Fractal and Mathematical Morphology in Intricate Comparison Between Tertiary Protein Structures

Ranjeet Kumar Rout,Pabitra Pal Choudhury, B. S. Daya Sagar, *Senior Member, IEEE,* and Sk. Sarif Hassan

Abstract—Intricate comparison between two given tertiary structures of proteins is as important as the comparison of their functions. In literature, several algorithms have been devised to compute the similarity and dissimilarity among protein structures. But, these algorithms compare protein structures by structural alignment of the protein backbones which are usually unable to determine precise differences. In this paper, an attempt has been made to compute the similarities and dissimilarities among tertiary protein structures using the fundamental mathematical morphology operations and fractal geometry which can decipher the intricate differences. In doing so, two techniques have been used here in determining the superficial structural (global similarity) as well as local similarity in atomic level of the protein molecules. This intricate structural differences would provide an insight to biologists to understand the protein structures and their functions in more precise fashion.

Index Terms—Tertiary Protein Structure, Geodesic Dilation, Skeleton and Fractal Dimension.

1 Introduction

Proteins are made of amino acids chain with its length typically ranging from 50 to more than 3000. A carbon atom $(called\, C_\alpha)$ is connected to a carboxyl (-COOH) group, an amine (-NH2) group, a hydrogen atom and a residue (which depends on the specific amino acid) to formulate a single amino acid. The amine group of an amino acid is covalently bonded by polypeptide bond with the carboxyl group of another amino acid to form a protein. The sequence of carbon atoms forms the backbone of the protein. Whenever the protein is left in its natural environment, it folds to a specific 3D structure. This is due to the forces between the amino acids such that the total free energy is minimized [6]. This renders a stable 3D protein structure. Thus, a protein can either be considered as polypeptides sequence of 20 amino acids occurring naturally which popularly knowns as primary structure of protein or as a tertiary structure into which a particular protein folds [9]. It is a known fact that the two almost similar amino acid sequences (primary structures) would not necessarily mean protein functional similarity and vice-versa [8, and, 9]. Therefore it is evident that in comparison of protein functionality requires the intricate comparison of tertiary structures. The search for an effective solution for tertiary protein structural similarity is justified because such tools can be of aid to scientists for prediction of the functions of a hypothetical protein, which would enable to drug design and understand evolutionary network among proteins[12] [13]. In the literature, as in many cases there is not even a single superposition that reveals all regions of similarity between compared proteins (RMSD, DALI, ProSup) [1]. Also, there are many conceptual difficulties associated with various methods (RMSD, ad hoc scores based on local secondary structure, hydrogen bonding pattern, burial status, or interaction environment) which have not been resolved [2]. Classical criteria such as the Root Mean Square Deviation (RMSD) fail to identify similar shapes in a consistent way [3]. To add on various systems have been proposed for structural classification, such as Structural Classification of Proteins (SCOP), Class Architecture Topology Homology (CATH), Families of Structurally Similar Proteins (FSSP), and others. The similarity in their cases is computed using structural alignment algorithms such as DALI, CE, VAST, SSAP and others. Most of these methods are computationally intensive and time-consuming, especially when searching large databases due to intrinsic complexity of structural alignment [7]. Also, the prevailing practice in the protein crystallographic community for computing structural differences is highly inappropriate, in particular when medium- and low-resolution structures are involved [4]. Geometrical feature like Fractal dimension of C_α of the backbone structure of one peptide chain proteins are considered in [17]. Petros [18] applied the spherical trace transform to a protein shape to produce a rotation invariant shape descriptor. In order to address the problem of structure reconstruction a method has been developed COMAR [19]. In the approach given in [20], the proteins are represented by their fractal dimension which represents the self-similarity of the protein structure. Since the fractal dimension of the protein, as

- *Ranjeet Kumar Rout and Pabitra Pal Choudhury are with the Applied Statistics Unit, Indian Statistical Institute(ISI), Kolkata,* 700018 *India. E-mail: ranjeetkumarrout@gmail.com and pabitrapalchoudhury@gmail.com*
- *B. S. Daya Sagar is with the System Science and Informatics Unit, Indian Statistical Institute (ISI), Bangalore,* 560059 *India. E-mail:bsdsagar@isibang.ac.in.*
- *Sk. sarif Hassan is with the International Centre for Theoretical Sciences, Tata Institute of Fundamental Research(TIFR), Bangalore,* 560012 *India. E-mail:sarif.hassan@icts.res.in.*

single feature in the structural level, cannot be a robust representative of its geometric features, the geodesic distance between the whole structures in different faces and fractal dimension in atomic level have been introduced as different features. There are a huge number of algorithms for retrieving protein tertiary structures which extract some features from the alternative representatives of the polypeptide chain such as 3D Spline [21, 22]. Obviously, a more objective method is highly desirable. In this paper, these problems have been tried to resolve by introducing two different methods using Mathematical Morphology and Fractals which would yield desired output.

The organization of the paper is as follows: in section 2, the basic review of Mathematical Morphology operations; in section 3, the result and analysis based on the two methods are discussed; conclusion has been made in section 4.

2 Basics of Mathematical Morphology and Fractal Dimension

Mathematical Morphology is a widely used paradigm in the field of image processing. Morphological tools are already very popular for image segmentation, image decomposition etc. Morphological operations like erosion, dilation, opening, closing are used for processing images frequently and produce results with high accuracy. Let A be a slice of protein structure decomposed from tertiary protein structure and S is the structuring element with certain characteristic information. In this technical note, each slice is subject to use some morphological transformation. The points "." in each slice are in black color pixels (A) and the background is in with white pixel (A^c). The definitions of these basic morphological operators are as follows [10],

Erosion: Erosion transformation of the slice A by S is defined in equation (1).

$$A \ominus S = \{a - s : a \in A, s \in S\} = \bigcap_{s \in S} A_s \tag{1}$$

Dilation: Dilation transformation of the atoms in the slice A by S increases the size of atoms, which is shown in equation (2).

$$A \oplus S = \{a + s : a \in A, s \in S\} = \bigcup_{s \in S} A_s \tag{2}$$

Opening: Opening transformation of the slice A by S is shown in equation (3).

$$A \circ S = (A \ominus S) \oplus S \tag{3}$$

Multiscale Openning: Multi scale morphological opening can be performed by increasing the size of S. The n^{th} size of S is shown in equation (4).

$$\underbrace{S \oplus S \oplus \ldots \oplus S}_{\text{n-times}} = nS \tag{4}$$

$A \ominus nS$, $A \oplus nS$ and $A \circ nS$ denotes morphological erosion, dilation and closing by S of size n.

2.1 Morphological Skeleton

Morphological skeleton of every geometrical structure is a subset of the original structure which has the same connectivity as the original structure from which inference can be drawn. From each point of the skeleton the distance to the boundary of the original set is the radius of a maximal circle (whose center is at a point of the skeleton) which touches the boundary at least two different points. The skeleton of an object gives a clear idea about the shape of the object. For the shape A, and the structuring element S, the skeleton [14][15] can be constructed through the operation as shown in equation (5).

$$Sk_n = (A \ominus nS) \setminus (A \ominus nS) \circ S, for\ n = 1, 2, \ldots, N \tag{5}$$

And the reverse process is defined in equation (6), where N is the number of performed iterations. Dilating the skeleton N times iteratively using the multi-scale structuring S elements a shape that interpolate the original shape.

$$A^{'} = \bigcup_{n=0}^{N} Sk_n \oplus nS \ , where\ nS = \underbrace{S \oplus S \oplus \cdots \oplus S}_{\text{n-times}} \tag{6}$$

2.2 Geodesic Dilation

Geodesic dilation is a morphological transformation to operate only some part of the image (as marker) to grow until the boundary of the image on which it is applied. This transformation is used to find out similarity between two $2D$ images. The advantages of this transformation is that the structuring element can vary at each pixel, according to the image. The Geodesic Dilation δ_x of an image Y inside X is defined as the intersection of the dilation of Y (with respect to a structuring element S) with the image X as shown in equation (7).

$$\delta_x^n = (Y \oplus nS) \cap X \ \ where\ n = 1, 2, \cdots, n \tag{7}$$

So *Geodesic dilation* terminates when all the connected components of X are constructed i.e. idem potency is reached and defined in equation (8).

$$\forall n > n_0,\ \delta_x^{(n)}(Y) = \delta_x^{(n_0)}(Y) \tag{8}$$

2.3 Fractal Dimension

A fractal dimension is an index for characterizing fractal patterns or sets. The patterns illustrate self-similarity and the fractal dimension indicates the extent to which the fractal objects fills a particular Euclidean space in which it is embedded. The most commonly used technique in determining the fractal dimension is *Box Counting Method* which is briefly stated as follows.

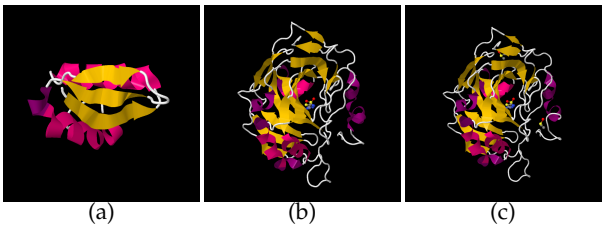
Box-Counting Method: This method computes the number of cells required to entirely cover an object, with grids of cells of varying size. Practically, this is performed by superimposing regular grids over an object and by counting the number of occupied cells. The

logarithm of $N(r)$, the number of occupied cells, versus the logarithm of $\frac{1}{r}$, where r is the size of one cell, gives a line whose gradient corresponds to the box dimension [15, 16]. To calculate the dimension for a fractal Sk_n , the Box-Counting dimension is defined as shown in equation (9).

$$Dim_{box}(Sk_n) = \lim_{n \to 0} \frac{logN(r)}{log\frac{1}{r}} \tag{9}$$

3 Data used and Specification

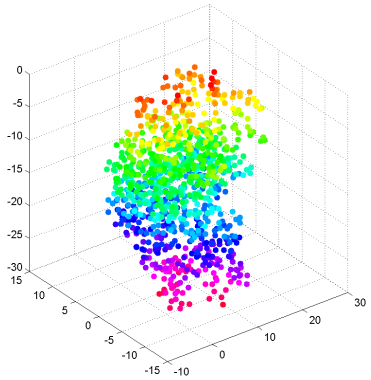
The Protein Data Bank PDB (http://www.rcsb.org/pdb/home/home.do) is the largest and most commonly used repository for any kind of information regarding proteins.The most popular techniques in obtaining protein tertiary structure are the X-Ray crystallography and Nuclear Magnetic Resonance (NMR). From the PDB database three proteins viz. $2LEP$, $3V2J$ and $3V2M$ in the standard *.pdb* format are collected and are shown in Fig. 1.



(a) (b) (c)

Fig. 1. Tertiary structure of three protein (a)$-2LEP$, (b)$-3V2J$ and (c)$-3V2M$.

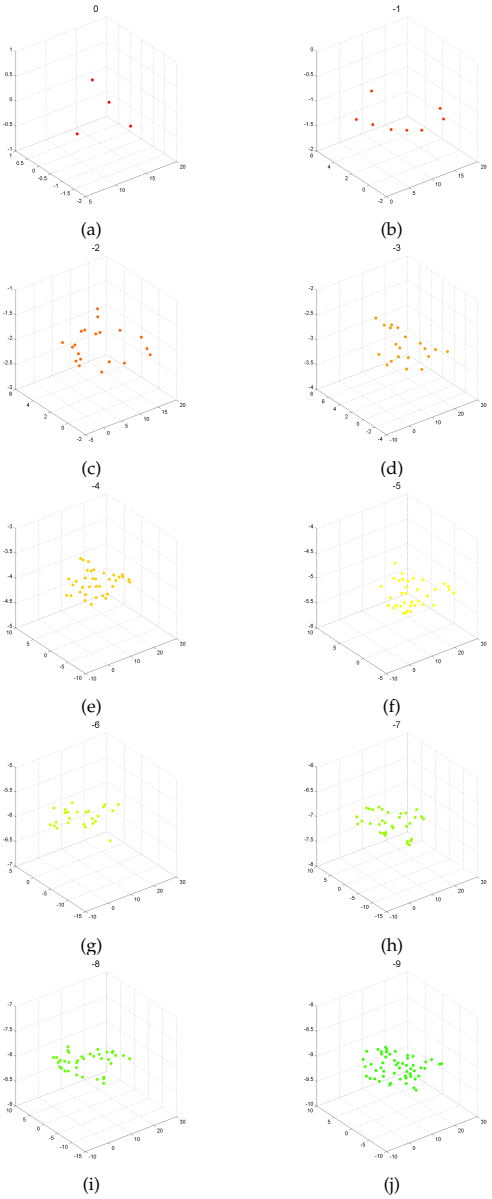
The tertiary structure is represented in a standard Cartesian coordinates of the atoms presented in the protein $2LEP$ is shown in Fig. 2.



2LEP

Fig. 2. Atoms of protein $2LEP$ having different color with respect to $z-$axis.

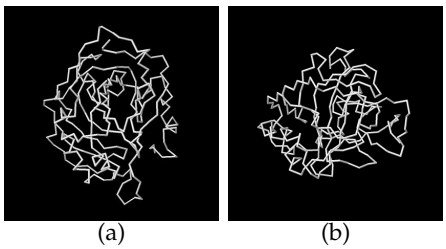
In the present work, the tertiary structure of protein has been sliced with reference to z-axis into non-overlapping slices. The methodology of the slice decomposition is adumbrated in Section 4. Some of the slices of protein $2LEP$ are depicted in Fig. 3.



(a) (b) (c) (d) (e) (f) (g) (h) (i) (j)

Fig. 3. (a-j) shows atoms of protein $2LEP$ at different $z-$axis.

Each of the tertiary structure of the protein Fig.1b and Fig.1c have been converted into backbone structure by $JMol$ viewer [http://www.rcsb.org/pdb/explore/jmol.do?structureId=3V2J&bionumber=1]. The backbone structure of the above two protein structures $3V2J$ and $3V2M$ are shown Fig. 4.



(a) (b)

Fig. 4. 3D Backbone structures of proteins (a) for $3V2J$ and (b) for $3V2M$.

The three tertiary protein structure shown in Fig.1a, Fig.1b, and Fig.1c are eventually used to demonstrate the two approaches respectively based on fractal dimension and geodesic dilation. First approach is demonstrated on the thirty slices decomposed from $2LEP$(Fig. 2) are shown in Fig. 3. The $3V2J$ and $3V2M$ protein structure and their corresponding backbone structure Fig.4a and Fig.4b were used to demonstrate geodesic base dilation.

4 THE INTENDED METHODS

In this section, two different methods are proposed to compare between two tertiary structures in intricate level on the basis of mathematical morphology and fractal geometry. These methods are explained in this section have been discussed for the shake of clarity on three tertiary protein structure acquired from PDB database.

4.1 Tertiary Structure Skeleton and Its Fractal Dimension

The detail method is described in different steps as follows:

Step I. The slices (non-overlapping)of the tertiary structure of protein f is obtained by slicing the tertiary structure along z-axis on its m ordinates with m planes $P_1, P_2, \cdots P_m$ that are perpendicular to the z-axis. These slices are represented as different functions $X^1, X^2, \cdots, X^m$, such that $\cup_{i=1}^{m} X^i = f$. The slices are made in such a manner so that the union of m slices contains almost all atoms of the tertiary protein structure as shown in Fig. 3. Each slice contains the protein atoms that share the same z-coordinate.

Step II . The slices consisting of atoms of the tertiary structure of $2LEP$ is represented by f. The function f could be thirty slice like $X^0, X^1, \cdots, X^{29}$ and only some are shown in Fig. 3. By performing the multi-scale opening on X^7(Fig.3h) slice until the N^{th} cycle yields a single connected component, such that all the atoms are contained by the connected component. Now $X^7 \circ nS$ is the connected component of X^7 as shown in Fig. 6k.

Step III. The morphological skeleton of the X^7 connected component

$$Sk_n(X^7) = (X^7 \ominus nS) \setminus (X^7 \ominus nS) \circ S,$$

for $n = 0, 1, 2, 3, \cdots, N$ and $i = 1, 2, 3, \cdots, m$ is obtained.

Step IV. Similar approach has been followed for generating connected components for rest of the slices decomposed from $2LEP$ protein structure. The skeleton network of the m^{th} level connected components of all the slices are generated and have been stacked as shown in Fig 7.

$$Sk = \bigcup_{i=1}^{m} X^i \ , where \ i = 1, 2, 3, \ldots, m.$$

Step V.The fractal dimension $D(p)$ of the Skeleton Sk is determined by the Box counting method.

The local similarity of the tertiary structure is obtained through the steps as mentioned earlier in this section. Each " $\cdot$ " represents an atom. The slice contains all the atoms whose z coordinate is at -0.700 . After getting a slice, a connected component has been obtained by the multi-scale opening of the atoms presented in that slice. For each iteration of multi-scale opening the size of structuring element increased by one. And for the opening, a primitive structuring element of size $n \times n$ where $n = 1, 2, \cdots, N$ is used. In doing so, the tertiary structural atomic information has been transformed into two dimensional plane without loosing much information. The iterations for the slice in Fig. 3. of $2LEP$ protein is obtaining a connected component is shown in Fig. 6. In

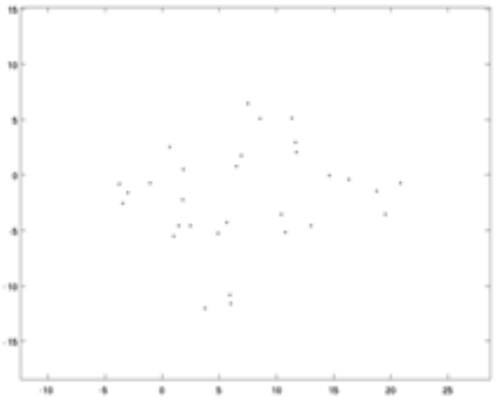


Fig. 5 Slice of protein $2LEP$ at z-axis= -7.

Fig. 6 from a to j shows the multi-scale opening using primitive structuring elements. Starting from size one, each figure shows the iteration with structuring element larger by one units from the previous one. The iterative opening may take a large number of iterations to contain all the atoms in a particular slice. And there may be more than one plane for a slice. So we dilate the plane with a primitive structuring element. This reduces the number of planes for each slice. The example of the plane after dilating with a disk shape structuring element of size 25, the resulting plane becomes a single connected component as shown in Fig. 6k. After acquiring all the planes for a particular protein structure, our next aim is to find the skeletons for each of the plane shapes. Example of the skeleton for the plane shown in Fig. 6l. If we stack the skeletons for all the planes over each other, then the resulting image gives us an idea of how the atoms form the overall protein structure in terms of the planes, that are formed by the coordinates of the atoms. For the protein *2LEP* the skeleton structure is shown Fig. 7, from the skeleton we have an idea of fractal-like distribution of protein atoms of the tertiary protein structure in the from of plane. Now we can compute the fractal dimension D_p of the skeleton of the corresponding tertiary protein structure and use it as the

feature of tertiary protein structure. Fractal Dimension D_p of a group of protein molecules are given in **Table 1** irrespective of their residue.

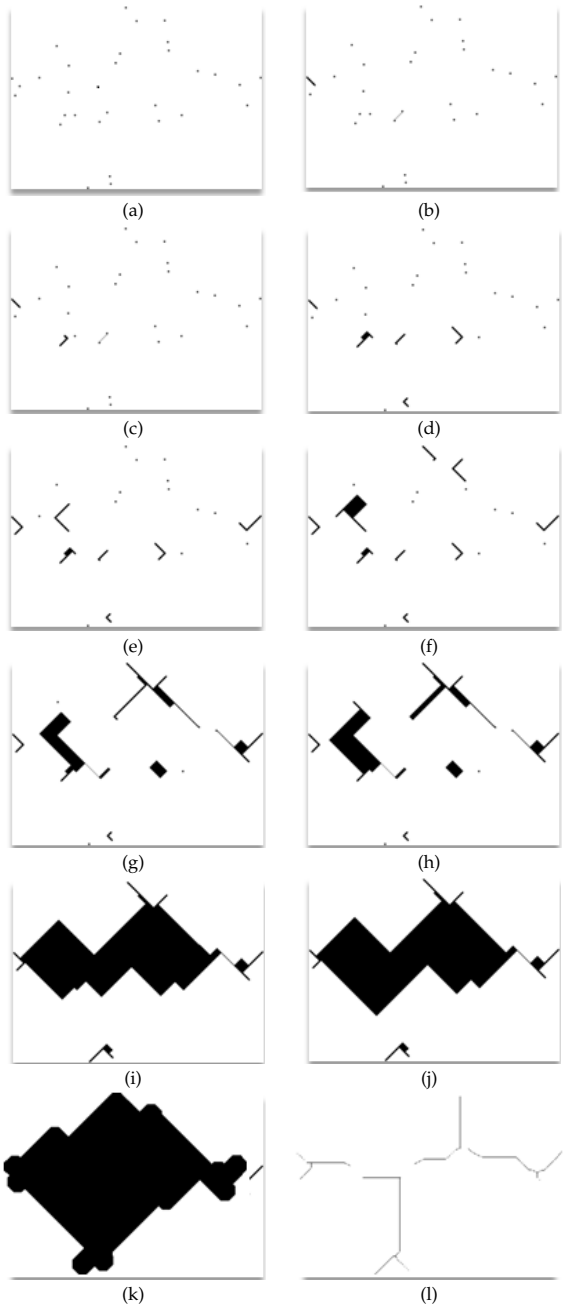
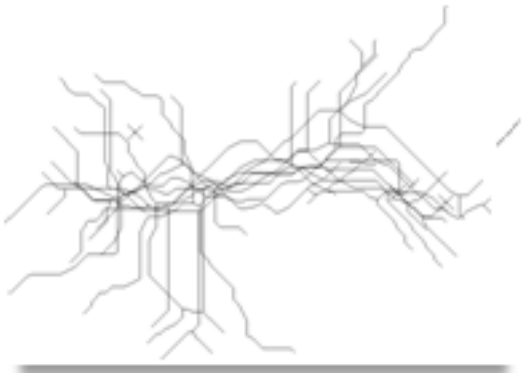


Fig 6. (a) shows a slice of protein $2LEP$. Fig.6(b-k) shows different iterations of multi-scale opening of Fig.6a, and Fig.6l shows the skeleton of the connected component of fig.6k.

4.2 Geodesic Dilation and Its Quantification

Unlike the existing algorithms [6,7,8,13,17, and 20], this method is only take care of tertiary structure of proteins. To apply Mathematical Morphology comfortably, the tertiary structure of proteins are converted into a collection of $2D$ objects. From the PDB database the protein structures $3V2J$ and $3V2M$ are viewed by using



2LEP

Fig. 7. The skeleton network constructed for all the 30 slices of $2LEP$ protein.

TABLE 1
Fractal Dimension of Protein Molecules

Protein ID	Residue Number	$D(P)$
3v2j	260	1.6612
3smk	236	1.6201
3t0o	238	1.6465
4ecs	435	1.6495
3v2m	260	1.6561
3sy1	190	1.6054
4ag2	226	1.6959
1cah	259	1.6611
1cai	259	1.6605
4bij	476	1.6802
2lep	69	1.5496
1cgi	245	1.6359
4eym	371	1.6495
2cbc	260	1.6614

JMol as shown in Fig. 1. and the protein structures are rotated depending on the 3-axis, from which the 6-faces or views (front, left, right, top, bottom, and back) of each tertiary protein structures are collected respectively as shown in Fig. 8. The following step are used for calculating Geodesic Dilation δ_p of the corresponding faces of protein molecules.

Let the six faces of each of the backbone structure are denoted by X^1(front), X^2(left), X^3(right), X^4 (top), X^5(bottom), and X^6(back). The six faces for each of the protein structure ($3V2J$ and $3V2M$) are shown respectively in Fig. 8(a-l). The aim of this section is to compute the similarity index between the backbone structure faces from source structure X_s^i and the target structure X_t^j. Here both the source backbone structure and target backbone is equals to

$$\bigcup_{i=1}^{6} X_s^i \; and \; \bigcup_{j=1}^{6} X_t^j.$$

Similarity index for X_s^i and X_t^j computed by geodesic dilation operation is as follows:

Step 1.If $X_s^i \cap X_t^j = X_s^i \cup X_t^j$, then there exists exact

similarity. In such cases, the dilation distance between X_s^i and X_t^j, and between X_t^j and X_s^i would be zero.

Step 2. By performing dilation operation $min\{n : X_s^i \subseteq (X_s^i \cap X_t^j) \oplus nS\} = \triangle_s$, where S is a structuring element,

Step 3. Similarly, By performing dilation operation $min\{n : X_t^j \subseteq (X_s^i \cap X_t^j) \oplus nS\} = \triangle_t$

Step 4. Difference between the number of Dilation with respect to all faces from source protein to target protein structure is defined as

$$\triangle = \sum_{i=1}^{6} \mid \triangle_s - \triangle_t \mid$$

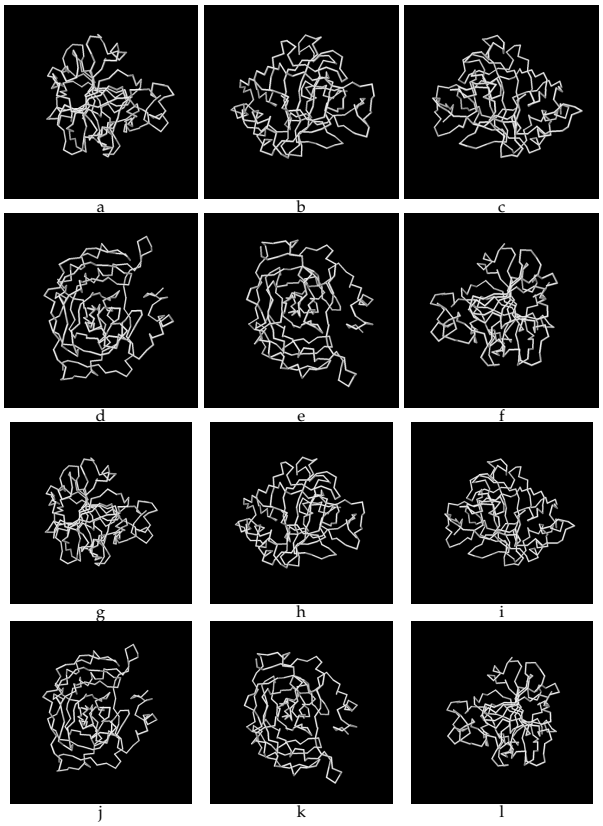


Fig. 8. (a-f) shows Six different 2D-faces of $3V2J$ protein and (g-l) shows Six different 2D-faces of $3V2M$ protein.

Here the front view of two different proteins are considered to compute the global similarity between them. For this purpose, specifically the front view of $3V2J$ and $3V2M$ as shown in the Fig. 8a and Fig. 8g. The common structural part of both the protein molecules is given in Fig.9. The Geodesic dilation as a device to determine the structural similarity between those protein molecules. The number of dilations required from the intersection part $(X_s^i \cap X_t^j)$ to both the protein molecules towards constructing the similar structure, i.e. the number of geodesic dilation from $X_s^i \cap X_t^j$ towards $3V2J$ and $3V2M$ is four for each. Similarly for other faces are given in Table 2. So the number of geodesic dilation $\triangle = 2$

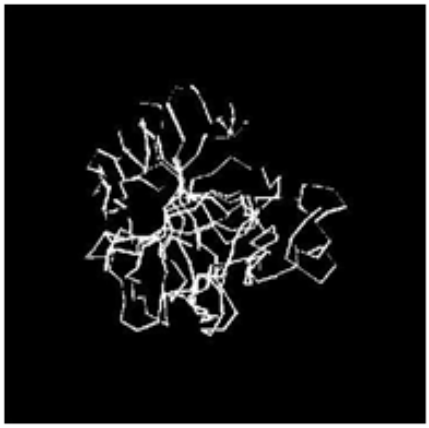


Fig. 9. intersection parts between front face of $3V2J$(Fig.8a) and front face of $3V2M$(Fig.8g) protein structure.

TABLE 2
Geodesic dilation of Different Faces

Protein ID	$X_s^i \cap X_t^j$	Geodesic Dilation $\triangle$ of different faces						$\triangle$
		Front	Left	Right	Top	Bottom	Back	
3V2J	3V2J ∩ 3V2M	4	4	5	6	5	3	2
3V2M		4	4	4	6	4	3	
2LE8	2LE8 ∩ 2LLS	10	6	6	7	8	10	64
2LLS		19	19	21	18	15	19	
1CAH	1CAH ∩ 2CBC	1	1	1	1	1	1	1
2CBC		1	1	1	2	1	1	
1CAH	1CAH ∩ 4EYM	10	10	9	18	18	11	35
4EYM		20	6	6	9	11	9	
3V2J	3V2J ∩ 3T0O	12	8	9	12	12	11	35
3T0O		4	5	5	5	11	5	

with respect to each faces, which shows that $3V2J$ and $3V2M$ have more structural similar, whereas for $2LE8$ and $2LLS$ have less structural similarity as $\triangle = 64$. Table 2 shows the different geodesic dilation of different protein molecules.

4.3 Result and Discussion

In our experiment, we downloaded about two hundred protein molecules from PDB and the result shows that our methodology performance quite well for comparing tertiary protein structure in intricate level. The similarity between two protein structures i and j can be computed by using the following equation:

$$\rho = |D_p(i) - D_p(j)|$$

and Geodesic Dilation $\triangle$, Where ρ is the difference between the fractal dimensions of any two protein molecules and some experimental results are shown in *Table 3*. The difference between the fractal dimension is essentially measure the difference between the structural complexities. As ρ approaches to Zero, the structures are closed to be similar. The experimental result shows that if $\rho \leq 0.008$ and $\triangle \leq 12$, two protein molecules are similar in structures and functions. Thus, lower difference between fractal dimensions and Geodesic dilation will ensure high similarity between the proteins which are being compared. This would become clearer with few examples. For the same from the **Table 3** we conclude that the proteins molecule **1CAH** is more similar to

1CHI and **2CBC** as their fractal dimension difference $\rho = 0.000604, 0.000314$ and geodesic dilation $\triangle = 5 and 1$ respectively.

TABLE 3
Difference between Fractal Dimension of Compared Proteins Pairs

Protein ID-1	FD- $D_i(p)$	Protein ID-2	FD- $D_j(p)$	ρ	δ_p	PDB Result
		3SMK	1.6201	0.0411	24	12 %
		3T0O	1.6465	0.0147	35	18 %
3V2J	***1.6612***	4ECS	1.6494	0.0118	31	2 %
		3V2M	**1.6561**	***0.0051***	**2**	**100 %**
		3SV1	1.6053	0.0559	28	28 %
		4AG2	1.6959	0.0347	39	31 %
		1CAI	***1.6605***	***0.0006***	***5***	***100***%
		4BIJ	1.6802	0.0191	21	41 %
1CAH	***1.6611***	2LEP	1.5497	0.1114	43	57 %
		1CGI	1.6359	0.0252	35	50 %
		4EYM	1.6495	0.0117	35	39 %
		2CBC	***1.6613***	***0.0002***	***1***	***100*** %

5 Conclusions

In this work, two technique have been devised to compute the structural similarity of protein structure. Unlike the other methods the indented methods are computationally efficient as well as very objective in catching the structural difference. In the experiment, atoms of all the protein structures are divided into slices by fixing the z co-ordinate values. It is needless to mention that the present work is based on area based morphological analysis, which can be extended towards the volumetric analysis of the tertiary protein structure in near future.

Appendix A
Symbol Table

TABLE 4
Symbol Table

Symbols Used	Propose of symbol
Slice of Tertiary Protein structure	A
structuring Element	S
Erosion Operator	$\oplus$
Erosion Operator	$\ominus$
Opening	$\circ$
Skeleton of a Slice of a Tertiary protein structure	$\ominus$
Connected component of i^{th} slice	C_i
Skeleton of the i^{th} connected component	$Sk_n(C_i)$
Skeleton of the tertiary protein structure	Sk_p
Fractal dimension of the Skeleton of the tertiary protein structure Sk_p	$D(p)$
i^{th} face/view of the source tertiary protein structure	X_i
i^{th} face/view of the targeted tertiary protein structure	Y_i
i^{th} marker of the two given tertiary protein structure	I_i
Number of dilation required from marker of the two given tertiary protein structure to Source protein structure	$\delta_{p_s}^n$
Number of dilation required from marker of the two given tertiary protein structure to target protein structure	$\delta_{p_t}^n$
Difference between $\delta_{p_s}^n$ and $\delta_{p_t}^n$	δ_p^n
Difference between the fractal dimension of source and targeted protein structure	ρ

References

[1] A. Zemla, *LGA: a method for finding 3D similarities in protein structures*; Nucleic Acids Research, Vol.-31, Issue-13, pp. 3370-3374 (2003).

[2] D. Goldman, C. H. Papadimitriouz, S. Istraily, *Algorithmic Aspects of Protein Structure Similarity*. FOCS 99 Proceedings of the 40th Annual Symposium on Foundations of Computer Science, pp. 512.(1999)

[3] F Guyon, P. Tuffry, *Assessing 3D scores for protein structure fragment mining*. Open Access Bioinformatics, Vol.-2, pp-6777(2010).

[4] G. J. Kleywegt, *Experimental assessment of differences between related protein crystal structures*. Acta Crystallogr. D Biol. Crystallogr, Vol.-55, pp. 1878-1884(1999).

[5] G. Klette, *Skeletons in digital image processing*. July, (2002).

[6] G. Lancia, S. Istrail, *Protein Structure Comparison: Algorithms and Applications*. In: Guerra, C., Istrail, S. (eds.) Mathematical Methods for Protein Structure Analysis and Design. LNCS (LNBI), Vol. 2666, pp. 133. Springer, Heidelberg (2003).

[7] I.G. Choi, J. Kwon, S. H. Kim, *Local feature frequency profile: A method to measure structural similarity in proteins*. PNAS, Vol. 101; pp.-11(2004).

[8] J. Galgonek, D. Hoksza, T. Skopal, *SProt: sphere-based protein structure similarity algorithm*. Galgonek et al. Proteome Science, Vol.-9(Suppl 1):S20 (2011).

[9] L. P. Chew, *Exact Computation of Protein Structure Similarity*, Proceeding SCG '06 Proceedings of the twenty-second annual symposium on Computational geometry pp.468-474(2006).

[10] L. L. Teo, B. S. Daya Sagar, *Modeling, description, and characterization of fractal pore via mathematical morphology*. Discrete Dynamics in Nature and Society, Vol. 2006, Article ID 89280; Pages 124 (2006).

[11] L. Holm, C. Ouzounis, C. Sander, G. Tuparev, G. Vriend, *A database of protein structure families with common folding motifs*. Vol.-1, Issue-12, pp.1691-1698(1992).

[12] N. Krasnogor, D. A. Pelta; *Measuring the similarity of protein structures by means of the universal similarity metric*, Bioinformatics, Vol.- 20,pp.1015-1021(2003).

[13] P. Koehl, *Protein structure similarities*. Curr Opin Struct Biol, Vol.-11, pp.348-353(2001).

[14] Petros A. Maragos, Ronald W. Schafer; *Morphological skeleton representation and coding of binary images*. IEEE Transactions on Acoustics, Speech and Signal processing, Vol. assp-34; pp. 5(1986).

[15] D. Avnir, O. Biham, D. Lidar, O. Malcai, *Is the geometry of Nature fractal*. Science Vol.-279, pp. 39-40(1998).

[16] K Develi, T Babadagli, *Quantification of natural fracture surfaces using fractal geometry*. Math. Geology, Vol.- 30 , pp. 971-998(1998).

[17] C Cui, D. Wang, X. Yuan, *3D protein structures similarity matching based on fractal features*. SPIE Vol. 5637, pp-567-572(2005).

[18] P Daras, D Zarpalas, A Axenopoulos, D Tzovaras, MG Strintzis, *Three-Dimensional Shape-Structure Comparison Method for Protein Classification*. IEEE/ACM Transactions on Computational Biology and Bioinformatics 2006, 3:193-207.

[19] M Vassura, L Margara, P Di Lena, F Medri, P Fariselli, R Casadio, *"Reconstruction of 3D Structures From Protein Contact Maps"*. IEEE/ACM Transactions on Computational Biology and Bioinformatics 2008, 5:3. July-September 2008.

[20] C. Cui, D. Wang, and J. Shi, *Comparing 3-D Protein Structures Similarity by Using Fractal Features,* Proc. IEEE Computational Systems Bioinformatics Conf. (CSB 04), pp. 698-699, Aug. 2004.

[21] T. Can and Y.F. Wang, *CTSS: A Robust and Efficient Method for Protein Structure Alignment Based on Local Geometrical and Biological Features,* Proc. IEEE CS Conf. Bioinformatics (CSB), pp. 169-179, Aug. 2003.

[22] A. Bhattacharya, T. Can, T. Kahveci, A.K. Singh, and Y.F. Wang, *ProGreSS: Simultaneous Searching of Protein Databases by Sequence and Structure*, Proc. Pacific Symp. Biocomputting, pp. 264-275, Jan. 2004.