ISSN: 2348 - 2117

International Journal of Engineering Technology and Computer Research (IJETCR)
Available Online at www.ijetcr.org
Volume 4; Issue 1; January-February-2016; Page No. 24-30

Study of Selfsimilarity in Biomolecules

Kapsae S.K [1], Yusuf H Shaikh[2]*, Wasim Ahmed Hyderi [3], Gulam Rabbani[4]

[1] S.B. Science College Aurangabad

[2] Shivaji Arts, Commerce and Science College Kannad.

[3] Maharashtra College Mumbai

[4] Dr. Rafiq Zakaria Campus, Dr. Rafiq Zakaria marg, Rauza Bagh, Aurangabad.

Abstract

Biomolecules have complex shapes and quantification of complex random shapes poses difficulties as they cannot be modeled using Euclidian geometry. We used box counting technique and the concept of fractals and fractal geometry for the purpose of quantification of complexity of shape and structure associated with shape of large biomolecules using their project in two dimensions. For this purpose two colour bitmap of two dimensional projection of wire-frame of biomolecules is used and box counting technique is implemented for quantifying the complexity of shape and structure. It is demonstrated that the concept of fractals and fractal dimension can be used for quantification of complex structures like biomolecules or porous materials. Details are presented and findings discussed.

Keywords: Biomolecule, Self-similarity, Scale invariance, Fractal, Fractal Dimension

1. Introduction

The task of investigation of complex biomolecules is challenging as these molecules have more than tens of thousands of atoms and thus that many bonds and related complex structure. Large molecules like proteins and enzymes have an involved structure with associated complexity of shape and structure. Vijay Natrajan et al[1]. discussed the significance of study of shapes of biomolecules also it is relevant from the point of view of docking and other properties like hydrophobicity and hydrophilicity[2].

With the advancement in technology and availability of sophisticated instrumentation and new computational techniques, the most important area of life sciences i.e. the biotechnology has seen new horizons. Never before techniques have been evolved and deeper insight into phenomena could be realized that revolutionized the field of healthcare and medicine. Newer and potential biological techniques are made available in a variety of forms, this was a result of rapid research in the field of microbiology, bioinformatics etc. These days almost all new techniques rely, one way or the other on computers, microcontrollers or advanced electronics.

Fractal geometry [3 to 5] is a branch of science and technology that proved to be very useful in characterizing complex shapes and structures where other techniques are either insufficient or fail to yield reliable results. We tried to use the concept of fractals and fractal dimension [6, 7] to characterize the complex shapes of biomolecules using a simple approach employing the two dimensional projection of the shape of biomolecules.

2. Methodology:

For the purpose of characterization of complex shapes of biomolecules we used a technique of extracting a two dimensional projection of biomolecule. Standard biomolecules and their details of chains and sequences are available in the form of open source data bases at number of places[8–10]. These data bases of desired molecules can be downloaded and visualized in three dimensions using rendering software like arguslab, rasmol, raswin etc [11, 12]. These are open source freeware applications made available by institutions and agencies for public use. Online tools for rendering and some standard data processing are also available [13].

The 3-D structure of large molecules result in highly complex structures and the .PDB file for the molecules contains x, y and z coordinates of each atom. Reconstruction of the entire 3D structure of the molecule and implementation based on occupied

Corresponding author: Yusuf H Shaikh

cells is tedious task. The task is much simplified if the projection is obtained in 2D and the structure analyzed using box-counting, this can be used for comparison of relative complexity of molecules. However absolute quantifications would demand for 3D analysis of the molecule. The two dimensional projection of the biomolecule is obtained by rendering the molecule of interest in a suitable software and capturing the view of the molecule from certain angle. These images are stored in an image or picture file, preferably in bitmap file with .BMP extension. These images are then converted into two colour bitmap images that can be used by other software for further analysis. We developed programs for reading the two colour bitmap images and converting them into a matrix form that is representative of that image. Box counting algorithm is implemented on this matrix representing the biomolecule. The program used boxes of different sizes (r) and scans the whole image to find out the number of boxes required (N) to completely cover the image. This process is repeated for different selected sizes of boxes and the results are recorded into a text file in the form of table of r, N, log(r) and log (N). Most of the naturally occurring shapes and patterns exhibit Self-similarity and scale invariance [14–16], almost all the molecules exhibit self similarity and scale invariance over certain range of scale.

To establish scale invariance a graph is plotted using log (N) at the y-axis and log(r) on the axis of x. The presence of Self-similarity is ascertained from the resulting graph. If the power law holds and scale invariance is present, the resulting graph is a straight line. Slope of this straight line gives the power law exponent from which the fractal dimension can be found. The power law exponent or the fractal dimension is related to the degree of complexity of structure and texture associated with the pattern under study [17, 18]. A higher fractal dimension implies a higher level of complexity associated with the shape and a lower fractal dimension corresponds to a lower degree of complexity. Thus the fractal analysis of the pattern i.e. two dimensional projection of the image of a complex biomolecule[19 to 21] reveals the associated degree of complexity of shape quantitatively and from the fractal dimension or the power law exponent information related to the complexity of shape can be derived.

Rendering of three typical biomolecules used i.e. 3BA8, 4HK4 and 3RIR are shown in the form of ball and stick in figure 1, 2 and 3.

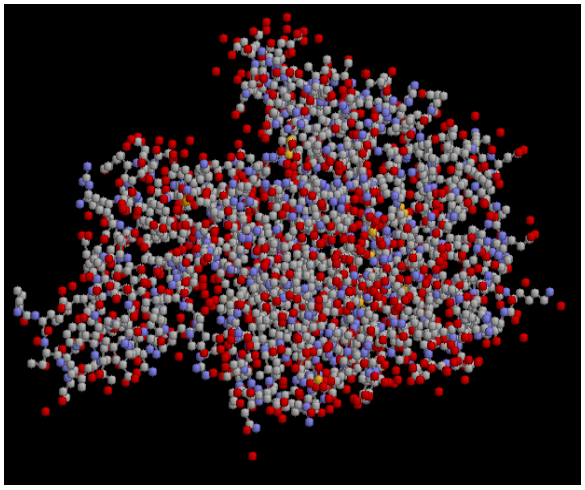


Figure 1: Enzyme 3BA8 that is a relatively large molecule and is displayed in Ball and Stick form.

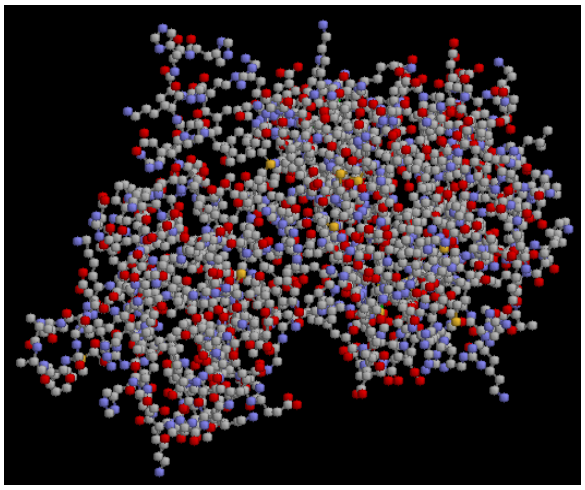


Figure 2: Enzyme 4HK4 that is a relatively small molecule and is displayed as Ball and Stick.

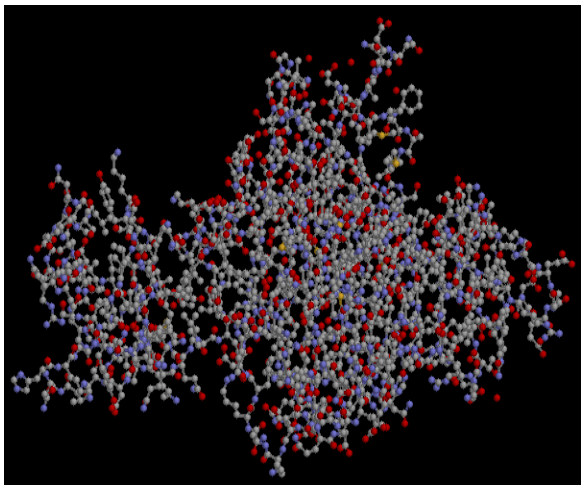


Figure 3: Enzyme 3RIR that is a relatively large molecule and is displayed as Ball and Stick.

Comparison of the three molecules in Fig. 1 to 3 clearly indicate that the structure of the molecules is sufficiently complicated and the molecule in Fig. 1 and 3 are relatively bigger in size in terms of physical size and number of atoms contained in, this is also evident from sizes of balls displayed. To have a feel of the complexity of structure and shape of the molecules the three molecules rendered in the form of wire frame are shown in Fig. 4, 5 and 6.

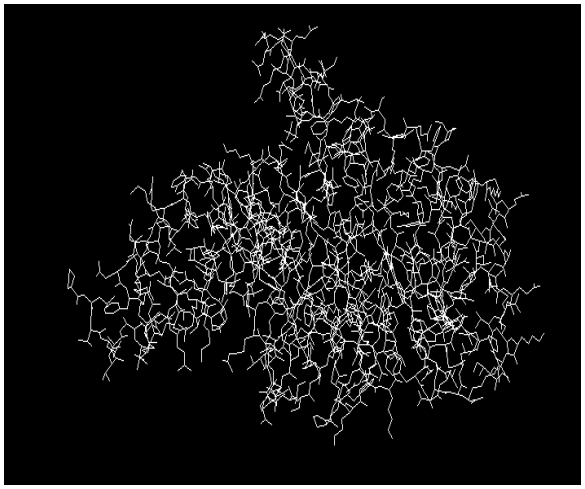


Figure 4: Enzyme 3BA8 showing wire frame for comparison with fig. 1.

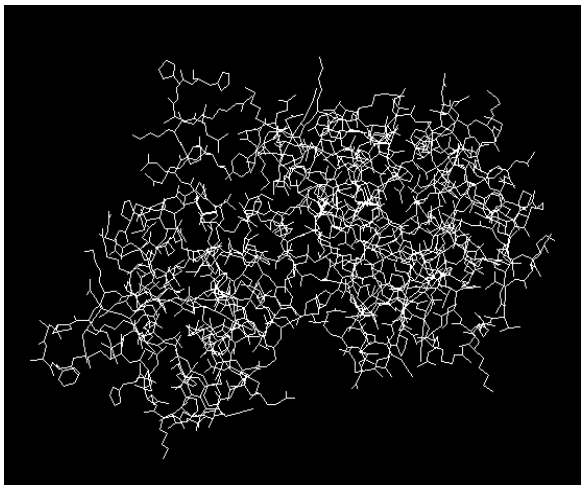


Figure 5: Enzyme 4HK4 showing wire frame for comparison with fig. 2.

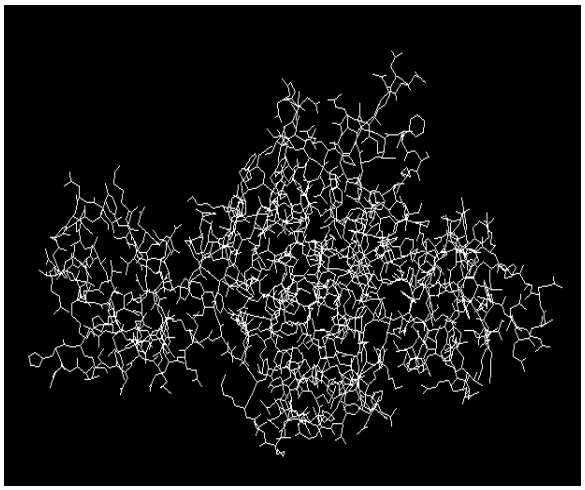


Figure 6: Enzyme 4HK4 showing wire frame for comparison with fig. 3.

For better comparison Figure 1 and 4 have to be magnified or Figures 2 and 4 reduced in size (to make the balls look identical or the wire frame sides become equal) for comparison of Fig 1 and 2 or 3 and 4.

Figure 4, 5 and 6 are the projections of the three molecules selected as seen from x direction i.e. projection on yz plane. For the three molecules, projections are obtained along x, y and z direction, i.e. in the yz, zx and xy planes respectively. The spatial distribution of atoms in the molecules varies from region to region and the associated complexity of shape varies. To demonstrate this, the complexity of structure and texture is studied in the three

directions and resulting fractal dimensions are compared. It is interesting to note that almost all the molecules possess fractal character exhibiting self similarity and scale invariance. All fractal shapes show self similarity and scale invariance which is found from the power law by using box counting technique. For the purpose of fractal analysis we implemented box-counting technique using the wire frames of the selected molecules that served as a two dimensional projection of the image of the skeleton of the molecule. This technique makes use of square boxes of different sizes (r) and scans the entire image to determine the total number of boxes (N) required covering the entire image. Computer programs are developed for the implementation of the box-counting technique the program performs the entire work of counting the number of boxes required (N) using boxes of different sizes (r) and finally saves the results in a file in tabular form for further processing and analysis.

To establish fractal character, using value of r and N from results of box counting a graph is plotted using log (N) on the y-axis and log(r) on the axis of x. If the power law is obeyed all the points lie along a straight line with a constant slope and the slope is the power law exponent. This also confirms the presence of self similarity and scale invariance. The presence of scale invariance provides the flexibility and freedom to use the images at a suitable (available) size thus it is not necessary to scale up or down the patterns to come to the same scale.

Figures 7, 8 and 9 present the result box counting implemented to the images shown in Fig. 4, 5 and 6 in the form of a graph of log(N) versus log(r). Figures 4, 5 and 6 are the projections of the images of the three molecules (3BA8, 4HK4 and 3RIR respectively) on the yz plane.

This plot shows if there is scale invariance and self similarity in the pattern. The presence of self similarity and scale invariance is indicated by the line joining these points. If the points lie along a straight line, power law is obeyed and Self-similarity and scale invariance is present.

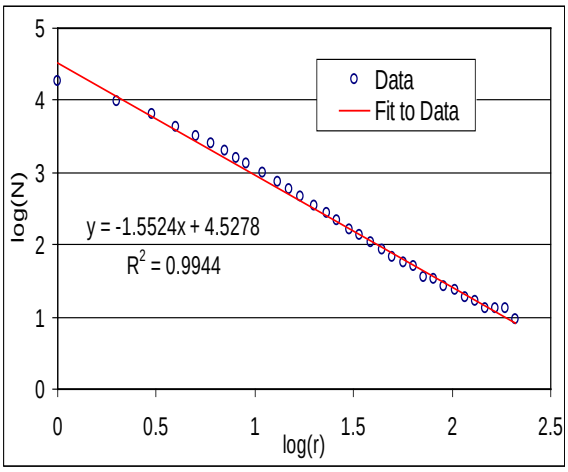


Figure 7: Log (N) versus log(r) plot for pattern shown in Fig. 4.

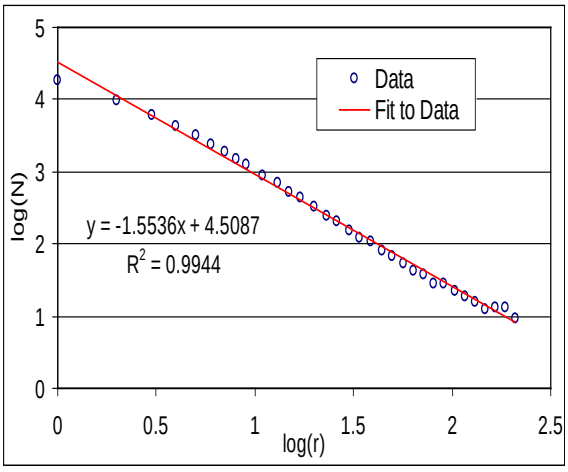


Figure 8: Log (N) versus log(r) plot for pattern shown in Fig. 5.

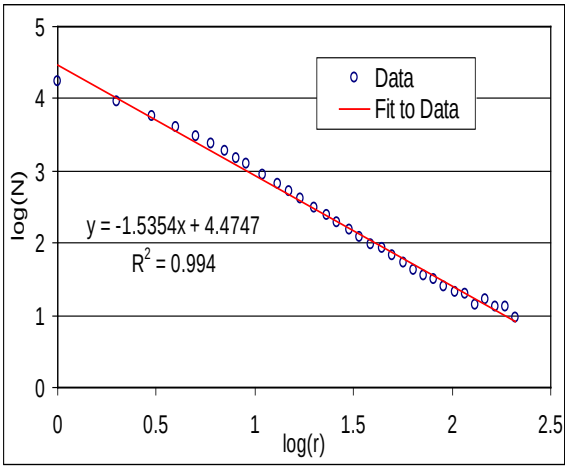


Figure 9: Log (N) versus log(r) plot for pattern shown in Fig. 6.

Fig. 7, 8 and 9 are the log(N) versus log(r) plots for Fig. 4, 5 and 6, the data plotted is from actual results of box counting and the straight line joining these point is the least square fit straight line joining these points. It is clearly seen from the plots that all the data points lie along a straight line which is also seen from the value of R^2 close to unity. The equation of the straight line best fitting the data points and the resulting value of R^2 are also shown on the graph in the inset. From the equation of the straight line, the slope can be found which the power law exponent is showing scale invariance.

Fractal dimension obtained from the slope of the straight line for molecule 3BA8 is 1.5524, that for 4HK4 is 1.5536 and that for 3RIR is 1.5354. A higher fractal dimension indicated higher degree of complexity with the pattern, this implies that the complexity of shape associated with the patter for 3BA8 and 4HK4 is more of less identical (as the fractal dimensions are close). The complexity of shape, structure and texture associated with shape of Fig. 6 i.e. projection of 3RIR on the yz plane is much less as compared to the earlier two (3BA8 and 4HK4).

For the sake of brevity all the tables and plots are not included and the findings are summarized in Table – 1 below. This gives the slope of straight line of the power law plot, fractal dimension and the value of R^2 that indicates how good the fitting is. A value or R^2 closer to unity indicates a good fit, it is clearly seen from the table that all the values of R^2 lie close to unity (greater than 0.99) indicating that the power law holds and the patterns are fractals with Self-similarity and scale invariance.

Table 1:

Molecule	Image	Slope	FD	R2
3BA8	3BA8_X	-1.5524	1.5524	0.9944
	3BA8_Y	-1.5442	1.5442	0.9923
	3BA8_Z	-1.5636	1.5636	0.9935
4HK4	4HK4_X	-1.5536	1.5536	0.9944
	4HK4_Y	-1.5515	1.5515	0.9937
	4HK4_Z	-1.5574	1.5574	0.9941
3RIR	3RIR_X	-1.5354	1.5354	0.9940
	3RIR_Y	-1.5329	1.5329	0.9873
	3RIR_Z	-1.5306	1.5306	0.9933

Comparison of fractal dimensions from the table for different projections of the three molecules selected is found to change slightly from projection to projection.

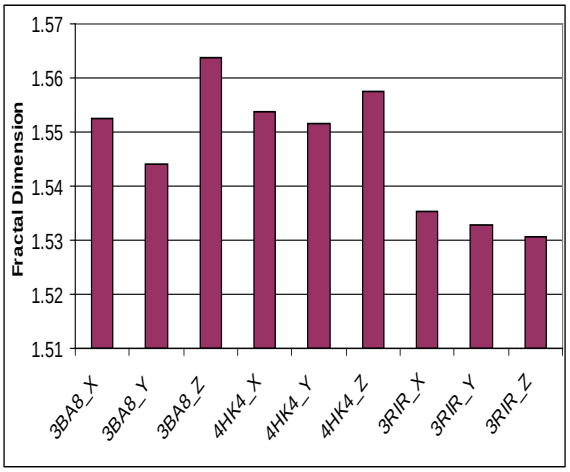
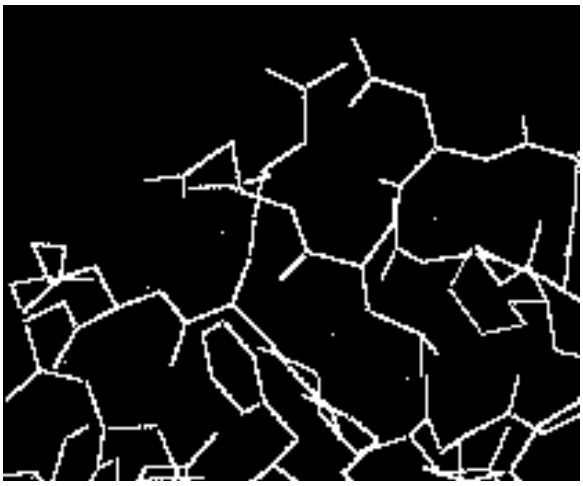
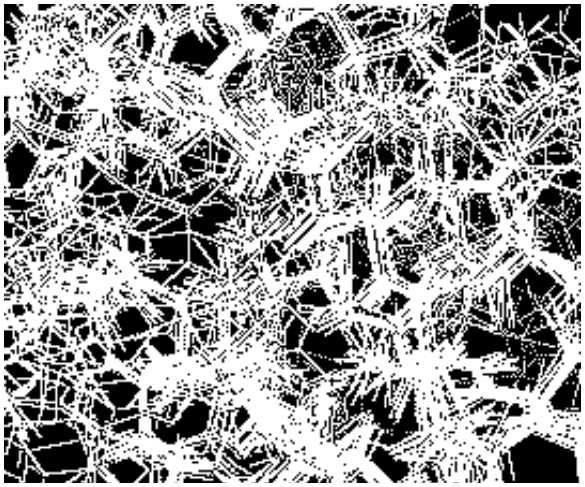


Figure 10: Comparison of Fractal dimensions of the nine images for the three molecules used.

For the first two molecules, the fractal dimension is highest for the projection in the xy plane, next lower fractal dimension is for the yz plane and the lowest fractal dimension is found for the projection on the zx plane. Most of the molecules exhibit this character, however for the third molecule i.e. 3RIR it is seen that fractal dimension is highest for projection on the yz plane then comes the zx plane and the lowest value of fractal dimension is found for the projection on the xy plane. Of course these variations are marginal and much less than one percent and thus can be ignored in qualitative comparisons. Two examples of low complexity images are shown below, the box counting of these two patterns yields a fractal dimension of 1.365 and 1.871 respectively.



A



B

Figure 11-A,B: Showing two patterns with different level of complexity.

Pattern with lower complexity have lower fractal dimensions and those with higher complexity show higher fractal dimension.

3. Conclusion

Characterization of complex structures like those of porous materials and complex molecules in terms of morphology poses difficulties. We used a simple technique base on the concept of fractals and fractal dimensions for characterization of complex shapes in a quantitative manner. Box counting technique is implemented on two dimensional projections of three biomolecules with lot of complexity of structure and successfully demonstrated that the technique is capable of quantifying complexity of structure in such cases. We present results of characterization of irregular shape of three molecules using projections of the molecule structure on three planes i.e. xy, yz and zx planes. It is established that the shapes of these molecules are fractals and obey power law proving that the self similarity and scale invariance exists. It is also shown that the degree of complexity slightly varies from projection to projection. Table – 1 summarizes the fractal dimensions obtained for the three molecules using three projections.

It is seen from Table – 1 and Fig. 10 that the molecule 3RIR exhibits lower fractal dimensions (in the range of 1.530 to 1.535) which is much lower as compared to the rest of the images. This is because of the fact that basically it is a molecule with limited features of shape associated with its structure. Work is in progress for implementation of 3D box counting of biomolecules and the findings are expected to provide a deeper insight into the shape, structure and texture of complex molecules.

REFERENCES:

1. Vijay Natarajan, Patrice Koehl, Yusu Wang, Bernd Hamann, Visual Analysis of Biomolecular Surfaces, *Mathematics and Visualization*, Springer, 237-255(**2008**)
2. Onda T., Shibuichi S., Satoh N., and K. Tsujii, 'Super Water-Repellent Fractal Surfaces' *Langmuir*, 12 (9), 2125(**1996)**(American Chemical Society)
3. Benoit B. Mandelbrot, the Fractal Geometry of Nature, W. H. Freeman and Company New York (**1983**)
4. B.B.Mandelbrot, Fractal geometry:what does it do ? Proc. R.Soc. Lond. A 423, 3 -16 (**1989**)
5. Paul S. Addison, Fractals and Chaos: An Illustrated Course Published by the Institute of Physics (IOP), London Ltd (**1997**)
6. Walter Beneson, John W.Harris, Horst Stocker, Holger Lutz, Hand book of Physics, published by Springer Verlag,Germany (**2000**)
7. Kenneth Falconer, 'Fractal Geometry: Mathematical Foundations and Application, John Wiley & Sons, Ltd United Kingdom(**2003**)
8. http://www.academia.edu/1114817/Biomolecules_Introduction_Structure_and_Functions_-Carbohydrates
9. Biomolecules **2014**, 4, 56- 75; doi: 10. 3390/ biom 4010056
10. Brian Maher, Andreas A. Albrecht, Martin Loomes, Xin- She Yang and Kathleen Steinh,biomolecules, 'A Firefly Inspired Method for Protein Structure Prediction in Lattice Models':www.mdpi.com /journal/ biomolecules
11. Software Arguslab: http://arguslab.en. softonic. com
12. Software Rasmol: http://www.chemistryteaching. com /rasmol.htm
13. Standard data processing tool: http://www. ehow.com/facts_5030382_methods-data-processing.html
14. Benoit B. Mandelbrot, *Science*, 637 (**1967**).
15. Witten T. & Sander L. M., *Phys Rev lett*. 47, 1400 (**1981**)
16. Zakde K. R., Munde S.V, Yusuf H.Shaikh,' Self-similarity and Fractal Character in Leaves, '*Journal*

of Chemical, Biological and Physical Sciences, 4,3; 3587- 3592(May2014– July**2014**).
17. Sander LM, *Scientific American*, 256, 94(**1987**)
18. Morency C, Chapleau R. *Harmonic and Fractal Image Analysis*;30,(**2003)**
19. Krishna Gopalakrishnan; Najarian, K., "Prediction of protein function using signal processing of biochemical properties", Bioinformatics Conference, **2003**. CSB **2003**. Proceedings of the 2003IEEEIssue,11-14 Page(s):536–538 (Aug.**2003)**
20. Armin B, Shlomo H. 'Fractals in science'_. Berlin Heidelberg: Springer-Verlag; **(1994)**
21. Nazneen Akhter, Yusuf Talib, Shaikh Yusuf H, **'**Fractal application for characterization of structure of macromolecules like carboh-ydrate 'J. Microbiol. Biotech.Res., 2 (1): 108 (**2012**)