

# Identification of Protein Interaction Networks of Gankyrin in Cancer

*By*

**Padma P. Nanaware**

**[LIFE09200804009]**

**Tata Memorial Centre  
Mumbai**

*A thesis submitted to the  
Board of Studies in Life Sciences  
In partial fulfillment of requirements  
For the Degree of*

**DOCTOR OF PHILOSOPHY**

*Of*

**HOMI BHABHA NATIONAL INSTITUTE**

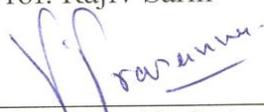
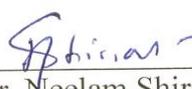
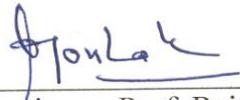


**October, 2014**

# HOMI BHABHA NATIONAL INSTITUTE

## Recommendations of the Viva Voce Board

As members of the Viva Voce Board, we certify that we have read the dissertation prepared by Ms. Padma P. Nanaware entitled "Identification of Protein Interaction Networks of Gankyrin in Cancer" and recommend that it may be accepted as fulfilling the thesis requirements for the award of Degree of Doctor of Philosophy.

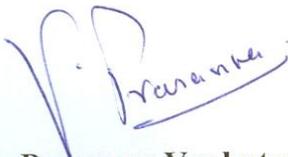
	21-10-2014
Chairperson - Prof. Rajiv Sarin	Date:
	21-10-2014
Convener - Dr. Prasanna Venkatraman	Date:
	21-10-2014
Member 1- Dr. Neelam Shirsat	Date:
	31.10.2014
Member 2- Dr. Tejpal Gupta	Date:
	21.10.2014
Invitee - Dr. Anjali Shiras, NCCS	Date:
	21/10/14
External Examiner - Prof. Rajesh Gokhale, IGIB	Date:

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: 21/10/2014

Place: Kharghar

  
(Dr. Prasanna Venkatraman)  
Guide

## STATEMENT BY AUTHOR

This dissertation has been submitted in partial fulfillment of requirements for an advanced degree at Homi Bhabha 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 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.



Date: 21<sup>st</sup> October 2014

Place: ACTREC, Kharghar

Padma P. Nanaware

## 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.



Date: 21<sup>st</sup> October 2014

Place: ACTREC, Kharghar

Padma P. Nanaware

## List of Publications arising from the thesis

### Journal

Discovery of multiple interacting partners of gankyrin, a proteasomal chaperone and an oncoprotein -Evidence for a common hot spot site at the interface and its functional relevance.

**Nanaware PP**, Ramteke MP, Somavarapu AK, Venkatraman P.

Proteins. 2014 Jan 15. doi: 10.1002/prot.24494

Identification of a novel ATPase activity in 14-3-3 proteins - Evidence from enzyme kinetics, structure guided modeling and mutagenesis studies.

Ramteke MP, Shelke P, Ramamoorthy V, Somavarapu AK, Gautam AK, **Nanaware PP**, Karanam S, Mukhopadhyay S, Venkatraman P.

FEBS Lett. 2013 Nov 20. doi: 10.1016/j.febslet.2013.11.008



Date: 21<sup>st</sup> October 2014

Place: ACTREC, Kharghar

Padma P. Nanaware

***Dedicated To My Loving Parents***

## ***Acknowledgements***

*I would like to take this opportunity to thank everybody who has helped me in all terms to make this journey possible.*

*First and foremost I would express my deepest gratitude to my guide, Dr. Prasanna Venkatraman without whom this thesis would have just remained a dream. Her persistent guidance, intelligent suggestions, critical evaluation helped to maintain stimulation and momentum of project. In hard times she stood beside me and supported me like a strong pillar and at times she played a role of good friend and philosopher. Mam, thanks for being so patient with me and tolerating my ample number of mistakes and correcting me again and again to make me a better researcher and a better person.*

*I sincerely thank Dr. S. Chiplunkar (Director, ACTREC), Prof. Rajiv Sarin (former Director, ACTREC), Dr. Surekha Zingde (former Dy. Director, ACTREC) for providing the entire research infrastructure.*

*I owe my sincere thanks to my doctoral committee members Prof. Rajiv Sarin, Dr. Neelam Shirsat, Dr. Anjali Shiras (NCCS), Dr. Tejpal Gupta for their support, timely and valuable suggestions which has helped me to mature into a better researcher.*

*I would like to convey my sincere thanks to my seniors Amit and Manoj to teach me and act my shield when I was scared to face the tough situations. I would like to thank both of you for being a great teacher and guiding me throughout my tenure. I am really grateful to you both for being so calm, highly patient with me and helping me whenever I asked for it.*

*I would like to thank my labmate and batchmate Nikhil for being so ever ready to help me and support me all time. I want to express my warmest thanks to Ludbe sir for providing the unconditional help whenever needed. I would like to thank all Prasanna*

*lab members Indrajit, Dr. Vinita, Dr. Priya, Kamlesh, Mahalaxmi, Burhan, Saim for providing such an excellent environment to work and making my learning experience a memorable one. I have really learned a lot from you all and thank a lot for it.*

*I would like to thank my trainees Namit, Manjit, Tarun, Sheena and all others who helped me and became my 25<sup>th</sup> hour when actually there were time constraints.*

*Thanks to all of you for helping me.*

*I am thankful to all CIR staff especially to Mr. Dandekar, Microscopy facility, photography, library, administration, dispatch, steno pool and account section for their constant help and support.*

*I would also like to thank my batch mates Manohar, Surya, Monica, Peeyush, Lalit, Zahid, Lumbini, Nitu, Vikrant, Anjana, Shibi, Rashmi and other ACTREC students (Juniors and Seniors) for healthy discussion and critical comments throughout my tenure.*

*Finally, I would like to dedicate my warmest thanks to my parents and brothers for their unconditional love, support and patience with me throughout my life. Without you both nothing could have been achieved.*

Padma P. Nanaware

<b>CONTENTS</b>	<b>Page No.</b>
<b>Synopsis</b>	1
<b>List of Figures</b>	16
<b>List of Tables</b>	17
<b>List of Abbreviations</b>	18
<b>Chapter 1: Introduction</b>	20
<b>Chapter 2: Review Of Literature</b>	26
2.1 Understanding biology	27
2.2 Hub Proteins	29
2.3 Structural Characteristics of Hub proteins	30
2.4 Oncogenes/Tumor suppressors as hub proteins	32
2.5 Gankyrin- An Oncoprotein	34
2.6 Gankyrin- non-ATPase subunit of 26S Proteasome	37
2.7 Gankyrin- Dual negative regulator of tumor suppressor p53 and Retinoblastoma (Rb)	37
2.8 Gankyrin – Negative regulator of NF- $\kappa$ B activity	38
2.9 Ras induced activation of Gankyrin	39
2.10 Gankyrin interacts with S6 ATPase of the 26S Proteasome	39
2.11 Gankyrin interacts with CDK4- Counteraction of gankyrin to regulate INK4-CDK4-Rb pathway	40
2.12 Gankyrin- Hub Protein	40
2.13 Identification of Protein-Protein Interactions	42
2.14 Hot spot sites	44
<b>Objectives:</b>	46
<b>Chapter 3: Material And Methods</b>	
3.1 Buffers and Reagents:	48
3.1.1 Luria-Bertani (LB) Medium (for 1 L)	48
3.1.2 LB-Ampicillin Agar Plates (for 1 L)	48
3.1.3 Tris-EDTA (TE) Buffer (for 50 ml)	48
3.1.4 Ampicillin Stock	49
3.1.5 50X TAE Buffer (for 1 L)	49
3.1.6 6X Gel Loading Buffer for DNA (for 100 ml)	49
3.1.7 Ethidium Bromide (EtBr)	49
3.1.8 Buffers for Ni NTA Column Purification and Gel Filtration:	49
3.1.8.1 Ni-NTA Lysis Buffer (for 1 L)	49
3.1.8.2 Ni-NTA Binding/Washing Buffer (for 1 L)	50
3.1.8.3 Ni-NTA Elution Buffer (for 1 L)	50
3.1.8.4 Running Buffer for Gel Filtration (for 1 L)	50

3.1.9 NP-40 Lysis buffer	51
3.1.10. 1X Transfer buffer	51
3.1.11 10 mM Sodium bicarbonate buffer pH 9.3	51
3.1.12 10 mM Phosphate Buffer pH 7.5	51
3.1.13 1X Phosphate buffer saline pH 7.5	51
3.1.14 Tissue Culture Media and Reagents	52
3.1.14.1 10X Phosphate buffered saline (PBS)	52
3.1.14.2 10X Trypsin	52
3.1.14.3 2X BES Buffered Saline	53
3.1.14.4 0.5 M CaCl <sub>2</sub>	53
3.1.15 Other Reagents	53
<b>Experimental Protocol:</b>	<b>54</b>
3.2.1 Primer Reconstitution	54
3.2.2 Determination of Nucleic Acid Concentration	54
3.2.3 PCR Amplification	54
3.2.4 Site Directed Mutagenesis	55
3.2.5 Restriction Digestion Reaction	56
3.2.6 Agarose Gel Electrophoresis of DNA	56
3.2.7 Recovery of DNA from Low Melting Agarose Gel	56
3.2.8 Ligation/Cloning	57
3.2.9 Transformation	57
3.2.10 Plasmid Mini Preparation	57
3.2.11 Plasmid Construction	58
3.2.12 Confirmation of Positive Clones or Mutation	60
3.2.13 Protein Expression	60
3.2.14 Ni-NTA Agarose Affinity Chromatography	61
3.2.15 Glutathione-S-Transferase purification affinity Chromatography	61
3.2.16 Gel Filtration Chromatography	61
3.2.17 Protein Estimation using Bradford Assay	62
3.2.18 SDS PAGE	62
3.2.19 Preparation of Glycerol Stocks	62
3.2.20 GST Pull down assay	63
3.2.21 Western blotting	63
3.2.22 Routine Maintenance of Cell Lines	64
3.2.23 Freezing and Revival of Cell cultures	65
3.2.24 Transfection	65
3.2.24.1 Transfection of plasmid DNA in HEK 293 cells	65
3.2.24.2 Transfection of siRNA in HEK 293 cells	66
3.2.24.3 Transfection of plasmid or siRNA in MDA-MB-231 and MDA-MB-435 cells	67
3.2.25 Generation of gankyrin over expressing stable HEK 293 cells	67
3.2.26 Proliferation assay-MTT assay	68
3.2.27 Etoposide induced apoptosis assay- MTT assay	68
3.2.28 NF- $\kappa$ B activation- Luciferase assay	68
3.2.29 Soft agar assay	68
3.2.30 Affinity pull down assay	69
3.2.31 Immunoprecipitation assay	69

3.2.32 RNA extraction and cDNA synthesis	69
3.2.33 Real Time PCR	71
3.2.34 Peptide inhibition assay	73
3.2.35 Isothermal calorimetry studies to quantitate the Interaction between CLIC1 and gankyrin	73
3.2.36 ELISA	74
3.2.37 Migration using wound healing assay	75
3.3.38 Invasion using Boyden chamber assay	76

#### **Chapter 4: Generation and Characterization of gankyrin overexpressing stable HEK 293 cells**

4.1 Introduction	79
4.2 Results and Discussion	82
4.3 Summary	86

#### **Chapter 5: Bioinformatic Identification Of Novel Interacting Partners Of Gankyrin**

5.1 Introduction	88
5.2 Results and Discussion	91
5.3 Summary	97

#### **Chapter 6: Choice Of Proteins And Experimental Validation Of Putative Interacting Partners Of Gankyrin**

6.1 Introduction	99
6.2 Results and Discussion	99
6.3 Summary	111

#### **Chapter 7: Evidence For The EEVD As The Hot Spot Site: Direct Interaction Between Gankyrin And Its Client Proteins**

7.1 Introduction	113
7.2 Results and Discussion	114
7.3 Summary	121

#### **Chapter 8: Interaction between Gankyrin and CLIC1 is Important for cell Migration and Invasion in MDA-MB-231 and HEK 293 cells**

8.1 Introduction	124
8.2 Results and Discussion	125
8.3 Summary	135

#### **Chapter 9: Expansion of Network to EEXD Containing Proteins**

9.1 Introduction	137
9.2 Results and Discussion	142
9.3 Summary	148

<b>Chapter 10: Conclusions and Significance of Study</b>	149
<b>References</b>	153
<b>Appendix</b>	174
<b>Reprints Of Published Articles</b>	177

## Homi Bhabha National Institute



### Ph. D. PROGRAMME

- 1. Name of the Student:** Ms. Padma P. Nanaware
- 2. Name of the Constituent Institution:** Tata Memorial Centre (TMC), Advanced Centre for Treatment, Research and Education in Cancer (ACTREC)
- 3. Enrolment No. :** LIFE09200804009
- 4. Title of the Thesis:** Identification of Protein Interaction Networks of Gankyrin in Cancer.
- 5. Board of Studies:** Life Sciences (LS)

## *Synopsis*

Protein-protein interactions (PPI) form an important functional network. They maintain spatio-temporal regulation inside the cellular system and their deregulation can manifest in the form of diseases like cancer. A network of PPI consists of nodes connected by links (edges or interactions) (Koutsogiannouli et al., 2013a). Oncogenes and tumor suppressors are critical nodes and act as hubs in clusters/modules that have the propensity of making multiple physical or functional connections with other nodes capable of controlling tumor cell behavior. Hub proteins are found to be co-expressed with their interacting partners in a tissue restricted manner. Because of their crucial role in maintaining the structural and therefore the functional network, hub proteins are recognized as druggable targets.

Our aim in this thesis was to characterize protein- protein interaction network mediated by gankyrin which we believe is a key hub protein that connects critical nodes in time and space. This interactome is likely to have common and unique partners under normal physiological conditions vis a vis diseased state such as in cancer. By identifying gankyrin specific functional network specific to cancer cells we aim to identify novel therapeutic targets.

Gankyrin (also known as PSMD10, p28) is an oncoprotein and a non-ATPase subunit of the proteasome. Gankyrin is a chaperone that transiently associates with the proteasome and assists in the assembly (Dawson et al., 2002). It is found to be consistently overexpressed in human liver (Higashitsuji et al., 2000b), pancreatic (Meng et al., 2010b), esophageal squamous cell carcinoma (ESCC) (Ortiz et al., 2008), colorectal cancer (Tang et al., 2010b), breast cancer (Zhen et al., 2012a) and glioma (Yang et al., 2012). This protein contains nuclear localization signal at the N terminus and acts a nuclear-cytoplasmic shuttling protein (Chen et al., 2007c).

Gankyrin carries seven ankyrin repeats which are well characterized as protein interaction domains. In accordance gankyrin is found to interact with MDM2 resulting in ubiquitination of p53 leading to its degradation (Higashitsuji et al., 2005). It also interacts with Rb and subjects it for degradation, resulting in the loss of cell cycle check point control leading to uncontrolled cellular proliferation (Higashitsuji et al., 2000b). Thus gankyrin functions as a dual negative regulator of two important tumor suppressor proteins p53 and Rb. Gankyrin interacts with RelA /NF- $\kappa$ B retaining the latter in the cytoplasm which inhibits the transcriptional activity (Chen et al., 2007c). It interacts with C-terminal half of MAGE-A4 and suppresses its tumorigenic activity (Nagao et al., 2003). Gankyrin is known to be a key regulator of Ras-mediated activation of Akt through the inhibition of the downstream RhoA/ROCK pathway (Man et al., 2010a). In pancreatic cancer downregulation of gankyrin induces cell cycle arrest (G1/S transition) by downregulating cyclin D1, cyclin E, cyclin A, CDK4, CDK2, PCNA and p-Rb expression (Meng et al., 2010b). Knock down of hepatoma upregulated protein (HURP) is shown to inhibit the proliferation of hepatocellular carcinoma cells via downregulation of gankyrin and accumulation of p53 (Kuo et al., 2012). Gankyrin is also known to establish a positive feedback loop in  $\beta$  –catenin signaling (Dong et al., 2011). Gankyrin in breast cancer tissues promotes migration through Rac1 activity (Zhen et al., 2012a) whereas in colorectal carcinoma it promotes metastasis via IL8 activation (Bai et al., 2013a).

These relatively new observations about gankyrin raise fundamental questions about its physiological role. We speculate that this protein may have multiple functions that are dependent and independent of its association with 26S Proteasome and role in protein-degradation. We believe that as an oncoprotein, it may directly or indirectly modulate many more pathways to bring about cellular transformation.

Among the pathways influenced by gankyrin as described above, direct interactions has been observed between gankyrin and **Rb** (Higashitsuji et al., 2000b), **MDM2** (Higashitsuji et al., 2005), **CDK4** (Dawson et al., 2002), **MAGE-A4** (Nagao et al., 2003), **p65** (RelA) (Chen et al., 2007c) and **FIH-1** (Wilkins et al., 2012). However no systematic efforts have been employed to map the interactome of gankyrin. Therefore, systematically identifying more interacting partners of gankyrin is a relevant and important step in the recognition of novel functions both under normal physiological conditions and in malignancy. Furthermore it would be important to identify cancer specific networks mediated by gankyrin that can then be manipulated either to reverse the phenotype or induce cell death.

## **Rationale**

Protein-protein interaction surfaces are large, but only few residues contribute to the maximum binding energy. Such residues are called as the hot spot sites. They are conserved across interfaces and can interact via same residues with other interacting proteins. However, the utility of this concept to identify novel interacting proteins is unexplored. Here, in this thesis we describe an approach to identify functionally relevant PPI and potentially vulnerable targets using the principles of structural bioinformatics.

## **Strategy**

Our strategy lies in the recognition of a hot spot site at the interface in a known complex of gankyrin-S6 ATPase. Identification of putative interacting partners of proteins that share the hot spot site (Short Linear Sequence Motif) in an accessible region of the protein, validating the interactions and demonstrating the functional

relevance of the hot spot site using the combination of biochemical, molecular and cell biological techniques.

With this purpose we define the following objectives.

**Objectives:**

1. Identifying the novel interacting partners of gankyrin and validating them in epithelial cancer cell lines.
2. Validating functional relevance of at least two of the interactions.

**Results and Discussions**

**Model system**

In order to map the differential protein interactome in cancer cells and provide a causal link with gankyrin we generated a model system in which gankyrin levels can be regulated. Stable clones of gankyrin were created and the oncogenic phenotypes like anchorage independent growth and resistance to cell death in response to apoptotic stimuli were tested.

**A hot spot site at the gankyrin interface and Prediction of Interacting partners**

We used the crystal structure of mouse (m) gankyrin complexed with human (h) S6 ATPase (Nakamura et al., 2007b) to identify a potential hot spot site at the interface. The human and mouse gankyrin share 98% sequence identity and the crystal structure of unliganded h-gankyrin is completely super imposable on the m-gankyrin-S6 ATPase structure. A linear stretch of four amino acids EEVD (aa 356- 359) makes dominant contribution to the interaction. Complex formation as seen by co-expression was abolished by any one of the double mutations- E356/E357A or D359/362A in S6 ATPase. To confirm the importance of these residues in interaction, we mutated these residues in S6 ATPase and co-expressed the mutant with human gankyrin. Interaction

was abolished and we identified E, E and D residues within the Short Linear Sequence Motif EEVD as the hot spot site. Based on this finding we hypothesized that proteins in the human proteome with EEVD in the accessible region have the potential of being putative interacting partners of gankyrin. A total of 264 unique proteins with EEVD in their primary sequence were found and for 34 of them structural information was available using which solvent accessible surface area (SASA) values were calculated. Among the 34 proteins we chose to test eight proteins NCK2, G-rich RNA sequence binding factor 1 (GRSF1), Chloride intracellular channel protein 1 (CLIC1), Eukaryotic initiation factor 4A-III (EIF4A3), dimethylarginine dimethylaminohydrolase 1 (DDAH1) and mitogen-activated protein kinase 1 (MAP2K1), Heat shock protein 70 (Hsp70), Heat shock protein 90 (Hsp90) as putative interacting partners of gankyrin. EEVD in all of these sequences were either well exposed or disordered. All these proteins like gankyrin are expressed in the cytoplasm or both in cytoplasm and nucleus. Six out of eight proteins carried hall mark cancer properties.

### **Validation**

Six proteins Hsp70, Hsp90, CLIC1, GRSF1, DDAH1 and MAP2K1 were present in the immune complex isolated from gankyrin overexpressing HEK 293 cells implicating the involvement of predicted residues at the interface. Some interactions may happen constitutively inside the cell whereas some might happen when gankyrin is overexpressed or in malignant conditions where PPI is deregulated. To test the generality of the interactions, we tested all the interactions in HEK 293 Wt cells. Three proteins (Hsp70, Hsp90 and GRSF1) were found in the immune complex while DDAH1 was barely detectable. MAP2K1 did not interact with gankyrin. These

interactions seen only in gankyrin overexpressing cells were also observed in cancer cells like MDA-MB-435, a breast tumor derived cell line and gankyrin is seen to be overexpressed in breast cancer tissues. MAP2K1, DDAH1 and CLIC1 were found to interact with gankyrin at its endogenous levels in MDA-MB-435 cells. These observations imply several possibilities a) these interactions are weak and therefore occur only when gankyrin is overexpressed or b) represent a deregulated phenotype where new networks are created on demand.

### **EEVD “the common hot spot site at the interface”**

Recombinant proteins Hsp70 Wt and CLIC1 Wt were found to interact directly with gankyrin and this interaction was confirmed to be mediated by EEVD by site directed mutagenesis. These findings were further confirmed in mammalian cells by expressing HA-CLIC1 or flag-Hsp70, and their corresponding AAVA mutants in MDA-MB-435 cell line. In the case of NCK2 and GRSF1 the interactions through EEVD was confirmed in MDA-MB-435 cell line. Short naked peptide EEVD inhibits gankyrin-CLIC1 interaction with an IC 50 value of 50  $\mu$ M. All these results confirm that ‘EEVD’ indeed acts as a hot spot site both in vitro and within the physiological milieu.

### **Expansion of the network to EEXD containing proteins**

Based on the crystal structure of gankyrin-S6 ATPase, Val within the SLIM, EEVD, seems to have the least influence on interaction. Moreover, when val in EEVD of CLIC1 was mutated to Glu the interaction by and large was unaffected. Thus, proteins carrying other EEXD residues may act as interacting partners of gankyrin. To test this possibility, we selected calreticulin, an endoplasmic reticulum protein known to be overexpressed in nuclear matrix of hepatocellular carcinoma tissues as compared

to normal tissues (Yoon et al., 2000). Calreticulin carries motif EEMD in the accessible region of the protein. Calreticulin Wt and the P domain which carries EEMD interact directly with gankyrin whereas their corresponding AAMA mutants did not show any interaction. This finding was replicated in HEK 293 stable clones overexpressing gankyrin. Interestingly this interaction was not observed in HEK 293 Wt cells but again like DDAH1, MAP2K1, and CLIC1, calreticulin-gankyrin interaction was observed in MDA-MB-435 cells. It is likely that gankyrin interactome may include atleast some if not all of EEXD containing proteins.

### **Functional relevance of the hot spot site – a case study with CLIC1.**

We chose to study the functional importance of gankyrin-CLIC1 interaction because of the following reasons 1) Endogenous levels of CLIC1 was barely detectable in HEK 293 cells, but upon gankyrin overexpression, CLIC1 protein levels were found to be upregulated. 2) Overexpression of CLIC1 is used as a prognostic marker for colorectal cancer (Petrova et al., 2008a) and gastric cancer (Chen et al., 2007a). Hence, we tested the affect the gankyrin-CLIC1 interaction in MDA-MB-231 cells and HEK 293 cells. We independently transfected a smartpool of specific siRNA against gankyrin or CLIC1 in MDA-MB-231 cells. Down regulation of either proteins showed marked reduction in the motility (wound healing) and invasion properties (measured using boyden chamber) of MDA-MB-231 cells. Down regulation either of the two mRNAs reduced the migratory potential of these cells.

In order to evaluate the effect of the gankyrin-CLIC1 interaction on the motility and invasion, CLIC1 Wt and CLIC1\_AAVA were overexpressed in MDA-MB-231 cells and HEK 293 cells. Overexpression of CLIC1 Wt resulted in an increase in the percentage of wound closed and enhanced the invasive potential of both the cell lines

when compared to their respective vector controls. Conversely, the mutant CLIC1\_AAVA cells behaved like those of the vector control cells. These observations were further confirmed by silencing endogenous CLIC1 and overexpressing CLIC1 Wt and CLIC1\_AAVA in MDA-MB-231 cells. The cells over expressing CLIC1 Wt were able to rescue the migratory potential of MDA-MB 231 whereas cells overexpressing CLIC1\_AAVA were unable to do so.

### **Summary**

In our study, we provide evidence that short linear sequence motif at protein-protein interfaces can be potentially used to identify novel functionally relevant protein complexes formed by key hub proteins. By proposing a common hot spot site at the interface of gankyrin interacting partners, we identify Hsp70, Hsp90, GRSF1, CLIC1, MAP2K1, NCK2, DDAH1 as novel interacting partners. These proteins share EEVD, a four residue SLiM, at the interface and by mutagenesis we confirm that that EEXD is a conserved hot spot site. In addition, we demonstrate the functional relevance of these interactions using CLIC1 as an example, where EEVD mediated interaction with gankyrin increases the migratory potential of MDA-MB-231 cells. Our studies also inform that variants of EEVD such as EEXD (X is any residue) can constitute a hot spot site which may help in the identification of other novel interacting partners.

## **Future Prospective**

The concept of short linear motifs at the interface is of great importance for the mapping of interactome of key regulatory proteins and hubs. The results presented here suggest that targeting PPIs at the hot spot site is a real possibility and would render these otherwise shallow surfaces potentially druggable. The study presented here forms the platform for future investigation on the ultra structural details of the interaction, biophysical characteristics and the generation of polypharmacological drugs by rational design or small molecule screening.

## **References**

- Bai, Z.-f., Tai, Y., Li, W., Zhen, C., Gu, W., Jian, Z., Wang, Q., Lin, J.E., Zhao, Q., and Gong, W. (2013). Gankyrin activates IL-8 to promote hepatic metastasis of colorectal cancer. *Cancer research*.
- Chen, C.D., Wang, C.S., Huang, Y.H., Chien, K.Y., Liang, Y., Chen, W.J., and Lin, K.H. (2007a). Overexpression of CLIC1 in human gastric carcinoma and its clinicopathological significance. *Proteomics* 7, 155-167.
- Chen, Y., Li, H.H., Fu, J., Wang, X.F., Ren, Y.B., Dong, L.W., Tang, S.H., Liu, S.Q., Wu, M.C., and Wang, H.Y. (2007b). Oncoprotein p28 GANK binds to RelA and retains NF-kappaB in the cytoplasm through nuclear export. *Cell research* 17, 1020-1029.
- Dawson, S., Apcher, S., Mee, M., Higashitsuji, H., Baker, R., Uhle, S., Dubiel, W., Fujita, J., and Mayer, R.J. (2002). Gankyrin is an ankyrin-repeat oncoprotein that interacts with CDK4 kinase and the S6 ATPase of the 26 S proteasome. *Journal of Biological Chemistry* 277, 10893-10902.

Dong, L.-w., Yang, G.-z., Pan, Y.-f., Chen, Y., Tan, Y.-x., Dai, R.-y., Ren, Y.-b., Fu, J., and Wang, H.-y. (2011). The oncoprotein p28GANK establishes a positive feedback loop in  $\beta$ -catenin signaling. *Cell research* 21, 1248-1261.

Higashitsuji, H., Higashitsuji, H., Itoh, K., Sakurai, T., Nagao, T., Sumitomo, H., Masuda, T., Dawson, S., Shimada, Y., and Mayer, R.J. (2005). The oncoprotein gankyrin binds to MDM2/HDM2, enhancing ubiquitylation and degradation of p53. *Cancer cell* 8, 75-87.

Higashitsuji, H., Itoh, K., Nagao, T., Dawson, S., Nonoguchi, K., Kido, T., Mayer, R.J., Arii, S., and Fujita, J. (2000). Reduced stability of retinoblastoma protein by gankyrin, an oncogenic ankyrin-repeat protein overexpressed in hepatomas. *Nature medicine* 6, 96-99.

Koutsogiannouli, E., Papavassiliou, A.G., and Papanikolaou, N.A. (2013). Complexity in cancer biology: is systems biology the answer? *Cancer medicine* 2, 164-177.

Kuo, T.-C., Chang, P.-Y., Huang, S.-F., Chou, C.-K., and Chao, C.C.-K. (2012). Knockdown of HURP inhibits the proliferation of hepatic carcinoma cells via downregulation of gankyrin and accumulation of p53. *Biochemical Pharmacology* 83, 758-768.

Man, J.-H., Liang, B., Gu, Y.-X., Zhou, T., Li, A.-L., Li, T., Jin, B.-F., Bai, B., Zhang, H.-Y., and Zhang, W.-N. (2010). Gankyrin plays an essential role in Ras-induced tumorigenesis through regulation of the RhoA/ROCK pathway in mammalian cells. *The Journal of clinical investigation* 120, 2829.

Meng, Y., He, L., Guo, X., Tang, S., Zhao, X., Du, R., Jin, J., Bi, Q., Li, H., Nie, Y., *et al.* (2010). Gankyrin promotes the proliferation of human pancreatic cancer. *Cancer letters* 297, 9-17.

Nagao, T., Higashitsuji, H., Nonoguchi, K., Sakurai, T., Dawson, S., Mayer, R.J., Itoh, K., and Fujita, J. (2003). MAGE-A4 interacts with the liver oncoprotein gankyrin and suppresses its tumorigenic activity. *Journal of Biological Chemistry* 278, 10668-10674.

Nakamura, Y., Nakano, K., Umehara, T., Kimura, M., Hayashizaki, Y., Tanaka, A., Horikoshi, M., Padmanabhan, B., and Yokoyama, S. (2007). Structure of the oncoprotein gankyrin in complex with S6 ATPase of the 26S proteasome. *Structure* 15, 179-189.

Ortiz, C.M., Ito, T., Tanaka, E., Tsunoda, S., Nagayama, S., Sakai, Y., Higashitsuji, H., Fujita, J., and Shimada, Y. (2008). Gankyrin oncoprotein overexpression as a critical factor for tumor growth in human esophageal squamous cell carcinoma and its clinical significance. *International journal of cancer* 122, 325-332.

Petrova, D.T., Asif, A.R., Armstrong, V.W., Dimova, I., Toshev, S., Yaramov, N., Oellerich, M., and Toncheva, D. (2008). Expression of chloride intracellular channel protein 1 (CLIC1) and tumor protein D52 (TPD52) as potential biomarkers for colorectal cancer. *Clinical biochemistry* 41, 1224-1236.

Tang, S., Yang, G., Meng, Y., Du, R., Li, X., Fan, R., Zhao, L., Bi, Q., Jin, J., Gao, L., *et al.* (2010). Overexpression of a novel gene gankyrin correlates with the malignant phenotype of colorectal cancer. *Cancer biology & therapy* 9, 88-95.

Taylor, I.W., Linding, R., Warde-Farley, D., Liu, Y., Pesquita, C., Faria, D., Bull, S., Pawson, T., Morris, Q., and Wrana, J.L. (2009). Dynamic modularity in protein interaction networks predicts breast cancer outcome. *Nature biotechnology* 27, 199-204.

Wilkins, S.E., Karttunen, S., Hampton-Smith, R.J., Murchland, I., Chapman-Smith, A., and Peet, D.J. (2012). Factor Inhibiting HIF (FIH) Recognizes Distinct Molecular

Features within Hypoxia-inducible Factor- $\alpha$  (HIF- $\alpha$ ) versus Ankyrin Repeat Substrates. *Journal of Biological Chemistry* 287, 8769-8781.

Yang, Y., Zhang, C., Li, L., Gao, Y., Luo, X., Zhang, Y., Liu, W., and Fei, Z. (2012). Up-regulated oncoprotein P28GANK correlates with proliferation and poor prognosis of human glioma. *World journal of surgical oncology* 10, 1-7.

Yoon, G.-S., Lee, H., Jung, Y., Yu, E., Moon, H.-B., Song, K., and Lee, I. (2000). Nuclear matrix of calreticulin in hepatocellular carcinoma. *Cancer research* 60, 1117-1120.

Zhen, C., Chen, L., Zhao, Q., Liang, B., Gu, Y., Bai, Z., Wang, K., Xu, X., Han, Q., and Fang, D. (2012). Gankyrin promotes breast cancer cell metastasis by regulating Rac1 activity. *Oncogene*.

## Accepted Publications

Discovery of multiple interacting partners of gankyrin, a proteasomal chaperone and an oncoprotein -Evidence for a common hot spot site at the interface and its functional relevance.

**Nanaware PP**, Ramteke MP, Somavarapu AK, Venkatraman P.

Proteins. 2014 Jan 15. doi: 10.1002/prot.24494

Identification of a novel ATPase activity in 14-3-3 proteins - Evidence from enzyme kinetics, structure guided modeling and mutagenesis studies.

Ramteke MP, Shelke P, Ramamoorthy V, Somavarapu AK, Gautam AK, **Nanaware PP**, Karanam S, Mukhopadhyay S, Venkatraman P.

FEBS Lett. 2013 Nov 20. doi: 10.1016/j.febslet.2013.11.008

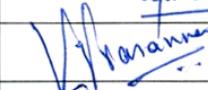
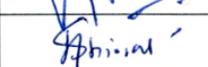
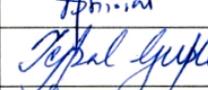
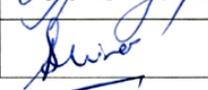


Signature of Student

Ms. Padma P. Nanaware

Date: 3<sup>rd</sup> February 2014

**Doctoral Committee:**

S. No.	Name	Designation	Signature	Date
1.	Prof. Rajiv Sarin	Chairman		07/02/14
2.	Dr. Prasanna Venkatraman	Convener		04/02/14
3.	Dr. Neelam Shirsat	Member		07/02/14
4.	Dr. Tejpal Gupta	Member		07/02/14
5.	Dr. Anjali Shiras (NCCS)	Member		06/02/14

Forwarded through:



Dr. S.V. Chiplunkar  
(Director, ACTREC,  
Chairperson,  
Academics and Training Programme, ACTREC)

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,  
TMC)



## LIST OF FIGURES

Sr. No.	Title	Page No.
1.	Fig. 2.1 Technology advances to study biology of life	28
2.	Fig. 2.2 Characteristics affecting the interaction promiscuity in hub proteins	32
3.	Fig. 2.3 Structure of Gankyrin	41
4.	Fig. 4.1 Generation of gankyrin overexpressing stable HEK 293 clones	83
5.	Fig. 4.2 Characterization of stable HEK 293 cells overexpressing gankyrin	85
6.	Fig. 5.1 Crystal structure of human gankyrin (1UOH) was aligned with mouse gankyrin- human S6 ATPase (2DVW) complex using PyMol	91
7.	Fig.5.2 Ligplot showing 2-Dimensional view of the interaction between gankyrin and S6 ATPase	93
8.	Fig. 5.3 Gankyrin interacts with S6 ATPase through E, E and D residues	94
9.	Fig. 6.1 Pie Chart showing functional classification of putative partners of gankyrin	100
10.	Fig. 6.2 Oncogenic or tumor suppressive influence of molecular co-chaperones	102
11.	Fig. 6.3 Key feature of the network with gankyrin as the hub and the putative interacting partners harboring EEVD hotspot site as potential nodes	104
12.	Fig. 6.4 Within the cellular milieu, gankyrin interacts with CLIC1, Hsp70, Hsp90, GRSF1, DDAH1, MAP2K1, and does not interact with EIF4A3	105
13.	Fig. 6.5 Transcripts levels of gankyrin, CLIC1, Hsp70 in gankyrin overexpressing HEK 293 clones	106
14.	Fig. 6.6 Within the cellular milieu, gankyrin interacts with Hsp70, Hsp90, GRSF1	107
15.	Fig. 6.7 Within the cellular milieu of MDA-MB-435 gankyrin interacts with CLIC1, DDAH and, MAP2K1	109
16.	Fig. 6.8 Confirming the specificity of interaction of gankyrin with its EEVD containing interacting partners	110
17.	Fig. 7.1 Purification of His-Gankyrin, GST-Hsp70, GST-Hsp70_AAVA, GST-CLIC1, GST-CLIC1_AAVA	114
18.	Fig. 7.2 Gankyrin binds to Hsp70, CLIC1 through E, E and D residues	115
19.	Fig 7.3 Gankyrin binds to Hsp70, CLIC1, NCK2 and GRSF1 through E, E and D residues	116
20.	Fig.7.4 Interface peptide EEVD inhibits interaction of full length protein with gankyrin	118
21.	Fig. 7.5 Isothermal calorimetry studies to study His-gankyrin interaction with CLIC1 Wt or CLIC1_AAVA	120
22.	Fig. 7.6 ELISA to study the interaction of gankyrin_Wt and its mutants, gankyrin_K116A or gankyrin_R41A with CLIC1	121
23.	Fig. 8.1 Silencing of gankyrin and CLIC1 affects migration in	

	MDA-MB-231 cells	126
24.	Fig. 8.2 Silencing of gankyrin and CLIC1 affects invasion in MDA-MB-231 cells	127
25.	Fig. 8.3 Overexpression of CLIC1 and CLIC1_AAVA affects migration of MDA-MB-231	129
26.	Fig. 8.4 Overexpression of CLIC1 and CLIC1_AAVA affects invasion of MDA-MB-231 cells	130
27.	Fig. 8.5 Interaction of gankyrin with CLIC1 enhances migration	131
28.	Fig. 8.6 Interaction of gankyrin with CLIC1 enhances migration - Reverse experiment	133
29.	Fig. 8.7 Effect of gankyrin and CLIC1 on cellular proliferation	134
30.	Fig. 9.1 Gankyrin does not require valine in EEVD for interaction with CLIC1	142
31.	Fig. 9.2 Comparison of protein profiles using two dimensional gel electrophoresis	144
32.	Fig. 9.3 Gankyrin interacts with calreticulin within the cellular milieu and <i>in vitro</i>	146
33.	Fig. 9.4 Calreticulin with EEMD (variant of EEVD) in the accessible region interacts with gankyrin only in gankyrin overexpressing stable HEK 293 clones and MDA-MB-435	147

## LIST OF TABLES

Sr. No.	Title	Page No.
1.	Table 2.1 Experimental approaches used for identifying protein-protein interactions	42
2.	Table 3.1 Primers used for Cloning and site directed mutagenesis	59
3.	Table 3.2 Primers used for Real time PCR	72
4.	Table 3.3 The list of siRNA sequences	76
5.	Table 5.1 Solvent accessible surface area values for EEVD containing proteins	96
6.	Table 6.1 Proteins for which interactions have been validated in this study show hall mark of cancer properties	101
7.	Table 7.1 Summary of protein complexes with known interface residues and peptides or small molecular inhibitors designed to inhibit the respective complexes	119
8.	Table 9.1 List of the sequences of known gankyrin interacting partners and those identified in this study	137

## **Abbreviations**

BME	$\beta$ - Mercaptoethanol
BSA	Bovine Serum Albumin
cDNA	complimentary Deoxyribonucleic acid
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
kDa	Kilodalton
M	Molar
mM	Millimolar
MALDI	Matrix assisted laser desorption ionization
MS	Mass spectrometry
MQ	Milli Q (water)
NCBI	National Centre for Biotechnology Information
Ni-NTA	Nickel-nitriloacetic acid
GST	Glutathione S-transferase
PAGE	Polyacrylamide gel electrophoresis
PCR	Polymerase chain reaction
PDB	Protein Data Bank
SDS	Sodium dodecyl sulphate
TEMED	N,N,N',N' Tetramethyl ethylene diamine
TBST	Tris buffered saline with 0.1% Tween-20
WT	Wild-type
CLIC1	Chloride Intracellular channel protein 1
PANTHER	Protein analysis Through Evolutionary Relationships
SASA	Solvent Accessible Surface Area

Hsp70	Heat shock protein 70kDa
Hsp90	Heat shock protein 90kDa
GRSF1	G-rich sequence factor 1
DDAH1	N(G)N(G)-dimethylarginine dimethylaminohydrolase 1
MAP2K1	Dual specificity mitogen-activated protein kinase kinase 1
EIF4A3	Eukaryotic initiation factor 4A-III
PNSAS	Prediction of Natural Substrates from Artificial Substrate of Proteases
SliM	Short linear Motif
WISE	Whole interactome scanning experiment
Dsg2	Desmoglein

# ***CHAPTER 1.***

## ***Introduction***

The phenotype of an organism is an outcome of a complex and yet well integrated intracellular processes like DNA replication, protein synthesis and their trafficking, chromosomal segregation, cell growth, motility etc and the influence of environment. Physical and functional interactions among the various players and the network of communications between them regulate these various processes and any aberration may result in disease. Direct protein-protein interactions (PPI) are central to such integrated functioning as they provide the vital link between divergent networks such as the signaling networks, transcriptional networks etc which modulate cellular behaviour providing elasticity to adapt to various stress and environmental perturbations. Due to the importance of PPI, many approaches have been developed to screen and identify them both as one on one interactions or as a complexes (reviewed under Table 2.1). Due to such efforts, it has been possible to map many if not all interactions which are regulated by physiological demands and are time dependent. Even more challenging has been the ability to extract biological insights from these vast depositaries of data. Analyzing these interactions, building statistically relevant functional network and validating these networks are some of the major challenges that are currently being addressed.

A basic network of PPI consists of biomolecules (nodes) connected by links (edges or interactions) (Ekman et al., 2006; Koutsogiannouli et al., 2013b). Oncogenes and tumor suppressors are critical nodes and act as hubs in clusters/modules that have the propensity of making multiple physical or functional connections with other nodes capable of controlling tumor cell behavior. Oncogenic addiction leads to either stabilization or rewiring of the network resulting in an altered phenotype (Kar et al., 2009). These oncogenic proteins acting as hubs, play a decisive role in maintaining

the structural and therefore the functional network and hence are recognized as therapeutically vulnerable targets (Koutsogiannouli et al., 2013b).

The current thesis focus is on Gankyrin, an oncoprotein overexpressed in many epithelial cancers. There are no mutations or translocations reported till date for gankyrin. Gankyrin was initially identified as a non-ATPase subunit of the proteasome (Dawson et al., 2002). Gankyrin is known to negatively regulate two important tumor suppressors p53 and Rb leading to loss of cell cycle check points and finally the uncontrolled cell proliferation (Higashitsuji et al., 2005; Higashitsuji et al., 2000a). Gankyrin is known to be a key regulator of Ras-mediated activation of Akt through the inhibition of the downstream RhoA/ROCK pathway (Man et al., 2010a). It establishes a positive feedback loop in  $\beta$ -catenin signaling (Dong et al., 2011). In breast cancer tissues, gankyrin promotes migration through Rac1 activity (Zhen et al., 2012a) whereas in colorectal carcinoma it promotes metastasis via IL8 activation (Bai et al., 2013b). Thus, gankyrin is known to affect many crucial signaling pathways and behaves like a key hub protein that connects many nodes in time and space. However it is still unclear as to what drives the oncogenic properties of gankyrin. And although gankyrin seems to be a major hub protein, the PPI network of this protein is uncharacterized. Gankyrin overexpression alone is enough to cause malignant transformation and impart invasive properties. Therefore PPI in the normal and cancer cells are likely to be different. The network is likely to be rewired with new interactions formed in the cancer cells and interaction seen in normal cells strengthened further. Therefore, the aim of this thesis is to characterize the PPI network of gankyrin using a novel structural bioinformatics method as described further below. Structurally, Gankyrin is characterized by seven ankyrin repeats and ankyrin repeats are protein interaction motifs that have the capacity to bind with

multiple proteins (Li et al., 2006; Mosavi et al., 2004). Generally protein-protein interactions involve large surface areas, but bulk of the binding energy is contributed by few key residues at the interaction surface (Bogan and Thorn, 1998; DeLano, 2002). These residues when mutated result in either rapid dissociation of the complex or prevent stable association and are called as the ‘hot spot’ sites. Such hot spot sites are conserved across interfaces (Hu et al., 2000).

In order to map the gankyrin interaction network, we capitalized on available crystal structure of a mouse gankyrin-human S6 ATPase complex (Nakamura et al., 2007b). We recognised a potential hot spot site made of short linear sequence motif ‘EEVD’ at the interface. We hypothesized that this hot spot site may be conserved and gankyrin may interact with other proteins that carry EEVD/EEXD motif. Using a simple bioinformatics tool, 34 proteins with EEVD motif in well exposed region on their surface were identified. We chose to test eight proteins- NCK2, G-rich RNA sequence binding factor 1 (GRSF1), Chloride intracellular channel protein 1 (CLIC1), Eukaryotic initiation factor 4A-III (EIF4A3), dimethylarginine dimethylaminohydrolase 1 (DDAH1) and mitogen-activated protein kinase 1 (MAP2K1), Heat shock protein 70 (Hsp70), Heat shock protein 90 (Hsp90) for further validation. Three of these interactions occur in HEK 293 cells only when gankyrin is overexpressed but occur in breast cancer cells at endogenous levels. Mutagenesis confirms that these interactions involve predicted residues which form the hot spot sites at the shared interface and could be a potential drug target.

In the Short Linear sequence Motif EEVD we found that Val has the least influence in interaction with gankyrin. Hence, we expanded our prediction to EEXD containing proteins. To test this, we selected calreticulin, an endoplasmic reticulum protein (Zamanian et al., 2013) known to be overexpressed in nuclear matrix of hepatocellular

carcinoma tissues as compared to normal tissues (Yoon et al., 2000). Calreticulin carries motif EEMD in the accessible region of the protein. Calreticulin Wt and the P domain which carries EEMD interact directly with gankyrin whereas their corresponding AAMA mutants did not show any interaction. This finding was replicated in HEK 293 stable clones overexpressing gankyrin. Interestingly this interaction was not observed in HEK 293 cells but again like DDAH1, MAP2K1, and CLIC1 was observed in MDA-MB-435 cells. It is likely this interactome is expanded to include at least some if not all of EEXD containing proteins.

In gankyrin overexpressing HEK 293 cells, CLIC1 levels were found to be overexpressed as compared to vector control cells. CLIC1 is a voltage gated channel protein and is known to be overexpressed in lung carcinoma (Wang et al., 2011), colorectal carcinoma (Petrova et al., 2008a), gastric carcinoma (Chen et al., 2007a), glioblastoma (Wang et al., 2012a), hepatocellular carcinoma (Huang et al., 2004) and known to alter invasion and migration properties of the cells (Wang et al., 2012b). Coincidentally, there are reports stating gankyrin is overexpressed in these cancers (Higashitsuji et al., 2000a; Man et al., 2010a; Tang et al., 2010a; Yang et al., 2012). Due to these reasons we tested the functional significance of gankyrin-CLIC1 interaction. The interaction of gankyrin through EEVD of CLIC1, increases the migratory potential of MDA-MB-231 cells.

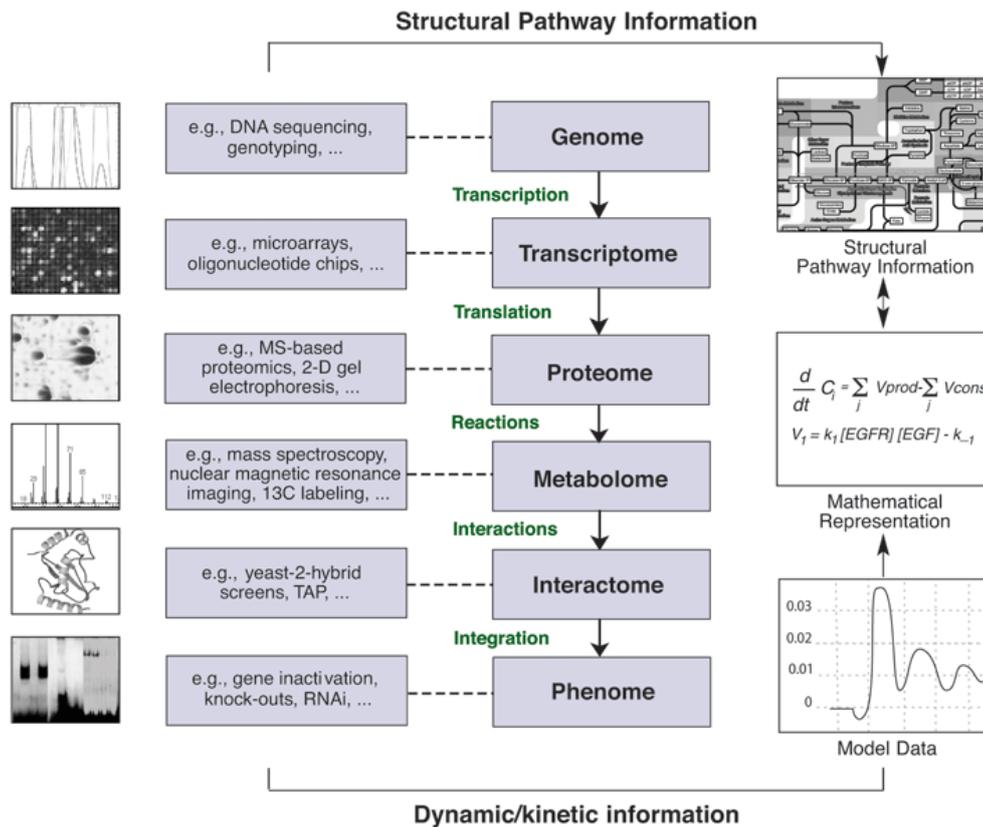
In summary, we provide a structural bioinformatics method to predict novel interacting partners of a protein involved in multiple interactions. In particular, we substantiate that short linear motif at protein-protein interfaces can be potentially used to identify novel functionally relevant protein complexes formed by key hub proteins. Further, we provide evidence for the potential use of protein interfaces as drug targets to inhibit protein-protein interaction mediated by gankyrin. In the case of gankyrin we

predict that due to the functional importance of gankyrin-CLIC1 interaction via EEVD in cell migration, this interface may very relevant for understanding metastasis and for the design and screen of cancer specific small molecular inhibitors of gankyrin.

**CHAPTER 2.**  
***Review Of Literature***

## **2.1 Understanding biology**

Biological research, for many years, relied on several independent approaches to characterize proteins, nucleic acid, lipids as individual molecules and processes such as DNA replication, transcription, protein synthesis, protein degradation and phenomena such as cell division, cell cycle, development and differentiation. Aware that a better understanding of biological complexity will require studying the system as an entity, technological platforms were developed towards a comprehensive analysis of the genetics, biochemistry and cell biology of a cell, tissues and organisms. Phenotypes are the resultant of the interactions of thousands of macromolecules that are the part of several inter-related pathways and the interacting functional networks (Vidal et al., 2011) and therefore diseases such as cancer are a manifestaion of a systemic problem. While many attempts are made to decode the underlying reasons, recent highthroughput experiments like protein-protein interaction assays, global mRNA expression analysis, systematic protein localization studies, structural studies by the medium of consortium and integration of mathematics, computer science and bioinformatics has provided a platform for a more comprehensive knowledge to test how a system may work. For example, it is clear that many oncogenes, tumor suppressors and other cancer related proteins act as crucial hubs in functional networks made of different circuits. These hubs are known to play a pivotal role in maintaining the integrity of the system as compared to non-hub proteins (Mackay et al., 2007; Patil et al., 2010b). PPI are dynamic processes and many oncogenes, tumor suppressors that are known to act as hubs in such network alter the biochemical wiring of the entire cell resulting in to phenotypic transformations (Koutsogiannouli et al., 2013b; Taylor et al., 2009).



**Figure.2.1**

*Technology advances to study biology of life. The highthroughput technologies gather information on numerous levels, including the genome, transcriptome (entirety of all genes that are converted into transcripts [i.e., mRNA molecules]), proteome (entirety of all proteins found in a given cell or tissue), metabolome (entirety of all metabolism products and intermediates in a cell or tissue), interactome (set of molecules, such as biologically active metabolism products, that interact with a given protein), and phenome (entirety of all observable characteristics of an organism) levels. The experimental data provide the structural and dynamic information that can then be used to generate mathematical formulas representing the observed reactions, leading to the development of comprehensive models and pathway maps (Fischer, 2005).*

## 2.2 Hub Proteins

Hub proteins are those proteins which are highly connected to other proteins and regulate several pathways. The connectivity ( $k$ ) of a protein is defined as the number of proteins with which it interacts and  $k$  values of hub proteins are in the range of 3-8 proteins (Ekman et al., 2006). Hubs can be classified as party hubs (PHs) and date hubs (DHs). Party hubs are believed to interact most of their partners at the same time where as date hubs interact with their partners at different times (Patil et al., 2010b). The distinction between party and date hubs were identified in the aspects of their expression correlation with their partners, topological positions in networks, genetic connectivity and evolution rates. Party hubs are found to be more common in eukaryotic species than date hubs (Fraser, 2005). Hub proteins are also classified based on the number of binding sites. They can interact with other proteins through the single binding sites or one or two binding interfaces (non-sociable/singlish interface) or through three or more interfaces (sociable/multi-interface hubs) (Patil et al., 2010b).

Genomic studies show that deleting a highly connected protein node (hub) is more likely to be lethal to an organism than deleting a low connected node (non-hub). This phenomenon is known as the centrality-lethality rule (He and Zhang, 2006). So the hubs are more important than non-hubs in organizing the global network structure. Some studies state that hub proteins are central to all functions and stability of PPI. They are physiologically more important and evolutionary more conserved than non-hub protein. Few examples include p53, p21, p27, BRCA1, ubiquitin, calmodulin etc which are well studied proteins which play a central role in various cellular mechanisms and acts as hub proteins. So dysfunction of these (hub) proteins may lead to several diseases such as cancer (Patil et al., 2010b).

### 2.3 Structural Characteristics of Hub proteins

The potentiality of the hub proteins lies in understanding its ability to recognize other proteins with required specificity, the requirement of structural or sequential characteristics that can identify multiple proteins through same or different binding surfaces. One of the important criteria in recognizing the interacting partners are the structural characteristics that support the process. The structural properties of hub proteins are varied than non-hub proteins.

**Intrinsic disorder** is an important factor or the characteristic of being hub proteins (Haynes et al., 2006). Structural flexibility or the ability of a protein to fold into an ensemble of conformations is one of the most significant factors affecting its binding ability. Different conformations of the proteins allow the flexibility to interact with different interacting partners. Local flexibility is dictated by small loops and coils in the proteins whereas global flexibility is dictated by large disordered regions inside the proteins (Patil et al., 2010b; Patil and Nakamura, 2006). Disordered regions are large unfolded regions in a protein that have no tertiary structure and little or no secondary structures. The hypothesis that hub proteins acquire more flexibility as compared to non hub proteins was studied by Dunker et al., (Dunker et al., 2005).

The disordered region in the hub is known to be present in two forms. Firstly, flexible linker that connects two ordered domains allowing the unrestricted movement with reach other. Ubiquitin-conjugating enzyme (Ubc1), an E2 ubiquitin ligase are hub proteins with a 22 residue disorder region acts as flexible linker (Merkley and Shaw, 2004). Calmodulin has 36 residue long disordered region that connects its two  $\text{Ca}^{2+}$  binding domains which allows it to bind different targets (Wilson and Brunger, 2000). Secondly, the disordered region itself acts the binding region. For example., p53 the

well known player in cell cycle is the important transcription factor and tumor suppressor binds to MDM2 through the disordered N-terminal region (Kussie et al., 1996). Cyclin-CDK complex binds to N-terminal disordered region of their inhibitors p21 and p27 (Kriwacki et al., 1996; Lacy et al., 2004). BRCA1 protein has a large central disordered region of approximately 1500 residues which binds to DNA and several proteins, but also acts as a flexible linker between its N-terminal RING domain and two C-terminal BRCT domains (Mark et al., 2005). Disordered regions bind with their target proteins with high specificity and high efficiency.

### **Surface Charge**

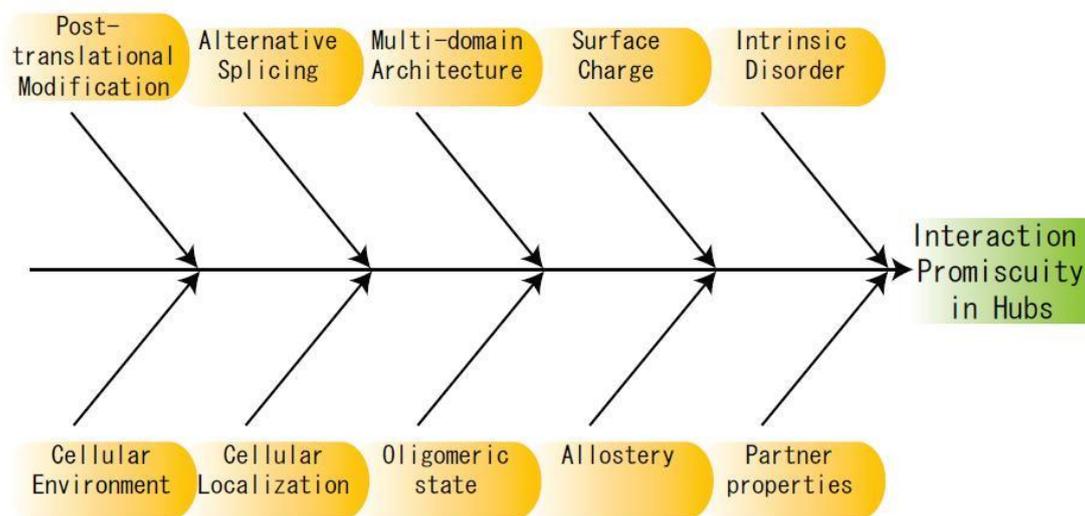
There are many other hub proteins which do not have any disordered residues. Proteins like Ubiquitin, Ferredoxin, Ras, small GTPases, Cofilin show highly charged surface area and have no large disordered regions. Electrostatic interactions are known to play an important role in possessing binding specificity, stable complex formation (Patil and Nakamura, 2007; Sheinerman et al., 2000). Stable complex formation includes the specific charged interactions by the residues at the interface area which contributes to the maximum binding energy. The stable complex formation is also contributed by the interactions outside the interface area which allows the stability of the complex. The interface of hub are enriched in residues like Arginine (Arg), Tyrosine (Tyr), Histidine (His) and Methionine (Met) which facilitates different types of interactions. The enrichment of these residues at the hot spots enhances the ability of the protein to form multiple interactions.

### **Domain Distribution and Enrichment**

Proteins enriched in multiple domains are known to act as potential hub proteins. Many proteins interact with their interacting proteins through single domain whereas several hubs interact with different interacting partners through different domains

(Patil et al., 2010a). Importin-beta, Elastase, Thioredoxin are few examples of proteins that interact with their interacting partners through single interface area whereas ubiquitin, Ras, Cdc42 interact with multiple interacting partners through multiple interfaces. Signaling proteins act as hub proteins and are enriched in domains like SH2, SH3 domains etc.

The other characteristics of hub proteins are mentioned in the diagram below (adapted by Ishikawa).



**Figure 2.2**

*Characteristics affecting the interaction promiscuity in hub proteins*

### **2.4 Oncogenes/Tumor suppressors as hub proteins**

Oncogenes and tumor suppressors are the proteins which target multiple signaling pathways at different subcellular levels. They bind to different proteins and stabilize the cellular network. During an external stimulus, they stabilize or rewire the network that leads to an altered phenotype. Few examples are cited below.

**Mitogen activated protein kinases (MAPKs)** are ubiquitously expressed in all cell types. MAPK activity is activated in various cellular signaling pathways like cell

proliferation, survival, motility, metabolism, transcription and translation. MAP kinases are terminal kinases and are activated on phosphorylation. There are various forms of MAP kinases involved in various signaling pathways. Hence, MAPKK acts an important hub that integrates responses to provide specificity in MAPK activation (Dhanasekaran and Reddy, 1998).

### **Tumor suppressor protein-p53**

p53 protein creates a fine balance between proliferation and apoptosis during oncogenic addiction. p53 activity is present at negligible levels inside the cell but is upregulated upon DNA damage. The induction of p53 is achieved through a post-translational mechanism that reduces p53 turnover. Induced p53 functions as a transcription factor for downstream genes that function in pathways of cell cycle regulation, apoptosis and DNA repair. It acts as a damage sensor, DNA repair effector, controls apoptotic checkpoints and in doing so it interacts with multiple proteins to bring about the effector function. Inside the nucleus it interacts with different proteins like Smads (Cordenonsi et al., 2003), p300(Shikama et al., 1999), different cofactors (JMY-junction mediating and regulatory protein) (Zuchero et al., 2009), axin (Li et al., 2009) etc to mediate its transcriptional activity. In the cytoplasm, it interacts with MDM2 an E3 ubiquitin ligase which induces monoubiquitination of p53 leading to its nuclear export whereas polyubiquitination of p3 protein leads to its degradation (Li et al., 2003). Zinc finger protein E4F1 binds to p53 and stimulates its recruitment on chromatin which leads to expression of target genes (Le Cam et al., 2006). It is known to form complexes with Bcl2, Bcl-xL, BAX and BAK to promote apoptosis (Green and Kroemer, 2009).

Hence, p53 is a well known and a well studied hub protein (Collavin et al., 2010).

## **β-catenin**

The Wnt/β-catenin signaling pathway plays a pivotal role in development and in maintenance of organs and tissues (Polakis, 2000). β-catenin is known to be activated in most of the cancers like colorectal cancer, medulloblastoma, glioma, hepatocellular carcinoma, melanoma etc. Upon activation it increases the transcriptional activity of target genes like Myc and TCF. β-catenin binds to E-cadherin which connects the adherens junction complex with the actin cytoskeleton and helps in cell-cell adhesion (Behrens et al., 1993). β-catenin complexed with E-cadherins is crucial for cell adhesion whereas uncomplexed β-catenin activates its transcriptional activity (Jeanes et al., 2008). Hence, it acts as a hub protein and is a potential therapeutic target (Polakis, 2000).

Since the present thesis focus is on a major oncoprotein gankyrin, a background literature of gankyrin is provided below.

### **2.5 Gankyrin- An Oncoprotein**

Fujita and colleagues (Higashitsuji et al., 2000) identified gankyrin as a gene that was consistently overexpressed in **human liver cancers** using cDNA subtractive hybridization. Gankyrin overexpressing NIH3T3 cells forms colonies in soft agar and are tumorigenic in nude mice. Gankyrin overexpression in mouse hepatocytes cells forms tumors with more vascularity. Gankyrin binds to and sequesters factor inhibiting hypoxia-inducible factor-1 (FIH-1), which results in decreased interaction between FIH-1 and hypoxia-inducible factor-1α (HIF-1α) and increased activity of HIF-1 to promote VEGF production which suggests that gankyrin might have the physiological roles in hypoxia conditions apart from its role as an oncoprotein (Liu et al., 2013). Panobinostat (LBH589), a hydroxamic acid-derived histone deacetylase inhibitor has shown promising anticancer effects on hepatocellular carcinoma by inhibiting

gankyrin/STAT3/Akt pathway (Song et al., 2013). LBH589 inhibited metastasis in vitro via down-regulation of N-cadherin, vimentin, TWIST1, VEGF and up-regulation of E-cadherin. It also induces apoptosis and G1 phase arrest in HCC cell lines.

Gankyrin overexpression was found in human **esophageal squamous cell carcinoma** (ESCC) tissues as compared to normal corresponding epithelia. Gankyrin overexpression is correlated with low survival rate, extent of primary tumor, lymph node metastasis, distant lymph node metastasis and is associated with poor prognosis (Ortiz et al., 2008; Ortiz et al., 2006).

Gankyrin expression was increased significantly in **pancreatic cancer**. Gankyrin expression was seen higher in poorly differentiated tumor tissues than in highly and moderately differentiated tumor tissues. Gankyrin downregulation in pancreatic cancer cell lines SW1990 decreases the cell growth, cellular proliferation and correspondingly their growth on soft agar. Downregulation of gankyrin expression caused an increase in the number of cells in the G1 phase, whereas the number of cells in the S phase was proportionally reduced. Gankyrin overexpression showed increase in tumor size in nude mice whereas gankyrin downregulation decreases the tumor size *in vivo* (Meng et al., 2010a).

Gankyrin is seen to be overexpressed in **human oral cancer** and occurs during the early stages of oral carcinogenesis. Thus it is considered as the chemopreventive target in oral cancer and is considered as the biomarker for epithelial carcinogenesis (Li et al., 2011).

Gankyrin is found to be highly expressed in **human lung cancers** that have Ras mutations and increased gankyrin expression is required for the constitutive activation of Akt and tumorigenesis in lung cancers (Man et al., 2010a).

Gankyrin was found to be overexpressed in **breast cancer** tissues as compared to normal breast epithelial cells. Gankyrin overexpression was associated with extensive intraductal carcinoma and ErbB2 positivity in invasive ductal carcinoma. It promotes breast cancer metastasis by regulating Rac1 activity. Overexpression of gankyrin accelerates focal adhesion turnover and increases cell migration. It was also observed that downregulation of gankyrin in mouse mammary tumor cells significantly decreases tumor metastasis to lung in animal models (Zhen et al., 2012a). It was also observed that gankyrin overexpression promotes metastasis in breast cancer cell lines under hypoxia conditions by regulating hypoxia inducible factor-1, hypoxia-inducible factor-1 $\alpha$  and partly through regulating E-cadherin (Gao et al., 2014).

**Gliomas** are the common primary intracranial tumor which has high proliferation and invasive potential. Gankyrin expression was found to be high in gliomas as compared to paracancerous tissues. Downregulation of gankyrin in U251 decreased the tumorigenicity *in vitro* and *in vivo*.

In **colorectal carcinoma (CRC)**, gankyrin is found to be overexpressed as compared to controls and the gankyrin expression correlated with Tumor Node Metastasis (TNM) stages and metastasis of CRC. Overexpression of gankyrin in LOVO cells promoted the cellular proliferation and tumorigenicity whereas the downregulation of gankyrin exerted the opposite effects. Co-expression of cyclin D1 and  $\beta$ -catenin positively correlated with gankyrin expression (Tang et al., 2010a). Gankyrin promotes metastasis in colorectal carcinoma via activation of IL8 (Bai et al., 2013b). In colorectal carcinoma, expression of vascular endothelial growth factor (VEGF) and stem cell markers (CD133, Nanog, Oct-4) expression was correlated with gankyrin expression (Mine et al., 2013).

Owing to the important role of gankyrin in cancer it is one of the potential therapeutic target.

## **2.6 Gankyrin- non-ATPase subunit of 26S Proteasome**

Gankyrin is a non-ATPase subunit of proteasome. Gankyrin is a chaperone and known to be transiently associated with 26S Proteasome. It is detected in free form inside cytoplasmic and the nuclear compartment of the cell and is also found to be associated with 19S regulatory particle (Dawson et al., 2002). Nas6P is the yeast homologue of gankyrin which acts as a chaperone and is required for the assembly of the 19S regulatory particles (Bedford et al., 2010; Roelofs et al., 2009). Gankyrin is highly conserved throughout evolution (~40% identity to yeast Nas6P). It is localized on human chromosome Xq22.3 in a region where DNA gains are frequently detected in kidney and colon carcinomas. Gankyrin consists of seven ankyrin repeats. Ankyrin repeats are known to be involved in protein-protein interactions (Nakamura et al., 2007b).

## **2.7 Gankyrin- Dual negative regulator of tumor suppressor p53 and Rb**

### **Retinoblastoma (Rb)**

Retinoblastoma (Rb) is a well studied tumor suppressor protein. Rb confers its function by regulating crucial proteins like the transcription factors of the E2F family. E2F1 proteins are known to bind to Rb1 and are released when Rb1 is inactivated allowing it to facilitate the transcription of genes important for progression into late G1 and S phases. Gankyrin binds to Rb through retinoblastoma binding motif LXCXE leading to the phosphorylation of Rb and subsequently degradation. This leads to increase in E2F1 activity and subsequent increase in cell cycle progression by escape of G1 to S

cell cycle check points. This suggests that increased expression of gankyrin promotes tumorigenicity by targeting Rb to the proteasome (Higashitsuji et al., 2000a).

### **P<sup>53</sup>**

Under normal conditions, p53 binds to E3 ubiquitin ligase MDM2 and remains in an inactive stage. During any stress or DNA damage, p53 gets dissociated from MDM2 leading to cell cycle arrest to allow DNA repair and survival of the cell or leads to apoptosis of the cell. It is known to control apoptosis by transactivation of many pro-apoptotic genes and repressing the transcription of antiapoptotic genes. Gankyrin binds to MDM2, an E3 ubiquitin ligase leading to ubiquitination of p53 and subsequently leading to degradation of p53 (Higashitsuji et al., 2005).

Thus gankyrin acts a dual negative regulator of two tumor suppressor proteins Rb and p53.

### **2.8 Gankyrin – Negative regulator of NF-κB activity**

NF-κB is a transcription factor that regulates a large number of genes involved in cell proliferation, differentiation, apoptosis and immune and inflammatory responses. Under basal conditions, NF-κB is regulated by the IκB family of proteins. The IκB family of proteins contains ankyrin repeats which facilitate binding to the the Rel Homology domain and in turn, inhibit the nuclear localization signal for NF-κB. Under stress conditions, IκB is phosphorylated leading to the polyubiquitylation and degradation of IκB and hence leading to the accumulation of NF-κB in the nucleus allowing transcriptional control (Chen et al., 2007d; Higashitsuji et al., 2007a).

Gankyrin interacts with RelA and suppresses NF-κB activity by sequestering NF-κB inside the cytoplasm. It also suppresses its transcriptional activity by regulating acetylation through SIRT1 which is a class 3 histone deacetylase (Higashitsuji et al., 2007a).

## **2.9 Ras induced activation of Gankyrin**

The Ras subfamily of proteins is a family of proteins in which each member belongs to the small GTPase class. These proteins are involved in cellular signal transduction and they directly affect proteins involved in cell growth and differentiation. It is well studied that uncontrolled Ras signal leads to malignancy. Ras induces tumorigenesis by regulating its large number of downstream effectors of which are linked to many different signalling pathways. PI3K is one such effector of Ras and is crucial in Ras induced tumorigenesis by activating Akt which is involved in inhibiting apoptosis. Mutations or deletions in phosphatase and tensin homolog (PTEN) which is a tumor suppressor leads to activation of Akt. RhoA protein belonging to the Rho GTPase family and its effector kinase ROCK have been reported to be essential for PTEN activity. Recently it has been shown that gankyrin is induced by active Ras G12V mutation which induces cellular transformation and tumorigenesis by indirectly effecting the inhibition of ROCK and in turn, facilitating the activation of Akt through this RhoA/ROCK/PTEN pathway (Man et al., 2010a).

## **2.10 Gankyrin interacts with S6 ATPase of the 26S Proteasome**

The yeast two hybrid assay showed that gankyrin interacts with free S6 ATPase and S6 ATPase associated with 26S Proteasome (Dawson et al., 2002). The C-terminal 78 amino acid residues of S6 ATPase are necessary and sufficient to mediate the interaction with gankyrin. The three dimensional structure of the mouse gankyrin- C terminal S6 ATPase (S6-C) has been solved (Nakamura et al., 2007b). The interface between gankyrin and S6-C buries a solvent-accessible surface area of about 2418 Å<sup>2</sup>. The electrostatic surface potential of the complex shows that the interactions are

dominated mainly by polar interactions. There are 25 electrostatic intermolecular interactions between gankyrin and S6-C. The detailed study showed that S6-C mutants E356A/E357A, D359A/D362A, R338A/R339A/R342A and gankyrin mutants R41A and R41A/K116A significantly disrupted the interaction in the complex as compared to wild-type gankyrin-S6C.

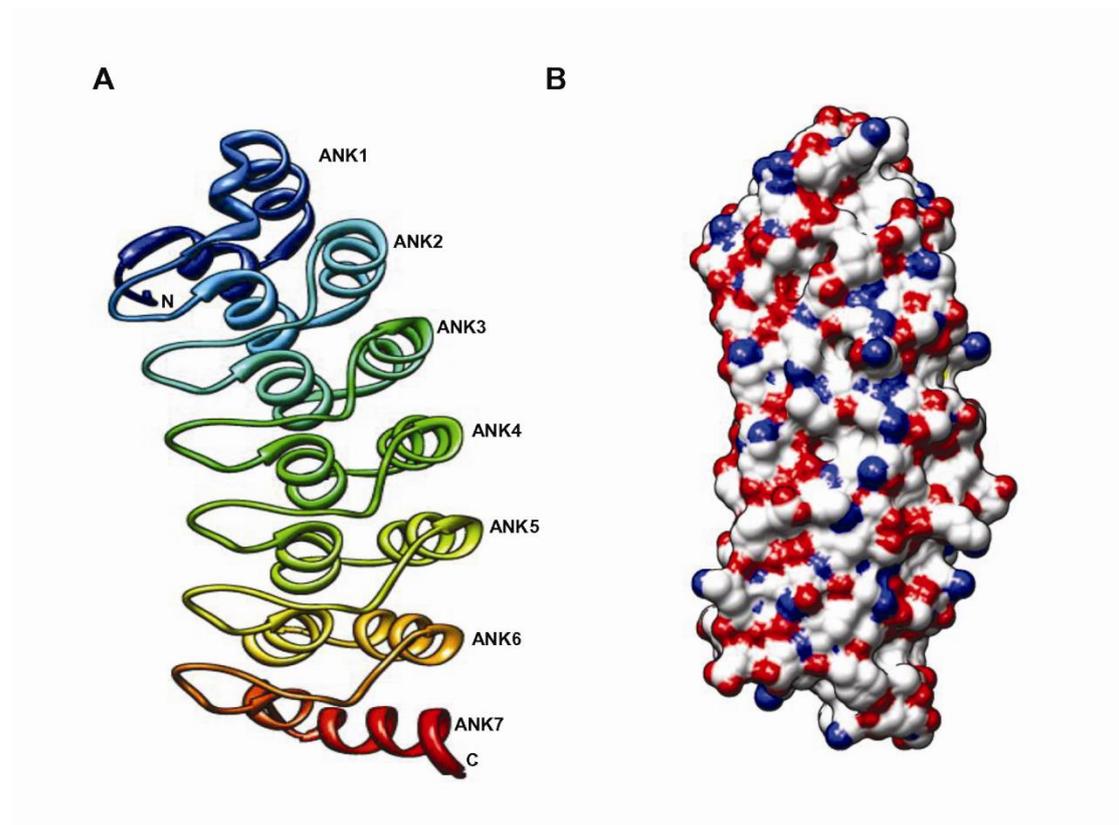
### **2.11 Gankyrin interacts with CDK4- Counteraction of gankyrin to regulate INK4-CDK4-Rb pathway**

Gankyrin interacts with cyclin-dependent kinase 4 (CDK4) resulting in a gankyrin-CDK4-Cyclin D2 ternary complex. CDK4 and Rb binding are independent and involve distinct structural regions on gankyrin. In doing so, gankyrin competes with INK4A, an inhibitor of cyclin kinases, for binding to CDK4. Based upon these findings, gankyrin appears to indirectly activate CDK4, resulting in the hyperphosphorylation of Rb and concomitant deregulation of E2F1-mediated transcription and cell cycle progression. Thus gankyrin is shown to play a major role in regulating cell cycle (Dawson et al., 2002).

### **2.12 Gankyrin-Hub protein**

Gankyrin seems to have been well studied with respect to its relationship to tumor suppressor proteins, some signaling pathways and association with other oncoproteins. However, the potential number of proteins that gankyrin may interact with seems less well studied. Gankyrin is made of seven ankyrin repeats which are well known to be involved in protein-protein interactions (Fig. 2.3). Gankyrin has a charged surface and electrostatic interactions mediated by such charged surfaces are known to play a role in specificity of the interactions and the stabilization of complex.

A key oncoprotein is likely to be involved in multiple cellular processes, such as resistance to apoptosis, metastasis and invasion through a number of mechanisms. These have not been completely realized. Hence, gankyrin seems to have all the characteristics of hub proteins the physical and functional interactions of which have to be validated in great detail. Mapping such interactions forms the major focus of the current thesis.



**Figure 2.3**

**Structure of Gankyrin.** (A) Ribbon representation of human gankyrin. Seven ankyrin repeats are indicated (ANK1-ANK7). The N-termini and C-termini are labeled. Figures are generated using PyMOL. (B) A surface potential representation of gankyrin. Red, blue, and white represent acidic, basic, and neutral, respectively.

### 2.13 Identification of Protein-Protein Interactions

PPI is defined as the interactions which are intentional, specific and involved in biomolecular events (De Las Rivas and Fontanillo, 2010). For example, interactions that occur during protein synthesis, their folding and degradation should be excluded under this definition (De Las Rivas and Fontanillo, 2010). During setting experimental assays such generic interactions should be filtered.

PPIs can be static or transient (Byrum et al., 2012). Certain interactions are stable or permanent- such as proteins involved in ATP synthase which constitutes macromolecular protein complexes and cellular machines or interactions between different subunits of 26S Proteasome, the cellular degradation machinery. Some interactions do not occur in all cells at any time and are called as transient- such as activation of gene expression by binding of transcription factors and signaling. Interactions are also cell type specific, cell cycle specific, specific to developmental stages, protein modifications (phosphorylation, acetylation) etc (De Las Rivas and Fontanillo, 2010; Perkins et al., 2010). An analysis of different PPI networks identified through large scale efforts suggest that the common nodes are few and those that can be reproduced in first or second experiment are increasingly small. Therefore, mapping functionally crucial physical, specific, protein-protein interactions continue to remain a major challenge. Some of the approaches used to screen for PPI complexes is summarized below.

***Table 2.1 Experimental approaches used for identifying protein-protein interactions***

Methods	Advantages	References
Mass spectrometry MAP-SILAC	1. Identifies dynamic interactors of the protein complexes.	(Abu-Farha et al., 2008; Link et al., 1999)

SILAC combined with affinity matrices (bead proteomes)	<ol style="list-style-type: none"> <li>1. Removes non-specific binders.</li> <li>2. Specificity filter to distinguish specific protein binding partners in both quantitative and nonquantitative pull-down and immunoprecipitation experiments</li> </ol>	(Oeljeklaus et al., 2009)
Fluorescent based assays	<ol style="list-style-type: none"> <li>1. Sensitive, rapid, detection at low concentration (picomolar) and single molecules</li> </ol>	(Royer and Scarlata, 2008)
Yeast two-hybrid system	<ol style="list-style-type: none"> <li>1. Simple to set up and inexpensive.</li> <li>2. Transient and weak interactions which are often important in signaling cascade are detected.</li> <li>3. This method generates large number of false positives and 50% of the interactions may be unreliable.</li> </ol>	(Byrum et al., 2012)
Chemical cross linking	<ol style="list-style-type: none"> <li>1. Used to capture transient or low affinity interactions.</li> </ol>	(Back et al., 2003; Tang and Bruce, 2009)
Tandem affinity purification	<ol style="list-style-type: none"> <li>1. Nonspecificity is low.</li> <li>2. Highly sensitive and selective.</li> <li>3. Low affinity interactions are missed.</li> </ol>	(Jessulat et al., 2010)

Peptide based approach have also been used for identification of interacting partners. Strategies like proteome peptide scanning and whole interactome scanning experiments (WISE) (Landgraf et al., 2004) which rely on short peptide sequences

within proteins (Short Linear Motifs or SLiM) and DOMINO for domain-peptide interactions have been used to identify new interactions (Ceol et al., 2007; Edwards et al., 2007). However, identifying functionally relevant SLiMs from random occurrences in eukaryotes does not seem straightforward. This is because many such SLiMs seem to be part of unstructured or disordered regions in proteins with large variability in sequence making it difficult to accurately predict or use this information.

### **2.14 Hot spot sites**

A concept that is well known in studies that encompass analysis of various individual protein complexes is called the hot spot site. It has been long realized that although that protein-protein interactions are spread through a large interface area, there are few groups or the cluster of residues in the interface which contributes to the maximum binding energy. When mutated to alanine these residues are shown to disrupt the interactions (Ofra and Rost, 2007). These are called the 'hot spot sites' as they provide bulk of the binding energy. Hot spot sites are conserved across interfaces and represent very few residues, they may also be vulnerable for interventions. Hot spots are typically defined as those residues for which  $\Delta\Delta G \geq 2 \text{kcal/mol}$  and they are known as the indicator of specificity (Bogan and Thorn, 1998). Three amino acids-tryptophan, arginine and tyrosine appear frequently in hot spots whereas leucine, methionine, serine, threonine and valine residues are seen rarely in hot spots (Bogan and Thorn, 1998). It has been shown that most hydrophobic residues are predominantly found in the interior of proteins, while polar and charged residues are preferred on surfaces (Tsai *et al.*, 1997). Due to differences in side-chain conformational entropy aspartate is favoured more over glutamate whereas asparagine is favoured more over glutamine (Bogan and Thorn, 1998).

Despite such structural and molecular details that can define the specificity in PPI, hot spot sites have not been widely utilized for the identification of PPI. There are many algorithms which help in predicting hot spot residues like iPRED, CPORT for predicting protein-protein interface residues, PPIcons helps in identifying PPI in selected proteomes using local sequence segments depending upon their physio-chemical characteristics, PRISM, SLiMFinder etc. Nevertheless there is no comprehensive report of the utility of hot spot residues to map a physical and functional network of a hub protein. This thesis predicts and utilizes a potential hot spot site for the identification of gankyrin interacting partners.

# *Objectives*

With the above background, the objectives for the proposed topic are -

1. Identifying the novel interacting partners of gankyrin and validating them in epithelial cancer cell lines.
2. Validating functional relevance of at least two of the interactions.

## **CHAPTER 3.**

### ***Materials and Methods***

## **MATERIALS**

### **3.1 Buffers and Reagents:**

#### **3.1.1 Luria-Bertani (LB) medium (for 1 L)**

NaCl	10 g
Tryptone	10 g
Yeast extract	5 g

Deionized water (MQ) was added to a final volume of 1 litre and pH was adjusted to 7.0 with 1 M NaOH and autoclaved.

Or

25 g of LB powder (Merck) was dissolved in 1 L of MQ water and autoclaved.

#### **3.1.2 LB-Ampicillin Agar Plates (for 1 L)**

NaCl	10 g
Tryptone	10 g
Yeast extract	5 g
Agar	20 g

Deionized water was added to a final volume of 1 litre and pH was adjusted to 7.0 with 10 N NaOH and autoclaved. It was cooled to about 55 °C and 1 ml of 100 mg/ml ampicillin was added. The media was poured into Petri dishes (~25 ml/100 mm plate).

#### **3.1.3 Tris-EDTA (TE) Buffer (for 50 ml)**

Tris	60.66 mg (10 mM)
EDTA	14.62 (1 mM)

pH was adjusted to 7.5 with 10 N NaOH and autoclaved.

### **3.1.4 Ampicillin Stock**

Stock concentration : 100 mg/ml (Filter sterilized using 0.2  $\mu$ m membrane)

Working concentration: 100  $\mu$ g/ml

### **3.1.5 50X TAE Buffer (for 1 L)**

Tris base 242 g

Glacial acetic acid 57.1

0.5 M EDTA (pH 8.0) 100 ml

Prepare 1X TAE buffer for agarose gel electrophoresis.

### **3.1.6 6X Gel Loading Buffer for DNA (for 100 ml)**

Xylene Cyanol FF 0.25 g (migrates at 4160 bp with TAE)

Bromophenol blue 0.25 g (migrates at 370 bp with TAE)

Glycerol 30 ml

### **3.1.7 Ethidium Bromide (EtBr)**

Stock concentration 10 mg/ml (20000X)

Working concentration 0.5  $\mu$ g/ml

### **3.1.8 Buffers for Ni NTA Column Purification and Gel Filtration:**

#### **3.1.8.1 Ni-NTA Lysis Buffer (for 1 L)**

Tris 6.06 g (50 mM, pH 8.0)

NaCl 29.22 g (500 mM)

Imidazole 0.6 g (10 mM, reduces non-specific binding of proteins)

Glycerol 100 ml (10%)

TritonX-100	1 g (0.1%)
Protease inhibitor (10X)	1X
BME (14.3 M)	3.5 ml (50 mM)
Lysozyme	1 g (1 mg/ml)

#### **3.1.8.2 Ni-NTA Binding/Washing Buffer (for 1 L)**

Tris	6.06 g (50 mM, pH 8.0)
NaCl	29.22 g (500 mM)
Imidazole	0.6 g (10 mM, reduces non-specific binding of proteins)
Glycerol	100 ml (10%)
TritonX-100	1 g (0.1%)
Protease inhibitor (10X)	1X
BME (14.3 M)	3.5 ml (50 mM)

#### **3.1.8.3 Ni-NTA Elution Buffer (for 1 L)**

Tris	6.06 g (50 mM, pH 8.0)
NaCl	29.22 g (500 mM)
Imidazole	30.04 g (500 mM, for elution)
Glycerol	500 ml (10%)
TritonX-100	1 g (0.1%)
Protease inhibitor (10X)	1X
BME (14.3 M)	3.5 ml (50 mM)

#### **3.1.8.4 Running Buffer for Gel Filtration (for 1 L)**

Tris	6.06 g (50 mM, pH 7.5) or 25 ml of 2 M Tris pH 7.5
------	--

NaCl	17.53 g (300 mM) or 60 ml of 5 M NaCl
BME (14.3 M)	3.5 ml (50 mM)

### **3.1.9 NP-40 Lysis Buffer (for 50 ml)**

Tris	0.121 g (20 mM, pH 7.5) or 0.5 ml of 2 M Tris pH 7.5
NaCl	0.4383 g (150 mM) or 1.5 ml of 5M NaCl
NP-40	250µl of absolute solution
DTT	0.0077 g (1mM DTT) or 50 µl of 1M DTT solution

### **3.1.10 1X Transfer Buffer (for 1 L)**

Glycine	14.4 g
Tris Base	3.02 g
Milli-Q	0.8 L
Methanol	200 ml

### **3.1.11 10mM Sodium bicarbonate buffer pH 9.3 (50 ml)**

Sodium bicarbonate	0.42 g
Sodium carbonate	0.17 g

### **3.1.12 10 mM Phosphate Buffer pH 7.5 (1 L)**

Monosodium phosphate	0.2596 g
Disodium phosphate	2.1758 g

### **3.1.13 1X Phosphate buffer saline pH 7.5 (1 L)**

Sodium phosphate	1.4 g
------------------	-------

NaCl                    8.76 g

### **3.1.14 Tissue Culture Media and Reagents**

Tissue Culture Medium: Commercially available powdered medium, Dulbecco's Modified Eagle Medium (DMEM) containing high glucose, pyridoxine hydrochloride and sodium pyruvate (Invitrogen) was prepared as per the manufacturer's instructions. Powdered medium was reconstituted in 800 ml autoclaved Milli-Q water under sterile conditions. 3.5 g sodium carbonate was added and pH was adjusted to 7.2 using 1 N HCl. The volume was made up to 1 L in a sterile volumetric flask and the medium was filter sterilized through a 0.22 µm pore size membrane which was fitted in the sterile filter assembly. The filtered medium was stored at 4 °C as 500 ml aliquots. 10% FBS was added to make up the complete medium.

#### **3.1.14.1 10X phosphate buffered saline (PBS)**

NaCl                    80.8 g

KCl                     2.0 g

Na<sub>2</sub>HPO<sub>4</sub> · 2H<sub>2</sub>O    12.6 g

KH<sub>2</sub>PO<sub>4</sub>                2.0 g

Glucose                10.0 g

The components were dissolved in autoclaved Milli-Q water and the volume made up to 1 L. The solution was filter sterilized and stored at 4 °C.

#### **3.1.14.2 10X Trypsin (0.25%)**

2.75 g of trypsin was added to 110 ml of autoclaved Milli-Q water and allowed to dissolve. The solution was sterilized by filtering through a 0.22 µm pore size filter.

The solution was stored as 10 ml aliquots at -20 °C. 10X stocks were diluted to 1X working solution with 1X PBS. Working solution was stored at 4 °C.

#### **3.1.14.3 2X BBS (BES Buffered Saline) (for 50 ml)**

50 mM BES	0.533 g
280 mM NaCl	0.818 g
1.5 mM Na <sub>2</sub> HPO <sub>4</sub> ·2H <sub>2</sub> O	0.0134 g

All the reagents were dissolved in 45 ml of Milli-Q water. The volume was made up to 50 ml and filter sterilized using 0.22 µm filter (Millipore) and stored as 0.5 ml aliquots at -20°C.

#### **3.1.14.4 0.5M CaCl<sub>2</sub> (for 50 ml)**

CaCl <sub>2</sub> ·2H <sub>2</sub> O	3.675 g
--------------------------------------	---------

CaCl<sub>2</sub>·2H<sub>2</sub>O was dissolved in 45 ml of Milli-Q water. The volume was made up to 50 ml and filter sterilized using 0.22 µm filter (Millipore) and stored as 0.5 ml aliquots at -20°C.

#### **3.1.15 Other Reagents (for 100 ml)**

2 M Tris	24.22 g (pH 7.5, adjust pH with concentrated HCl)
1 M MgCl <sub>2</sub>	9.52 g
1 M DTT	15.42 g
5 M NaCl	29.22 g
10 M NaOH	58.44 g

All the reagents were filtered with 0.22 µm filter membrane (Millipore).

## **3.2 Experimental Protocol:**

### **3.2.1 Primer Reconstitution**

All the primers used were obtained from Sigma and were received as a nucleic acid pellets in powder form (not visible with naked eyes). The primer pellet was centrifuged at 10,000 rpm (Eppendorf centrifuge, Model - 5417C) for 1 min. Primers were suspended in 10 mM Tris, pH 7.5 to obtain a stock concentration of 100  $\mu$ M as per the manufacturer's instruction written on the primer tube. The stock primers of 100  $\mu$ M were allowed to suspend in 10 mM Tris, pH 7.5 on ice bath for about 1 hour with intermittent vortexing after every 15 min. Reconstituted primers were stored at -20° C. The working stock of 10  $\mu$ M was prepared.

### **3.2.2 Determination of Nucleic Acid Concentration**

The concentration of the nucleic acid in solution was estimated using a spectrophotometer (NanoDrop, Model - ND 1000). The absorbance of the solution was measured at 260 nm and concentration was calculated using the following formula:

$$1 \text{ OD}_{260} = 50 \text{ } \mu\text{g/ml for double stranded DNA}$$

$$1 \text{ OD}_{260} = 40 \text{ } \mu\text{g/ml for RNA}$$

### **3.2.3 Polymerase Chain Reaction (PCR)**

PCR amplification was done using pfu PCR kit (Fermentas), 25 mM dNTPs, DMSO, DpnI (Fermentas). The proofreading activity of pfu enzyme provided error free amplification. A 50  $\mu$ l PCR reaction containing 50-60 ng template plasmid and control reaction (without pfu) was set-up.

5X – reaction Buffer	10 $\mu$ l
Template DNA	50 -100 ng
25 mM dNTP Mix	1 $\mu$ l
MQ Water	Variable
Primer (Forward 10 $\mu$ M)	1 $\mu$ l
Primer (Reverse 10 $\mu$ M)	1 $\mu$ l
DNA polymerase	1 U
Reaction volume	50 $\mu$ l

The cycling steps used were initial denaturation- 95 °C for 5 min, denaturation- 95 °C for 1 min, annealing - 55-65 °C for primer annealing for 1 min, extension- 72 °C for 1-4 min (depending upon PCR product size) and final extension for 10 min at 72 °C.

### **3.2.4 Site Directed Mutagenesis**

For site directed mutagenesis total 18 cycles were repeated without any final extension. The extension time of 30 sec per Kb of PCR product amplification was used for high fidelity fusion DNA polymerase (Thermo Scientific). The following PCR condition was used: initial denaturation – 95 °C for 5 min, denaturation- 95 °C for 1 min, annealing – 50 °C for 1 min, extension- 72 °C for 1min/kbp and number of cycle 19. The 10  $\mu$ l PCR product and control reaction (every component except pfu enzyme) were resolved on 0.8% agarose gel. After confirming amplification, DpnI digestion was setup in a 25  $\mu$ l reaction using 20  $\mu$ l of PCR product, 2  $\mu$ l of 10X Tango buffer and 10 units of DpnI for at least 8 h at 37 °C. DpnI would digest the parental plasmid (cleave adenomethylated dam sites). The digested product was then transformed in XL1 blue cells. The colonies were screened, plasmids isolated and finally sequence verified.

### **3.2.5 Restriction Digestion Reaction**

10X buffer (Tango/FD/FD Green)	2X
DNA template	250-300 ng
Restriction enzyme	1 unit/ 1µl for FD enzymes
Autoclaved MQ water	variable
Final Reaction Volume	20 µl

Restriction reaction was carried out at 37° C for 1-4 hours.

### **3.2.6 Agarose Gel Electrophoresis of DNA**

Agarose of 0.8 – 1 % was prepared in 1X TAE and heated to boil using microwave oven (MS-2342-AE, GE). Agarose solution was allowed to cool to ~50-60 °C (5 min) and ethidium bromide (EtBr) was added to a final concentration of 0.5 µg/ml. The solution was mixed thoroughly and poured in a gel casting tray with comb. The gel was allowed to polymerize for 20-30 min. The samples were loaded along with the 1X DNA loading buffer. The samples were resolved at 120 mA for 30-50 min. The gel was then documented (UVP, Bioimaging Systems) and viewed using LaunchVision Works LS software.

### **3.2.7 Recovery of DNA from Low Melting Agarose Gel**

Low melting agarose gel of 0.8% with ethidium bromide (0.5 µg/ml) was casted using 1X TAE buffer. The samples were loaded with DNA gel loading buffer and allowed to resolve at 120 mA for 30-50 min. The gel was then viewed under trans-illuminator and the band of interest was excised using a clean scalpel blade. The agarose gel piece containing DNA was allowed to melt at 65°C and DNA was extracted using gel elution columns (Sigma). In final step DNA was eluted in 20 µl of elution buffer.

### **3.2.8 Ligation/Cloning**

Ligation was set up using Fermentas rapid ligation kit. For cloning pJET cloning kit (Fermentas) was used. 20 µl of ligation reaction was set up with 80 ng of vector DNA, 30-40 ng of insert, 1U of T4 DNA ligase and 10 µl of 2X rapid ligation buffer. The ligation reaction was incubated at 22 °C for 4-5 hours.

### **3.2.9 Transformation**

The ultra-competent cells were thawed on ice [*E. coli* DH5α or *E. coli* XL1 for mutagenesis and cloning and *E. coli* Bl 21 (DE3), Origami DE3 for protein expression]. 10 µl of the Ligation mixture or 100 ng of plasmid DNA was added to an aliquot of 100 µl ultra-competent cells. It was tapped gently and incubated in ice for 30 min. The cells were heat shocked at 42°C for 90 sec and incubated in ice for 2 min. 500 µl of LB medium was added to the tube and kept for outgrowth at 37 °C for 45 min with vigorous shaking (200-250 rpm). The cells were centrifuged at 5000 rpm for 3 min. The cells were suspended in 50 µl of the medium and plated on LB ampicillin (100 µg/ml) plates and incubated overnight (16 hours) at 37 °C.

### **3.2.10 Plasmid Mini Preparation**

For the analysis of desired site directed mutagenesis or to select the desired clone, a single colony of bacteria was picked up with the help of autoclaved tooth pick and inoculated in 5 ml of LB medium containing 100 µg/ml of ampicillin. The culture was allowed to grow overnight at 37 °C shaker incubator (Lab companion, Model – SIF6000R). Bacteria were pelleted down by centrifugation at 5000 rpm (Plasto Craft, Model - Rota 4R). Plasmid miniprep was done using miniprep spin columns (Sigma).

### **3.2.11 Plasmid Construction**

Gankyrin pBluescript II SK (1) construct (kind gift from Dr. Jun Fujita, Kyoto University) was subcloned into mammalian expression vector p3XFLAG-CMVTM-10 (Sigma) or in prokaryotic expression vector pRSETA-TEV. Hsp70 in pGEX4T1 vector was received as a kind gift from Dr. Surekha Zingde, ACTREC. CLIC1, NCK2, GRSF1, Calreticulin, P domain of calreticulin, S6 ATPase cDNA were generated by RT-PCR of RNA extracted from HEK 293 cells. CLIC1, Calreticulin and P domain of calreticulin was cloned in pGEX4T1 (GE Amersham) and pCDNA 3.1 with HA tag at the N terminus (pCDNA3.1 was received as a kind gift from Dr. Sorab Dalal, ACTREC), while NCK2 and GRSF1 were cloned in pCDNA3.1. Hsp70 was also cloned in p3XFLAG-CMVTM-10. Mutations in Hsp70 (aa638EEVD641 to AAVA), CLIC1 (aa 150EEVD153 to AAVA), CLIC1 (aa 152V to aa 152E), NCK2 (aa 167EEVD170 to AAVA), GRSF1 (aa 144EEVD147 to AAVA), Calreticulin and P domain of calreticulin (aa 255EEMD258-255AAMA258) S6 ATPase (aa 356EEVD359 to AAVA) were generated using PCR based site directed mutagenesis with the help of Phusion high fidelity DNA polymerase (Finnzymes- Thermo Scientific) and were further confirmed by sequencing.

**Table 3.1 Primers used for Cloning and site directed mutagenesis**

<b>Construct</b>	<b>Forward Primer</b>	<b>Reverse Primer</b>
pCMV 10-3X-FLAG - gankyrin	5'AAGCTTATGGAGGGGTGTGTGTCTAACCC 3'	5'GAATTCTTAACCTTCCACCATTCTCTTGAG 3'
pRSETA-TEV-gankyrin	5'GGATCCATGGAGGGGTGTGTGTCTAACCC 3'	5'GAATTCTTAACCTTCCACCATTCTCTTGAG 3'
pGEX4T1-CLIC1	5'GGATCCATGGCTGAAGAACAAC 3'	5'GCGGCCGCTTATTTGAGGGCCTTTGCCACTTGCT 3'
pCDNA 3.1-NCK2	5'GGATCCATGACAGAAGAAGTTATTGTGAT 3'	5'CTCGAGTCACTGCAGGGCCCTGACGAGGTA 3'
pCDNA 3.1-GRSF1	5'GGATCCATGGCCGGCACGCGCTGGGTA 3'	5'CTCGAGTTATTTTCCTTTTGGACATGAA 3'
pCDNA 3.1-CLIC1	5'GGATCCATGGCTGAAGAACAAC 3'	5'CTCGAGTTATTTGAGGGCCTTTGCCACTTGCT 3'
pGEX4T1-calreticulin	5'GAATTCATGCTGCTATCCGTGCCG 3'	5'GCGGCCGCCTACAGCTCGTCCTTGGCC 3'
pGEX4T1-P Domain (calreticulin)	5'GAATTCGACGATTGGGACTTCTCTG 3'	5'GCGGCCGCATAGGCATAGATACTGGGG 3'
pCMV 10-3X-FLAG-Hsp70	5'AAGCTTATGGCCAAAGCCCGCGCGAT 3'	5'GAATTCCTAATCTACCTCCTCAATGGTG 3'
pCMV 10-3X FLAG-S6 ATPase	5'AAGCTTATGGAGGAGATAGGCATC 3'	5'GAATTCTCACTTGTA AAACTCATGCC 3'

<b>Mutation</b>	<b>Forward Primer</b>	<b>Reverse Primer</b>
pGEX4T1-CLIC1_AAVA and pCDNA3.1-CLIC1_AAVA	5'TCCCCCCTCCCAGCCGCCGTGGCCGAAACCAGTGCT3'	5'AGCACTGGTTTCGGCCACGGCGGGCTGGGAGGGGGGA3'
pGEX4T1-CLIC1_EEED	5'TCCCCCCTCCCAGAAGAAG AAGATGAAACCAGTGCT3'	5'AGCACTGGTTTCATCTTCTTCTTCTGGAGGGGGGA3'
pGEX4T1-Hsp70_AAVA	5'AGGCCCCACCATTGAGGAGGTAGATTAGGAATTCATCGATA 3'	5'TATCGATGAATTCCTAATCTACCTCC TCAATGGTGGGGCCT 3'
pGEX4T1-calreticulin_AAMA	5'AGGACTGGGATGCCGCCATGGCCGGAGAGTGGGAA3'	5'TTCCCACTCTCCGGCCATGGCGGCATCCCAGTCCT3'
pGEX4T1-P Domain_AAMA	5'AGGACTGGGATGCCGCCATGGCCGGAGAGTGGGAA3'	5'TTCCCACTCTCCGGCCATGGCGGCATCCCAGTCCT3'
pCDNA 3.1-GRSF1_AAVA	5'TCCAAGTTAGAAGCGGCGGTGGCGGATGTCTTTCTCATTCGA3'	5'TCGAATGAGAAAGACATCCGCCACCGCCGTTCTAACTTGGA3'
pCMV 10-3X FLAG S6 ATPase_AAVA	5'ATGAACCTCTCTGCGGCGGT TGCGTTGGAAGACTATGT3'	5'ACATAGTCTTCCAACGCAACCGCCGCAGAGAGGTTTCAT3'

### **3.2.12 Confirmation of Positive Clones or Mutation**

The positive clone or the desired mutation was confirmed by sequencing (3500 Genetic Analyzer, Applied Biosystem).

### **3.2.13 Protein Expression**

Gankyrin were expressed and purified by using *Escherichia coli* BL21 (DE3) strain. Gankyrin R41A and K116A mutant were expressed and purified using *Escherichia coli* origami strain. A single, transformed, isolated colony of *E. coli* BL 21 (DE3) was inoculated in 10 ml LB medium and grown overnight at 37° C with vigorous shaking (200-250 rpm). Inoculum of 10 ml was made in 1 litre LB broth and allowed to reach 0.8- 0.9 O.D.<sub>600</sub> (Biophotometer, Eppendorf). Protein was induced with 100 µM isopropyl-D-thiogalactoside (IPTG) and growth was continued at 20 °C for 16 hours. Cells were lysed by sonication in lysis buffer (50 mM Tris (pH 7.5), 50 mM β-mercaptoethanol (BME), 500 mM NaCl, 10% glycerol, 0.1% Triton X-100) with protease inhibitor cocktail (Sigma). The culture was transferred into HS50 tubes (Tarson) and centrifuged at 15000 rpm for 30 min at 4 °C using SS-34 rotor in Sorvall RC5C Plus centrifuge. The supernatant containing soluble protein was used for further purification. Individual protein was purified by nickel-nitriloacetic acid (Ni-NTA) agarose affinity chromatography (Invitrogen).

### **Cell Density Measurement**

The O.D. was measured at 600 nm with LB medium as blank using with the help of Biophotometer (Eppendorf).

### **3.2.14 Ni-NTA Agarose Affinity Chromatography**

Ni-NTA agarose beads of 1-2 ml from Invitrogen was liquated in 1X 30 cm econo column (Bio-Rad). Beads were washed with 1X washing/equilibration buffer with at least two column volumes under native conditions. Equilibrated Ni-NTA beads were incubated with protein lysate at 4 °C for about 30 min. After incubation unbound lysate (flow through) were collected separately. Beads were washed with washing buffer with 2-3 column volume. 6X His tagged protein were eluted with elution buffer containing imidazole (500 mM).

### **3.2.15 Glutathione-S-Transferase purification affinity Chromatography**

For purification of GST fusion proteins, glutathione beads (GE, Healthcare) were used. 200 µl of beads were used for purification for 250 ml induced culture. Beads were equilibrated with 1X PBS buffer containing 10 mM DTT. Equilibrated beads were incubated with protein lysate at 4 °C for about 30 min. After incubation unbound lysate (flow through) were collected separately. Beads were washed thoroughly 6-7 times with 1X PBS buffer. Bound protein was eluted using 10mM Glutathione in 50mM Tris pH 8.0. For quantitative studies we have cleaved GST tag form the GST-fusion proteins. After the beads are washed in 1X PBS buffer, beads were incubated with thrombin for overnight at 4 °C. For 100 µl of beads, the 10 units of thrombin were used to cleave the GST tag from the GST-fusion protein.

### **3.2.16 Gel Filtration Chromatography**

For further purification of protein, it was subjected for gel filtration chromatography using sephadex G-75 beads (GE Healthcare Life Sciences). Initially the gel filtration column was equilibrated with gel filtration running buffer with the flow rate of 0.5

ml/min using HPLC system (Bio-Rad). 2 ml (3-4 mg) of total protein volume injected in the gel filtration column and eluted under native conditions. Peak fraction were collected either separately or pulled together, dialyzed against Tris buffer pH 7.5 and used for further experiments. All the proteins were stored at – 20 °C.

### **3.2.17 Protein Estimation using Bradford Assay**

BSA standards (1, 0.5, 0.25, and 0.125 mg/ml) were prepared from 30 mg/ml of BSA stock. Unknowns (protein samples) were taken in various dilutions. 5 µl of the standards and the unknowns were taken in duplicates in a 96-well plate and 200 µl of Bradford reagent (1:4 diluted, Bio-Rad) was added to each well. Readings were taken with ELISA plate reader (Spectra Max 790) at 595 nm using SoftMaxPro 4.6 software. Protein concentration for unknown was determined with the help of standard graph generated with BSA.

### **3.2.18 SDS PAGE**

Protein samples were boiled using digital dry bath (JENCON-PLS) at 100 °C with 1X Laemmli buffer for 10 min before loading. The samples were loaded to the gel placed in the tank containing 1X SDS-PAGE running buffer. The gel was resolved at 150 V for 1:30 hours. The gel was stained with 0.25% coomassie brilliant blue R (Sigma) for 15-30 min. The gel was then destained overnight in the destainer (50% Methanol, 10% Acetic acid) with 2-3 changes at regular interval. The gel was finally preserved in 10% acetic acid and documented.

### **3.2.19 Preparation of Glycerol Stocks**

100 µl of the overnight culture was transferred into autoclaved 1.7 ml microfuge tube (Axygen). To it 100 µl of 30% glycerol (autoclaved) was added to the tube and gently mixed. When the solution was homogenous glycerol stocks were stored at -80 °C.

### **3.2.20 GST pull down assay**

GST-fusion proteins and their mutants in NP-40 lysis buffer were allowed to immobilize on GST beads for 1 hour at 4 °C. Beads were washed 2-3 times with the same buffer containing 300 mM NaCl. His-gankyrin was allowed to bind for 2 hours at 4 °C. Beads were washed thoroughly (5-6 times) and boiled in the presence of Laemmli buffer. Proteins were resolved on a 12% SDS PAGE and western blotting was performed using gankyrin antibody.

### **3.2.21 Western blotting**

For western blotting samples were resolved on 12% SDS PAGE. Proteins were transferred on polyvinylidene difluoride (PVDF) membrane (Hybond, GE Healthcare). PVDF membrane was blocked with 3% BSA in TBST at room temperature for 1 hour on rocker. PVDF membrane was incubated with 1:1000 dilution of anti-gankyrin (Sigma) or 1:4000 dilution of anti-flag (Sigma) or 1:500 dilution of anti-Hsp70 (Abcam) or 1:1000 dilution of anti-CLIC1 (Abcam) or 1:500 dilution of anti-DDAH1 (sigma) or 1:1000 dilution of anti-GRSF1 (Sigma) or 1:1000 dilution of anti-EIF4A3 (Sigma) or 1:1000 dilution of anti-Hsp90 (Santacruz) or 1:500 dilution of anti-MAP2K1 (Sigma) or 1:1000 dilution of anti-HA (Abcam) or 1:1000 dilution of anti-Calreticulin (Abcam) antibody (rabbit polyclonal, Santa Cruz Biotechnology) for 1 hour at room temperature on rocker. Antibodies were diluted in TBST (TBST - 50 mM Tris, 150 mM NaCl, 0.1% Tween 20) containing 3% BSA.

After 1 hour, primary antibody was removed and membrane was washed with TBST at least four times, 15 min each time at room temperature on rocker. Membrane was incubated with corresponding secondary antibody (1:5000 dilution; Sigma) at room temperature for 1 hour on rocker. Secondary antibody was removed and membrane was washed with TBST at least four times, 15 min each time on rocker at room temperature. Membrane was incubated with ECL plus reagent (GE Healthcare) was exposed to X-ray film (Kodak) and was developed using automated developer machine (Optimax 2010, Protec GmbH & Co.). Developed X-rays was analyzed for the corresponding western blotting.

### **3.2.22 Routine Maintenance of Cell Lines**

All glassware and plastic-ware used for tissue culture work were sterile. For maintenance and experimental use, all adherent cells (HEK 293, gankyrin over expressing stable HEK 293, MDA-MB-231, MDA-MB-435) were trypsinized and passaged as follows. Spent medium was aspirated out using a pasteur pipette and the cells in the plate were washed twice with 1X PBS. 1X trypsin was added to the cells, and was removed after the cells rounded up but just before cells start detaching. To inhibit the trypsin activity 1ml of complete medium was added to the cells. The cells were collected in 1X PBS and the cell suspension was transferred to the centrifuge tube and tightly corked. The cell suspension was centrifuged for approximately 3-4 min at 1000rpm in REMI bench top centrifuge. The supernatant was discarded and the cell pellet was loosened by tapping the tube gently. The cells were suspended in an appropriate volume of complete medium, cell count was taken using a haemocytometer, and the required cell number was seeded in tissue culture dishes, and incubated at 37 °C in a humidified 5% CO<sub>2</sub> incubator. The cultures were passaged

at around 70-80% confluency or were frozen using 90% FBS in 10% DMSO when required. The cells were cultured up to 5-6 passages and fresh vial of frozen cells was revived at regular time intervals. Gankyrin over expressing stable HEK 293 cells and the vector control cells were maintained in 800 µg/ml of G418 (Sigma).

### **3.2.23 Freezing and Revival of Cell cultures**

**Freezing:** For freezing of cells, 80-90% confluent cultures were trypsinized as described above. The cell pellet was loosened by tapping the tube gently and the centrifuge tube was placed on ice for 1-2 min. Pre-chilled freezing medium (90% FBS+ 10% DMSO) was added drop wise to the cell pellet ( $\sim 1 \times 10^6$  cells/ml of freezing mixture) on ice, with constant shaking to ensure even cell suspension and transferred to pre-chilled vials. These vials were cooled gradually at 4 °C, -20 °C, -80 °C and then stored in liquid nitrogen.

**Revival:** To revive the frozen cells, a vial containing frozen cells was removed from liquid nitrogen and immediately thawed in a water-bath at 37 °C. As soon as the cell suspension thawed, it was transferred to a centrifuge tube containing 5 ml complete medium and centrifuged at 1000 rpm for 2 min. The supernatant was discarded, cell pellet was loosened by tapping the tube gently and re-suspended in an appropriate volume of complete medium and added to the 55 mm of tissue culture plate. The medium was replaced after the cells had adhered to the tissue culture dish on next day of revival.

### **3.2.24 Transfection**

#### **3.2.24.1 Transfection of plasmid DNA in HEK 293 cells**

HEK 293 cells were transfected with pCMV10-3Xp3XFLAG-CMV<sup>TM</sup>-10-gankyrin or pCDNA<sup>TM</sup>3.1 (+) or vector alone using calcium phosphate method.

One day prior to transfection, the cells were trypsinized and  $5 \times 10^5$  cells were seeded in a 55 mm plate and  $5 \times 10^4$  cells were seeded for 6 well plate. The plate should be around 70-80% confluent at the time of transfection. Next day, 4 hr before transfection, medium was replaced with fresh complete medium. The cells were transfected using 12  $\mu$ g plasmid DNA (55 mm plate) and 6  $\mu$ g plasmid DNA (6 well plate) using BBS. A total of 12  $\mu$ g plasmid was diluted to 100  $\mu$ l in autoclaved Milli-Q in a sterile tube. 100  $\mu$ l of CaCl<sub>2</sub> was added drop wise to the DNA in an autoclaved eppendorf. Then, 200  $\mu$ l of 2X BBS was added dropwise and mixed gently by pipetting 3-4 times. The mix was incubated at RT for 20 min. After 20 min, the DNA complexes were mixed gently again and added drop wise over the cells, mixed gently by swirling the medium in the plate and incubated at 37 °C in CO<sub>2</sub> incubator for 16 hrs. After 16 hrs, the transfection medium was replaced with fresh complete medium.

#### **3.2.24.2 Transfection of siRNA in HEK 293 cells**

One day prior to transfection, the cells were trypsinized and  $5 \times 10^4$  cells were seeded in a 6 well plate. The plate should be around 50-60% confluent at the time of transfection. Next day, 4 hr before transfection, medium was replaced with fresh complete DMEM medium. Lipofectamine 2000 (Invitrogen) was used for transfection. In one eppendorf, 5  $\mu$ l of 20  $\mu$ M siRNA was diluted in 50  $\mu$ l of plane DMEM and in another eppendorf 3  $\mu$ l of lipofectamine was diluted in 50  $\mu$ l of plane DMEM. They were incubated for 5 min at RT. After 5 min, the mix from both the eppendorfs were mixed. The mix was mixed gently using pipetting and this 100  $\mu$ l of mix was incubated for 20 mins at RT. After 20 min, 400  $\mu$ l of plane DMEM was

added to the 100 µl of complex mixture and mixed gently by pipetting. The complete medium was removed from the cells completely and the transfection complex of 500 µl was added drop wise on to the cells. After 10 hrs, the medium was changed with the complete medium.

### **3.2.24.3 Transfection of plasmid or siRNA in MDA-MB-231 and MDA-MB-435 cells**

One day prior to transfection, the cells were trypsinized and  $5 \times 10^4$  cells were seeded in a 6 well plate. The plate should be around 50-60% confluent at the time of transfection. Next day, 4 hr before transfection, medium was replaced with fresh complete DMEM medium. Lipofectamine 2000 (Invitrogen) was used for transfection. In one eppendorf, 5 µl of 20 µM siRNA or 3 µg of plasmid was diluted in 50 µl of plane DMEM and in another eppendorf 3 µl of lipofectamine or 9 µl of lipofectamine 200 respectively was diluted in 50 µl of plane DMEM. They were incubated for 5 min at RT. After 5 min, the mix from both the eppendorfs were mixed. The mix was mixed gently using pipetting and this 100 µl of mix was incubated for 20 mins at RT. After 20 min, 400 µl of plane DMEM was added to the 100 µl of complex mixture and mixed gently by pipetting. The complete medium was removed from the cells completely and the transfection complex of 500 µl was added drop wise on to the cells. After 10 hrs, the medium was changed with the complete medium.

### **3.2.25 Generation of gankyrin over expressing stable HEK 293 cells**

HEK 293 cells were transfected with pCMV10-3Xp3XFLAG-CMV<sup>TM</sup>-10-gankyrin or vector alone using calcium phosphate method. Clonal transformants were selected in presence of 800 µg/ml G418 (Sigma).

### **3.2.26 Proliferation assay-MTT assay**

HEK 293 cells were transiently transfected with gankyrin or vector alone. 48h after transfection,  $5 \times 10^3$  cells were seeded in a 96 well plate and 20 µl of (3-[4,5-dimethylthiazol-2-yl]-2,5-diphenyltetrazolium bromide (MTT) (5mg/ml) (Sigma) was added. After 2 hours, 100µl of 10%SDS in 0.01 N HCl was added and absorbance was read at 550 nm.

### **3.2.27 Etoposide induced apoptosis assay- MTT assay**

Vector alone and gankyrin expressing stable HEK 293 clones were subjected to etoposide (Cipla) treatment.  $2 \times 10^3$  cells were seeded in 96 well and treated with 50µM etoposide. Cell viability was checked after 72 hours using MTT assay as described above.

### **3.2.28 NF-κB activation- Luciferase assay**

Stable clones of HEK 293 harboring pCMV 10-3X-FLAG or pCMV 10-3X-FLAG - gankyrin were transfected with 12 µg of ConA control or 3 kb enhancer ConA luciferase construct (a kind gift from Dr. Neil D. Perkins, UK). After 48 hours, cells were treated with 10ng/ml TNF-α (Invitrogen), lysed and luciferase assay was performed using Promega Luciferase Assay System in triplicates.

### **3.2.29 Soft agar assay**

Gankyrin overexpressing HEK 293 stable clones or vector alone clones obtained by the above method were compared for their ability to exhibit anchorage independent growth using soft agar colony formation assay.  $2 \times 10^3$  cells in each case were overlaid on a thin layer of agar (0.4% over 1% agarose). Colonies formed were counted after 7 days.

### **3.2.30 Affinity pull down assay**

Cells were harvested in NP-40 lysis buffer [20 mM Tris pH 7.5, 150 mM NaCl, 0.5% NP-40 and 1mM dithiothreitol (DTT) containing 1X protease inhibitor cocktail (Sigma)]. 1mg of total cell lysate was incubated with 10 $\mu$ l of M2 agarose anti-flag beads (Sigma) for 4 hours at 4°C. Beads were washed extensively with the NP-40 lysis buffer containing 300 mM NaCl were used for washes to increase the stringency for pull down experiments.

### **3.2.31 Immunoprecipitation assay**

For immunoprecipitation experiments, cells were harvested in NP-40 lysis buffer. 15 $\mu$ l of Sepharose G beads were incubated overnight with 3 $\mu$ g of respective antibody in same buffer. 1mg of precleared lysate was added to the antibody bound beads and incubated for 4 hours. These immune complexes were washed thoroughly for 4-5 times using NP-40 lysis buffer containing 300 mM NaCl. In some cases RIPA buffer (25mM Tris-HCl pH 7.5, 150mM NaCl, 1% NP-40, 1% sodium deoxycholate, 0.1% SDS) was used during washes to increase the stringency in immunoprecipitation experiments.

### **3.2.32 RNA extraction and cDNA synthesis**

Cell contains three types of RNA, Trizol based method extracts total RNA from the lysed cells, followed by amplification using only mRNA to synthesize cDNA using an enzyme- reverse transcriptase.  $1 \times 10^6$  cells were suspended in 1 ml of Trizol and the sample was either processed immediately or stored at  $-80^\circ\text{C}$  till further use. For RNA extraction the cells were thawed at RT and the cell pellet was dissolved completely by vortex mixing and repeated pipetting. 200  $\mu\text{l}$  of chloroform was added and the mixture was vortex mixed for 5 min, the mixture was kept on the bench top till two phases could be distinguished and then centrifuged for 10 min at  $12000 \times g / 4^\circ\text{C}$ . The aqueous phase was carefully transferred to a fresh tube without disturbing the interphase and the RNA was precipitated using 500  $\mu\text{l}$  isopropanol at RT/10 min and spun 20 min at  $12000 \times g / 4^\circ\text{C}$ . The isopropanol was gently removed and pellet was washed with 500  $\mu\text{l}$  75% ethanol; pellet was semi dried and dissolved in DEPC treated D/W (DEPC D/W) at  $55^\circ\text{C}$ ; quality and quantity of RNA was assessed by measuring O.D. 260/280. First strand cDNA synthesis was carried out using First-Strand cDNA synthesis kit (Invitrogen).

Components	Volume	Final Concentration
RNA		2 $\mu\text{g}$
10mM dNTP Mix	1 $\mu\text{l}$	1 mM
50 $\mu\text{M}$ oligo(dT)	1 $\mu\text{l}$	5 mM
DEPC-treated water	to 10 $\mu\text{l}$	

Incubate the tube at  $65^\circ\text{C}$  for 5 min, then place on ice for at least 1 min. Prepare the following cDNA synthesis mix, by adding each component in the indicated order.

Component	Volume
10X RT buffer	2 $\mu\text{l}$

25 mM MgCl <sub>2</sub>	4 µl
0.1 M DTT	2 µl
RNaseOUT (40 U/µl)	1 µl
SuperScript III RT (200U/µL)	1 µl

Add 10 µl of cDNA Synthesis Mix to each RNA/primer mixture, mix gently, and collect by brief centrifugation. Incubate for 50 min at 50°C and terminate the reactions at 85°C for 5 min. Chill on ice and collect the reactions by brief centrifugation. Add 1 µl of RNase H to each tube and incubate the tubes for 20 min at 37°C. cDNA synthesis reaction can be stored at -20°C to -10°C or used for PCR immediately.

### 3.2.33 Real Time PCR

2X SYBR Green Master Mix (Saf Labs) was used for real time pcr studies. cDNA obtained from the reverse transcription reaction was diluted to 10ng/µl with DEPC-treated Milli-Q water for use in the PCR reaction. A total of 10ng cDNA was used per reaction. PCR reaction was set up as follows:

Components	Volume per well
2X SYBR Green Master Mix	2.5 µl
Forward Primer 10 µM	0.5 µl
Reverse Primer 10 µM	0.5 µl
cDNA (10 ng/µl)	1.0 µl
Total volume with DEPC treated Milli-Q water	5.0 µl

Non-template PCR reaction was also set up which do not have template cDNA to ensure there is no primer-primer dimer formation. The 5 µl PCR reactions were

loaded into the 384-well microtiter optical plate. The plate was covered with an optical cover sheet and sealed with the help of a plastic applicator. The applicator was pressed evenly over the optical cover sheet several times to ensure proper sealing of the wells. The sealed plate was then centrifuged briefly at 2000 rpm for 2 min to spin down the reactions to ensure no air bubbles are left. The plate was loaded in the Real Time PCR machine (Avant 7900HT Sequence Detection machine, Applied Biosystems, NY, USA) with default cycling parameters. The PCR cycling parameters are:

Sr. No.	Temperature	Time	No. of Cycles
1	50 °C	2 min	1
2	95 °C	10 min	1
3	95 °C	15 sec	40
4	60 °C	1 min	

After completion of 40 cycles, the dissociation curve step of the amplified products was performed for all the reactions. The amplification data was collected in real time by the machine and stored in the SDS 2.1 software. After the completion of run, the data was analyzed using SDS 2.1 analysis software (Applied Biosystems, NY, USA). HPRT1 was used as the house keeping gene. The expression of the gene of interest was quantified and expressed as Relative Quantity (RQ) by comparative Ct method (Ct= threshold cycle as automatically determined by the SDS 2.1 software). The RQ was calculated as  $[2^{-(\Delta\Delta Ct)}]$ , where  $\Delta Ct = Ct(\text{gene}) - Ct(\text{HPRT1})$  and  $\Delta\Delta Ct = \Delta Ct(\text{gene levels in stable clones}) - \Delta Ct(\text{gene levels in vector control cells})$ .

**Table 3.2 Primers used for Real time PCR**

QRT-PCR Primers	Forward Primer	Reverse Primer

Gankyrin	5'GCAGCTTCGAAAAACAGGCA 3'	5'GGATGTTTGTGGATGCTTTGT 3'
HPRT	5'GAACCAGGTTATGACCTTGA 3'	5'GAGATCATCTCCACCAATTA 3'
CLIC1	5'GAATTCAAACCCAGCACTC 3'	5'CAGCACTGGTTTCATCCACTT 3'

### 3.2.34 Peptide inhibition assay

30 µg of His-gankyrin was incubated with different concentrations (1 µM- 500 µM) of EEVD peptide. Peptides GRRF and GRRR were included as controls. Preincubated complex were added to 30 µg of purified GST-CLIC1 bound to glutathione beads and incubated for 2 hours at 4°C. Beads were washed thoroughly for 5-6 times with NP-40 lysis buffer and boiled in the presence of Laemmli buffer. Proteins were resolved on a 12% SDS PAGE and western blotting was performed using gankyrin antibody. Band intensities were measured using ImageJ software. To obtain an approximate Kd value of the peptide, we made the following assumptions. We assume that band intensity at zero peptide represents total CLIC1 bound gankyrin. Band intensity in presence of varying concentrations of the peptide is then subtracted from this value. We consider this subtracted value to correspond to peptide bound fraction of gankyrin. These values are then plotted against peptide concentration to obtain ~ Kd of the peptide.

### 3.2.35 Isothermal calorimetry studies to quantitate the interaction between CLIC1 and gankyrin

To check the binding affinity between CLIC1 WT, CLIC1\_AAVA and gankyrin interaction isothermal calorimetry experiments were performed using MicoCal iTC200 (GE, Healthcare). CLIC1 Wt, CLIC1\_AAVA, His-gankyrin were purified and dialyzed in 10 mM phosphate buffer pH 7.5 to the final concentration of 100 µM. Before performing experiments, cell and syringe were washed properly. CLIC1 WT

or CLIC1\_AAVA of 50  $\mu\text{M}$  concentration were added to the syringe. His-gankyrin was added inside the cell to the final concentration of 5  $\mu\text{M}$ . The reaction was carried at 4  $^{\circ}\text{C}$ , 20 injections (2  $\mu\text{l}$ /injection) with the spacing of 200 sec. Appropriate blank experiments were performed with 10  $\mu\text{M}$  phosphate buffer pH 7.5 buffer inside the syringe and 5  $\mu\text{M}$  of His-gankyrin inside the cell. During data analysis blank experiment was subtracted from the main reaction to correct for dilution effects of the buffer in the protein. Analysis was performed using Origene software and data was fit using one site model.

### **3.2.36 ELISA**

200 ng of CLIC1 Wt was diluted in sodium carbonate buffer pH 9.0 and coated on maxisorp ELISA plates (NUNC) for overnight at 4  $^{\circ}\text{C}$ . Next day, wells were washed with TBST buffer for 3-4 times to remove the unbound CLIC1 Wt protein. Protein coated wells were blocked with 100  $\mu\text{l}$  of 2% BSA in TBST for 1 hr at 37  $^{\circ}\text{C}$ . Wells were washed with TBST buffer for 3- 4 times. After blocking, cells were incubated with His-gankyrin Wt, gankyrin K116A, gankyrin R41A of varying concentration from 100  $\mu\text{M}$  to 0.01  $\mu\text{M}$  for 1 hr at 37  $^{\circ}\text{C}$ . Wells were washed with TBST buffer for 3-4 times. Wells were incubated with gankyrin antibody (1:1000 dilution, sigma) for 1 hr at 37  $^{\circ}\text{C}$ . After washes, the wells were incubated with anti-rabbit secondary antibody (GE, Healthcare) for 1 hr at 37  $^{\circ}\text{C}$ . After vigorous washing the reaction was detected using TMB substrate (Genei, Bangalore). After 20 min, reaction was stopped using 2M  $\text{H}_2\text{SO}_4$  and read using spectrophotometer at 450 nm.

### 3.2.37 Migration using wound healing assay

Directional cell migration was studied using an in vitro wound healing assay (Rodriguez et al., 2005). Cells were seeded in 6 well plates and allowed to reach 70-90% confluence. To monitor the independent role of gankyrin or CLIC1, MDA-MB-231 cells were transfected with 100 nM gankyrin siRNA or CLIC1 siRNA or control siRNA and incubated for 72 hours before setting up the wound healing assay. To examine the effect of protein-protein interaction (PPI) on the motility of the cells, HEK 293 and MDA-MB-231 cells were transfected with CLIC1 Wt or CLIC1\_AAVA mutant or vector control and incubated for 48 hours. These experiments were conducted in the presence of endogenous CLIC1. Parental HEK 293 cells however do not show detectable endogenous CLIC1. To further confirm, the importance of PPI in this functional assay, endogenous CLIC1 in MDA-MB-231 cells were silenced using smartpool of UTR specific siRNA (Dharmacon) and co transfected with either CLIC1 Wt or CLIC1\_AAVA mutant. Control experiments included non-target siRNA with or without the cDNA for CLIC1 wt or AAVA mutant as the case may be. For transfection in MDA-MB-231 cells, 3 µg DNA and 9µl of lipofectamine 2000 was used. These transfected cells were grown to 100% confluence and were treated with 10 µg/ml of mitomycin C for 3 hours. Cell monolayer was then wounded with a plastic tip, washed with PBS to remove cell debris. Wound healing capacity was monitored for 16 or 24 hours using phase contrast microscopy (Axiovert 200M, Zeiss, Germany). Area under each wound was calculated using Image J software. Percentage of wound healed was calculated using the formula  $[(\text{Initial wound area}-\text{Final wound area})/\text{Initial wound area}] \times 100$ .

**Table 3.3 The list of siRNA sequences.**

	<b>Target Sequence (5'-3')</b>
ON-TARGET plus SMARTpool siRNA Gankyrin	GGGUGUGUGUCUAACCUAA
	GUUCUACUGUUGUCGUAUA
	GAGUAUUCUGGCCGAUAAA
	AGGCCUACGCCAAACGUUU
ON-TARGET plus SMARTpool siRNA CLIC1	GGAGAUCGAGCUCGCCUAU
	CAUCGGUACUUGAGCAAUG
	GGCAAAGGCCCUCAAUAA
	GGACCGAGACAGUGCAGAA
ON-TARGET plus SMARTpool 3' AND 5' UTR siRNA CLIC1	GAGCUUGUGUUGUGCUGAAUU
	AGAGAGUCCCUGAGUGUGAUU
	GAAGAGGGGAUGAGGGAAAUU
	GGACAACAUAUUUCAGUAAUU
ON-TARGET plus Non-targeting Pool	UGGUUUACAUGUCGACUAA
	UGGUUUACAUGUUGUGUGA
	UGGUUUACAUGUUUCUGA
	UGGUUUACAUGUUUCCUA

### **3.3.38 Invasion using Boyden chamber assay**

Invasion assay was performed using 24 well transwell culture inserts of 8µm pore size (BD Biosciences). Matrigel at a concentration 300 µg/ml diluted in 1X DMEM was added to the inner side of the chamber and allowed to polymerize at 37°C for 1 hour. Unpolymerized matrigel was removed and  $5 \times 10^4$  cells were seeded in 200 µl of DMEM. To the lower chamber complete medium (DMEM+10% FBS) was added as chemoattractant. Cells were then incubated for 24 hours at 37°C in the humidified incubator with 5% CO<sub>2</sub>. After incubation, inserts are removed and the non-invading cells on the upper surface of the insert were scraped. Lower surface of the insert

which harbors the invaded cells was fixed using cold methanol and stained using crystal violet. All images were taken at 5X magnification using an upright microscope (Axio Imager. Z1, Zeiss, Germany).

## **CHAPTER 4.**

### ***Generation and Characterization of gankyrin overexpressing stable HEK 293 cells***

## **4.1 Introduction:**

Gankyrin is an oncoprotein and well known to be overexpressed in many epithelial cancers (Fu et al., 2002; Higashitsuji et al., 2000a; Li et al., 2011; Man et al., 2010a; Meng et al., 2010a; Tang et al., 2010a). Recent reports show its prominent role in other neuroectodermal tumors like glioma (Yang et al., 2012). Gankyrin overexpression in NIH3T3 (mouse) cells results in malignant transformation and transfected cells can form colonies on soft agar and seed tumors in nude mice (Higashitsuji et al., 2000b). While the literature provides ample evidence for the oncogenic properties of gankyrin, the exact mechanism and its role in the characteristic cancer hall mark properties which are crucial for disease intervention remain unclear. Moreover, key oncoproteins act as hubs in protein-protein interaction network. While gankyrin seems to act like one by deregulating signaling pathways, the scope and vastness of gankyrin interactome is unclear. In order to better define the oncogenic properties of gankyrin and identify its hub properties, we developed a HEK 293 cellular model in which gankyrin expression can be regulated. Using such a model system where gankyrin is transexpressed will allow us to establish a causal relationship to gankyrin and attribute the changes in the network to overexpression of gankyrin. This chapter describes how we established the stable clones and characterized the oncogenic properties of gankyrin.

### **Establishing stable cell lines**

There are different protocols used for establishment of stable cell lines, depending upon the type of expression (inducible/constitutive) and the presence of selection marker eg. G418, puromycin, hygromycin B in the plasmid construct that is incorporated. Each cell line has the different sensitivity towards different selection

marker. Hence, it is necessary to determine the optimal concentration of drug required for the selection of stable clones. Generally, the concentration of drug selected is such that it gives the massive death in 3-4 days and kills all the untransfected cells in 2 weeks (Rosato, 2007). Several methods are used for the generation of a stable cell line, depending on the scope of the experiment. A mixed population of antibiotic resistant cells can be used directly for experimental analysis (batch culture) with the advantage of generating fast results. This method however suffers from the disadvantage of undefined and genetically mixed cell population. This can be overcome by generating clonal cells. This can be achieved by diluting antibiotic resistant cells in such a way that culture as single, isolated cells is achieved e.g., by plating in 6-well plates or other methods. Subsequently, the selection process is applied to the single cell cultures. The procedure of single cell cloning may be repeated several times to obtain 100% clonal purity. In contrast to transient expression, stable expression allows long term, as well as defined and reproducible, expression of the gene of interest. We overexpressed gankyrin by transexpression in p3XFLAG-CMV<sup>TM</sup>-10. Stable clones were selected as described below.

### **Characterization of gankyrin overexpressing cells**

Gankyrin overexpressing clones were characterized using proliferation assay, NF- $\kappa$ B assay (luciferase assay), soft agar assay and resistance to apoptosis.

**Proliferation assay:** Cellular proliferation was monitored using 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT assay). The MTT assay is a colorimetric assay for assessing cell viability (Van Meerloo et al., 2011). It is a very simple and non-hazardous method of observing for cellular proliferation. MTT can be metabolized by all cells and hence can be used by all cell types. The NAD(P)H-dependent cellular oxidoreductase enzymes reflects the number of viable

cells present. These enzymes are capable of reducing the tetrazolium dye MTT to its insoluble formazan, which has a purple color. Other assays which can be used for checking cellular proliferation are [<sup>3</sup>H]-Tdr assay, BrdU incorporation assay (Crane and Bhattacharya, 2013) and cell labeling using CFSE (carboxyfluorescein diacetate succinimidyl ester) (Fulcher and Wong, 1999).

### **NF-κB activity:**

NF-κB (Nuclear Factor-KappaB, NF-κB) is a heterodimeric protein composed of different combinations of members of the Rel family of transcription factors. The NF-κB /Rel family of transcription factors (p50, p65, c-Rel, etc.) are involved in stress, immune, and inflammatory responses. It's a key protein modulating the apoptotic response. Phosphorylation of IκB leads to its degradation, whereas free NF-κB complex translocates to the nucleus, binds to NF-κB DNA response elements, and induce the transcription of the target genes (Higashitsuji et al., 2007a).

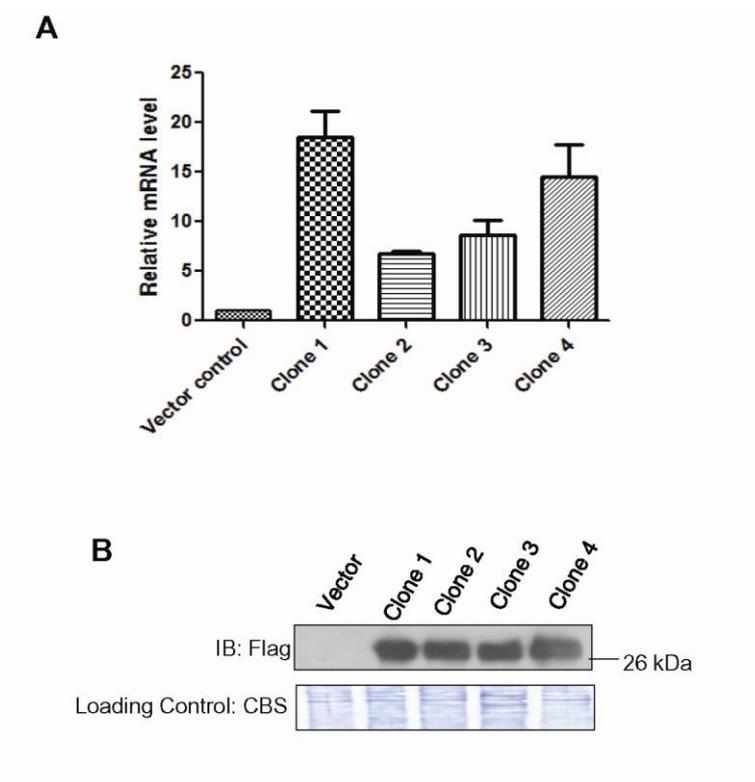
NF-κB activity was monitored using luciferase reporter assay. Firefly and Renilla luciferase reporters are widely used to monitor transcriptional activities in cell biology as they are highly sensitive, flexible, and easily quantified. In dual-luciferase reporter assays, both luciferase reporter enzymes (the firefly luciferase and the Renilla luciferase) are expressed simultaneously in a cell. The firefly luciferase reporter is the experimental reporter and denotes the effect of specifically designed experimental conditions whereas the constitutively expressing Renilla luciferase acts as an internal control which normalizes for unwanted variability caused by well-to-well or plate-to-plate differences in cytotoxicity, transfection efficiency, technical variability, and off-target effects (Naylor, 1999).

**Soft agar assay:** Cellular transformation occurs via a series of genetic and epigenetic alterations that yield a cell population that is capable of proliferating independently of both external and internal signals that normally control growth. Anchorage-independent growth is one of the hallmarks of cell transformation, which is considered the most accurate and stringent *in vitro* assay for detecting malignant transformation of cells (Hanahan and Weinberg, 2011b). The soft agar colony formation assay is a common method to monitor anchorage-independent growth, which measures proliferation in a semisolid culture media after 3-4 weeks.

**Resistance to apoptosis:** Tumorigenic cells are known to be resistant to apoptosis (Hanahan and Weinberg, 2011b). To test, if gankyrin overexpressing cells we treated cells with etoposide, an apoptosis inducing agent. Etoposide is a cytotoxic agent belonging to category of topoisomerase inhibitor. Etoposide forms a ternary complex with DNA and the topoisomerase II enzyme (which aids in DNA unwinding), prevents re-ligation of the DNA strands, and by doing so causes DNA strands to break. Therefore, this causes errors in DNA synthesis and promotes apoptosis of the cell (Van Maanen et al., 1988).

#### **4.2 Results and Discussion:**

Gankyrin was overexpressed as a flag tag fusion and stable colonies were selected using G418 as the selection marker. Expression levels of gankyrin in the stable clones were tested using qRT-PCR. Total gankyrin (using gankyrin specific primers) is expressed at 5 to 15 fold higher levels in stable clones as compared to the vector alone cells (Fig. 4.1A). Western blot using anti-flag antibody shows specific expression of flag-gankyrin in transfected cells (Fig. 4.1B).

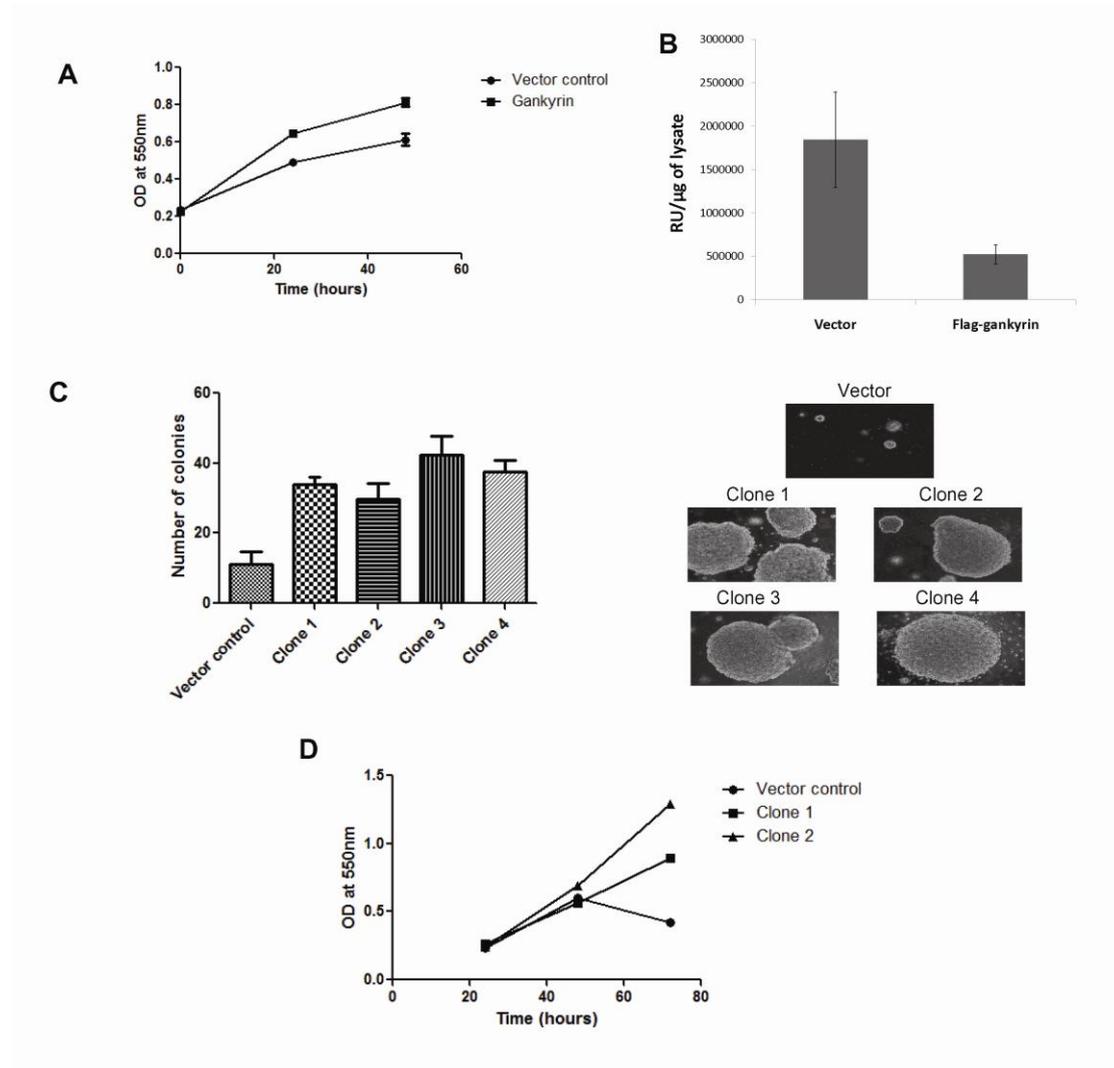


**Figure 4. 1**

**Generation of gankyrin overexpressing stable HEK 293 clones.** (A) HEK 293 stable clones overexpressing gankyrin were generated. Total levels of gankyrin were quantitated using gankyrin specific primers and real time PCR. HPRT was used as the internal control. Data represent mean  $\pm$ SD of one experiment done in duplicate; (B) Western blot using anti-flag antibody shows overexpression of transfected gankyrin in stable clones. Coomassie brilliant blue stained bots (CBS) indicates equal loading.

Characterization of gankyrin overexpressing cells were performed systematically. Initially, some of the reported phenotypic characteristics seen in gankyrin overexpressing cells, like increase in proliferation (Kim et al., 2009) and reduction in NF- $\kappa$ B activity (Higashitsuji et al., 2007b) were recapitulated. Gankyrin over expressing HEK 293 cells showed increase in cellular proliferation after 48 hours

(Fig. 4.2A). Gankyrin is known to interact with RelA and suppress NF- $\kappa$ B activity. Decrease in NF- $\kappa$ B activity upon gankyrin overexpression was also recapitulated as reported (Fig. 4.2B). The ability of gankyrin overexpressing cells to grow on soft agar was tested. This anchorage independent growth assay is considered as a surrogate assay for *in vivo* tumorigenesis. The stable clones or vector control cells were allowed to grow on soft agar for 7 days and colonies formed were counted manually. Gankyrin overexpression in HEK 293 cells resulted in enhanced growth of colonies on soft agar (Fig. 4.2C). Number and size of colonies were much larger upon gankyrin overexpression, additional malignant features are only introduced in these transformed HEK 293 cells in response to gankyrin overexpression. Gankyrin overexpressing cells and vector control cells were subjected to etoposide treatment for different time intervals and the cell viability was checked using MTT assay. Gankyrin overexpressing cells showed resistant to death whereas vector control cells were less viable after 72 hours (Fig. 4.2D).



**Figure 4. 2**

**Characterization of stable HEK 293 cells over expressing gankyrin.** (A) HEK 293 cells over expressing gankyrin show increased cell proliferation as compared to vector alone transfected HEK 293 cells. Cell proliferation was checked using MTT assay after 72 hrs. Data represents mean  $\pm$  SD of one experiment done in triplicates; (B) NF- $\kappa$ B activation is inhibited in stable clones over expressing gankyrin as compared to vector alone cells after TNF- $\alpha$  treatment. Data represent mean  $\pm$  SD of one experiment done in triplicates; (C) HEK 293 stable clones (four clones) show anchorage independent growth as compared to the vector alone cells;  $2 \times 10^3$  cells were seeded and number of colonies formed after 7 days are plotted. The data represents mean  $\pm$ SD of one experiment done in triplicate. Three independent experiments were performed in duplicate to confirm these observations. A

*representative 10x image of one of the fields from each plate is also shown; (D) HEK 293 vector alone and gankyrin over expressing cells were treated with 50  $\mu$ M etoposide and cell viability was checked after 72 hrs. Clones overexpressing gankyrin were resistant to apoptosis as compared to the vector alone cells. Data represent mean  $\pm$ SD of one experiment done in duplicate.*

### **4.3 Summary**

The HEK 293 stable clones simulate oncogenic properties of gankyrin like resistant to apoptosis and enhancement of growth in colonies on soft agar. Thus the functions of gankyrin specific interacting partners can be attributed to these cells.

## **CHAPTER 5.**

### ***Bioinformatic identification of novel interacting partners of gankyrin***

## 5.1 Introduction

Many approaches have been used in the past to identify novel interacting partners. Few highthroughput experimental approaches widely used are mentioned in the table 2.1. Many computational approaches have also been put forward where protein-protein interactions are predicted based on experimentally generated data stored in databases (Pagel et al., 2005; Stark et al., 2006; Von Mering et al., 2005). For example kinases and phosphatases interact with other proteins through motifs or domains which show sequence conservation or well defined properties (Patil et al., 2010b; Taylor et al., 2009). In such cases identification starts with scanning for such domains, motifs in the primary sequence of proteins under study. There are many proteins which do not have any known motifs or domains but are capable of multiple interactions through some important regions present in the interface region. These regions are short, linear segments are 3-10 residues long and play crucial role in protein interactions, post translational modifications and protein trafficking (Edwards et al., 2007; Lam et al., 2010). There are many algorithms available which scans for SLiMs: 1) **SLiMDisc** (Davey et al., 2006) server predicts such regions depending upon the common ancestry in sharing motifs by concentrating on convergently evolved motifs which are based on common attributes like subcellular localization or shared interacting partner. 2) **Interactions of Eukaryotic Linear Motif (iELM)** (Weatheritt et al., 2012) server helps in predicting the function and positional interface for the subset of interactions mediated by short linear motifs. 3) **Scansite 2.0** (Obenauer et al., 2003) web server identifies short protein sequence motifs that help in predicting cell signaling interactions. They are represented as a position specific scoring matrix (PSSM) based on results from oriented peptide library and phage display experiments. 4) **DILIMOT** (Neduva and Russell, 2006) web server predicts

short linear motifs if given the set of sequences sharing a common functional feature like interacting partner or the localization.

Other bioinformatic analysis indicate that sequences of about 3-10 residues are signatures of unique biological function (Jacob and Unger, 2007). Short linear motifs primarily in disordered regions of the proteins, terminal residues and three amino acid motifs can also assign functions to proteins (Dror and Ivett, 2007; Edwards et al., 2007). These motifs have been used to identify novel substrates of kinases harboring SH2-SH3 domains and proteins like 14-3-3 which recognize such short sequences in phosphorylated proteins (Edwards et al., 2007).

Studies in our lab suggest that judicious combination of functionally important short linear sequences with filters based on structure and other biological properties can help in the identification of truly positive interactions in proteome wide screening. PNSAS - 'Prediction of Natural Substrates from Artificial Substrate of Proteases', an algorithm developed by us for predicting natural substrates of endo proteases uses such principles (Venkatraman et al., 2009). Power of the method lies in the use of short sequence motifs coupled with physiologically relevant filters namely, accessibility of the cleavage site in the folded protein and subcellular co-localization to reduce false positive identification (Venkatraman et al., 2009). We have validated this algorithm by identifying Dsg-2, a desmosomal protein involved in cell-cell junction as a physiologically relevant novel substrate of a transmembrane protease called matriptase. Ability of matriptase to regulate levels of Dsg-2 at cell surface is likely to play a crucial role in cell invasion and metastasis (Wadhawan et al., 2012). In addition, we recently demonstrated that a 13 residue peptide based on the sequence of apomyoglobin can inhibit interaction of the full length protein with the proteasome (Singh Gautam et al., 2012).

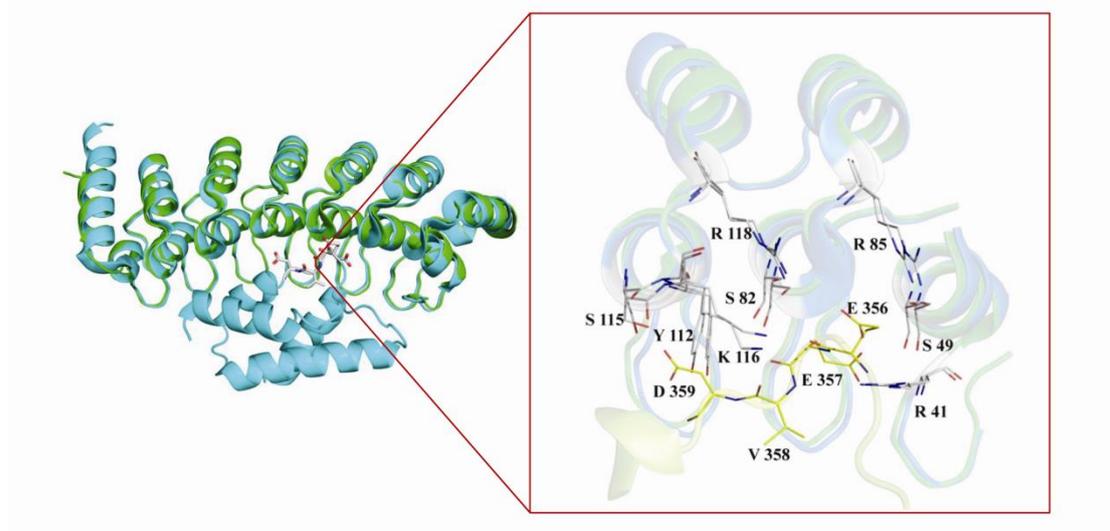
Yet another property seen with protein-protein interactions is the presence of what are called the hot spot sites. When two proteins interact with each other, they share a large interface area, but only few key residues contribute to the maximum binding energy. These are called as the hot spot residues (Bogan and Thorn, 1998). Hot spot sites can also be defined as the functional residues that mediate substrate binding, transition-state stabilization or product release, i.e. the residues located in the active site. Often these hot spot residues are also functionally conserved (Lichtarge et al., 1996). Identification of such hot spot residues however remains a major challenge (DeLano, 2002). Again some attempts have been made using Hot spot wizards which helps in identification of functionally conserved residues by integrating the knowledge of databases and the computational tools (Pavelka et al., 2009).

Here, using structural bioinformatics, we exploit the concept of hot spot sites to identify putative interacting partners of gankyrin. Our strategy lies in the recognition of a hot spot site at the interface in a known complex of gankyrin-S6 ATPase and predicting proteins that would share the hot spot site (Short Linear Sequence Motif) present in an accessible region of the protein. Protein sequences from human proteome data was downloaded from the Uniprot website (URL:<ftp://ftp.ebi.ac.uk/pub/databases/uniprot>; March 21, 2012). This dataset contains a total of 81,194 sequences in FASTA format and has two parts, one containing manually annotated data (UniProtKB/Swiss-Prot) and the other a computationally analyzed data awaiting manual annotation (Uni-ProtKB/TrEMBL). Manually annotated 35,961 entries were used for further analysis. Proteins containing EEVD in their primary sequence were extracted using perl scripts. From this list, proteins with three-dimensional structures were short listed and their Uniprot IDs were submitted to the Protein Data Bank (PDB) and structures were downloaded.

Solvent Accessible Surface Area (SASA) of the EEVD was calculated in the context of the tetra peptide using Parameter optimized surfaces (POPS) stand alone application (Fraternali and Cavallo, 2002).

## 5.2 Results and Discussion

The crystal structural of mouse (m-gankyrin) in complex with human S6 ATPase subunit of the proteasome is solved by Nakamura et al. and available in protein database bank. We took advantage of the published crystal structure of mouse gankyrin (m-gankyrin) in complex with human S6 ATPase subunit of the proteasome (Nakamura et al., 2007a) to predict novel interacting partners of human gankyrin. The complex buries 2418 Å<sup>2</sup> of surface area. The human (h) and m-gankyrin share 98% sequence identity and the crystal structure of unliganded h-gankyrin is completely super imposable on the m-gankyrin-S6 ATPase structure (Fig. 5.1).

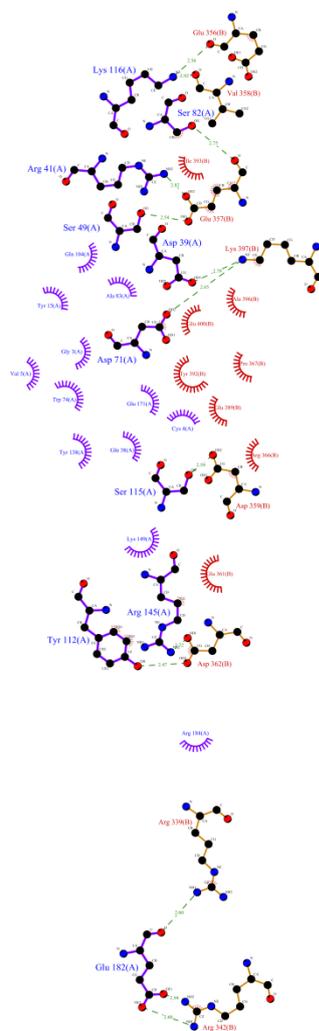


**Figure 5.1**

*Crystal structure of human gankyrin (1UOH) was aligned with mouse gankyrin-human S6 ATPase (2DVW) complex using Pymol. Residues from human gankyrin*

*superpose on the interacting residues of the mouse complex. Image was generated using PyMol (The PyMOL Molecular Graphics System, Version 1.5.0.1 Schrödinger, LLC).*

In this structure, the interface is enriched in polar and charged residues. A linear sequence of four amino acids EEVD (aa 356-359) in S6 ATPase (Fig. 5.1) makes dominant contribution to the interaction. Residue E356 is hydrogen bonded to Lys116. E357 forms a salt bridge with Arg41 and is also hydrogen bonded to Ser49 and Ser82. Val358 main chain carbonyl group is engaged in a hydrogen bond with Lys116. Asp359 is in polar contact with Ser115. Bridging this contiguous stretch of contact are two arginine residues Arg339 and Arg342 at one end and Lys397 at the other and these interact with residues in gankyrin through hydrogen bond and salt bridges (Fig. 5.2).



**Key**

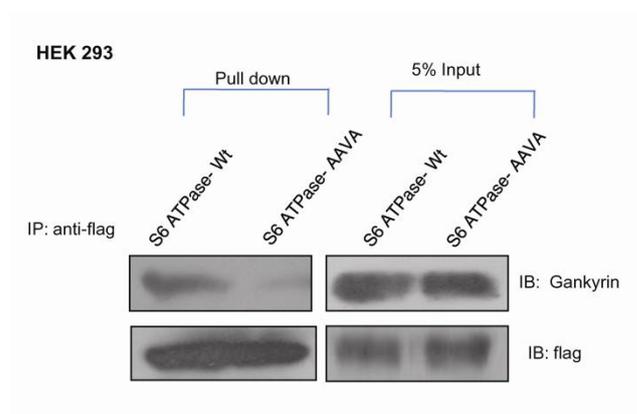
- ● Residues of first surface
- ● Residues of second surface
- ● Hydrogen bond and its length
- Residues involved in hydrophobic contact(s)
- Corresponding atoms involved in hydrophobic contact(s)

**Figure 5.2**

***Ligplot showing 2-Dimensional view of the interaction between gankyrin and S6 ATPase.***

Complex formation was abolished by any one of the double mutations of E356/E357A or D359/362A, whereas triple mutation of arginine residues,

R338/R339/R342A was required to abolish complex formation. Although these results have not been proven in direct interaction assays using purified components, these observations indicate that residues E356, E357 and D359 within EEVD are crucial for interaction. While the Arg residues are present in a helix, E356/E357/D359 residues are present in a loop and E362 is present in a helix. We tested the role of EEVD in complex formation between S6 ATPase and gankyrin by overexpressing Flag-S6 ATPase Wt and its corresponding AAVA mutant in HEK 293 cells followed by immunoprecipitation. Results indicate that the mutations significantly abrogate interaction (Fig. 5.3).



**Figure 5.3**

***Gankyrin interacts with S6 ATPase through E, E and D residues.*** *Lysates of HEK 293 cells expressing flag-S6 ATPase Wt or flag-S6 ATPase\_AAVA mutant were immobilized on M2 agarose anti-flag beads followed by immunoblotting with anti-flag and anti-gankyrin antibody. Gankyrin interacts with S6 ATPase Wt but not with the corresponding mutant.*

Based on these results and evidence cited above, we chose to test the short linear stretch of EEVD present in the loop as a probable hot spot site at this interface. We

hypothesized that other proteins within the human proteome which contain a contiguous stretch of EEVD may also interact with gankyrin. Short sequences such as EEVD are customarily believed to be promiscuous.

We decided to test whether proteins from the human proteome which harbor EEVD in the accessible region of the protein can interact with gankyrin. Ability of gankyrin to interact with multiple proteins would lead to a system wide network some of which may dictate cancer specific phenotypes. A total of 264 unique proteins with EEVD in their primary sequence were found and for 34 of them structural information was available using which solvent accessible surface area (SASA) values were calculated. For comparative analysis SASA values were normalized either against EEVD from S6 ATPase (considered as 1) or they were obtained from the POPS server. For 4 proteins the corresponding structure did not have any electron density for the peptide sequence and these were verified to be disordered (Table 5.1).

Uniprot ID	Protein Name	PDB ID	rSASA from reference: 2DVW	rSASA from pops program
NCK2_HUMAN	Cytoplasmic protein NCK2	4E6R	1.24	0.8
GRSF1_HUMAN	G-rich sequence factor 1	2LMI	1.18	0.71
BPNT1_HUMAN	3'(2')5'-bisphosphate nucleotidase 1	2WEF	1.15	0.67
DNMT1_HUMAN	DNA (cytosine-5)-methyltransferase 1	3SWR	1.03	0.62
CLIC1_HUMAN	Chloride intracellular channel protein 1	3UVH	1.02	0.62
PRS6B_HUMAN	26S protease regulatory subunit 6B	2DVW	1	0.58
IF4A3_HUMAN	Eukaryotic initiation factor 4A-III	2J0S	0.97	0.56
RIR2B_HUMAN	Ribonucleoside-diphosphate reductase subunit M2 B	4DJN	0.92	0.54
DDAH1_HUMAN	N(G)N(G)-dimethylarginine dimethylaminohydrolase 1	3I2E	0.92	0.53
MP2K2_HUMAN	Dual specificity mitogen-activated protein kinase kinase 2	1S9I	0.91	0.54
MP2K1_HUMAN	Dual specificity mitogen-activated protein kinase kinase 1	3EQC	0.91	0.53
RIR2_HUMAN	Ribonucleoside-diphosphate reductase subunit M2	3OLJ	0.84	0.49
REV1_HUMAN	DNA repair protein REV1	3GQC	0.81	0.47
PPIE_HUMAN	Peptidyl-prolyl cis-trans isomerase E	3MDF	0.80	0.46
CALM_HUMAN	Calmodulin	2LL6	0.79	0.46
HLTF_HUMAN	Helicase-like transcription factor	2L1I	0.78	0.45
HECW2_HUMAN	E3 ubiquitin-protein ligase HECW2	2LFE	0.77	0.44
SYT1_HUMAN	Synaptotagmin-1	2R83	0.69	0.40
CALL3_HUMAN	Calmodulin-like protein 3	1GGZ	0.66	0.38
FBXL5_HUMAN	F-box/LRR-repeat protein 5	3U9J	0.66	0.38
DYN1_HUMAN	Dynamin-1	3SNH	0.65	0.38
SIR5_HUMAN	NAD-dependent protein deacylase sirtuin-5	3RIY	0.60	0.35
LMNA_HUMAN	Prelamin-A/C	1IVT	0.59	0.35
SCO1_HUMAN	Protein SCO1 homolog	2HRN	0.57	0.34
SETB1_HUMAN	Histone-lysine N-methyltransferase SETDB1	3DLM	0.54	0.32
ROA1_HUMAN	Heterogeneous nuclear ribonucleoprotein A1	1U1P	0.53	0.31
PZRN3_HUMAN	E3 ubiquitin-protein ligase PDZRN3	1WH1	0.52	0.31
SH3L2_HUMAN	SH3 domain-binding glutamic acid-rich-like protein 2	2CT6	0.42	0.25
OTC_HUMAN	Ornithine carbamoyltransferase	1OTH	0.42	0.25
AIMP1_HUMAN	Aminoacyl tRNA synthase complex-interacting multifunctional protein 1	1FL0	0.28	0.17
HS90A_HUMAN	Heat shock protein HSP 90-alpha	3Q6M	Disordered	
SLD5_HUMAN	DNA replication complex GINS protein SLD5	2E9X	Disordered	
HSP71_HUMAN	Heat shock 70 kDa protein 1A/1B	3LOF	Disordered	
TB22A_HUMAN	TBC1 domain family member 22A	2QFZ	Disordered	

**Table 5.1**

*Solvent accessible surface area values for EEVD containing proteins. rSASA in each sequence was calculated using the POPS program. Values were normalized (relative*

*SASA or rSASA) using the SASA values for EEVD in the gankyrin-S6 ATPase complex. The values calculated from POPS program are also tabulated. The value for S6 ATPase was obtained in its uncomplexed form.*

### **5.3 Summary:**

We predict novel putative interacting partners of gankyrin using a) structural details of gankyrin-S6 ATPase complex b) by recognizing a short linear sequence motif and proving that it is a hot spot site at the interface, c) by extrapolating that important hot spot sites are generally conserved and d) by using the principle that interaction sites are generally in an accessible region.

## **CHAPTER 6.**

### ***Choice of Proteins and experimental validation of putative interacting partners of gankyrin***

## 6.1 Introduction

Bioinformatics based screening identified 264 proteins which carry the motif EEVD in their primary sequence which is expected to serve as the docking site for gankyrin. Using solvent accessibility as criteria we short listed 34 proteins. For further experimentation, eight proteins we selected based on the following criteria: a) proteins involved in functions that can modulate oncogenesis and b) a stringent cut off for relative solvent accessibility (rSASA).

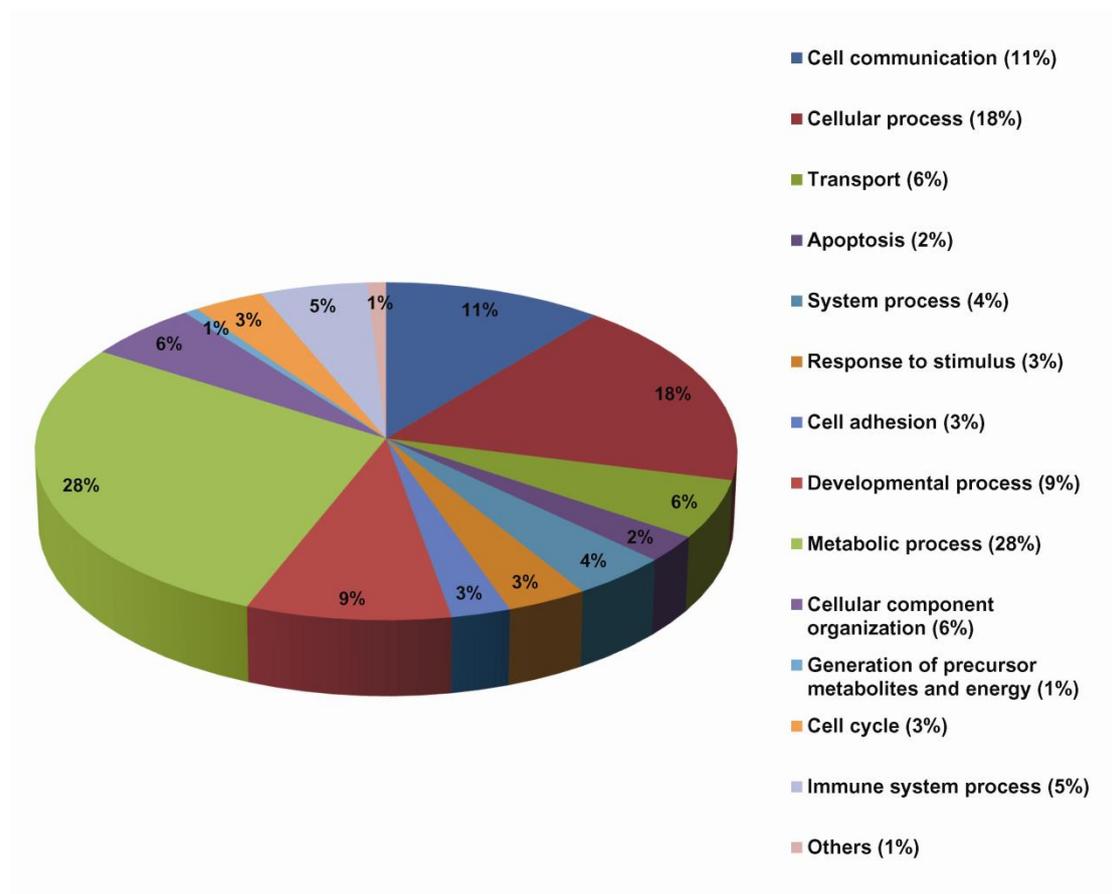
Protein-protein Interaction between two suspected proteins can be studied by any of the following methods. In vitro binding, cross-linking, **biochemical chromatography methods** (affinity column, co-purification, gel filtration chromatography), **direct immunological methods** (e.g. co-immunoprecipitation), or **indirect immunological methods** (e.g. immunostaining, immunolocalization, far western blots) (Berggård et al., 2007). We experimentally validated the predicted interactions using affinity pull down assay and immunoprecipitation studies.

## 6.2 Results and Discussion

### Choice of proteins

To select proteins for experimental validation, we used a stringent cut off for relative solvent accessibility (rSASA). Among the 34 proteins with available crystal structure, we chose proteins with rSASA >0.5 as determined by POPS server. Among the 11 proteins with rSASA >0.5, we chose to test NCK2, G-rich RNA sequence binding factor 1 (GRSF1), Chloride intracellular channel protein 1 (CLIC1), Eukaryotic initiation factor 4A-III (EIF4A3), dimethylarginine dimethylaminohydrolase 1 (DDAH1) and mitogen-activated protein kinase 1 (MAP2K1) as putative interacting partners. Two proteins (Heat shock protein 70-Hsp70, Heat shock protein-Hsp90) out

of four where EEVD was present in the disordered region were included in the study. All these proteins are expressed in the cytoplasm or both in cytoplasm and nucleus. Gankyrin is a nuclear cytoplasmic shuttling protein (Higashitsuji et al., 2007b). Hsp70 and Hsp90 proteins carrying a putative signature motif for gankyrin binding intrigued us and it seemed important to ask if native gankyrin would actually bind these chaperones which may directly influence their function by behaving as a common co-chaperone to Hsp70 and Hsp90. Interestingly, Hsp90 plays a role in the assembly of 26S Proteasome (Imai et al., 2003). Functions of these proteins were discerned from Panther data base (Thomas et al., 2003) and their association with cancer was verified using the classic eight hall mark properties of cancer (Hanahan and Weinberg, 2011a) . Figure 6.1 shows the functional classification of all 264 proteins containing EEVD in their primary sequence based upon the Panther database. Six out of eight proteins that we chose carried hall mark cancer properties (Table 6.1).



**Figure 6.1**

*Pie Chart showing functional classification of putative partners of gankyrin classified based on the panther database biological processes using PANTHER classification system.*

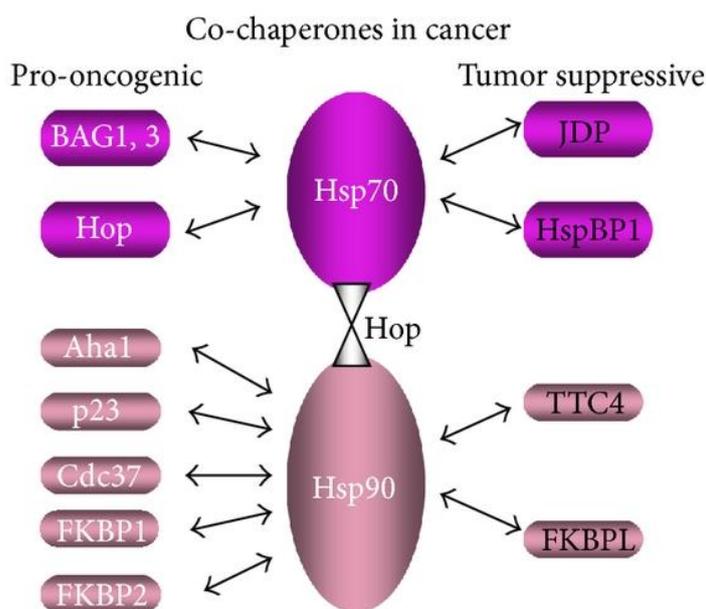
**Table 6.1**

**Proteins for which interactions have been validated in this study show hall mark of cancer properties.**

Uniprot ID	Protein Name	Panther protein class	Hall mark of cancer
NCK2_HUMAN	Cytoplasmic protein NCK2	Signaling molecule	Sustaining proliferative signalling
CLIC1_HUMAN	Chloride intracellular channel protein 1	Cytoskeletal protein	Activating invasion and Metastasis
MP2K1_HUMAN	Dual specificity mitogen-activated protein kinase kinase 1	Kinase, Transferase	Sustaining proliferative signaling
DDAH1_HUMAN	N(G)N(G)-dimethylarginine dimethylaminohydrolase 1	Hydrolase	Inducing Angiogenesis
HS90A_HUMAN	Heat shock protein HSP 90-alpha	Chaperone	Resisting Cell Death
HSP71_HUMAN	Heat shock 70 kDa protein 1A/1B	Chaperone	Resisting Cell Death

Hsp70 and Hsp90 are very well established for their role as chaperones. They play an essential role in conferring resistant to cell death thus leading in malignant transformation and in generation of antigenic peptides (Jolly and Morimoto, 2000). While their well characterized role seems to be in binding to unfolded proteins via degenerate but specific sequence motifs these chaperones also recognize native or near native proteins under unique circumstances (Carrello et al., 2004; Pang et al.,

2001). These chaperones also bind to kinases (Jolly and Morimoto, 2000). Heat shock proteins along with their co-chaperones assist in protein folding and make tumorigenic cells resistant to death. These co-chaperones are found to be over expressed in cancer. Hsp70 co-chaperones are Bag1, Bag3 and Hop whereas co-chaperones of Hsp90 are Aha1, p23, Cdc37 and FkBP1 (Calderwood, 2013). Some of their functions and the fate of assembly are schematically represented in figure. 6.2

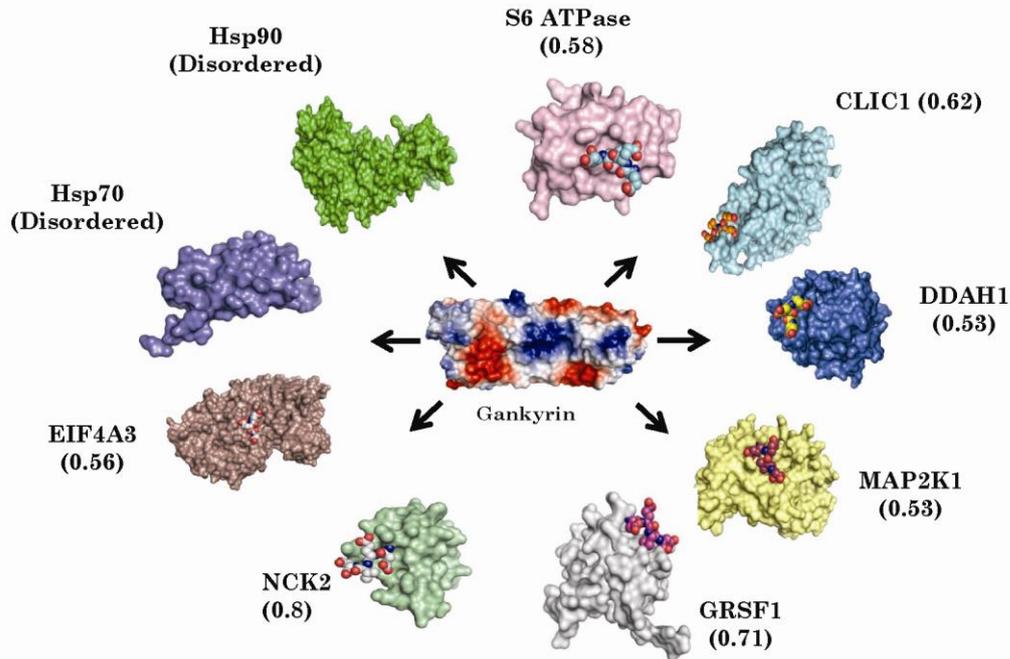


**Figure 6.2**

*Oncogenic or tumor suppressive influence of molecular co-chaperones. (Hop is shown twice in the cartoon to depict its role in bridging Hsp70 and Hsp90 as well as its pro-oncogenic influence) (Calderwood, 2013).*

These co-chaperones along with Hsp70 and Hsp90 are believed to be very potential druggable targets. Geldanamycin a known anti-cancer agent targets Hsp90 in cancer cells. CLIC1 a chloride intracellular channel protein is involved in metastasis and invasion (Chen et al., 2007b; Petrova et al., 2008b). In many cancer cells, voltage gated channels show this property and are considered attractive targets for cancer therapy (Le Guennec et al., 2007). NCK2 is a non-catalytic tyrosine kinase adaptor

protein, family of Src homology adaptor protein (Lehmann et al., 1990). They modulate many downstream events through their various SH domains. NCK2 is found to be overexpressed in many malignant conditions. In human metastatic melanoma cell line, NCK2 is shown to be involved in enhancing proliferation, migration and invasion (Labelle-Côté et al., 2011). NCK2 promotes the cellular apoptosis by mediating phosphorylation of p53 on serine 15 and hence activating the p53 which further causes cell cycle arrest and eventually apoptosis (Errington and Macara, 2013). NCK2 binds to growth factor receptors or their substrates and modulates gene expression in response to kras signaling (Braverman and Quilliam, 1999; Latreille and Larose, 2006). Gankyrin is a key regulator of Ras mediated activation and seems to specifically activate PI3K/Akt pathway. Evidence suggests that gankyrin is highly expressed in human lung cancer having kras mutations which results in Akt activation (Man et al., 2010b). MAP2K1 belongs to the family of dual specificity kinases which has the potential of phosphorylating serine/threonine kinases as well as tyrosine kinases (Tyr) (Dhanasekaran and Reddy, 1998). MAP2K1 phosphorylate ERK1 and ERK2 and the activation of these two proteins leads to the cellular proliferation. Thus, they are known to play a central role in regulation of cell proliferation, differentiation and apoptosis (Liu et al., 2004). DDAH1 is a positive regulator of angiogenesis (Kostourou et al., 2004). It is involved in vascular permeability and in NO synthesis. EIF4A3 is an abundant DEAD-box RNA helicase and a member of the eIF4A family of translation initiation factors (Chan et al., 2004). Presence of EEVD in a well accessible region and the functions of these proteins make them good candidates to test our hypothesis. Structure of these proteins with EEVD in the accessible region (where applicable) along with their rSASA values are shown in Figure 6.3.



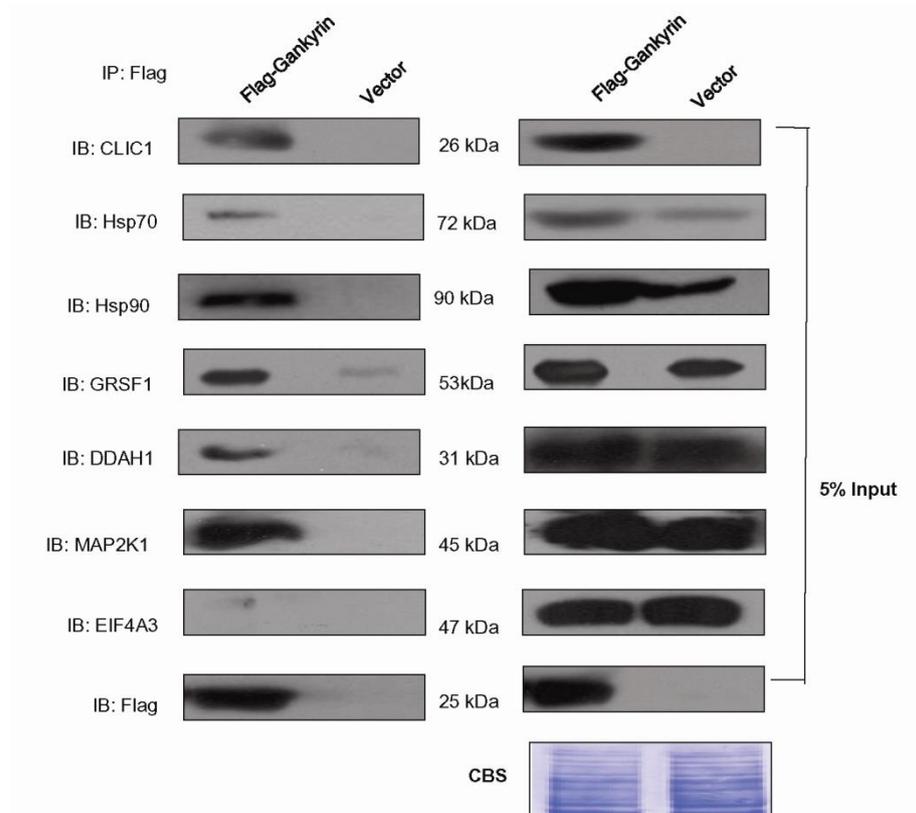
**Figure 6.3**

*Key feature of the network with gankyrin as the hub and the putative interacting partners harbouring EEVD hot spot site as potential nodes. Gankyrin (centre) is represented using surface electrostatic potential and interacting partners with rSASA value  $>0.5$  are represented as protein surfaces. ‘Hot spot’ site of interaction EEVD is represented in spheres. We believe that gankyrin is the hub protein and the putative interacting partners harbouring the hot spot site EEVD act as potential nodes to create a system wide network.*

### **Experimental validation of predicted interactions**

For experimental validation of predicted interactions we used affinity pull down of gankyrin expressed as a flag tagged fusion in HEK 293 cells. Six proteins Hsp70, Hsp90, CLIC1, GRSF1, DDAH1 and MAP2K1 were present in the pull down complex indicating that the predicted sequence EEVD may be involved in the interaction between each of these client proteins and gankyrin (Fig. 6.4). However

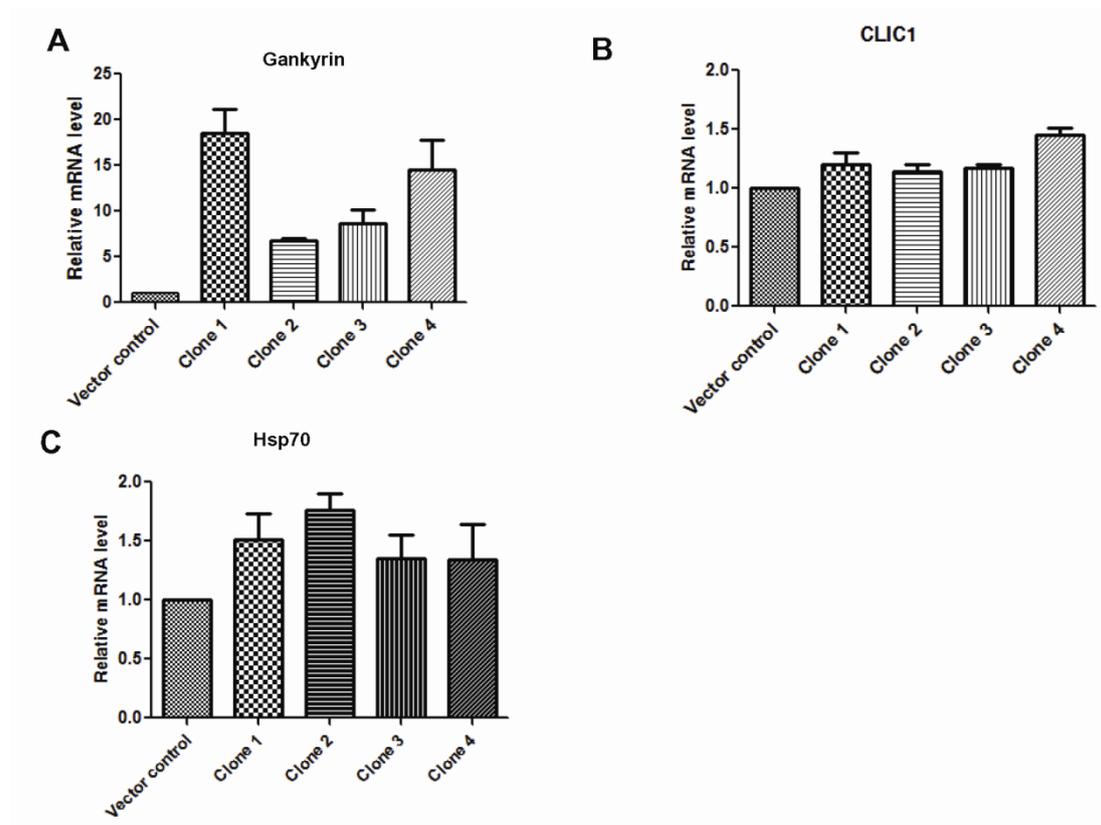
EIF4A3 was not found in the pull down complex. NCK2 antibody failed to detect protein from cell lysate even when used at lower dilutions than recommended. Owing to the important role of this protein in cancer we expressed this protein in mammalian cells and showed that interaction is mediated through EEVD (see below).



**Figure 6.4**

*Within the cellular milieu, gankyrin interacts with CLIC1, Hsp70, Hsp90, GRSF1, DDAH1, MAP2K1, and does not interact with EIF4A3. Lysate of HEK 293 cells expressing flag-gankyrin or flag alone (vector) were immobilized on M2 agarose anti-flag beads followed by immunoblotting with anti-Hsp70, anti-Hsp90, anti-GRSF1, anti-DDAH1, anti-CLIC1, anti-MAP2K1, anti-EIF4A3 antibody; 5% input represents loading control. CBS indicates equal loading. Note that some proteins are upregulated upon gankyrin overexpression.*

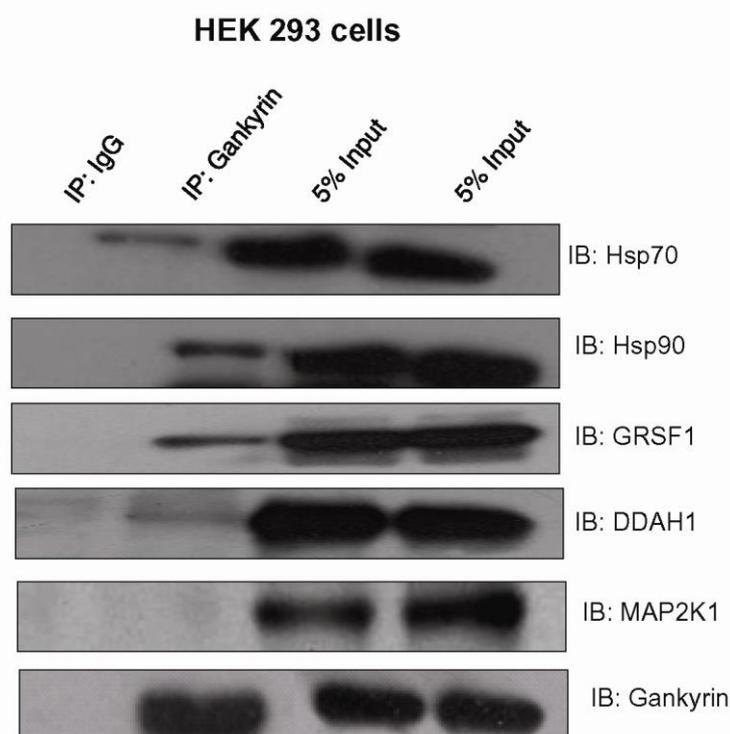
We observed higher levels of gankyrin and CLIC1. No detectable changes have been seen at the transcript levels using real time PCR. It may be due to 1) Gankyrin may bind to Hsp70 and CLIC1 and prevent it from proteasomal degradation. 2) Due to any changes at the protein synthesis. 3) Other post translational modification. (Fig 6.5).



**Figure 6.5**

**Transcripts levels of gankyrin, CLIC1, Hsp70 in gankyrin overexpressing HEK 293 clones.** Real Time PCR data showing (A) Levels of gankyrin in vector control HEK 293 cells and gankyrin over expressing stable HEK 293 clones. CLIC1 (B) and Hsp70 (C) mRNA levels are not altered in gankyrin over expressing HEK 293 cells.

Due to similar subcellular localization of gankyrin and the client proteins it is possible that some of these interactions may occur constitutively and at endogenous levels. It is also possible that some of these interactions may occur in response to gankyrin overexpression and therefore may be unique to malignant phenotype. To test these possibilities we used anti-gankyrin antibody immobilized on protein G Sepharose to pull down the endogenous complexes in HEK 293 cells. Three proteins (Hsp70, Hsp90 and GRSF1) were found in the immune complex while DDAH1 was barely detectable and MAP2K1 did not interact with gankyrin (Fig. 6.6).

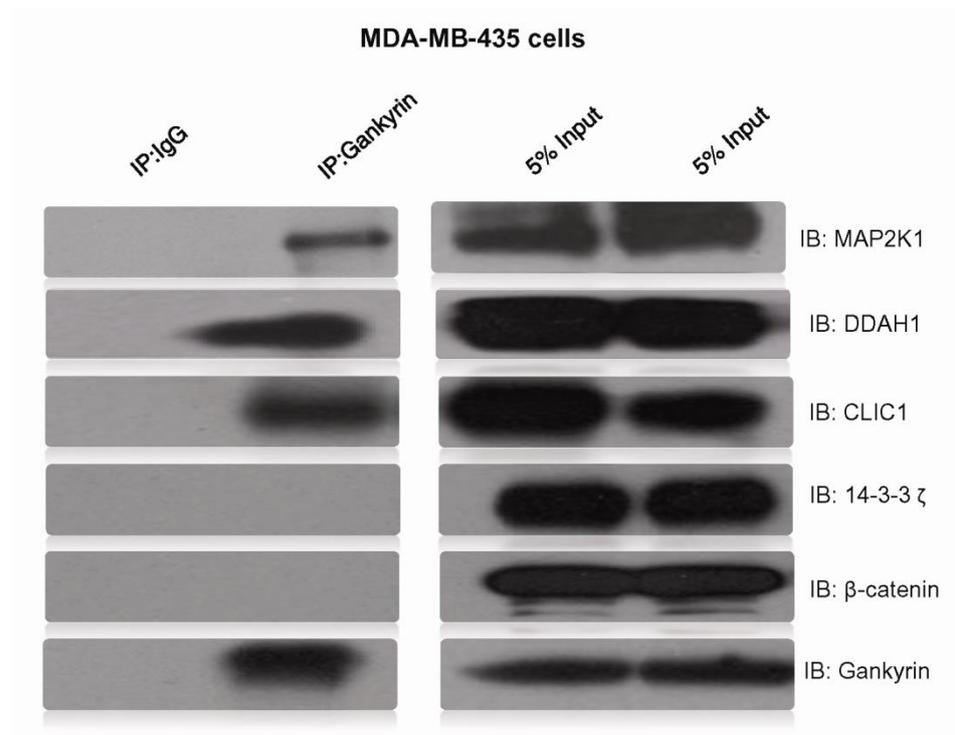


**Figure 6.6**

*Within the cellular milieu, gankyrin interacts with Hsp70, Hsp90, GRSF1. Lysate of HEK 293 cells were added to gankyrin bound and IgG bound Sepharose G beads followed by immunoblotting with anti-Hsp70, anti-Hsp90, anti-GRSF1, anti DDAH1, anti-MAP2K1 and anti-gankyrin antibody. MAP2K1 does not interact with gankyrin at endogenous levels.*

CLIC1 was hardly detectable in lysate of HEK 293 cells but was several fold overexpressed in HEK 293 stable clones (Figure 6.4 input). Levels of MAP2K1 and DDAH1 were not very different between vector control and gankyrin overexpressing HEK 293 cells. Enough evidence exists in literature to demonstrate the role of heat shock proteins in non chaperone function which may go beyond their ability to recognize misfolded proteins. Chaperones also interact with their co-chaperones through specific interactions between native proteins to execute their function (Jolly and Morimoto, 2000). It is interesting to note that the EEVD in Hsp70 and Hsp90 is involved in interaction with other proteins (Carrello et al., 2004).

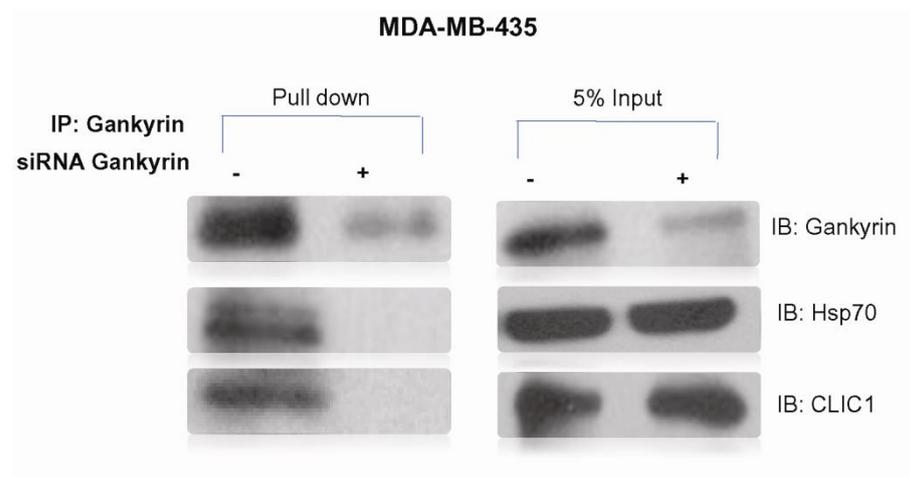
To test if interactions seen in response to gankyrin overexpression in HEK 293 also occur in cancer cells where gankyrin is known to play a role in the deregulated phenotype, we used a breast cancer cell line. Endogenous levels of gankyrin mRNA reportedly is less in normal and other tumor tissues as compared to breast cancer tissues (Zhen et al., 2012b). Therefore we used MDA-MB-435, a breast tumor derived cell line to check if the interactions observed upon gankyrin overexpression in HEK 293 cells are also seen in this cell line. MAP2K1, DDAH1, and CLIC1 were found to interact with gankyrin at endogenous levels in MDA-MB-435 cells (Fig. 6.7). Proteins which do not have EEVD like 14-3-3 $\zeta$  or  $\beta$ -catenin which carries a variant EEED do not interact with gankyrin under the same conditions (Fig. 6.7).



**Figure 6.7**

*Within the cellular milieu of MDA-MB-435 gankyrin interacts with CLIC1, DDAH1 and, MAP2K1. Lysate of MDA-MB-435 cells were added to gankyrin bound or IgG bound Sepharose G beads followed by immunoblotting with anti-MAP2K1, anti-DDAH1, anti-CLIC1, anti-14-3-3 zeta, anti- $\beta$ -catenin and anti-gankyrin antibody.*

Interestingly we found that  $\beta$ -catenin levels increase in HEK 293 cells upon gankyrin overexpression but this protein as in MDA-MB-435 cells does not interact with gankyrin (data not shown). Specificity of interaction between gankyrin and EEVD containing proteins CLIC1 and Hsp70 in MDA-MB-435 cells was further confirmed by repeating IP upon gankyrin knock down using specific siRNA (Fig. 6.8).



**Figure 6.8**

*Confirming the specificity of interaction of gankyrin with its EEVD containing interacting partners. Lysate of MDA-MB-435 cells expressing siRNA against gankyrin and MDA-MB-435 cells were added to gankyrin bound Sepharose G beads followed by immunoblotting with anti-Hsp70, anti-CLIC1 and anti-gankyrin antibody.*

Normal cellular homeostasis requires turnover of proteins and spatio-temporal regulation of protein interactions. By stabilizing existing networks or creating new networks gankyrin may induce deregulated phenotype that may be partly responsible for malignancy.

EIF4A3 carries EEVD sequence in the accessible region but does not show any interaction with gankyrin. Apart from sequence and solvent accessibility there are many factors such as sub cellular compartmentalization, post translational modifications, upstream signaling events that govern interaction. Hence, EIF4A3 protein not interacting with gankyrin can be due to other factors governing interaction.

### **6.3 Summary**

Gankyrin interacts with proteins containing EEVD in their solvent accessible region. As a proof of principle, we have demonstrated the interaction of gankyrin with Hsp70, Hsp90, NCK2, GRSF1, CLIC1, DDAH1 and MAP2K1. Gankyrin interaction with MAP2K1, CLIC1 and DDAH1 seems to be unique to cancerous cells or cells over expressing gankyrin and hence may have a functionally important role in altering phenotype. Other interactions with gankyrin might be important to maintain the physiological functions of gankyrin inside the normal cellular milieu which are as yet unclear. Their role in cancer if any is also unclear.

## **CHAPTER 7.**

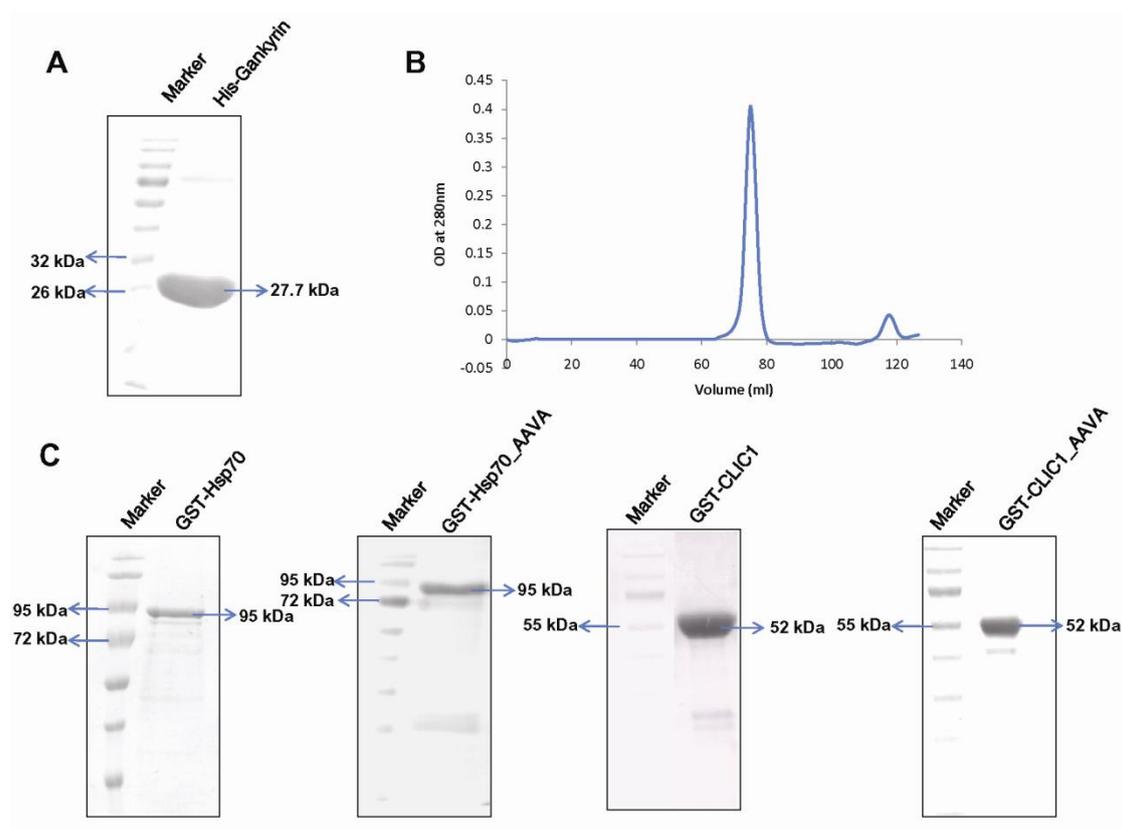
***Evidence for the EEVD as the hot spot site: Direct interaction between gankyrin and its client proteins***

## 7.1 Introduction

Interactome is defined as the complete map of protein-protein interactions that occur in a living organism. The critical step in exploring the complexity of the interactome is to map physical protein-protein interactions. Two proteins can be functionally linked but may not necessarily be directly interacting with each other. For eg., proteins in the ribosome or in the basal transcriptional apparatus which are part of a multimeric complex are physically linked to the neighboring subunits or proteins but only share ‘functional interactions’ with others. Proteins linked in the functional context can be identified using immunoprecipitation (IP), Co-IP or tandem affinity purification coupled to mass spectrometry (TAP-MS) (Berggård et al., 2007; Phizicky and Fields, 1995). Binary physical interactions can be mapped using yeast two hybrid assay, GST-pull down assay, far-western blotting. Biophysical techniques like Surface plasma resonance, Isothermal calorimetry titration, ELISA, Protein microarray can be used to identify the binary interactions quantitatively (Berggård et al., 2007; Espina et al., 2004; Pierce et al., 1999). When methods used to identify protein-protein interactions were compared, techniques which demonstrate physical association seem to be more faithful representatives of functionally relevant interactions. 90-95% interactions established by direct /indirect immune precipitation studies and 80% of *in vitro* biochemical studies correlate with the cellular role (Sprinzak et al., 2003). Direct interactions are further substantiated by identifying the few residues that are actually in contact across the interface. This is again a major challenge and crucial for building a valid PPI. Here we establish that the interactions seen in chapter 6 are mediated by the predicted EEVD motif and this forms the shared hot spot site at the interface of gankyrin interacting partners.

## 7.2 Results and Discussion

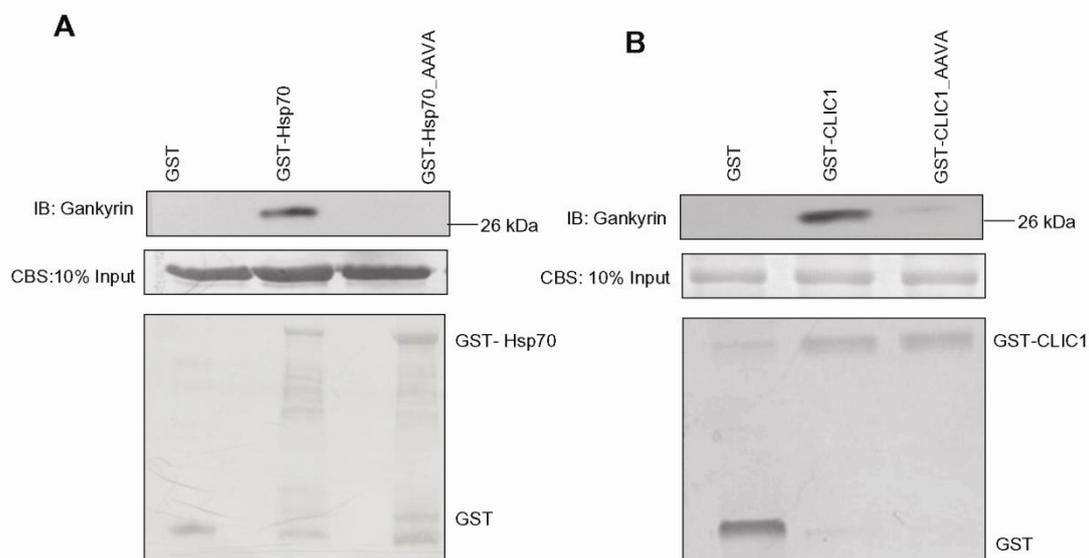
Affinity pull down assays with *in vitro* biochemical experiments was used to demonstrate direct interactions between gankyrin and the client proteins identified in this study. Recombinant form of gankyrin was expressed as a His-tagged fusion protein while the client proteins Hsp70 and CLIC1 were expressed as GST-fusions (Fig. 7.1).



**Figure 7.1**

**Purification of His-Gankyrin, GST-Hsp70, GST-Hsp70\_AAVA, GST-CLIC1, GST-CLIC1\_AAVA.** (A and B) SDS PAGE showing purified His-gankyrin at ~28kDa. Purified His-gankyrin was further purified using HPLC column which elutes as monomer at 76 ml of column volume. (C) SDS PAGE showing purified GST-Hsp70 (95 kDa), GST-Hsp70\_AAVA (95 kDa), GST-CLIC1 (52 kDa), GST-CLIC1\_AAVA (52 kDa).

Gankyrin was present only in fractions containing GST-fusion proteins and not in GST alone control (Fig. 7.2 A and B). To confirm that these direct interactions are mediated through residues E, E and D in each case, the sequence EEVD was mutated to AAVA and GST pull down assay was repeated. Mutations abolished interaction of each of these proteins with gankyrin (Fig. 7.2 A and B).

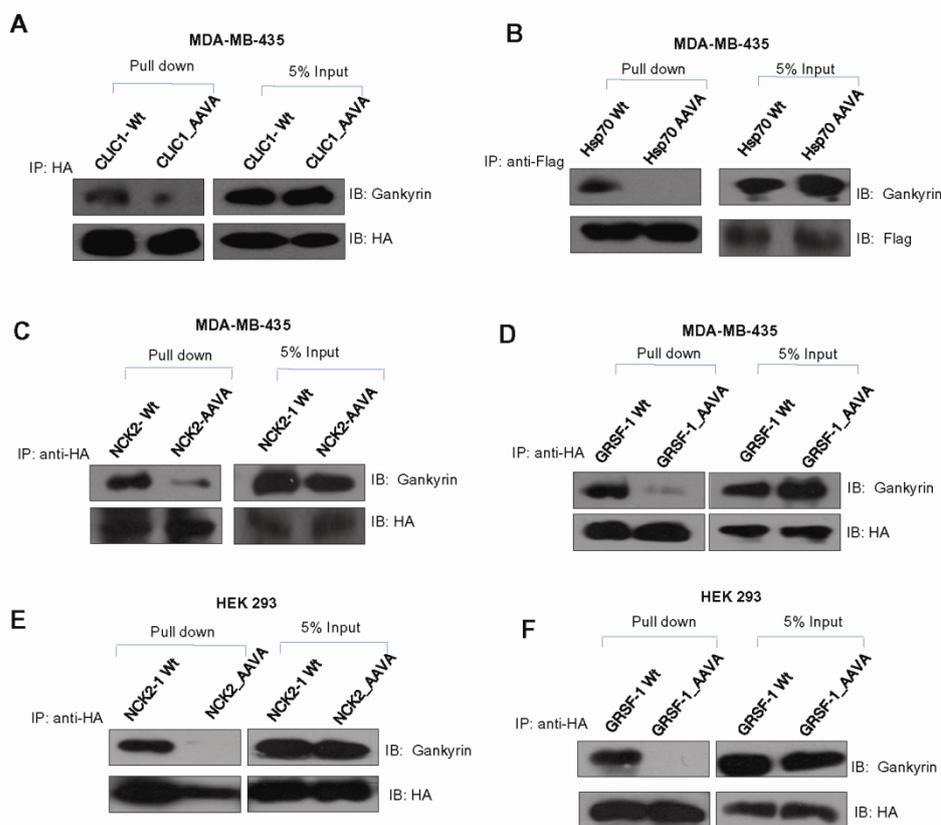


**Figure 7.2**

***Gankyrin binds to Hsp70, CLIC1 through E, E and D residues. (A and B) Purified GST-Hsp70, GST-CLIC1 and their respective AAVA mutants were independently immobilized on glutathione-Sepharose 4B and incubated with 6X His tagged gankyrin. Bound His-gankyrin was detected using anti-gankyrin antibody. Wt proteins interact with His-gankyrin while the mutants are unable to do so. GST alone does not interact with gankyrin. The input for gankyrin and GST-fusion proteins are shown.***

To verify whether the in vitro direct interaction between recombinant proteins is also sensitive to mutations within EEVD inside the cellular milieu, HA-CLIC1, Flag-Hsp70, HA-NCK2 and HA-GRSF1 and their corresponding AAVA mutant were

cloned and expressed in MDA-MB-435 cells. While all the Wt proteins showed definitive interaction with endogenous gankyrin (Fig. 7.3 A, B, C and D), the corresponding mutant proteins failed to interact. Interaction of endogenous gankyrin with Wt NCK2 and Wt GRSF1 and not the corresponding AAVA mutant were also detected in HEK 293 cells (Fig. 7.3 E and F). These results establish that observed interactions within the cells are direct and EEVD is the shared hot spot site at the interface.



**Figure 7.3**

***Gankyrin binds to Hsp70, CLIC1, NCK2 and GRSF1 through E, E and D residues.***

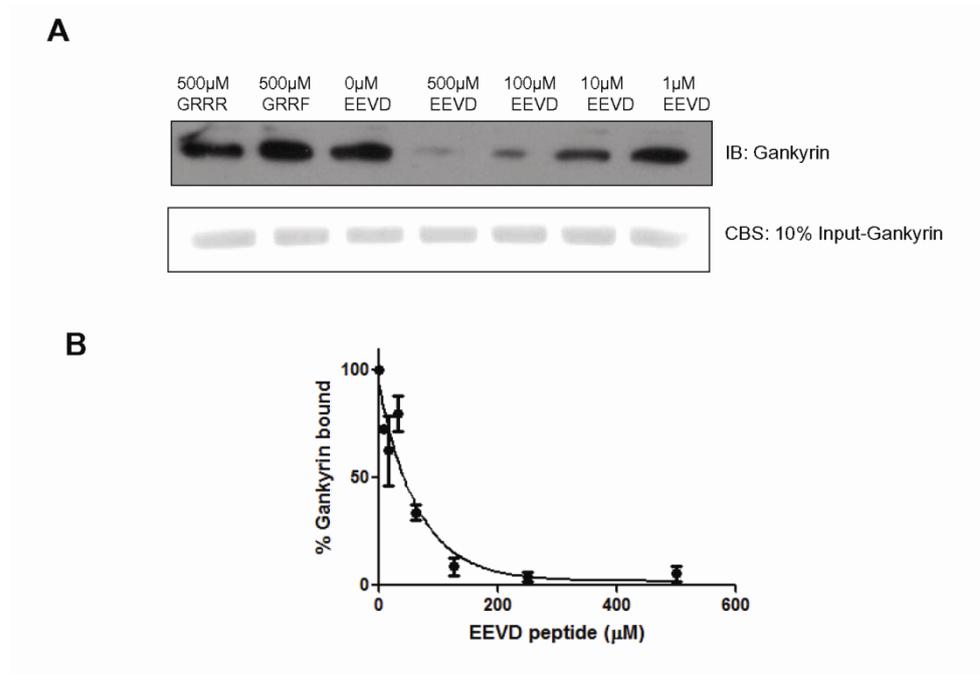
(A) Lysate of MDA-MB-435 cells over expressing HA-CLIC1 Wt or CLIC1\_AAVA;

(B) Flag-Hsp70 Wt or Hsp70\_AAVA; (C) HA-NCK2 Wt or NCK2\_AAVA; (D) HA-

GRSF1 Wt or GRSF1\_AAVA (E) Lysate of HEK 293 cells over expressing HA-NCK2

or NCK2. (F) HA-GRSF1 or GRSF1\_AAVA were independently added to anti-HA antibody bound to Sepharose G beads followed by immunoblotting with anti-gankyrin antibody. Gankyrin interacts with CLIC1 Wt, Hsp70 Wt, NCK2 Wt, GRSF1 Wt but not with the corresponding AAVA mutants.

To further illustrate our point, we tested the ability of naked short peptide sequence EEVD to inhibit binding of CLIC1 (used as a representative example) to gankyrin. In the *in vitro* pull down assay, peptide EEVD prevented complex formation between the two proteins in a concentration dependent manner (Fig. 7.4 A). Control peptides that do not interact with gankyrin but interact with another subunit of the proteasome (unpublished work) did not prevent interaction even at 0.5 mM, the highest concentration tested for EEVD. Using a semiquantitative approach, peptide concentration required for 50% inhibition in binding was estimated to be approximately 50  $\mu$ M (Fig. 7.4 B).



**Figure 7.4**

***Interface peptide EEVD inhibits interaction of full length protein with gankyrin.***

*(A) His-Gankyrin was incubated with different concentrations of EEVD peptide. Peptides GRRF and GRRR were used as controls. These preincubated complexes were added to the GST-CLIC1 bound GST beads. EEVD peptide and the two control peptides inhibited binding of His-Gankyrin to GST-CLIC1 in a concentration dependent manner. (B) Complex formation between gankyrin-CLIC1 was tested in presence of varying concentrations of peptide. Band intensities of CLIC1 bound gankyrin following IP and western blotting was quantitated using Image J. 50% inhibition in binding of two proteins was observed at 50  $\mu$ M peptide concentration.*

This value for a short peptide lies within the range reported for other peptides used to inhibit protein-protein interactions (Koch et al., 1993) (Table 7.1).

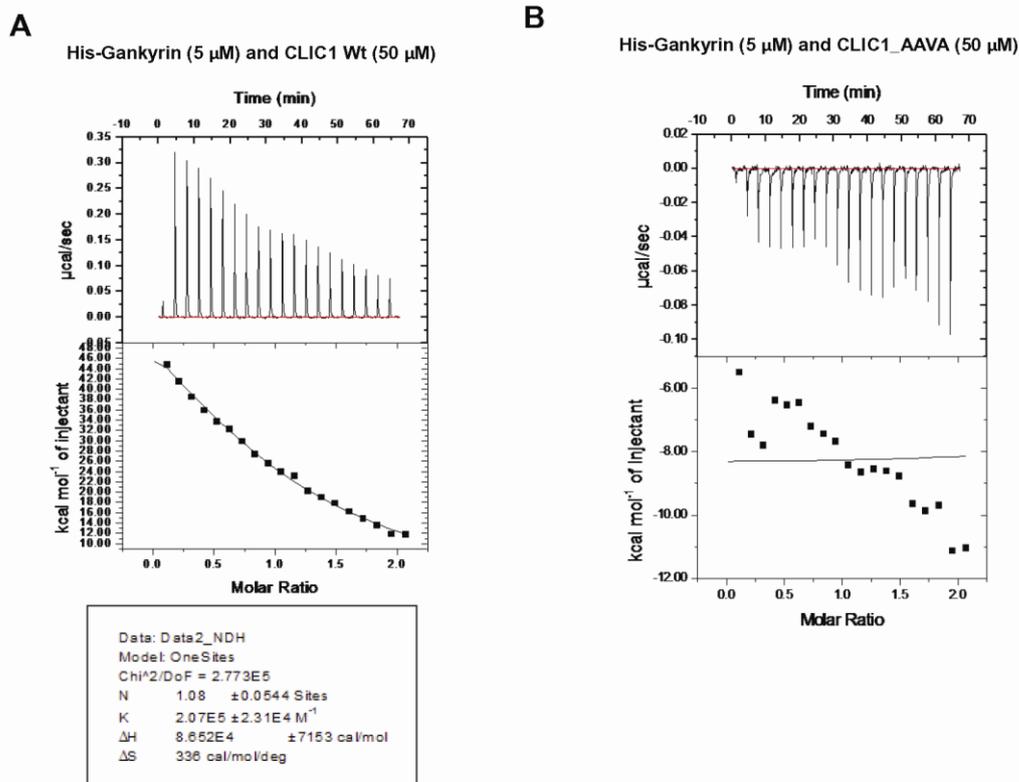
***Table 7.1***

***Summary of protein complexes with known interface residues and peptides or small molecular inhibitors designed to inhibit the respective complexes.***

Protein-protein complex	Interface/hot spot sites known	Peptide Inhibitors/ Small molecule inhibitors	Dissociation constant	Ref.*
hDM2-p53 known complex	p53 (activation domain)- F19, W23, L26	14-helical $\beta$ -peptide	368 and 583nM	Kritzer JA et., 2005
EphB4-EphrinB2 interaction	G-H loop from Ephrin B2	Computationally designing inhibitory peptide based on the protein-protein docking Benchmark 3.0 and CAPRI	-	London N et al., 2010
$\beta$ -Adrenergic Receptor Kinase-G $\beta$ $\gamma$	$\beta$ ARK2 - Gln546 to Ser670	28 amino acid peptide-(Trp643-Ser670)	76 $\mu$ M	Koch WJ et al., 1993
Annexin1-PKC PCKI-PKC RACKs- PKC	-	15 amino acid peptide derived from annexin1(aa332-aa346) inhibited the binding of PKC with both RACKs (30kDa and 33kDa)	90% inhibition at 10 $\mu$ M peptide 1	Mochly-Rosen D et al., 1991
RACK1-V5 region of $\beta$ II protein Kinase C	V5 $\beta$ II protein Kinase C	6 amino acid peptide derived from aa645-aa650 in betaIIPKC	Half maximal binding: 500nM	Stebbins & Mochly-Rosen. 2001
IL2-IL2R alpha chain	B alpha-helix in IL-2 (Lys-35, Arg-38, Phe-42, Lys-43)	acylphenylalanine derivatives, Ro26-4550, compound 1 are designed as a small molecular inhibitor	Compound 1: 8.2 $\mu$ M Ro26-4550: 19 $\mu$ M Acylphenylalanine derivatives Derivative 23: IC50-45 $\mu$ M, Derivative 24: IC50-3 $\mu$ M	Emerson SD et al., 2003 Sauve K et al., 1991 Zurawski SM et al., 1993 Arkin & Wells., 2003 Arkin & Wells 2004
Human growth hormone (hGH)-extracellular domain of its receptor (hGHbp)	hGHbp (W104, W169, I103, I105, P106, I165)	-	-	Clackson T et al., 1995
a. B7-CD28 b. B7-CTLA4	CD28 and CTLA4 binds to GFCC'C" face of the N-terminal V-set domain of human B7.1	dipyrazolo[3,4-b:3',4'-d]pyridin-3 series and other small molecule ligand for B7-1	CD28-B7.1 Compound 1: 60 $\pm$ 17 Compound 3: 4.0 $\pm$ 1.0 dipyrazolo[3,4-b:3',4'-d]pyridin-3 series: nanomolar to low micromolar range	Erbe DV et al., 2002 Green NJ et al., 2003
Bcl-x <sub>L</sub> and Bak peptide (aa72-aa87)	Bak (Leu78, Ile85, Ile81, Val 74 and Asp83) Bcl-x <sub>L</sub> (Tyr101, Leu108, Val126, Phe146)	Mutations in Bak peptide disrupts the ability of Bak to heterodimerize with Bcl-x <sub>L</sub>	0.34 $\pm$ 0.03 $\mu$ M	Sattler M et al., 1997
Human thymidylate synthase (hTS)- LR peptide	loops 188–194 of subunit A, the $\beta$ -strands 175–181 of subunit B, and loops 142–157 of both subunits	LR peptide (aa208-aa215)	45% percentage inhibition at 100 $\mu$ m peptide concentration.	Prasanna V et al., 1998 Cardinale D et al., 2011

\*- References are mentioned in detail in the main text

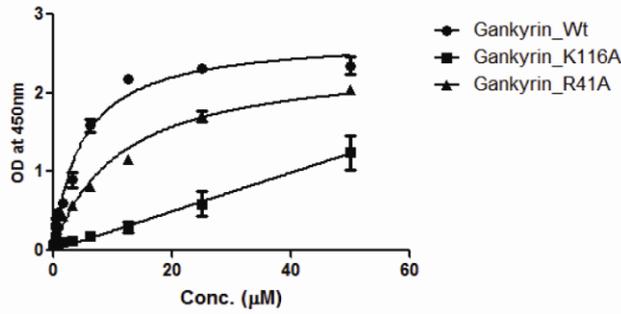
Isothermal calorimetry experiments were done to characterize the binding affinity between gankyrin and CLIC1\_Wt or CLIC1\_AAVA. His-gankyrin at 5  $\mu$ M was added to the cell and 50  $\mu$ M of CLIC1 Wt or CLIC1\_AAVA was added to the syringe as the ligand. The blank reaction was set as buffer (syringe) and was subtracted from the main reaction. His-gankyrin showed interaction with CLIC1 Wt with  $K_d$  of 5  $\mu$ M and stoichiometric ratio (N) of 1 whereas His-gankyrin did not show any interaction with the mutant CLIC1\_AAVA (Fig. 7.5).



**Figure 7.5**

*Isothermal calorimetry studies to study His-gankyrin interaction with CLIC1 Wt or CLIC1\_AAVA. (A) His-gankyrin of 5  $\mu\text{M}$  shows the interaction with 50  $\mu\text{M}$  of CLIC1 Wt with the dissociation constant of 5  $\mu\text{M}$  whereas (B) His-gankyrin did not show any interaction with CLIC1\_AAVA at the similar concentration.*

Gankyrin R41A and R41A/K116A showed the reduction in complex formation with S6 ATPase Wt (Nakamura et al., 2007b). We tested the ability of gankyrin mutants (R41A or K116A) to bind to CLIC1 using ELISA. ELISA experiments showed that gankyrin Wt interacts with CLIC1 with  $K_d$  of 4.6  $\mu\text{M}$  which is approximately similar as ITC experiments. Gankyrin\_K116A drastically reduced its ability of binding to CLIC1 whereas R41A showed approximately two fold decrease in binding affinity (Fig. 7.6).



	Gankyrin_Wt	Gankyrin_K116A	Gankyrin_R41A
Dissociation constant ( $K_d$ )	4.6 $\mu$ M	Ambiguous	10.76 $\mu$ M

**Figure 7.6**

*ELISA to study the interaction of gankyrin\_Wt and its mutant, gankyrin\_K116A or gankyrin\_R41A with CLIC1. Gankyrin Wt shows the interaction with CLIC1 with the dissociation constant of 4.6  $\mu$ M whereas Gankyrin\_K116A shows drastic reduction in interaction and gankyrin\_R41A shows the 2.5 fold reduction in binding (10.76  $\mu$ M).*

These results provide unequivocal evidence for the presence of a hot spot motif shared by proteins in the gankyrin hub network.

### 7.3 Summary

Here, we report that gankyrin interacts directly with Hsp70, CLIC1 through E, E and D residues under *in vitro* conditions. In MDA-MB-435 cells and HEK 293 cells CLIC1, Hsp70, NCK2, GRSF1 interacts with gankyrin through E, E and D residues. We also use a semiquantitative approach which shows that EEVD peptide concentration required for 50% inhibition of CLIC1 binding is approximately 50  $\mu$ M. Isothermal calorimetry studies further proved that gankyrin interacts with CLIC1 Wt with the  $K_d$  of 5  $\mu$ M whereas CLIC1\_AAVA mutant failed to interact. Gankyrin

mutant K116A showed the reduced affinity to CLIC1 Wt whereas gankyrin\_K41A mutant showed two fold reduction in binding indicating that CLIC1 binds to gankyrin in the same pocket where S6 ATPase binds to gankyrin. This proves that 'EEVD' sequence is the hot spot site and is a common conserved recognition motif among gankyrin interacting proteins.

## **CHAPTER 8.**

***Interaction between Gankyrin and  
CLIC1 is important for cell migration  
and invasion in MDA-MB-231 and HEK  
293 cells***

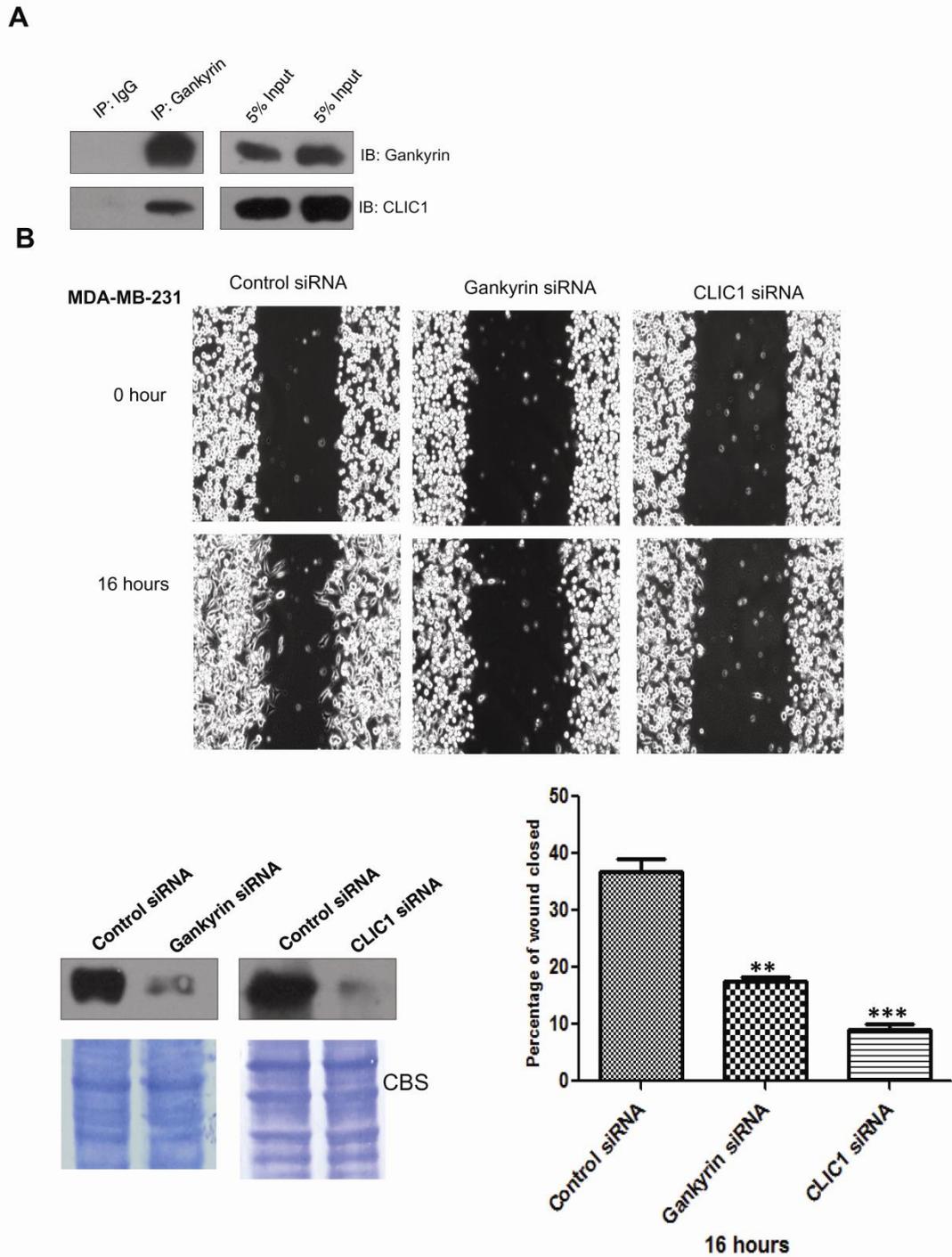
## 8.1 Introduction

Not only is it important to come up with a robust method to predict interacting partners and validate them, it is also important to establish that the predicted interactions are functionally relevant. Many prediction methods fall short of doing this and therefore become suspects with the fear of remaining as prediction methods that are of little practical relevance. In this chapter we validate our experimental proof of principle study by demonstrating that the interactions through predicted motif are functionally relevant and the hot spots site could be a potential drug target. We use Chloride intracellular channel 1 (CLIC1) as an example.

CLIC1 belongs to the class of chloride channels but does not behave like a classical ion channel protein. CLIC1 exists both as a soluble globular protein and also as an integral membrane protein which functions as an ion channel protein (Valenzuela et al., 1997). Induction of stress translocates CLIC1 from cytoplasm to membrane to execute its function as a chloride ( $\text{Cl}^-$ ) channel. CLIC1 helps in the regulation of cell cycle and selectively expresses on the plasma membrane of the cells during G2/M phase (Valenzuela et al., 2000). CLIC1 is known to be overexpressed in many solid tumors like colorectal cancer (Petrova et al., 2008a), lung adenocarcinoma (Wang et al., 2011), gastric carcinoma (Chen et al., 2007a) including recent reports in glioblastoma which suggests its role as a potential target and prognostic marker (Setti et al., 2013). CLIC1 overexpression is shown to increase cell migration and invasion. Gankyrin being an oncoprotein drives many signalling pathways haywired. Hence, we speculated that gankyrin interaction with CLIC1 may be directly associated with the oncogenic property of gankyrin.

## **8.2 Result and Discussion**

In order to evaluate the functional relevance of the interaction between gankyrin and proteins containing EEVD, we chose to study function of gankyrin-CLIC1 interaction because of the following reasons a) CLIC1 was barely detectable in HEK 293 Wt cells and levels increase in response to gankyrin overexpression; b) Chloride intracellular channel proteins are well known for their role in invasion and metastasis (Chen et al., 2007b; Le Guennec et al., 2007; Petrova et al., 2008b). We initially tested the independent role of gankyrin and CLIC1 in the metastatic potential of MDA-MB-231 by performing surrogate wound healing and invasion assays. Knock down of either of the two proteins decreased the ability of these cells to migrate and close the wound. CLIC1 seems to be more effective than gankyrin in healing the wound (Fig. 8.1).

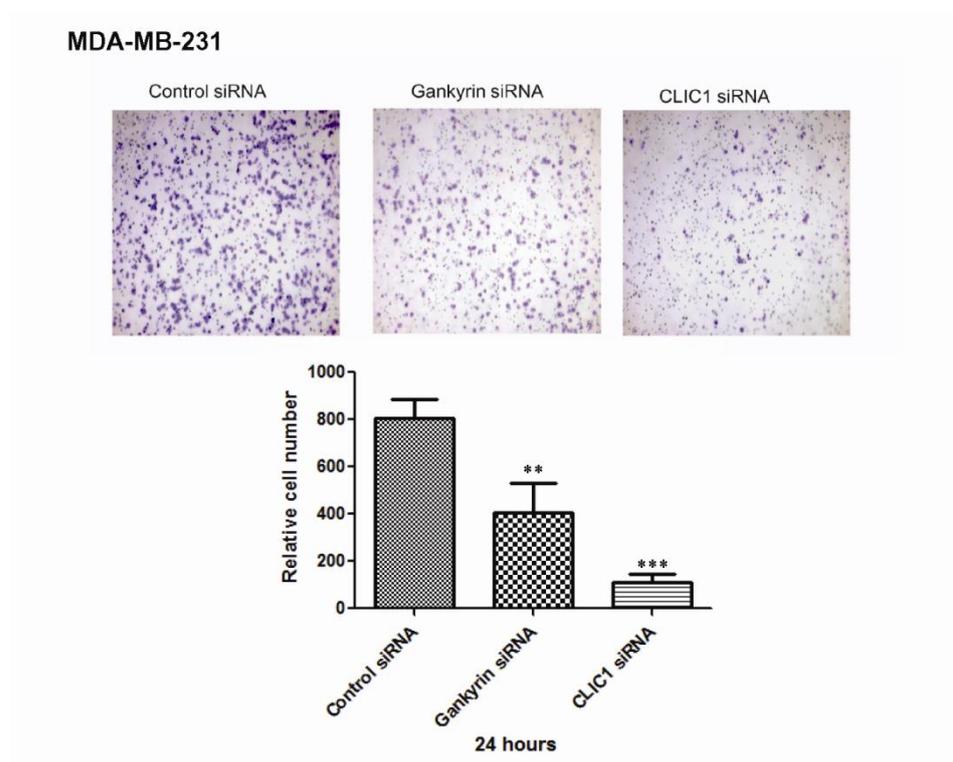


**Figure 8.1**

*Silencing of gankyrin and CLIC1 affects migration in MDA-MB-231 cells. (A) Lysate of MDA-MB-231 cells were added to gankyrin bound or IgG bound Sepharose G beads followed by immunoblotting with anti-CLIC1 and anti-gankyrin antibody; (B) MDA-MB-231 cells were treated with 100 nM of gankyrin specific siRNA or*

*CLIC1* siRNA or control siRNA. After 72 hrs, cells at 90% confluence were treated with 10 µg/ml of mitomycin C for 3 hrs, wounded and healing monitored for 16 hrs. Percentage wound healed at the end of 16 hrs were calculated and the data are represented as mean ±SEM of three experiments. Statistical analysis is done using unpaired t-test (\*\*P=0.0013, \*\*\*P=0.0004). Western blot confirms down regulation of gankyrin and *CLIC1* in the respective experiments. CBS shows equal loading.

These results were recapitulated in the invasion assay using Boyden Chamber (Fig. 8.2). These results confirmed the role of these two proteins in migration and invasion.

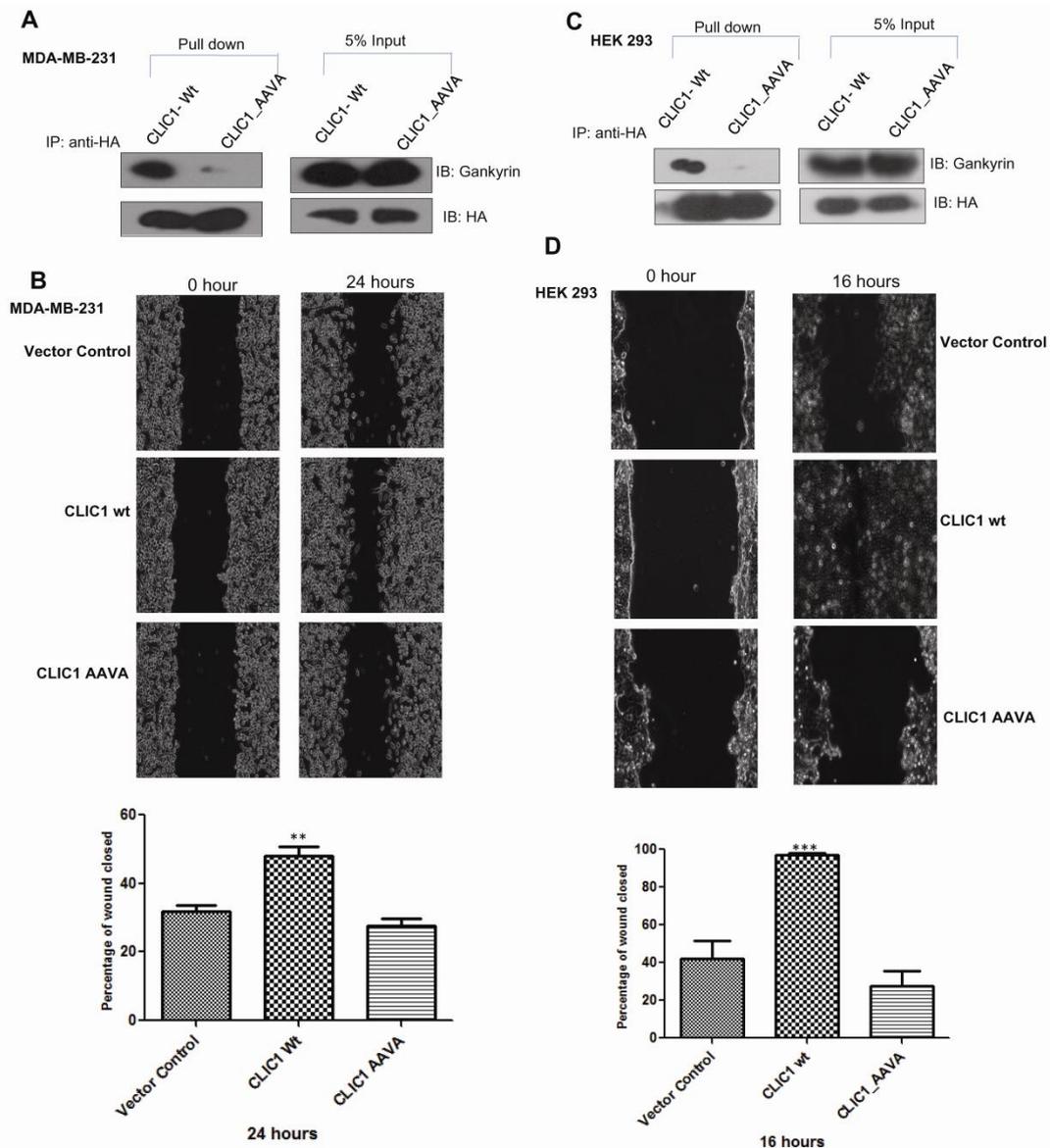


**Figure 8.2**

**Silencing of gankyrin and *CLIC1* affects invasion in MDA-MB-231 cells.** Cells transfected with siRNA gankyrin or siRNA *CLIC1* show reduction in invasive properties of MDA-MB-231 cells as compared to control siRNA transfected cells. A

*representative image for each is shown and the data are represented as mean  $\pm$ SEM of three experiments.*

To check if the interaction is important for driving this phenotype CLIC1 Wt or the AAVA mutant of CLIC1 were overexpressed in MDA-MB-231 cells as well as in HEK 293 cells and their ability to migrate and close the wound was monitored. In MDA-MB-231 cells, over expression of CLIC1 Wt protein resulted in an increase in the percentage of wound closed (48%) as compared to vector control (31%) while the mutant cells (27%) behaved like those of the vector control cells (Fig. 8.3). HEK 293 cells overexpressing CLIC1 Wt showed ~98% wound closure as compared to vector control cells (42%) and cells overexpressing CLIC1\_AAVA mutant showed 27% wound closure.

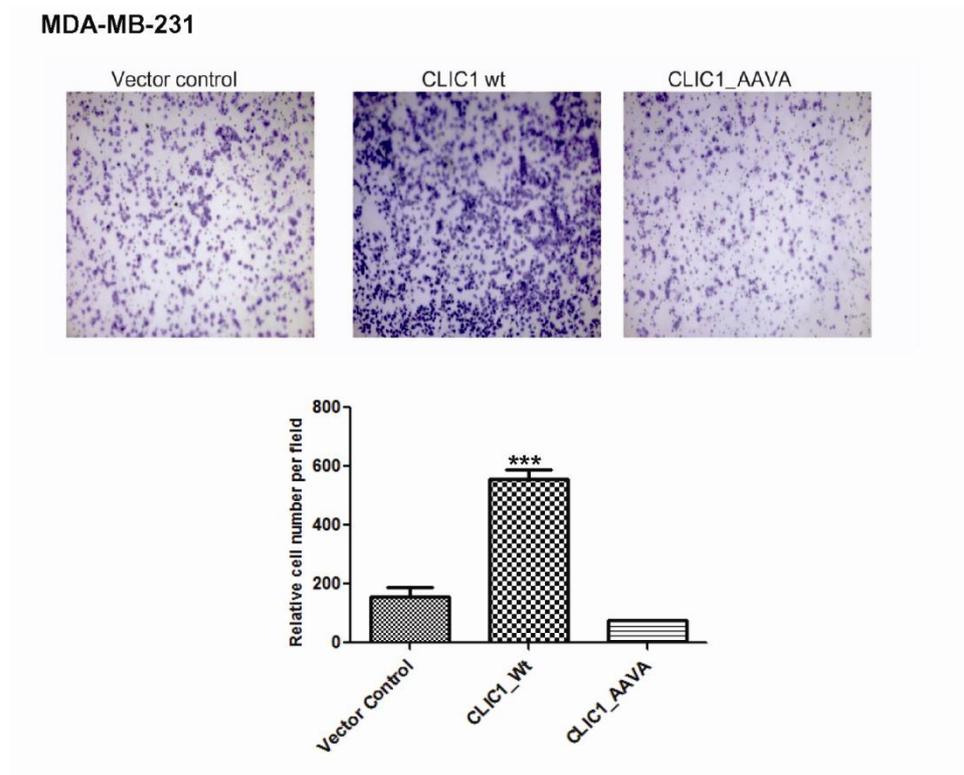


**Figure 8.3**

**Overexpression of CLIC1 and CLIC1\_AAVA affects migration of MDA-MB-231 cells.** (A) Lysate of MDA-MB-231 cells expressing HA-CLIC1 Wt and CLIC1\_AAVA were independently added to anti-HA antibody bound to Sepharose G beads followed by immunoblotting with anti-gankyrin antibody. Gankyrin interacts with CLIC1 Wt but not with its corresponding mutant; (B) MDA-MB-231 cells transiently overexpressing CLIC1 Wt or CLIC1\_AAVA were treated with 10  $\mu$ g/mL of mitomycin C for 3 hrs, wounded and healing monitored for 24 hrs. Percentage of wound healed

was calculated and the data are represented as mean±SEM of three independent experiments (\*\*P = 0.0059); (C) Lysate of HEK 293 cells expressing HA-tagged CLIC1 Wt and CLIC1\_AAVA were independently added to anti-HA antibody bound to Sepharose G beads followed by immunoblotting with anti-gankyrin antibody. Gankyrin interacts with CLIC1 Wt but not with its corresponding mutant; (D) HEK 293 cells transiently overexpressing CLIC1 Wt and CLIC1\_AAVA were treated with 10 µg/ml of mitomycin C for 3 hrs, wounded and healing monitored for 16 hrs. Data were processed as above (\*\*\*P=0.0002).

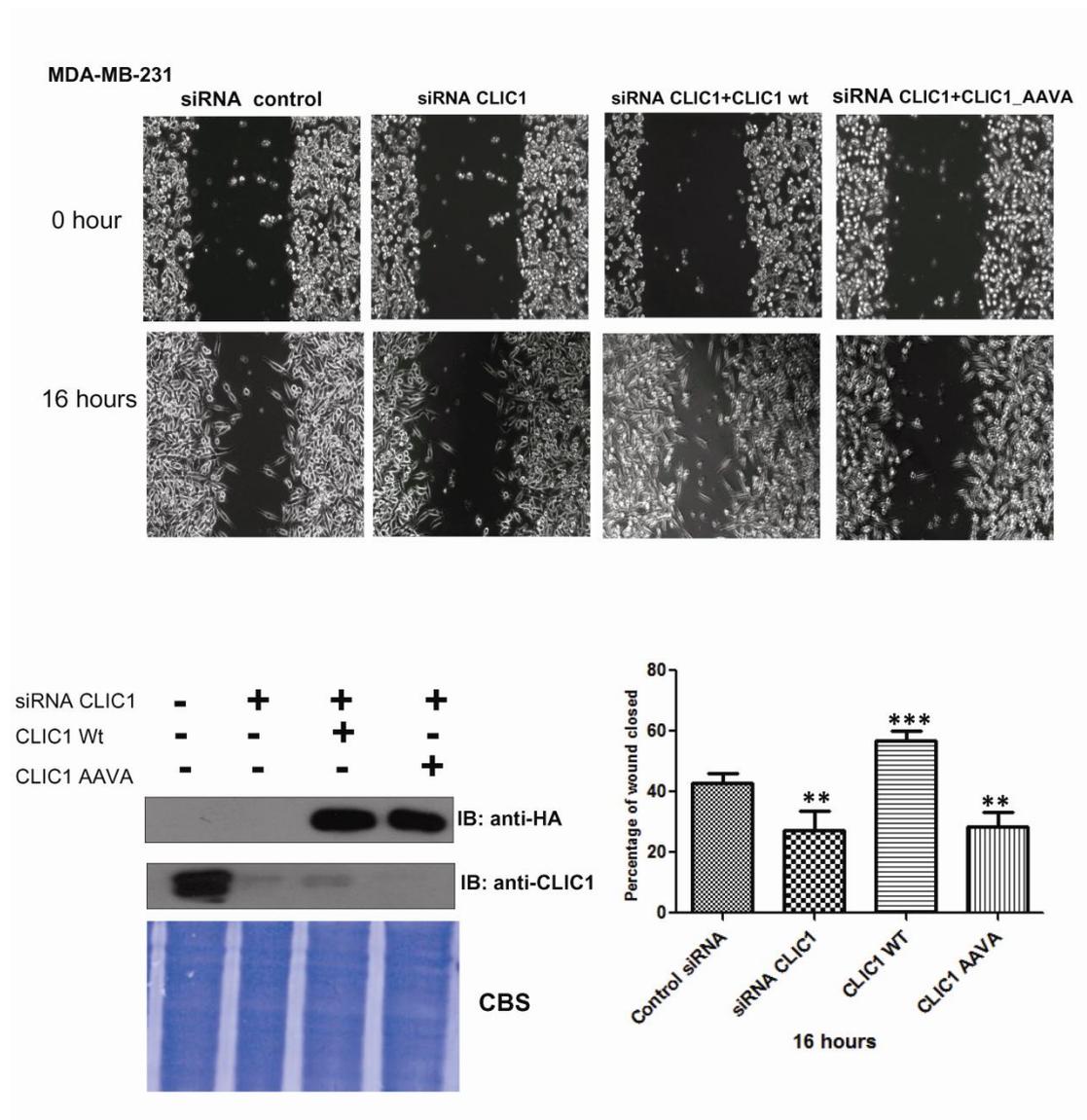
MDA-MB-231 cells overexpressing CLIC1 Wt showed increase in their invasive potential as compared to vector control and CLIC1\_AAVA overexpressing cells (Fig. 8.4) behaved like the vector control cells.



**Figure 8.4**

**Overexpression of CLIC1 and CLIC1\_AAVA affects invasion of MDA-MB-231 cells.** Cells transfected with CLIC1 Wt shows increase in invasion as compared to CLIC1\_AAVA transfected MDA-MB-231 cells. A representative image for each is shown and the data are represented as mean  $\pm$ SEM of three experiments.

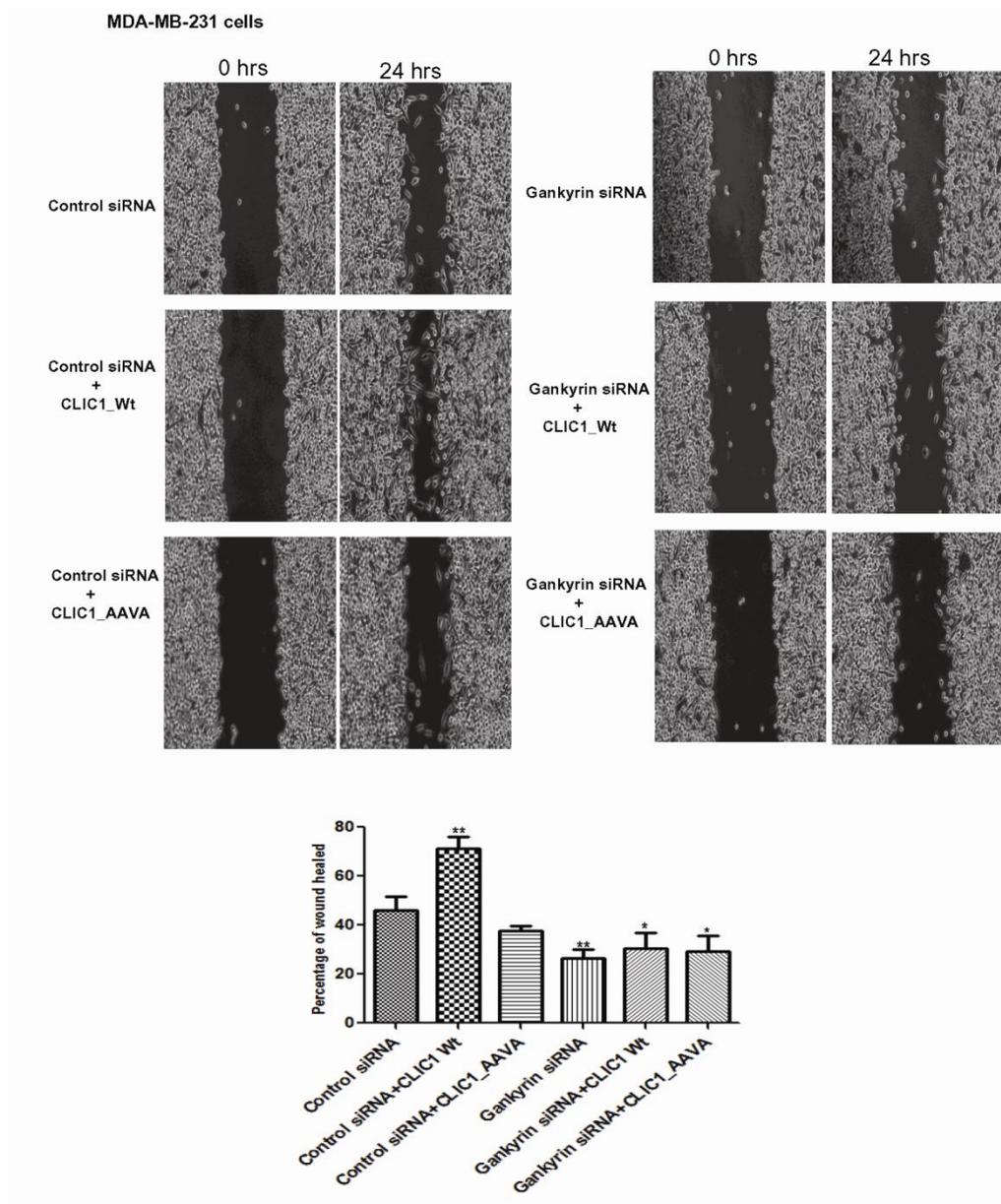
We further confirmed these observations by specifically silencing endogenous CLIC1 by using smart pool of UTR specific siRNA and demonstrating that only Wt CLIC1 and not the mutant protein can rescue the ‘phenotype’ in MDA-MB-231 cells (Fig. 8.5).



**Figure 8.5**

**Interaction of gankyrin with CLIC1 enhances migration.** MDA-MB-231 cells were transfected with control siRNA or CLIC1 siRNA (UTR region) or CLIC1 siRNA+cDNA for CLIC1 Wt or CLIC1 siRNA+cDNA for CLIC1\_AAVA. After 72 hrs these cells were treated with 10 µg/ml of mitomycin C for 3 hrs, wounded and healing monitored for 16 hrs. Data from two independent experiments were processed as described before (\*\*P = 0.0042, \*\*\*P = 0.0006, \*\*P = 0.0020). Cells transfected with CLIC1 Wt show significant increase in the percentage of wound healed as compared with that of CLIC1\_AAVA transfected cells which behave like the vector control cells.

We also did the reverse experiment where we downregulated endogenous gankyrin using smart pool of siRNA against gankyrin and upregulated CLIC1 Wt or CLIC1\_AAVA in MDA-MB-231 cells. CLIC1 Wt transfected cells shows 30% migration whereas CLIC1\_AAVA transfected cells show 29% migration showing no significant difference. The controls were as expected i.e control siRNA transfected cells showed 46% migration, CLIC1 Wt transfected cells showed enhancement in migration to 72% and CLIC1\_AAVA transfected cells behaved approximately like control siRNA transfected cells 38% (Fig. 8.6).

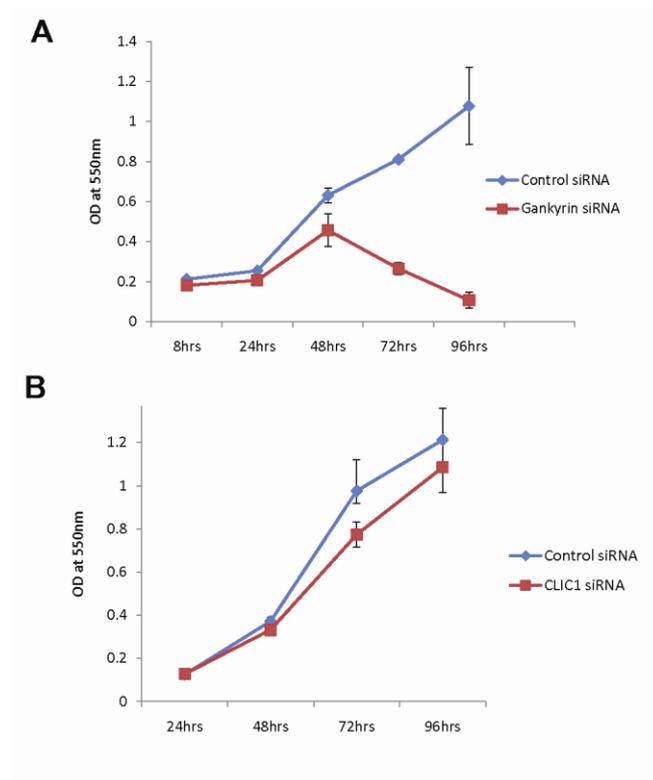


**Figure 8.6**

***Interaction of gankyrin with CLIC1 enhances migration-Reverse experiment.***

*MDA-MB-231 cells were transfected with control siRNA or gankyrin siRNA or control siRNA+CLIC1 Wt or control siRNA+CLIC1\_AAVA or gankyrin siRNA+CLIC1 Wt or gankyrin siRNA+CLIC1\_AAVA. After 72 hrs these cells were treated with 10 µg/mL of mitomycin C for 3 hrs, wounded and healing monitored for 16 h. Data from two independent experiments were processed as described before (\*\*P = 0.0038, \*\*P = 0.0073, \*P = 0.0367, \*P=0.0293).*

As gankyrin is known to play a vital role in proliferation, we checked if gankyrin-CLIC1 interaction is necessary to enhance or decrease the cellular proliferation. We downregulated gankyrin or CLIC1 using smartpool of gankyrin or CLIC1 specific siRNA and checked for proliferation using MTT assay. Downregulation of gankyrin showed decrease in proliferation whereas downregulation of CLIC1 had no effect on cellular proliferation (Fig. 8.7).



**Figure 8.7**

**Effect of gankyrin and CLIC1 on cellular proliferation.** (A) Silencing of gankyrin in MDA-MB-231 cells shows reduction in cellular proliferation whereas (B) Silencing of CLIC1 in MDA-MB-231 does not show any effect on cellular proliferation.

### **11.3 Summary**

These results taken together indicate that interaction of CLIC1 with gankyrin through EEVD is a major determinant in the function of this protein and drives the migratory/invasive property of the breast cancer cell line.

## **CHAPTER 9.**

### ***Expansion of Network to EEXD containing proteins***

## 9.1 Introduction

Structural analysis of gankyrin-S6 ATPase interaction suggests that Val358 in ‘EEVD’ is hydrogen bonded with Lys116 of human gankyrin. They do not share any other bonding with gankyrin. The Nakamura et al., group also showed that mutation of E, E and D to Ala (A) abrogated the interaction between gankyrin and S6 ATPase. This provoked us to hypothesize that not only EEVD containing proteins but other variants ‘EEXD’ might interact with gankyrin. Among the proteins that were reported to interact with gankyrin such as Rb, CDK4, MDM2, MAGE-A4, p65 (RelA) of NF- $\kappa$ B and FIH-1, EEND sequence is present in MDM2 and Rb carries the tetrapeptide motif ‘EEPD’ (Table 9.1). While there is no structural information available for MDM2 in this region, in Rb ‘EEPD’ sequence forms the N-terminus of the solved crystal structure (aa 53EEPD 56). The involvement of these residues in interaction with gankyrin however remains to be tested.

**Table 9.1** List of the sequences of known gankyrin interacting partners and those identified in this study.

<b>Known gankyrin interactors reported in Literature</b>	
<b>Cyclin-dependent kinase 4 (Ref)</b> >gi 4502735 ref NP_000066.1  (Dawson S et al., 2002)	MATSRYEPVAEIGVGAYGTVYKARDPHSGHFVALKSVRVPNGGGGGGLPIST VREVALRRLEAFEHPNVRLMDVCATSRTDREIKVTLVFEHVDQDLRTYLDK APPPGLPAETIKDLMRQFLRGLDFLHANCIVHRDLKPENILVTSGGTVKLADFG ARIYSYQMALTPVVVTLWYRAPEVLLQSTYATPVDMWSVGCIFAEMFRRKPLF CGNSEADQLGKIFDLIGLPPEDDWPRDVSLPRGAFPPRGPRPVQSVVPEMEESGA QLLLEMLTFNPHKRISAFRALQHSYLHKDEGNPE
<b>Melanoma-associated antigen 4</b>	MSSEQKSQHCKPEEGVEAQEEALGLVGAQAPTTEEQEAAVSSSSPLVPGTLEEV PAAESAGPPQSPQGASALPTTISFTCWRQPNEGSSSQEEEGPSTSPDAESLFREAL

<p>&gt;gi 58530865 ref NP_002353.3  (Nagao T et al., 2003)</p>	<p>SNKVDELAHFLLRKYRAKELVTKAEMLERVIKKNYKRCFPVIFGKASESLKMIFGI DVKEVDPASNTYTLVTCLGLSYDGLLGNNQIFPKTGLLIIVLGTIAMEGDSASEE EIWEELGVMGVYDGREHTVYGEPRKLLTQDWVQENYLEYRQVPGSNPARYEF LWGPRALAETS YVKVLEHVVRVNARVRIAYPSLREAALLEEEEGV</p>
<p><b>E3 ubiquitin-protein ligase Mdm2 isoform a</b> &gt;gi 89993689 ref NP_002383.2  (Higashitsuji H et al., 2005)</p>	<p>MVRSRQMCNTNMSVPTDGA VTT SQIPASEQETLVRPKPLLLKLLKSVGAQKDT YTMKEVLFYLGQYIMTKRLYDEKQQHIVYCSNDLLGDLFGVPSFSVKEHRKIYT MIYRNLVVVNQQESSDSGTSVSENRC HLEGGSDQKDLVQELQEEKPSSSHLVSR PSTSSRRRAISETEENSDEL SGERQKRKHSISLSFDESLALCVIREICCERSSSS ESTGTSPNPDL DAGVSEHSGDWLDQDSVSDQFSVEFEVESL DSEDYSLSEEGQE LSDEDEVEYQVTVYQAGESD TDSFEEDPEISLADYWKCTSCNEMNPPLPSHCNR CWALRENWLPEDK GKDKGEISEKAKLENSTQAEEGFDVPDCKKTIVNDSRESC <b>VEEND</b>DKITQASQSQESEDYSQPSTSSSIYSSQEDVKEFEREETQDKESVESLPL LNAIEPCVICQGRPKNGCIVHGKTGHLMACFTCAKCLKKRNKPCPVC RQPIQMI VLTYFP</p>
<p><b>Retinoblastoma-associated protein</b> &gt;gi 108773787 ref NP_00312.2  (Higashitsuji H et al., 2000)</p>	<p>MPPKTPRKAATAAAAAA EPPAPPPPPPEEDPEQDSGPEDLPLVRLEFEET<b>EEP</b> <b>D</b>F TALCQKLKIPDHVRERAWLTWEKVSSVDGVLGGYIQKKKELWGICIFIAAVD LDEMSFTFTELQKNIEISVHKFFNLLKEIDTSTKVDNAMSRLKKYDVL FALFSK LERTCELIYLTQPSSSISTEINSALVLKVS WITFLAKGEVLQMEDDLVISFQLML CVLDYFIKLSPPMLLKEPYKTAVIPINGSRPTPRRGQNR SARIAKQLENDTRIIEVL CKEHECNIDEVKNVYFKNFIPFMNSLGLVTSNGLPEVENLSKRYEEIYLKNKDL DARLFLDHDKTLQTDSIDSFETQRTPRKSNLDEEVNVIPHTPVRTVMNTIQQLM MILNSASDQPS ENLISYFNNCTVNPKESILKRVKDIGYIFKEKFAKAVGQGCVEIG SQRYKLGVRLYYRVMESMLKSEEERLSIQNFSKLLNDNIFHMSLLACALEVVM ATYSRSTSQNLDSGTDLSFPWILNVLNLKAFDFYKVIESFIKAEGNLTREMIKHL ERCEHRIMESLAWLSDSPLFDLIKQSKDREGPTDHLESACPLNPLQNNHTAAD MYLSPVRS P KKKGSTTRVNSTANAETQATSAFQTQKPLKSTSLSLFYKKVYRLA YLRNLTL CERLLSEHPELEHIIWTLFQHTLQNEYELMRDRHLDQIMMCSMYGIC KVKNIDLKFKIIVTAYKDLPHAVQETFKRVLIKEEEYDSIIVFYNSVFMQR LKTNI LQYASTRPPTLSPIPHIPRSPYKFPSSPLRIPGGNIYISPLKSPYKISEGLPTPTKMTP</p>

	RSRILV SIGESFGTSEKFQKINQMVCNSDRVLKRSAEGSNPPKPLKCLRFDIEGSD EADGSKHLPGESKFFQKLAEMTSTRTRMQKQKMNDMSMDTSNKEEK
<b>Transcription factor p65 isoform 1</b> >gi 223468676 ref NP_068810.3  (Chen Y et al., 2007)	MDELFPLIFPAEPAQASGPYVEIIEQPKQRGMRFYKCEGRSAGSIPGERSTDTTK THPTIKINGYTGPGTVRISLVTKDPPHRPHHELVGKDCRDGFYEALCPDRCIHS FQNLGIQCVKKRDLEQAISQRIQTNNNPFQVPIEEQRGDYDLNAVRLCFQVTVR DPSGRPLRLPPVLSHPIFDNRAPNTAELKICRVNRNSGSLGGDEIFLLCDKVQKE DIEVYFTGPGWEARGSFSQADVHRQVAIVFRTTPPYADPSLQAPVRVSMQLRRPS DRELSEPMEFQYLPDTPDRHRIEEKRKRKYETFKSIMKKS PFSGPTDPRPPPRRIA VPSRSSASVPKPAPQYPFTSSLSTINYDEFPTMVFPSPGQISQASALAPAPPQVLPQ APAPAPAPAMVSALAQAPAPVPVLAPGPPQAVAPPAPKPTQAGEGTLSEALLQL QFDDDLGALLGNSTDPVFTDLASVDNSEFQQLLNQGIPVAPHTTEPMLMEYYP EAITRLVTGAQRPPDPAPAPL GAPGLPNGLLSGDEDFSSIADMDFSALLSQISS
<b>Hypoxia-inducible factor 1-alpha inhibitor</b> >gi 148596936 ref NP_060372.2  (Liu Y et al., 2013)	MAATAAEAVASGSGEPREEAGALGPAWDESQLRSYSFPTRPIRLSQSDPRAEE LIENEPVVLTDTNLVYPALKWDLEYLQENIGNGDFSVYSASTHKFLYDEKKM ANFNQFKPRSNREEMKFHEFVEKQLDIQQRGGEERLYLQQLNDTVGRKIVMD FLGFNWNWINKQQGKRGWGQLTSNLLIGMEGNVTPAHYDEQQNFFAQIKGY KRCILFPDQFECLYPYPVHHPCDRQSQVDFDNPDIYERFPNFQNVVGYETVVGP GDVLYIPMYWVHHIESLLNGGITITVNFVYKGAPTPKRIEYPLKAHQKVAIMRN IEKMLGEALGNPQEVGPLLNTMIKGRYN
<b>Novel Interactors reported in this study</b>	
<b>Heat shock protein 70 kDa 1A/1B</b> >gi 194248072 ref NP_05336.3	MAKAAAIGIDLGTTYSCVGVFQHGKVEIANDQGNRTTPSYVAFTDTERLIGDA AKNQVALNPQNTVFDAKRLIGRKFQVQSDMKHWPVQVINDGDKPKVQVS YKGETKAFYPPEISSMVLTKMKEIAEAYLGYPVTNAVITVPAYFNDSQRQATKD AGVIAGLNLRIINEPTAAAIAYGLDRTGKGERNVLIFDLGGGTFDVSILTIDGDI FEVKATAGDTHLGGEDFDNRLVNHVVEEFKRKHKKDISQNKRAVRRLRTACER AKRTLSSSTQASLEIDSLFEGIDFYTSITRARFEELCSDLFRSTLEPVEKALRDAKL DKAQIHDLVLVGGSTRIPKVQKLLQDFFNGRDLNKSINPDEAVAYGAAVQAAIL MGDKSENVQDLLLLDVAPLSLGLTAGGVMTALIKRNSTIPTKQTQIFTTYSDN

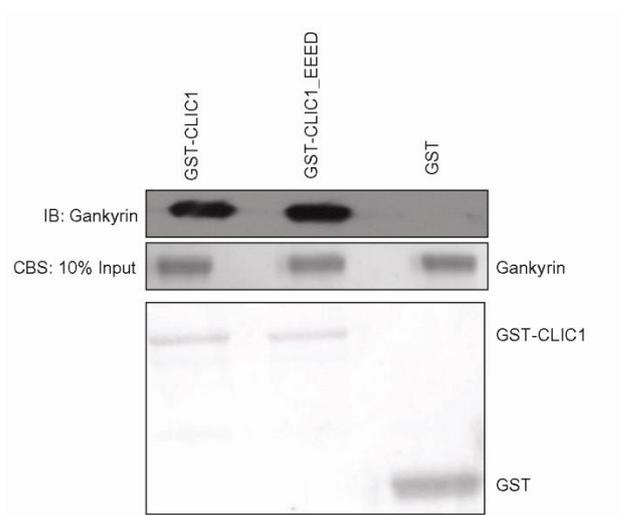
	<p>QPGVLIQVYEGERAMTKDNNLLGRFELSGIPPAPRGVVPQIEVTFDIDANGILNVT  ATDKSTGKANKITITNDKGRLSKEEIERMVQEAKEYKAEDDEVQRERVSNAKNALE  SYAFNMKSAVEDEGLKKGISEADKKKVLDKCQEVISWLDANTLAEKDEFEHKR  KELEQVCNPIISGLYQGAGGPGGGFGAQQPKGGSGSGPTI<b>EEVD</b></p>
<p><b>Heat shock protein</b>  <b>HSP 90-beta isoform a</b>  &gt;gi 431822403 ref NP_0  01258899.1 </p>	<p>MPEEVHHGEEEVETFAFQAEIAQLMSLIINTFYSNKEIFLRELISNASDALDKIRY  ESLTDPKLDGKELKIDIIPNPQERTLTLVDTGIGMTKADLINNLGTIAKSGTKA  FMEALQAGADISMIGQFGVGFYSAYLVAEKVVVITKHNDDEQYAWESSAGGSF  TVRADHGEPIGRGTKVILHLKEDQTEYLEERRVKEVVKKHSQFIGYPITLYLEKE  REKEISDDEAEEEEKGEKEEEDKDDEEKPKIEDVGSDEEDDSGDKKKKTKKIKE  KYIDQEELNKTKPIWTRNPDDITQEEYGEFYKSLTNDWEDHLAVKHFSVEGQLE  FRALLFIPRRAPFDLFENKKKNNIKLYVRRVFIMDSCDELIPEYLNFIKRVVDSE  DLPLNISREMLQQSKILKVIKKNIVKKCLELFSLEAEDKENYKFFYEAFSKNLKL  GIHEDSTNRRRLSELLRYHTSQSGDEMTSLSEYVSRMKETQKSIYYITGESKEQV  ANSAFVERVRKRGFEVVYMTPEIDEYCVQQLKEFDGKSLVSVTKEGLELPEDEE  EKKKMEESKAKFENLCKLMKEILDKKVEKVTISNRLVSSPCCIVTSTYGWTANM  ERIMKAQALRDNSTMGYMMAKKHLEINPDHPIVETLRQKAEADKNDKAVKDL  VVLLFETALLSSGFSLEDPQTHSNRIYRMIKLGGLGIDEDEVAAEENAAVPDEIPP  LEGDEDASRM<b>EEVD</b></p>
<p><b>CLIC1</b>  &gt;gi 49457095 emb CAG  46868.1 </p>	<p>MAEEQPQVELFVKAGSDGAKIGNCPFSQRLFMVLWLKGVTFNVTVDTKRRT  TVQKLCPPGGQLPFLLYGTEVHTDTNKIEEFLEAVLCPPRYPKLAALNPESNTAGL  DIFAKFSAYIKNSNPALNDNLEKGLLKALKVLDNYLTSPLP<b>EEVD</b>ETSAAEDEV  QRKFLDGNELTLADCNLLPKLHIVQVCKKYRGFTIPEAFRGVHRYLSNAYARE  EFASTCPDDEEIELAYEQVAKALK</p>
<p><b>Dual specificity</b>  <b>Mitogen-activated</b>  <b>protein kinase kinase 1</b>  &gt;gi 5579478 ref NP_002</p>	<p>MPKKKPTPIQLNPAPDGSVNGTSSAETNLEALQKKLEELDEQQRKRLEAFL  TQKQKVGELKDDDFEKISELGAGNGGVVFKVSHKPSGLVMARKLIHLEIKPAIR  NQIIRELQVLHECNSPYIVGFYGFYSDGEISICMEHMDGGSLDQVLKAGRIPE  QILGKVSIAVIKGLTYLREKHKIMHRDVKPSNILVNSRGEIKLDFGVSGQLIDS</p>

746.1	<p>MANSFVGTRSYMSPERLQGTHYSVQSDIWSMGLSLVEMAVGRYPIPPDAKEL  ELMFGCQVEGDAAETPPRPRTPGRPLSSYGMDSRPPMAIFELLDYIVNEPPPCLP  SGVFSLEFQDFVNKCLIKNPAERADLKQLMVHAFIKRSDA<b>EEVD</b>FAGWLCSTIG  LNQPSTPTHAAGV</p>
<b>DDAH1</b> protein >gi 21707415 gb AAH3 3680.1	<p>MAGLGHPAAFGRATHAVVRALPESLGQHALRSAKGE<b>EEVD</b>VARAERQHQLYVG  VLGSKLGLQVVELPADESLPDCVFVEDVAVVCEETALITRPGAPSRRKEVDMM  KEALEKLQLNIVEMKDENATLDGGDVLFTGREFFVGLSKRTNQRGAIEILADTFK  DYAVSTVPVADGLHLKSFCSMAGPNLIAIGSSESAQKALKIMQMSDHRYDKL  TVPDDIAANCIYLNIPNKGHVLLHRTPEEYPESAKVYEKLDHMLIPVSMSELEK  VDGLLTCCSVLINKKVDS</p>
<b>G-rich sequence factor</b> <b>1 isoform 1</b> >gi 149193321 ref NP_0 02083.3	<p>MAGTRWVLGALLRGCNCSSCRRTGAACLPFYSAAGSIPSGVSGRRRLLLL  AAAAAASQTRGLQTGPVPPGRLAGPPAVATSAAAAAASYPALRASLLPQSLA  AAAAVPTRSYSQESKTTYLEDLPPPPEYELAPSKLE<b>EEVD</b>DVFLIRAQGLPWSC  MEDVLNFFSDCRIRNGENGIHFLNRDGRGDALIEMESEQDVQKALEKHRM  YMGQRYVEVYEINNEDVDALMKSLQVKSSPVVNDGVVRLRGLPYSCNEKDIV  DFFAGLNIVDITFVMDYRGRKKTGEAYVQFEEPAMANQALLKHREEIGNRYIEIF  PSRRNEVRTHVGSYKGGKIASFPTAKYITEPEMVFEHEVNEDIQPMTAFESEKE  IELPKEVPEKLPEAADFGTSSLHFVHMRGLPFQANAQDIINFFAPLKPVRITMEY  SSSGKATGEADVHFETHEDAVAAMLKDRSHVHHR YIELFLNSCPKGGK</p>
<b>Cytoplasmic protein</b> <b>NCK2 isoform A</b> >gi 52630425 ref NP_00 1004720.1	<p>MTEEVIVIAKWDYTAQQDQELDIKKNERLWLLDDSKTWWRVNAANRTGYVP  SNYVERKNSLKKGSLVKNLKD TLGLGKTRRKT SARDASPTPSTDAEYPANGSG  ADRIYDLNIPAFVKFAYVAEREDEL SVKGSRTVMEKCS DGWWRG SYNGQIG  WFPSNYVL<b>EEVD</b>EAAAESPSFLSLRKGASLSNGQGSRLHV VQTLYPFSSVTEE  ELNFEKGETMEVIEKPENDPEWWKCKNARGQVGLVPKNYVVVLS DGPALHPA  HAPQISYTG PSSGRFAGREWYYGNVTRHQAECALNERGVEGDFLIRDSESSPS  DFSVSLKASGKNKHFKVQLVDNVY CIGQRRFHTMDELVEHYKKAPIFTSEHGE  KLYLVRALQ</p>

It is interesting to note two key proteins known to interact with gankyrin carry the EEVD motif. It seems therefore that other sequence variants of EEVD such as EEXD (X is any residue) may also interact with gankyrin. Here we test this hypothesis using calreticulin as a representative of EEXD proteins. This protein was seen to be upregulated when gankyrin was transexpressed in HEK-293 cells. Calreticulin was one of the proteins that were repeatedly identified by MS-MS as a differentially expressed protein by 2D electrophoresis.

## Results and Discussion

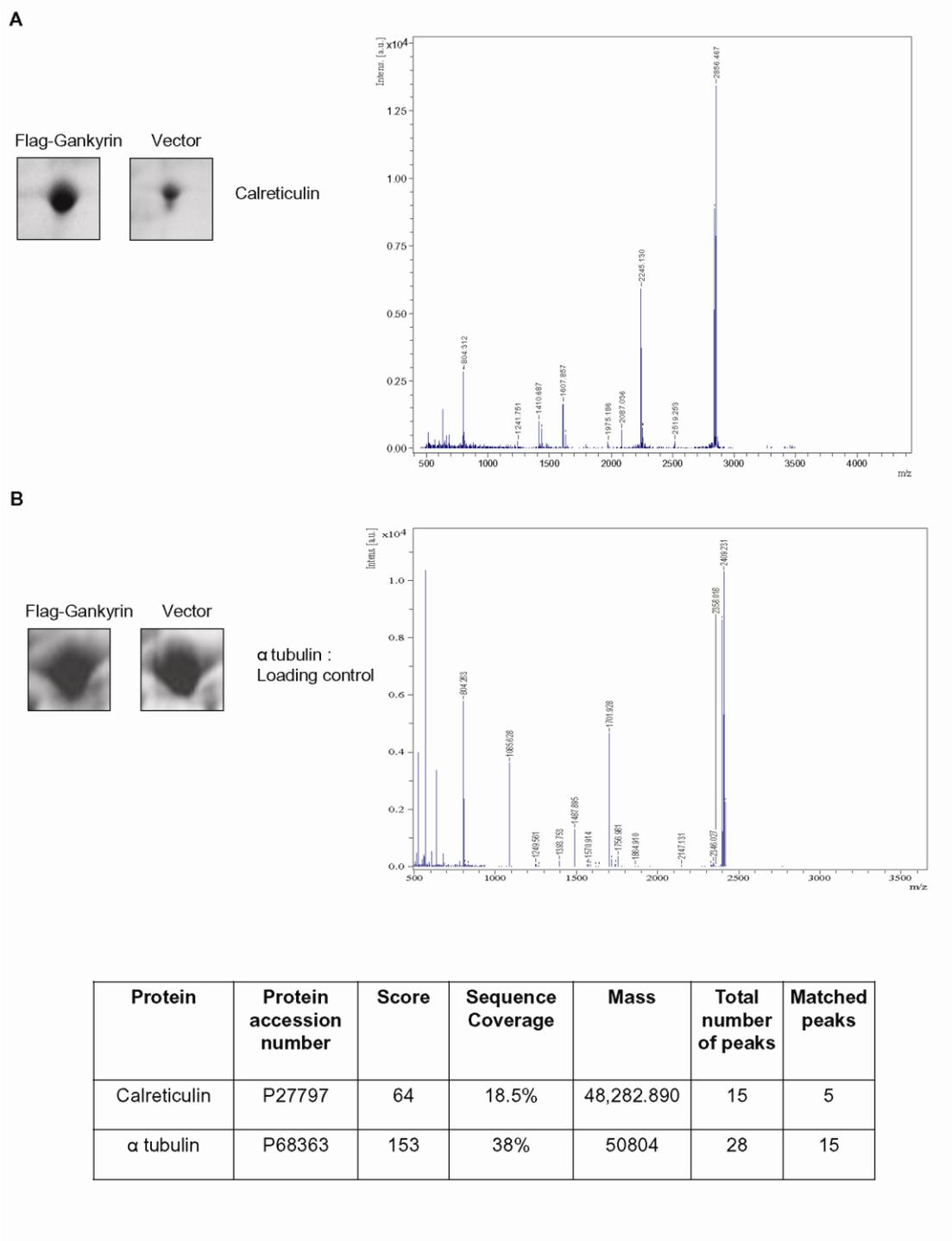
Our experimental data suggests that valine in EEVD is not important for binding with gankyrin. When Val is mutated to Glu in CLIC1 the mutant protein still binds to gankyrin (Fig 9.1). On the other hand  $\beta$ -catenin which carries EEED does not interact with gankyrin though we observed the stabilization of  $\beta$ -catenin which is not surprising because interactions are not dictated by sequence or accessibility alone. Subcellular localization and other regulatory mechanisms are well known to play a role in determining interactions.



**Figure 9.1**

*Gankyrin does not require Valine in EEVD for interaction with CLIC1. Purified GST, GST-CLIC1, GST-CLIC1\_EEED were independently immobilized on glutathione-Sepharose 4B and incubated with 6X-His tagged gankyrin. Bound His-gankyrin was detected using anti-gankyrin antibody. CLIC1 Wt and CLIC1\_EEED interacts with gankyrin whereas GST alone did not interact. The input for gankyrin and GST-fusion proteins are shown.*

We chose to test calreticulin, as it was found to be overexpressed in gankyrin over expressing HEK 293 stable clones (Fig. 9.2).



**Figure 9.2**

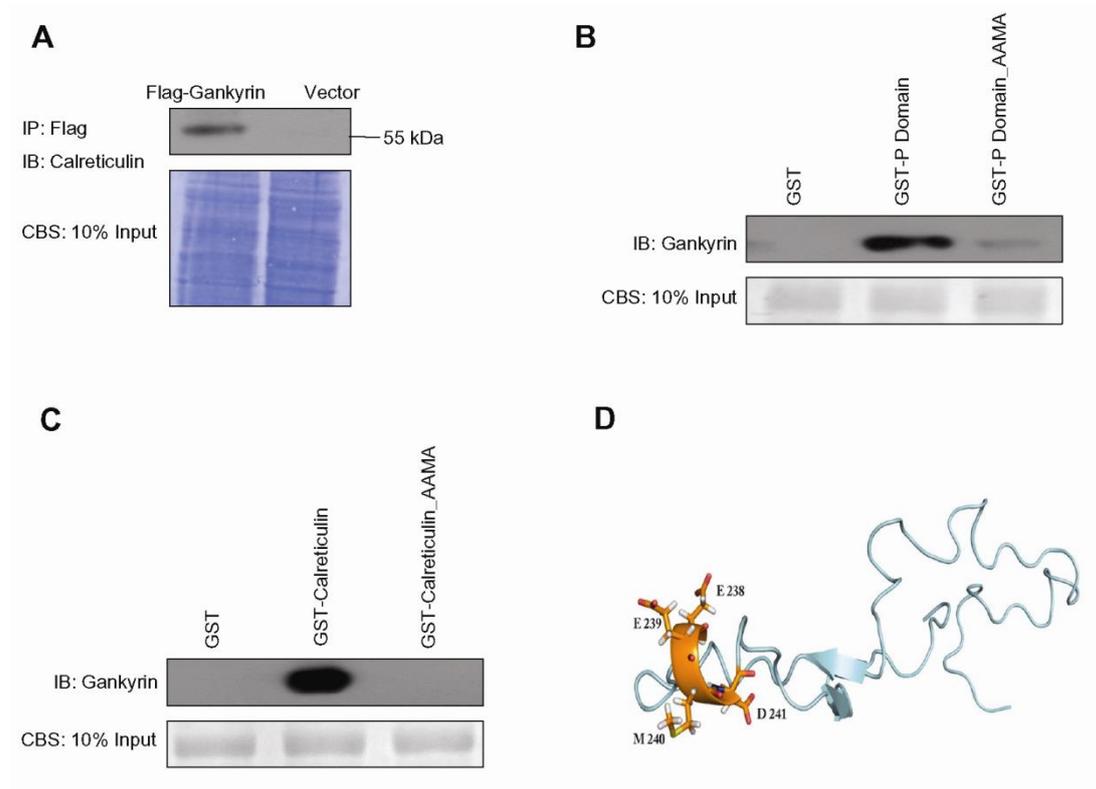
*Comparison of protein profiles using two dimensional gel electrophoresis. Lysates from gankyrin over expressing cells or vector control were separated by 2D gel electrophoresis. Differentially expressed protein (A) calreticulin was identified by MALDI-TOF-TOF. Loading control is represented by (B)  $\alpha$  tubulin which was also*

*identified by MALDI-TOF-TOF. Peptide mass tolerance was set at  $\pm 100$  and only one missed cleavage was allowed.*

When the sequence and structure were analyzed, calreticulin was found to carry motif EEMD sequence in the solvent accessible region. Calreticulin is an endoplasmic reticulum protein and a chaperone that helps in folding, oligomeric assembly and quality control in the endoplasmic reticulum (ER) via the calreticulin/calnexin cycle (Bedard et al., 2005; Zamanian et al., 2013). However, there are reports which show that calreticulin is expressed in nuclear matrixes of hepatocellular carcinoma cells (Yoon et al., 2000). Therefore we tested the interaction between calreticulin and gankyrin via the EEMD motif.

Affinity pull down of gankyrin using flag antibody in lysate over expressing HEK 293 cells was performed and probed for the presence of calreticulin. Calreticulin was found to be present in the immune-complex indicating that EEMD containing variant (calreticulin) might be responsible for interacting with gankyrin (Fig. 9.3A). Calreticulin has 3 domains- N terminal domain, P domain containing EEMD sequence and C-terminal domain. The available NMR structure of the P-loop of rat calreticulin shows that this sequence EEMD in calreticulin present in the P-loop is accessible to the solvent (Fig. 9.3 D). To confirm the direct interaction between calreticulin, P domain of calreticulin and gankyrin, we purified gankyrin as his-gankyrin fusion protein and calreticulin, the P domain of calreticulin as GST fusion protein followed by GST pull down assay. GST pull down assay showed that gankyrin interacts with GST-calreticulin and GST-P domain whereas GST control does not show any interaction (Fig. 9.3 B and C). To further prove that E, E and D in EEMD are important for interacting with gankyrin, we made mutant constructs GST-

calreticulin\_AAVA and GST-P domain\_AAVA and performed GST pull down assay (Fig. 9.3 B and C). Mutants do not show any interaction whereas Wt proteins were found to be interacting.

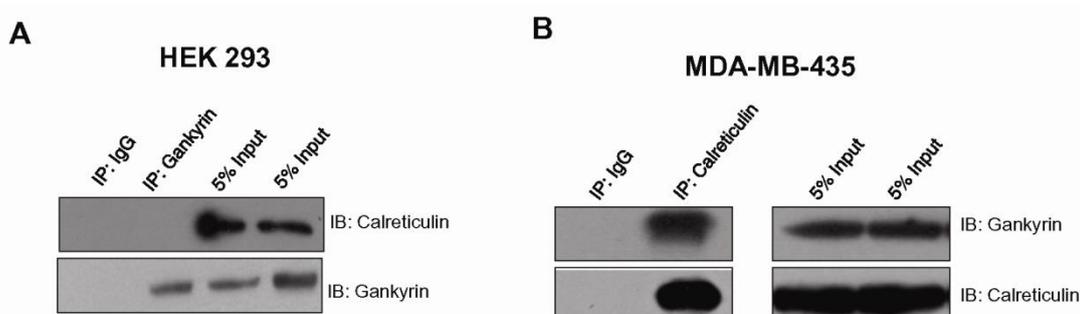


**Figure 9.3**

***Gankyrin interacts with calreticulin within the cellular milieu and in vitro.*** (A) Lysates of HEK 293 cells expressing flag-gankyrin or flag alone (vector) were immobilized on M2 agarose anti-flag beads followed by immunoblotting with anti-calreticulin antibody. (B) Purified GST-P loop, GST-P loop\_AAVA or GST alone were independently immobilized on glutathione-Sepharose 4B and incubated with 6X-His tagged gankyrin. Bound His-gankyrin was detected using anti-gankyrin antibody. GST-P loop interacts with His-gankyrin but the mutant is unable to do so. (C) Purified GST-calreticulin, GST-calreticulin\_AAVA or GST alone were independently immobilized on glutathione-Sepharose 4B and incubated with 6X-His tagged gankyrin. Bound His-gankyrin was detected using anti-gankyrin antibody.

*GST-calreticulin interacts with His-gankyrin but the mutant is unable to do so. (D) NMR Structure of the rat calreticulin P-loop (1HHN) sequence EEMD shown as stick is accessible to the solvent. Image was generated using PyMol (The PyMOL Molecular Graphics System, Version 1.5.0.1 Schrödinger, LLC).*

We further tested the interaction at normal physiological levels inside HEK 293 cells. Immunoprecipitation assay using gankyrin antibody was performed and probed for the calreticulin in the immune complex. Calreticulin was not found to be interacting with gankyrin in HEK 293 Wt cells but was detected in gankyrin transexpressing stable clones (Fig. 9.4 A). Hence, we checked for the interaction in breast cancer cell line MDA-MB-435 cells. Calreticulin was detected to be interacting in MDA-MB-435 cells at their normal physiological conditions (Fig. 9.4 B).



**Figure 9.4**

*Calreticulin with EEMD (variant of EEVD) in the accessible region interacts with gankyrin only in gankyrin overexpressing stable HEK 293 clones and MDA-MB-435 cells.*

*(A) Lysates of HEK 293 cells were added to gankyrin bound and IgG bound sepharose G beads followed by immunoblotting with anti-calreticulin and anti-gankyrin antibody. Results show that gankyrin does not interact endogenously with calreticulin. (B) Lysates of MDA-MB-435 cells were added to calreticulin bound and IgG bound sepharose G beads followed by immunoblotting with gankyrin antibody*

*showing that gankyrin interacts with calreticulin inside the cellular milieu of MDA-MB-435 cells.*

It confirms that this interaction is unique to gankyrin overexpressing cancer cells or cancerous cell lines. One caveat of the study is that the subcellular distribution of calreticulin and colocalization with gankyrin has not been investigated.

### **Summary**

Calreticulin contains EEMD sequence, a variant of EEXD and is shown to interact with gankyrin through E, E and D residues. It is a proof of principle that valine in EEVD sequence is not important for interaction and therefore variants of EEXD may be the potential interactors of gankyrin. Calreticulin interaction with gankyrin is unique to gankyrin over expressing cells or cancerous cells. Moreover, calreticulin is observed to be stabilized upon gankyrin overexpression in stable HEK 293 clones over expressing gankyrin. Hence, we believe that this interaction might have the role in causing malignancy in tumor cells where gankyrin acts as an oncoprotein.

## **CHAPTER 10.**

### ***Conclusion and Significance of Study***

Protein-protein interactions are central to cellular communication (Kar et al., 2009). These interactions are stabilized or new interactions are created in abnormal conditions like cancer. Therefore identifying system wide interactions of a key regulatory protein like an oncoprotein or a tumor suppressor protein is crucial for understanding the deregulated phenotype. Many studies aimed at characterizing network of interactions show that oncoproteins act as hubs. Perturbation of key interactions of such hubs with other proteins in the node would render the network unstable. Therefore these interactions are vulnerable and may be utilized for therapeutic intervention (Tsai et al., 2009).

In the present study, we use the concept of Short Linear Sequence Motifs and the residue level knowledge of a known protein interface of gankyrin-S6 ATPase complex to predict unknown complexes of gankyrin. The premise was based on the following observations. 1) Oncogenes as hub proteins that can interact with multiple proteins and altering or destabilizing the network. 2) Interaction of hub protein with multiple partners can be mediated by a common recognition motif. 3) Hot spots are present on the interface and when mutated abrogate the interaction with the  $\Delta\Delta G \geq 2\text{kcal}$ . 4) Hot spot residues are conserved across the interface (Bogan and Thorn, 1998).

Using the structural details of Gankyrin-S6 ATPase interface, seven novel gankyrin interacting proteins. Biochemical, biophysical studies followed by mutagenesis experiments confirmed predicted SLIM and the hot spot site, EEVD, as the common recognition motif among gankyrin interacting proteins. CLIC1, DDAH1, MAP2K1 interactions are unique to gankyrin over expressing HEK 293 clones and malignant cancer cells (MDA-MB-231, MDA-MB-435) and therefore unique to gankyrin mediated network.

Gankyrin may deregulate normal cellular homeostasis by stabilizing the basal network with Hsp70, Hsp90, GRSF1, MAP2K1, CLIC1, DDAH1, NCK2 or rewire the network by introducing new players such as CLIC1, DDAH1, MAP2K1 and shift the nodes and edges of the functional networks. We further demonstrate that interaction of CLIC1 with gankyrin through EEVD is mandatory for the ability of MDA-MB-231 cells to migrate and invade as demonstrated by the wound healing assay and invasion assay. The short peptide, EEVD is able to inhibit the binding of CLIC1 to gankyrin suggesting that EEVD peptide mimetics as the potential inhibitor of the migratory potential and perhaps invasive properties of cancer.

### *Significance of the Study*

Our study integrates various key important principles described above for accurately identifying novel protein protein interactions. Judging by the functional annotations of the proteins identified it seems that this network of interactions is likely to bridge at least six of the eight hall mark properties of cancer. Our ability to inhibit these interactions implies that this network of interactions can be perturbed. Disruption of interaction between gankyrin and CLIC1 by mutating the hot spot site inhibits the ability of the cancer cell line MDA-MB-231 cells to migrate or invade thus validating functional relevance of these interactions. Taken together the concept of short linear sequence motifs at protein interfaces can be used to identify novel functionally relevant protein complexes formed by key hub proteins. The study presented here also forms the platform for future investigation on the ultra structural details of the interaction, biophysical characteristics and the generation of polypharmacological drugs by rational design or small molecule screening.

## References:

Abu-Farha, M., Elisma, F., and Figeys, D. (2008). Identification of protein–protein interactions by mass spectrometry coupled techniques. In *Protein–Protein Interaction* (Springer), pp. 67-80.

Back, J.W., de Jong, L., Muijsers, A.O., and de Koster, C.G. (2003). Chemical cross-linking and mass spectrometry for protein structural modeling. *Journal of molecular biology* 331, 303-313.

Bai, Z.-f., Tai, Y., Li, W., Zhen, C., Gu, W., Jian, Z., Wang, Q., Lin, J.E., Zhao, Q., and Gong, W. (2013a). Gankyrin activates IL-8 to promote hepatic metastasis of colorectal cancer. *Cancer research*.

Bai, Z., Tai, Y., Li, W., Zhen, C., Gu, W., Jian, Z., Wang, Q., Lin, J.E., Zhao, Q., and Gong, W. (2013b). Gankyrin activates IL-8 to promote hepatic metastasis of colorectal cancer. *Cancer research* 73, 4548-4558.

Bedard, K., Szabo, E., Michalak, M., and Opas, M. (2005). Cellular functions of endoplasmic reticulum chaperones calreticulin, calnexin, and ERp57. *International review of cytology* 245, 91-121.

Bedford, L., Paine, S., Sheppard, P.W., Mayer, R.J., and Roelofs, J. (2010). Assembly, structure, and function of the 26S proteasome. *Trends in cell biology* 20, 391-401.

Behrens, J., Vakaet, L., Friis, R., Winterhager, E., Van Roy, F., Mareel, M.M., and Birchmeier, W. (1993). Loss of epithelial differentiation and gain of invasiveness correlates with tyrosine phosphorylation of the E-cadherin/beta-catenin complex in

cells transformed with a temperature-sensitive v-SRC gene. *The Journal of cell biology* 120, 757-766.

Berggård, T., Linse, S., and James, P. (2007). Methods for the detection and analysis of protein–protein interactions. *Proteomics* 7, 2833-2842.

Bogan, A.A., and Thorn, K.S. (1998). Anatomy of hot spots in protein interfaces. *Journal of molecular biology* 280, 1-9.

Braverman, L.E., and Quilliam, L.A. (1999). Identification of Grb4/Nck $\beta$ , a src homology 2 and 3 domain-containing adapter protein having similar binding and biological properties to Nck. *Journal of Biological Chemistry* 274, 5542-5549.

Byrum, S., Smart, S.K., Larson, S., and Tackett, A.J. (2012). Analysis of stable and transient protein–protein interactions. In *Chromatin Remodeling* (Springer), pp. 143-152.

Calderwood, S.K. (2013). *Molecular Cochaperones: Tumor Growth and Cancer Treatment*. Scientifica 2013.

Carrello, A., Allan, R.K., Morgan, S.L., Owen, B.A., Mok, D., Ward, B.K., Minchin, R.F., Toft, D.O., and Ratajczak, T. (2004). Interaction of the Hsp90 cochaperone cyclophilin 40 with Hsc70. *Cell Stress Chaperones* 9, 167-181.

Ceol, A., Chatr-aryamontri, A., Santonico, E., Sacco, R., Castagnoli, L., and Cesareni, G. (2007). DOMINO: a database of domain–peptide interactions. *Nucleic acids research* 35, D557-D560.

CHAN, C.C., DOSTIE, J., DIEM, M.D., FENG, W., Mann, M., RAPPSILBER, J., and DREYFUSS, G. (2004). eIF4A3 is a novel component of the exon junction complex. *Rna* 10, 200-209.

Chen, C.D., Wang, C.S., Huang, Y.H., Chien, K.Y., Liang, Y., Chen, W.J., and Lin, K.H. (2007a). Overexpression of CLIC1 in human gastric carcinoma and its clinicopathological significance. *Proteomics* 7, 155-167.

Chen, C.D., Wang, C.S., Huang, Y.H., Chien, K.Y., Liang, Y., Chen, W.J., and Lin, K.H. (2007b). Overexpression of CLIC1 in human gastric carcinoma and its clinicopathological significance. *Proteomics* 7, 155-167.

Chen, Y., Li, H.H., Fu, J., Wang, X.F., Ren, Y.B., Dong, L.W., Tang, S.H., Liu, S.Q., Wu, M.C., and Wang, H.Y. (2007c). Oncoprotein p28 GANK binds to RelA and retains NF-kappaB in the cytoplasm through nuclear export. *Cell research* 17, 1020-1029.

Chen, Y., Li, H.H., Fu, J., Wang, X.F., Ren, Y.B., Dong, L.W., Tang, S.H., Liu, S.Q., Wu, M.C., and Wang, H.Y. (2007d). Oncoprotein p28GANK binds to RelA and retains NF-κB in the cytoplasm through nuclear export. *Cell research* 17, 1020-1029.

Collavin, L., Lunardi, A., and Del Sal, G. (2010). p53-family proteins and their regulators: hubs and spokes in tumor suppression. *Cell Death & Differentiation* 17, 901-911.

Cordenonsi, M., Dupont, S., Maretto, S., Insinga, A., Imbriano, C., and Piccolo, S. (2003). Links between tumor suppressors: p53 is required for TGF-β gene responses by cooperating with Smads. *Cell* 113, 301-314.

Crane, A.M., and Bhattacharya, S.K. (2013). The Use of Bromodeoxyuridine Incorporation Assays to Assess Corneal Stem Cell Proliferation. In *Corneal Regenerative Medicine* (Springer), pp. 65-70.

Davey, N.E., Shields, D.C., and Edwards, R.J. (2006). SLiMDisc: short, linear motif discovery, correcting for common evolutionary descent. *Nucleic acids research* *34*, 3546-3554.

Dawson, S., Apcher, S., Mee, M., Higashitsuji, H., Baker, R., Uhle, S., Dubiel, W., Fujita, J., and Mayer, R.J. (2002). Gankyrin is an ankyrin-repeat oncoprotein that interacts with CDK4 kinase and the S6 ATPase of the 26 S proteasome. *Journal of Biological Chemistry* *277*, 10893-10902.

De Las Rivas, J., and Fontanillo, C. (2010). Protein–protein interactions essentials: key concepts to building and analyzing interactome networks. *PLoS computational biology* *6*, e1000807.

DeLano, W.L. (2002). Unraveling hot spots in binding interfaces: progress and challenges. *Current opinion in structural biology* *12*, 14-20.

Dhanasekaran, N., and Reddy, E.P. (1998). Signaling by dual specificity kinases. *Oncogene* *17*.

Dong, L.-w., Yang, G.-z., Pan, Y.-f., Chen, Y., Tan, Y.-x., Dai, R.-y., Ren, Y.-b., Fu, J., and Wang, H.-y. (2011). The oncoprotein p28GANK establishes a positive feedback loop in  $\beta$ -catenin signaling. *Cell research* *21*, 1248-1261.

Dror, T., and Ivett, B. (2007). Recruitment of rare 3-grams at functional sites: Is this a mechanism for increasing enzyme specificity? *BMC Bioinformatics* *8*.

Dunker, A.K., Cortese, M.S., Romero, P., Iakoucheva, L.M., and Uversky, V.N. (2005). Flexible nets. *Febs Journal* 272, 5129-5148.

Edwards, R.J., Davey, N.E., and Shields, D.C. (2007). SLiMFinder: a probabilistic method for identifying over-represented, convergently evolved, short linear motifs in proteins. *PLoS One* 2, e967.

Ekman, D., Light, S., Björklund, Å.K., and Elofsson, A. (2006). What properties characterize the hub proteins of the protein-protein interaction network of *Saccharomyces cerevisiae*? *Genome biology* 7, R45.

Errington, T.M., and Macara, I.G. (2013). Depletion of the Adaptor Protein NCK Increases UV-Induced p53 Phosphorylation and Promotes Apoptosis. *PloS one* 8, e76204.

Espina, V., Woodhouse, E.C., Wulfkuhle, J., Asmussen, H.D., Petricoin III, E.F., and Liotta, L.A. (2004). Protein microarray detection strategies: focus on direct detection technologies. *Journal of immunological methods* 290, 121-133.

Fischer, H.P. (2005). Towards quantitative biology: integration of biological information to elucidate disease pathways and to guide drug discovery. *Biotechnology annual review* 11, 1-68.

Fraser, H.B. (2005). Modularity and evolutionary constraint on proteins. *Nature genetics* 37, 351-352.

Fraternali, F., and Cavallo, L. (2002). Parameter optimized surfaces (POPS): analysis of key interactions and conformational changes in the ribosome. *Nucleic acids research* 30, 2950-2960.

Fu, X.-Y., Wang, H.-Y., Tan, L., Liu, S.-Q., Cao, H.-F., and Wu, M.-C. (2002). Overexpression of p28/gankyrin in human hepatocellular carcinoma and its clinical significance. *World Journal of Gastroenterology* 8, 638-643.

Fulcher, D., and Wong, S. (1999). Carboxyfluorescein succinimidyl ester-based proliferative assays for assessment of T cell function in the diagnostic laboratory. *Immunology and cell biology* 77, 559-564.

Gao, L., Xie, H., Dong, L., Zou, J., Fu, J., Gao, X., Ou, L., Xiang, S., and Song, H. (2014). Gankyrin is essential for hypoxia enhanced metastatic potential in breast cancer cells. *Molecular medicine reports* 9, 1032-1036.

Green, D.R., and Kroemer, G. (2009). Cytoplasmic functions of the tumour suppressor p53. *Nature* 458, 1127-1130.

Hanahan, D., and Weinberg, R.A. (2011a). Hallmarks of cancer: the next generation. *Cell* 144, 646-674.

Hanahan, D., and Weinberg, R.A. (2011b). Hallmarks of cancer: the next generation. *Cell* 144, 646-674.

Haynes, C., Oldfield, C.J., Ji, F., Klitgord, N., Cusick, M.E., Radivojac, P., Uversky, V.N., Vidal, M., and Iakoucheva, L.M. (2006). Intrinsic disorder is a common feature of hub proteins from four eukaryotic interactomes. *PLoS computational biology* 2, e100.

He, X., and Zhang, J. (2006). Why do hubs tend to be essential in protein networks? *PLoS genetics* 2, e88.

Higashitsuji, H., Higashitsuji, H., Itoh, K., Sakurai, T., Nagao, T., Sumitomo, H., Masuda, T., Dawson, S., Shimada, Y., and Mayer, R.J. (2005). The oncoprotein gankyrin binds to MDM2/HDM2, enhancing ubiquitylation and degradation of p53. *Cancer cell* 8, 75-87.

Higashitsuji, H., Higashitsuji, H., Liu, Y., Masuda, T., Fujita, T., Abdel-Aziz, H.I., Kongkham, S., Dawson, S., John Mayer, R., and Itoh, Y. (2007a). The oncoprotein gankyrin interacts with RelA and suppresses NF- $\kappa$ B activity. *Biochemical and biophysical research communications* 363, 879-884.

Higashitsuji, H., Higashitsuji, H., Liu, Y., Masuda, T., Fujita, T., Abdel-Aziz, H.I., Kongkham, S., Dawson, S., John Mayer, R., Itoh, Y., *et al.* (2007b). The oncoprotein gankyrin interacts with RelA and suppresses NF-kappaB activity. *Biochemical and biophysical research communications* 363, 879-884.

Higashitsuji, H., Itoh, K., Nagao, T., Dawson, S., Nonoguchi, K., Kido, T., Mayer, R.J., Aii, S., and Fujita, J. (2000a). Reduced stability of retinoblastoma protein by gankyrin, an oncogenic ankyrin-repeat protein overexpressed in hepatomas. *Nature medicine* 6, 96-99.

Higashitsuji, H., Itoh, K., Nagao, T., Dawson, S., Nonoguchi, K., Kido, T., Mayer, R.J., Aii, S., and Fujita, J. (2000b). Reduced stability of retinoblastoma protein by gankyrin, an oncogenic ankyrin-repeat protein overexpressed in hepatomas. *Nature medicine* 6, 96-99.

Hu, Z., Ma, B., Wolfson, H., and Nussinov, R. (2000). Conservation of polar residues as hot spots at protein interfaces. *Proteins: Structure, Function, and Bioinformatics* 39, 331-342.

Huang, J.-S., Chao, C.-C., Su, T.-L., Yeh, S.-H., Chen, D.-S., Chen, C.-T., Chen, P.-J., and Jou, Y.-S. (2004). Diverse cellular transformation capability of overexpressed genes in human hepatocellular carcinoma. *Biochemical and biophysical research communications* 315, 950-958.

Imai, J., Maruya, M., Yashiroda, H., Yahara, I., and Tanaka, K. (2003). The molecular chaperone Hsp90 plays a role in the assembly and maintenance of the 26S proteasome. *EMBO J* 22, 3557-3567.

Jacob, E., and Unger, R. (2007). A tale of two tails: why are terminal residues of proteins exposed? *Bioinformatics* 23, e225-230.

Jeanes, A., Gottardi, C., and Yap, A. (2008). Cadherins and cancer: how does cadherin dysfunction promote tumor progression? *Oncogene* 27, 6920-6929.

Jessulat, M., Buist, T., Alamgir, M., Hooshyar, M., Xu, J., Aoki, H., Ganoza, M.C., Butland, G., and Golshani, A. (2010). In Vivo Investigation of Protein–Protein Interactions for Helicases Using Tandem Affinity Purification. In *Helicases* (Springer), pp. 99-111.

Jolly, C., and Morimoto, R.I. (2000). Role of the heat shock response and molecular chaperones in oncogenesis and cell death. *J Natl Cancer Inst* 92, 1564-1572.

Kar, G., Gursoy, A., and Keskin, O. (2009). Human cancer protein-protein interaction network: a structural perspective. *PLoS computational biology* 5, e1000601.

Kim, S.Y., Hur, W., Choi, J.E., Kim, D., Wang, J.S., Yoon, H.Y., Piao, L.S., and Yoon, S.K. (2009). Functional characterization of human oncoprotein gankyrin in Zebrafish. *Exp Mol Med* 41, 8-16.

Koch, W.J., Inglese, J., Stone, W.C., and Lefkowitz, R.J. (1993). The binding site for the beta gamma subunits of heterotrimeric G proteins on the beta-adrenergic receptor kinase. *J Biol Chem* 268, 8256-8260.

Kostourou, V., Troy, H., Murray, J.F., Cullis, E.R., Whitley, G.S., Griffiths, J.R., and Robinson, S.P. (2004). Overexpression of dimethylarginine dimethylaminohydrolase enhances tumor hypoxia: an insight into the relationship of hypoxia and angiogenesis in vivo. *Neoplasia* 6, 401-411.

Koutsogiannouli, E., Papavassiliou, A.G., and Papanikolaou, N.A. (2013a). Complexity in cancer biology: is systems biology the answer? *Cancer medicine* 2, 164-177.

Koutsogiannouli, E., Papavassiliou, A.G., and Papanikolaou, N.A. (2013b). Complexity in cancer biology: is systems biology the answer? *Cancer medicine* 2, 164-177.

Kriwacki, R.W., Hengst, L., Tennant, L., Reed, S.I., and Wright, P.E. (1996). Structural studies of p21Waf1/Cip1/Sdi1 in the free and Cdk2-bound state: conformational disorder mediates binding diversity. *Proceedings of the National Academy of Sciences* 93, 11504-11509.

Kuo, T.-C., Chang, P.-Y., Huang, S.-F., Chou, C.-K., and Chao, C.C.-K. (2012). Knockdown of HURP inhibits the proliferation of hepatocellular carcinoma cells via downregulation of gankyrin and accumulation of p53. *Biochemical Pharmacology* 83, 758-768.

Kussie, P.H., Gorina, S., Marechal, V., Elenbaas, B., Moreau, J., Levine, A.J., and Pavletich, N.P. (1996). Structure of the MDM2 oncoprotein bound to the p53 tumor suppressor transactivation domain. *Science* 274, 948-953.

Labelle-Côté, M., Dusseault, J., Ismaïl, S., Picard-Cloutier, A., Siegel, P.M., and Larose, L. (2011). Nck2 promotes human melanoma cell proliferation, migration and invasion in vitro and primary melanoma-derived tumor growth in vivo. *BMC cancer* 11, 443.

Lacy, E.R., Filippov, I., Lewis, W.S., Otieno, S., Xiao, L., Weiss, S., Hengst, L., and Kriwacki, R.W. (2004). p27 binds cyclin-CDK complexes through a sequential mechanism involving binding-induced protein folding. *Nature structural & molecular biology* 11, 358-364.

Lam, H.Y., Kim, P.M., Mok, J., Tonikian, R., Sidhu, S.S., Turk, B.E., Snyder, M., and Gerstein, M.B. (2010). MOTIPS: automated motif analysis for predicting targets of modular protein domains. *BMC bioinformatics* 11, 243.

Landgraf, C., Panni, S., Montecchi-Palazzi, L., Castagnoli, L., Schneider-Mergener, J., Volkmer-Engert, R., and Cesareni, G. (2004). Protein interaction networks by proteome peptide scanning. *PLoS biology* 2, e14.

Latreille, M., and Larose, L. (2006). Nck in a complex containing the catalytic subunit of protein phosphatase 1 regulates eukaryotic initiation factor 2 $\alpha$  signaling and cell survival to endoplasmic reticulum stress. *Journal of Biological Chemistry* 281, 26633-26644.

Le Cam, L., Linares, L.K., Paul, C., Julien, E., Lacroix, M., Hatchi, E., Triboulet, R., Bossis, G., Shmueli, A., and Rodriguez, M.S. (2006). E4F1 is an atypical ubiquitin

ligase that modulates p53 effector functions independently of degradation. *Cell* 127, 775-788.

Le Guennec, J.Y., Ouadid-Ahidouch, H., Soriani, O., Besson, P., Ahidouch, A., and Vandier, C. (2007). Voltage-gated ion channels, new targets in anti-cancer research. *Recent patents on anti-cancer drug discovery* 2, 189-202.

Lehmann, J.r.M., Riethmüller, G., and Johnson, J.P. (1990). Nck, a melanoma cDNA encoding a cytoplasmic protein consisting of the src homology units SH2 and SH3. *Nucleic acids research* 18, 1048.

Li, J., Knobloch, T.J., Kresty, L.A., Zhang, Z., Lang, J.C., Schuller, D.E., and WEGHORST, C.M. (2011). Gankyrin, a biomarker for epithelial carcinogenesis, is overexpressed in human oral cancer. *Anticancer research* 31, 2683-2692.

Li, J., Mahajan, A., and Tsai, M.-D. (2006). Ankyrin repeat: a unique motif mediating protein-protein interactions. *Biochemistry* 45, 15168-15178.

Li, M., Brooks, C.L., Wu-Baer, F., Chen, D., Baer, R., and Gu, W. (2003). Mono- versus polyubiquitination: differential control of p53 fate by Mdm2. *Science* 302, 1972-1975.

Li, Q., Lin, S., Wang, X., Lian, G., Lu, Z., Guo, H., Ruan, K., Wang, Y., Ye, Z., and Han, J. (2009). Axin determines cell fate by controlling the p53 activation threshold after DNA damage. *nature cell biology* 11, 1128-1134.

Lichtarge, O., Bourne, H.R., and Cohen, F.E. (1996). An evolutionary trace method defines binding surfaces common to protein families. *Journal of molecular biology* 257, 342-358.

Link, A.J., Eng, J., Schieltz, D.M., Carmack, E., Mize, G.J., Morris, D.R., Garvik, B.M., and Yates, J.R. (1999). Direct analysis of protein complexes using mass spectrometry. *Nature biotechnology* *17*, 676-682.

Liu, X., Yan, S., Zhou, T., Terada, Y., and Erikson, R.L. (2004). The MAP kinase pathway is required for entry into mitosis and cell survival. *Oncogene* *23*, 763-776.

Liu, Y., Higashitsuji, H., Higashitsuji, H., Itoh, K., Sakurai, T., Koike, K., Hirota, K., Fukumoto, M., and Fujita, J. (2013). Overexpression of gankyrin in mouse hepatocytes induces hemangioma by suppressing factor inhibiting hypoxia-inducible factor-1 (FIH-1) and activating hypoxia-inducible factor-1. *Biochemical and biophysical research communications* *432*, 22-27.

Mackay, J.P., Sunde, M., Lowry, J.A., Crossley, M., and Matthews, J.M. (2007). Protein interactions: is seeing believing? *Trends in biochemical sciences* *32*, 530-531.

Man, J.-H., Liang, B., Gu, Y.-X., Zhou, T., Li, A.-L., Li, T., Jin, B.-F., Bai, B., Zhang, H.-Y., and Zhang, W.-N. (2010a). Gankyrin plays an essential role in Ras-induced tumorigenesis through regulation of the RhoA/ROCK pathway in mammalian cells. *The Journal of clinical investigation* *120*, 2829.

Man, J.H., Liang, B., Gu, Y.X., Zhou, T., Li, A.L., Li, T., Jin, B.F., Bai, B., Zhang, H.Y., Zhang, W.N., *et al.* (2010b). Gankyrin plays an essential role in Ras-induced tumorigenesis through regulation of the RhoA/ROCK pathway in mammalian cells. *J Clin Invest* *120*, 2829-2841.

Mark, W.-Y., Liao, J.C., Lu, Y., Ayed, A., Laister, R., Szymczyna, B., Chakrabartty, A., and Arrowsmith, C.H. (2005). Characterization of segments from the central

region of BRCA1: an intrinsically disordered scaffold for multiple protein–protein and protein–DNA interactions? *Journal of molecular biology* 345, 275-287.

Meng, Y., He, L., Guo, X., Tang, S., Zhao, X., Du, R., Jin, J., Bi, Q., Li, H., and Nie, Y. (2010a). Gankyrin promotes the proliferation of human pancreatic cancer. *Cancer letters* 297, 9-17.

Meng, Y., He, L., Guo, X., Tang, S., Zhao, X., Du, R., Jin, J., Bi, Q., Li, H., Nie, Y., *et al.* (2010b). Gankyrin promotes the proliferation of human pancreatic cancer. *Cancer letters* 297, 9-17.

Merkley, N., and Shaw, G.S. (2004). Solution structure of the flexible class II ubiquitin-conjugating enzyme Ubc1 provides insights for polyubiquitin chain assembly. *Journal of Biological Chemistry* 279, 47139-47147.

Mine, H., Sakurai, T., Kashida, H., Matsui, S., Nishida, N., Nagai, T., Hagiwara, S., Watanabe, T., and Kudo, M. (2013). Association of Gankyrin and stemness factor expression in human colorectal cancer. *Digestive diseases and sciences* 58, 2337-2344.

Mosavi, L.K., Cammett, T.J., Desrosiers, D.C., and Peng, Z.y. (2004). The ankyrin repeat as molecular architecture for protein recognition. *Protein Science* 13, 1435-1448.

Nagao, T., Higashitsuji, H., Nonoguchi, K., Sakurai, T., Dawson, S., Mayer, R.J., Itoh, K., and Fujita, J. (2003). MAGE-A4 interacts with the liver oncoprotein gankyrin and suppresses its tumorigenic activity. *Journal of Biological Chemistry* 278, 10668-10674.

Nakamura, Y., Nakano, K., Umehara, T., Kimura, M., Hayashizaki, Y., Tanaka, A., Horikoshi, M., Padmanabhan, B., and Yokoyama, S. (2007a). Structure of the oncoprotein gankyrin in complex with S6 ATPase of the 26S proteasome. *Structure* 15, 179-189.

Nakamura, Y., Nakano, K., Umehara, T., Kimura, M., Hayashizaki, Y., Tanaka, A., Horikoshi, M., Padmanabhan, B., and Yokoyama, S. (2007b). Structure of the oncoprotein gankyrin in complex with S6 ATPase of the 26S proteasome. *Structure* 15, 179-189.

Naylor, L.H. (1999). Reporter gene technology: the future looks bright. *Biochemical pharmacology* 58, 749-757.

Neduva, V., and Russell, R.B. (2006). DILIMOT: discovery of linear motifs in proteins. *Nucleic acids research* 34, W350-W355.

Obenauer, J.C., Cantley, L.C., and Yaffe, M.B. (2003). Scansite 2.0: Proteome-wide prediction of cell signaling interactions using short sequence motifs. *Nucleic acids research* 31, 3635-3641.

Oeljeklaus, S., Meyer, H.E., and Warscheid, B. (2009). New dimensions in the study of protein complexes using quantitative mass spectrometry. *FEBS letters* 583, 1674-1683.

Ofran, Y., and Rost, B. (2007). Protein-protein interaction hotspots carved into sequences. *PLoS computational biology* 3, e119.

Ortiz, C.M., Ito, T., Tanaka, E., Tsunoda, S., Nagayama, S., Sakai, Y., Higashitsuji, H., Fujita, J., and Shimada, Y. (2008). Gankyrin oncoprotein overexpression as a

critical factor for tumor growth in human esophageal squamous cell carcinoma and its clinical significance. *International Journal of Cancer* 122, 325-332.

Ortiz, C.M., Tsunoda, S., Tanaka, E., Ito, T., Higashitsuji, H., Fujita, J., Shimada, Y., and Sakai, Y. (2006). Overexpression of the novel oncoprotein gankyrin in human esophageal squamous cell carcinoma (ESCC) and its clinical significance. *Proceedings of the American Association for Cancer Research* 2006, 997.

Pagel, P., Kovac, S., Oesterheld, M., Brauner, B., Dunger-Kaltenbach, I., Frishman, G., Montrone, C., Mark, P., Stümpflen, V., and Mewes, H.-W. (2005). The MIPS mammalian protein–protein interaction database. *Bioinformatics* 21, 832-834.

Pang, Q., Keeble, W., Christianson, T.A., Faulkner, G.R., and Bagby, G.C. (2001). FANCC interacts with Hsp70 to protect hematopoietic cells from IFN-gamma/TNF-alpha-mediated cytotoxicity. *EMBO J* 20, 4478-4489.

Patil, A., Kinoshita, K., and Nakamura, H. (2010a). Domain distribution and intrinsic disorder in hubs in the human protein–protein interaction network. *Protein Science* 19, 1461-1468.

Patil, A., Kinoshita, K., and Nakamura, H. (2010b). Hub promiscuity in protein–protein interaction networks. *International journal of molecular sciences* 11, 1930-1943.

Patil, A., and Nakamura, H. (2006). Disordered domains and high surface charge confer hubs with the ability to interact with multiple proteins in interaction networks. *FEBS letters* 580, 2041-2045.

Patil, A., and Nakamura, H. (2007). The role of charged surface residues in the binding ability of small hubs in protein-protein interaction networks. *Biophysics* 3, 27-35.

Pavelka, A., Chovancova, E., and Damborsky, J. (2009). HotSpot Wizard: a web server for identification of hot spots in protein engineering. *Nucleic acids research* 37, W376-W383.

Perkins, J.R., Diboun, I., Dessailly, B.H., Lees, J.G., and Orengo, C. (2010). Transient protein-protein interactions: structural, functional, and network properties. *Structure* 18, 1233-1243.

Petrova, D.T., Asif, A.R., Armstrong, V.W., Dimova, I., Toshev, S., Yaramov, N., Oellerich, M., and Toncheva, D. (2008a). Expression of chloride intracellular channel protein 1 (CLIC1) and tumor protein D52 (TPD52) as potential biomarkers for colorectal cancer. *Clinical biochemistry* 41, 1224-1236.

Petrova, D.T., Asif, A.R., Armstrong, V.W., Dimova, I., Toshev, S., Yaramov, N., Oellerich, M., and Toncheva, D. (2008b). Expression of chloride intracellular channel protein 1 (CLIC1) and tumor protein D52 (TPD52) as potential biomarkers for colorectal cancer. *Clinical biochemistry* 41, 1224-1236.

Phizicky, E.M., and Fields, S. (1995). Protein-protein interactions: methods for detection and analysis. *Microbiological reviews* 59, 94-123.

Pierce, M.M., Raman, C., and Nall, B.T. (1999). Isothermal titration calorimetry of protein-protein interactions. *Methods* 19, 213-221.

Polakis, P. (2000). Wnt signaling and cancer. *Genes & development* 14, 1837-1851.

Rodriguez, L.G., Wu, X., and Guan, J.L. (2005). Wound-healing assay. *Methods in molecular biology* 294, 23-29.

Roelofs, J., Park, S., Haas, W., Tian, G., McAllister, F.E., Huo, Y., Lee, B.-H., Zhang, F., Shi, Y., and Gygi, S.P. (2009). Chaperone-mediated pathway of proteasome regulatory particle assembly. *Nature* 459, 861-865.

Rosato, E. (2007). *Circadian rhythms: Methods and protocols*, Vol 362 (Springer).

Royer, C.A., and Scarlata, S.F. (2008). Fluorescence approaches to quantifying biomolecular interactions. *Methods in enzymology* 450, 79-106.

Setti, M., Savalli, N., Osti, D., Richichi, C., Angelini, M., Brescia, P., Fornasari, L., Carro, M.S., Mazzanti, M., and Pelicci, G. (2013). Functional role of clic1 ion channel in Glioblastoma-Derived Stem/Progenitor cells. *Journal of the National Cancer Institute* 105, 1644-1655.

Sheinerman, F.B., Norel, R., and Honig, B. (2000). Electrostatic aspects of protein-protein interactions. *Current opinion in structural biology* 10, 153-159.

Shikama, N., Lee, C.-W., France, S., Delavaine, L., Lyon, J., Krstic-Demonacos, M., and La Thangue, N.B. (1999). A novel cofactor for p300 that regulates the p53 response. *Molecular cell* 4, 365-376.

Singh Gautam, A.K., Balakrishnan, S., and Venkatraman, P. (2012). Direct ubiquitin independent recognition and degradation of a folded protein by the eukaryotic proteasomes-origin of intrinsic degradation signals. *PloS one* 7, e34864.

Song, X., Wang, J., Zheng, T., Song, R., Liang, Y., Bhatta, N., Yin, D., Pan, S., Liu, J., and Jiang, H. (2013). LBH589 Inhibits proliferation and metastasis of

hepatocellular carcinoma via inhibition of gankyrin/stat3/akt pathway. *Molecular cancer* 12, 114.

Sprinzak, E., Sattath, S., and Margalit, H. (2003). How reliable are experimental protein-protein interaction data? *Journal of molecular biology* 327, 919-923.

Stark, C., Breitkreutz, B.-J., Reguly, T., Boucher, L., Breitkreutz, A., and Tyers, M. (2006). BioGRID: a general repository for interaction datasets. *Nucleic acids research* 34, D535-D539.

Tang, S., Yang, G., Meng, Y., Du, R., Li, X., Fan, R., Zhao, L., Bi, Q., Jin, J., and Gao, L. (2010a). Overexpression of a novel gene gankyrin correlates with the malignant phenotype of colorectal cancer. *Cancer biology & therapy* 9, 88-95.

Tang, S., Yang, G., Meng, Y., Du, R., Li, X., Fan, R., Zhao, L., Bi, Q., Jin, J., Gao, L., *et al.* (2010b). Overexpression of a novel gene gankyrin correlates with the malignant phenotype of colorectal cancer. *Cancer biology & therapy* 9, 88-95.

Tang, X., and Bruce, J.E. (2009). Chemical cross-linking for protein-protein interaction studies. In *Mass Spectrometry of Proteins and Peptides* (Springer), pp. 283-293.

Taylor, I.W., Linding, R., Warde-Farley, D., Liu, Y., Pesquita, C., Faria, D., Bull, S., Pawson, T., Morris, Q., and Wrana, J.L. (2009). Dynamic modularity in protein interaction networks predicts breast cancer outcome. *Nature biotechnology* 27, 199-204.

Thomas, P.D., Kejariwal, A., Campbell, M.J., Mi, H., Diemer, K., Guo, N., Ladunga, I., Ulitsky-Lazareva, B., Muruganujan, A., Rabkin, S., *et al.* (2003). PANTHER: a

browsable database of gene products organized by biological function, using curated protein family and subfamily classification. *Nucleic Acids Res* 31, 334-341.

Tsai, C.-J., Ma, B., and Nussinov, R. (2009). Protein–protein interaction networks: how can a hub protein bind so many different partners? *Trends in biochemical sciences* 34, 594-600.

Valenzuela, S.M., Martin, D.K., Por, S.B., Robbins, J.M., Warton, K., Bootcov, M.R., Schofield, P.R., Campbell, T.J., and Breit, S.N. (1997). Molecular cloning and expression of a chloride ion channel of cell nuclei. *Journal of Biological Chemistry* 272, 12575-12582.

Valenzuela, S.M., Mazzanti, M., Tonini, R., Qiu, M.R., Warton, K., Musgrove, E.A., Campbell, T.J., and Breit, S.N. (2000). The nuclear chloride ion channel NCC27 is involved in regulation of the cell cycle. *The Journal of physiology* 529, 541-552.

Van Maanen, J., Retel, J., De Vries, J., and Pinedo, H. (1988). Mechanism of action of antitumor drug etoposide: a review. *Journal of the National Cancer Institute* 80, 1526-1533.

van Meerloo, J., Kaspers, G.J., and Cloos, J. (2011). Cell sensitivity assays: the MTT assay. In *Cancer Cell Culture* (Springer), pp. 237-245.

Venkatraman, P., Balakrishnan, S., Rao, S., Hooda, Y., and Pol, S. (2009). A sequence and structure based method to predict putative substrates, functions and regulatory networks of endo proteases. *PloS one* 4, e5700.

Vidal, M., Cusick, M.E., and Barabasi, A.-L. (2011). Interactome networks and human disease. *Cell* 144, 986-998.

Von Mering, C., Jensen, L.J., Snel, B., Hooper, S.D., Krupp, M., Foglierini, M., Jouffre, N., Huynen, M.A., and Bork, P. (2005). STRING: known and predicted protein–protein associations, integrated and transferred across organisms. *Nucleic acids research* 33, D433-D437.

Wadhawan, V., Kolhe, Y.A., Sangith, N., Gautam, A.K., and Venkatraman, P. (2012). From prediction to experimental validation: desmoglein 2 is a functionally relevant substrate of matriptase in epithelial cells and their reciprocal relationship is important for cell adhesion. *Biochem J* 447, 61-70.

Wang, L., He, S., Tu, Y., Ji, P., Zong, J., Zhang, J., Feng, F., Zhao, J., Zhang, Y., and Gao, G. (2012a). Elevated expression of chloride intracellular channel 1 is correlated with poor prognosis in human gliomas. *J Exp Clin Cancer Res* 31, 44.

Wang, P., Zhang, C., Yu, P., Tang, B., Liu, T., Cui, H., and Xu, J. (2012b). Regulation of colon cancer cell migration and invasion by CLIC1-mediated RVD. *Molecular and cellular biochemistry* 365, 313-321.

Wang, W., Xu, X., Wang, W., Shao, W., Li, L., Yin, W., Xiu, L., Mo, M., Zhao, J., and He, Q. (2011). The expression and clinical significance of CLIC1 and HSP27 in lung adenocarcinoma. *Tumor Biology* 32, 1199-1208.

Weatheritt, R.J., Jehl, P., Dinkel, H., and Gibson, T.J. (2012). iELM—a web server to explore short linear motif-mediated interactions. *Nucleic acids research* 40, W364-W369.

Wilkins, S.E., Karttunen, S., Hampton-Smith, R.J., Murchland, I., Chapman-Smith, A., and Peet, D.J. (2012). Factor Inhibiting HIF (FIH) Recognizes Distinct Molecular

Features within Hypoxia-inducible Factor- $\alpha$  (HIF- $\alpha$ ) versus Ankyrin Repeat Substrates. *Journal of Biological Chemistry* 287, 8769-8781.

Wilson, M.A., and Brunger, A.T. (2000). The 1.0 Å crystal structure of Ca<sup>2+</sup>-bound calmodulin: an analysis of disorder and implications for functionally relevant plasticity. *Journal of molecular biology* 301, 1237-1256.

Yang, Y., Zhang, C., Li, L., Gao, Y., Luo, X., Zhang, Y., Liu, W., and Fei, Z. (2012). Up-regulated oncoprotein P28GANK correlates with proliferation and poor prognosis of human glioma. *World journal of surgical oncology* 10, 1-7.

Yoon, G.-S., Lee, H., Jung, Y., Yu, E., Moon, H.-B., Song, K., and Lee, I. (2000). Nuclear matrix of calreticulin in hepatocellular carcinoma. *Cancer research* 60, 1117-1120.

Zamanian, M., Veerakumarasivam, A., Abdullah, S., and Rosli, R. (2013). Calreticulin and cancer. *Pathology & Oncology Research* 19, 149-154.

Zhen, C., Chen, L., Zhao, Q., Liang, B., Gu, Y., Bai, Z., Wang, K., Xu, X., Han, Q., and Fang, D. (2012a). Gankyrin promotes breast cancer cell metastasis by regulating Rac1 activity. *Oncogene*.

Zhen, C., Chen, L., Zhao, Q., Liang, B., Gu, Y.X., Bai, Z.F., Wang, K., Xu, X., Han, Q.Y., Fang, D.F., *et al.* (2012b). Gankyrin promotes breast cancer cell metastasis by regulating Rac1 activity. *Oncogene*.

Zuchero, J.B., Coutts, A.S., Quinlan, M.E., La Thangue, N.B., and Mullins, R.D. (2009). p53-cofactor JMY is a multifunctional actin nucleation factor. *Nature cell biology* 11, 451-459.

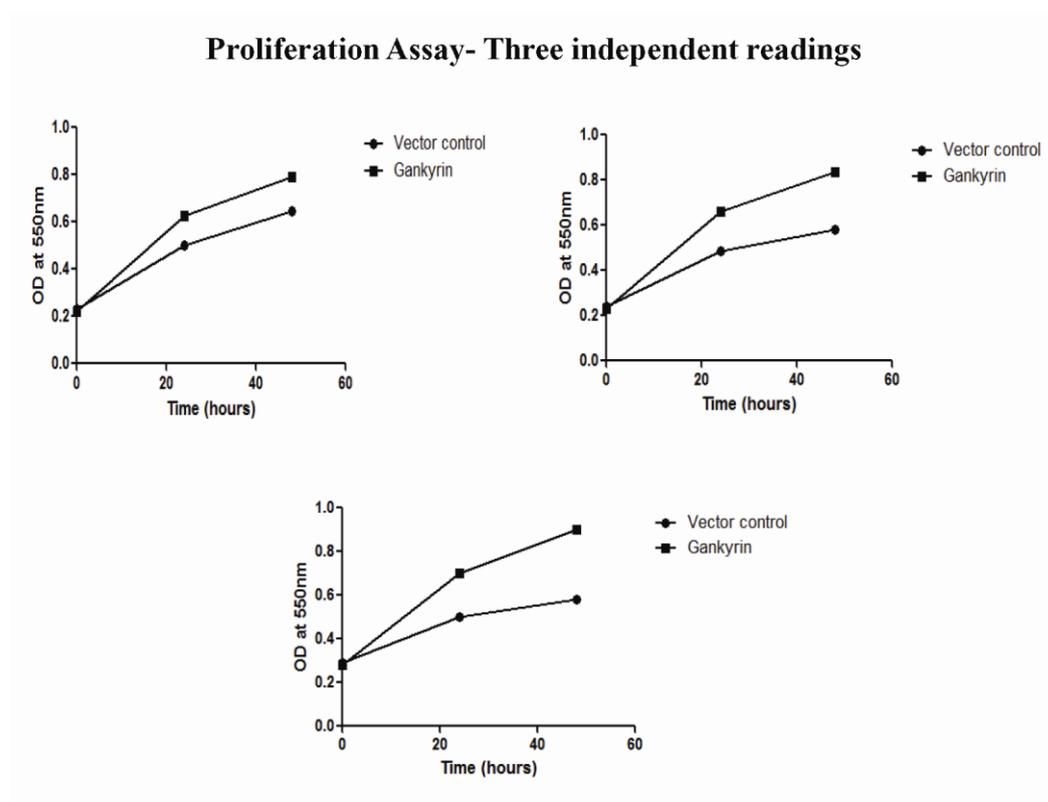
## ***Appendix***

## Reviewer's Comment

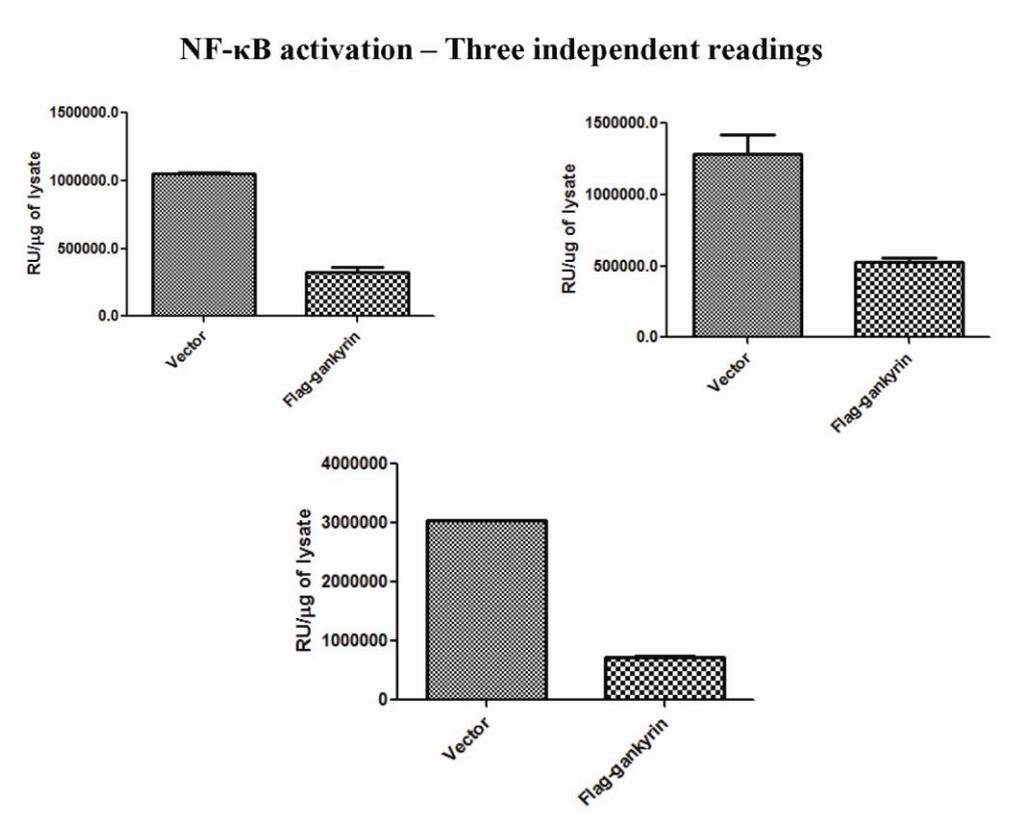
P83 Fig. 4.1 and numerous places elsewhere in thesis (e.g p 86 legend 4.2). The candidate presents means  $\pm$ SD on one experiment done in duplicate. This is not suitable for a number of reasons. First, one should do analysis on independent experiments. Second, one cannot derive meaningful SDs from two samples. In some cases in the thesis, it is really not really a problem, especially when validating previously described phenomena. I suggest that the candidate scrap SDs in such cases and replace with the raw data of the replicates.

Independent readings for assays in Fig. 4.2 are plotted below.

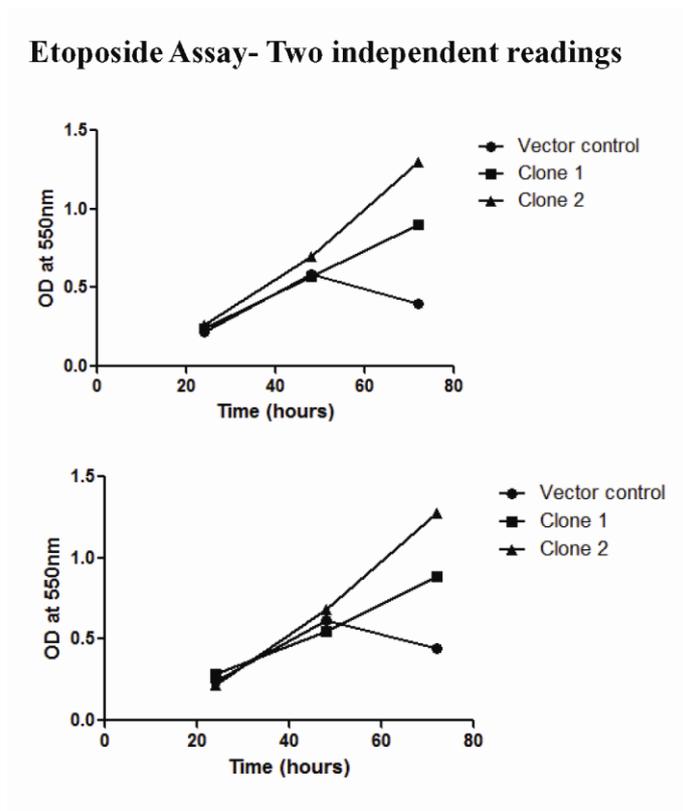
**Fig. 4.2A- Cell Proliferation Assay**



**Fig. 4.2B- NF- $\kappa$ B Activation**



**Fig. 4.2C- Etoposide Induced Cell Viability Assay**



## ***Reprints Of Published Articles***



## Identification of a novel ATPase activity in 14-3-3 proteins – Evidence from enzyme kinetics, structure guided modeling and mutagenesis studies



Manoj P. Ramteke<sup>a</sup>, Pradnya Shelke<sup>a</sup>, Vidhya Ramamoorthy<sup>a</sup>, Arun Kumar Somavarapu<sup>a</sup>, Amit Kumar Singh Gautam<sup>a</sup>, Padma P. Nanaware<sup>a</sup>, Sudheer Karanam<sup>b</sup>, Sami Mukhopadhyay<sup>b</sup>, Prasanna Venkatraman<sup>a,\*</sup>

<sup>a</sup>Advanced Centre for Treatment, Research and Education in Cancer (ACTREC), Tata Memorial Centre (TMC), Kharghar, Navi Mumbai 410210, India

<sup>b</sup>Vlife Sciences Technologies Pvt. Ltd., 2nd Floor, Plot No-05, Ram Indu Park, Baner Road, Pune 411045, India

### ARTICLE INFO

#### Article history:

Received 29 July 2013

Revised 16 October 2013

Accepted 5 November 2013

Available online 20 November 2013

Edited by Peter Brzezinski

#### Keywords:

14-3-3

ATP hydrolysis

Chaperone

Docking

Mutation

### ABSTRACT

**14-3-3 Proteins bind phosphorylated sequences in proteins and regulate multiple cellular functions. For the first time, we show that pure recombinant human 14-3-3  $\zeta$ ,  $\gamma$ ,  $\epsilon$  and  $\tau$  isoforms hydrolyze ATP with similar  $K_m$  and  $k_{cat}$  values. In sharp contrast the sigma isoform has no detectable activity. Docking studies identify two putative binding pockets in 14-3-3 zeta. Mutation of D124A in the amphipathic pocket enhances binding affinity and catalysis. Mutation of a critical Arg (R55A) at the dimer interface in zeta reduces binding and decreases catalysis. These experimental results coincide with a binding pose at the dimer interface. This newly identified function could be a moon lighting function in some of these isoforms.**

© 2013 Federation of European Biochemical Societies. Published by Elsevier B.V. All rights reserved.

### 1. Introduction

14-3-3 proteins ( $\beta$ ,  $\gamma$ ,  $\epsilon$ ,  $\eta$ ,  $\sigma$ ,  $\tau$  and  $\zeta$  isoforms) are an important family of highly conserved dimeric proteins. They bind to Ser/Thr phosphorylated proteins through two major consensus motifs, RSXpSXP (mode I) and RXY/FXpSXP (mode II), where pS represents phosphoserine [1–3] and a third minor binding motif pS/pT(X 1–2)-COOH (mode III) where pT represent phosphothreonine [4]. These proteins are involved in cell metabolism, signal transduction, cell cycle control, apoptosis, protein trafficking, transcription and stress response as well as in malignant transformation [5–9]. These functions have been attributed primarily to their ability to bind to

phosphorylated sequences within the client proteins [10–13]. In addition to these well known adaptor functions, 14-3-3 $\tau$  was shown to possess an ATP/ADP exchange activity [14]. In rat liver, mitochondrial import function of a cytosolic 14-3-3 seemed to require energy from ATP hydrolysis [15–18]. Of late *Drosophila* 14-3-3 $\zeta$  in conjunction with Hsp70 was shown to solubilize aggregated proteins in an ATP dependent manner although whether this function is dependent on energy from ATP hydrolysis has not been demonstrated so far [19]. This disaggregating function seems analogous to yeast Hsp104 which is probably the only other chaperone well known to solubilize preformed aggregates. No such activity has been identified in any mammalian proteins but mammalian 14-3-3 $\zeta$  seems to affect aggregation of tau and huntingtin [20–22].

These observations caught our attention as these ATP dependent activity of 14-3-3 is not a major focus of research and any enzymatic activity of 14-3-3 seems controversial [23]. Prompted, we asked if human 14-3-3 $\zeta$  has any detectable ATPase activity. If so identifying residues involved in ATP binding/hydrolysis is likely to better define the enzymatic property associated with the 14-3-3 protein which is otherwise well known for its scaffold and chaperone-like functions. Here we report the in vitro characterization of ATPase activity of 14-3-3 $\zeta$  and identify residues that are important

*Abbreviations:* WT, wild-type; Ni-NTA, nickel-nitriloacetic acid; MALDI, matrix assisted laser desorption ionization; TOF, tandem time of flight; PDB, protein data bank; ATP- $\gamma$ -S, adenosine 5'-(3-thiotriphosphate); AMP-PCP, adenyllylmethylenediphosphonate; HEPES, 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid; Ci, curie; RMSD, root mean-square deviation; PLP, piecewise linear pairwise potential

\* Corresponding author. Address: KS-244, ACTREC, Tata Memorial Centre (TMC), Kharghar, Navi Mumbai 410210, India. Fax: +91 022 27405085.

E-mail address: [vprasanna@actrec.gov.in](mailto:vprasanna@actrec.gov.in) (P. Venkatraman).

for binding/hydrolysis. By virtue of residue/structure conservation 14-3-3  $\gamma$ ,  $\epsilon$  and  $\tau$  isoforms also show similar ATPase activity while the sigma isoform which does not carry an ATP sensor sequence lacks any detectable activity.

## 2. Results and discussion

### 2.1. ATPase activity of human 14-3-3 zeta

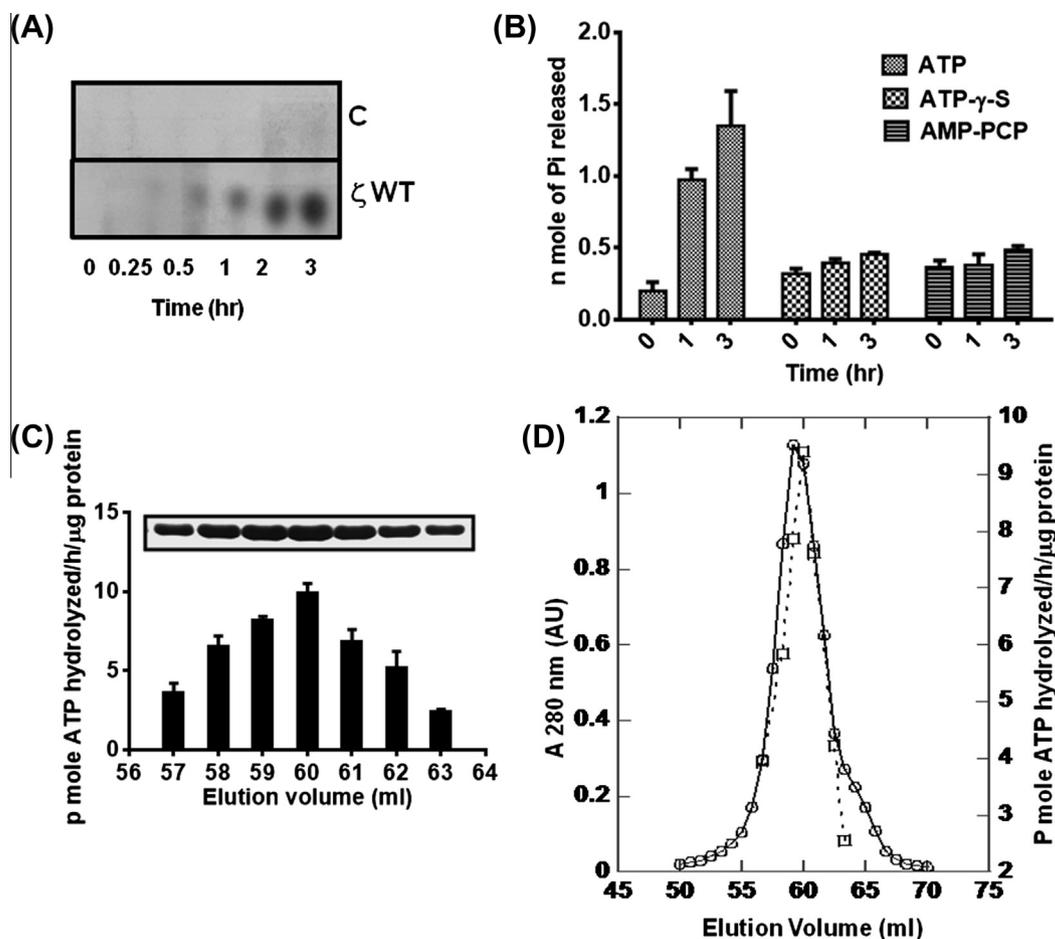
We expressed and purified 14-3-3 by affinity chromatography and gel filtration. Identity of the recombinant protein was confirmed by MALDI-TOF-TOF (data not shown) and by nano LC-MS<sup>E</sup> (Supplementary Table T1) as well as by Western blotting using 14-3-3 $\zeta$  specific antibody (Supplementary Fig. S1(A)). Mass spectrometric analysis reveals 89.39% sequence coverage with no contaminating proteins from *Escherichia coli* in which the protein was expressed (Supplementary Table T1). Three other identifications also belong to 14-3-3 proteins (Supplementary Table T1). Two other proteins identified are very different from 14-3-3 $\zeta$  in molecular weight and are undetectable by standard techniques. Considering 14-3-3 $\zeta$  used here is a recombinant protein, these identifications are clearly false positive identifications.

In fusion His-tag was cleaved and confirmed by western blotting using anti His-antibody (Supplementary Fig. S1(B)). Dimeric nature was confirmed by native PAGE (Supplementary Fig. S1(C)).

Ability of this pure 14-3-3 $\zeta$  WT to hydrolyze ATP was then tested using  $\gamma$ -<sup>32</sup>P labeled ATP (Fig. 1(A)). 14-3-3 $\zeta$  WT hydrolyses ATP and releases inorganic phosphate (Pi) in a time dependent manner. Similar results were obtained with the calorimetric assay (Fig. 1(B)) [24]. Absorbance values in presence of two non-hydrolyzable analogs ATP- $\gamma$ -S and AMP-PCP were not very different in the presence or absence of 14-3-3 $\zeta$  indicating that the calorimetric reaction is a true reflection of gamma phosphate hydrolysis (Fig. 1(B)). Due to the inherent low signal to noise ratio of the calorimetric assay, we standardized a luminescence based assay called the ADP-Glo<sup>TM</sup> max from Promega which has higher sensitivity, a large dynamic range and is designed to eliminate background from unhydrolyzed ATP. Using this assay we followed ATP hydrolysis of each fraction of the single peak of 14-3-3 $\zeta$  that elutes out from the gel filtration column and correlated it with UV absorption (Fig. 1(C and D)). Results show that activity is coincident with the protein fractions confirming that 14-3-3 $\zeta$  has detectable intrinsic ATPase activity.

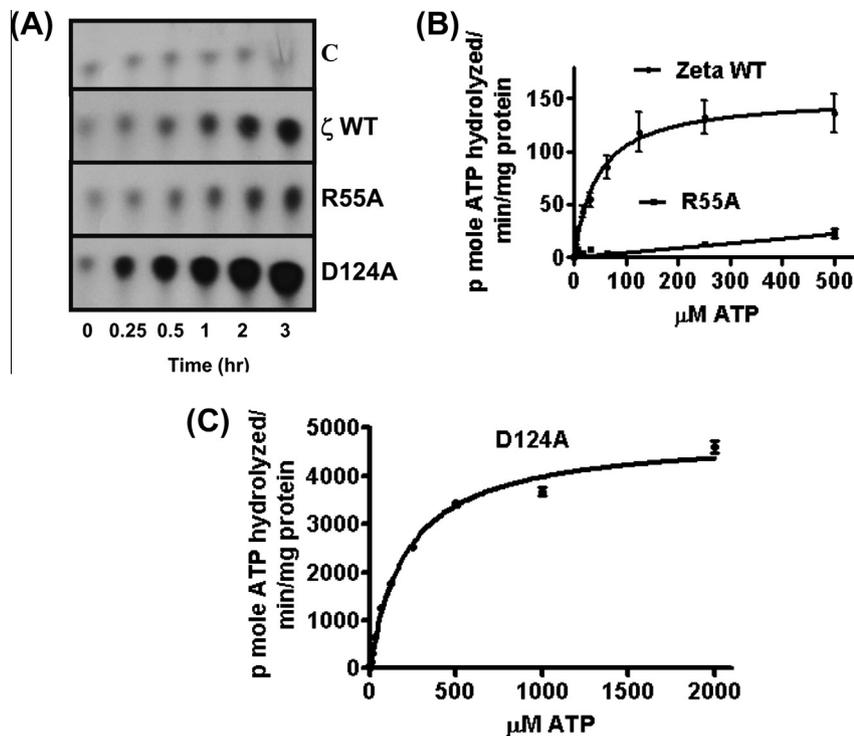
### 2.2. Prediction of putative residues involved in ATP binding/hydrolysis

To obtain clues towards the structural motif in 14-3-3 $\zeta$  involved in ATP binding/hydrolysis, we performed BLAST search of the protein sequence of 14-3-3 $\zeta$  (YWHAZ) against other ATP binding proteins (data not shown). No consensus motif such as the AAA



**Fig. 1.** Demonstration of ATPase activity in 14-3-3 $\zeta$  (A) Inorganic phosphate ( $\gamma$ -<sup>32</sup>Pi) released from ( $\gamma$ -<sup>32</sup>P) ATP hydrolysis by 14-3-3 $\zeta$  WT at different time points was monitored by PEI-TLC. Lane C represents control ( $\gamma$ -<sup>32</sup>P) ATP without any protein. (B) Pi release by 14-3-3 $\zeta$  WT in presence of cold ATP, ATP- $\gamma$ -S or AMP-PCP was monitored at different time points using malachite green. Data are represented as mean  $\pm$  S.D. ( $n = 3$ ). (C) ATPase activity of each fraction of 14-3-3 $\zeta$  WT eluted from gel filtration. Proteins from each elution fraction were run on 12% SDS-PAGE (top panel). (D) Correlation between ATPase activity and gel filtration profile of 14-3-3 $\zeta$  WT. Data represents gel filtration profile of 14-3-3 $\zeta$  WT (continuous line) and corresponding ATPase activity (dotted line).





**Fig. 3.** Mutations enhance or decrease ATPase activity of 14-3-3 $\zeta$  (A) ATPase activity of R55A and D124A mutants was monitored using radiolabeled ( $\gamma$ - $^{32}\text{P}$ ) ATP and compared with the WT protein. Lane C represents control ( $\gamma$ - $^{32}\text{P}$ ) ATP without any protein. (B) ATPase activity of 14-3-3 $\zeta$  WT, R55A and (C) D124A mutants was measured using ADP-Glo™ max assay. Two independent experiments in triplicates were conducted. Data are represented as mean  $\pm$  S.E.M. S.E.M. – standard error of mean.

**Table 1**

$V_{\max}$ ,  $K_m$  and  $k_{\text{cat}}$  for 14-3-3 active isoforms, D124A mutant and Hsp 70.

Sr. no.	Protein	$V_{\max}$ (p mole ATP hydrolyzed/min/mg protein)	$K_m$ ( $\mu\text{M}$ )	$k_{\text{cat}}$ ( $\text{min}^{-1}$ )
1	14-3-3 $\zeta$ WT	152 $\pm$ 10.19	44.33 $\pm$ 10.05	0.0087 $\pm$ 0.00058
2	14-3-3 $\gamma$ WT	149.3 $\pm$ 12.55	46.81 $\pm$ 13.15	0.0085 $\pm$ 0.00072
3	14-3-3 $\tau$ WT	152.4 $\pm$ 8.18	22.57 $\pm$ 4.72	0.0087 $\pm$ 0.00047
4	14-3-3 $\epsilon$ WT	146.9 $\pm$ 8.9	27.31 $\pm$ 6.22	0.0084 $\pm$ 0.00051
5	D124A mutant	4830 $\pm$ 80.22	214.5 $\pm$ 11.53	0.28 $\pm$ 0.0046
6	Hsp 70	2335 $\pm$ 54.22	51.07 $\pm$ 4.51	0.16 $\pm$ 0.0038

tions did not affect ability of any of these proteins to bind peptides (Supplementary Fig. S2(B)).

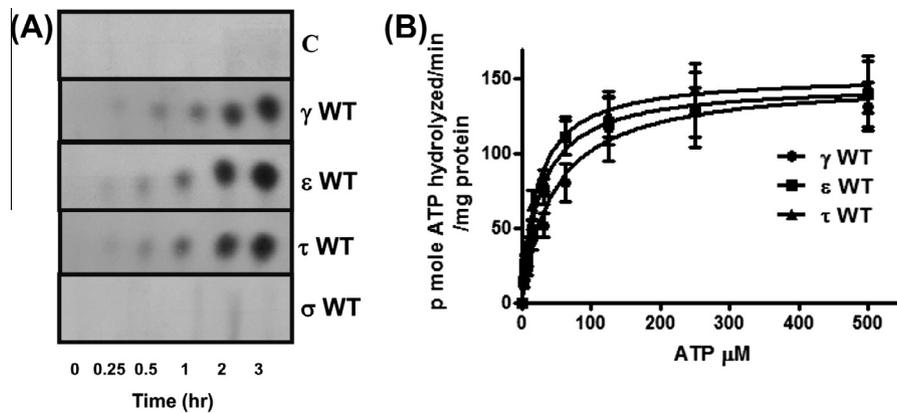
It is to be noted that 14-3-3 structure is very different from these chaperones and carries no structural or sequence homology with these or other ATPase family members. The ATPase activity of 14-3-3 $\zeta$  although slow, is nevertheless close to the classical chaperone Hsp 90 and the activity of D124A mutant ( $k_{\text{cat}}$  0.28  $\text{min}^{-1}$ ) is even better than that of Hsp 70 ( $k_{\text{cat}}$  0.16  $\text{min}^{-1}$ ). This enhancement in activity by single amino acid substitution is rare and provides the strongest support for the intrinsic ATPase activity of 14-3-3 $\zeta$ . Literature survey suggested that Hsp 104 protein carries RR and GAR motif which act as sensors for ATP [33]. R<sub>55</sub> at the dimer interface of 14-3-3 $\zeta$  is part of an identical Short Linear Sequence Motif G<sub>53</sub>A<sub>54</sub>R<sub>55</sub>R<sub>56</sub>. By extrapolation, one may presume that R<sub>55</sub> in this motif is probably acting as an ATP sensor in 14-3-3 $\zeta$ .

Prompted by this possibility we looked for the presence of GARR sequence in other 14-3-3 isoforms. We compared the reported structural alignment of 14-3-3 isoforms and found that while D124 is present in all isoforms, R<sub>55</sub> is conserved in all but the sigma protein (GGQR Supplementary Fig. S3). We expressed and purified 14-3-3  $\gamma$ ,  $\epsilon$ ,  $\tau$  and the  $\sigma$  isoform and found that all except sigma show similar ATPase activity measured using radioactive ATP and the ADP-Glo™ assay (Fig. 4(A and B)). The  $K_m$  and

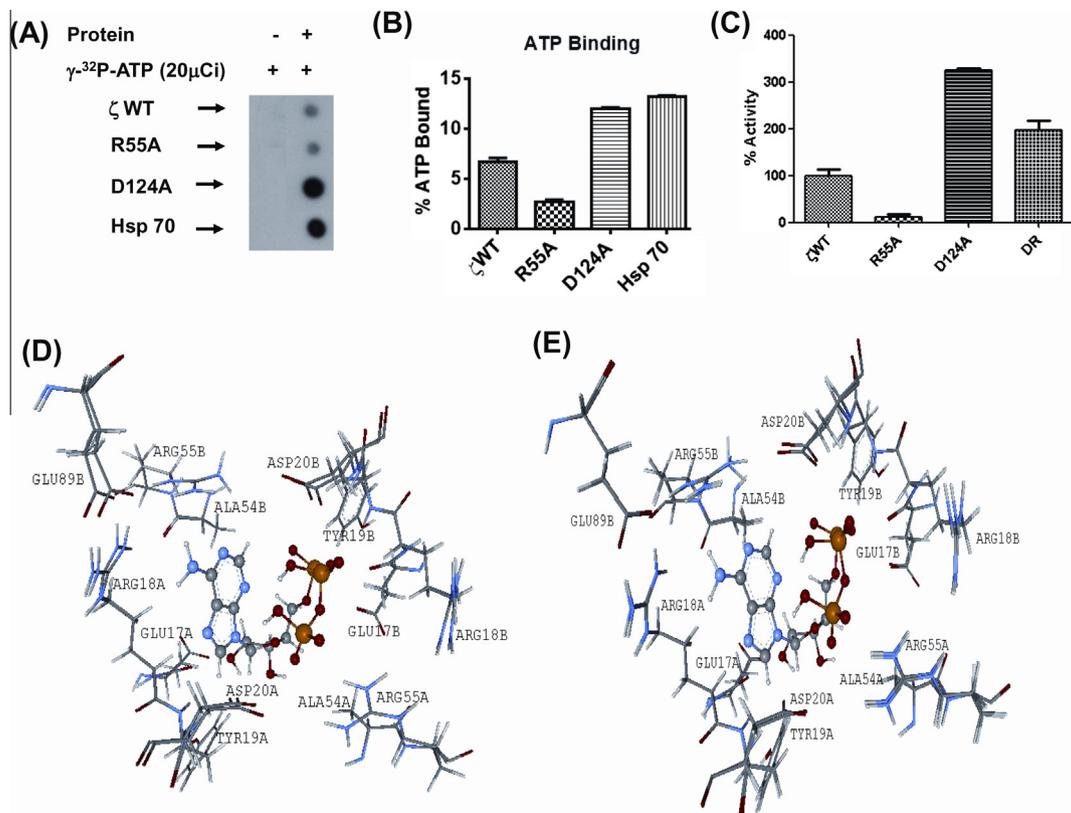
$k_{\text{cat}}$  values are also similar in the active isoforms (Table 1). Absence of the critical Arg may be the reason for the lack of activity in the sigma isoform. ATPase activity seen in these isoforms could be a moon lighting function. Such a function seems to correlate with the presence of the same ATP sensor sequence motif G(A/G)RR. This however needs to be tested by mutagenesis. Lack of ATPase activity in the sigma isoform which was expressed and purified identical to 14-3-3 $\zeta$  isoform, further supports our contention that observed ATPase activity in 14-3-3 preparations is intrinsic to the respective isoforms.

#### 2.4. Estimation of binding affinity of ATP to 14-3-3 zeta and the mutants

It is to be noted that  $V_{\max}$  of D124A mutant is  $\sim$ 30 times more than the WT protein and the  $K_m$  for ATP hydrolysis is approximately four times higher (Table 1). While  $K_m$  is routinely considered as a measure of binding affinity, the relationship holds true only under specific conditions. To obtain a measure of binding affinity of WT 14-3-3 $\zeta$  and two mutant proteins for ATP, we incubated  $\gamma$ - $^{32}\text{P}$  radiolabelled ATP with each protein and aliquots were spotted on to nitrocellulose membrane. ATP bound to protein was visualized using autoradiogram. Hsp 70 was used as a control (Fig. 5(A)). Results confirm that these proteins bind to ATP. ATP



**Fig. 4.** ATPase activity of other 14-3-3 isoforms (A) Inorganic phosphate ( $\gamma$ - $^{32}\text{P}$ ) released from ( $\gamma$ - $^{32}\text{P}$ ) ATP hydrolysis by 14-3-3 isoforms at different time points was monitored by PEI-TLC. There is no detectable  $\gamma$ - $^{32}\text{P}$  in lanes where ATP was incubated with the sigma isoform. Lane C represents control ( $\gamma$ - $^{32}\text{P}$ ) ATP without any protein. (B) ATP hydrolysis of active 14-3-3  $\gamma$ ,  $\epsilon$  and  $\tau$  was monitored by ADP-Glo™ max ATPase assay. Two independent experiments in triplicates were conducted. Data are represented as mean  $\pm$  S.E.M. S.E.M. – standard error of mean.



**Fig. 5.** Effect of mutations on ATP binding (A) ATP bound protein was trapped on nitrocellulose membrane (NC). 14-3-3 $\zeta$  WT or R55A or D124A mutants or Hsp 70 was incubated with radiolabeled ( $\gamma$ - $^{32}\text{P}$ ) ATP and spotted on NC. Free ATP was washed away (B) protein bound ATP ( $\gamma$ - $^{32}\text{P}$ ) was separated using desalting spin column and quantitated using liquid scintillation counter. Data are represented as mean  $\pm$  S.D. ( $n = 3$ ). S.D. – standard deviation. (C) 14-3-3 $\zeta$  WT, R55A, D124A and DR (R55A, D124A) double mutant was subjected for ATPase assay in the presence of 1 mM ATP (Sigma). Pi released was monitored using calorimetry assay. R55A mutation in D124A background resulted in reduction of ATPase activity of D124A mutant. Data are represented as mean  $\pm$  S.D. ( $n = 3$ ). S.D. – standard deviation. (D) An overlay of 14-3-3 $\zeta$  WT and R55A mutant docked poses and (E) is an overlay of D124A mutant and the 14-3-3 $\zeta$  WT docked poses (only the ligand and interacting residues are shown). D124 is not in view. Bound ATP is represented as a ball and stick model. Residues within 4 Å from ATP are shown and RMSD of superposition (of backbone atoms) is 0.15 Å.

bound protein was isolated from free ATP using a spin column and radioactivity was counted in a liquid scintillation counter. Approximately 6.7% of the total hot ATP added bound to WT protein, 2.66% in R55A and ~12% binding was seen in the D124A mutant (Fig. 5(B)). ~13% of hot ATP was bound to Hsp 70 under the same conditions. These results indicate that the D124 mutant has a better binding affinity for ATP while R55A mutant has less affinity as compared to the WT protein. ~2% binding could be seen with the sigma isoform as well (data not shown).

### 2.5. Computational analysis of binding energy and experimental constraints converge on one putative binding site

In the absence of high resolution structural information of bound ATP, it is difficult to explain why other amino acid substitutions in the vicinity of bound ATP in either binding pockets failed to show any effect. It is possible that ATP may bind in different orientations or binding may be accompanied by conformational changes that are not captured by docking algorithms. It is also intriguing

that mutation in one pocket resulted in enhancement of activity and in the other a partial loss. To substantiate the effect of these mutations, we engineered R55A mutation in the background of D124A mutant (DR mutant). Compared to the hyper active D124A mutant, DR mutant is only 60% active indicating that R55 near the interface is a key residue in binding/hydrolysis (Fig. 5(C)).

Based on these results we asked whether a single binding pocket can explain the effect of two mutations. We used ligand docking algorithm by VlifeMDS and imposed constraints that would mimic experimental observations in analyzing the docking results. The correct docked pose would be the one in which the D124A mutation results in a more negative binding energy and R55A mutation results in more positive binding energy for ATP. Mutants were engineered in silico using VlifeMDS software [34] and docking was carried out using VlifeMDS tools. Binding energy (BE) for ATP for all proteins was calculated for the same top 10 binding poses (Supplementary Table T2). Although ATP did bind to the same amphipathic pocket in the monomer as seen before (data not shown), no pose on the monomer confirmed to the applied constraints. In contrast a single pose at the dimer interface (pose 4 in Supplementary Table T2) showed (a) better binding of ATP in D124A mutant (BE  $-0.77$  kcal/mol vs  $4.04$  kcal/mol for the WT protein) and (b) weaker binding in R55A mutant (BE  $9.34$  kcal/mol). Detailed analysis of this pose indicates that ATP forms two hydrogen bonds with Arg55 in the B chain and also makes charged interactions with Arg55 in the A chain. In addition, Arg55 makes numerous van der Waal (vdW) interactions at the dimer interface (Fig. 5(D)). This docking pose clearly explains the importance of Arg55 in binding interactions (Fig. 5(D)). Mutation of this residue to Ala would result in loss of hydrogen bond interactions as well as the charge interaction between the positively charged guanidine group of arginine and negatively charged phosphate group of ATP, adversely affecting binding energy and therefore catalysis. Increased rate of hydrolysis of D124A is reflected in better binding energy of ATP in this mutant. This could be due to favorable steric interactions between ATP and the mutated protein (Fig. 5(E)) mediated by an increase in the vdW component of the overall binding energy.

## 2.6. Single turnover studies

To correlate ATP binding and ATP hydrolysis, we performed single turnover ATPase assay. WT 14-3-3 $\zeta$ , D124A and R55A mutant proteins and 14-3-3 $\sigma$  were independently incubated with radioactive ATP and the bound complex free from ATP was isolated. Protein fraction which has the maximum signal for ATP was chosen and hydrolysis was followed with time (Fig. 6). The data could be

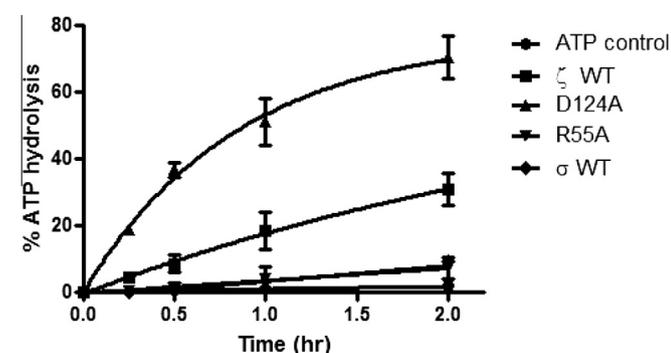


Fig. 6. Single turnover ATPase assay WT 14-3-3 $\zeta$ , R55A, D124A or WT 14-3-3 $\sigma$  bound-ATP complex was isolated and single turnover ATPase assay was performed. Data represents percentage of ATP hydrolysis with time from two independent experiments. Data are represented as mean  $\pm$  S.D. – standard deviation.

readily fit to a single exponential function and results indicate that 30% of bound ATP is hydrolyzed in 2 h by 14-3-3 $\zeta$  while D124A hydrolyzes almost 70% of bound ATP by this time (rate constant  $K = 1.17 \pm 0.21$ ). R55A and 14-3-3 $\sigma$  did not show any detectable hydrolysis of ATP under these conditions. Conversion of almost 70% of the bound ATP to product by D124A in a single exponential manner indicates that binding is likely to be stoichiometric ruling out the presence of any non-specifically bound ATP that cannot be hydrolyzed. Commensurate with other results 14-3-3 $\zeta$  WT is slow in hydrolyzing the bound ATP which is very likely due to the faster dissociation of the protein-ATP complex into the free forms.

These results in toto seem to define the pocket near the interface as the most likely binding site for ATP. Nevertheless it is possible that there are indeed two binding pockets for ATP. It is also possible that these mutations have a long range effect on a binding site located elsewhere in the protein or they affect equilibrium distribution between active and inactive conformers.

## 3. Conclusions

Our study assigns for the first time an ATP hydrolyzing activity to human 14-3-3  $\zeta$ ,  $\epsilon$  and  $\tau$  isoforms. Our mutagenesis results based on prediction tools show that this activity in zeta can be altered resulting in gain or loss of function. While the ultra-structural details of the binding pocket, catalytic residues and more importantly the physiological and functional relevance of this ATP hydrolyzing activity remains to be established, our results are likely to provide the necessary impetus for an in depth investigation of this enzymatic activity and its physiological significance. The ATPase activity in 14-3-3 $\zeta$  seems to stem from an unconventional binding pocket and could be a moon lighting function. The structure may represent a new fold among atypical ATPases. It remains to be seen whether the enzymatic activity of 14-3-3 proteins is preserved in cellular milieu and whether some of their known functions and perhaps many unknown functions are dependent on this activity. The effect of in vitro mutations may also be mimicked by regulatory mechanisms inside the cell. Notably the sigma isoform which differs from all other isoforms both in sequence, structure and other functions, failed to show any detectable ATPase activity. Whether the same sequence or structure plays a role in other active isoforms remains to be established.

## 4. Material and methods

Plasmids used in the study, protein purification and characterization by Western blot and nano LC-MS<sup>E</sup> analysis [35], ATP binding and single turnover ATPase assay are detailed in Supplementary methods.

### 4.1. ATPase assay

(A) *Calorimetry*: 7  $\mu$ M (calculated for dimeric protein) of protein was incubated with 1 mM of ATP or ATP- $\gamma$ -S or AMP-PCP (Sigma) at 37  $^{\circ}$ C in 50  $\mu$ l reaction buffer (RB) (RB: 20 mM HEPES buffer pH 7.5, 5 mM MgCl<sub>2</sub> and 2 mM DTT). Release of inorganic phosphate (Pi) was monitored at different time points using Malachite green assay with minor modifications. 450  $\mu$ l of assay mixture (3 part of 0.4% malachite green, 1 part 4.2% ammonium molybdate made in 5 N HCl, 0.056% polyvinyl alcohol) was added to 50  $\mu$ l reaction volume. Mixture was allowed to incubate for 10 min at room temperature and 200  $\mu$ l of the sample was read at 630 nm (Spectra Max 190, Molecular Devices) with appropriate blank. Amount of phosphate released was estimated with the help of standards generated using KH<sub>2</sub>PO<sub>4</sub>.

(B) *Luminescence assay*: ADP-Glo™ max ATPase assay kit (Promega) was used as per manufacturer's instructions. Luminescence was recorded using multimode microplate reader (Mitras LB 940, Berthold Technologies).

(C) *Radioactive assay*: Trace amounts of ( $\gamma$ - $^{32}$ P) ATP [10  $\mu$ Ci (3000 Ci/mmol), PerkinElmer] in RB (25  $\mu$ l) containing 100  $\mu$ M of cold ATP (Sigma) was incubated as before. Reaction was stopped by spotting 2  $\mu$ l of sample on *poly(ethyleneimine) cellulose* thin layer chromatographic (PEI-TLC) plates (Fluka) and developed with 0.75 M  $\text{KH}_2\text{PO}_4$  (pH 3.5). Plates were dried and exposed to X-ray film (Kodak) to monitor inorganic phosphate (Pi) release. The experiment was repeated at least two times with two independent protein preparations. All radioactive experiments were performed according to standard institutional guidelines.

#### 4.2. Docking of ATP using Schrödinger software and molecular dynamic simulation of the bound complex

ATP was docked to PDB ID: 2C1J which has no structure breakage in between (1–230) [36] using Schrödinger software (for details please see [Supplemental methods](#)).

#### 4.3. Computational analysis of binding energy and experimental constraints converge on one putative binding site

ATP was blind docked separately into monomer and dimer structures of 14-3-3 $\zeta$  (PDB ID: 2C1J) using VLifeMDS software. Top 10 poses based on PLP score were selected ([Supplementary Table T2](#)) [37]. Binding energy (kcal/mole) for each pose was calculated using the formula:  $E_{\text{bind}} = E_{\text{complex}} - (E_{\text{protein}} + E_{\text{ligand}})$ . R55A and D124A mutations were created, energy optimized and binding energies were calculated for the same ten docked poses of ATP as seen with the WT protein ([Supplementary Table T2](#)).

#### Author's contributions

M.R. purified all proteins and performed all experiments in this study. P.S. and V.R. were involved in initial studies. A.K. used Schrödinger for docking. A.K.S.G. assisted in radioactive experiments. P.N. purified Hsp 70 protein. S.K. and S.M. performed docking, in silico and binding energy calculations using VLifeMDS software. P.V. designed and directed the project, wrote the manuscript assisted by M.R., S.K. and S.M.

#### Competing interests

The authors declare that they have no competing interests.

#### Acknowledgments

We thank Dr. Sorab Dalal (ACTREC) for 14-3-3  $\zeta$ ,  $\gamma$ ,  $\epsilon$ , and  $\sigma$  and Dr. Surekha Zingde (ACTREC) for pGEX4T1-Hsp 70 constructs; Madhuja, Arvind and Dr. Patrick D'Silva (IISc, Bangalore) for their assistance in single turnover ATPase assay; Dr. Hemangi (Dr. Kalraiy Lab, ACTREC) and Dr. Mahesh Kulkarni (NCL, Pune) for nano LC-MS<sup>E</sup> experiments. ACTREC and Council of Scientific and Industrial Research, India for JRF and SRF fellowship to MR; Department of Biotechnology, Government of India, for Grant BT/PR11372/BRB/10/655/2008 and Department of Science and Technology, Government of India, for Grant SR/SO/BB-0031/2011.

#### Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.febslet.2013.11.008>.

#### References

- [1] Yaffe, M.B., Rittinger, K., Volinia, S., Caron, P.R., Aitken, A., Leffers, H., Gambin, S.J., Smerdon, S.J. and Cantley, L.C. (1997) The structural basis for 14-3-3: phosphopeptide binding specificity. *Cell* 91 (7), 961–971.
- [2] Zhao, J., Meyerkord, C.L., Du, Y., Khuri, F.R. and Fu, H. (2011) 14-3-3 proteins as potential therapeutic targets. *Semin. Cell Dev. Biol.* 22 (7), 705–712.
- [3] Gardino, A.K. and Yaffe, M.B. (2011) 14-3-3 proteins as signaling integration points for cell cycle control and apoptosis. *Semin. Cell Dev. Biol.* 22, 688–695.
- [4] Ganguly, S., Weller, J.L., Ho, A., Chemineau, P., Malpoux, B. and Klein, D.C. (2005) Melatonin synthesis: 14-3-3-dependent activation and inhibition of arylalkylamine N-acetyltransferase mediated by phosphoserine-205. *Proc. Nat. Acad. Sci. U.S.A.* 102 (4), 1222–1227.
- [5] Dalal, S.N., Yaffe, M.B. and DeCaprio, J.A. (2004) 14-3-3 family members act coordinately to regulate mitotic progression. *Cell Cycle* 3 (5), 672–677.
- [6] Hermeking, H. (2003) The 14-3-3 cancer connection. *Nat. Rev. Cancer* 3 (12), 931–943.
- [7] Hosing, A.S., Kundu, S.T. and Dalal, S.N. (2008) 14-3-3 gamma is required to enforce both the incomplete S phase and G2 DNA damage checkpoints. *Cell Cycle* 7 (20), 3171–3179.
- [8] Telles, E., Hosing, A.S., Kundu, S.T., Venkatraman, P. and Dalal, S.N. (2009) A novel pocket in 14-3-3 $\epsilon$  is required to mediate specific complex formation with cdc25C and to inhibit cell cycle progression upon activation of checkpoint pathways. *Exp. Cell Res.* 315 (8), 1448–1457.
- [9] Aitken, A. (2011) Post-translational modification of 14-3-3 isoforms and regulation of cellular function. *Semin. Cell Dev. Biol.* 22, 673–680.
- [10] Muslin, A.J., Tanner, J.W., Allen, P.M. and Shaw, A.S. (1996) Interaction of 14-3-3 with signaling proteins is mediated by the recognition of phosphoserine. *Cell* 84 (6), 889–897.
- [11] Obsil, T., Ghirlando, R., Klein, D.C., Ganguly, S. and Dyda, F. (2001) Crystal structure of the 14-3-3zeta: serotonin N-acetyltransferase complex. A role for scaffolding in enzyme regulation. *Cell* 105 (2), 257–267.
- [12] Gardino, A.K., Smerdon, S.J. and Yaffe, M.B. (2006) Structural determinants of 14-3-3 binding specificities and regulation of subcellular localization of 14-3-3-ligand complexes: a comparison of the X-ray crystal structures of all human 14-3-3 isoforms. *Semin. Cancer Biol.* 16, 173–182.
- [13] Forrest, A. and Gabrielli, B. (2001) Cdc25B activity is regulated by 14-3-3. *Oncogene* 20 (32), 4393–4401.
- [14] Yano, M., Mori, S., Niwa, Y., Inoue, M. and Kido, H. (1997) Intrinsic nucleoside diphosphate kinase-like activity as a novel function of 14-3-3 proteins. *FEBS Lett.* 419 (2–3), 244–248.
- [15] Komiya, T., Hachiya, N., Sakaguchi, M., Omura, T. and Mihara, K. (1994) Recognition of mitochondria-targeting signals by a cytosolic import stimulation factor, MSF. *J. Biol. Chem.* 269 (49), 30893–30897.
- [16] Hachiya, N., Komiya, T., Alam, R., Iwahashi, J., Sakaguchi, M., Omura, T. and Mihara, K. (1994) MSF, a novel cytoplasmic chaperone which functions in precursor targeting to mitochondria. *EMBO J.* 13 (21), 5146–5154.
- [17] Komiya, T. and Mihara, K. (1996) Protein import into mammalian mitochondria. Characterization of the intermediates along the import pathway of the precursor into the matrix. *J. Biol. Chem.* 271 (36), 22105–22110.
- [18] Komiya, T., Sakaguchi, M. and Mihara, K. (1996) Cytoplasmic chaperones determine the targeting pathway of precursor proteins to mitochondria. *EMBO J.* 15 (2), 399–407.
- [19] Yano, M., Nakamura, S., Wu, X., Okumura, Y. and Kido, H. (2006) A novel function of 14-3-3 protein: 14-3-3zeta is a heat-shock-related molecular chaperone that dissolves thermal-aggregated proteins. *Mol. Cell* 17 (11), 4769–4779.
- [20] Omi, K., Hachiya, N.S., Tanaka, M., Tokunaga, K. and Kaneko, K. (2008) 14-3-3zeta is indispensable for aggregate formation of polyglutamine-expanded huntingtin protein. *Neurosci. Lett.* 431 (1), 45–50.
- [21] Sluchanko, N.N. and Gusev, N.B. (2011) Probable participation of 14-3-3 in tau protein oligomerization and aggregation. *J. Alzheimers Dis.* 27 (3), 467–476.
- [22] Berg, D., Holzmann, C. and Riess, O. (2003) 14-3-3 proteins in the nervous system. *Nat. Rev. Neurosci.* 4 (9), 752–762.
- [23] Aitken, A. (2006) 14-3-3 proteins: a historic overview. *Semin. Cancer Biol.* 16 (3), 162–172.
- [24] Ames, B.N. (1966) Assay of inorganic phosphate, total phosphate and phosphatase. *Methods Enzymol.* 8, 115–118.
- [25] Chauhan, J.S., Mishra, N.K. and Raghava, G.P.S. (2009) Identification of ATP binding residues of a protein from its primary sequence. *BMC Bioinformatics* 10 (1), 434.
- [26] Liu, D., Bienkowska, J., Petosa, C., Collier, R.J., Fu, H. and Liddington, R. (1995) Crystal structure of the zeta isoform of the 14-3-3 protein. *Nature* 376, 191–194.
- [27] Glide v. Schrödinger, LLC, New York, NY, 2011, Version 5.7; <http://www.schrodinger.com>.
- [28] Jaiswal, H., Conz, C., Otto, H., Wölfe, T., Fitzke, E., Mayer, M.P. and Rospert, S. (2011) The chaperone network connected to human ribosome-associated complex. *Mol. Cell Biol.* 31 (6), 1160–1173.
- [29] Bimston, D., Song, J., Winchester, D., Takayama, S., Reed, J.C. and Morimoto, R.I. (1998) BAG-1, a negative regulator of Hsp70 chaperone activity, uncouples nucleotide hydrolysis from substrate release. *EMBO J.* 17 (23), 6871–6878.
- [30] Olson, C.L., Nadeau, K.C., Sullivan, M.A., Winquist, A.G., Donelson, J.E., Walsh, C.T. and Engman, D.M. (1994) Molecular and biochemical comparison of the 70-kDa heat shock proteins of *Trypanosoma cruzi*. *J. Biol. Chem.* 269 (5), 3868–3874.

- [31] Owen, B.A., Sullivan, W.P., Felts, S.J. and Toft, D.O. (2002) Regulation of heat shock protein 90 ATPase activity by sequences in the carboxyl terminus. *J. Biol. Chem.* 277 (9), 7086–7091.
- [32] McLaughlin, S.H., Smith, H.W. and Jackson, S.E. (2002) Stimulation of the weak ATPase activity of human hsp90 by a client protein. *J. Mol. Biol.* 315 (4), 787–798.
- [33] Bösl, B., Grimminger, V. and Walter, S. (2006) The molecular chaperone Hsp104—a molecular machine for protein disaggregation. *J. Struct. Biol.* 156 (1), 139–148.
- [34] VLifeMDS V, VLife Sciences Technologies Pvt. Ltd., Pune, India 2012, Version 4.2; <http://www.vlifesciences.com>.
- [35] Silva, J.C., Gorenstein, M.V., Li, G.-Z., Vissers, J.P. and Geromanos, S.J. (2006) Absolute quantification of proteins by LCMSE a virtue of parallel MS acquisition. *Mol. Cell. Proteomics* 5 (1), 144–156.
- [36] Macdonald, N., Welburn, J.P.I., Noble, M.E.M., Nguyen, A., Yaffe, M.B., Clynes, D., Moggs, J.G., Orphanides, G., Thomson, S. and Edmunds, J.W. (2005) Molecular basis for the recognition of phosphorylated and phosphoacetylated histone h3 by 14-3-3. *Mol. Cell* 20 (2), 199–211.
- [37] Gehlhaar, D.K., Verkhivker, G.M., Rejto, P.A., Sherman, C.J., Fogel, D.R., Fogel, L.J. and Freer, S.T. (1995) Molecular recognition of the inhibitor AG-1343 by HIV-1 protease: conformationally flexible docking by evolutionary programming. *Chem. Biol.* 2 (5), 317–324.

# Discovery of multiple interacting partners of gankyrin, a proteasomal chaperone and an oncoprotein—Evidence for a common hot spot site at the interface and its functional relevance

Padma P. Nanaware, Manoj P. Ramteke, Arun K. Somavarapu, and Prasanna Venkatraman\*

Advanced Centre for Treatment, Research and Education in Cancer, Navi Mumbai, India

## ABSTRACT

Gankyrin, a non-ATPase component of the proteasome and a chaperone of proteasome assembly, is also an oncoprotein. Gankyrin regulates a variety of oncogenic signaling pathways in cancer cells and accelerates degradation of tumor suppressor proteins p53 and Rb. Therefore gankyrin may be a unique hub integrating signaling networks with the degradation pathway. To identify new interactions that may be crucial in consolidating its role as an oncogenic hub, crystal structure of gankyrin-proteasome ATPase complex was used to predict novel interacting partners. EEVD, a four amino acid linear sequence seems a hot spot site at this interface. By searching for EEVD in exposed regions of human proteins in PDB database, we predicted 34 novel interactions. Eight proteins were tested and seven of them were found to interact with gankyrin. Affinity of four interactions is high enough for endogenous detection. Others require gankyrin overexpression in HEK 293 cells or occur endogenously in breast cancer cell line- MDA-MB-435, reflecting lower affinity or presence of a deregulated network. Mutagenesis and peptide inhibition confirm that EEVD is the common hot spot site at these interfaces and therefore a potential polypharmacological drug target. In MDA-MB-231 cells in which the endogenous CLIC1 is silenced, trans-expression of Wt protein (CLIC1\_EEVD) and not the hot spot site mutant (CLIC1\_AAVA) resulted in significant rescue of the migratory potential. Our approach can be extended to identify novel functionally relevant protein-protein interactions, in expansion of oncogenic networks and in identifying potential therapeutic targets.

Proteins 2014; 00:000–000.  
© 2013 Wiley Periodicals, Inc.

**Key words:** short linear sequence motif; binding; hub protein; mutation; function; inhibition.

## INTRODUCTION

Protein-protein interactions create complex functional networks in cells and these interaction networks coincide with the signaling networks that regulate cell behavior.<sup>1</sup> Detecting and describing these interaction networks is the goal of many large proteomics and bioinformatics studies.<sup>2–5</sup> Cancer associated proteins especially oncoproteins are identified as hubs in such networks<sup>1</sup> and key interactions within the network represent prime targets for therapeutic interventions.

Gankyrin, an oncoprotein overexpressed in many epithelial cancers<sup>6–12</sup> was initially identified as a non-ATPase subunit of the proteasome. While gankyrin is not considered as a stable component of the 26S proteasome, it is detected in its free form as well as in complex with 19S regulatory particles of the proteasome.<sup>13</sup> Nas6, the

homologue of gankyrin acts as chaperone required for the assembly of the 19S regulatory particles.<sup>14,15</sup>

Some of the known functions of gankyrin as an oncoprotein involve its ability to deregulate key signaling networks and/or influence degradation of crucial regulatory molecules by the proteasome. By directly binding to MDM2, gankyrin facilitates degradation of p53.<sup>16</sup> By

Additional Supporting Information may be found in the online version of this article.

Grant sponsor: SIA; grant number: 2691; Grant sponsor: TMC-ACTREC. Arun K. Somavarapu's current address is Protein Science and Engineering Department, Institute of Microbial Technology, Sec 39-A, Chandigarh 160036, India.

\*Correspondence to: Prasanna Venkatraman, KS 244 Advanced Centre for Treatment Research and Education in Cancer, Tata Memorial Centre, Sector 22, Kharghar, Navi Mumbai, Maharashtra 410210, India.

E-mail: vprasanna@actrec.gov.in

Received 5 June 2013; Revised 20 November 2013; Accepted 9 December 2013  
Published online 13 December 2013 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/prot.24494

directly binding to Rb, gankyrin increases Rb phosphorylation and degradation by the proteasome<sup>12</sup> resulting in the release of E2F a transcription factor, responsible for cellular proliferation. Gankyrin is also known to interact with CDK4 kinase which may be important for cell cycle regulation.<sup>13,17</sup> Interaction of gankyrin with MAGE-A4 seems to suppress tumor formation in athymic mice overexpressing gankyrin.<sup>18</sup> Gankyrin also binds to ankyrin repeats in NF- $\kappa$ B and inhibits its activity.<sup>19</sup> Gankyrin is also known to bind to NF- $\kappa$ B and suppresses the transcriptional activity by modulating acetylation of SIRT1.<sup>20</sup> Ras mediated oncogenic signaling is dependent on the presence of gankyrin.<sup>7</sup> NIH3T3 cells overexpressing gankyrin when injected into nude mice form tumors.<sup>12</sup> When gankyrin expression is silenced, cells undergo reduced proliferation and reduced colony formation on soft agar assay.<sup>7</sup> When gankyrin silenced pancreatic cancer cells were injected into nude mice, the tumors formed were of reduced size. In contrast, when gankyrin was overexpressed, these cells formed large size tumors.<sup>8</sup>

These findings clearly suggest that gankyrin connects multiple oncogenic pathways and therefore shows characteristics of a key hub protein. Being part of the ubiquitin-proteasome pathway gankyrin is uniquely positioned to link at least some of these pathways to the ubiquitin proteasome network which may play a decisive role in disease progression. As mentioned before, gankyrin is also overexpressed in many cancers. Hub proteins that are overexpressed, are estimated to be at least three times more essential than the non-hub counterparts.<sup>21</sup> They are also potential anti-cancer drug targets.

Challenge in understanding properties of such hub proteins and their utility as therapeutic targets lies in identifying key interactions among the many possible ones. With limited surface available for interactions, the question of how multiple interactions are achieved by a single protein is a matter of considerable debate.<sup>1,22</sup> Although protein-protein interactions may involve large surface areas, bulk of the binding energy is contributed by few key residues at the interaction surface. These when mutated result in either rapid dissociation of the complex or prevent stable association and are called as the 'hot spot' sites.<sup>23,24</sup> Such hot spot sites are conserved across interfaces.<sup>24</sup> A key regulatory molecule therefore may interact with multiple partners through such a common recognition motif.<sup>22</sup>

Our aim to establish protein interaction network of gankyrin are twofold (a) to better understand its role in oncogenesis and (b) in the future to be able to identify vulnerable nodes in the network that are amenable for therapeutic intervention. Since crystal structure of gankyrin in complex with an S6ATPase of the proteasome is known, we began by predicting a probable hot spot site at the S6ATPase interface formed by EEVD a linear short sequence of four amino acid residues. Since hot spot

sites are conserved, it is likely that gankyrin may recognize other proteins which also carry EEVD in an accessible region of the protein. Using a simple bioinformatics tool we identify 34 proteins with EEVD in well exposed region on their surface and experimentally prove interactions with seven out of the eight proteins tested. Three of these interactions occur in HEK 293 cells only when gankyrin is overexpressed but occur in breast cancer cells at endogenous levels. Mutagenesis confirms that these interactions involve predicted residues which form the hot spot sites at the shared interface and could be a potential drug target.

Our strategy may be extended to the identification of novel protein-protein interactions which share linear sequence of amino acids at the interaction sites, to expand functionally relevant interaction networks of key regulatory proteins and perhaps in the identification of therapeutically vulnerable targets.

## MATERIALS AND METHODS

### Bioinformatic detection of putative interacting partners of gankyrin

Protein sequences from human proteome data was downloaded from the Uniprot website (URL:ftp://ftp.ebi.ac.uk/pub/databases/uniprot; March 21, 2012). This data set contains a total of 81,194 sequences in FASTA format and has two parts, one containing manually annotated data (UniProtKB/Swiss-Prot) and the other a computationally analyzed data awaiting manual annotation (UniProtKB/TrEMBL). Manually annotated 35,961 entries were used for further analysis. Proteins containing EEVD in their primary sequence were extracted using perl scripts. From this list, proteins with three-dimensional structures were short listed and their Uniprot IDs were submitted to the Protein Data Bank (PDB) and structures were downloaded. Solvent Accessible Surface Area (SASA) of the EEVD was calculated in the context of the tetra peptide using Parameter optimized surfaces (POPS) stand alone application.<sup>25</sup>

### Clones and constructs

Gankyrin pBluescript II SK (+) construct (kind gift from Dr. Jun Fujita, Kyoto University) was subcloned into mammalian expression vector p3XFLAG-CMV<sup>TM</sup>-10 (Sigma) or in prokaryotic expression vector pRSETA-TEV. Hsp70 in pGEX4T1 vector was received as a kind gift from Dr. Surekha Zingde, ACTREC. CLIC1, NCK2, GRSF1, S6ATPase cDNA were generated by RT-PCR of RNA extracted from HEK 293 cells. CLIC1 was cloned in pGEX4T1 (GE Amersham) and pCDNA 3.1 with HA tag at the N terminus (pCDNA3.1 was received as a kind gift from Dr. Sorab Dalal, ACTREC), while NCK2 and GRSF1 were cloned in pCDNA3.1. Hsp70 was also cloned in

p3XFLAG-CMV<sup>TM</sup>-10. Mutations in Hsp70 (aa 638EEVD641 to AAVA), CLIC1 (aa 150EEVD153 to AAVA), CLIC1 (aa 152V to aa 152E), NCK2 (aa 167EEVD170 to AAVA), GRSF1 (aa 144EEVD147 to AAVA), S6ATPase (aa 356EEVD359 to AAVA) were generated using PCR based site directed mutagenesis with the help of Phusion high fidelity DNA polymerase (Finnzymes-Thermo Scientific) and were further confirmed by sequencing. Primers used for cloning and site directed mutagenesis are reported in Supporting Information Table S1.

### Cell culture, transfection, and generation of stable clones

HEK 293, MDA-MB-231, MDA-MB-435 cells were cultured in DMEM (GIBCO) supplemented with 10% FBS (GIBCO). Cells were maintained in 5% CO<sub>2</sub> incubator at 37°C. HEK 293 cells were transfected with pCMV10-3X-p3XFLAG-CMV<sup>TM</sup>-10-gankyrin or vector alone using calcium phosphate method. Clonal transformants were selected in presence of 800 µg/mL G418 (Sigma). Lipofectamine 2000 was used to transfect siRNA or gene of interest in MDA-MB-231 and MDA-MB-435 cells.

### Characterization of stable clones

#### Soft agar assay

Gankyrin overexpressing HEK 293 stable clones or vector alone clones obtained by the above method were compared for their ability to exhibit anchorage independent growth using soft agar colony formation assay.  $2 \times 10^3$  cells in each case were overlaid on a thin layer of agar (0.4% over 1% agarose). Colonies formed were counted after 7 days.

#### MTT assay for proliferation

HEK 293 cells were transiently transfected with gankyrin or vector alone. Forty-eight hours after transfection,  $5 \times 10^3$  cells were seeded in a 96-well plate and 20 µL of (3-[4,5-dimethylthiazol-2-yl]-2,5-diphenyltetrazolium bromide (MTT) (5 mg/mL) (Sigma) was added. After 2 h, 100 µL of 10% SDS in 0.01 N HCl was added and absorbance was read at 550 nm.

#### Etoposide treatment for apoptosis

Vector alone and gankyrin expressing stable HEK 293 clones were subjected to etoposide (Cipla) treatment;  $2 \times 10^3$  cells were seeded in 96-well and treated with 50 µM etoposide. Cell viability was checked after 72 h using MTT assay as described earlier.

#### Luciferase assay for NF-κB activity

Stable clones of HEK 293 harboring p3XFLAG-CMV<sup>TM</sup>-10 or p3XFLAG-CMV<sup>TM</sup>-10-gankyrin were transfected with 12 µg of ConA control or 3 kb enhancer

ConA luciferase construct (a kind gift from Dr. Neil D. Perkins, UK). After 48 h, cells were treated with 10 ng/mL TNF-α (Invitrogen), lysed and luciferase assay was performed using Promega Luciferase Assay System in triplicates.

### Affinity pull down and western blotting

Cells were harvested in NP-40 lysis buffer (20 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.5% NP-40 and 1 mM dithiothreitol (DTT) containing 1× protease inhibitor cocktail (Sigma)). One milligram of total cell lysate was incubated with 10 µL of M2 agarose anti-flag beads (Sigma) for 4 h at 4°C. Beads were washed extensively with the NP-40 lysis buffer containing 300 mM NaCl. In order to increase the stringency in pull down experiments (Figure 3(C-F) and Supporting Information Fig.S2(C)), RIPA buffer (25 mM Tris-HCl pH 7.5, 150 mM NaCl, 1% NP-40, 1% sodium deoxycholate, 0.1% SDS) was used during washes. Samples were boiled in the presence of Laemmli buffer. Proteins were resolved on a 12% SDS PAGE and western blotting was performed using standard protocol. Antibodies used in the analysis were against gankyrin (Sigma), Hsp70 (Abcam), CLIC1 (Abcam), DDAH1 (Sigma), GRSF1 (Sigma), EIF4A3 (Sigma), Hsp90 (Santacruz), MAP2K1 (Sigma), HA (Abcam). Immunoblots were visualized by enhanced chemiluminescence (ECL plus; Amersham). HEK 293 cells and MDA-MB-435 cells overexpressing HA-NCK2 or NCK2\_AAVA were treated with 5 µM MG132 for 6 h before cell harvesting.

### Immunoprecipitation and western blotting

For immunoprecipitation experiments, cells were harvested in NP-40 lysis buffer. 15 µL of Sepharose G beads were incubated overnight with 3 µg of respective antibody in same buffer. One milligram of precleared lysate was added to the antibody bound beads and incubated for 4 h. These immune complexes were washed thoroughly for four to five times using NP-40 lysis buffer containing 300 mM NaCl. In some cases RIPA buffer (25 mM Tris-HCl pH 7.5, 150 mM NaCl, 1% NP-40, 1% sodium deoxycholate, 0.1% SDS) was used during washes to increase the stringency in immunoprecipitation experiments [Fig. 2(D,E)]. Samples were boiled in the presence of Laemmli buffer. Proteins were resolved on 12% SDS PAGE and western blotting was performed.

### Expression and purification of proteins

Recombinant His-gankyrin, GST-fusions of Hsp70, CLIC1 and their corresponding AAVA mutants and EEED mutants were expressed in *Escherichia coli* BL21 DE(3) using 100 µM IPTG at 24°C for 16 h. His-gankyrin was purified using nickel-nitrilotriacetic acid (Ni-NTA) chromatography (Qiagen) as a His-tag protein.

GST and other fusion proteins listed above were purified using glutathione Sepharose (GE Amersham) following manufacturer's protocol. All proteins were dialyzed against 50 mM Tris-HCl pH 7.5 and stored at  $-20^{\circ}\text{C}$ .

#### **In vitro GST pull down assay**

GST-fusion proteins and their mutants in NP-40 lysis buffer were allowed to immobilize on GST beads for 1 h at  $4^{\circ}\text{C}$ . Beads were washed two to three times with the same buffer but containing 300 mM NaCl. His-gankyrin was allowed to bind for 2 h at  $4^{\circ}\text{C}$ . Beads were washed thoroughly (five to six times) and boiled in the presence of Laemmli buffer. Proteins were resolved on a 12% SDS PAGE and western blotting was performed using anti-gankyrin antibody.

#### **Peptide inhibition**

His-gankyrin (30  $\mu\text{g}$ ) was incubated with different concentrations (1–500  $\mu\text{M}$ ) of EEVD peptide. Peptides GRRF and GRRR were included as controls. Preincubated complex were added to 30  $\mu\text{g}$  of purified GST-CLIC1 bound to glutathione beads and incubated for 2 h at  $4^{\circ}\text{C}$ . Beads were washed thoroughly for five to six times with NP-40 lysis buffer and boiled in the presence of Laemmli buffer. Proteins were resolved on a 12% SDS PAGE and western blotting was performed using gankyrin antibody. Band intensities were measured using ImageJ software. To obtain an approximate  $K_d$  value of the peptide, we made the following assumptions. We assume that band intensity at zero peptide represents total CLIC1 bound gankyrin. Band intensity in presence of varying concentrations of the peptide is then subtracted from this value. We consider this subtracted value to correspond to peptide bound fraction of gankyrin. These values are then plotted against peptide concentration to obtain  $\sim K_d$  of the peptide.

#### **Wound healing assay**

Directional cell migration was studied using an *in vitro* wound healing assay.<sup>26</sup> Cells were seeded in six-well plates and allowed to reach 70 to 90% confluence. To monitor the independent role of gankyrin or CLIC1, MDA-MB-231 cells were transfected with 100 nM gankyrin siRNA or CLIC1 siRNA or control siRNA and incubated for 72 h before setting up the wound healing assay. To examine the effect of protein-protein interaction (PPI) on the motility of the cells, HEK 293 and MDA-MB-231 cells were transfected with CLIC1 Wt or CLIC1\_AAVA mutant or vector control and incubated for 48 h. These experiments were conducted in the presence of endogenous CLIC1. Parental HEK 293 cells however do not show detectable endogenous CLIC1. To further confirm, the importance of PPI in this functional assay, endogenous CLIC1 in MDA-MB-231 cells were

silenced using smartpool of UTR specific siRNA (Dharmacon) and co-transfected with either CLIC1 Wt or CLIC1\_AAVA mutant. Control experiments included non-target siRNA with or without the cDNA for CLIC1 Wt or AAVA mutant as the case may be. For transfection in MDA-MB-231 cells, 6  $\mu\text{g}$  DNA and 18  $\mu\text{L}$  of lipofectamine 2000 was used. These transfected cells were grown to 100% confluence and were treated with 10  $\mu\text{g}/\text{mL}$  of mitomycin C for 3 h. Cell monolayer was then wounded with a plastic tip, washed with PBS to remove cell debris. Wound healing capacity was monitored for 16 h or 24 h using phase contrast microscopy (Axiovert 200M, Zeiss, Germany). Area under each wound was calculated using Image J software. Percentage of wound healed was calculated using the formula [(initial wound area – final wound area)/initial wound area]  $\times 100$ . The siRNA sequences used are tabulated in Supporting Information Table S1.

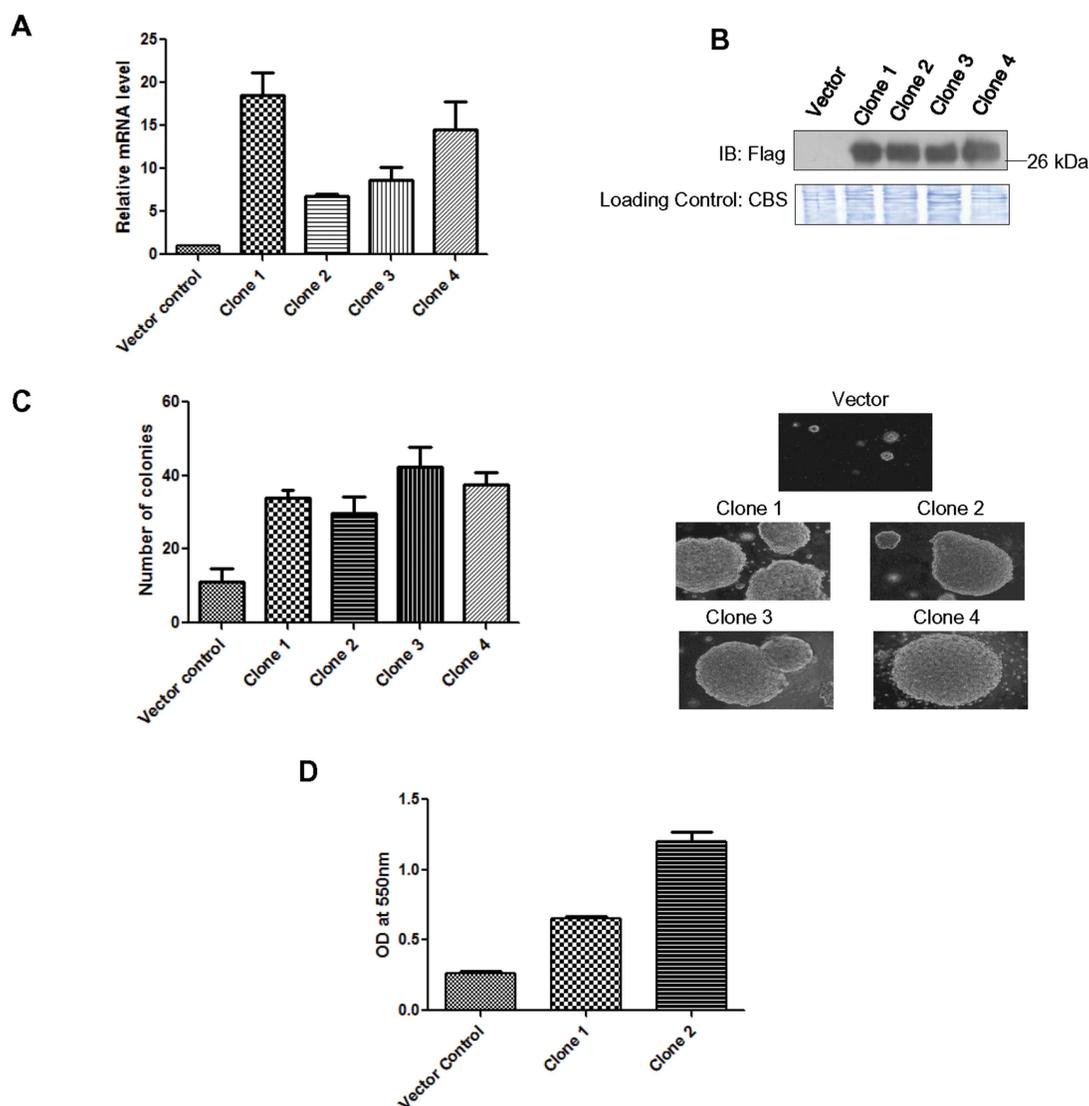
#### **In vitro invasion assay**

Invasion assay was performed using 24-well transwell culture inserts of 8  $\mu\text{m}$  pore size (BD Biosciences). Matrigel at a concentration 300  $\mu\text{g}/\text{mL}$  diluted in  $1\times$  DMEM was added to the inner side of the chamber and allowed to polymerize at  $37^{\circ}\text{C}$  for 1 h. Unpolymerized matrigel was removed and  $5 \times 10^4$  cells were seeded in 200  $\mu\text{L}$  of DMEM. To the lower chamber complete medium (DMEM + 10% FBS) was added as chemoattractant. Cells were then incubated for 24 h at  $37^{\circ}\text{C}$  in a humidified  $\text{CO}_2$  (5%) incubator. After incubation, inserts were removed and the noninvading cells on the upper surface of the insert were scraped. Lower surface of the insert which harbors the invaded cells was fixed using cold methanol and stained using crystal violet. All images were taken at  $5\times$  magnification using an upright microscope (Axio Imager. Z1, Zeiss, Germany).

## **RESULTS**

### **Overexpression of gankyrin in HEK 293 cells and generation of stable clones**

Gankyrin overexpression in NIH3T3 (mouse) cells results in malignant transformation and transfected cells can form colonies on soft agar and seed tumors in nude mice.<sup>12</sup> We replicated this model in HEK 293 cells with the major goal of identifying protein-protein interactions that may be functionally and therapeutically relevant. In addition, by using a model system where gankyrin is transexpressed allows us to establish a causal relationship to gankyrin and attribute the changes in the network to overexpression of gankyrin. Gankyrin was overexpressed as a flag tag fusion protein and stable colonies were selected. Expression levels of gankyrin in the stable clones were tested using qRT-PCR. Total gankyrin (using

**Figure 1**

Generation of gankyrin overexpressing stable HEK 293 clones. (A) HEK 293 stable clones overexpressing gankyrin were generated. Total levels of gankyrin were quantitated using gankyrin specific primers and real time PCR. HPRT was used as the internal control. Data represent mean  $\pm$  SD of one experiment done in duplicate; (B) Western blot using anti-flag antibody shows overexpression of transfected gankyrin in stable clones. Coomassie brilliant blue stained bots (CBS) indicates equal loading; (C) HEK 293 stable clones (four clones) show anchorage independent growth as compared to the vector alone cells;  $2 \times 10^3$  cells were seeded and number of colonies formed after 7 days are plotted. The data represents mean  $\pm$  SD of one experiment done in triplicate. Three independent experiments were performed in duplicate to confirm these observations. A representative  $10\times$  image of one of the fields from each plate is also shown; (D) HEK 293 vector alone and gankyrin over expressing cells were treated with  $50 \mu\text{M}$  etoposide and cell viability was checked after 72 h. Clones overexpressing gankyrin were resistant to apoptosis as compared to the vector alone cells. Data represent mean  $\pm$  SD of one experiment done in duplicate. [Color figure can be viewed in the online issue, which is available at [wileyonlinelibrary.com](http://wileyonlinelibrary.com).]

gankyrin specific primers) is expressed at 5 to 15-fold higher levels in stable clones as compared to the vector alone cells [Fig. 1(A)]. Western blot using anti-flag antibody shows specific expression of flag-gankyrin in transfected cells [Fig. 1(B)]. Some of the reported phenotypic characteristics seen in gankyrin overexpressing cells, like increase in proliferation<sup>27</sup> [Supporting Information Fig. S1(A)] and reduction in NF- $\kappa$ B activity<sup>20</sup> [Supporting Information Fig. S1(B)] were recapitulated. In addition

we found that gankyrin overexpression results in enhanced growth of colonies on soft agar [Fig. 1(C)]. This anchorage independent growth assay is considered as a surrogate assay for *in vivo* tumorigenesis.<sup>28</sup> Since number and size of colonies were much larger upon gankyrin overexpression, additional malignant features are only introduced in these transformed HEK 293 cells in response to gankyrin overexpression. Gankyrin overexpression also resulted in resistance to apoptosis upon

etoposide treatment [Fig. 1(D)]. Thus these HEK 293 stable clones simulate oncogenic properties of gankyrin.

### Bioinformatic identification of novel interacting partners of gankyrin

We took advantage of the published crystal structure of mouse gankyrin (m-gankyrin) in complex with human S6ATPase subunit of the proteasome<sup>29</sup> to predict novel interacting partners of human gankyrin. The complex buries 2418 Å<sup>2</sup> of surface area. The human (h) and m-gankyrin share 98% sequence identity and the crystal structure of unliganded h-gankyrin is completely superimposable on the m-gankyrin-S6ATPase structure [Supporting Information Fig. S2(A)]. In this structure, the interface is enriched in polar and charged residues. A linear sequence of four amino acids EEVD (aa 356–359) in S6 ATPase [Supporting Information Fig. S2(B)] makes dominant contribution to the interaction. Residue E356 is hydrogen bonded to Lys116. E357 forms a salt bridge with Arg41 and is also hydrogen bonded to Ser49 and Ser82. Val358 main chain carbonyl group is engaged in a hydrogen bond with Lys116. Asp359 is in polar contact with Ser115. Bridging this contiguous stretch of contact are two arginine residues Arg339 and Arg342 at one end and Lys397 at the other and these interact with residues in gankyrin through hydrogen bonds and salt bridges [Supporting Information Fig. S2(B)].

Complex formation was abolished by any one of the double mutations of E356/E357A or D359/362A, whereas triple mutation of arginine residues, R338/R339/R342A was required to abolish complex formation. Although these results have not been proven in direct interaction assays using purified components, these observations indicate that residues E356, E357 and D359 within EEVD are crucial for interaction. While the Arg residues are present in a helix, E356/E357/D359 residues are present in a loop and E362 is present in a helix. We tested the role of EEVD in complex formation between S6ATPase and gankyrin by overexpressing Flag-S6ATPase Wt and its corresponding AAVA mutant in HEK 293 cells followed by immunoprecipitation. Results indicate that the mutations significantly abrogate interaction [Supporting Information Fig. S2(C)]. Based on these results and evidence cited above, we chose to test the short linear stretch of EEVD present in the loop as a probable hot spot sites at this interface. We hypothesized that other proteins within the human proteome which contain a contiguous stretch of EEVD may also interact with gankyrin. Short sequences such as EEVD are customarily believed to be promiscuous. However, our experience suggest that judicious combination of functionally important short sequences with filters based on structure and other biological properties can reduce false positive identification in proteome wide screening. “Prediction of Natural Substrates from Artificial Substrate of Proteases,”

(PNSAS) an algorithm developed by us for predicting natural substrates of endo proteases uses such principles.<sup>30</sup> Power of the method lies in the use of short sequence motifs coupled with physiologically relevant filters namely, accessibility of the cleavage site in the folded protein and subcellular co-localization to reduce false positive identification.<sup>30</sup> We have validated this algorithm by identifying Dsg-2, a desmosomal protein involved in cell-cell junction as a physiologically relevant novel substrate of a transmembrane protease called matriptase. Ability of matriptase to regulate levels of Dsg-2 at cell surface is likely to play a crucial role in cell invasion and metastasis.<sup>31</sup> In addition, we recently demonstrated that a 13-residue peptide based on the sequence of apomyoglobin can inhibit interaction of the full length protein with the proteasome.<sup>32</sup> Other bioinformatic analysis indicate that sequences of about 3 to 10 residues are signatures of unique biological function.<sup>33</sup> Short linear motifs primarily in disordered regions of the proteins, terminal residues and three amino acid motifs can also assign functions to proteins.<sup>34,35</sup> These motifs have been used to identify novel substrates of kinases harboring SH2-SH3 domains and proteins like 14-3-3 which recognize such short sequences in phosphorylated proteins.<sup>34</sup>

With this background, we decided to test whether proteins from the human proteome which harbor EEVD in the accessible region of the protein can interact with gankyrin. Ability of gankyrin to interact with multiple proteins would lead to a system wide network some of which may dictate cancer specific phenotypes. A total of 264 unique proteins with EEVD in their primary sequence were found (Supporting Information Table S2) and for 34 of them structural information was available using which solvent accessible surface area (SASA) values were calculated. For comparative analysis SASA values were normalized either against EEVD from S6ATPase (considered as 1) or they were obtained from the POPS server. For 4 proteins the corresponding structure did not have any electron density for the peptide sequence and these were verified to be disordered (Table I).

### Choice of proteins

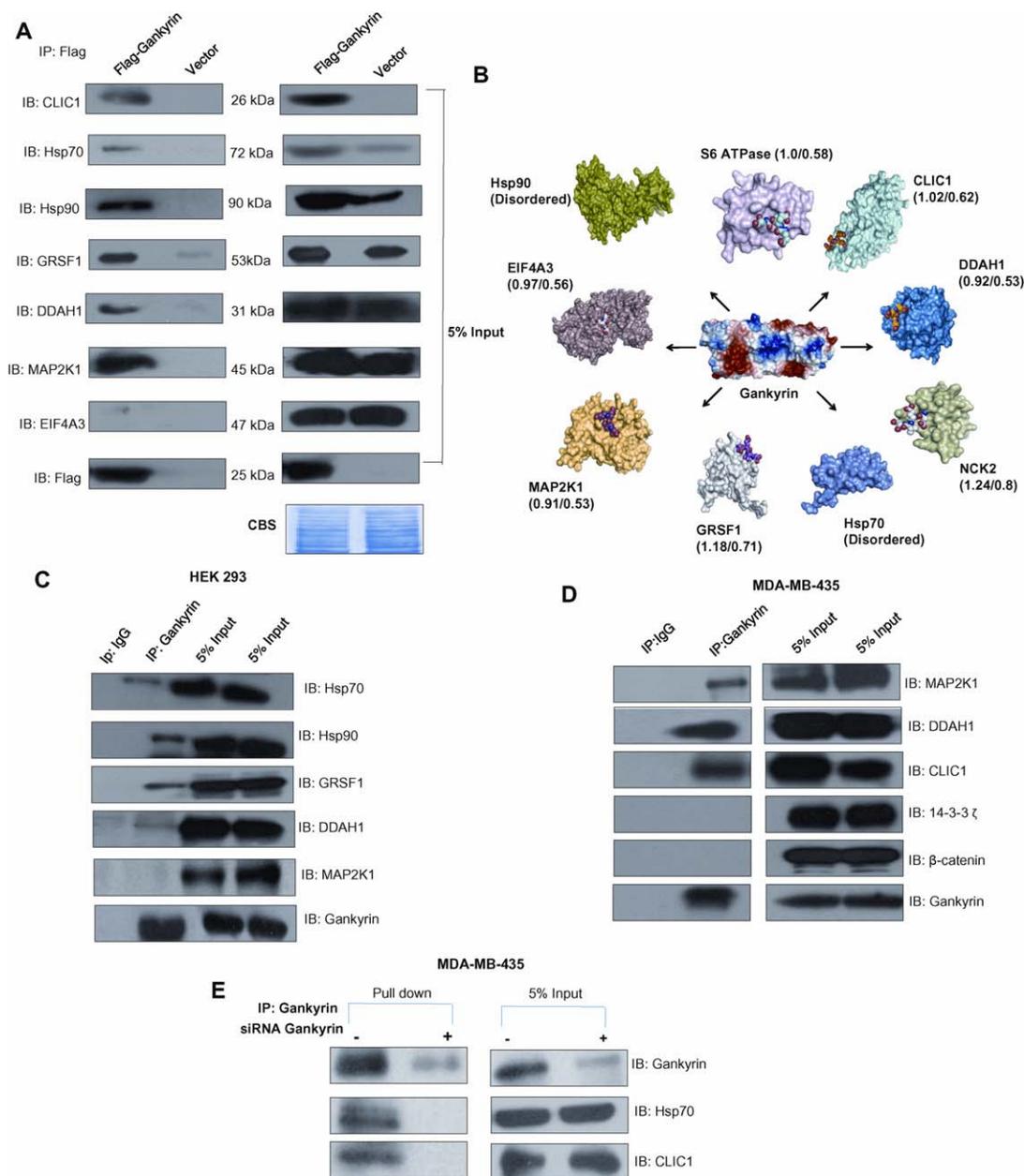
To select proteins for experimental validation, we used a stringent cut off for relative solvent accessibility (rSASA). Among the 34 proteins with available crystal structure, we chose proteins with rSASA > 0.5 as determined by POPS server. Among the 11 proteins with rSASA > 0.5, we chose to test NCK2, G-rich RNA sequence binding factor 1 (GRSF1), chloride intracellular channel protein 1 (CLIC1), eukaryotic initiation factor 4A-III (EIF4A3), dimethylarginine dimethylaminohydrolyase 1 (DDAH1), and mitogen-activated protein kinase 1 (MAP2K1) as putative interacting partners. We also included two proteins (heat shock protein 70-Hsp70,

**Table 1**

EEVD is Solvent Accessible in Putative Interacting Partners of Gankyrin

Uniprot ID	Protein Name	PDB ID	rSASA from reference: 2DVW	rSASA from pops program
NCK2_HUMAN	Cytoplasmic protein NCK2	4E6R	1.24	0.8
GRSF1_HUMAN	G-rich sequence factor 1	2LMI	1.18	0.71
BPNT1_HUMAN	3'/(2')5'-bisphosphate nucleotidase 1	2WEF	1.15	0.67
DNMT1_HUMAN	DNA (cytosine-5)-methyltransferase 1	3SWR	1.03	0.62
CLIC1_HUMAN	Chloride intracellular channel protein 1	3UVH	1.02	0.62
PRS6B_HUMAN	26S protease regulatory subunit 6B	2DVW	1	0.58
IF4A3_HUMAN	Eukaryotic initiation factor 4A-III	2JOS	0.97	0.56
RIR2B_HUMAN	Ribonucleoside-diphosphate reductase subunit M2 B	4DJN	0.92	0.54
DDAH1_HUMAN	N(G)N(G)-dimethylarginine dimethylaminohydrolase 1	3I2E	0.92	0.53
MP2K2_HUMAN	Dual specificity mitogen-activated protein kinase kinase 2	1S9I	0.91	0.54
MP2K1_HUMAN	Dual specificity mitogen-activated protein kinase kinase 1	3EQC	0.91	0.53
RIR2_HUMAN	Ribonucleoside-diphosphate reductase subunit M2	30LJ	0.84	0.49
REV1_HUMAN	DNA repair protein REV1	3GQC	0.81	0.47
PPIE_HUMAN	Peptidyl-prolyl cis-trans isomerase E	3MDF	0.80	0.46
CALM_HUMAN	Calmodulin	2LL6	0.79	0.46
HLTF_HUMAN	Helicase-like transcription factor	2L1I	0.78	0.45
HECW2_HUMAN	E3 ubiquitin-protein ligase HECW2	2LFE	0.77	0.44
SYT1_HUMAN	Synaptotagmin-1	2R83	0.69	0.40
CALL3_HUMAN	Calmodulin-like protein 3	1GGZ	0.66	0.38
FBXL5_HUMAN	F-box/LRR-repeat protein 5	3U9J	0.66	0.38
DYN1_HUMAN	Dynamin-1	3SNH	0.65	0.38
SIR5_HUMAN	NAD-dependent protein deacetylase sirtuin-5	3RIY	0.60	0.35
LMNA_HUMAN	Prelamin-A/C	1IVT	0.59	0.35
SCO1_HUMAN	Protein SCO1 homolog	2HRN	0.57	0.34
SETB1_HUMAN	Histone-lysine N-methyltransferase SETDB1	3DLM	0.54	0.32
ROA1_HUMAN	Heterogeneous nuclear ribonucleoprotein A1	1U1P	0.53	0.31
PZRN3_HUMAN	E3 ubiquitin-protein ligase PDZRN3	1WH1	0.52	0.31
SH3L2_HUMAN	SH3 domain-binding glutamic acid-rich-like protein 2	2CT6	0.42	0.25
OTC_HUMAN	Ornithine carbamoyltransferase	10TH	0.42	0.25
AIMP1_HUMAN	Aminoacyl tRNA synthase complex-interacting multifunctional protein 1	1FLO	0.28	0.17
HS90A_HUMAN	Heat shock protein HSP 90-alpha	3Q6M		Disordered
SLD5_HUMAN	DNA replication complex GINS protein SLD5	2E9X		Disordered
HSP71_HUMAN	Heat shock 70 kDa protein 1 A/1 B	3LOF		Disordered
TB22A_HUMAN	TBC1 domain family member 22A	2QFZ		Disordered

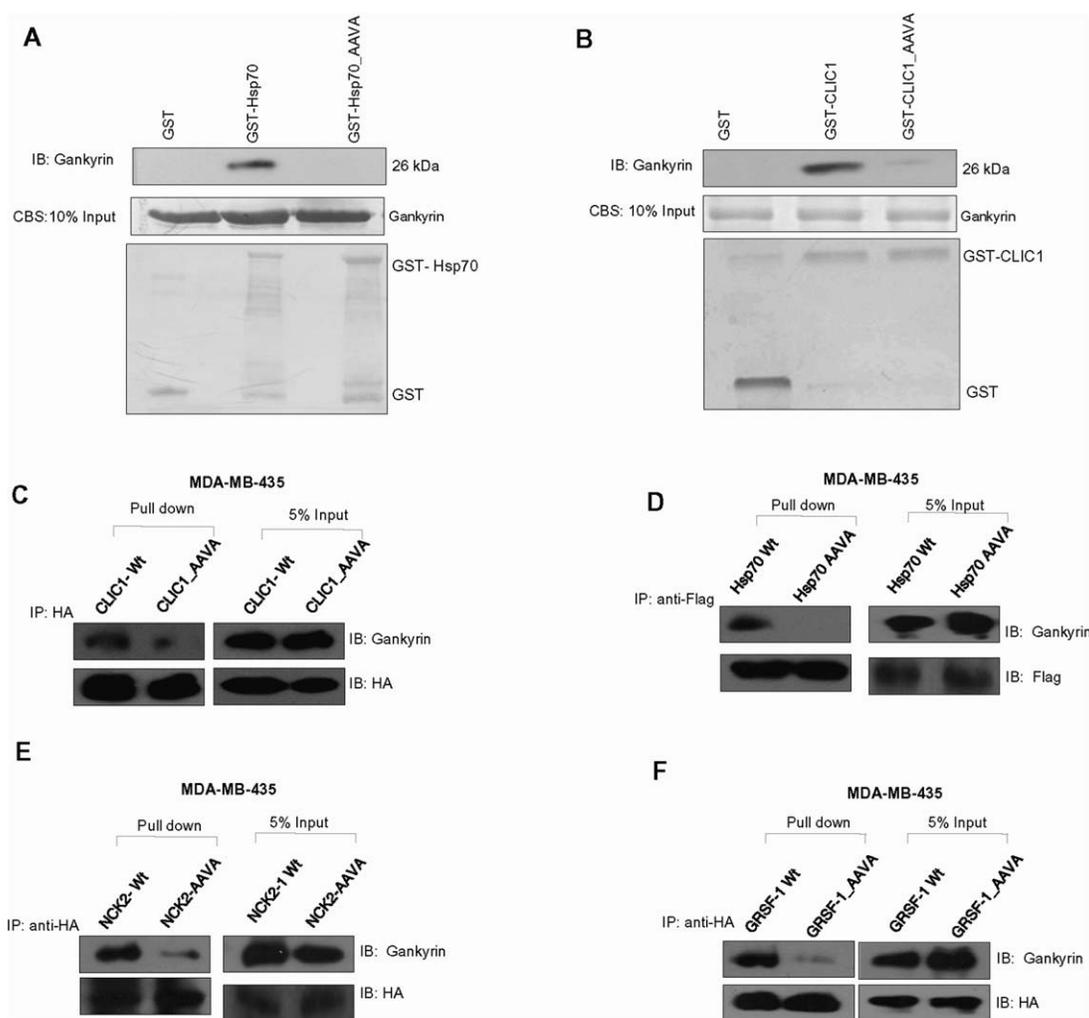
Solvent accessible surface area values for EEVD in each sequence was calculated using the POPS program. Values were normalized (relative SASA or rSASA) using the SASA values for EEVD in the gankyrin-S6 ATPase complex. The values calculated from POPS program are also tabulated. The value for S6 ATPase was obtained in its uncomplexed form.

**Figure 2**

Within the cellular milieu, gankyrin interacts with CLIC1, Hsp70, Hsp90, GRSF1, DDAH1, MAP2K1, and does not interact with EIF4A3. (A) Lysate of HEK 293 cells expressing flag-gankyrin or flag alone (vector) were immobilized on M2 agarose anti-flag beads followed by immunoblotting with anti-Hsp70, anti-Hsp90, anti-GRSF1, anti-DDAH1, anti-CLIC1, anti-MAP2K1, anti-EIF4A3 antibody; 5% input represents loading control. CBS indicates equal loading. Note that some proteins are upregulated upon gankyrin overexpression; (B) Gankyrin (centre) is represented using surface electrostatic potential and interacting partners with rSASA value  $>0.5$  are represented as protein surfaces. ‘Hot spot’ site of interaction EEVD is represented in spheres. We believe that gankyrin is the hub protein and the putative interacting partners harbouring the hot spot site EEVD act as potential nodes to create a system wide network; (C) Lysate of HEK 293 cells were added to gankyrin bound and IgG bound Sepharose G beads followed by immunoblotting with anti-Hsp70, anti-Hsp90, anti-GRSF1, anti-DDAH1, anti-MAP2K1 and anti-gankyrin antibody. MAP2K1 does not interact with gankyrin at endogenous levels; (D) Lysate of MDA-MB-435 cells were added to gankyrin bound or IgG bound Sepharose G beads followed by immunoblotting with anti-MAP2K1, anti-DDAH1, anti-CLIC1, anti-14-3-3 zeta, anti- $\beta$ -catenin and anti-gankyrin antibody; (E) Lysate of MDA-MB-435 cells expressing siRNA against gankyrin and MDA-MB-435 cells were added to gankyrin bound Sepharose G beads followed by immunoblotting with anti-Hsp70, anti-CLIC1 and anti-gankyrin antibody.

heat shock protein90-Hsp90) out of four where EEVD was present in the disordered region. All these proteins are expressed in the cytoplasm or both in cytoplasm and

nucleus. Gankyrin is a nuclear cytoplasmic shuttling protein.<sup>20</sup> Hsp70 and Hsp90 proteins carrying a putative signature motif for gankyrin binding intrigued us and it

**Figure 3**

Gankyrin binds to Hsp70, CLIC1, NCK2 and GRSF1 through E, E and D residues. (A and B) Purified GST-Hsp70, GST-CLIC1 and their respective AAVA mutants were independently immobilized on glutathione-Sepharose 4B and incubated with 6×His tagged gankyrin. Bound His-gankyrin was detected using anti-gankyrin antibody. Wt proteins interact with His-gankyrin while the mutants are unable to do so. GST alone does not interact with gankyrin. The input for gankyrin and GST-fusion proteins are shown; (C) Lysate of MDA-MB-435 cells expressing HA-CLIC1 Wt or CLIC1-AAVA; (D) Flag-Hsp70 Wt or Hsp70-AAVA; (E) HA-NCK2 Wt or NCK2-AAVA; (F) HA-GRSF1 Wt or GRSF1-AAVA were independently added to anti-HA antibody bound to Sepharose G beads followed by immunoblotting with anti-gankyrin antibody. Gankyrin interacts with CLIC1 Wt, Hsp70 Wt, NCK2 Wt, GRSF1 Wt but not with the corresponding AAVA mutants.

seemed important to ask if native gankyrin would actually bind these chaperones which may directly influence their function. Interestingly, Hsp90 plays a role in the assembly of 26S proteasome.<sup>36</sup> Functions of these proteins were discerned from Panther data base<sup>37</sup> and their association with cancer was verified using the classic eight hall mark properties of cancer.<sup>38</sup> Six out of eight proteins that we chose carried hall mark cancer properties (Supporting Information Table S3). Apart from their role as conventional chaperones, Hsp70 and Hsp90 in cancer cells are very important in conferring resistant to cell death, in malignant transformation and in generation of antigenic peptides.<sup>39</sup> While their well characterized role seems to be in binding to unfolded

proteins via degenerate but specific sequence motifs these chaperones also recognize native or near native proteins under unique circumstances.<sup>40,41</sup> These chaperones also bind to kinases.<sup>39</sup> Geldanamycin, a known anti-cancer agent targets Hsp90 in cancer cells. CLIC1, a chloride intracellular channel protein is involved in metastasis and invasion.<sup>42,43</sup> In many cancer cells, voltage gated channels show this property and are considered attractive targets for cancer therapy.<sup>44</sup> NCK2 binds to growth factor receptors or their substrates and modulates gene expression in response to Kras signaling.<sup>45,46</sup> Gankyrin is a key regulator of Ras mediated activation and seems to specifically activate PI3K/Akt pathway. Evidence suggests that gankyrin is highly expressed in human lung cancer

having Kras mutations which results in Akt activation.<sup>7</sup> MAP kinase pathway is involved in sustained proliferation.<sup>47</sup> DDAH1 is a positive regulator of angiogenesis.<sup>48</sup> EIF4A3 is an abundant DEAD-box RNA helicase and a member of the eukaryotic initiation factor-4A (eIF4A) family of translation initiation factors.<sup>49</sup> Presence of EEVD in a well accessible region and the functions of these proteins make them good candidates to test our presumptions. Structure of these proteins with EEVD in the accessible region (where applicable) along with their rSASA values are shown in Figure 2(B).

### Experimental validation of predicted interactions

For experimental validation of predicted interactions we used affinity pull down of gankyrin expressed as a flag tagged fusion in HEK 293 cells. Six proteins Hsp70, Hsp90, CLIC1, GRSF1, DDAH1 and MAP2K1 were present in the pull down complex indicating that the predicted sequence EEVD may be involved in the interaction between each of these client proteins and gankyrin [Fig. 2(A)]. However EIF4A3 was not found in the pull down complex. NCK2 antibody failed to detect protein from cell lysate even when used at lower dilutions than recommended. Owing to the important role of this protein in cancer we expressed this protein in mammalian cells and show that interaction is mediated through EEVD (see below).

Due to similar subcellular localization of gankyrin and the client proteins it is possible that some of these interactions may occur constitutively and at endogenous levels. It is also possible that some of these interactions may occur in response to gankyrin overexpression and therefore may be unique to malignant phenotype. To test these possibilities we used anti-gankyrin antibody immobilized on protein G Sepharose beads to pull down the endogenous complexes in HEK 293 cells. Three proteins (Hsp70, Hsp90 and GRSF1) were found in the immune complex while DDAH1 was barely detectable [Fig. 2(C)]. MAP2K1 did not interact with gankyrin [Fig. 2(C)]. CLIC1 was hardly detectable in lysate of HEK 293 cells but was several fold overexpressed in HEK 293 stable clones [Fig. 2(A); input]. Levels of MAP2K1 and DDAH1 were not very different between vector control and gankyrin overexpressing HEK 293 cells. Enough evidence exists in literature to demonstrate the role of heat shock proteins in non chaperone function which may go beyond their ability to recognize mis-folded proteins. Chaperones also interact with their co-chaperones through specific interactions between native proteins to execute their function.<sup>39</sup> It is interesting to note that the EEVD in Hsp70 and Hsp90 is involved in interaction with other proteins.<sup>40</sup>

To test if interactions seen in response to gankyrin overexpression in HEK 293 also occur in cancer cells

where gankyrin is known to play a role in the deregulated phenotype, we used a breast cancer cell line. Endogenous levels of gankyrin mRNA reportedly is less in normal and other tumor tissues as compared to breast cancer tissues.<sup>6</sup> Therefore we used MDA-MB-435, a breast tumor derived cell line to check if the interactions observed upon gankyrin overexpression in HEK 293 cells are also seen in this cell line. MAP2K1, DDAH1, and CLIC1 were found to interact with gankyrin at endogenous levels in MDA-MB-435 cells [Fig. 2(D)]. Proteins which do not have EEVD like 14-3-3 $\zeta$  or  $\beta$ -catenin which carries a variant EEED do not interact with gankyrin under the same conditions [Fig. 2(D)]. Interestingly we found that  $\beta$ -catenin levels increase in HEK 293 cells upon gankyrin overexpression but this protein as in MDA-MB-435 cells does not interact with gankyrin (data not shown). Specificity of interaction between gankyrin and EEVD containing proteins CLIC1 and Hsp70 in MDA-MB-435 cells was further confirmed by repeating IP upon gankyrin knock down using specific siRNA [Fig. 2(E)]. Normal cellular homeostasis requires turnover of proteins and spatiotemporal regulation of protein interactions. By stabilizing existing networks or creating new networks gankyrin may induce deregulated phenotype that may be partly responsible for malignancy.

### Direct interactions of client proteins and evidence for EEVD as the hot spot site of interaction

When methods used to identify protein-protein interactions are compared, techniques which demonstrate physical association seem to be more faithful representatives of functionally relevant interactions; 90 to 95% interactions established by direct/indirect immune precipitation studies and 80% of *in vitro* biochemical studies correlate with the cellular role.<sup>50</sup>

We followed affinity pull down assays with *in vitro* biochemical experiments to demonstrate direct interactions between gankyrin and the client proteins identified in this study. Recombinant form of gankyrin was expressed as a his-tagged fusion protein while the client proteins Hsp70 and CLIC1 were expressed as GST-fusions. Gankyrin was present only in fractions containing GST-fusion proteins and not in GST alone control [Fig. 3(A,B)]. To confirm that these direct interactions are mediated through residues E, E and D in each case, the sequence EEVD was mutated to AAVA and GST pull down assay was repeated. Mutations abolished interaction of each of these proteins with gankyrin [Fig. 3(A,B)]. To verify whether the *in vitro* direct interaction between recombinant proteins is also sensitive to mutations within EEVD within the cellular milieu, HA-CLIC1, Flag-Hsp70, HA-NCK2 and HA-GRSF1 and their corresponding AAVA mutant were cloned and expressed in MDA-MB-435 cells. While all the Wt proteins showed

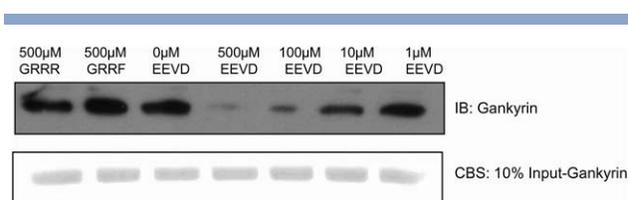
definitive interaction with endogenous gankyrin [Fig. 3(C–F)], as expected the corresponding mutant proteins failed to interact. Interaction of endogenous gankyrin with Wt NCK2 and Wt GRSF1 and not the corresponding AAVA mutant were also detected in HEK 293 cells [Supporting Information Fig. S3(A,B)]. These results establish that observed interactions within the cells are direct and EEVD is the shared hot spot site at the interface.

To further illustrate our point, we tested the ability of naked short peptide sequence EEVD to inhibit binding of CLIC1 (used as a representative example) to gankyrin. In the *in vitro* pull down assay, peptide EEVD prevented complex formation between the two proteins in a concentration dependent manner [Fig. 4]. Control peptides that do not interact with gankyrin but interact with another subunit of the proteasome (unpublished work) did not prevent interaction even at 0.5 mM, the highest concentration tested for EEVD. Using a semiquantitative approach, peptide concentration required for 50% inhibition in binding was estimated to be approximately 50  $\mu$ M (Supporting Information Fig. S4). This value for a short peptide lies within the range reported for other peptides used to inhibit protein-protein interactions<sup>51</sup> (Supporting Information Table S4). These results provide unequivocal evidence for the presence of a hot spot motif shared by proteins in the gankyrin hub network.

#### Interaction between Gankyrin and CLIC1 is important for cell migration and invasion in MDA-MB-231 and HEK 293 cells

In order to evaluate the functional relevance of the interaction between gankyrin and proteins containing EEVD, we chose to study function of gankyrin-CLIC1 interaction because of the following reasons: (a) CLIC1 was barely detectable in HEK 293 Wt cells and levels increase in response to gankyrin overexpression; (b) Chloride intracellular channel proteins are well known for their role in invasion and metastasis.<sup>42–44</sup> We initially tested the independent role of gankyrin and CLIC1 in the metastatic potential of MDA-MB-231 by performing surrogate wound healing and invasion assays. Knock down of either of the two proteins decreased the ability of these cells to migrate and close the wound. CLIC1 seems to be more effective than gankyrin in healing the wound (Fig. 5). These results were recapitulated in the invasion assay using Boyden Chamber [Supporting Information Fig. S5(A)]. These results confirmed the role of these two proteins in migration and invasion.

To check if the interaction is important for driving this phenotype CLIC1 Wt or the AAVA mutant of CLIC1 were overexpressed in MDA-MB-231 cells as well as in HEK 293 cells and their ability to migrate and close the wound was monitored. In MDA-MB-231 cells, overexpression of CLIC1 Wt protein resulted in an increase



**Figure 4**

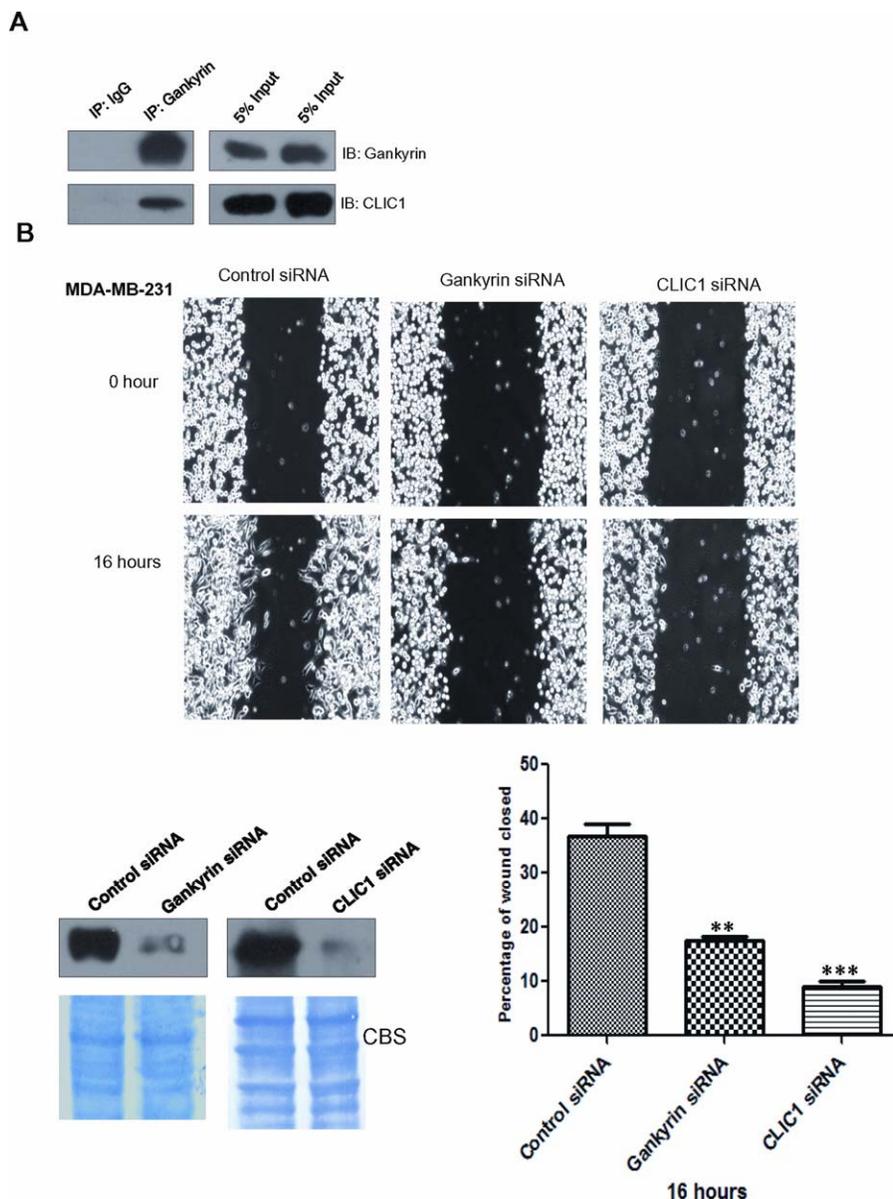
Interface peptide EEVD inhibits interaction of full length protein with gankyrin. His-Gankyrin was incubated with different concentrations of EEVD peptide. Peptides GRRF and GRRR were used as controls. These preincubated complexes were added to the GST-CLIC1 bound GST beads. EEVD peptide and the two control peptides inhibited binding of His-Gankyrin to GST-CLIC1 in a concentration dependent manner.

in the percentage of wound closed (48%) as compared to vector control (31%) while the mutant cells (27%) behaved like those of the vector control cells (Fig. 6). HEK 293 cells overexpressing CLIC1 Wt showed ~98% wound closure as compared to vector control cells (42%) and cells overexpressing CLIC1\_AAVA mutant showed 27% wound closure. MDA-MB-231 cells overexpressing CLIC1 Wt showed increase in their invasive potential as compared to vector control and CLIC1\_AAVA overexpressing cells [Supporting Information Fig. S5(B)] behaved like the vector control cells. We further confirmed these observations by specifically silencing endogenous CLIC1 using smart pool of UTR specific siRNA and demonstrating that only Wt CLIC1 and not the mutant protein can rescue the “phenotype” in MDA-MB-231 cells (Fig. 7). These results taken together indicate that interaction of CLIC1 with gankyrin through EEVD enhances the migratory potential of CLIC1 in this *ex vivo* model.

## DISCUSSION

Protein-protein interactions are central to cellular communication.<sup>1</sup> These interactions are stabilized or new interactions created in abnormal conditions like cancer.<sup>1</sup> Therefore identifying system wide interactions of a key regulatory protein like an oncoprotein or a tumor suppressor protein is crucial for understanding the deregulated phenotype. Many studies aimed at characterizing network of interactions show that oncoproteins act as hubs. Perturbation of key interactions of such hubs with other proteins in the node would render the network unstable.<sup>22</sup> Therefore these interactions are vulnerable for therapeutic intervention.

In the present study we illustrate one method to identify multiple interacting partners of a hub protein. We use residue level knowledge of a known protein interface of gankyrin-S6ATPase complex to predict unknown complexes of gankyrin.<sup>1,52–54</sup> We asked whether some of the basic principles common to protein-protein interactions identified by independent investigators can be used

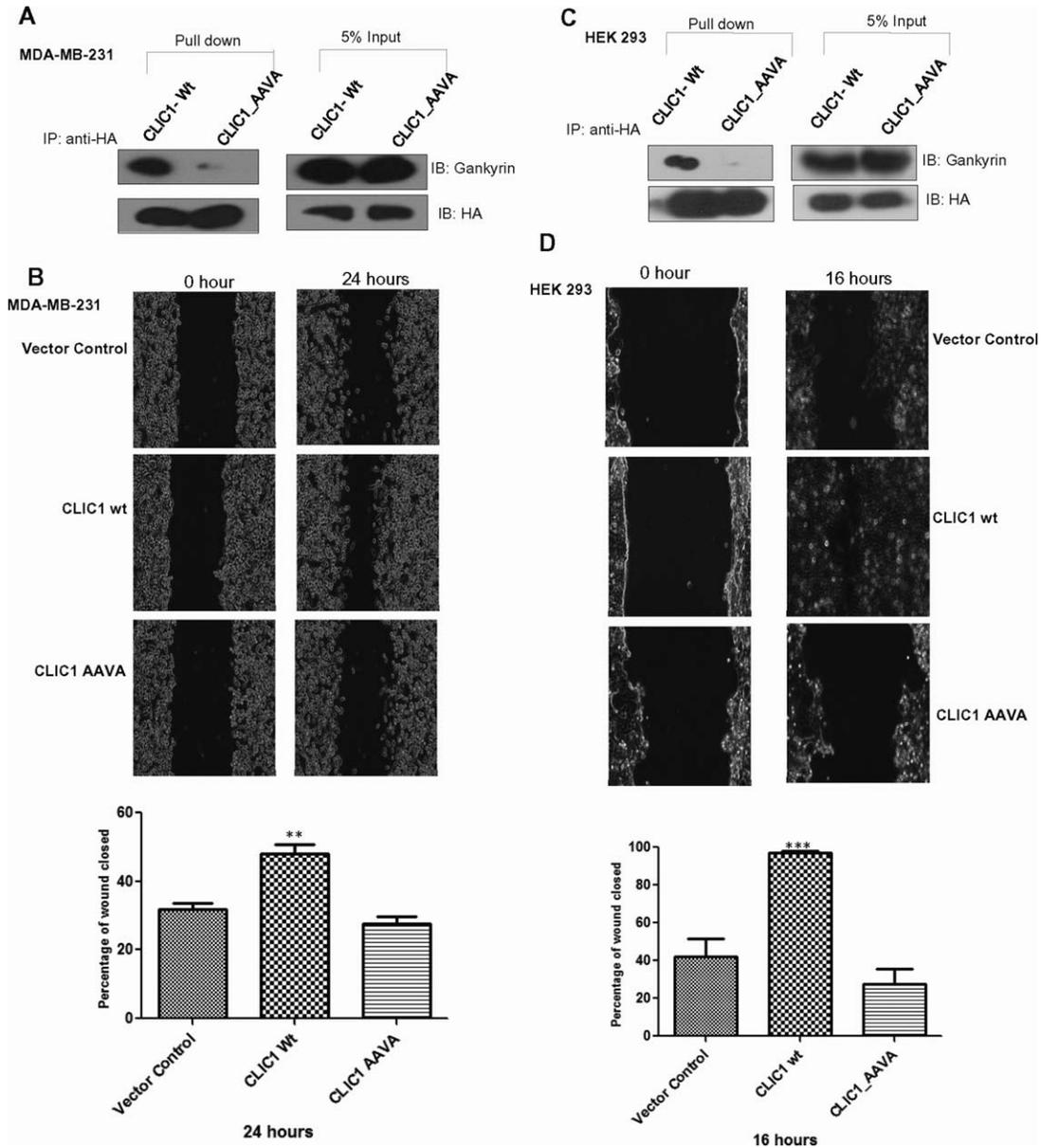
**Figure 5**

Silencing of gankyrin and CLIC1 affects migration in MDA-MB-231 cells. (A) Lysate of MDA-MB-231 cells were added to gankyrin bound or IgG bound Sepharose G beads followed by immunoblotting with anti-CLIC1 and anti-gankyrin antibody; (B) MDA-MB-231 cells were treated with 100 nM of gankyrin specific siRNA or CLIC1 siRNA or control siRNA. After 72 h, cells at 90% confluence were treated with 10  $\mu$ M of mitomycin C for 3 h, wounded and healing monitored for 16 h. Percentage wound healed at the end of 16 h were calculated and the data are represented as mean  $\pm$  SEM of three experiments. Statistical analysis is done using unpaired *t*-test (\*\* $P$  = 0.0013, \*\*\* $P$  = 0.0004). Western blot confirms down regulation of gankyrin and CLIC1 in the respective experiments. CBS shows equal loading. [Color figure can be viewed in the online issue, which is available at [wileyonlinelibrary.com](http://wileyonlinelibrary.com).]

to predict large number of novel interactions? Basic assumptions of our method are the following: (a) interactions of a protein with multiple partners may involve a common recognition motif (b) this recognition motif may be independent of sequence homology among the client proteins (c) hub protein like gankyrin whose deregulated functions are dependent on overexpression are likely to recognize such common motifs in non-

homologous proteins. This may explain their ability to alter multiple pathways and drive the oncogenic phenotype (d) since hot spot sites are conserved across interfaces and represent very few residues, they may also be vulnerable for interventions.

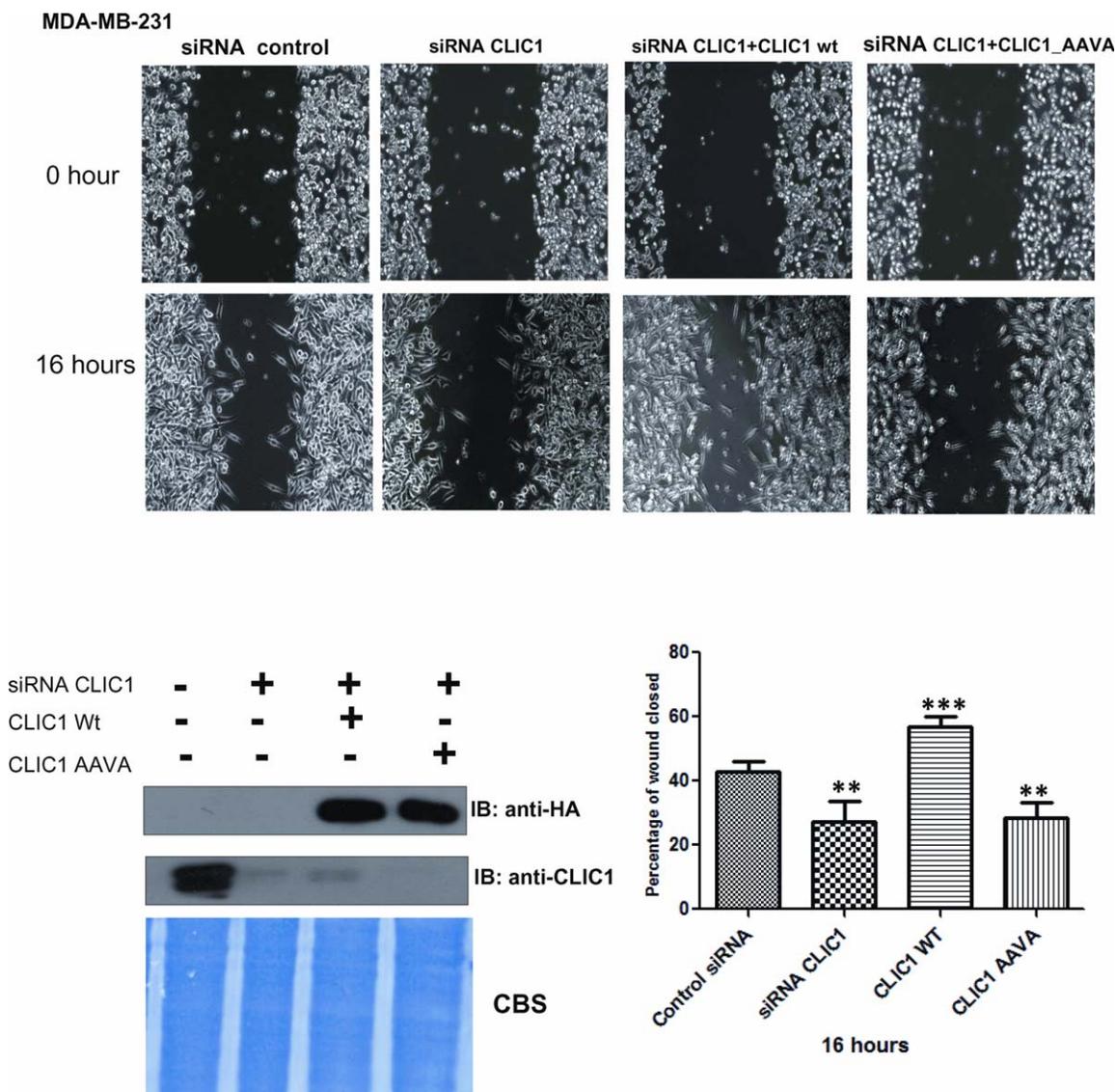
Protein interfaces are determined by shape complementarity, enriched in patches of buried salt bridges and pairing of hydrophobic and hydrophilic hot spots which

**Figure 6**

Overexpression of CLIC1 and CLIC1\_AAVA affects migration of MDA-MB-231 cells. (A) Lysate of MDA-MB-231 cells expressing HA-CLIC1 Wt and CLIC1\_AAVA were independently added to anti-HA antibody bound to Sepharose G beads followed by immunoblotting with anti-gankyrin antibody. Gankyrin interacts with CLIC1 Wt but not with its corresponding mutant; (B) MDA-MB-231 cells transiently overexpressing CLIC1 Wt or CLIC1\_AAVA were treated with 10  $\mu\text{g}/\text{mL}$  of mitomycin C for 3 h, wounded and healing monitored for 24 h. Percentage of wound healed was calculated and the data are represented as mean  $\pm$  SEM of three independent experiments (\*\* $P = 0.0059$ ); (C) Lysate of HEK 293 cells expressing HA-tagged CLIC1 Wt and CLIC1\_AAVA were independently added to anti-HA antibody bound to Sepharose G beads followed by immunoblotting with anti-gankyrin antibody. Gankyrin interacts with CLIC1 Wt but not with its corresponding mutant; (D) HEK 293 cells transiently overexpressing CLIC1 Wt and CLIC1\_AAVA were treated with 10  $\mu\text{g}/\text{mL}$  of mitomycin C for 3 h, wounded and healing monitored for 16 h. Data were processed as above (\*\* $P = 0.0002$ ).

are juxtaposed on the opposite sides of the interface.<sup>55</sup> As described above, gankyrin-S6 interface is enriched in charged residues at the interface and buries a large surface area upon complex formation. Residues like tryptophan, glycine, proline, cysteine, tyrosine, and glutamate are apparently more conserved at hot spot sites.<sup>55</sup> Aspar-

tate is also considered as a relatively common hot spot site residue.<sup>24</sup> Besides as described above, we found that in mammalian cells mutation of EEVD to AAVA in S6ATPase prevents interaction with gankyrin [Supporting Information Fig. S2(C)]. Based on the strength of these observations using a linear stretch of a four amino acid

**Figure 7**

Interaction of gankyrin with CLIC1 enhances migration. MDA-MB-231 cells were transfected with control siRNA or CLIC1 siRNA (UTR region) or CLIC1 siRNA + cDNA for CLIC1 Wt or CLIC1 siRNA + cDNA for CLIC1\_AAVA. After 72 h these cells were treated with 10  $\mu\text{g}/\text{mL}$  of mitomycin C for 3 h, wounded and healing monitored for 16 h. Data from two independent experiments were processed as described in Figure 5 (\*\* $P = 0.0042$ , \*\*\* $P = 0.0006$ , \*\* $P = 0.0020$ ). Cells transfected with CLIC1 Wt show significant increase in the percentage of wound healed as compared with that of CLIC1\_AAVA transfected cells which behave like the vector control cells. [Color figure can be viewed in the online issue, which is available at [wileyonlinelibrary.com](http://wileyonlinelibrary.com).]

sequence EEVD at the S6ATPase gankyrin interface, 34 novel interacting partners from the PDB data of human proteins were predicted. Stringent cut offs were used for testing our hypothesis. Proteins in which EEVD was present in an accessible region (with rSASA values  $>0.5$ ) were chosen. Interactions within the cell would be dictated by factors other than sequence specificity viz. subcellular distribution, post translational modifications and signaling cues.<sup>30</sup> Proteins present in the cytoplasm like gankyrin were short listed. Some of these may also be resident in other subcellular compartments like nuclei as well.

Among the eight proteins tested, seven interactions were seen in at least two stable clones and in three independent experiments. Four of these interactions occur at endogenous levels indicating that these are of sufficient affinity to be detected whereas three are unique to gankyrin overexpression either in stable clones or in breast cancer cells. These are probably of lower affinity or occur in response to oncogene induced signaling. Not all proteins containing EEVD in an accessible region interact with gankyrin as seen by the failure of EIF4A3 to be detected in the pull down complex [Fig. 2(A)]. This is

not surprising since there are additional levels of regulation that govern PPI. Interaction between recombinant proteins and failure of AAVA mutants to interact establishes that, observed interactions are in fact through predicted EE and D residues. Ability of EEVD peptide to compete with full length protein for interaction further substantiates that bulk of the binding energy indeed comes from this short sequence motif. Therefore this motif is a conserved hot spot site interaction shared among the client proteins. Although it is unclear if interacting residues in gankyrin are the same in all cases, interaction interfaces of the same protein, mediating different interactions are more likely to cause distinct interruptions in the overall interactome. This has different biological consequences leading to pleiotropic effects.<sup>56</sup> We believe that one of the major reasons for the observed oncogenic property of gankyrin in multiple cell types is its ability to interact with different partners through the conserved hot spot sites on the client proteins. There are many more EEVD containing proteins in the human proteome that remain to be tested. Moreover based on our mutagenesis experiments valine at the EEVD hot spot site does not seem to have a major contribution to binding affinity. It seems therefore that other sequence variants of EEVD such as EEXD (X is any residue) may also interact with gankyrin. This is supported by the fact that when Val is mutated to Glu in CLIC1 the mutant protein still binds to gankyrin [Supporting Information Fig. S3(C)]. On the other hand  $\beta$ -catenin which carries EEED does not interact with gankyrin which is not surprising because interactions are not dictated by sequence or accessibility alone. Subcellular localization and other regulatory mechanisms are well known to play a role in determining interactions.

Among the proteins that were reported to interact with gankyrin such as Rb, CDK4, MDM2, MAGE-A4, p65 (RelA) of NF- $\kappa$ B and FIH-1,<sup>57</sup> EEND is present in MDM2 and Rb carries the tetrapeptide motif EEPD (Supporting Information Table S5). While there is no structural information available for MDM2 in this region, in RB EEPD forms the N-terminus of the solved crystal structure (aa 53EEPD56). Whether these residues are actually involved in interaction with gankyrin however remains to be tested. But it is interesting to note two key proteins known to interact with gankyrin carry the EEXD motif.

#### **How does identification of these novel interactions help in better understanding of the role of gankyrin in oncogenesis and as a key hub protein?**

Based on its role in Rb, p53 degradation, involvement of gankyrin in cell proliferation seems understandable. However, its role in other oncogenic properties like can-

cer metastasis is still unclear. A very recently study show that gankyrin is involved in tumor metastasis.<sup>6</sup>

Two novel interacting proteins CLIC1 and DDAH1 identified here are important in metastasis and angiogenesis.<sup>48,58</sup> DDAH1 is involved in vascular permeability and in NO synthesis.<sup>59,60</sup> It is especially involved in inflammation induced cancer. Increased DDAH1 expression has direct consequence on vascularization and tumor growth.<sup>48</sup> DDAH1 is also predicted to methylate arginine residues in proteins which are likely to affect protein turnover. Here is another example wherein gankyrin via interaction with DDAH1 may cross talk with the proteasome pathway.

CLIC1 protein levels are more when gankyrin is over-expressed. This protein is seen in complex with gankyrin in breast cancer cells [Fig. 2(D)]. Voltage gated channels such as CLIC1 induce metastatic phenotype.<sup>58</sup> Overexpression of CLIC1 is a potential prognostic marker for colorectal carcinoma,<sup>58</sup> gallbladder carcinoma,<sup>61</sup> gastric cancer,<sup>42</sup> lung adenocarcinoma<sup>62</sup> and glioma.<sup>63</sup> Both DDAH1 and CLIC1 play a universal role in metastasis and angiogenesis in many cancer types. In this context it will be interesting to note that these interactions with gankyrin are conserved in both glioma cells and in DAOY medulloblastoma cells (data not shown). These results implicate a more universal role for gankyrin as an important hub in protein-protein interaction network in several different cancer types and therefore a potentially vulnerable target for therapeutic intervention.

#### **What would be the consequence of interaction of gankyrin with these proteins?**

Gankyrin may deregulate normal cellular homeostasis by stabilizing the basal network. It may rewire the network by introducing new players and shift the nodes and edges of the functional networks. Establishing this concept through experimentation is a difficult but an important and mandatory task. We have initiated these studies and in this report we clearly demonstrate that interaction of CLIC1 with gankyrin is mandatory for the ability of MDA-MB-231 cells to migrate/invade as demonstrated by the wound healing assay. Since this property is dependent on the motility of the cells, we believe that this interaction is likely to be crucial in the invasive properties of these metastatic cells in breast cancer.

#### **How does the approach used here compare with similar or different approaches used in the identification of protein-protein interactions?**

Several methods like yeast two hybrid, phage display or mass spectrometry based high throughput detection of affinity enriched complexes are used to identify unknown protein-protein interactions.<sup>2,3,64,65</sup> While first two methods suffer from accuracy, MS based

approaches tend to be more accurate. Experimental strategies like proteome peptide scanning and whole interactome scanning experiments (WISE) which rely on short peptide sequences within proteins (Short Linear Motifs or SLiM) have been used to identify new interactions. However, identifying functionally relevant SLiMs from random occurrences in eukaryotes is not straightforward.<sup>4,34,66–69</sup> This is because many such SLiMs seem to be part of unstructured or disordered regions in proteins with large variability in sequence making it difficult to accurately predict or use this information. Experimental proof from a wide variety of examples such as ours will help in enhancing the reliability and accuracy of these predictions.

All such studies can be reliable and accurate if residue-level structural information on protein-protein interfaces is incorporated in prediction programs or other detection algorithms. However, high-resolution structural information, especially of protein complexes, grows at a much lower pace and therefore there is paucity of information that can be readily extended to large scale identifications. Nevertheless ability to do so will go a long way to better understand deregulated physical and functional networks mediated by oncoproteins in diseases such as cancer. Recently published article makes a bold attempt in predicting novel interactions by modelling structures of unknown complexes.<sup>5</sup>

Some of the well-known interactions with information on hot spot sites are derived from single complex analysis driven by investigator driven interest. These have been deposited in different data bases.<sup>23,24</sup> In addition there are many examples in literature where peptides and small molecules are used as inhibitors of protein-protein interaction emphasizing again the importance of structural information in the success of such approaches.<sup>51,70–85</sup> Supporting information Table S4 lists examples of protein complexes with known interface and peptides/small molecules to inhibit such interactions. It is in this context that the approach presented here gains importance. Our study integrates various key important principles described above for accurately identifying novel protein-protein interactions. We use the concept of (a) SLiM across secondary structural elements including disordered regions in proteins, (b) the concept of sequence conservation at protein interfaces and (c) presence of hot spot sites to identify novel interactions of a hub protein. Judging by the functional annotations of the proteins identified it seems that this network of interactions is likely to bridge at least six of the eight hall mark properties of cancer. Our ability to inhibit these interactions implies that this network of interactions can be perturbed. Disruption of interaction between gankyrin and CLIC1 by mutating the hot spot site inhibits the ability of the cancer cell line MDA-MB-231 cells to migrate or invade thus validating functional relevance of these interactions. Taken together the concept of short linear

sequence motifs at protein interfaces can be used to identify novel functionally relevant protein complexes formed by key hub proteins.

## ACKNOWLEDGMENTS

The authors thank Amit Kumar Singh Gautam for the S6ATPase cDNA.

## REFERENCES

1. Kar G, Gursoy A, Keskin O. Human cancer protein-protein interaction network: a structural perspective. *PLoS Comput Biol* 2009;5:e1000601.
2. Vasilescu J, Figeys D. Mapping protein-protein interactions by mass spectrometry. *Curr Opin Biotechnol* 2006;17:394–399.
3. Vermeulen M, Hubner NC, Mann M. High confidence determination of specific protein-protein interactions using quantitative mass spectrometry. *Curr Opin Biotechnol* 2008;19:331–337.
4. Weatheritt RJ, Luck K, Petsalaki E, Davey NE, Gibson TJ. The identification of short linear motif-mediated interfaces within the human interactome. *Bioinformatics* 2012;28:976–982.
5. Zhang QC, Petrey D, Deng L, Qiang L, Shi Y, Thu CA, Bisikirska B, Lefebvre C, Accili D, Hunter T, Maniatis T, Califano A, Honig B. Structure-based prediction of protein-protein interactions on a genome-wide scale. *Nature* 2012;490:556–560.
6. Zhen C, Chen L, Zhao Q, Liang B, Gu YX, Bai ZF, Wang K, Xu X, Han QY, Fang DF, Wang SX, Zhou T, Xia Q, Gong WL, Wang N, Li HY, Jin BF, Man JH. Gankyrin promotes breast cancer cell metastasis by regulating Rac1 activity. *Oncogene* 2013;32:3452–3460.
7. Man JH, Liang B, Gu YX, Zhou T, Li AL, Li T, Jin BF, Bai B, Zhang HY, Zhang WN, Li WH, Gong WL, Li HY, Zhang XM. Gankyrin plays an essential role in Ras-induced tumorigenesis through regulation of the RhoA/ROCK pathway in mammalian cells. *J Clin Invest* 2010;120:2829–2841.
8. Meng Y, He L, Guo X, Tang S, Zhao X, Du R, Jin J, Bi Q, Li H, Nie Y, Liu J, Fan D. Gankyrin promotes the proliferation of human pancreatic cancer. *Cancer Lett* 2010;297:9–17.
9. Fu XY, Wang HY, Tan L, Liu SQ, Cao HF, Wu MC. Overexpression of p28/gankyrin in human hepatocellular carcinoma and its clinical significance. *World J Gastroenterol* 2002;8:638–643.
10. Li J, Knobloch TJ, Kresty LA, Zhang Z, Lang JC, Schuller DE, Weghorst CM. Gankyrin, a biomarker for epithelial carcinogenesis, is overexpressed in human oral cancer. *Anticancer Res* 2011;31:2683–2692.
11. Tang S, Yang G, Meng Y, Du R, Li X, Fan R, Zhao L, Bi Q, Jin J, Gao L, Zhang L, Li H, Fan M, Wang Y, Wu K, Liu J, Fan D. Overexpression of a novel gene gankyrin correlates with the malignant phenotype of colorectal cancer. *Cancer Biol Ther* 2010;9:88–95.
12. Higashitsuji H, Itoh K, Nagao T, Dawson S, Nonoguchi K, Kido T, Mayer RJ, Arai S, Fujita J. Reduced stability of retinoblastoma protein by gankyrin, an oncogenic ankyrin-repeat protein overexpressed in hepatomas. *Nat Med* 2000;6:96–99.
13. Dawson S, Apcher S, Mee M, Higashitsuji H, Baker R, Uhle S, Dubiel W, Fujita J, Mayer RJ. Gankyrin is an ankyrin-repeat oncoprotein that interacts with CDK4 kinase and the S6 ATPase of the 26 S proteasome. *J Biol Chem* 2002;277:10893–10902.
14. Bedford L, Paine S, Sheppard PW, Mayer RJ, Roelofs J. Assembly, structure, and function of the 26S proteasome. *Trends Cell Biol* 2010;20:391–401.
15. Roelofs J, Park S, Haas W, Tian G, McAllister FE, Huo Y, Lee BH, Zhang F, Shi Y, Gygi SP, Finley D. Chaperone-mediated pathway of proteasome regulatory particle assembly. *Nature* 2009;459:861–865.
16. Higashitsuji H, Higashitsuji H, Itoh K, Sakurai T, Nagao T, Sumitomo Y, Masuda T, Dawson S, Shimada Y, Mayer RJ, Fujita J.

- The oncoprotein gankyrin binds to MDM2/HDM2, enhancing ubiquitylation and degradation of p53. *Cancer Cell* 2005;8:75–87.
17. Krzywda S, Brzozowski AM, Higashitsuji H, Fujita J, Welchman R, Dawson S, Mayer RJ, Wilkinson AJ. The crystal structure of gankyrin, an oncoprotein found in complexes with cyclin-dependent kinase 4, a 19 S proteasomal ATPase regulator, and the tumor suppressors Rb and p53. *J Biol Chem* 2004;279:1541–1545.
  18. Nagao T, Higashitsuji H, Nonoguchi K, Sakurai T, Dawson S, Mayer RJ, Itoh K, Fujita J. MAGE-A4 interacts with the liver oncoprotein gankyrin and suppresses its tumorigenic activity. *J Biol Chem* 2003;278:10668–10674.
  19. Chen Y, Li HH, Fu J, Wang XF, Ren YB, Dong LW, Tang SH, Liu SQ, Wu MC, Wang HY. Oncoprotein p28 GANK binds to RelA and retains NF-kappaB in the cytoplasm through nuclear export. *Cell Res* 2007;17:1020–1029.
  20. Higashitsuji H, Higashitsuji H, Liu Y, Masuda T, Fujita T, Abdel-Aziz HI, Kongkham S, Dawson S, John Mayer R, Itoh Y, Sakurai T, Itoh K, Fujita J. The oncoprotein gankyrin interacts with RelA and suppresses NF-kappaB activity. *Biochem Biophys Res Commun* 2007;363:879–884.
  21. Ekman D, Light S, Björklund ÅK, Elofsson A. What properties characterize the hub proteins of the protein-protein interaction network of *Saccharomyces cerevisiae*? *Genome Biol* 2006;7:R45.
  22. Tsai C-J, Buyong M, Nussinov R. Protein-protein interaction networks: how can a hub protein bind so many different partners? *Trends Biochem Sci* 2009;34:594–600.
  23. Kortemme T, Baker D. A simple physical model for binding energy hot spots in protein-protein complexes. *Proc Natl Acad Sci USA* 2002;99:14116–14121.
  24. Bogan AA, Thorn KS. Anatomy of hot spots in protein interfaces. *J Mol Biol* 1998;280:1–9.
  25. Fraternali F, Cavallo L. Parameter optimized surfaces (POPS): analysis of key interactions and conformational changes in the ribosome. *Nucleic Acids Res* 2002;30:2950–2960.
  26. Rodriguez LG, Wu X, Guan JL. Wound-healing assay. *Methods Mol Biol* 2005;294:23–29.
  27. Kim SY, Hur W, Choi JE, Kim D, Wang JS, Yoon HY, Piao LS, Yoon SK. Functional characterization of human oncoprotein gankyrin in Zebrafish. *Exp Mol Med* 2009;41:8–16.
  28. Mani SA, Guo W, Liao MJ, Eaton EN, Ayyanan A, Zhou AY, Brooks M, Reinhard F, Zhang CC, Shipitsin M, Campbell LL, Polyak K, Brisken C, Yang J, Weinberg RA. The epithelial-mesenchymal transition generates cells with properties of stem cells. *Cell* 2008;133:704–715.
  29. Nakamura Y, Nakano K, Umehara T, Kimura M, Hayashizaki Y, Tanaka A, Horikoshi M, Padmanabhan B, Yokoyama S. Structure of the oncoprotein gankyrin in complex with S6 ATPase of the 26S proteasome. *Structure* 2007;15:179–189.
  30. Venkatraman P, Balakrishnan S, Rao S, Hooda Y, Pol S. A sequence and structure based method to predict putative substrates, functions and regulatory networks of endo proteases. *PLoS One* 2009;4:e5700.
  31. Wadhawan V, Kolhe YA, Sangith N, Gautam AK, Venkatraman P. From prediction to experimental validation: desmoglein 2 is a functionally relevant substrate of matriptase in epithelial cells and their reciprocal relationship is important for cell adhesion. *Biochem J* 2012;447:61–70.
  32. Singh Gautam AK, Balakrishnan S, Venkatraman P. Direct ubiquitin independent recognition and degradation of a folded protein by the eukaryotic proteasomes-origin of intrinsic degradation signals. *PLoS One* 2012;7:e34864.
  33. Jacob E, Unger R. A tale of two tails: why are terminal residues of proteins exposed? *Bioinformatics* 2007;23:e225–e230.
  34. Edwards RJ, Davey NE, Shields DC. SLiMfinder: a probabilistic method for identifying over-represented, convergently evolved, short linear motifs in proteins. *PLoS One* 2007;2:e967.
  35. Dror T, Ivet B. Recruitment of rare 3-grams at functional sites: is this a mechanism for increasing enzyme specificity? *BMC Bioinformatics* 2007;8:226.
  36. Imai J, Maruya M, Yashiroda H, Yahara I, Tanaka K. The molecular chaperone Hsp90 plays a role in the assembly and maintenance of the 26S proteasome. *EMBO J* 2003;22:3557–3567.
  37. Thomas PD, Kejariwal A, Campbell MJ, Mi H, Diemer K, Guo N, Ladunga I, Ulitsky-Lazareva B, Muruganujan A, Rabkin S, Vandergriff JA, Doremieux O. PANTHER: a browsable database of gene products organized by biological function, using curated protein family and subfamily classification. *Nucleic Acids Res* 2003;31:334–341.
  38. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell* 2011;144:646–674.
  39. Jolly C, Morimoto RI. Role of the heat shock response and molecular chaperones in oncogenesis and cell death. *J Natl Cancer Inst* 2000;92:1564–1572.
  40. Carrello A, Allan RK, Morgan SL, Owen BA, Mok D, Ward BK, Minchin RF, Toft DO, Ratajczak T. Interaction of the Hsp90 co-chaperone cyclophilin 40 with Hsc70. *Cell Stress Chaperones* 2004;9:167–181.
  41. Pang Q, Keeble W, Christianson TA, Faulkner GR, Bagby GC. FANCC interacts with Hsp70 to protect hematopoietic cells from IFN-gamma/TNF-alpha-mediated cytotoxicity. *EMBO J* 2001;20:4478–4489.
  42. Chen CD, Wang CS, Huang YH, Chien KY, Liang Y, Chen WJ, Lin KH. Overexpression of CLIC1 in human gastric carcinoma and its clinicopathological significance. *Proteomics* 2007;7:155–167.
  43. Petrova DT, Asif AR, Armstrong VW, Dimova I, Toshev S, Yaramov N, Oellerich M, Toncheva D. Expression of chloride intracellular channel protein 1 (CLIC1) and tumor protein D52 (TPD52) as potential biomarkers for colorectal cancer. *Clin Biochem* 2008;41:1224–1236.
  44. Le Guennec JY, Ouadid-Ahidouch H, Soriani O, Besson P, Ahidouch A, Vandier C. Voltage-gated ion channels, new targets in anti-cancer research. *Recent Pat Anticancer Drug Discov* 2007;2:189–202.
  45. Braverman LE, Quilliam LA. Identification of Grb4/Nckβ, a src homology 2 and 3 domain-containing adapter protein having similar binding and biological properties to Nck. *J Biol Chem* 1999;274:5542–5549.
  46. Latreille M, Larose L. Nck in a complex containing the catalytic subunit of protein phosphatase 1 regulates eukaryotic initiation factor 2α signaling and cell survival to endoplasmic reticulum stress. *J Biol Chem* 2006;281:26633–26644.
  47. Liu X, Yan S, Zhou T, Terada Y, Erikson RL. The MAP kinase pathway is required for entry into mitosis and cell survival. *Oncogene* 2004;23:763–776.
  48. Kostourou V, Troy H, Murray JF, Cullis ER, Whitley GS, Griffiths JR, Robinson SP. Overexpression of dimethylarginine dimethylaminohydrolase enhances tumor hypoxia: an insight into the relationship of hypoxia and angiogenesis in vivo. *Neoplasia* 2004;6:401–411.
  49. Chan CC, Dostie J, Diem MD, Feng W, Mann M, Rappsilber J, Dreyfuss G. eIF4A3 is a novel component of the exon junction complex. *RNA* 2004;10:200–209.
  50. Sprinzak E, Sattath S, Margalit H. How reliable are experimental protein-protein interaction data? *J Mol Biol* 2003;327:919–923.
  51. Koch WJ, Ingles J, Stone WC, Lefkowitz RJ. The binding site for the beta gamma subunits of heterotrimeric G proteins on the beta-adrenergic receptor kinase. *J Biol Chem* 1993;268:8256–8260.
  52. Ogmen U, Keskin O, Aytuna AS, Nussinov R, Gursoy A. PRISM: protein interactions by structural matching. *Nucleic Acids Res* 2005;33(Suppl 2):W331–W336.
  53. Tuncbag N, Gursoy A, Nussinov R, Keskin O. Predicting protein-protein interactions on a proteome scale by matching evolutionary and structural similarities at interfaces using PRISM. *Nat Protoc* 2011;6:1341–1354.
  54. Guney E, Tuncbag N, Keskin O, Gursoy A. HotSpring: database of computational hot spots in protein interfaces. *Nucleic Acids Res* 2008;36(Database issue):D662–D666.

55. Ma B, Elkayam T, Wolfson H, Nussinov R. Protein–protein interactions: structurally conserved residues distinguish between binding sites and exposed protein surfaces. *Proc Natl Acad Sci USA* 2003; 100:5772–5777.
56. Wang X, Wei X, Thijssen B, Das J, Lipkin SM, Yu H. Three-dimensional reconstruction of protein networks provides insight into human genetic disease. *Nat Biotechnol* 2012;30:159–164.
57. Liu Y, Higashitsuji H, Higashitsuji H, Itoh K, Sakurai T, Koike K, Hirota K, Fukumoto M, Fujita J. Overexpression of gankyrin in mouse hepatocytes induces hemangioma by suppressing factor inhibiting hypoxia-inducible factor-1 (FIH-1) and activating hypoxia-inducible factor-1. *Biochem Biophys Res Commun* 2013;432:22–27.
58. Wang P, Zhang C, Yu P, Tang B, Liu T, Cui H, Xu J. Regulation of colon cancer cell migration and invasion by CLIC1-mediated RVD. *Mol Cell Biochem* 2012;365:313–321.
59. MacAllister RJ, Parry H, Kimoto M, Ogawa T, Russell RJ, Hodson H, Whitley GS, Vallance P. Regulation of nitric oxide synthesis by dimethylarginine dimethylaminohydrolase. *Br J Pharmacol* 1996; 119:1533–1540.
60. Leiper J, Nandi M, Torondel B, Murray-Rust J, Malaki M, O’Hara B, Rossiter S, Anthony S, Madhani M, Selwood D, Smith C, Wojciak-Stothard B, Rudiger A, Stidwill R, McDonald NQ, Vallance P. Disruption of methylarginine metabolism impairs vascular homeostasis. *Nat Med* 2007;13:198–203.
61. Wang J-W, Peng S-Y, Li J-T, Wang Y, Zhang Z-P, Cheng Y, Cheng D-Q, Weng W-H, Wu X-S, Fei X-Z. Identification of metastasis-associated proteins involved in gallbladder carcinoma metastasis by proteomic analysis and functional exploration of chloride intracellular channel 1. *Cancer Lett* 2009;281:71–81.
62. Wang W, Xu X, Wang W, Shao W, Li L, Yin W, Xiu L, Mo M, Zhao J, He Q. The expression and clinical significance of CLIC1 and HSP27 in lung adenocarcinoma. *Tumor Biol* 2011;32:1199–1208.
63. Wang L, He S, Tu Y, Ji P, Zong J, Zhang J, Feng F, Zhao J, Zhang Y, Gao G. Elevated expression of chloride intracellular channel 1 is correlated with poor prognosis in human gliomas. *J Exp Clin Cancer Res* 2012;31:1–7.
64. Dengiel J, Hoyer-Hansen M, Nielsen MO, Eisenberg T, Harder LM, Schandorff S, Farkas T, Kirkegaard T, Becker AC, Schroeder S, Vanselow K, Lundberg E, Nielsen MM, Kristensen AR, Akimov V, Bunkenborg J, Madeo F, Jødtteid M, Andersen JS. Identification of autophagosom-associated proteins and regulators by quantitative proteomic analysis and genetic screens. *Mol Cell Proteomics* 2012;11:M111014035.
65. Rodi DJ, Makowski L, Kay BK. One from column A and two from column B: the benefits of phage display in molecular-recognition studies. *Curr Opin Chem Biol* 2002;6:92–96.
66. Neduva V, Russell RB. Peptides mediating interaction networks: new leads at last. *Curr Opin Biotechnol* 2006;17:465–471.
67. Ceol A, Chatr-aryamontri A, Santonico E, Sacco R, Castagnoli L, Cesareni G. DOMINO: a database of domain-peptide interactions. *Nucleic Acids Res* 2007;35(Database issue):D557–D560.
68. Kuzuoglu-Ozturk D, Huntzinger E, Schmidt S, Izaurralde E. The *Caenorhabditis elegans* GW182 protein AIN-1 interacts with PAB-1 and subunits of the PAN2-PAN3 and CCR4-NOT deadenylase complexes. *Nucleic Acids Res* 2012;40:5651–5665.
69. Landgraf C, Panni S, Montecchi-Palazzi L, Castagnoli L, Schneider-Mergener J, Volkmer-Engert R, Cesareni G. Protein interaction networks by proteome peptide scanning. *PLoS Biol* 2004;2:E14.
70. Kritzer JA, Stephens OM, Guarracino DA, Reznik SK, Schepartz A.  $\beta$ -Peptides as inhibitors of protein–protein interactions. *Bioorg Med Chem* 2005;13:11–16.
71. London N, Raveh B, Movshovitz Attias D, Schueler Furman O. Can self inhibitory peptides be derived from the interfaces of globular protein–protein interactions? *Proteins* 2010;78:3140–3149.
72. Mochly-Rosen D, Khaner H, Lopez J, Smith BL. Intracellular receptors for activated protein kinase C. Identification of a binding site for the enzyme. *J Biol Chem* 1991;266:14866–14868.
73. Mochly-Rosen D, Qvit N. Peptide inhibitors of protein-protein interactions. *Trends Endocrinol Metab* 2009;20:25–33.
74. Stebbins EG, Mochly-Rosen D. Binding specificity for RACK1 resides in the V5 region of beta II protein kinase C. *J Biol Chem* 2001;276:29644–29650.
75. Emerson SD, Palermo R, Liu CM, Tilley JW, Chen L, Danho W, Madison VS, Greeley DN, Ju G, Fry DC. NMR characterization of interleukin-2 in complexes with the IL-2Ralpha receptor component, and with low molecular weight compounds that inhibit the IL-2/IL-Ralpha interaction. *Protein Sci* 2003;12:811–822.
76. Sauve K, Nachman M, Spence C, Bailon P, Campbell E, Tsien WH, Kondas JA, Hakimi J, Ju G. Localization in human interleukin 2 of the binding site to the alpha chain (p55) of the interleukin 2 receptor. *Proc Natl Acad Sci USA* 1991;88:4636–4640.
77. Zurawski SM, Vega F, Jr., Doyle EL, Huyghe B, Flaherty K, McKay DB, Zurawski G. Definition and spatial location of mouse interleukin-2 residues that interact with its heterotrimeric receptor. *EMBO J* 1993;12:5113–5119.
78. Arkin MR, Randal M, DeLano WL, Hyde J, Luong TN, Oslob JD, Raphael DR, Taylor L, Wang J, McDowell RS, Wells JA, Braisted AC. Binding of small molecules to an adaptive protein-protein interface. *Proc Natl Acad Sci USA* 2003;100:1603–1608.
79. Arkin MR, Wells JA. Small-molecule inhibitors of protein-protein interactions: progressing towards the dream. *Nat Rev Drug Discov* 2004;3:301–317.
80. Clackson T, Wells JA. A hot spot of binding energy in a hormone-receptor interface. *Science* 1995;267:383–386.
81. Erbe DV, Wang S, Xing Y, Tobin JF. Small molecule ligands define a binding site on the immune regulatory protein B7.1. *J Biol Chem* 2002;277:7363–7368.
82. Green NJ, Xiang J, Chen J, Chen L, Davies AM, Erbe D, Tam S, Tobin JF. Structure-activity studies of a series of dipyrzolo[3,4-b:3’,4’-d]pyridin-3-ones binding to the immune regulatory protein B7.1. *Bioorg Med Chem* 2003;11:2991–3013.
83. Sattler M, Liang H, Nettlesheim D, Meadows RP, Harlan JE, Eberstadt M, Yoon HS, Shuker SB, Chang BS, Minn AJ, Thompson CB, Fesik SW. Structure of Bcl-xL-Bak peptide complex: recognition between regulators of apoptosis. *Science* 1997;275:983–986.
84. Prasanna V, Bhattacharjya S, Balaran P. Synthetic interface peptides as inactivators of multimeric enzymes: inhibitory and conformational properties of three fragments from *Lactobacillus casei* thymidylate synthase. *Biochemistry* 1998;37:6883–6893.
85. Cardinale D, Guaitoli G, Tondi D, Luciani R, Henrich S, Salo-Ahen OM, Ferrari S, Marverti G, Guerrieri D, Ligabue A, Frassinetti C, Pozzi C, Mangani S, Fessas D, Guerrini R, Ponterini G, Wade RC, Costi MP. Protein-protein interface-binding peptides inhibit the cancer therapy target human thymidylate synthase. *Proc Natl Acad Sci USA* 2011;108:E542–E549.