Available online at www.sciencedirect.com
SCIENCE DIRECT®
Physica A 361 (2006) 250–254
PHYSICA A
www.elsevier.com/locate/physa

Protein chain packing and percolation threshold

M.A. Moret[a,b,*], M.C. Santana[b], E. Nogueira Jr.[c], G.F. Zebende[a,b]

[a]*CEPPEV-Fundação Visconde de Cairu, 40226-900, Salvador, Bahia, Brazil*
[b]*Departamento de Física, Universidade Estadual de Feira de Santana, 44031-460, Feira de Santana, Bahia, Brazil*
[c]*Instituto de de Física, Universidade Federal da Bahia, 40210-340 Salvador, Bahia, Brazil*

Received 26 May 2005
Available online 28 September 2005

Abstract

The major factor that drives a protein chain toward collapse and folding is the hydrophobic effect. Amino acids with apolar side chains join forming a solvent-shielded hydrophobic core. This process pack the structure in the native one. Here we investigate the average packing density of 5526 protein chains deposited in the Brookhaven Protein Data Bank. This analysis is carried out from the scaling analysis of the mass-size exponent and it shows that the exponent is $\delta = 2.47$. This fractal dimension of the protein chain is close to the one obtained in the randomly packed spheres near their percolation threshold. The present findings supply a measure of the protein compactness, that tends to a constant protein chain packing. The average packing density tends to $\rho = 0.86\,\text{a.u.}/\text{Å}^3$.

Keywords: Protein folding; Self-similarity; Average packing density

It is well known that macromolecular systems have a great number of stable conformations in the energy hyper-surface [1–9] that increases with the number of degrees of freedom in the molecular system. The biological activity depends on the spatial conformation acquired by the macromolecules in the physiological medium.

*Corresponding author.
E-mail address: moret@cairu.br (M.A. Moret).

0378-4371/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.physa.2005.08.001

The action of hormones and drugs is also dependent on the molecular three-dimensional structure of the target molecules. In recent years, the atomic description of biological molecules in computational simulations have promoted significant advances in the comprehension of the biological process as well as proposed new insights in design of molecules to satisfy specific properties.

Protein folding is driven by hydrophobic forces [10]. In this process, apolar amino acids associate to form a hydrophobic core, packing the protein in the folded state. Analyzing globular proteins, Lattman and Rose concluded that the native fold determines the packing but packing does not determine the native fold [11]. This view is corroborated by the widespread occurrence of protein families whose members assume the same fold without having sequence similarity. Evidently, there are a myriad number of ways in which the internal residues can pack together efficiently. Packing in proteins was first studied by using a Voronoi analysis for proteins in a space-filling model [12], where each atom is taken to be a sphere with a fixed radius given by the van der Waals one. As consequence of this analysis, the average packing density in proteins is as high as that inside crystalline solids [10].

The average packing density has been studied by the most varied approaches as, for instance, by a method based on Delaunay tesselation [13], a coarse-grained scale [14], among other approaches. The first approach to analyze the protein packing is based on Voronoi tesselation methods that have been widely used for examining packing, volume and surface area [15,16]. A major difficulty in this approach is the assignment of the Voronoi polyhedra to surface residues, which may necessitate including or modelling solvent molecules. An elegant way to avoid ambiguity in the surface residue analysis is based on Delaunay triangulation, where for each Voronoi edge, connect the corresponding two atom centers with a line segment, and for each Voronoi vertex, place a triangle spanning the three atom centers of the three Voronoi cells. Completing this, for all Voronoi edges and Voronoi vertices gives a collection of line segments and triangles. Together with the vertices representing atom centers, they form the Delaunay complex, which is the underline structure of Delaunay triangulation. This approach allows computing the packing densities, that the mean interior packing density is 0.74 (van der Waals volume per total volume). The Delaunay tesselation allows to calculate the relation between the molecular surface volume and the average radius of the protein and it supplies the following fractal dimension $\delta = 2.47$ [13].

Another interesting approach to evaluate the packing of residues is based on a coarse-grained scale. In this method a Monte Carlo algorithm is utilized for superimposing residue clusters collected from known protein structures. A residue cluster is composed of a central residue, and the set of all neighboring residues located within a first coordinate shell. A constant radius is used for defining the first coordinate volume [14]. From this approach is possible approximates two-third of the protein packing as a fcc geometry on a coarse-grained scale. The remaining one-third refer to residues which are more loosely or randomly packed.

In this paper we show a very simple methodology to measure the average packing density that does not present ambiguities, because it does not depend on stochastic optimization methods or it does not depend of probe atoms to evaluate voids in the

proteins. Our strategy amounts to measure the average radius of each protein chain as a function of the mass of this molecule. This approach allows to measure the high compact of those molecules. We describe this compactness via the mass-size exponent (fractal dimension) [17], i.e., the relation between the average radius and the mass of protein chains, as following

$$M_i^j \propto \langle R_i^j \rangle^{\delta}, \tag{1}$$

where M_i^j is the mass of the jth protein chain belonging of the ith protein and $\langle R_i^j \rangle$ is the correspondent average radius (the average distance from the geometric center for all coordinates).

We propose to measure the average packing density from the relation

$$\rho = \frac{\Delta M_i^j}{\Delta V_i^j}, \tag{2}$$

where V_i^j is the volume of the sphere with $\langle R_i^j \rangle$ as radius. Fig. 1 depicts the behavior of the packing density as a function of average radius of 5526 protein chains analyzed. From this figure we note that protein chains with $\geqslant 15$ residues ($\langle R_i^j \rangle \geqslant 4.0\,\text{Å}$) present a tendency to stabilize the average packing density to $\rho = 0.86\,\text{a.u.}/\text{Å}^3$. The histogram of the average packing density behaves as a normal distribution, as shown Fig. 2. Therefore this study shows that the packing of the protein chains behaves as an uniform packing density. We recall that these protein chains were deposited in the Brookhaven Protein Data Bank. Furthermore, the compact folded structure presents an elegant fractal behavior, that is depicted by the average radius as function of the mass of protein chains [18]. The fractal dimension

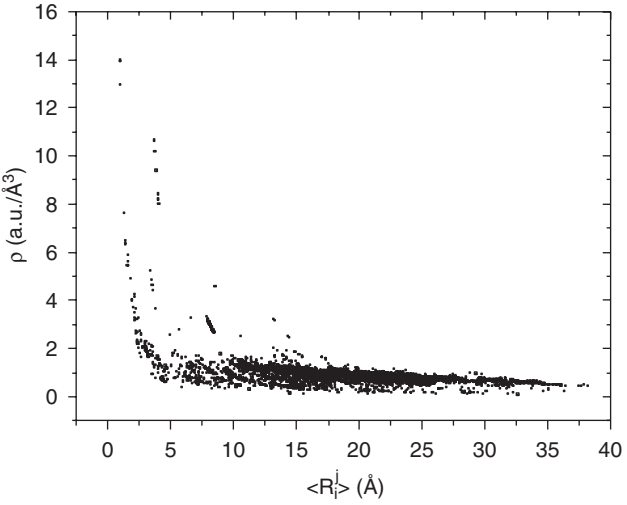

Fig. 1. The behavior of the average packing density of 5526 protein chains in function of average radius. Each chain with average radius $\geqslant 4.0\,\text{Å}$ tends to stabilizes the density ($\rho = 0.86\,\text{a.u.}/\text{Å}^3$, with $\sigma = 0.25\,\text{a.u.}/\text{Å}^3$ as standard deviation and $err = 0.003\,\text{a.u.}/\text{Å}^3$ as standard error).

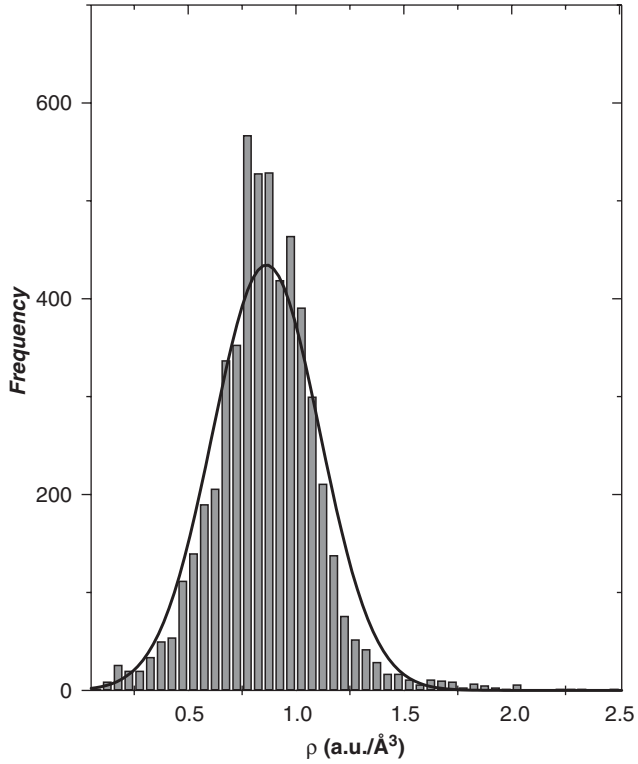

Fig. 2. The histogram of the average packing density of protein chains. This histogram obeys a normal distribution (continuous line), with $\langle\rho\rangle = 0.86\,\text{a.u.}/\text{Å}^3$, $\sigma = 0.25$ and $err = 0.003\,\text{a.u.}/\text{Å}^3$.

of the protein chains is $\delta = 2.47 \pm 0.03$, that is exactly the same dimension obtained by the analysis of the molecular surface volume [13]. Therefore, the atoms that compose the protein are distributed in the macromolecule volume as a fractal object. In this context, the interactions among these atoms are made among the most different scales of the object, taking with that to a multifractality in the energy hypersurface [4]. On the other hand, these interactions among atoms take the packing of the protein chain a constant.

From the percolation theory [19] it is possible to obtain the property of randomly packed spheres. At the percolation threshold the volume of a cluster of spheres scales with the length of the cluster with a characteristic exponent $\delta = 2.50$ [20,21]. In this sense, the fractal dimension obtained so much by the relation between the volume [13] and the radius as for the mass and radius [18], it supplies us a very close dimension value of the percolation threshold.

Concluding, the compactness of the protein chains turns the geometry of these molecules a fractal object. It is interesting to point out that independent of the origin of the molecule, its family, organism where it was expressed, each chain of a protein has the same fractal dimension [13,18]. Thus, analyzing the packing density of the

protein chains, we observed this density independent of the number of degrees of freedom, biological activity, primary structure, and other characteristics of the proteins. This landscape suggests us that the atoms that compose a protein are packed in a similar way as those random spheres in the percolation threshold and this scenario corroborates to the fact that only correct folded structures are able to packing efficiently [16]. Then, there is a large number of ways in which the atoms can pack together, but the physiological way behaves like spheres in their percolation threshold. In this sense, it is necessary to take into account not only the amino acid sequences, but also the packing density in order to obtain accurate predictions in protein folding studies.

Another interesting result that we point out is that the protein chains present a compactness reasonably uniform adjusted in a normal distribution. Unfortunately, this result cannot be extended for the proteins with more than one chain, because voids among chains can take to a decrease of the average packing density.

One of us wishes to thank P.M. Bisch for suggestions and very useful discussions (M.A. Moret). The partial financial support from Conselho Nacional de Desenvolvimento Científico e Tecnológico, Fundação de Amparo a Pesquisa do Estado da Bahia.

References

[1] J.M. Yon, Cell. Mol. Life Sci. 53 (1997) 557.
[2] M.A. Moret, P.G. Pascutti, P.M. Bisch, K.C. Mundim, J. Comput. Chem. 19 (1998) 647.
[3] D.J. Wales, H.A. Scheraga, Science 285 (1999) 1368.
[4] M.A. Moret, P.G. Pascutti, K.C. Mundim, P.M. Bisch, E. Nogueira Jr., Phys. Rev. E 63 (2001) 020901(R).
[5] M.A. Moret, P.M. Bisch, K.C. Mundim, P.G. Pascutti, Biophys. J. 82 (2002) 1123.
[6] K.A. Dill, Biochemistry 29 (1990) 7133.
[7] P.G. Wolynes, J.N. Onuchic, D. Thirumalai, Science 267 (1995) 1619.
[8] Y. Zhou, M.J. Karplus, Mol. Biol. 293 (1999) 917.
[9] Y. Zhou, M. Karplus, Nature 401 (1999) 400.
[10] D. Voet, J. Voet, Biochemistry, second ed., Wiley, New York, 1995.
[11] E.E. Lattman, G.D. Rose, Proc. Nat. Acad. Sci. 90 (1993) 439.
[12] F. Richards, J. Mol. Biol. 82 (1974) 1.
[13] J. Liang, K. Dill, Biophys. J. 81 (2001) 751.
[14] Z. Bagci, R.L. Jernigan, I. Bahar, J. Chem. Phys. 116 (2002) 2269.
[15] J. Finney, J. Mol. Biol. 96 (1975) 721.
[16] B. Honig, J. Mol. Biol. 293 (1999) 283.
[17] M.A.F. Gomes, J. Phys. A 20 (1987) L283.
[18] M.A. Moret, J.G.V. Miranda, E. Nogueira Jr., M.C. Santana, G.F. Zebende, Phys. Rev. E 71 (2005) 012901.
[19] D. Stauffer, Phys. Rep. 54 (1979) 1.
[20] D. Stauffer, Introduction to Percolation Theory, Taylor & Francis, London, 1985.
[21] T. Vicsek, Fractal Growth Phenomena, second ed., World Scientific, Singapore, 1992.