Design, Synthesis and Evaluation of Novel Anti-rheumatic Agents

THESIS

Submitted in partial fulfillment

of the requirements for the degree of

DOCTOR OF PHILOSOPHY

by

VADIRAJ KURDEKAR

Under the Supervision of

PROF. HEMANT R. JADHAV

BITS Pilani

Pilani | Dubai | Goa | Hyderabad

BIRLA INSTITUTE OF TECHNOLOGY AND SCIENCE, PILANI

2014

BIRLA INSTITUTE OF TECHNOLOGY AND SCIENCE, PILANI

CERTIFICATE

This is to certify that the thesis entitled "**Design, Synthesis and Evaluation of Novel Anti-rheumatic Agents**" and submitted by Mr. **Vadiraj Kurdekar** ID No **2008PHXF026P** for award of Ph.D. of the Institute embodies original work done by him under my supervision.

Signature of the Supervisor

Name in capital letters : **Prof. HEMANT R. JADHAV**

Designation   : **Associate Professor, Department of Pharmacy,**
        Associate Dean, Academic Research Division,
        Birla Institute of Technology and Science Pilani, Pilani
        Campus, Vidya Vihar, Pilani - 333031
        India.

Date:

Dedicated to My Parents

ABSTRACT

Rheumatoid Arthritis (RA) an autoimmune disease leading to bone and joint damage. With availability of small molecules only for symptomatic treatment of RA, research and development of small molecular inhibitor against promising targets of RA has attracted the attention of scientist's world-over. Of many targets identified for RA, TNF-α Converting Enzyme (TACE) enhances secretion of TNF-α by cleaving pro-TNF-α and Peptidyl Arginine Deiminase 4 (PAD4) is involved in production of auto-antigens of RA.

TNF-α is a key cytokine in RA. The disease severity is well correlated with TNF-α level in RA patients. With the success of anti-TNF-α antibodies in RA treatment, development of TACE inhibitors is a promising strategy to treat RA. From the available crystal structure of TACE (2OI0), virtual database of 70 TACE inhibitors of benzophenone and urea series was designed using structure based drug design approach. Promising designed molecules were synthesized and enzyme inhibition assay was carried out. Urea derivatives showed relatively higher activity compared to benzophenone derivatives in in vitro assay. In vitro-in silico correlation showed that solubility is an important factor in determining in vitro activity.

With the understanding of PAD4 mechanism of action and other arginine mimetic inhibitors, attempt was made to design reversible arginine mimetic inhibitors. Using computational tools and with knowledge base of 642 arginine mimetic compounds, a new pharmacophore was proposed, based on which possible reversible PAD4 inhibitors belonging to benzamide and benzamidine series were designed. After in silico evaluation, active compounds were synthesized and screened for in vitro PAD4 inhibitory potential. In *p*-(aminomethyl)-benzamide series many compounds showed moderate activity and *N*-[4-(aminomethyl)phenyl]-2-(2-methoxyphenyl)acetamide (M2MBN) was identified as most potent reversible PAD4 inhibitor. Reason for enhanced binding of M2MBN to PAD4 was further studied by molecular simulation. 16 ns simulation showed that protein ligand complex is stable and benzyl group of M2MBN forms additional pi-pi interaction with Histidine 471 leading to enhanced activity.

Keywords: Arginine mimetic, Benzamide, Benzamidine, Benzophenone, Hydroxamic acid, Rheumatoid Arthritis, PAD4, TACE.

ACKNOWLEDGEMENTS

Date: Vadiraj Kurdekar

LIST OF TABLES

LIST OF ILLUSTRATIONS

Figure

Schemes

LIST OF ABBREVIATIONS

+ve	positive
-ve	negative
5-LOX	5-lipoxygenase
°C	degree centigrade
μl	micro liter
Abs	antibodies
ACPAs	Anti Citrullinated Peptide Antibodies
ACR	American College of Rheumatology
ADAM17	A Disintegrin and metalloproteinase domain-containing protein 17
Ala	Alanine
$AlCl_3$	Aluminum chloride
APC	Antigen Presenting Cells
Arg	Arginine
Asp	Aspartic acid
BAA	Benzoyl-L-Arginine Amide
BOC	t- butoxycarbonyl
CD40	Cluster of Differentiation 40
CD40L	Cluster of Differentiation 40 Ligand
CD138+	Cluster of Differentiation 138 positive
CRP	C-reactive protein
CPK model	Robert Corey, Linus Pauling and Walter Koltun model
COX-2	Cyclooxygenase-2
CTLA4	Cytotoxic T-Lymphocyte Antigen 4
DALY	Disability-Adjusted Life Year
DCM	Dichloro methane
DCs	Dendritic Cells
DIAD	Diisopropyl azodicarboxylate
DMARDs	Disease-Modifying Anti-Rheumatic Drugs
DMSO	Dimethyl sulfoxide
EGFR	Epidermal Growth Factor Receptor
ESR	Erythrocyte Sedimentation Rate
Ex/Em	Excitation/Emission
FLSs	Fibroblasts like Synoviocytes
FDA	Food and Drug Administration
FRED	Fast Rigid Exhaustive Docking
FRET	Fluorescence Resonance Energy Transfer
g	gram
GA	Genetic Algorithm
GA-LS	Genetic Algorithm and Lamarckian Search
GI	GastroIntestine
GM-CSF	Granulocyte Macrophage Colony-Stimulating Factor
GMSF	Guanidine-group modifying enzyme superfamily
GSK	GlaxoSmithKline
iNOS	inducible Nitric Oxide Synthase

JCLogP	Jchem log partition coefficient
HBA	hydrogen bond acceptor
HBD	hydrogen bond donor
HBr	Hydrogen Bromide
His	Histidine
HLA	Human Leukocyte Antigen
HLA-DRB	HLA-DR beta chain
hr	hour
IC_{50}	half maximal inhibitory concentration
ICAM	IL-6 receptor and Intercellular Adhesion Molecule
IFN-γ	interferon-γ
Ig G	Immunoglobulin G
Ig M	Immunoglobulin M
IL	interleukin
IR	Infra Red
JNK	c-Jun N-terminal kinase
K	kelvin
kcal	kilocalorie
kDa	kilo Dalton
ki	kinetics of inhibition
Lts	Liters
mg	milligram
min	minutes
ml	milli Liters
mM	milli Moles
MMFF	Merck Molecular Force Field
MM-GBSA	Molecular mechanics with generalised Born and surface area solvation
MMPs	Matrix metalloproteinases
mol	mole
MP	Melting Point
mRNA	Messenger RNA
mTNF-α	membrane bound form of TNF-α
MW	Molecular Weight
$NaHCO_3$	Sodium bicarbonate
NF-κB	Nuclear Factor kappa-light-chain-enhancer of activated B cells
Nk cells	Natural killer cells
nM	nano mole
NMR	Nuclear Magnatic Resonance
NPT	constant number (N), pressure (P), and temperature (T)
NVT	constant number (N), volume (V), and temperature (T)
ns	nanosecond
NSAIDs	Non-steroidal anti-inflammatory drugs
OA	Osteoarthritis
OPLS	Optimized Potentials for Liquid Simulations
PAD	Peptidyl Arginine Deiminase
PDB	Protein Data Bank

PEI	Potency Efficacy Index
PK/PD	Pharmacokinetic/Pharmacodynamic
ppm	part per million
PTPN22	Protein Tyrosine Phosphatase, non-receptor type 22
RA	Rheumatoid Arthritis
RAF	Rapidly Accelerated Fibrosarcoma
RF	Rheumatoid Factor
RMSD	Root Mean Square Deviation
s	second
SBDD	Structure Based Drug Design
SNP	Single Nucleotide Polymorphism
TACE	TNF-α Converting Enzyme
TGF-α	Transforming growth factor alpha
Th cells	helper T cells
THF	tetrahydrofuran
TLC	Thin Layer Chromatography
TMS	trimethylsilane
TNF-α	Tumor Necrosis Factor Alpha
TNF-R1	TNF receptor type 1
TNF-R2	TNF receptor type 2
TPSA	Topological Polar Surface Area
TRADD	TNF receptor-Associated Death Domain
TRAF2	TNF receptor-associated factor 2
sTNF-α	soluble form of TNF-α
US	United States
VSGB	Variable Dielectric Surface Generalized Born
WB assay	Whole Blood assay
WHO	World Health Organization

TABLE OF CONTENTS

Chapter 1

Introduction to Rheumatoid Arthritis (RA)

Recognition of rheumatic diseases may have begun as long as 2,400 years ago when the writings of Hippocrates suggested ailments of the joints. The term 'rheumatos' means - 'flowing', '-oids' means 'resembling', 'arthro' means joint and 'itis' means 'inflammation'. These terms were introduced in the first century AD and came from an ancient concept that ailments were caused by substances that flowed through the body, or 'humours'. Both Sushruta (11th century BC) and Charaka (8th century BC) described this disease as '*Amavata*'. In sanskrit '*Ama*' means 'undigested food'. '*Vata*' is one of the dosha associated with air [1]. The disease condition in *Ayurveda* was described as

" *Sa Kashtaha Sarvarogaanaam yadaa prakupito Bhavet |*
Hastha paada shiro gulpha trika jaanu uru sandhishu ||
Karoti sarujam shopham yatra doshaha prapadyate |
Sa desho rujate atyatham vyaviddha iva vrishchkaihi || "

It means when *Amavata* gets aggravated the joints (*sandhi*) of hands (*hastha*), feet (*paada*), ankles and elbow (*gulpha*), lower back (*trika*), knee (*jaanu*) and hip (*uru*) becomes inflamed and painful. The pain in the affected joints resembles the pain of a scorpion's sting [2].

Rheumatoid Arthritis (RA) is an autoimmune systemic inflammatory disease, characterized by destruction of cartilage and inflammation of the synovium and substantial disability, if not treated early [3]. The metacarpophalangeal and wrist joints are primarily affected followed by knees. As the disease progresses, osteoporosis, damage to joint cartilage, bone displacement and reabsorption occurs.

1.1 Global epidemiology

The estimated global prevalence is 1% [4]. In developed countries it varies from 0.3 to 1% with average of 0.8% in adults. Native Americans have a high prevalence rate of 5.3-6.0% and Africans have low rate of 0-0.3% [5] which clearly supports genetic as well as life-style related origin of this disease. In India, prevalence is about 0.8% for females and 0.13% for male with total accounting for 0.93% in adult population [6]. In China, Indonesia, and the Philippines, the rate is at 0.4% both in urban and rural population [7].

Actual cause for RA is still unknown but it has been hypothesized that environmental factors triggers RA in genetically susceptible individuals. RA susceptibility and severity are also linked with Human Leukocyte Antigen (HLA)-DRB1 and Protein Tyrosine Phosphatase,

non-receptor type 22 (PTPN22) gene [8]. The environmental factors include smoking, viral infections (Epstein-Barr virus, Parvo virus), bacterial infections (Proteus, Mycoplasma species) and stress factors that affect immunity [9].

1.2   Mortality in RA

RA is a chronic disorder, with poor prognosis, there is increased mortality and morbidity and life expectancy decreases by 3-10 years [10]. Mortality is associated with co-morbid conditions like cardiovascular diseases [11], respiratory diseases [12] and depression. The pain, economical burden, severe side effects of drugs and loss of work capacity leads to depression followed by suicide. It is reported that 52.6% of the female RA patients committed suicide and 90% of them suffered by co-morbid depressive disorders [13].

1.3   Socio economical problems associated with RA

RA is recognized internationally not only as problem that impairs individual lives, but also as problem that affects national economies. The WHO and 37 countries had proclaimed year 2000 to 2010 as the bone and joint decade. The total cost of treatment increases with duration of disease [14]. Not only treatment is costly, patients of working age experience a diminished work capacity followed by total work disability [15]. Age-standardized disability-adjusted life year (DALY) rates for RA from developed countries are high than developing countries [16] (as given in **Figure 1-1**).
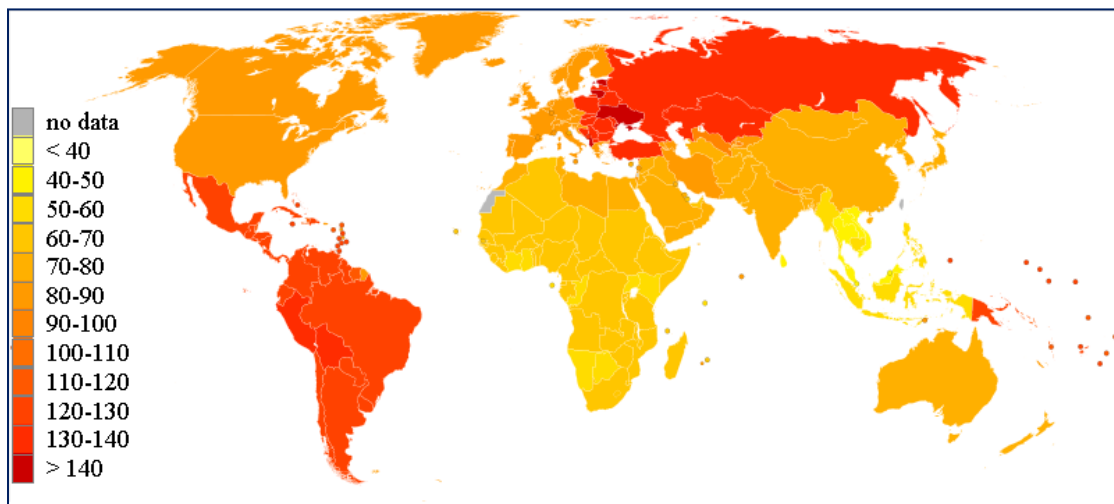
Figure 1-1: Age-standardized disability-adjusted life year (DALY) rates from Rheumatoid arthritis by country (per 100,000 inhabitants) in 2004 [17].

The DALY rate is lowest in Africa due to low prevalence rate. The DALY is a measure of overall disease burden, expressed as the number of years lost due to ill-health, disability or early death. Physical disability and chronic pain have profound effect on psychological

behavior as well as on social and personal aspects of life by impacting individual performance [18].

1.4 Pathophysiology of RA

In genetically susceptible individuals, self proteins are recognized by immature Dendritic Cells (DCs) as antigen. It has been reported that citrullinated proteins- the catalytic products of Peptidyl Arginine Deiminase (PAD) can also act as self antigen [19]. After antigen engulfment, DCs mature into Antigen Presenting Cells (APCs) and activate T cells. T cells, in turn, activate B cells and trigger autoimmunity against these antigens, which are bodies own proteins. The activated B cells in synovium enhance adaptive immunity in joints. The acute inflammation in synovium then leads to chronic inflammation, a pathological condition in RA. The RA synovium contains lymphoid centers at perivascular region comprising of mainly T cells and B cells [20].

For immune reaction, the trigger can be antigen specific or antigen nonspecific. In antigen nonspecific mechanism, activated APCs in the lympoidal centers trigger CD4+ T cells to secret Interleukin 2 (IL-2). Secreted IL-2 causes T cell differentiation and proliferation to helper T (Th) cells such as Th1, Th2 and Th17 [21]. Matured Th cells secrete cytokines that influence angiogenesis, migration, apoptosis, cell growth and differentiation in synovium [22]. Th1 cells express proinflammatory cytokines like TNF-α, IL-2, IL-12, Granulocyte Macrophage Colony-Stimulating Factor (GM-CSF), interferon-γ (IFN-γ) and Cluster of Differentiation 40 Ligand (CD40L). Th2 cells produce anti-inflammatory cytokines like IL-4, IL-5, IL-10, IL-13, CD40L and stimulate B cells to produce antibodies. Th17 cells secret mainly IL-17 [23, 24]. In RA, balance between Th1 and Th2 cytokines gets disturbed. Uncontrolled production of IL-17 along with IL-1 and TNF-α lead to stimulation of synovial cells to produce Matrix Metalloproteases (MMPs) and other proinflammatory cytokines which acts on osteoclast leading to bone erosion [25]. In an antigen specific mechanism, B cells in synovial lympoidal center, on activation, get differentiated into plasma cells (Cluster of Differentiation 138 positive (CD138+)) and produce antibodies such as Rheumatoid Factor (RF) or Anti Citrullinated Peptide Antibodies (ACPAs) [26]. These auto-antibodies bind to specific antigens forming immune complex, which binds with complement or Fc receptors on macrophages triggering the release of cytokines (like TNF-α and interleukin-1, 6, 12, 15, 18, and 23), reactive oxygen intermediates (ROS) and phagocytosis of immune complex. Released cytokines also activate neutrophils, mast cells and Natural killer (Nk) cells contributing towards synovitis [27-29]. These cytokines also activate Fibroblasts like Synoviocytes (FLSs), chondrocytes and osteoclasts. FLSs express high level of chemokines,

adhesion molecules and MMPs particularly MMP-1, 3, 8, 13, 14, and 16 leading to cartilage destruction and severity of synovitis [30]. IL-1 and IL-17 secreted by FLSs and macrophages activate chondrocytes, activation further amplified by TNF-α [31]. Activated osteoclast releases several proteases of cathepsin and MMP family and hydrogen ions leading to destruction of mineral tissue of bone. The resorption pits created by chondrocytes get filled by inflammatory mediators and demineralization slowly progresses up to bone marrow allowing access of synovial tissues to bone marrow [32]. This leads to chronic condition of RA.

1.5   Diagnosis of RA

American College of Rheumatology (ACR) in 1958 presented diagnosis criteria for RA that was revised in 1987 to prevent misdiagnosis of other form of arthritis as RA [33]. ACR-1987 diagnostic criteria has low sensitivity and specificity for early RA, while in established RA it has low sensitivity and but high specificity. Higher specificity was achieved with use of radiography techniques. With identification of the role of ACPAs, measurement of its level in serum has increased sensitivity and specificity in early RA there by it has become one of the laboratory test to classify RA in ACR-2010 report [34]. Patients who have at least 1 joint with definite synovitis or a patient with synovitis that is not better explained by another disease has to be diagnosed for RA. The criteria given by ACR-2010 report and European League Against Rheumatism (EULAR) help to identify patients having definite RA (see **Table 1-1**).

Table 1-1: The 2010 ACR/European League Against Rheumatism (EULAR) classification criteria for RA [34].

Observations	Score
A. Joint involvement	
One Large joint	0
2-10 large joints	1
1-3 small joints (with or without involvement of large joints)	2
4-10 small joints (with or without involvement of large joints)	3
>10 joints with at least 1 small joint	5
B. Serology (at least one positive test result needed)	
Negative Rheumatoid Factor (RF) and negative Anti-Citrullinated Protein Antibody(ACPA)	0
Low positive RF or low Positive ACPA	2

High positive RF or high Positive ACPA	3
C. Acute phase reactants (at least one test result needed)	
Normal C-reactive protein (CRP) and normal Erythrocyte Sedimentation Rate (ESR)	0
Abnormal CRP or abnormal ESR	1
D. Duration of symptoms	
< 6 weeks	0
≥ 6 weeks	1

Add scores of categories A to D. Score of ≥6/10 is needed to classify patient as having definite RA. Large joint means joint of shoulders, elbows, hips, knees, and ankles; Small joint means Proximal Interphalangeal (PIP), 2nd to 5th Metatarsophalangeal (MCP) joints, thumb interphalangeal joints and wrists joints.

1.6 WHO International Classification of Diseases-10 (ICD-10) of RA

The International Statistical Classification of Diseases and Related Health Problems 10th Revision (ICD-10) is a coding of diseases and signs, symptoms, abnormal findings, complaints, social circumstances and external causes of injury or diseases, as classified by the World Health Organization (WHO) [35].

RA is given coding in ICD-10 from M05-M06. Some of the codes are

- M05   Seropositive RA
    - M05.1   Rheumatoid lung disease
    - M05.2   Rheumatoid Vasculitis
    - M05.3   RA with involvement of other organs and systems
    - M05.9   Seropositive RA, unspecified
- M06   Seronegative RA
    - M06.1   Adult onset still disease
    - M06.3   Rheumatoid nodule
    - M06.4   Inflammatory Polyarthropathy
    - M06.9   RA, unspecified

1.7 Approved Drugs for the treatment of RA

Table 1-2 and **1-3** lists currently FDA approved drugs for RA along with their side effects. Non-steroidal anti-inflammatory drugs (NSAIDs) and glucocorticoids are the drugs of choice for the symptomatic treatment of RA. However major side effects associated with NSAIDs, like gastric ulcer, GastroIntestine (GI) bleeding, etc limits their use in chronic treatment of RA. COX-2 inhibitors, though devoid of GI side effects, have a concern that their

long term use may increase the risk of stroke and heart attack. Low dose of glucocorticoids controls inflammation and pain but long term use causes bone loss. Disease-Modifying Anti-Rheumatic Drugs (DMARDs) are the first line treatment for RA. DMARDs, act slowly, but substantially reduce inflammation, reduce or prevent joint damage and enable patients to continue daily activity. Based on disease severity and patient response, a second line treatment with biological anti-rheumatic agents is used to prevent progress of the disease. Biological agents, mainly monoclonal antibodies (MAbs), act within 2 weeks of medication. However they are expensive and bring about severe reduction in immunity leading to susceptibility to infections.

Table 1-2 : List of currently approved non biological drugs for RA [36]**.**

Drug	Action	Dose	Side effects
DMARDS (Disease-Modifying Anti-Rheumatic Drugs)			
Sulphasalazine	Anti-inflammatory	1,000–1,500 mg twice daily (oral)	Hypersensitivity, Thrombocytopenia
Methotrexate (MTX)	Anti-metabolite	10–25 mg weekly (oral or subcutaneous)	Hepatotoxicity, Neutropenia
Hydroxychloroquine (HCQ)	Interferes with antigen-processing	200–400 mg daily (oral)	Retinopathy
Leflunomide (LEF)	Anti-metabolite	10–20 mg daily (oral)	Hepatotoxicity, Myelotoxicity, Hypertension
Gold salts (sodium aurothiomalate)	unknown	50 mg weekly (intramuscular)/ 3–6 mg daily (oral)	Hypersensitivity, Diarrhoea, Nephritis
Cyclosporine A	T-cell activation inhibitor	2.5–5.0 mg/kg per day (oral)	Nephrotoxicity, Hypertension
Azathioprine	Cytostatic	50–200 mg daily (oral)	Hepatotoxicity, Myelotoxicity
Minocycline	Anti-inflammatory /antibiotic	200 mg daily (oral)	Sensitizes skin to sunlight

Table 1-3: List of currently approved biological drugs for RA [36]**.**

Drug	Action	Dose	Side effects
Biological Agents: Anti TNF-α agents			
Infliximab (Chimeric MAb)	TNF-α blockade	5 mg/kg as intravenous infusion regimen at 0, 2, and 6 weeks followed by a maintenance regimen of 5 mg/kg every 8 weeks	Infection, headache, hepatotoxicity
Adalimumab (Human MAb)	TNF-α blockade	40 mg biweekly (subcutaneous)	Injection site reaction, infection
Golimumab (Human MAb)	TNF-α blockade	100 mg every 4 weeks (subcutaneous)	Injection site reaction, infection
Certolizumab (Pegylated Fab' fragment from humanized MAb)	TNF-α blockade	200 mg biweekly or 400 mg monthly (subcutaneous)	Injection site reactions, infection
Etanercept (Recombinant TNF receptor (p75) dimerized on Ig frame)	TNF-α blockade	50 mg weekly (subcutaneous)	Injection site reaction, infection
Biological Agents: Non Anti-TNF-α agents			
Rituximab (Chimeric MAb)	B-cell depletion	1,000 mg, 2 intravenous infusions 2 weeks apart	Injection site reactions, infections
Abatacept (Recombinant CTLA4 molecule dimerized on Ig frame)	T-cell co-stimulation blocker	500–1,000 mg monthly (intravenous)	Injection site reaction, infection
Tocilizumab (Humanized MAb)	IL-6-receptor blockade	8 mg/kg every 4 weeks (intravenous)	Injection site reaction, infection, elevated cholesterol
Anakinra (Recombinant human Interleukin-1 receptor antagonist)	IL-1 receptor blockade	100 mg daily (subcutaneous)	Injection site reaction, infection, neutropenia

1.8   Current Treatment strategy of RA

Currently there is no complete cure for RA, current therapy only reduces or stops progress of the disease by symptomatic treatment of pain, stiffness, swelling and preventing structural damage, thereby extending the term of disability and death. Based on disease stage, prognosis, patient response, cost and patient's susceptibility to opportunistic infections, different treatment strategies are used. ACR-2010 has recommended and listed the facts to be considered while treating RA as elucidated in **Figure 1-2** [37]. Initial therapy includes DMARDs like Hydroxychloroquine, Leflunomide, Methotrexate, Sulfasalazine, etc. Due to lack of new data, Gold salts, Cyclosporine, Azathioprine, Minocycline are not recommended. Depending on patient's condition, mono or double or triple (Methotrexate, Hydroxychloroquine, Sulfasalazine) DMARDs with or without biologicals like Abatacept, Rituximab, Tocilizumab, Adalimumab, Etanercept, Infliximab, Certolizumab, Golimumab are recommended. Consideration has to be given towards the cost, patient's preference in decision making for biological agents, side effects of DMARDs and biological agents and patient's response to DMARDs. Committee also recommended for Tuberculosis screening and vaccination (for Pneumococcal, Influenza, Hepatitis, Human Papilloma virus and Herpes Zoster) of patient before starting biologicals. It also suggested restricted use of biological agents in patients suffering from hepatitis, congestive heart failure and malignancy.
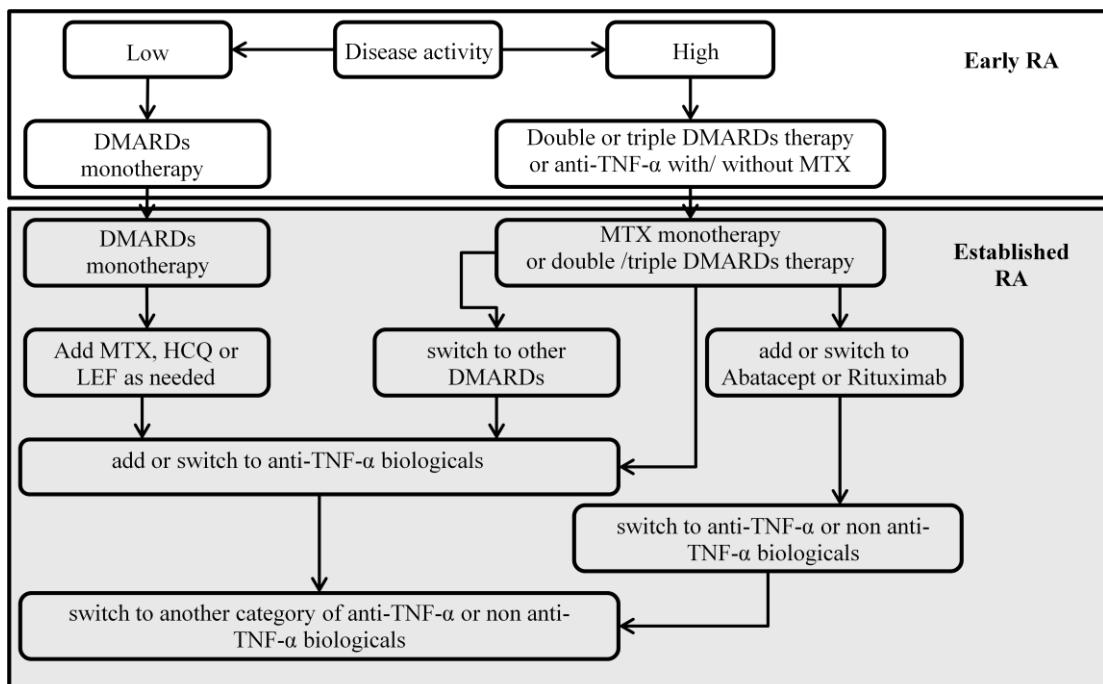
Figure 1-2: ARC-2010 recommended treatment strategy for RA. (LEF= Leflunomide; MTX= Methotrexate; HCQ= Hydroxychloroquine)

1.9   References:

[1]   Lekurwale, P.; Pandey, K.; Yadaiah, P. Management of Amavata with '*Amrita Ghrita*': A clinical study. *Ayu.* **2010,** 31(4), 430-435.

[2]   Rheumatoid Arthritis (*Ama Vata*). http://www.charaka.org/rheumatoid-arthritis.html (Accessed June 28, 2013).

[3]   Abbas, A.K.; Lichtman, A.H.; Pillai, S. *Cellular and molecular immunology*. 6th ed.; Saunders Philadelphia: 2005.

[4]   Dowman, B.; Campbell, R.M.; Zgaga, L.; Adeloye, D.; Chan, K.Y. Estimating the burden of rheumatoid arthritis in Africa: A systematic analysis. *J Glob Health.* **2012,** 2(2),020406.

[5]   Drosos, A.; Alamanos, I.; Voulgari, P.; Psychos, D.; Katsaraki, A.; Papadopoulos, I.; Dimou, G.; Siozos, C. Epidemiology of adult rheumatoid arthritis in northwest Greece 1987-1995. *J Rheumatol.* **1997,** 24(11), 2129-2133.

[6]   Symmons, D.; Mathers, C.; Pfleger, B., *WHO, Global burden and incidence of rheumatoid arthritis* WHO Press: Geneva, 2006.

[7]   Mijiyawa, M. Epidemiology and semiology of rheumatoid arthritis in Third World countries. *Rev Rhum Engl Ed.* **1995,** 62(2), 121-126.

[8]   Silman, A.J.; Pearson, J.E. Epidemiology and genetics of rheumatoid arthritis. *Arthritis Res.* **2002,** 4(Suppl 3), S265-S272.

[9]   Gibofsky, A. Overview of epidemiology, pathophysiology, and diagnosis of rheumatoid arthritis. *Am J Manag Care.* **2012,** 18(Suppl 13), s295-s302.

[10]   Wasserman, A.M. Diagnosis and management of rheumatoid arthritis. *Am Fam Physician.* **2011,** 84(11), 1245-1252.

[11]   Sacks, J.J.; Helmick, C.G.; Langmaid, G. Deaths from arthritis and other rheumatic conditions, United States, 1979-1998. *J Rheumatol.* **2004,** 31(9), 1823-1828.

[12]   Carmona, L.; Hernández-García, C.; Vadillo, C.; Pato, E.; Balsa, A.; González-Alvaro, I.; Belmonte, M.A.; Tena, X.; Sanmartí, R. Increased risk of tuberculosis in patients with rheumatoid arthritis. *J Rheumatol.* **2003,** 30(7), 1436-1439.

[13]   Timonen, M.; Viilo, K.; Hakko, H.; Särkioja, T.; Ylikulju, M.; Meyer-Rochow, V.B.; Väisänen, E.; Räsänen, P. Suicides in persons suffering from rheumatoid arthritis. *Rheumatology (Oxford).* **2003,** 42(2), 287-291.

[14]   Lapsley, H.M.; March, L.M.; Tribe, K.L.; Cross, M.J.; Brooks, P.M. Living with osteoarthritis: patient expenditures, health status, and social impact. *Arthritis Care Res.* **2001,** 45(3), 301-306.

[15]   Katz, P.P. The impact of rheumatoid arthritis on life activities. *Arthritis Care Res.* **1995,** 8(4), 272-278.

[16]   WHO, Disease and injury country estimates, Burden of Disease. http://www.who.int/healthinfo/global_burden_disease/estimates_country/en/index. html (Accessed July 30, 2013).

[17]   Wikipedia. Rheumatoid arthritis world map. http://en.wikipedia.org/wiki/File:Rheumatoid_arthritis_world_map_-_DALY_-_WHO2004.svg (Accessed July 30, 2013).

[18] Newman, S.; Mulligan, K. The psychology of rheumatic diseases. *Baillieres Best Pract Res Clin Rheumatol.* **2000,** 14(4), 773-786.

[19] Lakos, G.; Soós, L.; Fekete, A.; Szabó, Z.; Zeher, M.; Horváth, I.; Dankó, K.; Kapitány, A.; Gyetvai, A.; Szegedi, G. Anti-cyclic citrullinated peptide antibody isotypes in rheumatoid arthritis: association with disease duration, rheumatoid factor production and the presence of shared epitope. *Clin Exp Rheumatol.* **2008,** 26(2), 253-260.

[20] McInnes, I.B.; Schett, G. The pathogenesis of rheumatoid arthritis. *N Engl J Med.* **2011,** 365(23), 2205-2219.

[21] Firestein, G.S. *Etiology and Pathogenesis of Rheumatoid Arthritis*. Elsevier Health Sciences. **2012**; Vol. 2(9), Chapter 65, 1035-1086.

[22] Niu, X.; Miasnikova, G.Y.; Sergueeva, A.I.; Polyakova, L.A.; Okhotin, D.J.; Tuktanov, N.V.; Nouraie, M.; Ammosova, T.; Nekhai, S.; Gordeuk, V.R. Altered cytokine profiles in patients with Chuvash polycythemia. *Am J Hematol.* **2009,** 84(2), 74-78.

[23] Groom, J.R.; Richmond, J.; Murooka, T.T.; Sorensen, E.W.; Sung, J.H.; Bankert, K.; von Andrian, U.H.; Moon, J.J.; Mempel, T.R.; Luster, A.D. CXCR3 Chemokine Receptor-Ligand Interactions in the Lymph Node Optimize CD4+ T Helper 1 Cell Differentiation. *Immunity.* **2012,** 37(6), 1091-1103.

[24] Wahlstrom, J.; Katchar, K.; Wigzell, H.; Olerup, O.; Eklund, A.; Grunewald, J. Analysis of intracellular cytokines in CD4+ and CD8+ lung and blood T cells in sarcoidosis. *Am J Respir Crit Care Med.* **2001,** 163(1), 115-121.

[25] Lubberts, E.; Joosten, L.A.; Chabaud, M.; van den Bersselaar, L.; Oppers, B.; Coenen-de Roo, C.J.; Richards, C.D.; Miossec, P.; van den Berg, W.B. IL-4 gene therapy for collagen arthritis suppresses synovial IL-17 and osteoprotegerin ligand and prevents bone erosion. *J Clin Invest.* **2000,** 105(12), 1697-1710.

[26] Ohata, J.; Zvaifler, N.J.; Nishio, M.; Boyle, D.L.; Kalled, S.L.; Carson, D.A.; Kipps, T.J. Fibroblast-like synoviocytes of mesenchymal origin express functional B cell-activating factor of the TNF family in response to proinflammatory cytokines. *J Immunol.* **2005,** 174(2), 864-870.

[27] Cascao, R.; Rosario, H.; Souto-Carneiro, M.; Fonseca, J. Neutrophils in rheumatoid arthritis: More than simple final effectors. *Autoimmun Rev.* **2010,** 9(8), 531-535.

[28] Nigrovic, P.A.; Lee, D.M. Synovial mast cells: role in acute and chronic arthritis. *Immunol Rev.* **2007,** 217(1), 19-37.

[29] Shegarfi, H.; Naddafi, F.; Mirshafiey, A. Natural killer cells and their role in rheumatoid arthritis: friend or foe? *ScientificWorldJournal.* **2012,** 2012, 491974.

[30] Bradfield, P.F.; Amft, N.; Vernon-Wilson, E.; Exley, A.E.; Parsonage, G.; Rainger, G.E.; Nash, G.B.; Thomas, A.M.; Simmons, D.L.; Salmon, M.; Buckley, C.D. Rheumatoid fibroblast-like synoviocytes overexpress the chemokine stromal cell-derived factor 1 (CXCL12), which supports distinct patterns and rates of CD4+ and CD8+ T cell migration within synovial tissue. *Arthritis Rheum.* **2003,** 48(9), 2472-2482.

[31] Kim, H. TNF-alpha induced chondrocyte apoptosis in NF-κB suppression is augmented by inhibition of p38 mitogen-activated protein kinase or phosphatidylinositol 3-kinase. *Arthritis Res Ther.* **2003,** 5(Suppl 3), 58.

[32] Jimenez-Boj, E.; Redlich, K.; Türk, B.; Hanslik-Schnabel, B.; Wanivenhaus, A.; Chott, A.; Smolen, J.S.; Schett, G. Interaction between synovial inflammatory tissue and bone marrow in rheumatoid arthritis. *J Immunol.* **2005,** 175(4), 2579-2588.

[33] Arnett, F.C.; Edworthy, S.M.; Bloch, D.A.; Mcshane, D.J.; Fries, J.F.; Cooper, N.S.; Healey, L.A.; Kaplan, S.R.; Liang, M.H.; Luthra, H.S. The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. *Arthritis Rheum.* **1988,** 31(3), 315-324.

[34] Aletaha, D.; Neogi, T.; Silman, A.J.; Funovits, J.; Felson, D.T.; Bingham, C.O.; Birnbaum, N.S.; Burmester, G.R.; Bykerk, V.P.; Cohen, M.D. 2010 rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. *Arthritis & Rheum.* **2010,** 62(9), 2569-2581.

[35] WHO, ICD-10 Version:2010. http://apps.who.int/classifications/icd10/browse/2010/en (Accessed june 13 2013).

[36] Van Vollenhoven, R.F. Treatment of rheumatoid arthritis: state of the art 2009. *Nat Rev Rheumatol.* **2009,** 5(10), 531-541.

[37] Singh, J.A.; Furst, D.E.; Bharat, A.; Curtis, J.R.; Kavanaugh, A.F.; Kremer, J.M.; Moreland, L.W.; O'Dell, J.; Winthrop, K.L.; Beukelman, T.; Bridges, S.L.; Chatham, W.W.; Paulus, H.E.; Suarez-almazor, M.; Bombardier, C.; Dougados, M.; Khanna, D.; King, C.M.; Leong, A.L.; Matteson, E.L.; Schousboe, J.T.; Moynihan, E.; Kolba, K.S.; Jain, A.; Volkmann, E.R.; Agrawal, H.; Bae, S.; Mudano, A.S.; Patkar, N.M.; Saag, K.G. 2012 Update of the 2008 American College of Rheumatology recommendations for the use of disease-modifying antirheumatic drugs and biologic agents in the treatment of rheumatoid arthritis. *Arthritis Care Res.* **2012,** 64(5), 625-639.

Chapter 2

Design, Synthesis and Evaluation of TACE

Inhibitors

2.1 Introduction to TACE

The term Tumor Necrosis Factor (TNF) was first coined in 1960s with observation of induced tumor regression. However, in 1984, for the first time the cytotoxic factor isolated from macrophages (MW 17kDa) was named as TNF, later called as TNF-α. Another cytotoxic factor isolated from lymphocytes (MW 20 kDa) was named as lymphotoxin and later called as TNF-β [1]. Till early 1990s, TNF-α was considered as poor therapeutic target due to cytokine redundancy. However, with identification of anti-TNF antibodies and its role in suppression of inflammation [2], TNF-α was used as therapeutic target for drug development in inflammation mediated diseases like RA and Crohn`s disease [3, 4].

2.1.1 Physiological Role of TNF-α

TNF-α is initially synthesized as transmembrane protein called mTNF-α or pro TNF-α (26 kDa) which is cleaved by TNF-α Converting Enzyme (TACE). Under normal physiology, TNF-α exhibits dual functionality by involving in tissue regeneration and destruction [5]. It plays a crucial role in innate and adaptive immunity. TNF-α action is mediated by acting on two distinct receptors namely - TNF receptor type 1 (TNF-R1) or p55 and TNF receptor type 2 (TNF-R2) or p75. On activation of TNF-R1, apoptotic pathways are triggered via TNF receptor-associated death domain (TRADD) protein or the NF-κB (Nuclear Factor kappa-light-chain-enhancer of activated B cells) pathway mediated proinflammatory mediators (Inter leukine-6 (IL-6), IL-8, IL-18, inducible nitric oxide synthase (iNOS), cyclooxygenase-2 (COX-2), and 5-lipoxygenase (5-LOX)) synthesis [1]. TNF-R2 are expressed in few cells like T-cells, some neural cells like astrocytes, microglia and cardiac myocytes [6]. On activation they recruit TNF receptor-associated factor 2 (TRAF2) and activate c-Jun N-terminal kinase (JNK), NF-κB pathways leading to cell survival and proliferation [7].

2.1.2 Physiological Role of TACE

TACE cleaves its major substrate mTNF-α at the residues Ala76 and Val77 to release soluble form of TNF-α (sTNF-α) [8] (**Figure 2.1-1**). TACE also acts on multiple substrates like Epidermal Growth Factor Receptor (EGFR) ligands (Transforming growth factor-α (TGF-α), amphiregulin, epicellulin, epigen) which are necessary for cell growth, Heregulin, CD44, CD40, L-selectin, Hypoxia-induced carbonic anhydrase IX, amyloid precursor protein, IL-6

receptor and Intercellular Adhesion Molecule 1 (ICAM) and others [9, 10] and thus controls various processes involved in cell proliferation, differentiation and inflammation.
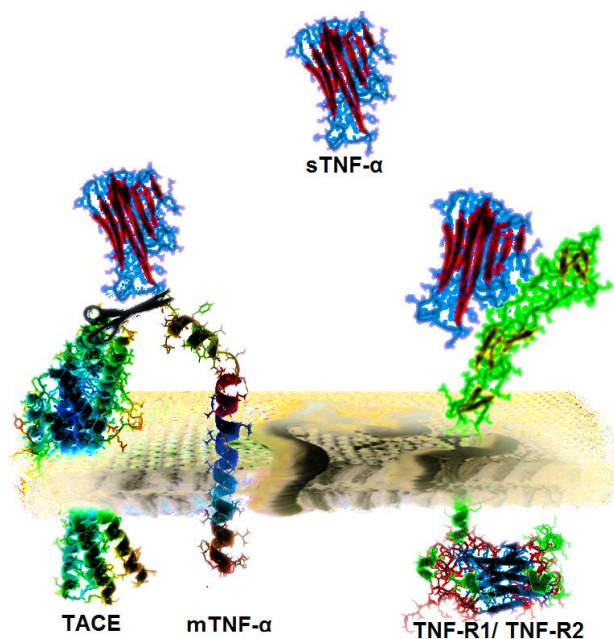
Figure 2.1-1: Mechanism of TNF-α release by TACE

2.1.3   TACE and RA

TACE enhances the release of TNF-α, a major proinflammatory mediator. Both TNF-α and TACE concentrations in synovium of RA patients are increased [11-13]. Severity of RA is correlated with TNF-α level. Animal models provided key evidence that antagonizing TNF-α is a viable therapeutic strategy for RA [14]. Anti TNF-α antibodies reduced the inflammation and joint destruction in RA patients [15].

2.1.4   TACE Enzymology

Table 2.1-1: Biochemical properties of TACE [16, 17].

Protein Name	A Disintegrin and metalloprotease domain-containing protein 17 (ADAM17), Snake venom-like protease, TNF-α convertase, ADAM17 endopeptidase, CD_antigen (CD156b), ADAM18
Enzyme Code	3.4.24.86 Hydrolase. Acting on peptide bonds (peptide hydrolase), metalloendopeptidases, ADAM 17 endopeptidase. Belongs to peptidase family M12

Sequence length	824 amino acids
Catalytic activity	26 kDa proTNF-α + H2O = 17 kDa soluble TNF-α Cleaves Pro-Leu-Ala-Gln-Ala-|-Val-Arg-Ser-Ser-Ser in the membrane-bound, 26-kDa form of TNF-α. Similarly cleaves other membrane-anchored, cell-surface proteins to 'shed' the extracellular domains
Cofactor	Zinc
Tissue specificity	Expressed at highest levels in adult heart, placenta, skeletal muscle, pancreas, spleen, thymus, prostate, testes, ovary and small intestine, and in fetal brain, lung, liver and kidney
Post-translational modification	The precursor is cleaved by a furin endopeptidase. Stimulation by growth factor or phorbol 12-myristate 13-acetate.
Gene Location	Homo sapiens chromosome 2p25.1
Uniplot ID	P78536
Pfam annotation	Pep_M12, Reprolysin_5, Disintegrin
SCOP domain	class: α and β proteins; Fold: Zincin-like; Superfamily: Metalloproteases ("zincins"); Family: TNF-α converting enzyme, TACE

Table 2.1-2: List of crystal structures available from protein database [18]**.**

PDB ID	Method	Resolution (Å)	Remarks
1BKC	X-ray	2.00	Catalytic domain of h-TACE with peptidomemitic inhibitor (INN)
1ZXC	X-ray	2.28	Catalytic domain of h-TACE with Acetylenic sulfonamide hydroxymate (IH6) inhibitor

2A8H	X-ray	2.30	Catalytic domain of h-TACE with thiomorpholine sulfonamide hydroxymate (4NH) inhibitor
2DDF	X-ray	1.70	Catalytic domain of h-TACE with TAPI-2
2FV5	X-ray	2.10	Catalytic domain of h-TACE with IK682 inhibitor
2FV9	X-ray	2.02	Catalytic domain of h-TACE with peptidomemitic inhibitor (INN)
2I47	X-ray	1.90	Catalytic domain of h-TACE with Acetylenic sulfonamide hydroxymate (KGY) inhibitor
2M2F	NMR	-	membran-proximal domain of ADAM17
2OI0	X-ray	2.00	Catalytic domain of h-TACE with Acetylenic sulfonamide thiol (283) inhibitor
3B92	X-ray	2.00	Catalytic domain of h-TACE with Acetylenic sulfonamide thiol (440) inhibitor
3CKI	X-ray	2.30	Catalytic domain of h-TACE with TIMP-3
3E8R	X-ray	1.90	Catalytic domain of h-TACE with 2-phenyl-quinoline hydroxymate (615) inhibitor
3EDZ	X-ray	1.90	Catalytic domain of h-TACE with 2-methyl-quinoline hydroxymate (550) inhibitor
3G42	X-ray	2.10	Catalytic domain of h-TACE with tryptophan sulfonamide (792) inhibitor
3KMC	X-ray	1.80	Catalytic domain of h-TACE with tartrate (403) inhibitor
3KME	X-ray	1.85	Catalytic domain of h-TACE with phenyl- pyrrolidinyl tartrate (Z59) inhibitor
3L0T	X-ray	1.92	Catalytic domain of h-TACE with hydantoin (Z94) inhibitor

3L0V	X-ray	1.75	Catalytic domain of h-TACE with hydantoin (724) inhibitor
3LE9	X-ray	1.85	Catalytic domain of h-TACE with phenyl hydantoin (727) inhibitor
3LEA	X-ray	2.00	Catalytic domain of h-TACE with biphenyl hydantoin (Z93) inhibitor
3LGP	X-ray	1.90	Catalytic domain of h-TACE with benzimidazole thienyl-tartrate based (50X) inhibitor
3O64	X-ray	1.88	Catalytic domain of h-TACE with 2-(2-amin 4-yl)pyrrolidine-based tartrate diamides (786)

2.1.5   Structure of TACE

TACE consists of prodomain, Metalloprotease domain, disintegrin domain, transmembrane domain and cytoplasmic domain [19] (**Figure 2.1-2**). Pro-TACE is synthesized and matures in trans-golgi complex and then migrates to cell membrane. Pro-TACE is inactive and consists of 824 amino acids (134 kDa).

Prodomain acts as chaperone during TACE synthesis and gets cleaved mainly by furin [20]. Metalloprotease domain comprises of active site. Center of the catalytic domain consists of highly twisted β sheets, with larger α helices on both the sides. Catalytic zinc atom forms coordinate complex with imidazole nitrogens of HIS405, HIS409 and HIS415 and forms a conserved zinc binding motif HEXXHXXGXXH followed by "Met turn" a character of Metazincin protein family [21].

At right side of zinc atom in active site is a hydrophobic S1' pocket bridged to S3' pocket as shown in **Figure 2.1-3**. The S1' pocket serves as specificity site for designing TACE inhibitors. Disintegrin domain is involved in interaction with integrins, thereby aiding in cell migration and cell-cell interaction [22]. Cys-rich domain helps in substrate recognition and TACE maturation. Cytoplasmic domain carries tyrosine phosphorylation site and interacts with signaling proteins like Protein Kinase C (PKC) and Extracellular signal-regulated kinase (ERK) [23].

Domain	Residues	Function
Signal peptide	1-17	
Prodomain	18-214	chaperone
Metalloprotease	223-474	catalysis
Disintegrin	475-563	Interaction with integrins
CYS-rich	564-602	Substrate recognition
Transmembrane	672-692	
Cytoplasmic	693-824	Interaction with signaling proteins
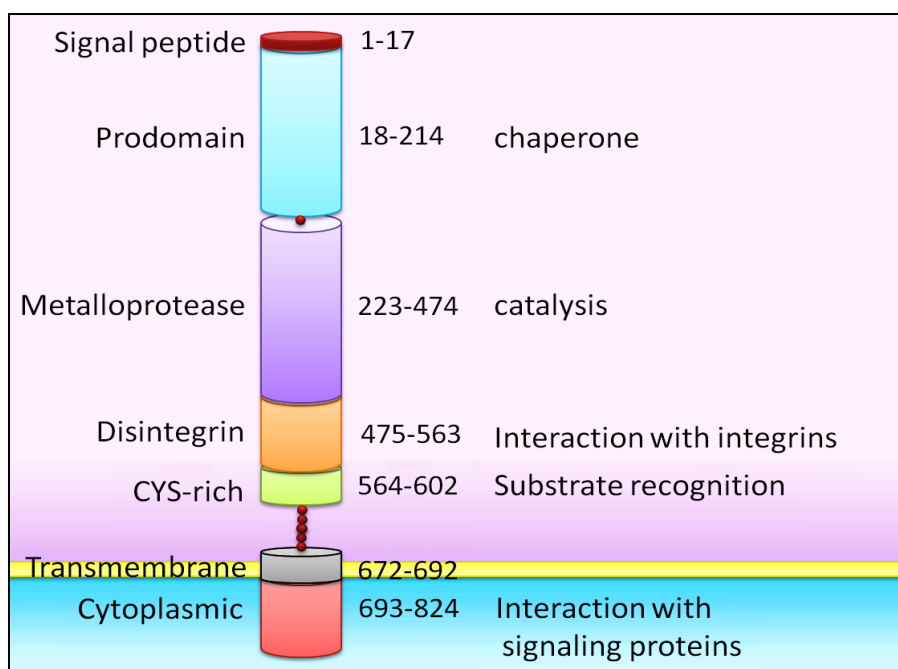
Figure 2.1-2: Domains of TACE
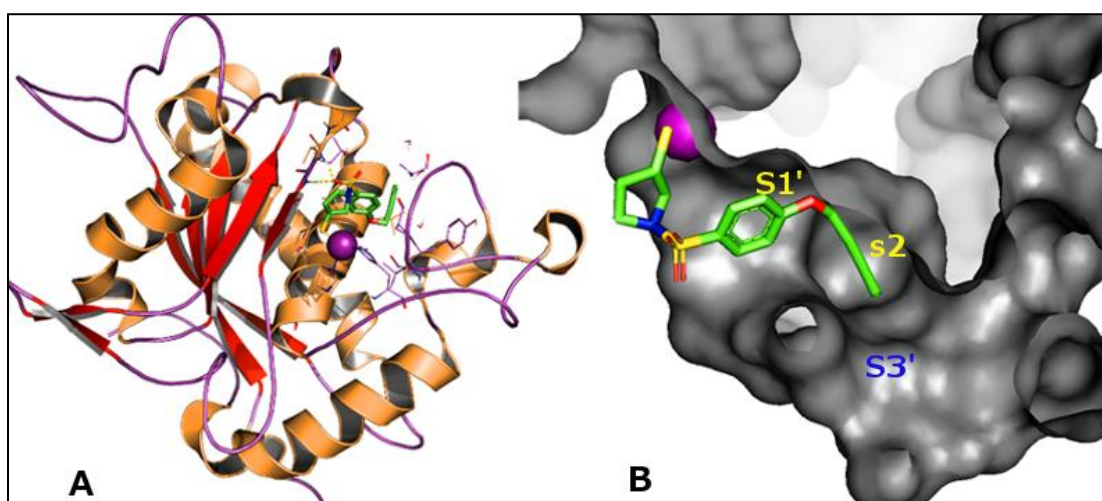
Figure 2.1-3: A: structure of TACE catalytic site domain (PDB ID: 2OI0); **B:** "L" shaped cavity of active site occupied by cocrystalised ligand.

2.1.6   TACE Inhibitor development

In early 1990s, many peptide mimetic broad spectrum Matrix metalloproteinases (MMPs) inhibitors were screened for their anticancer activity. In 1995, based on preclinical and clinical trial results for cA2, a chimeric (human IgG1/K, mouse Fv) anti-TNF antibody, Maini et al. concluded that TNF-α is a critical mediator in inflammation [24]. McGeehan et al., while studying the in vitro and in vivo effects of hydroxamate MMP inhibitors, found that inhibitors reduced the TNF-α secretion from cells without affecting secretion of other cytokines and thus hypothesized that a zinc dependent endopeptidase cleaves TNF-α from its precursor [25]. Later, as TACE involvement in RA was discovered, research was directed

towards screening and enhancing potency of known MMP inhibitors while retaining activity towards other MMPs. During this process, many analogues earlier synthesized were screened for TACE with slight modifications [26]. But with failure of broad spectrum MMPs in clinical trials due to musculoskelectal adverse effects, research has now concentrated towards the design of TACE specific inhibitors. The 1st generation TACE inhibitors were peptide mimetic and they lacked selectivity, the 2nd generation inhibitors showed improved selectivity while in 3rd generation the hydroxamic acid group was replaced with other zinc chelators

Scientists at AstraZeneca developed marimastat (**201**) analogues. Enhanced in vitro potency was observed in **202** but it lacked selectivity and in vivo efficacy [27].

Researchers at DuPont Pharmaceuticals, cyclised P1 and P2 residues of **201** and **203** (BB16) to form a macrocyclic hydroxamic acids. Compound **204** and **205** showed activity in nM concentration in whole blood (WB) assay. However they lacked oral bioavailability [28].

201 (Marimastat)
TACE IC$_{50}$:   3.8nM
WB Assay IC$_{50}$: 4.1µM
lack invivo efficacy

202
TACE IC$_{50}$:   0.57 nM
WB Assay IC$_{50}$: 0.28µM
lack invivo efficacy

203 (BB16)
WB Assay IC$_{50}$: 1.8µM
lack invivo efficacy

204 (SP057)
TACE Ki :   4.2 nM
WB assay: 67 nM
ED$_{50}$:   2.6mg/kg

205 (SL422)
TACE Ki :   12nM
WB assay: 220nM
ED$_{50}$:   15mg/kg
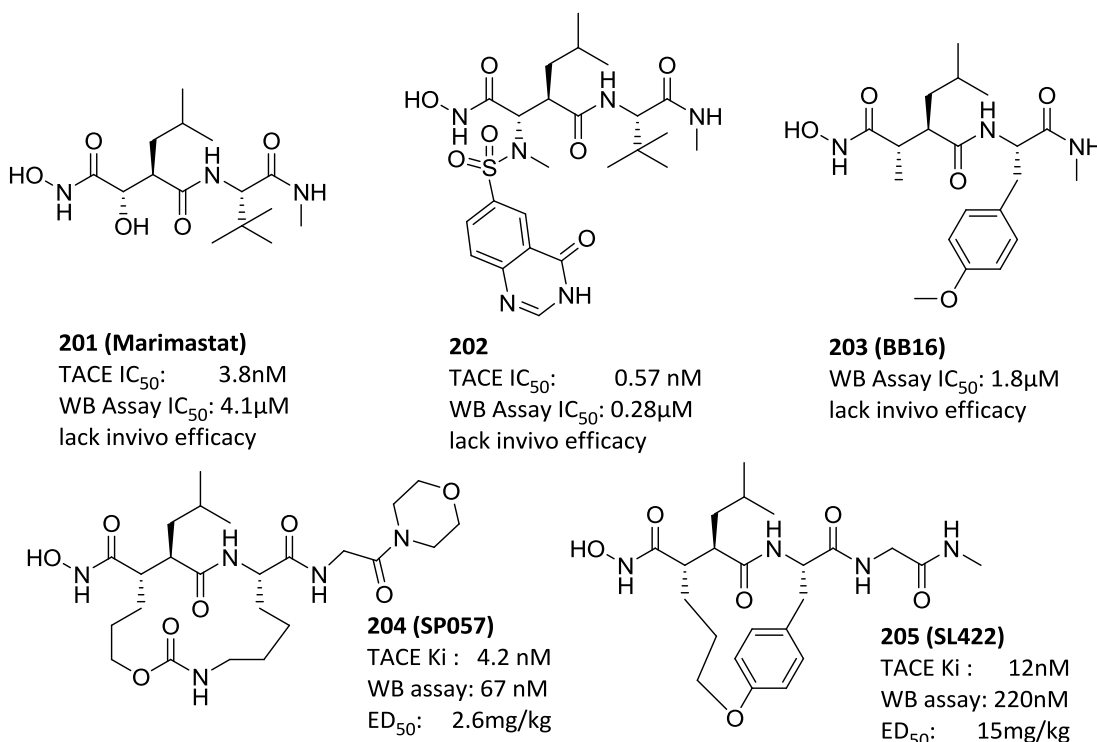
Researchers at Abbott tried to address the selectivity by introducing bulky group in compound **205** and increased macrocyclic ring size by another carbon atom. Compound **206** showed TACE IC$_{50}$ 7.6 nM and 14 fold selectivity over MMP1 and MMP2, as they lack S2 and S3 cavity [29].

Park et al. prepared hydroxamate analogues of natural product gelastatin. Compound **207** showed TACE IC$_{50}$ value of 1.98 µM but lacked selectivity [30]. Chun et al. prepared a

series of coumarins based on gelastatin hydroxymate. Compound **209** had increased selectivity for TACE as compared to MMP2 and MMP9 indicating the role of bulky group (methyl (**208**) to benzyl (**209**)) at Cα carbon atom [31].

Yang et al. tried to explore the effect of different phenyl groups separated by alkyl chain and a tertiary amino group at the other end. Compound **210** with 2-(4-(2-Dimethylamino-ethoxy)phenyl)ethyl substitution showed TACE IC_{50} at 2.8 nM and showed selectivity over MMPs [32].

207
TACE IC_{50}: 1.98μM
RAW264.7 cell Assay IC_{50}: 8.8μM
MMP2 IC_{50}: 6nM
MMP9 IC_{50}: 23nM

208
TACE IC_{50}: 3nM
RAW264.7 cell Assay IC_{50}: 0.32μM

206
TACE IC_{50}: 7.6nM
THP-1 assay IC_{50}: 140nM
MMP1 IC_{50}: 2300nM
MMP2 IC_{50}: 110nM

209
TACE IC_{50}: 5nM
RAW264.7 cell Assay IC_{50}: 0.54μM

210
TACE IC_{50}: 2.8nM
RAW264.7 cell Assay IC_{50}: 0.06μM
MMP1 IC_{50}: 557μM
MMP3 IC_{50}: 8.7μM
MMP9 IC_{50}: 23μM
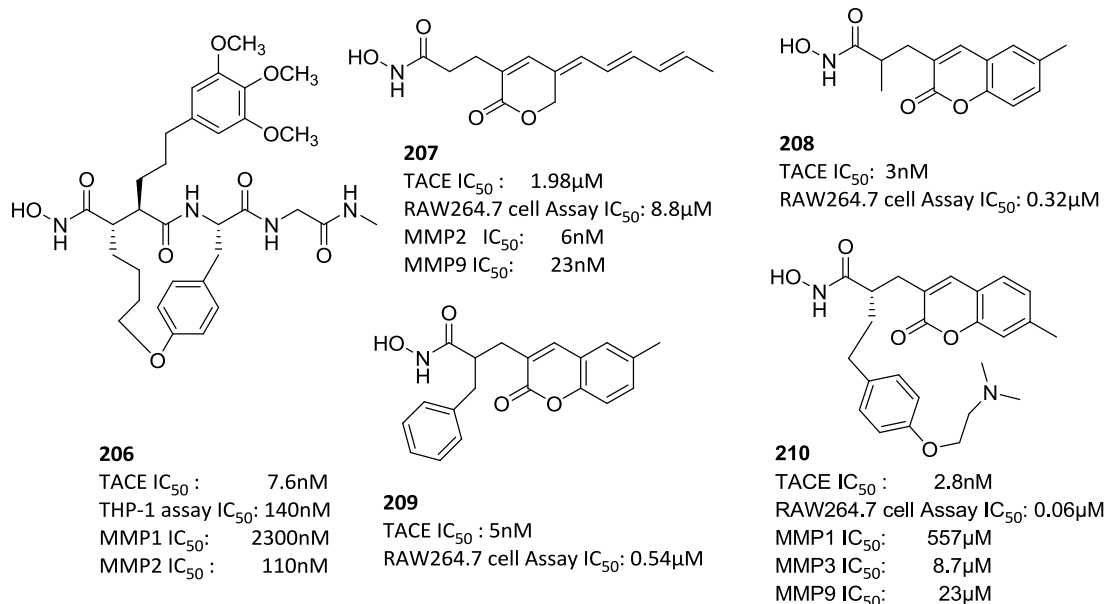
Research at Wyeth: Initially researchers at Wyeth prepared homology model of TACE using PDB structure 2AIG. They found that S1 pocket of TACE resembles like other MMPs, but in TACE it is extended till S3 cavity and the channel resembles L shape. This particular structural observation was used to design selective TACE inhibitors and later it was confirmed with crystal structure 1BKC. Initially, Anthranilate-Sulfonamide-Hydroxamate derivatives were prepared. Various alkoxy and aryloxy substitutions were done at *para* position of phenyl ring. When methoxy was replaced with propargyloxy group (**211**), a 2 fold increase in activity as well as selectivity against other MMPs was observed (MMP1 - 101 fold, MMP9- 19 fold, MMP-13, 10 fold) [33, 34]. When anthranilic acid of **211** was replaced with glycine (**212**), a moderate cellular activity and 2 fold raise in enzymatic activity was observed while retaining selectivity over other MMPs [35]. Later conformational rigidity was introduced by cyclizing sulfonamide amine and Cα atom of glycine with macrocyclic ring. With benzodiazepine ring and cyclopropyl substitution compound **213** retained enzymatic activity and showed enhanced cellular activity but selectivity was slightly reduced [36]. Compound **214** was orally active and showed good selectivity over MMP-1, 2,7,8, 9,13 and 14 [37].

211
TACE IC$_{50}$: 16 nM
MMP1 IC$_{50}$: 1616 nM
MMP9 IC$_{50}$: 304 nM
MMP13 IC$_{50}$: 154 nM

212
TACE IC$_{50}$: 7 nM
THP-1 assay
%inhib: 33 (at 3µM)
MMP1 IC$_{50}$: 1895 nM
MMP9 IC$_{50}$: 309 nM
MMP13 IC$_{50}$: 99 nM

213
TACE IC$_{50}$: 10 nM
THP-1 assay
%inhib: 59 (at 3µM)
MMP1 IC$_{50}$: 841 nM
MMP9 IC$_{50}$: 33 nM
MMP13 IC$_{50}$: 29 nM

214
TACE IC$_{50}$: 1.4 nM
WB assay IC$_{50}$: 3.3 µM
MMP1 IC$_{50}$: 6600 nM
MMP2 IC$_{50}$: 345 nM
MMP13 IC$_{50}$: 74 nM
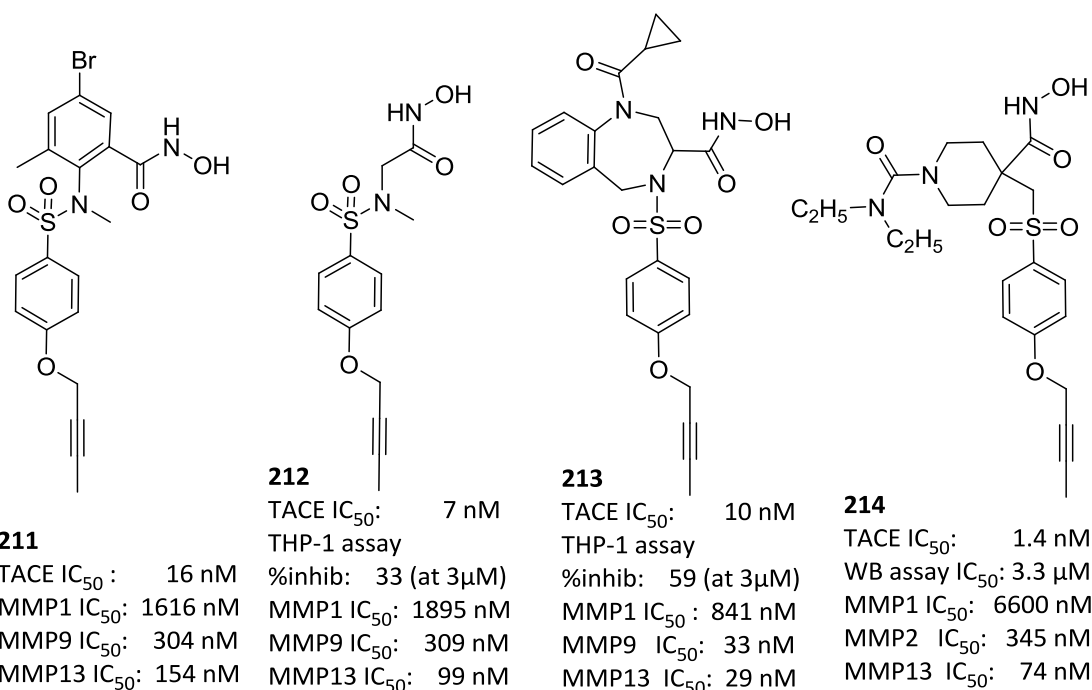
Research at Bristol-Myers Squib: Scientists at BMS believed that non-drugability of 2nd generation TACE inhibitors are due to chelation to metals other than zinc thereby causing complex Pharmacokinetic/Pharmacodynamic (PK/PD) profile and toxicity [38]. Poor absorption, high renal clearance, O-glucuronidation and amide hydrolysis were thought to be associated with hydroxamic acid [39]. Therefore, attempts were made to replace hydroxamic acid with different zinc binders such as hydantoin, pyrimidinetrione, triazolone, imidazolone and tartrates.

215
TACE IC$_{50}$: 1.03 µM
MMP1 Ki: > 4.95 µM
MMP2 Ki: > 3.33 µM
MMP9 Ki: > 2.13 µM
MMP13 Ki:>5.02 µM

216
TACE IC$_{50}$: 0.081 µM
MMP1 Ki: > 4.95 µM
MMP2 Ki: > 3.33 µM
MMP9 Ki: > 2.13 µM
MMP13 Ki:>5.02 µM

217
TACE IC$_{50}$: 170 nM

218
TACE IC$_{50}$: 11 nM

219
TACE IC$_{50}$: 26 nM
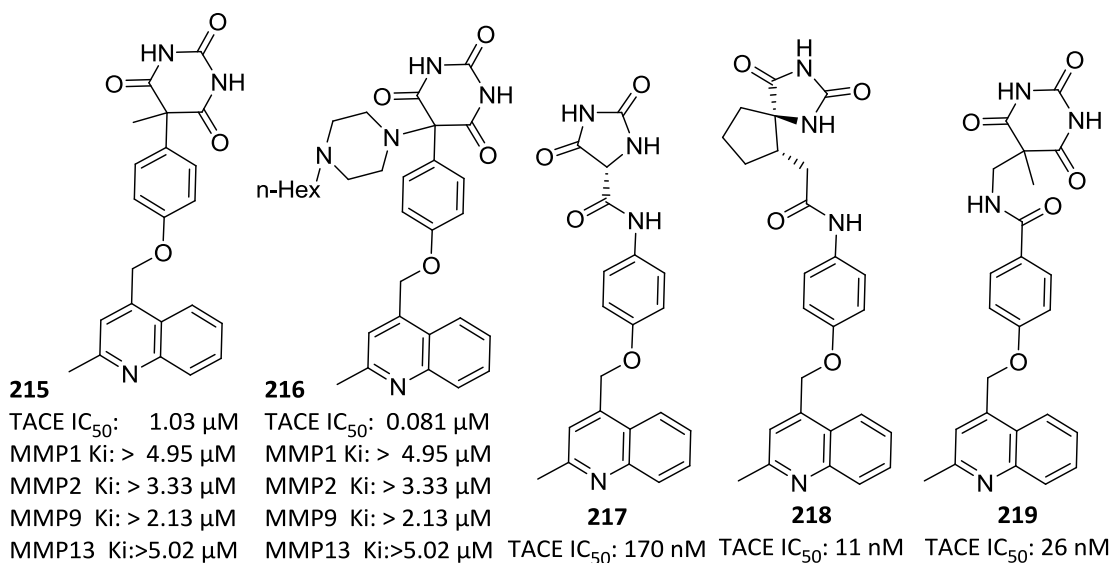
Initially they prepared pyrimidinetrione ring containing compound **215** (TACE IC$_{50}$= 1.03 µM). A twelve fold increase in potency was observed when methyl group present in **215** at 5th position of pyrimidinetrione was replaced with 4-*N*-hexyl-piperazin-1-yl group (**216**) [38].

Significant increase in potency was observed when compound **217**, containing hydrantoin, was conformationally restricted with spirocyclopentane to yield compound **218** [40]. Later researchers tried to introduce linker connecting pyrimidinetrione and phenyl ring of compound **215**. A 4 fold increase in potency was observed when linked with –CH$_2$NHC(O)- as in compound **219**. While with –CH$_2$C(O)NH- linker potency decreased by 12 folds [41]. With optimized linker, effect of substitution at 5th position of pyrimidinetrione ring was then studied. Activity was increased with 4-*N*-substituted piperazin-1-yl group. 4-*N*-mesyl-piperazin-1-yl substitution increased the potency to 2 nM against TACE but IC$_{50}$ value by WB assay was more than 50 µM, whereas 4-*N*-pivaloyl-piperazin-1-yl substituted compound **220** showed good activity in WB assay but lost selectivity against MMP2 and 12. The WB assay potency was significantly boosted when pyrimidinetrione ring was conformationally restricted with spiro ring. N-(Boc)-pyrrolidine-derived spiro analogue **221** showed significant potency gain in WB assay while selectivity against MMP 12 was only 8 fold [41]. When linker of **218** was then replaced with optimized –CH$_2$NHC(O)- to design compound **222**, a 2 fold increase in potency was observed. While replacing pyrimidinetrione ring of **221** with hydantoin (**223**) increased the potency by 10 folds [42].

220
TACE IC$_{50}$:   29 nM
WB assay :   38 µM
MMP2 Ki: 0.02 µM
MMP3 Ki: >4.50 µM
MMP12 Ki: 0.05 µM

221
TACE IC$_{50}$:   128 nM
WB assay :   14 µM
MMP2 Ki: > 3.33 µM
MMP3 Ki: > 4.50 µM
MMP12 Ki: 1.03 µM
MMP13 Ki: >5.21 µM

222
TACE IC$_{50}$:   11 nM
MMP1 Ki: >4946 nM
MMP2 Ki: >3333 nM
MMP9 Ki: >2128 nM
MMP13 Ki:>5.02 µM

223
TACE IC$_{50}$:   13 nM
MMP1 Ki: >4946 nM
MMP2 Ki: >3333 nM
MMP9 Ki: >2128 nM
MMP13 Ki:>5.02 µM
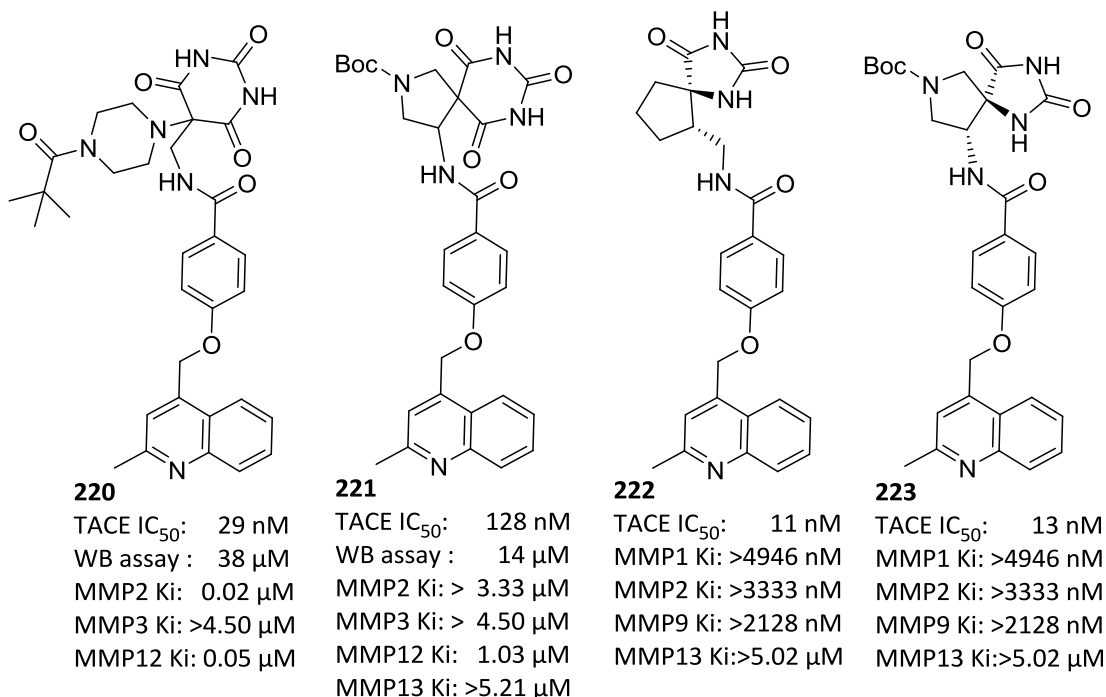
Research at Merck: Researchers at Merck used tartrate moiety as zinc binder. Three alcoholic hydroxyl groups of tartarate were found to interact in tridentate fashion since zinc is present at the entrance of the TACE cavity. Researchers at Schering-Plough Research Institute initially prepared tartaric acid di-amide derivatives (Compound **224** Ki = 11 nM).

When ethyl linker was replaced with methyl between tartrate amide and thiophenyl ring, a significant loss in activity was observed [43]. When Merck scientists introduced bulkier substitution on 2-methyl thiophenyl ring activity was significantly improved. Benzimidazole substituted compound **225** exhibited potent activity (Ki = 2 nM) but poor rat PK profile. Oral bioavailability and activity was then improved in compound **226** (TACE Ki : 0.86 nM, WB assay IC_{50} 0.886 µM) where thiophenyl-benzimidazole ring of **225** was replaced with phenyl ring substituted with pyrazole moiety [44]. Replacement of *m*-chloro phenyl moiety with 2-(*N*-mono alkylated)-aminothiazole moiety improved the oral bioavailability. Methyl substitution on amine of thiazole leads to reduced enzyme activity compared to **226** but cellular activity was increased by 2 folds (Ki = 7 nM, WB assay IC_{50} 5.8 µM). Further replacing methyl with cyclopropyl group (**227**) displayed good PK profile and activity (Ki = 5 nM, WB assay IC_{50} = 3.02 µM). Significant reduction in oral bioavailability and increase in potency was observed when pyrrolidine ring was fused with phenyl (compound **228**, Ki 1 nM, WB assay IC_{50} = 0.416 µM) [39].

224
TACE Ki: 11 nM
MMP1 Ki: >50 µM
MMP2 Ki: 33 µM
MMP3 Ki: 17 µM
MMP13 Ki: 20 µM

225
TACE Ki: 2 nM

226
TACE Ki: xxx 0.86 nM
WB assay IC_{50}: 0.89µM
MMP1 Ki: > 100 µM
MMP2 Ki: 43.6 µM
MMP9 Ki: 88.4 µM
MMP13 Ki: 3.3 µM

227
TACE Ki: 5 nM
WB assay IC_{50}: 3.02 µM
MMP1 Ki: > 40 µM
MMP2 Ki: 20.5 µM
MMP3 Ki: > 40 µM
MMP9 Ki: > 40 µM
MMP13 Ki: > 40 µM

228
TACE Ki: 1 nM
WB assay IC_{50}: 0.42 µM
MMP1 Ki: > 40 µM
MMP2 Ki: > 40 µM
MMP3 Ki: > 40 µM
MMP9 Ki: 10.1 µM
MMP13 Ki: 14.1 µM

2.1.7 References

[1] Aggarwal, B.B.; Gupta, S.C.; Kim, J.H. Historical perspectives on tumor necrosis factor and its superfamily: 25 years later, a golden journey. *Blood.* **2012,** 119(3), 651-665.

[2] Piguet, P.; Grau, G.; Vesin, C.; Loetscher, H.; Gentz, R.; Lesslauer, W. Evolution of collagen arthritis in mice is arrested by treatment with anti-tumour necrosis factor (TNF) antibody or a recombinant soluble TNF receptor. *Immunology.* **1992,** 77(4), 510-514.

[3] Moreland, L.W.; Baumgartner, S.W.; Schiff, M.H.; Tindall, E.A.; Fleischmann, R.M.; Weaver, A.L.; Ettlinger, R.E.; Cohen, S.; Koopman, W.J.; Mohler, K. Treatment of rheumatoid arthritis with a recombinant human tumor necrosis factor receptor (p75)–Fc fusion protein. *N Engl J Med.* **1997,** 337(3), 141-147.

[4] Targan, S.R.; Hanauer, S.B.; van Deventer, S.J.; Mayer, L.; Present, D.H.; Braakman, T.; DeWoody, K.L.; Schaible, T.F.; Rutgeerts, P.J. A short-term study of chimeric monoclonal antibody cA2 to tumor necrosis factor α for Crohn's disease. *N Engl J Med.* **1997,** 337(15), 1029-1036.

[5] Haider, S.; Knöfler, M. Human tumour necrosis factor: physiological and pathological roles in placenta and endometrium. *Placenta.* **2009,** 30(2), 111-123.

[6] Naudé, P.J.; Den Boer, J.A.; Luiten, P.G.; Eisel, U.L. Tumor necrosis factor receptor cross-talk. *FEBS J.* **2011,** 278(6), 888-898.

[7] Faustman, D.; Davis, M. TNF receptor 2 pathway: drug target for autoimmune diseases. *Nat Rev Drug Discov.* **2010,** 9(6), 482-493

[8] Newton, R.; Solomon, K.; Covington, M.; Decicco, C.; Haley, P.; Friedman, S.; Vaddi, K. Biology of TACE inhibition. *Ann Rheum Dis.* **2001,** 60(suppl 3), iii25-iii32.

[9] Moss, M.L.; Sklair-Tavron, L.; Nudelman, R. Drug insight: tumor necrosis factor-converting enzyme as a pharmaceutical target for rheumatoid arthritis. *Nat Clin Pract Rheumatol.* **2008,** 4(6), 300-309.

[10] Saito, K.; Horiuchi, K.; Kimura, T.; Mizuno, S.; Yoda, M.; Morioka, H.; Akiyama, H.; Threadgill, D.; Okada, Y.; Toyama, Y.; Sato, K. Conditional inactivation of TNFα-converting enzyme in chondrocytes results in an elongated growth plate and shorter long bones. *PLoS One.* **2013,** 8(1), e54853.

[11] Saxne, T.; Palladino, M.; Heinegard, D.; Talal, N.; Wollheim, F. Detection of tumor necrosis factor α but not tumor necrosis factor β in rheumatoid arthritis synovial fluid and serum. *Arthritis Rheum.* **1988,** 31(8), 1041-1045.

[12] Chu, C.; Field, M.; Feldmann, M.; Maini, R. Localization of tumor necrosis factor α in synovial tissues and at the cartilage–pannus junction in patients with rheumatoid arthritis. *Arthritis Rheum.* **1991,** 34(9), 1125-1132.

[13] Brennan, F.M.; Green, P.; Amjadi, P.; Robertshaw, H.J.; Alvarez-Iglesias, M.; Takata, M. Interleukin-10 regulates TNF-alpha-converting enzyme (TACE/ADAM-17) involving a TIMP-3 dependent and independent mechanism. *Eur J Immunol.* **2008,** 38(4), 1106-1117.

[14] Williams, R.O.; Feldmann, M.; Maini, R.N. Anti-tumor necrosis factor ameliorates joint disease in murine collagen-induced arthritis. *Proc Natl Acad Sci.* **1992,** 89(20), 9784-9788.

[15] Feldmann, M. Development of anti-TNF therapy for rheumatoid arthritis. *Nat Rev Immunol.* **2002,** 2(5), 364-371.

[16] Enzyme Structures Database. http://www.ebi.ac.uk/thornton-srv/databases/cgi-bin/enzymes/GetPage.pl?ec_number=3.4.24.86. (accessed on Decenber 24th, 2013).

[17] NCBI. ADAM metallopeptidase domain 17 [Homo sapiens (human)]. http://www.ncbi.nlm.nih.gov/gene/6868. (accessed on December 24th 2013).

[18] De, B.; Natchus, M.G.; Cheng, M.; Pikul, S.; Almstead, N.G.; Taiwo, Y.O.; Snider, C.E.; Chen, L.; Barnett, B.; Gu, F. The next generation of MMP inhibitors: design and synthesis. *Ann N Y Acad Sci.* **1999,** 878(1), 40-60.

[19] Gooz, P.; Dang, Y.; Higashiyama, S.; Twal, W.O.; Haycraft, C.J.; Gooz, M. A disintegrin and metalloenzyme (ADAM) 17 activation is regulated by α5β1 integrin in kidney mesangial cells. *PloS one.* **2012,** 7(3), e33350.

[20] Schlondorff, J.; Becherer, J.; Blobel, C. Intracellular maturation and localization of the tumour necrosis factor α convertase (TACE). *Biochem J.* **2000,** 347, 131-138.

[21] Bode, W.; Gomis-Rüth, F.X.; Stöckler, W. Astacins, serralysins, snake venom and matrix metalloproteinases exhibit identical zinc-binding environments (HEXXHXXGXXH and Met-turn) and topologies and should be grouped into a common family, the 'metzincins'. *FEBS letters.* **1993,** 331(1), 134-140.

[22] Bax, D.V.; Messent, A.J.; Tart, J.; van Hoang, M.; Kott, J.; Maciewicz, R.A.; Humphries, M.J. Integrin α5β1 and ADAM-17 interact in vitro and co-localize in migrating HeLa cells. *J Biol Chem.* **2004,** 279(21), 22377-22386.

[23] Arribas, J.; Esselens, C. ADAM17 as a therapeutic target in multiple diseases. *Curr Pharm Des.* **2009,** 15(20), 2319-2335.

[24] Feldmann, M.; Brennan, F.M.; Elliott, M.J.; Williams, R.O.; Maini, R.N. TNF alpha is an effective therapeutic target for rheumatoid arthritis. *Ann N Y Acad Sci.* **1995,** 766(1), 272-278.

[25] McGeehan, G.M.; Becherer, J.D.; Bast, R.C.; Boyer, C.M.; Champion, B.; Connolly, K.M.; Conway, J.G.; Furdon, P.; Karp, S.; Kidao, S. Regulation of tumour necrosis factor-α processing by a metalloproteinase inhibitor. *Nature.* **1994,** 370(6490), 558-561.

[26] De, B.; Natchus, M.G.; Cheng, M.; Pikul, S.; Almstead, N.G.; Taiwo, Y.O.; Snider, C.E.; Chen, L.; Barnett, B.; Gu, F. The next generation of MMP inhibitors: design and synthesis. *Ann N Y Acad Sci.* **1999,** 878(1), 40-60.

[27] Barlaam, B.; Bird, T.G.; Lambert-van der Brempt, C.; Campbell, D.; Foster, S.J.; Maciewicz, R. New α-substituted succinate-based hydroxamic acids as TNFα convertase inhibitors. *J Med Chem.* **1999,** 42(23), 4890-4908.

[28] Xue, C.B.; Voss, M.E.; Nelson, D.J.; Duan, J.J.W.; Cherney, R.J.; Jacobson, I.C.; He, X.; Roderick, J.; Chen, L.; Corbett, R.L. Design, synthesis, and structure-activity relationships of macrocyclic hydroxamic acids that inhibit tumor necrosis factor α release in vitro and in vivo. *J Med Chem.* **2001,** 44(16), 2636-2660.

[29] Holms, J.; Mast, K.; Marcotte, P.; Elmore, I.; Li, J.; Pease, L.; Glaser, K.; Morgan, D.; Michaelides, M.; Davidsen, S. Discovery of selective hydroxamic acid inhibitors of tumor necrosis factor-α converting enzyme. *Bioorg Med Chem Lett.* **2001,** 11(22), 2907-2910.

[30] Park, S.K.; Han, S.B.; Lee, K.; Lee, H.J.; Kho, Y.H.; Chun, H.; Choi, Y.; Yang, J.Y.; Yoon, Y.D.; Lee, C.W. Gelastatins and their hydroxamates as dual functional inhibitors for TNF-α converting enzyme and matrix metalloproteinases: Synthesis, biological evaluation, and mechanism studies. *Biochem Biophys Res Commun.* **2006,** 341(2), 627-634.

[31] Chun, K.; Park, S.K.; Kim, H.M.; Choi, Y.; Kim, M.H.; Park, C.H.; Joe, B.Y.; Chun, T.G.; Choi, H.M.; Lee, H.Y. Chromen-based TNF-α converting enzyme (TACE) inhibitors: design, synthesis, and biological evaluation. *Bioorg Med Chem.* **2008,** 16(1), 530-535.

[32] Yang, J.S.; Chun, K.; Park, J.E.; Cho, M.; Seo, J.; Song, D.; Yoon, H.; Park, C.H.; Joe, B.Y.; Choi, J.H. Structure based optimization of chromen-based TNF-α converting enzyme (TACE) inhibitors on S1' pocket and their quantitative structure–activity relationship (QSAR) study. *Bioorg Med Chem.* **2010,** 18(24), 8618-8629.

[33] Chen, J.M.; Jin, G.; Sung, A.; Levin, J.I. Anthranilate sulfonamide hydroxamate TACE inhibitors. Part 1: Structure-based design of novel acetylenic P1' groups. *Bioorg Med Chem Lett.* **2002,** 12(8), 1195-1198.

[34] Levin, J.; Chen, J.; Du, M.; Nelson, F.; Wehr, T.; DiJoseph, J.; Killar, L.; Skala, S.; Sung, A.; Sharr, M. The discovery of anthranilic acid-based MMP inhibitors. Part 3: incorporation of basic amines. *Bioorg Med Chem Lett.* **2001,** 11(22), 2975-2978.

[35] Levin, J.; Chen, J.; Cheung, K.; Cole, D.; Crago, C.; Santos, E.; Du, X.; Khafizova, G.; MacEwan, G.; Niu, C. Acetylenic TACE inhibitors. Part 1. SAR of the acyclic sulfonamide hydroxamates. *Bioorg Med Chem Lett.* **2003,** 13(16), 2799-2803.

[36] Nelson, F.C.; Delos Santos, E.; Levin, J.I.; Chen, J.M.; Skotnicki, J.S.; DiJoseph, J.F.; Sharr, M.A.; Sung, A.; Killar, L.M.; Cowling, R. Benzodiazepine inhibitors of the MMPs and TACE. *Bioorg Med Chem Lett.* **2002,** 12(20), 2867-2870.

[37] Park, K.; Aplasca, A.; Du, M.T.; Sun, L.; Zhu, Y.; Zhang, Y.; Levin, J.I. Design and synthesis of butynyloxyphenyl β-sulfone piperidine hydroxamates as TACE inhibitors. *Bioorg Med Chem Lett.* **2006,** 16(15), 3927-3931.

[38] Duan, J.J.W.; Lu, Z.; Wasserman, Z.R.; Liu, R.Q.; Covington, M.B.; Decicco, C.P. Non-hydroxamate 5-phenylpyrimidine-2, 4, 6-trione derivatives as selective inhibitors of TNF-α converting enzyme. *Bioorg Med Chem Lett.* **2005,** 15(12), 2970-2973.

[39] Dai, C.; Li, D.; Popovici-Muller, J.; Zhao, L.; Girijavallabhan, V.M.; Rosner, K.E.; Lavey, B.J.; Rizvi, R.; Shankar, B.B.; Wong, M.K. 2-(2-Aminothiazol-4-yl) pyrrolidine-based tartrate diamides as potent, selective and orally bioavailable TACE inhibitors. *Bioorg Med Chem Lett.* **2011,** 21(10), 3172-3176.

[40] Sheppeck II, J.E.; Tebben, A.; Gilmore, J.L.; Yang, A.; Wasserman, Z.R.; Decicco, C.P.; Duan, J.J.W. A molecular modeling analysis of novel non-hydroxamate inhibitors of TACE. *Bioorg Med Chem Lett.* **2007,** 17(5), 1408-1412.

[41] Duan, J.J.W.; Chen, L.; Lu, Z.; Jiang, B.; Asakawa, N.; Sheppeck II, J.E.; Liu, R.Q.; Covington, M.B.; Pitts, W.; Kim, S.H. Discovery of low nanomolar non-hydroxamate inhibitors of tumor necrosis factor-α converting enzyme (TACE). *Bioorg Med Chem Lett.* **2007,** 17(1), 266-271.

[42] Sheppeck II, J.E.; Gilmore, J.L.; Yang, A.; Chen, X.T.; Xue, C.B.; Roderick, J.; Liu, R.Q.; Covington, M.B.; Decicco, C.P.; Duan, J.J.W. Discovery of novel hydantoins as selective non-hydroxamate inhibitors of tumor necrosis factor-α converting enzyme (TACE). *Bioorg Med Chem Lett.* **2007,** 17(5), 1413-1417.

[43] Rosner, K.E.; Guo, Z.; Orth, P.; Shipps Jr, G.W.; Belanger, D.B.; Chan, T.Y.; Curran, P.J.; Dai, C.; Deng, Y.; Girijavallabhan, V.M. The discovery of novel tartrate-based TNF-α converting enzyme (TACE) inhibitors. *Bioorg Med Chem Lett.* **2010,** 20(3), 1189-1193.

[44] Li, D.; Popovici-Muller, J.; Belanger, D.B.; Caldwell, J.; Dai, C.; David, M.; Girijavallabhan, V.M.; Lavey, B.J.; Lee, J.F.; Liu, Z. Structure and activity relationships of tartrate-based TACE inhibitors. *Bioorg Med Chem Lett.* **2010,** 20(16), 4812-4815.

2.2 Objectives

Clinical success of anti TNF-α biological in RA validates TNF-α level reduction in RA as important strategy to treat RA. However high cost, route of administration, risk of infections caused by suppressed immunity, ineffectiveness in some population and hypersensitivity reactions [1] of these antibodies encouraged design and development of small molecules that inhibit TNF-α release on oral delivery.

Aim 1: To design and perform virtual screening of virtual compounds database created by modifying lead compound 2-(4-(but-2-ynyloxy)-N-methylphenylsulfonamido)-N-hydroxybenzamide (**ASH**) in a manner to

ASH
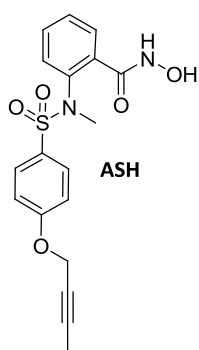
- vary the distance between two phenyl rings
- change the orientation of two phenyl rings by modifying coupler
- replace sulfonamide with other linkers as sulfonamides are associated with allergic reactions [2]

Then substituent selection was done based on following criteria

- should not violate Lipinski rule of oral bioavailability
- should have optimal complementary interactions with receptor
- should be devoid of amino acids and any stereo centers if possible

Aim 2: To synthesize the in silico hits by different synthetic approaches followed by purification and characterization

Aim 3: To screen synthesized compounds for their potential TACE inhibitory activity using in vitro assay.

2.2.1 Reference

[1]   Ikeda, K.; Cox, S.; Emery, P. Biological therapy in early arthritis-overtreatment or the way to go? *Arthritis Res Ther.* **2007,** 9(3), 211.

[2]   Ghimire, S.; Kyung, E.; Lee, J.; Kim, J.; Kang, W.; Kim, E. An evidence-based approach for providing cautionary recommendations to sulfonamide-allergic patients and determining cross-reactivity among sulfonamide-containing medications. *J Clin Pharm Ther.* **2013,** 38(3), 196-202.

2.3 Design of TACE Inhibitors

A database of 70 compounds was designed in silico by modifying the lead 2-(4-(but-2-ynyloxy)-N-methylphenylsulfonamido)-N-hydroxybenzamide (**ASH**), a known TACE inhibitor with IC_{50} of 27 nM [1]. The modification was done as per the objectives specified in **Section 2.2**. Designed Virtual database was then subjected to in silico screening. The active molecules were selected for further study. The modifications in active (in silico) compounds are illustrated in **Table 2.3-1**.

Table 2.3-1: Designed TACE inhibitors
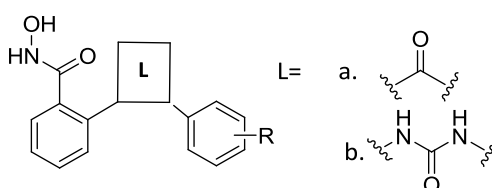
Benzopheone derivatives			Urea derivatives		
code	Linker	R	code	Linker	R
2.5ME_N	a	-OCH$_2$CH$_2$OCH$_3$	3.4CLPN	b	*p*-Cl
2.5PRO_N	a	-O(CH$_2$)$_2$CH$_3$	3.2MXPN	b	*o*-OCH$_3$
2.5PENT_N	a	-O(CH$_2$)$_4$CH$_3$	3.3MXPN	b	*m*-OCH$_3$
2.5ALLY_N	a	- OCH$_2$CHCH$_2$	3.4MXPN	b	*p*-OCH$_3$
2.5HEXA_N	a	-O(CH$_2$)$_5$CH$_3$	3.4NTPN	b	*p*-NO$_2$
2.52B_N	a	-OCH$_2$CCCH$_3$	3.3NTPN	b	*m*-NO$_2$
2.5ETHA_N	a	-OCH$_2$CH$_3$	3.4FLPN	b	*p*-F
2.5MP_N	a	-OCH$_2$ CH$_2$(CH$_3$) CH$_3$	3.2FLPN	b	*o*-F
2.5tBA_N	a	-OC(CH$_3$)$_3$	3.3FLPN	b	*m*-F
2.5CYH_N	a	-O-*c*-Hexyl	3.2TOPN	b	*o*-CH$_3$
2.53B_N	a	-OCH$_2$CH$_2$CCH	3.4TOPN	b	*p*-CH$_3$
2.5MBA_N	a	-OCH$_2$[m-Methoxy-*Ph*]	3.3TOPN	b	*m*-CH$_3$
2.5ANS_N	a	-OCH$_3$			

2.3.1 Methodology

2.3.1.1 Oral bioavailability prediction

Various physicochemical properties of ligands that influence oral bioavailability like molecular weight, hydrogen bond donor (HBD) count and hydrogen bond acceptor count (HBA), topological polar surface area (TPSA), molar refractivity (MR), and log partition coefficient (logP) were determined using Chemaxon Jchem for Excel [2] and aqueous solubility was determined with AlogPS v2.1 algorithm [3]. HBD is defined as number of

oxygen or nitrogen atoms with at least one hydrogen attached. Whereas HBA is defined as number of oxygen and nitrogen atoms present in the molecule with at least one lone pair of electrons. TPSA is defined as the sum of polar atom surfaces in the molecule [2]. MR is a measure of total polarizability of a mole of a substance. logP is the logarithm of partition coefficient between octanol and water. All these parameters affect solubility and partitioning between biological barriers which can have direct correlation with oral bioavailability. Based on observation of approved drugs with these properties various rules were developed to predict oral bioavailability. One such rule is Rule of five or Lipinsky rule [4], which states that for good oral bioavailability given molecule should not violate more than one of the following rules:

- Molecular weight should be less than 500 kDa.

- LogP should be less than 5

- Should have less than 10 hydrogen-bond acceptors (sum of oxygen and nitrogen atoms)

- Should have less than 5 hydrogen bond donors (sum of hydroxyl and amine groups)

This rule is applicable for the drugs absorbed by passive diffusion through cell membrane. Another rule called GSK oral bioavailability rule considers TPSA and number of rotatable bonds of the molecule [5]. It states that for good oral bioavailability molecule should have:

- Less than 10 rotatable bonds

- TPSA should be equal to or less than 140 Å^2

Oral bioavailability is also affected by dissolution of drug which in turn depends on solubility. Aqueous solubility at 25 °C was determined by AlogPS 2.1 algorithm.

2.3.1.2   Toxicity prediction

All compounds should be tested for biological safety in order to minimize the risk of elimination at the later phase of clinical development. Some of the common toxicity screening tests involve study of mutagenicity, tumorigenicity, skin irritation, reproductive toxicity, cardiotoxicity, etc. In this study, the toxicity risk assessment was carried using OSIRIS property explorer [6], a software of Actelion Pharmaceuticals Ltd., Switzerland, hosted on http://www.organic-chemistry.org website. The Drug score is calculated directly from cLogP, logS, molecular weight while drug likeness takes into account the drug score and risk factors [7]. Risk alerts are calculated based on presence of fragments with known risk factors. However, these risk alerts are not fully reliable. And absence of risk alerts does not indicate that compounds are free of toxicity.

2.3.1.3   In silico binding energy and pose prediction

Protein and ligand preparation: pdb structure 2OI0 [8], an human TACE crystal structure in complex with 3s-1-{[4-(but-2-yn-1-yloxy)phenyl]sulfonyl}pyrrolidine-3-thiol retrieved from Brookhaven protein database was used for docking study. 2OI0 chain A was superimposed on chain A of 1BKC using SuperPose v1.0 [9]. 2OI0 pdb file was prepared using protein preparation wizard of Maestro [10]. Chain A was selected and all water molecules were removed. Ionization states of amino acids and metal zinc were determined using Epik at pH 7.0 ± 2.0 [11]. Hydrogens were added and minimized using OPLS 2005 force field. All the compounds were drawn using Chemaxon Marvin Sketch [12] and prepared with ligand preparation wizard of Maestro. All possible Ionization states at pH 7.0±2.0 were enumerated using Epik and minimized. For Autodock, ligands were further processed by adding Gasteiger charges using MGL tools v1.5.4 [13].

Docking with Glide: Grid of 2OI0 active site was created using Maestro. A centroid at 8.0, 6.0, 25.0 on x, y, z axis respectively was used to define the grid box. Grid box length was set to 6, 14 and 10 Å along x, y and z directions respectively. Docking of Maestro processed ligand database was performed using Glide SP (standard precession) V5.0, and Glide XP (Extra Precision) [12].

Docking with FRED: PDB structure file prepared by Maestro was selected and active site grid was generated using FRED receptor setup program [8]. A grid box of volume 7360 $Å^3$ was created around cocrystalised ligand. Tautomeric states of His405, His409 and His415 were adjusted in order to form optimal coordination complex with zinc. Inner and outer contours were set at 88 Å and 2053 Å respectively. After getting satisfactory results in trial dock, grid file was saved for future use. FRED does not produce ligand conformers on fly therefore conformers were generated first using Omega [14]. Maximum of 1000 conformers falling in the energy window of 25 Kcal/mol from its minima were generated for each ligand with 0.5 RMSD between each conformers. Search force field was set to Merck Molecular Force Field 94 (MMFF94). Multiconformer database was then docked with receptor grid generated earlier. FRED was allowed to assign protein charges. Initial optimization was done with chemgauss3 empirical scoring function then local minimization using MMFF 94 force field was carried out and final poses were scored with chemgauss3, chemgauss2, shapegauss, Piecewise Linear Potential (PLP), Chemical Gaussian Overlay (CGO), Chemical Gaussian Tanimoto (CGT) scores and ranked. While reporting only chemgauss3 values are given.

Docking with Autodock: Autodock V4.2 employing genetic algorithm as search program was used for the study [13]. 2OI0 Pdb structure file prepared by Maestro was selected, kollman charges were assigned to the protein. Energy grid of the active site was created using Autogrid 4.2. A box with npts (number of points in xyz) of 42-44-36 Å with center at 8.0, 6.0 and 25.0 (x,y,z) coordinate. The npts spacing was kept at 0.375 Å. Energy map for electrostatic potential, desolvation potential, atom type maps like carbon, oxygen acceptor, nitrogen, hydrogen bond donor, sulfur, fluorine, chlorine and was created. Ligands were then docked to the active site with genetic algorithm and Lamarckian search (GA-LS) parameter producing 27000 generations in each run. 15 such runs were performed for each ligand. Docked poses were clustered and analyzed for interaction with the protein. Lowest energy binding pose with favorable interactions is reported.

2.3.2   Results and Discussion

2.3.2.1   Predicted oral bioavailability of the designed molecules:

Designed molecules were subjected to in silico oral bioavailability prediction. The descriptors used for prediction are given in **Table 2.3-2** and **Table 2.3-3** for benzophenone and Urea series respectively. All the compounds from benzophenone and urea series fulfilled the Lipinski rule of five criteria. JClogP of benzophenone derivatives varied from 2.08 to 4.34 while Urea derivatives showed between 2.11 and 3.22. JClogP of designed compounds were significantly lower than literature compounds (> 5 JClogP). Since the log P values of reported inhibitors are above 5, it was one of the aim to design compounds with lower log P value. All the compounds of benzophenone series had 2 hydrogen bond donors where as Ureas series had 4 hydrogen bond donor atoms. While hydrogen bond acceptor in benzophenone series varied between 5-6, for urea series it varied between 6-9. Number of rotatable bonds in benzophenone series varied between 4-9, while urea series was relatively rigid with 3-4 numbers of rotatable bonds. Topological polar surface area in benzophenone series was comparatively less than urea series. Number of rotatable bonds and topological polar surface area of all the compounds in benzophenone and urea series were well within the criteria of GSK oral bioavailability prediction rule (number of rotatable bonds < 10, topological polar surface area < 140 Å^2) [5].

Table 2.3-2: Physicochemical properties of Benzophenone series

code	MW (Da)	HBD	HBA	TPSA (Å^2)	Rot	MR (cm^3/mol)	JCLogP	Lipinsky violation
2.5ME_N	315.32	2	6	84.86	7	84.26	2.08	0
2.5PRO_N	299.32	2	5	75.63	6	82.51	3.00	0
2.5PENT_N	327.37	2	5	75.63	8	91.78	3.89	0
2.5ALLY_N	297.30	2	5	75.63	6	82.24	2.85	0
2.5HEXA_N	341.40	2	5	75.63	9	96.41	4.34	0
2.52B_N	309.32	2	5	75.63	6	85.15	3.19	0
2.5ETHA_N	285.29	2	5	75.63	5	77.88	2.48	0
2.5MP_N	313.35	2	5	75.63	6	87.10	3.42	0
2.5tBA_N	313.35	2	5	75.63	5	87.11	3.18	0
2.5CYH_N	339.38	2	5	75.63	5	94.28	3.92	0
2.53B_N	309.32	2	5	75.63	7	84.97	2.56	0
2.5MBA_N	377.39	2	6	84.86	7	104.41	3.69	0
2.5ANS_N	271.27	2	5	75.63	4	73.24	2.12	0

MW: molecular weight; Da: Dalton; HBD: hydrogen bond donor count; HBA: hydrogen bond acceptor count; TPSA: topological polar surface area; Rot: number of rotatable bonds; MR: molar refractivity; JCLogP: Jchem log partition coefficient.

Table 2.3-3: Physicochemical properties of Urea series

code	MW (Da)	HBD	HBA	TPSA (Å^2)	Rot	MR (cm^3/mol)	JCLogP	Lipinsky violation
3.4CLPN	305.72	4	6	90.46	3	80.49	3.22	0
3.2MXPN	301.30	4	7	99.69	4	82.27	2.46	0
3.3MXPN	301.30	4	7	99.69	4	82.27	2.46	0
3.4MXPN	301.30	4	7	99.69	4	82.27	2.46	0
3.4NTPN	316.27	4	9	139.29	4	82.14	2.56	0
3.3NTPN	316.27	4	9	139.29	4	82.14	2.56	0
3.4FLPN	289.26	4	6	90.46	3	75.59	2.76	0
3.2FLPN	289.26	4	6	90.46	3	75.59	2.11	0
3.3FLPN	289.26	4	6	90.46	3	75.59	2.76	0
3.2TOPN	285.30	4	6	90.46	3	80.42	2.48	0
3.4TOPN	285.30	4	6	90.46	3	80.42	3.13	0
3.3TOPN	285.30	4	6	90.46	3	80.42	3.13	0

MW: molecular weight; Da: Dalton; HBD: hydrogen bond donor count; HBA: hydrogen bond acceptor count; TPSA: topological polar surface area; Rot: number of rotatable bonds; MR: molar refractivity; JCLogP: Jchem log partition coefficient.

2.3.2.2 Toxicity prediction of the designed molecules:

Toxicity risk assessment was carried using OSIRIS property explorer and is given in **Table 2.3-4** and **Table 2.3-5** for benzophenone and urea series respectively. All compounds in both the series showed high risk of mutagenicity and mild risk of tumorigenicity due to hydroxamic acid. It has been proposed that enzymatic acylation of hydroxamic acid leads to formation of isocynates by Lossen rearrangement (**Figure 2.3-1**). Resultant isocynates can undergo carbamoylation with the target molecules of the cell.
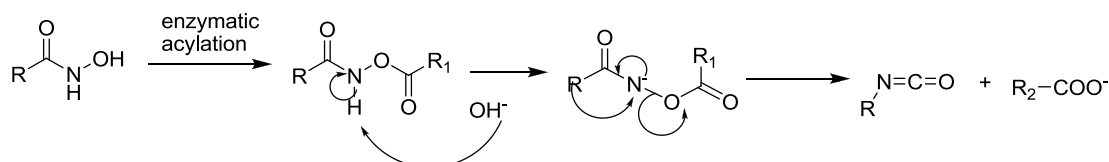
Figure 2.3-1: Mechanism of isocynate formation from Hydroxamic acid (Lossen Rearrangement)

However, a detailed study of mutagenicity property of hydroxamic acid was studied by Paul et.al and it showed that bacterial cells are 5 times more susceptible for mutagenesis than human cells [15]. The activity also varies with substitution and some of the hydroxamic acid derivatives are non mutagenic. Recently, many of the hydroxamic acid derivatives reported as MMP inhibitors were reported to be devoid of any mutagenic evidence on human cells [16].

Some of the designed alkyl and aryl ether containing compounds like 2.5PENT_N, 2.5ALLY_N and 2.5HEX_N showed high risk of skin irritation. While only 3-methoxy phenyl ring containing compounds 2.5MBA_N and 3.3MXPN showed high risk of affecting reproductive organs.

Drug likeliness score takes into account physicochemical properties and groups present in marketed drugs. Whereas Drug score considers toxicity risk and drug likeliness score. Drug likeliness should be a positive value. Compounds with higher values represent the significant PK similarity with ideal drug. Urea derivatives showed positive drug likeliness score except for nitro compounds. While most of benzophenone series showed non drug likeliness except for compound 2.5MP_N and 2.5MBA_N. Reason for non drug likeliness was due to low solubility and presence of benzophenone ring. Drug score varies between 0 to 1 representing bad and good respectively. Drug score of benzophenones was significantly lower compared to urea series mostly due to low solubility. All the compounds scored below 0.5 due to their potential risk of mutagenicity and tumorigenicity. Compounds with potential risk of skin irritation and reproductive effect were scored lowest in the series. Although

OSIRIS risk alerts are not fully reliable but are theoretical toxicity guides. If we neglect hydroxamic acid group related mutagenicity and tumorigenic alerts, many compounds qualify for development of novel TACE inhibitors.

Table 2.3-4: Predicted toxicity and drug score of Benzophenone series

Code	MUT	TUM	Irritant	REP	cLogP	LogS	Drug likeliness	Drug score
2.5ME_N	0.6	0.8	1	1	1.62	-3.44	-4.37	0.21
2.5PRO_N	0.6	0.8	1	1	2.58	-4.09	0.19	0.29
2.5PENT_N	0.6	0.8	0.6	1	3.51	-4.63	-7.17	0.1
2.5ALLY_N	0.6	0.8	0.6	1	2.29	-4.03	-5.01	0.12
2.5HEXA_N	0.6	0.8	0.6	1	3.97	-4.9	-12.1	0.09
2.52B_N	0.6	0.8	1	1	2.59	-6	-2.38	0.15
2.5ETHA_N	0.6	0.8	1	1	2.12	-3.82	-1.28	0.24
2.5MP_N	0.6	0.8	1	1	2.98	-4.46	1.13	0.31
2.5tBA_N	0.6	0.8	1	1	2.67	-4.3	-12.3	0.18
2.5CYH_N	0.6	0.8	1	1	3.27	-5.09	-4.32	0.15
2.53B_N	0.6	0.8	1	1	2.04	-4.41	-1.22	0.22
2.5MBA_N	0.6	0.8	1	0.6	2.94	-4.86	0.4	0.15
2.5ANS_N	0.6	0.8	1	1	1.68	-3.52	-0.34	0.29

MUT = mutagenesicity, TUM = tumorigenic, REP = reproductive effect. Values in column representing MUT, TUM, Irritant and REP represents 1= no indication, 0.8 = medium risk and 0.6 = high risk., dark to light color code in column Druglikeliness and Drug score indicates unfavorable to favorable values.

Table 2.3-5: Predicted toxicity and drug score of Urea series

Code	MUT	TUM	Irritant	REP	cLogP	LogS	Drug likeliness	Drug score
3.4CLPN	0.6	0.8	1	1	2.11	-4.16	3.82	0.37
3.2MXPN	0.6	0.8	1	1	1.39	-3.44	2.67	0.4
3.3MXPN	0.6	0.8	1	0.6	1.39	-3.44	2.66	0.24
3.4MXPN	0.6	0.8	1	1	1.39	-3.44	1.77	0.38
3.4NTPN	0.6	0.8	1	1	1.36	-3.89	-10.32	0.2
3.3NTPN	0.6	0.8	1	1	1.36	-3.89	-2.4	0.21
3.4FLPN	0.6	0.8	1	1	1.55	-3.74	1.72	0.37
3.2FLPN	0.6	0.8	1	1	1.55	-3.74	1.13	0.35
3.3FLPN	0.6	0.8	1	1	1.55	-3.74	0.06	0.24
3.2TOPN	0.6	0.8	1	1	1.81	-3.77	2.65	0.39
3.4TOPN	0.6	0.8	1	1	1.81	-3.77	1.53	0.36
3.3TOPN	0.6	0.8	1	1	1.81	-3.77	2.51	0.39

MUT = mutagenesicity, TUM = tumorigenic, REP = reproductive effect. Values in column representing MUT, TUM, Irritant and REP represents 1= no indication, 0.8 = medium risk and 0.6 = high risk., dark to light color code in column Druglikeliness and Drug score indicates unfavorable to favorable values.

2.3.2.3 Docking scores and binding patterns of designed molecules

Docking study was performed on TACE crystal structure 2OI0 using Autodock 4.2, Glide SP, Glide XP and FRED algorithm. Binding scores were then compared with compound 2-(4-(but-2-ynyloxy)-N-methylphenylsulfonamido)-N-hydroxybenzamide (**ASH**), a known TACE inhibitor with IC_{50} of 27 nM [1].

ASH
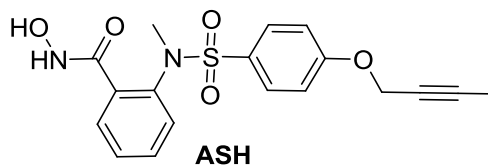
2-(4-(but-2-ynyloxy)-*N*-methylphenylsulfonamido)-*N*-hydroxybenzamide

Benzophenone series: Binding scores of Benzophenone series are given in **Table 2.3-6**. Compound 2.5MBA_N bearing *m*-methoxy benzyloxy group showed highest binding score compared to others while methoxy group containing 2.5ANS_N showed relatively lower binding score in all docking algorithms. On visual analysis It was found that binding pattern of all the derivatives was similar to compound ASH. The deviation in binding scores was found to be mainly contributed by bulky groups attached to phenoxy ring and similar scoring trend was observed across all software's. Binding score variation was mainly contributed by van der waals interaction energy, which increased with increased bulkiness. ASH and 2.5B_N shares same acetylenic group, but molecule 2.52B_N showed higher binding score compared to ASH in all the docking software's. The increased score was mainly due to orientation of keto oxygen leading to enhanced hydrogen bonding interaction with Leu348 and Gly349. The docking score confirmed that designed benzophenones were strong binders than literature reported compound ASH.

More tightly binding compounds 2.5PENT_N, 2.5MBA_N, 2.5CYH_N and 2.5HEXA_N showed zinc chelation, polar and hydrophobic interactions with TACE. **Figure 2.3-2** illustrates the interaction pattern of some predicted potent inhibitors with TACE. For compound 2.52B_N, an interaction of hydroxamic acid with zinc was observed in bidentate fashion (C=O, 3...Zn; 2.16 Å, N-O, 4 ... Zn; 2.33 Å). Carbonyl oxygen of benzophenones formed hydrogen bond with backbone nitrogens of Leu348 (C=O, 1 ...NH, L348; 2.93 Å) and Gly349 residues (C=O, 1 ...NH, G349; 2.92 Å). Another hydrogen bond was formed between hydroxyl hydrogen with Glu406 (O-H, 12...COO, E406; 2.19 Å) and amine hydrogen with backbone carbonyl group of Gly349 (N-H, 11...CO, G349; 1.96 Å). Hydrophobic interactions were seen between second phenyl and acetylenic moiety with Leu348, Val402, Leu401, Val434 and with carbon chain of Glu398. A pi-pi stacking interaction of second phenyl moiety with His405 was also observed.

Table 2.3-6: In silico screening results of Benzophenone series

Code	Autodock	Glide SP		Glide XP		FRED
	Score	Score	Energy	Score	Energy	chemscore3
2.5ME_N	-9.10	-7.68	-56.10	-7.58	-50.92	-121.06
2.5PRO_N	-9.19	-6.55	-53.50	-7.72	-45.67	-117.78
2.5PENT_N	-10.06	-8.64	-53.31	-8.42	-57.49	-126.51
2.5ALLY_N	-9.53	-6.23	-53.87	-7.63	-55.58	-117.55
2.5HEXA_N	-10.42	-7.95	-59.52	-8.86	-58.19	-127.25
2.52B_N	-9.68	-8.36	-54.15	-7.52	-56.12	-118.20
2.5ETHA_N	-9.47	-7.56	-51.34	-7.17	-49.49	-112.02
2.5MP_N	-9.29	-7.73	-51.64	-7.76	-56.40	-118.80
2.5tBA_N	-9.52	-8.63	-55.38	-8.33	-51.59	-115.77
2.5CYH_N	-11.29	-9.05	-62.24	-8.68	-61.04	-126.03
2.53B_N	-9.87	-7.83	-55.94	-8.19	-53.77	-119.88
2.5MBA_N	-11.15	-9.87	-66.00	-8.94	-58.45	-134.15
2.5ANS_N	-7.67	-7.61	-49.53	-6.74	-46.98	-106.83
ASH	-9.31	-7.54	-59.05	-7.06	-52.93	-111.68

All values are expressed in Kcal/mol
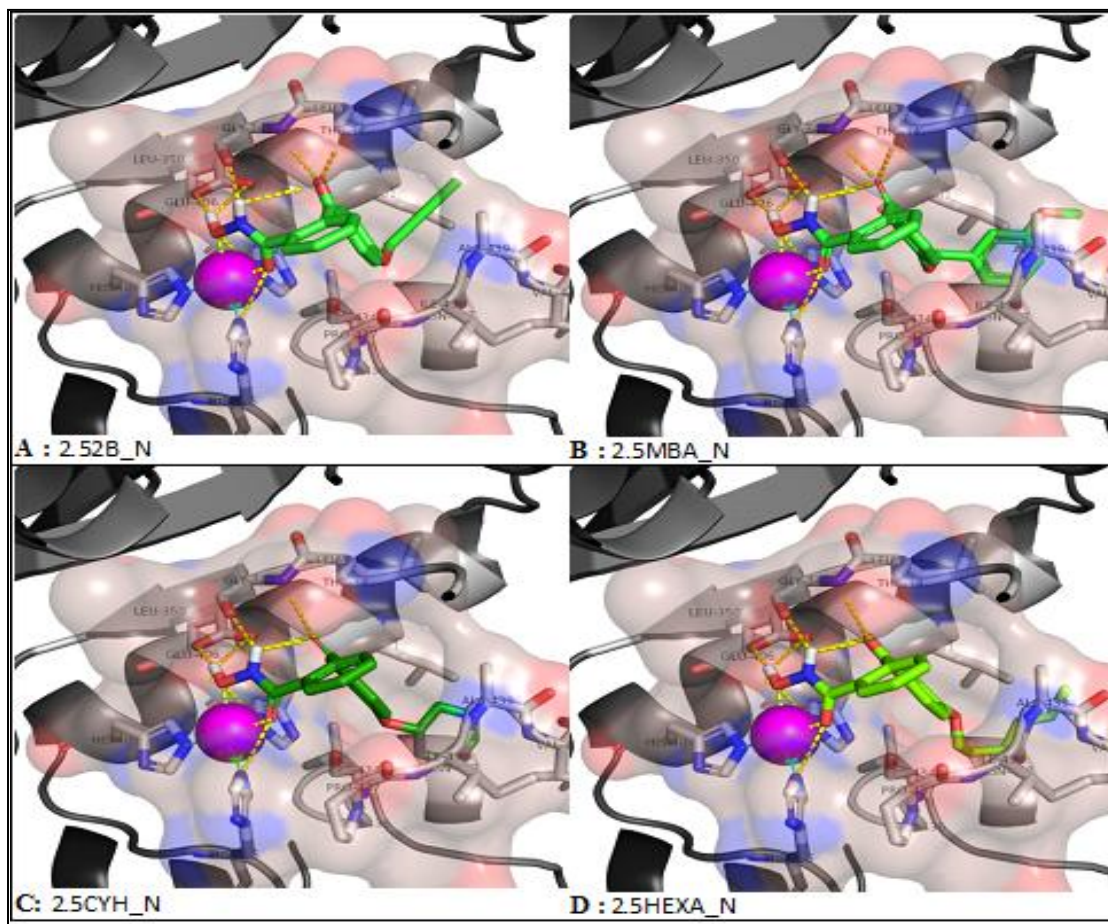
Figure 2.3-2: Binding pose of potent Benzophenone derivatives against TACE, plotted with Pymol [17]. Inhibitors are represented in green. Important binding residues of TACE are represented in gray. Zinc atom is presented as pink sphere and rest of the protein as cartoon model. Yellow dotted line represents hydrogen bond.

As given in **Figure 2.3-2**, in 2.5MBA_N oxygen atoms of hydroxamic acid interact with zinc in bidentate fashion (C=O, 4...Zn; 2.19 Å, N-O, 5 ... Zn; 2.36 Å). Carbonyl oxygen of benzophenone formed hydrogen bond with backbone nitrogens of Leu348 (C=O, 3 ...NH, L348; 2.16 Å) and Gly349 residues (C=O, 3 ...NH, G349; 2.07 Å). Another hydrogen bond was formed between hydroxyl hydrogen with GLU406 (O-H, 19...COO, E406; 2.33 Å) and amine hydrogen with backbone carbonyl group of Gly349 (N-H, 18...CO, G349; 2.07 Å). Hydrophobic interactions were seen between second phenyl and m-(methoxy)benzyl moiety with LEU348, Val402, Leu401, Val434 and with carbon chain of Glu398. A pi-pi stacking interaction of second phenyl moiety with HIS405 was also observed. Similar type of zinc chelating, hydrogen bonding and pi-pi stacking interaction were observed in compounds 2.5CYH_N and 2.5HEXA_N while cyclo-hexyl and n-hexyl group groups were found in hydrophobic pocket formed by Leu348, Val402, Leu401, Val434 and carbon chain of Glu398.

Urea series: Binding scores of urea series are given in **Table 2.3-7**. Compound 3.3NTPN with *m*-nitro substitution showed highest binding score compared to others in all the software. *o*-methoxy substituted compound 3.2MXPN scored low in Autodock (-8.06 Kcal/mol) and FRED (-94.06), *o*-methyl substituted compound 3.2TOPN scored low in Glide SP (-6.0 Kcal/mol), and second lowest in FRED (-98.66). Methoxy substituted compounds 3.2MXPN, 3.3MXPN and 3.4MXPN scored low in Glide XP. It was found that *ortho* substituted compounds are less tolerated mainly due to steric clash in S2 pocket therefore scored low in all software compared to *para* and *meta* counterparts. The *meta* substitution tends to enhance binding than *para*. Electron withdrawing or releasing group had no significant effect on binding but bulkiness at *meta* or *para* position enhanced binding scores.

Figure 2.3-3 illustrates the interaction pattern of some TACE inhibitors of urea series. For compound 3.3NTPN, an interaction of hydroxamic acid with zinc was observed in bidentate fashion (C=O, 3...Zn; 2.10 Å, N-O, 4 ... Zn; 2.36 Å). Carbonyl oxygen (C=O, 2) of urea did not interact with Leu348 or Gly349 residue. Instead it faced hydrophobic side chain of Leu348 and Ala439. A hydrogen bond was formed between hydroxyl hydrogen with Glu406 (O-H, 10...COO, E406; 2.43 Å) and amine hydrogen with backbone carbonyl group of Gly349 (N-H, 5...CO, G349; 1.94 Å). Hydrophobic interactions were seen by phenyl moiety with Ala439, Leu348, Val402 and Leu401. No pi-pi stacking interaction of another phenyl moiety with His405 was observed.

Table 2.3-7: In silico screening results of Urea series

Code	Autodock	Glide SP		Glide XP		FRED
	Score	Score	Energy	Score	Energy	chemscore3
3.4CLPN	-9.06	-6.61	-46.5	-5.78	-48.56	-107.13
3.2MXPN	-8.06	-7.6	-55.58	-5.71	-48.62	-94.06
3.3MXPN	-9.35	-7.6	-56.45	-5.54	-56.91	-110.58
3.4MXPN	-8.82	-7.08	-52.06	-5.21	-54.37	-106.99
3.4NTPN	-10.22	-7.06	-57.25	-6.8	-56.91	-100.16
3.3NTPN	-12.17	-7.46	-57.68	-6.71	-59.4	-104.7
3.4FLPN	-8.86	-6.64	-48.32	-6.68	-51.66	-106.2
3.2FLPN	-8.71	-7.31	-47.86	-6.61	-53.33	-100.5
3.3FLPN	-8.84	-6.33	-47.12	-6.72	-54.51	-105.64
3.2TOPN	-8.81	-6	-53.16	-6.33	-50.48	-98.66
3.4TOPN	-9.04	-6.55	-54.88	-7.06	-52.79	-101.8
3.3TOPN	-9.44	-6.2	-56.72	-7.11	-55.39	-111.43
ASH	-9.31	-7.54	-59.05	-7.06	-52.93	-111.68

All values are expressed in Kcal/mol
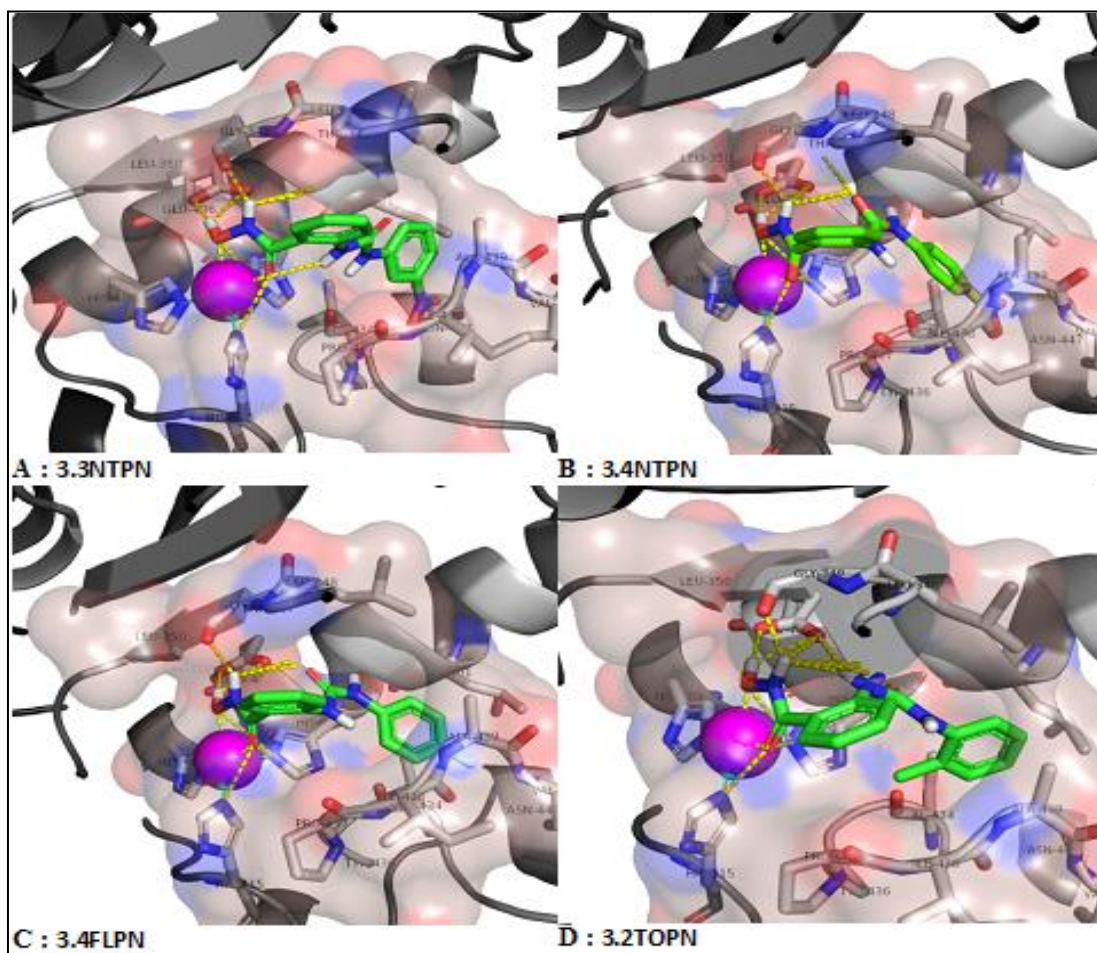
Figure 2.3-3: Binding pose of potent Urea derivatives against TACE, plotted with Pymol [17]. Inhibitors are represented in green. Important binding residues of TACE are represented in gray. Zinc atom is presented as pink sphere and rest of protein as cartoon model. Yellow dotted line represents hydrogen bond.

The *m*-Nitro moiety fits in the pocket below S2 site instead of S3 site. Pocket below S2 site is relatively polar as it is formed mainly by backbone amide, while the hydrophobic side chains are oriented away from the pocked. In compound 3.4NTPN, hydroxamic acid interaction with zinc in bidentate fashion was observed (C=O, 1...Zn; 2.16 Ǻ, N-O, 3 … Zn; 2.34 Ǻ). Carbonyl oxygen (C=O, 2) of urea did not interact of Leu348 or Gly349 residue. A hydrogen bond was formed between hydroxyl hydrogen with Glu406 (O-H, 10...COO, E406; 2.43 Ǻ) and amine hydrogen with backbone carbonyl group of Gly349 (N-H, 5...CO, G349; 2.07 Ǻ). Hydrophobic interactions were seen by phenyl moiety with Leu401, Val434, and Ala439. A very poor pi-pi stacking interaction of other phenyl moiety with His405 was observed. Binding pattern of 3.4FLPN was similar to 3.4NTPN. Urea carbonyl carbon orientation was same in both cases. The distance between urea oxygen and backbone amine of Leu348 or Gly349 was more than 3.5 Å so no hydrogen bond was formed between them. In compound 3.2TOPN, hydroxamic acid interacted with zinc in bidentate fashion (C=O, 1...Zn; 2.16 Ǻ, N-O, 3 … Zn; 2.31 Ǻ). Carbonyl oxygen (C=O, 2) of urea did not interact with Leu348 or Gly349 residue but one of the amine of urea moiety made hydrogen bonding interaction with carboxylic group of Glu406 (N-H, 6...COO, E406; 2.33 Ǻ). Also a hydrogen bond was formed between hydroxyl hydrogen with Glu406 (O-H, 10...COO, E406; 2.45 Ǻ) and amine hydrogen with backbone carbonyl group of Gly349 (N-H, 5...CO, G349; 2.03 Ǻ). Hydrophobic interactions were seen between phenyl moiety and Ala439, Leu401, Leu348 and Val402. No pi-pi stacking interaction of second phenyl moiety with His405 was observed. Methyl group at *ortho* position of phenyl group was not favored in polar environment.

Compared to compound ASH, many urea derivatives ranked low in docking mainly due to planar nature of urea which leads to steric clash when carbonyl group of urea tried to interact with Gly349 amine. Therefore, carbonyl group of urea series was unable to form hydrogen bond with Gly349 amine while benzophenone series showed this interaction. Another reason may be the entropic penalty which is high in urea series than benzophenone series as relatively polar urea group disfavors entry into hydrophobic S2 cavity. It was observed that phenoxy group in benzophenone helped the molecule to bend and explore S3 pocket. While in urea series such L bend was possible with *meta* substitution only. Due to increased length in urea series, the phenyl ring was pushed further in S2 cavity resulting in steric clash if *para* position is substituted. The binding pattern of ASH, 2.52B_N and 3.4NTPN are represented in **Figure 2.3-4** clearly indicating possible reason differences in binding scores.
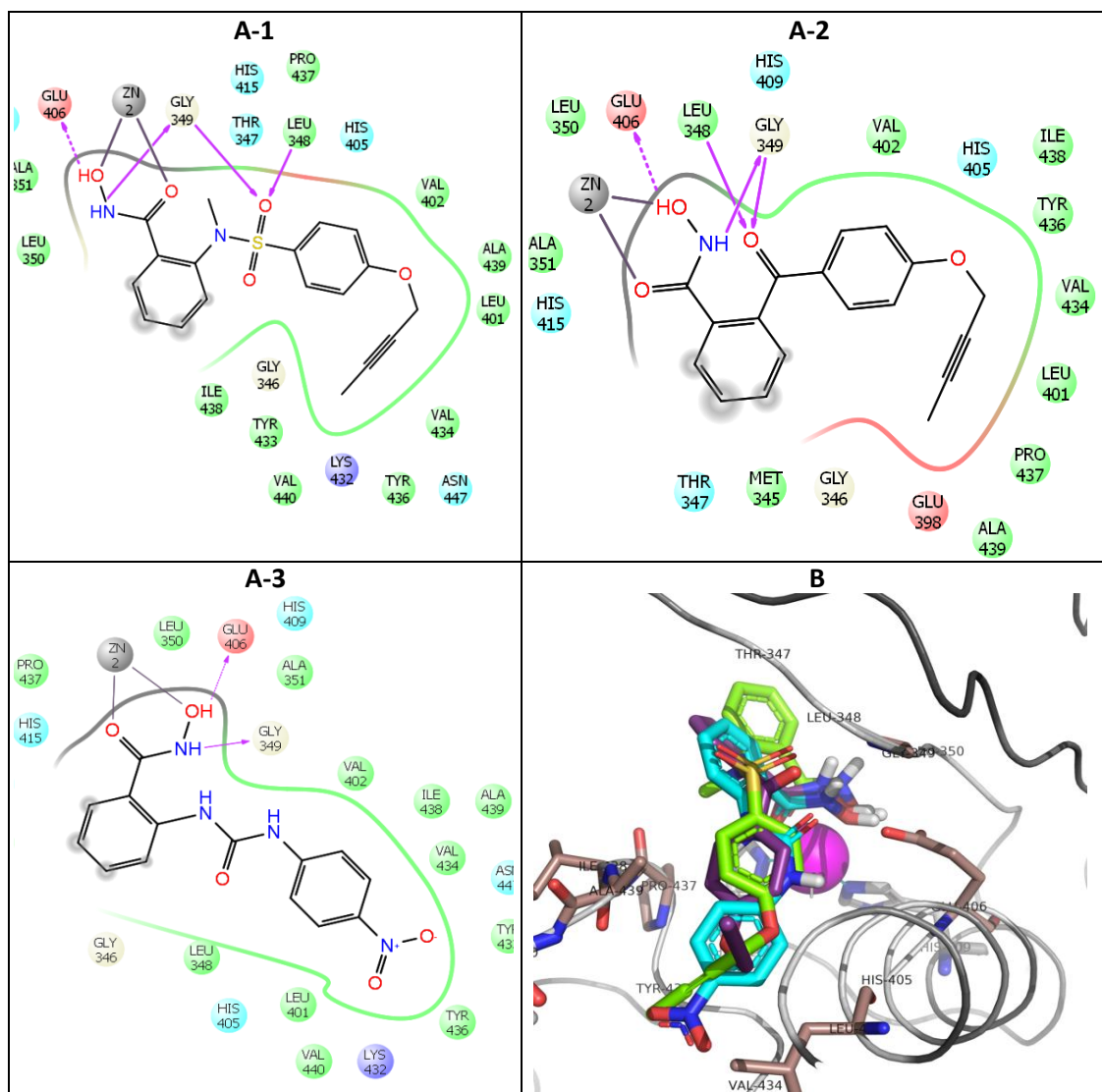
Figure 2.3-4: A. 2D interaction plot of (**1**) ASH with TACE, **(2)** 2.52B_N with TACE, **(3)** 3.4NTPN with TACE; **B**. Docked pose overlay of ASH (magenta), 2.52B_N (green) and 3.4NTPN (cyano)

2.3.3 References

[1] Levin, J.; Chen, J.; Du, M.; Nelson, F.; Killar, L.; Skala, S.; Sung, A.; Jin, G.; Cowling, R.; Barone, D. Anthranilate sulfonamide hydroxamate TACE inhibitors. Part 2: SAR of the acetylenic P1' group. *Bioorg Med Chem Lett.* **2002,** 12(8), 1199-1202.

[2] Fernandes, J.; Gattass, C.R. Topological polar surface area defines substrate transport by multidrug resistance associated protein 1 (MRP1/ABCC1). *J Med Chem.* **2009,** 52(4), 1214-1218.

[3] Tetko, I.V.; Tanchuk, V.Y. Application of associative neural networks for prediction of lipophilicity in ALOGPS 2.1 program. *J Chem Inf Comput Sci.* **2002,** 42(5), 1136-1145.

[4] Leeson, P. Drug discovery: Chemical beauty contest. *Nature.* **2012,** 481(7382), 455-456.

[5] Veber, D.F.; Johnson, S.R.; Cheng, H.Y.; Smith, B.R.; Ward, K.W.; Kopple, K.D. Molecular properties that influence the oral bioavailability of drug candidates. *J Med Chem.* **2002,** 45(12), 2615-2623.

[6]   Sander, T.; Freyss, J.; von Korff, M.; Reich, J.R.; Rufener, C. OSIRIS, an entirely in-house developed drug discovery informatics system. *J Chem Inf Model.* **2009,** 49(2), 232-246.

[7]   Sander, T.L. Osiris Property Explorer (2001-2013). http://www.organic-chemistry.org/prog/peo/ (Accessed November 18 2013).

[8]   Govinda Rao, B.; Bandarage, U.K.; Wang, T.; Come, J.H.; Perola, E.; Wei, Y.; Tian, S.K.; Saunders, J.O. Novel thiol-based TACE inhibitors: rational design, synthesis, and SAR of thiol-containing aryl sulfonamides. *Bioorg Med Chem Lett.* **2007**,17(8), 2250-2253.

[9]   Maiti, R.; Van Domselaar, G.H.; Zhang, H.; Wishart, D.S. SuperPose: a simple server for sophisticated structural superposition. *Nucleic Acids Res.* **2004,** 32(suppl 2), W590-W594.

[10] Marvin 5.2.6, ChemAxon (http://www.chemaxon.com), **2009**.

[11] Shelley, J.C.; Cholleti, A.; Frye, L.L.; Greenwood, J.R.; Timlin, M.R.; Uchimaya, M. Epik: a software program for pK a prediction and protonation state generation for drug-like molecules. *J Comput Aid Mol Des.* **2007,** 21(12), 681-691.

[12] Halgren, T.A.; Murphy, R.B.; Friesner, R.A.; Beard, H.S.; Frye, L.L.; Pollard, W.T.; Banks, J.L. Glide: a new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening. *J Med Chem.* **2004,** 47(7), 1750-1759.

[13] Morris, G.M.; Huey, R.; Lindstrom, W.; Sanner, M.F.; Belew, R.K.; Goodsell, D.S.; Olson, A.J. AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J Comput Chem.* **2009,** 30(16), 2785-2791.

[14] OMEGA version 2.3.2, OpenEye Scientific Software, Inc., Santa Fe, NM, USA, www.eyesopen.com, **2010**.

[15] Skipper, P.L.; Tannenbaum, S.R.; Thilly, W.G.; Furth, E.E.; Bishop, W.W. Mutagenicity of hydroxamic acids and the probable involvement of carbamoylation. *Cancer Res.* **1980,** 40(12), 4704-4708.

[16] Pal, D.; Saha, S. Hydroxamic acid–A novel molecule for anticancer therapy. *J Adv Pharm Technol Res.* **2012,** 3(2), 92-99.

[17] The PyMOL Molecular Graphics System, Version 1.6 Schrödinger, LLC, **2012**.

2.4 Synthesis of TACE Inhibitors

Based on different synthetic strategies, synthesis of molecules given in **section 2.3** was carried out using two different schemes. All benzophenone derivatives were synthesized according to **Scheme 2.4-1** using phthalic anhydride and anisol as starting materials whereas all urea derivatives were synthesized according to **Scheme 2.4-2** using anthranilic acid as starting material.

For Benzophenone series, anisol undergoes Friedel craft acylation with phthalic anhydride in presence of Lewis acid (Aluminum chloride, $AlCl_3$). Initially acylium ion is formed by action of $AlCl_3$ on phthalic anhydride, which acts as an electrophile and attacks on electrons of aromatic C=C. As methoxy group of anisol is activator and *ortho/para* director, electrophilic aromatic substitution can occur on *ortho* or *para* position but due to steric clash, *para* substitution takes place. Later, methyl moiety of acylated anisol was cleaved in presence of strong acid (HBr). Free carboxylic acid of formed phenol was protected and phenolic hydroxyl was coupled with different alkyl and aryl alcohols by Mitsunobu reaction. Later carboxylic group was deprotected, activated to acyl chloride and coupled with hydroxyl amine to yield desired hydroxamic acid.
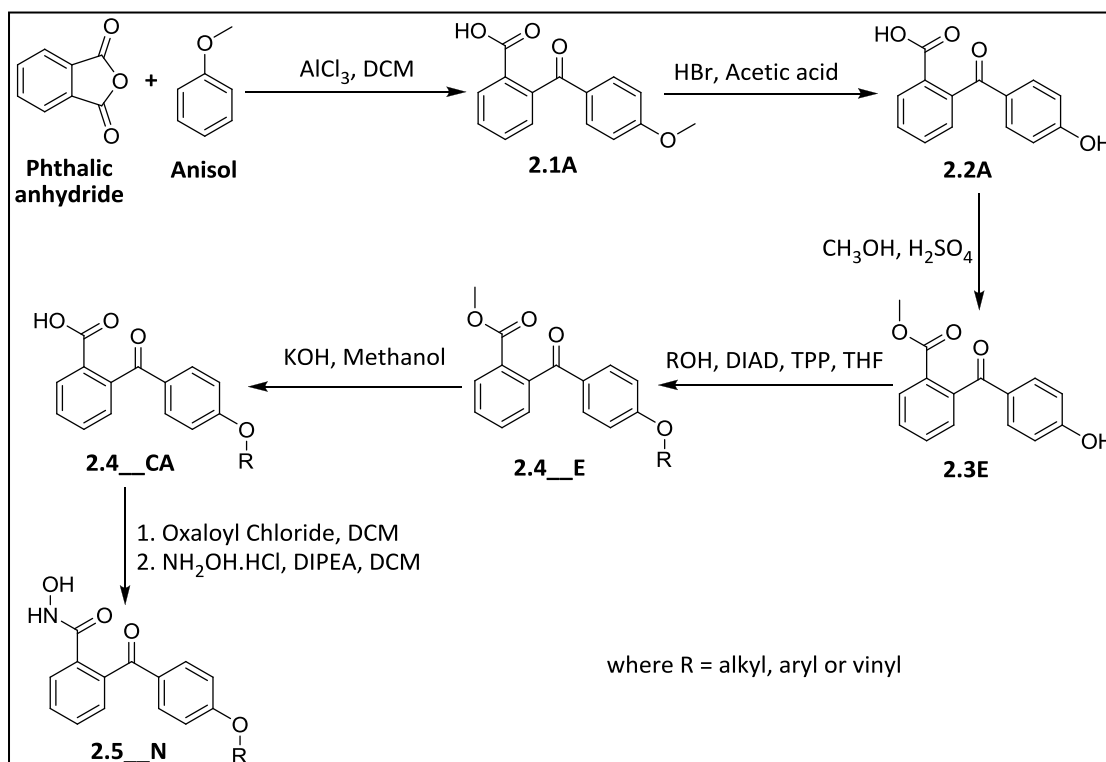
Scheme 2.4-1: Synthetic route of Benzophenone series

For urea derivatives, initially carboxylic acid moiety of anthranilic acid was protected by converting it into ester and then amino group was reacted with various phenyl isocynates to form corresponding urea derivatives. Later hydroxamic acid was formed by reacting esters with hydroxyl amine under basic condition.
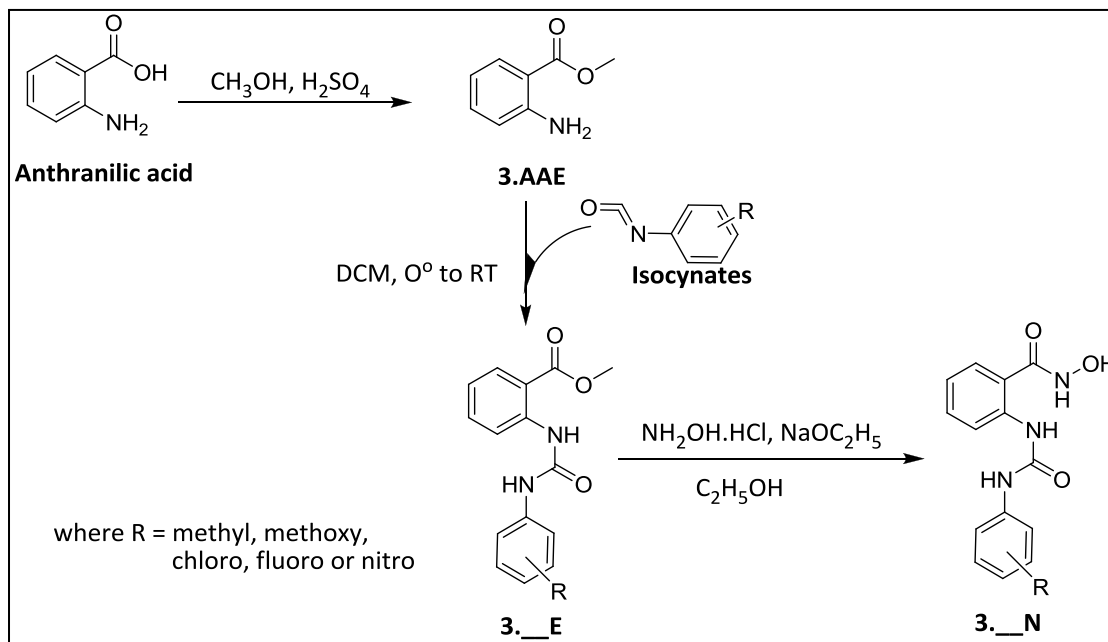
Scheme 2.4-2: Synthetic route of urea series

2.4.1 Methodology

2.4.1.1 Synthesis of Benzophenone derivatives

Synthesis of 2-[(4-methoxyphenyl)carbonyl]benzoic acid (2.1A) [1]: Solution of phthalic anhydride (14.81 g, 100 mM) in anisol (27.17 ml, 250 mM) was cooled in ice bath. Aluminium chloride (20.00 gm, 150 mM) was added in portion with stirring. After complete addition the reaction mixture was continued to stir for 1 hr. Later ice bath was removed and reaction was further stirred for 1 hr at 60°C. Progress of the reaction was monitored with TLC using ethyl acetate, n-hexane (3:1) as solvent system. Reaction mixture was then cooled and excess aluminium chloride was decomposed with 500 ml of ice and 1N hydrochloric acid mixture (50:50). Product was then extracted with dichloromethane (DCM) (3 X 200 ml). Organic layer was combined and washed with water. Product was later extracted from organic phase with 1N sodium hydroxide solution (4 X 100 ml). Aqueous phase was then combined and residual amount of anisol was removed with DCM (2 X 100 ml). Cold aqueous phase was then acidified with concentrated hydrochloric acid (HCl) (~75 ml). Solid separated was filtered and washed with ice cold water till free from acid. Product 2.1A was recrystalized with aqueous acetic acid water (50%) and crystals were washed with cold water till free from acetic acid.

Synthesis of 2-[(4-hydroxyphenyl)carbonyl]benzoic acid (2.2A) [1]: 48% aqueous hydrogen bromide (HBr) (75 ml) was added to the solution of compound 2.1A (15.37 g, 60 mM) in glacial acetic acid (40 ml) and refluxed (115-120°C) for 26 hrs. Reaction was monitored with TLC using ethyl acetate, n-hexane (1:1) as solvent system. On completion, reaction mixture was cooled to room temperature and carefully poured into ice cold water (250 ml). Product was later extracted with ethyl acetate (3 X 200 ml). Combined organic layer was then washed with water (100 ml) and dried over anhydrous sodium sulfate. Concentration of organic layer by vacuum distillation yielded intermediate 2.2A as white solid.

Synthesis of methyl 2-[(4-hydroxyphenyl)carbonyl]benzoate (2.3E) [1]: To a solution of compound 2.2A (14.05 g, 58 mM) in methanol (150 ml) was added concentrated sulfuric acid (2 ml). The reaction mixture was refluxed for 20 hrs. Reaction progress was monitored with TLC using ethyl acetate, n-hexane (1:1) as solvent system. On completion, reaction mixture was concentrated and residue was diluted with ice cold water (200 ml), solid sodium bicarbonate (NaHCO$_3$) was added until the solution became slightly basic and later extracted with ethyl acetate (2 X 100 ml). Organic layer were combined, washed with water (100 ml), dried over anhydrous sodium sulfate and concentrated to get crystals of compound 2.3E.

Synthesis of methyl 2-[(4-aryl/alkyloxyphenyl)carbonyl]benzoate (2.4__E) [1]: To a stirred, chilled solution of 2.3E (1.02 g, 4 mM) and triphenyl phosphine (TPP) (1.57 g, 6 mM) in dry tetrahydrofuran (THF) (10 ml) under nitrogen environment, was added drop wise with diisopropyl azodicarboxylate (DIAD) (1.18 ml 6 mM). Reaction mixture was stirred in ice bath for 30 min, later ice bath was removed and stirring was continued for additional 15 min. Later to this reaction mixture, different alcohols (6 mM) in dry THF (5 ml) were added. The reaction was allowed to stir for 55°C for 2 days and progress of the reaction was monitored with TLC using ethyl acetate, n-hexane (3:7) as solvent system. On completion of the reaction, solvent was evaporated under reduced pressure. Concentrated reaction mass was then quenched with 5 % HCl (10 ml) and extracted with diethyl ether (3 X 10ml). Combined ether layer was evaporated and product 2.4__E was isolated by column chromatography using silica gel (230-400 mesh) and 5% ethyl acetate / n-hexane mixture as mobile phase.

Synthesis of 2-[(4-aryl/alkyloxyphenyl)carbonyl]benzoic acid (2.4__CA): Obtained esters (2.4_E) was dissolved in 10 ml methanol with potassium Hydroxide (0.45 g, 8 mM) and stirred at 35°C for 6 hrs. Reaction progress was monitored using TLC with ethyl acetate, n-hexane (3, 7) as solvent system. On completion, reaction mass was concentrated under reduced pressure. Residues were dissolved in 25 ml water and unreacted ester was removed

by extraction with 10 ml diethyl ether. Aqueous extract was acidified and extracted with diethyl ether (2 X 15 ml). On concentrating ether layer under vacuum yielded product 2.4__CA as white solid in practically pure form.

Synthesis of 2-[(4-aryl/alkyloxyphenyl)carbonyl]benzohydroxymic acid (2.5__N) [2]: To 1 mM of compound 2.4_CA dissolved in 10 ml of DCM, added 5 mM (0.43 ml) of oxalyl chloride and drop of dimethyl formamide (DMF). Reaction mixture was refluxed till the completion of the reaction (~ 30 to 60 min) and progress was monitored using TLC with ethyl acetate, n-hexane (3, 7) as solvent system. Excess oxalyl chloride was removed by distillation. Crude acyl chloride was then dissolved in 10 ml of DCM. Solution of hydroxylamine hydrochloride (0.09 g, 1.3 mM) and N,N-Diisopropylethylamine (Hünig's base, DIPEA) (0.31 ml, 3 mM) in 5 ml DCM was added drop wise to stirring solution of acyl chloride kept in ice bath. Reaction mass was allowed to stir for 6 hrs. On completion of the reaction confirmed by TLC with ethyl acetate, n-hexane (3, 7) as solvent system, reaction mass was diluted with 15 ml water and extracted with ethyl acetate (2 X 10 ml). Separated organic layers were combined dried over anhydrous sodium sulfate and concentrated under reduced pressure. The residue was purified via column chromatography on silica gel (230-400 mesh) eluting with 5% ethyl acetate / n-hexane mixture to get compound 2.5_N.

2.4.1.2   Synthesis of urea derivatives

Synthesis of Methyl 2-aminobenzoate (3.AAE) [1]: To a solution of antranilic acid (20.57 g, 150 mM) in methanol (150 ml) was added concentrated sulfuric acid (2 ml). The reaction mixture was refluxed for 24 hrs and progress was monitored with TLC using ethyl acetate, n-hexane (1:1) as solvent system. After the completion of the reaction, the reaction mixture was concentrated. Residue was diluted with ice cold water (200 ml) and solid sodium bicarbonate ($NaHCO_3$) was added until the solution becomes slightly basic. The aqueous layer was then extracted with ethyl acetate (2 X 100 ml). Combined organic layer was washed with water (100 ml), dried over anhydrous sodium sulfate and concentrated to get the compound 3.AAE in liquid form.

Synthesis of Methyl 2-[(substituted-phenylcarbamoyl)amino]benzoate (3.__E) [3]: Solution of compound 3.AAE (1.0 g, 6.6 mM) in DCM (15 ml) was stirred at 0°C. Isocynates (6.6 mM) were added and reaction mixture was stirred for 4-6 hrs. Progress of the reaction was monitored by TLC using ethyl acetate, n-hexane (1:1) as solvent system. After the completion of the reaction solid precipitated were filtered and washed with 20 ml diethyl ether. Formed urea derivatives (3.__E) were used without further purification.

Synthesis of N-hydroxy-2-[(substituted-phenylcarbamoyl)amino]benzamide (3.__N) [4]: To the solution of sodium ethoxide (0.27 g, 4 mM) in ethanol 5 ml was added finely powdered hydroxylamine hydrochloride (0.28 g, 4 mM) and stirred for 15 min. To this reaction mixture compound 3.__E (3mM) was added and refluxed for 6-8 hrs. On completion of the reaction, confirmed by TLC with ethyl acetate, n-hexane (1:2) as solvent system, reaction mixture was diluted with 10 ml of water. Precipitated product was collected dried and purified by column chromatography using silica gel (60-120 mesh) as stationary phase and 5% ethyl acetate / n-hexane mixture as mobile phase.

2.4.2 Results and Discussion

2.4.2.1 Synthesis of Benzophenone derivatives

2-[(4-methoxyphenyl)carbonyl]benzoic acid (2.1A) was prepared by Friedel craft acylation of anisol with phthalic anhydride in presence of $AlCl_3$. This reaction is an example for electrophilic aromatic substitution reaction. Intermediate acylium ion attacks aromatic electrons. Out of possible *ortho* or *para* substitution only *para* substitution occurs due to steric clash with methoxy group and is confirmed by aromatic proton splitting pattern in ^{1}H-NMR (doublet peaks at 7.70 and 6.90 δ (ppm). Formed compound showed carbonyl bond stretching at 1693.50 and 1666.50 cm^{-1} in IR spectroscopy for carboxylic acid and ketone moiety respectively.

Cleavage of ether bond of 2.1A was carried out using HBr. The reaction follows SN2 mechanism. Disappearance of 3 protons of methoxy group (3.86 ppm (s, OCH_3, 3H), present of 2.1A in ^{1}H-NMR clearly indicates cleavage of ether bond to form phenol. Obtained phenol 2.2A was practically pure and used without further purification. Carboxylic acid OH group of 2.2A was required to be protected in order to react phenolic OH by Mitsunobu reaction. Carboxylic group was protected by converting to ester by Fisher esterification which is an example for nucleophilic acyl substitution reaction. Disappearance of broad OH stretching peak of carboxylic acid and presence of sharp OH stretch of phenol at 3336.8 cm^{-1}, carbonyl stretching of ester at 1718.58 cm^{-1} confirmed the compound 2.3E. Further it was confirmed by a singlet peak at 3.64 δ (ppm) correspond 3 protons of methyl ester in ^{1}H NMR. After carboxylic acid protection, phenolic OH was coupled with various alkyl and aryl alcohols under Mitsunobu reaction condition. Intermediate betaine formed by nucleophilic attack of Triphenyl phosphine (TPP) on Diisopropyl azodicarboxylate (DIAD), deprotonates phenolic OH and forms intermediate ion complex. On addition of another alcohol, intermediate ion complex gets reduced to hydrazine by deprotonating second alcohol. Alkoxide ion then

attacks oxyphosphonium ion followed by resonance leading to product 2.4__E and triphenylphosphine oxide. Product was separated by column chromatography. Yield of the product 2.4__E varied from 0-98% and is tabulated in **Table 2.4-1**. Reactivity found to depend on nature of alcohol and bulkiness. Alkyl alcohols were more reactive than aryl alcohol, while reactivity was high for primary alkyl alcohols followed by secondary and tertiary. 1(2-Hydroxy Ethyl) Morpholine and 1(2-Hydroxy Ethyl) Phthalimide were found to be unreactive. Disappearance of hydroxyl stretch (3336.85 cm^{-1}) in IR and presence of proton peaks corresponding to second alcohol confirmed the product formation.

Table 2.4-1: Isolated yield of methyl 2-{4-(alkyl/aryl)oxybenzoyl}benzoates (2.4__E)

Sl.no	Second Alcohol used	Code of product	% yield
1	Allyl Alcohol	2.4ALL_E	92.81
2	t-Butanol	2.4TtBA_E	23.21
3	2-Butyn-1-ol	2.42B_E	84.33
4	3-Butyn-1ol	2.43B_E	38.11
5	Cinnamoyl Alcohol	2.4CINN_E	11.01
6	Ethanol	2.4ETHAN_E	96.73
7	Furfuryl Alcohol	2.4FUR_E	3.10
8	cyclo Hexanol	2.4CYH_CE	39.89
9	n-Hexanol	2.4HEX_E	92.53
10	1(2-Hydroxy Ethyl) Morpholine	2.4HEM_E	0
11	1(2-Hydroxy Ethyl) Phthalimide	2.4HEPH_E	0
12	1(2-Hydroxy Ethyl) Piperidine	2.4HEP_E	31.29
13	2-Methoxy Ethanol	2.4ME_E	98.62
14	3-Methoxy Benzyl Alcohol	2.4MBA_E	74.39
15	2-Methyl Propanol	2.4MP_E	91.24
16	4-Nitro Benzyl Alcohol	2.44NBA_E	6.39
17	n-Pentanol	2.4PENT_E	92.68
18	n-Propanol	2.4PRO_E	98.05

Table 2.4-2: Isolated yield of 2-{4-(alkyl/aryl)oxybenzoyl}benzoic acids (2.4__CA)

Code of product	% yield	Code of product	% yield
2.4ALL_CA	83.02	2.4ME_CA	81.04
2.4tBA_CA	83.04	2.4MBA_CA	94.60
2.42B_CA	87.64	2.4MP_CA	95.52
2.43B_CA	82.48	2.44NBA_CA	82.97
2.4ETHAN_CA	97.54	2.4PENT_CA	74.27
2.4CYH_CA	88.87	2.4PRO_CA	86.10
2.4HEX_CA	97.68		

Obtained ester 2.4__E was hydrolyzed using methanolic potassium hydroxide. Hydrolyzed compound 2.4__CA was isolated by diethyl ether on acidification. Yield of

hydrolyzed products are reported in **Table 2.4-2.** Disappearance of 3 protons of methyl ester [3.64 δ (ppm)] in ^{1}H NMR, and broad OH stretch of carboxylic acid in IR spectra confirms the the formation of 2.4_CA.

Due to presence of unprotected keto group, direct coupling of ester with hydroxyl amine was not feasible as keto group is more reactive than ester. Even free carboxylic acids are less reactive than keto therefore activation of carboxylic group to more reactive acyl chlorides was necessary for selective reaction with amine. Compounds 2.4_CA were converted to acyl chlorides with oxalyl chloride. Hydroxylamine hydrochloride was solubilised in DCM with the aid of DIPEA. The free base hydroxylamine was added drop by drop to acyl chloride solution in DCM kept in ice bath. Nitrogen of hydroxylamine makes nucleophilic attack on carbonyl carbon of acyl chloride. Followed by elimination of chloride ion to yield hydroxamic acids (2.5__N). The reaction follows SN2 mechanism. After completion of the reaction, water was added to remove unreacted hydroxyl amine and DIPEA. Formed hydroxamic acid was isolated with ethyl acetate and purified with column chromatography. Yield of the hydroxamic acid are reported in **Table 2.4-3**. Compounds were characterized by ferric chloride test and by ^{1}H NMR spectroscopy.

Table 2.4-3: Isolated yield of N-hydroxy-2-{4-(alkyl/aryl)oxybenzoyl}benzamide (2.5__N)

Code of product	% yield	Code of product	% yield
2.4ALL_N	26.91	2.5ME_N	44.40
2.5tBA_N	4.56	2.5MBA_N	29.15
2.52B_N	16.16	2.5MP_N	22.34
2.53B_N	18.36	2.5PENT_N	25.21
2.5ETHAN_N	31.55	2.5PRO_N	22.13
2.5CYH_N	47.14	2.5ANS_N	28.12
2.5HEX_N	49.79		

Spectral Information of intermediates and products

2-[(4-methoxyphenyl)carbonyl]benzoic acid (2.1A): Yield 60.31%, MP 126-128°C; IR (cm^{-1}) (KBr): 3012.97(s,b, OH carboxylic acid), 1693.50(S, C=O carboxylic acid), 1666.50 (S, C=O keto). ^{1}H-NMR δ (ppm) (400Mhz, CDCl$_3$, TMS): 3.86 (s, OCH$_3$, 3H), 6.90 (d, ArH, 2H), 7.36 (d, ArH, 1H), 7.55 (t, ArH, 1H), 7.65 (t, ArH, 1H), 7.70 (d, ArH, 2H), 8.08 (d, ArH, 1H).

2-[(4-hydroxyphenyl)carbonyl]benzoic acid (2.2A): Yield: 78.85%, MP 198-202°C; IR (cm^{-1}) (KBr): (KBr):3145.90 (s,b, aromatic OH), 1693.50 (S, C=O acid), 1647.21 (S, C=O keto). ^{1}H-NMR δ (ppm) (400Mhz, CDCl$_3$, TMS): 6.80 (d, ArH, 2H) 7.30 (d, ArH, 1H), 7.53-7.65 (m, ArH, 4H) 7.98 (t, ArH, 1H), 10.05 (s, phenolic OH, 1H), 12.75 (s, carboxylic OH, 1H)

Methyl 2-[(4-hydroxyphenyl)carbonyl]benzoate (2.3E): Yield: 98.64%, MP 132-134°C; IR (cm^{-1}) (KBr):3336.85 (s, aromatic OH), 1718.58 (S, C=O ester), 1643.35 (S, C=O keto). ^{1}H-NMR δ (ppm) (400Mhz, CDCl$_3$, TMS): 3.64 (s, COOCH$_3$, 3H), 6.82-6.84 (m, ArH, 2H) 7.34-7.37 (m, ArH, 1H), 7.55-7.68 (m, ArH, 4H) 7.98- 8.00 (m, ArH, 1H), 10.08 (s, OH, 1H)

Methyl 2-(4-isobutoxybenzoyl)benzoate (2.4MP_E): IR (cm^{-1}) (KBr): 1737.45 (S, C=O ester), 1693.50 (S, C=O keto); ^{1}H-NMR δ (ppm) (400Mhz, CDCl$_3$, TMS): 0.98 (d, CH$_3$, 6H), 2.03 (m, CH, 1H), 3.59 (s, OCH$_3$, 3H), 3.83 (d, CH$_2$, 2H), 7.03 (d, ArH, 2H), 7.44 (d, ArH, 1H), 7.59 (d, ArH, 2H), 7.70 (m, ArH, 2H), 7.99 (d, ArH, 1H).

2-(4-ethoxybenzoyl)-benzoic acid (2.4ETHAN_CA): IR (cm^{-1}) (KBr): 2974.23 (s,b OH carboxylic acid), 1691.57 (s, C=O carboxylic acid), 1662.64 (s, C=O keto). ^{1}H-NMR δ (ppm) (400Mhz, CDCl$_3$, TMS): 1.38 (t, CH$_3$, 3H), 4.09 (q, OCH$_2$-, 2H), 6.94 (d, ArH, 2H), 7.33 (d, ArH, 1H), 7.58 (t, ArH, 1H), 7.61 (d, ArH, 2H), 7.66 (t, ArH, 1H), 7.89 (d, ArH, 1H)

2-(4-(but-2-ynyloxy)benzoyl)-benzoic acid (2.42B_CA): IR (cm^{-1}) (KBr): 2870.08 (s,b OH carboxylic acid), 1687.71 (s, C=O carboxylic acid), 1660.71 (s, C=O keto).

2-(4-(allyloxy)benzoyl)-*N*-hydroxybenzamide (2.5ALL_N): IR (cm^{-1}) (KBr): 3082.25 (s, OH hydroxamic acid), 1737.86, 16 (S, C= O hydroxamic acid), 1716.65 (S, C=O keto). ^{1}H-NMR δ (ppm) (400Mhz, CDCl$_3$, TMS): 1.88 (s, CH$_2$, 2H), 4.79 (d, CH$_2$, 2H), 6.08 (m, CH, 1H), 7.17 (d, ArH, 2H), 7.55 (d, ArH, 2H), 7.61 (d, ArH, 1H), 7.97-8.00 (m, ArH, 2H), 8.37(d, ArH, 1H)

2-(4-(but-2-ynyloxy)benzoyl)-*N*-hydroxybenzamide (2.52B_N): IR (cm^{-1}) (KBr): 3053.91(s, OH hydroxamic acid), 1732.08 (S, C=O hydroxamic acid), 1716.65 (S, C=O keto). ^{1}H-NMR δ (ppm) (400Mhz, CDCl$_3$, TMS): 1.88 (t, CH$_3$, 1H), 4.79 (d, CH$_2$, 2H), 7.17 (d, ArH, 2H), 7.55 (d, ArH, 2H), 7.61 (m, ArH, 1H), 7.97-8.00 (m, ArH, 2H), 8.36-8.38 (m, ArH, 1H)

2-(4-(but-3-ynyloxy)benzoyl)-*N*-hydroxybenzamide (2.53B_N): IR (cm^{-1}) (KBr): 3041.74(s,b, OH hydroxamic acid), 1732.08 (S, C=O hydroxamic acid), 1716.65 (S, C=O keto). ^{1}H-NMR δ (ppm) (400Mhz, CDCl$_3$, TMS): 1.25 (s, CH, 1H), 2.72 (t, CH$_2$, 2H), 4.19 (t, CH$_2$, 2H), 7.14 (d, ArH, 2H), 7.55 (d, ArH, 2H), 7.59-7.61(m, ArH, 1H), 7.97-7.99 (m, ArH, 2H), 8.36-8.38 (m, ArH, 1H)

2-(4-ethoxybenzoyl)-*N*-hydroxybenzamide (2.5ETHAN_N): IR (cm^{-1}) (KBr): 3049.46 (s, OH hydroxamic acid), 1732.08 (S, C=O hydroxamic acid), 1716.65 (S, C=O keto).

2-(4-methoxybenzoyl)-*N*-hydroxybenzamide (2.5ANS_N): IR (cm^{-1}) (KBr): 3078.39 (s, OH hydroxamic acid), 1745.58 (S, C=O hydroxamic acid), 1732.08 (S, C=O keto).

2-(4-propoxybenzoyl)-*N*-hydroxybenzamide (2.5PRO_N): IR (cm^{-1}) (KBr): 3068.75 (s, OH hydroxamic acid), 1770.65 (S, C=O hydroxamic acid), 1737.86 (S, C=O keto).

2-(4-pentoxybenzoyl)-*N*-hydroxybenzamide (2.5PENT_N): IR (cm^{-1}) (KBr): 3043.23 (s, OH hydroxamic acid), 1778.37 (S, C=O hydroxamic acid), 1745.58 (S, C=O keto).

2-(4-(2-methyl)propoxybenzoyl)-*N*-hydroxybenzamide (2.5MP_N): IR (cm^{-1}) (KBr): 3048.29 (s, OH hydroxamic acid), 1766.61 (S, C=O hydroxamic acid), 1738.71 (S, C=O keto). ^{1}H-NMR δ (ppm) (400Mhz, CDCl$_3$, TMS): 1.04 (d, CH3, 6H), 2.09 (m, CH, 1H), 3.84 (d, CH2, 2H), 7.11 (d, ArH, 2H), 7.54 (d, ArH, 2H), 7.61 (d, ArH, 1H), 7.99 (m, ArH, 2H), 8.34 (d, ArH, 1H).
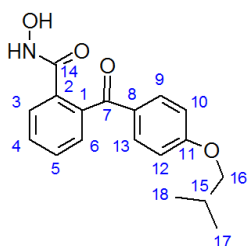
^{13}C NMR δ (ppm) (100Mhz, DMSO and CDCl$_3$): 162.93 (C7), 160.30 (C14), 156.61 (C11), 135.64 (C1), 133.88 (C2), 130.59 (d, C9 and C13d), 128.00 (C8), 127.44 (C5), 126.89 (C4), 122.55 (C3), 122.09 (C6), 114.57 (d, C10, C12), 73.89 (C16), 27.66 (C15), 18.93 (d, C17 and C18)

2-(4-(cyclohexyloxy)benzoyl)-*N*-hydroxybenzamide (2.5CYH_N): IR (cm^{-1}) (KBr): 3066.82 (s, OH hydroxamic acid), 1735.93 (S, C=O hydroxamic acid), 1714.55 (S, C=O keto). ^{1}H-NMR δ (ppm) (400Mhz, CDCl$_3$, TMS):1.45 (t,CH, 1H), 1.50-1.61 (m, CH$_2$, 4H), 1.78-1.99 (m, CH$_2$, 4H), 2.54-2.55 (m, CH$_2$, 2H), 7.09 (d, ArH, 2H), 7.51 (d, ArH, 2H), 7.61-7.63 (m, ArH, 1H), 7.97-8.00 (m, ArH, 2H), 8.35-8.37 (m, ArH, 1H)

***N*-hydroxy-2-{4-[(3-methoxybenzyl)oxy]benzoyl}benzamide (2.5MBA_N):** IR (cm^{-1}) (KBr): 2926.16 (s, OH hydroxamic acid), 1769.15 (s, C=O hydroxamic acid), 1736.17 (s, C=O keto) ^{1}H-NMR δ (ppm) (400Mhz, CDCl$_3$, TMS): 3.80 (s, CH$_3$, 3H), 5.18 (s, CH$_2$, 2H), 6.90 (d, ArH, 1H), 7.15 (d, ArH, 1H), 7.21 (d, ArH, 2H), 7.23 (t, ArH, 1H), 7.04 (s, ArH, 1H), 7.56 (d, ArH, 2H), 7.60 (d, ArH, 1H), 7.98-8.00 (m, ArH, 2H), 8.35 (d, ArH, 1H)

2-(4-(hexyloxy)benzoyl)-*N*-hydroxybenzamide (2.5HEX__N): IR (cm^{-1}) (KBr): 3048.78 (s, OH hydroxamic acid), 1734.01 (s, C=O hydroxamic acid), 1716.65 (S, C=O keto). ^{1}H-NMR δ (ppm) (400Mhz, CDCl$_3$, TMS): 0.92 (t, CH$_3$, 3H), 1.341.37 (m, CH$_2$, 4H), 1.81 (p, CH$_2$, 2H), 2.54 (p, CH$_2$, 2H), 4.06 (t, CH$_2$, 2H), 7.10 (d, ArH, 2H), 7.53 (d, ArH, 2H), 7.59-7.61 (m, ArH, 1H), 7.97-8.00 (m, ArH, 2H), 8.35-8.37 (m, ArH, 1H)

***N*-hydroxy-2-(4-(2-methoxyethoxy)benzoyl)benzamide (2.5ME__N):** IR (cm^{-1}) (KBr): 3038.55 (s, OH hydroxamic acid), 1731.93 (S, C=O hydroxamic acid), 1714.06 (S, C=O keto). ^{1}H-NMR δ (ppm) (400Mhz, CDCl$_3$, TMS): 3.40 (s, CH$_3$, 3H), 3.76 (t, CH$_2$, 2H), 4.21 (t, CH$_2$, 2H), 7.14 (d, ArH, 2H), 7.54 (d, ArH, 2H), 7.58-7.61 (m, ArH, 1H), 7.97-8.00 (m, ArH, 2H), 8.35-8.37 (m, ArH, 1H)

2.4.2.2 Synthesis of Urea derivatives

Carboxylic acid moiety of antranilic acid was protected by esterification using methanol and sulfuric acid as catalyst. Disappearance of broad OH Stretching peak of carboxylic acid and presence of sharp $-NH_2$ stretch of amine at 3483.44 and 3369.64 cm^{-1}, carbonyl ester stretch at 1687.71 cm^{-1} confirmed the compound 3.AAE. Free amine was then reacted with various isocynates. Amine nitrogen acts as nucleophile and attacks on carbonyl carbon of isocynates followed by internal proton transfer from amine to anion nitrogen of isocynate to form urea. Formed urea derivatives get precipitated from reaction mixture and on washing with ether yield practically pure product 3.__E. Yields of urea esters are presented in **Table 2.4-4.** Presence of ester carbonyl stretching peak at around 1705-1737 and amide carbonyl stretching peak at around 1593-1697 cm^{-1} in IR spectroscopy confirmed the compound formation.

Table 2.4-4: Isolated yield of urea esters

Isocynate used	Code of product	% yield
4-ChloroPhenyl isocynate	3.4CLPE	81.10
2-Methoxy Phenyl isocynate	3.2MXPE	0.00
3-Methoxy Phenyl isocynate	3.3MXPE	56.51
4-Methoxy Phenyl isocynate	3.4MXPE	65.08
4-NitroPhenyl isocynate	3.4NTPE	61.99
3-NitroPhenyl isocynate	3.3NTPE	38.93
4-FluoroPhenyl isocynate	3.4FLPE	49.93
2-FluoroPhenyl isocynate	3.2FLPE	60.97
3-FluoroPhenyl isocynate	3.3FLPE	58.34
2-TolylPhenyl isocynate	3.2TOPE	9.06
4-TolylPhenyl isocynate	3.4TOPE	46.90
3-TolylPhenyl isocynate	3.3TOPE	67.68

Table 2.4-5: Isolated yield of urea hydroxamates

Code of product	% yield	Code of product	% yield
3.4CLPN	65.42	3.4FLPN	76.06
3.2MXPN	0	3.2FLPN	62.45
3.3MXPN	79.66	3.3FLPN	82.97
3.4MXPN	43.15	3.2TOPN	31.55
3.4NTPN	79.05	3.4TOPN	91.13
3.3NTPN	72.72	3.3TOPN	80.62

Ester compounds 3.__E were then converted to hydroxamic acid by coupling with hydroxylamine. Amine of Hydroxylamine in presence of mild base acts as nucleophile and attacks on carbonyl carbon of ester to form amide bond. Yields are listed in **Table 2.4-5** and the formation was confirmed by disappearance of singlet peak of 3 protons of methyl ester

at around 3.53 δ (ppm) in ^{1}H NMR and by observing formation of orange to red colored ferric hydroxamate complex on treatment with ferric chloride.

Spectral Information of intermediates and products

Methyl 2-aminobenzoate (3.AAE): IR (cm^{-1}) (KBr): 3483.44 and 3369.64 (s, NH2)), 1687.71 (s, C=O ester); ^{1}H-NMR δ (ppm) (400Mhz CDCl$_3$, DMSO TMS): 3.82 (s, -OCH$_3$, 3H), 3.85 (-COOCH$_3$, 3H), 6.90(d, Ar, 2H), 7.01 (t, Ar, 1H), 7.33 (d, ArH, 2 H), 7.51 (t, ArH, 1H), 7.98 (d, ArH, 1H), 8.54, ArH, 1H), 10.40 (s, -NH, 1H)

Methyl 2-{[(4-chlorophenyl)carbamoyl]amino}benzoate (3.4CLPE): IR (cm^{-1}) (KBr): 3307.92, 3290.56 (s, NH), 1714.72 (s, C=O ester), 1651.12 (S, C=O amide)

Methyl 2-{[(3-methoxyphenyl)carbamoyl]amino}benzoate (3.3MXPE): IR (cm^{-1}) (KBr): 3290.56 (s, NH)), 1714.72 (s, C=O ester), 1651.07 (S, C=O amide)

Methyl 2-{[(4-methoxyphenyl)carbamoyl]amino}benzoate (3.4MXPE): IR (cm^{-1}) (KBr): 3292.49 (s, NH)), 1707.00 (s, C=O ester), 1658.78 (S, C=O amide)

Methyl 2-{[(4-nitrophenyl)carbamoyl]amino}benzoate (3.4NTPE): IR (cm^{-1}) (KBr): 3354.21, 3317.56 (s, NH)), 1730.15 (s, C=O ester), 1697.36 (S, C=O amide)

Methyl 2-{[(4-fluorophenyl)carbamoyl]amino}benzoate (3.4FLPE): IR (cm^{-1}) (KBr): 3307.92, 3278.99 (s, NH)), 1712.79 (s, C=O ester), 1660.71 (s, C=O amide)

Methyl 2-{[(2-methylphenyl)carbamoyl]amino}benzoate (3.2TOPE): IR (cm^{-1}) (KBr): 3153.61, 3109.25 (s, NH)), 1693.50 (s, C=O ester), 1593.20 (s, C=O amide)

Methyl 2-{[(4-methylphenyl)carbamoyl]amino}benzoate (3.4TOPE): IR (cm^{-1}) (KBr): 3309.85 (s, NH)), 1708.93 (s, C=O ester), 1658.78 (s, C=O amide)

Methyl 2-{[(3-methylphenyl)carbamoyl]amino}benzoate (3.3TOPE): IR (cm^{-1}) (KBr): 3319.49, 3284.77 (s, NH)), 1705.07 (s, C=O ester), 1605.67 (s, C=O amide)

N-hydroxy-2-{[(4-chlorophenyl)carbamoyl]amino}benzamide (3.4CLPN): MP > 250°C; IR (cm^{-1}) (KBr): 3066.82 (s, OH), 3196.05, (s, NH), 1732.08 (s, C=O hydroxamic acid), 1651.07 (s, C=O amide); ^{1}H-NMR δ (ppm) (400Mhz CDCl$_3$, DMSO TMS): 7.18-7.22(t, Ar, 1H), 7.27-7.25 (d, Ar, 1H), 7.31-7.29 (d, Ar, 2H), 7.48-7.50 (d, Ar, 2H), 7.65-7.61 (t, Ar, 1H), 8.00-7.97 (d, Ar, 1H), 8.01 (s, NH, 1H), 11.60 (s, NH, 1H).

N-hydroxy-2-{[(3-methoxyphenyl)carbamoyl]amino}benzamide (3.3MXPN): MP > 250°C; IR (cm^{-1}) (KBr): 3226.91 (s, OH), 3196.05, 3140.11 (s, NH), 1732.08 (s, C=O hydroxamic acid),

1651.07 (s, C=O amide) ; ^{1}H-NMR δ (ppm) (400Mhz CDCl$_3$, DMSO TMS): 3.79 (s, OCH$_3$, 3H), 6.85-6.87 (d, Ar, 1H), 6.89 (t, Ar, 1H), 6.98 (d, Ar, 1H), 7.20 (t, Ar, 1H), 7.24 (d, Ar, 1H), 7.38 (t, Ar, 1H), 7.65 (t, Ar, 1H), 7.95(d, Ar, 1H), 11.55 (s, NH, 1H).

N-hydroxy-2-{[(4-methoxyphenyl)carbamoyl]amino}benzamide (3.4MXPN): MP > 250°C; IR (cm^{-1}) (KBr): 3066.82 (s, OH), 3194.12, 3126.61 (s, NH), 1722.43 (s, C=O hydroxamic acid), 1660.71 (s, C=O amide) ; ^{1}H-NMR δ (ppm) (400Mhz CDCl$_3$, DMSO TMS): 3.83 (s, OCH$_3$, 3H), 7.00 (d, Ar, 2H), 7.16-7.20 (m, Ar, 3H), 7.24 (d, Ar, 1H), 7.62 (t, Ar, 1H), 7.96 (d, Ar, 1H), 11.52 (s, NH, 1H).

N-hydroxy-2-{[(4-nitrophenyl)carbamoyl]amino}benzamide (3.4NTPN): MP > 250°C; IR (cm^{-1}) (KBr): 3078.39 (s, OH), 3196.05, 3118.90 (s, NH), 1728.22 (s, C=O hydroxamic acid), 1681.98 (s, C=O amide) ; ^{1}H-NMR δ (ppm) (400Mhz CDCl$_3$, DMSO TMS): 7.22 (t, Ar, 1H), 7.27 (d, Ar, 1H), 7.62-7.69 (m, Ar, 3), 7.98 (d, Ar, 1H), 8.35 (d, Ar, 2H), 11.70 (s, NH, 1H).

N-hydroxy-2-{[(3-nitrophenyl)carbamoyl]amino}benzamide (3.3NTPN): MP > 250°C; IR (cm^{-1}) (KBr): 3061.03 (s, OH), 3196.05, 3126.61 (s, NH), 1732.08 (s, C=O hydroxamic acid), 1672.28 (s, C=O amide) ; ^{1}H-NMR δ (ppm) (400Mhz CDCl$_3$, DMSO TMS): 7.22 (t, Ar, 1H), 7.27 (d, Ar, 1H), 7.67 (t, Ar, 1H), 7.76-7.83 (m, Ar, 2H), 7.99 (d, Ar, 1H), 8.29-8.35 (m, Ar, 2H), 11.69 (s, NH, 1H).

N-hydroxy-2-{[(4-fluorophenyl)carbamoyl]amino}benzamide (3.4FLPN): MP > 250°C; IR (cm^{-1}) (KBr): 3246.20 (s, OH), 3205.69, 3145.90 (s, NH), 1732.08 (s, C=O hydroxamic acid), 1651.07 (s, C=O amide) ; ^{1}H-NMR δ (ppm) (400Mhz CDCl$_3$, DMSO TMS): 7.17-7.33(m, Ar, 6H), 7.63 (t, Ar, 1H), 7.978 (d, Ar, 1H), 11.58 (s, NH, 1H).

N-hydroxy-2-{[(2-fluorophenyl)carbamoyl]amino}benzamide (3.2FLPN): MP 248-250°C; IR (cm^{-1}) (KBr): 3286.7 (s, OH), 3232.70 (s, NH), 1730.15 (s, C=O hydroxamic acid), 1658.78 (s, C=O amide) ; ^{1}H-NMR δ (ppm) (400Mhz CDCl$_3$, DMSO TMS): 7.20-7.39(m, Ar, 5H), 7.45-7.51 (m, Ar, 1H), 7.65 (t, Ar, 1H), 8.01(d, Ar, 1H), 11.64 (s, NH, 1H).

N-hydroxy-2-{[(3-fluorophenyl)carbamoyl]amino}benzamide (3.3FLPN): MP > 250°C; IR (cm^{-1}) (KBr): 3238.48 (s, OH), 3199.91, 3142.04 (s, NH), 1737.86 (s, C=O hydroxamic acid), 1666.50 (s, C=O amide) ; ^{1}H-NMR δ (ppm) (400Mhz CDCl$_3$, DMSO TMS): 7.16-7.26 (m, Ar, 5H), 7.48-7.50 (m, Ar, 1H), 7.66 (t, Ar, 1H), 7.95 (d, Ar, 1H), 11.61 (s, NH, 1H).

3.3FLPN
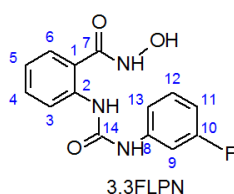
^{13}C NMR δ (ppm) (100Mhz, DMSO and CDCl$_3$): 162.05 (d, *J*= 243 Hz, C10), 161.99 (C7), 149.94 (C14), 139.73 (C2), 137.08 (d, *J*= 10.5 Hz C8), 134.99 (C4), 130.00 (d, *J*= 8.9 Hz C12), 127.47 C6, 125.21 (C13), 122.36

C5, 116.45 (d, *J*= 23.1 Hz C11), 115.24 C1, 115.02 (d, *J*= 20.8 Hz C9), 114.17 (C3)

N-hydroxy-2-{[(2-methylphenyl)carbamoyl]amino}benzamide (3.2TOPN): MP 238-240°C; IR (cm^{-1}) (KBr): 3278.99 (s, OH), 3240.41, 3215.53 (s, NH), 1681.93 (s, C=O hydroxamic acid), 1660.71 (s, C=O amide) ; ^{1}H-NMR δ (ppm) (400Mhz CDCl$_3$, DMSO TMS): 2.14(s, CH$_3$, 3H), 7.12 (d, Ar, 1H), 7.29-7.50 (m, Ar, 4H), 7.49 (d, Ar, 1H), 7.69 (t, Ar, 1H), 8.04 (d, Ar, 1H), 12.97 (s, NH, 1H),

N-hydroxy-2-{[(4-methylphenyl)carbamoyl]amino}benzamide (3.4TOPN): MP 244 °C; IR (cm^{-1}) (KBr): 3199.91 (s, OH), 3238.48, 3142.04 (s, NH), 1730.15 (s, C=O hydroxamic acid), 1666.50 (s, C=O amide) ; ^{1}H-NMR δ (ppm) (400Mhz CDCl$_3$, DMSO TMS): 2.47 (s, CH$_3$, 3H), 6.86 (d, Ar, 1H), 7.21-7.25 (m, Ar, 3H), 7.37 (d, Ar, 2H), 7.51 (t, Ar, 1H), 8.16 (d, Ar, 1H), 10.17 (s, NH, 1H) disappearance of 10.17 on D2O exchange.

N-hydroxy-2-{[(3-methylphenyl)carbamoyl]amino}benzamide (3.3TOPN): MP 220-222°C; IR (cm^{-1}) (KBr): 3242.38 (s, OH), 3201.38, 3142.04 (s, NH), 1730.15 (s, C=O hydroxamic acid), 1658.78 (s, C=O amide) ; ^{1}H-NMR δ (ppm) (400Mhz CDCl$_3$, DMSO TMS): 2.38 (s, CH$_3$, 3H), 7.06-7.25 (m, Ar, 5H), 7.37 (t, Ar, 1H), 7.63 (t, Ar, 1H), 7.96 (d, Ar, 1H), 11.55 (s, NH, 1H).

3.3TOPN
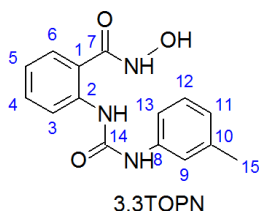
^{13}C NMR δ (ppm) (100Mhz, DMSO and CDCl$_3$): 162.13 (C7), 150.19 (C14), 139.68 (C2), 138.20 (C8), 135.35 (C10), 134.77 (C4), 129.16 (C1), 128.65 (C12), 128.48 (C6), 127.45 (C11), 125.71 (C5), 122.23 (C3), 115.17 (C9), 114.18 (C13), 20.81 (C15).

2.4.3   References

[1] Wyss, D.F.; Arasappan, A.; Senior, M.M.; Wang, Y.-S.; Beyer, B.M.; Njoroge, F.G.; McCoy, M.A. Non-peptidic small-molecule inhibitors of the single-chain hepatitis C virus NS3 protease/NS4A cofactor complex discovered by structure-based NMR screening. *J Med Chem.* **2004,** 47(10), 2486-2498.

[2] Wagh, R.; Mahajan, H.S.; Kaskhedikar, S.G. Synthesis and Antioxidant Activity of Substituted Phenylalkene Hydroxamic Acids. *Asian J Chem.* **2007,** 19(6), 4188-4192.

[3] Tale, R.H.; Rodge, A.H.; Hatnapure, G.D.; Keche, A.P. The novel 3, 4-dihydropyrimidin-2 (1-*H*)-one urea derivatives of N-aryl urea: Synthesis, anti-inflammatory, antibacterial and antifungal activity evaluation. *Bioorg Med Chem Lett.* **2011,** 21(15), 4648-4651.

[4] Cherney, R.J.; King, B.W.; Gilmore, J.L.; Liu, R.-Q.; Covington, M.B.; Duan, J.J.-W.; Decicco, C.P. Conversion of potent MMP inhibitors into selective TACE inhibitors. *Bioorg Med Chem Lett.* **2006,** 16(4), 1028-1031.

2.5 In Vitro Screening of TACE Inhibitors

TACE is a metalloprotease, involved in cleavage of peptide bond of membrane bound pro-proteins to release soluble ectodomain of the same. TACE is an important target in RA due to its role in cleavage of pro TNF-α to release soluble TNF-α. TACE inhibitor assay was performed according to the procedure given in BioVision's TACE inhibitor screening Kit [1]. In vitro activity of TACE can be measured in terms of fluorescent intensity which is directly proportional to amount of FRET (Fluorescence Resonance Energy Transfer) substrate hydrolyzed by TACE. Free enzyme cleaves FRET substrate and releases Fluorescent substance. The amount of fluorescent product released is measured by fluorimetry at Ex/Em = 318/449 nm.

2.5.1 Methodology

TACE enzyme was dissolved in 220 µl assay buffer and stored at -20°C. FRET substrate was diluted in 2:23 proportion with assay buffer. A10 mM primary stock of inhibitors was prepared in DMSO from which 1 mM and 400 µM secondary and tertiary stock was prepared respectively. Initially all the reagents except enzyme solution were equilibrated to room temperature. Amount of each reagent added to wells of black 96 well plate is given in **Table 2.5-1**. After addition of inhibitor, reaction mixture was incubated at 37°C for 5 min followed by addition of 25 µl of substrate solution. Plate was covered and incubated for 30 min at 37°C. Plate cover was removed and fluorescence was monitored using excitation wavelength of 318 nm and emission wavelength of 449 nm. Percentage inhibition of the TACE was then calculated by **Equation 2.5-1**.

Table 2.5-1: Well content in in vitro TACE activity screening assay

well	DMSO with Buffer	Inhibitor	TACE	Solvent	Substrate
100% activity (PC)	25	-	50	-	25
Background (NC)	25	-	-	50	25
Inhibitor well	-	25	50	-	25

PC. Positive control, NC: negative control; all values are expressed in µl

Equation 2.5-1:
$$\% \text{ Inhibition} = \frac{\Delta \text{ RFU of PC} - \Delta \text{ RFU of EI}}{\Delta \text{RFU of PC}} \times 100$$

Where RFU= Relative Fluorescence Units, PC = positive control or enzyme control, EI = enzyme inhibitor solution.

Percentage inhibition generally indicates binding ability of the inhibitors with the enzyme. But by increasing size of the molecule using mainly hydrophobic groups, binding tends to increase. In order to evaluate actual performance increase in the molecule normalization of activity data is needed. Potency Efficacy Index (PEI) is a normalized scale based on molecular weight and helps in relative comparison of compounds tested under similar condition and at same concentration only [2].

Equation 2.5-2: $$PEI = \frac{\% \text{ inhibition at a given concentration as fraction between 0-1.0}}{\text{Molecular Weight (Da)}}$$

2.5.2   Results and Discussion
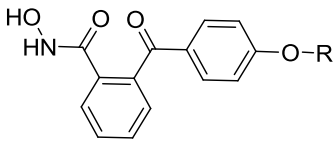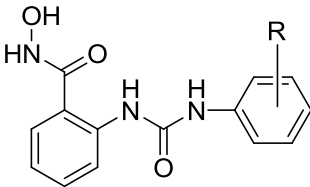
2.5.2.1   In vitro assay

Synthesized compounds were screened for in vitro h-TACE inhibition potential at 100 µM concentration using FRET assay method and Infinite M1000 PRO (Tecan) microplate reader. Percentage inhibitory potential and PEI of each compound is given in **Table 2.5-2**. Six urea analogues showed more than 50% inhibition against TACE while benzophenone analogues showed mild to moderate activity.

Benzophenone series: Compared to urea series Benzophenone derivatives were having lower inhibitory potential against TACE. Most active compounds in the series were 2.5ANS_N and 2.5ETHA_N with methyl and ethyl substitution at *para* hydroxy of phenyl ring oriented at S1 cavity of TACE with 42.31 and 42.73% inhibition respectively at 100µM concentration. Further increasing linear chain length with propyl (2.5PRO_N) or pentyl (2.5PENT_N) lead to nearly 56% and 52% reduction in the activity compared to 2.5ETHA_N. Replacing pentyl with a 2-(methoxy)ethyl group increased inhibitory activity from 20.37 to 29.09%. The n-hexyl substitution (2.5HEX_N) significantly increased the potency to 39.77% inhibition while cyclohexyl substitution (2.5CYH_N) reduced the activity to 28.8%. Replacing linear chain with open branched substitution did not significantly affect the activity. 2-(methyl)-propyl substitution in 2.5MP_N showed 21.87% inhibition while t-butyl substitution as in 2.5tBA_N showed 19.03% inhibition. Introduction of unsaturation in alkyl chain increased the potency slightly by increasing the rigidity of the chain. The allyl substitution (2.5ALLYL_N) raised the potency to 25.85% from 18.74% of 2.5PRO_N, while most rigid linear aryl group 2-buten-1yl (2.52B_N) showed 37.97% inhibition. Introducing partial flexibility with 3-buten-1yl (2.53B_N) reduced the activity to 26.73%. 3-methoxy benzyl substitution (2.5MBA_N) did not show significant improvement compared to 2.53B_N.

Table 2.5-2: In vitro TACE Inhibitory potential of synthesized TACE inhibitors

Benzophenone series				Urea series			
code	R	% inhib (100µM)	PEI X 10⁻³	code	R	% inhib (100µM)	PEI X10⁻³
2.5ME_N	2-(Methoxy) ethyl-	29.09	0.923	3.4CLPN	*p*-Chloro-	17.5	0.572
2.5PRO_N	n-Propyl-	18.74	0.626	3.2MXPN	*o*-Methoxy-	NT	
2.5PENT_N	n-Penyl-	20.37	0.622	3.3MXPN	*m*-Methoxy-	68.37	2.269
2.5ALLY_N	Allyl-	25.85	0.869	3.4MXPN	*p*- Methoxy-	59.29	1.968
2.5HEXA_N	n-Hexyl-	39.77	1.165	3.4NTPN	*p*-Nitro-	78.63	2.486
2.52B_N	2-Butyn-1-yl-	37.97	1.228	3.3NTPN	*m*-Nitro-	75.62	2.391
2.5ETHA_N	Ethyl-	42.73	1.498	3.4FLPN	*p*-Fluoro-	73.99	2.558
2.5MP_N	2-(methyl) propyl-	21.87	0.698	3.2FLPN	*o*-Fluoro-	5.82	0.201
2.5tBA_N	*ter*-Butyl-	19.03	0.607	3.3FLPN	*m*-Fluoro-	18.23	0.63
2.5CYH_N	Cyclo-Hexyl-	28.8	0.849	3.2TOPN	*o*-Methyl-	73.85	2.589
2.53B_N	3-Butyn-1-yl-	26.73	0.864	3.4TOPN	*p*-Methyl-	14.27	0.5
2.5MBA_N	m-(methoxy)benzyl-	29.83	0.79	3.3TOPN	*m*-Methyl-	24.8	0.869
2.5ANS_N	Methyl -	42.31	1.56				

% inhib: percentage inhibition; PEI: potency efficacy index; NT: not tested.

Urea series: Nitro substition on *para* position of phenyl ring oriented at S1 cavity (3.4NTPN) was the most potent compound with 78.63% inhibition at 100µM concentration followed by *meta* nitro substituted compound 3.3NTPN with 75.62% inhibition. *para* substitution with fluorine group slightly reduced the activity to 73.99% while *meta* fluoro substitution (3.3FLPN) reduced the activity drastically to 18.23% while *ortho* fluoro substitution (3.2FLPN) rendered the compound almost inactive with 5.82% inhibition. Changing the *para* Fluoro to Chloro (3.4CLPN) also reduced activity to 17.50%. Changing electron withdrawing group from *para* position to electron releasing group significantly reduced the activity. *para* methoxy substitution (3.4MXPN) showed 59.29% inhibition while *meta* methoxy substitution (3.3MXPN) raised the potency to 68.37%. Similar trend was observed with methyl substitution. *para* methyl substitution (3.4TOPN) showed weak activity (14.27% inhibition) while activity was raised significantly to 24.80% with *meta* methyl substitution (3.3TOPN)

and *ortho* methoxy substitution lead to potent compound 3.2TOPN with 73.85% TACE inhibition.

2.5.2.2   In silico in vitro screening correlation

Since no correlation between docking score and inhibitory potency represented as percentage inhibition was established, effect of solubility in biochemical environment was investigated theoretically. Predicted solubility of benzophenone series and Urea series using AlogPS algorithm are presented in **Table 2.5-3** and **Table 2.5-4**. Benzophenone derivatives are relatively less soluble than Urea derivatives. Many of the benzophenone derivatives showed predicted solubility below 15mg/Lts. While urea derivatives were predicted to have more than 28mg/Lts aqueous solubility. When docking scores and biological activity data were evaluated, no correlation was established. High dock scoring Benzophenone series showed lesser biological activity compared to Urea analogues, which may be due to their low solubility. One of the main parameter influencing biochemical assay is the solubility. Though low soluble leads may have acceptable oral bioavailability and higher potency but in biochemical assay it may lead to false negatives by underestimating potencies. Low solubility of compound in Drug Discovery Bioassay can lead to [3]:

- Low HTS hit rates

- Precipitation from dimethyl sulfoxide (DMSO) stocks

- Activity appears (falsely) lower or higher than the correct value

- Discrepancies between enzyme and cell assays

- Erratic assay results

As the compounds were tested at 100 µM concentration it is possible that compounds may have been precipitated while diluting with assay buffer. As this fact was not considered and to compensate the low solubility mediated false negative results, Solubility Efficacy Index (SEI) term was used. SEI helps in identifying potent compound among the analogues tested under identical condition at same but high concentration of the compound in DMSO. SEI of the screened compounds was calculated as per the **Equation 2.5-3** and tabulated in **Table 2.5-3** and **Table 2.5-4**.

Equation 2.5-3 $$\text{SEI} = \frac{\%\text{ inhibition at a given concentration as fraction between 0-1.0}}{\text{Aqueous Solubility (g/Lts)}}$$

Where aqueous solubility was predicted using AlogPS [4] program.

Table 2.5-3: Solubility Efficacy Index of Benzophenone series

code	Autodock Score	Glide SP Score	Glide XP Score	FRED chemscore3	% inhibition	Predicted solubility	SEI
	Kcal/mol				(100 µM)	g/Lts	
2.5ME_N	-9.1	-7.68	-7.58	-121.06	29.09	0.018	16.16
2.5PRO_N	-9.19	-6.55	-7.72	-117.78	18.74	0.022	8.52
2.5PENT_N	-10.06	-8.64	-8.42	-126.51	20.37	0.005	40.74
2.5ALLY_N	-9.53	-6.23	-7.63	-117.55	25.85	0.043	6.01
2.5HEXA_N	-10.42	-7.95	-8.86	-127.25	39.77	0.003	132.57
2.52B_N	-9.68	-8.36	-7.52	-118.2	37.97	0.011	34.52
2.5ETHA_N	-9.47	-7.56	-7.17	-112.02	42.73	0.031	13.78
2.5MP_N	-9.29	-7.73	-7.76	-118.8	21.87	0.01	21.87
2.5tBA_N	-9.52	-8.63	-8.33	-115.77	19.03	0.012	15.86
2.5CYH_N	-11.29	-9.05	-8.68	-126.03	28.8	0.004	72
2.53B_N	-9.87	-7.83	-8.19	-119.88	26.73	0.013	20.56
2.5MBA_N	-11.15	-9.87	-8.94	-134.15	29.83	0.002	149.15
2.5ANS_N	-7.67	-7.61	-6.74	-106.83	42.31	0.026	16.27

Table 2.5-4: Solubility Efficacy Index of Urea series

code	Autodock Score	Glide SP Score	Glide XP Score	FRED chemscore3	% inhibition	Predicted solubility	SEI
	Kcal/mol				(100 µM)	g/Lts	
3.2MXPN	-8.06	-7.6	-5.71	-94.06	NT	0.066	NT
3.3MXPN	-9.35	-7.6	-5.54	-110.58	68.37	0.065	10.52
3.4MXPN	-8.82	-7.08	-5.21	-106.99	59.29	0.064	9.26
3.4NTPN	-10.22	-7.06	-6.8	-100.16	78.63	0.028	28.08
3.3NTPN	-12.17	-7.46	-6.71	-104.7	75.62	0.031	24.39
3.4FLPN	-8.86	-6.64	-6.68	-106.2	73.99	0.076	9.74
3.2FLPN	-8.71	-7.31	-6.61	-100.5	5.82	0.061	0.95
3.3FLPN	-8.84	-6.33	-6.72	-105.64	18.23	0.071	2.57
3.2TOPN	-8.81	-6	-6.33	-98.66	73.85	0.052	14.2
3.4TOPN	-9.04	-6.55	-7.06	-101.8	14.27	0.05	2.85
3.3TOPN	-9.44	-6.2	-7.11	-111.43	24.8	0.051	4.86

NT: not tested

Since all the derivatives lack ionisable groups therefore change in solubility with buffer may be considered similar in all the compounds. Thus, it is assumed that predicted aqueous solubility correlates well with solubility in assay buffer. Higher SEI does not indicate the compounds are more potent but directs the medicinal chemist to check those compounds at

lower concentration in order to avoid possible false negatives. With SEI the correlation between docking score and activity was established. Compounds with SEI greater than 25 showed higher docking scores calculated from Autodock, Glide SP, Glide XP and FRED. Benzophenone series showed higher SEI than urea series and the same is reflected in docking scores, while percentage inhibition of urea series was greater which can be attributed to their greater solubility compared to Benzophenone series. Since all the compounds were tested at 100μM concentration, SEI is more meaningful in detecting false negatives. Some of the compounds with possible false negative result or underestimated potency are 2.5PENT_N, 2.5HEX_N, 2.5B_N, 2.5CYH_N and 2.5MBA_N. From the SEI a good correlation with docking scores was established

2.5.3   References

[1]   http://www.biovision.com/tace-inhibitor-screening-assay-kit-fluorometric-5587.html (Accessed on Aug 2012).

[2]   Abad-Zapatero, C.; Metz, J.T. Ligand efficiency indices as guideposts for drug discovery. *Drug Discov Today.* **2005,** 10(7), 464-469.

[3]   Di, L.; Kerns, E.H. Biological assay challenges from compound solubility: strategies for bioassay optimization. *Drug Discov Today.* **2006,** 11(9), 446-451.

[4]   Tetko, I.V.; Tanchuk, V.Y. Application of associative neural networks for prediction of lipophilicity in ALOGPS 2.1 program. *J Chem Inf Comput Sci.* **2002,** 42(5), 1136-1145.

Chapter 3

Design, Synthesis and Evaluation of PAD4

Inhibitors

3.1 Introduction to PAD4

Citrullination of proteins was first reported by Roger and Simmonds in 1958 [1]. However, its physiological importance was neglected till late ninety's. Later with identification of auto antibodies against citrullinated proteins in RA patients its physiological role was investigated. Initially it was found that citrullination of peptidyl arginine leads to loss of positive charges and thereby alters protein-protein interaction, destabilizes the protein and enhances its susceptibility for degradation by proteases [2]. Recently it has been identified that such post-translational modification of proteins can be a cause for numerous diseases like RA, Cancer, Alzheimer's disease and Multiple Sclerosis [3-5].

3.1.1 Physiological Role of PAD

Citrullination is an enzymatic conversion of peptidyl arginine to peptidyl citrulline by Peptidyl Arginine Deiminase (PAD) resulting in loss of molecular weight by less than 1 Da and a positive charge (**Figure 3.1-1**) [6].
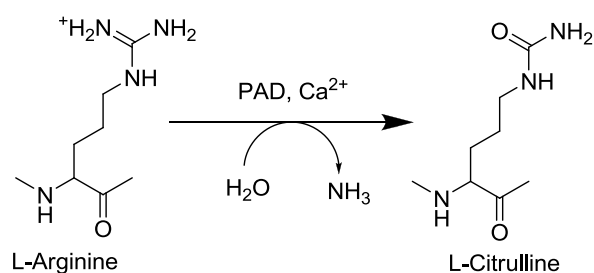
Figure 3.1-1: Catalytic reaction of PAD

Five isoforms of PAD enzyme have been identified namely PAD1, PAD2, PAD3, PAD4 and PAD6. PAD1 is expressed mainly in epidermis, hair follicles, sweat glands and uterus. It citrullinates cytokeratin, which modulates cornification of epidermis. PAD2 is expressed in brain, skeletal muscles, uterus, and pancreas and citrullinates mainly myelin and also plays a key role in neurodegenerative disease. PAD3 is expressed in epidermis and controls the citrullination of trichohyaline in fair follicles. PAD4 is expressed in hematopoietic tissue (bone marrow and spleen) and in neutrophils, eosinophils and monocytes [7] and is the only type present in nucleus of these cells. PAD6 is expressed specifically in embryonic stem cells, oocytes, peripheral blood leukocytes, and testis. It is mainly involved in folliculogenesis and spermatogenesis and is also responsible for development of zygote [4, 8, 9].

Under normal physiological conditions, citrullination

a. destabilizes the protein and makes it more susceptible for enzymatic hydrolysis (**Figure 3.1-2**) [10].

b. alters protein-protein interactions e.g. citrullination of Inhibitor of growth 4 (ING4), a tumor suppressor protein, causes loss of affinity for p53 [11].

c. induces depolymerisation as observed in vimentin filament leading to disruption of cytoskeletal network [6].

d. renders some protein as inactive e.g. citrullination of antithrombin suppresses its ability to inhibit thrombin leading to increased thrombin activity [4].

e. controls gene expression by blocking methylation of Histone arginine involved in gene expression [12].
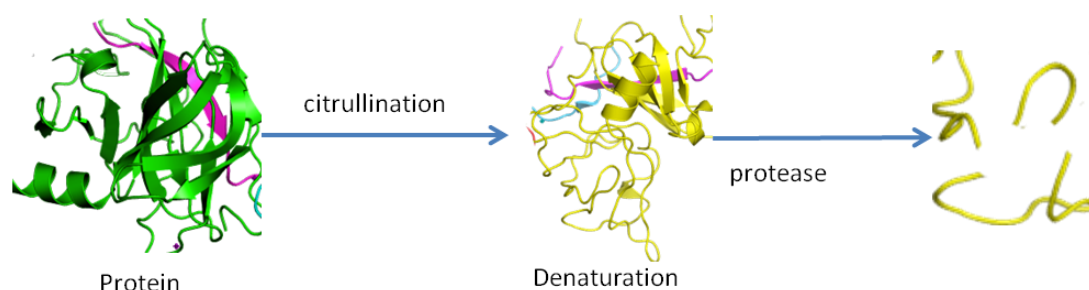
Figure 3.1-2: Protein hydrolysis aided by citrullination

3.1.2 PAD4 and RA

Citrullinated proteins act as autoantigens in RA [13]. Compared to RA synovium, PAD4 expression in Osteoarthritis (OA) and normal synovium is far less with undetectable levels of citrulinated proteins [14]. RA patients (45.0%) showed higher levels of Anti Citrullinated Peptide Antibodies (ACPAs) as compared to OA (5.9%) and systemic lupus erythematosus (9.0%) patients. There is high correlation between erythrocyte sedimentation rate, IgM, and IgG with ACPAs [15] and levels of ACPAs strongly correlate with bone destruction [16]. Failure of ACPAs induction by citrullinated type II collagen in collagen induced arthritis mice model indicates that citrullinated proteins can only trigger RA in humans [17]. Genetically, production of ACPAs was found to be associated with HLA-DRB1 shared epitopes, a known risk factor for RA [18]. In a study of 268 RA patients in Netherlands, HLA-DRB1*0401, DRB1*1001, DQB1*0302, and DQB1*0501 genes were found to be associated with ACPA+ve RA. Also, significant increased joint damage rate were observed in ACPA+ve patients compared to ACPA-ve patients [19]. In another study using 110 Japanese RA patients, HLA-DRB1*0405 association was significantly high in ACPA+ve RA patients [20]. In Korean

population, DRB1*0901 was found to be associated with ACPA and smoking increased possibility of ACPA +ve RA by 36.11 folds [21]. Out of 17 identified Single-nucleotide polymorphisms (SNPs) in PAD4 gene located in 1p36 chromosome, eight have been found to have significant correlation with RA [22]. Also, some of these single point mutations make PAD4 mRNA more stable leading to enhanced production of PAD4 and citrullination in RA [23].

3.1.3   PAD4 Enzymology

Table 3.1-1: Biochemical properties of PAD4 [24, 25]**.**

Protein Name	Protein-arginine deiminase type-4 (PAD-IV or PAD4)
Enzyme Code	3.5.3.15 Hydrolase. Acting on carbon-nitrogen bonds, other than peptide bonds. In linear amidines. Protein-arginine deiminase
Gene Name	PADI4 (PADI5, PDI5)
Sequence length	663 Amino acids
Catalytic activity	Protein L-arginine + H_2O = protein L-citrulline + NH_3.
Cofactor	Calcium ions (Ca^{++})
Tissue specificity	hematopoietic cells like eosinophiles neutrophils, and macrophages
Post-translational modification	Autocitrullination at Arg372 and Arg374 leading to inactivation
Optimum pH	6.5-9.0
Gene Location	Homo sapiens chromosome 1p36.13
Uniplot ID	Q9UM07
Pfam annotation	PAD N-terminal domain (PAD_N), PAD middle domain (PAD_M), PAD
SCOP domain	Cupredoxins, Peptidylarginine deiminase Pa, Pentein

Table 3.1-2: List of X-ray crystal structures available on protein database [26]**.**

PDB ID	Resolution (Å)	Remarks
1WD8	2.8	Calcium free form of hPAD4
1WD9	2.6	Calcium bound form of hPAD4
1WDA	2.3	hPAD4 and benzoyl-L-arginine amide (BAA) complex
2DEW	2.1	complex with histone H3 N-terminal tail including Arg8
2DEX	2.1	complex with histone H3 N-terminal peptide including Arg17
2DEY	2.25	complex with histone H4 N-terminal tail including Arg3
2DW5	2.3	complex with N-alpha-benzoyl-N5-(2-fluoro-1-iminoethyl)-L-ornithine amide
3APM	2.5	human SNP PAD4
3APN	2.7	human wild-type PAD4
3B1T	2.5	complex with o-Cl-amidine
3B1U	2.1	complex with o-F-amidine
4DKT	2.98	hPAD4 and N-acetyl-L-threonyl-L-alpha-aspartyl-N5-[(1E)-2-fluoroethanimidoyl]-L-ornithinamide complex

3.1.4 Structure of PAD4

PAD4 is a 663 amino acid 74 kDa protein belonging to guanidino-group modifying enzyme superfamily (GMSF) where a basic catalytic motif Cys-His-Asp(Glu) is conserved [27]. PAD4 exists as a dimer and each monomer consists of two domains. Residues from Met1 to Pro300 form N terminal domain made up of two immunoglobulin like subdomains. Sub domain 1 consists of nine β strands (β1- β9) with nuclear localization signal ([56] PPAKKKST [6]) on surface. Sub domain 2 consists of 10 β strands (β10- β19), four α helices and three calcium ions. Residues from Asn301 to Pro663 form C terminal domain consisting of five ββαβ modules arranged circularly and called as α/β propeller and one β strand from each module contributes for active site formation (**Figure 3.1-3**) [28-29].
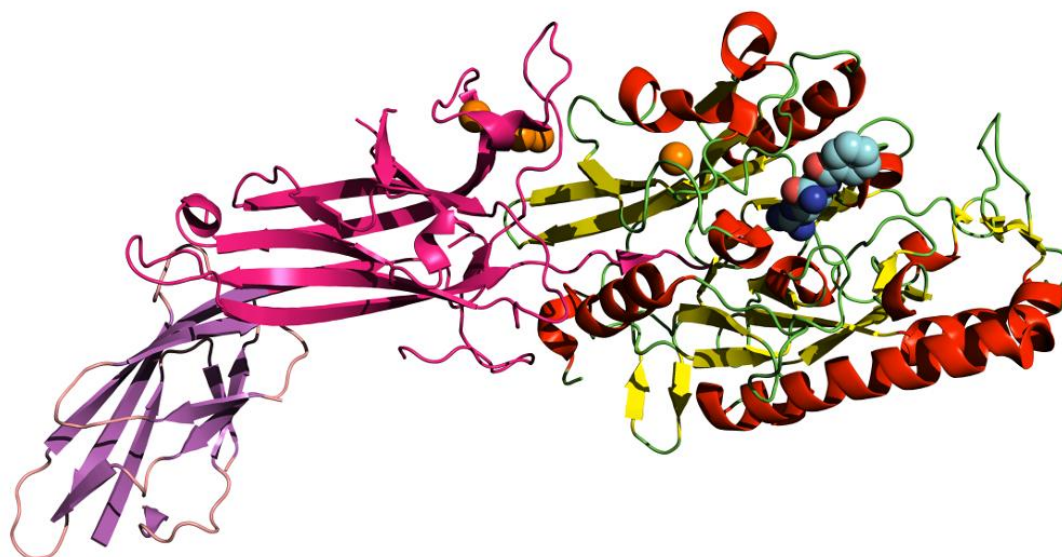
Figure 3.1-3: PAD4 structure (PDB 1WDA). Magenta and pink colored portions form N terminus domain and represent subdomain 1 and subdomain 2 respectively. Rest of the portion is C terminus domain. Ligand (BAA) bound to active site is represented as CPK model and 5 calcium ions are represented as orange spheres.

Catalytic site present at C-terminal domain consists of charged and polar residues like Asp350, His471, Asp473, His644, and Cys645. Two calcium ions near to the active site affect conformation and activity of the enzyme [27]. During catalysis, Asp350 and Asp473 act as base contributing for substrate binding while His471 acts as acid and base allowing proton exchange during the catalysis where as Cys645 acts as nucleophile and forms covalent bond with imine carbon of the substrate. According to mechanism of catalysis, proposed by Ke et al. [27] (**Figure 3.1-4**), Cys645 and His471 initially exist in neutral form. On substrate binding, guanidine group of substrate causes deprotonation and activation of Cys645. Activated Cys645 then makes nucleophilic attack on imine carbon of substrate forming transition state. Ammonia released from intermediate transition state gets exchanged with water. Initial

Cys645 deprotonation and release of ammonia are found to be the rate determining step with energy barrier of 20.9 Kcal/mol. His471 and Asp473 enhances the nucleophilicity of the water by hydrogen bonding. Activated water molecule then makes a nucleophilic attack on imine carbon followed by release of Cys645 with its bonding pair of electrons. The overall energy barrier for hydrolysis was found to be 16.5 Kcal/mol [27].
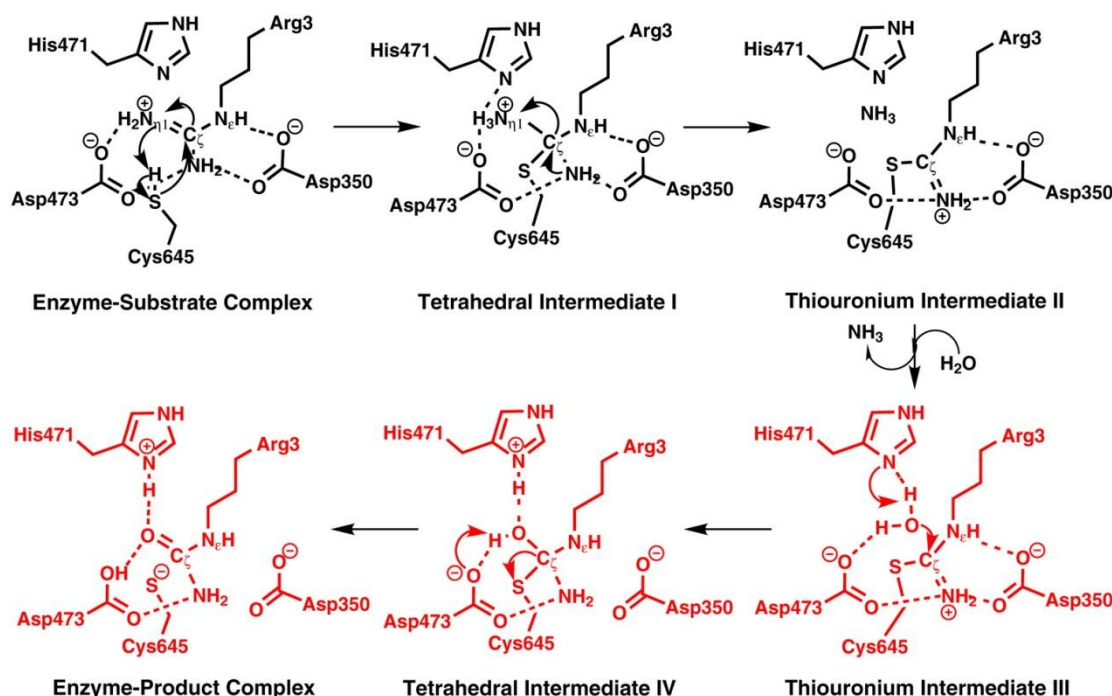
Figure 3.1-4: Mechanism of PAD4 catalysis as proposed by Ke et al. [27].

3.1.5   PAD4 Inhibitor development

Luo and coworkers in 2006 developed arginine mimetic irreversible PAD4 inhibitors [30]. A good leaving group (halogen) at position X in structure **301** aided in covalent bond formation of inhibitor with sulfur of Cys645. A chlorine at position 'X' was found to be more active (Cl-amidine, IC_{50} 5.9 µM) and selective for PAD4 compared to Fluorine moiety (F-amidine, IC_{50} 21.6 µM). Hydrogen substitution rendered the compound inactive (H-amidine, IC_{50} >1000 µM) as it does not form any covalent bond with enzyme. Strong receptor binding of these ligands was also due to associated hydrogen bonding of ligand with Arg374 and Arg639. A 100 fold increase in the activity was observed when chain length was increased from 2 to 3 carbon atoms and drastic decrease was observed when chain length was further increased.

Effect of substitution at position 2 and linker length was further explored by Causey et al. [31]. Compounds with mono-amide bond as observed in **302** (IC_{50} >1000 µM) were less active compared to diamide counterparts (F-amidine). Sulfonamide derivative of F-amidine

also caused reduction in activity by 22 times (compound **303** IC$_{50}$ 500 µM). Nature of carbonyl substituent at amine was found to have strong influence on the activity. Activity increased from 230 µM to 19 µM by changing imidazole(**304**) to pyrrole (**305**).

Bicker et al. [32] prepared D-isomers of compounds (**301**) prepared earlier [30]. D-isomers had increased selectivity towards PAD1 but PAD4 inhibitory activity was decreased drastically. They also studied toxicity of potent PAD4 inhibitor L-Cl-amidine on mice. Results showed serious adverse effects such as anxiety lasting for 1-4 hrs followed by diarrhea lasting for 2-3 days along with significant weight loss after 4 days of administration.
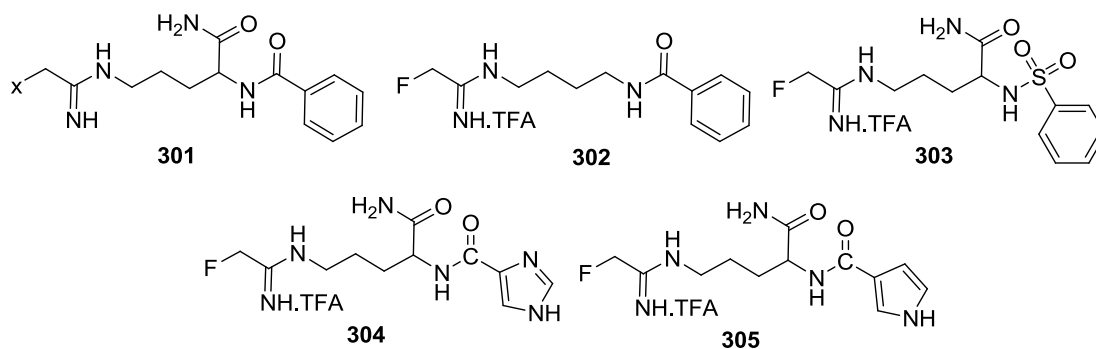
Jones et al. [33] prepared a library of 264 F-amidine analogues by substitution at terminal amino and carboxylic group of amino acid. Thr-Asp F-amidine (TDFA structure **306**) was more potent than F-amidine. Activity further increased when 2-mercaptoethylamine was used instead of amine to form amide bond with carboxylic group of amino acid (**307**).
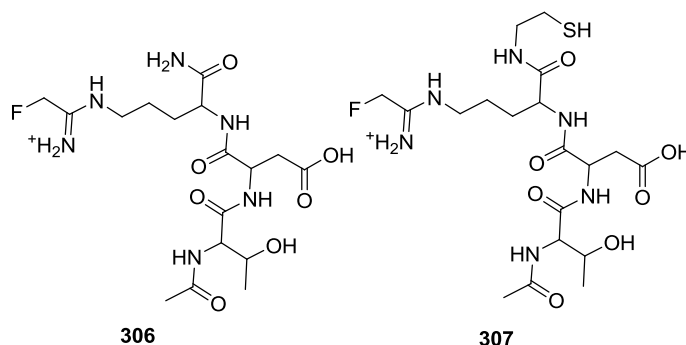
Bello et al. [34] replaced benzamide of Cl-amidine with acetamide and carboxylic group of amino acid was esterified. Compound **308** showed significantly weak activity compared to Cl-amidine. Confirming earlier report [30], 3 carbon atom chain was also found to be optimum.
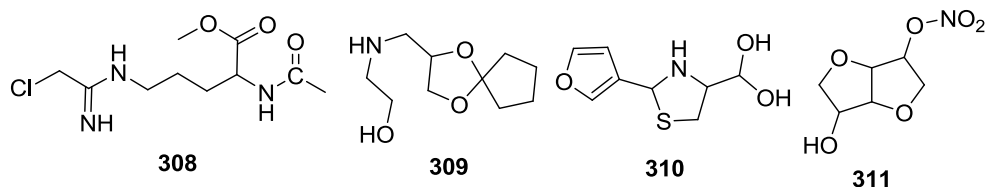
Teo et al. [35] virtually screened around 1 million compounds and tested 22 hits as reversible inhibitors. The reported IC_{50} values for compound **309**, **310** and **311** were 2.9, 1.6, and 1.5 mM respectively.

Bozdag et al. [36] investigated PAD4 inhibitory activity of compounds with primary and secondary guanidine moiety. Compounds **312** and **313** bearing guanidine and aromatic group showed inhibition between 4-6% at 1 and 10 μM concentrations. Cimetidine (**314**) was found to be inactive mainly because of presence of linear cyano group on guanidine, which hinders its interaction with catalytic site residue. Disubstituted guanidine compound **315** showed 34% inhibition at 10 μM concentration. In silico study concluded that disubstitution on guanidine group restricted its entry in active site, rendering it inactive but in vitro results indicated that the activity might be due to its binding to other site blocking active site cavity.
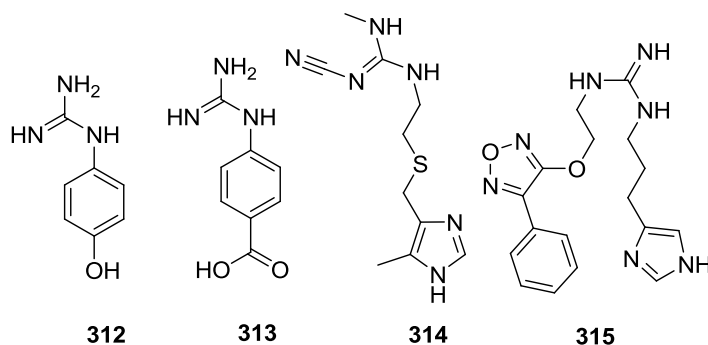
Wei et. al. [37] performed in silico virtual screening of commercial library and 10 compounds out of 800 hits were selected for in vitro testing at 30 μM concentration. Four compounds (**316-319**) showed activity with IC_{50} values ranging from 1.32-1.76 mM.
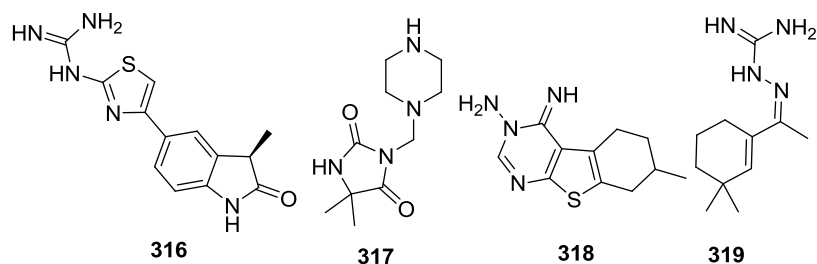
3.1.6   References

[1] Rogers, G.E.; Simmonds, D.H. Content of citrulline and other amino-acids in a protein of hair follicles. *Nature.* **1958,** 182(4629), 186-187.

[2] Vossenaar, E.R.; Zendman, A.J.; van Venrooij, W.J.; Pruijn, G.J. PAD, a growing family of citrullinating enzymes: genes, features and involvement in disease. *Bioessays.* **2003,** 25(11), 1106-1118.

[3] Mastronardi, F.G.; Wood, D.D.; Mei, J.; Raijmakers, R.; Tseveleki, V.; Dosch, H.-M.; Probert, L.; Casaccia-Bonnefil, P.; Moscarello, M.A. Increased citrullination of histone H3 in multiple sclerosis brain and animal models of demyelination: a role for tumor

necrosis factor-induced peptidylarginine deiminase 4 translocation. *J Neurosci.* **2006,** 26(44), 11387-11396.

[4] Mohanan, S.; Cherrington, B.D.; Horibata, S.; McElwee, J.L.; Thompson, P.R.; Coonrod, S.A. Potential role of peptidylarginine deiminase enzymes and protein citrullination in cancer pathogenesis. *Biochem Res Int.* **2012,** 2012, Article ID: 895343.

[5] Meamar, R.; Askari, G.; Ghasemi, M.; Ghazvini, M.R.A.; Vesal, S.; Sharifkhah, M.; Faradonbeh, N.A.; Dehghani, L. Is anti-cyclic citrullinated Peptide antibody a good value biomarker for Alzheimer disease? *Int J Prev Med.* **2013,** 4(Suppl 2), S201-S204.

[6] van Venrooij, W.J.; Pruijn, G.J. Citrullination: a small change for a protein with great consequences for rheumatoid arthritis. *Arthritis Res.* **2000,** 2(4), 249-251.

[7] Nakashima, K.; Arai, S.; Suzuki, A.; Nariai, Y.; Urano, T.; Nakayama, M.; Ohara, O.; Yamamura, K.-i.; Yamamoto, K.; Miyazaki, T. PAD4 regulates proliferation of multipotent haematopoietic cells by controlling c-myc expression. *Nature Commun.* **2013,** 4, 1836.

[8] Nachat, R.; Méchin, M.-C.; Takahara, H.; Chavanas, S.; Charveron, M.; Serre, G.; Simon, M. Peptidylarginine deiminase isoforms 1–3 are expressed in the epidermis and involved in the deimination of K1 and filaggrin. *J Invest Dermatol.* **2005,** 124(2), 384-393.

[9] Foulquier, C.; Sebbag, M.; Clavel, C.; Chapuy-Regaud, S.; Al Badine, R.; Méchin, M.C.; Vincent, C.; Nachat, R.; Yamada, M.; Takahara, H. Peptidyl arginine deiminase type 2 (PAD-2) and PAD-4 but not PAD-1, PAD-3, and PAD-6 are expressed in rheumatoid arthritis synovium in close association with tissue inflammation. *Arthritis Rheum.* **2007,** 56(11), 3541-3553.

[10] Harauz, G.; Ishiyama, N.; Hill, C.; Bates, I.R.; Libich, D.S.; Fares, C. Myelin basic protein—diverse conformational states of an intrinsically unstructured protein and its roles in myelin assembly and multiple sclerosis. *Micron.* **2004,** 35(7), 503-542.

[11] Guo, Q.; Fast, W. Citrullination of inhibitor of growth 4 (ING4) by peptidylarginine deminase 4 (PAD4) disrupts the interaction between ING4 and p53. *J Biol Chem.* **2011,** 286(19), 17069-17078.

[12] Wang, Y.; Wysocka, J.; Sayegh, J.; Lee, Y.-H.; Perlin, J.R.; Leonelli, L.; Sonbuchner, L.S.; McDonald, C.H.; Cook, R.G.; Dou, Y. Human PAD4 regulates histone arginine methylation levels via demethylimination. *Science.* **2004,** 306(5694), 279-283.

[13] Suzuki, A.; Yamada, R.; Chang, X.; Tokuhiro, S.; Sawada, T.; Suzuki, M.; Nagasaki, M.; Nakayama-Hamada, M.; Kawaida, R.; Ono, M. Functional haplotypes of PADI4, encoding citrullinating enzyme peptidylarginine deiminase 4, are associated with rheumatoid arthritis. *Nat Genet.* **2003,** 34(4), 395-402.

[14] Kinloch, A.; Lundberg, K.; Wait, R.; Wegner, N.; Lim, N.H.; Zendman, A.J.; Saxne, T.; Malmstr, V.; Venables, P.J. Synovial fluid is a site of citrullination of autoantigens in inflammatory arthritis. *Arthritis Rheum.* **2008,** 58(8), 2287-2295.

[15] Zhao, J.; Zhao, Y.; He, J.; Jia, R.; Li, Z. Prevalence and significance of anti-peptidylarginine deiminase 4 antibodies in rheumatoid arthritis. *J Rheumatol.* **2008,** 35(6), 969-974.

[16] Syversen, S.W.; Goll, G.L.; Van der Heijde, D.; Landewé, R.; Lie, B.A.; odegård, S.; Uhlig, T.; Gaarder, P.I.; Kvien, T.K. Prediction of radiographic progression in rheumatoid

arthritis and the role of antibodies against mutated citrullinated vimentin: results from a 10-year prospective study. *Ann Rheum Dis.* **2010,** 69(2), 345-351.

[17] Cantaert, T.; Teitsma, C.; Tak, P.P.; Baeten, D. Presence and role of anti–citrullinated protein antibodies in experimental arthritis models. *Arthritis Rheum.* **2013,** 65(4), 939-948.

[18] Wiik, A.S. The immune response to citrullinated proteins in patients with rheumatoid arthritis. *Clin Rev Allergy Immunol.* **2007,** 32(1), 13-21.

[19] Van Gaalen, F.A.; Van Aken, J.; Huizinga, T.W.; Schreuder, G.M.T.; Breedveld, F.C.; Zanelli, E.; Van Venrooij, W.J.; Verweij, C.L.; Toes, R.E.; De Vries, R.R. Association between HLA class II genes and autoantibodies to cyclic citrullinated peptides (CCPs) influences the severity of rheumatoid arthritis. *Arthritis Rheum.* **2004,** 50(7), 2113-2121.

[20] Furuya, T.; Hakoda, M.; Ichikawa, N.; Higami, K.; Nanke, Y.; Yago, T.; Kobashigawa, T.; Tokunaga, K.; Tsuchiya, N.; Kamatani, N. Differential association of HLA-DRB1 alleles in Japanese patients with early rheumatoid arthritis in relationship to autoantibodies to cyclic citrullinated peptide. *Clin Exp Rheumatol.* **2007,** 25(2), 219-224.

[21] Bang, S.Y.; Lee, K.H.; Cho, S.K.; Lee, H.S.; Lee, K.W.; Bae, S.C. Smoking increases rheumatoid arthritis susceptibility in individuals carrying the HLA–DRB1 shared epitope, regardless of rheumatoid factor or anti–cyclic citrullinated peptide antibody status. *Arthritis Rheum.* **2010,** 62(2), 369-377.

[22] Vossenaar, E.R.; Zendman, A.J.; Van Venrooij, W.J. Citrullination, a possible functional link between susceptibility genes and rheumatoid arthritis. *Arthritis Res Ther.* **2004,** 6(1), 1-5.

[23] Horikoshi, N.; Tachiwana, H.; Saito, K.; Osakabe, A.; Sato, M.; Yamada, M.; Akashi, S.; Nishimura, Y.; Kagawa, W.; Kurumizaka, H. Structural and biochemical analyses of the human PAD4 variant encoded by a functional haplotype gene. *Acta Crystallogr, Sect D: Biol Crystallogr.* **2011,** 67(2), 112-118.

[24] Consortium, T.U. Update on activities at the Universal Protein Resource (UniProt) in 2013. *Nucleic Acids Res.* **2013,** 41(D1), D43-D47.

[25] PADI4 peptidyl arginine deiminase, type IV [Homo sapiens (human)]. http://www.ncbi.nlm.nih.gov/gene?term=45426862[PUID]&RID=5TW9RS5G016&log$= genealign&blast_rank=5 (Accessed Oct. 10th, 2013).

[26] Protein Data Bank. http://www.rcsb.org/pdb.

[27] Ke, Z.; Wang, S.; Xie, D.; Zhang, Y. Born–Oppenheimer ab Initio QM/MM molecular dynamics simulations of the hydrolysis reaction catalyzed by protein arginine deiminase 4. *J Phys Chem B.* **2009,** 113(52), 16705-16710.

[28] Arita, K.; Hashimoto, H.; Shimizu, T.; Nakashima, K.; Yamada, M.; Sato, M. Structural basis for Ca2+-induced activation of human PAD4. *Nat Struct Mol Biol.* **2004,** 11(8), 777-783.

[29] Arita, K.; Shimizu, T.; Hashimoto, H.; Hidaka, Y.; Yamada, M.; Sato, M. Structural basis for histone N-terminal recognition by human peptidylarginine deiminase 4. *Proc Natl Acad Sci.* **2006,** 103(14), 5291-5296.

[30] Luo, Y.; Arita, K.; Bhatia, M.; Knuckley, B.; Lee, Y.-H.; Stallcup, M.R.; Sato, M.; Thompson, P.R. Inhibitors and inactivators of protein arginine deiminase 4: functional and structural characterization. *Biochemistry.* **2006,** 45(39), 11727-11736.

[31] Causey, C.P.; Jones, J.E.; Slack, J.L.; Kamei, D.; Jones Jr, L.E.; Subramanian, V.; Knuckley, B.; Ebrahimi, P.; Chumanevich, A.A.; Luo, Y. The Development of N-α-(2-Carboxyl) benzoyl-N 5-(2-fluoro-1-iminoethyl)-l-ornithine Amide (o-F-amidine) and N-α-(2-Carboxyl) benzoyl-N 5-(2-chloro-1-iminoethyl)-l-ornithine Amide (o-Cl-amidine) as second generation protein arginine deiminase (PAD) inhibitors. *J Med Chem.* **2011,** 54(19), 6919-6935.

[32] Bicker, K.L.; Anguish, L.; Chumanevich, A.A.; Cameron, M.D.; Cui, X.; Witalison, E.; Subramanian, V.; Zhang, X.; Chumanevich, A.P.; Hofseth, L.J. d-Amino acid-based protein arginine deiminase inhibitors: synthesis, pharmacokinetics, and in cellulo efficacy. *ACS Med Chem Lett.* **2012,** 3(12), 1081-1085.

[33] Jones, J.E.; Slack, J.L.; Fang, P.; Zhang, X.; Subramanian, V.; Causey, C.P.; Coonrod, S.A.; Guo, M.; Thompson, P.R. Synthesis and screening of a haloacetamidine containing library to identify PAD4 selective inhibitors. *ACS Chem Biol.* **2011,** 7(1), 160-165.

[34] Bello, A.M.; Wasilewski, E.; Wei, L.; Moscarello, M.A.; Kotra, L.P. Interrogation of the active sites of protein arginine deiminases (PAD1,-2, and-4) using designer probes. *ACS Med Chem Lett.* **2013,** 4(2), 249-253.

[35] Teo, C.Y.; Shave, S.; Chor, A.L.T.; Salleh, A.B.; Rahman, M.B.B.A.; Walkinshaw, M.D.; Tejo, B.A. Discovery of a new class of inhibitors for the protein arginine deiminase type 4 (PAD4) by structure-based virtual screening. *BMC Bioinformatics.* **2012,** 13(Suppl 17), S4.

[36] Bozdag, M.; Dreker, T.; Henry, C.; Tosco, P.; Vallardo, M.; Fruttero, R.; Scozzafava, A.; Carta, F.; Supuran, C.T. Novel small molecule protein arginine deiminase 4 (PAD4) inhibitors. *Bioorg Med Chem Lett.* **2013,** 23(3), 715-719.

[37] Wei, L.; Wasilewski, E.; Chakka, S.K.; Bello, A.M.; Moscarello, M.A.; Kotra, L.P. Novel inhibitors of protein arginine deiminase with potential activity in multiple sclerosis animal model. *J Med Chem.* **2013,** 56(4), 1715-1722.

3.2 Objectives

C itrullinated proteins act as self antigen and trigger immune reaction. PAD4 citrullinates proteins in synovium therefore inhibition of PAD4 can serve as target for drug development against RA. Reported PAD4 inhibitors are irreversible inhibitors binding covalently to the enzyme. It poses threat by binding to other enzymes and leading to toxicity which was evidenced in one of the case [1]. As non covalently binding reversible PAD4 inhibitors were not reported at the time when project was started, we aimed to design and develop non covalently binding inhibitors that mimic arginine.

Aim 1: To design database of virtual compounds using the knowledge of receptor structure and existing arginine mimetic inhibitors. Also to screen the database in silico to select hits by following criteria:

- Hit docking score should be comparatively high
- Hit should have optimal complementary interactions with receptor
- Hit should be devoid of amino acids and any stereocenter if possible
- Hit should be comparatively rigid than arginine
- Hit should not violate Lipinski rule of oral bioavailability
- Hit synthesis should be feasible in our laboratory conditions
- Hit should have sites for derivatization

Aim 2: To synthesize the in silico hits by different synthetic approaches followed by purification and characterization

Aim 3: To screen synthesized compounds for their potential PAD4 inhibitory activity using in-vitro assay

3.2.1 Reference

[1] Bicker, K.L.; Anguish, L.; Chumanevich, A.A.; Cameron, M.D.; Cui, X.; Witalison, E.; Subramanian, V.; Zhang, X.; Chumanevich, A.P.; Hofseth, L.J. D-amino acid-based protein arginine deiminase inhibitors: synthesis, pharmacokinetics, and in cellulo efficacy. *ACS Med Chem Lett.* **2012,** 3(12), 1081-1085.

3.3 Structure Based Design of PAD4 Inhibitors

Structure Based Drug Design (SBDD) utilizes the knowledge of 3 dimensional structure of the receptor obtained by X-ray or NMR or by homology modeling. In SBDD, by analyzing the active site, ligands are designed that have shape and chemical complimentarity with the receptor. The docking program estimates their binding affinity in two steps. First step involves generation of different conformers and positioning these conformers in the active site and then second step scores the pose. The energetically favored poses with shape and electrostatic complimentarity are scored as hits [1].

With the available crystal structure of PAD4, SBDD of PAD4 inhibitors was performed. Since PAD4 acts on arginine residues of the protein, arginine mimetic ligands were designed using existing knowledge base. Arginine mimetic inhibitors are reported in literature for targets like Thrombin, Factor Xa, and Histamine H2 receptors. By docking these inhibitors against PAD4, hits were identified by analyzing binding patterns and docking scores.

3.3.1 Methodology

Preparation of protein:

Human PAD4 crystal structure 1WDA pdb [2] in complex with benzoyl-L-arginine amide (BAA) was used for docking study. In 1WDA residue Ala645 was mutated from Cys645 to aid crystallization of PAD4 with its substrate BAA. Using Schrodinger Maestro [3] program, Ala645 was mutated back to Cys645 and appropriate rotamer was selected based on position of Cys645 in 1WD9 (human PAD4 pdb structure). Using protein preparation wizard, bond orders were assigned for all atoms. Ionization and metal binding modes were detected using Epik at pH 7.0±2.0 [4]. All water molecules were removed except HOH941, as it is conserved in PAD family. Hydrogens were added and protein was minimized to the extent of 0.3 Å RMSD using Optimized Potentials for Liquid Simulations (OPLS) 2005 force field.

Virtual database creation:

A virtual database of 642 arginine mimetic compounds was drawn from 26 articles reporting inhibitors against Thrombin, Factor Xa, Histamine H2 receptor, RAF kinase and Factor VIIa. For virtual screening (VS), database was standardized using ligand preparation wizard of Maestro. All possible ionization states at pH 7.0±2.0 were enumerated using Epik

and maximum of 2 stereoisomers were enumerated by Stereoizer (a module of Schrodinger Maestro [3] program), when chirality was not specified. The ligands were then minimized using OPLS 2005 force field. For VS by openeye software, the database was passed through Pkatyper [5] to enumerate different ionization states of the molecule. Then using Flipper [6], stereoisomers were enumerated for the molecules when chirality was not specified and maximum of 2 chiral centers were considered per molecules.

Docking with Glide

Grid-based Ligand Docking with Energetic (Glide) performs flexible docking in reference to ligand. Grid of 1WDA active site was created using Maestro. A centroid of cocrystalised ligand was selected to define gride box and size was determined with default settings. Validation of Glide-SP docking was done using cocrystalised ligand BAA. After satisfactory validation results, virtual database of ligands were docked to 1WDA using Glide SP (standard precision) [7].

Docking with FRED

FRED is a Fast Rigid Exhaustive Docking algorithm developed by Openeye [8]. FRED performs rigid docking of multi conformer file of a molecule. Since conformers are not generated on fly, another program called OMEGA is used [9]. PAD4 pdb structure file prepared by Maestro was selected and active site grid was generated using FRED receptor setup program [10]. All water molecules were removed and a grid box of volume 6935 Å was created around cocrystallised ligand. Tautomeric state of His471 and His644 was decided based on best possible interaction with ligand and neighboring carboxylate ions of aspartic acids. Inner and outer contour was set at 64 Å and 3436 Å respectively. After getting satisfactory results in trial dock, grid file was saved for future use. Maximum of 1000 conformers falling in the energy window of 25 Kcal/mol from its minima were generated for each ligand with 0.8 RMSD between each conformer. Search force field was set to MMFF94. Multiconformer database was then docked with receptor grid generated earlier.

3.3.2 Results and discussion

Virtual database of arginine mimetic compounds were created, standardized and minimized using appropriate force field. PDB structure 1WDA containing bound substrate was mutated from Ala645 to Cys645. Hydrogen were added and their position was optimized based on ligands and water molecules. Ionization pattern of PAD4 at pH around 7

was assigned and visually confirmed in the active site. Conserved water was retained in Glide docking and striped off for FRED docking.
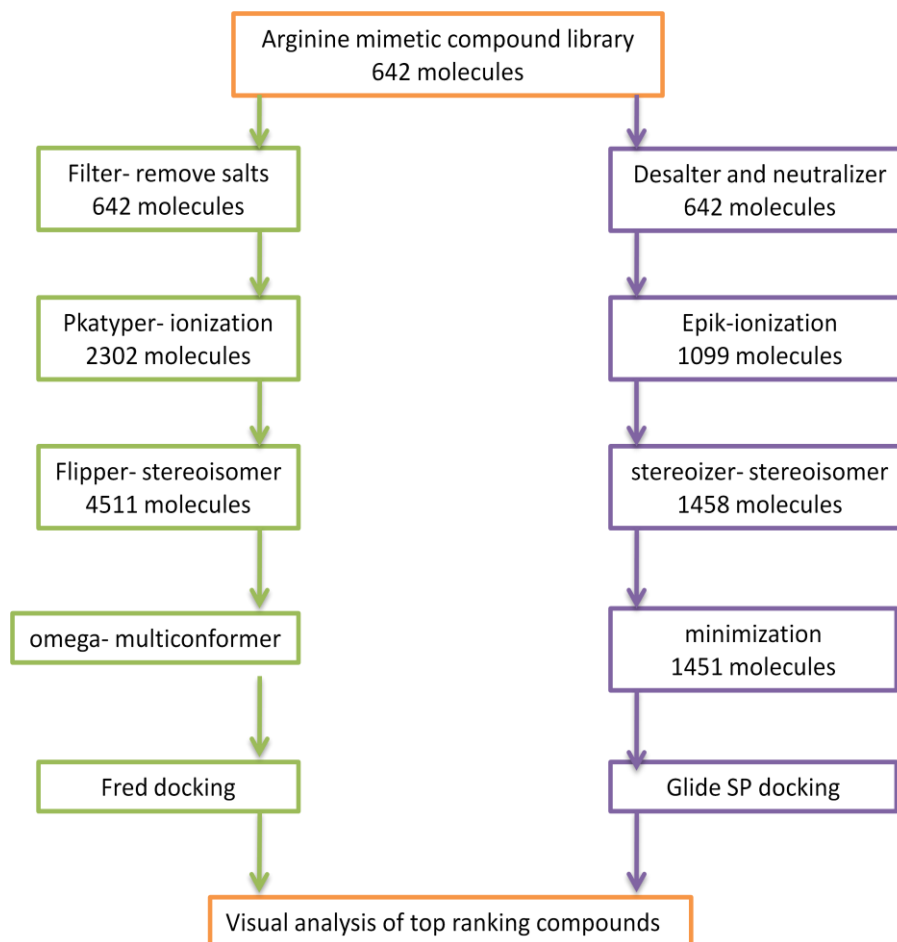
Figure 3.3-1: Schematic representation of different steps of virtual screening of arginine mimetic compounds against PAD4 (1WDA) using FRED and Glide docking software.

As Epik identifies physiological ionization state of molecules using Hammett and Taft based pKa prediction, small number of ionization states were produced compared to pkatyper. Number of compounds was increased again when possible stereoisomers were created for undefined chiral centers. The RMSD between docked pose and cocrystallised pose of the substrate BAA form both software was well below 2 Å. Using glide standard precision mode and FRED, docking was performed with virtual database. Various stages of VS are illustrated in **Figure 3.3-1**. Docking results were sorted according to the scores. Top 100 compounds were selected and visually analyzed for their binding modes. Glide docking retrieved hits having binding pattern similar to that of substrate BAA as it appears that glide docking score is highly influenced by number of hydrogen bonding interactions. Structures of some of the selected hits are represented in **Figure 3.3-2**.
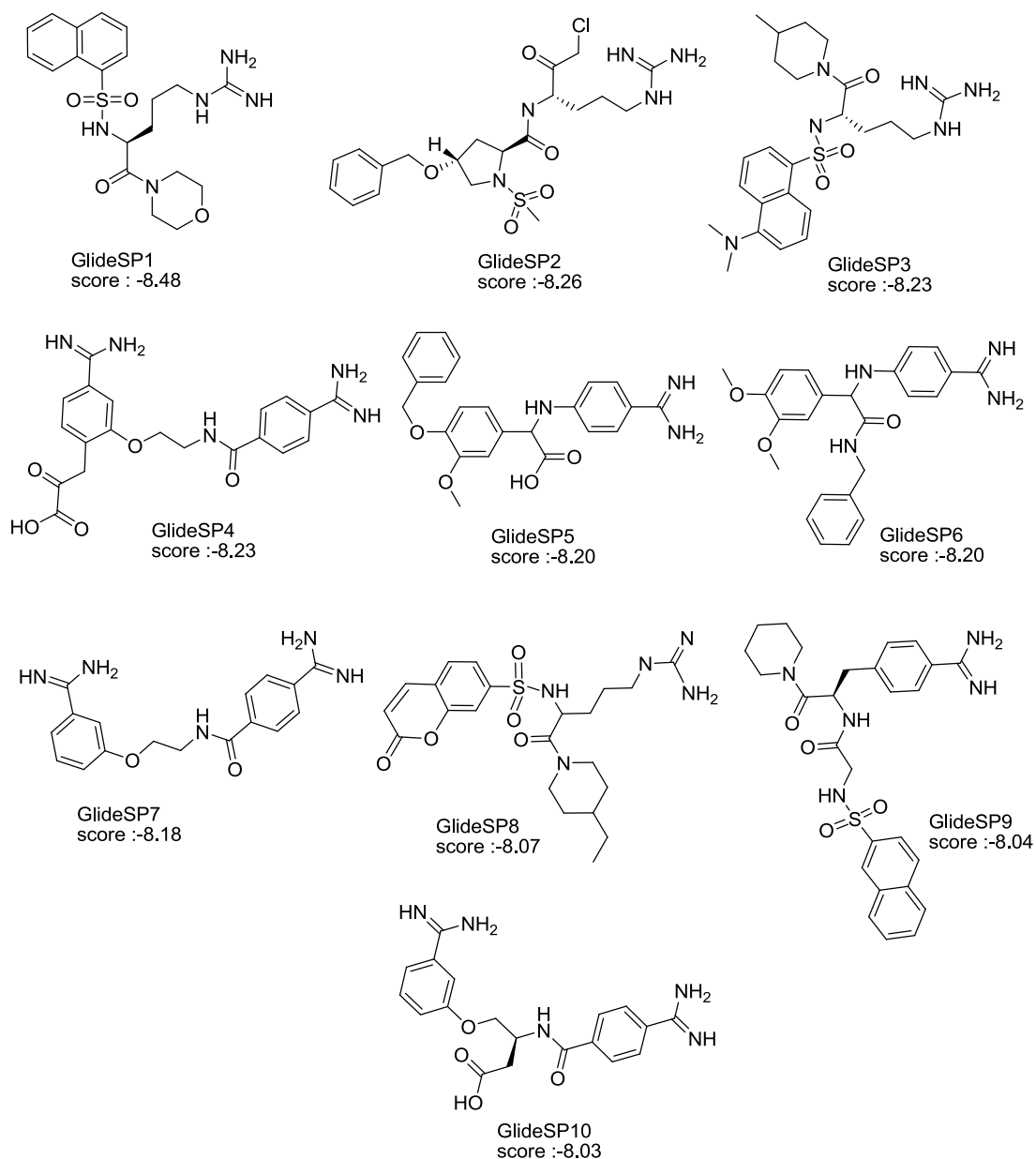
Figure 3.3-2: Structures of 10 selected Glide hits having good score.

In FRED Chemgauss3, many top ranked poses were different from the ones produced by Glide. Chemgauss3 scoring function which is composed by Chemgauss3 steric + Chemgauss3 desolvation + Chemgauss3 donor + Chemgauss3 acceptor + Chemgauss3 chelator terms. Chemgauss3 steric score was found to have strong influence on total score than others. Therefore, poses with less number of hydrogen bonding interaction but good van der waals interaction were ranked at top. Some of the compounds having good score in FRED are given in **Figure 3.3-3**. Arginine analogues showed greater binding energy than benzamidines in both docking protocols and its frequency of being hit was high in FRED.
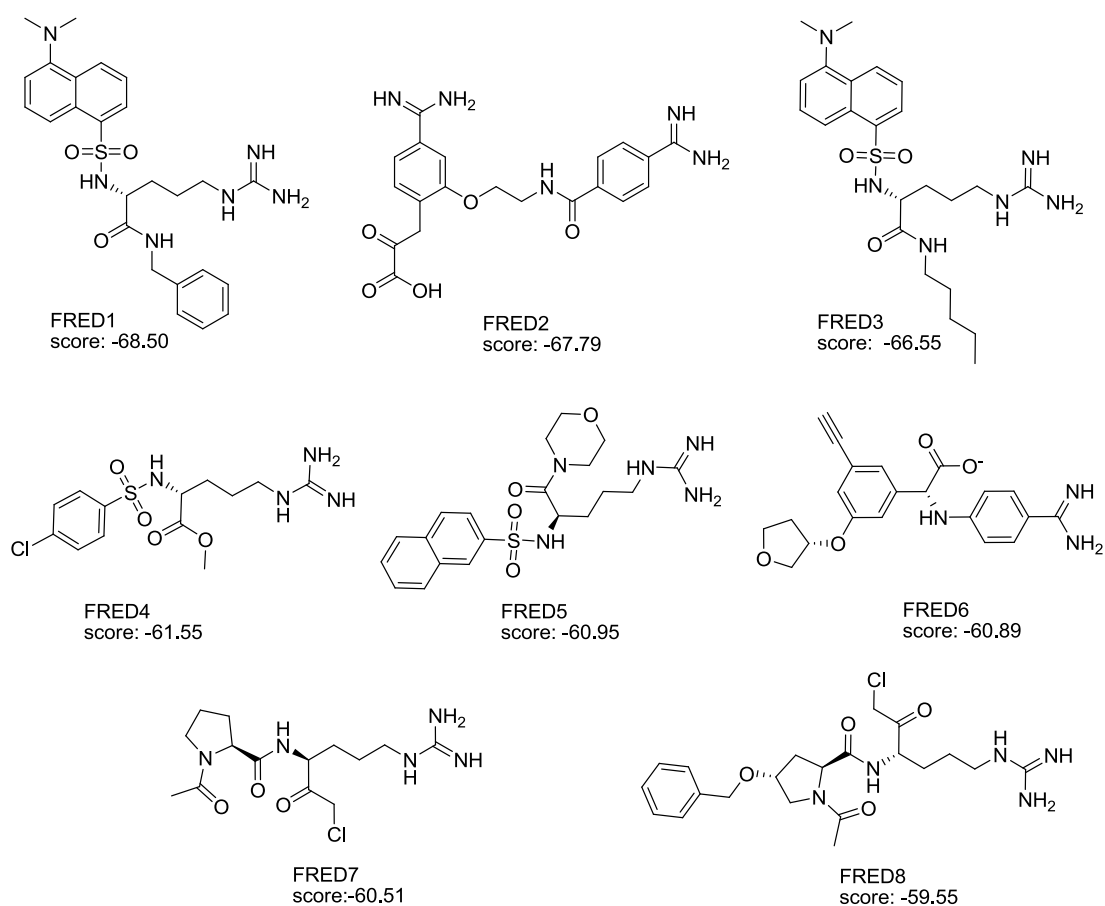
Figure 3.3-3: Structures of 8 selected FRED hits having good score.

Visual analysis of docking poses indicated that amidine can acts as bioisosteric replacement for guanidine group and benzene ring for 3 carbon atom chain of substrate BAA. Hits GlideSP1 and GlideSP3 docked with greater binding energy (-8.48 and -8.23 Kcal/mol respectively) and 8 hydrogen bonding interactions were observed with the protein (**Figure 3.3-4**). Guanidine head made electrostatic interaction with Asp473, Asp350 and conserved water. Carbon chain of ligands was in hydrophobic region formed by Val469 and Trp347 where as carbonyl and sulfonyl oxygen interacted with Arg374 by H-bond, amine of sulfonamide group made H-Bond interaction with carbonyl group of Arg639 in protein backbone. Overall binding pattern was similar to that of substrate BAA. In GlideSP4 and GlideSP6 hits, terminal basic group is amidine instead of guanidine. Amidine group made lesser number of H-bond interactions compared to guanidine. Benzene ring replaced for carbon chain made good van der waals interactions. Orientation of ring in GlideSP6 hit was much favorable for pi-pi interaction with Trp347.
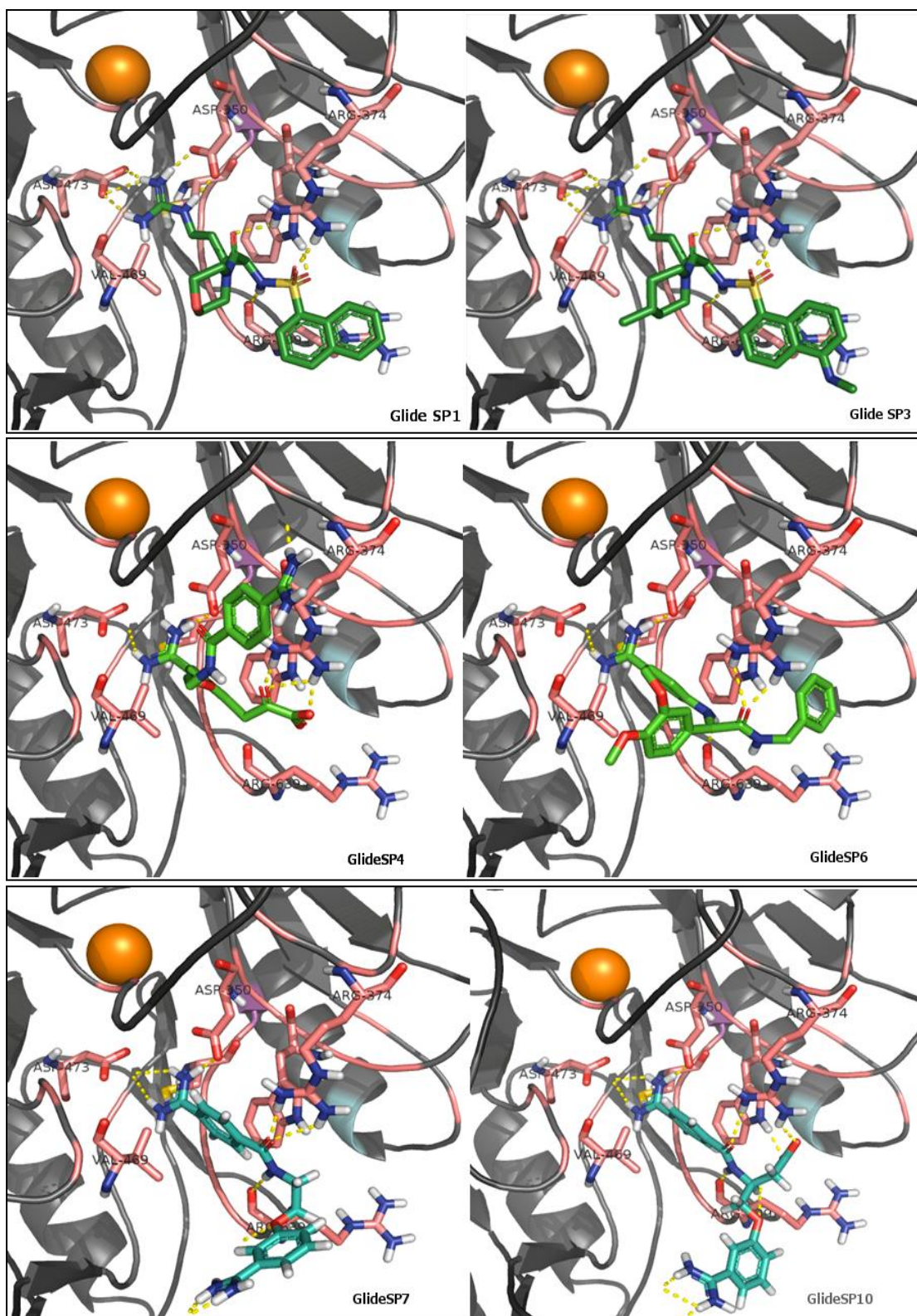
Figure 3.3-4: Interaction diagram of GlideSP1, GlideSP3, GlideSP4, GlideSP6, GlideSP7 and GlideSP10 hits with 1WDA structure. Residues in the active site are shown in pink and hits in green. Rest of the protein is represented as cartoon. Hydrogen bonds are represented as yellow broken lines. Orange colored sphere near Asp473 and Asp350 amino acids is calcium ion.

A contact with Arg374 was made by terminal keto carboxylic acid in GlideSP4 and by carbonyl group of benzylamide in GlideSP6. Due to increased length in GlideSP4 amide carbonyl group was not involved in interaction with Arg374. GlideSP4 lacks H-bond donor group near the carbonyl group of Arg639, whereas this interaction was retained in GlideSP6. Effect of less H-bond interactions was negligible on binding energy due to enhanced hydrophobic interactions compared to GlideSP1, and GlideSP3 hits (**Figure 3.3-4**).

Hits GlideSP7 and GlideSP10 are 4-carbamimidoylbenzoic acid derivatives. Amidine group made lesser H-bonding interactions with Asp473 and Asp350 as compared to guanidine (**Figure 3.3-4**). Benzene ring had favorable van der waals interaction with Val469 and pi-pi stacking interaction with Trp347. Amide carbonyl oxygen and amine made hydrogen bonding interaction with Arg374 and Arg639 respectively. Free carboxylic group of GlideSP10 interacted with ring nitrogen of Trp347 and Arg374. Ether oxygen and terminal amidine groups in both the hits had favorable interactions.

3.3.3   Conclusion

Analysing binding pattern of the hits in Glide and FRED docking and comparing structural similarity of the hits and PAD4 substrate (BAA), it was planed to design compounds with base at position 1, hydrophobic group at position 2 and hydrogen bond donor/acceptor at position 3 of the ligand skeleton given in **Table 3.3-1**.

Table 3.3-1: Proposed PAD4 inhibitor structures with possible variations.
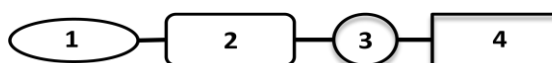
Pharmacophoric group 1	Pharmacophoric group 2	Pharmacophoric group 3	Non pharmacophoric group 4
NH_2-		- CO- NH_2-	-alkyl
		$-NH_2-CO-$	-aryl
		$-NH_2-SO_2-$	
		$-NH_2-CO-NH_2-$	

The molecules designed had amine or amidine or guanidine group at position 1, a benzyl or benzene ring with para or meta substitution at position 2 and amide or sulfonamide or urea group at position 3. These structural variations can explain how basicity,

hydrophobicity, distance between position1 and 3 changed by para or meta substitution, and position 3 groups affect orientation of aromatic ring and activity of PAD4 enzyme. Since position 4 is outside the cavity, different alkyl and aryl substitutions can help to identify possiable residue that enhances binding and favours tighter binding of the ligand with the enzyme.

3.3.4 References

[1]   Kroemer, R.T. Structure-based drug design: docking and scoring. *Curr Protein Pept Sci.* **2007,** 8(4), 312-328.

[2]   Arita, K.; Hashimoto, H.; Shimizu, T.; Nakashima, K.; Yamada, M.; Sato, M. Structural basis for Ca(2+)induced activation of human PAD4. *Nat Struct Mol Biol.* **2004,** 11(8), 777-778.

[3]   Schrodinger Suite 2009: Maestro, version 9.0, Schrödinger, LLC, New York, NY, **2009**.

[4]   Shelley, J.C.; Cholleti, A.; Frye, L.L.; Greenwood, J.R.; Timlin, M.R.; Uchimaya, M. Epik: a software program for pK a prediction and protonation state generation for drug-like molecules. *J Comput Aided Mol Des.* **2007,** 21(12), 681-691.

[5]   QUAPAC version 1.3.1, OpenEye Scientific Software S F, NM, USA, www.eyesopen.com, **2010**.

[6]   Flipper version 2.3.2, OpenEye Scientific Software S F, NM, USA, www.eyesopen.com, **2010**.

[7]   Halgren, T.A.; Murphy, R.B.; Friesner, R.A.; Beard, H.S.; Frye, L.L.; Pollard, W.T.; Banks, J.L. Glide: a new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening. *J Med Chem.* **2004,** 47(7), 1750-1759.

[8]   Sagi, K.; Nakagawa, T.; Yamanashi, M.; Makino, S.; Takahashi, M.; Takayanagi, M.; Takenaka, K.; Suzuki, N.; Oono, S.; Kataoka, N. Rational design, synthesis, and structure-activity relationships of novel Factor Xa inhibitors:(2-Substituted-4-amidinophenyl) pyruvic and-propionic Acids. *J Med Chem.* **2003,** 46(10), 1845-1857.

[9]   Groebke Zbinden, K.; Banner, D.W.; Ackermann, J.; D'Arcy, A.; Kirchhofer, D.; Ji, Y.-H.; Tschopp, T.B.; Wallbaum, S.; Weber, L. Design of selective phenylglycine amide tissue factor/factor VIIa inhibitors. *Bioorg Med Chem Lett.* **2005,** 15(3), 817-822.

[10] FRED version 2.2.5, OpenEye Scientific Software S F, NM, USA, www.eyesopen.com, **2010**.

3.4 In silico screening of designed PAD4 Inhibitors

Virtual database of the molecules comprising three series viz. Benzamidine, *p*-amino-benzamide and *p*-(methylamino)-benzamide was designed as per the template mentioned in **section 3.3.3**. Database was then screened for in silico oral bioavailability and toxicity prediction followed by virtual screening against PAD4.

3.4.1 Methodology

3.4.1.1 Oral bioavailability prediction

Various physicochemical properties of ligands that influence oral bioavailability, like Molecular Weight (MW), hydrogen bond donor count (HBD) and hydrogen bond acceptor count (HBA), topological polar surface area (TPSA), Molar Refractivity (MR) and log partition coefficient (logP) were determined using Chemaxon Jchem for excel [1] and aqueous solubility with AlogPS v2.1 algorithm [2]. All these parameters affect solubility and partitioning between biological barriers which can have a direct correlation with oral bioavailability. Based on observation of approved drugs, various rule were developed to predict oral bioavailability. One such rule is Rule of five or Lipinsky rule [3] which is applicable for the drugs absorbed by passive diffusion through cell membrane. Another rule called GSK oral bioavailability rule considers TPSA and number of rotatable bonds of the molecule (number of rotatable bonds <10, topological polar surface area < 140 $Å^2$) [4].

3.4.1.2 Toxicity prediction

All molecules should be tested for biological safety in order to minimize the risk of elimination at the later phase of clinical development. Some of the common toxicity screening tests involve study of mutagenicity, tumorigenicity, skin irritation, reproductive toxicity, cardiotoxicity, etc. In this study, the toxicity risk assessment was carried using OSIRIS property explorer [5]. The Drug score is calculated directly from cLogP, logS, molecular weight while drug likeness takes into account the drug score and risk factors [6]. Risk alerts are calculated based on presence of fragments with known risk factors. However, these risk alerts are nor fully reliable. And absence of risk alerts does not indicate that molecules are free of toxicity.

3.4.1.3 In silico binding energy and pose prediction

Protein and ligand preparation: Human PAD4 crystal structure 1WDA pdb [7], in complex with benzoyl-L-arginine amide (BAA) was used for docking study. 1WDA is a mutated protein where active site residue Cys645 was mutated to Ala645. Using Schrodinger Maestro [8] program, Ala645 was mutated back to Cys645 and appropriate rotamer was selected based on position of Cys645 in 1WD9, another human PAD4 pdb structure. Using protein preparation wizard, bond orders were assigned for all atoms. Ionization and metal binding modes were detected using Epik at pH 7.0±2.0 [9]. All water molecules were removed except HOH941 as it is conserved in PAD family. Hydrogens were added and protein was minimized to the extent of 0.3 Å RMSD using Optimized Potentials for Liquid Simulations (OPLS) 2005 force field. All the molecules were drawn using Chemaxon Marvin Sketch [10] and prepared with ligand preparation wizard of Maestro. All possible Ionization state at pH 7.0±2.0 were enumerated using Epik [11] and minimized. For Autodock, ligands were further processed by adding Gasteiger charges using MGL tools v1.5.4 [12].

Docking with Glide: Grid of 1WDA active site was created using Maestro. A centroid of cocrystallised ligand was selected (x,y,z coordinates: 24.117, 49.335, 22.86) to define gride box and grid box length was set to 20, 20 and 20 Å along x, y and z direction respectively. Docking of Maestro processed ligand database was performed using Glide SP (Standard Precision) and Glide XP (Extra Precision) [13].

Docking with FRED: Pdb structure file prepared by Maestro was selected and active site grid was generated using Fred receptor setup program [14]. Water molecules were removed and a grid box of volume 6935 $\mathring{A}^3$ was created around cocrystallised ligand. Tautomeric state of His471 and His644 was decided based on best possible interaction with ligand and neighboring carboxylate ions of aspartic acids. Inner and outer contours were set at 64 Å and 3436 Å respectively. After getting satisfactory results in trial dock, grid file was saved for future use. FRED does not produces ligand conformers on fly therefore conformers were generated using OMEGA [15]. Maximum of 1000 conformers falling in the energy window of 25 Kcal/mol from its minima were generated for each ligands with 0.5 RMSD between each conformer. Search force field was set to Merck Molecular Force Field 94 (MMFF94). Multiconformer database was then docked with receptor grid generated earlier. FRED was allowed to assign protein charges and initial optimization was done with chemgauss3 empirical scoring function then local minimization using MMFF 94 force field was carried out and final poses were scored with chemgauss3, chemgauss2, shapegauss, Piecewise Linear

Potential (PLP), Chemical Gaussian Overlay (CGO), Chemical Gaussian Tanimoto (CGT) scores and ranked. While reporting only chemgauss3 values are given.

Docking with Autodock: Autodock V4.2 employing genetic algorithm as search program was used for the study [12]. 1WDA pdb structure file prepared by Maestro was selected, water molecules were removed and Kollman charges were assigned to the protein. Energy grid of the active site was created using Autogrid 4.2. A box with npts (number of points in xyz) of 54-54-54 Å with center at 24.117, 49.335 and 22.86 (x,y, z) coordinate. The npts spacing was kept at 0.375 Å. Energy map for electrostatic potential, desolvation potential, atom type maps like carbon, oxygen as acceptor, nitrogen, hydrogen bond donor, sulfur, fluorine, chlorine and was created. Ligands were then docked to the active site with genetic algorithm and lamarckian search (GA-LS) parameter producing 27000 generations in each run. And 15 such runs were performed for each ligand. Docked poses were clustered and analyzed for interaction with the protein. Lowest energy binding poses are reported.

MM-GBSA Binding energy estimation: Molecular mechanics with generalized Born and surface area solvation (MM-GBSA) is a method to estimate free energy of solute-solvent interactions using implicit (continuum) solvation model [16]. MM-GBSA energy estimation includes OPLS molecular mechanics energy (E_{MM}), SGB solvation model for polar solvation (G_{SGB}) and non polar salvation term (G_{NP}). Therefore total free energy of binding is expressed as **Equation 3.4-1**.

Equation 3.4-1 :
$$\Delta G_{bind} = G_{complex} - (G_{protein} + G_{ligand})$$

Where ΔG_{bind} is change in the free energy on binding, $G_{complex}$ is a calculated average free energy of the complex and $G_{protein}$ and G_{ligand} are calculated average free energy of protein and ligands respectively. Free energy G is calculated from the **Equation 3.4 2.**

Equation 3.4-2:
$$G = E_{MM} + G_{SGB} + G_{NP}$$

Glide XP poses were selected for binding energy estimation. All the protein atoms were kept frozen while ligand was made flexible. The minimization was carried out using Prime MM-GBSA [17] for 100 iterations. Ligands in unbound state were minimized with variable dielectric surface generalized Born (VSGB) solvation model and free energy was calculated.

3.4.2 Results and Discussion

3.4.2.1 Predicted oral bioavailability of designed molecules

The descriptors used for prediction of oral bioavailability are given in **Table 3.4-1, Table 3.4-2** and **Table 3.4-3** for Benzamidine, *p*-amino-benzamide and *p*-(methylamino)-benzamide series respectively. All the designed molecules fulfilled the Lipinski rule of five criteria for oral bioavailability.

Benzamidine series: JClogP of the derivatives varied from -0.06 to 2.43. Hydrogen bond donor atoms in benzamidine series varied between 4 and 5, while hydrogen bond acceptor atoms were between 5 and 8. Number of rotatable bonds were between 3 and 6. Topological polar surface area of sulfonamide benzamidine derivatives was higher than amide and urea analogues. Number of rotatable bonds and topological polar surface area of all the molecules were well within the criteria of GSK oral bioavailability prediction rule except for A4S2NA, A4S3NA and A4S4N [4]. Ghose rule which proposed that a drug like molecule have LogP between -0.4 and 5.6, MW of 160-480 Da, MR of 40-130 and number of atoms between 20-70, is also followed by many molecules [18]. Many molecules were predicted to have good aqueous solubility except for A4ACNA and A4S3TFA.

***p*-amino-benzamide series:** JClogP of the derivatives varied from -0.22 to 2.84. Hydrogen bond donor atoms in this series varied between 2-3 and hydrogen bond acceptor atoms varied between 3 and 6. Number of rotatable bonds were between 1 and 4. Topological polar surface area was between 55.12 and 88.62. Number of rotatable bonds and topological polar surface area of all the molecules were well within the criteria of GSK oral bioavailability prediction rule [4]. Ghose rule was also obeyed by all designed molecules of this series. Many molecules were predicted to have good aqueous solubility except for molecule 4A4BRN and 4ANAPN.

***p*-(methylamino)-benzamide series:** JClogP of the derivatives varied from -0.64 to 3.18. Same number (3) of Hydrogen bond donor atoms were observed in all the molecules but hydrogen bond acceptor count varied between 3 and 6. Number of rotatable bonds were between 3 to 5. Topological polar surface was between 55.12 to 103.95. Number of rotatable bonds and topological polar surface area of all the molecules in this series were well within the criteria of GSK oral bioavailability prediction rule [4]. Ghose rule was also obeyed by all designed molecules. Many molecules of this series were predicted to have low aqueous solubility. Molecule MNAPN was predicted to have very low aqueous solubility.

Table 3.4-1: Physicochemical properties of Benzamidine series

code	Solubility (g/Lts)	MW(Da)	HBD	HBA	TPSA (Å^2)	Rot	MR	JCLogP
A4A4MA	0.14	269.3	4	5	88.2	4	75.51	1.4
A4A3MA	0.14	269.3	4	5	88.2	4	75.51	1.24
A4ACNA	0.015	310.31	4	7	127.8	5	75.51	2.43
A4ABEA	0.11	253.3	4	4	78.97	4	84.58	1.97
A4ACPA	0.54	203.24	4	4	78.97	3	55.87	0.91
A4A4NA	0.071	284.27	4	7	127.8	4	75.36	1.11
A4A2FA	0.029	257.26	4	4	78.97	3	69.57	1.22
A4A3FA	0.028	257.26	4	4	78.97	3	69.57	1.48
A4A4FA	0.025	257.26	4	4	78.97	3	69.57	1.64
A4A2FuA	0.2	229.23	4	5	92.11	3	61.87	-0.06
A4A3NA	0.073	284.27	4	7	127.8	4	75.36	1.12
A4A2NA	0.072	284.27	4	7	127.8	4	75.36	0.88
A4ABNA	0.16	239.27	4	4	78.97	3	69.7	1.47
A4S4MA	0.072	305.35	4	6	113.65	5	79.8	0.39
A4S2NA	0.027	320.32	4	8	153.25	5	79.65	0.5
A4S3NA	0.033	320.32	4	8	153.25	5	79.65	0.5
A4S4NA	0.034	320.32	4	8	153.25	5	79.65	0.5
A4S2TA	0.05	289.35	4	5	104.42	4	78.41	1.07
A4S3TA	0.049	289.35	4	5	104.42	4	78.41	1.07
A4S4TA	0.049	289.35	4	5	104.42	4	78.41	1.07
A4S3TFA	0.016	343.32	4	5	104.42	5	78.74	1.44
A4S4CA	0.035	309.77	4	5	104.42	4	78.59	1.16
A4I4NA	0.088	299.28	5	8	139.83	4	78.41	1.07
A4I2FA	0.033	272.28	5	5	91	3	72.62	0.19
A4I3FA	0.034	272.28	5	5	91	3	72.62	1.43
A4I4FA	0.039	272.28	5	5	91	3	72.62	1.49
A4I2MA	0.16	284.31	5	6	100.23	4	78.56	0.72
A4I3MA	0.17	284.31	5	6	100.23	4	78.56	1.15
A4I4MA	0.16	284.31	5	6	100.23	4	78.56	1.21
A4I3TA	0.11	268.31	5	5	91	3	77.18	1.86
A4I4TA	0.1	268.31	5	5	91	3	77.18	1.88
A4I4CA	0.093	288.73	5	5	91	3	77.35	1.94
A4IBKA	0.059	330.77	5	6	108.07	5	87.42	1.28
A4I2ANA	0.11	326.35	5	7	117.3	6	88.64	1.84
A4I3ANA	0.098	326.35	5	7	117.3	6	88.64	1.66
A4IA4NA	0.095	326.35	5	7	117.3	6	88.64	1.5

MW: molecular weight; Da: Dalton; HBD: hydrogen bond donor count; HBA: hydrogen bond acceptor count; TPSA: topological polar surface area; Rot: number of rotatable bonds; MR: molar refractivity (cm^3/mol); JCLogP: Jchem log partition coefficient.

Table 3.4-2: Physicochemical properties of *p*-amino benzamide series

code	Solubility (g/Lts)	MW (Da)	HBD	HBA	TPSA (Å^2)	Rot	MR	JCLogP
4A4TON	0.07	226.27	3	3	55.12	2	69.88	2.75
4A3TON	0.08	226.27	3	3	55.12	2	69.88	2.75
4A2TON	0.09	226.27	3	3	55.12	2	69.88	2.75
4A4BRN	0.01	291.14	3	3	55.12	2	72.75	3
4A4CLN	0.04	246.69	3	3	55.12	2	69.95	2.84
4A3CLN	0.04	246.69	3	3	55.12	2	69.95	2.84
4A2CLN	0.05	246.69	3	3	55.12	2	69.95	2.84
4A4ASN	0.05	242.27	3	4	64.35	3	71.74	2.08
4A3ASN	0.06	242.27	3	4	64.35	3	71.74	2.08
4A2ASN	0.06	242.27	3	4	64.35	3	71.74	2.08
4A3NTN	0.02	257.24	3	6	103.95	3	71.61	2.18
4A4MBN	0.04	256.3	3	4	64.35	4	75.27	1.79
4A3MBN	0.04	256.3	3	4	64.35	4	75.27	1.79
4AANIN	0.19	212.25	3	3	55.12	2	65.06	2.24
4ANAPN	0.01	262.31	3	3	55.12	2	82.9	3.23
4AMORN	52.52	206.24	2	4	55.56	1	57.5	0.22
4APPN	0.66	281.35	2	4	49.57	2	83.9	2.18
4APBN	0.65	295.38	2	4	49.57	3	88.62	2.01

MW: molecular weight; HBD: hydrogen bond donor count; HBA: hydrogen bond acceptor count; TPSA: topological polar surface area; Rot: number of rotatable bonds; MR: molar refractivity (cm^3/mol); JCLogP: Jchem log partition coefficient.

Table 3.4-3: Physicochemical properties of *p*-(methylamino)-benzamide series

code	Solubility (g/Lts)	MW(Da)	HBD	HBA	TPSA (Å^2)	Rot	MR	JCLogP
M4APN	0.38	227.26	3	4	68.01	3	67.37	0.81
M3APN	0.35	227.26	3	4	68.01	3	67.37	0.75
M2APN	0.32	227.26	3	4	68.01	3	67.37	1.34
M4TON	0.0309	240.3	3	3	55.12	3	74.1	2.7
M3TON	0.0321	240.3	3	3	55.12	3	74.1	2.7
M2TON	0.0328	240.3	3	3	55.12	3	74.1	2.05
M4BRN	0.0328	305.17	3	3	55.12	3	76.97	2.96
M4CLN	0.0166	206.72	3	3	55.12	3	74.17	2.79
M3CLN	0.0172	206.72	3	3	55.12	3	74.17	2.79
M2CLN	0.0183	206.72	3	3	55.12	3	74.17	2
M4ASN	0.0398	256.3	3	4	64.35	4	75.95	2.03
M3ASN	0.0399	256.3	3	4	64.35	4	75.95	2.03
M2ASN	0.0425	256.3	3	4	64.35	4	75.95	1.88
M3NTN	0.0131	271.27	3	6	103.95	4	75.82	1.64
M2NTN	0.0145	271.27	3	6	103.95	4	75.82	1.1
M4MBN	0.0223	270.33	3	4	64.35	5	79.49	1.74

M3MBN	0.0223	270.33	3	4	64.35	5	79.49	1.74
M2MBN	0.0232	270.33	3	4	64.35	5	79.49	1.74
MANIN	0.094	226.27	3	3	55.12	3	69.28	2.19
MNAPN	0.0025	276.33	3	3	55.12	3	87.12	3.18
MPRON	0.7	192.26	3	3	55.12	4	57.59	1.05
MBUTN	0.41	206.28	3	3	55.12	5	62.22	1.5
MCPRN	1.11	190.24	3	3	55.12	3	55.22	0.64

MW: molecular weight; HBD: hydrogen bond donor count; HBA: hydrogen bond acceptor count; TPSA: topological polar surface area; Rot: number of rotatable bonds; MR: molar refractivity (cm^3/mol); JCLogP: Jchem log partition coefficient.

3.4.2.2 Toxicity prediction of the designed molecules

Toxicity risk assessment was carried using OSIRIS property explorer and represented in **Table 3.4-4, Table 3.4-5** and **Table 3.4-6** for Benzamidine, *p*-amino benzamide and *p*-(methylamino)-benzamide series respectively.

Benzamidine series: Benzamidine amide analogues with *m*-nitrocinnamoyl (A4ACNA), *p*-nitrophenyl (A4A4NA), *m*-nitrophenyl (A4A3NA), moieties were predicted to have high risk of mutagenicity while molecules with *o*-fluorophenyl (A4A2FA), *o*-nitro-phenyl (A4A2NA) substitution were found to have medium risk of mutagenicity. *m*-methoxy phenyl (A4A3MA) and *o*-fluoro phenyl (A4A2FA) substituted molecules were predicted to have high and medium risk of reproductive toxicity respectively. Drug likeliness score which takes into account physicochemical properties and groups most likely to be present in marketed drugs was positive for molecules A4A4MA, A4ABEA, A4ACPA, A4A2FA, A4A4fA, A4A2FuA and A4ABNA and these molecules also showed high drug score. Sulfonamide derivatives with *o*-toluoyl phenyl (A4S2TA) moiety were predicted to have high risk of mutagenicity while *o*-and *m*-nitrophenyl moieties (A4S2NA, A4S3NA) were found to have medium risk of mutagenicity. Sulfonamide derivatives did not showed any alert for tumorigenicity, skin irritation and reproductive toxicity. Molecule A4S4MA and A4S4CA showed greater drug score and remaining sulfonamide derivatives showed mild drug score. Urea derivatives with *p*-methoxyphenyl (A4I4MA) and *p*-toluoyl (A4I4TA) substitution were predicted to have high risk of mutagenicity while *m*-methoxyphenyl (A4I3MA) and *p*-nitrophenyl (A4I4NA) substituted molecules were predicted to have medium risk of mutagenicity. A4I4MA and A4I2ANA were predicted to pose mild risk of tumorigenicity and skin irritation respectively. *m*-methoxyphenyl (A4I3MA) and *m*-fluorophenyl (A4I3FA) substituted molecules were predicted to have high and medium risk of reproductive toxicity respectively. A4I2MA and A4I4CA showed good drug score and other molecules showed low to medium drug score.

Table 3.4-4: Predicted toxicity and drug score of Benzamidine series

Code	MUT	TUM	Irritant	REP	cLogP	LogS	Drug likeliness	Drug score
A4A4MA	1	1	1	1	1.34	-2.42	2.01	0.87
A4A3MA	1	1	1	0.6	1.34	-2.42	0.45	0.45
A4ACNA	0.6	1	1	1	1.44	-0.23	-9.7	0.26
A4ABEA	1	1	1	1	1.24	-2.36	2.85	0.91
A4ACPA	1	1	1	1	0.23	-1.74	3.67	0.95
A4A4NA	0.6	1	1	1	1.31	-2.86	-10.4	0.27
A4A2FA	0.8	1	1	1	1.5	-2.71	0.68	0.61
A4A3FA	1	1	1	0.8	1.5	-2.71	-2.15	0.4
A4A4FA	1	1	1	1	1.5	-2.71	1.34	0.82
A4A2FuA	1	1	1	1	0.55	-2.08	1.65	0.87
A4A3NA	0.6	1	1	1	1.31	-2.86	-4.63	0.27
A4A2NA	0.8	1	1	1	1.31	-2.86	-5.15	0.36
A4ABNA	1	1	1	1	1.44	-2.4	1.83	0.87
A4S4MA	1	1	1	1	0.59	-2.24	2.39	0.88
A4S2NA	0.8	1	1	1	0.57	-2.68	-5.09	0.36
A4S3NA	0.8	1	1	1	0.57	-2.68	-6.83	0.36
A4S4NA	1	1	1	1	0.57	-2.68	-6.65	0.45
A4S2TA	0.6	1	1	1	1.01	-2.56	1.89	0.51
A4S3TA	1	1	1	1	1.01	-2.56	-1.92	0.52
A4S4TA	1	1	1	1	1.01	-2.56	-4.29	0.56
A4S3TFA	1	1	1	1	1.46	-30	-8.73	0.43
A4S4CA	1	1	1	1	1.31	-2.95	3.08	0.87
A4I4NA	0.8	1	1	1	1.2	-3.37	-12	0.35
A4I2FA	1	1	1	1	1.39	-3.23	0.26	0.69
A4I3FA	1	1	1	0.8	1.39	-3.23	-1.25	0.43
A4I4FA	1	1	1	1	1.39	-3.23	0	0.67
A4I2MA	1	1	1	1	1.23	-2.93	1.84	0.84
A4I3MA	0.8	1	1	0.6	1.23	-2.93	1.39	0.39
A4I4MA	0.6	0.8	1	1	1.23	-2.93	0.1	0.33
A4I3TA	1	1	1	1	1.65	-3.26	1.21	0.78
A4I4TA	0.6	1	1	1	1.65	-3.26	-0.22	0.38
A4I4CA	1	1	1	1	1.95	-3.65	2.23	0.8
A4IBKA	1	1	1	1	1.44	-4	2.37	0.77
A4I2ANA	1	1	0.8	1	1.74	-3.35	-2.52	0.37
A4I3ANA	1	1	1	1	1.74	-3.35	-3.37	0.44
A4IA4NA	1	1	1	1	1.74	-3.35	-2.23	0.47

MUT = mutagenesicity, TUM = tumorigenic, REP = reproductive effect. Values in column MUT, TUM, Irritant and REP represents 1= no indication, 0.8 = medium and 0.6 = high risk.

p-amino benzamide series: Molecules with *p*-toluoyl (4A4TON), *p*-anisyl (4A4ASN) and α-naphthyl (4ANAPN) moieties were predicted to have high risk of mutagenicity while *m*-pyridyl (4A3APN), *m*-anisyl (4A3ASN) and *m*-nitrophenyl (4A3NTN) substituted molecules

were predicted to have medium risk of mutagenicity. Molecule 4ANAPN was also predicted to have high risk of tumorigenicity, while 4A4ASN and 4A3NTN were shown to have mild risk of tumorigenicity. *o*-chloro substituted molecule 4A2CLN was predicted to have mild risk of skin irritation while molecules with 3-methoxy benzyl (4A3MBN) and *m*-anisyl (4A3ASN) substitution were predicted to have high risk of reproductive toxicity. Drug likeliness score which takes into account physicochemical properties and groups more likely to be present in marketed drugs, were positive except for 4A3NTN. Many molecules showed good drug score except for molecules with high risk of toxicity.

Table 3.4-5: Predicted Toxicity and drug score of *p*-amino benzamide series

Code	MUT	TUM	Irritant	REP	cLogP	LogS	Drug likeliness	Drug score
4A4APA	1	1	1	1	1.27	-2.41	1.54	0.86
4A3APA	0.8	1	1	1	1.27	-2.41	1.97	0.71
4A2APN	1	1	1	1	1.74	-2.93	2.51	0.88
4A4TON	0.6	1	1	1	2.66	-3.55	0.6	0.42
4A3TON	1	1	1	1	2.66	-3.55	1.6	0.78
4A2TON	1	1	1	1	2.66	-3.55	2.24	0.81
4A4BRN	1	1	1	1	3.04	-4.04	0.01	0.59
4A4CLN	1	1	1	1	2.96	-3.94	3	0.78
4A3CLN	1	1	1	1	2.96	-3.94	2.07	0.76
4A2CLN	1	1	0.8	1	2.96	-3.94	2.37	0.61
4A4ASN	0.6	0.8	1	1	2.24	-3.22	0.84	0.36
4A3ASN	0.8	1	1	0.6	2.24	-3.22	1.75	0.39
4A2ASN	1	1	1	1	2.24	-3.22	2.24	0.84
4A3NTN	0.8	0.8	1	1	2.21	-3.66	-3.32	0.28
4A4MBN	1	1	1	1	2.03	-3.04	3.17	0.87
4A3MBN	1	1	1	0.6	2.03	-3.04	2.06	0.5
4AANIN	1	1	1	1	2.34	-3.2	0.32	0.7
4ANAPN	0.6	0.6	1	1	3.53	-4.81	1.47	0.22
4AMORN	1	1	1	1	0.9	-1.4	3.36	0.95
4APPN	1	1	1	1	2.28	-2.75	8.74	0.89
4APBN	1	1	1	1	2.39	-2.23	8.31	0.9

MUT = mutagenesicity, TUM = tumorigenic, REP = reproductive effect. Values in column MUT, TUM, Irritant and REP represents 1= no indication, 0.8 = medium and 0.6 = high risk.

***p*-(methylamino)-benzamide series:** Similar to *p*-amino benzamide series, molecules with *p*-toluoyl (M4TON), *p*-anisyl (M4ASN), *o*-nitro-phenyl (M2NTN) and α-naphthyl (MNAPN) moieties were predicted to have high risk of mutagenicity. While molecules with *m*-pyridyl (M3APN), *m*-anisyl (M3ASN) and *m*-nitro-phenyl (M3NTN) substitution were found to have medium risk of mutagenicity. Molecule MNAPN was also predicted to have high risk of

tumorigenicity while M4ASN and M3NTN were shown to have mild risk of tumorigenicity. All the molecules were predicted to be devoid of skin irritation property. Molecule with 3-methoxy benzyl substitution (M3MBN) was predicted to have high risk of reproductive toxicity. Drug likeliness score was in negative values mainly due to less likelihood of benzylamine presence in marketed molecules. Due to low drug likeness molecules showed drug score below 0.6 and molecules with high risk of toxicity were scored below 0.3.

Table 3.4-6: Predicted Toxicity and drug score of *p*-(methylamino)-benzamide series

Code	MUT	TUM	Irritant	REP	cLogP	LogS	Drug likeliness	Drug score
M4APN	1	1	1	1	0.83	-2.29	-2.7	0.5
M3APN	0.8	1	1	1	0.83	-2.29	-2.26	0.41
M2APN	1	1	1	1	1.3	-2.82	-1.67	0.53
M4TON	0.6	1	1	1	2.22	-3.43	-3.59	0.27
M3TON	1	1	1	1	2.22	-3.43	-2.6	0.46
M2TON	1	1	1	1	2.22	-3.43	-1.95	0.49
M4BRN	1	1	1	1	2.6	-3.92	-4.18	0.41
M4CLN	1	1	1	1	2.52	-3.82	-1.14	0.51
M3CLN	1	1	1	1	2.52	-3.82	-2.12	0.46
M2CLN	1	1	1	1	2.52	-3.82	-1.81	0.38
M4ASN	0.6	0.8	1	1	1.8	-3.11	-3.29	0.22
M3ASN	0.8	1	1	1	1.8	-3.11	-2.42	0.23
M2ASN	1	1	1	1	1.8	-3.11	-1.93	0.5
M3NTN	0.8	0.8	1	1	1.77	-3.55	-7.52	0.28
M2NTN	0.6	1	1	1	1.77	-3.55	-9.19	0.26
M4MBN	1	1	1	1	1.59	-2.92	-0.93	0.58
M3MBN	1	1	1	0.6	1.59	-2.92	-2.09	0.3
M2MBN	1	1	1	1	1.59	-2.92	-1.66	0.52
MANIN	1	1	1	1	1.9	-3.09	-3.9	0.46
MNAPN	0.6	0.6	1	1	3.08	-4.69	-2.74	0.14
MPRON	1	1	1	1	1.22	-2.15	-1.76	0.54
MBUTN	1	1	1	1	1.69	-2.42	-3.67	0.48
MCPRN	1	1	1	1	0.96	-2.35	-1.28	0.58

MUT = mutagenesicity, TUM = tumorigenic, REP = reproductive effect. Values in column representing MUT, TUM, Irritant and REP represents 1= no indication, 0.8 = medium risk and 0.6 = high risk.

3.4.2.3 Docking scores and binding pattern of designed molecules

PAD4 crystal structure 1WDA was used for docking after mutation of Ala645 to Cys645. Docking study was performed using Autodock, Glide SP, Glide XP and FRED algorithm after validation using cocrystalised ligand BAA.

Benzamidine series: Binding scores of Benzamidine Series are represented in **Table 3.4-7**.

Table 3.4-7: In silico screening results of Benzamidine series

| code	Autodock	Glide SP		Glide XP		MM-GBSA	FRED
	Score	Score	Energy	Score	Energy	Energy	chemscore3
A4A4MA	-5.60	-8.26	-35.98	-0.69	-18.46	-21.67	-32.29
A4A3MA	-6.02	-8.17	-35.17	-2.48	-26.52	-22.55	-35.17
A4ACNA	-6.58	-7.82	-35.93	ND	ND	ND	-24.42
A4ABEA	-6.35	ND	ND	ND	ND	ND	-46.28
A4ACPA	-5.97	-7.92	-31.42	-1.10	-18.59	-30.95	-46.25
A4A4NA	-5.77	-8.20	-36.27	ND	ND	ND	-24.55
A4A2FA	-6.10	-8.45	-35.86	-2.73	-15.79	-37.61	-36.23
A4A3FA	-6.39	-8.30	-35.33	-2.26	-19.36	-31.81	-42.98
A4A4FA	-6.32	-8.21	-34.79	-2.58	-23.05	-23.21	-35.28
A4A2FuA	-5.78	-7.78	-28.37	ND	ND	ND	-39.12
A4A3NA	-6.89	-7.76	-31.79	ND	ND	ND	-36.40
A4A2NA	-5.70	-8.54	-38.80	-2.59	-25.99	-29.33	-38.72
A4ABNA	-5.84	-8.23	-34.31	-1.86	-22.23	-23.52	-41.75
A4S4MA	-6.20	-8.49	-40.29	-2.32	-19.30	-39.44	-42.25
A4S2NA	-5.97	-7.89	-41.79	-1.67	-27.58	-29.09	-39.17
A4S3NA	-5.58	-8.53	-40.77	ND	ND	ND	-39.02
A4S4NA	-6.86	-8.36	-39.38	ND	ND	ND	-42.19
A4S2TA	-5.99	-8.61	-39.53	0.72	-25.84	-16.01	-49.36
A4S3TA	-6.61	-8.45	-36.94	-1.33	-22.79	-41.03	-48.85
A4S4TA	-6.80	-8.52	-38.90	-2.85	-17.75	-36.97	-38.12
A4S3TFA	-6.33	-8.48	-40.66	-1.56	-26.39	-21.96	-40.07
A4S4CA	-6.76	-7.97	-34.73	-3.01	-21.34	-24.29	-37.71
A4I4NA	-6.36	-7.58	-33.13	ND	ND	ND	-35.17
A4I2FA	-6.20	-7.77	-33.25	-2.32	-18.24	-35.10	-45.38
A4I3FA	-6.38	-7.80	-32.76	-2.15	-16.60	-23.87	-39.32
A4I4FA	-6.29	-7.79	-32.85	-1.64	-16.88	-27.21	-39.03
A4I2MA	-6.06	-8.24	-38.87	-2.01	-19.95	-40.00	-42.35
A4I3MA	-6.22	-6.41	-39.29	ND	ND	ND	-39.26
A4I4MA	-6.37	-5.01	-39.52	-1.69	-21.52	-25.45	-38.35
A4I3TA	-6.68	-7.70	-33.49	-1.98	-21.09	-39.73	-41.01
A4I4TA	-6.58	-7.64	-32.88	-2.14	-18.70	-27.22	-48.99
A4I4CA	-5.52	-7.38	-29.39	-2.11	-19.73	-31.13	-47.40
A4IBKA	-5.53	-7.40	-33.84	-1.44	-25.83	-35.99	-31.37
A4I2ANA	-6.66	-8.06	-42.00	-2.07	-17.61	-40.59	-38.16
A4I3ANA	-6.20	-6.82	-36.43	ND	ND	ND	-46.06
A4IA4NA	-5.16	-7.21	-33.17	-0.88	-26.11	-44.62	-32.98

All values are expressed in Kcal/mol; ND: not docked; Energy: binding energy

No significant differences in Autodock, Glide XP and FRED docking scores were observed with amide, sulfonamide and urea linker, however scores for urea linker containing derivatives were relatively low in Glide SP docking. After detailed visual analysis it was observed that some of the derivatives bind differently than the substrate BAA. In case of molecules A4ABEA and A4A2NA of Benzamidine amide series it was observed that amidine was outside the cavity in Autodock and in A4ACPA and A4A2FUA, amidine moiety was outside the cavity in Glide SP while maintaining hydrophobic interaction and hydrogen bond interaction with ARG374. In FRED, molecules A4ABNA, A4A3TA, A4ATCA, A4A4TA, A4A2TA, A4A2CA and A4A3TA did not form hydrogen bonding interaction with Arg374 (**Figure 3.4-1**).
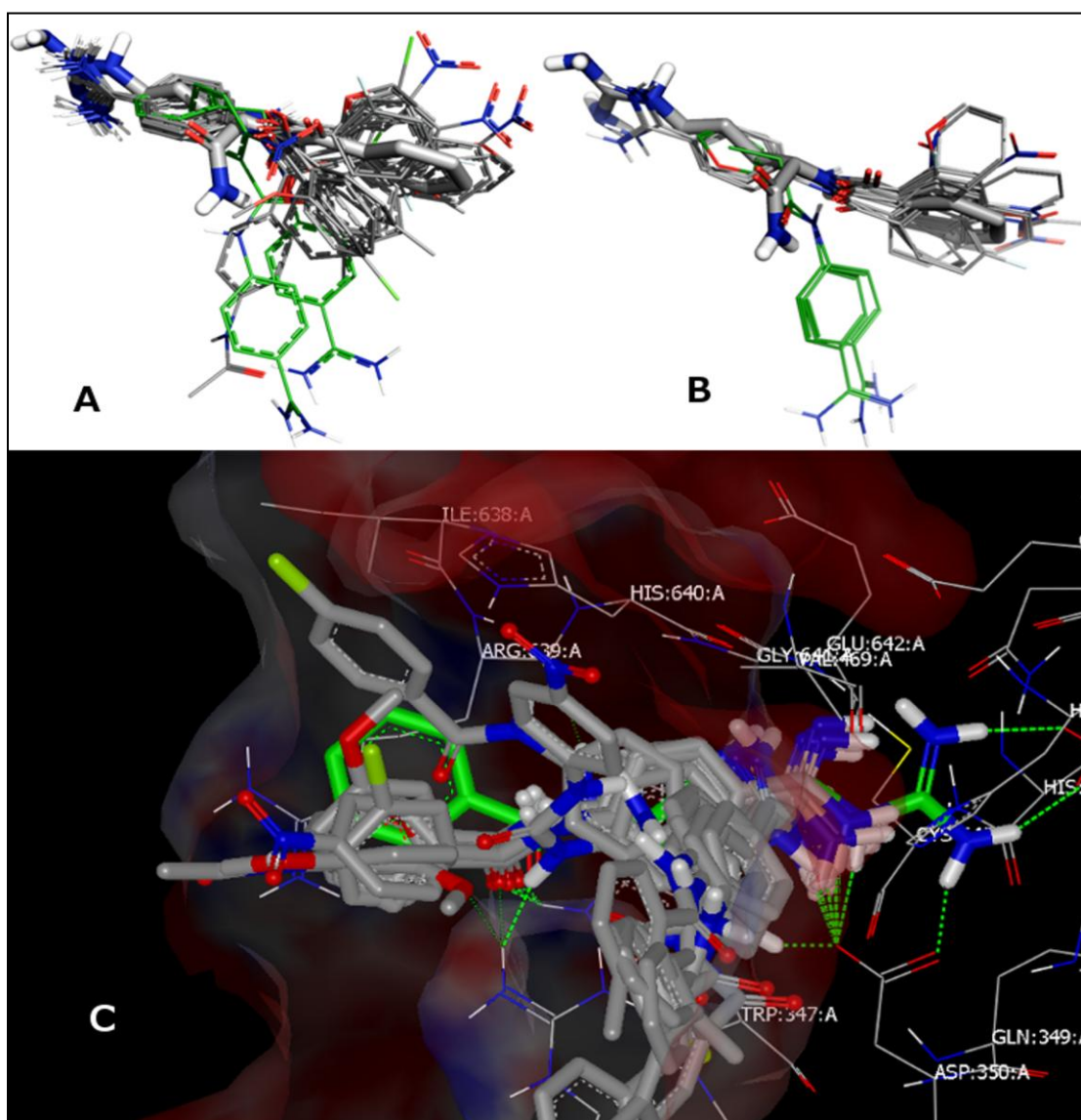
Figure 3.4-1: Binding poses of Benzamidine amide derivatives. A: Autodock binding pose; B: Glide SP binding pose (substrate BAA represented as stick model, designed ligands as wireframe model. Green wireframe are the ligands with amidine moiety outside the cavity); C: FRED binding pose (Dotted line represents hydrogen bonds)

In case of Benzamidine sulfonamide series, it was observed that in molecules A4S3TFA, A4S4AA, A4S4CA, A4S4TA and A4SBA, the amidine was outside the cavity in Autodock and in AS4MA, A4S2TA, A4S3TA, A4S4TA and A4SBA also the amidine moiety was outside the cavity in Glide SP while maintaining hydrophobic interaction and hydrogen bond interaction with Arg374. In FRED docking all the poses maintained hydrogen bonding interaction with Asp350 and Arg374 and poses were similar to substrate BAA **(Figure 3.4-2)** however it lacked hydrogen bonding interaction with carbonyl oxygen of Arg639.
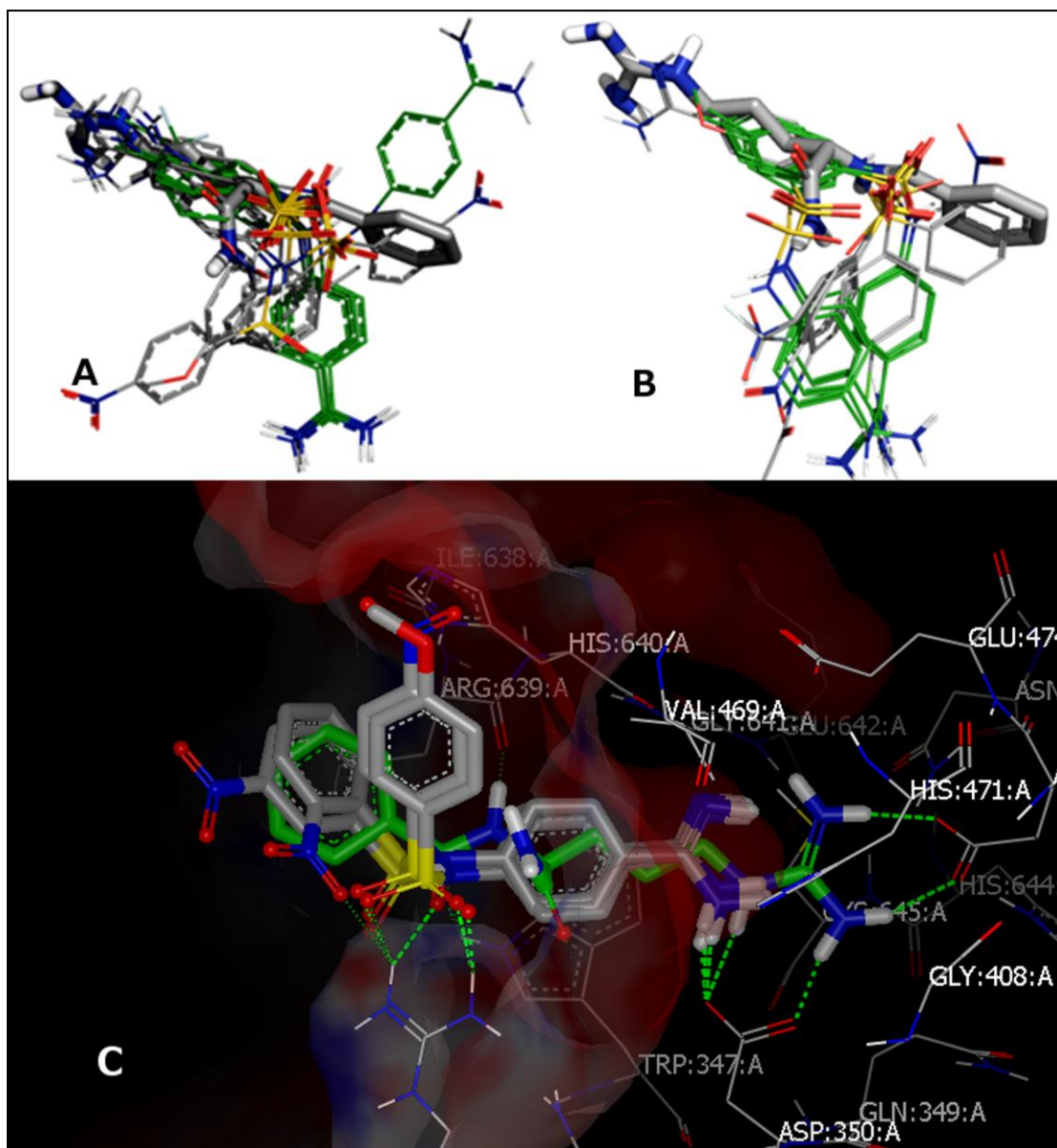
Figure 3.4-2: Binding poses of Benzamidine sulfonamide derivatives. A: Autodock binding pose; B: Glide SP binding pose (substrate BAA represented as stick model, designed ligands as wireframe model. Green wireframe are the ligands with amidine moiety outside the cavity); C: FRED binding pose (Dotted line represents hydrogen bonds)

In case of Benzamidine urea series, it was observed that in molecules A4I2ANA, A4I2MA, A4I2TA, A4I3ANA, A4I3NA and A4IA4NA, the amidine was outside the cavity in Autodock and in Glide SP docking all the molecule exhibited docking poses with amidine group outside the cavity but maintained hydrophobic interaction and hydrogen bond interaction with Arg374 or Arg639. In FRED docking all the poses maintained H bond interaction with Asp350 and hydrophobic interaction but failed to interact with Arg374 or Arg639 (**Figure 3.4-3**).
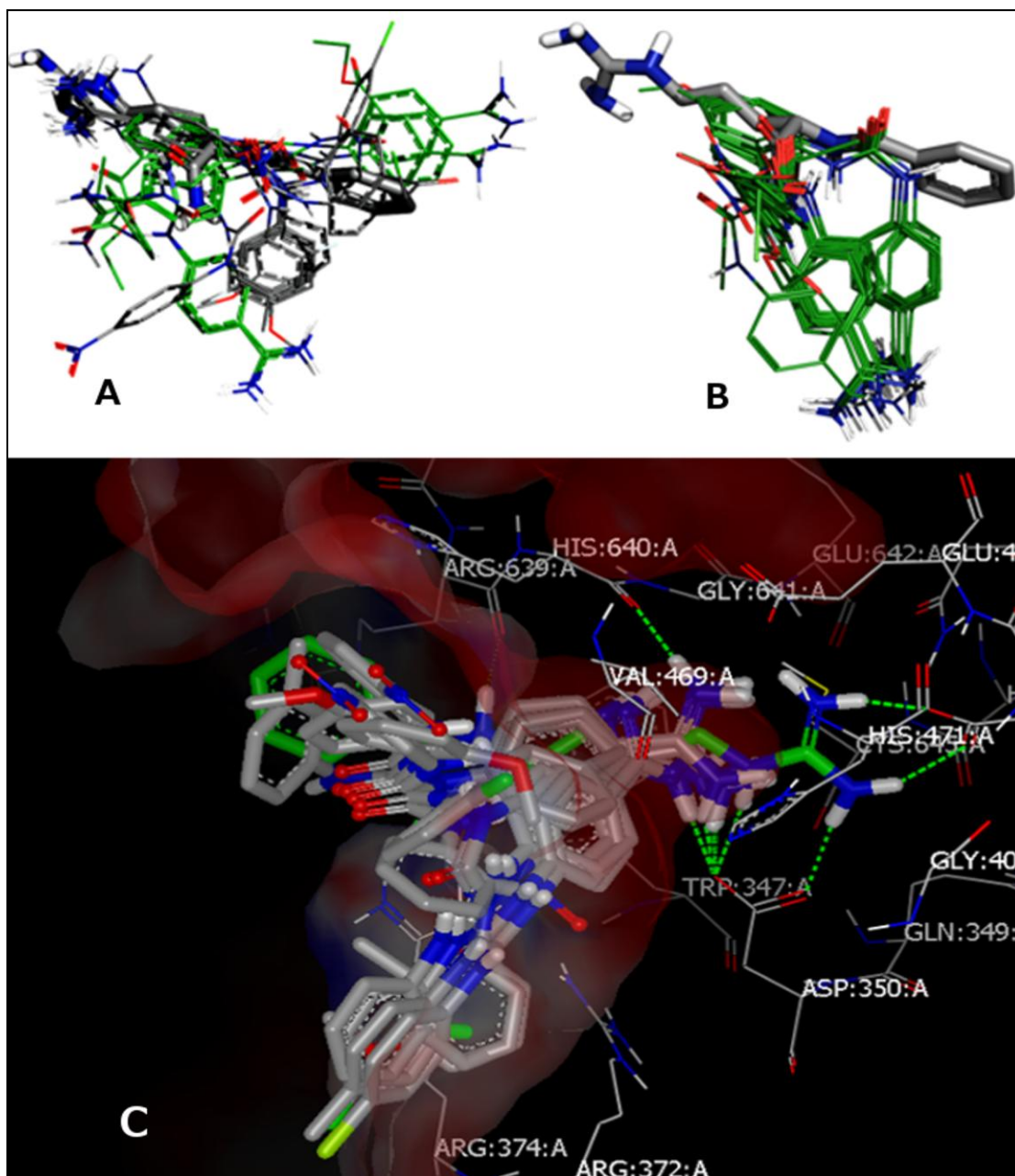
Figure 3.4-3: Binding poses of Benzamidine urea derivatives. A: Autodock binding pose; B: Glide SP binding pose (substrate BAA represented as stick model, designed ligands as wireframe model. Green wireframe are the ligands with amidine moiety outside the cavity); C: FRED binding pose (Dotted line represents hydrogen bonds)

p-amino benzamide series: Binding scores of *p*-amino benzamide series are represented in Table 3.4-8. Autodock scores varied between -6.95 and -4.62 kcal/mol. Glide SP scores varied between -5.83 and -4.03 kcal/mol. Glide XP score varied between -5.54 and -1.11 kcal/mol and FRED chemscore3 varied between -48.44 and -35.17 kcal/mol.

Table 3.4-8: In silico screening results of *p*-amino benzamide series

Code	Autodock	Glide SP		Glide XP		MM-GBSA	FRED
	Score	Score	Energy	Score	Energy	Energy	chemscore3
4A4APN	-6.01	-4.54	-29.00	-2.54	-23.98	-26.81	-40.93
4A3APN	-6.04	-5.09	-32.56	-4.77	-22.92	-42.16	-39.36
4A2APN	-6.40	-5.83	-32.00	-5.43	-23.92	-41.15	-45.16
4A4TON	-6.37	-4.53	-29.03	-3.56	-22.91	-28.52	-46.60
4A3TON	-6.60	-4.94	-30.48	-2.72	-23.49	-27.97	-43.60
4A2TON	-6.47	-4.81	-30.77	-2.74	-24.28	-27.66	-43.76
4A4BRN	ND	-4.53	-29.78	-3.95	-24.75	-19.93	-42.93
4A4CLN	-6.35	-4.62	-29.61	-3.77	-20.31	-29.27	-42.49
4A3CLN	-6.64	-4.75	-30.83	-2.77	-25.04	-31.59	-47.85
4A2CLN	-6.57	-4.82	-31.75	-3.50	-23.89	-43.03	-45.40
4A4ASN	-6.07	-4.79	-29.85	-2.40	-23.18	-28.34	-45.18
4A3ASN	-6.56	-4.56	-30.40	-3.48	-24.65	-28.80	-46.18
4A2ASN	-6.22	-5.48	-33.58	-3.56	-25.97	-35.87	-42.66
4A3NTN	-6.82	-4.85	-30.57	ND	ND	ND	-45.16
4A4MBN	-6.26	-4.21	-32.67	-2.77	-25.91	-28.20	-47.04
4A3MBN	-6.63	-4.73	-33.24	-2.90	-28.01	-36.12	-46.44
4AANIN	-6.26	-4.63	-30.22	-4.17	-29.59	-36.08	-42.93
4ANAPN	-6.95	-5.11	-32.79	-3.17	-27.67	-30.02	-48.44
4AMORN	-4.62	-3.90	-25.03	-3.29	-21.29	-17.45	-44.50
4APPN	-5.46	-4.03	-32.20	ND	ND	ND	-35.17
4APBN	-5.71	-4.31	-35.31	-1.14	-30.69	-23.84	-40.24

All values are expressed in Kcal/mol; ND: not docked; Energy: binding energy

On visualizing binding pattern, it was found that top ranking poses generated by Autodock satisfied all the pharmacophoric requirements except for 4APBN and 4AMORN which made slightly tilted orientation thereby missing hydrogen bond interaction with Asp350. In Glide SP docking also, molecules 4AMORN, 4APPN, and 4APBN made slightly tilted orientation there by missing hydrogen bond interaction with Asp350. Molecules 4A3APN, 4A2APN, 4A3TON, 4A2TON, 4A3CLN, 4A2CLN, 4A3ASN, 4A3NTN, 4AANIN and 4ANAPN had amine moiety outside the cavity while maintaining hydrophobic and hydrogen bonding interaction with Arg374 and Arg639. FRED docking pose of molecules occupied all the pharmacophoric sites with hydrogen bond interaction with Asp350 only. All the

molecules showed pi-pi stacking interaction with Trp347 residue and hydrogen bond interaction with Arg374 and Arg639 (**Figure 3.4-4**).
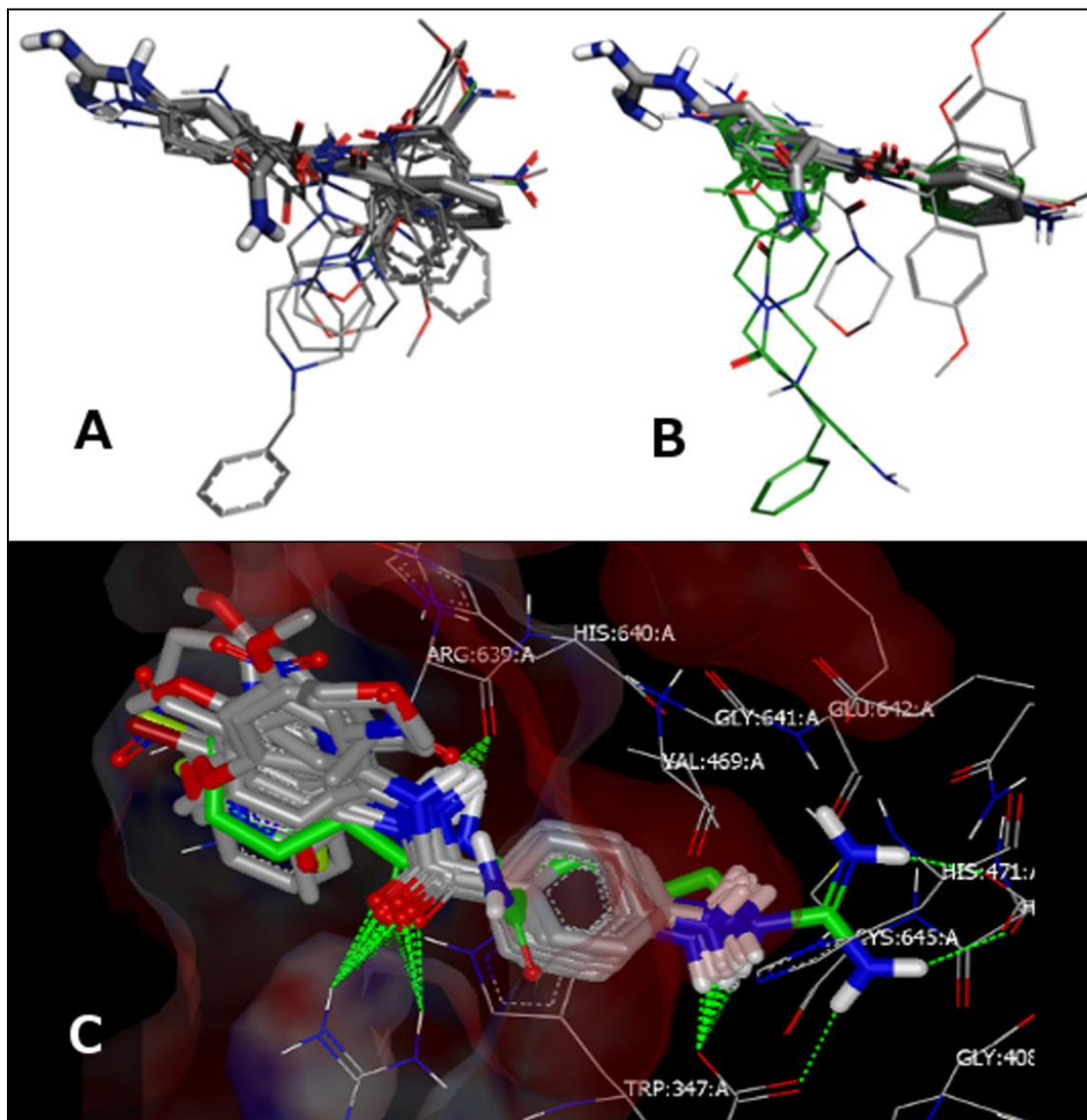
Figure 3.4-4: Binding poses of p-amino-Benzamide derivatives. A: Autodock binding pose; B: Glide SP binding pose (substrate BAA represented as stick model, designed ligands as wireframe model); C: FRED binding pose (Dotted line represents hydrogen bonds)
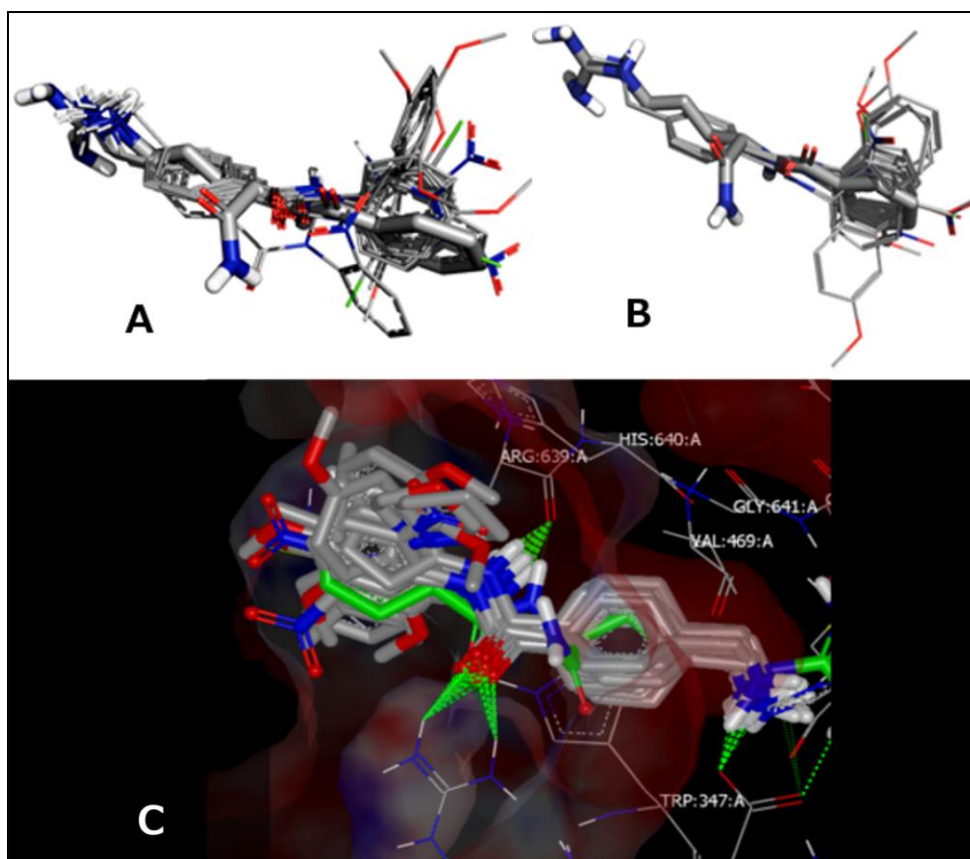
***p*-(methylamino)-benzamide series:** Binding scores of *p*-(methylamino)-benzamide series are represented in **Table 3.4-9**. Autodock scores varied from -8.68 to -7.40 kcal/mol. Glide SP scores ranged between -7.72 to -7.07 kcal/mol. Glide XP score varied from -6.53 to -1.53 kcal/mol while FRED chemscore3 varied between -48.68 to -33.33 kcal/mol. The Autodock scores and Glide SP scores were comparatively less deviating than Glide XP, FRED chemscore3 and free energy of binding calculated by MM-GBSA method.

Table 3.4-9: In silico screening results of *p*-(methylamino)-benzamide series

Code	Autodock	Glide SP		Glide XP		MM-GBSA	FRED
	Score	Score	Energy	Score	Energy	Energy	chemscore3
M4APN	-7.80	-7.04	-35.14	-6.53	-35.74	-41.41	-36.88
M3APN	-7.75	-7.04	-35.44	-4.65	-22.52	-43.18	-38.98
M2APN	-8.06	-6.97	-35.12	-5.34	-22.50	-39.55	-44.56
M4TON	-8.07	-7.10	-35.94	-5.61	-14.27	-43.23	-33.33
M3TON	-8.30	-7.21	-36.72	ND	ND	ND	-39.35
M2TON	-8.21	-7.12	-36.75	-1.53	-23.41	-36.98	-43.43
M4BRN	ND	-6.85	-35.96	-3.63	-19.44	-20.70	-39.93
M4CLN	-8.09	-7.02	-35.73	-3.55	-18.62	-30.04	-40.61
M3CLN	-8.30	-7.20	-36.87	-5.04	-15.48	-47.18	-43.00
M2CLN	-8.23	-7.20	-37.70	-1.43	-20.48	-26.12	-40.97
M4ASN	-7.77	-7.09	-36.38	-5.23	-16.19	-43.13	-39.18
M3ASN	-8.29	-7.05	-36.22	ND	ND	ND	-39.02
M2ASN	-7.91	-7.72	-39.67	-4.30	-20.34	-39.72	-41.80
M3NTN	-8.20	-7.07	-37.91	ND	ND	ND	-33.89
M2NTN	-8.11	-7.42	-39.18	ND	ND	ND	-37.87
M4MBN	-7.98	-6.88	-38.16	-5.19	-14.09	-49.59	-39.95
M3MBN	-8.46	-6.94	-39.39	-5.27	-20.21	-51.58	-47.10
M2MBN	-7.44	-6.78	-37.87	-4.40	-18.34	-46.52	-48.68
MANIN	-8.03	-7.07	-35.05	-3.54	-20.24	-37.05	-38.80
MNAPN	-8.68	-7.09	-38.03	-1.61	-22.56	-29.53	-33.88
MPRON	-7.40	ND	ND	ND	ND	ND	-40.15
MBUTN	-7.51	-6.75	-33.89	-5.49	-10.12	-41.25	-43.30
MCPRN	-7.52	-6.84	-32.79	-3.59	-19.59	-33.79	-40.23

All values are expressed in Kcal/mol; ND: not docked; Energy: binding energy

All pharmacophoric sites were occupied by all the molecules of p-(methylamino)-benzamide series. Binding pattern followed similar trend in all docking programs and is represented in **Figure 3.4-5**. Derivatives showed hydrogen bonding interaction with Arg374. Arg639 and Asp350 and hydrophobic interaction with Val469 and Trp347 residues of PAD4. **Figure 3.4-6** illustrates the interaction pattern of active molecule M2MBN with PAD4. A hydrogen bond interaction of ligand amino moiety with imidazole ring nitrogen of His471 (N-H, 22…N (3409), H471; 1.88 Å) and carbonyl oxygen of Asp350 (N-H, 32…CO, D350; 1.65 Å). Amide moiety of ligand formed hydrogen bonding interaction with Arg639 (N-H, 28…CO, R639; 2.00 Å) and Arg374 (C=O, 11…N-H, R374; 1.71 Å). Aromatic ring made hydrophobic interaction with Val469 and Trp347. While *o*-methoxy substituted benzyl aromatic ring was involved in pi-pi stacking interaction with imidazole ring of His640.

Figure 3.4-5: Binding poses of *p*-(methylamino)-Benzamide derivatives. A: Autodock binding pose; B: Glide SP binding pose (substrate BAA represented as stick model, designed ligands as wireframe model); C: FRED binding pose (Dotted line represents hydrogen bonds)
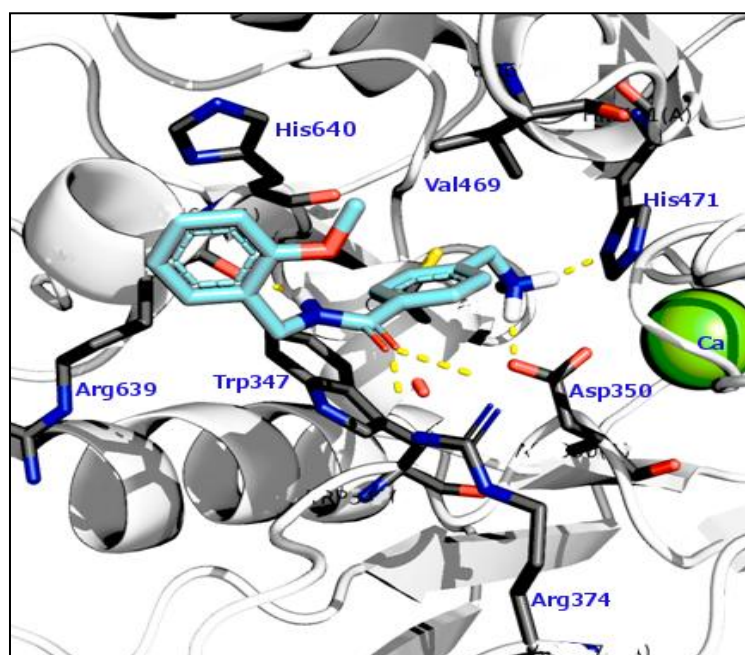

Figure 3.4-6: Binding interactions of M2MBN in PAD4 active site (energy minimization was carried out using Macromodel V10.0 [19] and plotted with Pymol [20]. M2MBN is represented in cyano. Active site residues of PAD4 are represented in black. Calcium ion is presented as green sphere and rest of protein as cartoon model. Yellow dotted line represents hydrogen bond).

3.4.3   References

[1]   JChem for excel 1.1.2, ChemAxon (http://www.chemaxon.com), **2009**.

[2]   Tetko, I.V.; Tanchuk, V.Y. Application of associative neural networks for prediction of lipophilicity in ALOGPS 2.1 program. *J Chem Inf Comput Sci.* **2002,** 42(5), 1136-1145.

[3]   Leeson, P. Drug discovery: Chemical beauty contest. *Nature.* **2012,** 481(7382), 455-456.

[4]   Veber, D.F.; Johnson, S.R.; Cheng, H.Y.; Smith, B.R.; Ward, K.W.; Kopple, K.D. Molecular properties that influence the oral bioavailability of drug candidates. *J Med Chem.* **2002,** 45(12), 2615-2623.

[5]   Sander, T.; Freyss, J.; von Korff, M.; Reich, J.R.; Rufener, C. OSIRIS, an entirely in-house developed drug discovery informatics system. *J Chem Inf Model.* **2009,** 49(2), 232-246.

[6]   Sander, T.L. Osiris Property Explorer (2001-2013). http://www.organic-chemistry.org/prog/peo/ (Accessed November 18 2013).

[7]   Arita, K.; Hashimoto, H.; Shimizu, T.; Nakashima, K.; Yamada, M.; Sato, M. Structural basis for Ca^{2+} induced activation of human PAD4. *Nat Struct Mol Biol.* **2004,** 11(8), 777-783.

[8]   Schrödinger suite 2009: Maestro, version 9.0, Schrödinger, LLC, New York, NY, **2009**.

[9]   Shelley, J.C.; Cholleti, A.; Frye, L.L.; Greenwood, J.R.; Timlin, M.R.; Uchimaya, M. Epik: a software program for pK a prediction and protonation state generation for drug-like molecules. *J Comput Mol Des.* **2007,** 21(12), 681-691.

[10] Marvin 5.2.6, ChemAxon (http://www.chemaxon.com), **2009**.

[11] Schrödinger suite 2009: Epik, version 2.0, Schrödinger, LLC, New York, NY, **2009**.

[12] Morris, G.M.; Huey, R.; Lindstrom, W.; Sanner, M.F.; Belew, R.K.; Goodsell, D.S.; Olson, A.J. AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J Comput Chem.* **2009,** 30(16), 2785-2791.

[13] Halgren, T.A.; Murphy, R.B.; Friesner, R.A.; Beard, H.S.; Frye, L.L.; Pollard, W.T.; Banks, J.L. Glide: a new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening. *J Med Chem.* **2004,** 47(7), 1750-1759.

[14] FRED version 2.2.5, OpenEye Scientific Software, Inc., NM, USA, **2010**.

[15] OMEGA version 2.3.2, OpenEye Scientific Software, Inc., NM, USA, **2010**.

[16] Hou, T.; Wang, J.; Li, Y.; Wang, W. Assessing the performance of the MM/PBSA and MM/GBSA methods. The accuracy of binding free energy calculations based on molecular dynamics simulations. *J Chem Inf Model.* **2010,** 51(1), 69-82.

[17] Schrödinger suite 2013: Prime MM-GBSA, version 3.0, Schrödinger, LLC, New York, NY, 2013.

[18] Ghose, A.K.; Viswanadhan, V.N.; Wendoloski, J.J. A knowledge-based approach in designing combinatorial or medicinal chemistry libraries for drug discovery. 1. A qualitative and quantitative characterization of known drug databases. *J Comb Chem.* **1999,** 1(1), 55-68.

[19] Schrödinger suite 2013-1: MacroModel, version 10.0, Schrödinger, LLC, New York, NY, **2013**.

[20] The PyMOL Molecular Graphics System, Version 1.6 Schrödinger, LLC. New York, NY, **2012**.

3.5 Synthesis of PAD4 Inhibitors

Based on different synthetic strategies, synthesis of designed molecules was carried out using two different schemes. All benzamidines were synthesized according to **Scheme 3.5-1** using 3/4-amino substituted benzonitriles. Free amino group of benzonitriles was reacted with different carboxylic acids or acyl chlorides, sulfonyl chlorides and isocynates to yield corresponding amides, sulfonamides and urea derivatives. Nitrile group was then converted to aldoxime followed by reduction to yield amidine as final product.
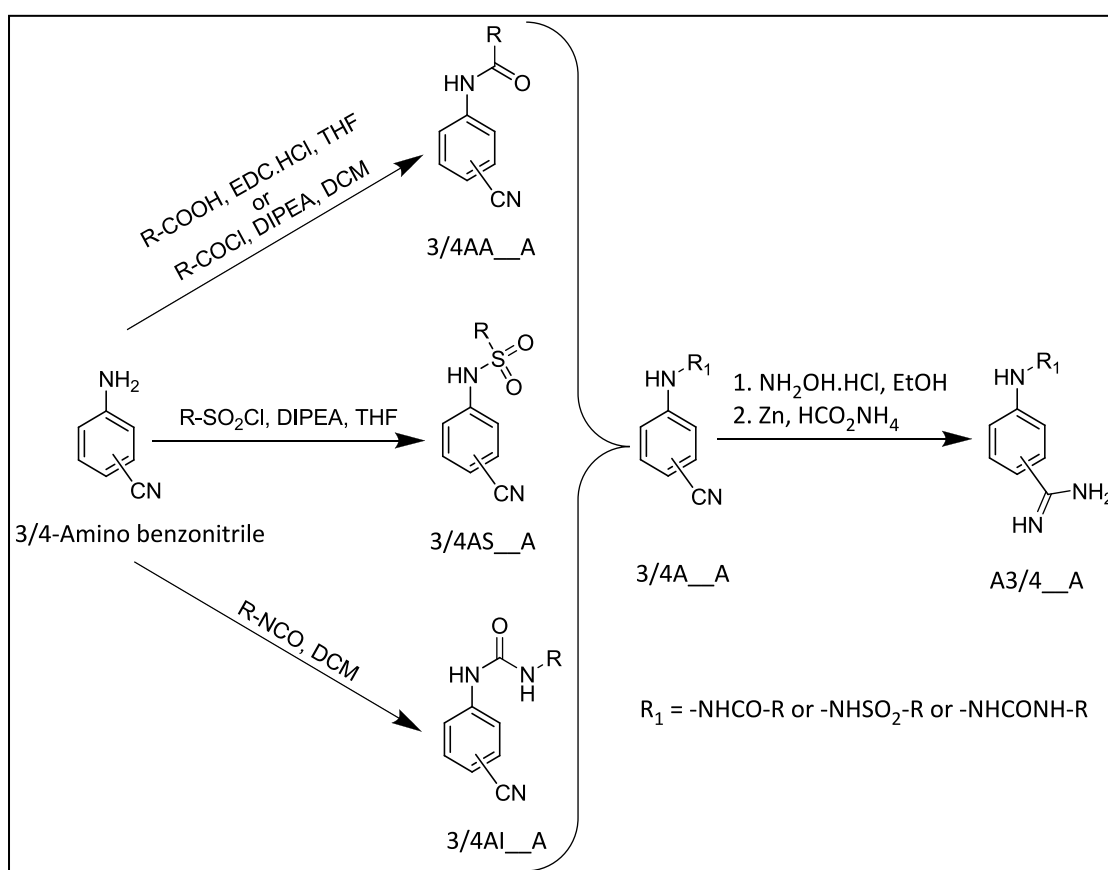
Scheme 3.5-1: Synthetic route of Benzamidine series

All benzamides were synthesized according to **Scheme 3.5-2** using starting material 4-amino-benzoic acid or 4-(aminomethyl) benzoic acid. Amino group was initially protected and then carboxylic acid moiety of starting material was made to react with different alkyl and aryl amines. Protection of amino group was then removed to produce final compound.
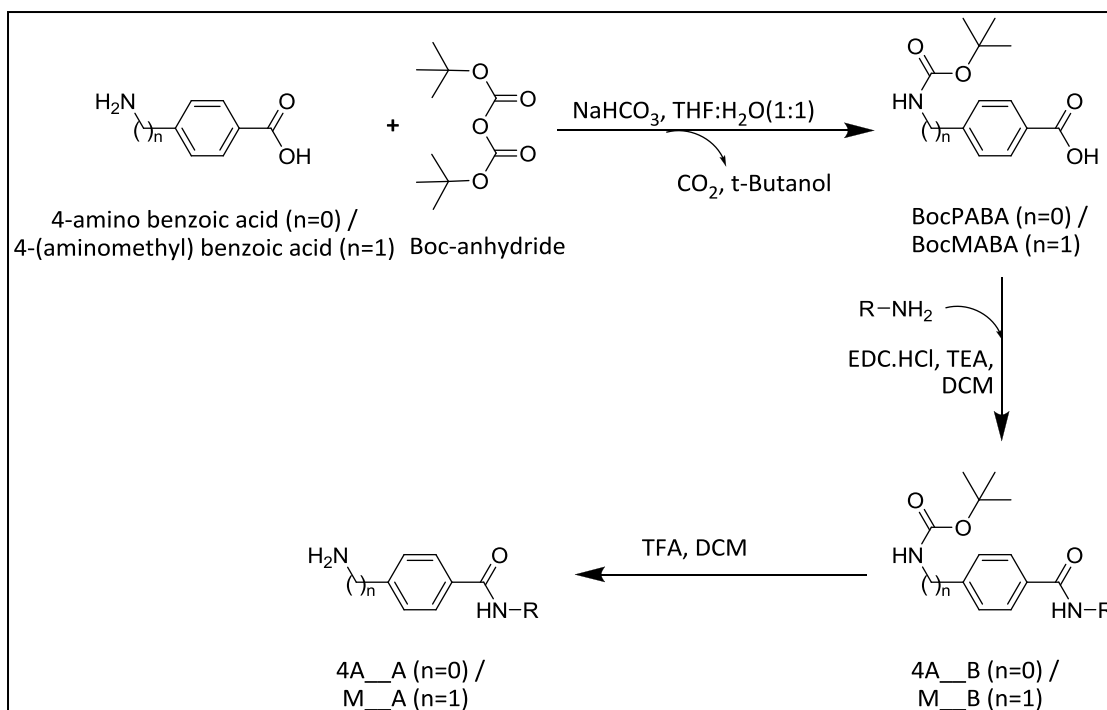
Scheme 3.5-2: Synthetic route of Benzamide series

3.5.1 Methodology

3.5.1.1 Synthesis of Benzamidines

Synthesis of N-(3/4-cyanophenyl) substituted-amides (4AA__A/3AA__A) from carboxylic acid [1]: To the ice cold solution of aromatic acid (5 mM) in 10 ml dry tetrahydrofuran (THF), N,N'-Diethylcarbodiimide hydrochloride (EDC.HCl) (1.05g, 5.5 mM) was added. The reaction mixture was allowed to stir for 30 minutes followed by addition of 4-amino benzonitrile or 3-amino benzonitrile (0.60g, 5.1 mM). Reaction mixture was stirred for 24 hrs at room temperature and progress was monitored with TLC using ethyl acetate, n-hexane (1:1) as solvent system. After completion of the reaction, solvent was evaporated by vacuum distillation. To the crude mass, 20 ml of chilled 0.1 N hydrochloric acid (HCl) was added and the product was extracted with ethyl acetate (2 X 10 ml). Combined organic layer was washed with 20 ml sodium bicarbonate solution (10%) followed by 20 ml of water. Organic phase was separated and dried over anhydrous sodium sulphate followed by evaporation to yield amide derivatives (4AA__A/ 3AA__A).

Synthesis of N-(3/4-cyanophenyl) substituted-amide (4AA__A/3AA__A) from acyl chlorides [2]: Acyl chloride (6 mM) was added to 10 ml of Dichloromethane (DCM) maintained at ice cold condition under stirring. To this, solution of 4-amino benzonitrile or 3-amino benzonitrile (0.60 g, 5.1 mM) and N,N-Diisopropylethylamine (DIPEA) (0.61 ml, 6 mM) in 5 ml DCM was added drop wise kept in ice bath. The reaction was stirred for 6 hrs. Reaction was

monitored by TLC using ethyl acetate, n-hexane (3:7) as solvent system. After completion of the reaction, the solvent was evaporated by vacuum distillation. 20 ml of ice cold 0.1 N HCl was then added to the crude mass and the product was extracted with ethyl acetate (2 X 10 ml). Combined organic layer was washed with 20 ml sodium bicarbonate solution (10%) followed by 20 ml of water. Organic phase was separated and dried over anhydrous sodium sulphate followed by evaporation to yield amide derivatives (3/4AA__A).

Synthesis of N-(3/4-cyanophenyl) substituted-sulfonamide (3/4AS__A) from sulfonyl chlorides [3]: To solution of sulfonyl chloride (6.4 mM) in 10 ml THF, 4-amino benzonitrile or 3-amino benzonitrile (0.71 g, 6 mM) and DIPEA (0.61 ml, 6 mM) was added. Reaction mixture was refluxed for 4-6 hrs. Progress of the reaction was monitored by TLC using ethyl acetate, n-hexane (3:7) as solvent system. After completion of the reaction, the solvent was evaporated by vacuum distillation. 20 ml of ice cold 0.1 N HCl was added to the reaction mass and the product was extracted with ethyl acetate (2 X 10 ml). Combined organic layer was washed with 20 ml sodium bicarbonate solution (10%) followed by 20 ml of water. Organic phase was separated and dried over anhydrous sodium sulphate followed by evaporation to yield sulfonamides from there corresponding amines (4AS__A/ 3AS__A).

Synthesis of 1-{(3/4)-cyanophenyl}-3-substituted urea (3/4AI__A) from isocynates [4]: To the solution of 4-amino benzonitrile or 3-amino benzonitrile (0.77 g, 6.5 mM) in DCM (15 ml) stirred at 0° C, was added isocynates (6 mM). Reaction mixture was stirred for 4-6 hrs. Progress of the reaction was monitored by TLC using ethyl acetate, n-hexane (1:2) as solvent system. After the completion of reaction, solid precipitated was filtered and washed with 20 ml of diethyl ether. Formed urea derivatives were used without further purification.

Synthesis of 1-{(3/4)-amidine phenyl}-3-substituted derivatives (A(3/4)___A) [1]: The solution of benzonitrile derivative synthesized above (3/4)A(A/S/I)__A (1 mM) was refluxed with hydroxylamine hydrochloride (0.33 g, 10 mM) and triethyl amine (TEA) (1.40 ml, 10 mM) in 15 ml methanol for 4-5 hrs. Reaction progress was monitored by TLC using ethyl acetate, n-hexane (1:1) as solvent system. After completion, zinc dust (0.65g, 10 mM), ammonium formate (0.63 g, 20 mM) and 20 ml methanol was added and refluxing was continued for another 8-12 hrs. Progress of the reaction was monitored by TLC using ethyl acetate, n-hexane (1:1) as solvent system. Methanol was distilled off after completion of the reaction. 100 ml water was added, amidines were extracted with ethyl acetate (2 X 50 ml). Organic phase was separated and dried over anhydrous sodium sulphate followed by evaporation to yield corresponding amidines (A(3/4)___A).

3.5.1.2 Synthesis of Benzamide derivatives

Synthesis of {4-(BOC-amino)/4-[(BOC-amino) methyl]} benzoic acid [5]: 4-amino benzoic acid (13.17 g, 100 mM) or 4-(aminomethyl) benzoic acid (15.12 g, 100 mM) and sodium bicarbonate (NaHCO$_3$) (9.24 g, 110 mM) were dissolved in 100 ml of THF: water solution (50%). To this, BOC anhydride (25.27 ml, 110 mM) was added with stirring. Stirring was continued for 20 hrs. Completion of the reaction was confirmed by TLC using ethyl acetate, n-hexane (1:1) as solvent system. Later to the reaction mass was diluted with 200 ml of ice cold water and acidified with concentrated HCl (~4 ml). The precipitate formed were filtered, washed with ice cold water till free from acid and dried at 60° C till weighed constant (~5 hrs) to yield Boc protected amine (BocPABA or BocMABA).

Synthesis of {4-(BOC-amino)/4-[(BOC-amino) methyl]}-*N*-substituted benzamide (4A__B/M__B) [1]: 6 mM of Boc-protected amino acid (BocPABA: 1.42 g, BocMABA: 1.51 g) was added to 10 ml ice cold dry THF followed by EDC.HCl (1.38 g, 7.2 mM) with stirring. Stirring was continued for 30 min and amine (7.2 mM) was added at the end. Stirring was then continued at room temperature for 24 hrs. Reaction progress was monitored with TLC using ethyl acetate, n-hexane (1:2) solvent system. After completion of the reaction, product was concentrated and diluted with 10 ml of ice cold 0.1 N HCl. Using ethyl acetate (2 X 10 ml) product was extracted and washed with 10 ml of sodium bicarbonate solution (10%) followed by 10 ml of water. Organic phase was then dried over anhydrous sodium sulphate and concentrated to yield amide derivatives (4A__B or M__B).

Synthesis of Synthesis of [4-(amino)/4-(aminomethyl)]-*N*-substituted benzamide (4A__N/M__N) [6]: To a solution of Boc-protected amine 4A__B or M__B (4 mM) in DCM (10 ml), was added trifluoro acetic acid (TFA) (2.45 ml, 32 mM) and stirred for 1 hrs at room temperature. Reaction was monitored with TLC using ethyl acetate, n-hexane (1:1) as solvent system. After completion of the reaction, excess TFA was removed by repeated vacuum distillation using 40 ml DCM. The residue left was taken in ice cold water (15 ml) and neutralized with sodium carbonate solution (20%). Solid separated was filtered and washed with cold water and dried in oven at 45° C.

3.5.2 Results and Discussion

3.5.2.1 Synthesis of Benzamidines

4-amino benzonitrile or 3-amino benzonitrile was coupled with various carboxylic acids and acyl chlorides. On reaction completion, *N,N*-diisopropylethylamine (DIPEA) and unreacted amine were removed by adding water and extracting product with ethyl acetate.

Organic phase was washed with dilute HCl followed by sodium bicarbonate solution to remove unreacted carboxylic acid as sodium salt. Ethyl acetate layer was then dried and concentrated to yield amide with yield ranged from 17.48 to 90.25% (**Table 3.5-1**). Product formation was confirmed by observation of characteristic NH stretching peak at around 3210 to 3410 cm^{-1}, CN stretching peak at 2227 to 2239 cm^{-1} and carbonyl stretching peak at 1651 to 1691 cm^{-1} in IR spectroscopy

Table 3.5-1: Isolated yield of N-3/(4-cyanophenyl) amides ((4AA__A/3AA__A)

Carboxylic acid/ acyl chloride used	*p*-Cyano		*m*-Cyano	
	Code	% Yield	Code	% Yield
4 Methoxy Benzoic acid	4AA4MA	26.16	3AA4MA	51.53
3 Methoxy Benzoic acid	4AA3MA	42.81	3AA3MA	66.60
2 Methoxy Benzoic acid	4AA2MA	NS	3AA2MA	51.53
3 Nitro Cinnamic Acid	4AACNA	46.37	3AACNA	NS
Phenyl Acetic Acid	4AABEA	57.56	3AABEA	NS
Cyclo Propionic Acid	4AACPA	90.22	3AACPA	NS
t-Cinnamic acid	4AATCA	NS	3AATCA	41.89
4 Acetamido Benzoic acid	4AAABA	NS	3AAABA	63.73
4 Nitro Benzoic acid	4AA4NA	47.90	3AA4NA	84.57
2 Fluoro Benzoic acid	4AA2FA	17.48	3AA2FA	51.84
3 Fluoro Benzoic acid	4AA3FA	29.97	3AA3FA	NS
4 Fluoro Benzoic acid	4AA4FA	97.41	3AA4FA	NS
2 Furoic acid	4AA2FUA	49.95	3AA2FUA	NS
3 Nitro Benzoic acid	4AA3NA	26.94	3AA3NA	NS
2 Nitro Benzoic acid	4AA2NA	47.90	3AA2NA	NS
Benzoic acid	4AABNA	49.50	3AABNA	NS
4 chloro benzoyl chloride	4AA4CA	NS	3AA4CA	77.92
3 chloro benzoyl chloride	4AA3CA	NS	3AA3CA	82.46
2 chloro benzoyl chloride	4AA2CA	NS	3AA2CA	90.25
2 Toluoyl Chloride	4AA2TA	NS	3AA2TA	71.95
4 Toluoyl Chloride	4AA4TA	NS	3AA4TA	74.07
3 Toluoyl Chloride	4AA3TA	NS	3AA3TA	84.65
Benzoyl chloride	4AABNA	NS	3AABNA	83.24

NS: not synthesized

Sulfonamide derivatives of 4-amino benzonitrile or 3-amino benzonitrile were prepared with different sulfonyl chlorides. The isolated percentage yield of sulfonamides ranged from 17.48 to 90.25 % (**Table 3.5-2**). Presence of sharp NH stretching at around 3169 - 3327 cm^{-1}, absence of O-H stretching peak, presence of CN stretching at 2233 to 2245 cm^{-1} and sulfonyl stretching peak corresponding to sulfonamide at 1174 to 1175 cm^{-1} confirmed the product formation.

Table 3.5-2 : Isolated yield of N-(3/4-cyanophenyl) sulfonamides ((4SA__A/3AS__A)

Sulfonyl chloride used	*p*-Cyano		*m*-Cyano	
	Code	% Yield	Code	% Yield
4 Methoxy Benzene Sulfonyl Chloride	4AS4MA	49.71	3AS4MA	76.30
2 Nitro Benzene Sulfonyl Chloride	4AS2NA	61.55	3AS2NA	62.10
3 Nitro Benzene Sulfonyl Chloride	42S3NA	57.70	32S3NA	75.28
4 Nitro Benzene Sulfonyl Chloride	4AS4NA	53.85	3AS4NA	58.80
2 Toluene Sulfonyl Chloride	4AS2TA	63.04	3AS2TA	NS
3 Toluene Sulfonyl Chloride	4AS3TA	51.41	3AS3TA	NS
4 Toluene Sulfonyl Chloride	4AS4TA	58.75	3AS4TA	NS
4 Acetamido Benzene Sulfonyl Chloride	4AS4AA	54.44	3AS4AA	NS
3 Trifluro Methyl Benzene Sulfonyl Chloride	4AS3TFA	60.27	3AS3TFA	NS
4 Chloro Benzene Sulfonyl Chloride	4AS4CA	51.81	3AS4CA	NS
Benzene Sulfonyl Chloride	4ASBA	57.43	3ASBA	NS

NS: not synthesized

Urea derivatives were synthesized by reacting different isocynates with 4/3-amino benzonitrile. Amine acts as a nucleophile and attacks on carbonyl carbon of isocynate, followed by internal rearrangement leading to urea derivatives which gets precipitated in the reaction mixture. Washing of the precipitate with ether yields practically pure product. Yield of the products are reported in **Table 3.5-3**.

**Table 3.5-3: Isolated yield of 1-(3/4-cyanophenyl)-3-substituted ureas ((4SA__A/3AS__A)**

Isocynate used	*p*-Cyano		*m*-Cyano	
	Code	% Yield	Code	% Yield
4 Nitro Phenylisocynate	4AI4NA	83.61	3AI4NA	54.92
3 Nitro Phenylisocynate	4AI3NA	0	3AI3NA	99.79
2 Nitro Phenylisocynate	4AI2NA	0	3AI2NA	0
2 Fluoro Phenylisocynate	4AI2FA	95.59	3AI2FA	85.54
3 Fluoro Phenylisocynate	4AI3FA	94.03	3AI3FA	83.58
4 Fluoro Phenylisocynate	4AI4FA	51.71	3AI4FA	98.60
2 Methoxy Phenylisocynate	4AI2MA	86.80	3AI2MA	64.85
3 Methoxy Phenylisocynate	4AI3MA	52.38	3AI3MA	0
4 Methoxy Phenylisocynate	4AI4MA	74.83	3AI4MA	72.33
2 tolyl Phenylisocynate	4AI2TA	0	3AI2TA	0
3 tolyl Phenylisocynate	4AI3TA	58.90	3AI3TA	87.55
4 tolyl Phenylisocynate	4AI4TA	97.10	3AI4TA	88.88
4 Chloro Phenylisocynate	4AI4CA	83.92	3AI4CA	84.65
4 (Chloro Acetyl) Phenylisocynate	4AIBKA	75.22	3AIBKA	NS
2-(EthoxyCarbonyl) Penylisocynate	4AI2ANA	49.14	3AI2ANA	NS
3-(EthoxyCarbonyl) Penylisocynate	4AI3ANA	99.57	3AI3ANA	NS

NS: not synthesized

Formation of the desired compounds was confirmed by IR spectroscopy. Compounds showed NH stretching at around 3321 to 3286 cm^{-1}, CN stretching at 2227 to 2238 cm^{-1} and

carbonyl stretching peak corresponding to amide at 1716 to 1633 cm^{-1}.

Benzonitriles were then condensed with hydroxylamine hydrochloride to form amidoxime. Amino group makes a nucleophilic attack on carbon of nitrile group and the intermediate amino alcohol rearranges to form amidoxime. After completion of the reaction as monitored by TLC (using ethyl acetate, n-hexane (1:1) as solvent system), amidoxime was reduced using zinc and ammonium formate (which acts as proton donor). Possible mechanism of reaction is illustrated in **Figure 3.5-1**.
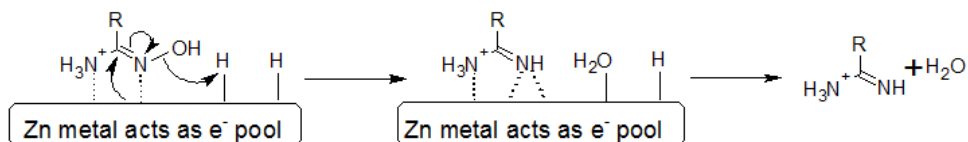
Figure 3.5-1: Possible reaction mechanism for reduction of amidoxime

Ammonium formamate protonates the substrate and forms a complex with zinc. Zinc donates electron to carbon, followed by elimination of hydroxyl group as water. Formed amidines were extracted with ethyl acetate and purified by column chromatography. Yields of *p*-benzamidines are tabulated in **Table 3.5-4Error! Reference source not found.**. *m*-benzamidines were not formed since the intermediates amidoximes failed to undergo reduction by zinc. Formation of the desired compounds was confirmed by presence of N-H stretch at 3325 to 3479 cm^{-1} in IR spectra, disappearance of CN stretching peak and characteristic peaks in ^{1}H NMR spectra, confirmed the product.

Table 3.5-4: Isolated yield of 1-{(3/4)-amidine phenyl}-3-substituted ureas (A(3/4)___A)
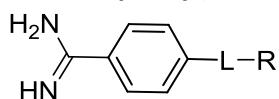
R	Code	% yield	R	Code	% yield
-*Ph*-4-OCH$_3$	A4A4MA	8	-*Ph* -3-CH$_3$	A4S3TA	1.97
-*Ph* -3-OCH$_3$	A4A3MA	15	-*Ph* -4-CH$_3$	A4S4TA	3.95
trans-Cinnamoyl-3-NO$_2$	A4ACNA	4	-*Ph* -3-CF$_3$	A4S3TFA	21.64
-CH$_2$-*Ph*	A4ABEA	16.66	-*Ph* -4-Cl	A4S4CA	3.69
-*cyclo*-Propyl	A4ACPA	13.33	-*Ph* -4-NO$_2$	A4I4NA	4.2
-*Ph* 4-NO$_2$	A4A4NA	7.14	-*Ph* -2-F	A4I2F	4.54
-*Ph* -2-F	A4A2FA	13.15	-*Ph* -3-F	A4I3F	3.12
-*Ph* -3-F	A4A3FA	31	-*Ph* -4-F	A4I4F	2.85
-*Ph* -4-F	A4A4FA	4.76	-*Ph* -2-OCH$_3$	A4I2MA	3.86
2-Furoyl	A4A2FuA	10.52	-*Ph* -3-OCH$_3$	A4I3MA	13.63
-*Ph* -3-NO$_3$	A4A3NA	11.7	-*Ph* -4-OCH$_3$	A4I4MA	2.43
-*Ph* -2-NO$_2$	A4A2NA	9.52	-*Ph* -3-CH$_3$	A4I3TA	2.7
-*Ph*	A4ABNA	28.57	-*Ph* -4-CH$_3$	A4I4TA	4.76
-*Ph* -4-OCH$_3$	A4S4MA	11.23	-*Ph* -4-Cl	A4I4CA	4.67

-Ph -2-NO$_2$	A4S2NA	26.76	*-Ph* -4-COCH$_2$Cl	A4IBKA	14.1
-Ph -3-NO$_2$	A4S3NA	23.19	*-Ph* -2-COOC$_2$H$_5$	A4I2ANA	2.55
-Ph -4-NO$_2$	A4S4NA	71.36	*-Ph*-3-COOC$_2$H$_5$	A4I3ANA	4.16
-Ph -2-CH$_3$	A4S2TA	1.97	*-Ph* -4-COOC$_2$H$_5$	A4I4ANA	3.19

L= -NHCO- in Compound with code starting A4A, -NHSO$_2$- in compounds with code starting with A4S, -NHCONH- in Compound with code starting A4I

Spectral Information of intermediates and products

N-(3-cyanophenyl)-4-methoxybenzamide (3AA4MA): IR (cm^{-1}) (KBr): 3410.15 (s, NH, amide), 2227.78 (s, CN), 1670.35 (s, C=O amide).

N-(3-cyanophenyl)-3-methoxybenzamide (3AA3MA): IR (cm^{-1}) (KBr): 3240.41 (s, NH, amide), 2227.78 (s, CN), 1651.07 (s, C=O amide).

N-(3-cyanophenyl)-2-methoxybenzamide (3AA2MA): IR (cm^{-1}) (KBr): 3323.35 (s, NH, amide), 2231.64 (s, CN), 1676.14 (s, C=O amide).

2-chloro-N-(3-cyanophenyl)benzamide (3AA2CA): IR (cm^{-1}) (KBr): 3240.41 (s, NH, amide), 2227.78 (s, CN), 1658.78 (s, C=O amide).

N-(3-cyanophenyl)-2-nitrobenzene-1-sulfonamide (3AS2NA): IR (cm^{-1}) (KBr): 3327.21 (s, NH, amide), 2233.57 (s, CN), 1583.56, 1174.65 (s, S=O sulfonamide).

N-(3-cyanophenyl)-3-nitrobenzene-1-sulfonamide (3AS3NA): IR (cm^{-1}) (KBr): 3170.97 (s, NH, amide), 2245.14, 1583.56, 1178.51 (s, S=O sulfonamide).

N-(3-cyanophenyl)-4-nitrobenzene-1-sulfonamide (3AS4NA): IR (cm^{-1}) (KBr): 3169.04 (s, NH, amide), 2245.14 (s, CN), 1175.38 (s, S=O sulfonamide).

3-(3-cyanophenyl)-1-(4-fluorophenyl)urea (3AI4FA): IR (cm^{-1}) (KBr): 3286.31 (s, NH, amide), 2227.89 (s, CN), 1634.71 (s, C=O amide).

3-(3-cyanophenyl)-1-(4-methoxyphenyl)urea (3AI4MA): IR (cm^{-1}) (KBr): 3296.44 (s, NH, amide), 2227.89 (s, CN), 1639.77 (s, C=O amide).

3-(3-cyanophenyl)-1-(4-nitrophenyl)urea (3AI4NA): IR (cm^{-1}) (KBr): 3331.89(s, NH, amide), 2232.95 (s, CN), 1723.49 (s, C=O amide).

3-(3-cyanophenyl)-1-(3-methylphenyl)urea (3AI3TA): IR (cm^{-1}) (KBr): 3331.89 (s, NH, amide), 2238.02 (s, CN), 1662.61 (s, C=O amide).

N-(4-carbamimidoylphenyl)cyclopropanecarboxamide (A4ACPA): IR (cm^{-1}) (KBr): 3361.93, 3300.20, 3176.76 (s, NH/ NH$_2$, amide/amidine), 1660.71 (s, C=O amide).

N-(4-carbamimidoylphenyl)-4-fluorobenzamide (A4A4FA): IR (cm^{-1}) (KBr): 3358.07, 3313.71, 3163.26 (s, NH/ NH$_2$, amide/amidine), 1660.71 (s, C=O amide).

N-(4-carbamimidoylphenyl)-3-fluorobenzamide (A4A3FA): IR (cm^{-1}) (KBr): 3325.28 (s, NH/ NH$_2$, amide/amidine), 1647.21 (s, C=O amide).

N-(4-carbamimidoylphenyl)-2-fluorobenzamide (A4A2FA): IR (cm^{-1}) (KBr): 3356.14, 3284.77, 3165.19 (s, NH/ NH$_2$, amide/amidine), 1653.00 (s, C=O amide).

N-(4-carbamimidoylphenyl)-4-nitrobenzamide (A4A4NA): IR (cm^{-1}) (KBr): 3387.00, 3311.78, 3211.48 (s, NH/ NH$_2$, amide/amidine), 1645.28 (s, C=O amide).
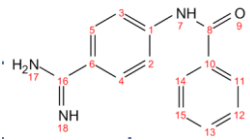
N-(4-carbamimidoylphenyl)-3-nitrobenzamide (A4A3NA): IR (cm^{-1}) (KBr): 3396.64, 3319.49 (s, NH, NH$_2$ amide/amidine), 1649.14 (s, C=O amide).

N-(4-carbamimidoylphenyl)-2-nitrobenzamide (A4A2NA): IR (cm^{-1}) (KBr): 3379.29, 3317.56, 3179.98 (s, NH/ NH$_2$, amide/amidine), 1645.28 (s, C=O amide).

N-(4-carbamimidoylphenyl)-3-methoxybenzamide (A4A3MA): IR (cm^{-1}) (KBr): 3358.07, 3298.28, 3161.33 (s, NH/ NH$_2$, amide/amidine), 1647.21 (s, C=O amide).

N-(4-carbamimidoylphenyl)-4-methoxybenzamide (A4A4MA): IR (cm^{-1}) (KBr): 3385.07, 3300.20, 3180.62 (s, NH/ NH$_2$, amide/amidine), 1649.14 (s, C=O amide).

N-(4-carbamimidoylphenyl)benzamide (A4ABNA): IR (cm^{-1}) (KBr): 3358.07, 3305.99, 3165.19 (s, NH/ NH$_2$, amide/amidine), 1645.28 (s, C=O amide); ^{1}H-NMR δ (ppm) (400Mhz CDCl$_3$, DMSO TMS): 7.52 (t, ArH, 2H), 7.58 (t, ArH, 1H), 7.84-7.87 (m, ArH, 4H), 7.97 (d, ArH, 2H), 10.38 (s, NH, 1H).^{13}C NMR δ (ppm) (100Mhz, DMSO and CDCl$_3$): 167.56 (C8), 156.41 (C16), 141.80 (C1), 134.75 (C10), 131.44 (C6), 128.97 (C13), 128.13 (C11, C14), 128.04 (C12, C15), 127.64 (C4, C5), 119.18 (C2, C3).

4-[3-(trifluoromethyl)benzenesulfonamido]benzene-1-carboximidamide (A4S3TFA): IR (cm^{-1}) (KBr): 3361.93, 3244.27, 3172.90 (s, NH/ NH$_2$, amide/amidine), 1649.14 (s, C=O amide).

4-(4-chlorobenzenesulfonamido)benzene-1-carboximidamide (A4S4CA): IR (cm^{-1}) (KBr): 3377.36, 3244.27, 3178.69 (s, NH/ NH$_2$, amide/amidine), 1569.62, 1171.34 (s, S=O).

4-(4-methylbenzenesulfonamido)benzene-1-carboximidamide (A4S2TA): IR (cm^{-1}) (KBr): 3479.58, 3365.78, 3111.18 (s, NH/ NH$_2$, amide/amidine), 1578.49, 1179.67 (s, S=O).

4-(4-nitrobenzenesulfonamido)benzene-1-carboximidamide (A4S2NA): IR (cm^{-1}) (KBr): 3479.58, 3365.778, 3109.25 (s, NH/ NH$_2$, amide/amidine), 1572.18, 1173.45 (s, S=O).

3.5.2.2 Synthesis of Benzamides

Amino group of amino acid (4-amino benzoic acid / 4-(aminomethyl) benzoic acid) was protected with Di-tert-butyl dicarbonate (Boc anhydride). Reaction follows nucleophilic substitution reaction (SN2) mechanism. Sodium bicarbonate was used to solubilize amino acid in the solvent. After stirring for overnight, the reaction mixture was diluted with water and acidified. Sodium salt of the product gets converted to free carboxylic acid and gets precipitated. Precipitate was collected and washed with water until free from acid. Isolated yield of Boc protected compounds BocPABA and BocMABA were 92.93 and 96.82% respectively. Amine protection was confirmed by disappearance of one of the two NH stretch peak and shift of carbonyl stretch of amide to 1687 cm^{-1} in IR spectroscopy. Amide NH proton peak at around 12.38 ppm and peak around 1.42 ppm in ^{1}H NMR spectra due to 9 protons of methyl group of t-butoxide confirms the Boc protection.

Boc Protected amino acid was then coupled with different amines using coupling reagent N,N'-diethylcarbodiimide (EDC). After evaporation of THF, reaction mass was diluted with water and acidified to remove unreacted amines. Byproduct diethyl urea also gets removed in water. Using ethyl acetate, product was extracted and washed with sodium bicarbonate solution to remove and unreacted Boc-amino acid as sodium salt. Ethyl acetate layer was then dried and concentrated to yield amide in practically pure form and yields of the product are tabulated in **Table 3.5-5**. Disappearance of broad -OH stretching due to carboxylic acid -OH and shift of C=O stretching from 1680-1682 cm^{-1} in IR indicates product formation.
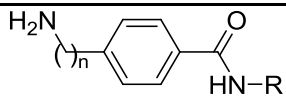
After acid amine coupling, Boc protected amine was cleaved to amine using Tri fluoro acetic acid. After completion of the reaction excess TFA was removed by azeotropic distillation with DCM and then protonated anime was basified with sodium bicarbonate solution. Amine precipitated was filtered and washed with water. Products were column purified using 10-25 % ethyl acetate in hexane over silica gel grade 100-120. Percentage yield of the product is reported in **Table 3.5-6**. Formation of desired product was then confirmed using IR and ^{1}H NMR spectroscopy.

Table 3.5-5: Isolated yield of 4-{(BOC-amino)/ [(BOC-amino) methyl]} benzamides (4A___B/M___B)

Amines used	4-(BOC-amino) benzoic acid series		4-[(BOC-amino) methyl] benzoic acid series	
	Code	% Yield	Code	% Yield
4-Amino Pyridine	4A4APB	55.85	M4APB	21.38
3-Amino Pyridine	4A3APB	84.04	M3APB	69.03
2-Amino Pyridine	4A2APB	45.45	M2APB	51.93
4-Toluidine	4A4TOB	68.42	M4TOB	79.31
3-Toluidine	4A3TOB	47.50	M3TOB	82.25
2-Toluidine	4A2TOB	45.45	M2TOB	76.38
4-Bromo aniline	4A4BRB	77.95	M4BRB	86.36
4-Chloro aniline	4A4CLB	86.50	M4CLB	81.48
3-Chloro aniline	4A3CLB	76.89	M3CLB	80.37
2-Chloro aniline	4A2CLB	45.65	M2CLB	33.26
4-Anisidine	4A4ASB	34.56	M4ASB	72.01
3-Anisidine	4A3ASB	40.27	M3ASB	82.07
2-Anisidine	4A2ASB	31.25	M2ASB	54.01
4-Nitro Aniline	4A4NTB	NS	M4NTB	21.54
3-Nitro Aniline	4A3NTB	44.44	M3NTB	72.03
2-Nitro Aniline	4A2NTB	NS	M2NTB	45.10
4-Fluoro aniline	4A4FLB	70.13	M4FLB	NS
4-Methoxy Benzyl amine	4A4MBB	86.04	M4MBB	83.68
3-Methoxy Benzyl amine	4A3MBB	87.44	M3MBB	0.00
2-Methoxy Benzyl amine	4A2MBB	83.24	M2MBB	84.13
Aniline	4AANIB	69.36	MANIB	74.30
α-Naphthyl amine	4ANAPB	30.84	MNAPB	67.07
Propyl amine	4APROB	75.04	MPROB	75.82
Butyl amine	4ABUTB	80.95	MBUTB	72.35
Cyclopropyl amine	4ACPRB	87.46	MCPRB	89.54
Furfuryl amine	4AFURB	80.49	MFURB	66.91
Morpholine	4AMORB	97.93	MMORB	NS
1-(2Amino ethyl) Pyrrolidine	4AEPYB	94.47	MEPYB	81.55
4-(2 Amino Ethyl) Morpholine	4AEMB	79.18	MEMB	79.33
1-(2 Amino Ethyl) Piperidine	4AEPIB	66.68	MEPIB	81.15
1-Phenyl piperazine	4A1PPB	80.83	M1PPB	NS
1-Benzyl piperazine	4A1BPB	50.99	M1BPB	NS

NS: not synthesized

Table 3.5-6: Isolated yield of 4-(amino/ aminomethyl) benzamides (4A___N/M___N)

R	4-amino benzamide series (n=0)		4-(aminomethyl) benzamide series (n=1)	
	Code	% Yield	Code	% Yield
p-pyridyl	4A4APN	0	M4APN	0.00
m-pyridyl	4A3APN	0	M3APN	39.60
o-pyridyl	4A2APN	0	M2APN	5.28
p-tolyl	4A4TON	72.46	M4TON	97.50
m-tolyl	4A3TON	27.27	M3TON	98.49
o-tolyl	4A2TON	43.02	M2TON	66.58
p-bromophenyl	4A4BRN	94.65	M4BRN	97.21
p-chlorophenyl	4A4CLN	78.64	M4CLN	89.50
m-chlorophenyl	4A3CLN	80.06	M3CLN	97.17
o-chlorophenyl	4A2CLN	43.78	M2CLN	34.52
p-anisyl	4A4ASN	39.90	M4ASN	82.91
m-anisyl	4A3ASN	64.63	M3ASN	94.94
o-anisyl	4A2ASN	49.20	M2ASN	1.95
p-nitrophenyl	4A4NTN	NS	M4NTN	0.00
m-nitrophenyl	4A3NTN	26.92	M3NTN	97.32
o-nitrophenyl	4A2NTN	NS	M2NTN	11.06
p-fluorophenyl	4A4FLN	NS	M4FLN	NS
p-methoxybenzyl	4A4MBN	83.24	M4MBN	83.23
m-methoxybenzyl	4A3MBN	90.17	M3MBN	NS
o-methoxybenzyl	4A2MBN	0	M2MBN	0.92
phenyl	4AANIN	88.63	MANIN	88.39
α-naphthyl	4ANAPN	69.11	MNAPN	99.88
propyl	4APRON	0	MPRON	9.10
butyl	4ABUTN	0	MBUTN	55.75
cyclo-propyl	4ACPRN	0	MCPRN	72.28
furfuryl	4AFURN	0	MFURN	0
morphonyl	4AMORN	86.48	MMORN	NS
pyrrolidine N-ethyl	4AEPYN	0	MEPYN	0
morpholine N-ethyl	4AEMN	0	MEMN	0
piperidine N-ethyl	4AEPIN	0	MEPIN	0
4-phenyl-1- piperazinyl	4A1PPN	69.90	M1PPN	NS
4-benzyl-1- piperazinyl	4A1BPN	85.71	M1BPN	NS

NS: not synthesized

Spectral Information of intermediates and products

4-{[(tert-butoxy)carbonyl]amino}benzoic acid (BocPABA): IR (cm^{-1}) (KBr): 3356.14 (S, NH amide), 2981.95 (s, b, OH, carboxylic acid) , 1687.71(s, C=O amide). ^{1}H-NMR δ (ppm) (400 Mhz, CDCl$_3$, TMS): 1.51 (s, CH$_3$, 9H), 7.56 (d, ArH, 2H), 7.83 (d, ArH, 2H), 12.39 (s, CONH, 1H).

4-({[(tert-butoxy)carbonyl]amino}methyl)benzoic acid (BocMABA): IR (cm^{-1}) (KBr): 3356.14 (s, NH amide), 2981.95 (s, b , OH, carboxylic acid) , 1687.71(s, C=O amide). ^{1}H-NMR δ (ppm) (400 Mhz, CDCl$_3$, TMS): 1.42 (s, CH$_3$, 9H), 4.21 (d, CH$_2$, 2H), 7.33 (d, ArH, 2H), 7.90 (d, ArH, 2H), 12.38 broad (s, CONH, 1H).

tert-butyl N-{4-[(4-methylphenyl)carbamoyl]phenyl}carbamate (4A4TOB) : IR (cm^{-1}) (KBr): 3344.57 and 3325.28 (s, NH amide), 1697.36 (s, C=O amide), 1643.35 (s, C=O amide). ^{1}H-NMR δ (ppm) (400 Mhz, CDCl$_3$, TMS): 1.52 (s, CH$_3$, 9H), 2.30 (s, CH$_3$, 3H), 7.10 (d, ArH, 2H), 7.58 (d, ArH, 2H), 7.63 (d, ArH, 2H), 7.88 (d, ArH, 2H), 7.88 (d, ArH, 2H), 9.46 (s, CONH, 1H), 9.87 (s, CONH, 1H).

tert-butyl N-({4-[(4-methylphenyl)carbamoyl]phenyl}methyl)carbamate (M4TOB): IR (cm^{-1}) (KBr): 3356.14 (s, NH amide), 1687.71 (s, C=O amide). ^{1}H-NMR δ (ppm) (400 Mhz, CDCl$_3$, TMS): 1.43 (s, CH$_3$, 9H), 2.33 (s, CH$_3$, 3H), 4.22 (d, CH$_2$, 2H), 6.89 (d, ArH. 1H), 7.19 (d, ArH, 1H), 7.33 (d, ArH, 1H), 7.36 (d, ArH, 2H), 7.61 (s, ArH, 1H), 7.90 (d, ArH, 2H), 10.00 (s, CONH, 1H).

4-amino-N-(4-methylphenyl)benzamide (4A4TON): IR (cm^{-1}) (KBr) : 3366.08, 3360.00 and 3211.48 (s, NH of amide and amine), 1647.21 (s, C=O amide).

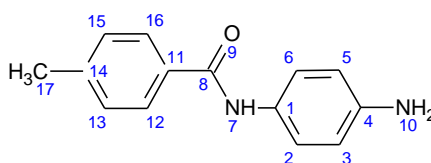**4-amino-N-(4-methoxyphenyl)benzamide (4A4ASN):** IR (cm^{-1}) (KBr): 3327.21, 3219.19, 3005.10 (s, NH of amide and NH$_2$ of amine), 1647.21 (s, C=O amide). ^{1}H-NMR δ (ppm) (400 Mhz, CDCl$_3$, TMS): 3.74 (s, OCH$_3$, 3H), 6.60 (d, ArH, 2H), 6.85 (d, ArH, 2H), 7.64 (d, ArH, 2H), 7.71 (d, ArH, 2H), 8.22 (s, NH$_2$, 2H), 9.62 (s, CONH, 1H).

4-amino-N-(3-chlorophenyl)benzamide (4A3CLN): IR (cm^{-1}) (KBr): 3315.63, 3215.34, 3039.81 (s, NH of amide and NH$_2$ of amine), 1645.28 (s, C=O amide). ^{1}H-NMR δ (ppm) (400 Mhz, CDCl$_3$, TMS): 6.62 (d, ArH, 2H), 7.03 (d, ArH, 1H), 7.27 (t, ArH, 1H), 7.67 (d, ArH, 1H), 7.72 (d, ArH, 2H), 7.96 (s, NH$_2$, 1H), 9.86 (s, CONH, 1H).

4-amino-N-[(3-methoxyphenyl)methyl]benzamide (4A3MBN): IR (cm^{-1}) (KBr): 3352.28, 3319.49, 3234.62 (s, NH of amide and NH$_2$ of amine), 1637.56 (s, C=O amide). ^{1}H-NMR δ (ppm) (400 Mhz, CDCl$_3$, TMS): 3.74 (s, -OCH$_3$, 3H) 6.57 (d, ArH, 2H), 6.75 (d, ArH, 1H), 6.86 (d, ArH, 1H), 7.19 (t, ArH, 1H), 7.63 (d, ArH, 2H), 8.49 (s, NH$_2$, 1H), 9.86 (s, CONH, 1H).

4-amino-N-(pyridin-4-yl)benzamide (4APPN): IR (cm⁻¹) (KBr): 3373.50, 3331.07, 3230.77 (s, NH of amide and NH_2 of amine), 1629.85 (s, C=O amide). ^{1}H-NMR δ (ppm) (400 Mhz, CDCl₃, TMS): 3.17 (s, CH₂, 4H), 3.78 (3, CH₂, 4H), 6.65 (d, ArH, 2H), 6.92 (m, ArH, 3H), 7.27 (m, ArH, 4H).

4-amino-N-(4-methylphenyl)benzamide (4A4TON): IR (cm⁻¹) (KBr): 3366.08, 3360.00, 3211.48 (s, NH of amide and NH_2 of amine), 1647.21 (s, C=O amide).^{1}H-NMR δ (ppm) (400 Mhz, CDCl₃, TMS): 2.28 (s, CH₃, 3H), 6.62 (d, ArH, 2H), 7.08 (d, ArH, 2H), 7.62 (d, ArH, 2H), 7.72 (d, ArH, 2H), 9.61 (s, CONH, 1H). ^{13}C NMR (100Mhz, DMSO and CDCl₃): δ (ppm) 165.23 (C8), 151.55 (C4), 136.98 (C14), 131.68 (C11), 129.09 (C13, C15), 128.59 (C12, C16), 121.60 (C1), 120.14 (C2 and C6), 112.59 (C3 and C5), 20.45 (C17).

4-amino-N-phenylbenzamide (4ANIN): IR (cm⁻¹) (KBr): 3346.50, 3317.56, 3217.27 (s, NH of amide and NH_2 of amine), 1643.35 (s, C=O amide). ^{1}H-NMR δ (ppm) (400 Mhz, CDCl₃, TMS): 6.61-6.63 (d, ArH, 2H), 7.00-7.03 (d, ArH, 2H), 7.25-7.29 (t, ArH, 1H), 7.72-7.76 (m, ArH, 4H), 9.71(s, CONH, 1H)

4-amino-N-(naphthalen-1-yl)benzamide (4ANAPN): IR (cm⁻¹) (KBr): 3473.80, 3358.07, 3234.62 (s, NH of amide and NH_2 of amine), 1687.71(s, C=O amide).

4-amino-N-(2-methoxyphenyl)benzamide (4A2ASN): IR (cm⁻¹) (KBr): 3441.01, 3346.50, 3230.77 (s, NH of amide and NH_2 of amine), 1641.42(s, C=O amide).

4-amino-N-(2-chlorophenyl)benzamide (4A2CLN): IR (cm⁻¹) (KBr): 3437.15, 3369.64, 3319.49 (s, NH of amide and NH_2 of amine), 1678.07 (s, C=O amide).

4-amino-N-(4-bromophenyl)benzamide (4A4BRN): IR (cm⁻¹) (KBr): 3334.92 (s, NH of amide and NH_2 of amine), 1658.78 (s, C=O amide).

4-amino-N-(3-methylphenyl)benzamide (4A3TON): IR (cm⁻¹) (KBr): 3417.86, 3332.99, 3234.62, (s, NH of amide and NH_2 of amine), 1641.42 (s, C=O amide).

4-amino-N-(2-methylphenyl)benzamide (4A2TON): IR (cm⁻¹) (KBr): 3458.37, 3373.50, 3344.57, 3238.48 (s, NH of amide and NH_2 of amine), 1641.42 (s, C=O amide).

4-amino-N-(3-methoxyphenyl)benzamide (4A3ASN): IR (cm⁻¹) (KBr): 3473.80, 3367.71, 3238.48 (s, NH of amide and NH_2 of amine), 1641.42 (s, C=O amide).

4-amino-N-[(4-methoxyphenyl)methyl]benzamide (4A4MBN): IR (cm^{-1}) (KBr): 3415.93, 3334.92, 3217.27 (s, NH of amide and NH$_2$ of amine), 1627.92 (s, C=O amide).

4-amino-N-(3-nitrophenyl)benzamide (4A3NTN): IR (cm^{-1}) (KBr): 3423.65, 3329.14, 3232.70 (s, NH of amide and NH$_2$ of amine), 1660.71 (s, C=O amide).

4-(aminomethyl)-N-(4-methylphenyl)benzamide (M4TON): IR (cm^{-1}) (KBr): 3296.35, 3020.53 (s, NH of amide and NH$_2$ of amine), 1651.07 (s, C=O amide). ^{1}H-NMR δ (ppm) (400 Mhz, CDCl$_3$, TMS):2.32 (s, -CH$_3$, 3H), 4.11 (s, -CH$_2$-, 2H), 7.12 (d, ArH, 2H), 7.58 (d, ArH, 2H), 7.65 (d, ArH, 2H), 8.02 (ArH, 2H), 10.06 (s, CONH, 1H).

4-(aminomethyl)-N-(4-chlorophenyl)benzamide (M4CLN): IR (cm^{-1}) (KBr): 3300.20, 3030.17 (s, NH of amide and NH$_2$ of amine), 1643.35 (s, C=O amide). ^{1}H-NMR δ (ppm) (400 Mhz, CDCl$_3$, TMS): 4.25 (d, -CH$_2$-, 2H), 7.32 (d, ArH. 2H), 7.40 (d, ArH, 1H), 7.47 (d, ArH, 1H), 7.83 (d, ArH, 2H), 7.92 (d, ArH, 2H), 10.24 (s, CONH, 1H).

4-(aminomethyl)-N-(3-chlorophenyl)benzamide (M3CLN): IR (cm^{-1}) (KBr): 3217.27, 3020.53, 3001.24 (s, NH of amide and NH$_2$ of amine), 1651.07 (s, C=O amide). ^{1}H-NMR δ (ppm) (400 Mhz, CDCl$_3$, TMS): 4.27 (d, CH$_2$, 2H), 7.08 (d, ArH. 1H), 7.30 (d, Ar, 1H), 7.41 (d, ArH, 1H), 7.47 (d, ArH, 1H), 7.70 (d, ArH, 1H), 7.93 (d, ArH, 2H), 7.98 (s, ArH, 1H), 10.25 (s, CONH, 1H).

4-(aminomethyl)-N-cyclopropylbenzamide (MCPN): IR (cm^{-1}) (KBr): 3358.07, 3232.70, (s, NH of amide and NH$_2$ of amine), 1633.71(s, C=O amide). ^{1}H-NMR δ (ppm) (400 Mhz, CDCl$_3$, TMS): 0.62 (d, CH$_2$, 2H), 2.78 (m, CH, 1H), 5.43 (s, -CH$_2$-, 2H), 6.53 (d, ArH, 2H), 7.55 (d, ArH. 2H), 7.92 (d, CONH, 1H).

4-(aminomethyl)-N-[(4-methoxyphenyl)methyl]benzamide (M4MBN): IR (cm^{-1}) (KBr): 3292.49, 3005.10 (s, NH of amide and NH$_2$ of amine), 1633.71 (s, C=O amide). ^{1}H-NMR δ (ppm) (400 Mhz, CDCl$_3$, TMS): 3.74 (s, OCH$_3$, 3H), 4.42 (d, CH$_2$, 2H), 6.84 (d, ArH, 2H), 7.25 (d, ArH, 2H), 7.33 (d, ArH, 1H), 7.39 (d, ArH, 1H), 7.84 (d, ArH, 2H), 8.84 (s, CONH, 1H).
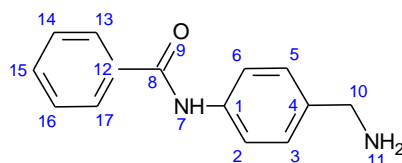
4-(aminomethyl)-N-butylbenzamide (MBUTN): IR (cm^{-1}) (KBr): 3332.99, 3302.13 (s, NH of amide and NH$_2$ of amine), 1631.78 (s, C=O amide). ^{1}H-NMR δ (ppm) (400 Mhz, CDCl$_3$, TMS): 0.91-0.95 (t, CH$_3$, 3H), 1.36 (m, CH$_2$, 2H), 1.54 (m, CH$_2$, 2H), 3.28 (t, CH$_2$, 2H), 4.21-4.23 (d, CH$_2$, 2H), 7.30-7.41 (d, ArH. 2H), 7.79-7.81 (d, ArH, 2H), 8.28-8.29 (s, CONH, 1H).

4-(aminomethyl)-N-(2-methylphenyl)benzamide (M2TON): IR (cm^{-1}) (KBr): 3296.35, 3026.31 (s, NH of amide and NH$_2$ of amine), 1645.28 (s, C=O amide). ^{1}H-NMR δ (ppm) (400 Mhz, CDCl$_3$, TMS): 2.26 (s, CH$_3$, 3H), 4.08 (s, CH$_2$, 2H), 7.17(d, ArH, 1H), 7.23 (d, ArH, 1H), 7.26 (d, ArH, 1H), 7.34 (d, ArH, 1H), 7.57 (d, ArH, 2H), 8.02 (d, ArH. 2H), 9.87 (s, CONH, 1H).

4-(aminomethyl)-N-(3-methylphenyl)benzamide (M3TON): IR (cm^{-1}) (KBr): 3232.70, 3007.02 (s, NH of amide and NH$_2$ of amine), 1651.07 (s, C=O amide). ^{1}H-NMR δ (ppm) (400 Mhz, CDCl$_3$, TMS): 2.33 (s, CH$_3$, 3H), 4.26 (d, CH$_2$, 2H), 6.89 (d, ArH, 1H), 7.19 (t, ArH, 1H), 7.39 (d, ArH, 1H), 7.46 (d, ArH, 1H), 7.56 (d, ArH, 1H), 7.62(d, ArH, 1H), 7.92 (d, ArH. 2H), 10.03 (s, CONH, 1H). ^{13}C NMR (100Mhz, DMSO and CDCl$_3$): δ (ppm) 165.27 (C8), 138.98 (C14), 137.50 (C4), 133.07 (C1), 128.10 (C12), 127.49 (C15), 126.73 (C17), 126.57 (C13), 124.07 (C16), 120.86 (C3, C5), 117.47 (C2, C6), 45.14 (C10), 21.19 (C18).
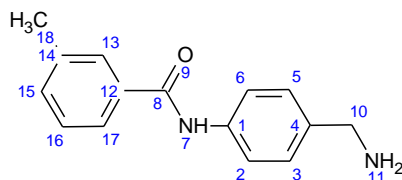
4-(aminomethyl)-N-phenylbenzamide (MANIN): IR (cm^{-1}) (KBr): 3325.28, 3030.17 (s, NH of amide and NH$_2$ of amine), 1651.07 (s, C=O amide). ^{1}H-NMR δ (ppm) (400 Mhz, CDCl$_3$, TMS): 4.26 (d, CH$_2$, 2H), 7.07 (t, ArH, 1H), 7.32 (d, ArH, 2H), 7.40 (d, ArH, 1H), 7.47 (d, ArH, 1H), 7.78 (d, ArH, 2H), 7.93 (d, ArH. 2H), 10.11 (s, CONH, 1H). ^{13}C NMR (100Mhz, DMSO and CDCl$_3$): δ (ppm) 165.34 (C8), 139.08 (C4), 133.05 (C1), 128.25 (C16, C14), 127.51(C15), 126.74 C13, C17, 126.57 (C5, C3), 123.31 (C12), 120.27 C6, C2 , 45.13 C10

4-(aminomethyl)-N-(4-methoxyphenyl)benzamide (M4ASN): IR (cm^{-1}) (KBr): 3329.49, 3005.10 (s, NH of amide and NH$_2$ of amine), 1643.35 (s, C=O amide). ^{1}H-NMR δ (ppm) (400 Mhz, CDCl$_3$, TMS): 3.76 (s, O CH$_3$, 3H), 4.24-4.25 (d, CH$_2$, 2H), 6.88 (d, ArH. 2H), 7.38 (d, ArH, 1H), 7.47 (d, ArH, 1H), 7.68 (d, ArH, 2H), 7.92 (d, ArH, 2H), 10.02 (s, CONH, 1H).

4-(aminomethyl)-N-(pyridin-3-yl)benzamide (M3APN): IR (cm^{-1}) (KBr): 3344.57, 3005.10 (s, NH of amide and NH$_2$ of amine), 1681.93 (s, C=O amide).

4-(aminomethyl)-N-(pyridin-2-yl)benzamide (M2APN): IR (cm^{-1}) (KBr): 3302.13, 3242.34, 3062.96 (s, NH of amide and NH$_2$ of amine), 1633.71 (s, C=O amide).

4-(aminomethyl)-N-(3-methoxyphenyl)benzamide (M3ASN): IR (cm^{-1}) (KBr): 3298.28, 3041.74 (s, NH of amide and NH$_2$ of amine), 1643.35 (s, C=O amide).

4-(aminomethyl)-N-(2-methoxyphenyl)benzamide (M2ASN): IR (cm^{-1}) (KBr): 3332.99, 3317.56, 3064.89 (s, NH of amide and NH$_2$ of amine), 1680.00 (s, C=O amide).

4-(aminomethyl)-N-(2-chlorophenyl)benzamide (M2CLN): IR (cm^{-1}) (KBr): 3431.36, 3294.42, 3037.89 (s, NH of amide and NH$_2$ of amine), 1681.93 (s, C=O amide).

4-(aminomethyl)-N-(4-bromophenyl)benzamide (M4BRN): IR (cm^{-1}) (KBr): 3298.28, 3037.89 (s, NH of amide and NH$_2$ of amine), 1643.35 (s, C=O amide).

4-(aminomethyl)-N-[(2-methoxyphenyl)methyl]benzamide (M2MBN): IR (cm^{-1}) (KBr): 3323.35, 3064.89 (s, NH of amide and NH$_2$ of amine), 1639.49 (s, C=O amide).

4-(aminomethyl)-N-(3-nitrophenyl)benzamide (M3NTN): IR (cm^{-1}) (KBr): 3369.64, 3078.39 (s, NH of amide and NH$_2$ of amine), 1680.00 (s, C=O amide).

4-(aminomethyl)-N-(2-nitrophenyl)benzamide (M2NTN): IR (cm^{-1}) (KBr): 3298.28 (s, NH of amide and NH$_2$ of amine), 1730.15 (s, C=O amide).

4-(aminomethyl)-N-(naphthalen-1-yl)benzamide (MNAPN): IR (cm^{-1}) (KBr): 3199.91, 3007.02 (s, NH of amide and NH$_2$ of amine), 1643.35 (s, C=O amide).

4-(aminomethyl)-N-propylbenzamide (MPRON): IR (cm^{-1}) (KBr): 3296.35 (s, NH of amide and NH$_2$ of amine), 1633.71 (s, C=O amide).

3.5.3 References

[1] Groebke Zbinden, K.; Banner, D.W.; Ackermann, J.; D'Arcy, A.; Kirchhofer, D.; Ji, Y.-H.; Tschopp, T.B.; Wallbaum, S.; Weber, L. Design of selective phenylglycine amide tissue factor/factor VIIa inhibitors. *Bioorg Med Chem Lett.* **2005,** 15(3), 817-822.

[2] Kotian, P.L.; Krishnan, R.; Rowland, S.; El-Kattan, Y.; Saini, S.K.; Upshaw, R.; Bantia, S.; Arnold, S.; Sudhakar Babu, Y.; Chand, P. Design, parallel synthesis, and crystal structures of biphenyl antithrombotics as selective inhibitors of tissue factor FVIIa complex. Part 1: Exploration of S2 pocket pharmacophores. *Bioorg Med Chem Lett.* **2009,** 17(11), 3934-3958.

[3] Rossi, C.; Fincham, C.I.; D'Andrea, P.; Porcelloni, M.; Ettorre, A.; Mauro, S.; Bigioni, M.; Binaschi, M.; Maggi, C.A.; Nardelli, F. 4-*N*-Hydroxy-4-[1-(sulfonyl) piperidin-4-yl]-butyramides as HDAC inhibitors. *Bioorg Med Chem Lett.* **2011,** 21(22), 6767-6769.

[4] Tale, R.H.; Rodge, A.H.; Hatnapure, G.D.; Keche, A.P. The novel 3, 4-dihydropyrimidin-2 (1*H*)-one urea derivatives of N-aryl urea: Synthesis, anti-inflammatory, antibacterial and antifungal activity evaluation. *Bioorg Med Chem Lett.* **2011,** 21(15), 4648-4651.

[5] Roy, G.; Placzek, E.; Scanlan, T.S. ApoB-100-containing lipoproteins are major carriers of 3-iodothyronamine in circulation. *J Biol Chem.* **2012,** 287(3), 1790-1800.

[6] Hartwig, S.; Nguyen, M.M.; Hecht, S. Exponential growth of functional poly (glutamic acid) dendrimers with variable stereochemistry. *Polym Chem.* **2010,** 1(1), 69-71.

3.6 In Vitro Screening of PAD4 Inhibitors

P AD4, a guanidine modifying enzyme, converts arginine residue to citrulline and ammonia. In vitro activity of PAD4 can be measured in terms of amount of citrulline or ammonia formed during citrullination of the substrate N-α-benzoyl-L-arginine ethyl ester (BAEE). Initially, PAD4 was incubated with inhibitor for certain period followed by addition of substrate BAEE. Amount of ammonia released by catalyzing the substrate is directly proportional to amount of free active enzyme. Released ammonia was detected using Cayman Chemical`s proprietary assay method [1]. The potency of inhibitor and ammonia release are inversely proportional.

3.6.1 Methodology

PAD4 enzyme was diluted to 600 µl with 50 mM borate assay buffer of pH 8.0 containing 10 mM calcium chloride. PAD4 substrate N-α-benzoyl-L-arginine ethyl ester (BAEE) was reconstituted with 600 µl of assay buffer. A 10 mM primary stock of inhibitors was prepared in DMSO from which 2.1 mM secondary stock was prepared. Initially all the reagents except enzyme solution were equilibrated to room temperature. Amount of reagents added to each well of black 96 well plate is given in **Table 3.6-1**. After addition of inhibitor, reaction mixture was incubated at 37° C for 20 min. followed by addition of 10 µl of substrate in each well. Plates were covered and incubated for 30 min at 37° C. The reaction was stopped at the end by addition of 20 µl of stop solution (a calcium chelator). 10 µl of ammonia detecting reagent was added, plated were again covered and incubated at 37°C for 15 min. Plate cover was removed and fluorescence was monitored using Tecan`s Infinite M1000 instrument at excitation wavelength of 405 nm and emission wavelength of 470 nm. Percentage inhibition of the PAD4 can be calculated by **Equation 3.6-1**.

Table 3.6-1: Well content in in vitro PAD4 activity screening assay

well	Buffer	Inhibitor	PAD4	DMSO with Buffer	Substrate	Stop reagent	Ammonia detector
100% activity (PC)	150	-	10	10	10	20	10
Background (NC)	160	-	-	10	10	20	10
Inhibitor well	150	10	10	-	10	20	10

PC. Positive control, NC: negative control; all values are expressed in µl

Equation 3.6-1: $$\% \text{ Inhibition} = \frac{\Delta \text{ RFU of PC} - \Delta \text{ RFU of EI}}{\Delta \text{RFU of PC}} \times 100$$

Where RFU= Relative Fluorescence Units, PC = positive control or enzyme control, EI = enzyme inhibitor solution.

Percentage inhibition generally indicates binding ability of the inhibitors with the enzyme. But by increasing size of the molecule, binding tends to increase. In order to evaluate actual performance increase in the molecule normalization of activity data is needed. Potency Efficacy Index (PEI) is a normalized scale based on molecular weight and helps in relative comparison of compounds tested under similar condition and at same concentration only [2].

Equation 3.6-2: $$\text{PEI} = \frac{\%\text{inhibition at a given concentration as fraction between } 0-1.0}{\text{Molecular Weight (Da)}}$$

3.6.2 Results and Discussion

3.6.2.1 In Vitro Assay

Synthesized compounds after characterization were screened for in vitro PAD4 inhibition at 100 µM concentration using FRET assay method and Infinite M1000 PRO (Tecan) microplate reader.

Benzamidine Series: Percentage PAD4 inhibitory potential of benzamidine series compounds is given in **Table 3.6-2**. Compared to amide and sulphonamide analogues most of the urea analogues were found to have mild to moderate activity.

Benzamidine amide series: In benzamidine amide series, 3-Nitro cinnamoylamide substituted at *para* position of benzamidine (A4ACNA) is the most potent compound with 51.77% inhibition at 100µM concentration. Whereas benzylamide substitution (A4ABEA), 2 Furoamide (A4A2FuA) and 3-nitro benzamide (A4A3NA) showed weak activity with % inhibition of 18.12, 17.74 and 12.82 respectively. Benzamide substitution (A4ABNA) showed weak activity with 8.72% inhibition. 4-Methoxy benzamide (A4A4MA) and 3-Methoxy benzamide (A4A3MA) were inactive to inhibit the enzyme. Substitution with electron withdrawing group as in 4-Nitro benzamide (A4A4NA), 2-Nitro benzamide (A4A2NA), 4-Fluoro benzamide (A4A4FA), 3-Fluoro benzamide (A4A3FA) and 2-Fluoro benzamide (A4A2FA) with the exception of 3-Nitro benzamide (A4A3NA) lead to loss of inhibitory potential. Alkyl substitution like cylcopropyl amide (A4ACPA) did not contribute to the potency of the inhibitor. Alkylaryl or unsubstituted aryl group seemed to enhance activity. 3-Nitro cinnamoylamide (A4ACNA), benzylamide (A4ABEA), 2-Furoamide (A4A2FuA) and benzamide (A4ABNA) derivatives showed potent to moderate activity. Activity was increased

with substitution between aromatic and amide group. The increase in activity was seen from phenyl to 2-furoyl to benzyl to cinnamoyl derivative.

Table 3.6-2: In vitro PAD4 inhibitory potential of synthesized Benzamidine series
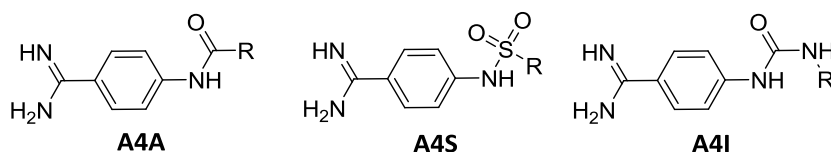
code	R	% inhib	PEI X 10^3	code	R	% inhib	PEI X 10^3
Benzamidine amides (A4A)				**Benzamidine sulphonamides (A4S)**			
A4A4MA	-*Ph*-4-OCH$_3$	IA		A4S4MA	-*Ph*-4-OCH$_3$	20.07	0.657
A4A3MA	-*Ph*-3-OCH$_3$	IA		A4S4TA	-*Ph*-4-CH$_3$	22.12	0.764
A4A4NA	-*Ph*-4-NO$_2$	IA		A4S3TA	-*Ph*-3-CH$_3$	IA	
A4A3NA	-*Ph*-3-NO$_2$	12.82	0.451	A4S2TA	-*Ph*-2-CH$_3$	IA	
A4A2NA	-*Ph*-2-NO$_2$	IA		A4S4NA	-*Ph*-4-NO$_2$	IA	
A4A4FA	-*Ph*-4-F	IA		A4S3NA	-*Ph*-3-NO$_2$	IA	
A4A3FA	-*Ph*-3-F	IA		A4S2NA	-*Ph*-2-NO$_2$	IA	
A4A2FA	-*Ph*-2-F	IA		A4S4CA	-*Ph*-4-Cl	19.3	0.623
A4ABEA	-CH$_2$-*Ph*	18.12	0.715	A4S3TFA	-*Ph*-3-CF$_3$	IA	
A4ABNA	-*Ph*	8.73	0.365				
A4A2FuA	2-Furoyl-	17.74	0.774				
A4ACNA	-*trans*-Cinnamoyl-3-NO$_2$	51.77	1.668				
A4ACPA	-*cyclo*-Propyl	IA					
Benzamidine ureas (A4I)							
A4I4MA	-*Ph*-4-OCH$_3$	IA		A4I4TA	-*Ph*-4-CH$_3$	36.22	1.350
A4I3MA	-*Ph*-3-OCH$_3$	IA		A4I3TA	-*Ph*-3-CH$_3$	23.96	0.893
A4I2MA	-*Ph*-2-OCH$_3$	31.18	1.097	A4I4NA	-*Ph*-4-NO$_2$	IA	
A4IBKA	-*Ph*-4-COCH$_2$Cl	30.5	0.922	A4I4FA	-*Ph*-4-F	19.9	0.731
A4I2ANA	-*Ph*-2-COOC$_2$H$_5$	16.68	0.511	A4I3FA	-*Ph*-3-F	IA	
A4I3ANA	-*Ph*-3-COOC$_2$H$_5$	25.78	0.79	A4I2FA	-*Ph*-2-F	IA	
A4I4ANA	-*Ph*-4-COOC$_2$H$_5$	18.79	0.576	A4I4CA	-*Ph*-4-Cl	33.07	1.145

% Inhib : percentage inhibition at 100 µM concentration; IA: inactive.

Benzamidine sulfonamide series: In benzamidine sulfonamides, only *para* substituted compounds showed moderate activity. Electron withdrawing and electron releasing group at *para* position did not alter the potency of the compound. Nitro- substituted aryl sulfonamide derivatives A4S2NA A4S3NA A4S4NA were inactive. *o*-toludine sulfonamide and *m*-toludine

sulfonamide derivatives A4S3TA A4S4TA were also inactive. Compound A4S3TFA, a *m*-trifuoro methyl benzenesulfonamide derivative also failed to inhibit the enzyme. *p*-methoxy benzene sulfonamide derivative A4S4MA showed 20.07% inhibition and it increased to 22.12% by replacing methoxy with methyl group (A4S4TA). On replacing methoxy with electron withdrawing chlorine atom (A4S4CA), the activity was decreased to 19.3%.
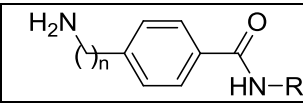
Benzamidine urea series : In benzamidine urea series, *p*-tolyl phenyl urea derivative A4I4TA showed potent activity with 36.22% inhibition at 100 µM concentration. Whereas *meta* methyl substitution rendered the compound A4I3TA inactive. Electron donating methoxy substitution at *meta* and *para* position (compounds A4I3MA and A4I4MA), rendered the compounds inactive towards PAD4 inhibition, while *ortho* substitution of methoxy group in A4I2MA caused moderate inhibition of PAD4 enzyme. Electron withdrawing groups such as nitro at *para* position made compound A4I4NA inactive towards PAD4. Fluoro substitution at *ortho* and *meta* position also rendered compounds A4I2FA and A4I3FA as inactive but *para* substitution made the compounds A4I4FA moderately active with 19.9% inhibition. Activity was further enhanced to 33.07% in compound A4I4CA, when *para* fluoro was replaced with chloro group. Activity was decreased slightly from 33.07 to 30.5% when *para* chloro was replaced with *para* CO-CH$_2$Cl (chloro acetyl) group in compound A4IBKA. Activity was further decreased to 18.79% by replacing chloroacetyl with COOCH$_2$CH$_3$ (ethoxy carbonyl) group in compound A4I4ANA. Substitution of COOCH$_2$CH$_3$ (ethoxy carbonyl) group at *ortho* position as in A4I2ANA and at *meta* position as in compound A4I3NA showed 16.68 and 26.78% inhibition respectively.

Benzamide Series: Percentage PAD4 inhibitory potential of benzamidine series is given in **Table 3.6-3**. Compared to 4-amino benzamide analogues, 4-(aminomethyl) benzamide analogues are more potent.

***p*-amino benzamide series:** The morpholin-4-yl derivative 4AMORN, showed potent activity with 49.79% inhibition. Activity was decreased to 27.17% with 4-phenylpiperazin-1-yl substitution in compound 4APPN where as compound was inactive when phenyl group on piperazine of 4APPN was replaced with benzyl group in 4APBN. Substitution with phenyl group in compound 4AANIN showed moderate activity with 23.8% inhibition where as substitution with α-naphthoyl group in 4ANAPN rendered the compound feebly active with 2.41% inhibition. Replacing phenyl with *m*-tolulyl group increased the potency of the compound 4A3TON slightly to 25.26% while replacing with *o*-toluyl and *p*-toluyl group

reduced inhibitory potential of compounds 4A4TON and 4A2TON to 16.93 and 16.67% respectively.

Table 3.6-3: In vitro PAD4 inhibitory potential of synthesized Benzamide series

R	4-amino benzamide series (n=0)			4-(aminomethyl) benzamide series (n=1)		
	code	% inhib	PEI X10^{-3}	code	% inhib	PEI X10^{-3}
p-Tolyl	4A4TON	16.93	0.748	M4TON	23.69	0.986
m-Tolyl	4A3TON	25.26	1.116	M3TON	28.2	1.174
o-Tolyl	4A2TON	16.67	0.737	M2TON	16.63	0.692
p-Bromophenyl	4A4BRN	21.73	0.746	M4BRN	21.8	0.714
p-Chlorophenyl	4A4CLN	15.55	0.630	M4CLN	36.41	1.761
m-Chlorophenyl	4A3CLN	4.59	0.186	M3CLN	24.63	1.191
o-Chlorophenyl	4A2CLN	31.18	1.264	M2CLN	39.95	1.933
p-Anisyl	4A4ASN	22.54	0.930	M4ASN	40.22	1.569
m-Anisyl	4A3ASN	24.37	1.006	M3ASN	42.77	1.669
o-Anisyl	4A2ASN	2.75	0.114	M2ASN	45.47	1.774
m-Nitrophenyl	4A3NTN	3.2	0.124	M3NTN	IA	
o-Nitrophenyl		NS		M2NTN	37.83	1.395
p-Methoxybenzyl	4A4MBN	IA		M4MBN	IA	
m-Methoxybenzyl	4A3MBN	42.46	1.657	M3MBN	NS	
o-Methoxybenzyl		NS		M2MBN	72.9	2.697
Phenyl	4AANIN	23.8	1.121	MANIN	4.7	0.208
α-Naphthyl	4ANAPN	2.41	0.092	MNAPN	46.3	1.676
Propyl		NS		MPRON	19.32	1.005
Butyl		NS		MBUTN	11.21	0.543
cyclo-Propyl		NS		MCPRN	IA	
p-Pyridyl		NS		M4APN	NS	
m-Pyridyl		NS		M3APN	IA	
o-Pyridyl		NS		M2APN	8.55	0.376
4-Phenyl-1- piperazinyl	4APPN	27.17	0.966		NS	
4-Benzyl-1- piperazinyl	4APBN	IA			NS	

% Inhib : percentage inhibition at 100 μM concentration; IA: inactive; NS: not synthesized

Substitution of phenyl ring with *p*-anisyl (4A4ASN) and *m*-anisyl (4A3ASN) did not alter the activity significantly but *o*-anisyl substitution in compound 4A2ASN drastically reduced inhibitory potential to 2.75%. Replacing *p*-anisyl with *p*-Methoxy benzyl rendered the compound 4A4MBN as inactive while *m*-methoxy benzyl substitution made the compound 4A3MBN as potent against PAD4 with 42.46% inhibition. Replacing electron releasing group

methoxy from *meta* position of compound 4A3ASN with electron withdrawing groups like chlorine as in 4A3CLN and nitro as in 4A3NTN lead to drastic reduction in potency to 4.59 and 3.2% respectively. *Para* substitution of bromine (4A4BRN) and chlorine (4A4CLN) showed 21.73 and 15.55% inhibition. Potency was raised to 31.18 with chlorine substitution at *ortho* position in compound 4A2CLN.

p-methyl amino benzamide series: Many of the compounds in this series showed potent activity compared to their *p*-amino benzamide analogues due to optimal distance between position 1 and 3 of pharmacophoric groups. Also the ionizability of amine of phenyl amine at assay pH (8) is low (pKa ~4.6) compared to benzyl amine (pKa= 9.51), causing loss of positive charge at assay pH which is necessary to have hydrogen bonding interactions with Asp350, Asp473 and His47.

Derivative with *o*-methoxy benzyl substitution (M2MBN) showed greater activity against PAD4 with 72.9% inhibition. While *p*- methoxy benzyl substitution (M4MBN)) rendered the compound inactive. Position of methoxy substitution did not affect the activity as observed in anisyl substitution (compounds M2ASN, M3ASN and M4ASN). *p*-anisyl derivative M2ASN showed 40.22% inhibition while substitution at *meta* position (M3ASN) increased the potency slightly to 42.77% and *ortho* substitution (M2ASN) lead to 45.47% inhibition. Replacing electron releasing methoxy with methyl reduced the activity significantly. *p*-methyl derivative M4TON showed 23.69% inhibition and increase in potency to 28.2% was observed with *m*-methyl substitution (M3TON), while *ortho* substitution(M2TON) reduced the activity significantly (16.63%). Relatively increased potency of *ortho* derivative was also observed in chloro and nitro substitution. *p*-chloro phenyl derivative M4CLN showed 36.41% inhibition and activity was raised to 39.95% with *ortho* substituted derivative M2CLN while *meta* substituted compound M3CLN was moderately active (24.63%). Replacement of *p*-chloro by *p*-bromo reduced the activity to 21.8%. *m*-nitro derivative M3NTN was inactive while *o*-nitro derivative M2NTN showed 37.83% inhibition. Unsubstituted phenyl ring (MANIN) rendered the compound weakly active (4.7%) while 2-pyridyl derivative M2APN also displayed weak activity (8.55%) and 3-pyridyl derivative M3APN was inactive towards PAD4. Replacement of phenyl ring in MANIN with 1-naphthyl ring (MNAPN) made weakly active compound MANIN as potent inhibitor (46.3%). When aromatic ring was replaced by alkyl chain activity was decreased with increase in carbon chain. When *n*-propyl chain was replaced with *n*-butyl activity was reduced from 19.32 to 11.21% while cyclopropyl derivative MCPRN was inactive.

3.6.2.2 In silico, in vitro screening correlation

No correlation between docking scores and in vitro activity was observed. In benzamidine series no significant differences in Autodock, Glide XP and FRED docking scores were observed with amide, sulfonamide and urea linker. While urea linker containing compounds scored relatively low in Glide SP docking, the in vitro activity data of these compounds showed mild to strong activity. The lack of correlation may be ascribed to the poor ability of docking software to differentiate actives from inactives (**Figure 3.6-1**).
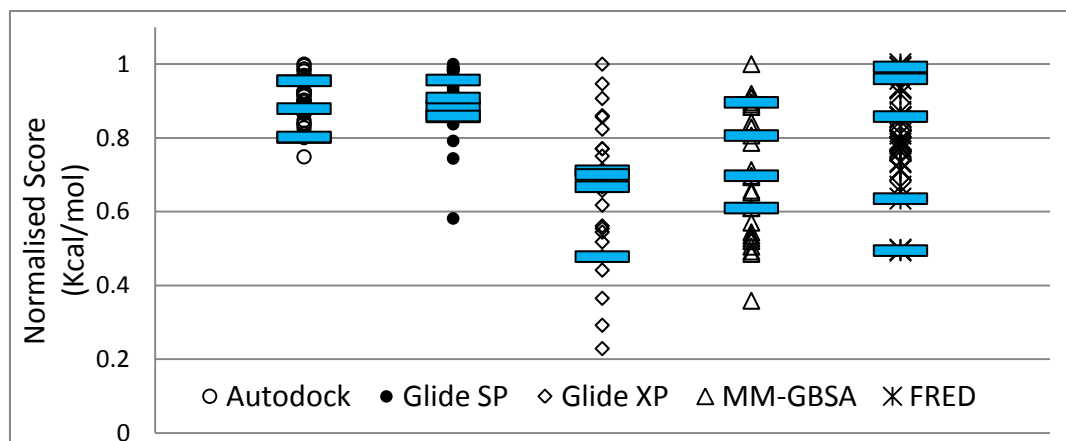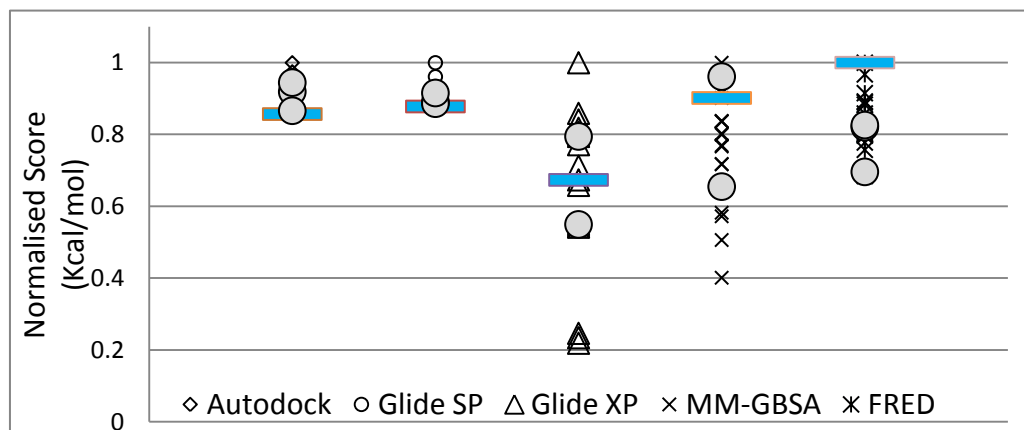

Figure 3.6-1 : Normalized docking scores of benzamidine series**.** A bar represents active compound.

In *p*-Amino-benzamide series, FRED chemgauss3 score and Autodock scores were comparatively less deviating than Glide XP scores and ΔG of Glide-XP pose estimated by MM-GBSA method. Autodock and Glide SP scored least for in vitro hit 4AMORN with docking score -4.62 and -3.90 Kcal/mol respectively; while Glide XP and FRED gave average docking score of -3.29 and -48.44 Kcal/mol respectively. Deviations of docking scores on normalized scale are represented in **Figure 3.6-2**.
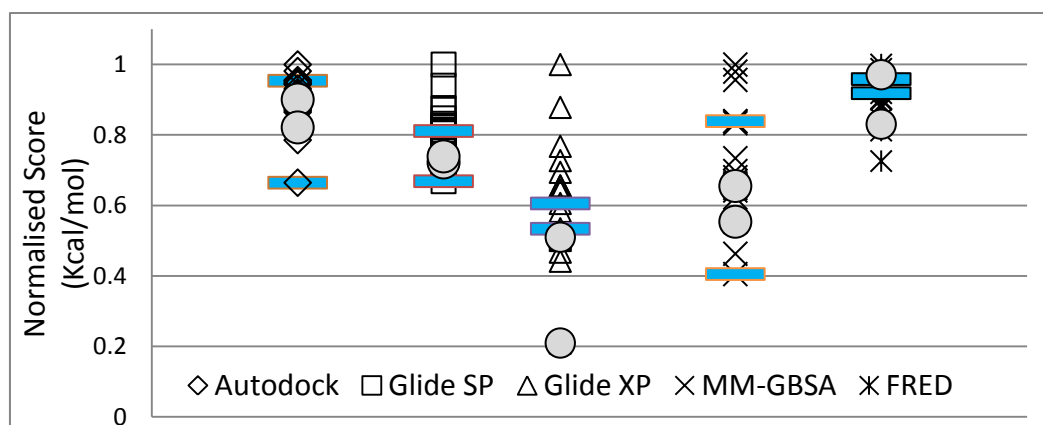
Figure 3.6-2: Normalized docking scores of *p*-amino-benzamide Series**.** Gray circles indicate inactive compounds (4A4MBN, 4APBN), while bars indicate most active compounds (4A3MBN, 4AMORN) identified by in vitro assay.

In *p*-(methylamino)-benzamide series, Autodock scores and Glide SP scores were comparatively less deviating then Glide XP scores FRED chemscore3 and Δ G of Glide-XP pose estimated by MM-GBSA method. FRED has given highest rank for in vitro hit M2MBN. Deviations of docking scores on normalized scale are represented in **Figure 3.6-3**.

Figure 3.6-3: Normalized docking scores of *p*-(methylamino)-benzamide Series. Gray circles indicate inactive compounds (M4MBN and MCRPN), while bars indicate most active compound (M2MBN) identified by in vitro assay.

3.6.2.3   Molecular simulation study

Potent PAD4 inhibitor M2MBN (72.9% inhibition) was selected for molecular simulation to investigate the factors responsible for enhanced binding and overall stability of predicted docking pose. Glide SP docked complex of M2MBN was selected. N-terminal domain of PAD4 was deleted and end terminals were capped with protein preparation wizard. With Desmond [3] system builder a periodic boundary model was set and solvated with Transferable Intermolecular Potential 3P (TIP3P) water model in a orthorhombic box of dimension 10 Å X 10 Å X 10 Å dimensions and α, β, γ angles of 90°, 90°, 90° respectively. System was neutralized by adding eight chloride ions, excluding from 7 Å region of docked ligand. Using OPLS 2005 force field system was then subjected to different steps of relaxation run prior to production quality. To remove bad contacts, to allow solvent to equilibrate and to hydrate the solutes in the system following different phases of relaxation protocol of Desmond was executed.

- Initial minimization for 10 steps of steepest-descent, followed by minimization over a maximum of 2000 steps of solute and ions with position restraint on protein and heavy atoms. The convergence criterion on the solute was set to 50.0 kcal mol^{-1} Å^{-1}.

- In the next step, all positional restraints were removed, and the system was minimized with 10 steps of steepest-descent followed by minimization over a maximum of 2000 steps with convergence criterion of 5 kcal mol^{-1} Å^{-1}.

- later a 12 ps simulation was performed using the NVT ensemble using a Berendsen thermostat, at 10 K temperature, 50.0 kcal mol^{-1} Å^{-1} force constant, velocity rescaling at every 1 ps and with restraints on solute heavy atoms.

- Followed by 12 ps simulation was performed at 1 atmospheric pressure using NPT ensemble with Berendsen thermostat, 10 K temperature, 50.0 kcal mol^{-1} Å^{-1} force constant, velocity rescaling at every 1 ps and with restraints on solute heavy atoms.

- The simulation was continued for another 24 ps under same condition as mentioned in 4th step but temperature was raised slowly from 10 K to 300 K.

- In last step of relaxation protocol all the restraints were removed and simulation was performed using NPT ensemble with Berendsen thermostat for 24 ps with 300 K temperature and 1 atmospheric pressure and relaxation time of 2.0 ps

Equilibrated system was then subjected to production phase using periodic boundary condition and NPT ensemble for 16 ns. The pressure of the cell was maintained at 1.013 bar with Martyna-Tobias-Klein pressure control method and temperature was maintained at 300K using maintained Nose-Hoover thermostat method. Covalent bonds with Hydrogen atom and heavy atoms were constrained with SHAKE algorithm with 2-fs integration step. Short range coulombic interaction cutoff was set at 9.0 Å. While long range coulombic interactions were handled with smooth Particle-Mesh Ewald method. The trajectories and the energies were recorded at every 4.8 and 1.2 ps, respectively. Using simulation quality analysis tool energetics of the simulation trajectories was analyzed. RMSD fluctuations, Hydrogen bonding and van der waals interactions were analyzed and plotted.

The overall stability of the simulation was evaluated by plotting total energy of the system as a function of time and RMSD measurement over time by superimposing trajectories into 0 ns system conformation. RMSD fluctuation of backbone heavy atoms was stable with RMSD fluctuation less than 2.4 Å and complex was energetically stable with <1% fluctuations of total energy of the system over the period of the simulation **Figure 3.6-4**.
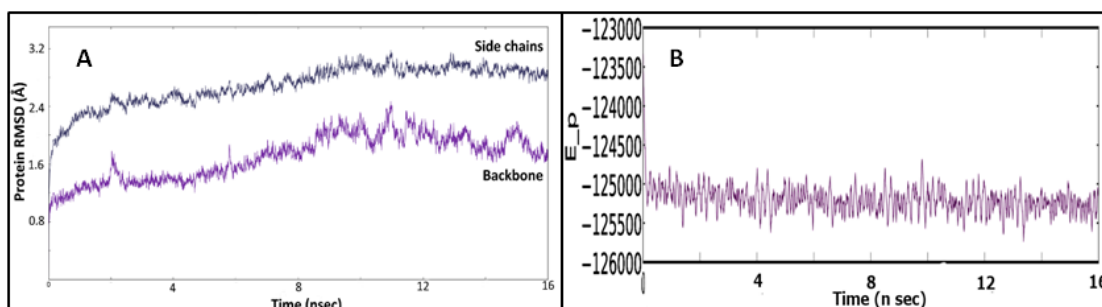
Figure 3.6-4: A. RMSD fluctuation of backbone and side chain heavy atoms and **B.** Change in total energy of the system during the course of simulation
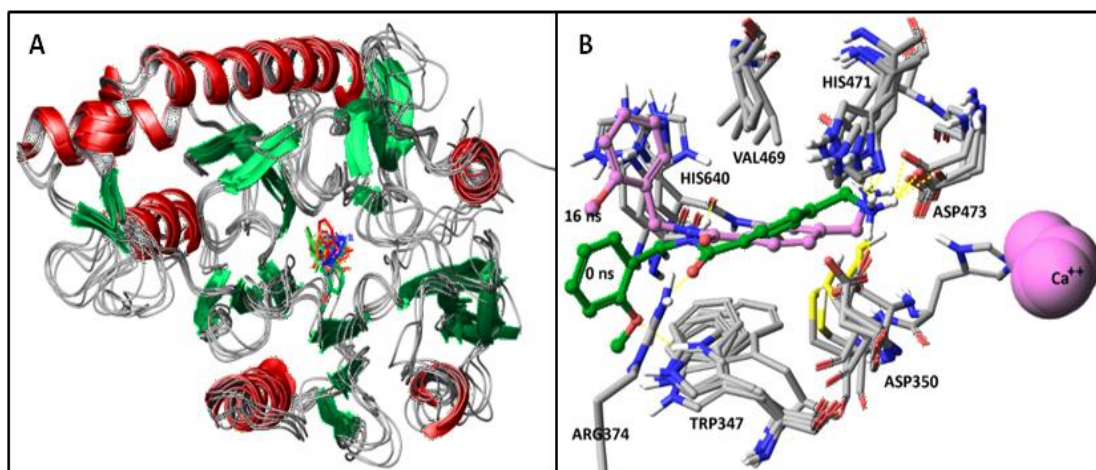
Figure 3.6-5: Superimposed 0, 2,5,10 and 16 ns trajectories **A.** total system, **B.** active site residues
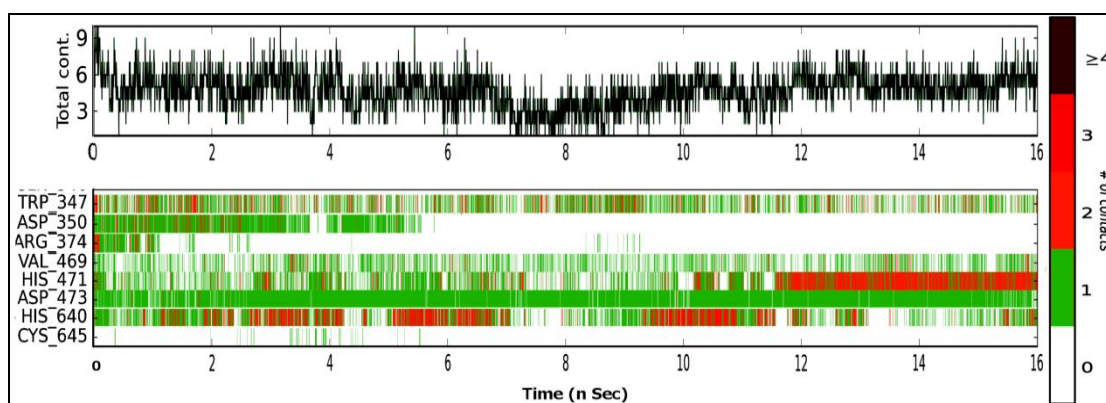
Figure 3.6-6: Interaction fluctuation plot of active site residues with ligand M2MBN.

Visual comparison of 0, 2, 5, 10 and 16 ns trajectories on analysis indicated the stability of the complex. Overall 3D conformation of the protein was stable. Side chain of Arg374 present at the opening of the cavity was moving away within a few ps of simulation. However for other active site amino acids, residual fluctuation was minimal as given in **Figure 3.6-5**. Terminal amine maintained hydrogen bonding interaction with Asp473 all the time however after 5 ns methyl amine part reorients and looses hydrogen bonding interaction with Asp350 and forms new Hydrogen bondinginteraction with adjacent His471. Hydrophobic interaction of phenyl ring was alternating between Val469 and Trp347. A variable pi-cationic (pi-pi) interaction with benzyl moiety and His640 was also observed throughout the simulation **Figure 3.6-6**. The simulation pose was similar to the pose produced during MM-GBSA energy estimation in **Figure 3.4-6**. The additional interaction due to benzyl group of M2MBN with His640 and optimal length between terminal amine and H-bond donating/ acceptor group could account for higher inhibitory activity against PAD4 in comparison to the other compounds.

3.6.3 References

[1] https://www.caymanchem.com/pdfs/700560.pdf (Accessed on dec 2012).

[2] Abad-Zapatero, C.; Metz, J.T. Ligand efficiency indices as guideposts for drug discovery. *Drug Discov Today.* **2005,** 10(7), 464-469.

[3] Desmond Molecular Dynamics System, version 3.6, D. E. Shaw Research, New York, NY, **2013**. Maestro-Desmond Interoperability Tools, version 3.6, Schrödinger, New York, NY, **2013**.

Chapter 4

Summary and Conclusion

With the intention to develop orally bioavailable new small molecular anti-rheumatic agents, two enzymes namely TACE and PAD4, which play a critical role in pathology of RA, were explored. Ligand and structure based approaches were followed to identify and develop inhibitors against these targets.

TNF-α Converting Enzyme (TACE)

TACE, a transmembrane protein, catalyzes synthesis of soluble TNF-α. To enhance potency and overcome toxicity of reported compound 2-(4-(but-2-ynyloxy)-N-methylphenylsulfonamido)-N-hydroxybenzamide (ASH), sulfonamide linker was replaced by urea and keto group.

To meet this objective, in-house virtual database was created and was subjected to oral bioavailability and toxicity prediction to ensure that the rejection at later stages is avoided. All the compounds were predicted to have good oral bioavailability according to Lipinsky rule of 5 and GSK oral bioavailability criteria. TACE database was then subjected for toxicity prediction using OSIRIS property explorer. All compounds were flagged with high risk of mutagenicity and medium risk of tumerogenecity due to presence of hydroxamic acid. However, literature reports indicate that it lacks toxicity in humans. *meta* and *ortho* methoxy substituted compounds were also flagged with high risk of reproductive toxicity although the group is normally present in most of the drugs used in clinic. Drug likeliness score of benzophenone derivatives was lower than diphenyl urea derivatives due to their low solubility.

Virtual database compounds were then subjected to in silico screening against TACE crystal structure 2OI0, using Autodock, Glide SP, Glide XP and FRED docking software. Docking scores of benzophenone derivatives were comparatively higher than urea derivatives because in benzophenones, carbonyl oxygen had interaction with Gly349 which was not observed in urea series. *o*-methoxy benzyl ether (2.5MBA_N) and *cyclo*-Hexyl ether (2.5CYH_N) substituted benzophenone derivative had higher docking scores in all the software. *m*-nitro phenyl substituted urea derivative (3.3NTPN) showed relatively higher docking score in its series but lower than reported molecule ASH in Glide and FRED docking.

All the virtual database compounds were synthesized using two different schemes. Benzophenone analogues were synthesized form phthalic anhydride using Fridel Craft acylation. Overall yield of the benzophenone series was determined by Mitsunobu reaction involving phenol and alcohol coupling. Rate of the reaction varied with bulkiness of second alcohol. Primary alcohol reactivity was higher compared to secondary and tertiary alcohols. Aryl alcohol also showed weak reactivity. Urea derivatives were synthesized from anthranilic acid and isocynates. Reactivity of isocynates varied with substitution. *para* substituted compounds showed higher reactivity compared to *meta* and *ortho.*

Biochemical TACE inhibition assay was carried out using FRET technique. Amount of FRET substrate hydrolysis by TACE was monitored by fluorescence based microplate reader. Synthesized TACE inhibitors were screened at 100 µM concentration. The activity was expressed as percentage inhibition and contribution of substitution was assessed using Potency Efficacy Index (PEI). Urea derivatives showed higher activity compared to benzophenone derivatives. Six urea derivatives 3.3MXPN, 3.4MXPN, 3.4NTPN, 3.3NTPN, 3.4FLPN, 3.2TOPN showed more than 50% inhibition, while benzophenone derivatives showed activity lesser than 50% inhibition. In terms of PEI *o*-Methoxy substituted urea derivative 3.2TOPN was most active and showed 73.85% inhibition of enzyme.

The in vitro in silico correlation was found to be very poor. This was interpreted mainly because of poor solubility of benzophenone analogues and testing the compounds at higher concentration. Effect of solubility on activity was evaluated by SEI. With SEI low soluble benzophenones were indicated to have false negative results due to poor solubility.

Peptidyl Arginine Deiminase 4 (PAD4)

PAD4 enzyme is identified as responsible factor for the generation of auto-antigens in RA by citrullination of proteins. Therefore inhibition of PAD4 serves as important target for blocking RA. To overcome toxicity issues associated with reported irreversible PAD4 inhibitors, small molecular arginine mimetic compounds which inhibit PAD4 in reversible fashion were planned to design.

Since PAD4 acts on arginine group of the proteins, a virtual database of 642 arginine mimetic compounds was created. Using structure based drug design, all molecules in virtual database were docked on PAD4 crystal structure. The chemical cluster of top scoring compounds in Glide and FRED was totally different although the pharmacophoric requirements were matching. Therefore, pharmacophoric features of both the clusters were

identified and molecules were designed by combining molecular features identified by both the software.

Designed molecules were predicted for oral bioavailability, toxicity and inhibitory potential using in silico methods. Benzamide and benzamidine series compound were predicted to have drug like bioavailability, however molecules with bulkier hydrophobic groups were predicted to have low aqueous solubility. Risk of mutagenicity, tumorigenic and reproductive toxicity was found in compounds with nitro, methoxy and methyl groups. Drug score of amidine series compounds were high except for those flagged with toxicity risks. Over all drug score for the *p*-amino benzamide analogues were higher compared to *p*-(aminomethyl)-benzamide analogues.

Designed molecules were subjected for virtual screening against PAD4 crystal structure 1WDA using Autodock, Glide SP, Glide XP and FRED docking software. Docking scores of *p*-(aminomethyl)-benzamide analogues were higher as compared to *p*-amino benzamide followed by amidines. Amidine group of benzamidines did not enter inside the cavity in Glide SP docking while in FRED, the amidine group only interacted with Asp350. In *p*-amino benzamide series many derivatives showed hydrophobic interaction with Val469, Trp347 and H-bond interaction with Arg374 and Arg639. However, the compounds did not interact with Asp473. Derivatives of *p*-(aminomethyl) benzamide series showed pose similar to substrate BAA and showed higher binding scores mainly due to optimal length between amide and amine group that helps for hydrogen bonding interaction with Asp350, Asp473, Arg374 and Arg639.

Molecules with free terminal amino or amidine groups were prepared using different synthetic strategies. All the benzamidines were synthesized using *meta* or *para* amino substituted benzonitriles. Nitrile group was later modified to amidine by reduction of intermediate amidoxime. Benzamidine series compounds were synthesized from *p*-amino benzoic acid and *p*-(aminomethyl) benzoic acid. Overall yield was determined by final deprotection step.

Biochemical PAD4 activity of synthesized PAD4 inhibitors was carried out using FRET technique at 100 µM concentration. Amount of ammonia released by citrullination of proteins was measured by coupling with FRET substrate. The activity was expressed as percentage inhibition. Amidine derivatives showed weaker activity and many were inactive. Highly active compound of the series A4ACNA (*m*-nitro cinnamoyl substituted) showed 51.77% inhibition. In *p*-amino benzamide series many derivatives showed weak to moderate

activity and the most active compound of the series 4AMORN (morpholin-4-yl moiety containing) showed 49.79% inhibition. In *p*-(aminomethyl)-benzamide series many compounds showed moderate activity while *o*-methoxy benzyl substituted compound M2MBN showed 72.9% inhibition.

No correlation was established between docking score and in vitro activity. However on visual observation of docking poses it was observed that many of benzamidine derivatives, amidine group was outside the cavity and in *p*-amino benzamide series the distance between amine and amide moiety was lesser compared to *p*-methylamino benzamide series molecules. Also the ionizability of amine of aniline at assay pH (8) is low (pKa ~4.6) compared to benzyl amine (pKa= 9.51) which causes loss of positive charge at assay pH which is necessary to have hydrogen bonding interactions with Asp350, Asp473 and His471.

To study the factors responsible for enhanced binding and stability of M2MBN – PAD4 complex, a 16 ns simulation study was carried out with TIP3P water model and OPLS force field 2005 using Desmond. Throughout the simulation terminal methylamine of M2MBN maintained hydrogen bonding interaction with Asp473 and after 5 ns a new H-bonding interaction was made with His471 instead of Asp350. The additional pi-pi or pi-cationic interaction of His470 with benzyl group clearly explains higher activity of M2MBN compared to other derivatives of this series.

Future Perspective

P resent study involved identification and lead development of TACE and PAD4 enzyme using computational techniques.

TNF-α Converting Enzyme (TACE)

Since TNF-α is an important chemokines in pathogenesis of RA, control of elevated TNF-α level in plasma by inhibiting TACE using small molecule served as attractive target. However, failure of compounds to show proof of concept in recent clinical trials, raised questions regarding the validity of this target. If target validity is addressed in future using more sophisticated techniques then once again the development of small molecular inhibitor against TACE can be revisited. In the present study, urea based TACE inhibitors were active and these need to be subjected to in vivo assays. Failure of benzophenone derivatives raised a question regarding contribution of solubility in biochemical assay, which needs to be addressed to minimize false negatives in drug discovery programs.

Peptidyl Arginine Deiminase 4 (PAD4)

Inhibition of PAD4 enzyme can block root cause of RA that is generation of citrullinated proteins, the auto-antigens of RA. Designed and synthesized reversible PAD4 inhibitors were found to be potent than any reported reversible PAD4 inhibitors. These inhibitor need to be tested in vivo to substantiate the reported in vitro activity. The approach used in design of these inhibitors, that is use of existing knowledge base of arginine mimetic compounds to develop novel inhibitor against PAD4, can be used in future design of new hits or in the development of hits to lead against various other targets such are proteases. Further, based on the present study, modification in the active structures can be done to enhance potency and reduce failures.

Journal Publications

- **Vadiraj Kurdekar**, Hemant R Jadhav (2014). A new open source data analysis python script for QSAR study and its validation. Medicinal Chemistry Research. pp. 1-9.

- **Vadiraj D Kurdekar**, Hemant R Jadhav. Small Molecular Inhibitors for the Treatment of Rheumatoid Arthritis: Progress so far. Mini-Reviews in Medicinal Chemistry. revision submitted

- Dipanwita Pati, Dilip Kumar Pandey, Radhakrishnan Mahesh, **Vadiraj Kurdekar** and Hemant R. Jadhav (2010). "Anti-Depressant-Like Activity of *Mucuna Pruriens*; A Traditional Indian Herb in Rodent Models of Depression" Pharmacologyonline, Vol. 1, pp. 537-551.

- Deepti Dongarwar, Inamdar Afeez, M. Dinakaran, **Vadiraj D. Kurdekar** and Hemant R. Jadhav (2010). "Synthesis and antioxidant potential of novel 2-isoxazole substituted chromenones synthesized by microwave irradiation". Asian Journal of Research in Chemistry, Vol. 3, Issue 1, pp. 106-109.

Manuscripts submitted:

- **Vadiraj D Kurdekar**, Hemant R Jadhav. Design, Molecular Simulation Study, Synthesis and Screening of Novel Benzophenone Hydroxamate Derivatives as TACE Inhibitors.

- **Vadiraj D Kurdekar**, Anupama Devarajan, Pankaj Wadhwa, Hemant R Jadhav. Synthesis and biological activity of Anthranilate urea hydroxamate TACE inhibitors.

- **Vadiraj D Kurdekar**, Hemant R Jadhav. Design, synthesis, in vitro screening and molecular simulation studies of benzylamine derivatives as potent reversible PAD4 inhibitors.

- **Vadiraj Kurdekar**, Hemant R Jadhav. Synthesis, Docking and In Vitro Screening of 4-amino Benzamides as Peptidyl Arginine Deiminase 4 (PAD4) Inhibitors.

- Satish N Dighe, **Vadiraj D Kurdekar**, Hemant R Jadhav. Synthesis, Docking and In Vitro Screening of 4-Benzamidine Sulfonamide Analogues as Peptidyl Arginine Deiminase 4 (PAD4) Inhibitors.

- Satish N Dighe, **Vadiraj D Kurdekar**, Hemant R Jadhav. Study of Arginine Mimetic Benzamidine Derivatives as a Lead for PAD4 Inhibitor Design.

  - Satish N Dighe, **Vadiraj D Kurdekar**, Hemant R Jadhav. Synthesis, Docking and In Vitro Screening of 4-Benzamidine Amide Analogues as Peptidyl Arginine Deiminase 4 (PAD4) Inhibitors.

Conference Publications

1. **Vadiraj D Kurdekar** and Hemant R. Jadhav, "Hit identification using molecular modelling approach: Case study of reversible PAD4 inhibitor", Poster presented at 'Overcoming the Bottlenecks in Drug Discovery and Development' RSC conference. Mar. 20-21, 2014, DSIN/Ranbaxy, Gurgaon.

2. **Vadiraj Kurdekar** and Hemant R. Jadhav, "Selection of pdb structure for docking: A case study with TACE enzyme', Poster presented at 'International Conference on Biomolecular Forms and Functions', Jan. 8-11, 2013, IISc Bangalore.

3. **Vadiraj Kurdekar** and Hemant R. Jadhav, "A New Python Script for QSAR study", Poster presented at 2nd International conference on open source for computer-aided Translational medicine, Feb. 22-25, 2012, IMTECH, Chandigarh.

4. Satish N. Dighe, **Vadiraj Kurdekar**, Hemant R. Jadhav, "Green chemical approach for N-tert-Butyloxycarbonylation of Amines", Poster presented at APTI 17th Annual National Convention 2012, Oct. 12-14, 2012, Manipal College of Pharmaceutical Sciences, Manipal.

5. **Vadiraj Kurdekar** and Hemant R. Jadhav, "Study on selectivity of TACE inhibitors against MMP13 and TACE: By Docking and QSAR approach", Poster presented at Contemporary Trends in Biological and Pharmaceutical Research (CTBPR -2011), Mar. 12-13, 2011, BITS, Pilani. Received Second Best Poster Prize.

6. Hemant R. Jadhav, **Vadiraj Kurdekar**, "Predicting Caco-2 cell Permeation Coefficient of Hydroxymate TACE Inhibitors by QSAR Analysis", Poster presented at 14th International Conference on Chemical Biology for Discovery: Perspectives and Challenges, Jan. 15-18, 2010, CDRI, Lucknow.

7. **Vadiraj Kurdekar** and Hemant R. Jadhav, "Role of Protein Arginine Deiminase-4 (PAD 4) in Rheumatoid Arthritis", National Conference on Emerging Trends in Life Sciences Research, Mar. 6-7, 2009, BITS, Pilani.

BIOGRAPHY OF VADIRAJ KURDEKAR

Mr. Vadiraj Kurdekar has completed his Bachelor`s degree in Pharmacy from Bapuji Pharmacy College, Davangere in the year 2005. He acquired his Master`s degree in Pharmacy in Pharmaceutical Chemistry from Government College of Pharmacy, Bangalore in year 2008. He had been working as research scholar at BITS-Pilani from 2008-2014. During his post-graduation studies he availed GATE scholarship and during his doctoral studies he was awarded with UGC-BSR Fellowship from University Grants Commission, India. He has presented his research finding in various national and internal conferences. He have also attended various workshops related to drug discovery and synthetic chemistry held in India and abroad.

BIOGRAPHY OF PROF. HEMANT R JADHAV

Dr. Hemant R Jadhav is presently working in the capacity of Associate Professor, Department of Pharmacy, Associate Dean, Academic Research (Ph. D. Programme) Division, Birla Institute of Technology and Science Pilani. He received his PhD degree for the thesis "Studies on Immunomodulatory potential of *Ocimum* species and functional interactions of constituents" in the year 2004 from National Institute of Pharmaceutical Educations and Research (NIPER), India. He is been involved in the research for last 14 years and in teaching for last 11 years. He has authored more than 15 research papers and 2 book chapters. He has presented his work in more than 30 various national and international conferences. He is an expert reviewer of many international journals like Journal of Pharmacy and Pharmacology, Medicinal Chemistry research, RSC Advances, Current Computer Aided Drug Design, Anticancer Agents in Medicinal Chemistry, Journal of Pharmacology and Pharmacotherapeutics, Toxicology and Industrial Health and Natural Product Research. He is a lifetime member of Association of Pharmacy Teachers of India. He has successively completed two projects, one of DST Fast Track under Young Scientist Scheme and other one of UGC. He is currently guiding four PhD students.

This Page is intentionally left blank