

RESEARCH ARTICLE

Selective DNA-binding of SP120 (rat ortholog of human hnRNP U) is mediated by arginine-glycine rich domain and modulated by RNA

Mary Miyaji^{1*}, Shinji Kawano², Ryohei Furuta¹, Emi Murakami², Shogo Ikeda², Kimiko M. Tsutsui¹, Ken Tsutsui¹

1 Department of Neurogenomics, Graduate School of Medicine, Dentistry and Pharmaceutical Sciences, Okayama University, Okayama, Japan, **2** Faculty of Science, Department of Biochemistry, Okayama University of Science, Okayama, Japan

* mmiyaji@okayama-u.ac.jp



OPEN ACCESS

Citation: Miyaji M, Kawano S, Furuta R, Murakami E, Ikeda S, Tsutsui KM, et al. (2023) Selective DNA-binding of SP120 (rat ortholog of human hnRNP U) is mediated by arginine-glycine rich domain and modulated by RNA. *PLoS ONE* 18(8): e0289599. <https://doi.org/10.1371/journal.pone.0289599>

Editor: Hari S. Misra, Gandhi Institute of Technology and Management, INDIA

Received: March 15, 2023

Accepted: July 20, 2023

Published: August 4, 2023

Copyright: © 2023 Miyaji et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its [Supporting Information](#) files.

Funding: This work was supported by JSPS KAKENHI Grant Numbers JP23710216 (MM) and JP21K06122 (MM). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

Abstract

A human protein heterogeneous ribonucleoprotein U (hnRNP U) also known as Scaffold attachment factor A (SAF-A) and its orthologous rat protein SP120 are abundant and multi-functional nuclear protein that directly binds to both DNA and RNA. The C-terminal region of hnRNP U enriched with arginine and glycine is essential for the interaction with RNA and the N-terminal region of SAF-A termed SAP domain has been ascribed to the DNA binding. We have reported that rat hnRNP U specifically and cooperatively binds to AT-rich DNA called nuclear scaffold/matrix-associated region (S/MAR) although its detailed mechanism remained unclear. In the present study analysis of hnRNP U deletion mutants revealed for the first time that a C-terminal domain enriched with Arg-Gly (defined here as 'RG domain') is predominantly important for the S/MAR-selective DNA binding activities. RG domain alone directly bound to S/MAR and coexistence with the SAP domain exerted a synergistic effect. The binding was inhibited by netropsin, a minor groove binder with preference to AT pairs that are enriched in S/MAR, suggesting that RG domain interacts with minor groove of S/MAR DNA. Interestingly, excess amounts of RNA attenuated the RG domain-dependent S/MAR-binding of hnRNP U. Taken together, hnRNP U may be the key element for the RNA-regulated recognition of S/MAR DNA and thus contributing to the dynamic structural changes of chromatin compartments.

Introduction

hnRNP U is a highly abundant nuclear protein that binds to both DNA and RNA. This protein was originally isolated from HeLa cells as one of the components of hnRNPs and shown to be associated with poly(A)⁺ RNA [1]. On the other hand, a human nuclear protein termed scaffold attachment factor A (SAF-A) was discovered to bind specifically to scaffold/matrix attachment region (S/MAR) DNA [2]. Sequence determination of SAF-A demonstrated that these are identical proteins [3]. S/MARs are generally AT-rich noncoding sequences that facilitate the formation of specific functional loci in the nucleus [4,5]. We independently identified a rat

ortholog of SAF-A in the nuclear scaffold preparation that selectively binds to S/MAR DNA and named it SP120 (scaffold protein 120 kDa) [6,7]. We use the term “hnRNP U” instead of SP120 throughout the present report although all the experiments were done with rat SP120. “SAF-box”, a short stretch of 31 amino acids at the amino-terminal region of SAF-A, was demonstrated to be responsible for the selective binding to S/MAR [8]. They immobilized SAF-box peptide (amino acid 1–45 of SAF-A) onto Sepharose beads and showed that the S/MAR-specific DNA binding to this peptide is highly dependent on its surface density, which suggested involvement of some cooperative mechanism. In a broader perspective study, homology search with amino acid sequences revealed a conserved motif called “SAP” in vast varieties of DNA-binding proteins including SAF-A, in which the SAF-box coincides with SAP [9]. We present here some evidence that the SAP domain (SAF-box) is also involved in the S/MAR-selective binding of rat hnRNP U, but its C-terminal region appears more essential for the full-range affinity.

As a characteristic feature located in the C-terminal region of human hnRNP U, an RNA binding motif, termed RGG-box was specified within the C-terminal arginine and glycine-rich 112 amino acids [10]. We define here a wider region called ‘RG domain’ that is composed of symmetrically arranged 8x2 RGs flanking a glycine-rich spacer (Fig 1A). RG domain is highly conserved among mammals and harbors the RGG-box (Fig 1B). We have shown previously that a region slightly wider than the RG domain of hnRNP U is required for RNA-mediated interaction with DNA topoisomerase IIβ [11]. In general, RGG/RG motifs are wide-spread in more than 1,000 human proteins. As many proteins with RGG/RG motifs are RNA-binding

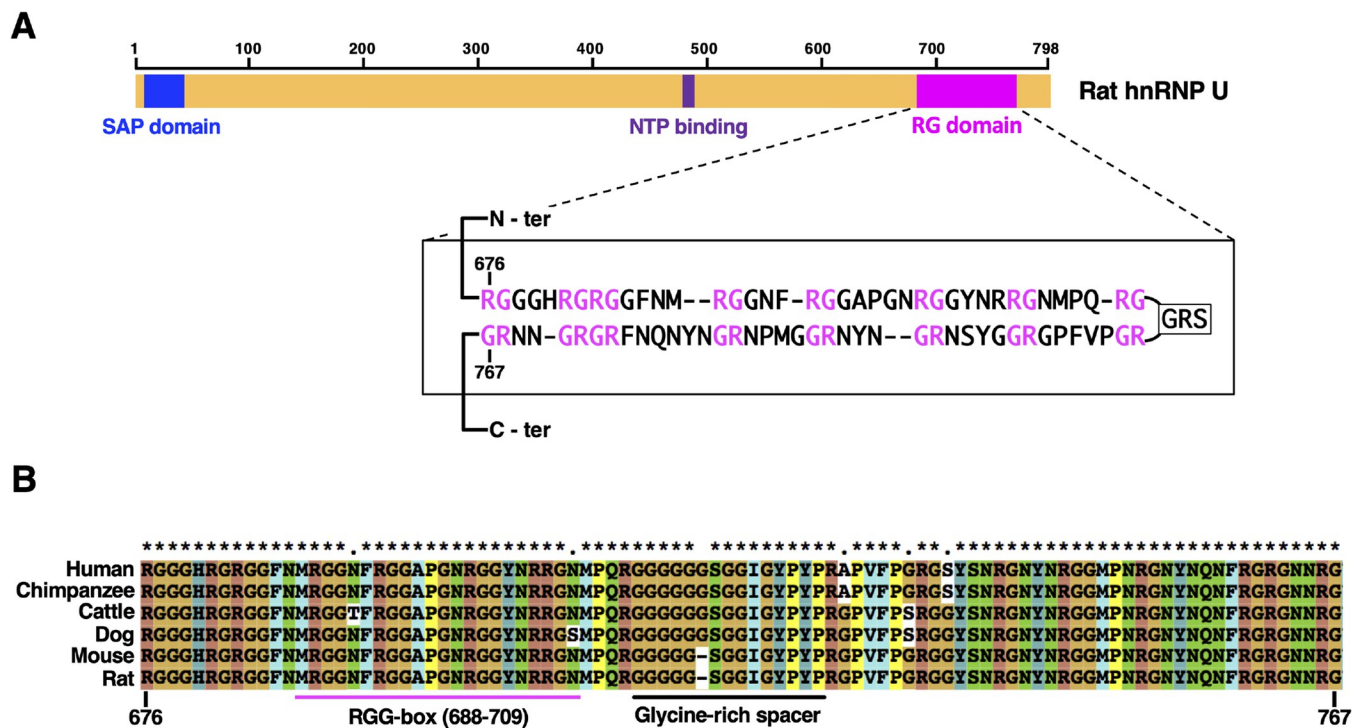


Fig 1. Domain structure of rat hnRNP U. (A) Domain diagram of rat hnRNP U used in this study. SAP (amino acids 8–42), NTP-binding (amino acids 478–485), and RG (amino acids 676–767) domains are indicated as colored boxes on full-length rat hnRNP U (amino acids 1–798). The RG domain is newly defined in this report whose amino acid sequence is shown in the box below. It contains a cluster of 16 RG dipeptides interrupted by the central glycine-rich spacer (GRS). The sequence is folded like a hairpin to highlight the symmetrical distribution of RGs. (B) Sequence alignment of the RG domain which is highly conserved in higher eukaryotes. RGG-box is originally described as a conserved motif in RNA binding proteins [10]. Glycine-rich spacer (GRS) is marked by a black line.

<https://doi.org/10.1371/journal.pone.0289599.g001>

proteins, this motif is considered to be involved in RNA-binding [12], although it is also reported to be involved in binding with double stranded DNAs [13].

Structurally RG domain is an intrinsically disordered region (IDR) [14], and can cause IDR-mediated phase separation [15]. Phase separation has been shown to be closely involved in cellular organization and regulation of nuclear structures [16,17]. It was suggested that microphase separation in the nuclear microenvironment can be switched between transcriptionally active and inactive states, that are reflected by 'A' and 'B' compartments observed in HiC experiments [18]. In the active microphase, sequence-specific multivalent DNA binding proteins are engaged in binding equilibrium with DNA to support transcription. Interestingly, several reports suggest that hnRNP U is required for maintaining the active (A) compartments [19,20]. In this context, RNA molecules and RNA-binding proteins may be main players in maintaining active compartments.

In this study, we constructed several deletion mutant proteins of hnRNP U and examined their DNA-binding and S/MAR-selective binding activities in a S/MAR vs non-S/MAR competitive binding scheme. Interestingly, deletion of RG domain disrupted the S/MAR selective DNA-binding of hnRNP U in soluble conditions. We also found a synergism between SAP and RG domains for the S/MAR-selective binding, which is interfered by excess amounts of RNA. Relationship between SAP and RG domains in the recognition of S/MAR by hnRNP U and a regulatory involvement of RNA will be discussed in view of the structural dynamism of nuclear matrix [20].

Materials and methods

Plasmid constructions

PCR primers for constructing plasmids used in this study are listed in [S1 Table](#). His-Myc constructs of hnRNP U Δ SAP, Δ RG, NC, and NC Δ RG were amplified with Phusion Hot Start DNA polymerase (NEB) by using pCold II-Myc-hnRNP U plasmid and its derivatives as templates [7]. The PCR products were cloned into pCold II-Myc vectors in-frame by swapping the full-length hnRNP U region of pCold II-Myc-hnRNP U. To amplify the DNA fragment encoding Δ RG and NC, PCR was performed in two steps. For the first step, DNA fragments encoding the N-terminal and the C-terminal parts of these mutants were amplified and the products were gel-purified. For the second step, PCR was performed using the primers in [S1 Table](#) and the mixtures of purified N-terminal and the C-terminal fragments prepared by the first step as templates. To amplify the DNA fragment encoding NC Δ RG, PCR was performed in the same way as amplifying the DNA fragment encoding Δ RG by using pColdII-Myc-NC as a template. The DNA fragments encoding His-Myc-NC or His-Myc-NC Δ RG were amplified using the primers listed in [S1 Table](#) with pColdII-Myc-NC or pColdII-Myc-NC Δ RG as templates then re-cloned into pST39 vector [21]. pGEX3X-RG and pGEX3X-SAP, and pGEX3X-SAP-RG were prepared by cloning the PCR-amplified products in-frame into pGEX3X vector. The R to K substitution mutant, pCold II-Myc-RK, was constructed by PCR in three steps. For the first step, N-terminal and C-terminal parts of DNA fragments were amplified using primers and pColdII-Myc-hnRNP U as a template. Middle part containing RG domain was amplified from an artificial gene (supplied by Life Technologies) in which all arginine residues in RG domain were changed to lysine residues. For the second step, PCR was performed using primers and the mixture of purified 2 fragments, N-terminal, middle as templates, and then an amplified fragment (the N- and middle parts were fused) was gel purified. For the third step, PCR was performed using primers and the mixture of the fused fragment and C-terminal parts as templates.

To prepare S/MAR DNA and vector fragments for EMSA, plasmid DNAs were digested with restriction endonucleases under the conditions recommended by the supplier (New England Biolabs) and purified by phenol/chloroform extraction and ethanol precipitation and

dissolved in TE buffer (10 mM Tris-HCl: pH8.0, 0.1 mM EDTA). The DNA fragments (pBS/*FTZ*) were generated by *Eco* RI digestion of pFTZ1 [7]. The plasmid, pBS-*Kcnd2* intergenic region, was constructed as follows. *Kcnd2* gene intergenic region was amplified on rat genomic DNA with PCR primers shown in S1 Table. Rat genomic DNA was prepared from purified brain nuclei of two-week-old Wistar rats as described [22]. The experiment was conducted in accordance with protocols approved by the Animal Care and Use Committee of the Okayama University Medical School to minimize suffering in the animals. Rats were euthanized by exsanguination under anesthetic (intraperitoneal injection of pentobarbital sodium). The product DNA was purified and phosphorylated by T4 polynucleotide kinase (NEB) according to the manufacturer's protocol and cloned into *Eco* RV site of pBlueScript II (SK+). The DNA fragments used for EMSA assay (S1 Fig) were generated by *Bam* HI and *Xho* I digestions of pBS-*Kcnd2* intergenic region. A topoisomerase II β target #31 cloned in pZerO vector [7] was used to generate pZerO/*Sat* I DNA. The plasmid was digested with *Bam* HI and *Xho* I to separate vector and insert. A randomly cloned fragment of *E. coli* DNA (pBS-*E. coli* DNA) was digested with *Eco* RI and *Xho* I to generate pBS/*E. coli* DNA.

Recombinant proteins

Preparations of His-Myc tagged recombinant proteins were performed as previously described [7], except that the lysates were incubated with RNase A (10 μ g/ml, QIAGEN) and DNase I (40 u/ml, Takara) on ice for 30 min and the wash buffer containing 1 M NaCl instead of 0.5 M NaCl was used. To express His-Myc-NC and NC Δ RG, *E. coli* cells BL21(DE3) were transformed with pST39-His-Myc-NC and pST39-His-Myc-NC Δ RG and the transformed cells were grown in LB medium supplemented with 100 μ g/ml ampicillin at 37°C. When the culture reached OD 0.3 (600 nm), 1mM IPTG was added and continued to cultivate at 37°C for 3 hours. Protein purification was conducted as described [7]. GST-fusion proteins were expressed in *E. coli* (Rosetta Blue (DE3) (Novagen)) and *E. coli* cells were recovered by centrifugation. The pellet was resuspended in an appropriate volume of lysis buffer (10 mM Tris-HCl (pH8.0), 500 mM NaCl, 0.5 mM PMSF, 1 mg/ml lysozyme) and incubate on ice for 15 min. Cells were disrupted by sonication (TOMY UD-201 with microtip). RNase A and DNase I were added to the lysate at a final concentration of 10 μ g/ml and 40 u/ml respectively and incubated on ice for 30 min. Triton X-100 and DTT were added at final concentrations of 2% and 5 mM, respectively. The cleared lysates were obtained by centrifugation and stored into several aliquots at -80°C. Protein amounts on beads were estimated by calibration to BSA standards on SDS-PAGE-stained gels. Raw images are shown in S4 Fig.

EMSA

Recombinant hnRNP U-WT and its mutants were incubated at 37°C for 30 min with 25 ng of equimolar mixture of vector and insert DNA fragments in 15 μ l of reaction mixture containing 35 mM Tris-HCl (pH 8.0), 72 mM KCl, 5 mM MgCl₂, 5 mM dithiothreitol (DTT), 5 mM spermidine, and 50 μ g/ml BSA. Samples were then subjected to electrophoresis on a 1% SeaKem GTG agarose gel containing 10 mM MgCl₂ in 0.5 \times TBE (Tris-borate-EDTA) buffer. The tank buffer (0.5 \times TBE) did not contain MgCl₂. After staining with 0.5 μ g/ml EtBr, DNA bands were recorded on VersaDocTM (Bio-rad) and quantified by densitometry. Uncropped original agarose gel images are also shown in S5 Fig.

On-bead binding assay

Bead-bound GST-fusion proteins were prepared by incubating aliquots of glutathione-Sepharose beads with the lysate containing varying amounts of GST-fusion proteins at 25°C for 1 h.

Beads were washed 3 times with wash buffer (10 mM Tris-HCl (pH 8.0), 500 mM NaCl, 0.1% Triton X-100) and then three times with binding buffer (50 mM Tris-HCl (pH8.0), 120 mM KCl, 10 mM MgCl₂, 30 µg/ml BSA, 0.5 mM EDTA, 0.5 mM DTT, 0.1% Triton X-100). DNA fragments (50 ng or 100 ng of pBS/*FTZ*) in binding buffer were added and incubated at 25°C for 1hr with mixing. The unbound DNA was separated by centrifugation. Beads were washed 3 times with the binding buffer and deproteinized by proteinase K (Roche) treatment. The released DNA was separated by centrifugation and recovered as bound fraction. Unbound and bound DNAs were purified and analyzed by 1% agarose gel electrophoresis. To examine the inhibitory effects of netropsin, pBS/*FTZ* (100 ng) was preincubated with fixed amounts of netropsin in the binding buffer at 25°C for 30 min. The preincubated DNA fragments were added to 10 µl of beads coupled with 16.5 pmol of GST-RG. The subsequent processes were performed as described above.

Aggregation assay

DNA fragments (amounts of pBS/*FTZ* indicated in Figure legend) were incubated with purified His-Myc-hnRNP U proteins in 10 µl of aggregation buffer containing 50 mM Tris-HCl (pH 8.0), 120 mM KCl, 10 mM MgCl₂, 0.5 mM EDTA, and 30 µg/ml BSA at 37°C for 30 min. Samples were separated into pellet and supernatant by centrifugation at 20,000 xg for 10 min at 4°C. Pellets were resuspended in 10 µl of aggregation buffer. Two µl of reaction mixture containing 4% SDS and 2 mg/ml Proteinase K (Roche) were added to the supernatants and the pellets and incubated at 55°C for 30 min. In the reactions containing RNA, samples were treated with 10 µg/ml RNase A at 37°C for 10 min to degrade RNA before the addition of SDS and Proteinase K. The samples were subjected to 1.0% agarose gel. DNA bands were detected by staining with GelRed Nucleic Acid Gel Stain (Biotium) and quantified by band densitometry using Image J software. Total RNA was purified from HEK293 cells by using RNeasy Mini kit (Qiagen). The total RNA was fractionated into poly A (+/-) RNA using Dynabeads mRNA Purification Kit (Invitrogen, DYNAL). To investigate the effect of salt concentration on the complex formation, pBS/*FTZ* (25 ng) was incubated with His-Myc-hnRNP U proteins in 25 µl of buffer containing 10 mM Tris-HCl (pH8.0), 10 mM MgCl₂, 1 mM DTT, 1% glycerol and varying NaCl concentrations (Fig 6A and 6B).

Results and discussion

Detection of S/MAR-selective binding activities of hnRNP U by mobility shift assay (EMSA)

To analyze the S/MAR-selective binding of hnRNP U, we employed magnesium-agarose electrophoretic mobility shift assay (EMSA) [23]. Plasmids harboring S/MARs or *E. coli* DNA were digested with restriction enzyme to liberate inserts from the vector portion, resulting in equimolar mixtures of inserts and vector DNA fragments. Vector DNA serves as a non-S/MAR internal control. After purification, the DNA mixture was incubated with purified recombinant hnRNP U protein and separated by agarose gel electrophoresis in the presence of 10 mM magnesium chloride, which is required for the selective binding of S/MAR with hnRNP U [6]. hnRNP U complexed with DNA forms large aggregates that fails to migrate into gel. Therefore, the amount of bound DNA was estimated from the decreased band densities when compared to protein-free DNA bands (leftmost lanes in S1A Fig).

FTZ is a classical S/MAR isolated from the upstream region of *Drosophila Fushitarazu* gene [24]. The *FTZ* insert disappeared completely in the presence of 40 nM hnRNP U whereas the vector band (V) persisted, indicating that the protein binds to *FTZ* with high selectivity

(S1B Fig). Upstream intergenic region of *Kcnd2* gene and rat *Sat I* repeats were obtained previously from DNA clones for topoisomerase II β action sites and found to be AT-rich S/MARs that bind to hnRNP U [7]. These DNA fragments also showed selective binding to hnRNP U as compared to the vector DNA. The C_{50} value reflecting relative affinities of S/MARs toward hnRNP U was estimated from the densitometric profiles (S1B Fig). As shown under the gel images of S1A Fig, C_{50} values roughly increase along with the GC-content of insert DNA fragments. This is consistent with our previous finding that the S/MAR-selective binding of hnRNP U positively correlates with the coverage rate of short A/T-tracts [7]. It is possible that not just GC-content, but also local density of A/T-tracts affects the affinity. The reason why C_{50} values for the same vector fragments vary somewhat in different insert combinations is not clear. These results indicated that magnesium-agarose EMSA is a useful tool for evaluating the AT-rich S/MAR-selective binding of hnRNP U in soluble conditions.

RG domain is required for proper S/MAR-selective binding of hnRNP U in soluble conditions

To determine the domain (s) of hnRNP U required for the S/MAR-selective binding, we constructed several deletion mutants (Fig 2A). SAP domain is a well-conserved region that specifically recognizes S/MAR DNA [8]. RG domain of the rat hnRNP U is a newly defined region (amino acid 676–767) that contains 16 Arg-Gly dipeptide sequences symmetrically positioned

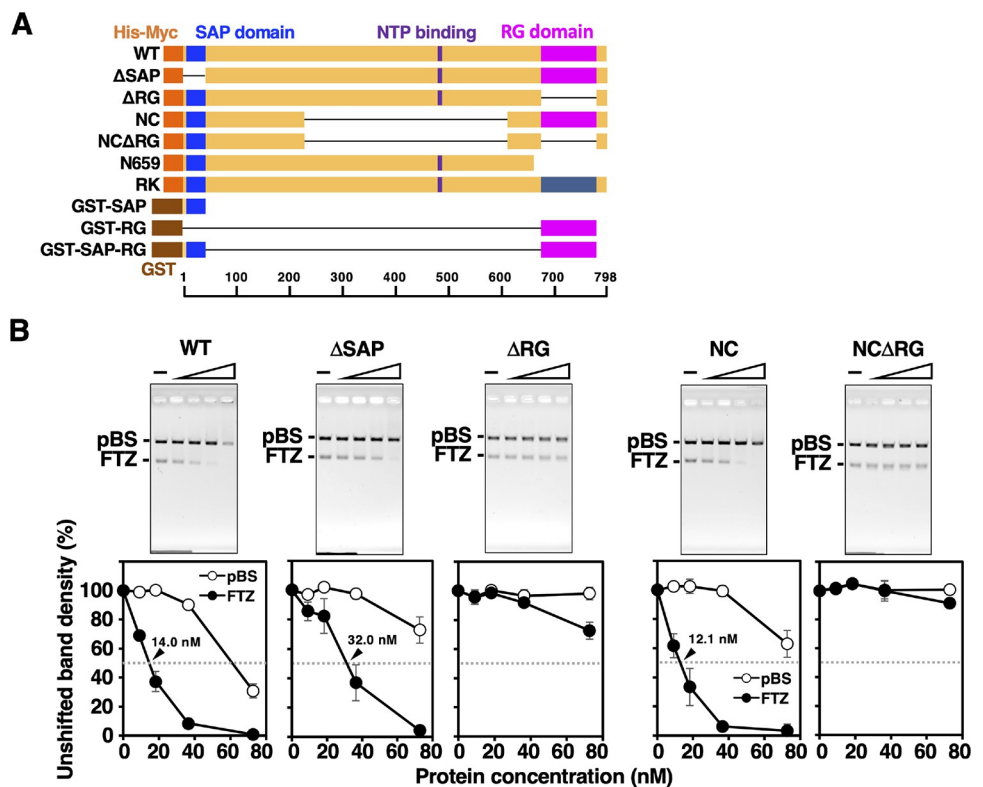


Fig 2. DNA-binding activities of domain deletion mutants. (A) Illustration of wild type (WT) and deletion mutants. The orange and brown boxes at the N-terminus represent His-Myc and GST tags, respectively. (B) Gel images from competitive magnesium-agarose EMSA are shown with band density plots. Concentrations for WT, Δ SAP, Δ RG, NC, and NC Δ RG were varied. Error bars indicate mean \pm standard deviation (SD) from three independent experiments. Half-maximal concentrations (C_{50}) for WT, Δ SAP, and NC are indicated by arrowheads.

<https://doi.org/10.1371/journal.pone.0289599.g002>

around the central glycine-rich spacer (Fig 1A). RG domain is highly conserved in mammals and contains the motif termed RGG-box that interacts with RNA [10] (Fig 1B). RGG or RG motifs exist in more than 1,000 human proteins and claimed to be involved in numerous physiological processes such as transcription, pre-mRNA splicing, DNA damage signaling, mRNA translation, and regulation of apoptosis [12]. We speculate that the whole structure of RG domain with regularly spaced RGs, not just repetitive short RGG/RG motifs, is important for the function of hnRNP U.

We compared the S/MAR-selective binding activities of hnRNP U deletion mutants using *FTZ* as a typical S/MAR (Fig 2B). A deletion mutant of SAP domain (Δ SAP in Fig 2A) had a slightly weaker but significant S/MAR-selective binding activity (C_{50} : 32 nM) as compared to WT (C_{50} : 14 nM). Although the SAP domain has been regarded as the region responsible for the S/MAR-selective binding of SAF-A by using immobilized proteins [8], these results indicate that other region(s) within Δ SAP is involved in the S/MAR-selective binding. To this end, we first examined a deletion mutant of RG domain (Δ RG). Δ RG showed a significantly decreased DNA-binding activity, being unable to determine C_{50} within the range of protein concentrations tested. Thus, RG domain is essential for the proper S/MAR-selective binding of hnRNP U in solution. To assess the role of the central region, we constructed the NC mutant in which the central region with putative ATPase domain (shown by 'NTP binding') was removed (Fig 2A). Remarkably, NC mutant had almost the same activity as WT, indicating that the central region is not required for S/MAR-selective binding of hnRNP U. On the other hand, NCARG showed a very low binding activity like Δ RG. Therefore, we reaffirmed that RG domain is essential for the proper S/MAR DNA binding of hnRNP U under soluble conditions.

RG domain itself has selective and strong S/MAR-binding activity

The DNA-binding activity of the SAP domain is weak in solution but when immobilized on beads, it exhibited a high affinity toward S/MAR, depending positively on its surface density [8]. Our observation with Δ RG and NCARG mutants may suggest that the role of RG domain is simply increasing the local concentration of SAP domain to exert S/MAR-selective binding. Otherwise, it is also possible that the RG domain itself binds directly to S/MAR DNA under these conditions. Although RG domain can be regarded as an RNA-binding domain, gene ontology analysis showed that about half of the proteins with clustered RGG motifs are associated with both RNA- and DNA-binding functions [12].

To compare DNA-binding activities of RG and SAP domains directly, they were expressed as GST-fusion proteins and immobilized at varying concentrations on glutathione-Sepharose beads. Using this on-bead binding assay, we employed the competitive binding scheme with *FTZ* S/MAR plasmid as in Fig 2B (Fig 3). The immobilized SAP domain showed a selective binding toward S/MAR (Fig 3A), which is consistent with the previous report [8]. The immobilized RG domain also showed selective binding to S/MAR, indicating that this domain alone can discriminate S/MAR from non-S/MAR DNAs (Fig 3B). Although RG domain showed higher affinity to *FTZ* compared to SAP as judged from the lower half saturation dose (13.0 pmol vs. 22.5 pmol), overall selectivity for S/MAR appears to be lower than SAP, whose binding is cooperative in nature as reflected by the sigmoidal binding curve (Fig 3A). Interactions between SAP domains bound to neighboring sites on DNA may cause a strong selectivity by a mass binding mode [4]. The dissociation constant (K_d) obtained from a nonlinear regression analysis on Fig 3A and 3B coincided with the C_{50} , and the Hill coefficients were 4.6 and 1.8 for SAP and RG, respectively. This indicates that the binding affinity of GST-RG is higher than GST-SAP but with less cooperativity.

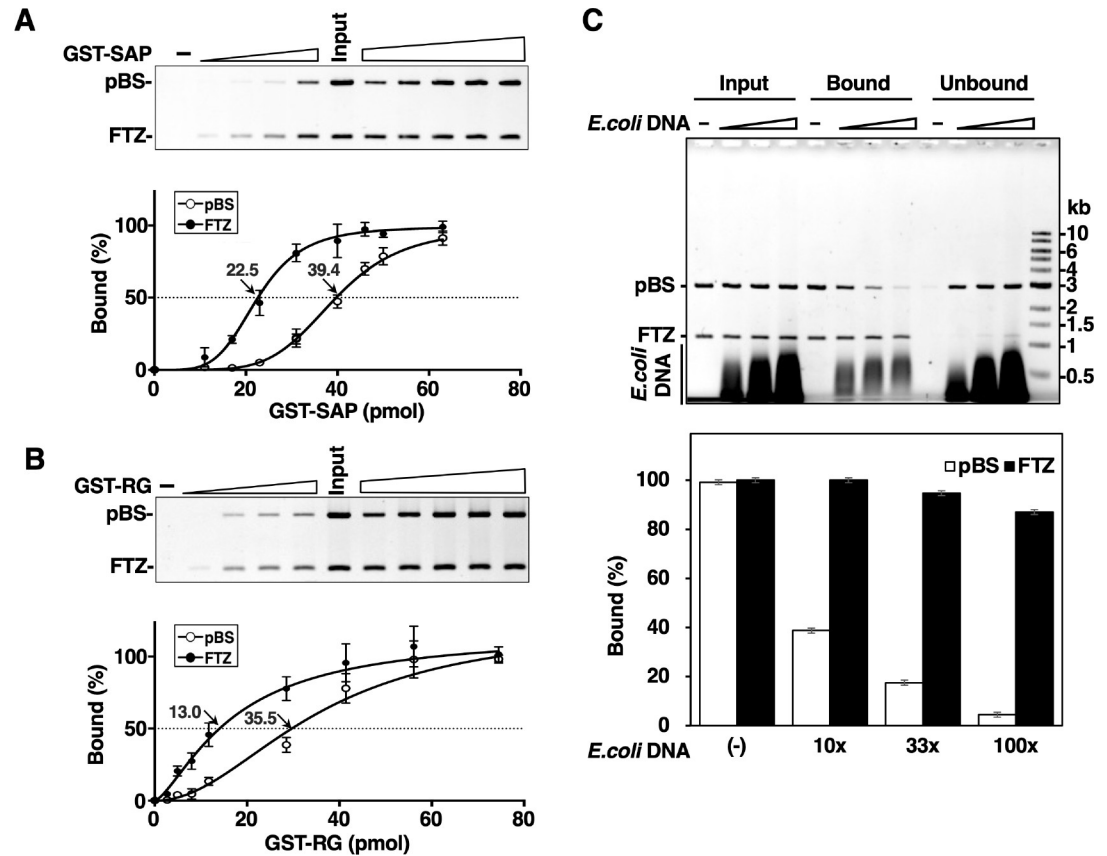


Fig 3. Comparison of DNA-binding activities of SAP and RG domains by on-bead DNA binding assay. (A) On-bead binding assay of SAP domain (amino acids 1–44). Fixed volume (2 μ l) of glutathione-Sepharose beads coupled with increasing amounts of GST-SAP protein were incubated with 100 ng (37.3 fmol) of pBS/FTZ. Bead-bound DNA was analyzed by agarose gel electrophoresis. The quantified band densities relative to those of input are plotted against the amounts of GST-SAP on beads. Error bars indicate mean \pm SD from at three independent experiments. In drawing the regression curve, a nonlinear curve fitting program (Graphpad prism) was applied to mean values. (B) On-bead binding assay of RG domain (amino acids 676–767). The same procedure as above was employed. Error bars indicate as described above. (C) On-bead binding assay of RG domain with or without excess amounts of *E. coli* DNA against 50 ng of pBS/FTZ (10, 33, and 100-folds). GST-RG (56 pmol) on 7.5 μ l beads was used for these experiments. The GC content of pBS and *E. coli* DNA fragments used as competitor was 50.4% and 50.8%, respectively. Error bars indicate mean \pm SD from three independent experiments.

<https://doi.org/10.1371/journal.pone.0289599.g003>

To further analyze the S/MAR-selective binding of RG domain, excess amounts of short *E. coli* DNA fragments (<1 kb, GC% = 50.8) were added as a competitor to the on-bead binding assay with pBS/FTZ (Fig 3C). A saturating dose of RG domain fixed on beads bound all input DNA fragments. As increasing the added competitor, binding of non-S/MAR DNA (pBS, GC% = 50.4) decreased whereas S/MAR (FTZ, GC% = 32.7) remained bound to GST-RG even in the presence of 100-fold excess amount of *E. coli* DNA against pBS/FTZ (Fig 3C). Taken together, RG domain itself has a strong DNA-binding activity with significant S/MAR-selectivity. This may be the reason why hnRNP U selectively binds to S/MARs under soluble conditions even if SAP domain was deleted (Fig 2B).

RNA modulates the DNA selectivity of hnRNP U

It is commonly accepted that the C-terminal region of hnRNP U containing RGG-box has an RNA-binding activity [10]. We next investigated whether RNA affects the DNA-binding and S/MAR-selectivity of hnRNP U. To examine the effect(s) of RNA on the hnRNP U-S/MAR

interaction, we employed aggregation assays in which large complexes formed by hnRNP U and nucleic acids were separated by centrifugation into pellet and supernatant fractions, where the binding of hnRNP U to DNA but not RNA forms large pelletable aggregates [3]. Total cellular RNA purified from HEK293 cells was applied as an RNA source (Fig 4). We have shown

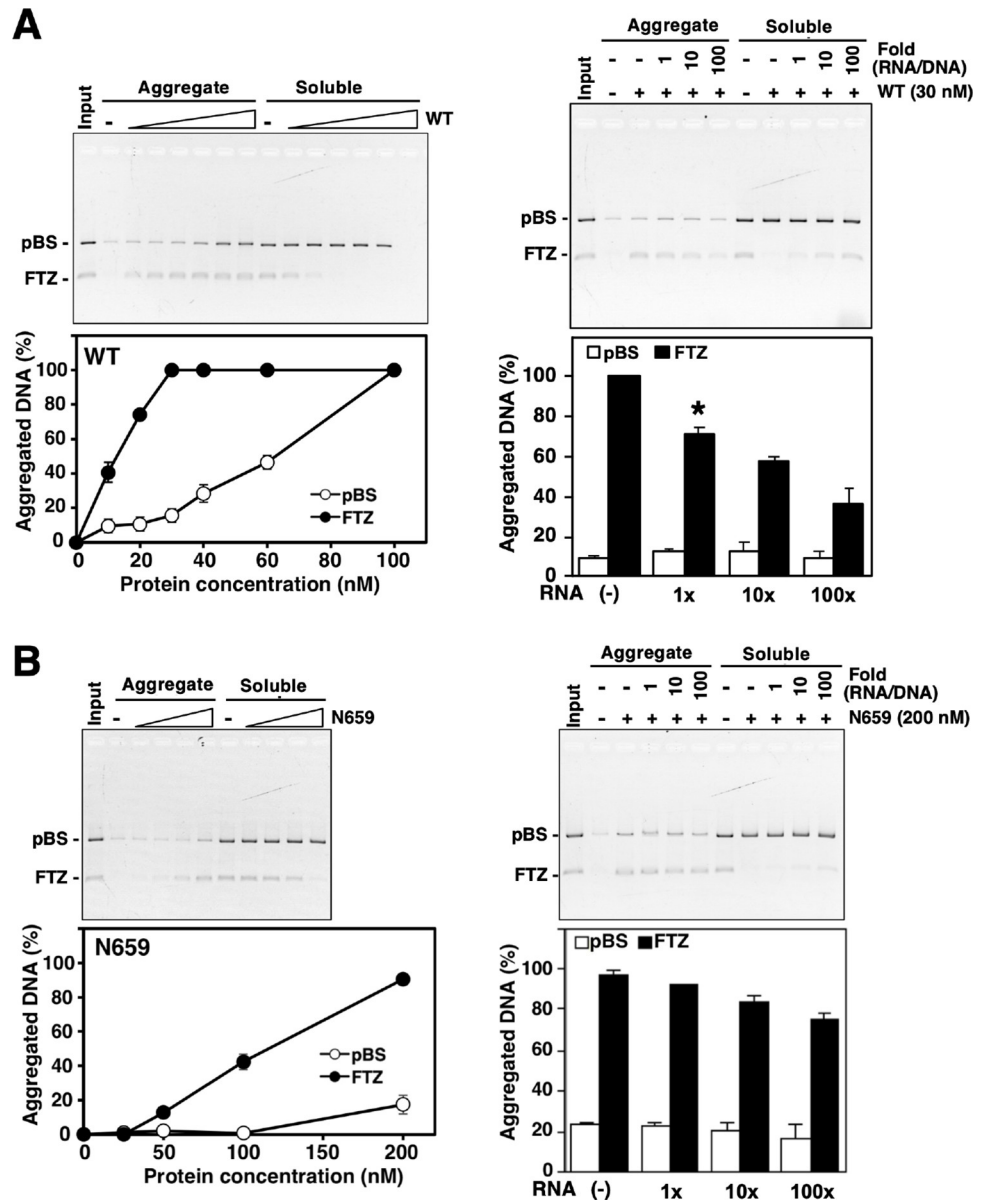


Fig 4. Effects of RNA on S/MAR-selective binding of hnRNP U. Aggregation assays using 20 ng of pBS/FTZ incubated with varying dose of recombinant hnRNP U proteins in the presence or absence of RNA. After the binding reaction, reaction mixture was centrifuged to separate into pellet (aggregate) and supernatant (soluble) fractions. Gel images and densitometric plots are shown in pairs. (A) Increasing concentrations of hnRNP U (WT) alone (left panels). The same experiment was done at 30 nM (right panels) of WT with or without excess amounts of total RNA (indicated by 'fold' against pBS/FTZ DNA). Bars plotted are average of 3 experiments with standard deviation. The decrease in aggregated DNA at 1x fold RNA was statistically significant by Student's *t* test (marked by asterisk, $p < 0.01$). (B) Increasing concentrations of N659 alone (left panels). The same experiment was done at 200 nM of N659 with or without excess amounts of total RNA (right panels). Error bars indicate mean \pm SD from three independent experiments. No statistical significance in the reduction of aggregated DNA was detected at 1x RNA ($p < 0.05$).

<https://doi.org/10.1371/journal.pone.0289599.g004>

previously that hnRNP U and topoisomerase II β forms a complex in the presence of total RNA, in which C-terminal region of hnRNP U was essential [11].

In the absence of hnRNP U, input DNA was recovered mainly in supernatant (Fig 4A, left panels). Amount of aggregated FTZ fragment increased sharply with increasing concentration of hnRNP U-WT. At a low concentration of hnRNP U-WT (30 nM), almost all FTZ but very little pBS fragments were aggregated into pellets, indicating the binding occurs in a S/MAR-selective manner. At a higher concentration of hnRNP U-WT (100 nM), however, almost all input DNA (both FTZ and pBS) were aggregated without any selectivity. In the presence of 30 nM WT (Fig 4A, right panels), the amount of aggregated FTZ fragment decreased with increasing amounts of RNA added, implying that the effect of RNA is competitive in nature since the binding sites for DNA and RNA overlap in the RG domain of hnRNP U.

To confirm that RG domain is responsible for the RNA effects, we prepared N659 that lacks the C-terminal region including RG domain (Fig 2A). N659 required much higher concentration to attain the same level of FTZ aggregation, almost ten times compared to WT (Fig 4B, left panels). This is consistent with the EMSA result with the deletion mutant lacking RG domain (Fig 2B). Therefore, the RNA effect had to be tested at a higher concentration of 200 nM (Fig 4B right panels). Aggregated FTZ by binding to N659 was only slightly reduced even at 100-fold excess amount of RNA. As the binding of WT to FTZ was significantly suppressed with the equal amount of RNA (Fig 4A, right panels), these results support the notion that RG domain is involved in the selective binding of S/MAR DNA, which is attenuated by the competitive binding of RNA on the same domain.

We have shown previously that C-terminal domain (amino acid 612–798) of hnRNP U is associated with poly(A)⁺ RNA [11]. Random cloning and sequencing of the associated RNA revealed partial sequences of various mRNAs. Interestingly, poly(A)⁺ RNA was replaceable by synthetic homopolymer poly (G), but not by other homopolymers, poly (A/C/U). Recently RGG domain of hnRNP U was reported to bind specifically to G-quadruplex RNA [25]. Poly (G) sequences have a high propensity to form G-quadruplex structure [26], and a transcriptome-wide search detected frequent G-quadruplex structures in poly(A)⁺ RNA [27]. In fact, poly (G) most efficiently inhibited the S/MAR-binding of hnRNP U in the aggregation assay (Fig 5). Although effective RNA components *in vivo* remain to be determined, the competitive binding between S/MAR DNA and RNA with respect to the RG domain of hnRNP U may play regulatory roles in the nuclear environment through microphase separation. As both S/MAR DNA and RNA are composed of multiple molecular species, their enormous combination would substantiate subtle levels of regulation.

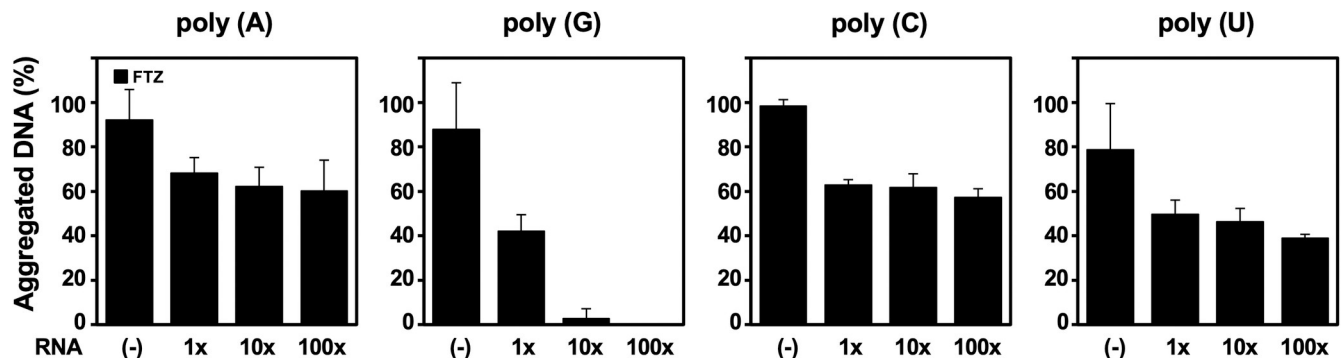


Fig 5. Effects of synthetic homopolymers on S/MAR-selective binding of hnRNP U. Aggregation assays using 20 ng of pBS/FTZ incubated with 20 nM of WT protein in the presence and absence of synthetic homopolymers that are indicated on top of the Figure. Added RNA amounts shown on bottom are expressed by 'fold' against pBS/FTZ DNA. Error bars indicate mean \pm SD from three independent experiments.

<https://doi.org/10.1371/journal.pone.0289599.g005>

Arginine residues in RG domain are important for proper S/MAR-binding activity of hnRNP U

Abundance and regular positioning of arginine residues in RG domain suggests that the guanidino group (or its positive charge) of arginine is responsible for the selective binding of hnRNP U to S/MAR DNAs. To test this, a mutant protein named RK was constructed, in which all arginines were substituted with lysines (a positive charge analogue of arginine), and the aggregation assay was performed at increasing salt concentrations to compare the stability of binding (Fig 6). At 50 mM NaCl, RK mutant also showed S/MAR-selectivity but its affinity toward FTZ was lower than WT as the protein concentration required to aggregate the same amount of DNA was doubled. RK mutant showed much lower affinity at higher NaCl concentrations (compare the aggregated DNA with 30 nM protein at 150 mM NaCl, for instance). These data may indicate that in RK mutant the interaction between DNA is mainly ionic through the ϵ -amino group of lysine but in WT protein hydrogen bonding of the guanidino group of arginine between base pairs may enable the strong and specific interaction. Thus, in physiological ionic conditions, arginine residues in RG domain are likely to be important for stable interaction between hnRNP U and S/MAR DNA.

Netropsin inhibits the S/MAR-selective binding of RG domain

S/MARs are generally AT-rich sequences enriched with short A/T-tracts that are recognized by S/MAR-binding proteins. This interaction is suppressed specifically by distamycin A, an

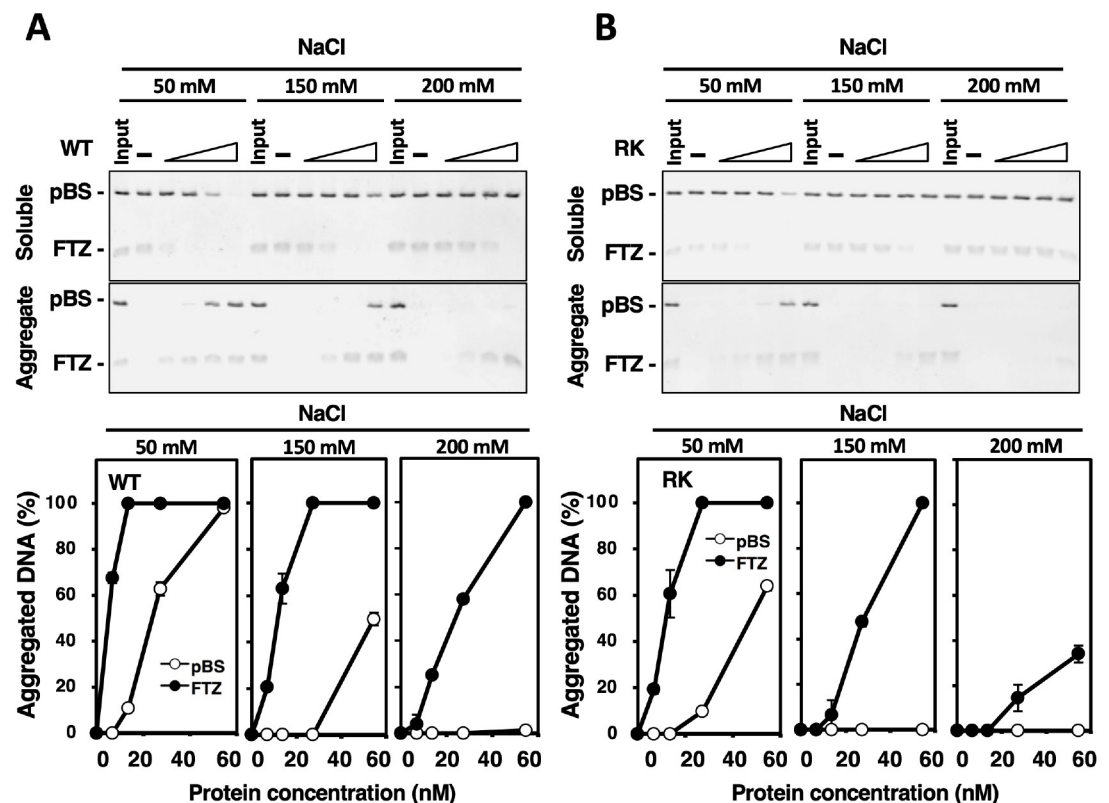


Fig 6. Effects of total substitution of arginines to lysines in the RG domain (RK mutant) on S/MAR-binding activity of hnRNP U. Aggregation assays with WT (A) and RK mutant (B) using 25 ng of pBS/FTZ at different NaCl concentrations (50, 150 and 200 mM). Gel images from aggregation assays are shown with band density plots below. Error bars indicate mean \pm SD from three independent experiments.

<https://doi.org/10.1371/journal.pone.0289599.g006>

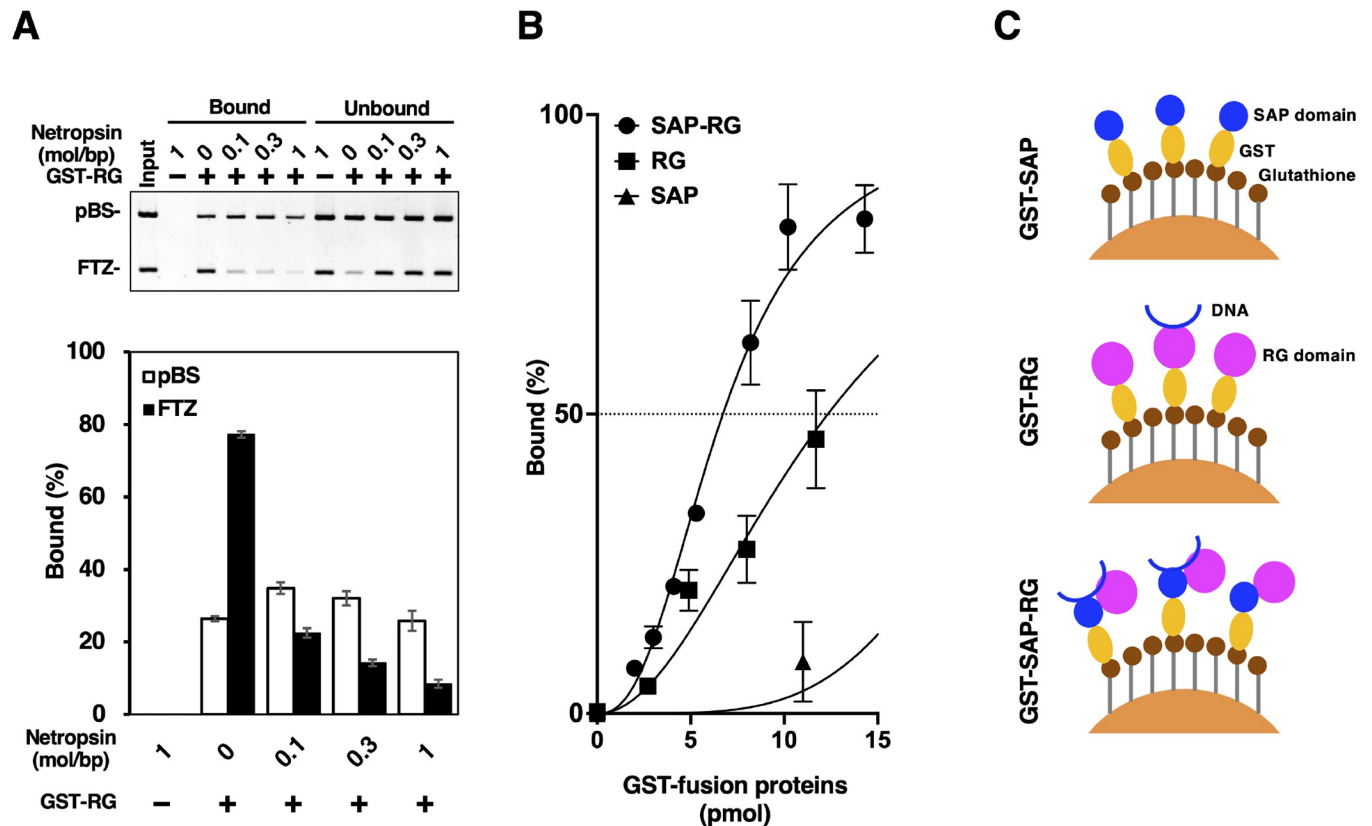


Fig 7. Effects of netropsin on the binding of FTZ to RG domain and possible synergism between SAP and RG domains. (A) Effects of netropsin on the selective binding of FTZ. For on-bead binding assays, GST-fusion proteins immobilized on glutathione-Sepharose beads (7.5 μ l) was incubated with 100 ng of pBS/FTZ (37.3 fmol) and the bead-bound DNA fragments were analyzed. pBS/FTZ was pre-incubated with or without netropsin at the molar ratio indicated in the Figure (drug vs base pair in DNA). GST-RG (15 pmol) was used for the assays. Error bars indicate mean \pm SD from three independent experiments. (B) On-bead binding assay of GST-SAP-RG. The same procedure as Fig 3A and 3B was employed. Data points for GST-SAP and GST-RG are transcribed from Fig 3A and 3B, respectively, and overlaid on the same Figure. Error bars indicate mean \pm SD from three independent experiments. In drawing the regression curve, a nonlinear curve fitting program (Graphpad prism) was applied to mean values. (C) Schematic representation of the mechanism for synergistic binding.

<https://doi.org/10.1371/journal.pone.0289599.g007>

antibiotic known to inhibit transcription [28]. Distamycin has a protonated nitrogen atom as in the guanidino group of arginine. To see whether similar mechanism is involved in the S/MAR-selective binding of hnRNP U, we used netropsin, a compound analogous to distamycin. Netropsin was shown to be a minor groove binder with AT pair preference [29]. It was also reported that narrow minor grooves of DNA are often associated with A/T-tracts and thus expressing enhanced negative electrostatic potentials, which stabilize the insertion of arginine residues in varieties of protein-DNA complexes [30].

It is likely, therefore, that arginine residues within RG domain interact with minor grooves of S/MAR DNA at the A/T-tracts. If so, it is expected that netropsin would antagonize the S/MAR-selective binding of RG domain. The on-bead binding assay demonstrated that the S/MAR-binding to RG domain (15 pmol) was selectively inhibited in a dose-dependent manner by pretreatment of DNA with netropsin (Fig 7A). The result supports the notion that RG domain of hnRNP U recognizes the minor groove of S/MAR through its arginine residues.

A synergistic action between RG and SAP domains

Since DNA-binding activities of SAP and RG domains were assessed separately (Fig 3A and 3B), we next investigated the activity when these two domains coexist on the same molecule

(GST-SAP-RG) to see whether synergistic effects are observed under the same conditions (Fig 7B). Regression analysis showed that K_d and Hill coefficient for GST-SAP-RG are 6.7 and 2.5, respectively. If SAP and RG domains bind DNA independently in GST-SAP-RG, the *FTZ* DNA binding will be additive. However, the result indicated that the level of bound DNA far more exceeds the sum of bound DNA with these domains alone and, more importantly, with higher binding affinity as reflected by the smaller K_d value, which is 13.0 and 22.5 for GST-RG and GST-SAP, respectively. This is a good indication that a synergistic interaction between SAP and RG domains is operating in the same molecule (shown schematically in Fig 7C). These domains may occupy close positions intramolecularly in native hnRNP U molecule although their direct physical association remains to be shown. A 3D structure of rat hnRNP U predicted by AlphaFold, a highly accurate protein structure prediction tool [31], showed that SAP forms a compact domain with two α -helices, but RG domain takes a disordered and extended structure (S3 Fig). These domains indeed can come close together because of the presence of a long stretch of IDR between SAP and central globular domains.

It is important to discriminate these domains with respect to the mode of interaction with S/MAR DNA. SAP domain requires multiple copies to form a stable association with single molecule of S/MAR (thus autosynergistic), but RG domain may require much less copies. The difference should be related to their domain size and amino acid composition. RG domain is about twice the size of SAP domain and is much enriched with basic amino acids, especially arginine. Arginine residues are supposed to interact directly with short A/T-tracts that are abundant in S/MAR DNA [7]. In contrast, molecular mechanism for the S/MAR-selectivity of SAP domain has never been proposed at this level. We speculate that in native hnRNP U molecule RG domain is the primary determinant of S/MAR recognition and intramolecular SAP domain ensures the binding. A plausible binding mechanism involved here is shown graphically in Fig 8.

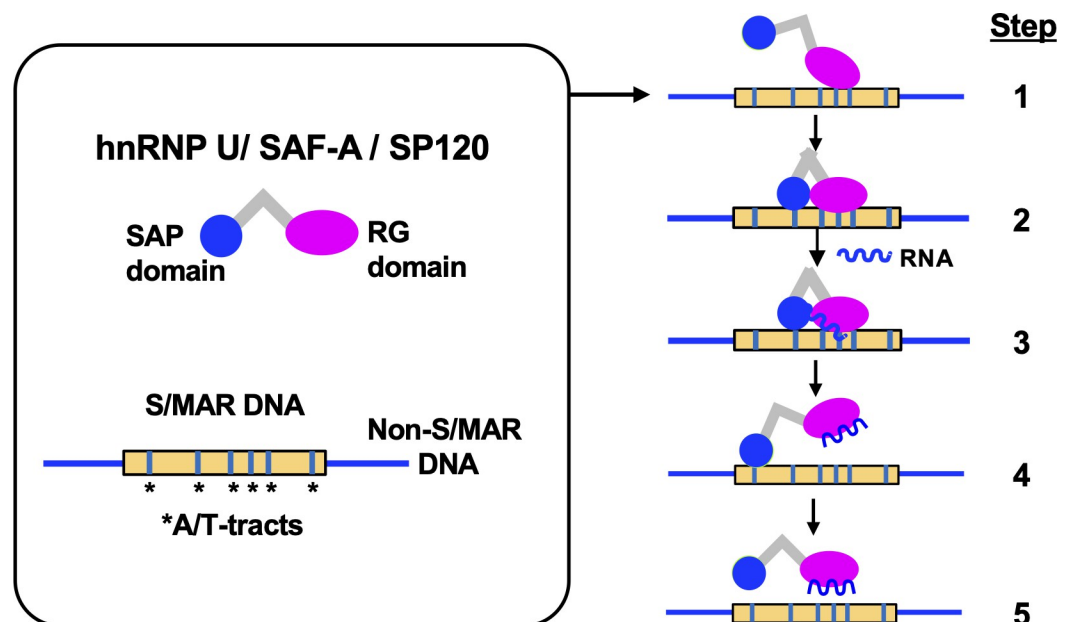


Fig 8. Schematic representation of the binding mechanism between hnRNP U and S/MAR DNA. To aid understanding, the binding process is expressed stepwise, which may not be the case in actual environment. Step1: RG domain recognizes and binds to A/T-tracts of S/MAR DNA. Step2: SAP domain interacts with RG domain and secure the binding. Step3: Ambient RNA binds to RG domain. Step4: RG domain is released from DNA and the interaction between SAP domain is disrupted. Step5: SAP domain with limited affinity is dissociated from DNA and hnRNP U entire molecule is released.

<https://doi.org/10.1371/journal.pone.0289599.g008>

Conclusions and perspective

In this study, we demonstrated unequivocally for the first time that the RG domain of hnRNP U possesses S/MAR-selective DNA binding activity by itself, which had been attributed solely to the SAP domain. Previously known functions of hnRNP U RG domain are localization of hnRNP U to the nuclear matrix [32], interaction with topoisomerase II β [11], and association with inactive X chromosome [33,34], for which all RNA binding is considered important. The present study thus adds one more function to RNA, that is the modulation of S/MAR-selective DNA binding of hnRNP U.

A significant proportion of noncoding RNAs are highly repetitive RNAs like in LINE-1 transcripts and are associated with nuclear scaffold/matrix in euchromatin regions [35]. They presented evidence for involvement of SAF-A in the retention of RNA. Recent report showed that SAF-A interacts with chromatin-associated RNAs via its “RGG domain” (amino acid #695–797) to regulate interphase chromatin structures [20]. ATP binding/hydrolysis by the ATPase domain (designated ‘NTP binding’ in Fig 1A) and RNA-binding state of “RGG domain” were thought to be involved in oligomerization of SAF-A and the interconversion of a large-scale chromatin folding that affects the transcription of neighboring genes. More recently, it is proposed that a phenomenon called ‘phase separation’ is operating in the formation of subnuclear compartments like nuclear bodies and nucleoli, where RNA is one of the major role-players. In view of this emerging concept of nuclear organization, the multifunctional protein hnRNP U/SAF-A/SP120 can now be regarded as a leading player of dynamic nuclear transactions [20]. Our finding that RNA attenuates the interaction between hnRNP U and S/MAR DNA should be important in its regulatory function.

Supporting information

S1 Fig. S/MAR-selective binding activity of hnRNP U detected by magnesium-agarose EMSA. (A) Competitive EMSA analysis with equimolar mixtures of vectors and featured inserts. Twenty-five ng each of DNA were cleaved into vector (V) and insert (I) fragments, incubated with varying amounts of tag-purified hnRNP U, and analyzed by EMSA. Inserts were derived from FTZ S/MAR, *Kcnd2* upstream intergenic region, Satellite I repeat (*Sat I*), and a randomly cloned *E. coli* DNA fragment (used as non-S/MAR control). To characterize the inserts, whose GC content, fragment length and C_{50} values are summarized at the bottom of the gel image. GC% for the vectors, pBS and pZErO, are 50.4 and 54.1, respectively. (B) Relative band densities for the vector and the insert bands were plotted against the concentration of hnRNP U. The dashed lines show the 50% level of unshifted band densities and arrow heads show the C_{50} of inserts. Maps of the recombinants used here is summarized in S2 Fig. (PDF)

S2 Fig. Map of plasmids used in S1 Fig. (A) pBS-FTZ (B) pBS-*Kcnd2* intergenic region (C) pZErO-*Sat I* DNA (D) pBS-*E. coli* DNA are shown. Restriction enzyme sites to separate the vector and the insert are indicated with nucleotide number. The insert region and the selection marker in vectors are painted blue and yellow, respectively. (PDF)

S3 Fig. Analysis of rat hnRNP U 3D structure on AlphaFold program (v2.0 pipeline). The sequence used for analysis is rat hnRNP U in UniProt database (Q6IMY8). Shown here is a “predicted alignment error plot”. The shade of green indicates expected distance error in Ångströms. The color at (x, y) corresponds to the expected distance error in residue x’s position, when the prediction and true structure are aligned on residue y. Namely, dark green is good (low error) and light green is bad (high error). This diagram clearly shows that hnRNP U

has folded structure in two positions. The small region close to N terminus coincides with SAP domain, whereas RG domain is located within the C-terminal disordered region. Presence of a long stretch of disordered segment between SAP domain and central structured domain may increase the mobility of SAP domain and thus facilitates the interaction with RG domain. (PDF)

S4 Fig. Purified recombinant proteins analyzed on SDS-PAGE. His-Myc-tagged or GST-fused proteins illustrated in Fig 2A were run on SDS-PAGE and stained with Coomassie Brilliant Blue. In each panel, gel images of proteins labeled on top are shown on the left, accompanied by uncropped images on the right. The cropped areas are boxed on the raw SDS-PAGE images. Amounts of proteins applied on to gels were 330, 770, 510, 300, 510, 360 and 100 ng for WT, RK, ΔRG, ΔSAP, NC, NCΔRG and N659, respectively. Proteins applied were 3.0 mg, 2.5 mg and 410 ng for GST-fused SAP, RG, and SAP-RG, respectively. M stands for molecular weight marker. (PDF)

S5 Fig. Uncropped agarose gel images used in this study. Cropped regions are boxed with black lines. For Fig 4A and 4B, 'Left' and 'Right' on top of the images indicate their positions in the figure. (PDF)

S1 Table. List of PCR primers used for plasmid construction. (XLSX)

Acknowledgments

We thank Kazuko Kiyama for technical assistance and Yuta Ueda and Makoto Yasui for the construction of expression vector, pColdII-Myc-hnRNP U N659.

Author Contributions

Conceptualization: Mary Miyaji, Kimiko M. Tsutsui, Ken Tsutsui.

Data curation: Mary Miyaji, Kimiko M. Tsutsui, Ken Tsutsui.

Formal analysis: Mary Miyaji, Shinji Kawano, Kimiko M. Tsutsui, Ken Tsutsui.

Funding acquisition: Mary Miyaji.

Investigation: Mary Miyaji, Shinji Kawano, Ryohei Furuta, Emi Murakami.

Project administration: Mary Miyaji, Kimiko M. Tsutsui, Ken Tsutsui.

Supervision: Shogo Ikeda, Kimiko M. Tsutsui, Ken Tsutsui.

Validation: Mary Miyaji.

Visualization: Mary Miyaji, Shinji Kawano, Ryohei Furuta, Kimiko M. Tsutsui, Ken Tsutsui.

Writing – original draft: Mary Miyaji.

Writing – review & editing: Kimiko M. Tsutsui, Ken Tsutsui.

References

1. Dreyfuss G, Choi YD, Adam SA. Characterization of heterogeneous nuclear RNA-protein complexes in vivo with monoclonal antibodies. *Mol Cell Biol.* 1984; 4: 1104–1114. <https://doi.org/10.1128/mcb.4.6.1104-1114.1984> PMID: 6204191

2. Romig H, Fackelmayer FO, Renz A, Ramsperger U, Richter A. Characterization of SAF-A, a novel nuclear DNA binding protein from HeLa cells with high affinity for nuclear matrix/scaffold attachment DNA elements. *EMBO Journal*. 1992; 11: 3431–3440. <https://doi.org/10.1002/j.1460-2075.1992.tb05422.x> PMID: 1324173
3. Fackelmayer FO, Dahm K, Renz A, Ramsperger U, Richter A. Nucleic-acid-binding properties of hnRNP-U/SAF-A, a nuclear-matrix protein which binds DNA and RNA in vivo and in vitro. *Eur J Biochem*. 1994; 221: 749–757. <https://doi.org/10.1111/j.1432-1033.1994.tb18788.x> PMID: 8174554
4. Tsutsui KM, Sano K, Tsutsui K. Dynamic view of the nuclear matrix. *Acta Med Okayama*. 2005; 59: 113–120. <https://doi.org/10.18926/AMO/31953> PMID: 16155636
5. Liebich I, Bode J, Reuter I WE. Evaluation of sequence motifs found in scaffold/matrix-attached regions (S/MARs). *Nucleic Acids Res*. 2002; 30: 3433–3442. <https://doi.org/10.1093/nar/gkf446> PMID: 12140328
6. Tsutsui K, Tsutsui KM, Okada S, Watarai S, Seki S, Yasuda T, et al. Identification and characterization of a nuclear scaffold protein that binds the matrix attachment region DNA. *Journal of Biological Chemistry*. 1993; 268: 12886–12894. [https://doi.org/10.1016/S0021-9258\(18\)31469-8](https://doi.org/10.1016/S0021-9258(18)31469-8) PMID: 8509422
7. Miyaji M, Furuta R, Sano K, Tsutsui KM, Tsutsui K. Genomic regions targeted by DNA topoisomerase II β frequently interact with a nuclear scaffold/matrix protein hnRNP U/SAF-A/SP120. *J Cell Biochem*. 2015; 116: 677–685. <https://doi.org/10.1002/jcb.25024> PMID: 25418483
8. Kipp M, Göhring F, Ostendorp T, van Drunen CM, van Driel R, Przybylski M, et al. SAF-Box, a conserved protein domain that specifically recognizes scaffold attachment region DNA. *Mol Cell Biol*. 2000; 20: 7480–7489. <https://doi.org/10.1128/MCB.20.20.7480-7489.2000> PMID: 11003645
9. Aravind L, Koonin E V. SAP—a putative DNA-binding motif involved in chromosomal organization. *Trends Biochem Sci*. 2000; 25: 112–114. [https://doi.org/10.1016/s0968-0004\(99\)01537-6](https://doi.org/10.1016/s0968-0004(99)01537-6) PMID: 10694879
10. Kiledjian M, Dreyfuss G. Primary structure and binding activity of the hnRNP U protein: binding RNA through RGG box. *EMBO J*. 1992; 11: 2655–2664. <https://doi.org/10.1002/j.1460-2075.1992.tb05331.x> PMID: 1628625
11. Kawano S, Miyaji M, Ichiyasu S, Tsutsui KM, Tsutsui K. Regulation of DNA Topoisomerase II β through RNA-dependent association with heterogeneous nuclear ribonucleoprotein U (hnRNP U). *J Biol Chem*. 2010; 285: 26451–26460. <https://doi.org/10.1074/jbc.M110.112979> PMID: 20554522
12. Thandapani P O'Connor TR, Bailey TL, Richard S. Defining the RGG/RG motif. *Mol Cell*. 2013; 50: 613–623. <https://doi.org/10.1016/j.molcel.2013.05.021> PMID: 23746349
13. Adams MM, Wang B, Xia Z, Morales JC, Lu X, Donehower LA, et al. 53BP1 oligomerization is independent of its methylation by PRMT1. *Cell Cycle*. 2005; 4: 1854–1861. <https://doi.org/10.4161/cc.4.12.2282> PMID: 16294047
14. Ozdilek BA, Thompson VF, Ahmed NS, White CI, Batey RT, Schwartz JC. Intrinsically disordered RGG/RG domains mediate degenerate specificity in RNA binding. *Nucleic Acids Res*. 2017; 45: 7984–7996. <https://doi.org/10.1093/nar/gkx460> PMID: 28575444
15. Chong PA, Vernon RM, Forman-Kay JD. RGG/RG motif regions in RNA binding and phase separation. *J Mol Biol*. 2018; 430: 4650–4665. <https://doi.org/10.1016/j.jmb.2018.06.014> PMID: 29913160
16. Yoo H, Triandafillou C, Drummond DA. Cellular sensing by phase separation: Using the process, not just the products. *J Biol Chem*. 2019; 294: 7151–7159. <https://doi.org/10.1074/jbc.TM118.001191> PMID: 30877200
17. Feric M, Misteli T. Phase separation in genome organization across evolution. *Trends Cell Biol*. 2021; 31: 671–685. <https://doi.org/10.1016/j.tcb.2021.03.001> PMID: 33771451
18. Brackley CA, Marenduzzo D. Bridging-induced microphase separation: photobleaching experiments, chromatin domains and the need for active reactions. *Brief Funct Genomics*. 2020; 19: 111–118. <https://doi.org/10.1093/bfgp/elz032> PMID: 31971237
19. Fan H, Lv P, Huo X, Wu J, Wang Q, Cheng L, et al. The nuclear matrix protein HNRNPU maintains 3D genome architecture globally in mouse hepatocytes. *Genome Res*. 2018; 28: 192–202. <https://doi.org/10.1101/gr.224576.117> PMID: 29273625
20. Nozawa R-S, Boteva L, Soares DC, Naughton C, Dun AR, Buckle A, et al. SAF-A regulates interphase chromosome structure through oligomerization with chromatin-associated RNAs. *Cell*. 2017; 169: 1214–1227.e18. <https://doi.org/10.1016/j.cell.2017.05.029> PMID: 28622508
21. Tan S. A modular polycistronic expression system for overexpressing protein complexes in *Escherichia coli*. *Protein Expr Purif*. 2001; 21: 224–234. <https://doi.org/10.1006/prep.2000.1363> PMID: 11162410
22. Sano K, Miyaji-Yamaguchi M, Tsutsui KM, Tsutsui K. Topoisomerase II β activates a subset of neuronal genes that are repressed in AT-rich genomic environment. *PLoS One*. 2008; 3: e4103. <https://doi.org/10.1371/journal.pone.0004103> PMID: 19116664

23. Tsutsui KM, Sano K, Hosoya O, Miyamoto T, Tsutsui K. Nuclear protein LEDGF/p75 recognizes supercoiled DNA by a novel DNA-binding domain. *Nucleic Acids Res.* 2011; 39: 5067–5081. <https://doi.org/10.1093/nar/gkr088> PMID: 21345933
24. Gasser SM, Laemmli UK. Cohabitation of scaffold binding regions with upstream/enhancer elements of three developmentally regulated genes of *D. melanogaster*. *Cell.* 1986; 46: 521–530. [https://doi.org/10.1016/0092-8674\(86\)90877-9](https://doi.org/10.1016/0092-8674(86)90877-9) PMID: 3015418
25. Izumi H, Funa K. Telomere Function and the G-quadruplex formation are regulated by hnRNP U. *Cells.* 2019;8. <https://doi.org/10.3390/cells8050390> PMID: 31035352
26. Sengar A, Heddi B, Phan AT. Formation of G-quadruplexes in poly-G sequences: structure of a propeller-type parallel-stranded G-quadruplex formed by a G₁ stretch. *Biochemistry.* 2014; 53: 7718–7723. <https://doi.org/10.1021/bi500990v> PMID: 25375976
27. Kwok CK, Marsico G, Sahakyan AB, Chambers VS, Balasubramanian S. rG4-seq reveals widespread formation of G-quadruplex structures in the human transcriptome. *Nat Methods.* 2016; 13: 841–844. <https://doi.org/10.1038/nmeth.3965> PMID: 27571552
28. Käs E, Izaurralde E, Laemmli UK. Specific inhibition of DNA binding to nuclear scaffolds and histone H1 by distamycin. The role of oligo(dA).oligo(dT) tracts. *J Mol Biol.* 1989; 210: 587–599. [https://doi.org/10.1016/0022-2836\(89\)90134-4](https://doi.org/10.1016/0022-2836(89)90134-4) PMID: 2614835
29. Coll M, Aymami J, van der Marel GA, van Boom JH, Rich A, Wang AH. Molecular structure of the netropsin-d(CGCGATATCGCG) complex: DNA conformation in an alternating AT segment. *Biochemistry.* 1989; 28: 310–320. <https://doi.org/10.1021/bi00427a042> PMID: 2539859
30. Rohs R, West SM, Sosinsky A, Liu P, Mann RS, Honig B. The role of DNA shape in protein-DNA recognition. *Nature.* 2009; 461: 1248–1253. <https://doi.org/10.1038/nature08473> PMID: 19865164
31. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, et al. Highly accurate protein structure prediction with AlphaFold. *Nature.* 2021; 596: 583–589. <https://doi.org/10.1038/s41586-021-03819-2> PMID: 34265844
32. Taniura H, Yoshikawa K. Necdin interacts with the ribonucleoprotein hnRNP U in the nuclear matrix. *J Cell Biochem.* 2002; 84: 545–555. PMID: 11813259
33. Helbig R, Fackelmayer FO. Scaffold attachment factor A (SAF-A) is concentrated in inactive X chromosome territories through its RGG domain. *Chromosoma.* 2003; 112: 173–182. <https://doi.org/10.1007/s00412-003-0258-0> PMID: 14608463
34. Hasegawa Y, Brockdorff N, Kawano S, Tsutsui KM, Tsutsui K, Nakagawa S. The matrix protein hnRNP U is required for chromosomal localization of Xist RNA. *Dev Cell.* 2010; 19: 469–476. <https://doi.org/10.1016/j.devcel.2010.08.006> PMID: 20833368
35. Hall LL, Carone DM, Gomez A V, Kolpa HJ, Byron M, Mehta N, et al. Stable C₀T-1 repeat RNA is abundant and is associated with euchromatic interphase chromosomes. *Cell.* 2014; 156: 907–919. <https://doi.org/10.1016/j.cell.2014.01.042> PMID: 24581492