A Framework for Structural Similarity Search in Proteins

Shi-Hua Zhang* De-Gang Liu Xiang-Sun Zhang

Institute of Applied Mathematics
Academy of Mathematics and Systems Science
Chinese Academy of Sciences, Beijing 100080, China

Abstract *The expanding protein sequence and structure databases await methods allowing rapid similarity search. A novel framework for structural similarity search in proteins is developed. It can not only distinguish the local similarity case, but also can detect the overall similarity in some sense. First, geometrical distance of C_α-C_α is used in evaluating structural similarity in proteins. Then a simple score metric is given out according to the geometrical distance for local similar fragment pairs. The rigid superimposition algorithm is used for computation of RMSD and rotation matrix. In order to detect the overall structure similarity, the notion of spacial compatibility is introduced using rotation matrix. Experimental analysis verifies that the novel framework is effective.*

Keywords structure similarity search, similarity score metric, consensus structure, spacial compatibility

1 Introduction

Structure genomics initiatives are set to produce a large amount of data, so there is a clear need for novel, fast data analysis strategies to extract biologically relevant similarity information. Many methods have been designed for searching structural similarity in proteins. And various properties of polypeptide chains in proteins can be used as criteria for similarity estimation such as structure alignment, RMSD (the root mean square deviation of all aligned C_α atoms) calculation (Hubbard, 1999, Guda et al., 2001), contact maps expressing the inte-residue distance (Ortiz et al., 1999), geometrical parameters-dihedral angle and radius of curvature (Leluk et al., 2003, Dua et al., 2004), C_α-C_α distance frequency (Carugo et al., 2002), fractal features (Cui et al., 2004), environmental properties (Jung et al., 2002) such as solvent-accessible surface, and conformational properties such as dihedral angles or the mutual orientation of centers of masses (Shindyalor and Bourne, 1998).

*Corresponding author. Fax: +86-10-62561963. E-mail address: zsh@amss.ac.cn

However, the quantitative assessment of structural similarity is problematic in many respect. First of all, the comparison of three-dimensional (3D) structures is very computation-intensive, which is partly due to the nature of structural alignments. Second, some properties such as the RMSD calculation work well as indicator of similarity only if the structures are closely related. Clearly, distantly related structures may only share a small segment that can be structurally aligned. Third, even for distantly overall related structures, diverse methods will produce diverse results e.g., diverse correspondence, RMSD value and aligned residue number.

The present work aims to define some simple linear profiles of the geometrical distance parameters of the 3D protein structure and a simple similarity score metric that can detect the local similarity situation fast and effectively for overall related structures so as to allow the searching for more detailed local similarity information. A rigid superimposition algorithm (Schwartz and Sharir, 1987) is used for calculating of RMSD and rotation matrix. RMSD value indicates that a few fragment pairs may be overlarge, and those are removed. A grouping algorithm based on the demographic clustering technique of data mining (Cabena et al., 1997) is used for clustering all the fragment pairs according to their relevant rotation matrices. In conclusion, it retrieves the overall similarity in a special way by a novel notion of spacial compatibility. The proposed method may be used for comparison of structures or as a seed for that. They could be used for classification of structures in proteins and so on. A detailed analysis and complex comparison of the members of serpine family is performed on the basis of the presented parameters with the noted DALI program (Holm and Sander, 1993) for the same purpose. A notion of "consensus structure", analogous to "consensus sequence" that has been introduced by Leluk et al.(2003) is suitable for our similarity search.

2 Methods

Our first goal is to find a profile of proteins and a similarity score metric that could be used to quickly detect local structural similarities of two proteins. We can image that two similar structure must have mostly the same relative distance between C_αs. Further, the distance between two consecutive C_αs is always around 3.8 Åand the angles between three consecutive C_αs vary lightly around 100°. So the distance between C_αs can be as a profile for the structure of protein. We define a distance vector DV_{N+1}:

$$\begin{aligned} DV_{N+1} &= [d(C_\alpha(i), C_\alpha(i+N))]_{i=1,\cdots,n-N} \\ &= [d_{1,1+N}, d_{2,2+N}, ..., d_{n-N-1,n-1}, d_{n-N,n}], \end{aligned}$$

where $d(C_\alpha(i), C_\alpha(i+N))$ indicates the distance between the ith C_α atom and the $(i+N)$th C_α atom in a protein backbone chain $\{C_\alpha(1), \cdots, C_\alpha(n)\}$ and N is equal to 3, 4.

We call a fragment which has k consecutive C_α atoms as a k-C_α fragment. A similarity score $S(i, j)$ (simply, the S score) is defined for two fragments i and j in

two proteins respectively according to the distance vector DV_4 and DV_5. For k-C_α fragments where k equal to $4, 5$, $S_4(i,j)$ and $S_5(i,j)$ are defined as follows:

$$S_4(i,j) = D(d^A_{i,i+3}, d^B_{j,j+3}),$$

$$S_5(i,j) = \min\{D(d^A_{i,i+3}, d^B_{j,j+3}), D(d^A_{i+1,i+4}, d^B_{j+1,j+4}), D(d^A_{i,i+4}, d^B_{j,j+4})\},$$

where

$$D(d_1, d_2) = C_1 - \frac{|d_1 - d_2|}{d_1 + d_2},$$

where C_1 is a positive constant, $d^A_{i,i+N}$ is the ith component of the distance vector DV^A_{N+1} of the protein A and likewise for $d^B_{j,j+N}$. $S_k(i,j)$ value represents the similarity situation of the ith k-C_α fragment in protein A to the jth k-C_α fragment in protein B i.e., the S score is more closer to C_1, the two fragments are more similar. We then prune S_{ij} in favor of consecutive and high-scoring segment: (a) Pairs with negative S scores are eliminated. (b) Pairs with isolated high S scores, i.e., those that can not form a stretch of four high-scoring pairs are also eliminated. The remainder S scores can form various corresponding pair-wise fragments of protein A and B respectively. The similarity dot-matrix of two proteins can be quickly computed.

As a matter of fact, similar similarity dot-matrix of two proteins can also be computed by using other measures such as (1) RMSD of 4 or 5-C_α fragments pairs; (2) the root mean square (RMS) value of the 4 or 5 residue (ϕ,φ) torsion angles (Ramachandran et al., 1963); (3) the α angles (the torsion angle defined by four consecutive C_α atoms) (Levitt, 1976); (4) other dihedral angles introduced in (Leluk et al., 2003, Dua et al., 2004). All these measures can obtain similar similarity dot-matrix, but the RMSD is expensive in numerical computation; the (ϕ,φ) torsion angles and other dihedral angles are too noisy and reflective of local small change for their flexible non-C_α atoms; the α angles and the $d(C_\alpha(i), C_\alpha(i+3))$ distance have one-to-one relationship in some sense.

Although the S score means some similarity, few remainder S score corresponding pair-wise fragments may have high RMSD value. There is the superimposition problem: given the corresponding C_α-atom set find a rotation and translation which superimposes one set to the other with minimal RMSD. The problem has been dealt with intensive and efficient solution. Many methods have been developed (Schwartz and Sharir, 1987; Kabsch, 1978; Horn, 1987; Besl and Mckay, 1992). We first compute the pair-wise fragments' RMSD value and it's rigid transform T consisting of a rotation matrix R and a translation vector a by using the algorithm of Schwartz and Sharir (1987) whose complexity is linear in the number of the matched C_α-atom pairs. Then we prune the fragments whose RMSD value is more than a certain threshold (C_2, e.g. 3.5 Å). Thus many local similar fragment pairs are given out.

Two overall structural similar proteins will also have similar local fragments and all those local similar fragments will be compatible in space i.e., at least they will have a roughly equal rotation matrix. Then we can detect local similarity

fragments whether they are subjected to "spacial compatibility". The concept of spacial compatibility is defined as follows: given a set of similar fragment pairs and their corresponding rotation matrix set $\mathbb{R} = \{R_1, R_2, ..., R_L\}$, we define the distance between two rotation matrices R_i and R_j, $D_F(R_i, R_j)$, as

$$D_F(R_i, R_j) = \|R_i - R_j\|_F,$$

where $\|\cdot\|_F$ is frobenius norm. The center R_0 of rotation matrix set $\mathbb{R}$ is defined as

$$R_0 = \sum_{i=1}^{L} R_i/L,$$

if

$$\|R_i - R_0\|_F \leq C_3, i = 1, 2, \cdots, L.$$

where C_3 is a constant (e.g. 0.4), we call such set of similar fragment pairs are subjected to spacial compatibility. In order to find out such groups, we apply a grouping algorithm (Table 1), which is based on the demographic clustering technique of data mining (Cabena et al., 1997). Intuitively, a set of fragment pairs subjected to spacial compatibility whose fragments' total length (repeated was counted once) is the largest is expected, we call it the optimal group.

3 Experiment Analysis

3.1 Data Collection

In order to analyze various cases and compare with other methods easily, a data set including 30 proteins belonging to five different families randomly selected from Alpha, Beta, Alpha and Beta(alpha/beta) and Multi-domain proteins (alpha and beta) classes is selected. Proteins representing serpine family: 1ATTa, 7APIa, 1AZXi, 2ACHa, 2ANTl, 1OVAa whose number of C_α are 420, 339, 417, 337, 398, 385 respectively which were used in Leluk et al.(2003) are taken as main examples for analyzing.

The proposed method was compared with the famous DALI program (Holm and Sander, 1993) used for the rigid structure alignment. In our experiment we always take $S_5(i, j)$ as the similarity score, the parameter C_2 which comes from practice is always set to 3.5 Åand C_3 is always set to 0.40 experientially.

3.2 Structural Similarity

We use DV_4 as a parameter profile, calculated according to the procedure presented in **Methods** characterize the structure of 1ATTa and 2ACHa. The result is shown in Figure 1. An interesting thing is that two regions of 1ATTa profile is higher than 12, but this is impossible for consecutive protein chain. So we can infer that the two regions are disconnected for 1ATTa which can be verified by inspecting its data and figuring its protein backbone of C_αs. We can easily find

Table 1: Grouping Algorithm based on demographic clustering technique of data mining

Grouping Algorithm:
Input: A set of rotation matrices and a distance measure C_3.
Output: A set of groups into which rotation matrices have been divided, where every rotation matrices in a group is within the distance C_3 of the group center.
Begin:
Step 0: Given a set $\mathbb{R}$ of rotation matrices.
Step 1: Take a rotation matrix R_1 from $\mathbb{R}$ Randomly, create group 1, with center $R_{01} = R_1$, set $N_1 = 1$.
Step 2: While ($\mathbb{R}$ is not empty)
{
a. Take R_p from $\mathbb{R}$.
b. Compute the distances d_j between R_p and existing group center R_{0j} (suppose we have k groups now, then $1 \leq j \leq k$).
c. Suppose $j_{min} = arg\min d_j$ is the minimum. if $d_{j_{min}} = D_F(R_{0j_{min}}, R_p) > C_3$, then create a new group k+1, with center $R_{0k+1} = R_p$, set $N_{k+1} = 1$. Else
1. Insert R_p into group j_{min}, $N_{j_{min}} = N_{j_{min}} + 1$.
2. Compute the new center $R'_{0j_{min}}$ of group j_{min}.
3. For $i = 1, 2, \cdots, N_{j_{min}}$
{
i. Re-compute the distance $D_F(R_{j_{min},i}, R'_{0j_{min}})$ between the rotation matrix $R_{j_{min},i}$ in group j_{min} and the new group center $R'_{0j_{min}}$.
ii. If $D_F(R_{j_{min},i}, R'_{0j_{min}}) > C_3$, put $R_{j_{min},i}$ into the set $\mathbb{R}$, $N_{j_{min}} = N_{j_{min}} - 1$, go to 2.
}
}
Step 3: For each group, re-calculate the distance between the contained rotation matrices and all of the group centers. If there is any Rotation Matrix that has a shorter distance with another group center than with its own group center, move it to the other group where the distance is shorter. If there are no such rotation matrix, go to **END**.
Step 4: Re-compute all the group centers. If any rotation matrix is no longer within distance C_3 of the center of its group, put it into the set $\mathbb{R}$. If $\mathbb{R}$ is not empty, go to step 3, else go to step 4.
END

that their similarity were reflected well. Continuous high values of DV_4 are present in β-sheet structure segment and continuous low values are present in α-helics structure segment commonly. In general secondary structure always corresponds to distinct feature. The similarity score dot-matrix for two similar protein 1ATTa and 2ACHa is presented in Figure 2 (where $C_1 = 0.12$). The inter-molecular similarity is very easily distinguishable. A rough diagonal can be seen in the map, showing the overall inter-molecular structure similarity.

Different values of parameter C_1 will produce different matched fragments. Figure 3 pictures the similar fragment pairs subjected to spacial compatibility under different parameter C_1 of two pair proteins: 1ATTa and 2ACHa, 1LW6 and 1SUC respectively. Table 2 shows the local similar fragment pairs of 1ATTa and 2ACHa with various C_1 values which satisfy the spacial compatibility. More local similar fragments were given out such as {[179–295, 224–340], 2.98} and {[1–60, 42–101], 2.26}, but DALI program can not obtain these results.

Table 2: Proteins and their respective families

FamililyName	Proteins selected from the family (PDB id)					
Serpins	1ATTa,	7APIa,	1AZXi,	2ACHa,	2ANTl,	1OVAa
Flavodoxin-related	1C7E,	1C7F,	1J9G,	1J8Q,	1J9E,	1AZL
Monodomain-cytochrome c	1B7V,	1K3G,	1K3H,	1KIB,	1N9C,	1CED
V set domains	1BJM,	2FB4,	2IG2,	3BJL,	4BJL,	1MCOl
Subtilases	1SEL,	1OYV,	1SCJ,	1CSE,	1SUC,	1LW6

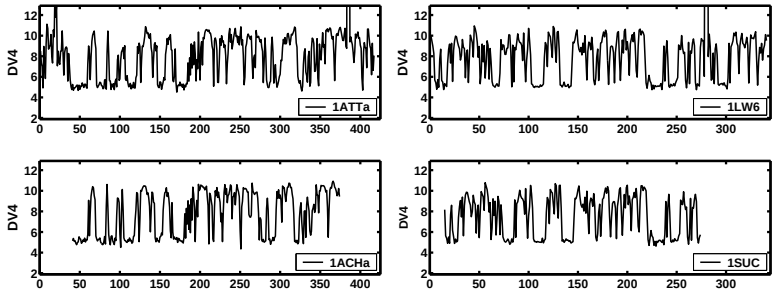

Figure 1: Profile of protein structure: the left is DV_4 profile of protein 1ATTa and 2ACHa from serpins family and likewise the right is DV_4 profile of protein 1LW6 and 1SUC from subtilases family.

3.3 Comparative Experiment

In order to well verify the effectivemess of overall similarity search of our method, a comparative experiment has been done with DALI program by analyzing serpine family members. The protein 1ATTa was taken as the template protein structure to which others were compared pair-wise. When C_1 take different value, the similar fragment pairs is some different (Figure 3). We take the combination of the results of our method when $C_1 = 0.12$ and $C_1 = 0.15$. This is rational, because their respective optimal group can form a larger group subjected to spacial compatibility. Figure 4 indicates similar fragments in some serpine family member and 1ATTa by using our method and DALI program respectively. In the comparative experiment the instance of similarity can be depicted alike and it is so efficient. The covered range of our method shows the well similarity between proteins 1ATTa and other serpine family members. In another way, the rotation matrices of two methods are quite close (the rotation matrices of our method is the center of the optimal group which satisfy spacial compatibility under some parameter C_1). For example, the rotation matrices of 2ACHa and 1ATTa are $\begin{pmatrix} -0.3761 & 0.3825 & 0.8379 \\ -0.2013 & -0.9198 & 0.3269 \\ 0.9002 & -0.0466 & 0.4250 \end{pmatrix}$ and $\begin{pmatrix} -0.3513 & 0.3666 & 0.8615 \\ -0.2033 & -0.9281 & 0.3121 \\ 0.9139 & -0.0655 & 0.4006 \end{pmatrix}$ respectively. This all mean the conception of spacial compatibility can well represent the overall structural similarity in proteins.

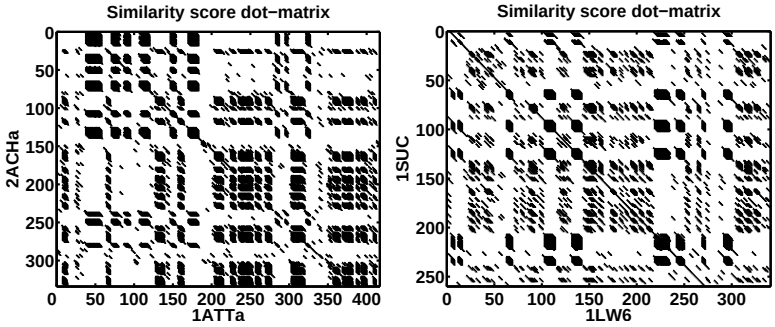

Figure 2: Similarity score dot-matrix: the left is similarity score dot-matrix of 1ATTa and 2ACHa and the right is similarity score dot-matrix of 1LW6 and 1SUC.

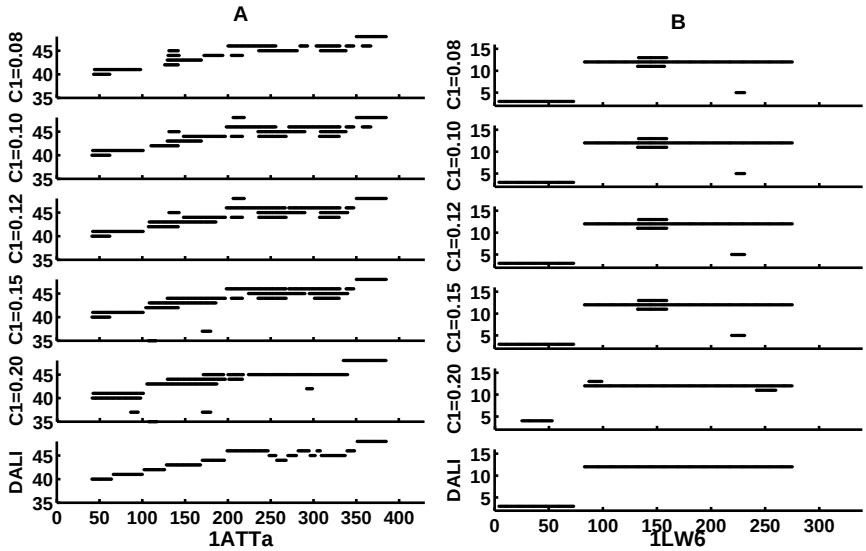

Figure 3: A: Similar fragments between 2ACHa and 1ATTa by using our method with five different parameters C_1=0.08, 0.10, 0.12, 0.15, 0.20 and the DALI program. 1ATTa structure is taken as template structure for comparison in two methods. Horizontal axis represents sequence of the template protein molecule 1ATTa. Vertical axis represents sequence of protein 2ACHa in relative numbers versus sequence of target protein molecule. Line parallel to horizontal axis represents the situation when two fragments are similar without any shift in their sequences. B: similarly for 1SUC and 1LW6

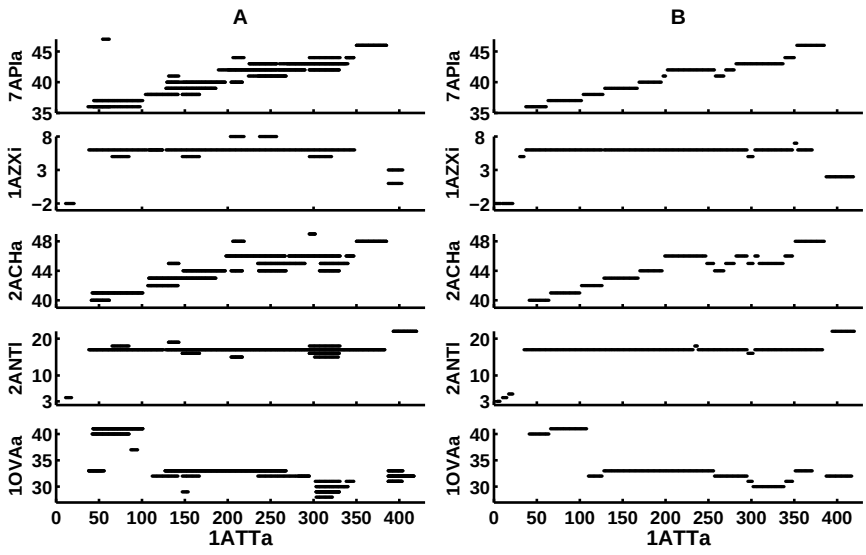

Figure 4: Structure similarity in serpine family members:(A) Our method: the result is the combination of C_1=0.12 and 0.15. (B) DALI program. 1ATTa protein structure is taken as template structure for comparison in two methods. Horizontal axis represents sequence of template protein molecule 1ATTa. Vertical axis represents sequence of compared protein chain in relative numbers versus sequence of target protein molecule. Line parallel to horizontal axis represents the situation when two fragments are similar without any shift in their sequences.

4 Discussion and Future work

The experimental analysis shows that the proposed geometric distance parameters can represent the protein well and can be used to search for structural similarity in proteins, although the criteria selected for this search is simple and can be implemented easily. The profiles of DV_4 and DV_5 really express the visual characteristics of the structure chain in protein. The method can quickly find local similar fragments of proteins and can generate data that provide detailed information about regions of local similarity in proteins structures. At the same time, the notion of spacial compatibility not only enriches the annotation of overall structural similarity, but also provides more detailed similar local fragments(we can call it "consensus structure") of overall structure similarity which is significant because different structure alignment methods produce various alignment results.

An effective profile of proteins is given out. DV_4 and DV_5 profile comparison between the target molecule and the predicted form of protein structure could be useful in the CASP project (Hubbard, 1999). Random coiled fragments, which usually are difficult to identify, can also be easily analyzed uniformly with secondary structure fragments when the proposed parameters are used. It also can

be employed as a tool of classification of proteins and be able to identify sequence structure patterns that would represent local structural motifs. Further, it can be developed into structure alignment method or provide initiative solution for some unstable structure alignment method such as the method designed by Chen et al(2004). The most important is that it can be designed for large-scale and more detailed similarity search in database.

References

[1] Besl, P. J., and McKay, N. D. (1992) A method for registration of 3-D shapes. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 14, 239-256.

[2] Cabena, P., Hadjnian, Stadler, Verhees and Zanasi (1997) Discovering Data Mining-From Concept to Implementation. Prentice Hall PTR.

[3] Carugo, O. and Pongor, S. (2002) Protein fold similarity estimated by a probabilistic approach based on C_α-C_α distance comparison. *J. Mol. Biol.*, 315, 887-898.

[4] Chen, L. N., Zhou, T. S. and Tang, Y. (2004) Protein structure alignment by deterministic annealing. *Bioinformatics*, 20, 1-12.

[5] Chew, L. P. and Kedem, K. (2002) Finding consensus shape for a protein family. *18th ACM Symposium on Computational Geometry*, ACM Press, Barcelona, Spain.

[6] Cui, C. Y., Wang, D. H. and Shi, J. Y. (2004) Comparing 3-D protein structures similarity by using fractal features. *Proceeding of the 2004 IEEE Computational Systems Bioinformatics Conference (CSB 2004).*

[7] Dua, S. and Kandiraju, N. (2004) A novel computational framework for structural classification of proteins using local geometric parameter matching. *Proceeding of the 2004 IEEE Computational Systems Bioinformatics Conference (CSB 2004).*

[8] Flocco, M. M. and Mowbray, S. L. (1995) C_α based torsion angles: A simple tool to analyze protein conformational changes. *Protein Sci*, 12, 2118-2122.

[9] Guda, C., Scheeff, E. D., Bourne, P. E. and Shindyalow, I. N. (2001) A new algorithm for the alignment of multiple protein structure using Monte Carlo optimization. *Proceedings of Pacific Symposia on Biocomputing*, 6, 25-286.

[10] Holm, L. and Sander, C. (1993) Protein-structure comparison by alignment of distance matrices. *J. Mol. Biol.*, 233, 123-138.

[11] Horn, B. K. P. (1987) Closed form solution of absolute orientation using unit quaternions. *J. Optical Society A*, 4(4), 629-642.

[12] Hubbard, T. J. P. (1999) RMS/Coverage Graphs: a quantitative method for comparing three-dimensional protein structure prediction. *Proteins*, Suppl. 3, 15-21.

[13] Jung, J. and Lee, B. (2000) Protein structure alignment using environmental profiles. *Protein Eng.*, 13, 535-543.

[14] Kabsch, W. (1978) A solution for the best rotation to relate two sets of vectors. *Acta Crystal A*, 32, 922-923.

[15] Leluk, J., Konieczny, L. and Roterman, I. (2003) Search for structural similarity in proteins. *Bioinformatics*, 19, 117-124.

[16] Levitt, M (1976) A simplified representation of protein conformations for rapid simulation of protein folding. *J. Mol. Biol.* 104, 59-107.

[17] Ortiz, A. R., Kolinski, A., Rotkiewicz, P., Ilkowski, B. and Skolnick, J. (1999) Ab initio folding of proteins using restraints derived from evolutionary information. *Proteins*, Suppl. 3, 177-185.

[18] Ramachandran, G. N., Ramakrishnan, C. and Sasisekharan, V. Conformation of polypeptides and proteins. *J. Mol. Biol.*, 7, 95-99, 1963.

[19] Schwartz, J. T. and Sharir, M. (1987) Identification of partially obscured objects in two dimensions by matching of noisy characteristic curves. *Int. J. Robotics Res.*, 6, 29-44.

[20] Shindyalov, I. N. and Bourne, P. E. (1998) Protein structure alignment by incremental combinatorial extension (CE) of the optimal path. *Protein. Eng.*, 11, 739-747.

Table 3: The Fragment pairs of 1ATTa and 2ACHa subjected to spacial compatibility. The table has shown the result of five different parameter C_1. {[S1–E1, S2–E2], R} represent a pair similar fragments, where S1, S2 are the start atoms' position of fragments in protein 2ACHa and 1ATTa, E1, E2 are the end atoms' position and R represents their RMSD.

C_1	**Matched Rigid Fragment Pairs**		
	{[3– 22, 43– 62], 0.63}	{[3– 57, 44– 98], 2.13}	{[84–100, 126–142], 2.45}
	{[86–126, 129–169], 1.18}	{[86– 99, 130–143], 2.04}	{[128–150, 172–194], 0.69}
0.08	{[154–210, 200–256], 1.95}	{[160–173, 204–217], 1.77}	{[191–236, 236–281], 2.75}
	{[257–285, 303–331], 2.66}	{[263–293, 308–338], 1.16}	{[302–337, 350–385], 1.56}
	{[1– 11, 52– 62], 0.61}	{[86– 97, 131–142], 1.25}	{[238–247, 284–293], 0.67}
	{[292–301, 338–347], 1.46}	{[311–321, 357–367], 1.03}	
	{[1– 22, 41– 62], 0.67}	{[1– 60, 42–101], 2.26}	{[68–100, 110–142], 2.17}
	{[86–126, 129–169], 1.18}	{[104–153, 148–197], 2.60}	{[152–210, 198–256], 2.04}
0.10	{[158–171, 206–219], 1.56}	{[160–173, 204–217], 1.77}	{[190–245, 235–290], 3.25}
	{[192–224, 236–268], 2.76}	{[225–285, 271–331], 2.79}	{[263–286, 307–330], 2.93}
	{[263–293, 308–338], 1.16}	{[302–337, 350–385], 1.56}	{[1– 11, 52– 62], 0.61}
	{[86– 98, 131–143], 1.55}	{[246–253, 295–302], 0.96}	{[292–301, 338–347], 1.46}
	{[311–321, 357–367], 1.03}		
	{[1– 22, 41– 62], 0.67}	{[1– 60, 42–101], 2.26}	{[65–100, 107–142], 2.13}
	{[65–143, 108–186], 2.83}	{[104–153, 148–197], 2.60}	{[152–221, 198–267], 3.20}
0.12	{[158–171, 206–219], 1.56}	{[160–173, 204–217], 1.77}	{[190–245, 235–290], 3.25}
	{[192–224, 236–268], 2.76}	{[225–285, 271–331], 2.79}	{[263–286, 307–330], 2.93}
	{[263–295, 308–340], 1.40}	{[302–337, 350–385], 1.56}	{[1– 11, 52– 62], 0.61}
	{[86– 98, 131–143], 1.55}	{[246–253, 295–302], 0.96}	{[292–301, 338–347], 1.46}
	{[1– 22, 41– 62], 0.67}	{[1– 60, 42–101], 2.26}	{[62–100, 104–142], 2.50}
0.15	{[65–143, 108–186], 2.83}	{[85–153, 129–197], 2.95}	{[152–222, 198–268], 3.25}
	{[160–173, 204–217], 1.77}	{[179–245, 224–290], 3.46}	{[191–224, 235–268], 2.79}
	{[225–285, 271–331], 2.79}	{[250–295, 295–340], 1.69}	{[257–286, 301–330], 3.12}
	{[302–337, 350–385], 1.56}	{[1– 12, 52– 63], 0.92}	{[72– 81, 107–116], 1.15}
	{[133–143, 170–180], 0.59}	{[292–301, 338–347], 1.46}	
	{[1– 22, 41– 62], 0.67}	{[1– 60, 42–101], 2.26}	{[62–100, 104–142], 2.50}
0.20	{[65–143, 108–186], 2.83}	{[85–153, 129–197], 2.95}	{[152–222, 198–268], 3.25}
	{[160–173, 204–217], 1.77}	{[179–245, 224–290], 3.46}	{[191–224, 235–268], 2.79}
	{[225–285, 271–331], 2.79}	{[250–295, 295–340], 1.69}	{[257–286, 301–330], 3.12}
	{[302–337, 350–385], 1.56}	{[1– 12, 52– 63], 0.92}	{[72– 81, 107–116], 1.15}
	{[133–143, 170–180], 0.59}	{[292–301, 338–347], 1.46}	