UNDERSTANDING THE ANTI-INFLUENZA ACTIVITY OF BIOLOGICALLY ACTIVE COMPOUNDS USING PHARMACOPHORE, MOLECULAR DOCKING AND QSAR

Researcher:

Mehwish Huma

REG NO: 19-FBAS/MSBI/F08

Supervisor:

Dr. Naveeda Riaz

Ms. Saima Kalsoom

Department of Bioinformatics and Biotechnology
Faculty of Basic and Applied Sciences
INTERNATIONAL ISLAMIC UNIVERSITY, ISLAMABAD

Accession No TH-9304

DATA ENTERED 7/14/2

MS
572.80285
MELI
methodology
(1)

Understanding the Anti-Influenza Activity of Biologically Active Compounds Using Pharmacophore, Molecular Docking and QSAR

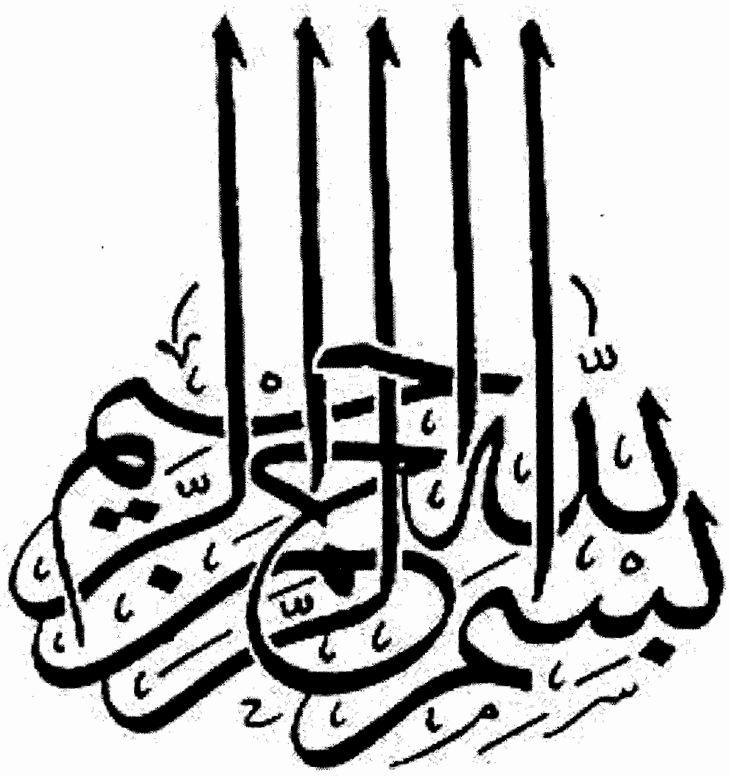
Researcher

Mehwish Huma

19-FBAS/MSBI/F08

Submitted as partial requirement for the fulfillment of MS in Bioinformatics at Department of Bioinformatics and Biotechnology, Faculty of Basic & Applied Sciences, International Islamic University, Islamabad.

Supervisor: Dr. Naveeda Riaz.
Co-Supervisor: Miss Saima Kalsoom.



In the name of Allah, the Most Beneficent, the Most Merciful.

DECLARATION

I hereby affirm that the work presented in the following thesis is my own endeavor, except where accredited otherwise, and that the thesis is my own composition. No part of the thesis has been beforehand presented for any other degree.

Mehwish Huma.

Dedicated to Allah (S.W.T), The Creator of Everything that Exists under the Sun and above it…the Lightning Bolt that Destroys a House and the Hand of a Man that Rebuilds it.

Certificate

Title of Thesis: Understanding the Anti-Influenza Activity of Biologically Active Compounds Using Pharmacophore, Molecular Docking and QSAR

Name of Student: Mehwish Huma

Registration No: 19/FBAS/MSBI/F08

Accepted by the Department of Bioinformatics and Biotechnology, Faculty of Basic and Applied Sciences, International Islamic University Islamabad as a requirement for the fulfillment for Master of Science in Bioinformatics

Viva Voce Committee

Chairman/Director/Head

Dr. Irfan Khan
Professor Dept BI& BT

External Examiner

Internal Examiner

Supervisor

Dr. Naveeda Riaz
Assistance Professor Dept. BI & BT

Mrs. Saima Kalsoom
PhD Scholar, QAU

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

2D: Two dimension

3D: Three Dimension

Å: Angstrom

AFMOC: Adaptation of Fields for Molecular Comparison

ALA: Alanine

ARG: Arginine

ASN: Aspargine

ASP: Aspartic Acid

CADD: Computer Aided Drug Design

CDC: U.S. Centers for Disease Control and Prevention

COMBINE: Comparative Binding Energy Analysis

CoMFA: Comparative Molecular Field Analysis

CoMMA: Comparative Molecular moment Analysis

CoMSA: Comparative Molecular Surface Analysis

CoMSIA: Comparative Molecular Similarity Indices Analysis

CoRIA: Comparative Residue Interaction Analysis

DNA: Deoxyribose Nucleic Acid

FDA: Food and Drug Administration

GALAHAD: A Genetic Algorithm with Linear Assignment for the Hypermolecular Alignment of Datasets

GASP: Genetic Algorithm Superimposition Program

GERM: Genetically Evolved Receptor Models

GLU: Glutamic Acid

GLY: Glycine

HA: Hemagglutanin

HASL: Hypothetical Active Site Lattice

HBA: Hydrogen Bond Acceptor

HBD: Hydrogen bond Donor

HCV: Hepatitis C Virus

H_f: Heat of Formation

HOMO: Highest Occupied Molecular Orbital

IAV: Influenza A Virus

IC_{50}: Inhibitory Concentration

ILE: Isoleucine

IUPAC: International Union of Pure and Applied Chemistry

Log P: Partition Coefficient

LEU: Leucine

LUMO: Lowest Unoccupied Molecular Orbital

M1: Matrix Protein 1

M2: Matrix Protein 2

MOE: Molecular Orbital Environment

MR: Molecular Refractivity

MSA: Molecular Shape Analysis

NA: Neuraminidase

NI: Negative Ionizable

NP: Nucleocapsid Protein

NS: Non Structural protein

PA: Active Polymerase

PB1: Catalytic Subunit of Polymerase

PB2: Host Cap Binding Polymerase

PDB: Protein Data Bank

PI: Positive Ionizable

PRO: Proline

QSAR: Quantitative Structure Activity Relationship

QSPR: Quantitative Structure Property Relationship

QSTR: Quantitative Structure Toxicity Relationship

RNA: Ribonucleic Acid

RNP: Ribonucleocaspid

SER: Serine

TRP: Tryptophan

TYR: Tyrosine

VAL: Valine

VMD: Visual Molecular Dynamics

ACKNOWLEDGMENTS

At the outset I express thanks to Allah (S.W.T), the Most Knowledgeable and the Al Mighty. He is the creator of the sun as a source of a radiant light and the moon shining, the day to find sustenance and the night for stillness, and it is Him who created everything in the best of forms and perfected them. All my thanks goes to Him who created me as a human being and imparted unto me articulate thought and speech without which I could not be coherent on my basics lest perform this work of significance.

I m grateful to the loving buttress of my father and mother who encouraged me when I felt low, who became the ray of hope at times of despair, who showed me light in the darkest hours and most of all who forfeited their rest, time, desires, money to make me a strong being I am today. In addition I thank my sisters and brothers to help me where ever and whenever I needed it and always advised me to think positive when negativity plagued me inside out.

Thanks and appreciation seems like small words when it comes to **Miss. Saima Kalsoom**, a PHD scholar of Quaid-e-Azam University, Islamabad. She is the exclusive initiator of the outline I worked on, to reach the goal and present this work with full confidence. I am indebted to her for showing me the path when I went astray and treating me with affection even when I was being unreasonable. I admire her undaunted support during my research work.

Furthermore **Dr. Naveeda Riaz**, assistant professor, International Islamic University, Islamabad, is also among the list of people without whom I would not be able to complete this job in due course of time.

Last but not the least I credit the whole badge of **MSBI-2**, especially Maimoona Ali to be my cohort at good times and forte at worst ones during my time at International Islamic University, Islamabad.

In the end, I pray to Allah (SWT) to forgive me for my lacking and bless me and all those who supported me with success in this world and the hereafter (Ameen).

Mehwish Huma.

ABSTRACT

An influenza outbreak takes its toll on population count as well as economic stability of a region. The first line of defense in any influenza outbreak is the FDA approved anti influenza drugs. Neuraminidase, a protein target, has been studied for drug designing potential in recent years that resulted in the form of Oseltamivir and Zanamivir. Both drugs tested on the crystallographic structures of N2 and N9 are potent for all types due to conservation of active site residues. Influenza inhibitors of four different classes showing activity in biological assays have been studied to exploit their potential for Computer Aided Drug Designing (CADD) by means of Pharmacophore Generation, Molecular Docking. Flavonoids, a novel class of compounds, have been scrutinized for Quantative Structure Activity Relationship (QSAR). Conclusively, a 2-D pharmacophore has been proposed that a compound should have at least one Hydrogen Bond Donor (HBD) and two Hydrogen Bond Acceptors (HBA), to show activity as a neuraminidase inhibitor. In addition, analogue compounds 1 and 2 are proposed to be tested in biological assays as they show maximum interactions with the target site in molecular docking studies. Lowest Unoccupied Molecular Orbital energy (LUMO), Heat of Formation (H_f) and Molar Refractivity (MR) are pin pointed to have a direct correlation with activity in Flavonoids.

CHAPTER 1

INTRODUCTION

Death and disease are a part of human life since the dawn of history. Human health is constantly at the peril of being impaired with different external factors (viruses, bacteria) as well as some internal factors (genetic changes). These impairments lead to a number of diseases. Some of these diseases are still considered very hazardous (such as cancer, hepatitis and AIDS) while there are others that once were thought out as deadly are curable now (such as malaria and tuberculosis) with the advancement in therapeutic methods.

In this plethora of diseases, Influenza is unique in the sense that it is lethal as well as harmless at the same time. Influenza, commonly called 'flu', is a viral infection that affects mainly nose, throat, bronchi and occasionally lungs. Infection usually lasts for about a week, and is characterized by sudden onset of high fever, aching muscles, headache and severe malaise, non-productive cough, sore throat and rhinitis[1].

Influenza continues to affect a large portion of population in modern times, irrespective of age and previous infection history. Moreover, there have been periodic influenza epidemics that have devastated otherwise healthy populations worldwide. Some pandemics like 1917-1919 "Spanish flu" were responsible for wiping out tens of millions of people from the face of earth (Kilbourne, 1987). Since then influenza

[1] http://www.who.int/

epidemics keep recurring from time to time as is evident from the statistics brought to notice in table 1.

Table 1: Morality Associated with Influenza Events

Mortality Associated with Influenza Pandemics and Selected Seasonal Epidemic Events, 1918-2009*		
Years	**Circulating virus (genetic mechanism)**	**Excess Deaths from Any Cause** *No. per 100,000 persons/yr*
1918-1919	H1N1 (viral introduction), pandemic	598.0
1928-1929	H1N1 (drift)	96.7
1934-1936	H1N1(drift)	52.0
1947-1948	H1N1 A (intrasubtypic reassortment)	8.9
1951-1953	H1N1 (intrasubtypic reassortment)	34.1
1957-1958	H2N2 (antigenic shift), pandemic	40.6
1968-1969	H3N2 (antigenic shift), pandemic	16.9
1972-1973	H3N2 A Port Chalmers (drift)	11.8
1975-1976	H3N2 (drift) and H1N1 ("swine flu" outbreak)	12.4
1977-1978	H3N2 (drift) and H1N1 (viral return)	21.0
1997-1999	H3N2 A Sydney (intrasubtypic reassortment) and H1N1 (drift)	49.5
2003-2004	H3N2 A Fujian (intrasubtypic reassortment) and H1N1 (drift)	17.1
2009	H3N2 and H1N1 (drift) and swine-origin H1N1 (viral introduction), pandemic	?

*Mortality data includes deaths associated with all influenza A and B viruses combined (D. M. Morens, 2009).

According to statistics provided by Centre for Disease Control and prevention (CDC), the influenza associated deaths between 40th week of 2006 and March 6, 2010 are shown in Figure 1. The hike in mortality rate, pink peak, shows the Swine Influenza outbreak in 2009. Death rate per infection as reported in Pakistan is 1 in 12 cases[2]. The first line of defense in any influenza pandemic is over the counter anti influenza drugs.

[2] http://www.flucount.org/

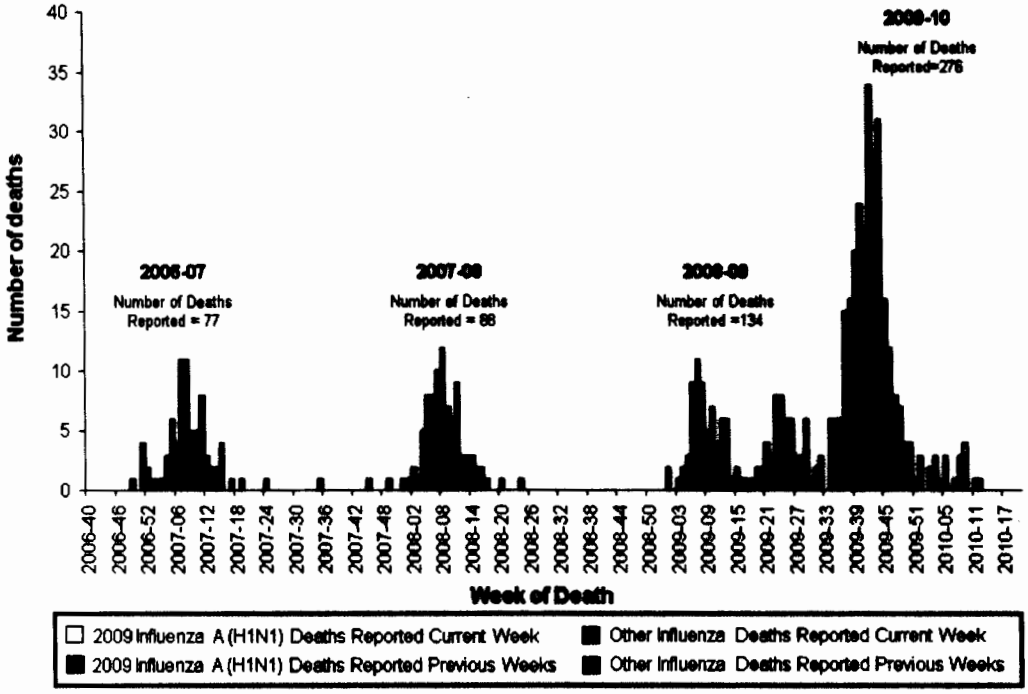


Figure 1: Number of Influenza Associated Deaths from 2006-2007 Season to present (http://www.cdc.gov/).

Extensive research has been conducted over the past century to deal with unpredictable influenza outbreaks. The line of investigation resulted into several US Food and Drug Administration (FDA) approved drugs as Amantadine, Rimantadine, Zanamivir and Oseltamivir (Richard J Sugrue *et al.*, 2008). Inadequacy of these drugs such as side effects, problems in routes of administration and emergence of drug resistant strains opens up the horizon for the development of new antiviral strategies to thwart the next pandemic.

Zanamivir and Oseltamivir are two classic examples of Computer Aided Drug Designing (CADD) against Neuraminidase (NA) protein of Influenza A virus. NA plays a very important role in life cycle of influenza virus by helping in budding of the progeny viruses (Cobb, 2007). Structural insights have shown that active site of NA protein is highly conserved in its all subtypes (currently 9 subtypes of NA of

Influenza A Virus are known). This high level safeguard of active site from mutations points to the importance of NA.

It is evident from table 1 that N1 and N2 have stably adapted to the human population (D. M. Morens, 2009). Currently available drugs against NA are designed by conducting studies on crystal structures of N2 and N9 but were found potent for other NA's as well (Xu *et al.*, 2008). This confirmed the conservation of active site amino acids in all subtypes of NAs and emphasized the importance of NA in cell cycle of Influenza virus. N1 has remained a major contributor in the viruses (e.g. Spanish Influenza Virus H1N1, Avian Influenza Virus H5N1 and Swine Influenza Virus H1N1) that wiped millions of people from the face of earth. Moreover, current events have revealed that drug resistant mutants can be subtype specific which points towards modifications in the active site, its geometry or inhibitor binding fashion in the different subtypes (Xu *et al.*, 2008).

Recently a study showed that NA of 2009 avian flu virus (H5N1) has a cleft in the vicinity of substrate binding site of NA. This cavity was also found in 1918 Spanish flu virus (H1N1). This cavity was conserved among Group 1 NAs (N1, N4, N5 and N8). The cleft appears from an open conformation of loop 150 (Gly 147 to Asp 151) of NA and closes upon Zanamivir binding (Xu *et al.*, 2008). The potential of this cavity for drug designing has not been manipulated yet and can be important for designing subtype specific drugs.

Bearing in mind the above stated facts, need of the hour is to check anti influenza potential of loop 150 and surrounding amino acids of Group 1 NAs. This study is targeted to manipulate the potential of biologically active compounds to work as NA inhibitors to come up with a subtype specific drug. The work is intended to gain

understanding of the inhibitory potential of biologically active compounds using three techniques of CADD that are:

- Pharmacophore modeling,
- Molecular Docking and
- Quantitative Structure Activity Relationship (QSAR).

As an upshot, it is anticipated to concoct a new ray of hope for the society that is constantly at the risk of viral pandemics. This study is also optimistic about the outcome in order to come up with following features:

- A new drug with more effectiveness and less administration problems.
- A pharmacophore model integrating the classes of compounds under discussion.
- Descriptors directly influencing the biological activity of compounds.
- Amino acids more important in drug substrate interactions in loop 150 and its neighborhood.

A growing interest has been observed in finding new anti influenza drugs due to three influenza outbreaks in last two decades, which extended from poultry to mammals and most significantly to humans. It will be of assistance to the society if one comes up with subtype specific as well as general anti influenza inhibitors that can be used in case of any forth coming pandemic or epidemic to lessen the damage anticipated.

CHAPTER 2

LITERATURE REVIEW

2.1 Influenza Virus

Influenza virus belongs to the family orthomyoxyviridae[3]. The family contains five genera, three of which cause influenza like symptoms in mammals, birds and human. These genera are Influenza A virus, Influenza B virus and Influenza C. Of the three types only type A and B are cause of extensive upsurge of disease. Type A virus has been extensively studied for a vast variety of its genome segments (K. G. Nicholson, 2003). Influenza A is a major cause of disease in humans in the form of annual epidemics and sporadic pandemics (C. Pappas *et al.*, 2008).

2.1.1 Influenza Virus Genetics

Influenza A virus (IAV) contains a linear, negative sense, single stranded RNA (ribonucleic acid). The viral genome is segmented into eight parts. Each genome segment encodes one protein except two segments that encode two proteins each via alternate splicing. The virion is enveloped and has two integral membrane glycoproteins hemagglutinin (HA) and neuraminidase (NA) (Brown, 2000). As is evident from Figure 2, genomic RNA segments are encapsulated with nucelocapsid protein (NP) and three connected subunits of viral polymerases PB1 (catalytic subunit

[3] International Committee on Taxonomy of Viruses Index of Viruses - Orthomyoxyviridae (2006). In: ICTVdB - The Universal Virus Database, version 4. Buchan-Osmond, C (Ed), Columbia University, New York, USA.

of polymerase), PB2 (involved in host cap binding and endonuclease) and PA (subunit of polymerase active in viral RNA synthesis) to form the coiled ribonucleocaspids (RNP) (Cox *et al.*, 1998; Huang *et al.*, 1990).

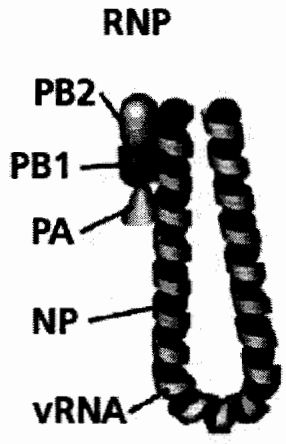


Figure 2: Ribonucleoprotein (RNP) consisting of RNA segments associated with Nucleoprotein (NP) and the PA, PB1 and PB2 Polymerase Proteins (Cox *et al.*, 1998).

There are two non structural (NS) proteins coded by the same region of RNA through alternate splicing. NS1 protein is involved in the RNP nuclear export while NS2 protein is present in the cytoplasm and acts as interferon response inhibitor (Brown, 2000). M2 protein is excessively expressed on the envelope and acts as ion channels permitting ions to enter the virion and maintain the PH (Holsinger *et al.*, 1993). The matrix protein (M1) is present in complex with RNP and NS2 (involved in export of RNP on infecting the host cell) at the interior surface of the envelope to interface with the cytoplasmic projections of HA, NA and M2 protein. M1 and M2 proteins are coded by same RNA segment (Brown, 2000; Cox *et al.*, 1998).

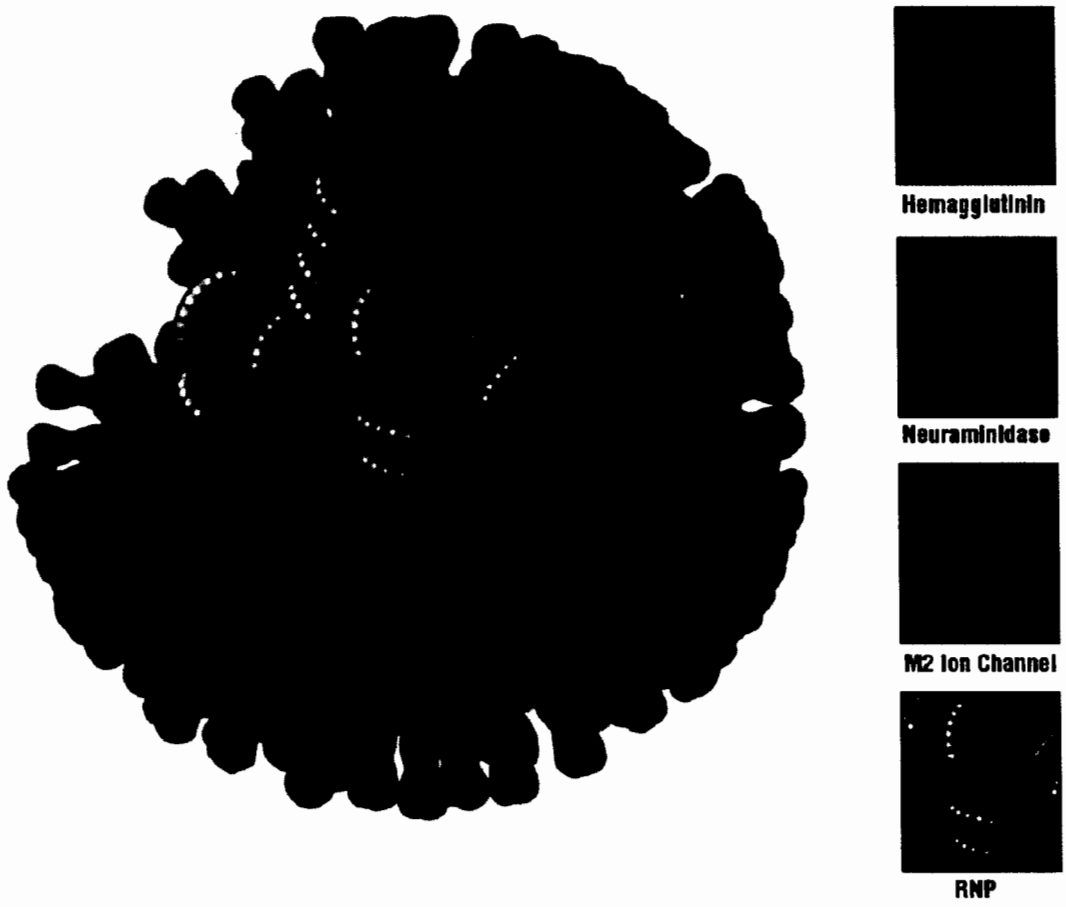


Figure 3: Schematic Diagram of Influenza A Virion[4].

IAV are further classified into various subtypes on the basis of antigenic differences between these two surface proteins. Currently, there are 16 known types of HA and 9 known types of NA (K. G. Nicholson *et al.*, 2003; R. A. M. Fouchier, 2005).

2.1.2 Life Cycle

Influenza virus attacks the host by attaching to the host cell receptors through HA binding. HA binds the host cell by interacting with the receptors containing a terminal neuramic acid residue (Gubareva *et al.*, 2000). Virion is engulfed by the host cell via

[4] http://www.cdc.gov/h1n1flu/images/3D_Influenza_transparent_key_pieslice_lrg.gif

the process of endocytosis to be uncoated in the endosomes (Brown, 2000). Two significant events occur on uncoating:

- M2 channels transport protons to the interior of the virion resulting in low PH. This low PH mediates the separation of M1 matrix from RNP.
- Low PH mediated conformational change on HA leading to the fusion of viral envelope with endosomal membrane. Conclusively, RNP is released into the cytoplasm (Cox *et al.*, 1998; Brown, 2000).

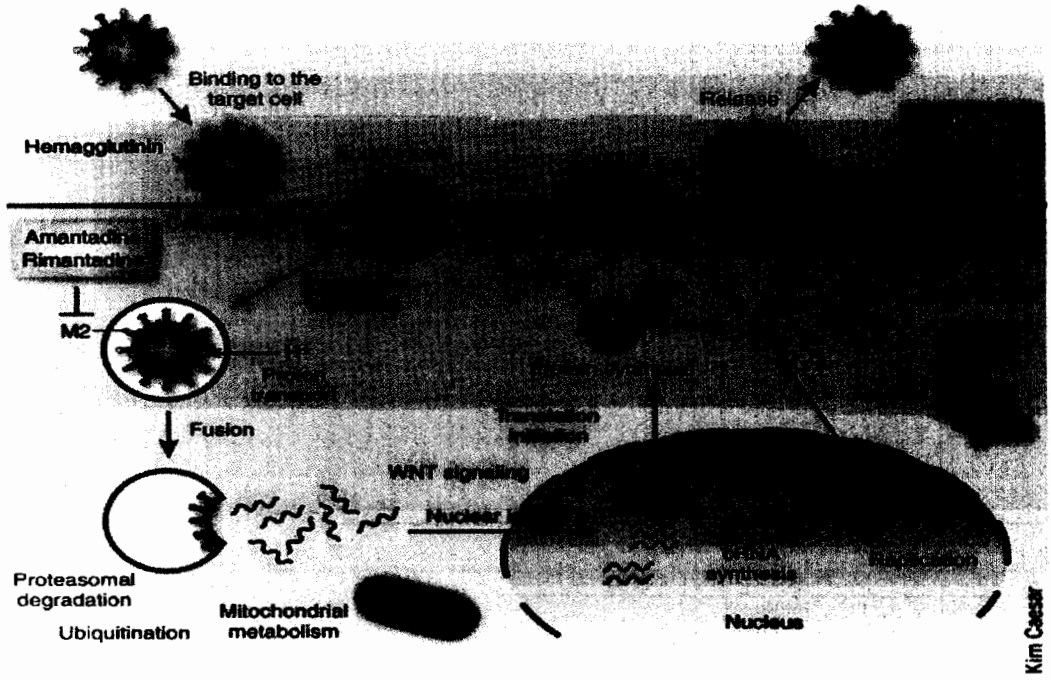


Figure 4: Influenza virus replication cycle (Ji-Young Min, 2010).

RNPs once free from viral envelope migrate to the host cell nucleus through nuclear pore (Moorman, 2003). Replication and transcription of RNP starts in the nucleus. The primary transcripts thus formed are transported back to cytoplasm to be translated by the host cell machinery (Cox *et al.*, 1998). After being translated, the newly formed proteins HA, NA and M2 are translocated to the host cell endoplasmic reticulum. After the glycocylation of these proteins they are transported to Golgi apparatus from where they are transported to the plasma membrane (Cox *et al.*, 1998).

Apparently, interactions between M1 protein and the protrusions of HA and NA proteins are involved in directing and compelling the final stage of virion assembly i.e. budding (Brown, 2000). NA plays its role in budding of the infant virions by destroying the receptors recognized HA by cleaving α-ketosidic bond linkage between HA and host cell receptors having terminal neuramic acid (Gubareva, 2000). A balance between HA activity and NA activity is necessary for the docking of virion to the host cell and release of progeny virus from the host (Mitnaul *et al.*, 2000).

2.1.3 Existing Drug Targets

A number of targets have been identified and studied for antiviral therapy in the past years. These targets can be classified on the basis of their involvement in the life cycle as:

1. Binding to the target cell
2. Uptake into the cell
3. Transcription and replication
4. Assembly and budding (Beagle and Bray, 2008).

At the binding stage infection is countered with the drugs effective against HA by either blocking the HA binding to the host cell sialic acid or by inducing the antibody response against HA (Beagle and Bray, 2008). M2 ion channels help in the uptake of virion into the cell by creating the low PH. Till date two drugs have been approved against M2 protein, amantadine and rimantadine, which work by blocking the M2 channel. Conclusively virion uncoating is halted (Richard J Sugrue *et al.*, 2008). Drugs resistant viruses are readily generated against these FDA approved drugs, hence they are not recommended for use in seasonal influenza (Richard J Sugrue *et al.*, 2008).

Ribavirin, Viramidine and T-705 have shown meaning full results by terminating the viral replication and transcription. Ribavirin's clinical utility is restricted due to the risk of hemolytic anemia in patients while Viramidine and T-705 are in their testing phases. Partly as a result of amantadine resistant viruses surfacing other antiviral strategies are being considered. Consequently, Zanamivir and Oseltamivir are two FDA approved drugs targeted to influenza virus neuraminidase which play its prime role in budding (Richard J Sugrue *et al.*, 2008).

Recent studies have shown that other viral components such as NS1, M1, polymerase and NP are considered as future targets for antiviral drugs (Basu *et al.*, 2009; Nakazawa *et al.*, 2008; Das *et al.*, 2010; Krug and Aramini, 2009; Vivek Darapaneni *et al.*, 2009).

Even with the two classes of drugs available for influenza, human race is still immunologically unprepared to counter with the attack of seasonal influenza (Basu *et al.*, 2009). Moreover, there is a constant threat of recurring epidemics that arise from time to time in one part of the world or other. In fact we are living in a flu emergency area which is always red alert for new influenza virus strains that arise either by antigenic shift or genetic drift (D. M. Morens *et al.*, 2009). There is a need for novel inhibitors that can be used either alone or in combination with other therapies.

2.1.4 Neuraminidase- A Promising Drug Target

Though several anti viral targets are under study to make an efficient drug against influenza viruses, NA has its own prominent place in the field of drug discovery for IAV. NA is a hydrolase that has enzymatic activity in cleaving α ketosidic bond linkage (Nguyen *et al.*, 2010). This enzymatic activity is involved in destroying the receptors that bind to influenza virus HA. When influenza viruses lacks the NA

activity, the infant virus aggregate on the cell surface and are not able to bud off and the viral spread is rigorously impaired (J. B. Finley *et al.*, 1999).

It is vital for the virus to move to and from the site of infection (Hayden and Ison, 2001). Cleavage of HA receptors by NA is crucial to avert attachment of infant viruses with another and also with the cell membrane (Hayden and Ison, 2001). It is also believed that NA has its role in helping the virus penetrate the thick mucosal layer and reach the epithelial cells (Colman, 1994).

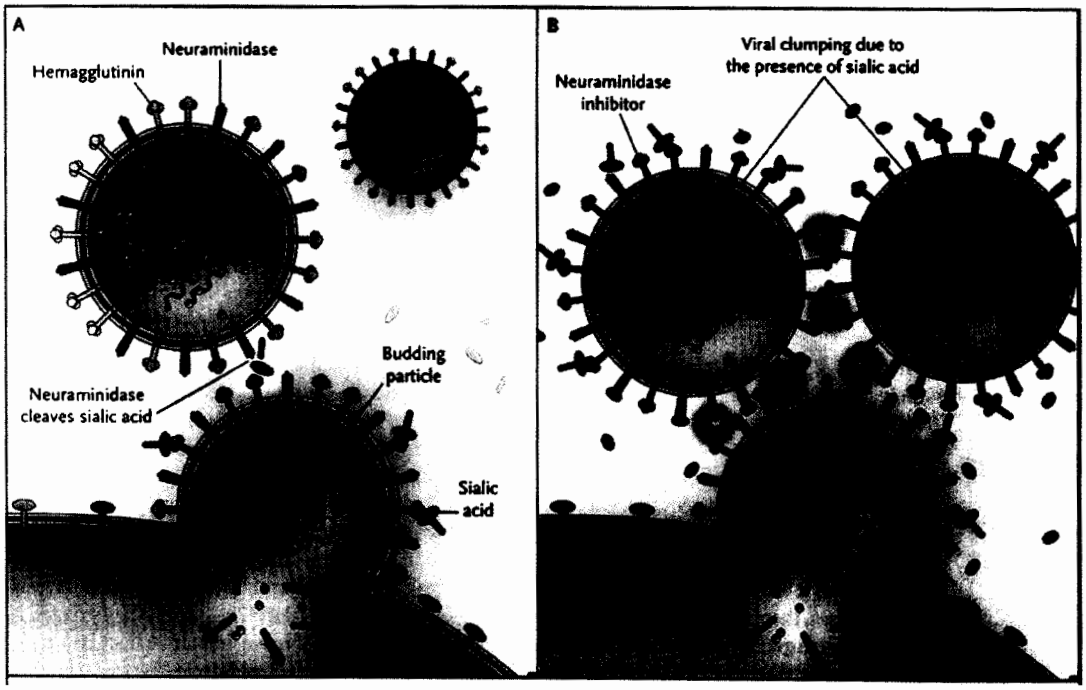


Figure 5: Normal Budding and Release of Influenza Virus from an Infected Cell (Panel A) and Release Restricted by a Neuraminidase Inhibitor (Panel B) (Monto, 2005).

NA is a mushroom like tetramer consisting of four more or less spherical subunits, 60 KDa each, having the centrally located binding site of the natural ligand sialic acid (Colman, 1995). Protein can be divided into a head having the carboxylic end, a stalk that is attached to the head by means of hydrophobic residues, a transmembrane portion and an intracellular amino terminal with six highly conserved polar residues as shown in Figure 6 (Cox *et al.*, 1998).

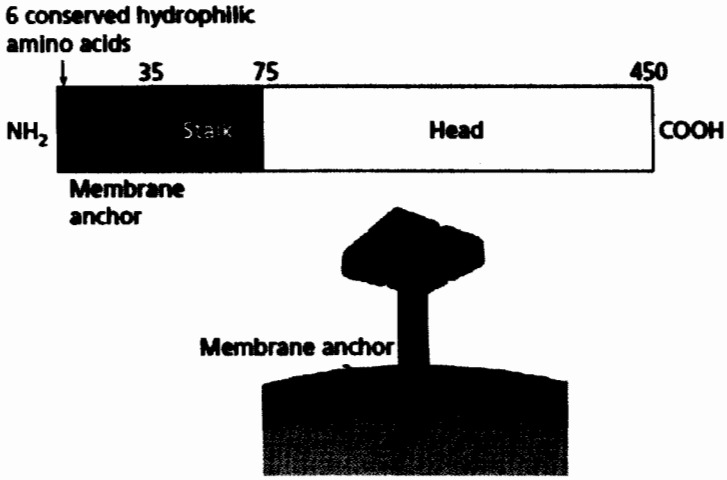


Figure 6: Schematic Diagram of Neuraminidase Molecule (Cox *et al.*, 1998).

Each monomeric subunit has six twisted β sheets consisting of four anti parallel strands arranged in a propeller style as seen in Figure 7 (J. B. Finley *et al.*, 1999). Active site can be seen in each monomer with an attached ligand in the figure below.

Figure 7: Schematic Illustration of Neuraminidase Tetramer[5].

[5] http://www.pdb.org/

Each monomer has an active site lined with 10 conserved residues that are common to the active site of all Neuraminidase subtypes of IAV as shown in Figure 8 (J. Zhang *et al.*, 2007). The inhibitors bind the active site by means of a negative ionizable group to any of the three arginine residues, Arg118, Arg 371 and Arg 292, present at one side of the active site. On the other side of the binding cavity hydrophobic pocket is formed by the residues Trp178 and Ile 222. A negatively charged pocket is presented by the presence of Glu119, Asp151 and Glu227 on the third side. The final part consists of Glu276 and Glu277 (J. Zhang *et al.*, 2007; J. Gong *et al.*, 2007).

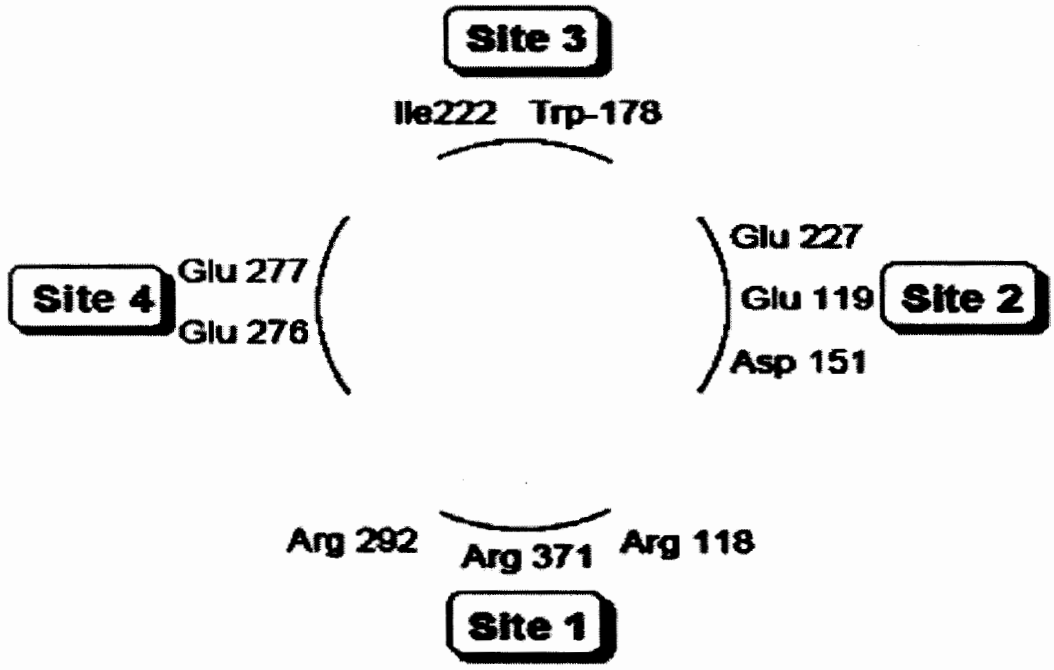


Figure 8: Air Plane Model of Active Site of Neuraminidase (J. Zhang *et al.*, 2007).

Out of the nine subtypes of NA, only N1 and N2 have truly become human adapted viruses (D. M. Morens *et al.*, 2009). Zanamivir and Oseltamivir designed against the crystallographic structure of N2 and N9 have successfully inhibited other types of NA including N1 due to the striking similarity in all active site of all NA types (Xu *et al.*, 2008). In this study N1 is considered the potential protein for

antiviral therapy as it was part of the first most devastating pandemic in 1918 and has remained a part of recurrent epidemics in recent years.

The H1N1 virus associated with 2009 swine flu epidemic is a 4th generation descendant of the H1N1 1918 virus (D. M. Morens *et al.*, 2009). The protein considered for the current study was the first recombinant reported from the 1918 virus and has a binding cavity, 150 loop, close to the active site as shown in Figure 9. This cavity is present in all N1 neuraminidases as well as N4, N5 and N8 (Xu *et al.*, 2008). It has also been reported that this cleft closes upon Zanamivir binding to the active site proposing that it could be further studied as a probable drug target (Xu *et al.*, 2008).

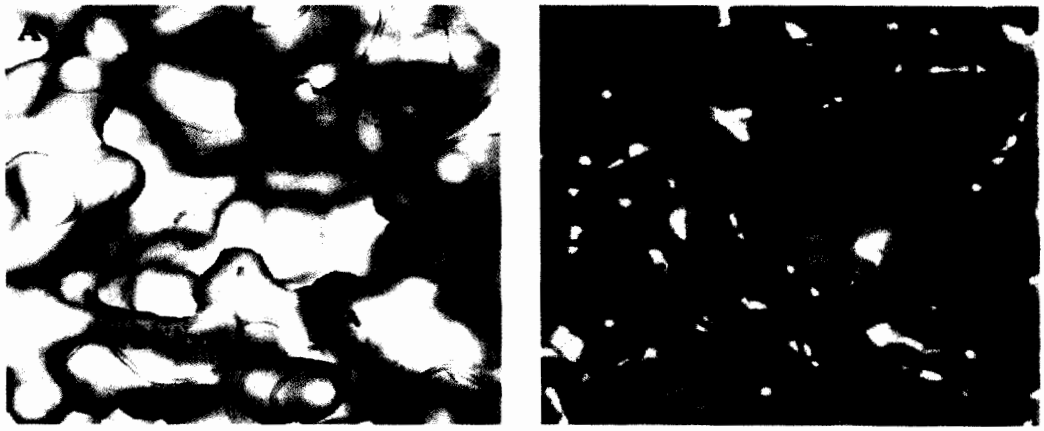


Figure 9: Molecular surface of the NA active site with and without Zanamivir (A) Native structure showing the large 150 cavity that is close to the unoccupied Zanamivir binding site (B) Inhibitor bound structure, in which structural changes bring the 150 loop proximal to the Zanamivir binding site and, therefore, close the 150 cavity (Xu *et al.*, 2008).

2.1.5 Neuraminidase Inhibitors

FDA has till date approved two drugs, Zanamivir and Oseltamivir, for neuraminidase target to be used for preventive use and uncomplicated illness by Influenza A and B virus (Richard J Sugrue *et al.*, 2008). Ribavirin, a nucleoside analogue, hinders replication of either DNA or RNA and has shown evocative results

in tissue culture studies of IAV, but results in influenza patients were not consistent (Richard J Sugrue *et al.*, 2008). Ribavirin is FDA approved for prevention again HCV (Hepatitis C Virus) but not for Influenza (Richard J Sugrue *et al.*, 2008; Beigel and Bray, 2008).

Zanamivir was the first NA inhibitor commercially developed with the market name Relenza® (FDA approval: 1999). The drug is given through intranasal passage to the patients above 7 years of age (Cox *et al.*, 1998). The low bioavailability of drug, small volume of distribution and speedy renal elimination as well as uselessness for severely ill patients who are unable to inhale has opened horizon for discovery of new NA inhibitors (T. Honda *et al.*, 2002).

Oseltamivir developed under the market name Tamiflu® (FDA approval: 1999) was the first orally active drug for influenza commercially prepared. Oseltamivir though having higher bioavailability than zanamivir has some serious side effects including vomiting and nausea (H. J. Jeong *et al.*, 2009). In some studies neurological disturbances have also been reported (Richard J Sugrue *et al.*, 2008). Moreover, some recent reports have also suggested resistance of Oseltamivir in 2009 swine flu (T. Rungrotmongkol *et al.*, 2009).

The route of administration for Zanamivir has been a major reason for its limited use by the public. Oseltamivir though orally active has other complications. Need of the hour, in view of the problems with standard drugs, is to put efforts in designing a new drug that show better results and prepares the human race for any upcoming epidemic by IAV (Hammad and Taha, 2009). The problem thus stated is being addressed by this study.

2.1.6 Classes of Neuraminidase Inhibitors

Several classes of compounds have shown their activity against IAV Neuraminidase and are potential candidates to be considered for drug designing against IAV. Four main structural classes of compounds are identified:

1. Zanamivir derivatives,
2. Thiourea analogues,
3. Polyphenols and
4. Carbocyclic compounds.

Zanamivir, a heterocyclic compound, derivatives are a constantly in the spot light for drug design against IAV to counteract with the low availability problem of zanamivir (T. Honda *et al.*, 2002). Thiourea is a group of compounds structurally similar to urea but differ a lot in their significance. Thiourea is being used as a drug in treating hyperthyroidism as well as Herpes Simplex Virus induced disease (Zeijl *et al.*, 2000). Potential of Thiourea to inhibit NA has also been reported (P. C. Nair and M. E. Sobhia, 2008). Polyphenols is a class of compounds which occur largely in herbs and nutritional products (Wang *et al.*, 2005). These compounds are studied at large scale for their therapeutic potential and are found to have antibacterial, antifungal, antiviral, antioxidant, antidiabetic, antithrombogenic and antineoplastic activity (Tapas *et al.*, 2008; W. Bylka *et al.*, 2004). The possibility of these compounds to act as NA inhibitors has been tested in recent years (H. J. Jeong *et al.*, 2009; Y. B. Ryu *et al.*, 2010; A. G. Mercader and A. B. Pomilio, 2010; A. L. Liu *et al.*, 2008). Carbocyclic compounds are being considered for potential inhibitors of NA even before the approval of Zanamivir and Oseltamivir. The activity of these compounds has been reported in a few studies (M. A. Williams *et al.*, 1997; W. Lew

et al., 2000). Keeping in view the therapeutic potential of each class of compounds, a hybrid data set was considered for the study having several compounds from each class.

2.2 Computer Aided Drug Designing

In the past times, people tried to treat diseases by herbal products that were found to have therapeutic benefits by fortunate accidents until Paul Ehlrich gave the postulate that cells have selective sets of 'chemoreceptors' and screened hundreds of phenyl arsenic acid derivatives in animal models to find treatment for Syphilis in rabbits in early nineteenth century (Jarrott, 2004; Tsinopoulos and McCarthy, 2002). Later on, Gerhard Domagk, in 1932, found that red dye, Prontosil Rubrum™, is effective against *Streptococcus* bacteria (Jarrott, 2004). These two discoveries gave the idea that chemical compounds can be therapeutically used to discover new drugs. But drug discovery is not an easy task. According to an estimate only one compound out of about 5000 novel compounds make its way to the clinic as a useful market drug, taking at least 10 years in progressive phases and costs about US$ 900 million (Jarrott, 2004). To deal with this problem time and progress in every discipline put forth computational methods to minimize the number of unsuccessful chemicals. These methods are collectively known as 'Computer Aided Drug Design (CADD)' in today's world.

With the progress in scientific fields, drug development process is no longer an unsystematic, hit and miss procedure. Instead, with more information and more avant-garde tools at hand, drug design and development has become a systematic process based on three strategic assets: knowledge, technology and available resources (Tsinopoulos and McCarthy, 2002). CADD incorporates the knowledge of receptor

ligand interaction in order to hypothesize the best refined ligand (Lead) that can bind a target structure (protein or DNA) to show some therapeutic benefit in comparatively much less time and cost then traditional drug discovery as shown in Figure 10 (Richards, 1994).

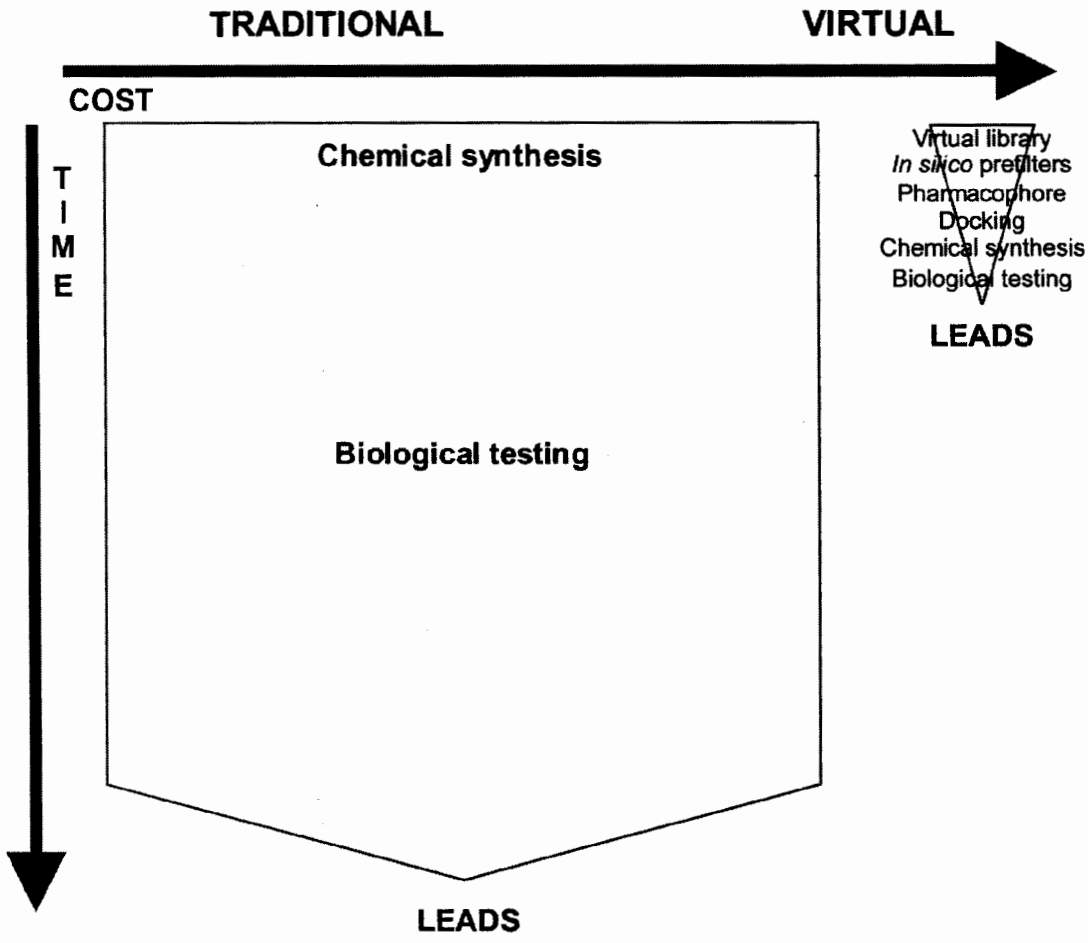


Figure 10: Comparison of traditional and virtual screening in terms of expected cost and time requirments.

First step in CADD is recognition of the target site which is preferably a ligand binding pocket on an aimed molecule, a protein or an enzyme, lined with a multiplicity of probable Hydrogen Bond Donors (HBD), Hydrogen Bond Acceptors (HBA) and entities having hydrophobic characteristics (Anderson, 2003). During target selection, three different situations can come up:

1. Receptor structure and its bioactive conformation are yet not established.

2. Only the bioactive conformation of the ligand is known.
3. Structure is known but bioactive conformation of the ligand is or is not well-known.

Computational methods are then used to identify hits (chemical compounds found active in biological assays), pick lead (most suitable candidate for further study) on the basis of some criteria and optimize lead (manipulate lead to convert it into active drug by refining its physiochemical, pharmaceutical and pharmacokinetic properties) (Kapetanovic, 2008). The lead is the preferred compound but is not perfect in all aspects. It may show toxicity, have administrative problems or solubility issues. To transform this compound into a registered drug, it is optimized and checked in clinical trials (Daharwal, 2006). Generally used line of action to reach from hits to lead compound are pharmacophore generation (ligand based drug design requiring a 3D spatial layout of chemical attributes that relate to biological activity), ligand- target docking (structure based drug design) and quantitative structure activity relationship (Kapetanovic, 2008).

2.2.1 Pharmacophore

The term pharmacophore was primarily established by Monty Kier in his research articles published in 1967-1971 while he was serving at University of Michigan (Drie, 2007). International Union of Pure and Applied Chemistry (IUPAC) define pharmacophore as:

"An ensemble of sterric and electronic features that is necessary to ensure the optimal supramolecular interactions with a specific biological target and to trigger (or block) its biological response" (C. R. Ganellin *et al.*, 1998).

Generally speaking Pharmacophore models are subdivided into two distinct classes that are receptor based pharmacophore and ligand based pharmacophore (A. Yamaguchi). Receptor based pharmacophore refers to designing a collection of features using the information imparted by active site based on the well-known lock and key concept. But ligand based pharmacophore models are used in case when 3D structure of the receptor protein is unknown (Y. Patel *et al.*, 2002). The features included to generate a pharmacophore model are HBD, HBA, hydrophobic, aromatic, positive ionizable and negative ionizable groups. Pharmacophore itself is just a blue print but does not represent an actual compound (Kapetanovic, 2008).

There are a number of software's used for 3D pharmacophore generation such as Catalyst, DISCO, GASP, GALAHAD, Phase, MOE and LigandScout (Kapetanovic, 2008; Y. Patel *et al.*, 2002; Nicola J. Richmond, 2006). Most of these work by analyzing the pharmacophoric features of active compounds and aligning multiple ligands that have active conformations to find the best overlay of the analyzed features (Y. Patel *et al.*, 2002).

Pharmacophore models for influenza neuraminidase have been generated using ligand based approach (Hammad and Taha, 2009; J. Zhang *et al.*, 2006). Both studies are carried out via Catalyst software using a number of compound classes. The best model thus generated had five features that is, a positive ionizable, a negative ionizable, a hydrophobic point and two hydrogen bond donors as shown in Figure 11.

2.2.2 Molecular Docking

Molecular Docking is a process in which favored orientation of two macromolecules usually a ligand and receptor protein or enzymes is predicted on the basis of shape complementarity and the interaction among the macromolecules by the

use of computational methods (A. Kaapro and J. Ojanen, 2002; A. May and M. Zacharias, 2005). Interactions between the macromolecules are described on the basis of attractive and repulsive forces between units that together make up the macromolecules (A. Kaapro and J. Ojanen, 2002). These forces include forces with electrostatic origin (dipole-dipole, charge-charge and charge-dipole), forces with electrodynamic origin (charge-induced dipole and Vander Waals), sterric forces (entropy) and solvent related forces (due to structural changes of solvent on addition of ions, colloids or proteins into the solvent) (A. Kaapro and J. Ojanen, 2002).

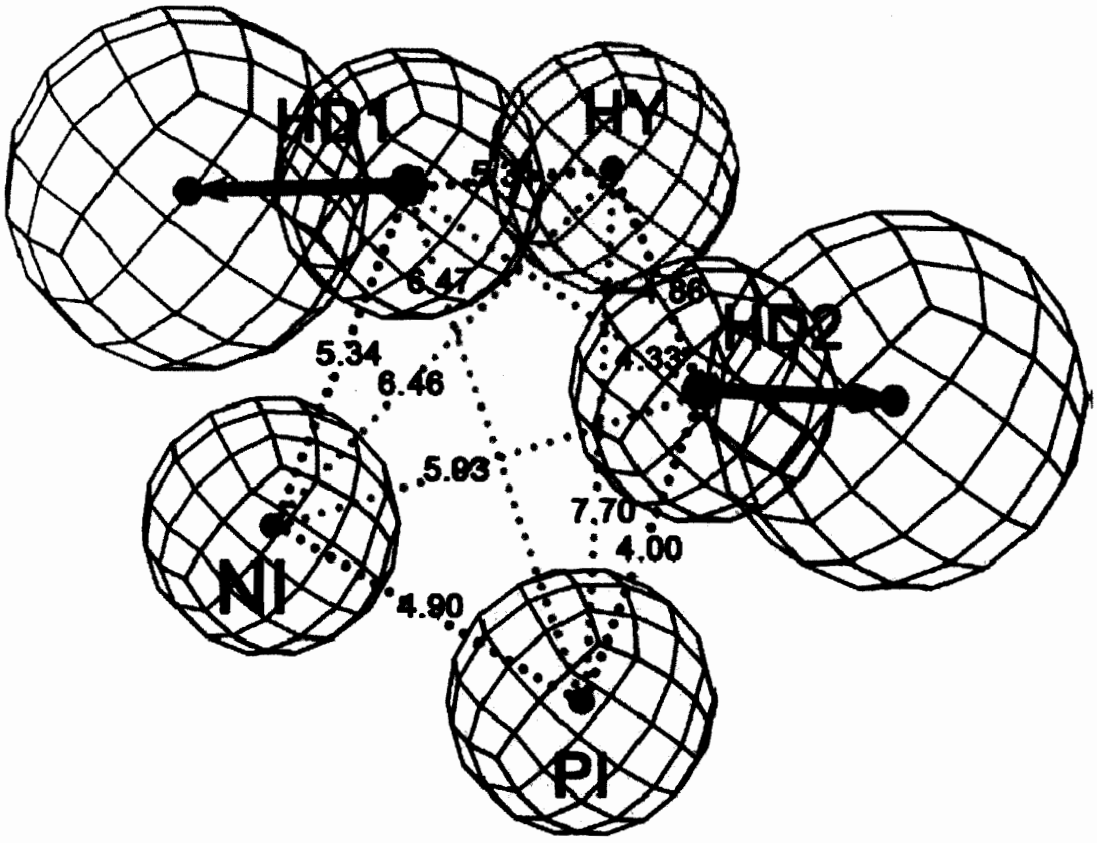


Figure 11: The best hypothesis model produced by the HypoGen module in Catalyst 4.10 software package. Pharmacophore features are color coded with light blue for Hydrophobic, red for Positive Ionizable Group, blue for Negative Ionizable Group and magenta for Hydrogen Bond Donors. Distance between pharmacophore features is reported in Armstrong. HY, hydrophobic group; HD1, hydrogen bond donor 1; HD2, hydrogen bond donor 2; PI, positive ionizable group and NI, negative ionizable group.

The docking process consists of two steps: firstly, a number of configurations of ligand with the target molecule, including the binding modes, are generated using a search algorithm. Secondly, the resulting configurations are assessed by a scoring function to come up with highly stable experimental binding configurations (A. Kaapro and J. Ojanen, 2002; Kapetanovic, 2008). Some common search algorithms are Molecular dynamics, Monte Carlo methods, Genetic algorithms, Fragment-based methods, Point complementary methods, Distance geometry methods, Tabu searches and Systematic searches (A. Kaapro and J. Ojanen, 2002).

Two approaches are currently being used to evaluate a ligand protein conformation. First approach uses the whole scoring function to rank a ligand protein conformation and the same function is then used to rank a new compound bound to the same protein. In the second method, an abridged function is first implemented to come up with the potential structures for docking studies and then a precise and laborious function is applied to rank the follow on structures. Generally used scoring methods are Force-field methods, Empirical free energy scoring functions and Knowledge-based potential of mean force (A. Kaapro and J. Ojanen, 2002).

For a program to work effectively it needs not only to generate the right conformation but also be able to recognize it by the use of appropriate scoring method (Y. Fukunishi, H. Nakamura, 2008). There have been a number of software packages now available for running the molecular docking job but neither one has gained the status of being a standard (Kapetanovic, 2008). Commonly used docking softwares are AutoDock, DOCK, FlexX and Gold (A. Kaapro and J. Ojanen, 2002). Recently, an update of AutoDock called AutoDock Vina is introduced which is claimed to be more efficient in calculation time and accuracy (Trott and Oslon, 2009).

2.2.3 QSAR (Quantitative Structure Activity Relationship)

Quantitative structure activity relationship, in simplest terms, is a statistical method of finding correlation between structure and function of a set of compounds by computationally intense methods. Structure here points to the properties or descriptors of the molecules, their substituent or interaction energy fields, function corresponds to an experimental biological/biochemical endpoint like binding affinity, activity, toxicity or rate constants (Verma *et al.*, 2010). Normally the term, QSAR, is used when some computational or mathematical method, such as linear regression analysis, multivariant analysis and pattern recognition, is applied to model biological activity (Winkler, 2002). When the same method is applied to model toxicological data, it is called Quantitative Structure Toxicity Relationship (QSTR) and to model physicochemical properties it is called Quantitative Structure Property Relationship (QSPR) (Winkler, 2002; Kapetanovic, 2008). QSAR approaches are frequently used in predictive toxicology (Kapetanovic, 2008). Recently, QSAR studies have become an essential part of drug design and development and play a vital role in analysis by presenting cheaper and rapid alternatives that were usually restricted to later stages in discovery cascade (C. H. Andrade *et al.*, 2010).

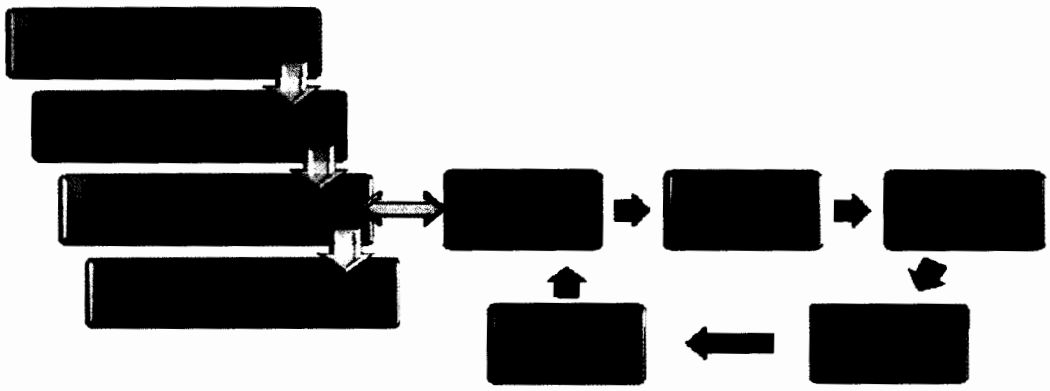


Figure 12: Schematic representation of the processes included in a lead optimization from the hit identification. QSAR methods are essential to reach this goal (C. H. Andrade *et al.*, 2010).

QSAR methodolgy is based on recognition that a molecule is a collection of three dimensional distribution of properties or descriptors which are sterric, electronic and lipophillic properties (Winkler, 2002). QSAR methods focus generally on quantitative correlation and recapitulating trends between chemical structure modifications and endpoint biological activities highlighting the key determinants of biological activity. Moreover, these methods are also availed for lead optimization and prediction of biological activities of untested compounds (Verma *et al.*, 2010).

Out of the classes based on dimensionality, 3D QSAR has gained the most popularity since it advent by Hansch and Free Wilson in 1960's (Verma *et al.*, 2010). The eminent approaches are Comparative Molecular Field Analysis (CoMFA), Molecular Shape Analysis (MSA), Hypothetical Active Site Lattice (HASL), Comparative Molecular Similarity Indices Analysis (CoMSIA), Genetically Evolved Receptor Models (GERM), Comparative Binding Energy Analysis (COMBINE), Comparative Molecular Moment Analysis (CoMMA), Comparative Molecular Surface Analysis (CoMSA), Adaptation of Fields for Molecular Comparison (AFMoC) and Comparative Residue Interaction Analysis (CoRIA) (Verma *et al.*, 2010; Wrinkler, 2002; Kubinyi, 1997).

Several QSAR studies on influenza neuraminidase inhibitors have been conducted in recent years (Hammad and Taha, 2009; W.J. Lu *et al.*, 2008; R. P. Verma and C. Hansch, 2006). QSAR models for neuraminidase inhibitors have suggested that hydrophobicity and hydrogen bond interactions play a key role in NA activity (W.J. Lu *et al.*, 2008). Additionally, molar volume is also indicated as a prime factor for NA activity (R. P. Verma and C. Hansch, 2006). HOMO and LUMO has also indicated a certain role in ligands' electronic characteristics for neuraminidase binding (Hammad and Taha, 2009).

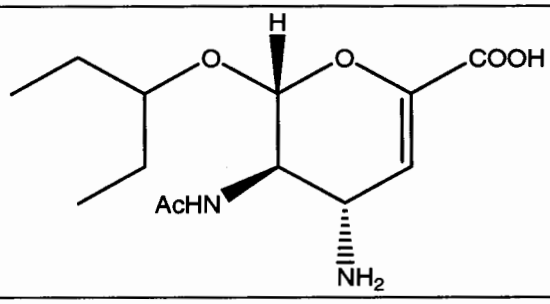
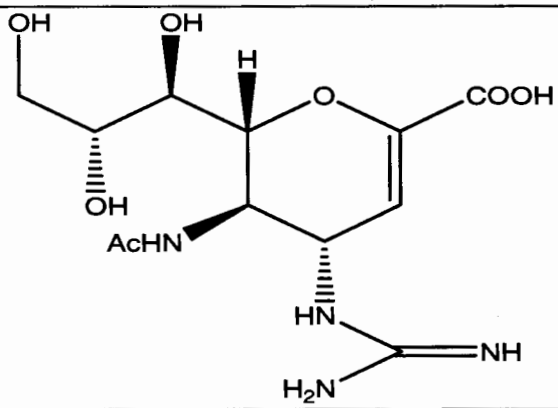
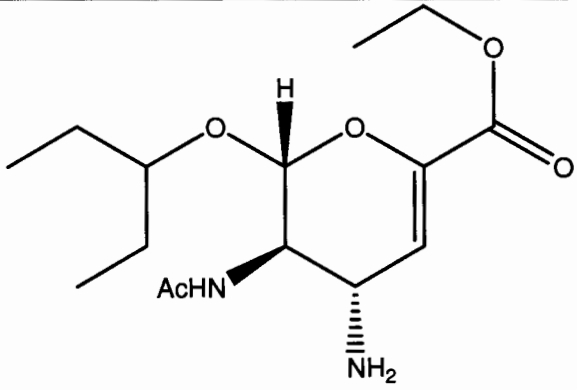
CHAPTER 3

METHODOLOGY

3.1 Anti Influenza Virus Drugs

Sickness and death are parts of life, but you can optimize your life. You can make progress as you strive toward perfection. This progress in case of fighting influenza epidemics, that keep arising from time to time, has led to the approval of two drugs by FDA (Food and Drug Administration) for neuraminidase protein target. These drugs are Zanamivir (Market Name: Relenza) and Oseltamivir (Market name: Tamiflu) and are shown in table 2. Another compound Oseltamivir Carboxylate, GS4071, is also considered equivalent to standard drugs because it is a prodrug of Oseltamivir. These market drugs are considered as paradigm compounds for this study. A number of other compounds known to have shown anti influenza activity in different experimental studies belonging to different classess are also considered. Few compounds found to be inactive against influenza virus are also part of the final dataset. The final dataset comprise of 41 compounds (T. Honda *et al.*, 2002; P. C. Nair & M. E. Sobhia, 2008; J. B. Finley *et al.*, 1999; H. J. Jeong *et al.*, 2009; Y. B. Ryu, 2010; A. G. Mercader & A. B. Pomilio, 2010; Ai-Lin Liu *et al.*, 2008; W. Lew *et al.*, 2000; M. A. Williams *et al.*, 1997) along with FDA approved standards and other biological active and inactive compounds. The concluding data set is made known in Table 3 at the end of this chapter.

Table 2: Standard Drugs and Prodrug for Treating Influenza.

Serial No.	Compound Name	Compound Structure	Inhibitory Concentration IC50(μM)
1.	Oseltamivir Carboxylate GS4071*		0.001
2.	Zanamivir		0.0007
3.	Oseltamivir		0.015

* Prodrug for Oseltamivir

3.2 Compound Portrayal

The chemical structures of selected compounds along with the FDA approved standard compounds are made using the molecule editor ChemDraw Ultra 8.0 (Mendelsohn, 2004). It is the ultimate structure drawing tool used by professional scientists, science students and scientific authors to communicate chemical structures in their research and published work. The structures are saved with .cdx extension

which is the native ChemDraw file format. The chemical structures made thereby are also saved with .mol extension, MDL MolFile file format.

3.3 Generation of 3D Pharmacophore

Pharmacophore is the minimum set of essential structural features such as HBD, HBA, hydrophobic, aromatic, positive ionizable and negative ionizable groups that are required for the biological activity of a compound at a specific receptor location. In this study, LigandScout 2.03 and 3.0 versions (Langer, 2005) are used to determine the 3-D pharmacophore of the dataset. LigandScout 3.0 has the option of ligand based as well as structure based pharmacophore modeling and a separate module for aligning a no of selected compounds; however for this study structure based modeling option has been used.

The mol files created using ChemDraw are used as an input to the software for pharmacophore generation via LigandScout 2.03 and pdb files created using Chem3D Ultra are used for the later version of LigandScout. After importing the file to the working environment of the software, create pharmacophore command is used to generate set of all pharmacophore features of the compound. The process of pharmacophore creation is repeated for all the compounds in the dataset as well as for the standard drugs. After studying the pharmacophore features of all the compounds, a set of common features is selected and further studied for distance approximation between the predicted pharmacophore features. Distance between the selected pharmacophore features is calculated for each candidate compound so that a correlation between feature distance and activity can be found. These common features are taken by the superimposition of ligands in LigandScout. Conclusively a

pharmacophore model has been predicted for anti influenza Neuraminidase Inhibitors on the basis of the chosen dataset.

3.4 Molecular Docking

Docking is performed to foretell the preferred orientation of a ligand with a macromolecule such as a protein in bound form to make a stable complex. Hence two essential elements required for a docking study are a protein and a ligand. An apt target protein was selected for anti influenza neuraminidase inhibitors (Xu *et al.*, 2008). An IAV neuraminidase (pdb id: 3B7E) with an additional cavity for ligand binding near its natural substrate binding cavity was considered for docking studies. Pdb file of 3B7E was acquired from protein data bank (rcsb.org).

3.4.1 AutoDock Vina

For running a docking job on AutoDock Vina (Olson, 2010) we need pdbqt files for both protein and ligand. These files were prepared using AutoDock 4.0 (Ewing *et al.*, 2001) by passing through following steps.

3.4.2 Preparing the Protein

The pdb file of the selected protein was imported into the software environment by read molecule command. Once the protein was visible, polar hydrogens were added to the protein through edit menu. To increase the readability of atoms the protein was colored according to atoms so that carbons were visible in cyan color, nitrogen elements were in blue color, oxygen elements were in red etc. Then the macromolecule was saved as a pdbqt file after selecting the molecule via grid menu. Finally, we have to choose part of the receptor where we want our ligand to bind while the docking job is executed. This part was chosen using grid box menu. The values of the grid box were selected according to the active site of the macromolecule.

The grid box position was adjusted using centre grid box option, which gives the freedom to centre the grid box through adjusting x, y and z coordinates. The size of the grid box was also attuned by adjusting the no. of points for the grid box in all three dimensions. The pdbqt file prepared so was used for all the ligands in the docking studies.

3.4.3 Preparing the Ligand

To make the ligand ready for the final execution, the ligand pdb was brought into the working environment of AutoDock 4.0. AutoDock automatically calculates the Gasteiger charges, finds aromatic carbons, detects rotatable bonds and assigns an AutoDock type to each atom of the ligand. The rotatable and non rotatable bond count was attuned using the choose torsion option in torsion tree menu of the ligand. Once the torsions were set the file was saved as pdbqt file of the ligand. The course of action was repeated for all the compounds in the dataset and their pdbqt files were prepared.

3.4.4 Operating the Docking Task

The pdbqt files of the ligand and the macromolecule were placed in a new folder. A new text document was made by the name "conf.txt". Name of the pdbqt files of ligand and the receptor and the dimensions of the grid box were recorded into this txt file in a specific format. The prepared file was also placed in the same folder where pdbqt files were previously stored.

The docking job was executed by AutoDock Vina by writing the command given below in command prompt after opening the folder where all the input files are located. The execution command was:

"\program files\The Scripps research institute\vina\vina.exe" --config conf.txt --log log.txt

After the job was executed, output files were written to the same folder. A log file was also part of the output which had list of all the possible orientations of the ligand with the protein along with their binding affinity. The docking conformation having the lowest energy was selected for the further analysis. Procedure was repeated for all the candidate compounds and the standard compounds.

3.5 Ligand Protein Interactions

The ligand protein interactions were predicted using the molecular graphics software VMD (Visual Molecular Dynamics), which is a tool for analyzing and visualizing macromolecules (Humphrey, 1996). The lowest energy conformation of each ligand obtained during docking studies and the target macromolecule was launched in VMD. Protein was analyzed with in 5 Å of the docked position of ligand to find the probable interaction. Possible ionic, hydrogen, hydrophobic and vanderwaal interactions were recorded for each ligand individually.

3.6 Identifying the Lead

Binding interactions of each compound were studied meticulously and the most active neuraminidase inhibitor with highest possible interactions in a close distance range with the macromolecule was chosen as Lead.

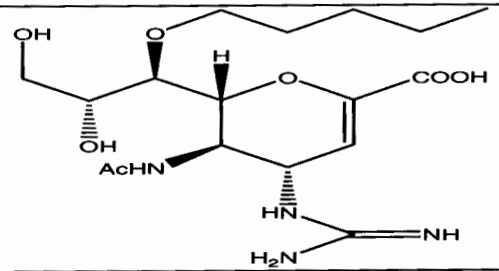
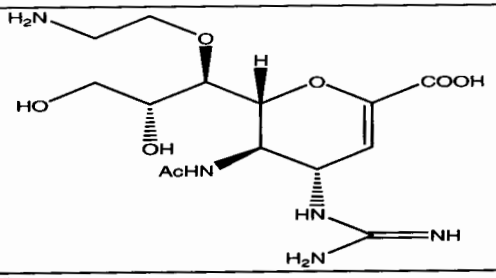
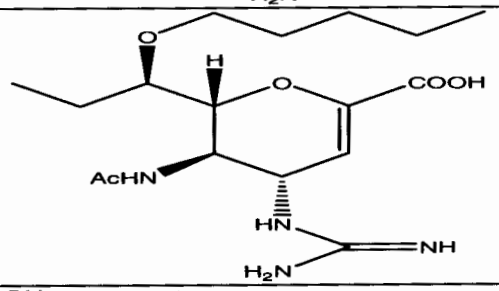
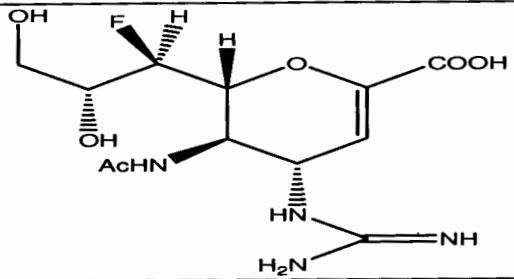
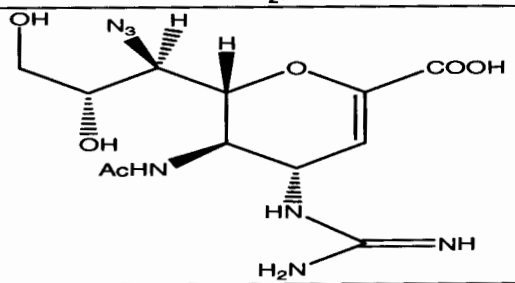
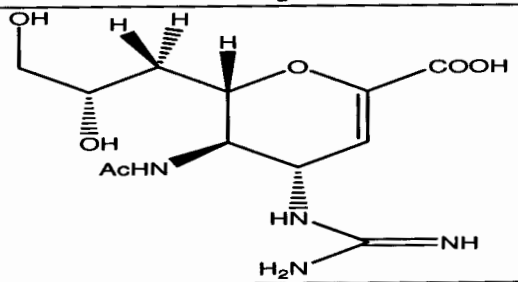
3.7 Analogue Scheming

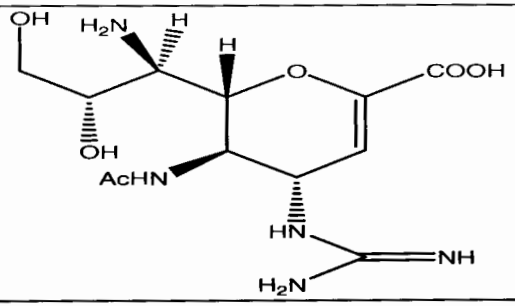
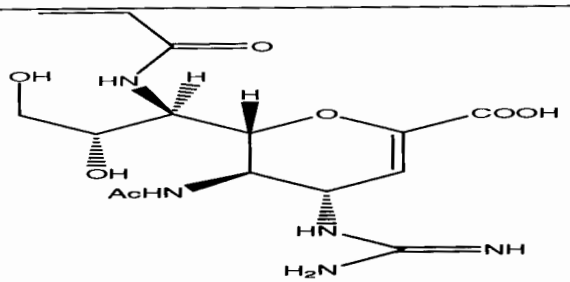
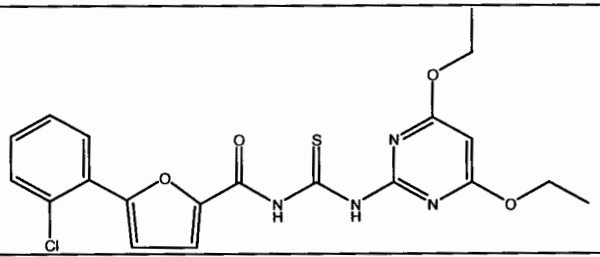
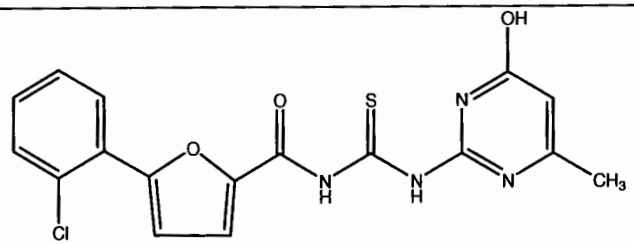
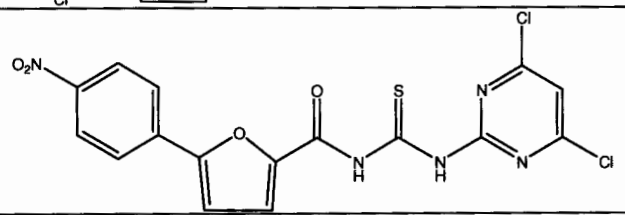
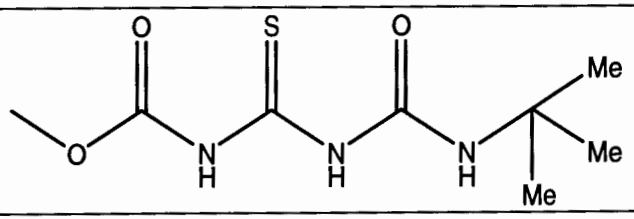
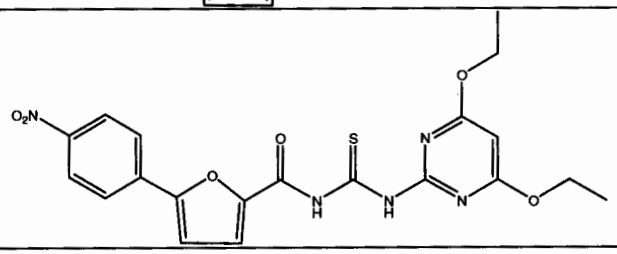
Using the Lead compound as a template, four structural analogues were formulated. These analogues were devised by adding, deleting or replacing a functional group keeping in view the observed relationship between functional characteristic and activity of compounds. Process of docking was repeated for the analogues and their low energy conformations were studied for possible binding interactions.

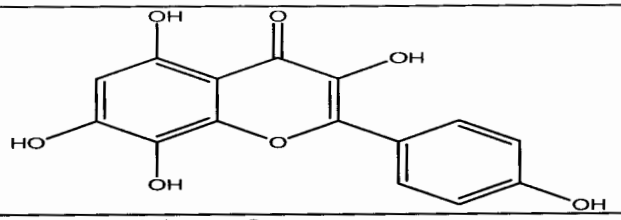
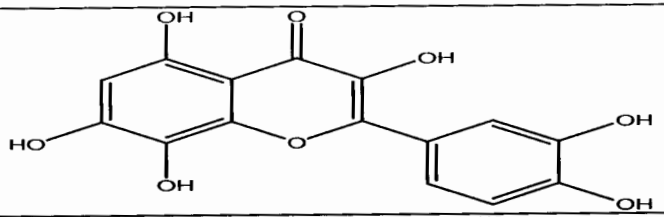
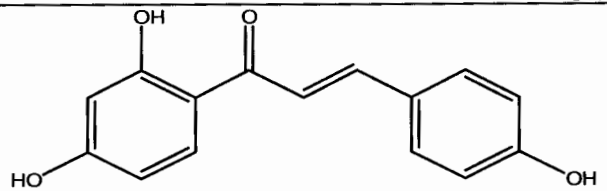
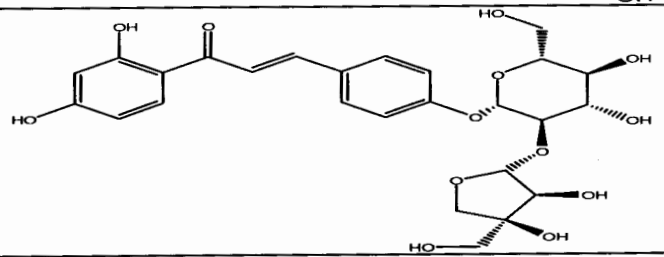
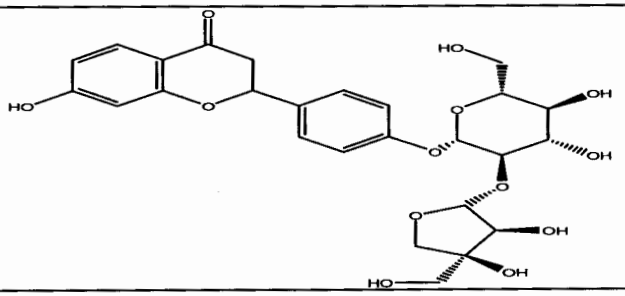
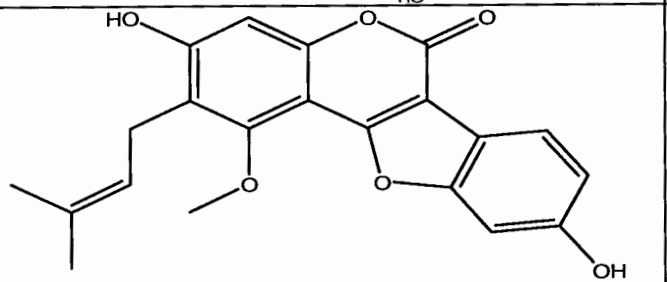
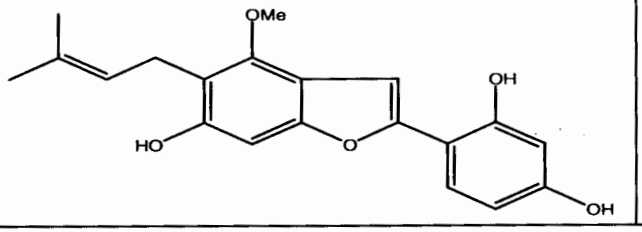
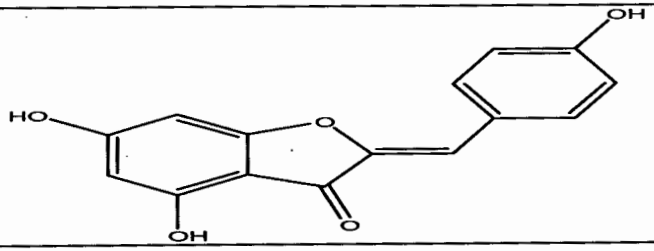
3.8 Quantative Structure Activity Relationships (QSAR)

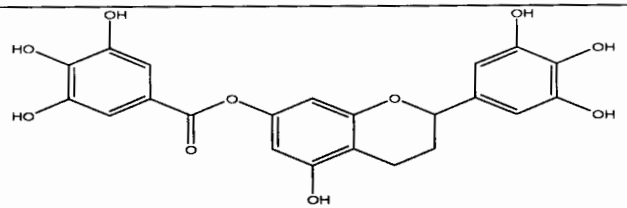
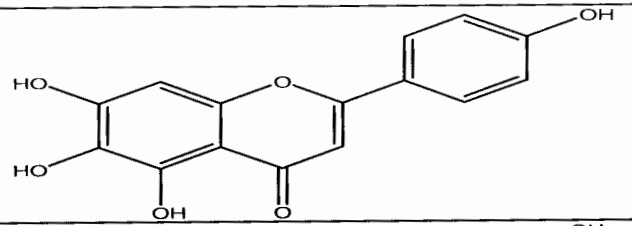
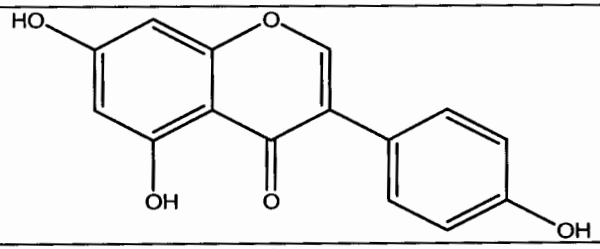
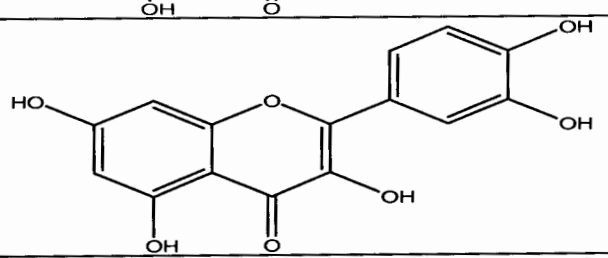
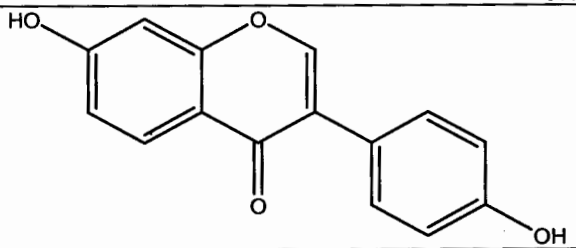
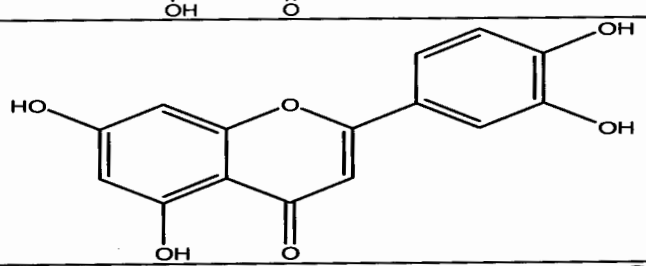
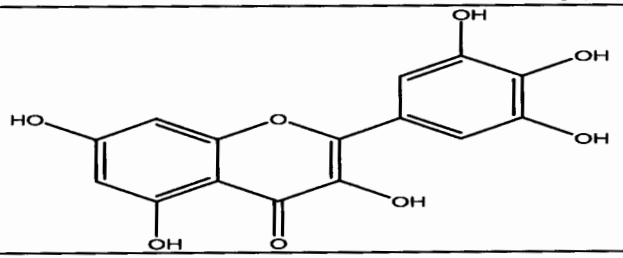
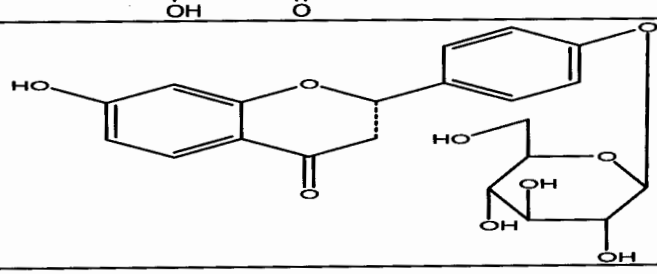
For performing QSAR a different set of compounds, all belonging to the same group i.e. polyphenols, were selected including some compounds from the original data set (H. J. Jeong *et al.*, 2009; Y. B. Ryu, 2010). To correlate property descriptors of compounds with their biological activity QSAR was performed. A number of thermodynamic and electronic descriptors were calculated using ChemDraw and HyperChem. Single point calculations were done using HyperChem that determine the molecular energies. The values thus calculated by above stated tools were plotted against biological activity of these compounds using Microsoft Excel 2007.

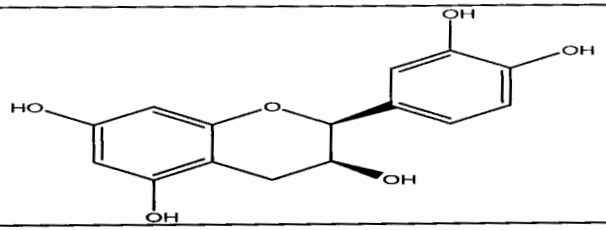
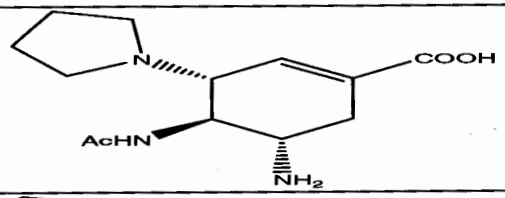
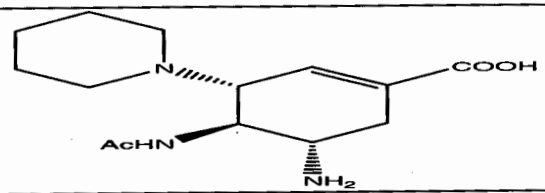
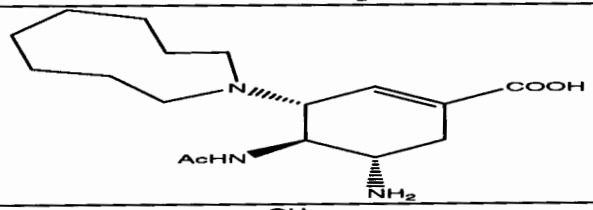
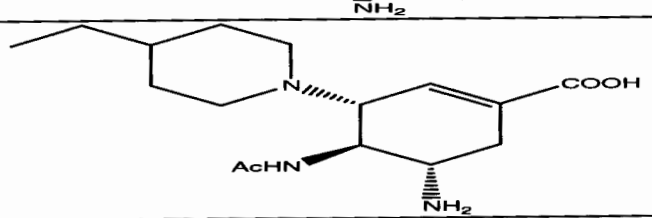
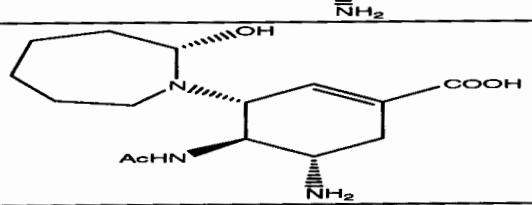
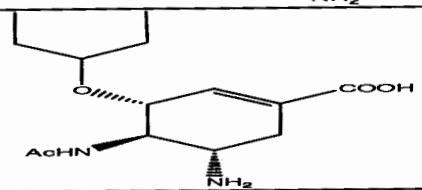
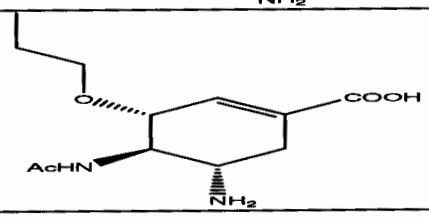
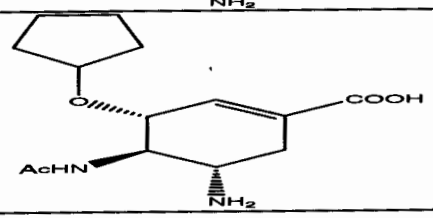
Table 3: Selected Dataset for Study of Influenza Treatment.

C. No	Compound Structure	IC 50	C. No	Compound Structure	IC50
ZANAMIVIR DERIVATIVES					
4.		0.035	5.		0.127
6.		0.024	7.		0.002
8.		0.0044	9.		0.0316

10.		0.112	11.		0.144
THIOUREA ANALOGUES					
12.		0.008	13.		0.32
14.		1.66	15.		1.83
16.		>20*			

POLYPHENOLS					
17.		8.9	18.		2.6
19.		9.0	20.		12.9
21.		18.2	22.		3.1
23.		27.8	24.		25.6

25.		35.49	26.		48.56
27.		77.1	28.		58.39
29.		37.1	30.		33.7
31.		82.6	32.		>100*

33.		>100*			
CARBOCYCLIC COMPOUNDS					
34.		0.310	35.		0.025
36.		0.075	37.		0.052
38.		0.008	39.		0.001
40.		0.130	41.		0.022

*inactive compounds in experimental studies.

one NI (red 3D asterisk) and one PI (blue 3D asterisk). This Pharmacophore model of standard drug is deficient in hydrophobic feature predicted by J. Zhang *et al.*, 2006 and has additional HBAs in the model. This difference points to the need of more general pharmacophore models to be predicted for anti influenza compounds.

Figure 13: Pharmacophore Model of Zanamivir Predicted By LigandScout 3.0.

Pharmacophore features of each candidate compound are shown in table 4. Zanamivir derivatives chiefly consist of HBAs, HBDs, PI and NI residues. Thiourea analogues are deficient in NI residues but rich in all other pharmacophore traits. Polyphenols are predominantly lacking ionizable residues, whereas, Carbocyclic compounds, like Zanamivir derivatives, were rich in HBAs, HBDs, PI and NI residues but lack hydrophobic residues and aromatic rings.

Each class of compounds has shown significant difference in features from other classes chosen for the study. Moreover, there is noticeable deviation in pharmacophore models when compared to the pharmacophore model of standard drugs. General comparison of table 2, 3 and 4 leads us to draw noteworthy conclusions, such as; Zanamivir derivatives and carbocyclic compounds have more significant biological activities then the other two classes due to the presence of ionizable residues. Hydrophobic residues though part of standard drugs cannot be

connected to biological activity in this dataset, as in some compounds the presence of hydrophobic residue did not affected biological activity considerably.

Table 4: Pharmacophore Features of Compounds considered for treating Influenza.

Compound No.	Hydrophobic Residues	Hydrogen Bond Acceptor	Hydrogen Bond Donor	Negative Ionizable Residues	Positive Ionizable Residues	Aromatic Rings
1.	2	5	2	1	1	-
2.	-	7	7	1	1	-
3.	3	4	2	-	1	-
4.	1	7	6	1	1	-
5.	-	7	7	1	2	-
6.	2	3	4	-	2	-
7.	1	7	6	1	1	-
8.	-	6	9	1	3	-
9.	-	6	6	1	1	-
10.	-	6	7	1	2	-
11.	-	7	7	1	1	-
12.	5	4	4	-	2	2
13.	3	3	5	-	2	2
14.	1	4	4	-	2	2
15.	1	3	3	-	-	-
16.	3	6	4	-	2	2
17.	1	7	6	-	-	2
18.	1	8	7	-	-	2
19.	1	4	1	-	-	1
20.	1	11	8	-	-	2
21.	1	11	8	-	-	2
22.	2	4	3	-	-	3
23.	2	5	3	-	-	1
24.	1	5	1	-	-	1
25.	-	10	4	-	-	2
26.	1	6	4	-	-	2
27.	1	5	1	-	-	1
28.	1	7	6	-	-	2
29.	1	4	1	-	-	1
30.	1	6	2	-	-	1
31.	-	8	3	-	-	1
32.	1	8	6	-	-	1
33.	1	6	5	-	-	2
34.	-	3	3	1	2	-
35.	-	3	3	1	2	-
36.	-	3	3	1	2	-
37.	1	3	3	1	2	-
38.	-	4	4	1	2	-
39.	2	4	2	1	1	-
40.	1	4	2	1	1	-
41.	-	4	2	1	1	-

To come up with a pharmacophore common to all these classes and responsible for showing activity against influenza target, NA, a shared pharmacophore is generated using one compound from each candidate class and the standard drug, Zanamivir. The representative 2D and 3D pharmacophore of each group is given in figures 14, 15, 16 and 17, showing the pharmacophore features as predicted by LigandScout 3.0.

Figure 14: Pharmacophore Model of Compound No. 4 (Zanamivir Derivative).

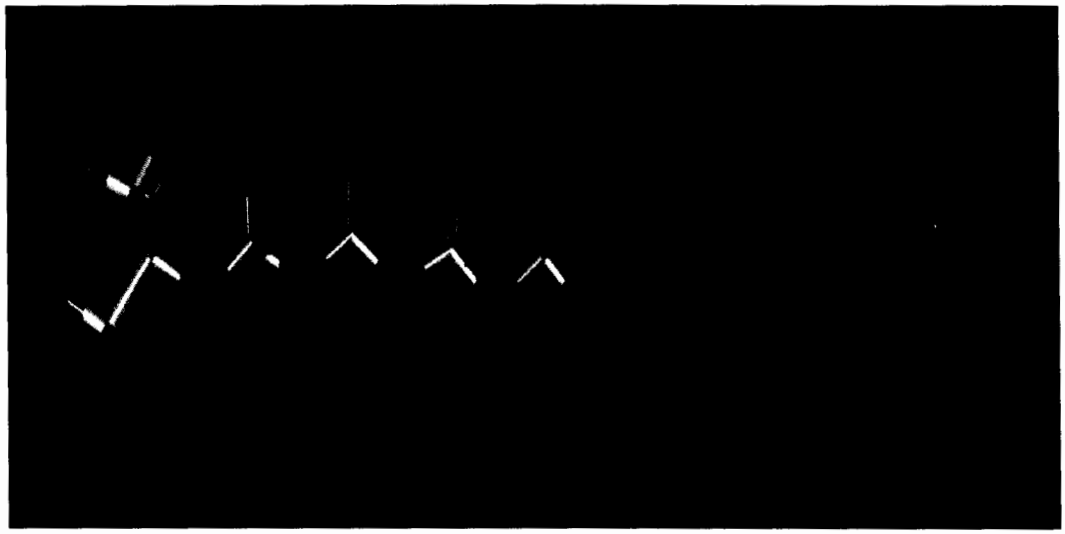


Figure 15: Pharmacophore Model of Compound No. 15 (Thiourea Analogue).

Figure 16: Pharmacophore Model of Compound No. 23 (PolyPhenol).

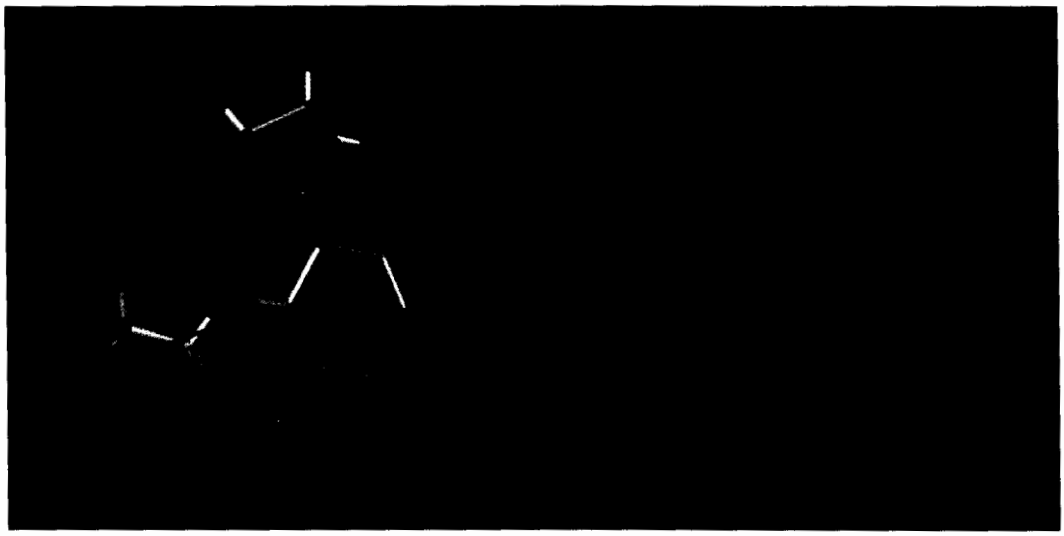


Figure 17: Pharmacophore Model of Compound No. 34 (Carbocyclic Compound).

The standard compound Zanamivir has been set as the reference point to generate the shared pharmacophore model of the selected compounds of each class. Predicted shared pharmacophore as shown by Figure 18 has seven HBAs, four HBDs, one PI and a NI. It is evident from table 4 that PI and NI residues are not present in all Thiourea analogues and polyphenolic compounds. Considering the absence of these features in two classes and consistency of HBAs and HBDs in all candidate classes,

only HBAs and HBDs were included in distance calculation for predicted pharmacophore model.

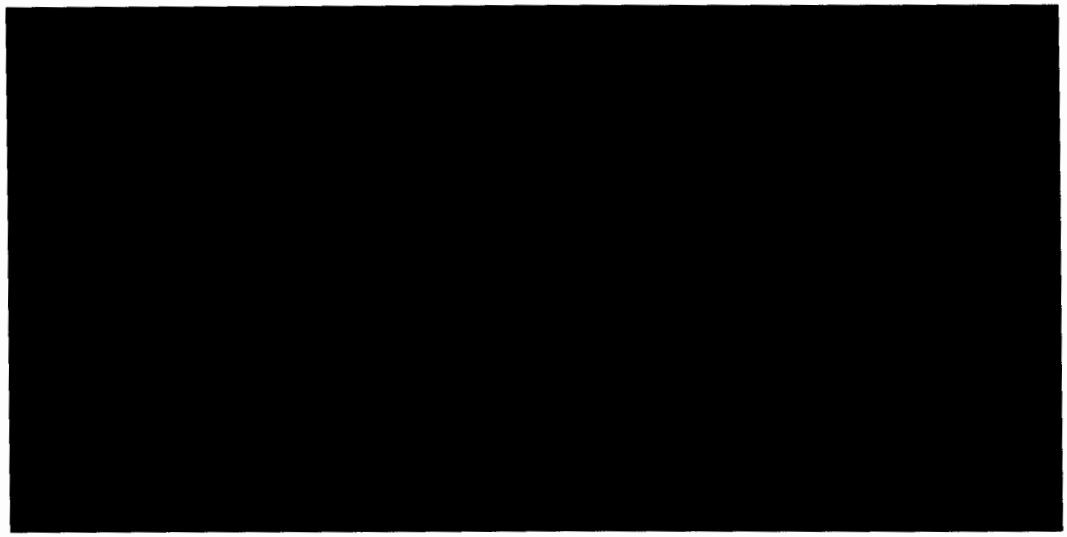


Figure 18: Shared Pharmacophore Model with Zanamivir Set as Reference Point.

For the sake of simplicity, two HBAs and one HBD are considered here for distance calculation. The distances measured among the predicted pharmacophore features of each class of compounds and the standard drugs using VMD are made known in table 5. The distance ranges from minimum to maximum were measured for each class of compounds among HBA_1 to HBA_2, HBA_1 to HBD and HBA_2 to HBD.

Table 5: Pharmacophore Triangle Distance for Each Class of Compounds.

Ligands	HBA_1-HBA_2 (Å)	HBA_1-HBD (Å)	HBA_2-HBD (Å)
Zanamivir Derivatives	1.55-1.97	1.43-2.47	1.97-2.18
Thio Urea Analogues	1.43-1.55	1.43	1.43-2.14
Polyphenols	1.43-2.86	1.43-2.86	2.18-3.78
Carbocyclic Compounds	1.43-2.18	1.43-2.98	2.97-3.60
Standard Drugs	1.43-1.65	1.65-2.47	2.18

The distances among the common pharmacophore features between the predicted pharmacophore points are as shown in figure 7. The distances between HBA_1 and HBA_2, HBA_1 and HBD and HBA_2 to HBD range from 1.43 to 2.86Å, 1.43 to 2.98Å and 1.43 to 3.78Å respectively.

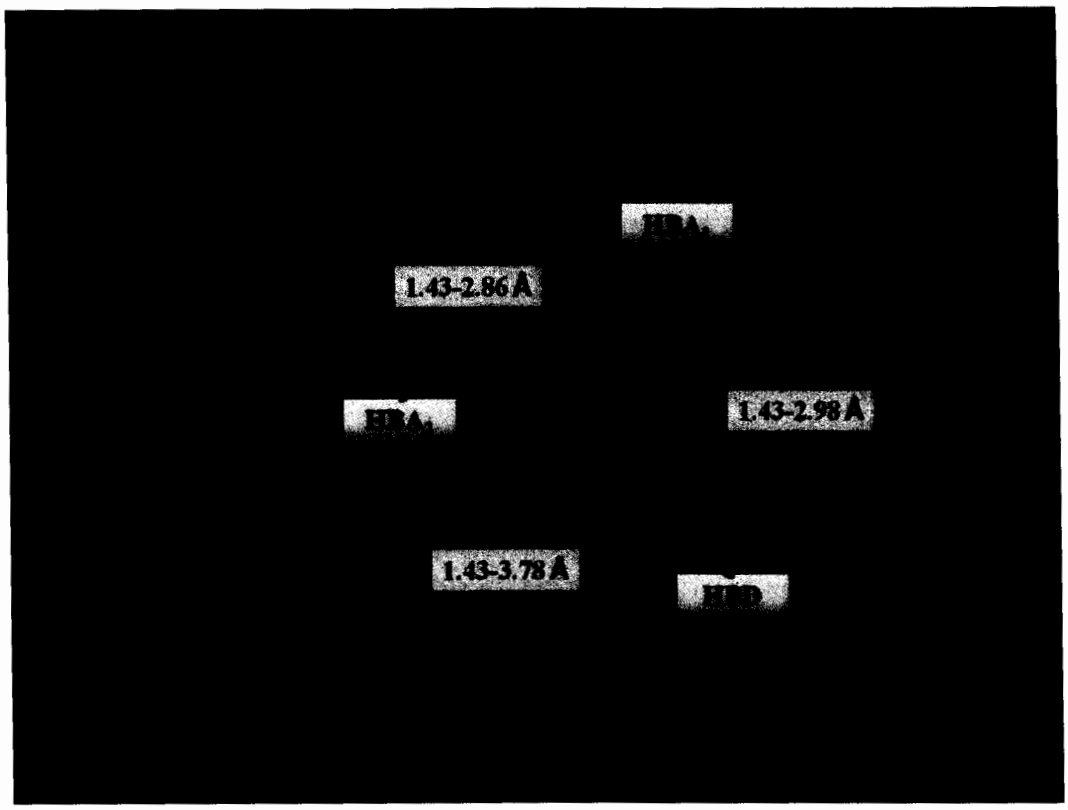


Figure 19: Distance Ranges among Pharmacophore features in Predicted Pharmacophore Model.

4.2 Docking Studies

Molecular docking is carried out in order to predict the best orientation of a molecule when bound to another molecule. AutoDock 4.0 and AutoDock Vina are used to conduct the docking job for influenza virus neuraminidase with the intention of coming up with a ligand that shows the best orientation and binding affinity. The binding models of all candidate classes i.e. Zanamivir Derivatives, Thio Urea Analogues, Polyphenols and Carbocyclic Compounds are obtained here using the most efficient program till date, AutoDock vina. The structure of target protein, Neuraminidase (pdb id: 3B7E), was acquired from protein data bank which is a reservoir of 3-D structures of biological molecules. The 3-D structure of target protein as viewed by VMD, colored according to Structure and NewCartoon drawing method,

is shown in Fig 20 below along with the active site of Zanamivir highlighted in both chains using Surf drawing style and colored according to name of the residues.

Figure 20: Neuraminidase Protein represented in NewCartoon style and the active site of Zanamivir in both chains is represented as Surf drawing method.

Docking calculations were performed using biologically active compounds listed in Table 3 along with Standard compounds and inactive compounds used as controls for the docking studies. The strategic results in a molecular docking job are the docked structures, the binding energies of these docked structures and the similarities of results with each other for a specific target protein. AutoDock Vina yields tens to hundred conformations for one structure while docking, depending upon the complexity and rotatable bonds allowed. The results are ordered according to their binding affinities, starting at the minimum in the output folder. Preferred conformations are with least energies; hence minimum energy conformations are selected and are further studied for the amino acids in drug vicinity and possible interactions.

4.2.1 Active Site of Neuraminidase

Active site is a cavity or a cleft in the structure of a protein that allows ligands of specific shape to fit in and interact with the target protein. It is termed active because usually this interaction results in a chemical change or a reaction. Active site of NA has been described in an earlier chapter. A recent research article has indicated the presence of a small loop near the active site of NA and suggested that its ligand binding potential be studies (Xu *et al.*, 2008). Docking is performed here with intention to manipulate the discretionary prospective.

On the basis of docking results, the amino acids within 5.0Å of ligand binding site are ILE 108, GLY 112, ASP 113, VAL 114, LEU 127, GLU 128, GLU 165, ALA 166, SER 168, TYR 169, ASN 170, ARG 172, ASN 208 and GLY 209 as is evident from table 6. A profound observation of the interactions illustrate that ASP 113, GLU 165, SER 168, TYR 169 and ARG 172 are the major players of binding in NA. Binding site with labeled amino Acids is shown in Figure below.

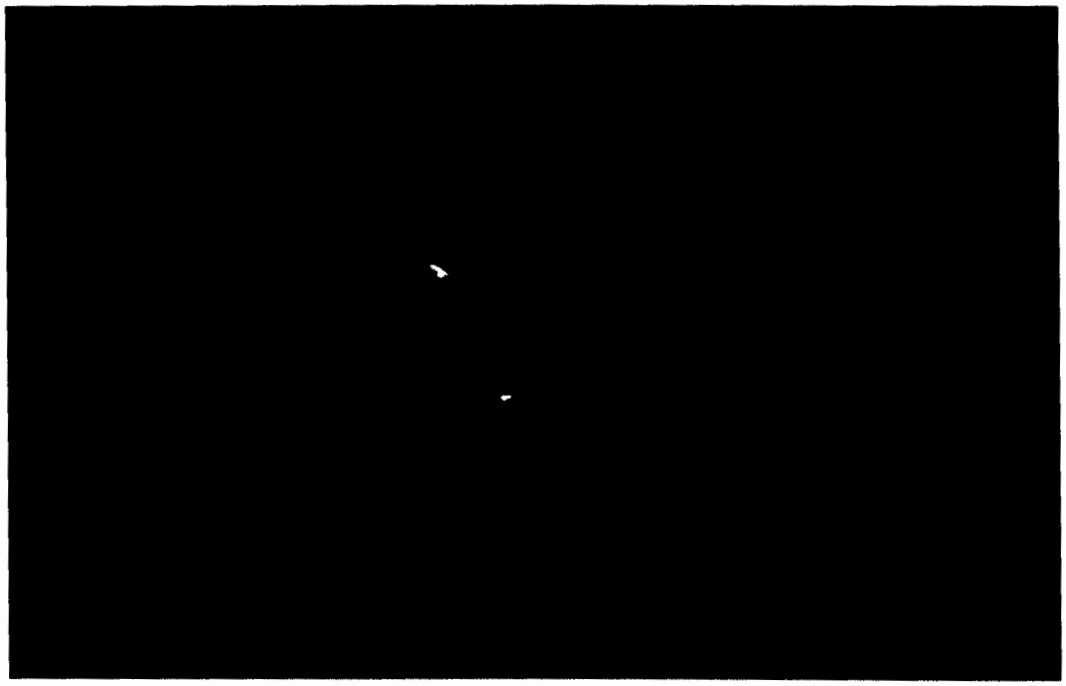


Figure 21: Binding Site of Neuraminidase Showing Amino Acids within 5.0Å.

Table 6: Amino Acids Present in the 5Å.

C. No	TYR: 169	ASN: 170	GLU: 165	ASP: 113	SER: 168	ILE: 108	GLY: 112	VAL: 114	ALA: 166	GLU: 174	GLU: 128	ARG: 172	GLY: 209	ASN: 208	LEU: 127	PHE: 173	LYS: 111	ILE: 210	PRO: 162
1	✓	✓	✓	✓	✓	✓	✓	✓	✓	×	×	×	×	×	×	×	×	×	×
2	×	×	×	×	×	×	×	×	×	×	✓	✓	×	✓	×	×	×	×	✓
3	✓	✓	✓	✓	✓	✓	✓	✓	✓	×	×	×	×	×	×	×	×	×	×
4	✓	✓	✓	✓	✓	✓	✓	✓	✓	×	×	×	×	×	×	×	×	×	×
5	×	✓	✓	×	×	×	×	×	×	×	✓	✓	×	×	×	×	×	×	✓
6	✓	✓	✓	✓	✓	✓	✓	✓	✓	×	×	×	×	×	×	×	✓	×	×
7	✓	×	×	✓	×	×	✓	×	×	×	×	×	×	×	×	×	✓	×	×
8	✓	×	×	✓	×	×	✓	×	×	×	×	×	×	×	×	×	✓	×	×
9	✓	×	×	✓	×	×	✓	✓	×	×	×	×	×	×	×	×	✓	×	×
10	×	×	✓	×	×	×	×	×	×	×	✓	✓	×	×	×	✓	×	×	✓
11	✓	×	×	✓	×	×	✓	×	×	×	×	×	×	×	×	×	✓	×	×
12	✓	✓	✓	✓	✓	✓	✓	✓	✓	×	×	×	×	×	×	×	×	×	×
13	✓	×	×	✓	✓	✓	✓	✓	✓	×	×	×	×	×	×	×	×	×	×
14	✓	×	✓	✓	✓	✓	✓	✓	✓	×	×	×	×	×	×	×	×	×	×
15	×	×	×	×	×	×	×	×	×	✓	✓	✓	✓	✓	✓	✓	×	✓	×
16	✓	✓	×	✓	✓	✓	✓	✓	✓	×	×	×	×	×	×	×	×	×	×
17	×	×	✓	×	×	×	×	×	×	×	✓	✓	✓	✓	✓	✓	×	×	✓
18	×	×	×	×	×	×	×	×	×	✓	×	✓	✓	✓	✓	✓	×	✓	×
19	×	×	×	×	×	×	×	×	×	×	✓	✓	×	✓	×	×	×	×	✓
20	✓	✓	×	×	×	×	×	×	×	✓	✓	✓	✓	✓	✓	✓	×	✓	×
21	×	×	×	×	×	×	×	×	×	✓	✓	✓	✓	✓	✓	✓	×	✓	✓
22	✓	✓	×	✓	✓	✓	✓	×	×	×	×	✓	×	×	×	×	×	×	×
23	×	×	×	×	×	×	×	×	×	×	✓	✓	×	✓	×	×	×	×	✓
24	×	×	×	×	×	×	×	×	×	✓	×	✓	✓	✓	×	✓	×	✓	×
25	✓	×	×	×	×	×	×	×	×	×	✓	✓	✓	✓	✓	✓	×	✓	✓
26	×	×	×	×	×	×	×	×	×	×	✓	✓	×	✓	×	×	×	×	✓
27	×	×	×	✓	×	×	✓	×	×	×	×	×	×	×	×	×	✓	×	×
28	×	×	×	×	×	×	×	×	×	✓	×	✓	✓	✓	✓	✓	×	✓	×

29	×	×	×	×	×	×	×	×	×	×	✓	✓	×	×	✓	×	×	×	✓
30	×	×	×	×	×	×	×	×	×	✓	✓	✓	✓	✓	✓	✓	×	×	✓
31	×	×	×	×	×	×	×	×	×	×	✓	✓	×	✓	×	×	×	×	✓
32	×	×	✓	×	×	×	×	×	×	✓	×	✓	✓	✓	✓	✓	×	×	×
33	✓	✓	×	✓	✓	✓	✓	✓	✓	×	×	×	×	×	×	×	×	×	×
34	×	×	✓	×	×	×	×	×	×	×	✓	✓	✓	✓	✓	✓	×	×	✓
35	✓	✓	×	✓	✓	✓	✓	✓	✓	×	×	×	×	×	×	×	×	×	×
36	✓	✓	✓	✓	✓	✓	✓	✓	✓	×	×	×	×	×	×	×	×	×	×
37	✓	✓	×	✓	✓	✓	✓	✓	✓	×	×	×	×	×	×	×	×	×	×
38	✓	✓	×	✓	✓	✓	✓	✓	✓	×	×	×	×	×	×	×	×	×	×
39	✓	✓	×	✓	✓	✓	✓	✓	✓	×	×	×	×	×	×	×	×	×	×
40	×	×	✓	×	×	×	×	×	×	✓	✓	✓	✓	✓	✓	✓	×	×	✓
41	✓	✓	×	✓	✓	✓	✓	✓	✓	×	×	×	×	×	×	×	✓	×	×

4.2.2 Molecular Docking of Standard Drugs

Standard drugs, Oseltamivir and Zanamivir, and a prodrug, GS 4071, are docked with NA macromolecule. The aptitude of the standard drugs to bind the macromolecule is measured by selecting the best confirmation on the basis of binding affinity (Kcal/mol). Comprehensive analysis of the amino acid within 5Å of the binding site has shown that the common amino acids are around the loop 150. Prodrug, GS4071, is shown here in the figure below along with its most significant interactions.

Figure 22: Predicted Interactions of Prodrug, GS4071, within Binding Site of NA.

It is evident that GS4071 showed two hydrophobic interactions and three hydrogen bond interactions. The hydrophobic interactions are between hydrophobic C-11 of the ligand and 2nd Epsilon Carbon of TYR 169 with a distance measurement 3.95Å and hydrophobic C-3 of the prodrug with the Gamma Carbon of TYR 169. Moreover O-38, N-36 and O-42 interact with 2nd Delta Nitrogen of ASN170, Gamma oxygen of

SER168 and 2nd Delta Oxygen of ASP113 with 3.74, 4.30 and 4.31Å distances respectively. Standard Drug, Oseltamivir, showed majorly hydrophobic interactions with the binding site as illustrated in Figure 23.

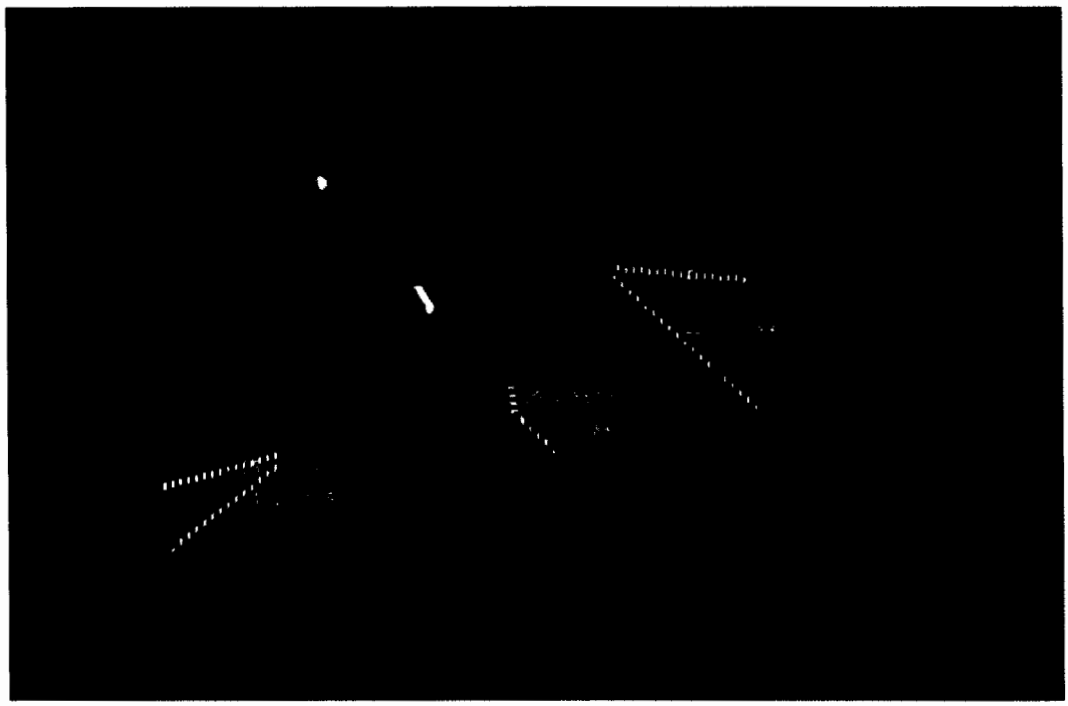


Figure 231: Predicted Interactions of Standard Drug, Oseltamivir, within the Binding Site of NA.

From the Figure above it is evident that C-14 of Oseltamivir shows hydrophobic interactions with 2nd Gamma Carbon of ILE 108 and VAL 114 with the distance 3.74 and 4.03Å respectively. Similarly C-12 interacted with 2nd Epsilon Carbon of TYR169 with 4.18Å distance and with Gamma Carbon of TYR169 at a distance range of 3.75Å in hydrophobic manner. Moreover C-3 interacted in a hydrophobic fashion with 2nd Epsilon Carbon and 2nd Delta Carbon of TYR169 at 3.82 and 3.89Å correspondingly. No hydrogen interactions are found due to large distance between HBA and HBDs of ligand and target site. Similarly, no ionic interactions are found in a distance range reasonable enough to be considered important, despite of the presence of ionizable residues in the vicinity.

4.2.3 Molecular Docking of Selected Data Set

In order to get full insight of the activity of ligands, all the compounds in the data set are docked into the target site using the same procedure as used for the standard drugs. Best conformation of each docked ligand is selected on the basis of affinities of the docked conformations. Affinity is a measure of how tightly ligands bind to the target site on the macromolecules, and lower the affinity value more stable the docked conformation will be. The docked conformations are viewed and processed in VMD to get the optimal interactions of ligands with the target site.

On the basis of pharmacophore features of the ligands and the presence of type of amino acids in the vicinity of target site where ligands are docked, possible binding interaction such as hydrophobic, hydrogen bond, ionic and vanderwaal interactions are found within the reasonable distance range of up till 5Å. The binding interactions along with the binding affinity of all the ligands and the standard drugs are made known in Table 7.

A compound should have low binding energy as compared to standard drugs, enough possible interactions with the neighboring amino acids at the target site that keep the ligand bound to the target site once it reach there and also as good experimental biological activity as a standard to be considered a lead compound. On a substantial analysis of interactions of ligands with the amino acids in their vicinity, their binding energies and experimental data of their biological activity, it is noticed that compound 4, 8, 22 and 39 show more interactions then the rest of the ligands. Moreover these specified compounds have similar amino acids in their 5Å area and they showed interactions with similar amino acids as can be seen in table 7.

Table 7: Interactions of Ligands within 5Å and Their Affinity with Binding Site.

C. No	IC 50 (μM)	Affinity (Kcal/Mol)	Hydrophobic Interactions	Vander Waal Interactions	Hydrogen Bonding	Ionic Interactions
1	0.001	-5.7	C(3) -TYR169:CB [3.81] C(3)-TYR169: CD1 [4.06] C(11)-TYR169:CD2 [3.95] C(11)-TYR169:CE2 [3.95]	None	N(37)-SER168:OG [4.30] O(38)-ASN170:ND2 [3.743]	None
2	0.0007	-5.8	None	None	N(35)-GLU165:OE2[4.12] O(39)-ARG130:NH2[4.13] O(37)-ARG130:NH1[3.10] O(40)-ARG172:NH1[2.80]	C(7)-ARG130:NH2[4.00]
3	0.015	-5.1	C(12)-TYR169:CG [3.75] C(3)-TYR169:CE2 [3.82] C(3)-TYR169:CD2 [3.89] C(14)-ILE108:CG2 [3.74]	None	None	None
4	0.035	-5.4	C(17)-TYR169:CG [3.83] C(17)-TYR169:CD1 [3.67] C(3)-TYR169:CZ [3.83] C(9)-TYR169:CD2[3.55]	None	O(55)-ASP113:OD2[4.73] O(55)-SER168:OG[2.98] O(56)-SER168:OG[3.77]	None
5	0.127	-5.3	None	None	N(43)-GUI128:OE2 [4.94] N(42)-GLU128:OE1 [3.25] O(46)-ARG172:NH1 [4.31]	N(42)-GLU128:CD [4.09]
6	0.024	-5.8	None	None	N(50)-GLU165:OE1 [3.25] N(51)-ASN170:OD1 [4.70]	N(50)-GLU165:CD [4.13]
7	0.002	-6.2	None	None	O(38)-ASP113:OD1 [2.98] O(40)-ASP113:OD2 [2.72]	N(35)-ASP113:CG [4.00]
8	0.0044	-6.2	None	None	N(33)-ASP113:OD1 [2.80] O(41)-ASP113:OD1 [2.98]	N(33)-ASP113:CG [4.00]
9	0.0316	-6.0	None	None	O(42)-ASP113:OD2 [2.70] N(35)-ASP113:OD1 [2.81]	N(36)-ASP113:CG[4.01]
10	0.0112	-6.0	None	None	O(39)-ARG130:NE [3.05] O(40)-GLU128:OE1 [2.81] N(34)-GLU165:OE2 [3.36]	N(34)-GLU165:CD [4.55]
11	0.144	-6.2	None	None	O(46)-ASP113:OD2 [2.71]	N(41)-ASP113:CG [3.98]

					O(50)-ASP113:OD1 [2.93] N(41)-ASP113:OD1 [2.80]	
12	0.08	-6.6	C(8)-TYR169:CE2 [3.70] C(8)-TYR169:CD2 [3.52]	C(7)-TYR169:CG [3.58]	N(44)-GLU165:OE1 [4.43] N(41)-SER168:OG [4.37]	None
13	0.32	-6.9	C(6)-TYR169:CG [3.76] C(7)-TYR169:CG [3.54] C(4)-TYR169:CD2 [3.82]	None	N(33)-SER168:OG[4.22]	None
14	1.66	-7.0	C(8)-TYR169:CE2 [3.52] C(11)-TYR169:CE2 [3.80]	C(6)-TYR169:CG [3.70]	N(30)-GLU165:OE2 [4.60] N(28)-SER168:OG [4.40]	None
15	1.83	-4.8	None	None	N(24)-ASN208:OD1[4.19] O(27)-ASN208:ND2[3.98]	None
16	>20	-6.3	C(2)-TYR169:CD2 [3.50] C(8)-TYR169: CE2 [3.62]	C(2) -TYR169:CD2 [3.50] C(7)-TYR169:CG [3.56]	N(44)-SER168:OG [4.37]	None
17	8.9	-6.9	C(15)-PRO162:CG [3.37] C(13)-PRO162:CB [3.85]	None	O(31)-ASN208:OD [4.48] O(30)-ARG172:NH1 [4.21]	None
18	2.6	-7.4	None	None	O(26)-TYR100:OH [3.50] O(29)-ARG172:NH1 [3.02]	None
19	9.0	-6.7	C(11)-TRP403:CZ3 [4.29] C(15)-TRP403:CH2 [4.05]	C(15)-TRP403:CH2 [4.05] C(11)-TRP403:CZ2 [4.73]	O(29)-ASN347:OD1 [2.87] O(30)-ARG371:NE [3.07] O(31)-ARG371:NE [4.65]	None
20	12.9	-7.0	None	None	O(67)-ARG172:NH1 [3.70] O(66)-ARG172:NE [3.65] O(58)-GLU174:OE1 [2.91]	None
21	18.9	-7.6	C(25)-PRO162:CG [4.52]	None	O(63)-GLU128:OE1 [2.95] O(58)-GLU174:OE1 [3.12] O(66)-ARG172:NE [3.76]	None
22	3.1	-7.6	C(6)-TYR169:CZ [3.63] C(7)-TYR169:CE2 [4.39] C(20)- ILE108:CG2 [3.94]	C(5)-TYR169:CE2 [3.83] C(6)-TYR169:CZ [3.63] C(10)-TYR169:CE1 [3.56]	O(42)-TYR169:OH [4.63] O(41)-SER168:OG [3.13]	None
23	27.8	-6.1	None	None	O(41)-GLU463:OE1 [3.76]	None
24	25.6	-6.3	None	None	O(27)-GLU174: OE1 [4.58] O(30)-ARG172:NH1 [4.48]	None
25	35.4	-7.8	None	None	O(49)-ARG172: NH1 [3.30] O(49)-ARG172:NE [3.67]	None
26	48.56	-6.7	C(3)-TRP403:CH2 [4.56]	C(3)-TRP403:CH2 [4.56]	O(28)-ASN325:OD1 [2.91]	None

			C(7)-TRP403:CH2 [3.91] C(7)-TRP403:CZ [4.29]	C(6)-TRP403:CH2 [4.55] C(7)-TRP403:CH2 [3.91]	O(26)-ASN347:ND2 [3.06] O(27)-ASN347:OD1 [3.30]	
27	77.1	-6.6	None	None	O(26)-ASN141:ND2 [2.81] O(27)- ASP142:OD2 [3.63]	None
28	58.39	-7.1	None	None	O(27)-TYR100:OH [3.49] O(30)-ARG172:NH2 [4.59]	None
29	37.1	-6.7	C(12)-TRP403:CH2 [4.28] C(14)-TRP403:CH2 [3.79]	C(12)-TRP403:CZ3 [4.42] C(15)-TRP403:CH2 [4.70]	O(27)-ASN347:OD1 [2.78]	None
30	33.7	-6.5	None	None	O(26)-ARG171:NE [4.70] O(28)-ASN208:ND2 [3.61]	None
31	82.6	-6.9	None	None	O(27)-GLU463:OE2 [4.19] O(28)-ASN146:ND2 [4.13]	None
32	>100	-7.5	None	None	O(47)-GLU128:OE1 [4.38] O(49)-ARG172:NE [3.91] O(52)-GLU174:OE1 [2.70]	None
33	>100	-6.6	C(4)-TRY169:CD2 [3.38] C(7)-TRY169:CD2 [4.21]	C(1)-TRY169:CD2 [4.32] C(4)-TRY169:CD2 [3.38] C(7)-TRY169:CD2 [4.21]	O(32)-SER168:OG [4.68] O(35)-SER168:OG [4.55]	None
34	0.310	-5.8	None	None	O(39)-GLU165:OE2 [3.04]	None
35	0.025	-5.9	None	None	N(38)-SER168:OG [3.88] O(42)-SER168:OG [3.59]	None
36	0.075	-6.9	None	None	N(48)-SER168:OG [4.99] O(51)-GLU165:OE1 [3.01]	None
37	0.052	-6.1	C(16)-TYR169:CD1 [3.75]	None	N(44)-SER168:OG [3.87] O(48)-SER168:OG [3.62]	None
38	0.008	-6.2	None	None	N(41)-SER168:OG [3.85] O(46)-SER168:OG [3.59]	None
39	0.001	-5.2	C(8)-TYR169:CD1 [3.71] C(14)-TYR169:CE2 [3.66]	None	N(40)-SER168:OG [3.81] N(43)-SER168:OG [3.82]	None
40	0.130	-5.3	None	None	O(35)-ARG172:NE [4.50]	None
41	0.022	-5.8	None	None	N(37)-SER168:OG [3.98] N(38)-SER168:OG [3.00] O(39)-SER168:OG [3.01]	None

It was so observed after the detailed 3D analysis of the best conformation of each ligand receptor complex that ASP 113, GLU 165, SER 168, TYR 169 and ARG 172 participated actively in the binding interactions. TYR 169 is observed as the major contributor in hydrophobic interactions while ASP 113, GLU 165, SER 168 and ARG 172 chipped in hydrogen bonding. Ionic and vanderwaal interactions are also seen in some compounds.

4.2.4 Lead Compound Identification

Analysis of compounds in terms of their binding energies and binding interactions direct to the selection of possible lead compound. As identified prior that the standard drugs interacted with the target primarily via hydrogen bonding and hydrophobic interactions, as shown in table 7, these two types of interactions are given prime importance during lead selection along with the other parameters defined earlier. Compound 4, 8, 22 and 39 are considered as candidates for becoming the lead. Compounds 7, 8 and 9 showed same types of interactions with nearly equivalent distances so only compound 8 is considered as a lead candidate on the basis of its biological activity and low binding energy.

Selected compounds showed following interactions: Compound 4 has many hydrophobic interactions out of which the 4 with lowest distance measures are chosen to be presented here. The distances are 3.55, 3.67, 3.83 and 3.83Å. In addition, hydrogen bond interactions with distance measures 2.98, 3.71 and 4.73Å are also observed. Three hydrogen bond interactions measuring 2.72, 2.80 and 2.98Å and an ionic interaction measuring 4.01Å are found in compound 8. Compound 22 showed three hydrophobic and vanderwaal interactions with distance measures 3.63, 3.94 and 4.39Å along with two hydrogen bond interactions that measured 3.13 and 4.63Å. Two

hydrophobic and two hydrogen bond interactions with distances 3.66, 3.71, 3.81 and 3.82Å are observed in compound 39. Considering the interactions above compound 4 is selected as a lead compound. Compound 8 though has smaller distance measures but showed no hydrophobic interaction and hence is not selected as a lead.

As declared earlier compound 4 selected as a lead showed four hydrophobic interactions including C-9 of ligand with 2nd Delta Carbon with a distance measure of 3.55, C-17 with 1st Delta Carbon with 3.67, C-17 with Gamma Carbon with 3.83 and C-3 with terminal Carbon CZ of TYR 169 ring with a distance of 3.83Å. In addition three hydrogen bond interactions take account of O-55 with Gamma Oxygen of SER 168 with 2.98 distance measure, O-56 with Gamma Oxygen of SER 168 with 3.71 and O-55 with 2nd Delta Oxygen of ASP 113 distanced at 4.73Å. Figure 24 shows hydrophobic and hydrogen bond interactions respectively.

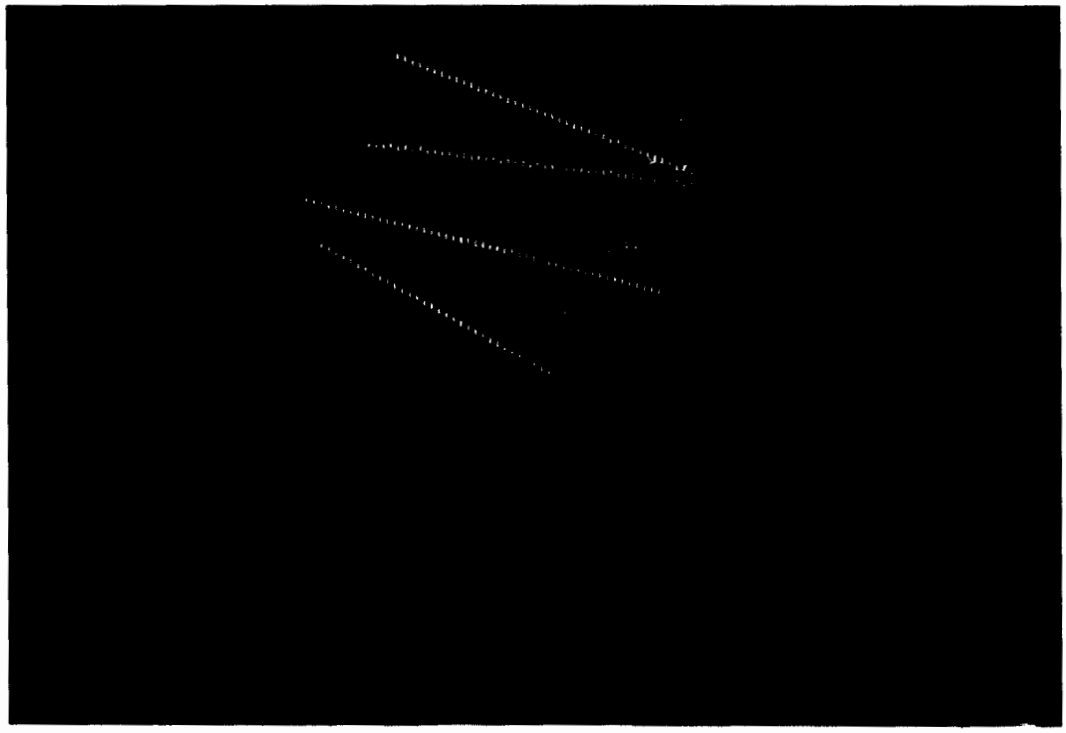


Figure 24: Predicted Hydrophobic Interactions of Lead Compound with Target Binding Site of NA.

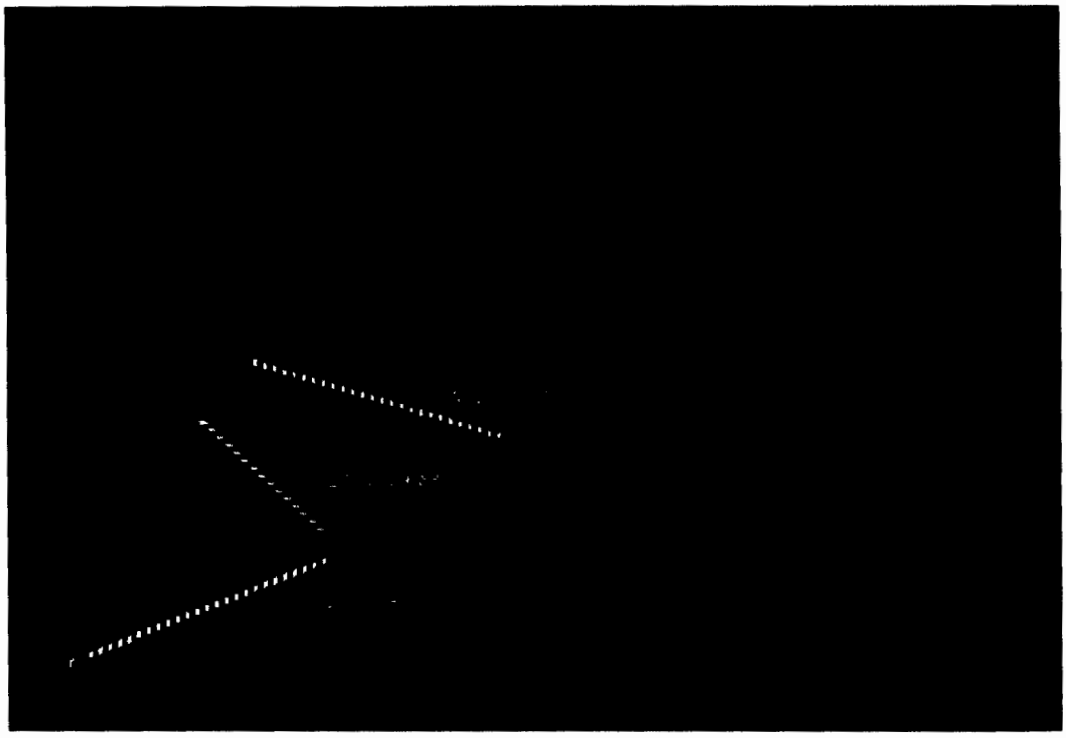


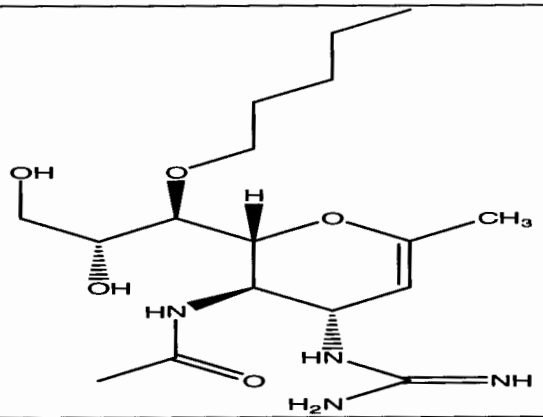
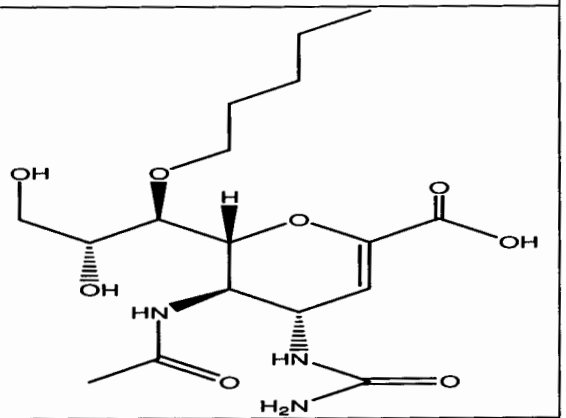
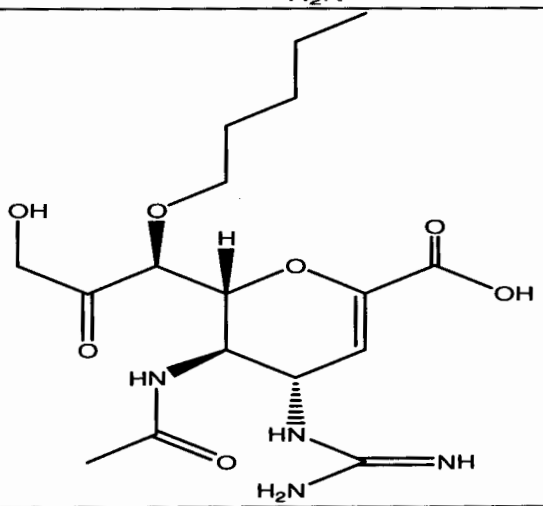
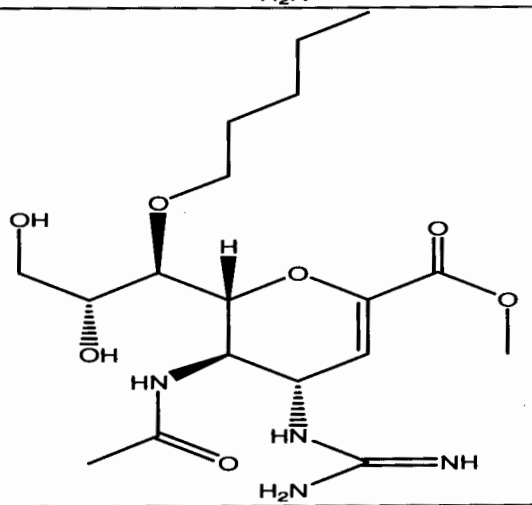
Figure 25: Predicted Hydrogen Bond Interactions of Lead Compound with Target Binding Site of NA.

4.2.5 Analogue Designing of Lead Compound:

After selecting compound 4 as a lead, its analogues are designed in order to maximize the interactions of the resulting compounds leading into more efficient and active compounds that can act as drugs. Analogues are designed by means of adding, deleting or replacing a functional group considering the information of their interactions. Analogues designed for compound 4 are given in table 8.

Analogue 1 is designed by means of retro reactions in order to make the compound more hydrophobic. Hydrogen bond interactions have been found very important in standard as well as lead compound so an effort to increase the hydrogen bond interactions of compound has been done via Ketone formation in 2nd analogue. Similarly, 3rd and 4th analogues are made via oxidation of alcohol and ester formation respectively to increase hydrogen bonding.

Table 8: Analogues of Lead Compound.

Analogue No.	Action Taken	Chemical Structure	Analogue No.	Action Taken	Chemical Structure
1.	Retro Reaction		2.	Ketone Formation	
3.	Oxidation of Alcohol		4.	Ester Formation	

All the analogues thus made are then docked into the NA binding site using the same course of action as for other compou nds mentioned earlier. The best conformation of each analogue is selected and its PDB file studied for binding interactions with the target site in VMD. The best possible interactions are made known in table 9.

Table 9: Binding Interactions of Analogues with Amino Acids Present in 5.0Å.

Analogue No.	Hydrophobic Interactions	Hydrogen Bonds	Ionic Interaction
1	C(3)-ALA138:CA[4.60] C(9)-LEU139:CD1[3.98] C(11)-LEU139:CD1[3.71] C(12)-LEU139:CD1[4.42] C(14)-ALA138:CA[3.72]	N(50)-ASP113:OD1[4.31] N(53)-ASP113:OD1[2.87] O(57)-ASP113:OD1[2.95] O(54)-ASP142:OD1[4.04]	N(53)-ASP113:CG[4.06]
2	C(3)-LEU139:CD1[4.11] C(9)-ALA138:CA[4.36] C(10)-LEU139:CD1[3.96] C(15)-LEU139:CD1[3.65] C(17)-LEU139:CD1[4.32]	N(48)-ASP113:OD1[4.21] N(49)-ASP113:OD1[2.91] O(50)-ASP113:OD2[2.74] O(54)-ASP113:OD1[3.85] O(55)-ASP113:OD1[2.85] O(56)-ASN141:ND2[3.73]	N(49)-ASP113:CG[4.11]
3	C(3)-PRO162:CB[4.32]	O(52)-ARG172:NE[4.97] O(53)-ARG172:NH1[3.04] O(54)-ARG172:NH1[4.03] O(56)-ARG172:NH1[4.29]	None
4	C(3)-TYR169:CZ[3.74] C(9)-TYR169:CD1[3.78] C(10)-TYR169:CD2[3.43] C(13)-TYR169:CD2[3.80] C(14)-TYR169:CD2[3.98]	O(56)-ASN170:ND2[3.83]	None

As is evident from the table over all three types of interactions are studied here i.e. hydrophobic, hydrogen and ionic interactions. Analogues showed more potential according to the predicted interactions as 1st analogue has five hydrophobic, four hydrogen bonds and one ionic interaction. Likewise 2nd analogue has five hydrophobic, six hydrogen bonds and one ionic interaction. 3rd analogue has one hydrophobic, four hydrogen bonds and no ionic interaction. In addition, 4th analogue has five hydrophobic and one hydrogen bond with no ionic interaction.

It is evident from Figure 26 that ALA 138 is involved in two hydrophobic interactions with distance measures 3.72 and 4.60Å while LEU 139 shows three

hydrophobic interactions 3.71, 3.98 and 4.42Å. Substantial hydrogen bond interactions have been found with ASP 113 and ASP 142 measuring 4.31, 2.87 and 2.95Å with ASP 113 and 4.04Å with ASP 142. ASP 113 is also drawn in an ionic interaction measuring 4.06Å that will give more stability to ligand protein complex.

Figure 261: Predicted Interaction of Analogue 1 with Binding Site of NA.

Analogue 2 also showed significant interactions with the amino acids present in the binding site. Here too, ALA 138 and LEU 139 participated actively in hydrophobic interactions resulting in a 4.36Å interaction with ALA 138 and four interactions measuring 3.65, 3.96, 4.11 and 4.32Å concerning LEU 139. ASP 113 showed five possible hydrogen bond interactions with small distance measures, 2.74, 2.85, 2.91, 3.85 and 4.21Å, pointing towards their significance. ASN 141 is engaged in one hydrogen bond interaction measuring 3.73Å. One ionic interaction with a distance measure of 4.11Å has also been predicted with ASP 113.

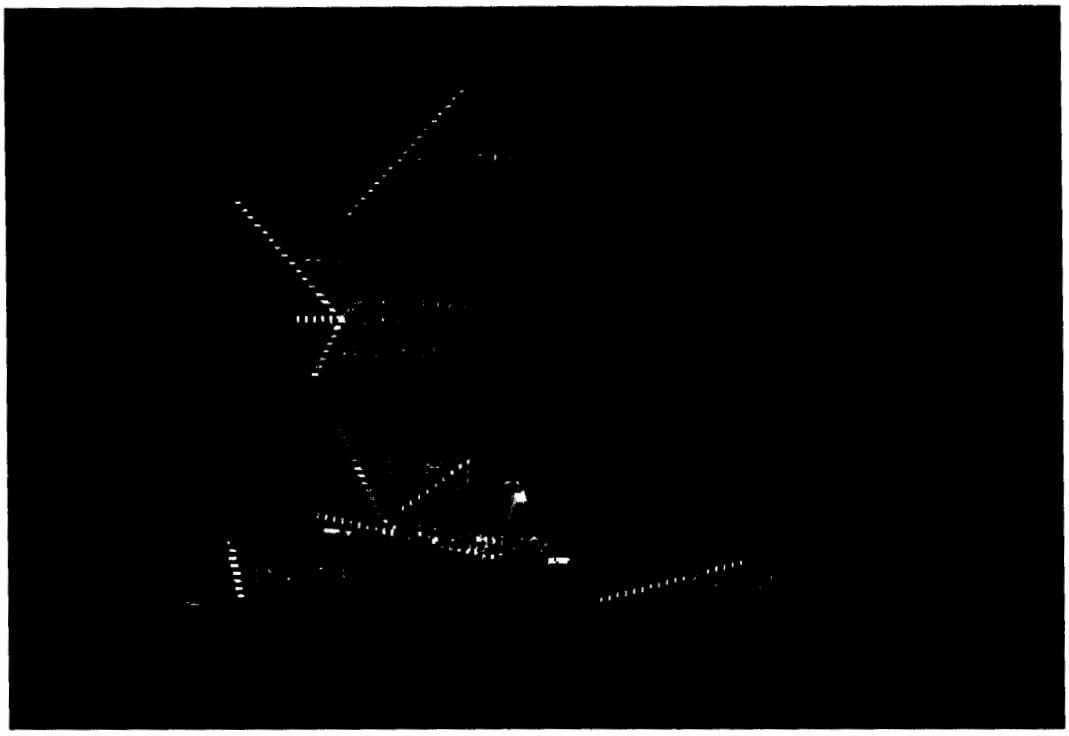


Figure 272: Predicted Interaction of Analogue 2 with Binding Site of NA.

On the basis of the above stated results shown by analogue 1 and 2, it is suggested here that these compound have an increased potential for being the next FDA approved standard in the line of CADD for influenza treatment.

4.3 Quantitative Structure Activity Relationship (QSAR)

In structure based drug designing, to correlate the biological activity (IC50) of a ligand with the chemical structure of a molecule by means of its molecular descriptors, QSAR is performed as a part of Computer Aided Drug Designing (W.J. Lu *et al.*, 2008).

Molecular descriptors are mathematical values signifying that whether a certain physiochemical property is playing part in imparting biological activity to a compound or not. The positive aspect of this scheme is that descriptor calculation can be done without taking into account any experimental techniques, based on the

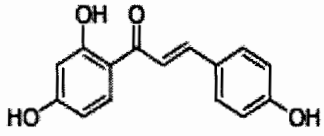
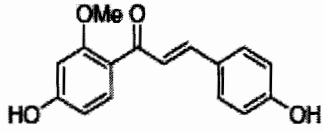
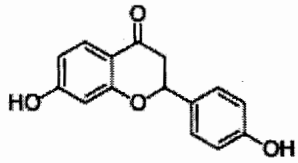
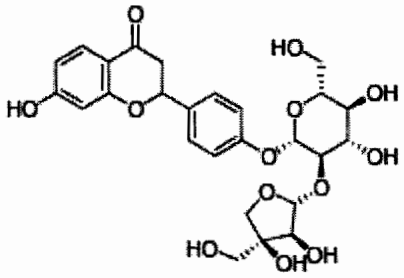
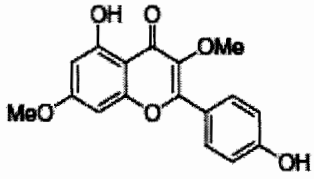
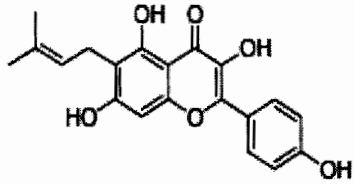
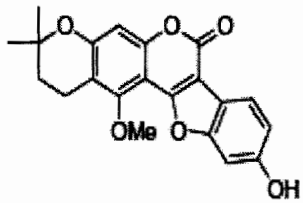
chemical structure of the compounds, using commercially available softwares such as ChemDraw, HyperChem and Dragon (W.J. Lu *et al.*, 2008).

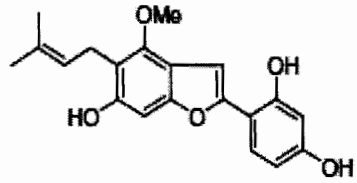
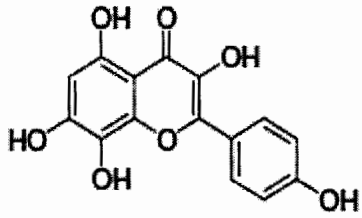
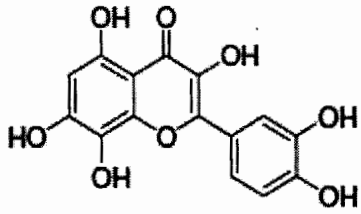
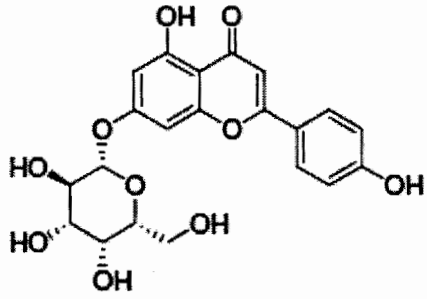
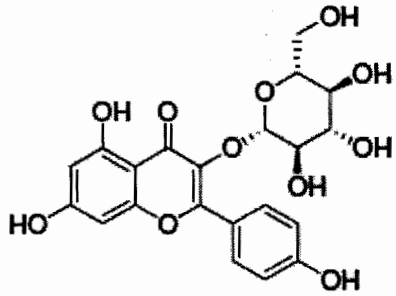
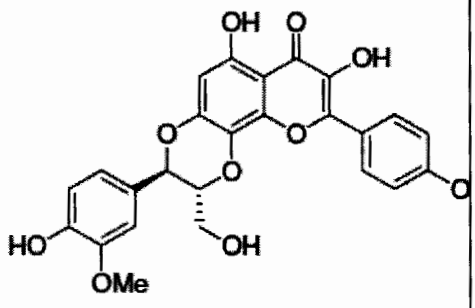
It is usually narrated that alike compounds having similar structures share comparable physiochemical properties, are assumed to have similar biological properties and activity values. This assumption is the basis of nearly all quantitative structure activity relationships (Hugo Kubinyi, 1997). Based on this postulation only one class of compounds, Polyphenols are selected for QSAR study here. A QSAR study conducted in recent past has shown that Flavonoids and biflavonoids, sub groups of polyphenols, can be manipulated for their potential to act as influenza virus H1N1 neuraminidase inhibitors (A.G. Mercader and A.B. Pomilio, 2010).

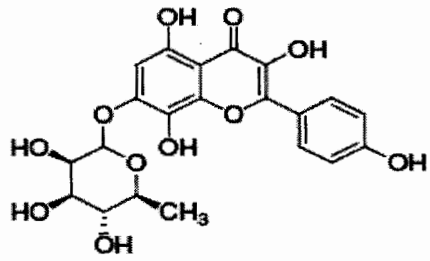
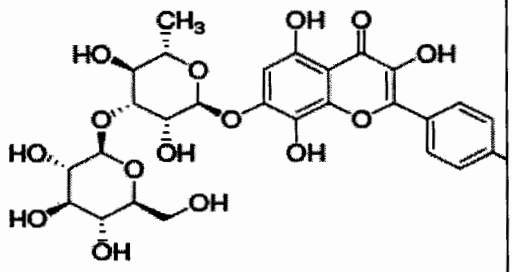
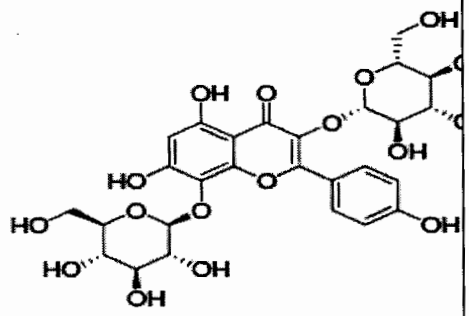
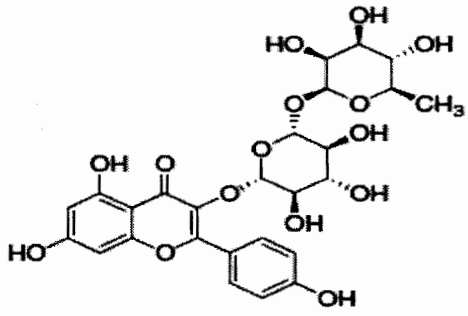
In nearly the same era, more polyphenols popped up in literature proved to have anti-influenza viral activities against neuraminidase protein target (H. J. Jeong *et al.*, 2009; Y. B. Ryu *et al.*, 2010). Most of these compounds are not considered for the former mentioned study of QSAR model of Flavonoids and biflavonoids. An attempt has been made here to include more polyphenols into the QSAR study including just two from the original data set, Gossypetin and Herbacetin, considered for pharmacophore generation and molecular docking. The selected data set for QSAR study has been shown in table 10.

The compounds of the dataset selected for QSAR study belong to a number of groups such as flavones, flavonols, isoflavones, Flavonoids and biflavonoids that belong to a major class of compounds called polyphenols, which is also a part of the classes under discussion earlier for pharmacophore generation and molecular docking studies. The compounds selected for QSAR study here have a large variation in their biological activity, ranging from 2.6 to 92.4μM, so that the results thus derived are consistent with any activity range in a scientific study.

Table 10: Selected Dataset for Quantative Structure Activity Relationship

C. No.	Compound Name	Compound Structure	IC 50 (μM)
1	isoliquiritigenin		9.0
2	20-methoxyisoliquiritigenin		24.3
3	Liquiritigenin		46.8
4	liquiritin apioside		18.2
5	Kumatakenin		36.4
6	Licoflavonol		20.6
7	Isoglycyrol		92.4

8	Licocoumarone		27.8
9	Herbacetin		8.9
10	Gossypetin		2.6
11	Cosmosiin		46.9
12	Astragalin		38.4
13	Rhodiolinin		10.3

14	Rhodionin		32.2
15	Rhodiosin		56.2
16	Linocinamarin		44.2
17	Nicotiflorin		31.2

A number of electronic and thermodynamic descriptors are selected for the QSAR model and parameter calculation is done using ChemDraw and HyperChem. Some of the thermodynamic descriptors chosen for the study, Molar Refractivity (MR) and logP, are computed using ChemDraw. While the electronic parameters such as Highest Occupied Molecular Orbital energy (E_{HOMO}) and Lowest Unoccupied Molecular Orbital energy (E_{LUMO}) and rest of the thermodynamic descriptors such as

Total Energy (E_{TOTAL}) and Heat of Formation (H_f), are determined via HyperChem. The statistics of the calculated parameters are made known in table 11.

Table 11: Thermodynamic and Electronic Descriptors

C. No	log P	MR (cm3/ mol)	E_{TOTAL} (kcal/mol)	H_f (kcal/mol)	E_{HOMO} (kcal/mol)	E_{LUMO} (kcal/mol)
		Thermodynamic Descriptors			Electronic Descriptors	
1	2.42	72.86	-13477.8019	51489.95241	-4.081966	3.470395
2	2.68	78.3	-13477.801	51489.95241	-4.081966	3.470395
3	2.02	109.86	-15847.5908	57904.39023	-10.579349	8.555182
4	-1.13	130.87	-28035.1414	144386.415	-2.765894	4.505488
5	1.6	83.81	2142.549058	73024.359	-10.074473	9.072817
6	2.57	98.91	-17865.7234	85369.67866	-50.033956	0.20711
7	3.08	199.49	-16555.3194	373503.7427	-0.947393	20.928692
8	3.68	97.6	-29953.5149	54191.4916	-29.335367	0.894139
9	0.35	76.51	-24970.5907	57652.47365	-0.516766	2.417364
10	-0.04	78.33	-30405.8481	59566.35735	-1.70081	1.12415
11	0.1	110.47	-16456.4702	114740.4108	-6.945042	8.497634
12	-1	107.97	-31053.1745	111884.9609	-3.176157	7.238313
13	1.29	125.02	-39885.5019	95790.34906	-10.947391	4.317102
14	-0.63	109.1	-9549.56210	115820.12	-32.507429	4.962665
15	-2.38	142.37	-35596.0243	153045.7566	-11.108196	4.97908
16	-3.23	143.73	-31924.9834	172850.1652	-3.558804	8.602381
17	-1.42	135.23	-48245.9056	134481.8191	-2.077337	3.550373

To come up with a parameter that is a considerable player in the biological activity of a compound, the mathematical values attained for parameters, taken as independent variables, are plotted against IC50, a dependent variable, to calculate the value of regression constant. Value of regression constant above 6.0 points to a correlation where as a lower value signifies that there is no correlation between activity and that specific parameter.

4.3.1 Correlation of Thermodynamic Descriptors

LogP, partition coefficient, is a measure of hydrophobicity of a compound where as MR is a critical piece of information to estimate the molecular polarizability. The values of regression constants of logP and MR are 0.0008 and 0.6005 respectively as

is evident from the Figures 28. LogP shows no correlation with the biological activity of a compound but regression value of MR shows that molecular polarizability is playing an imperative role in biological activity of a compound to act as a NA inhibitor.

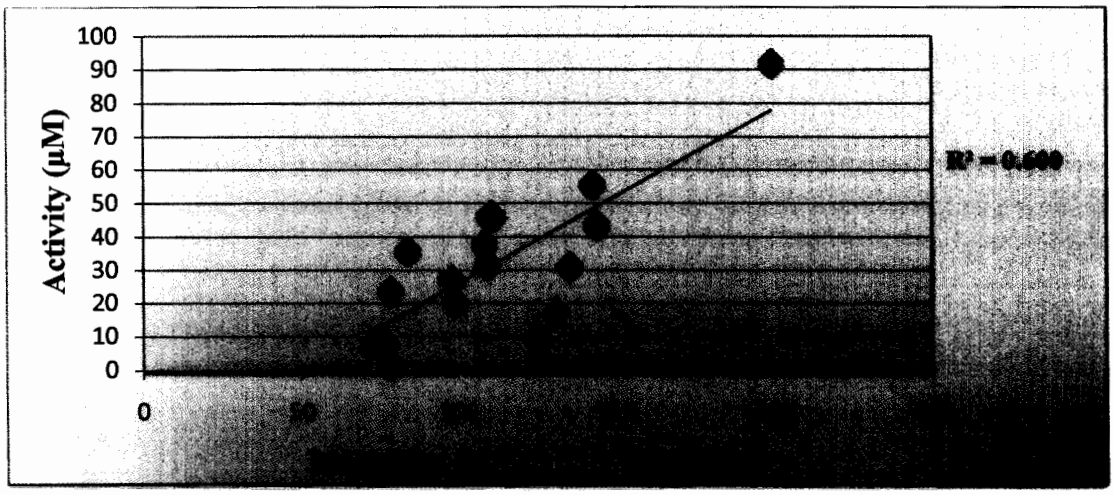


Figure 28: Scatter plot of IC50 and Molar Refractivity.

More over E_{TOTAL} has regression value 0.0204, showing that E_{TOTAL} has no correlation with the biological activity. Heat of formation is the enthalpy for creating a molecule from its component atom and plays role in the thermal constancy and stability of a molecule. The value of regression constant for H_f is 0.6312 that illustrates its role in the biological activity.

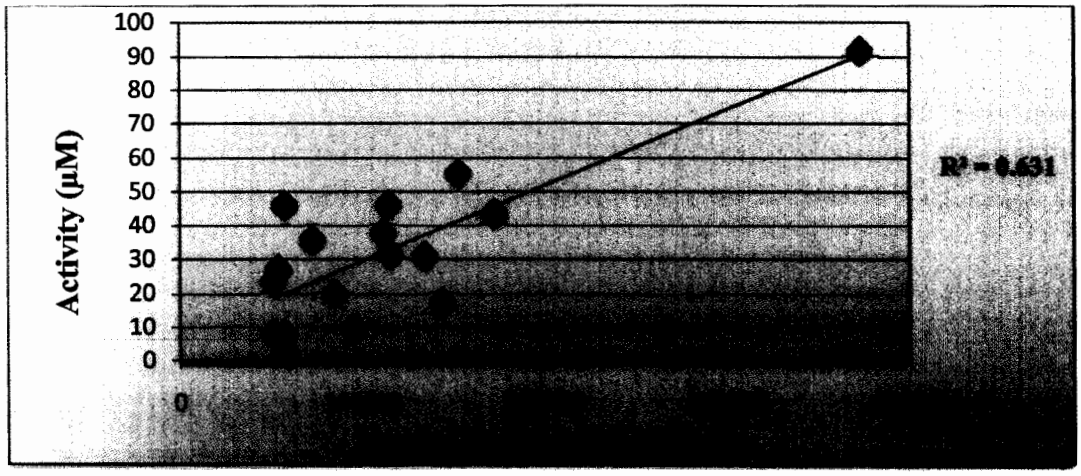


Figure 29: Scatter Plot of IC50 and Heat of Formation.

4.3.2 Correlation of Electronic Descriptors:

E_{HOMO} and E_{LUMO} are determinants of excitability of a compound. The regression values of E_{HOMO} and E_{LUMO} are 0.0103 and 0.7372 respectively as shown in Figure 30. The values are evidence that E_{HOMO} has no correlation with biological activity in this case while E_{LUMO} is an important factor in imparting biological activity to a NA inhibitor. Lowest Unoccupied Molecular Orbital is the low energy orbital that has room to receive an electron and act as an electron acceptor in an interaction between ligand and protein residues. Thus E_{LUMO} has its share in governing the reactivity of a molecule by measuring the electrophilicity.

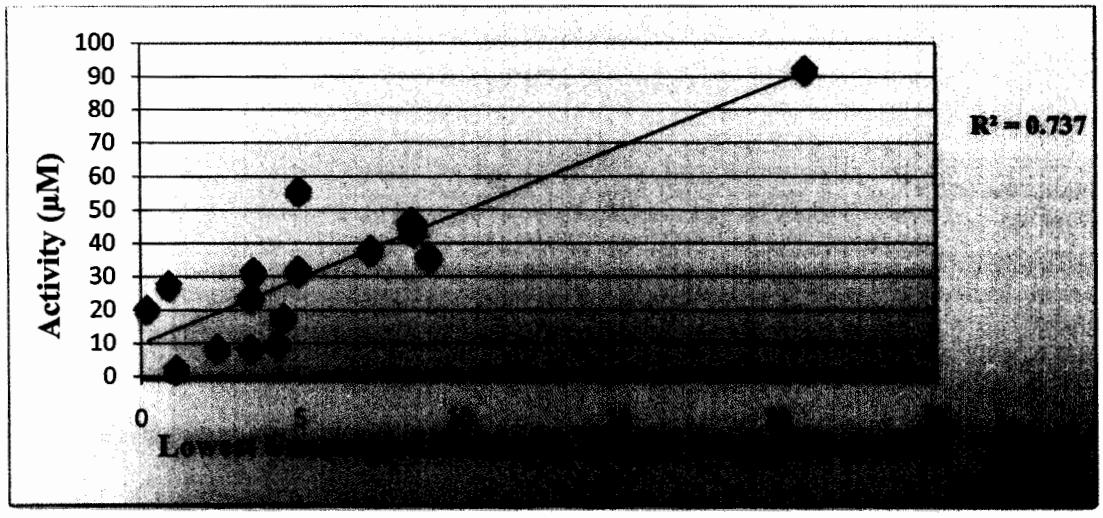


Figure 30: Scatter Plot of IC50 and E_{LUMO}.

CONCLUSION AND FUTURE ENHANCEMENT

The conducted study is optimistic about the upshot that it would be very helpful for the betterment of the world that is always a red zone of Influenza epidemic. A pharmacophore consisting of 2 HBs and 1 HBD has been suggested. It is also found that ASP 113, ALA 138, ASP 142, TYR169 and ARG 172 are very important for binding interactions. Moreover, the identified lead along with analogue 1 and 2 on the basis of their interactions are anticipated to be manipulated in clinical trials to bring more effective and potent drugs against seasonal and epidemic influenza in upcoming years. In addition, during QSAR studies correlation coefficients' value lead to the consequence that MR, H_f and E_{LUMO} are important players in biological activity of polyphenols. The identified lead and proposed analogues can be further studied in wet lab and clinical trials to bring a better drug to market. In addition, more classes can be incorporated into the study for more general results of QSAR.

REFERENCES

Aatu Kaapro and Janne Ojanen, (Nov 27, 2002) "Protein docking", URL: http://www.lce.hut.fi/teaching/S-114.500/k2002/Protdock.pdf.

Ai-Lin Liu, Hai-Di Wang, Simon MingYuen Lee, Yi-Tao Wang, Guan-Hua Du, (2008) "Structure–activity relationship of flavonoids as influenza virus neuraminidase inhibitors and their in vitro anti-viral activities", Bioorganic & Medicinal Chemistry 16, 7141–7147.

Amy C. Anderson, (2003) "The Process of Structure-Based Drug Design", Chemistry and Biology 10, 787-797.

Andreas May and Martin Zacharias, (2005) "Accounting for global protein deformability during protein–protein and protein–ligand docking", Biochimica et Biophysica Acta 1754, 225 – 231.

Andrew G. Mercader, Alicia B. Pomilio, (2010) "QSAR study of flavonoids and biflavonoids as influenza H1N1 virus neuraminidase inhibitors", European Journal of Medicinal Chemistry, 1–7.

AR Tapas, DM Sakarkar and RB Kakde, (2008) "Flavonoids as Nutraceuticals: A Review", Tropical Journal of Pharmaceutical Research 7 (3), 1089-1099.

Areej M. Abu Hammad, and Mutasem O. Taha, (2009) "Pharmacophore Modeling, Quantitative Structure Activity Relationship Analysis, and Shape-Complemented in Silico Screening Allow Access to Novel Influenza Neuraminidase Inhibitors", J. Chem. Inf. Model. 49 (4), 978-996.

Atsuko Yamaguchi, Hiroyasu Shimada, Ken Horiuchi and Tetsuo Nishikawa, "A New Method of Computing Ligand-Based Pharmacophore Models for Flexible Chemical Compounds".

Bevyn Jarrott, (2004) "Contributions to drug design and development", Chemistry in Austeralia, 14-17.

C. Tsinopoulos and I. P. McCarthy (2002) "An evolutionary classification of the strategies for drug discovery", Manufacturing Complexity Network Conference, Cambridge, 373-385.

Carolina H. Andrade, Kerly F. M. Pasqualoto, Elizabeth I. Ferreira and Anton J. Hopfinger, (2010) "4D-QSAR: Perspectives in Drug Design", Molecules 15, 3281-3294.

Claudia Pappas, Patricia V. Aguilar, Christopher F. Basler, Alicia Solo´ rzano, Hui Zeng, Lucy A. Perrone, Peter Palese, Adolfo Garcı´a-Sastre, Jacqueline M. Katz, and Terrence M. Tumpey, (2008) "Single gene reassortants identify a critical role for PB1, HA, and NA in the high virulence of the 1918 pandemic influenza virus", PNAS 105, 3064-3069.

David M. Morens, Jeffery K. Taubenberger, and Anthony S. Fauci, (2009) "The Persistent Legacy of the 1918 Influenza Virus", N Engl J Med, 225-229.

David A. Winkler, (2002) "The role of Quantitative Structure Activity Relationships (QSAR) in biomolecular discovery", Briefings in Bioinformatics 3 (1), 73-86.

Dipanwita Basu, Marcin P. Walkiewicz, Matthew Frieman, Ralph S. Baric, David T. Auble and Daniel A. Engel, (Feb. 2009) "Novel Influenza Virus NS1 Antagonists Block Replication and Restore Innate Immune Function", Journal Of Virology 83(4), 1881-1891.

E. D. Kilbourne, (1987) "History of influenza, In Influenza", New York: Plenum Press, 3-22.

E. G. Brown, (2000) "Influenza virus genetics", Biomed & Pharmacother 54, 196-209.

Ewing TJ, Makino S, Skillman AG, Kuntz ID (2001), "DOCK 4.0: search strategies for automated molecular docking of flexible molecule databases", J Computer Aided Mol Des 15, 411–428.

F. G. Hayden and M. G. Ison, (2002) "Viral infections in immunocompromised patients: what's new with respiratory viruses", Curr. Opin. Infect. Dis. 15, 355-367 (review).

G. W. Langer, (2005) "Ligandscout: 3-d pharmacophores derived from protein-bound ligands and their use as virtual screening filters", Journal of Chemical Information and Modeling, 160-169.

Ganellin CR, Wermuth CG, Lindberg P and Mitscher LA, (1998) "Glossary of terms used in medicinal chemistry", Journal of Pure and Applied Chemistry, 1129-1143.

Hugo Kubinyi, (1997) "QSAR and 3D QSAR in Drug Design Part 2: Applications and Problems", DDT 2, 538-546.

Hypercube, Inc., HyperChem® Release 7 for Windows®, (Jan 2002)

Hyung Jae Jeong, Young Bae Ryu, Su-Jin Park, Jang Hoon Kim, Hyung-Jun Kwon, Jin Hyo Kim, Ki Hun Park, Mun-Chual Rho and Woo Song Lee, (2009) "Neuraminidase inhibitory activities of flavonols isolated from Rhodiola rosea roots and their in vitro anti-influenza viral activities", Bioorganic & Medicinal Chemistry 17, 6816–6823.

I.M. Kapetanovic, (2008) "Computer-Aided Drug Discovery and Development (Caddd): In Silico-Chemico-Biological Approach", Chem Biol Interact. 171 (2), 165-176.

James B. Finley, Venkatram R. Atigadda, Franco Duarte, James J. Zhao, Wayne J. Brouillette, Gillian M. Air and Ming Luo, (1999) "Novel Aromatic Inhibitors of Influenza Virus Neuraminidase Make Selective Interactions with Conserved Residues and Water Molecules in the Active Site", J. Mol. Biol., 1107-1119.

Jianzhi Gong, Wenfang Xu and Jie Zhang, (2007) "Structure and Functions of Influenza Virus Neuraminidase", Current Medicinal Chemistry 14, 113-122.

Jie Zhang, Qiang Wang, Hao Fang, Wenfang Xu, Ailin Liub and Guanhua Du, (2006) "Design, synthesis, inhibitory activity, and SAR studies of pyrrolidine derivatives as neuraminidase inhibitors", Bioorganic and Medicinal Chemistry 15, 2749-2758.

Jie Zhang, Qiang Wang, Hao Fang, Wenfang Xu, Ailin Liu and Guanhua Du, (2007) "Design, synthesis, inhibitory activity, and SAR studies of pyrrolidine derivatives as neuraminidase inhibitors", Bioorganic and Medicinal Chemistry 15, 2749-2758.

Jitender Verma, Vijay M. Khedkar and Evans C. Coutinho, (2010) "3D-QSAR in Drug Design - A Review", Current Topics in Medicinal Chemistry 10, 95-115.

John Beigel and Mike Bray, (2008) "Current and future antiviral therapy of severe seasonal and avian influenza", Antiviral Res. 78(1), 91-102.

John H. Van Drie, (2007) "Monty Kier and the Origin of the Pharmacophore Concept", Internet Electronic Journal of Molecular Design 6, 271–279.

Jonathan P. Moorman, (2003) "Viral Characteristics of Influenza", Southern Medical Journal, 758-761.

Kalyan Das, James M Aramini, Li-Chung Ma, Robert M Krug and Eddy Arnold, (2010) "Structures of influenza A proteins and insights into antiviral drug targets", Nature Structural and Molecular Biology, 17, 530 – 538.

Karl G Nicholson, John M. Wood, Maria Zambon, (2003) "Influenza", Lancet 362, 1733-45.

Kristin Cobb, (2007) "DOCK THIS: In Silico Drug Design Feeds Drug Development", Biomedical Computation Review, 20-30.

Larisa V Gubareva, Laurent Kaiser and Frederick G Hayden, (2000) "Influenza virus neuraminidase inhibitors", Lancet 355, 827–35.

L. D. Mendelsohn, (2004) "ChemDraw 8 Ultra, Windows and Macintosh Versions", J. Chem. Inf. Comput. Sci., 2225-2226.

Leslie J. Holsinger, Deepali Nichani, Lawrence H. Pinto and Robert A. Lamb, (1994) "Influenza A Virus M2 Ion Channel Protein: a Structure-Function Analysis", Journal of Virology, 1551-1563.

Lyndon J. Mitnaul, Mikhail N. Matrosovich, Maria R. Castrucci, Alexander B. Tuzikov, Nikolai V. Bovin, Darwyn Kobasa and Yoshihiro Kawaoka, (July 2000) "Balanced Hemagglutinin and Neuraminidase Activities Are Critical for Efficient Replication of Influenza A Virus", Journal Of Virology 74, 6015-6020.

Marja Van Zeijl, Jeanette Fairhurst, Thomas R. Jones, Steven K. Vernon, John Morin, James Larocque, Boris Feld, Bryan O'hara, Jonathan D. Bloom and Stephen V. Johann, (2000) "Novel Class of Thiourea Compounds That Inhibit Herpes Simplex Virus Type 1 DNA Cleavage and Encapsidation: Resistance Maps to the UL6 Gene", Journal Of Virology, 9054-9061.

Matthew A. Williams, Willard Lew, Dirk B. Mendel, Chun Y. Tai, Paul A. Escarpe, W. Graeme Laver, Raymond C. Stevens, and Choung U. Kim, (1997) "Structure-activity relationships of Carbocyclic influenza neuraminidase inhibitors", Bioorganic & Medicinal Chemistry Letters 7, 1837-1842.

Misako Nakazawa, Shin-etsu Kadowaki, Izumi Watanabe, Youko Kadowaki, Masaya Takei and Hideyuki Fukuda, (2008) "PA subunit of RNA polymerase as a promising target for anti-influenza virus agents", Antiviral Research 78, 194-201.

Nancy J. Cox, Gabriele Neumann, Ruben O. Donis and Yoshihiro Kawaoka, (1998) "Orthomyxoviruses: influenza", In B. W. (ed.), Microbiology and microbial infections (9th ed., Vol. 1, 385-433). London, England.

Nicola J. Richmond, Charlene A. Abrams, Philippa R. N. Wolohan, Edmond Abrahamian, Peter Willett and Robert D. Clark, (2006) "GALAHAD: 1. Pharmacophore identification by hypermolecular alignment of ligands in 3D", J Comput Aided Mol Des 20, 567-587.

Oleg Trott and Arthur J. Olson, (2010) "AutoDock Vina: Improving the Speed and Accuracy of Docking with a New Scoring Function, Efficient Optimization, and Multithreading", J Comput Chem 31: 455–461.

O. T. Olson, (2010) "Autodock vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading", Journal of Computational Chemistry, 455-461.

P. M. Colman, (1994) "Influenza virus neuraminidase: Structure, antibodies, and inhibitors", Protein Science 3, 1687-169, Cambridge University Press, Printed in the USA.

P. M. Colman, (1995) "Design and antiviral properties of influenza virus neuraminidase inhibitors", Pure & Appl. Chem 67 (10), 1683-1688.

Pramod C. Nair, M. Elizabeth Sobhia, (2008) "Quantitative structure activity relationship studies on thiourea analogues as influenza virus neuraminidase inhibitors", European Journal of Medicinal Chemistry 43, 293-299.

Rajeshwar P. Verma and Corwin Hansch, (2006) "A QSAR study on influenza neuraminidase inhibitors", Bioorganic & Medicinal Chemistry 14, 982–996.

Richard J Sugrue, Boon-Huan Tan, Dawn SY Yeo and Richard Sutejo, (2008) "Antiviral Drugs for the Control of Pandemic Influenza Virus", Ann Acad Med Singapore 37, 518-24.

Robert M. Krug and James M. Aramini, (2009) "Emerging antiviral targets for influenza A virus", Trends Pharmacol Sci. 30(6), 269-277.

Ron A. M. Fouchier, Vincent Munster, Anders Wallensten, Theo M. Bestebroer, Sander Herfst, Derek Smith, Guus F. Rimmelzwaan, Bjorn Olsen and Albert D. M. E. Osterhaus, (2005) "Characterization of a Novel Influenza A Virus Hemagglutinin Subtype (H16) Obtained from Black-Headed Gulls", Journal of Virology, 2814-2822.

S. J. Daharwal, (2006) "Computer aided drug design and bioinformatics: A current tool for designing", Latest Reviews 4 (4).

Takeshi Honda, Takeshi Masuda, Shuku Yoshida, Masami Arai, Satoru Kaneko and Makoto Yamashita, (2002) "Synthesis and Anti-Influenza Virus Activity of 7-O-Alkylated Derivatives Related to Zanamivir", Bioorganic & Medicinal Chemistry Letters 12, 1925-1928.

Takeshi Honda, Takeshi Masuda, Shuku Yoshida, Masami Arai, Yoshiyuki Kobayashi and Makoto Yamashita, (2002) "Synthesis and Anti-Influenza Virus Activity of 4-Guanidino-7-substituted Neu5Ac2en Derivatives", Bioorganic & Medicinal Chemistry Letters 12, 1921-1924.

Tao-Sheng Huang, Peter Palese and Mark Krystal, (1990) "Determination of Influenza Virus Proteins Required for Genome Replication", Journal of Virology, 5669-5673.

Thanyada Rungrotmongko, Thanyarat Udommaneethanakit, Maturos Malaisree, Nadtanet Nunthaboot, Pathumwadee Intharathep, Pornthep Sompornpisut and Supot Hannongbua, (2009) "How does each substituent functional group of oseltamivir lose its activity against virulent H5N1 influenza mutants?", Biophysical Chemistry 145, 29-36.

Thi Ngoc Anh Nguyen, Trong Tuan Dao, Bui Thanh Tung, Hwanwon Choi, Eunhee Kim, Junsoo Park, Seong-IL Lim and Won Keun Oh, (2010) "Influenza A (H1N1) Neuraminidase Inhibitors from Vitis amurensis", Food Chemistry, doi: 10.1016/j.foodchem.2010.06.049.

Vivek Darapaneni, Varun K. Prabhaker and Andreas Kukol, (2009) "Large-scale analysis of influenza A virus sequences reveals potential drug target sites of non-structural proteins", Journal of General Virology 90 (9), 2124-2133.

Willard Lew, Huiwei Wu, Xiaowu Chen, Bradford J. Graves, Paul A. Escarpe, Holly L. MacArthur, Dirk B. Mendel and Choung U. Kim, (2000) "Carbocyclic Influenza Neuraminidase Inhibitors Possessing a C3-Cyclic Amine Side Chain: Synthesis and Inhibitory Activity", Bioorganic & Medicinal Chemistry Letters 10, 1257-1260.

W. Bylka, I. Matlawska and N.A. Pilewski, (2004) "Natural Flavonoids as Antimicrobial Agents", JANA 7 (2), 24-31.

W. Graham Richards, (1994) "Computer-aided drug design", Pure and Appl. Chern. 66 (8), 1589-1596.

W. J. Lu, Y. L. Chen, W. P. Ma, X. Y. Zhang, F. Luan, M. C. Liu, X. G. Chen and Z. D. Hu, (2008) "QSAR study of neuraminidase inhibitors based on heuristic method and radial basis function network", European Journal of Medicinal Chemistry 43, 569-576.

William Humphrey, Andrew Dalke and Klaus Schulte (February 1996), "VMD: Visual molecular dynamics," Journal of Molecular Graphics Volume 14 (1), 33-38

Xiaodong Wang, Allan W. Wolkoff, and Marilyn E. Morris, (2005) "Flavonoids as A Novel Class Of Human Organic Anion-Transporting Polypeptide Oatp1b1 (Oatp-C) Modulators", Drug Metabolism and Disposition 33, 1666-1672.

Xiaojin Xu, Xueyong Zhu, Raymond A. Dwek, James Stevens and Ian A. Wilson, (2008) "Structural Characterization of the 1918 Influenza Virus H1N1 Neuraminidase", Journal of Virology, Vol. 82, No. 21, 10493-10501.

Yogendra Patel, Valerie J. Gillet, Gianpaolo Bravi and Andrew R. Leach, (2002) "A comparison of the pharmacophore identification programs: Catalyst, DISCO and GASP", Journal of Computer-Aided Molecular Design 16, 653–681.

Yoshifumi Fukunishi and Haruki Nakamura, (2008) "Prediction of protein–ligand complex structure by docking software guided by other complex structures", Journal of Molecular Graphics and Modelling 26, 1030–1033.

Young Bae Ryu, Jang Hoon Kim, Su-Jin Park, Jong Sun Chang, Mun-Chual Rho, Ki-Hwan Bae, Ki Hun Park and Woo Song Lee, (2010) "Inhibition of neuraminidase activity by polyphenol compounds isolated from the roots of Glycyrrhiza uralensis", Bioorganic & Medicinal Chemistry Letters 20, 971-974.