[DOI: 10.24214/jcbps.A.9.1.03554]

Journal of Chemical, Biological and Physical Sciences



CODEN (USA): JCBPAT

An International Peer Review E-3 Journal of Sciences Available online atwww.jcbsc.org Section A: Chemical Sciences

Research Article

Conformational Study on Various Dipeptides And Tripeptides Containing Cysteine: A DFT Approach

S. Subedi¹ and H. P. Lamichhane^{*,2}

¹Department of Physics, Amrit Campus, Tribhuvan University, Kathmandu, Nepal

²Central Department of Physics, Tribhuvan University, Kirtipur, Nepal

Received: 25 August 2018; Revised: 15 October 2018; Accepted: 08 November 2018

Abstract: Density Functional Theory (DFT) calculations have been carried out on a series of thirty two dipeptides and eleven tripeptides. Dipeptides were constructed by fixing cysteine at N-terminus/C-terminus and varying the remaining terminus with sixteen different amino acids. Tripeptides were constructed adding N- and C-terminus different amino acids in two sides of a cysteine. From the geometry optimization of these dipeptides and tripeptides, NT-CT combination (cysteine fixed at N-terminus and C-terminus group varying with different amino acids) of dipeptides, and the basis set 6-311G* gave more precise data than CT-NT combination (cysteine fixed at C-terminus and N-terminus group varying with different amino acids) of dipeptides and the basis set 6-31G*. To study the structural stability and sequence of amino acids in dipeptides and tripeptides, we have investigated the bond lengths and bond angles of amide plane. The absolute deviation obtained in the bond lengths of amide plane for dipeptides and tripeptides were calculated. The analysis of α -carbon bond angle angle resulted that the bond angle around the α -carbon of cysteine residue does not vary significantly, as only the maximum deviation of very small angle was seen in the case of dipeptide, but the bond angle around the α -carbon of varied amino acids showed a significant deviation. And, the bond angle around the α -carbon of tripeptide showed the significant deviation in the α -carbon bond angle of X- and Y-amino group. In conclusion, there is deviation of amide plane from planarity, which was drawn from the investigation of dihedral angle analysis of dipeptides and tripeptides. This deviation has been explained in terms of the combined effect of the hydrogen bonding within the dipeptide or tripeptide and the steric hindrance of the -R group of X-position or Y-position amino acid. In order to study internal barriers to the rotation, we have performed the potential energy scan of the optimized structure of cysteine residue by rotating three different groups separately: (a) amino (-NH2) group (b) carboxyl (-COOH) group and (c) -R (-CH2SH) group.

Keywords: Density Functional Theory (DFT), Dipeptides, Tripeptides, N-terminus, C-terminus

INTRODUCTION

As its name suggests, an amino acid is an organic compound containing an amino group and a carboxyl group¹. The molecules containing both an amino (- NH2) group and a carboxyl (- COOH) group with a side chain (–R group) joined to the α -carbon atom is appeared as **in Figure 1**.

Proteins are by a wide margin the most essential of every organic compound. The very word "protein" is derived from the Greek proteios, which means "of first importance," and the researchers who named these compounds over 100 years back picked a fitting term. Many sorts of proteins exist, and they play out an assortment of functions, including the variety of roles in structural building, catalytic role, movement, transport, hormones, protection, storage, and regulation, and so on in bio life¹. However, the investigation of proteins is deficient without the investigation of amino acids; the essential building blocks of proteins. The astonishing range of physical and chemical properties of proteins is the consequence of the changing composition of amino acids in them^{2.3}.



Figure 1. Representation of α–amino acid.

To form a dipeptide, the amino group of one amino acid forms a peptide bond(C-N bond in **Figure 2**) with the carboxyl group of another amino acid. Because each amino acid has both an amino group and a carboxy group, two different dipeptides can be formed. Plane around this peptide bond is normally referred as the amide plane, which is appeared in **Figure 2**. In a similar way, as dipeptides made, tripeptides and polypeptides can be formed using three and more amino acid residues respectively. A tripeptide has two amide planes and peptide bonds each.



Figure 2. The representation of amide plane in a dipeptide when two amino acids are joined.

The cysteine molecule which is taken in our study, plays a fundamental role in cells since it is a part of proteins and in light of the fact that it acts as a reduced sulfur donor molecule. What's more, the cysteine molecule may likewise play a part in the redox motioning of various stress processes. Cysteine possesses a focal position in the plant essential and auxiliary metabolism because of its biochemical functions. It works in plants as a forerunner for a great number of biomolecules, for example, many plant defense compounds formed in response to various environmental unfriendly conditions. These bio-molecules contain sulfur moieties that play as functional groups and are derived from cysteine, and consequently, are intimately connected by means of their biosynthetic pathways ⁴.

The contribution of "computational" has contributed at first to its advancement; notwithstanding, as the field widens and develops in its significance, the inclusion of "biochemistry" increases conspicuously. In its initial stage, computational biochemistry has been exclusively the area of the individuals who are learned in programming. This impeded the appreciation about computational organic chemistry in the early days. The wide accessibility of cheap microcomputers and application programs in organic chemistry has made a difference to calm these confinements. Entrenched procedures have been reformulated to make more productive utilization of the new PC innovation. New and powerful calculations have been effectively executed. Besides, it is winding up progressively important that scientists are exposed to databases and database management systems because of exponential increment in the information of biochemical significance. Visual displaying of biochemical structures and phenomena can give a more natural comprehension of the procedure being assessed. Simulation of biochemical frameworks gives the biochemist control over the conduct of the model. Molecular displaying of biomolecules empowers researchers not just to anticipate and refine three-dimensional structures additionally to associate structures with their properties and functions⁵.

The objective of the conformational analysis is to reveal insight into conformational attributes of flexible biomolecules and to pick up understanding into the connection between their flexibility and their function. As a result of the significance of this approach, the conformational analysis plays a great part in numerous computational projects running from PC helped drug design to the investigation of molecular dynamics calculations and protein folding. Actually, most structure based medication design projects today utilize conformational investigation strategies as a component of their toolchest^{6,7}.

2. COMPUTATIONAL DETAILS

Density functional technique is picking up prominence in the investigation of electronic structure and estimations of different physical properties of molecules with overwhelming atoms. The outcomes obtained by this technique are viewed as more exact. We have utilized this technique to study the geometry of dipeptides and tripeptides with cysteine and energy barriers of cysteine residue using GAUSSIAN 03 suits of programs⁸.

Thirty two different dipeptide combinations and eleven different tripeptide combinations have been optimized using GAUSSIAN 03 suits of programs at DFT level with Becke's three parameters hybrid functional using LeeYang-Parr⁹ correlation function [B3LYP]. Because of the broad success to optimize large molecules, basis sets 6-31G* and 6-311G* have been used. Dipeptides were constructed in two different ways. At first (in NT-CT combination), cysteine was fixed at the N-terminus position, and other varying amino acids were used in C-terminus position which is hereafter named X-position. In the second way (in CT-NT combination), the terminus of of the amino acids in dipeptides were altered and N-terminus position is named X-position in this case. The X-position amino acids used in both cases are alanine(Ala), arginine(Arg), asparagine(Asn), cysteine(Cys), glutamine(Gln), glycine(Gly), isoleucine(Ile), Ieucine(Leu), lysine(Lys), Methionine(Met), phenylalanine(Phe), serine(Ser), threonine(Thr), tryptophan(Trp), tyrosine(Tyr) and valine(Val). Tripeptides were constructed adding N-

terminus and C-terminus different amino acids in two sides of cysteine. N-terminus amino acids and Cterminus amino acids added in cysteine are hereafter named as X- and Y-amino acids respectively in the case of tripeptides. After optimization of dipeptides and tripeptides by using GAUSSIAN-03 suits of programs with DFT-B3LYP correlation function applying the basis sets 6-31G* and 6-311G*, different parameters were calculated and analyzed.



Figure 3: Optimized structures of dipeptides and tripeptides with different basis sets.

The optimized structures of dipeptides and tripeptides with DFT-B3LYP correlation function applying the basis sets 6-31G* and 6-311G* are shown in the **Figure 3**. The optimized energies of sixteen dipeptides each having two combinations NT-CT and CT-NT and eleven tripeptides using DFTB3LYP method with basis sets 6-31G* and 6-311G* were calculated and listed in **Table 1 and Table 2**.

Table 1 shows that the dipeptide Cysteine-Methionine has the lowest optimization energy -907401 kcal mol⁻¹ and -907538 kcal mol⁻¹ for basis sets $6-31G^*$ and $6-311G^*$ in NT-CT combination and its CT-NT combination has optimization energy -907399 kcal mol⁻¹ and -907537 kcal mol⁻¹ for basis sets $6-31G^*$ and $6-311G^*$ respectively. Similarly, Cysteine-Glycine has the highest optimization energy -583534 kcal mol⁻¹ and -583639 kcal mol⁻¹ for basis sets $6-31G^*$ and $6-311G^*$ in NT-CT combination and its CT-NT combination has optimization energy -583525 kcal mol⁻¹ and -583631 kcal mol⁻¹ for basis sets $6-31G^*$ and $6-311G^*$ respectively. **Table 2** shows that the tripeptide Methionine-Cysteine-PhenylAlanine has the lowest optimization energy -1207581 kcal mol⁻¹ and -1207786 kcal mol⁻¹ for basis sets $6-31G^*$ and $6-311G^*$ respectively. On the other hand tripeptide Glycine-Cysteine-Isoleucine has the highest optimization energy -812737 kcal mol⁻¹ and -812896 kcal mol⁻¹ for basis sets $6-31G^*$ and $6-311G^*$ respectively which is just unit value greater than energy of Valine-CysteineAlanine for both basis sets. **Table 1** and **Table 2** resembles that NT-CT combination of dipeptides and basis set $6-311G^*$ give the stable data after the comparison of obtained energies of dipeptides and tripeptide. So, we will use NT-CT combination of dipeptides and basis set $6-311G^*$ in our further research.

Dipeptide	DFT-B3L	YP/6-31G*	DFT-B3LYP/6-3	11 G *
Combination	NT-CT	CT-NT	NT-CT	CT-NT
Cys-Ala	-608205	-608198	-608315	-608308
Cys-Arg	-785986	-785984	-786142	-786140
Cys-Asn	-714059	-714060	-714200	-714200
Cys-Cys	-858060	-858058	-858187	-858188
Cys-Gln	-738730	-738728	-738876	-738874
Cys-Gly	-583534	-583525	-583639	-583631
Cys-Ile	-682211	-682204	-682336	-682330
Cys-Leu	-682212	-682205	-682337	-682331
Cys-Lys	-716938	-716934	-717073	-717070
Cys-Met	-907401	-907399	-907538	-907537
Cys-Phe	-753187	-753182	-753326	-753322
Cys-Ser	-655392	-655391	-655517	-655517
Cys-Thr	-680069	-680061	-680199	-680191
Cys-Trp	-835747	-835745	-835904	-835903
Cys-Tyr	-800385	-800382	-800539	-800535
Cys-Val	-657542	-657536	-657663	-657656

Table 1: Calculated energy (kcal mol⁻¹) of all the dipeptides.

Table 2: Calculated energy (kcal mol⁻¹) of all the tripeptides

Tripeptide	Calculated Energy(kcal	mol-1)
Combination	DFT-B3LYP/6-31G*	DFT-B3LYP/6-311G*
Cys-Cys-Gln	-1143787	-1143988
Gln-Cys-Gly	-869254	-869434
Gly-Cys-Ile	-812737	-812896
Ile-Cys-Leu	-911415	-911594
Leu-Cys-Lys	-946142	-946331
Met-Cys-Phe	-1207581	-1207786
Phe-Cys-Ser	-955573	-955766
Ser-Cys-Thr	-882458	-882642
Thr-Cys-Trp	-1062806	-1063021
Tyr-Cys-Val	-1004921	-1005124
Val-Cys-Ala	-812738	-812897

RESULTS AND DISCUSSION

3.1. Bond Length and Bond Angle: The general geometrical structures of dipeptide and tripeptide which give the numbering scheme of atoms is shown in **Figure 4 and Figure 5** respectively. These numbering schemes have been followed for all further calculations. Exceptional cases have been written where they present. **Table 3** summarizes five bond lengths 4C-7C, 7C-11O, 7C-13N, 13N-14H, 13N-15C in Å unit and six bond angles 4C-7C-11O, 4C-7C-13N, 11O-7C-13N, 7C-13N-14H, 7C-13N- 15C, 14H-13N-15C in related to the amide plane for the sixteen dipeptides studied. **Table 4** shows five bond lengths 13C-12C, 12C-140, 12C-1N, 1N-2H, 1N-3C in Å unit related to the amide plane of the tripeptide towards X-amino group and five bond lengths 26C-24N, 24N-25H, 24N-6C, 6C-10O, 6C-3C in Å unit related to the amide plane of tripeptide towards Y-amino group.



Figure 4: Geometrical scheme with general atom numbering for dipeptides studied. Atoms 4, 6, 7, 15, 17, 18, 20 and 21 are carbon; 1 and 13 are nitrogen; 11, 22, and 29 are oxygen ;10 is sulphur. All other are hydrogen.

Table 3: The calculated bond lengths and angles of the amide plane for sixteen dipeptides studied. [For bond angles, $a \rightarrow 4C-7C-110$, $b \rightarrow 4C-7C-13N$, $c \rightarrow 110-7C-13N$, $d \rightarrow 7C-13N-14H$, $e \rightarrow 7C-13N-15C$, $f \rightarrow 14H-13N-15C$, the numbering of the atoms is based on **Figure 4**.]

V Amino acid	Bond lei	ngth/ Å				Bond angle/degrees					
A-Annio aciu	4C-7C	7C-110	7C-13N	13N-14H	13N-15C	а	b	с	d	e	f
Ala	1.541	1.221	1.367	1.011	1.449	121.2	116.3	122.5	117.7	120.7	118.1
Arg	1.538	1.226	1.357	1.010	1.456	120.7	116.4	122.9	120.7	122.1	116.6
Asn	1.539	1.225	1.358	1.011	1.456	120.4	116.3	123.2	119.7	122.8	115.5
Cys	1.537	1.225	1.360	1.010	1.453	121.0	116.2	122.7	120.7	122.1	116.3
Gln	1.540	1.222	1.365	1.010	1.451	121.3	116.5	122.2	118.3	120.6	118.2
Gly	1.541	1.220	1.369	1.010	1.441	121.3	116.2	122.5	117.7	120.5	118.4
Ile	1.537	1.226	1.358	1.010	1.456	120.7	116.4	122.9	120.4	121.8	116.6
Leu	1.540	1.222	1.365	1.010	1.453	121.2	116.5	122.3	117.9	120.5	117.9
Lys	1.540	1.222	1.365	1.010	1.453	121.3	116.5	122.2	117.7	120.5	117.9
Met	1.542	1.222	1.364	1.010	1.453	121.2	116.7	122.1	118.7	120.9	118.8
Phe	1.540	1.221	1.366	1.011	1.453	121.4	116.4	122.3	117.4	120.5	117.8
Ser	1.541	1.220	1.368	1.011	1.452	121.4	116.1	122.4	116.7	120.3	117.6
Thr	1.541	1.222	1.365	1.010	1.444	121.0	116.4	122.6	119.4	121.3	116.8
Trp	1.540	1.222	1.365	1.011	1.454	121.3	116.4	122.3	117.5	120.7	117.7
Tyr	1.540	1.221	1.365	1.011	1.453	121.3	116.4	122.3	117.5	120.6	117.8
Val	1.538	1.226	1.358	1.010	1.456	120.7	116.3	122.9	120.3	121.9	116.5
Average	1.540	1.223	1.363	1.010	1.452	121.1	116.4	122.5	118.6	121.1	117.4
Max. Deviation	0.003	0.003	0.006	0.001	0.011	0.6	0.2	0.4	1.9	0.8	1.9



Figure 5: Geometrical structure and general atom numbering scheme for tripeptides studied. Atoms 3,5,6,12,13,17,22,26,28,29,32,38,41,42,44,45,and 48, are carbon; 1,14 and 24 are nitrogen; 10,14,39, and 40 are oxygen ;9 and 32 are Sulphur. All other are hydrogen.

Table 3: The calculated bond lengths and angles of the amide plane for sixteen dipeptides studied. [For bond angles, $a \rightarrow 4C-7C-110$, $b \rightarrow 4C-7C-13N$, $c \rightarrow 110-7C-13N$, $d \rightarrow 7C-13N-14H$, $e \rightarrow 7C-13N-15C$, $f \rightarrow 14H-13N-15C$, the numbering of the atoms is based on **Figure 4**.]

V Amino soid	Bond ler	ngth/ Å				Bond an	Bond angle/degrees					
A-Amino acid	4C-7C	7C-110	7C-13N	13N-14H	13N-15C	а	b	с	d	e	f	
Ala	1.541	1.221	1.367	1.011	1.449	121.2	116.3	122.5	117.7	120.7	118.1	
Arg	1.538	1.226	1.357	1.010	1.456	120.7	116.4	122.9	120.7	122.1	116.6	
Asn	1.539	1.225	1.358	1.011	1.456	120.4	116.3	123.2	119.7	122.8	115.5	
Cys	1.537	1.225	1.360	1.010	1.453	121.0	116.2	122.7	120.7	122.1	116.3	
Gln	1.540	1.222	1.365	1.010	1.451	121.3	116.5	122.2	118.3	120.6	118.2	
Gly	1.541	1.220	1.369	1.010	1.441	121.3	116.2	122.5	117.7	120.5	118.4	
Ile	1.537	1.226	1.358	1.010	1.456	120.7	116.4	122.9	120.4	121.8	116.6	
Leu	1.540	1.222	1.365	1.010	1.453	121.2	116.5	122.3	117.9	120.5	117.9	
Lys	1.540	1.222	1.365	1.010	1.453	121.3	116.5	122.2	117.7	120.5	117.9	
Met	1.542	1.222	1.364	1.010	1.453	121.2	116.7	122.1	118.7	120.9	118.8	
Phe	1.540	1.221	1.366	1.011	1.453	121.4	116.4	122.3	117.4	120.5	117.8	
Ser	1.541	1.220	1.368	1.011	1.452	121.4	116.1	122.4	116.7	120.3	117.6	
Thr	1.541	1.222	1.365	1.010	1.444	121.0	116.4	122.6	119.4	121.3	116.8	
Trp	1.540	1.222	1.365	1.011	1.454	121.3	116.4	122.3	117.5	120.7	117.7	
Tyr	1.540	1.221	1.365	1.011	1.453	121.3	116.4	122.3	117.5	120.6	117.8	
Val	1.538	1.226	1.358	1.010	1.456	120.7	116.3	122.9	120.3	121.9	116.5	
Average	1.540	1.223	1.363	1.010	1.452	121.1	116.4	122.5	118.6	121.1	117.4	
Max. Deviation	0.003	0.003	0.006	0.001	0.011	0.6	0.2	0.4	1.9	0.8	1.9	

Observed magnitude of maximum deviations for five bonds 4C-7C, 7C-11O, 7C-13N, 13N-14H and 13N-15C related to the amide plane of dipeptide in **Table 3** are 0.003Å, .003Å, .006Å, 0.001Å and 0.011Å. The maximum of these five magnitude of maximum deviations observed in bond length is 0.011Å for 13N-15C of Cysteine-Glycine. These data of magnitude of maximum deviation indicate that there is a very small change in bond lengths related to amide plane with the variation of X-amino group in the dipeptide combinations with different amino acids. Also, the maximum deviations for six bond

angles 4C-7C-11O, 4C-7C-13N, 11O-7C-13N 7C-13N14H, 7C-13N-15C and 14H-13N-15C indicated in the **Table 3** by a, b, c, d, e and f are 0.6° , 0.2° , 0.4° , 1.9° , 0.8° and 1.9° respectively. Among them, the angle 7C-13N-14H of dipeptide Cysteine-Serine and 14H-13N-15C of dipeptide Cysteine-Asparagine has the same maximum deviation of bond angle with the value of 1.9° . In the same footing to the case of bond lengths, these data of very small maximum deviations indicate that there is a very small change in bond angles related to the amide plane with the variation of X-amino group with different amino acids in the dipeptides.

Table 4:	The calculated	bond lengths	of amide	planes	for e	eleven	tripeptides	studied.	[The	numbering	g of
the atom	s is based on Fi g	gure 5.]		_						-	

	Bond lengt	h/ Å					Bond lengt	h/ Å			
X-amino acid	13C-12C	12C-140	12C-1N	1N-2H	1N-3C	Y-amino acid	26C-24N	24N-25H	24N-6C	6C-10O	6C-3C
Cys	1.537	1.226	1.357	1.011	1.452	Gln	1.453	1.010	1.357	1.223	1.535
Gln	1.532	1.224	1.359	1.010	1.451	Gly	1.445	1.008	1.359	1.222	1.536
Gly	1.538	1.224	1.360	1.012	1.449	Ile	1.461	1.012	1.354	1.225	1.543
Ile	1.534	1.225	1.359	1.011	1.450	Leu	1.455	1.010	1.357	1.223	1.535
Leu	1.538	1.225	1.361	1.011	1.451	Lys	1.454	1.010	1.357	1.223	1.535
Met	1.533	1.223	1.358	1.010	1.452	Phe	1.462	1.010	1.357	1.223	1.536
Phe	1.538	1.226	1.358	1.011	1.452	Ser	1.459	1.009	1.358	1.222	1.535
Ser	1.544	1.227	1.353	1.017	1.448	Thr	1.446	1.013	1.358	1.223	1.540
Thr	1.551	1.221	1.363	1.011	1.450	Trp	1.454	1.010	1.357	1.223	1.536
Tyr	1.537	1.226	1.358	1.010	1.453	Val	1.461	1.008	1.353	1.225	1.535
Val	1.534	1.225	1.359	1.011	1.451	Ala	1.451	1.010	1.359	1.223	1.537
Average	1.538	1.225	1.359	1.011	1.451	Average	1.455	1.010	1.357	1.223	1.537
Max. deviation	0.006	0.004	0.006	0.001	0.003	Max. deviation	0.010	0.002	0.003	0.001	0.002

The maximum deviations observed in five bond lengths 13C-12C, 12C-140, 12C-1N, 1N-2H, 1N-3C related to the amide plane of tripeptide in **Table 4** towards the X-amino group are 0.006Å, 0.004Å, 0.006Å, 0.001Å, 0.003Å and that in five bond lengths 26C-24N, 24N-25H, 24N-6C, 6C-10O, 6C-3C related to the amide plane towards the Y-amino group are 0.010 Å, 0.002Å, 0.003Å, 0.001Å, 0.002Å respectively. Among them, the maximum of these maximum deviations is 0.006 Å for the bonds 13C-12C and 12C-1N related to the amide plane towards X-amino group of tripeptides Glutamine-Cysteine-Glycine and Serine-Cysteine-Threonine respectively and 0.010 Å for the bond 26C-24N related to the amide plane towards Y-amino group of tripeptide Glutamine-Cysteine-Glycine. These data of maximum deviation being very small values indicate that there is a very small change in bond lengths related to the amide planes with the variation of X-amino group and Y-amino group with different amino acids in tripeptide combinations. The maximum deviations observed for six bond angles 13C-12C-14O, 13C-12C-1N, 14O-12C-1N, 12C-1N-2H, 12C-1N-3C, 2H-1N-3C related to the amide plane of tripeptide towards X-amino group in Table 5 are 0.6°, 0.7°, 0.5°, 1.8°, 0.8°, 0.8° and that for six bond angles 26C-24N-25H, 26C-24N-6C, 25H-24N-6C, 24N-6C-10O, 24N-6C-3C, 10O-6C-3C related to the amide plane towards Y-amino group are 1.6°, 0.8°, 0.8°, 0.4°, 0.4°, 0.5° respectively. Among them, the largest values of maximum deviation are 1.8° for the bond angle 12C-1N-2H related to the amide plane towards Xamino group of tripeptide Serine-Cysteine-Threonine and 1.6° for the bond angle 26C-24N-25H related to the amide plane towards Y-amino group of the same tripeptide. These data of maximum deviation being very small values indicate that there is a very small change in bond angles related to the amide planes with the variation of X-amino group and Y-amino group with different amino acids in tripeptide combinations.

Table 5: The calculated bond angles of amide planes for all the tripeptides studied. [For bond angles of amide plane towards X-amino group, $a \rightarrow 13C-12C-14O$, $b \rightarrow 13C-12C-1N$, $c \rightarrow 14O-12C-1N$, $d \rightarrow 12C-1N-2H$, $e \rightarrow 12C-1N-3C$, $f \rightarrow 2H-1N-3C$ and for bond angles of amide plane towards Y-amino group, $g \rightarrow 26C-24N-25H$, $h \rightarrow 26C-24N-6C$, $i \rightarrow 25H-24N-6C$, $j \rightarrow 24N-6C-10O$, $k \rightarrow 24N-6C-3C$, $l \rightarrow 10O-6C-3C$. The numbering of the atoms is based on **Figure 5**.]

]	Bond ang	le/ degree	e]	Bond ang	le/ degree	e	
X-amino acid	а	b	с	d	е	f	Y-amino acid	g	h	i	j	k	1
Cys	120.6	116.6	122.7	121.8	122.5	115.2	Gln	118.8	120.6	118.4	122.6	116.1	121.3
Gln	121.0	116.5	122.5	122.2	122.4	115.4	Gly	119.3	122.1	118.4	123.3	115.5	121.2
Gly	121.8	115.7	122.5	122.3	121.9	114.6	Ile	119.2	120.8	119.5	123.0	116.3	120.7
Ile	121.6	116.2	122.2	122.1	122.5	115.4	Leu	118.7	120.8	118.1	122.6	116.0	121.4
Leu	121.2	116.6	122.2	122.1	122.5	115.2	Lys	118.7	120.9	118.0	122.6	115.9	121.4
Met	120.9	116.5	122.6	122.2	122.4	115.4	Phe	118.5	122.0	119.5	123.3	115.7	121.1
Phe	120.9	116.5	122.6	121.6	122.1	115.2	Ser	118.4	122.3	119.3	123.5	115.4	121.1
Ser	120.6	116.7	122.7	120.1	121.4	118.0	Thr	117.0	121.2	119.1	123.2	115.6	121.3
Thr	121.5	116.5	122.0	122.6	122.1	114.7	Trp	118.6	121.1	118.0	122.6	115.9	121.4
Tyr	121.1	116.3	122.7	121.6	122.5	115.0	Val	118.8	121.3	119.8	123.1	116.1	120.8
Val	121.7	116.1	122.2	122.1	122.4	115.4	Ala	118.9	122.1	118.0	123.3	115.5	121.2
Average	121.2	116.4	122.4	121.9	122.2	115.4	Average	118.6	121.4	118.7	123.0	115.8	121.2
Max. deviation	0.6	0.7	0.5	1.8	0.8	0.8	Max. deviation	1.6	0.8	0.8	0.4	0.4	0.5

From the comparison of the data obtained for the bond lengths and angles related to the amide plane of dipeptides and tripeptide, the conclusion can be drawn that there is not any significant changes in amide plane while changing the X-amino group and Y-amino group with the different amino acids. The small changes in the parameters like bond lengths and angles might be due to the variation of -R group of amino acids used in X- and Y-positions and due to the steric interaction of local species or those directly bonded to the atom which is connected to the α -carbon atom and H-bonding. Hence, the amide planes remain more or less rigid on the entire chain of proteins and polypeptides.

3.2. *a*-**Carbon Geometry:** When the geometries around the α -carbon atom vary significantly throughout a series of amino acid residue, they play very important role for peptide structure of the protein. The geometry of the α -carbon atom may contain large number of amino acid residue as protein could consists of thousands of residues. Because of this reason even a slight deviations in this geometry should have a big impact on protein structure. Ideally, the bond angles about an sp³ hybridized carbon^{10, 11} should be 109.5°. Because of the stereogenic nature of the α -carbon atoms, this ideal condition is not expected here. Here the important consideration is that how bond angles around α -carbon atoms change with the variation of X-and Y-amino group with different amino acids in the peptide formation.

In the case of dipeptides, we varied the X-amino group by sixteen different amino acid residues as they have different R-group. And, α -carbon bond angles were measured for both of the residues i.e. for fixed cysteine residue at N-terminus and varying residue(X-group) at C-terminus of dipeptides with respect to the R-group of each other. We have two α -C centers 4C and 15C as shown in **Figure 4** for each of the sixteen dipeptides studied. The measured angles in degree are 6C-4C $_{\alpha}$ -5H, 6C-4C $_{\alpha}$ -1N, 6C-4C $_{\alpha}$ 7C, 17C-15C $_{\alpha}$ -16H, 17C-15C $_{\alpha}$ -18C and 17C-15C $_{\alpha}$ -13N which are represented by alphabets a, b, c, d, e and f respectively in **Table 6**. The α -C bond angles in cysteine residue and that in X-amino acid residues are in the left and right portion respectively in this table.

Table 6: Calculated α -carbon bond angle (in degree) in both amino acid residue of all the dipeptides studied. [$a \rightarrow 6C-4C\alpha-5H$, $b \rightarrow 6C-4C\alpha-1N$, $c \rightarrow 6C-4C\alpha-7C$, $d \rightarrow 7C-15C\alpha-16H$, $e \rightarrow 17C-15C\alpha-18C$, $f \rightarrow 17C-15C\alpha-13N$. The atom numbering is based on **Figure 4.** *For Cys-Gly 17C corresponds to 17H, 16H corresponds to 18H and 18C corresponds to 16C*.]

X Amino acid	Cystein	ne(NT)		X-amir	no acid(C	CT)
A-Amino aciu	а	b	с	d	e	f
Ala	108.0	108.0	110.8	109.7	110.6	111.0
Arg	108.7	107.9	110.4	108.5	112.1	111.3
Asn	108.7	108.1	110.3	108.1	112.3	110.4
Cys	108.7	107.9	110.2	106.1	112.7	113.0
Gln	108.0	108.0	110.6	110.1	111.8	109.6
Gly	107.9	108.1	110.9	107.3	108.0	108.6
Ile	108.7	107.9	110.5	107.5	113.3	111.8
Leu	108.0	108.0	110.6	110.6	111.4	109.1
Lys	108.0	108.0	110.5	110.1	111.3	109.3
Met	108.0	107.9	110.8	109.6	110.6	109.4
Phe	108.0	108.0	110.5	110.2	110.8	108.9
Ser	107.9	108.0	110.7	109.5	110.4	108.5
Thr	107.9	108.1	110.8	109.4	111.1	111.0
Trp	108.1	108.0	110.3	110.4	110.3	108.7
Tyr	108.0	108.0	110.5	110.2	110.7	108.8
Val	108.7	107.9	110.4	107.5	113.3	111.8
Average	108.2	108.0	110.5	109.0	111.3	110.1
Max. Deviation	0.3	0.1	0.4	3.0	3.3	1.6

In **Table 6**, the maximum deviations observed in α -C bond angles 6C-4C_{α}-5H, 6C-4C_{α}-1N, 6C-4C_{α}-7C for cysteine residue are 0.3°, 0.1°, 0.4° and that in α -C bond angles 17C-15C_{α}-16H, 17C-15C_{α}-18C, 17C-15C_{α}-13N for varying X-amino group are 3.0°, 3.3°, 1.6° respectively. The range of the angles for cysteine residue is very small resulting the small values of maximum deviation. This resembles that there is not significant change in the α -carbon bond angles with the variations of X-amino group with different amino acids. This result was expected because the R-group of X-amino acid that affects the geometry of α -carbon of cysteine residue is at four bonds far from it, giving very small effect in any condition. But, significant variations are seen around the α -carbon atom of X-position residue. This result can be verified by the data of larger values of maximum deviation in the α -carbon bond angles 28C-26C_{α}-27H, 28C-26C_{α}-29C, 28C-26C_{α}-24N as 3.0°, 3.3°, 1.6° respectively. The varying -R group being in the nearest position resulted these values of larger values of maximum deviation. From this result we came to the conclusion that the α -carbon geometry is not retained throughout the sequence of amino acid, this fact is important to be considered for the larger peptides.

In the Cysteine-Cysteine combination, we have observed the small variations of α -carbon bond angles about N-terminus cysteine while considerable variations were seen in the case of C-terminus cysteine. This difference even for the same amino acid with different terminus combination can be attributed to the fact that there is a significant difference between the Nitrogen of the amino group of the first amino acid and the Nitrogen of the amino group of the X-amino acid and similarly with carboxyl carbon atom¹⁰.

In the case of tripeptides, α -C bond angles were measured with respect to the R-group for both varying residues (X- and Y-group). There are two α -C centers 13C and 26C for each of eleven tripeptides studied as shown in **Figure 5**. The measured angles in degree are $17C-13C_{\alpha}-16H$, $17C-13C_{\alpha}-16H$, $17C-13C_{\alpha}-16H$, $17C-13C_{\alpha}-12C$, $28C-26C_{\alpha}-27H$, $28C-26C_{\alpha}-29C$, $28C-26C_{\alpha}-24N$. The α -C bond angles for the X- and Y-amino group residues fixed at N- and C-terminus are at the left and right portion respectively in the **Table 7**.

The maximum deviations about the α -C bond angles $17C-13C_{\alpha}-16H$, $17C-13C_{\alpha}-15N$, $17C-13C_{\alpha}-12C$ of X-position are 2.1°, 3.0°, 1.3° and that about the α -C bond angles $28C-26C_{\alpha}-27H$, $28C-26C_{\alpha}-29C$, $28C-26C_{\alpha}-24N$ of Y-position are 1.9°, 3.2°, 2.7°. In the case of tripeptides, we have seen the significant variation in the α -C geometry of both X-position as well as Y-position because of the variation of R-group in every combinations. This result is due to the fact that the varying R-group of both X- and Y-amino acid that affects the geometry of α -carbon is at nearest position. This result of tripeptide supports the conclusion drawn from the α -C geometry analysis of dipeptides that the α -carbon geometry is not retained throughout the sequence of amino acid.

Table 7. Calculated α -carbon bond angle (in degree) in both amino acid residue of all the tripeptides studied. [$a \rightarrow 17C-13C_{\alpha}-16H$, $b \rightarrow 17C-13C_{\alpha}-15N$, $c \rightarrow 17C-13C_{\alpha}-12C$, $d \rightarrow 28C-26C_{\alpha}-27H$, $e \rightarrow 28C-26C_{\alpha}-29C$, $f \rightarrow 28C-26C_{\alpha}-24N$]

V Amino acid	Bond a	ngle / de	egree	Y-amino acid	Bond a	Bond angle / degree			
A-Amino aciu	a	b	c		d	e	f		
Cys	108.8	108.0	110.3	Gln	110.0	110.6	109.5		
Gln	108.7	114.4	110.5	Gly	107.1	107.5	107.6		
Gly	105.8	109.2	109.7	Ile	107.1	111.8	112.1		
Ile	106.5	115.2	112.2	Leu	110.2	110.9	108.9		
Leu	111.4	108.5	110.0	Lys	109.6	110.8	109.5		
Met	108.8	114.4	110.1	Phe	109.7	112.6	111.4		
Phe	108.7	108.7	109.8	Ser	109.2	111.8	110.9		
Ser	106.4	109.2	114.1	Thr	109.3	109.1	111.3		
Thr	107.1	109.6	111.8	Trp	109.9	110.5	108.8		
Tyr	108.7	108.6	109.8	Val	107.1	112.0	111.9		
Val	106.6	115.1	112.0	Ala	109.4	110.2	111.5		
Average	107.9	111.0	110.9	Average	109.0	110.7	110.3		

3.3 Dihedral Angle: The valuable information regarding the peptide bond and the planarity of the amide plane can be known by investigating the dihedral angles in the dipeptide and tripeptide. For dihedral analysis of dipeptides, the dihedral angle between atoms 15C and 14H of the amide plane with respect to the peptide bond (bond angle between 7C and 13N) is more specific and will be symbolized as ${}^{\circ}D_{N}$ hereafter. If Amide plane is planar, this dihedral angle ' D_N ' needs to be 180°. The other dihedral angle were studied are: the angle between 13N and 11O with respect to the bond joining 4C and 7C (referred as (D_1) , the angle between 4C and 15c with respect to the bond joining 13N and 7C (referred as (D_2)) and the angle between 4C and 14H with respect to the bond joining 13N and 7C (referred as 'D₃'). For a planar structure, these dihedral angles D_1 , D_2 and D_3 should be 180°, 180° and 0° respectively ^{8, 9}. The dihedral angle with the value -180° also represents the planar structure. The dihedral angles D₁, D₂, D₃ and D_N considered above are listed in **Table 8**. The values of these angles if they represented a planar structure are shown in brackets. The negative values of 180° also represents the planar structure. None of the dihedral angles taken into considerations in our analysis except D₁ for Gln, Gly and Leu in X-position have the perfect angle $(180^{\circ} \text{ or } 0^{\circ})$ which can correspond to a perfect planar amide plane. All other dihedrals have certain value of deviations from their expected values i.e. these dihedrals donot correspond to the planar amide plane. The conclusion can be drawn from this result that the geometry about the amide plane Nitrogen (atom 13N) is not planar.

Table 8: Calculated dihedral angles (in degree) of the amide plane for all the dipeptides studied. [The atom numbering is based on Figure 4.]

X-amino Dihedral angle /degree

acid	D ₁ (180)	D ₂ (180)	$D_{3}(0)$	D _N (180)
Ala	179.5	172.0	13.7	158.3
Arg	177.6	179.3	8.4	170.9
Asn	177.5	175.0	12.1	162.9
Cys	177.1	177.9	9.2	168.7
Gln	-180.0	172.7	12.5	160.1
Gly	-180.0	172.7	12.5	160.1
Ile	179.3	171.5	12.8	158.7
Leu	180.0	171.5	13.5	158.0
Lys	-179.8	171.0	13.6	157.4
Met	179.6	176.5	11.6	164.9
Phe	-179.8	170.7	14.3	156.4
Ser	-179.6	169.6	15.9	153.8
Thr	179.4	172.6	11.2	161.4
Trp	-179.4	170.0	13.2	156.8
Tyr	-179.7	170.6	13.9	156.7
Val	177.1	178.4	11.9	166.5

The deviations of the dihedral angle D_N from 180° have been listed in **Table 9**. Dipeptide Cysteine-Serine showed the maximum deviation of 26.2° in the dihedral D_N , which suggests that the geometry about amide plane nitrogen (13 N) is not planar. This table also presents the values of Φ i.e the dihedral angle between atoms 7C and 18C about the bond 15C-13N and the values of Ψ i.e. the dihedral angle between atoms 1N and 13N about the bond 7C-4C. The table shows the variations of angles Φ and Ψ with respect to various -R group. However, the specific trend of variation of these angles with the variation of –R group cannot be seen here.

As cysteine residue consists of -CH₂SH as the -R group, multiple hydrogen bonding is seen for 10S atom with the 5H and 14H hydrogen atom and also for the 1N atom with 8H and 9H atoms¹². The observed deviation in dihedral angles might be as a result of the presence of the hydrogen bonding within the atoms of cysteine residue as well as with the various atoms of the -R group. It is additionally reasonable to assume that these deviations are due to the steric interferences between the atoms of -R group and H-atoms of amide plane. The maximum deviation of D_N in our study is seen in Cysteine-Serine dipeptide. In this combination, the distances 11O-2H, 10S-5H, 10S-14H and 22O-12H are 2.260Å, 2.986 Å, 2.829Å and 2.294Å respectively. These inter-atomic distances clearly indicate the presence of good hydrogen bonding in the cysteine-Serine dipeptide which resulted the maximum deviation¹³. These inter-atomic distances are consistent with the weak hydrogen bonding¹⁴.

Table 10 shows the dihedral of the amide planes for the tripeptides studied. For the tripeptides studied, there are two amide planes, one along the side of X-amino group and another along Y-amino group. For the dihedral along X-amino group, the dihedral angle between atoms 3C and 2H of the amide plane with respect to the peptide bond(bond angle between 12C and 1N) will be symbolized as 'D_{Nx}' and the dihedral angle between atoms 26C and 25H of the amide plane with respect to the peptide bond(bond angle between 12C and 1N) will be symbolized as 'D_{Ny}' for the dihedral analysis along Y-amino group hereafter. Other dihedrals along X-amino groups are: 'D₁', the dihedral angle between 13C and 3C with respect to the bond 1N-12C and 'D₃', the dihedral angle between the atoms 13C and 2H with respect to the bond 1N-12C. Similarly, other dihedrals along Y-amino group are: 'D₄', the dihedral angle between 24N and 10O with respect to the bond 3C-6C, 'D₅', the dihedral angle between 3C and 26C with respect to the bond 12N-6C and 'D₆', the dihedral angle between the atoms 3C and 25H with respect to the bond 12N-6C. For a planar structure, these dihedrals should be 180° except D₃ and D₆ to be zero degree for that case ^{10,11}.

Table 9: Calculated deviation from 180° in 'D' and the corresponding value of Φ and Ψ in all the dipeptides studied. [The atom numbering is based on **Figure 4**.]

X-amino acid	-R group	D(deviation from 180°)	φ(in degree)	ψ(in degree)
Ala	-CH ₃	21.7	-75.0	149.8
Arg	-(CH ₂) ₃ (NH)C(NH ₂) ₂	9.1	-155.2	117.3
Asn	-CH ₂ (CONH ₂)	17.1	-166.5	121.6
Cys	-CH ₂ SH	11.3	-160.1	113.5
Gln	-(CH ₂) ₂ CONH ₂	19.9	-69.7	149.8
Gly	-Н	19.9	-85.4	150.7
Ile	-CHCH ₃ CH ₂ CH ₃	21.3	-153.7	115.3
Leu	$-CH_2CH(CH_3)_2$	22.0	-65.9	149.1
Lys	$-(CH_2)_4NH_2$	22.6	-64.0	149.6
Met	$-(CH_2)_2SCH_3$	15.1	-64.4	147.7
Phe	-CH ₂ Ph	23.6	-64.3	149.8
Ser	-CH ₂ OH	26.2	-66.9	152.0
Thr	-CHCH ₃ OH	18.6	-87.2	149.0
Trp	-CH ₂ (CCHNH)Ph	23.2	-61.6	149.6
Tyr	-CH ₂ Ph(OH)	23.3	-63.4	149.4
Val	-CH(CH ₃) ₂	13.5	-152.9	114.9

Table 10: Calculated dihedral angles (in degree) of the amide plane for all the tripeptides studied. [The atom numbering is based on **Figure 5**.]

X-amino	nino Dihedral angle (in degree)					Dihedral	angle (in d	egree)	
acid	D ₁ (180)	D ₂ (180)	D ₃ (0)	D _{Nx} (180)	acid	D ₁ (180)	D ₂ (180)	D ₃ (0)	D _{Ny} (180)
Cys	177.3	179.2	7.2	172.0	Gln	180.0	176.2	7.4	168.8
Gln	-179.6	-179.8	-2.1	-177.7	Gly	-178.8	-177.5	-4.6	-172.9
Gly	178.4	176.4	8.8	167.7	Ile	-179.9	176.9	5.4	171.4
Ile	-179.6	177.8	-0.9	178.7	Leu	-179.7	174.5	8.2	166.3
Leu	179.5	176.7	1.4	175.3	Lys	-179.9	174.9	8.5	166.5
Met	-179.4	-179.9	-3.6	-176.3	Phe	-179.3	177.9	-1.3	179.2
Phe	179.3	175.5	3.6	171.9	Ser	-179.4	177.4	-0.8	178.3
Ser	178.4	177.8	5.5	172.3	Thr	179.7	170.7	9.8	160.9
Thr	177.4	175.7	5.1	170.6	Trp	-180.0	175.2	8.1	167.1
Tyr	179.3	175.3	2.1	173.2	Val	180.0	179.1	2.5	176.5
Val	-179.4	177.1	-0.6	177.7	Ala	179.5	177.1	6.0	-171.1

Table11: Calculated deviation from 180° in 'D' and the corresponding value of Φ and Ψ in all the tripeptides studied

X-amino acid	-R group	D(deviation from 180°)	ψ(in degree)	Y-amino acid	-R group	D(deviation from 180°)	ф(in degree)
Cys	-CH2SH	8.0	113.6	Gln	-(CH2)2CONH2	11.2	-71.0
Gln	-(CH2)2CONH2	2.3	123.4	Gly	-H	7.1	109.4
Gly	-H	12.3	-171.4	Ile	-CHCH3CH2CH3	8.6	-148.3
Ile	-CHCH3CH2CH3	1.3	137.7	Leu	-CH2CH(CH3)2	13.7	-65.3
Leu	-CH2CH(CH3)2	4.7	130.8	Lys	-(CH2)4NH2	13.5	-66.4
Met	-(CH2)2SCH3	3.7	125.1	Phe	-CH2Ph	0.8	-137.3
Phe	-CH2Ph	8.1	127.6	Ser	-CH2OH	1.7	-139.9
Ser	-CH2OH	7.7	-165.1	Thr	-CHCH3OH	19.1	-91.9
Thr	-CHCH3OH	9.4	177.1	Trp	-CH2(CCHNH)Ph	12.9	-67.0
Tyr	-CH2Ph(OH)	6.8	125.8	Val	-CH(CH3)2	3.5	-154.2
Val	-CH(CH3)2	2.3	138.6	Ala	-CH3	8.9	-92.1

The deviations of the dihedral angles D_{Nx} and D_{Ny} from 180° have been listed in **Table 11**. Tripeptide Gly-Cys-Ilo showed the maximum deviation of 12.3° in the dihedral D_{Nx} along X-amino acid and Ser-Cys-Thr showed the maximum deviation of 19.1° in the dihedral D_{Ny} along Y-amino acid, which suggests that the geometry about amide plane nitrogens (1N and 24N) is not planar. Table 11 also presents the values of Ψ i.e the dihedral angle between atoms 1N and 15NC about the bond 13C-12N and the alues of Φ i.e. the dihedral angle between atoms 6C and 29C about the bond 26C-24N. The table shows the variations of angles Φ and Ψ with respect to various -R group. However, the specific trend of variation of these angles with the variation of -R group cannot be seen here. As in the case of dipeptides, we find deviations in the dihedral angles D_{Nx} and D_{Ny} from 180° for tripeptides. But these deviations are in a lower range than in the case of dipeptides. These deviations may be due to the presence of hydrogen bonding between different R-groups atoms and atoms in cysteine. It is additionally reasonable to assume that these deviations are due to the steric interferences between the atoms of -R group and Hatoms of amide plane. The maximum deviation of D_{Nx} and D_{Ny} in our study is seen in Gly-Cys-Ilo and Ser-Cys-Thr tripeptides. In Gly-Cys-Ilo combination, the distances 9S-21H, 9S-4H, 14O-7H and 14O-18H are 2.603 Å, 2.895 Å, 2.602 Å and 2.583 Å respectively. In Ser- Cys-Thr combination, the distances 33O-25H, 9S-25H and 10O-27H are 2.832 Å, 2.562 Å and 2.371 Å respectively. These inter-atomic distances clearly indicate the presence of good hydrogen bonding in tripeptides Gly-Cys-Ilo and Ser-CysThr, which resulted the maximum deviations¹³.

3.4 Potential Energy Scan (PES) of Cysteine (*Barriers to Rotation*) **:** To study the internal energy barriers to the rotation in the cysteine residue, we performed the potential energy scan of three different groups:(a) by rotating amino (-NH₂) group (b) by rotating carboxyl (-COOH) group and (c) by rotating - R (-CH₂SH) group. We can get the minimum energy conformation and important structural information about protein from the rotation of different groups. We first optimized cysteine residue at DFT-B3LYP level using the basis set 6-311G* to get optimized geometry and then performed potential energy scan on the same optimized geometry by rotation of -NH₂, -COOH and -CH₂SH groups within the range of -180° to +180° with the increment of 10° intervals by the same method (DFT-B3LYP/6-311G*). The dihedral angles for the rotation of -NH₂, -COOH and -CH₂SH groups are the angle of atoms 6C and 2H with respect to the bond 1N-3C, the angle of atoms 1N and 10O with respect to the bond 6C3C and the angle of atoms 6C and 9S with respect to the bond 5C-3C respectively keeping the rest of the part of cysteine

molecule fixed^{10,15}. The atom numbering of calculations of potential energy scan is based on the **Figure 9.**

The calculated energies for the rotation of $-NH_2$ group vs. different dihedral values have been presented graphically in the **Figure 6**, which has four maximums $A(E_A=-453072.571 \text{ kcal mol}^{-1})$, $C(E_C=-$ 453074.009 kcal mol⁻¹), $E(E_{E}=-453074.214 \text{ kcal mol}^{-1})$, and $G(E_{G}=-453074.557 \text{ kcal mol}^{-1})$ representing the highest energy conformers respectively. These highest conformers A, C, E and G correspond to the dihedral (angle of atoms 6C and 2H with respect to the bond 1N-3C) values 130° , -110° , -170° and 0° respectively. Similarly the lowest energy conformers are $B(E_B=-453077.279 \text{ kcal mol}^{-1})$, $D(E_D=-453077.279 \text{ kcal mol}^{-1})$ 453076.066 kcal mol⁻¹) and H(E_H=-453075.713 kcal mol⁻¹) representing three energy well, correspond to the respective dihedral (angle of atoms 6C and 2H with respect to the bond 1N-3C) values 50°, -50° and - 160° . Two barriers of rotation for the conformer B are 4.708 kcal mol⁻¹ (E_B to E_A) and 2.722 kcal mol⁻¹ (E_B to E_G), barriers for the conformer D are 2.057kcal mol⁻¹ (E_D to E_C) and 1.509 kcal mol⁻¹ (E_D to E_G) and barriers for the conformer H are 1.704kcal mol⁻¹ (E_H to E_C) and 1.499 kcal mol⁻¹ (E_H to E_E). These barriers indicate the significantly hindered rotation of -NH₂ group in cysteine. N-H/O-H bond distances which are responsible for different conformers during potential energy scan have been observed. The highest energy conformers 'A' and 'G' for cysteine arises due to gauche conformation between 1N of -NH₂ group and 120 of -COOH group¹¹. As these highly electronegative atoms are in close vicinity, the repulsion is maximum, which resulted the maximum energy. The lowest energy conformer 'B' is due to the presence of hydrogen bonding between atoms 12O and 14H with bond distance 2.855Å, which lowered the energy giving the maximum stable state. Also, the presence of hydrogen bonding between atoms 12O and 2H of cysteine with bond distance 2.339Å resulted the second lowest conformer 'D'. The bond distance for the bond between atoms 10O and 11H for the conformer 'E' is 4.428Å and that for the conformer 'H' is 2.793Å. The presence of H-bond between 10O and 11H for conformer 'H' resulted the sharp decrement in the energy going from the conformer 'E' to conformer 'H'. The H-bond between 12O and 2H with bond distance 2.339Å is the greatest contributory factor for lowering energy of the conformer 'D' than conformer 'C'.



Figure 6: Energy curve for the rotation of -NH₂ group in cysteine.



Figure 7: Energy curve for the rotation of -COOH group in cysteine



Figure 8: Energy curve for the rotation of -CH₂SH group in cysteine

The calculated energies for the rotation of -COOH group *vs.* different dihedral values have been presented graphically in the **Figure 7**, which has three maximums $I(E_1=-453072.677 \text{ kcal mol}^{-1})$, $K(E_K=-453073.042 \text{ kcal mol}^{-1})$ and $M(E_M=-453074.112 \text{ kcal mol}^{-1})$ representing the highest energy conformers respectively. These highest conformers I, K and M correspond to the dihedral(angle of atoms 1N and 100 with respect to the bond 6C and 3C) values 170° , -180° and 20° respectively. Similarly, the lowest energy conformers are $J(E_J=-453076.667 \text{ kcal mol}^{-1})$ and $L(E_L=-453076.114 \text{ kcal mol}^{-1})$ representing two energy well, correspond to the respective dihedral(angle of atoms 1N and 100 with respect to the bond 6C and 3C) values -60° and 110° . Two barriers of rotation for the conformer J are $3.625 \text{ kcal mol}^{-1}$ (E_J to E_K) and $2.555 \text{ kcal mol}^{-1}(E_J \text{ to } E_M)$ and that for the conformer L are $3.436 \text{ kcal mol}^{-1}(E_L \text{ to } E_I)$ and $2.001 \text{ kcal mol}^{-1}(E_L \text{ to } E_M)$. The highest conformers 'I' and 'K' for cysteine is due to gauche conformation between 12O of -COOH group and 1N of $-NH_2$ group and 'M' is due to gauche conformations between 10O of -COOH group and 1N of $-NH_2$ group 1'. As these highly electronegative atoms are in close vicinity the repulsion is maximum, which resulted the maximum energy or the maximum unstable state. The lowest energy conformer 'J' is due to the presence of hydrogen bonding between atoms 10O and 8H

with bond distance 2.761Å, which lowered the energy giving maximum stable state. Also, the presence of hydrogen bonding between atoms 12O and 8H of cysteine with bond distance 2.627Å resulted the second lowest conformer 'L'.



Figure 9. Conformations obtained for the potential energy scan while rotating different groups

The calculated energies for the rotation of -CH₂SH group vs. different dihedral values have been presented graphically in the **Figure 8**, which has three maximums $P(E_P = -453072.032 \text{ kcal mol}^{-1})$, $R(E_R=-453072.898 \text{ kcal mol}^{-1})$ and $T(E_T = -453073.684 \text{ kcal mol}^{-1})$ representing the highest energy conformers respectively. These highest conformers P, R and T correspond to the dihedral (angle of atoms 6C and 9S with respect to the bond 5C and 3C) values -120°, -0° and 120° respectively. Similarly the lowest energy conformers are Q(E_Q =-453077.911 kcal mol⁻¹), S(E_S =-453077.374 kcal mol⁻¹) and U(E_U =-453076.667 kcal mol⁻¹) representing two energy well, correspond to the respective dihedral(angle of atoms 6C and 9S with respect to the bond 5C and 3C) values -60°, 170° and 60°. Two barriers of rotation for the conformer Q are 5.879 kcal mol⁻¹ (E_Q to E_P) and 5.013 kcal mol⁻¹(E_Q to E_R) that for the conformer U are 3.769 kcal mol⁻¹(E_U to E_R) and 2.982 kcal mol⁻¹(E_U to E_T). The strong hydrogen bond between atoms 10O and 14H with bond distance 2.599Å caused maximum stable state giving the lowest energy conformer 'Q'. Similarly, second lowest conformer 'S' is the resultant given by the strong H-bonds 1N-11H, 12O-2H and 10O-4H with the respective bond distances 2.460Å, 2.431Å and 2.662Å. Another lower conformer 'U' is due to the presence of H-bonds 1N-7H, 1N-8H and 10O-8H with the bond distances 2.606Å, 2.711Å and 2.758Å respectively. These lowest conformers represents the stable states and the highest conformers 'P', 'R' and 'T' represents unstable states.

4. CONCLUSIONS

The different structural parameters studied in this thesis work give a valuable understanding of the structural properties of amino acid sequences and ultimately to protein structure in general. This valuable structural informations of small amino acid sequences enlightens the structural stability of a protein chain. Geometry optimization of the sixteen dipeptides and eleven tripeptides studied in this research by applying DFTB3LYP / 6-311G* level of theory gives more stable results than that obtained by DFT-B3LYP / 6-31G* method. The dipeptide combinations with N-terminus Cysteine and C-terminus other varying amino acids are more stable than that with C-terminus cysteine and N-termination other varying amino acids. The comparative study of the bond lengths and bond angles for the amide plane and the bond angles around the α -carbon atoms shows that these parameters do not vary significantly with the change of residues. It follows that these parameters are essentially fixed and need not be accounted in the entire protein chain. There is no considerable deviation in the geometry around first α -carbon of fixed cysteine residue while the geometry around α -carbon of varying amino acid significantly changes in the case of dipeptides. And there is significant deviations in the geometry around both α -carbons of varying X- and Y-amino group. This result indicates that the geometry around α -carbon at a particular position changes when the amino acid occupying that position is changed. Thus, the geometry about the α -carbon atom is not retained in an amino acid sequence which needs to be considered in the study of the protein chain. The conclusion that the amide plane is not actually planar can be drawn from the study of dihedral angles measured with respect to peptide bond in all sixteen dipeptides and eleven tripeptides. The presence of hydrogen bonding within the atoms of cysteine residue and the atoms of the -R group of the X-position and Y-position residues is one possible reason for this deviation. In addition, the steric hindrance of -R group and hydrogen bonding between amide plane hydrogen and oxygen of carboxylic acid terminus of dipeptides also have caused these deviations. These results are consistent with the results obtained by Keefe and Pearson on dipeptides containing alanine¹⁰, Dalai on dipeptides containing glycine¹¹. Some interesting result obtained from the potential energy scan carried out in order to examine the rotational obstacles in cysteine residue. We have rotated -NH₂, -COOH, and -CH₂SH groups one at a time within the range of -180° to 180° with the increment of 10° by maintaining all other coordinates fixed. The minimum energy conformers obtained from the energy curves obtained by plotting dihedral angle vs. calculated energy were used to calculate the energy barriers for rotation and are considered as the most probable conformers. The energy barrier for -NH₂, -COOH, and -CH₂SH groups are 4.708 kcal mol^{-1} , 3.625 kcal mol^{-1} and 5.879 kcal mol^{-1} respectively. Higher energy barrier was seen in -CH₂SH and -NH₂ group than in –COOH group, which indicates that the structure of cysteine molecule is more rigid during the rotation of -CH₂SH and -NH₂ group. The occurrence of multiple minima with higher barriers in -CH₂SH makes it more rigid to rotation. Large barriers to the rotation in cysteine molecule also suggest the rigidity in the protein chain. This rigidity will be present not only in the backbone but also on the orientation of the –R group.

5. FUTURE WORKS

There are various reasonable ways that this investigation could take later on. Geometry about α -carbon atom would be explained properly from the conformational study of polypeptides. An examination of environmental impacts, for example, a water solvent would likewise be fascinating to complete. It has been recommended that such impacts assume a substantial part in the determination of protein structure¹⁰. A simple technique to explain the hydrogen bonding in dipeptides and tripeptides of cysteine has been carried out in this thesis work but to achieve a whole picture, the study of hydrogen bonding within the polypeptides of cysteine would be a lot beneficial. The result obtained from dipeptides and tripeptides can be compared with the new experiment taking polypeptides. To develop the conclusions in

the form of theory, it would be essential to examine the experimental data with more dipeptides, tripeptides and polypeptides. Large barriers to the rotation in cysteine molecule has been assumed as rigidity in the protein chain. It would be the best to compare this result with PES of dipeptide and polypeptide chain.

ACKNOWLEDGEMENTS

The authors would like to thank the Physics Department of Amrit Campus for providing computer laboratory and Central Department of Physics, Tribhuvan University for providing Gaussian 03 program. We would like to thank Associate prof. Dr. Rajendra Parajuli, Amrit Campus, Tribhuvan University for the wonderful support in this research work.

REFERENCES

- 1. F. A. Bettelheim, W. H. Brown, M. K. Campbell, S. O. Farrell and O. Torres, *Introduction to General, Organic and Biochemistry*, Brooks Cole, Belmont, USA, 2012.
- 2. L. G. Wade, Organic Chemistry, Pearson Education Inc., Massachusetts, USA, 2010
- 3. D. Voet, J. G. Voet, & C. W. Pratt, *Fundamentals of biochemistry*, Wiley, New York, USA, 1999.
- 4. C. Gotor, C. Alvarez, M. A. Bermudez, I. Moreno, I. Garcla and L. C. Romero, Plant signaling and behavior, 2010, 5, 1028-1030.
- 5. C. S.Tsai, An Introduction to Computational Biochemistry, Wiley-Liss, New York, USA ,2002.
- 6. O. M. Becker, A. D. MacKerell Jr., B. Roux and M. Watanabe, *Computational Biochemistry and Biophysics*, M. Dekker, New York, USA,2001.
- 7. J. R. Wagner, C. T. Lee, J. D. Durrant, R. D. Malmstrom, V. A. Feher, & R. E. Amaro, Chemical reviews, 2016, 116, 6370-6390.
- M. J. Frisch, G.W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery, Jr., T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J.Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G.A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A.J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. ZakrzewskiS. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G.Baboul, S., J. Cioslowski, B.B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez and J. A. Pople, Gaussian 03, Revision E.01, Gaussian Inc., Wallingford CT ,2004.
- 9. C. Lee, W. Yang and R. G. Parr, Phys. Rev. B 37, 1988,785.
- 10. C. D. Keefe and J. K Pearson, J. Mol. Struct. 2004, 679, 65-72.
- 11. S. Dalai, J. Phy. Sci.2011, 15, 193-208.
- 12. S. Gronert and A. J. O'Hair, J. Am. Chem. Soc.1995, 117, 2071-2081.

- 13. S. Mondal, D. S. Chowdhuri, S. Ghosh, A. Mishra and S. Dalai, THEOCHEM, 2007, 810, 81-89.
- 14. Y. L. Zhao and Y. D. Wu, J. Am. Chem. Soc.2002, 124, 1570-1571.
- 15. M. N. Barroso, E. S. Cerutti, A. M. Rodriguez, E. A. Jauregui, O. Farkas, A. Perczel and R. D. Enriz, J. Mol. Struct.2001, 584, 21-37.

Corresponding author: H. P. Lamichhane,

1Department of Physics, Amrit Campus, Tribhuvan University, Kathmandu, Nepal, 2Central Department of Physics, Tribhuvan University, Kirtipur, Nepal. Online publication Date: 08.11.2018