Another Piece of the HIV Puzzle Falls into Place

Alexander Wlodawer

The human immunodeficiency virus (HIV), a retrovirus, is generally accepted as the causative agent of AIDS. Retroviruses encode their genetic information as RNA, and thus must undergo reverse transcription by an enzyme called reverse transcriptase (RT) into a DNA copy, which is then integrated into the genome of the host. Because human cells do not need RT, this enzyme is an inviting target in a search for antiviral compounds.

The current AIDS drugs are RT inhibitors. These nucleoside drugs, such as AZT and ddI, resulted from the usual drug search. However, biological resistance usually develops when they are used continuously; also the mode of binding to the target enzyme remains to be established. Recently, non-nucleoside inhibitors of RT have been identified, but the initial excitement was dampened because resistance can develop in as little as a month. Very little was known about the mode of inhibition by these compounds, and hence the need for a structural understanding of RT became even more critical if drugs to inhibit RT are to be designed rationally.

The process of "rational" drug design requires accurate information concerning the atomic structure of the target molecule. In a cycle of drug design (*1*), the structure of the target is identified by x-ray crystallography, nuclear magnetic resonance, or molecular modeling; promising compounds are designed and synthesized; their binding and inhibition properties are analyzed; and the structures of the complexes are again determined. If the inhibitors have appropriate properties, they can be further analyzed for behavior that would mark them as potential drugs, and if not, the cycle is repeated. Rational drug design has been attempted with some success in the synthesis of inhibitors of targets such as renin for treatment of hypertension and thymidylate synthase for cancer.

The goal of elucidating the structure of RT has now been reached. In this issue of *Science*, Steitz and his colleagues at Yale University describe the 3.5 Å structure of a complex of RT with a non-nucleoside inhibitor, Nevirapine (Boehringer Ingelheim) (*2*). This structure is an eagerly awaited event in protein crystallography, and the solution came about after several years of difficult work in a number of laboratories. A description of the structure of a ternary complex of RT with DNA and an antibody fragment (Fab) has also just been published (*3*), but because of very low resolution (7 Å) of the phased data, the structure cannot be assessed in detail. In contrast, the 3.5 Å structure from Yale has yielded a wealth of unexpected details.

The difficulty in solving the RT structure can be traced largely to obtaining well-diffracting crystals. Different approaches to improve diffraction were tried, such as cocrystallization with DNA, antibody fragments, or both (*3*), but what worked for Steitz and his colleagues was co-crystallization with the non-nucleoside inhibitor. The presence of the inhibitor has made it possible to extend the resolution to almost 3 Å. Perseverance and patience paid off.

As seen in the maps that are of excellent quality after having been improved by almost every procedure on record, the general features of the molecule are not in doubt. Probably most intriguing is the large difference between the two polypeptide chains p51 and p66, which share a common primary structure, with the addition of ribonuclease H (RNase H) in p66. While each of the four domains in both chains is similar (they were averaged in the map improvement steps), the relative orientation differs in the two chains. Only one active site can be seen on the p66 chain, while the same residues are deeply buried on p51 and thus unable to contribute to the catalytic activity. Interestingly, the structure explains quite well why protease is unable to cleave the RNase H–polymerase junction in the p66 chain after cleaving the first one. The first cut results in a rearrangement that in turn shields the equivalent region from further proteolytic attack.

RT is not the first HIV-encoded enzyme of a known structure. The first to be solved, was the aspartyl protease (PR), which is necessary for virion maturation. The structure of the apoenzyme was solved by a group from Merck early in 1989 (*4*). Soon thereafter, the structure of a complex with a substrate-based inhibitor was published (*5*).

These structures were used in academic and industrial laboratories to assist in the design of new and improved inhibitors. Investigators in that work are determining their own crystal structures as well; at least 120 different PR-inhibitor complexes have been examined in 17 different laboratories. A potent inhibitor of HIV PR is now in phase II clinical trials in England. That compound, Ro 31-8959, was designed and synthesized at Roche (Welwyn) on the basis of analyses of PR-inhibitor crystal structures. Its mode of binding to the target enzyme corresponded to the predictions based on previous structural data although some unexpected features for parts of the structure were also reported (*6*).

More recently, the structure of the RNase H domain of HIV RT was published (*7*). While HIV RNase H most probably does not exist by itself and this form of the enzyme had no detectable activity the structure became important in the later interpretation of the structure of RT, since the coordinates of the domain provided guideposts in the interpretation of the electron density maps. Even at the very low resolution investigation of the RT-DNA-Fab complex (*3*), it was possible to identify the part of the map corresponding to RNase H. The Steitz group replaced parts of the experimental map in the area corresponding to RNase H with one with calculated electron density and then used the edited map in the cyclic refinement.

The structure of a complex of RT with a non-nucleoside inhibitor brings us closer to having sufficient information for the next cycle of drug design. However, because the refinement data (coordinates) are only for the Cα, we should be aware that, except for some regions in which other data are available, the accuracy may be limited to within two or three positions in the chain. Thus, assessing the probability that any particular compound might be a good inhibitor, whether by binding in the active site or in the area where Nevirapine binds, is also limited. Nor can the current structure be used to explain the molecular basis of drug resistance, although it can certainly help in the interpretation of such data. Nevertheless, the important first step has been successfully taken.

The author is at the Macromolecular Structure Laboratory, National Cancer Institute–Frederick Cancer Research and Development Center, ABL-Basic Research Program, Frederick, MD 21702.

REFERENCES AND NOTES

1. K. Appelt *et al.*, *J. Med. Chem.* **34**, 1925 (1991).
2. L. A. Kohlstaedt *et al.*, *Science* **256**, 1783 (1992).
3. E. Arnold *et al.*, *Nature* **357**, 85 (1992).
4. M. Navia *et al.*, *ibid.* **337**, 615 (1989).
5. M. Miller *et al.*, *Science* **246**, 1149 (1989).
6. A. Krohn *et al.*, *J. Med. Chem.* **34**, 3340 (1991).
7. J. F. Davies *et al.*, *Science* **252**, 88 (1991).
8. I thank C. Michejda, R. Rubin, and A. Swain for their comments.