Investigation of the Three-Dimensional Macromolecular Structures

Ján Buša\textsuperscript{a,b,c}, Jozef Džurina\textsuperscript{a,b}, Edik Hayryan\textsuperscript{a,b,c}, Ján Plavka\textsuperscript{a,b}, Imrich Pokorný\textsuperscript{a,b}, Jaroslav Skřivánek\textsuperscript{a,b,c}, Shura Hayryan\textsuperscript{c}, Chin-Kun Hu\textsuperscript{c}, Ming-Chya Wu\textsuperscript{c}

\textsuperscript{a} Technical University in Košice, Slovak Republic
\textsuperscript{b} Laboratory of Information Technologies, JINR
\textsuperscript{c} Institute of Physics, Academia Sinica, Nankang, Taipei, Taiwan

Solvent environment plays a crucial role on the structure and function of biological macromolecules, such as DNA, RNA and proteins. Exact determination of direct interactions between the macromolecules and the solvent molecules still remains a very hard mathematical and computational problem. For this reason different kinds of approximations are usually done which allow description of these extremely complicated interactions through physically averaged macroscopic parameters. Such effective parameters include, for example, solvent accessible area of the solute molecule (in the calculation of thermodynamic properties and three-dimensional structure of macromolecules) and the so called excluded volume (see Figure 1). Computation of these parameters requires evaluation of complicated algorithms. For this reason, any effort to development an effective analytical methods is always welcome.

Figure 1: A molecular surface

Lee and Richards [4], and Chothia [5], introduced a solvent-accessible surface. The accessible surface is traced out by the probe sphere center as it rolls over the protein. It is a kind of expanded van der Waals surface. The molecular surface was first computed by Greer and Bush [8], Richmond [10] has defined the solvent-excluded volume to mean the volume contained within the solvent accessible surface, i.e. the volume which is inaccessible to the centers of solvent particles. That is the union of the expanded atom spheres. The excluded volume is an important quantity in the theory of gases and liquids [9]. The exploration of molecular volume and surface is essential for the understanding of drug action since short range dispersion forces play a major role in the binding of drug molecules to receptors (http://server.ccl.net/cca/documents/molecular-modeling).

The problem of the computation of volume and the surface area of the union of overlapping spheres has been the subject of methods both numerical (Rowlinson [3]; Pavani and
Figure 2: Stereographic projection of the spherical surface points onto the tangential plane


In [1] a new analytical method for computing solvent-accessible surface area of macromolecules and its gradients is presented. New approach, based on the stereographic projection (see Figure 1) by which the surface integrals are transformed to a sum of double integrals which are reduced to the curve integrals. This approach was further extended for the calculation of excluded volume, too. In [2] a Fortran package based on the new exact analytical methods for computing volume and surface area of overlapping spheres is presented. MPI Fortran version is described as well there. The package is also useful for computing the order parameter of continuum percolation models.

Analytical methods for computing the excluded volume and the accessible surface area of a macromolecule

We describe the molecule $M$ as a union of $n$ spheres (atoms) $S_1, \ldots, S_n$, i.e. $M = \bigcup_{j=1}^{n} S_j$. Let $(x_i, y_i, z_i)$ be Cartesian coordinates of the center of the $i$-th sphere and $r_i$ be its radius, where $1 \leq i \leq n$. For $j \neq i$ we say that $S_j$ is a neighbor of $S_i$ if $\text{In}(S_i) \cap \text{In}(S_j) \neq \emptyset$, where $\text{In}(S)$ denotes the interior of the set $S$. 

239
Stereographic projection

The points \((x, y, z)\) on the surface of \(i\)-th sphere satisfy the equation

\[
(x - x_i)^2 + (y - y_i)^2 + (z - z_i)^2 = r_i^2.
\]  

(1)

One can easily calculate from Fig. 2 that the point \((x_i, y_i, z_i)\) on the \(i\)th sphere is projected from the top point (NP-North Pole) of the sphere onto the point \((t, s)\) through the following equations

\[
t = -\frac{2r_i(x - x_i)}{z - z_i - r_i},
\]

\[
s = -\frac{2r_i(y - y_i)}{z - z_i - r_i}.
\]

(2)

This is one-to-one transformation except the top point \((x_i, y_i, z_i + r_i)\) itself. It follows from Eqs. (1) and (2) that the inverse transformation can be written as

\[
x = x_i + \frac{4r_i^2 t}{t^2 + s^2 + 4r_i^2},
\]

\[
y = y_i + \frac{4r_i^2 s}{t^2 + s^2 + 4r_i^2},
\]

\[
z = z_i + r_i - \frac{8r_i^3}{t^2 + s^2 + 4r_i^2}.
\]

(3)

The volume evaluation

Using Gauss-Ostrogradsky’s theorem one can reduce the volume \(V(\mathcal{M})\) evaluation to the surface integrals of the second kind

\[
V(\mathcal{M}) = \iiint_{V(\mathcal{M})} dx dy dz = \iint_{B(\mathcal{M})} z dx dy = \sum_{i=1}^{n} \iint_{B_i(\mathcal{M})} z dx dy,
\]

(4)

where \(B(M)\) is the boundary (surface) of \(M\) and \(B_i(M)\) is a part of the surface of \(S_i\) which is outside of all its neighbors (“free” surface of the sphere \(S_i\)).

Next step is the transformation of the surface integral over the particular surface \(B_i(M)\) into the double integral. This can be done by projecting the surface \(B_i(M)\) from some top point of the sphere (North Pole) into the domain \(\Omega_i\) in the plane tangent to the sphere at the diametrically opposite point (the South Pole of \(S_i\)) [1], see Figures 3 and 4.

For bounded \(\Omega_i\) we arrive to

\[
\mathcal{I}_i = \iint_{B_i(\mathcal{M})} z dx dy = \frac{128r_i^7}{3} \iint_{B(\Omega_i)} \frac{t ds - s dt}{(t^2 + s^2 + 4r_i^2)^3} + \frac{2r_i^3}{3} \iint_{B(\Omega_i)} \frac{t ds - s dt}{(t^2 + s^2 + 4r_i^2)} - \frac{8r_i^4(3z_i + 2r_i)}{3} \iint_{B(\Omega_i)} \frac{t ds - s dt}{(t^2 + s^2 + 4r_i^2)^2},
\]

(5)
where $B(\Omega_i)$ is the boundary of $\Omega_i$, which consists of the circular arcs (see Fig. 3). All line integrals along the circular arcs can be calculated exact.

Both, bounded and unbounded cases are represented by the formula

$$V = \sum_{i=1}^{n} [\chi_V(\Omega_i) + \mathcal{I}_i],$$

where $\mathcal{I}_i$ is defined above by Eq. (5), and

$$\chi_V(\Omega_i) = \begin{cases} 0, & \Omega_i \text{ is bounded}, \\ \frac{4}{3}\pi r_i^3, & \Omega_i \text{ is all plane except} \\ & \text{the union of several disks}, \end{cases}$$

Remark. If $\Omega_i$ is unbounded, the corresponding sum of integrals $\mathcal{I}_i$ will be negative and we get a correct value for the integral along $B(\Omega_i)$ given by Eq. (5).

The surface area evaluation

For the area $A(\mathcal{M})$ we have

$$A(\mathcal{M}) = \iint_{B(\mathcal{M})} |d\sigma| = \sum_{i=1}^{n} \iint_{B_i(\mathcal{M})} |d\sigma|. \quad (7)$$

Like in Eq. (5) if we denote

$$\mathcal{J}_i = \oint_{B(\Omega_i)} \frac{t \, ds - s \, dt}{t^2 + s^2 + 4r_i^2}, \quad (8)$$

then the general formula for surface area is similar to the formula (6) for the volume

$$A = \sum_{i=1}^{n} [\chi_A(\Omega_i) + 2r_i^2 \mathcal{J}_i], \quad (9)$$
where $\mathcal{J}_i$ is defined above by Eq. (8), and

$$\chi_A(\Omega_i) = \begin{cases} 
0, & \Omega_i \text{ is bounded}, \\
4\pi r_i^2, & \Omega_i \text{ is all plane except the union of several disks}.
\end{cases}$$

**Remark.** The integrals in formula (9) are calculated in the volume computation. So we can use about the same computing time to get both the volume and the area values.

The existence of cavities (see Figure 5) should be taken into account by solving different problems connected to the molecular properties. A special triangulation which includes inside possible cavities points and the algorithm of the construction of such triangulation are developed by the research team. This triangulation may be used for making a decision about the point location and for calculation of the solvent accessible surface area and inaccessible volume of a molecule in the case of possible molecular cavities existence.

This work was supported in part by the Russin Foundation for Basic Research (N 04-01-00490, N 05-01-00645).

**Publications:**


**References**


