

Database of three-dimensional structures of HIV proteinases

Sir—In only seven years since the first structure of an inhibitor complex of an aspartic proteinase encoded by human immunodeficiency virus type 1 (HIV-1) was published^{1,2}, a new and very promising class of anti-AIDS drugs has been developed. Three of them (saquinavir, indinavir, and zidovudine) have been approved within the past year for clinical use, while others are in advanced clinical studies^{3,4}. Introduction of HIV PR inhibitors as drugs is the first unambiguous and significant success of structure-assisted drug design. Many HIV PR-inhibitor complexes have been subjects of X-ray crystallographic analyses (reviewed in refs 5,6). Since the vast majority of crystal structures of HIV PR were solved in pharmaceutical companies, only a small fraction of them are publicly available. With much of the initial development effort having been completed, there is some danger that many of these structures could ultimately be lost. An effort to collect all of the structures of HIV-1, HIV-2, and SIV PRs into a single database aims to prevent such an outcome and to help in future drug design efforts. This database provides a unique source of

information about ligand-enzyme interactions, and it is freely available to the scientific community.

The database is designed as a WWW service accessible through the Internet; the homepage URL is <http://www-ffsc.ncifcrf.gov/HIVdb>. The information therein consists of two main parts—the informal part and the analytical and modelling part. The informal part of the database contains information about complexes, inhibitors, and enzymes, as well as the original sets of coordinates as obtained from the Brookhaven PDB, or directly from individual depositors.

All of the original coordinate files have a unified annotation, and they are stored in the PDB format widely used as input for molecular modelling programs. To visualize the structural information contained in these files, we have included a PDB-fluent Java applet. Alternatively, more complex renderings are available with simple molecular graphics packages such as RasMol⁷. All of these programs are publicly available and they work on many different computing platforms. For those users who have no access to professional molecular modelling software, the facility will be the only way to manipulate the structures in real time.

The second part of the database—analytical and modelling—is organized by services, giving access to various tools. The user will be able to immediately view the results of most of the analytical programs. Some of the computationally complex calculations, including the measurement of distances, angles and torsion angles, will be run in batch mode and the results will be returned by e-mail. This part also contains coordinates of the complexes trans-

formed into a common frame of reference, as well as the separate files of the PRs, inhibitors, and water taken from the coordinates of the complexes. Finally, since we assume that many users will be interested in different modes of superpositioning of the structures, we provide the program ProFit⁸ to compute different sets of transformations.

Among the different quantitative measures that characterize PR-inhibitor interactions, we paid particular attention to volume and surface calculations. Such analysis can aid in the process of understanding the properties of binding sites and moreover, the volumetric criteria (Fig.1) combined with reasonable template structures, could be used to direct a search for novel inhibitors⁹.

While the database, in its current form, is composed primarily of previously released structures in the Brookhaven PDB, new structures can be deposited using the WWW service. Depositors can use a prepared form for filling out all of the necessary information. The coordinates (in PDB format) can then be copied to the online form itself, or they can be sent by way of anonymous ftp. The address of our ftp server is <mars.ncifcrf.gov> and the structures can be placed in the <pub/hivpr/incoming> directory. As we mentioned in the original announcement of the database¹⁰, we expect that the user community will raise additional questions and suggestions of topics for inclusion.

Jiri Vondrasek¹, Christopher P. van Buskirk² and Alexander Wlodawer¹

¹Macromolecular Structure Laboratory, NCI-Frederick Cancer Research and Development Center, ABL-Basic Research Program, Frederick, Maryland 21702, USA

²The Krieger Mind/Brain Institute, Johns Hopkins University, Baltimore, Maryland 21218-2685, USA

Correspondence should be addressed to J.V. vond@proxima.ncifcrf.gov

Received 17 September; accepted 3 October 1996.

- Navia, M.A. *et al.* *Nature* **337**, 615–620 (1989).
- Wlodawer, A. *et al.* *Science* **245**, 616–621 (1989).
- Kempf, D.J. *et al.* *Proc. Natl. Acad. Sci. USA* **92**, 2484–2488 (1995).
- Reich, S.H. *et al.* *Proc. Natl. Acad. Sci. USA* **92**, 3298–3302 (1995).
- Wlodawer, A. & Erickson, J.W. *Annu. Rev. Biochem.* **62**, 543–585 (1993).
- Winslow, D.L. & Otto, M.J. *AIDS* **9**, S183–S192 (1995).
- Sayle, R.A. & Milner-White, E.J. *Trends Biochem. Sci.* **20**, 374–376 (1995).
- Martin, A.C.R. ProFit. <http://www.biochem.ucl.ac.uk/~martin/#profit> (1996).
- Shoichet, B.K. & Kuntz, I.D. *Protein Eng.* **6**, 223–232 (1993).
- Vondrasek, J. & Wlodawer, A. *Science* **272**, 337–338 (1996).

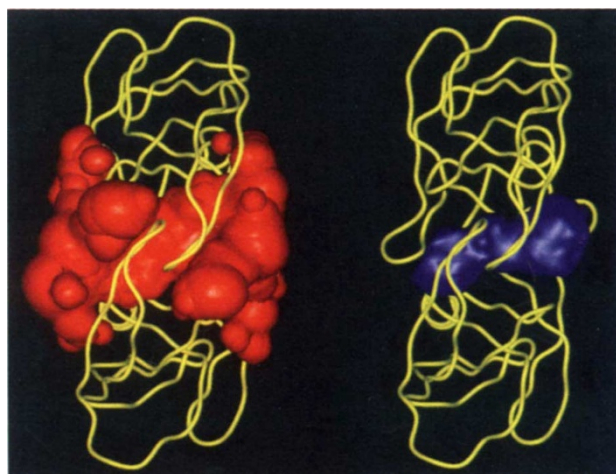


Fig. 1 Information about cavity volumes in the individual structures can be combined to derive maximal and minimal volumes of the inhibitor-binding site. After superimposing 48 structures of HIV-1 PR and calculating volumes not occupied by the protein, we have constructed by summation a maximal volume (left), and by intersection a minimal one (right). The minimal volume is most probably the limiting sterical requirement for potential inhibitors.