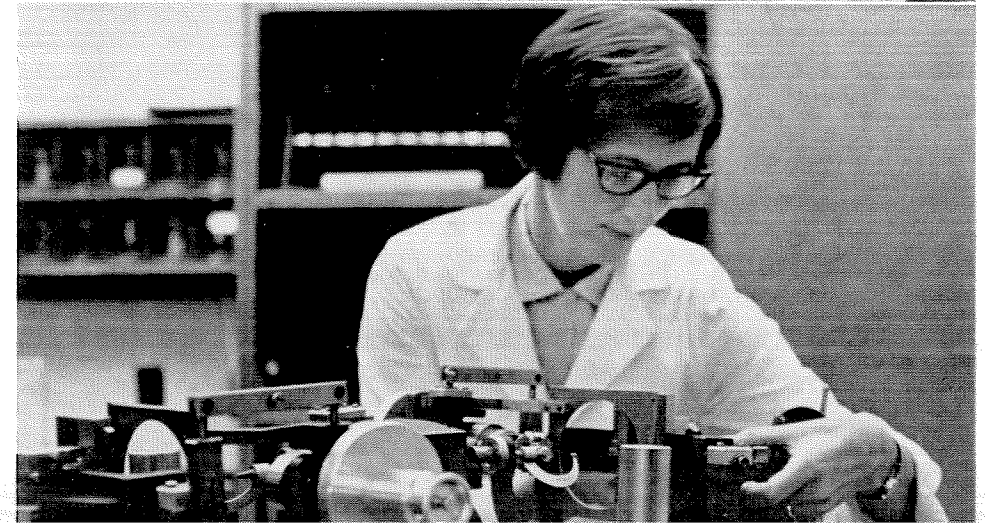
The richest source of cytochrome-containing mitochondria in mammals is the heart, the muscle that never stops working and the one that has the most continuous need for a power supply. The cytochrome for the Caltech structural work comes from horse hearts—about two and a half grams of crystalline cytochrome *c* from 35 pounds, or seven good-sized hearts. The protein must be separated, purified on a chromatographic column, and crystallized from ammonium sulfate solution. Crystals of oxidized horse-heart cytochrome are easy to grow in a matter of weeks. Those of other species are more difficult, and to date no species of the 30 or so which have been purified and analyzed at the Abbott Laboratories has crystallized well in the reduced form. This is an obstacle which will have to be overcome, of course, if the mechanism of oxidation and reduction is to be understood.

The first step in an x-ray diffraction structure analysis is to mount the crystal in the path of a narrow x-ray beam and then to move the crystal and some radiation detection system, either photographic film or counter, so as to collect the entire diffraction pattern of the crystal. On film the pattern will be a regular array of spots. The positions of these spots tell the investigator how the molecules are packed in the crystal (in this case a matter of only minor interest). But the *intensities* of diffraction in all of the different directions of diffraction lead to the important information—the atomic structure of the molecules doing the diffracting.

A complete structure analysis of a small organic compound may require measurement of only one or two thousand x-ray diffraction intensities. But the same analysis for a protein such as cytochrome *c* would involve 200,000 data. This means that for practical reasons a protein structure analysis is carried out in stages with increasing amounts of data and with successively higher resolution of de-



Cytochrome c is extracted from horse heart and is purified and fractionated on a chromatograph column. Here laboratory assistant Joan Varrum runs an assay on the fractionated samples.



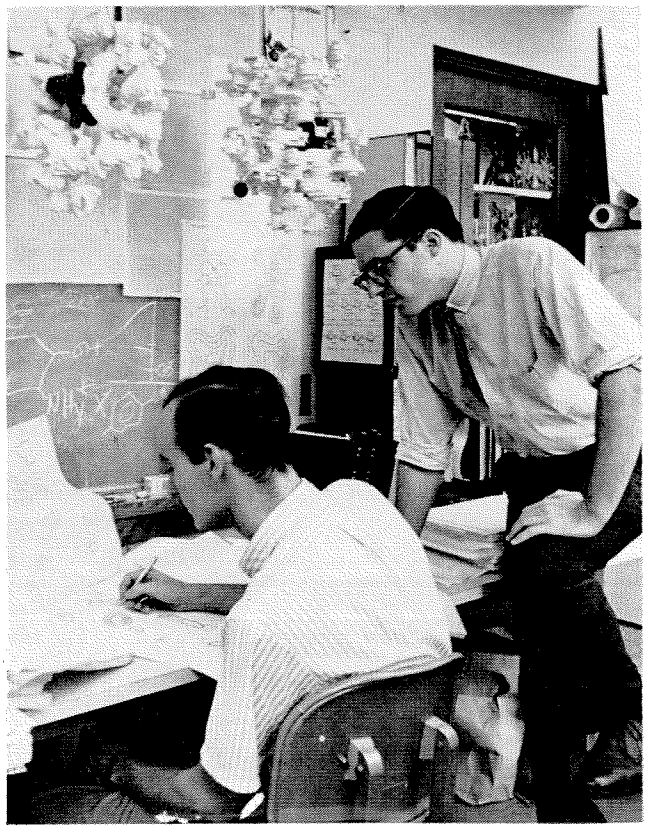
Crystallized protein is mounted on glass capillaries on an x-ray camera, and its diffraction pattern is collected. The crystal lies just above the crescent shaped beam stop near the center of the camera. Olga Battfay, laboratory assistant, aligns the precession camera.

Richard Dickerson assembles the plexiglass sheets which show the electron density of sections of the molecule. These sheets are used to plot the low resolution map of cytochrome c.

tail. It also means that automation, which is becoming more and more desirable in smaller structure analysis, is mandatory for proteins.

Richard Stanford, senior research fellow in chemistry and one of my collaborators in the Caltech protein structure group, has put together an automated, punched-tape-controlled, single-crystal diffractometer for collecting protein data. The paper master tape, produced at the chemistry division's remote computer console by the IBM 7094, tells the diffractometer where to position the crystal and proportional counter, and how long and under what conditions to measure diffraction intensity from the crystal. The results are punched out on paper tape for input to the remote console and subsequent analysis. The diffractometer can collect about 400 diffraction intensities per day and is similar in principle, if not in design, to the small-molecule diffractometer developed in chemistry by senior research fellow Sten Samson.

Each of the diffracted rays from the crystal has a phase as well as an intensity, but, as always with electromagnetic waves, all that we can observe with ease is the intensity. The phase information, unfortunately, is essential for the structure analysis. If a physicist wants phase information, he builds an interferometer. We do much the same thing, but in this case the chassis on which the interferometer is built is the cytochrome molecule itself. Extensive diffusion and reaction experiments have to be run until some way is found of labeling each molecule in the crystal at the same point with a heavy atom such as platinum or mercury. The interferometry experiment is then the comparison of total diffraction patterns of the protein with and without the heavy atom. This isomorphous replacement method, developed for proteins only 14 years ago by M. F. Perutz, is the key step which makes analysis



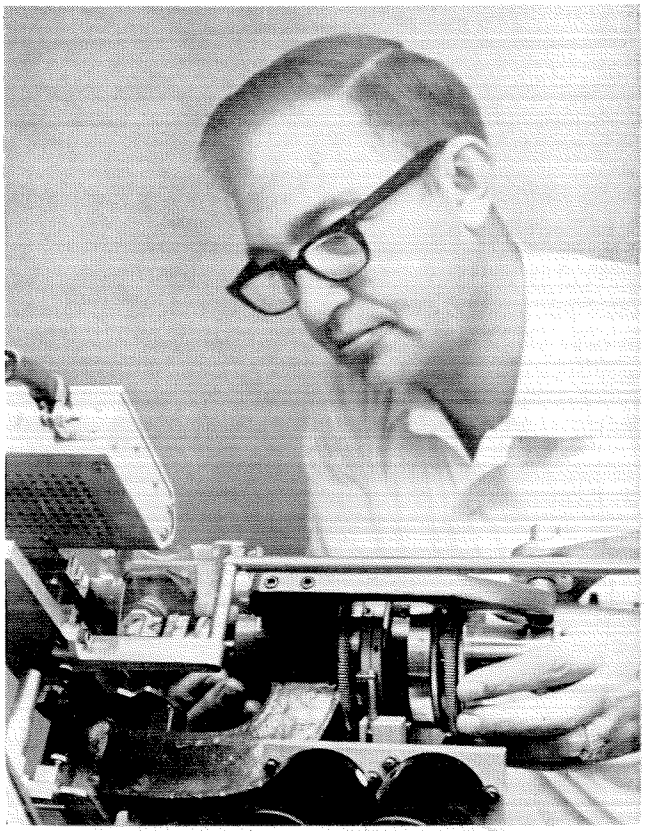
The electron density in the form of computer output is ultimately translated into solid models of molecules such as the ones suspended from the ceiling of the laboratory above the heads of Jon Weinzierl and David Eisenberg.

of such large molecules possible. Two such tagging groups have been found for cytochrome *c*, one containing platinum, the other, mercury.

The result of the x-ray analysis is the electron density at thousands of points through the molecule. This map must be plotted and interpreted in terms of a meaningful chemical structure. The current picture of the molecule is at a resolution of four angstrom units, which means that two features about four angstrom units (or 4×10^{-8} cm) apart would just be resolvable. At this resolution, benzene rings become shapeless blobs, and polypeptide chains become snaking ropes of density through the molecule. The over-all plan of the molecule is visible, though the details are not. But it appears even at this stage that the center of the molecule is filled with atoms of packed hydrocarbon-like amino acid side chains, forming a molecular "oil drop" in which the heme is immersed, sitting in a crevice in one side of the molecule. Enzymes and globular proteins that we find today have evolved to operate well in an aqueous environment. A common feature of those other proteins whose x-ray structures have been deciphered to date (myoglobin, hemoglobin, ribonuclease, chymotrypsin, and lysozyme) seems to be a local region of less polar, non-aqueous surroundings for the reactions with which it is involved. The chemistry of the heme iron will be greatly affected by the polarity of the medium around it.

The main feature of the egg-shaped cytochrome *c* molecule is a crevice, running lengthwise, into which the heme group is fitted. The heme is firmly attached to the protein on one side of the crevice by two covalent bonds and by coordination of the iron with the nitrogen atom of a histidine side chain. On the other side of the heme, a polypeptide chain sweeps in a curve past the iron atom and extends another coordinating side chain to it. Some chemists think that this side chain is a particular methionine. If so, then this structurally vital part of the chain would correspond to that part of the chemical sequence which has been absolutely unvarying through a billion years of evolution.

The present incomplete picture has encouraged us and has made us anxious to get on to the next stage, the high resolution analysis. In the coming year David Eisenberg, research fellow in chemistry, Dr. Stanford, the rest of the team, and I will be involved in collecting the data required for a map which will show the details of the path and the folding of the polypeptide chain from which the molecule is built. When the crystallization problem is solved, the reduced molecule analysis will follow.



Richard Stanford adjusts the automatic diffractometer which will be used for high resolution analysis.