been worked through. Based on the results from the first 500 files, Weiss thinks that a full 10% of the 3000 faculty and staff members to be investigated will turn out to have been collaborators, much more than the national average of 2% to 3%. The effort to cleanse the university of Stasi collaborators is seen as an unpleasant but necessary task, administered by clear rules; there is no punishment for those who simply went along with the old regime, but everyone who did "active harm" to others must go, so that an attitude of trust can be rebuilt in the university.

Among those who come before Weiss are colleagues who once mocked him for his refusal to join the party or accept its favors and so make his way upwards in the East German system. "'Good day, little man,' they would call as they rushed past me in the hallways," Weiss recalls with a trace of bitterness. But despite the mockery, Weiss refused to fit in, perhaps because his childhood was spent behind barbed wire in a Soviet gulag.

Weiss' father, Karl-Friedrich Weiss, was a scientist with knowledge of nuclear physics considered so valuable that at the end of the Second World War he was immediately taken into custody by U.S. forces. He escaped (on a stolen bicycle), only to be captured again, this time by the Soviets. The next 10 years were spent in a special camp built for nuclear researchers and their families. After being allowed back to Germany in the 1950s, Weiss—by then a committed Christian—says he could not follow his father's career because it would have meant "giving up my conscience." Success in Leipzig could come only by full participation in the totalitarian state. So, Weiss says, he abandoned thoughts of doing outstanding research and accepted the fate of an undistinguished lecturer.

Weiss sees the wounds of former times everywhere in the university. For example, the East German education system eliminated almost everyone who, like Weiss, thought differently. Even the students became a force of conservatism, which became obvious when "astonishingly few" of them took place in the crucial demonstrations, says Konrad Taut, a theologian at the university. "It is scary how effectively all critical opinions were silenced" under the communists, Taut says. "Their ideological machinery worked even better than that of the Nazis. Students were 'sieved' so carefully that they were politically paralyzed and unwilling or unable to think by the time they reached the universities. Virtually anyone critical was filtered out at age 15."

Weiss says he saw this filtering process in action time and again. For example, one student he knows was kept out of a state research program for molecular biology because she had named her daughter Sarah, a name which to the authorities "sounded suspiciously Zionist." Weiss himself says that he could have been arrested at any time and imprisoned for 3 years just for having a copy of a Western newsmagazine lying around his house. The charge would have been "possessing morally repugnant literature."

This atmosphere, says Weiss, has resulted in an inability to think independently. In ordinary life, "you had to be schizophrenic in order to survive," he says. "There was not a single person in this country who was not schizophrenic. Even the highest Party officials had antennas on their roofs so they could watch West German TV," but no one could admit to this in public, he says. Often, the state demanded that researchers, especially in physical sciences and computing, just copy what was already being done in the West. As an example, Weiss points to a state-ordained research topic in the chemistry department to determine how Kodak film was put together.

> ***"You have to be very tough when someone's guilt has been proven. This has made me lots of enemies."***
> —Cornelius Weiss

Progress in the ideological struggle is hard to gauge, but at least Weiss can point to one area in which he has reawakened the critical responses of his faculty members: He says he has the feeling that hostility toward him is growing "from both the left and the right. In order to be fair to the victims, you have to be very tough when someone's guilt has been proven. This has made me lots of enemies, especially among former Party members. At the same time, I am making no sweeping judgments—to fire all the old communists, for example—which leaves the anti-communists dissatisfied."

Indeed, Weiss says he expects to be voted out at the end of his first term in 2 years, a prospect he views with equanimity. He is confident that this will be enough time to institutionalize the mechanisms of renewal that he has set in motion. But he predicts that the process itself—like coming to terms with the country's Nazi past—will "take decades." ■ STEVEN DICKMAN

Steven Dickman is a free-lance science writer based in Munich.

Coping With

Two and a half years ago, the structural biology community found itself caught in an unaccustomed furor. Charging that many of their colleagues weren't playing by the rules, disgruntled researchers had mounted a protest against those who published papers on the structure of protein or nucleic acid molecules without archiving the underlying data in a national databank established 18 years earlier at Brookhaven National Laboratory (*Science*, 15 September 1989, p. 1179). These holdouts, the scientists argued, were slowing the course of legitimate inquiry by withholding the core of their research work from those who wanted to verify or extend it. "The fact is, there were people not submitting data," says Helen Berman of Rutgers University. "The protein databank had to beg people to put their data in."

Since then, the situation has been reversed dramatically—but in some ways, not much has improved. Where once Brookhaven database director Thomas Koetzle could point to more than 70 protein molecules whose structures had been published but whose coordinates had not been deposited in the database, he now has to admit he has a backlog of 330 protein structures just waiting to be processed for the archives—about a year's worth of work. The irony is not lost on those who brought the original complaints. "There is an embarrassment of riches," says Richard Dickerson, a nucleic acid crystallographer at the University of California at Los Angeles and one of the first researchers to raise the alarm in 1989. "It really is a disgrace."

Over the past several months, crystallographers have called a flurry of meetings to correct the situation, and many prominent researchers in the field seem to agree that the prospects for eliminating the backlog within a year or so are bright. Along the way, however, structural biologists and other scientists in data-intensive specialties may find some broader lessons in the sudden reversal of the databank's fortunes. The story of the Brookhaven database is more than just another example of the law of unintended consequences. It is a cautionary tale for the managers of large scientific databases, who—especially when they are not active researchers in the field—can find themselves suddenly overrun by data when the underlying technology of their field advances rapidly.

Many crystallographers agree that

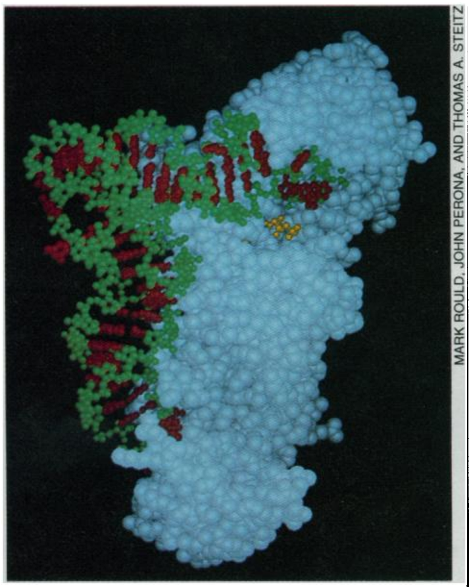
an "Embarrassment of Riches"

Brookhaven's problem started with policy changes spurred by the debate in the late 1980s. At that time, several prominent journals, including *Science*, began requiring authors to archive structure data upon publication, and the International Union of Crystallographers passed a similar resolution, although it allowed researchers to delay public release of archived data for up to a year. "Everybody complied, and then there were a ton of coordinates," says Berman.

But the root causes of the backlog appear to predate the controversy over withheld data, going back to technological and sociological changes that have vastly boosted research productivity in structural biology. For many years, until about 1987, submissions to the database trickled in at a rate of 30 to 40 structures a year. "Overnight, that number exploded," explains University of Alabama crystallographer Charles Bugg. By 1991, the database was receiving 500 submissions annually, or just about the same number every month it had once received in an entire year.

Much of the surge has come from improvements in data collection techniques and the increasing popularity of structural biology. The advent in the mid-1980s of "area detectors," which allowed researchers to take hundreds of x-ray diffraction data at once, rather than one at a time, sped up data collection dramatically, says Keith Watenpaugh, an Upjohn researcher and president of the American Crystallographic Association. Now, he says, crystallographers can get structure data in a few days that, 10 years ago, might have taken months. (Even further advances are expected as synchrotron radiation sources, which produce radiation thousands of times more intense than sources currently in use, become more widely available.) Meanwhile, from 1985 to 1991 the number of protein crystallographers worldwide grew dramatically. By one crude measure—the number of subscribers on the database's mailing list—the number quadrupled.

Koetzle, who has run the database since 1973, says it would have been hard to predict the sudden surge in structure data that has resulted from the growth of the field and its improved technology. "Up to about 5 years ago, submissions increased at a constant rate for many years," he says. "We always thought it would take off, but it was difficult to say when that would be."



MARK ROULD, JOHN PERONA, AND THOMAS A. STEITZ

Molecular embrace. ***Structures like this enzyme-RNA complex are backlogged at the database.***

Koetzle says he did try to keep pace by continuing to hire steadily over the past 5 years—over that time, his staff grew from three to 10—and installing a new computer system in 1987.

But some members of the crystallographic community say that Koetzle didn't do enough to prepare for the flood of new data. "I think he's tried to be very responsible, but he was clearly caught by a sudden avalanche of data without the pieces in place to handle it all," says Berman. And one of Koetzle's former Brookhaven colleagues, who now manages a protein engineering section for a major biotechnology company, complains that the database management lacks the experience in protein crystallography that would make it more responsive to the community's needs. For instance, this researcher points out, Koetzle himself has no working experience as a protein crystallographer. "At the very top, there's no leadership in terms of how this database can be developed further to serve the community," says the industry scientist. "It's all extracted from outside advisers."

Koetzle admits that he hasn't ever worked on a protein structure himself, although he argues that he's familiar with crystallography and its techniques, having published more than 100 papers in inorganic crystallography. He also notes that his co-investigator at the database, Enricci Abola, was trained in protein crystallography. Even so, he agrees that his database could stay in closer touch with trends in research, and says part of his long-term plan is to hire several young researchers who could provide the "fertilization" he and his staff need—a move the industry scientist applauds as "a great step forward."

Active research experience can certainly help a database manager, to judge from the experience of the Small Molecule Database in Oxford, England, which archives structure data for a variety of organic molecules. Growth in the number of submissions has been "exponential," says director Olga Kennard, but the database has kept up through incremental improvements in staffing and computer technology. Despite the fact that its staff must key in most of its submissions, instead of receiving them electronically as Brookhaven does, the Oxford database has a backlog of only 1.2 months, or 1000 structures out of the 10,000 submitted annually. Kennard herself credits her background. "Through my own research group, I remained part of the research community as a user," she says. "Perhaps one has a better feel for trends in that situation."

Before Koetzle can turn his attention to developing that kind of intimacy with the field he needs to deal with the backlog, which he says he can eliminate by mid-1992. The key elements of his strategy are a $770,000 capital upgrade of the database's aging computer system and several additions to his staff of 10—two additional database production assistants this year, and three or four more in 1993. At the same time, he has also recently introduced a "preprocessing" routine intended to catch obvious errors in structure data within a few weeks of its receipt—essentially a "quick and dirty" version of the elaborate data checking normally performed at the database. This procedure should allow the database to release—with appropriate disclaimers—all the backlogged structures as preprocessed data within a month or so.

Even if the predicament of the protein database is nearing a solution, it's not likely to remain an isolated case, as public databases spring up throughout science, from NASA's proposed Earth Observation System to the human genome project. Planning for data explosions may well become one of the challenges for database managers in the 1990s. ■ DAVID P. HAMILTON