2018

A Comprehensive Analysis of Aromatic-Proton Mediated Hydrogen Bonds

Mona S. Alshamrani
South Dakota State University

A COMPREHENSIVE ANALYSIS OF *AROMATIC*-PROTON MEDIATED HYDROGEN BONDS

BY

MONA ALSHAMRANI

A thesis submitted in partial fulfillment of the requirements for the

Master of Science

Major in Chemistry

South Dakota State University

2018

A COMPREHENSIVE ANALYSIS OF *AROMATIC*-PROTON MEDIATED HYDROGEN BONDS

This thesis is approved as a creditable and independent investigation by a candidate for Master of Science in Chemistry degree and is acceptable for meeting the thesis requirements for this degree. Acceptance of this does not imply that the conclusions reached by the candidate are necessarily the conclusions of the major department.

Suvobrata Chakravarty, Ph.D. Date
Thesis Advisor

Douglas Raynie, Ph.D. Date
Head, Department of Chemistry
and Biochemistry

Dean, Graduate School Date Date

This thesis is dedicated to my father without you I would never be where I stand today.

ACKNOWLEDGEMENTS

I would like to thank numerous people in my life who have supported me. First, I would like to thank my mother for her love, patience, and her prayers for me. Many thanks to my husband for his help and encouragement. I would like to thank my two brothers, who have been always proud of me. I also would like to thank all my sisters for all the prayers, encouragement, and support. I would like to thank my mother-in-law for her support. Many thanks to my adviser Dr. Suvobrata Chakravarty for his patience, management, and care to complete this project successfully. Finally, I would like to thank Saudi Arabia, for their academic support and financial support.

CONTENTS

LIST OF FIGURES:

LIST OF TABLES:

ABSTRACT

A COMPREHENSIVE ANALYSIS OF *AROMATIC*-PROTON MEDIATED HYDROGEN BONDS

MONA ALSHAMRANI

2018

Hydrogen bonds play critical role in folding, structure and recognition of biological macromolecules (e.g., proteins, RNA, DNA). In addition to classical hydrogen bonds (e.g., ─OH---O=, ─OH---O─, ─NH---O─ etc.), structural analysis of protein and nucleic acids, almost a decade ago, showed that hydrogen bonds (e.g., CH---O) with hydrogen atoms on aliphatic carbon atoms (hereafter, *aliphatic*-protons) also play very important role in the structure and function of biomolecules. Even though, protons of aromatic ring systems (hereafter, *aromatic*-protons) are more polar than the *aliphatic*-protons, systematic analysis of hydrogen bonds of *aromatic*-protons have not been carried out. Therefore, I carried out a systematic analysis of hydrogen bonds that are made with *aromatic*-protons of tryptophan, tyrosine and phenylalanine amino acids in high-resolution structures of proteins and their complexes using a computer program that I created in R- and Perl languages. The analysis showed that *aromatic* CH---O are very common in proteins and tryptophan-CD1 is the most frequent participant in *aromatic* CH---O hydrogen bonds. The normalized frequency of occurrences of *aromatic* CH---O hydrogen bonds are greater than that of the *aliphatic* CH---O hydrogen bonds. Therefore, like *aliphatic* CH---O hydrogen bonds we anticipate that *aromatic* CH---O hydrogen bonds are likely to play equally, if not more, important role in macromolecular folding, stability and recognition. The numerous

examples of *aromatic* CH---O hydrogen bonds observed here provides a good source of information that can be probed further by experiments to validate their important contribution.

CHAPTER 1. Introduction& Overview

1.1. Significance of CH---O hydrogen bond in biological system.

Hydrogen bonds (H-bonds) happen in the hydrogen of a highly electronegative atoms such as nitrogen, oxygen, fluorine etc.[1].H-bonds are naturally weaker than most covalent bonds, but collectively they can make a huge contribution in the function and stability of biological macromolecules (e.g., all H-bonds of double stranded DNA holds the entire genome). Occasionally, carbon atoms have not been measured a conventional hydrogen bond donor because of carbon has low electronegativity compared to oxygen and nitrogen. In addition to electronegative atoms (e.g., N, O, Cl, F etc.), several studies over the past two decades have established that carbon atoms are also capable of forming weak hydrogen bonds (CH---O hydrogen bonds)[1, 2]. These studies emphasize the important role of CH---O bonds in protein stability and function such as activation, recognition and enzyme catalysis[2, 3].

1.1.1 Role of hydrogen bonds in protein:

In structural characterizations of proteins and protein–protein interfaces[4], analysis of hydrogen bonds is usually the first step of understanding how the protein might function. In general, similar analysis of CH---O H-bonds suggested that they could contribute to the protein stability and protein-protein recognition [5, 6]. There are many studies described the CH---O hydrogen bonding interaction at the protein-protein interface in the *aliphatic*-protons[7], and the calculations showed that the average energy contribution of H-bond is 30%; that of a CH---O hydrogen bonds is 17%; and that of a hydrophobic interactions 50%[7, 8]. Values of CH---O hydrogen bonds can reach as high as ~ 40–50%

in some complexes[7]. Also, this study have reported these results, CH---O hydrogen bonds repeatedly exist between adjacent strands in both parallel and antiparallel orientations, having the obvious structural motif of bifurcated H-bonds[5, 7]. Additionally, this study submits that the weak CH---O hydrogen bonds makes an important influence to the association and stability of protein complexes [3, 7-10]. Therefore, the weak CH---O hydrogen bonds makes an important contribution to the association and stability of protein complexes and protein-protein interaction[5, 7, 8].In addition, In protein–protein binding hydrogen bonds often possibly show a powerful role and hence their designs can be used more directly to recognize native from nonnative binding orientations. Their patterns of interactions not only dictate specific recognition, but also provide much of the stability[3, 5]. Moreover, a recent study of USP7 catalytic domain for activation shows that CH---O hydrogen bond is one of the important factors that modulate USP7cd activation[11], comparative analysis of Inactive and active USP7 states shows that CZ2-HZ2---O in W285 makes hydrogen bond and this is strongly accepted that CH---O hydrogen bond is very important to activation of protein[11]. Furthermore, a comprehensive database analysis of CH---O hydrogen bonds in 3124 alpha-helices and their corresponding helix stations investigate the occurrence of CH---O hydrogen bonds in alpha-helices and helix termini in globular proteins. This analysis shows that CH---O hydrogen bonds contribute to the classical ─NH---O─ hydrogen bonds and other interactions to the overall strength and stability of helix and therefore of proteins [3].

1.1.2 The function of hydrogen bonds in protein-DNA, and RNA stability.

More importantly, C-H---O hydrogen bonds play very effective roles in protein-DNA recognition, as well as DNA and RNA stability[12]. Indeed, CH---O hydrogen bonds contribute to the stability of Z-DNA[13]. The stability of A- and B-DNA comes from the stacking of the bases on top of each other. Otherwise, left-handed Z-DNA is stabilized by 04'---H6-C6 hydrogen bond[8, 14] [13].A strong polarization of the guanine bases in Z-DNA is consistent with the Z-DNA specific guanine 06 and N7 coordination to metal and organic cations. The proximity of its N2 and C8 positions are neighboring phosphate groups, as well as several other Z-DNA-specific conformational features are there [13, 14]. Therefore, CH---O hydrogen bond weak interaction, but it still play significant roles in biological systems. For instance, in protein-DNA complexes, hydrogen bond play role in protein-DNA recognition[15]. A study shows that, on analysis of protein-DNA interfacial hydrogen bonds in 107 crystal structures using the in-house program "pyrHBfind" indicates that an overall smaller number of hydrogen bonds led to the consideration of alternative mechanisms playing a decisive role in stabilizing such complexes[16, 17]. These include CH---O hydrogen bonds[15]. Thus, in binding of proteins to DNA regulatory CH---O interactions play powerful role in the control of many cell processes [4, 13, 17, 18]. Group of researchers investigated and analyzed the CH---O interactions in the protein-DNA interface, based on 43 crystal structures of protein-DNA complexes[18]. They were surprising about the results, they have established that, the number of close intermolecular CH---O hydrogen bond contacts involving the thymine methyl group and position C5 of cytosine is equal to the number of protein-DNA hydrogen bonds involving nitrogen and oxygen atoms as donors and acceptors[18]. After they did a comprehensive

analysis they have reported that C5 of cytosine and C5-Met of thymine form fairly weak CH---O hydrogen bonds with Asp, Asn, Glu, Gln, Ser, and Thr, contributing to the specificity of recognition[18]. Therefore, the classical protein-DNA hydrogen bonds enable to enhance the stability of such complexes[19]. Likewise, there are several factors that can affect selectivity of nucleotide incorporation by a DNA polymerase, base pair hydrogen bonding, base pair size and shape, and base stacking interactions[15, 20]. However, all of these factors are necessary but not enough to achieve optimal efficiency for DNA polymerase comparing to hydrogen bond[12]. Therefore, CH---O hydrogen bond is weak bonds but in this case it is very important factor for efficiency and specificity of the human mitochondrial DNA polymerase[15, 18].

1.1.3 The effectiveness of hydrogen bond and in medical side:

Significantly, several studies have revealed that CH---O hydrogen bonds are responsible for determining of folded protein including enzymes and antibodies[4, 21]. As a result, CH---O bonds will be useful interaction in designing effective antibiotics, antibody, and antigen[21] [9]. In 2013 Lavanya and co-workers have studied the role of CH---O interactions in β-lactamases and the Influence of this bond on the structural stability[22]. In this study, they have studied the effect of CH---O interactions on major specificity and stability of β-lactamases in relation to environmental favorites. They have found that, all the residues located in the active site of β-lactamases are involved in CH---O interactions. These results used in understanding of the stability of β-lactamases[22]. In addition, CH---O hydrogen bonds act as an important driving force in ligand selectivity in the conformational stability of penicillin binding proteins[23], CH---O interactions contribute about 20-25% of total hydrogen bonds and this result will be useful for

researcher to explore penicillin binding proteins[22]. In addition to all the above, recent studies have proved the significance, and effectiveness of CH---O hydrogen bonds in develop new antigens for the design of HIV vaccine in the near future. Wilson and colleagues show that 4E10 recognizes an extended epitope that Includes gp41 MPER and viral membrane lipids and identify specific lipid-binding regions[24].In this study there are many factors led to identify the specific lipid-binding areas, however hydrogen bonds where the TRP residue HD1-atom is the hydrogen bond donor was one of these factors in this study[4, 21]. In Fig (1.1), they have seen that superposition of the 4E10 Fabs bound to 06:0 PA and to peptide epitope respectively, revealed that CDRH3 in its lipid-bound conformation would form unfavorable steric clashes with the gp41 peptide in the joining site[21, 24]. Although Lys100E (H) and Lys683 (P) could easily adopt different conformations, Trp680 (P) located in a restricted pocket, was within 1.3–1.7A° of Leu100C[24].

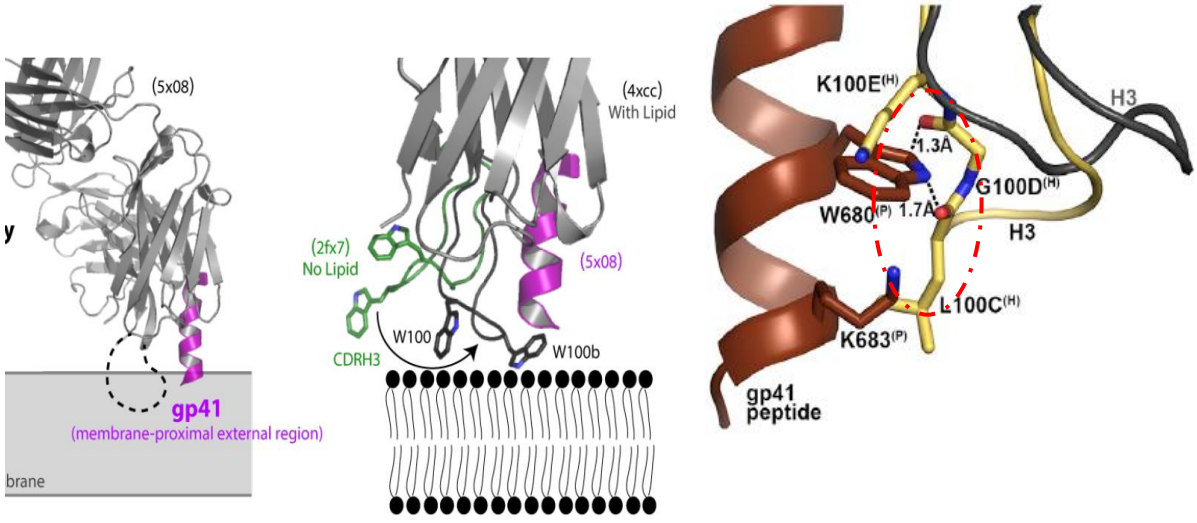


Figure1.1 model for the contact between anti-HIV-1 the antibody (4E10) and the membrane-proximal external region (MPER) of gp41[25].The model showing the unusually long CRDH3 with the two Trp residues (W100 and W110b) that engage the antigen was inspired from the recent study of Rujas [24], and was created using the antibody pdb-ids, 5x08 (in complex with the mutant antigen), 2fx7 (absence of lipids) and 4xcc (in presence of phosphatidylcholine). The proximity of HCDR3 Trp residues to the lipid head groups could be facilitated by anion–quadrupole interactions[25].

Supplementary, CH---O hydrogen bond plays significant role in developing a new antigens for the design of HIV vaccine[21]. In the study of Peptidomimetics built inhibitor design for HIV on of the results in this study examine the role of hydrogen bond which is the cutting results that show six hydrogen bonds formation between the ligand and binding interface of HIV-gp120.According to all of these studies [21, 23, 26], CH---O H-bonds are typically weaker than conventional hydrogen bonds, but this interaction has been considered very important. Hence, there is a critical requirement of studying this type of interaction and expose the common occurrence of CH---O weak hydrogen bonds in protein structures.

1.2 Objective:

Protons of aromatic ring systems (hereafter, *aromatic*-protons) are more polar than the *aliphatic*-protons, systematic analysis of hydrogen bonds of *aromatic*-protons have not been carried out**.** As a result, the objective of this project was to study and analysis the

Aromatic-proton mediated hydrogen bonds, particularly Trp-CD1-HD---O and then we apply this study to other aromatic amino acids (Tyr, and Phe). Trp contains an α-amino group, an α-*carboxylic acid* group, and a side chain *indole*, making it a non-polar *aromatic* amino acid [6], also phenylalanine, tyrosine are all aromatic amino acids as the side chains are formed by aromatic rings. In order to accomplish this objective, we examine the criteria of hydrogen bond between the proton in Trp, Tyr, and Phe with acceptor atoms. In Fig (1.2) shows ring atom of W,Y, and F participate in hydrogen bonds interaction that we will consider in this project.

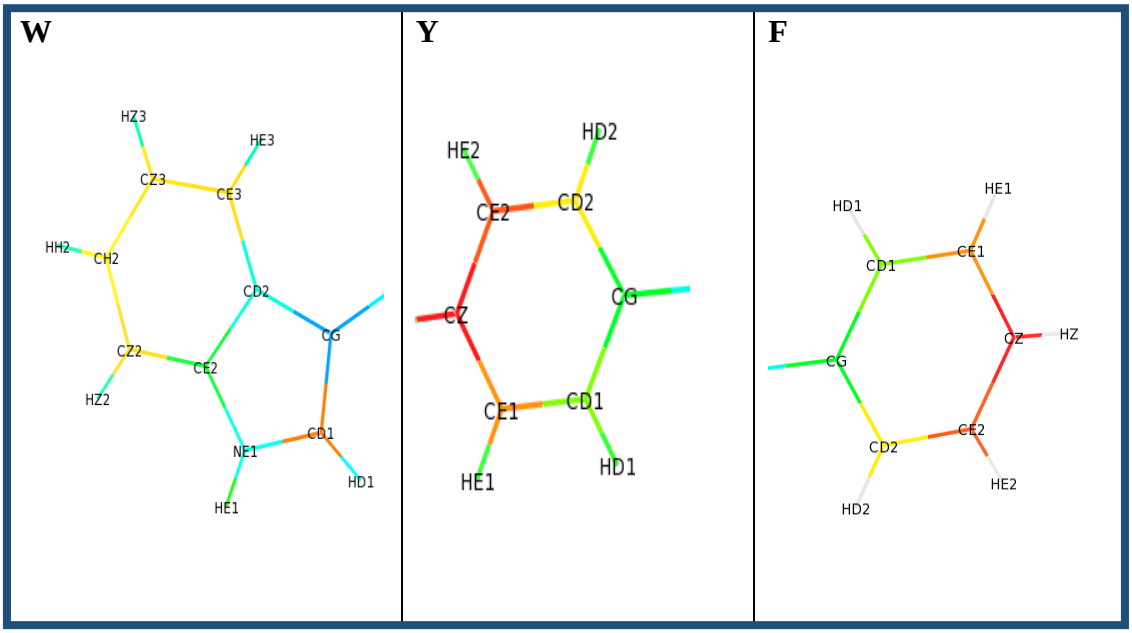


Figure1. 2 Ring atoms of Trp,Tyr, and Phe participate in hydrogen bonds interaction.

*1.2.1 Choice of the dataset, and HB Geometry***.**

A dataset involving 360 PDB files of very high resolution crystal structure of with resolution ≤1 Å, and a non-redundant set of 3700 high-resolution protein structures SCD dataset. We consider this data because with resolution values of 1 Å or so, are highly ordered and it is easy to see every atoms in the electron density map also we will be able to see the protons [6]. Lower resolution structures, with resolution of 3 Å or higher, show only the basic outlines of the protein chain, and the atomic structure must be inferred. Then, we extract aromatic amino acids from the data set because we were only looking for hydrogen bond in this group. Next step, we characterize CH---O hydrogen bonds in the ring atoms of aromatic amino acids by application of one of the most commonly used methods. It has been used for measuring CH---O parameters (angles and distances) [11, 27]. In many studies, almost all of hydrogen bond calculations have been done by using different software programs such as HBPLUS. However, this program has not been used in the study of CH---O *Aromatic*-proton hydrogen bonds. HPBLUS calculates hydrogen bonds of O, N atoms. In this project in order to accomplish this objective of studying the hydrogen bond's parameters, I have developed a new method of studying the parameters of CH---O hydrogen bonds in the ring of aromatic amino acids. Although, in order to identify hydrogen bonds we used the suggested geometric criteria of parameters that are accepted for CH---O hydrogen bonding by various people who have worked in this field[28], and that summarized in the Fig (1.3),Table 1.1

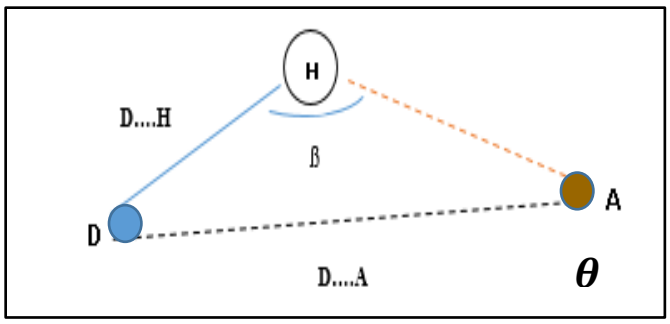


Figure 1.3 Distance and angular parameters used when defining CH---O hydrogen bonds.

Typical van der Waals distances *d* (2.7 Å) and *D* (3.7 Å) are frequently used as distance cutoffs for hydrogen bond identification. Example, the accepted definitions of distance and angle accepted for CH---O hydrogen bonding summarized in the table below[29].

Table 1.1 CH---O hydrogen bond parameters defined in the literature to identify C-H…O Interactions.

Distance C_α…O (Å)	Distance H…O(Å)	Angle (C_α-H…O)	Angle (H…O=C)	Reference
3.4 ± 0.2	2.6 ± 0.2	134 ± 22	-	Chakrabarti et al (protein)1998
3.15	2.63 to 2.79	100-109	91-110	Berman et al (collagen)1996.
3.43	2.50	>90	-	Steiner (small molecules) 1994
<3.5	-	>130	-	Fabiola et al (β sheets) 1997
3.0 to 4.0	2.2 to 2.5	>90	-	Wahl et al. (DNA/RNA) 1997
3.0 to 4.0	<2.8	>110	120 to 140	Desiraju (small molecules)1996

The parameters which were calculated include the following:

- The length of bond between the C atom and the H atom. (D-H)
- The distances between the Carbon atom and the O atom (DA)
- The angle at H, the C-H…O angle (DAH)

Different geometric formation standards have been developed to predict hydrogen bonds occurrence in the protein molecules by [28].

1.2.2 Hydrogen Bond Energy.

Hydrogen bond energy depends on the chemistry of the donor and acceptor atoms as well as their location, based on the potential function[30, 31]. Many studies have considered the characteristics of hydrogen bond using ab initio methods. It is necessary to determine the energy of hydrogen bond because energy considered to be one of the important factors for deciding a good/strong hydrogen bond [30, 32, 33]. Most of quantum chemical calculations and computations were performed by the Gaussian program.

Capabilities of Gaussian (similar to other large- scale computational chemistry packages[31, 34, 35]:

- Determine most stable (optimum) molecular geometry and energy. (Used to calculate ΔE and Keq.)
- Define a potential energy surface by stepping through a range of values for a geometry coordinate, such as bond distance or torsion angle.
- Predict IR, Raman, UV, NMR, and other spectra.
- Optimize transition states.
- Solvate molecules using the “polarized continuum (PCM)” or other models.

- Special tools for optimizing transition metal complexes, and other molecules containing large atoms.
- Model surfaces using a 2D periodic boundary condition (PBC) method, or crystals using 3D PBC.

In addition, the energy scale has been developed to revision hydrogen bonds occurrence in the molecules, they have defended the energy for weak CH---O interactions (−1.0 kcal/mol) and strong (−10.0kcal/mol) [31, 33]. In a comperhasive analysis of anion quderpol interaction in aromatic-proton we have calculated the energy of anion quderpole hydrogen bonds[36].The distributions of the energies suggest that the strength of the interactions, on average, is similar for all of the aromatic residues and that the average energy of Trp interactions is greater (more negative) than that of the others. On the basis of the calculated energy, we loosely defined weak (at least −3.0 kcal/mol) and strong (at most −8.0 kcal/mol) interactions Fig (1.4). This calculation was a driving force to study CH---O *aromatic*-proton mediated hydrogen bonds**.**

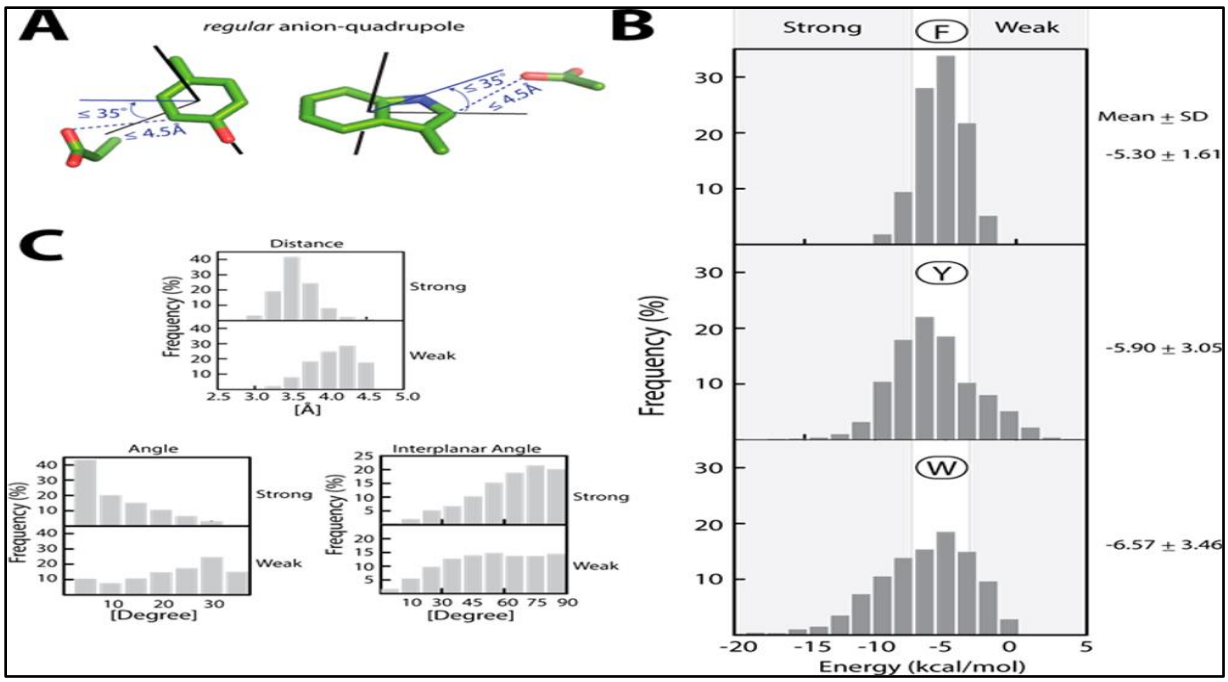


Figure 1.4 Regular anion−quadrupole interactions and their energies. (A) Distance (≤4.5 Å) and angle (≤35°) detection criteria. (B) Distribution of the interaction energies for Phe (top), Tyr (middle),and Trp (bottom). Energy ranges for the strong and weak anion−quadrupole interactions are indicated by the gray background. (C)Comparison of the distribution of properties between strong and weak interactions: distances (top), angles (left, bottom), and interplanar angles (right, bottom)[25].

2. Chapter 2. A Comprehensive Analysis of Aromatic-proton Mediated Hydrogen Bonds

Abstract:

In the biological macromolecules, CH---O hydrogen bonds make a contribution to the protein stability and protein-protein recognition and illustrate protein–protein interfaces[3, 4]. The study and analysis of this interaction is significant, possibly because it has been less investigated since CH---O interactions have been considered weak interactions. A comprehensive statistical analysis of C-H---O hydrogen in a very high-resolution 360 PDB with resolution ≤1 Å, and A non-redundant set of 3700 high-resolution protein structures dataset in order to investigate the *aromatic*-proton creating hydrogen bonds**.** In this study, the CH---O hydrogen bonded of tryptophan interactions with aromatic rings in proteins have been studied by calculating DH, DA, and DAH parameters see Fig (2.1). Systematic analysis of hydrogen bonds of *aromatic*-protons have been accomplished in this study using three different geometrical parameters: the C─H distance (DH), C─O distance (DA) and C—H—O angle (DAH). I studied CH---O interactions of the *aromatic*-proton, and I have observed that CD1-HD1…O interactions show the highest propensity, with Trp, Tyr, and Phe rings. In this analysis, only those contacts with the angle ($90° \leq \emptyset \leq 180°$) were accepted. In addition, bonds associated within a distance 2.3 - 3.5 Å were established as a hydrogen bond. Furthermore, In a comprehasive analysis of an anion quderpol interaction in an aromatic-proton we have calculated the energy of anion quderpole hydrogen bonds[25, 36].The distributions of the energies suggest that the strength of the interactions, on average, is similar for all of the aromatic residues and that the average energy of Trp interactions is greater (more negative) than that of the others. On

the basis of the calculated energy, we loosely defined weak (at least −3.0 kcal/mol) and strong (at most −8.0 kcal/mol) interactions. The energy calculation shows the polarity of the Trp-CD1 atom anions. Consequently, this consists of the CD1-HD1---O interaction geometry and the energy calculation of anion quderpol interaction showing the capability of Trp-CD1 to make a hydrogen bond after Trp- NE1 atom, and the polarity of HD1-atoms to be a H-bond donor in the *aromatic* ring.

2.1 Introduction.

CH---O is one of the non-covalent bonds and it is defined as a special type of weak hydrogen bond [1, 33, 37]. These interactions often happen in the structures of biomolecule like amino acid, protein, DNA and RNA. A number of reports indicate the significance of these type of hydrogen bonds in stabilizing nucleic acid, protein structures, and protein -DNA interface. Furthermore, the weak CH---O bond makes an influential role to the relationship and stability of protein complexes and protein-protein interaction [17, 30].On the medical side, CH---O interactions can be important in drug design; for example, this interactions contribution in the direction of the structural stability and specificity of therapeutic proteins [9, 35, 38].CH---O interactions provide strong proof of the impact of this interaction in protein modeling, ligand design, and structure-activity analysis as well as information on the relative strength of various aromatic CH donors [25, 39].More significantly, the importance in the study of the role of anion–quadrupole interactions in aromatic rings of (Phe),(Tyr),(Trp),in molecular recognition, folding, and assembly [25]. We have looked how the role of anions hydrogen bonds of tryptophan and tyrosine side chains and Trp-CD1 shows highest percentages in the energy calculation. Also, in the

regular anion−quadrupole interactions the frequency of a ring-atom participating in an anion–quadrupole and HB-anion–quadrupole interactions of Trp-CD1 shows highest percentages, see Fig (2.1)[25].

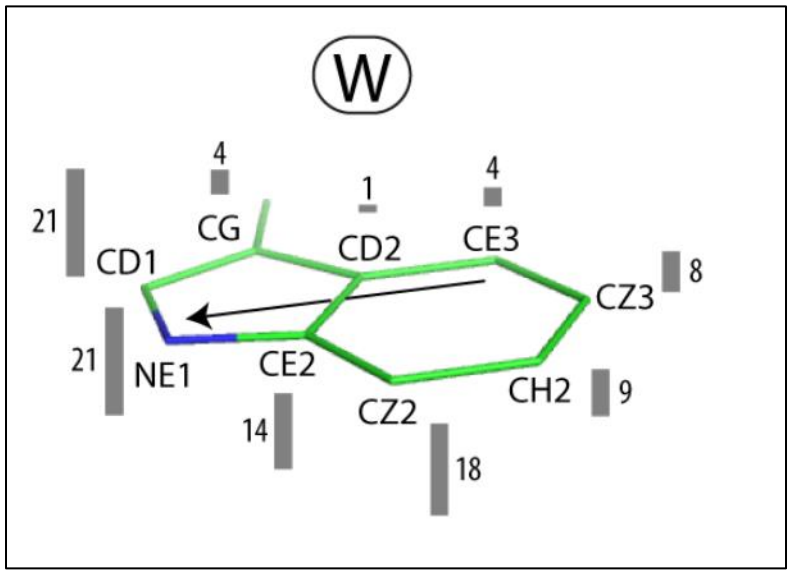


Figure2.1 The frequency of ring-atom participating in anion–quadrupole and HB-anion–quadrupole interactions of the Trp[25].

However, we still need serious study of CH---O interaction in the rings of aromatic amino acids, particularly in the *aromatic*-protons where the HD1-atom is the hydrogen bond donor. Therefore, in this research I scientifically accomplished a systematic analysis of hydrogen bonds that are made with *aromatic*-protons of tryptophan, tyrosine and phenylalanine amino acids in high-resolution structures of proteins and their complexes using a computer program that I created in R- and Perl languages in protein and the ring-atoms participating in HB interactions of the Trp that I consider in this study are showing in Fig (2.2).Thus, this study would also be beneficial in the study of other weak, non-

covalent interactions in biological macromolecules and the program that I have developed will be useful in the study of CH---O geometry for al researchers.

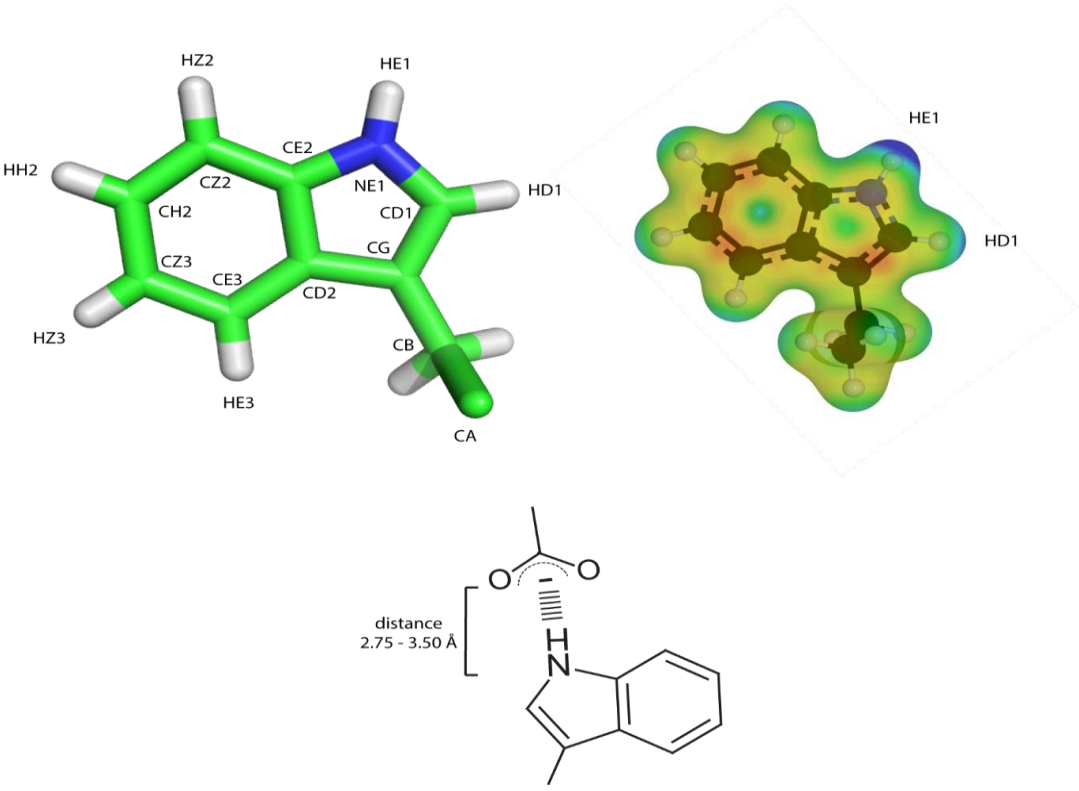


Figure2.2 The ring-atoms participating in HB interactions of the Trp.

2.2 Materials and methods:

2.2.1 Identification of hydrogen bond geometry:

We selected a very high-resolution 360 PDB with resolution ≤1 Å, and A non-redundant set of 3700 high-resolution protein structures SCD dataset in order to investigate the CH---O in *aromatic* rings. The principles used for the selection of the high-resolution

360 PDB data set were solved with resolution ≤1 Å because we need to see the protons to be able to investigate the standard H-bonding criteria for CD1-HD1…O Interactions and other carbon atoms in the *aromatic* ring which are C---H distance (D-H), C---O distance (DA) and C—H—O angle (DAH). Then I have calculated all expected HB in A non-redundant set of 3700 high-resolution protein structures SCD dataset by study and analysis the CD1-HD1…O of Trp, Tyr, and Phe to validate the method that I developed to study CD1-HD1….O geometry .Importantly, I developed a program in R and Perl languages to recognize the CH---O in the aromatics amino acids in the data set that I have described above by determine this H-bonding criteria, and These parameters are represented in Fig (2.2).

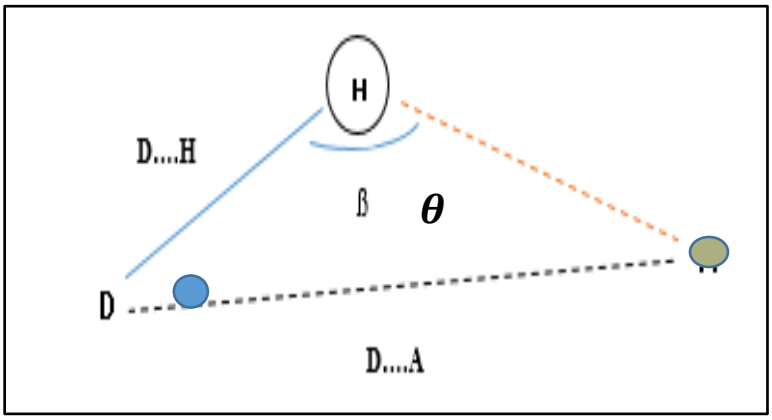


Figure 2.3 The geometrics criteria of hydrogen bond identification, A: Acceptor, D: Donor,

Ø Angle DHA.

To conclude, the geometrics criteria of hydrogen bond identification that I have followed were: D-A≤3.5 Å, D-H≤1.5 Å, and 90°≤ ∅ ≤180°, and only the bonds that satisfy these conditions were accepted as hydrogen bond.

2.2.2 Calculation of hydrogen bond criteria:

Hydrogen donor protein atoms analyzed in this study were HD1 of Trp CD1, Tyr CD1, and Phe CD1; acceptor atoms were oxygen, and carboxyl oxygen of the main chain and side chain of amino acids in the protein structure. Mainly, I used the 3d coordinate X, Y and Z in the PDB files of CD1, and HD1 atoms and then I have created a script in R program to calculate the distance (DH) using the coordinate geometry formula of the distance. The distance between points (x_1, y_1, z_1) and (x_2, y_2, z_2) is:

$$d=\sqrt{(x_2-x_1)^2+(y-y_1)^2+(z_2-z_1)^2}$$

*Where **d** is the distance.*

Similarly, we calculate the distance (DA) of CD1, and acceptor atoms were *oxygen*, and *carboxyl oxygen* of the main chain and side chain. Furthermore, I have written script in R program to find the angle (DHA) between three atoms. I use the formula below:

$$\theta = cos^{-1}\frac{\underset{a}{\rightarrow}\ \underset{b}{\rightarrow}}{|a|\ |b|}$$

Where $\underset{a}{\rightarrow}$ is the vector between hydrogen atom and the acceptor, $\underset{b}{\rightarrow}$is the vector between hydrogen atom and carbon atom, and θ is the angle between $\underset{a}{\rightarrow}$ and $\underset{b}{\rightarrow}$.

2.2.3 Distance from point to the plane. (Highest of acceptor atoms).

Likewise, I have written a script in R to calculate the shortest distance in space between given point and a plane for NE1 and CD1 in TRP amino acid to see the distribution of acceptor atoms, and I use the formula in Fig (2.4) [29],to calculate the distance from point which are acceptor atoms to plane which are CD1,and NE1. Finally, I have looked at the distribution of these acceptors around each plane.

The distance D between a point $P_0(x_0,y_0,z_0)$
and the plane $ax + by + cz + d = 0$
is

$$D=\frac{|ax_0+by_0+cz_0+d|}{\sqrt{a^2+b^2+c^2}}$$

Figure 2.4 Formula of Distance from point to plane

2.2.4. Energy Calculations

The energy of interactions was computed using NWChem quantum mechanical calculations[31]. For computing the energy of interaction (e.g., for an anion–quadrupole pair) the coordinates of side chain atoms of the interacting pair were first extracted from a *p*rotein *d*ata *b*ank (PDB) file. Hydrogen atoms for these side chains were added using the Open Babel version 2.3.1 chemical toolbox. Each CB atom had three protons and setting the pH to 7.0 (in Open Babel) ensured that Asp/Glu carboxyl groups remained deprotonated. An interacting pair of the type **A–B** (e.g., Asp–Phe) had three input files: (i) **A**, (ii) **B**, and (iii) **A–B**, where the **A** input file has the coordinates of the atoms of **A**. Each of the three input files for a pair was imported to the NWChem 6.6 program for energy calculations. Single-point energies were calculated using DFT with the Grimme DFT-D3 dispersion correction*43* with the AUG-CC-PVDZ basis set and B3LYP functional[40]. The energy (in EH units) was extracted from the corresponding NWChem log files[33]. The

energy of interaction (_E**AB**) was obtained as, _EAB = E(**A–B**) – [E(**A**) + E(**B**)], and _E**AB** in EH units was then converted to kcal/mole. There were 6,256 HB-anion–quadrupole **A–B** pairs in total for the energy calculations[30] [25].

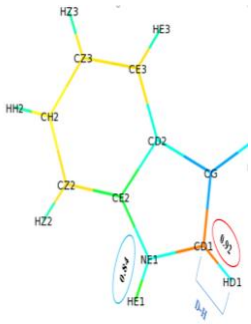
2.3 Results and discussion.

2.3.1 Calculation of hydrogen bond criteria:

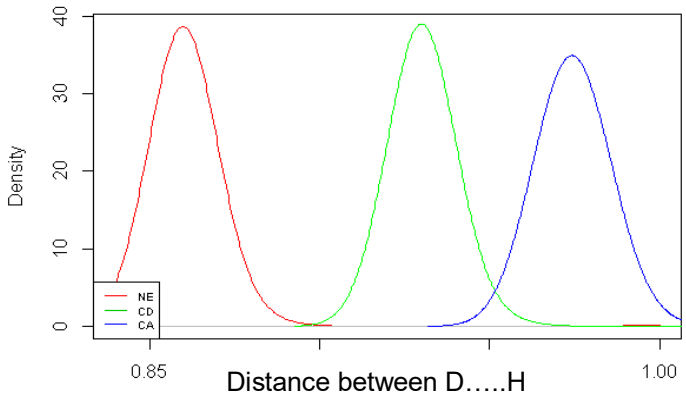
Using a very high-resolution (≤1Å) 360 crystal structures PDB files. A simple process was developed for identifying CD1-HD1…O interactions in proteins. I run the script that I have written (see the method) to calculate the first geometric parameter of HB which is the distance between CD1…..HD1 (the distance (D-H)).I have calculated this distances for all ring atoms in Trp see Table (2.1).

Table2.1 results of D-H in Trp

Atoms	Distance D---H
NE1-HE1	0.84 - 0.99
CD1-HD1	0.92 - 0.99
CZ2-HZ2	0.92 - 0.98
CZ3-HZ3	0.92 - 0.99
CE3-HE3	0.92- 0.98
CH2-HH2	0.91- 0.98
CA-HA	0.97 – 0.99



From *CHARMM36*, the charge of NE1 in the nature is (- 0.51) and the second higher charge is CD1 atom (-0.15) see Fig (2.4). I expect that, the length of CD1- HD1 will be similar or almost close to NE1-HE1. However, the result was disappointed CD1 atom shows longer distance than NE1 atom furthermore CD1 atom shows length close to the other carbon atoms, and the reason was because of the polarity and the charge of NE1 is more than CD1 .In addition, CA---HA shows the longer distance because CA the charge of alpha carbon is very weak comparing to NE1 and CD1 see Fig (2.4). The conclusion, *aromatic*-protons in the ring system are higher charge than the *aliphatic*-protons Fig (2.4).



CHARMM36

NE1= -0.51 **HE1= +0.37**

CD1= -0.15 **HD1=+ 0.22**

CA= +0.07 **HA= +0.09**

Figure 2.5 (**A**) Density plot to show the distributions of distance of NE1-HE1, CD1-HD1, and CA-HA bonds, (**B**) Charge of the atoms ring and there protons.

After that, I used the program that I developed (see the method) to calculate and analysis the second geometric parameter of HB, the distance (DA) which is the distance between ring atoms and the acceptor CD1---O. Adding, all acceptor atoms were (e.g. ─OH, ─COO^-, C=O) of the main chain and side chain that I considered hydrogen bonds only the acceptors with $2.3 \leq DA \leq 3.58$ Å Fig (2.5).

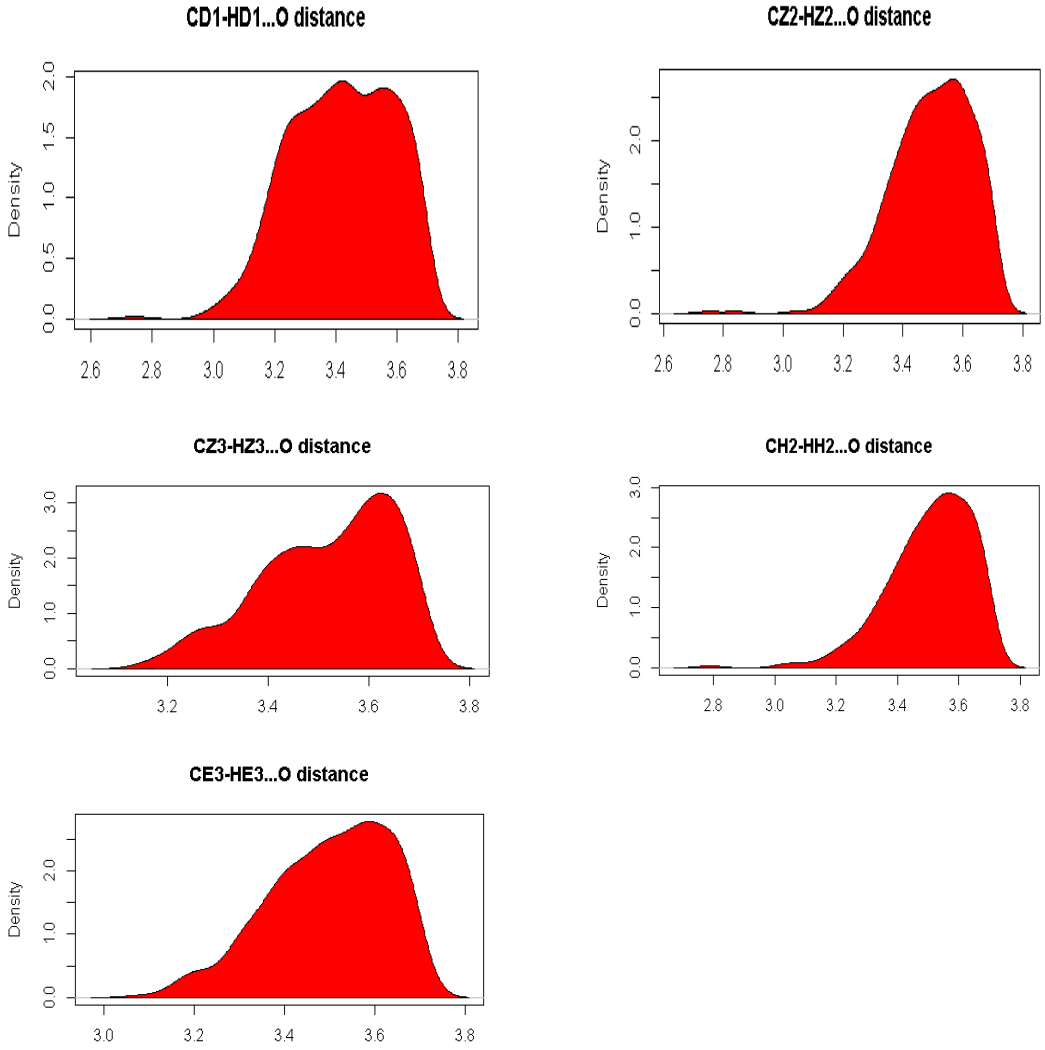
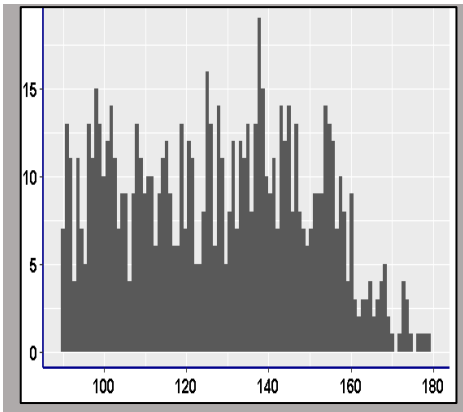
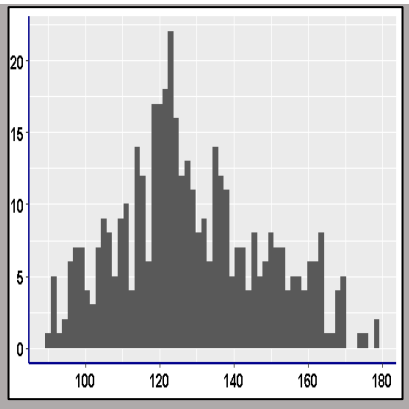


Figure 2.6 Density plot of distribution of DA of the atoms ring of Trp,CD1, CZ2,CZ3,CE3, and CH2.

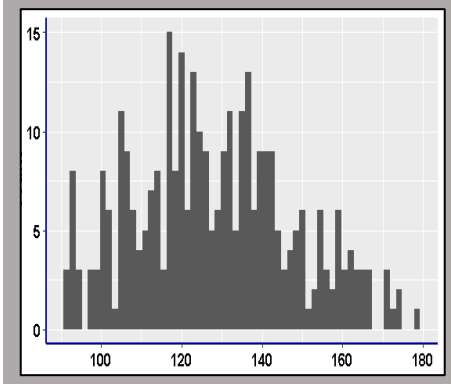
In the analysis of 360 very high resolution protein structure, I have calculated the third geometric parameter of HB which is C—H—O angle (DAH) in Trp-CD1 and other carbon atoms in the ring. Only the acceptor atoms under this condition ($90° \leq \theta \leq 180°$) were accepted as hydrogen bonds acceptor atoms Fig (2.6).



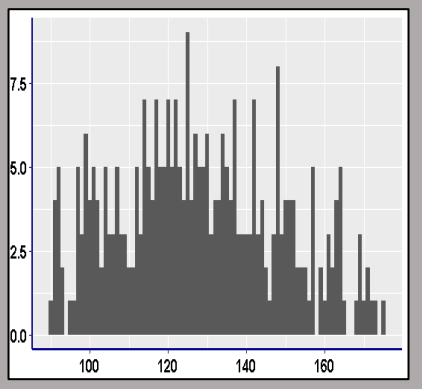
Angle CD1-HD1…O



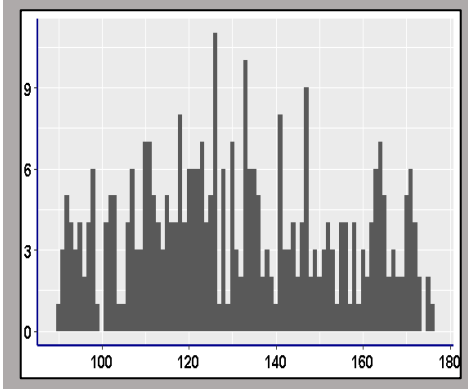
Angle CZ2-HZ2…O



Angle CZ3-HZ3…O



Angle CH2-HH2…O



Angle CE3-HE3…O

90°≤ D-H-A ≤180°

Figure 2.7 Distributions of DHA the angle of all expected HB in Trp-CD1, Trp-CZ2, Trp-CZ3, Trp-CH2 and Trp-CE3.

Successively, I make comparison in terms of total number of acceptor atoms for each atoms in the ring of Trp, CD1, CZ2, CZ3, CE3, and CH2. I have found the total number of hydrogen bond of acceptor atoms under the two conditions: ($2.3 \leq DA \leq 3.58$ Å) and ($90° \leq$ D-H-A $\leq 180°$) for CD1 more than other carbon atoms see Fig (2.8) .

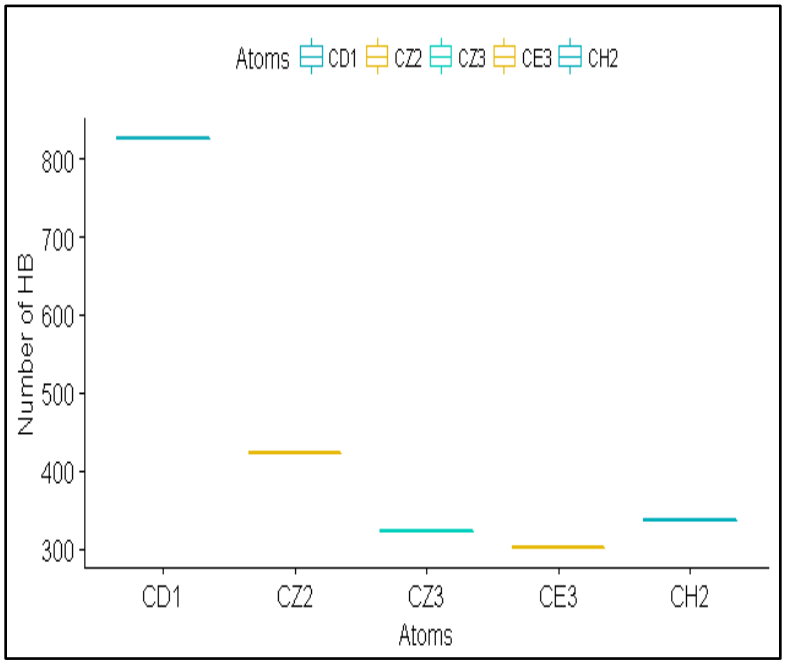


Figure 2.8 Total number of hydrogen bond In CD1, CZ2, CZ3, CE3, and CH2.

Then, I normalize the data and the percentage of total expected HB and in Trp-CD1 44% which is 20% more than Trp-CZ2 .Then I apply CD1 criteria to other carbon atoms in the ring just to make the result more accurate Fig (2.9).

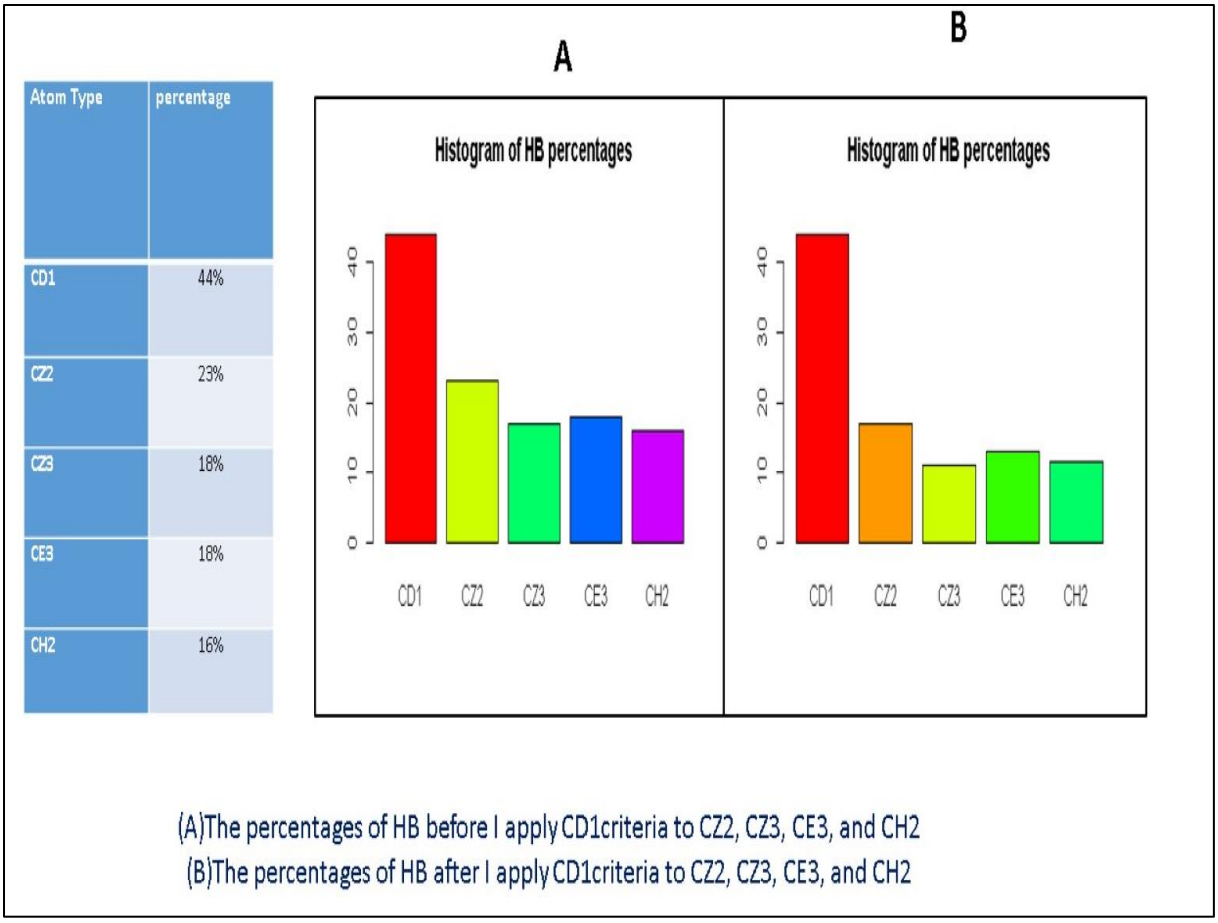


Figure 2.9(**A**) Total number of all expected HB In Trp-CD1 comparing to Other carbon atoms in the ring with (DA 2.3 ≤ DA ≤ 3.58 Å) ,and (90°≤ D-H-A ≤180°) : (**B**) Total number of all expected HB In Trp-CD1 comparing to Other carbon atoms with (DA 2.71 ≤ DA ≤ 3.58 Å.), and (90°≤ D-H-A ≤179°) which is CD1 criteria.

2.3.2 Analysis of CD1-HD1…O hydrogen bond.

To provide more evidence to the previous results which is the capability of CD1 to make hydrogen bond, I have done some test based on the distance (DA) and the angle (DHA) of this bonds. First test, I compare the result of Trp-CD1 mediate hydrogen bond with a weak hydrogen bond (*alpha*- carbon) Trp-CA with a strong hydrogen bond Trp-NE1.I have selected alpha-carbon because many report have showed the ability of alpha

carbon (CA) to make hydrogen bond and that was surprising because the charge of alpha carbon in nature is very small (0.09) comparing to CD1(0.25) and NE1(0.5).Hence, I calculated DA and DAH of NE1, and CA atoms from the same PDB files that I used to study CD1 . After that, I have calculated the percentages of all expected hydrogen bond of Trp-NE1, and Trp-CA and compare the result with hydrogen bond in Trp-CD1 Fig (2.10).

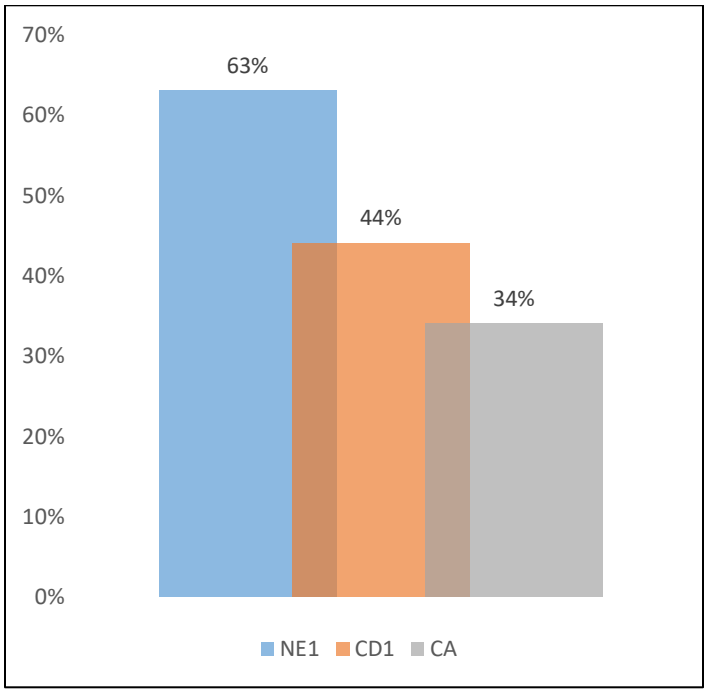


Figure 2.10 Histogram of the percentages of all expected hydrogen bond in Trp-NE1, CD1, and CA.

From the result of the first test above I have proofed the ability of *aromatic*- protons to form CH---O a hydrogen bond in particular Trp-CD1-HD1. Second test, analysis acceptor atoms which are oxygen, and carboxyl oxygen of the main chain and side chain of amino acids in the protein structure in CD1-HD1….O bonds. Conventionally, carbon has not been considered a predictable hydrogen bond donor due to its reasonably low

electronegativity compared with oxygen and nitrogen. In the meantime NE1 have been well studied, and there are some software available to calculate hydrogen bond in nitrogen. Therefore, I got the data of hydrogen bond of NE1 that have been calculated from HBPLUS and then I applied NE1 criteria to CD1 and other carbon atoms in the ring of Trp. Moreover, I applied the criteria of acceptor atoms with sp3, and sp2 hybridization in Trp-NE1….O hydrogen bond to create truthful comparison. In terms of accepter atoms with Sp3 hybridizations: DA of HB of NE1 atoms were 2.3-3.43 Å, and DAH the angle of HB of NE1 atoms were between 98.7°- 179.4°. Accepter atoms with Sp2 hybridizations: DA of HB of NE1 atoms were 2.42-3.43 Å, and DAH of HB of NE1 atoms were 98.7°-179.4°. In the two case sp3 and sp2 hybridizations. CD1 shows the highest percentage comparing to other carbon atoms in the Trp ring Fig (2.11).

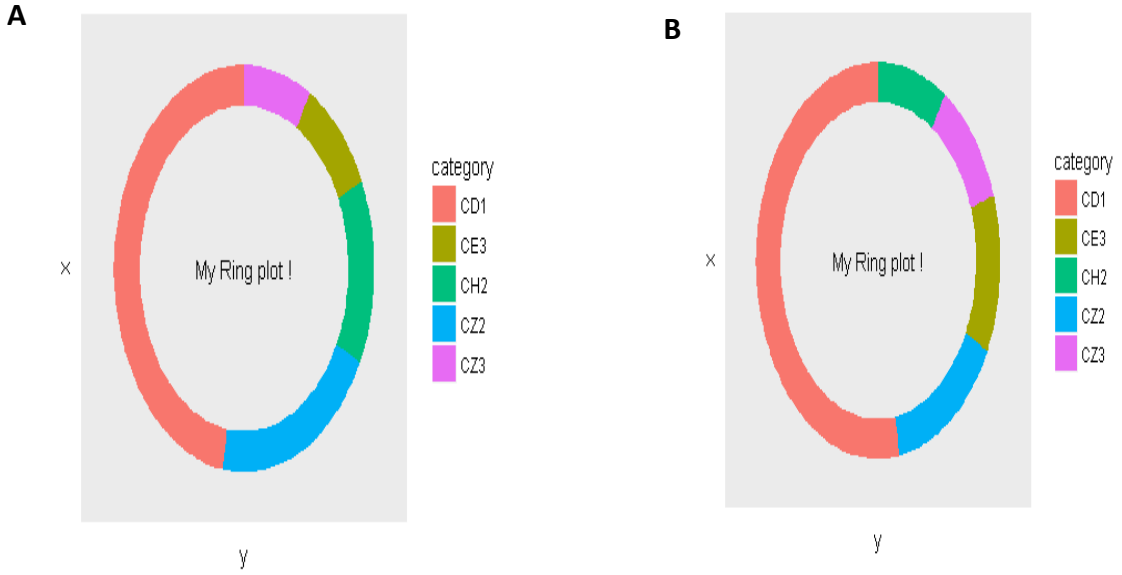


Figure 2.11 Ring plot of percentages of total Hydrogen bond of CD1, CZ2, CZ3, CE3, and CH2 after applying NE1 criteria to them. A: accepter atoms with Sp3 hybridizations, B: Accepter Atoms with Sp2 hybridizations.

Second test, I have done cut off to see where the most population will be for each of DA, and DHA In Trp-NE1 bonds when Sp2, and Sp3 hybridization. I have found that from75 % CUTOFF most of the bonds almost have DA between 2.82 - 3.09 Å and for the DHA 135 °- 170° in Sp2, in terms of Sp3 most of the bonds almost have DA between 2.80 - 3.06Å, and DHA between 130° -175° Fig (2.12). Besides, I use quantile function in R Program to get the exact 75% cutoff values.

A **B**

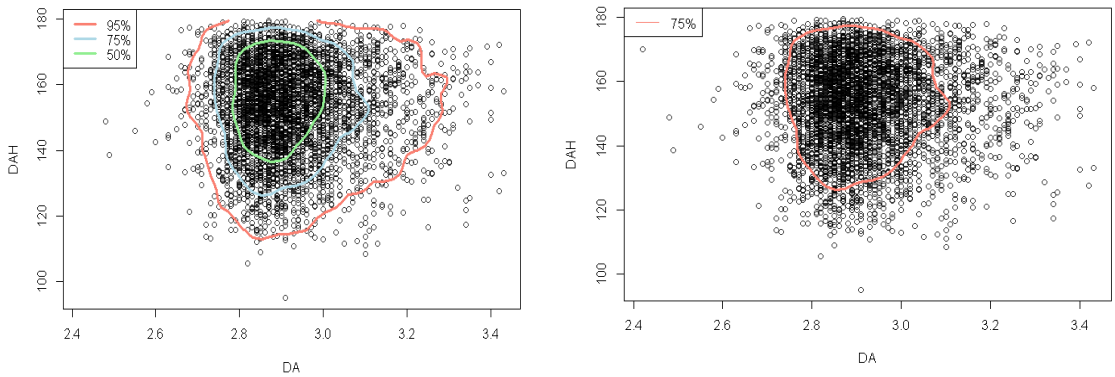


Figure 2.12 The 75% cutoff of DA, DHA values in Trp-NE1….O bond, (A): Sp2 hybridization,(B): Sp3 hybridization.

After that, I use the criteria of 75% cutoff of DA, DHA of Trp-NE1….O and I apply it to DA and DHA of Trp-CD1, Trp-CZ2, Trp-CZ3, Trp-CE3, and Trp-CH2. as a result, CD1 shows the highest percentages comparing to other carbon atoms in the ring Fig (2.13).

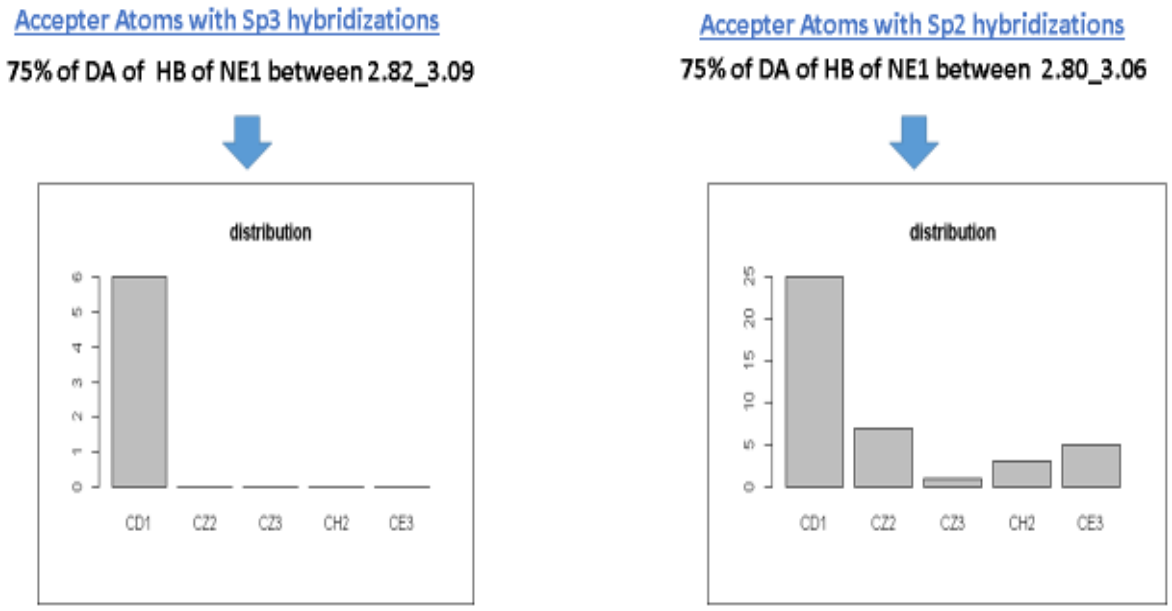
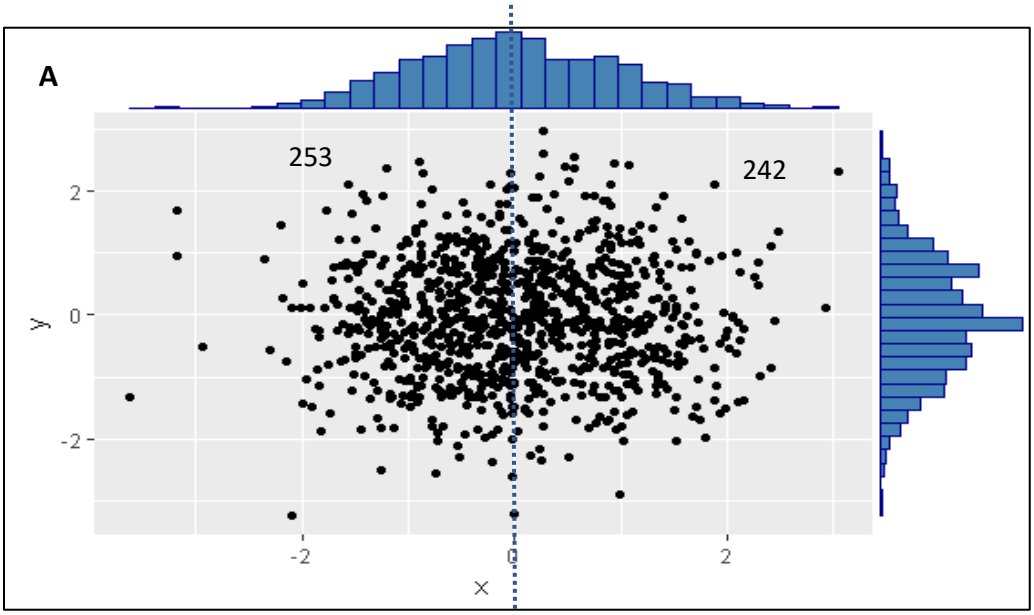


Figure 2.13 Histogram of total hydrogen bond in CD1 and other carbon atoms in the ring after applying NE1 75% cutoff criteria.

2.3.3 Analysis of acceptor atoms in CD1-HD1…O hydrogen bond.

In the calculation of the shortest distance in space between given point (acceptor), and plane which are Trp-NE1, and Trp-CD1, the highest of all acceptor to the two plan:Trp-NE1,and Trp-CD1 were from 0.002 to 2. Before I take the absolute values as the formula require see Fig (2.14) in the method, I use the result and make plot to look at the distributions of acceptor atoms around the plan. Also, I have decided a vector in the plan and see the distribution of the acceptor of CD1 and NE1 around this vector because CD1

atom close to NE1 atom in the ring ,and NE1 has high electronegativity I want to identify if the polarity of CD1 because of this feature. The distribution and the number of acceptor atoms around each plane (NE1, CD1) were equal. As a result, the polarity of CD1 is not because of the position of this atom in the ring Fig (2.14).



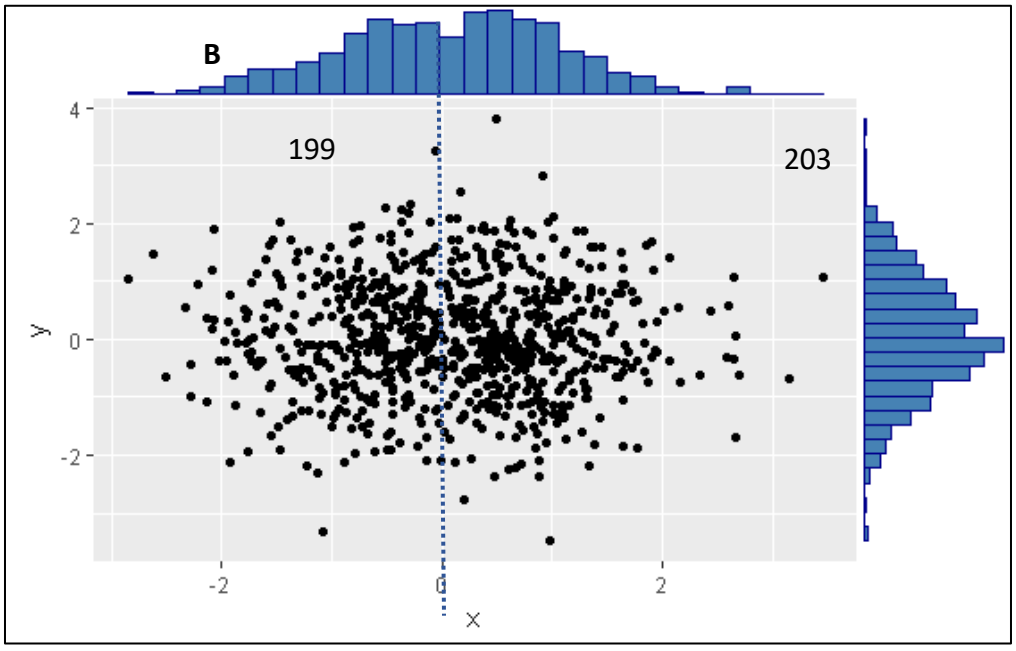


Figure 2.14 (A) distribution of Acceptor atoms around NE1, (B) distribution of Acceptor atoms around CD1, The number of acceptor atoms are equal around each plane.

In the analysis of acceptor atoms that I have found in hydrogen bond calculation of CD-HD1…O bond, I analyze all side chain acceptor atoms that are making hydrogen bond with CD1 Fig (2.15) and main chain Fig (2.16). CD1 show the highest percentage in all acceptor type in main chain, however, in side chain CD1 was the highest in all of them but when the acceptor is Gln, Trp-CZ2 make hydrogen bonding with this type of acceptor more than CD1. In general, in both case side chain and main chain CD1 show higher number in making hydrogen bonding.

Acceptor (Side Chain).

Atoms	ASP	GLU	ASN	GLN	SER	THR	TYR
CD1	57	33	31	11	37	25	17
CZ2	25	31	9	19	21	16	21
CZ3	6	15	5	8	13	9	16
CH2	13	28	5	11	11	12	10
CE3	14	11	17	7	13	4	11

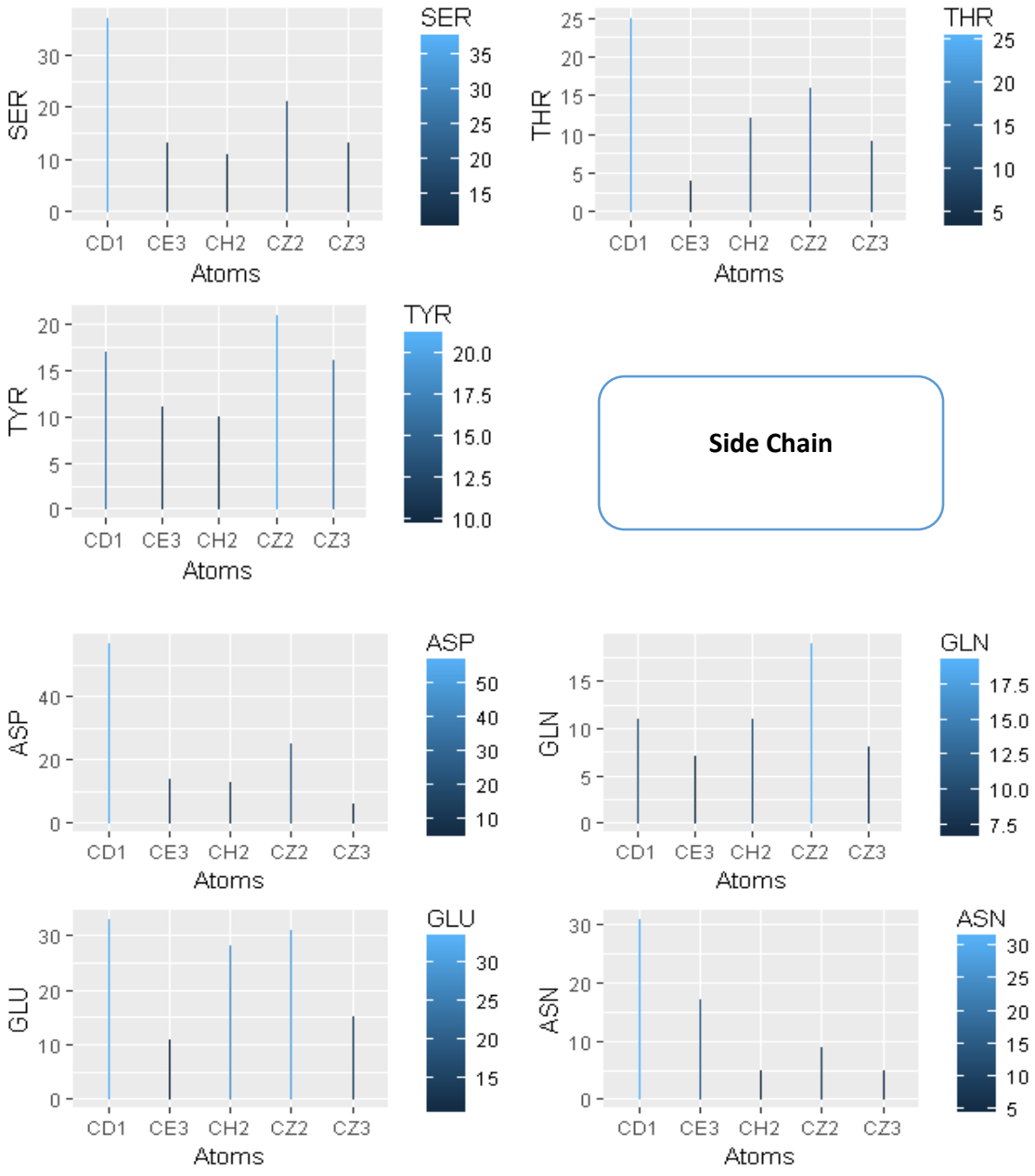


Figure 2.15 acceptor atoms side chain that are making hydrogen bond with CD1.

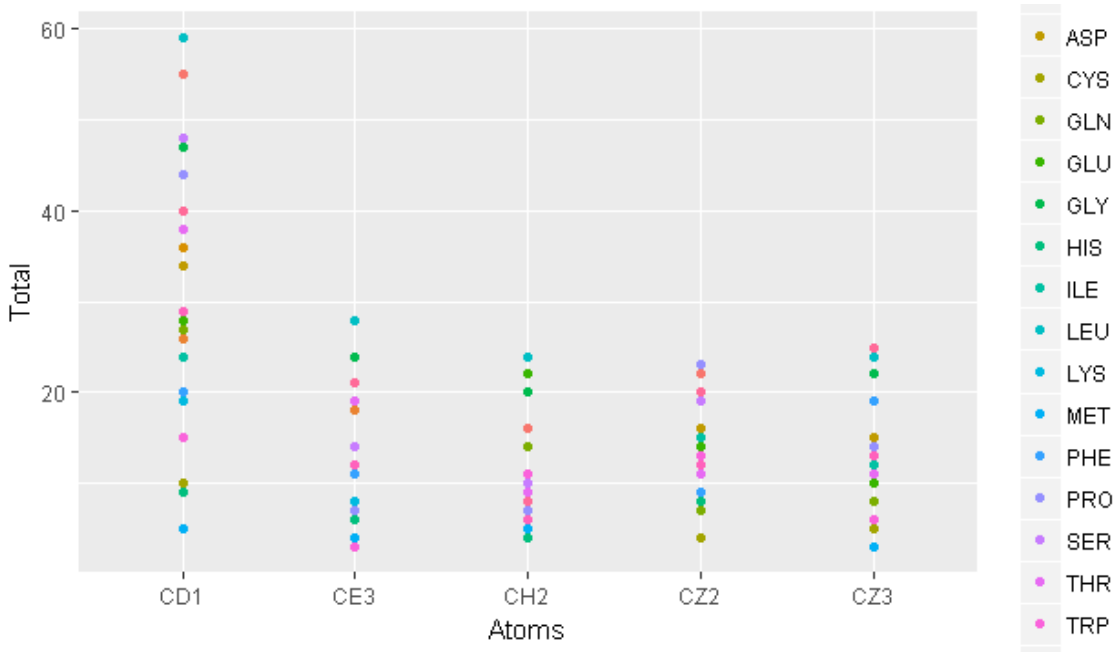


Figure 2.16 acceptor atoms main chain that are making hydrogen bond with CD1.

2.3.4 Case study.

First, In order to examine CD-HD1…O bonding, I have repeated all the study and the calculation I have done to analysis Trp-CD1 in 360 PDB files to Tyr-CD1, and Phe-CD1. After I run the program and calculated all expected hydrogen bond in Tyr, total number of hydrogen bonding in Tyr-CD1 52%, and Tyr-CE1 40% Fig (2.17).

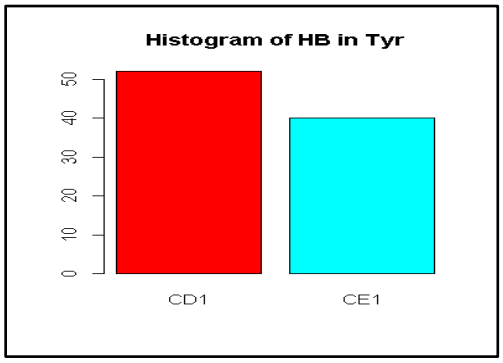


Figure 2.17 Histogram of the percentages of all expected hydrogen bonding I Tyr in the high resolution 360 PDB files.

In addition, in the study of hydrogen bonding in the atoms ring of Phe, total number of hydrogen bond in Phe-CD1 45%, Phe-CE1 27%, and Phe-CZ 26% see Fig (2.18).

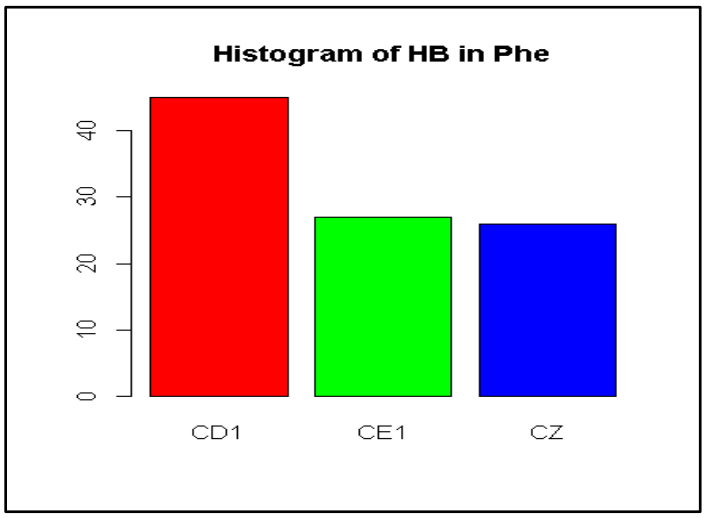
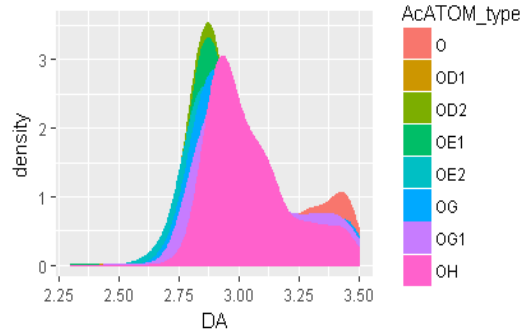


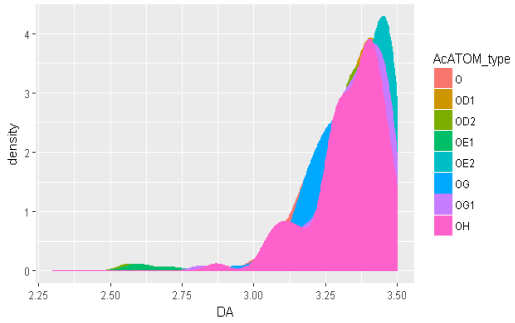
Figure 2.18 Histogram of the percentages of all expected hydrogen bonding Phe in the high resolution 360 PDB files.

Second, I selected a non-redundant set of 3700 high-resolution protein structures SCD dataset in order to study the CD1-HD1…O interactions of Trp rings. I calculated all expected hydrogen bond Trp-CD1-HD1….O in this data set using the same method to study the distance between CD1---O which is (DA) and the only acceptors satisfy the condition of regular hydrogen bond DA ≤3.5Å accepted as hydrogen bond Fig (2.19).

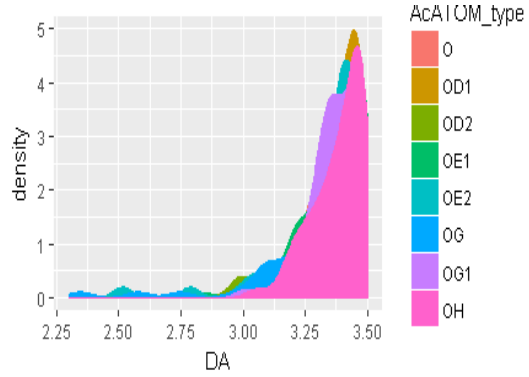
TRP-NE1…O



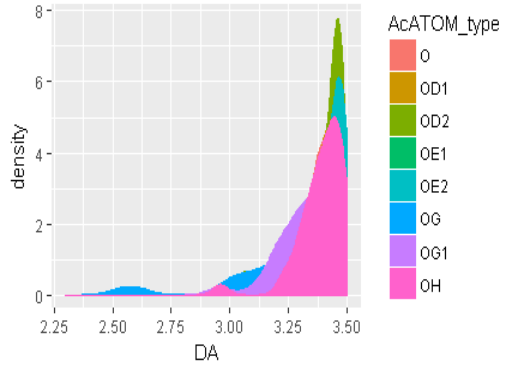
TRP-CD1…O



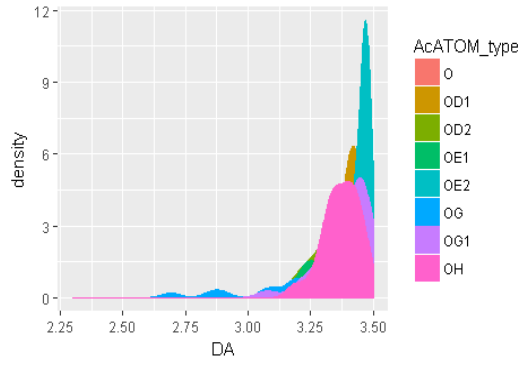
TRP-CZ2…O



TRP-CZ3…O



TRP-CE3…O



TRP-CH2…O

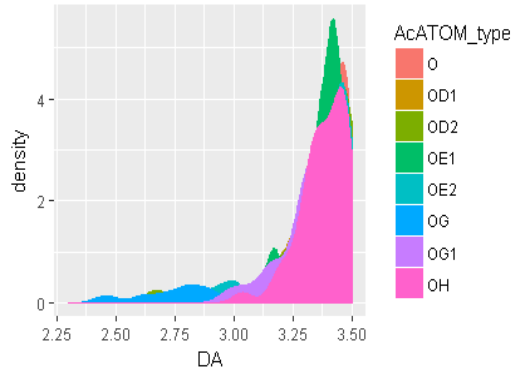


Figure 2.19 The DA of NE1-O, CD1-O, CZ2-O, CZ3-O, CE3, and CH2 in 3700 high-resolution protein structures SCD dataset.

As well, I have calculated all expected HB in NE1 and other carbon atoms in the ring to make comparison with CD1, CD1 make more hydrogen bond after NE1 and this support our previous results Table (2.2). Moreover, HBPLUS calculate DA of NE1 between **2.3-3.5Å**.I apply this condition to CD1 and other Carbon atoms in Trp and after apply this condition CD1 still show the higher number of hydrogen bond after NE1 Table (2.3), Fig (2.20).

Table 2.2 All expected HB in Trp from 3700 non-redundant single chains dataset.

Atoms	Total HB DA≤3.5	Percentages
TRP-NE1	6000	59%
TRP-CD1	5442	53%
TRP-CZ2	2212	22%
TRP-CZ3	1585	15%
TRP-CH2	1616	16%
TRP-CE3	1870	18%

Table 2.3 Condition of DA of NE1 in HBPLUS calculation which is ≤ **3.5 Å**.I apply this condition to CD1 and other Carbon atoms in Trp.

Atoms	Percentages
TRP-NE1	56%
TRP-CD1	39%
TRP-CZ2	14%
TRP-CZ3	9%
TRP-CH2	10%
TRP-CE3	11%

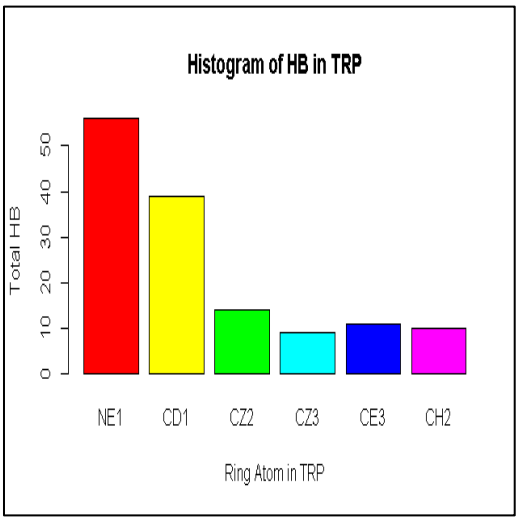


Figure 2.20 Distribution of percentages of expected hydrogen bond in ring of Trp from 3700 non-redundant single chains dataset.

In addition, alpha carbon considered as weak atom, and it can make hydrogen bond as many reports have been done in the past. Because of that, I have discovered all estimated HB in CA from3700 non-redundant single chains dataset, and then I compare the result with Trp-NE1 and Trp-CD1. CD1 was capable to make hydrogen bond after NE1 and more than CA. Therefore,CD1 second high charge and second high electronegativity atom in the ring comparing to other atoms in the ring of Trp Table(2.4), Fig(2.21).

Table 2.4 Percentages of all expected HB of Trp-CA, Trp-NE1, Trp-CD1 from3700 non-redundant single chains dataset

Atoms	*Total HB*	*Percentages*
TRP-NE1	5693	56%
TRP-CD1	3972	39%
CA of all amino acids	1625	17%

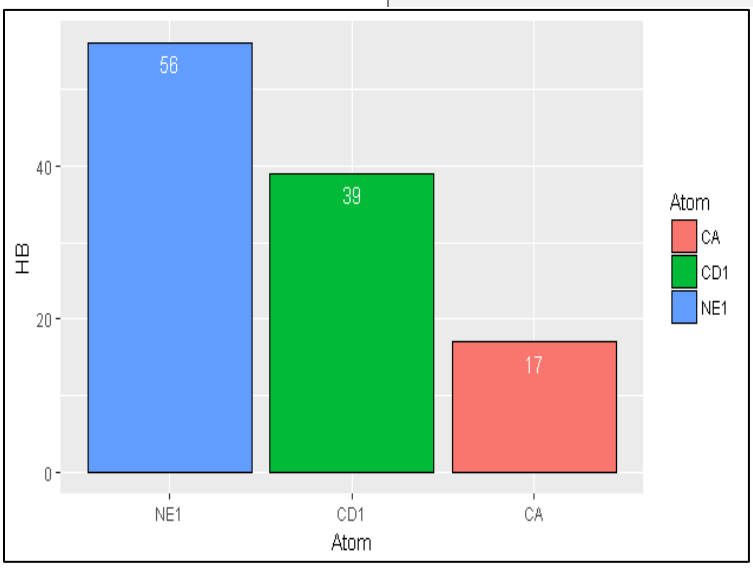


Figure 2.21 Histogram of all expected HB of Trp-CA, Trp-NE1, Trp-CD1 from3700 non-redundant single chains dataset.

Finally, In the analysis and study of hydrogen bonding geometry, I have repeated all the calculation I have done to analysis Trp-CD1 in the data 3700 non-redundant single chains dataset to look at Tyr-CD1, and Phe-CD1. After I run the program and calculated all expected hydrogen bond in Tyr, total number of hydrogen bonding in Tyr-CD1 23%, and Tyr-CE1 11% Fig (2.22).

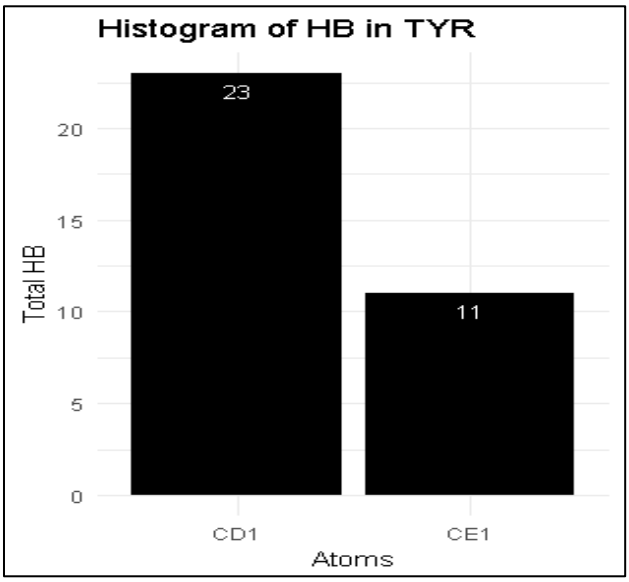


Figure 2.22 The percentages of all expected HB in CD1 and Other carbon atoms in Tyr.

Furthermore, I have calculated all expected HB in Phe from 3700 non-redundant single chains dataset. I have seen that Trp-CD1 has highest percentages in forming hydrogen bond, and that give more evidence to our results see Fig (2.23).

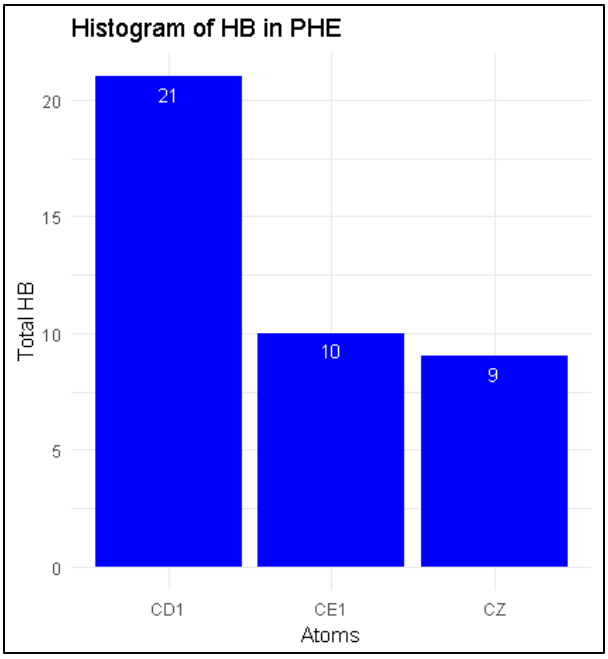
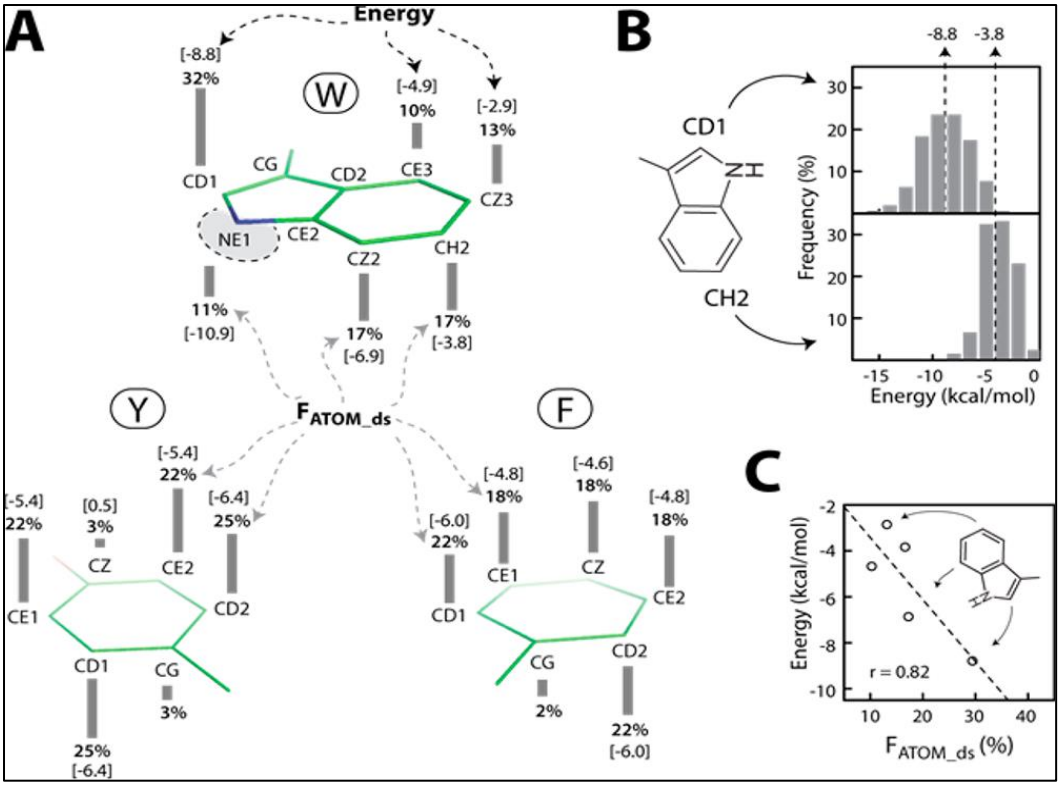


Figure 2.23 The Percentages of all HB in CD1 and Other carbon Atoms in Phe.

2.3.5 Interaction Energy.

In the study of anion−quadrupole interactions of the tyrosine (Tyr) and tryptophan (Trp) rings [25].We examined whether the computed energies are consistent with the frequency with which interactions occur in protein crystal structures[30, 33, 40]. For example, the FATOM_ds values for tyrosine, phenylalanine, and tryptophan atoms are consistent with the computed energy(e.g. FATOM_ds for Phe-CZ is 18% while that for Phe-CD1 is 22%) see Figures (2.24 A, B, C)[25].



Figures 2.24 Ring atom participation frequency and interaction energy.(A) The height of the bar (FATOM_ds) adjacent to the name of each atom represents the frequency (%) of its participation in anion− quadrupole interactions for that residue type, Trp (top), Tyr (bottom left), and Phe (bottom right). The number within the bracket above or below each bar represents the interaction energy (kilocalories per mole) when the indicated atom is closest to the interacting anion. (B) The distribution represents the energy of anion−quadrupole interactions when Trp-CD1 (top) and Trp-CH2 (bottom) atoms are closest to the interacting anion. The dotted lines with an arrow represent the mean interaction energy in kilocalories per mole. (C) Linear correlation (r = 0.82) between the energy and FATOM_ds for tryptophan ring carbon atoms. For a qualitative agreement between the energy and the frequency, we use FATOM_ds in place of log(FATOM_ds)[25]

In addition, the energy of an anion−quadrupole interaction based on quantum mechanical calculations depends not only on the interaction geometry but also on the ring atom[31, 37],the frequency of aromatic ring participation in anion−quadrupole. In The aspartate/glutamate anion− quadrupole interactions, Asp/Glu side chain carboxylates were the most frequent hydrogen bond acceptors for each type of side chain hydrogen bond donor in proteins [25].Since the energy calculation takes time to finish, I have showed the statistical analysis of the frequent hydrogen bond acceptors for each type of side chain, main chain hydrogen bond donor in proteins in the 3700 dataset that I have been used to study the bond geometry. In the statistical analysis, I have recognized that: first, most frequent hydrogen bond side chain and main chain showing in CD1 atom for Trp, Tyr, and Phe. Second, the aspartate/glutamate CD1 hydrogen bond interactions side chain carboxylates were the most frequent hydrogen bond acceptors see Fig (2.23) .Therefore, I have estimated that the energy calculation should consist with this statistical analysis as shown Fig (2.25).

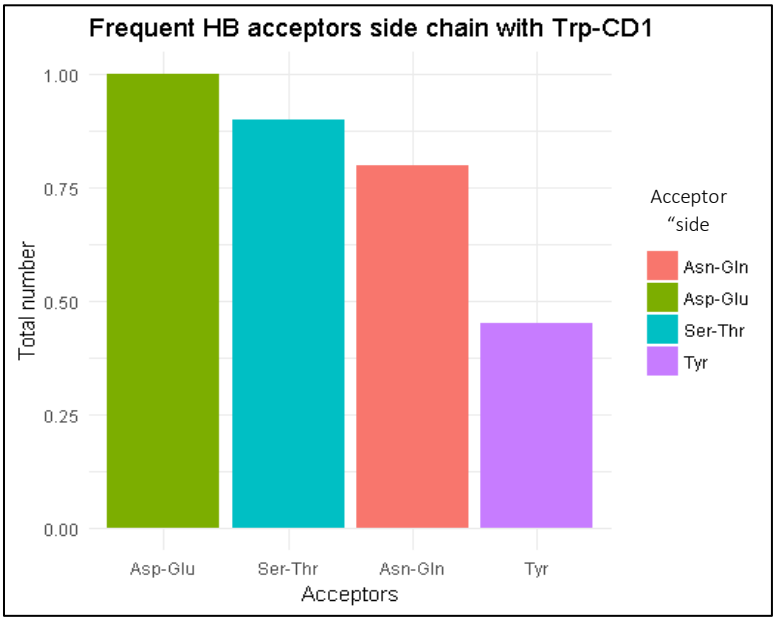


Figure 2.25 The most frequent hydrogen bond acceptors side chain

2.4. Conclusion.

The high-resolution set of 360 protein with a resolution ≤ 1 Å, and a non-redundant set of 3700 high-resolution protein structures SCD dataset were statistically analyzed by focusing mainly on the study of *aromatic*-proton mediated hydrogen bonds in terms of geometry; distance, and angle of the bond also, the statistical analysis of the frequent hydrogen bond, acceptors and the energy calculation of the hydrogen bond in anion quadrupole interaction. Therefore, the comprehensive analysis of *aromatic*-proton have been completed in this study illustrate the capability of Trp-CD1 to make a hydrogen bond after a Trp- NE1 atom. Additionally, this analysis validate the polarity of a HD1-atoms to be H-bond donor in CD1-HD1….O interaction.

Acknowledgments

The authors would like to acknowledge the South Dakota State University academic computing facility for the access to the Linux cluster for all the energy calculations. The authors would like to thank Suvobrata Chakravarty, Brian Moore, Jay Shore, and Adron R. Ung for all the energy calculation have been use in this study.

CHAPTER 3

3. Conclusions and Future Work

3.1 Conclusions:

In this study, the method was a program in R and Perl to study the geometries of CH---O hydrogen bonds in *aromatic* amino acids in a very high-resolution 360 protein structure with resolution ≤ 1 Å, and A non-redundant set of 3700 high-resolution protein structures SCD dataset. This method successfully finds all expected hydrogen bond contacts within distance DA ≤ 3.5 Å, and angle DHA ≤ 90°. This method with high simplicity and efficiency studied CH---O hydrogen bonds in *aromatic*-protons in a very big dataset. Additionally, the method reasonably give results that display the ability of a CD1 atom in the ring of tryptophan to make hydrogen bonds, were other available software to do that have difficulties in accomplishing that for two main reasons: (1) the data of protein that we use is very large, (2) many programs find the parameter of bonds in high electronegative atoms such as oxygen and nitrogen. However, the script still has small drawbacks that need to be overcome, such as timing the code takes 2 to 3 hours to finish the calculation when datasets are very large, and R software works very well with txt file or csv files. Therefore, we need to resave the PDB file in the appropriate file. In general, the program shows the capability to calculate all expected hydrogen bonds and I can just make a small change in the function that I used to speed up or convert the script that I use in this program to another language programs such as Perl, or C^{++}.

3.2. Future Work

I scientifically studied CH---O hydrogen bonding in Trp, Tyr, and Phe in a high-resolution, and non-redundant protein single chains, and I clearly recognize the capability of CD-HD1 in the rings of aromatic amino acids to make hydrogen bond. In addition, we have observed the usefulness of Trp-CD1-HD1 in designing effective antibiotics and antibodies. For example, in HIV structure[41], see Fig (3.1)[42], one study reported that the 4E10 was able to bind with the gp41 in HIV, see Fig(3.2)[24].Also, in the study of Peptidomimetics built inhibitor design for HIV, one of the results in this study examine the role of hydrogen bonds which have critical results that show six hydrogen bonds forming between the ligand and binding interface of HIV-gp120 [26].Therefore, my upcoming goal is to look at the interfaces protein−protein, DNA/RNA−protein, and membrane−protein.

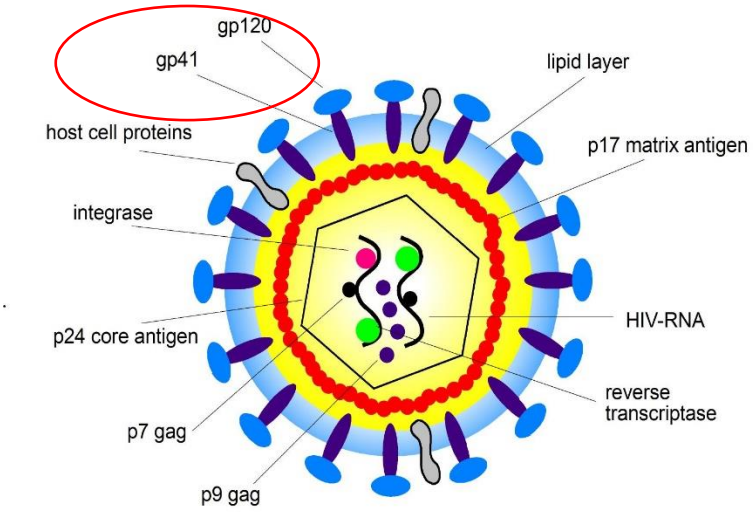


Figure 3.1 Structure of an HIV[42].

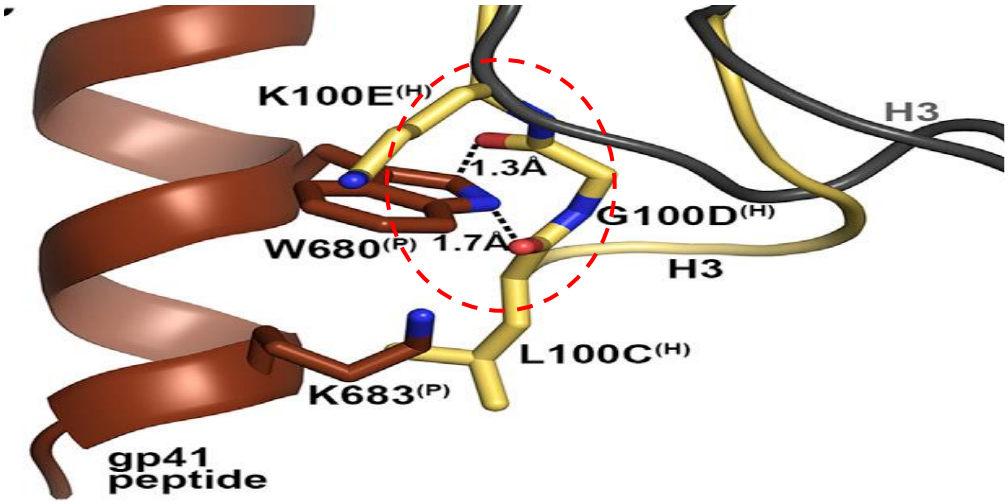


Figure 3. 2 Lipid Binding Site in 4E10 CDRH1 Region in Co-crystal Structures [24, 25].

In addition, building a webserver is very important since there is a need for programs to calculate the parameter of CH---O hydrogen bond in a large datasets. Then, the webserver will help many researchers to analyze the CH---O hydrogen bonds in ether aromatic or aliphatic protons.

References:

1. Chopra, P.P.a.D., *Understanding of Noncovalent Interactions Involving Organic Fluorine.* 2015: p. 37-61.
2. Seiji Tsuzuki, K.H., Tadafumi Uchimaru, Masuhiro Mikami, and and K. Tanabe, *The Magnitude of the CH/π Interaction between Benzene and Some Model Hydrocarbons.* ACS, 2000. **122**: p. 3746-3753.
3. Sethumadhavan*, V.S.a.R., *Effect of C-H...O Interactions in the Structural Stability of "All-Alpha" Proteins.* ChemTech, 2011. **Vol.3**: p. PP 348-356.
4. Daniel Boamah, T.L., Franchesca A. Poppinga, Shraddha Basu,† Shahariar Rahman,† and a.S.C. Francisca Essel, *Characteristics of a PHD Finger Subtype.* ACS, 2018. **57**: p. 525−539.
5. C. Nick Pace, 2* Hailong Fu,1 Katrina Lee Fryar,2 John Landua,2, et al., *Contribution of hydrogen bonds to protein stability.* the protein society March 2014.
6. Alok Jain, V.R., 2, a.R. Sankararamakrishnan1, and *, *Lone pair ··· π interactions between water oxygens and aromatic residues: Quantum chemical studies based on high-resolution protein structures and model compounds.* PROTEIN SCIENCE, 2009. **18**: p. 595-605.
7. Lai, L.J.a.L., *CH---O Hydrogen Bonds at Protein-Protein Interfaces.* JBC 2002. **277**(4): p. 37732–37740.
8. Lai‡, L.J.a.L., *CH....O Hydrogen Bonds at Protein-Protein Interfaces.* BIOLOGICAL CHEMISTRY, 2002. **Vol. 277**(40): p. 37732–37740.

9. SANDRO MECOZZI, A.P.W., JR., AND DENNIS A. DOUGHERTY*, *Cation-1 interactions in aromatics of biological and medicinal interest: Electrostatic potential surfaces as a useful qualitative guide.* PNAS. **Vol. 93**: p. 10566-10571.

10. Thomas Steiner , G.K., *Hydrogen bonds with π-acceptors in proteins: frequencies and role in stabilizing local 3D structures1.* 2001. **305**(3): p. 535-557.

11. Ays¸ eg€ul O¨ zen, 2 Lionel Rouge´ ,3 Charlene Bashore,1 Brian R. Hearn,4 Nicholas J. Skelton,2,* and Erin C. Dueber1,5,*, *Selectively Modulating Conformational States of*

USP7 Catalytic Domain for Activation. cell press, January 2018. **Volume 26**(issue 1): p. p72–84.e7, 2.

12. Hisae Tateishi-Karimata1, Tatsuya Ohyama1,†, Takahiro Muraoka2, Peter Podbevsek3,, S.T. Adam M. Wawro2, Shu-ichi Nakano5, Kazushi Kinbara2, Janez Plavec3, and and Naoki Sugimoto1, *Newly characterized interaction stabilizes DNA structure: oligoethylene glycols stabilize G-quadruplexes CH-π interactions.* Nucleic Acids Res, 2017. **45**(12): p. 7021–7030.

13. Rich, A.H.a.A., *The Biology of Left-handed*

Z-DNA. JBC, 1996. **271**(17): p. 11595–11598.

14. GESSNER§, M.E.A.R.V., *Stereoelectronic effects of deoxyribose 04' on DNA conformation.* Proc. Natl. Acad. Sci. USA January 1995. **Vol. 92**: p. 180-184.

15. Shayantani Mukherjee, S.M., † and Dhananjay Bhattacharyya*, *Role of Hydrogen Bonds in Protein-DNA Recognition: Effect of Nonplanar Amino Groups.* J. Phys, March 17, 2005. **109**: p. 10484-10492.

16. Yael Mandel-Gutfreund1, H.M., Robert L. Jernigan2 and a.V.B. Zhurkin2*, *A Role for CH .....O Interactions in*

Protein-DNA Recognition. JMB, 1998. **277**: p. 1129-1140.

17. Yael Mandel-Gutfreund1, H.M., Robert L. Jernigan2 and a.V.B. Zhurkin2*, *A Role for CH.....O Interactions in*

Protein-DNA Recognition. JMB, 1998. **277**: p. 1129-1140.

18. BZhurkin2, Y.-G.H.R.L.V., *A role for CH…O interactions in protein-DNA recognition.* Journal of Molecular Biology, 1998 **Volume 277**(5): p. Pages 1129-1140.

19. Modan K Das 1, a.H.-K.D., *A survey of DNA motif finding algorithms.* BMC Bioinformatics 2007.

20. Xing Du 1, Yi Li 1,†, Yuan-Ling Xia 1, et al., *Insights into Protein–Ligand Interactions:Mechanisms, Models, and Methods.* Int. J. Mol. Sci, 2016. **17**(2).

21. Tomonaga Ozawa *, E.T., Motoyasu Ozawa, Chiaki Handa, Harunobu Mukaiyama, and S.K. Toshihiro Nishimura, Kosuke Okazaki, *The importance of CH/p hydrogen bonds in rational drug design: An ab*

initio fragment molecular orbital study to leukocyte-specific protein

tyrosine (LCK) kinase. 2008. **16**: p. 10311–10318.

22. Anbarasu, P.L.S.R.A., *Influence of C-H. . .O interactions on the structural stability of β-lactamases.* JBP, 25 June 2013. **39**: p. 649–663.

23. Hang Yin1, a.A.D.F., 3, *Drugging Membrane Protein Interactions.* PMC, 2016 Feb 5.

24. Adriana Irimia, A.S., Robyn L. and I.A.W. Stanfield, *Crystallographic Identification of Lipid as an Integral*

Component of the Epitope of HIV Broadly

Neutralizing Antibody 4E10. cell press, 2016 January. **44**: p. 21–31.

25. Suvobrata Chakravarty, A.R.U., Brian G Moore, Jay S. Shore, and Mona Alshamrani, *A Comprehensive Analysis of Anion–Quadrupole Interactions in Protein Structures.* acs.biochem, 2018.

26. Umashankar Vetrivel1* , P.S. and G.S. Naveen kumar Nagarajan2, *Peptidomimetics Based Inhibitor Design*

for HIV – 1 gp120 Attachment Protein. JPB, November 2009. **Vol.2**.

27. Shahbazi*, Z., *Mechanical Model of Hydrogen Bonds in Protein Molecules.* American Journal of Mechanical Engineering,, 2015. **3**: p. 47-54.

28. Shahbazi*, Z., *Mechanical Model of Hydrogen Bonds in Protein Molecules.* American Journal of Mechanical Engineering,, 2015,. **Vol.3**: p. 47-54.

29. Rusczyk, R., *Introduction to Geometry.* 2007. **1**.

30. Mutasem Omar Sinnokrot†, a.C.D.S., *High-Accuracy Quantum Mechanical Studies of $\pi-\pi$ Interactions in Benzene Dimers.* ACS, 2006. **110**: p. 10656–10668.

31. Andrei V. Afonin, a.A.V.V.a.M.V.S.b., *Estimating the energy of intramolecular hydrogen bonds from 1H NMR and QTAIM calculations.* 2016(Issue 47).

32. Mondal A1, D.S., *Quantum mechanical electronic structure calculation reveals orientation dependence of hydrogen bond energy in proteins.* Proteins, 2017. **85**: p. 1046–1055.

33. Kwang S. Kim*, P.T., and Jin Yong Lee, *Molecular Clusters of π-Systems: Theoretical Studies of Structures, Spectra, and Origin of Interaction Energies.* ACS, 2000. **100**: p. 4145-4148.

34. Keller, J., *Introduction to Gaussian.* 2008: p. 1-37.

35. Chenzhong Liao1, Markus Sitzmann1, Angelo Pugliese1, and Marc C Nicklaus1, *Software and resources for computational medicinal chemistry.* NIH, 2011. **8**: p. 1057–1085.

36. Houk, S.E.W.a.K.N., *Are Anion/π Interactions Actually a Case of Simple Charge-Dipole Interactions?* ACS, 2010. **114**: p. 8658–8664.

37. Sławomir J. Grabowski, †,‡ W. Andrzej Sokalski,§ Edyta Dyguda,§ and Jerzy Leszczyn´ski, *Quantitative Classification of Covalent and Noncovalent H-Bonds.* ACS, 2006. **110**: p. 6444-6446.

38. Sciences, I.J.o.P.a.P., *CH...O INTERACTIONS STABILIZE THE STRUCTURE OF THE THERAPEUTIC PROTEINS: A COMPUTATIONAL STUDY.* International Journal of Pharmacy and Pharmaceutical Sciences, 2011. **3**(3): p. ISSN- 0975-1491.

39. Pierce AC1, S.K., Bemis GW., *Kinase inhibitors and the case for CH...O hydrogen bonds in protein-ligand binding.* Proteins: Structure, Function, and Bioinformatics, 2002. **49**(4): p. 567–576.

40. Jr, J.H.A.D.M., *CHARMM36 all-atom additive protein force field: validation based on comparison to NMR data.* Journal of Computational Chemistry, 2013. **34**(25): p. 2135-2145.

41. Yunji W. Davenport, A.P.W., Jr., Pamela J. Bjorkman, *Structure of an HIV-2 gp120 in Complex with CD4.* Journal of Virology, 2016. **90**.

42. Andrea Rubbert, G.B.a.M.O., *Pathogenesis of HIV-1 Infection.* 2011.

Appendix 1 :

A.1.1 Input file extract from (PDB file)

Residue	Atom name	Amino	x	y	z	ATOM	Filename
1	N	VAL	-5.066	0.058	13.305	N	1AHOA.pdb
1	O	VAL	-3.354	2.058	13.22	O	1AHOA.pdb
2	N	LYS	-2.935	1.92	11.013	N	1AHOA.pdb
2	O	LYS	-2.14	3.055	8.683	O	1AHOA.pdb
2	NZ	LYS	3.015	-0.908	9.879	N	1AHOA.pdb
3	N	ASP	-0.952	4.628	9.7	N	1AHOA.pdb
3	O	ASP	1.633	5.085	9.122	O	1AHOA.pdb
3	OD1	ASP	-3.28	6.409	8.885	O	1AHOA.pdb
3	OD2	ASP	-2.501	8.311	9.661	O	1AHOA.pdb
4	N	GLY	1.236	5.367	6.913	N	1AHOA.pdb
4	O	GLY	2.152	4.86	4.32	O	1AHOA.pdb
5	N	TYR	4.317	5.073	4.948	N	1AHOA.pdb
5	O	TYR	5.649	2.562	4.175	O	1AHOA.pdb
5	OH	TYR	6.666	10.964	3.42	O	1AHOA.pdb
5	HD1	TYR	5.81	7.066	1.498	H	1AHOA.pdb
6	N	ILE	4.167	2.716	2.493	N	1AHOA.pdb
6	O	ILE	6.11	1.571	0.867	O	1AHOA.pdb
6	HD11	ILE	2.075	0.281	-1.143	H	1AHOA.pdb

6	HD12	ILE	0.885	1.098	-0.401	H	1AHOA.pdb
6	HD13	ILE	1.693	1.807	-1.637	H	1AHOA.pdb
7	N	VAL	5.886	-0.382	1.947	N	1AHOA.pdb
7	O	VAL	5.968	-3.045	1.368	O	1AHOA.pdb
8	N	ASP	7.852	-2.885	0.152	N	1AHOA.pdb
8	O	ASP	8.789	-4.536	2.05	O	1AHOA.pdb
8	OD1	ASP	10.659	-3.821	-0.483	O	1AHOA.pdb
8	OD2	ASP	10.724	-4.16	-2.677	O	1AHOA.pdb
9	N	ASP	8.899	-6.313	0.69	N	1AHOA.pdb
9	O	ASP	11.365	-7.556	2.999	O	1AHOA.pdb
9	OD1A	ASP	10.052	-8.373	-1.046	O	1AHOA.pdb
9	OD1B	ASP	7.259	-9.001	-0.022	O	1AHOA.pdb
10	N	VAL	11.539	-5.98	1.374	N	1AHOA.pdb
10	O	VAL	14.076	-3.57	1.917	O	1AHOA.pdb
11	N	ASN	11.936	-3.517	2.483	N	1AHOA.pdb
11	O	ASN	12.527	0.005	2.272	O	1AHOA.pdb
11	OD1	ASN	11.592	-0.56	5.466	O	1AHOA.pdb
11	ND2	ASN	13.78	-0.071	5.346	N	1AHOA.pdb
12	N A	CYS	11.777	-1.274	0.601	N	1AHOA.pdb
12	N B	CYS	11.503	-1.503	0.812	N	1AHOA.pdb
12	O A	CYS	9.398	-0.462	-0.917	O	1AHOA.pdb
12	O B	CYS	9.28	-0.529	-0.459	O	1AHOA.pdb
13	N	THR	10.311	1.539	-0.709	N	1AHOA.pdb

13	O	THR	9.341	1.771	-3.313	O	1AHOA.pdb
13	OG1	THR	10.462	4.274	-1.095	O	1AHOA.pdb
14	N	TYR	7.335	2.541	-2.624	N	1AHOA.pdb
14	O	TYR	6.256	4.851	-4.232	O	1AHOA.pdb
14	OH	TYR	3.108	-2.87	-2.523	O	1AHOA.pdb
14	HD1	TYR	5.689	0.639	-1.651	H	1AHOA.pdb
15	N	PHE	7.615	3.844	-5.714	N	1AHOA.pdb
15	O	PHE	5.604	4.519	-7.468	O	1AHOA.pdb
15	HD1	PHE	10.063	7.004	-7.486	H	1AHOA.pdb
16	N	CYS	6.165	6.679	-7.332	N	1AHOA.pdb
16	O	CYS	6.005	9.231	-8.253	O	1AHOA.pdb
17	N	GLY	4.113	8.823	-9.417	N	1AHOA.pdb
17	O	GLY	2.387	11.926	-9.6	O	1AHOA.pdb
18	N	ARG	1.558	9.847	-9.618	N	1AHOA.pdb
18	O	ARG	-0.307	8.714	-7.772	O	1AHOA.pdb
18	NE	ARG	-1.042	8.024	-13.088	N	1AHOA.pdb
18	NH1	ARG	1.237	7.847	-13.407	N	1AHOA.pdb
18	NH2	ARG	-0.022	5.96	-13.116	N	1AHOA.pdb
19	N	ASN	-1.363	10.688	-7.674	N	1AHOA.pdb
19	O	ASN	-3.066	8.242	-5.771	O	1AHOA.pdb
19	OD1	ASN	-0.84	12.35	-4.997	O	1AHOA.pdb
19	ND2	ASN	-2.493	13.807	-5.344	N	1AHOA.pdb
20	N	ALA	-3.65	9.042	-7.798	N	1AHOA.pdb

20	O	ALA	-4.528	5.608	-7.577	O	1AHOA.pdb
21	N	TYR	-2.707	6.534	-8.541	N	1AHOA.pdb
21	O	TYR	-1.952	3.665	-6.735	O	1AHOA.pdb
21	OH	TYR	2.947	1.179	-8.287	O	1AHOA.pdb
21	HD1	TYR	-0.564	2.998	-10.246	H	1AHOA.pdb
22	N	CYS	-1.3	5.749	-6.206	N	1AHOA.pdb
22	O	CYS	-2.425	4.122	-3.294	O	1AHOA.pdb
23	N	ASN	-3.474	5.741	-4.437	N	1AHOA.pdb
23	O	ASN	-5.504	3.201	-3.219	O	1AHOA.pdb
23	OD1	ASN	-6.971	6.627	-2.18	O	1AHOA.pdb
23	ND2	ASN	-8.149	5.68	-3.809	N	1AHOA.pdb
24	N	GLU	-4.881	3.516	-5.341	N	1AHOA.pdb
24	O	GLU	-4.678	0.155	-4.494	O	1AHOA.pdb
24	OE1A	GLU	-4.535	1.98	-10.096	O	1AHOA.pdb
24	OE1B	GLU	-5.56	1.17	-9.874	O	1AHOA.pdb
24	OE2A	GLU	-5.987	3.569	-10.363	O	1AHOA.pdb
24	OE2B	GLU	-5.825	-0.966	-9.404	O	1AHOA.pdb
25	N	GLU	-2.903	1.441	-4.966	N	1AHOA.pdb
25	O	GLU	-2.198	-0.652	-2.205	O	1AHOA.pdb
25	OE1	GLU	0.121	-1.596	-5.871	O	1AHOA.pdb
25	OE2	GLU	-0.773	-0.812	-7.683	O	1AHOA.pdb
26	N	CYS	-2.686	1.527	-2.178	N	1AHOA.pdb
26	O	CYS	-4.387	-0.002	0.502	O	1AHOA.pdb

27	N	THR	-5.286	0.822	-1.381	N	1AHOA.pdb
27	O	THR	-7.026	-2.268	-0.904	O	1AHOA.pdb
27	OG1	THR	-7.464	0.694	-3.28	O	1AHOA.pdb
28	N	LYS	-5.405	-1.773	-2.375	N	1AHOA.pdb
28	O	LYS	-4.919	-5.054	-1.109	O	1AHOA.pdb
28	NZ	LYS	-2.174	-3.122	-8.245	N	1AHOA.pdb
29	N	LEU	-3.945	-3.106	-0.474	N	1AHOA.pdb
29	O	LEU	-4.122	-3.584	3.124	O	1AHOA.pdb
29	HD11	LEU	-0.091	-1.441	0.995	H	1AHOA.pdb
29	HD12	LEU	0.92	-2.715	0.884	H	1AHOA.pdb
29	HD13	LEU	0.543	-1.865	-0.455	H	1AHOA.pdb
30	N	LYS	-5.687	-2.978	1.623	N	1AHOA.pdb
30	O	LYS	-7.351	-1.4	4.412	O	1AHOA.pdb
31	N	GLY	-5.875	-0.579	2.961	N	1AHOA.pdb
31	O	GLY	-7.834	1.093	2.161	O	1AHOA.pdb
32	N	GLU	-7.359	2.624	3.705	N	1AHOA.pdb
32	O	GLU	-9.139	4.301	1.198	O	1AHOA.pdb
32	OE1	GLU	-8.544	7.631	4.503	O	1AHOA.pdb
32	OE2	GLU	-10.213	7.269	5.813	O	1AHOA.pdb
33	N	SER	-7.085	4.864	1.908	N	1AHOA.pdb
33	O	SER	-4.61	5.66	1.789	O	1AHOA.pdb
33	OG	SER	-7.135	7.787	2.147	O	1AHOA.pdb
34	N	GLY	-4.831	6.815	-0.123	N	1AHOA.pdb

34	O	GLY	-4.036	8.514	-2.049	O	1AHOA.pdb
35	N	TYR	-1.902	8.672	-1.268	N	1AHOA.pdb
35	O	TYR	0.677	8.786	-1.814	O	1AHOA.pdb
35	OH	TYR	1.521	12.888	2.359	O	1AHOA.pdb
35	HD1	TYR	-1.956	10.61	0.829	H	1AHOA.pdb
36	N	CYS	0.459	10.184	-3.56	N	1AHOA.pdb
36	O	CYS	2.291	12.516	-3.602	O	1AHOA.pdb
37	N	GLN	3.423	11.07	-2.271	N	1AHOA.pdb
37	O	GLN	6.363	11.613	-2.559	O	1AHOA.pdb

A.1.2 output file(hydrogen bond parameters)

Residue	Amino	ATOM	Ac residue	Ac amino	Ac ATOM	DH	DAH	DA	file
318	TRP	CD1	322	GLU	OE1A	0.922665	128.0838	2.713077	4N8GA.pdb
318	TRP	CD1	322	GLU	OE2B	0.923985	107.1814	2.772068	4N8GA.pdb
311	TRP	CD1	307	THR	O	0.925476	117.3123	2.965026	4YEPA.pdb
234	TRP	CD1	237	ASP	O	0.925804	138.9424	2.990718	4IILA.pdb
257	TRP	CD1	255	VAL	O	0.925864	96.87643	2.996869	4HNOA.pdb
32	TRP	CD1	31	SER	OG B	0.926417	108.0759	2.999742	3I94A.pdb
669	TRP	CD1	716	SER	O B	0.926901	120.4121	3.035548	4QLPB.pdb
7	TRP	CD1	168	VAL	O B	0.926983	139.1768	3.036643	2XU3A.pdb
155	TRP	CD1	151	LEU	O	0.927043	103.4604	3.039667	4PSRA.pdb
348	TRP	CD1	344	ASN	O	0.927089	110.574	3.041127	4QRNA.pdb
314	TRP	CD1	310	ALA	O	0.927101	98.6395	3.05029	4UABA.pdb
157	TRP	CD1	153	ALA	O	0.927389	102.864	3.055033	4FK9A.pdb
153	TRP	CD1	328	TYR	OH	0.927496	119.0355	3.061346	4JN7A.pdb
269	TRP	CD1	265	ARG	O A	0.927572	107.9995	3.068336	1MUWA.pdb
94	TRP	CD1	90	ASP	O	0.927648	133.9276	3.070831	4JN7A.pdb
84	TRP	CD1	76	LEU	O	0.927677	121.3262	3.075234	4HNOA.pdb
149	TRP	CD1	102	ALA	O	0.927744	120.7251	3.078645	5C5GA.pdb
266	TRP	CD1	261	PRO	O	0.927766	115.59	3.078825	5EJ8H.pdb
81	TRP	CD1	106	ALA	O	0.927809	157.3049	3.083514	4C08A.pdb
150	TRP	CD1	157	GLY	O	0.927911	112.2491	3.084567	4Z0YB.pdb
238	TRP	CD1	233	GLN	O	0.927916	97.90527	3.090767	5FLYB.pdb
305	TRP	CD1	307	ASN	OD1	0.927935	124.2581	3.093184	4QP5A.pdb
669	TRP	CD1	716	SER	O A	0.928119	119.6893	3.098021	4QLPB.pdb
109	TRP	CD1	108	ASP	OD2A	0.928175	122.8109	3.113903	1MJ5A.pdb
251	TRP	CD1	250	THR	OG1	0.928265	145.1888	3.114068	5JDAA.pdb
119	TRP	CD1	115	ASP	O	0.928291	140.6662	3.119354	4KEMA.pdb
108	TRP	CD1	56	LEU	O	0.928382	101.6619	3.121414	2VB1A.pdb
159	TRP	CD1	157	SER	OG B	0.928476	153.601	3.125962	4C08A.pdb

340	TRP	CD1	337	GLN	O	0.928484	119.1491	3.126779	4WFOA.pdb
402	TRP	CD1	398	TRP	O	0.92857	150.7976	3.127858	4G1QB.pdb
165	TRP	CD1	164	LEU	O	0.928632	134.4854	3.12955	5INJA.pdb
156	TRP	CD1	152	TYR	O	0.92864	126.3306	3.133462	4PP4A.pdb
406	TRP	CD1	408	ALA	O	0.928664	107.6677	3.135051	4G1QB.pdb
194	TRP	CD1	190	THR	O	0.9287	126.4022	3.135355	4XQCA.pdb
343	TRP	CD1	362	THR	OG1	0.9287	113.4852	3.137228	4U9HL.pdb
218	TRP	CD1	216	ASP	O	0.928777	137.8681	3.137605	4ZBOA.pdb
261	TRP	CD1	294	THR	O	0.928813	137.6067	3.139696	5E9NA.pdb
97	TRP	CD1	96	LEU	O	0.92882	105.5705	3.143592	1W7CA.pdb
235	TRP	CD1	232	ARG	O	0.928821	91.21711	3.143779	4BJIA.pdb
118	TRP	CD1	112	LEU	O	0.928914	112.0337	3.143984	3WA2X.pdb
88	TRP	CD1	85	LYS	O	0.928971	138.9742	3.145527	4Z3GA.pdb
222	TRP	CD1	220	TYR	O	0.92898	98.72851	3.149459	4QOSA.pdb
292	TRP	CD1	288	GLU	O	0.928984	118.3828	3.151632	3U65B.pdb
284	TRP	CD1	282	THR	OG1	0.928989	123.2822	3.153396	2FVYA.pdb
574	TRP	CD1	570	SER	O	0.929013	101.2095	3.154491	4C6FA.pdb
153	TRP	CD1	81	ASN	OD1	0.929027	133.9615	3.155272	4G1QB.pdb
32	TRP	CD1	168	TYR	O	0.929047	154.7447	3.160221	4EYZA.pdb
299	TRP	CD1	295	THR	O	0.929052	153.0318	3.160936	4JCCA.pdb
570	TRP	CD1	219	ASP	OD2	0.929069	102.6756	3.163705	4UDXX.pdb
59	TRP	CD1	55	TYR	O	0.929097	114.3212	3.164244	5COFA.pdb
223	TRP	CD1	181	LEU	O	0.929097	127.7496	3.164362	5HWQA.pdb
223	TRP	CD1	181	LEU	O	0.929098	128.048	3.165723	5HWOA.pdb
891	TRP	CD1	886	ARG	O B	0.929115	135.0887	3.167322	5C30A.pdb
284	TRP	CD1	282	ASP	O	0.92912	124.9162	3.168328	5DN9A.pdb
72	TRP	CD1	71	PHE	O	0.929122	143.1288	3.168702	4QQSA.pdb
169	TRP	CD1	165	ARG	O	0.929125	137.3148	3.170214	5CTMA.pdb
149	TRP	CD1	145	THR	O	0.929147	98.36793	3.172853	5C5GA.pdb
71	TRP	CD1	69	ASP	O	0.929147	103.0402	3.173134	2AXWA.pdb
94	TRP	CD1	90	ASP	OD1	0.929151	125.886	3.174327	4JN7A.pdb

69	TRP	CD1	105	THR	OG1	0.929151	108.6059	3.175796	5CL8A.pdb
233	TRP	CD1	228	SER	O	0.929156	118.4408	3.176611	5A99A.pdb
120	TRP	CD1	117	GLU	O	0.929156	161.4718	3.176729	4CDJA.pdb
133	TRP	CD1	129	GLU	O	0.929157	154.8317	3.177471	4PSSA.pdb
147	TRP	CD1	179	SER	OG	0.929158	113.6795	3.179276	2YN0A.pdb
227	TRP	CD1	223	ASN	O	0.929159	140.3251	3.183185	4JN7A.pdb
455	TRP	CD1	451	ASN	O	0.929174	91.73961	3.18376	1LLFA.pdb
305	TRP	CD1	87	SER	OG	0.929183	108.1094	3.184883	4FK9A.pdb
318	TRP	CD1	314	ARG	O	0.929186	135.5485	3.186379	2JLIA.pdb
308	TRP	CD1	304	THR	O	0.929187	155.4781	3.18752	4MCOA.pdb
337	TRP	CD1	335	ARG	O	0.929201	104.9628	3.189624	1JFBA.pdb
45	TRP	CD1	43	GLN	O	0.929201	112.6999	3.189906	4REKA.pdb
239	TRP	CD1	235	ASN	O	0.929202	154.9397	3.189961	5FC1A.pdb
389	TRP	CD1	325	THR	OG1	0.929208	108.5691	3.191436	4BQYA.pdb
683	TRP	CD1	696	SER	O	0.929213	163.3947	3.193117	4N5UA.pdb
118	TRP	CD1	72	TRP	O	0.929214	109.6235	3.193426	4L9OA.pdb
455	TRP	CD1	451	ASN	OD1	0.929218	104.8101	3.194255	1LLFA.pdb
28	TRP	CD1	24	SER	O A	0.929219	123.9703	3.194823	2VB1A.pdb
79	TRP	CD1	75	ARG	O	0.929234	155.6359	3.195251	4PJ2C.pdb
206	TRP	CD1	207	PRO	O	0.929235	141.0161	3.196062	4MNCA.pdb
203	TRP	CD1	124	ASP	OD2	0.929235	128.9655	3.196912	1Y93A.pdb
59	TRP	CD1	54	LEU	O	0.929239	108.2549	3.198957	5E75A.pdb
243	TRP	CD1	240	ASN	O	0.929241	137.5971	3.198983	3ZOJA.pdb
250	TRP	CD1	252	ILE	O	0.929252	117.2502	3.19957	1IXHA.pdb
71	TRP	CD1	68	SER	O	0.929258	139.5513	3.200212	4D0QA.pdb
148	TRP	CD1	146	ARG	O	0.929261	99.80256	3.204978	2NLRA.pdb
215	TRP	CD1	214	SER	OG	0.929263	168.5827	3.20589	5JBBS.pdb
27	TRP	CD1	23	GLN	O	0.929263	124.0881	3.207185	1TT8A.pdb
212	TRP	CD1	196	GLY	O	0.929265	119.0956	3.208561	4XDUA.pdb
76	TRP	CD1	74	ASN	O	0.929267	114.7568	3.20933	4UE8A.pdb
215	TRP	CD1	214	SER	OG	0.929272	167.5711	3.209393	5JB9S.pdb

344	TRP	CD1	355	ASN	O	0.929272	110.2114	3.212218	4X9YA.pdb
529	TRP	CD1	316	THR	O	0.929272	109.509	3.213004	4D7CA.pdb
119	TRP	CD1	115	ASP	OD1	0.929273	132.1677	3.213836	4KEMA.pdb
327	TRP	CD1	313	LEU	O	0.929274	118.4685	3.214053	4FK9A.pdb
17	TRP	CD1	13	PRO	O	0.929278	152.6694	3.214665	3NZLA.pdb
196	TRP	CD1	181	GLN	O	0.929278	118.7033	3.215138	4MTMA.pdb
294	TRP	CD1	306	SER	OG	0.929281	115.7075	3.215153	4X33B.pdb
192	TRP	CD1	150	ASP	O	0.929286	146.8763	3.21578	4YAPA.pdb
303	TRP	CD1	300	ASN	O	0.929287	111.0369	3.218121	1W7CA.pdb