UNDER PEER REVIEW

Research paper

Importance of N-terminal glycine-arginine rich (GAR) domain in *Giardia lamblia* fibrillarin in formation of snoRNA-fibrillarin complex

Sumallya Karmakar, Dibyendu Raj and Sandipan Ganguly*

National Institute of Cholera and Enteric Diseases, Indian Council of Medical Research, P33 CIT Road, Scheme – XM, Beliaghata, Kolkata – 700010.

ABSTRACT

Aims: Study the role of GAR domain of fibrillarin in *Giardia lamblia*.
Study design: Identifying a specific glycine arginine rich (GAR) sequence of *Giardia* fibrillarin which plays an important role in ribonucleoprotein particles complex formation with snoRNA during post transcriptional modifications of rRNA. The present study uses Electrophoretic Mobiliy Shift Assay (EMSA), to detect protein-RNA interactions through changes in migration speed during gel electrophoresis. ^{32}P labeled snoRNA incubated with purified fibrillarin and GAR domain truncated fibrillarin protein were separated by a non denaturing polyacrylamide gel where the band patterns suggested the interaction of the RNAs with proteins.
Place and Duration of study: Department of Parasitology, National Institute of Cholera and Enteric Diseases, Indian Council of Medical Research, Kolkata, between January 2011 and July 2012.
Methodology: Homology analysis of *G. lamblia* fibrillarin was performed with fibrillarin from *Homo sapiens*, *Mus musculus* and *Arabidopsis thaliana* to determine the conserved domain by using different softwares. Cloning, expression and purification of fibrillarin and GAR domain truncated fibrillarin were done to study the role of GAR domain in snoRNA-fibrillarin binding by EMSA and Gradient EMSA. Immunoblot assay of purified GAR domain truncated fibrillarin was performed and obtained a distinct band with polyclonal antibody against full length purified fibrillarin.
Results: Gel retardation assay showed that snoRNA does not bind with GAR domain truncated fibrillarin to form any RNP complex. Similarly in gradient EMSA the GAR domain truncated fibrillarin does not bind with any of the *in vitro* transcribed snoRNA even at much higher concentration than they bind in case of full length purified recombinant fibrillarin.
Conclusion: The amino terminal conserved glycine arginine rich (GAR) domain is present in *Giardia lamblia* fibrillarin and is involved in formation of RNPP complex with snoRNA of this organism.

Keywords: Giardia lamblia, small nucleolar RNA, fibrillarin, RNA-Protein complex, glaycine-arginine rich domain.

1. INTRODUCTION

Giardia lamblia (also known as *Giardia intestinalis* and *Giardia duodenalis*) is an amitochondrial, unicellular, gastrointestinal flagellated protozoan parasite causing one of the most common parasitic infections worldwide (*Adam R. D.*, 2001). It imparts to an

UNDER PEER REVIEW

estimated 300 million symptomatic human infections (called Giardiasis) per year and has been included as part of WHO Neglected Disease Initiative since 2004.

The diplomonad protist *Giardia lamblia* is a prevalent human enteric parasite that displays a highly reduced compact genome and somewhat limited metabolic capacity (Morrison H.G. et al. 2007). Eukaryotic precursor (pre-) RNA processing often requires ribonucleoprotein (RNP) complexes consisting of conserved and essential non-coding (nc) RNAs. Notable examples are the small nucleolar (sno) RNPs that participate in eukaryotic ribosome biogenesis through structural modification of specific nucleotides in ribosomal RNA (rRNA) and/or targeting cleavage of the pre-rRNA [reviewed in (Henras A. K. et. al. 2008). Although it was previously considered that *Giardia* lacks a nucleolus, recent studies have reported the organism to be nucleolate (Tian et al, 2010).

It has been demonstrated recently that numerous small nucleolar RNAs found in the nucleolus of eukaryotic cells, can be sorted into two classes (Balakin *et al*. 1996). Both of them are involved in pre rRNA processing. Members of one class containing structural motifs called box C/D are involved in 2'-O-ribose methylation of pre rRNA. The other class of snoRNAs are characterized by the conserved motifs called box H/ACA which converts the Uridine (U) residue of the pre rRNA into pseudouridine (ψ) (Maden et al, 1997; Bachellerie, et al, 2000). Recently, 20 snoRNA-like RNA has been identified in *Giardia* out of which 16 belong to C/D box class and the rest are H/ACA type (Cheng-Yong Yang et al. 2005). Fibrillarin is a methyl transferase (Reichow et al. 2007) that is required for pre rRNA processing. Association of C/D box snoRNA with fibrillarin has been reported previously (Tycowski et al., 1989 and Schimmang, et al. 1989) and in *Giardia*, the protein associates with RNA D, RNA J, RNA H, both *invitro* and *invivo* (Ghosh et al, 2001). Deletion analysis of fibrillarin in other organisms indicated that a major RNA binding element, which is extremely well conserved throughout evolution, lies in the amino terminal end of the protein (Bagni *et al*. 1998). The sequence is highly conserved between different species and confirms the previously observed similarity between different mammalian nucleolar proteins. Interestingly, this repetitive amino acid sequence motif begins near the amino terminus of fibrillarin. The motif is approximately 80 residues long and rich in glycine and arginine (GAR domain) (Henriquez *et al*. 1990). A model of box C/D snoRNP assembly proposes that this conserved GAR domain is necessary for fibrillarin to bind with snoRNA. The domain carries the pre-rRNA base-pairing interactions in the functional center of the U3 snoRNP (Wormsley *et al*. 2001).

It has been proposed that specific interaction of snoRNA with this amino terminal domain of fibrillarin in the nucleolus occurs for rRNA precursor processing (Yoo *et al.* 2003). Here in this study we describe the identification of the GAR domain of fibrillarin of *Giardia lamblia* and the essential role it plays for snoRNPP formation. This is the first demonstration of the importance of GAR domain in *Giardia* snoRNA binding for rRNA processing complex formation.

Understanding of the cell biology of this very complex and highly developed cell will greatly increase our knowledge about this interesting and enigmatic organism as well as promise to yield novel insight into the evolution of key pathways and structures of higher cells. Thus, what commenced as an interest in the molecular biology of medically important protozoa has developed into a fascinating expedition through the unusual eukaryote, *Giardia*.

2. MATERIAL AND METHODS

2.1 Homology analysis to find out a conserved N terminal Glycine Arginine rich domain of fibrillarin in *G. lamblia*

Homology analysis of *G. lamblia* fibrillarin was performed with fibrillarin from *Homo sapiens*, *Mus musculus* and *Arabidopsis thaliana* to determine the conserved domain. The analysis was performed using different softwares namely DNAsis and Omiga.

UNDER PEER REVIEW

2.2 Cloning of fibrillarin (pET33b-Fb) and N terminal GAR domain truncated fibrillarin (pET33b-δGFb)

The fibrillarin gene (pET33b-Fb) was cloned in pET 33b vector as described previously (Ghosh et al., 2001). For the cloning of pET33b-δGFb two primers were designed, with restriction sites engineered into them that precisely excluded the GAR domain of the protein (Table 1). The desired fragment was PCR amplified from the cDNA of *Giardia* and was used for cloning in pET 33b vector. Briefly, the PCR products and the plasmid were digested with Bam HI and Nde I restriction enzymes, purified and then ligated with T4 DNA ligase at 16^0C overnight. After the reaction was over, the ligase was heat inactivated by placing the tube at 65^0C water bath for 10 mins. The ligation mixture was transformed into competent *E.coli* XL1 Blue cells. Ligated positive clones of GAR domain truncated fibrillarin were selected randomly from LB plates containing 50µg/ml Kanamycin. The plasmid isolated from each selected clone was double digested by Bam HI and NdeI restriction enzymes, separated on a 1% agarose gel and photographed.

2.3 Overexpression and Purification of recombinant fibrillarin and GAR domain truncated fibrillarin protein using metal chelating resin column

Recombinant fibrillarin gene was expressed in *E.coli* BL21 (DE3) cells and purified as explained previously (Ghosh et al., 2001). For bacterial expression of GAR domain truncated fibrillarin gene, pET33b-δGFb was transformed to *E.coli* BL21 (DE3) cells. These cells were grown in LB supplemented with 50 µg/ml kanamycin and 0.3% glucose and incubated with shaking at 25^0C until OD_{600} reached 0.6, followed by induction by addition of IPTG (Sigma) to a final concentration of 1.0 mM. The cells were induced for 5 hr and harvested by centrifugation at 12,000g for 10 mins at 4^0C. The pellet was washed with normal saline and suspended in 2 ml of ice-cold binding buffer (20mM Tris-HCl, pH 7.9, 5mM imidazole and 0.5 M NaCl).

The bacterial pellet was dissolved in 1 ml lysis buffer (10X lysis buffer - 0.5 M Tris-HCl (pH 8.0), 10 mM EDTA, 1M NaCl) and sonicated in an ultrasound sonicator by 10-15 30-seconds bursts at 20 kilocycles at an interval of 1 min at 4^oC. Sonicated material was centrifuged at 12000 g for 10 min at 4^oC to remove the debris and the supernatant was used for purification. Purification was done by His tag column (HiTrap Chelating®, Amersham Pharmacia) which was equilibrated with $NiSO_4$ and binding buffer as per manufacturer's protocol.

The crude soluble GAR domain truncated fibrillarin protein was loaded onto this pre equilibrated His Tag column for binding of the target protein to Ni^{+2} ion immobilized on the His binding metal chelating resin. The unbound proteins were washed with 10 column volumes of binding buffer with 5mM imidazole and then with 6-column volume of wash buffer containing 20mM imidazole. The target protein was eluted with 6 volumes of elution buffer (20mM Tris-HCl (pH7.9) 250mM imidazole and 0.5M NaCl). Subsequently the imidazole and NaCl was dialyzed out against binding buffer using centricon (poresize 30000 daltons). The purified protein fractions were analyzed by SDS-PAGE.

2.4 Western Blot Analysis

Purified GAR domain truncated fibrillarin and *E. coli* CSA were separated by 10% SDS PAGE. Polypeptides from this SDS PAGE were transferred onto nitrocellulose membranes (pore size 0.45 µm, Bio-Rad, USA) using a Trans-blot apparatus (Bio-Rad, USA) in presence of transfer buffer (1.162gm Tris-HCl, 0. 0585gm glycine 0. 075gm SDS in 160ml of water and 40ml of methanol) at a constant voltage of 100 volts for 2 hours at 4°C.

After the protein was transferred the nitrocellulose strips were washed thoroughly in Tris buffered saline (25mM Tris-HCl (pH 7.4), 150mM NaCl) and incubated with blocking buffer (3% BSA in TBS) at 4°C overnight.

The nitrocellulose strips were washed with TBST (TBS + 0.05% Tween 20) and incubated with 1: 50 dilution of anti fibrillarin rabbit sera and allowed to react at 37°C for 90 minutes on a rotatest shaker.

UNDER PEER REVIEW

The strips were washed again (as above) and incubated with 1:5000 [diluted in TBST containing 1.5% BSA (w/v)] dilution of anti-rabbit IgG conjugated with horseradish peroxidase (Jackson Immunoresearch Lab, USA). The reaction was allowed to proceed at room temperature for 45 minutes.

The washed nitrocellulose strips were developed in 0.05% solution of 3,3'-diaminobenzidine tetrahydrochloride (DAB, Sigma, USA) and 0.03% hydrogen peroxide in TBS. The reaction was allowed to proceed at dark for 4 to 5 minutes at room temperature and was stopped by repeated washing of the strips in distilled water. The coloured immunoreactive polypeptide patterns were observed visually.

2.4 Role of GAR domain in snoRNA-fibrillarin binding

2.4.1 Electrophoresis mobility shift assay (EMSA)

To study the role of GAR domain for interaction with snoRNA, EMSA was performed with GAR domain truncated fibrillarin and snoRNA of *Giardia lamblia*. 40,000cpm ^{32}P labeled snoRNA and 2µg purified GAR domain truncated fibrillarin protein were mixed with binding buffer (10mM HEPES-KOH pH 7.9, 10% glycerol, 0.05% NP40, 1mM EDTA pH 8.0, 0.5mM DTT, 100mM KCl and 5mM $MgCl_2$). The total reaction volume was 20 µl and the reaction mixture was incubated at 0°C for 30 min. After incubation samples were electrophoresed on a native 5% polyacrylamide gel (29:1) in TAE buffer at 100 volts at 4°C. Different control reactions were also performed using full length recombinant fibrillarin protein, BSA, labeled RNAs both specific (different snoRNAs of *Giardia*) and non-specific (a 82 nucleotide leader RNA of Chandipura virus, a negative strand RNA virus).

2.4.2 Gradiant Electrophoresis Mobility Shift Assay (GEMSA) to show the loss of affinity of binding at higher concentrations

20µL binding reaction contained binding buffer (10mM HEPES-KOH pH 7.9, 10% glycerol, 0.05% NP40, 1mM EDTA pH 8.0, 0.5mM DTT, 100mM KCl and 5mM $MgCl_2$), 40,000cpm ^{32}P labeled snoRNA and different amount (i.e. 10-100ng) of purified GAR domain truncated fibrillarin protein. The reaction mixes were incubated at 0°C for 30 min. After incubation samples were electrophoresed on a 5% native polyacrylamide gel (29:1) in TAE buffer at 100 volts and 4°C. Positive control reactions using same concentrations of full-length recombinant fibrillarin protein and labled snoRNA were used.

3. RESULTS

3.1 Homology analysis of fibrillarin in *G. lamblia*

Homology analysis of *G. lamblia* fibrillarin sequence has revealed that N terminal Glycine Arginine Rich (GAR) domain of *Giardia lamblia* fibrillarin is largely homologus and conserved with fibrillarin from different sources, *viz. Homo sapiens, Mus masculus,* and *Arabidopsis thaliana* (Fig. 1).

3.2 Overexpression and Purification of recombinant GAR domain truncated fibrillarin protein

Overexpression of GAR domain truncated fibrillarin was standardized in LB containing 0.3% glucose and 50µg/ml kanamycin was the optimal growth medium and it was used for overexpression of δGFb. 0.6 OD_{600} of bacterial cell suspension to add IPTG for starting induction was found to be optimum for the yield as well as lower background of host proteins. IPTG concentration of 1.0mM was optimum for good production of GAR domain truncated fibrillarin protein. Higher IPTG concentration did not increase the production of δGFb protein. Optimum time was taken as 5 hours and 25^{0}C was chosen as the optimum temperature for induction.

UNDER PEER REVIEW

Purified fraction of recombinant GAR domain truncated fibrillarin and wash outs were run in 10% SDS PAGE which reveals a distinct band of purified GAR domain truncated fibrillarin (Fig. 2). The purified fraction comes out at 250mM Imidazole concentration.

3.3 Western blot analysis

Immunoblot assay of purified GAR domain truncated fibrillarin has produced a distinct band with polyclonal antibody against full length purified fibrillarin (fig. 3). This indicates that the 3 dimensional structure of the other domains of GAR truncated fibrillarin remains equally active as in full length fibrillarin. Thus, study with this protein will definitely explain the role of GAR domain in binding with snoRNAs as all other domains are active and could bind with the snoRNAs (if they possibly binds).

3.4 EMSA

Gel retardation assay showed that none of the three snRNA binds with GAR domain truncated fibrillarin to form any complex (fig. 4). However, full length purified recombinant fibrillarin that has been used as a positive control has showed clear shifting of band. Similarly in gradient GEMSA the GAR domain truncated fibrillarin does not bind with any of the *in vitro* transcribed snRNA even at much higher concentration than they bind in case of full length purified recombinant fibrillarin (fig. 5).

4. DISCUSSION

Fibrillarin (~35 kDa) comprises three structural domains, and is conserved among different organisms and species. Homology analysis of conserved domain in fibrillarin has suggested the importance of N terminal GAR domain. Fibrillarin have been identified in *Giardia* and it has been shown that it interacts with snRNA *in vitro* and *in vivo* for the formation of RNPP complex, but, how this interaction takes place and what are the binding domains for this complex formation has not yet been identified. Deletion analysis has revealed that the GAR domain contributes to the binding of snoRNA. In the EMSA study, absence of any proper shift clearly suggests the important role of this GAR domain in RNPP formation and maturation. This may be either due to the RNA binding role of the GAR domain or may be due to change in the activity of RNA binding domain due to absence of GAR domain.

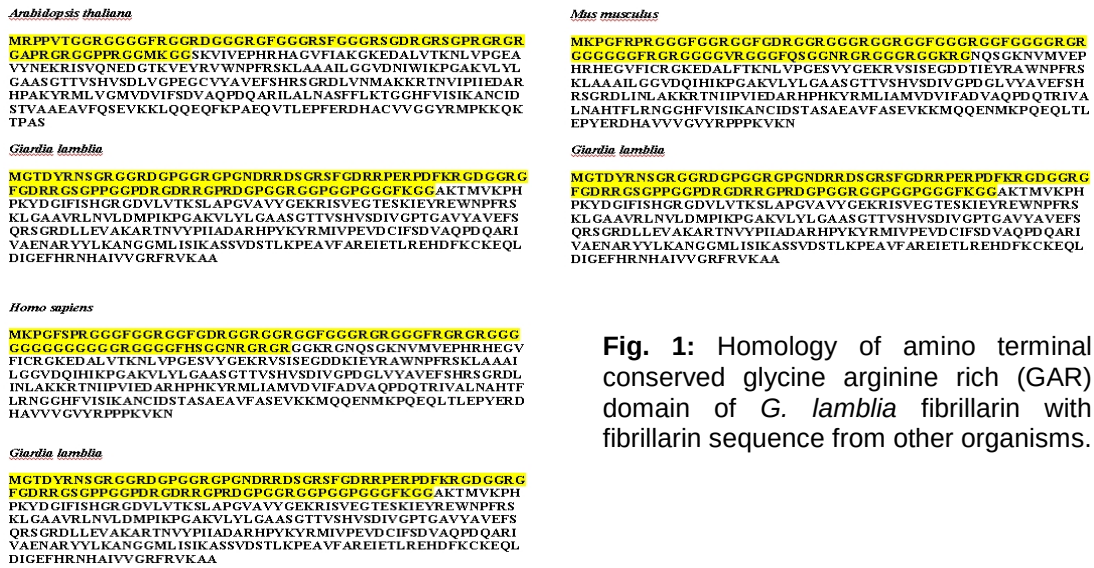

Fig. 1: Homology of amino terminal conserved glycine arginine rich (GAR) domain of *G. lamblia* fibrillarin with fibrillarin sequence from other organisms.

UNDER PEER REVIEW

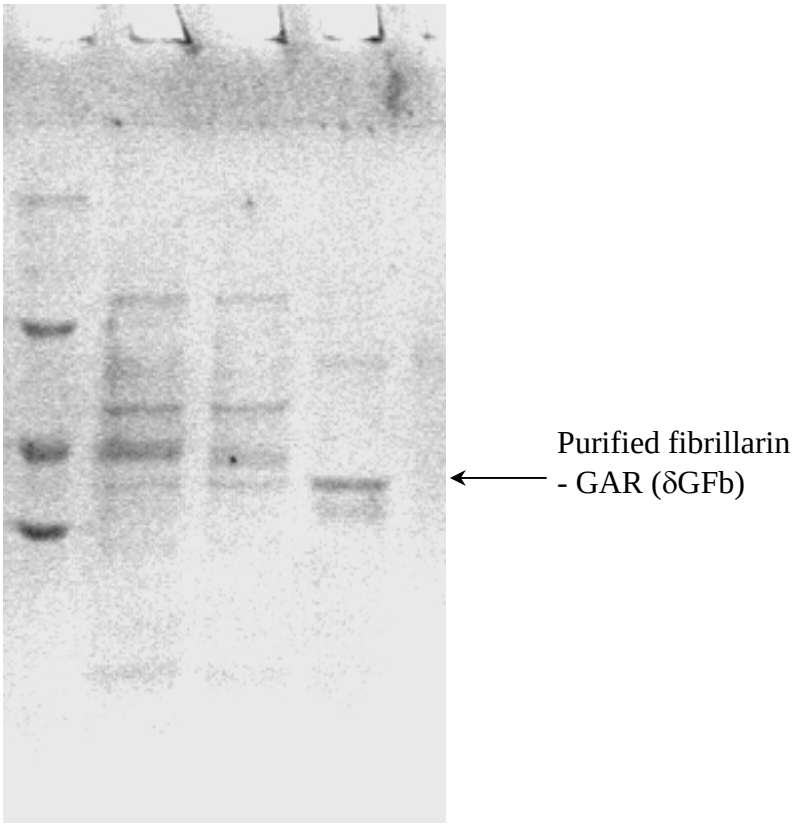

Fig. 2: Purification of GAR domain truncated fibrillarin : Lane 1 : protein molecular weight marker, Lane 2 & 3: fractions containing wash solutions. Lane 4: fraction containing purified GAR domain truncated fibrillarin

UNDER PEER REVIEW

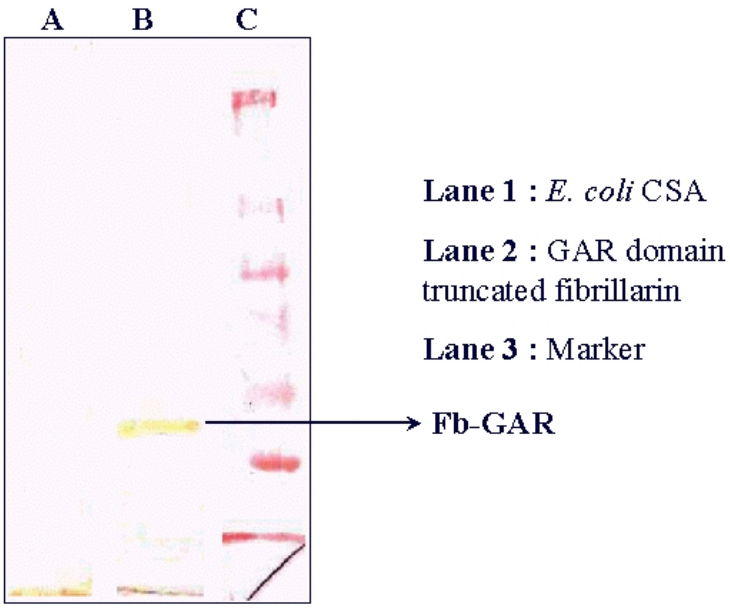

Fig. 3: Western blot assay with purified GAR domain truncated fibrillarin and antibody against full length recombinant purified fibrillarin. *E. coli* CSA was used as negative control.

UNDER PEER REVIEW

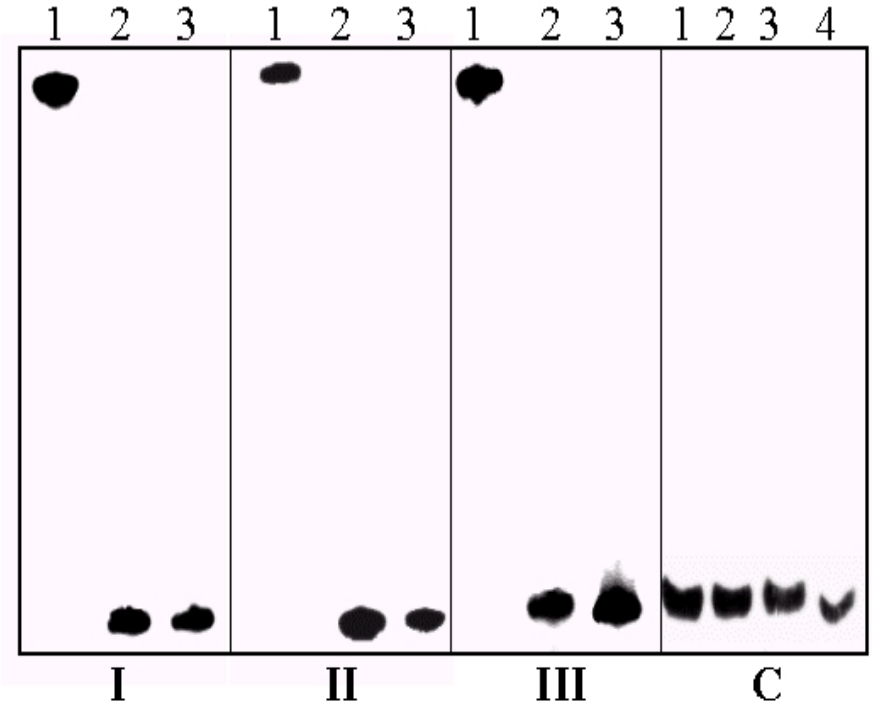

Fig. 4: Gel shift assay. for I, II and III Lane 1 : radiolabelled snRNA and fibrillarin, Lane 2 : radiolabelled snRNA. Lane 3 : radiolabelled snRNA and GAR truncated fibrillarin. For C : Lane 1 : radiolabelled non-specific RNA and fibrillarin, Lane 2 : Radiolabelled non-specific RNA. Lane 3 : radiolabelled non specific RNA and GAR truncated fibrillarin. Lane 4 : radiolabelled snRNA H and BSA. I : snRNA D, II : snRNA J and III : snRNA H.

UNDER PEER REVIEW

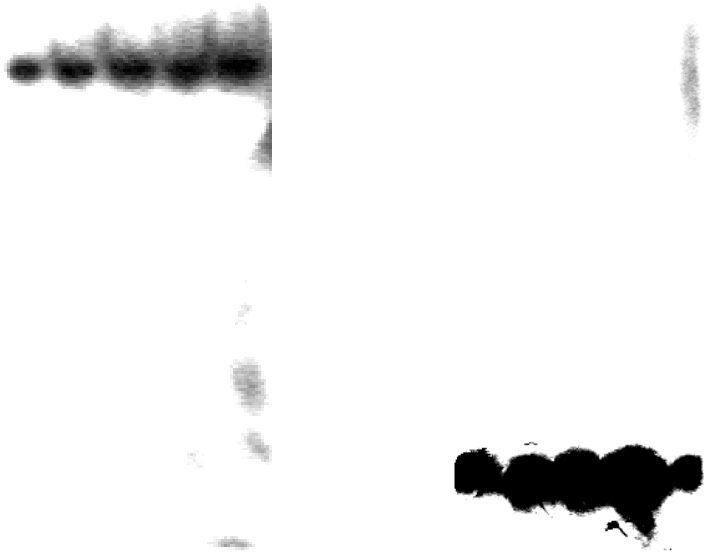

Fig. 5: Gradient GEMSA with full length fibrillarin (left panel) and GAR domain truncated fibrillarin (right panel). All the lanes containing snRNA D and variable amount of purified recombinant full length or truncated fibrillarin. Lane 1: 10µg Fb, Lane 2: 20µg Fb, Lane 3: 40µg Fb, Lane 4: 80µg Fb and Lane 5: 100µg Fb. In all the cases full length fibrillarin binds with snRNA D but

Table 1: Primers for PCR amplification of GAR domain truncated fibrillarin gene of *Giardia lamblia.*

Name	Sequence	Restriction Enzyme Site
GAR(TFb)_F	5'-C**GGATCC**GAUGCGGACCAAGGG-3'	BamHI
GAR(TFb)__R	5'-GGAATTC**CATATG**GAATTCCTCACGCTGCCTTG-3'	NdeI

5. CONCLUSION

Our findings clearly demonstrate how interaction between the core snRNP protein fibrillarin and the Box C/D snRNAs in *G. lamblia* takes place. This finding is of potential importance to our understanding of snoRNP biogenesis as well as in the etiology of rRNA maturation in *Giardia*.

UNDER PEER REVIEW

ACKNOWLEDGEMENTS

This study was supported by Indian Council of Medical Research, Govt. of India. The authors also thank University Grant Commission, India and Council of Scientific and Industrial Research, India for providing fellowship to DR and SK respectively during the study period and Mrs. Debarati Ganguly for her help in English language correction.

COMPETING INTERESTS

No competing interest exists.

AUTHORS' CONTRIBUTIONS

Author Sandipan Ganguly designed the study. Authors Sumallya karmakar and Dibyendu Raj carried out the necessary bench works and wrote the paper.

REFERENCES

Adam R. D. Biology of *Giardia lamblia.* Clin Microbiol Rev*. 2001; 14(3):447-75*

Bachellerie, J. P. et al. snoRNAs: a family of nucleolar RNAs with long complementarities to rRNA. *TIBS*. 1995. 20: 261-264.

Bagni C, Lapeyre B. Gar1p binds to the small nucleolar RNAs snR10 and snR30 in vitro through a nontypical RNA binding element. *J Biol Chem*. 1998; 273(18): 10868-73.

Balakin, A. G. ; Smith, L.; and Fournier, M. J. The RNA world of the nucleolus: two major families of small RNAs defined by different box elements with related functions. *Cell.* 1996. 86: 823-834.

Cheng-Yong Yang et al. Identification of 20 snoRNA-like RNAs from the primitive eukaryote, *Giardia lamblia*. BBRC, 2005. 328: 1224-1231.

Ghosh, et. al. Expression and purification of recombinant *Giardia* fibrillarin and its interaction with small nuclear RNAs. 2001, 21: 40-48.

Henras A. K. et. al. The post-transcriptional steps of eukaryotic ribosome biogenesis. Cell. Mol. Life Sci. 2008;65:2334–2359.

Henriquez, R.; Blobel, G. and Aris, J. P. Isolation and sequencing of NOPI. A yeast gene encoding a nucleolar protein homologues to a human autoimmune antigen. *J. Biol. Chem*. 1990. 265: 2209-2215.

Morrison H.G. et al. Genomic minimalism in the early diverging intestinal parasite *Giardia lamblia*.Science. 2007;317:1921–1926.

Reichow, S. L. et al. The structure and function of small nucleolar ribonucleoproteins. Nucleic Acid Research, 2007. 35(5): 1452-1464.

Schimmang, T. et. al. A yeast nucleolar protein related to mammalian fibrillarin is associated with small nucleolar RNA and is essential for viability. EMBO J. 8, 4015-4024.

Tian et. al. Identification of the nucleoli of Giardia lamblia with TEM and CFM. Parasitology Research. 2010; 106(4): 789-93.

Tycowoski, K., and Steitz, J. A. U3, U8 and U13 comprise a new class of mammalian snRNPs localised in the cell nucleolus. EMBO J. 8, 3113-3119.

Wormsley S, Samarsky DA, Fournier MJ, Baserga SJ. An unexpected, conserved element of the U3 snoRNA is required for Mpp10p association. *RNA*. 2001;7(6):904-19.

Yoo D, Wootton SK, Li G, Song C, Rowland RR. Colocalization and interaction of the porcine arterivirus nucleocapsid protein with the small nucleolar RNA-associated protein fibrillarin. *J Virol*. 2003;77(22):12173-83

UNDER PEER REVIEW