Developing Database of Inhibitors against JNK Isotypes for analyzing Specificity of Fragments against these Targets

Submitted by

PREETI RANA - 131501

ANKITA GUPTA - 131515

JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY - WAKNAGHAT

MAY 2017

Developing Database of Inhibitors against JNK Isotypes for analyzing Specificity of Fragments against these Targets

A PROJECT REPORT

Submitted by

PREETI RANA - 131501

ANKITA GUPTA - 131515

in partial fulfillment for the award of the degree

of

BACHELOR OF ENGINEERING

in

BIOINFORMATICS

JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY - WAKNAGHAT

MAY 2017

JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY - WAKNAGHAT

BONAFIDE CERTIFICATE

Certified that this project report **"Developing Database of Inhibitors against JNK Isotypes for analyzing Specificity of Scaffolds against these Targets"** is the bonafide work of **"PREETI RANA & ANKITA GUPTA"** who carried out the project under my supervision.

Dr. Chittaranjan Rout **Date :**

Associate Professor

Department of Biotechnology and Bioinformatics

ACKNOWLEDGMENTS

Working on this project was a source of immense knowledge to us. We would like to express our sincere gratitude to **Dr. Chittaranjan Rout** and Research Scholars **Ms. Nupur Munjal, Mr. Deepak Mishra** for their guidance and valuable support. A huge thanks to Research Scholars **Mrs. Deepshikha Rathore** and **Ms. Babita** for validating our data. A special thanks of gratitude to **Mr. Shubham Singla** for his technical support. We acknowledge many helpful comments received from other faculty members for improving the material and methods of our project. This work would not have been without the guidance and help of several individuals who in one way or another contributed and extended their valuable assistance in the preparation and completion of this study.

We also want to acknowledge Assistant Professor **Dr. Ram Singh** and Department of Applied Chemistry, **Delhi Technological University** for their technical support.

Thank You.

TABLE OF CONTENTS

CHAPTER NO. **TITLE** **PAGE NO.**

SUMMARY

The JNK inhibitors against Alzhmeir disease (AD) should be efficient when they modulate the JNK activity for treatment of AD without affecting other cells. Though some inhibitors were reported but using for them for treatment lead to side-effects and toxicity. Further, the current JNK inhibitors inhibit all JNKs because of highly similarity between the isotypes which may lead to further complications. Thus, we need specific JNK inhibitors that provide both cell type and signal specificity. Present study is intent to develope a database of inhibitors (AlzID) active against AD promising drug targets JNK isotypes collected from published literature. This database contains over 650 molecules and their activity data (IC_{50} values) against three JNK enzymes. Optimized 3D geometries are provided to allow virtual screening. Geometry of each inhibitor is optimized using B3LYP (Becke's Lee Yang and Parr correlation) approach. The inhibitors were annotated with their molecular properties such molecular weight, LogD, LogP, asymmetric atoms, number of rotatable bonds etc. Other information such as SMILES, IUPAC name, etc. were also provided. To determine common and specific fragments, each inhibtors were fragmented on the basis of Bemis Murcko method *i.e.* ring, side-chain, linker and framework (to find which fragment exist and how frequent are they). We identified the fragments that were occurring more than random in a dataset. Based on the frequencies of fragments in dataset, we identified common and unique fragments for JNK isotypes. The inhibitors data were provided with several common file formats including SMILES, SDF and mol2. A molecular drawing interface (JME) and R-Package 'rcdk' was incorporated into the database to facilitate searching of molecules on the basis of similarity. Text-based query is also available. Access and retrieval of data through similarity based searching and text-based method are available. This database also allow user to upload/draw desired molecules for searching. User can also get library of molecules for virtual screening that are specific against a particular JNK isotype.

CHAPTER - 1

INTRODUCTION

Alzheimer's disease is an irreversible, progressive brain disorder that slowly destroys memory and thinking skills, and eventually the ability to carry out the simplest tasks Alzheimer's is the most common cause of dementia among older adults**.** More than 4 million people in India have some form of dementia. Worldwide, at least 44 million people are living with dementia, making the disease a global health crisis. It is not a normal part of aging, although the greatest known risk factor is increasing age, and the majority of people with Alzheimer's are 65 and older. This is expected to double by 2030. Despite the magnitude, there is gross ignorance, neglect and services are scarce for people with dementia and their families.

1.1 Mechanism of Alzheimer Disease (AD)

AD is a neurological disease which causes death of brain cells that ultimately lead to memory loss and cognitive decline. The disease is characterized by cognitive impairment, progressive neurodegeneration and formation of amyloid-beta containing plaques & neurofibrillary tangles composed of hyperphosphorylated tau. The disease can initially be characterized by synaptic damage accompanied by neuronal loss.

The main thing lead to Alzheimer is the formation of a peptide known as **Amyloid beta** which clusters into **Amyloid Plaques** that leads to killing of neurons which further causes dementia (Fig 1.1).

There are two forms of amyloid beta peptides :- 40 and 42 amino acids. Enzymes that cleave Amyloid precursor protein (APP) are called Secretases. **Alpha-secretase** and **Beta-secretase** are two enzymes that initially compete to cleave the APP. Amyloid-beta formation does not take place if a-secretase cleaves APP. But if beta-secretase cleave APP then it can further be cleaved by gamma-secretase to form either a 40 amino acid Amyloid Peptide ($Aß_{40}$) which is soluble & mostly innocuous or a 42 amino acid peptide ($Aß_{42}$) which clumps together to form insoluble Amyloid Plaques.

Processes playing important role in causing death of neurons during Amyloid Plaque Formation as follows :

I. **Infammatory & oxidative damage**
II. **Neuro-fibrillary tangles**

Astrocytes and **Microglia** are two major types of brain cells that participate in inflammatory response. The former become more numerous in AD and produce Prostaglandin acid after activation which mediate inflammatory response. Microglial cells after activation produces damaging free radicals. Activities of both brain cells cause death of neurons.

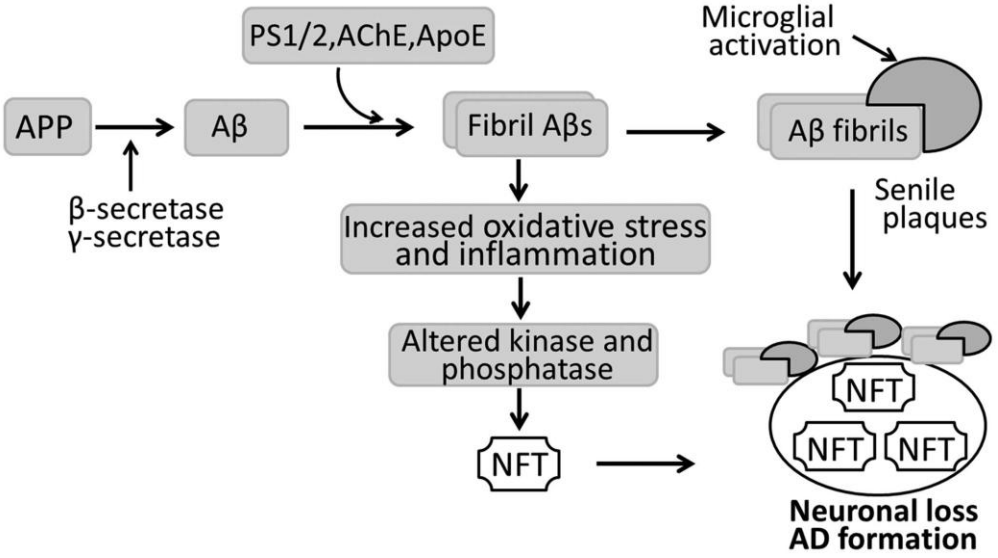


Fig. 1.1: Disease process of AD Formation of neuro-fibrillary tangles (NFTs) leads to memory loss

* Aß (Amyloid beta), *AChE (Acetylcholinestrase), *PS1/2 (Preseniline1/2), *ApoE (Apolipoprotein E), *NFT (Neuro-fibrillary tangles)

Amyloid precursor protein -> Fibrillar Amyloid-Beta and Oligomers -> Amyloid Plaques -> Inflammation & Neuro-Fibrillary Tangles -> Neuron Death and Synapse Loss

1.2 Treatment

AD is a complex disease and not one drug or single intervention can treat - it successfully. Current approaches are focusing on maintaining mental function, behavioural and slowing down the symptoms of AD. The following are few FDA approved drugs used for the treatment of AD :

FDA -Approved Drugs for Alzheimer's Disease

Drug	**Class & Indication**	**Mechanism of Action**	**Common Adverse Effects**
Donopezil	Acetyl Cholinesterase (Ach) inhibitor recommended to treat symptoms of mild to moderate and moderate to severe AD	Halts the breakdown of ACh in brain	Vomiting, Nausea , Diarrhoea;
Galantamine	ACh inhibitor prescribed to treat symptoms of mild-to-moderate AD	Halts the breakdown of ACh and thus release more acetylcholine in brain after stimulating nicotinic receptors	Nausea , Diarrhoea , vomiting , loss of appetite, weight loss;
Memantine	NMDA antagonist recommended to cure manifestation of moderate-to- severe of AD	Prevents the toxicity associated with excess glutamate and regulates its activation	Dizziness, confusion , headache, constipation;
Rivastigmine	Cholinesterase inhibitor recommended to cure manifestation of mild-to-moderate AD	Checks the breakdown of ACh and BCh in brain	Nausea , Diarrhoea , vomiting , loss of appetite, weight loss , Muscle weakness;

As given above these drugs have enough side-effects and sometimes these side-effects lead to other diseases.

1.3 Drugs in Research, Targets and their Shortcomings

The followings are few drugs that are in different stages of clinical trials. But they also produce major side-effects.

Drug	**Mechanism of Action**	**Stage of Development**	**Common Adverse Effects**
Solanezumab	Beta-Amyloid	Phase – 3	Nasopharnygitis, Diarrhea and increased blood Creatine Phosphokinase
Verubecestat	Beta-Secretase	Phase – 3	Signs of Liver Toxicity, Changes in Blood Glucose, Fur Depigmentation
AADvac1	Tau	Phase – 2	Allergy
CSP-1103	Beta-Amyloid, Inflammation	Phase - 2	Mild Diarrhea
Intepirdine	AChE Neurotransmitter	Phase – 3	Acute Toxicity

AChE: is the primary cholinesterase in the body. It is an enzyme that catalyzes the breakdown of acetylcholine and of some other choline esters that function as neurotransmitters.

BChE: is a nonspecific cholinesterase enzyme that hydrolyses many different choline-based esters. In humans, it is made in the liver - and found mainly in blood plasma.

1.4 JNKs are Considered Better Targets

The c-Jun N-terminal kinases belong to sub-family of MAPK (Mitogen-Activated Protein Kinase). Activation of JNK pathways causes natural death of cells during development as well as for pathological death that is associated with neurodegeneration disease. Side-effects of Alzheimer drugs are reported as major obstacle in the path of successful treatment. The c-Jun N-terminal kinases (JNKs) have major role in stress signalling pathways and the JNK3-isotype is expressed mainly in neuronal tissue. The JNK3 enzyme is highly expressed in post-mortem brains of individuals that suffered from AD. This enzyme has been found to play an upstream role in neuronal ischemic apoptosis. The loss of JNK3 protects the adult brain from glutamate-induced excitotoxicity and therefore, this enzyme is a promising drug target. However drugs targeting JNK3, due to similarity of this enzyme with other JNK-isotypes, are expected to produce side-effects. Consequently, the development of selective JNK3 inhibitors represents a useful approach may bring better treatment outcomes. Therefore, this resource has been developed to provide unique and common active fragments against drug targets of AD.
The selective expression of JNK3 in the brain, and the findings report that JNK3 knockout mice exhibit amelioration of neurodegeneration in animal models of AD, suggest that the inhibition of this isoform would be a promising therapeutic option.

Uncontrolled proliferation, cellular growth and relocation along with deregulated angiogenesis causes generation of malignant tumors. JNK signal reduction pathway may not act exclusively in apoptosis, sustained JNK activation leading to activation of AP1. This activation has recently been implicated to contribute to cellular survival of specific cancer types such as glial tumors and BCL-ABL transformed B-lymphoblasts.

In glial tumors, over-expressed JNK/AP1 activity was seen in most of tumor samples of primary brain. For transformed B lymphocytes, BCL-ABL was exhibit to activate the

JNK pathway which further up-regulate the expression of anti-apiptotic bcl-2 gene.

JNK isoforms have the following tissue distributions:

I. **JNK1** and **JNK2** are found in all cells and tissues.

II. **JNK3** is mainly expressed in the brain, but is also expressed in very less amount in the heart and the testis.

Therefore, highly selective drug developed against JNK3 can only bind to JNK3 expressed in brain. The selective drugs have lesser chance to interact with other proteins present in different tissues.

In brain tissue and cerebrospinal fluid collected from AD patient, JNK3 is highly expressed and activated. It is a crucial kinase for phosphorylation of Beta- APP at T668. Genetic depletion of JNK3 resulted in reduction of $Aß_{42}$ peptide level which leads to increase in number of neurons and improved cognition.

JNK3 activation is integrated with the increased levels of NFT's and senile plaques - (Fig. 1.1). This enzyme directly modulates the formation of NFTs by immediate phosphorylation of Tau. It has highest affinity toward phosphorylation at Ser, thus it can strongly autophosphorylate itself and contributes to hyperphosphorylation of Tau.

1.5 Difference between CNS and Non-CNS Drugs

Molecular Properties	**CNS**	**Non-CNS**
MW	319 (151–655)	330 (163–671)
ClogP	3.43- (0.16-6.59)	2.78- (-2.81-6.09)
ClogD	2.08 (−1.34–6.57	1.07 (−2.81–5.53)
H bond donors	0.85- (0-3)	1.56- (0-6)
H bond acceptors	3.56 (1–10)	4.51 (1–11)
Rotatable Bonds	1.27- (0-5)	2.18- (0-4)
Aromatic rings	1.92 (0–4)	1.93 (0–4)
Molecular Volume ($Å^3$)	800-1000	1000-1200
Topological Polar Surface Area ($Å^2$)	<76 (25-60)	>80 (80-140)

Table 1.5.1 : Difference in molecular properties of CNS & Non-CNS drugs

In contrast to other drugs, the CNS drugs should posses some unique characterstics due to blood brain barrier (BBB). For a CNS drug to achieve optimum therapeutic efficiency, it should possess high degree of potency and selectivity for interaction with the target.

1.6 JNKs Require Specificity

Although several JNK inhibitors have entered clinical trials, the initial pace of advancement and potential success seems to be limited for a variety of reasons. The fundamental reason is that these inhibitors cannot directly serve the purpose of modulating JNK activity for treatment of AD without affecting other cell types, as the JNKs mediate several signalling pathways in multiple cell types, thus, use of these compounds for treatment can lead to other side effects and toxicity. Further, the current JNK inhibitors inhibit all JNKs because of highly similarity between the isotypes which may lead to further complications (Fig. 1.6.1). Therefore, generation of JNK1, JNK2 and JNK3 specific inhibitors are required to trigger the problem of specificity. Thus we need specific JNK inhibitors that provide both cell type and signal specificity. This can be generated only if the specific interacting partners affecting specific cellular processes in disease are clearly defined. Hence, specificity is a critical issue that needs proper evaluation for successful JNK inhibition therapies.

The JNKs consist of ten isoforms obtained from three genes: JNK1 (four isoforms), JNK2 (four isoforms) and JNK3 (two isoforms). As there is highly similarity between sequences of JNK1, JNK2 and JNK3 (Fig. 1.2), therefore leads/drugs developed against JNK3 should be very specific (active against JNK3 at the same time inactive towards JNK1 & JNK2) require specificity to avoid side-effects of drugs. Because due to highly similar sequence there is highest probability of a drug to bind any of these targets which causes adverse side-effects.

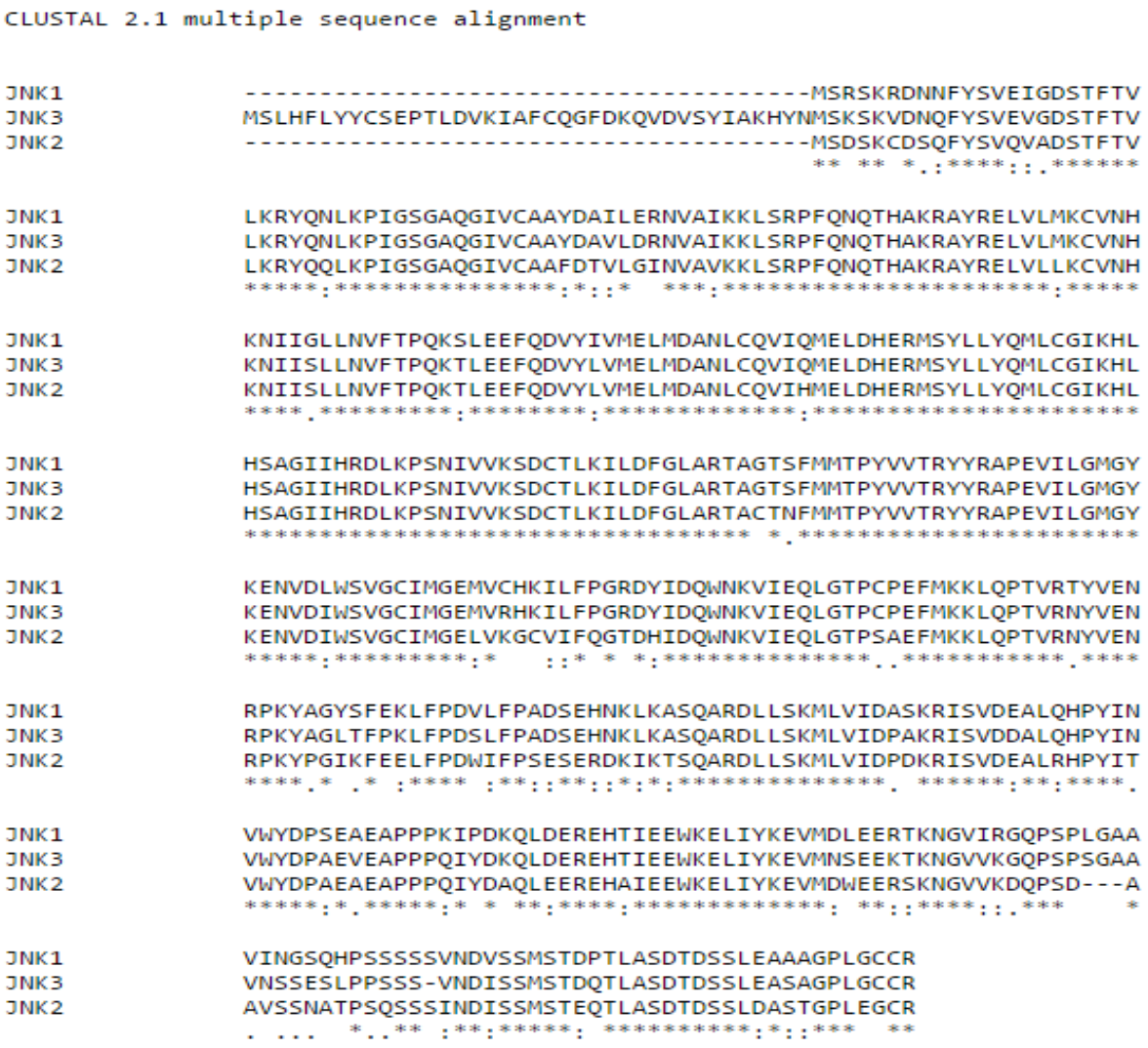


Fig. 1.6.1 : Multiple sequence alignment (MSA) of three JNK sequences determined using ClustalW. Important residues in JNK3 not common with JNK1 and JNK2 may be used for specific-JNK3 drug development

1.7 Scaffold Identification and its Impact on Drug Design

Scaffold is considered as a molecular fragment without any side-chain and a side chain is any acyclic chain or functional group with a single link point to remaining molecule. We can isolate scaffold and side-chains by detecting ring structure first. Starting are removed from the structure and retained as side-chains whereas remaining structure is stored as a scaffold.

Scaffold identification is process of finding relatively simple, often weakly potent, bioactive molecules that are **"ligand efficient"** i.e. most likely to bind a target. They possess a high binding affinity per heavy atoms and thus are ideal for optimization into clinical candidates with good drug-like properties. Because of the smaller and less complex nature of scaffolds, they increase the possibility of finding a match to the receptor. They have higher ligand efficiency thus provide greater scope for development when following a standard chemistry development strategy.

The four databases ACD, NCI, CMC and MDDR containing lead-like molecules were examined according to scaffold-based classification (SCA). The ACD was most diverse, followed by NCI, then CMC and finally MDDR. The objectives were to determine which fragments exist, how reccuring these are, how their occurrence related to one another and non-overlapping fragment in existing (drug) compound databases. Some scaffolds are found in pair in active compounds. Some scaffolds are not found in same compound. Two scaffolds may have same physiochemical properties therefore, they may not found in same molecule.

Some scaffolds occur more uncommonly (e.g. a phenyl ring) or chemical forms on which some drug classes are based. Some scaffolds have low occurence that might indicate rarely occuring parts of chemical space, potentially gripping for designing new compounds. Insights may provide preferences. Scaffolds that do not occur together, new chemical space can be analyzed. Their co-occurrences may be used to identify a replacement for a structural feature. Scaffold pairs barely occur together, possibly because of their similar physicochemical properties.

Fragment based screening have important role to find novel and patentable scaffolds. The core scaffold capable of binding to several target proteins may then further be optimized to a compound with appropriately balanced affinities between the target proteins.

We can group compounds into the same class if they have same topological scaffold in common. The aim backing for this was that chemists instinctively organize compounds based on scaffolds and functional groups.

1.8 Objectives

Keeping view of complexity of the disease and reported adverse side-effects of FDA approved drugs, the proposed study intend to use "Chemoinformatics" approach to suggest novel compounds against JNK3. So the followings are the objectives of the proposed work.

I. Developing database of inhibitors with their molecular data against JNK isotypes. 3D-structures of the compounds will also be provided to assist virtual screening.

II. To identify the fragments that are specific to our target and also the combination of scaffolds that are specific to one or more targets to be used for developing the most efficient drug against Alzheimer.

CHAPTER – 2

MATERIAL AND METHODS

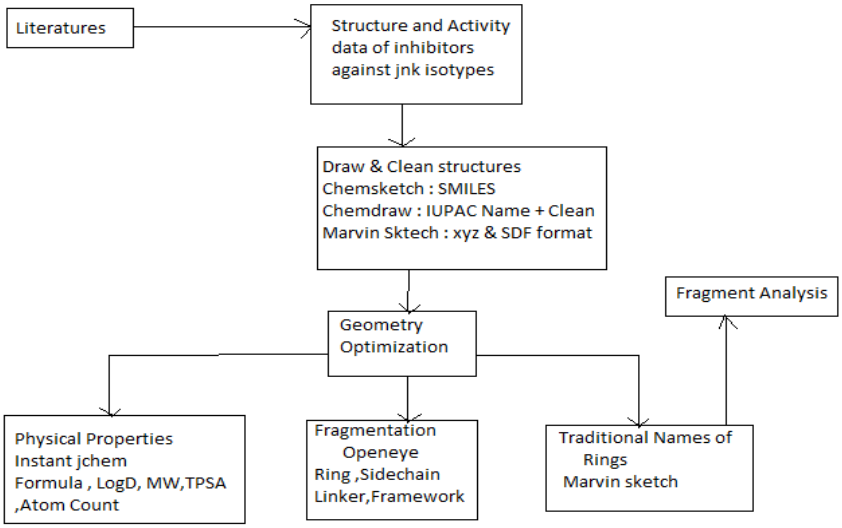
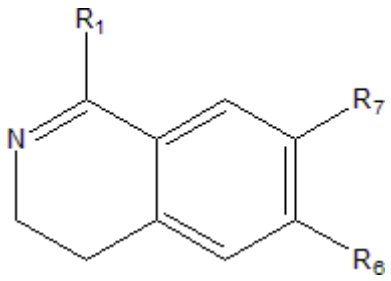


Fig. 2.1 : Flow diagram of proposed work

2.1 Compounds

Approximately 700 Inhibitor structures and their inhibition activity against the AD drug targets JNK-isotypes and/or MAO protein(s) are collected from the literature. Experimental inhibition activities are reported in around 30 published papers. The structure of these molecules and their activity data are given below :

Table 2.1.1 : 1-Aryl-3, 4-dihydroisoquinoline inhibitors of JNK3 and p38a of compounds 52–142001z, values in pIC50

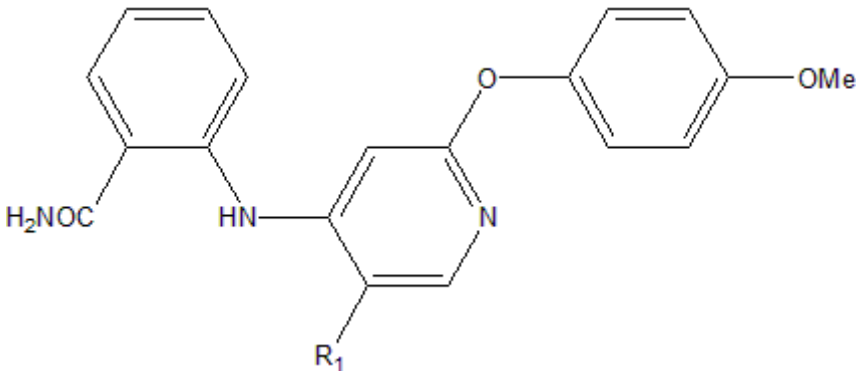


AlzID	R1	R6	R7	JNK3	p38(alpha)
52001z	3,4-Cl2-Phenyl	OMe	OMe	5.2	<4.8
62001z	Phenyl	OMe	OMe	<4.8	<4.8a
72001z	4-Cl-Phenyl	OMe	OMe	<4.8	<4.8
82001z	3-MeO-Phenyl	OMe	OMe	<4.8	<4.8
92001z	3-Cl-Phenyl	Cl	OMe	4.9	<4.8a
102001z	3-Cl-Phenyl	H	H	<4.8	<4.8
112001z	3-F-Phenyl	OMe	Cl	5.4	<4.8
122001z	3-F-Phenyl	Cl	OMe	<4.8	<4.8a
132001z	4-F-Phenyl	OMe	Cl	5.3	<4.8
142001z	4-F-Phenyl	Cl	OMe	<4.8	<4.8a

AlzID	R1	R6	R7	JNK3	p38a
163001z	3-Br-Phenyl	OMe	Cl	6.4	4.9
173001z	3-Br-Phenyl	Cl	OMe	5	<4.8a
183001z	2-Napthyl	OMe	Cl	6	<4.8
193001z	2-Napthyl	Cl	OMe	<4.8	<4.8a
203001z	3,4-Cl2Phenyl	OMe	Cl	6.5	5.1
213001z	3,4-Cl2Phenyl	Cl	OMe	4.8	<4.8a
223001z	Phenyl	OMe	Cl	5.2	<4.8
233001z	Phenyl	Cl	OMe	<4.8	<4.8a
243001z	3-Cl-Phenyl	OMe	Cl	6.6	5.1
253001z	3-Me-Phenyl	OMe	Cl	5.7	<4.8
263001z	2-F,3-Cl-Phenyl	OMe	Cl	6.2	4.9
273001z	-CH2Phenyl	OMe	Cl	<4.8	<4.8
283001z	Ethyl	OMe	Cl	<4.8	<4.8

AlzID	JNK1a	JNK2a	JNK3a	Erk-2a	p38ab
204001z	4	6.1	7.3	<4.0	5.1
244001z	<4.0	5.9	6.9	<4.0	5.1

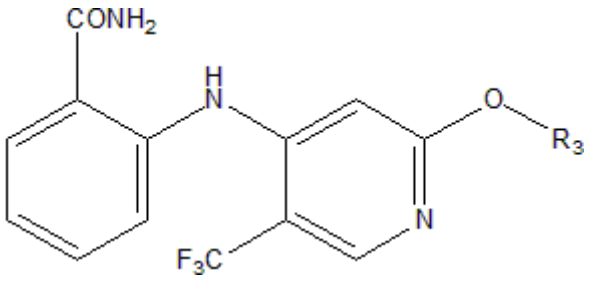
Table 2.1.2 : 5-Substituted 2-phenoxypyridines as JNK inhibitors



AlzID	R1	JNK1_IC50	JNK3_IC50
9a1002z	F	0.026	0.048
9b1002z	Cl	0.008	0.015
9c1002z	CF3	0.027	0.037

(Song, Xinyi, et al. "Synthesis and SAR of 2-Phenoxypyridines as novel c-Jun N-terminal kinase inhibitors." *Bioorganic & medicinal chemistry letters* 21.23 (2011): 7072-7075)

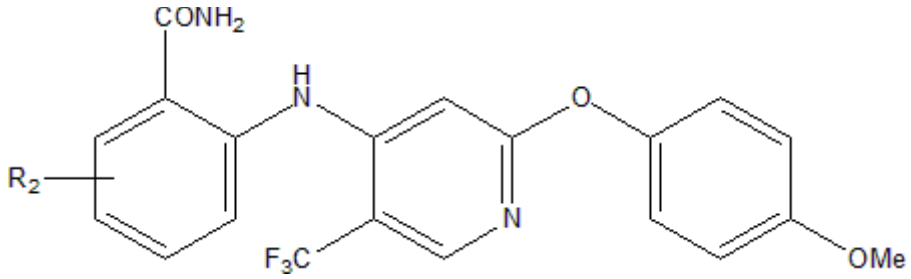
Table 2.1.3 : 2-Phenoxypyridine SAR as JNK1 and JNK3 inhibitors



AlzID	R3	JNK1	JNK3
9c2002z	4-MeOPh	0.027	0.037
10a2002z	3-MeOPh	0.038	0.07
10b2002z	2-MeOPh	3.2	3.5
10c2002z	4-CF3OPh	0.035	0.035
10d2002z	4-EtOPh	0.035	0.11
10e2002z	4-n-PrOPh	0.1	0.14
10f2002z	4-t-BuOPh	0.022	0.027
10g2002z	4-FPh	0.1	0.2
10h2002z	4-BrPh	0.48	0.38
10i2002z	3-ClPh	NT	0.2
10j2002z	3,4-ClPh	0.68	0.1
10k2002z	2-Cl,4-MeOPh	0.74	0.1
10l2002z	1-Naphthyl	0.65	0.99
10m2002z	2-Naphthyl	NT	0.32
10n2002z	3,4-MethylenedioxyPh	0.044	0.048
10o2002z	4-Triazole-Ph	NT	0.08
10p2002z	4-Piperazine-Ph	NT	0.034
10q2002z	3-Pyridyl-4-Ph	0.053	0.072

(Song, Xinyi, et al. "Synthesis and SAR of 2-Phenoxypyridines as novel c-Jun N-terminal kinase inhibitors." *Bioorganic & medicinal chemistry letters* 21.23 (2011): 7072-7075)

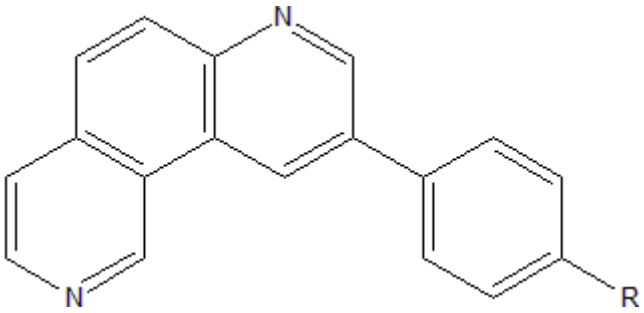
Table 2.1.4 : 2-Aminobenzamide ring SAR as JNK3 inhibitors



AlzID	R2	JNK3
9c3002z	H	0.037
11a3002z	2-Cl	2.2
11b3002z	2-OMe	2.2
11c3002z	3-F	0.045
11d3002z	3-Cl	0.03
11e3002z	4-F	0.049
11f3002z	4-Cl	0.078
11g3002z	4-OMe	0.017
11h3002z	5-F	0.06
11i3002z	5-Cl	0.16

(Song, Xinyi, et al. "Synthesis and SAR of 2-Phenoxypyridines as novel c-Jun N-terminal kinase inhibitors." *Bioorganic & medicinal chemistry letters* 21.23 (2011): 7072-7075)

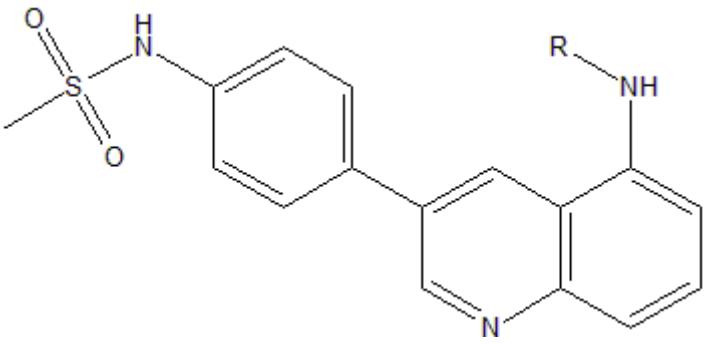
Table 2.1.5 : JNK3 inhibition by Phenanthroline derivatives 1003z

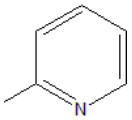
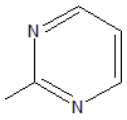
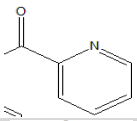
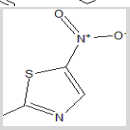
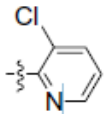


AlzID	R	JNK3	p38
11003z	OH	0.59	>20
1a1003z	H	1 ± 0.17	>20
1b1003z	Phenyl	>20	Nt
1c1003z	5-Pyrazolyl	3.6 ± 0.68	Nt
1d1003z	Morpholino	1.3 ± 0.22	Nt
1e1003z	NH2	0.51 ± 0.03	>20
1f1003z	NAc2	1.6 ± 0.27	Nt
1g1003z	NHAc	0.53 ± 0.08	>20
1h1003z	NHMs	0.93 ± 0.1	>20
1i1003z	NMs2	0.21 ± 0.02	>20

(Jiang, Rong, et al. "3, 5-Disubstituted quinolines as novel c-Jun N-terminal kinase inhibitors." *Bioorganic & medicinal chemistry letters* 17.22 (2007): 6378-6382)

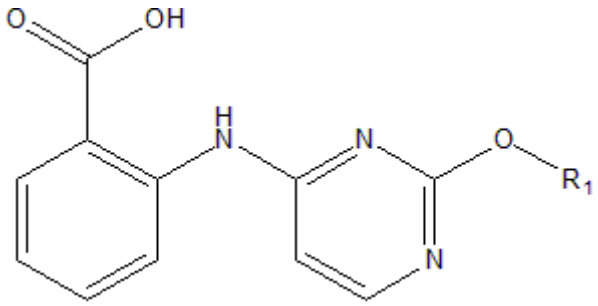
Table 2.1.6 : JNK3 inhibition by Phenanthroline derivatives 10003z



Compound	R	JNK3	p38
10a2003z		0.44 ± 0.09	>20
10b2003z		0.48 ± 0.06	>20
10c2003z		3.6 ± 0.32	Nt
10d2003z		0.76 ± 0.15	Nt
10e2003z		15 ± 0.17	Nt

(Jiang, Rong, et al. "3, 5-Disubstituted quinolines as novel c-Jun N-terminal kinase inhibitors." *Bioorganic & medicinal chemistry letters* 17.22 (2007): 6378-6382)

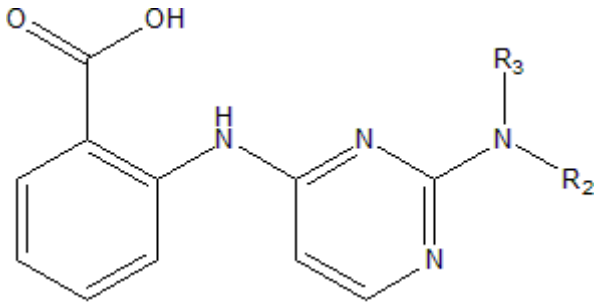
Table 2.1.6 : SAR summary of JNK inhibitors 1004z and 6004z



AlzID	R1	JNK1
11004z	Bu	1.9
6a1004z	Et	44.7
6b1004z	Pr	2.3
6c1004z	Pentyl	1.1
6d1004z	Hex	>100
6e1004z	i-Bu	0.7
6f1004z	CH2-c-Hex	0.7
6g1004z	Ph	0.3

(Liu, Mei, et al. "Discovery of a new class of 4-anilinopyrimidines as potent c-Jun N-terminal kinase inhibitors: synthesis and SAR studies." *Bioorganic & medicinal chemistry letters* 17.3 (2007): 668-672)

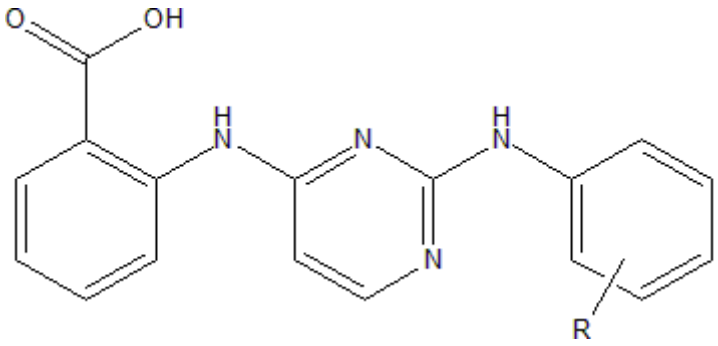
Table 2.1.7 : SAR summary of JNK inhibitors 1004z, 7004z



AlzID	R2	R3	JNK1
12004z	—	—	1.9
7a2004z	Bu	H	2.6
7b2004z	Bn	H	18.2
7c2004z	c-Hex	H	0.4
7d2004z	–(CH2)6–		15.4
7e2004z	Et	Et	>100
8a2004z	—	—	0.048

(Liu, Mei, et al. "Discovery of a new class of 4-anilinopyrimidines as potent c-Jun N-terminal kinase inhibitors: synthesis and SAR studies." *Bioorganic & medicinal chemistry letters* 17.3 (2007): 668-672)

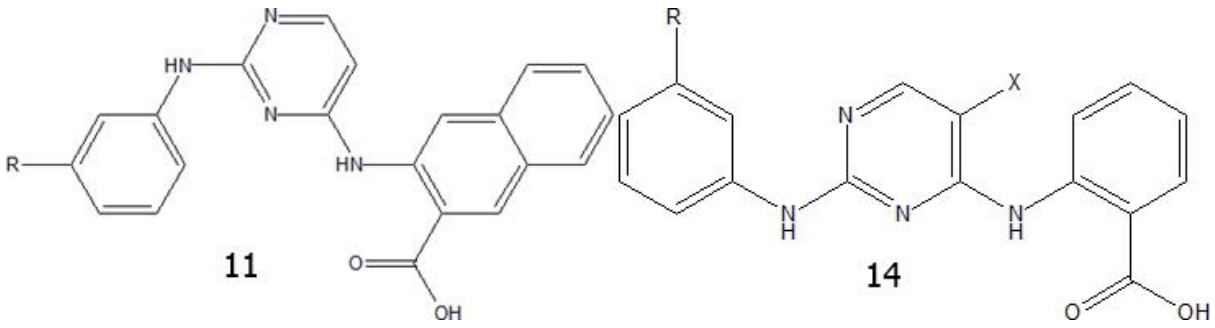
Table 2.1.8 : JNK inhibitor's SAR summary 8004z with 2-anilinosubstitutions



AlzID	R	JNK1(nM)
8a3004z	H	48
8b3004z	2-OH	76
8c3004z	3-OH	25
8d3004z	4-OH	35
8e3004z	2-F	35
8f3004z	3-F	28
8g3004z	4-F	29
8h3004z	3-Me	157
8i3004z	3-F,4-Me	93
8j3004z	2-Me,3-OH	278
8k3004z	3-CF3	391
8l3004z	4-CF3	186
8m3004z	4-NO2	33
8n3004z	4-Morpholine	37

(Liu, Mei, et al. "Discovery of a new class of 4-anilinopyrimidines as potent c-Jun N-terminal kinase inhibitors: synthesis and SAR studies." *Bioorganic & medicinal chemistry letters* 17.3 (2007): 668-672)

Table 2.1.9 : SAR summary for JNK inhibitors 11004z and 14004z



AlzID	R	X	JNK1
11a4004z	H	—	705
11b4004z	OH	—	204
14a4004z	H	F	82
14b4004z	OH	F	21
14c4004z	H	Br	32
14d4004z	OH	Br	20

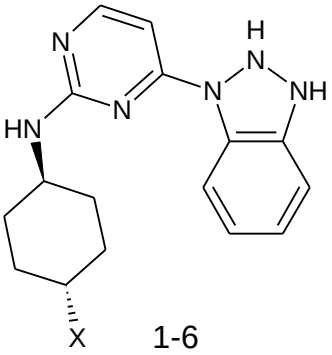
(Liu, Mei, et al. "Discovery of a new class of 4-anilinopyrimidines as potent c-Jun N-terminal kinase inhibitors: synthesis and SAR studies." *Bioorganic & medicinal chemistry letters* 17.3 (2007): 668-672)

Table 2.1.10 : Kinase selectivity profile of JNK1 inhibitor 2ba004z

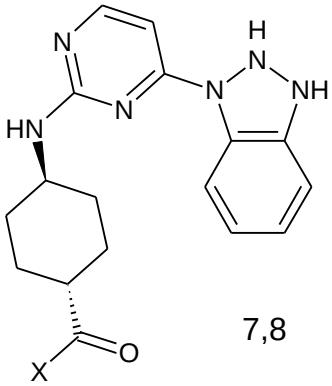
Kinases	2b004z	Kinases	2a004z
JNK1	0.009	PLK	0.73
P38	>50	CK2	0.73
ERK2	25	MEK	6.8
AKT1	15	CDK2	2.7
CHK1	0.82	MK2	>50
PAK4	5.5	COT2	>50

(Palmer, Wylie S., et al. "Development of amino-pyrimidine inhibitors of c-Jun N-terminal kinase (JNK): Kinase profiling guided optimization of a 1, 2, 3-benzotriazole lead." *Bioorganic & medicinal chemistry letters* 23.5 (2013): 1486-1492)

Table 2.1.11 : Biological effect of benzotriazoles20 with cyclohexyl-amine substituents



1-6



7,8

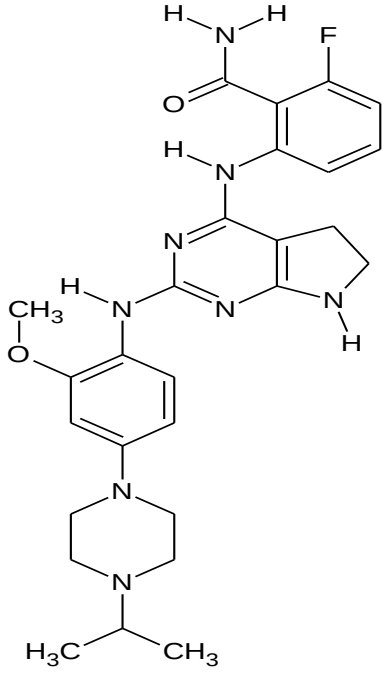
Compound	X	IC50(IM)				
		JNK1	JNK2	CDK2	c-Junb	HCT116c
11007z	NHSO2Me	0.024	0.097	0.3	1.3	2.3
21007z	H	0.1	0.39	2.7	nt	nt
31007z	OH	0.029	0.11	0.43	0.79	1.2
41007z	NH2	0.074	0.17	0.36	0.54	0.77
51007z	NHSO2NMe2	0.019	0.076	1.7	4.9	22
61007z	NH(C@O)CH3	0.039	0.23	0.57	1.5	16
71007z	1-Pyrrolidinyl	0.12	0.4	>6.2	3.3	>30
81007z	1-Morpholino	0.063	0.18	2	1.8	20

Table 2.1.12 : Biological activity of 4-alkoxy-substituted indazoles (18a–g007z, 19007z) and indole (21007z)

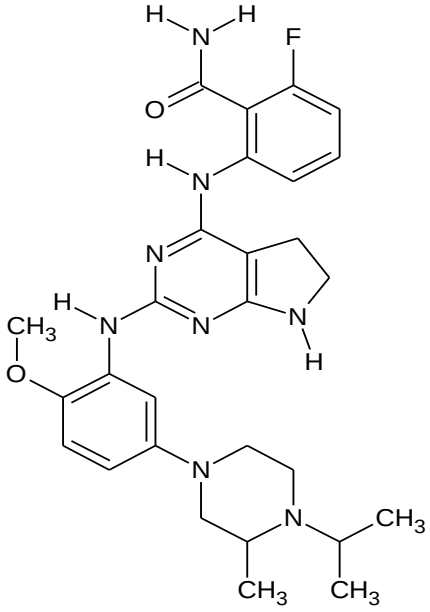
Compound	R-group	JNK1/2 (IC50(nM))	JNK2(Fold-shift)	c-Junc IC50

18a007z	CH3	23/178	37	3.1
18c007z	CH2CH2CH2OH	24/139	28	0.96
18d007z	CH2CH(OH)CH2OH	34/194	5.7	2
18e007z	CH2CH2CH2NHSO2CH3	56/246	17	1.1
18f007z	CH2CH2SO2CH3	33/151	4	0.98
18g007z	CH2CH2CH2SO2CH3	6.2/24	23	1

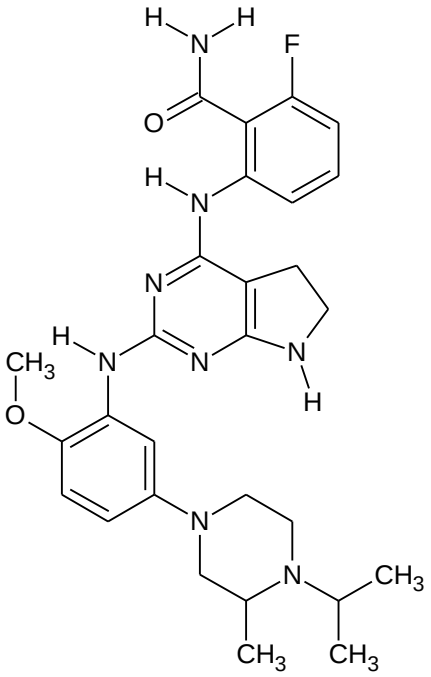
Table 2.1.13 : IGF-1R, JNK1, and JNK3 potencies for 1–3008z



1



2



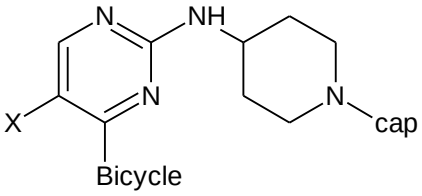
3

Compound	IGF-1R enzyme	JNK1 Enzyme IC50	JNK3 IC50	Phospho IGF-1R cellular
11008z	2	13	100	117
21008z	4	3162	5011	201
31008z	20	3891	6310	270

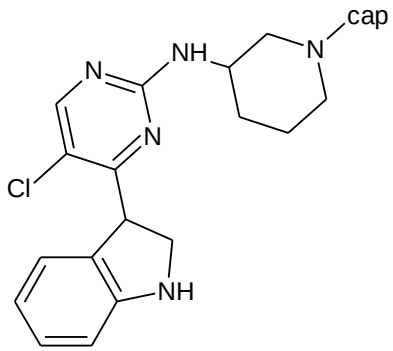
AlzID	5`X	Bicycle	Heterocycle	Cap	JNK1 IC50 (nm)	JNK2	JNK3	CDK2	c-Jun
51009z	CN	6-F-3-Indole	4-Piperidine	CONHEt	92	67	412	412	3700
61009z	Cl	3-Indole	4-Piperidine	CONHEt	13	25	57	1517	704
71009z	Me	3-Indole	4-Piperidine	CONHEt	320	250	410	IA	10,000
81009z	H	3-Indole	4-Piperidine	CONHEt	74	245	na	10,000	8091
91009z	Cl	3-Imidazopyridine	4-Piperidine	CONHEt	41	55	na	605	7723
101009z	Cl	1-Indole	4-Piperidine	CONHEt	457	709	na	4443	>10,000
111009z	H	3-Imidazopyridine	4-Piperidine	CONHEt	59	281	708	4219	19,331
121009z	H	1-Indole	4-Piperidine	CONHEt	71	512	na	10,000	29,891
131009z	H	Pyrazolopyridine	4-Piperidine	CONHEt	69	194	na	8663	4000
141009z	Cl	3-Indole	3-Pyrrolidine	CONHEt	360	177	582	2483	6268
151009z	Cl	3-Indole	Azetidine	CONHEt	1340	1551	5000	na	4272
161009z	Cl	3-Indole	3-(S)-Piperidine	CONHEt	29	15	32	555	6995
171009z	Cl	3-Indole	3-(S)-Piperidine	CH2CONHMe	139	267	na	>10,000	14,667
181009z	Cl	3-Indole	3-(R)-Piperidine	CONHEt	60	88	107	1264	3883
191009z	Cl	3-Indole	3-(R)-Piperidine	CH2CONHMe	15	31	31	612	2807
201009z	Cl	3-Indole	4-Piperidine	CH2CONHMe	13	22	14	123	1769
211009z	Cl	3-Indole	4-Piperidine	COOEt	37	49	82	na	na
221009z	Cl	3-Indole	4-Piperidine	CONMe2	15	37	na	4358	741
231009z	Cl	3-Indole	4-Piperidine	CO-(4Mepiperazine)	18	26	na	5281	2770
241009z	Cl	3-Indole	4-Piperidine	COCH2NHCOMe	28	46	na	551	1938
251009z	Cl	3-Indole	4-Piperidine	COCH2NHMe	67	85	179	2672	2028
261009z	Cl	3-Indole	4-Piperidine	COCH2NMe2	57	45	120	1126	2497
271009z	Cl	3-Indole	4-Piperidine	CO-(4Mepiperidine)	15	14	48	1895	813
281009z	Cl	3-Indole	4-Piperidine	CONH-(4-Mepiperidine)	47	62	na	6652	807

(Chamberlain, Stanley D., et al. "Optimization of 4, 6-bis-anilino-1H-pyrrolo [2, 3-d] pyrimidine IGF-1R tyrosine kinase inhibitors towards JNK selectivity." *Bioorganic & medicinal chemistry letters* 19.2 (2009): 360-364)

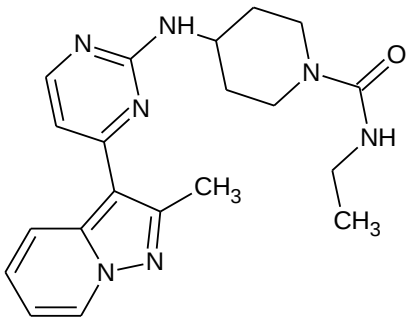
Table 2.1.14 : Enzymatic and cellular activity of aminopyrimidine analogue



5-12,20-28



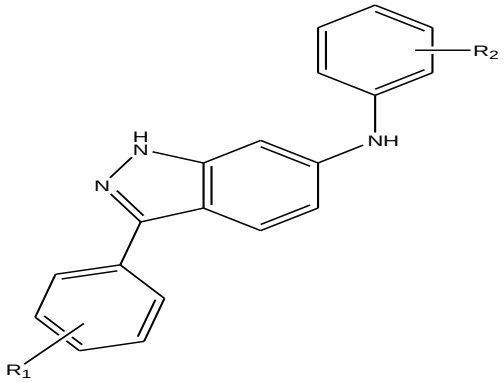
16-19



(13) , (29)-(30)

AlzID	5`X	Bicycle	R	JNK1 IC50 (nm)	JNK2	JNK3	CDK2	c-Jun
62009z	Cl	3-Indole	H	13	25	57	1517	704
132009z	H	Pyrazolopyridine	H	69	194	na	8663	4000
292009z	H	Pyrazolopyridine	iPr	520	698	na	>10,000	21,160
302009z	H	Pyrazolopyridine	Ph	22	5	5	>10,000	3845

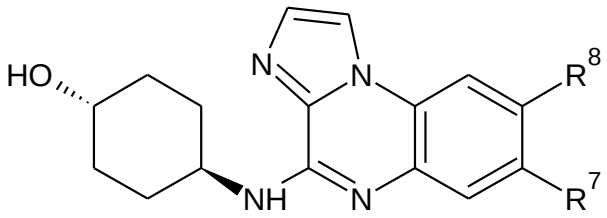
Table 2.1.15 : IC50 values for compounds 4a–j010z against JNK3, JNK1, and p38a24



AlzID	R1	R2	JNK3 IC50 (nM)	JNK1	p38a
4a1010z	H	H	48	>10,000	30
4b1010z	H	2-Cl	30	5350	18
4c1010z	H	2-OMe	202	1600	46
4d1010z	4-COOH	2-Cl	32	246	13
4e1010z	4-CONH(CH2)2N(CH3)2	2-Cl	1.9	45	8.7
4f1010z	3-CONH2	2-Cl	3.3	81	3.2
4g1010z	3-COOH	2-Cl	5.3	61	24
4h1010z	3-CONH(CH2)3-4-morfolinyl	2-Cl	3.4	228	3.7
4i1010z	3-CONH(CH2)2N(CH3)3	2-Cl	21	698	16
4j1010z	3-CONH-4-piperidinyl	2-Cl	1.4	71	4.4

(Swahn, Britt-Marie, et al. "Design and synthesis of 6-anilinoindazoles as selective inhibitors of c-Jun N-terminal kinase-3." *Bioorganic & medicinal chemistry letters* 15.22 (2005): 5095-5099)

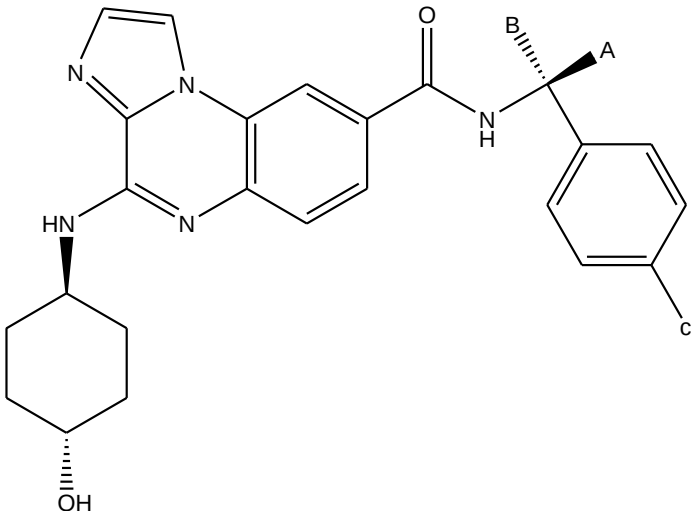
Table 2.1.16 : rhJNK1 IC50 values of 7- and 8-carboxy and carboxamide derivatives of compound 15-18011z



AlzID	R7	R8	rhJNK1
15a3011z	H	COOH	0.31
163011z	COOH	H	1.5
17a3011z	H	CONH2	0.57
183011z	CONH2	H	4.17

(Li, Bei, et al. "Hit-to-lead optimization and kinase selectivity of imidazo [1, 2-a] quinoxalin-4-amine derived JNK1 inhibitors." *Bioorganic & medicinal chemistry letters* 23.18 (2013): 5217-5222)

Table 2.1.17 : rhJNK1 IC50 values of 8-carboxamide derivatives of compound 23-41011z



AlzID	A	B	c	rhJNK1(IC50)
23a4011z	Me	H	H	1.98
19a4011z	H	Me	H	0.58
244011z	Et	H	H	1.7
25a4011z	H	Et	H	0.14

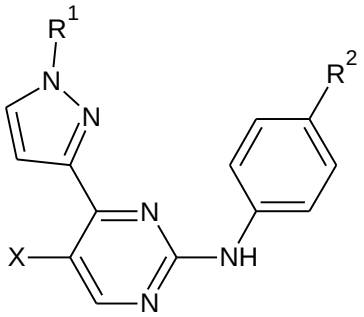
264011z	CH2OH	H	H	4.9
27a4011z	H	CH2OH	H	0.16
28a4011z	H	CH2CH2CH3	H	0.2
294011z	H	iPr	H	0.35
304011z	H	Ph	H	0.36
314011z	H	CH2CH2OH	H	0.62
324011z	H	CH2OCH3	H	0.69
33a4011z	Me	Me	H	0.27
344011z	Et	Et	H	0.16
35a4011z		CH2CH2	H	0.092
36a4011z		CH2CH2CH2	H	0.077
374011z		CH2CH2CH2CH2	H	0.18
384011z		CH2CH2CH2CH2CH2	H	1.22
394011z		CH2CH2CH2CH2CH2CH2	H	0.43
404011z		CH2CH2	Me	0.25
414011z		CH2CH2	Cl	0.19
AX135874011z		CH2CH2	F	0.16

Table 2.1.18 : SAR of 4-fluorophenyl isoxazolesa

AlzID	R	JNK3(IC50)	JNK1	p38
31012z	H	0.026	0.16	0.06
41012z	Ph	>20	nt	nt
61012z	CH3	0.032	nt	0.27
71012z	CN	0.126	nt	nt
91012z	N(Me)2	0.366	nt	5.16
101012z	NHMe	0.027	0.15	0.18
111012z	NHBn	0.134	nt	0.52
121012z	NH(CH2)2N(CH3)2	0.216	nt	0.17
131012z	OH	0.001	0.03	0.02

(He, Yuanjun, et al. "Synthesis and SAR of novel isoxazoles as potent c-jun N-terminal kinase (JNK) inhibitors." *Bioorganic & medicinal chemistry letters* 24.1 (2014): 161-164)

Table 2.1.18 : JNK3 inhibition by analogs of 7–9014z



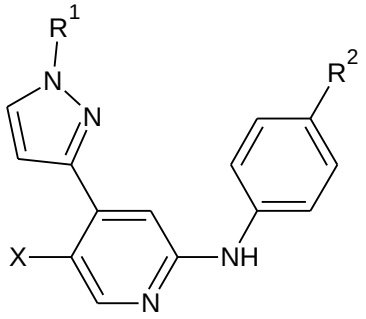
7-9 X=H,Cl

R1=H,Me

AlzID	X	R1	R2	JNK3 IC50	p38
11014z	H	H	A	0.63	>20
71014z	H	Me	A	1.45	nt
81014z	Cl	Me	A	0.86	nt
91014z	Cl	Me	B	0.73	nt

(Noël, Romain, et al. "Synthesis and SAR of 4-(pyrazol-3-yl)-pyridines as novel c-jun N-terminal kinase inhibitors." *Bioorganic & medicinal chemistry letters* 21.9 (2011): 2732-2735)

Table 2.1.19 : Inhibition of JNK3 by compounds 12–23014z



12-23 R^1=H,Me
X=H,Cl,F,Me
R^2=A -F

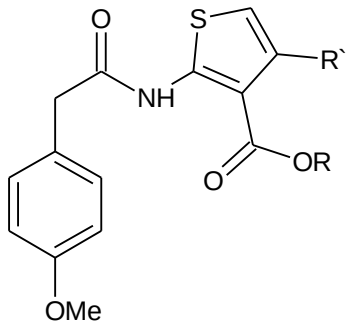
AlzID	X	R1	R2	JNK3 IC50	p38
122014z	H	H	E	0.16	>20
132014z	Cl	H	E	0.07	>20
142014z	Cl	Me	E	0.13	>20
152014z	Cl	Me	D	0.16	>20
162014z	Cl	Me	B	0.2	>20
172014z	Cl	Me	C	0.6	>20
182014z	F	Me	E	0.16	>20
192014z	F	Me	D	0.34	>20
202014z	F	Me	C	0.2	>20
212014z	F	Me	F	0.48	>20
222014z	F	Me	B	0.16	>20
232014z	Me	Me	B	0.75	>20

Table 2.1.20 : Inhibition of JNK3 by compounds 29–35014z

AlzID	X	R1	R2	JNK3 IC50	p38
293014z	F	Bn	E	0.1	>20
303014z	F	(CH2)2Ph	E	0.17	>20
323014z	H	3-MeOPh	D	0.65	>20
333014z	H	3-CNPh	D	0.24	>20
343014z	H	4-MeSPh	D	1.33	>20

353014z	H	2-Py	D	0.34	>20

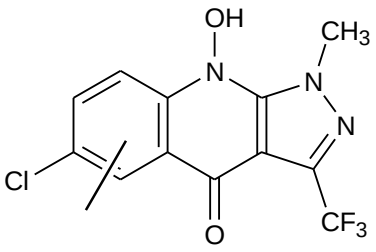
Table 2.1.21 : SAR of the 4-position of the thiophene



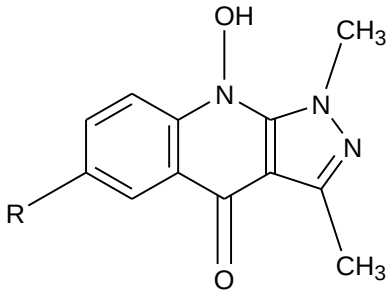
AlzID	R`	R	JNK3(IC50)	JNK1	JNK2
41015z	Me	Me	0.7	0.59	3.14
51015z	Et	Me	9.11	19.3	22.5
61015z	cPr	Et	9.62	15.58	6.29
71015z	CF3	Et	38.8	18.9	>50
81015z	CN	Et	0.31	0.35	0.93
91015z	CCH	Et	3.02	1.05	28.4

(Bowers, Simeon, et al. "Design and synthesis of a novel, orally active, brain penetrant, tri-substituted thiophene based JNK inhibitor." *Bioorganic & medicinal chemistry letters* 21.6 (2011): 1838-1843)

Table 2.1.22 : Enzymatic and cellular activity of analogs with aromatic ring modifications



9

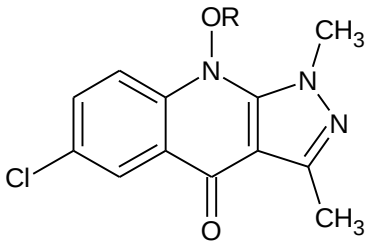


12

AlzID	R	JNK1	Pc-Jun
11017z	—	1.22	>30
21017z	—	0.98	16.4
101017z	—	0.92	>30
9a1017z	5-Cl	>10	NTa
9b1017z	7-Cl	5.14	>30
9c1017z	8-OMe	>10	NTa
9d1017z	7-N(Me)2	>10	NTa
12a1017z	H	>10	NTa
12b1017z	Ph	>10	NTa
12c1017z	1H-Pyrazol-3-yl	>10	NTa

(Liu, Mei, et al. "Synthesis and SAR of 1, 9-dihydro-9-hydroxypyrazolo [3, 4-b] quinolin-4-ones as novel, selective c-Jun N-terminal kinase inhibitors." *Bioorganic & medicinal chemistry letters* 16.10 (2006): 2590-2594)

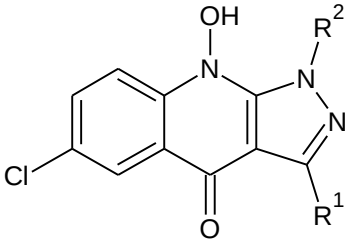
Table 2.1.23 : Enzymatic and cellular activity of 9-(alk)oxy analogs



14

AlzID	R	JNK1	Pc-Jun
12017z	H	1.22	>30
14a2017z	Me	4.59	19.4
14b2017z	Et	2.78	10.6
14c2017z	Pr	5.43	5.6
14d2017z	i-Pr	>10	NTa
14e2017z	Cyclopentyl	>10	NTa
14f2017z	Bn	>10	NTa

Table 2.1.24 : Enzymatic and cellular activity of analogs with C-1 and N-3 modifications

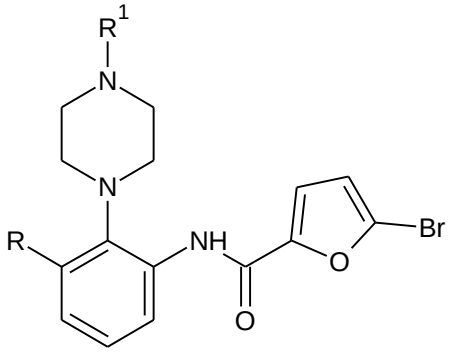


7

AlzID	R1	R2	JNK1	Pc-Jun
13017z	Me	Me	1.22	>30
7a3017z	Et	Me	0.96	32
7b3017z	n-Pr	Me	1.31	26.6
7c3017z	n-Bu	Me	0.52	>30
7d3017z	i-Pr	Me	2.64	>30
7e3017z	t-Bu	Me	>10	NTa

7f3017z	COOMe	Me	>10	NTa
7g3017z	4-NH2–Ph	Me	>10	NTa
7h3017z	CH2CH2OMe	Me	2.41	14.7
7i3017z	(CH2)2CO2H	Me	>10	NTa
7j3017z	(CH2)3CO2H	Me	2.74	>30
7k3017z	CH2CH2NH2	Me	2.83	>30
7l3017z	(CH2)3CONHMe	Me	0.75	>30
7m3017z	(CH2)2NHCOMe	Me	0.78	>30
7n3017z	(CH2)3NHCONHEt	Me	0.5	>30
7o3017z	(CH2)2NHCO2Et	Me	0.43	>30
7p3017z	Me	CH2CO2H	4.41	>30
7q3017z	Me	CH2CO2Et	7.64	>30
7r3017z	Me	CH2CONHMe	8.77	>30
7s3017z	Me	CH2CH2OH	7.67	>30
7t3017z	Me	Et	4.19	>30
7u3017z	Me	Ph	1.75	>30

Table 2.1.25 : N-Acyl-N-aryl piperazines

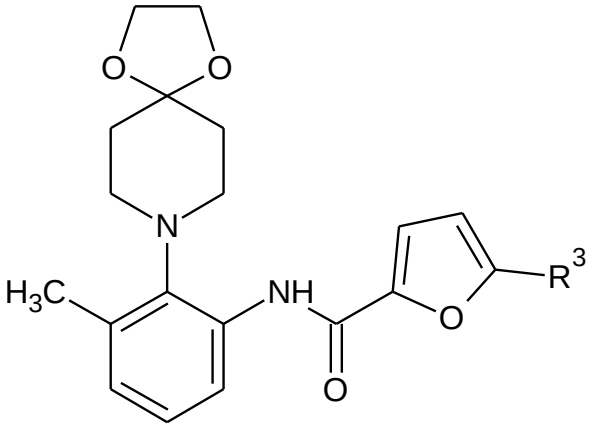


AlzID	R	R1	JNK3(IC50)	JNK1
4a1019z	Cl	H	9.9	4.1
4b1019z	Cl	Me	1.2	0.36
11019z	Cl	Ethyl	1.1	0.36

4c1019z	Cl	n-Pr	0.9	0.63
4d1019z	Cl	i-Pr	2.2	1.5
4e1019z	Cl	Allyl	0.33	0.24
4f1019z	Cl	2-Methallyl	0.54	0.4
4g1019z	Cl	Propargyl	0.16	0.14
4h1019z	Cl	Cyclopropyl	0.96	0.18
4i1019z	Cl	Furanylmethyl	0.25	0.27
4j1019z	Cl	Benzyl	1.1	0.88
4k1019z	Cl	Phenethyl	1.4	0.9
4l1019z	Cl	2-Pyridyl	1	1.2
4m1019z	Cl	Acetyl	0.81	0.6
4n1019z	Cl	Trifluoroacetyl	>20	>20
4o1019z	Cl	Boc	>20	6.3
4p1019z	Me	Allyl	0.2	0.18
4q1019z	F	Propargyl	0.29	0.11
4r1019z	Br	Allyl	>20	>20
4s1019z	CF3	Allyl	>20	>20
4t1019z	Ph	Allyl	>20*	>20
4u1019z	NHPh	Allyl	>20*	>20

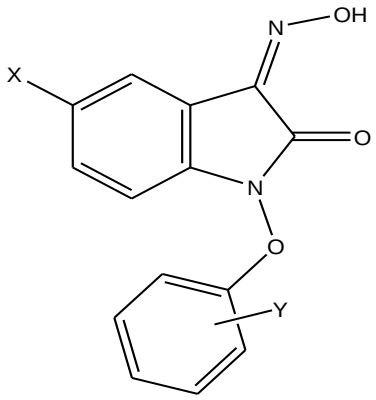
(Shin, Youseung, et al. "Synthesis and SAR of piperazine amides as novel c-jun N-terminal kinase (JNK) inhibitors." *Bioorganic & medicinal chemistry letters* 19.12 (2009): 3344-3347)

Table 2.1.26 : Aryl piperidines



AlzID	R3	JNK3(IC50)	JNk1
9b3091z	Br	0.06	0.09
9d3019z	Cl	0.08	0.04
9e3019z	F	0.41	0.21
9f3019z	CN	0.21	0.23
9g3019z	Me	0.53	0.33
9h3019z	CHF2	0.35	0.26
9i3019z	Et	1	1.4
9j3019z	Propynyl	0.11	0.11
9k3019z	Bn	4.8	3.2
9l3019z	OMe	0.62	0.45
9m3019z	NHAc	11	5.8
9n3019z	NHPh	>20	>20

Table 2.1.27 : JNK3 inhibition data for 5-substituted isatin derivatives

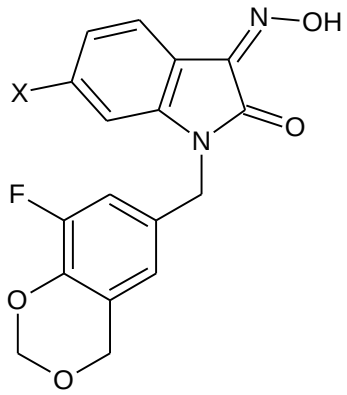


AlzID	X	Y	JNK3
31020z	H	H	>10
41020z	H	2-F	8.5
51020z	H	3-F	8.1
61020z	H	4-F	8.3
71020z	H	2-NO2	>10
81020z	H	3-NO2	0.94
91020z	H	4-NO2	7
101020z	H	3-Cl	5
111020z	H	4-Cl	6.7
121020z	H	3-Me	6.8
131020z	H	4-Me	9.9
141020z	H	3-CF3	5.2
151020z	H	3,5-(OMe)2	0.99
161020z	H	A	0.51
171020z	Me	H	>10
181020z	F	H	>10
191020z	Cl	H	12
201020z	Br	H	2.5
211020z	Me	3-CF3	2.7
221020z	Me	3-Me	>10
231020z	F	4-OMe	1.6
241020z	F	3,5-(OMe)2	0.79
251020z	F	A	0.44
261020z	Cl	A	0.74

271020z	OCF3	A	1.7

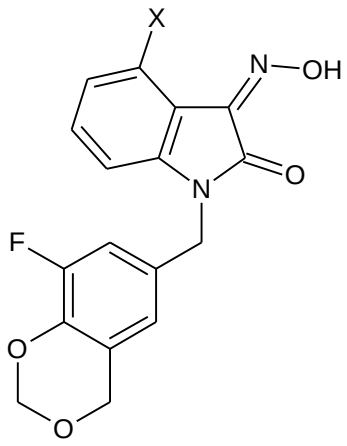
(Cao, Jingrong, et al. "Structure-based design and parallel synthesis of N-benzyl isatin oximes as JNK3 MAP kinase inhibitors." *Bioorganic & medicinal chemistry letters* 19.10 (2009): 2891-2895)

Table 2.1.28 : Data for 6-substituted isatins



AlzID	6-X	JNK3
162020z	H	0.51
282020z	Me	0.09
292020z	Br	0.1
302020z	Ph	0.2
312020z	OMe	0.24
322020z	CF3	0.33
332020z	NH2	0.4
342020z	MeSO2NH	0.46
352020z	EtCONH	0.61
362020z	NO2	0.64
372020z	COOMe	0.16
382020z	COOH	0.3
392020z	CONHEt	1

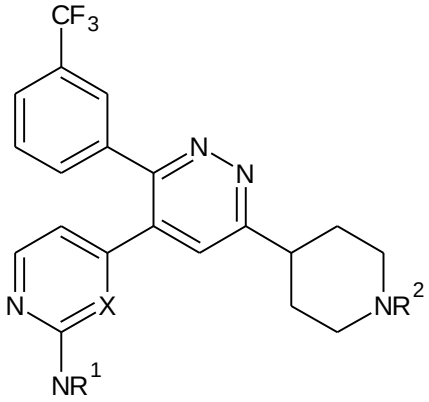
Table 2.1.29 : Data for 4-substituted isatins



AlzID	4-X	JNK3	p38a	ERK2
163020z	H	0.51		
403020z	Br	0.19	>10	>15
413020z	Ph	1.8	9.5	25
423020z	m-F–Ph	5.8	7.4	23
433020z	p-F–Ph	13	3.4	>15
443020z	CN	2		
453020z	COOMe	30		>15
463020z	Me	0.7		>15
473020z	CH@CHPh	0.74	0.23	
483020z	CH2CH2Ph	>15	0.03	>15
493020z	CH@CH2	0.14	8	>15
503020z	Et	1.4	>15	

(Cao, Jingrong, et al. "Structure-based design and parallel synthesis of N-benzyl isatin oximes as JNK3 MAP kinase inhibitors." *Bioorganic & medicinal chemistry letters* 19.10 (2009): 2891-2895)

Table 2.1.30 : SAR of piperidyl/pyridazines CF2



AlzID	X	R1	R2	p38	JNK3	THP-TNF
101021z	C	(S)-PhCH(CH3)	H	2.1±0.3	2970±31	21.5±0.6
151021z	N	MeCOPh	H	709±15	>10,000	1286±153
161021z	N	MeCO-pFPh	H	1772±82	>10,000	1245±54
171021z	N	(S)-PhCH(CH3)	H	1.6±0.2	1584±147	5.3±0.1
181021z	N	pF-PhCH(CH3)	H	10.2±0.5	>10,000	93.4±10.4
191021z	N	(S)-PhCH(CH3)	COCOHMe	6±0.9	556±67	15.4±1.1

(Tamayo, Nuria, et al. "Design and synthesis of potent pyridazine inhibitors of p38 MAP kinase." *Bioorganic & medicinal chemistry letters* 15.9 (2005): 2409-2413)

2.2 Enrichment of Inhibitor Data

The structure of inhibitors and their activity were collected from literature. The information such as Molecular Formula, SMILES, IUPAC Name, Composition, etc data was generated for all the molecules to facilitate searching. Further, the physiochemical properties : LogD, LogP, Molecular Weight, Strongest acidic pKa, strongest basic pKa, H-bond donor & acceptor, etc for all the inhibitors were calculated using ChemAxon software. To facilitate query based on common names, traditional names of all the rings attached to inhibitors were identified and incorporated to the database. To determine fragments specific to an enzyme,

the inhibitor structures were fragmented on the basis of Bemis Murcko (Ring, Linker, Side-Chain & Framework) method. All these data were processed and incorporated into AlzID database (http://14,139.240.55/AlzID/home.html).

Supplementary data will be provided in CD .

The following are the results :

Table. 2.2.1 : Data collection regarding the designed compounds (Ring, Formula, LogD)

AlzID	Ring	Formula	LogD
1101001z	indazole, benzene, piperidine	C25H24ClN5O	1.555543
2111001z	benzene, pyridine, tetrahydrofuran	C21H20N4O2	3.175707
3121001z	4H,5H,6H,7H-thieno[2,3-b]pyridine, naphthalene, cyclopropane	C23H19N3O2S	4.389167
4001z	benzene & 3,4-dihydroisoquinoline	C17H16ClNO2	3.991178
52001z	benzene & 3,4-dihydroisoquinoline	C17H15Cl2NO2	4.599198
62001z	benzene & 3,4-dihydroisoquinoline	C17H17NO2	3.368545
72001z	benzene & 3,4-dihydroisoquinoline	C17H16ClNO2	3.990408
82001z	benzene & 3,4-dihydroisoquinoline	C18H19NO3	3.221181
92001z	benzene & 3,4-dihydroisoquinoline	C16H13Cl2NO	4.757082
102001z	benzene & 3,4-dihydroisoquinoline	C15H12ClN	4.302024
112001z	benzene & 3,4-dihydroisoquinoline	C16H13ClFNO	4.295478
122001z	benzene & 3,4-dihydroisoquinoline	C16H13ClFNO	4.295992
132001z	benzene & 3,4-dihydroisoquinoline	C16H13ClFNO	4.294571
142001z	benzene & 3,4-dihydroisoquinoline	C16H13ClFNO	4.295278
163001z	benzene & 3,4-dihydroisoquinoline	C16H13BrClNO	4.92117
173001z	benzene & 3,4-dihydroisoquinoline	C16H13BrClNO	4.92176
183001z	1,4-dihydronaphthalene & 3,4-dihydroisoquinoline	C20H18ClNO	4.955815
193001z	1,4-dihydronaphthalene & 3,4-dihydroisoquinoline	C20H18ClNO	4.95993
203001z	benzene & 3,4-dihydroisoquinoline	C16H12Cl3NO	5.362153
213001z	benzene & 3,4-dihydroisoquinoline	C16H12Cl3NO	5.362394
223001z	benzene & 3,4-dihydroisoquinoline	C16H14ClNO	4.145591
233001z	benzene & 3,4-dihydroisoquinoline	C16H14ClNO	4.14762

243001z	benzene & 3,4-dihydroisoquinoline	C16H13Cl2NO	4.756499
253001z	benzene & 3,4-dihydroisoquinoline	C17H16ClNO	4.656175
263001z	benzene & 3,4-dihydroisoquinoline	C16H12Cl2FNO	4.901648
273001z	benzene & 3,4-dihydroisoquinoline	C17H16ClNO	4.073438
283001z	3,4-dihydroisoquinoline	C12H14ClNO	2.942131
1002z	benzene, pyrimidine	C15H18N4O2	4.079649
2002z	benzene, pyrimidine	C17H14N4O2	4.413818
3002z	benzene, pyridine	C18H15N3O2	4.439834
4002z	benzene, pyridine, piperazine	C22H23N5O2	2.515736
12a002z	benzene, pyridine	C20H18ClN3O3	3.810919
12b002z	benzene, pyridine	C20H18ClN3O3	5.110916
12c002z	benzene, pyridine & 1H-1,2,4-triazole	C20H16ClN5O2	4.478276
12d002z	benzene, pyrimidine	C18H15ClN4O3	4.860201
9a1002z	benzene, pyridine	C19H16FN3O3	4.425911
9b1002z	benzene, pyridine	C19H16ClN3O3	4.88724
9c1002z	benzene, pyridine	C20H16F3N3O3	5.160955
10a2002z	benzene, pyridine	C20H16F3N3O3	5.160957

Table. 2.2.2 : Data collection regarding the designed compounds (LogP, MW, Strongest acidic & basic pKa and TPSA)

AlzID	LogP	Molecular weight	Strongest acidic pKa	Strongest basic pKa	TPSA
1101001z	4.091428	445.95	14.2131933	10.0297269	81.84
2111001z	3.184022	360.417	11.96186876	5.700074101	76.14
3121001z	4.389167	401.48	13.59005026	-3.233704761	73.2
4001z	3.997527	301.77		5.568243363	30.82
52001z	4.601572	336.21		5.138987341	30.82
62001z	3.393482	267.328		6.171776942	30.82
72001z	3.997527	301.77		5.618395324	30.82
82001z	3.235811	297.354		5.934967697	40.05
92001z	4.759243	306.19		5.098200054	21.59
102001z	4.31287	241.72		5.803056412	12.36
112001z	4.2979	289.73		5.147845057	21.59
122001z	4.2979	289.73		5.043924812	21.59

132001z	4.2979	289.73		5.286391723	21.59
142001z	4.2979	289.73		5.182409824	21.59
163001z	4.923951	350.64		5.208038005	21.59
173001z	4.923951	350.64		5.104126167	21.59
183001z	4.975471	323.82		6.06573499	21.59
193001z	4.975471	323.82		5.961659109	21.59
203001z	5.363288	340.63		4.817892968	21.59
213001z	5.363288	340.63		4.714019974	21.59
223001z	4.155198	271.74		5.749803356	21.59
233001z	4.155198	271.74		5.645761755	21.59
243001z	4.759243	306.19		5.202133578	21.59
253001z	4.66862	285.77		5.863627657	21.59
263001z	4.901945	324.18		4.234778821	21.59
273001z	4.087833	285.77		5.927809624	21.59
283001z	2.954029	223.7		5.843800388	21.59
1002z	4.079773	286.335	13.86965873	3.870907569	90.13
2002z	4.413832	306.325	13.81379916	2.933885633	90.13
3002z	4.440884	305.337	14.68112539	4.798242566	77.24
4002z	4.013	389.459	14.6811271	8.890025769	92.51
12a002z	3.810934	383.83	14.63237138	2.930641621	77.68
12b002z	5.110934	383.83	14.91549421	3.01434495	72.48
12c002z	4.478674	393.83	10.45024635	3.031313161	84.95
12d002z	4.860206	370.79	12.42173864	1.127849361	99.36
9a1002z	4.425915	353.353	14.53379291	2.336020197	86.47
9b1002z	4.887258	369.81	14.61708707	3.014565119	86.47
9c1002z	5.161061	403.361	14.50501197	3.804169445	86.47
10a2002z	5.161061	403.361	14.50485682	3.793169201	86.47

Table. 2.2.3 : Data collection regarding the designed compounds (Composition and IUPAC Names)

AlzID	Composition	IUPAC Name

1101001z	C (67.33%), H (5.42%), Cl (7.95%), N (15.7%), O (3.59%)	3-{6-[(2-chlorophenyl)amino]-1H-indazol-3-yl}-N-(piperidin-4-yl)benzamide
2111001z	C (69.98%), H (5.59%), N (15.55%), O (8.88%)	(3S)-N-[2'-(phenylamino)-[4,4'-bipyridin]-2-yl]oxolane-3-carboxamide
3121001z	C (68.81%), H (4.77%), N (10.47%), O (7.97%), S (7.99%)	N-{3-cyano-7-cyclopropanecarbonyl-4H,5H,6H,7H-thieno[2,3-b]pyridin-2-yl}naphthalene-1-carboxamide
4001z	C (67.66%), H (5.34%), Cl (11.75%), N (4.64%), O (10.6%)	1-(3-chlorophenyl)-6,7-dimethoxy-3,4-dihydroisoquinoline
52001z	C (60.73%), H (4.5%), Cl (21.09%), N (4.17%), O (9.52%)	1-(3,4-dichlorophenyl)-6,7-dimethoxy-3,4-dihydroisoquinoline
62001z	C (76.38%), H (6.41%), N (5.24%), O (11.97%)	6,7-dimethoxy-1-phenyl-3,4-dihydroisoquinoline
72001z	C (67.66%), H (5.34%), Cl (11.75%), N (4.64%), O (10.6%)	1-(4-chlorophenyl)-6,7-dimethoxy-3,4-dihydroisoquinoline
82001z	C (72.71%), H (6.44%), N (4.71%), O (16.14%)	6,7-dimethoxy-1-(3-methoxyphenyl)-3,4-dihydroisoquinoline
92001z	C (62.76%), H (4.28%), Cl (23.16%), N (4.57%), O (5.23%)	6-chloro-1-(3-chlorophenyl)-7-methoxy-3,4-dihydroisoquinoline
102001z	C (74.53%), H (5%), Cl (14.67%), N (5.79%)	1-(3-chlorophenyl)-3,4-dihydroisoquinoline
112001z	C (66.33%), H (4.52%), Cl (12.24%), F (6.56%), N (4.83%), O (5.52%)	7-chloro-1-(3-fluorophenyl)-6-methoxy-3,4-dihydroisoquinoline
122001z	C (66.33%), H (4.52%), Cl (12.24%), F (6.56%), N (4.83%), O (5.52%)	6-chloro-1-(3-fluorophenyl)-7-methoxy-3,4-dihydroisoquinoline
132001z	C (66.33%), H (4.52%), Cl (12.24%), F (6.56%), N (4.83%), O (5.52%)	7-chloro-1-(4-fluorophenyl)-6-methoxy-3,4-dihydroisoquinoline

142001z	C (66.33%), H (4.52%), Cl (12.24%), F (6.56%), N (4.83%), O (5.52%)	6-chloro-1-(4-fluorophenyl)-7-methoxy-3,4-dihydroisoquinoline
163001z	C (54.81%), H (3.74%), Br (22.79%), Cl (10.11%), N (3.99%), O (4.56%)	1-(3-bromophenyl)-7-chloro-6-methoxy-3,4-dihydroisoquinoline
173001z	C (54.81%), H (3.74%), Br (22.79%), Cl (10.11%), N (3.99%), O (4.56%)	1-(3-bromophenyl)-6-chloro-7-methoxy-3,4-dihydroisoquinoline

Table. 2.2.4 : Data collection regarding the designed compounds (Assymetric atoms, Atom count, Bond count, Chiral atoms)

AlzID	Asymmetric atoms	Atom count	Bond count	Chiral atoms
1101001z	0	56	60	0
2111001z	1	47	50	1
3121001z	0	48	52	0
4001z	0	37	39	0
52001z	0	37	39	0
62001z	0	37	39	0
72001z	0	37	39	0
82001z	0	41	43	0
92001z	0	33	35	0
102001z	0	29	31	0
112001z	0	33	35	0
122001z	0	33	35	0
132001z	0	33	35	0
142001z	0	33	35	0
163001z	0	33	35	0
173001z	0	33	35	0
183001z	0	41	44	0
193001z	0	41	44	0

203001z	0	33	35	0
213001z	0	33	35	0
223001z	0	33	35	0
233001z	0	33	35	0
243001z	0	33	35	0
253001z	0	36	38	0
263001z	0	33	35	0
273001z	0	36	38	0
283001z	0	29	30	0
1002z	0	39	40	0
2002z	0	37	39	0
3002z	0	38	40	0
4002z	0	52	55	0
12a002z	0	45	47	0
12b002z	0	45	47	0
12c002z	0	44	47	0
12d002z	0	41	43	0
9a1002z	0	42	44	0
9b1002z	0	42	44	0
9c1002z	0	45	47	0
10a2002z	0	45	47	0

Table. 2.2.5 : Data collection regarding the designed compounds (H bond acceptors & donors, Ring count, Rotatable bonds)

AlzID	H bond acceptors	H bond donors	Ring count	Rotatable bonds
1101001z	4	4	5	5
2111001z	5	2	4	5
3121001z	3	1	5	3
4001z	3	0	3	3
52001z	3	0	3	3
62001z	3	0	3	3
72001z	3	0	3	3
82001z	4	0	3	4

92001z	2	0	3	2
102001z	1	0	3	1
112001z	2	0	3	2
122001z	2	0	3	2
132001z	2	0	3	2
142001z	2	0	3	2
163001z	2	0	3	2
173001z	2	0	3	2
183001z	2	0	4	2
193001z	2	0	4	2
203001z	2	0	3	2
213001z	2	0	3	2
223001z	2	0	3	2
233001z	2	0	3	2
243001z	2	0	3	2
253001z	2	0	3	2
263001z	2	0	3	2
273001z	2	0	3	3
283001z	2	0	2	2
1002z	5	2	2	7
2002z	4	2	3	5
3002z	3	2	3	5
4002z	5	3	4	6
12a002z	4	1	3	6
12b002z	4	2	3	6
12c002z	5	2	4	6
12d002z	5	2	3	6
9a1002z	4	2	3	6
9b1002z	4	2	3	6
9c1002z	4	2	3	7
10a2002z	4	2	3	7

2.3 Geometry Optimization

3D structures of these inhibitors are provided to assist virtual screening. The structures were drawn using MarvinSketch software available in ChemAxon.

Geometries were optimized using B3LYP (Becke's Lee Yang and Parr correlation) approach which is a hybrid functional algorithm that uses Becke's three parameters to mix in the exact Hartee-Fock Exchange correlation and Lee Yang and Parr (LYP) correlation functional that recovers dynamic electron correlation.

2.4 Fragmentation

Bemis Murcko Fragments i.e. ring, side-chain, linker and framework, of all the molecules were performed to facilitate faster analysis and co-occurrence studies. OpenEye's program OEChem and OEMedChem was used to generate fragments of each molecule.

```
1  package openeye.docexamples.oemedchem;
2
3  import openeye.oechem.*;
5  public class OEGetBemisMurcko {
6
7      public static void main(String[] args) {
8          // TODO Auto-generated method stub
9             OEGraphMol mol = new OEGraphMol();
10             oechem.OESmilesToMol(mol, "[H]c1ccc2c(c1)c(=NO)c(=O)n2Cc3cccc(F)c3");
11             for (OEAtomBondSet abset : oemedchem.OEGetBemisMurcko(mol)) {
12                 OEIsAtomMember fragatompred = new OEIsAtomMember(abset.GetAtoms());
13                 OEIsBondMember fragbondpred = new OEIsBondMember(abset.GetBonds());
14
15                 OEGraphMol fragment = new OEGraphMol();
16                 boolean  adjustHCount = true;
17                 oechem.OESubsetMol(fragment, mol, fragatompred, fragbondpred, adjustHCount);
18                 for (OERole role : abset.GetRoles()) {
19                     System.out.printf("%s %s%n", role.GetName(), oechem.OEMolToSmiles(fragment));
20                 }
21             }
22     }
23
24 }
25
```

Fig. 2.4.1 : Openeye's Fragmentation code

Table 2.4.3 : Fragment Collection (Ring, Linker)

AlzID	Ring_SMILES	Linker
1101001z	c1ccccc1.c1ccc(cc1)c2c3ccccc3[nH]n2.C1CCNCC1	CN.N
2111001z	c1ccccc1.c1cnccc1c2ccncc2.C1CCOC1	CN.N
3121001z	c1ccc2ccccc2c1.c1csc2c1CCCN2.C1CC1	C.CN
4001z	c1ccc(cc1)C2=NCCc3c2cccc3	-
52001z	c1ccc(cc1)C2=NCCc3c2cccc3	-
62001z	c1ccc(cc1)C2=NCCc3c2cccc3	-
72001z	c1ccc(cc1)C2=NCCc3c2cccc3	-
82001z	c1ccc(cc1)C2=NCCc3c2cccc3	-
92001z	c1ccc(cc1)C2=NCCc3c2cccc3	-
102001z	c1ccc(cc1)C2=NCCc3c2cccc3	-
112001z	c1ccc(cc1)C2=NCCc3c2cccc3	-
122001z	c1ccc(cc1)C2=NCCc3c2cccc3	-
132001z	c1ccc(cc1)C2=NCCc3c2cccc3	-
142001z	c1ccc(cc1)C2=NCCc3c2cccc3	-
163001z	c1ccc(cc1)C2=NCCc3c2cccc3	-
173001z	c1ccc(cc1)C2=NCCc3c2cccc3	-
183001z	c1ccc2c(c1)CCN=C2C3=CCc4ccccc4C3	-
193001z	c1ccc2c(c1)CCN=C2C3=CCc4ccccc4C3	-
203001z	c1ccc(cc1)C2=NCCc3c2cccc3	-
213001z	c1ccc(cc1)C2=NCCc3c2cccc3	-
223001z	c1ccc(cc1)C2=NCCc3c2cccc3	-
233001z	c1ccc(cc1)C2=NCCc3c2cccc3	-
243001z	c1ccc(cc1)C2=NCCc3c2cccc3	-
253001z	c1ccc(cc1)C2=NCCc3c2cccc3	-
263001z	c1ccc(cc1)C2=NCCc3c2cccc3	-
273001z	c1ccccc1.c1ccc2c(c1)CCN=C2	C
283001z	c1ccc2c(c1)CCN=C2	
1002z	c1ccccc1.c1cncnc1	N
2002z	c1ccccc1.c1ccccc1.c1cncnc1	N.O
3002z	c1ccccc1.c1ccccc1.c1ccncc1	N.O
4002z	c1ccccc1.c1ccc(cc1)N2CCNCC2.c1ccncc1	N.O
12a002z	c1ccccc1.c1ccccc1.c1ccncc1	N.O
12b002z	c1ccccc1.c1ccccc1.c1ccncc1	N.O

12c002z	c1ccccc1.c1ccc(cc1)c2nc[nH]n2.c1ccncc1	N.O
12d002z	c1ccccc1.c1ccccc1.c1cncnc1	N.O
9a1002z	c1ccccc1.c1ccccc1.c1ccncc1	N.O
9b1002z	c1ccccc1.c1ccccc1.c1ccncc1	N.O
9c1002z	c1ccccc1.c1ccccc1.c1ccncc1	N.O
10a2002z	c1ccccc1.c1ccccc1.c1ccncc1	N.O
10b2002z	c1ccccc1.c1ccccc1.c1ccncc1	N.O
10c2002z	c1ccccc1.c1ccccc1.c1ccncc1	N.O
10d2002z	c1ccccc1.c1ccccc1.c1ccncc1	N.O
10e2002z	c1ccccc1.c1ccccc1.c1ccncc1	N.O
10f2002z	c1ccccc1.c1ccccc1.c1ccncc1	N.O
10g2002z	c1ccccc1.c1ccccc1.c1ccncc1	N.O
10h2002z	c1ccccc1.c1ccccc1.c1ccncc1	N.O
10i2002z	c1ccccc1.c1ccccc1.c1ccncc1	N.O
10j2002z	c1ccccc1.c1ccccc1.c1ccncc1	N.O
10k2002z	c1ccccc1.c1ccccc1.c1ccncc1	N.O
10l2002z	c1ccccc1.c1ccncc1.c1ccc2ccccc2c1	N.O
10m2002z	c1ccccc1.c1ccncc1.c1ccc2ccccc2c1	N.O
10n2002z	c1ccccc1.c1ccncc1.c1ccc2c(c1)OCO2	N.O

Table 2.4.4 : Fragment Collection (Side - chain and SMILES)

AlzID	SideChain	Smiles
1101001z	O.Cl	Clc1ccccc1Nc1ccc2c(n[nH]c2c1)-c1cccc(c1)C(=O)NC1CCNCC1
2111001z	O	O=C(Nc1cc(ccn1)-c1ccnc(Nc2ccccc2)c1)[C@H]1CCOC1
3121001z	C#N.O.O	O=C(Nc1sc2N(CCCc2c1C#N)C(=O)C1CC1)c1cccc2ccccc12
4001z	CO.CO.Cl	COc1cc2CCN=C(c3cccc(Cl)c3)c2cc1OC
52001z	CO.CO.Cl.Cl	COc1cc2CCN=C(c3ccc(Cl)c(Cl)c3)c2cc1OC
62001z	CO.CO	COc1cc2CCN=C(c3ccccc3)c2cc1OC
72001z	CO.CO.Cl	COc1cc2CCN=C(c3ccc(Cl)cc3)c2cc1OC
82001z	CO.CO.CO	COc1cccc(c1)C1=NCCc2cc(OC)c(OC)cc12
92001z	CO.Cl.Cl	COc1cc2C(=NCCc2cc1Cl)c1cccc(Cl)c1
102001z	Cl	Clc1cccc(c1)C1=NCCc2ccccc12
112001z	CO.F.Cl	COc1cc2CCN=C(c3cccc(F)c3)c2cc1Cl

122001z	CO.F.Cl	COc1cc2C(=NCCc2cc1Cl)c1cccc(F)c1
132001z	CO.F.Cl	COc1cc2CCN=C(c3ccc(F)cc3)c2cc1Cl
142001z	CO.F.Cl	COc1cc2C(=NCCc2cc1Cl)c1ccc(F)cc1
163001z	CO.Cl.Br	COc1cc2CCN=C(c3cccc(Br)c3)c2cc1Cl
173001z	CO.Cl.Br	COc1cc2C(=NCCc2cc1Cl)c1cccc(Br)c1
183001z	CO.Cl	COc1cc2CCN=C(C3=CCc4ccccc4C3)c2cc1Cl
193001z	CO.Cl	COc1cc2C(=NCCc2cc1Cl)C1=CCc2ccccc2C1
203001z	CO.Cl.Cl.Cl	COc1cc2CCN=C(c3ccc(Cl)c(Cl)c3)c2cc1Cl
213001z	CO.Cl.Cl.Cl	COc1cc2C(=NCCc2cc1Cl)c1ccc(Cl)c(Cl)c1
223001z	CO.Cl	COc1cc2CCN=C(c3ccccc3)c2cc1Cl
233001z	CO.Cl	COc1cc2C(=NCCc2cc1Cl)c1ccccc1
243001z	CO.Cl.Cl	COc1cc2CCN=C(c3cccc(Cl)c3)c2cc1Cl
253001z	C.CO.Cl	COc1cc2CCN=C(c3cccc(C)c3)c2cc1Cl
263001z	CO.F.Cl.Cl	COc1cc2CCN=C(c3cccc(Cl)c3F)c2cc1Cl
273001z	CO.Cl	COc1cc2CCN=C(Cc3ccccc3)c2cc1Cl
283001z	CC.CO.Cl	CCC1=NCCc2cc(OC)c(Cl)cc12
1002z	CCCCO.C(=O)N	CCCCOc1nccc(Nc2ccccc2C(N)=O)n1
2002z	C(=O)N	NC(=O)c1ccccc1Nc1ccnc(Oc2ccccc2)n1
3002z	C(=O)N	NC(=O)c1ccccc1Nc1ccnc(Oc2ccccc2)c1
4002z	C(=O)N	NC(=O)c1ccccc1Nc1ccnc(Oc2ccc(cc2)N2CCNCC2)c1
12a002z	C.CO.C(=O)N.Cl	COc1ccc(Oc2cc(N(C)c3ccccc3C(N)=O)c(Cl)cn2)cc1
12b002z	CNC=O.CO.Cl	CNC(=O)c1ccccc1Nc1cc(Oc2ccc(OC)cc2)ncc1Cl
12c002z	CO.Cl	COc1ccc(Oc2cc(Nc3ccccc3-c3nc[nH]n3)c(Cl)cn2)cc1
12d002z	CO.C(=O)N.Cl	COc1ccc(Oc2ncc(Cl)c(Nc3ccccc3C(N)=O)n2)cc1
9a1002z	CO.C(=O)N.F	COc1ccc(Oc2cc(Nc3ccccc3C(N)=O)c(F)cn2)cc1
9b1002z	CO.C(=O)N.Cl	COc1ccc(Oc2cc(Nc3ccccc3C(N)=O)c(Cl)cn2)cc1
9c1002z	CO.C(=O)N.C(F)(F)F	COc1ccc(Oc2cc(Nc3ccccc3C(N)=O)c(cn2)C(F)(F)F)cc1

2.5 Identification of Common & Unique Fragments

Fragmentation dataset was analyzed to identify fragment patterns and their association with drug targets of Alzheimer Disease. We identified the traditional names of each ring of

molecules to analyze the frequency of fragments more than random in dataset by taking a particular threshold. We've also analyzed the frequencies of side-chain and linkers to get the pattern of fragments in dataset and to identify the association between the fragments for significant co-occurance. This analysis also assists to determine common and unique fragments.

2.6 Database Development

All the data were incorporated into "AlzID" Data resource (http://14,139.240.55/AlzID/home.html) to support access and retrieval of information. MySQL is used to generate queries and to retrieve information. Java Molecular Editor (JME) interface is provided to draw or upload structures of own choice. R's RCDK package is used for similarity searching based on fingerprints which further calculates similarity using "Tannimoto Coefficient".

2.7 Softwares Used

2.7.1 ChemSketch

It is software used for drawing chemical structures including organics, organometallics, polymers and Markush structures. It includes other features such as Calculation of molecular properties of molecules, their 2D and 3D structure cleaning and viewing and also the functionality for naming structures.

We used this software for designing our compounds and collecting SMILE Notation of respective compound. Compounds were stored in 2D (.png files) and 3D (.mol files) using this software.

2.7.2 ChemDraw

We used ChemDraw for cleaning the geometry of our structures & to generate the IUPAC Names of all the molecules.

2.7.3 MarvinSketch

MarvinSketch was used to generate xyz files of all the molecules for optimization & sdf files for generating the fingerprints of all the molecules for similarity searching. Traditional names of ring were calculated for fragment analysis using MarvinSketch.

2.7.4 Instant JChem

Instant JChem was used to calculate the physiochemical properties such as LogD, LogP, Molecular Weight, Strongest acidic pKa, Strongest basic pKa, H-bond donor & acceptor,

Topological polar surface area, rotatable bonds, etc of all the molecules.

2.7.5 PuTTy

PuTTY is an SSH and telnet client. We used PuTTy for optimizing the geometry of our compounds using B3LYP algorithm.

2.7.6 OpenEye

Openeye's OEMedChem and OEchem package was used for fragmentation of each molecule into Bemis Murcko Fragments i.e. ring, side-chain, linker & framework.

2.7.7 MySQL

MySQL is a Relational Database Management System (RDMS) which was used with PHP for the connectivity of our database to provide access and retrieval of information.

2.7.8 XAMPP

XAMPP is an integrated server package of Apache, MySQL, PHP and Perl. We used XAMPP to connect to localhost for executing our queries.

2.7.9 RCDK

RCDK is integrated CDK with R. CDK is Java library to support Chemoinformatics functionalities. We used RCDK to generate fingerprints of our molecules (approx. 700) and of query molecule to calculate the similarity between query and target molecules to provide desired library of molecules to the user.

CHAPTER – 3

RESULTS AND DISCUSSION

ALZHEIMERS DISEASE INHIBITORS DATABASE

(http://14,139.240.55/AlzID/home.html)

The database schema is organized relationally to make our database compatible with efficient loading, querying and updates. The resource presently contains data about 650 inhibitors. Data were prepared and validated for virtual screening. Inhibitors molecular properties were calculated so that user can get the “lead-like”, “fragment-like” and “drug-like” molecules based on Lipinski’s rule to evaluate drug-likeness or to determine if a molecule is biologically active or not. For this evaluation, we have provided molecular properties such as Molecular Weight, Hydrogen-bond donors, Hydrogen-bond acceptor, rotatable bonds, LogP etc.

ARCHITECTURE OF DATABASE

Fig. 3.1 : Architecture of AlzID

AlzID resource contains information about inhibitors active against the Alzheimer disease's (AD) promising JNK-isotypes drug targets.

• Information to facilitate searching: Molecular Formula, SMILES, IUPAC Name, Bemis Murcko Fragments (Ring, Linker, Side-chain, Framework)
• Physicochemical properties: LogD, LogP, Molecular Weight, Strongest acidic pKa, strongest basic pKa, H-bond donor & acceptor, etc
• Fragments that are unique or common to particular target
• Optimized 3D molecules (B3LYP/6-311G*) to assist computational drug discovery

All the data were incorporated into "AlzID" Database which uses the following architecture to store the data. Diverse queries are supported to access and retrieval of information.

Fig. 3.2 : Home Page of AlzID

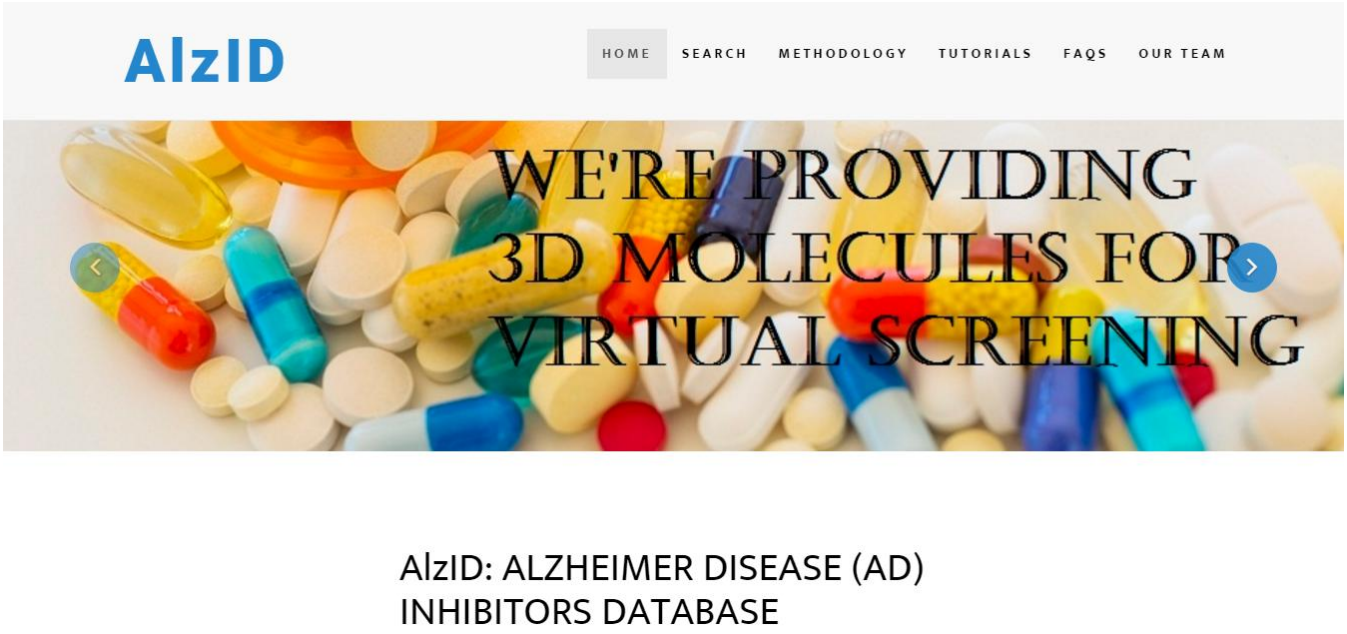


Fig. 3.3 : Front Page slider

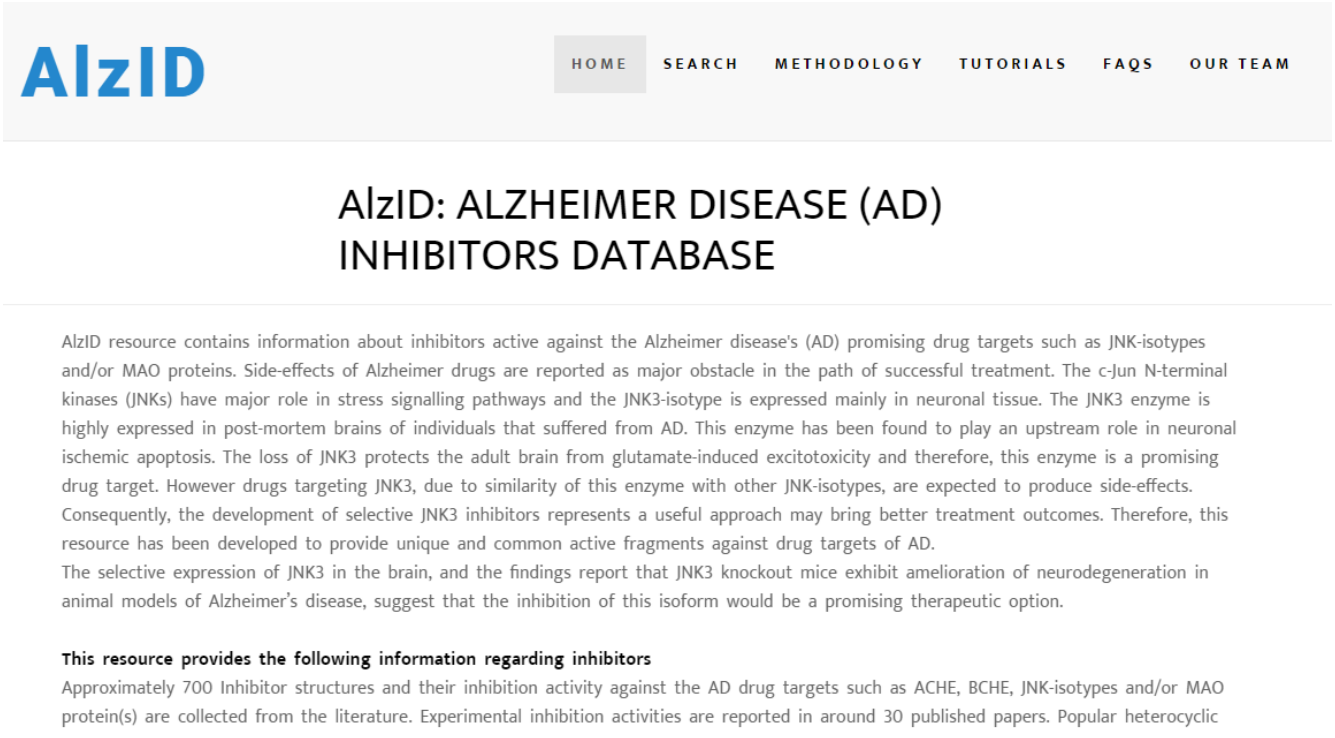


Fig. 3.4 : Front page content containing objective, data information and applications of AlzID

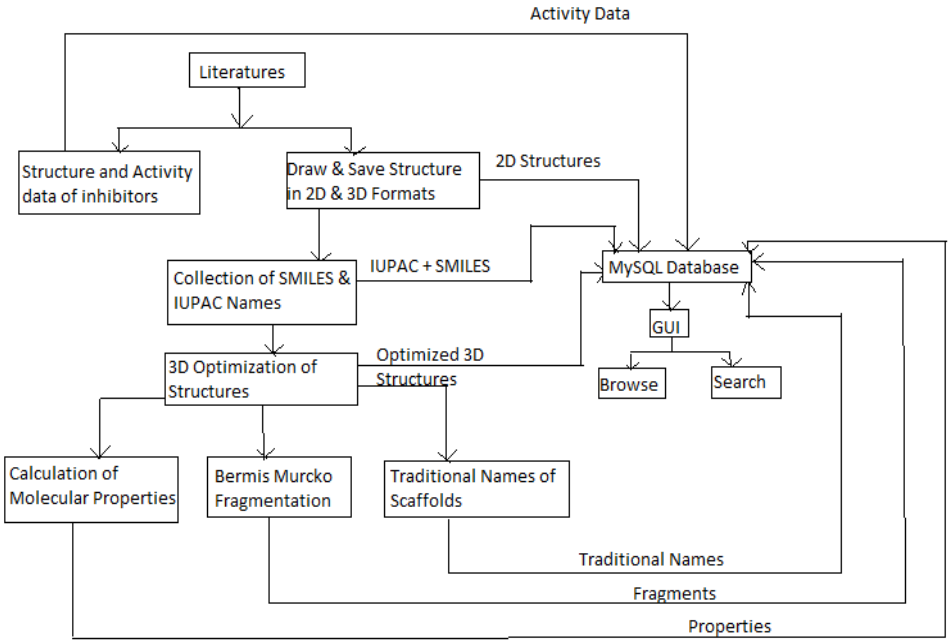


Fig. 3.1 : Architecture of AlzID

QUERY DESIGNING AND SEARCHING

The database provides users to search, browse and download molecules in pdb, SDF and mol formats. User can draw or upload the molecules in SMILES, SDF and mol2 format to get the libraries of their interest. Molecular structure can be drawn using Java Molecular Editor (JME) where user can click on "GET SMILES" button to get the SMILES of their molecule and choose the identity criteria like 90%, 80%, 70%, 60% or less for searching. SMILES may also be uploaded to search for library. We used R's RCDK package for similarity searching between the molecule or SMILES user upload and our dataset. We took input from users & execution is done on R. RCDK generates the fingerprints of user uploaded molecule or SMILES and our dataset to search for similarity. It calculates similarity based on "Tannimoto Coefficient".

```
library(rcdk)
a<- load.molecules(c('com002.sdf', 'C:/Users/HP/Pictures/xampp/htdocs/alzid/alzid/com002.sdf'))
filename <- "sm.txt"
sd <- readChar(filename, file.info(filename)$size)
sf <- toString(sd)
v <- parse.smiles(sf)[[1]]
query.fp <- get.fingerprint(v, type='maccs')
target.fps <- lapply(a, get.fingerprint, type='maccs')
sims <- unlist(lapply(target.fps,distance, fp2=query.fp, method='tanimoto'))
filei <- "iden.txt"
i <- readChar(filei, file.info(filei)$size)
hits <- which(sims > i)
write.table(hits, file = "C:/Users/HP/Pictures/xampp/htdocs/alzid/alzid/output.csv", sep = ",", col.names = NA,qmethod = "double")
hits
sims
```

Fig. 3.5 : RCDK code for Similarity Searching

The above code is used to generate fingerprints the molecules to calculate similarity based on tannimoto coefficient. Here ‘com002.sdf’ is combined sdf file of all the inhibitors in database, ‘sm.txt’ is a file containing the SMILES of compound user upload, whereas ‘iden.txt’ file contain the identity value i.e. 0.90, 0.80, 0.70, 0.60 so on user chooses to search for library of molecules.

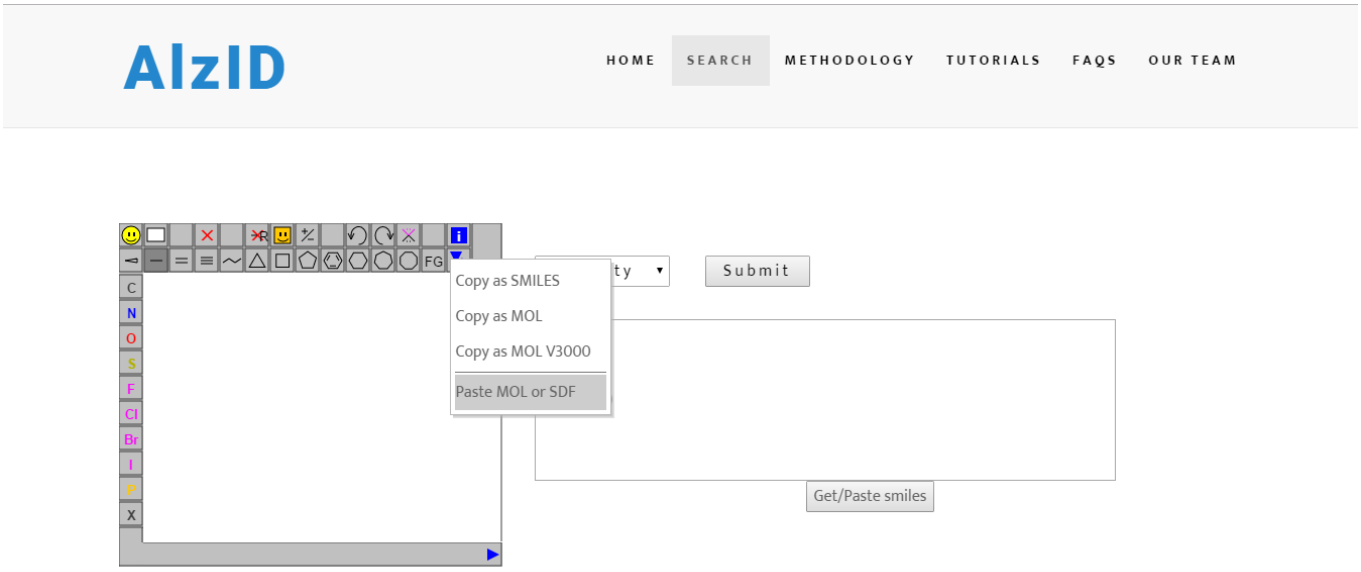


Fig. 3.6 : JME interface providing different file formats to upload file

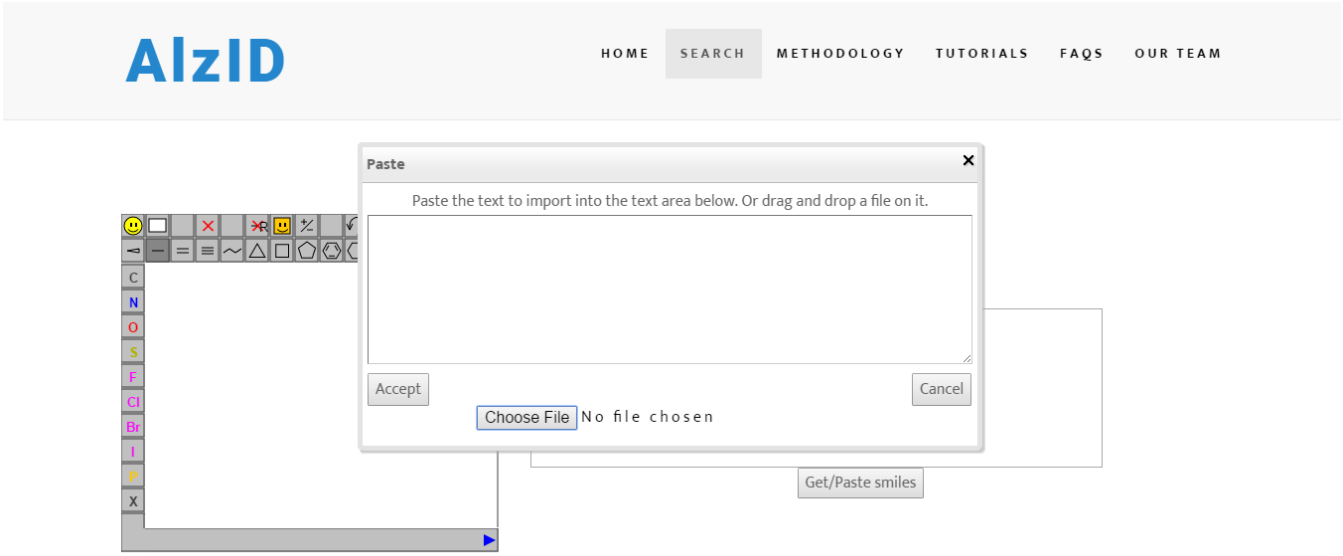


Fig. 3.7 : JME interface to choose file of own choice

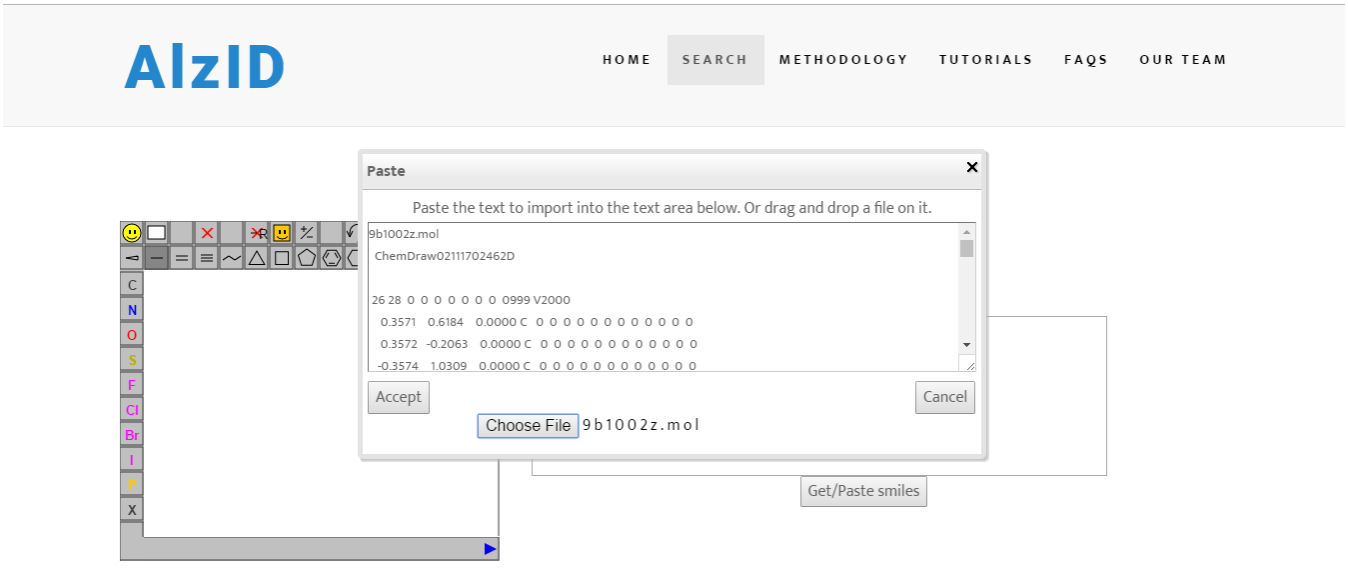


Fig. 3.8 : JME interface allowing mol file 9b1002z to upload

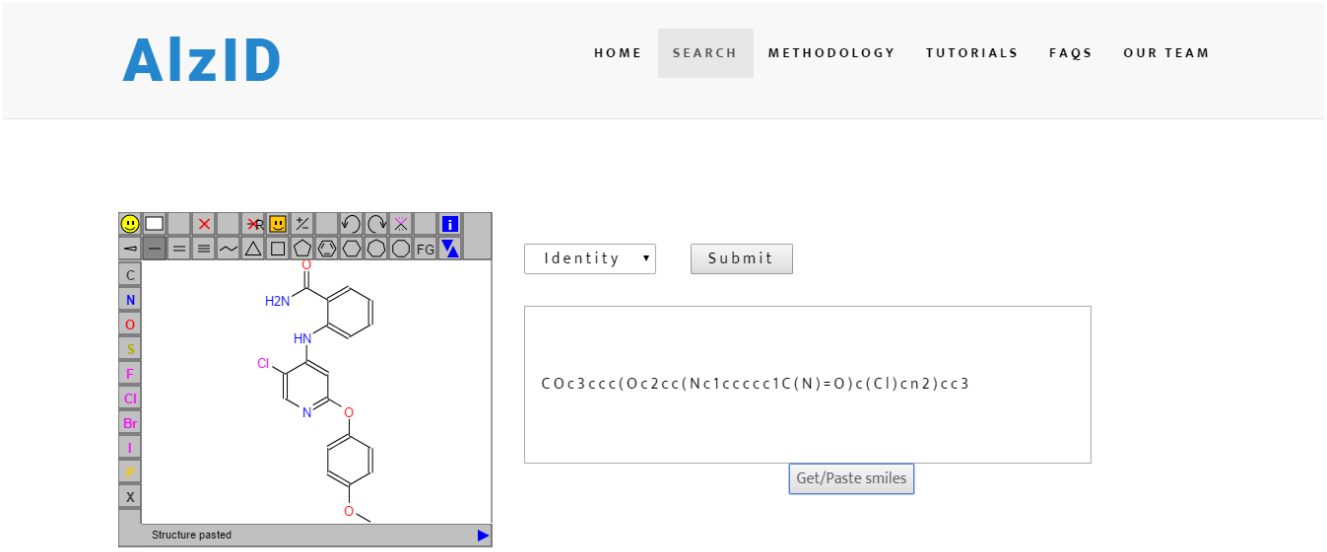


Fig. 3.9 : Uploading of mol file 9b1002z to get smiles on click over "Get/Paste smiles" button

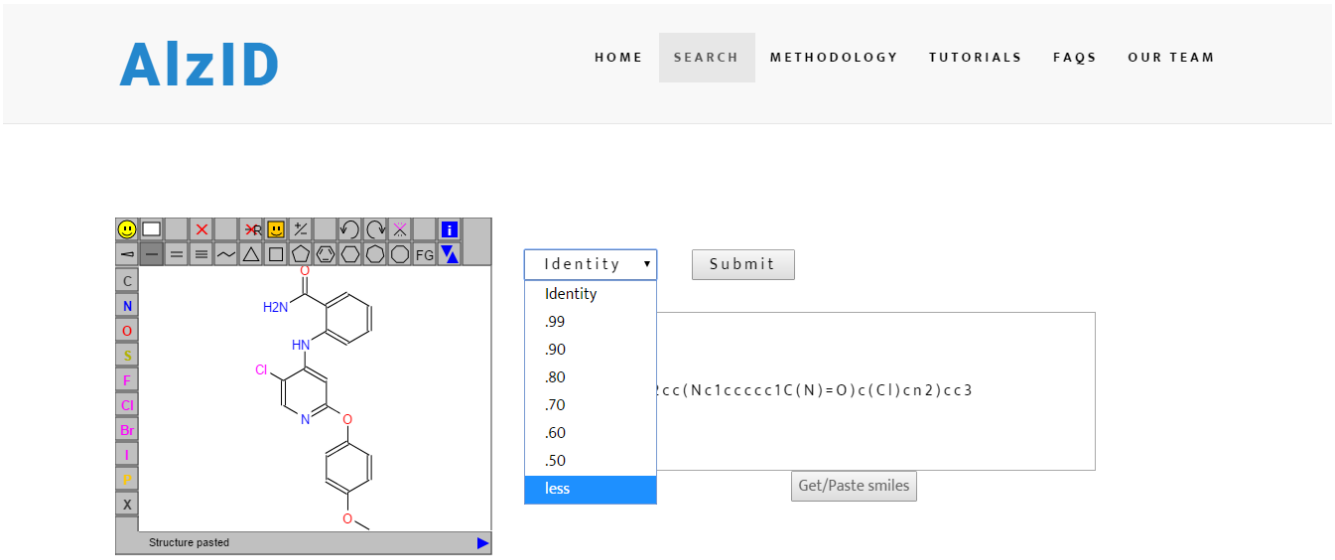


Fig. 3.10 : Draw Structure page providing "Identity Criteria" to make search more specific for library of molecules

The above figures showing steps of JME interface to upload file of own choice in different file formats and to allow searching based on Identity Criteria.

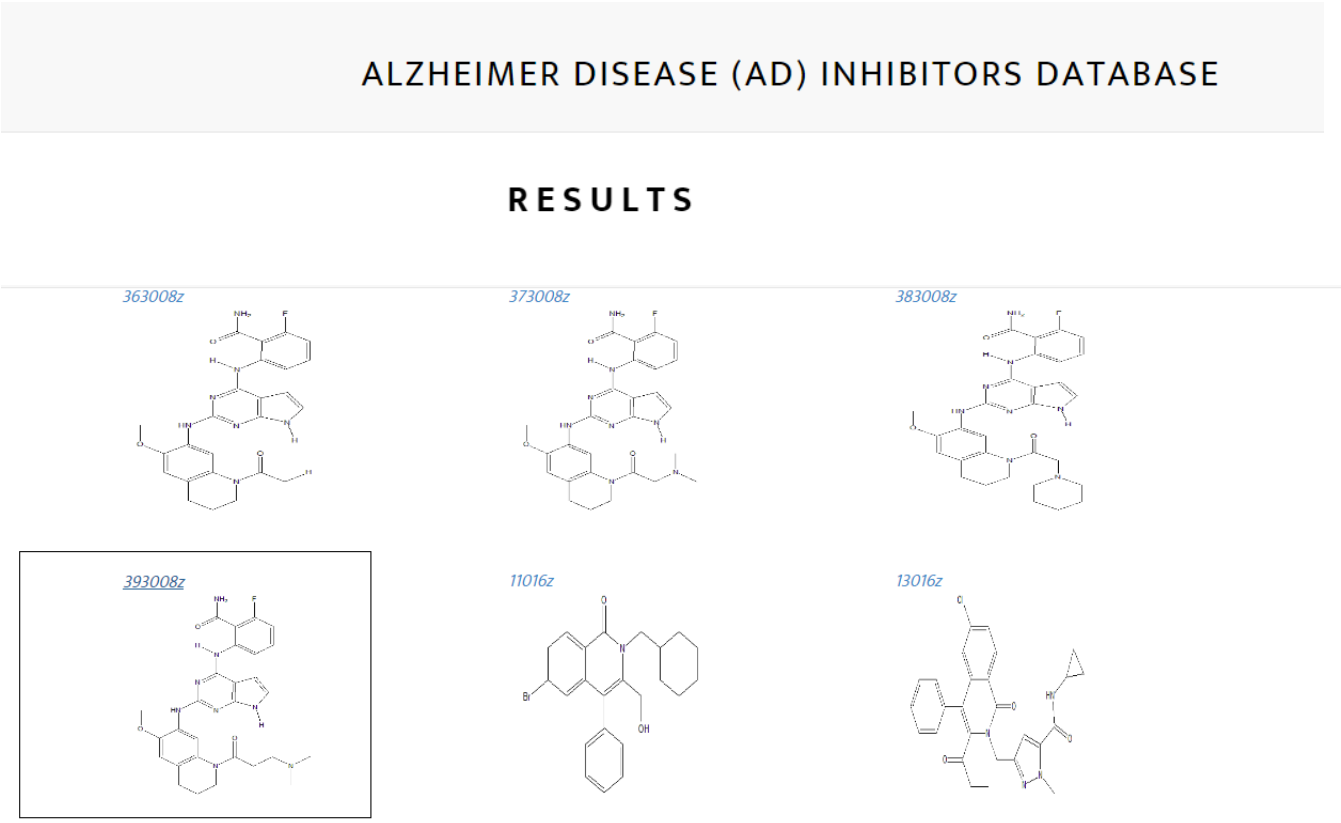


Fig. 3.11 : Result having molecules less than 0.50 identity with uploaded molecule

The above Fig. 3.11 showing the list of molecules having similarity lesser than .50 with the uploaded molecule after performing similarity searching between upload molecule and molecules in database based on tannimoto coefficient.

AlzID

ALZHEIMER DISEASE (AD) INHIBITORS DATABASE

RESULTS

Molecule's Structural IDs	393008z
Parent Structure Name	
JNK1 IC50	7.943
JNK1 pIC50	-
JNK1(Radiometric Filter Binding Assay)	
IGF-1R	0.01
Recombinant-Human JNK1	
JNK2 IC50	-
JNK2 pIC50	-
JNK2(Radiometric Filter Binding Assay)	
JNK2(Fold-shift b)	

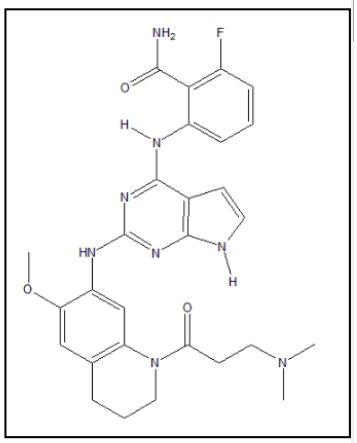


Fig. 3.12 : Results showing Activity data of selected inhibitor against various enzymes.

AlzID

ALZHEIMER DISEASE (AD) INHIBITORS DATABASE

PHYSICAL PROPERTIES

Formula	C28H31FN8O3
LogD	2.77270985
LogP	4.519470404
Molecular weight	546.607
Strongest acidic pKa	12.36736349
Strongest basic pKa	9.137255327
TPSA	141.5
Bemis Murcko Framework	C(C1CCC2CCCCC2C1)C1CC2CCCC2C(CC2CCCCC2)C1
Composition	C (61.53%), H (5.72%), F (3.48%), N (20.5%), O (8.78%)
IUPAC Name	2-{[2-({1-[3-(dimethylamino)propanoyl]-6-methoxy-1,2,3,4-tetrahydroquinolin-7-yl}amino)-7H-pyrrolo[2,3-d]pyrimidin-4-yl]amino}-6-fluorobenzamide
Smiles	COc1cc2CCCN(C(=O)CCN(C)C)c2cc1Nc1nc(Nc2cccc(F)c2C(N)=O)c2cc[nH]c2n1
Asymmetric atoms	0

Atom count	71
Bond count	75
Chiral atoms	0
H bond acceptors	8
H bond donors	4
Ring count	5
Rotatable bonds	9

Fig. 3.13 : Result showing Molecular Properties of selected molecule.

On QUICK SEARCH page users can search for the library of molecules along with the information about those molecules for their desired targets (eg. User can choose "JNK1 and JNK3 Specific" to search for the library of molecules that are specific to both JNK1 and JNK3 (Fig. 3.14)). We got 114 JNK1 specific, 134 JNK3 specific, 155 JNK1 AND JNK2 specific, 192 JNK1 AND JNK3 specific, 90 JNK2 AND JNK3 specific molecules.

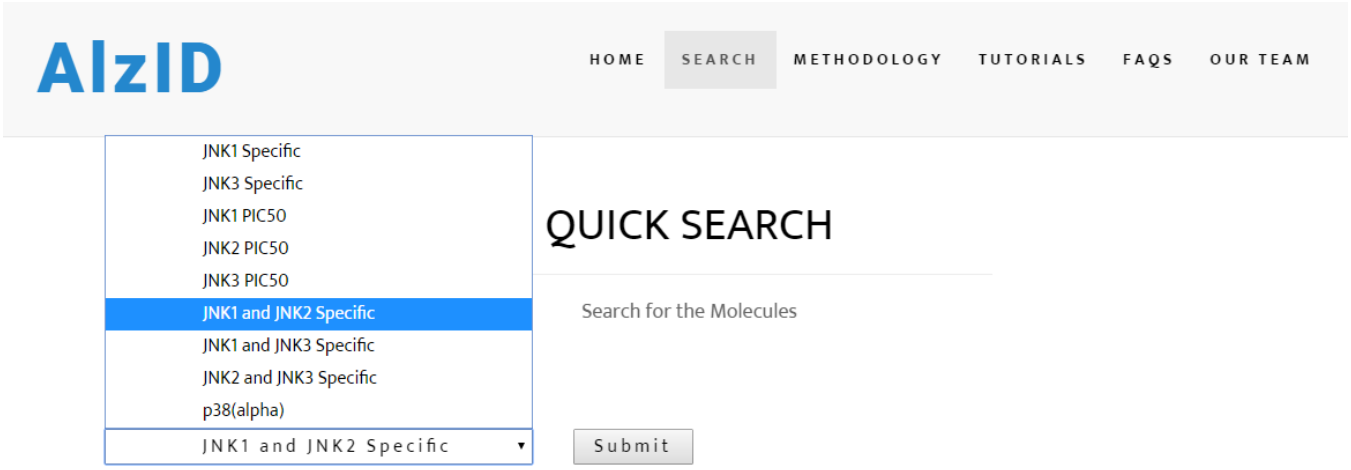


Fig. 3.14 : Quick search page providing library searching specific to particular enzyme

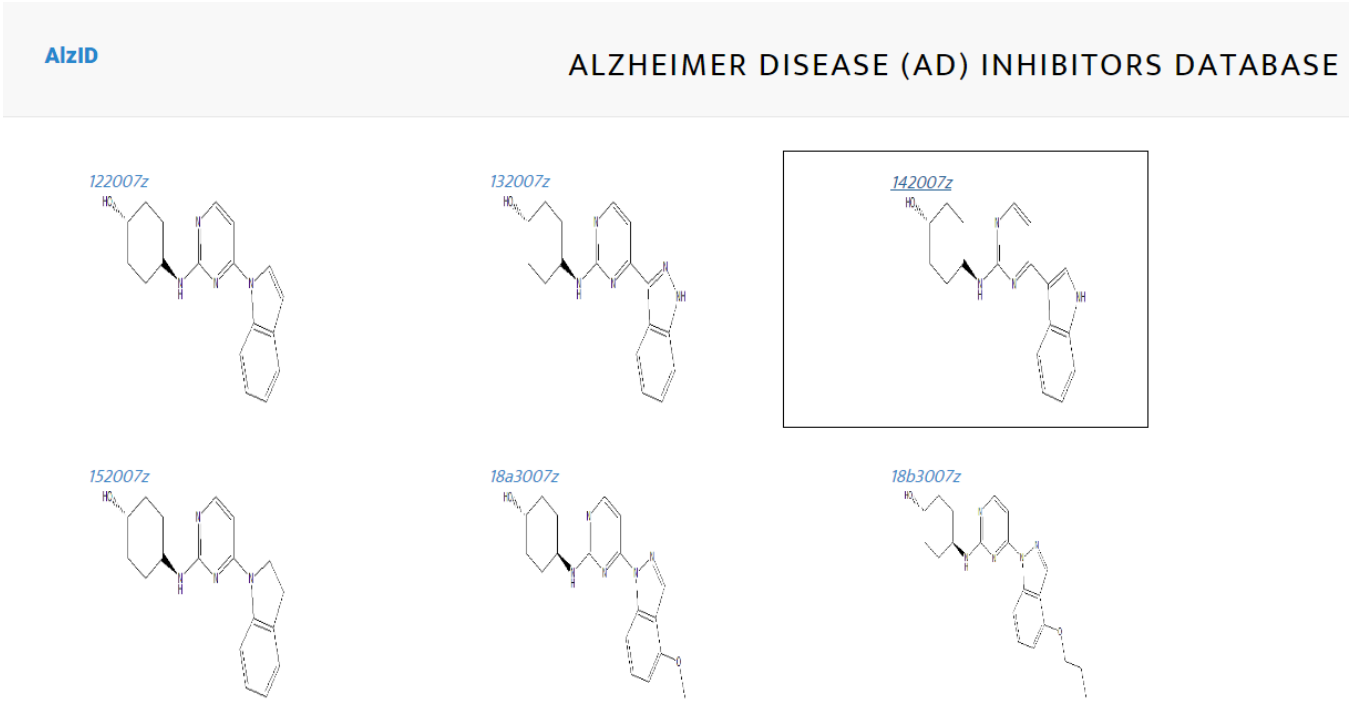


Fig. 3.15 : Result showing list of inhibitors specific to JNK1 AND JNK2

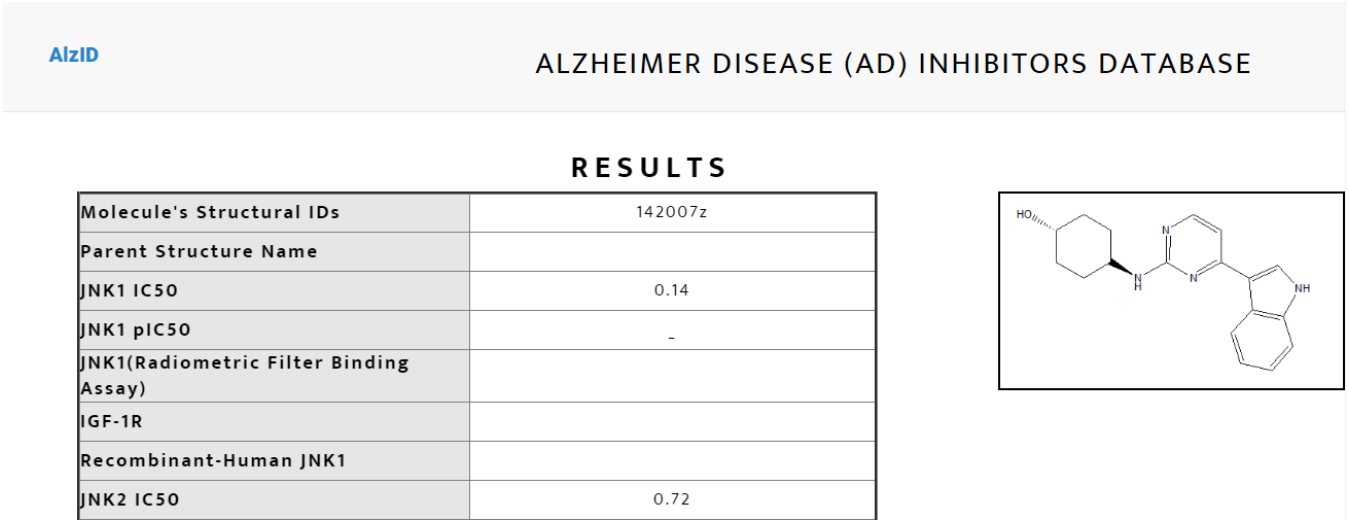


Fig. 3.16 : Activity data of selected inhibitor against different enzymes

PHYSICAL PROPERTIES

Formula	C18H20N4O
LogD	2.745633346
LogP	2.745689291
Molecular weight	308.385
Strongest acidic pKa	14.51663032
Strongest basic pKa	3.52428018
TPSA	73.83
Bemis Murcko Framework	C(C1CCCCC1)C1CCCC(C1)C1CCC2CCCCC12
Composition	C (70.11%), H (6.54%), N (18.17%), O (5.19%)
IUPAC Name	(1r,4r)-4-{[4-(1H-indol-3-yl)pyrimidin-2-yl]amino}cyclohexan-1-ol
Smiles	O[C@H]1CC[C@@H](CC1)Nc1nccc(n1)-c1c[nH]c2ccccc12
Asymmetric atoms	0

Atom count	43
Bond count	46
Chiral atoms	2
H bond acceptors	4
H bond donors	3
Ring count	4
Rotatable bonds	3

Fig. 3.17 : Result showing Molecular Properties of selected molecule

User can also search for library of molecules based on the ring of their choice.

METHODOLOGY

Methodology page contains information regading the steps used for generation and compilation of data to further integrate it to data resource “AlzID”. The steps include : structure & activity data collection, geometry optimization, fragmentation, fragment analysis and database development.

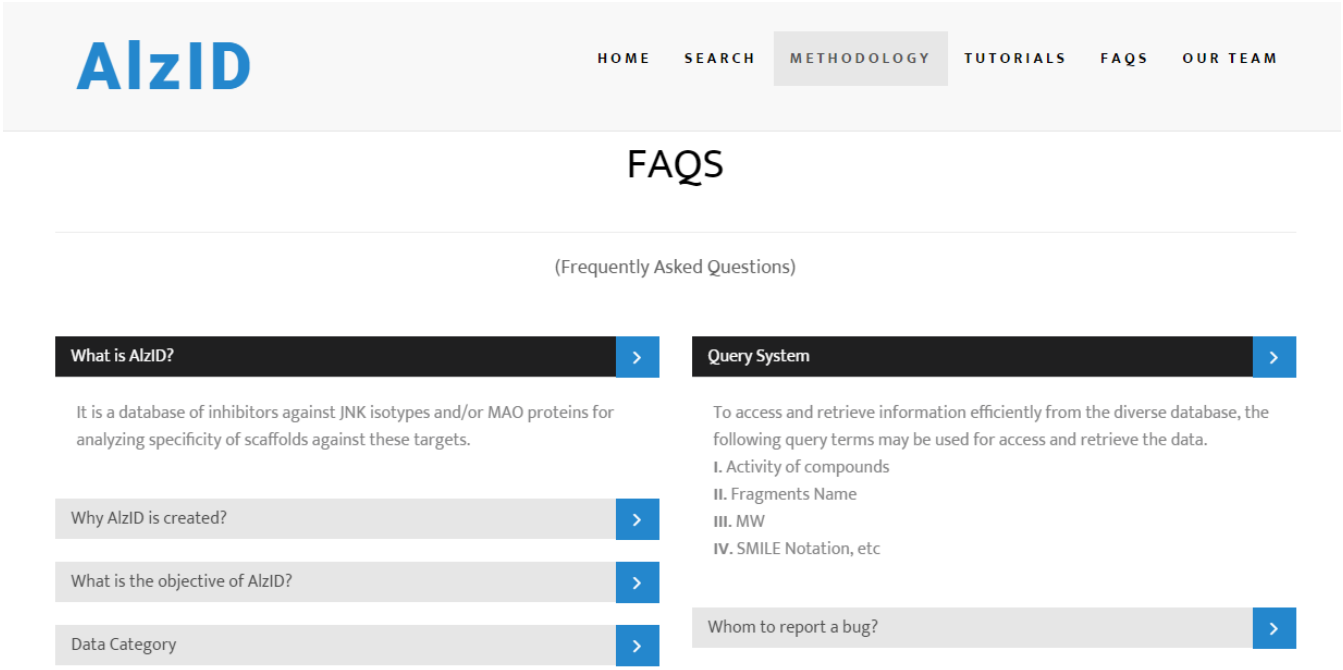


Fig. 3.18 : Frequently asked questions regarding AlzID

The above figure showing the "FAQS (Frequently Asked Questions)" page that contain list of questions with answers that can be asked.

CHAPTER – 4

FRAGMENT ANALYSIS

We mined our database of about 650 compounds which resulted over 2100 rings, 1300 linkers and 2400 side-chains. Fragment occurrence is very bias, with 70% of fragments occurring only once and a few fragments such as pyrimidine, piperazine, indoline, thiophene and piperidine etc being present in many molecules. We identified the fragments that were occurring more than random in dataset. We took 7 as threshold in ring system and we identified set of common and unique fragment that were occurring 7 or more times.

In case of side-chain and linker, we took 3 as threshold as the frequency of occurring unique fragments was very few. So based on this threshold we identified set of common and unique fragments for JNK isotypes.

Table 4.1 : Scaffolds found to be unique against JNKs

Scaffold	Specific Enzyme	Frequency
2,7-phenanthroline	JNK3	10
1H,4H,9H-pyrazolo[3,4-b]quinoline	JNK1	38
2,4-dihydro-1,3-benzodioxine	JNK3	12
Pyridazine	JNK3	16

Table 4.2 : Scaffolds found to be common against JNKs :

Scaffold	Specific Enzyme	Frequency
1,4-dihydroquinoline	JNK1/2	12
1,4-dihydro-1,8-naphthyridine	JNK1/2	8
Isoxazole	JNK1/3	20
Thiophene	JNK1/2/3	71
Isoquinoline	JNK1/2/4	10
Furan	JNK1/3	35
1,4-dioxa-8-azapiro[4.5]decan	JNK1/3	12

Table 4.3 : Side-chains that are unique against JNKs

Sidechain	Specific Enzyme	Frequency
NO	JNK3	50
N(=O)=O()	JNK3	4

Table 4.4 : Side-chains that are common against JNKs

Sidechain	Specific Enzyme	Frequency
C(CO)NC=O	JNK1/2	3
CS(=O)(=O)CCCO	JNK1/2	9
[C-]#[NH+]	JNK1/2/3	5
CCOC=O	JNK1/2/3	6
CC#C	JNK1/3	3

Table 4.5 : Linkers found to be unique against JNKs

Linker	Specific Enzyme	Frequency
CCCN	JNK3	7

Table 4.6 : Linkers found to be common against JNKs

Linker	specificity	Frequency
CCCO	JNK1/2/3	4

CHAPTER – 5

REFERENCES

I. Weber, Lutz. "JChem Base-ChemAxon." *Chemistry World* 5.10 (2008): 65-66.

II. Ultra, ChemDraw. "6.0 and Chem3D Ultra." *Cambridge Soft Corporation, Cambridge, USA* (2001).

III. OEChem, T. K. "OpenEye Scientific Software." *Inc., Santa Fe, NM, USA* (2012).

IV. Csizmadia, P. "MarvinSketch and MarvinView: molecule applets for the World Wide Web." *Proceedings of ECSOC-3, The Third International Electronic Conference on Synthetic Organic Chemistry, September 1ą30*. 1999.

V. Crews, Leslie, and Eliezer Masliah. "Molecular mechanisms of neurodegeneration in Alzheimer's disease." *Human molecular genetics*(2010): ddq160.

VI. Barres, Ben A. "The mystery and magic of glia: a perspective on their roles in health and disease." *Neuron* 60.3 (2008): 430-440.

VII. McKhann, Guy M., et al. "The diagnosis of dementia due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease." *Alzheimer's & dementia* 7.3 (2011): 263-269.

VIII. Bond, Mary, et al. "The effectiveness and cost-effectiveness of donepezil, galantamine, rivastigmine and memantine for the treatment of Alzheimer's disease (review of Technology Appraisal No. 111): a systematic review and economic model." (2012).

IX. Petersen, Ronald C., et al. "Vitamin E and donepezil for the treatment of mild cognitive impairment." *New England Journal of Medicine* 352.23 (2005): 2379-2388.

X. Pajouhesh, Hassan, and George R. Lenz. "Medicinal chemical properties of successful central nervous system drugs." *NeuroRx* 2.4 (2005): 541-553.

XI. Leeson, Paul. "Drug discovery: Chemical beauty contest." *Nature* 481.7382 (2012): 455-456.

XII. Borsello, Tiziana, and Gianluigi Forloni. "JNK signalling: a possible target to prevent neurodegeneration." *Current pharmaceutical design* 13.18 (2007): 1875-1886.

XIII. Früh, V., R. Heetebrij, and G. Siegal. "Fragment-Based Drug DiscoVery: A Practical approach." (2008): 135.

XIV. Spessard, Gary O. "ACD Labs/LogP dB 3.5 and ChemSketch 3.5." *Journal of chemical information and computer sciences* 38.6 (1998): 1250-1253.

XV. Visualizer, Discovery Studio. "Accelrys software inc." *Discovery Studio Visualizer* 2 (2005).

XVI. Stewart, James J. *MOPAC manual. A general molecular orbital package*. No. FJSRL-TR-90-0004. FRANK J SEILER RESEARCH LAB UNITED STATES AIR FORCE ACADEMY CO, 1990.

XVII. DeLano, Warren L. "The PyMOL molecular graphics system." (2002).

XVIII. Rahman, Syed Asad, et al. "Small molecule subgraph detector (SMSD) toolkit." *Journal of cheminformatics* 1.1 (2009): 1.

XIX. Moskovsky, A. A., et al. "WebQC: a web-interface for molecular modeling programs." *Biomeditsinskaya khimiya* 50.app1 (2004): 127-132.