

Molecular Design of HIV Protease Inhibitors Using a Dead End Elimination Algorithm

David J. Huggins & Bruce Tidor
Computer Science & Artificial Intelligence Laboratory
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

<http://mit.edu/tidor/>

Background

The immune deficiency AIDS was first identified during the early 1980s due to the loss of T4 helper cells in infected patients. The virus HIV was swiftly recognized as the cause. HIV has spread rapidly in both the developed and the developing world and is one of the most significant health concerns in the 21st century.

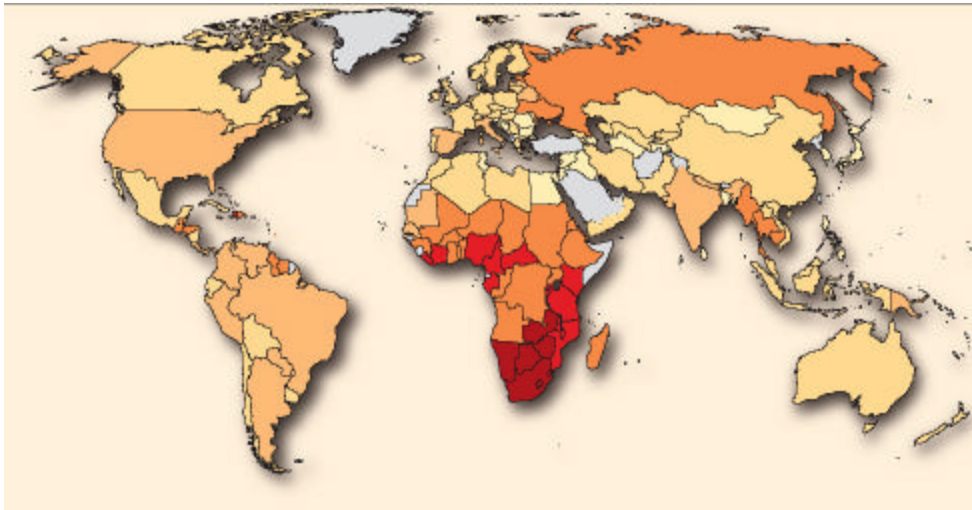


Figure 1. The prevalence of HIV infection around the world. Dark red countries are those with an infection rate over 15.0%, red countries are those with an infection rate over 5.0% and orange countries are those with an infection rate over 1.0%. There are approximately 39.4 million people infected with the HIV virus worldwide (Figure from the WHO <http://www.who.int/hiv/facts/en/>).

The virus itself is relatively simple and is comprised of a viral membrane that hosts a surface glycoprotein and encapsulates the viral RNA and a collection of functional proteins. The *pol* gene of the viral sequence encodes three very important viral proteins: HIV reverse transcriptase, HIV integrase and HIV protease. However, these three proteins are initially linked together and must be cut apart by the protease to yield

functional units. It has thus been suggested that a small molecule, which inhibits HIV protease, can retard or even stop the replication of the virus by interfering with this vital process.

Current Work

Recent work completed in this lab has focused on the creation of a molecular design algorithm (http://web.mit.edu/tidor/www/research/michael_altman2.pdf) for generating potential protein inhibitors. The algorithm requires a protein structure and a defined active site as well as one or many scaffolds for the inhibitor. A database a molecular fragments can then be connected to this scaffold and the interaction energy optimized by using dead-end elimination theory [1-5] to find and remove solutions that can be shown not to be part of the optimal solution. The remainder are then searched to find the optimal set of solutions using the A* algorithm [6, 7].

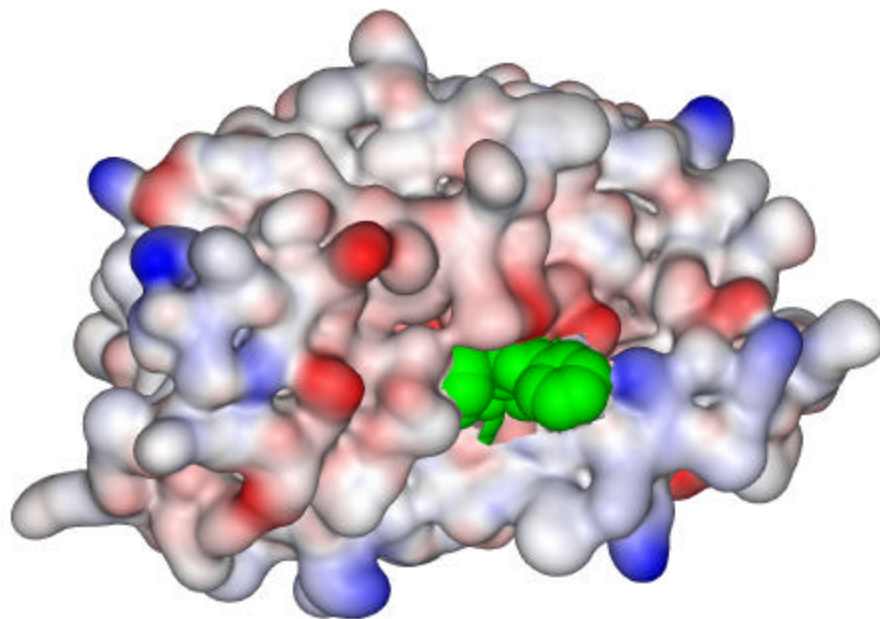


Figure 2. The molecular surface of HIV protease from the protein data bank file (1A8G) highlighting the regions of positive (blue) and negative (red) electrostatic potential. The inhibitor SDZ283-910 is shown in green.

The scoring function used in the algorithm includes a grid-based calculation of the van der Waals interaction energy and a combined calculation of the electrostatic interaction and the desolvation terms using charge optimization theory [8-10] (http://web.mit.edu/tidor/www/research/michael_altman.pdf). The algorithm has made some useful and experimentally tested predictions and the validation procedure is ongoing.

Future Work

My work will centre on the application of this algorithm to the case of HIV protease. This is a particularly important problem, due to the potential of a successfully designed inhibitor, but is also a difficult one. HIV protease, along with many of the other viral proteins, is able to mutate swiftly and easily due to the poor copying fidelity of Reverse Transcriptase. This allows the virus to adapt rapidly, providing resistance to many inhibitors.

HIV protease is also able to exist and operate in a variety of conformations and thus designing an inhibitor that will bind to these conformations becomes a difficult task. My work will be to attempt to circumvent these problems.

Research Support

This work is supported by a grant from the National Institutes of Health.

References

1. Desmet, J., et al., *The Dead-End Elimination Theorem and Its Use in Protein Side-Chain Positioning*. Nature, 1992. **356**(6369): p. 539-542.
2. Goldstein, R.F., *Efficient rotamer elimination applied to protein side-chains and related spin glasses*. Biophys J, 1994. **66**(5): p. 1335-40.
3. Gordon, D.B., et al., *Exact rotamer optimization for protein design*. J Comput Chem, 2003. **24**(2): p. 232-43.
4. Pierce, N.A., et al., *Conformational splitting: A more powerful criterion for dead-end elimination*. Journal of Computational Chemistry, 2000. **21**(11): p. 999-1009.
5. Gordon, D.B. and S.L. Mayo, *Radical performance enhancements for combinatorial optimization algorithms based on the dead-end elimination theorem*. Journal of Computational Chemistry, 1998. **19**(13): p. 1505-1514.
6. Leach, A.R. and A.P. Lemon, *Exploring the conformational space of protein side chains using dead-end elimination and the A* algorithm*. Proteins-Structure Function and Genetics, 1998. **33**(2): p. 227-239.
7. Leach, A.R., *Ligand Docking to Proteins with Discrete Side-Chain Flexibility*. Journal of Molecular Biology, 1994. **235**(1): p. 345-356.
8. Kangas, E. and B. Tidor, *Charge optimization leads to favorable electrostatic binding free energy*. Physical Review E, 1999. **59**(5): p. 5958-5961.
9. Lee, L.P. and B. Tidor, *Barstar is electrostatically optimized for tight binding to barnase*. Nature Structural Biology, 2001. **8**(1): p. 73-76.
10. Kangas, E. and B. Tidor, *Optimizing electrostatic affinity in ligand-receptor binding: Theory, computation, and ligand properties*. Journal of Chemical Physics, 1998. **109**(17): p. 7522-7545.