Molecular Characterization of Mitogen Activated Protein Kinase Kinase 3 (MKK3)

Researcher:
Awaisa Ghazal
Reg # 03/FBAS/MSBI/F07

Supervisors:
Dr. Asma Gul
Dr. Ishtiaq Qadri

Department of Bioinformatics &Biotechnology
Faculty of Basic and Applied Sciences

Medicine
Biochemistry
Heart failure, infarction
Mitogen activated Protein kinase
cycloheximide
Genetics

International Islamic university, Islamabad

Molecular Characterization of Mitogen Activated Protein Kinase Kinase 3 (MKK3)

Awaisa Ghazal
Reg #03/FBAS/MSBI/F07

DEDICATED TO

MY PARENTS AND MY SISTER ROMAISA GHAZAL

CERTIFICATE

Title of Thesis: Molecular Cloning and Characterization of Mitogen Activated protein Kinase kinase3 (MKK3)

Name of Student: Awaisa Ghazal

Registration No: 03/FBAS/MSBI/F07

Accepted by the Department of Bioinformatics, Faculty of Basic and Applied Sciences, International Islamic University Islamabad, in partial fulfillment of the requirements for the Master of Sciences in Bioinformatics.

Viva Voce Committee

Chairman _____

(Dr. Irfan Khan)

Supervisor _____

(Dr. Asma Gul)

External Supervisor _____

(Dr. Ishtiaq Qadri)

Internal Examiner _____

(Dr. Naveeda Riaz)

External Examiner _____

(Dr. Umer Khan Niazi)

(17th September, 2009)

ACKNOWLEDGEMENTS

All the acclamation is for all Almighty "**Allah**", Who bestowed the mankind with knowledge and wisdom. All the respect and honor for **Holy Prophet Muhammad (PBUH)** for enlightening us with the essence of faith in Allah and guiding the mankind, the true path for life.

These lines provide me the opportunity to rightly acknowledge, unmatched sincerity of Chairman Department of Bioinformatics Dr. Irfan khan, who encouraged and helped to conduct and complete this research work.

I feel pleasure to place on record my deep sense of thankfulness and indebtedness to my honorable supervisor **Dr. Ishtiaq Qadri** Director General NCVI (NUST Center Of Virology and Immunology) and **Dr. Asma Gul** Assistant Professor Islamic International University Islamabad for their kind supervision, dexterous, incentive teaching, scholarly and kind attitude and help through this research effort and in preparation of this manuscript.

I am honored to express my deep sense of gratitude and indebtedness to respected Dr. Sadaf Zaidi Post Doc. Fellow in NCVI and Dr. Umar Niazi Assistant Professor in NCVI for their supervision in my research work..

I feel to obligatory to express my thanks to my best friends Naghmana Kanwal, Farukh Javed Sheikh, Yasmeen Badshah, Tayyaba Bashir, Nadia Rasheed, Shazia Jabbar, Amara Jabin, Memoona Mushtaq, Sadia Pitafi, Lutuf Ali and Asim Ali. They are always there whenever I need a good friend.

I also would like to thanks my colleagues at NCVI for their cooperation during my research work.

And at last but not the least, I deem it a great honor and privilege to record profound sense of gratitude and extend my compliments to my affectionate **parents, sister (Romaisa Ghazal)** and brothers (**Affan Ahmed, Sufyan Ahmed, Salman Ahmed** and **Rizwan Ahmed**) for their encouragement and inspiration for higher ideas, moral and their love and affection..

"I do appreciate all of those who remembered me in their prayers and encouraged me through out my life and education career".

Awaisa Ghazal

LIST OF ABBREVIATIONS

ASK1	Apoptosis signal regulating kinase
BLAST	Basic local alignment Search tool
CO	Carbon monoxide
CSBP	Cytokine suppressive anti-inflammatory drug binding protein
EC	Enzyme Commission
ECM	Extracellular matrix
EGFR	Epidermal growth factor receptor
ERKs	Extra cellular Signal-regulated kinases
ExPASy	Expert protein analysis system
FLS	Fibroblast like synoviocytes
GO	Gene Ontology
HCC	Hepatocellular carcinoma
HO-1	Heme oxygenase
HOG	High Osmolarity glycerol
ICE-Protease	Aspartate specific protease
IFN	Interferon
IGF1R	Insulin like growth factor 1 receptor
IL-1	Interleukin 1
IL-1β	Interleukin-1β

IPTG	Isopropyl β-D-1-thiogalactopyranoside
JNK/SAPKs	C-Jun N-terminal kinases/stress-activated protein kinases
LB	Liquid broth
MAP	Mitogen-activated protein
MAPKs	Mitogen activated protein kinase
MEKs	Mitogen activated protein kinase kinases
MAPKK	MAPK Kinase
MAPKKK	MAPKK Kinase
MKK3	Mitogen activated protein kinase 3
MKK6	Mitogen activated protein kinase 6
MAPKAPKs	MAPK protein kinases
MKPs	MAPK-phosphatases
MLD-STZ	Multiple low doses of streptozotocin
MSA	Multiple sequence alignment
S-MSA	Structure based multiple sequence alignment
NCBI	National Center for Biotechnology Information
OA	Osteoarthritis
OMIM	Mendelian Inheritance in Man
PCR	Polymerase chain reaction
PDGFR	Platelet-derived growth factor receptor
PKC	Protein Kinase C

PSI BLAST	Position-Specific Iterated BLAST
PSSM	Position specific score matrix
RA	Rheumatoid arthritis
RASCAL	Rapid scanning and correction of multiple sequence alignment
ROS	Reactive oxygen species
RT	Reverse Transcription
SCF	Stem cell factor
TAK1	TGF-β-activated protein 1
TGFβ1	Transforming growth factor β1
TNFα	Tumor necrosis factor-α

TABLE OF FIGURES:

TABLE OF TABLES:

ABSTRACT

Mitogen activated protein kianses (MAPKs) comprise a family of well-conserved serine/threonine kinases. MAPK incite the amplification and specificity of the transmitted signals that eventually activate several regulatory molecules in the cytoplasm and in the nucleus to initiate cellular processes such as proliferation, differentiation, development, gene expression, metabolism to motility, survival and apoptosis depending upon cell type. MAPKs transduce signals in response to a wide variety of extracellular and intracellular stimuli, such as UV light, reactive oxygen species, osmotic and heat shock, cytokines, hormones, mitogenic stimuli etc. A MAPK signaling cascade is composed of a set of three evolutionarily conserved, sequentially acting kinases: a MAPK, a MAPK kinase (MAPKK), and a MAPKK kinase (MAPKKK). Five distinct groups of MAPKs have been characterized up till now: extracellular signal-regulated kinases (ERKs) 1 and 2 (ERKs1/2), C-Jun N-terminal kinase (JNK)/stress-activated protein kinase (SAPK), p38MAPKs isoforms α, β, γ, δ, ERKs 3 and 4, and ERK5. Activation of the p38 MAPK occurs through two different upstream kinases, Mitogen activated protein kinase kinase 3 (MKK3) and MKK6. MKK3 preferentially phosphorylates the p38α and p38β isoforms. HCV induced oxidative stress leads to activation of MAPKs especially p38 and JNK. To investigate role of MKK3 in activation of p38, MKK3 was cloned and sequenced. Based on the sequence, the predicted 3D model was proposed. 3D models were built for both activated and inactivated MKK3. Comparison of activated and inactivated models for MKK3 revealed some important differences in various loops of MKK3 predicted models. A critical stretch of three residues was found to be inserted at the carboxy tail of MKK3. The importance of these residues has yet to be elucidated. However these findings will help to identify HCV induced activators of MKK3 and its role in hepatocellular carcinoma.

Content

INTRODUCTION

1.1 Protein Kinase

Protein kinases transfer a phosphate group from ATP to an amino acid residue of a protein. Through changing enzyme activity, cellular location, or association with other proteins, phosphorylation is usually responsible for functional change in the target protein. Protein kinases constitute about 2% of human genome, 500 protein kinase genes are present in human genome. Bacteria and plants also contain protein kinases. Cellular pathways especially signal transduction and kinase activity pathways are regulated by kinases. A protein kinase may be a

- Tyrosine specific protein kinase such that Epidermal growth factor receptor (EGFR) Insulin receptor and insulin-like growth factor 1 receptor (IGF1R), Platelet-derived growth factor receptor (PDGFR),Stem cell factor (SCF) receptor etc.
- Serine-threonine protein kinase such as Mitogen activated protein kinase (MAPKs),
- Histidine specific protein kinase such as pyruvate dehydrogenase.
- Mixed kinase such as Mitogen activated protein kinase kinases (MEKs).

Integrated function of cells in an organism is achieved by highly interactive networks of protein kinases and other messenger systems. Mitogen-activated protein (MAP) kinase which is a family of protein kinase cascades, is a member of this repertoire of signaling molecules.

Phosphorylation by protein kinases is recognized as a major mechanism by which virtually every activity of eukaryotic cells is regulated, including gene expression, proliferation, motility, metabolism, apoptosis and membrane transport. Eukaryotic protein kinases share a conserved catalytic core. Eukaryotic protein kinases domains impart the catalytic activity. Three separate roles of the kinase domains are:

- Binding and orientation of the ATP (or GTP) phosphate donor.
- Binding and orientation of the protein (or peptide) substrate
- Transfer of the y-phosphate from ATP (or GTP) to the acceptor hydroxyl residue (Ser, Thr, or Tyr) of the protein substrate.

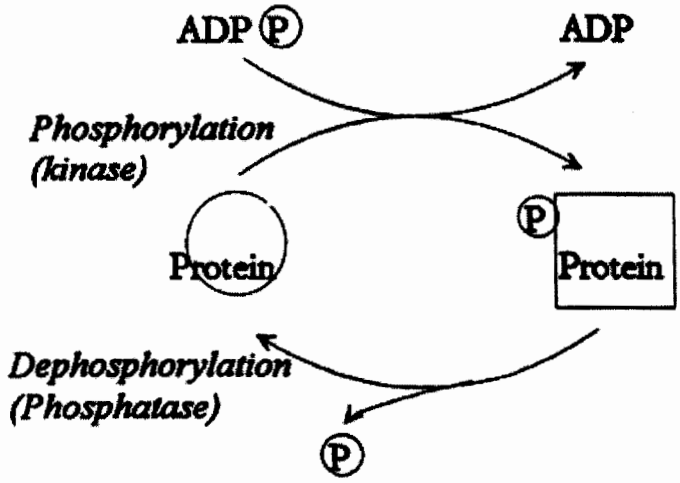


Figure 1.1: Protein Phosphorylation

Kinases use the y-phosphate of ATP (or GTP) to generate phosphate monoesters using protein alcohol groups (on Ser and Thr) and/or protein phenolic groups (on Tyr) as phosphate acceptors. There are a number of conserved regions in the catalytic domain of protein kinases. There is a glycine-rich stretch of residues which has been shown to be involved in ATP binding, at the N-terminal of the catalytic domain. There is a conserved aspartic acid in the central part of the catalytic domain which is important for the catalytic activity of the enzyme. Eukaryotic protein kinase (ePK) catalytic domain, revealed conserved subdomains that are never interrupted by amino-acid insertions, and identified highly conserved individual amino acids and motifs. 3dimensional structure of MKK3 is not available to elucidate its interaction with upstream, downstream kinases and to study its role in HCV induced pathogensis so there is a need to model 3D structure of MKK3.

feature of the MAP kinases, usually the kinase activation loop contains two activating phosphorylation sites, one a threonine and other a tyrosine, separated by a single variable residue. Mammalian MAP kinase phosphorylation motif is shown in table 1.1.

1.3 MAPKs Cascade

The description of additional pathways found initially in yeast gave the concept that there were multiple MAP kinases with individual regulation and functions i.e the cell wall pathway containing the kinase MPK1 and the high osmolarity glycerol (HOG) pathway containing the MAP kinase HOG1, and p38 enzymes, C-Jun N-terminal kinases/stress-activated protein kinases (JNK/SAPKs) and others in metazoans. Cluster of MAPKs which has been characterized till date is: extracellular signal-regulated kinases (ERKs) 1 and 2 (ERKs1/2), C-Jun N-terminal kinase (JNK)/stress-activated protein kinase (SAPK), p38MAPKs isoforms α, β, γ, δ, ERKs 3 and 4, and ERK5 (Chen *et al.*, 1992). A wide variety of different stimuli regulates activation of MAPKs: growth factor and probol esters preferentially activate ERK1 and 2, while stress stimuli ranging from osmotic shock and ionization radiation to cytokine stimulation activate the JNK and p38 kinases.

1.3.1 Mammalian MAP Kinases Phosphorylation Motif

MAP Kinase	Comments	P Site motif
ERK1	>80% identical to ERK2; abundant and ubiquitous	TEY
ERK2	Abundant and Ubiquitous	TEY
ERK3α	Immunobloting defects 63K and full length 95-100 K species;α is present in many species including human ~	SEG
ERK3β	75% identical to ERK3α	SEG
ERK1b	46K splice form of ERK1;comigrates with band originally named ERK4	TEY
JNK1	Multiple spliced forms	TPY

JNK2	Multiple spliced forms	TPY
JNK2	Multiple spliced forms	TPY
p38α	Sensitive to SB203580	TGY
p38β	Patially Sensitive to SB203580	TGY
p38β2	Sensitive to SB203580; lacks the 8 amino acid insertion unique to P38β	TGY
p38γ	Insensitive to SB203580	TGY
P38	Insensitive to SB203580	TGY
Mxi	P38α splice form lacking 80 C-t residues and containing 17 novel ones	TGY
EPK5	Involved in Proliferation	TEY
EPK7	May have a role in cell proliferation	TEY
NLK	Regulation of Wnt pathway; orthology of C elegans LIT-1; relative of Drosophila nemo	TEQ0
MAK	Expressed in cells undergoing meiosis in the testis but ovary	TDY
MRK	Expressed in embryonic myocardium; ubiquitious in adult tissues	TDY
MOK	Phorbol ester sensitive	TEY
KKIAKLRE	Cdc2-related kinase	TDY
KKIAMRE	T, Y mutants still activated in cells	TDY

Tabel 1.1: Different mammalian MAP Kinases and their phosphorylation motif (P site Motif).

Each family of MAPK is composed of a set of three evolutionary conserved sequentially acting kinases: a MAPK, a MAPK kinase (MAPKK, MEKs), a MAPK kinase kinase (MAPKKK, MEKKs). The MEKK activation leads to phosphorylation and activation of MEK, which then activates MAPK.

MAPK activity is stimulated by a MEK through dual phosphorylation on threonine and tyrosine residue located in activation loop of kinas. MEKs activity depends upon phosphorylation of serine or threonine residues in activation loops. The MEK kinase which are serine threonine kinases, are often activated through phosphorylation by small GTP-binding protein of Ras/Rho family in response to extracellular stimuli.(Ramer *et al.*, 1993, Rhodes *et al.*, 1990).

Dephosphorylation of various members of the MAP kinase pathway is carried out by phosphatases. MAP kinase phosphatase (MKP) family encompasses different dual-specificity phosphatases (Sun *et al.*, 1993). MAP kinases are inactivated at the appropriate times and locations by a substantial number of MKPs.

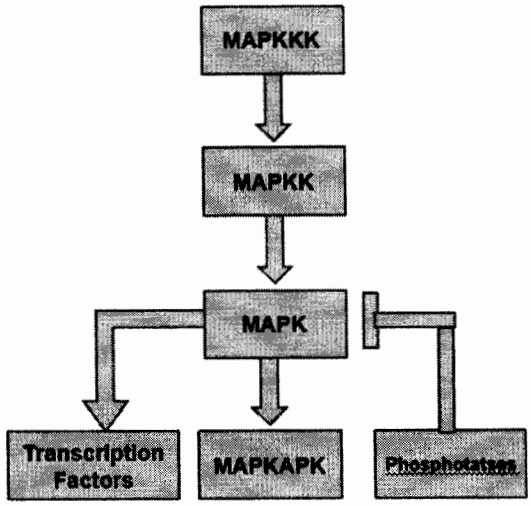


Figure 1.3: Diagrammatic representation of MAPK signaling cascade

1.4 Signal Integration and Specificity in MAPKs Cascade

Numerous ways of interactions among the cascades integrate responses and moderate outputs. MAP kinases exhibit overlapping substrate specificities (Lewis *et al.*, 1998, Waskiewicz *et al.*, 1997, Fukunaga & Hunter, 1993). Activation, localization, specificity, and targets of MAP kinases are facilitated by complexes of MAP kinases cascades. Integration may also occur early in the signaling pathway and at the top of the kinase module (Frost *et al.*, 1997, Mansour *et al.,* 1994, Rossomando *et al.,* 1994). Some MAP kinase cascades may be controlled by several, unrelated MEKKs and some MEKKs may regulate more than one cascade.

1.5 Mammalian MAPK Families

1.5.1 ERK1/2 Cascade

A classical mitogen activated protein cascade is mammalian ERK1/2, which encompasses MEKKs A-Raf, B-Raf and Raf-1, the MEKs MEK1, MEK2 and the MAPK ERK1 and ERK2. ERK1 (43kDa) and ERK2 (41 kDa) shows 85% identity,

with great identity in substrate binding region (Boulton *et al.,* 1990, Boulton *et al.,* 1991).In activation loop of both ERK1 and ERK2, Tyrosine and threonine, two phosphoacceptor sites, are separated by a glutamate residue to give the motif TEY and are phosphorylated to activate the kinases, (Payne *et al.,* 1991). Activity of the protein is not elevated by replacement of the two ERK2 phosphorylation sites with acidic residues (Zhang *et al.*, 1995). These are expressed to various extent in all tissues. For example, in several cells of neuroendocrine origin both may be equally expressed while in many immune cells ERK2 is the predominant species. Serum, growth factors, and probol ester strongly activate ERK1 and ERK2 but these are activated to lesser extent by cytokines, ligands of heteromeric G protein-coupled receptors, microtubule disorganization and osmotic stress. Upon stimulation ERK1 and ERK2 are accumulated in nucleus. Numerous substrates in all cellular compartment including cytoskeltal proteins (paxilin and neurofilaments), nuclear substrates (Pax 6, SRC-1, Elk-1, NF-AT, c-Fos, MEF2, STAT3 and C-Myc), membrane proteins (calnexin, CD120a and Syk), and several MKs are phosphorylated by activated ERK1 and ERK2. ERK1/2 signaling has been implicated a key regulator of cell proliferation and for this reason inhibitors of ERK pathway can be used as anticancer drugs.

1.5.2 C-Jun N-Terminal kinases/ Stress Activated Protein Kinases (JNK/ SAPK) Cascade

JNK1, JNK2, JNK3, also known as SAPKγ, SAPKα, SAPKβ respectively, are ubiquitously expressed. Like other MAP kinases, upon dual phosphorylation of tyrosine and threonine residues which are separated by a variable residue in activation loop, JNK/SAPKs are activated. Variable residue in all JNK/SAPKs is proline to give the motif TPY. UV irradiations, cytokines, growth factor deprivation, certain ligands for GPCRs, many other stresses, agents that interfere with DNA and protein synthesis, and to some extent growth factors, serum, and transforming agents regulate the activation of JNK/SAPKs. JNKs are stimulated by MEK4 and MEK7 which themselves are phosphorylated and activated by MLK2 and -3, MEKK1-4, TAO1, and -2, TPl-2, TAK1, ASK1, and -2 and DLK.

1.5.3 P38 Cascade

P38 (also known as SAPK2, RK, mHOG1 and CSBP), first described in 1994, is the classical member of the second MAPK pathway in mammalian cells. The p38 cascade consists of three conserved kinase modules that include several MEKKs: MEKK 1-4, MLK2 and -3, DLK, ASK1, TP12 and Tak1, the MEKS: MEK3, MKK6 (Enslen *et al.,* 2000, Han *et al.*, 1996 Stein *et al.,* 1996, De´rijard *et al.,* 1995) and MKK4, and the four known isoforms of P38 (α, β, γ, δ). Low molecular weight GTP-binding proteins from p21- activated kinases and the Rho family are the upstream of MEKs (Zarubin & Han, 2005). All p38 family members are well phosphorylated by MEK6 while p38α and p38β isoforms are preferentially phosphorylated by MEK3 (Enslen *et al.,* 2000). p38α and p38δ, in specific cell types, are activated by MKK4 which is an upstream kinase of JNK. MEK6 has a broader specificity than other MEKs because it can phosphorylate p38/ERK2 chimeras, and NLK *in vitro* (Ishitani *et al.,* 1999, Wilsbacher *et al.,* 1999). p38 isoforms have different affinity for upstream activators and downstream effectors and different tissue expression. Among p38 isoforms, p38α is the best characterized isoform. P38α was discovered independently in three context. It was found as: an activator for MAP kinase activated protein (MAPKAP) kinase-2 (Rouse *et al.,* 1994), a tyrosine phosphoprotein present in extracts of cells treated with inflammatory cytokines (Han *et al.,* 1994) and a target of a pyridinyl imidazole drug that blocked production of tumor necrosis factor-α (TNFα) and that is why it was named cytokine-suppressive antiinflammatory drug-binding protein or CSBP (Lee *et al.,* 1994). To identify the other three genes that encode members of the p38 subfamily: p38β, p38γ (SAPK3 or ERK6), and p38 δ (SAPK4), cloning strategies were used. The expression of p38*α* and p38*β* isoforms is ubiquitous in most tissues, whereas expression of p38*γ* is limited to skeletal muscle and that of p38*δ* to small intestine, lung, pancreas, testis and kidney. Activation loop in all of these kinases contains a TGY Motif which is phosphorylated by upstream kinases (MEKs). Pyridinyl imidazole inhibitors inhibit the activity of p38α and p38, but have no effect on p38γ- and p38δ. p38 family members are activated by a variety of agents including hormones, cytokines, osmotic and heat shock, GPCRs, and other stresses activate but mitogenic stimuli do not activate p38 cascade members.

In quiescent cells both the nucleus and cytoplasm contains p38, but the cellular localization of p38 upon cell stimulation is not well understood. p38 following the activation translocates from the cytoplasm to the nucleus according to some evidences, but other evidences suggests that activated p38 is also present in the cytoplasm of stimulated cells. p38 activity is critical for normal immune and inflammatory responses, indicated by a large body of evidence. Numerous extracellular mediators of inflammation including cytokines, chemoattractants, bacterial lipopolysaccharide and chemokines regulate p38 activity in macrophages, neutrophils, and T cells. Macrophage and neutrophil functional responses are also mediated by p38, these reponses include chemotaxis, respiratory burst activity, adherence, apoptosis and granular exocytosis. P38 also mediates apoptosis by regulating gamma interferon production and T-cell differentiation. Immune responses are also regulated by p38 through stabilizing specific cellular mRNAs involved in this process (Ono & Han, 2000).

Figure 1.4: Complexity of MAPKs signaling cascade. Green color block is representing the P38-MKK3 MAP kinase signaling cascade.

1.6 Mitogen Activated Protein Kinase Kinase 3 (MKK3)

MKK3 (also known as MAP2K3, MEK3) is an important kinase in p38 cascade as it is an upstream kinase which phosphorylates and activates p38 MAPKs. MEK3 is activated by a plethora of MEKKs: ASK1 (apoptosis signal regulating kinase), TAK1 (TGF-β-activated protein 1) and PTK1 and others, these MEKKs are activated in response to various physical and chemical stresses, such as UV irradiation, oxidative stress, ischemia, hypoxia and various cytokines, including tumor necrosis factor alpha (TNF-α) and interleukin-1 (IL-1). MEK3 does not activate ERK1/2 or JNK but it shows a high degree of specificity for p38.MKK3 and MKK6 both are regulator of p38. MKK6 and MKK3 have 82% identity at the amino acid level. However, at the DNA level especially at the C- and N-terminal regions there is less significant sequence homology. MKK3 and MKK6 also differ in cell and tissue expression. Within subdomain VIII of both MKK3 and MKK6, serine and threonine residues are activated upon phosphorylation by upstream MEK kinases (MAP3Ks). MEK3 preferentially phosphorylates the p38α and p38β isoforms. Selective recognition of the activation loop of p38 isoforms by MEK3 and formation of functional complexes between MEK3 and different p38 isoforms are thought to be the reasons for the specificity in P38 activation.

MEK3 is activated by dual phosphorylation of two residues, serine and threonine, in its activation loops. Activated MEK3 stimulate p38, p38 MAP kinase upon its activation regulates the activity of a wide range of transcription factors (p53, CHOP, ATF-1/2/6, MEF2, Sap1, NF-$_K$B, ELK1 and others), protein kinases (MNK1, MAPKAPK2, MSK1, PRAK), the microtubule associated protein Tau, cytosolic phospholipase A2, and some other proteins, which then further regulate the activity of their targets. In consequence, this complicated network of interacting proteins is responsible for different cell activities, like cytokine production, apoptosis, cell differentiation, cell-cycle arrest, tumour suppression and cell senescenc (Zarubin *et al.*, 2005, Olson & Hallahan, 2004).

MKK3 mediates Type I interferon signaling. MKK3 and MKK6 are rapidly activated during engagement of the Type I IFN receptor and play important roles in Type I IFN signaling and the generation of IFN responses. Type I interferons (IFNs) are

pleiotropic cytokines that mediate antitumor, antiviral, and immunomodulatory responses. These agents are used in clinical medicine for the treatment of certain malignancies, viral syndromes, andautoimmune disorders. IFNα results in activation of both MAP kinase kinase 3 (MKK3) and MAP kinase kinase 6 (MKK6). IFNα induces phosphorylation of MKK3 on serines 189. Such IFN-inducible activation of MKK3 and MKK6 is essential for downstream phosphorylation and activation of the p38 MAP kinase, Similarly, if MKK3 and MKK6 are deficient then IFN-dependent activation of the downstream effectors of p38, MAPKAPK-2 and MAPKAPK-3, is not induced, it demonstrate that the function of these MAP kinase kinases is required for full activation of the p38 pathway (Li *et al.*, 2006).

MKK3/P38 signaling plays an important role in mediating cytoprotection against oxidant-induced lung injury by carbon monoxide (CO). Lipids, proteins and DNA are demaged by increased oxidative burden that is mediated by exposure to high concentrations of O_2. In the lung this is followed by inflammation, fibrin deposition, and thickening of alveolar membranes. Edema formation, endothelial leakage, cytokine production, including interleukin (IL)-1 β, tumor necrosis factor (TNF) α, and IL-6, inflammatory cell influx into the airways, subsequent alveolar hemorrhage and diffuse alveolar damage are early evidences of injury. The stress-inducible gene heme oxygenase (HO-1) provides cytoprotection against oxidative stress. The precise mechanisms by which HO-1 provides cytoprotection were unclear. During the degradation of heme by HO, CO is released as a byproduct and plays a major role in providing the cytoprotection against oxidant-induced lung injury. CO also mediates cytoprotection of cultured epithelial cells against hyperoxic damage. CO mediates cytoprotective effects by selective activation of the MKK3/p38 protein MAP kinase signaling. CO preferentially activates MKK3 and p38β signaling while decreasing p38α expression. Anti-inflammatory effects of CO are mediated by the MKK3/p38 MAP Kinase signaling cascade as MAP kinase signaling cascade regulates the expression of many of pro-inflammatory cytokines like interleukin (IL)-1β , IL-6, and tumor necrosis factor (TNF)α. (Leo et al 2003).

MKK3 induces islet inflammation and hyperglycaemia in multiple low doses of streptozotocin (MLD-STZ). Hypoinsulinaema and hyperglycaemia are induced by destruction of pancreatic β-cells. In pancreatic islets, β-cell destruction is mediated

through the recruitment and activation of T cells and macrophages, and production of pro-inflammatory cytokines such as tumour necrosis factor-α (TNF-α), interferon-γ (IFN-γ) and interleukin-1β (IL-1β). The most ancient signal transduction pathways of MAPKs are widely used in many physiological processes. In pancreatic islets, IL-1β and TNF-α are known to activate p38 MAPK and JNK pathways. MKK3 signaling is responsible for the development of hyperglycaemia in the MLD-STZ mouse model. MCP-1 production is up-regulated by MKK3 signaling in damaged islets which in turn, induces macrophage infiltration and increased cytokines production causing destruction of the islet cells resulting in insulin deficiency and ultimately hyperglycaemia while leukocyte infiltration, production of pro-inflammatory cytokines and induction of hyperglycemia were not induced in MKK3 deficient islet cells, they show only initial phase of STZ induced islet cell apoptosis (Fukuda *et al.*, 2008).

MAPK are implicated in the regulation of many cellular processes in cardiac myocytes, including hypertrophy, apoptosis, and excitation-contraction coupling. In the heart, p38 is activated by a wide variety of contractile and hypertrophic agonists, chemokines, and reactive oxygen species (ROS). ANG II is a well-characterized inotropic and hypertrophic agonist that activates p38 via ANG II type 1 receptor (AT1R)-dependent signaling to tyrosine kinases, protein kinase C (PKC), and MKK3/6. MKK3/6/p38 signaling mediates cardic myocyte contractality on activation by ANG II (Sabri *et al.*, 2006).

p38 MAPK cascade, significantly contribute in transducing Transforming growth factor β1 (TGFβ1) actions. TGF β1, a 25-kDa homodimeric polypeptide, is a multifunctional cytokine that regulates a wide variety of cellular functions, including inhibition and stimulation of cell growth, as well as induction of cellular differentiation and extracellular matrix (ECM) protein synthesis. TGFβ1 also participates in pathological processes as a central mediator of tissue response to injury and progressive fibrosis in a variety of tissues, including the kidney. TGF β1 rapidly and strongly phosphorylates MKK3 and p38 MAPK and that activation of MKK3 is required for TGF β1-induced p38 MAPK activation in glomerular mesangial cells. TGF β1 selectively activates the p38 MAPK isoforms, p38 α and p38 γ, and both are dependent on the activation of MKK3 by TGF β1. Furthermore, MKK3 deficiency

selectively inhibited TGFβ1-stimulated up-regulation of pro-α1 collagen expression. MKK3 functions as a critical component of TGF β1 signaling pathway (Lin *et al.*, 2002).

Neonatal cardiomyocyte hypertrophic growth is regulated by MKK3/6-p38 signaling. Cardiomyopathy and dilated heart failure are mediated by stress-activated protein kinases (SAPKs) like c-Jun N-terminal kinase (JNK) and p38. Receptor tyrosine kinases, GPCRs, stretch, ischemia, and other general stressors each promote p38 and JNK activation in cardiomyocytes. The activation of p38 signaling cascade, at the level of the cell membrane, is initiated by MAPKKKs that induce the activation of the dual-specificity kinases, MKK3 and MKK6, in turn, these MAPKKs phosphorylate the p38 kinases. The major isoform of P38 expressed in the adult heart is p38 α. In cultured neonatal cardiomyocytes, hypertrophy and atrial natriuretic factor expression in vitro are induced by overexpression of activated MKK3 or MKK6, these MEKs further activate p38 which regulates myocyte growth response. So, neonatal myocytes are regulated by MKK3/6–p38 signaling (Liang & Molkentin, 2003).

Interleukin-1 (IL-1) activates p38 MAP kinase via the small G protein Ras, and this activity can be down regulated by another small G protein Rap. Transient transfection of cells with constitutively active forms of the known IL-1 signaling components MyD88, IRAK, and TRAF-6, or the upstream kinases MKK6 and MKK3,activated p38 MAPK. Dominant negative forms of these components inhibit activation of p38 MAPK by IL-1. The signaling pathway leading to the activation of p38 MAPK by IL-1 involves MyD88, IRAK, IRAK2, TRAF6, Ras, TAK1, MKK3 and MKK6, with a negative down-regulatory role for Rap, occurring at the level of Ras on the pathway .The study suggests that Ras takes part in the formation of a multiprotein signaling complex made up of IRAK, IRAK2, TRAF6, Ras and TAK1, which conveys the signal leading to the activation of p38 MAPK by IL-1 shown in Figure 1.3 (McDermott & Luke., 2001).

Rheumatoid arthritis (RA), a chronic inflammatory disease, marked with local invasion of bone and cartilage by synovial hyperplasia. Proinflammatory cytokines such as TNF-α and IL-1 that are responsible for activation of various intracellular signal transduction mechanisms are regulated by this disorder. Production of several mediators of inflammation and cartilage damage in synoviocytes and chondrocytes can be increased MAP kinases. Through a variety of transcriptional and translational

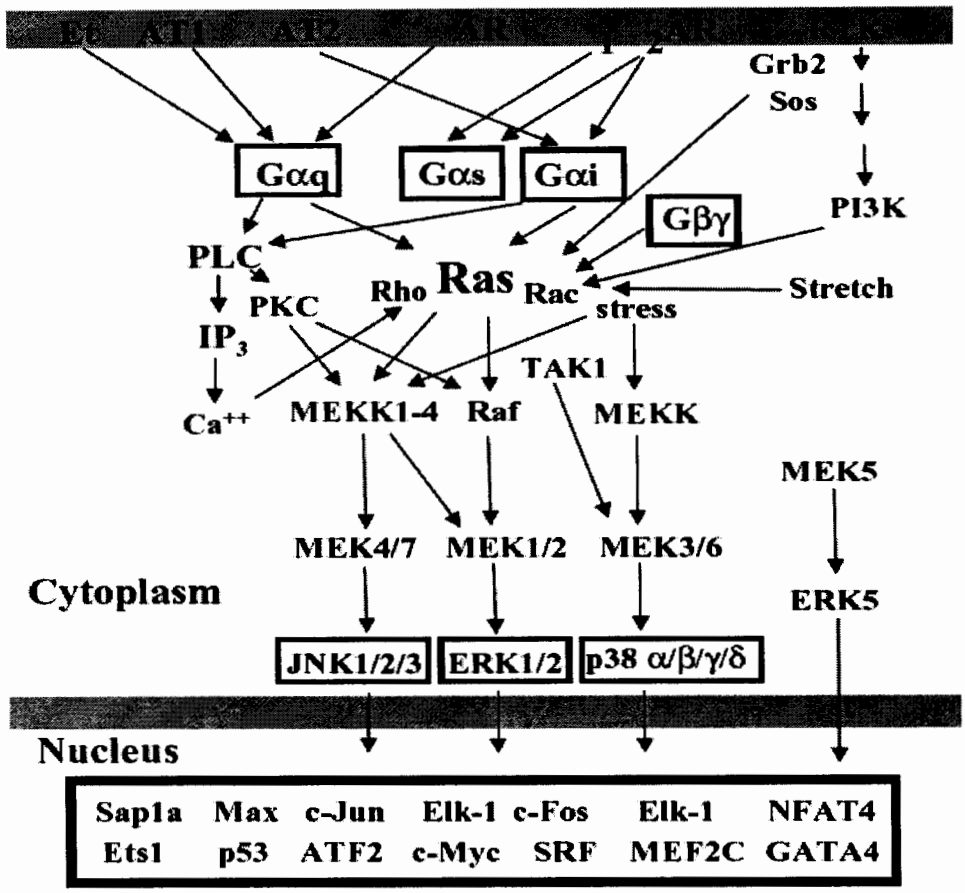


Figure 1.5: Diagrammatic representation of MAPK signaling pathways and their activating stimuli in a cardiomyocyte

mechanisms cytokine production can be regulated by p38 MAPK and it also involves in other inflammatory events, such as apoptosis, neutrophil activation, and nitric oxide synthase induction. Synovial inflammation and bone destruction in animal models of arthritis are also mediated by activation of p38. SB203580 inhibits the activity of p38 MAPK and reduces T lymphocytes, neutrophils and monocytes/macrophages proinflammatory cytokine production. p38 MAPK has two main upstream activators: MKK3and MKK6. RA synovial tissue and osteoarthritis (OA) activate MKK3 and MKK6. Western blot analysis and an anti-phospho-MKK antibody show rapidly increase in activation of MKK3 by IL-1 or TNF-α, while cytokine stimulation only modestly increases MKK6 phosphorylation. Therefore MKK3 and MKK6 are important regulators of p38 in fibroblast like synoviocytes and are activated in the synovium of patients (Chabaud-Riou & Firestein, 2004).

Apoptosis is a programmed cell death and it is an essential mechanism of eliminating aged or damaged cells and to maintain tissue integrity. Two ways for apoptosis are: ligands (death factors) binding to plasma membrane receptor (death factor receptor) and negative induction by the loss of suppressor activities. Apoptosis can be triggered by a variety of factors e.g. UV or γ radiation, heat shock, DNA injury, protease inhibitors, and factors. The common execution mechanism of apoptosis consists of the activation of cytosolic aspartate-specific protease (ICE-protease) termed caspases which can be regulated by various intracellular pathways. The activation of caspase is accompanied by characteristic morphologic changes of the cell, including cell shrinking, membrane blebbing, chromatin condensation, margination of nuclear matrix, formation of apoptotic bodies and DNA cleavage. Positive induction of apoptosis involves ligands such as tumor necrosis factor (TNF), TRAIL (TNFα-related apoptosis inducing ligands), or CD95 ligand/Fas ligand (CD95L/FasL), which are able to induce programmed cell death by acting on surface cell death factor receptors (TNF receptor-1, death receptor 3-6, and CD95/Fas recptor). The death ligand TNFα, can lead to activation of caspases in many cell types through the activiation of two members of mitogen activated protein kinase (MAPK) superfamily; cJUN N-terminal kinase or p38 family. The upstream mediator of p38 is MAPK kinase 3/6 (MKK3/MKK6) which can be activated by apoptosis signal regulating

kinase (ASK1). ASK1 also activates stress signaling kianse (SEK1/MKK4), the inducer of JNK after stress as shown in Figure 1.4 (Szabo & Tarnawsk, 2000).

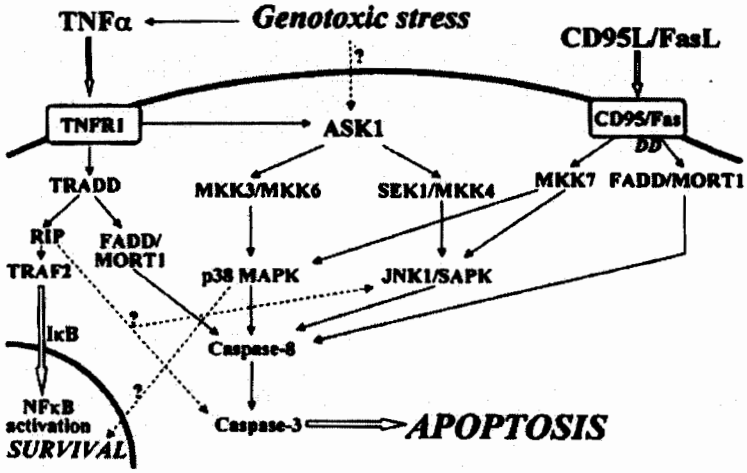


Figure 1.6: Diagrammatic presentation of general pathways of positive induction of apoptosis

MAPKs are activated by different type of viruses like HCV. P38 and JNK cascade of signaling events are triggered by HCV-induced oxidative stress. HCV-NS5A seems to play a critical role in the activation of p38 MAPK, JNK and AP-1 (Qadri *et al.*, 2004). MAPKs are induced by various HCV proteins. HCV-E2 induces activation of p38 MAPK and ERK. CD81 and low-density lipoprotein receptor (LDLR) are responsible for mediating this activation (Zhao et al., 2006). HCV-NS3 N-terminal peptide may up-regulate the activation of MAPK, but not affect the expression of MAPK (Feng et al., 2005). HCV-Core expression leads to deregulation of the mitotic checkpoint via a p38/PKR-dependent pathway thereby playing an important role in the viral pathogenesis of HCC (Spazian et al., 2006).

MAPKs induce various type of cancers. p38 has emerged as an important regulator of both embryonic development and cancer progression (Bradham *et al.*, 2006). HCV infects liver, once infection is established it leads to fibrosis, cirrhosis and hepatocellular carcinoma. Hepatocellular carcinoma (HCC) is one of the most common malignancies worldwide. Currently, no adjuvant treatment modalities have been conclusively shown to prolong survival in HCC. Through better understanding of the molecular basis of hepatocarcinogenesis, new preventative and treatment

modalities can be identified (Hung, 2004). p38 MAPK cascade may account to the unrestricted cell growth of human HCC (Iyoda *et al.*,2003). All of these studies are elucidating role of MAPKs in cancer but there is not a single study to reveal the role of MKK3 in induction of cancer specially hepatocellular carcinoma. The aim of study is to clone and sequence MKK3 and predict MKK3 structure that will facilitate to understand MKK3 interactions with its activators and substrates.

1.7 Protein Structure

Proteins fulfill several crucial functions, having catalytic, structural and regulatory roles in all organisms. Knowledge of the three dimensional structure of protein is a basic prerequisite for understanding their function. Protein three-dimensional structures are obtained using two popular experimental techniques, x-ray crystallography and nuclear magnetic resonance (NMR) spectroscopy.

1.7.1 X-Ray Crystallography

In x-ray protein crystallography, proteins need to be grown into large crystals in which their positions are fixed in a repeated, ordered fashion. The protein crystals are then illuminated with an intense x-ray beam. The x-rays are deflected by the electron clouds surrounding the atoms in the crystal producing a regular pattern of diffraction. The diffraction pattern is composed of thousands of tiny spots recorded on a x-ray film. The diffraction pattern can be converted into an electron density map using a mathematical procedure known as Fourier transform. To interpret a three-dimensional Structure from two-dimensional electron density maps requires solving the phases in the diffraction data. Once the phases are available, protein structures can be solved by modeling with amino acid residues that best fit the electron density map

1.7.2 Nuclear Magnetic Resonance Spectroscopy

NMR spectroscopy detects spinning patterns of atomic nuclei in a magnetic field. Protein samples are labeled with radioisotopes of C and N. A radio frequency radiation is used to induce transitions between nuclear spin states in a magnetic field. Interactions between spinning isotope pairs produce radio signal peaks that correlate with the distances between them. By interpreting the signals observed using NMR, proximity between atoms can be determined. Knowledge of distances between all

labeled atoms in a protein allows a protein model to be built that satisfies all the constraints.

1.7.3 Homology Modeling

Homology modeling techniques aim to build a 3D model of a protein of unknown structure (target) on the basis of sequence similarity to proteins of known structure (template). It is also known as comparative modeling. Homology modeling can be used in the interim to generate model (Mobarec *et al.,* 2009). Two proteins of similar function and origin are expected to have reasonable structural similarity. Therefore the known structure can be used as a template to model the structure of the unknown structure. Homology modeling produces an all-atom model based on alignment with template proteins.

Homology modeling is possible because a small change in the sequence usually results only a small change in the 3D structure (Chelvanayagam *et al.*, 1994). In general, key residues located in active site or important for the fold of homologues protein are highly conserved. Homology modeling can be helpful in progressing and testing hypothesis in structural biology, such as hypothesis about the location of ligand binding-site, substrate specificity (Klinkert *et al.,* 1994) and drug designing (Ring *et al.,* 1993).It can also provide starting model in X-ray crystallography and NMR spectroscopy. They can also be very valuable in situations where the protein structure is not available, or the quality of the available structure is not good.

AIMS AND OBJECTIVES

- **To amplify and clone MKK3 from human blood:** RNA will be extracted from whole blood of a healthy individual and will be subjected to cDNA synthesis. Thus cDNA obtained will be used for amplification. Amplified product will be cloned into vector.

- **To sequence cloned MKK3:** Positive clone will be screened and subjected to sequencing. Sequence will be analyzed and aligned with reported sequences.

- **To build 3D model on the basis of sequence through homology modeling:** Sequence obtained will be subjected to model building by using various bioinformatics tools. Models will be built for both activated and inactivated forms.

- **To reveal differences in activated and inactivated MKK3 models:** Models will be analyzed to investigate differences in different regions like ATP binding region, glycine rich region, activation loop and C terminal residues' conformation in activated and inactivated form.

MATERIAL AND METHODS

2.1 RNA Extraction

5CC Blood was extracted from a normal person and kept in EDTA vaccutainer. This blood was immediately proceeded to escape RNA degradation as half life of RNA is very short.

Total RNA from blood sample was extracted using TRIZOL reagent. To lyse the blood cells, 200µl blood was taken in 1.5ml eppendrof tubes containing 750 µl TRIZOL reagents. 20 µl of 5N glacial acetic acid was added and mixed vigorously for 15 seconds. For complete dissociation of nucleoprotein complexes, lysed samples were kept at room temperature for 5 minutes. Lysate were suspended with 200 µl chloroform and mixed vigorously for 15 seconds. The resulting mixture was kept at room temperature for 2-5 minutes and centrifuged at 14000 rpm for 15 minutes at 4°C in Centrifuge 5810R (eppendrof). Following the centrifugation, the mixture separated into a brownish phenol-chloroform phase, an interpahse, and a colorless upper aqueous phase. RNA remained exclusively in the aqueous phase while DNA and protein were in the interphase and organic phase. Aqueous phase was transferred to a fresh 1.5 ml eppendrof tube and RNA was precipitated by mixing with 500 µl of isopropanol. Mixture was kept at room temperature for 5-10 minutes and RNA pelleted by centrifugation at 14000 rpm for 8 minutes at 4°C in Centrifuge 5810R(eppendrof). RNA precipitate formed a white pellet at the bottom of tube. The supernatant was removed and RNA pellet was resuspended in 1 ml of 70% ethanol by vortexing. RNA suspension was centrifuged at 8000 rpm for 5 minutes at 4°C in Centrifuge 5810R (eppendrof). Ethanol was removed and RNA pellet was briefly air dried for 5 minutes. RNA was dissolved in diethyl pyrocarbonate (DEPC) treated water by passing the solution for a few times through a pipette tip and incubation for 10-15 minutes at 55-60°C. Extracted RNA was quantitated on BioPhotometer plus (eppendrof).

2.2 Oligonucleotides for MKK3

Oligonucleotides were designed by analyzing the homology in different reported MKK3 sequences in Genbank (NCBI) with following accession numbers: NM_002756, NM_145109, L36719, BC032478. The primers were designed to amplify complete coding sequence of MKK3 including initiation as well as termination codon. For identification of zero cutter restriction endonucleases NEBcutter V2.0 was used. Melting temperature, length and GC content of oligonucleotides were determined by OligoCalc.

EcoR1

MKK3F *AAA AAGCTT GCC GCC ATG GAG TCG CCC GCC TCG A*

HindIII

MKK3R *GGG GAATTC CTA TGA GTC TTC TCC CAG GAT CTC CTT CAC GAA G*

2.3 Reverse Transcription

RNA extracted was immediately subjected to cDNA synthesis by the technique of Reverse Transcription (RT). For cDNA synthesis, 7 µl RNA, DEPC H_2O and 1 µl MKK3 specific reverse oligonucleotides (10 picomole) were initially incubated at 70°C for 10 minutes followed by cooling on ice for 5 minutes. RT reaction was performed with 10µl of 5X RT buffer, 10mMdNTPs (Fermentas), 40U of RNAse inhibitors (Fermentas), 200U RT enzyme (Fermentas) in a final reaction volume of 50µl. The reaction was prepared on ice and mixed well (no vortex) and transferred directly to 2720 Thermal Cycler system (Applied Biosystem). The temperature profile for RT reaction was 42°C for 60 minutes. RT reaction was stored at -20°C prior to PCR.

2.4 Reverse Transcription Polymerase Chain Reactions (RT-PCR)

cDNA obtained was subjected to reverse transcription polymerase chain reactions. PCR reaction was performed with 5µl of RT-cDNA, 2mMdNTPs, 10 pmole forward and reverse MKK3 oligonucleotides, 2.5U of Taq polymerase (Fermentas), 1X Taq buffer (Fermentas), 25mM MgCl2 in a final reaction volume of 50µl. The reaction

was prepared on ice and mixed well (no vortex). Reaction was amplified in 2720 Thermal Cycler system (Applied Biosystem) for 35 cycles at the following thermal cycler profile of 94°C: 5 minutes, 94°C: 45 seconds, 66°C: 30 second, 72°C: 1 minute with a final elongation time of 10 minutes at 72°C.

2.5 Agarose Gel Electrophoresis:

The DNA was analysed on 1% agarose gel, prepared by dissolving 0.36 g agarose in 36ml of TBE (Tris Base, Boric Acid, EDTA pH:8.3) and boiling. 1.8ul of Ethidium bromide was added to it. The gel upon cooling to 50°C was poured into the gel caster of electrophoresis apparatus by Wealtec. The gel was then allowed to solidify. TBE buffer was filled in the chambers of the electrophoretic apparatus and the gel was placed in it allowing the buffer to make a thin film covering the gel. The DNA samples 5ul were prepared by mixing with 6x loading dye (5:1 ratio). Samples were loaded and the gel was electrophorased at a voltage of 80V, DNA bands were visualizd on Wealtec UV transilluminator.

2.6 Gel Purification:

DNA fragments were isolated from agarose gels by using Qiagen Gel Purification Kit. Briefy the PCR product was run on 1% agarose gel. The required fragment was cut out of the Gel under UV light and transferred to an eppendrof tube. Gel solublization buffer was added to the eppendrof tube containing the gel slice at a concentration of 30µl/10mg and was melt on Heat block (Wealtec) at 50°C for 15 minutes and was mixed after every 3 minutes to ensure gel dissolution. When gel slice was appeared to dissolve completely it was incubated for additional 5 minutes. The mixture was loaded into gel extraction column and centrifuged at 12000xg for 1 minute on table top centrifuge by Sigma. This step was repeated with remaining solution. Then 700µl of the wash buffer was added to column and incubated for 5 minutes. Column was centrifuged at 12000xg for 1 minute. Discarded the flow through and placed the column into new labeled eppendrof tube. 50 µl of pre heated double autoclaved water was added to the column for elution of purified DNA. Column with water was incubated for 1 minute and centrifuged at 12000xg for 2 minutes. The DNA was stored at -20°C until further processing.

2.7 Ligation

Ligation of PCR amplified fragment was done in pTZ57R/T vector supplied in InsTAclone ™ PCR cloning Kit (Fermentas) according to instructions supplied with kit. For cloning experiments, molar ratio for insert to vector was 3:1. The ligation reaction was carried out in a total volume of 30 µl at 4°C overnight.

2.8 Preparation of Competent Cells:

10 ml of LB medium was inoculated with cells from a single colony of E.coli Top10 host strain and incubated at 37°C overnight. 50ml LB medium was inoculated with 0.5ml of cells from the overnight culture. The culture was incubated at 37°C in shaker until the absorbance reached 0.6 at 600nm. The culture was transferred to a screw cap falcon tube and chilled on ice for 10 minutes. It was then centrifuged at 4000rpm at 4°C for 15 minutes in Centrifuge 5810R (Eppendrof) and the supernatant was discarded. The cells were suspended in 20ml ice cold sterile 50mM $CaCl_2$ solution and left on ice for 15 minutes. The suspension was then centrifuged at 4000rpm at 4°C for 15 minutes. The supernatant was discarded again and the cells were resuspended in 2ml 50mM $CaCl_2$ solution. The cells were stored at 4°C for further use.

2.9 Transformation:

10µl of the ligation mixture was added to 200µl of E.coli Top10 competent cells and incubated on ice for 40 minutes. It was then transferred to water bath at 40°C for 90 seconds for heat shock and immediately transferred to ice. After 5 minutes incubation on ice 800µl of LB medium was added to it and incubated at 37°C for 2 hours, without shaking. Then centrifuged at 13000rpm for I minute and cells were pelleted down. 700µl of culture was removed and pellet was re-suspended in remaining media. 150 µl of cells were spread on LB agar plate containing 50µg/ml Ampicillin, 20% IPTG and 20% X-GAL. The plated were incubated at 37°C overnight. Appearance of white colonies is measure of cells have been transformed with chromosomal insert in the Lac Z operon region of pTZ vector, while the blue colonies indicate non recombinant cells.

2.10 Screening of Transformants:

Single white colonies were picked and plated in duplicate on LB agar plates containing Ampicillin. After overnight incubation at 37°C clones were confirmed by colony PCR.

2.10.1 Colony PCR

Positive clones were confirmed by colony PCR. In this PCR, template was colony instead of cDNA or DNA. White colonies were picked through 10 µl pipette tip and dissolved scrupulously in 15µl double autoclaved H_2O. Colonies were denatured in 2720 Thermal Cycler system (Applied Biosystem) at the following thermal cycler profile of 95°C: 10 minutes, 25°C:10 minutes. Denatured mixture was centrifuged at 13000 rpm for 1 minute in table top centrifuge (Sigma) and supernatant was employed as template for colony PCR. Colony PCR reaction was performed with 5µl of template, 2mMdNTPs, 10 pmole forward and reverse MKK3 oligonucleotides, 2.5U of Taq polymerase (Fermentas), 1X Taq buffer (Fermentas), 25mM $MgCl_2$ in a final reaction volume of 20µl. The reaction was prepared on ice and mixed well (no vortex). Reaction was amplified in 2720 Thermal Cycler system (Applied Biosystem) for 35 cycles at the following thermal cycler profile of 94°C: 5 minutes, 94°C: 45 seconds, 66°C: 2720 second, 72°C: 10 minutes with a final elongation time of 10 minutes at 72°C.

2.10.2 Midi Preparation of Plasmid DNA by Alkaline Lysis Method:

Plasmid DNAs were isolated by alkaline lysis Method (Sambrook). Briefly a single Bacterial colony was used to inoculate 50ml of LB containing the Ampicillin and grown at 37°C with vigorous shaking for 12-16 hours (overnight). Shifted 30ml of the overnight culture into a Falcon tube and centrifuged at 4000rpm, 4°C for 5 minutes. Decanted the supernatant and resuspended the pellet in 600µl of ice cold alkaline lysis solution 1(1M Tris, 0.5M glucose, 0.5 M EDTA pH 8.0), incubated for 5 minutes at room temperature, then added 1200 µl of freshly prepared alkaline lysis solution 11(1% SDS, 1M NaOH). After 5 minutes incubation of lysate on ice and mixing by inverting the tube for several times, added 900 µl of alkaline lysis solution 111(5M potassium Acetate, Glacial Acetic Acid) and again mixed by inversion followed by 5 minutes incubation on ice. Centrifugation was done at 4000rpm, at 4°C for 5 minutes

and supernatant was extracted with an equal volume of phenol: chloroform (1:1). Extracted aqueous phase was precipitated in a fresh tube by adding 2ml of absolute ethanol. Precipitation was carried out at -20°C for one hour followed by centrifugation at 4000rpm at 4°C for 10 minutes. Rinsed the pellet with 1ml of ice-cold70% ethanol, air dried and resuspended in 200µl of double autoclaved water.

2.10.3 Restriction Endonuclease Digestion of DNA

DNA was digested by using restriction endonucleases EcoR1 (Fermentas) and HindIII (Promega) in 1X Fast digest reaction buffer. Digestion was carried out in a total volume of 50 µl at 37°C in incubator (Memmert) for two hours. The digested DNA was checked by agarose gel electrophoresis as described in Molecular Cloning (A Laboratory Manual) book.

2.10.4 Sequencing

Sequencing of cloned MKK3 was performed by CEQ 8000 Genetic Analysis System (Beckman Coulter).

2.11 Homology Modeling

Homology modeling/comparative modeling (Martí *et al.,* 2000) techniques aspire to build a 3D model of a protein of unknown structure (target) using the crystal structure information of corresponding template structure. In Homology modeling one or more known protein structures (template) similar to query sequence structure are identified then these template sequences are aligned with the query sequence. The template structure and sequence alignment are then used to produce a structural model of the target. The modeling process involved the following steps: identification of related structures (template), alignment between target and template structure, model building and refining and finally the evaluation of model. The homology modeling was accomplished by employing a number of methods.

2.12 Protein Sequence and Structure Data Resources

DNA sequence of MKK3 was obtained from sequencing of cloned MKK3 in pTZ57R/T vector. DNA sequence was subjected to translation through ExPAsy

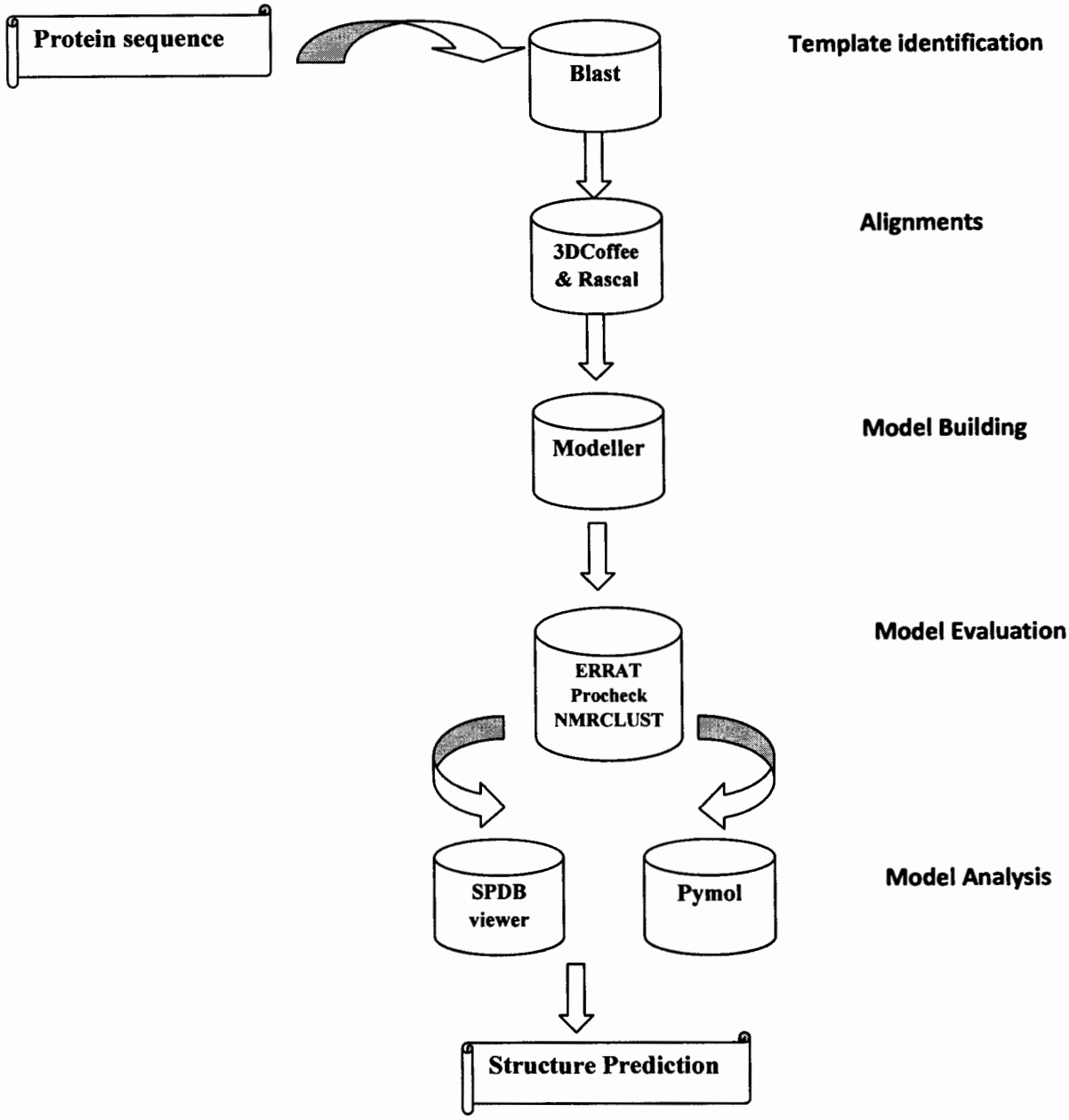


Figure 2.1: Methods involved in homology modeling

Translate. The ExPASy (**Expert Protein Analysis System**) proteomics server of the Swiss Institute of Bioinformatics (SIB) offers different tools and software packages dedicated to the analysis of protein sequences and structures as well as 2-D PAGE. ExPAsy Translate is a dedicated for the translation of a nucleotide (DNA/RNA) sequence to a protein sequence.

Translate is hosted by ExPASy proteomic server at: http://au.expasy.org/

Protein Data Bank is a comprehensive, public structural data repository for protein structures. It provides an up-to-date collection of three-dimensional structures of protein determined by X-ray crystallography and NMR with complete information. PDB is a worldwide central repository of structural information of biological macromolecules and is currently managed by the Research Collaboratory for Structural Bioinformatics (RCSB). In addition, the PDB website provides a number of services for structure submission and data searching and retrieval. RCSB PDB offers interconnection to external data pools e.g., Gene Ontology (GO), Enzyme Commission (EC), Online Mendelian Inheritance in Man (OMIM), National Center for Biotechnology Information (NCBI) repositories, etc (Kirchmair *et al.*,2008). PDB contains almost 58747 structures till date.

PDB is hosted by RCSB at: http://www.rcsb.org/pdb/home/home.do

2.13 Template Identification

Template identification is first step of homology modeling. The identification of all protein structures related to target sequence is greatly facilitated by the databases of protein sequences and structures along with software for scanning those databases. The suitability of the templates depends on the homology of the amino acid sequences between the target and templates which in turn affects the structure homology. As a rule of thumb, a database protein should have at least 40% sequence identity with the query sequence to be selected as template. Occasionally, a 20% identity level can be used as threshold as long as the identity of the sequence pair falls within the "safe zone" (Figure 2.2). The main tool used for comparing target protein with each sequence in database was BLAST (Basic local alignment search algorithm) which is

most commonly used and important program for searching homology between sequences (Altschul, 1990).

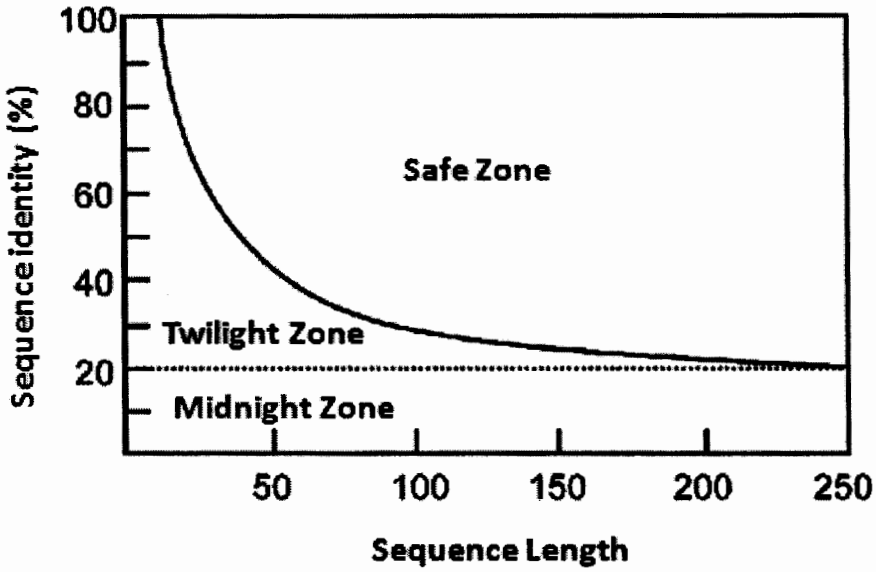


Figure 2.2: The three zones of protein sequence alignments. Two protein sequences can be regarded as homologous if the percentage sequence identity falls in the safe zone. Sequence identity values below the zone boundary, but above 20%, are considered to be in the twilight zone, where homologous relationships are less certain. The region below 20% is the midnight zone, where homologous relationships cannot be reliably determined (Krieger *et al.*, 2003).

2.13.1 BLAST

Various software tools are available for searching sequence databases; they all use some measure of similarity to distinguish between biologically significant and chance similarities. In *global alignment*, two sequences to be aligned are assumed to be generally similar over their entire length. Alignment is carried out from beginning to end of both sequences to find the best possible alignment across the entire length between the two sequences. *Local alignment*, on the other hand, does not assume that the two sequences in question have similarity over the entire length. It only finds local regions with the highest level of similarity between the two sequences and aligns these regions without regard for the alignment of the rest of the sequence regions e.g Needleman and Wunch algorithm (Needleman & Wunsch, 1970) and Smith Waterman algorithm (Smith & Waterman, 1981). These algorithms are impractical due to their computational cost. So these are replaced by heuristic algorithms like BLAST.

BLAST tools are used for searching DNA and protein databases for sequence similarities. BLAST uses BLOSUM-62, BLOSUM-45, PAM-120 etc scoring matrix. Position-Specific Iterated BLAST (PSI-BLAST) (Stephen *et al.,* 1997) methods are introduced to combine BLAST produced statistically significant alignments into a position-specific score matrix, and searching the database using this matrix. The PSI BLAST run speed per iteration is three times of the original speed and it is much more sensitive to weak but biologically relevant sequence similarities. Various versions of the BLAST algorithm are hosted by the National Center for Biotechnology Information (NCBI) at http://www.ncbi.nlm.nih.gov/blast. These can be used to search the various protein and nucleotide databases. Tamplate for tareget sequence was searched by using PSI-BLAST.

2.14 Sequence Alignment of Template and Target

Once the structure with the highest sequence similarity is identified as a template, the full-length sequences of the template and target proteins need to be realigned using refined alignment algorithms to obtain optimal alignment. This realignment is the most critical step in homology modeling, which directly affects the quality of the final

model. This is because incorrect alignment at this stage leads to incorrect designation of homologous residues and therefore to incorrect structural models. Errors made in the alignment step cannot be corrected in the following modeling steps. Therefore, the best possible multiple alignment algorithms should be used for this purpose. For closely related sequences with identity over 40%, the alignment is almost always correct. If the sequence identity between the template and target protein is low, especially in the "twilight zone" of less than 25% identity, then it can be difficult to obtain a good alignment (Krieger *et al.,* 2003). In that case more homologous templates can be added to the alignment to generate a multiple sequence alignment (MSA). Furthermore, using structural information can increase the accuracy of the MSA. To improve the quality of the MSA, structure-based comparison methods are used to derive structure-based MSA (S-MSA) or structural alignments. The MSA and S-MSA studies in this thesis were generated using the web based version of 3DCoffee (Olivier *et al.,* 2004).

2.14.1 3DCoffee

In many sequences analysis procedures key step is assembly of an accurate multiple sequence alignment (MSA). The web-based version of 3DCoffee is a tool for measuring high-quality multiple sequence alignments (MSAs). In order to increase the accuracy of the alignments 3DCoffee combines protein structures and sequences. Structures can be either directly uploaded into the server or provided as PDB identifiers. It does so by integrating three types of methods:

1) Sequence alignment methods.

2) Methods to compare two or more structures to generate a sequence alignment.

3) Threading methods to compare sequences and structures.

When one or more structures are used, then alignments produced are more accurate than similar alignments based on sequence information alone. In 3DCoffee structure–structure alignments are performed through SAP while structure–sequence comparisons are carried out through Fugue (Olivier *et al.,* 2004). The result is a structure based multiple sequence alignment if the right templates have been found. If some sequence has no templates, it will be entertained like regular sequence and

appropriate methods will be used for its alignments. 3DCoffee is hosted by TCoffee server at: http://tcoffee.vital-it.ch/cgibin/Tcoffee/

2.14.2 RASCAL

Errors in alignment were fixed by using RASCAL (Rapid Scanning and Correction of Multiple Sequence Alignments) to avoid any impediment in generating high quality model. RASCAL (Thompson *et al.,* 2003) is a knowledge-based approach which tries to correct potential errors remaining in the alignment. Indeed, building a multiple sequence alignment is based on a heuristic, and even if incorporation of prior information in the first steps of the alignment process helps the progressive alignment algorithm, some initial errors may remain. RASCAL uses a two step refinement process, it divides the alignment horizontally and vertically to formulate a lattice which will provide differentiation of well aligned regions (Thompson *et al.,* 2003). More efficient and reliable refinement strategy can be obtained by restricting the alignment correction to the less reliable regions.

To find badly aligned regions, the alignment is divided horizontally by clustering the sequences into subfamilies using Secator, a program based on a phylogenetic guide tree, and vertically using a sliding window in which a MD score is computed, for the whole alignment and for each subfamily. This identifies global and family-specific 'core block' regions corresponding to well-aligned zones in the alignment. A refinement strategy is then used to realigned inter core block regions, starting inside each family, then in a two-by-two family step, and finally between the global core blocks. RASCAL is hosted by PipeAlign server at: http://bips.ustrasbg.fr/PipeAlign

2.15 Model Building

Once optimal alignment is achieved, residues in the aligned regions of the target protein can assume a similar structure as the template proteins, meaning that the coordinates of the corresponding residues of the template proteins can be simply copied onto the target protein. There are various methods for building the model, e.g. modeling by segment matching, modeling by satisfaction of spatial restraints and rigid-body assembly (Martí-Renom *et al.,* 2000). It is important that a good modeling method should allow a degree of flexibility and automation, e.g. if the alignment is

changed a good modeling method should permit an easy recalculation of a model; it should capable of calculating models based on several templates; and the method should provide tools to incorporate restraints into the model based (e.g. predicted secondary structure features). In this study, models were built by using Modeller V 9.0 (Sali, 1993).

2.15.1 Modeller

A number of comprehensive modeling programs are able to perform the complete procedure of homology modeling in an automated fashion. The automation requires assembling a pipeline that includes target selection, alignment, model generation, and model evaluation. Some freely available protein modeling programs and servers are Modeller, 3D-JIGSAW and Swiss-Model (Schwede *et al.,* 2003).

Modeller is a tool for homology modeling. The user provides a predetermined sequence alignment of a template(s) and a target to allow the program to calculate a model containing all of the heavy atoms (non hydrogen atoms). The program models the backbone using a homology-derived restraint method, which relies on multiple sequence alignment between target and template proteins to distinguish highly conserved residues from less conserved ones. Conserved residues are given high restraints in copying from the template structures. Less conserved residues, including loop residues, are given less or no restraints, so that their conformations can be built in a more or less abinitio fashion. The entire model is optimized by energy minimization and molecular dynamics procedures.

Modeller is available free of charge for academic and non-profit organizations, and is available online at: http://salilab.org/modeller.

2.16 Model Evaluation

Modeling is an iterative process and model evaluation and refinement is an important part of it, as it helps in assessing the quality of the models produced and in selection of appropriate ones. Modeller generates any number of models for the target sequence. Models were evaluated by using: NMRCLUST and ERRAT

2.16.1 NMRCLUST

There were two criteria for selecting the high quality models, one was energy function and other was NMRCLUST (Lawrence *et al.,* 1996). Modeller ranks each model on its free energy, which can be used to choose models with lowest energies for further analysis. Modeller generates an ensemble of models so there is a need to select a single representative structure, or a subset of structures, from such an ensemble. NMRCLUST (Lawrence *et al.,* 1996) is a clustering program that clusters an ensemble of structures into a set of conformationaly related subfamilies and selects a representative structure from each cluster.

NMRCLUST takes the models generated by Modeller as input, and divides the models into clusters with a representative model for each cluster. So it helps in picking out the representative models for further evaluations.

2.16.2 ERRAT

ERRAT (Colovos & Yeates, 1993) is statistical tool that verifies protein structure and is excellent for evaluating the progress and refinement of generated crystallographic model. In fact, due to energetic and geometric effects, different types of atoms are non-randomly distributed with respect to each other in proteins. Three type of atoms: carbon (C), nitrogen (N), and oxygen (O) are present in protein with six different combinations of noncovalently bonded interactions (CC, CN, CO, NN, NO, and OO). The quality of PDB is determined by characterizing the set of pairwise interactions from nine-residue sliding windows. More randomized distributions of the different atom types is a result of errors in model building, statistical methods can be used to distinguish it from correct distributions. Errors in protein structures may be:

1. Mistracings of residues due to backbone connectivity error

2. Misalignments of residues

3. Misplacements of side chains

ERRAT works by analyzing the statistics of non-bonded interactions between different atom types. A single output plot is produced that gives the value of the error function *vs.* position of a 9-residue sliding window (Colovos & Yeates, 1993). ERRAT can identify residues with errors, which can be fixed by improving the

alignment in those areas. ERRAT is a sensitive technique and is good for identifying incorrectly folded regions in protein models.

ERRAT is available at: http://nihserver.mbi.ucla.edu/ERRATv2/

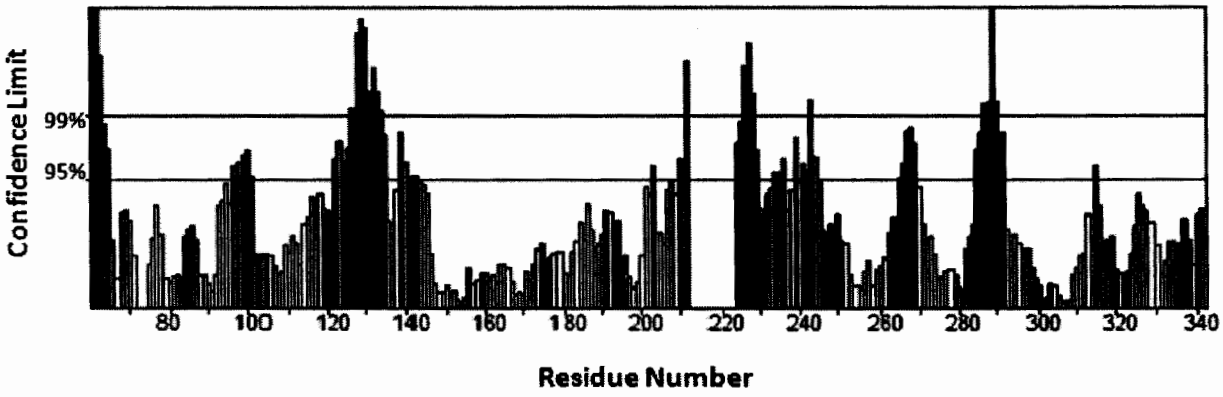


Figure 2.3: ERRAT plot for a MKK3 model. The black columns identify the problem areas.

2.17 Model Analysis

After model refinements, models were visually inspected by SwissPDB viewer (Guex & Peitsch, 1997) and PyMol (DeLano, 2002).

Swiss-PDB Viewer is a strauctur viewer for multiple platforms. It is essentially a Swiss-Army knife for structure visualization and modeling because it incorporates so many functions in a small shareware program. It is capable of structure visualization, analysis, and homology modeling. It allows display of multiple structures at the same time in different styles, by charge distribution, or by surface accessibility. It can measure distances, angles, and even mutate residues. In addition, it can calculate molecular surface, electrostatic potential, Ramachandran plot, and so on. The homology modeling part includes energy minimization and loop modeling. Swiss PDB viewer is hosted by ExPAsy server at: www.expasy.ch/spdbv/

PyMol is a molecular visualization system created by Warren Lyford DeLano and commercialized by DeLano Scientific LLC. It is a useful tool and is available freely to educational and scientific communities. It is capable of producing high quality 3D structure of small as well as biological macromolecules like proteins. PyMol is excellent for Viewing 3D Structures, giving presentations, rendering figures, sharing visualizations, exporting geometry and animating molecules. Almost a quarter of the scientific published images were made by using PyMol. PyMol is available at: www.pymol.org

RESULTS AND DISCUSSIONS

MKK3 is an important MAPKK of p38 signaling cascade. It is physphorylated in its activation loop by specific MAPKKKs and it preferentially phosphorylate p38 alpha. It may have role in induction of various types of cancers. For investigating role of MKK3, there was a need to clone it. For cloning RNA was extracted from a healthy individual blood and this RNA was subjected to reverse transcription polymerase chain reactions.

3.1 RNA Extraction and RT-PCR

Total RNA was extracted from the blood of healthy individual by using Trizol reagent. Trizol is phenol-based reagent contains a unique combination of denaturants and RNase inhibitors. The blood cells were lysed in the Trizol Reagent, and the mixture was separated into three phases by mixing with the chloroform followed by centrifugation. Isopropanol was used to precipitate the RNA from aquous phase. The entire procedure was completed in one hour and produce high yields of total RNA. RNA isolated from blood was subjected to the RT-PCR assay. The correctly sized DNA product (1044bp) was amplified by using specific primers. The amplified product of MKK3 was analyzed on 1% agarose gel as shown in Figure 3.1. MKK3 was amplified in bulk (100µl) and purified by using Qiagen Gel purification kit for cloning. The purified product of MKK3 was analyzed on 1% agarose gel Figure 3.2.

3.2 Cloning

The purified product was then used to clone into pTZ vector provided with Fermentas TA cloning Kit and transformed into E-coli Top10 competent cells. TA cloning of PCR product is based on the presence of 3'A overhangs. *Taq* polymerase has a non template-dependent terminal transferase activity that adds a single deoxyadenosine (A) to the 3′ ends of PCR products. The linearized vector supplied in the kit has single, overhanging 3′ deoxythymidine (T) residues. This allows the ligation of PCR Product with the vector. Recombinant clones were selected based on blue/white

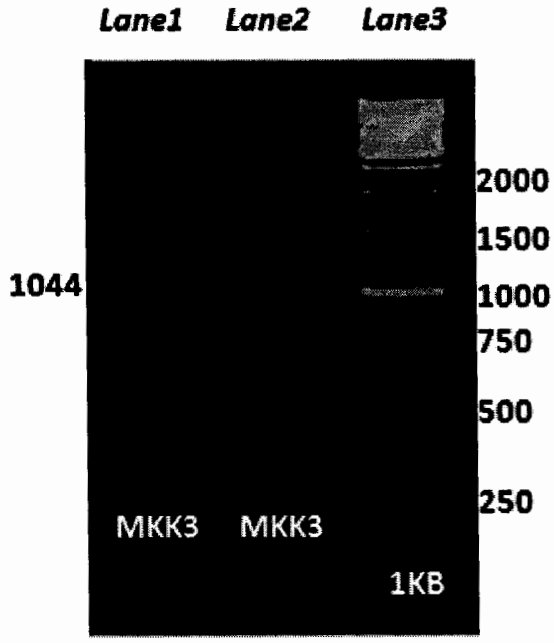


Figure 3.1: Agarose gel electrophoresis of RT PCR products obtained with specific oligonucleotides for MKK3. Lane 1 and lane 2: PCR product of MKK3 gene amplified from cDNA of same individual; lane 3: 1KB DNA Ladder (Fermentas).

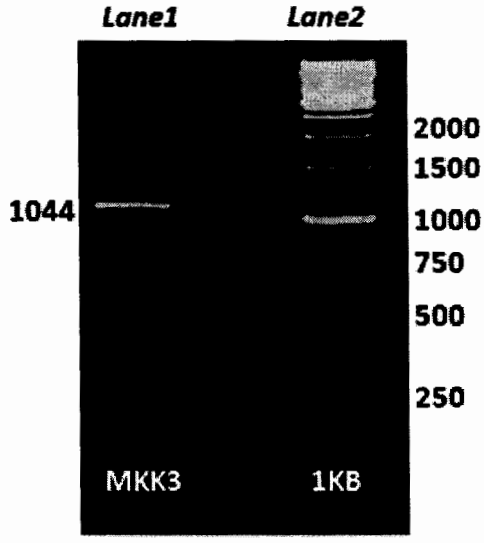


Figure 3.2: Agarose gel electrophoresis of RT-PCR purified product of MKK3. Lane 1: Purified PCR product of MKK3 gene; lane 2: 1KB DNA Ladder (Fermentas).

screening. The presence of *lac* operon in the vector is the molecular basis of blue/white screening. The LacZ protein is present with an internal multiple cloning site (MCS). The MCS can be cleaved by different restriction enzymes so that the foreign DNA can be inserted within the *lacZ*-α gene, thus disrupting the production of functional β-galactosidase. X-gal, a colorless modified galactose sugar is used for screening which metabolized by β-galactosidase(α subunit encodes by vector (pTZ57R/T) while Ω subunit encoded by the chromosome of the host strain to form functional β-galactosidase enzyme) & form an bright blue product (5-bromo-4 chloroindole). Isopropyl β-D-1-thiogalactopyranoside (IPTG), used as the inducer of the. The presence of blue color in the colonies shows that the colonies contain vector without insert. While white colonies shows the insertion of foreign DNA in the vector (Figure 3.3).

3.2.1 Confirmation of Positive clones by Colony PCR

White colonies were replicated (Figure 3.4) and colony number 7, 23, 39 and 51 were selected and colony PCR was performed by using MKK3 specific primers. Colony PCR is the useful tool for the screening of bacterial clones for correct ligation products. Selected colonies of bacteria were picked and denatured in autoclaved water and serve as DNA template in PCR reaction. Thus colonies which give rise to an amplification product of the expected size were likely to contain the DNA sequence of MKK3. The PCR amplified product of the above colonies was loaded on 1% agarose gel and visualized under UV light. Except colony number 23 all the other colonies 7, 39 and 51 were likely to contain MKK3 as they showed the specific size bands (Figure 3.5).

3.2.2 Confirmation of Positive clones by Restriction Endonuclease Digestion

For further confirmation of positive clones plasmid was isolated from the colonies (7, 39, and 51) which shows positive result after colony PCR by alkaline lysis method and verified on agarose gel (Figure 3.6). The plasmids were then digested with Restriction endonucleases BamH1 and EcoR1 which recognize a short DNA sequence specific for them. The digested products were electrophorased on 0.8 % agarose gel to determine the size of DNA fragments generated from digestion. The plasmid isolated

from colony number 51 clearly shows two distinct fragments of MKK3 (1044bp) and pTZ vector (3930bp) on agarose gel (Figure 3.7) and was named as MKpTZ51. These results confirmed that MKpTZ51 contains the positive clone of MKK3.

3.2.3 Sequencing

Plasmid was isolated from MKpTZ51 construct by using Invitrogen Plasmid preparation kit (Figure 3.8) and was proceed for sequencing.

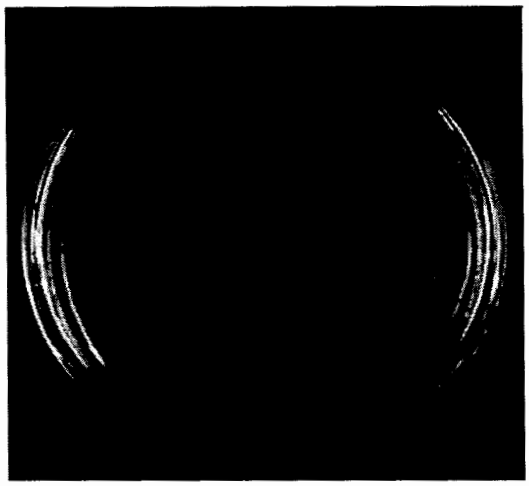


Figure 3.3: Agar plate with transformed and untransformed cells

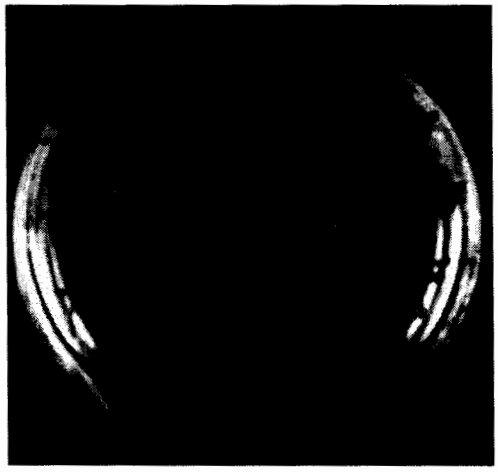


Figure 3.4: Agar plate representing replica of white colonies.

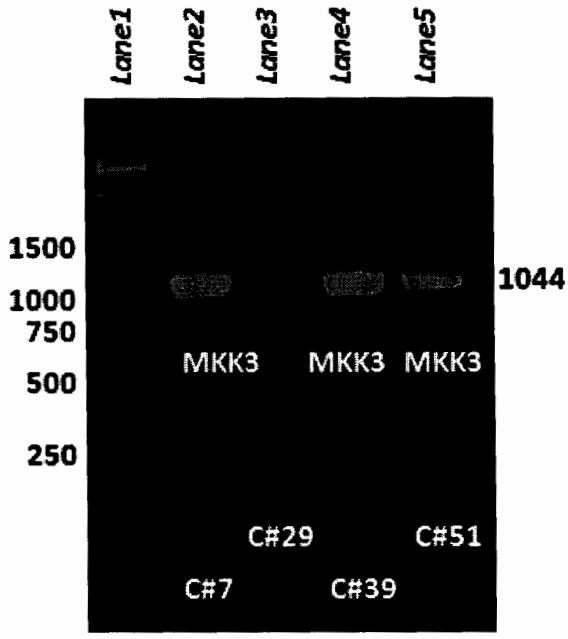


Figure 3.5: Agarose gel Electrophoresis of colony PCR Products obtained with specific oligonucleotides for MKK3.Colony PCR products were amplified by using colonies as template. Lane 1: 1KB DNA Ladder (Fermentas); Lane 2, 3 4 5: colony PCR product using colony No.7, 23, 39, and 51 as a template respectively.

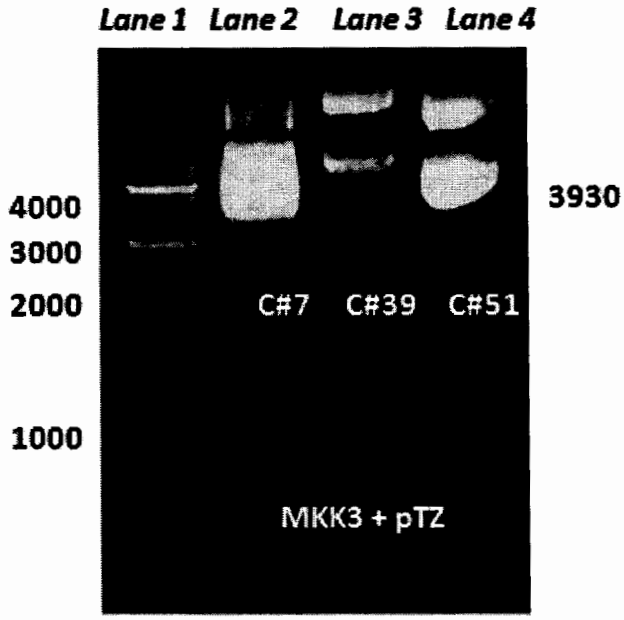


Figure 3.6: Agarose gel electrophoresis of Plasmid Preparation. Lane 1: 1KB DNA Ladder (Fermentas). Lane2, 3, 4: plasmids isolated from C# 7, 39, 51 respectively.

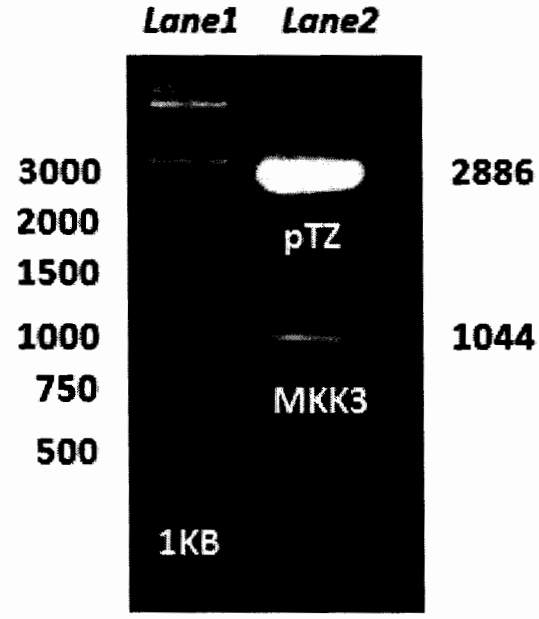


Figure 3.7: Agarose gel electrophoresis of restriction endonucleases digestion of MKpTZ51. Lane1: 1KB DNA Ladder (Fermentas). Lane2: digested product.

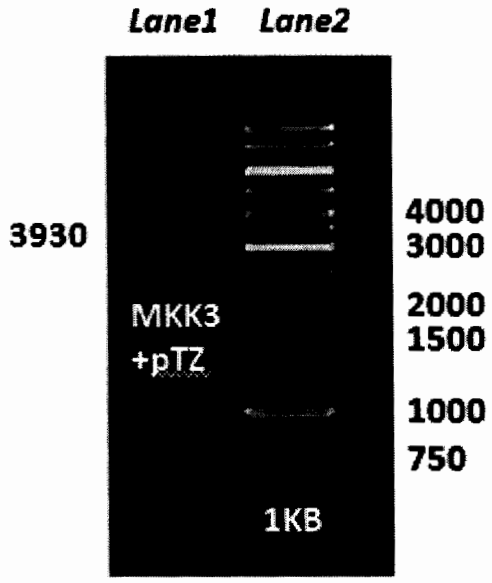


Figure 3.8: Agarose gel electrophoresis of Plasmid Preparation. Lane 2: 1KB DNA Ladder (Fermentas). Lane1: plasmid isolated from C# 51 (mKpTZ51).

MKpTZ51 construct was subjected to sequencing. The complete coding sequence of MKK3 obtained from sequencing was named MKK3gzl and is presented below:

>MKK3gzl

ATGGAGTCGCCCGCCTCGAGCCAGCCCGCCAGCATGCCCCAGTCCAAAGGAAAATCCAAGA
GGAAGAAGGATCTACGGATATCCTGCATGTCCAAGCCACCCGCACCCAACCCCACACCCCCC
CGGAACCTGGACTCCCGGACCTTCATCACCATTGGAGACAGAAACTTTGAGGTGGAGGCTG
ATGACTTGGTGACCATCTCAGAACTGGGCCGTGGAGCCTATGGGGTGGTAGAGAAGGTGC
GGCACGCCCAGAGCGGCACCATCATGGCCGTGAAGCGGATCCGGGCCACCGTGAACTCAC
AGGAGCAGAAGCGGCTGCTCATGGACCTGGACATCAACATGCGCACGGTCGACTGTTTCTA
CACTGTCACCTTCTACGGGGCACTATTCAGAGAGGGAGACGTGTGGATCTGCATGGAGCTC
ATGGACACATCCTTGGACAAGTTCTACCGGAAGGTGCTGGATAAAAACATGACAATTCCAG
AGGACATCCTTGGGGAGATTGCTGTGTCTATCGTGCGGGCCCTGGAGCATCTGCACAGCAA
GCTGTCGGTGATCCACAGAGATGTGAAGCCCTCCAATGTCCTTATCAACAAGGAGGGCCAT
GTGAAGATGTGTGACTTTGGCATCAGTGGCTACTTGGTGGACTCTGTGGCCAAGACGATGG
ATGCCGGCTGCAAGCCCTACATGGCCCCTGAGAGGATCAACCCAGAGCTGAACCAGAAGG
GCTACAATGTCAAGTCCGACGTCTGGAGCCTGGGCATCACCATGATTGAGATGGCCATCCT
GCGGTTCCCTTACGAGTCCTGGGGGACCCCGTTCCAGCAGCTGAAGCAGGTGGTGGAGGA
GCCGTCCCCCCAGCTCCCAGCCGACCGTTTCTCCCCCGAGTTTGTGGACTTCACTGCTCAGTG
CCTGAGGAAGAACCCCGCAGAGCGTATGAGCTACCTGGAGCTGATGGAGCACCCCTTCTTC
ACCTTGCACAAAACCAAGAAGACGGACATTGCTGCCTTCGTGAGAGATCTGGGAGAAGACT
CATAG

Alignment of MKK3gzl with reported MKK3 mRNA sequences obtained from NCBI was carried out through CLUSTALW. MKK3gzl nucleotide sequence showed 99-100% homology with reported MKK3 sequences and is shown below. At C terminal of MKK3gzl there is a deletion of stretch of nucleotides from 1023-1025.

3.3 Alignment Score of MKK3gzl Nucleotide Sequence

SeqA	Name	Len(nt)	SeqB	Name	Len(nt)	Score
1	MKK3gzl	1041	2	HMKK3	957	99
1	MKK3gzl	1041	3	MKK3VB	1044	100
1	MKK3gzl	1041	4	MKK3clone	1044	99
2	HMKK3	957	3	MKK3VB	1044	99
2	HMKK3	957	4	MKK3clone	1044	99
3	MKK3VB	1044	4	MKK3clone	1044	99

Table 3.1: This table is illustrating alignment score of MKK3gzl nucleotide sequence with different MKK3 mRNA sequences obtained from NCBI. SeqA and SeqB are showing the nucleotide sequence names, Len is showing the length of sequences while score column is illustrating the homology score of nucleotide sequences.

```
MKK3gzl      ATGGAGTCGCCCGCCTCGAGCCAGCCCGCCAGCATGCCCCAGTCCAAAGGAAAATCCAAG 60
MKK3clone    ATGGAGTCGCCCGCCTCGAGCCAGCCCGCCAGCATGCCCCAGTCCAAAGGAAAATCCAAG 60
MKK3VB       ATGGAGTCGCCCGCCTCGAGCCAGCCCGCCAGCATGCCCCAGTCCAAAGGAAAATCCAAG 60
HMKK3        ------------------------------------------------------------

MKK3gzl      AGGAAGAAGGATCTACGGATATCCTGCATGTCCAAGCCACCCGCACCCAACCCCACACCC 120
MKK3clone    AGGAAGAAGGATCTACGGATATCCTGCATGTCCAAGCCACCCGCACCCAACCCCACACCC 120
MKK3VB       AGGAAGAAGGATCTACGGATATCCTGCATGTCCAAGCCACCCGCACCCAACCCCACACCC 120
HMKK3        -------------------------ATGTCCAAGCCACCCGCACCCAACCCCACACCC 33
                                       **********************************

MKK3gzl      CCCCGGAACCTGGACTCCCGGACCTTCATCACCATTGGAGACAGAAACTTTGAGGTGGAG 180
MKK3clone    CCCCGGAACCTGGACTCCCGGACCTTCATCACCATTGGAGACAGAAACTTTGAGGTGGAG 180
MKK3VB       CCCCGGAACCTGGACTCCCGGACCTTCATCACCATTGGAGACAGAAACTTTGAGGTGGAG 180
HMKK3        CCCCGGAACCTGGACTCCCGGACCTTCATCACCATTGGAGACAGAAACTTTGAGGTGGAG 93
             ************************************************************

MKK3gzl      GCTGATGACTTGGTGACCATCTCAGAACTGGGCCGTGGAGCCTATGGGGTGGTAGAGAAG 240
MKK3clone    GCTGATGACTTGGTGACCATCTCAGAACTGGGCCGTGGAGCCTATGGGGTGGTAGAGAAG 240
MKK3VB       GCTGATGACTTGGTGACCATCTCAGAACTGGGCCGTGGAGCCTATGGGGTGGTAGAGAAG 240
HMKK3        GCTGATGACTTGGTGACCATCTCAGAACTGGGCCGTGGAGCCTATGGGGTGGTAGAGAAG 153
             ************************************************************

MKK3gzl      GTGCGGCACGCCCAGAGCGGCACCATCATGGCCGTGAAGCGGATCCGGGCCACCGTGAAC 300
MKK3clone    GTGCGGCACGCCCAGAGCGGCACCATCATGGCCGTGAAGCGGATCCGGGCCACCGTGAAC 300
MKK3VB       GTGCGGCACGCCCAGAGCGGCACCATCATGGCCGTGAAGCGGATCCGGGCCACCGTGAAC 300
HMKK3        GTGCGGCACGCCCAGAGCGGCACCATCATGGCCGTGAAGCGGATCCGGGCCACCGTGAAC 213
             ************************************************************
```

```
MKK3gzl      TCACAGGAGCAGAAGCGGCTGCTCATGGACCTGGACATCAACATGCGCACGGTCGACTGT 360
MKK3clone    TCACAGGAGCAGAAGCGGCTGCTCATGGACCTGGACATCAACATGCGCACGGTCGACTGC 360
MKK3VB       TCACAGGAGCAGAAGCGGCTGCTCATGGACCTGGACATCAACATGCGCACGGTCGACTGT 360
HMKK3        TCACAGGAGCAGAAGCGGCTGCTCATGGACCTGGACATCAACATGCGCACGGTCGACTGT 273
             ************************************************************

MKK3gzl      TTCTACACTGTCACCTTCTACGGGGCACTATTCAGAGAGGGAGACGTGTGGATCTGCATG 420
MKK3clone    TTCTACACTGTCACCTTCTACGGGGCACTATTCAGAGAGGGAGACGTGTGGATCTGCATG 420
MKK3VB       TTCTACACTGTCACCTTCTACGGGGCACTATTCAGAGAGGGAGACGTGTGGATCTGCATG 420
HMKK3        TTCTACACTGTCACCTTCTACGGGGCACTATTCAGAGAGGGAGACGTGTGGATCTGCATG 333
             ************************************************************

MKK3gzl      GAGCTCATGGACACATCCTTGGACAAGTTCTACCGGAAGGTGCTGGATAAAAACATGACA 480
MKK3clone    GAGCTCATGGACACATCCTTGGACAAGTTCTACCGGAAGGTGCTGGATAAAAACATGACA 480
MKK3VB       GAGCTCATGGACACATCCTTGGACAAGTTCTACCGGAAGGTGCTGGATAAAAACATGACA 480
HMKK3        GAGCTCATGGACACATCCTTGGACAAGTTCTACCGGAAGGTGCTGGATAAAAACATGACA 393
             ************************************************************

MKK3gzl      ATTCCAGAGGACATCCTTGGGGAGATTGCTGTGTCTATCGTGCGGGCCCTGGAGCATCTG 540
MKK3clone    ATTCCAGAGGACATCCTTGGGGAGATTGCTGTGTCTATCGTGCGGGCCCTGGAGCATCTG 540
MKK3VB       ATTCCAGAGGACATCCTTGGGGAGATTGCTGTGTCTATCGTGCGGGCCCTGGAGCATCTG 540
HMKK3        ATTCCAGAGGACATCCTTGGGGAGATTGCTGTGTCTATCGTGCGGGCCCTGGAGCATCTG 453
             ************************************************************

MKK3gzl      CACAGCAAGCTGTCGGTGATCCACAGAGATGTGAAGCCCTCCAATGTCCTTATCAACAAG 600
MKK3clone    CACAGCAAGCTGTCGGTGATCCACAGAGATGTGAAGCCCTCCAATGTCCTTATCAACAAG 600
MKK3VB       CACAGCAAGCTGTCGGTGATCCACAGAGATGTGAAGCCCTCCAATGTCCTTATCAACAAG 600
HMKK3        CACAGCAAGCTGTCGGTGATCCACAGAGATGTGAAGCCCTCCAATGTCCTTATCAACAAG 513
             ************************************************************

MKK3gzl      GAGGGCCATGTGAAGATGTGTGACTTTGGCATCAGTGGCTACTTGGTGGACTCTGTGGCC 660
MKK3clone    GAGGGCCATGTGAAGATGTGTGACTTTGGCATCAGTGGCTACTTGGTGGACTCTGTGGCC 660
MKK3VB       GAGGGCCATGTGAAGATGTGTGACTTTGGCATCAGTGGCTACTTGGTGGACTCTGTGGCC 660
HMKK3        GAGGGCCATGTGAAGATGTGTGACTTTGGCATCAGTGGCTACTTGGTGGACTCTGTGGCC 573
             ************************************************************

MKK3gzl      AAGACGATGGATGCCGGCTGCAAGCCCTACATGGCCCCTGAGAGGATCAACCCAGAGCTG 720
MKK3clone    AAGACGATGGATGCCGGCTGCAAGCCCTACATGGCCCCTGAGAGGATCAACCCAGAGCTG 720
MKK3VB       AAGACGATGGATGCCGGCTGCAAGCCCTACATGGCCCCTGAGAGGATCAACCCAGAGCTG 720
HMKK3        AAGACGATGGATGCCGGCTGCAAGCCCTACATGGCCCCTGAGAGGATCAACCCAGAGCTG 633
             ************************************************************

MKK3gzl      AACCAGAAGGGCTACAATGTCAAGTCCGACGTCTGGAGCCTGGGCATCACCATGATTGAG 780
MKK3clone    AACCAGAAGGGCTACAATGTCAAGTCCGACGTCTGGAGCCTGGGCATCACCATGATTGAG 780
MKK3VB       AACCAGAAGGGCTACAATGTCAAGTCCGACGTCTGGAGCCTGGGCATCACCATGATTGAG 780
HMKK3        AACCAGAAGGGCTACAATGTCAAGTCCGACGTCTGGAGCCTGGGCATCACCATGATTGAG 693
             ************************************************************

MKK3gzl      ATGGCCATCCTGCGGTTCCCTTACGAGTCCTGGGGGACCCCGTTCCAGCAGCTGAAGCAG 840
MKK3clone    ATGGCCATCCTGCGGTTCCCTTACGAGTCCTGGGGGACCCCGTTCCAGCAGCTGAAGCAG 840
MKK3VB       ATGGCCATCCTGCGGTTCCCTTACGAGTCCTGGGGGACCCCGTTCCAGCAGCTGAAGCAG 840
HMKK3        ATGGCCATCCTGCGGTTCCCTTACGAGTCCTGGGGGACCCCGTTCCAGCAGCTGAAGCAG 753
             ************************************************************
```

```
MKK3gzl         TTCACTGCTCAGTGCCTGAGGAAGAACCCCGCAGAGCGTATGAGCTACCTGGAGCTGATG 960
MKK3clone       TTCACTGCTCAGTGCCTGAGGAAGAACCCCGCAGAGCGTATGAGCTACCTGGAGCTGATG 960
MKK3VB          TTCACTGCTCAGTGCCTGAGGAAGAACCCCGCAGAGCGTATGAGCTACCTGGAGCTGATG 960
HMKK3           TTCACTGCTCAGTGCCTGAGGAAGAACCCCGCAGAGCGTATGAGCTACCTGGAGCTGATG 873
                ************************************************************

MKK3gzl         GAGCACCCCTTCTTCACCTTGCACAAAACCAAGAAGACGGACATTGCTGCCTTCGTGAGA 1020
MKK3clone       GAGCACCCCTTCTTCACCTTGCACAAAACCAAGAAGACGGACATTGCTGCCTTCGTGAAG 1020
MKK3VB          GAGCACCCCTTCTTCACCTTGCACAAAACCAAGAAGACGGACATTGCTGCCTTCGTGAAG 1020
HMKK3           GAGCACCCCTTCTTCACCTTGCACAAAACCAAGAAGACGGACATTGCTGCCTTCGTGAAG 933
                ************************************************************

MKK3gzl         GA---TCTGGGAGAAGACTCATAG 1041
MKK3clone       GAGATCCTGGGAGAAGACTCATAG 1044
MKK3VB          GAGATCCTGGGAGAAGACTCATAG 1044
HMKK3           AAGATCCTGGGAGAAGACTCATAG 957
                 *    ********************
```

MKK3gzl coding sequence was translated into protein sequence by using ExPASy and EMBL-EBI Translator. '*' at the end of protein sequence is indicating stop codon.

>MKK3gzl protein

MESPASSQPASMPQSKGKSKRKKDLRISCMSKPPAPNPTPPRNLDSRTFITIGDRNFEVEADDL
VTISELGRAYGVVEKVRHAQSGTIMAVKRIRATVNSQEQKRLLMDLDINMRTVDCFYTVTFYGA
LFREGDVWICMELMDTSLDKFYRKVLDKNMTIPEDILGEIAVSIVRALEHLHSKLSVIHRDVKPSN
VLINKEGHVKMCDFGISGYLVDSVAKTMDAGCKPYMAPERINPELNQKGYNVKSDVWSLGIT
MIEMAILRFPYESWGTPFQQLKQVVEEPSPQLPADRFSPEFVDFTAQCLRKNPAERMSYLELME
HPFFTLHKTKKTDIAAFVRDLGEDS*

This MKK3gzl protein was aligned with reported protein sequences obtained from NCBI and Swiss-Prot by using CLUSTALW. MKK3gzl protein showed 97-100% homology with reported MKK3 protein sequences. A critical stretch of three residues were found to be mutated in MKK3gzl protein. These residues are involved in interactions with upstream MEKKs (Takekawa *et al.*, 2005).

3.4 Alignment Score of MKK3gzl Protein Sequence

SeqA Name	Len(aa)	SeqB Name	Len(aa)	Score

1	Mkk3IB	347	2	MKK3	347	100
1	Mkk3IB	347	3	MKK3I3	347	100
1	Mkk3IB	347	4	MKK3gzl	346	97
2	MKK3	347	3	MKK3I3	347	100
2	MKK3	347	4	MKK3gzl	346	97
3	MKK3I3	347	4	MKK3gzl	346	97

Table 3.2: This table is illustrating alignment score of MKK3gzl protein with different MKK3 protein sequences obtained from NCBI and Swiss-Prot. SeqA and SeqB are showing the protein name, Len is showing the length of protein while score column is illustrating the homology score of protein sequences.

```
Mkk3IB          MESPASSQPASMPQSKGKSKRKKDLRISCMSKPPAPNPTPPRNLDSRTFITIGDRNFEVE 60
MKK3            MESPASSQPASMPQSKGKSKRKKDLRISCMSKPPAPNPTPPRNLDSRTFITIGDRNFEVE 60
MKK3I3          MESPASSQPASMPQSKGKSKRKKDLRISCMSKPPAPNPTPPRNLDSRTFITIGDRNFEVE 60
MKK3gzl         MESPASSQPASMPQSKGKSKRKKDLRISCMSKPPAPNPTPPRNLDSRTFITIGDRNFEVE 60
                ************************************************************

Mkk3IB          ADDLVTISELGRGAYGVVEKVRHAQSGTIMAVKRIRATVNSQEQKRLLMDLDINMRTVDC 120
MKK3            ADDLVTISELGRGAYGVVEKVRHAQSGTIMAVKRIRATVNSQEQKRLLMDLDINMRTVDC 120
MKK3I3          ADDLVTISELGRGAYGVVEKVRHAQSGTIMAVKRIRATVNSQEQKRLLMDLDINMRTVDC 120
MKK3gzl         ADDLVTISELGRGAYGVVEKVRHAQSGTIMAVKRIRATVNSQEQKRLLMDLDINMRTVDC 120
                ************************************************************

Mkk3IB          FYTVTFYGALFREGDVWICMELMDTSLDKFYRKVLDKNMTIPEDILGEIAVSIVRALEHL 180
MKK3            FYTVTFYGALFREGDVWICMELMDTSLDKFYRKVLDKNMTIPEDILGEIAVSIVRALEHL 180
MKK3I3          FYTVTFYGALFREGDVWICMELMDTSLDKFYRKVLDKNMTIPEDILGEIAVSIVRALEHL 180
MKK3gzl         FYTVTFYGALFREGDVWICMELMDTSLDKFYRKVLDKNMTIPEDILGEIAVSIVRALEHL 180
                ************************************************************

Mkk3IB          HSKLSVIHRDVKPSNVLINKEGHVKMCDFGISGYLVDSVAKTMDAGCKPYMAPERINPEL 240
MKK3            HSKLSVIHRDVKPSNVLINKEGHVKMCDFGISGYLVDSVAKTMDAGCKPYMAPERINPEL 240
MKK3I3          HSKLSVIHRDVKPSNVLINKEGHVKMCDFGISGYLVDSVAKTMDAGCKPYMAPERINPEL 240
MKK3gzl         HSKLSVIHRDVKPSNVLINKEGHVKMCDFGISGYLVDSVAKTMDAGCKPYMAPERINPEL 240
                ************************************************************


Mkk3IB          NQKGYNVKSDVWSLGITMIEMAILRFPYESWGTPFQQLKQVVEEPSPQLPADRFSPEFVD 300
MKK3            NQKGYNVKSDVWSLGITMIEMAILRFPYESWGTPFQQLKQVVEEPSPQLPADRFSPEFVD 300
MKK3I3          NQKGYNVKSDVWSLGITMIEMAILRFPYESWGTPFQQLKQVVEEPSPQLPADRFSPEFVD 300
MKK3gzl         NQKGYNVKSDVWSLGITMIEMAILRFPYESWGTPFQQLKQVVEEPSPQLPADRFSPEFVD 300
                ************************************************************


Mkk3IB          FTAQCLRKNPAERMSYLELMEHPFFTLHKTKKTDIAAFVKEILGEDS 347
MKK3            FTAQCLRKNPAERMSYLELMEHPFFTLHKTKKTDIAAFVKEILGEDS 347
MKK3I3          FTAQCLRKNPAERMSYLELMEHPFFTLHKTKKTDIAAFVKEILGEDS 347
MKK3gzl         FTAQCLRKNPAERMSYLELMEHPFFTLHKTKKTDIAAFVRD-LGEDS 346
                ********************************************:: *****
```

3.5 Template Identification

Molecular model of MKK3 was predicted by using the crystallographic structural data of mitogen activated protein kinase kinase 6 (MKK6). Template identification of human MKK3 protein was carried out by BLAST. In BLAST, PSI-BLAST was run against Protein Data Bank. PSI-BLAST results showed that 3FME and 3ENMA had sequence similarity >80% with MKK3 protein sequence (Table 3.3). So these two structures were selected as template for homology modeling of MKK3. Structures of active and inactive MKK6 having PDB IDs 3FME and 3ENMA were downloaded from Protein Data Bank (PDB). 3FME is crystal structure of human MEK6 activated mutant (S207D, T211D), while 3ENMA is structure of the kinase domain of the MAP2K MEK6 with phosphorylation site mimetic aspartic acid mutations (MEK6/DeltaN/DD). Phylogenetic tree was constructed by using neighbor joining method.

3.5.1 MKK3gzl Related Structures

PDB ID	Description	E-value	Aligned length	Sequence Identity
3FMEA	Chain A, Crystal Structure Of Human Mitogen-Activated Protein Kinase Kinase 6 (Mek6) Activated Mutant (S207d, T211d)	1.00E-91	290	83%
3ENMA	Chain A, The Structure Of The Map2k Mek6 Reveals An Autoinhibitory Dimer	1.00E-85	286	81%
3ENMB	Chain B, The Structure Of The Map2k Mek6 Reveals An Autoinhibitory Dimer	1.00E-85	286	81%
3ENMC	Chain C, The Structure Of The Map2k Mek6 Reveals An Autoinhibitory Dimer	1.00E-85	286	81%
3ENMD	Chain D, The Structure Of The Map2k Mek6 Reveals An Autoinhibitory Dimer	1.00E-85	286	81%
2DYLA	Chain A, Crystal Structure Of Human Mitogen-Activated Protein Kinase Kinase 7 Activated Mutant (S287d, T291d)	1.00E-80	294	44%

3EQBA	Chain A,X-Ray Structure Of The Human Mitogen-Activated Protein Kinase Kinase 1 (Mek1) In A Complex With Ligand And Mgatp	1.00E-72	283	40%
1S9JA	Chain A,X-Ray Structure Of The Human Mitogen-Activated Protein Kinase Kinase 1 (Mek1) In A Complex With Ligand And Mgatp	1.00E-72	283	40%
1S9IA	Chain A,X-Ray Structure Of The Human Mitogen-Activated Protein Kinase Kinase 2 (Mek2)in A Complex With Ligand And Mgatp	8.00E-72	293	39%
3EQIA	Chain A,X-Ray Structure Of The Human Mitogen-Activated Protein Kinase Kinase 1 (Mek1) In A Binary Complex With Adp And Mg2p	5.00E-74	299	39%
3COMB	Chain B, Crystal Structure Of Mst1 Kinase	1.00E-79	284	34%
3HCKA	Chain A,Nmr Ensemble Of The Uncomplexed Human Hck Sh2 Domain, 20 Structures	4.1	18	33%
2J7TA	Chain A, Crystal Structure Of Human Serine Threonine Kinase-10 Bound To Su11274	1.00E-77	286	31%
2C30A	Chain A, Crystal Structure Of The Human P21-Activated Kinase 6	1.00E-75	257	31%
2B9FA	Chain A, Crystal Structure Of Non-Phosphorylated Fus3	7.00E-73	276	29%
2F9GA	Chain A, Crystal Structure Of Fus3 Phosphorylated On Tyr182	1.00E-72	276	29%
2J51A	Chain A, Crystal Structure Of Human Ste20-Like Kinase Bound To 5-Amino-3-((4-(Aminosulfonyl)phenyl) N- amino)-(2,6- Difluorophenyl)-1h-1,2,4-Triazole-1-Carbothioamide	5.00E-83	304	29%
2JFMA	Chain A, Crystal Structure Of Human Ste20-Like Kinase (Unliganded Form)	8.00E-83	304	29%
2JFLA	Chain A, Crystal Structure Of Human Ste20-Like Kinase (Diphosphorylated Form) Bound To 5- Amino-3-((4- (Amino sulfonyl)phenyl)amino)-N-(2,6-Difluorophenyl)-1h-1, 2, 4- Triazole-1-Carbothioamide	2.00E-82	304	29%

1XO2B	Chain B, Crystal Structure Of A Human Cyclin-Dependent Kinase 6 Complex With A Flavonol Inhibitor, Fisetin	3.00E-71	271	29%

Table 3.3: Table showing structures related to MKK3gzl.

3.5 Sequence Alignment of Template and Target

MKK3gzl was aligned with two templates, 3FME and 3ENMA by using 3D Coffee. Alignment score with 3FME and 3ENMA was 65% and 68% respectively. Refinement in alignment was carried out through RASCAL. After fixing error in alignments, formatting of alignments to a color postscript file, was obtained through ESPript by using BLOSUM 62 matrix.

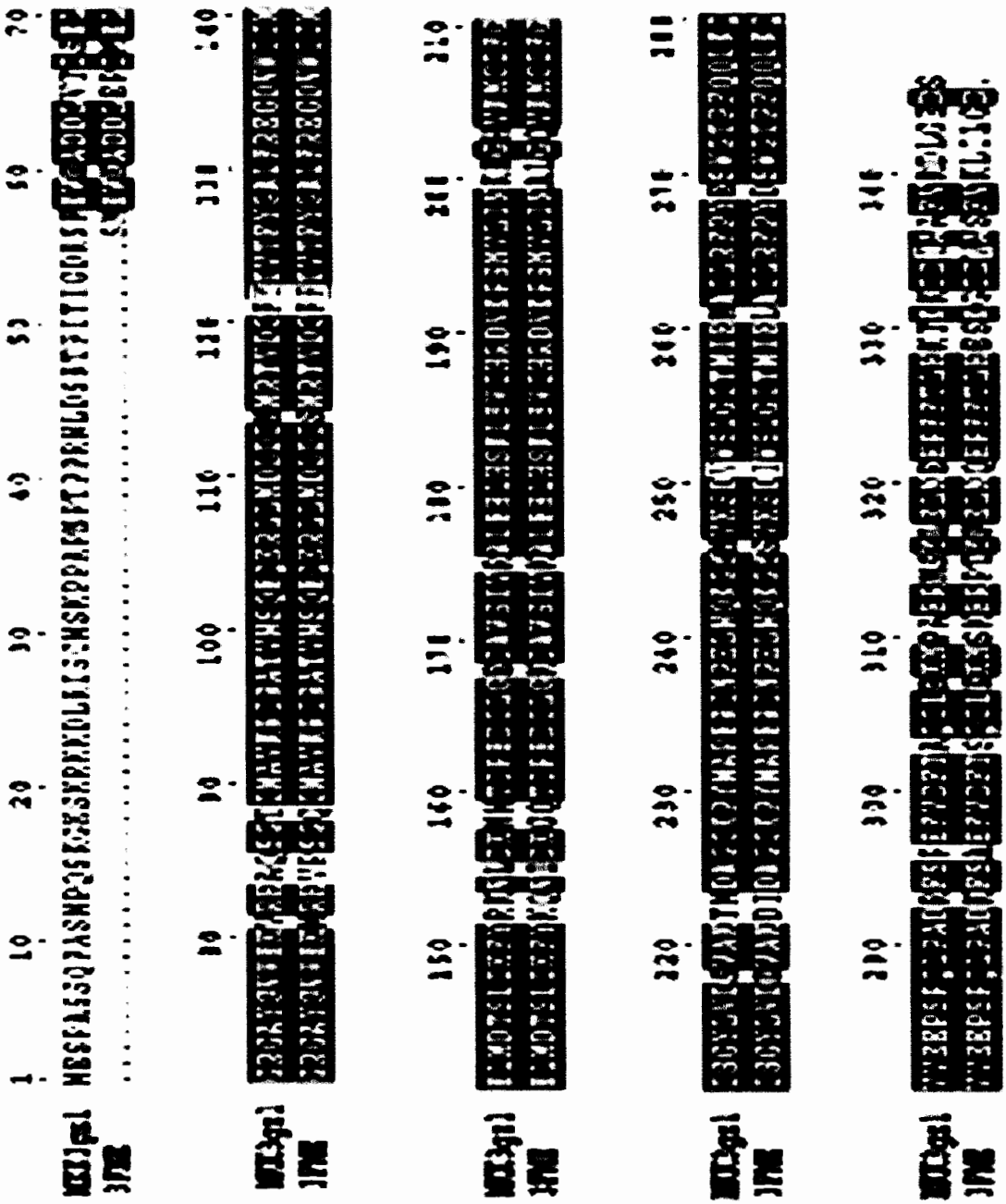


Figure 3.10: Result of ESPript illustrating alignment between MKK3gzl and 3FME.

3.6 Model Building and Evaluation

These alignments of MKK3gzl with 3FME and 3ENMA were used for model building through Modller. Modeller was provided three input files alignment file, script file and PDB file of template. 64 models were generated for each template (3FME, 3ENMA) and named MKK3gzl-3FME, MKK3gzl-3ENMA. Models were evaluated by using NMRCLUST and ERRAT. From ensemble of models, single representative model was selected through NMRCLUST. 11 representative models were selected for MKK3gzl-3FME and MKK3gzl-3ENMA models from output of NMRCLUST. These models were subjected to ERRAT to evaluate the quality of models. Best scoring models selected from MKK3gzl-3FME cluster had ERRAT score of 78% (Figure 3.11) while model selected from MKK3gzl-3ENMA cluster had ERRAT score of 90% (Figure 3.12). MKK3gzl-3FME model was named as activated MKK3 and MKK3gzl-3ENMA model was named as inactivated MKK3.

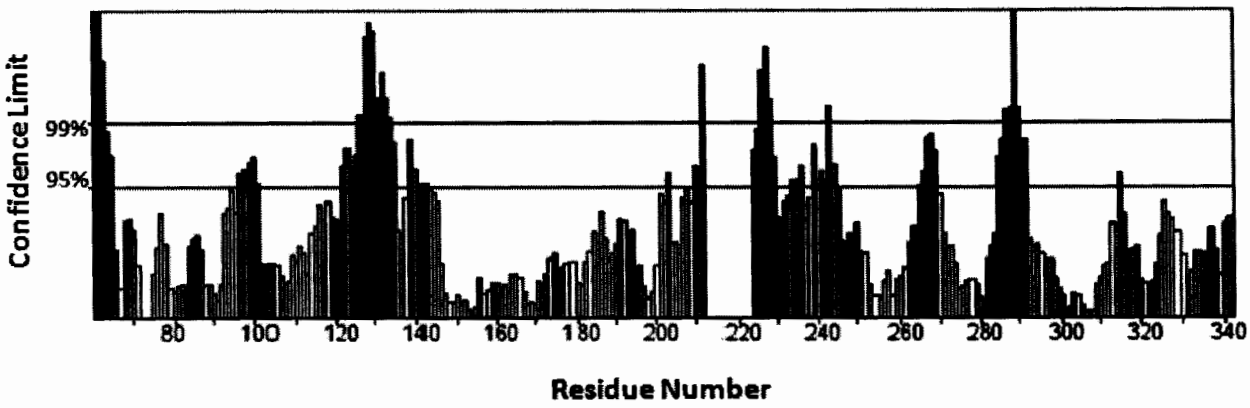


Figure 3.11: ERRAT result for MKK3 Model build by using 3FME as template. ERRAT score for this model was 73%. Here black columns are identifying the problem areas. First second and forth black columns have no bad effect on overall structure of MKK3 as they are not residing in impotant domains while third black column is in activiation domain.

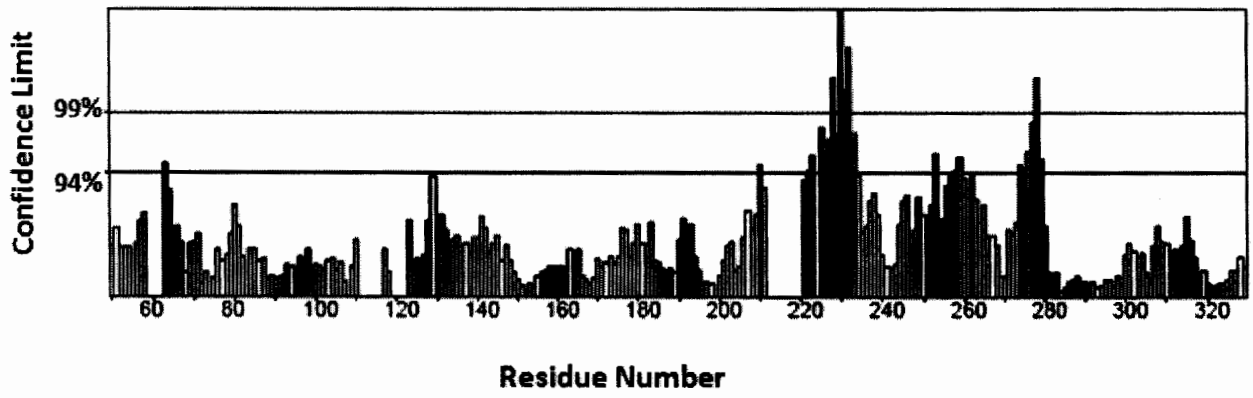


Figure 3.12: ERRAT result for MKK3 Model build by using 3ENMA as template. ERRAT score for this model was 90%. Here black columns are identifying problem areas. Here the first black column resides in activation loop while other black column does not reside in important domains.

3.7 Model Analysis

Activated and inactivated MKK3 models were analyzed by using SwissProt and PyMol. MKK3 belongs to STE group of protein kianses (Hanks, 2003). MKK3 adopts a bilobal fold similar to other kianses. The N-terminal lobe of MKK3 is composed of one long α-C helix and five β-sheets: β1, β2, β3, β4, β5 whereas C-Terminal is helical and encompasses helices D, E, F, G, H, I, J, K (Figure 3.15). All helices and sheets of active MKK3gzl are shown in Figure 3.14. Binding of ATP is largely the function of the N-terminal lobe, while peptide-substrate binding is mediated by the C-terminal lob (Steven, 2003).

The structures were predicted by using two templates, 3FME (active MKK6) and 3ENMA (inactive MKK6). At N-terminal of MAPKKs there is a docking motif known as D domain which mediates interaction with CD docking motif of MAPKs (Takekawa *et al.,* 2005). Models predicted lack the docking motif for p38 substrate. Active and inactive MKK3 kinase domain embraces 12 smaller subdomains like other kinases of kinase supefamily (Hanks & Hunter, 1995). In our models of MKK3, subdomain I, at the N-terminal of the kinase domain, contains the consensus motif Gly-x-GIy-x-x-Gly (GRGAYG) and encompasess β1, β2. Subdomain II encompasses β3 which contains the invariant Lys93, which has long been recognized as being essential for maximal enzyme activity. Subdomain III represents a large α-helix C in the N-terminal lobe. In the center of this helix, nearly invariant Glu residue (Glu103) is located which facilitate the interactions between Lys93 and phosphates of bound ATP. No invariant residues in subdomain IV, V, VIA, these subdomains represent β4, β5; helix D; and helix E respectively. Subdomain VIB folds into the small hydrophobic β6 and β7 with an intervening loop. Included here are two invariant residues Asp190 and Asnl95 in motif DVKPSN. The loop has been termed the catalytic loop because Asp190 within the loop is a candidate for the catalytic base, during an in- line phosphotransfer mechanism it accepts the proton from the attacking substrate hydroxyl group. Subdomain VII contains highly conserved DFG motif. Subdomain VIII appears to play a major role, because it includes residues Ser218, Thr222. Phosphorylation of

Ser218 and Thr222 is required for activation of MKK3. Subdomain IX corresponds to the large α- helix F of the large lobe while Subdomain X corresponds to the small α-helix G. Subdomain XI extends to the COOH-terminal end of the kinase domain. Here is the nearly invariant Arg313 which lies between a-helices H and I. The COOH-terminal boundary of the kinase domain is still poorly defined. In short, smaller N-terminal lobe, which includes subdomains I-IV, is primarily involved in anchoring and orienting the nucleotide. The larger COOH-terminal lobe, which includes subdomains VIA-XI, is largely responsible for binding the peptide substrate and initiating phosphotransfer (Figure 3.13).

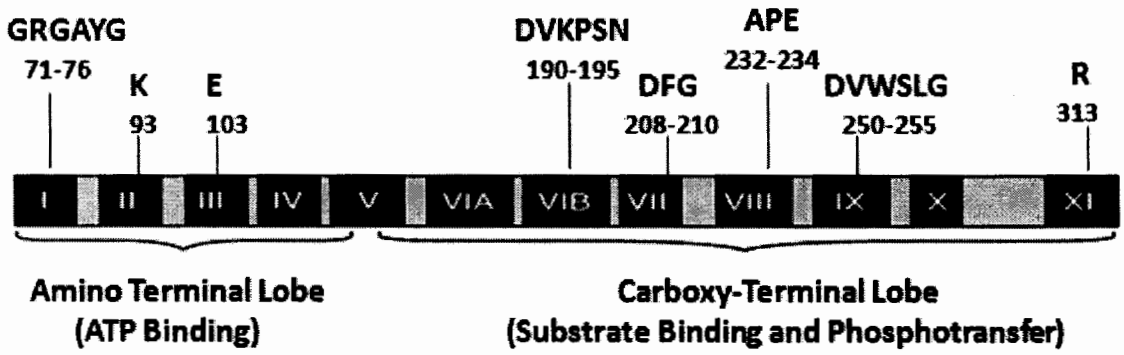


Figure 3.13: Conserved subdomains of kinase domain are indicated by Roman numerals. The positions of amino-acid residues and motifs highly conserved are indicated above the subdomains, using the single-letter amino-acid code, amino and carboxy terminal lobe are indicated below subdomains. Amino and carboxy terminal domians are responsible for ATP binding and substrate binding/ phosphotransfer respectively.

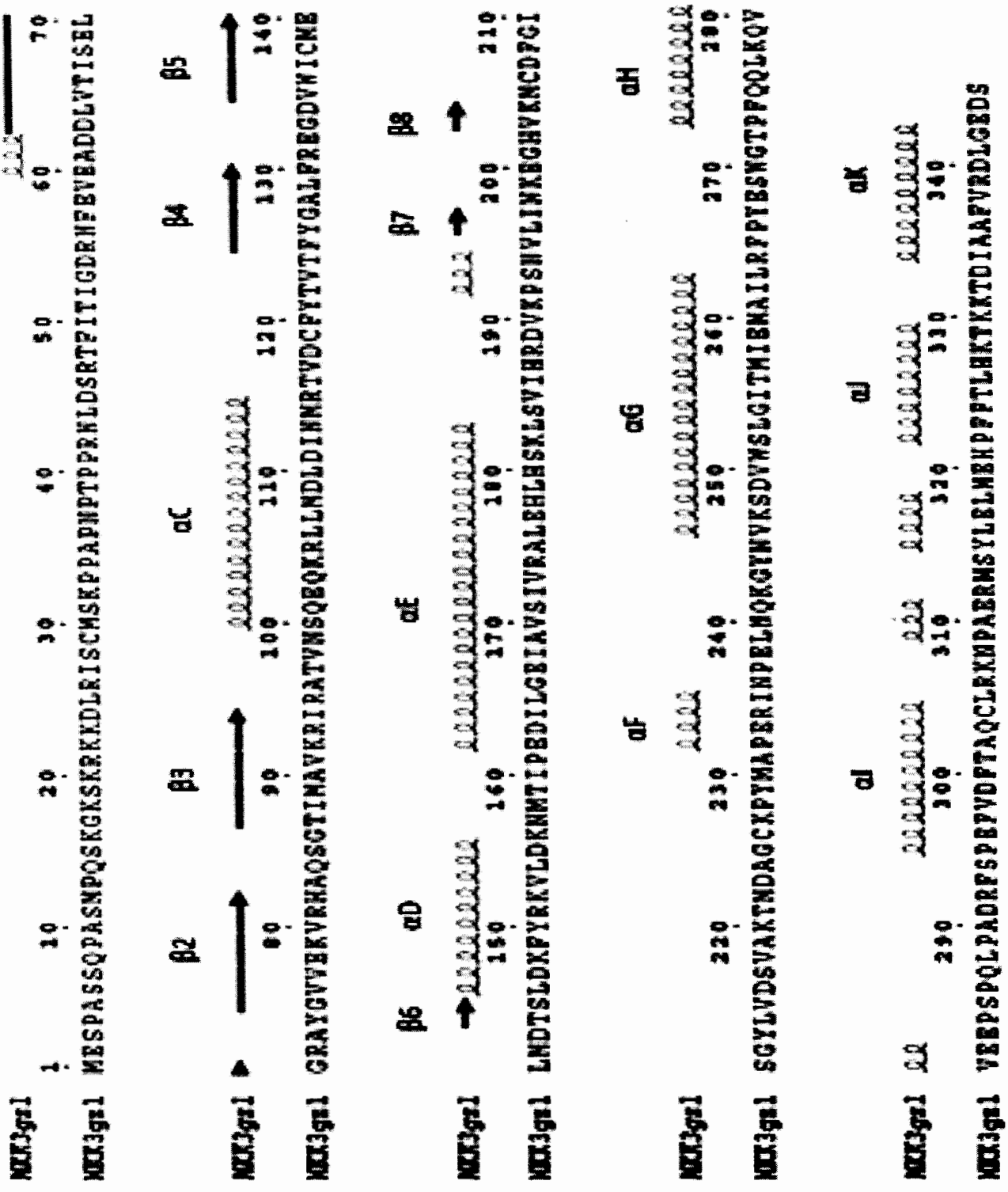


Figure 3.14: Diagrammatic representation of secondary structure of active MKK3gzl.

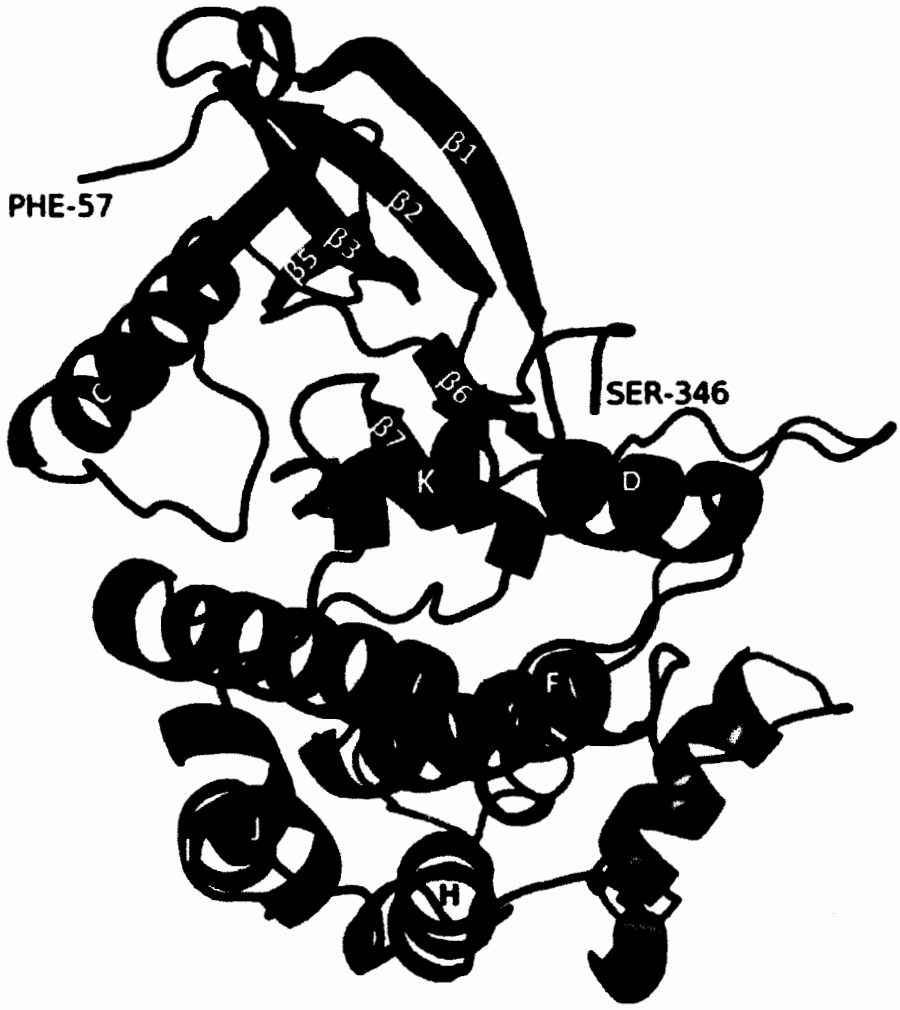


Figure 3.15: Diagrammatic representation of human mitogen active protein kinase kinase 3 predicted by homology modeling. The model depicts arrangement of α-helices, β-sheets and loop regions, including first and last residue of the conformation

Comparison of active and inactive models for MKK3 reveals some main differences in various parts of MKK3 models.

Inside the kinase domain ATP binding region is located at N-terminal and results of Prosite showed that it is from Leu70-Lys93. In active MKK3 structure, ATP binding region adopts a different confirmation as compared to inactive MKK3. It encompasses β1, β2 and β3 sheets. β1 and β2, referred as ATP phosphate binding ribbon or P loop, act as a flexible flap or clamp that covers and anchors the nontransferable phosphates of ATP. In active MKK3 structure, Glysine rich loop (GRGAYG) in subdomain I faces towards activation loop in subdomain VIII and is important for maintaining efficient phosphoryl transfer to substrates. While phosphate binding ribbon in inactive MKK3 structure is in the form of a large loop, a reminiscent of Ω loop (omega loop) (Leszczynski *et al.,* 1986), encompasses only β2 sheet, while lacking β1. Inactive MKK3 Glysine rich loop (GRGAYG) adopts a modified GRGAYG- up confirmation and is away from subdomain VIII (activation loop).

ATP pocket is formed between ATP binding region and activation loop. Activation loop that contains the potential phosphorylation sites of Ser218 and Thr222 is from Ile200-Ile236. At N-terminal of activation loop DFG motif is located. Active MKK3 DFG motif, in subdomain VII, lies in the loop that can be stabilized by a possible hydrogen bond between Asp208 and G1y210. Asp208 helps to orient the phosphate of ATP for transfer. In activation loop, Ser218 and Thr222 (potential phosphorylation sites) are exposed for phosphorylation by upstream MEKKs. Lys93 in subdomainII forms a salt bridge with the carboxyl group of the nearly invariant Glu103 in subdomain III and Lys93 helps to anchor and orient ATP so active MKK3 ATP pocket (wheat color) is stable as shown in Figure 3.15 and can be visualized more clearly from Figure 3.16.

Inactive MKK3 DFG motif, adopts a modified DFG-out confirmation, inactive kinases and inhibitor complexes also show this confirmation (Wrobleski & Doweyko, 2005). Asp208 of DFG resides in the pocket which is normally occupied by Phe209. Residues Ile211- Thr222 of activation loop is modeled into a twisted helix. This helix, along with Phe209 of DFG motif, is wedge between β sheets and helix c of N-terminal lobe. The location of this helix prevents the ion pair between catalytic Lys93

in β3 and Glu103 in α-C helix. Inactive MKK3 ATP pocket (wheat color) has loose conformation as shown in figure 3.16 and can be visualized more clearly from Figure 3.17. Active MKK3 P loop and activation loop are well configured and they form a stable pocket which is ready for ATP binding as shown in figure 3.17 while inactive MKK3 Ω loop and activation loop have loose confirmation so it cannot offer a stable pocket for ATP binding as shown in figure 3.17.

In C-terminal lobe, the, MKK3 kinase conserved helix I is followed by two further helices, J and K. J and K helix are a part of DVD (domain for versatile docking) docking site in MEKs which mediates interactions with upstream MEKKs (Takekawa *et al.*, 2005).

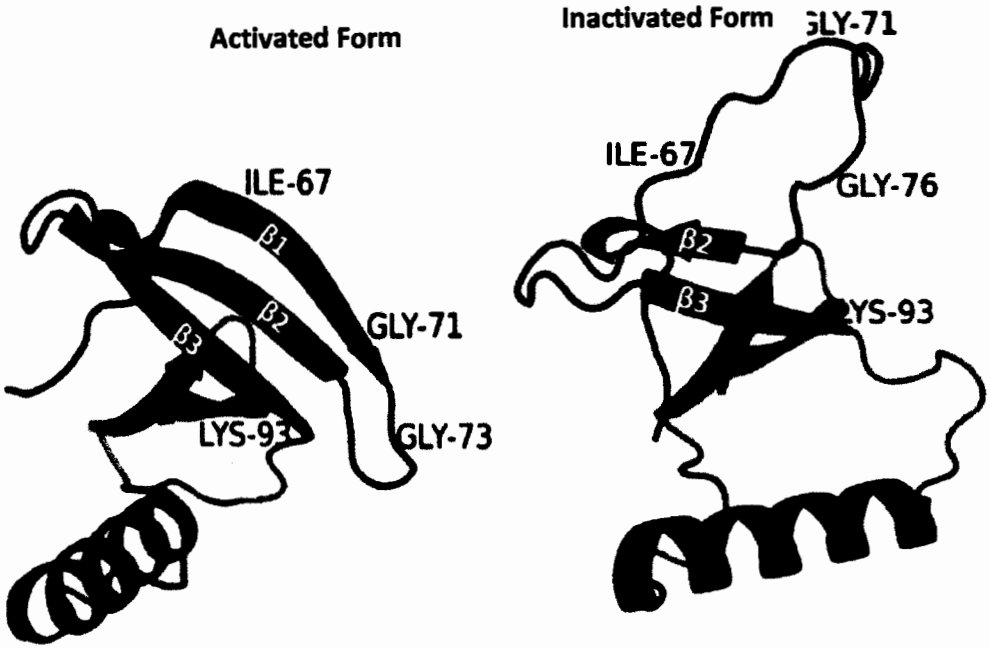


Figure 3.16: ATP binding region (red color) of active and inactive form from Ile67-lys93. In active from ATP binding ribbon (P loop) encompasses β1 and β2; and includes a glycine rich loop (green color), facing downwards. While in inactive form ATP binding ribbon (Ω loop known as omega loop) is deviated from its normal conformation and encompasses only β2, glycine rich loop (green color) adopts a modified conformation.

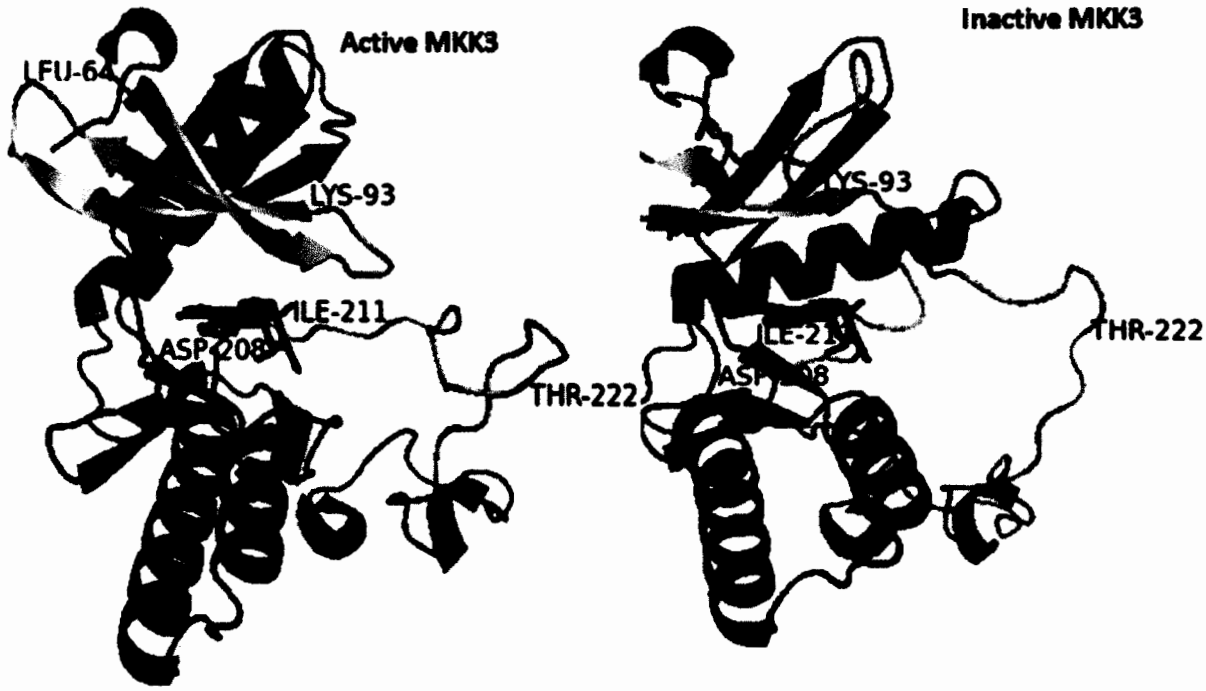


Figure 3.17: Ribbon view of active and inactive form showing ATP pocket formed between ATP binding region and activation loop (wheat color). Activation loop is from 1le211 to Ile236 while at its N-terminal DFG motif Asp208-Gly210 is located. STU (staurosporine in red color) is superimposed from 3FME to visualize the binding site of ATP.

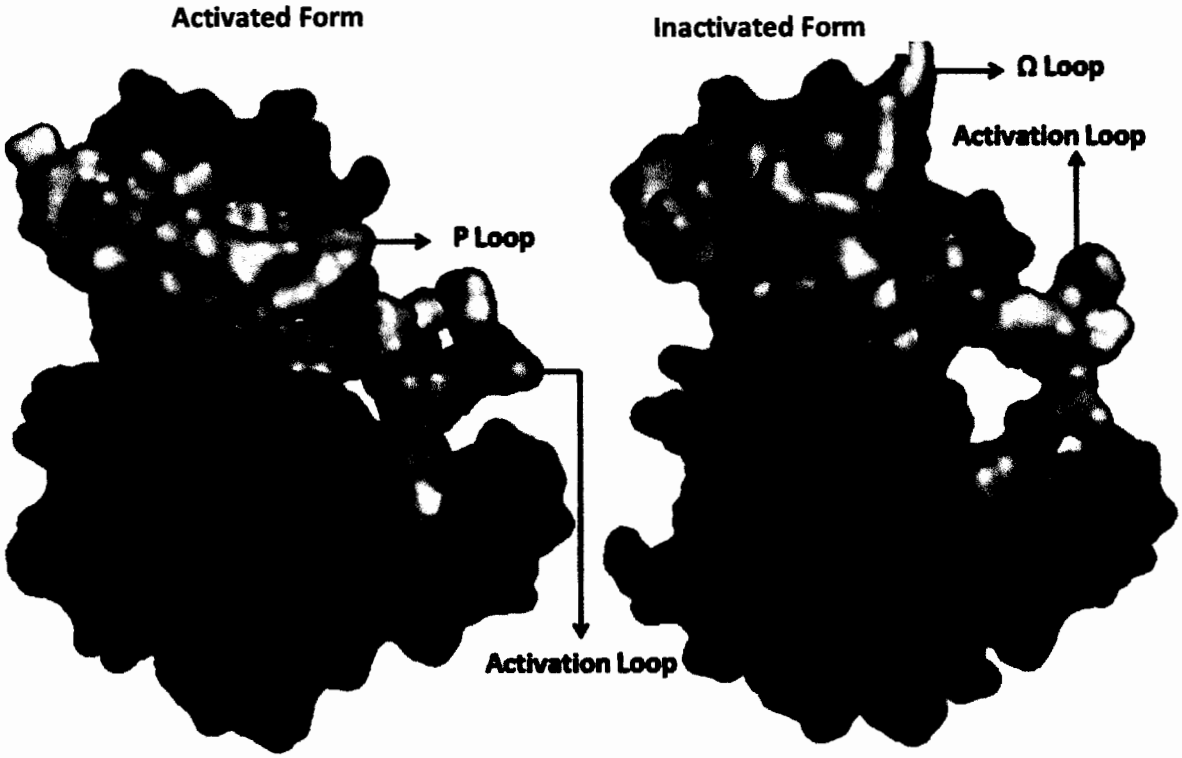


Figure 3.18: Surface view of active and inactive form showing ATP pocket (wheat color). STU ((staurosporine in red color) is superimposed from 3FME to visualize the binding of ATP.

CONCLUSION

MKK3 is a member of p38MAPK signaling cascade and is an upstream kinase of p38. It preferentially phosphorylates p38α and p38β. Many studied elucidate role of different HCV proteins in P38 activation but there are no clues about upstream MAKPs (MAPKKs) involved in HCV induced activation of P38. Some studies also reveal that p38 may account to induction of hepatocellular carcinoma but again there is no study to disclose role of upstream MAKPs involved in p38 induced hepatocellular carcinoma. In order to investigate the role of MKK3 in HCV induced hepatocellular carcinoma there was a need to clone it. For cloning, whole blood was taken from a normal individual and RNA was extracted from it. RNA was subjected to cDNA synthesis; this cDNA was used as template in PCR. PCR was carried out by using specific primers. The amplified product was ligated into pTZ vector and was transformed into competent cells. One of positive clone named mKpTZ51 was subjected to sequencing. Sequence obtained was named MKK3gzl and it was aligned with reported MKK3 mRNA sequences. Alignment revealed that there was some mutation in MKK3gzl. MKK3gzl nucleotide sequence was translated into protein named MKK3gzl Protein. MKK3gzl protein was aligned with reported protein sequence. The alignment revealed that a critical stretch of residues is mutated at C terminal. Role of these residues has yet to be elucidated.

In order to investigate role of MKK3 there was a need to predict its structure as its structure was not available. MKK3gzl protein structure was predicted by homology modeling. Two models were predicted for this protein one was active and other was inactive model. Active model was predicted using template 3FME and inactive model was predicted by using 3ENM. Both of these templates are structure of MKK6. The N-terminal lobe of MKK3 is composed of one long α-C helix and five β-sheets: β1, β2, β3, β4, β5 whereas C-Terminal is helical and encompasses helices D, E, F, G, H, I, J, K. Comparison of active and inactive models for MKK3 reveals some main differences in various parts of MKK3 models. ATP binding region of active and inactive is form from Ile67-lys93. In activated from ATP binding ribbon (P loop) encompasses β1 and β2; and includes a glycine rich loop, facing downwards. While in

inactivated form ATP binding ribbon (Ω loop known as omega loop) is deviated from its normal conformation and encompasses only β2, glycine rich loop adopts a modified conformation. ATP pocket is formed between ATP binding region and activation loop. Activation loop that contains the potential phosphorylation sites of Ser218 and Thr222 is from Ile200-Ile236. At N-terminal of activation loop DFG motif is located. Active MKK3 ATP pocket (wheat color) is stable while Inactive MKK3 ATP pocket (wheat color) has loose conformation. In active form, Ser218 and Thr222 (potential phosphorylation sites) are exposed in activation loop for phosphorylation by upstream MEKKs. Lys93 in subdomainII forms a salt bridge with the carboxyl group of the nearly invariant Glu103 in subdomain III and Lys93 helps to anchor and orient ATP so active MKK3 ATP pocket is stable. Inactive MKK3 DFG motif, adopts a modified DFG-out confirmation. Asp208 of DFG resides in the pocket which is normally occupied by Phe209. Residues Ile211- Thr222 of activation loop is modeled into a twisted helix. This helix, along with Phe209 of DFG motif, is wedge between β sheets and helix c of N-terminal lobe. The location of this helix prevents the ion pair between catalytic Lys93 in β3 and Glu103 in α-C helix it gives a loose confirmation to inactive MKK3. Both active and inactive models encompass J and K helices which are a part of DVD (domain for versatile docking) docking site in MEKs which mediates interactions with upstream MEKKs.

FUTURE PERSPECTIVES

- To clone and sequence MKK3 from liver biopsies of HCV infected persons, and identify putative mutations at C terminal.

- To investigate HCV induced interaction of DVD Motif with upstream kinases through docking and yeast two hybrid system.

- To investigate HCV induced interaction of CD Motif with upstream kinases through docking and yeast two hybrid system.

- To identify role of MKK3 in HCV induced pathogenesis specially hepatocellular carcinoma (HCC).

- To identify specific MKK3 inhibitors to prevent onset of pandemic caused by MKK3 in various contagions.

REFERENCES

Altschul SF(1990) Basic local alignment search tool. *J Mol Biol.* **215:** 403-410.

Bradham C, McClay D (2006) p38 MAPK in development and cancer. *Cell Cycle.* **5:** 824-828.

Boulton T, Yancopoulos G, Gregory J, Slaughter C, Moomaw C, Hsu J, Cobb M (1990) An insulin-stimulated protein kinase similar to yeast kinases involved in cell cycle control. *Science.* **249:** 64–67.

Boulton T, Nye S, Robbins D, Ip Y, Radziejewska E, Morgenbesser S, DePinho R, Panayotatos N, Cobb M, Yancopoulos G (1991) ERKs: a family of protein serine/threonine kinases that are activated and tyrosine phosphorylated in response to insulin and NGF. *Cell.* **65:** 663–675.

Chabaud-Riou M, Firestein (2004) Expression and Activation of Mitogen-Activated Protein Kinase Kinases-3 and -6 in Rheumatoid Arthritis *American Journal of Pathology*, **164**.

Chelvanayagam G, Roy G and Argos P (1994) Easy adaptation of protein structure to sequence. *Protein Engineering.* **7:** 173-184.

Colovos C and Yeates T (1993) Verification of protein structures: patterns of nonbonded atomic interactions. Protein *Sci.* **2**: 1511-1519.

Courchesne W, Kunisawa R, Thorner J (1989) A putative protein kinase overcomespheromone-induced arrest of cell cycling in S.cerevisiae. *Cell.* **58:** 1107–1119.

DeLano WL (2002). The PyMOL Molecular Graphics System. DeLano Scientific, San Carlos, CA, USA. http://www.pymol.org.

Derijard B, Raingeaud J, Barrett T, Wu I-J, Han J, Ulevitch R, Davis R (1995) Independent human MAP kinase signal transduction pathways defined by MEK and MKK isoforms. *Science.* **267:** 682–685.

Enslen H, Brancho D, Davis R (2000) Molecular determinants that mediate selective activation of p38 MAP kinase isoforms. *EMBO J.* **19:** 1301–1311.

Feng DY, Sun Y, Cheng RX, Ouyang X, Zheng H (2005) Effect of hepatitis C virus nonstructural protein NS3 on proliferation and MAPK phosphorylation of normal hepatocyte line. *World J Gastroenterol.* **11:** 2157-2161

Frost J, Steen H, Shapiro PS, Lewis R, Ahn J, Shaw P, Cobb MH (1997). Cross-cascade activation of ERKs and ternary complex actors by Rho family proteins. *EMBO J.* **16:** 6426–6438.

Fukuda K, Tesch H, Yap Y, Forbes M, Flavell A, Davisand J and Nikolic-Paterson J (2008). MKK3 signalling plays an essential role in leukocyte-mediated pancreatic injury in the multiple low-dose streptozotocin model. *Laboratory Investigation.* **88:** 398–407.

Fukunaga R, Hunter T (1997) MNK1, a new MAP kinase-activated protein kinase, isolated by a novel expression screening method for identifying protein kinase substrates. *EMBO J.* **16:** 1921–1933.

Guex N, Peitsch MC (1997). SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis.* **18:** 2714-2723

Han J, Lee J, Bibbs L, Ulevitch R (1994) A MAP kinase targeted by endotoxin and hyperosmolarity in mammalian cells. *Science.* **265:** 808–811.

Han J, Lee J, Jiang Y, Li Z, Feng L, Ulevitch R (1996) Characterization of the structure and function of a novel MAP kinase kinase (MKK6). J *Biol Chem.* **271:** 2886–2891.

Hanks S, Hunter T: (1995) The eukaryotic protein kinase superfamily: kinase (catalytic) domain structure and classification. *FASEB J.* **9:** 576-596.

Hanks S (2003). Genomic analysis of the eukaryotic protein kinase superfamily:a perspective *Genome Biology*, **4**:111.

Hung H (2004*)* Overexpression of tumour suppressor retinoblastoma 2 protein (pRb2/p130) in hepatocellular carcinoma. ***Carcinogenesis.* 25:** 1485-1494.

Ishitani T, J Ninomiya-Tsuji J, Nagai S, Nishita M, Meneghini M, Barker N, Waterman M, Bowerman B, Clevers H, Shibuya H, Matsumoto K (1999) The TAK1-NLK-MAPK-related pathway antagonizes signalling between b-catenin and transcription factor TCF. *Nature.* **399:** 798–802

Iyoda K, Sasaki Y, Horimoto M, Toyama T, Yakushijin T, Sakakibara M, Takehara T, Fujimoto J, Hori M, Wands JR, Hayashi N (2003) Involvement of the p38 mitogen-activated protein kinase cascade in hepatocellular carcinoma. *Cancer.* **97:** 3017-3026.

Kirchmair J, Markt P, Distinto S, Schuster D, Spitzer G, Klaus R. Liedl, Langer T and Wolber G (2008) The Protein Data Bank (PDB), Its Related Services and Software Tools as Key Components for In Silico Guided Drug Discovery *Journal of Medical Chemistry* .**51**: 7021–7040

Klinkert M, Cioli D, Shaw E, Turk V, Bode W, Butler R (1994) A Sequenceand structure similarities of cathepsin B from the parasite Schistosoma mansoni and human liver.*FEBS Lett.* **351:** 397-400.

Krieger, E., S.B. Nabuurs, and G. Vriend (2003) Homology Modelling, in Structural Bioinformatics, P.E. Bourne and H. Weissig, Editors., Wiley-Liss, Inc. 509-523.
Keisuke K, Diane MB (2004), Functions of MAP kinases: insights from gene-targeting studies. *Journal of biochemistry.* **135:** 653-656.

Lawrence A, Kelley, Stephen P, Gardner, Michael J Sutcliffe (1996) An automated approach for clustering an ensemble of NMR-derived protein structures into conformationally related subfamilies. *Protein Eng Des Sel.* **9:** 1063-1065.

Lee J, Laydon J, McDonnell P, Gallagher T, Kumary S, Green D, McNulty D, Blumenthal M, Heys J, Landvatter S, Strickler J, McLaughlin M, Siemens I, Fisher S, Livi G, White J, Adams J, Young P (1994). A protein kinase involved in the regulation of inflammatory cytokine biosynthesis. *Nature.* **372:** 739–746.

Lenassi M, Plemenitas A (2006). The role of p38 MAP kinase in cancer cell apoptosis *Radiol Oncol* **40:** 51-6.

Leo O, Sherrie L. Otterbein, Emeka Ifedigbo, Fang Liu, Danielle E. Morse, Colleen Fearns, Richard J. Ulevitch, Roy Knickelbein, Richard A. Flavell, and Augustine M. Choi (2003) MKK3 Mitogen-Activated Protein Kinase Pathway Mediates Carbon Monoxide-Induced Protection Against Oxidant-Induced Lung Injury. *American Journal of Pathology.* **163**: 2555-2563.

Leszczynski F, Rose G (1986) Loops in globular proteins: a novel category of secondary structure. *Science.* **234:** 849–855.

Lewis T, Shapiro P, Ahn N (1998) Signal transduction through MAP kinase cascades. *Adv Cancer Res.* **74:** 49–139.

Li Y, Batra S, Sassano A, Majchrzak B, Levy D, Gaestel M, Fish E, Davis R, Platanias L (2000). Activation of Mitogen-activated Protein Kinase Kinase (MKK) 3 and MKK6 by Type I Interferons. *JBC.* **281:** 10001-10010

Liang Q, Molkentin J (2003) Redefining the roles of p38 and JNK signaling in cardiac hypertrophy: dichotomy between cultured myocytes and animal models. *Journal of Molecular and Cellular Cardiology* **35:** 1385–1394.

Lin Wang, Rui Ma, Richard A. Flavell, and Mary E. Choi (2002) Requirement of Mitogen-activated Protein Kinase Kinase 3 (MKK3) for Activation of p38_ and p38_ MAPK Isoforms by TGF-_1 in Murine Mesangial Cells. *The journal of biological chemistry.* **277:** 47257–47262.

Marchler-Bauer, Panchenko A, Ariel N, Bryant S (2002) Comparison of sequence and structure alignments for protein domains. *Proteins*. **48:** 439.

Manning, M., and Davis, R.J. (2003). Targeting JNK for therapeutic benefit: from junk to gold? Nat. Rev. *Drug Discov.* **2**: 554–565.

Mansour S, Resing K, Candi J, Hermann A, Gloor J, Herskind K, Wartmann M, Davis R, Ahn N (1994). Mitogen activated protein (MAP) kinase phosphorylation of MAP kinase kinase: determination of phosphorylation sites by mass spectrometry and site-directed mutagenesis. *J Biochem* **116:** 304–314.

Martí-Renom, M, Stuart A., Fiser A., Sanchez R., Melo F., Sali A (2000). Comparative protein structure modeling of genes and genomes. *Annu Rev Bioph Biom*. **29**: 291.

McDermott E, Luke O (2001) Ras participates in the activation of p38 MAPK by Interleukin-1 by associating with IRAK, IRAK2, TRAF6 and TAK-1. *JBC.*

Mobarec J, Sanchez R, and Filizola M (2009) Modern Homology Modeling of G-Protein Coupled Receptors: Which Structural Template to Use *J. Med. Chem*

Needleman SB, Wunsch CD (1970). A general method applicable to the search for similarities in the amino acid sequence of two proteins. *JMol Biol.* **48**: 443-453.

Olivier P, Karsten S, Chantal A, Eamonn O and Cedric N (2004) 3DCoffee@igs: a web server for combining sequencesand structures into a multiple sequence alignment Nucleic Acids Research, Vol. **32.**

Olson JM, Hallahan AR (2004). p38 MAP kinase: a convergence point in cancer therapy. *Trends Mol Med*; **10:** 125-9.

Ono K, Han J. (2000). The p38 signal transduction pathway: activation and function. *Cell Signal*. **12:** 1–13.

Payne D, Rossomando A, Martino P, Erickson A, Her J, Shananowitz J, Hunt DF, Weber M, Sturgill T (1991). Identification of the regulatory phosphorylation sites in pp42/mitogen activated protein kinase (MAP kinase). *EMBO J.* **10:** 885–892.

Qadri I, Iwahashi M, Capasso J, Hopken M, Flores S, Schaack J, Simon F (2004) Induced oxidative stress and activated expression of manganese superoxide dismutase during hepatitis C virus replication: role of JNK, p38 MAPK and AP-1 *Biochem. J.* **378**: 919–928

Ramer S, Davis R (1993) A dominant truncation allele identifies a gene, *STE20* that encodes a putative protein necessary for mating in Saccharomyces cerevisiae. *Proc Natl Acad Sci.* **90:** 452–456.

Rhodes N, Connell L, Errede B (1990) STE11 is a protein kinase required for cell-type-specific transcription and signal transduction in yeast. *Genes Dev.* **4:** 1862–1874.

Ring C , Sun E , McKerrow J , Lee G , Rosenthal P , Kuntz I , and Cohen F (1993) Structure-based inhibitor design by using protein models for the development of antiparasitic agents. *PNAS.* **90:** 3583-3587.

Rossomando A, Dent P, Sturgill T, Marshak D (1994) Mitogen activated protein kinase kinase 1 (MKK1) is negatively regulated by threonine phosphorylation. *Mol Cell Biol.* **14:** 1594–1602.

Rossomando A, Payne D, Weber M, Sturgill T (1989) Evidence that pp42, a major tyrosine kinase target protein, is a mitogen-activated serine/threonine protein kinase. *Proc Natl Acad Sci.* **86:** 6940–6943.

Rouse J, Cohen P, Trigon S, Morange M, Alonso-Llamazares A, Zamanillo D, Hunt T, Nebreda AR (1994) A novel kinase cascade triggered by stress and heat shock that stimulates MAPKAP kinase-2 and phosphorylation of the small heat shock proteins. *Cell.* **78:** 1027–1037.

Sabri A, Pamela A, Lucchesi (2006) ANG II and cardiac myocyte contractility: p38 is not stressed out. *Am J Physiol Heart Circ Physiol.* **290:** 72–73.

Sali A (1993). Comparative protein modelling by satisfaction of spatialrestraints. *J Mol Biol.* **234**: 779.

Schwede T, Kopp J, Guex N, Peitsch MC (2003). SWISS-MODEL: An automated protein homology-modeling server. *Nucleic Acids Res.* **13:** 3381-3385.

Smith TT, Waterman MM (1981).Identification of common molecular subsequences. *J Mol Biol.* **147:** 195-197.

Spazian A, Alisi A, Sanna D, and Balsano C (2006) Role of p38 MAPK and RNA-dependent Protein Kinase (PKR) in Hepatitis C Virus Core-dependent Nuclear Delocalization of Cyclin B1. *THE JOURNAL OF BIOLOGICAL CHEMISTRY.* 281: 10983–10989

Stein B, Brady H, Yang M, Young D, Barbosa M (1996) Cloning and characterization of MEK6, a novel member of the mitogen activated protein kinase kinase cascade. *J Biol Chem.* ***271*****:** 11427– 11433.

Stephen F. Altschul, Thomas L. Madden, Alejandro A. Schaffer, Jinghui Zhang, Zheng Zhang, Webb Miller and David J. Lipman (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs *Nucleic Acids Research.* **25:** 3389–3402.

Sun H, Charles C, Lau L, Tonks N (1993) MKP-1 (3CH134), an immediate early gene product, is a dual specificity phosphatase that dephosphorylates MAP kinase in vivo. *Cell.* **75:** 487-93.

Szabo I, Tarnawski A (2000). Apoptosis in the gastric mucosa: molecular mechanisms basic and clinical implication *Jounrnal of physiology and pharmacology.* **51:** 3-15.

Takekawa M, Tatebayashi K, Saito H (2005). Conserved docking site is essential for activation of mammalian MAP kinase kinases by specific MAP kinase kinase kinases. Mol. *Cell.* **18:** 295–306.

Tanoue T, Eisuke N (2003) Molecular recognitions in the MAP kinase cascades. *Cellular Signalling.* **15:** 455–462.

Thompson J, Plewniak F, Thierry, Poch O (2003) RASCAL: Rapid scanning and correction of multiple sequence alignment programs. *Bioinformatics.*

Toda T, Shimanuki M, Yanagida M (1991). Fission yeast genes that confer resistance to staurosporine encode an AP-1-like transcription factor and a protein kinase related to the mammalian ERK1/ MAP2 and budding yeast FUS3 and KSS1 kinases. *Genes Dev.* **5:** 60–73.

Waskiewicz A, Flynn A, Proud C, Cooper J (1997) Mitogen activated protein kinases activate the serine-threonine kinase Mnk1 and Mnk2. *EMBO J.* **16:** 1909–1920

Wilsbacher JL, Goldsmith EJ, Cobb MH (1999). Phosphorylation of MAP kinases by MAP/ERK kinases involves multiple regions of the MAP kinases. *J Biol Chem.* **274:** 16988–16994.

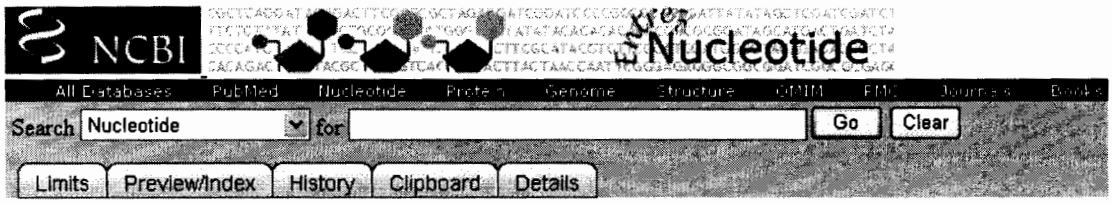
Wrobleski, Stephen T, Doweyko, Arthur M (2005). Structural comparison of p38 inhibitor-protein complexes: a review of recent p38 inhibitors having unique binding interactions. *Curr. Top. Med. Chem.* **5:** 1005-1016

Zarubin T, Han J (2005). Activation and signaling of the p38 MAP kinase pathway. *Cell Research.* **15:** 11-18.

Zeng J, Wang H, Chen Z, Ullrich A, Wu M (2002) Molecular cloning and characterization of a novel gene which is highly expressed in hepatocellular carcinoma. *Oncogene.* **21:** 4932-4943.

Zhang J, Zhang F, Ebert D, Cobb MH, Goldsmith EJ (1995) Activity of the MAP kinase ERK2 is controlled by a flexible surface loop. *Structure.* **3:** 299–307.

Zhao L, Zhang X, Zhao P, Cao J, Cao M, Zhu S, Liu H, Zhong-Tian Q (2006) Up-regulation of ERK and p38 MAPK signaling pathways by hepatitis C virus E2 envelope protein in human T lymphoma cell line. *Journal of Leukocyte Biology.*



Format: GenBank FASTA Graphics More Formats▾ Download▾

Homo sapiens mitogen activated protein kinase kinase 3 (MAP2K3) mRNA, complete cds

```
LOCUS       GQ225578                1041 bp    mRNA    linear   PRI 14-JUL-
2009
DEFINITION  Homo sapiens mitogen activated protein kinase kinase 3 (MAP2K3)
            mRNA, complete cds.
ACCESSION   GQ225578
VERSION     GQ225578.1  GI:253584714
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata;
Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;
            Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1041)
  AUTHORS   Ghazal,A., Zaidi,N.U.S., Kanwal,N., Javed,F., Fatima,K. and
            Qadri,I.
  TITLE     Cloning and sequencing of MKK3 gene from blood of normal
            individuals
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1041)
  AUTHORS   Ghazal,A., Zaidi,N.U.S., Kanwal,N., Javed,F., Fatima,K. and
            Qadri,I.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-MAY-2009) Nust Center for Virology and Immunology,
            National University of Science and Technology, Tmaiz Ud Din
Road,
            Rawalpindi, Punjab 54300, Pakistan
FEATURES             Location/Qualifiers
     source          1..1041
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /cell_type="leukocytes"
                     /country="Pakistan"
                     /collection_date="Dec-2008"
     gene            1..1041
                     /gene="MAP2K3"
                     /gene_synonym="MEK3"
                     /gene_synonym="MKK3"

     CDS             1..1041
                     /gene="MAP2K3"
                     /gene_synonym="MEK3"
                     /gene_synonym="MKK3"
                     /note="involved in the p38 cascade; preferentially
                      phosphorylates p38alpha and p38beta isoforms;
                      activated by a plethora of MEKKs which become activated
                      in response to various physical and chemical stresses;
```

activated by dua phosphorylation of serine and
threonine in the activation loops"
/codon_start=1
/product="mitogen activated protein kinase kinase 3"
/protein_id="ACT33197.1"
/db_xref="GI:253584715"

/translation="MESPASSQPASMPQSKGKSKRKKDLRISCMSKPPAPNPTPPRNL

DSRTFITIGDRNFEVEADDLVTISELGRGAYGVVEKVRHAQSGTIMAVKRIRATVNSQ

EQKRLLMDLDINMRTVDCFYTVTFYGALFREGDVWICMELMDTSLDKFYRKVLDKNMT

IPEDILGEIAVSIVRALEHLHSKLSVIHRDVKPSNVLINKEGHVKMCDFGISGYLVDS

VAKTMDAGCKPYMAPERINPELNQKGYNVKSDVWSLGITMIEMAILRFPYESWGTPFQ

QLKQVVEEPSPQLPADRFSPEFVDFTAQCLRKNPAERMSYLELMEHPFFTLHKTKKTD
IAAFVRDLGEDS"
ORIGIN

1 atggagtcgc ccgcctcgag ccagcccgcc agcatgcccc agtccaaagg aaaatccaag
61 aggaagaagg atctacggat atcctgcatg tccaagccac ccgcacccaa ccccacaccc
121 ccccggaacc tggactcccg gaccttcatc accattggag acagaaactt tgaggtggag
181 gctgatgact tggtgaccat ctcagaactg ggccgtggag cctatggggt ggtagagaag
241 gtgcggcacg cccagagcgg caccatcatg gccgtgaagc ggatccgggc caccgtgaac
301 tcacaggagc agaagcggct gctcatggac ctggacatca acatgcgcac ggtcgactgt
361 ttctacactg tcaccttcta cggggcacta ttcagagagg gagacgtgtg gatctgcatg
421 gagctcatgg acacatcctt ggacaagttc taccggaagg tgctggataa aaacatgaca
481 attccagagg acatccttgg ggagattgct gtgtctatcg tgcgggccct ggagcatctg
541 cacagcaagc tgtcggtgat ccacagagat gtgaagccct ccaatgtcct tatcaacaag
601 gagggccatg tgaagatgtg tgactttggc atcagtggct acttggtgga ctctgtggcc
661 aagacgatgg atgccggctg caagccctac atggcccctg agaggatcaa cccagagctg
721 aaccagaagg gctacaatgt caagtccgac gtctggagcc tgggcatcac catgattgag
781 atggccatcc tgcggttccc ttacgagtcc tggggaccc cgttccagca gctgaagcag
841 gtggtggagg agccgtcccc ccagctccca gccgaccgtt tctcccccga gtttgtggac
901 ttcactgctc agtgcctgag gaagaacccc gcagagcgta tgagctacct ggagctgatg
961 gagcacccct tcttcacctt gcacaaaacc aagaagacgg acattgctgc cttcgtgaga
1021 gatctgggag aagactcata g

//