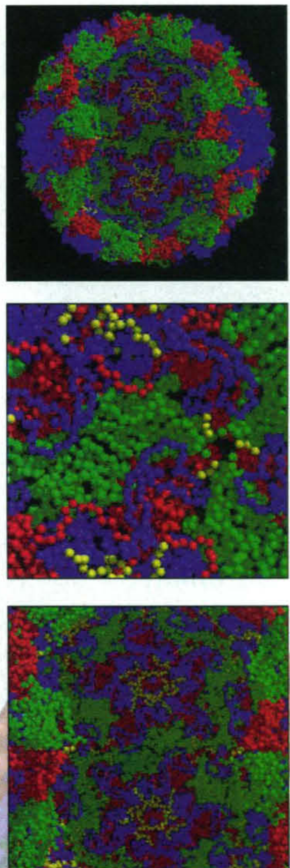
PAGE 1747
Biodiversity in the Smokies

1748
On the verge of a blowup?

BIOCOMPUTING

# NIH Urged to Fund Centers to Merge Computing and Biology

The National Institutes of Health (NIH) should fund a new network of interdisciplinary research centers that would boost the number of computer-savvy biologists, an advisory panel recommended last week. The advice, which drew a positive reaction from NIH chief Harold Varmus, is contained in a report that lays out a road map for the $16 billion agency in biocomputing, an emerging interdisciplinary field that the panel said has been neglected by academia. But it does not spell out how much NIH should invest in the field, nor how the money should be spent.



**Close ties.** Computer-generated images of a rhinovirus whose structure was revealed by x-ray crystallography illustrate the teamwork envisioned in NIH report.

The use of computers in biomedical science has grown explosively over the last decade, with researchers relying on the machines to do everything from browse the technical literature to model protein folding. But few biologists have the computing expertise needed to fully tap the flood of new data being generated by gene studies and clinical trials. A typical gene lab, for instance, can produce 100 terabytes of information a year—a database equivalent to 1 million encyclopedias. At the same time, few computer scientists know enough biology to figure out the best ways of sifting valuable nuggets from the data deluge. "You can count on the fingers of one hand" the number of researchers with topflight training in both fields, says geneticist David Botstein of Stanford University in California, a co-chair of the advisory panel.

To help NIH figure out how to attack the problem, Varmus last year appointed a 16-member working group led by Botstein and computer scientist Larry Smarr, director of the National Center for Supercomputing Applications at the University of Illinois, Champaign. After polling researchers working in disciplines from neuroscience to population genetics, the panel came up with four recommendations in a 20-page report* that it hopes will influence the 2001 budget request now being drawn up.

The panel's cornerstone recommendation is that NIH create between five and 20 biocomputing centers at universities and independent research institutes as part of a "national program of excellence." The competitively funded centers would range in size from a handful of researchers at a single institution to a multi-institutional consortium with a budget of up to $8 million a year. The funding would flow from a new NIH biocomputing program that would make research grants through one or more of NIH's disease-oriented institutes, with contributions from the host institution.

A second recommendation calls for NIH to take a more active role in shepherding the growing flock of biomedical databases, which hold everything from gene sequences to drug trial results. Although the agency funds a variety of bioinformatics research efforts, the panel noted, none is dedicated to organizing and curating databases, which researchers believe could yield important insights if indexed and integrated. "The goal is a system of interoperable databases," the panel wrote.

Panelists also asked NIH reviewers to be more supportive of requests from individual scientists for funds to hire biocomputing talent, including highly priced programmers. "It is not a pretty picture down in the study sections when you [ask for] three programmers at $85,000 each," bemoans Botstein. In the future, says the panel, NIH needs to ensure that a larger share of its bread-and-butter R01 grants to individuals "may be used for biomedical computation."

The fourth recommendation is for NIH to help build all kinds of computing resources. In particular, Smarr believes that midlevel networks—computing systems more capable than a single desktop machine but less powerful than a supercomputer—would be a good way to test experimental software. The panel warned, however, that more scientists are needed who know how to set up and operate such networks.

Indeed, Smarr says the panel "kept coming back to people as the limiting factor." He sees the proposed training and research centers as "watering holes" that will use research funds as a lure to attract both biologists and computer scientists to core problems, such as database organization. He and others also hope that NIH's backing will knock down disciplinary walls and bureaucratic obstacles to producing—and hiring—more biocomputing faculty. Funding the centers "will send a powerful message, both in academe and within the NIH community itself, about the importance of computation," the panel wrote.

Some biocomputing researchers say an NIH endorsement would help stanch the flow of academic biocomputing talent to industry. The report is "good news" and "should cause universities to pay attention," says computational biologist Larry Hunter, a researcher at NIH's National Cancer Institute in Bethesda, Maryland, and president of the International Association for Computational Biologists. "It could help reverse a situation in which the [academic] rewards for doing technology development are not very good," adds Chris Overton of the University of Pennsylvania, Philadelphia, one of a handful of schools to take steps to create new biological computing programs.

Because money is a sure-fire attention getter, the academic community is already eager to see how much NIH officials will decide to invest. Although some funds may be forthcoming in the 2000 fiscal year that begins in October, NIH's real response will be contained in its 2001 budget request that emerges early next year after negotiations with the White House. **–DAVID MALAKOFF**

* The Biomedical Science and Technology Initiative, prepared by the Working Group on Biomedical Computing, 3 June.