IDENTIFICATION OF PROTEIN PARTNERS FOR NIBP, A NOVEL NIK-AND IKKB-BINDING PROTEIN THROUGH EXPERIMENTAL, COMPUTATIONAL AND BIOINFORMATICS TECHNIQUES

A Thesis
Submitted to
the Temple University Graduate Board

In Partial Fulfillment
of the Requirements for the Degree
MASTER OF SCIENCE

---

by
Sombudha Adhikari
August 2013

Thesis Approvals:

Mohammad. F. Kiani, Ph.D., Thesis Advisor, College of Engineering
Wenhui Hu, M.D., Ph.D., School of Medicine
Roland Dunbrack, Ph.D, Fox Chase Cancer Center
Peter Lelkes, Ph.D., College of Engineering

ABSTRACT

Identification of protein partners for NIBP, a novel NIK- and IKKβ-binding protein through experimental, computational and bioinformatics techniques

NIBP is a prototype member of a novel protein family. It forms a novel subcomplex of NIK-NIBP-IKKβ and enhances cytokine-induced IKKβ-mediated NFκB activation. It is also named TRAPPC9 as a key member of trafficking particle protein (TRAPP) complex II, which is essential in *trans*-Golgi networking (TGN). The signaling pathways and molecular mechanisms for NIBP actions remain largely unknown.

The aim of this research is to identify potential proteins interacting with NIBP, resulting in the regulation of NFκB signaling pathways and other unknown signaling pathways. At Dr. Wenhui Hu's lab in the Department of Neuroscience, Temple University, sixteen partner proteins were experimentally identified that potentially bind to NIBP.

NIBP is a novel protein with no entry in the Protein Data Bank. From a computational and bioinformatics standpoint, we use prediction of secondary structure and protein disorder as well as homology-based structural modeling approaches to create a hypothesis on protein-protein interaction between NIBP and the partner proteins. Structurally, NIBP contains three distinct regions. The first region, consisting of 200 amino acids, forms a hybrid helix and beta sheet-based domain possibly similar to Sybindin domain. The second region comprised of approximately 310 residues, forms a

tetratrico peptide repeat (TPR) zone. The third region is a 675 residue long all beta sheet and loops zone with as many as 35 strands and only 2 helices, shared by Gryzun-domain containing proteins. It is likely to form two or three beta sheet sandwiches. The TPR regions of many proteins tend to bind to the peptides from disordered regions of other proteins. Many of the 16 potential binding proteins have high levels of disorder. These data suggest that the TPR region in NIBP most likely binds with many of these 16 proteins through peptides and other domains. It is also possible that the Sybindin-like domain and the Gryzun-like domain containing beta sheet sandwiches bind to some of these proteins.

DEDICATION

To my daughter, Paula and my wife Gargi.

ACKNOWLEDGEMENTS

I gratefully acknowledge the help and support of my advisor, Dr. Mohammad F Kiani, coadviser Dr. Wenhui Hu, coadvisor Dr. Roland Dunbrack, and Dr. Peter Lelkes. Their knowledge, experience and enthusiasm gave me the opportunity to complete this project. I thank Dr. Yonggang Zhang, postdoctoral fellow, and others in Dr. Hu's laboratory for their support on experiments and experimental data. I also thank Mr. Maxim Shapovalov, Dr. Qifang Xu, and others in Dr. Dunbrack's laboratory. In addition, I would like to acknowledge support and help from all the faculty and graduate students in engineering school, and the neuroscience department of Temple University School of Medicine.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

Chronic diseases, cancers and diabetes are associated with dysregulation of many biochemical cues. These biochemical cues are proteins that regulate cellular activity, migration and death [1, 2]. The synthesis of these proteins is regulated by nuclear transcription factors. One of the most studied transcription factors is nuclear factor kappa B (NFκB), which plays an important role in regulating the expression of various inflammatory mediators [3,4]. Genes that cause inflammation, immunity, cell survival and neural plasticity are regulated by the signaling of NFκB. Many different proteins have been identified that regulate the activity of NFκB via canonical and non-canonical signal transduction pathways. Yet, how these proteins regulate NFκB signaling is still unclear.

The self-renewal/proliferation, survival, migration and lineage differentiation of neural stem/progenitor cells are a series of processes for both embryonic and adult neurogenesis. Complex signal transduction pathways, inherent, and extrinsic factors firmly regulate neurogenesis. A large variety of neurodevelopmental defects and neurodegenerative diseases are associated with dysfunctional regulation of neurogenesis. NFκB signaling regulates various stages of neurogenesis and mediates interactions with other signal transduction pathways such as Notch, Shh, Wnt/β-catenin during the dynamic process of neurogenesis. In addition, NFκB signaling mediates the effect of numerous factors such as chemokines, cytokines, various growth factors, and

extracellular matrices contributing to neurogenesis. Understanding the process of regulation of NFκB and identification of associated proteins is important for developing drugs to treat various chronic diseases.

NIBP is a prototype member of a novel protein family. It regulates NFκB signaling in both canonical and non-canonical pathways. It forms a novel subcomplex of NIK-NIBP-IKKβ and enhances cytokine-induced IKKβ-mediated NFκB activation. It is also named TRAPPC9 as a key member of trafficking particle protein (TRAPP) complex II, which is essential in *trans*-Golgi networking (TGN). Both NFκB and TGN are critical in many physiological processes and pathological diseases.

NIBP mutation and deletion have been shown to be closely correlated with autosomal-recessive mental retardation, autism and stroke. Previous studies have shown that NIBP/NFκB signaling plays key roles in neurogenesis. NIBP is also highly expressed in cancer cells and regulates tumorigenesis. However, the signaling pathways and molecular mechanisms for NIBP actions remain largely unknown.

Regulation of NFκB activity

Many different proteins and other compounds like small molecules have been identified as having the ability to regulate NFκB activity. Yet, how these proteins and compounds regulate NFκB is still unclear. Moreover, the regulatory mechanisms may be different in different cell types. The different dynamics of NFκB activity may be caused by different signaling pathways. A canonical pathway and an alternative non-

canonical pathway for NFκB activation have been identified [5]. The canonical pathway is triggered by stimuli such as TNFα and IL-1β and depends on the IκB kinase (IKK), which consists of two catalytic subunits (IKKα and IKKβ) and a regulatory IKK subunit. The nuclear translocation of NFκB dimers (mainly p65/p50) and the activation of target genes is caused by IKK-mediated phosphorylation of the inhibitor proteins of NFκB (IκBs) to induce their ubiquitination and degradation. The alternative non-canonical pathway relies on the phosphorylation of IKKα by NFκB inducing kinase (NIK) to induce p100 processing into p52 and the nuclear translocation of RelB/p52 dimer. The non-canonical pathway is regulated by destabilization of NIK through TRAF-cIAP complex.

NIBP

The transcription factor NFκB plays an important role in both physiological and pathological events in the central nervous system. Nevertheless, the mechanisms of NFκB-mediated regulation of gene expression, and the signaling molecules participating in the NFκB pathway in the central nervous system are, to date, poorly understood. To identify such molecules, Dr. Hu conducted a yeast two-hybrid screen of a human brain cDNA library using NIK as bait. As a result, a novel NIK and IKKβ binding protein designated NIBP was identified that is mainly expressed in brain, muscle, heart, and kidney. Interestingly, low levels of expression were detected in immune tissues such as

spleen, thymus, and peripheral blood leukocytes, wherein NFκB is known to modulate immune function. Dr. Hu's laboratory demonstrated that NIBP expression in the brain is localized to neurons. NIBP physically interacts with NIK, IKKβ, but not IKKα or IKKγ. NIBP overexpression potentiates tumor necrosis factor-alpha-induced NFκB activation through increased phosphorylation of the IKK complex and its downstream IκBα and p65 substrates. Finally, knockdown of NIBP expression by small interfering RNA reduces tumor necrosis factor-alpha-induced NFκB activation, prevents nerve growth factor-induced neuronal differentiation, and decreases Bcl-xL gene expression in PC12 cells. The data from Dr. Hu's laboratory demonstrates that NIBP, by interacting with NIK and IKKβ, is a new enhancer of the cytokine-induced NFκB signaling pathway. Because of its neuronal expression, Dr. Hu's laboratory proposed that NIBP may be a potential target for modulating the NFκB signaling cascade in neuronal pathologies dependent upon abnormal activation of this pathway.

NIBP is a prototype member of a novel protein family. It forms a novel subcomplex of NIK-NIBP-IKKβ without IKKα and IKKγ and enhances cytokine-induced IKKβ-mediated NFκB activation. It is known to control NFκB activation via canonical and noncanonical pathways. It is also named TRAPPC9 as a key member of trafficking particle protein (TRAPP) complex II, which is essential in *trans*-Golgi networking (TGN). Both NFκB and TGN are critical in many physiological processes and pathological diseases. NIBP mutation or deletion has been shown to be closely correlated with autosomal-recessive mental retardation, autism and stroke. Previous

studies at Dr. Hu's laboratory have shown that NIBP/NFκB signaling plays key roles in neurogenesis. However, the signaling pathways and molecular mechanisms for NIBP actions remain largely unknown.

NIBP is a novel protein. It is not available in the Protein Data Bank (PDB) for the experimentally determined protein structures. It has various isoforms in different species. The most common form of human NIBP protein consists of 1148 amino acids (aa) and weighs 128.53 kDa. It is present mainly in the brain, muscles, heart and kidney and in limited amounts in the immune organs. NIBP is confined to neurons in the central nervous system, possibly conferring specificity to the NFκB pathway and playing a part in neuronal cell differentiation. NIBP is a subunit of the multiprotein complex TRAPP (transport particle protein), and is encoded by the gene TRAPPC9. It might play the role of a transport protein, responsible for transport from the ER to the Golgi apparatus, based on observations on other members of the TRAPP complex. Mutations in TRAPPC9 have been shown to cause mental retardation and diseases like hypoplasia and microcephaly, suggesting that the protein might play a role in human brain development.

CHAPTER 2

EXPERIMENTAL DATA

All biological experiments were done by Dr. Yonggang Zhang, postdoctoral fellow, and others in Dr. Wenhui Hu's lab, Department of Neuroscience, Temple University School of Medicine, Philadelphia, PA, USA. I participated as a trainee and as an observer.

In Dr. Hu's laboratory, Flag-tagged human NIBP(1148aa) was expressed in HEK293T cells and purified through immunoprecipitation with Flag antibody. ImaGenes' high-density protein macroarrays (UniPex-1) were incubated with Flag-NIBP(1148) protein and detected by standard Western blot with anti-Flag antibody and SuperSignal West Femto Substrate. The resulting images were analyzed manually. The interaction of positive clones with NIBP was further confirmed by co-immunoprecipitation and Western blot analysis (Figure 3). Sixteen partner proteins were detected by ImaGenes' high-density protein macroarrays (UniPex-1) that bind with NIBP. Five NIBP partner proteins are known for stem cell self-renewal and differentiation, seven are known for cellular trafficking, and four for heme signaling.

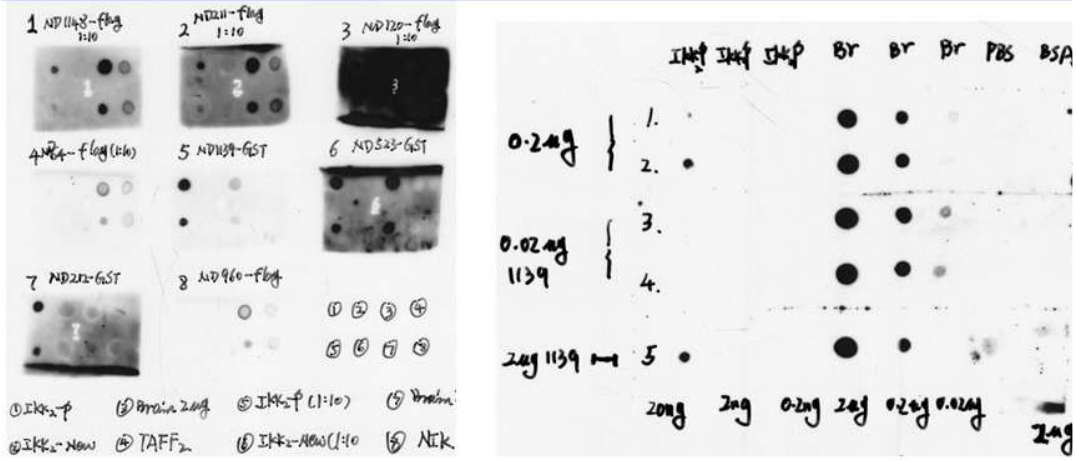


Figure 1. Efficient interaction of purified Flag- or GST-tagged NIBP fusion proteins with positive IKKβ and brain protein extracts determined by dot blotting. Data collected and recorded by Dr. Yonggang Zhang, postdoctoral fellow, and others in Dr. Wenhui Hu's lab, Department of Neuroscience, Temple University School of Medicine, Philadelphia, PA, USA.

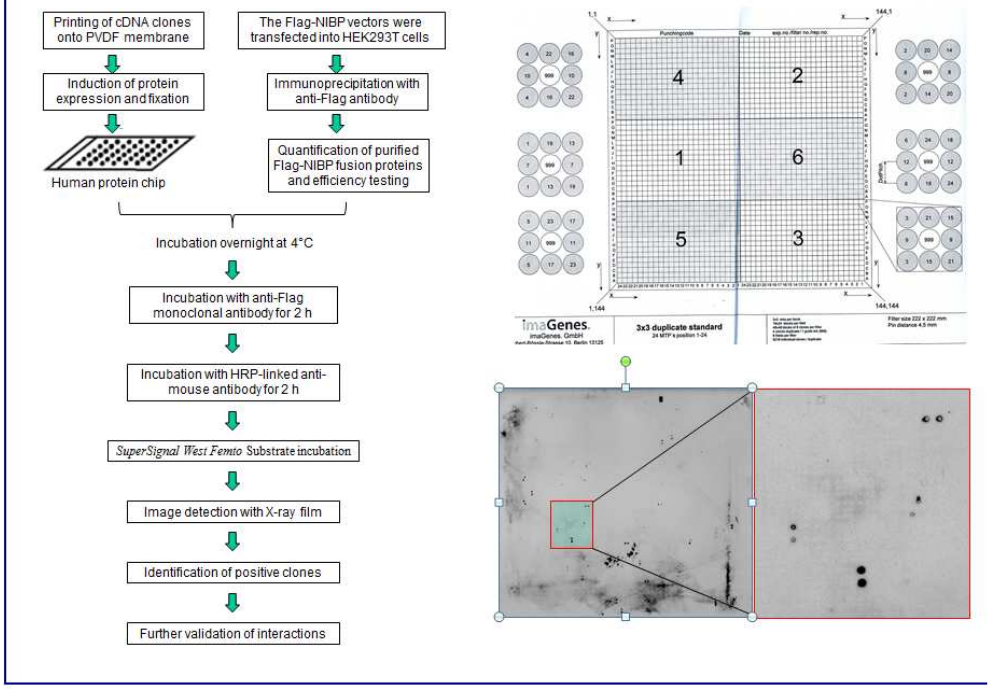


Figure 2. Protocol and representative data of protein macroarray. Data collected and recorded by Dr. Younggang Zhang, postdoctoral fellow, and others in Dr. Wenhui Hu's lab, Department of Neuroscience, Temple University School of Medicine, Philadelphia, PA, USA.

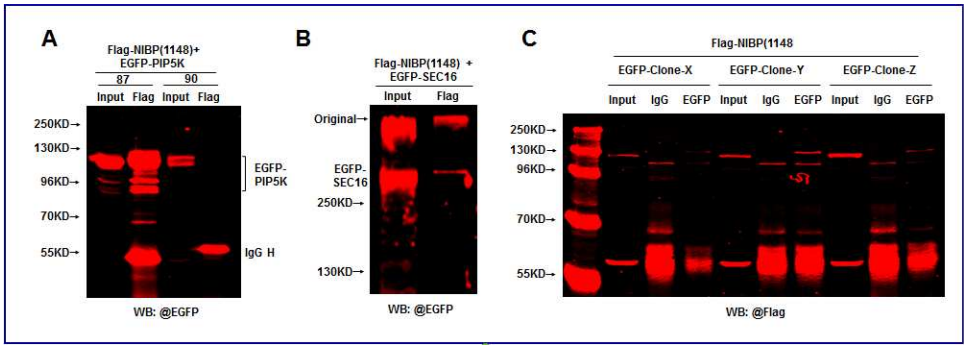


Figure 3. Co-IP and Western Blot. Data collected and recorded by Dr. Younggang Zhang, postdoctoral fellow, and others in Dr. Wenhui Hu's lab, Department of Neuroscience,Temple University School of Medicine, Philadelphia, PA, USA.

.

CHAPTER 3

STRUCTURAL MODELING OF NIBP

There is no entry for NIBP in Protein Data Bank (PDB). In order to explore or map the interaction between NIBP and the sixteen potential partner proteins, multiple approaches of structural modeling of NIBP are utilized. Protein disorder prediction modeling provides disordered regions within NIBP. Secondary structure prediction depicts the pattern of helix and beta sheet strands. Pfam and clan analyses coupled with hidden Markov modeling (HMM)-HMM comparison (HHSearch) provide further indication of possible domains and details on topology in this novel protein.

3.1 Disorder Prediction of NIBP

Protein disorder prediction using Neural Network based Iupred program identified that NIBP is a highly ordered protein. Figure 4 shows that almost the whole protein is ordered. Full NIBP protein sequence was obtained from Uniprot. The sequences were input in neural network based Iupred Website[6]. The resultant graphs of predicted disorder and structured regions were obtained. It outputs a disorder value between 0 and 1. Above the 0.5 line, especially for extended sequence regions greater than 30 aa, means an intrinsically disordered region of protein. Short regions above the line may be long loops in otherwise ordered proteins.

GI: 238624124 - NIBP 1148

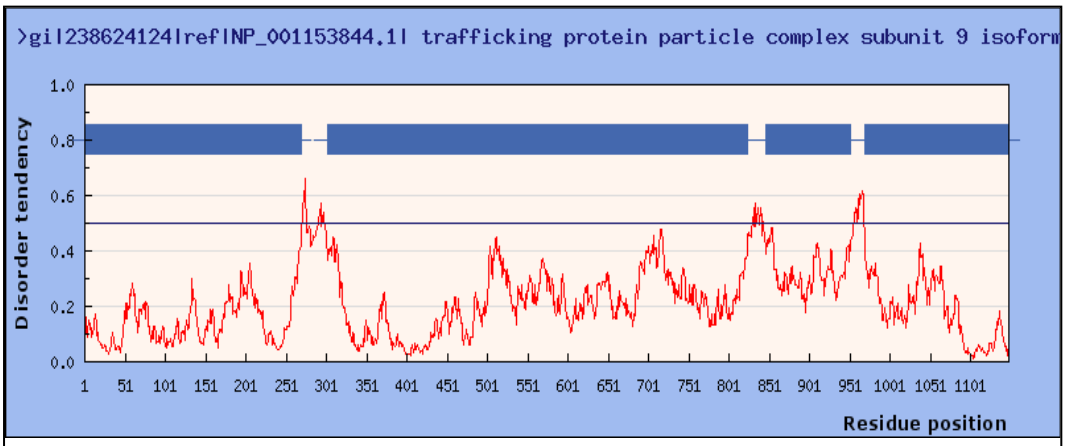


Figure 4. Disorder tendency of NIBP in Neural Network based Iupred

3.2 Secondary Structure Prediction

Secondary structure prediction tools provide the topology of the protein. In essence it predicts the pattern of the helix and betasheet strands in different regions. Bioassembly Modeler from Dr. Roland Dunbrack's lab was used for secondary structure prediction. Second iteration result from PsiBlast was used in PsiPred against Uniref 90 database to obtain high quality secondary structure prediction. In the resulting figures, red stand for helix and green stands for beta sheets. Analysis of the secondary structure prediction of NIBP clearly divides the protein into three distinct regions based on topology.

The region 1 consists of residues 1-170 with a topology of E-H-E-E-E-H-E-E-H**.** E stands for betasheets and H stands for alpha helix.

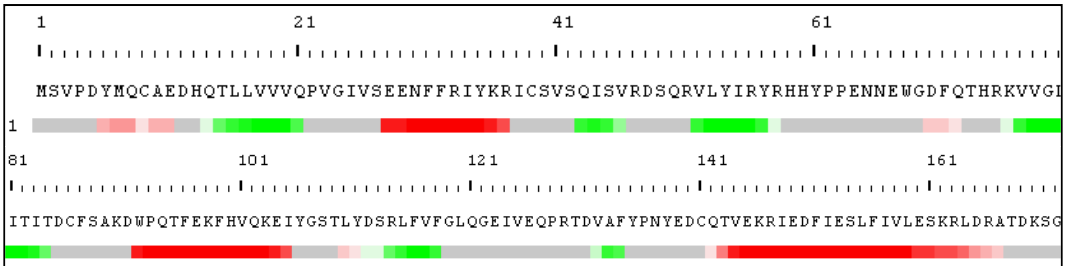


Figure 5. Region 1: residues 1-170, topology = E-H-E-E-E-H-E-E-H
The region2 consists of residues 200-510. In this region all are H-loop-H-loop H-loop. There are 14 helices in total.

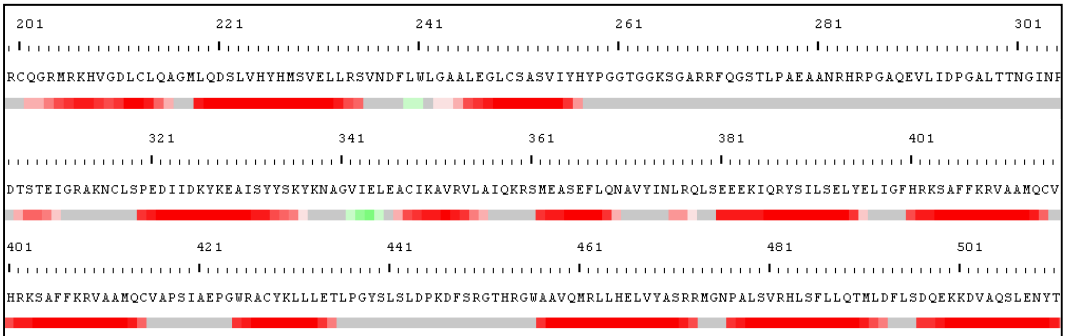


Figure 6. Region2: residues 200-510: all H-loop-H-loop H-loop = 14 helices.
Region 3 consists of residues 510-1148. In this region nearly all are beta sheet and loops with as many as 35 strands and only 2 helices (at 970 and 1010).

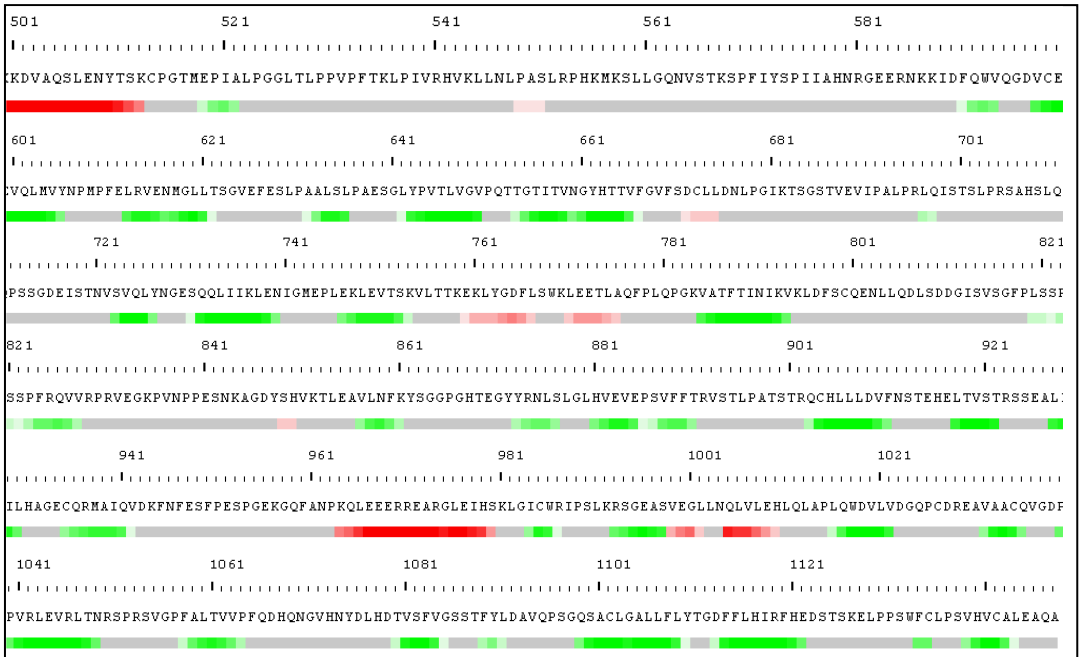


Figure 7. Region 3: residues 510-1185: nearly all beta sheet and loops with as many as 35 strands and only 2 helices (at 970 and 1010).

3.3 Pfam and the Clan

Proteins contain one or more functional regions, commonly known as domains. The function of the protein becomes clear when the domains present in a protein can be identified.

The Pfam database from the Sanger Institute is a large collection of protein domains and related families. Multiple sequence alignments and HMMs represent each protein family. NIBP has one distinctive Pfam called TRAPPC9-Trs120.

Figure 8. NIBP Pfam

The Pfam for NIBP is TRAPPC9-Trs120; this Pfam is in a clan with Sybindin, Clat_adaptor_s, and sedlin_N, Gryzun, Gryzun-like, and TRAPPC10. Three of these proteins are small 140 aa proteins of similar known structures. These are shown in Figure 11. The HMMs are all about 140 aa as well. These proteins have secondary structure E-E-H-E-E-E-H-H, and resemble Region 1 secondary structure of NIBP (strands are hard to predict to there could be another one in NIBP). The three remaining Pfam domains – Gryzun, Gryzun-like, and TRAPPC10 will be discussed further below.

Figure 9. TRS 120 Clan

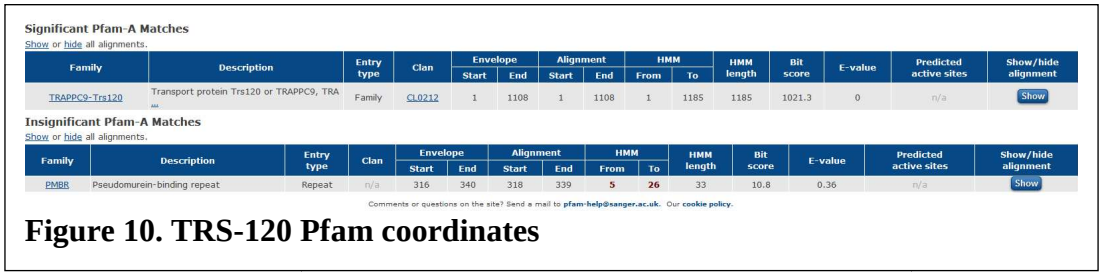
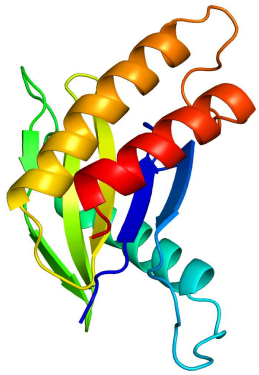
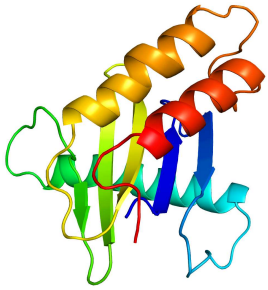


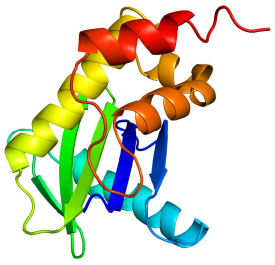
Figure 10. TRS-120 Pfam coordinates



Clan0212 Sedlin



Clan0212 Sybindin



Clan0212 Clat_adaptor_s

Figure 11. Small domains from the NIBP Snare Clan (CL0212)

3.4 Analysis with HHSearch

Homology detection & structure prediction by HMM-HMM comparison, HHPred server was used to align known Pfam domain to the NIBP protein. [7]

Pfam Gryzun aligns to 595-1122 of NIBP, and this is almost all of the Gryzun Pfam (25-554 of 554 lengths). This region coincides with almost the entire beta sheet predicted region of NIBP, and is presumably why TRAPPC9-Trs120 is in this clan.

Figure 12. HHSearch Results from NIBP sequence

Pfam TRAPPC-Trs85 aligns to 80-448 of NIBP, and corresponds to part of Region 1 and most of region 2, and is mostly predicted to be alpha helical. TPR-repeat containing proteins of known structure align to 200-500 of NIBP, consistent with secondary structure prediction and nature of NIBP as a potential scaffold protein for other proteins.

3.5 TPR Repeat Proteins

The tetratrico peptide repeat (TPR) proteins tend to bind other proteins or peptide regions from disordered regions of other proteins. As a structural motif, it mediates protein–protein interactions and the assembly of multiprotein complexes [9]. Proteins containing TPRs have been reported to be involved in a variety of biological processes, such as neurogenesis, mitochondrial and peroxisomal protein transport, transcriptional control, and protein folding [10,11]. TPR domains have been associated with molecular recognition and protein–protein interactions [9]. The structural and thermodynamic

studies of the TPR domains of Hsp70/Hsp90 organizing protein (Hop) [12] show a detailed description of protein–protein interactions mediated by TPR domains. TPR domains mediate protein–protein interactions in a variety of ways. Three tandem TPR motifs are the smallest functional unit that is widely used. Concave face of the repeat receives binding residues and binds the residues with high specificity.

3.6 HHSearch to Build Model Based on 3ro3 TPR

The following figure shows the alignment of 3ro3 TPR present in NIBP protein.

```
>3ro3_A PINS homolog, G-protein-signaling modulator 2; asymmetric cell division, protein binding; 1.10A {Mus
musculus}
Probab=97.86  E-value=0.00051  Score=69.99  Aligned_cols=152  Identities=15%  Similarity=0.198  Sum_probs=0.0

Q ss_pred              hcchhhhhhhhhhHHHHhCCHHHHHHHHHHHHHHHHhcCchHHHHHHHHHHHHHHHHHHhCcCCcccccccccccccc
Q sp|Q96Q05|TPPC  199 KRCQGRMRKHVGDLCLQAGMLQDSLVHYHMSVELLRSVNDFLWLGAALEGLCSASVIYHYPGGTGGKSGARRFQGSTLPA  278 (1148)
Q Consensus       199 kR~~GR~~K~lGDl~LlaG~~~dAL~~Y~eA~~~lks~~D~lW~A~ALEg~~~a~~ll~~~~~~~~~~~~~~r~~nssl~~  278 (1148)
                        ++..||...-||.+|+..|+|+.|+++|..|+++++..+|..+.+.++-.++.+...++
T Consensus         5 ~~~~~~~~~~~lg~~y~~~g~~~~A~~~~~~~al~i~~~~~~~~~~~~~~~~~~~lg~~y~~~~--------------------  63 (164)
T 3ro3_A           5 RAAQGRAFGNLGNTHYLLGNFRDAVIAHEQRLLIAKEFGDKAAERIAYSNLGNAYIFLG--------------------  63 (164)
T ss_dssp             HHHHHHHHHHHHHHHHHHTTCHHHHHHHHHHHHHHHHHHHTCHHHHHHHHHHHHHHHHHHTT--------------------
T ss_pred             HHHHHHHHHHHHHHHHHHcCCHHHHHHHHHHHHHHHHHHHcCCChHHHHHHHHHHHHHHHHcC--------------------


Q ss_pred              cccccCCCcccccccccccccccccccchhhhhccCCHHHHHHHHHHHHHHHHhhccCccEEhHHHHHHHHHHHhhcc
Q sp|Q96Q05|TPPC  279 EAANRHRPGAQEVLIDPGALTTNGINPDTSTEIGRAKNCLSPEDIIDKYKEAISYYSKYKNAGVIELEACIKAVRVLAIQ  358 (1148)
Q Consensus       279 ~~~~k~r~~~~~~~~~~~~p~~~~~~~~~~~~~~~~~~~~~~~~l~~~~~i~~~y~~Ai~~Y~k~~~~~~~le~ea~~k~~r~l~~~  358 (1148)
                                                          ..++.+.++++|+..+.+..... ....++...+.++..+
T Consensus        64 ---------------------------------------~~~~A~~~~~~~al~~~~~~~~~~~~~~~~l~~~~~~~~  102 (164)
T 3ro3_A          64 ---------------------------------------EFETASEYYKKTLLLARQLKDRA-VEAQSCYSLGNTYTLL  102 (164)
T ss_dssp             ---------------------------------------CHHHHHHHHHHHHHHHHHTTCHH-HHHHHHHHHHHHHHHHT
T ss_pred             ---------------------------------------CHHHHHHHhHHHHhhhhhhhccc-ccchhhhHHHHHHHHH


Q ss_pred              cchHHHHHHHHHHHhhhcchhhCcHHHHHHHHHHHHHHHHHhcCCchHhhHHHHHHHHH
Q sp|Q96Q05|TPPC  359 KRSMEASEFLQNAVYINLRQLSEEEKIQRYSILSELYELIGFHRKSAFFKRVAA  412 (1148)
Q Consensus       359 ~~~~eaa~~L~~a~~~~l~~l~~~~eq~~~~~~~La~vy~~lG~~RK~AF~lRla~  412 (1148)
                        +.+.++.+.+.+.++....-+.-........++..||.+|..+|=..|+--+.+.|.
T Consensus       103 ~~~~~A~~~~~~~~a~~~~~~~~~~~~~~~~~~~~~~la~~~~~~~~~~~~~~~~A~~~~~~~al  156 (164)
T 3ro3_A         103 QDYEKAIDYHLKHLAIAQELKDRIGEGRACWSLGNAYTALGNHDQAMHFAEKHL  156 (164)
T ss_dssp             TCHHHHHHHHHHHHHHHHHHHTTCHHHHHHHHHHHHHHHHHHHTCHHHHHHHHHHHH
T ss_pred             hhHHHHHHHHHHHHHHHHHHHcCchHHHHHHHHHHHHHHHHHcCCHHHHHHHHHHHH
```

Figure 13. Alignment of 3ro3 TPR present in NIBP protein

The following figure shows the model of the 3ro3 TPR present in NIBP protein build with the program MODELLER[8].

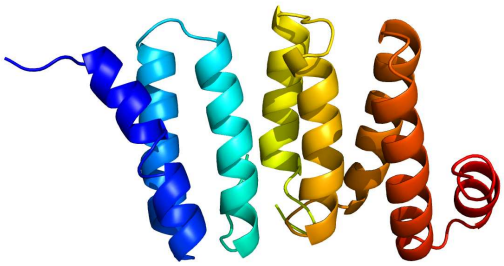


Figure14. Model of the 3ro3 TPR present in NIBP protein

3.7 NIBP Structural Model

NIBP structural modeling shows three distinct regions. The first region: residues 1-170, topology is E-H-E-E-E-H-E-E-H, which seems to be similar to Sybindin and Sedlin domains. The second region consists of residues 200-510: all H-loop-H-loop H-loop = 14 helices. It is the TPR region known for binding with other proteins through peptides and other domains. The third region is the residues 510-1185: nearly all beta sheet and loops with as many as 35 strands and only 2 helices (at 970 and 1010). It resembles a bunch of beta sheet sandwiches. It resembles the Gryzun domain.

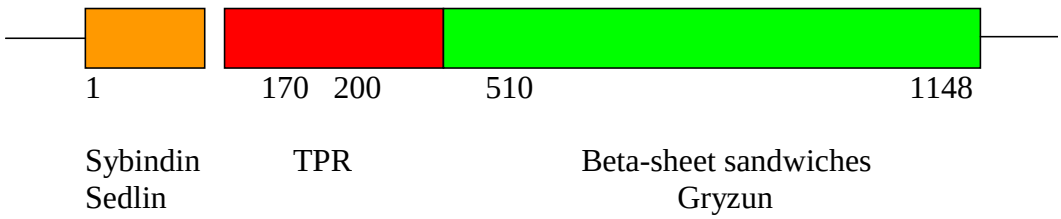


Figure15. NIBP structural model

Sybindin is associated with spinogenesis. It is a physiological syndecan-2 ligand. It is found on. small protrusions on the surface of dendrites, called dendritic spines that

receive the vast majority of excitatory synapses. Syndecan-2 induces spine formation by recruiting intracellular vesicles toward postsynaptic sites through the interaction with synbindin[13,14,15,16,17,18,19]. Sedlin is a 140 aa protein with a commonly accepted role in endoplasmic reticulum-to-Golgi transport. Spondyloepiphyseal dysplasia tarda, a progressive skeletal disorder is caused by several missense mutations and deletion mutations in the SEDL gene, which result in protein truncation by frame shift[20,21,22,23,24,25,26]. Gryzun is distantly related to the Trs130 subunit of the TRAPP complex. RNAi of human Gryzun (Q7Z392) putatively blocks Golgi exit. As part of the TRAPP complex, the Gryzun family is likely to be involved with trafficking of proteins through membranes [27,28,29,30,31,32]. All these features in NIBP structure support previous finding that NIBP is a key member of TRAPP complex involved in trans-Golgi networking. We predict that NIBP plays important role in regulating spinogenesis and axonal transport.

3.8 TRAPP Complex Proteins

We also examined the proteins in the TRAPP complex [33,34,35] using HHSearch to determine what the topology of the various components is. The results are shown in Table 1.

The smaller proteins (TRAPPC2, TRAPPC2-like, TRAPPC3, TRAPPC3-like, TRAPPC4, and TRAPPC6) are all small proteins of 140-219 amino acids. Three of them contain the small members of Pfam Clan0212, which resemble each other (Figure 11). An additional three proteins contain the TRAPP domain fold (TRAPPC3, TRAPPC3-

Table 1. TRAPP complex proteins

Protein	Uniprot	Length	Domains/Secondary structure
TRAPPC2	TPC2A_HUMAN	140	Sedlin_N
TRAPPC2L	TPC2L_HUMAN	140	Sedlin_N
TRAPPC3	TPPC3_HUMAN	180	TRAPP
TRAPPC3L	TPC3L_HUMAN	181	TRAPP
TRAPPC4	TPPC4_HUMAN	219	Sybindin
TRAPPC5	TPPC5_HUMAN	188	TRAPP
TRAPPC8	TPPC8_HUMAN	1435	Alpha-beta (1-230); TPRs (330-660); beta-sheet-region (650-1435) (Gryzun)
TRAPPC9	TPPC9_HUMAN	1148	Alpha-beta (1-160); TPRs (200-510); beta-sheet-region (520-1148) (Gryzun)
TRAPPC10	TPC10_HUMAN	1259	Alpha-beta (1-170); TPRs (170-570); beta-sheet region (570-1259) (TRAPPC10)
TRAPPC11	TPC11_HUMAN	1133	Alpha-beta (1-180); TPRs (180-570); beta-sheet region (570-1133) (Gryzun-like)
TRAPPC12	TPC12_HUMAN	735	TPRs (270-735)

like and TRAPPC5), which is also an alpha-beta fold although with a different topology than Sedlin_N and sybindin.

Four members of this complex, including NIBP/TRAPPC9, resemble each other in their secondary structure predictions and the proteins and Pfams that may be aligned to the sequences. These are TRAPPC8, TRAPPC9, TRAPPC10, and TRAPPC11. All four contain a region of 160-230 amino acids that alternate between two or three sheet strands and alpha helices, thus resembling sybindin and sedlin_N. Then there is a region of 300-400 amino acids in each protein of long helices broken up by predicted coil regions, aligning to TPR-containing proteins. The last region consists entirely of beta sheet strands separated by coil regions. The Pfams Gryzun, Gryzun-like, and TRAPPC10 all align to these regions in one or more of these four proteins, as demonstrated by HHSearch[7]. Thus we propose that these four proteins are homologous, arising from a single common ancestor.

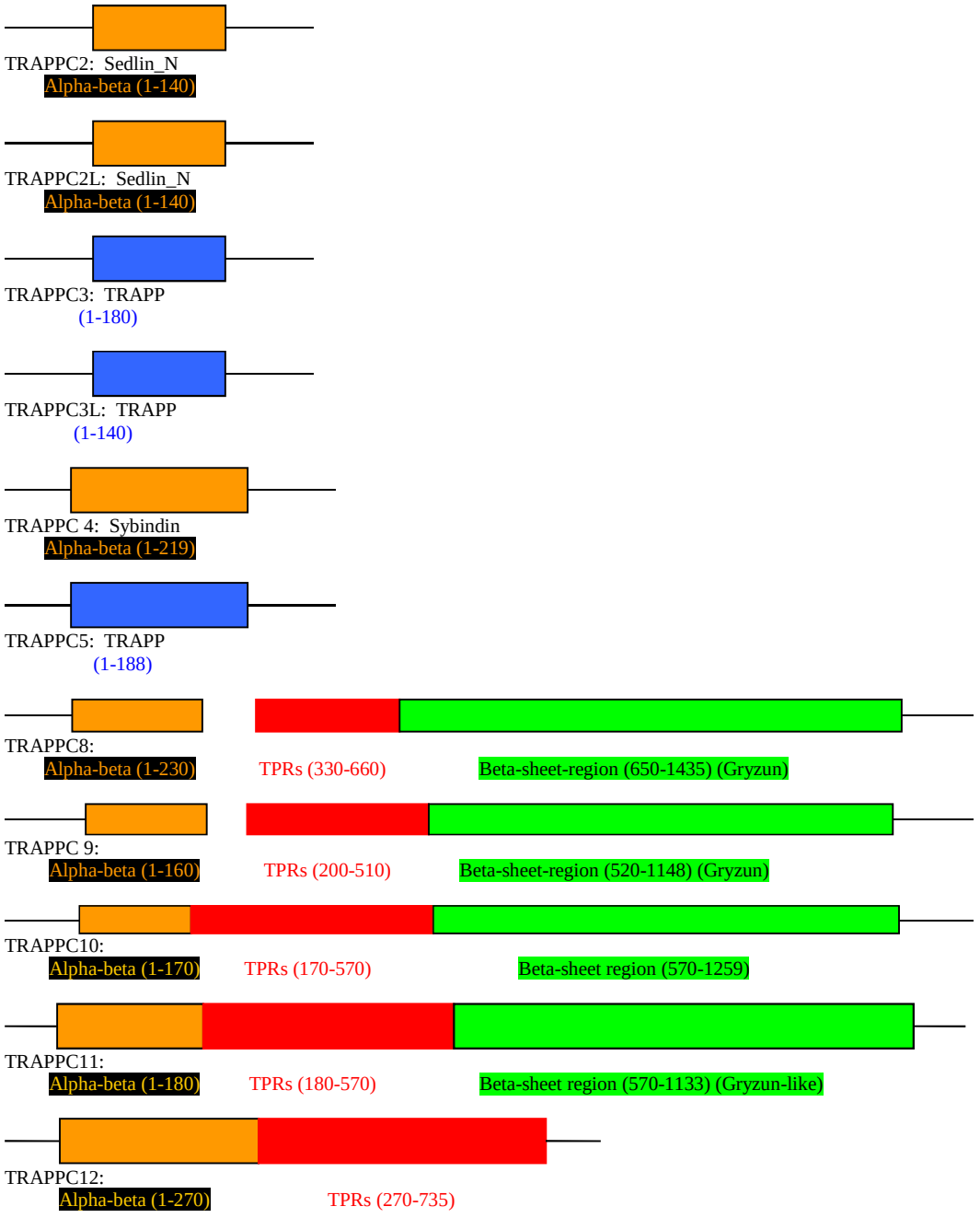


Figure16. TRAPPC Complex Proteins

CHAPTER 4

PARTNER PROTEINS

4.1 Disordered Regions in Partner Proteins

Many of the NIBP partner proteins identified by Protein macroarray are extremely disordered. For example, in humans, REST gene encodes RE1-Silencing Transcription factor (REST), also known as Neuron-Restrictive Silencer Factor (NRSF). It acts as a silencer. REST represses neural genes in non-neuronal cells. Alterations in the REST expression pattern putatively cause many genetic disorders. Huntington Disease, neuroblastomas, and the effects of epileptic seizures and ischaemia are also associated with REST [36,37]. PIP5K1C or Phosphatidylinositol-4-phosphate 5-kinase type-1 gamma is an enzyme encoded by the PIP5K1C gene in humans [38]. Transport protein Sec16A is required for secretory cargo traffic to the Golgi apparatus from the endoplasmic reticulum [39].

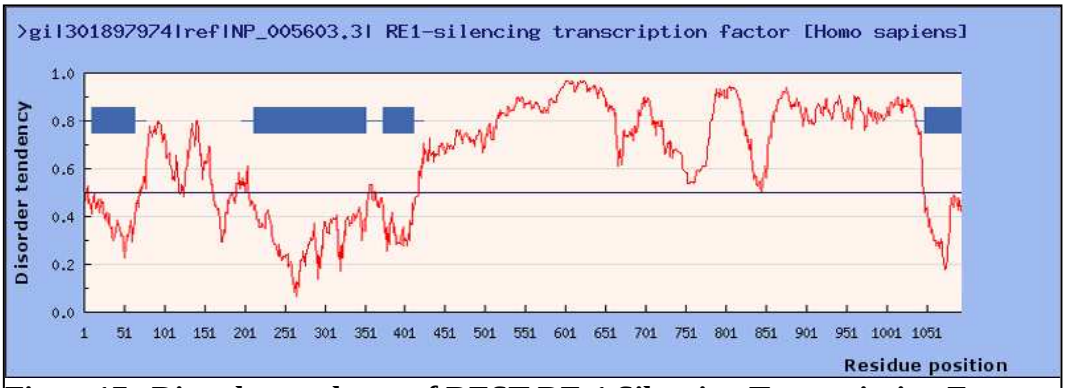


Figure17. Disorder tendency of REST RE-1 Silencing Transcription Factor

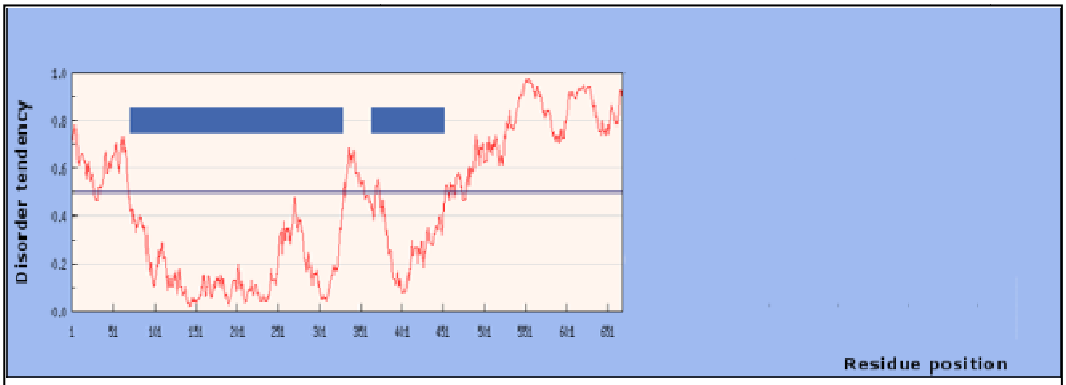


Figure18. Disorder tendency of PIP5K1C

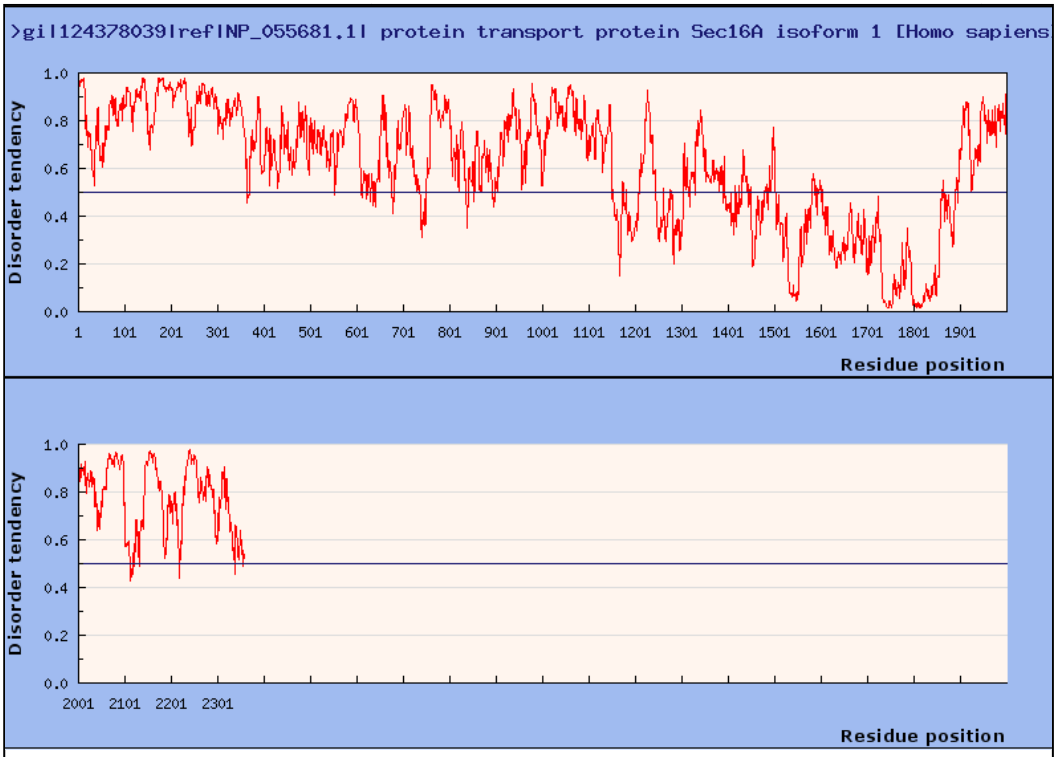


Figure19. Disorder tendency of Transport protein Sec16A

4.2 What TPRs in Pfam bind to:

The following table shows what TPRs in Pfam bind to. TPRs are known to bind to peptides, HSP90, SCP2 and so on. NIBP contains a distinct region of TPR repeats. From the HHSearch results of NIBP, the following TPRs were found as shown below in Tables 2. Protcid database was used to browse and find the Pfams that bind to these specific TPRs. It was interesting to observe if these Pfams are in common with partner protein Pfams.

Table 2. Pfams that bind to NIBP specific TPRs

TPR	Pfam id	Name	Pfam Domains									
TPR_1	PF00515	Tetratricop-eptide repeat	peptide	Hsp90	Ras	SCP2	YopD	GerE	Clathrin_lg_ch	Aminotran_5	Pfam-B_6614	Response_reg
TPR_11	PF13414	TPR repeat	peptide	HSP90	SCP2	YopD	APC_CDC26	Aminotran_5	(fn3)_(BRCT)			
TPR_12	PF13424	Tetratricop eptide repeat	peptide	Response_reg	Pfam-B_6614	Pfam-B_11735	Bacillus_PapR					
TPR_14	PF13428	Tetratricop eptide repeat										
TPR_16	PF13432	Tetratricop eptide repeat	peptide	HSP90	APC_CDC26							
TPR_17	PF13431	Tetratricop eptide repeat	peptide	HSP90								
TPR_2	PF07719	Tetratricop eptide repeat	peptide	Mdv1	DUF3249	Response_reg	Aminotran_5	APC_CDC26				
TPR_3	PF07720	Tetratricop eptide repeat	peptide	Ras	YopD	UQ_con						

4.3 Possibilities of binding with other partner proteins

Our structural models show that NIBP has three distinct regions. The TPR region is well known for mediating binding with other proteins. However, it is possible that the first region resembling Sybindin domain and the beta sheet sandwich region can also be involved in binding. Figure 20 shows the model of the TPR region in NIBP. The model is done with HHPred[7].

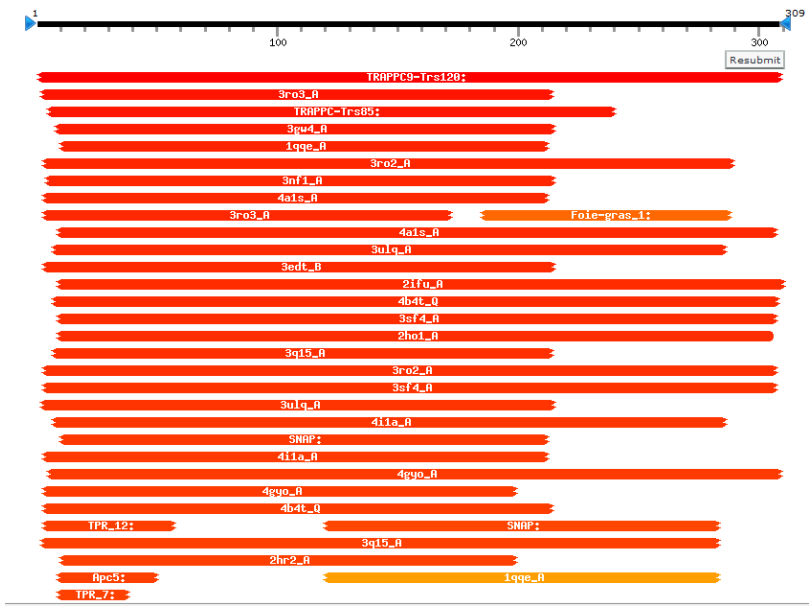


Figure 20. Model of the TPR region in NIBP with HHPred.

The TPRs bind to other proteins through Peptides and other domains. Figure 21 shows TPR 3ro3 of NIBP binding with a peptide.

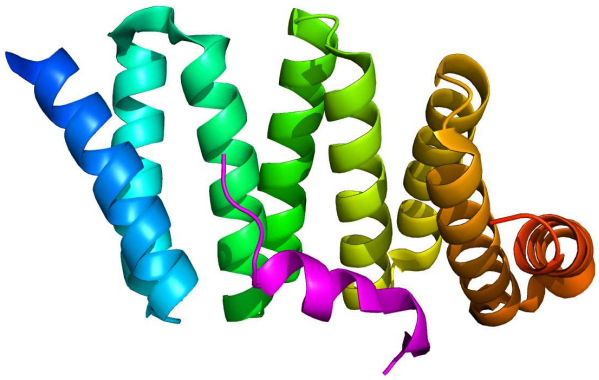


Figure 21. 3ro3 TPR present in NIBP binding with Peptide.

Figure 22 shows the model of the sandwiched beta sheet region of NIBP obtained through HHPred[7].

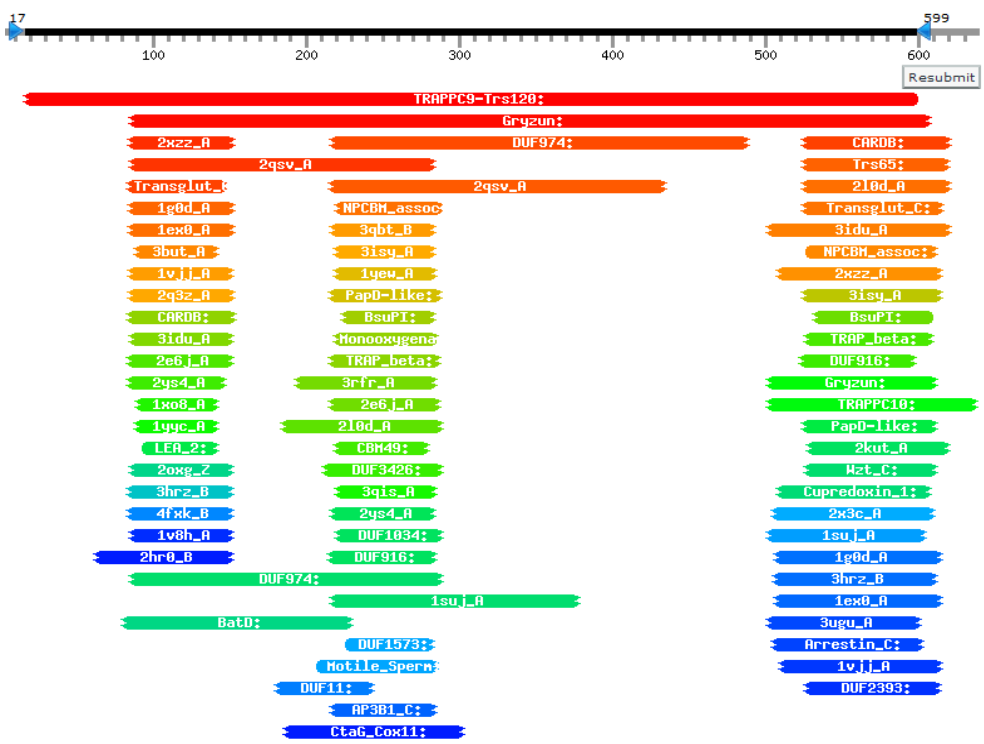


Figure 22. Model of NIBP beta sheet domains in HHPred.

CHAPTER 5

CONCLUSION

5.1 Regulation of NFκB activity via NIBP and Partner Protein Interactions

The novel proteins, NIBP (NIK and IKK2 binding protein), has been demonstrated to increase IKK2-mediated NFκB activation and be required for growth and differentiation of neuronal cell line PC12. NIBP is renamed TRAPPC9 because it is a key member of trafficking protein particle (TRAPP) complex II, implying its importance in regulating trans-Golgi networking and the TRAPP family of protein complexes.

NIBP apparently binds with many partner proteins which is evident from the Protein Macroarray experimental results in Dr. Hu's lab. There is further evidence of binding from Co-IP and Western Blot results in Dr. Hu's lab although further investigation is warranted.

For example, Homo sapiens RE1-silencing transcription factor (REST) binds with NIBP to control neural stem cell self-renewal. PIP5K1C regulates spermogenesis and stem cell differentiation. Similarly, based on macroarray data, sixteen different proteins bind to NIBP to perform certain functions in stem cell selfrenewal and differentiation, cellular trafficking, or heme signaling.

5.2 Correlation of Experimental, Computational and Bioinformatics Models

The experimental and bioinformatics analysis show possibilities of docking between NIBP and some of the partner proteins. Experimentally Dr. Hu' lab demonstrated that there is initial evidence of interaction between NIBP and the partner proteins. NIBP structural modeling shows three distinct regions. The first region: residues 1-170, topology (= E-H-E-H-E-E-H) seems to be similar to Sybindin domain. The second region consists of residues 200-510: all H-loop-H-loop H-loop = 14 helices. It is the TPR region known for binding with other proteins through peptides and other domains. The third region is the residues 510-1185: nearly all beta sheet and loops with as many as 35 strands and only 2 helices (at 970 and 1010). It resembles a bunch of beta sheet sandwiches. It resembles the Gryzun domain. The TPR region is well known for mediating binding with other proteins, especially with peptides. However, it is possible that first region resembling Sybindin domain and the beta sheet sandwich region resembling Gryzun domain can also be involved in binding.

5.3 Uncertainties in Experimental, Computational, and Bioinformatics Models

There are uncertainties in Macroarray and Co-IP/Western Blot experiments mainly from lack of reliable antibodies. Five antibodies were bought and tested in co-IP but only two confirmed the interaction. The other three antibodies did not work well by themselves for Western Blot and IP. Further experimental evidence is needed to validate the interactions of NIBP with its partner proteins.

Computational and bioinformatics analysis are based upon probabilities. These uncertainties require further careful analysis. Bioinformatics use many algorithms and software tools that are untested in specific protein or conditions in question. Machine learning approaches sometime help in making the software tool or algorithm learn. Machine learning methodologies were not used in the project.

Positive controls using IKKβ or NIK as the established partners for NIBP should be picked up for the computational modeling in the future.

5.4 Future Work:

The data obtained in this research effort can be used to further investigate the intricacies of protein-protein interactions between NIBP and the partner proteins. First, experimentally each domain may be expressed by itself and tested for their function in mammalian cell system. The first domain may include 1-170 or 1-250 residues. Second, it may be worthwhile to run ab initio structure prediction like I-Tasser on the first predicted domain (sybindin/Sedlin), the second predicted domain or region of TPRs, and the third predicted domain (Gryzun/Gryzun-like). Finally, mutations of conserved regions within each domain can be performed to disrupt the interactions with partner proteins.

BIBLIOGRAPHY

[1] Pikarsky E, Porat R M, Stein I, Abramovitch R, Amit S, Kasem S, Gutkovich-Pyest E, Urieli-Shoval S, Galun E and Ben-Neriah Y, NF-kB functions as a tumour promoter in inflammation-associated cancer. NATURE, VOL. 431, 2004, 461-466.

[2] Bierhaus A, Schiekofer S, Schwaninger M, et al, Diabetes-Associated Sustained Activation of the Transcription Factor Nuclear Factor-κB. DIABETES, VOL. 50, 2001, 2792-2808.

[3] Karin M and Greten F R, NF kappa B: Linking inflammation and immunity to cancer development and progression. NATURE REVIEWS IMMUNOLOGY, Vol. 5, 2005, Iss.:10, 749-759.

[4] Karin M, Nuclear factor-κB in cancer development and progression. NATURE, Vol. 441, 2006, 431-436.

[5] Bonnizzi G and Karin M, The two NF-kappaB activation pathways and their role in innate and adaptive immunity. Trends Immunol. 2004 Jun, 25(6):280-8.

[6] Prediction of Intrinsically Unstructured Proteins, IUPred server, http://iupred.enzim.hu/.

[7] HHpred – Homology detection & structure prediction by HMM-HMM comparison, HHpred webserver, http://toolkit.tuebingen.mpg.de/hhpred

[8] Modeller Program for Comparative Protein Structure Modellin by Satisfaction of Spatial Restraints, website, http://salilab.org/modeller/.

[9] Luca D. D'Andrea and Lynne Regan, TPR proteins: the versatile helix. TRENDS in Biochemical Sciences, Vol.28, No.12 December 2003.

[10] X Blatch, G.L. and Lassle, M. (1999) The tetratricopeptide repeat: a structural motif mediating protein–protein interactions. BioEssays 21, 932–939;

[11] Y Goebl, M. and Yanagida, M. (1991) The TPR snap helix: a novel protein repeat motif from mitosis to transcription. Trends Biochem. Sci. 16, 173–177.

[12] Scheufler, C. *et al.* (2000), Structure of TPR domain–peptide complexes: critical elements in the assembly of the Hsp70–Hsp90 multichaperone machine. Cell 101, 199–210.

[13] Ethel, I.M., *et al.*, Synbindin, A novel syndecan-2-binding protein in neuronal dendritic spines. J Cell Biol. 2000 Oct 2;151(1):53-68.

[14] Cole G.J., Halfter W. Agrinan extracellular matrix heparan sulfate proteoglycan involved in cell interactions and synaptogenesis. Perspect. Dev. Neurobiol. 1996;3:359–371.

[15] Ethell I.M., Yamaguchi Y. Cell-surface heparan sulfate proteoglycan syndecan-2 induces the maturation of dendritic spines in rat hippocampal neurons. J. Cell Biol. 1999;144:575–586.

[16] Schubert D. The possible role of adhesion in synaptic modification. Trends Neurosci. 1991;14:127–130.

[17] Shi S.-H., Hayashi Y., Petralia R.S., Zaman S.H., Wenthold R.J., Svoboda K., Malinow R. Rapid spine delivery and redistribution of AMPA receptors after synaptic NMDA receptor activation. Science. 1999;284:1811–1816.

[18] Spacek J., Harris K.M. Three-dimensional organization of smooth endoplasmic reticulum in hippocampal CA1 dendrites and dendritic spines of the immature and mature rat. J. Neurosci. 1997;17:190–203.

[19] Srivastava S., Osten P., Vilim F.S., Khatri L., Inman G., States B., Daly C., DeSouza S., Abagyan R., Valtschanoff J.G., Weinberg R.J., Ziff E.B. Novel anchorage of GluR2/3 to the postsynaptic density by the AMPA receptor-binding protein ABP. Neuron. 1998; 21:581–591.

[20] Gedeon, A.K., et al., The molecular basis of X-linked spondyloepiphyseal dysplasia tarda, American Journal of Human Genetics,2001, 68(6):1386-1397.

[21] Maquat, L.E., (1996) Defects in RNA splicing and the consequence of shortened translational reading frames. Am J Hum Genet 59:279–286.

[22] Miller, A.P., Willard H.F., (1998) Chromosomal basis of X chromosome inactivation: identification of a multigene domain in Xp11.21-p11.22 that escapes X inactivation. Proc Natl Acad Sci USA 95:8709–8714.

[23] Mumm, S, Christie, P.T., Finnegan, P., Jones, J., Dixon, P.H., Pannett AAJ, Harding B, Gottesman GS, Thakker RV, Whyte MP (2000) A five-base pair deletion in the sedlin gene causes spondyloepiphyseal dysplasia tarda in a six-generation Arkansas kindred. J Clin Endocrinol Metab 85:3343–3347.

[24] Nagaraja, R., MacMillan, S., Kere, J., Jones, C., Griffin, S., Schmatz, M., Terrell, J., Shomaker, M., Jermak, C., Hott C, Masisi, M., Mumm S, Srivastava, A., Pilia, G., Featherstone, T., Mazzarella, R., Kesterson, S., McCauley, B., Railey, B., Burough, F., Nowotny, V., D'Urso, M, States, D., Brownstein, B., Schlessinger, D., (1997) X chromosome map at 75-kb STS resolution, revealing extremes of recombination and GC content. Genome Res 7:210–222.

[25] Tiller, G.E., Hannig, V.L., Dozier, D., Carrel, L., Trevarthen, K.C., Wilcox, W.R., Mundlos, S., Haines, J.L., Gedeon, A.K., Gecz, J., A recurrent RNA-splicing mutation in the SEDL gene causes X-linked spondyloepiphyseal dysplasia tarda. Am J Hum Genet 68:1398–1407.

[26] Shapiro, M.B., Senepathy, P., RNA splice junctions of different classes of eukaryotes: sequence statistics and functional implications in gene expression. Nucleic Acids Res 15:7155–7175.

[27] Wendler, F., Gillingham, A.K., Sinka, R., Rosa-Ferreira, C., Gordon, D.E., Franch-Marro, X., Peden, A.A., Vincent, J.P., Munro, S.,A genome-wide RNA interference screen identifies two novel components of the metazoan secretory pathway,EMBO J. 2010 Jan 20,29(2):304-14.

[28] Bard, F., Casano, L., Mallabiabarrena, A., Wallace, E., Saito, K., Kitayama, H., Guizzunti, G., Hu, Y., Wendler, F., Dasgupta, R., Perrimon, N., Malhotra, V., Functional genomics reveals genes involved in protein secretion and Golgi organization. Nature 439: 604–607.

[29] Lee, M.C., Miller, E.A., Goldberg, J., Orci L, Schekman, R., Bi-directional protein transport between the ER and Golgi. Annu Rev Cell Dev Biol 20: 87–123.

[30] Levine, T.P., Munro, S., The pleckstrin homology domain of oxysterol-binding protein recognises a determinant specific to Golgi membranes. Curr Biol 8: 729–739.

[31] Adams, M..D., Celniker, S.E., Holt, R.A., Evans, C.A., Gocayne, J.D., Amanatides, P.G., Scherer, S.E., Li, P.W., Hoskins, R.A., Galle, R.F., George, R.A., Lewis, S.E., Richards, S., Ashburner, M., Henderson, S.N., Sutton, G.G., Wortman, J.R., Yandell, M.D., Venter, J.C.,The genome sequence of Drosophila melanogaster, Science 287:2185-2195.

[32] Misra, S., Crosby, M.A., Mungal,l C.J., Matthews, B.B., Campbell, K.S., Hradecky, P., Huang, Y., Kaminker, J.S., Millburn, G.H., Prochnik, S.E., Smith, C.D., Tupy, J.L., Whitfied, E.J., Bayraktaroglu, L., Berman, B.P., Bettencourt, B.R., Celniker, S.E., de Grey, A.D., Lewis, S.E.,Annotation of the Drosophila melanogaster

euchromatic genome: a systematic review, Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).

[33] Kim, Y.G., Raunser, S., Munger, C., Wagner, J., Song, Y.L., Cygler, M., Walz, T., Oh, B.H., Sacher, M.,The architecture of the multisubunit TRAPP I complex suggests a model for vesicle tethering, Cell, 2006 Nov 17, 127(4): 817-30.

[34] Kim, M.S., Yi, M.J., Lee, K.H., Wagner, J., Munger, C., Kim, Y.G., Whitewa, M., Cygler, M., Oh, B.H., Sacher, M.,Biochemical and crystallographic studies reveal a specific interaction between TRAPP subunits Trs33p and Bet3p, Traffic, 2005 Dec, 6(12): 1183-95.

[35] Sacher, M.,Kim,Y.,Lavie,A., Oh, B., Segev,N., The TRAPP Complex: Insights into its Architecture and Function, Traffic 2008 December, 9(12): 2032–2042.

[36] Schoenherr,C.J., Anderson, D.J., The neuron-restrictive silencer factor (NRSF): a coordinate repressor of multiple neuron-specific genes, Science 267 (5202): 1360–3.

[37] Chong, J.A., Tapia-Ramírez, J., Kim, S., Toledo-Aral, J.J., Zheng, Y., Boutros, M.C., Altshuller, Y.M., Frohman, M.A., Kraner, S.D., Mandel, G.,REST: a mammalian silencer protein that restricts sodium channel gene expression to neurons, Cell 80 (6): 949–57.

[38] Ishihara, H., Shibasaki, Y., Kizuki, N., Wada, T., Yazaki, Y., Asano, T., Oka, Y., Type I phosphatidylinositol-4-phosphate 5-kinases. Cloning of the third isoform and deletion/substitution analysis of members of this novel lipid kinase family, J Biol Chem 273 (15): 8741–8.

[39] Bhattacharyya, D.., Glick, B.S.,Two mammalian Sec16 homologues have nonredundant functions in endoplasmic reticulum (ER) export and transitional ER organization,Mol. Biol. Cell 18:839-849(2007).