CHARACTERIZATION OF METAL IONS INTERACTIONS IN IMMOBILIZED METAL ION AFFINITY CHROMATOGRAPHY BY USING COMPUTATIONAL TOOLS

İMMOBİLİZE METAL İYON AFİNİTE KROMATOGRAFİSİNDE KULLANILAN METAL İYONLARIN BİLGİSAYARLI ORTAMDA ETKİLEŞİM ÖZELLİKLERİNİN BELİRLENMESİ

Dima SALHA

PROF. DR. Adil DENIZLI
Supervisor

Submitted to Graduate School of Science and
Engineering of Hacettepe University
as a Partial Fulfillment to the Requirements
for the Award of the degree of Master of Science in Bioengineering
2017

This work named "**Characterization of Metal Ions Interactions in Immobilized Metal Ion Affinity Chromatography by Using Computational Tools**" by **DIMA SALHA** has been approved as a thesis for the Degree of **MASTER OF SCIENCE IN BIOENGINEERING** by the below mentioned Examining Committee Members.

Prof. Dr. Handan YAVUZ ALAGÖZ
Head .....

Prof. Dr. Adil DENİZLİ
Supervisor .....

Assoc. Prof. Dr. Nilay BERELİ
Member .....

Assist. Prof. Dr. Fatma YILMAZ
Member .....

Assist. Prof. Dr. Eda ÇELİK AKDUR
Member .....

This thesis has been approved as a thesis for the Degree of **MASTER OF SCIENCE IN BIOENGINEERING** by Board of Directors of the institute for Graduate School of Science and Engineering.

Prof. Dr. Menemşe GÜMÜŞDERELİOĞLU
Director of the Institute of
Graduate School of Science and Engineering

YAYINLAMA VE FİKRİ MÜLKİYET HAKLARI BEYANI

Enstitü tarafından onaylanan lisansüstü tezimin/raporumun tamamını veya herhangi bir kısmını, basılı (kağıt) ve elektronik formatta arşivleme ve aşağıda verilen koşullarla kullanıma açma iznini Hacettepe üniversitesine verdiğimi bildiririm. Bu izinle Üniversiteye verilen kullanım hakları dışındaki tüm fikri mülkiyet haklarım bende kalacak, tezimin tamamının ya da bir bölümünün gelecekteki çalışmalarda (makale, kitap, lisans ve patent vb.) kullanım hakları bana ait olacaktır.

Tezin kendi orijinal çalışmam olduğunu, başkalarının haklarını ihlal etmediğimi ve tezimin tek yetkili sahibi olduğumu beyan ve taahhüt ederim. Tezimde yer alan telif hakkı bulunan ve sahiplerinden yazılı izin alınarak kullanması zorunlu metinlerin yazılı izin alarak kullandığımı ve istenildiğinde suretlerini Üniversiteye teslim etmeyi taahhüt ederim.

- **Tezimin/Raporumun tamamı dünya çapında erişime açılabilir ve bir kısmı veya tamamının fotokopisi alınabilir.**
  (Bu seçenekle teziniz arama motorlarında indekslenebilecek, daha sonra tezinizin erişim statüsünün değiştirilmesini talep etseniz ve kütüphane bu talebinizi yerine getirse bile, tezinin arama motorlarının önbelleklerinde kalmaya devam edebilecektir.)

- **Tezimin/Raporumun** 17/5/2018 **tarihine kadar erişime açılmasını ve fotokopi alınmasını (İç Kapak, Özet, İçindekiler ve Kaynakça hariç) istemiyorum.**
  (Bu sürenin sonunda uzatma için başvuruda bulunmadığım taktirde, tezimin/raporumun tamamı her yerden erişime açılabilir, kaynak gösterilmek şartıyla bir kısmı ve ya tamamının fotokopisi alınabilir)

☐ **Tezimin/Raporumun ..... tarihine kadar erişime açılmasını istemiyorum, ancak kaynak gösterilmek şartıyla bir kısmı veya tamamının fotokopisinin alınmasını onaylıyorum.**

☐ **Serbest Seçenek/Yazarın Seçimi**

16/6/2017

(İmza)

Öğrencinin Adı Soyadı

Dima Salha

ETHICS

In this thesis study, prepared in accordance with the spelling rules of Institute of Graduate Studies in Science of Hacettepe University,
I declare that

- all the information and documents have been obtained in the base of the academic rules
- all audio-visual and written information and results have been presented according to the rules of scientific ethics.
- in case of using others works, related studies have been cited in accordance with the scientific standards
- all cited studies have been fully referenced
- I did not do any distortion in the data set
- and any part of this thesis has not been presented as another thesis study at this or any other university.

17/5/2017

Dima SALHA

ÖZET

İMMOBİLİZE METAL İYON AFİNİTE KROMATOGRAFİSİNDE KULLANILAN METAL İYONLARIN BİLGİSAYARLI ORTAMDA ETKİLEŞİM ÖZELLİKLERİNİN BELİRLENMESİ

Dima SALHA
YÜKSEK LİSANS TEZİ, Biyomühendislik Bölümü

Danışman: Prof. Dr. Adil DENİZLİ
Eş-Danışman: Doç. Dr. A. Müge ANDAÇ ÖZDİL
Mayıs 2017

Terapötik proteinlerin, peptitlerin, nükleik asitlerin, hormonların ve enzimlerin saflaştırılması veya ayrılması için etkin olarak kullanılan immobilize metal afinite kromatografisinde (IMAK), çeşitli fonksiyonel ligandlar ile metal iyonlarının etkileşimleri, bu yöntem ile ayrılması sağlanacak olan biyomakromoleküllerin bağlanma davranışlarını etkilemektedir. Birinci sıra geçiş metal iyonları (Zn^{2+}, Ni^{2+}, Cu^{2+}, ve Fe^{3+}), elektronca zengin (O, N, S içeren) iminodiasetik asit, nitrilotriasetik asit ve triskarboksimetil etilen-diamin gibi moleküllerle iyon dipol etkileşimleri üzerinden koordinasyon bağı kurarlar ve bunun sonucunda kararlı kompleksler oluştururlar.

Bu tez çalışmasında, immobilize metal afinite kromatografisinde yer alan etkileşimler dikkate alınarak, bilgisayarlı ortamda fonksiyonel ligand ve seçilen metal iyonları etkileşimlerinin modellenmesi ve etkin ayırma yöntemlerinin geliştirilmesi amaçlanmıştır. N-metakriloil-L-histidin metil ester (MAH) fonksiyonel ligand ve Zn^{2+} iyonları şelatlayıcı iyon olarak seçilmiştir. MAH monomeri Avogadro

programı kullanılarak bilgisayarlı ortamda çizilmiştir. İnsan insulin ve at kalbi sitokrom c, seçilen metal-ligand ile etkileşime girecek hedef moleküller olarak seçilmiştir. Otomatik moleküler yaklaşım yazılımı olan AutoDock 4.2, MAH monomerinin ve Zn^{2+} ile şelatlanmış MAH monomerinin, insan insulin ve at kalbi sitokrom C ile moleküler yaklaşım etkileşimlerinin incelenmesi için kullanılmıştır. Bu amaçla, sırasıyla, Zn^{2+} iyonlarına bir, iki ve üç molekül MAH bağlanmıştır. Yapılan moleküler yaklaşım çalışmaları sonucunda, bir molekül MAH monomerine bağlı Zn^{2+} iyonları ile en düşük bağlanma enerjisi elde edilmiştir. İnsan insülini için en düşük bağlanma enerjisi (- 4.14) kcal/mol ve at kalbi sitokrom C için en düşük bağlanma enerjisi (-4.92) kcal/mol olarak hesaplanmıştır.

Anahtar Kelimeler: IMAK, MAH, insulin, sitokrom C, bilgisayarlı yöntemler.

ABSTRACT

CHARACTERIZATION OF METAL IONS INTERACTIONS IN IMMOBILIZED METAL ION AFFINITY CHROMATOGRAPHY BY USING COMPUTATIONAL TOOLS

Dima SALHA

MASTER OF SCIENCE, Department of Bioengineering

Supervisor: PROF. DR. Adil DENIZLI

Co- Supervisor: Dr. A. Müge ANDAÇ ÖZDİL

May 2017

Immobilized metal ion affinity chromatography (IMAC) has become a widespread analytical and preparative separation method for therapeutic proteins, peptides nucleic acids, hormones, and enzymes. Many transition metals can perform stable complexes with electron rich compounds and may coordinate molecules containing O, N, and S by ion dipole interactions. Metal ion ligands are first-row transition metal ions (Zn^{2+}, Ni^{2+}, Cu^{2+}, and Fe^{3+}) incorporated by iminodiacetic acid, nitrilotriacetic acid, and tris (carboxymethyl) ethylene-diamine.

In this study we applied computational docking method on the interactions that occur in immobilized metal ion affinity chromatography. N-Methacryloyl-L-histidine Methyl Ester (MAH) is used as functional ligand. Then Zn^{2+} ions is selected to be chelated through imidazole groups on the MAH. N-Methacryloyl-L-histidine Methyl Ester (MAH) was drawn and created using Avogadro which has auto optimization tool. Human insulin molecule and horse heart cytochrome C are selected as targets to be interacted with our functional ligand. Automated docking software AutoDock

4.2 was used for docking of MAH molecules with human insulin and cytochrome C respectively. Thereafter, Zn^{2+} ion bound to one, two and three N-methacryloyl-L-histidine methyl ester (MAH) are studied and compared separately. By chelating Zn^{2+} ion to one MAH molecule and using human insulin as a target protein, the lowest binding energy (- 4.14) kcal/mol is found. Similarly, when horse heart cytochrome C is used as a target protein the lowest binding energy is (-4.92) kcal/mol by chelating Zn^{2+} ion to one MAH molecule.

Keywords: IMAC, MAH, insulin, cytochrome C, computational methods

ACKNOWLEDGEMENT

I would first like to thank my thesis supervisors, Prof Dr. Adil DENİZLİ, and Dr. A. Müge ANDAÇ ÖZDİL. I would like to express my deepest gratitude to them for their excellent guidance, caring and patience. I would never have been able to finish my dissertation without the guidance of them. They consistently allowed this thesis to be my own work, but steered me in the right direction whenever he thought I needed it.

I would also like to thank Uğur AYDIN for providing me with main concepts of the programs I used in my study.

I must express my very profound gratitude to my parents and to my husband for providing me with unfailing support and continuous encouragement throughout my years of study and through the process of researching and writing this thesis. This accomplishment would not have been possible without them. Thank you.

Index

Page No

ÖZ i
ABSTRACT..... iii
ACKNOWLEDGMENT v
INDEX vi
ABBREVIATIONS …….x

1. INTRODUCTION.....1
2. GENERAL INFORMATION 4
2.1. Affinity Chromatography.....4
2.1.1. History of affinity chromatography.....4
2.1.2. Fundamental principles of affinity chromatography 4
2.1.3. General formats of affinity chromatographic applications.....5
2.1.4. Biomolecules purified by affinity chromatography 7
2.1.5. Immobilization of affinity ligands.. 8
2.2. Immobilized Metal Ion Affinity Chromatography…. 10
2.2.1. General Concepts of IMACزز.....10
2.2.2. Components of IMAC…. 11
2.2.2.1. Metal Ions.. 11
2.2.2.2. Chelating Ligand.. 13
2.2.2.2.1. Synthesis of N-Methacryloyl-L-histidine Methyl Ester (MAH) 16
2.2.2.3. Media 17
2.3. IMAC Applications…….....17
2.4. Importance of Using Computational Simulation…. 18

2.5. Importance of Docking Studies for Immobilized Metal Ion Affinity Chromatography……20
2.6. AutoDock Program…….....21
2.7. Avogadro Program…….....23
2.8. Chimera program……....24
2.9. lPyMOL Program…….....24
2.10. Insulin……. 25
2.10.1. Structure of Insulin.. 26
2.11. Cytochrome C……....27
2.11.1. Cytochrome C Structure.. 28
3. EXPERIMENTAL……....30
3.1. Hardware and software 30
3.2. Preparation of macromolecules.....30
3.3. Preparation of ligand.....31
3.4. AutoDock.....32
3.4.1. Editing a PDB file.....32
3.4.2. Preparing a ligand file for AutoDock. 33
3.4.3. Docking.....33
3.4.4. Analyzing AutoDock Results-Reading Docking Logs 34
3.4.5 Analyzing AutoDock Results-Visualizing Docked Conformations 34
3.4.6. Analyzing AutoDock Results-Clustering Conformations.....35
3.5. Visualization.....35
4. RESULTS AND DISCUSSION.....36
4.1. Docking of One molecule of N-Methacryloyl-L-histidine Methyl Ester (MAH) at Human Insulin Molecule 36
4.1.1. Cluster Analysis of Conformations.....36
4.1.1.1. Clustering Histogram 36
4.1.1.2. RMSD Table 37
4.1.2. Information Entropy Analysis for This Clustering.....37
4.1.3. Statistical Mechanical Analysis.....38
4.1.4. Lowest Energy Docked Conformation 38

4.2. Docking of One molecule of N-Methacryloyl-L-histidine Methyl Ester (MAH) Bound to Zn^{2+} Ion at Human Insulin Molecule 40
4.2.1. Cluster Analysis of Conformations.....40
4.2.1.1. Clustering Histogram 40
4.2.1.2. RMSD Values 41
4.2.2. Information Entropy Analysis for This Clustering.....41
4.2.3. Statistical Mechanical Analysis.....42
4.2.4. Lowest Energy Docked Conformation 42
4.3. Docking of Two molecules of N-Methacryloyl-L-histidine Methyl Ester (MAH) Bound to Zn^{2+} Ion at Human Insulin Molecule 44
4.3.1. Cluster Analysis of Conformations.....44
4.3.1.1. Clustering Histogram 44
4.3.1.2. RMSD Values 45
4.3.2. Information Entropy Analysis for This Clustering.....46
4.3.3. Statistical Mechanical Analysis.....46
4.3.4. Lowest Energy Docked Conformation 46
4.4. Docking of Three molecules of N-Methacryloyl-L-histidine Methyl Ester (MAH) Bound to Zn2+ Ion at Human Insulin Molecule 48
4.4.1. Cluster Analysis of Conformations.....48
4.4.1.1. Clustering Histogram 48
4.4.1.2. RMSD Values 49
4.4.2. Information Entropy Analysis for This Clustering.....50
4.4.3. Statistical Mechanical Analysis.....50
4.4.4. Lowest Energy Docked Conformation 50
4.9. Comparison between the lowest binding energy of MAH-Insulin dockings. 52
4.5. Docking of One molecule of N-Methacryloyl-L-histidine Methyl Ester (MAH) at Horse Heart Cytochrome C.....53
4.5.1. Cluster Analysis of Conformations.....53
4.5.1.1. Clustering Histogram 54
4.5.1.2. RMSD Values 54
4.5.3. Statistical Mechanical Analysis.....55

4.5.4. Lowest Energy Docked Conformation 55
4.6. Docking of One molecule of N-Methacryloyl-L-histidine Methyl Ester (MAH) Bound to Zn^{2+} Ion at Horse Heart Cytochrome C 58
4.6.1. Cluster Analysis of Conformations 58
4.6.1.1. Clustering Histogram 58
4.6.1.2. RMSD Values 58
4.6.2. Information Entropy Analysis for This Clustering 59
4.6.3. Statistical Mechanical Analysis 59
4.6.4. Lowest Energy Docked Conformation 59
4.7. Docking of Two Molecules of N-Methacryloyl-L-histidine Methyl Ester (MAH) Bound to Zn^{2+} Ions at Horse Heart Cytochrome C 62
4.7.1. Cluster Analysis of Conformations 62
4.7.1.1. Clustering Histogram 63
4.7.1.2. RMSD Values 63
4.7.2. Information Entropy Analysis for This Clustering 64
4.7.3. Statistical Mechanical Analysis 64
4.7.4. Lowest Energy Docked Conformation 64
4.8. Docking of Three Molecules of N-Methacryloyl-L-histidine Methyl Ester (MAH) Bound to Zn^{2+} Ions at Horse Heart Cytochrome C 66
4.8.1. Cluster Analysis of Conformations 66
4.8.1.1. Clustering Histogram 66
4.8.1.2. RMSD Values 67
4.8.2. Information Entropy Analysis for This Clustering 68
4.8.3. Statistical Mechanical Analysis 68
4.8.4. Lowest Energy Docked Conformation 68
4.9. Comparison between the lowest binding energy of MAH- Horse Heart Cytochrome C dockings 70
5. CONCLUSION 72

Abbreviations

IMAC	Immobilized metal ion affinity chromatography
MAH	N-Methacryloyl-L-histidine Methyl Ester
ADT	AutoDock Tools
MD	Molecular Docking
LGA	Lamarckian genetic algorithm

1. INTRODUCTION

There is an ever-increasing necessity for studying structures and interactions of proteins which aim to achieve protein production in industrial and academic scales for a variety of applications. These include biopharmaceutical production, exploratory research, drug discovery initiatives, biotechnology products, target validation, and high-throughput screening. Consequently, there is an absolute requirement for the development of rapid, cost-effective methodologies which facilitate the purification of such products in the absence of contaminants, such as superfluous proteins and endotoxins [1]. Immobilized metal ion affinity chromatography (IMAC) is an efficient method for purification of biomolecules in both analytical and large-scale modes [2-5]. This kind of chromatography is based on the selective interaction of immobilized metal ions with proteins through their surface-exposed amino acid residues such as histidine, cysteine and tryptophan [6-9]. Many transition metals can form stable complexes with electron-rich compounds and may coordinate molecules containing O, N and S by ion dipole interactions [10]. As ligands for affinity separation, there are many benefits offered by immobilized metal ions. They are robust, small, inexpensive, physically and chemically stable, and can be easily coupled to matrices at high density resulting in high-capacity adsorbents [11]. The overall research strategy for the design of sterilizable, durable and highly selective affinity ligands contains five parts (Figure 1):

1) Identification of a target site and design of appropriate ligand which is complementary to the target based on X-ray crystallographic studies of complexes between the natural target protein and the biological ligand.
2) Solid phase synthesis and evaluation of related ligands.
3) Screening of the ligand library for binding the target protein by affinity chromatography.
4) Using *in silico* molecular modeling and docking of the ligand into the target protein to support selection and characterization of the lead ligand, (an affinity constant K_d in the range of 10^{-3} to 10^{-8} M between the protein and the immobilized ligand generally proves usable).

5) Optimization of the adsorbent and chromatographic parameters for the purification of the target protein.

By using this strategy, a nontoxic, chemically defined, fully synthetic and inexpensive affinity ligand can be obtained and used for the purification of high-value biopharmaceutical products [12].

The computational prediction of molecular complexes (molecular docking) is an important element for the understanding of functional relationships on molecular level [13]. In the field of molecular modeling, docking is a method which predicts the preferred orientation of one molecule to a second when bound to each other to form a stable complex. Knowledge of the preferred orientation in turn may be used to predict the strength of association or binding affinity between two molecules using for example scoring function [14]. Various software packages are commercially available to perform molecular modeling and docking. These software packages have a choice of energy minimization and automated docking programs that permit one to calculate, visualize and hypothesize about the energy and orientation of molecules in their three-dimensional state and when complexed with putative ligands [12].

The realistic prediction of protein–ligand complex structures (protein–ligand docking) is of major importance because only a small fraction of real and putative protein–ligand interactions in a cell can be determined experimentally. Modeling the interaction of two molecules is not an easy task. Many forces are involved in the intermolecular association, including hydrophobic, van der Waals, or stacking interactions between aromatic amino acids, hydrogen bonding, and electrostatic forces. Modeling the intermolecular interactions in a ligand-protein complex is difficult since there are many degrees of freedom as well as insufficient knowledge of the effect of solvent on the binding association. The process of docking a ligand to a binding site tries to mimic the natural course of interaction of the ligand and its receptor via the lowest energy pathway [15].

In this study, we applied computational docking method on the interactions that occur in immobilized metal ion affinity chromatography. N-Methacryloyl-L-histidine Methyl Ester (MAH) is used as functional ligand. Then Zn^{2+} ions is selected to be chelated through imidazole groups on the MAH. Human insulin molecule and horse

heart cytochrome c are selected as targets to be interacted with our functional ligand. This docking study is carried out by using automated docking software AutoDock 4.2., and the results are visualized by using Pymol and Chimera programs.

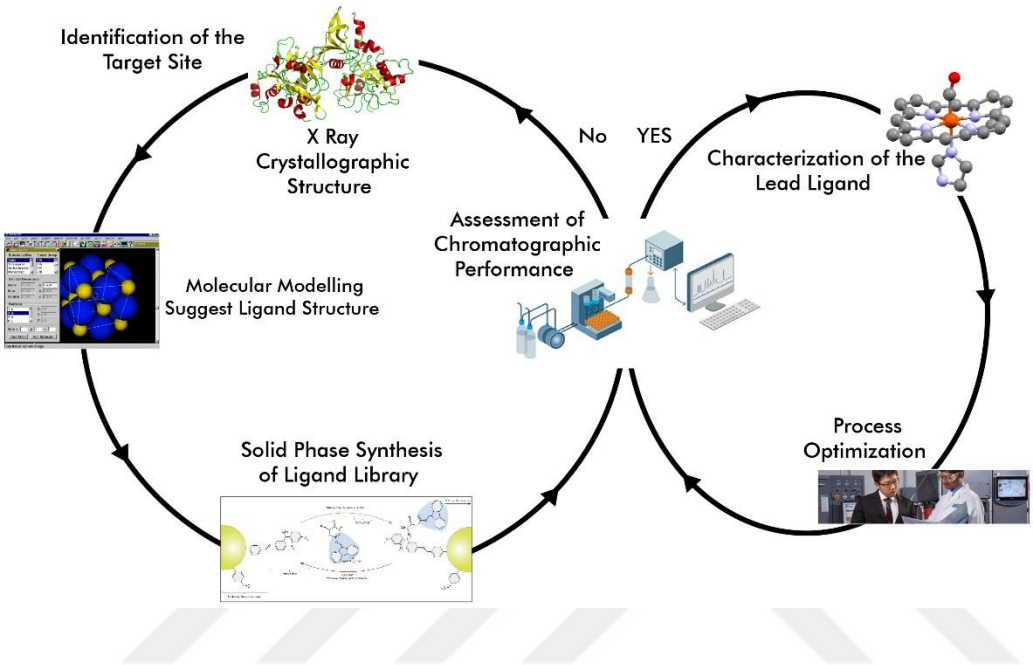

Figure 1. Research strategy for the design of de novo affinity ligands.

2. GENERAL INFORMATION

2.1. Affinity Chromatography

2.1.1. History of affinity chromatography

Affinity chromatography is a particular variant of chromatography in which the unique biological specificity and reversibility of the target analyte and ligand interaction is utilized for the separation [16].

The first article which published by the German scientist, Emil Starkenstein in 1910 described the concept of resolving macromolecule complexes via their interactions with an immobilized substrate. This manuscript discussed the influence of chloride on the enzymatic activity of liver α-amylase and was the corner stone of the early beginnings of this approach by several researchers [17].

Affinity chromatography was first used in the isolation of enzymes in 1953 by Lerman, who isolated tyrosinase on a column of cellulose with ethereally bound resorcinol residues. In subsequent year's affinity chromatography was employed only rarely, the reason clearly being the character of the insoluble supports that did not offer sufficient possibilities for complex formation between the product to be isolated and the attached affinant [18]. Affinity chromatography is still developing. It has played a central role in many " Omics" technologies, such as genomics, proteomics and metabolomics [17].

2.1.2. Fundamental principles of affinity chromatography

Affinity chromatography separates proteins on the basis of a reversible interaction between a protein (or group of proteins) and a specific ligand coupled to a chromatography matrix. The technique offers high selectivity, hence high resolution, and usually high capacity for the protein(s) of interest. Purification can be in the order of several thousand-fold and recoveries of active material are generally very high [19].

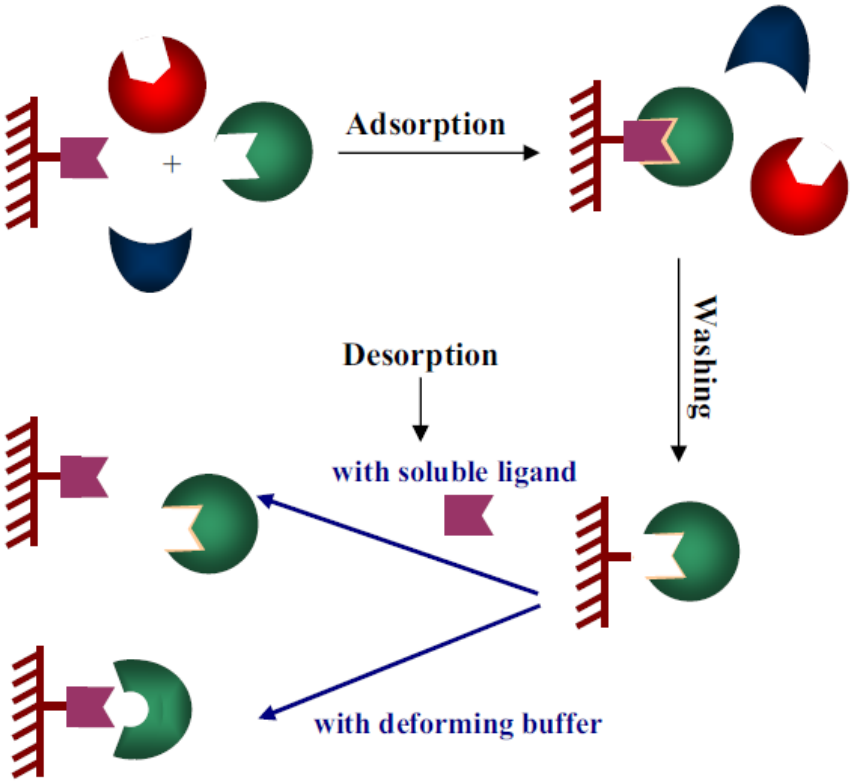

Figure 2. Schematic representation of the main steps in affinity chromatography [19].

2.1.3. General formats of affinity chromatographic applications

In the following figure, we can see the format that is used for both traditional and high pressure affinity chromatographic applications, in order to separate biomolecules.

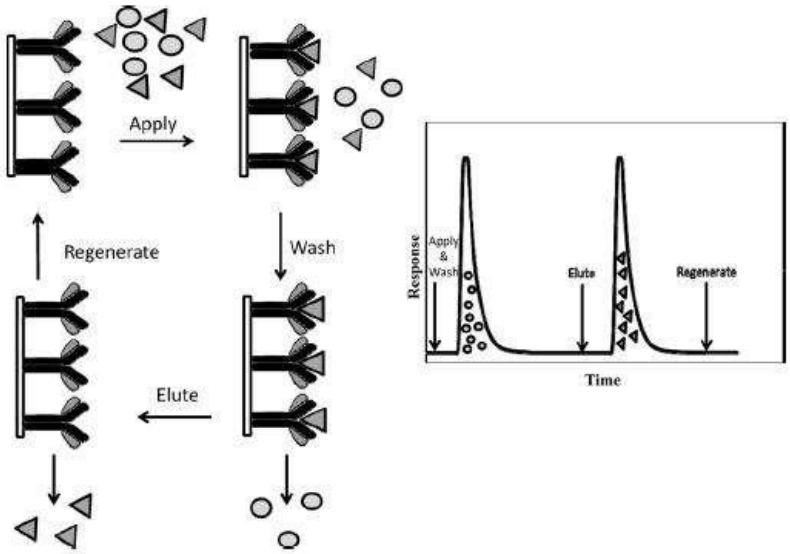

Figure 3 The format that is used for both traditional and high pressure affinity chromatographic applications [20].

The first step includes the injection of the sample in the affinity column, under conditions that let the target or the analyte of interest to bind to the immobilized ligand in a strong way. An aqueous buffer with pH and ionic strength, that simulates the inherent conditions of the ligand and its target, is regularly used in this kind of applications. During the use of the buffer, the elution of compounds in the sample, which do not interact much or at all with the ligand, is occurring and as a result a non-retained peak can be obtained [20].

Later, the dissociation of the target from the affinity ligand takes place by using an elution buffer. An essential requirement of this step is the alteration of the sample composition in the mobile phase in order to enable the elution of the target. This can be achieved either by adding a competing agent which is responsible for the displacement of the target from the column or by changing the pH conditions. The collection of the released target during this elution procedure is also possible and this target can be analyzed later in order to provide more information. The direct monitoring of the elution target, by using an HPLC support in the affinity column, is also feasible by an on-line method.

Combinations of both on-line and off-line methods with detection methods like absorbance, fluorescence or mass spectrometry are also possible. The

regeneration of the column before the next application can be made by passing through the original application buffer, after the elution of the target [21].

The scheme in the figure above is known as the on/off or step elution mode of affinity chromatography and it is responsible for the capture and the elution of the target [9]. This has been used in a wide range of applications not only in order to isolate compounds but also for the preparation procedure of the sample, especially in applications regarding biomedical and pharmaceutical analyses. The main reasons for this choice are the fact that this issue can be characterized as simple, selective, flexible, and relatively easy to use [22]. In addition, the automation of the format is easy when affinity columns, which are suitable for HPLC applications or as a part of HPLC systems, are used. Ultimately, this issue can be also used in order to detect analytes directly. In the step of elution mode, on-line absorbance or fluorescence detectors are used for this reason. Apart from these, mass spectrometric applications and post column reactors can also be used [23].

2.1.4. Biomolecules purified by affinity chromatography

In 1951, antibodies were first purified using affinity chromatography when Campbell et al. used affinity chromatography to isolate rabbit anti-bovine serum albumin antibodies. For their purification, bovine serum albumin was used as the affinity ligand on a cellulose support. Two years later, this technique was expanded to purify mushroom tyrosinase using an immobilized inhibitor of the enzyme (azophenol). Since then, affinity chromatography is commonly used to purify biomolecules such as enzymes, antibodies, recombinant proteins, and other biomolecules [17].

Affinity chromatography is a powerful tool for the purification of substances in a complex biological mixture. It can also provide separation of denatured and native forms of the same substance. Consequently, biomolecules which are difficult to purify have been obtained using bioselective adsorbents, e.g. immobilized metal ions (Ni^{2+} and Zn^{2+}) used to purify proteins containing zinc finger domains with natural affinity to divalent ions. The relative specificity degree of the affinity chromatography is due to the exploitation of biochemical properties inherent in certain molecules, instead of using small differences in physicochemical properties

(such as size, form and ionic charge, which are employed by other chromatographic methods) [24].

2.1.5. Immobilization of affinity ligands

When designing an affinity chromatography method for biomolecule purification, immobilization of the affinity ligand is also very important. Care must be taken when immobilizing an affinity ligand to ensure that the affinity ligand can actively bind the desired target after the immobilization procedure. Activity of the affinity ligand can be affected by multi-site attachment, orientation of the affinity ligand, and steric hindrance [17]. See Figure 4.

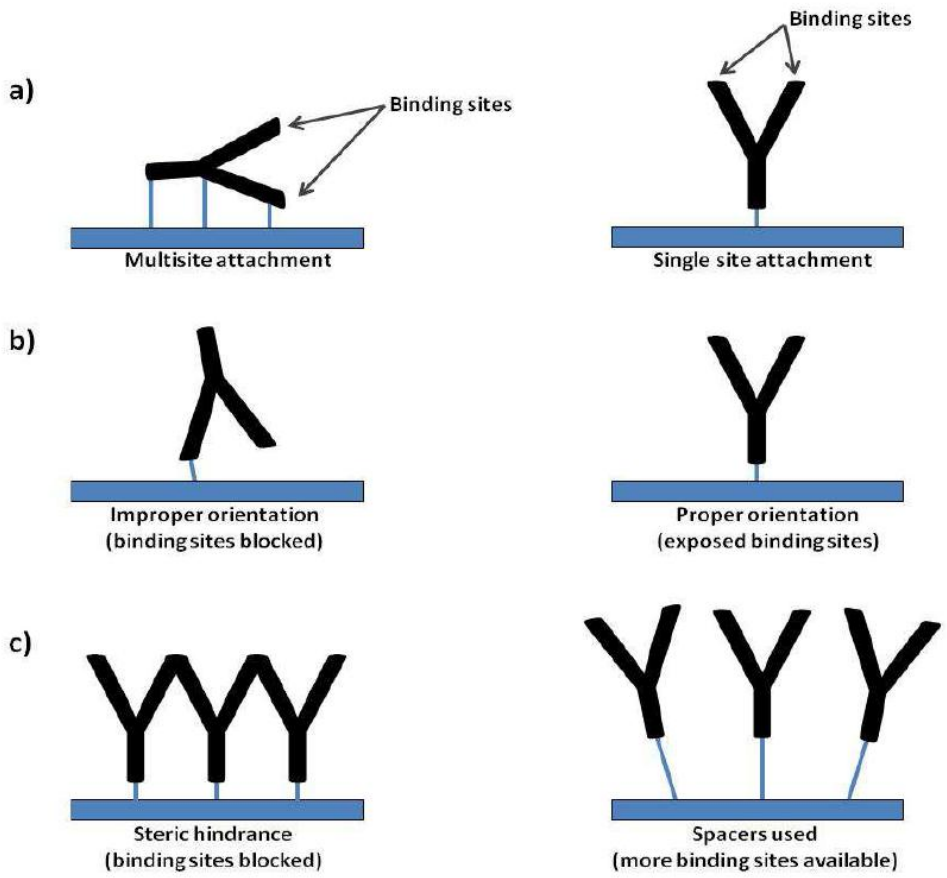

Figure 4. Immobilization of affinity ligands [17].

By contrast, most affinity purification strategies – especially those involving antibodies and other proteins – depend upon covalent chemical conjugation of ligands to the solid support matrix. Affinity ligands that have broad applicability are

commercially available in a variety of ready-to-use, pre-immobilized forms. Examples include Protein A agarose resin for general antibody purification and streptavidin magnetic beads for purifications involving biotinylated molecules.
Affinity chromatography utilizes the specific interactions between two molecules for the purification of a target molecule. A ligand having affinity for a target molecule is covalently attached to an insoluble support and functions as bait for capturing the target from complex solutions. The affinity ligand can be any molecule that will bind the target without also binding other molecules in the solution [25].

Ideally, immobilization methods which specifically avoid attaching the affinity ligand via functional groups within the binding site(s) are used. Undoubtedly, when performing affinity purifications, it is important to ensure the affinity ligands are immobilized so that the binding regions are exposed and free to interact and bind with the target molecules [17].

Affinity ligands have evolved from antibodies, enzymatic substrates, nucleic acids, cofactors, coenzymes, lectins, hormones, effectors, and inhibitors to a great diversity of small, low molecular weight peptides, polypeptides, and other organic structures. These newer classes of ligands can be made using biosynthetic and wholly synthetic methods. Common to all selection strategies is the need to begin with a variety of structures from which to discover candidate affinity ligands [26].

Affinity ligands can be covalently immobilized, adsorbed onto a surface via biospecific or nonspecific interactions, entrapped within a pore, or coordinated with a metal ion as in metal-ion affinity chromatography (IMAC).
One of the most common methods of attaching an affinity ligand to a solid support material is covalent immobilization. There is a wide range of coupling chemistries available when considering covalent immobilization methods. Amine, hydroxyl, sulfhydryl, carboxyl, and aldehyde groups have been used to link affinity ligands onto support materials [17].

2.2. Immobilized Metal Ion Affinity Chromatography

2.2.1. General Concepts of IMAC

Immobilized metal ion affinity chromatography (IMAC) firstly introduced by Porath, Carlsson, Olsson and Belfrage [27], is another type of affinity chromatography which has been showing huge growth when used in biomedical analysis [20]. IMAC was first designated as “metal chelate chromatography” [28] and later metal ion interaction chromatography and ligand-exchange chromatography before gaining its current term as immobilized metal ion affinity chromatography [29].

IMAC introduces a new approach for selectively interacting materials on the basis of their affinities for chelated metal ions [30]. Metal ion coordination with biological molecules is well suited to affinity adsorption due to its specificity and stability [31]. IMAC is centered on the interactions between immobilized metal ions and specific target groups from the protein surface, such as amino acids, peptides, proteins, and nucleic acids [20]. Moreover, it has been widely approved that concerning amino acids, tryptophan, histidine and cysteine residues play the most crucial role in the binding of proteins in IMAC, due to their strong interaction with metal-ions [32]. In IMAC a matrix is used and the metal ions are immobilized within a column through the use of chelating groups like iminodiacetic acid, nitrilotriacetic acid, carboxymethylatedaspartic acid, and L-glutamic acid. Metal ions that are often chelated to these groups include Ni^{2+}, Zn^{2+}, Cu^{2+} and Fe^{2+} [32].

IMAC was originally named as “metal chelate chromatography” because it was first developed for the separation and isolation of metal and histidine containing proteins. In this case, the sample is passed through an IMAC column, and firstly the targets that can bind to the immobilized metal ions are retained and later a competing agent is added or the pH is changed in order for the targets to be eluted. In spite of the use of IMAC for the purification of proteins, recently there is a wide variety of other areas where it can be applied [20].

To begin with, Takeda, Matsuoka and Gotoh report the employment of IMAC for the detection of drugs such as tetracyclines, quinolones, macrolides, β-lactams,

and aminoglycosides. Furthermore, Felix et al. investigate the use of IMAC for detection of biomarkers in the serum, urine, and tissues in order for diseases to be diagnosed. In this last case, the use of IMAC has been combined with mass spectrometry in surface-enhanced laser desorption/ionization (SELDI). Finally, in their work Sun, Chiu and He examine IMAC prior to mass spectrometry analysis for the enrichment of phosphoproteins [20].

Another recent interest of the researchers towards the immobilized metal ion affinity chromatography includes the refinement of IMAC for protein purification. IMAC has not been used a lot for large scale protein purification, thus researchers are highly interested in the investigation of the above and more specifically of unexpected conditions, for example what happens if there is a leakage of metal ions from IMAC columns. In this case, the metal ions may interfere with the purity of the protein raising concern for the researchers. Another issue of concern is the removal of the histidine tag from the recombinant protein when IMAC is used to isolate recombinant histidine-tagged proteins, and its possible co-elution with other proteins from the specific sample. As a result of the above, there is a great potential for improvements as far as IMAC applications are concerned. The development of new chelating ligands for IMAC, as well as methods to control better the selectivity of IMAC while proteins are isolated from samples, are two examples of new issues that should be investigated towards IMAC applications [20].

2.2.2. Components of IMAC

2.2.2.1. Metal Ions

A search of the literature on IMAC reveals a bewildering array of metal ions that have been used in this technique (e.g. Ag^{+}, Al^{3+}, Ca^{2+}, Co^{2+}, Cr^{3+}, Cu^{2+}, Eu^{3+}, Fe^{3+}, Hg^{2+}, La^{3+}, Mn^{2+}, Nd^{3+}, Ni^{2+}, Yb^{3+}, Zn^{3+}). The reason for this is that the nature of the metal ion (and indeed its chelator) influences the selectivity and affinity of the protein interaction [33].

The choice of the metal ion immobilized on the IMAC ligand relies on the application. While trivalent cations such as Al^{3+}, Ga^{3+}, and Fe^{3+} or tetravalent Zr^{4+} are preferred for capture of phosphoproteins and phosphopeptides, divalent Cu^{2+}, Ni^{2+}, Zn^{2+}, and Co^{2+} ions are used for purification of His-tagged proteins. Combinations of a tetradentate ligand that ensures strong immobilization, and a metal ion that leaves two coordination sites free for interaction with biopolymers (Ni^{2+}, Co^{2+}) has gained most acceptance and produces similar recovery and purity of eluted protein [34].

Although still not fully understood, the interaction used in IMAC depends on the formation of coordinated complexes between metal ions and electron donor groups on the protein surface. Some amino acids are especially suitable for binding and histidine is the one that exhibits the strongest interaction, as electron donor groups on the imidazole ring in histidine readily form coordination bonds with the immobilized transition metal. Cysteines can also contribute to binding if free sulfhydryl group are available in the appropriate, reduced state. Although also aromatic side chains of Trp, Phe and Tyr can interact with metal ions the actual protein retention in IMAC is based primarily on the availability of histidyl residues. Since many proteins contain these amino acids, it might be expected that all proteins are capable of binding to metal chelate columns. However, the residues must be located at the surface of the protein for successful coordination and the strength of interaction will depend on the number of such coordinations.

The borderline acids, containing Co^{2+}, Zn^{2+}, Cu^{2+} and Ni^{2+}, coordinate favorably with aromatic nitrogen atoms "borderline bases" and also with sulfur atoms "soft bases". The retention strength of the borderline metal cations, as chelated by iminodiacetate (IDA), follows the order Cu(II)>Ni(II)>Zn(II) ~Co(II). It may be noted that use of chelated metal ions displaying the highest protein retention does not necessarily translate into the best protein separation, since very high retention could also lead to increased adsorption of impurities [35].

2.2.2.2. Chelating Ligand

Metal chelators bound to chromatographic media fix the metal ion to a solid support, enabling the separation process to take place [33]. Chelating ligands are relatively inexpensive and capable of high metal ion loadings, which permit high protein-binding capacities. Moreover, the matrices can be re-used many times and can easily be regenerated by adding a buffer containing the specific metal ion. The ligands are stable over a wide range of temperatures and solvent conditions although reducing and chelating agents must be avoided as these readily displace chelated metal ion. The loss of metal ions is more pronounced at lower pH values. Apart from leading to reduced adsorption capacity, metal ions that leak from the sorbent can cause damage to the target proteins by metal-catalyzed reactions [34].

Several factors have been taken into consideration during the design of chelating ligands. Increasing the dentation number of the chelator will increase its affinity and reduce unwanted metal ion leakage from the column. Comparing with this, the need to provide free coordination sites for the protein with binding capacity and affinity increasing as the number of these sites increases. Furthermore, metal ion transfer must be avoided, i.e. the chelating ligand must bind the metal ion sufficiently tightly so as not to be stripped by proteins in the mixture to be purified [33].

The greater number of the chelating groups used in IMAC are multidentate chelating compounds providing the strength of the complex formed by the protein, metal ion and chelating group. The composition of the eluent buffer employed varies greatly when one is trying to find optimal conditions for a given protein separation, and in many cases, is the main factor of the specificity reached in some IMAC based purification protocols. These chelating substances are attached on the sorbent surface via spacers "linkage groups" which can differ in length and composition.

The final structure formed after the metal ion is chelated by the chelating group must allow some free coordination sites in the metal ion for the adsorption or binding of proteins or solvent molecules. Variation in the number of free coordination sites

may in part expound why some chelating substances (IDA, Iminodiacetic acid; Nitrilotriacetic acid, NTA, etc.) exhibit different selectivities and adsorption activities towards a given target protein. In the three-dentate IDA, the metal will bind to the nitrogen atom and the two carboxylate oxygens, leaving three sites for protein or solvent molecules "using Ni". In the same way, the tetradentate NTA is supposed to bind the chelating resin can have an important effect on metal ion with an extra carboxylate oxygen; this could give it a superior metal chelating strength, but on the other hand a weaker protein retention power [36]. Another feature of the tetradentate chelator would be a minor risk of metal leaching [37]. Fig.5 shows a model of the most commonly used chelators, IDA and NTA bound to Ni (II) atoms.

TACN has recently been introduced and used with a range of ‘soft’ metal ions. This chelator exhibits remarkable metal-binding stability at low pH, where other chelators would exhibit loss of the metal. This extended pH range could be used to gain further selectivity. The pentadentate TED offers very tight metal ion binding and highly selective protein binding. Furthermore, the strength of metal ion binding to TED can be exploited as a second column to remove potentially leached metal ions from other IMAC eluates [33].

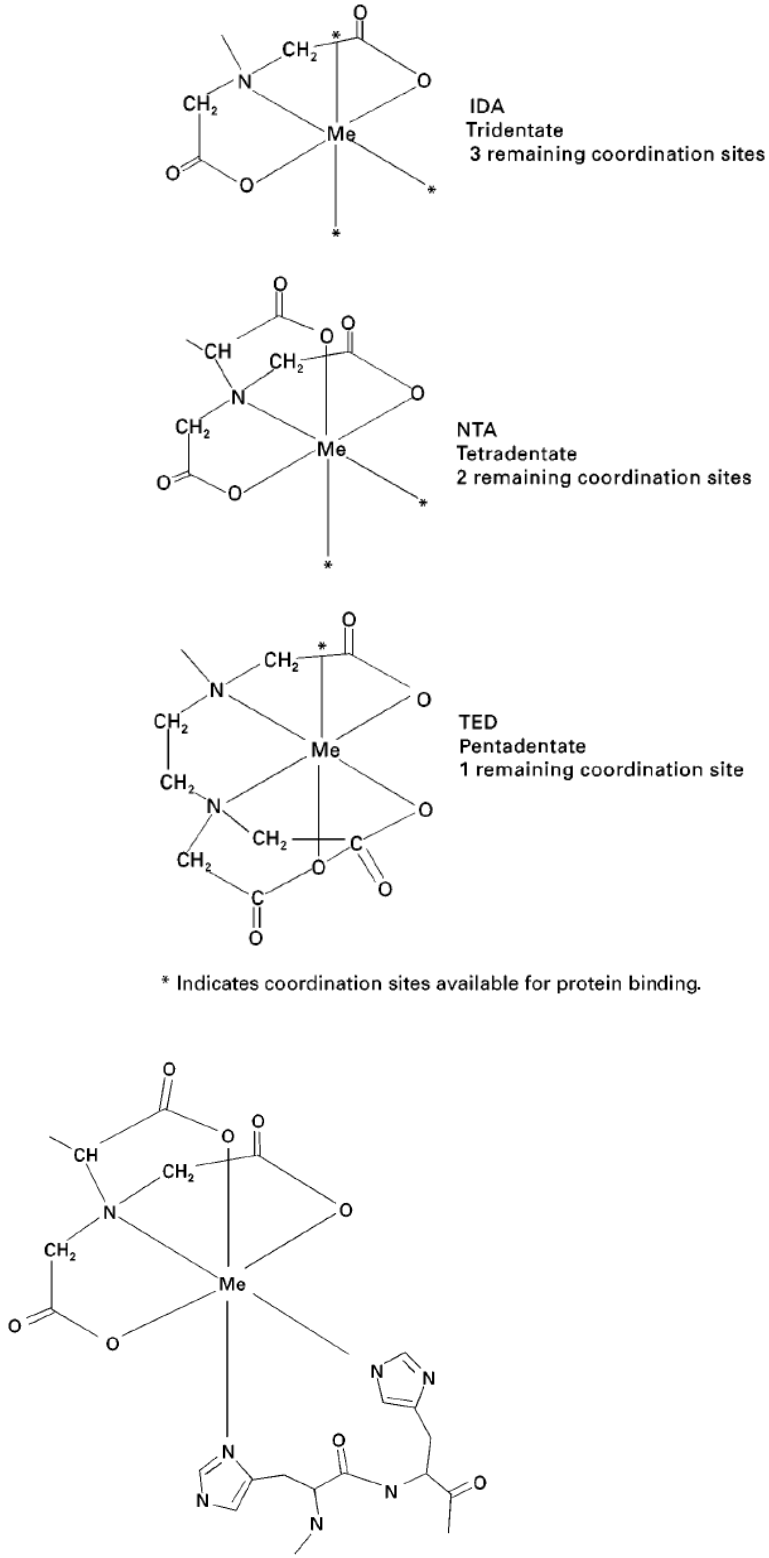

Figure 5. Schematic representation of IDA and NTA metal chelation [33].

Table 1. Abbreviation, names and functional structures of chelators.

Name	*Full name*	*Dentation*	*Structure*
IDA	Iminodiacetic acid	3	~ $\underline{N}(CH_2CO\underline{O}H)_2$
TACN	1,4,7-Triazocyclonanane	3	~ $N(CH_2CH_2\underline{N}H_2CH_2CH_2$ $\lfloor_____\rfloor$
NTA	Nitrilotriacetic acid	4	~ $CH(CO\underline{O}H)\underline{N}(CH_2CO\underline{O}H)_2$
TREN	Tris(2-aminoethyl)amine	4	~ $\underline{N}HCH_2CH_2\underline{N}(CH_2CH_2\underline{N}H_2)_2$
Talon	Proprietary	4	Proprietary
TED	Tris(carboxymethly)ethylenediamine	5	~ $\underline{N}(CH_2CO\underline{O}H)CH_2CH_2\underline{N}(CH_2CO\underline{O}H)_2$

~ Indicates chosen form of linkage to a chromatographic support, usually with a suitable spacer.
– Indicates atoms involved in metal ion coordination.

N-methacryloyl-L-histidine methyl ester (MAH) was developed as a functional ligand for affinity separation of proteins by introducing functional imidazole groups into polymer structure without any leakage [38]. MAH was used as the metal-coordinating monomer due to the affinity of imidazole nitrogen donor atoms towards metal ions [39].

2.2.2.2.1. Synthesis of N-Methacryloyl-L-histidine Methyl Ester (MAH)

The synthesis of MAH was adapted from the procedure reported by ref [35]. Briefly, 5.0 g of L-histidine methyl ester and 0.2 g of hydroquinone were dissolved in 100 mL of dichloromethane solution. This solution was cooled down to 8 °C, 12.7 g triethylamine was added followed by pouring slowly. 5.0mL of methacryloyl chloride and then the mixture was stirred magnetically for 2 h. At the end of this chemical reaction, amino acid methyl ester and triethylamine were extracted with 10% HCl. The product and methacryloyl chloride in dichloromethane were evaporated in a rotary evaporator and separated by column chromatography with chloroform/ethyl acetate mixture. The purified product was crystallized from ethyl acetate and cyclohexane. The structure of MAH was confirmed by NMR spectroscopy. The proton NMR spectrum of MAH monomer was taken in CDCl3 on a JEOL GX-400MHz instrument. The residual non-deuterated solvent tetramethylsilane (TMS) served as an internal reference. Chemical shifts are reported relative to TMS [38].

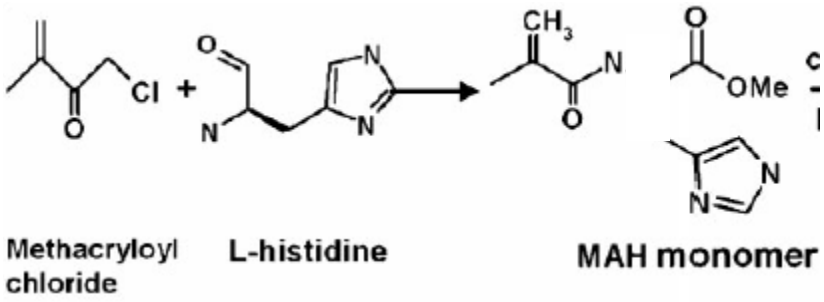

Figure 6. Synthesis of N-Methacryloyl-L-histidine Methyl Ester (MAH) monomer [38].

2.2.2.3. Media

Basically, the requirements of the support in IMAC are the same as also apply to affinity chromatography; easy to derivatize, exhibit no unspecific adsorption and have good physical, mechanical and chemical stability. Beaded agarose is the support predominantly used [34].

The first commercially available IMAC medium was IDA- Sepharose (AP Biotech). Today, many IDA chelating media are available, including modified forms of Sepharose (AP Biotech), agarose, polystyrene, polystyrene/divinylbenzene (Poros-Perseptive Biosystems), poly (alkalhydroxy-methacrylate), silica and even magnetic polystyrene beads (Dynabeads, Dynal Inc.). Among these types most are available as loose media and pre-packed columns for either low pressure or high performance liquid chromatography [33].

2.3. IMAC Applications

Initially developed for purification of native proteins with an intrinsic affinity to metal ions, IMAC has turned out to be a technology with a very big field of applications. On the chromatographic purification side, the range of proteins was expanded from the primary metalloproteins to antibodies, recombinant His-tagged proteins and phosphorylated proteins. IMAC is being used in proteomics approaches where fractions of the cellular protein pool are enriched and analyzed differentially (phosphoproteome, metalloproteome) by mass spectrometric techniques; here,

IMAC formats can be traditionally bead based or the ligand can be used on functionalized surfaces such as SELDI "surface-enhanced laser desorption/ionization" chips. Other chip-based applications include surface plasmon resonance "SPR" and allow the immobilization of His-tagged proteins for quantitative functional and kinetic investigations. Besides, the IMAC principle has been used as an inhibitor depletion step prior to PCR amplification of nucleic acids from complex samples such as blood in a technology called Chelex [40].

Nowadays, the His6 tag, consisting of six consecutive histidine residues, is one of the most commonly used tags for facilitated purification of recombinant proteins. The number of free coordination sites of the immobilized metal ion determines how many histidines that can bind concurrently. However, more histidine residues will enhance the probability for the correct orientation desired for coordination interactions. Since abundant neighboring histidine residues are uncommon among naturally occurring proteins, oligo-histidine affinity handles form the basis for high selectivity and efficiency, often providing a one-step isolation of proteins at over 90% purity. An ideal affinity tag should enable effective but not too strong binding, and allow elution of the desired protein under mild, nondestructive conditions. In the case of recombinant E. coli many host proteins strongly adhere to the IMAC matrices and are eluted with the target proteins. Consequently, new approaches for selecting improved histidine tags have focused on elution of the target protein in the contaminant-free window [34].

2.4. Importance of Using Computational Simulation

To better understand the basis of the activity of any molecule with biological activity, it is important to know how this molecule interacts with its site of action, more specifically its conformational properties in solution and orientation for the interaction. Molecular recognition in biological systems relies on specific attractive and/or repulsive interactions between two partner molecules. This study seeks to identify such interactions between ligands and their host molecules, typically proteins, given their three-dimensional (3D) structures.

Therefore, it is important to know about interaction geometries and approximate affinity contributions of attractive interactions. In addition, it is necessary to be aware of the fact that molecular interactions behave in a highly non-additive fashion.

Molecular interactions including protein-protein, protein-nucleic acid, enzyme-substrate, drug-protein, and drug-nucleic acid play significant roles in many essential biological processes, such as signal transduction, transport, gene expression control, cell regulation, antibody–antigen recognition, enzyme inhibition, and even the assembly of multi-domain proteins.
These interactions very often cause the formation of stable protein–protein or protein-ligand complexes that are essential to achieve their biological functions. The tertiary structure of proteins is necessary to understand the binding mode and affinity between interacting molecules. In the other hand, it is often hard and expensive to obtain complex structures by experimental methods, such as X-ray crystallography or NMR. Therefore, docking computation is considered a significant approach for understanding the protein-protein or protein-ligand interactions [41-43]. The number of three-dimensional protein structures determined by experimental techniques grows concurrently with structure databases such as Protein Data.

Molecular modeling encompasses all theoretical methods and computational technics used to model or mimic the behavior of molecules. The technics are used in the fields of computational chemistry, computational biology and materials science for studying molecular systems ranging from small chemical systems to large biological molecules and material assemblies. The simplest calculations can be performed by hand, but inevitably computers are required to perform molecular modeling of any reasonably sized systems the common feature of molecular modeling technics is the atomistic level description of the molecular systems, the lowest level of information is individual atoms (or a small group of atoms). This is in contrast in quantum chemistry (also known as electronic structure calculations) where electrons are considered explicitly. The benefit of molecular modeling is that

it reduces the complexity of the system allowing many more particles (atoms) to be considered during simulations [44].

Bank (PDB) and Worldwide Protein Data Bank (wwPDB) have over 88000 protein structures, many of which play important roles in critical metabolic pathways that may be considered as potential therapeutic targets and specific databases containing structures of binary complexes become available, in addition to information about their binding affinities, such as in PDBBIND [45], PLD [46], AffinDB [47] and BindDB [48], molecular docking procedures improve, getting more significance than ever [49].

Molecular docking is regarded as a widely-used computer simulation procedure to predict the conformation of a receptor-ligand complex, where the receptor is usually a protein or a nucleic acid molecule and the ligand is either a small molecule or another protein (Figure 7).

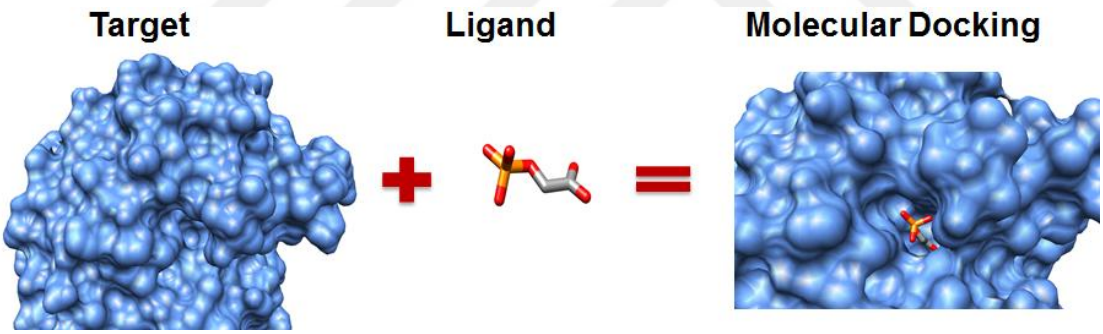

Figure 7. Elements in molecular docking.

2.5. Importance of Docking Studies for Immobilized Metal Ion Affinity Chromatography

The focus of molecular docking is to computationally simulate the molecular recognition process. Molecular docking aims to achieve an optimized confirmation for both the ligand and protein and relative orientation between protein and ligand such that the free energy of the overall system is minimized [44].

In Immobilized metal ion affinity chromatography (IMAC) the electron-rich compounds can form stable complexes with many transition metal ions, and O, N

and S containing molecules can bind to those ions by ion dipole interactions. Given these molecules with 3D conformations in atomic detail, it is important to know if the molecules bind to each other and, if it is so, what does the formed complex look like("docking") and how strong is the binding affinity (that can be related to the scoring functions.

In this context, computational modelling of interactions between functional ligands and metal ions has a great importance in terms of effectiveness and productivity of chromatographic separation system. It can contribute to the improvement of the efficiency of immobilized affinity chromatography methods. In addition, it can open the way to applicability of other molecules in affinity methods.

2.6. AutoDock Program

AutoDock is a suite of automated docking tools. It is designed to predict how small molecules, such as drug candidates or substrates, bind to a receptor of known 3D structure. AutoDock uses Monte Carlo simulated annealing and Lamarckian genetic algorithm to create a set of possible conformations. LGA is used as a global optimizer and energy minimization as a local search method. Current distributions of AutoDock consist of two generations of software: AutoDock 4 and AutoDock Vina [50].

AutoDock 4 actually consists of two main programs:

1) *AutoDock* performs the docking of the ligand to a set of grids describing the target protein;
2) *autogrid* pre-calculates these grids. In addition to using them for docking, the atomic affinity grids can be visualized. This can help, for example, to guide organic synthetic chemists design better binders [24].

A graphical user interface called AutoDockTools, or ADT for short, is developed in order to allow to set up which bonds will be treated as rotatable in the ligand and to analyze dockings [50].

AutoDock has applications in:

1. X-ray crystallography
2. structure-based drug design
3. lead optimization
4. virtual screening (HTS)
5. combinatorial library design
6. protein-protein docking
7. Chemical mechanism studies [50].

AutoDock 4 is free and is available under the GNU General Public License. AutoDock 4.2 is faster than earlier versions, and it allows side chains in the macromolecule to be flexible. As before, rigid docking is blindingly fast, and high quality flexible docking can be achieved in around a minute. Up to 40,000 rigid dockings can be done in a day on one CPÚ [50].

AutoDock 4.2 now has a free energy scoring function that is centered on a linear regression analysis, the AMBÉR force field, and an even larger set of various protein-ligand complexes with known inhibition constants than it is used in AutoDock 3.0. The best model was cross-validated with a separate set of HIV-1 protease complexes, and confirmed that the standard error is around 2.5 kcal/mol. This is enough to discriminate between leads with milli-, micro- and nano-molar inhibition constants [50].

AutoDock is free software. The introduction of AutoDock 4 includes three major improvements:

1- The docking results are more reliable and accurate.
2- It can optionally model flexibility in the target macromolecule.
3- It enables AutoDock 's use in evaluating protein-protein interactions.

AutoDock 4.0 can be compiled to take advantage of new search methods from the optimization library, ACRO, developed by William E. Hart at Sandia National Labs [50].

AutoDock is being used in academic, governmental, non-profit and commercial settings. It has now been distributed to more than 29000 users around the world. In January of 2011, a search of the ISI Citation Index showed more than 2700 publications have cited the primary AutoDock methods papers [50].

AutoDock is now distributed under the GPL open source license and is freely available for all to use. Some companies may wish to license AutoDock under a separate license agreement because of the restrictions of incorporating GPL licensed software into other codes for the purpose of redistribution, [50].

AutoDock has been widely-used and there are many examples of its successful application in the literature; in 2006, AutoDock was the most cited docking software. It is very fast, provides high quality predictions of ligand conformations, and good correlations between predicted inhibition constants and experimental ones. AutoDock has also been shown to be useful in blind docking, where the location of the binding site is not known. Plus, AutoDock is free software and version 4 is distributed under the GNU General Public License; it easy to obtain, too [50].

2.7. Avogadro Program

Avogadro is an advanced molecule editor and visualizer designed for cross-platform use in computational chemistry, bioinformatics, molecular modeling, materials science, and related areas. It offers flexible high quality rendering and a strong plugin architecture [51].

- Cross-Platform: Molecular builder- editor for Windows, Línux, and Mac OS X.
- Free, Open Source: Easy to install and all source code is available under the GNU GPL.
- International: Translations into Chinese, French, German, Italian, Russian, Spanish, and others.
- Intuitive: Built to work easily for students and advanced researchers both.
- Fast: Supports multi-threaded rendering and computation.

- Extensible: Plugin architecture for developers, including rendering, interactive tools, commands, and Python scripts.
- Flexible: Features include Open Babel import of chemical files, input generation for multiple computational chemistry packages, biomolecules, and crystallography [51].

2.8. Chimera program

UCSF Chimera is a highly extensible, interactive molecular visualization and analysis system. Chimera can read molecular structures and associated data in a large number of formats, display the structures in a variety of representations, and produce high quality images and animations suitable for publication and presentation. Furthermore, Chimera offers tools to show density maps and analyze microscopy data, display multiple sequence alignments, with crosstalk between the sequences and structures; utilize symmetry information for the display of higher order structures, and enable analysis of molecular dynamics trajectories and docking results [52].

Chimera is distributed with full documentation and a number of tutorials, and can be downloaded free of charge for academic, government, nonprofit, and personal use. Chimera is available for several platforms, including Windows, MacOS X, and Linux [52].

Chimera is supported and developed by the Resource for Biocomputing, Visualization, and Informatics and is funded by the NIH National Center for Research Resources [52].

2.9. PyMOL Program

PyMOL is a free cross-platform molecular graphics system made possible through recent advances in hardware, internet, and software development technology. PyMOL provides most of the capabilities and performance of traditional molecular graphics packages written in C or FORTRAN.

PyMOL was originally designed to:

1) visualize multiple conformations of a single structure [trajectories or docked ligand ensembles]

2) interface with external programs,

3) provide professional strength graphics under both Windows and Unix,

4) prepare publication quality images,

5) and suit into a tight budget.

PyMOL is one lone scientist's answer to the frustration he encountered with existing visualization and modeling software as a practicing computational scientist. Anyone who has studied the remarkable complexity of a macromolecular structure will likely agree that visualization is essential to understanding structural biology. Nevertheless, most researchers who use visualization packages ultimately run up against limitations inherent in them which make it difficult or impossible to get exactly what you need. Such limitations in a closed−source commercial software package cannot be easily surmounted, and the same is still true for free programs which aren't available in source form.

Although PyMOL is far from perfect and lacks such desirable features like a general (undo) capacity, it now has many helpful capabilities for the practicing research scientist.

PyMOL was created in an efficient but highly pragmatic manner, with heavy emphasis on delivering powerful features to end users. Expediency has almost always taken precedence over elegance, and adherence to established software development practices is inconsistent. PyMOL is about getting the job done now, as fast as possible, by whatever means were available. PyMOL succeeds in meeting important needs today [53].

2.10. Insulin

Insulin (from the Latin, insula meaning island) is a peptide hormone produced by beta cells in the pancreas, and by Brockmann body in some teleost fish [54]. It regulates the metabolism of carbohydrates and fats by promoting the absorption of

glucose from the blood to skeletal muscles and fat tissue and by causing fat to be stored rather than used for energy. Insulin also inhibits the production of glucose by the liver [55].

Except in the presence of the metabolic disorder diabetes mellitus and metabolic syndrome, insulin is provided within the body in a constant proportion to remove excess glucose from the blood, which otherwise would be toxic. When blood glucose levels fall below a certain level, the body begins to use stored glucose as an energy source through glycogenolysis, which breaks down the glycogen stored in the liver and muscles into glucose, which can then be utilized as an energy source. As a central metabolic control mechanism, its status is also used as a control signal to other body systems (such as amino acid uptake by body cells). In addition, it has several other anabolic effects throughout the body.

When control of insulin levels fails, diabetes mellitus can result. As a consequence, insulin is used medically to treat some forms of diabetes mellitus. Patients with type 1 diabetes depend on external insulin (most commonly injected subcutaneously) for their survival because the hormone is no longer produced internally [56]. Patients with type 2 diabetes are often insulin resistant and, because of such resistance, may suffer from a "relative" insulin deficiency. Some patients with type 2 diabetes may eventually require insulin if dietary modifications or other medications fail to control blood glucose levels adequately. Over 40% of those with Type 2 diabetes require insulin as part of their diabetes management plan.

2.10.1. Structure of Insulin

Like most of the other hormones, insulin is a protein comprising of 2 polypeptides chains A (with 21 amino acid residues) and B (with 30 amino acid residues) [Figure 8]. Chains A and B are linked by disulphide bridges. In addition, A-chain contains an intra-chain disulphide bridge linking residue 6 and 11. The structure of insulin is shown in the figure below. C-chain, which connects A and B chains is liberated along with insulin after breakdown of proinsulin. Insulin monomers aggregate to form dimers and hexamers [57]. Zn hexamer is composed of three insulin dimmers associated in threefold symmetrical pattern.

Insulin is produced and stored in the body as a hexamer (a unit of six insulin molecules), while the active form is the monomer. The hexamer is an inactive form with long-term stability, which serves as a way to keep the highly reactive insulin protected, yet readily available. The hexamer-monomer conversion is one of the central aspects of insulin formulations for injection. The hexamer is far more stable than the monomer, which is desirable for practical reasons; however, the monomer is a much faster-reacting drug because diffusion rate is inversely related to particle size. A fast-reacting drug means insulin injections do not have to precede mealtimes by hours, which in turn gives people with diabetes more flexibility in their daily schedules [58].

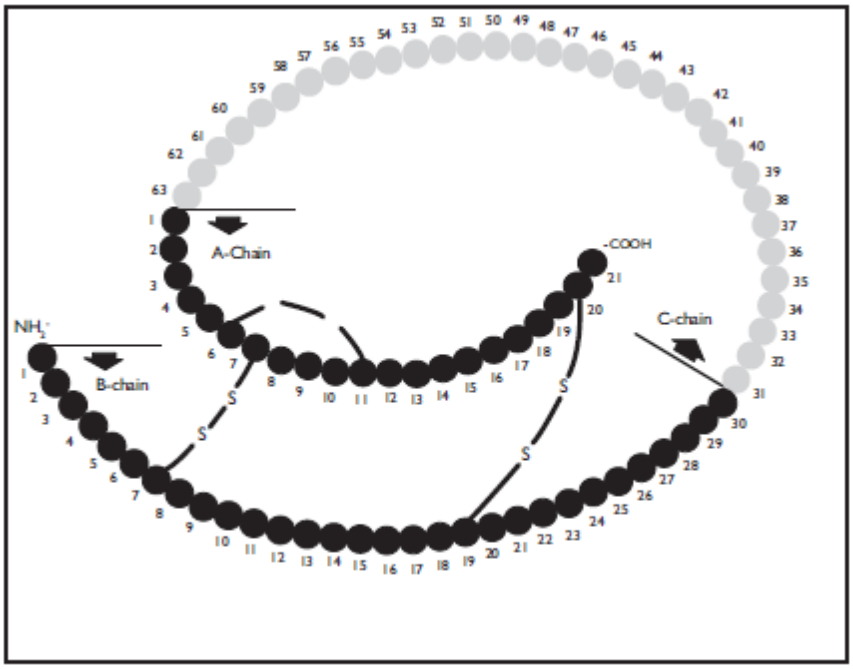

Figure 8. Structure of Insulin.

2.11. Cytochrome C

The cytochrome complex, or cyt C is a small hemeprotein found loosely associated with the inner membrane of the mitochondrion. It belongs to the cytochrome c family of proteins. Cytochrome C is a highly water soluble protein, unlike other cytochromes and is an essential component of the electron transport chain, where it carries one electron. It is capable of undergoing oxidation and reduction, but does not bind oxygen. It transfers electrons between Complexes III (Coenzyme Q – Cyt

C reductase) and IV (Cyt C oxidase). In humans, cytochrome c is encoded by the CYCS gene [59].

Ambler [60] recognized four classes of cyt C:

• **Class I** includes the lowspin soluble cyt C of mitochondria and bacteria, with the haem attachment site towards the Nterminus, and the sixth ligand provided by a methionine residue about 40 residues further on towards the Cterminus. The proteins contain three conserved "core" helices which form a "basket" around the haem group with one haem edge exposed to the solvent.

• **Class II** includes the highspin cyt C' and a number of lowspin cytochromes, e.g. cyt c556. The haem attachment site is close to the Cterminus. The protein fold comprises a four-helix bundle [61].

• **Class III** comprises the low redox potential multiplehaem cytochromes: cyt c7 (trihaem), c3 (tetrahaem), and high molecular weight cyt C (HMC; hexadecahaem), with only 30-40 residues per haem group. The haem c groups, all bisHis coordinated, are structurally and functionally nonequivalent and present different redox potentials in the range 0 to -400 mV [62].

• **Class IV** was originally created to hold the complex proteins that have other prosthetic groups as well as haem c, e.g. flavocytochrome c and cytochromes cd. Alternatively, Moore and Pettigrew [60] have suggested that Class IV cyt C are tetrahaem proteins containing both bisHis and HisMet coordinated haems, with a 3D structure exemplified by that of the photosynthetic reaction center (PRC) cyt C, and form a structurally homogeneous family.

2.11.1. Cytochrome C Structure

Mitochondrial cytochromes c is the most extensively studied electron-transfer proteins. For example, the cytochrome c amino acid sequences have been determined for organisms including humans, chimps, rhesus monkeys, spider monkeys, horses, donkeys, zebras, cows, pigs, sheep, camels, great whales, elephant seals, dogs, hippos, bats, rabbits, bull frogs, starfish and fruit flies.

Cytochrome c is easily separated from its mitochondrial environment because of its solubility in water. Furthermore, cytochrome c is weakly associated with the inner

membrane space. Cytochrome c is readily available in pure and native form, although it is expensive. The crystal structures of mitochondrial cytochrome c from several sources have been determined to atomic resolution (Figure 9). The interpretation of the physiochemical properties of cytochrome c is facilitated by the knowledge of conformation [63].

Cytochrome c (MW 12, 400) consists of a single polypeptide chain of 104 amino acid residues and covalently attached to a heme group. Cytochrome c has 19 positively charged lysine residues, plus two arginines also positively charged, but only 12 acidic residues (aspartic or glutamic acids). Cytochrome c is very basic with an isoelectric point near pH 10. Isoelectric point is the pH at which the number of positive charges and the number of negative charges of a compound are equal [64].

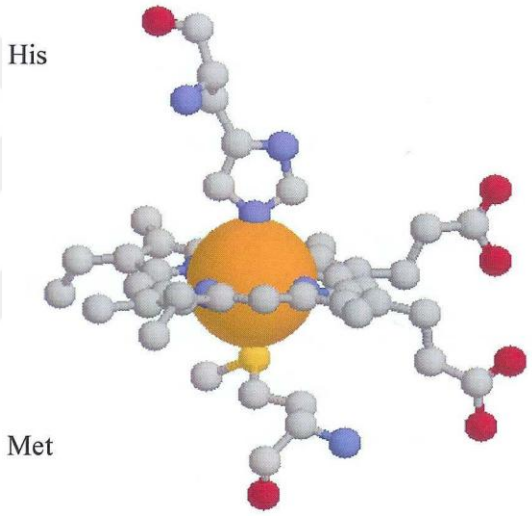

Figure 9. The Crystal Structures of Mitochondrial Cytochrome C [63].

3. EXPERIMENTAL

3.1. Hardware and software

The study was carried out on a Lenovo workstation with a 1.70 GHz processor, 4-GB RAM, and 500-GB hard drive running a Windows operating system. Bioinformatics software, such as AutoDock4.2, and other programs were used in this study.

3.2. Preparation of macromolecules

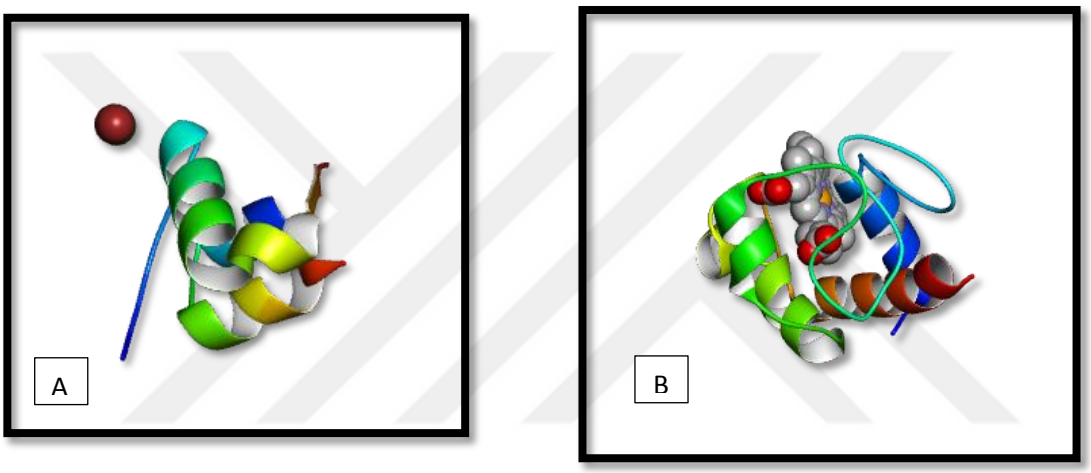

Figure 10. Ribbon Structures of Human Insulin (A), and Cytochrome C (B) [49].

ADT checks that the molecule has charges. If not, it adds Gasteiger charges to each atom. Remember that all hydrogens must be added to the macromolecule before it is chosen. If the molecule already had charges, ADT would ask if you want to preserve the input charges instead of adding Gasteiger charges.

ADT also determines the types of atoms in the macromolecule. AD4 can accommodate any number of atom types in the macromolecule.

3.3. Preparation of ligand

N-Methacryloyl-L-histidine Methyl Ester (MAH) is used as functional ligand. The 3D structure of MAH was not available in any databases. Hence, this molecule was drawn and created using Avogadro which has auto optimization tool.

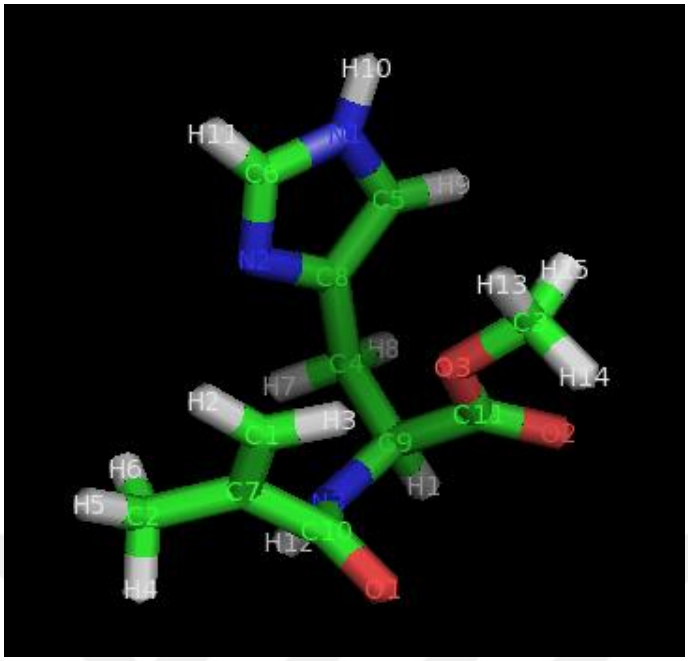

Figure 11. N-Methacryloyl-L-histidine Methyl Ester (MAH) drawn by Avogadro.

3.4. AutoDock

Graphical User Interface called AutoDockTools, or ADT, is used to help us easily set up the two molecules for docking, and when the dockings are completed also lets the user interactively visualize the docking results in 3D.

3.4.1. Editing a PDB file

Protein Data Bank (PDB) files can have a variety of potential problems that need to be corrected before they can be used in **AutoDock**. These potential problems include missing atoms, added waters, more than one molecule, chain breaks, alternate locations *etc.*

AutoDockTools (ADT) is built on the Python Molecule Viewer (PMV), and has an evolving set of tools designed to solve these kinds of problems. In particular, two modules, **editCommands** and **repairCommands**, contain many useful tools which permit you to add or remove hydrogens, repair residues by adding

missing atoms, modify histidine protonation, modify protonation of intrachain breaks, *etc.*

3.4.2. Preparing a ligand file for AutoDock.

AutoDock ligands have partial atomic charges for each atom. We also distinguish between aliphatic and aromatic carbons: names for aromatic carbons start with 'A' instead of 'C'. AutoDock ligands are written in files with special keywords recognized by AutoDock. The keywords ROOT, ENDROOT, BRANCH, and ENDBRANCH establish a "torsion tree" object or **torTree** that has a **root** and **branches**. The root is a rigid set of atoms, while the branches are rotatable groups of atoms connected to the rigid root. The keyword TORSDOF signals the number of torsional degrees of freedom in the ligand. The TORSDOF for a ligand is the total number of possible torsions in the ligand minus the number of torsions that only rotate hydrogens. TORSDOF is used in calculating the change in free energy caused by the loss of torsional degrees of freedom upon binding.

3.4.3. Docking

Automated docking was used to determine appropriate binding orientations and conformations of various ligands at the macromolecule. Autodock 4.2 was used for docking of MAH molecules with human insulin and cytochrome c respectively, and Lamarckian Genetic Algorithm (LGA) was used to determine the globally optimized confirmation. Polar hydrogen atoms were added, and Kollman charge, atomic solvation parameters, and fragmental volumes were assigned to the protein using Autodock tools. The grid spacing was 0.375 Å for each spacing; each grid map consisted of 126 × 126 × 126 grid points. During each docking experiment, 10 runs were performed, and the population size was set at 150; maximum number of evaluation, 2,500,000; maximum number of generations, 27,000; rate of gene mutation, 0.02; and cross-over rate, 0.8. The remaining parameters were set as default. A root mean square deviation (RMSD) tolerance for each docking was set at 2.0 Å. Every ligand molecule had 0.2983 coefficients of torsional degrees of freedom for docking. At the end

of docking, a cluster analysis was performed. For docking of each ligand, all the confirmations were clustered together and ranked by the lowest binding energy. These docked complexes were subjected to further analysis.

3.4.4. Analyzing AutoDock Results-Reading Docking Logs

Reading a docking log or a set of docking logs is the first step in analyzing the results of docking experiments. (By convention, these results files have the extension ".dlg".) During its automated docking procedure, AutoDock outputs a detailed record to the file specified after the –l parameter. The output includes many details about the docking which are output as AutoDock parses the input files and reports what it finds. For example, for each AutoGrid map, it reports opening the map file and how many data points it read in.

When it parses the input ligand file, it reports building various internal data structures. After the input phase, AutoDock begins the specified number of runs. It reports which run number it is starting; it may report specifics about each generation. After completing the runs, AutoDock begins an analysis phase and records details of that process. At the very end, it reports a summary of the amount of time taken and the words 'Successful Completion'. The level of output detail is controlled by the parameter (outlev) in the docking parameter file. For dockings using the GA-LS algorithm, outlev 0 is recommended. The key results in a docking log are the docked structures found at the end of each run, the energies of these docked structures and their similarities to each other. The similarity of docked structures is measured by computing the root-mean-square-deviation, **rmsd**, between the coordinates of the atoms. The docking results consist of the PDBQ of the Cartesian coordinates of the atoms in the docked molecule, along with the state variables that describe this docked conformation and position.

3.4.5 Analyzing AutoDock Results-Visualizing Docked Conformations

The 'best' docking result can be considered to be the conformation with the lowest (docked) energy. Alternatively, it can be selected based on its rms deviation from a reference structure. At the end of each docking run, AutoDock outputs a result which is the lowest energy conformation of the ligand it found

during that run. This conformation is a combination of translation, quaternion and torsion angles and is characterized by intermolecular energy, internal energy and torsional energy. The first two of these combined give the '**docking energy**' while the first and third give '**binding energy**.' AutoDock also breaks down the total energy into a vdW energy and an electrostatic energy for each atom.

3.4.6. Analyzing AutoDock Results-Clustering Conformations

An AutoDock docking experiment usually has several solutions. The reliability of a docking result depends on the similarity of its final docked conformations. One way to measure the reliability of a result is to compare the rmsd of the lowest energy conformations and their rmsd to one another, to group them into families of similar conformations or "clusters."

The dpf keyword, **analysis**, determines whether clustering is done by AutoDock. It is also possible to cluster conformations with ADT. By default, AutoDock clusters docked results at 0.5 Å rmsd. This process involves ordering all of the conformations by docked energy, from lowest to highest. The lowest energy conformation is used as the seed for the first cluster.

Next, the second conformation is compared to the first. If it is within the rmsd tolerance, it is added to the first cluster. If not, it becomes the first member of a new cluster. This process is repeated with the rest of the docked results, grouping them into families of similar conformations.

3.5. Visualization

The visualization of structure files was carried out using the graphical interface of the ADT program, PyMol molecular graphics system and Chimera molecular visualization and analysis system.

4. RESULTS AND DISCUSSION

AutoDock 4.2 was used to dock ligands to identify the active entities and determine the binding sites in target proteins. Lamarckian Genetic Algorithm (LGA) for docking was implemented with defined parameters for determining the docking performance. The output of molecular docking was clustered to determine the binding free energy (BE) and optimal docking energy conformation that is considered as the best docked structure, as well as to elucidate their binding state in the receptor.

4.1. Docking of One molecule of N-Methacryloyl-L-histidine Methyl Ester (MAH) at Human Insulin Molecule

4.1.1. Cluster Analysis of Conformations

Number of conformations = 10

RMSD cluster analysis will be performed using the ligand atoms only (19 / 19 total atoms).

Outputting structurally similar clusters, ranked in order of increasing energy.

The docking experiments were carried out at 298.15 K.

4.1.1.1. Clustering Histogram

Table 4.1. Clustering Histogram of docking of one molecule of (MAH) to human insulin molecule.

Cluster Rank	Lowest Binding Energy (Kcal/mol)	Run	Mean Binding Energy (Kcal/mol)	Number in Cluster
1	-3.43	7	-3.43	1
2	-3.35	5	-3.35	1
3	-3.23	2	-3.23	1
4	-3.12	3	-3.12	1
5	-2.87	8	-2.87	1
6	-2.81	1	-2.81	1
7	-2.80	4	-2.80	1

8	-2.77	9	-2.77	1
9	-2.75	6	-2.75	1
10	-2.74	10	-2.74	1

4.1.1.2. RMSD Table

Table 4.2. RMSD Values and Binding Energies of Runs.

Rank	Sub-Rank	Run	Binding Energy(Kcal/mol)	Cluster RMSD	Reference RMSD
1	1	7	-3.43	0.00	19.05
2	1	5	-3.35	0.00	16.09
3	1	2	-3.23	0.00	16.68
4	1	3	-3.12	0.00	11.47
5	1	8	-2.87	0.00	18.15
6	1	1	-2.81	0.00	18.01
7	1	4	-2.80	0.00	12.87
8	1	9	-2.77	0.00	16.66
9	1	6	-2.75	0.00	12.10
10	1	10	-2.74	0.00	13.87

The key results in a docking log are the docked structures found at the end of each run, the energies of these docked structures and their similarities to each other. The similarity of docked structures is measured by computing the root-mean-square-deviation, **rmsd**, between the coordinates of the atoms.

In Table 4.2, the similar rmsd values were obtained for 10 runs, in which they were grouped into families of similar clusters. The lowest binding energy of MAH-Insulin docking (-3.43 kcal/mol) was obtained for the 7th rank.

4.1.2. Information Entropy Analysis for This Clustering

Information entropy for this clustering = 1.00 (rmstol = 2.00 Angstrom)

4.1.3. Statistical Mechanical Analysis

Partition function: Q = 10.05 at Temperature, T = 298.15 K

Free energy: A ~ -1367.22 kcal/mol at Temperature, T = 298.15 K

Internal energy: U = -2.99 kcal/mol at Temperature, T = 298.15 K

Entropy: S = 4.58 kcal/mol/K at Temperature, T = 298.15 K

4.1.4. Lowest Energy Docked Conformation

Table 4.3. Lowest Energy Docked Conformation**.**

Run	7
Cluster Rank	1
Number of conformations in this cluster	1
RMSD from reference structure	19.050 A
Estimated Free Energy of Binding	**-3.43 kcal/mol**
Estimated Inhibition Constant, Ki	3.07 uM (micro molar) [Temperature = 298.15 K]
Final Intermolecular Energy	-5.22 kcal/mol
vdW + Hbond + desolv Energy	-5.11 kcal/mol
Electrostatic Energy	-0.11 kcal/mol
Final Total Internal Energy	-1.30 kcal/mol
Torsional Free Energy	+1.79 kcal/mol
Unbound System's Energy [= (2)]	-1.30 kcal/mol

Table 4.3 listed the lowest energy docked conformation. The lowest binding free energy was estimated as -3.43 kcal/mol. The negative (-) sign of estimated free binding energy indicates that the binding of MAH monomer to insulin is favorable.

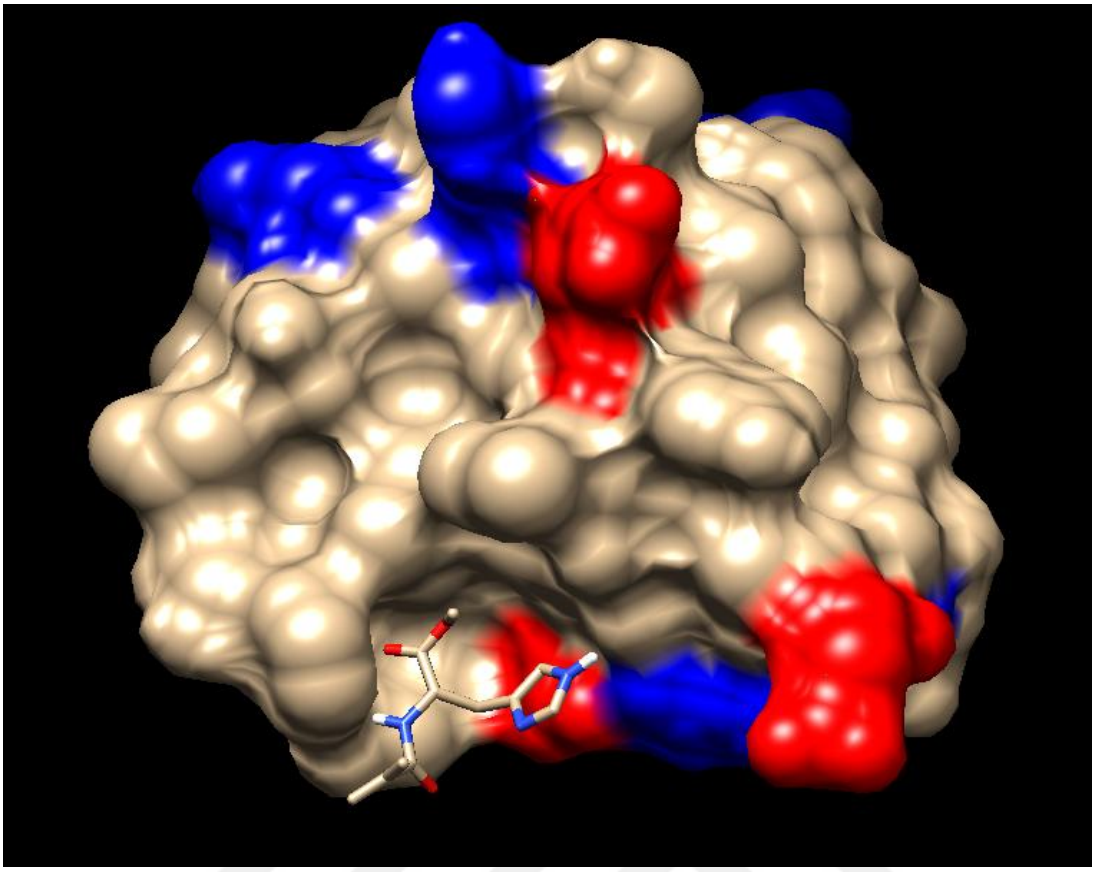

Figure 4.1. Visualization of docking one molecule of MAH at Insulin by using Chimera software. The electropositive regions are in blue, the electronegative regions are in red.

In Figure 4.1, one molecule of MAH docking at Insulin was visualized by using Chimera software. The electropositive and electronegative regions of insulin are colored in blue and red, respectively.

It was determined that MAH monomer well-fitted in the binding site of insulin and make two specific H-bonding interactions with insulin amino acid moieties. One H-bonding interaction mainly occur between the imidazole group of MAH monomer and the carboxyl backbone of neighboring valine (VAL-18) amino acid of insulin. The other H-bonding interaction mainly occur between the amide group of MAH monomer and hydroxyl groups of neighboring tyrosine (TRY-14) residues of insulin (Figure 4.2).

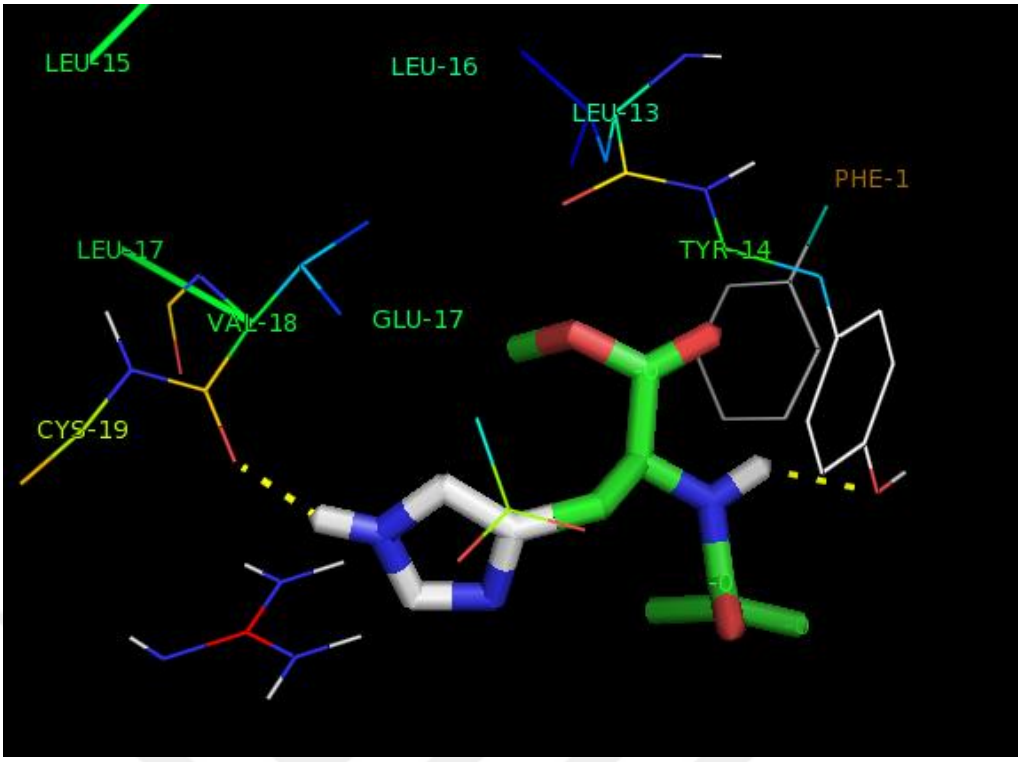

Figure 4.2. Visualization of docking one molecule of MAH at Insulin by using Pymol software.

4.2. Docking of One molecule of N-Methacryloyl-L-histidine Methyl Ester (MAH) Bound to Zn^{2+} Ion at Human Insulin Molecule

4.2.1. Cluster Analysis of Conformations

Number of conformations = 10

RMSD cluster analysis will be performed using the ligand atoms only (20 / 20 total atoms).

Outputting structurally similar clusters, ranked in order of increasing energy.

4.2.1.1. Clustering Histogram

Table 4.4. Clustering histogram of docking one molecule of MAH-Zn^{2+} at human insulin molecule.

Cluster Rank	Lowest Binding Energy(kcal/mol)	Run	Mean Binding Energy(kcal/mol)	Number in Cluster
1	-4.14	7	-4.14	1

2	-3.84	6	-3.84	1
3	-3.70	4	-3.50	3
4	-3.59	10	-3.59	1
5	-3.49	1	-3.49	1
6	-3.37	3	-3.37	1
7	-3.32	5	-3.32	1
8	-3.18	2	-3.18	1

Number of multi-member conformational clusters found = 1, out of 10 runs**.**

4.2.1.2. RMSD Values

Table 4.5. RMSD Values and Binding Energies of Runs.

Rank	Sub-Rank	Run	Binding Energy(kcal/mol)	Cluster RMSD	Reference RMSD
1	1	7	-4.14	0.00	10.68
2	1	6	-3.84	0.00	15.06
3	1	4	-3.70	0.00	16.84
3	2	9	-3.50	1.31	16.77
3	3	8	-3.30	1.59	16.94
4	1	10	-3.59	0.00	16.39
5	1	1	-3.49	0.00	15.74
6	1	3	-3.37	0.00	12.19
7	1	5	-3.32	0.00	17.06
8	1	2	-3.18	0.00	11.87

In Table 4.5, the similar RMSD values were obtained for 10 runs, in which they were grouped into families of similar clusters. The lowest binding energy of one molecule of MAH-Zn^{2+}-Insulin docking (-4.14 kcal/mol) was obtained for the 7th rank.

4.2.2. Information Entropy Analysis for This Clustering

Information entropy for this clustering = 0.86 (rmstol = 2.00 Angstrom)

4.2.3. Statistical Mechanical Analysis

Partition function:	Q = 10.06	at Temperature,	T = 298.15 K
Free energy:	A ~ -1367.78 kcal/mol	at Temperature,	T = 298.15 K
Internal energy:	U = -3.54 kcal/mol	at Temperature,	T = 298.15 K
Entropy:	S = 4.58 kcal/mol/K	at Temperature,	T = 298.15 K

4.2.4. Lowest Energy Docked Conformation

Table 4.6. Lowest Energy Docked Conformation

Run	7
Cluster Rank	1
Number of conformations in this cluster	1
RMSD from reference structure	10.679 A
Estimated Free Energy of Binding	**- 4.14 kcal/mol**
Estimated Inhibition Constant, Ki	925.09 uM (micro molar) [Temperature = 298.15 K]
Final Intermolecular Energy	-5.93 kcal/mol
vdW + Hbond + desolv Energy	-5.76 kcal/mol
Electrostatic Energy	-0.17 kcal/mol
Final Total Internal Energy	-1.23 kcal/mol
Torsional Free Energy	+1.79 kcal/mol
Unbound System's Energy [= (2)]	-1.23 kcal/mol

Table 4.6 listed the lowest energy docked conformation. The lowest binding free energy was estimated as -4.14 kcal/mol. The negative (-) sign of estimated free binding energy indicates that the docking of one molecule of MAH monomer chelated with Zn^{2+} ions at insulin is favorable.

In Figure 4.3, one molecule of MAH bound to Zn^{2+} ion docking at Insulin was visualized by using Chimera software. The electropositive and electronegative regions of insulin are colored in blue and red, respectively.

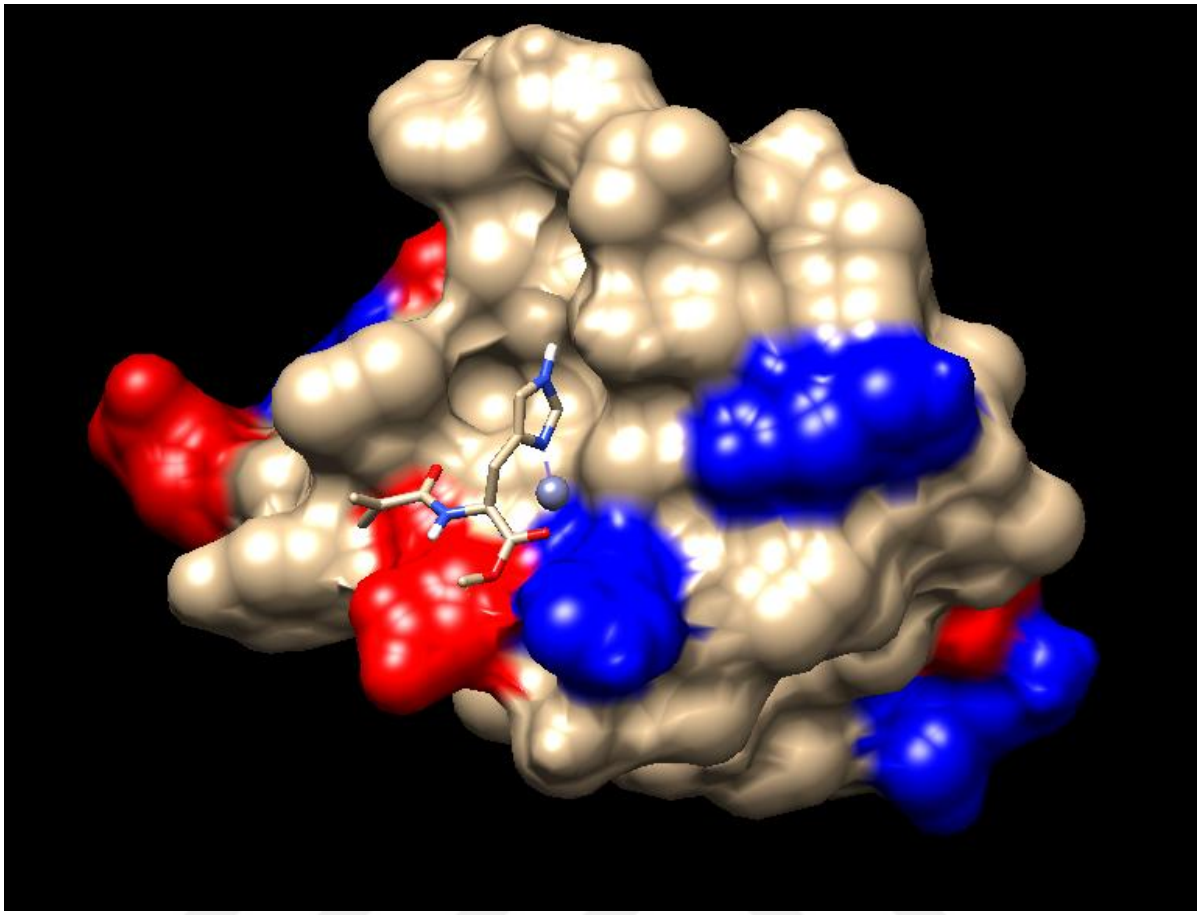

Figure 4.3. Visualization of docking one molecule of MAH bound to Zn^{2+} ion at Insulin by using Chimera software. The electropositive regions are in blue, the electronegative regions are in red.

It was determined that one molecule of MAH monomer bound to Zn^{2+} ion well-fitted in the binding site of insulin and make four specific (three H-bonding, one metal coordinating) interactions with insulin amino acid moieties. One H-bonding interaction mainly occur between the imidazole group of MAH monomer and the carboxyl backbone of neighboring glutamine (GLN-4) amino acid of insulin. The other two H-bonding interaction mainly occur between the methyl carboxyl group of MAH monomer and amide groups of neighboring histidine (HIS-10) residues of insulin. The metal coordinating interaction through Zn^{2+} ion occur between the imidazole group of MAH monomer and imidazole group of neighboring histidine (HIS-10) residues of insulin (Figure 4.4).

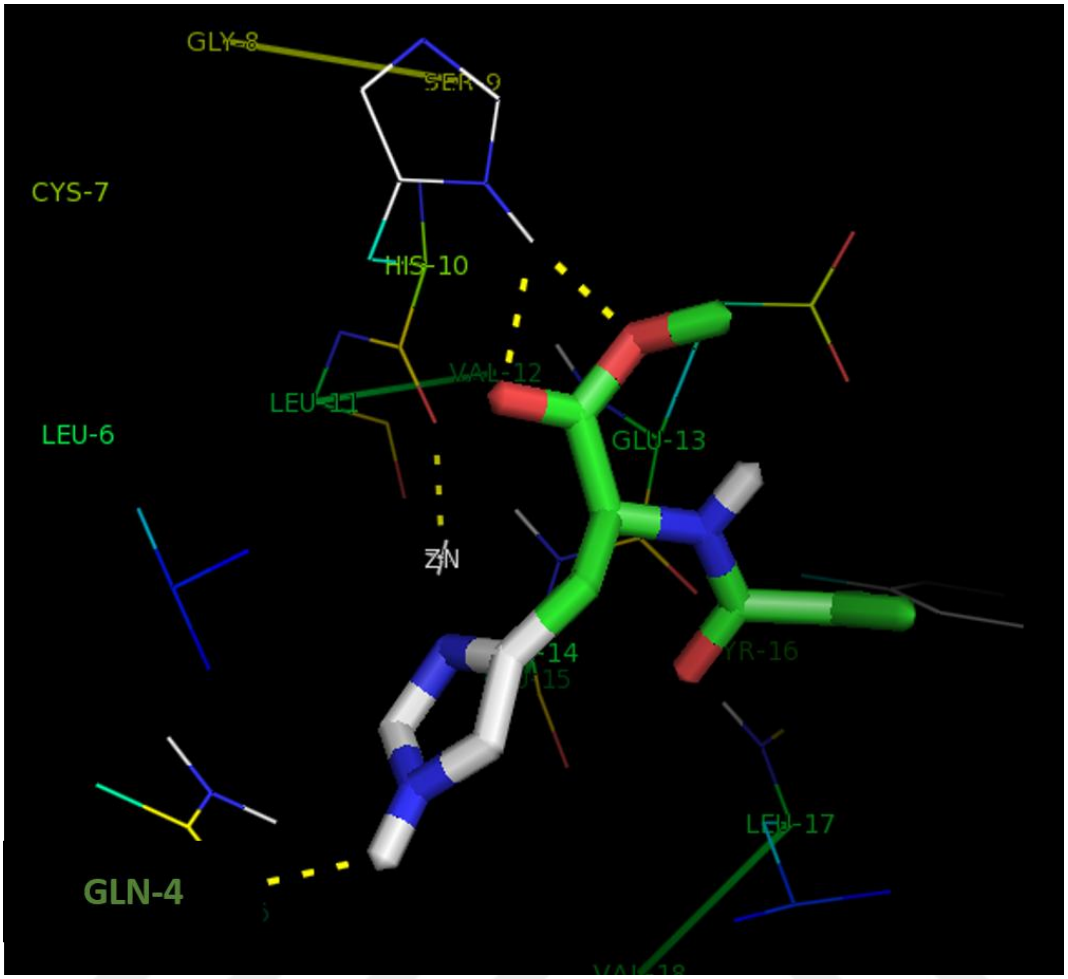

Figure 4.4. Visualization of docking of one molecule of MAH bound to Zn^{2+} ion at human Insulin molecule using Pymol software.

4.3. Docking of Two molecules of N-Methacryloyl-L-histidine Methyl Ester (MAH) Bound to Zn^{2+} Ion at Human Insulin Molecule

4.3.1. Cluster Analysis of Conformations

Number of conformations = 10

RMSD cluster analysis will be performed using the ligand atoms only (39 / 39 total atoms).

Outputting structurally similar clusters, ranked in order of increasing energy.

4.3.1.1. Clustering Histogram

Table 4.7. Clustering Histogram of docking two molecules of MAH-Zn^{2+} at human insulin molecule.

Cluster Rank	Lowest Binding Energy(kcal/mol)	Run	Mean Binding Energy(kcal/mol)	Number in Cluster
1	-2.84	5	-2.84	1
2	-2.44	9	-2.44	1
3	-1.61	7	-1.61	1
4	-1.55	1	-1.55	1
5	-1.41	4	-1.41	1
6	-1.37	2	-1.37	1
7	-1.16	10	-1.16	1
8	-0.84	8	-0.84	1
9	-0.55	6	-0.55	1
10	-0.52	3	-0.52	1

4.3.1.2. RMSD Values

Table 4.8. RMSD Values and Binding Energies of Runs.

Rank	Sub-Rank	Run	Binding Energy(kcal/mol)	Cluster RMSD	Reference RMSD
1	1	5	-2.84	0.00	12.09
2	1	9	-2.44	0.00	9.19
3	1	7	-1.61	0.00	18.75
4	1	1	-1.55	0.00	9.09
5	1	4	-1.41	0.00	15.09
6	1	2	-1.37	0.00	12.33
7	1	10	-1.16	0.00	8.20
8	1	8	-0.84	0.00	7.33
9	1	6	-0.55	0.00	8.85
10	1	3	-0.52	0.00	18.39

In Table 4.8, the similar rmsd values were obtained for 10 runs, in which they were grouped into families of similar clusters. The lowest binding energy of two molecules of MAH-Zn^{2+}-Insulin docking (-2.84 kcal/mol) was obtained for the 5th rank.

4.3.2. Information Entropy Analysis for This Clustering

Information entropy for this clustering = 1.00 (rmstol = 2.00 Angstrom)

4.3.3. Statistical Mechanical Analysis

Partition function: Q = 10.02 at Temperature, T = 298.15 K

Free energy: A ~ -1365.67 kcal/mol at Temperature, T = 298.15 K

Internal energy: U = -1.43 kcal/mol at Temperature, T = 298.15 K

Entropy: S = 4.58 kcal/mol/K at Temperature, T = 298.15 K

4.3.4. Lowest Energy Docked Conformation

Table 4.9. Lowest Energy Docked Conformation

Run	5
Cluster Rank	1
Number of conformations in this cluster	1
RMSD from reference structure	12.087 A
Estimated Free Energy of Binding	**-2.84 kcal/mol**
Estimated Inhibition Constant, Ki	8.23 uM (micro molar) [Temperature = 298.15 K]
Final Intermolecular Energy	-7.02 kcal/mol
vdW + Hbond + desolv Energy	-6.89 kcal/mol
Electrostatic Energy	-0.13 kcal/mol
Final Total Internal Energy	-3.33 kcal/mol
Torsional Free Energy	+4.18 kcal/mol
Unbound System's Energy [= (2)]	-3.33 kcal/mol

Table 4.9 listed the lowest energy docked conformation. The lowest binding free energy was estimated as -2.84 kcal/mol. The negative (-) sign of estimated free binding energy indicates that the docking of two molecules of MAH monomer chelated with Zn^{2+} ions at insulin is favorable.

In Figure 4.5, two molecules of MAH bound to Zn^{2+} ion docking at Insulin was visualized by using Chimera software. The electropositive and electronegative regions of insulin are colored in blue and red, respectively.

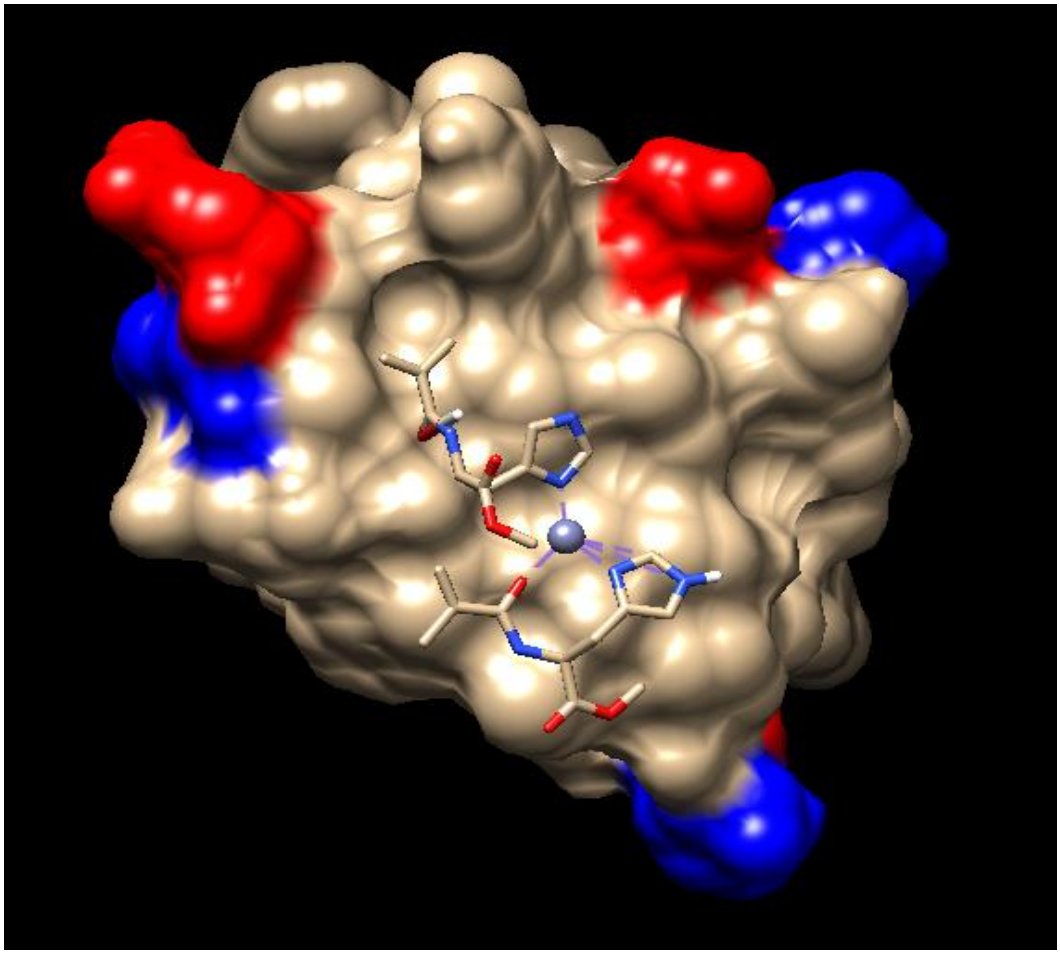

Figure 4.5. Visualization of docking two molecules of MAH bound to Zn^{2+} ion at Insulin by using Chimera software. The electropositive regions are in blue, the electronegative regions are in red.

It was determined that two molecules of MAH monomer bound to Zn^{2+} ion well-fitted in the binding site of insulin and make specific H-bonding interactions with insulin amino acid moieties. The H-bonding interaction mainly occur between the imidazole group of MAH monomer bound to Zn^{2+} ion and the carboxyl backbone and benzyl groups of neighboring threonine (THR-27) and phenyl alanine (PHE-25) amino acid residues of insulin (Figure 4.6).

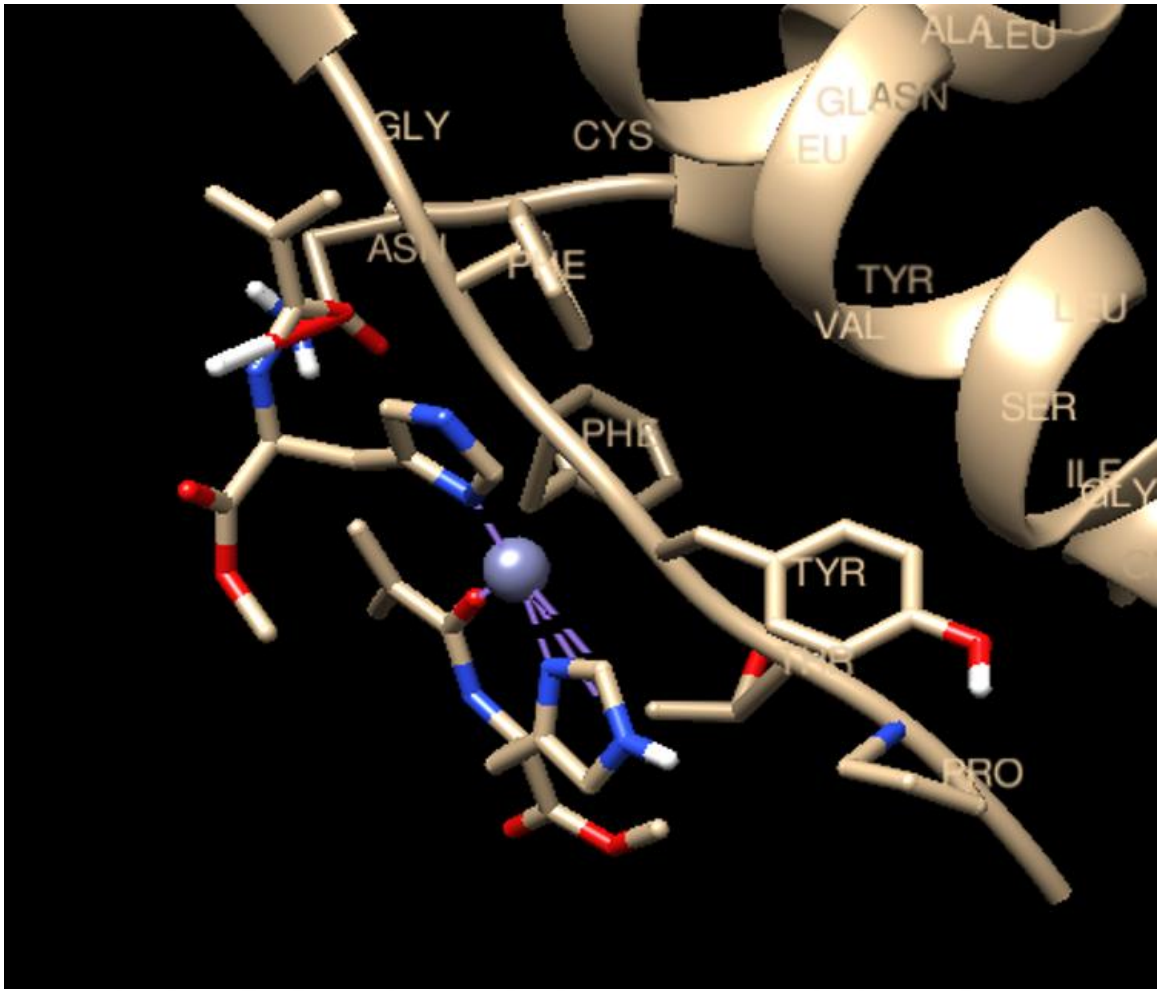

Figure 4.6. Docking of two molecules of MAH bound to zn^{2+} ion at human insulin molecule.

4.4. Docking of Three molecules of N-Methacryloyl-L-histidine Methyl Ester (MAH) Bound to Zn2+ Ion at Human Insulin Molecule

4.4.1. Cluster Analysis of Conformations

Number of conformations = 10

RMSD cluster analysis will be performed using the ligand atoms only (58 / 58 total atoms).

Outputting structurally similar clusters, ranked in order of increasing energy.

4.4.1.1. Clustering Histogram

Table 4.10. Clustering histogram docking of three molecules of (MAH) bound to zn^{2+} ion at human insulin molecule

Cluster Rank	Lowest Binding Energy(kcal/mol)	Run	Mean Binding Energy(kcal/mol)	Number in Cluster
1	-0.92	6	-0.92	1
2	-0.57	7	-0.57	1
3	-0.55	2	-0.55	1
4	+0.54	5	+0.54	1
5	+0.85	8	+0.85	1
6	+0.97	4	+0.97	1
7	+1.71	9	+1.71	1
8	+1.73	3	+1.73	1
9	+1.89	10	+1.89	1
10	+2.20	1	+2.20	1

4.4.1.2. RMSD Values

Table 4.11. RMSD Values and Binding Energies of Runs.

Rank	Sub-Rank	Run	Binding Energy(kcal/mol)	Cluster RMSD	Reference RMSD
1	1	6	-0.92	0.00	13.32
2	1	7	-0.57	0.00	10.96
3	1	2	-0.55	0.00	11.87
4	1	5	+0.54	0.00	12.18
5	1	8	+0.85	0.00	7.20
6	1	4	+0.97	0.00	19.35
7	1	9	+1.71	0.00	10.10
8	1	3	+1.73	0.00	11.40
9	1	10	+1.89	0.00	25.34
10	1	1	+2.20	0.00	21.16

In Table 4.11, the similar rmsd values were obtained for 10 runs, in which they were grouped into families of similar clusters. The lowest binding energy of three molecules of MAH-Zn^{2+}-Insulin docking (-0.92 kcal/mol) was obtained for the 6th rank.

4.4.2. Information Entropy Analysis for This Clustering

Information entropy for this clustering = 1.00 (rmstol = 2.00 Angstrom)

4.4.3. Statistical Mechanical Analysis

Partition function: Q = 9.99 at Temperature, T = 298.15 K

Free energy: A ~ -1363.45 kcal/mol at Temperature, T = 298.15 K

Internal energy: U = 0.79 kcal/mol at Temperature, T = 298.15 K

Entropy: S = 4.58 kcal/mol/K at Temperature, T = 298.15 K

4.4.4. Lowest Energy Docked Conformation

Table 4.12. Lowest Energy Docked Conformation

Run	6
Cluster Rank	1
Number of conformations in this cluster	1
RMSD from reference structure	13.316 A
Estimated Free Energy of Binding	**-0.92 kcal/mol**
Estimated Inhibition Constant, Ki	210.98 mM (millimolar) [Temperature = 298.15 K]
Final Intermolecular Energy	-7.19 kcal/mol
vdW + Hbond + desolv Energy	-6.83 kcal/mol
Electrostatic Energy	-0.35 kcal/mol
Final Total Internal Energy	-6.80 kcal/mol
Torsional Free Energy	+6.26 kcal/mol
Unbound System's Energy [= (2)]	-6.80 kcal/mol

Table 4.12 listed the lowest energy docked conformation. The lowest binding free energy was estimated as -0.92 kcal/mol. The negative (-) sign of estimated free binding energy indicates that the docking of three molecules of MAH monomer chelated with Zn^{2+} ions at insulin is favorable.

In Figure 4.7, three molecules of MAH bound to Zn^{2+} ion docking at Insulin was visualized by using Chimera software. The electropositive and electronegative regions of insulin are colored in blue and red, respectively.

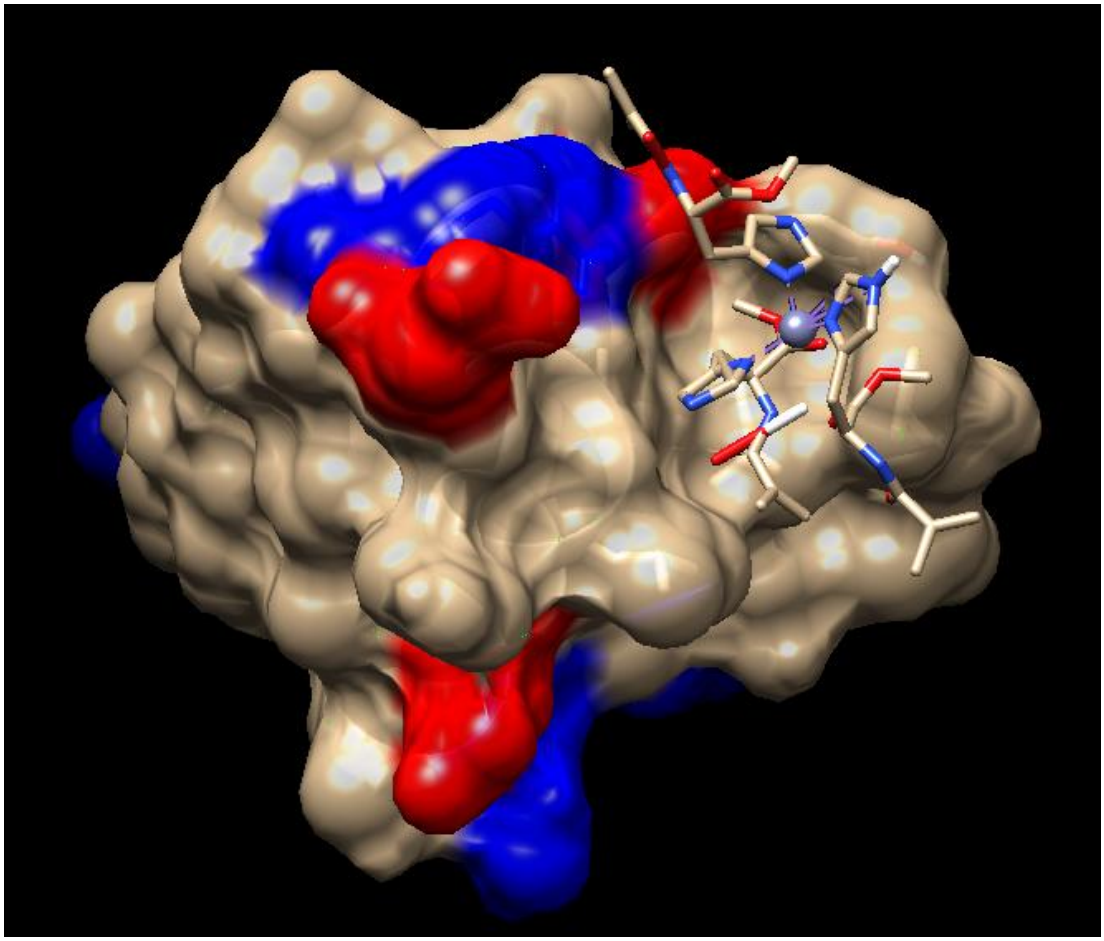

Figure 4.7. Visualization of docking three molecules of MAH bound to Zn^{2+} Ion at Insulin by using Chimera software. The electropositive regions are in blue, the electronegative regions are in red.

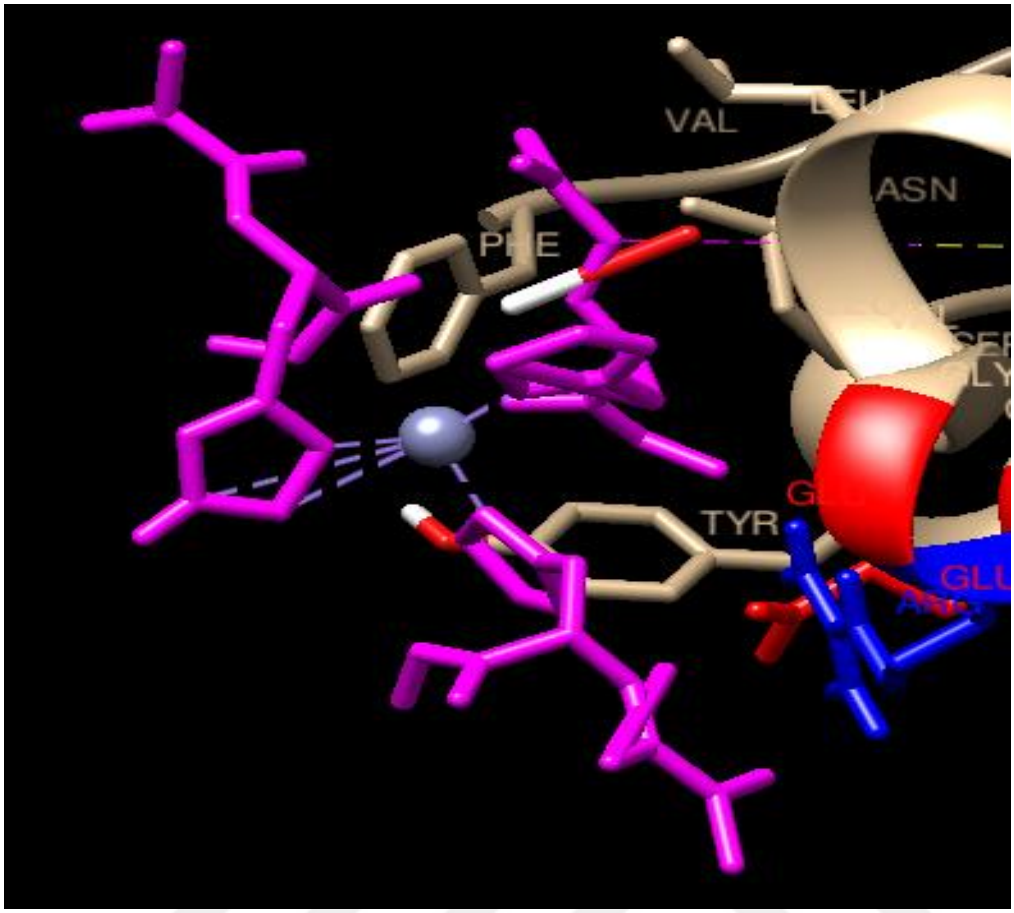

Figure 4.8. Docking of three molecules of MAH bound to zn^{2+} ion at human insulin molecule.

It was determined that two molecules of MAH monomer bound to Zn^{2+} ion well-fitted in the binding site of insulin and make specific hydrophobic and metal chelating interaction with insulin amino acid moieties. The metal chelating interaction mainly occur between the imidazole groups of MAH monomer bound to Zn^{2+} ion and the phenyl groups of neighboring tyrosine (TYR-14) amino acid residues of insulin. The hydrophobic interactions occurred between the methyl groups of MAH monomer and benzyl group of neighboring phenyl alanine (PHE-1) (Figure 4.8).

4.9. Comparison between the lowest binding energy of MAH-Insulin dockings.

AutoDock generates a set of different ligand binding poses and use a scoring function to estimate binding affinities for the generated ligand poses in order to

determine the best binding mode. Low binding energy has better docking score. Obviously, when Zn^{2+} ion is bound to one molecule of N-Methacryloyl-L-histidine Methyl Ester (MAH), binding energy declines and becomes the lowest binding energy (-4.14 kcal/mol) of all dockings obtained in this study. The lower the free binding energy, the higher the binding affinity of MAH-Zn^{2+} towards insulin. It was concluded that one molecule of MAH monomer bound to Zn^{2+} is the most favorable metal chelated monomer docked at insulin (Diagram 4.1).

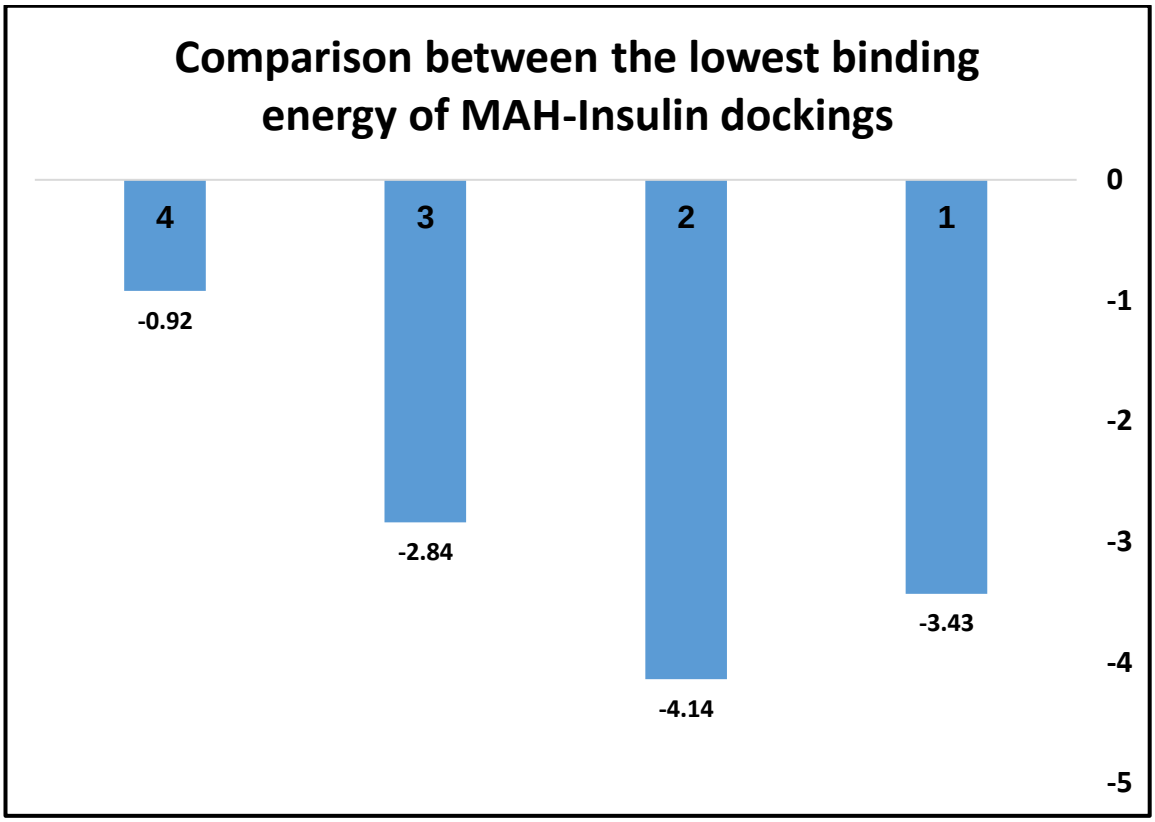

Diagram 4.1. Column1 (MAH-Insulin). Column2 (MAH-Zn^{2+}-Insulin)**.** Column3 (2MAH-Zn^{2+}-Insulin). Column4 (3MAH-Zn^{2+}-Insulin).

4.5. Docking of One molecule of N-Methacryloyl-L-histidine Methyl Ester (MAH) at Horse Heart Cytochrome C

4.5.1. Cluster Analysis of Conformations

Number of conformations = 10

RMSD cluster analysis will be performed using the ligand atoms only (19 / 19 total atoms).

Outputting structurally similar clusters, ranked in order of increasing energy.

4.5.1.1. Clustering Histogram

Table 4.13. Clustering histogram of docking of one molecule of (MAH) at horse heart Cytochrome C

Cluster Rank	Lowest Binding Energy(kcal/mol)	Run	Mean Binding Energy(kcal/mol)	Number in Cluster
1	-3.91	4	-3.63	2
2	-3.84	10	-3.84	1
3	-3.64	1	-3.64	1
4	-3.62	8	-3.22	2
5	-3.57	5	-3.57	1
6	-3.34	2	-3.34	1
7	-3.22	6	-3.22	1
8	-2.66	9	-2.66	1

Number of multi-member conformational clusters found = 2, out of 10 runs.

4.5.1.2. RMSD Values

Table 4.14. RMSD Values and Binding Energies of Runs.

Rank	Sub-Rank	Run	Binding Energy(kcal/mol)	Cluster RMSD	Reference RMSD
1	1	4	-3.91	0.00	48.48
1	2	3	-3.34	-3.22	49.10
2	1	10	-3.84	0.00	43.01
3	1	1	-3.64	0.00	50.41
4	1	8	-3.62	0.00	31.66
4	2	7	-2.82	1.53	31.26
5	1	5	-3.57	0.00	52.42
6	1	2	-3.34	0.00	58.03
7	1	6	-3.22	0.00	60.31
8	1	9	-2.66	0.00	60.68

In Table 4.14, the similar rmsd values were obtained for 10 runs, in which they were grouped into families of similar clusters. The lowest binding energy of MAH-Insulin docking (-3.91 kcal/mol) was obtained for the 4th rank.

4.5.2. Information Entropy Analysis for This Clustering

Information entropy for this clustering = 0.88 (rmstol = 2.00 Angstrom)

4.5.3. Statistical Mechanical Analysis

Partition function:	Q = 10.06	at Temperature,	T = 298.15 K
Free energy:	A ~ -1367.64 kcal/mol	at Temperature,	T = 298.15 K
Internal energy:	U = -3.40 kcal/mol	at Temperature,	T = 298.15 K
Entropy:	S = 4.58 kcal/mol/K	at Temperature,	T = 298.15 K

4.5.4. Lowest Energy Docked Conformation

Table 4.15. Lowest Energy Docked Conformation

Run	4
Cluster Rank	1
Number of conformations in this cluster	2
RMSD from reference structure	48.481 A
Estimated Free Energy of Binding	**-3.91 kcal/mol**
Estimated Inhibition Constant, Ki	1.35 mM (milli molar) [Temperature = 298.15 K]
Final Intermolecular Energy	-5.70 kcal/mol
vdW + Hbond + desolv Energy	-5.60 kcal/mol
Electrostatic Energy	-0.11 kcal/mol
Final Total Internal Energy	-1.11 kcal/mol
Torsional Free Energy	+1.79 kcal/mol
Unbound System's Energy	-1.11 kcal/mol

Table 4.15 listed the lowest energy docked conformation. The lowest binding free energy was estimated as -3.91 kcal/mol. The negative (-) sign of estimated free binding energy indicates that the binding of one molecule of MAH monomer to Cytochrome C is favorable.

In Figure 4.9, one molecule of MAH docking at Cytochrome C was visualized by using Chimera software. The electropositive and electronegative regions of Cytochrome C are colored in blue and red, respectively.

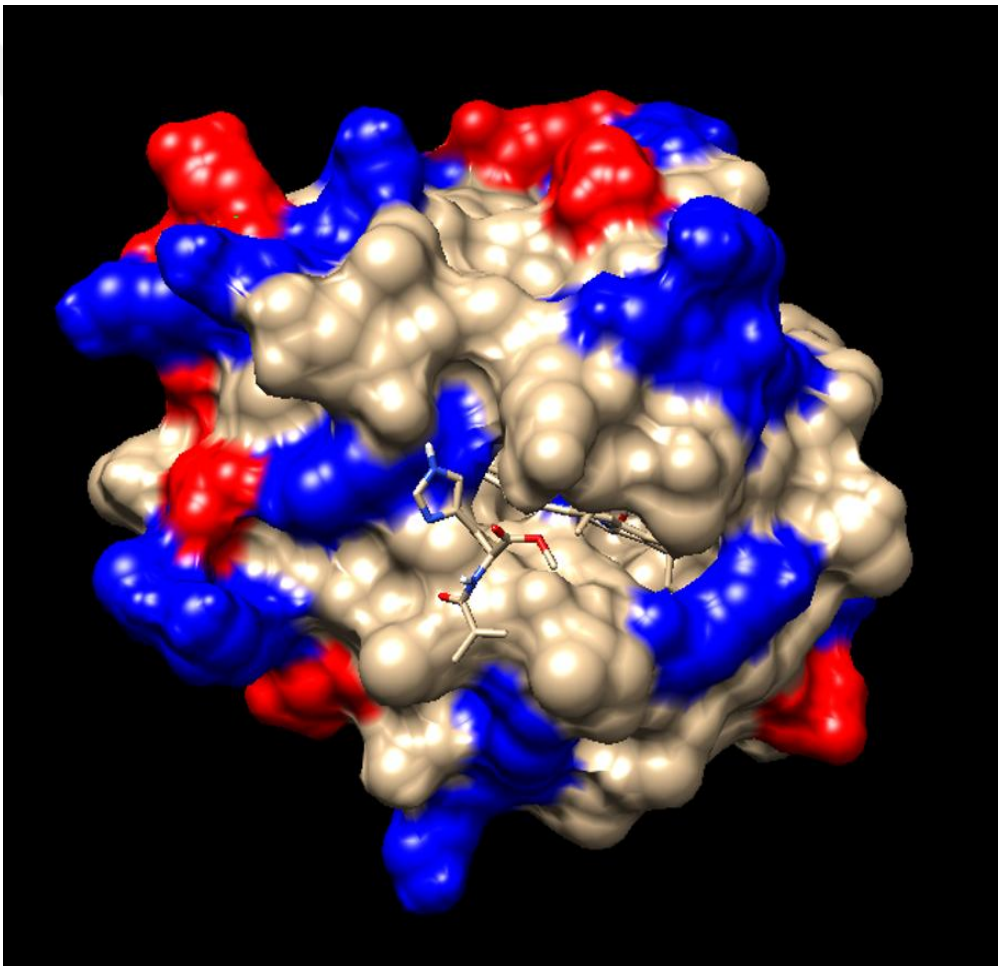

Figure 4.9. Visualization of docking one molecule of MAH at Horse Heart Cytochrome C by using Chimera software. The electropositive regions are in blue, the electronegative regions are in red.

It was determined that MAH monomer well-fitted in the binding site of Cytochrome C and make three specific H-bonding interactions with

Cytochrome C amino acid moieties. One H-bonding interaction mainly occur between the imidazole group of MAH monomer and the carboxyl backbone of neighboring glutamine (GLN-12) amino acid of Cytochrome C. The other H-bonding interaction occur between the methyl carboxyl group of MAH monomer and amine groups of neighboring GLN-16 residues of Cytochrome C. The third H-bonding interaction occur between oxygen of MAH monomer and neighboring alanine (ALA-83) residues of Cytochrome C (Figure 4.10).

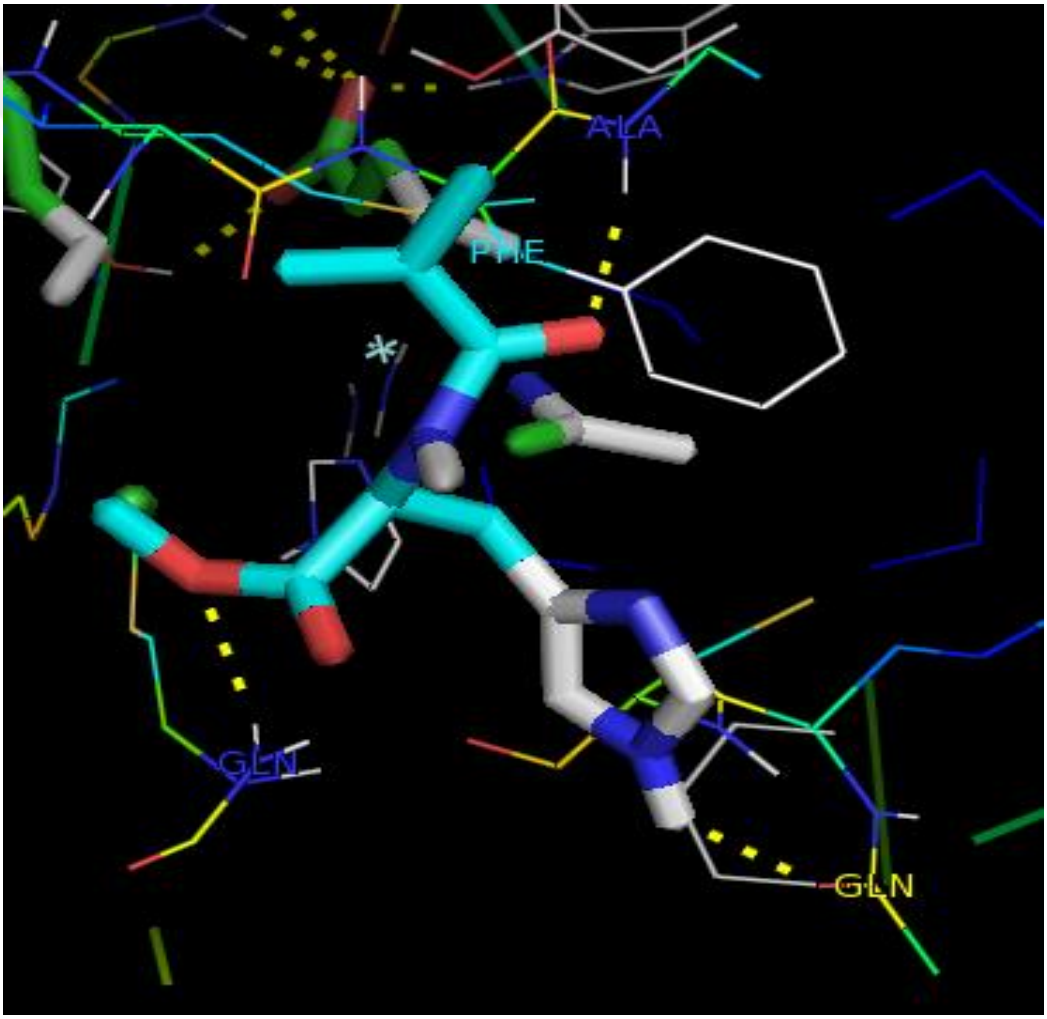

Figure 4.10. Docking of one molecule of MAH at Horse Heart Cytochrome C.

4.6. Docking of One molecule of N-Methacryloyl-L-histidine Methyl Ester (MAH) Bound to Zn^{2+} Ion at Horse Heart Cytochrome C

4.6.1. Cluster Analysis of Conformations

Number of conformations = 10

RMSD cluster analysis will be performed using the ligand atoms only (20 / 20 total atoms).

Outputting structurally similar clusters, ranked in order of increasing energy.

4.6.1.1. Clustering Histogram

Table 4.16. Clustering histogram docking of one molecule of n-methacryloyl-l-histidine methyl ester (MAH) bound to zn2+ ion at horse heart Cytochrome C

Cluster Rank	Lowest Binding Energy(kcal/mol)	Run	Mean Binding Energy(kcal/mol)	Number in Cluster
1	-4.92	1	-4.92	1
2	-4.72	2	-4.72	1
3	-4.54	8	-4.54	1
4	-4.27	9	-4.27	1
5	-4.15	3	-4.15	1
6	-3.76	6	-3.69	2
7	-3.72	10	-3.72	1
8	-3.63	5	-3.63	1
9	-3.54	4	-3.54	1

Number of multi-member conformational clusters found = 1, out of 10 runs.

4.6.1.2. RMSD Values

Table 4.17. RMSD Values and Binding Energies of Runs.

Rank	Sub-Rank	Run	Binding Energy(kcal/mol)	Cluster RMSD	Reference RMSD
1	1	1	-4.92	0.00	53.32
2	1	2	-4.72	0.00	43.30

3	1	8	-4.54	0.00	50.20
4	1	9	-4.27	0.00	42.89
5	1	3	-4.15	0.00	52.37
6	1	6	-3.76	0.00	56.97
6	2	7	-3.62	1.97	57.47
7	1	10	-3.72	0.00	30.92
8	1	5	-3.63	0.00	48.47
9	1	4	-3.54	0.00	44.83

In Table 4.17, the similar rmsd values were obtained for 10 runs, in which they were grouped into families of similar clusters. The lowest binding energy of one molecule of MAH-Zn^{2+}-Cytochrome C docking (-4.92 kcal/mol) was obtained for the 1st rank.

4.6.2. Information Entropy Analysis for This Clustering

Information entropy for this clustering = 0.94 (rmstol = 2.00 Angstrom)

4.6.3. Statistical Mechanical Analysis

Partition function: Q = 10.07 at Temperature, T = 298.15 K

Free energy: A ~ -1368.32 kcal/mol at Temperature, T = 298.15 K

Internal energy: U = -4.09 kcal/mol at Temperature, T = 298.15 K

Entropy: S = 4.58 kcal/mol/K at Temperature, T = 298.15 K

4.6.4. Lowest Energy Docked Conformation

Table 4.18. Lowest Energy Docked Conformation

Run	1
Cluster Rank	1
Number of conformations in this cluster	1
RMSD from reference structure	53.325 A
Estimated Free Energy of Binding	**-4.92 kcal/mol**

Estimated Inhibition Constant, Ki	248.69 uM (micromolar) [Temperature = 298.15 K]
Final Intermolecular Energy	-6.71 kcal/mol
vdW + Hbond + desolv Energy	-6.26 kcal/mol
Electrostatic Energy	-0.45 kcal/mol
Final Total Internal Energy	-0.79 kcal/mol
Torsional Free Energy	+1.79 kcal/mol
Unbound System's Energy	-0.79 kcal/mol

Table 4.18 listed the lowest energy docked conformation. The lowest binding free energy was estimated as -4.92 kcal/mol. The negative (-) sign of estimated free binding energy indicates that the docking of one molecule of MAH monomer chelated with Zn^{2+} ions at Cytochrome C is favorable.

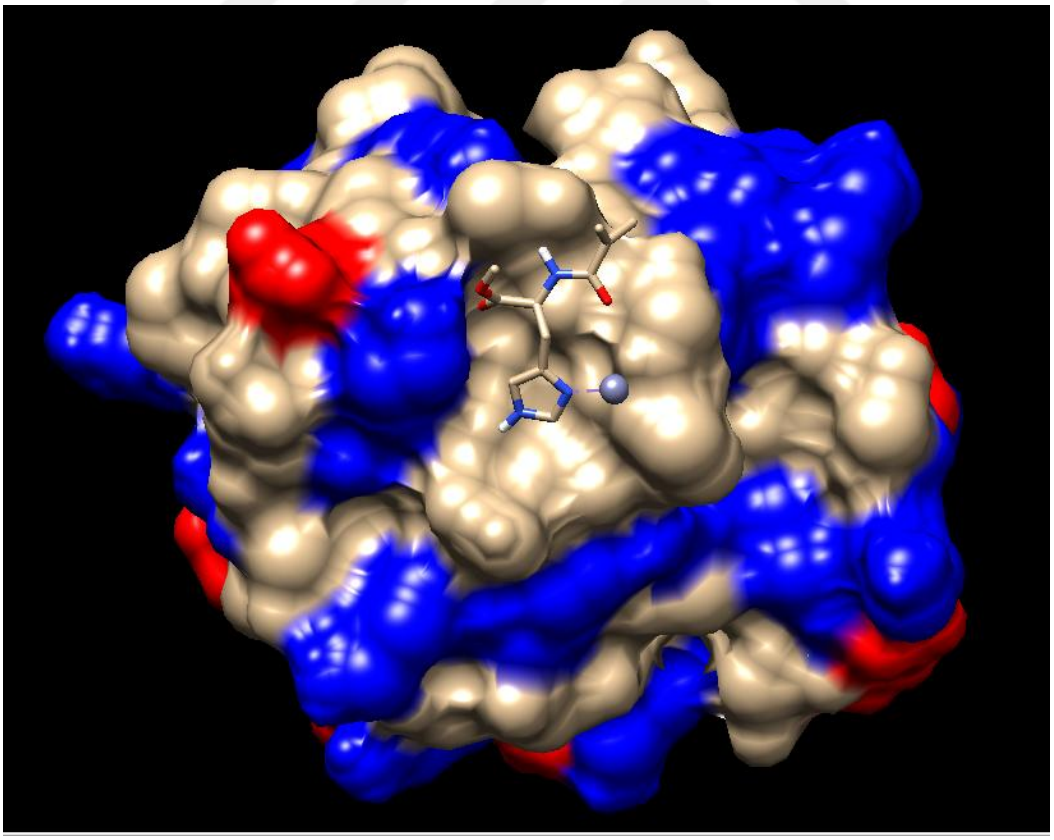

Figure 4.11. Visualization of docking one molecule of MAH bound to Zn^{2+} Ion at Horse Heart Cytochrome C by using Chimera software. The electropositive regions are in blue, the electronegative regions are in red.

In Figure 4.11, one molecule of MAH bound to Zn^{2+} ion docking at Cytochrome C was visualized by using Chimera software. The electropositive and electronegative regions of Cytochrome C are colored in blue and red, respectively.

It was determined that one molecule of MAH monomer bound to Zn^{2+} ion well-fitted in the binding site of Cytochrome C and make specific H-bonding and metal coordinating interactions with Cytochrome C amino acid moieties. The H-bonding interactions mainly occur between the imidazole, carboxyl and oxygen group of MAH monomer and the neighboring threonine (THR-47), tyrosine (TRY-48), lysine (LYS-53), glycine (GLY-41, GLY-45) amino acid of Cytochrome C, respectively. The metal coordinating interaction through Zn^{2+} ion occur between the imidazole group of MAH monomer and amide backbone of neighboring alanine (ALA-43) residues of Cytochrome C (Figure 4.12).

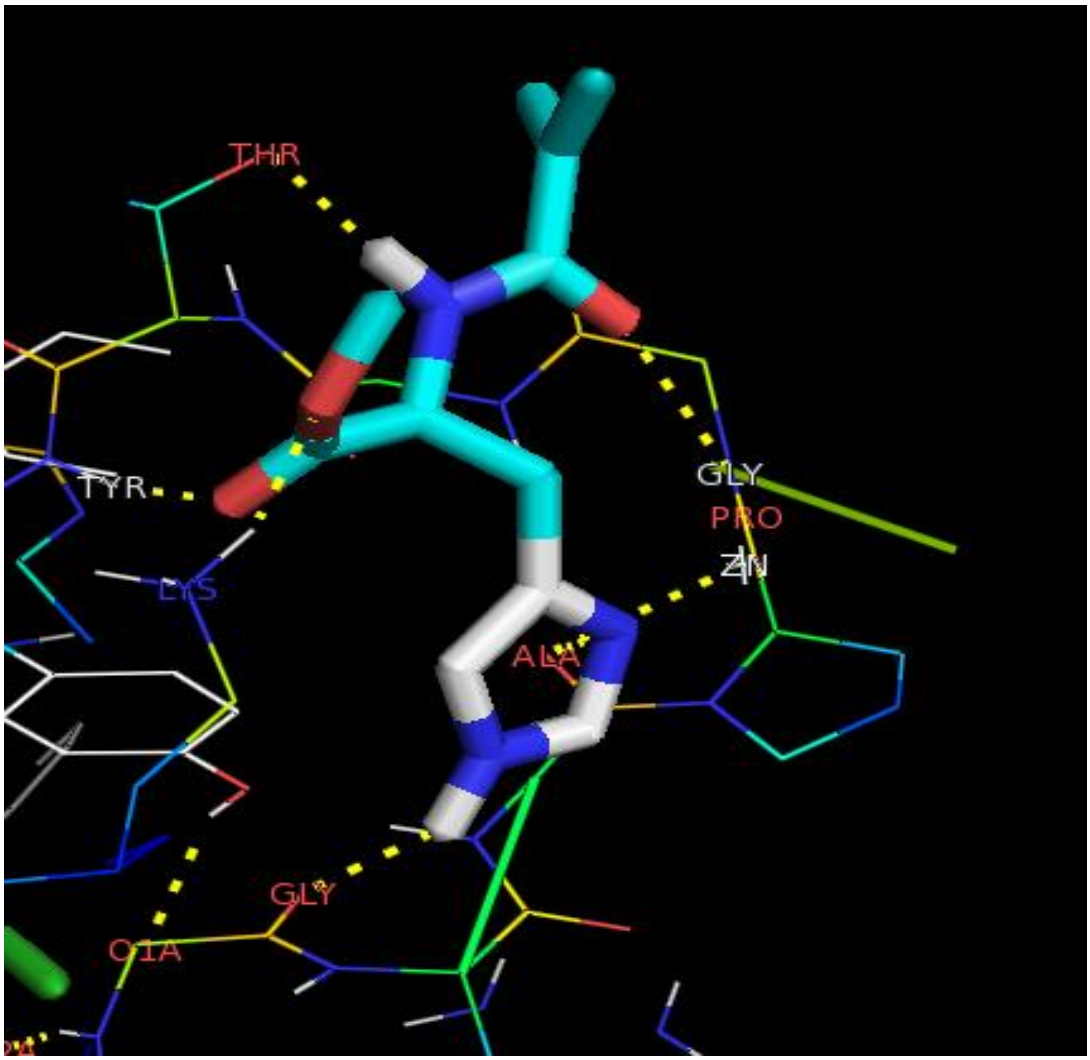

Figure 4.12. Docking of one molecule of MAH bound to zn^{2+} ion at horse heart Cytochrome C.

4.7. Docking of Two Molecules of N-Methacryloyl-L-histidine Methyl Ester (MAH) Bound to Zn^{2+} Ions at Horse Heart Cytochrome C

4.7.1. Cluster Analysis of Conformations

Number of conformations = 10

RMSD cluster analysis will be performed using the ligand atoms only (39 / 39 total atoms).

Outputting structurally similar clusters, ranked in order of increasing energy.

4.7.1.1. Clustering Histogram

Table 4.19. Clustering histogram docking of two molecules of (MAH) bound to zn^{2+} ions at horse heart Cytochrome C

Cluster Rank	Lowest Binding Energy(kcal/mol)	Run	Mean Binding Energy(kcal/mol)	Number in Cluster
1	-2.88	10	-2.88	1
2	-1.52	2	-1.52	1
3	-1.48	1	-1.48	1
4	-1.38	3	-1.38	1
5	-1.33	4	-1.33	1
6	-1.26	6	-1.26	1
7	-1.15	7	-1.15	1
8	-1.01	5	-1.01	1
9	-0.98	9	-0.98	1
10	-0.27	8	-0.27	1

4.7.1.2. RMSD Values

Table 4.20. RMSD Values and Binding Energies of Runs.

Rank	Sub-Rank	Run	Binding Energy(kcal/mol)	Cluster RMSD	Reference RMSD
1	1	10	-2.88	0.00	51.05
2	1	2	-1.52	0.00	60.26
3	1	1	-1.48	0.00	32.73
4	1	3	-1.38	0.00	56.03
5	1	4	-1.33	0.00	48.56
6	1	6	-1.26	0.00	60.15
7	1	7	-1.15	0.00	50.04
8	1	5	-1.01	0.00	48.42
9	1	9	-0.98	0.00	47.92
10	1	8	-0.27	0.00	60.65

In Table 4.20, the similar rmsd values were obtained for 10 runs, in which they were grouped into families of similar clusters. The lowest binding energy of two molecules of MAH-Zn^{2+}-Cytochrome C docking (-2.88 kcal/mol) was obtained for the 10th rank.

4.7.2. Information Entropy Analysis for This Clustering

Information entropy for this clustering = 1.00 (rmstol = 2.00 Angstrom)

4.7.3. Statistical Mechanical Analysis

Partition function:	Q = 10.02	at Temperature,	T = 298.15 K
Free energy:	A ~ -1365.56 kcal/mol	at Temperature,	T = 298.15 K
Internal energy:	U = -1.33 kcal/mol	at Temperature,	T = 298.15 K
Entropy:	S = 4.58 kcal/mol/K	at Temperature,	T = 298.15 K

4.7.4. Lowest Energy Docked Conformation

Table 4.21. Lowest Energy Docked Conformation

Run	10
Cluster Rank	1
Number of conformations in this cluster	1
RMSD from reference structure	51.054 A
Estimated Free Energy of Binding	**-2.88 kcal/mol**
Estimated Inhibition Constant, Ki	7.74 mM (millimolar) [Temperature = 298.15 K]
Final Intermolecular Energy	-7.06 kcal/mol
vdW + Hbond + desolv Energy	-6.91 kcal/mol
Electrostatic Energy	-0.14 kcal/mol
Final Total Internal Energy	-2.71 kcal/mol
Torsional Free Energy	+4.18 kcal/mol
Unbound System's Energy	-2.71 kcal/mol

Table 4.21 listed the lowest energy docked conformation. The lowest binding free energy was estimated as -2.88 kcal/mol. The negative (-) sign of estimated

free binding energy indicates that the docking of two molecules of MAH monomer chelated with Zn^{2+} ions at Cytochrome C is favorable.

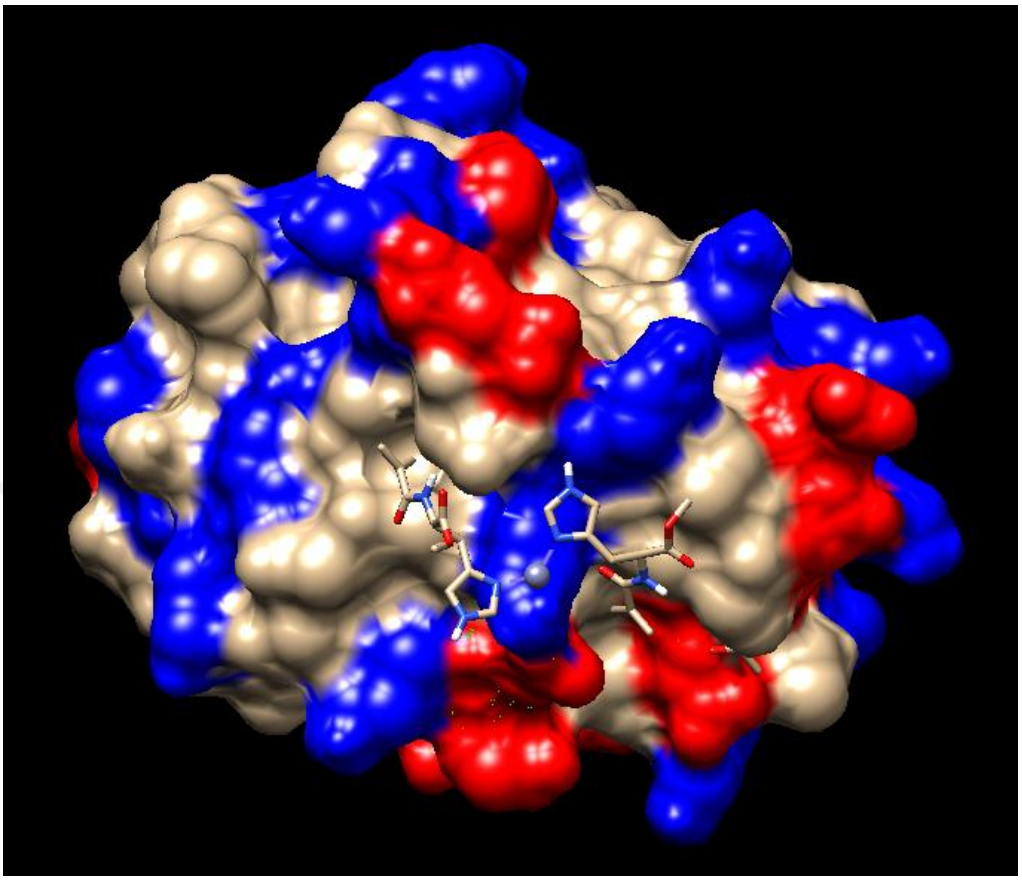

Figure 4.13. Visualization of docking two molecules of MAH bound to Zn^{2+} Ion at Horse Heart Cytochrome C by using Chimera software, the electropositive regions are in blue. The electronegative regions are in red.

In Figure 4.13, two molecules of MAH bound to Zn^{2+} ion docking at Cytochrome C was visualized by using Chimera software. The electropositive and electronegative regions of Cytochrome C are colored in blue and red, respectively.

It was determined that two molecules of MAH monomer bound to Zn^{2+} ion well-fitted in the binding site of Cytochrome C and make one specific H-bonding interaction with Cytochrome C amino acid moiety. The H-bonding interaction mainly occurs between the amide of MAH monomer and the neighboring $\sigma 1-$oxygen of asparagine (ASN-103) residue of Cytochrome C (Figure 4.14).

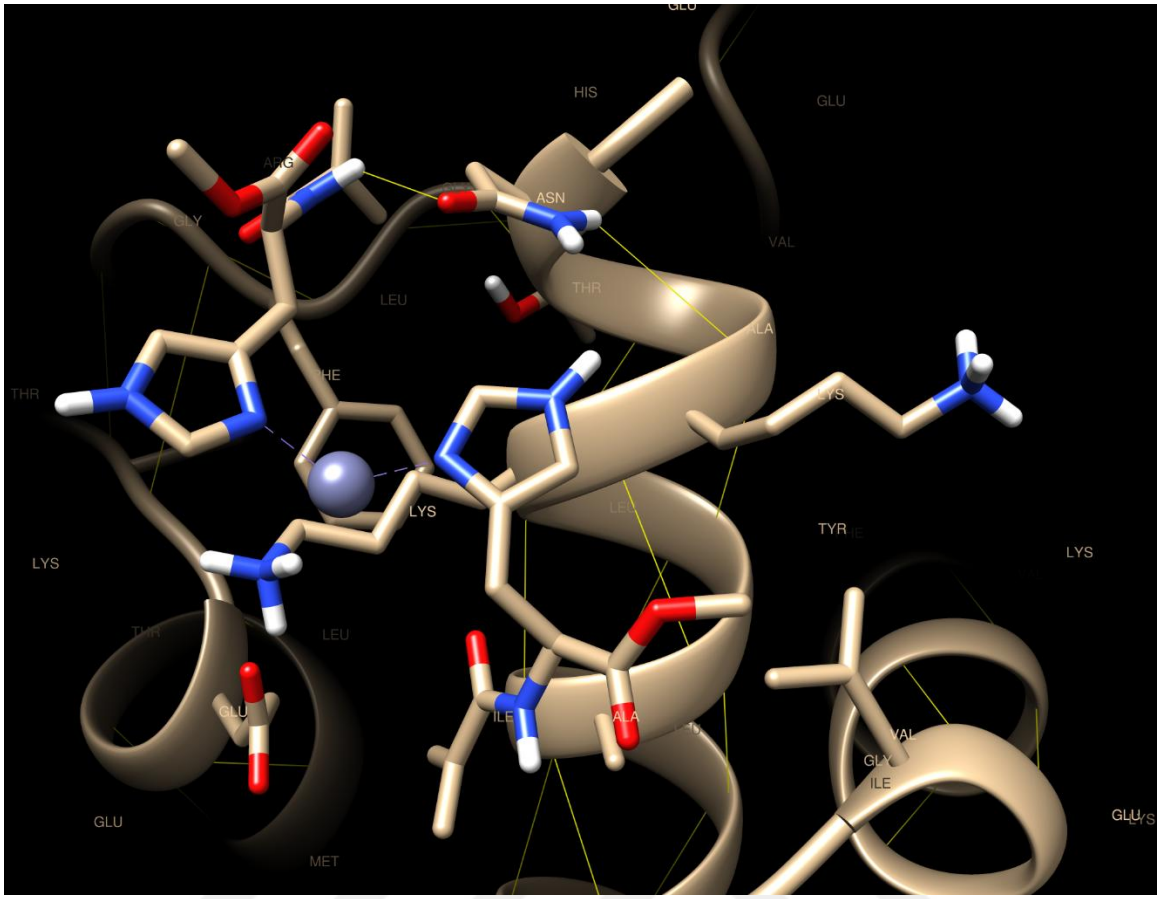

Figure 4.14. Docking of two molecules of mah bound to zn^{2+} ions with horse heart Cytochrome C (yellow lines show the H-bonding interactions).

4.8. Docking of Three Molecules of N-Methacryloyl-L-histidine Methyl Ester (MAH) Bound to Zn^{2+} Ions at Horse Heart Cytochrome C

4.8.1. Cluster Analysis of Conformations

Number of conformations = 10

RMSD cluster analysis will be performed using the ligand atoms only (58 / 58 total atoms).

Outputting structurally similar clusters, ranked in order of increasing energy.

4.8.1.1. Clustering Histogram

Table 4.21. Clustering histogram of docking of three molecules of (mah) bound to zn2+ ions at horse heart cytochrome C.

Cluster Rank	Lowest Binding Energy(kcal/mol)	Run	Mean Binding Energy(kcal/mol)	Number in Cluster
1	-0.37	1	-0.37	1
2	-0.15	7	-0.15	1
3	+0.34	10	+0.34	1
4	+0.49	5	+0.49	1
5	+0.51	6	+0.51	1
6	+0.56	3	+0.56	1
7	+0.83	8	+0.83	1
8	+0.86	2	+0.86	1
9	+1.11	4	+1.11	1
10	+2.44	9	+2.44	1

4.8.1.2. RMSD Values

Table 4.23. RMSD Values and Binding Energies of Runs.

Rank	Sub-Rank	Run	Binding Energy(kcal/mol)	Cluster RMSD	Reference RMSD
1	1	1	-0.37	0.00	32.77
2	1	7	-0.15	0.00	49.10
3	1	10	+0.34	0.00	52.83
4	1	5	+0.49	0.00	61.83
5	1	6	+0.51	0.00	53.07
6	1	3	+0.56	0.00	55.53
7	1	8	+0.83	0.00	55.28
8	1	2	+0.86	0.00	53.51
9	1	4	+1.11	0.00	36.38
10	1	9	+2.44	0.00	49.72

In Table 4.23, the similar rmsd values were obtained for 10 runs, in which they were grouped into families of similar clusters. The lowest binding energy of three molecules of MAH-Zn^{2+}-Cytochrome C docking (-0.37 kcal/mol) was obtained for the 1st rank.

4.8.2. Information Entropy Analysis for This Clustering

Information entropy for this clustering = 1.00 (rmstol = 2.00 Angstrom)

4.8.3. Statistical Mechanical Analysis

Partition function: Q = 9.99 at Temperature, T = 298.15 K

Free energy: A ~ -1363.58 kcal/mol at Temperature, T = 298.15 K

Internal energy: U = 0.66 kcal/mol at Temperature, T = 298.15 K

Entropy: S = 4.58 kcal/mol/K at Temperature, T = 298.15 K

4.8.4. Lowest Energy Docked Conformation

Table 4.24. Lowest Energy Docked Conformation

Run	1
Cluster Rank	1
Number of conformations in this cluster	1
RMSD from reference structure	**32.772 A**
Estimated Free Energy of Binding	**-0.37 kcal/mol**
Estimated Inhibition Constant, Ki	534.38 mM (millimolar) [Temperature = 298.15 K]
Final Intermolecular Energy	-6.64 kcal/mol
vdW + Hbond + desolv Energy	-6.26 kcal/mol
Electrostatic Energy	-0.37 kcal/mol
Final Total Internal Energy	-6.53 kcal/mol
Torsional Free Energy	+6.26 kcal/mol
Unbound System's Energy	-6.53kcal/mol

Table 4.24 listed the lowest energy docked conformation. The lowest binding free energy was estimated as -0.37 kcal/mol. The negative (-) sign of estimated free binding energy indicates that the docking of three molecules of MAH monomer chelated with Zn^{2+} ions at Cytochrome C is favorable.

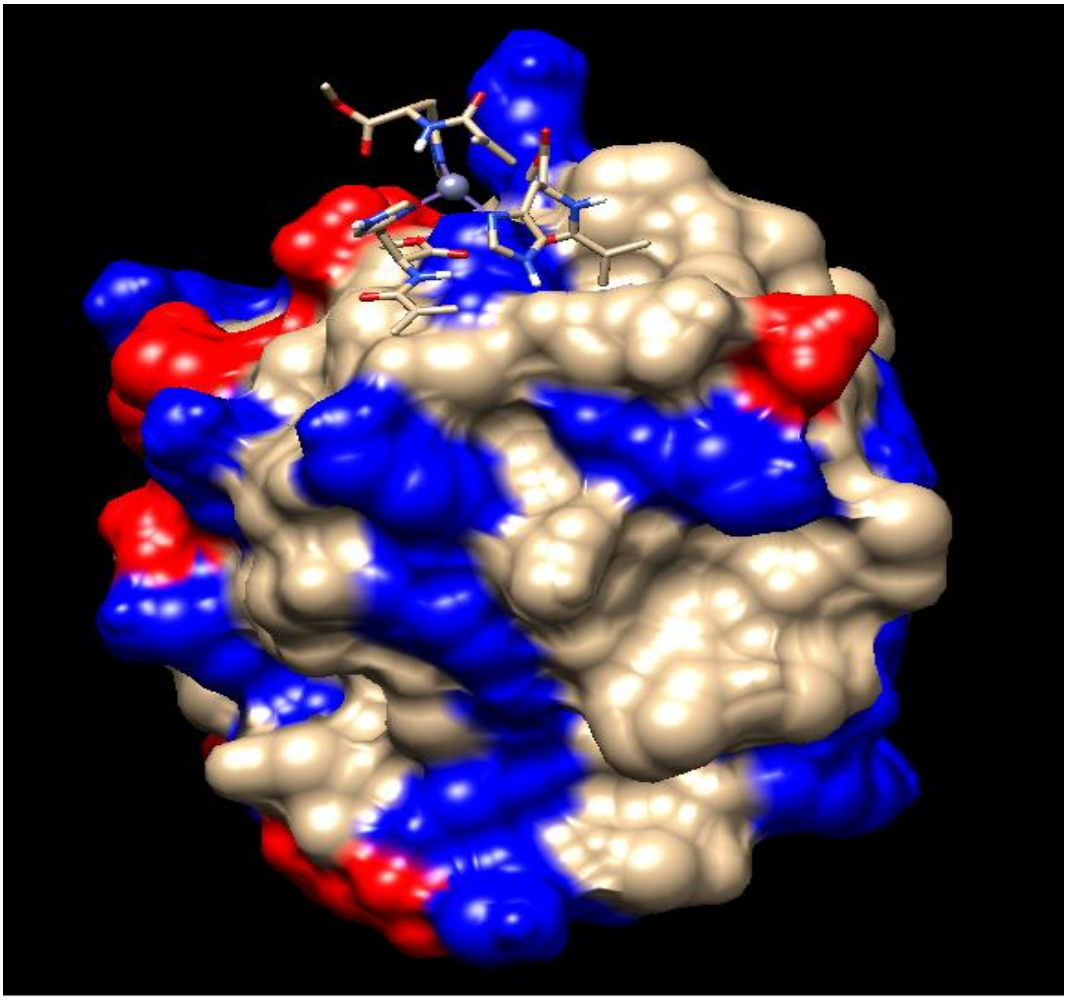

Figure 4.15. Visualization of docking three molecules of MAH bound to Zn2+ Ion at Horse Heart Cytochrome C by using Chimera software, the electropositive regions are in blue. The electronegative regions are in red.

In Figure 4.15, three molecules of MAH bound to Zn^{2+} ion docking at Cytochrome C was visualized by using Chimera software. The electropositive and electronegative regions of Cytochrome C are colored in blue and red, respectively.

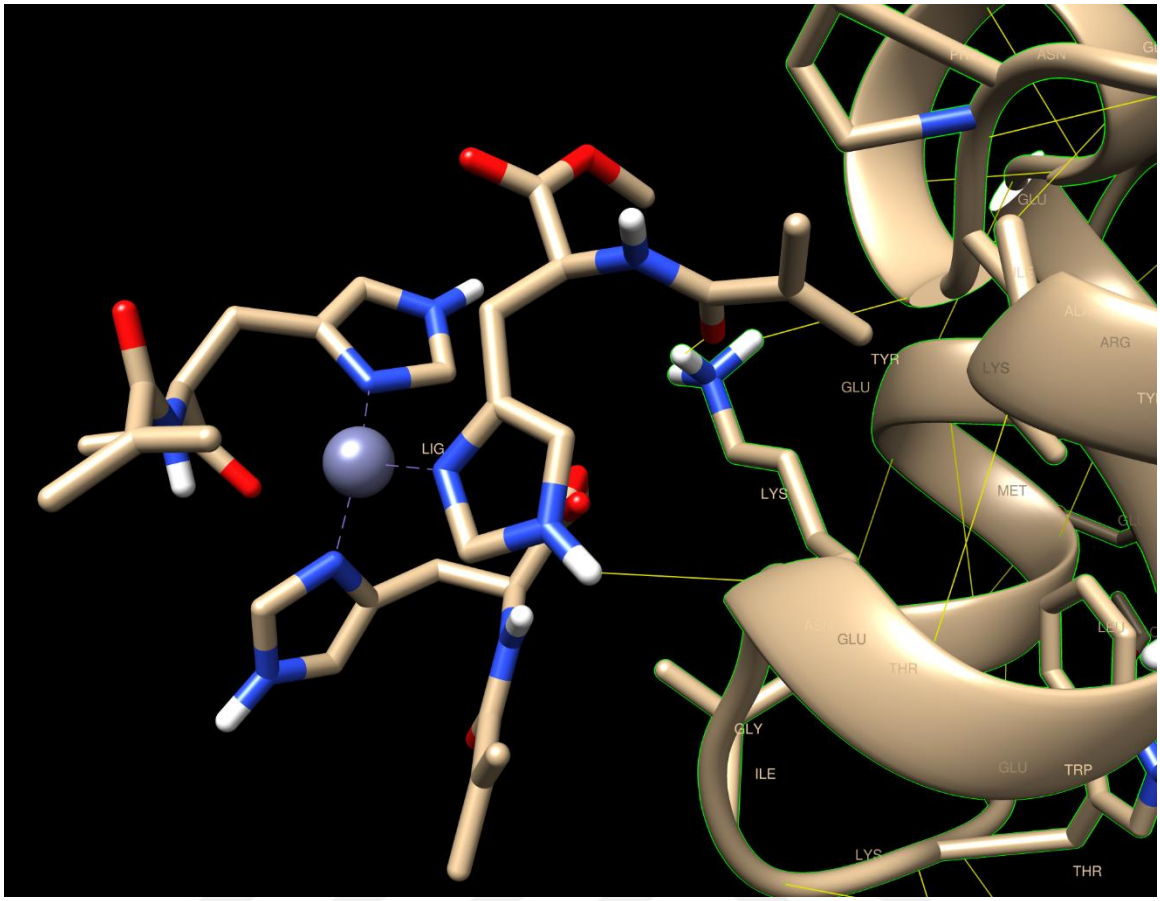

Figure 4.16. Docking of three molecules of mah bound to zn^{2+} ions with horse heart Cytochrome C (yellow lines show the H-bonding interactions).

It was determined that three molecules of MAH monomer bound to Zn^{2+} ion well-fitted in the binding site of Cytochrome C and make two specific H-bonding interaction with Cytochrome C amino acid moiety. One H-bonding interaction mainly occurs between the imidazole group of MAH monomer and the neighboring $\sigma 1-$oxygen of asparagine (ASN-54) residue of Cytochrome C. The other H-bonding interaction occurs between the oxygen of MAH monomer and the neighboring amine group of (LYS-55) residue of Cytochrome C. (Figure 4.16).

4.9. Comparison between the lowest binding energy of MAH- Horse Heart Cytochrome C dockings

AutoDock generates a set of different ligand binding poses and use a scoring function to estimate binding affinities for the generated ligand poses in order to

determine the best binding mode. Low binding energy has better docking score. Obviously, when Zn^{2+} ion is bound to one molecule of N-Methacryloyl-L-histidine Methyl Ester (MAH), binding energy declines and becomes the lowest binding energy of all dockings obtained in this study. The lower the free binding energy, the higher the binding affinity of MAH-Zn^{2+} towards Cytochrome C. It was concluded that one molecule of MAH monomer bound to Zn^{2+} is the most favorable metal chelated monomer docked at Cytochrome C (Diagram 4.2).

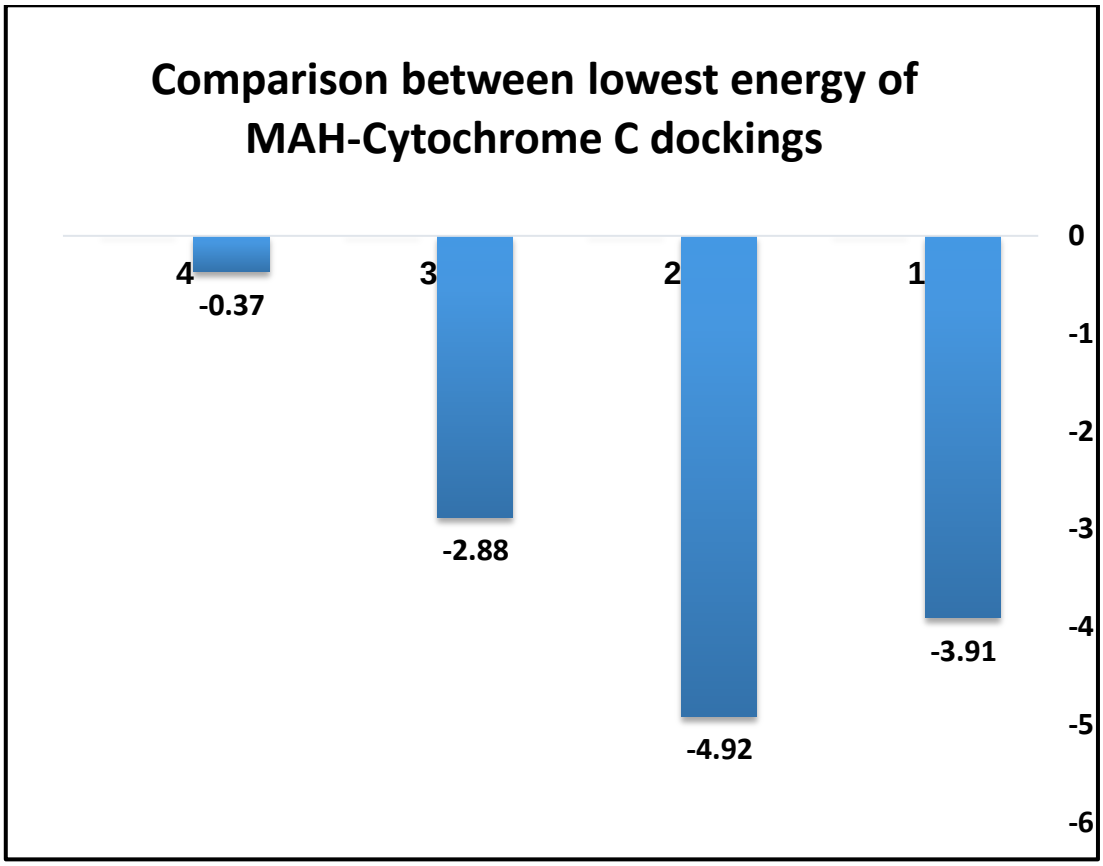

Diagram 4.2. Column1 (MAH- Cytochrome C). Column2 (MAH-Zn^{2+}-Cytochrome C)**.** Column3 (2MAH-Zn^{2+}- Cytochrome C). Column4 (3MAH-Zn^{2+}- Cytochrome C).

5. CONCLUSION

- In this study, we applied computational docking method on the interactions that occur in immobilized metal ion affinity chromatography. N-Methacryloyl-L-histidine Methyl Ester (MAH) is used as functional ligand. Then Zn^{2+} ions are selected to be chelated through imidazole groups on the MAH. N-Methacryloyl-L-histidine Methyl Ester (MAH) was drawn and created using Avogadro which has auto optimization tool.
- Human insulin molecule and horse heart Cytochrome C are selected as targets to be interacted with our functional ligand. The protein human insulin (PDB ID: 3E7Y) and horse heart Cytochrome C (PDB ID: 1HRC) were retrieved from the RCSB protein databank (www.rcsb.org/pdb) and saved in pdb file format.
- Automated docking software AutoDock 4.2 was used for docking of MAH molecules with human insulin and cytochrome c respectively.
- Docking of one molecule of N-methacryloyl-L-histidine methyl ester (MAH) with human insulin molecule showed that the lowest binding energy is -3.43 kcal/mol, whereas docking of one molecule of N-methacryloyl-L-histidine methyl ester (MAH) bound to Zn^{2+} ion with human insulin molecule is reduced the lowest binding energy. By chelating Zn^{2+} Ion, the lowest binding energy became - 4.14 kcal/mol.
- Docking of two molecules of N-methacryloyl- L -histidine methyl ester (MAH) bound to Zn^{2+} ion with human insulin molecule showed that the lowest binding energy is -2.84 kcal/mol, whereas the lowest binding energy in docking of three molecules of MAH bound to Zn^{2+} ion to human insulin molecule is -0.92 kcal/mol.
- Docking MAH bound to Zn^{2+} ion to human insulin molecule is the best result of dockings due to the lowest binding energy - 4.14 kcal/mol and the bonds between the ligand and the protein.
- Docking of one molecule of n-methacryloyl- L -histidine methyl ester (MAH) with cytochrome c molecule showed that the lowest binding energy is -3.91 kcal/mol, whereas docking of one molecule of N-methacryloyl- L -histidine methyl ester (MAH) bound to Zn^{2+} ion with

Cytochrome C molecule is reduced the lowest binding energy. By chelating Zn^{2+} Ion, the lowest binding energy became -4.92 kcal/mol.

- Docking of two molecules of N-methacryloyl- L -histidine methyl ester (MAH) bound to Zn^{2+} ion with cytochrome c molecule showed that the lowest binding energy is -2.88 kcal/mol, whereas the lowest binding energy in docking of three molecules of MAH bound to Zn^{2+} ion to Cytochrome C molecule is -0.37 kcal/mol.
- Docking MAH bound to Zn^{2+} ion to Cytochrome C molecule is the best result of dockings due to the lowest binding energy -4.92 kcal/mol and the bonds between the ligand and the protein.
- This docking analysis reveals that the lowest binding energy increased by using more than one molecule of MAH bound to Zn^{2+} ion as functional ligand.

The binding energy results are inconformity with the related literature [65-67]. In conclusion, this study is promising for improving IMAC research in the field of protein purification and separation. In addition to, the Autodock results of metal chelated and non-chelated ligands, which are MAH and MAH-Zn^{2+}, can also be applied to affinity and molecular imprinted based purification and separation systems.

REFERENCES

[1] Pina, Ana S., et al. "A Tailor-Made "Tag–Receptor" Affinity Pair for the Purification of Fusion Proteins." *ChemBioChem* 15.10 (**2014**): 1423-1435.

[2] Willoughby, N. A., et al. "Immobilised metal ion affinity chromatography purification of alcohol dehydrogenase from baker's yeast using an expanded bed adsorption system." *Journal of Chromatography A* 840.2 (**1999**): 195-204.

[3] Hemdan, Esmat Sayeda, and Jerker Porath. "Development of immobilized metal affinity chromatography: I. Comparison of two iminodiacetate gels." *Journal of Chromatography A* 323.2 (**1985**): 247-254.

[4] Gelūnait, Lina, et al. "Chelated mercury as a ligand in immobilized metal ion affinity chromatography of proteins." *Journal of Chromatography A* 904.2 (**2000**): 131-143.

[5] Nakagawa, Yasuo, et al. "High-performance immobilized metal ion affinity chromatography of peptides: analytical separation of biologically active synthetic peptides." *Analytical biochemistry* 168.1 (**1988**): 75-81.

[6] Porath, Jerker, et al. "Metal chelate affinity chromatography, a new approach to protein fractionation." *Nature* 258 (**1975**): 598-599.

[7] Arnold, Frances H. "Metal-affinity separations: a new dimension in protein processing." *Nature biotechnology* 9.2 (**1991**): 151-156.

[8] Vijayalakshmi, M. A. "Pseudobiospecific ligand affinity chromatography." *Trends in Biotechnology* 7.3 (**1989**): 71-76.

[9] Lönnerdal, B., and C. L. Keen. "Metal chelate affinity chromatography of proteins." *J. Appl. Biochem* 4 (**1982**): 203-8.

[10] Çimen, Duygu, and Adil Denizli. "Immobilized metal affinity monolithic cryogels for cytochrome c purification." *Colloids and Surfaces B: Biointerfaces* 93 (**2012**): 29-35.

[11] O'Brien, Susan M., et al. "Characterisation of non-porous magnetic chelator supports and their use to recover polyhistidine-tailed T4 lysozyme from a crude E. coli extract." *Journal of biotechnology* 54.1 (**1997**): 53-67.

[12] Subramanian, Ganapathy. *Biopharmaceutical Production Technology, 2 Volume Set*. Vol. 1. John Wiley & Sons, **2012**.

[13] http://www.zbh.uni-hamburg.de/en/research/computational-molecular-design-old/research/molecular-docking.html(**2016**)

[14] Mukesh, B., and K. Rakesh. "Review on Molecular docking." *Ijrap* 2.6 (**2011**): 1746-51.

[15] Sousa, Sergio Filipe, Pedro Alexandrino Fernandes, and Maria Joao Ramos. "Protein–ligand docking: current status and future

challenges." *Proteins: Structure, Function, and Bioinformatics* 65.1 (**2006**): 15-26.

[16] McNaught, Alan D., and Alan D. McNaught. *Compendium of chemical terminology*. Vol. 1669. Oxford: Blackwell Science, **1997**.

[17] Magdeldin, Sameh, and Annette Moser. "Affinity chromatography: Principles and applications." *S. Magdeldin, InTech, Croatia* (**2012**): 1-28.

[18] Mohr, Peter. *Affinity chromatography: practical and theoretical aspects*. Vol. 33. CRC Press, **1985**.

[19] Hage, David S., and Jack Cazes, eds. *Handbook of affinity chromatography*. CRC Press, **2005**.

[20] Hage, David S., et al. "Pharmaceutical and biomedical applications of affinity chromatography: recent trends and developments." *Journal of pharmaceutical and biomedical analysis* 69 (**2012**): 93-105.

[21] Hage, David S. "Survey of recent advances in analytical applications of immunoaffinity chromatography." *Journal of Chromatography B: Biomedical Sciences and Applications* 715.1 (**1998**): 3-28.

[22] Kim, Hee Seung, and David S. Hage. "Immobilization methods for affinity chromatography." *Handbook of affinity chromatography* 2 (**2006**): 35-78.

[23] Kool, Jeroen, et al. "Advances in mass spectrometry-based post-column bioaffinity profiling of mixtures." *Analytical and bioanalytical chemistry* 399.8 (**2011**): 2655-2668.

[24] Coelho, Luana CBB, et al. *Protein purification by affinity chromatography*. INTECH Open Access Publisher, **2012**.

[25] https://www.thermofisher.com/tr/en/home/life-science/protein-biology/protein-biology-learning-center/protein-biology-resource-library/pierce-protein-methods/covalent-immobilization-affinity-ligands.html (**2016**)

[26] Glaser, V. "Combinatorial chemistry evolves into strategic technology for optimizing lead compounds." *GENETIC ENGINEERING NEWS* 17.10 (**1997**): 1.

[27] Porath, Jerker, et al. "Metal chelate affinity chromatography, a new approach to protein fractionation." *Nature* 258 (**1975**): 598-599.

[28] Ueda, E. K. M., P. W. Gout, and L. Morganti. "Current and prospective applications of metal ion–protein binding." *Journal of Chromatography A* 988.1 (**2003**): 1-23.

[29] Yip, Tai-Tung, and T. William Hutchens. "Immobilized metal ion affinity chromatography." *Molecular biotechnology* 1.2 (**1994**): 151-164.

[30] Baydemir, Gözde, Ali Derazshamshir, Müge Andaç, Cenk Andaç, and Adil Denizli. "Reversible immobilization of glycoamylase by a variety of

Cu2+-chelated membranes." *Journal of Applied Polymer Science* 126, no. 2 (**2012**): 575-586.

[31] Bakhshpour, Monireh, Emel Tamahkar, Müge Andaç, and Adil Denizli. "Affinity binding of proteins to the modified bacterial cellulose nanofibers." *Journal of Chromatography B* (**2017**).

[32] Gutierrez, R., E. M. Martín del Valle, and M. A. Galan. "Immobilized metal-ion affinity chromatography: status and trends." *Separation & Purification Reviews* 36.1 (**2007**): 71-111.

[33] Blowers, D. P. "Immobilized Metal Ion Chromatography." (**2000**).

[34] Karlström, Amelie Eriksson, and Sophia Hober. "Chromatographic methods for protein purification." (**2006**).

[35] Ueda, E. K. M., P. W. Gout, and L. Morganti. "Current and prospective applications of metal ion–protein binding." *Journal of Chromatography A* 988.1 (**2003**): 1-23.

[36] Janson, Jan-Christer, ed. *Protein purification: principles, high resolution methods, and applications*. Vol. 151. John Wiley & Sons, **2012**.

[37] Porath, Jerker. "Immobilized metal ion affinity chromatography." *Protein expression and purification* 3.4 (**1992**): 263-281..

[38] Derazshamshir, Ali, Gözde Baydemir, Müge Andac, Rıdvan Say, Igor Yu Galaev, and Adil Denizli. "Molecularly Imprinted PHEMA-Based Cryogel for Depletion of Hemoglobin from Human Blood." *Macromolecular Chemistry and Physics* 211, no. 6 (**2010**): 657-668.

[39] Bereli, Nilay, Müge Andaç, Gözde Baydemir, Ridvan Say, Igor Yu Galaev, and Adil Denizli. "Protein recognition via ion-coordinated molecularly imprinted supermacroporous cryogels." *Journal of Chromatography A* 1190, no. 1 (**2008**): 18-26.

[40] Walsh, P. Sean, David A. Metzger, and R. Higushi. "Chelex 100 as a medium for simple extraction of DNA for PCR-based typing from forensic material. BioTechniques 10 (4): 506-13 (April 1991)." *Biotechniques* 54.3 (2013): 134.

[41] Huang, Sheng-You, and Xiaoqin Zou. "Advances and challenges in protein-ligand docking." *International journal of molecular sciences* 11.8 (**2010**): 3016-3034.

[42] Halperin, Inbal, et al. "Principles of docking: An overview of search algorithms and a guide to scoring functions." *Proteins: Structure, Function, and Bioinformatics* 47.4 (**2002**): 409-443.

[43] Sousa, Sergio Filipe, Pedro Alexandrino Fernandes, and Maria Joao Ramos. "Protein–ligand docking: current status and future challenges." *Proteins: Structure, Function, and Bioinformatics* 65.1 (2006): 15-26.

[44] Molecular modelling from Wikipedia, the free encyclopedia https://en.wikipedia.org/wiki/Molecular_modelling (**2016**)

[45] Wang, Renxiao, et al. "The PDBbind database: methodologies and updates." *Journal of medicinal chemistry* 48.12 (**2005**): 4111-4119.

[46] Puvanendrampillai, Dushyanthan, and John BO Mitchell. "Protein Ligand Database (PLD): additional understanding of the nature and specificity of protein–ligand complexes." *Bioinformatics* 19.14 (**2003**): 1856-1857.

[47] Block, Peter, et al. "AffinDB: a freely accessible database of affinities for protein–ligand complexes from the PDB." *Nucleic acids research* 34.suppl 1 (**2006**): D522-D526.

[48] Liu, Tiqing, et al. "BindingDB: a web-accessible database of experimentally determined protein–ligand binding affinities." *Nucleic acids research* 35.suppl 1 (**2007**): D198-D201.

[49] Dias, Raquel, Jr de Azevedo, and F. Walter. "Molecular docking algorithms." *Current Drug Targets* 9.12 (**2008**): 1040-1047.

[50] http://AutoDock.scripps.edu/(**2016**)

[51] http://avogadro.cc/wiki/Main_Page(**2016**)

[52] https://www.cgl.ucsf.edu/chimera/(**2016**)

[53]DeLano, Warren L., and Sarina Bromberg. "PyMOL user's guide." *DeLano Scientific LLC, San Carlos, California, USA* (**2004**).

[54] Wright, James R., et al. "A review of piscine islet xenotransplantation using wild-type tilapia donors and the production of transgenic tilapia expressing a "humanized" tilapia insulin." *xenotransplantation* 21.6 (**2014**): 485-495.

[55] Sonksen, P., and J. Sonksen. "Insulin: understanding its action in health and disease." *British journal of anaesthesia* 85.1 (**2000**): 69-79.

[56] Ibrahim, Ammal Esmaeel. "Investigate the level of IL-8 in patients with diabetes mellitus and its relation with lipid profile and the incidence of atherosclerosis." *European Journal of Chemistry* 5.3 (**2014**): 529-531.

[57] Bell, Graeme I., et al. "Sequence of the human insulin gene." *Nature* 284.5751 (**1980**): 26-32.

[58] Dunn, Michael F. "Zinc–ligand interactions modulate assembly and stability of the insulin hexamer–a review." *Biometals* 18.4 (**2005**): 295-303.

[59] Tafani, Marco, et al. "Cytochrome c release upon Fas receptor activation depends on translocation of full-length bid and the induction of the mitochondrial permeability transition." *Journal of Biological Chemistry* 277.12 (**2002**): 10073-10082.

[60] Ambler, R. P. "Sequence variability in bacterial cytochromes c." *Biochimica et Biophysica Acta (BBA)-Bioenergetics* 1058.1 (**1991**): 42-47.

[61] Moore, Geoffrey R. "Bacterial 4-α-helical bundle cytochromes." *Biochimica et Biophysica Acta (BBA)-Bioenergetics* 1058.1 (**1991**): 38-41.

[62] Coutinho, Isabel B., and António V. Xavier. "[9] Tetraheme cytochromes." *Methods in enzymology* 243 **(1994)**: 119-140.

[63] Bowden, Edmond F., Fred M. Hawkridge, and Henry N. Blount. "Interfacial electrochemistry of cytochrome c at tin oxide, indium oxide, gold, and platinum electrodes." *Journal of Electroanalytical Chemistry and Interfacial Electrochemistry* 161.2 (**1984**): 355-376.

[64] Dickerson, Richard E., and Russell Timkovich. "7 Cytochromes c." *The enzymes* 11 (**1975**): 397-547.

[65] Singh S., Gupta Kumar S., Nischal A., Khattri S., Nath R., Pant Kumar K., Seth Kishore P. "Identification and Characterization of Novel Small-Molecule Inhibitors against Hepatitis Delta Virus Replication by Using Docking Strategies" *Hepatitis Monthly* 11 **(2011)**: 803-809.

[66] Abraham R., Varghese G.K., Nisha N., Sreekumar S. "Molecular Docking of Terminalia Cuneata on Cholesteryl-Esterase" *International Journal of Computational Bioinformatics and In Silico Modeling* 3 **(2014)**: 327-331.

[67] Jena L., Deshmukh S., Waghmare P., Kumar S., Harinath B.C. "Study of mechanism of interaction of truncated isoniazid–nicotinamide adenine dinucleotide adduct against multiple enzymes of Mycobacterium tuberculosis by a computational approach" *International Journal of Mycobacteriology* 4 **(2015)**: 276–283.

CURRICULUM VITAE

Credentials

Name, Surname: Dima Salha

Place of Birth: Aleppo- Syria

Marital Status: Married

E-mail: biodima87@hotmail.com

Address:

Education

BSc: Aleppo University

MSc: Hacettepe University

Foreign Languages

English

Turkish

Arabic (native language)

Work Experience

Lecturer: 2011-2013 Aleppo University

Research assistance: Summer 2009 Molecular Biology Dept. in Dababo Labs.

Area of Experiences

Molecular modelling, PCR techniques, Real Time PCR techniques.

Projects and Budgets

Graduation Project: Detection of Familial Mediterranean Fever by PCR technique.

Publication

Oral and Poster Presentations

Genetic Code

Genome Sequencing Projects

Protein Purification

Replication & Mutations

Photosynthesis

Animal Vectors

Bioremediation & Biopesticides

Using Techniques of Radiation in Immunology System

Monoclonal Antibodies

RNA Transcription in Eukaryotes

Cell Communication

Southern Blotting

Recombinant Phi29 DNA Polymerase

Computational Methods for MicroRNA Target Prediction

Cell Adhesion Molecules

X Ray

Mouse Model of Hepatitis B Virus Infection

HACETTEPE ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ
YÜKSEK LİSANS/DOKTORA TEZ ÇALIŞMASI ORJİNALLİK RAPORU

HACETTEPE ÜNİVERSİTESİ
FEN BİLİMLER ENSTİTÜSÜ
BİYOMÜHENDİSLİK ANABİLİM DALI BAŞKANLIĞI'NA

Tarih: 08/06/2017

Tez Başlığı / Konusu: İmmobilize metal iyon afinite kromatografisinde kullanılan metal iyonların bilgisayarlı ortamda etkileşim özelliklerinin belirlenmesi

Yukarıda başlığı/konusu gösterilen tez çalışmamın a) Kapak sayfası, b) Giriş, c) Ana bölümler d) Sonuç kısımlarından oluşan toplam 74 sayfalık kısmına ilişkin, 06/06/2017 tarihinde şahsım/tez danışmanım tarafından *Turnitin* adlı intihal tespit programından aşağıda belirtilen filtrelemeler uygulanarak alınmış olan orijinallik raporuna göre, tezimin benzerlik oranı % 6'dır.

Uygulanan filtrelemeler:

1- Kaynakça hariç
2- Alıntılar hariç/~~dâhil~~
3- 5 kelimeden daha az örtüşme içeren metin kısımları hariç

Hacettepe Üniversitesi Fen Bilimleri Enstitüsü Tez Çalışması Orjinallik Raporu Alınması ve Kullanılması Uygulama Esasları'nı inceledim ve bu Uygulama Esasları'nda belirtilen azami benzerlik oranlarına göre tez çalışmamın herhangi bir intihal içermediğini; aksinin tespit edileceği muhtemel durumda doğabilecek her türlü hukuki sorumluluğu kabul ettiğimi ve yukarıda vermiş olduğum bilgilerin doğru olduğunu beyan ederim.

08/06/2017

Gereğini saygılarımla arz ederim.

Tarih ve İmza
08/06/2017

Adı Soyadı: DIMA SALHA
Öğrenci No: N13127429
Anabilim Dalı: Biyomühendislik
Programı: Yüksek Lisans
Statüsü: ■ Y.Lisans ☐ Doktora ☐ Bütünleşik Dr.

DANIŞMAN ONAYI

UYGUNDUR.

Prof. Dr. Adil DENİZLİ