REPUBLIC OF TURKEY

BAHÇEŞEHİR UNIVERSITY

MACHINE LEARNING ALGORITHMS AND COMBINED MULTI-SCALE MOLECULAR MODELING SIMULATIONS AGAINST NADPH OXIDASE (NOX) ENZYMES FOR DESIGNING OF SMALL MOLECULE THERAPEUTICS

Master's Thesis

ASENA HİMMETOĞLU

İSTANBUL, 2020

THE REPUBLIC OF TURKEY

BAHÇEŞEHİR UNIVERSITY

GRADUATE SCHOOL

MACHINE LEARNING ALGORITHMS AND COMBINED MULTI-SCALE MOLECULAR MODELING SIMULATIONS AGAINST NADPH OXIDASE (NOX) ENZYMES FOR DESIGNING OF SMALL MOLECULE THERAPEUTICS

Master's Thesis

ASENA HİMMETOĞLU

SUPERVISOR: Prof. Dr. SERDAR DURDAĞI

CO-SUPERVISOR: Assoc. Prof. YEŞİM NEĞİŞ

İSTANBUL, 2020

THE REPUBLIC OF TURKEY

BAHÇEŞEHİR UNIVERSITY

GRADUATE SCHOOL

Name of the thesis: MACHINE LEARNING ALGORITHMS AND COMBINED MULTI-SCALE MOLECULAR MODELING SIMULATIONS AGAINST NADPH OXIDASE (NOX) ENZYMES FOR DESIGNING OF SMALL MOLECULE THERAPEUTICS
Name/ Last Name of the Student: Asena HİMMETOĞLU
Date of the Defense of Thesis: 17/09/2020
The thesis has been approved by the Graduated School.

Director Signature

Assoc. Prof. Burak KÜNTAY

This is to certify that we have read this thesis and we find it fully adequate in scope, quality and content, as a thesis for degree of Master of Sciences.

Examining Committee Members

Thesis Supervisor

Prof. Dr. Serdar DURDAĞI ……

Thesis Co-Supervisor

Assoc. Prof. Yeşim NEĞİŞ ……

Member

Assist. Prof. Berna DOĞAN ……

Member

Assist. Prof. Mert GÜR …….

Member

Assist. Prof. Özge ŞENSOY …….

ACKNOWLEDGEMENTS

First of all, I would like to thank Dr. Timuçin Avşar, who opened my way for a chance to take part in Bahçeşehir University Computational Biology and Molecular Simulations Laboratory, and Dr. Yeşim Neğiş and Dr. Serdar Durdağı for being able to take part in this laboratory. I could not imagine more than one year of experience I gained in the laboratory a year ago.

I am grateful to Dr. Serdar Durdağı for his technical and academic support. Dr. Serdar Durdağı drew a magnificent and gigantic scene that I could take part in, I learnt to be disciplined and principled with our work with him.

I would also like to extend my thanks to Dr. Yeşim Neğiş who provided the most support in being motivated, she was the knight in this chess play. Whenever I had a hard time she became a great mentor.

I am also grateful to Dr. Türker Kılıç who has remembered us with his attention and support, motivating us.

I would also thank to all laboratory members, especially; İsmail Erol, Busecan Aksoydan, Berna Doğan, Lalehan Oktay, İlayda Tolu and my project partner Saima İkram.

İsmail Erol was like a queen I changed to pawn in chess. He always helped me, never made me feel like a pawn with limited movement. And my labmates made me move widely like a bishop in chess. I would like to thank all the laboratory members for being part of the whole with such joy and learning.

Finally, of course I would like to thank my family for this chess play, I would not have been able to play without them.

ABSTRACT

MACHINE LEARNING ALGORITHMS AND COMBINED MULTI-SCALE MOLECULAR MODELING SIMULATIONS AGAINST NADPH OXIDASE (NOX) ENZYMES FOR DESIGNING OF SMALL MOLECULE THERAPEUTICS

Asena Himmetoğlu

Neuroscience Master Program

Thesis Supervisor: Prof. Dr. Serdar DURDAĞI
Thesis Co-Supervisor: Assoc. Prof. Yeşim NEĞİŞ

September 2020, 74 pages

NOX is a multi-subunit enzyme that catalyzes the reduction of molecular oxygen to form superoxide. There are 7 isoforms in mammals among which NOX2 is the one that is highly expressed in phagocytes, but expression has also been detected in neurons, microglia, endothelial cells, cardiomyocytes and vascular smooth muscle cells. While superoxide produced by NOX2 is vital to the immune system, cellular signaling, regulation of gene expression and cell differentiation, increased NOX activity promotes to a wide range of pathological processes including neurotoxicity, neurodegeneration, cardiovascular diseases and organ failure. Thus, modulation of NOX2 activity may prove to be a promising treatment for neurodegenerative diseases as well as an adjunctive agent to prevent its secondary complications. Throughout the thesis study, action mechanism is investigated using multi-scale molecular modeling simulations techniques and small molecule therapeutics are screened and designed using target-driven drug screening approaches.

Keywords: Molecular Docking, Molecular Dynamics (MD) Simulations, NOX2, AutoQSAR

ÖZET

NADPH OKSİDAZ (NOX) ENZİMLERİNE KARŞI KÜÇÜK MOLEKÜL TERAPÖTİKLERİNİN TASARIMI İÇİN MAKİNE ÖĞRENMESİ ALGORİTMALARI ve KOMBİNE ÇOK ÖLÇEKLİ MOLEKÜLER SİMÜLASYONLAR

Asena Himmetoğlu

Sinirbilim Yüksek Lisans Programı

Tez Danışmanı: Prof. Dr. Serdar DURDAĞI
Tez Eş Danışmanı: Doç. Dr. Yeşim NEĞİŞ

2020, 74 sayfa

NOX süperoksit oluşturmak için moleküler oksijenin azaltılmasını katalizleyen çok alt birimli bir enzimdir. Memelilerde bu enzim için yedi isoform bulunur. NOX2 bugüne kadar en çok çalışılmış izoformdur ve fagositlerde yüksek oranda eksprese edilir ancak nöronlar, mikroglia, endotel hücreleri, kardiyomiyositler ve vasküler düz kas hücrelerinde de tespit edilmiştir. NOX2 enzimi tarafından üretilen süperoksit, bağışıklık sistemi, hücresel sinyalleşme, gen ekspresyonunun düzenlenmesi ve hücre farklılaşması için hayati öneme sahiptir. Ancak artan NOX aktivitesi, nörotoksisite, nörodejenerasyon, kardiyovasküler hastalıklar ve organ yetmezliği gibi çok çeşitli patolojik süreçlere katkıda bulunur. Bu nedenle NOX2 aktivitesinin modülasyonu nörodejeneratif hastalıklar için umut verici bir tedavi ve ikincil komplikasyonları önlemek için önemli bir yardımcı ajan olabilir. Tez çalışması boyunca, etki mekanizması; çok ölçekli simülasyon teknikleri kullanılarak incelenmiş küçük molekül terapötikleri taranarak hedefe yönelik ilaç tarama yaklaşımları kullanılarak tasarımları yapılması hedeflenmiştir.

Anahtar Kelimeler: Moleküler Kenetlenme, Moleküler Dinamik (MD) Simülasyonları, NOX2, AutoQSAR

CONTENTS

TABLES

FIGURES

ABBREVATIONS

2D	: Two Dimensional
3D	: Three Dimensional
ALA	: Alanine
ARG	: Arginine
ASN	: Asparagine
ASP	: Aspartic Acid
C-TERMINAL	: Carboxy-Terminal
CADD	: Computer Aided Drug Design
D	: H-bond Donor Site
FDA	: Food and Drug Administration
GBSA	: Generalized Born Solvent Accessible Surface Area
GLU	: Glutamine
H	: Hydrophobic Site
HIS	: Histidine

HTVS	: High Throughput Virtual Screening
IFD	: Induced Fit Docking
L	: Ligand
LBDD	: Ligand Based Drug Design
LYS	: Lysine
MD	: Molecular Dynamics
MET	: Methionine
MM	: Molecular Dynamics
N	: Negatively Charged
N-TERMINAL	: Amino-Terminal
OPLS	: Optimized Potential for Liquid Simulations
P	: Positively Charged
PDB	: Protein Databank
PHE	: Phenylalanine
P	: Protein
PRO	: Proline

QM	: Quantum Mechanics
QSAR	: Quantitative Structure Activity Relationship
RMSF	: Root Mean Square Fluctuation
SBDD	: Structure Based Drug Design
SER	: Serine
SP	: Standard Precision
THR	: Threonine
TRP	: Tryptophan
TYR	: Tyrosine
VAL	: Valine

1. INTRODUCTION

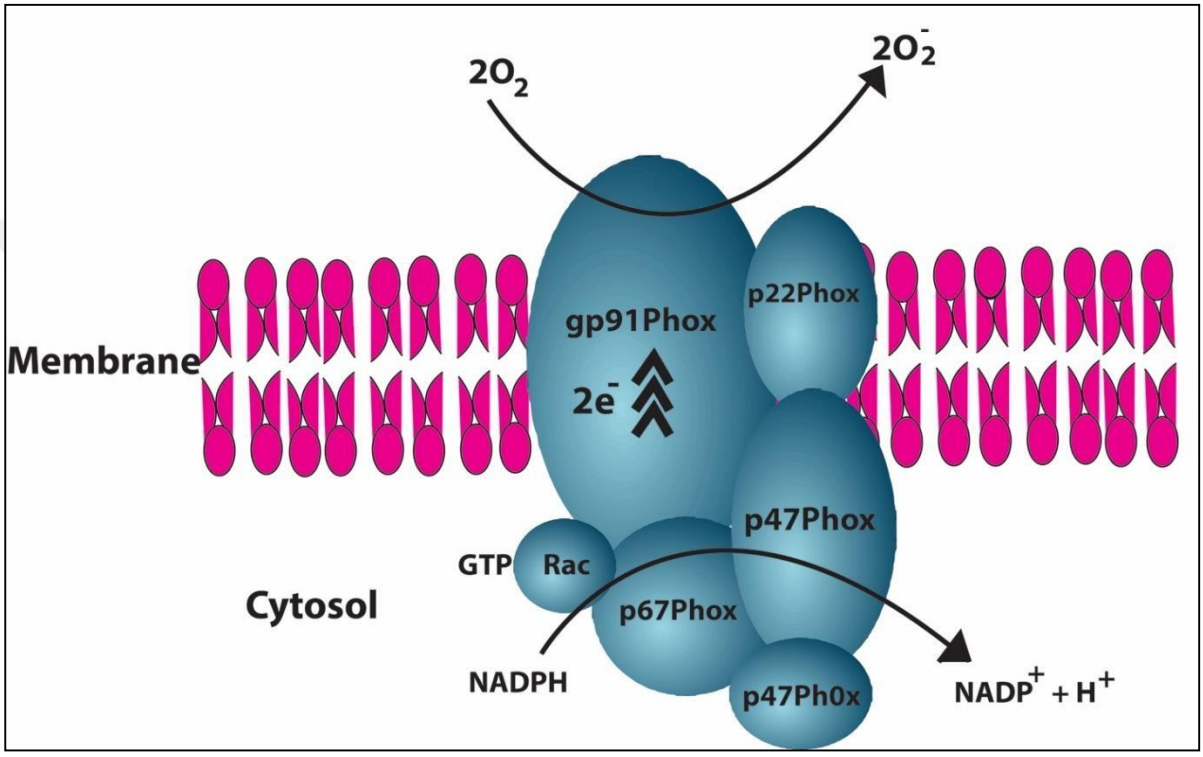
NADPH oxidase (NOX) family of enzymes are multi-subunit enzyme complexes that transport electrons across biological membranes. The electron acceptor is oxygen and the product of the electron transfer reaction is superoxide ($O_2^{-\cdot}$) or hydrogen peroxide (H_2O_2). Reactive oxygen species (ROS) are small, reactive molecules derived from oxygen that are used as cell intermediates at low concentrations. One of the most important physiological functions of ROS produced by NOXs is the destruction of foreign microorganisms and abnormal cells in phagocytes (Bedard & Krause 2007). The second effect is on signal transmission pathways; $O_2^{-\cdot}$ and H_2O_2 interact with proteins, change their function, position, half-life, regulate gene expression and many other various fundamental processes such as cell development, differentiation, apoptosis, etc. (Brown & Griendling 2009). In addition to NOXs, other enzymes such as xanthine oxidase, cytochrome P450 oxidases, mitochondrial electron transport chain enzymes, oxidases may produce ROS, but in all these systems ROS production occurs either as a by-product of the key catalytic functions of these enzymes or as a result of damage to the function of enzymes. Whereas, producing ROS is the primary and only feature of NOXs (Di Meo et al. 2016).

NOXs belong to the NOX enzymes family that has seven members, called NOX1, NOX2 ($gp91_{phox}$), NOX3, NOX4, NOX5, DUOX1 and DUOX2. These homologues have been identified in the mammalian genome and plants and some lower organisms also express NOX enzymes. The distribution of the isoforms varies between species and tissues. Although NOX family enzymes resemble each other structurally and functionally, there are differences in how they form functional enzymes complexes and the products produced. NOX1, NOX2, NOX3 and NOX5 generate $O_2^{-\cdot}$ while NOX4, DUOX1, and DUOX2 produce H_2O_2.

NOX enzyme complexes are comprised of at least seven membrane-bound and cytosolic proteins in which cytosolic subunits are called activators or regulators. While activator proteins are involved in directly modulating catalytic activity, regulatory proteins determine the structural requirements required for activation. All NOX family members have a common

and conserved catalytic subunit, called NOX. This region consists of six helical transmembrane segments, two iron-heme regions as prosthetic sections, and the cytosolic C-terminal binding sites FAD and NADPH. NOX1–4 isoforms include the $p22_{phox}$ membrane protein, which is absent in NOX5, DUOX1 and DUOX2. $p22_{phox}$ plays a role in the stabilization of the NOX enzyme complex and the assembly of other regulatory subunits (Brandes et al. 2014). The protein components are located separately in the inactive process and when stimulated by a stimulus, the subunits combine to form a functional enzyme complex (Rastogi et al. 2017). Among the members of the NOX family, NOX2 is the most studied isoform in terms of its biochemical properties and constitutes the prototype of all NOX enzymes. The enzyme complex consists of membrane-bound NOX2 (also known as $gp91_{phox}$) and $p22_{phox}$; cytosolic $p47_{phox}$ (regulator), $p67_{phox}$ (activator), $p40_{phox}$, and Rho-family GTP-binding Rac proteins (Figure 1.1). When the enzyme is activated, GDP-GTP exchange stimulates the cytosolic Rac, after cytosolic Rac stimulated the $p47_{phox}$ phosphorylation induces a conformational transition, and together with the other subunits, locates to the membrane and interacts with $p22_{phox}$ to form the active NOX2 enzyme complex and with electron transfer, $O_2^{\bullet-}$ is formed (El-Benna et al. 2009).

Figure 1.1: Schematic representation of NOX2 activation, GTPstimulates the cytosolic Rac, and $p47_{phox}$ phosphorylation induces a conformational transition, then together with the other subunits, they locate to the membrane and interact with $p22_{phox}$ to form the active NOX2 enzyme complex



NOX family members differ narrowly across function, regulation and localization (Sumimoto et al. 2005). Due to the subunits with which NOX1 interacts, it is constitutively active. With the homologues NOXO1 and NOXA1 instead of p47-phox and p67-phox, respectively, it is found to contribute to its constitutive activity as, unlike p47-phox, NOXO1 does not require phosphorylation to translocate and enable NOX1 (Sumimoto et al. 2005; Takeya et al., 2003). NOX3 is dependent upon p22phox, just like NOX1 and NOX2, but it does not bind to Rac. p22phox and POLDIP2 are involved in activation of NOX4. At N-terminus NOX5, DUOX1 and DUOX2 have calcium-binding regions that make them different from other NOXs. DUOX1 and DUOX2 have a domain that has a structure similar to that of the active site (Skonieczna et al. 2017).

Increased ROS due to NOX overactivation is documented to be involved in the pathogenesis of diabetes, myocardial infarction, nephropathy, hypertension, aneurysm and neurodegenerative diseases (Forrester et al. 2018). As a significant producer of central nervous system ROS, NOX also contributes significantly to the pathogenesis of neurodegenerative diseases (Tarafdar & Pula, 2018). Crucial molecular targets for the bioactive agents ingested by diet to prevent and treat degenerative diseases have been reported to be NOXs. The development or redox signaling network of specific and non-toxic NOX inhibitors can provide useful therapeutic strategies for the treatment of processes linked to oxidative stress such as cancer and other degenerative diseases (Altenhöfer et al. 2015; Jaquet et al. 2009).

Preserving the integrity and effective functioning of the nervous system is a very complicated process which involves multiple mechanisms to operate in a coordinated manner. In this process, a number of molecules play important roles, one of the important molecules important for neuron function is vitamin E. Vitamin E is present in nature in 8 different types ; α, β, δ, γ tocopherols and tocotrienols. α-tocopherol is the most common type found in nature (Brigelius-Flohé & Traber 1999). In humans, vitamin E deficiency causes neuromuscular abnormalities characterized by spinocerebellar ataxia and is called AVED (Ataxia with vitamin E deficiency). Cerebellar atrophy and motor neuron damage are observed in these patients (Yokota et al. 2000; Sokol 1988). It has been reported that neurological symptoms that develop due to vitamin E deficiency can be prevented by vitamin E supplementation and, in some cases could be reversible (Meydani 2001). It has been shown that the vitamin E stores in the brains of mice fed with a diet lacking vitamin E for a long time are depleted later compared to other tissues. This finding emphasizes that vitamin E plays a critical role in normal neuron development and function (Cuddihy et al. 2008).

In the regulation of inflammatory signal transduction, ROS plays a major role. In the pathogenesis of many inflammatory illnesses, ROS plays a central role. In response to inflammatory stimuli proinflammatory cytokines are synthesized and released from cells. ROS released both from cytosol and mitochondria plays a significant role in the release of proinflammatory cytokines, redox responsive transcription factors (such as NF-κB and HIF-

1α) and NLRP3 (NLR family pyrin domain containing 3) (Forrester et al., 2018; Mittal et al. 2014).

Intensive studies are carried out for developing agents that specifically inhibit NOX enzymes. In experimental studies, although various substances are used to suppress the activity of NOX enzymes, it has been reported that these substances do not exhibit NOX specific effects, their bioavailability is poor, and they cause highly toxic side effects. Potential pharmaceuticals in inhibiting NOX enzymes and their mechanism of action are still under investigation. The peculiarly regulated activation mechanism of NOX enables the specificity of cytosolic subunit-controlled NOX isoforms to be increased. Furthermore, it is possible to develop inhibitors that specifically target sequences specific to NOX subunits to understand the biochemical interactions between subunits (Altenhöfer et al. 2015; Cifuentes-Pagano et al. 2015).

1.1 THEORETICAL BACKGROUND

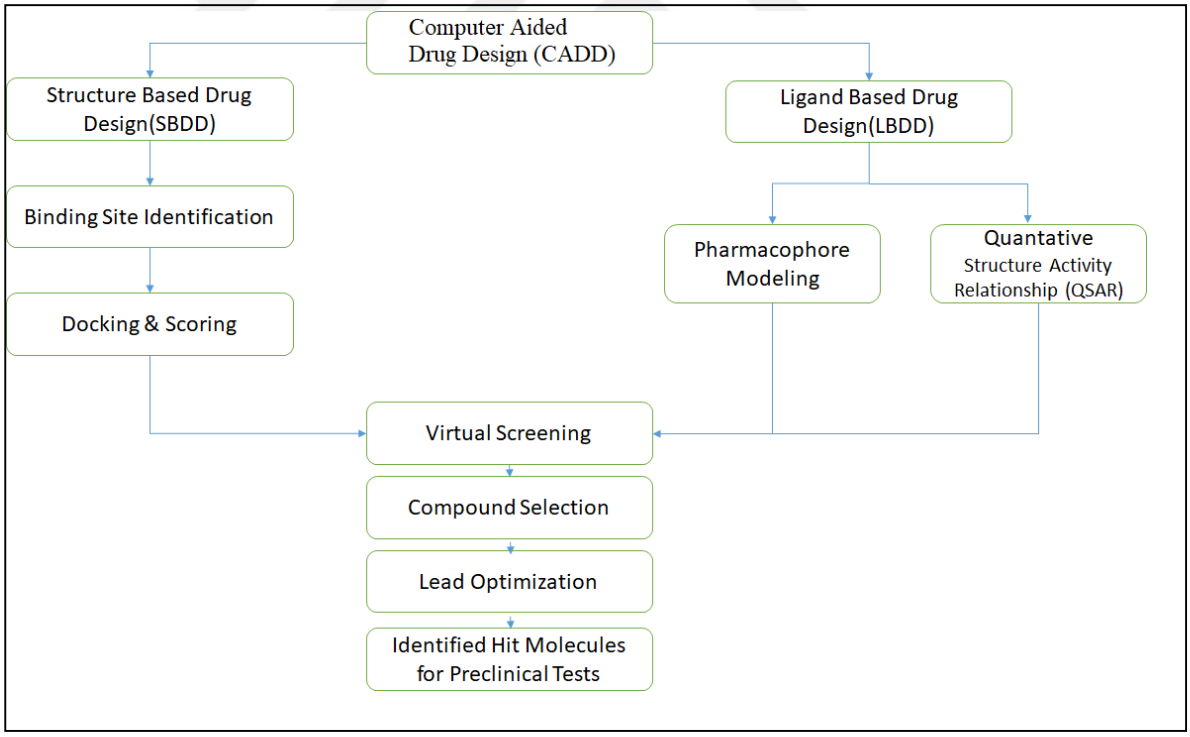
1.1.1 Computer-Based Drug Design

Introducing a new drug into the market is a complicated and expensive process from time, energy, and labour point of view. In general, the drug research and development process take 10-14 years and more than USD 1 billion of funding in total (Prieto-Martínez et al. 2019). There are four steps that the novel drug candidate must move through during this conventional, costly process. First step is the discovery work; candidate molecules are chosen to target either a gene or protein with a drug. It is then confirmed for the specified disease in question. After that, preclinical testing comes for drug development, which is divided into two *in vitro* and *in vivo*. The interactions of drug molecules are studied *in vitro*. The drug molecules are tested *in vivo*, i.e. on animal models and other living organisms. While efficacy tests begin in this phase, safety is of utmost importance because the Food and Drug Administration (FDA) does not allow preclinical studies to progress into human trials without comprehensive safety data. The number of drug-molecule candidates could be reduced from thousands to a few after this stage depending on the safety data. Before starting the human clinical trials, candidate drug must be investigated by the responsible organizations (i.e., FDA) for the analysis of the outcomes of the preclinical testing, investigating side effects and other safety features of an experimental drug, inspecting the chemical structure and mechanism of drug action. If the responsible organizations approve the investigational, experimental product, then clinical trials could be started. Phase I is the first phase of human clinical trials which requires a relatively small group of healthy individuals. Phase I focuses on safety while developers were able to see early signs of effectiveness. If Phase I is completed, the clinical Phase II studies will begin. In this stage, it is applied to one hundred or more patients suffering from the disease in question. The focus remains on safety in Phase II trials with close monitoring of short-term impact, though that priority will start to be put upon whether or not drug performs as intended. If the concepts are related, Phase III will follow. Security continues to be a priority in Phase III, but efficacy also plays a critical role. Phase III trials include over a few hundreds to thousand patients suffering from the disease

for which drug is being developed. Phase IV should be also completed for long-term protection studies following Phase III.

Today, computational approaches have become popular for implementation in the drug design, discovery, and production processes. Whereas the discovery and development of a drug are known for its difficulty that takes a lot of time and money, solutions with computer-aided drug design (CADD) minimize time, expense, and risk-borne factors. It has been shown that using CADD will reduce costs by around 50%. There are two primary forms of CADD; (i) direct, also known as structure-based drug design, (ii) indirect, also known as ligand-based drug design (Prieto-Martínez et al., 2019) (Figure 1.2).

Figure 1.2: Workflow of Computer-Aided Drug Design



Adapted from (Macalino et al. 2015).

1.1.1.1 Structure-Based Drug Design

When investigating the protein-ligand binding process, structural computational methods are commonly applied. If an experimental structure of the protein is not available, theoretical approaches such as homology modeling, threading, or *ab initio* prediction allow structural models to be created, which can be used to predict protein-ligand binding (Du et al. 2016). Otherwise, if data on the experimental protein structure is released, its positional coordinates can be obtained and acquired through the protein data bank (PDB) ("The Protein Data Bank." 2003). If the protein is obtained in holo form, i.e., with ligand molecule in its active site, then the information from its active site can be used to dock with other small molecules. Information on the binding site can be obtained from computational methods and algorithms or experimental mutation studies if the protein acquired is in the apo form, i.e. the form that has no ligand with it.

The structure-based drug design is an iterative process that often continues through multiple cycles before an established lead could reach phase I of clinical trials. In the first cycle, computer algorithms are used to position compounds or fragments from a database into a selected active region of the target structure. Such compounds are scored and classified on the basis of their steric and electrostatic interactions with the target site, and biochemical assays are used to test the selected best compounds. In the second cycle, to increase the potency the complex structure of target with a promising lead from the first step, one with at least micromolar inhibition *in vitro*, new fragments could be added to the promising compound such that interactions between target and compound could be enhanced. Anderson describes additional cycles include synthesizing the optimized lead, deciding the new target's structure: lead complex, and further optimizing the lead compound; after several cycles of the drug design process (Anderson 2003).

In terms of structural drug design, the most common method is molecular docking, utilized since the early 1980s. For molecular docking studies, programs based on different algorithms have been developed, which have made docking an increasingly attractive tool in

pharmaceutical research. The docking process involves two initial steps: prediction of the ligand conformation, including its position and orientation at the target binding site within certain poses and evaluation of the binding affinity, i.e. scoring of the binding poses. These two steps apply the methods of sampling and scoring schemes. Knowing where the binding site is located before docking processes significantly increases the effectiveness of the docking. The binding site is already known in some cases before docking process. In addition, one can get information about the binding site by comparing the target protein with a family of proteins that share a similar feature or with proteins co-crystallized with other ligands. Without knowledge of the binding sites, cavity detection programs or online servers, e.g., GRID, POCKET, SurfNet, PASS, and MMC, can be used to classify sites within proteins that are putative active. Blind docking process in which the entire surface of the target protein is considered for ligand docking can also be implemented (Hassan et al., 2017; Meng et al. 2012).

Numerous docking algorithms and programs, such as DOCK, AutoDock, AutoDock Vina, FlexX, GOLD, Glide, are available. Each of these programs have their own algorithms and scoring functions to rank molecules according to their predicted binding affinities to search different conformations of ligands at the binding site. The Glide docking algorithm, which uses a hierarchical filter to predict bioactive ligand conformations at the binding pocket, was considered in this thesis.

Docking algorithms and programs include the static picture of protein-ligand complexes and do not provide physiological conditions, including temperature and pH level (Sherman et al. 2006). Since all biological processes are dynamic and occur under physiological conditions, these conditions are essential for correct representation of the existence of dynamic cellular processes. Molecular Dynamics (MD) simulations can be applied to solve this looming problem with the usage of Newton's law of motion. MD simulations help to calculate the potential energy of every atom in the complex. Potential energy term is calculated including non-bonded interactions, bond lengths and angles, and dihedral angles. MD method allows the estimation of atomic positions and offers insights into the dynamic existence of protein-

ligand complexes over a while, producing MD trajectories.

1.1.1.2 Ligand –Based Drug Design

In the case where a target molecule lacks 3D structural information, ligand-based approaches could be convenient and advantageous. Ligand-based approaches can be classified under three main divisions: similarity search, pharmacophore modeling, and QSAR.

Pharmacophores are 3D spatial representations of chemical functionalities of a molecule that are known to induce biological activity. According to the IUPAC's definition, a pharmacophore is an assembly of steric and electronic characteristics required to ensure supramolecular interactions with a particular biological target structure and trigger or block its biological response. Pharmacophore models are a preferred tool when it comes to qualitative and quantitative characterization of pharmacological activity.

Standard pharmacophoric features include hydrogen bond donors, hydrogen bond acceptors, aromatic rings, hydrophobic sites, positive ionizable groups, and negatively ionizable groups (Chandrasekaran et al. 2018). Additionally, pharmacophore models may also include directionality instructions for chemical sites they contain, such as hydrogen bonds and aromatic interaction orientation (Schaller et al. 2020). To construct a pharmacophore model of a specified biomolecular target, a set of known active compounds are selected. Then, pharmacophore features or sites are generated by superpositioning shared chemical functionalities of all the possible conformations generated from the active compounds (Schaller et al. 2020). Following this, an understanding of the binding pattern of the target in question is roughly estimated. After a pharmacophore model is constructed, it can be used for 3D chemical database searching, virtual screening, ligand profiling, *de novo* fragment design and activity prediction. The pharmacophore model will serve as a template for hit identification, and ligands with the best fit will be assumed to have higher activity (Chandrasekaran et al., 2018). In the ligand-based drug design, the unknown activestructure of the ligand molecules is to be predicted depending on the ligands with known biological activity on a specific target (Chandrasekaran et al., 2018).

In the 3D structure similarity search, the aim is to find compounds from a compound library that resembles the three-dimensional form of a molecule (Willett et al. 1998). Iyer and colleagues categorized the techniques into the following categories, based on shape interpretations: global feature-based, recognition-based manufacturing feature, graph-based, histogram-based, product-based information, and 3D object-based recognition (Iyer et al.2005). Molecular 3D similarity search techniques have focused on 3D object recognition since non-3D object-recognition-based techniques cannot provide exact matching or superimposing of shape when measuring the similarity in 3D. 2D fingerprinting approaches are faster but restricted in finding novel scaffolds. On the other hand, 3D shape recognition approaches may be capable of finding novel scaffolds but determined instability for certain shape groups (Yan et al. 2016).

Quantitative structure-activity relationship (QSAR) analysis is the most effective method due to its high and fast productivity and good hit ratio. The physicochemical and geometric features of the molecule which are called chemical descriptors are measured on different levels of molecular structure representation, ranging from 1D to nD for a QSAR model, and then compared with biological properties using machine learning approaches after collection of concerned chemogenomics data from databases and literature. After constructing and validating a QSAR model, it can be applied to predict the biological property of new compounds (Neves et al. 2018) or design a new one. As a ligand-based drug design analysis QSAR developed by Hansch and Fujita in 1964, since then it keeps its efficiency among models for finding a statistically significant correlation between the molecular structure and continuous or categorical/binary biological/toxicological property using regression and classification techniques (Cherkasov et al., 2014). Development of a QSAR model can be provided by establishing empirical, linear, or nonlinear relationships between the values of chemical descriptors and experimentally measured properties or bioactivities of those molecules, followed by applying these models to predict or design novel chemicals with desired properties.

A QSAR model should be associated with five principles: i) identified endpoint, ii) precise algorithm, iii) identified domain of applicability, iv) suitable measurements of fitness,

robustness and predictability, v) mechanical interpretation if it is possible ("OECD Princ. Corp. Gov. 2004," 2004). Today with pioneering machine learning, QSAR approaches are having benefits of it. Machine learning makes it easier to comply with above QSAR principles. The automation of routine tasks with well-defined inputs, outputs, and performance requirements has significantly improved productivity within several scientific disciplines. In the quest to achieve more productivity gains in pharmaceutical science, advances such as the development of chemical synthesis machines to build molecules using well-understood reactions and purification approaches, and the creation of synthetic pathway prediction tools are being developed. Several advances in the fields of drug discovery and QSAR modeling have been combined to achieve productivity improvements in automation. The introduction of a generally recognized, best practices approach for the creation of validated models means that the automation cycle is well established.

In recent years growing volumes of public data have become available for pharmaceutical-interest databases such as ChEMBL, PubChem, and ChemSpider in particular. This enables more endpoint models to be developed with broader applicability domains than ever before. A longitudinal analysis of the QSAR model performance over time in drug development projects has shown that updating QSAR models leads to higher accuracy and usefulness as more data becomes available (Gawehn et al. 2016). The increasingly decentralized nature of drug discovery includes QSAR models, which can be easily and quickly implemented through project groups and organizations. Lastly, evidence indicates that there is no single best set of descriptors and technologies that fits all datasets. Automation can help to determine the optimum combination of descriptors and appropriate techniques by determining the brute force of multiple possibilities in a way that is beneficial in terms of the human labour required to produce a high-quality QSAR models (Dixon et al. 2016). In this thesis, AutoQSAR from Schrödinger's Maestro suite is preferred to use. (Dixon et al. 2016).

2. LITERATURE REVIEW

2.1 NOX2 and Vitamin E Molecules

Several vitamin E forms, and their metabolites have been suggested as signaling molecules, but their physiological function still remains unknown. In this thesis, it has been aimed to investigate the effect of vitamin E isoforms and natural metabolites on NOX enzyme activity using in silico molecular modelling approaches. The structural motifs of the vitamin E family, and particularly the chroman ring portion are highly suitable for making various modifications to increase their therapeutic target biological activities. Natural vitamin E analogues may cause changes in cellular response and gene expression by either directly binding to the active site of enzymes, transcription factors and receptors involved in signal transduction or by affecting their intracellular / membrane translocation (Pein et al. 2018).

One of the most powerful micronutrients known to modulate the immune system is vitamin E. It is present in the cells of the immune system at a higher concentration than in other blood cells. In animals and humans, vitamin E deficiency has been shown to affect the normal functioning of the immune system, and this can be treated with vitamin E supplements. Vitamin E supplementation has been shown to improve the function of the immune system and reduce the risk of infection, particularly in older people, according to current dietary recommendations. The mechanisms responsible for the effect of vitamin E on the immune system and inflammation have been investigated in cell-based, preclinical and clinical intervention studies, and vitamin E has been shown to modulate T-cell activity, T-cell membrane integrity, signal transduction, and cell division, directly or indirectly, by affecting inflammatory regulators derived from immune cells (Lewis et al. 2019). When investigating the function of vitamin E in the immune system, studies focus primarily on the form of alpha-tocopherol. It has been shown that other vitamin E analogues can also have important immunomodulatory roles, including other tocopherols and tocotrienols. Similar to the effects

of vitamin E on other cells, reducing ROS and oxidative stress plays an essential role in the function of immune system, since the processes such as preserving membrane integrity, inflammation, signal transduction modulation, and cell division are highly susceptible to oxidative stress (Azzi 2019; Jiang 2014).

Studies have shown that different vitamin E isoforms and their metabolites inhibit lipid-metabolizing enzymes such as phospholipase A2, lipoxygenases (5-12-15-LOX), HMG-CoA reductase, cyclooxygenase (especially COX-2), and hence suppress the synthesis of inflammatory mediators such as eicosanoids, lipoxins, prostaglandins, thromboxanes, leukotrienes and NF-κB and JAK-STAT signal transduction pathways and through these mechanisms exhibit anti-inflammatory effects (Zingg 2015; Z. Jiang, Yin, & Jiang 2011). It has been observed that different vitamin E isoforms have different effects on modulation of enzyme activity, and γT, δT and γTE forms are more effective than others in the inhibition of lipoxygenases. It has been reported that not only isoforms of vitamin E but also products released from their metabolism have roles in the regulation of enzyme activity. It was shown that δ-carboxychromanols, which are long-chain metabolites of δ-tocopherol competitively inhibited COX-1, COX-2 and 5-LOX enzymes (Pein et al. 2018). However, because these metabolites are not observed in plasma unless a pharmacological dosage of vitamin E is supplemented, their physiological functions and mode of action are not completely established (Zingg 2019). It has been reported that α-T-13'-COOH, a metabolite of α-tocopherol is found in the serum of humans and mice at the nanomolar concentrations and selectively inhibits 5-LOX activity and hence suppresses the inflammation (Pein et al. 2018).

Besides its role in immune defense and signal transduction, NOX is an important generator of oxidative stress. Increases in lipid peroxidation, cell death, matrix metalloprotease (MMP) production, DNA damage and activation of signal transmission pathways occur if the increased ROS levels can not be balanced for by the body's antioxidative defense demands, and this will trigger cell damage. In many diseases such as diabetes and cancer where inflammation plays a role, excessive development of ROS by NOX enzymes, as well as NOX upregulation are seen. It also plays a pivotal role in neurological pathologies. As the major

ROS producer of central nervous system, NOX is significantly involved in the pathogenesis of neurodegenerative diseases and brain injury caused by ischemia/reperfusion (Rastogi et al. 2017). As a result of its widespread action, NOX has become an attractive target for therapies and drugs, but understanding the activation mechanism is crucial for designing therapeutic agents of appropriate specificity (Sorce et al. 2017; Ajayi et al. 2013).

As the phagocytic isoform of NOX and best characterized member of NOX family, NOX2 plays a role in destroying pathogens that invade the body. However, the expression of NOX2 is not limited to phagocytic cell types, it is also expressed in various tissues, including vascular smooth muscle cells, adventitial fibroblasts, cardiomyocytes and microglial cells. Increased NOX2 activity plays a vital role in pathogenesis of many diseases, particularly diseases of the cardiovascular and cerebrovascular systems. In most of these disease models, the role of NOX2 in innate and acquired immune response can help to explain the involvement of NOX2 to the disease processes. The association of NOX2 activation with many disease models is due to the fact that most of them contain a significant inflammatory component (Diebold et al. 2015). Protective effects of the inhibition of NOX2 catalytic activity have been observed in *in vivo* experiments performed in mice, however as complete suppression of NOX2 activation will suppress the immune response, particularly to pathogens (Liu et al., 2019). This should be held in mind when considering the therapeutic inhibition of NOX2.

In 2010, in the Protein Data Bank (PDB), a crystal structure of the NADPH binding domain of gp91phox (PDB 3A1F) was published by H. Sumimoto (Beaumel et al. 2017). The crystal structure of the NADPH binding domain of gp91phox 3A1F is also known as Cytochrome b-245 heavy chain. However, when the protein is searched by UniProt, it is shown that this PDB contains the sequence 385-570, and there is a shortage of specific essential components for NOX2 activation such as p47phox and p67phox (Bateman et al. 2017).

Recent NOX studies also provided a deeper understanding of the NOX complex (Nguyen et al., 2017; Panday et al. 2015). The NOX complex needs a specific activation mechanism through subunit interaction, comprising independent cytosolic subunits; p47phox, p67phox,

p40phox and Rac, and membrane subunits; gp91phox and p22phox (Figure 1.1). Especially, p47phox plays the most important role in activating, binding and translocating the cytosolic subunits to the membrane and anchoring to p22phox to coordinate the NOX activation and form functional complex. Furthermore, these interactions, particularly between p47phox and p22phox, rely on phosphorylation started by upstream processes involving protein kinase C (PKC) (Rastogi et al. 2017). Structural and biochemical data show that p47phox phosphorylation triggers a molecular switch that relieves intramolecular interaction. This helps p47phox to interact with p22phox's cytoplasmic tail and initiate the active, membrane-bound enzyme complex formation. Groemping et al. reported the crucial amino acids for the binding cavity of the p47-p22phox complex (PDB;1OV3); TRP193, TRP204, PRO206, PHE209, LEU210, GLU241, ASP243, TRP263, TYR279, ARG296, SER303 (Groemping et al. 2003). However, this structure still has missing residues. The crystal structure only contains the sequence of 156-285, however for assembly of p22, p47and p67phox 1-390 sequence interval is needed.

In the sequence of 1-390, there are three PDB in RCSB Protein Databank; 1KQ6 (1-141), 1UEC (151-340) and 1K4U (359-390) (Bateman et al. 2017). As these structures are representing the individual domains, when merging, gaps occur between the domains. In the cases where the 3D structures are missing it difficult to understand the binding specificities of protein ligands, homology modeling is increasingly becoming the tool of choice for obtaining 3D protein coordinates. Homology modeling is a description of the similarity of environmental residues at topologically corresponding locations in the reference proteins (Vyas et al. 2012). However, sometimes as in this case, 3D structures exist but there may have been missing residues in the domain. In that case, there is a method called AIDA: ab initio domain assembly server (http://aida.godziklab.org/). It is a method capable of defining domains in multi-domain proteins and predicting their 3D structures and relative spatial arrangements (Xu et al. 2015).

Understanding the biochemistry of mechanism of action and properties of the NOX enzymes will provide information about the modulation of the enzymatic activity. The specific and non-toxic inhibitors of NADPH can provide beneficial therapeutic strategies for oxidative

stress-related diseases for degenerative diseases because NOX overproduction can lead to neurotoxicity, CNS degeneration, and cardiovascular disorders. In chronic CNS diseases such as Parkinson's disease (PD), Alzheimer's disease (AD), Huntington 's disease (HD), multiple sclerosis (MS), and amyotrophic lateral sclerosis (ALS), NOX has been shown to play a critical role (Barua et al. 2019).

Because of the complex mechanisms involved in NOX activation, these enzymes can be targeted at various levels. Their gene expression can be downregulated. Also, NOX activation can be minimized by blocking the translocation of its cytosolic subunits to the membrane. Another way is inhibiting the activity of p47phox subunit, either by preventing its phosphorylation or by blocking its binding to other subunits. Diminishing signal transduction and inhibition of Rac1 translocation were also shown to reduce ROS production. A variety of compounds, including apocynin, diphenylene iodonium (DPI), and benzensulfonylfluoride 4-(2-aminoethyl) - (AEBSF) for have been used for inhibiting NOX2. However, it has become clear that these inhibitors are not unique to NOX and are not limited to specific isoforms of NOX (Maraldi 2013). Many small-molecule inhibitors have been reported to date (especially for NOX2) however, none of them showed *in vivo* inhibitory activity, nor they exhibit specificity. As such, the identification of improved NOX inhibitors is necessary in order to enable further evaluation of the biological functions of NOX enzymes as well as the therapeutic potential of NOX inhibition. In the study of Hirano et al., the pharmacological profiles of GSK2795039, a novel NOX2 inhibitor, have been characterized both in *in vitro* and *in vivo* studies, in comparison with other NOX-inhibitors that have been released so far (Hirano et al. 2015).

On the other hand, the mechanism shows vitamin E can modulate NOX activity, are not well defined but may play a role in regulating the protein expression of the enzyme at transcriptional or at post-translation levels. Vitamin E may also directly influence the biological activity of the enzymes. Hydrophobic vitamin E within the cell membrane may alter the cell membrane-associated NOX by inhibiting or interrupting the formation of NOX complex (Chen et al. 2001). Natural vitamin E analogues can cause cellular response and gene expression changes by either directly binding to or activating intracellular/membrane

translocations or activity of enzymes, transcription factors, and receptors involved in signal transduction. The term vitamin E is used for lipid-soluble compounds that are closely related to a group of molecules called tocopherols and tocotrienols. The molecules belonging to these groups contain a chromane ring containing hydroxyl and a fragment isoprenoid. Depending on the location and amount of methyl groups that are present in the chromane chain, the group members are referred to as α, β, δ and γ. Of these isoforms, α-tocopherol is the most common type in nature exhibit the highest biological activity. While tocopherols and tocotrienols have similar *in vitro* antioxidant activity, they have different *in vivo* biological effects (Brigelius-Flohé et al. 2002; Brigelius-Flohé & Traber 1999). Tocopherols, tocopherol derivatives, tocotrienols, and their metabolites control the gene expression of inflammatory, lipid metabolic, and neuroactive proteins and are selective on enzymes with major immunoregulation functions (Azzi et al. 2004). Studies show that α-tocopherol exhibits its physiological effects by explicitly modulating the expression of different genes and cellular signaling pathways unrelated with its antioxidant properties (Azzi 2007; Azzi & Stocker 2000). α-Tocopherol can be phosphorylated to α-tocopherol phosphate (α-TP) in *in vitro* and *in vivo* models. This non-antioxidant physiological derivative shows similar cellular functions to α-tocopherol, but it is generally more potent (Negis et al. 2005). It has been reported that not only isoforms of vitamin E but also metabolites which occur physiologically have anti-inflammatory effects (M. Cook-Mills & A. McCary 2012). Since NOX2 is the most critical isoform in regulating inflammatory processes, novel modulators for this isoform have always been studied. Given the widespread involvement of NOX2 activation in various disease processes, targeting NOX2 with drugs or modulators has been suggested to show a great potential to decrease the progression of various inflammatory diseases (Diebold et al. 2015). NOX isoforms have a variety of physiological functions, especially in immunity; the primary therapy goal should specifically inhibit or modulate the isoforms. Further clarification of the biochemical properties and mechanisms of action of NOX enzymes will enable specific inhibitors with lower side effects to be identified (Cifuentes-Pagano et al. 2015).

3. METHODS

3.1 AutoQSAR Numeric Modeling

The deep-learning-based AutoQSAR method enables to create predictive QSAR models on datasets of hundreds of thousands of molecules or more. The DeepChem approach to QSAR is to create a convolutional neural network – similar to those that are highly successful in image and video processing, but "convolution" is applied to atoms instead of pixels. Switching from image processing to the chemical field, DeepChem treats small molecules as graphs where nodes are atoms and edges are bonds. The chemical features of molecule (atom type, valence, charge, etc.) are added to each node in graph. The convolution, in this case, is to apply filters to neighboring atoms (as images) instead of neighboring pixels (de Souza et al. 2017; Dixon et al. 2016). In a single workflow, descriptor generation, feature selection and development of a huge number of QSAR models was automated within AutoQSAR (Dixon et al. 2016) (Figure 3.1).

Topology-based descriptors are produced from AutoQSAR and these descriptors are made out of physiochemical properties, graph-theoretical indices and functional category numbers. For almost all compounds, descriptors that measure scarce properties that take a single value, such as 0, are discarded. Because these descriptors do not contain a statistically significant amount of information. Also, AutoQSAR discards compounds when, at least, ninety percent of the compounds of learning set have the same value. A descriptor is eliminated when its variance is less than "$1-12(^{(\Sigma\ x2)}/N)$" over the learning set where Σ stands for the sum and N stands for the number of descriptors (Dixon et al. 2016).

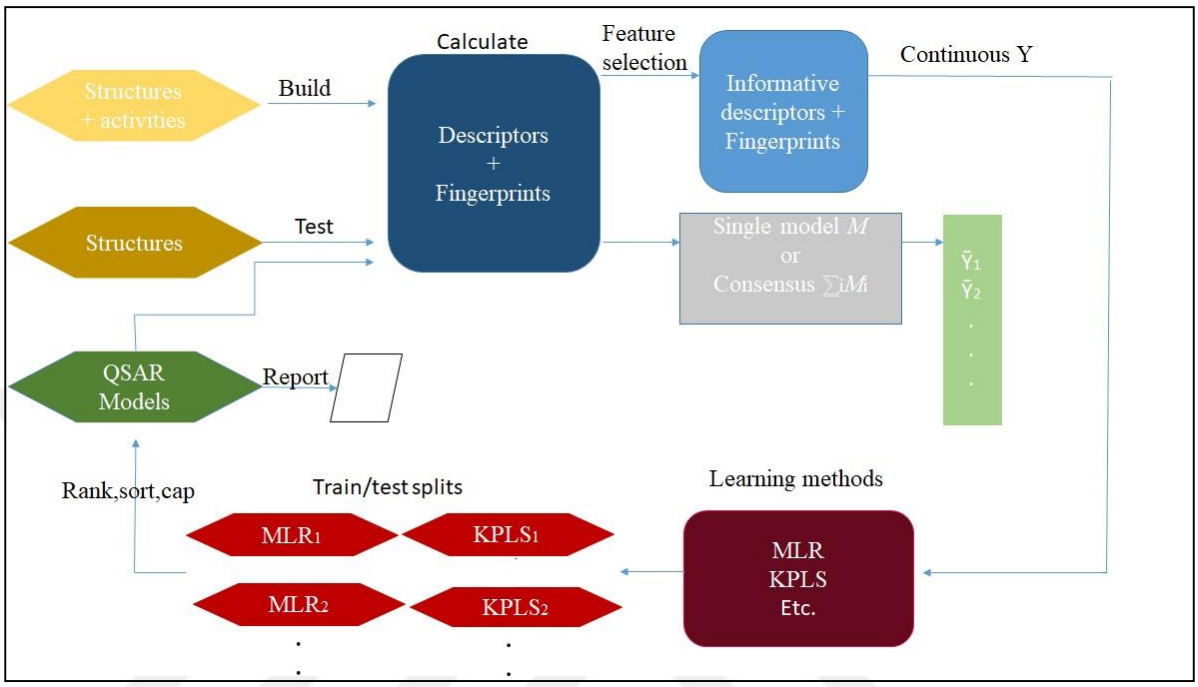
In this work, we constructed our model with numeric descriptors. To calculate conditional probabilities, a natural distribution of probability is accepted when it is an ordinary numeric descriptor. Subsequently, according to the resulted values of bits or descriptors, the compound with highest probability is distinguished.

AutoQSAR builds a powerful score function to specify the effective and weak models. By using this method, and according to the outcome obtained from the training and test set of the selected model, a quality score will be concluded. The minimum quality score is “0” and the maximum is “1”; if the score is 1 then the model made a perfect prediction; however, if the score is 0 then the predictions are completely incorrect. The score is calculated by the function of accuracy test of subtracting the absolute value of the difference between the accuracy of test set and the accuracy of training set from 1.0 (Dixon et al. 2016).

AutoQSAR generates and advances four types of 2D fingerprints; dendritic, linear, MOLPRINT2D and radial fingerprints. Dendritic fingerprint stands for linear and branched fragments, linear fingerprint stands for linear fingerprints and ring closures, MOLPRINT2D fingerprint is radial-like fingerprint that encodes atom environments using lists of atom types located at different topological distances. Radial fingerprint is fragments that grow radially from each atom, also known as extended connectivity fingerprints (Dixon et al. 2016).

To develop models, AutoQSAR uses machine-learning methods including Multiple Linear Regression (MLR), Partial Least Square regression (PLS), Principal Components Regression (PCR), and Kernel-based Partial Least Square regression (KPLS). These methods are employed against variables (continuous/ categorical-dependent). KPLS exploits both descriptors and fingerprints whereas other methods make use of only the descriptors. As a conclusion, this method produces a total of five models for a given training set of compounds (Dixon et al. 2016). Figure 3.1 demonstrates the workflow of the AutoQSAR model process for continuous variables.

Figure 3.1: Workflow of the AutoQSAR model process for continuous variables



Adapted from (Dixon et al. 2016)

To create our AutoQSAR Numeric Model, we used 15 NOX2 inhibitor molecules from the ChEMBL database (Cifuentes-Pagano et al. 2013). Random training set was selected as 70% thus ten molecules were in the training set, and five molecules were in the test set. Property type was set to Numerical. Prediction property was selected as "r_user_Binding_Affinity". Validation set was not used. Number of models to keep was selected as 10.

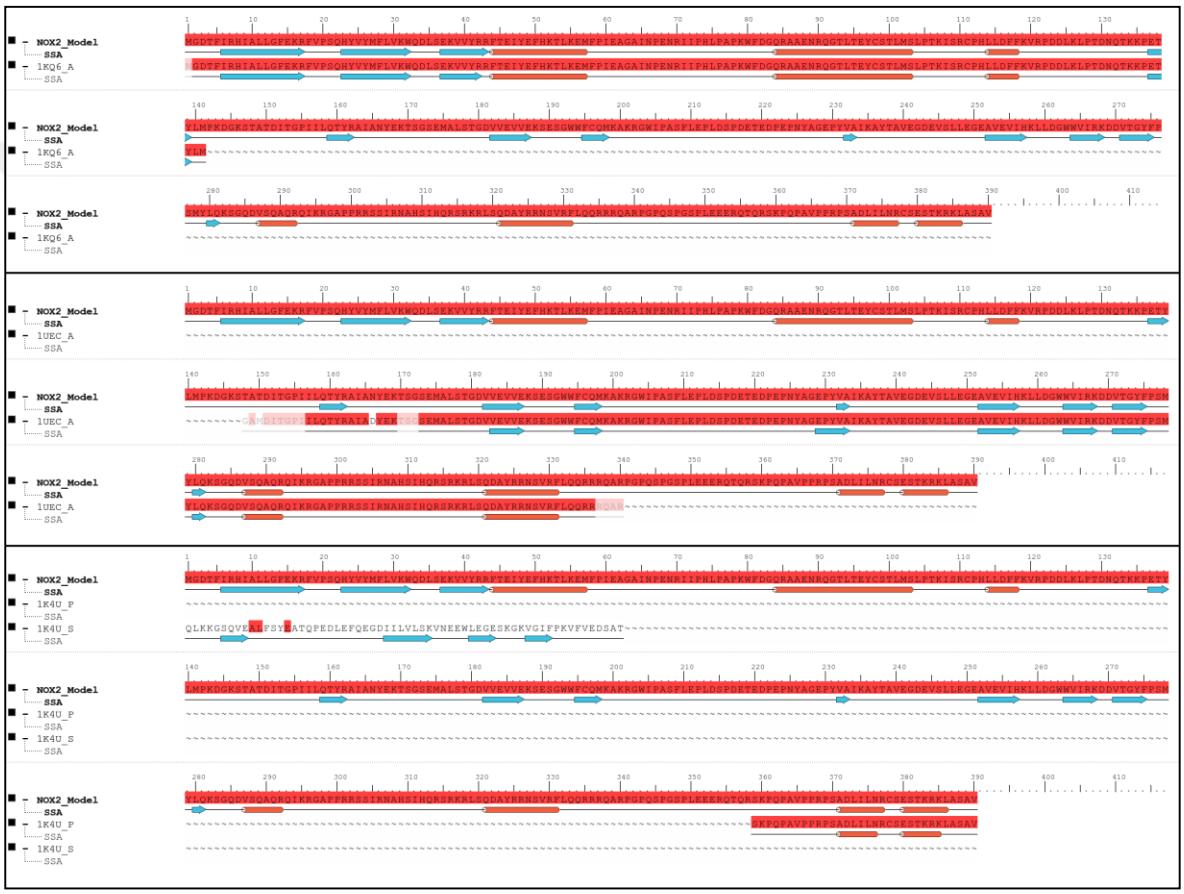
3.2 Domain Assembly

NOX2 enzyme has different components that are important for its activation, e.g. $p22_{phox}$, $p67_{phox}$ and $p47_{phox}$. However, the full structure of the enzyme complex is not determined experimentally with all components that are important for NOX2 activation. Crystal structures of components are present in protein databank individually (PDB; 1KQ6, 1UEC, 1K4U). We assemble the three domains through online AIDA (Ab Initio Domain Assembly) Server (D. Xu et al. 2015). It is an online tool that can be used for the prediction of their 3D

structure and spatial arrangements.

In NOX2, the N-terminal residues of the complex with resiues between 1-141 belongs to the $p47_{phox}$ PX domain (PDB; 1KQ6). The crystal structure of cytosolic p22 and autoinhibited form of the tandem SH3 domain of $p47_{phox}$ with residues between 151-340 (PDB;1UEC) (Yuzawa et al. 2004). It has the essential part of NOX2; $p47_{phox}$ that act as an organizer and involve in phosphorylation. The solution structure of the $p67_{phox}$ C-terminal SH3 domain complexed with the $p47_{phox}$ C-terminal tail region has redidues between 359-390 (PDB;1K4U) (Kami et al. 2002). $p67_{phox}$ is the activator of the system that interacts via its C-terminal SH3 domain with the C-terminal proline-rich region of $p47_{phox}$ which acts as an adaptor protein. NOX domains were assembled using AIDA for later validation. All domains that are present in the PDB are aligned with our NOX2 model (Figure 3.2).

Figure 3.2: Pairwise alignment of the domains that are present in protein databank and our NOX2 model, our NOX2 model has a complete "1-390" sequence and it matches up with the domains that are present in protein databank; 1KQ6, 1UEC, 1K4U



3.3 Model preparation

Assembled domain model was conducted using Schrödinger Maestro's protein preparation wizard; the process consists of adding hydrogen atoms to protein, fixing side chains and loops and developing disulfide bridges. While using PROPKA, protonation states of amino acids were produced at pH 7.4 to mimic the physiological conditions. For Protein Structural

Optimization, the OPLS2005 force field was employed (Madhavi Sastry et al. 2013; Bas et al. 2008 ; H. Li et al. 2005).

3.4 Ligand preparation

For this study, eight vitamin E isoforms, α-T-13'-COOH and α-tocopherol phosphate have been used. All molecules were prepared using Schrödinger Maestro's LigPrep module with OPLS2005 force field. The total number of molecules exceeded 64 after ligand preparation due to the generation of stereoisomers and different molecular protonation states. Epik was used to protonate the ligands at pH 7.4 (Shivakumar et al., 2010; Shelley et al., 2007).

3.5 Molecular Docking

In Maestro, various docking algorithms such as Glide / SP, Glide / XP and Glide / IFD were used for molecular docking analysis, and blind docking methods (Halgren et al. 2004; Friesner et al. 2004). Glide docking algorithms with different accuracies are available in the Maestro molecular modeling package, namely starting from the most complicated one; quantum-polarized ligand docking (QPLD)(Cho et al. 2005), induced-fit docking (IFD) (Sherman et al. 2006), extra precision (XP) (Friesner et al. 2006), Standard Precision (SP) and high-throughput virtual screening (HTVS). In Glide, there are three choices of docking precision, given under Precision in the Docking section. The HTVS module is intended to screen vast numbers of ligands as quickly as possible. HTVS has much less conformational sampling than SP docking, and may not be used with score-in-place sampling. SP module is ideal for screening large numbers of ligands of unknown nature. Docking and scoring with XP is a more robust and discriminating procedure that takes longer to run than SP. In the IFD algorithm, initial docking is performed using SP or XP mode, and then poses generated are refined using Maestro's Prime module to provide flexibility in the binding pocket for the surrounding residues (Jacobson et al. 2004). Thus it is possible to adopt more favourable orientations for active site residues to optimize ligand-protein interactions. Eventually, target structures with optimized binding site are used to re-dock and rank ligands back into the

binding pocket.

Grid box for receptor was built around the catalytic residues TRP193, TRP204, PRO206, PHE209, LEU210, GLU241, ASP243, TRP263, TYR279, ARG296, SER303 and the rotation in side chains of certain amino acids was allowed so that the structure used is flexible enough. To generate docking poses, we used flexible and rigid docking algorithms. The force field OPLS2005 was used for generating different ligand docking poses. The IFD was carried out in several steps. For the receptor and ligand respectively, (M. Xu & Lill 2013) the ligands were initially docked to rigid protein using soft-potential Glide docking with a van der Waals radii scaling of 0.50/0.50. Residues with at least one atom within 5 Å of any of the 20 ligand poses have been selected for conformational search and minimization while residues outside this range have been considered as rigid. Hence the protein versatility was considered as a cycle in this manner. Thereafter, the backbone, side chains, and ligand were subjected to energy minimization. Redocking of the ligands was performed, selected within 30.0 kcal/mol of their minimal energy structure.

3.6 Molecular Dynamics (MD) simulations

MD simulations were carried out using the Desmond (Bowers et al. 2006) program with the OPLS2005 force field and the RESPA integrator to investigate the structural and dynamic profiles of the ligands at the binding pocket of the protein (Predescu et al. 2012). Top docking poses obtained with IFD were immersed in orthorhombic water box (distance between edge box and complex set at 10 Å). The preparation of the system involved the explicit addition of water molecules (SPC solvent model), and the system 0.15 M NaCl was added to neutralize the system. The long-range electrostatic interactions were calculated by the particle mesh Ewald method (Essmann et al. 1995). A cutoff radius of 9.0 Å was used for both van der Waals and Coulombic interactions. All production MD simulations were run under NPT ensemble, at 310 K (maintained though a Nose-Hoover thermostat)(Nosé 1984) and a constant pressure of 1.01 bar (maintained through Martyna-Tobias-Klein thermostat and with isotropic pressure coupling) (Lippert et al. 2013; Martyna et al. 1994). No pressure

was exerted on the generated system, and initial velocity values are used based on Maxwell-Boltzmann distribution. The time-step was assigned as 2.0 fs.

Three separate output run times were considered in MD simulations: i) 1 ns short MD simulations, ii) 10 ns short MD simulations iii) 100 ns long MD simulations). While for 1 ns simulations, all 10 top docking complexes were considered, for 10 and 100 ns simulations 6 top docking complexes were used. Apo form and α-tocopherol bound protein form simulations were performed as positive control. The simulation box was individually prepared using an in-house script for the top 10 protein-ligand complexes. In the course of the 1 ns MD simulations, 100 trajectory frames were included and MM/GBSA calculations were performed for each these trajectory frames. The average values as well as standard deviations of MM/GBSA values of separate frames were calculated. Based on average MM/GBSA scores, six compounds were selected for longer MD simulations of 10 and 100 ns.

Ligand Root Mean Square Deviation (RMSD) was investigated with two different aspects. (1) The Lig-fit prot (2) Lig-fit Lig. Lig-fit prot measure the translational motion of the ligand and the binding pocket of protein, Lig-fit-lig measures the rotational motion of the ligand. Non-hydrogen atoms of the ligands were considered in the RMSD calculations.

3.7 Mutational analysis of active site residues

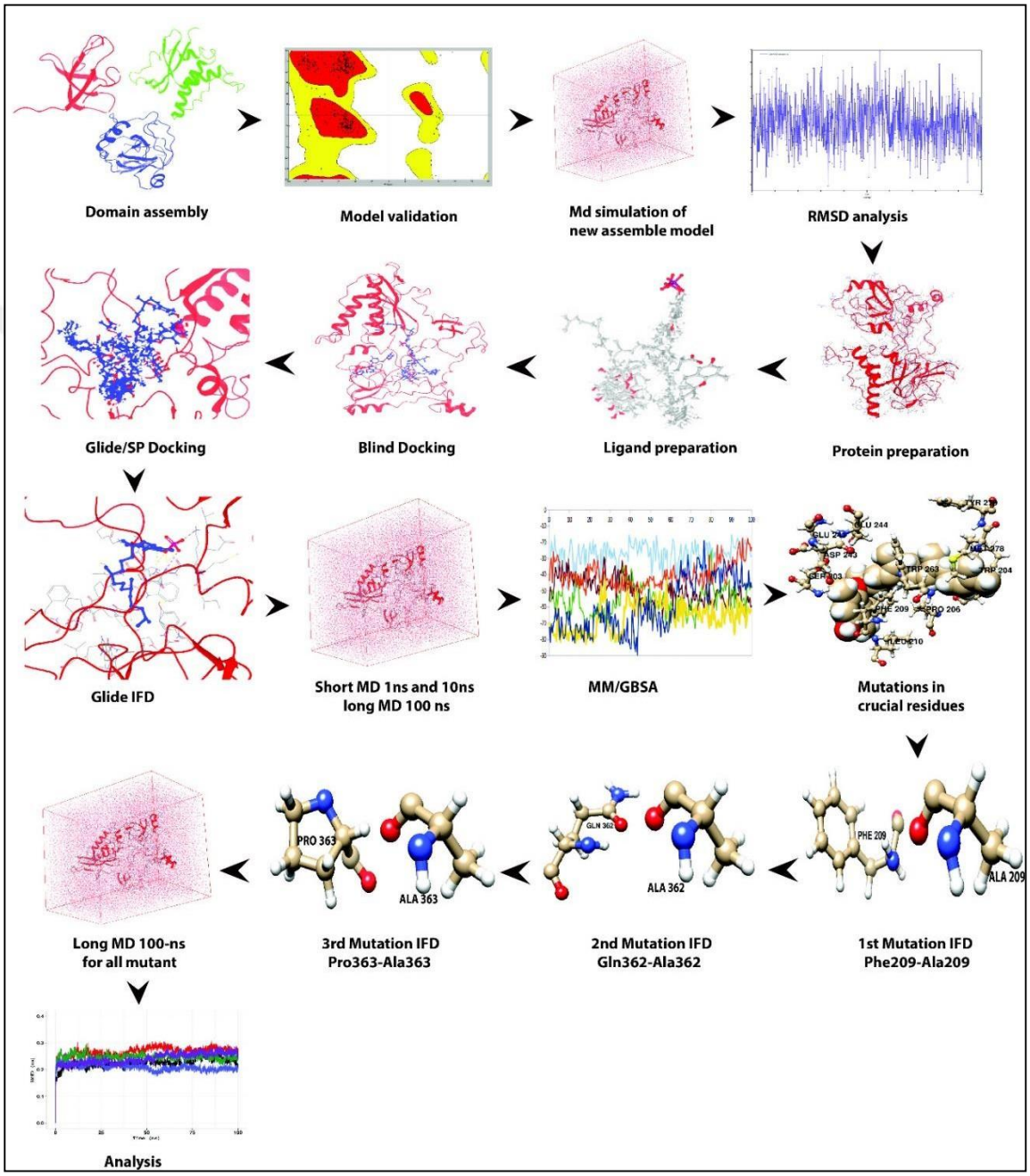
The top-scoring complex showing interactions with binding pocket residues at the end of long 100 ns MD simulation was selected to determine the effect of the mutation on the binding efficiency of the ligand in the binding pocket. PHE209, GLN362 and PRO363 were substituted with ALA. After substitution, the ligand was re-docked in the binding pocket. Glide/IFD docking algorithm was used and 100 ns MD simulations were performed for the top docking pose again.

3.8 The Molecular Mechanics/Generalized Born Surface Area (MM/GBSA) Analysis for Binding Free Energies

The ligand binding affinities were determined using the MM/GBSA system by integrating molecular mechanics and continuum solvent models (Rastelli et al. 2010). The Prime module of Schrodinger (Jacobson et al. 2004) was used in binding free energy calculations of complexes by Molecular Mechanics/Generalized Born Surface Area (MM/GBSA) approach (Bashford and Case 2000). VSGB 2.0 implicit solvation model (Li et al. 2011) were utilized during MM/GBSA calculations. In this model, on the basis of the polarization of protein side chains, a variable dielectric constant approach is used. Thus, the internal dielectric constant ranged from 1.0 to 4.0 for complex systems, while the outside of the complex system was considered to be a water environment with a dielectric constant of 80.0.

Analysis of MM/GBSA was conducted for six molecules. 100 trajectory frames were select ed for short MD simulations and 100 trajectory frames for long MD simulations, respective ly. VSGB2.0 solvation model was used to calculate the free binding energy of complexes, using OPLS2005 (J. Li et al. 2011) (Figure 3.3).

Figure 3.3: Workflow of combined multi-scale molecular modeling simulations of the study



4. RESULTS

4.1 3D QSAR Modeling by Using Machine Learning Algorithms

A numeric model is constructed with 15 NOX2 inhibitor molecules from ChEMBL database (Cifuentes-Pagano et al. 2013). 10 of the molecules were used for training set, 5 of the molecules were used for test set. The score of the model was 0.87. In AutoQSAR, a model is appointed a quality score based on its performance with respect to the selected training and test sets for that model. The accuracy is defined as a value on the interval (0,1), with 1 corresponding to perfect predictions and 0 corresponding to entirely incorrect predictions, and a model *M* is then scored as follows:

$$\text{scoreM} = \text{accuracy}_{\text{test}} (1.0 - | \text{accuracy}_{\text{train}} - \text{accuracy}_{\text{test}}|) \quad \textbf{(4.1)}$$

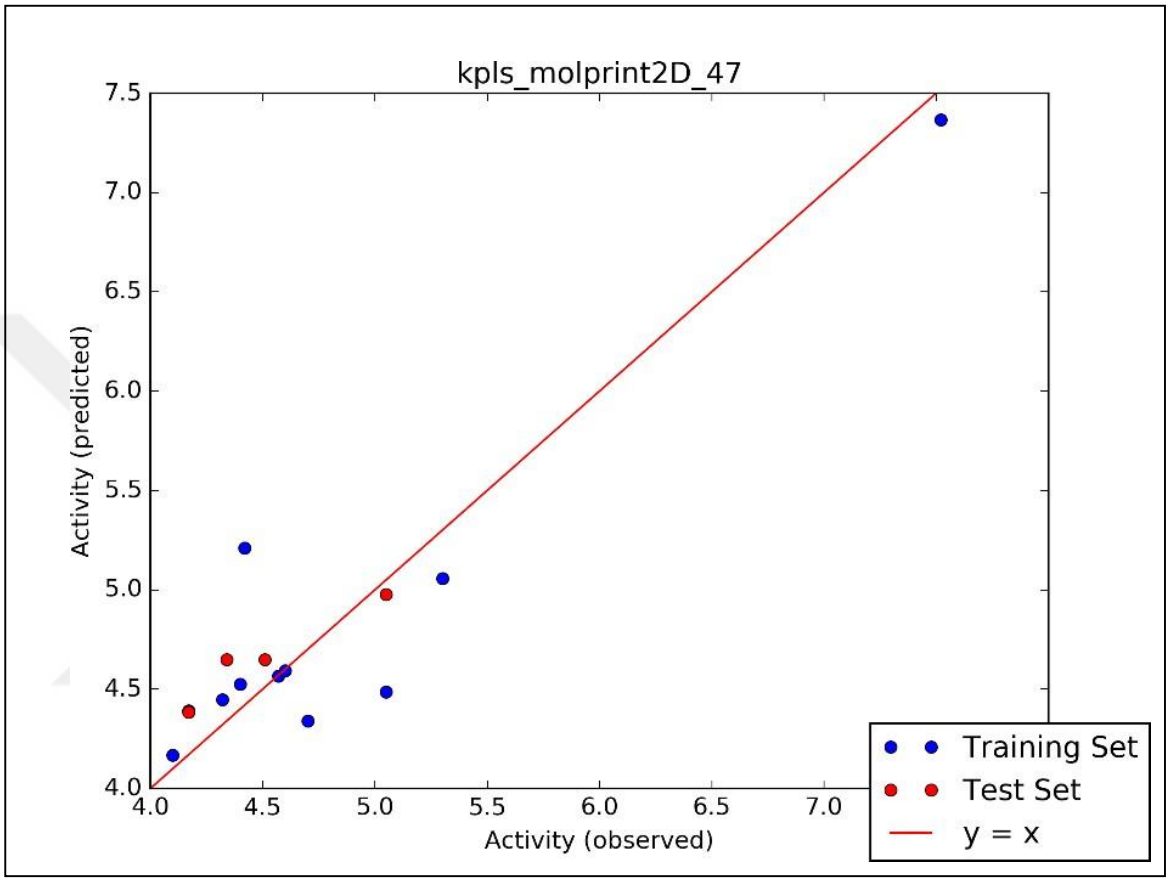
This formula rewards models that display high accuracy for both training and test sets, and penalizes models where accuracy is low for both sets, or where accuracy is very different for both sets. To evaluate the performance of the model, the statistical results were given in Table 4.1.

Table 4.1: AutoQSAR Numeric Model Results, RMSE stands for the Root Mean Square Error, and it is the standard deviation of residuals (prediction errors), residuals are a measure of the difference between the data points and the regression line; RMSE is a measure of how these residuals are distributed, it shows how the data is clustered along the best fit axis, SD stands for Standard Deviation, The standard deviation measures data spread over the mean value, R^2 is a statistical indicator that reflects the proportion of the variance for a dependent variable explained in a regression model by an independent variable or variables; it can also be referred to as the determining coefficient

SCORE	0.87
R^2	0.87
Q^2	0.62
SD	0.37
RMSE	0.20

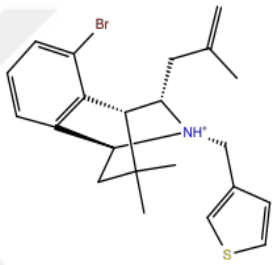
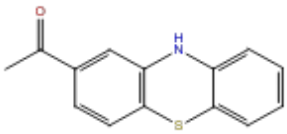
The model is constructed by using KPLS_Molprint2D_47 hybrid algorithm (Figure 4.1).

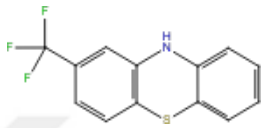
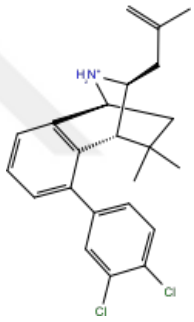
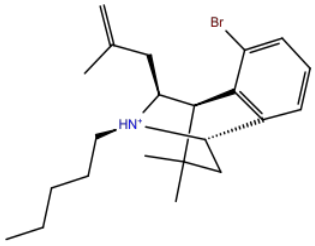
Figure 4.1: Demonstration of the internal predictions, activity predicted versus activity observed

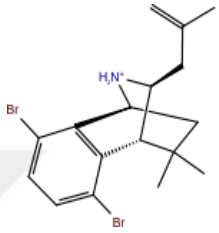
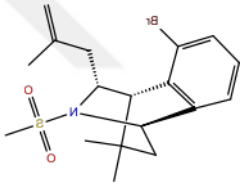
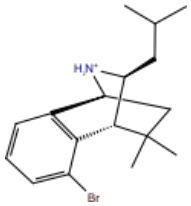


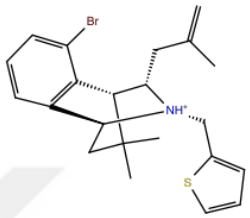
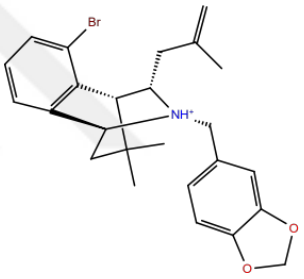
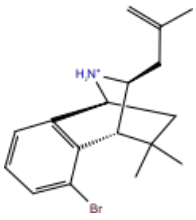
To test the model, we predicted the activity of the internal test set. The model has 0.242 error rate per ligand and has no error more than 1 logarithmic unit (Table 4.2).

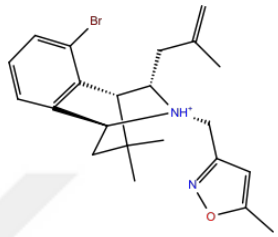
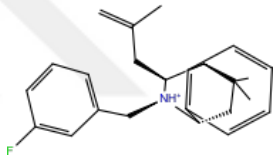
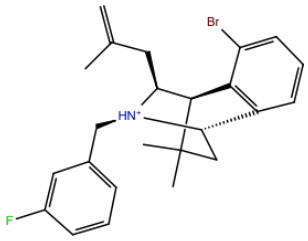
Table 4.2: Internal predictions of the dataset that was used in the model construction

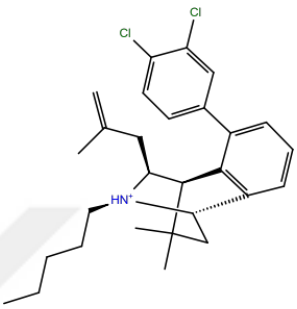
Molecule Name	Binding Affinity Experimental Values (pIC_{50})	Internal Predictions (pIC_{50})	2D Structure
CHEMBL3733336	7.52	7.25	
CHEMBL407734	5.30	4.90	

Molecule Name	Binding Affinity Experimental Values (pIC_{50})	Internal Predictions (pIC_{50})	2D Structure
CHEMBL541018	5.05	4.90	
CHEMBL3347554	5.05	4.07	
CHEMBL3347550	4.70	5.29	

Molecule Name	Binding Affinity Experimental Values (pIC_{50})	Internal Predictions (pIC_{50})	2D Structure
CHEMBL3347492	4.60	4.55	
CHEMBL3347583	4.57	4.80	
CHEMBL3347502	4.51	4.43	

Molecule Name	Binding Affinity Experimental Values (pIC_{50})	Internal Predictions (pIC_{50})	2D Structure
CHEMBL3347551	4.42	4.93	
CHEMBL3347542	4.40	4.39	
CHEMBL3347491	4.34	4.36	

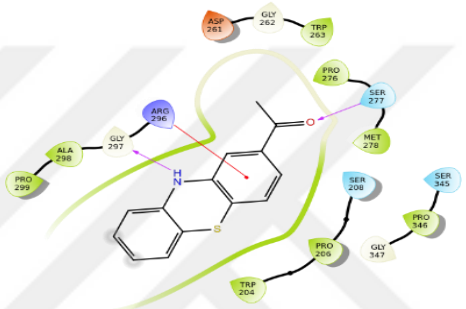
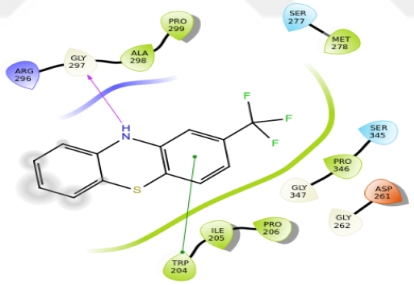
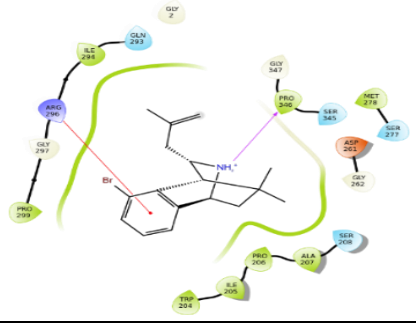
Molecule Name	Binding Affinity Experimental Values (pIC_{50})	Internal Predictions (pIC_{50})	2D Structure
CHEBL3347543	4.32	4.26	
CHEMBL3347503	4.17	4.36	
CHEMBL3347549	4.17	4.25	

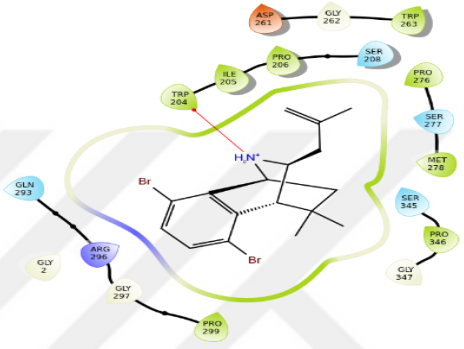
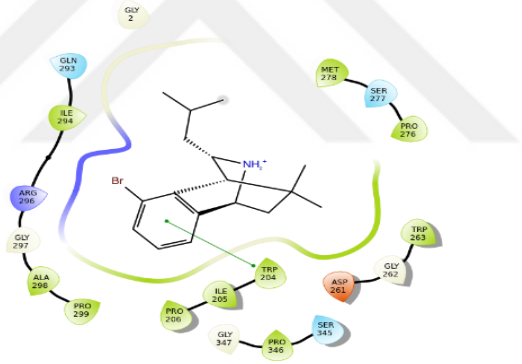
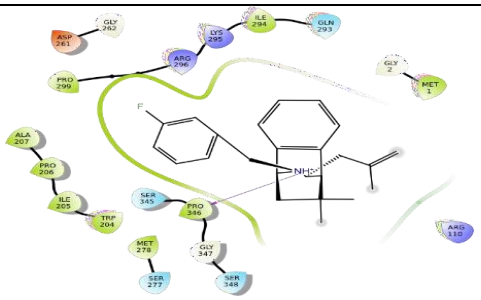
Molecule Name	Binding Affinity Experimental Values (pIC_{50})	Internal Predictions (pIC_{50})	2D Structure
CHEMBL3347567	4.10	4.01	

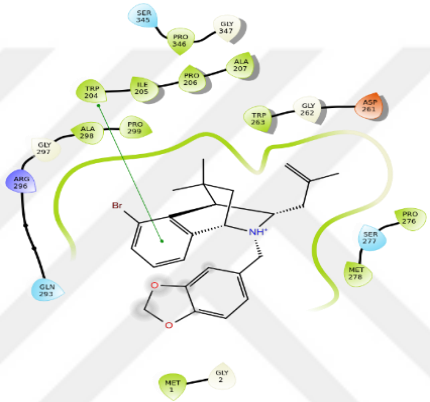
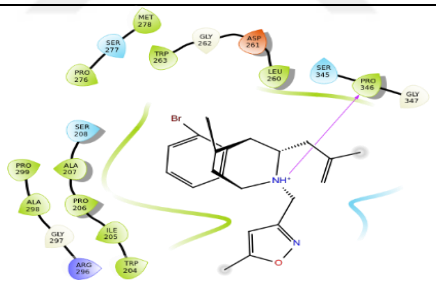
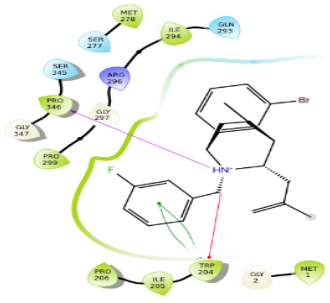
Molecules are retrieved from ChEMBL database (Cifuentes-Pagano et al. 2013).

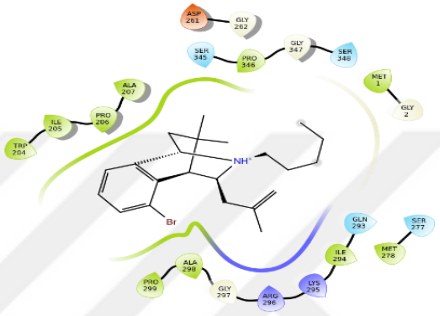
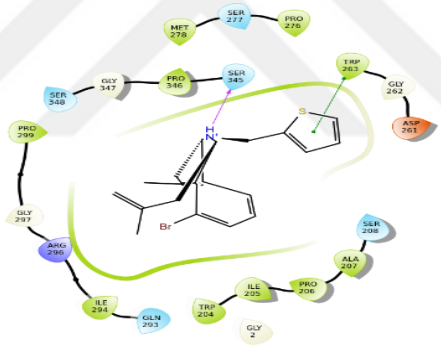
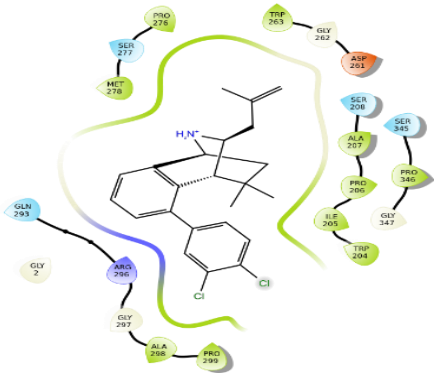
After predicting the pIC_{50} values of the internal test set, we have applied molecular docking with SP and IFD docking for the constructed NOX2 structure and the internal test set (Table 4.3).

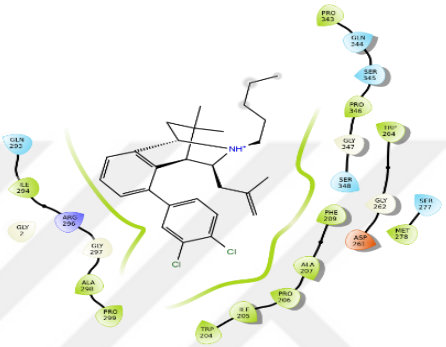
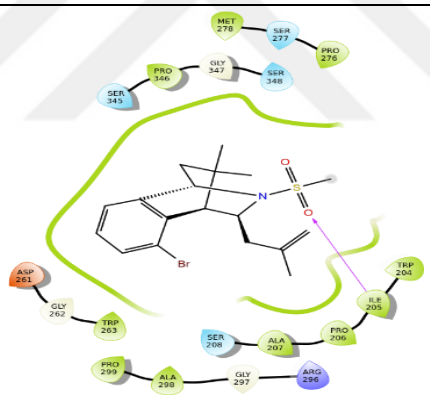
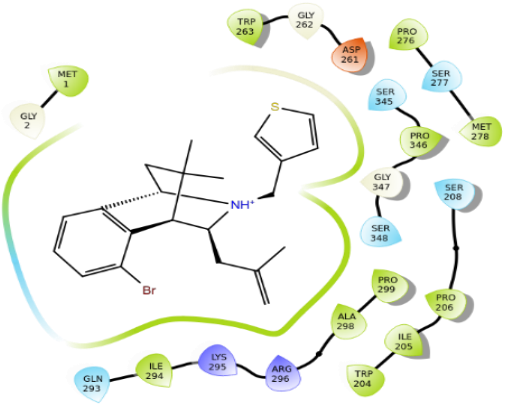
Table 4.3: Molecular docking results of the internal test set

Molecule name	SP Docking Scores (kcal/ mol)	2D Interactions (IFD)	IFD Scores (kcal/ mol)
CHEMBL407734	-4.14		-6.81
CHEMBL541018	-4.50		-7.85
CHEMBL3347491	-3.94		-6.02

Molecule name	SP Docking Scores (kcal/mol)	2D Interactions (IFD)	IFD Scores (kcal/ mol)
CHEMBL3347492	-3.91		-6.08
CHEMBL3347502	-4.27		-6.24
CHEMBL3347503	-3.30		-4.94

Molecule name	SP Docking Scores (kcal/mol)	2D Interactions (IFD)	IFD Scores (kcal/ mol)
CHEMBL3347542	-3.16		-5.37
CHEMBL3347543	-4.23		-5.57
CHEMBL3347549	-2.70		-6.74

Molecule name	SP Docking Scores (kcal/mol)	2D Interactions (IFD)	IFD Scores (kcal/ mol)
CHEMBL3347550	-4.41		-6.19
CHEMBL3347551	-3.05		-5.67
CHEMBL3347554	-2.98		-6.41

Molecule name	SP Docking Scores (kcal/mol)	2D Interactions (IFD)	IFD Scores (kcal/ mol)
CHEMBL3347567	-3.94		-7.58
CHEMBL3347583	-2.97		-6.77
CHEMBL3733336	-3.45		-6.30

15 NOX2 inhibitors that are used for AutoQSAR model were analyzed within the reported and predicted binding pocket through hydrogen bonding and other potential non-covalent interactions. CHEMBL541018 was docked in the binding pocket with the docking score of -7.85 kcal/mol. Moreover, formed Π-Π stacking interactions with TRP204 and H-bonding interactions with GLY297. Furthermore, other non-covalent interactions were observed with hydrophobic polar and non-polar residues. CHEMBL3347567 has the docking score of -7.58 kcal/mol. It formed interactions with TRP204, PRO206 and PHE209.

Among these 15 NOX2 inhibitors, the one with the highest pIC_{50} value is CHEMBL3733336. Thus, we applied MD simulation for this molecule for 100 ns (Table 4.4)

Table 4.4: Average MM/GBSA result of the NOX2 inhibitor CHEMBL3733336

Molecule Name	MM/GBSA (kcal/mol)
CHEMBL3733336	-32.43

At last, we applied screening for the natural Vitamin E isoforms, α-tocopherol phosphate and α -T-13′-COOH to predict their pIC_{50} values by using the AutoQSAR model we developed (Table 4.5).

Table 4.5: External predictions of AutoQSAR model for ten molecules

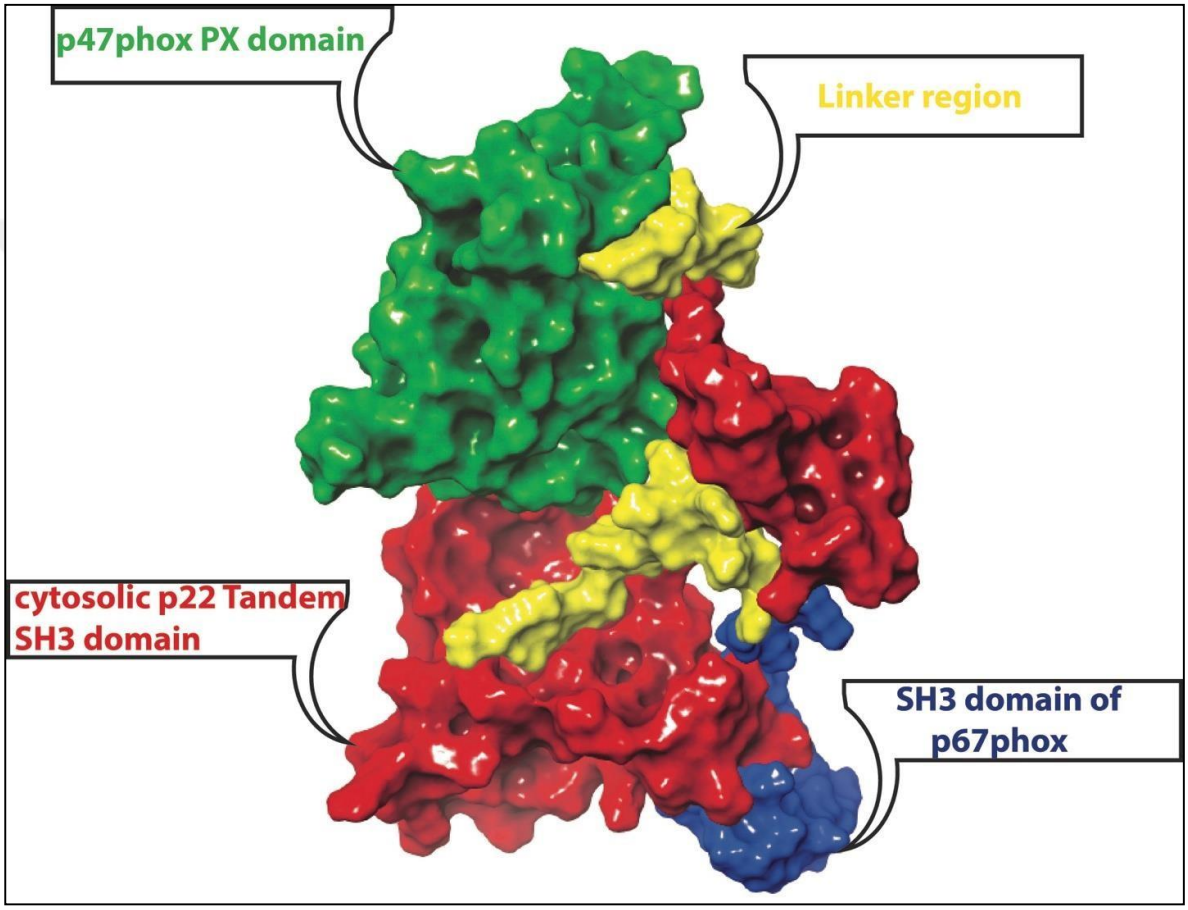
Molecule name	**External Predictions (pIC_{50})**
α –tocopherol	5.04
α –tocotrienol	5.18

Molecule name	**External Predictions (pIC_{50})**
β-tocopherol	4.95
β-tocotrienol	5.09
δ-tocopherol	4.93
δ-tocotrienol	5.08
γ-tocopherol	4.95
γ-tocotrienol	5.09
α-tocopherol phosphate	5.04
α -T-13′-COOH	5.07

4.2 NOX2 Model Construction

In this thesis, three domains which are critical for NOX2 activation have been assembled with an online server called AIDA (ab initio domain assembly) was utilized. This server is an online tool that is performed for identification of multi-domain proteins through their domains. Thus it can be used to predict their 3D structure and spatial arrangements (D. Xu et al. 2015). We linked three domains of NOX2 (PDB IDs, 1KQ6, 1UEC, 1K4U) with two linkers. $p47_{phox}$ PX domain (residue numbers, 1-141) is vital for anchoring of $p47_{phox}$ in the membrane and taken from crystal structure PDB ID, 1KQ6. The cytosolic p22 domain (residue numbers, 151-340) in autoinhibited form of the tandem SH3 domain of $p47_{phox.}$ $p47_{phox}$ that act as an organizer and involve in phosphorylation is the most critical part of NOX2 (PDB; 1UEC) (Yuzawa et al. 2004). The third portion comprises the "359-390" sequence and consists of the solution structure of the $p67_{phox}$ C-terminal SH3 domain complexed with the $p47_{phox}$ C-terminal tail region (Kami et al. 2002). $P67_{phox}$ is the system activator which engages with the C-terminal proline-rich region of $p47_{phox}$ which acts as an adaptor protein via its C-terminal SH3 domain. We linked three domains with two linker sequences [linkers between 1K4U and 1UEC is PGPQSPGSPLEEE, between 1UEC and 1KQ6 PKDGKSTAT]. The newly constructed model was subjected to 100 ns MD simulation to equilibrate the system and dispose of the possible steric clashes in the structure. Then, based on RMSD calculation, quadrant of the stable structure was saved and Ramachandran plot of the protein structure was plotted to determine the structural quality of the model as well as for further validation. The minimized model structure was used for molecular docking and MD simulations to analyze the protein structure behaviour and ligand effect under the physiological condition (Figure 4.2).

Figure 4.2: Structure model of NOX2, for the newly constructed structure, three domains are linked with two linker sequences [linkers in yellow colour are in between 1K4U and 1UEC; PGPQSPGSPLEEE and in between 1UEC and 1KQ6; PKDGKSTAT]



4.3 Docking Analysis of Vitamin E Molecules

The family of vitamin E is composed of eight isoforms; four tocopherols and four tocotrienols. Owing to its antioxidant and non-antioxidant properties, α-tocopherol is the most studied member of this family over other tocotrienols and tocopherols (Koufaki 2016).

α-Tocopherol can be phosphorylated to tocopherol phosphate (α-TP) in *in vitro* and *in vivo* models. α-T-13'-COOH, a metabolite of α-tocopherol is found in the serum of humans and mice at the nanomolar concentrations and selectively inhibits 5-LOX activity and hence suppresses the inflammation (Pein et al. 2018).

For this study eight vitamin E isoforms, α-T-13'-COOH and α-tocopherol phosphate were used for molecular docking. First of all, Blind Docking was applied to observe where the ligands bind (Figure 4.3), then ligands are applied to SP docking (Table 4.6).

Figure 4.3: Blind docking of eight vitamin E isoforms, α-T-13'-COOH and α-tocopherol phosphate

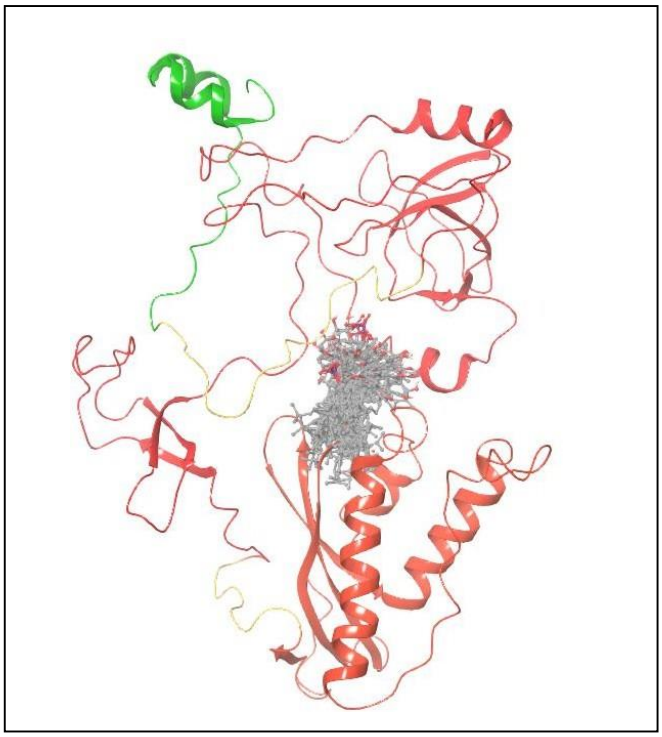
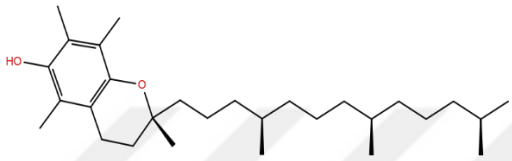
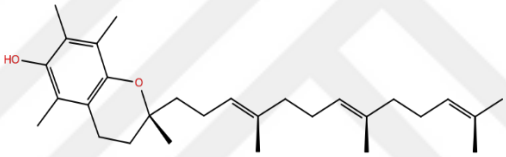
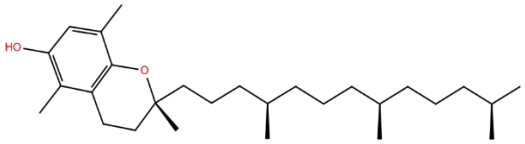
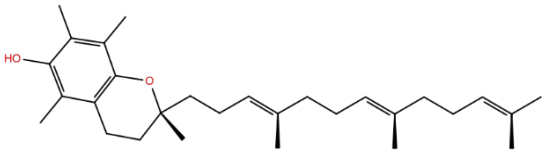
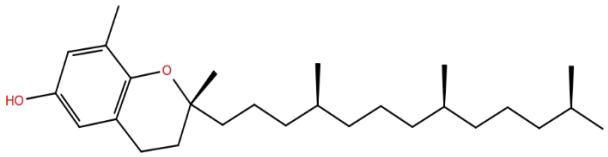
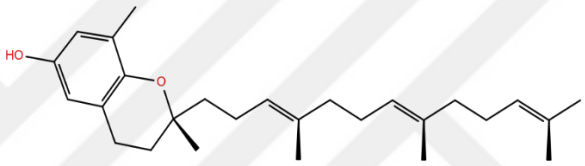
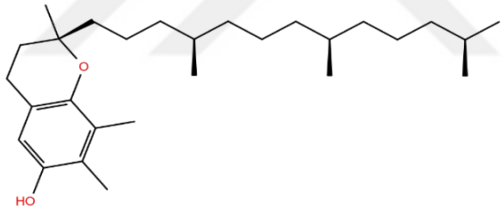
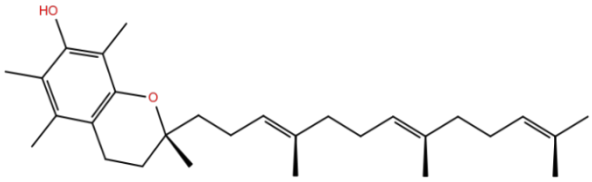
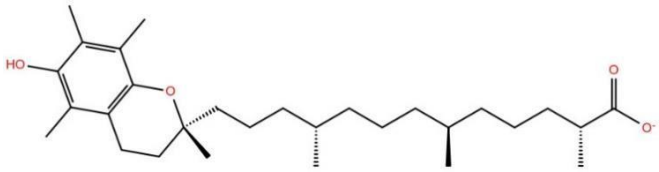
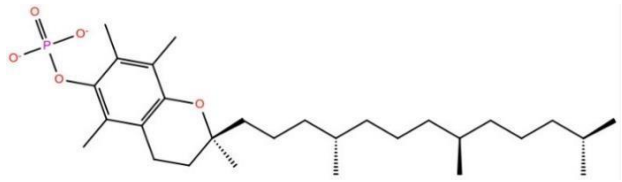


Table 4.6: SP Docking results of eight vitamin E isoforms, α-tocopherol phosphate and α-T-13'-COOH

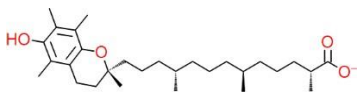
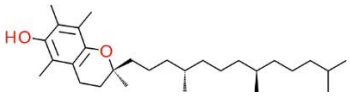
Molecule Name	Structure	SP Docking Scores (kcal/mol)
α-tocopherol		-4.24
α-tocotrienol		-4.09
β-tocopherol		-3.89
β-tocotrienol		-3.94

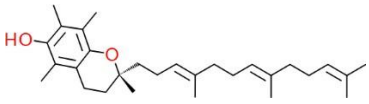
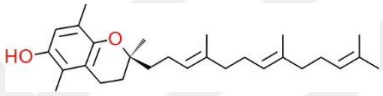
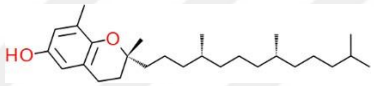
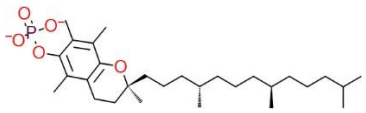
Molecule Name	Structure	SP Docking Scores (kcal/mol)
δ-tocopherol		-3.77
δ-tocotrienol		-4.17
γ-tocopherol		-4.30
γ-tocotrienol		-4.62
α-T-13'-COOH		-4.34

Molecule Name	Structure	SP Docking Scores (kcal/mol)
α-tocopherol phosphate		-4.39

Top six molecules were analyzed based on docking score, ligand efficiency and interactions with crucial residues and then were applied to IFD Docking (Table 4.7).

Table 4.7: **The docking scores of selected molecules with different protocols, ligand efficiency and 2D structures**

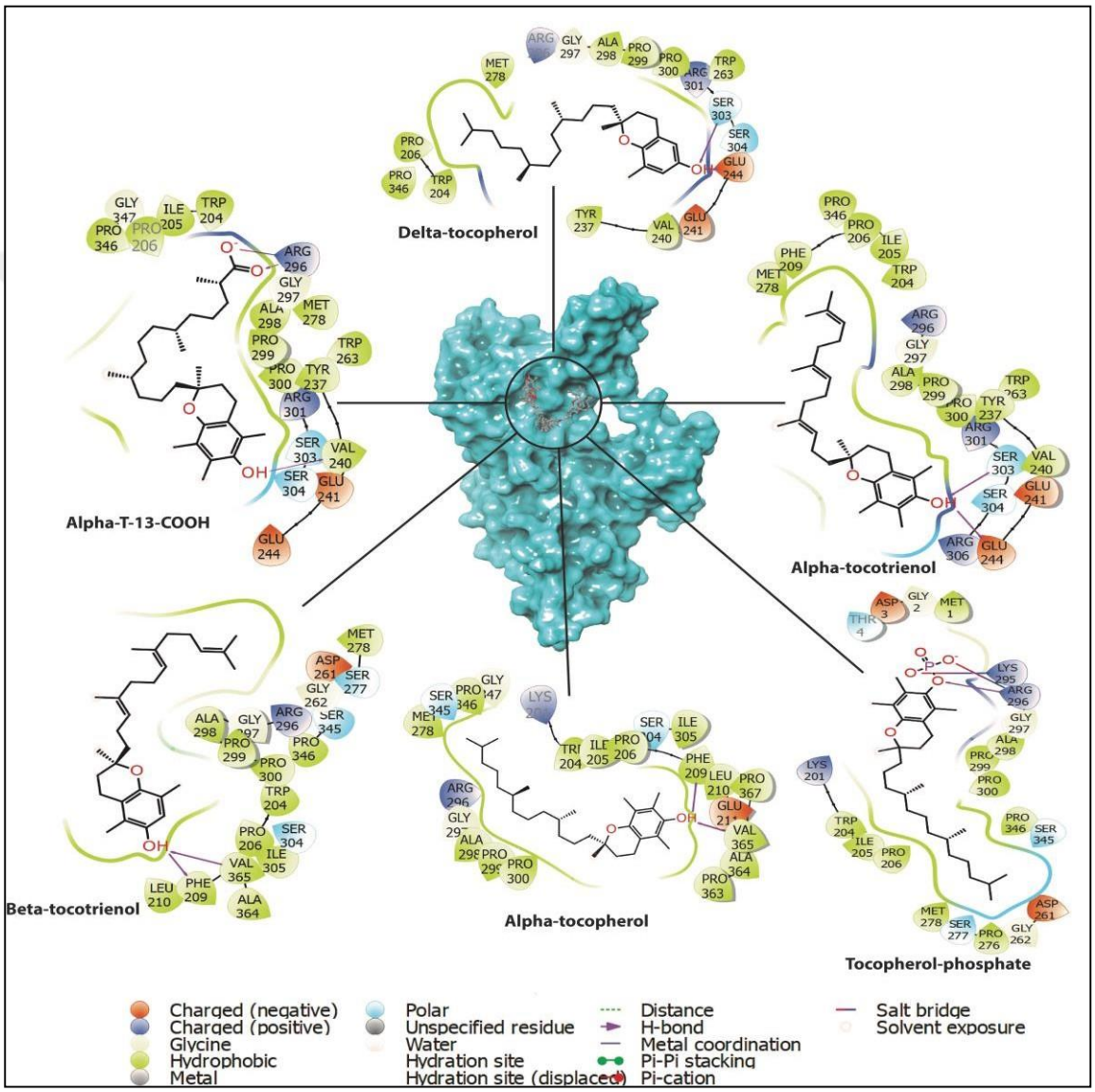
Molecules	2D Structures	SP Docking Scores (kcal/mol)	IFD Docking Scores (kcal/mol)	Blind Docking (kcal/mol)
α-T-13'-COOH		-4.34	-9.41	-4.93
α-tocopherol		-4.24	-7.48	-3.70

Molecules	2D Structures	SP Docking Scores (kcal/mol)	IFD Docking Scores (kcal/mol)	Blind Docking (kcal/mol)
α-tocotrienol		-4.09	-6.19	-4.37
β-tocotrienol		-3.94	-8.51	-3.90
δ-tocopherol		-3.77	-5.85	-4.08
α-tocopherol phosphate		-4.39	-9.83	-4.58

Top six molecules were analyzed within the reported and predicted binding pocket through hydrogen bonding and other potential non-covalent interactions. α-T-13-COOH was docked in the binding pocket with the docking score of -9.41 kcal/mol and formed hydrogen bond and salt bridge with ARG296 (Figure 4.4). Moreover, other non-covalent interactions were observed with hydrophobic polar and non-polar residues. α-tocopherol made significant interaction with PHE209 and VAL365 with the docking score of -7.48 kcal/mol (Figure 4.4). We found a similar type of interaction in β-tocotrienol with PHE209 and VAL365 with increased docking score -8.51 kcal/mol (Figure 4.4). Two hydrogen bonds with GLU244 and

SER303 and other non-covalent bond interactions were observed for α-tocotrienol with a docking score of -6.19 kcal/mol. δ-tocopherol made two hydrogen bonds with GLU244 and SER303 with a docking score of -5.85 kcal/mol (Figure 4.4). The phosphate group of α-tocopherol phosphate made two salt bridges and one hydrogen bond with ALA298 and ARG296 with the docking score of -9.86 kcal/mol (Figure 4.4). Vitamin E molecules with the desired interaction profile and docking scores were selected to perform further MD simulations.

Figure 4.4: Docking analyzes of top six molecules within the reported and predicted binding pocket through noncovalent interactions



4.4 Molecular Dynamics (MD) Simulations Analysis

MD simulations help us to investigate the time-dependent changes in the various complexes such as α-T-13'-COOH, α-tocopherol, α-tocotrienol, β-tocotrienol, δ-tocopherol and α-tocopherol phosphate.

In each MD simulation, RMSD was calculated to determine the stability of all the ligands in the domains binding pocket. One of the first parts of these research practices is to use the RMSD values to annotate structural and dynamic attributes that enable us to evaluate the structural stability of binding cleft residues and the ligands used during the MD simulations. RMSD for carbon alpha (Cα) atoms of each residue is used from their original positions is used to define the versatility of possible protein conformations along with MD simulations.

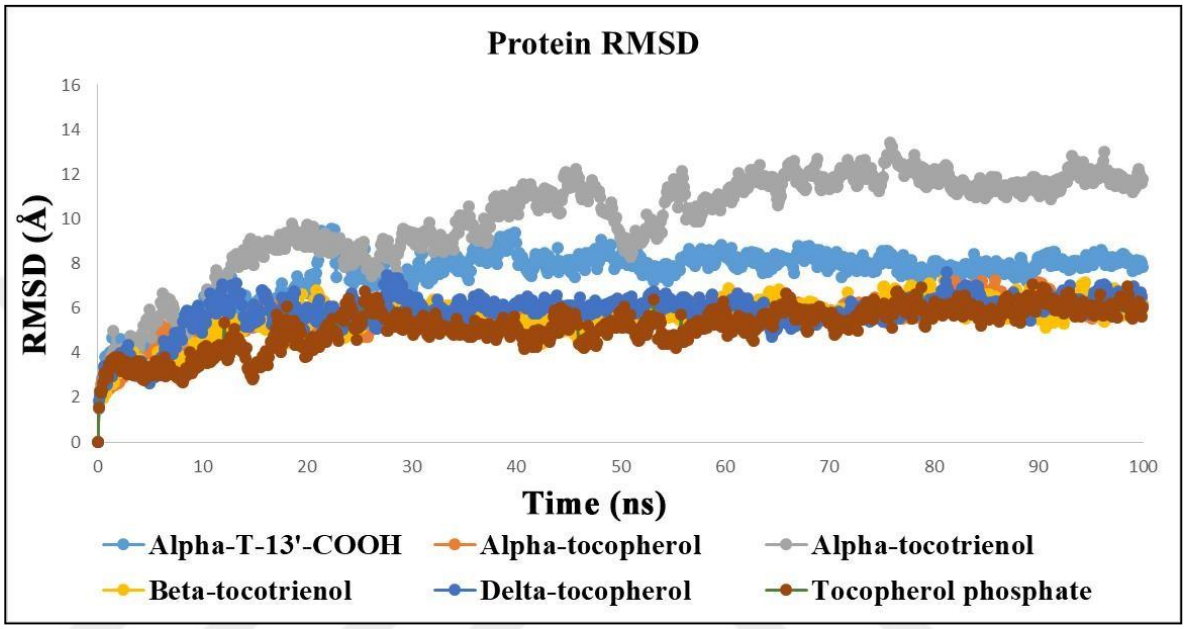
The Lig-fit-prot and Lig-fit-Lig represent the ligand motion in the binding pocket. Mean values of protein RMSD, Lig-fit prot, Lig-fit Lig and protein RMSF are shown in Table 4.8.

Table 4.8: Average RMSDs of top six vitamin E isoforms, α-tocopherol phosphate and α-T-13'-COOH during 100 ns simulation time

Molecules	**Protein RMSD (Å)**	**Lig-fit-prot RMSD (Å)**	**Lig-fit-Lig RMSD (Å)**	**Protein RMSF (Å)**
α-T-13'-COOH	7.42	6.90	2.54	3.23
α-tocopherol	5.60	5.56	1.90	2.73
α-tocotrienol	9.94	4.98	1.65	3.42
β-tocotrienol	5.53	4.17	1.44	2.57
δ-tocopherol	5.82	3.33	1.35	3.03
α-tocopherol phosphate	5.13	3.31	1.03	2.23

From the RMSD graph, we can see that α-tocopherol phosphate, α-T-13'-COOH and β-tocotrienol shows stable RMSD. α-tocotrienol slightly destabilized the protein complex. There were slight fluctuations observed in δ-tocopherol (Figure 4.5).

Figure 4.5: Protein RMSD graph of top six molecules of vitamin E isoforms, α- T-13'-COOH and α-tocopherol phosphate, protein backbone RMSD graph of selected molecules during the 100 ns simulation time



Average of Lig-fit Lig ligand RMSDs were shown in Table 4.8. α-T-13'-COOH that may show the higher torsional motion in the binding pocket. α-tocopherol phosphate shows minimum motion that shows its stability. Nevertheless, the other two molecules, α-tocopherol and β-tocotrienol, shows rotational motion in starting 38 ns after that they were stable. δ-tocopherol also shows less rotational motion in the binding pocket (Figure 4.6). Selected molecules are shown in LigFitProt RMSD plots in Figure 4.7.

Figure 4.6: LigFitLig RMSD plots of selected molecules, the mean values of LigFitProt RMSDs for α -T-13'-COOH, α -tocopherol, α-tocotrienol, β-tocotrienol, δ-tocopherol and α-tocopherol phosphate were found to be respectively 6.90, 5.56, 4.98, 4.17, 3.33 and 3.31 Å as shown in Table 4.8

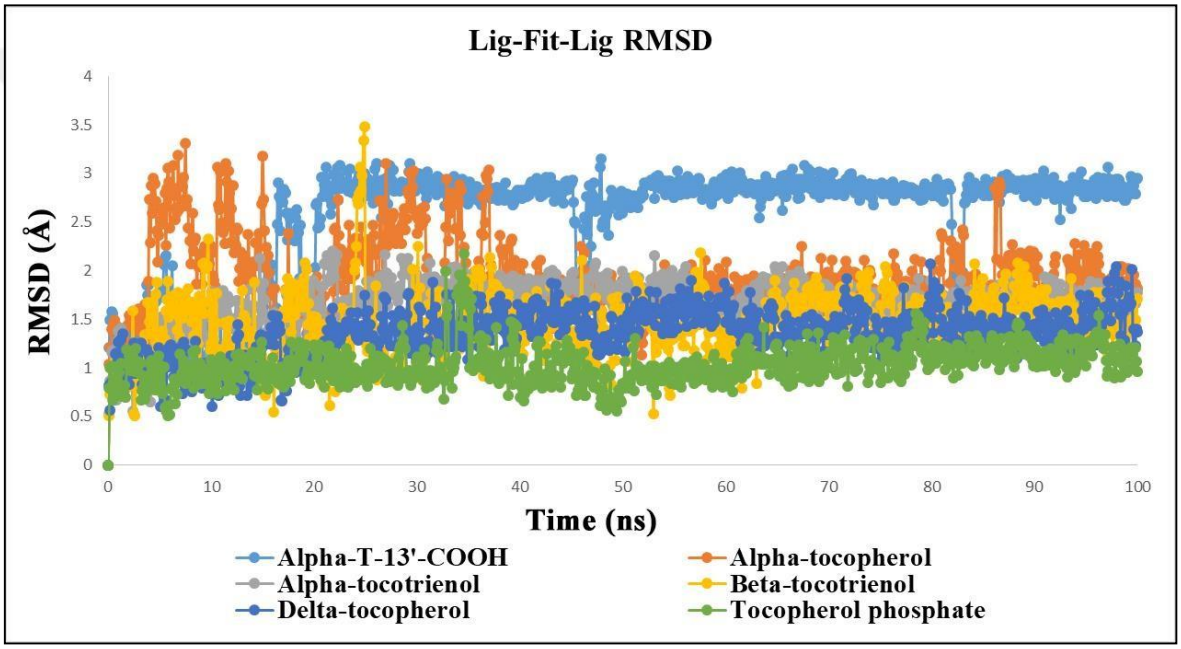
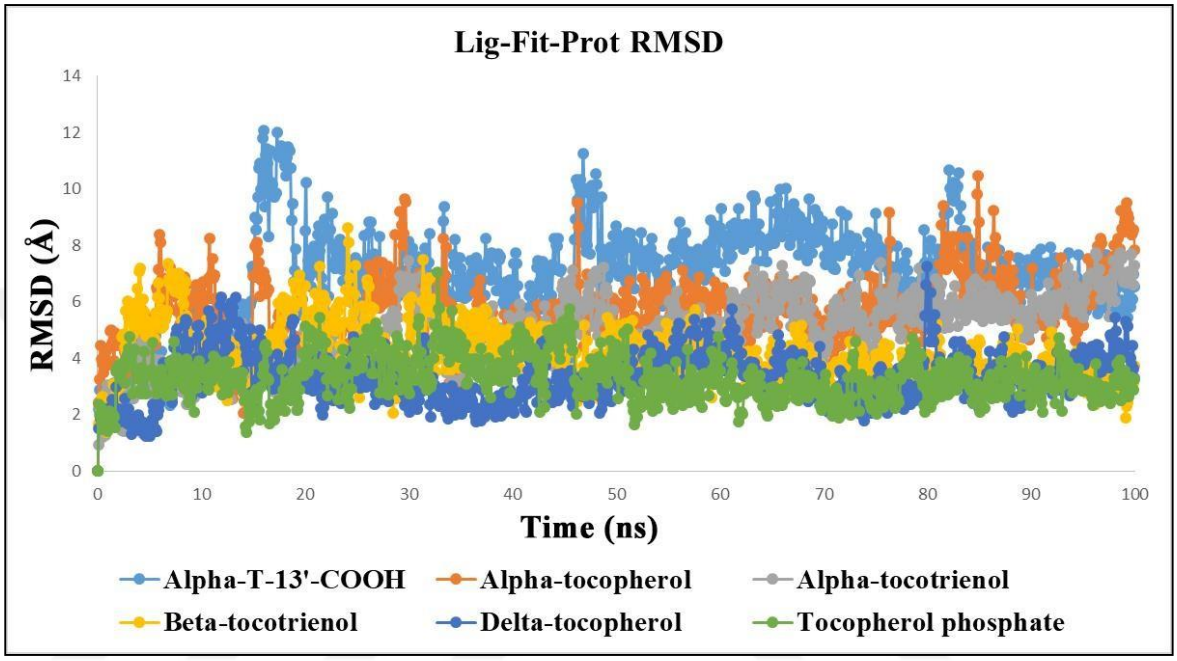


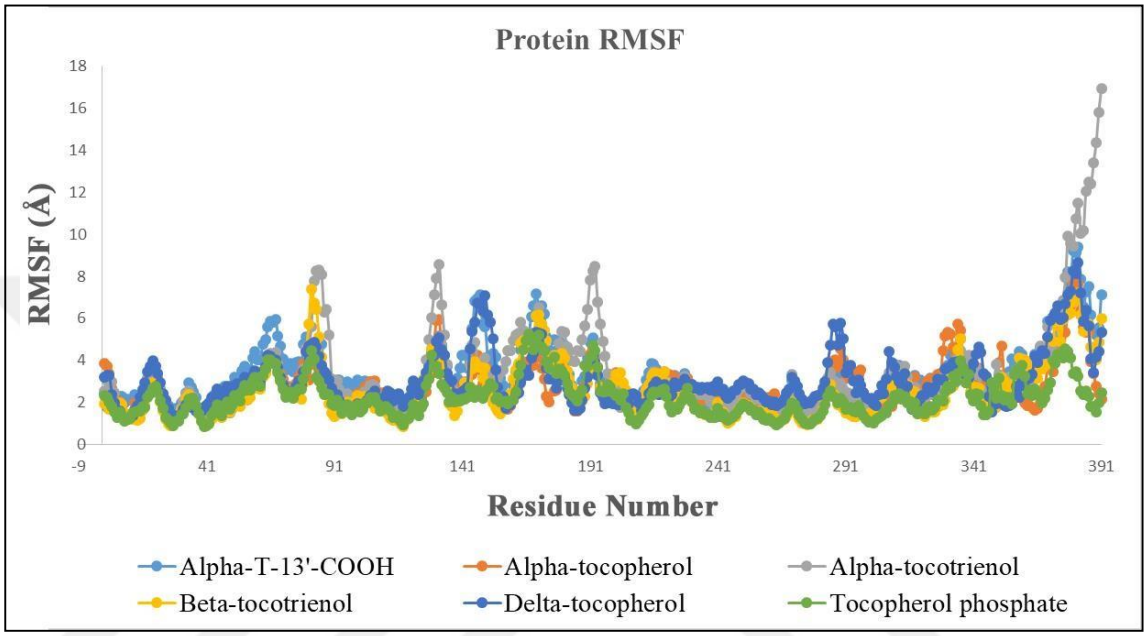
Figure 4.7: LigFitProt RMSD plots of selected molecules



δ-tocopherol and α-tocopherol phosphate have stable conformation in the binding pocket of the protein. Whereas other selected compounds have high fluctuations in RMSD values, indicating its lower stability.

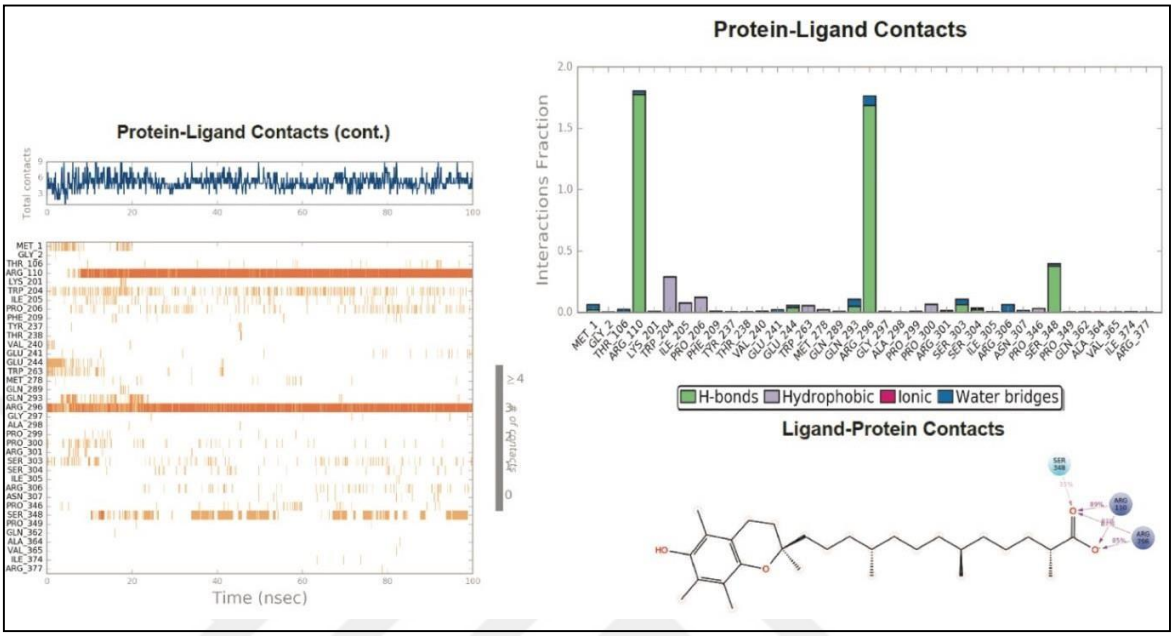
RMSF (Root Mean Square Fluctuation) indicated that overall fluctuation pattern for the NOX2 and its complexes. α-tocotrienol showed higher fluctuations at the binding pocket and significant higher fluctuation compared to other molecules with the average value of 3.42 Å. Residues in the binding region show smaller fluctuation. α-Tocopherol phosphate and δ-tocopherol show a similar pattern of fluctuations (Figure 4.8).

Figure 4.8: Overall fluctuation pattern for the NOX2 and its complexes



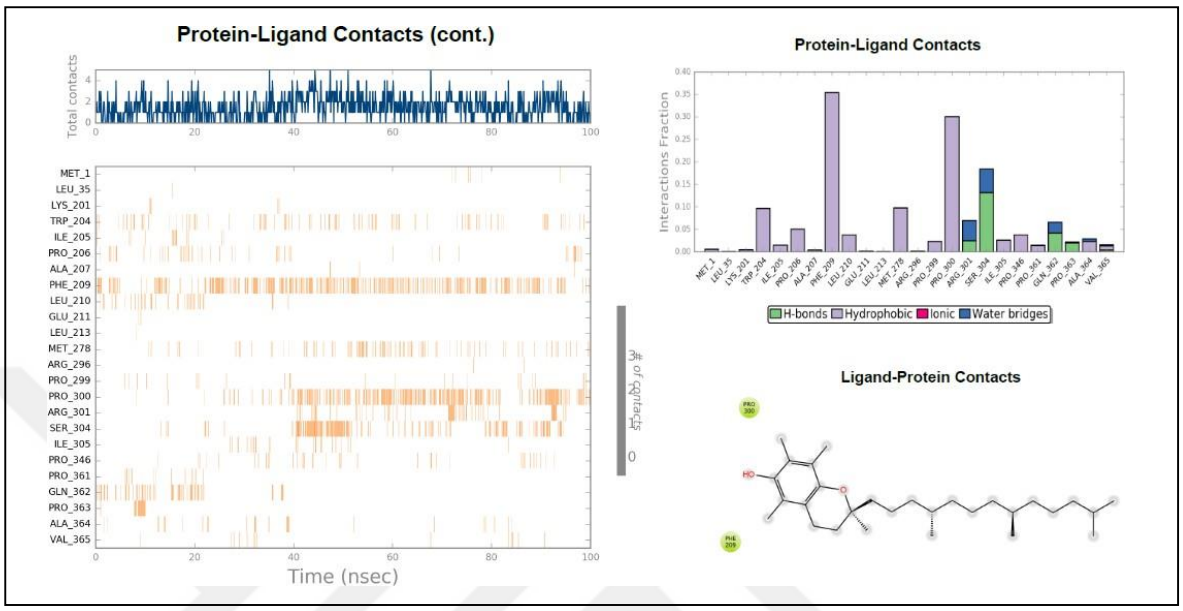
α-T-13'-COOH showed significant interactions with SER348, ARG110 and ARG296. These interactions were observed for 35%, 89.87% and 85% through hydrogen bonds as well as through water bridges (Figure 4.9).

Figure 4.9: Contact of NOX2 and α-T-13-COOH during 100 ns time period



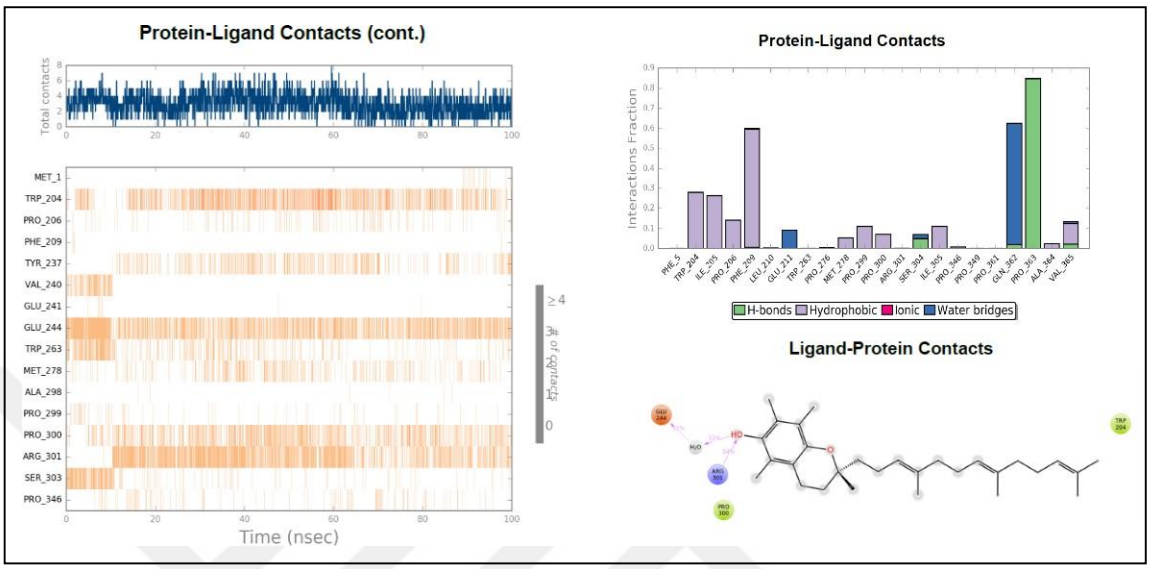
α-Tocopherol showed interactions with PRO300 and PHE209 through hydrophobic interaction. Other crucial residues made hydrogen bond and water bridges with the ligand, but that was not maintained during the simulation time (Figure 4.10).

Figure 4.10: Contact of NOX2 and α-tocopherol during 100 ns time period



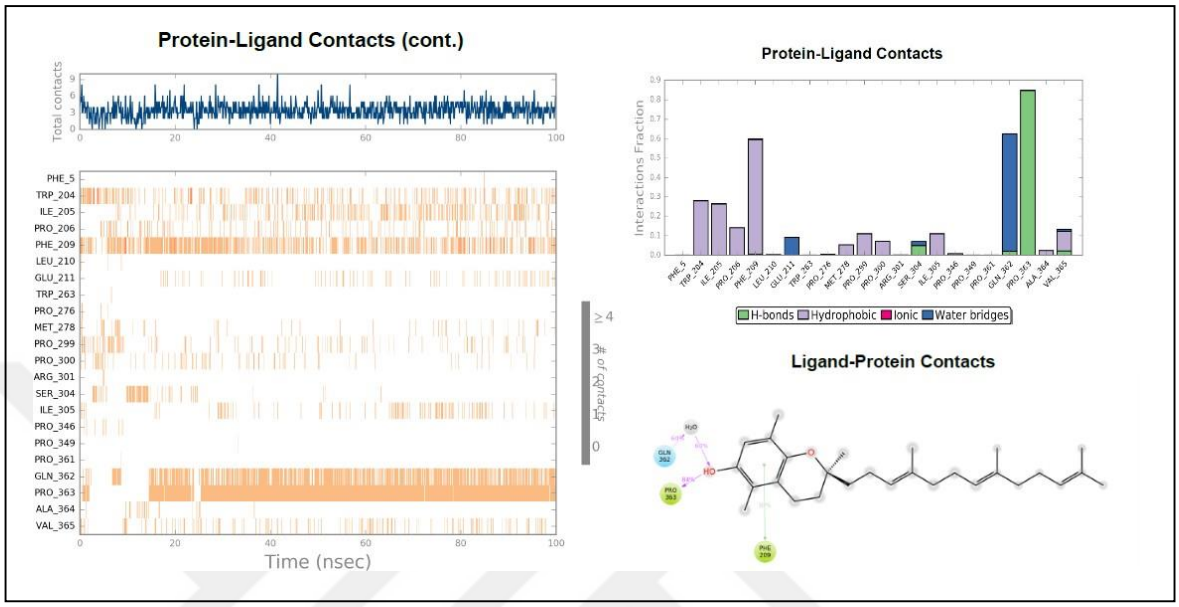
In 2D simulation interaction map an α-tocotrienol showed a substantial interaction with GLU244, ARG301, PRO300 and TPR204 through hydrogen bonds, hydrophobic interactions and water bridges that were present for 51% and 54 (Figure 4.11).

Figure 4.11: Contact of NOX2 and α-tocotrienol during 100 ns time period



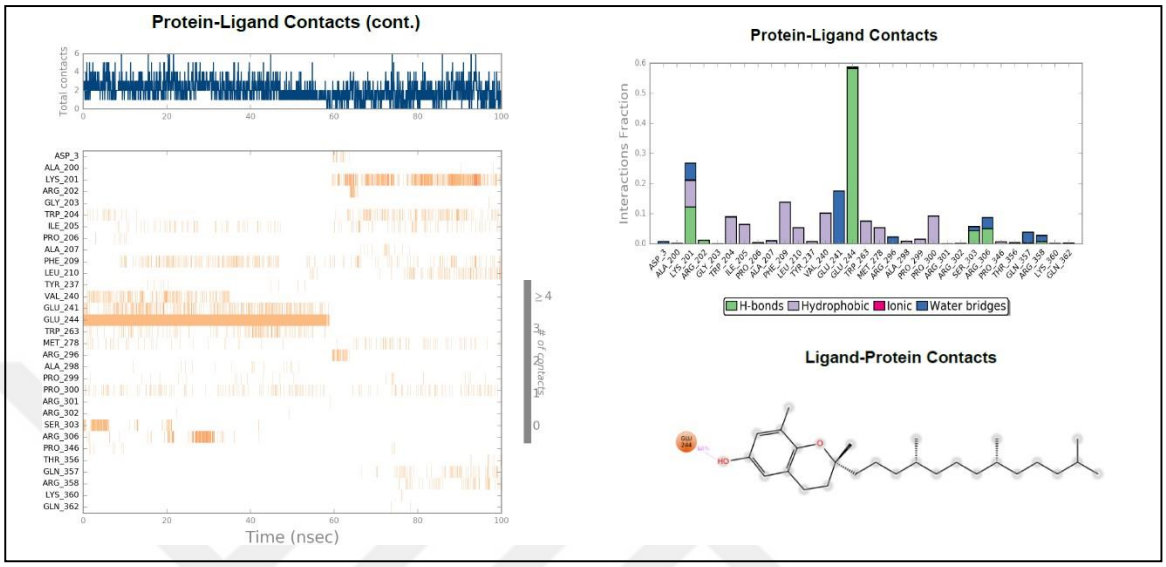
β-tocotrienol maintained substantial contacts with GLU362, PRO363 and PHE209 via hydrogen bonding in the 100 ns MD simulations that maintained 60%, 84%, and 30% (Figure 4.12).

Figure 4.12: Contact of NOX2 and β-tocotrienol during 100 ns time period



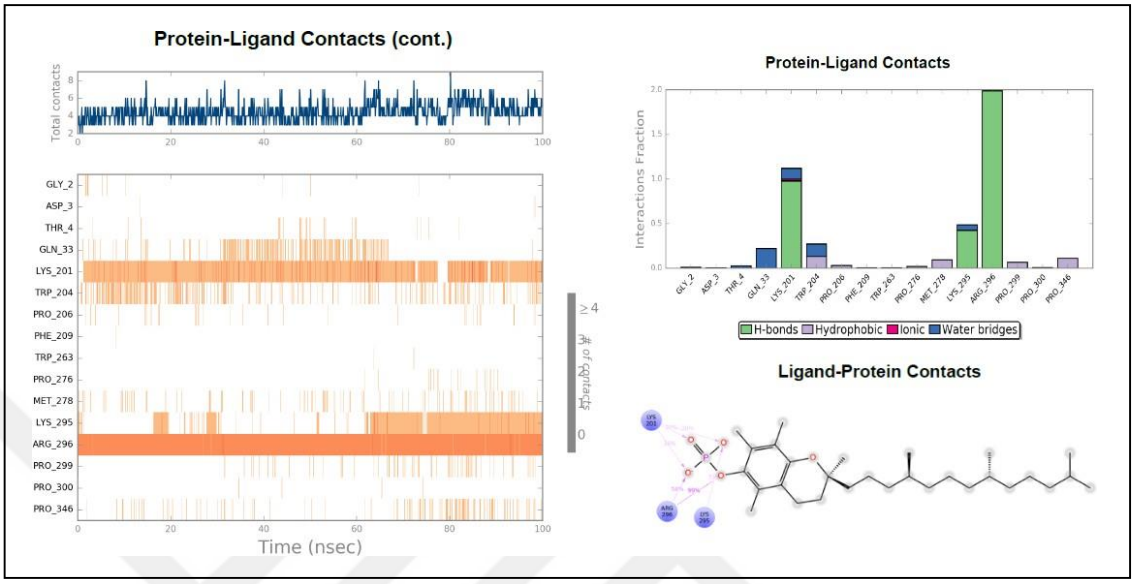
δ-tocopherol maintains hydrogen bond only with GLU244 that was present 58% through simulation time (Figure 4.13).

Figure 4.13: Contact of NOX2 and δ-tocopherol during 100 ns time period



The phosphate group of α-tocopherol phosphate made 99% interaction with ARG296 via hydrogen bond. LYS201 and LYS295 made interactions with ligand through hydrogen bonds and water bridge. It was present for 30% and 35% through simulation time (Figure 4.14).

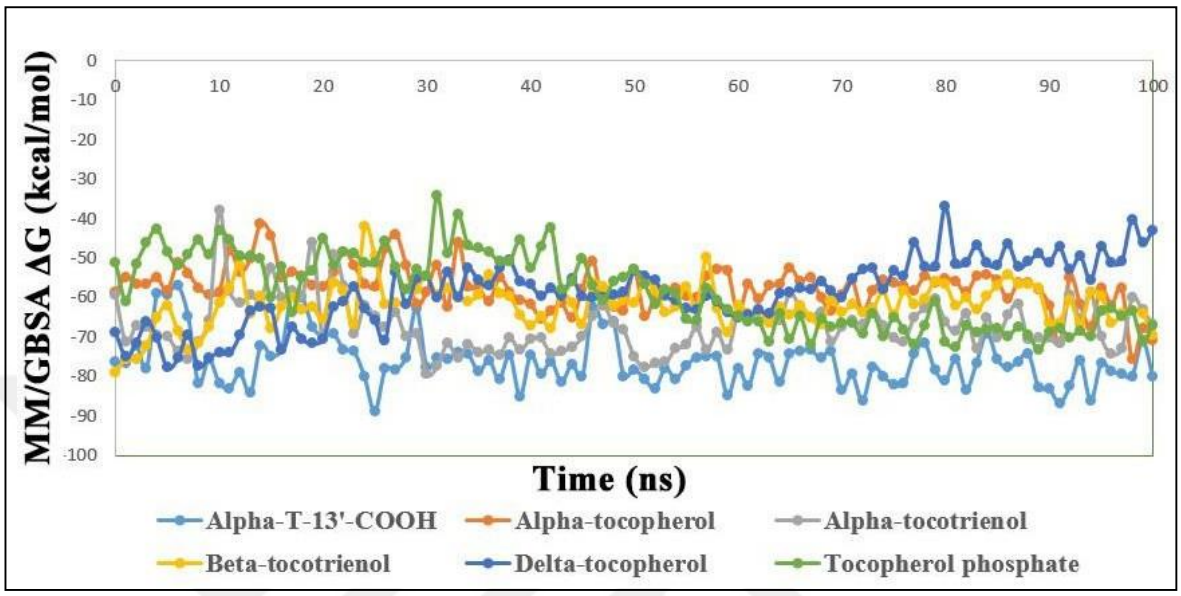
Figure 4.14: Contact of NOX2 and α-tocopherol phosphate



The selected molecules remained stable throughout the simulations and showed interactions with NOX2. Most of the interacting residues observed in docking and simulations were the catalytic residues.

MM/GBSA was performed for analyzing the binding free energy of the selected ligands. 100 frames were taken from the whole trajectory for each of the complexes. Figure 4.15 shows the plot of binding free energy versus the simulation time.

Figure 4.15: Plots of binding free energy values versus the simulation time



α -T-13'-COOH has the lowest average ΔG of -76.17 kcal/mol over the simulation time signifying a more stable ligand conformation and higher binding affinity. Similarly, α-tocotrienol also showed a strong binding affinity with ΔG of -66.95 kcal/mol, β-tocotrienol with -61.94 kcal/mol, δ-tocopherol with -58.72 kcal/mol. α-Tocopherol phosphate and α-tocopherol show the highest binding energy values, ΔG of -58.41, -57.39 kcal/mol, respectively (Table 4.9).

Table 4.9: Average MM/GBSA results of top six molecules

Molecules	MM/GBSA (kcal/mol)	Standard Deviation
α-T-13'-COOH	-76.17	6.38
α-tocopherol	-57.39	5.50
α-tocotrienol	-66.95	5.86
β-tocotrienol	-61.94	5.48
δ-tocopherol	-58.72	10.02
α-tocopherol phosphate	-58.41	9.35

4.5 Mutations analysis

β-tocotrienol is selected for mutational analysis. At the end of the simulation, β-tocotrienol maintains significant contacts with crucial residues. PHE209, GLN362 and PRO363 were replaced with ALA to analyze the drug binding affinity. Table 4.10 shows that ALA209 increase predicted binding affinity.

Table 4.10: Mutation analysis results of molecular docking

Mutations	Glide/IFD (kcal/mol)	MM/GBSA (kcal/mol)
β-tocotrienol (Wild Type)	-8.51	-61.94
Phe209Ala	-7.28	-73.29
Gln362Ala	-8.93	-63.86
Pro363Ala	-8.56	-58.94

4.6 Comparison of IFD Docking Scores, MM/GBSA Results and AutoQSAR Predictions

The predicted binding affinity values of the vitamin E molecules are predicted by AutoQSAR. Predictions are given in pIC_{50} in Table 4.11.

The docking scores and MM/GBSA results were in ΔG level, to compare the docking scores and MM/GBSA results with AutoQSAR predictions, pIC_{50} values are converted to IC_{50} values since $\Delta G = R.T \ln (IC_{50})$.

$$pIC_{50} = \log_{10}(IC_{50}) \quad \textbf{(4.2)}$$

$$\Delta G = -RT \ln K \quad \textbf{(4.3)}$$

Here it is used IC_{50} instead K.

$$\Delta G = -R.T \ln (IC_{50}) \quad \textbf{(4.4)}$$

The converted pIC_{50} values and ΔG values are shown in Table 4.11.

Table 4.11: AutoQSAR predictions converted into ΔG

Molecule Name	pIC_{50}	IC_{50} (μM)	ΔG (kcal/mol)
α-T-13'-COOH	5.07	8.51	-6.91
α-tocopherol	4.95	9.12	-6.87
α-tocotrienol	5.18	6.61	-7.06
β-tocotrienol	5.09	8.13	-6.94
δ-tocopherol	4.93	11.75	-6.72
α-tocopherol phosphate	5.04	9.12	-6.87

After the conversion is completed to compare the predictions that are calculated with different strategies ranking is applied. The values for comparison are given in Table 4.12.

Table 4.12: IFD Docking scores and MM/GBSA results of top six molecules and AutoQSAR predictions converted into ΔG

Molecule Name	**SP Docking Scores (kcal/mol)**	**IFD Docking Scores (kcal/mol)**	**MM/GBSA Results (kcal/mol)**	**AutoQSAR Predictions Converted into ΔG (kcal/mol)**
α -T-13'-COOH	-4.34	-9.41	-76.17	-6.91
α -tocopherol	-4.24	-7.48	-57.39	-6.87
α -tocotrienol	-4.09	-6.19	-66.95	-7.06
β-tocotrienol	-3.94	-8.51	-61.94	-6.94
δ-tocopherol	-3.77	-5.85	-58.72	-6.72
α-tocopherol phosphate	-4.39	-9.83	-58.41	-6.87

5. DISCUSSION

We started with ligand-based drug design strategy to construct the 3D QSAR model for predictions of binding activity of natural vitamin E isoforms and derivatives. In this thesis, it was preferred to use AutoQSAR from Schrödinger's Maestro suite (Dixon et al. 2016). To predict the pIC_{50} values of vitamin E molecules, an AutoQSAR model was constructed, we screened for the known NOX2 inhibitors from the ChEMBL database for the dataset. However, the data was limited to 15 molecules. Nevertheless, the score of the model was 0.87. For these 15 NOX2 inhibitors, the model predicts an internal test set with an average 0.242 error rate, also with no error for more than one logarithmic unit. This was promising. Then we changed our strategy to the structure-based drug design and performed docking for these 15 NOX2 inhibitors with our NOX2 model. These NOX2 inhibitors have different structures from the vitamin E molecules. All 15 NOX2 inhibitors show interaction with TRP204, ARG296, PRO206, and MET278. On the other hand, some of the molecules show interaction with some novel residues. CHEMBL3347543 and CHEMBL33474549 show also interaction with PRO346; CHEMBL3347551 shows interaction with SER345 and TRP263; CHEMBL3347583 interact with ILE205; CHEMBL407734 has interactions with GLY297, ARG296, and SER277. CHEMBL541018 has interactions with GLY297. Among these residues, only ARG296 is mentioned in the literature as being a crucial residue in the binding of NOX2 (Groemping et al. 2003). However, since these 15 NOX2 inhibitors are experimentally known molecules, these residues may have an essential role in the binding of NOX2. Because the literature study was done with the PDB of 1OV3 that comprises of structure of the $p22_{phox}$- $p47_{phox}$ complex (Groemping et al. 2003). However, we construct our model in a way that comprises the structures of $p22_{phox}$, $p47_{phox}$ and $p67_{phox}$.

We applied molecular dynamics simulations for the best-known molecule of 15 NOX2 inhibitors that are acquired from the ChEMBL database. We were able to compare the results with vitamin E molecules. The MM/GBSA result of the best NOX2 inhibitor

(CHEMBL3733336) was -32.43 kcal/mol, while vitamin E molecules MM/GBSA results range from -57.39 to -76.17 kcal/mol.

NOX2 structure is constructed according to the previous studies, with the binding pocket of TRP193, TRP204, PRO206, PHE209, LEU210, GLU241, ASP243, TRP263, MET278, TYR279, ARG296, SER303. Among ten natural vitamin E molecules, the top six were analyzed within the reported and predicted binding pocket through hydrogen bonding and other potential non-covalent interactions. The interaction with TRP204, PRO206, PRO299, PRO300, and ARG296 was frequent for the top six molecules. Among these interactions, TRP204, PRO206, and ARG296 were mentioned in the literature (Groemping et al. 2003). Nonetheless, six vitamin E molecules also show interactions with PRO299 and PRO300 according to our findings; thus, these amino acids may also play a crucial role in binding.

Based on IFD Docking results of six molecules, best-resulted molecules were α-tocopherol phosphate, α -T-13-COOH, and β-tocotrienol with -9.83, -9.41 and -8.51 kcal/mol, docking scores, respectively. α -T-13-COOH formed a hydrogen bond and salt bridge with ARG296. Other non-covalent interactions were observed with hydrophobic, polar and non-polar residues. For MD simulations α -T-13'-COOH showed significant interactions with SER348, ARG110, and ARG296 that were observed through hydrogen bonds and water bridges. Ligand interactions with ARG296 is known for NOX2 (Groemping et al. 2003); however, SER348, ARG110 are first detected in this study, and it may be an essential interaction for NOX2. The phosphate group of α-tocopherol phosphate made two salt bridges and one hydrogen bond with ALA298 and ARG296. Here ALA298 interaction is detected; it may have importance for binding NOX2. β-tocotrienol has interactions with LEU210, ARG296, MET278, PHE209, PRO206, TRP204 that all these residues are mentioned as crucial in the literature (Groemping et al. 2003). However, interactions of β-tocotrienol with VAL365, ILE305, ALA364, SER304, PRO300, PRO299, ALA298, SER345, and ASP261 have been observed. These interactions have not been reported in the literature; however, they may play an essential role in the binding of NOX2.

Our model comprises of three domains of $p47_{phox}$ PX domain, cytosolic $p22_{phox}$ tandem SH3 domain and SH3 domain of $p67_{phox}$. However previous studies include one or two domains(Magnani et al. 2017; Groemping et al. 2003). For example, the study of Megnani et al. includes cytosolic $p22_{phox}$ tandem SH3 domain and the study of Groemping et al. includes both cytosolic $p22_{phox}$ tandem SH3 domain and $p47_{phox}$ PX domain. Differently we construct a model of three crucial domains that are needed for the activation of NOX2 in this study.

After model construction by AutoQSAR, docking, MD simulations and post- processing MD analyses were performed. In order to compare the predicted binding affinities of vitamin E molecules with ligand and target-driven based approaches, pIC_{50} values were converted to binding free energies (ΔG).

Among MM/GBSA predictions, α-T-13′-COOH gave the best ranking and the worst result was taken from δ-tocopherol.

With this study we constructed a NOX2 model that consists of all three crucial domains of $p47_{phox}$ PX domain, cytosolic $p22_{phox}$ tandem SH3 domain and SH3 domain of $p67_{phox}$. For the further studies it may help for the more accurate predictions for computational studies.

6. CONCLUSIONS

The crystal structure of the NADPH binding domain of $gp91_{phox}$ (3A1F) has first been published in PBD in 2010 (Beaumel et al. 2017). Although in recent years, many *in vitro* and *in vivo* studies have been carried out for identifying promising compounds against NOX2, in silico studies are limited in number. In this thesis, it is aimed to investigate direct interaction and binding mode of vitamin E molecules to NOX2. The already available 3A1F structure provides only a limited sequence and only consists of cytosolic p22phox tandem SH3 domain. On the other hand, the other crystal structure that the p22phox-p47phox complex (PDB; 1OV3) has both of $p47_{phox}$ PX domain and cytosolic $p22_{phox}$ tandem SH3 domain. Groemping et al. mentioned the crucial residues for the binding pocket in the study of the crystal structure of the p22phox-p47phox complex (PDB; 1OV3) (Groemping et al. 2003). This crystal structure consists of two domains of $p47_{phox}$ PX domain and cytosolic $p22_{phox}$ tandem SH3 domain. However, for the activated NOX2 it is needed all three domains of $p47_{phox}$ PX domain, cytosolic $p22_{phox}$ tandem SH3 domain and SH3 domain of $p67_{phox}$. Although we took advantage of this crystal structure (PDB; 1OV3) for specifying the binding pocket, it is lack of SH3 domain of $p67_{phox}$ and this absence may even effective on the possible binding sites of the vitamin E derivatives against NOX2. We have constructed a novel model that includes p22phox-p47phox and p67phox subunits by using the corresponding PBD data 1KQ6, 1UEC and 1K4U.

As in the literature, there is no complete 3D structure of NOX2 available; the model we constructed provides essential insights for further evaluation of NOX2 mechanism. After validation of the new model, docking algorithms were applied to observe the binding affinity of natural vitamin E molecules at NOX2. Among ten vitamin E molecules analyzed, six molecules were selected based on docking score, ligand efficiency, and interactions with crucial residues. However, since there is no experimental data of binding affinity of these natural vitamin E molecules in literature, we created a new 3D QSAR model by using machine learning algorithms that can predict the binding affinity of these natural vitamin E

molecules. This model is not only able to predict the pIC_{50} values of these natural vitamin E molecules but also predict novel candidates to inhibit NOX2. The results will further be analyzed in *in vitro* studies. This model can be useful for screening massive datasets of small molecules for NOX2 inhibitors.

REFERENCES:

Periodicals

Ajayi, A., Yu, X., & Ström, A. L. (2013). The role of NADPH oxidase (NOX) enzymes in neurodegenerative disease. *Frontiers in Biology*. https://doi.org/10.1007/s11515-012-1250-y

Altenhöfer, S., Radermacher, K. A., Kleikers, P. W. M., Wingler, K., & Schmidt, H. H. H. W. (2015). Evolution of NADPH oxidase inhibitors: Selectivity and mechanisms for target engagement. *Antioxidants and Redox Signaling*. https://doi.org/10.1089/ars.2013.5814

Azzi, A. (2019). Tocopherols, tocotrienols and tocomonoenols: Many similar molecules but only one vitamin E. *Redox Biology*. https://doi.org/10.1016/j.redox.2019.101259

Bas, D. C., Rogers, D. M., & Jensen, J. H. (2008). Very fast prediction and rationalization of pKa values for protein-ligand complexes. *Proteins: Structure, Function and Genetics*. https://doi.org/10.1002/prot.22102

Bedard, K., & Krause, K. H. (2007). The NOX family of ROS-generating NADPH oxidases: Physiology and pathophysiology. *Physiological Reviews*. https://doi.org/10.1152/physrev.00044.2005

Bowers, K. J., Chow, E., Xu, H., Dror, R. O., Eastwood, M. P., Gregersen, B. A., … Shaw, D. E. (2006). Scalable algorithms for molecular dynamics simulations on commodity clusters. *Proceedings of the 2006 ACM/IEEE Conference on Supercomputing, SC'06*. https://doi.org/10.1145/1188455.1188544

Brandes, R. P., Weissmann, N., & Schröder, K. (2014). Nox family NADPH oxidases: Molecular mechanisms of activation. *Free Radical Biology and Medicine*. https://doi.org/10.1016/j.freeradbiomed.2014.07.046

Brigelius-Flohé, R., & Traber, M. G. (1999). Vitamin E: function and metabolism. *The FASEB Journal*. https://doi.org/10.1096/fasebj.13.10.1145

Brown, D. I., & Griendling, K. K. (2009). Nox proteins in signal transduction. *Free Radical Biology and Medicine*. https://doi.org/10.1016/j.freeradbiomed.2009.07.023

Cho, A. E., Guallar, V., Berne, B. J., & Friesner, R. (2005). Importance of accurate charges in molecular docking: Quantum Mechanical/Molecular Mechanical (QM/MM) approach. *Journal of Computational Chemistry*. https://doi.org/10.1002/jcc.20222

Cifuentes-Pagano, E., Saha, J., Csányi, G., Al Ghouleh, I., Sahoo, S., Rodríguez, A., Skoda, E. M. (2013). Bridged tetrahydroisoquinolines as selective NADPH oxidase 2 (Nox 2) inhibitors. *MedChemComm*. https://doi.org/10.1039/c3md00061c

Cifuentes-Pagano, M., Meijles, D., & Pagano, P. (2015). Nox Inhibitors & Therapies: Rational Design of Peptidic and Small Molecule Inhibitors. *Current Pharmaceutical Design*. https://doi.org/10.2174/1381612821666151029112013

Cuddihy, S. L., Ali, S. S., Musiek, E. S., Lucero, J., Kopp, S. J., Morrow, J. D., & Dugan, L. L. (2008). Prolonged α-tocopherol deficiency decreases oxidative stress and unmasks α-tocopherol-dependent regulation of mitochondrial function in the brain. *Journal of Biological Chemistry*. https://doi.org/10.1074/jbc.M702572200

de Souza, A. S., de Oliveira, M. T., & Andricopulo, A. D. (2017). Development of a pharmacophore for cruzain using oxadiazoles as virtual molecular probes: quantitative structure–activity relationship studies. *Journal of Computer-Aided Molecular Design*. https://doi.org/10.1007/s10822-017-0039-0

Di Meo, S., Reed, T. T., Venditti, P., & Victor, V. M. (2016). Role of ROS and RNS Sources in Physiological and Pathological Conditions. *Oxidative Medicine and Cellular Longevity*. https://doi.org/10.1155/2016/1245049

Diebold, B. A., Smith, S. M. E., Li, Y., & Lambeth, J. D. (2015). NOX2 as a target for drug development: Indications, possible complications, and progress. *Antioxidants and Redox Signaling*. https://doi.org/10.1089/ars.2014.5862

Dixon, S. L., Duan, J., Smith, E., Von Bargen, C. D., Sherman, W., & Repasky, M. P. (2016). AutoQSAR: An automated machine learning tool for best-practice quantitative structure-activity relationship modeling. *Future Medicinal Chemistry*. https://doi.org/10.4155/fmc-2016-0093

El-Benna, J., Dang, P. M. C., Gougerot-Pocidalo, M. A., Marie, J. C., & Braut-Boucher, F. (2009). p47phox, the phagocyte NADPH oxidase/NOX2 organizer: Structure, phosphorylation and implication in diseases. *Experimental and Molecular Medicine*. https://doi.org/10.3858/emm.2009.41.4.058

Forrester, S. J., Kikuchi, D. S., Hernandes, M. S., Xu, Q., & Griendling, K. K. (2018). Reactive oxygen species in metabolic and inflammatory signaling. *Circulation Research*. https://doi.org/10.1161/CIRCRESAHA.117.311401

Friesner, R. A., Banks, J. L., Murphy, R. B., Halgren, T. A., Klicic, J. J., Mainz, D. T., Shenkin, P. S. (2004). Glide: A New Approach for Rapid, Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy. *Journal of Medicinal Chemistry*. https://doi.org/10.1021/jm0306430

Friesner, R. A., Murphy, R. B., Repasky, M. P., Frye, L. L., Greenwood, J. R., Halgren, T. A., … Mainz, D. T. (2006). Extra precision glide: Docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes. *Journal of Medicinal Chemistry*. https://doi.org/10.1021/jm051256o

Gawehn, E., Hiss, J. A., & Schneider, G. (2016). Deep Learning in Drug Discovery. *Molecular Informatics*. https://doi.org/10.1002/minf.201501008

Groemping, Y., Lapouge, K., Smerdon, S. J., & Rittinger, K. (2003). Molecular basis of phosphorylation-induced activation of the NADPH oxidase. *Cell*. https://doi.org/10.1016/S0092-8674(03)00314-3

Halgren, T. A., Murphy, R. B., Friesner, R. A., Beard, H. S., Frye, L. L., Pollard, W. T., & Banks, J. L. (2004). Glide: A New Approach for Rapid, Accurate Docking and Scoring. 2. Enrichment Factors in Database Screening. *Journal of Medicinal Chemistry*. https://doi.org/10.1021/jm030644s

Hassan, N. M., Alhossary, A. A., Mu, Y., & Kwoh, C. K. (2017). Protein-Ligand Blind Docking Using QuickVina-W with Inter-Process Spatio-Temporal Integration. *Scientific Reports*. https://doi.org/10.1038/s41598-017-15571-7

Jacobson, M. P., Pincus, D. L., Rapp, C. S., Day, T. J. F., Honig, B., Shaw, D. E., & Friesner, R. A. (2004). A Hierarchical Approach to All-Atom Protein Loop Prediction. *Proteins: Structure, Function and Genetics*. https://doi.org/10.1002/prot.10613

Jaquet, V., Scapozza, L., Clark, R. A., Krause, K. H., & Lambeth, J. D. (2009). Small-molecule nox inhibitors: ROS-generating NADPH oxidases as therapeutic targets. *Antioxidants and Redox Signaling*. https://doi.org/10.1089/ars.2009.2585

Jiang, Q. (2014). Natural forms of vitamin E: Metabolism, antioxidant, and anti-inflammatory activities and their role in disease prevention and therapy. *Free Radical Biology and Medicine*. https://doi.org/10.1016/j.freeradbiomed.2014.03.035

Jiang, Z., Yin, X., & Jiang, Q. (2011). Natural Forms of Vitamin E and 13′-Carboxychromanol, a Long-Chain Vitamin E Metabolite, Inhibit Leukotriene Generation from Stimulated Neutrophils by Blocking Calcium Influx and Suppressing 5-Lipoxygenase Activity, Respectively. *The Journal of Immunology*. https://doi.org/10.4049/jimmunol.1002342

Kami, K., Takeya, R., Sumimoto, H., & Kohda, D. (2002). Diverse recognition of non-PxxP peptide ligands by the SH3 domains from p67phox, Grb2 and Pex13p. *EMBO Journal*. https://doi.org/10.1093/emboj/cdf428

Lewis, E. D., Meydani, S. N., & Wu, D. (2019). Regulatory role of vitamin E in the immune system and inflammation. *IUBMB Life*. https://doi.org/10.1002/iub.1976

Li, H., Robertson, A. D., & Jensen, J. H. (2005). Very fast empirical prediction and rationalization of protein pK a values. *Proteins: Structure, Function and Genetics*. https://doi.org/10.1002/prot.20660

Li, J., Abel, R., Zhu, K., Cao, Y., Zhao, S., & Friesner, R. A. (2011). The VSGB 2.0 model: A next generation energy model for high resolution protein structure modeling. *Proteins: Structure, Function and Bioinformatics*. https://doi.org/10.1002/prot.23106

Liu, F. C., Yu, H. P., Chen, P. J., Yang, H. W., Chang, S. H., Tzeng, C. C., Hwang, T. L. (2019). A novel NOX2 inhibitor attenuates human neutrophil oxidative stress and ameliorates inflammatory arthritis in mice. *Redox Biology*. https://doi.org/10.1016/j.redox.2019.101273

Madhavi Sastry, G., Adzhigirey, M., Day, T., Annabhimoju, R., & Sherman, W. (2013). Protein and ligand preparation: Parameters, protocols, and influence on virtual screening enrichments. *Journal of Computer-Aided Molecular Design*. https://doi.org/10.1007/s10822-013-9644-8

Magnani, F., Nenci, S., Fananas, E. M., Ceccon, M., Romero, E., Fraaije, M. W., & Mattevi, A. (2017). Crystal structures and atomic model of NADPH oxidase. *Proceedings of the National Academy of Sciences of the United States of America*. https://doi.org/10.1073/pnas.1702293114

Martyna, G. J., Tobias, D. J., & Klein, M. L. (1994). Constant pressure molecular dynamics algorithms. *The Journal of Chemical Physics*. https://doi.org/10.1063/1.467468

Meydani, M. (2001). Nutrition interventions in aging and age-associated disease. *Annals of the New York Academy of Sciences*. https://doi.org/10.1111/j.1749-6632.2001.tb05652.x

Mittal, M., Siddiqui, M. R., Tran, K., Reddy, S. P., & Malik, A. B. (2014). Reactive oxygen species in inflammation and tissue injury. *Antioxidants and Redox Signaling*. https://doi.org/10.1089/ars.2012.5149

Nguyen, G. T., Green, E. R., & Mecsas, J. (2017). Neutrophils to the ROScue: Mechanisms of NADPH oxidase activation and bacterial resistance. *Frontiers in Cellular and Infection Microbiology*. https://doi.org/10.3389/fcimb.2017.00373

Nosé, S. (1984). A unified formulation of the constant temperature molecular dynamics methods. *The Journal of Chemical Physics*. https://doi.org/10.1063/1.447334

Panday, A., Sahoo, M. K., Osorio, D., & Batra, S. (2015). NADPH oxidases: An overview from structure to innate immunity-associated pathologies. *Cellular and Molecular Immunology*. https://doi.org/10.1038/cmi.2014.89

Pein, H., Ville, A., Pace, S., Temml, V., Garscha, U., Raasch, M., … Koeberle, A. (2018). Endogenous metabolites of vitamin E limit inflammation by targeting 5-lipoxygenase. *Nature Communications*. https://doi.org/10.1038/s41467-018-06158-5

Predescu, C., Lippert, R. A., Eastwood, M. P., Ierardi, D., Xu, H., Jensen, M., Shaw, D. E. (2012). Computationally efficient molecular dynamics integrators with improved sampling accuracy. *Molecular Physics*. https://doi.org/10.1080/00268976.2012.681311

Rastelli, G., Del Rio, A., Degliesposti, G., & Sgobba, M. (2010). Fast and accurate predictions of binding free energies using MM-PBSA and MM-GBSA. *Journal of Computational Chemistry*. https://doi.org/10.1002/jcc.21372

Rastogi, R., Geng, X., Li, F., & Ding, Y. (2017). NOX activation by subunit interaction and underlying mechanisms in disease. *Frontiers in Cellular Neuroscience*. https://doi.org/10.3389/fncel.2016.00301

Shelley, J. C., Cholleti, A., Frye, L. L., Greenwood, J. R., Timlin, M. R., & Uchimaya, M. (2007). Epik: A software program for pKa prediction and protonation state generation for drug-like molecules. *Journal of Computer-Aided Molecular Design*. https://doi.org/10.1007/s10822-007-9133-z

Sherman, W., Beard, H. S., & Farid, R. (2006). Use of an induced fit receptor structure in virtual screening. *Chemical Biology and Drug Design*. https://doi.org/10.1111/j.1747-0285.2005.00327.x

Shivakumar, D., Williams, J., Wu, Y., Damm, W., Shelley, J., & Sherman, W. (2010). Prediction of absolute solvation free energies using molecular dynamics free energy perturbation and the opls force field. *Journal of Chemical Theory and Computation*. https://doi.org/10.1021/ct900587b

Skonieczna, M., Hejmo, T., Poterala-Hejmo, A., Cieslar-Pobuda, A., & Buldak, R. J. (2017). NADPH Oxidases (NOX): Insights into Selected Functions and Mechanisms of Action in Cancer and Stem Cells. *Oxidative Medicine and Cellular Longevity*. https://doi.org/10.1155/2017/9420539

Sokol, R. J. (1988). Vitamin E deficiency and neurologic disease. *Annual Review of Nutrition*. https://doi.org/10.1146/annurev.nu.08.070188.002031

Sorce, S., Stocker, R., Seredenina, T., Holmdahl, R., Aguzzi, A., Chio, A., Jaquet, V. (2017). NADPH oxidases as drug targets and biomarkers in neurodegenerative diseases: What is the evidence? *Free Radical Biology and Medicine*. https://doi.org/10.1016/j.freeradbiomed.2017.08.006

Sumimoto, H., Miyano, K., & Takeya, R. (2005). Molecular composition and regulation of the Nox family NAD(P)H oxidases. *Biochemical and Biophysical Research Communications*. https://doi.org/10.1016/j.bbrc.2005.08.210

Takeya, R., Ueno, N., Kami, K., Taura, M., Kohjima, M., Izaki, T., Sumimoto, H. (2003). Novel Human Homologues of p47phox and p67phox Participate in Activation of Superoxide-producing NADPH Oxidases. *Journal of Biological Chemistry*. https://doi.org/10.1074/jbc.M212856200

Tarafdar, A., & Pula, G. (2018). The role of NADPH oxidases and oxidative stress in neurodegenerative disorders. *International Journal of Molecular Sciences*. https://doi.org/10.3390/ijms19123824

Xu, D., Jaroszewski, L., Li, Z., & Godzik, A. (2015). AIDA: Ab initio domain assembly for automated multi-domain protein structure prediction and domain-domain interaction prediction. *Bioinformatics*. https://doi.org/10.1093/bioinformatics/btv092

Xu, M., & Lill, M. A. (2013). Induced fit docking, and the use of QM/MM methods in docking. *Drug Discovery Today: Technologies*. https://doi.org/10.1016/j.ddtec.2013.02.003

Yokota, T., Uchihara, T., Kumagai, J., Shiojiri, T., Pang, J. J., Arita, M., Mizusawa, H. (2000). Postmortem study of ataxia with retinitis pigmentosa by mutation of the alpha-tocopherol transfer protein gene. *Journal of Neurology, Neurosurgery, and Psychiatry*.

Yuzawa, S., Suzuki, N. N., Fujioka, Y., Ogura, K., Sumimoto, H., & Inagaki, F. (2004). A molecular mechanism for autoinhibition of the tandem SH3 domains of p47phox, the regulatory subunit of the phagocyte NADPH oxidase. *Genes to Cells*. https://doi.org/10.1111/j.1356-9597.2004.00733.x

Zingg, J. M. (2015). Vitamin E: A Role in Signal Transduction. *Annual Review of Nutrition*. https://doi.org/10.1146/annurev-nutr-071714-034347

Zingg, J. M. (2019). Vitamin E: Regulatory Role on Signal Transduction. *IUBMB Life*. https://doi.org/10.1002/iub.1986

CIRRICULUM VITAE

ASENA HİMMETOĞLU

Graduated Student

A brain learner that seeks for physicomathematics

01.09.1995/ İstanbul

asena.himmetoglu@bahcesehir.edu.tr

asenahim@gmail.com

gencyazi.com/author/asenahimmetoglu/

+905530472453

İstanbul, Türkiye

EDUCATION

MSc in Bahçeşehir University

NeuroscienceMasterProgram

09/2018 – Present ***GPA: 3.58***

Thesis

- Machine Learning Algorithms and Combined Multiscale Molecular Modeling Simulations Against NADPH OXIDASE (NOX) Enzymes for Designing of Small Molecule Therapeutics

Bachelor's in Computer Engineering

İstanbul Altınbaş University

09/2013 – 02/2018

Thesis

- Brain Imagery

WORK EXPERIENCE

Project Assistant

Bahçeşehir University Computational Biology & Molecular Simulations Laboratory

02/2019 – Present

Achievements/Tasks

- AutoQSAR Modeling
- Molecular Docking
- Molecular Dynamics

Intern

Hpc Bulut Bilişim İTU CBS ARGE Center

07/2016 – 08/2016

Achievements/Tasks

- Basic Android Programming

Intern

T-Soft /Davutpaşa Teknopark

06/2016 – 07/2016

Achievements/Tasks

- Web Programming
- PHP, XML,Array functions, Filtering,
- .htacces and Session function

PROJECTS

- Novel Fibroblast Growth Factor Receptor-2 (FGFR-2) Blockers Using Integrated
- Machine Learning and Molecular Modeling Approaches(10/2019)
  AutoQSAR, Molecular Docking, Molecular Dynamics

- Development of Machine Learning-Based Novel Anti- Cancer Therapeutics Against Malignant Glioma (04/2019)
  AutoQSAR modelling, Molecular Docking, Molecular Dynamics

- Comenius Interchange Project (05/2010 – 06/2011)
  "Cultural Interaction Between Greece and Turkey" Project

CERTIFICATES

- 7th International BAU Drug Design Congress (12/2019)

- İstanbul University Facultyof Medicine Neuroscience Club Neuroscience Congress - 6 (09/2018)

- Philosophy Conferences (04/2014)
  Altınbaş University- H.Kaan Ökten

- Psychology Days (05/2014)
  TPÖÇG ve Altınbaş University

POSTER PRESENTATION

Novel Fibroblast Growth Factor Receptor-2 (FGFR-2) Blockers Using Integrated Machine Learning and Molecular Modeling Approaches (12/2019)

7th International BAU Drug Design Congress

LANGUAGES

- English
  Full Professional Proficiency
- German
  Elementary Proficiency
- Russian
  Elementary Proficiency

INTERESTS

- Nature
- Psychology
- Cognition
- Philosophy
- Writing
- Photography
- Drawing and Painting