УДК 539.211

ROUGHNESS OF THE GLOBULAR PROTEIN SURFACE

A.A.Timchenko*, O.V.Galzitskaya*, I.N.Serdyuk*

Protein surface analysis using high resolution X ray shows that this surface has a two-level organization, on the micro- and macro-scales. On the micro-scale (2–7 Å), the surface is characterized by the $D = 2.1$ fractal dimension which is intrinsic to surface with weak deformation and reflects the local atomic group packing. On the macro-scale the large scale surface defects are revealed which are interpreted as the result of secondary structure elements packing.

The investigation has been performed at the Frank Laboratory of Neutron Physics, JINR and the Institute of Protein Research, RAS.

Неровность глобулярной поверхности протеина

Тимченко А.А. и др.

Проведено изучение структуры поверхности протеинов методом рассеяния рентгеновских лучей высокого разрешения. Показано, что эта поверхность имеет двухуровневую организацию. На микроуровне (2–7 Å) поверхность характеризуется $D = 2,1$ размерной фрактальной структурой. На макроуровне выявлены дефекты поверхности большого масштаба.

Работа выполнена в Лаборатории нейтронной физики им. И.М.Франка ОИЯИ и Институте белка РАН.

The technique of protein surface quantitative analysis using high resolution X-ray data was first proposed in [1], where the accessible surface area (A_s) was analysed. Such an analysis gives a possibility to study overall properties of the protein surface and its detailed structure. The deviation of the power law extent from 2/3 in the $A_s - M$ dependence was considered as an indication of the protein surface fractal structure [2,3]. Strictly speaking, a surface is fractal if the dependence of the number of probe bodies (balls, cubes, etc.) fully covering the surface on the probe size is a power law:

$$N(r) = \text{const}^* \, r^{-D} \tag{1}$$

with the extent $2 < D < 3$ not coinciding with the topological dimension ($D_{\text{top}} = 2$) and D being a fractal dimension [4]. A strict fractal dimension is determined at $r \to 0$. For self-

*Institute of Protein Research, Russian Academy of Sciences, Pushchino

similar bodies, the relationship between the fractal surface area and the value of confined volume (V) has the following power law [5]:

$$A_m = \text{const}^* V^{D/3}. \quad (2)$$

Qualitatively at $D > 2$ this means that the size of irregularities increases with the increase of the particle size. Calculation of D from X-ray scattering patterns of some proteins gave a broad range of D from 2 to 2.8. No systematic analysis of protein surface fractality was made. It has been shown [6] that the asymmetry extent of globular proteins does not change with increasing their sizes, and the observed non-trivial dependence of the protein accessible surface area on the molecular mass ($A_s - M$ dependence) is a reflection of the protein surface relief peculiarities. To clarify these peculiarities, an analysis of the molecular surface on the basis of high resolution X-ray data has been done for 25 globular proteins not containing prosthetic groups (Table). The procedure was based on studying the dependence of the minimal number (N) of probe bodies (here cubes) covering the entire protein surface, both on their size ($N - R$ dependence) and on the value of dry protein volume ($N - V$ dependence). The molecular and accessible surfaces of the protein were determined by rolling a ball (approximating a water molecule) of 2.8 Å in diameter along the protein surface [7]. Both surfaces consisted of an array of cubes with a 0.3 Å edge.

Fractal Properties of the Protein Surface

Relationship (2) is valid if the whole particle surface is fractal. At the same time, relationship (1) is applicable for analysis of any surface, but the value D in (1) depends on the range of probe body sizes. In other words, expression (1) indicates the levels of different surface organizations and scans it with different resolution (from the atom size to the overall dimension of a macromolecule). As an example, Figure 1 shows the log-log dependence of the number of cubes on their size (the $N - R$ dependence) for two proteins with a five-fold difference in their molecular mass and a sphere 24 Å in diameter tightly packed with balls. The course of the dependence is the same for other studied proteins and spheres of different sizes. One can see the nonlinear character of the dependence mainly due to the initial points. The D value calculated from the first points appeared to be

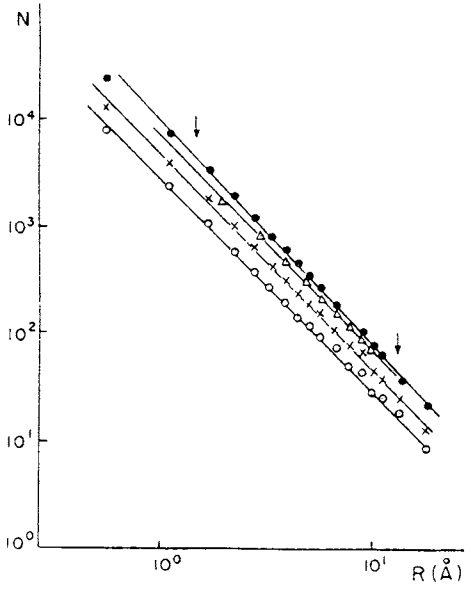

Fig.1. The dependence of the number of probe cubes fully covering a molecular surface on their size for proteins 1FDX (x), 4CHA (•), sphere 24 Å in diameter tightly packed with balls (o). The same dependence for the surface constructed of balls 12 Å in diameter centered in C_α atoms of 1LZT (△). The range for the straight line fitting is marked by arrows and the parameters of fitting are given in the Table

less than the topological dimensionality $D_{top} = 2$. The explanation of this behavior is given in [8].

This region corresponds to the probe cube sizes 2–10 Å and, to avoid ambiguity, it is the same for both proteins and spheres. The parameters of the obtained dependences and their errors for all studied proteins and spheres are given in the Table. A conclusion is drawn that the protein surfaces are similar and show a «graphite-like» [9] fractality on the 2–10 Å scale. The main reason for the observed fractality is the packing of atomic groups.

Table. The dependence of the number of cubes (N) surrounding the molecular surface on their edge size (R) for proteins of molecular mass (M) and spheres of different diameters filled with balls with the radius 1.6 Å

Protein	PDB index	Mol. mass (M)	$N - R$ dependence
Crambin	1CRN	4710	$4.15 \cdot 10^4 \cdot \mathbf{R}^{-2.031 \pm 0.009}$
Ferredoxin, *Paerogenes*	1FDX	5400	$5.72 \cdot 10^4 \cdot \mathbf{R}^{-2.113 \pm 0.017}$
Pancreatic trypsin inhibitor, bovine	4PTI	6490	$5.55 \cdot 10^4 \cdot \mathbf{R}^{-2.028 \pm 0.014}$
Neurotoxin B, sea snake	1NXB	6840	$5.90 \cdot 10^4 \cdot \mathbf{R}^{-2.048 \pm 0.019}$
Neurotoxin 3, scorpion	1SN3	7050	$6.64 \cdot 10^4 \cdot \mathbf{R}^{-2.085 \pm 0.024}$
Intestinal calcium-binding protein	1ICB	8470	$7.53 \cdot 10^4 \cdot \mathbf{R}^{-2.087 \pm 0.017}$
High potential iron protein	1HIP	8880	$7.18 \cdot 10^4 \cdot \mathbf{R}^{-2.052 \pm 0.021}$
Plastocyanin	1PCY	10450	$7.85 \cdot 10^4 \cdot \mathbf{R}^{-2.031 \pm 0.013}$
Parvalbumin, carp	3CPV	11400	$1.11 \cdot 10^5 \cdot \mathbf{R}^{-2.127 \pm 0.011}$
Ribonuclease A	1RN3	13670	$1.17 \cdot 10^5 \cdot \mathbf{R}^{-2.097 \pm 0.011}$
Azurin, *Alcaligenes denitrificans*	2AZA	13950	$1.10 \cdot 10^5 \cdot \mathbf{R}^{-2.098 \pm 0.014}$
Lysozyme, hen egg white	1LZT	14280	$1.11 \cdot 10^5 \cdot \mathbf{R}^{-2.079 \pm 0.017}$
Lysozyme, human	1LZ1	14670	$1.21 \cdot 10^5 \cdot \mathbf{R}^{-2.110 \pm 0.015}$
Nuclease, *Staphylococcus aureus*	2SNS	15980	$1.45 \cdot 10^5 \cdot \mathbf{R}^{-2.112 \pm 0.016}$
Dihydrofolate reductase, *E.coli*	4DFR	17960	$1.56 \cdot 10^5 \cdot \mathbf{R}^{-2.135 \pm 0.018}$
Lysozyme, bacteriophage T4	2LZM	18610	$1.64 \cdot 10^5 \cdot \mathbf{R}^{-2.129 \pm 0.017}$
β-Trypsin	1TPO	23190	$1.73 \cdot 10^5 \cdot \mathbf{R}^{-2.106 \pm 0.015}$
Actinidin	2ACT	23380	$1.67 \cdot 10^5 \cdot \mathbf{R}^{-2.009 \pm 0.016}$
Elastase	3EST	24840	$1.97 \cdot 10^5 \cdot \mathbf{R}^{-2.115 \pm 0.013}$
Chymotripsin	4CHA	25000	$1.97 \cdot 10^5 \cdot \mathbf{R}^{-2.118 \pm 0.014}$
Chymotripsinogen A	2CGA	25620	$2.20 \cdot 10^5 \cdot \mathbf{R}^{-2.133 \pm 0.014}$

Protein	PDB index	Mol. mass (M)	$N-R$ dependence
Subtilisin BPN'	1SBT	27480	$1.95 \cdot 10^5 \cdot \mathbf{R}^{-2.113 \pm 0.015}$
Carbonic anhydrase B	2CAB	28350	$1.98 \cdot 10^5 \cdot \mathbf{R}^{-2.095 \pm 0.009}$
Carbonic anhydrase form C	1CAC	28720	$2.84 \cdot 10^5 \cdot \mathbf{R}^{-2.183 \pm 0.018}$
Pepsin, *Penicillium*	2APP	33380	$2.38 \cdot 10^5 \cdot \mathbf{R}^{-2.113 \pm 0.018}$
sphere ($D = 22$ Å)			$2.80 \cdot 10^4 \cdot \mathbf{R}^{-1.960 \pm 0.016}$
sphere ($D = 24$ Å)			$3.87 \cdot 10^4 \cdot \mathbf{R}^{-1.987 \pm 0.030}$
sphere ($D = 26$ Å)			$4.24 \cdot 10^4 \cdot \mathbf{R}^{-2.010 \pm 0.023}$
sphere ($D = 28$ Å)			$5.69 \cdot 10^4 \cdot \mathbf{R}^{-2.027 \pm 0.033}$
sphere ($D = 30$ Å)			$5.11 \cdot 10^4 \cdot \mathbf{R}^{-1.948 \pm 0.022}$

The $N - V$ dependence of the number of probe cubes surrounding the molecular surface on the dry 5 volume value was analysed for each fixed probe size. The slope of the log-log $N - V$ dependence is, on the average, 0.76 for the probe cube size less than 7 Å. This value coincides with an analogous one for the log-log $A - V$ dependence. At larger probe cube sizes a smooth transfer to the 0.69 slope takes place (see Fig. 2 where the fractal dimension D is given according to Eq.(2)). This value is close to 0.67 which is specific for even self-similar bodies. An analogous dependence for spheres of different sizes tightly packed with balls has the 0.67 slope. Distinct dependence of this slope on the probe cube size is not observed. According to (2), the 0.76 slope for smaller probe cube sizes corresponds to the fractal dimension $D = 0.76^*3 = 2.28$, and the fractal dimension from (1) is on the average 2.10 (see the Table). Non-coincidence of these values can be evidence of some regular peculiarities of the protein surface. These peculiarities virtually do not affect the molecule asymmetry [6], but essentially contribute to the surface area. It is noteworthy that the fractal dimension from (2) is greater than that from (1). The higher value of the fractal dimension from the $A - V$ dependence can be interpreted as an increase in the number of large-scale irregularities on the protein surface following an increase in the protein size [8].

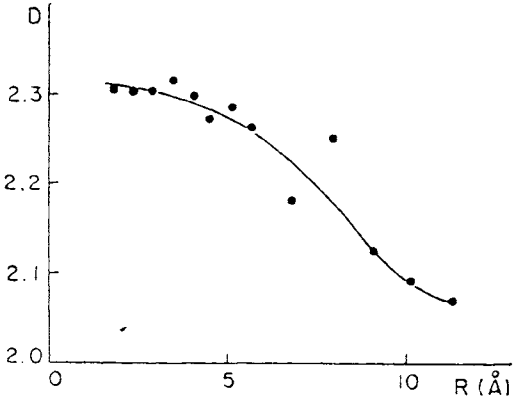

Fig.2. The fractal dimension D calculated from the $N - V$ dependence versus the probe cube size (R)

The outlined interpretation of the protein surface structure shows that this surface has a two-level organization, on the micro- and macro-scales. On the micro-scale (2–7 Å), the surface is characterized by the $D = 2.1$ fractal dimension which is intrinsic to surfaces with weak deformation and reflects the local atomic group packing. On the macro-scale the large-scale surface defects are revealed which are interpreted as the result of secondary structure elements packing.

References

1. Lee B., Richards F.M. — J. Mol. Biol., 1971, v.55, p.379.
2. Fushman D. — Biomol. Struc. Dynamics, 1990, N.7(6), p.1333.
3. Fedorov B.A., Fedorov B.B., Shmidt P.W. — J. Chem. Phys., 1993, v.99(5), p.4076.
4. Pfeifer P., Obert M. — The Fractal Approach to Heterogeneous Chemistry. Avnir D. (ed.). Wiley & Sons Ltd., 1989.
5. Pfeifer P., Avnir D. — J. Chem. Phys., 1983, v.79, p.3558 [Erratum: J. Chem. Phys., 1984, v.80, p.4573.]
6. Serdyuk I.N., Galzitskaya O.V., Timchenko A.A. — Biophizika, 1997, v.42, p.1197.
7. Pavlov M.Yu., Fedorov B.A. — Biopolymers, 1983, v.22, p.1507.
8. Timchenko A.A., Galzitskaya O.V., Serdyuk I.N. — Proteins, 1997, v.28, p.194.
9. Avnir D., Farin D., Pfeifer P. — Nature, 1984, v.308, p.261.

Received on July 28, 1998.