Lab Notes

The image-processing algorithms developed for the Voyager missions to the outer solar system are now being used to explore inner-space territories such as blood samples and cell cultures.

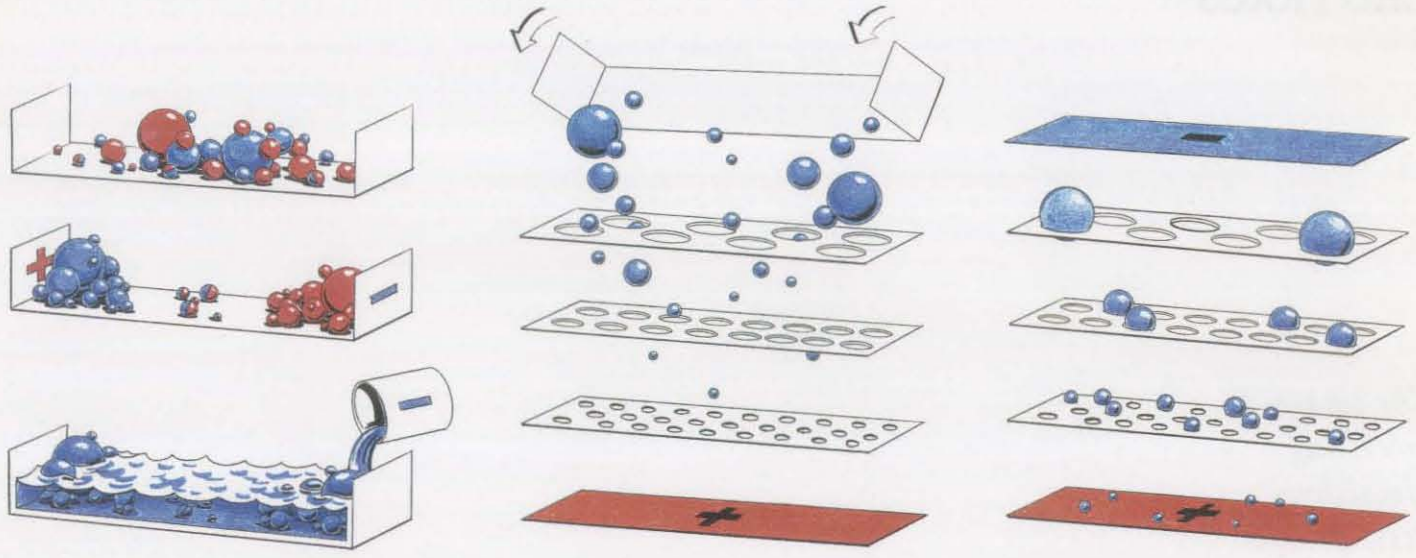
From the Great Red Spot to Protein Blots

Searching for volcanism on Triton and watching a cell develop are more closely related than you might think. The image-processing algorithms developed for the Voyager missions to the outer solar system are now being used to explore inner-space territories such as blood samples and cell cultures. But Jerry Solomon, a senior member of the technical staff at JPL, and biologist Michael Harrington, a member of the Beckman Institute at Caltech, aren't touching up mug shots of microbes—they're using the algorithms to read spots off a two-dimensional electrophoresis gel.

Two-dimensional gel electrophoresis is a powerful but underused technique for separating a complex mixture of very similar proteins such as plasma or cellular innards. The sample is applied to one end of a gel-filled capillary tube, and an electric field draws the proteins through it in proportion to their charges. Then the gel is extruded along the edge of a slab of a different gel, like toothpaste onto the bristles of a toothbrush, and treated with a mild detergent solution to give all the proteins a uniform charge. Another electric field, perpendicular to the previous one, draws the proteins onto the slab at migration rates inversely proportional to their sizes. The proteins are rendered visible either directly, by various stains, or indirectly, by radioactive labels. A one-microliter droplet of plasma will speckle a standard 20-centimeter-square gel with several thousand spots, each a different protein that can, if the sample is large enough, be analyzed and identified.

The technique has tremendous promise for basic biology and for medical diagnostics. A researcher studying cellular development could track the proteins involved at each stage by following the shifting patterns of spots on a series of gels. A disease caused by a missing or aberrant protein would reveal itself as a mismatched spot when gels from diseased and normal cell cultures are compared. Other deviations from the norm would be equally revealing—detecting brain-specific proteins in a cerebrospinal-fluid sample could indicate a brain injury, or some degenerative process like Alzheimer's disease.

The catch is that it's very hard to keep your eye on one spot in the crowd from gel to gel, even if nothing else is changing. Unfortunately, lots of things



Two-dimensional gel electrophoresis in cartoon form. The spheres are protein molecules of various sizes; each sphere's color symbolizes the protein's charge—red is positive, blue is negative. Neutral proteins are shown as half red and half blue. The top left-hand drawing shows the initial sample. Applying an electric field (middle) separates the proteins by their charge. Adding a detergent solution (bottom) gives the separated proteins a uniform charge. The proteins can now be sorted by size (center drawing) by applying a second electric field in a direction perpendicular to the first field. The right-hand drawing shows the result—the proteins are sorted by charge along one dimension and by size along the other, allowing very similar proteins to be separated from one another.

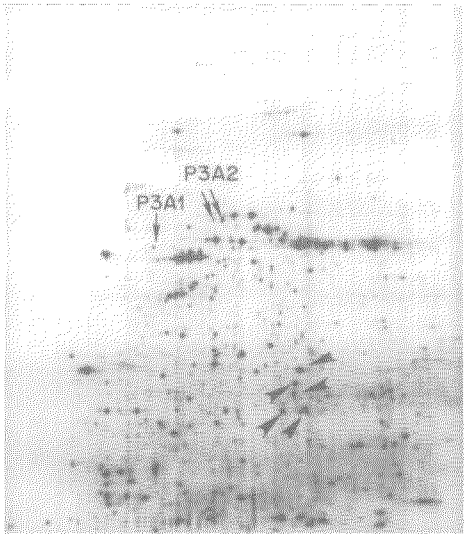
are. The extrusion step allows all sorts of positioning errors to creep in. (Harrington is working on hybrid gels that would eliminate this transfer, and is trying to automate the whole process.) Subtle variations in the gels or in the electrical fields can make spots wander from run to run, even when the gels are perfectly aligned. And to cap it all off, a spot can change its appearance by developing a tail, metamorphosing from an oval to a kidney shape or even to a seeming paw-print, or fading from a sharply defined blot to a fuzzy smudge that blurs into its surroundings.

Harrington's and Solomon's close collaboration was essential to understanding and coping with all these variations. Some were separation artifacts, curable with a little judicious tinkering—the tails, for example, turned out to be due to interactions between the protein and the gel matrix. Others proved to be biologically significant and had to be handled through image processing. The kidneys and paw-prints—"concavities"—were actually two or more oval spots overlapping. And spots grew fuzzy when the proteins that made them had been converted to glycoproteins (by reactions with sugar molecules) in the normal course of metabolism. "Jerry wrote some code to arbitrarily give the 'fuzzies' clean edges," says Harrington. "No other system has done that. And he can pull the concavities apart into their component ellipses, so that we can match a concavity on one gel to an ellipse on another. These are the kinds of issues that can't be addressed by a biologist and a programmer working in isolation. Jerry wouldn't know which features were important, like the fuzzies, and which were irrelevant, like the tails. And I wouldn't know how to impart my gel-reading experience to a computer. I don't know of anywhere else where one could tackle all the mutually dependent aspects of developing this technology."

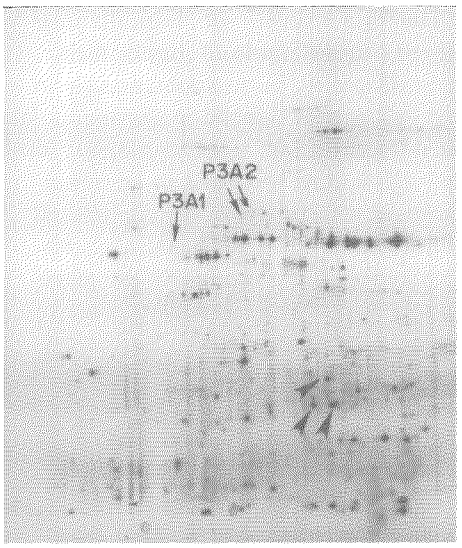
The pair has been working on the problem since the autumn of 1988, when Lee Hood, Bowles Professor of Biology and Director of the Center for Molecular Biotechnology, brought them together. They have gotten the system to the point where a gel can be scanned, its spots registered, and their pattern compared to previously scanned ones. But the biggest problem from the medical point of view is still ahead—building the data base of "normal" samples. For proteins as for people, "normal" encompasses many variations. In order to be of any diagnostic value, therefore, a "consensus normal" pattern must be assembled for each tissue type.

Doing basic biology doesn't have to wait, however. Eric Davidson, Chandler Professor of Cell Biology, studies the

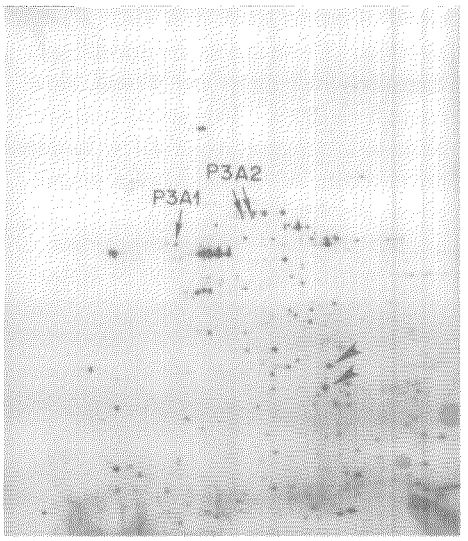
Pre-column



Post-column, low-salt fraction



Post-column, high-salt fraction



Two-dimensional electrophoresis gels of proteins extracted from sea-urchin embryo nuclei, before and after passing through the affinity-binding column. P3A1 and P3A2 are known DNA-binding proteins. The "post-column, low-salt fraction" contains the proteins that weren't trapped by the column, while the "post-column, high-salt fraction" contains the ones that were. The five bold arrowheads show how effective the separation is.

process of gene activation—"expression"—that controls embryonic development, using sea-urchin embryos. An embryo begins as a single cell, which divides repeatedly to form a ball of identical cells, which differentiates to become the various parts of the complete organism. But differentiation starts long before any outward transformation becomes visible, as different sets of genes begin to be expressed in cells in specific regions of the embryo. Davidson's group is particularly interested in the onset of this process, known as "molecular differentiation," which occurs when the embryo's cell population approaches 100. At this early stage, the sea-urchin embryo consists of five invisible territories, each of which will construct a different portion of the larva. Davidson's and Hood's laboratories are collaborating to analyze the regulatory molecules that control gene expression in each territory. Davidson had guessed that 50 to 100 DNA-binding proteins would be needed to control the process, and the new system provides a way to identify them amid the thousand or so proteins present in the cell nucleus at that stage. Using methods developed by James Coffman, a postdoc in Davidson's lab, and Frank Calzone, now an assistant professor of biology at UC Irvine, the team adsorbs an extract of these proteins onto an affinity-binding column that traps all proteins with a predisposition to bind to DNA. Rinsing all the unbound proteins off the column and desorbing what remains yields an extract that two-dimensional gel analysis shows to contain roughly 50 proteins, including the DNA-binding proteins conclusively identified to date. The group will now tackle the problem of identifying the rest of the proteins. "This is tremendously exciting work that we couldn't have done without this image-processing system," says Harrington. "And the same general strategy can be applied to any kind of cell—neural precursors or immune-system components, for instance. We can also do double comparisons, using activated and quiescent cells, to determine which DNA-binding proteins are involved in a particular activation step." □—*DS*