Comprehensive Analysis of Natural Disulfide Bridges in Proteins

Maryam Khoshnejat[a], Seyed Ahahriar Arab[a]*

[a] Department of Biophysics, Faculty of Biological Science, Tarbiat Modares University, Tehran, Iran

Abstract. Disulfide design is one of several approaches in protein engineering with the wide range of applications like as the study of protein conformation, stability, dynamics and function. Analysis of natural disulfide bonds will provide the basis of models for disulfide engineering and can extend insights to explore potential disulfide bond with higher accuracy. In this effort, analysis has carried out on 8868 non-redundant disulfide bonds. Several important parameters such as three distance parameters, surface accessibility, secondary structures, loop length, flexibility, geometry, dihedral strain energy distribution and conformations have been characterized. Moreover, the dataset was classified according to the 20 types of disulfide configurations and study was performed separately on each group. In brief, analysis revealed that natural disulfide bonds typically occur inaccessible from the solvent with low dihedral strain energy distribution and frequently emerge as –RHSpiral configuration. In addition, the coil is the most preferred secondary structure of cysteine residues involving in a disulfide bond and Coil-sheet peptide fragments have the most connection through disulfide bonds in compare to other 5 possible fragments.

Keywords: Disulfide bond; Dihedral strain energy; Configuration; Analysis; Geometry.

1. Introduction

The single covalent linkage derived by the coupling of two thiol (–SH) groups is called disulfide bond, S-S bond, or disulfide bridge. In protein structures, the thiol groups in two cysteine residues can form disulfide bond by oxidation of the sulfur atoms. Therefore, Disulfide bond has two states: (1) oxidized and (2) reduced form, reactions which change disulfide state to the reduced form can disrupt the linkage [1, 2]. Disulfide bond have been considered in the three different following types: structural, catalytic and allosteric. Structural bonds play an important role in protein folding and stability [3] whereas catalytic bonds mediate thiol-disulfide interchange reactions and allosteric bonds can trigger a protein conformational change and control its function through their formations and cleavages [4].

The geometry of disulfide bond can be described by its five $\chi1$, $\chi2$, $\chi3$, $\chi2'$ and $\chi1'$ dihedral torsion angles. The $\chi1$ torsion angle defined by the N–Cα–Cβ–Sγ bond and the $\chi2$ angle determined by the Cα-Cβ-Sγ-Sγ' bond. Rotation of the Cβ atoms around the S–S bond has been represented as $\chi3$ dihedral angle. In addition, the Cα-Cβ-Sγ and Cβ-Sγ-Sγ' are two valuable angles which generally found near 1140 and 1050, respectively [5]. There are 20 types of disulfide configurations based on the sign of five torsion angles [4, 6]. Previous studies have been shown that from the different possible configurations, allosteric disulfides emerge in the –RHStaple, –LHHook and -/+RHHook configurations and most of the catalytic bonds have been observed in the +/-RHHook group [4, 6, 7].

Disulfide bonds play an important role in protein structure, stability, dynamics and function [8, 9]. Experiments suggest that disulfide bridges can enhance stability of the native structure by decreasing the conformational entropy and enhancing the free energy of the denatured state [5, 10]. In protein engineering efforts disulfide bridges are appealing candidates to enhance protein stability and to assist in investigation of protein kinetics, conformation and function. Study of natural sulfides can be helpful in directed disulfide engineering as well as improvement of model building of protein structures. In the past, several analyses have been performed for characterization of protein disulfide bonds [11-17]. Dani et al. have studied the distribution of bond dihedral angles, distances, residue depth and surface accessibility on the 172 proteins which representing 730 disulfides. They illustrated that disulfides commonly occurred at the buried sites of proteins

* Corresponding author.
E-mail address: sh.arab@modares.ac.ir

with the peaks at ±90 for χ3 and the non-bonded distances 3.9-6.7 Å and 3.5-4.5 Å for Cα-Cα and Cβ-Cβ, respectively [17]. In the recent work, analysis on 1505 native disulfide bonds in 331 non-homologous proteins have been revealed that χ3 peaks at −87 and +97 degrees. In addition, the energy function based on the χ3, χ1 dihedrals and the two Cα–Cβ–Sγ angles has been used for calculation of energy distribution and an energy value less than 2.2 Kcal/mol has been observed for 90 percentages of native disulfide bonds [15].

As the protein structures database is expanding, the knowledge about the protein can increase. Analysis on protein structures lead to the better understanding of the relationship between protein structure and function thus improving the success in rational design of de novo proteins. It is essential to update our concepts about disulfides by analyzing new structures. In this attempt, a comprehensive analysis has been performed on the 8868 unique disulfide bonds and roughly all information has been extracted.

2. Experimental procedures

The analysis was performed on the non-redundant dataset of cysteine residues which involved in the disulfide bridge. Protein structures were selected from the 2014 release of Protein Data Bank [18] using three following criteria; (1) the structure should contains protein without any DNA and RNA molecules, (2) must have at least one disulfide bond and (3) sequence identity between structures should be less than 40 percentages. This yielded 3,686 proteins with 14,664 disulfide bonds. The filtering has been intended to remove redundancy; for this reason, only one set of those similar intra-chain disulfide bonds which repeat in all chains in homo-multimeric proteins have been kept and the other ones were eliminated, thus 8868 disulfide bonds represent the final non-redundant dataset which consists of 338 inter-chain disulfides. Distribution of distance parameters, secondary structure, surface accessibility, flexibility, sequential distance, angles, dihedrals, dihedral strain energies and distribution of all 20 types of configuration have been analyzed.

Surface accessibility (ACC) is defined as the accessible surface area (ASA) of the residue in native protein divided by the ASA of the residue in an extended tripeptide Ala-X-Ala conformation. Accessible surface area has been calculated by DSSP [19] software. The extended state ASA values with the units in A2 are as fallow: Ala-110.2; Arg-229.0; Asn-146.4; Asp-144.1; Cys-140.4; Gln-178.6; Glu-174.7; Gly-78.7; His-181.9; Ile-185.0; Lys-205.7; Leu-183.1; Met-200.1; Pro-141.9; Phe-200.7; Ser-117.2; Thr-138.7; Trp-240.5; Tyr-213.7; Val-153.7 [20].

Secondary structures have been determined by DSSP program, which assigns eight-state secondary structures to each residue. These eight-state structures have been classified to three structures: Helix (G, H and I), Coil (C, T, and S) and sheet (E and B).

The flexibility has been expressed with B-factor value in X-ray structures; since experimental conditions, crystal contacts, and the refinement procedures can affect the B-factor values, they have to be normalized to compensate the differences among structures for reasonable comparison [21, 22]. Three different criteria as B-norm of backbone, side chains and Cα atoms were examined to assess the flexibility. B-factors of atoms corresponding to each criterion for each structure were extracted from PDB files and normalized using Eq. (1) where B is the actual B-factor, μ and σ is the average and the standard deviation of B-factor for corresponding atoms in the particular PDB structure, respectively.

$$B\_norm = \frac{(B-\mu)}{\sigma} \quad (1)$$

The following formula has been used to calculate the dihedral strain energy (DSE) of each disulfide according to the five χ dihedral torsion angles which has been shown that it can provide good estimates into the quantity of strain in a disulfide bond [4, 23].

$$\text{DSE (kJ/mol)} = 8.37\,(1 + \cos 3\chi1) + 8.37\,(1 + \cos 3\chi1') + 4.18\,(1 + \cos 3\chi2) + 4.18\,(1 + \cos 3\chi2') + 14.64(1 + \cos 2\chi3) + 2.51\,(1 + \cos 3\chi3) \quad (2)$$

3. Results

At the first, analysis was performed on the whole dataset. The distribution of three distance parameters are shown in Fig. 1 which reflects distances Ciα_Cjα, Ciβ_Cjβ and centroid(i)_centroid(j) (centroid of side-chain atoms) where i and j are cysteine residues that are linked by a disulfide bridge. Ciα_Cjα distances have greater range with broader occurrence than the other two distance parameters. Ciα_Cjα distances vary from 3.4 to 8.8 Å with a peak at 6 Å. The highest distance has been observed in PDB-ID 1PS1 (inter-chain disulfide bond between cysteine 272 from chain A and cysteine 272 from chain B) and 3SZG (intra-chain disulfide bond between cysteine 17 and cysteine 40 from chain A) which both have –LHHook configuration. Approximately, about 90 percentages of Ciβ_Cjβ distances in disulfide bonds are within 3.5-4.5 Å and all centroid(i)_centroid(j) distances roughly lie between 2.3-3.3 Å.

The number of residues enclosed by the disulfide bond has been considered as the loop length. In 28 cases minimum loop length with 2 residues has been detected (e.g. PDB-ID: 1AHJ, 2BLF, 3LVT, 4MY2). The Hook is the dominant configuration in this group. The maximum loop length has been observed in 3BVU between cysteine 31 and 1032, which has -LHSpiral configuration with low energy (4.76 KJ/mol). 88.05 percentages of intra-chain disulfides have been distinguished with loop length less than 100 residues; the distribution has been shown in Fig. 2.

Average accessibility of two cysteine residues, which involved in the disulfide bridge has been assessed. Analysis shows that 93.35 percentages of naturally occurring disulfide bonds have ACC less than 36 percentages thus are located at buried and intermediate sites of proteins (Fig. 3). Three intra-chain disulfide bridges with ACC more than 80 percentages have been determined (PDB-ID: 1I8E, 1I93, 1MPV) with

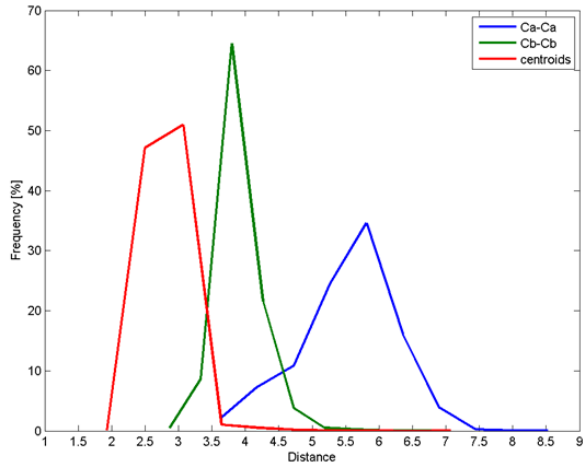

Fig. 1. Percentage distribution of distance parameters (Å), Distance between the two Cα atoms, the two Cβ atoms and center of side-chain atoms were calculated for each disulfide.

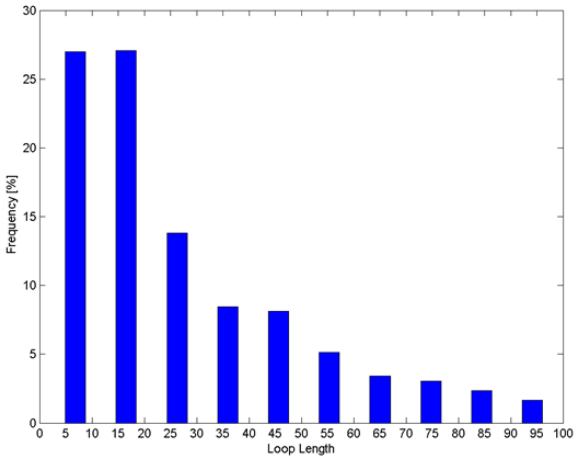

Fig. 2. Percentage distribution of loop length from those disulfides with loop length less than 100, The number of residues enclosed by the disulfide bond has been considered as the loop length. 88 percentage of disulfides show loop length less than 100 residues.

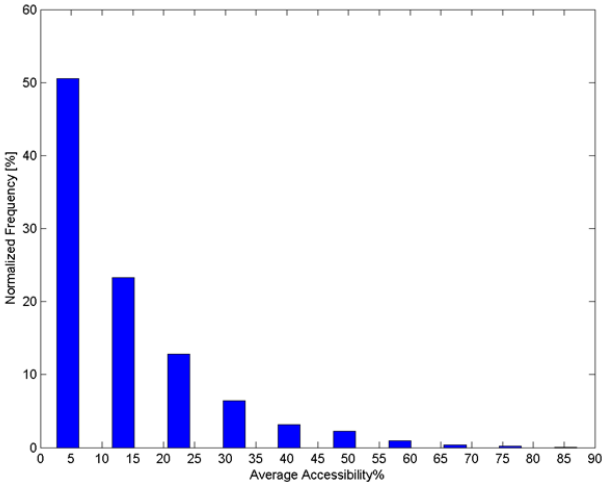

Fig. 3. Average accessibilities of two cysteine residues connected by disulfide bond, the average value of a pair has been considered for a disulfide bond.

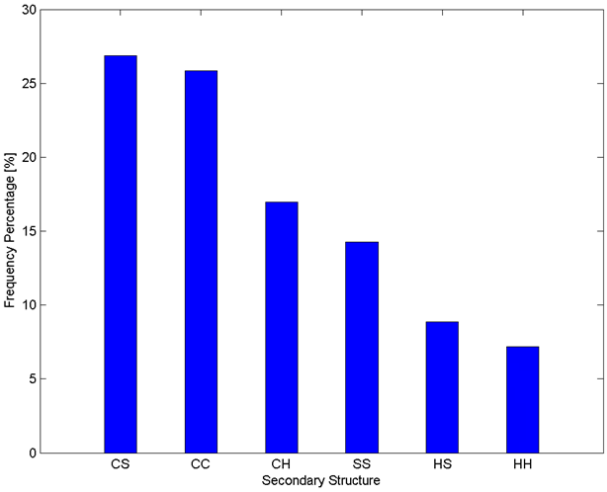

Fig. 4. Secondary structure frequency of two cysteine residues involving in the disulfide bond, CC: coil-coil, HH: helix-helix, SS: sheet-sheet, CH: coil-helix, HS: helix-sheet, CS: coil-sheet, e.g. CS means that in the disulfide bond one of two cysteine residues has secondary structure as coil and the other one has secondary structure as sheet.

relatively high strain energy, –RHStaple and –RHHook configurations and all linked coil-coil fragments to each other.

Secondary structures of cysteine residues in the dataset have been studied. It is clearly observed that the coil is the most common secondary structure of cysteine residues linking by the disulfide bond. Frequency percentages 20.09, 47.78, 32.13 have been observed for helix, coil and sheet, respectively. In addition, pair mode secondary structures of cysteine residues in the disulfide bonds have been studied to comprehend which peptide fragments connected through disulfide bond? The results are shown in Fig. 4. As is clear, disulfide bonds generally link coil-sheet and coil-coil fragments to each other. Helix-helix fragments have the lowest incidence in the dataset.

Since thus far, several criteria such as backbone, side-chain and Cα atoms have been applied for flexibility estimation [15, 17, 22, 24] in this analysis all three cases were examined and average value for each pair has been calculated (Fig. 5).

In addition, the average B-factor of the backbone and β-carbon atoms [25] has been assessed. The remarkable similarity with backbone curve in Figure 5 has been observed, so we refrain from showing it. It has been illustrated that the peak corresponding to each of the three curves has been diagnosed below zero value, which commonly assigned to the rigid region [21, 22]. It was surprising that in average, side-chain atoms which participating in disulfide connection has been found with lower flexibility than the backbone atoms; therefore, disulfide rolling as the strong bridge between

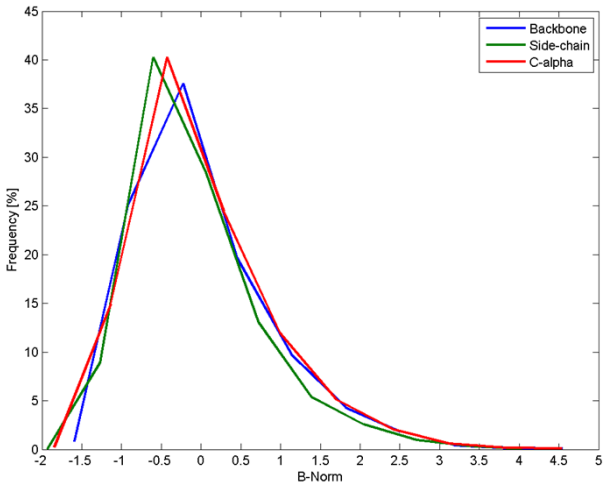

Fig. 5. Frequency distribution of normalized B-factor in three different sets of atoms. Flexibility of backbone, side-chain and Cα atoms has been assessed using normalized B-factor.

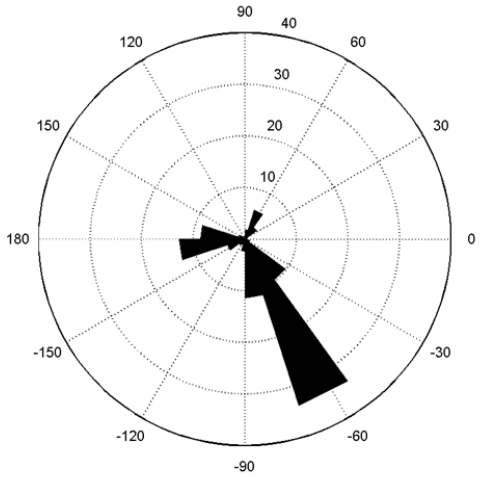

Fig. 6. Percentage distribution of χ_1 dihedral angle.

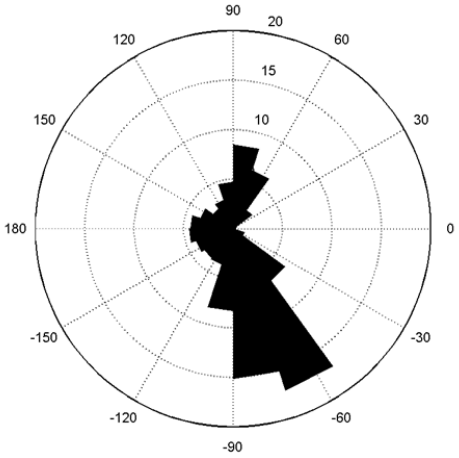

Fig. 7. Percentage distribution of χ_2 dihedral angle.

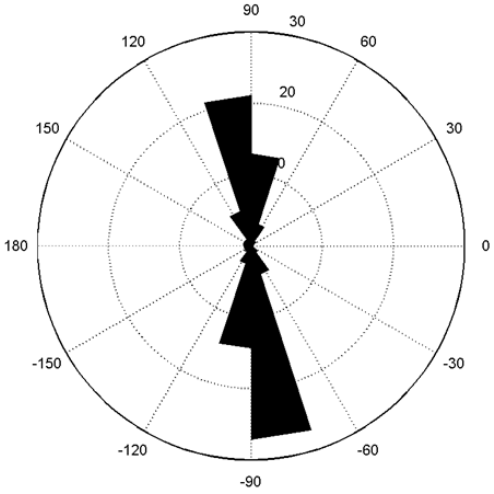

Fig. 8. Percentage distribution of χ_3 dihedral angle.

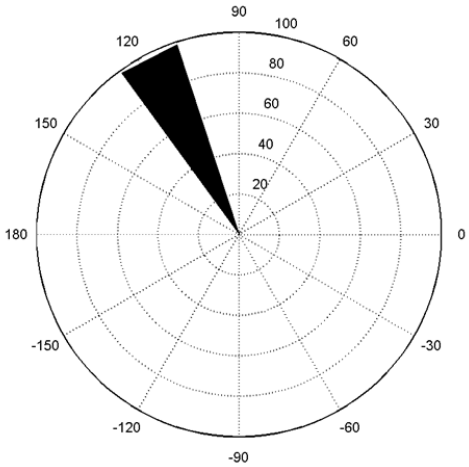

Fig. 9. Percentage distribution of Cα-Cβ-Sγ angle.

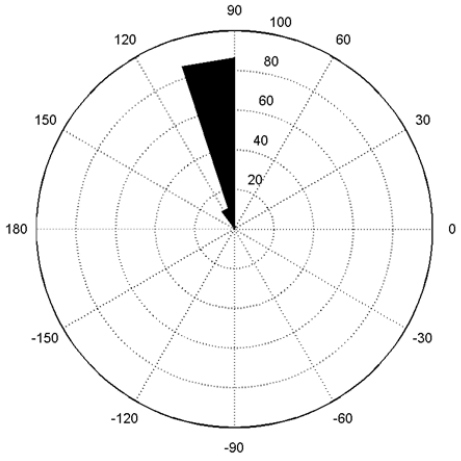

Fig. 10. Percentage distribution of Cα-Cβ-Sγ' angle.

peptide fragments and can strongly affect the local flexibility.

Distributions of χ1 and χ2 dihedral angles are illustrated in Fig. 6 and 7, respectively; χ1 has great majority at -65° whereas broader distribution has been observed for χ2 angles. Fig. 8 shows repartition of χ3 dihedral angles; the values are distributed around 90° with two peaks at +100° and -80° for right and left handed disulfide bonds, respectively. Left handed χ3 dihedral angles with 53 percentages have slightly higher frequency than the right-handed ones. In addition, Fig. 9 and 10 illustrate the Cα-Cβ-Sγ and Cβ- Sγ - Sγ' angles which are consistent with previous reports [5].

Dihedral strain energies have been calculated with respect to the five dihedral angles. Histogram of DSE (Fig. 11) represents the variation between 0-80 KJ/mol. DSE values less than 25 KJ/mol have been observed in 83.64 percentages of natural disulfide bridges. 48 disulfides have DSE more than 60 KJ/mol which respectively 26, 13 and 9 Hook, Spiral and Staple configurations have been found and from them 12 cases with Helix secondary structure have been distinguished.

According to the sign of five dihedral angles, distribution of disulfide configurations has been evaluated (Fig. 12). Approximately, configurations with the low energies are devoted to the higher frequency. –LHSpiral and +RHStaple is assigned to the highest frequent configuration with the lowest average DSE and the lowest frequent configuration with the highest average DSE, respectively. Commonly, average strain energy in right handed disulfides (17.51 KJ/mol) is slightly higher than that in left handed disulfides (14.82 KJ/mol).

After analysis of whole dataset, disulfides were further categorized into the 20 configuration groups and then surveys were conducted separately on each category. χ3 typically

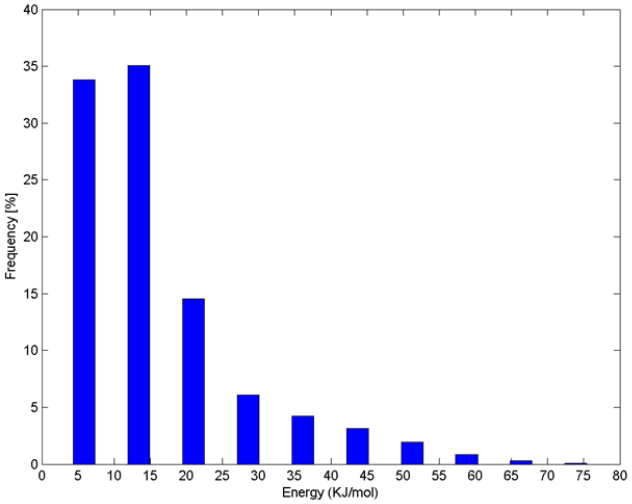

Fig. 11. Percentage distribution of dihedral strain energy in disulfide bonds.

peaks around ±90° whereas different peaks for χ2 and χ1 have been observed in different conformation. The narrow distribution of Ciβ_Cjβ distances and DSE both have been found in +LHSpiral. In average, higher distances correlated to the higher dihedral strain energies. Generally, Staple configurations have emerged with the highest average dihedral energy, Hook configurations have been observed with mediocre dihedral energy and Spiral configurations allocated to the lowest levels. In addition, Spiral configurations commonly have mediated the connection of the coil-coil fragments. Some data are given in Table 1. The results in details are available

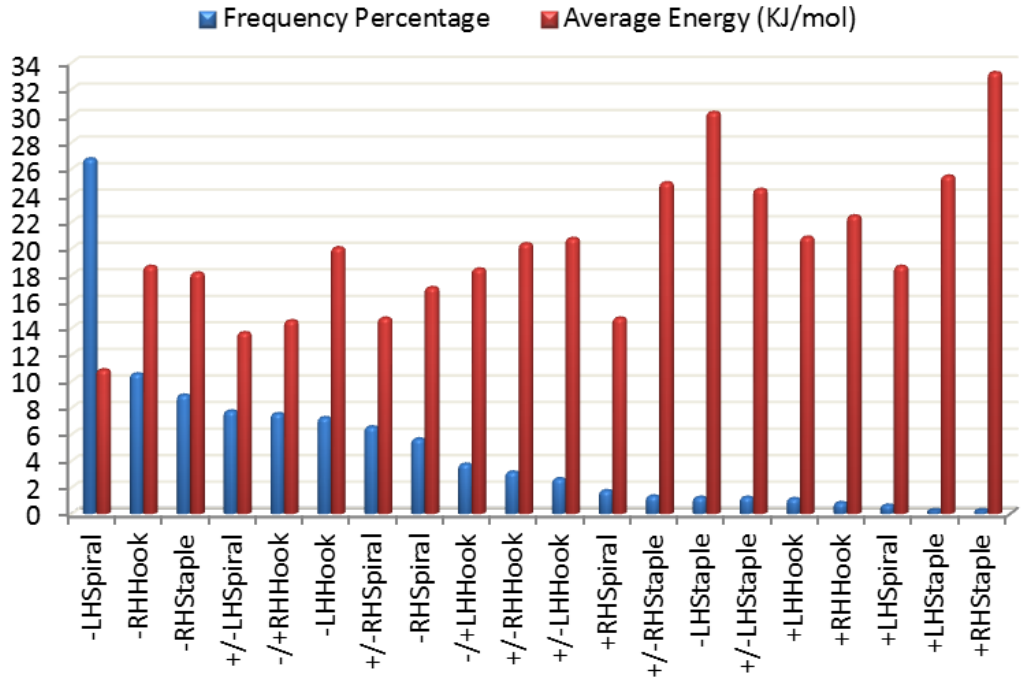

Fig. 12. Frequency percentage and Average dihedral strain energy in each specific disulfide configurations, 8868 disulfides have been categorized into the 20 configuration groups and Frequency percentage and Average dihedral strain energy in each group has been calculated.

through http://bioinf.modares.ac.ir/software/mpdb/analysis.

4. Discussion

The non-redundant dataset consisting of the 8868 disulfide bonds has been used for analysis. Several different parameters such as distance (Ciβ_Cjβ, Ciα_Cjα and centroids distance), surface accessibility, secondary structure (single and pair modes), flexibility, loop length, dihedrals, angles, energy distribution and configurations have been used for analysis. Moreover, dataset were categorized based on the 20 possible configurations and analysis was carried out separately on each specific group.

The distribution of three different distance parameters for native disulfides has been evaluated. As the results show, Ciβ_Cjβ and centroids distances have narrow range comparing to Ciα_Cjα, thus using these two distance parameters can be more confident than Ciα_Cjα to find potential disulfide bonds. In addition, far distances leading to the inappropriate dihedrals and can conduct more strain energy to the structure; this could be why the high DSE commonly has been found along with the high Ciβ_Cjβ distances.

Dani et al. have been proposed that in disulfide engineering, stabilizing mutations usually occur in regions have a loop length greater than 25. This item has been explored in natural disulfide bonds. It has been observed that 61 percentages of disulfides have loop length less than 30 residues; so, no particular consequence has been determined in this case.

Accessibility analysis verifies that natural disulfides mostly occurred at buried and intermediate sites of proteins, which can hide them from the solvent to limit unfavorable interactions [12, 17]. Oxidative conditions in the environment is critical for maintenance of the disulfide bond, therefore locating at the protein core makes it safe for disulfide bond to escape from the attacks of reductive functional groups in solvent.

Secondary structure has been studied to comprehend the preferred structure for disulfide bonding. It is obviously that the coil is the most prevalent secondary structure due to the conformational freedom, which gives the sulfur atoms opportunity to come close together without additional strain and form the disulfide bond [11]. Helix is the most regular structure, which makes it difficult to form a disulfide bond; thus, is the least preferred secondary structure for disulfide bonding. The Significant effect of the coil is to the extent that disulfide bonding between a cysteine in the coil and the other one in helix structure is preferable than the disulfide bond between cysteine residues, which are both in sheet structures. Also, high dihedral strain energy in configurations such as Staple can be compensated with structural freedom of coil-coil regions.

The knowledge of disulfides conformational energies can be useful in evaluating the stabilities of protein structures [13]. Energy distributions revealed that most of disulfide bonds have the low energy with –LHSpiral configuration. The more facility of disulfide cleavage is caused by the higher strain energy. Allosteric and catalytic disulfide bonds for their functional role need to disrupt easier than structural disulfide bond. 4 Allosteric bonds emerging in the –RHStaple, –LHHook or –/+RHHook configurations 4, 6, 7 which have detected with higher DSE than many structural disulfide bonds.

Table 1. Disulfide bonds were separated into the 20 configuration groups according to the sign of five dihedral angles. The DSE and distance parameters were calculated for each one, and the maximum, minimum and average are shown respectively for each group as the maximum-minimum (average). The column six shows the percentage distribution of each configuration in secondary structure. CC: coil-coil, HH: helix-helix, SS: sheet-sheet, CH: coil-helix, HS: helix-sheet, CS: coil-sheet, e.g. CS means that in the disulfide bond one of two cysteine residues has secondary structure as coil and the other one has secondary structure as sheet.

Whole dataset	8868	5.5 (3.4-8.8)	3.9 (2.6-7.3)	16.1 (2.1-77.7)	26.9, 25.9, 17.0, 14.3, 8.9, 7.2
-LHSpiral	2373	5.7 (4.1-7.3)	3.8 (2.7-5.1)	10.9 (2.2-68.6)	28.4, 21.0, 17.3, 9.1, 13.9, 10.4
-RHHook	938	5.4 (3.5-7.7)	4.0 (2.9-6.6)	18.7 (2.8-67.8)	30.5, 29.0, 22.1, 4.9, 7.4, 6.2
-RHStaple	795	4.3 (3.4-7.2)	4.1 (3.0-5.0)	18.2 (3.5-62.6)	12.8, 20.8, 3.9, 58.6, 1.1, 2.8
+/-LHSpiral	688	6.0 (4.5-7.5)	3.8 (3.0-5.2)	13.7 (2.8-65.8)	23.1, 23.7, 20.6, 8.9, 10, 13.7
-/+RHHook	678	5.2 (3.7-7.2)	4.0 (3.1-7.3)	14.6 (3.2-66.5)	24.5, 33.5, 20.8, 4.1, 6.6, 10.5
-LHHook	648	5.5 (3.4-8.8)	4.0 (2.9-6.6)	20.1 (2.5-75.7)	29.8, 28.7, 19.8, 6.6, 9.6, 5.6
+/-RHSpiral	585	6.0 (4.2-7.6)	3.8 (2.8-5.8)	14.8 (2.5-63.3)	38.1, 19.8, 9.9, 18.1, 10.9, 3.1
-RHSpiral	505	6.0 (3.8-8.2)	3.8 (2.8-6.2)	17.1 (3.0-72.5)	31.5, 28.7, 13.5, 9.9, 12.1, 4.4
-/+LHHook	338	5.4 (3.6-7.6)	4.0 (2.8-5.9)	18.5 (2.1-60.6)	31.1, 30.2, 16.9, 13, 3.8, 5.0
+/-RHHook	282	5.7 (3.9-8.5)	3.9 (2.6-6.3)	20.4 (2.4-59.8)	12.4, 27.3, 43.6, 9.9, 3.2, 3.5
+/-LHHook	236	5.9 (4.5-8.0)	4.0 (2.9-5.6)	20.8 (3.1-60.1)	26.7, 29.7, 15.3, 22.0, 3.8, 2.5
+RHSpiral	156	6.0 (4.9-8.2)	3.8 (2.9-6.5)	14.8 (2.7-68.1)	41.0, 17.9, 10.3, 13.5, 15.4, 1.9
+/-RHStaple	120	5.2 (3.7-8.0)	4.1 (3.0-6.3)	25.0 (3.2-77.7)	24.2, 39.2, 18.3, 10.0, 5.0, 3.3
-LHStaple	117	5.0 (3.5-7.3)	4.1 (3.1-5.1)	30.3 (2.7-66.2)	14.5, 35.9, 10.3, 28.2, 1.7, 9.4
+/-LHStaple	113	5.3 (3.5-7.5)	4.0 (3.0-5.8)	24.5 (3.9-54.1)	18.6, 48.7, 14.2, 10.6, 2.7, 5.3
+LHHook	102	5.7 (4.6-8.0)	4.0 (3.0-5.8)	20.9 (2.8-61.0)	35.3, 27.5, 13.7, 16.7, 4.9, 2.0
+RHHook	84	5.8 (4.6-8.3)	4.0 (2.8-5.8)	22.5 (3.1-71.1)	35.7, 40.5, 14.3, 4.8, 1.2, 3.6
+LHSpiral	59	6.3 (5.0-6.9)	3.9 (3.2-4.7)	18.7 (3.4-51.7)	22.0, 27.1, 11.9, 23.7, 6.8, 8.5
+LHStaple	28	5.6 (4.5-6.5)	4.1 (2.9-5.6)	25.5 (5.3-75.8)	32.1, 25.0, 3.6, 32.1, 3.6, 3.6
+RHStaple	23	5.7 (4.7-7.0)	4.2 (2.9-5.4)	33.3 (6.1-68.5)	4.3, 60.9, 17.4, 13.0, 4.3, 0.0

References

1. Gilbert, Hiram F. "Protein Disulfide Isomerase and the Catalysis of Oxidative Protein Folding." Folding of Disulfide Proteins. Springer New York, 2011. 133-149.
2. Savojardo, Castrense, et al. "Improving the prediction of disulfide bonds in Eukaryotes with machine learning methods and protein subcellular localization." Bioinformatics 27.16 (2011): 2224-2230.
3. Betz, Stephen F. "Disulfide bonds and the stability of globular proteins." Protein Science 2.10 (1993): 1551-1558.
4. Schmidt, Bryan, Lorraine Ho, and Philip J. Hogg. "Allosteric disulfide bonds." Biochemistry 45.24 (2006): 7429-7433.
5. Dombkowski, Alan A., Kazi Zakia Sultana, and Douglas B. Craig. "Protein disulfide engineering." FEBS letters 588.2 (2014): 206-212.
6. Schmidt, Bryan, and Philip J. Hogg. "Search for allosteric disulfide bonds in NMR structures." BMC structural biology 7.1 (2007): 49.
7. Butera, Diego, et al. "Control of blood proteins by functional disulfide bonds." Blood 123.13 (2014): 2000-2007.
8. Sevier, Carolyn S., and Chris A. Kaiser. "Formation and transfer of disulphide bonds in living cells." Nature reviews Molecular cell biology 3.11 (2002): 836-847.
9. Dubey, Anubha. "The Prediction of Disulphide Bonding in HIV and other lenti-viruses by Machine Learning Techniques." International Journal of Scientific Research in Knowledge 2.2 (2014): 57.
10. van Beek, Hugo L., et al. "Stabilization of cyclohexanone monooxygenase by a computationally designed disulfide bond spanning only one residue." FEBS open bio 4 (2014): 168-174.
11. Thornton, J. M. "Disulphide bridges in globular proteins." Journal of molecular biology 151.2 (1981): 261-287.
12. Petersen, Maria Teresa Neves, Per Harald Jonson, and Steffen B. Petersen. "Amino acid neighbours and detailed conformational analysis of cysteines in proteins." Protein engineering 12.7 (1999): 535-548.
13. Qian, Weili, and Samuel Krimm. "Energetics of the disulfide bridge: An ab initio study." Biopolymers 33.10 (1993): 1591-1603.
14. Creighton, Thomas E. Proteins: structures and molecular properties. Macmillan, 1993.
15. Craig, Douglas B., and Alan A. Dombkowski. "Disulfide by Design 2.0: a web-based tool for disulfide engineering in proteins." BMC bioinformatics 14.1 (2013): 346.
16. Srinivasan, N., et al. "Conformations of disulfide bridges in proteins." Int J Pept Protein Res 36.2 (1990): 147-155.
17. Dani, Vardhan S., C. Ramakrishnan, and Raghavan Varadarajan. "MODIP revisited: re-evaluation and refinement of an automated procedure for modeling of disulfide bonds in proteins." Protein engineering 16.3 (2003): 187-193.
18. Berman, Helen M., et al. "The Protein Data Bank, 1999–." International Tables for Crystallography Volume F: Crystallography of biological macromolecules. Springer Netherlands, 2006. 675-684.
19. Kabsch, Wolfgang, and Christian Sander. "Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features." Biopolymers 22.12 (1983): 2577-2637.
20. Ahmad, Shandar, et al. "ASAView: database and tool for solvent accessibility representation in proteins." BMC bioinformatics 5.1 (2004): 51.
21. Zhang, Hua, et al. "On the relation between residue flexibility and local solvent accessibility in proteins." Proteins: Structure, Function, and Bioinformatics 76.3 (2009): 617-636.
22. Sonavane, Shrihari, Ashok A. Jaybhaye, and Ajaykumar G. Jadhav. "Prediction of temperature factors from protein sequence." Bioinformation 9.3 (2013): 134-140.
23. Kuwajima, Kunihiro, et al. "Kinetics of disulfide bond reduction in. alpha.-lactalbumin by dithiothreitol and molecular basis of superreactivity of the Cys6-Cys120 disulfide bond." Biochemistry 29.36 (1990): 8240-8249.
24. Radivojac, Predrag, et al. "Protein flexibility and intrinsic disorder." Protein Science 13.1 (2004): 71-80.
25. Dombkowski, Alan A. "Disulfide by Design™: a computational method for the rational design of disulfide bonds in proteins." Bioinformatics 19.14 (2003): 1852-1853.