Display and Analysis of Protein Binding-site Topologies Using Accessible Surfaces

Regine S. Bohacek and Colin McMartin*
CIBA-GEIGY Corporation, Pharmaceuticals Division, Research Department, Summit, NJ 07901, USA

We have found that, with an appropriate selection of probe distance, the accessible surface, originally defined by Richards,[1] has a remarkable capacity for revealing the overall structure and fine details of binding site topology. A rapid algorithm has been developed for computing a contour representation of this surface: the surface of points that are at least a probe distance plus the van der Waals radius away from each atom of the binding site. When the probe distance is 1.4 to 1.6 Å, the surface develops extremely useful properties. It displays the accessible volume whose surface defines the optimum position for the centers of ligand atoms. Since the volume of accessible space is quite small, its surface gives a much sharper picture of the effective shape of the binding site than does the van der Waals surface. The surface can be colored to represent electrostatic potential or absolute field strength. The latter representation displays local polarity, an indication of hydrophylic and hydrophobic surface character. The utility of the surface in defining the space available to naturally aligned molecules and ligands has been examined using crystal data for 15 crystals of molecules of low molecular weight and seven protein ligand complexes. Different proteins were found to have striking and recurrent topological features; fine structured narrow shafts, tunnels, grooves and slit-like pockets that connect to and contrast with broad convex surfaces. By focusing attention on the spaces and surfaces accessible to ligand atoms, the surface is ideally suited to applications in computer-assisted molecular design and is fully compatible with a wide range of other graphical and computational tools.

REFERENCE

1 Lee, B. and Richards, F.M. *J. Mol. Biol.* 1971, **55**, 379–400

Large-scale Computational Methods and Visualization in Molecular and Materials Design

Erich Wimmer
Cray Research, Inc., Mendota Heights, MN 55120, USA

During the last decade we have witnessed a rapid increase in the use of computationally intensive methods in the study of biological systems such as nucleic acid and protein structures. In part, this development was driven by the increased availability of computational hardware that, in turn, has stimulated the development and use of more accurate theoretical/computational approaches. This lecture focuses on the two key approaches in theoretical molecular science: quantum mechanics and statistical mechanics. Quantum mechanical simulations are now possible for significant parts of biologically active molecules. For example, Hartree-Fock theory allows the calculation of molecular properties, such as stable conformations, relative heats of formation, charge distributions and vibrational properties of molecules, with about 10–15 first-row atoms with high-quality basis sets. Calculations on larger molecules leads to a disk-storage problem that can be circumvented with direct SCF schemes, albeit at the price of more computing time that is limited by an N^4 scaling problem.

An alternative quantum mechanical many-body technique, known as (spin) density functional theory, provides a surprisingly accurate and computationally efficient molecular orbital method that scales only with a third power. This molecular density functional approach allows the calculation of molecular geometries, charge distributions, electrostatic potentials, and relative energies of similar quality as Hartree-Fock theory with some correlation. Density functional methods can be readily applied to metallic and organometallic systems, thus opening up new possibilities in biochemical research. Calculations on molecular systems with up to 70 atoms (including Zn atoms) using polarized double-zeta basis sets illustrate the capabilities as well as current limitations of this approach.

The speech of supercomputers, combined with the convenience of graphics workstations, enables quantum mechanical calculations on the semiempirical level to be carried out in an interactive fashion. A prototype of such an integrated, distributed processing system will be discussed in conjunction with the semiempirical MOPAC code. Advances in the development of molecular force-field and statistical mechanics techniques, such as the free-energy perturbational approach, have opened up new ways of approaching the drug design problem. Here, sufficient computational power is of the essence, but certainly no guarantee for reliability, because many theoretical and computational assumptions and approximations, which are inherent in a force-field approach, have to be carefully checked before quantitative predictions can be made with chemical accuracy. The lecture will conclude with an outlook on future developments in theoretical/computational methodologies as well as computing technologies.

Studies on Polyfunctional DNA Intercalators

Michael Cory and Terri A. Fairley
Division of Organic Chemistry, Wellcome Research Laboratories, Burroughs Wellcome Co., Research Triangle Park, NC 27709, USA

Steven C. Zimmerman and Carol R. Lamberson
Department of Chemistry, University of Illinois, Urbana, IL 61801, USA

*Present address: CIBA-GEIGY Pharmaceuticals, Wimblehurst Road, Horsham RH12 4AB, West Sussex, UK.