In Silico and Experimental Determination of IL1β-Receptor Inhibitor Peptides

by

Fulya Akşit

A Dissertation Submitted to the

Graduate School of Sciences and Engineering

in Partial Fulfillment of the Requirements for

the Degree of

Master of Science

in

Chemical and Biological Engineering

KOÇ ÜNİVERSİTESİ

January 8, 2020

In Silico and Experimental Determination of IL1β-Receptor Inhibitor Peptides

Koç University

Graduate School of Sciences and Engineering

This is to certify that I have examined this copy of a master's thesis by

Fulya Akşit

and have found that it is complete and satisfactory in all respects,
and that any and all revisions required by the final
examining committee have been made.

Committee Members:

Assist. Prof. Cem ALBAYRAK (Advisor)

Prof. Burak ERMAN [Co-Advisor]

Prof. Türkan HALİLOĞLU

Prof. Özlem Keskin Özkaya

Assist. Prof. Hasan DEMİRCİ

Date:

To thrombocytes and "Synapse" 😊

ABSTRACT

***In Silico* and Experimental Determination of IL1β-Receptor Inhibitor Peptides**

Fulya Akşit

Master of Science in Chemical and Biological Engineering

January 8, 2020

Therapeutic peptides have become promising due to their high affinity and specificity to the target, deep tissue penetration, low toxicity, convenience for structural modifications, and cost effectiveness. For a lead molecule to reach the market, approximately a 15-year period and an investment of around 2.6 billion dollars are necessary. One of the most important parameters to evaluate the lead compounds in the process of Drug Discovery and Development is binding affinity, which is the strength of the non-covalent interactions between the ligand and the target molecules. This thesis study focuses on *in silico* design of peptides to block IL1β – Receptor instead of therapeutically used recombinant protein, Anakinra, which has the same structure as the human native IL1 – Receptor antagonist molecule. In order to verify the validity of the computationally designed peptides, surface plasmon resonance methods were used to measure the binding affinity of them.

9 different peptide sequences were designed from the three-dimensional structure of IL1-Receptor Antagonist protein by considering the type and number of non-covalent interactions that form with the receptor. In order to evaluate the binding affinities of these peptides, molecular docking simulations were performed for 100 runs by fixing the bonds. Peptide C, which is Arg – Ile – Trp – Asp – Val – Asn – Gln, showed the highest binding affinity with K_D of 7.91×10^{-8} M. The docking studies were repeated to confirm the binding affinity of Peptide C without fixing the bonds for 999 runs. Since Peptide C is consistent in terms of its binding affinity to the receptor, N- terminus and C- terminus modifications were applied to increase the conformational stability of the peptide. N- terminus modification resulted as Phe insertion to the amino group of Peptide C, while C- terminus modification requires Lys insertion. The modifications increased the number of hydrophobic interactions while also keeping hydrogen bonds between the peptide and the

receptor that both N- and C- termini modified peptide has K_D of $0.24x10^{-9}$ M. To understand unbinding pathway and binding affinities of our designed peptides in a more accurate and realistic environment, SMD simulations were performed for unmodified sequence which is Peptide C, N- terminus modified 8- amino acid peptide, and both N- and C- termini modified 9- amino acid peptide. The binding affinity values for the three of the peptides from SMD simulations are consistent with the average values obtained from the docking. Since water molecules are present in SMD simulations, they interfere between the peptide and the receptor molecules that the distance between the interactions increase thereby, causing the loss of some of the hydrogen bonds and most of the hydrophobic interactions. Due to this fact and realistic aspect of SMD, three of the peptides show binding affinities to the receptor at the average values of the docking results.

In Chapter 3, three of the peptides obtained as a result of the computational studies were synthesized using Fmoc Solid Phase Peptide Synthesis methods. High Performance Liquid Chromatography was used to analyze the synthesis. Even though, we showed that our first synthesis which is for the unmodified peptide was successful, due to a problem occurrence either in the column or in the instrument, after a while, we even could not obtain the chromatograms of the standard samples of the peptides. The chromatograms of the standard samples were obtained using a different instrument. However, due to TFA fragmentation of the peptides, our synthesis products might have corrupted that further studies continued with the custom synthesis peptides.

Finally, to validate the binding affinity of the peptides experimentally, Surface Plasmon Resonance (SPR) method was used. In Biacore T200 instrument, CM5 sensor chip surface was immobilized with IL1-Receptor using amine coupling chemistry. To verify the accuracy of the instrument, binding affinity of IL1-β to IL1-Receptor was measured performing kinetics analysis. Concentration dependency was obtained and K_D was found out as 2.8 nM which is a consistent result with the literature value, 2.7 nM. Although concentration dependency was also obtained for the peptides, binding affinity values could not be obtained yet due to the bulk effect of DMSO. The kinetics analysis to measure binding affinity of the peptides will be repeated by performing solvent correction prior to the analysis to prevent bulk effect of DMSO.

ÖZETÇE

Bilgisayar Ortamında ve Deneysel Olarak IL1-β Reseptör İnhibitörü Peptitlerin Belirlenmesi

Fulya Akşit

Kimya ve Biyoloji Mühendisliği, Yüksek Lisans

8 Ocak 2020

Terapötik peptitler yüksek bağlanma ve özgül bağ kurma yeteneleri, dokunun içerisine derinlemesine girebilme, düşük toksisite ve yapısal modifikasyonlarının kolay olması özelliklerinden dolayı ilaç tasarımı için gelecek vaat eden moleküller olmuşlardır. Bir ilaç adayı molekülün pazara girebilmesi için yaklaşık 15 senelik bir araştırma ve geliştirme süreci ile birlikte 2.6 trilyon dolarlık yatırım gerekmektedir. İlaç Tasarımı ve Geliştirme sürecinde aday molekülün incelenmesi için gerekli en önemli parametrelerden birisi aday molekülün hedefe olan bağlanma derecesidir. Bağlanma derecesi, ligand ve hedef molekül arasındaki kovalent olmayan bağların güçlülük değeridir. Bu tez çalışmaları IL1β-Reseptör proteinini bloke etmek için terapötik olarak kullanılan, insanda doğal olarak bulunan IL1-Reseptör antagonist molekülü ile aynı yapıya sahip ve bir rekombinant protein olan Anakinra yerine bilgisayar ortamında peptit tasarımına odaklanmıştır. Bilgisayar ortamında hesaplamalı olarak tasarımı yaılan peptitlerin geçerliliğini göstermek amacıyla da bağlanma derecesini ölçebilecek bir metot olan yüzey plazmon rezonans deneyleri yapılmıştır.

IL1-Reseptör proteininin üç boyutlu yapısı incelenerek ve kovalent olmayan bağların türü ve sayısı dikkate alınarak bu yapıdan 9 farklı peptit dizilimi tasarlanmıştır. Bu peptitlerin bağlanma derecelerini inceleyebilmek için moleküler seviyede yerleştirme (docking) çalışmaları her bir peptit için aralarındaki bağları sabit tutarak 100 kez yapılmıştır. Arg – Ile – Trp – Asp – Val – Asn – Gln dizilimine sahip Peptit C, 7.91×10^{-8} M değerindeki K_D değeri ile tasarlanan peptitler arasında en yüksek bağlanma derecesini göstermiştir. Peptit C'nin yüksek bağlanma derecesine sahip olduğunu teyit etmek için docking çalışmaları peptidin bağlarını sabitlemeden 999 defa yapılmış ve benzer sonuç elde edilmiştir. Peptit C'nin bağlanma derecesinin yüksekliğinin tutarlılığını gösterdikten

sonra yapısal kararlılığını arttırmak amacıyla peptidin amino ve karboksil grup uçlarına modifikasyon yapılması amaçlanmıştır. N- terminal modifikasyonu Phe eklenmesi sonucunu verirken C- terminal modifikasyonu da karboksil grubu ucuna Lys aminoasidi eklenmesi sonucunu vermiştir. Bu modifikasyonlar peptit ve reseptör arasındaki hidrojen bağlarını tutarak aradaki hidrofobik etkileşimleri arttırmış ve hem N- hem de C-terminalleri modifiye olmuş 9 aminoasitli peptit $0.24x10^{-9}$ M K_D değeri sonucunu vermiştir. Tasarlanan peptitlerin kopma yolunu anlamak ve bağlanma derecelerini daha doğru ve gerçekçi bir ortamda incelemek amacıyla SMD simülasyonları üç peptit için de yapılmıştır. SMD sonuçlarına göre üç peptit için de bağlanma dereceleri docking çalışmalarından alınan ortalama değerlerle tutarlılık göstermektedir. SMD simülasyonlarında ortamda su bulunduğu için peptit ve reseptör moleküllerinin arasına su molekülleri girerek etkileşimler ve bağlar arasındaki mesafelerin artmasına ve bu da bazı hidrojen bağlarının ve çoğu hidrofobik etkileşimlerin kaybolmasına neden olmaktadır. Bu nedenlerden dolayı da peptitler, docking çalışmalarından alınan en iyi sonuçlar yerine ortalama değerlere yakın bağlanma dereceleri göstermektedir.

Tezin 3. bölümünde bilgisayar ortamında hesaplamalı olarak tasarlanan peptitler Fmoc kimyasının kullanıldığı Katı Hal Peptit Sentezi metoduyla sentezlenmişlerdir. Yüksek Performanslı Sıvı Kromatografi metodu da yapılan sentezi incelemek için kullanılmıştır. Modifiye olmamış peptit ilk sentezimiz olmuş olup o sentezin başarılı olduğunu göstermemize rağmen kromatografi kolonunda veya cihazın kendisinde daha sonradan oluşan bir problemden dolayı peptitlere ait standard örenklerin dahi kromatogramları elde edilememiştir. Daha sonra standard örneklerin kromatogramları başka bir cihazda aynı metodu kullanarak elde edilmiştir. Ancak TFA'nın peptidi parçalara ayırma özelliğinden dolayı bu bekleme sürecinde peptitler bozulmaya uğramış olabilir ve gelecek çalışmalara standard olarak kullandığımız peptitlerle devam edilmiştir.

Son olarak, deneysel olarak peptitlerin bağlanma derecelerini ölçmek amacıyla Yüzey Plazmon Rezonans metodu kullanıldı. Biacore T200 cihazında CM5 sensör çipine amin birleştirme kimyası yoluyla IL1-Reseptör proteini yerleştirildi. Cihazın doğru çalıştığından emin olmak için IL1-β proteininin IL1-Reseptör Proteinine olan bağlanma derecesi ölçüldü ve K_D değeri 2.8 nM olarak elde edildi. Literatürde de bu değerin 2.7 nM olarak elde edildiği gösterilmiştir. Peptitler için de konsantrasyona bağlılık elde edilmiş olsa da DMSO'nun hacim etkisinden dolayı bağlanma dereceleri ölçülememiştir. Peptitlerin bağlanma derecelerini ölçmek için yapılacak olan kinetik analiz deneyleri

DMSO’nun hacim etkisini engellemek için çözücü doğrulama metodu kullanarak tekrar edilecektir.

ACKNOWLEDGEMENTS

During my master study, I have gained a lot of experience and information in both computational and experimental methods. I always wanted to study a research topic related to drug design and this thesis work gave that opportunity for me. Along with my advisors inspiring suggestions and motivations, I have learned how to progress in research.

First of all, I would like to present my appreciation to my advisor Assist. Prof. Cem Albayrak and my co-advisor Prof. Burak Erman for giving me the opportunity to work both *in silico* and experimentally on peptide design under their supervision with their precious help, support and motivations during my studies. After each weekly meeting with Cem Albayrak, I have noticed that I have learned new approaches and way of thinking from his experience and ideas. I have gained different perspectives from the discussions that we made. The meetings we had at Burak Erman's office are very memorable for me. There were moments that I was just watching the flow of different ideas which inspired me, and I have always felt grateful that I could get the opportunity to study under their light of inspiration. I believe that each discussion from those meetings will guide my way of becoming a researcher. In addition, I would also thank to both for their precious patience and understanding of my situation during my mother's treatment.

I would also present my appreciation to Prof. Batu Erman for allowing us to conduct surface plasmon resonance measurements with Biacore T200 and his valuable comments and ideas about our experiments and the analysis results. I would also thank to Sofia Piepoli for all her help during the experiments.

I also would like to thank Assist. Prof. Hasan Demirci for his precious support and motivations. Each discussion that I made with him has broadened my knowledge and reminded me to work with my passion. He has always inspired me with his knowledge, experience, ideas, and way of thinking to overcome the problems.

I would also present my appreciation to my committee members Prof. Türkan Haliloğlu and Prof. Özlem Keskin Özkaya for their valuable comments and time.

I also would like to thank Prof. Seda Keskin Avcı for providing me their licensed software, GOLD, to perform my molecular docking studies.

I also would like to present my appreciation to Aysima Hacısüleyman for becoming a dear friend. Beyond that, I would thank her for her very valuable help and support since the beginning of my master study. She is one of the greatest and dearest part of my graduate years.

I am also grateful both Koç University and TUBITAK - 217S655 project for funding me.

Finally, I would like to present my deepest appreciation to Gökhan Tanısalı for his unforgettable precious support, patience, and providing me with the benefit of his wisdom😊. I also would like to present my appreciation to my family for their permanent support. Especially, I am very grateful to my lovely brother, Fuat Akşit, for his valuable help, support, and friendship. I am very thankful for having a very friendly graduate school environment that I feel myself very lucky to have Yağmur Ersoy as one of my dearest friends. I also would like to thank Merve Yılmam, Houman Jalali, Beste Tunalı, Ehsan Sarayloo, Jerome Johnson and Selin Turan.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1 :
INTRODUCTION

In this thesis, the primary goal was to computationally design a peptide antagonist molecule which targets IL1β receptor (IL-1R1) protein and validate its binding affinity experimentally. In order to reach this objective, parameters such as type and number of noncovalent interactions for the computational design of peptides were determined. For the evaluation of designed peptides, molecular docking and steered molecular dynamics simulations were performed. One of the important results obtained from the computational studies is the dissociation constant of the ligand/antagonist molecule from the target/receptor, which is also referred to as the binding affinity of the ligand. In the following sections of this chapter, the process becoming a drug candidate and the phases of the process will be discussed. In addition, the important parameters to evaluate and to analyze the candidates such as the dissociation constant, its relation to the Gibbs free energy and thereby, how the strength of non-covalent interactions contribute to binding affinity are discussed.

1.1 Drug Discovery and Development

Drug discovery and development takes on average of 14 years, and includes five different phases (Figure 1). The process is named drug discovery, however the discovered molecules are generally not drug. They are referred to lead compounds. The lead compounds are discovered and they can become drugs at the end of the 5-phase process. After they are structurally modified and pass to the preclinical phase, they are called drug candidates. Basically, the lead molecules can be considered as prototype compounds with different attractive characteristics, such as biological and/or pharmacological activity which refer to low toxicity, high absorption, compatibility to metabolism, appropriate half-life, and high affinity toward the target molecule, respectively. However, the lead compounds might also have undesirable characteristics including high toxicity, absorption difficulties, and insolubility or metabolism problems [1].

Figure 1.1: The general flow path of drug discovery and development process. Adapted from [1].

Figure 1.1 shows the phases of drug discovery and development process which includes 5 phases for a drug molecule enter the market. The first phase, drug target selection, as its name suggests, includes the selection of the molecule from metabolism to be targeted and thereby the disease is either treated or prevented. The target molecules are usually proteins, however they can also be lipids, or nucleic acids such as DNA or RNA. For the drug target selection phase, the crucial step is to understand the metabolism causing the disease at the cellular level. The next important thing is to focus on understanding cellular networks of proteins or disease pathways, because most of the time, a molecule in the pathway is not only related to one disease or a protein. However, a single molecule may also transmit messages to several different proteins from different cellular networks, and thus, can be affected by functions of multiple pathways. Knowing and understanding the roles of proteins of a disease pathway makes it possible to identify and choose appropriate targets selection for drug discovery and development [2].

Once the drug target selection is complete, lead discovery phase initiates. This phase may involve the use of several methods, for instance, high-throughput screening and computer-based design to find molecules, either natural or chemically modified compounds or biologics, that will interact with the target molecule. Molecules identified in this phase and shown to bind the target molecule are called leads or lead compounds. High-throughput screening is a method to test a large number of compounds using a software at low concentrations in terms of their activity, at the end of which, the target molecules are categorized as either antagonists (which inhibit the target molecule) or agonists, i.e. activator molecules. The goal of this phase is to identify leads which confirm the alteration of the disease evaluating their binding affinity towards the target. In addition, computer-based design uses molecular docking simulations for the selected target and leads, thereby finding concentration values for leads to inhibit or activate the target molecule. The lower the concentration values, the better the binding properties the higher the affinity. As a result of the molecular docking simulations, the lead compounds with the highest affinity are chosen for future phases [3].

The chosen lead compounds are then structurally modified to improve its safety and effectiveness by lowering its toxicity, and preventing issues related to absorption, solubility, and/or metabolism. When the structural modification of the leads is complete, they can be tested in the preclinical environment. After the completion of the structural modification of lead, its terminology changes to drug candidate. Hence, it enters Phase 4, Preclinical and Clinical Development phase, as a drug candidate.

The Preclinical and Clinical Development phase is a very important part of the drug discovery and development process. On average, only one of 10,000 leads ultimately enters the market, and most of drug candidates are eliminated in this phase. Although Preclinical and Clinical Development involves computer-aided analysis methods and numerous experiments to analyze and understand the effects of the drug candidate in disease metabolism, these methods are insufficient. Animal studies are crucial to understand complex disease mechanisms. Moreover, governments and regulatory authorities require companies test the drug candidates in animals, prior to testing them in humans. If the safety and efficacy of a drug candidate can be proven in patients and has the required pharmacokinetic properties (such as appropriate absorption and metabolism in the patient's body), then it can be taken to the fifth and final phase for Regulatory Approval.

Each drug has to be registered for entry into the market. For registration, all the data from clinical trials including those that demonstrate the safety and efficacy of the drug candidate are compiled and submitted to the region's regulatory authority. Additionally, the safety of the drug is determined by post authorization and monitored at defined intervals conducting the clinical tests. After all steps are complete, the drug can be marketed and used by patients. The drug discovery and development process is designed to make innovative medicines effective and safe for patients, both for the patients to overcome the disease and to improve their quality of life. As mentioned above, this process lasts an average of 14 years and costs around 2.6 billion USD [4].

In the following section, one of the main important parameters for the design of a lead molecule, dissociation constant, its relationship to the Gibbs Free Energy, and how the non-covalent interactions contribute to the affinity are discussed.

1.1.1 Non-Covalent Interactions in Molecular Recognition Events

Cellular networks are intricate and involve complex interactions between proteins, sugars, lipids, nucleic acids and small molecules. Excess or insufficient production of and/or abnormality in the structure of the target molecule could cause malfunctioning of the cellular networks and thus, disease pathways. In addition to these reasons, the strength of the interactions in the pathways are very important for the network components to function properly. In a similar manner, the strength of the interactions forms the basis of drug discovery and the development process. It was previously stated that lead discovery phase includes high-throughput screening and/or computer-based design methods to obtain leads functioning as the antagonist or agonist molecules at low concentrations. The analysis criteria to select the lead compounds is the binding affinity which is measured as the strength of the interactions. It should also be noted that non-covalent interactions contribute the most to binding affinity, in other words, the strength of interactions. Therefore, they play important roles in the recognition events between biological molecules.

Binding of one molecule to another is generally known as molecular recognition events, examples include proper substrate recognition by enzymes, transmission of cellular signals through the pathways, and the control of transcription and translation. Non-covalent complexes form during the molecular recognition events with ionic, hydrogen bonding, and hydrophobic interactions that they hold such complexes together. As non-covalent interactions are much weaker than covalent bonds, a considerable amount of these complexes can dissociate even at room temperature, thereby forming unbound and bound species at equilibrium. The strength of the molecular interaction can be deduced from the concentration of unbound and bound molecules. Therefore, non-covalent interactions play a critical role to analyze the binding affinity of the leads by measuring the strength of the interactions between the two molecules. Furthermore, in biochemical terminology, lead compounds are referred to as ligands, which are typically small molecules, organic compounds that bind to a target macromolecule called a receptor, which is usually a protein. [5].

The dissociation constant, K_D determines the strength of non-covalent interactions between the receptor and the ligand. The value of the strength of the interactions is referred to as binding affinity. K_D is related to the standard Gibbs free energy, ΔG^0. As the free energy during binding decreases, the binding affinity increases. The standard

Gibbs free energy change during complex formation is called the binding free energy change [6]. The equation 1.1 gives the relation between the dissociation constant and the change in free energy upon binding while Equation 1.2 gives how to deduce dissociation constant from the concentration of the molecules:

$$\Delta G°_{bind} = -RT\ lnK_D \tag{1.1}$$

$$K_D = \frac{[P].[L]}{[P.L]} \tag{1.2}$$

$\Delta G°_{bind}$ has units of kJ x mol^{-1}. R is gas constant and has the unit of kJ x mol^{-1} x K^{-1}, while T has the unit of K and thus, RT also has units of energy, kJ x mol^{-1}. Then, lnK_D mathematically appears as a dimensionless number. However, the unit of K_D is in concentration units, Molar. There are also standard state concentration terms which take place in the expression of equilibrium constants. However, the values of the standard state concentrations are neglected. The complete expression of dissociation constant, K_D, is:

$$K_D = \frac{\frac{[P]}{[P]°}.\frac{[L]}{[L]°}}{\frac{[P.L]}{[P.L]°}} \tag{1.3}$$

, where P stands for protein, and L for ligand. [P]°, [L]° and [P.L]° are standard state concentrations and they are equal to 1 M. If the equation 1.3 is rewritten:

$$K_D = \left(\frac{[P.L]°}{[P]°.[L]°}\right).\frac{[P].[L]}{[P.L]} = \left(\frac{[P.L]°}{[P]°.[L]°}\right).K_D^* \tag{1.4}$$

K_D^* is a pseudo equilibrium constant:

$$K_D^* = \left(\frac{[P].[L]}{[P.L]}\right) \tag{1.5}$$

As can be seen from equations 1.4 and 1.5, K_D^* has units of concentration, Molar. The value of the term $\left(\frac{[P.L]°}{[P]°.[L]°}\right)$is equal to $\frac{1}{M}$ and as K_D^* is 1 M, K_D appears as a dimensionless number. Hence, the numerical values of K_D and K_D^* are the same but they have different units. In practice, K_D and K_D^* are used interchangeably that molar units are referred to K_D.

For the tightest non-covalent interactions in biochemistry, dissociation constants are in the range of nanomolar to femtomolar, whereas the weakest ones are in millimolar range. The dissociation constants of the tightest interactions correspond to approximately -50 kJ/mole binding free energy change, and -17 kJ/mole for weaker interactions. Table 1.1 presents biologically important noncovalent interactions, their approximate dissociation constants, and corresponding binding free energy change values [6].

Table 1.1: K_D and corresponding $\Delta G°_{bind}$ values of biologically important non-covalent interactions. Adapted from [6].

Type of Interaction	**K_D (Molar)**	**$\Delta G°_{bind}$(at 300K) (kJ/mol)**
Enzyme:ATP	10^{-3} to 10^{-6}	-17 to -35
Signaling protein binding to a target	10^{-6}	-35
Sequence-specific recognition of DNA by a transcription factor	10^{-9}	-52
Small molecule inhibitors of proteins (drugs)	10^{-9} to 10^{-12}	-52 to -69
Biotin binding to avidin protein (Strongest known non-covalent interaction)	10^{-15}	-86

Table 1.2: Approximate energy values of non-covalent interactions. Adapted from [7].

Type of Interaction	**Energy Values (kJ/mol)**	**Energy Values (kcal/mol)**
Electrostatic Interactions	5.9	1.4
Hydrogen Bonding	4-13	1-3
van der Waals Interactions	2-4	0.5 - 1

Macromolecules such as proteins have active binding sites and natural ligands that bind specifically at these sites. These natural ligands could be either naturally occurring molecule, or if there is no natural ligand for the receptor, ATP or GTP are always

necessary for the energy requirements of the macromolecules and cells. The drug molecules displace the natural ligands by binding to the active site of the protein. This working mechanism is called competitive inhibition. For competitive inhibitors, the effective strength of interactions depends on the concentration of natural ligands and the dissociation constant for an inhibitor binding to a protein's active site is referred to as K_I. The important parameter to understand and analyze the affinity and efficacy of drug molecules is IC_{50}, which is the corresponding value of the concentration of the inhibitor that reduces the activity of drug molecules by 50% in the presence of inhibitor molecules and the natural ligand [7]. The relation between IC_{50} value and K_I is given by the following equation:

$$K_I = IC_{50} \times \left(\frac{K_D}{K_D + [L]}\right) \tag{1.6}$$

In addition to all the mentioned parameters and criteria, conformational changes in the protein upon binding of the ligand is one of the major factors that govern the drug design. Analysis of the cavities of the protein structures, only the potential high affinity antagonists or agonists could be designed. This is the first step to design a drug molecule. However, the conformation of a protein is intrinsic to its function and its flexibility and hence, its conformation changes while a drug or its natural ligand enters or exits the active site. Although conformational changes of the protein upon ligand binding could be advantageous, they may also decrease the specificity of a drug. For instance, both imatinib and dasatinib are cancer drugs for chronic myelogenous leukemia. While imatinib penetrates deeply into the tyrosine kinase active site to inhibit uncontrolled phosphorylation, dasatinib cannot penetrate in that active site as much as imatinib. Even though dasatinib has 350 times higher binding affinity, due to the structural differences between the two molecules, it has lower specificity than imatinib [6, 8, 9].

The most significant contribution to binding free energy comes from hydrophobic interactions. Since, they can affect a lot of atoms forming a stack of interactions. In protein – ligand interactions, hydrophobic moieties plus aromatic groups of protein repel water molecules and polar side-chain groups of amino acids, thereby forming a stack of hydrophobic interactions affecting a lot of atoms which contributes to strength of interactions. Even though increasing hydrophobicity of the protein-ligand interface is a way to increase binding affinity, it is not a good option in practice for drug design. The

hydrophobic effect lowers the specificity of interactions while increases the strength of the strength [10].

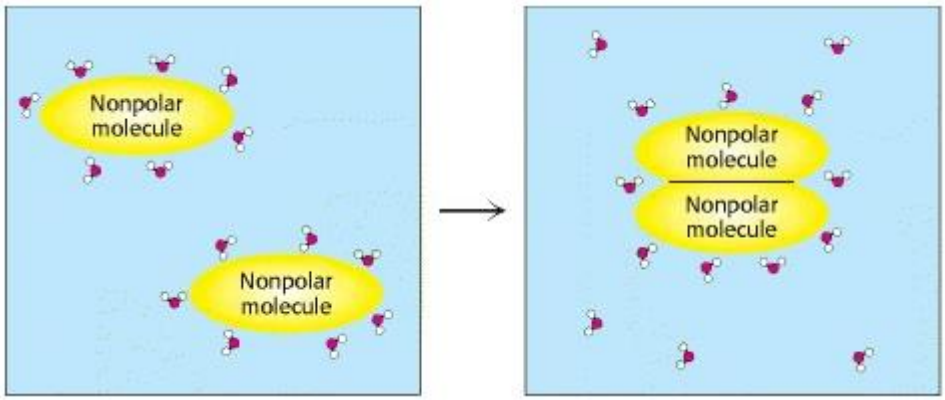

Figure 1.2: Formation of aggregation between non-polar molecules in water. Adapted from [11].

Parameters such as dissociation constant, binding affinity, and the strength of non-covalent interactions as well as understanding their relation among them are the key points to consider in *in silico* design of ligands. In this thesis, peptides were chosen as antagonists / inhibitory ligands due to their promising features, which will be discussed in the following section.

1.1.2 Therapeutic Peptides

Peptides are defined as short (<50) chains of amino acid monomers. As they have much fewer amino acids than proteins, they do not have higher order (tertiary or quaternary) structures. They only exhibit a high degree of secondary structure. Figure 3 shows the difference between an amino acid, peptides, and proteins.

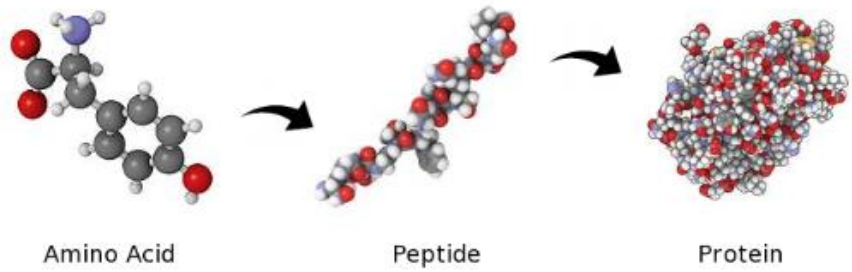

Figure 1.3: Amino acids come together to form peptides and, as the number of residues increase, proteins form. Adapted from [12].

Peptides have several advantages over both small molecules and therapeutic proteins. When compared to small molecules, peptides have higher binding affinity and

specificity to their target. In addition to them, their toxicity profiles are lower that peptides are better candidates for drug market. The binding affinity values of therapeutic peptides are less than 10 nM, thus, they are classified as high-affinity molecules. Moreover, due to their small size, peptides can penetrate more deeply into tissues than therapeutic proteins, which increases their efficacy in cancer treatment [13]. For instance, Lupron™, which is a cytotoxic therapeutic peptide, was approved to use against prostate cancer in 1985[14].

Their other advantages over proteins are that therapeutic peptides can be stored even at room temperature. Additionally, as peptides lack tertiary structure, modifying them is much easier than that of protein structures. As discussed in previous sections, structural modification is crucial for lead compounds to become drug candidates. A structural modification within a protein can cause conformational changes, thereby altering its binding properties. Thus, peptide-based lead compounds could be structurally modified easily that makes them promising for research and development of drug discovery [15, 16]. As peptides can be produced by chemical synthesis, unlike protein production, they can reach the market in a shorter time period while also lowering the costs. However, peptides have limitations for therapeutic applications: Short circulating half-life in bloodstream caused by rapid filtration through kidneys, physical and chemical instability.

Short circulating half-life is one of the major drawbacks for the peptide-based drug candidates reducing their chances to reach the market as they become less efficacious. Conjugation of / functionalization with various molecules, and substitution of non-canonical amino acids are known to extend the half-life of peptides [17, 18]. For example, polyethylene glycol (PEG) conjugation has been used to extend the half-life of a therapeutic peptide [18, 19]. Since, non-canonical aminoacids are unrecognizable by proteases, non-canonical amino acids substitution is a way to prevent rapid degradation of peptides [20]. As peptides can be chemically synthesized, conjugation of various molecules, and noncanonical amino acids substitution to peptides can be easily performed during the synthesis [20].

Until today, more than 7000 naturally occurring peptides have been discovered. They play diverse roles such as neurotransmitters, growth factors, signaling molecules, and hormones. Due to the promising features of peptides such as easier structural modifications and cost effective production, there is a growing market in drug discovery.

So far, there are 60 therapeutic peptides have reached the market and around 150 molecules are in clinical trial phase [21, 22]. Figure 1.4 presents a dataset until 2017. 68 of the 484 candidates were approved for medical use, while 8 were withdrawn from the market, and 155 are currently in clinical development [22].

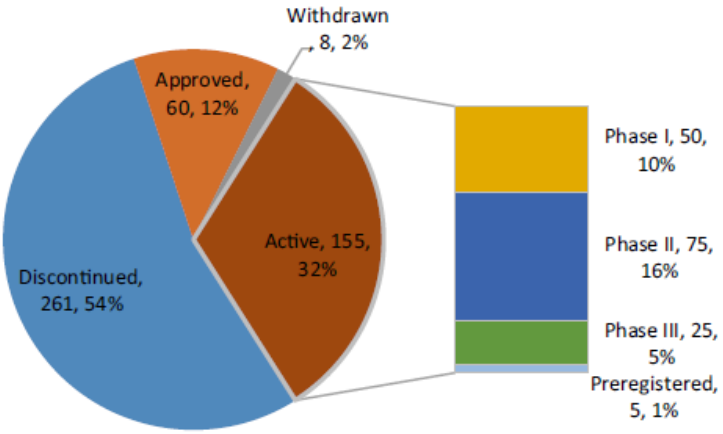

Figure 1.4: Therapeutic peptides currently under development since 2017. Adapted from [22].

Due to their promising features such as high affinity and specificity to target, lower toxicity profiles, cost effectiveness, and easier structural modifications, therapeutic peptides have become of interest for the studies of this thesis. In the first part of the thesis, the computational design of peptides was performed for Interleukin-1 beta (IL-1β) protein.

1.2 IL-1β MECHANISM

Interleukin-1 (IL-1) is a large proinflammatory cytokine family that includes 11 members from Interleukin-1 Family member 1 to 11 (IL-1F1 to IL-1F11) in human and mice [23]. IL-1 family has important roles in both autoinflammatory and autoimmune diseases such as Behcet's disease, Multiple Sclerosis, Alzheimer's disease, asthma, osteoarthritis, Rheumatoid Arthritis, Schnitzler Syndrome, Familial Mediterranean Fever, and Type I Diabetes but it should be noted that the list of diseases, which IL-1 pathway takes role, is not restricted to these examples [24].

IL-1α and IL-1β are the most predominantly secreted members of IL-1 family that they are the most intensely studied to understand their roles in a number of diseases [23, 25]. There are two types of receptors that IL-1α and IL-1β bind, which are IL-1 Receptor Type 1 (IL-1RI) and IL-1 Receptor Type 2 (IL-1RII). Both IL-1α and IL-1β bind to IL-

1RI and IL-1RII. However, IL-1RII lacks a signaling-competent part that cannot initiate signaling [23]. When either IL-1α and IL-1β binds to IL-1RI, this binding event induces the binding of a second receptor chain to the protein and receptor complex, which is IL-1 Receptor Accessory Protein(IL-1RAcP) [26]. Figure 1.5 presents the requirement of IL-1RAcP to initiate intracellular signaling.

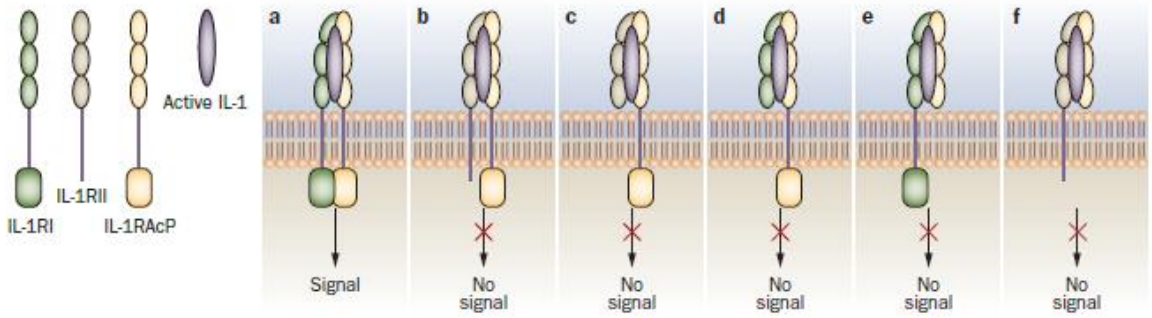

Figure 1.5: Requirement of IL-1RAcP for IL-1 signal transduction. Adapted from [26].

As IL-1RAcP is only necessary for the initiation of intracellular signaling, and also to computationally save time, binding simulations and experiments were performed using IL-1RI, while IL-1RAcP was omitted.

Figure 1.6 shows the pathway of IL-1α and IL-1β including their interaction with other pathways. IL-1α and IL-1β can rapidly induce gene expression within the cells. Once IL-1α or IL-1β binds to IL-1RI, and IL-1RAcP binds the two-protein complex followed by IL1-RAcP ligation, the resulting trimeric complex initiates a signaling cascade. Toll- and IL-1R-like domains on IL-1RI and IL-1RAcP, respectively, get close to each other and they assembly myeloid differentiation primary response gene 88 (MYD88), IL-1 receptor-associated kinase 4 (IRAK4), and Toll-interacting protein (TOLLIP) [27]. Following the formation of MYD88, TOLLIP, and IRAK4, autophosphorylation of IRAK4 starts which subsequently initiates phosphorylation of IRAK1 and IRAK2. After the phosphorylation, tumor necrosis factor-associated factor 6 (TRAF6) is recruited and oligomerized [28-30]. TRAF6 and phosphorylated IRAK1 and IRAK2 dissociate from the initial receptor complexes and they migrate to the membrane. At the membrane, they associate with TGF-β activated kinase 1 (TAK1), and TAB1 and TAB2, which are TAK1 binding proteins. Following that, the formed complexes migrate back to the cytosol [31]. After this point, there are 2 pathways for the signal propagation.

One of the pathways is to follow MAPK - JNK – ERK while the other one is IKK - IκB – NF-κβ. Both pathway routes are essential for apoptosis, cell survival, and inflammation.

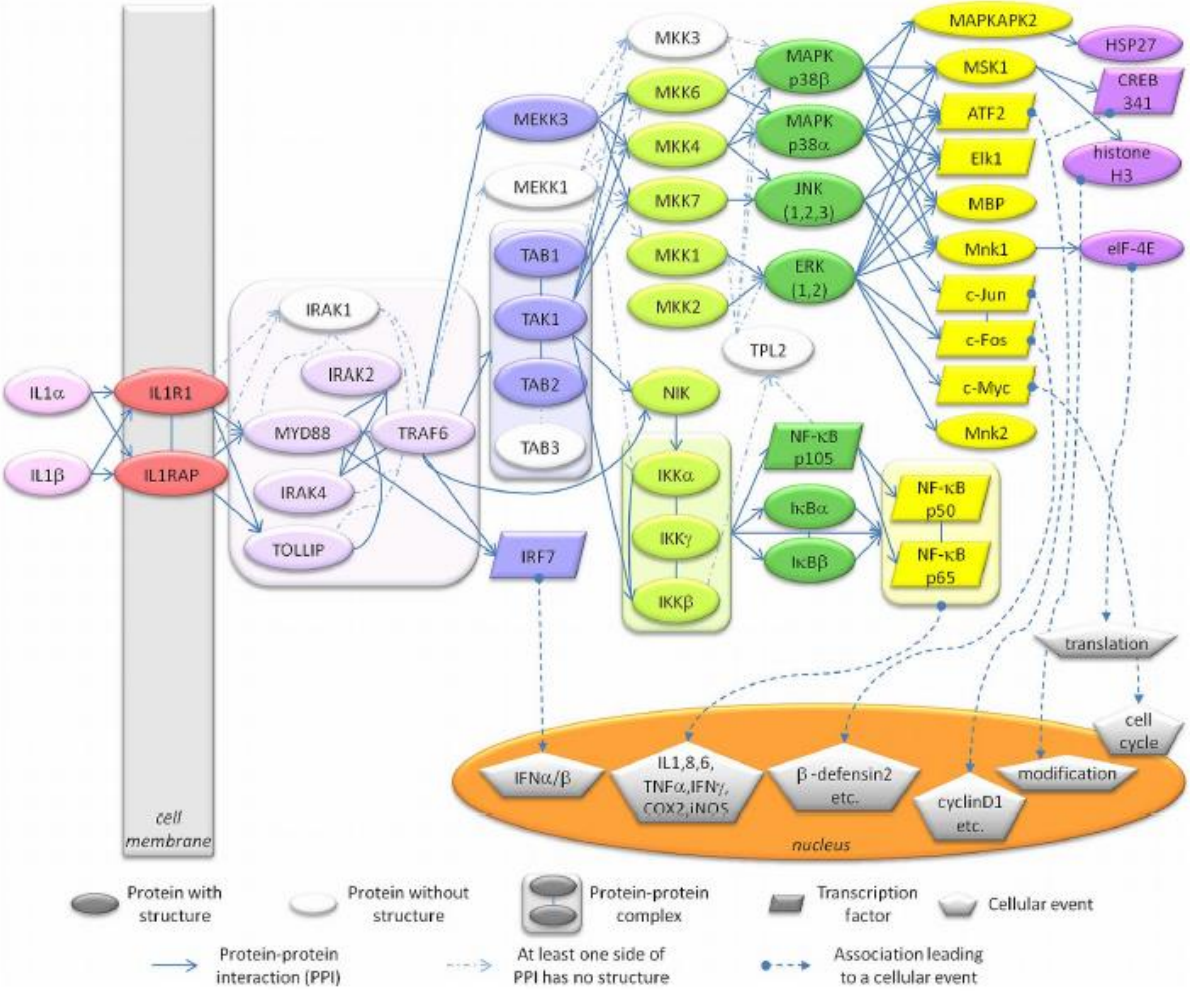

Figure 1.6: IL-1 signaling pathway. Adapted from [25].

When the levels of IL-1β was detected in the patients with rheumatoid arthritis, there is an increased production of IL-1β, or the natural inhibitor ligand of IL-1RI, which is IL-1 receptor antagonist (IL-1RA) molecule. Increased levels of either IL-1β or IL-1RA creates an imbalance in the pair of IL-1RA bound IL-1β complexes [32]. In addition, Rooney *et al.*, demonstrated that elevated levels of IL-1β is correlated with disease severity in *Rheumatoid Arthritis* patients [33]. Anakinra, or Kineret™ has been developed to treat *Rheumatoid Arthritis* patients by targeting IL-1β. Kineret™ has been produced by Amgen as a recombinant protein in the same structure of IL-1 Receptor Antagonist to block the increased activity of IL-1β [31]. As Anakinra is the leading therapeutic agent among IL-1 antagonists due to its safety profile, and long half-life, the main goal of this

thesis is to design peptide drugs, which have promising features, that may provide an alternative to Anakinra by taking Anakinra as the reference molecule. In order to reach this objective, we first designed peptides that mimicked the binding of Anakinra to IL-1RI in computer simulations. Detailed information about computational design of the IL-1R inhibitory peptides and experimental studies are given in the following chapters of the thesis, while the next section summarizes computational methods used in peptide design.

1.2.1 Computational Design of Peptide Ligands

As mentioned previously, drug discovery process starts with drug target selection and lead discovery. These two phases form the basis of structure-based drug design. After, the obtained ligand molecule is tested *in vitro* and *in vivo* both to confirm the results of *in silico* design, and to acquire information about its safety, efficacy, and other features such as its side effects [34, 35].

For *in silico* design, high-resolution structures of the target molecule which can be obtained from Protein Data Bank (PDB) are necessary [36]. Crystal structures are necessary to understand the structure of the protein, and its interactions with natural ligands. In order to design a peptide to target a protein, there are three approaches which can be followed. The first approach includes using a protein structure with a peptide ligand that the interactions within the complex are analyzed. According to the analysis, mutations are performed to increase peptide's specificity. In this case, the peptide backbone is kept fixed [37]. Subsequently, similar to the first approach, a structure with a peptide ligand is used and analyzed by homology modeling. However, during the homology modelling, the peptide backbone is flexible. Hence, this uses different domain – peptide modeling conformations to generate different structures which can be then superimposed on the target structure [38]. The last approach uses the structure of the ligand binding site and involves designing a peptide sequence for this binding site [39]. The strength of the interactions between the target and the ligand determines the binding affinity, which is the key parameter for *in silico* peptide design. If the approach does not include superimposing the ligand to target molecule, molecular docking simulations are performed to determine the best peptide sequence for that binding site of the target according to the binding affinity values of the poses. Steered Molecular Dynamics simulations also provide further analysis of the unbinding mechanism and binding affinity of the designed peptide to the target by applying rupture force [34]. Additionally, to analyze the designed ligand conformationally Molecular Dynamics simulations are

performed [35, 40]. Molecular Dynamics simulations show how the conformation of a ligand and/or receptor changes upon binding in the presence of water.

Chapter 2 :
IN SILICO DESIGN OF IL1β – RECEPTOR INHIBITOR PEPTIDES

The primary goal of this thesis is to design peptide lead molecules to treat IL1β-Receptor. As it was mentioned in Chapter 1, Anakinra has almost the same structure of the natural antagonist molecule of IL1-Receptor with only a Methionine addition at the amino terminal [41]. Anakinra is produced as a recombinant protein with trade name of Kineret™ and it is used as the antagonist of IL1-Receptor for the treatment of *Rheumatoid Arthritis* [31].

For the design of IL1β-Receptor inhibitor peptides, the structure of its natural ligand was used. The PDB ID of the structure of the complex used is 1IRA. Figure 2.1 shows the three dimensional structure of Anakinra bound to IL1β-Receptor [42].

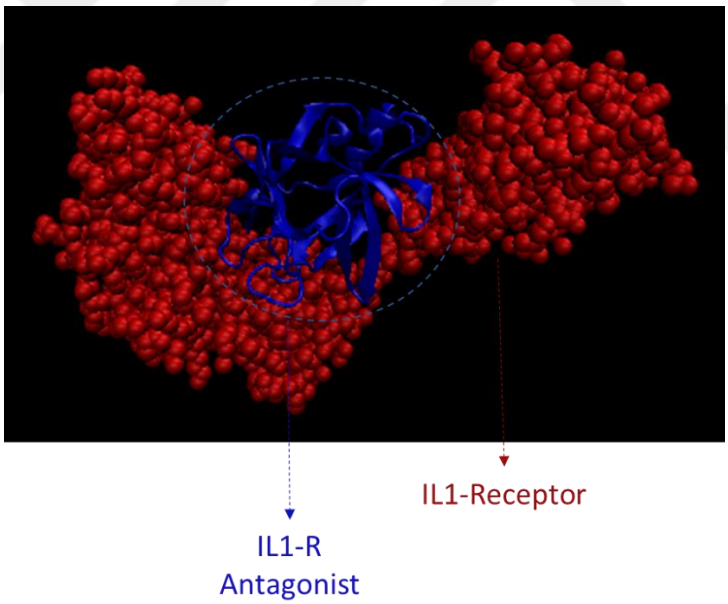

Figure 2.1:Three-dimensional structure of Anakinra – IL1β-Receptor complex. Obtained from the structure with PDB ID 1IRA.

As the aim of our research is to restore the balance between IL1-R antagonist and IL1-β by binding peptides instead of Anakinra, different peptide sequences were designed. For the design of the peptides, human native IL1-R Antagonist protein was used as the basis structure to understand the interactions between the inhibitor and the receptor. First, the important binding regions of IL1-Receptor to the antagonist molecule as well as

the residues of the antagonist molecule on the binding region were determined as shown in Figure 2.2.

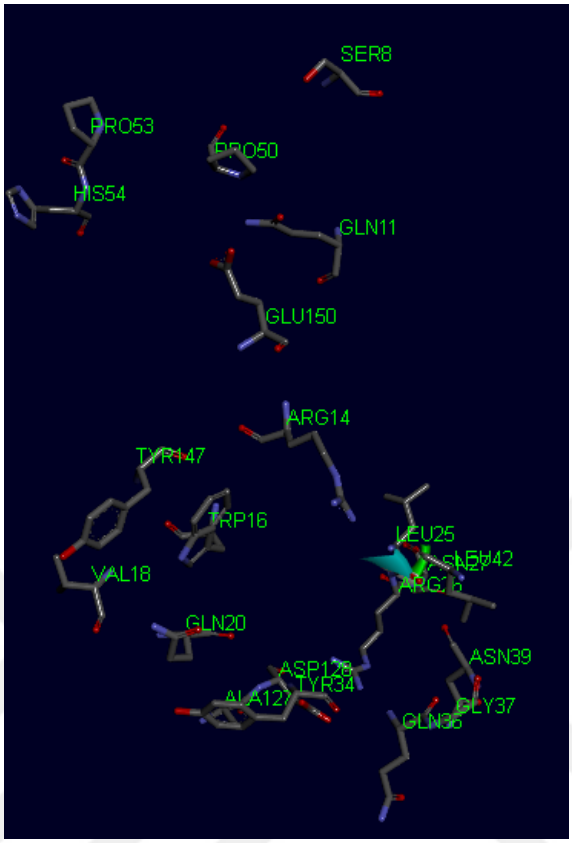

Figure 2.2: The important binding residues on the antagonist molecule.

In order to determine the important binding residues which are necessary for the peptide backbone to bind IL1-Receptor, the type and number of noncovalent interactions were considered. After the determination of important binding residues on the antagonist molecule, the close residues were put together to form a variety of peptide sequences in the binding region of the receptor protein. As a result, 9 different peptide sequences were obtained. The designed sequences are given in the Table 2.1.

Table 2.1:The peptide sequences designed from the important binding regions of IL1-R Antagonist and IL1-Receptor Proteins.

	The Amino Acid Sequences of The Peptides
Peptide A	Ser – Lys – Met – Gln – Ala – Phe
Peptide B	Ser – Gln – Arg – Trp – Val – Gln
Peptide C	Arg – Ile – Trp – Asp – Val – Asn – Gln
Peptide D	Leu – Arg – Asn
Peptide E	Tyr – Leu – Gln – Gly – Pro – Asn – Val – Asn – Leu
Peptide F	Pro – Ile – Glu – Pro – His
Peptide G	Pro – His
Peptide H	Ala – Asp
Peptide I	Tyr – Phe – Gln – Glu – Asp

In order to evaluate the binding affinities of these peptides, molecular docking simulations were performed by using GOLD – Protein Ligand Docking Software [43].

Molecular docking is an algorithm-based tool which aims to predict the predominant poses of a ligand bound to protein. It generates different conformations, and orientations of the ligand, referred to as poses of the ligand, to predict the best binding mode of the ligand in which the complex is stable. In short, molecular docking methods aim to predict the three dimensional structure of the complex to understand and analyze the interactions, and the binding mechanism [44]. Figure 2.3 shows the general scheme of molecular docking techniques.

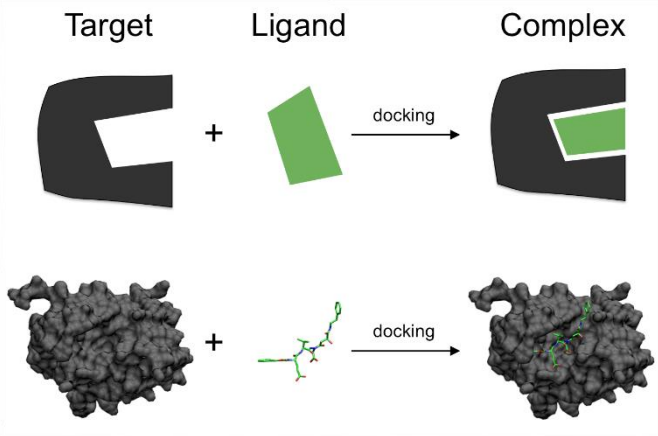

Figure 2.3: The molecular docking algorithms effectively search the conformational space to predict the binding modes of the ligand. Adapted from [45].

GOLD software uses Genetic Algorithm as the search algorithm. Genetic Algorithm was inspired by Darwin's natural selection theory. Each poses of the ligand with respect to the protein are the individuals and the individuals form the population. Each of the individuals has three degrees of freedom, which are the conformation, orientation, and translation movements of the ligand. The degrees of freedom are represented by chromosomes and the real valued genes constituting chromosomes represent the individuals. They are evaluated by using fitness score functions. According to the scores of the individuals, the genes are mutated, crossover, and reproduce to produce new individuals to find out the best position of the ligand [45].

Therefore, molecular docking simulations of the peptide sequences designed from the important binding regions of IL1-R Antagonist and IL1-Receptor proteins were performed by using GOLD software. As the aim of this step is to find out the sequence which shows the best binding affinity to the receptor and then the selected peptide would be designed further, the bonds between the residues were fixed. In this way, only the translational and orientational degrees of freedom are taken into the account to computationally save time. Additionally, the docking was performed for 100 runs for each of the sequence. The parameters such as binding region coordinates, and search space radius are as the following:

Binding Region Coordinates: (X: 41.50, Y: 9.36, Z: 49.83)

Search Space Radius: 17 Å

Molecular docking simulations evaluate the total interaction energy between the protein and the ligand, and the intramolecular ligand energy using fitness score functions.

In this study, Chemscore fitness function was used. The values obtained by Chemscore function were converted into the change in Gibbs Free energy so that binding affinities of the peptide sequences could be computed. The conversion to find out the change in Gibbs Free Energy from Chemscore values obtained as the results of docking is given in Eq. 2.1 [46]:

$$\Delta G_{binding} = 0.9111 \times (Chemscore) - 8.1901 \quad (2.1)$$

From the Equation 1.1, binding affinity can be found as; $K_D = \exp\left(-{}^{\Delta G_{binding}}/_{RT}\right)$.

R is ideal gas constant and taken as 0.008314 $kJ/mol \times K^{-1}$. T corresponds to the temperature and has the unit of Kelvin. It is taken as the body temperature, which is 36.5°C and corresponds to 309.5 K. Using this information, $\Delta G_{binding}$ and the binding affinity values of the peptide sequences were calculated. The highest binding affinity result for each of the sequences are presented in Table 2.2.

Table 2.2: $\Delta G_{binding}$ and the binding affinity values of the designed peptide sequences as result of molecular docking studies.

	ΔG (kJ/mol)	**K_D (M)**
Peptide A	37.28	5.03×10^{-7}
Peptide B	36.07	7.67×10^{-7}
Peptide C	42.03	7.91×10^{-8}
Peptide D	33.62	2.08×10^{-6}
Peptide E	36.34	7.22×10^{-7}
Peptide F	33.16	2.49×10^{-6}
Peptide G	28.02	1.84×10^{-5}
Peptide H	34.06	1.76×10^{-6}
Peptide I	36.18	7.67×10^{-7}

As binding affinity is inversely proportional to K_D value, Peptide C, which has the sequence of Arg – Ile – Trp – Asp – Val – Asn – Gln, has showed the highest binding

affinity to IL1-R among the 9 different sequences. It was concluded that Peptide C could be used for both further design simulations and the synthesis of the peptide. To confirm that Peptide C has high binding affinity to IL-1 Receptor, docking was repeated without fixing the bonds of the peptide for 999 runs. The result from the docking confirms Peptide C is a good candidate that could be used for further studies.

The lowest K_D value which gives the highest binding affinity, the average of K_D values of 999 runs, and the highest K_D value which gives the lowest binding affinity are presented in Table 2.3. Figure 2.4 shows the pose of the peptide C bound to IL1-Receptor with the highest binding affinity.

Table 2.3: Summary of K_D values as the result of 999 run docking of Peptide C.

	K_D (M)
The pose with the highest binding affinity	$0.89x10^{-8}$
Average of the poses	0.03
The pose with the lowest binding affinity	1.64

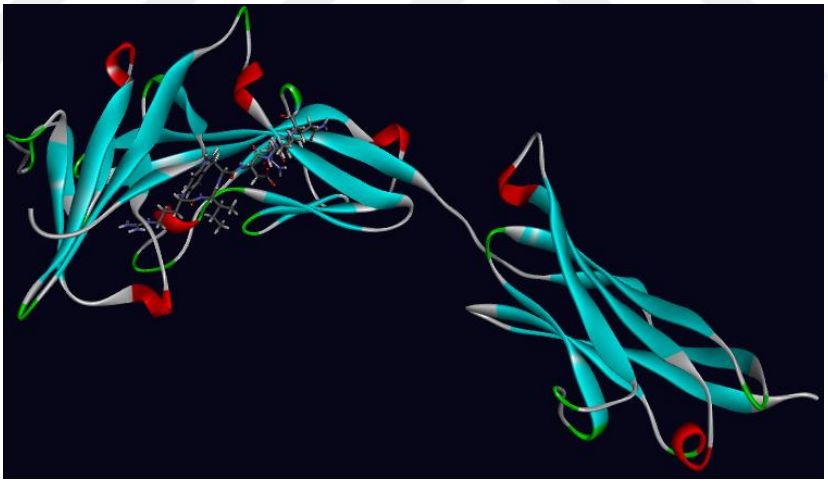

Figure 2.4: The pose of Peptide C bound to IL-1 Receptor, which has K_D = 0.89 x 10-8 M.

In order to increase conformational stability of Peptide C, N- terminus and C-terminus modifications were applied that we have obtained Second Generation Peptides. The modifications were performed by the insertion of 20 amino acids separately to each terminus of the peptide. At first, N- terminus modification was performed that 20 amino acids separately were added to the amino group of Peptide C. Thus, there were 20 different ligands which have the sequence of Peptide C in common and one different

amino acid at the N- terminus. The molecular docking simulations were performed for 999 runs for each of the 20 ligands without fixing the bonds. Phe insertion at amino group has showed the highest binding affinity. Summary of the results of Phe modified peptide are presented in Table 2.4:

Table 2.4: Summary of K_D values as the result of 999 run docking of Phe inserted peptide.

	K_D (M)
The pose with the highest binding affinity	$0.88x10^{-7}$
Average of the poses	0.0023
The pose with the lowest binding affinity	0.102

According to the results of the docking of Phe insertion at the amino group, the highest binding affinity got decreased one order of magnitude while the average value of the K_D decreased one order of magnitude showing better binding affinity in the average. The negative control pose which has the highest K_D also proves this situation that the value got closer to the average. Hence, the binding affinity values of the peptide have entered a smaller range which demonstrates the increased conformational stability of the peptide. Therefore, the sequence of N- terminus modified peptide is Phe – Arg – Ile – Trp – Asp – Val – Asn – Gln. Figure 2.5 shows the pose of the N- terminus modified peptide bound to IL1-Receptor with the highest binding affinity.

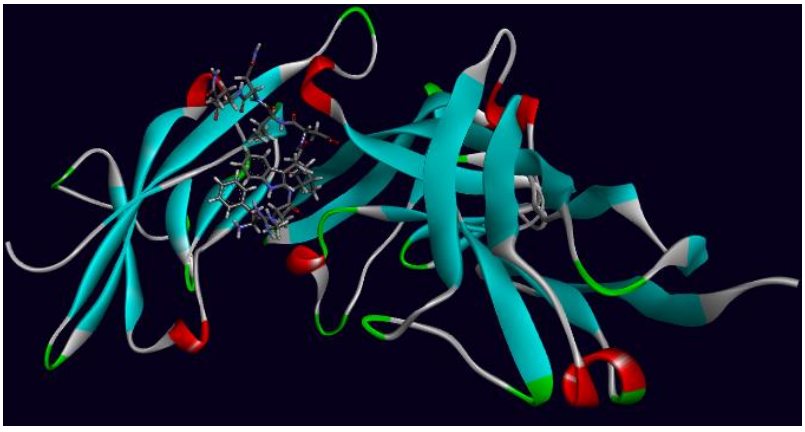

Figure 2.5: The pose of 8-amino acid sequence peptide bound to IL-1 Receptor, which has K_D of 0.88x10-7 M.

In a similar manner to N- terminus modification of the Peptide C, C- terminus modification was also applied. In this instance, 20 amino acids were separately added to

the carboxyl group of N- terminus modified 8- amino acid sequence peptide. Thus, another group of 20 different ligands were designed. They have the sequence of N-terminus modified peptide in common and one different amino acid at the C- terminus. The molecular docking simulations were performed for 999 runs for each of the 20 ligands without fixing the bonds. Lys insertion at carboxyl group of the N- terminus modified peptide has showed the highest binding affinity. Summary of the results of Lys modified peptide are given in Table 2.5:

Table 2.5: Summary of K_D values as the result of 999 run docking of Lys inserted peptide.

	K_D (M)
The pose with the highest binding affinity	$0.24x10^{-9}$
Average of the poses	0.0016
The pose with the lowest binding affinity	0.13

According to the results of the docking of Lys insertion at the carboxyl group of N-terminus modified peptide, the highest binding affinity increased one order of magnitude when compared with Peptide C and N- terminus modified peptide. The average value of the K_D is in the same range with the unmodified and N- terminus modified peptides. The negative control pose which has the highest K_D is also in the same range with only N-terminus modified 8- amino acid sequence peptide. Application of both N- and C-terminus modifications have generated a possibility of a pose of the ligand with stronger interactions, thereby binding better to inhibit IL1-Receptor. The sequence of N- and C-terminus modified peptide is Phe – Arg – Ile – Trp – Asp – Val – Asn – Gln – Lys. Figure 2.6 shows the pose of the both N- and C- termini modified peptide bound to IL1-Receptor with the highest binding affinity. In Figure 2.7, the images of the poses of peptide C and both N- and C- termini modified peptide are shown.

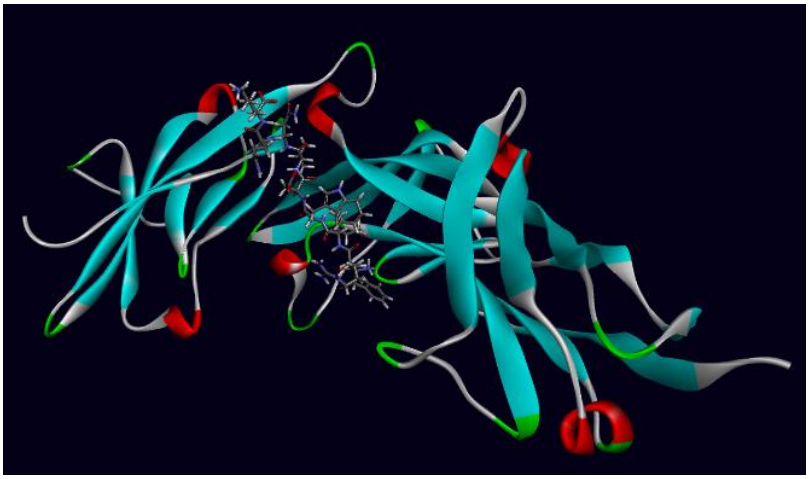

Figure 2.6: The pose of 9-amino acid sequence peptide bound to IL-1 Receptor, which has K_D of 0.24x10-9 M.

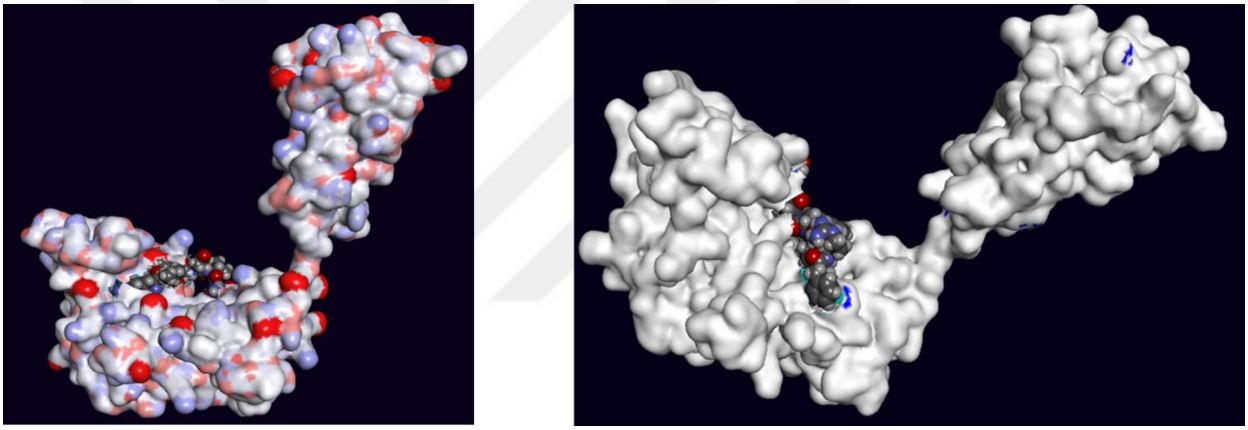

Figure 2.7: The image at left shows the pose of unmodified peptide to IL1-Receptor which has K_D of 0.89 x 10-8 M while the right side shows the pose of both N- and C-termini modified 9- amino acid sequence peptide with K_D of 0.24x10-9 M.

The unmodified 7- amino acid peptide sequence, which is Peptide C, binds to IL1-Receptor protein with 5 hydrogen bonds from the oxygen atoms of Ile and Gln on the peptide sequence. In addition to hydrogen bonds, there is also 1 hydrophobic interaction between Trp on the peptide and Lys on IL1-R. The application of N- terminus and C-terminus modifications bring new hydrophobic interactions for the binding of the peptides to IL1-R. In the case of N- terminus modified 8- amino acid sequence peptide has 3 hydrogen bonds and 10 hydrophobic interactions with the receptor. The residues which participate to hydrogen bonds changed to Trp and Arg residues of the peptide that they form the bonds with Lys, Tyr, and Phe residues of IL1-R. Moreover, both N- and C-termini modified 9- amino acid sequence peptide has showed the highest binding affinity

with 3 hydrogen bonds and 8 hydrophobic interactions. Instead of Arg, Asn participates to form hydrogen bonds that Trp and Asn are bonded to hydrogen atoms of Ile and Tyr residues of IL1-R. The hydrophobic interactions occur between Val and Ile residues of the peptide, and Val, Leu, Pro, and Phe residues of the protein. Figure 13 shows how well both N- and C- termini modified peptide fills the cavity when it is bound to the receptor. The decrease in K_D value also proves the situation quantitatively.

The molecular docking studies do not include water molecules in the simulation environment. However, the molecular binding events happen in the cells, which is composed of 70% of water molecules. Although molecular docking has become an essential tool to predict the free energy calculations upon the binding of a ligand to a protein, it neglects the presence of water which is the necessary molecule in the cells. Since water molecules move and rotate fast that they can form hydrogen bonds very easily. Thus, if a change occurs in a water molecule especially in the binding region of a protein, it does not only affect the neighboring molecules but also the surrounding layers that the effect due to a change in a water molecule results affecting the whole water network. Another simplification of docking studies is the lack of motion of the protein. There are also softwares which take both protein and the ligand as rigid molecules. Due to this reason we used GOLD for the docking studies. However, the lack of motion of protein still exists which decreases the reality adaptation of the generated poses. As a result, solvent neglection and the lack of motion of protein are the disadvantages of molecular docking simulations [47].

Since docking has low predictive power in structure-based drug design, Molecular Dynamics methods are also used to calculate free energy change upon binding. In MD simulations, the dynamics of both the ligand and the protein are considered and a cell environment is created by also solvating the complex. Therefore, they give more accurate values than docking methods, however, they are computationally expensive and MD simulations cannot be used for each target [48]. Steered Molecular Dynamics (SMD) has been developed by taking the basis of Atomic Force Microscopy (AFM) experiments to understand unfolding mechanism of the biomolecules [49]. SMD simulations are also used to study unbinding pathway of a ligand from its receptor that SMD has become a promising tool for structure based drug design due to its high predictive power and more efficiency than MD methods [48].

In order to investigate the binding affinity of a ligand to its receptor, as it is done in AFM experiments, an external force is applied in the environment of water molecules to pull the ligand from the receptor while also anchoring atoms in the receptor. The ligand is pulled from the receptor by applying rupture force to the pulling residues of the ligand through a spring. The general schematic illustration of SMD is shown in Figure 14. As it is known from the AFM experiments, the force applied to the ligand is proportional to the distance taken by the ligand. Figure 2.8 shows a general schematic for how SMD works.

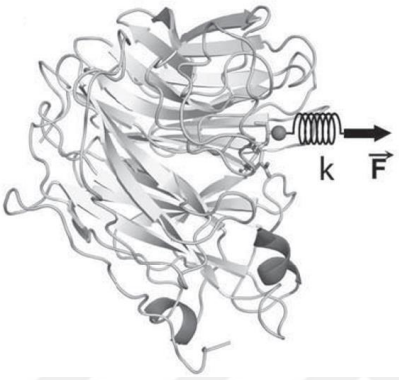

Figure 2.8: A general schematic illustration for how SMD works to understand unbinding pathway of a ligand from its receptor. Adapted from [48].

Under the force, total energy is calculated by using Equation 2.1 [50]:

$$e^{-\beta \Delta F} = \langle e^{-\beta W} \rangle \tag{2.1}$$

Total energy of the receptor – ligand complex is as the following [49]:

$$E = E_{receptor} + E_{ligand} + E_{receptor-ligand} + E_{force} \tag{2.2}$$

Consequently, in order to understand unbinding pathway and binding affinities of our designed peptides from IL1-Receptor in a more accurate way, we performed SMD simulations for unmodified sequence which is Peptide C, N- terminus modified 8- amino acid peptide, and both N- and C- termini modified 9- amino acid peptide. The simulations were performed applying rupture force to the whole residues of the peptides while anchoring Pro146 residue on the receptor. The purpose of pulling all residues of the peptide is to apply an external force to each amino acid of the peptide. When SMD simulations were performed by only pulling 1 or 3 residues, the shape of the peptide became deformed and it sticked to the protein. In order to prevent the deformation of the peptide, the rupture force was applied to whole residues of the peptide, thereby applying equal amount of force to each amino acid. As a result, the shape of the peptide stays in its

original form and unbinds from the protein. As the parameters of SMD simulations, periodic boundary conditions were applied, and counter-ions were added to neutralize the solvated system. Constant velocity pulling method was applied with velocity of 35 Å/ns. SMD simulations were performed after minimization, equilibration, and annealing steps for 500000 timesteps and therefore, totally for 1 ns.

The SMD results for peptide C are given in Figure 2.9.

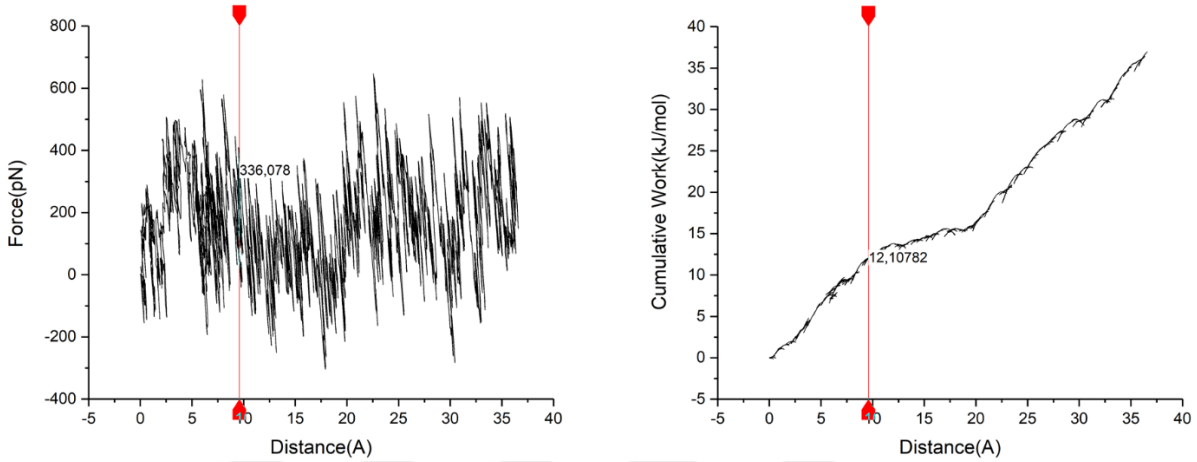

Figure 2.9: As a result of SMD simulations, (a) force vs. distance and (b) work vs. distance graphs were obtained for the unmodified peptide.

SMD results of 7- amino acid peptide – IL1-Receptor complex show that the strength of the interactions decreased when the simulation was performed in the presence of water. When the graphs and trajectory movie of the complex are analyzed, it is seen that there are 2 hydrogen bonds at the beginning, whereas the expected number of hydrogen bonds is 5 due to the docking results. From SMD results, only Gln participates in hydrogen bonding at 4.0 Å cutoff and most of the hydrophobic interactions disappeared because the water molecules interfered which results in increase of the distance between the bonds. The moment that peptide unbounds from the protein was determined according to the analysis of the hydrogen bonds. The frame in which hydrogen bonds break off completely between the peptide and protein was taken to determine the moment which the peptide unbounds. According to the analysis criteria mentioned, the peptide unbounds from the protein at 60th frame which corresponds to 0.24 ns and 9.6 Å in distance. However, as it is seen from the graphs, the system continues to apply an external force to the peptide even after it unbounded from the protein. Since, the peptide continues to form hydrogen bonds with another site of the binding region of the protein that the system continues to apply the force due to these interactions.

The cumulative work done on the peptide at the point the peptide unbounds gives the K_D value of the peptide from Equation 1.1.

$$K_D = \exp\left(-{}^{\Delta G_{binding}}/_{RT}\right)$$

R=0.008314 $kJ/mol \times K^{-1}$

T=309.5 K

As $\Delta G_{binding}$ equals to cumulative work value at the distance the peptide unbounds from the ligand, $\Delta G_{binding}$ of the 7- amino acid peptide is 12.10782 kJ/mol. Substituting the values into the equation, K_D value is found as 0.0090 M. If the result is compared with the docking results, it is not compatible with the binding affinity of the best pose. However, the result obtained from SMD is compatible with the average of the docking results. Water molecules interference, which is an important and realistic aspect of SMD, causes the peptide to move away from the position obtained from the docking thereby, causing the loss of some of the interactions. Hence, the peptide behaves and gives the binding affinity in the range of average values of the docking.

The SMD results for N- terminus modified peptide are given in Figure 2.10.

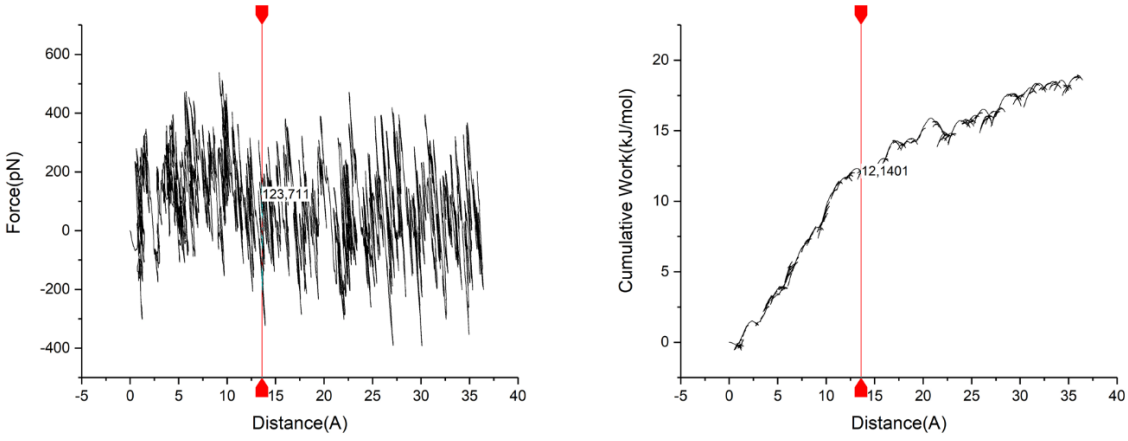

Figure 2.10: As a result of SMD simulations, (a) force vs. distance and (b) work vs. distance graphs were obtained for the N- terminus modified peptide.

The results for 8- amino acid peptide – IL1-Receptor complex show that the strength of the interactions again decreased when the simulation was performed in the presence of water. When the graphs and trajectory movie of the complex are analyzed, it is seen that there are 2 hydrogen bonds at the beginning, whereas the expected number of

hydrogen bonds is 3 due to the docking results. Arg and Trp participate in hydrogen bonding at 4.0 Å cutoff and most of the hydrophobic interactions disappeared due to the water molecules interference which results in increase of the distance between the bonds. The moment that peptide unbounds from the protein was again determined according to the split of hydrogen bonds between the peptide and the protein. Thus, the peptide unbounds from the protein at 85th frame which corresponds to 0.34 ns and 13.6 Å in distance. However, as it is seen from the graphs, the system continues to apply an external force to the peptide even after it unbounded from the protein. Since, the peptide continues to form hydrogen bonds with water that the system continues to apply the force.

The cumulative work done on the peptide at the point the peptide unbounds gives the K_D value of the peptide from Equation 1.1.

$$K_D = \exp\left(-{}^{\Delta G_{binding}}/_{RT}\right)$$

R=0.008314 $kJ/mol \times K^{-1}$

T=309.5 K

As $\Delta G_{binding}$ equals to cumulative work value at the distance the peptide unbounded from the ligand, $\Delta G_{binding}$ of the 8- amino acid peptide is 12.1401 kJ/mol. Substituting the values into the equation, K_D value is found as 0.0089 M. If the result is compared with the docking results, it is not compatible with the binding affinity of the best pose. However, the result obtained from SMD is compatible with the average of the docking results. Water molecules interference, which is an important and realistic aspect of SMD, causes the peptide to move away from the position obtained from the docking thereby, causing the loss of some of the interactions. Hence, the peptide behaves and gives the binding affinity in the range of average values of the docking.

The SMD results for both N- and C- termini modified peptide are given in Figure 2.11.

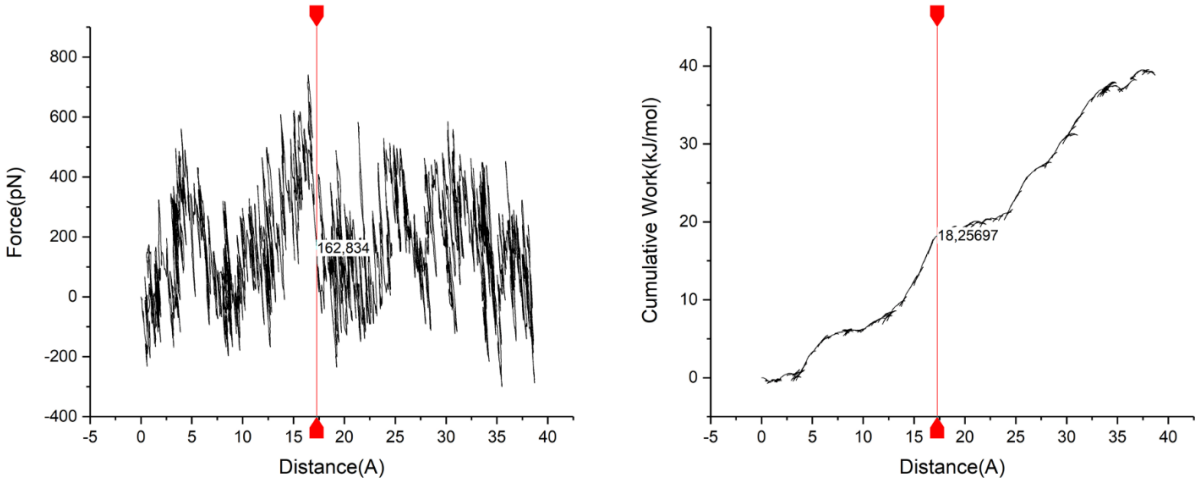

Figure 2.11: As a result of SMD simulations, (a) force vs. distance and (b) work vs. distance graphs were obtained for the both N- and C- termini modified peptide.

SMD results of 9- amino acid peptide – IL1-Receptor complex shows that the strength of the interactions decreased when the simulation was performed in the presence of water. When the graphs and trajectory movie of the complex are analyzed, it is seen that there are 2 hydrogen bonds at the beginning, whereas the expected number of hydrogen bonds is 3 due to the docking results. From SMD results, Asn and Trp residues participate in hydrogen bonding at 4.0 Å cutoff and most of the hydrophobic interactions disappeared because the water molecules interfered which results in increase of the distance between the bonds. Using the same analysis criteria to determine the moment the peptide unbounds from the protein at 108th frame which corresponds to 0.432 ns and 17.28 Å in distance. However, as it is seen from the graphs, the system continues to apply an external force to the peptide even after it unbounded from the protein. Since, the peptide forms hydrogen bonds between water molecules that the system applies the force to remove those interactions. The cumulative work done on the peptide at the point the peptide unbounds gives the K_D value of the peptide from Equation 1.1.

$$K_D = \exp\left(-{}^{\Delta G_{binding}}/_{RT}\right)$$

R=0.008314 $kJ/mol \times K^{-1}$

T=309.5 K

As $\Delta G_{binding}$ equals to cumulative work value at the distance the peptide unbounds from the ligand, $\Delta G_{binding}$ of the 9- amino acid peptide is 18.25697 kJ/mol. Substituting the values into the equation, K_D value is found as 0.00083 M. If the result is

compared with the docking results, it is not compatible with the binding affinity of the best pose. However, the result obtained from SMD is compatible with the average of the docking results. Water molecules interference, which is an important and realistic aspect of SMD, causes the peptide to move away from the position obtained from the docking thereby, causing the loss of some of the interactions. Hence, the peptide behaves and gives the binding affinity in the range of average values of the docking. Table 2.6 summarizes the K_D results of the three peptides.

Table 2.6: The summary of binding affinity results of the peptides.

	K_D value of the best pose of the docking (M)	K_D value of the average of the docking results (M)	K_D value obtained from SMD (M)	K_D value of the pose with least binding affinity obtained from docking
Unmodified peptide	$0.89x10^{-8}$	**0.03**	**0.0090**	1.64
N- terminus modified peptide	$0.88x10^{-7}$	**0.0023**	**0.0089**	0.102
N- & C- termini modified peptide	$0.24x10^{-9}$	**0.0016**	**0.00083**	0.13

Due to the interference of the water molecules between protein and the peptide complexes, binding affinity values of the peptides are not close to the values obtained from the best poses of the docking studies. However, SMD simulations are closer to mimic realistic environment due to the presence of water molecules and taking the protein flexibility into account. Additionally, the compatibility of the SMD results with the average values obtained from docking simulations shows that three of the designed peptides tend to behave and bind to the protein similarly as the poses which have K_D values around the average from the docking results. Furthermore, as the binding affinity increased with N- and, both N- and C- termini modifications compared to unmodified 7-

amino acid peptide in the docking studies, they also showed the same increase pattern in SMD simulations.

Chapter 3 :
THE SYNTHESIS AND ANALYSIS OF *IN SILICO* DESIGNED PEPTIDES

In this chapter, the peptide sequences obtained from computational studies were synthesized using Fmoc-based Solid-Phase Peptide Synthesis (SPPS) and the obtained products were evaluated by High Performance Liquid Chromatography. Hence, the following sections explain the methods used for the synthesis and analysis of the unmodified, N- terminally modified, and both N- and C- terminally modified peptide.

3.1 Solid-Phase Peptide Synthesis (SPPS)

The three peptide sequences obtained from computational design were synthesized by SPPS for use in further studies such as surface plasmon resonance (SPR) analysis.

SPPS was first developed by R. B. Merrifield in 1962. This is a method to chemically synthesize a peptide by the consecutive addition of amino acids. The idea behind the method is to synthesize the peptide on an insoluble solid particle, which is referred to as resin, and then to remove the desired sequence from the support. C- terminal amino acid in the peptide is covalently bonded to the resin and when the synthesis is complete, the peptide is cleaved by breaking the covalent bond between the resin and the peptide [51]. In addition, understanding of peptide bond formation is important for following the SPPS protocol.

Peptide bond is formed as a result of a condensation reaction which occurs between carboxyl group of one amino acid and α-amino group of other amino acid. Figure 3.1 shows the peptide bond formation in a dipeptide [52].

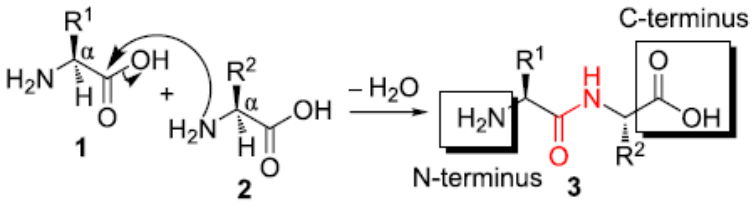

Figure 3.1:The reaction between the amino group and the carboxyl group leads a peptide bond formation, shown in red. Adapted from [52].

In Solid-Phase Peptide Synthesis, as the peptide is conjugated to the resin from its C- terminus, the synthesis is performed from C to N direction, whereas typically the sequence is written in the N to C direction. In order to obtain the desired sequence without forming any undesired side products, one amino acid at a time is attached. In order to prevent undesired products formation, amino groups of the amino acids are protected. However, protection of only the amino group is not enough to prevent unwanted reactions and side products. Since; amino acids also have other reactive side chains, they must be protected as well [53, 54]. The protecting groups can be either *tert*-butyloxycarbonyl (Boc) or 9-fluorenylmethoxycarbonyl (Fmoc). The SPPS was first developed using Boc-protected amino acids. As this group is labile to acids, the synthesis requires the use of toxic hydrogen fluoride [55]. In 1972, Carpino developed 9-fluorenylmethoxycarbonyl (Fmoc) as the amino protecting group [56]. Fmoc is labile to bases that it doesn't require the use of strong acids, hence, the synthesis can be performed under milder reaction conditions. Fmoc has other advantageous properties as well. For instance, Fmoc-protected amino acids preparation is easy that they can be produced in high yields, they are stable, and they are soluble in all solvents used for SPPS. Fmoc group is also stable in trifluoroacetic acid (TFA), which is used for the cleavage of the peptide from the resin [55]. Due to its advantageous properties for SPPS, Fmoc has become the standard protecting group for peptide synthesis adapting SPPS protocol. Figure 3.2 shows the chemical structure of Fmoc group while Figure 3.3 shows the use of a protecting group in an amino acid.

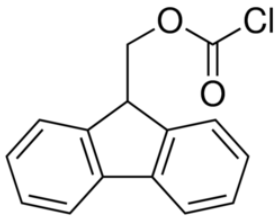

Figure 3.2: The chemical structure of 9- fluorenylmethoxycarbonyl chloride (Fmoc chloride). Adapted from [57].

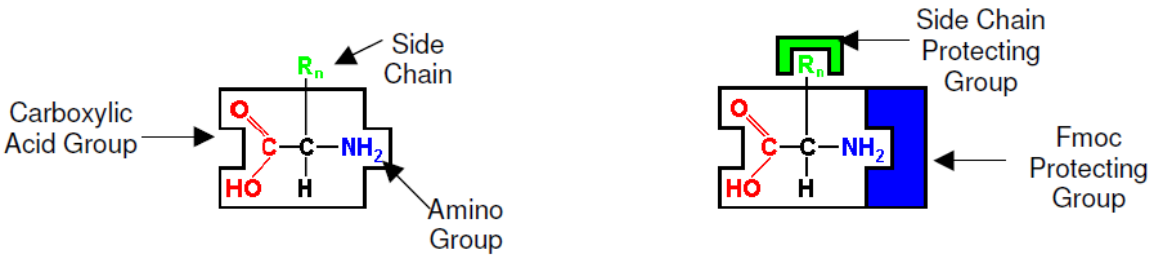

Figure 3.3: In order to conjugate one amino acid at a time, amino and the reactive side chain groups must be protected. As the peptide grows from C to N direction, carboxyl group must be activated. This figure shows the difference of the amino acid which will be used in SPPS. Adapted from [58].

The resin on which the desired sequence obtained must possess some features such as [58]:

- Insolubility in all solvents used in SPPS,
- Chemically and physically resistance,
- Low crosslinking, and good swelling.

A resin has its functional site, which is referred to as core, and each of the cores have linkers attached on them. According to the material of the core, the type of resin typically used in SPPS is polystyrene resin. In addition, the linkers determine the properties of the final product and which chemicals would be used for the reaction. Wang and Rink linkers are the most widely used ones for Fmoc-based SPPS. Wang resin produces peptide acid (COOH) as the final product and conjugates the first amino acid to the resin with a different protocol from the protocol of subsequent amino acids coupling to each other. This makes SPPS protocol more complicated and time consuming. Rink resin produces peptide amide ($CONH_2$) and the protocol to conjugate the first amino acid to the resin is the same as the subsequent ones that the use of Rink resin is more preferable [53, 54]. Figure 3.4 illustrates the resin and how a peptide is attached to the resin by linkers.

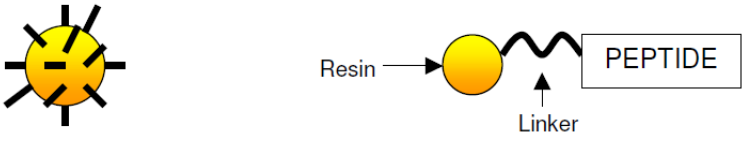

Figure 3.4: Peptides are attached to the resin bead by linkers. Adapted from [58].

Therefore, the sequences we obtained from the computational design studies for 7- amino acid, 8-amino acid, and 9- amino acid peptides were synthesized using Fmoc-based SPPS methods. The protocol followed during the synthesis is given below [59].

3.1.1 Fmoc Solid Phase Peptide Synthesis Protocol

1. Resin Preparation

In our protocol, due to its advantageous properties, Rink resin was used. The linkers of Rink resin have Fmoc groups. For the first amino acid to attach the resin, Fmoc group on the linker must be removed. Thus, deprotection of Fmoc group initiates the synthesis protocol. Prior to the first deprotection step, the resin should be swollen in dimethylformamide (DMF). SPPS is performed using polar aprotic solvents and the reactions rely on the swelling properties of the resin in these solvents. DMF is a polar aprotic solvent that is not capable of hydrogen bonding. Thus, swelling properties enable the coupling in SPPS. In other words, good swelling properties increase the accessibility of the coupling sites on the resin beads [60]. Due to the swelling properties of Rink resin, DMF is the most widely used solvent for the SPPS reactions [61]. The SPPS reactions were carried out in a 10-ml peptide synthesis syringe with a glass wool filter. In the beginning, the resin was loaded into the syringe and then 8 ml of DMF was added to the resin for swelling and stirred for 5 minutes. The resin loading and mixing steps were repeated 3 times.

2. Deprotection Cycle

Deprotection of Fmoc group occurs using 8ml of 20% (v/v) piperidine in DMF. Piperidine solution in DMF is so basic and strongly nucleophilic that it can even access hindered sites [62]. The first step was to add 8 ml of piperidine solution in DMF to the resin and stir for 1 minute. When the 1-minute stirring was complete, the solution was discharged, and the syringe was again loaded with 8 ml of 20% (v/v) piperidine in DMF. It was stirred for 10 minutes. In this way, all the Fmoc groups can be deprotected. Once a new solution is used in SPPS, the principle is to wash the resin with DMF to remove remaining molecules from the previous solution, thereby preventing unwanted side reactions. Hence, after the deprotection step was complete, the resin was washed with 8 ml of DMF for 5 minutes. This washing step was repeated 5 times. Figure 3.5 demonstrates the deprotection cycle.

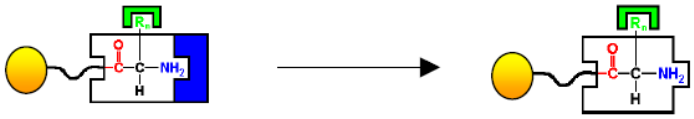

Figure 3.5: The illustration presents the Fmoc removal from the N- terminus of the resin or peptide. Adapted from [58].

3. Activation and Coupling Cycle

As the carboxyl group of an amino acid will attach to the resin from N- terminus, it must be activated. For this purpose, a basic activator solution is used. In our experiments, HBTU was used as the activator. The base was Diisopropylethylamine (DIEA). In addition, the amino acid was dissolved in DMF. Thus, 1 mmol Fmoc-protected amino acid and 1 mmol HBTU were dissolved in 8 ml DMF to activate the carboxyl group of the amino acid. After mixing the solution, 2 mmol DIEA was added into the activated amino acid solution. This mixture was then added to the resin inside the syringe. It was stirred for 2 hours, then the solvent was discharged. Once the coupling step was complete, the resin was washed with 8.0 ml of DMF for 30 seconds. This washing step was repeated 5 times. Figure 3.6 shows the coupling of an amino acid schematically.

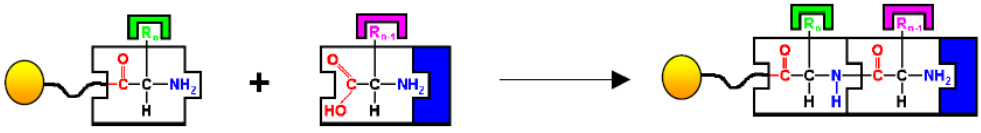

Figure 3.6: This illustration presents the coupling of an amino acid to another amino acid attached to the resin. Adapted from [58].

4. Capping Cycle

In order to prevent the formation of undesired peptides, a capping reaction was performed. Capping reaction blocks the unreacted amino groups. For this purpose, 10 ml of 10% (v/v) acetic anhydride in DMF solution was used. After the washing step, acetic anhydride in DMF solution was added to the resin and stirred for 30 minutes. Once the capping step was complete, the resin was washed with 8 ml of DMF for 30 seconds. This washing step is repeated 3 times.

After each coupling and capping steps, Kaiser test, which is a qualitative method to analyze the presence of free primary amine groups, was performed. If the coupling was successful, starting from deprotection cycle coupling and capping reactions of subsequent

amino acids were performed. After the desired peptide sequence was synthesized, it must be cleaved from the resin as a final step.

5. Final Deprotection and Cleavage Cycle

When the coupling of the last amino acid is complete, Fmoc protecting group remains in the amino acid. Before releasing the peptide from the resin, the last Fmoc group must be removed by using the same steps in the deprotection cycle. After the final deprotection was done, the synthesized peptide was cleaved from the resin using the mixture of trifluoroacetic acid (TFA)/phenol/distilled water/triisopropylsilane(TIPS) (88:5:5:2 v/v/v/v). The mixture was added to the resin and stirred for 3 hours. The discharged solution was precipitated using cold anhydrous ether, and the precipitate comprises the synthesized peptides. Then, the peptides were dissolved in acetonitrile-water mixture to be lyophilized and stored.

In summary, the computationally designed 3 peptide sequences were synthesized using Fmoc-based Solid Phase Peptide Synthesis (SPPS) protocol. However, during the synthesis, side products, including smaller peptides may have been produced as well. In order to purify the desired peptides, High Performance Liquid Chromatography (HPLC) is used. Prior to purification, we tested our synthesized peptides analytically to find out the exact retention time of the desired product. In the next section of this chapter, HPLC methodology and our HPLC results are given.

3.2 Peptide Analysis by High Performance Liquid Chromatography

High performance liquid chromatography, abbreviated as HPLC, is an instrumental method to be used in both analytical and preparative scales/experiments [63]. As in other liquid chromatography types, mobile and stationary phases are present. In HPLC, the pump pressure, which moves the mobile phase through a chromatography column, can go up to 500 bar. Other types of liquid chromatography columns can limit only up to 30 – 50 bar. This unique feature of HPLC makes the analysis much faster and allows for the use of a much smaller resin particles, thereby increasing the surface area for interactions between the sample and the resin. Due to this unique feature of HPLC, it can be used to separate and identify small molecules even in trace concentrations such as parts per trillion [63, 64]. A general flow scheme of HPLC is shown in Figure 3.7.

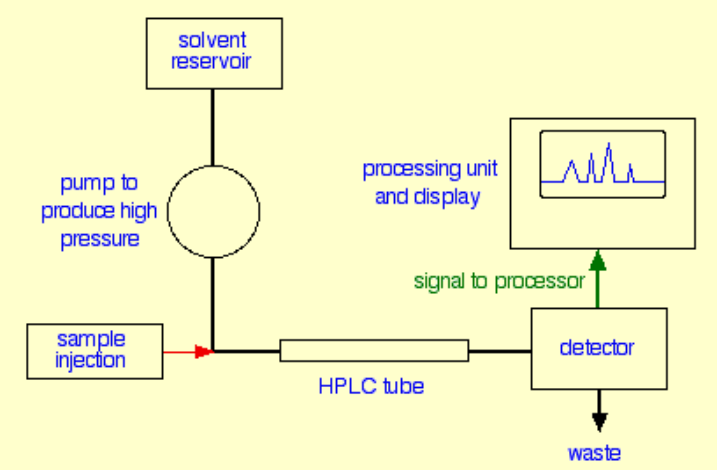

Figure 3.7: A general flow scheme of HPLC. The mobile phase is kept at the reservoir and it is moved to the column through the pump. Once the sample is injected, it dissolves in the mobile phase and they go through the column together. According to the strength of the interactions between the sample in the solvent and the stationary phase, retention time of the components is determined. In addition, once each of the component absorbs the wavelength coming from the detector, they give peak that these peaks determine the characteristics and the purity of the components. Adapted from [65].

The other major difference of HPLC is that it has two modes about the nature of stationary and mobile phases: It can be used both as normal phase and reversed phase. The criteria determining the mode of detection is the relative polarity of the mobile phase and the stationary phase. In normal phase HPLC, the stationary phase, which is the column packing material, consists of tiny polar silica particles. For normal phase, the mobile phase has less polarity than the stationary phase that the non-polar molecules will pass more quickly through the column. However, in reversed-phase HPLC (RP-HPLC), the column packing material is modified to become non-polar. Long hydrocarbon chains are attached to the silica particles. These hydrocarbon chains can have 4,8 or 18 carbon atoms, and the column types in reversed-phase HPLC are named according to the length of hydrocarbon chains. For instance, if the silica particles are attached to 18 carbon atoms, the column type is C18. As the name of the mode suggests, the mobile phase will be polar, unlike in normal phase. Thus, there will be a strong attraction between the sample in the solvent and the stationary phase that polar molecules will pass the column more quickly. Due to the strong interactions in reversed phase, the separation of the components occur in high sensitivity, versatility, and efficiency [63-65].

Reversed-phase HPLC has become an essential tool for the analysis and purification of peptides [66]. Therefore, for this thesis, reversed phase HPLC methods

were used to analyze whether peptide synthesis had been successful. In addition, as our peptide sequences are designed computationally, they are not routinely available, but could be custom synthesized. To ensure that our HPLC analysis are correct, we also obtained our peptide sequences synthesized by a company, TAG Copenhagen. Hence, these custom synthesis peptides were used as standard samples.

In general, the mobile phases for the peptide analysis and separation in HPLC were 0.1% trifluoroacetic acid (TFA) in water, and 0.1% TFA in Acetonitrile [66, 67]. TFA is added to fully protonate organic molecules [63]. Thus, TFA causes the formation of effective ion pairs on the peptide and decreases the pH below pK_a value of the peptide to increase retention [67].

The methodology of our reversed-phase HPLC studies is given below, and the duration of various mobile phases is given in Table 3.1.

Mobile phase A: 0.1% TFA in H_2O

Mobile phase B: 0.1% TFA in Acetonitrile

Stationary phase: ACE C8-300

Flow rate: 1 ml/min

Injection volume: 10 µl

Diode-Array Detector Wavelength: 220 ± 4 nm

Gradient Timetable

Table 3.1:Gradient Timetable of HPLC method used.

Time (min)	Mobile Phase A (%)	Mobile Phase B (%)
0	88	12
0.10	88	12
15.00	63	37
20.00	20	80
25.00	10	90
45.00	0	100
55.00	10	90
70.00	88	12

Using this methodology, both manually and custom synthesized peptides were analyzed in Agilent 1260 Infinity II HPLC located in Koç University Tüpraş Energy Center (KUTEM). First, the in-house and commercially synthesized 7 amino acid peptides were analyzed; their chromatograms are given in Figures 3.8 and 3.9, respectively.

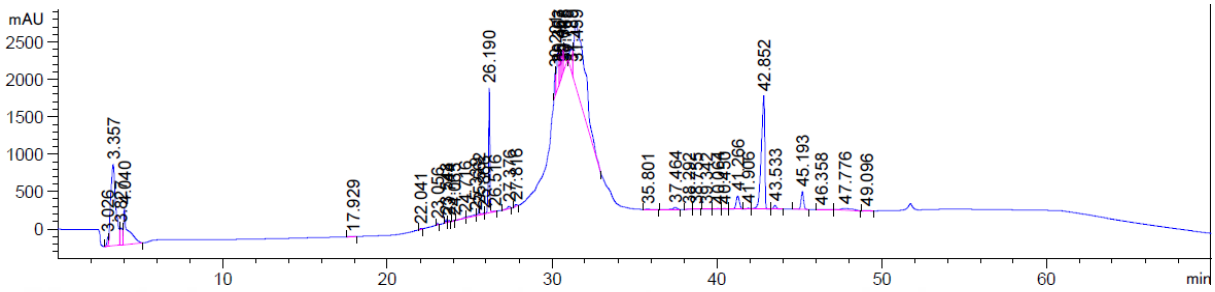

Figure 3.8: Chromatogram of manually synthesized peptide 7.

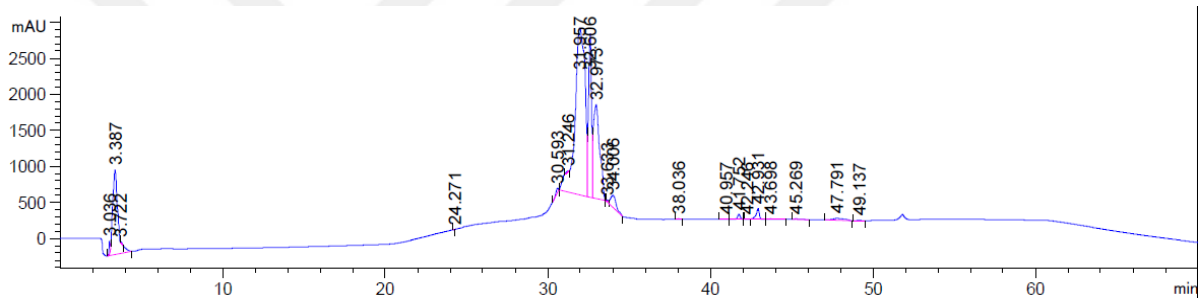

Figure 3.9: Chromatogram of standard sample, custom synthesized peptide 7.

From the commercial standard, it is seen that 7 amino acid peptide is eluted from the column between 30 and 34 minutes. When compared with manually synthesized peptide, the result is consistent, which shows our synthesis was successful.

Then, HPLC analysis was continued for 8 and 9 amino acid peptides. First of all, we tried to obtain the chromatograms of standard samples.

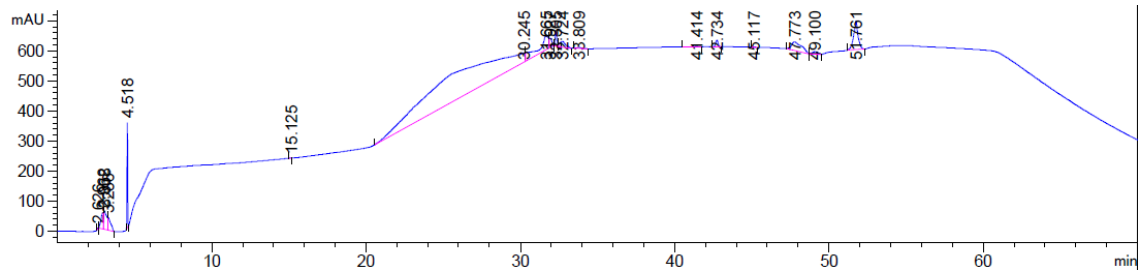

Figure 3.10: Chromatogram of standard sample, custom synthesized peptide 8.

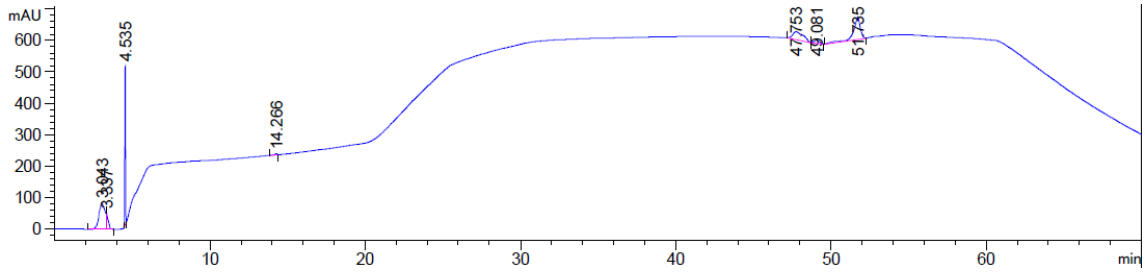

Figure 3.11: Chromatogram of standard sample, custom synthesized peptide 9.

As it is seen from the chromatograms of standard samples of 8- and 9- amino acid peptides, there is a broad plateau in both, and we could not obtain sharp peaks. Even though, we performed troubleshooting experiments, the chromatograms did not change. As we performed the analysis with the standard samples and have had the chromatograms of them sent by the company, we had a suspicion if there is any problem in the column and/or in the instrument. Then, we also performed the analysis of our standard samples using a different instrument (Thermo Scientific Ultimate 3000 HPLC) located in the Koç University Translational Medicine and Research Center (KUTTAM). The chromatograms of the standard samples are given in Figures 3.12, 3.13, 3.14, 3.15, 3.16, and 3.17.

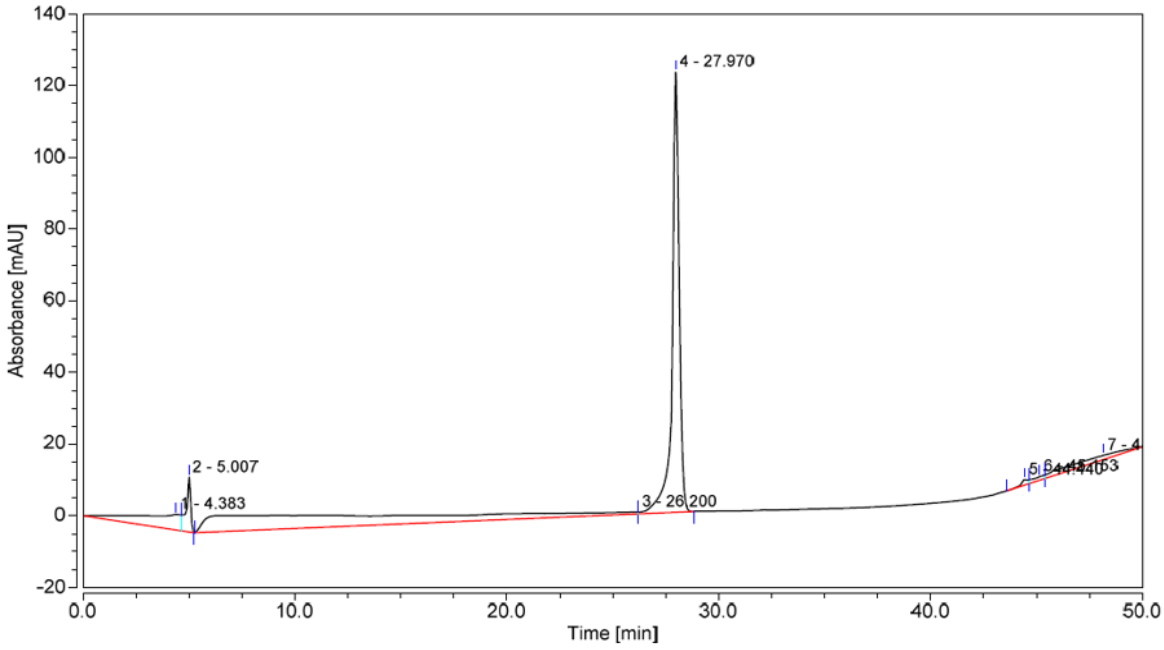

Figure 3.12: Chromatogram of standard sample, custom synthesized peptide 7.

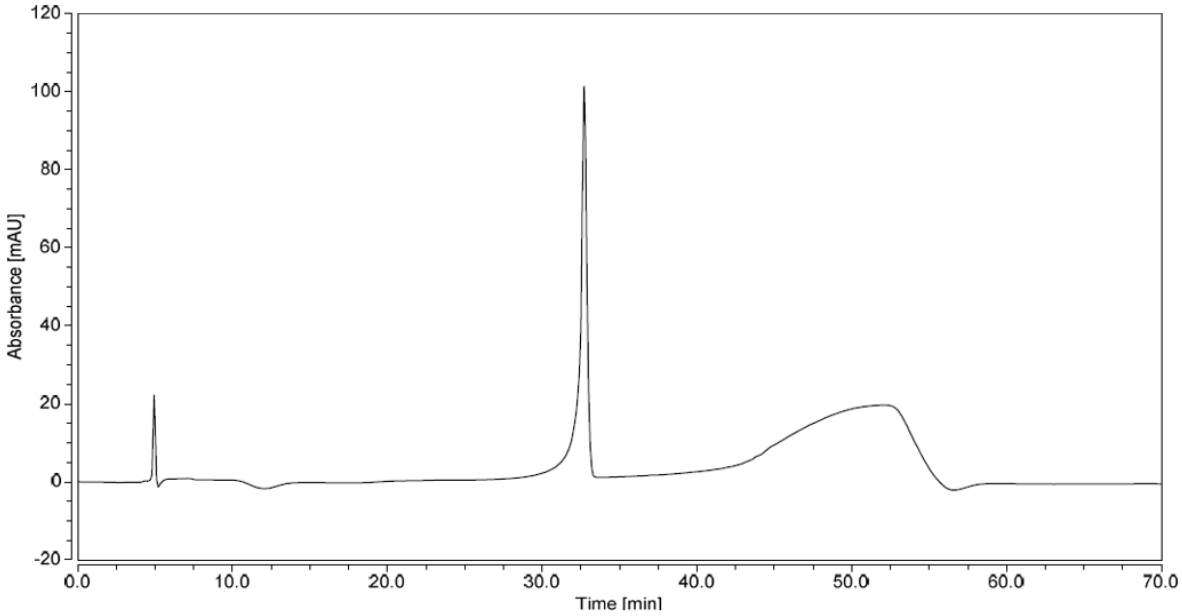

Figure 3.13: Chromatogram of standard sample, custom synthesized peptide 8.

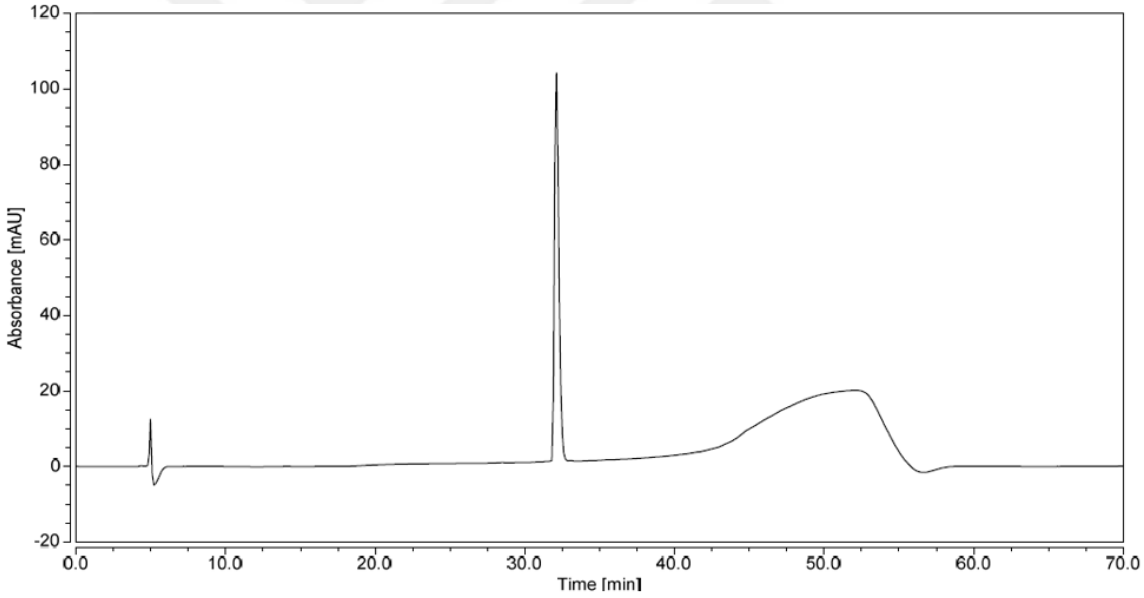

Figure 3.14: Chromatogram of standard sample, custom synthesized peptide 9.

Consequently, the expected chromatograms of the standard samples were obtained with this second-instrument at KUTTAM. This showed us that a problem occurred in the column or in the instrument at KUTEM. Therefore, we also analyzed our manual synthesis peptides using the device at KUTTAM.

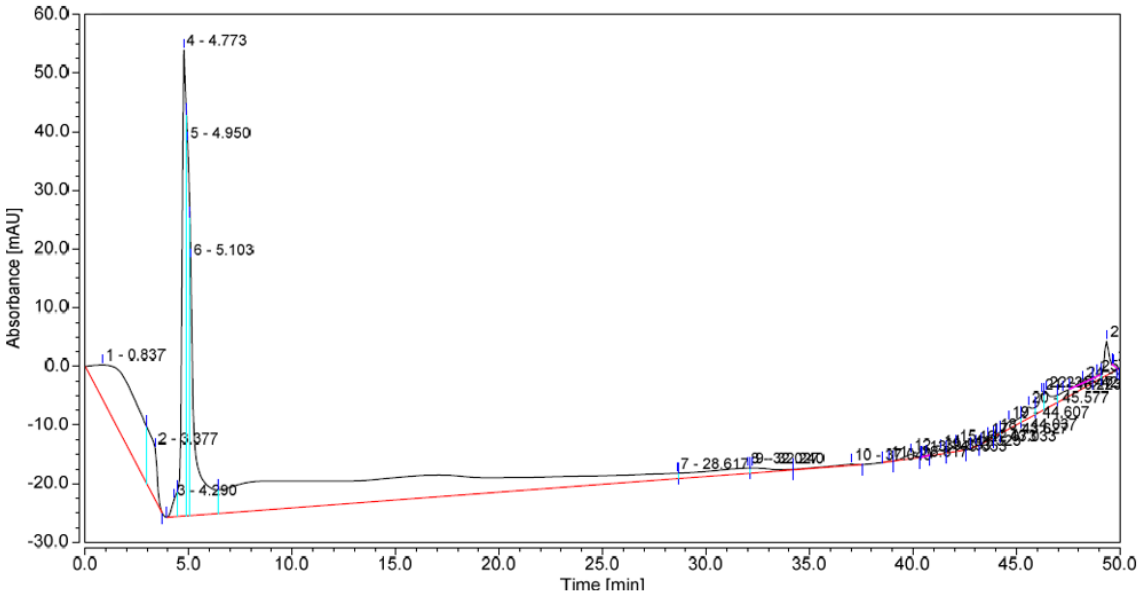

Figure 3.15: Chromatogram of manually synthesized peptide 7.

As can be seen from the Figures 3.8 and 3.9, we concluded our synthesis was successful for Peptide 7 when we performed the analysis at KUTEM. We had aimed to first purify the sample then to lyophilize it for the storage. Even though we were informed that there was a preparative C8 column at KUTEM, the column was not in the inventory. While we were waiting for the preparative column to arrive, we tried to solve the problem associated with the instrument at KUTEM by analyzing the 8- and 9-aa peptides. During that period, peptide solutions were formulated in 70% acetonitrile and 30% H_2O and stored at -20°C. However, during the cleavage from the resin, trifluoroacetic acid was used. It was shown previously that residual amount of TFA remains in the peptide solution and this might causes side reactions during the storage such as peptide fragmentation [68]. These chromatograms also show the situation that the sample was fragmented due to the storage in the solution during that period. The chromatograms of Peptide 8, and 9 also show similar results. Therefore, we could not show that our peptide synthesis was successful due to problems associated with the instrument at KUTEM.

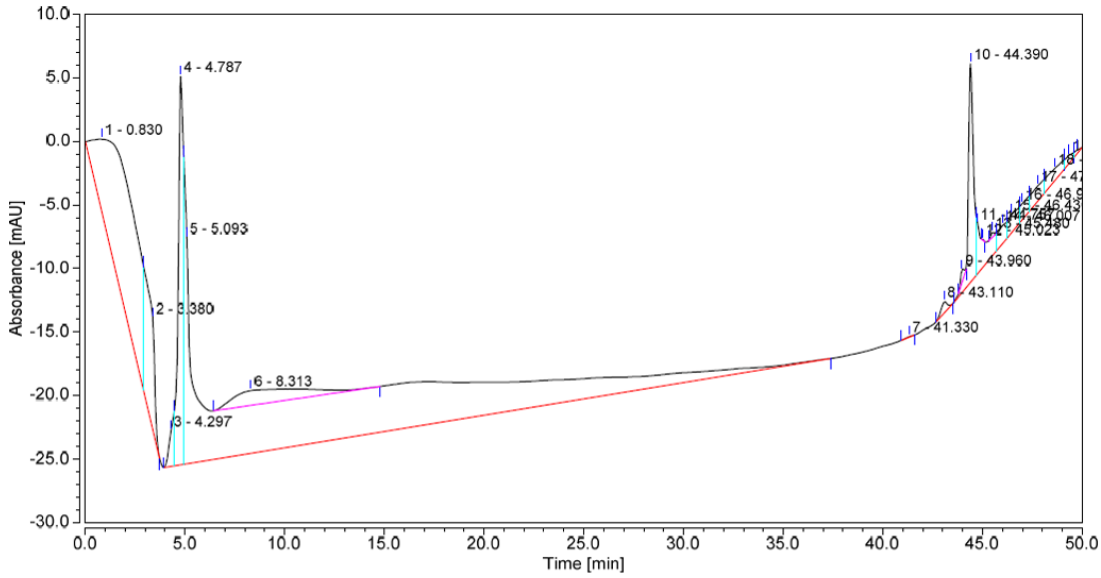

Figure 3.16: Chromatogram of manually synthesized peptide 8.

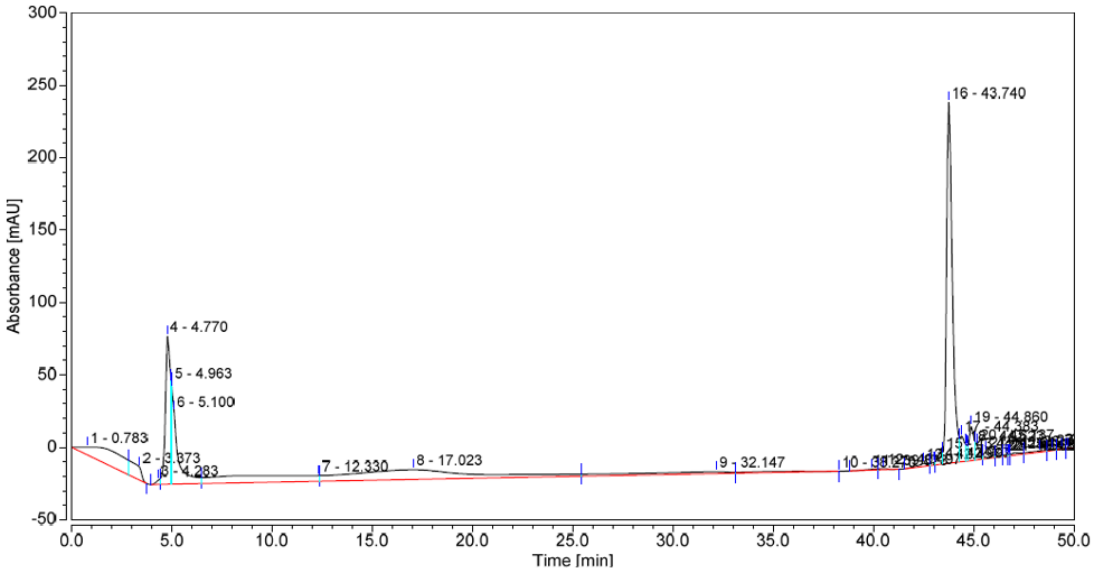

Figure 3.17: Chromatogram of manually synthesized peptide 9.

Figure 3.17 shows a sharp and split peak at around 43rd minute. However, as it is seen from the chromatogram of standard of peptide 9 (Figure 3.14), the expected peak is at 32nd minute. Thus, this peak does not correspond to the desired peptide. A larger molecule than the desired peptide might have formed as a result of side reactions with the TFA in the peptide solution.

Since we could not verify our in-house peptide synthesis protocol, surface plasmon resonance (SPR) experiments were conducted using the custom synthesized peptides obtained from TAG.

Chapter 4 :
BINDING AFFINITY MEASUREMENT BY SURFACE PLASMON RESONANCE

In this chapter, in order to experimentally validate the binding affinity of the *in silico* designed peptides, surface plasmon resonance measurements were performed using the Biacore T200 instrument in the laboratory of Prof. Batu Erman at Sabancı University.

Surface plasmon resonance (SPR) is an important optical detection method to understand and quantitatively analyze biomolecular interactions label-free and in real-time [69]. In addition to classic binding experiments, binding affinity can be measured using Scatchard Plot, Fluorescence Resonance Energy Transfer, and Isothermal Titration Calorimetry. Surface plasmon resonance is a method that gives accurate results due to its ability to label-free detection, handle complex samples, replicate measurements, real-time monitoring, and requirement of minimal amount of samples [70]. Importantly, both the association (k_{on}) and the dissociation rate constant (k_{off}) can be determined independently; and their ratio, which is k_{off}/k_{on}, gives the equilibrium dissociation constant, K_D [71]. Therefore, this method also helps us to analyze and understand the dynamics of the binding interaction instead of only measuring binding affinity. Therefore, the further studies can be designed in a better understanding of the interactions between the molecules.

Surface Plasmon Resonance is a phenomenon where a light source passes through a glass prism which reflects the backside of the sensor chip surface. This reflected beam then goes into a detector at a certain incident angle referred to as the resonance angle [72]. The light is absorbed by the electrons in the metal film of the sensor chip that they start to resonate. These resonating electrons are referred to as surface plasmons.

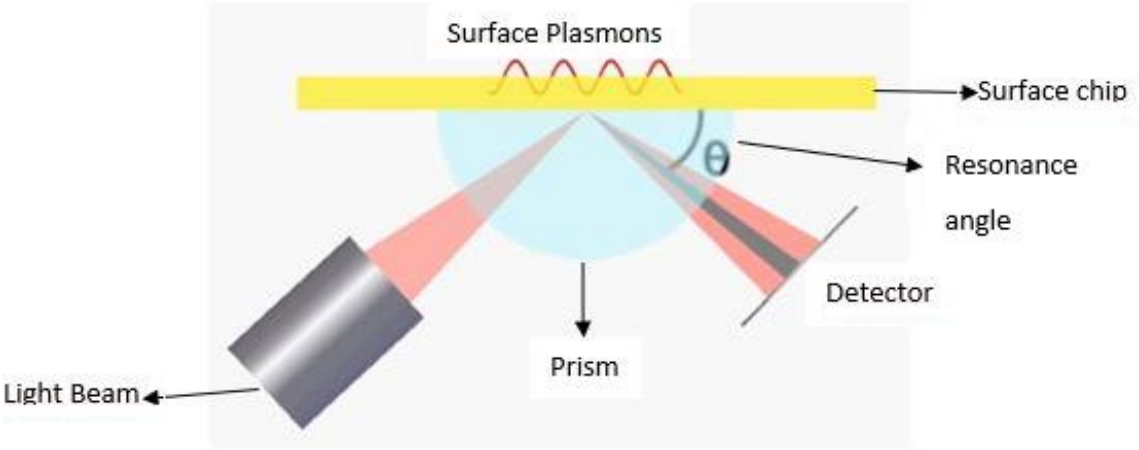

Figure 4.1: A general scheme of SPR. The light is passed through a prism into the surface chip that reflects the metal film of the chip surface. The reflected beam is collected and analyzed by the detector. Adapted from [70].

At resonance angle, surface plasmons start to resonate with light, which causes the absorption of the light at that incident angle. This creates an intensity loss in the reflected beam, which appears as a dark band. This band is seen as a dip in the SPR reflection intensity curve. The shape and location of the dip are used to analyze the interactions which occur at the sensor surface. As the molecular interactions take place, the angular position of the dark line shifts, and these shifts give the information about the molecular binding events on the surface and the interactions kinetics [71, 72]. Figure 4.2 shows the dark line in the reflected beam and its shifted position upon binding.

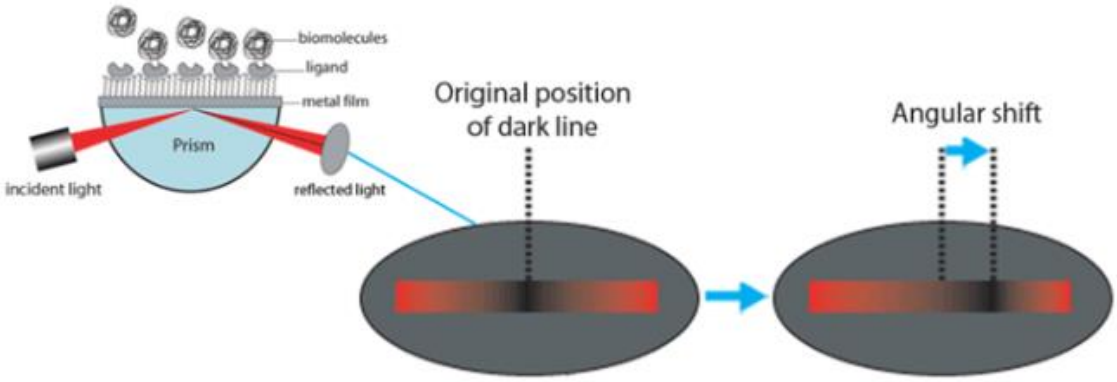

Figure 4.2: Once the surface plasmons resonate with light, they give a dark line in the reflected beam at their resonance angle and this dark line shifts as the molecular binding events occur in the surface. Adapted from [73].

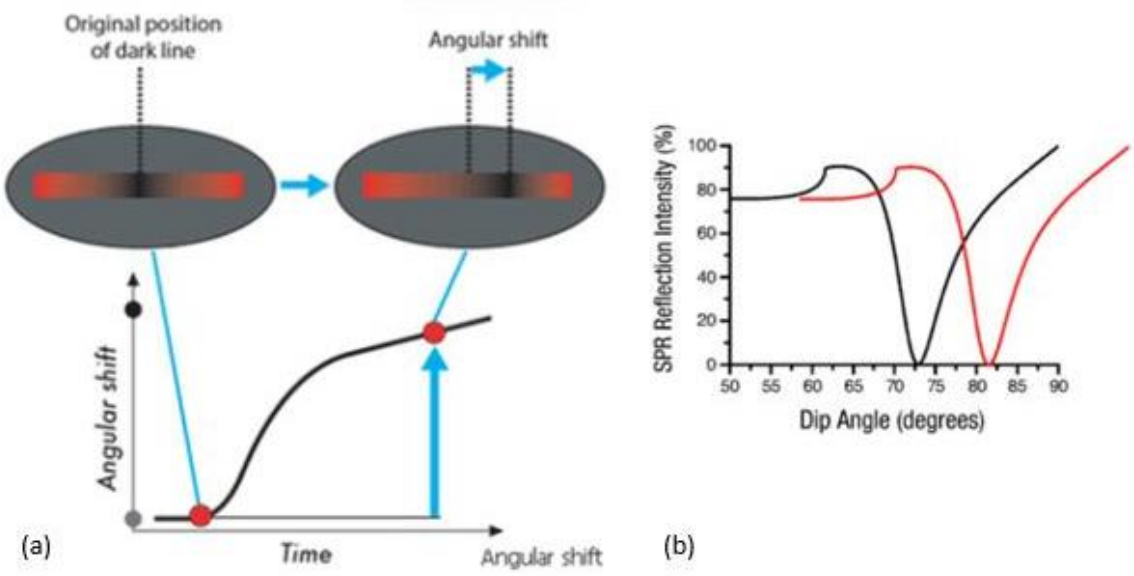

Figure 4.3: (a) As molecular binding events take place on the surface, the position of the dark line in the reflected beam changes. In order to study the kinetics of the interactions, the dark line is monitored and its shift vs. time is analyzed. (b) The angular shift in the position of the dark line is shown in red. The shape of the curve gives information about the surface. Adapted from [73].

In drug design terminology, ligand refers to a molecule that binds to a specific protein in the environment. However, in SPR terminology, the ligand actually refers to the molecules which are in the environment, in other words, proteins/receptors. The ligand is immobilized on the sensor chip surface then the analyte, peptides in our case, is flowed continuously over the immobilized ligand and binds them. As molecular binding events take place, mass of the sensor chip increases. This increase causes a proportional increase in the refractive index. Refractive index is an optical term that corresponds to a dimensionless number describing how fast light can go from one medium to another within a material. In SPR, the refractive index is measured as a change in the resonance angle. Therefore, as molecular interactions occur, both mass on the surface chip and change in the resonance angle increase. This increase is proportional to the analyte amount bound to the ligand in protein – protein interactions [74].

As mentioned above, when surface plasmons resonate, a dark line appears in the reflected beam. As the analyte binds the ligand molecules immobilized on the chip, mass of the chip and change of the resonance angle increase. Thus, angular shift occur in the dark line's position and this changes the signal coming from the detector. The response signal is measured in Response Units (RU). Experimental setup of SPR is given in Figure 4.4.

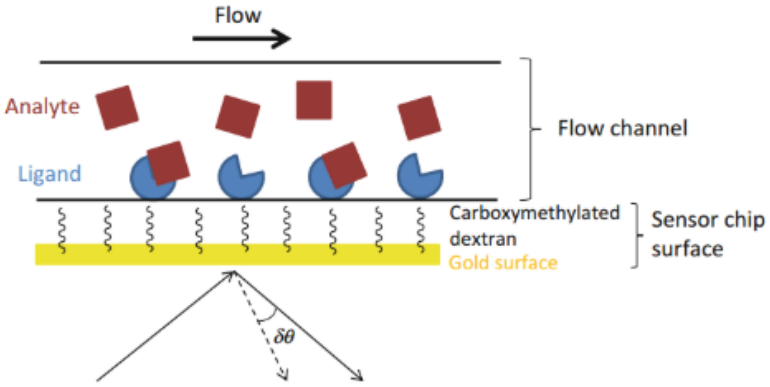

Figure 4.4: Experimental setup of Surface Plasmon Resonance. Adapted from [74].

Monitoring the changes in response signal during the experiment gives a graph, referred to as the sensorgram. The sensorgram shows how the response signals change during the experiment. Figure 4.5 presents the depiction of the different phases of an experiment in a sensorgram plot.

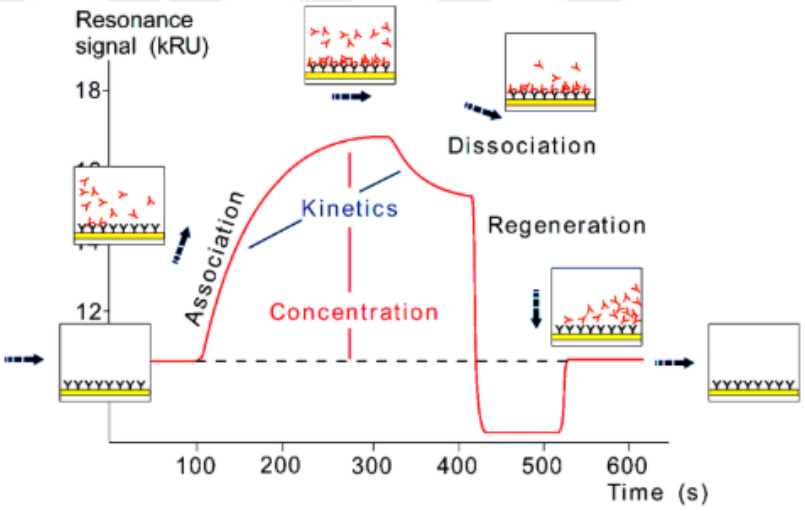

Figure 4.5: As the analyte is injected, analyte molecules are bound to the ligand that association phase dominates. Once the injection is complete, the analyte starts to dissociate from the ligand. During the dissociation phase, a regeneration buffer is loaded to the flow channel to reach the baseline again. Analysis of association and dissociation phases are named as kinetics. The change of response units during association and dissociation phases determine the concentration dependency. Adapted from [74].

There are different kinds of sensor chips used in SPR. All of them have gold surface but their coatings are different due to the various immobilization chemistries for specific molecules interactions. For protein-protein interactions, CM5 sensor chip is used [75] which has carboxymethylated dextran matrices on the gold surface that allows amine coupling chemistry for the immobilization. Amine coupling uses N- and ε- terminal lysine

residues of the ligand to immobilize them to the chip surface [76]. In our experiments, due to the nature of the molecules, CM5 chip was used.

We thus conducted our experiments on the Biacore T200 SPR instrument. The ligand is the IL1-R, and the analyte molecules are our computationally designed 7-amino acid, 8-amino acid, and 9-amino acid peptides. The purpose of this experiment was to validate the binding affinity of the designed peptides to block IL1-R, which is the primary goal of this thesis. In order to validate the instrument's accuracy, we also decided to test IL1-β binding to IL1-R, since, IL1-β is the natural ligand of IL1-R and its binding affinity is known from the literature. To perform the experiments, IL1-β and IL1-R proteins were bought from R&D Systems, while custom synthesized peptides used as analyte.

For the first step, IL1-R was immobilized to the sensor chip CM5. In order to validate immobilization of IL1-R to the sensor chip, we decided to first immobilize a small amount of the IL1-R protein to the surface. Figure 4.6 shows the immobilization of IL1-R protein to the chip.

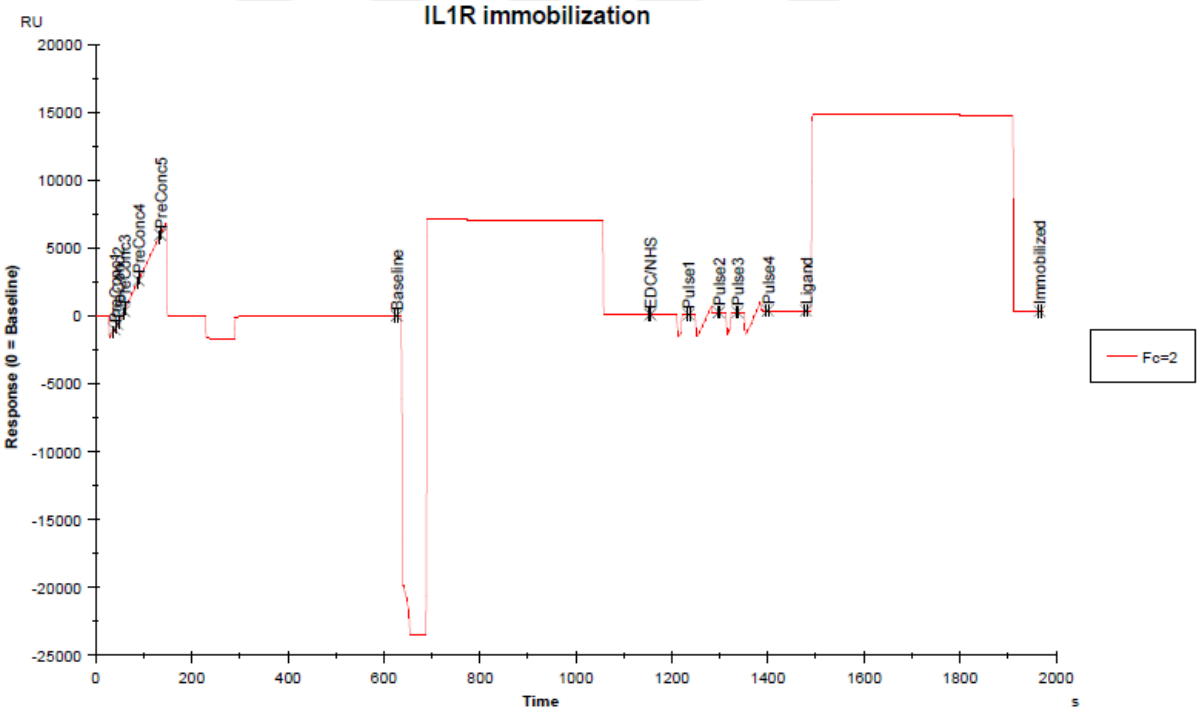

Figure 4.6: The sensorgram for IL1-R Immobilization to sensor chip CM5 into flow cell 2.

The target immobilization level was given to the instrument as 220 RU at pH 5. 278 nM of IL1-R solution in HBS-EP buffer (0.01 M HEPES, pH 7.4, 0.15 M NaCl, 3 mM EDTA, 0.005% v/v Surfactant P20) was prepared and given to the instrument. The

immobilization starts with the pre-concentration step. During pre-concentration, the instrument requires 5 different concentration steps and tries to obtain high protein concentration at the chip. This procedure hence, makes the immobilization more efficient [77]. If the pre-concentration step is not successful and it cannot obtain a linear increase from the first pre-concentration to the fifth step, then the immobilization procedure stops, and the experiment must be repeated with a concentrated ligand solution. In our experiment, pre-concentration steps were completed successfully that the system could reach the baseline. Once the baseline gets stable, EDC/NHS solution is injected to activate the chip surface by modifying the carboxymethyl groups to N-hydroxysuccinimide esters [76]. When EDC/NHS injection starts, the baseline may first change as shown in the graph (Figure 4.6) but then it increases due to surface activation. After activation is complete, ligand is injected, and it appears in the plot as pulses. The last pulse (Pulse 4 in this case, Figure 4.6) indicates the final immobilization level of the ligand. In this graph, Pulse 4 is 911 RUs. In order to prevent further undesired immobilization, the surface must be deactivated. For this purpose, ethanolamine hydrochloride solution is used. This is the phase between 1500 and 2000 seconds. The deactivation procedure does not change the amount of immobilized ligand. To find out the final immobilized ligand amount, the amount which EDC/NHS solution activated must be subtracted from the level of Pulse 4. For our experiment, EDC/NHS activation and the last pulse correspond to 618 and 911 Rus, respectively, thereby yielding an immobilized ligand level of 293 Rus. We had thus successfully immobilized the IL1-R ligand at low level on the CM5 chip. The next step was to measure the binding affinity of IL1-β to IL1-R to verify the accuracy of the instrument.

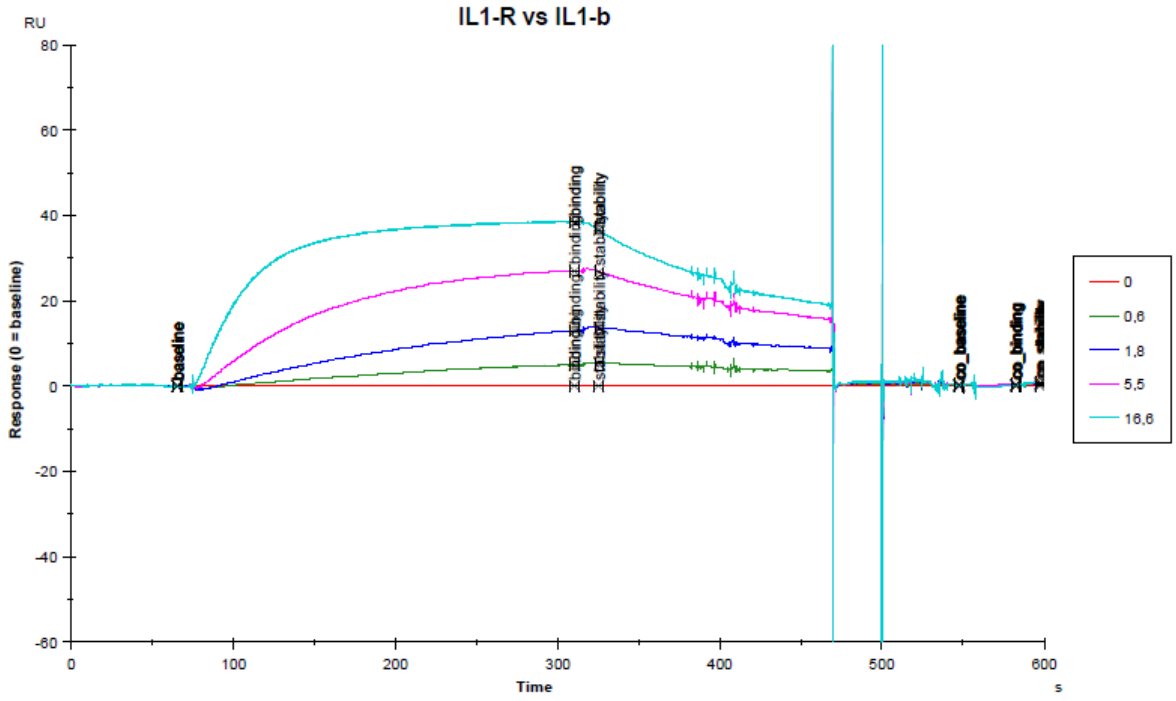

Figure 4.7: Sensorgram for IL1-β binding to IL1-R.

In order to measure the binding affinity of IL1-β to IL1-R, kinetics analysis was performed. As the flow cell 2 was immobilized with IL1-R, the experiment was again conducted using this flow cell. IL1-β solutions were prepared at 5 different concentrations, namely 0 µM, 0.6 µM, 1.8 µM, 5.5 µM, and 16.6 µM. The sensorgram (Figure 4.7) shows the expected result that as the concentration increases, the change in the response units increase as well. As a result of this experiment, K_D was measured as 2.8 nM which is consistent with the result of previous experimental studies; (2.7 nM) [78]. Accordingly, the experimental setup and conditions were adequate measuring the binding affinity of our peptides. Thus, the next step in our project was to immobilize a higher amount of IL1-R to the sensor chip and to measure the binding affinity of our computationally designed peptides to IL1-R.

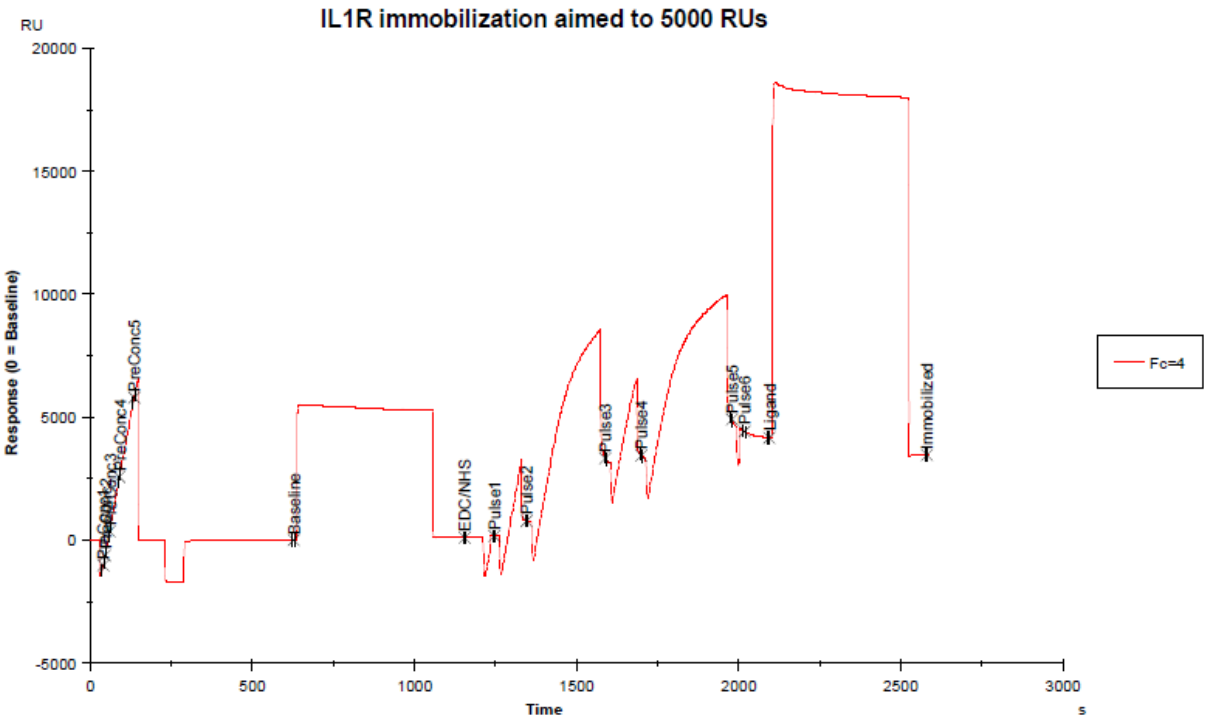

Figure 4.8: The sensorgram for IL1-R Immobilization to sensor chip CM5 into flow cell 4.

Along with the results obtained, we immobilized IL1-R at 3360 Response Units again using amine coupling chemistry as shown in Figure 4.8. The target value for immobilization was 5000 RU.

In order to conduct kinetics analysis for the peptides, the procedures in the literature include the use of dimethylsulfoxide (DMSO) for dissolving the peptides and hence; in the running buffer [75, 79].

According to this information, the peptides, which have 7 and 8 amino acids, were dissolved in 100% DMSO, while the 9 amino acids peptide was dissolved in 84% v/v DMSO solution in water. One of the protocols from the manufacture of Biacore (GE Healthcare) recommended including 2% DMSO in the running buffer. It should be noted that what is used to prepare the analyte solution must be included in the running buffer [75]. The protocol suggests using PBS-P 1X (Phosphate buffered saline, pH 7.4, supplemented with 2% DMSO) as the running buffer. To make the DMSO content of the peptide solutions consistent with the running buffer, the 7- and 8- amino acid peptide solutions were diluted 50-fold to 200 µM with 1X PBS-P solution. Since the 9 amino acid peptide has a different DMSO composition, it was first diluted 25-fold to 200 µM again with 1X PBS-P solution but, at this concentration, the DMSO concentration was 3.36%

v/v. As DMSO is highly hygroscopic, buffer mismatch is very important that DMSO concentration was reduced to 2% in 100 µM 9 amino acid peptide solution. After the said dilutions, the peptide stock solutions with 2% DMSO concentration and 1X PBS-P running buffer including 2% DMSO were ready for kinetics experiments. 100 µM, 50 µM, 25 µM, 12.5 µM, 6.25 µM, and 0 µM peptide solutions were used for the SPR measurement. By using these solutions and the running buffer, Figures 4.9, 4.10, and 4.11 which present the sensorgram graphs of 7, 8, and 9 amino acid peptides, respectively were obtained.

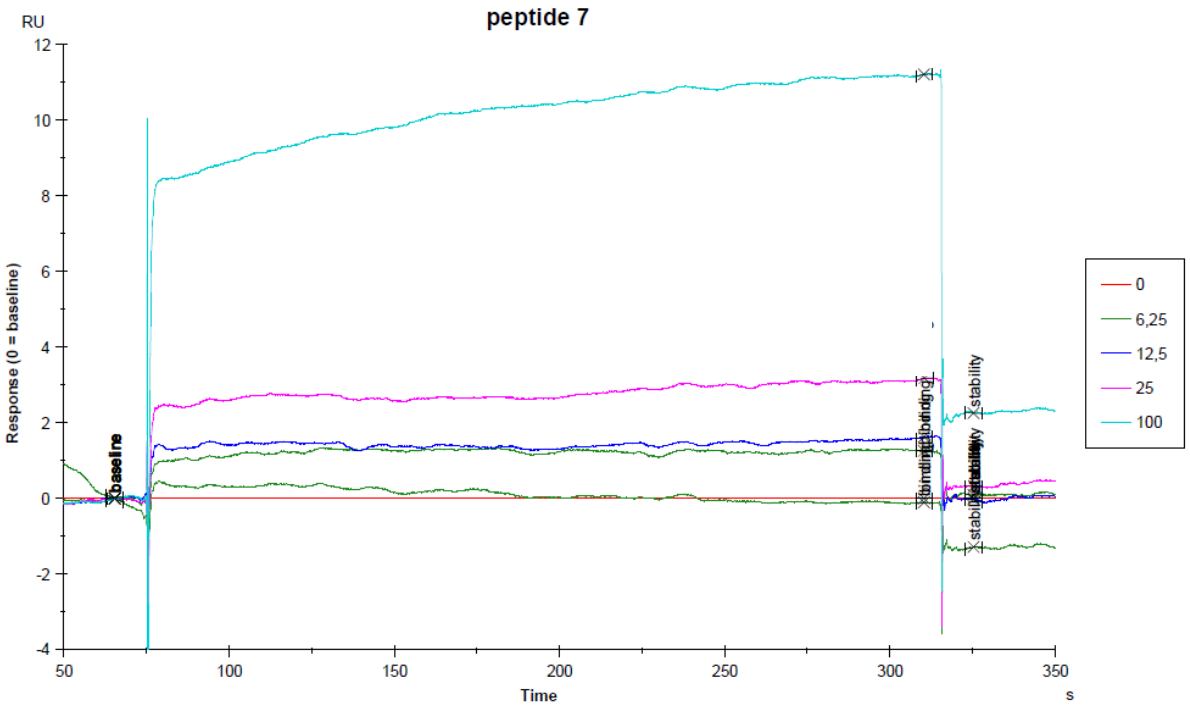

Figure 4.9: The sensorgram for kinetics analysis of 7 amino acid peptide.

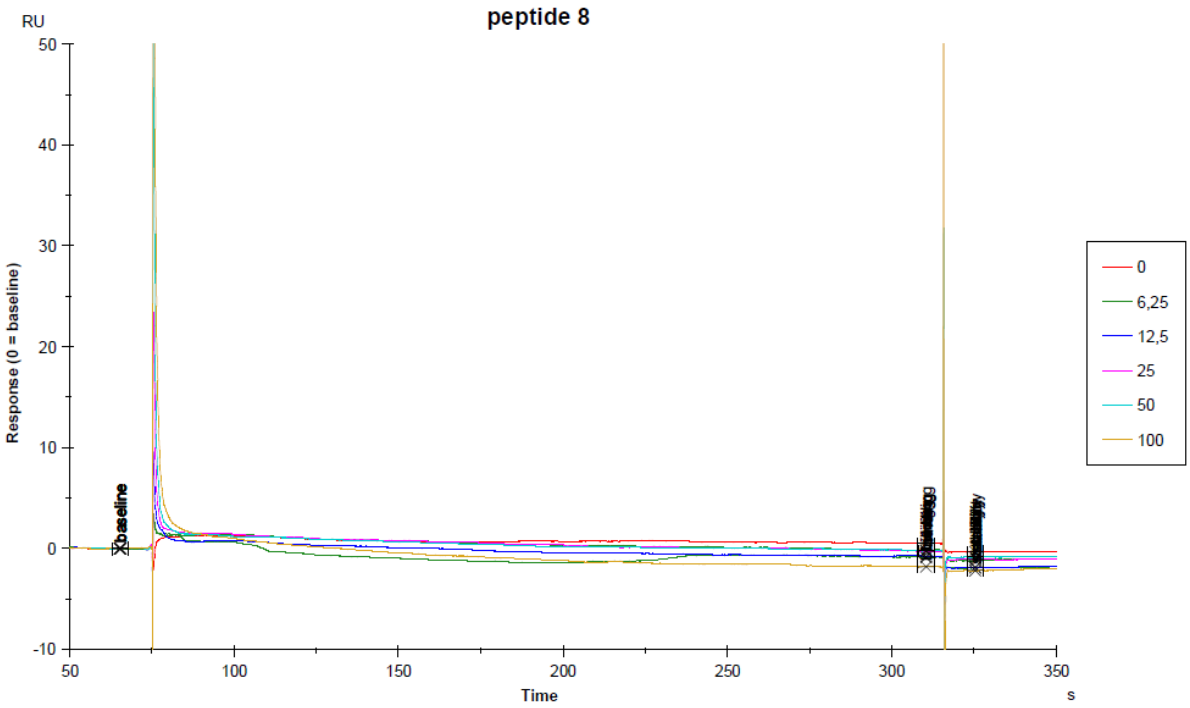

Figure 4.10: The sensorgram for kinetics analysis of 8 amino acid peptide.

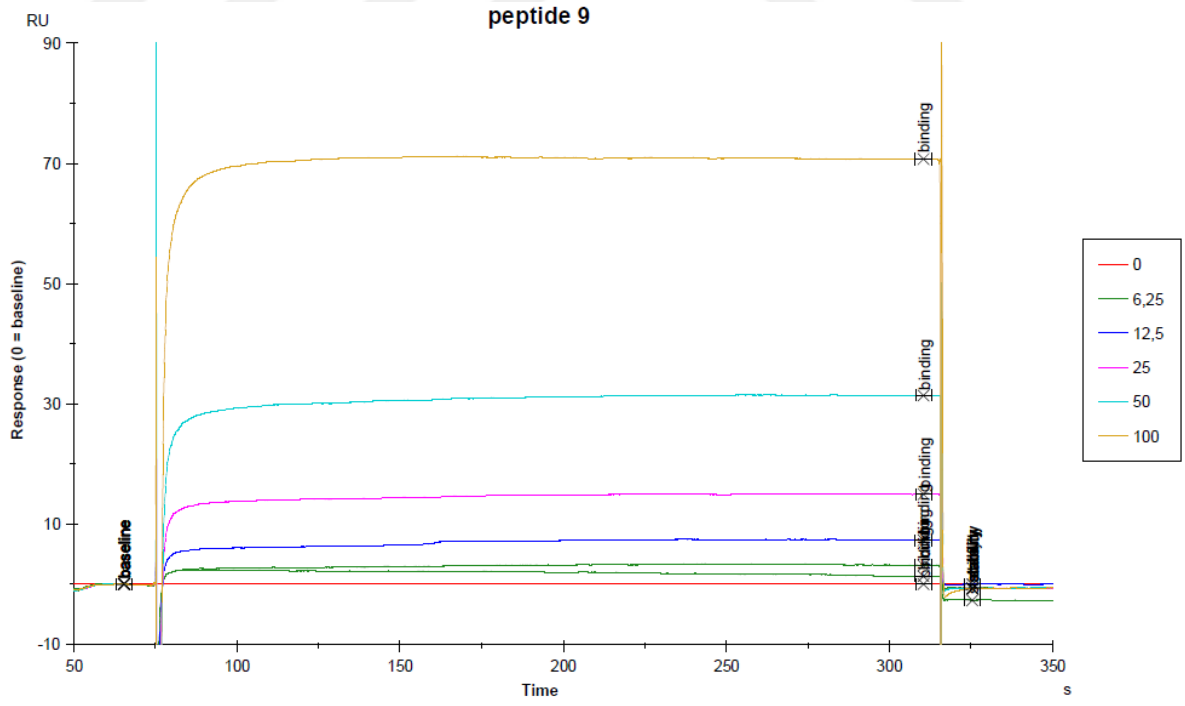

Figure 4.11: The sensorgram for kinetics analysis of 9 amino acid peptide.

We could not obtain affinity constants from the SPR experiments, but the sensorgrams from all three peptides showed concentration dependency in ligand binding. Moreover, these results indicate that the experimental setup must be improved to conduct the experiment. In the sensorgram of peptide 7, there is no line for 50 μM analyte. If there is no concentration dependency in one of the analytes, Biacore automatically exclude that sample. Furthermore, when the steepness of slopes was analyzed, it was understood that there was a huge bulk effect due to DMSO. Bulk effect, also named as solvent artifact, can be defined as the difference of the component composition in the analyte and the running buffer. As DMSO is highly hygroscopic, 1% buffer mismatch creates a difference of 1200 RUs. Due to the small molecular weight of our peptides, the maximum response units could be reached in our experiments were at around 100 RUs. As 1200 RUs is a much larger number, it might be masking the results we obtained from the experiments. In order to analyze our results, this effect must be removed. Additionally, while conducting the experiment the flow cell just next to the one the ligand was immobilized is used as a reference flow cell. When the analyte is injected, the samples go to both reference cell and the flow cell at which the experiment is conducted. As there is no immobilized ligand on the sensor surface of the reference cell, the response unit should be very low. Due to the difference in buffer composition, additives, pH and analyte concentrations, the reaction of the reference cell to the analyte might change. Even though the system gives the amount of immobilized ligand on the surface in Response Units, it occupies the carboxymethylated dextran matrices volumetrically. After the flow of the running buffer on the surfaces, the analyte is injected, and the matrices might shrink or expand due to the different volumes of the ligand molecules in the matrices. In those cases, the reaction of both reference and the active flow cell differ under ionic interactions and the use of highly hygroscopic solvents such as DMSO. Since, they can create non-specific interactions between the analyte and the surface which causes fallacy for the analysis of the binding interactions. Even though the aim of the reference flow cell is to correct this effect, 1 – 2% difference could happen which creates huge bulk effect especially, if DMSO is used. 1 – 2% difference causes approximately 4500 RUs change when DMSO is used [80, 81].

To prevent the bulk effect of DMSO, a solvent correction can be applied. This procedure basically uses a calibration curve to correct for the 1 – 2% difference in DMSO content between reference and active flow cells.

In our experiments, as we did not perform solvent correction, the bulk effect of DMSO may have masked the binding interactions and led to steep binding curves in the sensorgrams. As mentioned above, the steepness of the slopes is higher than expected. This might happen due to two reasons. One of them is the peptide binds and then unbinds very fast which also gives high affinity. The second reason might be due to the bulk effect. As discussed above, DMSO creates non-specific interactions with the analyte that causes misinterpretation to analyze the graphics. In order to understand and analyze the binding behavior of our computationally designed peptides to IL-1R, Biacore experiments must be repeated by performing solvent correction so that the bulk effect due to DMSO can be accounted for.

Chapter 5 :
CONCLUSIONS

Peptides are attractive therapeutic molecules due to their promising features, such as high affinity and specificity to the target, deep tissue penetration, low toxicity, convenience for structural modifications, and cost effectiveness. To date, 60 peptide-based drugs have entered in the pharmaceutical market. For a lead molecule to reach the market, approximately a 15-year period and an investment of around 2.6 billion dollars are necessary. One of the most important parameters for evaluating lead compounds in the process of Drug Discovery and Development is the binding affinity, which is the strength of the non-covalent interactions between the drug and its target. IL1-β and its receptor (IL1-Receptor) are attractive targets due to their highly importance in the metabolism of autoimmune and autoinflammatory diseases as well as Type II Diabetes and apoptosis, IL1 proteins have become of interest. When the levels of IL-1β were analyzed in the patients with *Rheumatoid Arthritis*, it was found that there is an increased production of IL1-β, or IL1-Receptor antagonist molecule. Elevated levels of IL-1β creates an imbalance in the pair of antagonist and IL1-β. Anakinra (Kineret™) is a recombinant protein that is very similar to the native IL1-Receptor Antagonist to block the increased activity of IL1-β. Thus, the main aim of this thesis was to design a peptide alternative to anakinra by computationally and then to validate the computational results experimentally.

First, peptides were designed *in silico* to bind IL1-R using the three-dimensional structure of IL1-Receptor antagonist molecule (Chapter 2). After the imbortant binding regions of the antagonist were determined, according to the type and number of non-covalent interactions between anakinra and the receptor, 9 different peptide sequences were obtained. To evaluate the binding affinities of these peptides, molecular docking simulations were performed using GOLD Software for 100 runs by fixing the bonds of the peptides. Peptide C, which has the sequence of Arg – Ile – Trp – Asp – Val – Asn – Gln, showed the highest binding affinity with K_D of 7.91×10^{-8} M (affinity for the best pose) (Figure 2.4). The docking studies were repeated to confirm the binding affinity of Peptide C without fixing the bonds for 999 runs. Since Peptide C has high affinity to the receptor, N- and C- terminal modifications were applied to increase the conformational

stability of the peptide. These modifications comprised inserting one of 20 natural amino acids successively to N- and C- termini. Phe insertion at the N- terminus resulted in the highest binding affinity due to its ability to form hydrophobic interactions between the receptor. Lys insertion at the C- terminus resulted in the highest binding affinity for N- and C- terminal modified peptide. As Lys is a charged amino acid, it could form additional hydrogen bonds between the receptor while also protecting the hydrophobic interactions. These modifications increased the number of hydrophobic interactions between the peptides and IL1-R, while also keeping the intermolecular hydrogen bonds intact. N-terminally modified 8- amino acid peptide has K_D of $0.88x10^{-7}$ M (affinity for the best pose) (Figure 2.5). Both N- and C- termini modified peptide has K_D of $0.24x10^{-9}$ M (affinity for the best pose) (Figure 2.6) which is the highest binding affinity among three of the peptides. To understand the dissociation pathway and calculate binding affinities of our designed peptides in a more accurate and realistic environment *in silico*, SMD simulations were performed for the unmodified, 7-aa sequence (Peptide C), N- terminally modified 8-amino acid peptide, and dual modified 9-amino acid peptide. The K_D values obtained from SMD simulations are $9x10^{-3}$ M, $8.9x10^{-3}$ M, and $8.3x10^{-4}$ M for Peptide C, N- terminus modified peptide, both N- and C- terminus modified peptide, respectively. The binding affinity values for all three of the peptides from SMD simulations are consistent with the average values obtained from the docking and the peptides show slightly elevated affinity consistent with the end modifications and the results of the docking studies. Since water molecules are present in SMD simulations, they come between the peptide and the receptor molecules, thereby increasing the distance between the interacting molecules (i.e. peptides and IL1-R) which, in turn causes the loss of some of the hydrogen bonds and most of the hydrophobic interactions. Three of the peptides show binding affinities to the receptor at the average values of the docking results due to the loss of hydrogen bonds in the presence of water.

After the computational studies, three of the peptides obtained as a result of the computational studies were synthesized using Fmoc-based solid phase peptide synthesis (SPPS) methods. High-performance liquid chromatography was used to analyze the synthesized molecules. Even though; we showed that our first synthesis (where the unmodified peptide produced) was successful, due to a problem occurrence either in the column or in the instrument, after a while, we could not even obtain the expected chromatograms for the custom-synthesized reference peptides. The chromatograms of

these standard samples were obtained using a different instrument. However, due to TFA fragmentation of the peptides, our synthesis products might have degraded; therefore, custom synthesized peptides were used in further (SPR) studies.

Finally, to validate the binding affinity of the peptides experimentally, Surface Plasmon Resonance (SPR) method was used (Chapter 4). In Biacore T200 instrument, IL1-R protein was immobilized on the at 3360 RU using amine coupling chemistry. To verify the accuracy of the instrument, binding affinity of IL1-β to IL1-Receptor was measured by kinetics analysis. Concentration dependent binding of IL1-β was observed, and K_D was found to be 2.8 nM which is consistent with the literature value-of 2.7 nM. Although concentration dependency was also observed for the peptides, binding affinity values could not be obtained from the sensorgrams due to the bulk effect of DMSO. DMSO creates non-specific interactions with the analyte that causes misinterpretation to analyze the results. Kinetic SPR experiments to measure binding affinity of the peptides will be repeated by performing solvent correction prior to the kinetics analysis to prevent bulk effect of DMSO from obscuring our binding measurements.

BIBLIOGRAPHY

1. Silverman, R.B., *Chem 316: The organic chemistry of drug design and drug action.* Abstracts of Papers of the American Chemical Society, 2006. **232**: p. 197-197.
2. Knowles, J. and G. Gromo, *Target selection in drug discovery.* Nature Reviews Drug Discovery, 2003. **2**(1): p. 63-69.
3. Broach, J.R. and J. Thorner, *High-throughput screening for drug discovery.* Nature, 1996. **384**(6604): p. 14-16.
4. Mullard, A., *New drugs cost US$2.6 billion to develop.* Nature Reviews Drug Discovery, 2014. **13**(877).
5. Hunter, C.A., *Meldola Lecture - the Role of Aromatic Interactions in Molecular Recognition.* Chemical Society Reviews, 1994. **23**(2): p. 101-109.
6. John Kuriyan, B.K.D.W., *The Molecules of Life: Physical and Chemical Properties*. 2013, New York and London: Garland Science. 1008.
7. Hulme, E.C. and M.A. Trevethick, *Ligand binding assays at equilibrium: validation and interpretation.* British Journal of Pharmacology, 2010. **161**(6): p. 1219-1237.
8. Noble, M.E.M., J.A. Endicott, and L.N. Johnson, *Protein kinase inhibitors: Insights into drug design from structure.* Science, 2004. **303**(5665): p. 1800-1805.
9. Shukla, D., et al., *Activation pathway of Src kinase reveals intermediate states as targets for drug design.* Nature Communications, 2014. **5**.
10. Du, X., et al., *Insights into Protein-Ligand Interactions: Mechanisms, Models, and Methods.* International Journal of Molecular Sciences, 2016. **17**(2).
11. Berg JM, T.J., Stryer L, *Biochemistry*. Fifth ed. 2002, New York: W H Freeman.
12. WILLIAMS, M. *Scientists Think They Found Missing Evidence That Explains How Life Started on Earth*. 2019 [cited 2019 21 July]; Available from: https://www.sciencealert.com/scientists-have-stumbled-upon-a-possible-missing-link-that-explains-the-origins-of-life.
13. Shockley, T.R., et al., *Penetration of Tumor-Tissue by Antibodies and Other Immunoproteins.* Annals of the New York Academy of Sciences, 1991. **618**: p. 367-382.
14. Gnanaraj, J. and M.W. Saif, *Hypersensitivity vasculitis associated with leuprolide (Lupron).* Cutaneous and Ocular Toxicology, 2010. **29**(3): p. 224-227.
15. Recio, C., et al., *The Potential Therapeutic Application of Peptides and Peptidomimetics in Cardiovascular Disease.* Frontiers in Pharmacology, 2017. **7**.
16. Sato, A.K., et al., *Therapeutic peptides: technological advances driving peptides into development.* Current Opinion in Biotechnology, 2006. **17**(6): p. 638-642.
17. Holmes, D.L., et al., *Site specific 1 : 1 opioid : albumin conjugate with in vitro activity and long in vivo duration.* Bioconjugate Chemistry, 2000. **11**(4): p. 439-444.
18. Kim, T.H., H. Lee, and T.G. Park, *Pegylated recombinant human epidermal growth factor (rhEGF) for sustained release from biodegradable PLGA microspheres.* Biomaterials, 2002. **23**(11): p. 2311-2317.
19. Still, J.G., *Development of oral insulin: progress and current status.* Diabetes-Metabolism Research and Reviews, 2002. **18**: p. S29-S37.
20. Lien, S. and H.B. Lowman, *Therapeutic peptides.* Trends in Biotechnology, 2003. **21**(12): p. 556-562.
21. Fosgerau, K. and T. Hoffmann, *Peptide therapeutics: current status and future directions.* Drug Discovery Today, 2015. **20**(1): p. 122-128.

22. Lau, J.L. and M.K. Dunn, *Therapeutic peptides: Historical perspectives, current development trends, and future directions.* Bioorganic & Medicinal Chemistry, 2018. **26**(10): p. 2700-2707.
23. Weber, A., P. Wasiliew, and M. Kracht, *Interleukin-1 (IL-1) Pathway.* Science Signaling, 2010. **3**(105).
24. Braddock, M. and A. Quinn, *Targeting IL-1 in inflammatory disease: New opportunities for therapeutic intervention.* Nature Reviews Drug Discovery, 2004. **3**(4): p. 1-10.
25. Ozbabacan, S.E.A., et al., *The Structural Pathway of Interleukin 1 (IL-1) Initiated Signaling Reveals Mechanisms of Oncogenic Mutations and SNPs in Inflammation and Cancer.* Plos Computational Biology, 2014. **10**(2).
26. Gabay, C., C. Lamacchia, and G. Palmer, *IL-1 pathways in inflammation and human diseases.* Nature Reviews Rheumatology, 2010. **6**(4): p. 232-241.
27. Brikos, C., et al., *Mass spectrometric analysis of the endogenous type I interleukin-1 (IL-1) receptor signaling complex formed after IL-1 binding identifies IL-1RAcP, MyD88, and IRAK-4 as the stable components.* Molecular & Cellular Proteomics, 2007. **6**(9): p. 1551-1559.
28. Cao, Z.D., W.J. Henzel, and X.O. Gao, *IRAK: A kinase associated with the interleukin-1 receptor.* Science, 1996. **271**(5252): p. 1128-1131.
29. Cao, Z.D., et al., *TARF6 is a signal transducer for interleukin-1.* Nature, 1996. **383**(6599): p. 443-446.
30. Kawagoe, T., et al., *Sequential control of Toll-like receptor-dependent responses by IRAK1 and IRAK2.* Nature Immunology, 2008. **9**(6): p. 684-691.
31. Dinarello, C.A., *Immunological and Inflammatory Functions of the Interleukin-1 Family.* Annual Review of Immunology, 2009. **27**: p. 519-550.
32. Eastgate, J.A., et al., *Correlation of Plasma Interleukin-1 Levels with Disease-Activity in Rheumatoid-Arthritis.* Lancet, 1988. **2**(8613): p. 706-709.
33. Rooney, M., J.A. Symons, and G.W. Duff, *Interleukin-1-Beta in Synovial-Fluid Is Related to Local Disease-Activity in Rheumatoid-Arthritis.* Rheumatology International, 1990. **10**(5): p. 217-219.
34. Vanhee, P., et al., *Computational design of peptide ligands.* Trends Biotechnol, 2011. **29**(5): p. 231-9.
35. Xu, D., et al., *Computational design of peptide ligands to target the intermolecular interaction between viral envelope protein and pediatric receptor.* Comput Biol Chem, 2017. **69**: p. 120-125.
36. Berman, H.M., et al., *The Protein Data Bank.* Nucleic Acids Res, 2000. **28**(1): p. 235-42.
37. Grigoryan, G., A.W. Reinke, and A.E. Keating, *Design of protein-interaction specificity gives selective bZIP-binding peptides.* Nature, 2009. **458**(7240): p. 859-64.
38. Fernandez-Ballester, G., et al., *Structure-based prediction of the Saccharomyces cerevisiae SH3-ligand interactions.* J Mol Biol, 2009. **388**(4): p. 902-16.
39. King, C.A. and P. Bradley, *Structure-based prediction of protein-peptide specificity in Rosetta.* Proteins, 2010. **78**(16): p. 3437-49.
40. Li, H.M., et al., *De Novo Computational Design for Development of a Peptide Ligand Oriented to VEGFR-3 with High Affinity and Long Circulation.* Mol Pharm, 2017. **14**(7): p. 2236-2244.
41. Fleischmann, R., R. Stern, and I. Iqbal, *Anakinra: an inhibitor of IL-1 for the treatment of rheumatoid arthritis.* Expert Opinion on Biological Therapy, 2004. **4**(8): p. 1333-1344.
42. Schreuder, H., et al., *A new cytokine-receptor binding mode revealed by the crystal structure of the IL-1 receptor with an antagonist.* Nature, 1997. **386**(6621): p. 194-200.
43. Jones, G., et al., *Development and validation of a genetic algorithm for flexible docking.* Journal of Molecular Biology, 1997. **267**(3): p. 727-748.
44. Dias, R. and W.F. de Azevedo, *Molecular Docking Algorithms.* Current Drug Targets, 2008. **9**(12): p. 1040-1047.

45. de Magalhaes, C.S., H.J.C. Barbosa, and L.E. Dardenne, *A genetic algorithm for the ligand-protein docking problem.* Genetics and Molecular Biology, 2004. **27**(4): p. 605-610.
46. Verdonk, M.L., et al., *Improved protein-ligand docking using GOLD.* Proteins-Structure Function and Genetics, 2003. **52**(4): p. 609-623.
47. Pantsar, T. and A. Poso, *Binding Affinity via Docking: Fact and Fiction.* Molecules, 2018. **23**(8).
48. Li, M.S. and B.K. Mai, *Steered Molecular Dynamics-A Promising Tool for Drug Design.* Current Bioinformatics, 2012. **7**(4): p. 342-351.
49. Isralewitz, B., et al., *Steered molecular dynamics investigations of protein function.* Journal of Molecular Graphics & Modelling, 2001. **19**(1): p. 13-25.
50. Park, S., et al., *Free energy calculation from steered molecular dynamics simulations using Jarzynski's equality.* Journal of Chemical Physics, 2003. **119**(6): p. 3559-3566.
51. Merrifield, R.B., *Solid-Phase Peptide Synthesis. 3. An Improved Synthesis of Bradykinin.* Biochemistry, 1964. **3**: p. 1385-90.
52. Made, V., S. Els-Heindl, and A.G. Beck-Sickinger, *Automated solid-phase peptide synthesis to obtain therapeutic peptides.* Beilstein Journal of Organic Chemistry, 2014. **10**: p. 1197-1212.
53. Sheppard, E.A.a.R.C., *Solid Phase Peptide Synthesis: A Practical*

Approach. 1989, Oxford: Oxford University Press.
54. W.C. Chan and P.D. White, E., *Fmoc Solid Phase Peptide Synthesis: A Practical*

Approach. 2000, Oxford: Oxford University Press.
55. Wenyi Li, N.M.O.B.-S. and a.J.D.W. Mohammed Akhter Hossain, *The 9-Fluorenylmethoxycarbonyl (Fmoc) Group in*

Chemical Peptide Synthesis – Its Past, Present, and Future. Australian Journal of Chemistry, 2019.
56. Han, L.A.C.a.G.Y., *The 9- Fluorenylmethoxycarbonyl Amino-Protecting Group.* Journal of Organic Chemistry, 1972. **37**(22): p. 6.
57. SigmaAldrich. *Fmoc chloride*. [cited 2019 December 18]; Available from: https://www.sigmaaldrich.com/catalog/product/aldrich/160512?lang=en&region=TR.
58. *Introduction to Fmoc Solid Phase Peptide Synthesis*. [cited 2019 March 16]; Available from: www.peptideinstruments.com.
59. Coin, I., M. Beyermann, and M. Bienert, *Solid-phase peptide synthesis: from standard procedures to the synthesis of difficult sequences.* Nature Protocols, 2007. **2**(12): p. 3247-3256.
60. Kitas, E.A., et al., *Alternative Strategies for the Fmoc Solid-Phase Synthesis of O-4-Phospho-L-Tyrosine-Containing Peptides.* Helvetica Chimica Acta, 1991. **74**(6): p. 1314-1328.
61. Santini, R., M.C. Griffith, and M. Qi, *A measure of solvent effects on swelling of resins for solid phase organic synthesis.* Tetrahedron Letters, 1998. **39**(49): p. 8951-8954.
62. Ralhan, K., V.G. KrishnaKumar, and S. Gupta, *Piperazine and DBU: a safer alternative for rapid and efficient Fmoc deprotection in solid phase peptide synthesis.* Rsc Advances, 2015. **5**(126): p. 104417-104425.
63. Crowley, T.E., *Chapter 5 - High-performance liquid chromatography*, in *Purification and Characterization of Secondary Metabolites*. 2020, Academic Press. p. 10.
64. Group, T.L. *High Performance Liquid Chromatography*. [cited 2019 December 19]; Available from: https://www.lindeus.com/en/industries/laboratories/analysis_and_measurement/high-performance_liquid_chromatography/index.html.
65. *High Performance Liquid Chromatography*. [cited 2019 December 20]; Available from: https://www.chemguide.co.uk/analysis/chromatography/hplc.html.

66. Carr, D., *A Guide to the Analysis and Purification of Proteins and Peptides by Reversed-Phase HPLC*, A.H. Columns, Editor., ACE HPLC.
67. *Improved Reversed-Phase Peptide Separations on High-Performance Silica Particles*, S. Aldrich, Editor. 2010.
68. Rubini, C., et al., *Mechanistic studies of amide bond scission during acidolytic deprotection of Pip containing peptide.* J Pept Sci, 2008. **14**(8): p. 989-97.
69. Tang, Y.J., X.Q. Zeng, and J. Liang, *Surface Plasmon Resonance: An Introduction to a Surface Spectroscopy Technique.* Journal of Chemical Education, 2010. **87**(7): p. 742-746.
70. *The Advantages of SPR Technology*. [cited 2019 December 20]; Available from: https://www.reichertspr.com/blog/2015/10/28/the-advantages-of-spr-technology/.
71. Oda, M., et al., *Kinetic analysis of DNA binding by the c-Myb DNA-binding domain using surface plasmon resonance.* Febs Letters, 1999. **454**(3): p. 288-292.
72. Akimoto, T., et al., *Effect of incident angle of light on sensitivity and detection limit for layers of antibody with surface plasmon resonance spectroscopy.* Biosensors & Bioelectronics, 2000. **15**(7-8): p. 355-362.
73. *Principle and Protocol of Surface Plasmon Resonance*. [cited 2019 December 20]; Available from: https://biosensingusa.com/technologies/surface-plasmon-resonance/surface-plasmon%20resonance-work/.
74. *Surface Plasmon Resonance*. 2006 [cited 2019 December 20]; Available from: https://www.sprpages.nl/spr-overview/spr-theory.
75. *Biacore Assay Handbook*, G.H.L. Sciences, Editor., GE Healthcare.
76. *Amine Coupling*. 2006 [cited 2019 December 20]; Available from: https://www.sprpages.nl/immobilization/immobilization-procedures/amine.
77. *Pre-Concentration*. 2006; Available from: https://www.sprpages.nl/immobilization/ligand-pre-concentration.
78. Symons, J.A., J.A. Eastgate, and G.W. Duff, *Purification and Characterization of a Novel Soluble Receptor for Interleukin-1.* Journal of Experimental Medicine, 1991. **174**(5): p. 1251-1254.
79. Eiji Ishii, Y.E.a.R.U., *Protein-peptide Interaction by Surface Plasmon Resonance.* Bio-Protocol, 2013. **3**(3): p. 7.
80. Karlsson, R., *Real-Time Competitive Kinetic-Analysis of Interactions between Low-Molecular-Weight Ligands in Solution and Surface-Immobilized Receptors.* Analytical Biochemistry, 1994. **221**(1): p. 142-151.
81. Karlsson, R., et al., *Biosensor analysis of drug-target interactions: Direct and competitive binding assays for investigation of interactions between thrombin and thrombin inhibitors.* Analytical Biochemistry, 2000. **278**(1): p. 1-13.