CHAPTER 3

MOLECULAR MODELING OF COLLAGEN CROSSLINKING WITH D-AMINO ACIDS

3.1 INTRODUCTION

Development of computational tools for rapid screening of potential candidates for novel polymeric biomaterials is very important for drug delivery and tissue engineering. Top pharmaceutical companies use both biophysical and computational methods for small ligand screening for crosslinker, and drug discovery and design. Using these computational approaches including molecular dynamic simulations combined with docking and data mining, new polymeric biomaterials are developed. This pre-screening approach of large polymer libraries helps in development of interesting “lead” biomaterials as drug carriers and for tissue engineering etc, reducing the time and cost efforts as part of rational design process. The Protein Data Bank and Pubchem Compound Database have been reliable, rich source of information on molecular interactions and their role in protein stability and structure.

In a typical virtual screening experiment, ligands can be generated de novo using combinatorial chemistry or taken from a library of chemical compounds. The resulting ensemble of thousands to millions of small molecules are then optimally docked to the target protein and subsequently ranked according to their calculated binding energies. There are two main bottlenecks in this procedure. The first one is related to the propensity to
generate native or native-like docking poses and understanding of biomolecular recognition and the methods used for sampling the conformations of proteins and low molecular weight molecules separately. The second bottleneck in virtual screening is related to the correlation between the experimental and predicted binding affinities. Numerous studies predict false-positives or compounds with poor affinities. In the second, particular emphasis is placed on advanced methods for predicting three-dimensional ligand-binding pockets, determining binding energies from physics based interactions and sampling the configurational space of protein–ligand complexes. The effectiveness of empirical scores and knowledge-based functions to evaluate the interaction between ligands and rigid proteins with a continuum description of solvent can be found. Finally, the last focuses on the field of binary and multi-component protein interactions. In many cases, the design of ligands must be envisioned in the context of a network of interacting molecules that have well-defined three-dimensional structures in isolation or become folded upon either binding to one partner or polymerization (Tuffery et al 2012).

3.1.1 Protein and Ligand Plasticity

It is well established that proteins, while adopting well-defined structures in aqueous solution, are in constant motion and display conformational heterogeneity. Experimental and theoretical studies also show that proteins can fluctuate between open and closed forms in the absence of ligand and protein domains are dynamic with movements including the hinge bending and the shear motions. Looking at molecular recognition, involving the non-covalent association of ligands (either low molecular weight molecules or macromolecules) to large macromolecules with high affinity and specificity, two mechanisms have been considered for a long time (Tuffery et al 2012).
3.1.2 Protein–ligand Interaction

The goal of protein docking with ligand is to predict the position and orientation of a ligand when it is bound to a protein receptor or enzyme active site. Recent pharmaceutical research employs docking studies for a variety of purposes, most notably in the virtual screening of drug candidates. Several protein–ligand docking software are available, such as AutoDock or EADock. There are also web services (Molecular Docking Server, SwissDock) that calculate the site, geometry and binding energy of ligands interacting with proteins. Computational capacity has increased dramatically over the last decade making possible the use of more sophisticated and computationally intensive methods in computer-assisted drug discovery and design. However, dealing with receptor flexibility in docking is still a problematic issue. The main reason behind this difficulty is the large number of degrees of freedom that have to be considered in this kind of calculations. However, neglecting it leads to poor docking results in terms of binding pose prediction in real-world settings. An important and very active issue in protein–ligand recognition, where the ligand is a small molecule such as drug, crosslinker, is to predict three-dimensional ligand-binding pockets. Numerous structure-based methods, in particular for in silico screening of small compounds, have been developed to detect pockets, clefts or cavities in proteins (Tuffery et al 2012).

3.1.3 Protein–protein and Multi-component Protein Interactions

Macromolecular docking is the computational modelling of the quaternary structure of complexes formed by two or more interacting biological macromolecules. Protein–protein complexes are the most commonly attempted targets of such modelling, followed by protein–nucleic acid, protein–carbohydrate and protein-lipid complexes. The field of protein–protein interactions has rapidly progressed in the past 12 years. Protein–
protein docking aims to predict the three-dimensional structure of a complex from the knowledge of the structure of the individual proteins in aqueous solution. Many docking methods are now able to predict the three-dimensional structure of binary assemblies if the protein partners do not display important conformational changes between their bound and unbound forms (Tuffery et al 2012).

The molecular docking is an excellent approach to the study of the effects of crosslinking agents, individually and collectively, on triple helical structure of collagen. Covalent and non covalent inter and intra molecular crosslinking, electrostatic interactions (salt bridges), hydrophobic interactions and H-bonds are among the contributing factors in the stabilization of collagen playing an important role in the preparation of collagen matrix since it can affect the final properties of the matrix (Ramshaw et al 2009; Place, 2009). The extent of interaction of the collagen in the interlayer spacing of matrix is affected by the interaction between the functional groups of collagen with the crosslinker (Berisio et al 2009; Thiagarajan et al 2010; Mogilner et al 2010). The chemical crosslinking involves the formation of either intra or intermolecular ionic or covalent bonds between the AAs residues of collagen. Condensation of AAs residues by both DHT and photochemical crosslinking with UV-light have resulted in successful collagen crosslinking. These physical crosslinking methods are convenient because they do not require any chemical additives, however, they commonly yield an insufficient degree of crosslinking. Other chemical crosslinking methods using inorganic polyvalent cations, such as trivalent chromium and aluminium etc., and organic aldehydes, carbodiimides and polyepoxy compounds, etc., have been used successfully to produce ionic crosslinks; however, these methods significantly interfere with the mobility and the strength of the material remained insufficient for many biological applications (Ashwin et al 2011; Ye et al 2010). Among them, the EDC/NHS cross-linking system is frequently used in the treatment of collagen scaffold, which has shown better biocompatibility.
than aldehydes (Everaerts et al 2008). The cross-linking takes place by reaction between carboxyl groups of glutamic/aspartic acid residues and amine groups of lysine of collagen to form amide bonds.

As an alternative to chemically altering collagen to improve the $T_s$, $T_d$ and resistance to biodegradation, while preserving the biocompatibility, EDC/NHS method used to cross-link the collagen matrix in the presence of D-AAs serving as cross-linking bridges has been studied in chapter 2. Recent computational study suggested that replacing AAs residues from L to D-conformation would stabilize the collagen triple helix. These D-AAs proteins and peptides have longer-acting anti-microbial, antioxidant, anti-mutagenic and anti-carcinogenic activity than their L-enantiomeric analogues. The D-AAs containing peptide and protein fragments are less digestible than the L-enantiomers. The D-AAs are most effective chemopreventive agents against the initiation, promotion and progression stages of multistage carcinogenesis. The D-AAs peptides contribute to innate host defense against a number of microbial pathogens. The D-AAs, not usually found in natural proteins, have conformational attributes that are useful for the imposition of conformational stability and as structural probes. D-AAs endow resistance to proteolytic degradation, an attribute important in chemotherapeutic application. Replacement glycine residues with D-AAs in several globular proteins and found that the incorporation of D-AAs can greatly enhance conformational stability (Anil et al 2004, 2006). In addition, D-AAs can be better than L-AAs as glycine surrogates, retaining protein function without perturbing protein structure (Valiyaveetil et al 2004; Xie et al 2005; Bang et al 2006). Indeed, many compounds with antimicrobial and antitumor activity contain D-AAs. D-AAs in animal peptides act as hormones and antimicrobial agents. D-AAs act as intermediates of semisynthetic antibiotics, bioactive peptides and other physiologically active compounds.
The use of natural and unnatural AAs for designing of nanocomposite and nanoclay biomaterials has been investigated (Katti et al 2010, MacPhee et al 2004; Chow et al 2008; Ambre et al 2011). The development of chiral compounds, especially in the pharmaceutical field and food, is placing increasing demands to achieve chiral purity control of therapeutic drugs (Annedi et al 2006; Friedman 2010). Because the enantiomers of the same therapeutic drug often show different pharmacological or bioactive effects. Poly-α-AAs have been investigated for possible use in a wide variety of biomedical applications (Sun et al 2011).

Nowadays, the molecular docking is an outstanding excellent approach to the study of the effects of crosslinking agent on stabilization of biomaterials for industrial and biomedical applications. Two approaches are particularly popular within the molecular docking community. One approach uses a matching that describes the protein and the ligand as complementary surfaces. The second approach simulates the actual docking process in which the ligand-protein pairwise interaction energies are calculated. Both approaches have significant advantages as well as some limitations. With regard to the bioinformatics on crosslinking chemistry of collagen, only very few reports are available. Applications of computer modelling for the design of orthopaedic, dental and cardiovascular biomaterials have been studied (Sloten et al 1998; Vaidyanathan et al 2001). The experimental and modeling study of collagen scaffolds with the effects of crosslinking and fiber alignment have been studied (Xu et al 2011). The aromatic interactions promoting self-association of collagen triple-helical peptides to higher-order structures have been studied (Kar et al 2006 and 2009). Therefore, an understanding of the role of D-AAs on collagen may help to design collagen matrix. The field has developed rapidly, permitting the construction of synthetic, protein based molecules. This has allowed protein chemists to explore the physical and biological effects of varying AAs stereochemistry. In this study, we use the Protein Data Base (PDB) and Pubchem Compound
Database as a source of structural information for specific D-AAs sidechain interactions with collagen. These works will provide a basis for stabilization of collagen for biomaterial applications.

3.1.4 Objective of the Work

The objectives of the present study is to evaluate the molecular level understanding of interactions of selected D-AAs with energy minimized triple helical structure of CLP for stabilization by molecular docking method for biomaterial applications. Three D-AAs such as D-alanine (D-Ala), D-glutamic acid (D-Glu) and D-lysine (D-Lys) were selected from Pubchem Compound Database. Energy minimized triple helical structures of the collagen like peptide (CLP) were collected from Protein Data Bank (PDB). The D-AAs were docked to CLP interlayer region by Autodock procedures and the receptor sites were evaluated. The binding energy of the D-AAs with CLP were evaluated and compared. H-bonds between D-AAs and CLP were detected in the complexes of several of the conformations of all the D-AAs and best conformation were analysed. These molecular docking studies have provided clear evidence for stabilization of collagen by unnatural stereochemistry of D-AAs for biomaterial applications. Molecular modelling studies can be used as a tool to study the interaction of CLP with D-AAs. In the present work molecular modelling tools are used.

- For computation of the docking of CLP with D-AAs.
- To estimate the binding energy of D-AAs with CLP.
- To identify the major binding sites in the CLP for interaction with D-AAs.
- To study the molecular docking of D-AAs docked CLP with ChC.
3.2 EXPERIMENTAL PROCEDURES

The computational calculations comprise the determination of the three-dimensional structure of the CLP and D-AAs by energy minimization methods, and the simulation of the interaction between CLP-D-AAs and ChC active site by a docking procedure. Therefore, the area of the surface of the collagenase capable to interact with CLP-D-AAs had to be obtained by simulation.

3.2.1 CLP Selection

The crystal structure of triple helical CLP was obtained from Protein Data Bank (PDB: 3A1H) for molecular docking studies. The AAs sequence of 3A1H is shown in Figure 3.1.

>3A1H:A|PDBID|CHAIN|SEQUENCE
>3A1H:B|PDBID|CHAIN|SEQUENCE
>3A1H:C|PDBID|CHAIN|SEQUENCE
>3A1H:D|PDBID|CHAIN|SEQUENCE
>3A1H:E|PDBID|CHAIN|SEQUENCE
>3A1H:F|PDBID|CHAIN|SEQUENCE

Figure 3.1 The AAs sequence of CLP

The OH, COOH and NH$_2$ in the AAs can be considered as the potential interacting sites for the formation of H-bonds with D-AAs. AAs having similar functional groups have not been repeated in the model peptide sequence. The single-crystal structures of three collagen-like host-guest peptides are (Pro-Pro-Gly)(4)-Hyp-Yaa-Gly-(Pro-Pro-Gly)(4) [Yaa=Thr, Val, Ser; Hyp(4R)]. Following AAs are sequences of CLP.
Table 3.1 List of AAs and their properties

<table>
<thead>
<tr>
<th>Amino Acid</th>
<th>3-Letter</th>
<th>1-Letter</th>
<th>Side-chain polarity</th>
<th>Side-chain charge (pH 7.4)</th>
<th>Hydropathy index</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alanine</td>
<td>Ala</td>
<td>A</td>
<td>nonpolar</td>
<td>Neutral</td>
<td>1.8</td>
</tr>
<tr>
<td>Arginine</td>
<td>Arg</td>
<td>R</td>
<td>polar</td>
<td>Positive</td>
<td>−4.5</td>
</tr>
<tr>
<td>Asparagine</td>
<td>Asn</td>
<td>N</td>
<td>polar</td>
<td>Neutral</td>
<td>−3.5</td>
</tr>
<tr>
<td>Aspartic acid</td>
<td>Asp</td>
<td>D</td>
<td>polar</td>
<td>Negative</td>
<td>−3.5</td>
</tr>
<tr>
<td>Cysteine</td>
<td>Cys</td>
<td>C</td>
<td>polar</td>
<td>Neutral</td>
<td>2.5</td>
</tr>
<tr>
<td>Glutamic acid</td>
<td>Glu</td>
<td>E</td>
<td>polar</td>
<td>Negative</td>
<td>−3.5</td>
</tr>
<tr>
<td>Glutamine</td>
<td>Gln</td>
<td>Q</td>
<td>polar</td>
<td>Neutral</td>
<td>−3.5</td>
</tr>
<tr>
<td>Glycine</td>
<td>Gly</td>
<td>G</td>
<td>nonpolar</td>
<td>Neutral</td>
<td>−0.4</td>
</tr>
<tr>
<td>Histidine</td>
<td>His</td>
<td>H</td>
<td>polar</td>
<td>positive (10%) neutral (90%)</td>
<td>−3.2</td>
</tr>
<tr>
<td>Isoleucine</td>
<td>Ile</td>
<td>I</td>
<td>nonpolar</td>
<td>Neutral</td>
<td>4.5</td>
</tr>
<tr>
<td>Leucine</td>
<td>Leu</td>
<td>L</td>
<td>nonpolar</td>
<td>neutral</td>
<td>3.8</td>
</tr>
<tr>
<td>Lysine</td>
<td>Lys</td>
<td>K</td>
<td>polar</td>
<td>positive</td>
<td>−3.9</td>
</tr>
<tr>
<td>Methionine</td>
<td>Met</td>
<td>M</td>
<td>nonpolar</td>
<td>neutral</td>
<td>1.9</td>
</tr>
<tr>
<td>Phenylalanine</td>
<td>Phe</td>
<td>F</td>
<td>nonpolar</td>
<td>neutral</td>
<td>2.8</td>
</tr>
<tr>
<td>Proline</td>
<td>Pro</td>
<td>P</td>
<td>nonpolar</td>
<td>neutral</td>
<td>−1.6</td>
</tr>
<tr>
<td>Serine</td>
<td>Ser</td>
<td>S</td>
<td>polar</td>
<td>neutral</td>
<td>−0.8</td>
</tr>
<tr>
<td>Threonine</td>
<td>Thr</td>
<td>T</td>
<td>polar</td>
<td>neutral</td>
<td>−0.7</td>
</tr>
<tr>
<td>Tryptophan</td>
<td>Trp</td>
<td>W</td>
<td>nonpolar</td>
<td>neutral</td>
<td>−0.9</td>
</tr>
<tr>
<td>Tyrosine</td>
<td>Tyr</td>
<td>Y</td>
<td>polar</td>
<td>neutral</td>
<td>−1.3</td>
</tr>
<tr>
<td>Valine</td>
<td>Val</td>
<td>V</td>
<td>nonpolar</td>
<td>neutral</td>
<td>4.2</td>
</tr>
</tbody>
</table>

3.2.2 Ligand Selection

The structures of three different selected D-AAs such as D-Alanine (CID:71080), D-Glutamic acid (CID:23327) and D-Lysine (CID:57449)) were retrieved from Pubchem Compound Database for this study. The details are shown in the Figure 3.2. The AAs properties are shown in the Table 3.1.
**Amino acids**

<table>
<thead>
<tr>
<th>Amino acids</th>
<th>L-Configuration</th>
<th>D-Configuration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alanine</td>
<td><img src="image" alt="Alanine L-Configuration" /></td>
<td><img src="image" alt="Alanine D-Configuration" /></td>
</tr>
<tr>
<td>Glutamic acid</td>
<td><img src="image" alt="Glutamic acid L-Configuration" /></td>
<td><img src="image" alt="Glutamic acid D-Configuration" /></td>
</tr>
<tr>
<td>Lysine</td>
<td><img src="image" alt="Lysine L-Configuration" /></td>
<td><img src="image" alt="Lysine D-Configuration" /></td>
</tr>
</tbody>
</table>

**Figure 3.2** Schematic representations of molecular structures of D-AAs

### 3.2.3 Collagenase Selection

The crystal structure of a collagen-binding domain (CBD) with an N-terminal domain linker from *Clostridium histolyticum* class I collagenase (ChC) was determined at 1.00 Å resolution in the absence of calcium (1NQJ) and at 1.65 Å resolution in the presence of calcium (1NQD) were retrieved from PDB for this study. The AAs sequence of 1NQD chain A and B is shown in the Figure 3.3.
The mature enzyme is composed of four domains: a metalloprotease domain, a spacing domain and two CBDs. A 12-residue-long linker is found at the N-terminus of each CBD. In the absence of calcium, the CBD reveals a beta-sheet sandwich fold with the linker adopting an alpha-helix. The addition of calcium unwinds the linker and anchors it to the distal side of the sandwich as a new beta-strand.

### 3.2.4 Molecular Docking of the Interaction of D-AAs with CLP

The molecular docking studies comprise the determination of the three dimensional structure of CLP interaction with D-AAs (Autodock 4.2).

### 3.2.5 Molecular Docking of D-AAs Docked CLP with Collagenase

The molecular docking studies comprising the determination of the three-dimensional structure of D-AAs docked CLP with ChC were carried out on a PatchDock (Duhovery et al 2002; Schneidman-Duhovery et al 2003 and 2005; Mashiach et al 2010). Orientations that placed CLP within 4.5Å of T957, Y970, L992 and Y996 were presented in the active site of the collagenase. In addition, a 6Å filtered was used for the weaker interacting H959 and R929.
3.3 RESULTS

Incorporating D-AAs in molecular design has the potential to improve existing protein stability and create new topologies inaccessible to homo chiral molecules (Annavarapu et al 2009; Woolfson et al 2010). D-AAs rarely occur naturally, making it difficult to infer general rules for how they would be tolerated in proteins through an analysis of existing protein structures. However, protein elements containing short left-handed turns and helices turn out to contain useful information. Molecular mechanisms used in proteins to stabilize left-handed elements by L-AAs are structurally enantiomeric to potential synthetic strategies for stabilizing right-handed elements with D-AAs. From a functional point of view, the D-residues are much more resistant to enzyme-catalyzed breakdown than L-residues, a property of considerable pharmaceutical importance (Friedman 2010; Martinez-Rodriguez et al 2010). It is of profound interest to establish how the D-AAs exhibit better interaction with collagen for stabilization in order to develop the collagen matrix for tissue engineering scaffold. Recent computational study suggested that replacement AAs residues L to D-conformation would stabilize the collagen triple helix (Shah et al 1999). To probe the stabilization of collagen matrix by D-AAs, the interaction of model of CLP with D-AAs has been studied.

A fundamental understanding of the interaction of D-AAs with CLP is important to unravel the nature of interactions that are required for the stabilization of collagen matrix for biomedical applications. The molecular docking studies are used to validate the stabilization of collagen.

3.3.1 CLP-D-AAs Interaction

The preparation and characterization of D-AAs modified collagen matrix by presence of EDC/NHS for design of scaffold for tissue engineering have been earlier studied (Ref Chapter 2). The treatment of collagen with D-AAs has been found to increase the thermal stability and provide stability
against collagenase activity. In this study, molecular docking of CLP interaction with D-AAs such as D-Ala, D-Glu and D-Lys were carried out. The energy minimized triple helical structure of CLP (PDB: 3A1H) is shown in Figure 3.4.

![The energy minimized triple helical structure of CLP](image)

**Figure 3.4 The energy minimized triple helical structure of CLP**

The structures of three different selected D-AAs such as D-Ala (CID: 71080), D-Glu (CID: 23327) and D-Lys CID: 57449) are shown on Figure 3.2. As would be expected, the potential binding sites was in the interlayer region than in the region between two CLP. The best geometries obtained through the docking represent the lowest binding energy between CLP and D-AAs, resulting in stronger H-bonds. Inter and intrahelical H-bonds and hydrophobic interactions with the involvement of functional groups of collagen might have been involved in the stabilization.

### 3.3.1.1 Interaction of D-Ala

The interaction of CLP with D-Ala binding at multiple sites and selected the best predicted H-bonds (dashed yellow lines) are shown on Figure 3.5 and Table 3.2. The binding energy of CLP interaction with D-Ala has been computed to be -3.56 Kcal/mol. Three potential sites viz., D-Ala (N-H...O)Thr14, D-Ala(O..H-N) Thr14 and D-Ala(O..H-N) Thr14 have been identified, which maintained hydrogen bonds length 2.78 Å, 3.18 and 3.06 Å respectively.
Figure 3.5  Binding site, H-bonding and hydrophobic interaction of CLP with D-Ala
Table 3.2 The binding energy and H-bonding interaction of CLP with D-Ala

<table>
<thead>
<tr>
<th>Site interaction</th>
<th>D-Ala complex with CLP</th>
<th>Types of interactions</th>
<th>Bond distance (Å)</th>
<th>Binding energy (Kcal/mol)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thr</td>
<td>D-Ala(N-H...O)Thr14</td>
<td>H-bond</td>
<td>2.78</td>
<td>-3.56</td>
</tr>
<tr>
<td>Thr</td>
<td>D-Ala(O..H-N)Thr14</td>
<td>H-bond</td>
<td>3.18</td>
<td>3.06</td>
</tr>
<tr>
<td></td>
<td>D-Ala(O..H-N)Thr14</td>
<td></td>
<td>3.06</td>
<td></td>
</tr>
</tbody>
</table>

Alanine is classified as a nonpolar AA with the chemical formula CH₃CH- (NH₂)COOH. D-Ala occurs in bacterial cell walls and in some peptide antibiotics. D-Ala is one of the essential components of the peptidoglycan in the bacterial cell wall. D-Ala is an important raw material for the production of medicines and agrochemicals and several fermentative or enzymatic methods have been developed to produce it. Tsai et al 2005 suggested that D-Ala residue stabilizes the collagen triple helix.

3.3.1.2 Interaction of D-Glu

The interaction of CLP with D-Glu binding at multiple sites and the best conformation predicted H-bonds (dashed yellow lines) formed are shown on Figure 3.6 and Table 3.3. The binding energy of CLP interaction with D-Glu has been computed to be -3.69Kcal/mol. Three potential sites viz., D-Glu(O-H...N)Pro11, D-Glu(O-H..O)Pro11, D-Glu(O...H-N)Thr14 and D-Glu(N-H...O)Thr14 have been identified, which maintained hydrogen bonds length 3.27 Å, 3.22 Å, 2.84 Å, 2.91 Å respectively.
Figure 3.6  Binding site, H-bonding and hydrophobic interaction of CLP with D-Glu
Table 3.3 The binding energy and H-bonding interaction of CLP with D-Glu

<table>
<thead>
<tr>
<th>Site interaction</th>
<th>D-Glu complex with CLP</th>
<th>Types of interactions</th>
<th>Bond distance (Å)</th>
<th>Binding energy (Kcal/mol)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pro</td>
<td>D-Glu(O−H...N)Pro11</td>
<td>H-bond</td>
<td>3.27</td>
<td>-3.69</td>
</tr>
<tr>
<td></td>
<td>D-Glu(O−H...O)Pro11</td>
<td></td>
<td>3.22</td>
<td></td>
</tr>
<tr>
<td>Thr</td>
<td>D-Glu(O...H-N)Thr14</td>
<td>H-bond</td>
<td>2.84</td>
<td></td>
</tr>
<tr>
<td>Thr</td>
<td>D-Glu(N−H...O)Thr14</td>
<td>H-bond</td>
<td>2.91</td>
<td></td>
</tr>
</tbody>
</table>

The carboxylate anions and salts of glutamic acid are known as glutamates. The side chain carboxylic acid functional group has pKa of 4.1 and exists in its negatively charged deprotonated carboxylate form at pH greater than 4.1. Therefore it is also negatively charged at physiological pH ranging from 7.35 to 7.45. D-Glu is an essential building block for peptidoglycan in bacterial cell walls.

3.3.1.3 Interaction of D-Lys

The interaction of CLP with D-Lys binding at multiple sites and the lowest energy predicted H-bonds (dashed yellow lines) formed are shown on Figure 3.7 and Table 3.4. The binding energy of CLP interaction with D-Lys has been computed to be -4.23 Kcal/mol. The potential sites viz., D-Lys(N-H...O)Pro11, D-Lys(O-H...O)Gly12, D-Lys(N-H...O)Gly12 and D-Lys(N-H...O)Thr14 have been identified, which maintained hydrogen bonds length 2.69 Å, 2.67 Å, 3.10 Å and 2.72 Å respectively.
Figure 3.7 Binding site, H-bonding and hydrophobic interaction of CLP with D-Lys
Table 3.4 The binding energy, H-bonding interaction of CLP with D-Lys

<table>
<thead>
<tr>
<th>Site interaction</th>
<th>D-Lys complex with CLP</th>
<th>Types of interactions</th>
<th>Bond distance (Å)</th>
<th>Binding energy (Kcal/mol)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pro</td>
<td>D-Lys(N-H...O)Pro11</td>
<td>H-bond</td>
<td>2.69</td>
<td>-4.23</td>
</tr>
<tr>
<td>Gly</td>
<td>D-Lys(O-H...O)Gly12</td>
<td>H-bond</td>
<td>2.67</td>
<td></td>
</tr>
<tr>
<td>Gly</td>
<td>D-Lys(N-H...O)Gly12</td>
<td>H-bond</td>
<td>3.10</td>
<td></td>
</tr>
<tr>
<td>Thr</td>
<td>D-Lys(N-H...O)Thr14</td>
<td>H-bond</td>
<td>2.72</td>
<td></td>
</tr>
</tbody>
</table>

The ε-amino group often participates in H-bonding and as a general base in catalysis. The ε-amino group, which is attached to the NH$_3^+$ group, is the fifth carbon down from the α-carbon, which is attached to the COOH group. These studies clearly show the possible anchoring points of D-Lys than D-Ala and D-Glu with the CLP. The stability of triple helical conformation of collagen is imparted by the intermolecular H-bonds. This may be due to a net increase in the number of inter and intra molecular crosslinks and interactions between the D-AAs and CLP. The hydrophobic interactions determined predominantly by glycine residues and H-bonds formed between Hyp residues play an important role in the stabilization of collagen. Presumably, the combination of high hydrophobicity and capacity to form peptide bonds permit these molecules incorporate in certain areas of collagen fibrils and promote stabilization of their structure. This could be possibly due to the decrease in the availability of the active sites in the matrix. The binding energies of CLP interaction with D-AAs such as D-Ala, D-Glu and D-Lys are -3.56, -3.69 and -4.23 Kcal/mole respectively (shown in the Figure 3.8). The D-Lys required lowest binding energy compared to D-Ala and D-Glu. These are rare due to the unfavorable steric interactions. Hence, the value to protein engineering and design is to realize that the structural enantiomer of such interactions would involve right-handed structures stabilized with D-AAs.
Figure 3.8 Binding energy of D-AAs with CLP

Left-handed helices are understandably rare in proteins due to the inherent conformational preferences dictated by backbone stereochemistry. Less than one percent of residues are found in contiguous left-handed turns or helices of length three or greater. The influence of D-AAs substitutions on the stability of an amphipathic, monomeric helix were studied (Krause et al 2000).

3.3.2 CLP-D-AAs-ChC Interaction

3.3.2.1 CLP-ChC interaction

CLP–ChC docking was used to predict possible orientations of ChC across CLP. The C-terminal collagen-binding domain (CBD) of these enzymes is the minimal segment required to bind to collagen fibril (Figure 3.9). Collagen binding efficiency of CBD is more pronounced in the presence of Ca$^{2+}$. CBD also binds to collagenous peptides with triple helical conformation but not to collagenous peptides that lack triple helix or to gelatin suggesting that the CBD-collagen interaction is conformation-specific. Calcium ions enhance the binding at physiological concentration.
Figure 3.9 A) The 3D structures of ChC viewed by cartoon representation and B) Five AAs residue namely Thr957, Tyr970, Leu992, Tyr994 and Tyr996 in line format are labelled in the active site of ChC.

The most complete and closest model CLP which was docked with ChC is shown in the Figure 3.10. The solutions were filtered to select binding models consistent with the five key residues known to interact with CLP.
Figure 3.10 Proposed structure of CLP-ChC interaction

Of the top ten solutions, the highest scoring orientation appeared eight times only with slight variations. The final orientation among the top ten allowed CLP to interact with His959 is skewed. Since the single-crystal structures of three CLP host-guest peptides (Pro-Pro-Gly)(4)-Hyp-Yaa-Gly-(Pro-Pro-Gly)(4) [Yaa=Thr, Val, Ser; Hyp(4R)] bind tightly to ChC, the OH group of Tyr994 is likely to be H-bonding to the main-chain atoms rather than to the OH group of Hyp as expected for a CLP-ChC complex. Based on these results, the experimental results can be discussed. Thr957, Tyr970, Leu992, Tyr994 and Tyr996 AAs residues are very important for bonding with CLP
(Tyr994 is probably the most important: its substitution reduces the activity more than 5 times). Tyr990, Tyr931, Tyr932, Asp966 and Asp974 are less important, because their loss or blockade have a small inhibition effect. CBD binds unidirectionally to the C-terminal region of collagenous peptides. Though CBD targets the least ordered region of the collagenous peptides, CBD could also target partly unwound regions even in the middle of a tropocollagen. The catalytic domain is situated at the N terminus of CBD. Bacterial collagenolysis may involve optimal orientation of the catalytic domain with respect to tropocollagen. Furthermore, the collagenolysis could initiate from partly unwound regions in tropocollagen.

3.3.2.2 CLP-D-AAs-ChC interaction

Molecular docking of CLP-D-AAs-ChC interaction was carried out by the action of ChC on D-AAs docked CLP. Figure 3.11-13 shows the complete and closest structure to the ChC structure, where the contact surfaces are coloured according to their distance. The best geometries obtained through the docking represent the high interaction energy between CLP and ChC, resulting in preventing the free access of ChC to reactive sites on the CLP chains. The degradation of D-AAs protein and protein based materials are not well equipped by the proteolytic machinery, however, it is this result above all else that has spurned research into stereochemical inversion and directional manipulation of peptide and polypeptide chains. The expectation has been that systematic inversion of the stereochemistry at the peptide backbone $\alpha$-carbon atoms, if accompanied by chain reversal, should yield proteolytically stable retro-inverso peptide isomers. These D-AAs derivatives form the basis of important families of enzyme inhibitors. Incorporating variable D-AAs in collagen scaffold has the potential to improve existing collagen stability and create new topologies inaccessible to homo chiral molecules.
3.3.2.2.1 CLP-D-Ala-ChC

The docking was carried out by the action of CLP-D-Ala with ChC. Figure 3.11 shows the complete and closest structure to the ChC with CLP-D-Ala, where the contact surfaces are colored according to their distance.

![Figure 3.11 The CLP-D-Ala binding to ChC active site](image)

Based on the reported molecular docking evidence, that D-Ala efficiently inhibit the ChC, these molecular docking representations clearly show there are no orientational changes of ChC on CLP-D-Ala site but D-Ala trigger the binding of ChC with CLP.
3.3.2.2 CLP-D-Glu-ChC

The docking was carried out by the action of CLP-D-Ala with ChC. Figure 3.12 shows the completed and closest structure to the ChC with CLP-D-Glu, where the contact surfaces are colored according to their distance.

Figure 3.12 The CLP-D-Glu binding to ChC active site

The dashed yellow line represents the possible H-bonds. These representations clearly show the possible anchoring points of Coll-D-Glu moieties with the ChC active site. The D-Glu efficiently changes binding site of ChC. These molecular docking representations clearly show there are
orientational changes of ChC on CLP-D-Glu site and trigger the binding of ChC with CLP.

### 3.3.2.2.3 CLP-D-Lys-ChC

The CLP-D-Lys-ChC interaction was carried out by the action of ChC with CLP-D-Lys. Figure 3.13 shows the closest structure to the ChC structure, where the contact surfaces of ChC with CLP are colored according to their distance. The dashed yellow line represents the possible H-bonds. These representations clearly show the possible anchoring points of CLP-D-Lys moieties with the ChC active site. Most importantly, CLP-D-Lys greatly impairs ChC activity. This representation clearly show there are orientational changes of ChC on CLP-D-Lys and trigger the binding of ChC with CLP. The stability of D-AAs crosslinked matrix against ChC would have been brought about by protecting the active sites in collagen through interaction with D-AAs not recognized by ChC and another practical application is the development of thermostable collagen that incorporates D-AAs. It has to be noted that a higher biostability was observed for D-Lys than D-Ala and D-Glu. The D-Lys can interact with collagen through H-bonding and hydrophobic interactions. The di-amino groups of the D-Lys molecule, likely to crosslink with collagen may establish multiple peptide and H-bonds with neighboring collagen molecules, resulting in improved stability of matrix and prevent the free access of ChC to reactive sites on the collagen chains. The stability of Coll-D-Lys against ChC would have been brought about by protecting the active sites in collagen.
The significant differences in the enzymatic stability offered by D-Lys could be due to the effectiveness of the latter in exhibiting better interaction with CLP through H-bonded inter and intramolecular crosslinks. The experiments show that the inhibiting effect of D-AAs on ChC decreases from D-Lys, D-Ala and D-Glu.

3.4 DISCUSSION

The increasing interest of the pharmaceuticals and fine chemicals industries in the production of drugs and intermediates is an incentive to devise new chiral technologies and to improve enantioselective processes and
D-AAs utilized as building blocks, drug discovery, molecular scaffolds and composite matrix representing a nearly infinite array of diverse structural entities (Latacz et al 2006; Caligiuri et al 2012). The D-AAs are used to design and obtain new peptidomimetics (Oh et al 1999). They have also been used to create drugs or other therapeutic agents that will not be degraded as quickly as natural AAs by biodegradation. Structure-based designs of non-natural AAs inhibitors of amyloid fibril formation have been studied (Sievers et al 2011). Fabrication of self-assembling D-form peptide nanofiber scaffold for improving the peptide stability and controlled drug release have been studied (Liang et al 2009; Luca et al 2010; Luo et al 2011). Molecular modeling expanding the family of recombinant bacterial collagens of varying composition form triple-helices of similar stability have been studied (Xu et al 2010). This work has been motivated by natural examples of CLP that combine D-AAs.

The effect of the D-AAs used as a novel crosslinking bridge on enhancing the ability of collagen matrix to TS, T_{D} and resistance to the collagenase biodegradation and biocompatibility is well documented in the previous chapter 2. Molecular dynamics studies of the interprotein interactions in collagen fibrils have been investigated (Streeter et al 2011). Molecular modeling of nature designs tough collagen: explaining the nanostructure of collagen fibrils have been studied (Buehler et al 2006). Unstable molecules forming stable tissues have been studied (Persikov et al 2002). The effect of pH on structure and stability of collagen-like peptide and molecular dynamics simulations have been studied (Pu 2011). Increase in T_{D} of biomaterial gained by crosslinking agent is related to long range ordering in collagen induced by small molecules. The ability of D-AAs to stabilize collagen through crosslinking may well result from their ability to reduce accessibility of solvent to molecular surface of collagen scaffold. The present investigation also offers the possibility to understand recognition of
phenomena associated with collagen-collagen interaction, in general, based on interfacial volume and contact surface area of collagen scaffold.

In this study, collagen was docked with D-AAs, i.e. D-Asp/Glu, D-Ala and D-Lys, that contain 1/2, 1/1 and 2/1 of NH₂/COOH groups respectively. The side chain interaction of D-AAs into CLP imposes H-bonding interaction. To devise strategies for using collagen in the development of advanced biomaterials for biomedical engineering, it is necessary to confer resistance to proteolytic (collagenase) degradation with chemical or physical crosslinking strategies. The D-AAs can crosslink with collagen through peptide and multipoint H-bond interactions resulting in improved stability of collagen scaffold that could prevent the free access of MMPs to reactive sites on the collagen chains. Based on the reported molecular docking evidence, interaction of D-Lys efficient with collagen can be evinced. These molecular docking representations clearly show the possible anchoring points of D-AAs with the collagen. Docking calculations defined the best geometry for ligand-enzyme interaction. The best poses between each ligand and collagenase active site were selected. D-Lys was the only ligand for which a unfavorable interaction with the enzyme. Most importantly, multipoint H-bond and hydrophobic interactions of D-Lys greatly interact with collagen than D-Ala and D-Glu. D-Lys has the lowest binding energy (-4.2 Kcal/mol, indicating strongest interactions) when compared to D-Ala and D-Glu (-3.6 and -3.7 Kcal/mol respectively) and also D-Lys(N-H...O)Pro11, D-Lys(N-H...O)Pro11, D-Lys(O-H...O)Gly12, D-Lys(N-H...O)Gly12 and D-Lys(N-H...O)Thr14, which maintained hydrogen bonds length 2.69 Å, 2.67 Å, 3.10 Å and 2.72 Å respectively. D-Lys has the lowest binding energy with collagen and in experimental studies the matrix prepared with this amino acid exhibits improved fibrillar assembly and staggered alignment without the undesired structural stiffness and aggregations. The D-AAs along a peptide chain may be less digestible than the L-forms, D-D, D-L and L-D form of peptide bonds partly or fully inaccessible to proteolytic activities (Hamamoto et al 2002; De Vrese et al 2000). The antimicrobial toxin, gramicidin, is a well studied example of such a molecule, containing alternating L and D-AAs. D-AAs can function as C-terminal helix caps. It is a general observation that the cyclization of linear peptides that do not contain
turn structure-inducing residues is improbable or a very slow process. However, successful cyclization reactions have been possible and it was observed that the cyclization is favored by C-terminal D-AAs. A common strategy employed to improve yields in cyclization reactions involves the incorporation of AAs residues such as proline, glycine and D-AAs, which induce turns in the conformations of peptides. In a CLP, adding D-Ala to the C-terminus of a helix resulted in no significant change in stability whereas D-Lys increased stability presumably due to stabilization of the helix. The proteolytic machinery is not well equipped to deal with D-AAs polypeptides. This finding has spurred research into stereochemical and directional manipulation of peptide chains. The expectation has been that systematic inversion of the stereochemistry at the peptide backbone α-carbon atoms, if accompanied by chain reversal, should yield proteolytically stable retro-inverso peptide isomers. The study and design of the collagen largely dependents on D-AAs possessing non-natural structures or non-natural absolute configurations. Orientational changes in ChC on CLP-D-Lys matrix is observed which may decrease its accessibility to the degradation sites in CLP. These orientational changes in ChC may be stabilizing this CLP against the action of the former. The information derived from the present study could help in the design of heterochiral collagen based biomaterial.

3.5 CONCLUSIONS

To make the rational engineering and design of heterochiral proteins tractable, the role of AAs stereochemistry in stability and structure needs to be better understood. This study presents potential rules based on insights gained from the analysis of natural collagen. Using left-handed helices in the database of existing CLP structures as a case study, several candidates have been found for docking conformation that could be applied to the thermo stabilization of collagen by D-AAs. As synthetic methods for building proteins continue to improve, the ability to construct larger molecules with mixtures of natural and unnatural AAs becomes increasing
interesting tool. Natural proteins can provide important insight into how designed proteins can take advantage of the increased chemical diversity made possible by unnatural AAs.

In this present work, molecular level understanding of the interaction of collagen with three D-AAs viz., D-Ala, D-Glu and D-Lys, which represent neutral, acidic and basic AAs, enhance stability of collagen. All the D-AAs have resulted in H-bonds with both the side chain functional groups as well as with the amide backbone of the D-AAs. Moreover oxygen of the the carbonyl groups has resulted as a H-bond acceptor for the hydrogen of the OH groups of the D-AAs. While the D-Ala and D-Glu have highest binding energies, it has been shown here that D-Lys interaction with CLP imparts the lowest binding energy. The docking observations have proved the structures of CLP are stabilized. From this docking the interaction ability of the chosen D-AAs with CLP has been clearly demonstrated. Among the tested D-AAs, D-Lys is the only efficient interaction between CLP and could not be recognized and degraded by collagenase. This is a convincing evidence to prove that more stable collagen can be produced with D-AAs, which will have antimicrobial and anti-proteolytic activities. In conclusion, the molecular docking studies provide the data for stabilization of collagen matrix for tissue engineering applications.