Structural and Molecular characterization of an Extracellular signal Regulated Kinase (ERK1/2) in association with Ribosomal S6 Kinases

By

Bhanu Prakash Jagilinki

LIFE 09200904001

Tata Memorial Centre-ACTREC
Navi Mumbai

A thesis submitted to the

Board of Studies in Life Sciences

In partial fulfilment of requirements

for the Degree of

DOCTOR OF PHILOSOPHY

of

HOMI BHABHA NATIONAL INSTITUTE

July, 2016

HomiBhabha National Institute

Recommendations of the Viva Voce Committee

As members of the Viva Voce Committee, we certify that we have read the dissertation prepared by Bhanu Prakash Jagilinki entitled **"Structural and Molecular characterization of an Extracellular signal Regulated Kinase (ERK1/2) in association with Ribosomal S6 Kinases"**and recommend that it may be accepted as fulfilling the thesis requirement for the award of Degree of Doctor of Philosophy.

Chairman - Prof. Vinay Kumar — **Date:** 18/7/16

Guide / Convener - Dr. Ashok Varma — **Date:** 18/7/16

Member 1 - Dr. Sanjay Gupta — **Date:** 18/7/16

Member 2-Dr.Murali Krishna — **Date:** 18/7/16

Invitee -Prof. C.P. Rao — **Date:** 18/7/16

External Examiner -Prof.Punit Kaur — **Date:** 18/7/2016

Final approval and acceptance of this thesis is contingent upon the candidate's submission of the final copies of the thesis to HBNI.

I hereby certify that I have read this thesis prepared under my direction and recommend that it may be accepted as fulfilling the thesis requirement.

Date: 18/7/16

Dr. Ashok Varma
(Guide)

Place:Navi Mumbai

STATEMENT BY AUTHOR

This dissertation has been submitted in partial fulfilment of requirements for an advanced degree at HomiBhabha National Institute (HBNI) and is deposited in the Library to be made available to borrowers under rules of the HBNI.

Brief quotations from this dissertation are allowable without special permission, provided that accurate acknowledgement of source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the Competent Authority of HBNI when in his or her judgment the proposed use of the material is in the interests of scholarship. In all other instances, however, permission must be obtained from the author.

Bhanu

Navi Mumbai, Bhanu Prakash Jagilinki

Date 18/07/2016

DECLARATION

I, hereby declare that the investigation presented in the thesis has been carried out by me. The work is original and has not been submitted earlier as a whole or in part for a degree / diploma at this or any other Institution / University.

Bhanu Prakash Jagilinki

List of Publications arising from the thesis

Journal

1."Conserved residues at the MAPKs binding interfaces that regulate transcriptional machinery", **Bhanu P. Jagilinki**, Nikhil Gadewal, Harshal Mehta, Hafiza Mahadik, Vikrant Pandey, Anamika, Ulka Sawant, Prasad A. Wadegaonkar, Peyush Goyal, Satish Kumar & Ashok K.Varma, Journal of biomolecular structure & dynamics, 2015, Vol. 33, 852-860.

2."Functional basis and biophysical approaches to characterize the C-terminal domain of human-Ribosomal s6 Kinase 3", **Bhanu P Jagilinki**, Rajan Kumar Choudhary, Pankaj S Thapa, Nikhil Gadewal, M.V. Hosur, Satish Kumar, Ashok K Varma, Cell Biochem Biophys, June, 2016, pp 1-9

3. Expression, purification and Biophysical characterization of p90-Ribosomal S6 Kinase-1 from Escherichia coli (manuscript under preparation)

Conferences

1. Poster presented at the 31st Annual convention of the Indian Association for Cancer Research (IACR), 26-29 January, 2012, ACTREC, Navi Mumbai, India. "Sequences of specific conserved residues at the RAF1-MEK1-ERK2-RSKs binding interfaces that regulate transcriptional machinery".

2. Poster presented at the 2nd Global Cancer Genomics Symposium (GCGC), 19-20th November, 2012, ACTREC, Navi Mumbai, India. "Expression and purification of human Ribosomal S6 Kinase-1 (hRSK1)".

3. Poster presented at the International Conference on Structural Genomics 2015 – Deep Sequencing Meets Structural Biology – (ICSG2015-DSMSB), 7 -11 June, 2015, Weizmann Institute of Science, Rehovot, Israel. "Protein folding pattern for RSK 3 CTKD purified from Escherichia coli inclusion bodies".

4. Poster presented at the Indo-French conference on 'Application of Structural Biology in Translational Research & Structure Guided Drug-Design' 19-20 November, 2015, ACTREC, Navi Mumbai, India. "Biophysical characterization of RSK 1 and RSK 3 CTKD from Escherichia coli".

Workshops

1. Modern Trends in Macromolecular structures (MTMS-2011), 21-24 February, 2011, IIT Bombay, Mumbai, India.

2. Protein structure determination using X-ray diffraction, at INDUS-2, 9-10 April, 2013, RRCAT, Indore, India.

Bhanu Prakash Jagilinki

Dedicated to my

DAD

ACKNOWLEDGEMENTS

In the first place, I would like to sincerely thank my guide Dr. Ashok K. Varma, for his immense support throughout my Ph.D., which helped me a lot in both scientific and personal development. His inspiration will definitely help me to go further ahead confidently in my career in the field of structural biology.

I thank my doctoral committee members Dr. Vinay Kumar, Dr. M.V. Hosur, Prof. C.P Rao, Dr. Sanjay Gupta and Dr. Murali Krishna for their suggestions during the doctoral committee meetings. I always had learned new things and developed new ideas from their suggestions.

I am very thankful to Dr. Shubhada Chiplunkar, Director, for providing me an excellent opportunity to work at ACTREC.

I also sincerely thank the ACTREC scientific faculty especially Dr. Bose, Dr. Prasanna and Dr. Venu for always being supportive.

I am very thankful to all scientific staff of all research facilities especially Mr. Dandekar and Mr. Kulkarni (CIR), Mr. Shashi Dolas (Mass-spectrometry), Mr. Naresh and Mrs. Sharda (DNA sequencing) and Mr. Nikhil (BTIS).

I am very happy to have wonderful lab members, Mrs. Ulka, Dr. Dilip, Dr. Vikrant, Lumbini, Mrs. Anamika, Mrs. Pradyna, Rajan, Quadir, Pankaj and Sopan, who supported me immensely during my journey. I also thank Varma lab trainees especially Harshal, Dhwani, Sanjivani, Priyancka, Hafiza, Sweta Raikundalia, Sneha, Jaspreet, Nishita and Swetha Phulmali for their help during my Ph.D.

I would also like to thank all my batch mates (batch 2009) at ACTREC for their continuous support and motivation. I am very thankful, to my best friend and batch mate Mr. Shafqat Ali Khan, for his support and motivation all through the good and bad times. I am also thankful to my friends Lalith and Raja for their help and support throughout my Ph.D.

CONTENTS

Homi Bhabha National Institute

SYNOPSIS OF Ph. D. THESIS

1.	**Name of the Student:**	Bhanu Prakash Jagilinki
2.	**Name of the Constituent Institution:**	Tata Memorial Centre, Advanced Centre for Treatment, Research and Education in Cancer
3.	**Enrolment No. :**	LIFE 09200904001
4.	**Title of the Thesis:**	"STRUCTURAL AND MOLECULAR CHARACTERIZATION OF AN EXTRACELLULAR SIGNAL REGULATED KINASE (ERK1/2) IN" ASSOCAITION WITH RIBOSOMAL S6 KINASEs"
5.	**Board of Studies:**	Life Science

Synopsis

Cancer is the disease of uncontrolled growth of cells. There are different kinds of cancers, classified on the basis of origin of malignancy, such as breast cancer, lung cancer, colon cancer, prostate cancer, oral cancer and ovarian cancer. If the malignant growth is unchecked then the abnormal cells invade onto other healthy organs of the body during a process described as metastasis. Cancer is considered to be one of the leading causes of mortality across the world. The major causes of human cancers are environmental factors such as tobacco, alcohol, diet, infections, radiation, stress, and environmental pollutants, however, these are manageable by taking care of personal life style. The another reason for cancer is hereditary, which is passed on from generation to generation. The characteristic features of all cancerous cells include over-sensitization to proliferation signals, evasion of apoptosis, insensitivity to anti-growth signals, indefinite replicative potential, the ability to metastasize, and the ability to induce angiogenesis [1]. The cancerous cells abrogate the signaling pathways, such as Mitogen-Activated Protein Kinase (MAPK) pathway involved in normal cell growth and proliferation [2]. Hence kinases, which are members of these signaling pathways, are often targeted for developing new small molecule drug leads for therapeutics purpose [3-6]. For the development of any new drugs, the three dimensional structure of the target protein plays very important role, and this atomic level molecular mechanisms would save lot of time and money in designing new drugs. The classical MAPK pathway is often highly unregulated in most of the known cancers. Extracellular signal-Regulated Kinases (ERKs) and Ribosomal s6 kinases (RSKs) which play major roles in many cellular functions such as growth, proliferation, differentiation and cell migration are the most important components of this pathway [7, 8]. So far, eight isoforms of ERK and four isoforms of RSK have been discovered. To better understand this pathway at the atomic level, the *in-vitro*, *in-silico* and biophysical based approach have been carried out with following objectives:

- Structural and biophysical characterization of ERK2 and RSK1 and 3.
- Investigation of the interaction of ERK2 with RSK3.
- *In-silico* modelling of ERK2 and RSK1/2/3/4, and docking studies to understand the protein-protein interactions.

Chapter 1: This chapter is about the introduction of MAPK pathway, particularly emphasizing the roles of ERKs and RSKs. ERKs play a major role in transducing the signal from cell membrane to the nucleus. They belong to CMGC group of protein kinases and are directly activated by MEK1/2 (MAPK/ERK kinase) [9]. ERKs form a sub group within the MAPKs which are activated upon growth factor signaling, and ERK1 and ERK2 are the major isoforms. ERK1 is also known as MAPK3 and its gene is found on chromosome 16 (16p11.2), whereas ERK2, known as MAPK1, is coded by a gene on chromosome 22 (22q11.21). These proteins are Ser/Thr kinases, and are of molecular weight of 43 and 41 kDa, respectively. Upon activation through phosphorylation by MEK, they bind and phosphorylate RSKs leading to the activation of RSKs. Knockout studies on mice has revealed that ERK2 is indispensable for survival, as ERK2 knockout mice die very early during development [10, 11].

Ribosomal s6 kinases (RSKs) are the terminal mediators of external signal to the nucleus of the cell [12]. There are two subfamilies within RSKs; $p90^{rsk}$, also known as MAPK-activated protein kinase-1 (MAPKAP-K1) and $p70^{rsk}$, also known as S6-H1 kinase. In the $p90^{rsk}$ sub family, there are four isoforms in humans, RSK1-4, and two structurally related homologues, RSK-like protein kinase (RLPK) or Mitogen- and Stress-activated kinase (MSK1/2). The $p70^{rsk}$ sub family has two isoforms S6K1 and S6K2 [13]. The four isoforms of RSKs share 75-80 % sequence identity [14]. Like ERKs, RSKs are also Ser/Thr kinases, but structurally distinct having two functional domains which belong to two completely different family of kinases. The N-Terminal Kinase Domain (NTKD) is separated from the C-Terminal

Kinase Domain (CTKD) by a linker of around 100 amino acids, whereas the other ends are flanked by N and C-terminal tails of around 60 amino acids each [15]. The linker region contains phosphorylation sites which are highly conserved. The NTKD shares a greater homology with AGC family of protein kinases, and the CTKD belongs to Calcium/calmodulin-dependent kinases (CaMK) family. The NTKD is involved in phosphorylation of various substrates both in cytoplasm and in nucleus, while CTKD is involved in auto-phosphorylation [16, 17]. The C-terminal tail of RSKs houses the ERK binding motif of consensus sequence Leu-Ala-Gln-Arg-Arg, known as D domain [18]. Upon complete phosphorylation of RSKs, they activate a number of substrates like transcription factors SRF, ERα, CREB, FOS, JuN, STAT, NFκB and TIF1A inside the nucleus [19-21]. In the cytoplasm they are responsible for functionality of the translation mechanism by activating rpS6 and mTORC1, and for cell cycle regulation inhibiting the MYT1 and $p27^{KPI}$ through phosphorylation [22, 23]. Thus, RSKs move into and out of the nucleus depending upon the requirement of the cell. These RSKs are being targeted by various researchers in designing new drugs which may be of therapeutic value in future cancer treatment [24]. However, No three dimensional structure of full length RSK isoforms is available as of now.

Chapter 2: This chapter includes a brief description of the materials and various methodologies used in this study. The biochemical materials used were a variety of restriction enzymes, HI fidelity Phusion polymerase, dNTPS, Ligase, Dpn1, commercially available DNA extraction kits, various plasmids, *E. coli* competent cells, Ni-NTA, GST and amylose resin for affinity chromatography, various crystallization screens for crystallization trials.

The molecular biology techniques used were Mini-prep and Maxi-prep for Plasmid isolation, DNA Agarose gel Electrophoresis for charecterzing DNA samples, Polymerase chain

reaction (PCR) for amplifying genes of interest, restriction digestion to doubly digest the Gel and PCR extracted DNA and ligation to seal the sticky ends. Genes were cloned into plasmid vectors by recombinant DNA technology using restriction enzymes cloning method. DNA sequencing was done for all the constructs to ensure gene of interest is in proper frame, and is without any mutations. Site-directed insertion was done for RSK1 full length construct to incorporate additional 6 HIS residues at the C-terminal end to increase the yields of the protein. pMT-wb-v5 vector was used for cloning of the functional domains of RSKs and transfection in Drosophila Schneider s2 insect cell lines. Stable insect cell lines were established by repeated passaging in the presence of Blasticidin.

The different methods for protein purification are described with emphasis on purification by affinity chromatography using Ni-NTA, GST and Amylose resins for ERK2 and RSKs proteins. Size Exclusion Chromatography (SEC) was performed using FPLC as the last step in purification for getting greater purity and homogeneity. SDS-PAGE gel was used to confirm the purity. MALDI/TOF-TOF and Western Blotting were used to confirm the nativity of proteins. Proteins assembled in inclusion bodies were extracted and purified under denaturation conditions using 8 M urea and especially RSK3 CTKD was successfully refolded using various additives like glycerol, *L*-arginine and Trition X-100. Some of these proteins when expressed in insect cell lines are found to be in soluble fraction. The techniques of CD-spectroscopy, Fluorescence-spectroscopy and limited proteolysis assay were performed on purified proteins to characterize them biophysically for their secondary and tertiary structures. The T_m of purified proteins were calculated by thermal denaturation studies using CD and Fluorescence spectroscopy. Dynamic Light Scattering (DLS) was performed to understand the oligomeric nature of protein. GST pull-down assay and Isothermal Titration Calorimetry were performed to study the interaction between ERK2 and RSK3 C-terminal tail.

Crystallization trials were done by hanging-drop and sitting-drop vapor diffusion methods at 22 °C. Various commercially available crystallization screens such as Hampton's Crystal Screen 1 and 2, SaltRx 1, PEG/Ion, PEGRx and Index, Qiagen PEGs suite (Nextal tubes) and Molecular Dimensions (JCSG-plus) screens were used by sitting drop method. Co-crystallization were tried between ERK2 and RSK3 C-terminal peptide using Hampton's Crystal Screen 1 and 2.

In-silico analyses were carried out to predict the Protein-Protein Interactions at the atomic level. Though MEK1 and ERK2 structures are available in PDB, the sequence (1-60) of MEK1 (PDB id: 3E8N), which interacts with ERK2 (PDB ID: 2WNT), is missing in the PDB structure. Therefore, I-tasser webserver was used to generate the 3D structure of the MEK1 including the missing fragment. Similarly for ERK2 and RSK1 CTKD docking, the C-terminal tail of RSK1 which binds to ERK2 was missing, therefore C-terminal tail of RSK1 was also modeled using the I-tasser server. Molecular Docking between ERK2 and four different interacting peptides of RSKs isoforms were carried out. The 14 amino acids C-terminal tails of all RSKs were modeled into tertiary structure using *ab-initio* based protein modeling. The generated models were subjected to side-chain refinement using side chain refinement module in Discovery studio 2.1. The entire simulation was carried under CHARMM force field. All the docking studies were carried out using ZDock in Discovery Studio 2.5 for predicting the binding interfaces.

Chapter 3: This chapter describes purification and biophysical characterization of ERK2. ERK2, which has a molecular mass of ~41 kDa including the 6 HIS tag was expressed and purified from *E.coli* BL21 (DE3) competent cells by using IMAC and Ni-NTA resin. The protein bound to the column was eluted by lysis buffer containing 300 mM Imidazole. The

protein was further purified by using SEC superdex-75 column connected to an FPLC system. The protein identity is confirmed as ERK2 by trypsin in-gel digestion and Mass spectrometry. CD data suggests that ERK2 is predominantly α-helical with a significant percentage of β-sheets, which is consistent with the crystal structure of ERK2. Since ERK2 is known to interact with C-terminal tail of RSKs, we have performed GST pull down assay. RSK3 C-terminal tail (669-733) was cloned in pGEX-KT vector between BamH1 and EcoR1. After TEV cleavage SDS-PAGE showed that the RSK3 C-terminal tail (669-733) undergoes rapid degradation at its C-terminal end. The un-cleaved GST-RSK3 C-terminal tail was used for pull down assay using purified ERK2.

However, the interaction between ERK2 and RSK3 C-terminal tail could not be seen on the SDS gel. ITC was then performed using ERK2 and RSK3 C terminal tail peptide of sequence 'PRLEPVLSSNLAQRRGMK' from 709-726 residues of RSK3. The ITC results have confirmed the interaction between the ERK2 and RSK3 C-terminal tail through the heat change from the thermogram. Attempts were made to co-crystallize ERK2 and RSK3 C terminal tail peptide.

This chapter also includes *in-silico* studies involving macromolecular docking to predict the Protein-Protein interactions (PPIs) between members of MAPK pathway, particularly emphasizing on MEK1 and ERK2 and also ERK2 and RSKs. These studies were done to reveal the amino acids that are critical to the interactions between MEK1, ERK2 and RSKs. Since N-terminal end of MEK1 and C-terminal end of RSK1, which binds the ERK2 are missing, they were modelled using the I-tasser server. The models were then evaluated using Procheck and Verify_3D server, and were found to have acceptable conformational properties. Docking results suggested that residues 314 to 329 of ERK2 interact with MEK1 (Lys4 and Asn12) and RSK1 (Arg725 and Arg726) in a two pronged manner. *In-silico* docking studies were also performed between ERK2 (PDB id: 1TVO) and the modelled C-terminal 14

residue oligopeptides of all RSK isoforms. A build-protein tool from Discovery studio 2.5.5 was used to construct the structures of the 14 residue oligopeptides from RSK1/2/3/4. These structures were found to adopt alpha helical conformation. The side chains were then subjected to refinement using CHARMM force field and the results have been published.

Chapter 4: This chapter describes the purification and biophysical characterization of RSK1. RSK1 full length (1-735) was cloned in pRSET-A vector between restriction sites Xho1 and EcoR1. After confirming the DNA sequence, it was over expressed in *E.coli* BL-21(pLysS) strain and further purified by using Ni-NTA resin. The initial yields were found to be very poor with the protein co-eluting with bacterial contaminants at low imidazole concentrations. We then analyzed the N-terminal region of RSK1 using KYTE DOOLITTLE Hydropathy plot and found it to be Hydrophobic in nature, probably causing partial burial of the 6 HIS tag. Since C-terminal end of RSK1 is highly hydrophilic, we then inserted another 6 HIS tag at carboxyl terminus by incorporating 18 nucleotide sequence coding for 6 HIS residues just after the 735^{th} residue, and before the stop codon by Site-directed insertion. This construct now has 6 HIS tags at both ends. After verifying the SDM insertion by DNA sequencing, the doubly 6 HIS tagged (N and C-terminus) RSK1 full length was purified to greater extent than with N-terminal His tag alone. The RSK1 full length was passed through desalting column and further biophysical characterization was done using CD and Fluorescence spectroscopy. The CD data suggests RSK1 full length is rich in α-helical content. The RSK1 full length has 7 Trp residues and the fluorescence emission maximum (λ_{max}) was at 337 nm suggesting that the protein is well folded into a tertiary structure. Further thermal denaturation studies revealed that the protein has a T_m value of around 45 °C which is relatively low indicating a lesser stability of the fold.

This chapter also describes the cloning of RSK1 CTKD (418-675) in pET-28 a (+) between EcoR1 and Sal1 restriction sites. RSK1 CTKD (418-675) was found to form mostly in inclusion bodies and it is therefore purified under denaturating conditions using 8 M Urea. Further this construct was also cloned in pMT-wb-v5 vector for purpose of expressing the protein in insect cells. A stable insect cell line was established by repeatedly passaging Drosophila Schneider s2 cell lines in presence of Blasticidin.

Chapter 5: Purification and biophysical characterization of three domains of RSK3 has been described in this chapter. The RSK3 CTKD (415-672) was cloned in pMAL-C2x (MBP tag) and pET-28a (6-His tag) vectors. The RSK3 CTKD from pMAL-C2x vector, when expressed in bacterial system, co-purified with GroEL chaperone. The RSK3 CTKD in pMAL-C2x vector purification was optimized, and more stringent washes with lysis buffer containing 1 M NaCl, 20 % glycerol, 5 mM $MgCl_2$ and 5 mM ATP yielded a single band on SDS-PAGE. The MBP tagged RSK3 CTKD fusion protein was further purified by gel filtration in superdex-200 column using FPLC. Since MBP tag could not be cleaved, this FPLC purified MBP-RSK3 CTKD was used for crystallization trials. Attempts were then initiated to make a 6-His tagged RSK3 CTKD.

The 6-His tagged RSK3 CTKD in pET-28a vector, when expressed in Rosetta 2 (DE3) competent cells was also found to accumulate mostly into inclusion bodies. RSK3 CTKD was then extracted under denaturation conditions using 8 M Urea and then refolded. The inclusion body pellet was washed thoroughly with detergents Triton X-100, sodium deoxycholate and sarkosyl to remove contaminating bacterial proteins in the supernatant. The protein of interest in the inclusion pellet was solubilized by adding 8 M urea to the lysis buffer. The supernatant containing solubilized RSK3 CTKD in denatured form was loaded onto Ni-NTA resin

containing column for further purification by IMAC. The RSK3 CTKD was then dialyzed in refolding buffer which is 50mM Tris, pH 8.0 containing 500mM NaCl, 50mM KCl, 10mM $MgCl_2$, 2mM DTT, 15 % Glycerol, 1 % Triton-X-100 and 500 mM *L*-arginine. The refolded RSK3 CTKD was allowed to pass through FPLC to remove the refolding buffer constituents and also to get homogenous population of the protein. CD Spectroscopy suggested the RSK3 CTKD to be predominantly α-helical. The observed emission maximum (λ_{max}) at 332 nm in Fluorescence spectroscopy confirmed the protein to be properly folded by burying the tryptophans into the hydrophobic core. DLS and FPLC elution profiles of RSK3 CTKD suggests the RSK3 CTKD to be homo-trimeric in nature at pH 8.0. Since the crystallization trials for RSK3 CTKD were unsuccessful, a three dimensional molecular model was built using the software Modeller 9.14 with RSK1 CTKD (PDB id: 2WNT) as template. The model was further validated using PROCHECK and ProSA webservers [25, 26].

This chapter also includes the cloning of RSK3 NTKD (59-318) in pET-28 a (+) between EcoR1 and Sal1 restriction sites. Since RSK3 NTKD (59-318) forms inclusion bodies, it is purified under denaturating conditions using 8 M Urea. Further the functional domains, the NTKD (59-318) and CTKD (415-672) of RSK3 were cloned also in pMT-wb-v5 vector between Ecor1 and Xho1 restriction sites for transfection in Drosophila Schneider s2 insect cell lines. For establishing stable cell lines after every 48 hours, Blasticidin that kills non-transfected cells was added along with the fresh sterile medium. After few cycles of seeding with fresh medium and Blasticidin, stable cell lines were finally established. RSK3 NTKD (59-318) and RSK3 CTKD (415-672) were expressed as soluble proteins in the insect system, and this was confirmed by Western blotting using anti-His antibodies.

Chapter 6: This chapter contains a summary of the work done on ERK and RSK kinases involved in MAPK pathway. A suggestions for future research on this subject is also given. ERK2 was purified and its interaction with multifunctional RSK3 tail was established by ITC.

The amino acids critical to the interactions between MEK1 and ERK2 and similarly between ERK2 and RSKs have been explored using *in-silico* methods. RSK1 full length was cloned and expressed in Bl-21(pLysS) strain and was purified by IMAC. It was found that measurable yields of the protein were obtained only when an additional 6 HIS tag was present at the carboxyl end of the protein. All the functional domains of RSK1 and RSK3 were cloned into pET-28 a (+) vector. However these proteins were mostly insoluble and formed inclusion bodies when expressed in bacterial system. RSK3 CTKD was successfully refolded and characterized biophysically using CD and Fluorescence spectroscopy. DLS suggested that RSK3 CTKD existed as a trimeric protein, in contrast to RSK1 CTKD, which is reported to be a dimer. Further RSK3 CTKD cloned in pMAL-C2x vector, co-purifies with GroEL chaperone. Protocol to get rid of these chaperone contaminants have been established by trial and error. Thus RSK3 CTKD and RSK3 NTKD have been expressed as soluble proteins for the first time in drosophila Schneider s2 insect cell lines. So far, no structure is available for RSK3 isolated and purified from any organism. It may now be possible to produce RSK3 functional domains in large amounts from drosophila Schneider s2 insect cell lines for crystallographic analysis. Crystal structures of either full length RSK3 or its domains will throw light on the reported multi-functionality of RSK3. It will also help us to understand how RSK3 functions as a tumour suppressor, particularly in ovarian cancer [27].

References:

1. Hanahan, D. and R.A. Weinberg, *The hallmarks of cancer.* Cell, 2000. **100**(1): p. 57-70.
2. Zhang, X., et al., *MAPK/ERK signaling pathway-induced hyper-O-GlcNAcylation enhances cancer malignancy.* Mol Cell Biochem, 2015.
3. Sebolt-Leopold, J.S. and R. Herrera, *Targeting the mitogen-activated protein kinase cascade to treat cancer.* Nat Rev Cancer, 2004. **4**(12): p. 937-47.
4. Morris, E.J., et al., *Discovery of a novel ERK inhibitor with activity in models of acquired resistance to BRAF and MEK inhibitors.* Cancer Discov, 2013. **3**(7): p. 742-50.
5. Herrero, A., et al., *Small Molecule Inhibition of ERK Dimerization Prevents Tumorigenesis by RAS-ERK Pathway Oncogenes.* Cancer Cell, 2015. **28**(2): p. 170-82.
6. Uehling, D.E. and P.A. Harris, *Recent progress on MAP kinase pathway inhibitors.* Bioorg Med Chem Lett, 2015. **25**(19): p. 4047-56.
7. Mooz, J., et al., *Dimerization of the kinase ARAF promotes MAPK pathway activation and cell migration.* Sci Signal, 2014. **7**(337): p. ra73.

8. Yan, Y., et al., *Extracellular regulated protein kinases 1/2 phosphorylation is required for hepatic differentiation of human umbilical cord-derived mesenchymal stem cells.* Exp Biol Med (Maywood), 2015. **240**(4): p. 534-45.
9. Crews, C.M., A. Alessandrini, and R.L. Erikson, *The primary structure of MEK, a protein kinase that phosphorylates the ERK gene product.* Science, 1992. **258**(5081): p. 478-80.
10. Saba-El-Leil, M.K., et al., *An essential function of the mitogen-activated protein kinase Erk2 in mouse trophoblast development.* EMBO Rep, 2003. **4**(10): p. 964-8.
11. Yao, Y., et al., *Extracellular signal-regulated kinase 2 is necessary for mesoderm differentiation.* Proc Natl Acad Sci U S A, 2003. **100**(22): p. 12759-64.
12. Romeo, Y., X. Zhang, and P.P. Roux, *Regulation and function of the RSK family of protein kinases.* Biochem J, 2012. **441**(2): p. 553-69.
13. Anjum, R. and J. Blenis, *The RSK family of kinases: emerging roles in cellular signalling.* Nat Rev Mol Cell Biol, 2008. **9**(10): p. 747-58.
14. Lara, R., M.J. Seckl, and O.E. Pardo, *The p90 RSK family members: common functions and isoform specificity.* Cancer Res, 2013. **73**(17): p. 5301-8.
15. Jones, S.W., et al., *A Xenopus ribosomal protein S6 kinase has two apparent kinase domains that are each similar to distinct protein kinases.* Proc Natl Acad Sci U S A, 1988. **85**(10): p. 3377-81.
16. Bjorbaek, C., Y. Zhao, and D.E. Moller, *Divergent functional roles for p90rsk kinase domains.* J Biol Chem, 1995. **270**(32): p. 18848-52.
17. Vik, T.A. and J.W. Ryder, *Identification of serine 380 as the major site of autophosphorylation of Xenopus pp90rsk.* Biochem Biophys Res Commun, 1997. **235**(2): p. 398-402.
18. Smith, J.A., et al., *Identification of an extracellular signal-regulated kinase (ERK) docking site in ribosomal S6 kinase, a sequence critical for activation by ERK in vivo.* J Biol Chem, 1999. **274**(5): p. 2893-8.
19. De Cesare, D., et al., *Rsk-2 activity is necessary for epidermal growth factor-induced phosphorylation of CREB protein and transcription of c-fos gene.* Proc Natl Acad Sci U S A, 1998. **95**(21): p. 12202-7.
20. Frodin, M. and S. Gammeltoft, *Role and regulation of 90 kDa ribosomal S6 kinase (RSK) in signal transduction.* Mol Cell Endocrinol, 1999. **151**(1-2): p. 65-77.
21. Zhao, J., et al., *ERK-dependent phosphorylation of the transcription initiation factor TIF-IA is required for RNA polymerase I transcription and cell growth.* Mol Cell, 2003. **11**(2): p. 405-13.
22. Fujita, N., S. Sato, and T. Tsuruo, *Phosphorylation of p27Kip1 at threonine 198 by p90 ribosomal protein S6 kinases promotes its binding to 14-3-3 and cytoplasmic localization.* J Biol Chem, 2003. **278**(49): p. 49254-60.
23. Fingar, D.C. and J. Blenis, *Target of rapamycin (TOR): an integrator of nutrient and growth factor signals and coordinator of cell growth and cell cycle progression.* Oncogene, 2004. **23**(18): p. 3151-71.
24. Nguyen, T.L., *Targeting RSK: an overview of small molecule inhibitors.* Anticancer Agents Med Chem, 2008. **8**(7): p. 710-6.
25. Laskowski, R.A., et al., *AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR.* J Biomol NMR, 1996. **8**(4): p. 477-86.
26. Wiederstein, M. and M.J. Sippl, *ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins.* Nucleic Acids Res, 2007. **35**(Web Server issue): p. W407-10.
27. Bignone, P.A., et al., *RPS6KA2, a putative tumour suppressor gene at 6q27 in sporadic epithelial ovarian cancer.* Oncogene, 2007. **26**(5): p. 683-700.

Publications in Refereed Journal:

a. Published:

1. Bhanu P. Jagilinki, Nikhil Gadewal, Harshal Mehta, Hafiza Mahadik, Vikrant Pandey, Anamika, Ulka Sawant, Prasad A. Wadegaonkar, Peyush Goyal, Satish Kumar & Ashok K.Varma. Conserved residues at the MAPKs binding interfaces that regulate transcriptional machinery. Journal of biomolecular structure & dynamics. 2015; Vol. 33, No. 4, 852–860. doi: 10.1080/07391102.2014.915764

b. Accepted: NA

c. Communicated:

1. Bhanu P Jagilinki, Rajan Kumar Choudhary, Pankaj S Thapa, Nikhil Gadewal, M.V.Hosur, Satish Kumar, Ashok K Varma. Functional basis and biophysical approaches to characterize the C-terminal domain of human-Ribosomal s6 Kinase 3 (Communicated).

d. Conference attended / poster presentations:

1. Poster presented at the 31st Annual convention of the Indian Association for Cancer Research (IACR), 26-29 January, 2012, ACTREC, Navi Mumbai. "Sequences of specific conserved residues at the RAF1- MEK1- ERK2 – RSKs binding interfaces that regulate transcriptional machinery".
2. Poster presented at the 2nd Global Cancer Genomics Symposium (GCGC), 19-20th November, 2012, ACTREC, Navi Mumbai. "Expression and purification of human Ribosomal S6 Kinase-1 (hRSK1)".
3. Poster presented at the International Conference on Structural Genomics 2015 – Deep Sequencing Meets Structural Biology – (ICSG2015-DSMSB), 7 -11 June, 2015, Weizmann Institute of Science, Rehovot, Israel. "Protein folding pattern for RSK 3 CTKD purified from Escherichia coli inclusion bodies"

Signature of Student:
Date: 6/10/2015

Doctoral Committee:

S. No.	Name	Designation	Signature	Date
1.	Dr. Vinay Kumar	Chairman		06/10/2015
2.	Dr. Ashok Varma	Guide		06/10/15
3.	Prof. C.P. Rao	Member		08/10/2015
4.	Dr. Sanjay Gupta	Member		07/10/15
5.	Dr. Murali Krishna	Member		6/10/15

Forwarded through:

Dr. S.V. Chiplunkar
Director ACTREC
Chairperson,
Academic & Training Programme

Dr. S. V. Chiplunkar
Director
Advanced Centre for Treatment, Research & Education in Cancer (ACTREC)
Tata Memorial Centre
Kharghar, Navi Mumbai 410210.

Dr. K. Sharma
Director, Academics
T.M.C.

PROF. K. S. SHARMA
DIRECTOR (ACADEMICS)
TATA MEMORIAL CENTRE,
PAREL, MUMBAI

Abbreviations

ACE: Atomic Contact Energy

AGC: PKA, PKG, PKC

AMP-PNP: 5'-Adenylyl-imidodiphosphate

ATP: Adenosine Triphosphate

BLAST: Basic Local Alignment Search Tool

BME: β-mercaptoethanol

bp: base pairs

BSA: Bovine serum albumin

CaMK: Calcium/calmodulin-dependent kianse

CAPRI: Critical Assessment of PRediction of Interactions

CASP: Critical Assessment of protein Structure Prediction

CCD: Charged Coupled Device

CD: Circular Dichroism

CDKs: Cyclin-dependent kinases

CHCA: α-cyano-4-hydroxycinnamic acid

CLS: Coffin–Lowry syndrome

CMA: CarboxylMethylAspartate

CREB: cAMP response element-binding protein

CTKD: C-terminal Kinase Domain

DAPK: Death-Associated Protein Kinase

ddNTPs: di-deoxy Nucleoside TriPhosphates

DEAE: Diethyl-aminoethyl

DI: Dimeric Interface

DLS: Dynamic Light Scattering

DNA: Deoxyribonucleic Acid

dNTPs: deoxy Nucleoside TriPhosphate

DOPE: Discrete Optimized Protein Energy

DUSPs: Dual-specificity MAP kinase phosphatases

EB: Elution Buffer

EDTA: Ethylenediaminetetraacetic acid

eIF4B: eukaryotic translation Initiation Factor-4B

eIF4E: eukaryotic translation Initiation Factor-4E

EM: Electron Microscopy

EMR: Electromagnetic Radiation

ERK: Extracellular signal-Regulated Kinase

EtBr: Ethidium bromide

FBS: Fetal Bovine Serum

Fc: Fragment crystallizable

FFT: Fast Fourier Transform

FOS: FBJ murine osteosarcoma viral oncogene homolog

FMK: Fluoro-methyl-ketone

FPLC: Fast protein liquid chromatography

GDP: Guanosine diphosphate

GPC: Gel Permeation Chromatography

GRB2: Growth factor Receptor-Bound protein 2

GSC: Grid-based Shape Complementarity

GSH: reduced Glutathione

GSK: Glycogen Synthase Kinase

GST: Glutathione S-Transferase

GTP: Guanosine triphosphate

HDAC: Histone Deacetylase

HIC: Hydrophobic interaction chromatography

HRP: HorseRadish Peroxidase

IC_{50}: half maximal Inhibitory Concentration

IEC: Ion Exchange chromatography

IGF: Insulin-like growth factors

IMAC: Immobilized Metal ion Affinity Chromatography

IPTG: isopropyl-β-D-thiogalactoside

I-tasser: Iterative Threading ASSEmbly Refinement

ITC: Isothermal Titration Calorimetry

JNKs: c-Jun N-terminal kinases

kb: kilo base pairs

K_d: Dissociation constant

kDa: kilo Dalton

LB: Luria-Bertani

LOMETS: Local meta-threading-server

MAPK: Mitogen-activated protein kinase

MAPKAPK-1: MAPK activated protein kinase-1

MAPKK: Mitogen-activated protein kinase kinase

MALDI-TOF/TOF: Matrix Assisted Laser Desorption Ionization

MBP: Maltose binding protein

MCS: Multiple cloning sites

MEK: MAPK/ERK kinase

MM: Multiple Myeloma

MS: Mass Spectrometry

MSK: Mitogen and stress activated kinase

MT: Metallothionein Promoter

mTOR: mammalian target of rapamycin

NES: Nuclear Export Signal

NFKB: nuclear factor kappa-light-chain-enhancer of activated B cells

NLS: Nuclear Localization Signal

NMR: Nuclear Magnetic Resonance

NTA: NitriloTriacetic Acid

NTKD: N-Terminal Kinase Domain

OD_{600}: Optical Density at 600 nm

OmpT: Outer membrane protease

ORF: Open Reading Frame

PAGE: Poly-Acrylamide Gel Electrophoresis

PCR: Polymerase Chain Reaction

PDB: Protein Data Bank

PDK: 3′-phosphoinositide-dependent kinase

PEA: Phosphoprotein Enriched in Astrocytes

PETG: Phenylethyl β-d-thiogalactopyranoside

pI: Isoelectric point

PI3K: Phosphoinositide 3-kinase

PKA: Protein Kinase A

PKB: Protein Kinase B

PKC: Protein Kinase C

PKG: Protein Kinase G

PML: ProMyelocytic Leukemia

PPIs: Protein-Protein Interactions

PSSpred: Protein Secondary Structure PREDiction

PTMs: Post-Translation Modifications

PVDF: Polyvinylidene fluoride

RAF: Rapidly Accelerated Fibrosarcoma

Ras: Rat sarcoma

R_h: Hydrodynamic radius

RHEB: Ras Homolog Enriched in Brain

RIPA: Radioimmunoprecipitation Assay

RMSD: Root-mean-square deviation

RLPK: RSK like protein kinase

RPM: Revolutions Per Minute

RPTOR: Regulatory associated Protein of mTOR

RSK: Ribosomal S6 Kinase

RTK: Receptor Tyrosine Kinase

SDM: Site-Directed Mutagenesis

SDS: Sodium dodecyl sulfate

SEC: Size Exclusion Chromatography

SHH: Sonic Hedgehog

SOS: Son Of Sevneless

SRF: Serum Response Factor

STAT: Signal Transducer and Activator of Transcription

TAE: Tris, acetic acid and EDTA

TBS: Tris-Buffered Saline

TBST: Tris-Buffered Saline and Tween 20

TD: Thermal denaturation

TEV: Tobacco Etch Virus

TFA: TrifluoroAcetic acid

TH: Tyrosine Hydroxylase

T_m: Melting Temperature

TSC: Tuberous Sclerosis Complex

UV: Ultraviolet

List of Figures

List of Tables

Chapter 1:

Introduction

1.1. Cancer

Cancers are the families of diseases in which cells grow and divide without any regulation. There are different kinds of cancers, such as breast cancer, lung cancer, colon cancer, prostate cancer, oral cancer and ovarian cancer depending on the organ first affected. Cancers are broadly classified into the following groups:

Carcinoma: Carcinoma is the cancer of epithelial cells. This group includes most of the cancers originating from breast, lung, pancreas and colon [1].

Sarcoma: Sarcoma refers to cancer of mesenchymal cells, which are present in the connective tissues like bone, cartilage and nerve cells [2].

Lymphoma: Lymphoma is the cancer of lymphatic cells present in lymph nodes [3].

Leukemia: Leukemia is the cancer of white blood cells produced in the bone marrow. It constitutes about 30 % of all childhood cancers [4].

Germ cell tumours: These cancers are derived from germ cells. Testicular cancer and ovarian cancer are the major cancers of this group [5].

Blastoma: Blastoma is the cancer of precursor cells. Neuroblastoma, nephroblastoma, medulloblastoma and retinoblastoma are examples of this group [6].

Some abnormally growing cancerous cells invade other healthy organs of the body and start colonizing that organ, which is described as metastasis. The majority of cancer deaths are due to metastasis. The occurrence of cancer is brought about by multiple steps which brings in dynamic changes by incorporating mutations within the genome [7]. These mutations disturb the balance between oncogenes and tumour suppressor genes by causing gain of function in oncogenes and loss of function in tumour suppressor genes, thereby converting the normal cells into cancerous cells [8]. Apart from hereditary reasons,

environmental and lifestyle factors are also responsible for the occurrence of cancer. The major environmental factors are tobacco, alcohol, diet, infections, radiation, stress and pollutants. However, many of these are manageable by taking care of personal life style. Cancers caused by environmental factors are not passed on from generation to generation.

1.2. Characteristic features of cancer

The characteristic features of all cancerous cells include: 1) over-sensitization to proliferation signals, 2) evasion of apoptosis, 3) insensitivity to anti-growth signals, 4) indefinite replicative potential, 5) the ability to metastasize and 6) induce angiogenesis. All these properties have been graphically depicted **Figure 1.1** [9]. Also these cancerous cells acclimatize very quickly to different micro-environments by undergoing different genetic mutations.

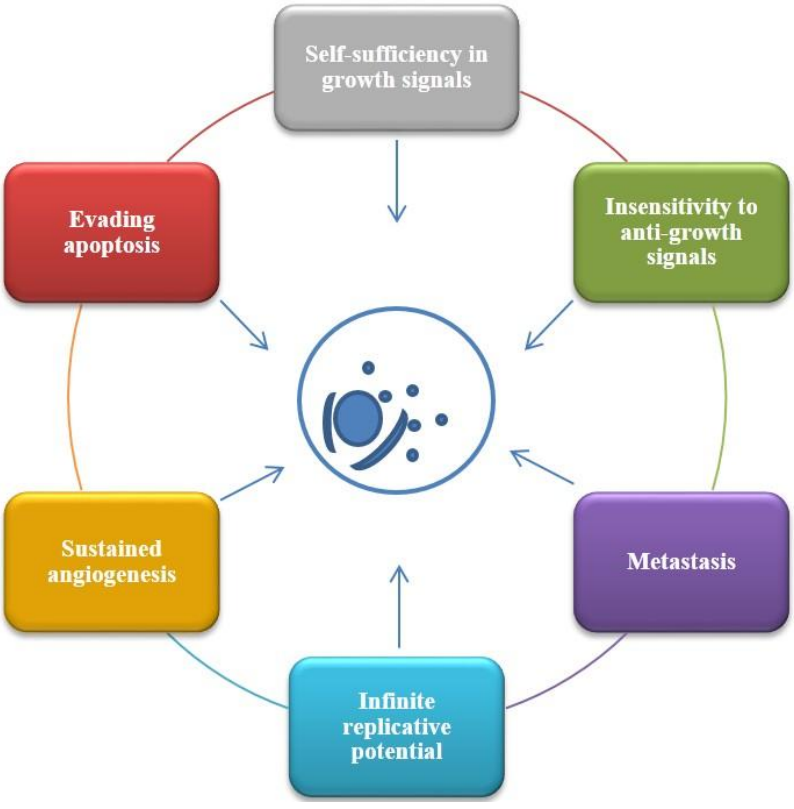
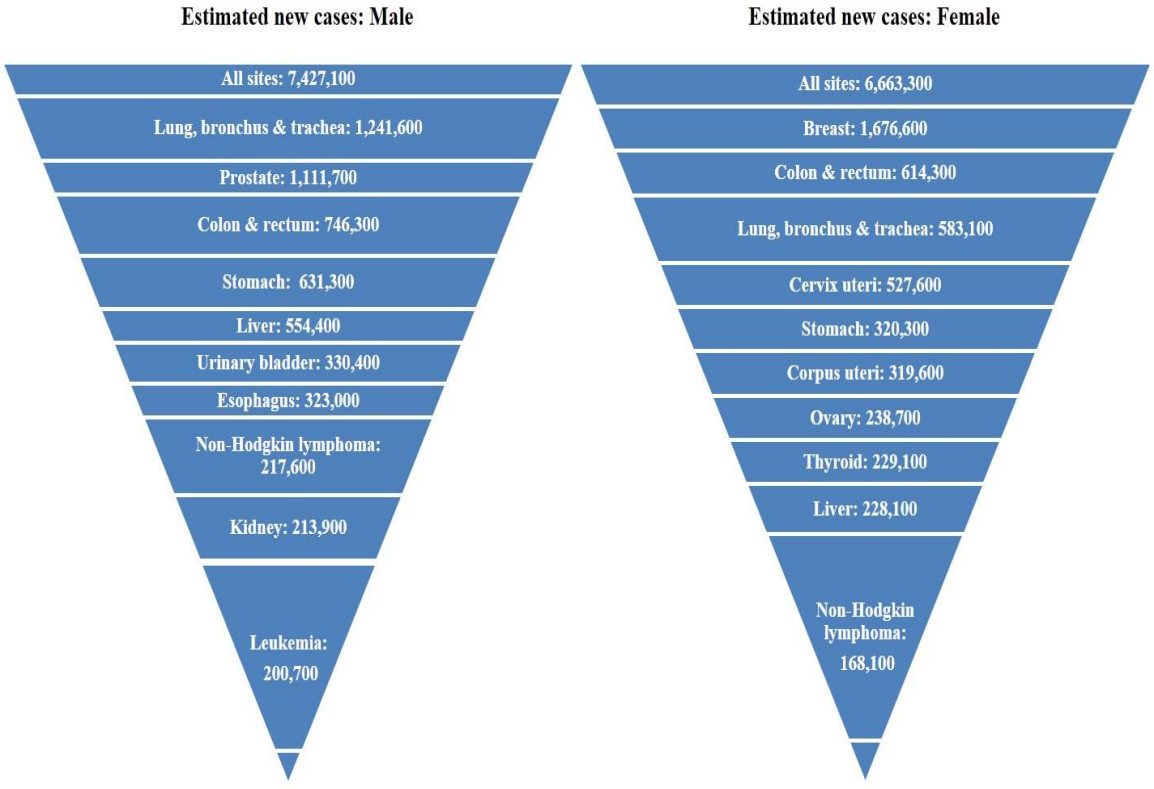


Figure 1.1: A schematic representation of progression of known cancers [9].

1.3. Worldwide cancer statistics

Cancer is considered to be one of the leading causes of human mortality around the world. It is a growing burden both in developed and in developing countries. The GLOBOCAN epidemiology study estimated the cancer worldwide for the year 2012, and predicted that there would be as many as 14.1 million new cancer cases and 8.2 million deaths [10]. **Figure 1.2 (A)** shows the estimated new cancer cases worldwide while **Figure 1.2 (B)** shows the estimated deaths worldwide for the year 2012. Lung cancer in men and breast cancer in women have the highest incidence and mortality rates among all known cancers, both worldwide and also in developing countries. However, in developed countries, prostate cancer in men and lung cancer in women are highly prevalent [11].



(A)

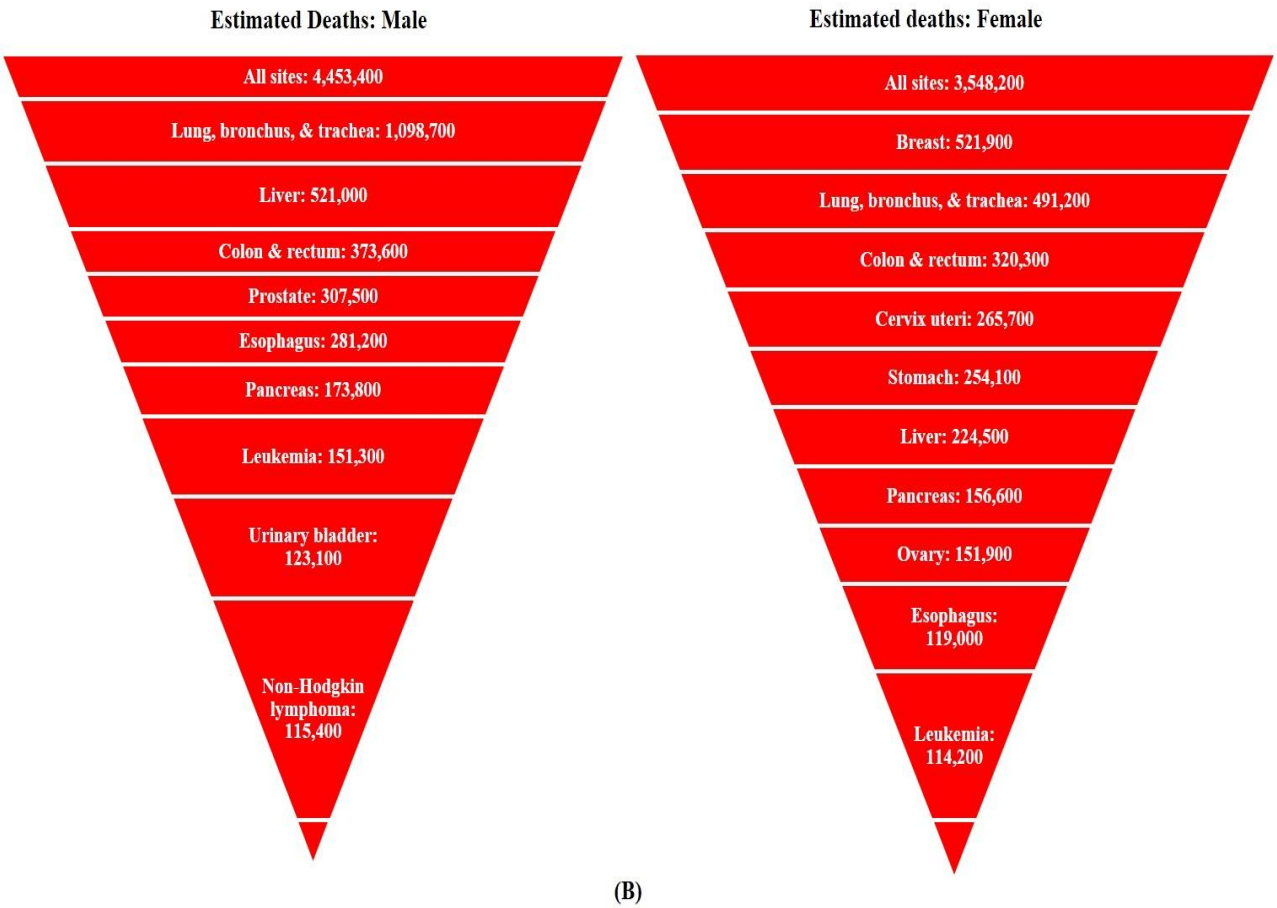
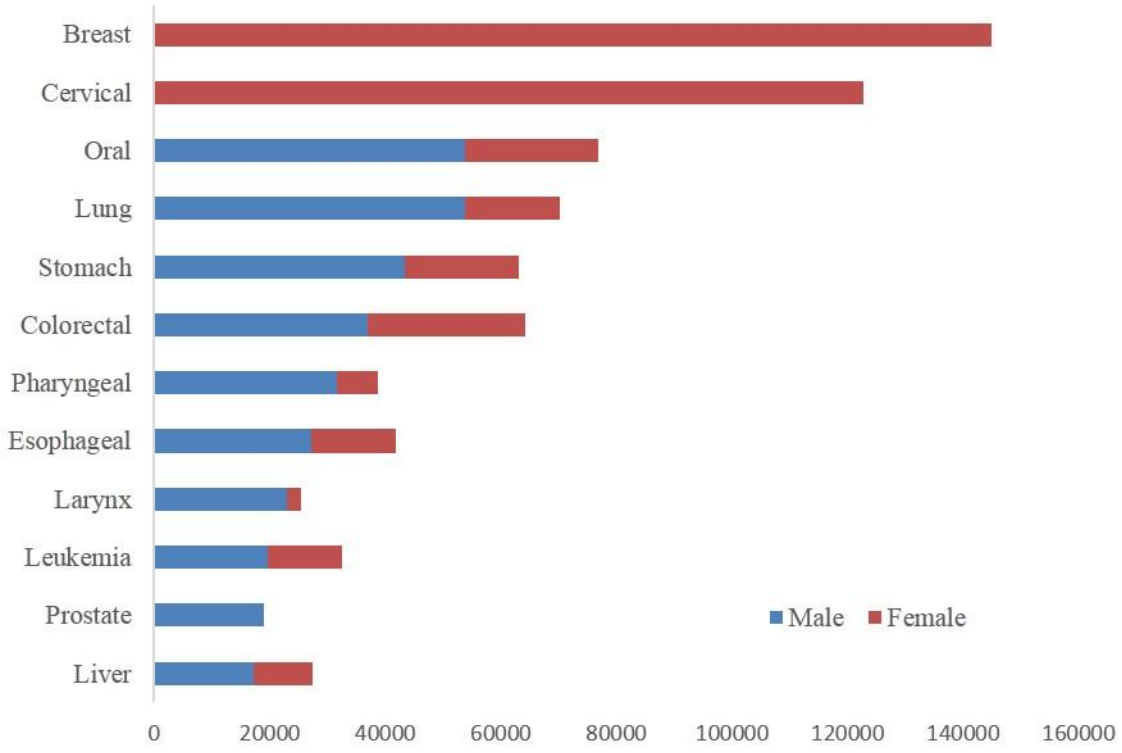


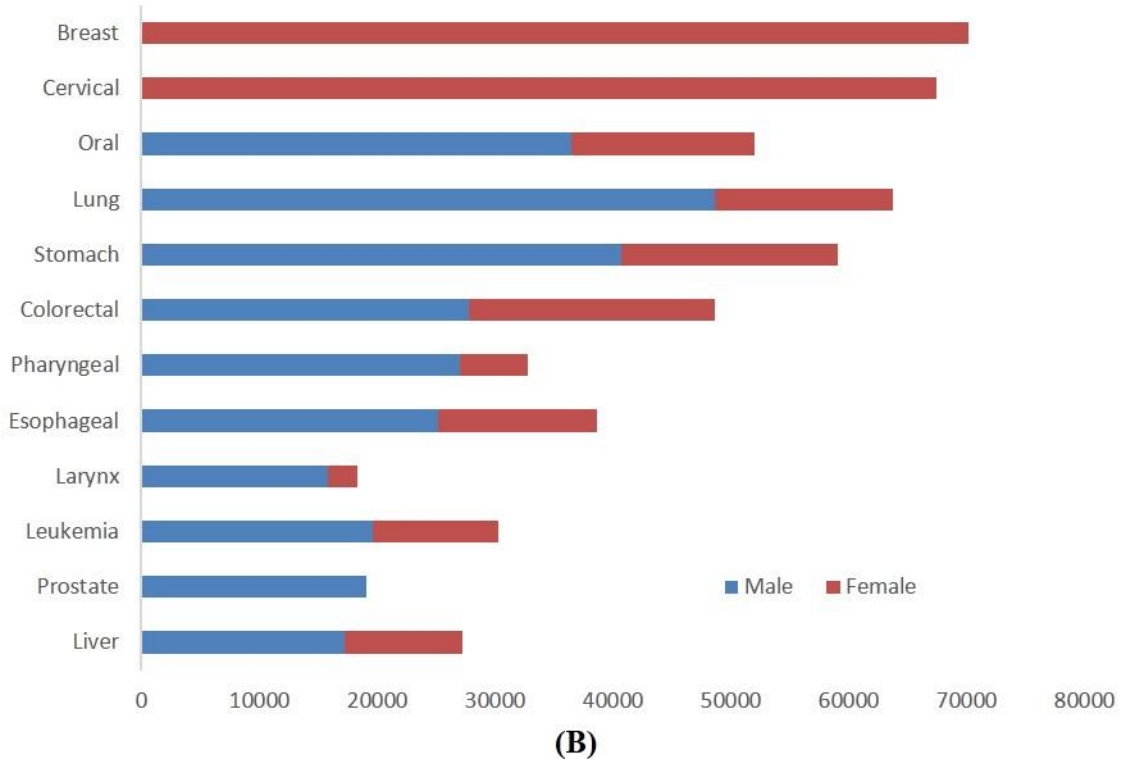
Figure 1.2: Worldwide scenario of (A) estimated new cancer cases and (B) deaths in the year 2012 [11].

1.3.1. Cancer statistics in India

Cancer is a huge burden in India also. Global GLOBOCAN 2012 estimated, one million of the new cases and nearly 0.7 million of the deaths were reported from India itself [12, 13]. In men, tobacco related cancers are more prevalent having high incidence and mortality as seen in **Figure 1.3 (A) and (B)**, whereas breast and cervical cancer have high incidence & mortality in women as seen in **Figure 1.3 (A) and (B)** [12]. Breast cancer remains the major threat among all women across the world and also in Indian population.



(A)



(B)

Figure 1.3: Statistics of Cancer in India (A) Incidence and (B) Mortality in the year 2012 (Blue represents Men and Red represents Women) [12].

1.4. Signalling pathways involved in cancer

Normal cells are using different pathways to carry out metabolic and catabolic activities which are mediated by different signalling pathways. The cancerous cells abrogate the signalling pathways such as Notch, Wnt, Sonic Hedgehog (SHH), Mitogen-Activated Protein Kinase (MAPK), involved in normal cell growth and proliferation [14-23]. Unlike biochemical pathways, the signalling pathways contain transducer protein as well as effector kinases. The members of these pathways, particularly the receptor kinases and the signalling kinases, are often targeted for developing new small molecule inhibitor leads for therapeutic purposes [24-28]. It is well known that, to develop new inhibitors, the three dimensional structure and the molecular mechanism of the target protein play very important roles. Therefore, the atomic level knowledge of the molecular mechanism would save lot of time and money in designing new small molecular inhibitors. MAPK pathway is often targeted as it has two main proto-oncogenes within the pathway, which are Ras (Rat Sarcoma) and Raf (Rapidly Accelerated Fibrosarcoma) [29, 30]. Furthermore, mutations identified in Receptor Tyrosine Kinases (RTKs) and ERK1/2 also leads to various human cancers [31, 32].

1.4.1. Mitogen-Activated Protein Kinase (MAPK) pathway

Since early 1990's MAPK pathway had been continuously studied by researchers across the world [33]. MAPK pathway is one of the cellular signalling pathways and is highly conserved throughout of the evolution. This pathway consists of eight proteins of which two are transducer proteins and six are kinases **Figure 1.4**. This pathway is activated by different growth factors, peptide hormones and neurotransmitters upon binding to Receptor Tyrosine Kinases (RTKs) [34]. RTKs are integral membrane proteins having an extracellular receptor binding domain and an intracellular kinase domain. Upon receptor binding to the extracellular domain, these RTKs undergo auto-phosphorylation and dimerization thereby

creating a docking site for Growth factor Receptor-Bound protein 2 (GRB2) at the cytoplasmic end of the protein. GRB2 is an adaptor protein which links RTKs with Son Of Sevneless (SOS). SOS is a Guanine nucleotide exchange factor which exchanges its bound Guanosine triphosphate (GTP) with the Guanosine diphosphate (GDP) bound on the inactive Ras. Ras is a membrane anchored cytoplasmic small GTPase, which is activated by GTP binding [35-37]. The GTP bound Ras further transduces the signal to the nucleus through a series of steps, the first of which is binding to Raf kinase in the cytoplasm **Figure 1.4**. Ras phosphorylates downstream Raf and activate it, and this activated Raf kinase further binds to MAPK and ERK Kinases (MEK1/2) and activates it by phosphorylation [38]. Activated MEK binds and phosphorylates Extracellular signal-Regulated Kinase (ERK1/2) thereby activating it [39, 40]. The activated ERK along with 3′-phosphoinositide-dependent kinase-1 (PDK1) binds and phosphorylates Ribosomal S6 Kinases (RSK1/2/3/4) [41, 42]. **Figure 1.4** shows the complete activation of MAPK pathway in detail. MAPK pathway, which plays a major role in most of the cellular functions such as growth, proliferation, differentiation and cell migration, is highly unregulated in a majority of the known human cancers [14, 23, 43, 44].

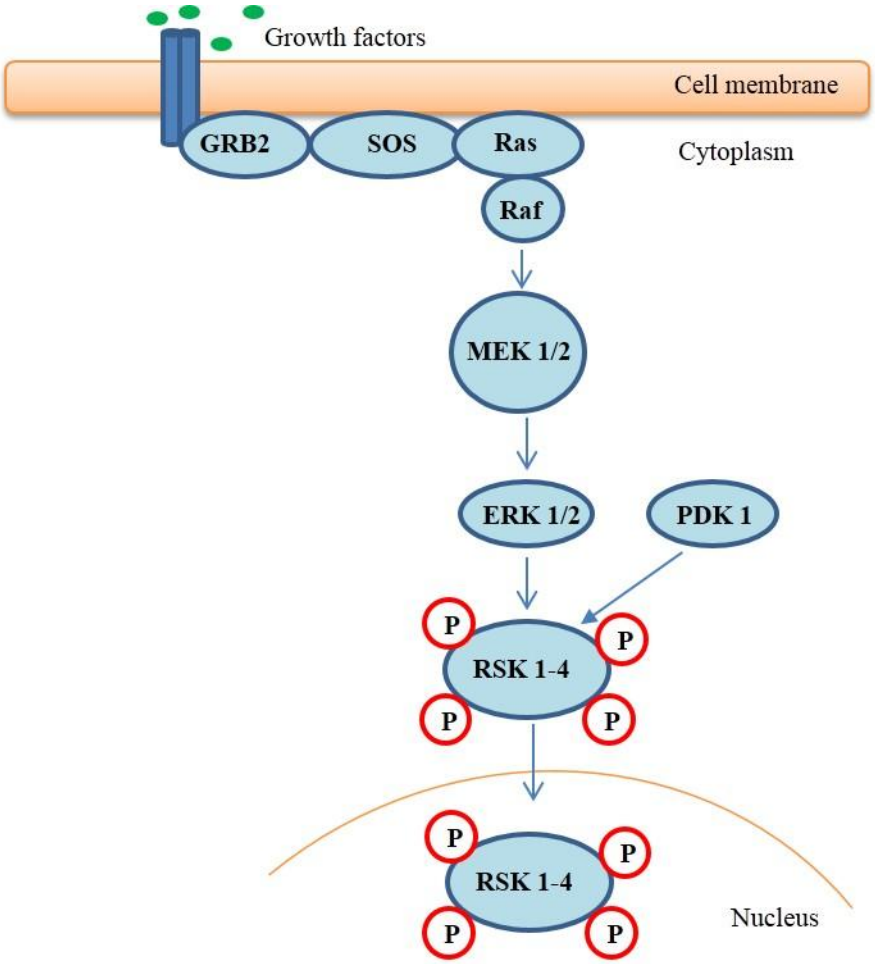


Figure 1.4: Mechanism of activation of MAPK pathway. Growth factors from extra cellular medium binds and activates receptors on plasma membrane which undergoes autophosphorylation, thereby creating docking site for GRB2. It further recruits SOS and Ras sequentially, in the inner side of membrane. The activated Ras, initiates a cascade of phosphorylation events starting with Raf, followed by MEK, then ERK which further activates RSKs leading to the complete activation of MAPK pathway. (GRB2: Growth factor Receptor-Bound protein 2, SOS: Son Of Sevneless, MEK: MAPK/ERK kinase) [71].

Though 3′-phosphoinositide-dependent kinase-1 (PDK1) is also required for RSK activation, ERK is still considered to be the direct activator of RSKs. Upon complete activation, RSKs have a number of substrates to activate like transcription factors such as SRF, ERα, CREB, FOS, JuN, STAT, NFκB and TIF1A inside the nucleus [45-48]. In cytoplasm they are responsible for translation by activating rpS6 and mTORC1, and regulate cell cycle by inhibiting MYT1 and $p27^{KPI}$ through phosphorylation [49-54]. Thus RSKs, self-coordinate their movement into and out of the nucleus depending upon the requirement of the

cell. Thus, the important members of MAPK pathway that transduce the signal from extra cellular cell surface to the nucleus of the cell are ERKs and RSKs. Therefore, these are the subjects of study in this thesis.

1.4.1.1. Extracellular signal-Regulated Kinase (ERK)

ERK belongs to CMGC family, which includes Cyclin-dependent kinases (CDKs), MAPKs, Glycogen Synthase Kinases (GSKs) and CDK-like kinases. ERKs form a sub group within the MAPK family, which also includes c-Jun N-terminal kinases (JNKs) and the p38 MAPKs [55]. So far, eight isoforms of ERK have been discovered while, ERK1 and ERK2 are the major isoforms [56]. ERK1 is also known as MAPK3 and its gene is found on chromosome 16 (16p11.2), whereas ERK2, known as MAPK1, is coded by a gene on chromosome 22 (22q11.21). ERKs discharge important functions both in the nucleus and in the cytoplasm. Interestingly, ERK2 functions as dimer when inside the nucleus and as monomer when in the cytoplasm. ERK1 and ERK2 shares almost 88 % sequence homology. However, knockout studies on mice has revealed that ERK2 is indispensable for survival, as ERK2 knockout mice die very early during development. On the other hand, ERK1 knockout mice do survive [57, 58].

1.4.1.1.1. ERK Domain organization

ERK2 is a serine/threonine kinase which is directly activated through phosphorylation by MEK1/2 (MAPK/ERK kinase) and also known as Mitogen-Activated Protein Kinase Kinase (MAPKK or MAP2K) [59-62]. The ERK2 contains 360 amino acids and its kinase domain comprises from 25-313 amino acids which also includes the signature motif Thr-X-Tyr from 185-187 amino acids. The activation of ERK requires dual

phosphorylation of Threonine and Tyrosine within the activation loop of signature motif TXY (X is any amino acid) [63].

1.4.1.1.2. The structural basis of ERK2

The three dimensional structure of human ERK2 have been determined using x-ray crystallography. The Structure of ERK2 is bilobal having one small N-lobe and a large C-lobe. The N-lobe has five antiparallel β-strands (β1-5) and one helix (αC). It contains a conserved glycine-rich (GXGXXG) ATP binding loop, also known as P-loop, between the β1 and β2 strands. The C-lobe has six α-helices (αD-I) and four short β-strands (β6-9) which occur between αE and αF helices [64, 65]. **Figure 1.5** shows the complete primary structure of human ERK2, and also indicating the minor and major segments and sub domains [65, 66].

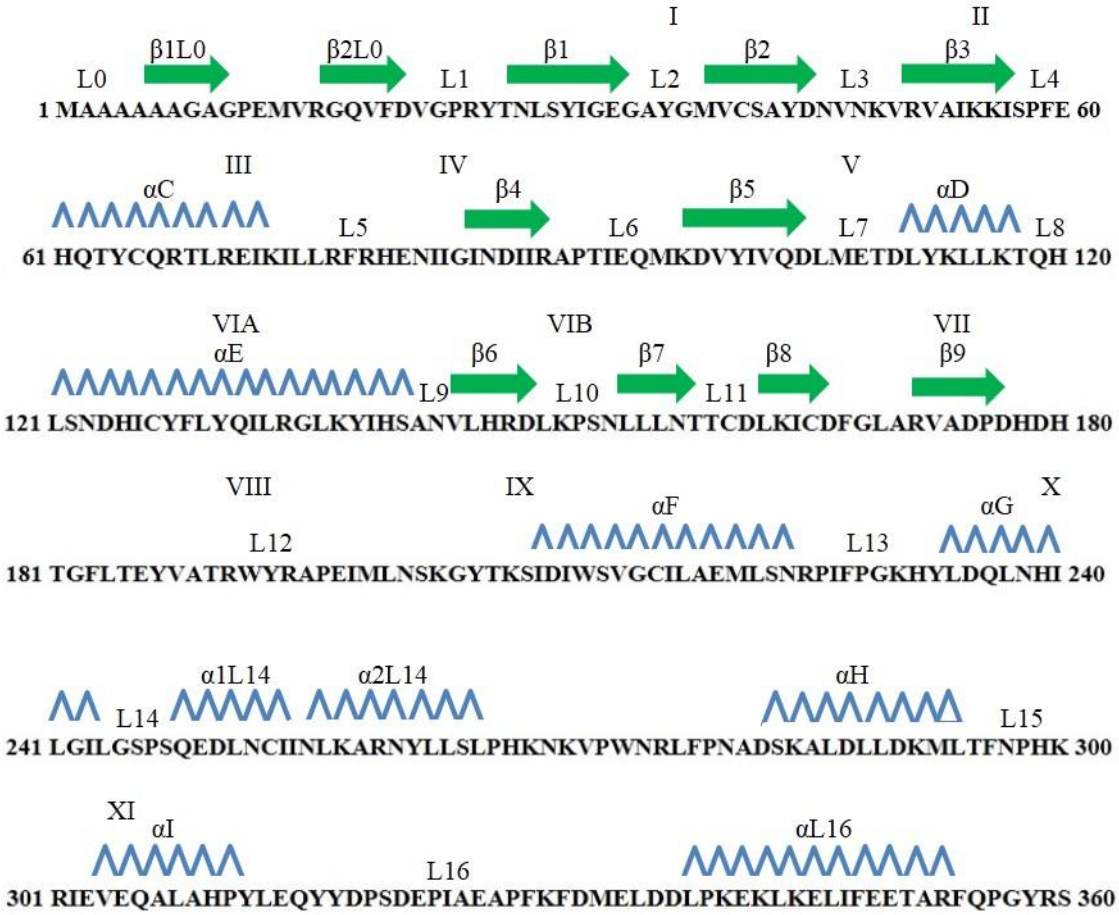


Figure 1.5: Primary structure of human ERK2. The positioning of α-helices (zigzag) and β-sheets (arrows) are labelled sequentially forming the secondary structure of ERK2. The roman numerals on the top indicates the sub-domains of ERK2 [61].

Figure 1.6 shows the structure of ERK2 based on PDB id: 1TVO as viewed in Pymol software. The amino acid residues form the catalytic triad of ERK2: Asp149, Lys151 and Thr190 are shown as green sticks. The residues Thr185 and Tyr187 which undergo phosphorylation are shown in red sticks. The L16 loop which contains both the Nuclear Export Signal (NES) and the dimeric interface, is shown in dark blue [67].

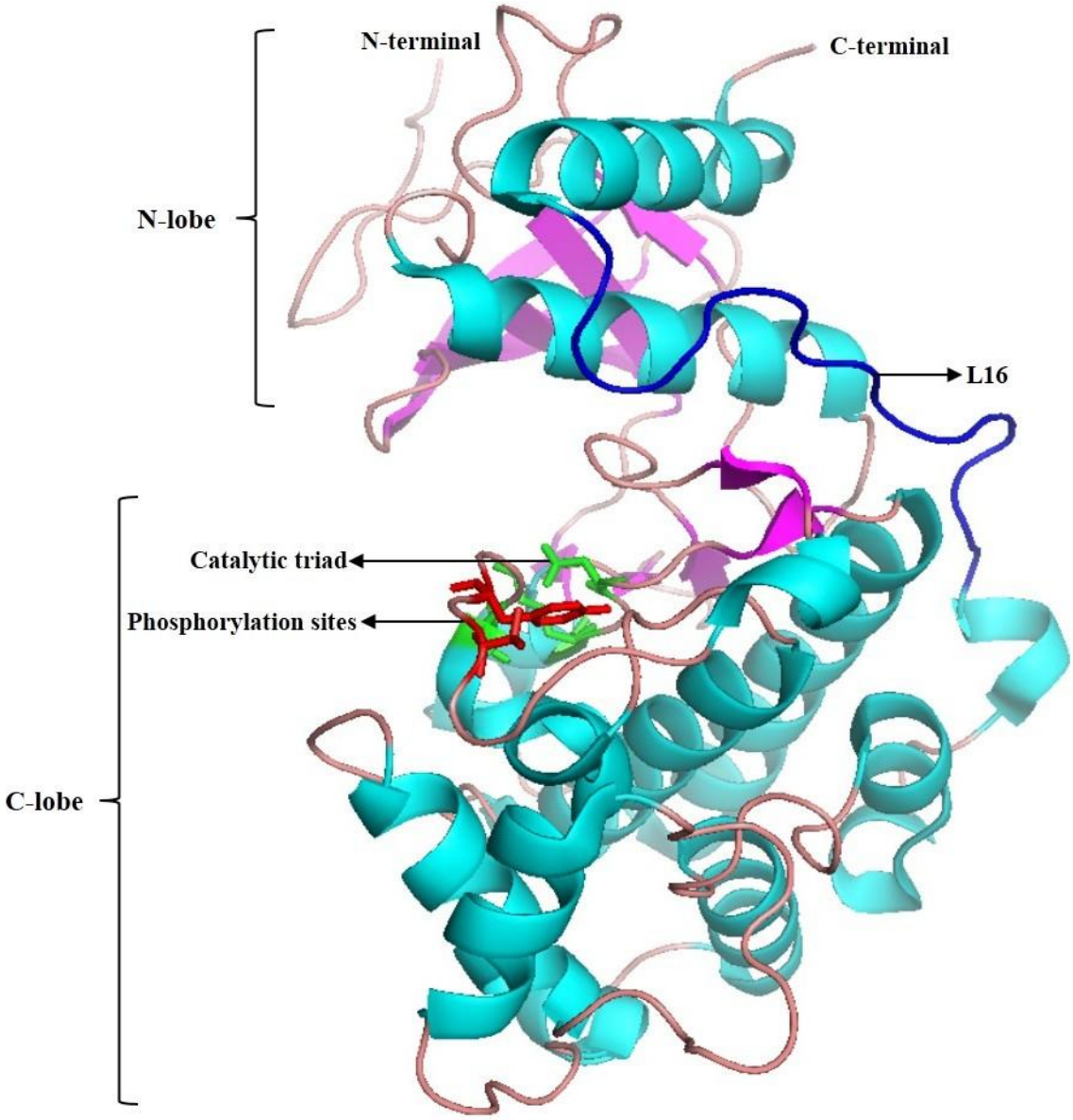


Figure 1.6: Structure of human ERK2. Cartoon diagram of human ERK2 showing it is bilobal based on PDB id: 1TVO. The catalytic triad is shown in green sticks. Thr185 and Tyr187 which undergoes phosphorylation are shown in red sticks. The L16 loop is shown in dark blue.

ERK2 shuttles between the nucleus and the cytoplasm by utilizing these Nuclear Export Signal (NES) and Nuclear Localization Signal (NLS) motifs [68-70]. The dimeric interface (DI) is close to the NES. Upon dimerization of ERK2, NES is completely masked as it is close to the dimeric interface and the ERK2 dimer which has two active NLS, gets translocated into the nucleus [71]. **Figure 1.7** shows the dimeric structure of ERK2.

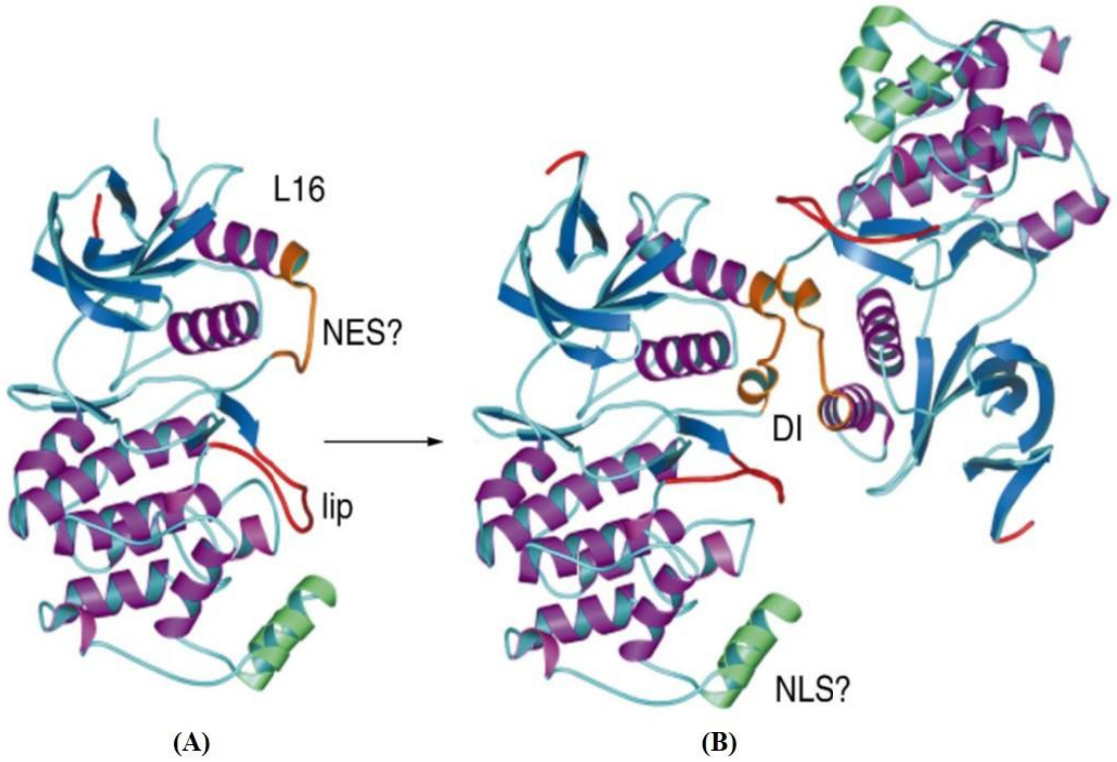


Figure 1.7: Dimeric structure of ERK2. The transition of ERK2 monomer (A) to ERK2 dimer (B). The NES within the L16 is close to the dimeric interface (DI) is being masked in the dimeric structure [71].

1.4.1.1.3. Negative control by dephosphorylation

The MEK1/2 activates ERK1/2, and the MAPK pathway is active as long as ERK2 is in bisphosphorylated state, and the transcription factors are continuously active throughout the period. The MAPK activity is however reversed by dual-specificity MAP kinase phosphatases (DUSPs) which negatively regulate ERK by dephosphorylation [72, 73]. Thus MEK phosphorylation and DUSP dephosphorylation regulate the overall functioning of ERK1/2 within the cell.

1.4.1.1.4. Chemical inhibitors of ERK1/2

Small molecules inhibitors have already been reported which specifically bind to ERK1/2. FR180204 has a pyrazolopyridazine parent ring and has a IC_{50} of 0.33 µM [67]. It is an ATP competitive inhibitor of ERK. FR148083, a benzoxa cyclotetradecine based inhibitor, irreversibly inhibits ERK2 and has a IC_{50} of 80 nM [74]. **Figure 1.8** shows these two inhibitors of ERK2.

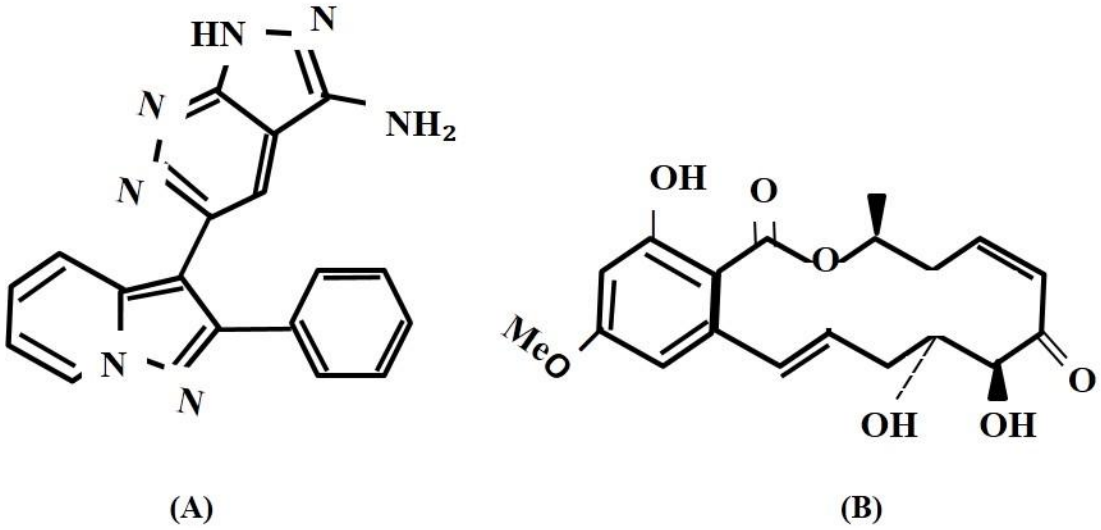


Figure 1.8: ERK2 inhibitors. (A) FR180204 and (B) FR148083 [65].

1.4.1.1.5. Substrates of ERK

Though there are many substrates of ERK, RSKs are considered very important among all, as they serve terminal mediators of external signal to the nucleus of the cell.

1.4.1.2. Ribosomal S6 Kinases (RSKs)

Ribosomal S6 Kinases belong to Ribosomal S6 Kinase (RSK) family of protein kinases involved in signal transduction [75]. There are two subfamilies within RSK family, and they are p90RSK and p70RSK. The p90RSK is also known as MAPK Activated Protein Kinase-1 (MAPKAPK-1) and has four isoforms of p90RSK in humans, RSK1/2/3/4. It also has two structurally related homologues known as RSK like protein kinases (RLPK) or Mitogen and stress activated kinases (MSK1 and MSK2) [75, 76]. The p70RSK is also known as S6 kinase which includes two isoforms S6K1 and S6K2. The main distinguishing feature between p90RSK and p70RSK is that the 90 kDa family contains two non-identical kinase domains, while the 70 kDa family contains only one kinase domain [77, 78]. **Table 1.1** shows the members of all the genes of RSK family, and their chromosomal location in the human genome.

Table 1.1: Overview of p90 and p70 RSKs of RSK family. The gene names and chromosomal location is indicated in the right panels.

Protein	Other names	Gene name	Location
RSK1	RSK 90 kDa, polypeptide 1	RPS6KA1	1p
RSK2	RSK 90 kDa, polypeptide 3	RPS6KA3	Xp22.2
RSK3	RSK 90 kDa, polypeptide 2	RPS6KA2	6q27
RSK4	RSK 90 kDa, polypeptide 6	RPS6KA6	Xq21.1
MSK1	RLPK, RSK 90 kDa, polypeptide 5	RPS6KA5	14q31
MSK2	RSKB, RSK 90 kDa, polypeptide 4	RPS6KA4	11q11
S6K1	S6Kα, RSK 70 kDa, polypeptide 1	RPS6KB1	17q23.1
S6K2	S6Kβ, RSK 70 kDa, polypeptide 2	RPS6KB2	11q13.2

1.4.1.2.1. Domain organization of RSKs

The four isoforms of RSKs display up to 75 % sequence identity. They contains eight phosphorylation sites and the two independent kinase domains toward each end of the single polypeptide chain as shown in **Figure 1.9**. The N-Terminal Kinase Domain (NTKD) is connected with the C-Terminal Kinase Domain (CTKD) by a conserved linker region of ~ 100 amino acids. The two kinase domains are flanked by N and C-terminal tails of ~ 60 amino acids each [77]. The NTKD shares a greater homology with AGC family of protein kinases [79]. It is the major kinase domain as it is involved in substrate phosphorylation by binding to various substrates, both in the cytoplasm and in the nucleus, however, the CTKD belongs to Calcium/calmodulin-dependent kianses (CaMK) family [80]. CTKD is involved in auto-phosphorylation of RSK, which is one of the mandatory steps in the complete activation of RSKs [81, 82]. The C-terminal tail of RSK is basic in nature and houses D-domain which binds to the ERK. The D-domain has a characteristic ERK binding motif containing NH_2-Leu-Ala-Gln-Arg-Arg-COOH sequences which is conserved in all RSK isoforms. **Figure 1.9** shows domain organization of RSK and conserved residues of all four human RSK isoforms. The percentage sequence homology by region is also indicated in **Figure 1.9**.

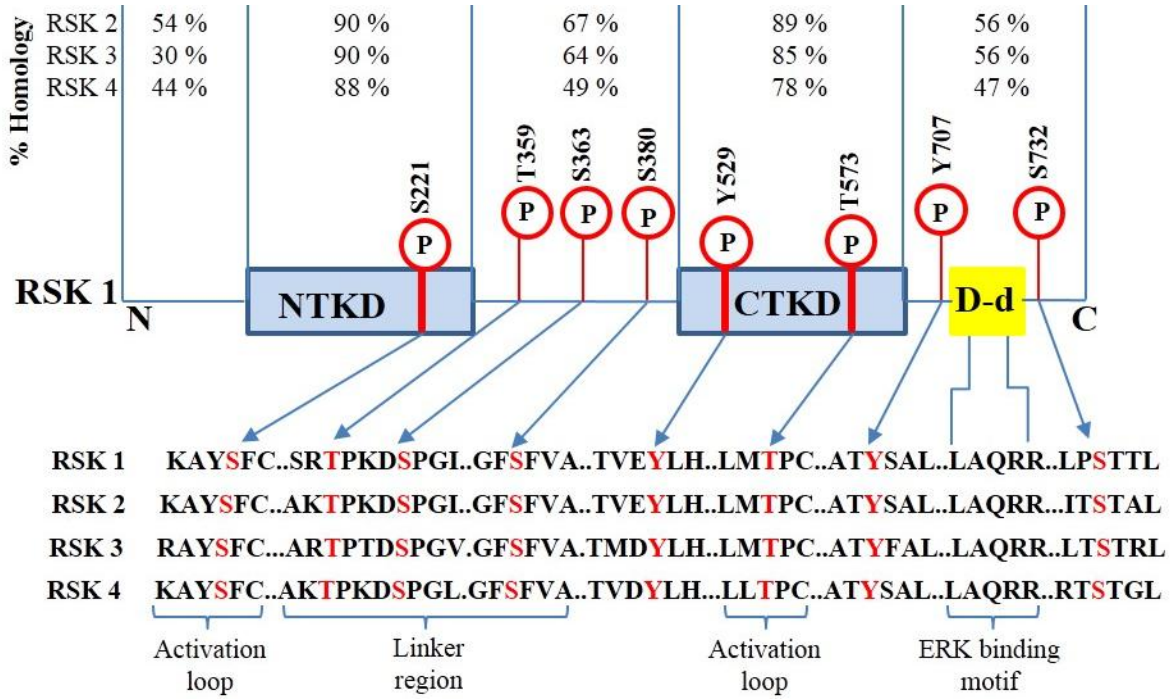


Figure 1.9: Domain organization of RSKs. The domain organization of RSK1 showing the NTKD and CTKD separated by linker region. The NTKD and CTKD is flanked by N and C-terminal tails on either ends. The D-domain within the C-terminal tail is shown in yellow box. The important phosphorylation residues of RSK1 are marked 'P' inside a circle. The percentage homology of RSK1 with respect to RSK2, 3 and 4 isoforms are marked above. Below the domain organization is the consensus sequence of all four isoforms with respect to important regions harboring the phospho-residue [82].

1.4.1.2.2. Mechanism of RSK1 activation

RSK activation is a complex process which involves multiple phosphorylation steps, and occur sequentially one after another. Upon ERK binding to RSK1 on the ERK binding D-domain at the C-terminal end, ERK phosphorylates Thr573 within the activation loop of CTKD. ERK also phosphorylates RSK1 at Thr359 and Ser363, whose function is not known. Upon Thr573 phosphorylation by ERK, the CTKD of RSK1 is activated which then auto phosphorylates Ser380 in the linker region, thereby creating a docking site for 3′-phosphoinositide-dependent kinase-1 (PDK1). PDK1 phosphorylates Ser221 within the activation loop of NTKD. The fully active RSK1 further phosphorylates Ser732 which

weakens the association of RSK1 with ERK at the C-terminal end thereby causing the dissociation of ERK [75, 83-87]. **Figure 1.10** shows the complete activation of RSK1 which requires multiple phosphorylation steps including auto-phosphorylation. Upon activation of RSKs a complete cycle of MAPK pathway is completed starting from the activation of Receptor Tyrosine Kinases (RTKs) on the cell membrane. The fully activated RSKs activate many substrates both in nucleus and in cytoplasm depending upon the requirement of the cell.

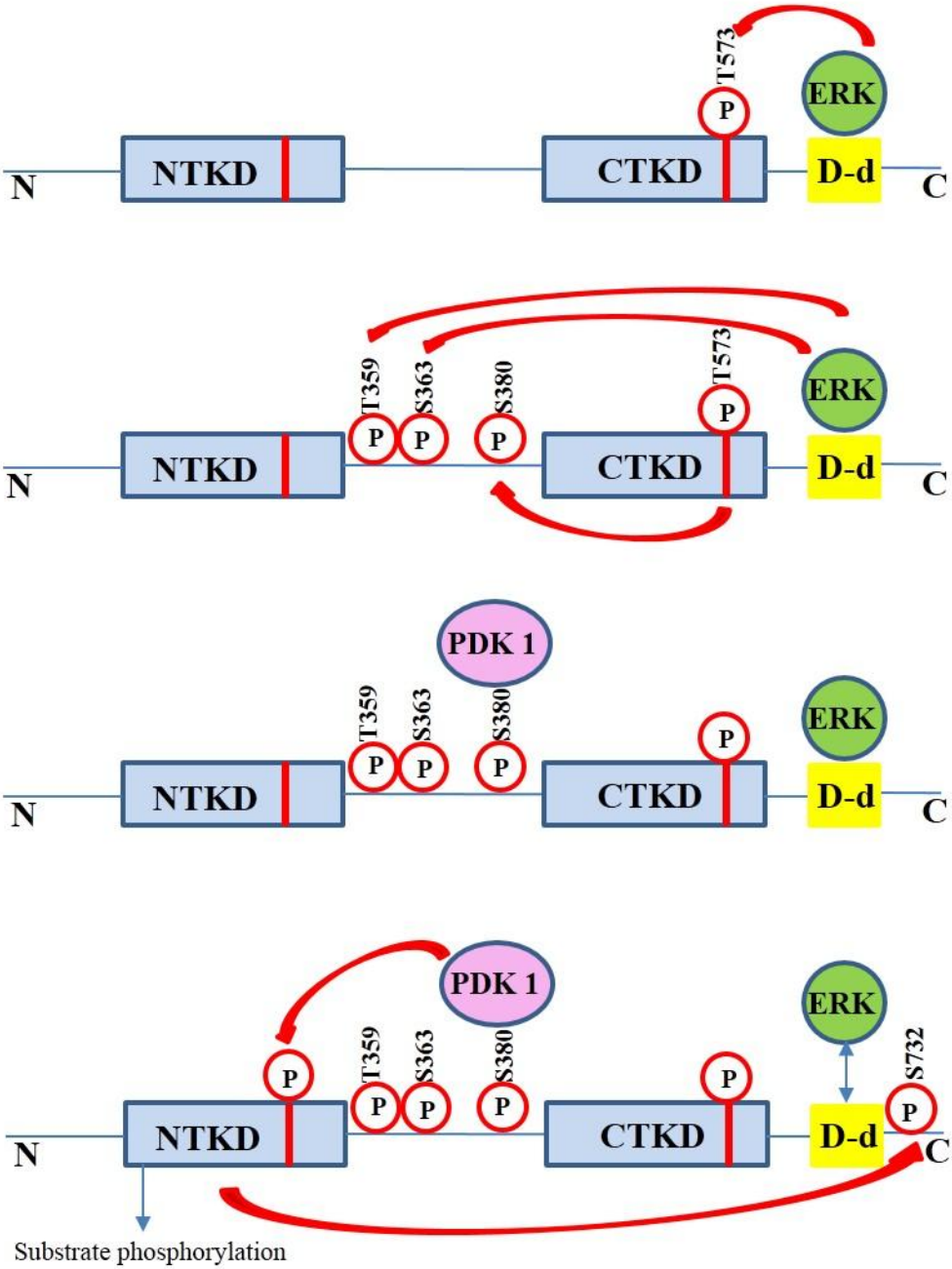


Figure 1.10: Mechanism of RSK1 activation. The sequential activation of RSKs by ERK, followed by auto-phosphorylation. Further phosphorylation by PDK1 activates RSK1. Flowing RSK1 activation the NTKD, further phosphorylates Ser732 which dissociates ERK at the D-domain within the C-terminal tail of the RSK1 [75].

1.4.1.2.3. Role of RSKs in Translation

Since Ribosomal S6 Kinases nomenclature is based on its ability to activate rpS6, which is the ribosomal smaller subunit 40S protein 6, they have a major role in translation. The major pathway involved in translation is the PI3K–mTOR pathway which is activated by binding of Insulin or Insulin-like growth factors (IGFs) on the plasma membrane receptors [88-91]. These receptors facilitate the inter-conversion of phosphatidylinositol-4,5-bisphosphate to phosphatidylinositol-3,4,5-trisphosphate, which is mediated by Phosphoinositide 3-kinase (PI3K) on the inner side of the membrane [92]. The Phosphatidylinositol-3,4,5-trisphosphate creates docking sites for PDK1 and AKT, also known as Protein Kinase B (PKB). The AKT negatively regulates Tuberous Sclerosis Complex (TSC) by phosphorylation [93, 94]. Activated RSKs are also required for the negative regulation of TSC [51]. This, in turn results in relieving of the inhibitory effect of TSC on Ras Homolog Enriched in Brain (RHEB) [95, 96]. Upon RHEB activation, it phosphorylates and activates mammalian Target Of Rapamycin complex-1 (mTORC1) [97]. The mTORC1 activates S6K1/2 (p70RSKs) which along with p90RSKs activates and initiates cap-dependent translation in eukaryotes. The mTORC1 negatively regulates 4EBP1, from which eukaryotic translation Initiation Factor-4E (eIF4E) is relived from inhibition. The eIF4E is a cap binding protein which initiates assembly of translation machinery at the ribosome which includes eIF4B and eIF3 [98, 99]. eIF4B is activated by phosphorylation by both p90RSKs and p70RSKs. An adaptor protein eIF4G binds to eIF4E and assembles eIF4A, which is a helicase, at the other end. This heterotrimeric complex of E, G and A

constitutes eIF4F, which regulates eukaryotic translation rates [90, 100]. **Figure 1.11** is a schematic diagram illustrating the role of p90RSKs and p70RSKs in Eukaryotic translation.

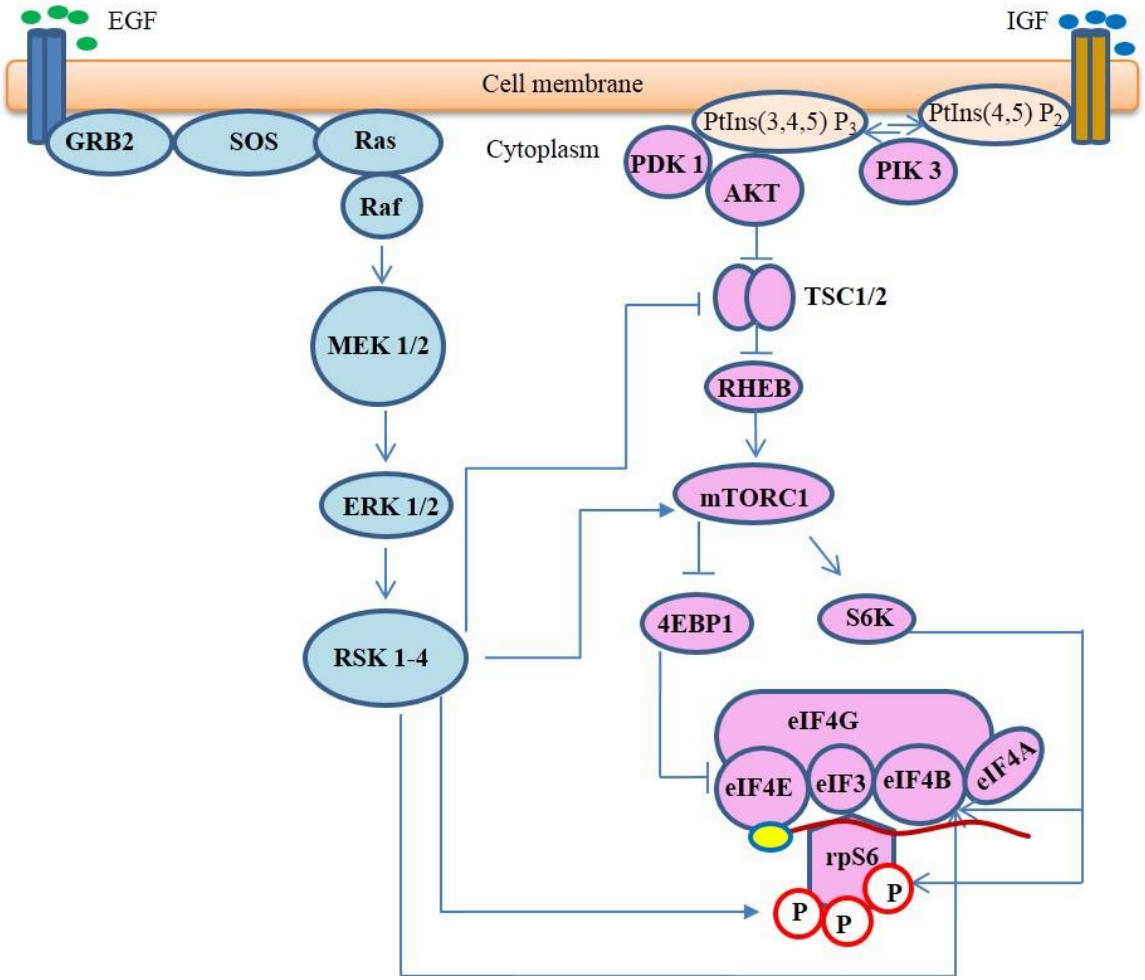


Figure 1.11: Role of RSK in translation. The MAPK and PI3K–mTOR pathways are required for initiating translation machinery. (PtdIns(4,5)P2: Phosphatidylinositol-4,5-bisphosphate; PtdIns(3,4,5)P3: Phosphatidylinositol-3,4,5-trisphosphate) [75].

1.4.1.2.4. Structure of RSKs

So far no full length structure of RSK has been reported for any of its isoforms. However, crystal structures of individual domains NTKD and CTKD have been determined for human RSK1 and RSK2 isoforms [101-105]. Further, our search could not find any

structure in the PDB for human RSK3 isoform. An attempt has been also made to crystallize the full length RSK.

1.4.1.2.4.1. Structure of NTKD of RSK

The NTKD of RSKs comprises of ~ 260 amino acids starting from ~ 58-315 amino acids depending on the full length isoform and the organism. As far as sequence is concern, it belongs to AGC family of kinases. **Figure 1.12** shows the RSK2 NTKD (PDB id: 3G51) which is bilobal containing a small N-terminal lobe which is β-rich and a large α-helical C-terminal lobe. The structure was solved as a complex with AMP-PNP (5'-Adenylyl-imidodiphosphate), which is a non-hydrolyzable ATP analogue, by molecular replacement using NTKD of Mitogen and stress activated kinase 1 (MSK1) with the PDB id: 1VZO [102, 106]. **Figure 1.13** shows the important interactions between RSK2 NTKD amino acids and AMP-PNP as analysed in Ligplot [107]. These interacting amino acids constitute the core ATP-binding site, placed between the lobes, and is required for transferring the γ-phosphate of the ATP to the downstream substrates. The activation loop of RSK2 NTKD lies within residues 205–232. **Figure 1.14** shows the zoomed view of Lys216 of the activation loop makes prominent contacts with β-phosphate and Asp211.

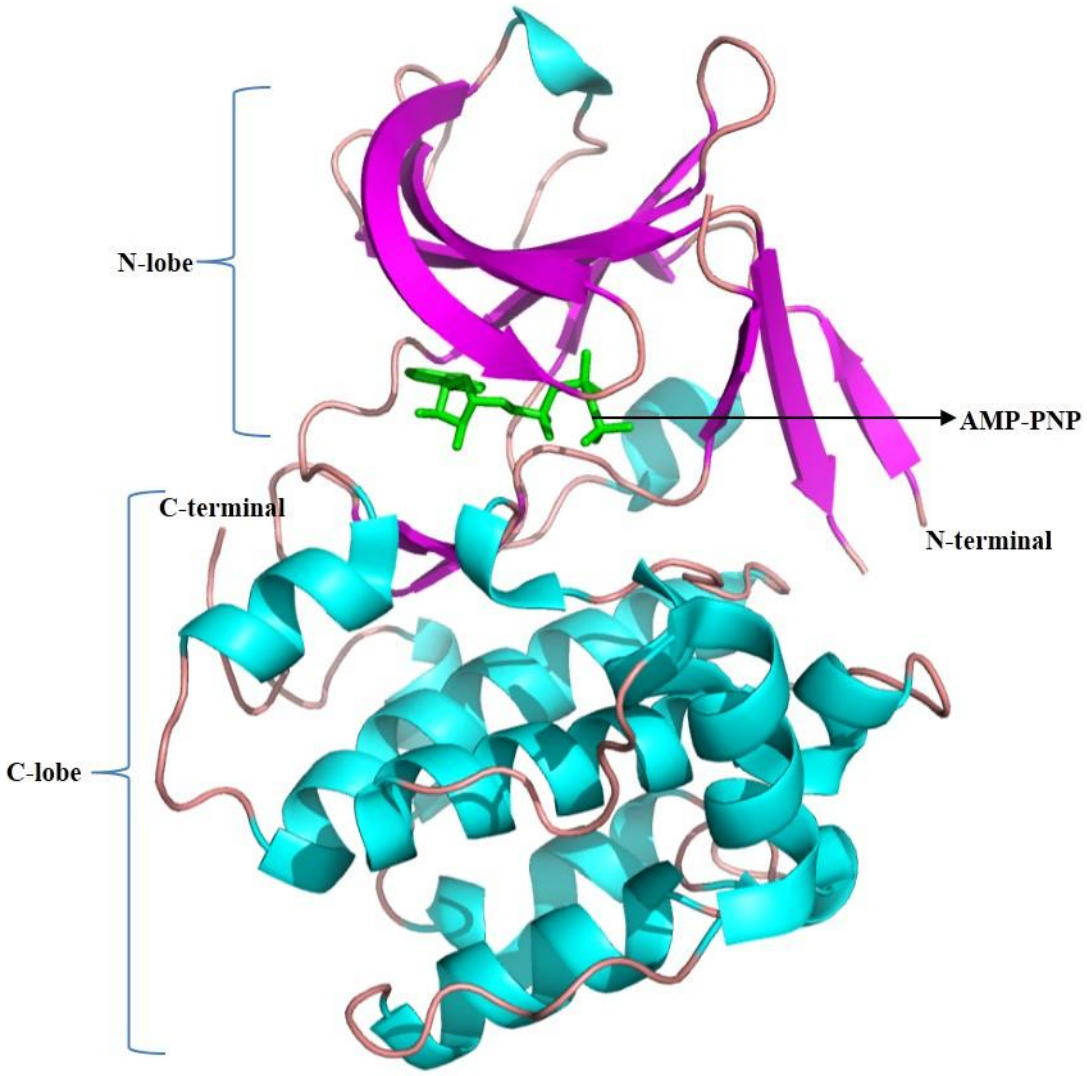


Figure 1.12: Structure of RSK2 NTKD. Ribbon diagram of RSK2 NTKD based on PDB id: 3G51, showing it is bilobal. The α-helices are shown in cyan and β-strands are in magenta colour. The AMP-PNP molecule is shown in green sticks [102].

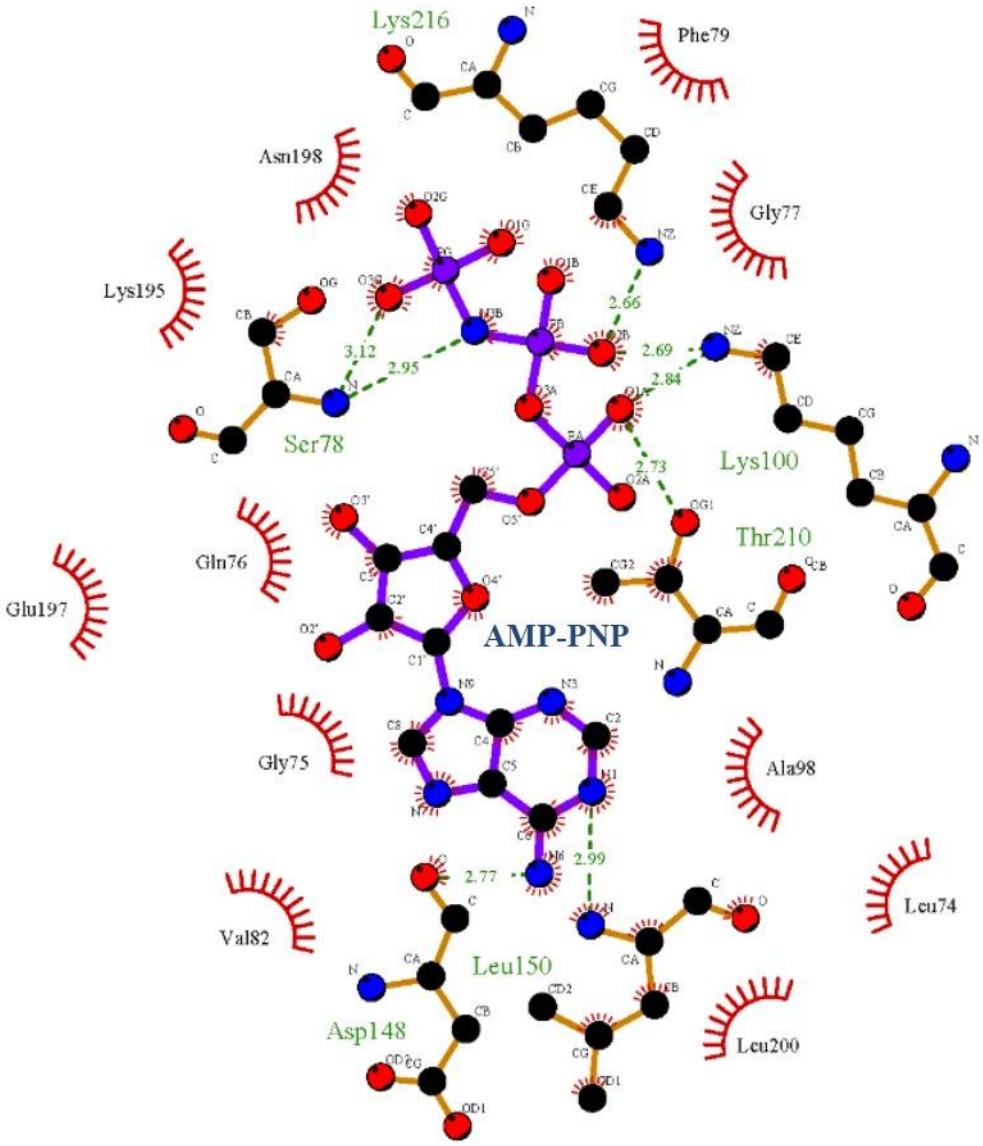


Figure 1.13: Binding association between RSK2 NTKD and AMP-PNP. Ligplot representation of important interactions between the amino acids of RSK2 NTKD and AMP-PNP. The ligand bond is shown in blue colour. The hydrogen bonds are shown in dashed green lines and the corresponding distance is indicated [102].

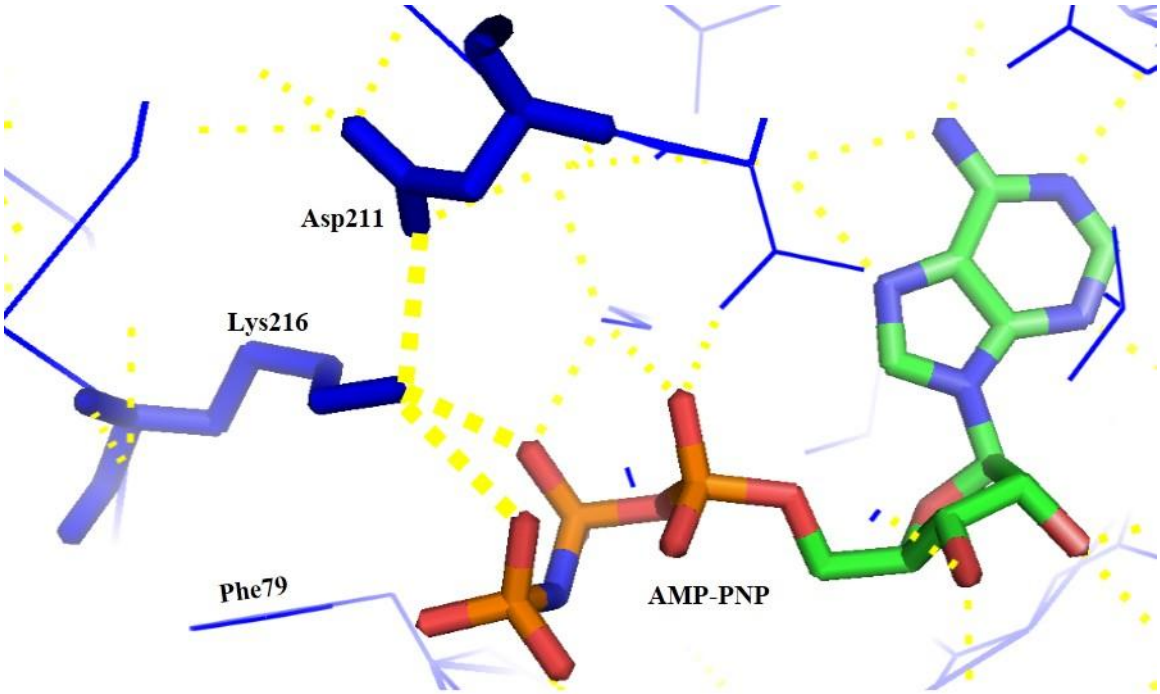


Figure 1.14: Interactions by Lys216 of RSK2 NTKD with AMP-PNP. Lys216 of the activation loop of RSK2 NTKD making prominent contacts with β-phosphate of AMP-PNP and Asp211 [102].

1.4.1.2.4.2. Structure of CTKD of RSK

The CTKD of RSKs comprises of ~ 260 amino acids, starting from ~ 415-672 amino acids depending on the full length isoform and the organism. Like NTKD, the RSK2 CTKD is also bilobal. The smaller N-lobe contains five antiparallel β-strands (β1-β5) and one α-helix (αC), whereas the larger C-lobe is α-rich. The main feature of the inactive structure is the large α-helix (αL), which is within the substrate binding grove, and keeps the domain inactive leading to auto-inhibition of this domain. Upon ERK activation, the αL is displaced thereby exposing the substrate binding groove which results the activation of the domain. **Figure 1.15** shows RSK3 CTKD inactive structure highlighting the important α-helices and β-strands [103].

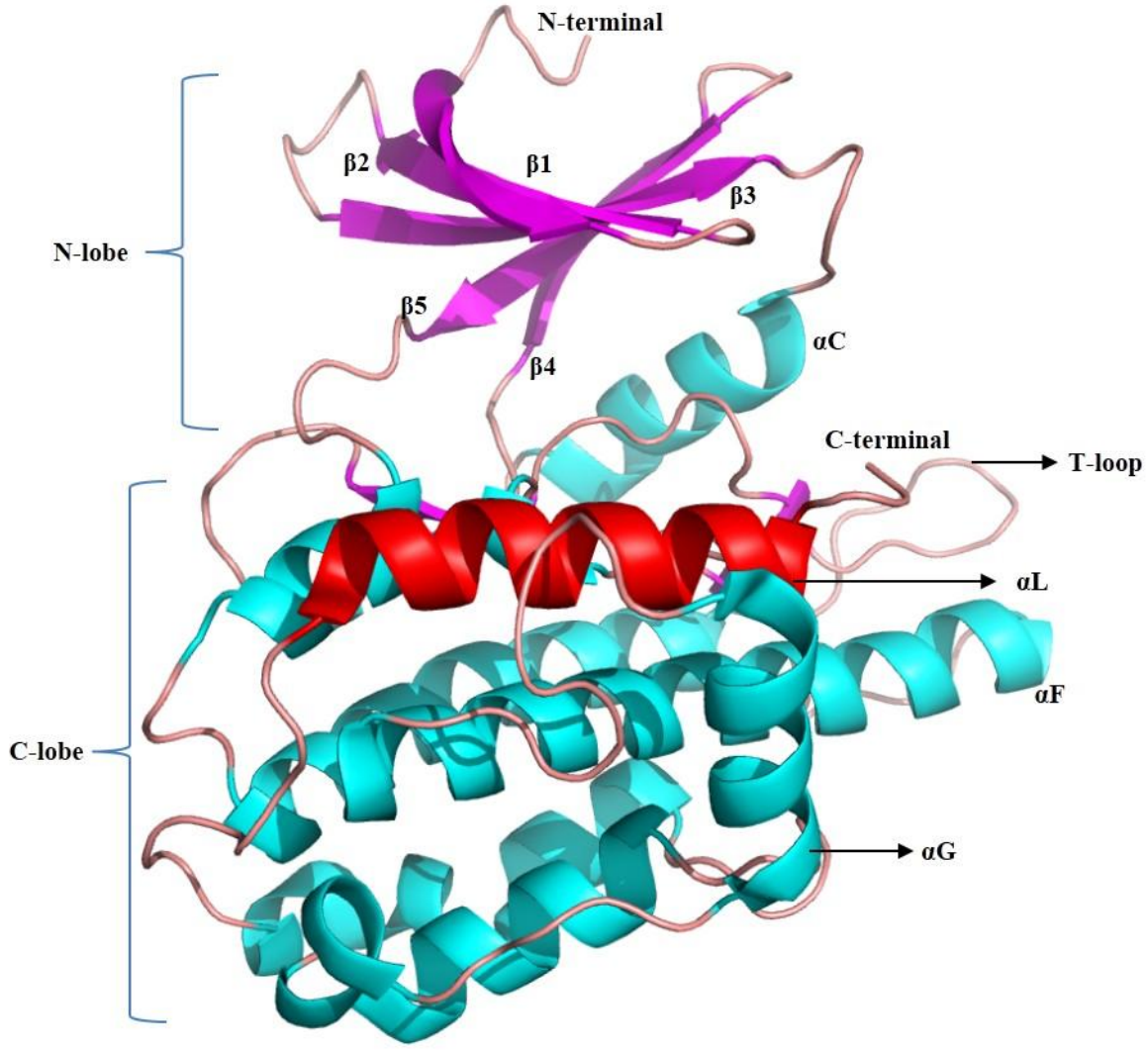


Figure 1.15: Structure of RSK2 CTKD in inactive confirmation. Ribbon diagram of RSK2 CTKD showing it is bilobal. The important α-helices and β-strands have been labelled. The auto-inhibitory αL helix is shown in red which is occupying the space near the substrate binding grove of the CTKD in the inactive conformation [103].

The three dimensional structure of inactive RSK1 CTKD has been determined in 2012. The structure is similar to that of RSK2 CTKD and the RMSD is 0.609 Å. **Figure 1.16** shows the structure of RSK1 CTKD depicting the position of αL resembling as that of a cradle, which is perfectly residing near the substrate binding grove, thereby auto-inhibiting the enzyme [104].

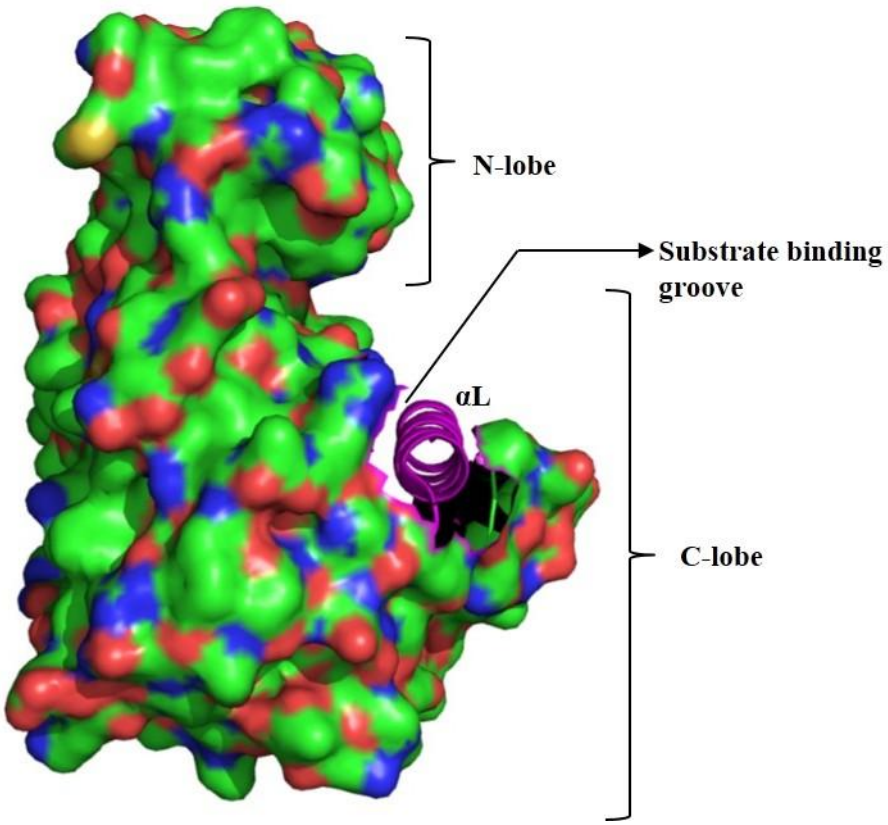


Figure 1.16: Surface view of RSK1 CTKD in inactive confirmation. The auto-inhibitory αL helix in ribbon is seen as a cradle occupying the substrate binding groove [104].

1.4.1.2.5. Role of RSKs in diseases

The mutations in RSK2 are associated with Coffin–Lowry syndrome (CLS), which is an X-linked syndrome [108]. RSK2 is encoded by RPS6KA3 gene and the locations is Xp22.2 in the human genome. The CLS patients have severe mental retardation and also have characteristic phenotypes which includes microcephalic, kyphoscoliosis (kyphosis and scoliosis), Micrognathia, stunted growth and facial dysmorphism. So far ~ 128 mutations had been discovered within the RSK2 gene isolated from the patients [109-111].

The role of RSKs in human cancers is well documented and plays a significant role in progression of most of the cancers especially prostate cancer, breast cancer and Multiple Myeloma (MM) [112-114].

1.4.1.2.6. Inhibitors of RSKs

There are several kinase inhibitors that non-specifically inhibit RSKs such as staurosporine, Indirubin-3'oxime and dihydroxylnapthalene (NSC356821) [115-119]. **Figure 1.17 (A)** shows structure of Staurosporine. The recent RSK specific inhibitors have been broadly classified into three classes on the basis of the structure of the parent molecule of the inhibitor. They are 1) kaempferol based glycosides, 2) dihydropteridinones and 3) pyrrolopyrimidines [119].

SL0101 is a kaempferol based glycoside, which is an acetylated flavonol and occurs naturally in a plant *Fosteronia refracta*. SL0101 is the first discovered specific inhibitor of RSKs and is widely used to specifically inhibit RSKs. It's IC_{50} for RSK2 is 90 nM in presence of ATP at 10 µM concentration. SL0101 is an ATP competitive reversible inhibitor for RSKs and binds to the ATP binding pocket within the NTKD [120, 121]. **Figure 1.17 (B)** shows structure of SL0101 which typically contains a kaempferol moiety and a sugar moiety which is rhamnose. The rhamnose moiety contributes maximum to the inhibitory effect of SL0101 on RSKs [122].

BI-D1870 is a Dihydropteridinone based specific inhibitor of RSKs and is currently one of the most potent RSK inhibitors with a IC_{50} of 15-30 nM in in presence of ATP at 100 µM concentration. **Figure 1.17 (C)** shows the structure of BI-D1870. Like SL010, BI-D1870 is also a reversible inhibitor for RSKs and binds to the ATP binding pocket within the NTKD [123, 124].

A Fluoro-methyl-ketone (FMK) derivative of pyrrolopyrimidine, known as Pyrrolopyrimidine FMK is the most potent RSK inhibitors with a IC_{50} of 15 nM. **Figure 1.17 (D)** shows the structure of FMK. Unlike SL010 and BI-D1870, Pyrrolopyrimidine FMK is an irreversible inhibitor of RSKs and binds covalently to CTKD [125, 126].

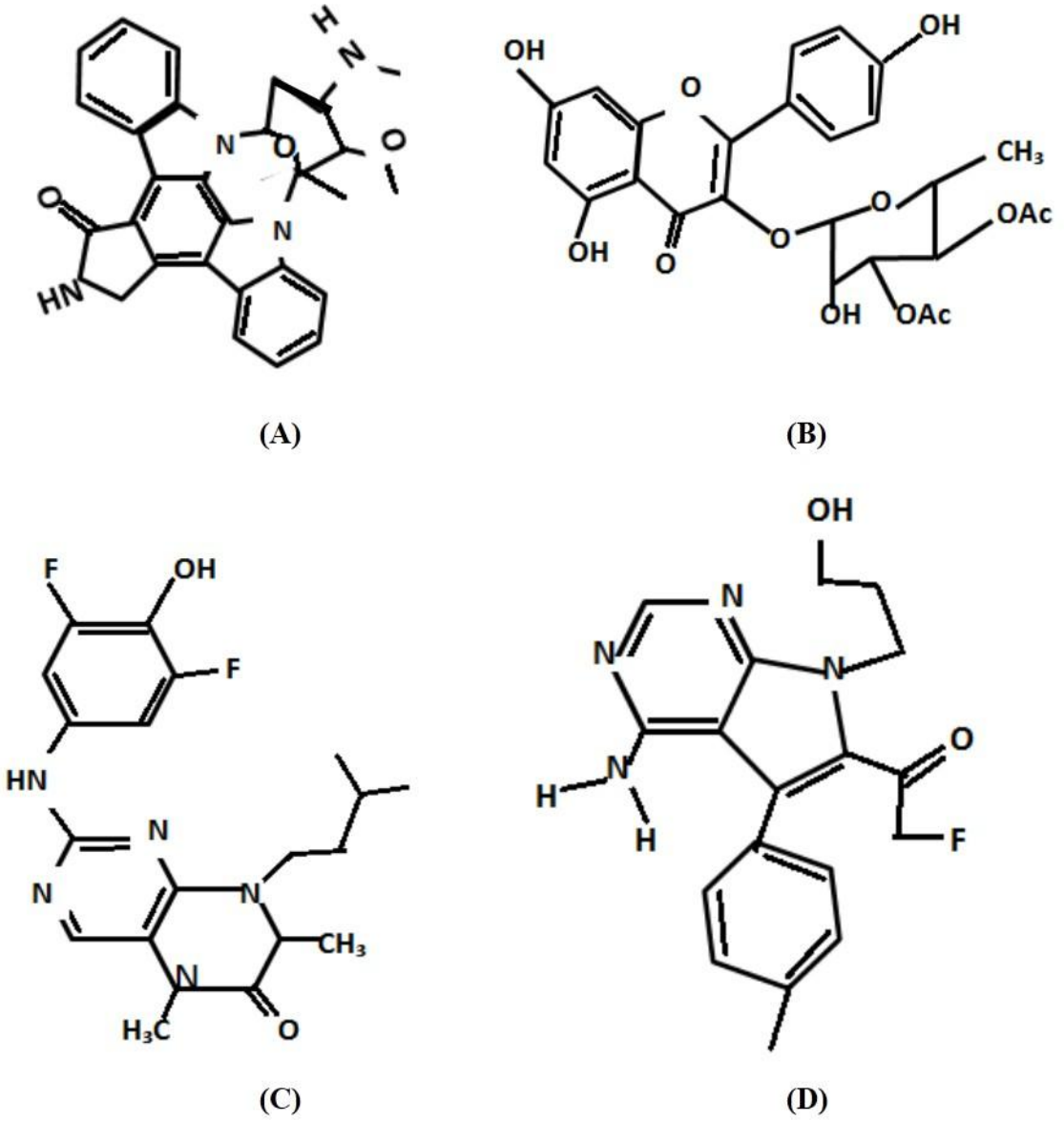


Figure 1.17: RSK inhibitors. (A) Staurosporine (B) SL0101 (C) BI-D1870 and (D) FMK [119].

1.4.1.2.7. Substrate recognition by RSKs

RSKs have many substrates both in the nucleus and in the cytoplasm which are either activated or inhibited by phosphorylation. All the RSK substrates have one of the following unique motifs, they are either Arg/Lys-Xaa-Arg-Xaa-Xaa-pSer/Thr or Arg-Arg-Xaa-pSer/Thr, where Xaa is any amino acid [127, 128].

1.5. Conclusion

Cancer is a life threatening disease, and is one of the leading causes of mortality across the world. In developing countries, the incidence rates of cancer are increasing with time, and it is also a major burden for Indian population. Cancerous cells abrogate cellular pathways which are required for the normal functioning of the cell, and they also show resistance to cell death. If unchecked, they can metastasize other normal tissues and organs far away from the site of cancer origin ultimately leading to death. In all known human cancers, many cellular pathways are highly unregulated and among them MAPK pathway is widely studied as almost all of its members are involved in cancerous cell survival and progression. Hence MAPK pathway members are currently being targeted for developing new cancer therapeutics.

ERKs and RSKs constitute the important members of MAPK pathway, and transduce the extra cellular signal into the nucleus of the cell. The molecular mass of ERK2 is around 42 kDa and the full length structure of ERK2 revealed that it has two lobes and the substrate binding groove is in between the lobes. Activation of ERK2 requires dual phosphorylation at Thr185 and Tyr187 residues, and is mediated by MEK. Activated ERK2 phosphorylates and activates RSKs.

RSKs are very large complex kinases having two functional kinase domains on a single polypeptide chain. Since they are the terminal mediators of the MAPK pathway, their activation is highly robust and involves multiple phosphorylations at various sites within the protein. Though structures of no full length protein are available for any of the RSK isoforms, the individual kinase domain structures have been determined. The structure of RSK2 NTKD has been determined as a complex with AMP-PNP [102]. The structure of RSK CTKD revealed the αL helix occupies the substrate binding grove thereby auto-inhibiting the enzyme.

Chapter 2:

Materials and methods

2.1. Sources for reagents and instruments

1. Bacterial strains: DH5α, BL21 (DE3), BL21 (DE3) pLysE and Rosetta 2 (DE3)

2. Plasmids: pGEX-KT, pMal-c2X, pRSET-A and pET-28 a (+)

3. Bacterial culture and chemicals: LB medium (Himedia, India, Cat.# M575), 90 mm (diameter) x 15 mm petri plates (Himedia, India, Cat.# PW008), Ampicillin (Himedia, India, Cat.# CMS645), Kanamycin (Himedia, India, Cat.# TC136), chloramphenicol (Himedia, India, Cat.# TC204) and IPTG (MP biomedicals, USA, Cat.# 102101)

4. Insect cell line culture: Drosophila Schneider s2 insect cell lines, Cellfectin II Reagent (Invitrogen, USA, Cat.# 10362-100), 6-well plate (Nunc, Denmark Cat.# 140675), 75 cm^2 tissue culture flasks (Corning, USA Cat.# CLS3290) and Schneider's Drosophila Medium (Thermo Fisher, USA Cat.# 21720-024)

5. Protein biochemistry: ATP (Himedia, India, Cat.# TC085), Protease inhibitor cocktail (Sigma, USA Cat.# P2714), snake skin dialysis bag (Thermo Fisher, USA, Cat.# 68035) and Trypsin (Sigma, USA Cat.# T7575)

6. Chemicals, salts, buffers, precipitants, detergents and organic solvents: Sigma (USA), Merck (Germany), Fluka (Germany), Himedia (India) and SRL (India)

7. DNA isolation kits: Mini-prep kit (Qiagen, Germany, Cat.# 27106), Maxi prep kit (Qiagen, Germany, Cat.# 12163), PCR extraction kit (Qiagen, Germany, Cat.# 28106) and Gel extraction kit (Qiagen, Germany, Cat.# 28706)

8. Affinity resins: Ni-NTA (Qiagen, Germany, Cat.# 32169), GST (Novagen, Germany Cat.#70541) and Amylose (New England Biolabs, USA, Cat.# E8021L)

9. Pre-packed gel filtration columns Superdex-200 and 75 GE Healthcare (Sweden)

10. Crystallization screens: Hampton research (USA), Qiagen (Germany)

11. Synthetic peptide: USV Biotech (India)

12. Centricons: Millipore (USA)

13. DNA electrophoresis: Agarose (Himedia, India, Cat.# MB002), Ethidium Bromide (Thermo Fisher USA, Cat.# 5585011), Bromophenol blue (Sigma-aldrich, USA, Cat.# B0126), Xylene cyanol (Sigma-aldrich, USA, Cat.# X4126) Transilluminator (UVP bioimaging systems, Canada) and Nanodrop (Thermo Fisher, USA)

14. PAGE: Acrylamide (G biosciences, USA, Cat.# RC-002), Bis-acrylamide (Sigma-aldrich, USA, Cat.# M2022), Bradford reagent (Sigma-aldrich, USA, Cat.# B6916), Prestained protein marker (Thermo Fisher, USA, Cat.# 26616), SDA-PAGE apparatus (Bio-Rad, USA).

15. All Restriction enzymes and Dpn1 Fermentas (USA), Phusion polymerase (Finnzymes, Finland, Cat.# M0530S), Quick ligase (NEB, USA, Cat.# M2200S)

16. FPLC ACTA purifier (GE healthcare, Sweden)

17. Vibration free 22 °C incubator (Sanyo, Japan)

18. Stereo microscope (Olympus, Japan)

19. -20 °C incubator (Sanyo, Japan) and -80 °C deep freezer (Thermo Fisher, USA)

20. Avanti J-26XP centrifuge (Beckman Coulter, USA), Heraeus multifuge X3 Centrifuge (Thermo scientific, USA) and Eppendorf centrifuge 5415R (Sigma-aldrich, USA)

21. DNA sequencer (Applied Biosystems, USA)

22. ITC-200 (GE Healthcare, Sweden)

23. DynaPro NanoStar DLS (Wyatt Technology, USA)

24. CD spectrophotometer (JASCO, Japan)

25. Fluorescence spectrometer (Horiba, Japan)

26. Ultraflex II MALDI-TOF/TOF mass spectrometer (Bruker Daltonics)

2.2. Materials

2.2.1. Luria Bertani (LB) medium

LB medium was used both in liquid broth as well as solid agar to grow different strains of *Escherichia coli* such as DH5α, BL21 (DE3), BL21 (DE3) pLysE and Rosetta 2 (DE3). For making LB broth, 20 g of powdered LB broth was dissolved in 1000 ml distilled water. Similarly for LB agar, 3.5 g of LB agar was added to 100 ml distilled water. Both LB broth and agar were sterilized by autoclaving them at 121 °C and 15 psi for 20 minutes. Following autoclaving when the temperature reached around 48 °C, the LB agar was added with appropriate antibiotics and poured on 90 mm x 15 mm petri plates.

2.2.2. Stock solutions

For bacterial culturing antibiotics such as ampicillin, kanamycin and chloramphenicol were used appropriately depending on the bacterial strain and the type of plasmid. Ampicillin was used for pRSET-A and pGEX-KT vectors whereas kanamycin was used for pET-28 a (+) vector. Additionally chloramphenicol was added for bacterial strains BL21 (DE3) pLysE and Rosetta 2 (DE3). The antibiotic stocks used for bacterial culturing are 100 mg/ml for ampicillin, 50 mg/ml for kanamycin and 34 mg/ml for chloramphenicol. Ampicillin and kanamycin were prepared by dissolving the respective antibiotics in distilled autoclaved water whereas chloramphenicol was prepared by dissolving in molecular biology grade alcohol (Merck, Germany, Cat.# 108543). All the bacterial antibiotics were used in 1000 times dilution during culturing. For inducing the bacterial culture a stock of 1 M IPTG was used.

For Drosophila Schneider s2 insect cell lines culturing, Blasticidin (Sigma-aldrich, USA, Cat.# 15205) antibiotic was used.

2.2.2.1. Agarose gel electrophoresis

1x running TAE (Tris, Acetic acid and EDTA) buffer: 40 mM Tris, 20 mM glacial acetic acid, and 1 mM EDTA

6x dye loading dye: 0.25 % Bromophenol blue, 0.25% xylene cyanol and 30 % glycerol.

2.2.2.2. SDS PAGE

1x running buffer: 25 mM Tris, 192 mM glycine and 0.2 % SDS

5x loading dye: 125 mM Tris (pH 6.8), 50 % glycerol, 4 % SDS, 0.25 % Bromophenol blue and 1.42 M BME (β-mercaptoethanol)

Destaining solution: 45 % Methanol and 10 % acetic acid

Staining solution: 0.5 % coomassie brilliant blue in destaining solution

2.2.2.3. Western blotting

Transfer buffer: 0.19 M Glycine, 25 mM Tris, 0.01 % SDS and 20 % metahanol

Ponceau stain: 0.5 % w/v Ponceau S in 1 % v/v acetic acid

TBS: 100 mM Tris-Cl, pH 7.5 and 0.9 % w/v NaCl

TBST: 0.1 % v/v Tween 20 in TBS

Blocking buffer: 5 % BSA in TBST

2.3. Methods

2.3.1. Molecular Cloning

Molecular cloning is also known as Recombinant DNA technology in which a target gene is cloned into a desired vector, thus making a chimeric or recombinant DNA. In the first step the target gene is amplified by Polymerase Chain Reaction (PCR) and the PCR product is recovered. The recipient vector and the PCR product are subjected to double restriction digestion, with same set of restriction enzymes independently in two separate reactions. These doubly digested recipient vector and PCR product are ligated using Ligase enzyme. **Figure 2.1** shows the graphical representation of molecular cloning. It includes many steps such as Polymerase Chain Reaction (PCR), PCR/Gel extraction, restriction digestion, ligation, transformation and DNA extraction (mini-prep or maxi-prep). The target gene insertion into the host vector is determined by gel shift assay, gene pop-out assay and by DNA sequencing.

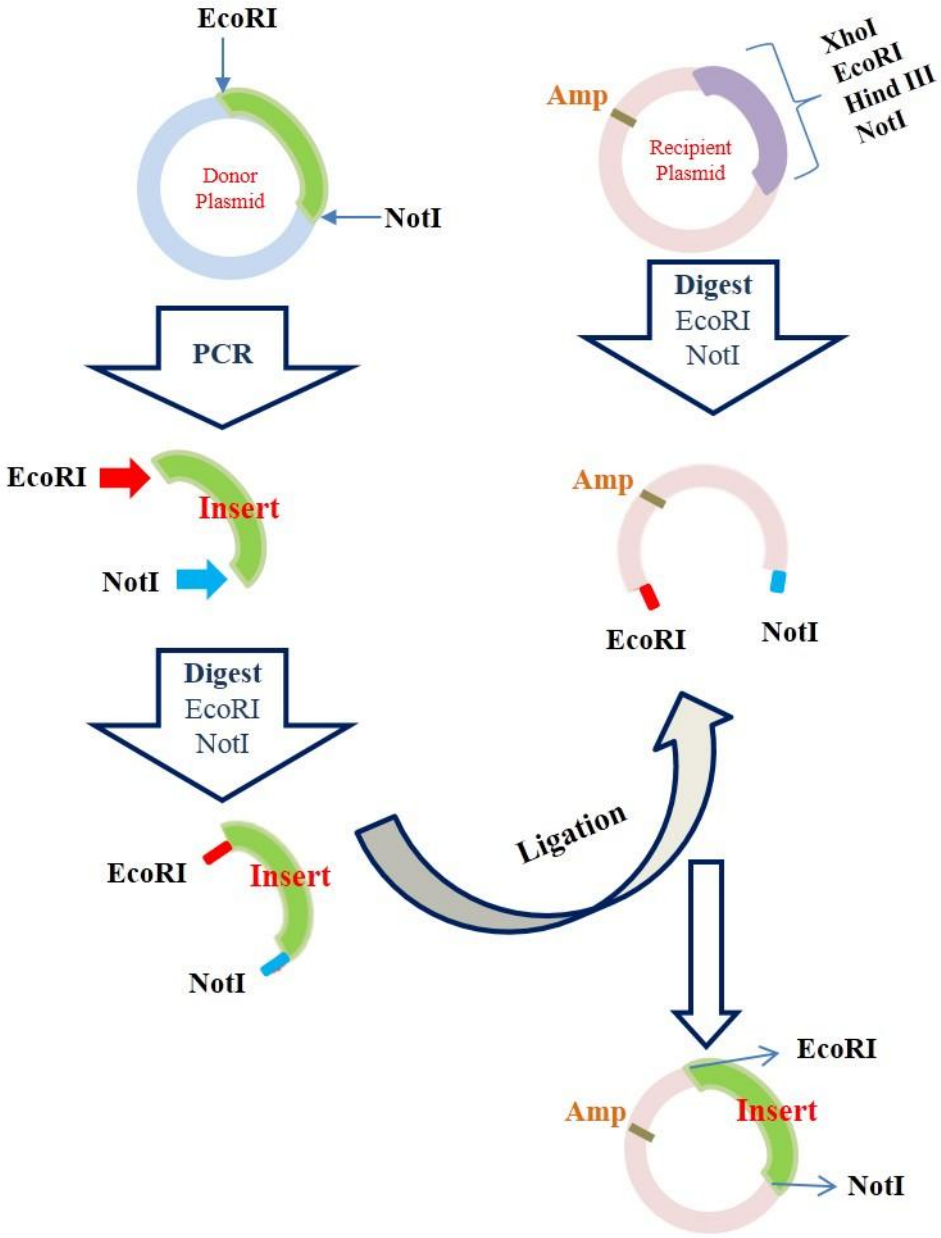


Figure 2.1: A graphical representation of molecular cloning

2.3.1.1. Polymerase Chain Reaction (PCR)

PCR, developed by Dr. Kary Mullis during early 1980's, is a process in which a piece of DNA can be amplified using specific primers complementary to a region of the DNA [129-131]. In 1993, he won the Nobel Prize in Chemistry. A typical PCR contains a thermally stable DNA polymerase, dNTPs, primers (forward and reverse), template DNA and polymerase buffer which generally contains salts and cofactors required for polymerization

reaction [132]. There are three major steps in the PCR reaction, 1) denaturation, 2) annealing and 3) elongation. The denaturation step is required to unwind the parental DNA and is performed at ~ 95 °C. It is followed by a annealing step and is usually between 65 to 68 °C [133]. The elongation is performed at 72 °C and duration of this step also depends on the length of the amplicon. **Figure 2.2** shows a general PCR reaction involving all the steps. **Table 2.1** shows the PCR experimental set up which was used to amplify RSK functional domains. **Table 2.2** shows a 40 µl PCR mix. Following PCR amplification, the amplicon is analysed by agarose gel electrophoresis.

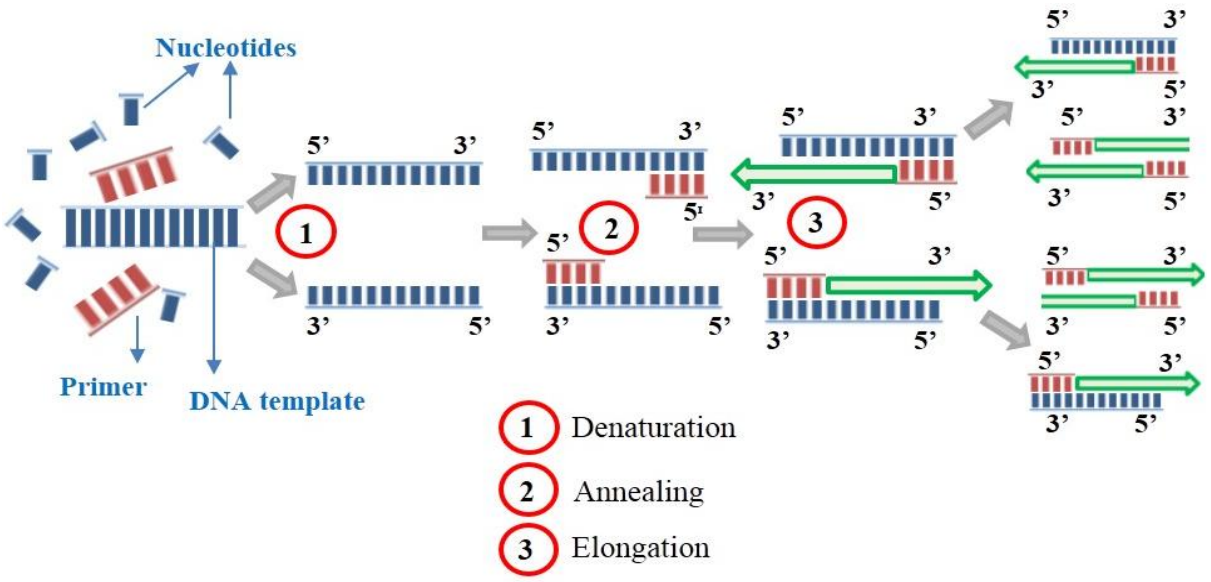


Figure 2.2: A graphical representation of PCR indicating its important steps

Table 2.1: PCR experimental conditions

Experimental parameters		
1.	95 ^{0}C	5 minutes
2.	95 ^{0}C	30 seconds
3.	65 ^{0}C	30 seconds
4.	72 ^{0}C	1 minute
5. Go to step 2 for 32 times		
6.	72 ^{0}C	10 minutes
7. End		

Table 2.2: A 40 µl PCR mixture constituents

PCR mix	
Template DNA (100 ng/µl)	1 µl
Forward primer (10 pmole/µl)	1 µl
Reverse primer (10 pmole/µl)	1 µl
dNTPs (10 mM)	2 µl
5X HF buffer	8 µl
Phusion polymerase	1 µl
Distilled water	26 µl

2.3.1.2. Agarose gel electrophoresis

Agarose gel electrophoresis is the most common molecular biology technique to view and analyse DNA samples [134-137]. The basic principle behind this method is that DNA possesses a net negative charge, and thus upon providing electricity (5 volts/cm), the DNA molecules tend to move to positive electrode (anode). Generally 0.5 to 2 % agarose gels are used and the percentage of agarose is inversely proportional to the size of the DNA being analysed. Agarose is thermally stable molecule and melts only at temperatures higher than 92 °C. After dissolving the agarose at higher temperature in TAE buffer, the mixture is allowed to cool until it reaches around 50 °C and Ethidium bromide (EtBr) is added so that the final concentration is 0.5 µg/ml, and is poured on a casting tray. EtBr is a fluorescent molecule used to stain DNA molecules. It fluoresces bright orange colour when exposed to UV light [138]. The cast agarose gel is gently placed inside the electrophoresis apparatus and enough 1X TAE buffer is added so that the entire gel is completely submerged. The DNA samples prior to loading on to agarose gels, are first mixed with 6x loading buffer. The loading buffer contains synthetic dyes such as bromophenol blue and xylene cyanol which will helps in monitoring the DNA migration through gel pores.

2.3.1.3. Extraction of the PCR product

After PCR, 1 to 2 µl of the PCR product is loaded on to the agarose gel. If the band is single, the amplified PCR product is recovered by PCR extraction kit (Qiagen, Germany, cat.# 28104). If there is any non-specific additional band apart from the band of interest, the entire PCR product is loaded on to gel and the target band is excised by viewing under transilluminator. The DNA from the excised band is recovered by using gel extraction kit (Qiagen, Germany, cat.# 28704).

2.3.1.3.1. Extraction using PCR extraction kit

Following confirmation of a single band of target gene on agarose gel electrophoresis, the PCR product can be extracted using PCR extraction kit. PCR products for all the functional domains of RSK had been extracted by using this method. The steps involved in this method are as follows.

1. To one volume of PCR product, add five volumes of PB buffer.
2. Load this mixture on to the DNA column and allow DNA to bind for at least 5 minutes.
3. Following incubation, spin the column at 10,000 rpm for one minute at room temperature.
4. After discarding flow through, add 750 µl of PE buffer and centrifuge at 13,000 rpm for one minute at room temperature.
5. An additional empty spin is given to remove any leftover traces of PE Buffer.
6. The column is then placed in fresh autoclaved 1.7 ml eppendorf (Axygen, USA, cat.# MCT-175-C). 10 to 30 µl of elution buffer is added to column and allowed to incubate for at least 5 minutes.

7. Following incubation, the column was centrifuged at 13,000 rpm for one minute at room temperature.

The eluted DNA concentration was measured by Nano drop and the samples were kept at -20 °C for further use.

2.3.1.3.2. Extraction using gel extraction kit

RSK1 full length PCR product was extracted using Qiagen gel extraction kit. Following gel electrophoresis, the target PCR product has to be carefully excised under UV transilluminator with protective eye and hand care. The steps involved in this method are as follows.

1. The excised DNA band is weighed and transferred into a clean 1.7 ml eppendorf. Add three volumes of QG buffer.
2. Incubate this mixture at 55 °C till the gel becomes completely dissolved in the QG buffer.
3. Once the agarose is dissolved, allow the mixture to cool down for few minutes and add one gel volume of Isopropanol.
4. This mixture is then loaded on to the spin column and allowed to bind. Rest of the steps are the same as mentioned from step 2 in **2.3.1.3.1.** section.

2.3.1.4. Restriction digestion of DNA

Restriction enzymes are unique bacterial endonucleases which recognize foreign DNA and neutralize them by digesting the phosphodiester bond [139, 140]. Type II restriction endonucleases are widely used in recombinant DNA technology as they

specifically recognize a unique specific DNA sequence, knows as restriction site, which is a palindromic sequence of 6-8 nucleotides [141, 142]. Type II restriction endonucleases such as EcoR1, BamH1, Xho1 and Sal1 had been used for cloning in this study.

A total of 10 µg PCR product was subjected to double restriction digestion. The reaction mixture contains 10X FD buffer (Fermentas) and two restriction enzymes each corresponding to the site within the forward and reverse primers. The restriction digestion is performed at 37 °C for two hours in a water bath. Following two hours digestion, the digested PCR product is extracted as mentioned in **2.3.1.3.1.** section. A 30 µl double restriction reaction, with the template DNA concentration of 500 ng/µl is shown in **Table 2.3**.

Table 2.3: A 30 µl double restriction digestion reaction constituents

Double restriction digestion	
Template DNA (500 ng/µl)	20 µl
Restriction enzyme 1	1 µl
Restriction enzyme 2	1 µl
10X FD buffer	4 µl
Distilled water	4 µl

2.3.1.5. Selection of backbone plasmid

Selection of plasmid vector is very important as it plays an important role in recombinant protein purification. The plasmid vectors used in recombinant protein purification are generally equipped with affinity tags such as 6-HIS, GST and MBP. 6-HIS based vectors like pRSET-A and pET-28 a (+), GST based pGEX-KT and MBP based pMal-c2X vectors had been used in this study.

Plasmids are covalently circularly bound double stranded DNA, which occur as extra-chromosomal genetic material [143]. A typical bacterial plasmid should contain these

following features: an origin of replication, antibiotic resistant gene, promoter and Multiple Cloning Sites (MCS). Bacterial expression plasmid vectors additionally contain DNA region which codes for a specific affinity tag. **Figure 2.3** shows backbone vector map of pET-28 a (+) which has kanamycin resistant gene.

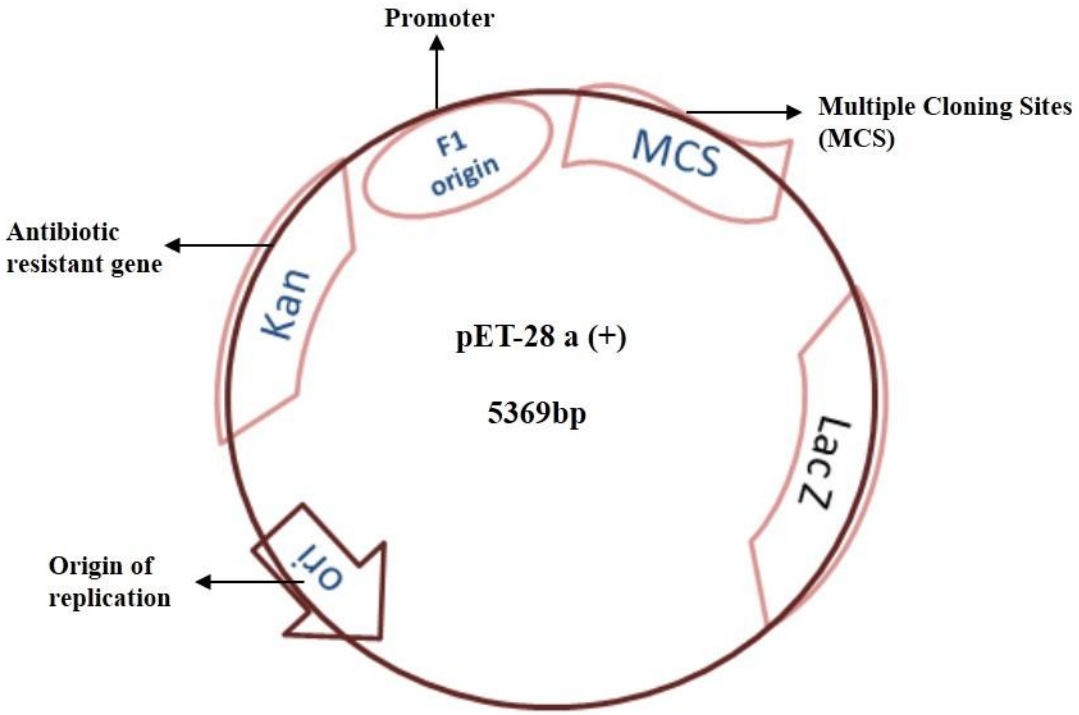


Figure 2.3: pET-28 a (+) vector backbone

After selecting the desired plasmid, the plasmid vector is also subjected to double restriction digestion with the same enzymes as mentioned in **2.3.1.4**.

2.3.1.6. Ligation of DNA

Ligation is a process in which the two DNA molecules are covalently joined by the formation of a phosphodiester bond between 3'-OH of DNA molecule and 5'-PO_3 group of another DNA molecule [144]. T4 DNA ligase is used for DNA ligation *in-vitro* experiments for making chimeric DNA [145]. One of the DNA molecule is the double digested plasmid vector whereas the other DNA molecule is the double digested PCR product. The plasmid

vector and PCR product are generally taken in the molar ratio of 1:3, and a minimum of 50 ng of plasmid had to be taken. Quick ligation kit from NEB is used for all ligation reactions used in this study. All the reactions were performed with a 20 µl total reaction volume and incubated at room temperature for 15 minutes.

2.3.1.7. Transformation of ligated DNA

Transformation is a typical phenomenon, in which bacterial cells tend to take in exogenous DNA from the surroundings through the plasma membrane [146, 147]. It was initially discovered by Frederick Griffith in 1928 and named the factor responsible for transformation as 'transforming principle'. The identity of transforming principle was proved to be a DNA molecule by Avery, MacLeod and McCarty in 1944. The DNA uptake is possible during a brief period of bacterial life cycle known as competence stage, which occurs between lag phase and log phase [148]. Transformation can be achieved *in-vitro* by providing heat shock to the bacterial cells. The principle behind DNA uptake is that during heat shock, the high temperature creates pores within the plasma membrane which enables the entry of plasmid in to the bacterial cells. Further, the $CaCl_2$ present in the medium helps in depolarizing the negative charge on the inner face of plasma membrane thus favouring the entry of negatively charged DNA into the cell. Once the heat shock is removed the cell membrane pores are resealed and the membrane potential is restored back to normalcy [149]. The steps involved in bacterial transformation by heat shock method is as follows.

1. Take the DH5α competent cells from -80 °C and thaw it on ice.
2. Add 100 ng DNA to the DH5α competent cells and incubate on ice for 45 minutes.
3. Following incubation, heat shock is given in water bath at 42 °C for 2 minutes.
4. After heat shock, the competent cells were incubated on ice for 5 minutes.

5. Following 5 minutes incubation, ~ 800 µl of autoclaved LB medium is added and further incubated at 37 °C for 45 minutes in a shaker incubator.
6. After 45 minutes, the bacterial cell culture was centrifuged at 4000 rpm.
7. The spent medium is removed leaving 100 µl of medium on the pellet for re-suspension.
8. The re-suspended culture is poured and plated uniformly all over onto a LB agar plate containing the appropriate antibiotic.
9. The plate is labelled appropriately and incubated at 37 °C for 14 hours in an incubator.

The transformed bacterial cells will form individual colonies following overnight incubation as the non-transformed cells are selectively killed by the antibiotic. To extract the DNA, a single isolated colony is picked and inoculated in 10 ml LB broth containing appropriate antibiotic. The LB broth is then incubated at 37 °C for 10 hours in shaker incubator. This culture is further used for isolating recombinant plasmid DNA by Mini-prep or Maxi-prep.

2.3.1.8. Plasmid DNA isolation by Mini-perp

The Plasmid isolation was done by using commercially available kits from Qiagen. The steps involved in plasmid isolation by Mini-prep are as follows.

1. Centrifuge the overnight grown 10 ml bacterial culture at 6000 rpm, 4 °C for 10 minutes.
2. Re-suspend the bacterial pellet gently in 250 µl of P1 buffer additionally containing RNase A (100 µg/ml).

3. Transfer the lysate into a fresh autoclaved 1.7 ml eppendorf and add 250 µl of P2 buffer and mix it by gently inverting the eppendorf 6-8 times. This is the lysis step and should not go beyond 5 minutes.
4. After lysis, add 350 µl of N3 buffer and mix it by gently inverting the eppendorf 6-8 times.
5. Centrifuge the eppendorf at 13000 rpm, room temperature for 10 minutes.
6. The clear supernatant is added on to the column and allow few minutes for DNA binding
7. Centrifuge the column at 10000 rpm, room temperature for 1 minute.
8. Discard the flow through and add 500 µl of PB buffer and centrifuge at 13000 rpm, room temperature for 1 minute.
9. Discard the flow through and add 750 µl of PE buffer and centrifuge at 13000 rpm, room temperature for 1 minute.
10. An additional empty spin is given at 13000 rpm, room temperature for 1 minute to remove any leftover traces of PE Buffer.
11. The column is then placed in freshly autoclaved 1.7 ml eppendorf. 10 to 30 µl of elution buffer (EB buffer) is added to column and allowed to incubate for at least 5 minutes.
12. Following incubation, the column is centrifuged at 13,000 rpm, room temperature for 1 minute.

The eluted DNA concentration was measured by Nano drop and the samples were kept in -20 °C for further use.

2.3.1.9. Confirmation of Molecular cloning

Following Mini-prep, the extracted DNA had to be verified weather the gene of interest had been ligated or not. To check this, load ~ 200 ng of DNA on 1 % agarose gel along with uncut empty vector. If the ligation is successful, there will be a shift of band and the band corresponding to ligated sample will be slightly above the band corresponding to empty vector. However, this method is useful if the insert size is ≥ 100 bp long. Smaller the insert, it would be difficult to see the possible shift on agarose gel. After analysing the gel shift assay, the potential clones will be subjected to double restriction digestion with the restriction enzymes. The ligated clones will pop-out the gene which can be detected by running an agarose gel. After confirming the insert release from the gene pop-out assay, the recombinant constructs will be subjected to DNA sequencing to check for occurrence of if any mutations or frame shifts in the recombinant constructs. DNA sequencing by di-deoxy Nucleoside Triphosphates (ddNTPs) chain terminators was initially developed by Frederick Sanger [150, 151]. Currently the ddNTPS are covalently labelled with four different fluorescent dyes each specific for A, T, G and C [152].

For DNA sequencing, 2 µl of 150 ng/µl plasmid DNA with 1 µl of 10 pmole/µl sequencing primer were used. The data obtained after DNA sequencing was analysed using BioEdit software.

2.3.2. Site-directed insertion

Site-directed insertion is a targeted nucleotide insertion in template DNA which is similar to the Site-Directed Mutagenesis (SDM) [153-155]. This method is useful in selectively mutating, inserting and deleting nucleotides at desired location within the template DNA and is widely used in studying mutational effects on proteins. For site-directed

insertion, a maximum of 21 nucleotides can be inserted in a single reaction. The primers are designed in such a way that the sequence of DNA to be incorporated is flanked by a segment of template DNA of at least 15-18 nucleotides both at 5' and 3' from the site of insertion. This will help in binding of the primer to the template DNA. During PCR, apart from whole plasmid duplication, the incorporating sequence will also be duplicated at the targeted site. The PCR is set similarly to the one mentioned in **Table 2.1.**, however the elongation step is extended to 3-5 minutes depending on the size of template plasmid. Also the number of cycles had to be between 18-20, otherwise the dNTPs will be exhausted during the reaction. The PCR mix is same as the one mentioned in **Table 2.2**.

Following PCR, the reaction mix will contain the methylated parental template and non-methylated duplicated plasmid with site directed insertion. To remove the methylated plasmid (*wild type*), the PCR amplified sample will be subjected to Dpn1 digestion at 37 °C for 20 minutes. Following digestion, the reaction mixture is added to DH5α competent cells and transformation was done as mentioned in **2.3.1.7.** section. Further mini-prep was done to isolate the plasmid as mentioned in **2.3.1.8.** section and inserted sequence were verified by DNA sequencing as mentioned in **2.3.1.9.** section. Site directed insertion was used to incorporate 18 nucleotides to generate an additional 6 HIS tag at the C-terminal end of RSK1 full length protein.

2.3.3. Transfection

Transfection is the phenomenon of uptake of naked DNA by the Eukaryotic cells from the surrounding media, whereas transformation is a bacterial phenomenon. Drosophila Schneider s2 cell lines are widely used insect cell line expression system for recombinant protein purification [156]. This expression system is advantageous over bacterial system in

retaining the critical Post-Translation Modifications (PTMs) which occur in eukaryotic proteins. The steps involved in transfection are as follows.

1. Drosophila Schneider S2 cells (10^6 cells/ml) are poured onto a 6 well plate (2 ml/well) and incubated overnight at 25 °C.
2. A DNA and transfecting agent mix is prepared by adding 2 µg DNA and 5 µl of Cellfectin to total 100 µl of serum free media. Incubate the DNA mix for 30 minutes at room temperature.
3. After 30 minutes incubation, 900 µl of serum free media is added and mixed gently.
4. This transfecting mix is poured gently on to 2 ml of Drosophila Schneider S2 cells in the 6 well plate.
5. After 6 hours, the media is aspirated and fresh complete media containing 10 % FBS is added along with blasticidin antibiotic and incubated at 25 °C for 48 hours.
6. For establishing stable cell lines after every 48 hours the spent medium is replaced with fresh complete media with 10 % FBS and with appropriate antibiotic. In about 10 such cycles (3 weeks), stable cell lines will be established.

RSK functional domains had been cloned into pMT-wb-v5 vector and transfected into Drosophila Schneider s2 insect cell lines and stable cell lines were established subsequently.

2.3.4. Culturing and expression of recombinant protein

2.3.4.1. Bacterial culturing and expression

Bacterial strains like DH5α were used to amplify the plasmid. However for over expression of recombinant proteins bacterial expression strains like BL21 (DE3), BL21 (DE3) pLysE and Rosetta 2 (DE3) were used. These bacterial strains have T7 polymerase which is under the control of Lac operon [157]. All these strains are deletion mutants of

proteases such as Lon and Outer membrane protease (OmpT). BL21 (DE3) pLysE and Rosetta 2 (DE3) strains additionally contain chloramphenicol resistant gene.

2.3.4.1.1. Starter culture

The recombinant plasmid is transformed into one of the bacterial expression strains with appropriate antibiotics as mentioned in **2.3.1.7.** section. A single isolateed colony is inoculated into 10 ml LB broth with appropriate antibiotic. The culture is incubated at 37 °C for 6-8 hours in shaker incubator. This is known as starter culture and is used to inoculate larger cultures for protein purification.

2.3.4.1.2. Dilution and induction

Following 6-8 hours, the bacterial starter culture is then transferred into 1 litre fresh autoclaved LB broth containing appropriate antibiotic and allowed to grow at 37 °C, 200 rpm till OD_{600} reached 0.5-0.6. Once the desirable OD_{600} is reached, the bacterial culture is kept at 4 °C for 20 minutes. Following this, add 0.3-0.4 mM isopropyl-β-D-thiogalactoside (IPTG) and allow the culture to grow at 20-22 °C, 200 rpm in a shaker incubator for 14-16 hours. Induction by IPTG helps in over-expression of target recombinant protein, as it is under the control of Lac operon.

2.3.4.1.3. Harvesting of bacterial culture

Following overnight induction, the bacterial culture is harvested by centrifuging at 6000 rpm, 4 °C for 10 minutes. The supernatant is decanted and the bacterial pellet is stored in -80 °C for further use.

2.3.4.2 Insect cell line culturing and expression

The transfected stable insect cell lines are grown in 75 cm^2 tissue culture flasks at 25 °C in presence of blasticidin. The target protein is under the control of Metallothionein (MT) promoter, which is induced by 0.5 mM cupric sulphate. Following induction, the insect cells are harvested by centrifuging at 5000 rpm, 4 °C for 10 minutes. The supernatant is discarded and the cell pellet is used for protein purification or can be stored at -80 °C till required.

2.3.5. Recombinant protein purification

There are several methods to purify proteins such as affinity chromatography, Size Exclusion chromatography (SEC), Ion Exchange chromatography (IEC) and Hydrophobic Interaction Chromatography (HIC).

2.3.5.1. Affinity chromatography

Affinity chromatography is the most preferred first choice method for the purification of recombinant proteins. The principle behind this method is based on the affinity between two molecules [158]. One of the interacting molecule known as bait is entrapped in a stationary phase whereas its interacting molecule known as prey is passed over it through mobile phase. The interacting partners bind to resin and become the part of stationary phase. After significant washes, the prey molecule is recovered by competitive analogue of prey molecule. Affinity chromatography employs a tag which is the part of plasmid vector. There are many affinity tags available for affinity chromatography and tags like poly-Histidine, Glutathione S-Transferase (GST) and Maltose binding protein (MBP) are commonly used.

2.3.5.1.1. Poly-Histidine tag

Affinity chromatography by Poly-Histidine tag is one of the most commonly used techniques for protein purification which contains either 6 HIS or 8 HIS resides as affinity tag. However, 6 HIS tag is most commonly used in most of the commercially available plasmid vectors. The 6 HIS tagged recombinant proteins are generally purified by Ni-NTA resin, where Ni is Nickel metal ion and NTA is nitrilotriacetic acid [159, 160]. The NTA is covalently attached to matrix and is used as stationary phase during purification. The coordination sphere of Nickel has six binding sites, out of which four are occupied by NTA while remaining two are available for binding histidine residues within the tag of the recombinant protein [161]. **Figure 2.4** shows the coordination sphere over Nickel metal ion. Alternatively to Ni-NTA resin, Co-CMA resin is also used, where the co-ordinating metal ion is Cobalt and CMA is carboxylmethylaspartate [162]. Since a metal ion is also immobilized along with the matrix, this method is also known as Immobilized Metal ion Affinity Chromatography (IMAC) [163]. The bound recombinant protein can be eluted either by histidine or imidazole. ERK2, RSK1 full length and all the functional domains of RSKs were purified by 6-HIS tag in this work.

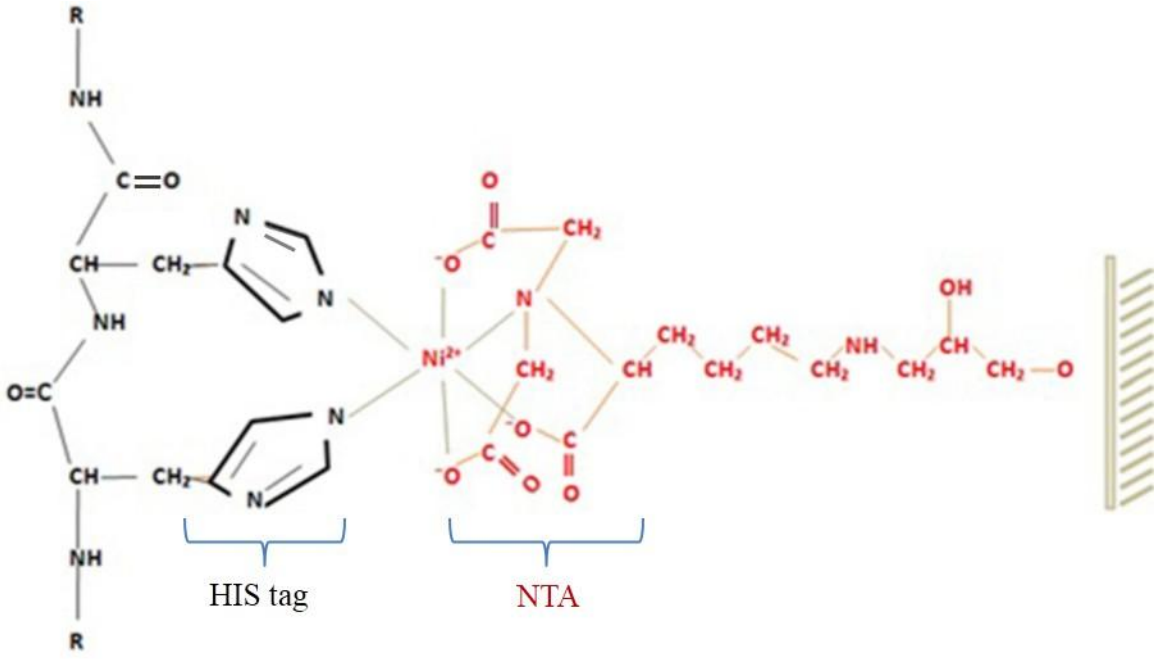


Figure 2.4: Coordination sphere around Ni metal ion formed between 6 HIS tag and Ni-NTA resin [34].

2.3.5.1.2. GST tag

Glutathione S-transferase (GST) tag is one of the widely used affinity tags to purify recombinant proteins. GST is a 26 kDa protein initially isolated from *Schistosoma japonicum* and used for protein purification as it has high affinity for reduced glutathione (GSH) [164, 165]. Following the binding of recombinant protein, significant buffer washes are given to remove non-specifically bound proteins. Since the tag is 26 kDa which is considerably large, the GST tag has to be removed, to get the protein in native form. For this, a protease cleavage sequence should be present between the GST sequence and target protein which is usually present in all the commercially available vectors or can be incorporated during cloning step. The most commonly used proteases to cleave the tags are Tobacco Etch Virus (TEV) protease, Thrombin and Factor Xa. Following protease cleavage, the target protein is eluted and the GST tag is still bound to GSH resin. The GSH resin can be regenerated by passing 20 mM reduced Glutathione (GSH), in 50 mM Tris, pH 8.0.

2.3.5.1.3. MBP tag

Maltose Binding Protein (MBP) is a highly soluble protein used in affinity chromatography, which also significantly increases the solubility of the protein it is tagged to [166-168]. MBP has high affinity for Amylose resin and can be eluted by 20 mM Maltose. The molecular weight of MBP is 43 kDa and is an even bigger tag than GST tag. Hence after purification, the MBP tag has to be cleaved. **Figure 2.5** shows the graphical representation of affinity chromatography using MBP tag.

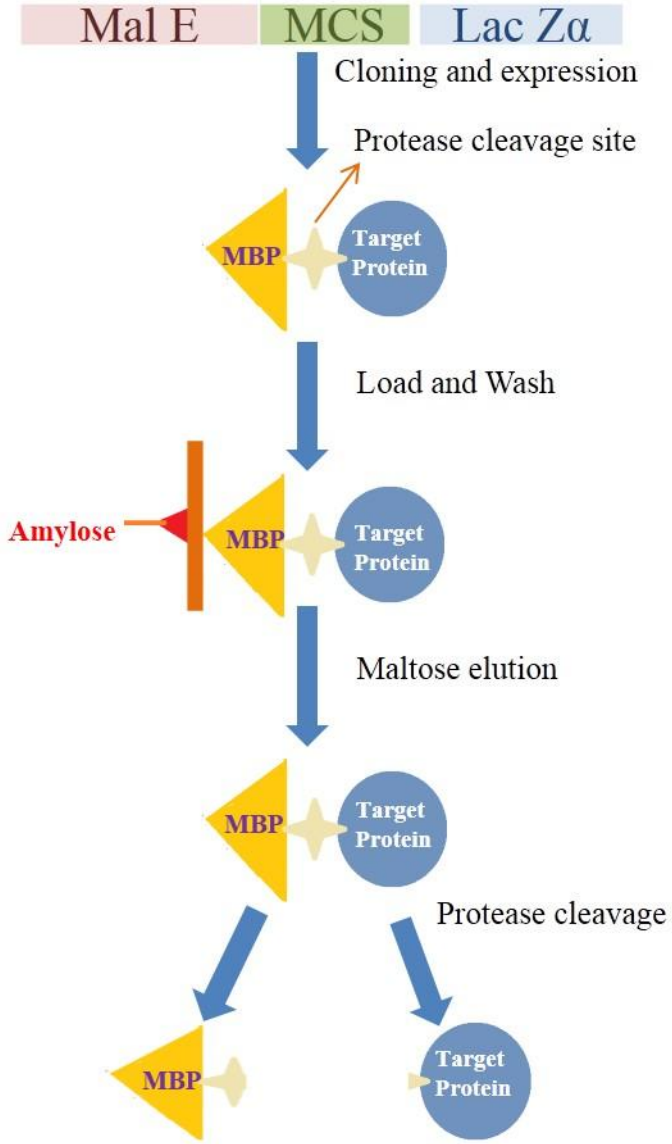


Figure 2.5: A graphical representation of purification of a recombinant protein with MBP tag.

2.3.5.2. Size Exclusion chromatography

Size Exclusion chromatography (SEC) is also known as Gel Permeation Chromatography (GPC). The principle here is, the separation of proteins is based on the size, in such a way that larger proteins elute first followed by smaller proteins [169, 170]. Superdex is a widely used matrix for performing SEC. The extent of cross linking of the dextran matrix determines the molecular exclusion limit of the matrix. Any molecule above this limit will not enter through the matrix, but will navigate through the interstitial spaces between the matrices and will be eluted in void volume, which is usually 1/3 of the column size. Superdex-200 and Superdex-75 columns are commercially available for SEC, whereas 200 and 75 indicates the molecular exclusion limits as 200 kDa and 75 kDa respectively. These columns are connected to Fast Protein Liquid Chromatography (FPLC) system, to increase the flow rates of elution [171]. The concentrated protein up to 2 ml (as per loop size) is injected into the injection loop. The buffer flow rate is maintained between 0.5-1 ml/minute. The FPLC system has integrated UV spectrophotometer which measures the absorbance at $\lambda = 280$ nm, which can be used to monitor the flow of protein eluting through the column. All the proteins purified by affinity chromatography in this work had been further purified to greater purity by SEC using either Superdex-200 or Superdex-75 columns.

2.3.5.3. Ion Exchange chromatography

Ion Exchange Chromatography (IEC) is used to separate proteins on the basis of net charge of the molecules. There are two types of IEC, one is Anion-exchange chromatography and Cation-exchange chromatography [172]. In Anion-exchange chromatography, the matrix contains positively charged groups and are used as anion exchangers and bind negatively charged proteins. In Cation-exchange chromatography, the

matrix is negatively charged and used to purify positively charged proteins. The commonly used functional groups in anion exchangers are Quaternary ammonium (strong) and DEAE (weak) whereas Sulfopropyl (strong) and Carboxymethyl (weak) are common cation exchangers. The bound protein is eluted from the matrix by passing higher concentrations of counter ions, usually 0.2 to 1 M NaCl.

2.3.5.4. Hydrophobic Interaction Chromatography

Protein purification by Hydrophobic Interaction Chromatography (HIC) is based on the hydrophobic interactions between the matrix and the non-polar groups on the surface of the protein molecule [173]. Butyl Sepharose resin is commonly used to purify proteins by HIC. The binding of target protein is performed in presence of high concentration of salting-out salts like ammonium sulphate, and elution is done by decreasing the concentration of salting-out salts in the elution buffer.

2.3.6. Confirmation of target protein

After protein purification, it is necessary to confirm the identity of the protein and should be done before further scaling up the protein purification. The nativity of protein can be confirmed by either trypsin in-gel digestion followed by Mass Spectrometry (MS) or Western blotting.

2.3.6.1. Mass Spectrometry

Mass Spectrometry is a method to quantify the amount and detect the type of biomolecules present in the given solution [174]. The biomolecules are converted to ions and

in presence of magnetic and electric fields, the individual ions behave differently with respect to their mass/charge ratio (m/z). A typical Mass spectrometer consists of three major parts:

1. Ion source: For charging of biomolecules to ionic species.
2. Mass analyser: For sorting out different ionic species under magnetic and electric fields.
3. Detector: For detecting the separated ions.

Matrix Assisted Laser Desorption Ionization (MALDI) is the most commonly used ionization method for ionising protein and peptide molecules [175]. In MALDI, the chemical matrix helps in ionization of biomolecules present in the solution upon bombardment with Laser. The most commonly used matrices are α-Cyano-4-HydroxyCinnamic Acid (CHCA) and sinapinic acid. To confirm the identity, the protein molecules are subjected to trypsin in-gel digestion before being analysed by Mass Spectrometry.

2.3.6.2. Western blotting

Western blotting is one of the most commonly used methods to detect the nativity of proteins, particularly when the concentrations of the target proteins is too low [176]. The first step involved in Western blotting is separation of proteins on the basis of their size by Poly-Acrylamide Gel Electrophoresis (PAGE). The separated protein is then laterally transferred either onto a nitrocellulose or Polyvinylidene fluoride (PVDF) membrane by passing electricity and this step is known as electro-blotting [177]. **Figure 2.6** shows the general setup for electro-blotting. The next step employs specific monoclonal (primary) antibodies against the target protein on the membrane. This immune blot is further incubated with secondary antibody which is specific for Fc (Fragment crystallisable) region of primary antibodies. The secondary antibodies are covalently linked with either alkaline phosphatase

or horseradish peroxidase, which helps in visualization of target protein by interacting with its substrates by a process known as chemiluminescence. The identity of RSK3 NTKD and RSK3 CTKD from insect cell lines were confirmed by Western blotting using Anti-HIS antibodies.

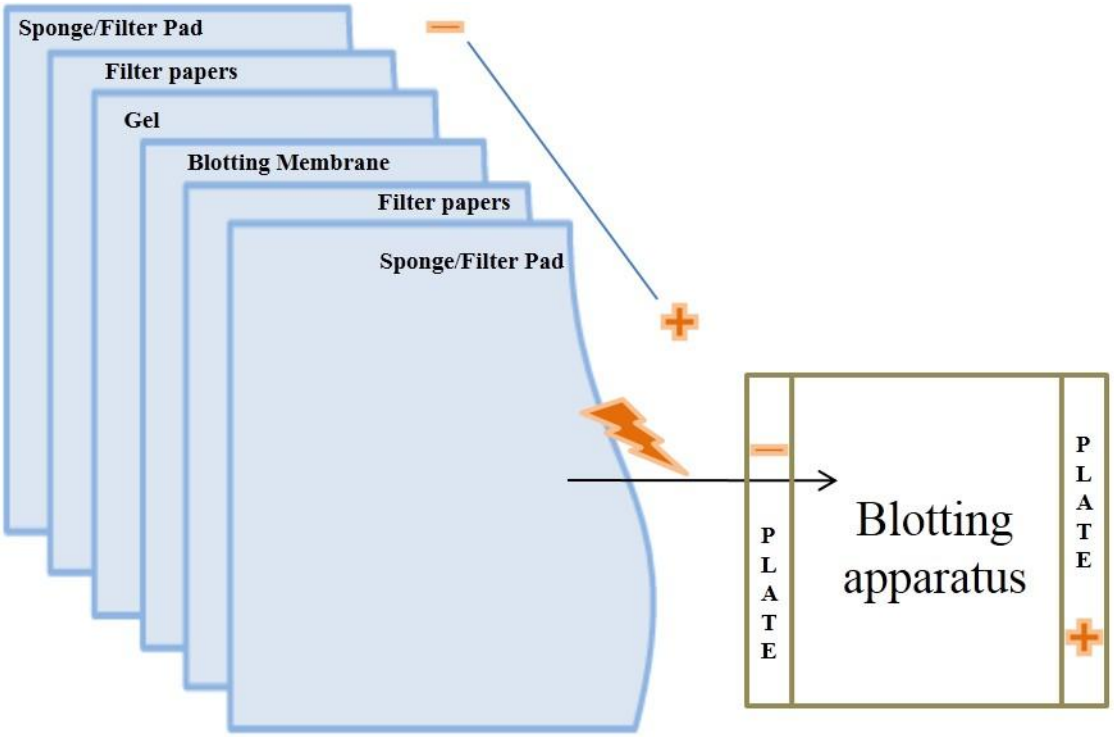


Figure 2.6: A general setup of layers for electro-blotting used in Western blotting.

2.3.7. Biophysical characterization of proteins

Biophysical characterisation is very essential for all purified proteins. For any functional and structural studies, it is necessary for the protein to be in its native secondary and tertiary conformation. There are various biophysical methods to characterize purified proteins such as Circular Dichroism (CD), Fluorescence spectroscopy, Thermal denaturation, limited proteolysis and Dynamic Light Scattering (DLS).

2.3.7.1. Circular-Dichroism

Circular-Dichroism (CD) is a very sensitive method used to characterize and estimate the secondary structure within the proteins [178, 179]. Proteins which are made of L-amino acids have the tendency to differentially absorb the right and left components of the circularly polarized light. CD employs circularly polarized light which is used to measure the differences in absorption of right polarized light with respect to left polarized light. In circularly polarized light, the electric component of the EMR rotates along the direction of propagation keeping the magnitude of the electric vector constant. In a CD spectrophotometer, the circularly polarized light is derived from a light source by passing it through a series of prisms, mirrors and slits. The circularly polarized light is then passed onto protein sample in quartz cuvette and the output signal is photo-multiplied and detected.

The differential absorption of right and left components of circularly polarized light is observed maximum in the far UV region of EMR between λ = 250-190 nm, which is mainly contributed by amide bonds. The CD spectra of proteins is characterized by two major electronic transitions which are $n \rightarrow \pi^*$ between λ = 230-215 nm and $\pi \rightarrow \pi^*$ between λ = 200-190 nm [180]. The type and amount of helices and sheets determine the CD pattern and a α rich protein has characteristic dips at λ = 208 and λ = 222 nm whereas a β rich protein has a single dip between λ = 215-218 nm. The minimum concentration of purified protein required for CD measurements is 10 μM. All the proteins which have been purified in this work had been characterized biophysically by CD.

2.3.7.2. Fluorescence spectroscopy

Fluorescence spectroscopy is widely used to determine the three dimensional conformation of the proteins by observing its tryptophan environment [181, 182]. A protein

containing at least one Trp residue when excited, reaches to higher energy state, and while coming back to ground state, it loses part of the energy due to collisions, and is emitted as fluorescence. Tryptophan is the most hydrophobic amino acid and is generally buried deep inside the core of the protein. As the Tryptophan residues are buried, the availability of polar solvent around them decreases due to internal quenching, hence the emission maxima (λ_{max}) is usually less [183]. A denatured protein has all its Tryptophans completely exposed and the λ_{max} increases and shifts to right side of the EMR, known as red shift. Although, Tyrosine and Phenylalanine fluoresce, their contributions is significantly less, due to low quantum yield. The fluorescence spectra can be recorded for proteins at excitation wavelengths λ = 280 nm and λ = 295 nm and the emission range is between λ = 310-400 nm. The minimum protein concentration required for Fluorescence spectroscopy is 1-2 µM. The λ_{max} was calculated by Fluorescence spectroscopy for all the purified proteins in this work.

2.3.7.3. Thermal Denaturation

Thermal denaturation is a method by which we can calculate the melting temperature (T_m) of the protein. With increase in temperature, the protein gradually loses its tertiary and secondary structure, and slowly unfolds. The temperature at which 50 % of the protein is unfolded is identified as the T_m [184]. A Thermal denaturation curve typically contains 3 zones, a pre transition, an exponential phase and a post transition. The midpoint of the line between the pre transition and post transition states, is considered as the T_m. There are various methods to calculate the T_m of a given protein, however, CD and Fluorescence spectroscopy are widely used.

Thermal denaturation by CD was performed using 10 µM protein between temperature ranges from 10 °C to 80 °C with an increment of 2 °C. The mean residual ellipticity

corresponding to λ = 222 nm was used to calculate the fractions unfolded. Similarly, T_m can also be calculated by Fluorescence spectroscopy using 1 µM protein. A spectra was also recorded using buffer alone for blank correction. Average emission wavelength was calculated as the ratio of the summation of product of blank corrected values and respective wavelength to the summation of respective wavelength. Form these values melting temperature (T_m) was calculated by plotting fraction unfolded on Y-axis with respect to temperature on X-axis.

2.3.7.4. Limited Proteolysis

Limited proteolysis is one of the best and most convenient methods to study protein folding behaviour of individual proteins. Further it will help to determine the molten globular state and the most stable domain within a given protein [185]. A fully folded and compact protein will have least number of protease cleavage sites accessible on its surface, and hence it would take more time for proteases to completely digest the protein. However, if the protein is not folded properly, all of its protease cleavage sites are easily accessible and the protein undergoes rapid digestion. This difference in time of digestion is used to characterize the most stable domains and the folded state of the proteins. Trypsin and chymotrypsin are widely used for limited proteolysis studies. The protease is added to protein of interest and immediately an aliquot of 30-60 µl was removed after adding the enzyme to record limited proteolysis at zero minute. Subsequently equal volume of aliquots will be removed at 5, 10, 15, 30, 60, 120 and 180 minutes respectively. These samples will be further analysed on 15% SDS-PAGE gel. The band corresponding to the most stable domain, which had shown resistance to protease digestion for larger time can be excised and its sequence can be detected by trypsin in-gel digestion followed by MS.

2.3.7.5. Dynamic Light Scattering

Dynamic Light Scattering (DLS) is a method to calculate the hydrodynamic radius of proteins and also to determine the oligomeric state of proteins [186]. Biomolecules like proteins in aqueous solution exhibit Brownian motion due to which there will be a fluctuation in the concentration locally. Hence, when light passes through them there will be fluctuations in the intensity of the light that is scattered. This scattered light gives us the decay time and from the time-dependent autocorrelation function, the diffusion coefficient (D_t) can be calculated by the following equation.

$$\Gamma = q^2 D_t$$

Where Γ is decay constant, D_t is diffusion coefficient and q is the scattering vector given by following equation.

$$q = (4\pi n/\lambda) \sin(\theta/2)$$

Where n is refractive index, λ is wavelength, and θ the scattering angle.

This diffusion coefficient is used to calculate the hydrodynamic radius by using Stokes-Einstein equation which is given as

$$R_h = k_B T/6\pi\eta D_t$$

Where R_h is hydrodynamic radius**,** k_B is Boltzman's constant, T is temperature (Kelvin) and η is viscosity.

This formula is used to calculate the hydrodynamic radius of the proteins. The graph depicting the autocorrelation function on Y axis and time on X axis is known as correlogram. This determines whether the distribution of protein molecules is monomodal or not. For proteins, the decay time is faster at the beginning of the correlogram, and the range of mean radius for proteins is usually between 1-10 micro seconds (μs). Further, the merger of graph at the end with basement suggests the protein solution is monodisperse in nature with the

absence of any aggregates. In DLS experiments, the percentage polydispersity should be ≤ 20 % for monodisperse solutions.

2.3.8. Protein-Protein interactions

Protein-Protein interactions (PPIs) are characterized by interactions between two or more proteins which are held by weak intermolecular interactions at the docking interface. When two proteins interact, they will form dimer and when three interact they form trimer, and with more proteins it is oligomers. All these monomers of proteins in a complex are held together as a unit by PPIs, which govern all known biophysical functions within a cell. There are various biophysical methods, to analyse these PPIs which are Pull-down assay, ITC, SPR, and *in-silico* docking [187, 188].

2.3.8.1. Pull-down assay

Pull-down assays are one of the classical methods to detect the protein-protein interactions. It is based on affinity chromatography as explained in **2.3.5.1.** section. All the tags used for affinity chromatography can be used for Pull-down assays. However, GST Pull-down assay is widely used [189]. Once the recombinant protein is bound and immobilized on the beads, its interacting partner or cell lysate is passed over the beads or incubated for 30-60 minutes. It is followed by washing with wash buffer to remove excess non-interacting proteins. The bound beads is then analysed by running SDS or native gel. In case of cell lysate, SDS-PAGE gel should be used as the individual proteins interacting with the immobilized fusion protein will be visualized as discrete bands, which can be excised and the nativity of the interacting partner can be confirmed by MS.

2.3.8.2. Isothermal Titration Calorimetry

Isothermal Titration Calorimetry (ITC) is the most commonly used method for detecting bimolecular interactions particularly Protein-Protein Interactions (PPIs) [190, 191]. This method is also used for analysing protein-DNA and protein-drug interactions. All known bimolecular interactions are accompanied with a slight heat change, which can be detected by a calorimeter in an adiabatic system. The stronger the interaction, the larger is the heat change. ITC also helps us to calculate the binding affinity (K_a), change in Enthalpy (ΔH) and also the stoichiometry involved during the interaction. The graph depicting the heat change of the interaction is known as thermogram.

2.3.8.3. Surface Plasmon Resonance

Surface Plasmon Resonance (SPR) employs a biosensor chip made of highly conductive metal sheets like gold or silver. The surface electrons on these metal sheets behave like a cloud of electrons over its surface creating an electronic plasmon. One side of this chip has a matrix to adhere the target protein (analyte) and its interacting ligand is passed over it in a mobile phase. The other side of the biosensor metal chip is exposed to the light source. When the ligand binds to the analyte, there is increase in the refractive index of the medium close to the metal surface, which alters the reflective properties of the metal chip. When the light rays strike the metal sensor chip, there will a difference in the intensity of the emitting light which is recorded in a sensorgram [192].

2.3.8.4. *In-silico* docking

Molecular docking is an *in-silico* method for predicting PPIs at atomic level. It is a powerful routinely used tool for drug discovery also [193]. It accurately predicts the docking interface and type of interactions between the interacting molecules. Macromolecular docking yields a vast number of plausible structures. The best possible structure of the interacting molecules is selected on the basis of scoring which depends on the type of software being used for docking studies. Critical Assessment of Prediction of Interactions (CAPRI) is initiated by a group of researchers from across the world to assess the efficiency of the available docking softwares by giving them the same number and type of protein ligands and receptors and was found that ZDOCK algorithm had performed relatively better than the others [194, 195]. ZDOCK algorithm was developed by Dr. Zhiping Weng group in 2003 at the Boston University. The docking scoring depends on three main principles, 1) desolvation based on Atomic Contact Energy (ACE), 2) Grid-based shape complementarity (GSC) and 3) Fast Fourier Transform (FFT) based electrostatics [196]. In 2007, Dr. Zhiping Weng group further developed ZRANK which included the van der Waals forces also [197]. The Macromolecular docking helps in supporting the PPIs confirmed by *in-vitro* experiments at atomic level.

2.3.9. Protein structure determination

The 3D structure of a protein can be determined by NMR, X-ray crystallography and cryo-EM. Depending on the size and amount of secondary and tertiary structure present, one of these methods is employed to solve the structure of the given protein. However, in certain complex proteins, a combination of these methods is employed to solve its 3D structure. The structures so far solved by these methods are deposited in the Protein Data Bank (PDB), an open repository database for all the known protein structures, which was

initiated in the early 1970's [198]. In May 2014, Nature editorials had published an article 'Hard data-It has been no small feat for the Protein Data Bank to stay relevant for 100,000 structures' suggesting the total number of PDB structures have crossed the 100 k mark.

2.3.9.1. NMR spectroscopy

In Nuclear magnetic resonance (NMR) spectroscopy the protein of interest is in solution and it is highly useful in solving the structures of flexible proteins, which are otherwise difficult to solve by X-ray crystallography [199]. However the size of protein is the major limitation of this method and is used for smaller proteins typically less than 200 amino acids. In NMR spectroscopy the protein is subjected in presence of strong magnetic field and the EMR emission due to resonance is analysed in the higher frequency range. The data yields a set of atomic distances between the nuclei known as restraints, which are further used to build the model structure of the protein. Higher the number of restraints, greater is the resolution of structure. A typical NMR structure contains a set of closely related structures know as ensemble. Highly flexible proteins have larger root-mean-square deviation (RMSD). Generally the average structure of ensemble is considered as the structure of protein.

2.3.9.2. X-ray crystallography

X-ray Crystallography is the widely used method for determining protein structures, in which protein had to be crystallized. For crystallization, huge amount of highly purified protein is required. Each protein has a tendency to crystalize in a specific condition. Hence protein crystallization should be tried with large variety of crystallizing conditions and each condition contains different combinations of a buffer, salt and precipitant. Once the crystal is formed, for determining its structure, the protein crystal is bombarded with X-rays. The

diffracting rays from the crystal are recorded on a charged coupled device (CCD) detector plate [200]. In order to obtain the completeness of the data, the crystal is rotated and diffraction data for each rotation should be collected. Once the complete data set (.MTZ files) is collected, which contains diffraction spots, their Miller indices and intensities are calculated using iMosflm within the CCP4 suite [201, 202]. These values are used to calculate the structure factor of the protein. Now, determination of the accurate phases is the key to attain the perfect structure of the protein from the data, which can be solved by employing various methods such as molecular replacement and isomorphous replacement [203, 204]. The amplitudes and data from the phase is further used to build the electron density maps. From these maps the final model structure of the protein is built, using Coot. Further the structure is refined and validated before submitting to the PDB.

X-ray crystallography method can be employed for solving protein structures of wide range of proteins ranging from few amino acids to many amino acids. Getting diffracting grade crystals is very challenging in this method and it is very difficult to crystalize flexible proteins.

2.3.9.3. Cryo-EM

Cryo-Electron Microscopy (EM) is the most recent among all the available methods to solve the protein structure. It is generally used to solve structures of very large proteins and complexes which are very difficult to crystalize, including membrane proteins, large enzymatic oligomeric complexes, viral capsid proteins and whole viral particles [205, 206]. With robust advances in EM, had seen an intense surge in the number of PDB structures submitted and the researchers are now more focussing on Cryo-EM [207]. The highest resolution attained so far from Cryo-EM is 2.2 Å for the β-galactosidase of *E. coli* in complex

with phenylethyl β-d-thiogalactopyranoside (PETG), a cell-permeant inhibitor [208]. In Cryo-EM the sample is vitrified over a thin layer of frozen layer of an inert solution and subjected to high vacuum at very low temperatures maintained by liquid nitrogen. Several projection images of the single molecule of the protein is recorded with different orientations. These multiple projection images are then assembled appropriately to align in a particular sequence and final structure will be averaged from these images [209].

2.3.9.4. Homology Modelling

Homology modelling is an *in-silico* method of determining protein structure. For homology modelling, a homologous structure of the target protein in the PDB is a prerequisite [210]. Since 1994, Critical Assessment of protein Structure Prediction (CASP) was initiated by researchers across the world to assess the performance of available protein structure prediction tools. After many rounds so far, I-tasser had been rated best for most of the times [211, 212].

Iterative Threading Assembly Refinement (I-tasser) was developed by Dr. Yang Zhang group at the University of Michigan, which is based on threading and assembly principles [213, 214]. The modelling involve different stages. Initially when a query sequence is submitted, I-tasser predicts its secondary structure using Protein Secondary Structure Prediction (PSSpred). Then a closest template is selected by fold recognition or threading using local meta-threading-server (LOMETS) [215]. There will be several locally aligned structures on the template and the unaligned regions can be built by *ab-initio* modelling. These fragments are then assembled by threading using replica-exchange Monte Carlo simulation to build the full model of the protein [216]. Since these simulations generate a large number of structures, the best possible structures are clustered out using SPICKER on

the basis of their lowest energies [217]. The selected structure will be subjected to ModRefiner for energy minimization and refinement from the Cα traces to build the final model [218].

Apart from I-tasser there are many other softwares and webservers to predict the protein structure and the most commonly used ones are MODELLER, ROBETTA, SWISS-MODEL and Rosetta@home [219-222].

Chapter 3:

Structural and functional basis to understand binding interface of ERK2-RSKs

3.1 Introduction:

Extracellular signal-Regulated Kinases (ERKs) are a subset of Mitogen-Activated Protein Kinases (MAPKs). So far, eight isoforms of ERKs have been identified, and these are characterized by the presence of a unique Thr-X-Tyr motif in their activation loop [56, 223]. ERK1/2 regulates most of the cellular activities including cell growth, motility, proliferation, differentiation and cell death [34, 224-229]. Extracellular signal-Regulated Kinase 2 encoded by MAPK1 gene, is present on the q arm of Chromosome 22 (22q11.21) in the human genome. The ERK2 comprises 360 amino acids and has a molecular weight of 41389.7 Da, and a calculated pI of 6.50 [230, 231]. ERK2 is reported to be expressed in almost every cell type. However, higher levels of expression are seen in brain and hematopoietic cells [232, 233]. Knockout studies on mice has revealed that ERK2 is indispensable for survival as ERK2 knockout mice die very early during development, and none of the other isoforms of ERKs compensate this loss [57, 58, 234].

ERK2 is a serine/threonine kinase which belongs to CMGC group of protein kinases. ERK2 is directly activated through phosphorylation by upstream MEK1/2 (MAPK/ERK kinase), also known as Mitogen-Activated Protein Kinase Kinase (MAPKK or MAP2K) [59-62]. The kinase domain of ERK2 comprises of 289 amino acids ranging from 25-313 amino acids. The signature motif of ERKs i.e. Thr-X-Tyr motif, is present at 185-187 in ERK2. ERK2 shuttles between nucleus and cytoplasm as it has both Nuclear Export Signal (NES) and Nuclear Localization Signal (NLS) [68-70]. However, ERK2 monomer remains in the cytoplasm, and dimerization favours translocation to the nucleus. Upon dimerization of ERK2, NES is completely masked as it is close to the dimer interface, and the ERK2 dimer has two active NLS which favours its nuclear translocation [71].

Apart from MEK and RSKs in the MAPK pathway, ERK2 has a number of interacting partners, both in the nucleus and cytoplasm which includes ELK1, DAPK1, DUSP1, GSK3B, STAT5A, PEA15, RPTOR, HDAC4, PML and TH [235-244]

ERK2 full length structure had been determined and there are many entries for ERK2 in the Protein Data Bank (PDB) such as 4QP9, 4ZZN, 4N0S and 4IZA [245-248]. The best resolved crystal structure for ERK2 (1-360) has been reported up to 1.45 Å (PDB: 4QTA) [249]. ERK2 is predominantly α-rich protein with a helical content of around 38 %. However, our literature search could not find any complex structure of ERK2 with binding partners like RSKs/MSKs. We wanted to crystallize the ERK2-RSKs binding interface to unravel the function associated to protein-protein interactions. The serious efforts were made to crystallize the ERK2-RSKs complex but, however, there was no success in getting crystals.

This chapter describes the purification of ERK2 protein and its interaction with RSK3. The interactions studies were performed using *in-vitro*, biophysical experiments and *in-silico* tools. The *in-vitro* experiments were GST pull down assay and, biophysical interactions using Isothermal Titration Calorimetry (ITC) and macromolecular docking, that was performed using *in-silico* tools.

3.2. Materials and methods

3.2.1. Expression and purification of 6 HIS-ERK2 full length protein

The ERK2 (1-360) cloned in the bacterial expression vector pRSET-A with 6 HIS tag, was transformed into BL21 (DE3) competent cells and plated on LB agar plate containing 100 µg/ml Ampicillin. A single isolated colony was inoculated in 10 ml of Luria Broth along with 10 µl of 100 mg/ml Ampicillin, and was allowed to grow in a shaker incubator (SIF-6000R, Medline scientific) operating at 37 °C and 200 rpm. After 8 hours of

incubation, the starter culture was diluted into one litre of Luria Broth with 100 µg/ml Ampicillin. This diluted culture was further incubated in a shaker incubator (Multitron standard, Infors HT) at 37 °C, 200 rpm, till the optical density (OD) at λ = 600 reached 0.5. To this culture isopropyl-β-D-thiogalactoside (IPTG) was then added to a concentration of 0.3 mM and further incubated at 20 °C, 200 rpm. Following 16 hours of induction, the bacterial culture was harvested by centrifugation at 4 °C, 5000 rpm for 20 minutes.

Bacterial pellet from one litre culture was resuspended in 10 ml of Buffer A (50 mM Tris-HCl pH 8.0, 500 mM NaCl), additionally containing protease inhibitor cocktail in 1:1000 dilution (Sigma-Aldrich). Further, sonication was performed at 4 °C for cell wall disruption by digital sonifier (Bransson). 5 cycles of sonication were performed, each cycle of one minute containing a repeated pulse of 4 seconds with an intermittent break of 2 seconds. After every cycle, one minute gap was given and the reaction mixture was maintained on ice throughout the process. The broken cells were centrifuged in a JA-25.50 sorvall rotor at 18,000 rpm, 4 °C for 45 minutes (Avanti J-26S XPI, Beckman coulter). The supernatant was used for IMAC purification of 6 HIS-ERK2 full length protein using Ni-NTA (Qiagen) affinity chromatography.

The Ni-NTA resin (Qiagen) was equilibrated with buffer A, and binding was performed by adding the supernatant containing ERK2 protein by gravity flow method. The flow through was collected in a beaker kept on ice, and was again passed back onto the Ni-NTA resin in order to get maximum binding. The protein bound resin was washed with 10 column volume of wash buffer i.e. buffer A with 10 mM Imidazole. Another 20 column volume wash were performed using buffer A with 20 mM Imidazole. Further 4 more column volume washes, each with buffer A containing 40 mM and 70 mM Imidazole respectively were performed and eluent were collected to check if there is any loss of ERK2. After these stringent washes, ERK2 full length protein was eluted in buffer A with 150 mM and 300 mM

Imidazole. The elution fractions were analysed on 10 % SDS-PAGE gel and visualized by Coomassie staining. The elution fractions containing ERK2 were pooled, and further concentrated up to 2 ml for purification using Size Exclusion Chromatography (SEC), AKTA-FPLC (GE) in the buffer B (10 mM Tris, 100 mM NaCl, pH 8.0). These FPLC purified fractions were analysed on 10 % SDS-PAGE gel.

3.2.2. Peptide mass fingerprinting by mass spectrometry

The ~ 41 kDa band of purified ERK2 full length protein was excised from the SDS-PAGE gel, and chopped into tiny pieces. In-Gel Trypsin digestion was performed as mentioned by Shevchenko et al [250]. The tryptic cleavage products were then extracted and finally reconstituted in 50 % Acetonitrile containing 0.1 % TFA. The mass spectra of these tryptic cleavage products were recorded by Ultraflex II MALDI-TOF/TOF mass spectrometer (Bruker Daltonics) and identified using Peptide mass fingerprinting by comparing with those in the SwissProt database using Mascot (Matrix Sciences) search engine.

3.2.3. CD Spectroscopy and Fluorescence spectroscopy of ERK2

For CD Spectroscopy (Jasco J-815 CD spectrophotometer), 300 µl of highly FPLC purified ERK2 full length protein in Buffer B at a concentration of 10 µM, was taken in quartz cuvette, and a spectrum was recorded between λ = 250 and 190 nm. To estimate the secondary structure the data obtained from CD spectroscopy was submitted to K2D3 online server [251].

For Fluorescence spectroscopy **(**Horiba scientific), 200 µl of 1 µM of ERK2 protein in buffer B was used. The spectra were recorded with an excitation wavelength λ = 295 nm, and emission spectrum was recorded between λ = 310 nm and 400 nm.

3.2.4. Cloning and ligation RSK3 C-terminal tail (669-733) in pGEX-KT vector

C-terminal tail (669-733) of RSK3 has been cloned in pGEX-KT vector between BamH1 and EcoR1 restriction sites. The PCR primers required for cloning were designed manually using the sequence information of RSK3 from (669-733) residues. The primer sequences are as follows. Forward primer:

5’ -GTCGGATCCGAGAACCTGTACTTTCAGGGTCACCCGTGGGTGGTCAAC- 3’

Reverse primer: 5’-GTCGAATTCCTATTACAGCCGCGTGGACCTGAGTCT- 3’. The tetrad PCR system (Bio-Rad), and the Phusion High-Fidelity DNA Polymerase enzyme (Finzymes) were used to perform the PCR. Initial denaturation step was at 95 °C for a duration of 5 minutes, PCR cycle comprised of denaturation at 95 °C for 30 seconds, primer annealing at 65 °C for 30 seconds and the polymerase extension at 72 °C for 30 seconds, and this cycle was repeated 32 times. The amplified PCR product was loaded onto 1 % agarose gel, and visualized by staining with ethidium bromide. The PCR product corresponding to 240 bp was excised under a UV filter on a Gel doc equipment (UVP bioimaging systems). The PCR product purified by gel extraction kit (Qiagen), was further treated with BamH1 and EcoR1 restriction enzymes (Fermentas) to generate sticky ends. This digested RSK3 C-terminal tail was then ligated (Quick Ligation kit from New England Biolabs) for 10 minutes at room temperature into pGEX-KT vector, already treated with BamH1 and EcoR1 restriction enzymes. The ligated product was transformed into DH5α competent cells and plated on Luria Agar containing 100 µg/ml Ampicillin. A single isolated colony from the

plate was then used to inoculate 10 ml of Luria Broth containing 100 µg/ml Ampicillin, and the culture was incubated at 37 °C for 12 hours. The bacterial culture was harvested by centrifugation at 4 °C and 6000 rpm. The recombinant plasmid was isolated using miniprep kit (Qiagen). The presence of RSK3 C-terminal tail insert was confirmed by agarose gel electrophoresis on the sample subjected to double digestion with BamH1 and EcoR1 restriction enzymes at 37 °C. The identity of the inserted PCR product in the recombinant plasmid was further confirmed by DNA sequencing (Applied Biosystems).

3.2.5. Expression and purification of RSK3 C-terminal tail (669-733)/pGEX-KT

The RSK3 (669-733)/pGEX-KT construct was transformed into BL21 (DE3) competent cells and plated on Luria agar containing 100 µg/ml Ampicillin. A 10 ml of Luria Broth medium containing 10 µl of Ampicillin (100 mg/ml) was inoculated with a single isolated colony and was further incubated in a shaker incubator at 37 °C and 200 rpm. After 8 hours of incubation, the starting culture was diluted into 1 litre of Luria Broth, with 1 ml Ampicillin (100 mg/ml) and was further allowed to grow in a shaker incubator at 37 °C, 200 rpm till the optical density (OD) at λ = 600 reached between 0.6-0.8. The culture was then induced with 0.3 mM isopropyl-β-D-thiogalactoside (IPTG), and incubated for 16 hours at 20 °C, at 200 rpm. The bacterial culture was harvested by centrifugation for 15 minutes at 4 °C and 6000 rpm.

The molecular mass of RSK3 C-terminal tail (669-733) is 7442.6 Da and the fusion protein is around 33.5 kDa along with GST tag. For purification, one litre bacterial pellet of RSK3 C-terminal tail (669-733) was lysed in buffer A, containing protease inhibitor cocktail in a ratio of 1:1000 by volume (Sigma-Aldrich). The bacterial cells were disrupted by sonication (Branson digital sonifier) for 5 one minute cycles separated by one minute

incubation on ice, keeping the pulse on for 4 seconds and off for 2 seconds in every cycle. The sample was maintained on ice throughout the process. The cell lysate was then centrifuged in a JA-25.50 sorvall rotor at 18,000 rpm, 4 °C for 45 minutes (Avanti J-26S XPI, Beckman coulter). The supernatant was used for affinity purification of RSK3 C-terminal tail (669-733) using GST (Novagen) resin.

The supernatant was allowed to bind on 1 ml of GST resin (Novagen) pre-equilibrated with buffer A. The binding was performed for 1 hour in gravity flow method. After 20 column washes with buffer A, TEV cleavage was performed by adding 100 µl of 1 mg/ml of TEV protease in 1 ml buffer A, and allowed cleavage reaction for 2 hours before elution. The elution fractions were loaded on 15 % SDS-PAGE gel and were visualized by coomassie staining.

3.2.6. GST Pull down assay

For GST pull down assay, GST tagged RSK3 C-terminal tail (669-733) and FPLC purified ERK2 were used. The GST tagged RSK3 C-terminal tail (669-733) fusion protein was allowed to bind on 1 ml of GST resin, packed in a column. Following significant buffer washes of the column, FPLC purified ERK2 was then loaded on this column. The binding was done for 1 hour by gravity flow method. After 20 column washes with buffer B, 20 µl of GST resin, removed from the column was loaded on 12 % SDS-PAGE and the protein bands were visualized using coomassie staining.

3.2.7. Binding analysis between ERK2 and RSK3 peptide

Isothermal Titration Calorimetry (ITC) was performed to study the binding affinity between ERK2 and RSK3 C-terminal tail. The RSK3 peptide with the residues from 709-726 (NH_2-PRLEPVLSSNLAQRRGMK-COOH) was commercially procured. This peptide also comprises the ERK2 binding motif which is NH_2-^{719}LAQRR^{723}-COOH. Both the ERK2 and RSK3 C-terminal tail peptide were in buffer B and the reaction was carried at 25 °C using ITC-200 (GE life sciences). The concentration of ERK2 was 30 µM, and that of the RSK3 peptide was 300 µM. The ITC reaction was performed keeping ERK2 in the sample cell and RSK3 peptide in the syringe. Total 20 injections, each of 2 µl was injected into the sample cell with a spacing of 120 seconds between successive injections. The detailed experimental parameters are mentioned in **Table 3.1.** Similarly another reaction was done with buffer B in the sample cell and RSK3 peptide to subtract the heat of dilution for the base line correction. The data was analysed using Origin software (OriginLab, Northampton, USA) to calculate the thermodynamic parameters and the stoichiometry of the reaction.

Table 3.1: ITC experimental parameters for ERK2 and RSK3 C-terminal tail peptide

Experimental parameters	
Total injections	20
Cell Temperature	25 °C
Reference power	8 µCal/sec
Initial Delay	60 sec
Syringe concentration	0.3 mM
Cell concentration	0.03 mM
Stirring speed	1000 RPM
Injection parameters	
Injection volume	2 µl
Duration	4 sec
Spacing	120 sec
Filter period	5 sec

3.2.8. *In-silico* modelling of MEK1 and RSK1

The N-terminus residues, 3-11 of MEK1, contains the ERK docking site. Since the reported structure of MEK1 (PDB id: 3E8N) does not contain 60 residues at N-terminal end, I-tasser web server was used to build the model of full length MEK1 protein [213, 252]. Furthermore, in the structure of C-terminal end of RSK1 (PDB id: 2WNT), the ERK docking site was missing, therefore C-terminal domain of RSK1 was also modelled using I-tasser server. The built models were then validated using Procheck and Verify_3D server [253, 254].

3.2.9. Modelling of RSK1/2/3/4 peptides

RSK sequences comprising of 14 amino acids at the C-terminal end of all RSK isoforms including the ERK binding motif 'NH_2-LAQRR-COOH' were selected and the sequences are as follows, RSK1: NH_2-IESSILAQRRVRKL-COOH, RSK2: NH_2-VGRSTLAQRRGIKK-COOH, RSK3: NH_2-VLSSNLAQRRGMKR-COOH and RSK4: NH_2-VAASSLAQRRSMKK-COOH. These sequences were then subjected to BLAST in order to find a suitable template of known structures from Protein Data Bank (PDB) [255]. A web-based programme, PHYRE was used for protein fold recognition, which is based on primary sequence, secondary structure and solvation potential [256]. To build the structures for the 14 amino acid sequences of RSK1/2/3/4, a built-in protein tool within the Discovery studio 2.5.5 (Accelrys: San Diego, USA) was used. CHARMM force field was used to refine the side chains and energy minimization [257]. KYTE-DOOLITTLE plot within the ExPASy tools was used to calculate the hydrophobic index of these 14 amino acid residues for all the four RSK isoforms, keeping the window size as 7 [258].

3.2.10. Protein-protein docking of MEK1 and ERK2

In-silico docking of modelled MEK1 and ERK2 was performed using ZDock from Discovery Studio 2.5. The hydrophobic (NH_2-^{9}IQL^{11}-COOH) and basic (NH_2-^{3}KKK^{5}-COOH) residues at N-terminal of MEK1 and acidic residues (Asp318 and Asp321) within the L16 loop near the C-terminal of ERK2 were assigned for directed docking. Of the many poses generated, the best docked pose was selected on rank basis which is derived from the combination of shape complementarity, the desolvation principle and electrostatic energies.

3.2.11. Protein-protein docking of ERK2 and RSK1 CTKD

Docking of ERK2 and modelled RSK1 CTKD was performed using ZDock from Discovery Studio 2.5. The acidic residues (Asp318 and Asp321) within the C-terminal L16 loop of ERK2 and the basic residues (Arg725 and Arg726) of RSK1 were filtered for docking. The best docked pose was selected on the basis of its scoring with respect to shape complementarity, the desolvation principle and electrostatic energies.

3.2.12. Protein-protein docking of RSK1 CTKD on MEK1-ERK2 complex

The RSK1 structure was docked onto the dimeric complex model of MEK1-ERK2. For docking, the RSK1 residues (NH_2-^{722}LAQRRVRKL^{730}-COOH) within the C-terminal region and acidic residues (Asp318 and Asp321) within the C-terminal L16 loop of ERK2 were selected. On the basis of the docking scores, the best pose of the hetero-trimeric complex was selected.

3.2.13. Docking of RSK1/2/3/4 peptides on ERK2

The modelled RSK peptides of 14 amino acids of all the four RSK isoforms were docked onto ERK2 molecule using the ZDock program keeping the angular step size as 15. The best pose was selected on the basis of the docking score and optimal Hydrogen bonding. Ligplot tool was then used to assess the atomic interactions among these selected top poses of all the four complexes [107].

3.3. Results and discussion

3.3.1. Purification of ERK2 full length

6 HIS tagged ERK2 full length protein (1-360) was purified from *E. coli* BL21 (DE3) competent cells. **Figure 3.1** shows purified 6 HIS tagged ERK2 full length protein. The elution fractions containing ERK2 were pooled and loaded onto FPLC Superdex-75 column to get highly purified protein. **Figure 3.2 (A)** shows FPLC purified ERK2 while **Figure 3.2 (B)** shows the FPLC chromatogram of the purification.

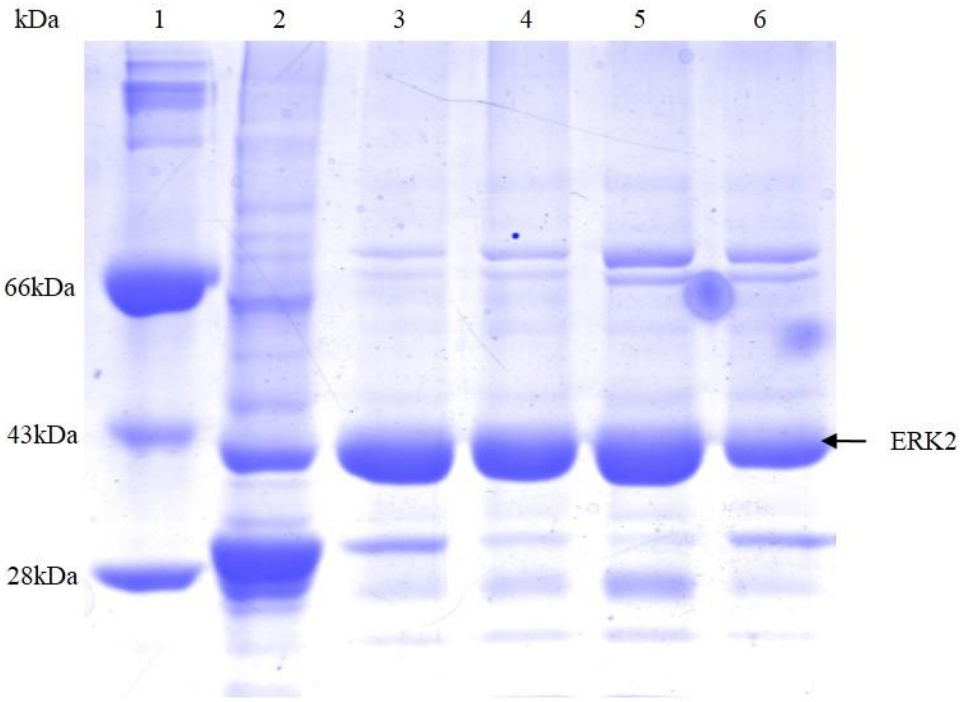
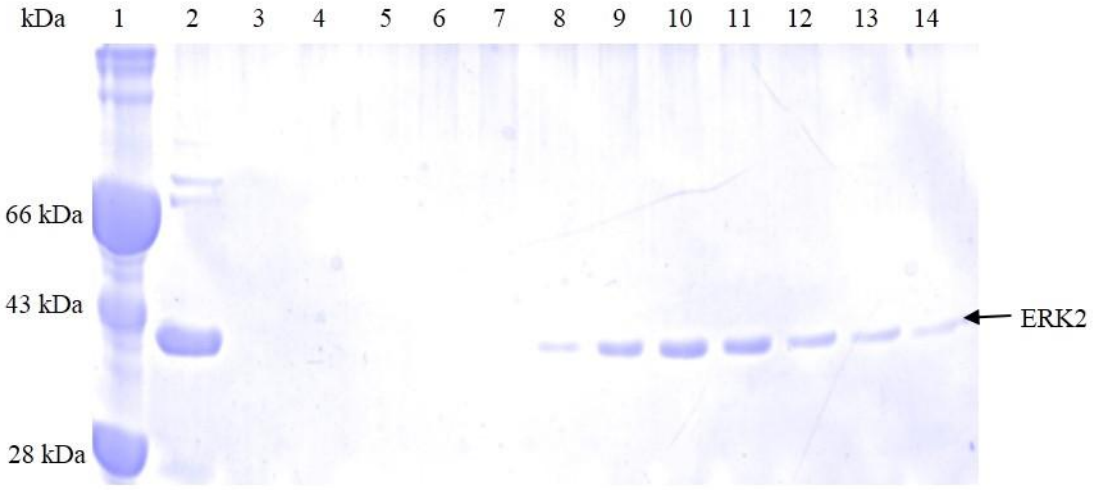
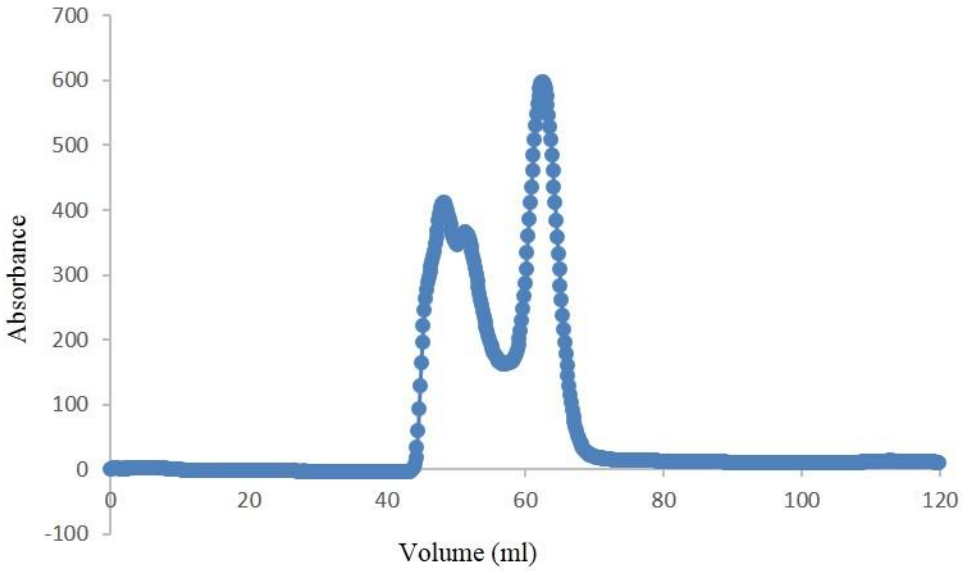


Figure 3.1: Purification of ERK2 full length. The 10 % SDS-PAGE gel shows the 41 kDa ERK2 full length protein eluting at 150 mM imidazole. (lane 1: Protein marker, Lanes 2: 70 mM Imidazole elution. Lanes 3-6: 150 mM Imidazole elutions)



(A)

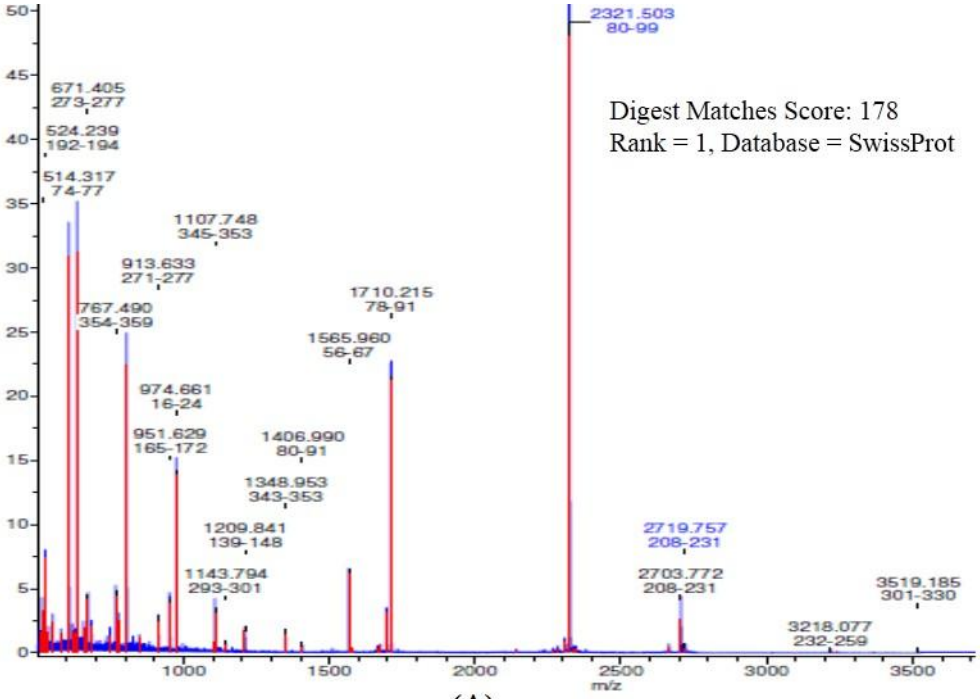


(B)

Figure 3.2 : FPLC Purification of ERK2 full length. (A) The 10 % SDS-PAGE gel shows the 41 kDa ERK2 full length purified from Superdex-75 column. (lane 1: Protein marker, Lane 2: Concentrated protein before FPLC, lanes 8-14: 60 ml fractions). (B) FPLC Chromatogram of ERK2 purification.

3.3.2. Detection of purified ERK2 protein

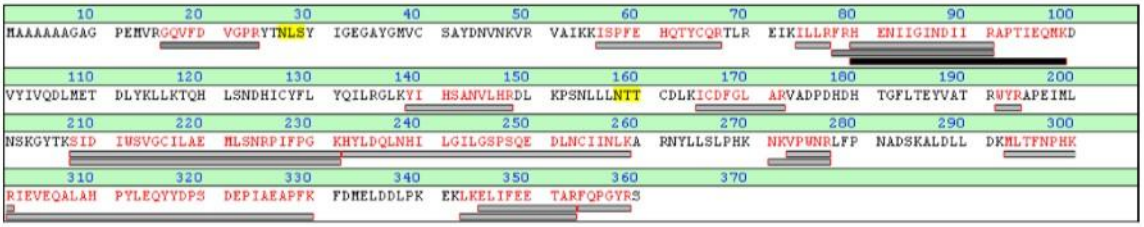
To check the nativity of purified protein, Trypsin in-gel digestion was performed. **Figure 3.3 (A)** shows the identified extracted peptides with their m/z values as determined by MALDI-TOF mass spectrometry. **Figure 3.3 (B)** shows the Peptide mass fingerprinting from Mascot search engine using SwissProt database confirming the peptides of ERK2.



(A)

Mitogen-activated protein kinase 1 OS=Homo sapiens GN=MAPK1 PE=1 SV=3 MK01_HUMAN

Intensity Coverage:	48.8 % (113788 cnts)	Sequence Coverage MS:	50.6%
Sequence Coverage MS/MS:	0.0%	pI (isoelectric point):	6.6



(B)

Figure 3.3: Peptide mass fingerprinting of in-gel tryptic digests of ERK2. (A) MALDI-TOF/mass spectrometry data showing the identified tryptic peptides. (B) The Mascot search engine confirmed the peptides were corresponding to ERK2.

3.3.3. Circular-Dichroism and Fluorescence spectroscopy of ERK2

To analyse the secondary structure of ERK2, Circular-Dichroism spectroscopy was performed. **Figure 3.4** shows the CD spectrum which suggests that the protein is predominantly α-helical in nature as per dips at wavelengths λ = 208 and 222 nm respectively. The CD spectrum is consistent with the ERK2 crystal structure. To estimate the

α-helices and β-sheets contents, the CD spectrum was submitted to K2D3 online server which predicted α-helical content to be around 42 % and β-sheets content to be around 15 % [251]. The 15 % β-sheets is corresponding to the small dip at λ = 218 nm.

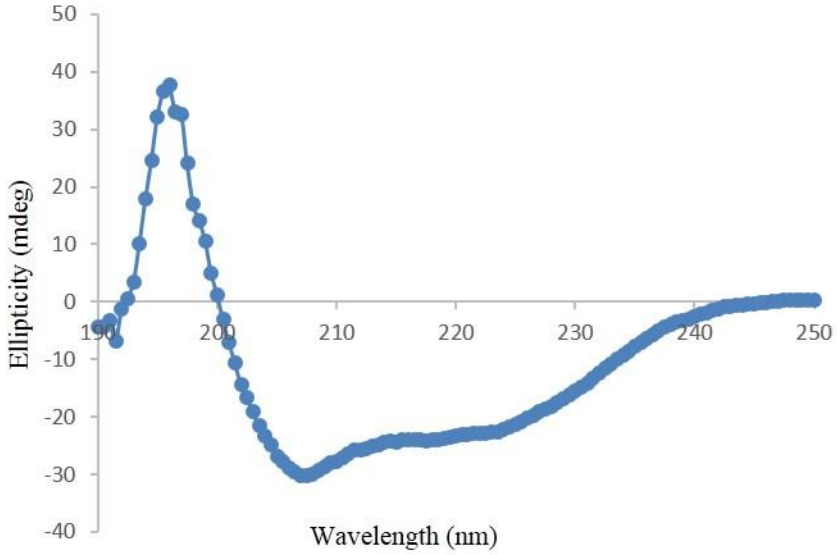


Figure 3.4: CD analysis of ERK2 full length. The CD Spectra of ERK2 full length shows the dips at λ = 208 nm and λ = 222 nm indicates it is predominantly α-helical which is in consistent with the crystal structure of ERK2.

Furthermore, to analyse the tertiary conformation of ERK2 full length protein, Tryptophan Fluorescence spectroscopy was performed. There are 3 tryptophan (Trp) residues in ERK2 full length protein (1-360) at positions Trp 192, Trp 212 and Trp 275. The fluorescence spectrum was recorded with an excitation wave length of λ = 295 nm and the emission spectra was collected from λ = 310 to 400 nm. **Figure 3.5** shows the emission maxima (λ_{max}) of ERK2 full length after blank correction at λ = 335 nm.

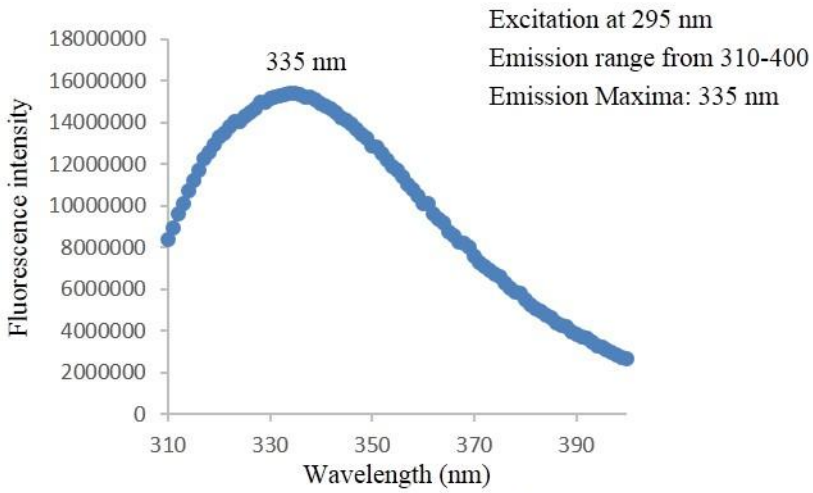


Figure 3.5: Fluorescence spectroscopy of ERK2 full length. The Fluorescence spectra of ERK2 recoded with a excitation wavelength λ = 295 nm shows the emission maxima (λ_{max}) is 335 nm.

3.3.4. Cloning and ligation of RSK3 C-terminal tail (669-733) in pGEX-KT vector

RSK3 C-terminal tail (669-733) was cloned into bacterial expression GST tag based pGEX-KT vector. **Figure 3.6 (A)** shows PCR amplified RSK3 C-terminal tail (669-733) product of 240 base pairs. Ligation was performed between RSK3 C-terminal tail (669-733) and pGEX-KT vector, both of which were doubly digested with BamH1 and EcoR1 restriction enzymes. **Figure 3.6 (B)** shows the 240 base pair RSK3 C-terminal tail (669-733) insert release, (which additionally includes the DNA sequence coding for TEV cleavage site, restriction sites, stop codons and the terminal clamps) from the recombinant pGEX-KT plasmid following a double restriction digestion with BamH1 and EcoR1 restriction enzymes. **Figure 3.7** shows the sequencing data as viewed in BioEdit software which confirms the 100 % sequence of RSK3 C-terminal tail (669-733) and is in proper frame.

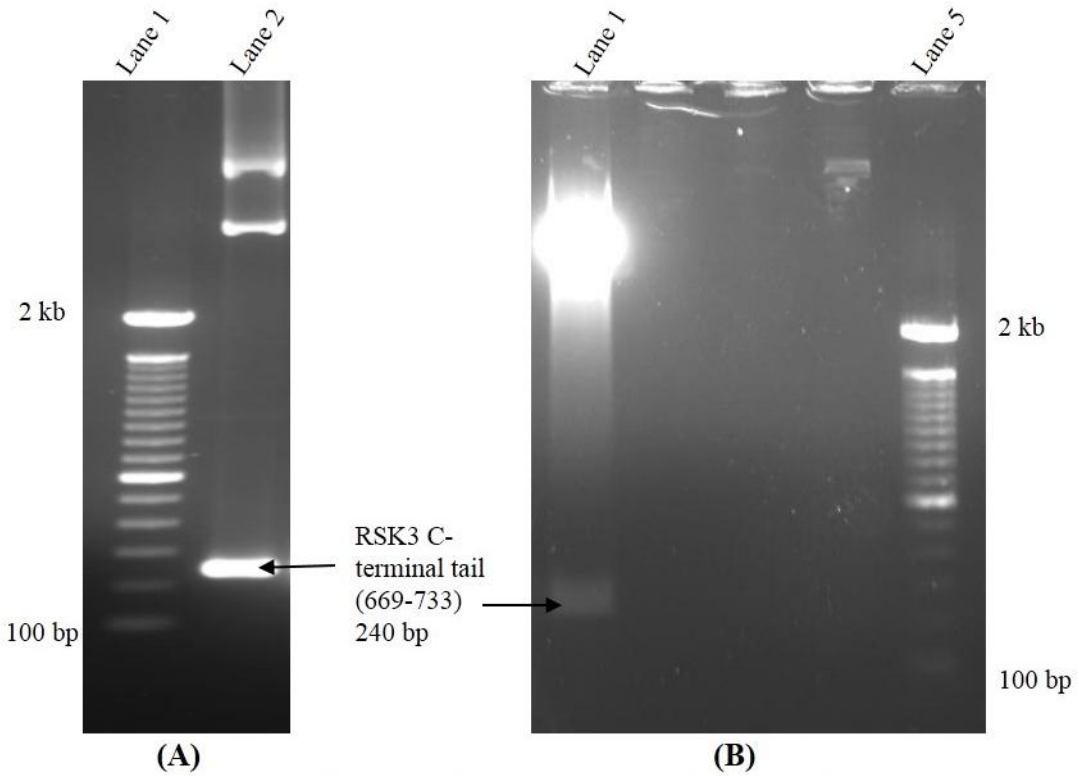


Figure 3.6: Cloning of RSK3 (669-733) C-terminal tail. (A) PCR amplification of RSK3 C-terminal tail (669-733) (Lane 1: 100 bp ladder, Lane 2: PCR product) (B) Insert release of RSK3 C-terminal tail (669-733) from the vector backbone following double restriction digestion. (Lane 1: RSK3 C-terminal tail (669-733)/pGEX-KT vector after double restriction digestion, Lane 5: 100 bp ladder)

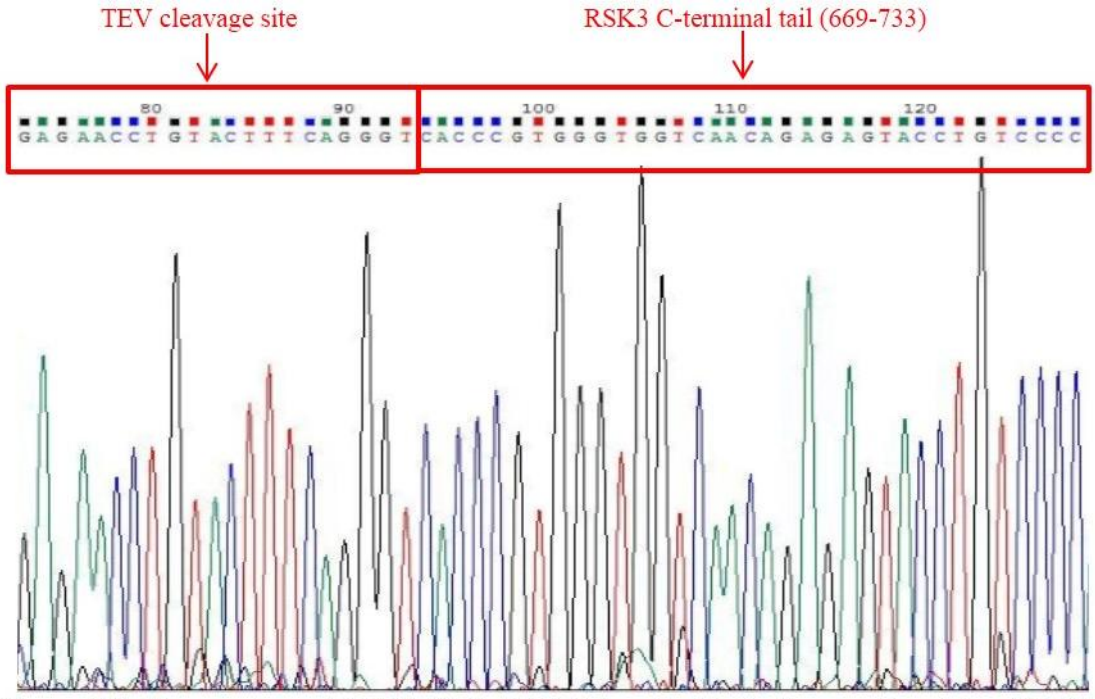


Figure 3.7: Sequencing of RSK3 (669-733) C-terminal tail. Sequencing data of RSK3 C-terminal tail (669-733)/pGEX-KT recombinant construct with pGEX-KT forward primer.

3.3.5. Expression and purification of RSK3 C-terminal tail (669-733)/pGEX-KT

The Molecular weight of RSK3 C-terminal tail (669-733) is ~ 7.0 kDa and the fusion protein is ~ 33.0 kDa. The RSK3 C-terminal tail (669-733) was expressed in BL21 (DE3) competent cells and purified by affinity chromatography using Glutathione S-Transferase (GST) resin and analysed on 15 % SDS-PAGE gel. **Figure 3.8** shows the purification profile of RSK3 C-terminal tail (669-733) of ~7 kDa which is being cleaved into 3 shorter polypeptides following two hours of TEV cleavage. The RSK3 C-terminal tail (669-733) is more stable only with the tag and undergoes rapid proteolysis when the tag is removed.

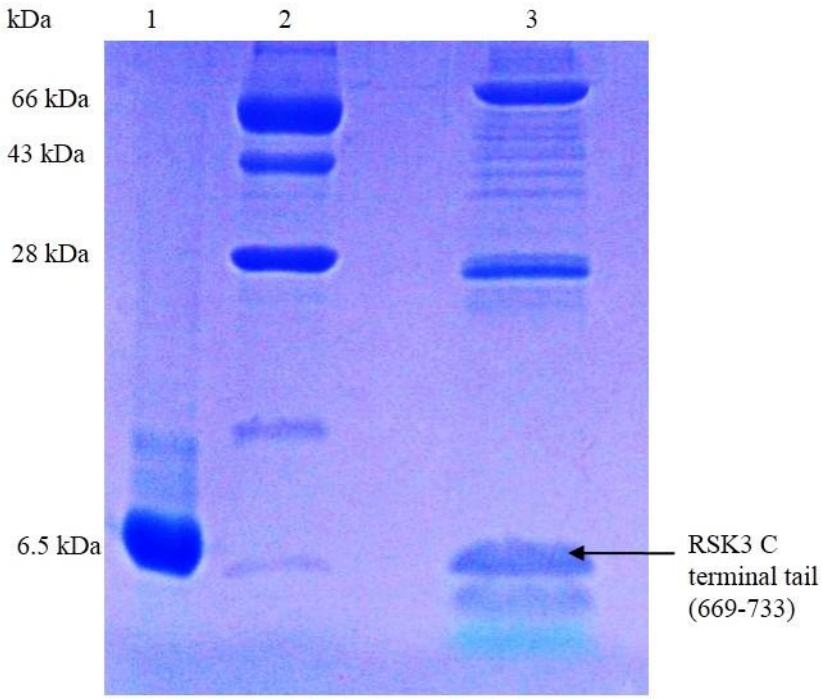


Figure 3.8: Purification profile of RSK3 C-terminal tail (669-733). The 15 % SDS-PAGE gel shows the RSK3 C-terminal tail (669-733) being cleaved into 3 fragments following TEV cleavage from GST tag. (Lane 1: Aprotinin, 6.5 kDa; Lane 2: Protein marker; Lane 3: Elution following 2 hours of TEV cleavage).

3.3.6. GST Pull down assay

For GST pull down assay, GST tagged RSK3 C-terminal tail (669-733) was bound on GST resin, to pull down FPLC purified ERK2. After incubation of GST-RSK3 C-terminal tail (669-733) with ERK2 full length for an hour, significant washes were given, and the protein bound on GST resin was analysed on 12 % SDS-PAGE Gel. The **Figure 3.9** shows ERK2 was washed out and could not bind to GST tagged RSK3 C-terminal tail (669-733).

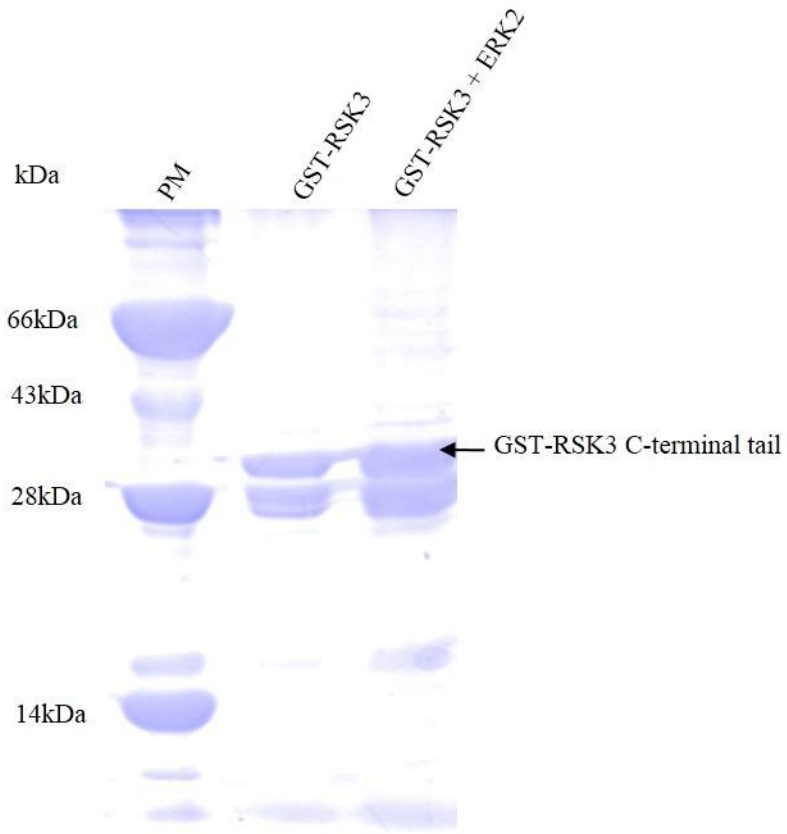


Figure 3.9: GST pull down assay. The 12 % SDS-PAGE gel shows the ERK2 (41 kDa) could not be pulled by GST-RSK3 C-terminal tail.

3.3.7. Binding affinity between ERK2 and RSK3 peptide

Noting the relatively low resolution of GST pull down assay, binding affinity between ERK2 with RSK3 peptide was performed using Isothermal Titration Calorimetry which is more sensitive. Purified ERK2 was taken in the sample cell and RSK3 C-terminal

peptide was used as titrant. **Figure 3.10** shows the heat change recorded between the ERK2 full length and RSK3 C-terminal peptide. After subtracting the heat of dilution, the data was fitted in to one-site binding model and the stoichiometry was calculated as N = 0.450 (±0.0827), with the dissociation constant of 1.54 µM. The Change in Enthalpy (H) was -4.487 kcalmol^{-1} and Entropy (S) was 11.5 calmol^{-1}K^{-1}. ERK2 is known to interact with 100-fold lower specificity (kcat/Km) towards peptide substrates than the full length protein substrates [259, 260]. Since ERK2 is purified from *E. coli* system, it is not phosphorylated and owing to these reasons, the affinity recorded is moderately low.

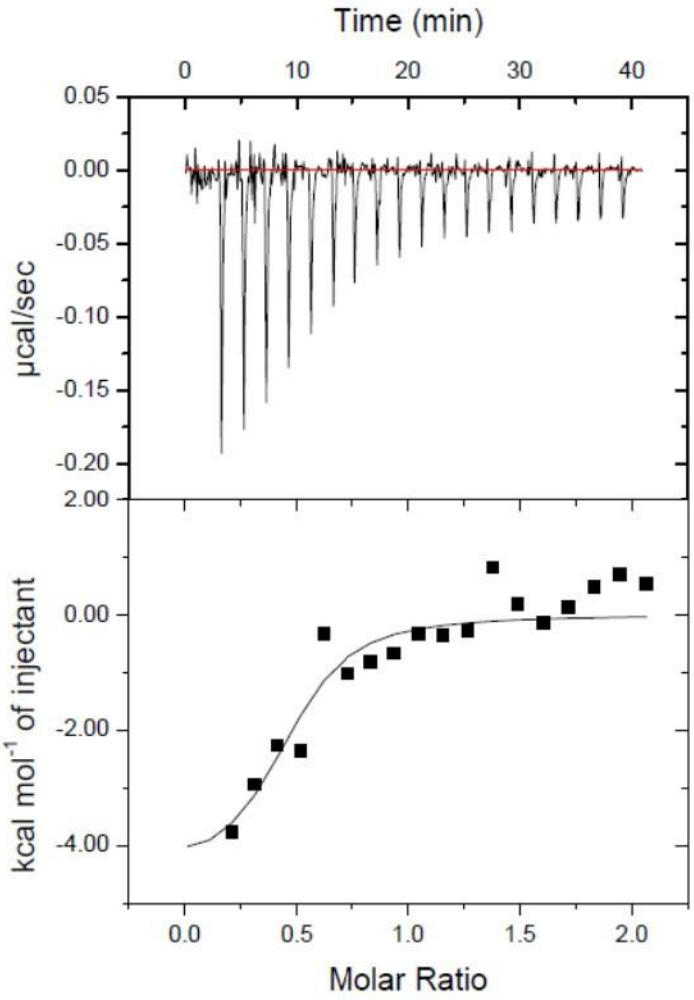
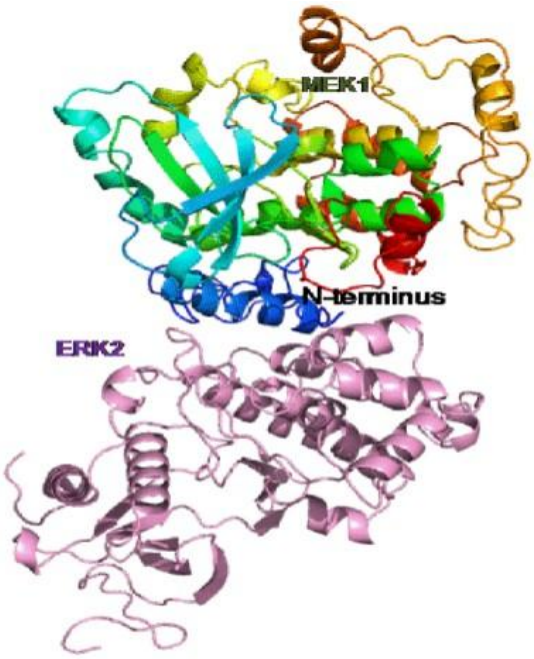


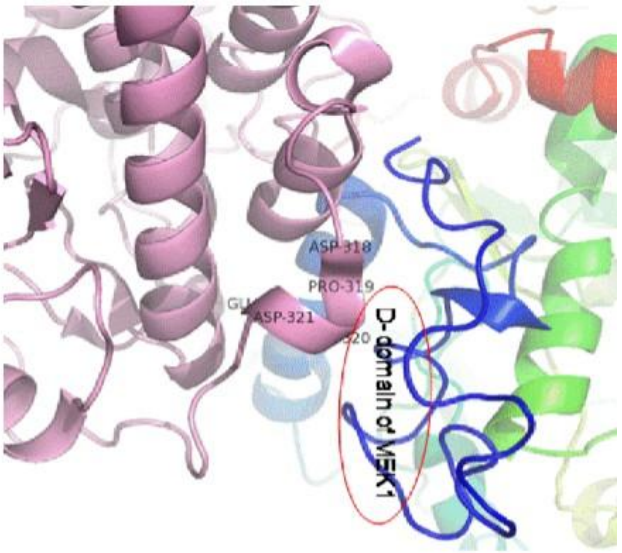
Figure 3.10: Binding analysis between ERK2 and RSK3 C-terminal tail peptide. Binding thermogram for ERK2 with RSK3 C-terminal tail peptide, shows it is an exothermic reaction (The upper panel shows the raw data, while the lower panel shows the heat of dilution corrected integrated binding isotherm along with the fitted binding curve).

3.3.8. Protein-protein docking of MEK1 and ERK2

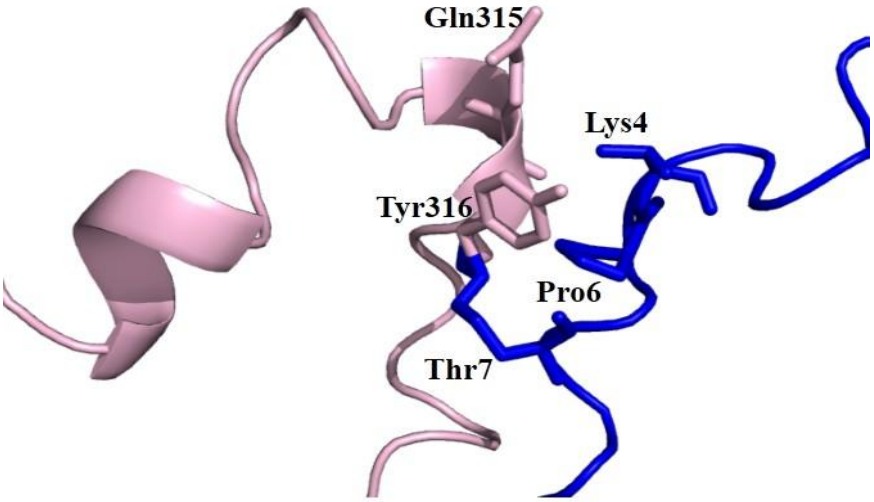
Since ERK2 exits both as a monomer and as a homodimer, we had explored the docking of ERK2 with MEK1. The N-terminus of MEK1 has mostly hydrophobic and basic amino acids which interact with the acidic amino acids within the C-terminal region of ERK2. **Figure 3.11 (A)** shows the molecular docking between the modelled MEK1 and ERK2 (PDB id: 1TVO) revealing the atomic orientations of amino acids at binding interface. Interestingly, the Clustalw scores for sequences of MEK1 and ERK2 scored as low as 18, yet retaining the 3D structure particularly at the ATP binding pocket. Major distortions were seen only at the N and C-terminal ends [67, 261, 262]. The important residues interacting and the type of interaction in this modelled 1:1 complex are listed in **Table 3.2**. Only the top 5 selected poses Z-rank scores are listed. **Figure 3.11 (B)** shows the ERK2 residues Tyr316, Tyr317, Asp318, Asp321, and Glu322 at the binding interface of MEK1. **Figure 3.11 (C)** shows the zoomed view of ERK2 residues Gln315 and Tyr316 interacting with Lys4, Pro6 and Thr7 of MEK1 through non bonded contacts. **Figure 3.11 (D)** shows the MEK1 docked onto the dimeric structure of ERK2. Due to steric hindrances the trimeric complex was not stable suggesting that, such a complex between MEK1 and ERK2 dimer may not be possible. This further suggests that the MEK1 remains in the cytoplasm itself, whereas the dimeric ERK2 translocates into the nucleus.


MEK1
N-terminus
ERK2

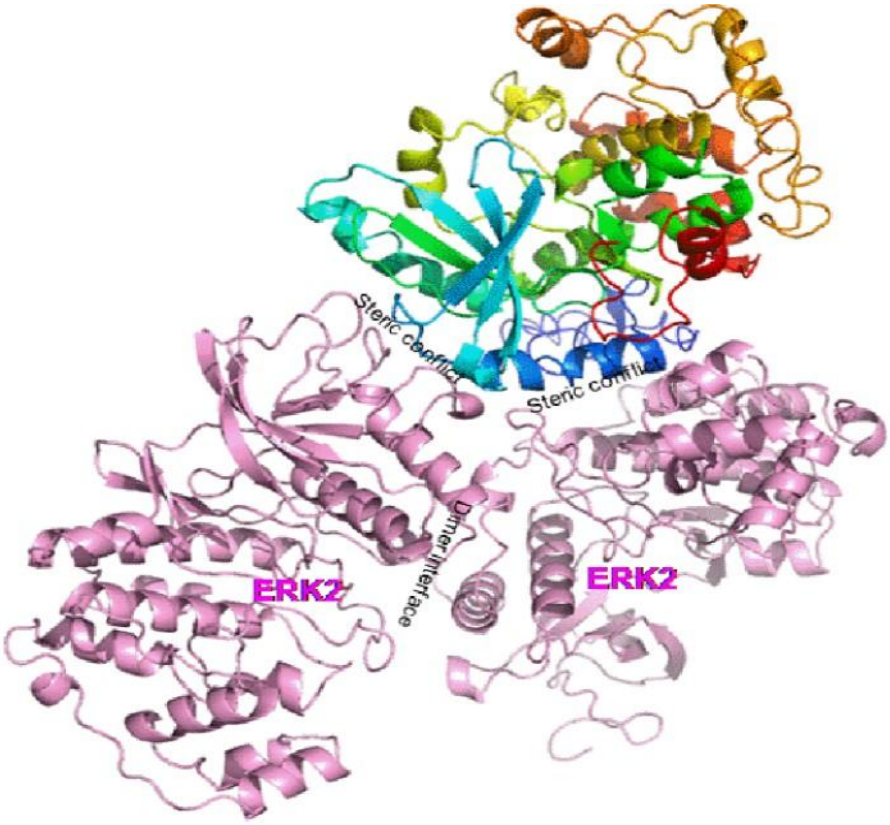
(A)


ASP-318
PRO-319
ASP-321
D-domain of MEK1

(B)


Gln315
Lys4
Tyr316
Pro6
Thr7

(C)



(D)

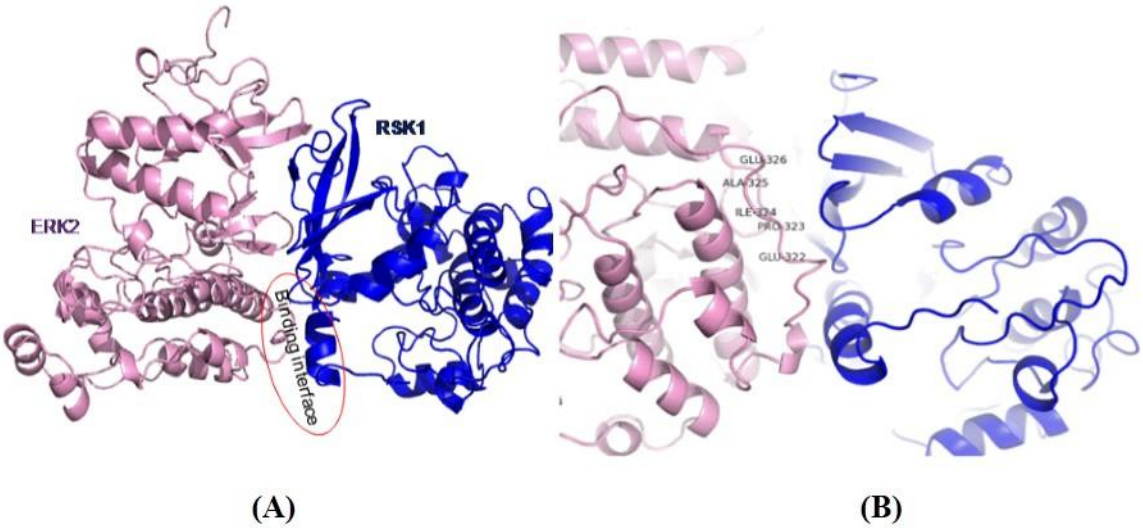
Figure 3.11: *In-silico* Docking of MEK1 model on ERK2 (A) N-terminus of MEK1 is forming complex with ERK2 molecule. (B) The important residues at the docking interface of ERK2 is labelled and the docking domain of MEK1 is encircled. (C) The important interacting MEK1 residues namely Lys4, Pro6 and Thr7 are shown in blue sticks whereas ERK2 residues Gln315 and Tyr316 are shown in pink sticks. There are steric clashes between the docked structure of p-ERK2 dimer and MEK1 (ERK2 is coloured in pink and MEK1 in rainbow).

Table 3.2: Molecular interactions between ERK2 (PDB id: 1TVO) and MEK1 (PDB id: 3E8N). (Bold letters indicate **ERK2** and italics indicate *MEK1*. The Pose finally selected was highlighted in Green colour).

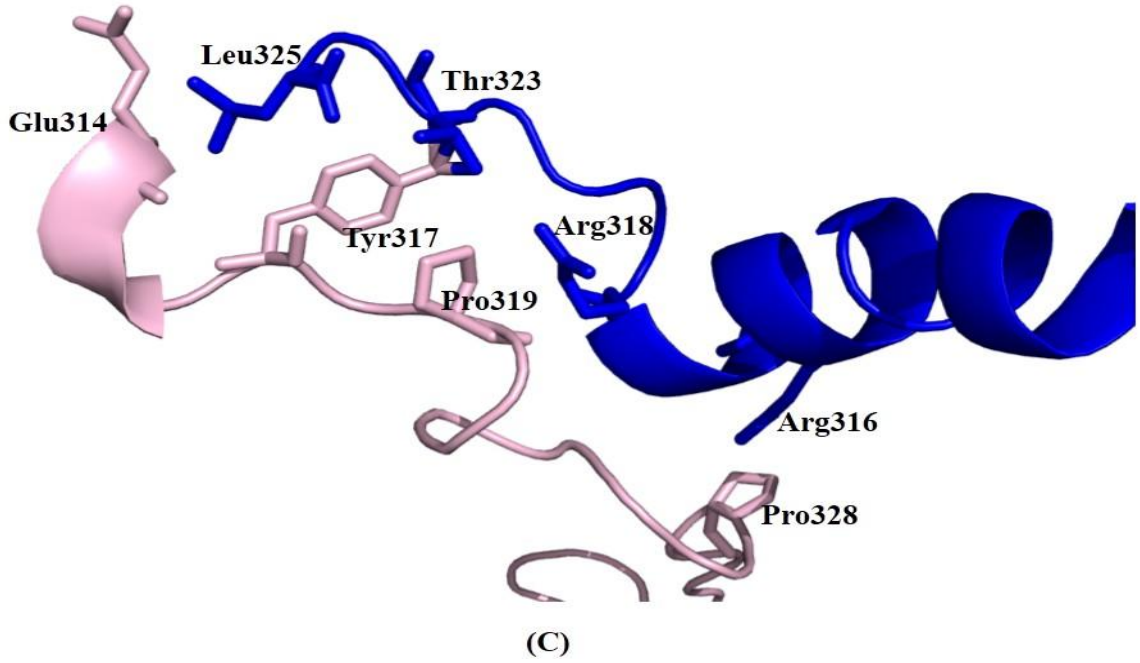
ZDock Pose	ZRank score	Molecular interactions	
		H-bonds	Non-bonded contacts
Pose 1	-83.028	**Ala309**: *Asn12*	**Tyr317**: *Pro8*; **Asp318**: *Lys4*; **Ser320**: *Pro6*
Pose 2	-78.508	**Leu313**:*Asn12*	**Tyr317**: *Pro8*; **Asp318**: *Thr7*; **Ser320**: *Pro6*
Pose 3	-76.976	**Ala309**: *Asn12*	**Asp318**: *Lys4*; **Ser320**: *Pro6*; **Glu326**: *His358*
Pose 4	-76.283	**Glu109**: *Lys4*	**Gln119**: *Pro8*; **His120**: *Asn12*; **Thr159**: *Pro6*
Pose 5	-75.260	Nil	**His180**: *Asn12*; **Asn253**: *Thr7*; **His299**: *Pro8*

3.3.9. Protein-protein docking of ERK2 and RSK1 CTKD

Figure 3.12 (A) shows the molecular docking between the ERK2 and RSK1 CTKD. **Figure 3.12 (B)** shows the important residues of ERK2 at the binding interface with RSK1. **Figure 3.12 (C)** shows the zoomed view of ERK2 residues Glu314, Tyr317, Pro319 and Pro328 interacting with Leu325, Thr323, Arg318 and Arg316 of RSK1 CTKD through non bonded contacts. **Table 3.3** shows the Z-rank scores and molecular interactions of top 5 poses.



(A) (B)



(C)

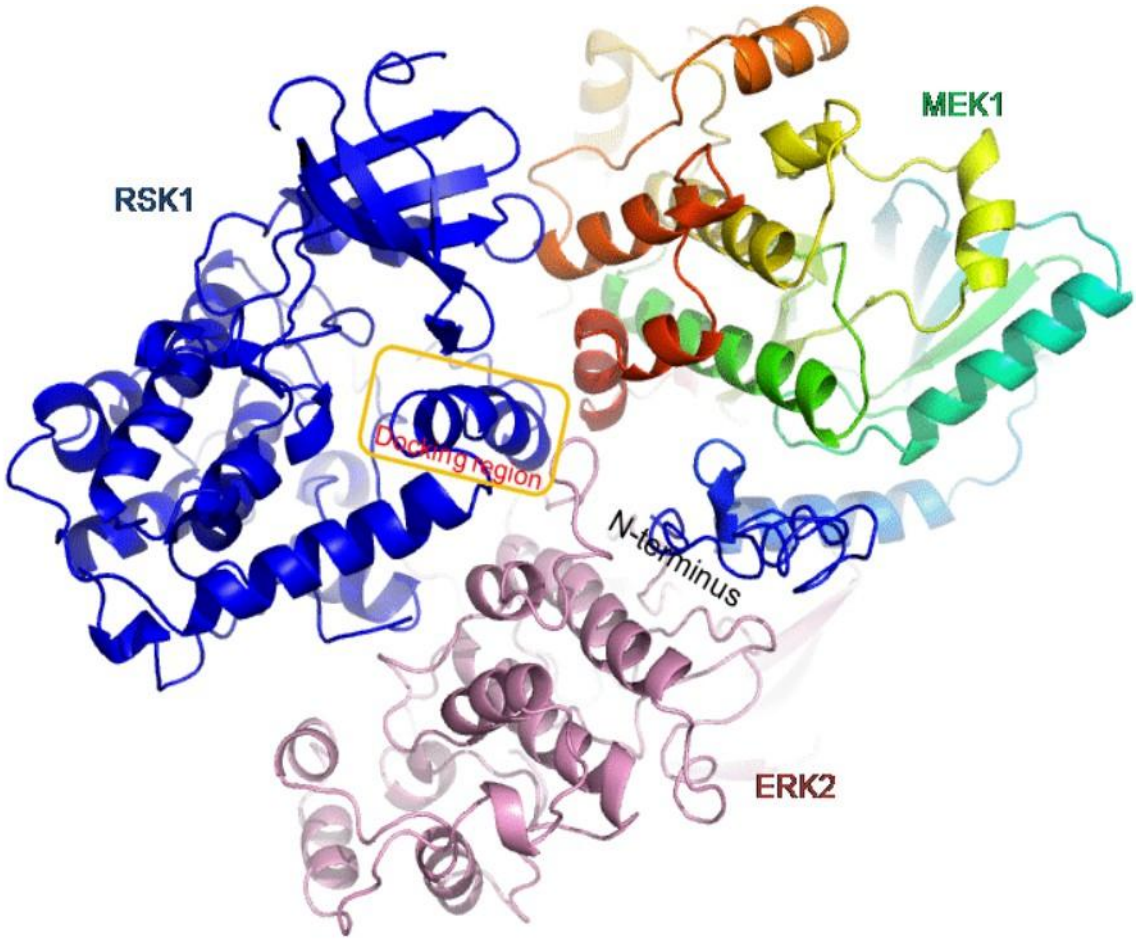
Figure 3.12: *In-silico* Docking of ERK2 on RSK1 CTKD model. (A) The ERK2 has a docking site at the C-terminus of RSK1. (B) The important residues at the binding interface has been shown. (C) The important interacting ERK2 residues namely Glu314, Tyr317, Pro319 and Pro328 are shown in pink sticks whereas RSK1 CTKD residues Leu325, Thr323, Arg318 and Arg316 are shown in blue sticks

Table 3.3: Molecular interactions between ERK2 (PDB id: 1TVO) and Modelled RSK1 C-terminal domain (PDB id: 2WNT). (Bold letters indicate **ERK2** and italics indicate Modelled *RSK1* C-terminal domain. The Pose finally selected was highlighted in Green colour.

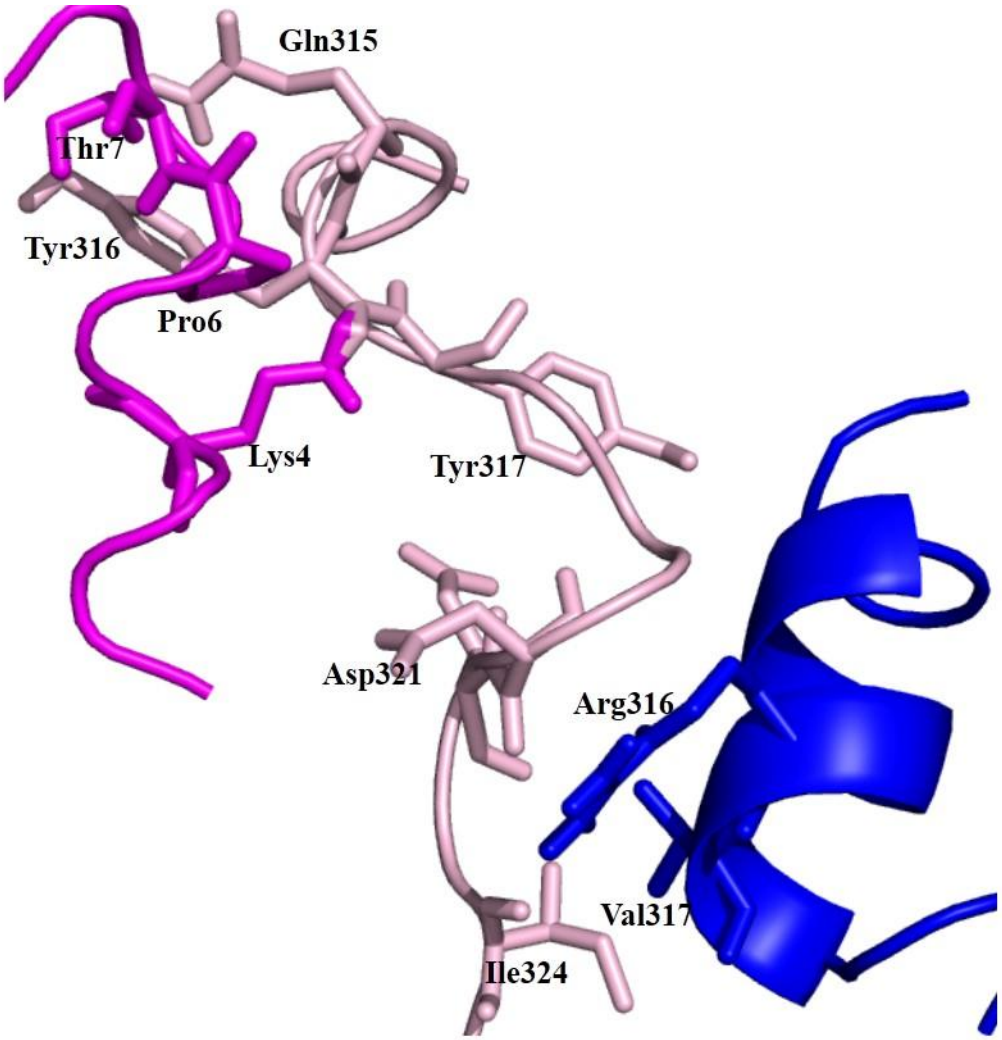
ZDock Pose	ZRank score	Molecular interactions	
		H-bonds	Non-bonded contacts
Pose 1	-80.543	**Tyr128**: *Lys177*	**Asp318**: *Leu325*; **Asp321**:*Thr323*; **Glu326**: *Val317*
Pose 2	-68.178	**Ala325**: *Arg316*	**Tyr317**: *Thr323*; **Asp318**: *Pro164*; **Ile324**: *Val317*
Pose 3	-67.066	**Tyr317**: *Ser322*	**Glu314**: *Arg315*; **Asp318**: *Arg44*; **Asp321**: *Lys41*
Pose 4	-64.586	**His125**: *Arg316*	**Tyr316**: *Val317*; **Asp318**: *Pro164*; **Asp321**: *Ser322*
Pose 5	-63.274	**Asp124**: *Arg316*	**His125**: *Leu312*; **Thr159**: *Val16*; **Thr160**: *Ser18*

3.3.10. Binding between RSK1 and MEK1-ERK2 complex

Figure 3.13 (A) shows the RSK1 docked on to the hetero dimer complex of MEK1 and ERK2. Aspartic acid at positions 318 and 321 of ERK2 play an important role in making atomic contacts with the MEK1 and RSKs. ERK2 residues from 314-329 of ERK2 has the necessary amino acid residues which are required for docking to both MEK1 and RSKs. From this trimeric complex, the MEK1 dissociates from the ERK2 and RSK1 heterodimer, making way for RSK1 to translocate to the nucleus. **Figure 3.13 (B)** shows the zoomed view of the trimeric interface of MEK1, ERK2 and RSK1 CTKD highlighting the important residues in the form of sticks



(A)



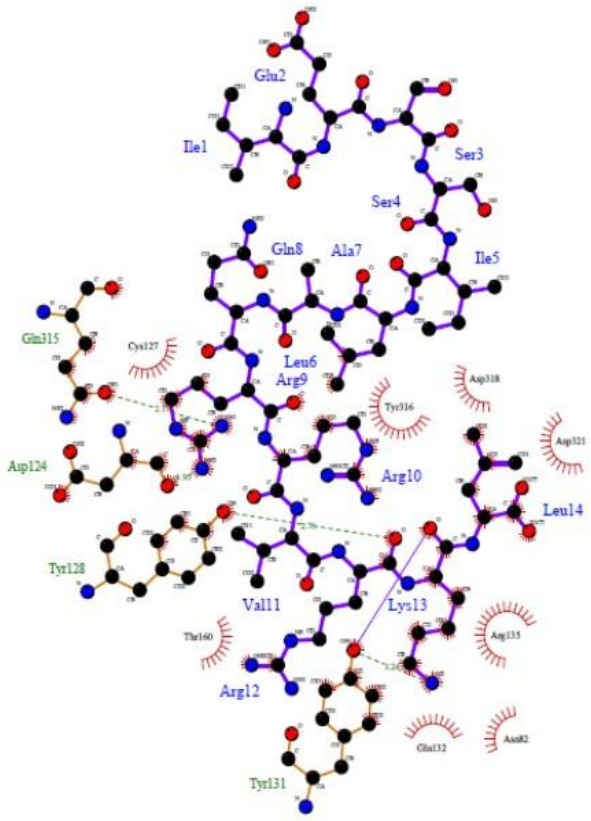
(B)

Figure 3.13: *In-silico* docking of RSK1 CTKD model on complex of MEK1-ERK2. (A) The N-terminus of MEK1 and C-terminus of RSK1 are interacting with ERK2. (ERK2 is coloured in pink, MEK1 in rainbow and RSK1 is coloured in blue). (B) The important interacting residues of MEK1, ERK2 and RSK1 CTKD at the binding interface (MEK1 is coloured in magenta, ERK2 is coloured in pink and RSK1 CTKD is coloured in blue).

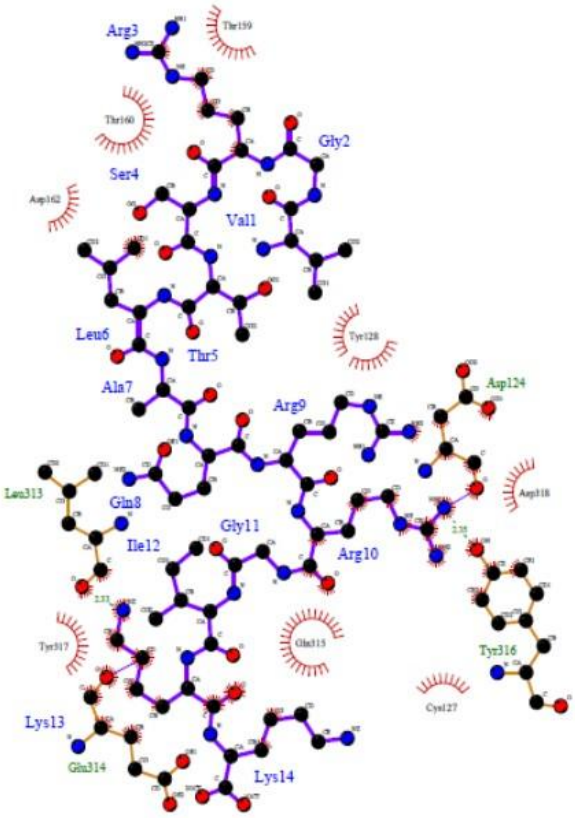
3.3.11. Binding between ERK2 and RSK1/2/3/4 peptides

The 14 amino acid sequences of all RSK isoforms (1-4) flanking the ERK binding motif 'NH_2-LAQRR-COOH' were evaluated on the basis of secondary structure, hydrophobic index, flexibility and accessibility for interactions with ERK2 residues. For homology modelling, >30 % sequence identity is a prerequisite. Our initial search results

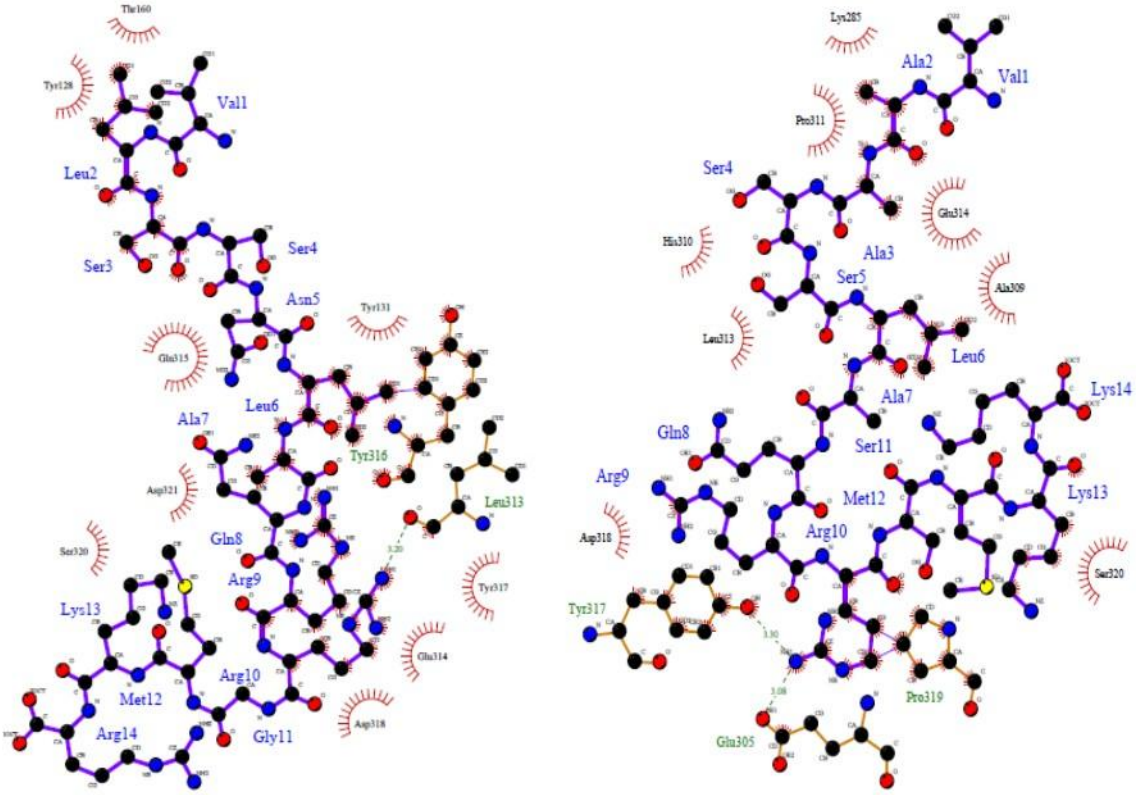
revealed relatively lower homology for the ERK Docking sequences within the C-terminal region of all four RSK isoforms. Owing to lower homology, Jpred3 was used to build these structures, which can correctly predict the secondary structures up to 81.5 % efficiency [263]. Jpred3 employs Jnet algorithm which is based on amino acid composition and similarity to already known secondary structures. Since *ab-initio* modelling has greater efficiency in building tertiary models for smaller peptides, it is therefore used to build these 14 amino acid sequences of all four RSK isoforms and the final poses were filtered on the basis of their minimum energy potential [264, 265]. These peptides were built into α-helical structures. All the RSK isoforms at the C-terminal tail have conserved "NH_2-LAQRR-COOH" amino acids. These conserved amino acids along with its flanking sequences on both ends, are necessary for the critical interactions with the residues of ERK2 protein. The sequences especially which are flanking the conserved "NH_2-LAQRR-COOH" motif have slight variations among all the four isoforms and hence each isoform of RSK interacts with ERK2 in different orientations. These differences in orientations play an important role in binding affinities and differential functionalities of the RSKs. **Figure 3. 14 (A-D)** shows the differences in the weak intermolecular interactions observed between ERK2-RSKs complexes as seen through LIGPLOT. This infers that the differences in amino acid sequences flanking the conserved "LAQRR" motif are responsible for the change in the conformation and binding affinities. So, within the limits of macromolecular docking, the important hydrophobic interactions and hydrogen bonding between all the four ERK2-RSK1/2/3/4 complexes had been predicted and analysed through LIGPLOT. **Table 3.4-3.7** shows the Z-rank scores and molecular interactions of top 5 poses.


Glu2
Ile1
Ser3
Ser4
Gln8
Ala7
Ile5
Gln315
Cys127
Leu6
Arg9
Tyr316
Asp318
Asp321
Asp124
Arg10
Leu14
Tyr128
Val11
Lys13
Thr160
Arg135
Arg12
Glu132
Asn82
Tyr131

(A)


Arg3
Thr159
Thr160
Gly2
Ser4
Asp162
Val1
Leu6
Thr5
Tyr128
Ala7
Asp124
Arg9
Leu313
Gln8
Asp118
Gly11
Ile12
Arg10
Tyr317
Gln315
Tyr316
Cys127
Lys13
Gln314
Lys14

(B)



(C) **(D)**

Figure 3.14: Binding association between ERK2 and C-terminus peptides of (A) RSK1 (B) RSK2 (C) RSK3 (D) RSK4.

Table 3.4: Molecular interactions between ERK2 (PDB id: 1TVO) and Modelled RSK1 C-terminal tail (Bold letters indicate **ERK2** and italics indicate *RSK1* peptide. The Pose finally selected was highlighted in Green colour).

ZDock Pose	ZRank score	Molecular interactions	
		H-bonds	Non-bonded contacts
Pose 1	-61.913	**Gln315**: *Arg9*	**Tyr316**: *Arg10*; **Asp318**: *Leu14*; **Asp321**: *Leu14*
Pose 2	-59.833	Nil	**Tyr316**: *Arg10*; **Asp318**: *Val11*; **Pro319**: *Gln8*
Pose 3	-59.084	**Gln315**: *Ala7*	**Tyr316**: *Glu2*; **Tyr317**: *Arg10*; **Asp318**: *Arg9*
Pose 4	-56.466	**Asp318**: *Arg10*	**Pro311**: *Lys13*; **Pro319**: *Leu6*; **Ser320**: *Ala7*
Pose 5	-56.310	**Asp318**: *Arg10*	**Tyr317**: *Arg10*; **Pro319**: *Leu6*; **Ser320**: *Ala7*

Table 3.5: Molecular interactions between ERK2 (PDB id: 1TVO) and Modelled RSK2 C-terminal tail (Bold letters indicate **ERK2** and italics indicate *RSK2* peptide. The Pose finally selected was highlighted in Green colour).

ZDock Pose	ZRank score	Molecular interactions	
		H-bonds	Non-bonded contacts
Pose 1	-63.733	**Tyr316**: *Arg10*	**Thr160**: *Arg3*; **Tyr317**: *Lys13*; **Asp318**: *Arg10*
Pose 2	-61.904	**Gln315**: *Arg9*	**Tyr316**: *Arg9*; **Tyr317**: *Leu6;* **Asp318**:*Arg9*
Pose 3	-61.160	**Tyr316**: *Arg9*	**Tyr317**: *Leu6*; **Asp318**: *Arg9*; **Asp321**: *Lys13*
Pose 4	-60.397	Nil	**Glu314**: *Arg10*; **Tyr317**: *Gln8;* **Asp318**: *Ile12*
Pose 5	-58.622	Nil	**Tyr317**: *Leu6*; **Pro319**: *Arg10*; **Ser320**: *Lys13*

Table 3.6: Molecular interactions between ERK2 (PDB id: 1TVO) and Modelled RSK3 C-terminal tail (Bold letters indicate **ERK2** and italics indicate *RSK3* peptide. The Pose finally selected was highlighted in Green colour).

ZDock Pose	ZRank score	Molecular interactions	
		H-bonds	Non-bonded contacts
Pose 1	-62.412	**Asp318**: *Arg10*	**Glu314**: *Arg14*; **Tyr316**: *Arg10*; **Ser320**: *Leu6*
Pose 2	-62.335	**Leu313**: *Arg10*	**Tyr128**: *Leu2*; **Asp318**: *Arg10*; **Asp321**: *Gln8*
Pose 3	-60.919	**Pro319**: *Arg10*	**Glu314**: *Leu2*; **Tyr317**: *Leu6*; **Ser320**: *Arg14*
Pose 4	-60.79	**Glu305**: *Arg10*	**Asp318**: *Lys13*; **Pro319**: *Arg10*; **Ser320**: *Arg14*
Pose 5	-60.261	**Tyr316:** *Arg10*	**Lys285**: *Arg9*; **Tyr317**: *Leu6*; **Asp318**: *Arg10*

Table 3.7: Molecular interactions between ERK2 (PDB id: 1TVO) and Modelled RSK4 C-terminal tail (Bold letters indicate **ERK2** and italics indicate *RSK4* peptide. The Pose finally selected was highlighted in Green colour).

ZDock Pose	ZRank score	Molecular interactions	
		H-bonds	Non-bonded contacts
Pose 1	-62.397	Nil	**Glu305**: *Gln8*; **Tyr317**: *Ala7*; **Ser320**: *Lys14*
Pose 2	-61.404	**Glu305**: *Arg10*	**Asp318**: *Arg9*; **Pro319**: *Arg10*; **Ser320**: *Lys13*
Pose 3	-60.759	Nil	**Glu314**: *Arg10*; **Gln315**: *Arg9*; **Asp318**: *Lys13*
Pose 4	-60.725	**Tyr317**: *Gln8*	**Pro311**: *Ser4*; **Glu314**: *Arg10*; **Tyr317**: *Ser11*
Pose 5	-57.519	**Asp318**: *Lys14*	**Pro311**: *Ala3*; **Glu314**: *Arg10*; **Tyr317**: *Gln8*

3.4. Conclusion

ERK2 full length protein was purified form *E. coli* bacterial system and the purified protein was used to study protein-protein interactions (PPIs) with 65 amino acids C-terminal tail of RSK3. No interactions were observed using GST pull down assay. However, we have then used the 14 amino acid long peptide to study interactions with ERK2. The interaction study performed using Isothermal Titration Calorimetry (ITC), we have observed binding between ERK2 and RSK3. This may be due to the fact that pull down assay is less sensitive, whereas ITC is very sensitive and can detect interactions up to nM range. It has been mentioned that ERK2 binds with its peptide substrates with an affinity as low as 100 times to that of full length substrates [259, 260]. The relative binding affinity K_d recorded is moderately low between ERK2 and RSK3 C-terminal tail peptide, which is in accordance with the earlier reports. Furthermore, ERK2 is not phosphorylated as it is purified from bacterial system and this fact, can also be the reason for the moderately low binding.

In-silico studies had also been carried out to unravel the macromolecular interactions at atomic level of the MAPK pathway particularly emphasizing on ERK2. In the MAPK pathway, ERK2 interacts with MEK1 and all the four RSK isoforms. Since the ERK2 docking regions were missing within the available structures for MEK1 and RSK1, the regions were built using I-tasser. Through macromolecular docking the critical Protein Protein Interactions (PPIs) between MEK1 and ERK2 and similarly between ERK2 and all the four RSK isoforms have been identified. The studies revealed that ERK2 residues from 314-329 make prominent atomic contacts with both MEK1 and RSKs, suggesting the interaction is mediated in a two-pronged manner. However, MEK1 cannot transduce into nucleus owing to steric hindrances with in the MEK-ERK-RSK complex, thus MEK is retained in the cytoplasm, whereas ERK-RSK enters the nucleus. The amino acid sequences flanking the ERK binding motif play an important role in determining the binding affinities

and orientations thus playing an important role in differential functionalities within the cell. This is essential for RSK isoforms to choose a subset of transcription factors to activate, depending upon the cell requirement. These studies also showed the importance of steric hindrance in MAPK pathway, playing a major role in dissociation of upstream kinases from the effective transducing signalling complex. Studying these PPIs at atomic level among the MAPK pathway members help researchers to design and develop new drugs which may be potentially therapeutic in nature.

Chapter 4:

Studies of folding patterns of Ribosomal

S6 Kinase 1

4.1 Introduction

Ribosomal S6 Kinases (RSKs) belong to the family of Ser/Thr kinases and constitutes the bottom end of the classical Mitogen-Activated Protein Kinase (MAPK) pathway [41]. Ribosomal S6 Kinase 1 protein is encoded by RPS6KA1 gene, present on the p arm of Chromosome 1 in the human genome. The RSK1 protein is of 735 amino acids long with the molecular weight of 82722.9 Da, and pI 7.68. RSK1 is expressed in different cell types, however higher expression levels are seen in kidneys, lungs and pancreas [266]. RSK1 protein expression is highly elevated in most of the cancer types, particularly prostate cancer [112]. RSK1 has two functional domains. The N-terminal Kinase Domain (NTKD) that comprises 260 residues from 62-321 amino acids, is homologous to AGC family of kinases. The C-terminal Kinase domain (CTKD) has 258 amino acids from 418-675. It is homologous to Calcium/calmodulin-dependent kinases (CaMK) family by sequence [79, 81]. These two functional kinase domains are joined by highly conserved linker region of 96 residues from 322-417 amino acids. The complete activation requires multiple phosphorylation steps including auto-phosphorylation [83, 85-87]. The activated ERK1/2, activates RSK1 by binding to its C-terminal tail which houses the D-domain (NH_2-Leu-Ala-Gln-Arg-Arg-COOH), i.e. ERK binding motif [86]. Upon activation, the CTKD of RSK1 gets auto-phosphorylated at Serine 380 position, and creates a docking site for PDK1. The PDK1 then phosphorylates Serine at 221 position thereby activating the NTKD of RSK1 [42]. The NTKD further phosphorylates Serine at 732 position, which weakens the complex interaction with ERK and hence, the complex dissociates and now the fully functional RSK1 has many substrates to activate both in the cytoplasm and in the nucleus.

Upon complete activation, RSK1 has a number of substrates to activate such as SRF, ERα, CREB, FOS, JuN, STAT, NFκB,TIF1A TSC1 and 2, CDKN1B, CARHSP1,

ESR1, MAPT and RPTOR [45-48, 51, 267-271]. The full length structure of RSK1 will not only help researchers to understand the molecular mechanism of MAPK pathway, but also help in designing new small molecule inhibitors [24].

Immobilized Metal ion Affinity Chromatography (IMAC) purification with 6-HIS tag is one of the most common affinity chromatography techniques used by protein biochemists [272]. Protein folding is a complex phenomenon which occurs inside the cell in such a way that hydrophobic amino acids are buried deep inside the three dimensional core whereas hydrophilic residues are on the surface [273]. Some proteins have highly hydrophobic tails at either or both the ends. When such proteins fold into native conformation it burry its hydrophobic ends (tails) deep into the core. These proteins give poor purification yields when purified with 6 HIS tag as it may also get fully or partially buried along with the hydrophobic tail thereby decreasing the surface availability of the tag. To overcome this, a dual affinity tag may be used to increase the protein yields. The best way to incorporate an additional sequence into the Open Reading Frame (ORF) is site-directed insertion which is very similar to Site-Directed Mutagenesis (SDM) [274]. However site-directed insertion has a limitation on number of nucleotides being incorporated in a single step. Hence, the best possible additional tag can be incorporated by site directed insertion would be 6-HIS tag. This dual (N and C-terminal) 6 HIS tag fusion protein will certainly have larger purification yields. This technique was successfully used in making dual 6 HIS tag RSK1 full length for its purification. This chapter also includes the purification of RSK1 CTKD.

4.2. Materials and methods

4.2.1. Cloning and Ligation of RSK1 full length

Ribosomal S6 Kinase 1 (Accession No; NM_002953) full length comprising of 735 amino acids was cloned into 6 HIS based pRSET-A vector between Xho1 and EcoR1 restriction enzyme sites. The cDNA sequence of RSK1 was downloaded from http://www.ncbi.nlm.nih.gov/nucleotide/ and primers were designed by keeping Xho1 in the forward primer and EcoR1 in the reverse primers which are as follows, Forward primer: 5' -GTCCTCGAGGAGAACCTGTACTTTCAGGGTATGCCGCTCGCCCAGCTC- 3'

Reverse primer: 5' -GTCGAATTCCTATTACAGGGTGGTGGATGGCAACTT-3'

The RSK1 full length was amplified by PCR (DNA engine tetrad 2, Bio-Rad) using Phusion High-Fidelity DNA Polymerase (Finnzymes). The PCR reaction mixture has 2 µl of 100 ng/µl template cDNA of RSK1, 1 µl each of 10 µM forward and reverse primers, 2 µl of 10 mM of dNTP stock, 10 µl of 5X Phusion HF buffer and 1 µl of Phusion Polymerase. The total volume was adjusted to 50 µl with distilled water. The PCR program had one total denaturation step at 95 °C for 5 minutes and a 30 cycle PCR was programmed as follows; denaturation at 95 °C for 30 seconds, primer annealing at 65 °C for 30 seconds and polymerization at 72 °C for 2 minutes. The amplification was verified by loading the PCR product on ethidium bromide (0.4 ug/ml) containing 1 % Agarose gel and visualized under a UV Gel doc (UVP bioimaging systems).

The PCR product corresponding to ~ 2.3 kb of RSK1 full length was excised from the agarose gel and recovered by gel extraction kit (Qiagen). It was doubly digested with Fast digest Xho1 and EcoR1 restriction enzymes (Fermentas) at 37 °C for 2 hours in a water bath. Similarly pRSET-A vector was also digested with Fast digest Xho1 and EcoR1 restriction enzymes. The doubly digested RSK1 PCR product and pRSET-A vector were individually recovered by PCR Purification Kit (Qiagen). Ligation was performed between RSK1 full length and pRSET-A vector using Quick Ligation kit (New England Biolabs) at room temperature for 15 minutes. The ligated mixture was then added into a vial of DH5α

competent cells for transformation and plated onto a LB agar plate containing 100 ug/ml Ampicillin. Following 16 hours of incubation at 37 °C a single isolated colony was picked and inoculated in 10 ml of Luria Broth containing 10 µl of 100 mg/ml Ampicillin and incubated in a shaker incubator (SIF-6000R, Medline scientific) at 37 °C, for 8 hours. Later the culture was pelleted down by centrifugation at 5000 rpm, 4 °C for 15 minutes. RSK1 recombinant plasmid was isolated from the pellet and treated with Xho1 and EcoR1. The digested product was confirmed as ~ 2.3 kb by running in 1 % agarose gel. This was further validated by DNA sequencing (Applied Biosystems).

4.2.2. Site-directed insertion of RSK1 full length

An additional 6 HIS tag nucleotide sequence was incorporated at the C-terminal end of RSK1 full length using following primers. Forward Primer:

5'-CCATCCACCACCCTG**CACCACCACCACCACCAC**TAATAGGAATTCGAAGC-3'

Reverse Primer:

5'-GCTTCGAATTCCTATTA**GTGGTGGTGGTGGTGGTG**CAGGGTGGTGGATGG-3'

(The 18 nucleotide sequence in bold translates for 6 HIS residues)

A gradient PCR between 50 °C and 55 °C was performed using Phusion High-Fidelity DNA Polymerase (Finnzymes) with an elongation time of 3 minutes in a tetrad PCR system (DNA engine tetrad 2, Bio-Rad). After 17 cycles, the PCR products were loaded on 1 % agarose gel, and the PCR amplification was observed at 51.7 °C. This amplified product was subjected to Dpn1 digestion to remove methylated DNA template. For Dpn1 digestion 7 µl of PCR amplified product at 51.7 °C, 2 µl of 10X Fast Digest Buffer (Fermentas) and 0.5 µl of Fast Digest Dpn1 (Fermentas) was added, and total reaction volume was made up to 20 µl with distilled water. The Reaction was carried out at 37 °C for 20 minutes, and further

transformed into DH5α competent cells. Plasmid isolation was performed as per Qiagen protocol. The insertion of 18 nucleotides coding for 6 HIS residues was confirmed by DNA sequencing (Applied Biosystems) using 2 µl of 150 ng/µl recombinant construct and 1 µl of 5 uM T7 terminator sequencing primer. The sequencing data was analysed using BioEdit software.

4.2.3. Expression and purification of dual 6 HIS-RSK1 full length

The RSK1/pRSET-A with 6 HIS tag at N and C-terminal construct was transformed into BL21 (DE3) pLysE competent cells and plated onto LB agar plate with 100 µg/ml Ampicillin and 34 µg/ml chloramphenicol. A single isolated colony was inoculated in 10 ml of Luria Broth along with 10 µl each of 100 mg/ml Ampicillin and 34 mg /ml chloramphenicol, and was grown in a shaker incubator at 37 °C, 200 rpm. After 8 hours of incubation, the starting culture was diluted in 1 litre Luria Broth along with 100 µg/ml Ampicillin and 34 µg/ml chloramphenicol, and further incubated in a shaker incubator (Multitron standard, Infors HT) at 37 °C, 250 rpm, till the optical density (OD) at λ = 600 reaches between 0.6-0.8. The culture was then induced with 0.3 mM isopropyl-β-D-thiogalactoside (IPTG) and incubated at 20 °C with 250 rpm. Following 16 hours of induction the bacterial culture was harvested by centrifugation at 4 °C, 5000 rpm for 20 minutes.

One litre of bacterial pellet was resuspended in 10 ml of buffer C (50 mM Tris-HCl pH 8.5, 500 mM NaCl, 4 mM $MgCl_2$), additionally containing protease inhibitor cocktail in 1:1000 dilution (Sigma-Aldrich). Sonication was performed for cell wall disruption by digital sonifier (Bransson). 5 cycles of sonication was performed, each of one minute containing a repeated pulse of 4 seconds with an intermittent break of 2 seconds. After each cycle of sonication one minute gap was given and the sample was maintained on

ice throughout the process. This sonicated sample was centrifuged in a JA-25.50 sorvall rotor at 18,000 rpm, 4 °C for 45 minutes (Avanti J-26S XPI, Beckman coulter). The supernatant was used for IMAC purification of dual 6 HIS-RSK1 using Ni-NTA resin (Qiagen).

The Ni-NTA column (Qiagen) was equilibrated with buffer C, and binding was performed by adding the supernatant containing RSK1 full length protein by gravity flow method. This protein bound Ni-NTA column was washed with 20 column volume wash buffer containing buffer C with 10 mM and 20 mM Imidazole. Further 4 more column volume washes each with buffer C containing 40 mM and 70 mM Imidazole respectively was given and collected to check if there is any loss of RSK1 protein. After these stringent washes, RSK1 full length was eluted in buffer C with 150 mM and 300 mM of Imidazole. These elution fractions were analysed on 8 % SDS-PAGE gel and visualized by coomassie staining.

4.2.4. Desalting of dual 6 HIS-RSK1 full length

To remove the excess salts from IMAC purification, desalting was performed using Sephadex G-25 M (GE) pre-packed columns. The elutions containing pure RSK1 protein were pooled and concentrated upto 2.5 ml and loaded on to the desalting column, pre-equilibrated with 10 column volumes of Buffer D (10 mM Tris-HCl pH 8.5, 100 mM NaCl, and 2 mM $MgCl_2$). The RSK1 was finally eluted in buffer D and purified protein was further used to perform *in-vitro* experiments.

4.2.5. Peptide mass fingerprinting by mass spectrometry

The 88 kDa band of purified RSK1 full length was excised from the SDS-PAGE gel and chopped into tiny pieces. Trypsin in-gel digestion was performed as mentioned by Shevchenko et al [250]. The tryptic digests were then extracted and

reconstituted in 50 % Acetonitrile containing 0.1 % TFA. The mass spectra of these tryptic digests were recorded by Ultraflex II MALDI-TOF/TOF mass spectrometer (Bruker Daltonics) and analysed using Mascot (Matrix Sciences) search engine.

4.2.6. CD Spectroscopy and Fluorescence spectroscopy of RSK1 full length

For CD Spectroscopy (Jasco J-815 CD spectrophotometer), 300 µl of 10 µM of purified RSK1 protein in Buffer D was taken in quartz cuvette, and a spectrum was recorded between λ = 250 and 190 nm. The spectrum obtained from CD spectroscopy was submitted to K2D3 online server to estimate the secondary structure [251].

Further Fluorescence spectroscopy **(**Horiba scientific) was performed using 200 µl of 1 µM of RSK1 protein in buffer D. The emission spectrum was recorded between λ = 310 nm to 400 nm for an excitation wavelength (λ) of 295 nm.

Thermal Denaturation by florescence spectroscopy was performed between 10 °C to 80 °C with an increment of 2 °C to study the thermal stability of RSK1 protein using 200 µl of 1 µM protein. In all, 36 spectra were collected between 10 °C to 80 °C. Similar spectra were recorded using buffer D for blank correction. From these blank corrected values, average emission wavelength was calculated as the ratio of the summation of product of blank corrected values at respective wavelength to the summation of blank corrected values. These values were plotted on graph by plotting fraction unfolded on Y-axis with respect to temperature on X-axis. To calculate the $T_{m,}$ the data was fitted into Linear Extrapolation graph [275].

4.2.7. Cloning of RSK1 CTKD (418-675) in pET-28 a (+)

Ribosomal S6 Kinase 1 C-Terminal Kinase Domain was cloned into 6 HIS based pET-28 a (+) vector between EcoR1 and Sal1 restriction enzyme sites. The primers were

designed by keeping EcoR1 in the forward primer and Sal1 in the reverse primer which are as follows, Forward Primer:

5'-GTCGAATTCGAGAACCTGTACTTTCAGGGTTACGTGGTAAAGGAGACAATT-3'

Reverse Primer: 5' -GTCGTCGACCTATTAGACCCATGGATGCTGCAGAAC -3'

Initial denaturation step was 95 °C and for a duration of 5 minutes, PCR cycle comprised of denaturation at 95 °C for 30 seconds, primer annealing at 65 °C for 30 seconds and the polymerase extension at 72 °C for 30 seconds and this cycle was repeated 32 times. Cloning was performed as mentioned in **4.2.1.** section. Instead of 100 µg/ml Ampicillin, 50 µg/ml kanamycin was used as pET-28 a (+) vector has kanamycin resistant gene.

4.2.8. Expression and purification of RSK1 CTKD/pET-28 a (+)

The RSK1 CTKD/pET-28 a (+) construct was transformed into Rosetta 2(DE3) competent cells and expressed as mentioned in **4.2.3.** section, using 50 µg/ml kanamycin. Since most of the RSK1 CTKD has been expressed in inclusion bodies, it is purified under denaturating conditions using 8 M Urea. One litre of bacterial pellet was resuspended in 10 ml of buffer E (50 mM Tris-HCl pH 8.0, 500 mM NaCl and 4 mM $MgCl_2$). Protease inhibitor cocktail was added in a ratio of 1:1000 by volume (Sigma-Aldrich). The cells were disrupted by sonication (Branson digital sonifier) for 5 one minute cycles separated by one minute incubation on ice, keeping the pulse on for 4 seconds and off for 2 seconds in every cycle. The sample was maintained on ice throughout the process. The cell lysate was then centrifuged for 20 minutes at 12,000 rpm and 4 °C. After separating the supernatant, the cell pellet was thoroughly washed thrice with the buffer E, containing 1 % Triton X-100, 1 % Sarkosyl and 1 % Sodium Deoxycholate. The cell pellet was given one extra wash with buffer E containing 4 M Urea, before resuspending it into 8 M Urea containing Buffer F (50

mM Tris-HCl pH 8.0, 500 mM NaCl, 4 mM $MgCl_2$ and 8 M Urea). The denatured protein was separated by centrifuging at 12,000 rpm for 20 minutes at 4 °C. The supernatant was used for purification of RSK1 CTKD by Immobilized Metal Ion Affinity Chromatography (IMAC).

Ni-NTA column (Qiagen) was pre-equilibrated with five column washes of buffer F (50 mM Tris-HCl pH 8.0, 500 mM NaCl, 4 mM $MgCl_2$ and 8 M Urea). Then the cell lysate containing denatured RSK1 CTKD was allowed to bind in gravity flow methods to the pre-equilibrated Ni-NTA beads. The 20 column washes were performed on the protein bound resin with buffer F containing 10 mM and 20 mM imidazole respectively. RSK1 CTKD was finally eluted in buffer F containing 150 mM and 300 mM imidazole.

4.2.9. Cloning of RSK1 CTKD in pMT-wb-v5 and transfection in insect cell lines

Noting the problem during purification of native protein in bacterial expression system, we decided to look for alternate expression system i.e. insect cell lines. Ribosomal S6 Kinase 1, C-Terminal Kinase Domain was cloned into pMT-wb-v5 vector between restriction enzymes Ecor1 and Xho1. Since Xho1 lies downstream of Sal1 in pET-28 a (+) vector, the RSK1 CTKD insert was excised from pET-28 a (+) vector by double restriction digestion using Ecor1 and Xho1 enzymes and cloned in pMT-wb-v5 vector.

For transfection, 2 ml of drosophila Schneider s2 insect cell lines (from Dr. Venu Gopal Lab, ACTREC) were plated into 6-well plate. Transfection reagent complex was made by mixing 5 µl X-treme GENE HP DNA Transfection Reagent, and 10 µg of RSK1 CTKD/pMT-wb-v5 plasmid DNA in 100 µl of serum free sterile medium and allowed to incubate for 30 minutes. Now, DNA complex was added into wells containing Schneider s2 cells. After gentle rocking, 6-well plate was incubated at 25 °C for 48 hours. To establish

stable cell lines, Blasticidin was added along with the fresh sterile medium at every 48 hours. Few cycles of seeding with fresh medium and Blasticidin, stable cell lines were established.

4.3. Results and discussion

4.3.1. Cloning and expression of N-terminal 6 HIS-RSK1 full length

RSK1 full length (1-735) amino acids was cloned into N-terminal 6 HIS tag based pRSET-A vector. **Figure 4.1 (A)** shows PCR amplified RSK1 product of 2250 base pairs. Ligation was performed between RSK1 and pRSET-A vector, both of which were doubly digested prior with Xho1 and EcoR1 restriction enzymes. **Figure 4.1 (B)** shows the 2250 base pair RSK1 full length insert (which additionally contains TEV protease cleavage sequence, restriction sites and stop codons) from the recombinant pRSET-A plasmid following a double restriction digestion.

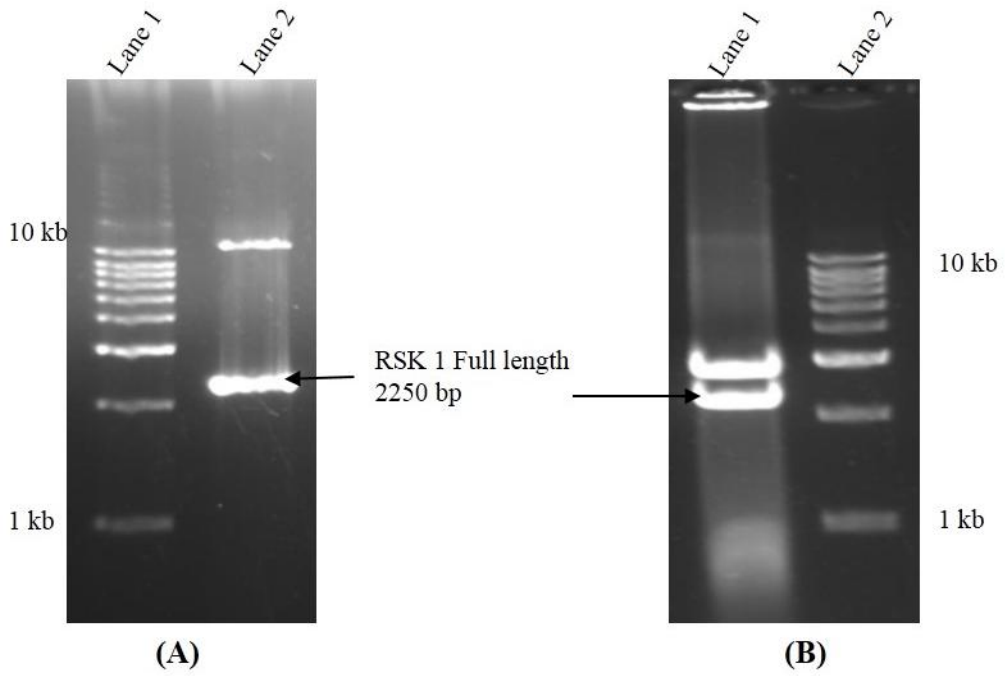


Figure 4.1: Cloning of RSK1 full length (1-735) (A) PCR amplification of RSK1 full length (Lane 1: 1 kb ladder, Lane 2: PCR product) (B) Insert release of RSK1 full length from the vector backbone following double restriction digestion. (Lane 1: RSK1 full length/pRSET-A after double restriction digestion, Lane 2: 1 kb ladder)

4.3.2. Expression and purification of 6 HIS-RSK1 full length

The Molecular weight of RSK1 (1-735) amino acids is ~ 82 kDa and the fusion protein in pRSET-A vector is ~ 87 kDa which includes the sequences of Xpress epitope, TEV cleavage sequence, and an additional 46 residues at the N terminal end. The RSK1 was expressed in BL21 (DE3) pLysE competent cells and purified by IMAC using Ni-NTA resin and analysed on 8 % SDS-PAGE gel. **Figure 4.2** shows the purification profile of RSK1 in which most of the RSK1 (up to 90 %) is lost as it is co-eluting with other bacterial contaminants as low as 40 mM and 70 mM Imidazole concentrations.

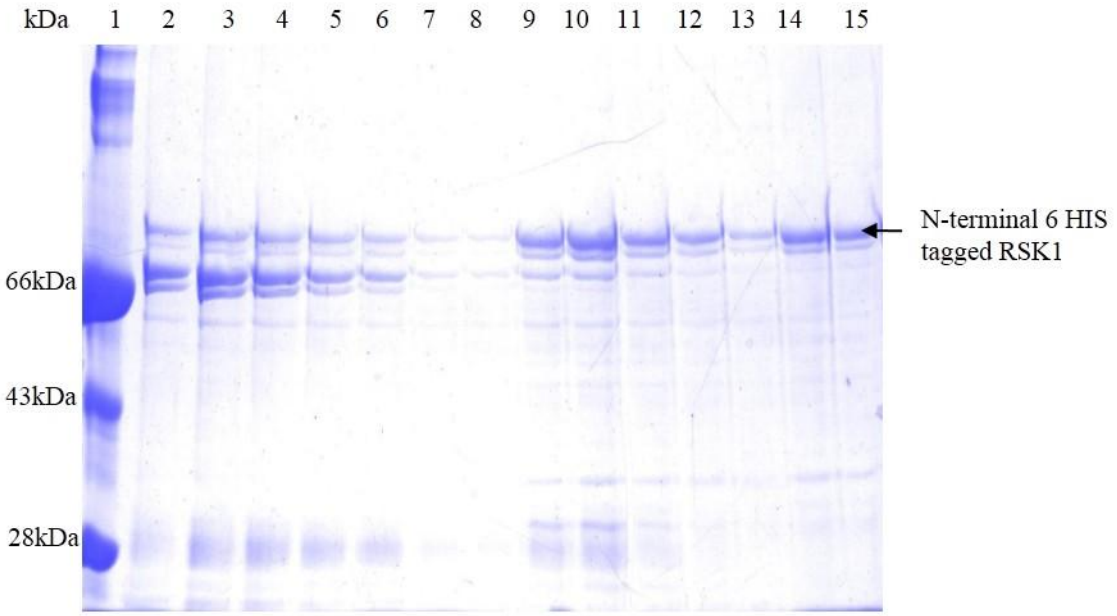


Figure 4.2: Purification of RSK1 full length. The 8 % SDS-PAGE gel shows the 87 kDa RSK1 protein with N-terminal HIS tag co-eluting with bacterial contaminants in lysis buffer containing 40 to 70 mM Imidazole. (lane 1: Protein marker, Lanes 2-8: 40 mM Imidazole elutions. Lanes 9-15: 70 mM Imidazole elutions).

4.3.3. N-terminal analysis and site directed insertion of RSK1 full length

Since the RSK1 full length protein is co-eluting at very low imidazole concentrations, the N-terminal tail of RSK1 was analysed for hydropathicity index using KYTE DOOLITTLE hydropathy plot embedded in the ProtScale tool of the ExPASy server [258].

Figure 4.3 shows the KYTE DOOLITTLE data, which suggests the stretch of first 18 amino acids of RSK1 "NH_2-MPLAQLKEPWPLMELVPL-COOH" is hydrophobic in nature with a hydrophobicity index or grand average of hydropathicity +0.39. This shows the N-terminal tail of RSK1 is hydrophobic in nature thereby decreasing the surface availability of N-terminal 6 HIS tag.

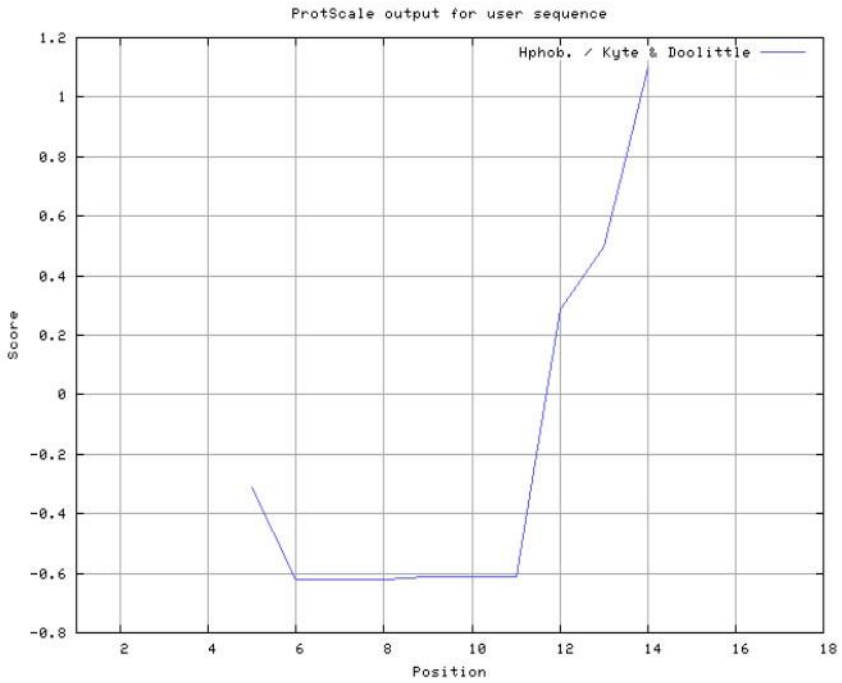


Figure 4.3: KYTE DOOLITTLE plot of N-terminal tail of RSK1. The hydropathy plot of first 18 amino acids of RSK1 shows that the N-terminal tail of RSK1 protein is hydrophobic in nature with a grand average of hydropathicity +0.39

Since the RSK1 C-terminal tail is hydrophilic in nature, an additional 18 nucleotides corresponding to 6 HIS residues were incorporated at the carboxyl end of C-terminal tail through site directed insertion of the recombinant RSK1/pRSET-A construct. A gradient PCR between 50 °C and 55 °C was performed. **Figure 4.4** shows the site directed insertion amplification at 51.7 °C. The amplified plasmid was subjected to Dpn1 digestion, and transformed into DH5α competent cells. DNA sequencing was performed to verify the insertion of 18 nucleotide. **Figure 4.5** shows sequencing data as viewed in BioEdit software

which confirms 100 % perfect insertion of 18 nucleotide sequence coding the 6 HIS residues after the C-terminal end of RSK1 full length, just before the stop codons.

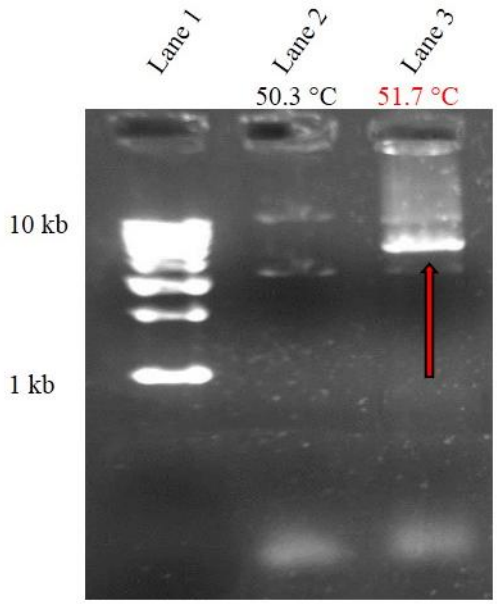


Figure 4.4: Site-directed insertion of RSK1 full length. 1 % agarose gel shows the gradient PCR products of the SDM for dual HIS tag of RSK1 full length which amplified at 51.7 °C.

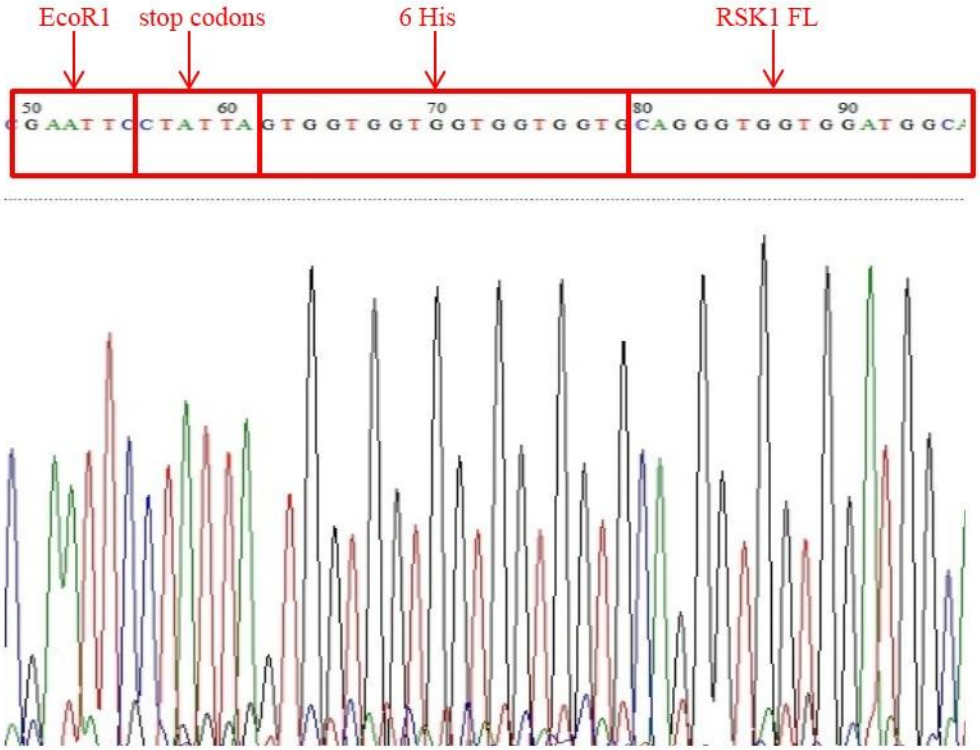


Figure 4.5: Sequencing of dual 6 HIS-RSK1 full length. Sequencing data of SDM amplified dual 6 HIS-RSK1 full length recombinant construct with T7 terminal primer shows the perfect incorporation of 18 nucleotides encoding 6 Histidine residues at the C-terminal end of protein just before the stop codons.

4.3.4. Expression and purification of dual 6 HIS-RSK1 full length

Dual 6 HIS tagged (at N and C-terminal) RSK1 (1-735) was purified from *E. coli* Bl-21(pLysS). **Figure 4.6** shows purified RSK1 full length which mostly elutes at 150 and 300 mM Imidazole as seen in lanes 5-12.

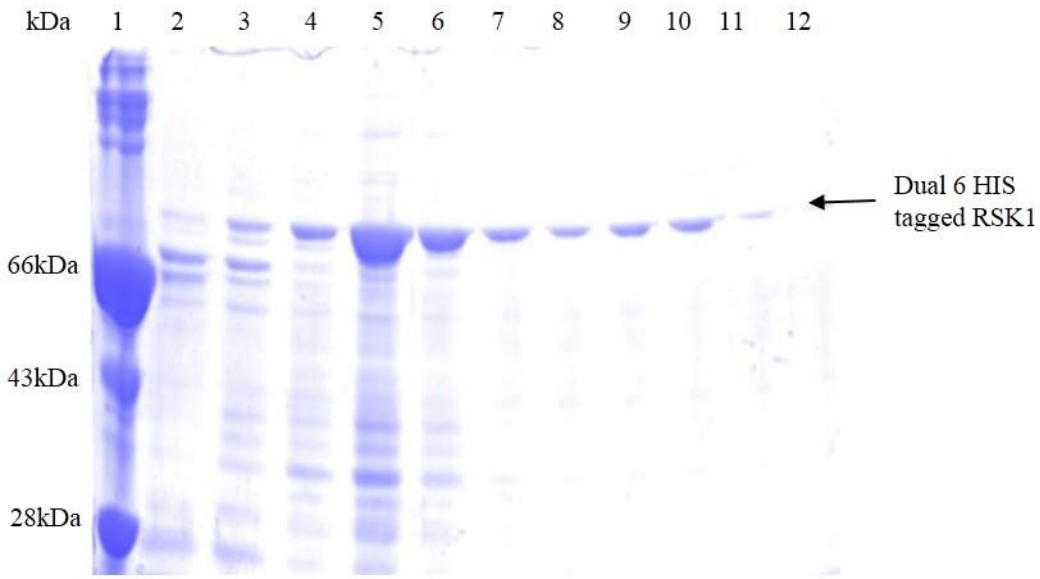


Figure 4.6: Purification of dual 6 HIS-RSK1 full length. The 8 % SDS-PAGE gel shows the 88 kDa dual 6 HIS-RSK1 full length protein with N and C-6 HIS tags eluting in lysis buffer containing 150 and 300 mM Imidazole with greater purity. (lane 1: Protein marker, Lane 2: 40 mM Imidazole elution. Lane 3: 70 mM Imidazole elution, Lanes 4-8: 150 mM Imidazole elutions, Lanes 9-12: 300 mM Imidazole elutions)

4.3.5. Detection of purified RSK1 full length protein

To check the nativity of purified protein, trypsin in-gel digestion was performed. **Figure 4.7 (A)** shows the identified extracted peptides with their m/z values as determined by MALDI-TOF/mass spectrometry. **Figure 4.7 (B)** shows the peptide mass fingerprinting by Mascot search engine using SwissProt database which confirms that the peptides are of RSK1 protein with a high score.

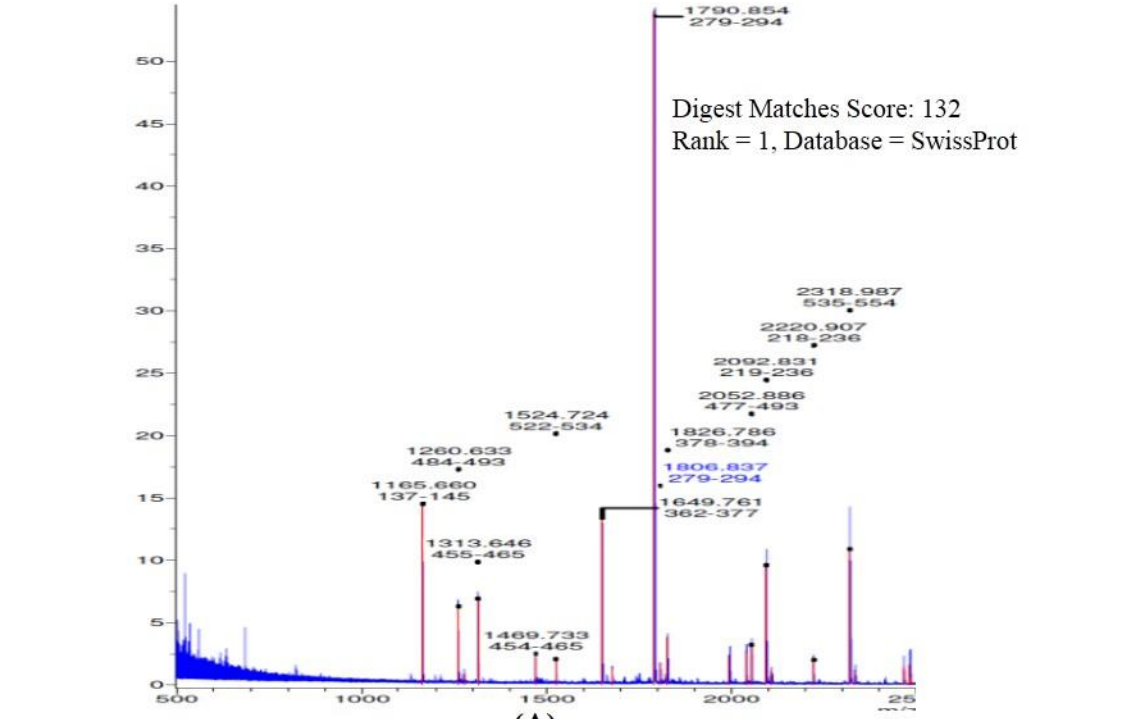


Figure 4.7: Peptide mass fingerprinting of in-gel tryptic digests of RSK1 full length. (A) MALDI-TOF/mass spectrometry data showing the identified tryptic peptides of RSK1 full length. (B) The Mascot search engine confirmed the peptides were corresponding to RSK1.

4.3.6. CD and Fluorescence spectroscopy of RSK1 full length

To analyse the secondary structure of RSK1 (1-735) protein, CD spectroscopy was performed. **Figure 4.8** shows the CD spectrum and the dips at wavelengths λ = 208 and 222 nm respectively suggest that the protein is predominantly α-helical in nature.

Further, to estimate the contents of α-helices and β-sheets, the CD spectra were submitted to K2D3 online server which predicted α–helical content of ~ 52 % and β–sheets content to be ~ 10 %.

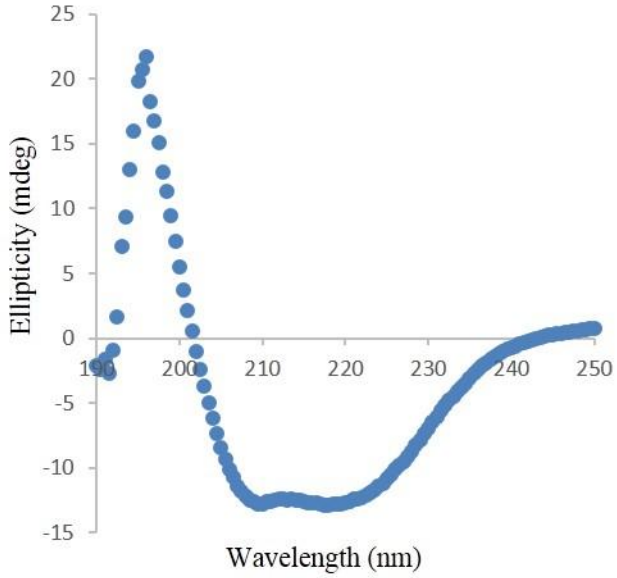


Figure 4.8: CD Spectra of RSK1 full length. The CD Spectra shows RSK1 protein is predominantly α-helical owing to dips at λ = 208 nm and λ = 222 nm. After matching with closest spectrum, the K2D3 server predicted the protein has around 52 % of α-helices and 10 % β-strands.

Furthermore, to analyse the tertiary conformation of RSK1 full length protein, Tryptophan Fluorescence spectroscopy was performed. There are 7 tryptophan (Trp) residues in RSK1 (1-735) amino acids at different positions Trp10, Trp245, Trp246, Trp326, Trp598, Trp641 and Trp674 respectively. In addition to this, there is one Trp in the N-terminal tag of pRSET-A vector. So there are total eight Trp residues in the RSK1 fusion protein. The fluorescence spectrum was recorded for an excitation wave length of λ = 295 nm and the emission spectrum was collected from λ = 310 to 400 nm. **Figure 4.9** shows that the emission maximum (λ_{max}) of RSK1 full length after blank correction is λ = 337 nm, suggesting the tryptophan residues are mostly buried inside the hydrophobic core of the protein.

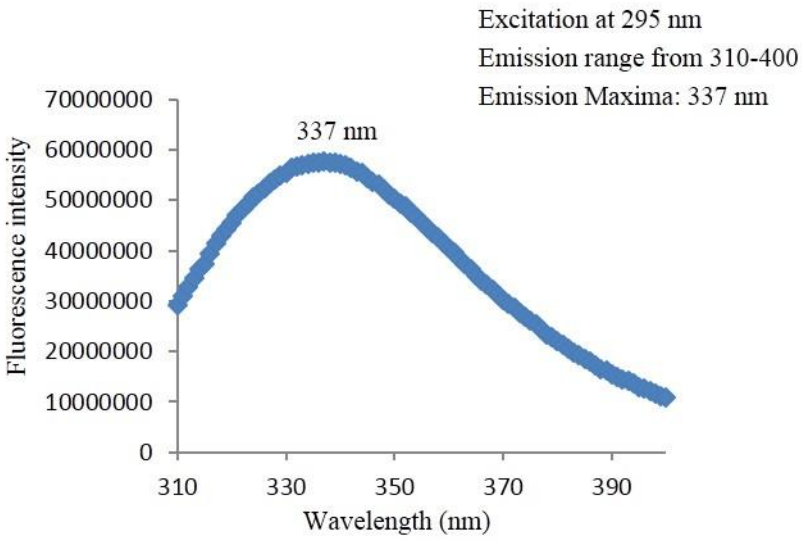
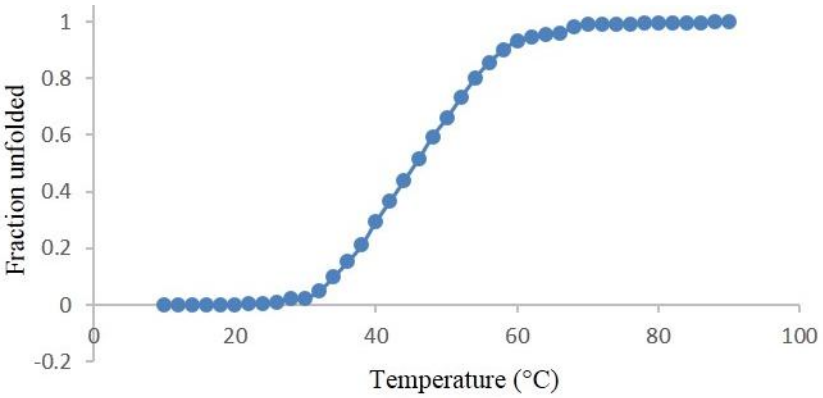


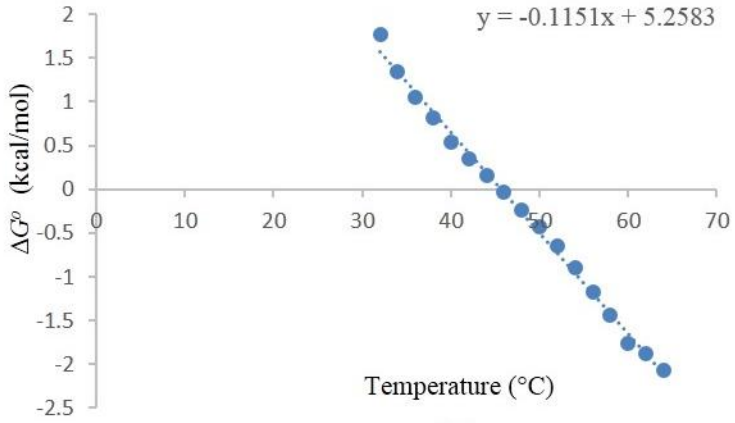
Figure 4.9: Fluorescence spectroscopy of RSK1 full length. The Fluorescence spectra of RSK1 full length recoded with a excitation wavelength λ = 295 nm shows the emission maxima (λ_{max}) is 337 nm.

4.3.7. Thermal denaturation (TD) of RSK1 full length by Fluorescence Spectroscopy

To determine the thermal stability of RSK1 full length protein, Thermal Denaturation was performed using Fluorescence spectroscopy at λ = 295 nm excitation wavelength. **Figure 4.10 (A)** shows the midpoint of the slope between the pre transition at the beginning of the graph and post transition at the end of the graph as the T_m. **Figure 4.10 (B)** shows the fitted Linear Extrapolation graph and the T_m was accurately calculated as X-axis intercept at 45.68 °C and the standard free energy was calculated as Y-axis intercept at 5.2583 kcal/mol by extrapolating the graph [275]. The Thermal denaturation study revealed that the protein is relatively less stable owing to its low T_m.



(A)



(B)

Figure 4.10: Thermal denaturation by Fluorescence spectroscopy of RSK1 full length. (A) The thermal denaturation profile of RSK1 full length depicting the fraction unfolded with function of temperature showing the pre transition and post transition states (B) The standard free energy of unfolding of RSK1 full length, graph fitted into Linear Extrapolation shows the T_m as the X-axis intercept which is 45.68 °C and the ΔG^o (H_2O) is 5.2583 kcal/mol.

4.3.8. Crystallization trials for RSK1

Figure 4.11 (A-F) shows precipitation of Ribosomal S6 Kinase 1 full length (1-735) in different buffer conditions during crystallization trials by sitting-drop vapour diffusion method using Hampton crystal screens 1 and 2.

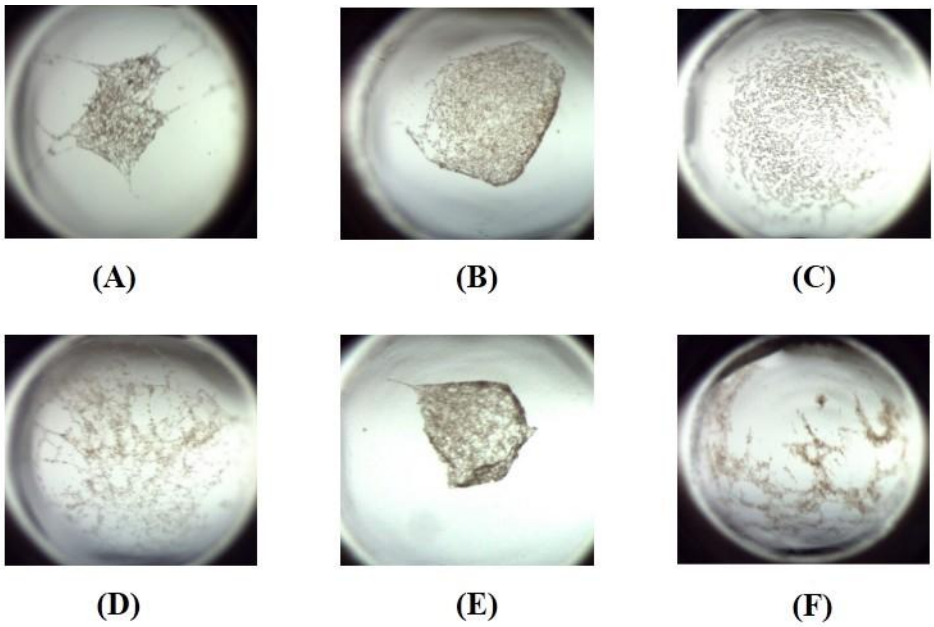


Figure 4.11: Crystallization trials for RSK1 full length. (A to F) shows the precipitated RSK1 full length during crystallization trials.

4.3.9. Cloning of RSK1 CTKD (418-675) in pET-28 a (+) and pMT-wb-v5

Figure 4.12 (A) shows Ribosomal S6 Kinase 1, C Terminal Kinase Domain (418-675) insert release of 819 base pairs from pET-28 a (+) vector digested with EcoR1 and Sal1 restriction enzymes. Similarly, Ribosomal S6 Kinase 1, C-Terminal Kinase Domain (418-675) was also cloned into pMT-wb-v5 vector between restriction enzymes Ecor1 and Xho1. **Figure 4.12 (B)** shows RSK1 CTKD insert release of 819 base pairs from pMT-wb-v5 vector back bone after double digestion with EcoR1 and Xho1 restriction enzymes. **Figure 4.13** shows the sequencing data of RSK1 CTKD/ pET-28 a (+) construct.

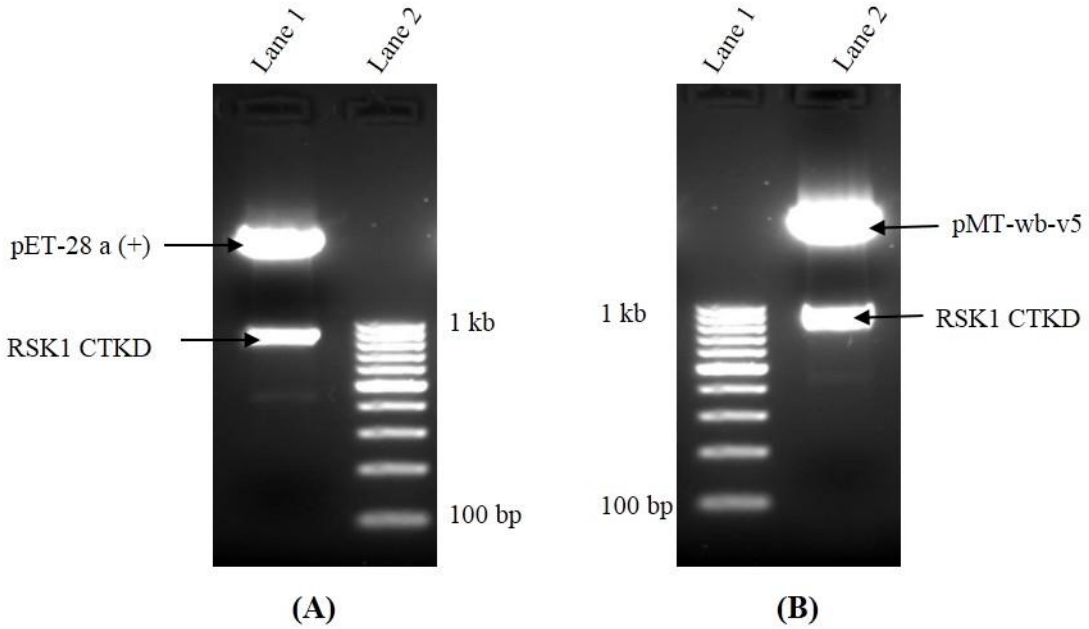


Figure 4.12: Cloning of RSK1 CTKD (418-675) (A) Insert release of RSK1 CTKD from the pET-28 a (+) vector backbone following double restriction digestion (Lane 1: RSK1 CTKD/pET-28 a (+) after double restriction digestion, Lane 2: 100 bp ladder) (B) Insert release of RSK1 CTKD from the pMT-wb-v5 vector backbone following double restriction digestion (Lane 1: 100 bp ladder, Lane 2: RSK 1 CTKD/pMT-wb-v5 vector after double restriction digestion).

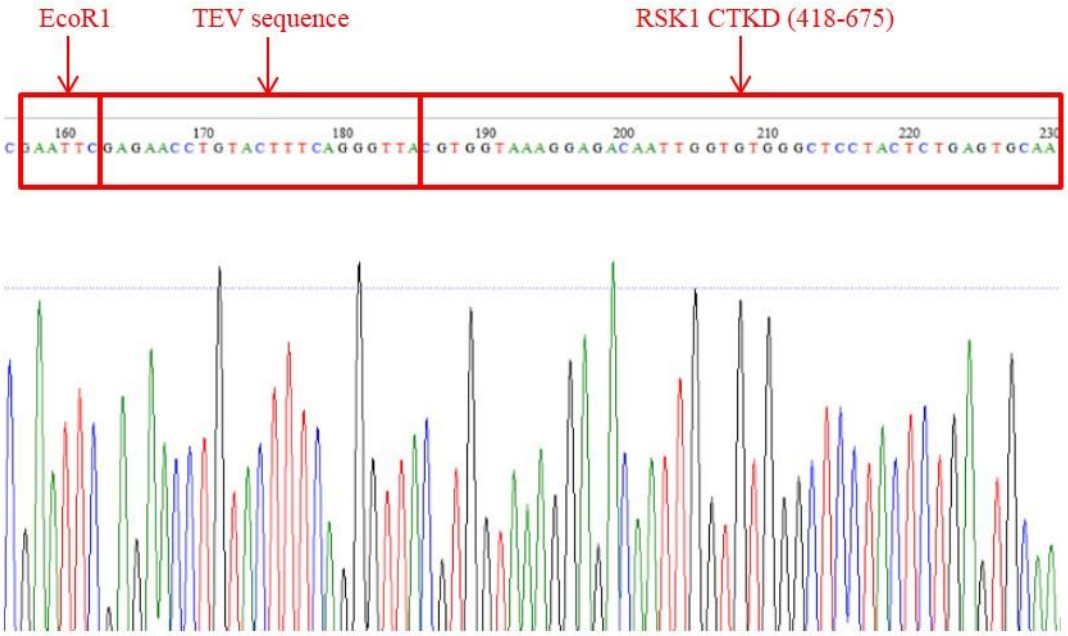


Figure 4.13: Sequencing of 6 HIS-RSK1 CTKD. Sequencing data of RSK1 CTKD (418-675) with T7 promoter primer shows the sequence in proper frame.

4.3.10. Purification of RSK1 CTKD (418-675)/pET-28 a (+)

The Molecular weight of RSK1 CTKD (418-675) is ~ 29 kDa and the fusion protein is ~ 33 kDa. The RSK1 CTKD (418-675) was expressed in Rosetta 2 (DE3) competent cells and purified by IMAC using Ni-NTA column. Since RSK1 CTKD (418-675) accumulates in inclusion bodies, it is purified under denaturating conditions using 8 M Urea. After separating the supernatant containing soluble fraction, the bacterial inclusion pellet was thoroughly washed thrice with buffer E, additionally containing 1 % Triton X-100, 1 % Sarkosyl and 1 % Sodium Deoxycholate. Following these washes, the pellet was solubilized with 8 M Urea containing buffer F, and RSK1 CTKD was purified under denaturating conditions using Ni-NTA resin. **Figure 4.14** shows purified RSK1 CTKD from inclusion bodies on a 12 % SDS-PAGE gel.

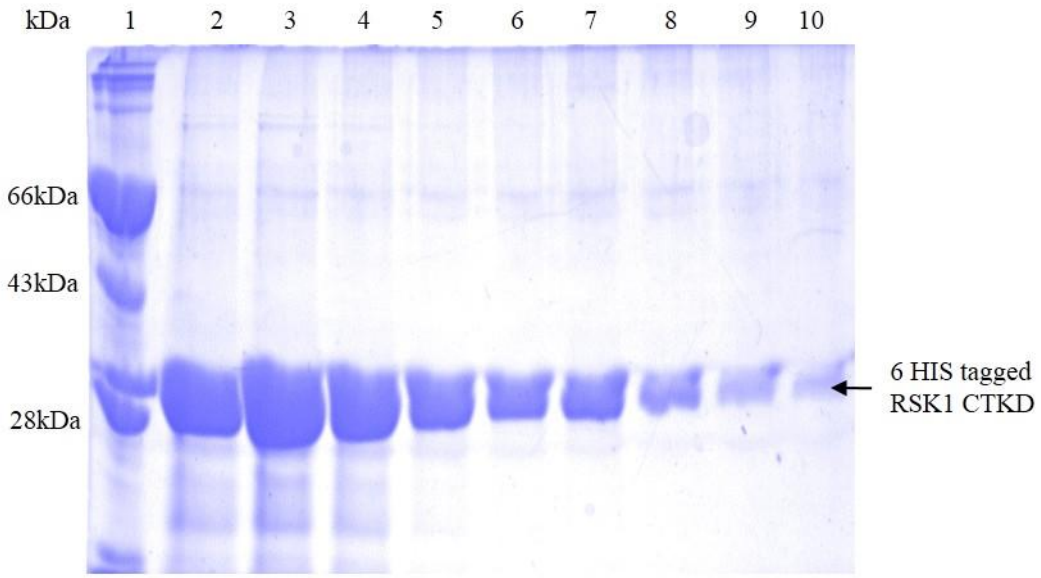


Figure 4.14: Purification of RSK1 CTKD (418-675) from inclusion bodies. The 12 % SDS-PAGE gel shows the 33 kDa RSK1 CTKD (418-675) with N-terminal HIS tag (lane 1: Protein marker, Lanes 2-6: 150 mM Imidazole elutions. Lanes 7-10: 300 mM Imidazole elutions).

4.4. Conclusion

Expressing large proteins with a small 6 HIS tag can sometimes yield poor results. It further depends on the surface accessibility of the tag. If the residues at terminal tails of proteins are hydrophobic, they are generally buried inside the core of the protein during protein folding. In such cases, the 6 HIS tag present at that end gets buried deep inside the hydrophobic core of the protein and produces problems during protein purification. Hence, the best alternative is to incorporate an additional tag at the other end of the protein by site-directed insertion. For this 6 HIS tag would be more stringent as it would require only 18 nucleotides for incorporation by site-directed insertion.

Similarly, upon analysing the KYTE DOOLITTLE data, the N-terminal end of RSK1 was found to be hydrophobic in nature while the C-terminal end is hydrophilic in nature. Therefore, human Ribosomal S6 kinase 1 (RSK1) full length cloned in pRSET-A vector with N-terminal 6 HIS tag, could not be purified efficiently. With this information, an additional 6 HIS tag was engineered at the C-terminal end by site-directed insertion. This dual 6 HIS tagged RSK1 full length protein was expressed and purified from BL21 (DE3) pLysE competent *E. coli* cells to higher purity. The purified protein was further confirmed by mass spectrometry. This strategy in incorporating an additional 6 HIS tag at the other end of protein by site-directed insertion helped in protein purification of RSK1 full length. Fluorescence and CD data suggests the purified protein is well folded, and had a compact structure which is predominantly α-helical in nature. However, thermal denaturation studies revealed it to be thermally less stable owing to its relatively low T_m.

RSK1 CTKD was cloned in pET-28 a (+) vector between EcoR1 and Sal1 and purified from bacterial inclusion bodies. RSK1 CTKD was also cloned in pMT-wb-v5 vector and transfected in drosophila Schneider s2 insect cell lines and stable cell lines were established. Since three decades into discovery of RSKs, the full length structures for any of

its isoforms are still not available. Moreover, a lot of research is still going on MAPK pathway because of its significant role in many reported cancers. Crystal structure of RSK1 full length would help in unravelling the molecular insights of RSK1. It would further help in designing structure based small molecular inhibitors for new cancer therapeutics.

Chapter 5:

Structural and functional basis to understand the folding patterns of Ribosomal S6 Kinase 3

5.1. Introduction

Ribosomal S6 Kinases are part of classical Mitogen-activated Protein Kinase (MAPK) pathway, which transduces the extra cellular signals to the nucleus [41]. In the human genome, Ribosomal S6 Kinase 3 full length protein is encoded by RPS6KA2 gene, present on the q arm of Chromosome 6. The RSK3 comprises of 733 amino acids, having molecular weight of 83238.7 Da, and the theoretical pI 8.82. RSK3 protein expression is very high in lungs, skeletal muscle, brain, uterus, prostate and ovaries [276, 277].

The molecular structure of RSK3 comprises of two functionally active domains. The N-terminal Kinase domain (NTKD) comprises of 260 amino acids from 59-318, whereas the C-terminal Kinase domain (CTKD) which is 258 amino acids from 415-672. The NTKD shares sequence homology with kinases of the AGC (PKA, PKG, PKC) family, while the CTKD is homologous to the calcium/calmodulin-dependent protein kinases (CaMKs) [79, 80]. The CTKD is involved in auto-phosphorylation of RSK whereas NTKD is responsible for various substrate phosphorylations [79]. These two distinct kinase domains are connected by a highly conserved linker region which is around 96 amino acids from 319-414 residues [77]. An N-terminal tail of 58 residues and a C-terminal tail of 61 residues are at either ends of NTKD and CTKD. The ERK binding region, known as the D domain (NH_2-^{719}Leu-Ala-Gln-Arg-Arg^{723} -COOH) is present within the C-terminal tail of RSK3 [84, 86].

The complete activation requires multiple phosphorylations at various conserved residues [83, 85-87]. ERK phosphorylates Threonine at position 570 thereby activating the CTKD of RSK3 [83]. Upon activation, the RSK3 CTKD auto-phosphorylates Serine 377, creating a docking site for PDK1 [42, 278]. The PDK1 then phosphorylates Serine 218 thereby activating the NTKD. The NTKD further phosphorylates Serine 730, which weakens the interaction with ERK, and hence the RSK3 becomes fully functional.

Upon complete activation, RSK3 has a number of substrates to activate, which include FOS, Jun, CREBBP, FLNC, NQO2 and ZMYM5 [276, 279-282]. In strong contrast to RSKs contribution to tumor progression, only RSK3 functions as a potential tumor suppressor in ovarian cancer. RSK3 was shown to inhibit the growth of epithelial ovarian carcinoma cell line UCI 101 [277].

The crystal structures of NTKD and CTKD are available for RSK1 and RSK2 isoforms [101-105]. However, no entry in the PDB is found for RSK3 from any organism till now. The individual domains and full length structures will help us to understand the overall folding pattern of RSK3.

This chapter explains the purification of RSK3 CTKD (415-672) both using pET-28 a (+) and pMal-c2X vectors. Further this chapter includes the biophysical characterization of RSK3 CTKD (415-672) purified from pET-28 a (+) construct, and cloning & purification of RSK3 NTKD (59-318).

5.2. Materials and Methods

5.2.1. Gene cloning of RSK3 CTKD (415-672) in pMal-c2X

C-terminal Kinase domain (415-672) of RSK3 (ACCESSION NM_021135) has been cloned in pMal-c2X vector between EcoR1 and Sal1 restriction sites. The PCR primers required for cloning were designed manually using the sequence information of RSK3 (415-672) residues. The primer sequences are as follows. Forward primer:

5’-GTCGAATTCGAGAACCTGTACTTTCAGGGTTACGAGATCAAGGAGGACATC-3’ Reverse primer: 5’–GTCGTCGACCTATTACACCCACGGGTGTTTGAGCAC-3’. The tetrad PCR system from Bio-Rad and the Phusion High-Fidelity DNA Polymerase enzyme (Finzymes) were used to perform the PCR. Initial denaturation step was at 95 °C for a

duration of 5 minutes, PCR cycle comprised of denaturation at 95 °C for 30 seconds, primer annealing at 65 °C for 30 seconds and the polymerase extension at 72 °C for 30 seconds, and this cycle was repeated 32 times. The amplified PCR product was loaded onto 1 % agarose gel and visualized by staining with ethidium bromide. The PCR product corresponding to ~ 800 bp was excised by visualizing under a UV filter on a Gel doc equipment (UVP bioimaging systems). The PCR product purified using gel extraction kit (Qiagen) was further treated with EcoR1 and Sal1 restriction enzymes (Fermentas) to generate sticky ends. This digested RSK3 CTKD was then ligated (Quick Ligation kit from New England Biolabs) for 10 minutes at room temperature into pMal-c2X vector also treated with EcoR1 and Sal1 restriction enzymes. The ligated product was transformed into DH5α competent cells, and plated on Luria Agar containing 100 µg/ml Ampicillin. A single isolated colony from the plate was then used to inoculate 10 ml of Luria Broth containing 100 µg/ml Ampicillin, and the culture was incubated at 37 °C for 12 hours. The bacterial culture was harvested by centrifugation at 4 °C and 6000 rpm. The recombinant plasmid was isolated using miniprep kit (Qiagen). The presence of RSK3 CTKD insert was checked by agarose gel electrophoresis on the sample subjected to double digestion with EcoR1 and Sal1 restriction enzymes at 37 °C. The identity of the inserted PCR product in the recombinant plasmid was further confirmed by DNA sequencing (Applied Biosystems).

5.2.2. Expression and purification of RSK3 CTKD (415-672)/pMal-c2X

The RSK3 CTKD/pMal-c2X construct was transformed into Rosetta 2 (DE3) competent cells, and plated on Luria agar containing 100 µg/ml Ampicillin and 34 µg/ml chloramphenicol. A 10 ml of Luria Broth medium containing appropriate antibiotics was inoculated with a single isolated colony, and further incubated in a shaker incubator at 37 °C, 200 rpm. After 8 hours of incubation, this starter culture was diluted into 1 litre of Luria

Broth, containing ampicillin and chloramphenicol, and was further allowed to grow in a shaker incubator at 37 °C, 200 rpm till the optical density (OD) λ= 600 reached a value of 0.6. The culture was then induced with 0.4 mM isopropyl-β-D-thiogalactoside (IPTG) for 16 hours at 20 °C, 200 rpm. The bacterial culture was harvested by centrifugation for 15 minutes at 4 °C and 6000 rpm.

One litre of bacterial pellet was resuspended in 10 ml of Buffer E (50 mM Tris-HCl pH 8.0, 500 mM NaCl and 4 mM $MgCl_2$). Protease inhibitor cocktail was added in a ratio of 1:1000 by volume (Sigma-Aldrich). The cells were disrupted by sonication (Branson digital sonifier) for five times, one minute cycles separated by one minute incubation on ice, keeping the pulse on for 4 seconds and off for 2 seconds in every cycle. The sample was maintained on ice throughout the process. Then the cell lysate was centrifuged in a JA-25.50 sorvall rotor at 18,000 rpm, 4 °C for 45 minutes (Avanti J-26S XPI, Beckman coulter). The supernatant was used for affinity purification of RSK3 CTKD (415-672) using amylose resin.

1 ml of amylose resin (New England Biolabs) was equilibrated with buffer E, and binding was performed by adding the supernatant containing RSK3 CTKD (415-672) on to the amylose resin by gravity flow method. Later, washing was done for 20 column volume washes with buffer E. Maltose Binding Protein (MBP) tagged RSK3 CTKD (415-672) fusion protein was eluted in buffer E, containing 20 mM Maltose. These elutions were analysed on 10 % SDS-PAGE gel and visualized by coomassie staining. There was an additional band of ~ 60 kDa co-eluting with MBP-RSK3 CTKD (415-672) fusion protein. The identity of ~ 60 kDa band was confirmed by Mass spectrometry as groL1 60 kDa chaperonin 1 of *E. coli*. To eliminate the bacterial groL1 60 kDa chaperonin 1, the RSK3 CTKD (415-672) was purified again from 1 litre of bacterial pellet, lysed in 10 ml of buffer G (50 mM Tris-HCl pH 8.0, 1 M NaCl, 20 % Glycerol, 5 mM $MgCl_2$). Further sonication, centrifugation and binding of RSK3 CTKD (415-672) were performed. To remove the groL1, 60 kDa chaperonin 1, protein

bound amylase resin was additionally washed with Buffer G containing 5 mM ATP. After washing with 40 column volumes, MBP-RSK3 CTKD (415-672) fusion protein was eluted in buffer G containing 20 mM Maltose. Since the MBP tag could not be cleaved by TEV protease, MBP-RSK3 CTKD (415-672) fusion protein was purified using Superdex-75 FPLC column. This fusion protein was further used for setting up crystallization trials.

5.2.3. Gene cloning of RSK3 CTKD (415-672) in pET-28 a (+)

C-terminal Kinase domain (415-672) of RSK3 (ACCESSION NM_021135) has been cloned in pET-28 a (+) vector between EcoR1 and Sal1 restriction sites. The same primers which were used for cloning in pMal-c2X, were further used for cloning RSK3 (415–672) in pET-28 a (+) vector. Cloning protocol was performed as mentioned in **5.2.1.** section. Since, pET-28 a (+) vector has kanamycin resistant gene, hence 50 µg/ml kanamycin was used instead of ampicillin.

5.2.4. Expression and purification of RSK3 CTKD (415-672)/pET-28 a (+)

The RSK3 CTKD/pET-28 a (+) construct was transformed into Rosetta 2 (DE3) competent cells and plated on Luria agar containing 50 µg/ml kanamycin and 34 µg/ml chloramphenicol. RSK3 CTKD (415–672) was expressed as mentioned in **5.2.2.** section, with 50 µg/ml kanamycin.

One litre of induced bacterial pellet of RSK3 CTKD/pET-28 a (+) was lysed in buffer E. Sonication and centrifugation was done as mentioned in **5.2.2.** section. Since RSK3 CTKD (415–672) forms inclusion bodies, the supernatant containing the soluble fraction was discarded and the cell pellet was thoroughly washed three times with buffer E, additionally containing 1 % Triton X-100, 1 % Sarkosyl and 1 % Sodium Deoxycholate. After each wash

the cell lysate was incubated at 4 °C for 30 minutes, and centrifuged for 20 minutes at 12,000 rpm, 4 °C to remove the supernatant. The cell pellet was given one extra wash with buffer E containing 4 M Urea, before resuspending it into buffer F (50 mM Tris-HCl pH 8.0, 500 mM NaCl**,** 8 M Urea). The denatured protein was separated by centrifuging at 12,000 rpm for 20 minutes at 4 °C. The supernatant was used for purification of 6 HIS-RSK3 CTKD by Immobilized Metal Ion Affinity Chromatography (IMAC).

For IMAC purification, Ni-NTA column (Qiagen) was pre-equilibrated with five column washes of buffer F. Then the cell lysate containing denatured RSK3 CTKD was allowed to bind on the pre-equilibrated Ni-NTA beads by gravity flow methods. The 20 column washes were performed on the protein bound resin with buffer F containing 10 mM and 20 mM imidazole respectively. RSK3 CTKD was eluted in buffer F with 100 mM to 250 mM imidazole. Elution fractions from each run were loaded on 12 % SDS-PAGE gel and visualized using coomassie staining.

5.2.5. Trypsin in-gel digestion and Mass spectrometry

The band corresponding to 33 kDa of purified RSK3 CTKD was excised from 12 % SDS-PAGE gel, and chopped into tiny pieces. These pieces were transferred into 1.7 ml eppendorf (Axygen) tube containing 1 ml de-staining solution which is water, methanol and glacial acetic acid in the ratio of 4.5:4.5:1, and vortexed overnight. The following day, the gel pieces were washed with Ammonium bicarbonate and Acetonitrile, and further treated with trypsin as described by Shevchenko et al [250]. The peptides were then extracted and lyophilized. The lyophilized sample was dissolved in 0.1 % TFA in 50 % of Acetonitrile for mass spectrometry, in which 1 µl of sample was mixed with 2 µl of α-cyano-4-hydroxycinnamic acid, and plated on MTP 384 target plate ground steel BC (Bruker) before allowed to dry by dried droplet method. The mass spectra were recorded by Ultraflex II

MALDI-TOF/TOF mass spectrometer (Bruker Daltonics). For Peptide mass fingerprinting, the spectra were compared with those in the database using Mascot (Matrix Sciences) search engine integrated in the flexAnalysis software (Bruker Daltonics) [283]. groL1 60 kDa chaperonin 1, which co-eluted with MBP-RSK3 CTKD (415-672) fusion protein was also identified by mass spectrometry.

5.2.6. Refolding of RSK3 CTKD (415-672)

The fractions containing RSK3 CTKD (415–672) were pooled down and transferred into Snakeskin dialysis bag (Thermo Scientific). The protein sample was dialysed for 72 hours at 4 °C with intermittent changes of refolding buffer, buffer H (50 mM Tris, pH 8.0, 500 mM NaCl, 50 mM KCl, 10 mM $MgCl_2$, 2 mM DTT, 15 % Glycerol, 1 % Triton-X-100 and 500 mM *l*-Arginine). The First two buffer changes were made after 6 hours of dialysis whereas remaining changes were after 12 hours each. This refolded RSK3 CTKD (415–672) protein was concentrated up to 2 ml and further purified by Size Exclusion Chromatography (SEC) using FPLC (AKTA-GE) in the buffer B (10 mM Tris, pH 8.0, 100 mM NaCl). The FPLC fractions were analysed on 12 % SDS-PAGE gel.

5.2.7. Far UV-CD spectroscopy of RSK3 CTKD (415-672)

10 µM of refolded RSK3 CTKD (415–672) in Buffer B was used to measure the Far UV-CD spectra using Jasco J-815 CD spectrophotometer. The spectrum was recorded between λ = 260 and 190 nm and submitted to K2D3 online server to estimate the content of α-helices and β-sheets [251]. To study the thermal stability using CD spectroscopy, thermal denaturation was performed for 300 µl of 10 µM proteins by varying the sample temperature from 10 °C to 80 °C with an increment of 2 °C.

5.2.8. Fluorescence spectroscopy RSK3 CTKD (415-672)

A fluorescence emission spectrum (Horiba scientific) was recorded from λ = 310 nm to 400 nm for an excitation wavelength of λ=280 nm. The denatured RSK3 CTKD (415–672) was in buffer containing 50 mM Tris-HCl pH 8.0, 500 mM NaCl, 8 M Urea and 250 mM Imidazole whereas refolded RSK3 CTKD (415–672) was in Buffer B. The protein concentration was 1 µM and blank correction was applied individually with respective buffers.

5.2.9. Limited Proteolysis of RSK3 CTKD (415-672) with Trypsin

10 µl of 10 pM Trypsin was added to 250 µl of 3 mg/ml FPLC purified RSK3 CTKD (415–672) protein. 30 µl aliquot was immediately removed after adding trypsin to record limited proteolysis at zero minute. Subsequently aliquots were removed at 5, 10, 15, 30, 60, 120 and 180 minutes and analysed on 15 % SDS-PAGE gel, which also has a lane containing 30 µl of protein without trypsin as a negative control.

5.2.10. Dynamic Light Scattering of RSK3 CTKD (415-672)

To check the oligomeric nature of RSK3 CTKD (415–672), Dynamic Light Scattering (DLS) was performed using 50 µl of refolded RSK3 CTKD at a concentration of 2 mg/ml in Buffer B and the hydrodynamic radius was calculated. (DynaPro NanoStar, Wyatt Technology).

5.2.11. Homology modelling of RSK3 CTKD

The homology model for C-terminal kinase domain of RSK3 was built using Modeller 9.14 software [219, 284]. Crystal structure of RSK1 CTKD (PDB ID: 2WNT) was

used as a template. The sequence identities between C-terminal end of RSK1 and RSK3 were found to be 85 %. The Modeller 9.14 generated five structures and ranked them based on DOPE (Discrete Optimized Protein Energy) score. The structure with least DOPE score – 34348.81 was selected and validated by PROCHECK and ProSA webservers [253] [285] [286].

5.2.12. Gene Cloning of RSK3 NTKD (59-318) in pET-28 a (+)

Ribosomal S6 Kinase 3 N-Terminal Kinase Domain was cloned into 6-HIS based pET-28 a (+) vector between EcoR1 and Sal1 restriction enzyme sites. The primers were designed by keeping EcoR1 in the forward primer and Sal1 in the reverse primers which are as follows, Forward Primer:

5’ -GTCGAATTCGAGAACCTGTACTTTCAGGGTTTTGAGCTGCTGAAGGTTTTA -3’

Reverse Primer: 5' -GTCGTCGACCTATTAAAAGAAGGGATGGCGCTTAATTTC -3’

The PCR amplification and cloning was done as mentioned in **5.2.1.** section. The identity of the inserted PCR product in the recombinant plasmid was further confirmed by DNA sequencing (Applied Biosystems).

5.2.13. Expression and purification of RSK3 NTKD/pET-28 a (+)

The RSK3 NTKD/pET-28 a (+) construct was transformed into Rosetta 2 (DE3) competent cells and plated on Luria agar containing 50 µg/ml kanamycin and 34 µg/ml chloramphenicol. RSK3 NTKD (59-318) was expressed as mentioned in **5.2.2.** section, using 50 µg/ml kanamycin. One litre bacterial pellet of RSK3 NTKD/pET-28 a (+) was lysed in buffer E. Sonication and centrifugation was done as mentioned in **5.2.2.** section. Since RSK3 NTKD (59-318) forms inclusion bodies, the protein was purified under denaturation

conditions as mentioned in **5.2.4.** section. Elution fractions containing RSK3 NTKD (59-318) were loaded on 10 % SDS-PAGE gel and visualized using coomassie staining.

5.2.14. Gene Cloning of RSK3 CTKD and RSK3 NTKD in pMT-wb-v5

Ribosomal S6 Kinase 3, C-Terminal Kinase Domain and N-Terminal Kinase Domain were cloned into pMT-wb-v5 vector between restriction enzymes EcoR1 and Xho1. Since Xho1 site lies downstream of Sal1 in pET-28 a (+) vector, the RSK3 CTKD and RSK3 NTKD inserts had been excised from their respective pET-28 a (+) constructs by double digestion using EcoR1 and Xho1 enzymes and cloned individually in pMT-wb-v5 vector digested with EcoR1 and Xho1.

5.2.15. Transfection of RSK3 CTKD and RSK3 NTKD in insect cell lines

Drosophila Schneider s2 insect cell lines were plated into 6-well plate, 2 ml in each well. DNA-transfection reagent complex was made by mixing 5 µl X-treme GENE HP DNA Transfection Reagent on 10 µg of RSK3 CTKD/pMT-wb-v5 plasmid DNA, and RSK3 NTKD/pMT-wb-v5 plasmid DNA separately in 100 µl of serum free sterile medium. After 30 minutes of incubation, the transfection DNA complexes containing RSK3 CTKD and RSK3 NTKD plasmids were added on separate wells containing Schneider s2 cells. After gentle shake, the 6-well plate was incubated at 25 °C for 48 hours. To establish the stable cell lines, after every 48 hours, Blasticidin was added along with the fresh sterile medium, and after few cycles of seeding with fresh medium and Blasticidin, stable cell lines were established. 100 ml each of Schneider s2 insect cell lines harbouring RSK3 CTKD and RSK3 NTKD were pelleted down at 6000 rpm, 4 °C for 10 minutes. The respective cell pellet was added with 500 µl of RIPA lysis buffer, buffer I (50 mM Tris-HCl pH 8.0, 150 mM NaCl, 5 mM EDTA, 1 % NP-40, 0.5 % Sodium Deoxycholate and 0.1 % SDS). After sonication and

centrifugation as mentioned in **5.2.2.** section, the supernatant was diluted 4 times in buffer E, to avoid the deleterious effects on binding efficiency to IMAC. This diluted supernatant of RSK3 CTKD and RSK3 NTKD were added on pre-equilibrated 100 µl of Ni-NTA resin separately and washed with wash buffer E. The protein bound Ni-NTA resin of RSK3 CTKD and RSK3 NTKD were loaded on to 12 % SDS–PAGE gel, and further western blotting was performed. For Western blotting, the 12 % SDS-PAGE gel without staining was electro-blotted on to PVDF membrane in an electro-blotting apparatus (Trans-blot cell, Bio-Rad) at 4°C in 1x transfer buffer; 0.19 M Glycine, 25 mM Tris-HCl, 0.01 % SDS and 20 % Methanol. The PVDF membrane was carefully incubated with blocking buffer; 5 % BSA in TBST (100 mM Tris-HCl, pH 7.5, 0.9 % w/v NaCl and 0.1 % v/v Tween20). After two hours of blocking and stringent wash at room temperature on a rocker, the blocking solution was replaced with anti HIS antibodies (primary) derived from mouse in the ratio of 1:1000. After incubation with primary antibody, this antibody was replaced by HRP conjugated secondary antibody derived from mouse in the ratio of 1:3000 and incubated for two hours. Soon after incubation with secondary antibody, the PVDF membrane was further incubated with immobilon Western HRP Substrate (Millipore). The Chemiluminescence was detected by exposing the PVDF membrane on to X-ray film inside a dark room and developed using X-ray film processor (Optimax, Protec).

5.3. Results

5.3.1. Gene Cloning of RSK3 CTKD (415-672)

RSK3 CTKD (415-672) was cloned into pMal-c2X vector. **Figure 5.1 (A)** shows PCR amplified RSK3 CTKD (415-672) product of 819 base pairs. Ligation was performed between RSK3 CTKD (415-672) and pMal-c2X vector, both of which were doubly digested

prior with EcoR1 and Sal1 restriction enzymes. **Figure 5.1 (B)** shows the 819 base pair RSK3 CTKD (415-672) insert release (which also includes 45 nucleotides corresponding to TEV cleavage site, restriction sites, stop codons and terminal clamps) from the recombinant pMal-c2X plasmid. This insert release was DNA sequenced and viewed in BioEdit software. **Figure 5.2** confirms the 100 % sequence of RSK3 CTKD (415-672) and in proper frame.

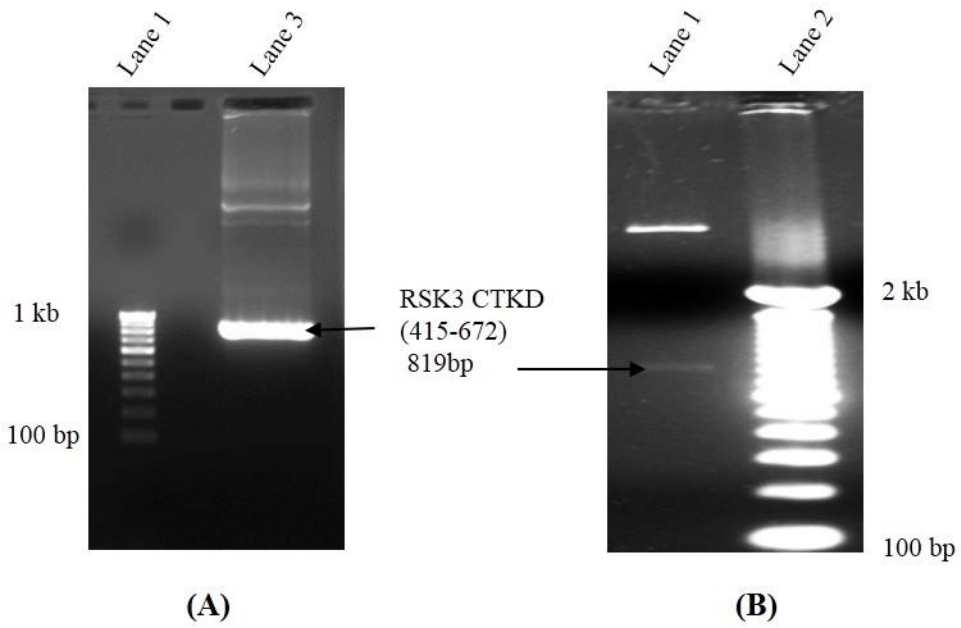


Figure 5.1: Cloning of RSK3 CTKD (415-672). (A) PCR amplification of RSK3 CTKD (415-672) (Lane 1: 100 bp ladder, Lane 3: PCR product) (B) Insert release of RSK3 CTKD (415-672) from the vector backbone following double restriction digestion. (Lane 1: RSK3 CTKD (415-672)/pMal-c2X after double restriction digestion, Lane 2: 100 bp ladder)

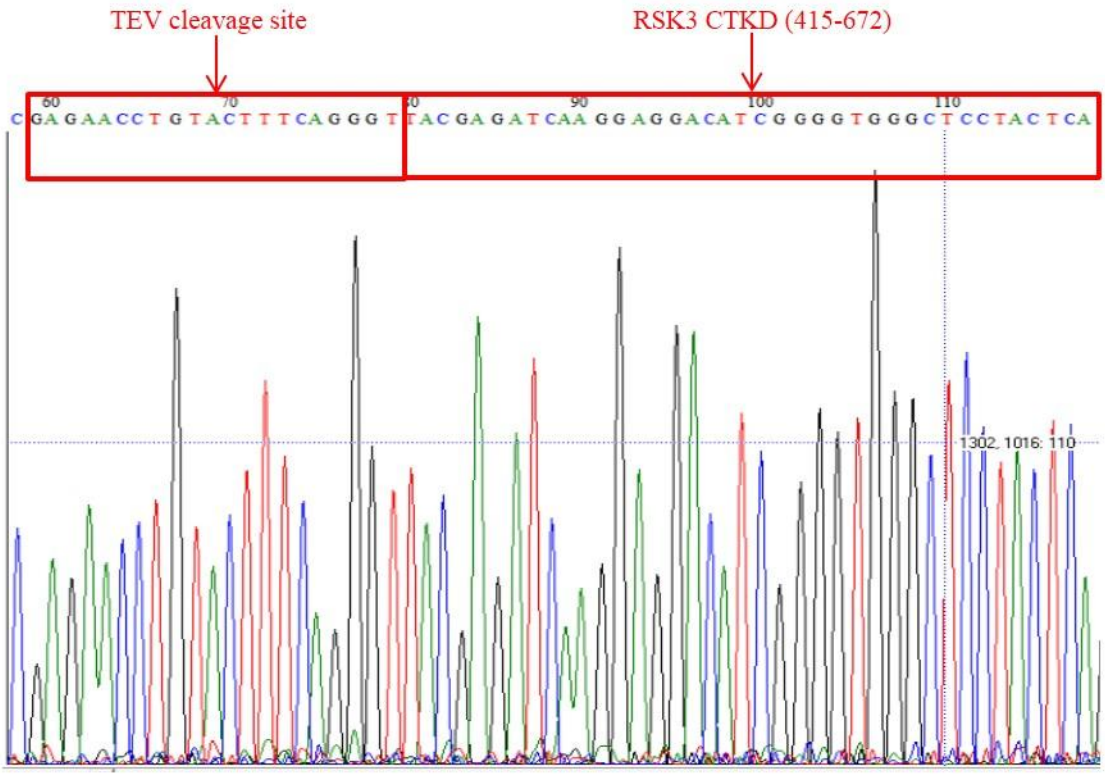


Figure 5.2: Sequencing of RSK3 CTKD (415-672). Sequencing data of RSK3 CTKD (415-672)/pMal-c2X recombinant construct with pMal-c2X forward primer.

5.3.2. Expression and purification of RSK3 CTKD (415-672)

The Molecular weight of RSK3 CTKD (415-672) is ~ 29 kDa and the fusion protein is ~ 72 kDa. The RSK3 CTKD (415-672) was expressed in Rosetta 2 (DE3) competent cells and purified by affinity chromatography using amylose resin. The affinity purified sample was analysed on 12 % SDS-PAGE gel. **Figure 5.3** shows the purification profile of RSK3 CTKD (415-672) which is co-eluting with bacterial chaperonin 60 kDa protein. To confirm identity of the contaminant 60 kDa protein, Trypsin in-gel digestion was performed. **Figure 5.4 (A)** shows the identified extracted peptides with their m/z values as determined by MALDI-TOF/mass spectrometry. **Figure 5.4 (B)** shows the Peptide mass fingerprinting by Mascot search engine using SwissProt database which confirms that the peptides are of groL1 60 kDa chaperonin [283].

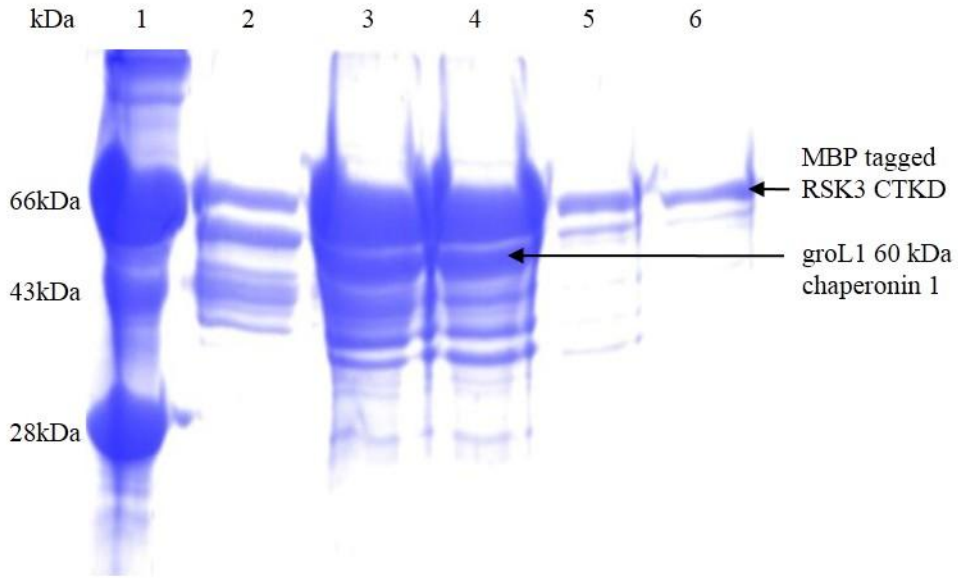
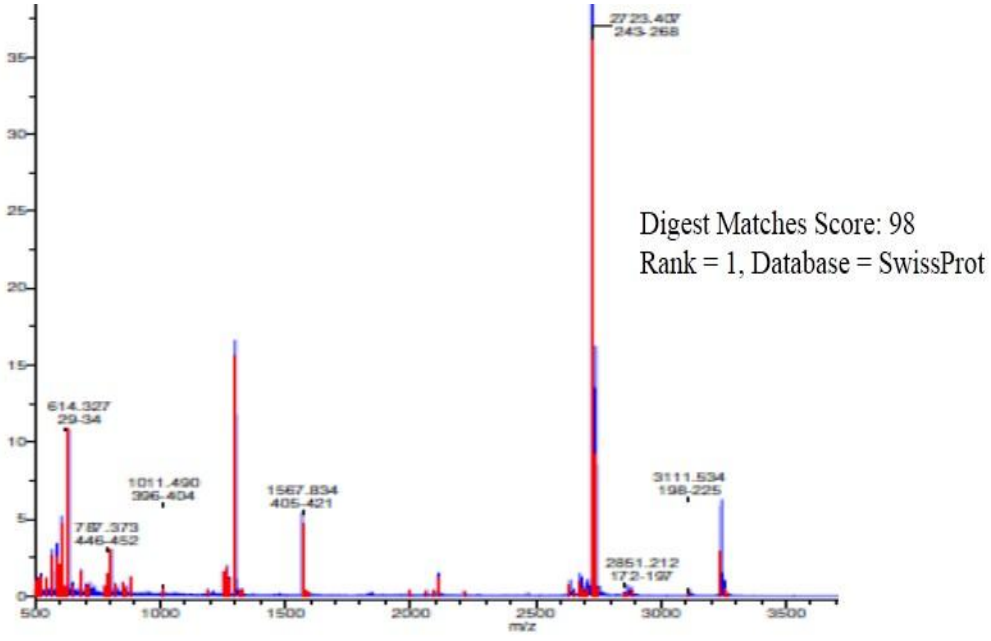


Figure 5.3: Purification of RSK3 CTKD (415-672). The 12 % SDS-PAGE gel shows the 72 kDa RSK3 CTKD (415-672) with N-terminal MBP tag co-eluting with bacterial groL1 60 kDa chaperonin 1 (lane 1: Protein marker, Lane 2: 0.05 mM maltose elution, Lanes 3-6: 20 mM Maltose elutions).



(A)

60 kDa chaperonin 1 OS=Escherichia coli O1:K1 / APEC GN=groL1 PE=3 SV=1 CH601_ECOK1

Intensity Coverage:	34.4 % (42126 cnts)	Sequence Coverage MS:	21.7%
Sequence Coverage MS/MS:	0.0%	pI (isoelectric point):	4.7

10	20	30	40	50	60	70	80
MAAKDVKFGN	DARVKMLRGV	NVLADAVKVT	LGPKGRNVVL	DKSFGAPTIT	KDGVSVAREI	ELEDKFENMG	AQMVKEVASK
90	100	110	120	130	140	150	160
ANDAAGDGTT	TATVLAQAII	TEGLKAVAAG	MNPMDLKRGI	DKAVTAAVEE	LKALSVPCSD	SKAIAQVGTI	SANSDETVGK
170	180	190	200	210	220	230	240
LIAEAMDKVG	KEGVITVEDG	TGLQDELDVV	EGMQFDRGYL	SPYFINKPET	GAVELESPFI	LLADKKISNI	REMLPVLEAV
250	260	270	280	290	300	310	320
AKAGKPLLII	AEDVEGEALA	TLVVNTMRGI	VKVAAVKAPG	FGDRRKAMLQ	DIATLTGGTV	ISEEIGMELE	KATLEDLGQA
330	340	350	360	370	380	390	400
KRVVINKDTT	TIIDGVGEEA	AIQGRVAQIR	QQIEEATSDY	DREKLQERVA	KLAGGVAVIK	VGAATEVEMK	EKKARVEDAL
410	420	430	440	450	460	470	480
HATRAAVEEG	VVAGGGVALI	RVASKLADLR	GQNEDQNVGI	KVALRAMEAP	LRQIVLNCGE	EPSVVANTVK	GGDGNYGYNA
490	500	510	520	530	540	550	
ATEEYGNMID	MGILDPTKVT	RSALQYAASV	AGLMITTECM	VTDLPKNDAA	DLGAAGGMGG	MGGMGGMM	

(B)

Figure 5.4: Peptide mass fingerprinting of in-gel tryptic digests of 60 kDa contamination band. (A) MALDI-TOF/mass spectrometry data showing the identified tryptic peptides of 60 kDa contamination band. (B) The Mascot search engine confirmed the peptides were corresponding to *E. coli* groL1 60 kDa chaperonin 1.

After confirming the groL1 60 kDa chaperonin 1, few more stringent washes were given to amylose resin with wash buffer containing ATP and $MgCl_2$ to purify MBP tagged RSK3 CTKD to greater extent. **Figure 5.5** shows highly purified MBP-RSK3 CTKD (415-672) protein. The MBP tag could not be cleaved with TEV protease and hence crystallization trials of RSK3 CTKD (415-672) were set along with MBP tag.

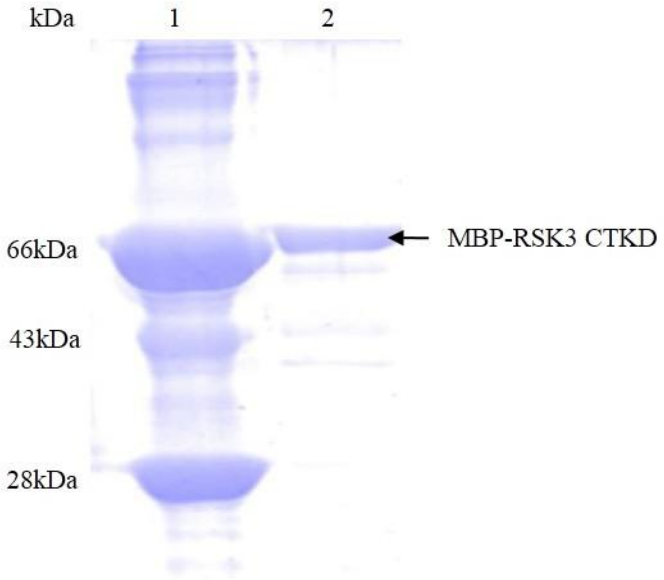


Figure 5.5: Purification of RSK3 CTKD (415-672). The 10 % SDS-PAGE gel shows the 72 kDa RSK3 CTKD (415-672) with N-terminal MBP tag (lane 1: Protein marker, Lane 2: Purified MBP-RSK3 CTKD).

5.3.3. Crystallization trials for RSK3 CTKD (415-672)

Efforts were not successful to cleave out the fusion MBP tag by TEV protease and Factor Xa. Hence, MBP-RSK3 CTKD (415-672) fusion protein was tried to crystallize both by hanging-drop and sitting-drop methods. However, no crystals were seen in any of the conditions. **Figure 5.6 (A-F)** shows precipitation of MBP-RSK3 CTKD (415-672) fusion protein in different conditions during crystallization using Hampton crystal screens 1 and 2.

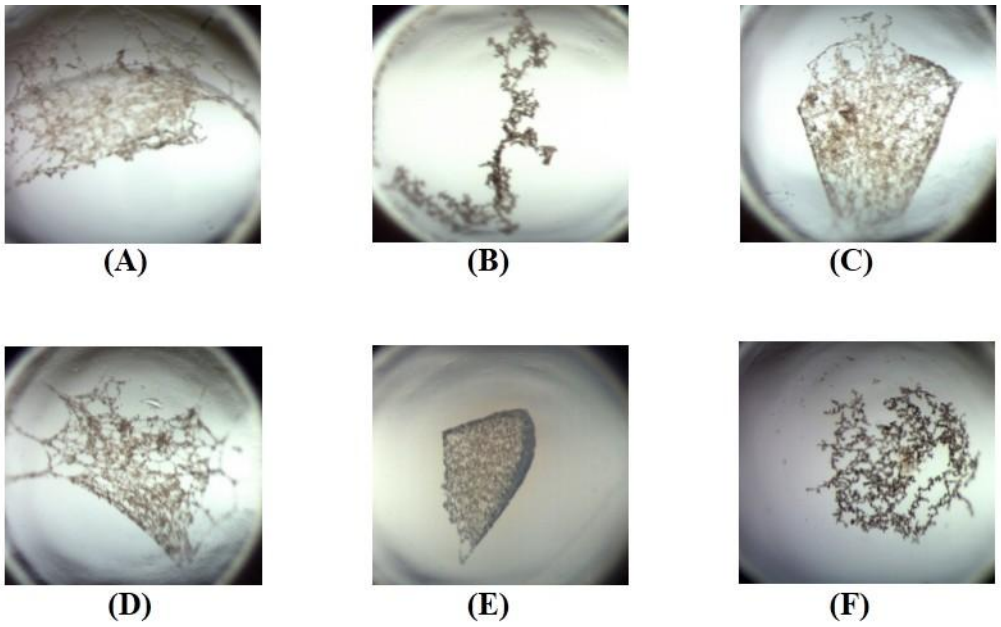


Figure 5.6: Crystallization trials for MBP-RSK3 CTKD (415-672). (A to F) shows the precipitated MBP-RSK3 CTKD (415-672) during crystallization trials.

5.3.4. Gene Cloning of RSK3 CTKD (415-672)

C-terminal Kinase domain of Ribosomal S6 Kinase 3 was cloned in pET-28 a (+) vector. The PCR product shown in **Figure 5.1 (A)** was used to clone in pET-28 a (+) vector. **Figure 5.7** shows the 819 base pair RSK3 CTKD (415-672) insert release from the recombinant pET-28 a (+) vector plasmid following a double restriction digestion with EcoR1 and Sal1 restriction enzymes.

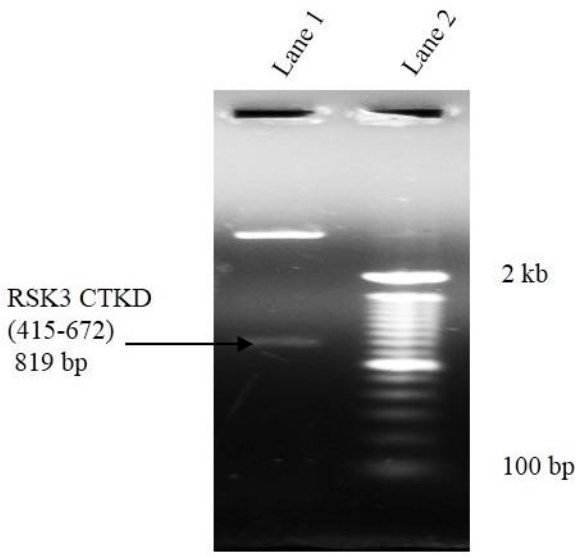


Figure 5.7: Cloning of RSK3 CTKD (415-672). Insert release of RSK3 CTKD (415-672) from the vector backbone following double restriction digestion. (Lane 1: RSK3 CTKD (415-672)/pET-28 a (+) after double restriction digestion, Lane 2: 100 bp ladder)

5.3.5. Expression and purification of RSK3 CTKD (415-672)/pET-28 a (+)

The predicted Molecular mass of RSK3 CTKD is ~ 29 kDa and of the fusion protein is around ~ 33 kDa which includes additional 43 amino acids at the N-terminus. RSK3 CTKD was expressed using Rosetta 2 (DE3) bacterial strain. **Figure 5.8** shows the over expression of RSK3 CTKD (415-672), which is the part of inclusion bodies. The band at 33 kDa is in agreement with the expected molecular weight of protein. Bacterial pellet containing RSK3 CTKD inclusion bodies were washed with detergents Triton X-100, Sodium Deoxycholate and Sarkosyl to remove contaminants before solubilizing in 8 M Urea. **Figure 5.9** shows the denatured RSK3 CTKD (415-672) eluted from Ni-NTA affinity resin and analysed on 12 % SDS-PAGE gel**.** The protein was then dialysed in refolding buffer listed in **Table 5.1** at 4 °C for 72 hours. This dialysed protein was concentrated and passed through Superdex-200 FPLC column to get highly purified protein. FPLC purified RSK3 CTKD (415-672) was analysed on 12 % SDS-PAGE gel, **Figure 5.10**.

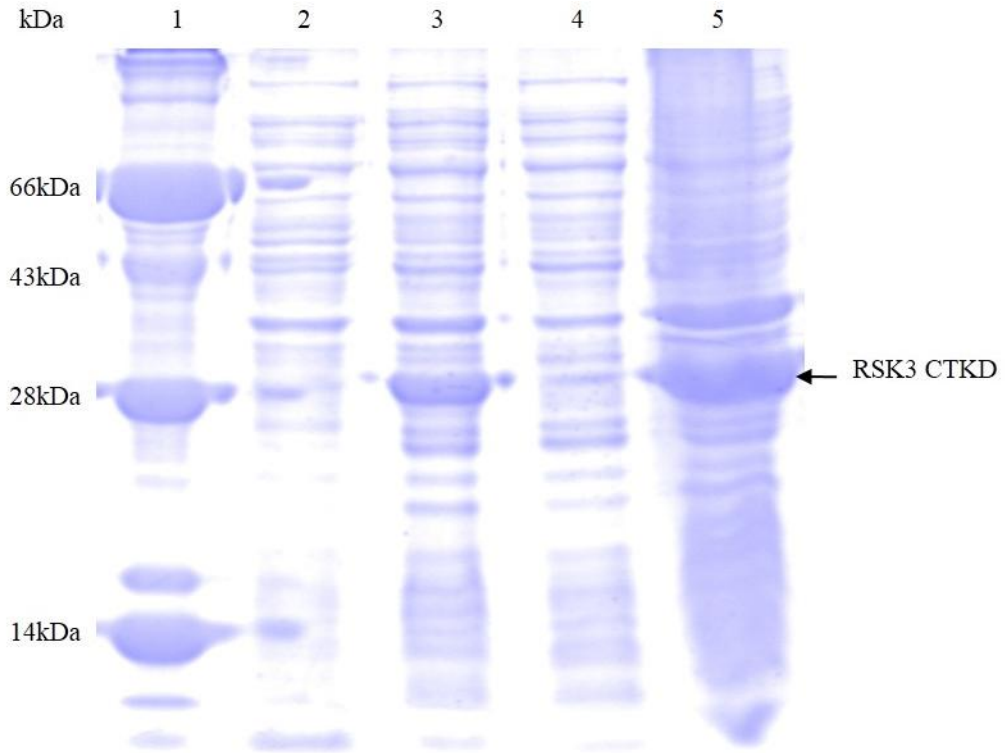


Figure 5.8: Expression profile of RSK3 CTKD (415-672). The 12 % SDS-PAGE gel shows the expression of RSK3 CTKD in the Induced whole fraction and pellet but not in soluble fraction suggesting the protein is insoluble (Lane 1: Protein marker, Lane 2: Uninduced whole cell; Lane 3: Induced whole cell; Lane 4: Induced Soluble fraction; Lane 5: Bacterial insoluble pellet).

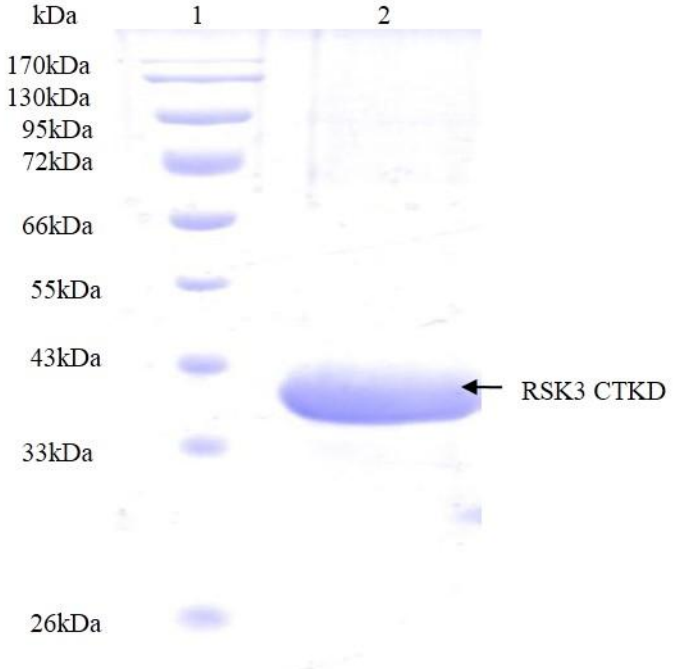


Figure 5.9: Purification of RSK3 CTKD (415-672). The 12 % SDS-PAGE gel shows Purified RSK3 CTKD from inclusion bodies in 8 M Urea eluted at 250 mM Imidazole (Lane 1: Protein marker; Lane 2: RSK3 CTKD in 8 M Urea).

Table 5.1: Refolding buffer for RSK3 CTKD (415-672)

Refolding buffer: pH 8.0	
Tris	50 mM
NaCl	500 mM
KCl	50 mM
DTT	2 mM
$MgCl_2$	10 mM
Glycerol	15 %
Triton-X-100	1 %
l-Arginine	500 mM

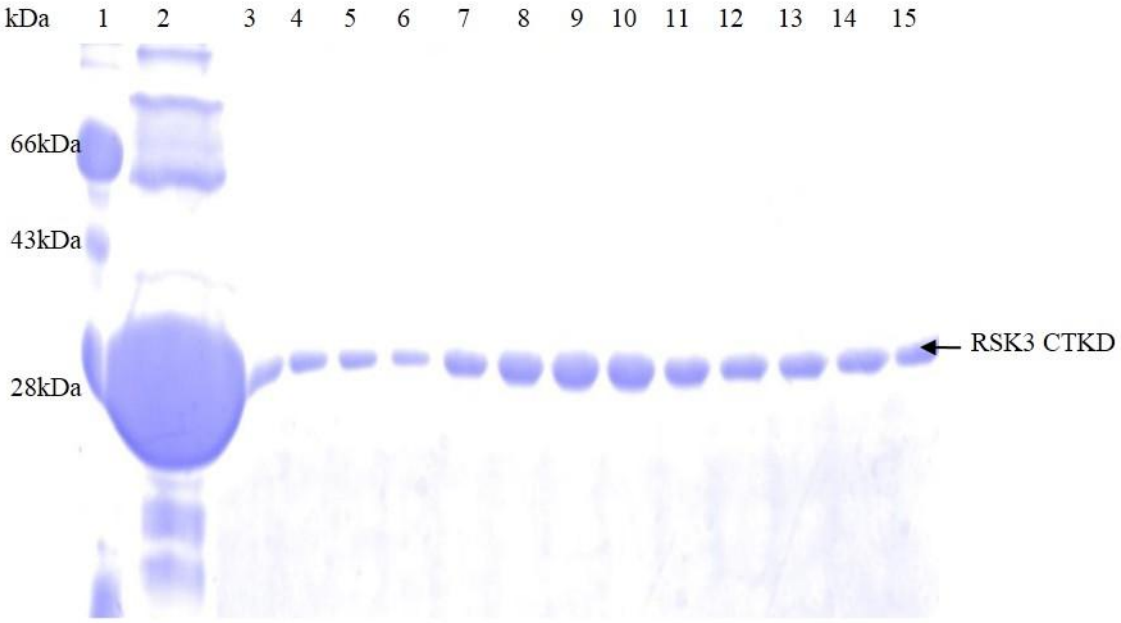
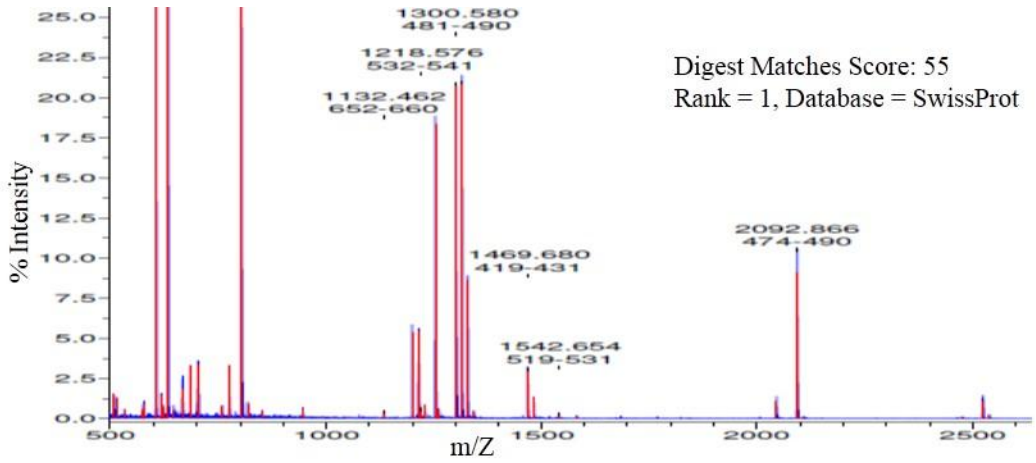


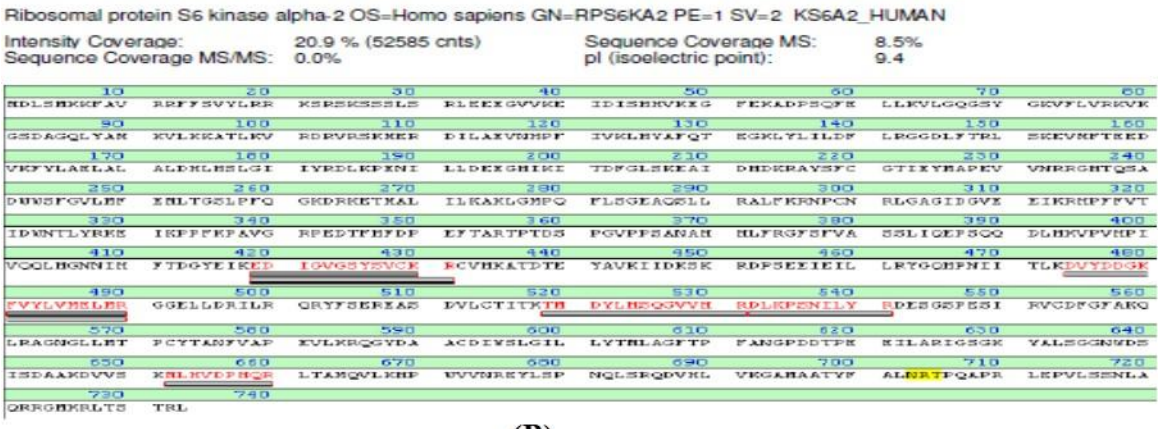
Figure 5.10: FPLC Purification of RSK3 CTKD. The 12 % SDS-PAGE gel shows the RSK3 CTKD purified from Superdex-200 column (Lane 1: Protein marker; Lane 2: Concentrated protein before FPLC; Lanes 3-6: void fractions; lanes 7-15: 80 ml fractions).

5.3.6. Confirmation of RSK3 CTKD (415-672)

To check the nativity of the purified protein, Trypsin in-gel digestion was performed. **Figure 5.11 (A)** shows distribution of masses of peptides generated by in-gel trypsin digestion. The peptide mass fingerprinting analysis has confirmed that the purified protein is Ribosomal S6 Kinase 3, and is within the cloned region of CTKD (415-672) **Figure 5.11 (B)**.



(A)



(B)

Figure 5.11: Peptide mass fingerprinting of in-gel tryptic digests of RSK3 CTKD (415-672). (A) The identified tryptic peptides of RSK3 as analysed by MALDI-TOF/mass spectrometry. (B) The Mascot search engine confirmed the peptides were corresponding to RSK3 at the desired region of cloning.

5.3.7. Secondary structure analysis of RSK3 CTKD (415-672)

CD spectrum was recorded in the far-UV region for the FPLC purified protein. **Figure 5.12 (A)** shows the characteristic dips at λ = 208 nm and at λ = 222 nm indicating the RSK3 CTKD (415-672) is predominantly α-helical. **Figure 5.12 (B)** shows K2D3 data which estimates by matching with the closest spectrum, the α-helical content to be around 52 % and β-sheets to be around 8 % [251].

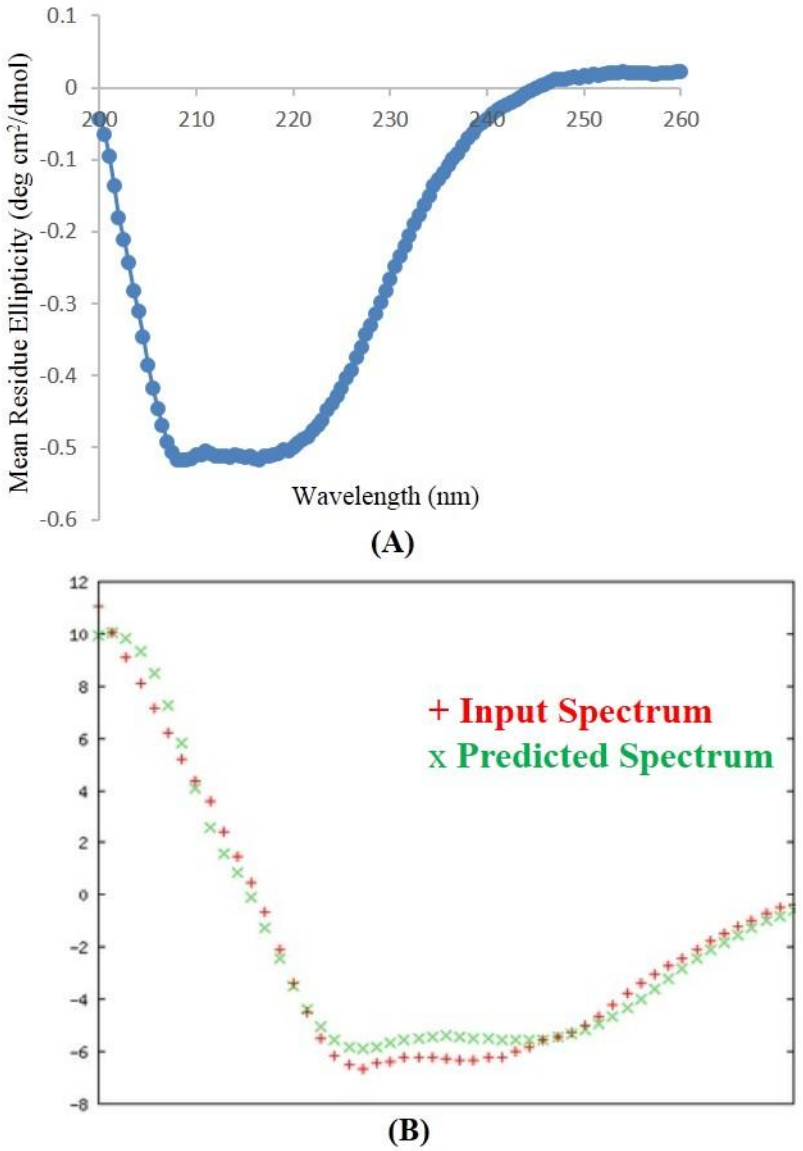


Figure 5.12: CD analysis of RSK3 CTKD (A) CD Spectra of refolded RSK3 CTKD shows the dips λ = 208 nm and λ = 222 nm indicates and it is predominantly α-helical. (B) The K2D3 server predicted the protein has around 52 % of α-helices and 8 % β-strands.

5.3.8. Tertiary structure analysis of RSK3 CTKD (415-672)

Tryptophan Fluorescence spectroscopy was performed to analyse the tertiary conformation of RSK3 CTKD (415-672). There are 3 tryptophan (Trp) residues in the CTKD (415-672) region of RSK3 at positions Trp595, Trp638 and Trp671 respectively. **Figure 5.13** shows the blank corrected fluorescence spectrum for unfolded RSK3 CTKD (in red) recorded for an excitation wavelength of λ = 280 nm, shows the emission maximum at λ = 347 nm. On the other hand, refolded RSK3 CTKD (in blue) has the emission maximum at λ = 332 nm. The 15 nm blue shift confirms the protein to be properly refolded into a tertiary structure burying the tryptophans deep into the hydrophobic core.

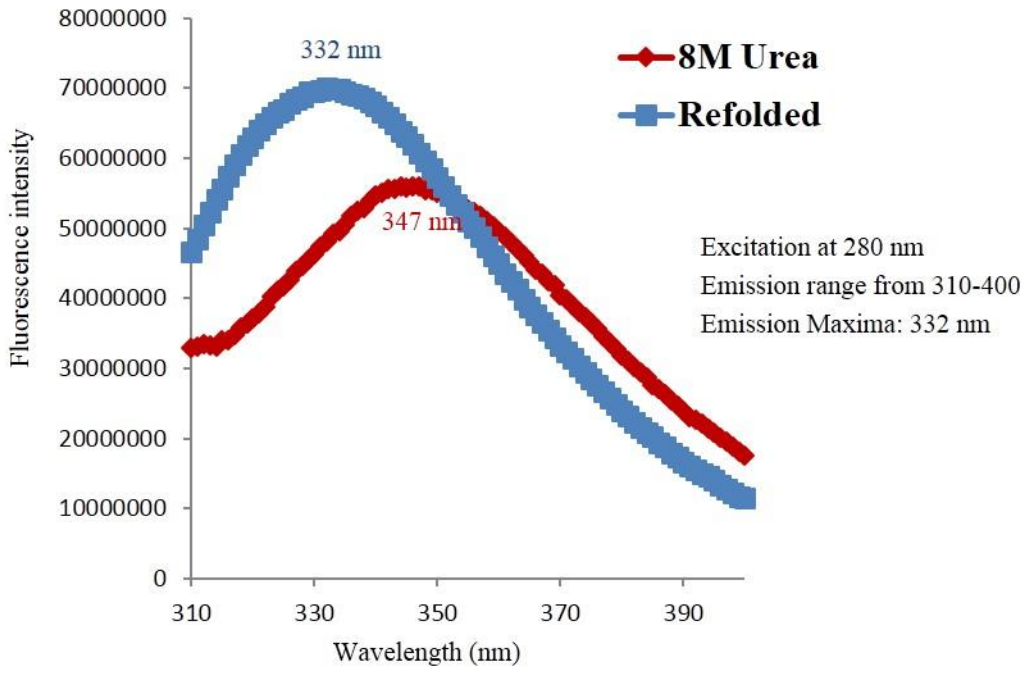
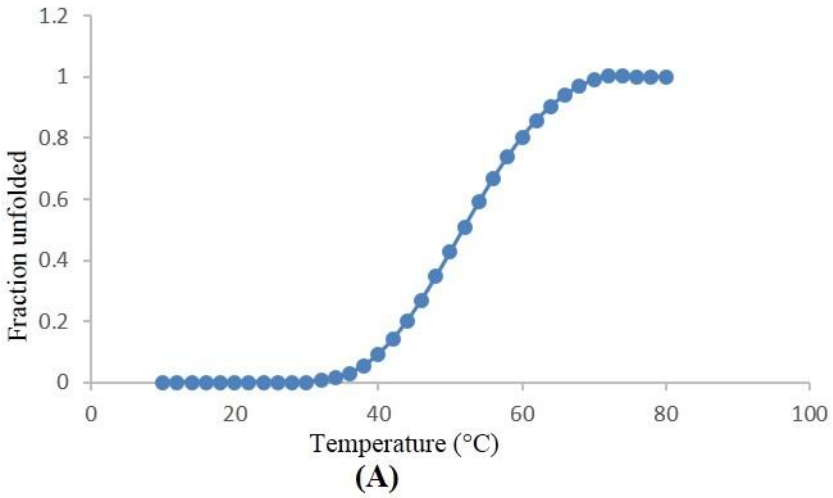


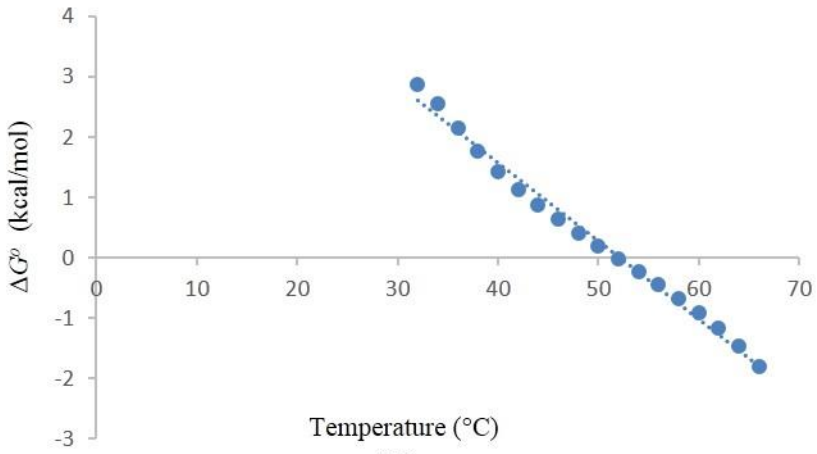
Figure 5.13: Fluorescence spectroscopy of RSK3 CTKD (415-672). Fluorescence spectroscopy of RSK3 CTKD with a excitation wavelength λ = 280 nm shows the emission maxima (λ_{max}) for unfolded protein is 347 nm whereas for the refolded protein it is 332 nm.

5.3.9. Thermal denaturation of RSK3 CTKD (415-672)

CD Spectroscopy was used to study the thermal denaturation of RSK3 CTKD (415-672) by recording the spectra at different temperatures ranging from 10 °C to 80 °C with an interval of 2 °C. **Figure 5.14 (A)** shows the mean residual ellipticity corresponding to λ = 222 nm, which was used to calculate the fraction unfolded, and was plotted with respect to temperature. The Thermal denaturation study revealed that the protein is relatively stable till 30 °C and starts losing its α-helical characteristics from 40 °C, and at 70 °C completely loses its secondary structure. **Figure 5.14 (B)** shows the fitted Linear Extrapolation graph [275] and the T_m was accurately calculated as X-axis intercept at 52 °C (±0.24 °C) and the standard free energy was calculated as Y-axis intercept at 6.9575 kcal/mol (±0.018kcal/mol) by extrapolating the graph.



(A)



(B)

Figure 5.14: Thermal denaturation of RSK3 CTKD (A) The thermal denaturation profile of RSK3 CTKD depicting the fraction unfolded with function of temperature showing the pre transition and post transition states (B) The standard free energy of unfolding of RSK3 CTKD graph fitted into Linear Extrapolation shows the Tm as the X-axis intercept which is around 52 °C and the ΔG^o (H2O) is 6.9575 kcal/mol.

5.3.10. Limited Proteolysis of RSK3 CTKD (415-672)

To confirm the folded state of RSK3 CTKD (415-672), limited proteolysis with trypsin was performed. 15 % SDS-PAGE gel suggests that RSK3 CTKD (415-672) has a compact structure. **Figure 5.15**. shows the slow rate of proteolysis, at the beginning is suggestive of folded nature of the protein making accessibility to cleavage sites difficult for Trypsin. However, an additional band is formed at ~ 9 kDa after 5 minutes which resists complete proteolysis till 120 minutes and starts degrading after three hours.

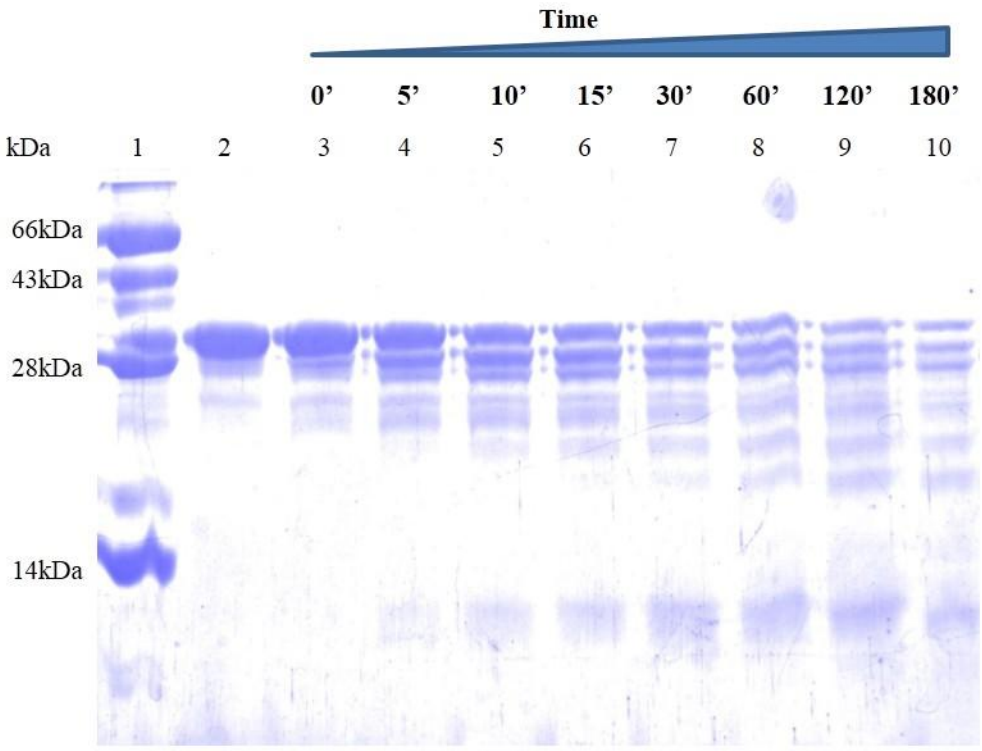
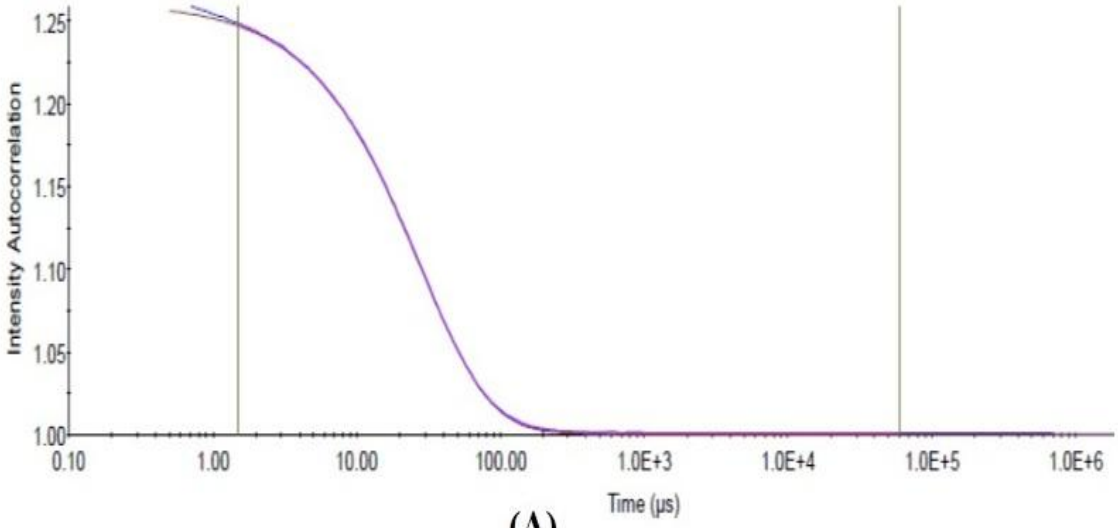


Figure 5.15: Limited proteolysis of RSK3 CTKD (415-672) with trypsin. 15 % SDS-PAGE gel shows relatively slow proteolysis of the RSK3 CTKD with trypsin (Lane 1: Protein marker; Lane 2: Negative control, Lanes 3-10: The time intervals of trypsin digestion of RSK3 CTKD are indicated in minutes above the gel).

5.3.11. Trimeric nature of RSK3 CTKD

Dynamic Light Scattering (DLS) experiments were performed to check the oligomeric nature of RSK3 CTKD (415-672) and calculate its hydrodynamic radius (R_h). **Figure 5.16 (A)** shows that the Correlogram obtained is consistent with a monomodal distribution. Fast decay time at the beginning of the graph suggests the mean radius to be within the range 1 to 10 micro seconds (μs) for proteins. Furthermore, merger of graph at the end with basement suggests the absence of any aggregates, suggesting purified RSK3 CTKD (415-672) is monodisperse in nature. In monodisperse solutions, the percentage of polydispersity is ≤ 20 %. **Figure 5.16 (B)** shows the Dynamic Light Scattering (DLS) data of RSK3 CTKD which was performed by taking 8 acquisitions, out of which four showed less

than 20 % polydispersity, whereas remaining four are ≤ 23.1 %. The derived R_h and Molecular weight of the monodisperse species of RSK3 CTKD are ~ 4.3 nm and ~ 100 kDa respectively. The Molecular weight of 100 kDa is approximately 3 times the size of monomer. Hence it can be concluded that RSK3 CTKD forms trimer in solution.



(A)

Item	Time	Temp	Intensity	Radius	%Pd	Mw-R
	(s)	(C)	(Cnt/s)	(nm)		(kDa)
Acq 1	237.4	25.0	521261	4.2	23.1	97
Acq 2	242.5	25.0	475630	4.3	18.2	100
Acq 3	247.5	25.0	470048	4.2	19.2	95
Acq 4	252.6	25.0	475164	4.2	21.6	98
Acq 5	257.6	25.0	483348	4.2	18.4	99
Acq 6	262.7	25.0	483281	4.3	18.9	100
Acq 7	267.7	25.0	481878	4.3	22.2	100
Acq 8	272.8	25.0	480873	4.3	20.1	103

(B)

Figure 5.16: Dynamic Light Scattering of RSK3 CTKD (415-672). (A) The Correlogram of RSK3 CTKD suggests the protein is monomodal and monodisperse. (B) The recorded 8 acquisitions suggests the R_h and relative molecular weight is ~ 4.3 nm and ~ 100 kDa confirming the trimeric nature of RSK3 CTKD.

5.3.12. Crystallization trials for RSK3 CTKD (415-672)

RSK3 CTKD (415-672) was purified to highest purity and crystallization trials were set at various protein concentrations from 10 mg/ml 12.5 mg/ml, 15 mg/ml and 20 mg/ml using commercially available Hampton and Qiagen screens. Phoenix robotics was also used for setting crystallization trials. **Figure 5.17 (A-E)** shows precipitation of RSK3 CTKD (415-672) protein with different buffer conditions during crystallization trials by sitting-drop vapour diffusion method.

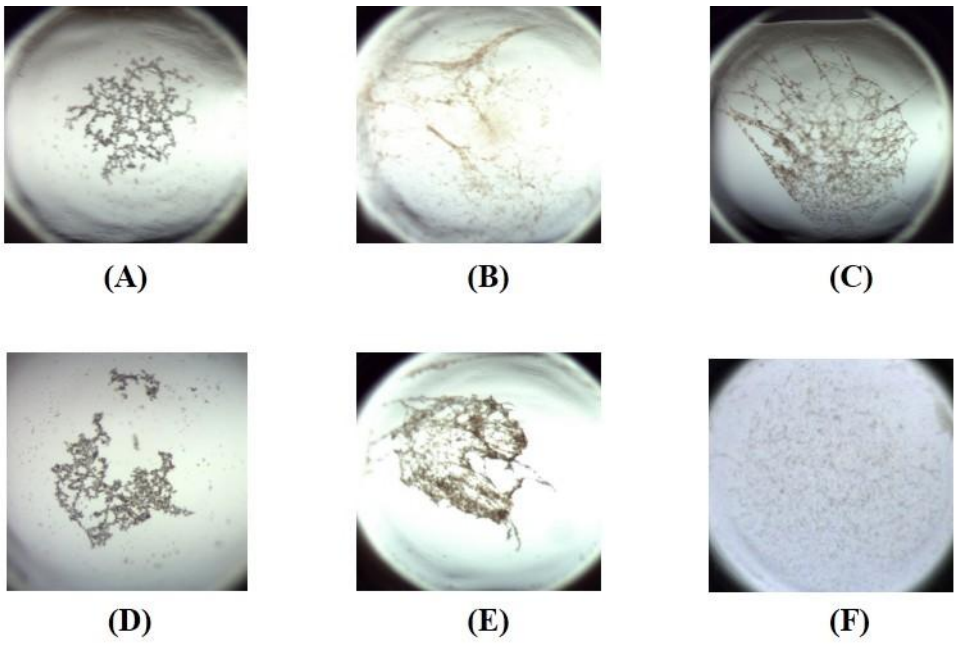


Figure 5.17: Crystallization trials for RSK3 CTKD (415-672). (A to F) shows the precipitated RSK3 CTKD (415-672) during crystallization trials.

5.3.13. *In-silico* modelling of RSK3 CTKD

In the absence of crystal structure of RSK3 CTKD, *in-silico* homology model was built to compare the *in-vitro* and biophysical results. Since the C-terminal regions of RSK1 and RSK3 are 85 % identical, RSK3 was modelled using Modeller 9.14 using the crystal structure of RSK1 (PDB ID: 2WNT) as a template. The modeller generated five

plausible structures which were ranked based on DOPE (Discrete Optimized Protein Energy) score. **Figure 5.18** shows the model structure with the least DOPE score of –34348.81. The model is predominantly α-helical in nature, which is in agreement with the CD spectroscopy results. The model structure was further validated by PROCHECK and ProSA web servers [253, 286]. **Figure 5.19** shows the Ramachandran plot having 213 residues in the most favoured regions (94.7 %), 10 residues in additional allowed regions (4.4 %) and 2 residues in generously allowed regions (0.9 %). No residues were found in the disallowed regions. Furthermore, to validate the *z*-score of the structure we had submitted the model structure to ProSA webserver. ProSA compares the *z*-score of the model structure with the already available *z*-scores of all the X-ray and NMR structures in the PDB. **Figure 5.20** shows the *z*-score for RSK3 CTKD, i.e. -7.24, which is a very good. Overall PROCHECK and ProSA suggests the RSK3 CTKD model is of a good and acceptable quality.

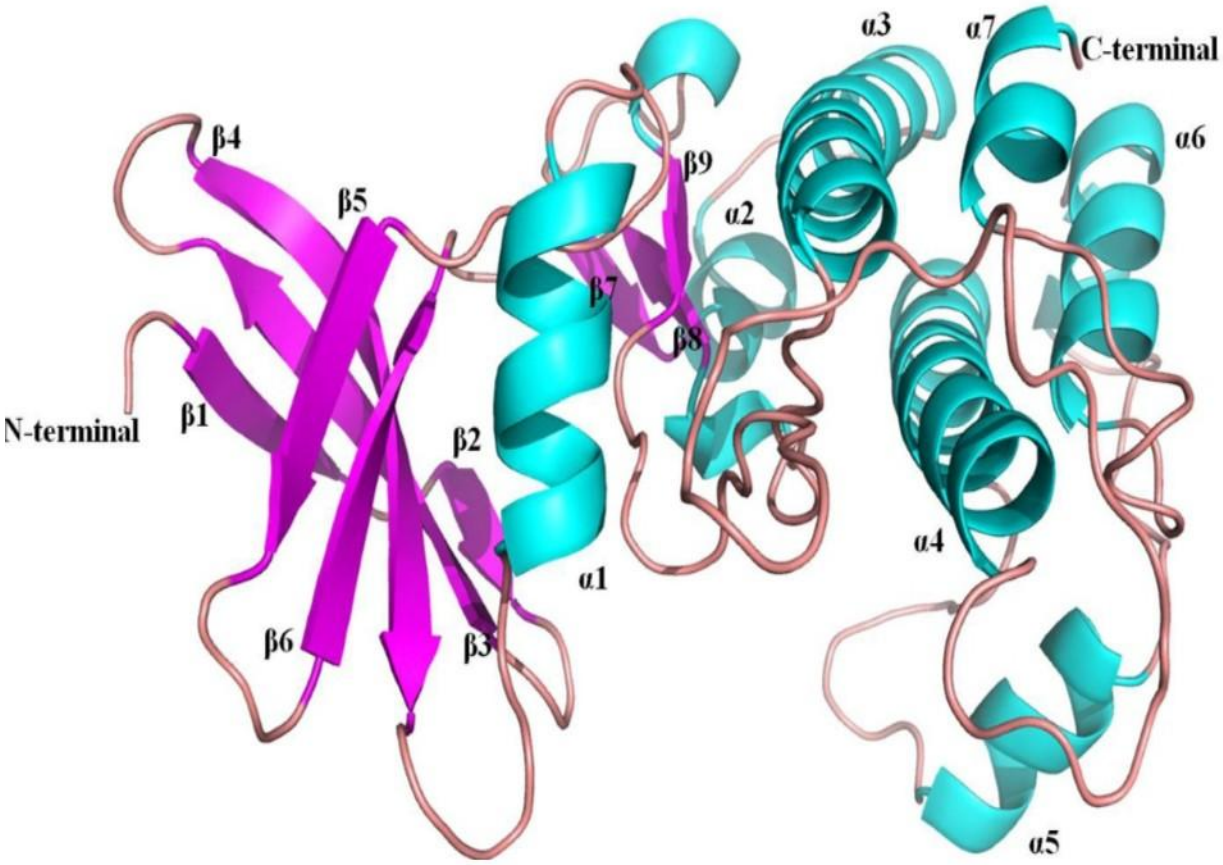


Figure 5.18: The model of RSK3 CTKD built on Modeller 9.14. RSK3 CTKD was modelled using the crystal structure of RSK1 (PDB id: 2WNT) as a template (The α-helices are shown in cyan and β-strands are shown in magenta).

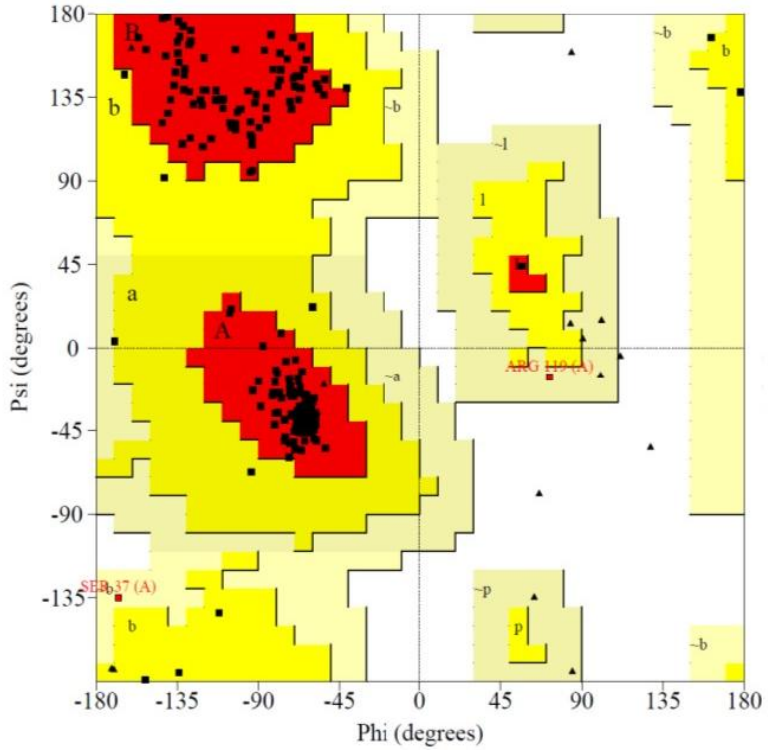


Figure 5.19: PROCHECK data of RSK3 CTKD. Ramachandran Plot of RSK3 CTKD from PROCHECK server shows that there are 213 residues in the most favoured regions (94.7 %), 10 residues in additional allowed regions (4.4 %), 2 Residues in generously allowed regions (0.9 %) and 0 residues in disallowed regions.

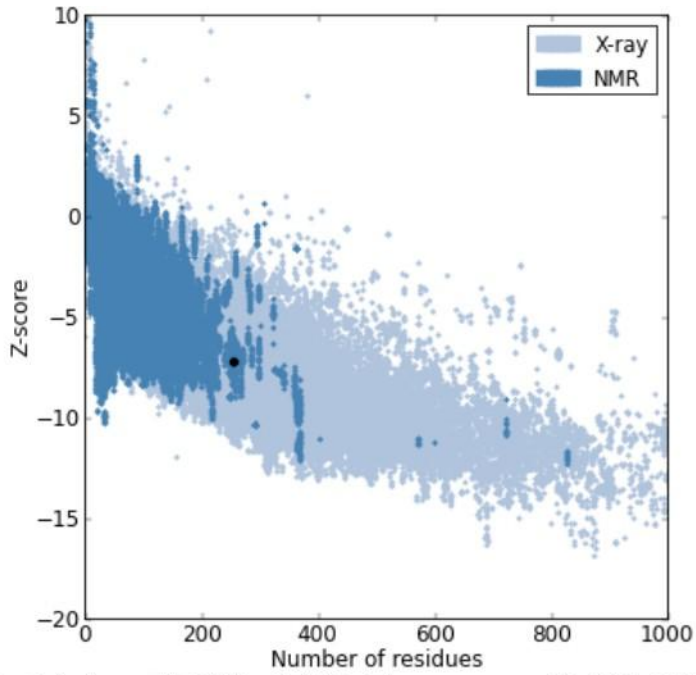
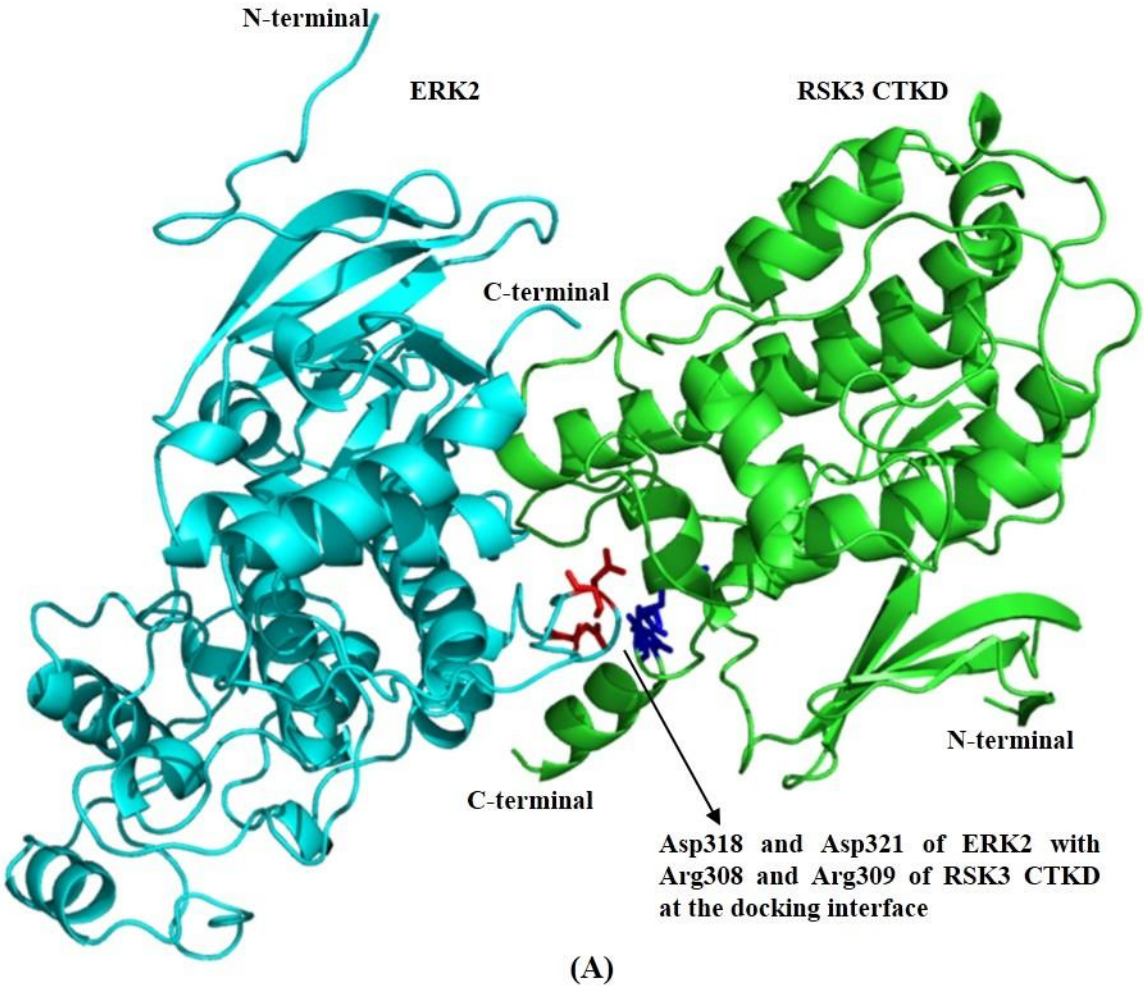
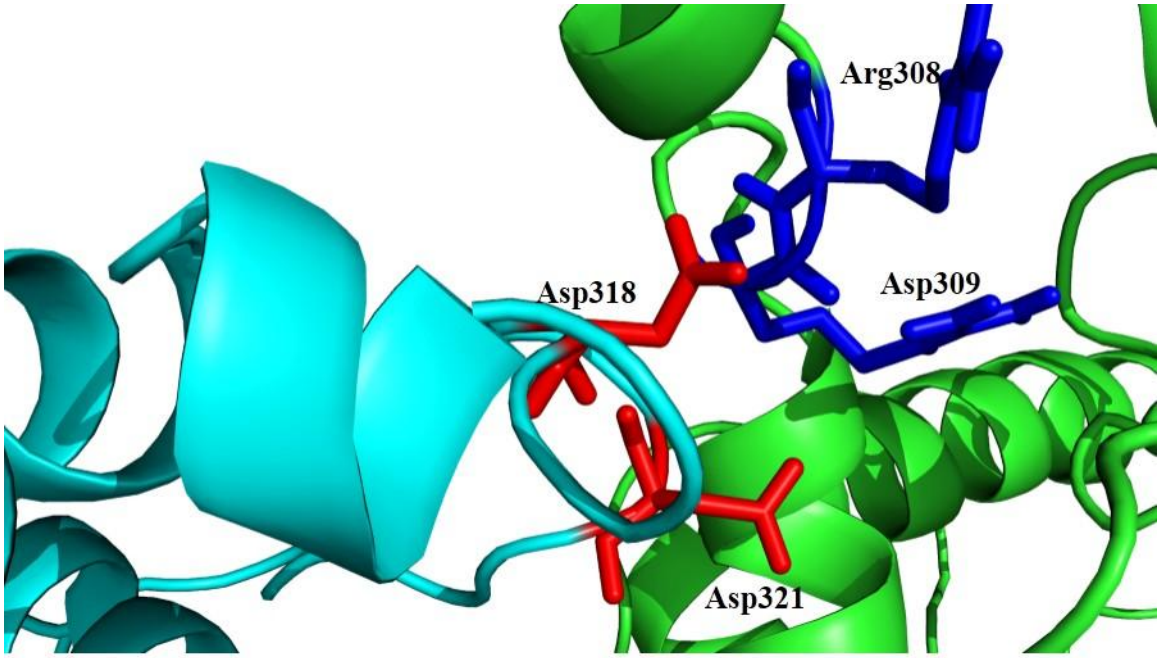


Figure 5.20: ProSA data of RSK3 CTKD: The *z*-score of RSK3 CTKD is -7.24 (black dot) which is well within the acceptable range of known *z*-scores from the PDB including both X-ray and NMR structures.

Since the ERK2 binding motif was missing in the model structure of RSK3 CTKD, another model was further built including the C-terminal tail 'NH_2-^{719}LAQRR^{723}-COOH' using ROBETTA server [220]. This model was then docked on to ERK2 (PDB id: 1TVO) using HADDOCK [287]. **Figure 5.21 (A)** shows the docked complex between ERK2 (cyan) and RSK3 CTKD (green). The important residues at the docking interface have been highlighted. **Figure 5.21 (B)** shows the zoomed view of docking interface highlighting Asp318 and Asp321 of ERK2 in red sticks whereas Arg308 and Arg309 of RSK3 CTKD in blue sticks.


N-terminal
ERK2
RSK3 CTKD
C-terminal
N-terminal
C-terminal
Asp318 and Asp321 of ERK2 with Arg308 and Arg309 of RSK3 CTKD at the docking interface

(A)


Arg308
Asp318
Asp309
Asp321

(B)

Figure 5.21: Docking of ERK2 on RSK3 CTKD model. (A) The ERK2 (cyan) has a docking site at the C-terminal end of RSK1 (green). (B) The important residues at the binding interface of ERK2 is shown in red sticks (Asp318 and Asp321) whereas RSK3 CTKD in blue sticks (Arg308 and Arg309).

After characterizing RSK3 CTKD, we decided to explore RSK 3 NTKD.

5.3.14. Gene Cloning of RSK3 NTKD (59-318)

RSK3 NTKD (59-318) was cloned into bacterial expression vector pET-28 a (+) vector. **Figure 5.22 (A)** shows PCR amplified 3 NTKD (59-318) product of 825 base pairs. Ligation was performed between RSK3 NTKD (59-318) and pET-28 a (+) vector, doubly digested with EcoR1 and Sal1 restriction enzymes. **Figure 5.22 (B)** shows the 825 base pair RSK3 NTKD (59-318) insert from the recombinant pET-28 a (+) plasmid following a double restriction digestion. **Figure 5.23** shows the sequencing data of RSK3 NTKD (59-318)/pET-28 a (+) construct as viewed in BioEdit software and it is in proper frame.

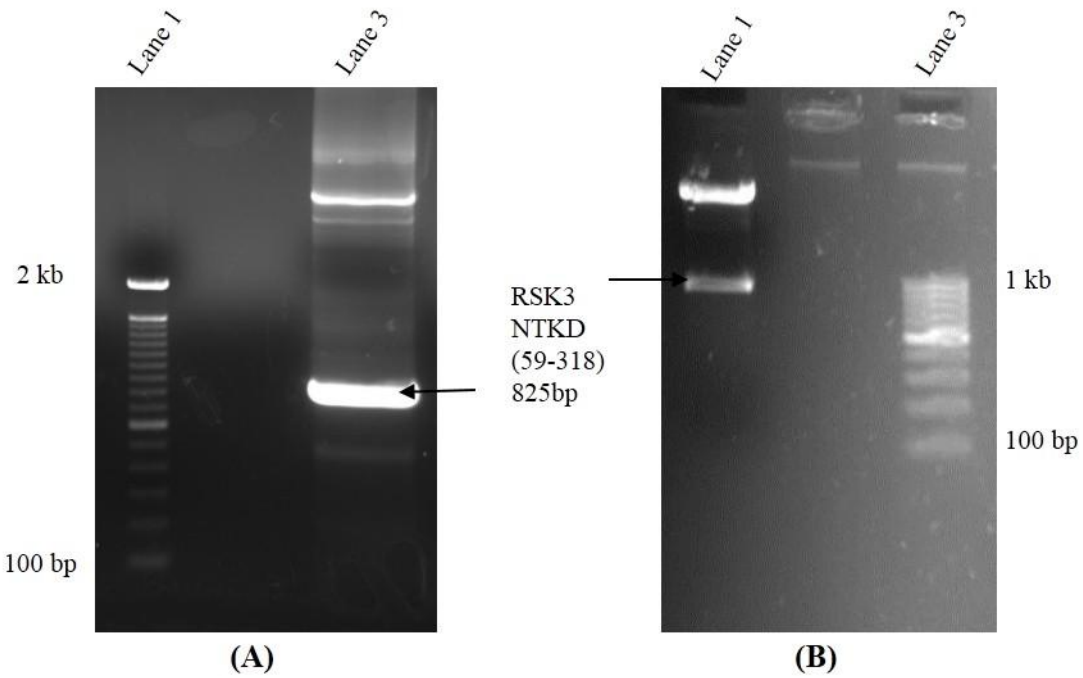


Figure 5.22: Cloning of RSK3 NTKD (59-318). (A) PCR amplification of RSK3 NTKD (59-318) (Lane 1: 100 bp ladder, Lane 3: PCR product of RSK3 NTKD (59-318)) (B) Insert release of RSK3 NTKD (59-318) from the vector backbone following double restriction digestion. (Lane 1: RSK3 NTKD (59-318)/pET-28 a (+) vector after double restriction digestion, Lane 3: 100 bp ladder)

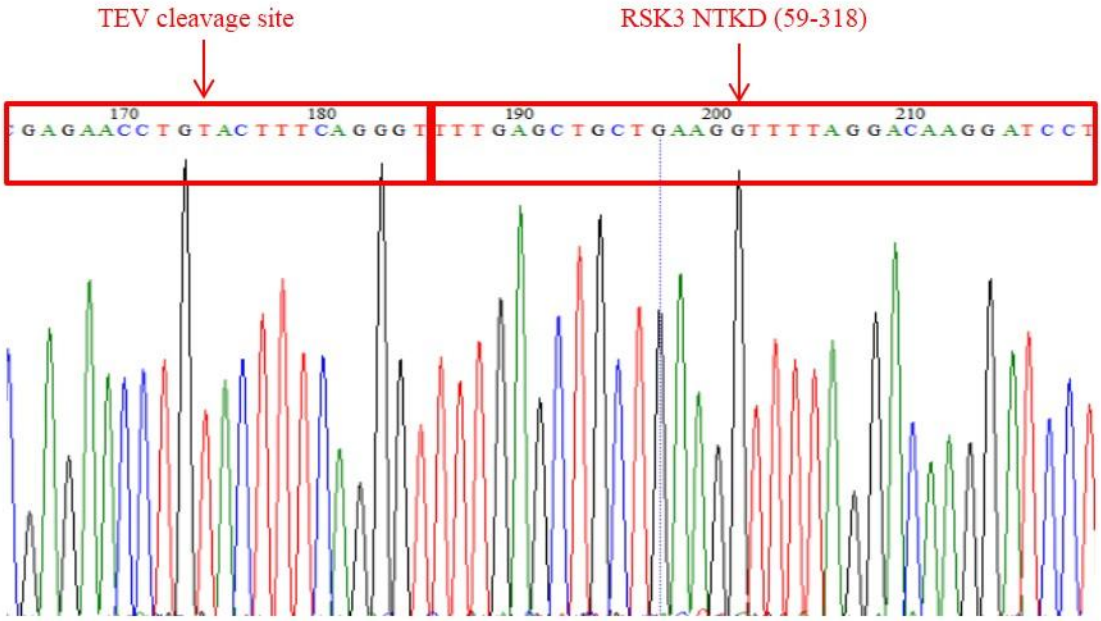


Figure 5.23: Sequencing of RSK3 NTKD (59-318). Sequencing data of RSK3 NTKD (59-318) /pET-28 a (+) recombinant construct with T7 promoter primer.

5.3.15. Expression and purification of RSK3 NTKD

The Molecular weight of RSK3 NTKD (59-318) is ~ 29 kDa and the fusion protein is ~ 33 kDa. The RSK3 NTKD (59-318) was expressed in Rosetta 2 (DE3) competent cells and purified by IMAC using Ni-NTA resin. Since RSK3 NTKD (59-318) goes into inclusion bodies, it is further purified under denaturating conditions using 8 M Urea. After separating the supernatant containing soluble fraction, the bacterial inclusion body pellet was thoroughly washed with buffer E additionally containing 1 % Triton X-100, 1 % Sarkosyl and 1 % Sodium Deoxycholate. Following these washes, the pellet was solubilized in 8 M Urea buffer (buffer F), and RSK3 NTKD (59-318) was purified under denaturating conditions using Ni-NTA resin. **Figure 5.24** shows purified RSK3 NTKD in 8 M Urea on a 10 % SDS-PAGE gel.

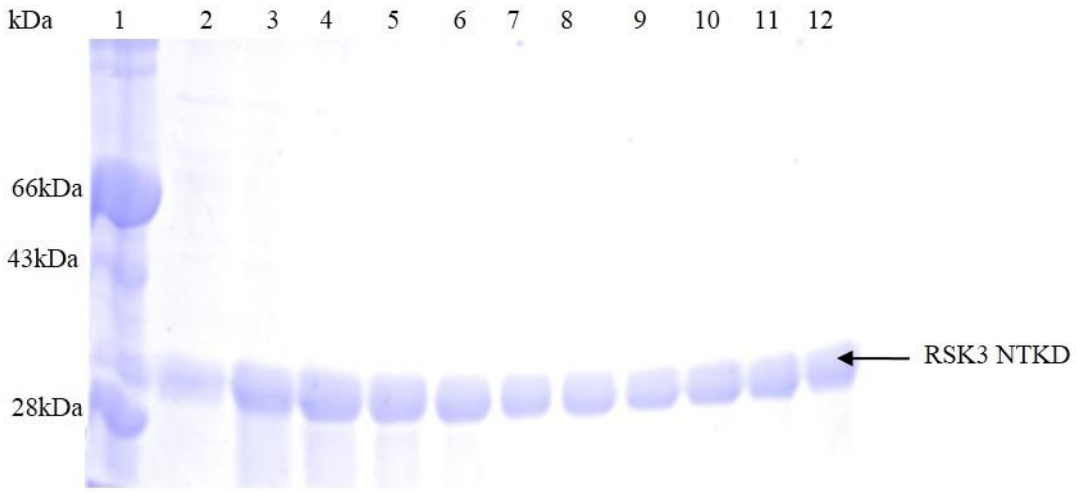


Figure 5.24: Purification of RSK3 NTKD (59-318) from inclusion bodies. The 10 % SDS-PAGE gel shows the 33 kDa RSK3 NTKD (59-318) with N-terminal HIS tag (lane 1: Protein marker, Lanes 2-8: 150 mM Imidazole elutions. Lanes 9-12: 300 mM Imidazole elutions).

5.3.16. Western blotting confirmation of expression of RSK3 CTKD and RSK3 NTKD in Drosophila Schneider s2 insect cell lines

The RSK3 CTKD (415-672) and RSK3 NTKD (59-318) had been cloned into pMT-wb-v5 vector, and transfected into Drosophila Schneider s2 insect cell lines. After inducing with cupric sulphate, the expression was checked by Western blotting using anti-HIS antibodies. **Figure 5.25** shows the X-ray film confirming the expression of RSK3 CTKD and RSK3 NTKD by Western blotting. The chemiluminescence was generated by the reaction between the HRP conjugated secondary antibody with its substrate.

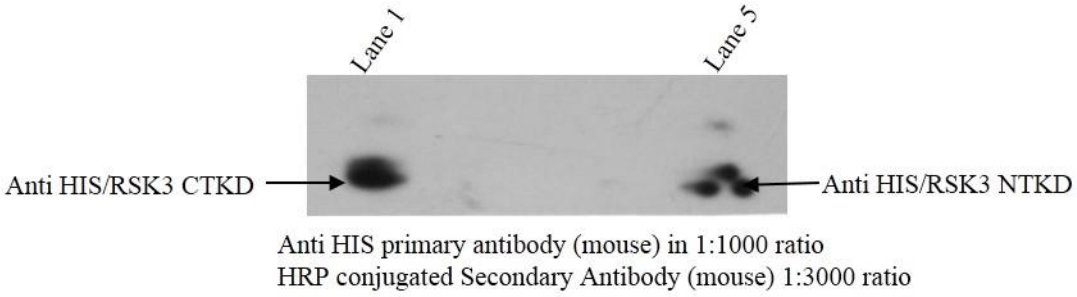


Figure 5.25 Western blot of RSK3 CTKD and RSK3 NTKD from Drosophila Schneider s2 Cells. The X-ray film shows the presence of 6 HIS-RSK3 CTKD (415-672) and 6 HIS-RSK3 NTKD (59-318) using anti HIS primary antibody and chemiluminescence was detected using HRP conjugated antibody and its substrate. (lane 1: RSK 3 CTKD, Lane 5: RSK 3 NTKD)

5.4. Conclusion

Expressing large proteins with soluble tags in *E. coli* can sometimes lead to co-expression of bacterial chaperones which firmly bind to the protein of interest. Various strategies including extensive stringent washes were employed to remove these chaperones. MBP tagged RSK3 CTKD (415-672) when expressed in *E. coli* system, co-expressed with groL1 60 kDa chaperonin 1. By stringent washes with wash buffer containing Glycerol, $MgCl_2$ and ATP, MBP-RSK3 CTKD (415-672) was successfully purified.

Kinases are very important proteins as they are the signal transducers within the cell. Expression and purification of complex human kinases in bacterial system is very challenging as they usually form inclusion bodies. Proteins from inclusion bodies need to be denatured first, and then refolded into native form. To prevent nonspecific aggregation and precipitation during refolding, various additives like glycerol, *l*-arginine, DTT, Triton X-100 etc. in various proportions were tried to determine the most suitable condition for refolding [288] [289]. RSK3 CTKD (415-672) was cloned, expressed and purified from *E. coli* inclusion bodies. The denatured protein was successfully refolded and was further validated

using CD, fluorescence spectroscopy and limited proteolysis. Using multi-disciplinary *in-vitro, in-silico* and biophysical approaches, it has been observed that the refolded protein is predominantly α-helical with a folded tertiary conformation. Furthermore, DLS suggest that RSK3 CTKD forms a homo-trimer. The thermal denaturation study revealed, that the T_m of RSK3 CTKD (415-672) is 52 °C, suggesting the protein is thermally stable. The homology model is in agreement with CD data as it is predominantly α-helical in nature. Further the docking experiment between the ERK2 and RSK3 CTKD including the C-terminal tail is in agreement with the earlier reported in-vitro results.

The RSK3 NTKD (59-318)/pET-28 a (+) was also insoluble when expressed in bacterial system. Hence, RSK3 NTKD (59-318) was purified under denaturing conditions using 8 M Urea.

Both the domains of RSK3, the CTKD and the NTKD were cloned into pMT-wb-v5 vector and transfected into Drosophila Schneider s2 insect cell lines. Stable cell lines were established by repeated cycles of adding fresh medium along with Blasticidin. The transfected insect cell lines were induced with cupric sulphate and the expression for these domains were found to be in the soluble fraction which was confirmed by western blotting analysis. Expressing these domains in insect system had provided the necessary PTMs at the required positions, thus making these proteins to be more stable and are expressed in the soluble fraction.

So far, no crystal structure is reported for RSK3 for any domain from any organism. Crystal structures of either full length RSK3 or its domains will help in understanding its complex structure.

Chapter 6:

Summary and future directions

6.1. Summary

MAPK pathway is one of the important signalling pathways which regulates most of the cellular functions within a eukaryotic cell. Kinases in this pathway are known to be mutated and highly unregulated in most of the human cancers. Different research groups are working relentlessly targeting to determine the atomic structures of each member of MAPK pathway to develop new potential small molecule drug leads. Since the three dimensional atomic structure of proteins plays an important role in drug discovery, it is worth to study the atomic structure of each protein of MAPK pathway. So far, crystal structures of MEK and ERK2 have been determined [67, 249, 262]. Albeit, the structures of individual functional domains of RSK for isoforms 1 and 2 had been determined. However, no full length structures are available for any of the RSK isoforms. Hence, full length structures of these RSKs will help in understanding the protein folding and inter domain movement that regulates the overall functionality of these proteins. Furthermore, studies of protein-protein interactions in this pathway will help us to understand the functions of different signalling cascade present in MAPK pathway.

In our studies, we have purified ERK2 from *E. coli* system and studied the folding patterns and protein-protein interactions using Circular Dichroism, fluorescence spectroscopy and Isothermal Titration Calorimetry (ITC). ITC was performed between purified ERK2 protein with RSK3 C-terminal tail peptide, which resulted revealed moderately low binding between them. This is in agreement with the previous reports that ERK2 binds with lower affinity with peptide substrates in comparison to full length substrates [259, 260]. Further, attempts were made to co-crystallize ERK2 with RSK3 C-terminal interacting peptide using Hampton crystallization screens 1 and 2. However, no crystals were obtained in any of the conditions. Hence, we decided to carry out *in-silico* docking study to visualise the important amino acids at the binding interface. The ERK binding motif in the available structures

within the N-terminal region MEK and in the C-terminal region of RSK1 CTKD were missing. Therefore, missing regions were modelled into tertiary structures. After validating these models, macromolecular docking studies were performed to assess the interacting residues between MEK1-ERK2 interface, and ERK2-RSKs interfaces. It has been observed that the residues from 314 to 329 of ERK2 interact with MEK1 and RSKs in a two pronged manner. These studies also showed the importance of steric hindrance as a major player in the dissociation of upstream kinases from the signalling complex in the MAPK pathway. ERK2 binds to all isoforms of RSKs with different binding affinities, and the differences in these binding affinities for various isozymes are required for triggering a specific type of interactions for transcription factor related cell signalling. Understanding these interactions between different MEK1-ERK2-RSKs complexes at atomic level is a subject of extensive studies for designing the small molecule inhibitors.

Ribosomal S6 Kinases are very large complex kinases in the human kinome, as they have two functionally distinct kinase domains, belonging to completely different families of protein kinases. The molecular mass of RSK1 is ~ 83 kDa and purification of such large proteins in bacterial *E. coli* system is possible with 6 HIS tag, as the molecular weights of GST and MBP fusion proteins are too large to express in bacterial system. Though there are many advantages of 6 HIS tag, its utility during protein purification depends up on its accessibility. The KYTE DOOLITTLE plot of first 18 amino acids of the N-terminal region of RSK1 full length revealed it to be hydrophobic, and hence it was eluting from Ni-NTA column at very low concentration of imidazole along with other bacterial contaminants, though the expression levels are good. To increase the binding affinity to Ni-NTA column and purify it to greater extent, primers were designed to engineer the 18 nucleotide sequence coding for 6 HIS residues at the C-terminal end. With this incorporation, considerable amount of RSK1 full length protein was purified. Further thermal denaturation studies

revealed that the protein has low T_m of 45.68 °C suggesting the proteins is less stable. After confirming its secondary and tertiary conformations by CD and fluorescence spectroscopy, attempts were made to crystallize the protein. However, attempts have not yielded crystals so far. Therefore RSK1 CTKD was cloned and expressed in Rosetta 2 (DE3) RSK1 CTKD/ pET-28 a (+) forms inclusion bodies when expressed in *E. coli* system and was purified under denaturation conditions using Ni-NTA column. RSK1 CTKD was also cloned in pMT-wb-v5 vector and transfected in drosophila Schneider s2 insect cell lines and subsequently stable cell lines were established.

The MBP tagged RSK3 CTKD was cloned in pMal-c2X vector and expressed in Rosetta 2 (DE3) competent cells. The MBP-RSK3 CTKD fusion protein, when expressed in *E. coli* system co-purifies with bacterial chaperones. Mass spectrometry data confirmed this chaperone as groL1 60 kDa chaperonin 1. Hence, a protocol to get rid of this chaperone contaminant was established by trial and error using ATP, $MgCl_2$ and glycerol in the wash buffer. It was fond that excessive washing by as many as 40 column washes were necessary to remove the contaminant. Crystallization attempts were performed to crystalize the fusion protein as the MBP tag could not be cleaved by the proteases.

The functional domains of RSK3 CTKD and RSK3 NTKD were cloned in pET-28 a (+) vector, and when expressed in *E. coli* system, they formed inclusion bodies. Proteins from inclusion bodies were denatured using 8 M urea, and then refolded back into native conformation. To prevent nonspecific aggregation and precipitation during refolding, various additives like glycerol, *l*-arginine, DTT, Triton X-100 etc. in various proportions were used to determine the most suitable condition for refolding. This refolding protocol can be employed to refold other complex kinases also. Using multi-model, *in-vitro, in-silico* and biophysical approaches, we have analysed the folding of RSK3 CTKD. Trypsin in-gel digestion and mass spectroscopy analysis confirmed the identity of the cloned region as CTKD of RSK3. The

Fluorescence spectroscopy data revealed the emission maximum at λ_{max} = 347 nm for denatured protein, whereas for refolded protein is emission maximum at λ_{max} = 332 nm. This blue shift of 15 nm suggests the RSK3 CTKD is properly folded into a compact tertiary structure, burying its tryptophans deep into the hydrophobic core of the protein. The CD data revealed that, the RSK3 CTKD is predominantly α-helical, which is characteristic of most of the kinases. Further, thermal denaturation studies revealed the T_m of RSK3 CTKD to be 52 °C (±0.24 °C), suggesting the protein is thermally stable. Slow rate of proteolysis in the limited proteolysis assay, suggests the tryptic cleavage sites were not easily accessible initially, which further strengthened the data that the protein is properly folded into a compact structure. DLS experiments after 8 acquisitions, revealed that RSK3 CTKD forms a homo trimer in solution with respect to its hydrodynamic radius and molecular mass. Further, refolded and FPLC purified RSK3 CTKD was used for setting crystallization trials manually, using sitting drop vapour diffusion method. Also, 600 different crystallization buffer conditions had been set for refolded RSK3 CTKD using Phoenix robotics (IIT-Bombay). But so far no crystal was seen in any of the conditions.

Hence, a model for RSK3 CTKD was built using the Modeller software. After selecting the best model, based on its DOPE scoring, the model was further validated by PROCHECK and ProSA webservers. The Ramachandran plot revealed, 94.7 % of the residues were in most favoured regions, 4.4 % residues in additionally allowed regions and 0.9 % in generously allowed regions, and no residue was found in disallowed regions. Further, the *z*-score for RSK3 CTKD model was calculated as -7.24 by the ProSA webserver, which is within the allowed range for correct structures. These PROCHECK and ProSA results validated the good quality of the model. Since the built model for RSK3 CTKD, does not include the C-terminal tail containing the ERK2 binding motif, another model was built for RSK3 CTKD including the entire C-terminal tail. This RSK3 CTKD and C-terminal tail

model was docked on ERK2 structure. These docking results revealed that the orientations of Asp318 and Asp321 of ERK2 and Arg308 and Arg309 of RSK3 CTKD are in close proximity at the docking interface.

Both the insoluble RSK3 CTKD (415-672) and RSK3 NTKD (59-318) were also cloned into pMT-wb-v5 vector and transfected into Drosophila Schneider s2 insect cell lines. Stable cell lines were established by repeated cycles of adding sterile medium along with Blasticidin. The transfected insect cell lines were induced with cupric sulphate, and the expression domains were found to be in the soluble fraction, which was further confirmed by western blotting analysis. This is, in sharp contrast to the *E. coli* system, where these domains form inclusion bodes. Expressing these domains in insect cell lines had provided the necessary PTMs at the required positions, thus making these proteins to be more stable in the soluble fraction.

6.2. Future directions

So far, no crystal structure is available for RSK3 from any organism. RSK3 is functionally distinct from other isoforms of RSK, because of its additional function as a tumour suppressor in ovarian cancer. Hence the three dimensional structure of this RSK3 will help us to better understand its overall structure and function. Since the functional domains NTKD and CTKD, expressed in drosophila Schneider s2 insect cells are soluble, these proteins can be produced in large amounts for crystallographic analysis. Hence, RSK3 CTKD and RSK3 NTKD should be expressed and purified from insect cell system and crystallization trials should be set. Crystal structures of either full length RSK3 or its domains will throw light on the reported multi-functionality of RSK3.

The RSK3 CTKD from inclusion bodies was purified and after refolding, could be obtained in concentrations as high as 30 mg/ml. Attempts to crystalize refolded RSK3 CTKD with inhibitors and ATP analogues can be made.

References:

1. Elenbaas, B. and R.A. Weinberg, *Heterotypic signaling between epithelial tumor cells and fibroblasts in carcinoma formation.* Exp Cell Res, 2001. **264**(1): p. 169-84.
2. Borden, E.C., et al., *Soft tissue sarcomas of adults: state of the translational science.* Clin Cancer Res, 2003. **9**(6): p. 1941-56.
3. Musilova, K. and M. Mraz, *MicroRNAs in B-cell lymphomas: how a complex biology gets more complex.* Leukemia, 2015. **29**(5): p. 1004-17.
4. Amitay, E.L. and L. Keinan-Boker, *Breastfeeding and Childhood Leukemia Incidence: A Meta-analysis and Systematic Review.* JAMA Pediatr, 2015. **169**(6): p. e151025.
5. Talerman, A. *Germ cell tumours*. in *Annales de pathologie*. 1984.
6. Jordan, C.T., M.L. Guzman, and M. Noble, *Cancer stem cells.* N Engl J Med, 2006. **355**(12): p. 1253-61.
7. Vogelstein, B. and K.W. Kinzler, *The multistep nature of cancer.* Trends Genet, 1993. **9**(4): p. 138-41.
8. Gao, C., et al., *Chromosome instability, chromosome transcriptome, and clonal evolution of tumor cell populations.* Proc Natl Acad Sci U S A, 2007. **104**(21): p. 8995-9000.
9. Hanahan, D. and R.A. Weinberg, *The hallmarks of cancer.* Cell, 2000. **100**(1): p. 57-70.
10. Ferlay, J., et al., *Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012.* Int J Cancer, 2015. **136**(5): p. E359-86.
11. Torre, L.A., et al., *Global cancer statistics, 2012.* CA Cancer J Clin, 2015. **65**(2): p. 87-108.
12. Mallath, M.K., et al., *The growing burden of cancer in India: epidemiology and social context.* Lancet Oncol, 2014. **15**(6): p. e205-12.
13. Dikshit, R., et al., *Cancer mortality in India: a nationally representative survey.* Lancet, 2012. **379**(9828): p. 1807-16.
14. Zhang, X., et al., *MAPK/ERK signaling pathway-induced hyper-O-GlcNAcylation enhances cancer malignancy.* Mol Cell Biochem, 2015.
15. Weng, A.P., et al., *Activating mutations of NOTCH1 in human T cell acute lymphoblastic leukemia.* Science, 2004. **306**(5694): p. 269-71.
16. Wang, Z., et al., *Down-regulation of Notch-1 contributes to cell growth inhibition and apoptosis in pancreatic cancer cells.* Mol Cancer Ther, 2006. **5**(3): p. 483-93.
17. Zhou, X.L., et al., *Downregulation of Dickkopf-1 is responsible for high proliferation of breast cancer cells via losing control of Wnt/beta-catenin signaling.* Acta Pharmacol Sin, 2010. **31**(2): p. 202-10.
18. Saito, T., et al., *Downregulation of sFRP-2 by epigenetic silencing activates the beta-catenin/Wnt signaling pathway in esophageal basaloid squamous cell carcinoma.* Virchows Arch, 2014. **464**(2): p. 135-43.
19. Garcia-Zaragoza, E., et al., *Intraepithelial paracrine Hedgehog signaling induces the expansion of ciliated cells that express diverse progenitor cell markers in the basal epithelium of the mouse mammary gland.* Dev Biol, 2012. **372**(1): p. 28-44.
20. Jeng, K.S., et al., *High expression of Sonic Hedgehog signaling pathway genes indicates a risk of recurrence of breast carcinoma.* Onco Targets Ther, 2013. **7**: p. 79-86.
21. Wolff, A.C., et al., *Recommendations for human epidermal growth factor receptor 2 testing in breast cancer: American Society of Clinical Oncology/College of American Pathologists clinical practice guideline update.* Arch Pathol Lab Med, 2014. **138**(2): p. 241-56.
22. Nwabo Kamdje, A.H., et al., *Signaling pathways in breast cancer: therapeutic targeting of the microenvironment.* Cell Signal, 2014. **26**(12): p. 2843-56.
23. Burotto, M., et al., *The MAPK pathway across different malignancies: a new perspective.* Cancer, 2014. **120**(22): p. 3446-56.
24. Sebolt-Leopold, J.S. and R. Herrera, *Targeting the mitogen-activated protein kinase cascade to treat cancer.* Nat Rev Cancer, 2004. **4**(12): p. 937-47.

25. Morris, E.J., et al., *Discovery of a novel ERK inhibitor with activity in models of acquired resistance to BRAF and MEK inhibitors.* Cancer Discov, 2013. **3**(7): p. 742-50.
26. Herrero, A., et al., *Small Molecule Inhibition of ERK Dimerization Prevents Tumorigenesis by RAS-ERK Pathway Oncogenes.* Cancer Cell, 2015. **28**(2): p. 170-82.
27. Uehling, D.E. and P.A. Harris, *Recent progress on MAP kinase pathway inhibitors.* Bioorg Med Chem Lett, 2015. **25**(19): p. 4047-56.
28. Regad, T., *Targeting RTK Signaling Pathways in Cancer.* Cancers (Basel), 2015. **7**(3): p. 1758-84.
29. Takashima, A. and D.V. Faller, *Targeting the RAS oncogene.* Expert Opin Ther Targets, 2013. **17**(5): p. 507-31.
30. Maurer, G., B. Tarkowski, and M. Baccarini, *Raf kinases in cancer-roles and therapeutic opportunities.* Oncogene, 2011. **30**(32): p. 3477-88.
31. Samatar, A.A. and P.I. Poulikakos, *Targeting RAS-ERK signalling in cancer: promises and challenges.* Nat Rev Drug Discov, 2014. **13**(12): p. 928-42.
32. Goetz, E.M., et al., *ERK mutations confer resistance to mitogen-activated protein kinase pathway inhibitors.* Cancer Res, 2014. **74**(23): p. 7079-89.
33. Malumbres, M. and M. Barbacid, *RAS oncogenes: the first 30 years.* Nat Rev Cancer, 2003. **3**(6): p. 459-65.
34. Lewis, T.S., P.S. Shapiro, and N.G. Ahn, *Signal transduction through MAP kinase cascades.* Adv Cancer Res, 1998. **74**: p. 49-139.
35. Wood, K.W., et al., *ras mediates nerve growth factor receptor modulation of three signal-transducing protein kinases: MAP kinase, Raf-1, and RSK.* Cell, 1992. **68**(6): p. 1041-50.
36. Geyer, M. and A. Wittinghofer, *GEFs, GAPs, GDIs and effectors: taking a closer (3D) look at the regulation of Ras-related GTP-binding proteins.* Curr Opin Struct Biol, 1997. **7**(6): p. 786-92.
37. Campbell, S.L., et al., *Increasing complexity of Ras signaling.* Oncogene, 1998. **17**(11 Reviews): p. 1395-413.
38. Chong, H., H.G. Vikis, and K.L. Guan, *Mechanisms of regulating the Raf kinase family.* Cell Signal, 2003. **15**(5): p. 463-9.
39. Terai, K. and M. Matsuda, *Ras binding opens c-Raf to expose the docking site for mitogen-activated protein kinase kinase.* EMBO Rep, 2005. **6**(3): p. 251-5.
40. Tarrega, C., et al., *ERK2 shows a restrictive and locally selective mechanism of recognition by its tyrosine phosphatase inactivators not shared by its activator MEK1.* J Biol Chem, 2005. **280**(45): p. 37885-94.
41. Poteet-Smith, C.E., et al., *Generation of constitutively active p90 ribosomal S6 kinase in vivo. Implications for the mitogen-activated protein kinase-activated protein kinase family.* J Biol Chem, 1999. **274**(32): p. 22135-8.
42. Jensen, C.J., et al., *90-kDa ribosomal S6 kinase is phosphorylated and activated by 3-phosphoinositide-dependent protein kinase-1.* J Biol Chem, 1999. **274**(38): p. 27168-76.
43. Mooz, J., et al., *Dimerization of the kinase ARAF promotes MAPK pathway activation and cell migration.* Sci Signal, 2014. **7**(337): p. ra73.
44. Yan, Y., et al., *Extracellular regulated protein kinases 1/2 phosphorylation is required for hepatic differentiation of human umbilical cord-derived mesenchymal stem cells.* Exp Biol Med (Maywood), 2015. **240**(4): p. 534-45.
45. Chen, R.H., C. Sarnecki, and J. Blenis, *Nuclear localization and regulation of erk- and rsk-encoded protein kinases.* Mol Cell Biol, 1992. **12**(3): p. 915-27.
46. De Cesare, D., et al., *Rsk-2 activity is necessary for epidermal growth factor-induced phosphorylation of CREB protein and transcription of c-fos gene.* Proc Natl Acad Sci U S A, 1998. **95**(21): p. 12202-7.
47. Frodin, M. and S. Gammeltoft, *Role and regulation of 90 kDa ribosomal S6 kinase (RSK) in signal transduction.* Mol Cell Endocrinol, 1999. **151**(1-2): p. 65-77.

48. Zhao, J., et al., *ERK-dependent phosphorylation of the transcription initiation factor TIF-IA is required for RNA polymerase I transcription and cell growth.* Mol Cell, 2003. **11**(2): p. 405-13.
49. Nielsen, P.J., G. Thomas, and J.L. Maller, *Increased phosphorylation of ribosomal protein S6 during meiotic maturation of Xenopus oocytes.* Proc Natl Acad Sci U S A, 1982. **79**(9): p. 2937-41.
50. Erikson, E. and J.L. Maller, *A protein kinase from Xenopus eggs specific for ribosomal protein S6.* Proc Natl Acad Sci U S A, 1985. **82**(3): p. 742-6.
51. Roux, P.P., et al., *Tumor-promoting phorbol esters and activated Ras inactivate the tuberous sclerosis tumor suppressor complex via p90 ribosomal S6 kinase.* Proc Natl Acad Sci U S A, 2004. **101**(37): p. 13489-94.
52. Fingar, D.C. and J. Blenis, *Target of rapamycin (TOR): an integrator of nutrient and growth factor signals and coordinator of cell growth and cell cycle progression.* Oncogene, 2004. **23**(18): p. 3151-71.
53. Palmer, A., A.C. Gavin, and A.R. Nebreda, *A link between MAP kinase and p34(cdc2)/cyclin B during oocyte maturation: p90(rsk) phosphorylates and inactivates the p34(cdc2) inhibitory kinase Myt1.* EMBO J, 1998. **17**(17): p. 5037-47.
54. Fujita, N., S. Sato, and T. Tsuruo, *Phosphorylation of p27Kip1 at threonine 198 by p90 ribosomal protein S6 kinases promotes its binding to 14-3-3 and cytoplasmic localization.* J Biol Chem, 2003. **278**(49): p. 49254-60.
55. Derijard, B., et al., *JNK1: a protein kinase stimulated by UV light and Ha-Ras that binds and phosphorylates the c-Jun activation domain.* Cell, 1994. **76**(6): p. 1025-37.
56. Bogoyevitch, M.A. and N.W. Court, *Counting on mitogen-activated protein kinases--ERKs 3, 4, 5, 6, 7 and 8.* Cell Signal, 2004. **16**(12): p. 1345-54.
57. Saba-El-Leil, M.K., et al., *An essential function of the mitogen-activated protein kinase Erk2 in mouse trophoblast development.* EMBO Rep, 2003. **4**(10): p. 964-8.
58. Yao, Y., et al., *Extracellular signal-regulated kinase 2 is necessary for mesoderm differentiation.* Proc Natl Acad Sci U S A, 2003. **100**(22): p. 12759-64.
59. Crews, C.M., A. Alessandrini, and R.L. Erikson, *The primary structure of MEK, a protein kinase that phosphorylates the ERK gene product.* Science, 1992. **258**(5081): p. 478-80.
60. Kosako, H., et al., *Xenopus MAP kinase activator is a serine/threonine/tyrosine kinase activated by threonine phosphorylation.* EMBO J, 1992. **11**(8): p. 2903-8.
61. Wu, J., et al., *Molecular structure of a protein-tyrosine/threonine kinase activating p42 mitogen-activated protein (MAP) kinase: MAP kinase kinase.* Proc Natl Acad Sci U S A, 1993. **90**(1): p. 173-7.
62. Zheng, C.F. and K.L. Guan, *Cloning and characterization of two distinct human extracellular signal-regulated kinase activator kinases, MEK1 and MEK2.* J Biol Chem, 1993. **268**(15): p. 11435-9.
63. Ferrell, J.E., Jr. and R.R. Bhatt, *Mechanistic studies of the dual phosphorylation of mitogen-activated protein kinase.* J Biol Chem, 1997. **272**(30): p. 19008-16.
64. Taylor, S.S. and A.P. Kornev, *Protein kinases: evolution of dynamic regulatory proteins.* Trends Biochem Sci, 2011. **36**(2): p. 65-77.
65. Roskoski, R., Jr., *ERK1/2 MAP kinases: structure, function, and regulation.* Pharmacol Res, 2012. **66**(2): p. 105-43.
66. Hanks, S.K., A.M. Quinn, and T. Hunter, *The protein kinase family: conserved features and deduced phylogeny of the catalytic domains.* Science, 1988. **241**(4861): p. 42-52.
67. Ohori, M., et al., *Identification of a selective ERK inhibitor and structural determination of the inhibitor-ERK2 complex.* Biochem Biophys Res Commun, 2005. **336**(1): p. 357-63.
68. Canagarajah, B.J., et al., *Activation mechanism of the MAP kinase ERK2 by dual phosphorylation.* Cell, 1997. **90**(5): p. 859-69.
69. Khokhlatchev, A.V., et al., *Phosphorylation of the MAP kinase ERK2 promotes its homodimerization and nuclear translocation.* Cell, 1998. **93**(4): p. 605-15.

70. Adachi, M., M. Fukuda, and E. Nishida, *Two co-existing mechanisms for nuclear import of MAP kinase: passive diffusion of a monomer and active transport of a dimer.* EMBO J, 1999. **18**(19): p. 5347-58.
71. Cobb, M.H. and E.J. Goldsmith, *Dimerization in MAP-kinase signaling.* Trends Biochem Sci, 2000. **25**(1): p. 7-9.
72. Owens, D.M. and S.M. Keyse, *Differential regulation of MAP kinase signalling by dual-specificity protein phosphatases.* Oncogene, 2007. **26**(22): p. 3203-13.
73. Bermudez, O., G. Pages, and C. Gimond, *The dual-specificity MAP kinase phosphatases: critical roles in development and cancer.* Am J Physiol Cell Physiol, 2010. **299**(2): p. C189-202.
74. Ohori, M., et al., *Role of a cysteine residue in the active site of ERK and the MAPKK family.* Biochem Biophys Res Commun, 2007. **353**(3): p. 633-7.
75. Anjum, R. and J. Blenis, *The RSK family of kinases: emerging roles in cellular signalling.* Nat Rev Mol Cell Biol, 2008. **9**(10): p. 747-58.
76. Deak, M., et al., *Mitogen- and stress-activated protein kinase-1 (MSK1) is directly activated by MAPK and SAPK2/p38, and may mediate activation of CREB.* EMBO J, 1998. **17**(15): p. 4426-41.
77. Jones, S.W., et al., *A Xenopus ribosomal protein S6 kinase has two apparent kinase domains that are each similar to distinct protein kinases.* Proc Natl Acad Sci U S A, 1988. **85**(10): p. 3377-81.
78. Blenis, J., et al., *Distinct mechanisms for the activation of the RSK kinases/MAP2 kinase/pp90rsk and pp70-S6 kinase signaling systems are indicated by inhibition of protein synthesis.* Cell Growth Differ, 1991. **2**(6): p. 279-85.
79. Bjorbaek, C., Y. Zhao, and D.E. Moller, *Divergent functional roles for p90rsk kinase domains.* J Biol Chem, 1995. **270**(32): p. 18848-52.
80. Fisher, T.L. and J. Blenis, *Evidence for two catalytically active kinase domains in pp90rsk.* Mol Cell Biol, 1996. **16**(3): p. 1212-9.
81. Vik, T.A. and J.W. Ryder, *Identification of serine 380 as the major site of autophosphorylation of Xenopus pp90rsk.* Biochem Biophys Res Commun, 1997. **235**(2): p. 398-402.
82. Lara, R., M.J. Seckl, and O.E. Pardo, *The p90 RSK family members: common functions and isoform specificity.* Cancer Res, 2013. **73**(17): p. 5301-8.
83. Dalby, K.N., et al., *Identification of regulatory phosphorylation sites in mitogen-activated protein kinase (MAPK)-activated protein kinase-1a/p90rsk that are inducible by MAPK.* J Biol Chem, 1998. **273**(3): p. 1496-505.
84. Roux, P.P., S.A. Richards, and J. Blenis, *Phosphorylation of p90 ribosomal S6 kinase (RSK) regulates extracellular signal-regulated kinase docking and RSK activity.* Mol Cell Biol, 2003. **23**(14): p. 4796-804.
85. Sutherland, C., D.G. Campbell, and P. Cohen, *Identification of insulin-stimulated protein kinase-1 as the rabbit equivalent of rskmo-2. Identification of two threonines phosphorylated during activation by mitogen-activated protein kinase.* Eur J Biochem, 1993. **212**(2): p. 581-8.
86. Smith, J.A., et al., *Identification of an extracellular signal-regulated kinase (ERK) docking site in ribosomal S6 kinase, a sequence critical for activation by ERK in vivo.* J Biol Chem, 1999. **274**(5): p. 2893-8.
87. Richards, S.A., et al., *Characterization of regulatory events associated with membrane targeting of p90 ribosomal S6 kinase 1.* Mol Cell Biol, 2001. **21**(21): p. 7470-80.
88. Reiter, A.K., et al., *The mTOR signaling pathway mediates control of ribosomal protein mRNA translation in rat liver.* Int J Biochem Cell Biol, 2004. **36**(11): p. 2169-79.
89. Holland, E.C., et al., *Signaling control of mRNA translation in cancer pathogenesis.* Oncogene, 2004. **23**(18): p. 3138-44.
90. Holz, M.K., et al., *mTOR and S6K1 mediate assembly of the translation preinitiation complex through dynamic protein interchange and ordered phosphorylation events.* Cell, 2005. **123**(4): p. 569-80.

91. Manning, B.D. and L.C. Cantley, *AKT/PKB signaling: navigating downstream.* Cell, 2007. **129**(7): p. 1261-74.
92. Vanhaesebroeck, B., et al., *The emerging mechanisms of isoform-specific PI3K signalling.* Nat Rev Mol Cell Biol, 2010. **11**(5): p. 329-41.
93. Gao, X. and D. Pan, *TSC1 and TSC2 tumor suppressors antagonize insulin signaling in cell growth.* Genes Dev, 2001. **15**(11): p. 1383-92.
94. Inoki, K., et al., *TSC2 is phosphorylated and inhibited by Akt and suppresses mTOR signalling.* Nat Cell Biol, 2002. **4**(9): p. 648-57.
95. Garami, A., et al., *Insulin activation of Rheb, a mediator of mTOR/S6K/4E-BP signaling, is inhibited by TSC1 and 2.* Mol Cell, 2003. **11**(6): p. 1457-66.
96. Inoki, K., et al., *Rheb GTPase is a direct target of TSC2 GAP activity and regulates mTOR signaling.* Genes Dev, 2003. **17**(15): p. 1829-34.
97. Dunlop, E.A. and A.R. Tee, *Mammalian target of rapamycin complex 1: signalling inputs, substrates and feedback mechanisms.* Cell Signal, 2009. **21**(6): p. 827-35.
98. Shahbazian, D., et al., *The mTOR/PI3K and MAPK pathways converge on eIF4B to control its phosphorylation and activity.* EMBO J, 2006. **25**(12): p. 2781-91.
99. Raught, B., et al., *Phosphorylation of eucaryotic translation initiation factor 4B Ser422 is modulated by S6 kinases.* EMBO J, 2004. **23**(8): p. 1761-9.
100. Rozen, F., et al., *Bidirectional RNA helicase activity of eucaryotic translation initiation factors 4A and 4F.* Mol Cell Biol, 1990. **10**(3): p. 1134-44.
101. Ikuta, M., et al., *Crystal structures of the N-terminal kinase domain of human RSK1 bound to three different ligands: Implications for the design of RSK1 specific inhibitors.* Protein Sci, 2007. **16**(12): p. 2626-35.
102. Malakhova, M., et al., *Structural diversity of the active N-terminal kinase domain of p90 ribosomal S6 kinase 2.* PLoS One, 2009. **4**(11): p. e8044.
103. Malakhova, M., et al., *Structural basis for activation of the autoinhibitory C-terminal kinase domain of p90 RSK2.* Nat Struct Mol Biol, 2008. **15**(1): p. 112-3.
104. Li, D., et al., *Structural basis for the autoinhibition of the C-terminal kinase domain of human RSK1.* Acta Crystallogr D Biol Crystallogr, 2012. **68**(Pt 6): p. 680-5.
105. Serafimova, I.M., et al., *Reversible targeting of noncatalytic cysteines with chemically tuned electrophiles.* Nat Chem Biol, 2012. **8**(5): p. 471-6.
106. Smith, K.J., et al., *The structure of MSK1 reveals a novel autoinhibitory conformation for a dual kinase protein.* Structure, 2004. **12**(6): p. 1067-77.
107. Laskowski, R.A. and M.B. Swindells, *LigPlot+: multiple ligand-protein interaction diagrams for drug discovery.* J Chem Inf Model, 2011. **51**(10): p. 2778-86.
108. Lowry, B., J.R. Miller, and F.C. Fraser, *A new dominant gene mental retardation syndrome. Association with small stature, tapering fingers, characteristic facies, and possible hydrocephalus.* Am J Dis Child, 1971. **121**(6): p. 496-500.
109. Abidi, F., et al., *Novel mutations in Rsk-2, the gene for Coffin-Lowry syndrome (CLS).* Eur J Hum Genet, 1999. **7**(1): p. 20-6.
110. Jacquot, S., et al., *X-linked Coffin-Lowry syndrome (CLS, MIM 303600, RPS6KA3 gene, protein product known under various names: pp90(rsk2), RSK2, ISPK, MAPKAP1).* Eur J Hum Genet, 2002. **10**(1): p. 2-5.
111. Delaunoy, J.P., et al., *Identification of novel mutations in the RSK2 gene (RPS6KA3) in patients with Coffin-Lowry syndrome.* Clin Genet, 2006. **70**(2): p. 161-6.
112. Clark, D.E., et al., *The serine/threonine protein kinase, p90 ribosomal S6 kinase, is an important regulator of prostate cancer cell proliferation.* Cancer Res, 2005. **65**(8): p. 3108-16.
113. Thakur, A., et al., *Aberrant expression of X-linked genes RbAp46, Rsk4, and Cldn2 in breast cancer.* Mol Cancer Res, 2007. **5**(2): p. 171-81.

114. Kang, S., et al., *FGFR3 activates RSK2 to mediate hematopoietic transformation through tyrosine phosphorylation of RSK2 and activation of the MEK/ERK pathway.* Cancer Cell, 2007. **12**(3): p. 201-14.
115. Lawrie, A.M., et al., *Protein kinase inhibition by staurosporine revealed in details of the molecular interaction with CDK2.* Nat Struct Biol, 1997. **4**(10): p. 796-801.
116. Alessi, D.R., *The protein kinase C inhibitors Ro 318220 and GF 109203X are equally potent inhibitors of MAPKAP kinase-1beta (Rsk-2) and p70 S6 kinase.* FEBS Lett, 1997. **402**(2-3): p. 121-3.
117. Brehmer, D., et al., *Proteome-wide identification of cellular targets affected by bisindolylmaleimide-type protein kinase C inhibitors.* Mol Cell Proteomics, 2004. **3**(5): p. 490-500.
118. Nguyen, T.L., et al., *Homology model of RSK2 N-terminal kinase domain, structure-based identification of novel RSK2 inhibitors, and preliminary common pharmacophore.* Bioorg Med Chem, 2006. **14**(17): p. 6097-105.
119. Nguyen, T.L., *Targeting RSK: an overview of small molecule inhibitors.* Anticancer Agents Med Chem, 2008. **8**(7): p. 710-6.
120. Smith, J.A., et al., *Identification of the first specific inhibitor of p90 ribosomal S6 kinase (RSK) reveals an unexpected role for RSK in cancer cell proliferation.* Cancer Res, 2005. **65**(3): p. 1027-34.
121. Smith, J.A., et al., *Structural basis for the activity of the RSK-specific inhibitor, SL0101.* Bioorg Med Chem, 2007. **15**(14): p. 5018-34.
122. Smith, J.A., et al., *Influence of rhamnose substituents on the potency of SL0101, an inhibitor of the Ser/Thr kinase, RSK.* Bioorg Med Chem, 2006. **14**(17): p. 6034-42.
123. Sapkota, G.P., et al., *BI-D1870 is a specific inhibitor of the p90 RSK (ribosomal S6 kinase) isoforms in vitro and in vivo.* Biochem J, 2007. **401**(1): p. 29-38.
124. Bain, J., et al., *The selectivity of protein kinase inhibitors: a further update.* Biochem J, 2007. **408**(3): p. 297-315.
125. Cohen, M.S., et al., *Structural bioinformatics-based design of selective, irreversible kinase inhibitors.* Science, 2005. **308**(5726): p. 1318-21.
126. Cohen, M.S., H. Hadjivassiliou, and J. Taunton, *A clickable inhibitor reveals context-dependent autoactivation of p90 RSK.* Nat Chem Biol, 2007. **3**(3): p. 156-60.
127. Leighton, I.A., et al., *Comparison of the specificities of p70 S6 kinase and MAPKAP kinase-1 identifies a relatively specific substrate for p70 S6 kinase: the N-terminal kinase domain of MAPKAP kinase-1 is essential for peptide phosphorylation.* FEBS Lett, 1995. **375**(3): p. 289-93.
128. Romeo, Y., X. Zhang, and P.P. Roux, *Regulation and function of the RSK family of protein kinases.* Biochem J, 2012. **441**(2): p. 553-69.
129. Saiki, R.K., et al., *Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia.* Science, 1985. **230**(4732): p. 1350-4.
130. Mullis, K.B., *The unusual origin of the polymerase chain reaction.* Sci Am, 1990. **262**(4): p. 56-61, 64-5.
131. Bartlett, J.M. and D. Stirling, *A short history of the polymerase chain reaction.* Methods Mol Biol, 2003. **226**: p. 3-6.
132. Saiki, R.K., et al., *Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase.* Science, 1988. **239**(4839): p. 487-91.
133. Rychlik, W., W.J. Spencer, and R.E. Rhoads, *Optimization of the annealing temperature for DNA amplification in vitro.* Nucleic Acids Res, 1990. **18**(21): p. 6409-12.
134. Stellwagen, N.C., *Accurate molecular weight determinations of deoxyribonucleic acid restriction fragments on agarose gels.* Biochemistry, 1983. **22**(26): p. 6180-5.
135. Aaij, C. and P. Borst, *The gel electrophoresis of DNA.* Biochim Biophys Acta, 1972. **269**(2): p. 192-200.

136. Smith, S.B., P.K. Aldridge, and J.B. Callis, *Observation of individual DNA molecules undergoing gel electrophoresis.* Science, 1989. **243**(4888): p. 203-6.
137. Meyers, J.A., et al., *Simple agarose gel electrophoretic method for the identification and characterization of plasmid deoxyribonucleic acid.* J Bacteriol, 1976. **127**(3): p. 1529-37.
138. Waring, M.J., *Complex formation between ethidium bromide and nucleic acids.* J Mol Biol, 1965. **13**(1): p. 269-82.
139. Meselson, M. and R. Yuan, *DNA restriction enzyme from E. coli.* Nature, 1968. **217**(5134): p. 1110-4.
140. Boyer, H.W. and D. Roulland-Dussoix, *A complementation analysis of the restriction and modification of DNA in Escherichia coli.* J Mol Biol, 1969. **41**(3): p. 459-72.
141. Roberts, R.J., *Restriction endonucleases.* CRC Crit Rev Biochem, 1976. **4**(2): p. 123-64.
142. Roberts, R.J., *How restriction enzymes became the workhorses of molecular biology.* Proc Natl Acad Sci U S A, 2005. **102**(17): p. 5905-8.
143. Lederberg, J., *Cell genetics and hereditary symbiosis.* Physiol Rev, 1952. **32**(4): p. 403-30.
144. Dugaiczyk, A., H.W. Boyer, and H.M. Goodman, *Ligation of EcoRI endonuclease-generated DNA fragments into linear and circular structures.* J Mol Biol, 1975. **96**(1): p. 171-84.
145. Raae, A.J., R.K. Kleppe, and K. Kleppe, *Kinetics and effect of salts and polyamines on T4 polynucleotide ligase.* Eur J Biochem, 1975. **60**(2): p. 437-43.
146. Hanahan, D., *Studies on transformation of Escherichia coli with plasmids.* J Mol Biol, 1983. **166**(4): p. 557-80.
147. Chen, I. and D. Dubnau, *DNA uptake during bacterial transformation.* Nat Rev Microbiol, 2004. **2**(3): p. 241-9.
148. Solomon, J.M. and A.D. Grossman, *Who's competent and when: regulation of natural genetic competence in bacteria.* Trends Genet, 1996. **12**(4): p. 150-5.
149. Panja, S., et al., *How does plasmid DNA penetrate cell membranes in artificial transformation process of Escherichia coli?* Mol Membr Biol, 2008. **25**(5): p. 411-22.
150. Sanger, F. and A.R. Coulson, *A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase.* J Mol Biol, 1975. **94**(3): p. 441-8.
151. Sanger, F., S. Nicklen, and A.R. Coulson, *DNA sequencing with chain-terminating inhibitors.* Proc Natl Acad Sci U S A, 1977. **74**(12): p. 5463-7.
152. Smith, L.M., et al., *Fluorescence detection in automated DNA sequence analysis.* Nature, 1986. **321**(6071): p. 674-9.
153. Flavell, R.A., et al., *Site-directed mutagenesis: effect of an extracistronic mutation on the in vitro propagation of bacteriophage Qbeta RNA.* Proc Natl Acad Sci U S A, 1975. **72**(1): p. 367-71.
154. Hutchison, C.A., 3rd, et al., *Mutagenesis at a specific position in a DNA sequence.* J Biol Chem, 1978. **253**(18): p. 6551-60.
155. Shortle, D., D. DiMaio, and D. Nathans, *Directed mutagenesis.* Annu Rev Genet, 1981. **15**: p. 265-94.
156. Bunch, T.A., Y. Grinblat, and L.S. Goldstein, *Characterization and use of the Drosophila metallothionein promoter in cultured Drosophila melanogaster cells.* Nucleic Acids Res, 1988. **16**(3): p. 1043-61.
157. Casali, N., *Escherichia coli host strains.* Methods Mol Biol, 2003. **235**: p. 27-48.
158. Uhlen, M., *Affinity as a tool in life science.* Biotechniques, 2008. **44**(5): p. 649-54.
159. Hochuli, E., H. Dobeli, and A. Schacher, *New metal chelate adsorbent selective for proteins and peptides containing neighbouring histidine residues.* J Chromatogr, 1987. **411**: p. 177-84.
160. Janknecht, R., et al., *Rapid and efficient purification of native histidine-tagged protein expressed by recombinant vaccinia virus.* Proc Natl Acad Sci U S A, 1991. **88**(20): p. 8972-6.
161. Hengen, P., *Purification of His-Tag fusion proteins from Escherichia coli.* Trends Biochem Sci, 1995. **20**(7): p. 285-6.

162. Bornhorst, J.A. and J.J. Falke, *Purification of proteins using polyhistidine affinity tags.* Methods Enzymol, 2000. **326**: p. 245-54.
163. Porath, J., *Immobilized metal ion affinity chromatography.* Protein Expr Purif, 1992. **3**(4): p. 263-81.
164. Smith, D.B. and K.S. Johnson, *Single-step purification of polypeptides expressed in Escherichia coli as fusions with glutathione S-transferase.* Gene, 1988. **67**(1): p. 31-40.
165. Harper, S. and D.W. Speicher, *Purification of proteins fused to glutathione S-transferase.* Methods Mol Biol, 2011. **681**: p. 259-80.
166. Kapust, R.B. and D.S. Waugh, *Escherichia coli maltose-binding protein is uncommonly effective at promoting the solubility of polypeptides to which it is fused.* Protein Sci, 1999. **8**(8): p. 1668-74.
167. Fox, J.D. and D.S. Waugh, *Maltose-binding protein as a solubility enhancer.* Methods Mol Biol, 2003. **205**: p. 99-117.
168. Kellermann, O.K. and T. Ferenci, *Maltose-binding protein from Escherichia coli.* Methods Enzymol, 1982. **90 Pt E**: p. 459-63.
169. Porath, J. and P. Flodin, *Gel filtration: a method for desalting and group separation.* Nature, 1959. **183**(4676): p. 1657-9.
170. Barth, H.G., B.E. Boyes, and C. Jackson, *Size exclusion chromatography.* Anal Chem, 1994. **66**(12): p. 595R-620R.
171. Verbeke, K. and A. Verbruggen, *Usefulness of fast protein liquid chromatography as an alternative to high performance liquid chromatography of 99mTc-labelled human serum albumin preparations.* J Pharm Biomed Anal, 1996. **14**(8-10): p. 1209-13.
172. Rossomando, E.F., *Ion-exchange chromatography.* Methods Enzymol, 1990. **182**: p. 309-17.
173. Queiroz, J.A., C.T. Tomaz, and J.M. Cabral, *Hydrophobic interaction chromatography of proteins.* J Biotechnol, 2001. **87**(2): p. 143-59.
174. Price, P., *Standard definitions of terms relating to mass spectrometry : A report from the committee on measurements and standards of the American society for mass spectrometry.* J Am Soc Mass Spectrom, 1991. **2**(4): p. 336-48.
175. Karas, M. and R. Kruger, *Ion formation in MALDI: the cluster ionization mechanism.* Chem Rev, 2003. **103**(2): p. 427-40.
176. Mahmood, T. and P.C. Yang, *Western blot: technique, theory, and trouble shooting.* N Am J Med Sci, 2012. **4**(9): p. 429-34.
177. Towbin, H., T. Staehelin, and J. Gordon, *Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications.* Proc Natl Acad Sci U S A, 1979. **76**(9): p. 4350-4.
178. Kelly, S.M., T.J. Jess, and N.C. Price, *How to study proteins by circular dichroism.* Biochim Biophys Acta, 2005. **1751**(2): p. 119-39.
179. Greenfield, N.J., *Using circular dichroism spectra to estimate protein secondary structure.* Nat Protoc, 2006. **1**(6): p. 2876-90.
180. Woody, R.W. and A. Koslowski, *Recent developments in the electronic spectroscopy of amides and alpha-helical polypeptides.* Biophys Chem, 2002. **101-102**: p. 535-51.
181. Andrews, L.J. and L.S. Forster, *Protein difference spectra. Effect of solvent and charge on tryptophan.* Biochemistry, 1972. **11**(10): p. 1875-9.
182. Vivian, J.T. and P.R. Callis, *Mechanisms of tryptophan fluorescence shifts in proteins.* Biophys J, 2001. **80**(5): p. 2093-109.
183. Chen, Y. and M.D. Barkley, *Toward understanding tryptophan fluorescence in proteins.* Biochemistry, 1998. **37**(28): p. 9976-82.
184. Freire, E., *Thermal denaturation methods in the study of protein folding.* Methods Enzymol, 1995. **259**: p. 144-68.
185. Fontana, A., et al., *Probing protein structure by limited proteolysis.* Acta Biochim Pol, 2004. **51**(2): p. 299-321.

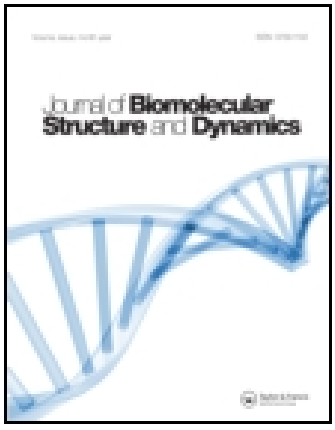
186. Lorber, B., et al., *Protein analysis by dynamic light scattering: methods and techniques for students.* Biochem Mol Biol Educ, 2012. **40**(6): p. 372-82.
187. Phizicky, E.M. and S. Fields, *Protein-protein interactions: methods for detection and analysis.* Microbiol Rev, 1995. **59**(1): p. 94-123.
188. Berggard, T., S. Linse, and P. James, *Methods for the detection and analysis of protein-protein interactions.* Proteomics, 2007. **7**(16): p. 2833-42.
189. Brymora, A., V.A. Valova, and P.J. Robinson, *Protein-protein interactions identified by pull-down experiments and mass spectrometry.* Curr Protoc Cell Biol, 2004. **Chapter 17**: p. Unit 17 5.
190. Pierce, M.M., C.S. Raman, and B.T. Nall, *Isothermal titration calorimetry of protein-protein interactions.* Methods, 1999. **19**(2): p. 213-21.
191. Zhou, X., R.M. Kini, and J. Sivaraman, *Application of isothermal titration calorimetry and column chromatography for identification of biomolecular targets.* Nat Protoc, 2011. **6**(2): p. 158-65.
192. Patching, S.G., *Surface plasmon resonance spectroscopy for characterisation of membrane protein-ligand interactions and its potential for drug discovery.* Biochim Biophys Acta, 2014. **1838**(1 Pt A): p. 43-55.
193. Kroemer, R.T., *Structure-based drug design: docking and scoring.* Curr Protein Pept Sci, 2007. **8**(4): p. 312-28.
194. Janin, J., et al., *CAPRI: a Critical Assessment of PRedicted Interactions.* Proteins, 2003. **52**(1): p. 2-9.
195. Hwang, H., et al., *Performance of ZDOCK and ZRANK in CAPRI rounds 13-19.* Proteins, 2010. **78**(15): p. 3104-10.
196. Chen, R., L. Li, and Z. Weng, *ZDOCK: an initial-stage protein-docking algorithm.* Proteins, 2003. **52**(1): p. 80-7.
197. Pierce, B. and Z. Weng, *ZRANK: reranking protein docking predictions with an optimized energy function.* Proteins, 2007. **67**(4): p. 1078-86.
198. Berman, H.M., *The Protein Data Bank: a historical perspective.* Acta Crystallogr A, 2008. **64**(Pt 1): p. 88-95.
199. Williamson, M.P., *NMR of proteins.* Nat Prod Rep, 1993. **10**(3): p. 207-32.
200. Ito, K., et al., *A 3 x 6 arrayed CCD X-ray detector for continuous rotation method in macromolecular crystallography.* J Synchrotron Radiat, 2007. **14**(Pt 1): p. 144-50.
201. Battye, T.G., et al., *iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM.* Acta Crystallogr D Biol Crystallogr, 2011. **67**(Pt 4): p. 271-81.
202. Winn, M.D., et al., *Overview of the CCP4 suite and current developments.* Acta Crystallogr D Biol Crystallogr, 2011. **67**(Pt 4): p. 235-42.
203. Ramakrishnan, V. and V. Biou, *Treatment of multiwavelength anomalous diffraction data as a special case of multiple isomorphous replacement.* Methods Enzymol, 1997. **276**: p. 538-57.
204. Rossmann, M.G., *The molecular replacement method.* Acta Crystallogr A, 1990. **46 (Pt 2)**: p. 73-82.
205. Wang, R.Y., et al., *De novo protein structure determination from near-atomic-resolution cryo-EM maps.* Nat Methods, 2015. **12**(4): p. 335-8.
206. Sun, L., et al., *Cryo-EM structure of the bacteriophage T4 portal protein assembly at near-atomic resolution.* Nat Commun, 2015. **6**: p. 7548.
207. Callaway, E., *The revolution will not be crystallized: a new method sweeps through structural biology.* Nature, 2015. **525**(7568): p. 172-4.
208. Bartesaghi, A., et al., *2.2 A resolution cryo-EM structure of beta-galactosidase in complex with a cell-permeant inhibitor.* Science, 2015. **348**(6239): p. 1147-51.
209. Kuhlbrandt, W., *Cryo-EM enters a new era.* Elife, 2014. **3**: p. e03678.

210. Krieger, E., S.B. Nabuurs, and G. Vriend, *Homology modeling.* Methods Biochem Anal, 2003. **44**: p. 509-23.
211. Moult, J., et al., *A large-scale experiment to assess protein structure prediction methods.* Proteins, 1995. **23**(3): p. ii-v.
212. Tramontano, A., et al., *The assessment of methods for protein structure prediction.* Methods Mol Biol, 2008. **413**: p. 43-57.
213. Yang, J., et al., *The I-TASSER Suite: protein structure and function prediction.* Nat Methods, 2015. **12**(1): p. 7-8.
214. Roy, A., A. Kucukural, and Y. Zhang, *I-TASSER: a unified platform for automated protein structure and function prediction.* Nat Protoc, 2010. **5**(4): p. 725-38.
215. Wu, S. and Y. Zhang, *LOMETS: a local meta-threading-server for protein structure prediction.* Nucleic Acids Res, 2007. **35**(10): p. 3375-82.
216. Zhang, Y., D. Kihara, and J. Skolnick, *Local energy landscape flattening: parallel hyperbolic Monte Carlo sampling of protein folding.* Proteins, 2002. **48**(2): p. 192-201.
217. Zhang, Y. and J. Skolnick, *SPICKER: a clustering approach to identify near-native protein folds.* J Comput Chem, 2004. **25**(6): p. 865-71.
218. Xu, D. and Y. Zhang, *Improving the physical realism and structural accuracy of protein models by a two-step atomic-level energy minimization.* Biophys J, 2011. **101**(10): p. 2525-34.
219. Fiser, A. and A. Sali, *Modeller: generation and refinement of homology-based protein structure models.* Methods Enzymol, 2003. **374**: p. 461-91.
220. Kim, D.E., D. Chivian, and D. Baker, *Protein structure prediction and analysis using the Robetta server.* Nucleic Acids Res, 2004. **32**(Web Server issue): p. W526-31.
221. Schwede, T., et al., *SWISS-MODEL: An automated protein homology-modeling server.* Nucleic Acids Res, 2003. **31**(13): p. 3381-5.
222. Das, R., et al., *Structure prediction for CASP7 targets using extensive all-atom refinement with Rosetta@home.* Proteins, 2007. **69 Suppl 8**: p. 118-28.
223. Gartner, A., K. Nasmyth, and G. Ammerer, *Signal transduction in Saccharomyces cerevisiae requires tyrosine and threonine phosphorylation of FUS3 and KSS1.* Genes Dev, 1992. **6**(7): p. 1280-92.
224. Mabuchi, S., et al., *Tamoxifen inhibits cell proliferation via mitogen-activated protein kinase cascades in human ovarian cancer cell lines in a manner not dependent on the expression of estrogen receptor or the sensitivity to cisplatin.* Endocrinology, 2004. **145**(3): p. 1302-13.
225. Igata, M., et al., *Adenosine monophosphate-activated protein kinase suppresses vascular smooth muscle cell proliferation through the inhibition of cell cycle progression.* Circ Res, 2005. **97**(8): p. 837-44.
226. Satomi, Y. and H. Nishino, *Implication of mitogen-activated protein kinase in the induction of G1 cell cycle arrest and gadd45 expression by the carotenoid fucoxanthin in human cancer cells.* Biochim Biophys Acta, 2009. **1790**(4): p. 260-6.
227. Johnson, G.L. and R. Lapadat, *Mitogen-activated protein kinase pathways mediated by ERK, JNK, and p38 protein kinases.* Science, 2002. **298**(5600): p. 1911-2.
228. Raman, M., W. Chen, and M.H. Cobb, *Differential regulation and properties of MAPKs.* Oncogene, 2007. **26**(22): p. 3100-12.
229. Yoon, S. and R. Seger, *The extracellular signal-regulated kinase: multiple substrates regulate diverse cellular functions.* Growth Factors, 2006. **24**(1): p. 21-44.
230. Cooper, J.A., et al., *Similar effects of platelet-derived growth factor and epidermal growth factor on the phosphorylation of tyrosine in cellular proteins.* Cell, 1982. **31**(1): p. 263-73.
231. Ray, L.B. and T.W. Sturgill, *Insulin-stimulated microtubule-associated protein kinase is phosphorylated on tyrosine and threonine in vivo.* Proc Natl Acad Sci U S A, 1988. **85**(11): p. 3753-7.
232. Adams, J.P. and J.D. Sweatt, *Molecular psychology: roles for the ERK MAP kinase cascade in memory.* Annu Rev Pharmacol Toxicol, 2002. **42**: p. 135-63.

233. Milella, M., S.M. Kornblau, and M. Andreeff, *The mitogen-activated protein kinase signaling module as a therapeutic target in hematologic malignancies.* Rev Clin Exp Hematol, 2003. **7**(2): p. 160-90.
234. Hatano, N., et al., *Essential role for ERK2 mitogen-activated protein kinase in placental development.* Genes Cells, 2003. **8**(11): p. 847-56.
235. Bourguignon, L.Y., et al., *Hyaluronan-CD44 interaction with IQGAP1 promotes Cdc42 and ERK signaling, leading to actin binding, Elk-1/estrogen receptor transcriptional activation, and ovarian cancer progression.* J Biol Chem, 2005. **280**(12): p. 11961-72.
236. Chen, C.H., et al., *Bidirectional signals transduced by DAPK-ERK interaction promote the apoptotic effect of DAPK.* EMBO J, 2005. **24**(2): p. 294-304.
237. Calvisi, D.F., et al., *Dual-specificity phosphatase 1 ubiquitination in extracellular signal-regulated kinase-mediated control of growth in human hepatocellular carcinoma.* Cancer Res, 2008. **68**(11): p. 4192-200.
238. Ding, Q., et al., *Erk associates with and primes GSK-3beta for its inactivation resulting in upregulation of beta-catenin.* Mol Cell, 2005. **19**(2): p. 159-70.
239. Pircher, T.J., et al., *Extracellular signal-regulated kinase (ERK) interacts with signal transducer and activator of transcription (STAT) 5a.* Mol Endocrinol, 1999. **13**(4): p. 555-65.
240. Formstecher, E., et al., *PEA-15 mediates cytoplasmic sequestration of ERK MAP kinase.* Dev Cell, 2001. **1**(2): p. 239-50.
241. Carriere, A., et al., *ERK1/2 phosphorylate Raptor to promote Ras-dependent activation of mTOR complex 1 (mTORC1).* J Biol Chem, 2011. **286**(1): p. 567-77.
242. Zhou, X., et al., *Histone deacetylase 4 associates with extracellular signal-regulated kinases 1 and 2, and its cellular localization is regulated by oncogenic Ras.* Proc Natl Acad Sci U S A, 2000. **97**(26): p. 14329-33.
243. Warren, D.T., et al., *Novel nuclear nesprin-2 variants tether active extracellular signal-regulated MAPK1 and MAPK2 at promyelocytic leukemia protein nuclear bodies and act to regulate smooth muscle cell proliferation.* J Biol Chem, 2010. **285**(2): p. 1311-20.
244. Haycock, J.W., et al., *ERK1 and ERK2, two microtubule-associated protein 2 kinases, mediate the phosphorylation of tyrosine hydroxylase at serine-31 in situ.* Proc Natl Acad Sci U S A, 1992. **89**(6): p. 2365-9.
245. Burdick, D.J., et al., *Fragment-based discovery of potent ERK2 pyrrolopyrazine inhibitors.* Bioorg Med Chem Lett, 2015. **25**(21): p. 4728-32.
246. Ward, R.A., et al., *Structure-Guided Design of Highly Selective and Potent Covalent Inhibitors of ERK1/2.* J Med Chem, 2015. **58**(11): p. 4790-801.
247. Yang, G., et al., *Caffeic acid directly targets ERK1/2 to attenuate solar UV-induced skin carcinogenesis.* Cancer Prev Res (Phila), 2014. **7**(10): p. 1056-66.
248. Mace, P.D., et al., *Structure of ERK2 bound to PEA-15 reveals a mechanism for rapid release of activated MAPK.* Nat Commun, 2013. **4**: p. 1681.
249. Chaikuad, A., et al., *A unique inhibitor binding site in ERK1/2 is associated with slow binding kinetics.* Nat Chem Biol, 2014. **10**(10): p. 853-60.
250. Shevchenko, A., et al., *In-gel digestion for mass spectrometric characterization of proteins and proteomes.* Nat Protoc, 2006. **1**(6): p. 2856-60.
251. Louis-Jeune, C., M.A. Andrade-Navarro, and C. Perez-Iratxeta, *Prediction of protein secondary structure from circular dichroism using theoretically derived spectra.* Proteins, 2012. **80**(2): p. 374-81.
252. Zhang, Y., *I-TASSER server for protein 3D structure prediction.* BMC Bioinformatics, 2008. **9**: p. 40.
253. Laskowski, R.A., et al., *AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR.* J Biomol NMR, 1996. **8**(4): p. 477-86.
254. Luthy, R., J.U. Bowie, and D. Eisenberg, *Assessment of protein models with three-dimensional profiles.* Nature, 1992. **356**(6364): p. 83-5.

255. Muller, A., R.M. MacCallum, and M.J. Sternberg, *Benchmarking PSI-BLAST in genome annotation.* J Mol Biol, 1999. **293**(5): p. 1257-71.
256. Bennett-Lovsey, R.M., et al., *Exploring the extremes of sequence/structure space with ensemble fold recognition in the program Phyre.* Proteins, 2008. **70**(3): p. 611-25.
257. Brooks, B.R., et al., *CHARMM: the biomolecular simulation program.* J Comput Chem, 2009. **30**(10): p. 1545-614.
258. Kyte, J. and R.F. Doolittle, *A simple method for displaying the hydropathic character of a protein.* J Mol Biol, 1982. **157**(1): p. 105-32.
259. Jacobs, D., et al., *Multiple docking sites on substrate proteins form a modular system that mediates recognition by ERK MAP kinase.* Genes Dev, 1999. **13**(2): p. 163-75.
260. Waas, W.F., et al., *A kinetic approach towards understanding substrate interactions and the catalytic mechanism of the serine/threonine protein kinase ERK2: identifying a potential regulatory role for divalent magnesium.* Biochim Biophys Acta, 2004. **1697**(1-2): p. 81-7.
261. Larkin, M.A., et al., *Clustal W and Clustal X version 2.0.* Bioinformatics, 2007. **23**(21): p. 2947-8.
262. Iverson, C., et al., *RDEA119/BAY 869766: a potent, selective, allosteric inhibitor of MEK1/2 for the treatment of cancer.* Cancer Res, 2009. **69**(17): p. 6839-47.
263. Cole, C., J.D. Barber, and G.J. Barton, *The Jpred 3 secondary structure prediction server.* Nucleic Acids Res, 2008. **36**(Web Server issue): p. W197-201.
264. Petsalaki, E., et al., *Accurate prediction of peptide binding sites on protein surfaces.* PLoS Comput Biol, 2009. **5**(3): p. e1000335.
265. Kihara, D., et al., *Ab initio protein structure prediction on a genomic scale: application to the Mycoplasma genitalium genome.* Proc Natl Acad Sci U S A, 2002. **99**(9): p. 5993-8.
266. Moller, D.E., et al., *Human rsk isoforms: cloning and characterization of tissue-specific expression.* Am J Physiol, 1994. **266**(2 Pt 1): p. C351-9.
267. Larrea, M.D., et al., *RSK1 drives p27Kip1 phosphorylation at T198 to promote RhoA inhibition and increase cell motility.* Proc Natl Acad Sci U S A, 2009. **106**(23): p. 9268-73.
268. Auld, G.C., et al., *Identification of calcium-regulated heat-stable protein of 24 kDa (CRHSP24) as a physiological substrate for PKB and RSK using KESTREL.* Biochem J, 2005. **389**(Pt 3): p. 775-83.
269. Joel, P.B., et al., *pp90rsk1 regulates estrogen receptor-mediated transcription through phosphorylation of Ser-167.* Mol Cell Biol, 1998. **18**(4): p. 1978-84.
270. Virdee, K., et al., *Phosphorylation of human microtubule-associated protein tau by protein kinases of the AGC subfamily.* FEBS Lett, 2007. **581**(14): p. 2657-62.
271. Carriere, A., et al., *Oncogenic MAPK signaling stimulates mTORC1 activity by promoting RSK-mediated raptor phosphorylation.* Curr Biol, 2008. **18**(17): p. 1269-77.
272. Schmitt, J., H. Hess, and H.G. Stunnenberg, *Affinity purification of histidine-tagged proteins.* Mol Biol Rep, 1993. **18**(3): p. 223-30.
273. Gething, M.J. and J. Sambrook, *Protein folding in the cell.* Nature, 1992. **355**(6355): p. 33-45.
274. Carter, P., *Site-directed mutagenesis.* Biochem J, 1986. **237**(1): p. 1-7.
275. Pace, C.N. and K.L. Shaw, *Linear extrapolation method of analyzing solvent denaturation curves.* Proteins, 2000. **Suppl 4**: p. 1-7.
276. Zhao, Y., et al., *RSK3 encodes a novel pp90rsk isoform with a unique N-terminal sequence: growth factor-stimulated kinase function and nuclear translocation.* Mol Cell Biol, 1995. **15**(8): p. 4353-63.
277. Bignone, P.A., et al., *RPS6KA2, a putative tumour suppressor gene at 6q27 in sporadic epithelial ovarian cancer.* Oncogene, 2007. **26**(5): p. 683-700.
278. Oppermann, F.S., et al., *Large-scale proteomics analysis of the human kinome.* Mol Cell Proteomics, 2009. **8**(7): p. 1751-64.
279. Nakajima, T., et al., *The signal-dependent coactivator CBP is a nuclear target for pp90RSK.* Cell, 1996. **86**(3): p. 465-74.

280. Wang, J., et al., *Toward an understanding of the protein interaction network of the human liver.* Mol Syst Biol, 2011. **7**: p. 536.
281. Havugimana, P.C., et al., *A census of human soluble protein complexes.* Cell, 2012. **150**(5): p. 1068-81.
282. Huttlin, E.L., et al., *The BioPlex Network: A Systematic Exploration of the Human Interactome.* Cell, 2015. **162**(2): p. 425-40.
283. Pappin, D.J., P. Hojrup, and A.J. Bleasby, *Rapid identification of proteins by peptide-mass fingerprinting.* Curr Biol, 1993. **3**(6): p. 327-32.
284. Eswar, N., et al., *Comparative protein structure modeling using Modeller.* Curr Protoc Bioinformatics, 2006. **Chapter 5**: p. Unit 5 6.
285. Sippl, M.J., *Recognition of errors in three-dimensional structures of proteins.* Proteins, 1993. **17**(4): p. 355-62.
286. Wiederstein, M. and M.J. Sippl, *ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins.* Nucleic Acids Res, 2007. **35**(Web Server issue): p. W407-10.
287. de Vries, S.J., M. van Dijk, and A.M. Bonvin, *The HADDOCK web server for data-driven biomolecular docking.* Nat Protoc, 2010. **5**(5): p. 883-97.
288. Armstrong, N., A. de Lencastre, and E. Gouaux, *A new protein folding screen: application to the ligand binding domains of a glutamate and kainate receptor and to lysozyme and carbonic anhydrase.* Protein Sci, 1999. **8**(7): p. 1475-83.
289. Holbourn, K.P. and K.R. Acharya, *Cloning, expression and purification of the CCN family of proteins in Escherichia coli.* Biochem Biophys Res Commun, 2011. **407**(4): p. 837-41.



Journal of Biomolecular Structure and Dynamics
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/tbsd20

Conserved residues at the MAPKs binding interfaces that regulate transcriptional machinery

Bhanu P. Jagilinki[a], Nikhil Gadewal[a], Harshal Mehta[a], Hafiza Mahadik[a], Vikrant Pandey[a], Anamika[a], Ulka Sawant[a], Prasad A. Wadegaonkar[b], Peyush Goyal[c], Satish Kumar[d] & Ashok K. Varma[a]

[a] Tata Memorial Centre, Advanced Centre for Treatment, Research and Education in Cancer, Kharghar, Navi Mumbai 410 210, Maharashtra, India

[b] Department of Biotechnology, Sant Gadge Baba Amravati University, Amravati 444602, Maharashtra, India

[c] Department of Biotechnology, Block - 2, 6th-8th floor, CGO Complex, Lodhi Road, New Delhi 110 003, India

[d] Department of Biochemistry & Bioinformatics Centre, Mahatma Gandhi Institute of Medical Sciences, Sevagram (Wardha) 442102, India

Accepted author version posted online: 16 Apr 2014.Published online: 14 May 2014.

To cite this article: Bhanu P. Jagilinki, Nikhil Gadewal, Harshal Mehta, Hafiza Mahadik, Vikrant Pandey, Anamika, Ulka Sawant, Prasad A. Wadegaonkar, Peyush Goyal, Satish Kumar & Ashok K. Varma (2015) Conserved residues at the MAPKs binding interfaces that regulate transcriptional machinery, Journal of Biomolecular Structure and Dynamics, 33:4, 852-860, DOI: 10.1080/07391102.2014.915764

To link to this article: http://dx.doi.org/10.1080/07391102.2014.915764

Journal of Biomolecular Structure and Dynamics, 2015
Vol. 33, No. 4, 852–860, http://dx.doi.org/10.1080/07391102.2014.915764

Conserved residues at the MAPKs binding interfaces that regulate transcriptional machinery

Bhanu P. Jagilinki[a], Nikhil Gadewal[a], Harshal Mehta[a], Hafiza Mahadik[a], Vikrant[a], Anamika Pandey[a], Ulka Sawant[a], Prasad A. Wadegaonkar[b], Peyush Goyal[c], Satish Kumar[d] and Ashok K. Varma[a]*

[a]*Tata Memorial Centre, Advanced Centre for Treatment, Research and Education in Cancer, Kharghar, Navi Mumbai 410 210, Maharashtra, India;* [b]*Department of Biotechnology, Sant Gadge Baba Amravati University, Amravati 444602, Maharashtra, India;* [c]*Department of Biotechnology, Block – 2, 6th–8th floor, CGO Complex, Lodhi Road, New Delhi 110 003, India;* [d]*Department of Biochemistry & Bioinformatics Centre, Mahatma Gandhi Institute of Medical Sciences, Sevagram (Wardha) 442102, India*

Communicated by Ramaswamy H. Sarma

(*Received 25 July 2013; accepted 13 April 2014*)

Signaling through c-Raf downstream pathways is the crucial subject of extensive studies because over expressed or mutated genes in this pathway lead to a variety of human cancers. On the basis of cellular localization, this pathway has been sub-divided into two cascades. The first RAF1-MEK1-ERK2 cascade which remains in the cytosol, whereas the second MEK1-ERK2-RSKs transduces into the nucleus and regulates the transactivation function. But how a few amino acids critically regulate the transcriptional function remains unclear. In this paper, we have performed *in silico* studies to unravel how atomic complexities at the MEK1-ERK2-RSKs pathways intercedes different functional responses. The secondary structure of the ERK, RSKs have been modeled using Jpred3, PSI-PHRED, protein modeler, and Integrated sequence analyzer from Discovery Studio software. Peptides of RSKs isozymes (RSK1/2/3/4) were built and docked on ERK2 structure using ZDOCK module. The hydropathy index for the RSKs molecules was determined using the KYTE–DOOLITTLE plot. The simulations of complex molecules were carried out using a CHARMM force field. The protein–protein interactions (PPIs) in different cascade of MAP kinase (MAPK) have been shown to be similar to those predicted *in vivo*. PPIs elucidate that the amino acids located at the conserved domains of MAPK pathways are responsible for transactivation functions.

Keywords: *In silico*; MAPKs; Protein–Protein Interactions; transactivation function; RSKs

Introduction

There are several members of the MAP kinase (MAPK) pathway, which feature Ser/Thr family of protein kinases (Hanks, Quinn, & Hunter, 1988; Roux & Blenis, 2004). Every kinase pathway is triggered by a specific activator (Ben-Levy et al., 1995). This activated kinase has long been studied for various responses like gene expression, cell cycle arrest, and cell proliferation (Gong et al., 2008; Igata et al., 2005; Mabuchi et al., 2004; Satomi & Nishino, 2009; Zhang & Kong, 2008). The well studied c-Raf downstream pathway leads to an attractive signaling mechanism which regulates cellular function such as cell growth and motility (Hilger, Scheulen, & Strumberg, 2002; Roux et al., 2007). The general layout of this pathway has been divided into two cascades, the first is RAF1-MEK1-ERK2 which remains in the cytosol, and the second MEK1-ERK2-RSKs is involved in the translocation of this complex into the nucleus. It is difficult to define the distinct function of both the interacting cascades, because the sequential activation of these kinases activates other kinases localized in different parts of cell such as the plasma membrane, cytosol, and nucleus. This pathway also helps to characterize the directionality of signals in the cellular environment. Despite a large volume of *in vitro* and *in vivo* data available for understanding the signal transduction mechanism of Ras pathways, there are still many unanswered questions. For example, (1) how does RAF1 translocate from the plasma membrane and activate the downstream kinases? (2) how do RAF1-MEK1-ERK2 complexes retain themselves in the cytosol? (3) What are the critical amino acids at the MEK1-ERK2-RSKs cascade that translocate this complex to the nucleus? (4) Is there any cross-talk between with the kinases other than in this pathway? There are contradictory reports about the cytosolic retention of RAF1-MEK1-ERK2 and nucleur translocation of MEK1-ERK2-RSKs complexes. However, our study explores the functional mechanism underlying MEK1-ERK2-RSKs using protein–protein interactions (PPIs).

The MAPK family of kinases is well known for mediating cellular responses to growth signals. The signal is transduced through phosphorylation, which occurs from the upstream kinase to the downstream kinase (Shi et al., 2007). Ras is a member of the G protein family,

*Corresponding author. Email: avarma@actrec.gov.in

while Raf belongs to the MAP3 K family, MEK belongs to the MAP2 K family, and ERK belongs to the MAPK family (Channavajhala et al., 2003; Ellinger-Ziegelbauer, Kelly, & Siebenlist, 1999; He, Cai, Yang, Liu, & Wan, 2009). RSKs are constituents of the AGC subfamily in the human kinome (Nguyen, 2008). The functional motif in MAPK pathways can be divided into two types of domains – the first type is the phosphorylation induced activation domains and the second type are the binding complex induced functional domains. The phosphorylated domain is buried inside the three-dimensional structural packing whereas the binding domain is located on the surface of the protein (Figure 1(A)).

Activated c-Ras interacts with the c-Raf at the plasma membrane which undergoes through conformational changes. Amino acids critical for the activation of Raf have been identified earlier (Terai & Matsuda, 2005). The conformational changes on c-Raf induce its binding to MEK1 kinases. The c-Raf-MEK1 complex activates MEK1 and phosphorylates it. This activated MEK1 binds to its downstream kinase ERK2 and phosphorylates it. The phosphorylated ERK2 binds to RSKs which activates the transcriptional and translational machinery by means of phosphorylation.

The extracellular signal regulated kinase ERK-2 is a 42 kDa protein which transduces the signal through the activation of MEK1/2 (Tarrega et al., 2005). MEK1/2 is an upstream signal transducer for ERK2 while ERK2 is an upstream signal transducer for RSKs. The RSKs (p90 ribosomal S6 family of kinases) are located downstream of the Ras/mitogen-activated protein kinase (MAPK) pathway (Poteet-Smith, Smith, Lannigan, Freed, & Sturgill, 1999). RSKs have two sequentially diverse and functionally distinct C-terminal (CTD) and N-terminal (NTD) protein kinase domains, with a linker at the interface, and tails of a few amino acids at the N and C-terminals. Sequentially, the N-terminal kinase domain bears a resemblance to the AGC family of kinases, whereas the C-terminal kinase domain resembles the calcium/calmodulin kinase family (Hanks et al., 1988; Jones, Erikson, Blenis, Maller, & Erikson, 1988). Functionally, the N-terminal kinase domain mediates substrate phosphorylation whereas C-terminal kinase domain is involved in autophosphorylation (Frödin & Gammeltoft, 1999; Poteet-Smith et al., 1999) The activation of RSKs is a complex process which involves multiple phosphorylations at various conserved amino acids (Dalby, Morrice, Caudwell, Avruch, & Cohen, 1998). These phosphorylation sites are highly conserved in RSK isozymes. The extended C-terminal kinase domains contain highly conserved phosphorylated serine, which is known to be one of the binding domains of ERK2. The fully activated P90RSK has four major phosphorylation sites (Ser 221, Ser 363, Ser 380, and Thr 573) and two additional sites (Thr 359 and Ser 732) with specific functions. The atomic association between different MAPKs pocket will help in finding the critical amino acids responsible for signal transduction and transactivation functions. An efficient signal transducer needs to avoid the cross-talk between the genes scaffold. Earlier data predicted that kinases in these pathways form robust complexes, but the molecular association of the proteins

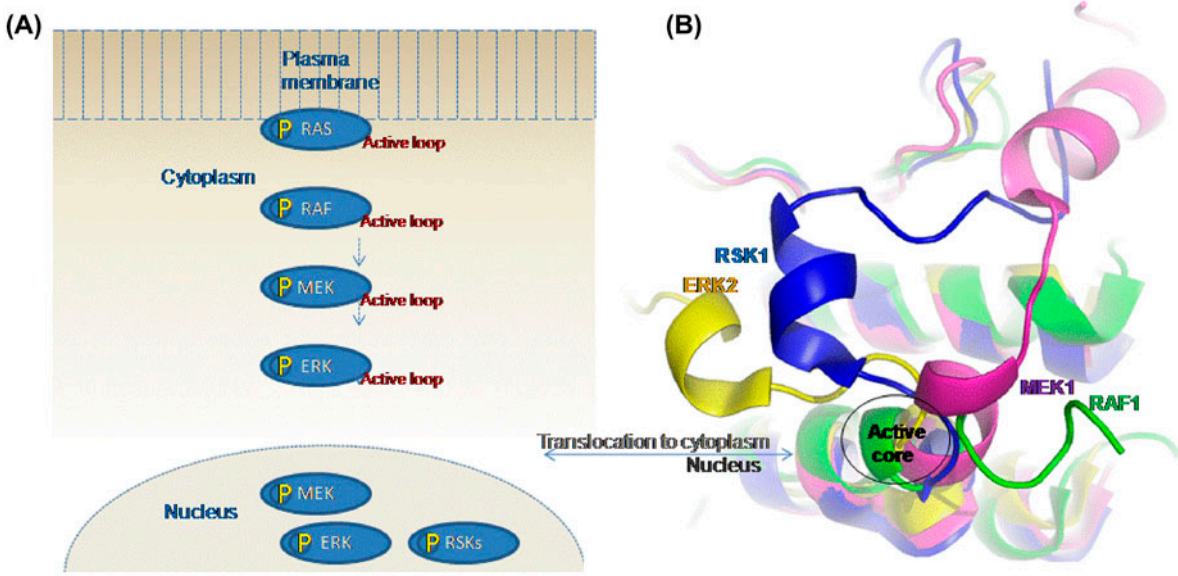


Figure 1. (A) The schematic representation of c-Raf downstream kinase pathway, the phosphorylation site (P) is buried inside the kinase domains and active binding domain is located at the molecular surface. (B) The superimposed structures of RAF1, MEK1, ERK2, and RSK which shows good structural similarity, however the bio-active core is conformationally diverse.

at the residual level is still implicit (Tian & Song, 2012). This paper provides structural insights due to conformational changes occurred in the pathways, and furthermore their effects on functional mechanism.

Results and discussion

It is well known that PPIs play an important role in coordinated signaling process of MAPK pathways. It has been reported that ERK2 has essential hydrophobic and charged amino acids that help to activate RSKs. But these sequences have not demonstrated a role to establish the signaling between the MAPKs pathways, which eventually leads to activation of the downstream protein kinase. Considering this fact, we looked at the upstream protein kinases in MAPK and found a docking site which is away from the enzymatic active site. The molecular *in silico* based approach has been applied to discover sequences of specific amino acid at the different MAPKs binding cascade which has not been evaluated extensively (Figure 1(B)).

MEK1-ERK2 binding interface

It has been reported that the MEK1-ERK2 binding interface is essential for the activation of RSKs. MEK1/2 in this pathway interact with ERK1/2 and transmit the signals to its downstream target. The N-terminus of MEK1 has mostly hydrophobic and basic residues which are essential for binding to ERK2 (Fukuda, Gotoh, & Nishida, 1997). Several lines of investigations have suggested that proper folded ERK2 dimer accumulate in the nucleus, whereas MEK1 associated ERK2 retained in the cytoplasm (Fukuda et al., 1997; Rubinfeld, Hanoch, & Seger, 1999; Xu, Stippec, Robinson, & Cobb, 2001). Hence, ERK2 is shuttling between cytoplasm and nucleus. To predict the atomic orientations of amino acids at the MEK1-ERK2 binding interface responsible for translocation of proper folded complex into nucleus, crystal structure of phospho-ERK2 which is found to be in twofold symmetry has been visualized. The sequence alignment of MEK1 (PDB ID: 3E8 N) and ERK2 (PDB ID: 1TVO) using Clustalw (Larkin et al., 2007) scores very less 18.0 while the overall three-dimensional folding and ATP binding site is similar except distortion at N and C-terminus. Few important domains of ERK2 have been characterized as binding domains of MEK1. The molecular docking between MEK1 and ERK2 has been performed on structurally determined molecule of ERK2 as well as symmetry related molecules (Figure 2(A) and (B)). It has been observed that ERK2 residues Tyr 316, Tyr 317, Asp 318, Asp 321, and Glu 322 are at the binding interface of MEK1 (Figure 2(C)) (Table 1, Supplementary material). However, MEK1 docked on the dimer structure of ERK2 is forming steric clashes (Figure 2(D)) which indicates the dimer form of ERK2 may not form complex with MEK1. This further elucidates that the MEK1 is not in a position to form a trimeric complex with ERK2 dimer and translocate into nucleus but rather remains in the cytosol. However, monomeric form of ERK2 is associated with the RSK1 and translocate to the nucleus.

Interactions between ERK2 and RSKs isoforms

The sequences of all RSK isoforms flanking the docking site were evaluated on the basis of secondary structure, hydrophobicity/hydrophilicity characteristics, flexibility, and accessibility of the sequences for ERK2 interactions. The homology model building, considering the appropriate sequence similarity (>30%) with the template, is critically important in analyzing the bioactive core of the protein structure. However, our initial search yielded ~30% sequence homology for the entire C-terminal tail region and significantly lower homology for the RSK CTD Docking sequences. To construct the sequences to be docked onto ERK2, a build protein tool integrated in the Discovery Studio2.1 was used. Due to lower sequence homology, an alternate method, Jpred3, was also tried for protein structure construction (Cole, Barber, & Barton, 2008). Jpred3 predicts secondary structure at an accuracy of 81.5%. Jpred3 derives alignment profiles and predicts secondary structure and solvent accessibility based on a Jnet algorithm. Both the programs observed similar predictions for the RSK isoforms docking sequences. The primary sequence was then modeled into a tertiary structure using *ab initio* based protein modeling. The RSK isoforms of 14 amino acid docking sequences were built into a model having minimum energy conformation. Since energy-based modeling works better for a smaller number of amino acids than a long peptide chain (Petsalaki, Stark, García-Urdiales, & Russell, 2009). The conserved docking sequence was predicted to be helical in structure and the flanking sequences were seen to be consisting of helices and coils. We have selected 14 amino acid sequences through observation of docking physicochemical attributes which includes conserved docking sequence as well as the flanking sequences. This selection has been analyzed using the Discovery Studio module for analyzing docked structures (Chen, Lyne, Giordanetto, Lovell, & Li, 2006). RSKs binding site has been found on the surface groove of ERK2, and far away (>20 Å) from the inhibitor binding sites, speculating inhibitor independent binding sites for each RSKs (Figure 3(A)). However, ERK2 residues from 323 to 329 responsible for nuclear translocation are located at the binding interface of RSK (Figure 3(B)) (Table 2), Supplementary material. It can be concluded here that the enzymatic site of ERK2 is buried inside of the three-dimensional structure, however

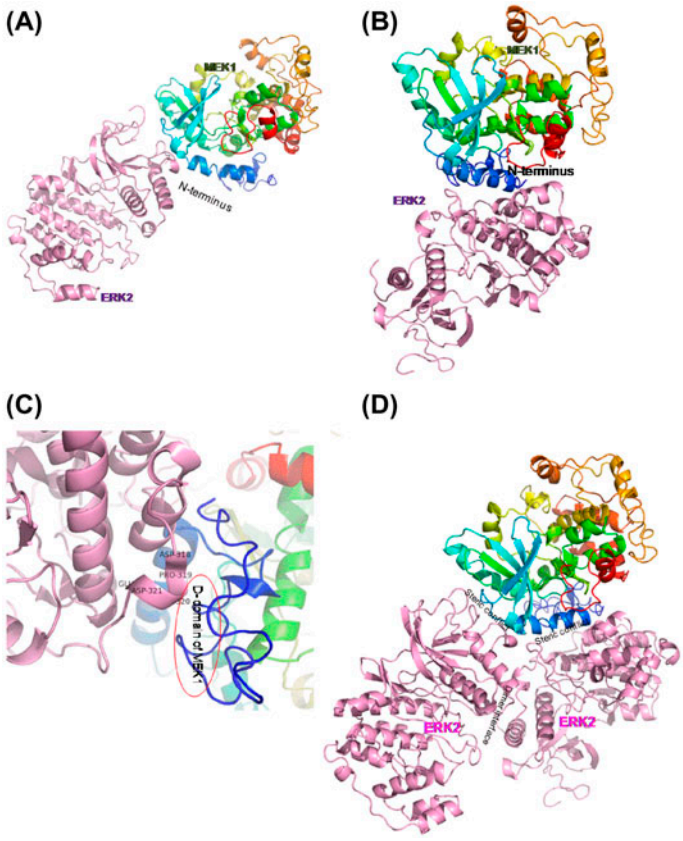


Figure 2. (A) Docking of MEK1 model on ERK2, N-terminus of MEK1 is forming complex with ERK2 molecule. (B) Docked structure of symmetry related molecules of ERK2 with MEK1. (C) The important residue at the docking interface of ERK2 is labeled and the docking domain of MEK1 is encircled. (D) There are steric clashes between the docked structure of p-ERK2 dimer and MEK1. ERK2 is colored in pink and MEK1 in rainbow.

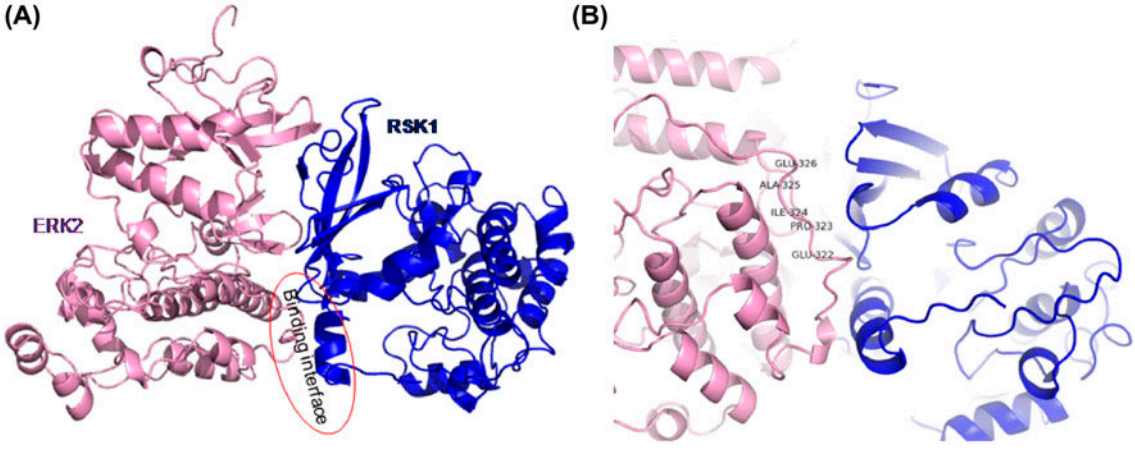


Figure 3. Complex model of ERK2-RSK1. (A) The ERK2 has a docking site at the C-terminus of RSK1. (B) The important residues at the binding interface has been shown.

binding domain being far from enzymatic sites helps in the translocation of this complex into the nucleus.

RSK isoforms at the C-terminal tail have conserved "LAQRR" amino acids. These conserved amino acids at

docking region, along with its flanking sequences, has been identified as interacting amino acids for the specific residue of an ERK2 molecule (Gavin & Nebreda, 1999; Smith, Poteet-Smith, Malarkey, & Sturgill, 1999). The sequences flanking around conserved "LAQRR" motif are variable in different RSKs isoforms which interacts with ERK2 with a specific docking orientation. These docking orientations determine the extent of association between the two molecules as well as variable functionalities. There are differences in the weak intermolecular interactions observed between ERK2-RSKs complexes (Figure 4(A)–(D)) (Tables 3–6, Supplementary material). It concludes that the differences of amino acids at the C-terminus and N-terminus of this conserved sequence are responsible for the change in the conformation and binding affinity. So within the limits of the docking software, to our conclusion the hydrophobic interactions and hydrogen bonding observed between ERK2-RSKs predict the folding behavior of complex molecules and also the directionality of signal transduction. It has been analyzed that the RSK1 has less binding than RSK2, RSK3, and RSK4. The *in silico* based results are in full agreement with what has already been observed *in vivo* (Smith et al., 1999).

RSK docking region is hydrophilic in nature

The RSKs isoforms sequences were tested for their hydrophobic/hydrophilic characteristics in order to predispose docking site for upstream kinase. For any residue to interact with a protein molecule, it has to be sufficiently exposed, i.e. accessible to the docking sequence of the approaching protein. The KYTE–DOOLITTLE (Kyte & Doolittle, 1982) plot in the ProtScale tool describes the sequences flanking at the end of the C-terminal of RSK isoforms to be particularly hydrophilic in nature. This is one of the rationale these sequences containing motifs are adopting the docking site at the surface of protein structures. The conserved residues within the carboxyl-terminal end of the RSK isozymes provides a specific ERK docking site (Biondi & Nebreda, 2003). The presence of an Arginine or a Lysine residue at the docking regions in all RSK isoforms raises the possibilities for such residues playing an important role in the binding of specific amino acid residues of ERK2 (Roux, Richards, & Blenis, 2003).

Transcriptional machinery in the nucleus at atomic level

Aspartic acid at positions 318 and 321, and residues from 323 to 329 in ERK2 makes prominent contacts with the MEK1 and RSKs, respectively. Therefore, it can be concluded that regions between 314 and 329 of ERK2 is a common docking domain for MEK1 and RSKs. The primary sequences of RSKs involved in binding are characterized by a cluster of chargedresidues surrounded by hydrophobic residues (Figure 5(A)). The PPIs between MEK1-ERK2-RSKs cascade are very robust which make this complex very intact (Figure 5(B)). However, there is conflict in the results on translocation of the MEK1-ERK2-RSKs complex into the nucleus. A large amount of data predicts that the properly folded MEK1-ERK2-RSKs transduce into the nucleus (Skarpen et al., 2008; Smith et al., 1999) and MEK1 return back to cytosol leaving ERK2-RSKs in the nucleus, but how long the MEK1-ERK2-RSKs complex remains there in the nucleus is vague. The results presented in this paper unravel the molecular association between ERK2-RSKs that may help in predicting the amino acids responsible for cellular translocation. This difference in the position of the inhibitor site and the RSKs docking site anticipates the dual function of ERK2, first activation of ERK2 due to an enzymatic site present in the core of the structure and the second translocation of the ERK2-RSKs complex into the nucleus present at the protein surface. The RSKs are large isozymes which form complexes with different components of transactivation domain. Within the limit of the *in silico* docking, it has been that the amino acids Asp 318, Asp 321 are critical for docking between the ERK2-RSKs interface and may play a role in regulating the transcription function.

Conclusion

Comparing the PPIs between different cascades of MEK1-ERK2-RSK1/2/3/4 complexes, it has been observed that MEK1 and ERK2 are key molecules regulating the functional activity in the MAPK pathway. However, residues from 314 to 329 of ERK2 interact with MEK1 and RSKs in two pronged manner. Binding association between MEK1-ERK2-RSKs complex has revealed that there are steric hindrances between RSKs isozymes in the complex which in turn disassociates MEK1 from ERK2 and RSK complex. The differences observed in docking orientations for each isozyme of RSKs indicate that the sequence present at the N and C-terminus of conserved "LAQRR" sequence motif of RSKs plays a major role in predicting the binding affinity between ERK2-RSKs molecule. The PPIs between ERK2-RSKs molecules unravel the role of the ERK2 molecule for different RSKs isozymes, and eventually conclude the catalog of important differences among the closely related RSK family of kinases. In our studies we have showed the macromolecular docking and interactions of amino acids at atomic level. We also showed the importance of steric hindrance which plays a major role in dissociation of upstream kinases from the signaling complex in the MAPK pathway. The differences in binding affinities for various isozymes are essential for triggering a specific

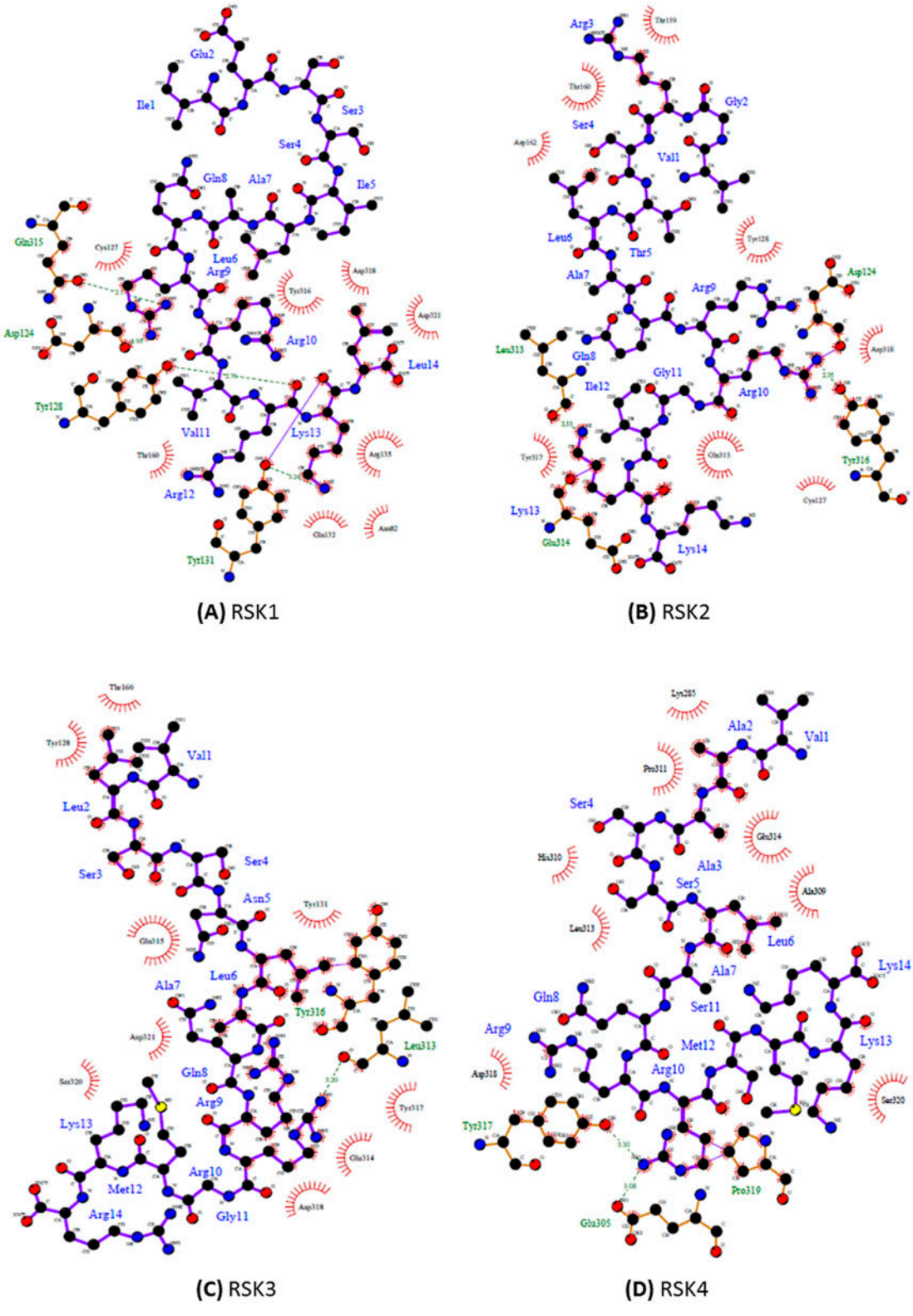


Figure 4. Binding association between ERK2 and C-terminus peptide of (A) RSK1, (B) RSK2, (C) RSK3, (D) RSK4.

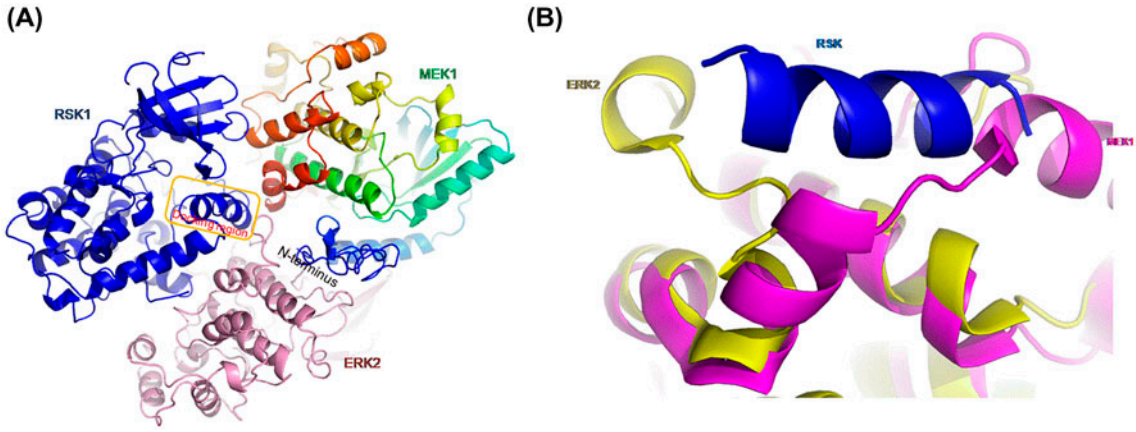


Figure 5. (A) Trimeric complex model of MEK1-ERK2-RSK1, N-terminus of MEK1 and C-terminus of RSK1 is interacting with interface of ERK2. The RSK1 is colored in blue. (B) N-terminus peptide motif of RSK1 docked to the model interface of ERK2-MEK1, it has been observed that peptide motif is also finding the same conformation what has been observed in *in vitro*.

type of transcription factors depending on the requirement of the cell. Atomic association between different MEK1-ERK2-RSKs complexes will be helpful in designing the small molecule inhibitors, which may regulate the structural basis of transactivation function.

Methodology

Model building of MEK1, ERK2, and RSK1

The MEK1 N-terminus contains the ERK docking site. Since the PDB structure of MEK1 (PDB ID: 3E8 N) does not contain 60 residues at N-terminal, I-tasser (Zhang, 2008)server was used to model the full length structure. At C-terminal end of RSK1 (PDB ID: 2WNT) the ERK docking site was missing, therefore C-terminal domain of RSK1 was also modeled using the I-tasser server. The models were then evaluated using Procheck (Laskowski, Rullmannn, MacArthur, Kaptein, & Thornton, 1996) and Verify_3D server (Lüthy, Bowie, & Eisenberg, 1992). The missing loops of ERK2 (PDB ID: 1TVO) were modeled and minimized using the Prime tool of Maestro 9.2 (Schrodinger).

Docking of MEK1 and ERK2

Docking of MEK1 and ERK2 was performed using ZDock from Discovery Studio 2.5. The hydrophobic (^{9}IQL^{11}) and basic (^{3}KKK^{5}) residues at N-terminal of MEK1 and acidic residues (Asp318 and Asp321) at C-terminal L16 segment of ERK2 are selected to filter the docking poses. In addition to the shape complimentarily, the desolvation and electrostatic energy were used for ranking of the docked poses. Based on the docking scores and evaluation of binding interface residue the best docked pose was selected.

Docking of RSK1 to docked complex of MEK1 and ERK2

The docked complex of MEK1 and ERK2 was considered as a receptor and docked with RSK1. The poses were filtered by specifying the C-terminal residues (^{722}LAQRRVRKL^{730}) in RSK1 and acidic residues (Asp318 and Asp321) at C-terminal L16 segment of ERK2. Based on the docking scores the best docked pose of RSK1 with MEK1 and ERK2 was selected.

Modeling of RSK1/2/3/4 peptides

RSKs sequences (RSK1: IESSILAQRRVRKL, RSK2: VGRSTLAQRRGIKK, RSK3: VLSSNLAQRRGMKR, RSK4: VAASSLAQRRSMKK) subjected to BLAST (Müller, MacCallum, & Sternberg, 1999) in order to find a suitable template of known structures from PDB. PHYRE web-based program (Bennett-Lovsey, Herbert, Sternberg, & Kelley, 2008) was used for protein fold recognition using 1D and 3D sequence profiles coupled with secondary structure and solvation potential information. A build protein tool from Discovery studio 2.5.5 was used to construct the structures of RSK1/2/3/4 in alpha helical form. The side chains were then subjected to refinement using CHARMM force field (Brooks et al., 2009). The hydrophilic nature of RSK was calculated by KYTE–DOOLITTLE scale in ProtScale. Window size of 7 was kept and is good for finding hydrophilic regions that are likely exposed on the surface.

Docking of RSK1/2/3/4 peptides on ERK2

The Modeled RSKs peptides were docked to ERK2 molecule using the ZDock program keeping angular step size as 15. Out of the 2000 poses generated, only top poses that showed optimal hydrogen bonding were

selected. The molecular interactions between RSK1/2/3/4 and ERK2 were plotted using Ligplot tool (Laskowski & Swindells, 2011).

Supplementary material

The supplementary material for this paper is available online at http://dx.doi.10.1080/07391102.2014.915764.

Acknowledgments

We thank DBT-BTIS facility at ACTREC for providing necessary software for this study.

References

Ben-Levy, R., Leighton, I. A., Doza, Y. N., Attwood, P., Morrice, N., Marshall, C. J., & Cohen, P. (1995). Identification of novel phosphorylation sites required for activation of MAPKAP kinase-2. *EMBO Journal, 14*, 5920–5930.

Bennett-Lovsey, R. M., Herbert, A. D., Sternberg, M. J., & Kelley, L. A. (2008). Exploring the extremes of sequence/structure space with ensemble fold recognition in the program Phyre. *Proteins, 70*, 611–625.

Biondi, R. M., & Nebreda, A. R. (2003). Signalling specificity of Ser/Thr protein kinases through docking-site-mediated interactions. *Biochemical Journal, 372*, 1–13.

Brooks, B. R., Brooks, C. L., 3rd, Mackerell, A. D., Jr., Nilsson, L., Petrella, R. J., Roux, B., … Karplus, M. (2009). CHARMM: The biomolecular simulation program. *Journal of Computational Chemistry, 30*, 1545–1614.

Channavajhala, P. L., Wu, L., Cuozzo, J. W., Hall, J. P., Liu, W., Lin, L. L., & Zhang, Y. (2003). Identification of a novel human kinase supporter of Ras (hKSR-2) that functions as a negative regulator of Cot (Tpl2) signaling. *Journal of Biological Chemistry, 278*, 47089–47097.

Chen, H., Lyne, P. D., Giordanetto, F., Lovell, T., & Li, J. (2006). On evaluating molecular-docking methods for pose prediction and enrichment factors. *Journal of Chemical Information and Modeling, 46*, 401–415.

Cole, C., Barber, J. D., & Barton, G. J. (2008). The Jpred 3 secondary structure prediction server. *Nucleic Acids Research, 36*, W197–W201.

Dalby, K. N., Morrice, N., Caudwell, F. B., Avruch, J., & Cohen, P. (1998). Identification of regulatory phosphorylation sites in mitogen-activated protein kinase (MAPK)-activated protein kinase-1a/p90rsk that are inducible by MAPK. *Journal of Biological Chemistry, 273*, 1496–1505.

Ellinger-Ziegelbauer, H., Kelly, K., & Siebenlist, U. (1999). Cell cycle arrest and reversion of Ras-induced transformation by a conditionally activated form of mitogen-activated protein kinase kinase kinase 3. *Molecular Cell Biology, 19*, 3857–3868.

Frödin, M., & Gammeltoft, S. (1999). Role and regulation of 90 kDa ribosomal S6 kinase (RSK) in signal transduction. *Molecular and Cellular Endocrinology, 151*, 65–77.

Fukuda, M., Gotoh, Y., & Nishida, E. (1997). Interaction of MAP kinase with MAP kinase kinase: Its possible role in the control of nucleocytoplasmic transport of MAP kinase. *The EMBO Journal, 16*, 1901–1908.

Gavin, A. C., & Nebreda, A. R. (1999). A MAP kinase docking site is required for phosphorylation and activation of p90rsk/MAPKAP kinase-1. *Current Biology, 9*, 281–286.

Gong, X. W., Wei, J., Li, Y. S., Cheng, W. W., Deng, P., & Jiang, Y. (2008). Effect of p38 mitogen-activated protein kinase gene knockout on cell proliferation of embryonic fibroblasts in mice. *Zhongguo Wei Zhong Bing Ji Jiu Yi Xue, 20*, 527–529.

Hanks, S. K., Quinn, A. M., & Hunter, T. (1988). The protein kinase family: Conserved features and deduced phylogeny of the catalytic domains. *Science, 241*, 42–52.

He, Y. Y., Cai, B., Yang, Y. X., Liu, X. L., & Wan, X. P. (2009). Estrogenic G protein-coupled receptor 30 signaling is involved in regulation of endometrial carcinoma by promoting proliferation, invasion potential, and interleukin-6 secretion via the MEK/ERK mitogen-activated protein kinase pathway. *Cancer Science, 100*, 1051–1061.

Hilger, R. A., Scheulen, M. E., & Strumberg, D. (2002). The Ras-Raf-MEK-ERK pathway in the treatment of cancer. *Onkologie, 25*, 511–518.

Igata, M., Motoshima, H., Tsuruzoe, K., Kojima, K., Matsumura, T., Kondo, T., … Araki, E. (2005). Adenosine monophosphate-activated protein kinase suppresses vascular smooth muscle cell proliferation through the inhibition of cell cycle progression. *Circulation Research, 97*, 837–844.

Jones, S. W., Erikson, E., Blenis, J., Maller, J. L., & Erikson, R. L. (1988). A Xenopus ribosomal protein S6 kinase has two apparent kinase domains that are each similar to distinct protein kinases. *Proceedings of the National Academy of Sciences USA, 85*, 3377–3381.

Kyte, J., & Doolittle, R. F. (1982). A simple method for displaying the hydropathic character of a protein. *Journal of Molecular Biology, 157*, 105–132.

Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., … Higgins, D. G. (2007). Clustal W and Clustal X version 2.0. *Bioinformatics, 23*, 2947–2948.

Laskowski, R. A., Rullmannn, J. A., MacArthur, M. W., Kaptein, R., & Thornton, J. M. (1996). AQUA and PROCHECK-NMR: Programs for checking the quality of protein structures solved by NMR. *Journal of Biomolecular NMR, 8*, 477–486.

Laskowski, R. A., & Swindells, M. B. (2011). LigPlot+: Multiple ligand-protein interaction diagrams for drug discovery. *Journal of Chemical Information and Modeling, 51*, 2778–2786.

Lüthy, R., Bowie, J. U., & Eisenberg, D. (1992). Assessment of protein models with three-dimensional profiles. *Nature, 356*, 83–85.

Mabuchi, S., Ohmichi, M., Kimura, A., Ikebuchi, Y., Hisamoto, K., Arimoto-Ishida, E., … Murata, Y. (2004). Tamoxifen inhibits cell proliferation via mitogen-activated protein kinase cascades in human ovarian cancer cell lines in a manner not dependent on the expression of estrogen receptor or the sensitivity to cisplatin. *Endocrinology, 145*, 1302–1313.

Müller, A., MacCallum, R. M., & Sternberg, M. J. (1999). Benchmarking PSI-BLAST in genome annotation. *Journal of Molecular Biology, 293*, 1257–1271.

Nguyen, T. L. (2008). Targeting RSK: An overview of small molecule inhibitors. *Anti-Cancer Agents in Medicinal Chemistry, 8*, 710–716.

Petsalaki, E., Stark, A., García-Urdiales, E., & Russell, R. B. (2009). Accurate prediction of peptide binding sites on protein surfaces. *PLoS Computational Biology, 5*, e1000335.

Poteet-Smith, C. E., Smith, J. A., Lannigan, D. A., Freed, T. A., & Sturgill, T. W. (1999). Generation of constitutively active p90 ribosomal S6 kinase *in vivo*. Implications for the mitogen-activated protein kinase-activated protein

kinase family. *Journal of Biological Chemistry, 274*, 22135–22138.

Roux, P. P., & Blenis, J. (2004). ERK and p38 MAPK-activated protein kinases: A family of protein kinases with diverse biological functions. *Microbiology and Molecular Biology Reviews, 68*, 320–344.

Roux, P. P., Richards, S. A., & Blenis, J. (2003). Phosphorylation of p90 ribosomal S6 kinase (RSK) regulates extracellular signal-regulated kinase docking and RSK activity. *Journal of Molecular Cell Biology, 23*, 4796–4804.

Roux, P. P., Shahbazian, D., Vu, H., Holz, M. K., Cohen, M. S., Taunton, J., … Blenis, J. (2007). RAS/ERK signaling promotes site-specific ribosomal protein S6 phosphorylation via RSK and stimulates cap-dependent translation. *Journal of Biological Chemistry, 282*, 14056–14064.

Rubinfeld, H., Hanoch, T., & Seger, R. (1999). Identification of a cytoplasmic-retention sequence in ERK2. *Journal of Biological Chemistry, 274*, 30349–30352.

Satomi, Y., & Nishino, H. (2009). Implication of mitogen-activated protein kinase in the induction of G1 cell cycle arrest and gadd45 expression by the carotenoid fucoxanthin in human cancer cells. *Biochimica et Biophysica Acta, 1790*, 260–266.

Shi, F., Chiu, Y. J., Cho, Y., Bullard, T. A., Sokabe, M., & Fujiwara, K. (2007). Down-regulation of ERK but not MEK phosphorylation in cultured endothelial cells by repeated changes in cyclic stretch. *Cardiovascular Research, 73*, 813–822.

Skarpen, E., Flinder, L. I., Rosseland, C. M., Orstavik, S., Wierod, L., Oksvold, M. P., … Huitfeldt, H. S. (2008). MEK1 and MEK2 regulate distinct functions by sorting ERK2 to different intracellular compartments. *FASEB Journal, 22*, 466–476.

Smith, J. A., Poteet-Smith, C. E., Malarkey, K., & Sturgill, T. W. (1999). Identification of an extracellular signal-regulated kinase (ERK) docking site in ribosomal S6 kinase, a sequence critical for activation by ERK *in vivo*. *Journal of Biological Chemistry, 274*, 2893–2898.

Tarrega, C., Rios, P., Cejudo-Marin, R., Blanco-Aparicio, C., van den Berk, L., Schepens, J., … Pulido, R. (2005). ERK2 shows a restrictive and locally selective mechanism of recognition by its tyrosine phosphatase inactivators not shared by its activator MEK1. *Journal of Biological Chemistry, 280*, 37885–37894.

Terai, K., & Matsuda, M. (2005). Ras binding opens c-Raf to expose the docking site for mitogen-activated protein kinase kinase. *EMBO Reports, 6*, 251–255.

Tian, T., & Song, J. (2012). Mathematical modelling of the MAP kinase pathway using proteomic datasets. *PLoS One, 7*, e42230.

Xu, B., Stippec, S., Robinson, F. L., & Cobb, M. H. (2001). Hydrophobic as well as charged residues in both MEK1 and ERK2 are important for their proper docking. *Journal of Biological Chemistry, 276*, 26509–26515.

Zhang, Y. (2008). I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics, 9*, 40.

Zhang, Y. X., & Kong, C. Z. (2008). The role of mitogen-activated protein kinase cascades in inhibition of proliferation in human prostate carcinoma cells by raloxifene: An *in vitro* experiment. *Zhonghua Yi Xue Za Zhi, 88*, 271–275.

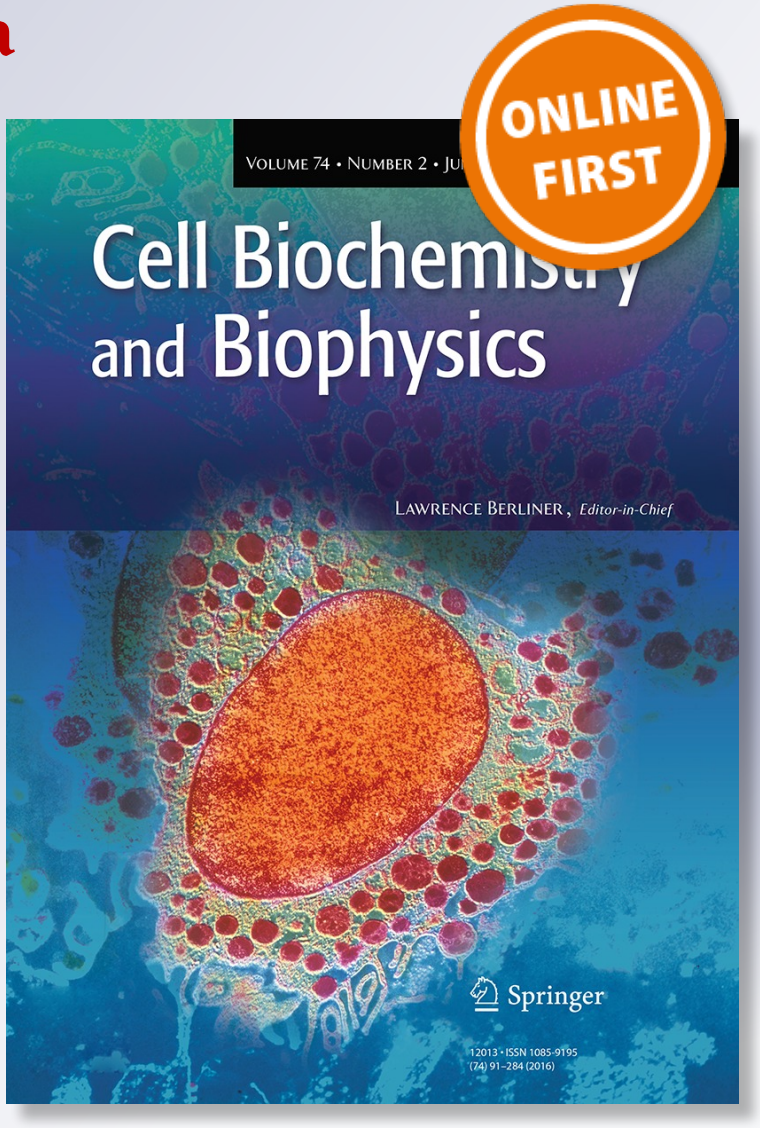
Functional Basis and Biophysical Approaches to Characterize the C-Terminal Domain of Human—Ribosomal S6 Kinases-3

Bhanu P. Jagilinki, Rajan Kumar Choudhary, Pankaj S. Thapa, Nikhil Gadewal, M. V. Hosur, Satish Kumar & Ashok K. Varma

Cell Biochemistry and Biophysics

ISSN 1085-9195

Cell Biochem Biophys
DOI 10.1007/s12013-016-0745-6



Springer

Cell Biochem Biophys
DOI 10.1007/s12013-016-0745-6

ORIGINAL PAPER

Functional Basis and Biophysical Approaches to Characterize the C-Terminal Domain of Human—Ribosomal S6 Kinases-3

Bhanu P. Jagilinki[1] · Rajan Kumar Choudhary[1] · Pankaj S. Thapa[1] · Nikhil Gadewal[1] · M. V. Hosur[1] · Satish Kumar[2] · Ashok K. Varma[1]

Received: 19 August 2015 / Accepted: 9 June 2016

Abstract Ribosomal S6 kinases (RSKs) are the major functional components in mitogen-activated protein kinase (MAPK) pathway, and these are activated by upstream Extracellular signal-regulated kinase. Upon activation, RSKs activate a number of substrate molecules involved in transcription, translation and cell-cycle regulation. But how cellular binding partners are engaged in the MAPK pathways and regulate the molecular mechanisms have not been explored. Considering the importance of protein–protein interactions in cell signalling and folding pattern of native protein, functional C-terminal kinase domain of RSK3 has been characterized using in vitro, in silico and biophysical approaches. RSKs discharge different functions by binding to downstream kinase partners. Hence, depending upon cellular binding partners, RSKs translocate between cytoplasm and nucleus. In our study, it has been observed that the refolded C-terminal Kinase domain (CTKD) of RSK 3 has a compact domain structure which is predominantly α-helical in nature by burying the tryptophans deep into the core, which was confirmed by CD, Fluorescence spectroscopy and limited proteolysis assay. Our study also revealed that RSK 3 CTKD was found to be a homotrimer from DLS experiments. A model was also built for RSK 3 CTKD and was further validated using PROCHECK and ProSA webservers.

Keywords CD spectroscopy · Dynamic light scattering · Fluorescence spectroscopy · Inclusion bodies · In silico modelling · MAPK/ERK and RSK

Abbreviation

RSK	Ribosomal s6 kinase
MAPK	Mitogen-Activated Protein Kinase
ERK	Extracellular signal-regulated kinase
PPI	Protein–Protein Interactions
CTKD	C-terminal kinase Domain
IPTG	Isopropyl-β-D-thiogalactoside
DLS	Dynamic Light Scattering
IMAC	Immobilized Metal ion Affinity Chromatography
RMSD	Root Mean Square Deviation

Introduction

The Ras–Raf mitogen-activated protein kinase (MAPK) pathway is a key signalling pathway, which is involved in the regulation of cell survival, growth, proliferation and differentiation. Kinases in the MAPK pathway get activated by phosphorylation at Ser/Thr residues. Ribosomal s6 kinase (RSK) activation requires multiple phosphorylations at different conserved Ser/Thr amino acids, and this occurs sequentially one after another by upstream MAPK kinases [1–3]. One of the important kinases in the MAPK Pathway is the ERK1/2 enzyme which phosphorylates at a specific C-terminal kinase domain (CTKD) and triggers phosphorylations, which in turn leads to final activation of RSKs. There are four reported isoforms of RSKs in

Electronic supplementary material The online version of this article (doi:10.1007/s12013-016-0745-6) contains supplementary material, which is available to authorized users.

✉ Ashok K. Varma
avarma@actrec.gov.in

[1] Advanced Centre for Treatment, Research and Education in Cancer, Tata Memorial Centre, Navi Mumbai, Kharghar 410 210, Maharashtra, India

[2] Department of Biochemistry and Bioinformatics Centre, Mahatma Gandhi Institute of Medical Sciences, Sevagram, Wardha 442102, India

Published online: 21 June 2016

humans, RSK 1/2/3/4 [4]. All the isozymes of RSKs share a high degree of sequence homology (75–80 % amino acid identity) [5, 6] and belong to the family of serine/threonine kinases. The activated RSKs regulate a number of transcription factors including CREB, SRF, ERα, FOS, JuN, STAT, MyoD, NFκB and TIF1A and also of coactivators of transcription such as CBP and p300 [7–12]. Different mutations discovered in these kinases are responsible for several reported cancers [13]. The signalling network in the downstream of Ras kinase involves extensively specific protein–protein interactions (PPIs). However, despite several reports, the atomic details of these PPIs and the associated molecular mechanisms are still unexplored. Our earlier in silico analysis had identified some of the PPIs involved at the binding interface of RSKs which regulate the activity of transcription factors [14].

The molecular structure of RSKs comprises of two functionally distinct N-terminal Kinase (NTKD) and the C-terminal Kinase (CTKD) domain. These two distinct kinase domains are connected by a linker region of ~100 amino acids, which has the highly conserved phosphoserine [5]. The NTKD shares homology with kinases of the AGC (protein kinase A/protein kinase G/protein kinase C) family, whereas the CTKD is homologous to the calcium/calmodulin-dependent protein kinases (CaMKs) [6, 15]. Further, there are N- and C-terminal tails of ~60 amino acids each flanking at both the ends. The CTKD is involved in autophosphorylation of RSK at the linker region, whereas NTKD is responsible for various substrate phosphorylations [15, 16]. Except ERK, RSK interacts with various substrates through its NTKD. The ERK-binding region, known as the D domain (Leu-Ala-Gln-Arg-Arg), is present within the C-terminal tail of RSKs [17]. The domain organization of RSK 3 is explained graphically in Fig. 1.

The crystal structures of NTKD and CTKD of RSK1/2 have been determined [18–22]. However, our search could not find any structures in the PDB for human-RSK 3 isoforms. Further, no crystal structure of even full length of RSKs isozymes has been determined. Considering the importance of RSKs isozymes, we have carried out the functional and biophysical characterization of the CTKD of RSK 3 isoform.

Materials and Methods

Cloning and Ligation

C-terminal Kinase domain of RSK 3 (ACCESSION NM_021135) (cDNA kind gift from John Blenis, Harvard Medical School, USA) has been cloned between EcoR1 and Sal1 restriction sites using pET-28a(+) vector. The PCR primers required for cloning were designed manually using the sequence information of RSK 3 (415–672) residues. The primer sequences are as follows: Forward primer: 5′-GTCGAATTCGAGAACCTGTAC TTTCAGGGT TACGAGATCAAGGAGGACATC-3′, Reverse primer: 5′-GTCGTCGAC CTATTACACCCACGGGTGTTTGAG CAC-3′. The tetrad PCR system from Bio-rad and the Phusion High-Fidelity DNA Polymerase enzyme (Finnzymes) were used to perform the PCR. Initial denaturation step was at 95 °C for a duration of 5 min, PCR cycle comprised of denaturation at 95 °C for 30 s, primer annealing at 65 °C for 30 s and the polymerase extension at 72 °C for 30 s, and this cycle was repeated 32 times. The amplified PCR product was loaded onto 1 % agarose gel and visualized by staining with ethidium bromide. The PCR product corresponding to ~800 bp was excised by visualizing under a UV filter on Gel doc equipment (UVP bioimaging systems). The PCR product recovered by gel extraction kit (Qiagen) was further treated with EcoR1 and Sal1 restriction enzymes (Fermentas) to generate sticky ends. This digested RSK 3 CTKD was ligated into pET-28a(+) vector using Quick Ligation kit (New England Biolabs). The ligated product was transformed into DH5α competent cells, and plated on Luria agar plate containing 50 μg/ml kanamycin. A single isolated colony was picked to inoculate with 10 ml of Luria Broth containing 50 μg/ml kanamycin, and the culture was allowed to grow overnight

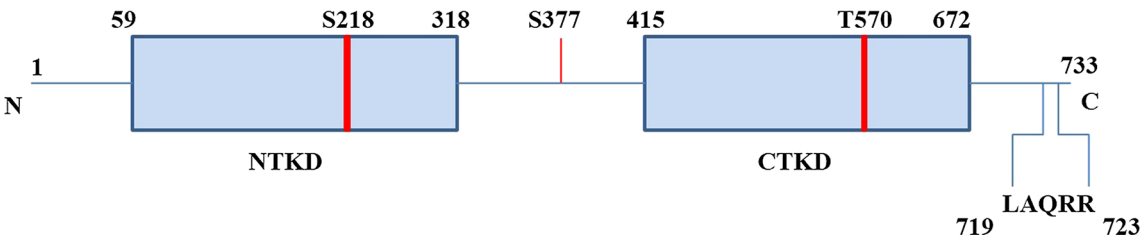


Fig. 1 RSK 3 (1–733) had two functional kinase domains, the NTKD (59–318) and the CTKD (415–672), connected by highly conserved linker region from 319 to 414. These two domains have N- and C-terminal flanking tails on either ends. The C-terminal tail has the ERK-binding motif 'LAQRR' (719–723) which is required for ERK docking. The important phosphorylation sites Ser218, Ser377 and Thr570 had been labelled

at 200 rpm and 37 °C. The bacterial culture was harvested by centrifugation at 4 °C and 6000 rpm. The recombinant plasmid was isolated using miniprep kit (Qiagen). The presence of RSK 3 CTKD insert was checked by agarose gel electrophoresis on the sample subjected to double digestion with EcoR1 and Sal1 restriction enzymes at 37 °C. The identity of the inserted PCR product in the recombinant plasmid was further confirmed by DNA sequencing (Applied Biosystems).

Expression and IMAC Purification

The RSK 3 CTKD/pET-28a(+) construct was transformed into Rosetta 2 (DE3) competent cells and plated on Luria agar containing 50 μg/ml Kanamycin and 34 μg/ml chloramphenicol. A 10 ml of Luria Broth medium containing 10 μl each of Kanamycin (50 mg/ml) and chloramphenicol (34 mg/ml) was inoculated with a single isolated colony and further incubated in a shaker at 37 °C and 200 rpm. After 8 h of incubation, the starting culture was diluted in 1 l of Luria Broth containing Kanamycin and chloramphenicol and was allowed to grow in a shaker incubator at 37 °C, 200 rpm till the optical density (OD_{600}) reached a value 0.6. The culture was then induced with 0.4 mM isopropyl-β-D-thiogalactoside (IPTG), and was further incubated on a shaker incubator for 16 h at 20 °C and 200 rpm. The bacterial culture was harvested by centrifugation for 15 min at 4 °C and 6000 rpm.

The bacterial pellet was resuspended in 10 ml of lysis buffer (50 mM Tris–HCl, pH 8.0, 500 mM NaCl and 4 mM $MgCl_2$). Protease inhibitor cocktail was added in a ratio of 1:1000 by volume (Sigma-Aldrich). The cells were disrupted by sonication (Branson digital sonifier) for 5 times, each 1-min cycle separated by 1-min incubation on ice, keeping the pulse on for 4 s and off for 2 s in every cycle. The sample was maintained on ice throughout the process. The cell lysate was then centrifuged for 20 min at 12,000 rpm and 4 °C. After separating the supernatant, the cell pellet was thoroughly washed thrice with the lysis buffer containing 1 % Triton X-100, 1 % Sarkosyl and 1 % Sodium Deoxycholate. The cell pellet was given one extra wash with lysis buffer containing 4 M Urea, before resuspending it into lysis buffer containing 8 M Urea. The denatured protein was separated by centrifuging at 12,000 rpm for 20 min at 4 °C. The supernatant was used for purification of 6His-RSK 3 CTKD by Immobilized Metal Ion Affinity Chromatography (IMAC).

For IMAC purification, Ni–NTA beads (Qiagen) were pre-equilibrated with five column washes of 50 mM Tris–HCl, pH 8.0; 500 mM NaCl; 8 M Urea. Then, the cell lysate containing denatured RSK 3 CTKD was allowed to bind on to the pre-equilibrated Ni–NTA beads using gravity flow methods. The 20-column washes were performed on the protein bound resin with lysis buffer containing 10 and 20 mM imidazole, respectively. RSK 3 CTKD was eluted in lysis buffers containing 100 and 250 mM imidazole. Elution fractions from each run were loaded on to 12 % SDS-PAGE and were visualized using coomassie staining.

Trypsin In-Gel Digestion and Mass Spectrometry

The protein band corresponding to 33 kDa of purified RSK 3 CTKD was excised from 12 % SDS-PAGE gel and was chopped into tiny pieces. These pieces were transferred into 1.7 ml eppendorf (Axygen) tube containing 1 ml of de-staining solution which is water, methanol and glacial acetic acid in the ratio of 4.5:4.5:1 and vortexed overnight. The following day, the gel pieces were washed with $(NH_4)HCO_3$ and Acetonitrile, and further treated with trypsin as described by Shevchenko et al. [23]. The peptides were then extracted and lyophilized. The lyophilized sample was dissolved in 0.1 % TFA in 50 % of acetonitrile for mass spectrometry, in which 1 μl of digested protein sample was mixed with 2 μl of α-cyano-4-hydroxycinnamic acid, and plated on MTP 384 target plate ground steel BC (Bruker) before allowed to dry using dried droplet method. The mass spectra were recorded using Ultraflex II MALDI-TOF/TOF mass spectrometer (Bruker Daltonics). For Peptide mass fingerprinting, the spectra were compared with those in the database using Mascot (Matrix Sciences) search engine integrated in the flexAnalysis software (Bruker Daltonics) [24].

Dialysis and Refolding

The fractions containing RSK 3 CTKD were transferred into SnakeSkin Pleated Dialysis Tubing (Thermo Scientific). The pooled protein sample was dialysed for 72 h at 4 °C with intermittent changes of refolding buffer 50 mM Tris, pH 8.0 containing 500 mM NaCl, 50 mM KCl, 10 mM $MgCl_2$, 2 mM DTT, 15 % Glycerol, 1 % Triton X-100 and 500 mM L-Arginine. The first two buffer changes were made after each 6 h of dialysis, whereas remaining changes were after 12 h. This refolded RSK 3 CTKD protein was concentrated up to 2 ml and further purified by size exclusion chromatography (SEC) using FPLC (AKTA-GE) in the Buffer A (10 mM Tris, 100 mM NaCl, pH 8.0). The FPLC fractions were analysed on 12 % SDS-PAGE gel.

Fluorescence Spectroscopy

A fluorescence spectrum (Horiba scientific) was recorded with excitation wavelength at $\lambda = 280$ nm, and the emission spectra were collected from $\lambda = 310$ to 400 nm. The

denatured RSK 3 CTKD was preserved in buffer containing 50 mM Tris–HCl, 500 mM NaCl, 8 M Urea and 250 mM Imidazole, pH 8.0, whereas refolded RSK 3 CTKD was in Buffer A. The protein concentration was 1 μM, and blank correction was performed individually with respective buffers.

Far UV CD Spectroscopy

For CD spectroscopy, 10 μM of refolded RSK 3 CTKD protein in Buffer A was used to record CD spectra (Jasco J-815 CD spectrophotometer) between $\lambda = 260$ and 200 nm. A total of five scans were recorded using slit width of 3 mm and the path length of 1 mm. The recoded CD spectra were submitted to K2D3 online server to estimate the content of α-helices and β-sheets [25]. To study the thermal stability, CD spectroscopy on 300 μl of 10 μM protein was performed from 10 to 80 °C with an increment of 2 °C.

Limited Proteolysis with Trypsin

10 μl of 10 pM Trypsin was added to 250 μl of FPLC purified RSK 3 CTKD (3 mg/ml) protein. 30 μl aliquot was immediately removed after adding trypsin to record limited proteolysis at 0 min. Subsequently, aliquots were removed at 5, 10, 15, 30, 60, 120 and 180 min, respectively. These samples were analysed on 15 % SDS-PAGE gel, which also has 30 μl of protein without trypsin as a negative control.

Dynamic Light Scattering (DLS)

To check the oligomeric nature of RSK 3 CTKD, Dynamic Light Scattering (DLS) was performed using 50 μl of refolded RSK 3 CTKD at a concentration of 2 mg/ml in Buffer A and the hydrodynamic radius was calculated (DynaPro NanoStar, Wyatt Technology). Similarly, DLS was also performed using 50 μl of refolded RSK 3 CTKD at acidic pH (sodium acetate buffer) and also at physiological pH 7.4. All the samples were centrifuged at 12,000 rpm for 10 min at 4 °C and also filtered with 0.45 μm filter before doing the DLS experiments.

Homology Modelling

The homology model for C-terminal end of RSK 3 was built using Modeller 9.14 software [26, 27]. Crystal structure of RSK 1 CTKD (PDB ID: 2WNT) was used as a template. The sequence identities between C-terminal end of RSK1 and RSK3 were 85 %, while positives were 92 %. The Modeller 9.14 generated five structures which were ranked based on DOPE (Discrete Optimized Protein Energy) score. The structure with least DOPE score—34348.81—was selected and validated by PROCHECK and ProSA webservers [28–30].

Results

Gene Cloning, Expression and Purification of RSK 3 CTKD (415–672)

C-terminal Kinase domain containing (415–672) amino acids of Ribosomal S6 Kinase 3 was cloned in pET-28a(+) vector. RSK 3 CTKD was expressed using Rosetta 2(DE3) bacterial strain and further purified (Fig. 2a). The predicted molecular mass of RSK 3 CTKD is ~29 kDa and that of the fusion protein is around ~33 kDa which includes the fusion tag of additional 43 amino acids at the N-terminus. As may be seen, the over expression of RSK 3 CTKD is a part of inclusion bodies. The band at 33 kDa is in agreement with the observed molecular weight. It was necessary to thoroughly wash the bacterial pellet containing RSK 3 CTKD inclusion bodies with detergents Triton X-100, Sodium Deoxycholate and Sarkosyl to remove contaminants before solubilizing in 8 M Urea. The denatured protein was purified using Ni–NTA affinity resin and analysed on 12 % SDS-PAGE gel (Fig. 2b) followed by refolding and purification in FPLC system and the fractions were analysed on 12 % SDS-PAGE gel (Fig. 2c).

The identity of the purified protein was confirmed by MALDI-TOF-TOF. Distribution of masses of peptides generated by In-gel trypsin digestion is shown in Supplementary data suppl. Fig. 1a. The peptide mass fingerprinting analysis which further confirms the purified protein to be (415–672) of Ribosomal S6 Kinase 3 protein (Supplementary data: suppl. Fig. 1b).

Secondary and Tertiary Structural Analysis of RSK 3 CTKD (415–672)

CD spectrum was recorded in the far UV region for the FPLC purified protein (Fig. 3a). The dips at $\lambda = 208$ nm and at $\lambda = 222$ nm indicated that the protein predominantly adopted α-helical form. To estimate the secondary structure, the data obtained from CD spectroscopy were submitted to K2D3 online server [25]. After matching with the closest spectrum, the server reported the majority of α-helical content, which is around 52 %, while β-strands was found to be around 8 % and random coil around 40 % at pH 8.0 (Fig. 3b). However, the secondary structure prediction server (SOPMA) predicted 44 % of α-helices, 15 % of β-strands and 41 % of random coil [31].

The blank corrected fluorescence spectrum for unfolded RSK 3 CTKD was recorded at an excitation wavelength of

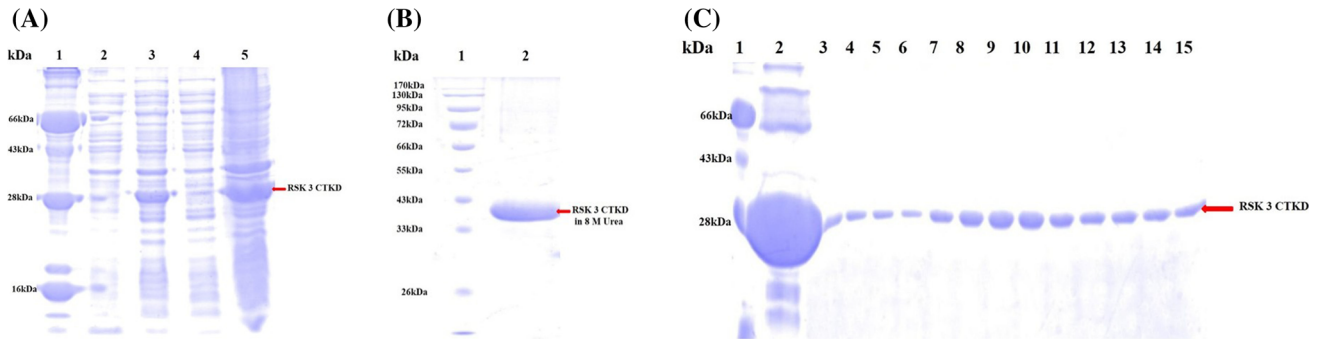


Fig. 2 RSK 3 CTKD purification and identification. **a** The 12 % SDS-PAGE gel shows the expression of RSK 3 CTKD in the induced whole fraction and pellet but not in soluble fraction suggestion the protein is insoluble. *Lane 1* Protein marker, *Lane 2* Uninduced whole cell, *Lane 3* Induced whole cell, *Lane 4* Induced soluble fraction, *Lane 5* Bacterial insoluble pellet. **b** The 12 % SDS-PAGE gel shows purified RSK 3 CTKD from inclusion bodies in 8 M Urea eluted at 250 mM Imidazole. *Lane 1* Protein marker, *Lane 2* RSK 3 CTKD in 8 M Urea. **c** The 12 % SDS-PAGE gel shows the RSK 3 CTKD purified from FPLC Superdex-200 column. *Lane 1* Protein marker, *Lane 2* Concentrated protein before FPLC, *Lanes 3–6* void fractions, *Lanes 7–15* 80 ml fractions

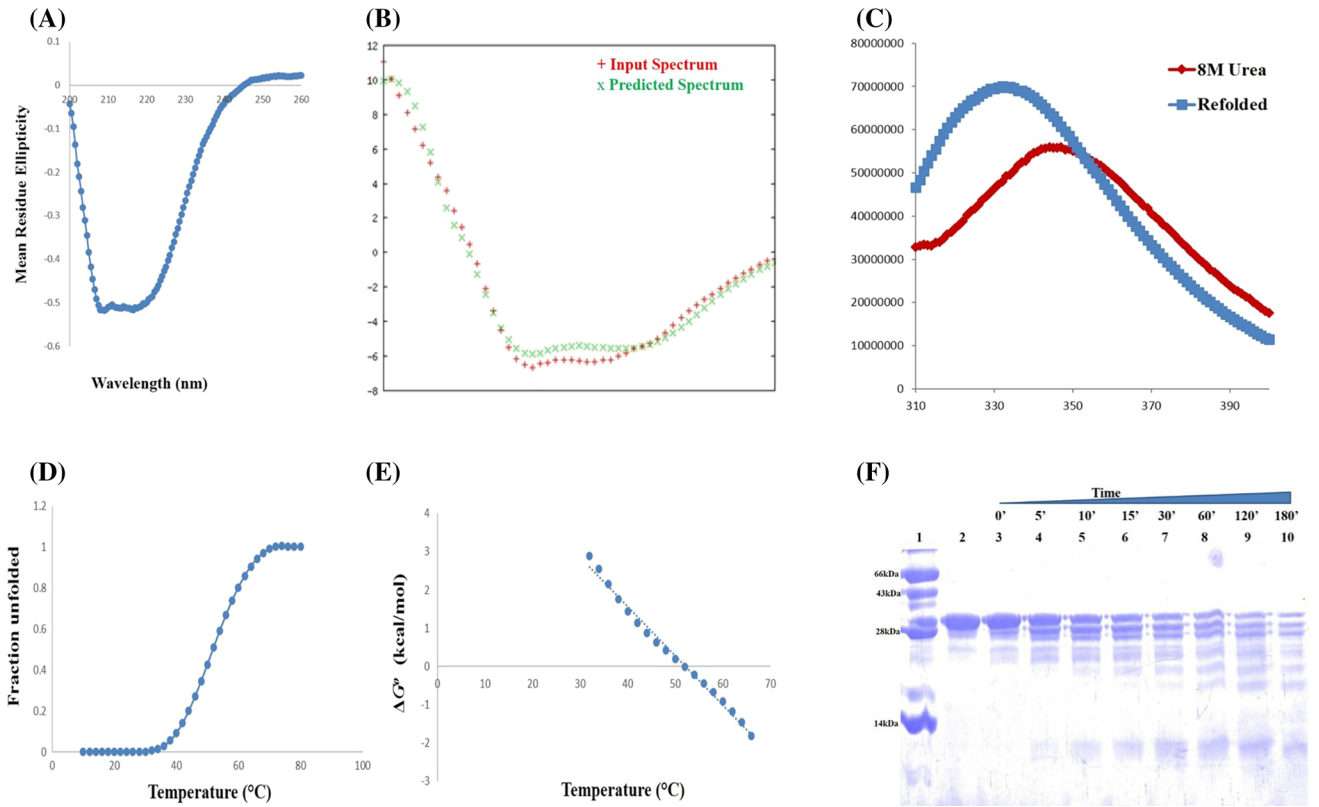


Fig. 3 Secondary and tertiary conformations of RSK 3CTKD. **a** CD Spectra of refolded RSK 3 CTKD show the dips at 208 and 222 nm indicating that it is predominantly in α-helical form. **b** The K2D3 server predicted that the protein has around 52 % α-helices and 8 % β-strands with reference to the closest matching spectrum. **c** Fluorescence spectroscopy of RSK 3 CTKD with a excitation wavelength of 280 nm shows that the absorbance maximum (λ_{max}) for unfolded protein is 347 nm, whereas for the refolded protein it is 332 nm. **d** The thermal denaturation profile of RSK 3 CTKD depicting the fraction unfolded with function of temperature showing the pre-transition and post-transition states. **e** The standard free energy of unfolding of RSK 3 CTKD graph fitted into linear extrapolation shows the T_m as the X-axis intercept which is 52 °C and the ΔG^o (H_2O) is 6.9575 kcal/mol. **f** 15 % SDS gel shows relatively slow proteolysis of the RSK 3 CTKD with trypsin-limited proteolysis, which suggests that the protein is properly folded. *Lane 1* Protein marker, *Lane 2* Negative control, *Lanes 3–10* The time intervals of trypsin digestion of RSK 3 CTKD are indicated in minutes above the gel

$\lambda = 280$ nm, which shows the emission maximum at $\lambda = 347$ nm (Fig. 3c). On the other hand, for refolded RSK 3 CTKD, the emission maximum is at $\lambda = 332$ nm (Fig. 3c). This blue shift of 15 nm confirms the protein to be properly refolded into a tertiary structure that buries the tryptophans into the hydrophobic core.

Thermal Denaturation of RSK 3 CTKD (415–672) by CD Spectroscopy

Thermal denaturation of RSK 3 CTKD was studied by CD Spectroscopy at temperatures ranging from 10 °C to 80 °C with an interval of 2 °C. The mean residue ellipticity corresponding to $\lambda = 222$ nm was used to calculate the fractions unfolded, which were then plotted with respect to temperature. The thermal denaturation study revealed that the protein is relatively stable till 30 °C and starts losing its helical characteristics from 40 °C, and at 70 °C it completely loses its secondary structure. The temperature at which 50 % of the protein is unfolded is identified as the melting temperature (T_m), which is also the midpoint of the line between the pre-transition and post-transition states (Fig. 3d). The values of T_m and ΔG^0 obtained by the linear extrapolation method [32] are 52 °C (±0.24 °C) and 6.9575 kcal/mol (±0.018 kcal/mol), respectively (Fig. 3e).

Limited Proteolysis of RSK 3 CTKD (415–672) with Trypsin

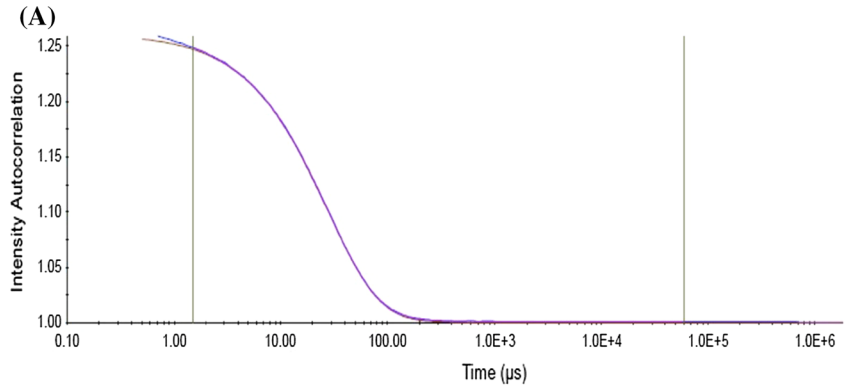
To confirm the folded state of RSK 3 CTKD, limited proteolysis with trypsin was performed (Fig. 3f). The gel pattern suggests that RSK 3 CTKD has a compact structure. The slow rate of proteolysis suggests well-folded pattern of protein making accessibility of cleavage sites for Trypsin difficult initially. However, an additional band is formed at ~9 kDa after 5 min which resists complete proteolysis till 120 min and starts degrading after three hours.

Trimeric Nature of RSK 3 CTKD

To check the oligomeric nature of RSK 3 CTKD and to calculate its hydrodynamic radius (R_h), DLS experiments were performed. The correlogram obtained is consistent with a monomodal distribution. Fast decay time at the beginning of the graph suggests the mean radius to be within the range for proteins which is usually between 1 and 10 micro seconds (μs). Furthermore, merger of graph at the end with basement suggests the absence of any aggregates. This confirms the purified RSK 3 CTKD is monodisperse in nature (Fig. 4a). In monodisperse solutions, the percentage polydispersity should be ≤20 %. The DLS of RSK 3 CTKD was performed by taking eight acquisitions out of which four showed less than 20 % polydispersity, whereas remaining four are ≤23.1 %. The derived R_h and molecular weight of the monodisperse species of RSK 3 CTKD are 4.25 ± 0.05 nm and 99 ± 4 kDa, respectively. The R_h value calculated for the 3D structure of RSK 1 CTKD is 1.5 nm which is much smaller than the value obtained. The molecular weight of 99 ± 4 kDa is approximately 3 times the size of monomer. Hence, it can be calculated that RSK 3 CTKD is a trimer in solution, at pH 8.0. (Fig. 4b), which is in agreement with the elution profiles of RSK 3 CTKD on AKTA FPLC Superdex-200 (GE) and Superdex-75 (GE) Size exclusion columns. (Fig. 4c, d). The DLS experiments were further performed at acidic pH and at physiological pH 7.4. For acidic pH, 20 mM Sodium acetate, pH 5.0 and 100 mM NaCl was used. Although there is no change in the oligomeric nature of RSK 3 CTKD, the protein tends to be less stable in acidic pH as the protein slowly precipitates with time. RSK 3 CTKD was found to be more stable at pH 8.0 and was concentrated up to 20 mg/ml concentration without precipitation for crystallization trials.

In Silico Modelling

In the absence of crystal structure of RSK 3 CTKD, in silico homology modelling was carried out to compare the in vitro and biophysical results. Since the C-terminal regions of RSK 1 and RSK 3 are 85 % identical, RSK3 was modelled using the crystal structure of RSK 1 (PDB ID: 2WNT) as a template. The modeller generated five plausible structures which were ranked based on DOPE score. The structure with the least DOPE score—34348.81—was found to be predominantly α-helical, which is in agreement with our CD spectroscopy results (Fig. 5). The model has 43 % α-helices, 15 % β-strands and 42 % random coil which is very close to the secondary structure prediction server values of 44 % α-helices, 15 % β-strands and 41 % random coil. However, the K2D3 analysis prediction of CD data suggested slightly higher α-helices of 52 % and lower β-strands of 8 %, while the percentage of random coils is almost the same. The RSK 3 CTKD model, when superimposed on the parent template 2WNT of RSK 1 CTKD, has RMSD of 0.392 Å. Like RSK 1 CTKD, RSK 3 CTKD model is also bilobal. The smaller N-lobe contains six antiparallel β-strands (1–6) and one α-helix (α1), which occurs between β4 and β5, whereas the larger C-lobe is α-rich. The substrate binding pocket is in between the lobes. We also found that there are more number of surfaces exposed hydrophobic residues in the RSK 3 CTKD model in comparison with the RSK 1 CTKD template. The relatively higher surface exposed hydrophobic residues suggests the possible trimeric behaviour of the RSK 3 CTKD. We further validated the model structure by PROCHECK



(B)

Item	Time	Temp	Intensity	Radius	%Pd	Mw-R
	(s)	(C)	(Cnt/s)	(nm)		(kDa)
Acq 1	237.4	25.0	521261	4.2	23.1	97
Acq 2	242.5	25.0	475630	4.3	18.2	100
Acq 3	247.5	25.0	470048	4.2	19.2	95
Acq 4	252.6	25.0	475164	4.2	21.6	98
Acq 5	257.6	25.0	483348	4.2	18.4	99
Acq 6	262.7	25.0	483281	4.3	18.9	100
Acq 7	267.7	25.0	481878	4.3	22.2	100
Acq 8	272.8	25.0	480873	4.3	20.1	103

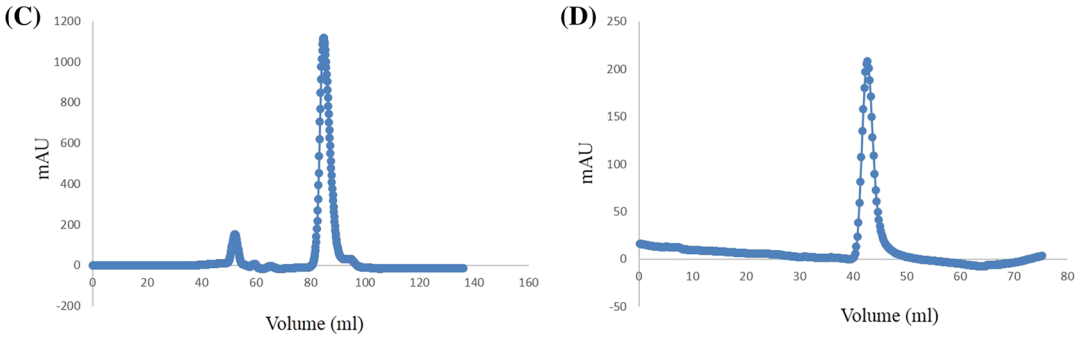


Fig. 4 Oligomeric nature of RSK 3 CTKD. **a** The correlogram of RSK 3 CTKD suggests that the protein is monomodal and monodisperse. **b** The recorded eight acquisitions suggests that the R_h and relative molecular weight are around 4.25 ± 0.05 nm and 99 ± 4, confirming the trimeric nature of RSK 3 CTKD. **c** FPLC elution profile of RSK 3 CTKD shows protein eluting at 80 from 150 ml Superdex-200 column and **d** RSK 3 CTKD eluting at 40 from 120 ml Superdex-75 column, suggesting that the protein size is larger than exclusion size which is 75 kDa

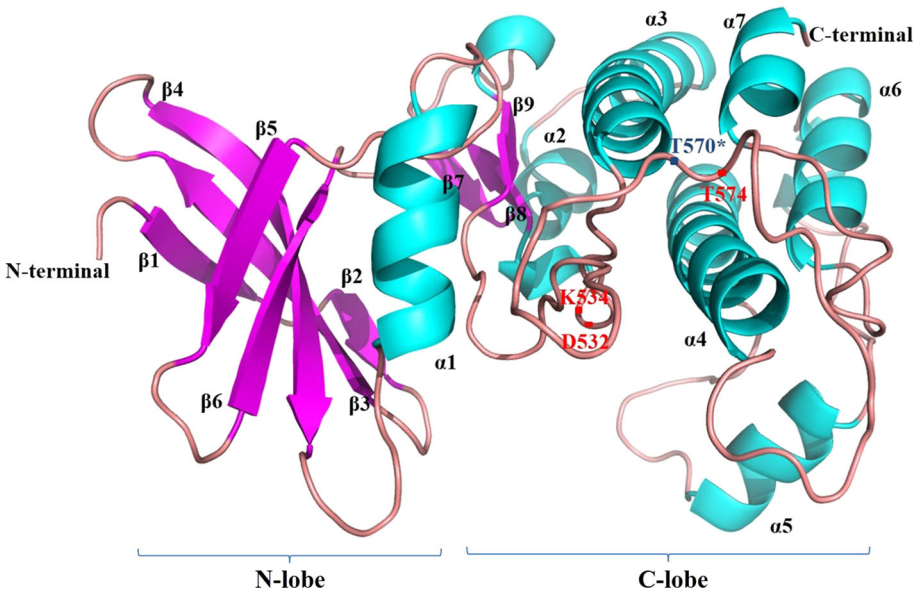


Fig. 5 Model building of RSK 3 CTKD: RSK 3 CTKD model was built on Modeller 9.14 using the crystal structure of RSK 1 (PDB ID: 2WNT) as a template. The model shows the bilobal nature of the protein and is predominantly α-helical in nature. The active site Asp532, Lys534 and Thr574 had been labelled in *red*, whereas the phosphorylation residue Thr570* had been labelled in *blue* (Color figure online)

and ProSA webservers. The Ramachandran plot obtained from PROCHECK server shows that there are 213 residues in the most favoured regions (94.7 %), 10 residues in additional allowed regions (4.4 %) and 2 residues in generously allowed regions (0.9 %). No residues were found in the disallowed regions (Supplementary data suppl. Fig. 2a). Furthermore, to validate the *z*-score of the structure we had submitted the model structure to ProSA server.

ProSA compares the z-score of the model structure with the already available z-scores of all the X-ray and NMR structures in the PDB. The z-score for RSK 3 CTKD was −7.24 which is a very good score (Supplementary data suppl. Fig. 2b). Overall PROCHECK and ProSA suggest that the RSK 3 CTKD model is of good and acceptable quality.

Discussion

Kinases are very important proteins as they are the signal transducers within the cell. Expression and purification of complex human kinases in bacterial system are very challenging as they usually form inclusion bodies. Proteins from inclusion bodies need to be denatured first, then refolded into native active form. To prevent nonspecific aggregation and precipitation during refolding, various additives like glycerol, L-arginine, DTT and Triton X-100 in various proportions were tried to determine the most suitable condition for refolding [33, 34]. In this study, we have cloned, expressed and purified RSK 3 CTKD (415-672) from *Escherichia coli* inclusion bodies. The denatured protein was successfully refolded and was further validated using CD spectroscopy, fluorescence spectroscopy and limited proteolysis. The blue shift in the emission maximum of 15 nm from 347 to 332 nm in the Fluorescence spectroscopy confirms the protein is in native condition and well folded. It has been observed that the refolded protein has proper secondary and tertiary conformation which is confirmed by multimodel in vitro, in silico and biophysical approaches. Also the thermal denaturation studies confirmed that the protein is thermally stable with T_m of 52 °C. Furthermore, DLS and FPLC elution profiles suggest that RSK 3 CTKD forms a homotrimer. This active-folded protein further can be used for protein crystallization. However, CTKD of RSK3 is phosphorylated by ERK protein. This phosphorylated CTKD further autophosphorylates its linker region substrate.

Considering the functional importance of RSK3, a model was also built for RSK 3 CTKD. The model structure has a very similar conformation as that of RSK 1 CTKD and the superimposed structures have acceptable RMSD of 0.392 Å. The model structure also retained the basic bilobal structure as seen in RSK 1 CTKD. The homotrimeric nature from our in vitro experiments is further supported by the relatively higher surface exposed hydrophobic residues in the RSK 3 CTKD model in comparison with the template. The model after validation with PROCHECK and ProSA webservers found that there are no Ramachandran outliers, and a z-score of −7.24 suggests that the model is well defined and is of good quality. So far there are no crystal structures available for RSK 3 from any organism. In contrast to other isoforms of RSK, only RSK 3 is reported to have tumour suppressor activity [35]. Hence, understanding its 3D structure will throw a greater light on the mechanism and reported multifunctionality of RSK 3, which would also address as how RSK 3 functions as tumour suppressor. In conclusion, the folding pattern of CTKD of RSK 3 is predominantly α-helical, and is a homotrimer, which is in sharp contrast to the behaviour of RSK 1 and RSK 2 CTKDs.

Acknowledgments We thank CRI common facility, DBT funded BTIS, Proteomics facility at ACTREC. Mr Bhanu thanks UGC, New Delhi for fellowship. Dr. M.V. Hosur is thankful to DAE for RRF award.

References

1. Sutherland, C., Campbell, D. G., & Cohen, P. (1993). Identification of insulin-stimulated protein kinase-1 as the rabbit equivalent of rskmo-2. Identification of two threonines phosphorylated during activation by mitogen-activated protein kinase. *European Journal of Biochemistry, 212*(2), 581–588.
2. Smith, J. A., et al. (1999). Identification of an extracellular signal-regulated kinase (ERK) docking site in ribosomal S6 kinase, a sequence critical for activation by ERK in vivo. *Journal of Biological Chemistry, 274*(5), 2893–2898.
3. Jensen, C. J., et al. (1999). 90-kDa ribosomal S6 kinase is phosphorylated and activated by 3-phosphoinositide-dependent protein kinase-1. *Journal of Biological Chemistry, 274*(38), 27168–27176.
4. Carriere, A., et al. (2008). The RSK factors of activating the Ras/MAPK signaling cascade. *Front Biosci, 13*, 4258–4275.
5. Jones, S. W., et al. (1988). A Xenopus ribosomal protein S6 kinase has two apparent kinase domains that are each similar to distinct protein kinases. *Proceedings of the National Academy of Sciences, 85*(10), 3377–3381.
6. Fisher, T. L., & Blenis, J. (1996). Evidence for two catalytically active kinase domains in pp90rsk. *Molecular and Cellular Biology, 16*(3), 1212–1219.
7. Frodin, M., & Gammeltoft, S. (1999). Role and regulation of 90 kDa ribosomal S6 kinase (RSK) in signal transduction. *Molecular and Cellular Endocrinology, 151*(1–2), 65–77.
8. Chen, R. H., Sarnecki, C., & Blenis, J. (1992). Nuclear localization and regulation of erk- and rsk-encoded protein kinases. *Molecular and Cellular Biology, 12*(3), 915–927.
9. De Cesare, D., et al. (1998). Rsk-2 activity is necessary for epidermal growth factor-induced phosphorylation of CREB protein and transcription of c-fos gene. *Proceedings of the National Academy of Sciences, 95*(21), 12202–12207.
10. Joel, P. B., et al. (1998). pp90rsk1 regulates estrogen receptor-mediated transcription through phosphorylation of Ser-167. *Molecular and Cellular Biology, 18*(4), 1978–1984.
11. Zhao, J., et al. (2003). ERK-dependent phosphorylation of the transcription initiation factor TIF-IA is required for RNA polymerase I transcription and cell growth. *Molecular Cell, 11*(2), 405–413.
12. Nakajima, T., et al. (1996). The signal-dependent coactivator CBP is a nuclear target for pp90RSK. *Cell, 86*(3), 465–474.
13. Roberts, P. J., & Der, C. J. (2007). Targeting the Raf-MEK-ERK mitogen-activated protein kinase cascade for the treatment of cancer. *Oncogene, 26*(22), 3291–3310.

14. Jagilinki, B. P., et al. (2014). Conserved residues at the MAPKs binding interfaces that regulate transcriptional machinery. *Journal of Biomolecular Structure and Dynamics, 33*(4), 1–9.
15. Bjorbaek, C., Zhao, Y., & Moller, D. E. (1995). Divergent functional roles for p90rsk kinase domains. *Journal of Biological Chemistry, 270*(32), 18848–18852.
16. Vik, T. A., & Ryder, J. W. (1997). Identification of serine 380 as the major site of autophosphorylation of Xenopus pp90rsk. *Biochemical and Biophysical Research Communications, 235*(2), 398–402.
17. Roux, P. P., Richards, S. A., & Blenis, J. (2003). Phosphorylation of p90 ribosomal S6 kinase (RSK) regulates extracellular signal-regulated kinase docking and RSK activity. *Molecular and Cellular Biology, 23*(14), 4796–4804.
18. Ikuta, M., et al. (2007). Crystal structures of the N-terminal kinase domain of human RSK1 bound to three different ligands: implications for the design of RSK1 specific inhibitors. *Protein Science, 16*(12), 2626–2635.
19. Malakhova, M., et al. (2009). Structural diversity of the active N-terminal kinase domain of p90 ribosomal S6 kinase 2. *PLoS One, 4*(11), e8044.
20. Malakhova, M., et al. (2008). Structural basis for activation of the autoinhibitory C-terminal kinase domain of p90 RSK2. *Nature Structural & Molecular Biology, 15*(1), 112–113.
21. Li, D., et al. (2012). Structural basis for the autoinhibition of the C-terminal kinase domain of human RSK1. *Acta Crystallographica. Section D, Biological Crystallography, 68*(Pt 6), 680–685.
22. Serafimova, I. M., et al. (2012). Reversible targeting of noncatalytic cysteines with chemically tuned electrophiles. *Nature Chemical Biology, 8*(5), 471–476.
23. Shevchenko, A., et al. (2006). In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nature Protocols, 1*(6), 2856–2860.
24. Pappin, D. J., Hojrup, P., & Bleasby, A. J. (1993). Rapid identification of proteins by peptide-mass fingerprinting. *Current Biology, 3*(6), 327–332.
25. Louis-Jeune, C., Andrade-Navarro, M. A., & Perez-Iratxeta, C. (2012). Prediction of protein secondary structure from circular dichroism using theoretically derived spectra. *Proteins, 80*(2), 374–381.
26. Fiser, A., & Sali, A. (2003). Modeller: generation and refinement of homology-based protein structure models. *Methods in Enzymology, 374*, 461–491.
27. Eswar, N., et al. (2006). Comparative protein structure modeling using Modeller. *Current Protocols in Bioinformatics, Chapter 5*, 5–6.
28. Laskowski, R. A., et al. (1996). AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *Journal of Biomolecular NMR, 8*(4), 477–486.
29. Sippl, M. J. (1993). Recognition of errors in three-dimensional structures of proteins. *Proteins, 17*(4), 355–362.
30. Wiederstein, M., & Sippl, M. J. (2007). ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Research, 35*(Web Server issue), W407–W410.
31. Geourjon, C., & Deleage, G. (1995). SOPMA: significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments. *Computer Applications in the Biosciences, 11*(6), 681–684.
32. Pace, C. N., & Shaw, K. L. (2000). Linear extrapolation method of analyzing solvent denaturation curves. *Proteins, Suppl 4*, 1–7.
33. Armstrong, N., de Lencastre, A., & Gouaux, E. (1999). A new protein folding screen: application to the ligand binding domains of a glutamate and kainate receptor and to lysozyme and carbonic anhydrase. *Protein Science, 8*(7), 1475–1483.
34. Holbourn, K. P., & Acharya, K. R. (2011). Cloning, expression and purification of the CCN family of proteins in Escherichia coli. *Biochemical and Biophysical Research Communications, 407*(4), 837–841.
35. Bignone, P. A., et al. (2007). RPS6KA2, a putative tumour suppressor gene at 6q27 in sporadic epithelial ovarian cancer. *Oncogene, 26*(5), 683–700.