BEYOND DISCOVERY™

THE PATH FROM RESEARCH TO HUMAN BENEFIT

DISARMING A DEADLY VIRUS

PROTEASES AND THEIR INHIBITORS

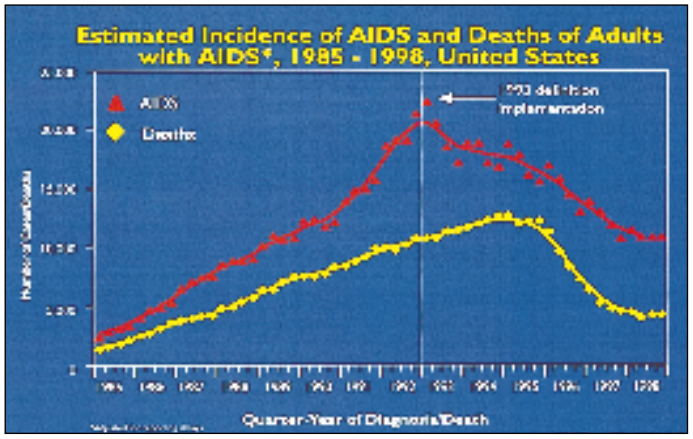
In the 1980s most people who contracted HIV, the human immunodeficiency virus that causes acquired immune deficiency syndrome (AIDS), could expect an inexorable destruction of their immune system, followed by an onslaught of appalling diseases, weakness, wasting, and death. Today, in the United States and other western nations, the fate of people who are HIV positive is dramatically different, thanks to the introduction in 1996 of a class of drugs called protease inhibitors. Between 1996 and 1998 the number of deaths attributable to HIV infection dropped more than 70 percent and AIDS ceased to rank among the top 10 leading causes of death in the United States. Indeed, the AIDS death rate in 1998 was the lowest since 1987, the first year such rates were available, and it is anticipated that the rate continues to fall.

Protease inhibitors, though a new weapon in the fight against AIDS, have actually been in medical use for almost 20 years, since the development of Captopril, a drug widely used to lower blood pressure. As doctors and patients are all too aware, protease inhibitors are not a cure for AIDS, and they also cause serious side effects. Moreover, because HIV tends to mutate readily, populations of drug-resistant virus will emerge sooner or later during treatment. Still, for those who were desperately ill only a few years ago, the advent of drugs specifically designed to act as protease inhibitors have meant nothing short of a new lease on life. The following article traces the basic research that led to the development of these life-lengthening drugs, beginning with scores of scientists seeking a deeper understanding of human biochemistry.

An Enzyme with Bite

All of the protease inhibitor drugs now in use against HIV, as well as those being developed to fight other infectious diseases and even cancer, owe their existence to a discovery made more than 120 years ago by the German scientist Wilhelm Kühne. In 1876 Kühne found a substance in pancreatic juice that degraded other biological substances. He also discovered that the substance, which he called trypsin, originates as an inactive precursor—trypsinogen—that is converted to active trypsin.

Kühne was working long before scientists had parsed out the chemicals of life. Today scientists know that the nucleic acids deoxyribonucleic acid (DNA) and ribonucleic acid (RNA) are the repositories of our genetic information.



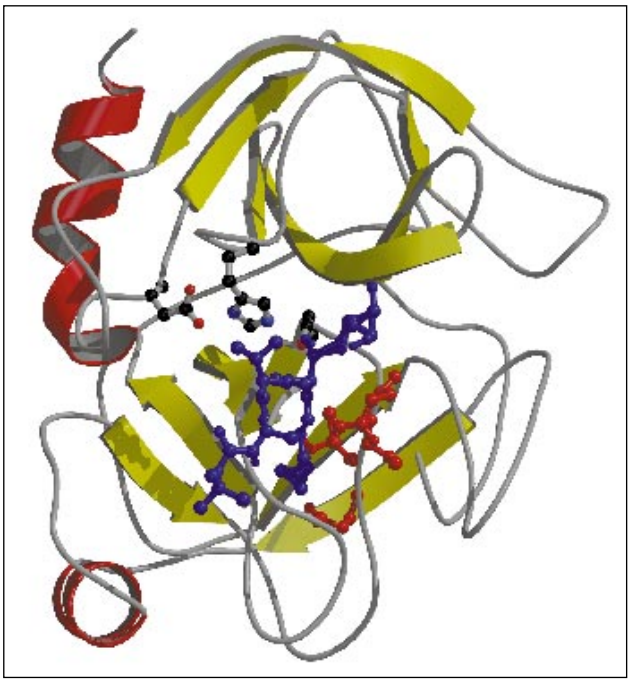
Estimated Incidence of AIDS and Deaths of Adults with AIDS, 1985-1998, United States (Centers for Disease Control/ Dr. Demetri Vacalis)

DNA and RNA specify the exact order in which chemical entities called amino acids must be strung together to build proteins, which in turn are essential to cell function. Some proteins, for example, form a skeletonlike network that gives a cell its structure. Others help the cell to read and replicate the genetic instructions coded in its nucleic acids. Still other proteins, called enzymes, act as catalysts for the biological chemistry that goes on within cells. Some enzymes help manufacture proteins; other enzymes cut them.

In his discovery of trypsin, Kühne was the first scientist to find an enzyme that degraded proteins. Trypsin turned out to be a protease, a class of enzymes whose function is now understood to be to cut proteins by breaking the bonds between amino acids. Indeed, proteases abound in human cells, for cells need the help of different proteases for many of their functions.

Kühne's finding was the first crucial step in the development of the protease inhibitors used to fight HIV, but more than half a century would pass before scientists gradually began to build on it. In the early 1930s John H. Northrop and Moses Kunitz, of the Rockefeller University in New York, isolated trypsin and another related digestive enzyme, chymotrypsin, and identified them as proteins. Two decades later, in 1952, Frederick Sanger, of the University of Cambridge in England, developed a method to determine the sequence of amino acids in proteins. In the 1960s other scientists determined the amino acid sequences of trypsin and chymotrypsin. These enzymes—the major workhorses of digestion—together degrade the proteins in food. By comparing the sequences of these two closely related but functionally different enzymes, scientists gained an early view of which parts of the proteases were important for their function.

Even as these efforts to decode the building blocks of enzymes were under way, other researchers had been taking a different approach to understanding enzymes using a technique called x-ray crystallography, which was developed in the 1910s. X-ray crystallography allows scientists to determine the positions of atoms in a molecule, producing a kind of picture of the molecule's three-dimensional structure. Proteins that can be induced to form crystals (and currently not all of them can) can be probed by x-ray crystallography. In the 1950s John Kendrew and Max Perutz, of the Medical Research Council's Laboratory of Molecular Biology in England, were the first to successfully apply the technique to the study of protein structure. By the late 1960s scientists had determined



Representation of the structure and sequence of the enzyme trypsin, a member of the protease class of enzymes. (Robert Stroud/University of California, San Francisco).

the three-dimensional structures of trypsin, chymotrypsin, and another digestive protein—elastase. Comparison of the structures of these closely related proteases yielded a wealth of information on how these enzymes function. Thus, by the late 1960s various scientists around the world were beginning to understand both the chemical makeup and the molecular structure of enzymes. During this same period, Christian Anfinsen and colleagues, of the National Institutes of Health in Bethesda, Maryland, thought to unite the two approaches by looking into the relationship between the sequence of an enzyme's amino acids and its shape. Anfinsen proved that the specific folds of a given protein molecule are uniquely determined by the sequence of its 20 different amino acids. For example, amino acids that repel water are likely to end up in a protein's interior, where they would not interact with the watery environment of a cell. Water-soluble amino acids, by contrast, are more likely to be found on the protein's exterior surface. The three-dimensional structures of trypsin, chymotrypsin, and elastase fit that model nicely. Anfinsen further proved that the amino acid sequence is in turn laid down by the genes that code for the protein.

In time researchers began the process of piecing together how enzymes work and the significance of their many folds and twists—investigations that continue today. Every enzyme, whether its function is to synthesize or cleave, acts on molecules of a

specific substance, called the substrate, which may be any one of a wide range of substances in the body, including proteins and nucleic acids. The work takes place in a region of the enzyme called the active site, a groove or pocket shaped specifically to accommodate one or at most a few similarly shaped substrates. One imperfect though helpful analogy likens the substrate to a hand and an enzyme's active site to a glove. Just as a glove can accept hands of many sizes—provided they are all basically hand shaped—it cannot accommodate more than one hand at a time. Once a given hand is in a glove, even if it does not fit perfectly, the glove is effectively unavailable for other hands. When the inhibiting compound binds to the active site, it prevents the enzymatic reaction from proceeding.

Deciphering both the structure of an enzyme and the sequences of its amino acids lets scientists see what the active site looks like and allows them to reconstruct the actual chemical reaction that takes place between the enzyme and its substrate. This in turn is a powerful tool for creating inhibiting compounds for specific enzymes. In fact, scientists try to design inhibitors that have an even greater affinity for the active site than does the natural substrate, so that the inhibitor will win the fight to bind with the enzyme.

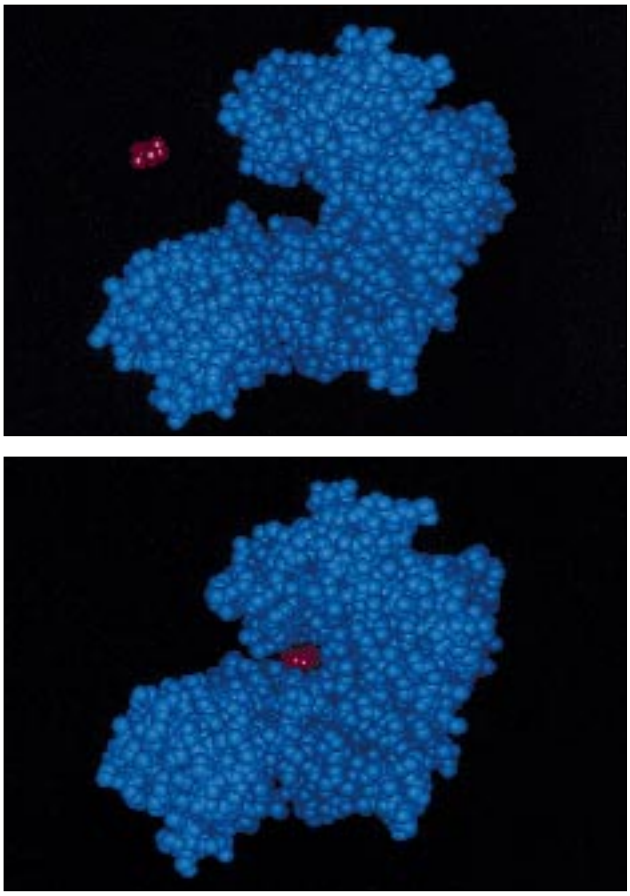


Diagram of enzyme/substrate binding (Thomas A. Steitz/ Yale University)

Inhibiting by Design

As might be expected, many proteases have natural inhibitors; a cell could not endure for long if these powerful enzymes were not regulated. Although the first natural protease inhibitors were identified by Northrop and Kunitz as part of their protease studies in the 1930s, the era of designing protease inhibitors in the laboratory did not get under way until a decade or so later, and even then progress came piecemeal. In 1949 Arnold Kent Balls, of the Western Regional Research Laboratory in Albany, California, discovered that a synthetic substance—a nerve gas in fact—could inactivate acetylcholine esterase, an enzyme involved in the processing of the neurochemical acetylcholine. This inhibitor worked by reacting with the amino acid serine in the enzyme's active site. It turned out that the proteases trypsin, chymotrypsin, and elastase also were inactivated by the nerve gas, suggesting that they too possessed the amino acid serine in their active sites. Scientists concluded that all these proteins formed a family related by mechanism, a concept of relatedness that would, in another 20 years, prove useful time and time again.

The first test of this concept came in the 1970s, in the course of efforts by David W. Cushman and Miguel A. Ondetti, of the Squibb Institute for Medical Research in Princeton, New Jersey, to find a means to control high blood pressure. Cushman and Ondetti were investigating ways to inhibit the angiotensin-converting enzyme (ACE), one of two proteases required to make angiotensin, a hormone found in the kidney that raises blood pressure. No crystallographic image existed of ACE, but scientists understood something of how it worked. For one thing, they knew of at least one inhibitor—a snake venom protein—that worked on ACE. For another, they knew that ACE required charged metal atoms, or ions, to function, making it a member of a family of proteases called metalloproteases because they contain metal ions in their active sites. Looking around for similar proteins whose structures were already known, Cushman and Ondetti turned to carboxypeptidase A, another metalloprotease whose structure had been solved a few years earlier by

W. Lipscomb and his colleagues at Harvard University. Cushman and Ondetti extrapolated from the shape of carboxypeptidase A a prediction as to which chemicals might alter the activity of ACE. The result of their effort was the development in 1979 of the ACE inhibitor Captopril, the first protease inhibitor developed through the process of "drug design," meaning a drug is fine-tuned by a kind of step-by-step modification and retesting. Captoptril has been in medical use ever since.

More solutions by analogy followed, as scientists interested in finding other substances to lower blood pressure looked to solve the structure of renin, a protein that works in concert with ACE to make angiotensin. Renin was difficult to crystallize, but clues to its structure came, albeit in roundabout fashion, through investigators who were looking into the structure of the digestive enzyme pepsin. Both renin and pepsin are members of a family known as aspartic proteases for the amino acid aspartate in their active sites. Although scientists took the first x-ray crystallography images of pepsin in the 1930s, they had a difficult time unraveling the structure of pepsin. It was not until the 1970s that research on the structure of fungal acid proteases served as a guide for solving the structure of pig pepsin.

The pepsin structure in turn was a leg up for scientists searching for a renin inhibitor. During the 1980s researchers used an ingenious technique to develop inhibitors that effectively halted renin's enzymatic reaction. In essence, scientists focused on the instant in the reaction process when the enzyme binds most tightly to its substrate. They examined the structure of the substrate during this critical transition state and designed an inhibitor that would mimic not the final form of the substrate-enzyme reaction but the substrate's structure during this highly attractive transition. As effective as these renin inhibitors were in the laboratory, though, they were metabolized in the liver before they could reach the kidneys, where renin is made. The renin inhibitors thus were not useful as blood pressure drugs and were shelved—at least for the moment.

The AIDS Challenge

Just as molecular biology and drug design were coming into their own, they were confronted in the early 1980s by a formidable new challenge. Doctors were starting to see patients with unusual diseases, including uncommon forms of pneumonia, rare fungal infections, and Kaposi's sarcoma. Physicians soon realized that these patients shared a common symptom—the progressive decline of their immune systems.

Time Line

This time line shows the chain of basic research that led to the development of protease inhibitor drugs for HIV.

1876
Wilhelm Kühne discovers the digestive enzyme trypsin, the first protease.

1930s
John H. Northrop and Moses Kunitz isolate trypsin and identify it as a protein. They also discover natural substances that inhibit proteases.

1949
Arnold Kent Balls discovers that a synthetic substance can inhibit a protease.

1952
Frederick Sanger develops a method to determine the sequence of the amino acid constituents of proteins.

1950s
John Kendrew and Max Perutz successfully use x-ray crystallography to solve the structure of myoglobin and hemoglobin, the first protein structures to be deciphered.

1960s
Scientists determine the amino acid sequence and structure of trypsin, chymotrypsin, and elastase.

1970
David W. Cushman and Miguel A. Ondetti develop Captopril, a drug to control blood pressure. An inhibitor of angiotensin-converting enzyme (ACE), Captopril is based on the closely related structure of carboxypeptidase A.

1970s
Scientists solve the structures of several proteases, including pepsin.

1970s
Pepsin structure guides the development of compounds to inhibit a structurally related protease, renin, in an effort to develop new medications to control blood pressure. The renin inhibitors are unsuitable as drugs and are shelved.

Whatever was responsible for this syndrome, which we now call AIDS, was killing infection-fighting immune cells, whose numbers dropped steadily as a person's disease worsened. Medicine had no way to stop the decline. But by 1983 researchers, led by Luc Montagnier at the Pasteur Institute in France, and by Robert Gallo at the National Cancer Institute in the United States, had isolated the responsible virus, HIV. The ensuing molecular studies of HIV by researchers around the world held the promise that a medical solution—a treatment, perhaps, if not a cure—would soon be at hand.

Montagnier had identified HIV as a retrovirus, which, like all viruses, must rely on the infected host cell for the machinery to carry out genetic instructions for viral reproduction. However, unlike other viruses, which carry viral DNA, retroviruses carry RNA and the enzyme reverse transcriptase. Retroviruses thus reverse the usual flow of information in the cell because they replicate by using RNA as a template, which reverse transcriptase converts into DNA. Scientists had been aware of retroviruses since 1911, when Peyton Rous, of the Rockefeller Institute for Medical Research, discovered a virus that could efficiently cause cancer in chickens. In 1966 he was awarded the Nobel Prize for Physiology or Medicine for his discovery. Retroviruses have now been associated with natural cancers in many animals, including birds, mice, cats, cattle, fish, and humans. Indeed, the first known human retrovirus, human T-lymphotrophic virus type-1, which causes adult T-cell leukemia, was discovered in 1981 in the laboratory of Robert Gallo.

During the so-called War on Cancer in the 1970s, researchers studying retroviruses learned that they share the same fundamental properties no matter what species they come from or what disease they cause. Scientists found, for example, that the retrovirus particle contains about 10 different proteins, which are involved in such functions as forming the virus structure, copying RNA into DNA, and inserting the viral DNA into the cell DNA. Researchers learned further that retrovirus proteins are made not as individual units but rather as polyproteins, precursors that consist of long strings containing several different proteins. For the individual virus proteins to become fully functional, they first must be cut out of the polyproteins. Scientists soon discovered that this excision was performed by the virus protease, one of the proteins in the retrovirus particle and itself a part of a polyprotein. They also found that the assembly of the retrovirus is arranged in such a way that the budding of new particles occurs before the protease becomes active. Particles newly released from infected cells thus contain only polyproteins and

1980s
AIDS is recognized as an epidemic.

1983
Luc Montagnier and Robert Gallo isolate the human immunodeficiency virus (HIV), the virus that causes AIDS.

1985
The drug AZT gains widespread acceptance as an effective anti-HIV therapy. AZT works to stop the virus from reproducing its genes.

1986
Several groups of scientists, working independently, isolate the gene for the HIV protease.

1988
Irving Sigal and colleagues show that a genetically mutated HIV protease produces immature viral progeny.

Late 1980s
Brendan Larder, Graham Darby, and Douglas Richman demonstrate that HIV becomes resistant to AZT after less than a year of treatment.

1989
Manual Navia and colleagues solve the structure of HIV protease, which is independently solved and subsequently refined by Alexander Wlodawer and colleagues.

1996
Researchers present findings that patients taking a combination therapy of protease inhibitors and other HIV drugs could reduce the amount of virus in their blood to undetectable levels.

cannot infect healthy cells until the protease cleaves itself to become active and then cleaves the polyproteins into their active forms.

By the time the AIDS epidemic struck in the early 1980s, a large cadre of scientists—veterans of the War on Cancer—had firsthand knowledge of retrovirus replication mechanisms. Thus, once HIV had been isolated and identified as a retrovirus, it did not take long for scientists to decipher its genes. Researchers used this knowledge to map many of the key details of the virus replication cycle and to identify the HIV proteins—including the all-important HIV protease. Four of the proteins made by HIV were enzymes—particularly promising targets for drug design. Scientists used the viral genes that encoded these enzymes to program bacteria to produce large quantities of the viral proteins so that they could be studied in detail. In particular, the viral protease was studied in much the same way that trypsin and other proteases had been studied 30 years before—only this time one goal was to learn how to stop its activity.

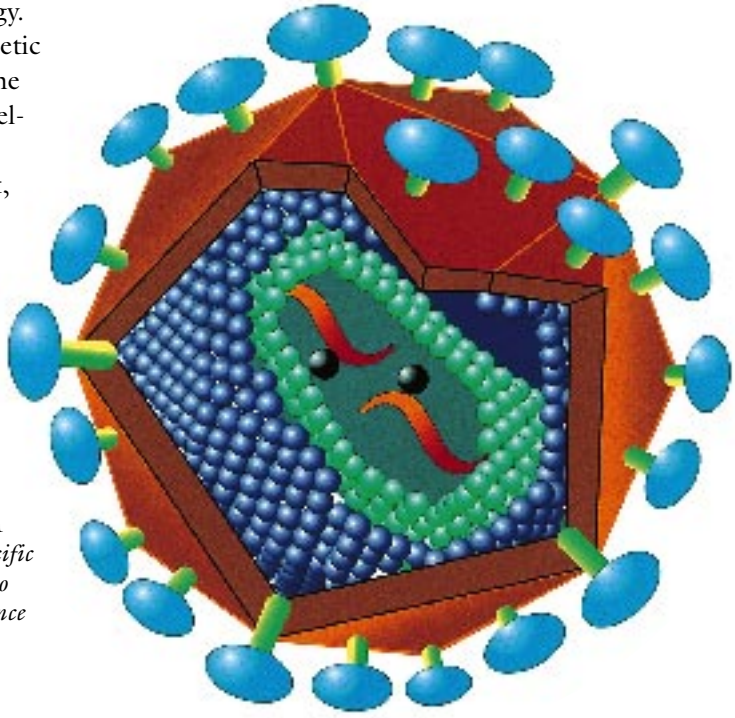
Halting Viral Replication

Broadly speaking, there are two ways to keep HIV from multiplying once it has entered the cell: stop the replication of viral genes (which involves halting reverse transcriptase) or stop the manufacture of viral proteins. Most early efforts at finding drugs effective against AIDS focused on the first strategy. By about 1985 the search for drugs to block genetic replication of HIV led scientists to azidothymidine (AZT), a reverse transcriptase inhibitor drug developed many years before and previously tested, unsuccessfully, as an anticancer drug. Simply put, AZT prevents the HIV enzyme reverse transcriptase from converting viral RNA into DNA. For a while it looked as though AZT might be sufficient to slow viral reproduction and halt progression of the disease; over the next few years many efforts were directed at developing more drugs like it.

However, some scientists began to look down the other path, exploring strategies to inhibit the manufacture of viral proteins. They began to look more closely at HIV protease and how it goes about its job of snipping out individual proteins from the polyprotein precursors. Reasoning that preventing this excision of viral proteins would either stop viral reproduction entirely or would yield crippled viral progeny, a number of researchers in both academia and industry began working to devise a therapeutic agent that could effectively block the HIV protease.

The intensified effort came none too soon. At about this time, Brendan Larder and Graham Darby, of the Wellcome Research Laboratories in England, and Douglas Richman, of the University of California, San Diego, reported devastating news: In AIDS patients taking AZT, HIV could became resistant to the drug after less than a year of treatment. Without AZT the medical community had virtually no ammunition against AIDS. Then a number of clinicians, perceiving that the virus would probably develop resistance to any one drug given alone, proposed the concept of using drug "cocktails." The idea was that different drugs would be designed to inhibit several different stages of the viral replication cycle simultaneously. Patients would take the drugs in combination, in the hopes that multiple drugs would be more potent.

At that time the only ingredients available for the cocktail were AZT and related drugs, all reverse



Artist's formulation of an HIV particle, depicting RNA and reverse transcriptase in the center, and the virus-specific proteins on the surface. Protease inhibitors are designed to interrupt and prevent the formation of these proteins, hence not allowing the HIV particle to mature and replicate. (Andrew Davies, National Institutes for Biological Standards and Control, UK)

transcriptase inhibitors. But researchers who had been studying HIV protease now hoped that a protease inhibitor would be an important addition. Encouragingly, several important breakthroughs on the HIV protease occurred in rapid succession. In the mid-1980s scientists demonstrated that protease associated with retroviruses shared a signature amino acid sequence with aspartic proteases such as renin and pepsin. This led researchers to propose that HIV might also use an aspartic protease to process its proteins. In 1986 several groups working independently isolated the gene for HIV protease and confirmed this hypothesis. By 1988 Irving Sigal and a team at Merck had used new gene technologies to create a mutant HIV protease and shown that this debilitated protease would, as known for other retroviruses, force viral replication to produce only immature viral progeny, incapable of infecting immune cells.

The following year the research community got its first look at the actual structure of the HIV protease. In 1989 two different research groups—Manuel Navia and colleagues at Merck and Alexander Wlodawer and colleagues at the Frederick Cancer Research Facility in Maryland—determined the three-dimensional structure of HIV protease. They observed that it was an aspartic protease like renin and pepsin but was much smaller and was made up of two identical protein chains instead of one. The two protein chains came together like two halves of a walnut, with the active site of the protease in the center of the two shells.

Hitting the Target

Investigators now set about to look for specific compounds that would inhibit HIV protease. Unfortunately, as scientists had already learned with renin, the characteristics that make a good inhibitor do not necessarily make a good drug. In many cases, for example, a good inhibitor may not be well tolerated by the body. Some compounds are toxic because they interfere with normal bodily functions. Others work only at dangerously high doses. And still others are not well absorbed by the intestine or are metabolized in the liver before they can reach their therapeutic target. The trick in drug design, then, is to find a compound that works specifically against the target, is nontoxic, can be given in low doses, and remains in the body long enough to be useful.

Many initial attempts to develop an HIV protease inhibitor yielded good inhibitors that turned out to be poor drugs. By 1992 prospects looked better. Because of the sequence similarities to the aspartic proteases, researchers had renewed their interest in the old renin inhibitors, modifying them for use against the HIV protease. Soon they reported the development of potent protease inhibitors based on renin inhibitors that were also promising drug candidates. In December 1995 the first protease inhibitor, Saquinavir, was approved by the U.S. Food and Drug Administration (FDA). By the spring of 1995 two more inhibitors—Ritonavir and Indinavir—were approved.

Then in July 1996 at the International Conference on AIDS in Vancouver, British Columbia, a number of presenters made announcements that sent a thrill of hope through the AIDS community. David Ho of the Aaron Diamond AIDS Research Center in New York, for example, described the remarkable results he and his colleagues had observed with a drug cocktail that mixed an HIV protease inhibitor with the older AZT-type antivirals, more of which had come on the market starting in 1991. He reported that, although the combination therapy was not a cure, it not only boosted AIDS patients' immune cell counts but also reduced the amount of virus in the blood to undetectable levels.

Today, in remarkably short order, five HIV protease inhibitors have made it through the development process, been approved by the FDA, and are included in the AIDS drug cocktail. All five of them—Amprenavir, Indinavir, Nelfinavir, Ritonavir, and Saquinavir—are the chemical descendants of, though structurally distinct from, renin inhibitors. The scientists who developed these drugs were guided by structural models of the HIV protease, which revealed that the key to activating the virus, enabling it to multiply and spread, was the breaking of the chemical bond between the amino acids phenylalanine and proline. Researchers were excited to find that, although the viral protease can break the bond between the amino acids, no mammalian proteases can do so selectively. Thus, any drug designed to act on the enzymes that break the peptide bond between phenylalanine and proline is very specific for the viral protein. Since the drug should not interfere with any other normal mammalian functions, it should have fewer side effects. The structural models also showed that HIV protease is a symmetrical molecule—that is, its two halves are symmetrically related. At first scientists tried to devise a symmetric inhibitor—until they realized

A common Highly Active Antiretroviral Therapy (HAART) regimen consisting of a combination of different drugs (National Institute of Allergy and Infectious Diseases/ National Institutes of Health)

that during the transition state, at the instant of tightest binding between the inhibitor and the protease, the inhibitor was actually asymmetric. Asymmetric inhibitors proved to be better in the long run, being more easily used by the body than symmetric compounds.

New Vistas

Despite their demonstrated effectiveness, the protease inhibitors have drawbacks. For one thing, they don't work for everyone. For another, because of HIV's rapid mutation rate, resistance has in fact become a problem. And even when protease inhibitors do work, they can have severe side effects. For example, protease inhibitors may dramatically alter the way the body metabolizes and stores fat. Some AIDS patients have to take other drugs to counter side effects, including nausea, diarrhea, kidney stones, depression, and anxiety.

Pharmaceutical companies continue to work to develop new generations of safer protease inhibitors, with an eye to figuring out how to prevent the virus from becoming resistant to them. Meanwhile, researchers are finding that the activity of existing inhibitors might be improved by combining them in different ways. For example, some inhibitors slow drug metabolism in the liver, and these might be used to prolong the time other protease inhibitors remain in the body. In effect, one protease inhibitor might help lower the required dose of another.

The success—however guarded—of protease inhibitors in the treatment of AIDS has ignited an explosive interest in their use for other diseases. Already, the structures have been solved for the proteases found in hepatitis C virus; cytomegalovirus, a herpes-type virus responsible for fatal organ infections in people with AIDS; and rhinovirus, better known as the cause of the common cold. Even as scientists look to solve more viral protease structures, they are starting to realize the importance of proteases beyond viral disease. For example, proteases may be involved in osteoporosis and inflammation, and they also may kill brain cells inappropriately, thereby contributing to stroke or Alzheimer's disease.

The real success of the HIV protease inhibitors is their revolutionary effect on modern drug design. HIV protease is now the model for structure-based drug design, including recent efforts to design vaccines against the virus. The techniques developed over the past 10 years in the search for HIV protease inhibitors built on the foundation laid by the basic amino acid and x-ray crystallography research. Thanks to scientists who were seeking a deeper understanding of biochemistry, patients with a deadly disease can begin to think about the future again.

This article, which was published in 2000 and has not been updated or revised, was written by science writer Roberta Conlan with the assistance of Drs. Charles Craik, David Davies, Anthony Fauci, Gary Felsenfeld, David Ho, Hans Neurath, Max Perutz, Robert Stroud, Nancy Thornberry, George Vande Woude, Ms. Michelle Hoffman and Mr. Sean O'Brien Strub for ***Beyond Discovery™: The Path from Research to Human Benefit****, a project of the National Academy of Sciences.*

The Academy, located in Washington, D.C., is a society of distinguished scholars engaged in scientific and engineering research, dedicated to the use of science and technology for the public welfare. For more than a century it has provided independent, objective scientific advice to the nation.

Funding for this article was provided by the Camille and Henry Dreyfus Foundation, the Pfizer Foundation, Inc., and the National Academy of Sciences.

 February 2000