Investigation of the Three-Dimensional Macromolecular Structures

Ján Buša[a,b,c], Jozef Džurina[a,b], Edik Hayryan[a,b,c], Ján Plavka[a,b], Imrich Pokorný[a,b], Jaroslav Skřivánek[a,b,c], Shura Hayryan[c], Chin-Kun Hu[c], Ming-Chya Wu[c]

[a] *Technical University in Košice, Slovak Republic*

[b] *Laboratory of Information Technologies, JINR*

[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.

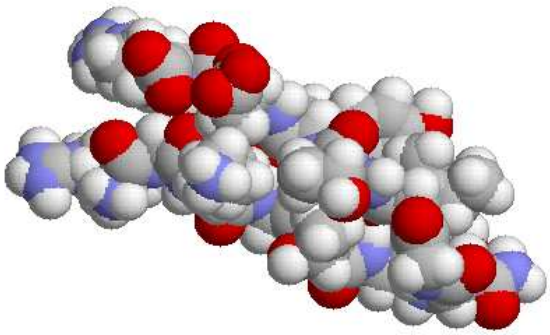

Figure 1: A molecular surface

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

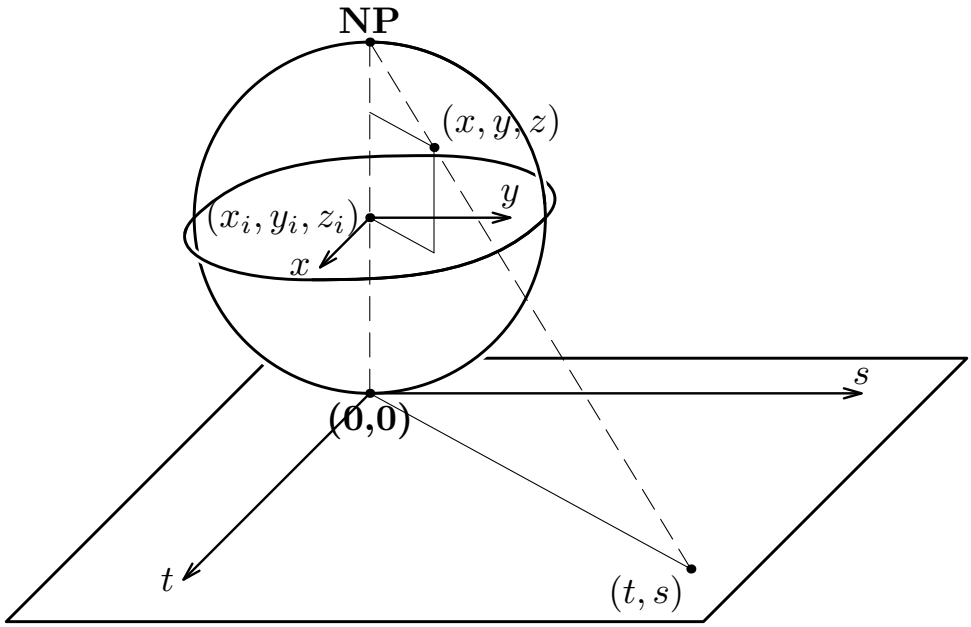

Figure 2: Stereographic projection of the spherical surface points onto the tangential plane

Ranghino [11]; Gavezzotti [12]) and analytic (Richmond [10]; Kang, Nemethy, and Scheraga [13]; Gibson and Scheraga [14, 15, 16]; Guerrero-Ruiz, Ocadiz-Ramirez, and Garduno-Juarez [17]; Petitjean [18]). More information can be found at `www.netsci.org/Science/Compchem/feature14.html`, where an excellent overview, written by author of [6, 7] M. L. Conolly, is presented. The E. Silla's, at al. package GEPOL [20, 19] for computing the molecular area and volume is referred there. At `http://www.biohedron.com` one can find the M. L. Conolly's molecular surface package presentation.

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 macromolecula

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 $\mathrm{In}\,(S_i) \cap \mathrm{In}\,(S_j) \neq \emptyset$, where $\mathrm{In}\,(S)$ denotes the interior of the set S.

Stereographic projection

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

$$(x - x_i)^2 + (y - y_i)^2 + (z - z_i)^2 = r_i^2. \tag{1}$$

One can easily calculate from Fig. 2 that the point (x_i, y_i, z_i) on the ith sphere is projetcted from the top point (NP-North Pole) of the sphere onto the point $(t, s) \in \mathbb{R}^2$ through the following equations

$$\begin{aligned} t &= -\frac{2r_i(x - x_i)}{z - z_i - r_i} \\ s &= -\frac{2r_i(y - y_i)}{z - z_i - r_i} \end{aligned} \tag{2}$$

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

$$\begin{aligned} x &= x_i + \frac{4r_i^2 t}{t^2 + s^2 + 4r_i^2} \\ y &= y_i + \frac{4r_i^2 s}{t^2 + s^2 + 4r_i^2} \\ z &= z_i + r_i - \frac{8r_i^3}{t^2 + s^2 + 4r_i^2} \end{aligned} \tag{3}$$

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

$$V(\mathcal{M}) = \iiint\limits_{V(\mathcal{M})} dx\, dy\, dz = \iint\limits_{B(\mathcal{M})} z\, dx\, dy = \sum_{i=1}^{n} \iint\limits_{B_i(\mathcal{M})} z\, dx\, dy, \tag{4}$$

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 Ω_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 Ω_i we arrive to

$$\begin{aligned} \mathcal{I}_i = \iint\limits_{B_i(\mathcal{M})} z\, dx\, dy = \frac{128r_i^7}{3} \oint\limits_{B(\Omega_i)} \frac{t\, ds - s\, dt}{(t^2 + s^2 + 4r_i^2)^3} + \\ \frac{2r_i^3}{3} \oint\limits_{B(\Omega_i)} \frac{t\, ds - s\, dt}{(t^2 + s^2 + 4r_i^2)} - \\ \frac{8r_i^4(3z_i + 2r_i)}{3} \oint\limits_{B(\Omega_i)} \frac{t\, ds - s\, dt}{(t^2 + s^2 + 4r_i^2)^2}, \end{aligned} \tag{5}$$

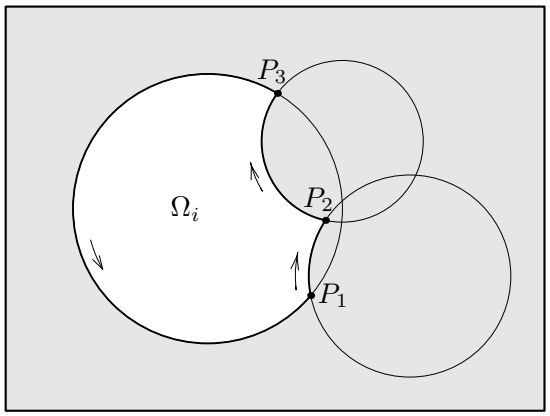

Figure 3: The bounded image Ω_i of the surface part $B_i(\mathcal{M})$

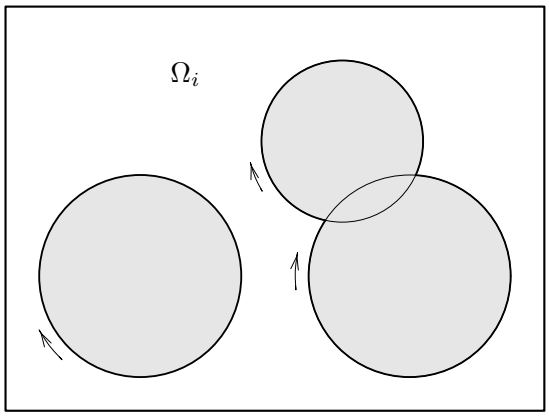

Figure 4: Unbounded domain Ω_i

where $B(\Omega_i)$ is the boundary of Ω_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

$$V = \sum_{i=1}^{n} \left[\chi_V(\Omega_i) + \mathcal{I}_i\right], \tag{6}$$

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

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

Remark. If Ω_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

$$A(\mathcal{M}) = \iint\limits_{B(\mathcal{M})} |d\boldsymbol{\sigma}| = \sum_{i=1}^{n} \iint\limits_{B_i(\mathcal{M})} |d\boldsymbol{\sigma}|. \tag{7}$$

Like in Eq. (5) if we denote

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

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

$$A = \sum_{i=1}^{n} \left[\chi_A(\Omega_i) + 2r_i^2 \mathcal{J}_i\right], \tag{9}$$

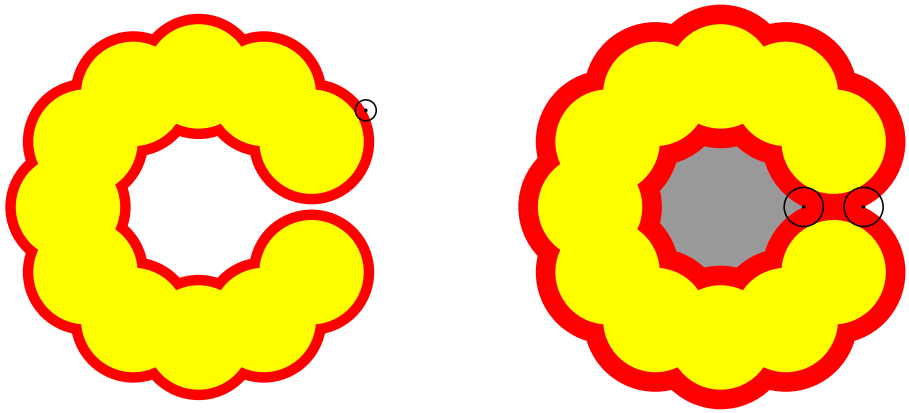

Figure 5: The molecule without (left) and with (right) a cavity

where $\mathcal{J}_i$ is defined above by Eq. (8), and

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

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:

[1] Sh. Hayryan, C.-K. Hu, J. Skřivánek, E. Hayryan, I. Pokorný: *A new analytical method for computing solvent-accessible surface area of macromolecules and its gradients*, Journal of Computational Chemistry, vol. 26, no. 4 (2005), 334–343.

[2] J. Buša, J. Džurina, E. Hayryan, Sh. Hayryan, Ch.-K. Hu, J. Plavka, I. Pokorný, J. Skřivánek, M.-Ch. Wu: *ARVO: A Fortran package for computing the solvent accessible surface area and the excluded volume of overlapping spheres via analytic equations*, Computer Physics Communications, vol. 165 (2005), 59–96.

References

[3] J. S. Rowlinson: *The triplet distribution function in a fluid of hard spheres,* Molec. Phys. 6, 517–524 (1963).

[4] B. Lee and F. M. Richards: *The interpretation of protein structures: Estimation of static accessibility,* J. Mol. Biol. 55, 379–400 (1971).

[5] C. Chothia: *Hydrophobic bonding and accessible surface area in proteins,* Nature 248, 338–339 (1974).

[6] M. L. Connolly: *Analytical molecular surface calculation,* J. Appl. Cryst. **16**, pp. 548–558 (1976).

[7] M. L. Connolly: *Computation of molecular volume,* J. Am. Chem. Soc. 107, 1118–1124 (1985b).

[8] J. Greer and B. Bush: *Macromolecular shape and surface maps by solvent exclusion,* Proc. Natl. Acad. Sci. USA 75, 303–307 (1978).

[9] T. L. Hill: *Statistical Mechanics,* McGraw-Hill, New York, pp. 122–285.

[10] T. J. Richmond: *Solvent accessible surface area and excluded volume in proteins,* Journal of Molecular Biology, No. 178, pp. 63–89 (1984).

[11] R. Pavani and G. Ranghino: *A method to compute the volume of a molecule,* Computers and Chemistry 6, 133–135 (1982).

[12] A. Gavezzotti: *The calculation of molecular volumes and the use of volume analysis in the investigation of structured media and of solid-state organic reactivity,* J. Am. Chem. Soc. 105, 5220–5225 (1983).

[13] Y. K. Kang, G. Nemethy, and H. A. Scheraga: *Free energy of hydration of solute molecules. 1. Improvement of the hydration shell model by exact computations of overlapping volumes,* J. Phys. Chem. 91, 4105–4109 (1987).

[14] K. D. Gibson and H. A. Scheraga: *Exact calculation of the volume and surface area of fused hard-sphere molecules with unequal atomic radii,* Molec. Phys. 62, 1247–1265 (1987).

[15] K. D. Gibson and H. A. Scheraga: *Volume of the intersection of three spheres of unequal size. A simplified formula,* J. Phys. Chem. 91, 4121–4122 (1987).

[16] K. D. Gibson and H. A. Scheraga: *Surface area of the intersection of three spheres with unequal radii, a simplified analytical formula,* Molec. Phys. 64, 641–644 (1988).

[17] G. Guerrero-Ruiz, A. Ocadiz-Ramirez, and R. Garduno-Juarez: *ESFERA: A program for exact calculations of the volume and surface area of fused hard-sphere molecules with unequal atomic radii,* Computers Chem. 15, 351–352 (1991).

[18] M. Petitjean: *On the analytical calculation of van der Waals surfaces and volumes: Some numerical aspects,* J. Comp. Chem. 15, 507–523 (1994).

[19] E. Silla, F. Villar, O. Nilsson, J. L. Pascual-Ahuir, and O. Tapia *Molecular volumes and surfaces of biomacromolecules via GEPOL: A fast and efficient algorithm,* J. Mol. Graphics 8, 168–172 (1990).

[20] E. Silla, I. Tun, and J. L. Pascual-Ahuir *GEPOL: An improved description of molecular surfaces. II. Computing the molecular area and volume,* J. Comp. Chem. 12, 1077–1088 (1991).