

1 Identification of the N-terminal glycine-arginine
2 rich (GAR) domain in *Giardia lamblia* fibrillarin
3 and evidence ~~for~~of its essentiality for ~~the~~binding
4 ~~with snoRNA of this protein~~snoRNA binding

Formatted: Left

5
6 Sumallya Karmakar⁺, Dibyendu Raj⁺ and Sandipan Ganguly^{*}

7
8 National Institute of Cholera and Enteric Diseases, Indian Council of Medical
9 Research, P33 CIT Road, Scheme – XM, Beliaghata, Kolkata – 700010.

10
11 + Authors have equal contribution

12
13 ABSTRACT

14
15
16
17
18
19
20
21
22
23
Aims: Study the role of glycine-arginine rich (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 Mobility Shift Assay (EMSA) to detect protein-RNA interactions. ³²P labeled snoRNA incubated with purified fibrillarin and/or GAR domain truncated fibrillarin protein were separated by a non denaturing polyacrylamide gel where the band patterns suggested the interaction of the RNAs with both the 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 fibrillarins from *Homo sapiens*, *Mus musculus* and *Arabidopsis thaliana* to determine the conserved domain by ClustalW analysis, using Seaview Graphical representation version 4. 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 by using polyclonal antibody of purified recombinant giardial fibrillarin protein.
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 do for full length purified recombinant fibrillarin.
Conclusion: The amino terminal conserved glycine arginine rich (GAR) domain is present in *Giardia lamblia* fibrillarin and is essential in for binding with snoRNA of this protein.

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

20 1. INTRODUCTION

21 *Giardia lamblia* (also known as *Giardia intestinalis* and *Giardia duodenalis*) is an
22 amitochondrial, unicellular, gastrointestinal flagellated protozoan parasite causing one of
23 the most common parasitic infections worldwide (1). It imparts to an 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 (2). 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 (3). Although it was previously considered that *Giardia* lacks a nucleolus, recent studies have reported the organism to be nucleolate (4).

It has been demonstrated recently that numerous small nucleolar RNAs (snoRNA) found in the nucleolus of eukaryotic cells, can be sorted into two classes (5). 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) (6, 7). Recently, 20 snoRNA-like RNAs have been identified in *Giardia* out of which 16 belong to C/D box class and the rest are H/ACA type (8). Fibrillarin (Fb) is a methyl transferase (9) that is required for pre rRNA processing. Association of C/D box snoRNA with fibrillarin has been reported previously (10,11) and in *Giardia*, the protein associates with RNA D, RNA J, RNA H, both *in vitro* and *in vivo* (12). 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 (13). The sequence is highly conserved between different species and confirms the previously observed similarity between different mammalian nucleolar proteins. The motif is approximately 80 residues long and rich in glycine and arginine (GAR domain) (14). A model of box C/D snoRNP assembly proposes that this conserved GAR domain is necessary for fibrillarin to bind snoRNA. The domain carries the pre-rRNA base-pairing interactions in the functional center of the U3 snoRNP (15).

It has been proposed that specific interaction of snoRNA with this amino terminal domain of fibrillarin in the nucleolus occurs for pre-rRNA processing (16). Here in this study we describe the identification of the GAR domain of *Giardia lamblia* fibrillarin protein and the essential role it plays for binding snoRNAs. 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 *G.lamblia* will greatly increase our knowledge about this 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 determine a conserved N terminal Glycine Arginine rich domain of fibrillarin in *G. lamblia*

Homology analysis of *G. lamblia* fibrillarin was performed with amino acid sequence of fibrillarins from *Homo sapiens*, *Mus musculus* and *Arabidopsis thaliana* to determine the conserved domain. The analysis was performed by ClustalW using SeaView Graphical representation version 4.

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

The *Giardia* fibrillarin gene was cloned in pET 33b vector as described previously (12). For the cloning of pET33b- δ GFb 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~~ *Giardia* cDNA and was cloned into 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 using T4 DNA ligase at 16°C overnight. After the reaction was over, the enzyme was inactivated at 65°C water-bath for 10 mins. The ligation mixture was transformed into competent *E.coli* XL1 Blue cells. Ligated pE positive clones of GAR domain truncated fibrillarin were selected randomly from LB plates containing 50 µg/ml Kanamycin. The plasmid isolated from each selected clone was double digested by BamHI and NdeI restriction enzymes, separated on a 1% agarose gel and photographed.

2.3 Recombinant proteins expression and purification

Recombinant fibrillarin gene was expressed in *E.coli* BL21 (DE3) cells and purified as explained previously (12). 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°C until OD₆₀₀ reached 0.6, followed by 1 mM IPTG (Sigma) 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°C. The pellet was washed with PBS and resuspended in 2 ml of ice-cold binding buffer (20mM Tris-HCl, pH 7.9, 5mM imidazole and 0.5 M NaCl) then it 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°C. Sonicated material was centrifuged at 12000 g for 10 min at 4°C to remove the debris and the supernatant was used for purification. Purification was done using His tag column (HiTrap Chelating®, Amersham Pharmacia) which was equilibrated with NiSO₄ 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. 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 (pore size 30000 daltons). The purified protein fractions were analyzed by SDS-PAGE.

Comment [u1]: Why NaCl dialysis if both concentrations are the same 0,5M?

2.4 Western Blot Analysis

Purified GAR domain truncated fibrillarin and *E. coli* crude soluble antigen (CSA) (prepared in the laboratory) 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 and 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 anti-Giardia fibrillarin (developed in rabbit) 1:50 dilution of giardial fibrillarin antibody raised in rabbit and allowed to react at 37°C for 90 minutes on a rotatest shaker.

The strips were washed again (as above) and incubated with anti-rabbit IgG-HRP (Horseradish peroxidase) (Jackson Immunoresearch Lab, USA), 1:5000 dilution, [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.

Formatted: Indent: First line: 0"

Formatted: Font: 10 pt

Formatted: Font: 10 pt

Formatted: Font: 10 pt

Formatted: Font: 10 pt

130 The washed nitrocellulose strips were developed in 0.05% solution of 3,3'-
131 diaminobenzidine tetrahydrochloride (DAB, Sigma, USA) and 0.03% hydrogen peroxide in
132 TBS. The reaction was allowed to proceed at in darkness for 4 to 5 minutes at room
133 temperature and was stopped by repeated washing of the strips in distilled water. The
134 coloured immunoreactive polypeptide patterns were observed
135

136 **2.4 Role of GAR domain in snoRNA fibrillarin binding**

137 **2.4.1 Electrophoresis mobility shift assay (EMSA)**

Formatted: Font: 11 pt

138 To study the interaction between the GAR domain and snoRNA of *Giardia lamblia*, a EMSA
139 was performed with GAR truncated fibrillarin and snoRNAs. 40,000cpm ³²P labeled
140 snoRNA and 2µg purified GAR domain truncated fibrillarin protein were mixed with binding
141 buffer (10mM HEPES-KOH pH 7.9, 10% glycerol, 0.05% NP40, 1mM EDTA pH 8.0, 0.5mM
142 DTT, 100mM KCl and 5mM MgCl₂). The total reaction volume was 20 µl and the reaction
143 mixture was incubated at 0°C for 30 min. After incubation samples were electrophoresed on
144 a native 5% polyacrylamide gel (29:1) in TAE buffer at 100 volts at 4°C. Different control
145 reactions were also performed using full length recombinant fibrillarin protein, Bovine Serum
146 Albumin (BSA), labeled RNAs both specific (RNA J and RNA H of *Giardia*) and non-specific
147 (a 82 nucleotide leader RNA of Chandipura virus, a negative strand RNA virus).
148

149 **2.4.26 Gradient Electrophoresis Mobility Shift Assay (GEMSA)**

Formatted: Font: 11 pt

150
151 20µL binding reaction containing contains binding buffer (10mM HEPES-KOH pH 7.9, 10%
152 glycerol, 0.05% NP40, 1mM EDTA pH 8.0, 0.5mM DTT, 100mM KCl and 5mM MgCl₂),
153 40,000cpm ³²P labeled snoRNA (RNA D) and different amount (i.e. 10-100ng) of purified
154 GAR domain truncated fibrillarin protein. The reaction mixes were incubated at 0°C for 30
155 min. After incubation samples were electrophoresed on a 5% native polyacrylamide gel
156 (29:1) in TAE buffer at 100 volts and 4°C. Positive control reactions using same
157 concentrations of full-length recombinant fibrillarin protein and labled snoRNA were used.
158

159 **3. RESULTS**

160 **3.1 Homology analysis of fibrillarin in *G. lamblia***

161 Homology analysis of *G. lamblia* fibrillarin sequence has revealed that N terminal Glycine
162 Arginine Rich (GAR) domain of *Giardia lamblia* fibrillarin is largely homologous and
163 conserved with fibrillarin from different sources, viz. *Homo sapiens*, *Mus masculus*, and
164 *Arabidopsis thaliana* (Fig. 1).
165

166 **3.2 Expression and Purification of recombinant GAR domain truncated** 167 **fibrillarin protein**

168
169 Purified fraction of recombinant GAR domain truncated fibrillarin and wash outs were run in
170 10% SDS PAGE which reveals a distinct band of purified GAR domain truncated fibrillarin
171 after staining with Coomassie blue (Fig. 2).
172

172 **3.3 Western blot analysis**

173 Immunoblot assay of purified GAR domain truncated fibrillarin has produced a distinct band
174 with polyclonal antibody against full length purified fibrillarin (fig. 3). This indicates that the 3
175 dimensional structure of the other domains of GAR truncated fibrillarin remains equally
176 active as in full length fibrillarin. Thus, study with this protein will definitely explain the role of
177 GAR domain in binding with snoRNAs as all other domains are active and could bind with
178 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 has been identified in *Giardia* and it has been shown that it interacts with snRNA *in vitro* and *in vivo* in 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.

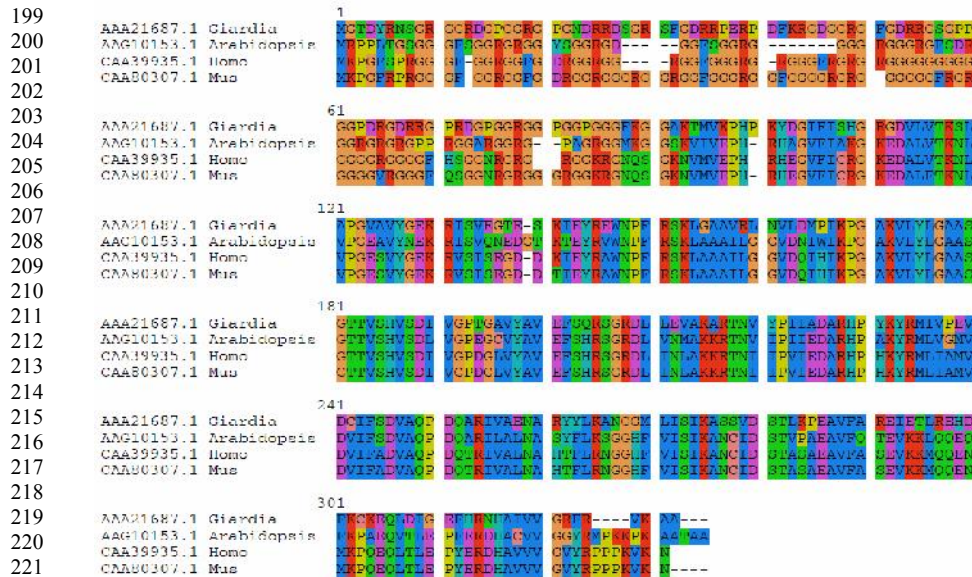
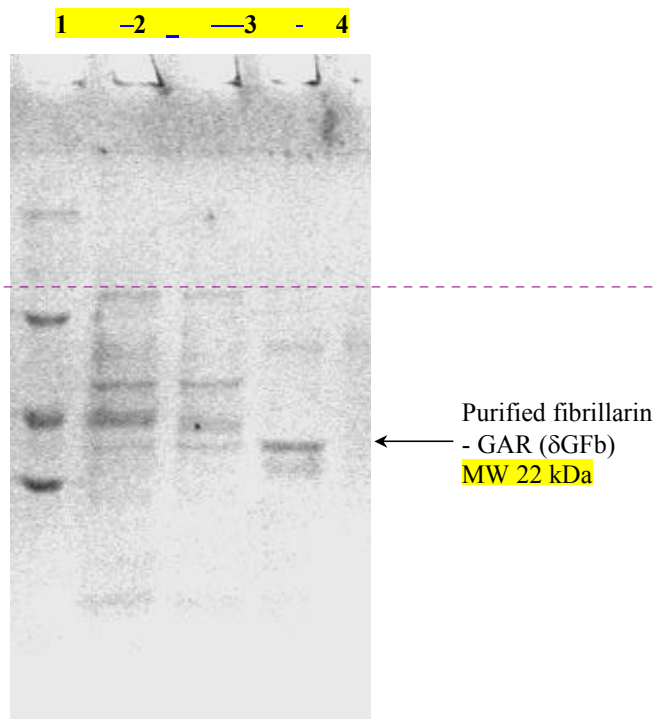


Fig. 1: Homology of amino terminal conserved glycine arginine rich (GAR) domain of *G. lamblia* fibrillarin with fibrillarin sequence from other organisms.

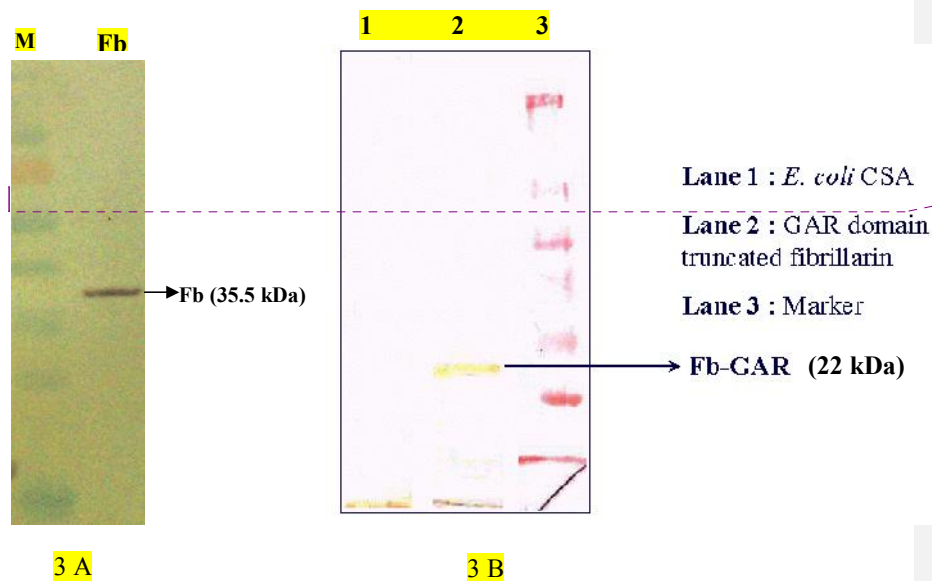
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287



Comment [u2]: MW marker numbers

Fig. 2: Coomassie blue stained SDS PAGE showing purification of GAR domain truncated fibrillar protein : Lane 1 : protein molecular weight marker, Lane 2 & 3: fractions containing wash solutions. Lane 4: fraction containing purified GAR domain truncated fibrillar protein.

288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343



Comment [u3]: MW markers numbers in both membranes.
Write indications for Lane 1, 2 and 3 with 3B
Legend below.

344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399

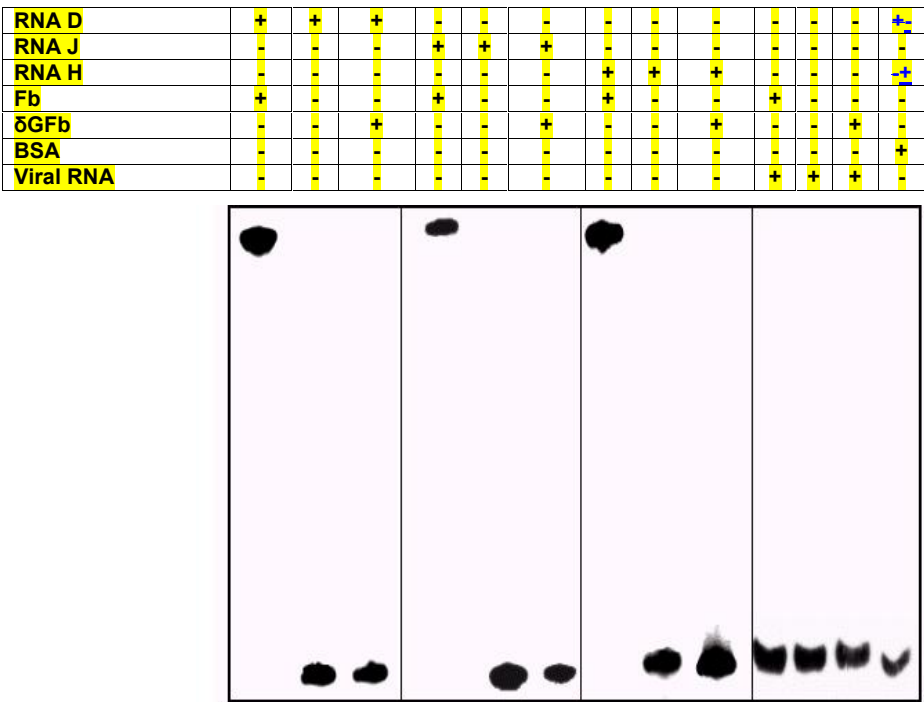
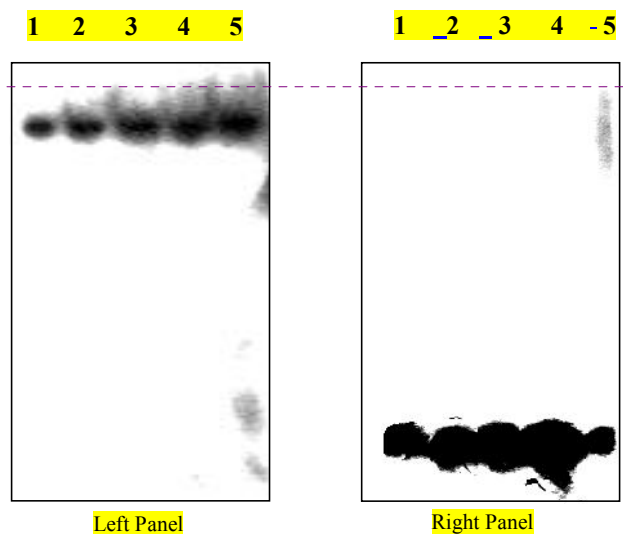


Fig. 4: Gel shift assay showing the interaction of *gGiardia* snoRNAs with fibrillarin (Fb) and GAR truncated fibrillarin (δGFb) proteins. BSA is used as a non-specific protein and viral RNA is used as non-specific RNA to serve as negative controls for this experiment.

Formatted: Font: Italic

400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451



Comment [u4]: Frames

Fig. 5 : Gradient GEMSA with using increasing amounts of full length fibrillarin (left panel) and GAR domain truncated fibrillarin (right panel) incubated with snRNA D. All the lanes containing snRNA D and variable amount of purified recombinant full length or truncated fibrillarin. Lane 1 : 10 μg Fb, Lane 2 : 20 μg Fb, Lane 3 : 40 μg Fb, Lane 4 : 80 μg Fb and Lane 5 : 100 μg Fb. In all the cases full length fibrillarin binds with snRNA D but truncated fibrillarin does not bind with snRNA D.

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

Name	Sequence	Restriction Enzyme Site
GAR(TFb) F	5'-CGGATCCGAUGCGGACCAAGGG-3'	BamHI
GAR(TFb) R	5'-GGAATTCCATATGGAATTCCTCAGCTGCCTTG-3'	NdeI

5. CONCLUSION

Our findings clearly demonstrate that the GAR domain of *Giardial* fibrillarin protein is essential for the binding of this protein with its sno-RNAs. This finding is of potential importance to our understanding of snoRNP biogenesis in *Giardia lamblia*.

Formatted: Font: Italic

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

Authors Sumaliya Karmakar and Dibyendu Raj carried out the necessary bench works and drafted the manuscript. Author Sandipan Ganguly designed the study, analysed the data and wrote the paper.

REFERENCES

1. Adam R. D. Biology of *Giardia lamblia*. Clin Microbiol Rev. 2001; 14(3):447-75
2. Morrison H.G, McArthur A.G, Gillin F.D, Aley S.B, Adam R.D, Olsen G.J. et al. Genomic minimalism in the early diverging intestinal parasite *Giardia lamblia*. Science. 2007;317:1921–1926.
3. Henras A. K, Soudet J, Gerus M, Lebason S, Caizerques-Ferrer M, Mouqin A, et al. The post-transcriptional steps of eukaryotic ribosome biogenesis. Cell. Mol. Life Sci. 2008;65:2334–2359.
4. Tian XF, Yang ZH, Shen H, Adam R.D and Lu SQ. Identification of the nucleoli of *Giardia lamblia* with TEM and CFM. Parasitology Research. 2010; 106(4): 789-93.
5. 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.
6. Bachellerie J.P, Michot B, Nicoloso M, Balakin A, Ni J and Fournier M.J. snoRNAs: a family of nucleolar RNAs with long complementarities to rRNA. TIBS. 1995. 20: 261-264.
7. Maden B.E. Eukaryotic ribosomal RNA. Guides to 95 new angles. Nature. 1997. 389 (6647): 129-131.
8. Yang, C.Y., Zhou, H. and Luo, J. Identification of 20 snoRNA-like RNAs from the primitive eukaryote, *Giardia lamblia*. BBRC, 2005. 328: 1224-1231.
9. Reichow SL, Hama T, Adrian R, D-Amare F and Varani G. The structure and function of small nucleolar ribonucleoproteins. Nucleic Acid Research, 2007. 35(5): 1452-1464.
10. Tycowski, 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.
11. Schimmang T, Tollervey D, Kern H, Frank R and Hurt E.C. A yeast nucleolar protein related to mammalian fibrillarin is associated with small nucleolar RNA and is essential for viability. EMBO J. 8, 4015-4024.

- 502 12. Ghosh S, Ghosh R, Das P and Chattopadhyay D. Expression and purification of
503 recombinant *Giardia* fibrillarin and its interaction with small nuclear RNAs. Protein
504 Expression and Purification. 2001, 21: 40-48.
- 505 13. Bagni C, Lapeyre B. Gar1p binds to the small nucleolar RNAs snR10 and snR30 in
506 vitro through a nontypical RNA binding element. *J Biol Chem*. 1998; 273(18):
507 10868-73.
- 508 14. Henriquez, R.; Blobel, G. and Aris, J. P. Isolation and sequencing of NOPI. A yeast
509 gene encoding a nucleolar protein homologues to a human autoimmune antigen. *J.*
510 *Biol. Chem*. 1990. 265: 2209-2215.
- 511 15. Wormsley S, Samarsky D.A, Fournier M.J and Baserga S.J. An unexpected,
512 conserved element of the U3 snoRNA is required for Mpp10p association. *RNA*.
513 2001;7(6):904-19.
- 514 16. Yoo D, Wootton S.K, Li G, Song C and Rowland R.R. Colocalization and
515 interaction of the porcine arterivirus nucleocapsid protein with the small
516 nucleolar RNA-associated protein fibrillarin. *J Virol*. 2003;77(22):12173-83

Fig. 6.1 : Homology of amino terminal conserved glycine arginine rich (GAR) domain of *G. lamblia* fibrillarin with fibrillarin sequence from other organisms