

Prediction of tertiary structure of NSSRs' RNA recognition motif and the RNA binding activity

To cite this article: Kazuo Fushimi *et al* 2005 *Sci. Technol. Adv. Mater.* **6** 475

View the [article online](#) for updates and enhancements.

You may also like

- [Wideband planar retro-reflective metasurfaces for backscattering enhancement under oblique incidence](#)
Yuxiang Jia, Jiafu Wang, Jie Yang et al.
- [Performing broadband and tunable mathematical operations based on acoustic reconfigurable metasurfaces](#)
Shuyu Zuo, Chengxin Cai, Xiaojun Li et al.
- [Bending and free vibration analysis of nanobeams by differential and integral forms of nonlocal strain gradient with Rayleigh–Ritz method](#)
Mahmood Fagher and Shahrokh Hosseini-Hashemi



Prediction of tertiary structure of NSSRs' RNA recognition motif and the RNA binding activity

Kazuo Fushimi^{a,b}, Shizuka Uchida^a, Ryoussuke Matsushita^a, Toshifumi Tsukahara^{a,b,*}

^aCenter for Nano Materials and Technology, Japan Advanced Institute of Science and Technology, 1-1, Asahidai, Tatsunokuchi, Ishikawa 923-1292, Japan

^bCore Research for Evolutional Science and Technology, Japan Science and Technology Agency, 4-1-8, Honcho, Kawaguchi, Saitama 332-0012, Japan

Received 12 November 2004; revised 2 February 2005; accepted 2 February 2005

Available online 28 June 2005

Abstract

RNAs possess potentials to become excellent bio-material because of their biochemical and biological activities. For instance, most RNA splicings are catalyzed by machinery including their own RNAs or other RNAs. The eukaryote machineries for splicing of pre-mRNA, which are called spliceosomes, are flexible and accurate for separating substrates. Although RNAs themselves catalyze the splicing, spliceosomes are supported by many proteins. Furthermore, a great accuracy is required for the alternative splicing because there are choices available, which must be regulated in tissue-specific and developmental manners. Neural-salient serine/arginine-rich (NSSR) proteins 1 and 2 are candidates for supporting the accuracy of the splicing. The features of their amino acid sequences suggest that NSSRs are SR proteins, which bind to pre-mRNA and determine the splicing site. Since SR proteins have a RNA recognition motif or motives (RRM or RRM), which binds to RNA, we predicted the secondary and tertiary structures of NSSRs' RRM by comparing them to RRM of other proteins. The predicted structure suggested that the RNA binding activity of NSSRs' RRM is similar to the poly A binding protein (PABP). Moreover, to detect the targets for NSSR, mRNAs were obtained by screening them from murine brains with bacterial recombinant NSSRs' RRM and microarray experiments were conducted using these mRNAs. The results suggested that NSSRs bind specifically to particular pre-mRNAs and regulate the alternative splicing of the binding pre-mRNA.

© 2005 Elsevier Ltd. All rights reserved.

Keywords: Splicing; SR protein; Brain; Pre-mRNA; Tertiary structure; RNA recognition motif (RRM); cDNA microarray

1. Introduction

Some RNAs have catalytic and biological activities; therefore, it is valuable to study them and to explore their applications for various fields, including material developments. Among these RNAs, self splicing is one of their catalytic activities. For example, many mRNAs in eukaryotic cells are produced when introns are spliced out from the pre-mRNAs, which are immediate copies of genomic DNA by transcription. This splicing is not an autocatalytic reaction, but catalytically active RNA components splice the pre-mRNA [1]. In this mechanism, many proteins cooperate to activate the catalytic activity. Among these proteins, SR proteins are included. The characteristics of the amino acid sequences

of SR proteins are the presence of the RNA recognition motif/motifs (RRM/RRMs) at the N-terminal and the serin-arginine dipeptide sequence at the C-terminal. They recognize exons and bring the splicing machinery to the correct position to perform splicing at exact sites [2]. In addition, more than one mRNAs are produced from a single gene, due to different choices of exons on the gene in metazoan, which is known as alternative splicing. This alternative splicing is regulated in a tissue and developmental specific manner in living organisms. SR proteins also contribute to regulation of alternative splicing.

We previously reported that two genes for SR protein like polypeptides, neural salient serine/arginine-rich 1 and 2 (NSSRs 1 and 2), were cloned from a subtraction library of neuroectodermal differentiated P19 cells by degenerative PCR [3]. To characterize these genes, the conserved domain search was performed, which suggested the presence of an RRM in the N-terminal region of the products. Furthermore, their C-terminals consist of SR repetitive sequences.

* Corresponding author. Tel.: +81 761 51 1440; fax: +81 761 51 1149.
E-mail address: tukahara@jaist.ac.jp (T. Tsukahara).

However, there exists a difference between them; that is, NSSR2 is shorter than NSSR1 due to alternative splicing. In the expression levels, mRNAs for NSSR 1 and 2 are present at higher levels in the brain and testis than in other tissues. Unlike NSSR2, which is expressed before the differentiation of the cells, NSSR1 is expressed in the neural stage during the neuroectodermal differentiation of embryonal carcinoma cells, P19, suggesting that it is engaged in the production of neuronal splicing products. In addition, the homology search for NSSRs revealed that there exist high similarities with many RNA binding proteins, including SR proteins. Currently, it is known that the overexpression of NSSRs regulate alternative splicing of GluR-B gene [3]; however, it is not yet clear how the RRM of NSSRs contributes to the regulation of its splicing activities.

In recent years, the amount of protein structures analyzed by NMR and X-ray crystallography has increased tremendously. Among them, RRM is one of the well characterized polypeptide structures, which consists of four β -sheets strands and α -helix in between the strands ($\beta/\alpha/\beta/\beta/\alpha/\beta$) [4–10]. A typical RRM is 90–100 amino acid long with two conserved sequence motifs, RNP-1 [(K/R)-G-(F/Y)-(G/A)-F-V-x-(F/Y)] and RNP-2 [(L/I)-(F/Y)-(V/L)-(G/K)-(G/N)-(L/M)] [6]. Aromatic amino acids on RNP-1 and 2 interact with bases of their target RNAs by stacking on the surface of their RRM. Their RNA binding activity is supported primarily on their β sheet surface.

Microarrays are a very useful tool to analyze relative amounts of mRNA in a high-throughput manner. They can be applied for analyses for mRNA expression to understand biological phenomena, diseases or diagnoses. Their basic procedural scheme is that nucleic acids are labeled by fluorescent dyes, such as Cy3 and Cy5, and the labeled nucleic acids are hybridized with spotted probes on the microarray slide to give fluorescent intensities, which represent the amount of nucleic acids contained in the cells.

In this paper, we predicted the structure of the RRM in NSSRs and showed its possibility of binding to an RNA. Furthermore, using combination of recombinant technology and microarray analysis, we successfully screened murine brain mRNAs to find NSSRs' targets.

2. Experimental procedure

2.1. Structure analyses of NSSRs' RRM

The comparison of amino acid sequences was performed using conserved domain database search at NCBI (<http://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml>) [11]. For the prediction of the secondary structure of NSSRs' RRM, PSIPRED (<http://bioinf.cs.ucl.ac.uk/psipred/psiform.html>) was utilized [12]. The tertiary structure of NSSRs' RRM was computed through SWISS MODEL (<http://swissmodel.expasy.org/SWISS-MODEL.html>) [13]. The structures

were visualized with Swiss PDB Viewer (<http://kr.expasy.org/>)

2.2. Production and extraction of the NSSRs' RRM recombinant protein

To produce an NSSRs' RRM recombinant protein, the DNA fragments that code the RRM were amplified by high fidelity PCR (pyrobest, TAKARA) using primers, RRMF 5'-ctggaattcccatgtcccatactgcgcc-3' and RRMR 5'-ccgaattctatatcggtcataatcgtcata-3'. These fragments were subcloned into the *EcoRI* site of pYesTrp3. The obtained plasmids were digested with *HindIII* and *XhoI* to insert into pET-34b(+), which carries a cellulose binding domain tag (CBD tag, Novagen).

The NSSRs' RRM recombinant protein was produced in the 100 ml culture of *E. coli*, Rosetta(DE3)pLysS, by induction with 1 mM IPTG in LB medium for three hours at 37 °C. The bacterial pellet was frozen once, resuspended with PBS containing the protease inhibitor cocktail (NACALAI TESQUE), and then lysed by sonication though the treatment on ice (VIBRA CELL, SONIC and MATERIAL). The lysate was concentrated by ultrafiltration using Amicon Ultra-4 (MWCO 10,000, MILLIPOR), and then applied onto cellulose resin to detect the RNA binding activity of NSSR's RRM-CBD fusion protein.

2.3. Messenger RNA binding assay

Total RNA was extracted from the bulk of murine brains obtained from 5-day-old to adult mice using TRIzol (Invitrogen) followed by purification using mTRAP total kit (ACTIVE MOTIF). The purified recombinant NSSRs' RRM-CBD fusion protein was mixed with mRNA fraction, then absorbed on 5 mg of cellulose resin (CBinD 100, Novagen). After four washes with 1xSSC containing 0.01% TRITON X-102 and 0.2 units/ μ l SUPERase \cdot In (Ambion), the resin was resuspended in 1000 μ l of the same buffer, and then 20 μ g of the murine brain mRNA were added. After incubation for 10 min, the resin was washed four times with the same buffer. Next, the resin was resuspended with 100 μ l of 0.5% SDS, and then 500 μ l of TRIzol were added into the resin to purify the mRNA on the resin.

2.4. SDS-PAGE and immunoblot analysis

SDS-PAGE was performed as described by Laemmli [14]. The proteins fractionated by SDS-PAGE were stained with CBB (Rapid Stain CBB kit, NACALAI TESQUE) or transferred onto PVDF membranes. Anti-CBD-Tag rabbit polyclonal antibody (Novagen) was used as a primary antibody (\times 5000 dilution), and alkaline phosphatase linked anti-rabbit IgG (Chemicon) was used as secondary antibody (\times 10,000 dilution). The CBD-tagged proteins were detected using BICP-NBT solution kit (NACALAI TESQUE).

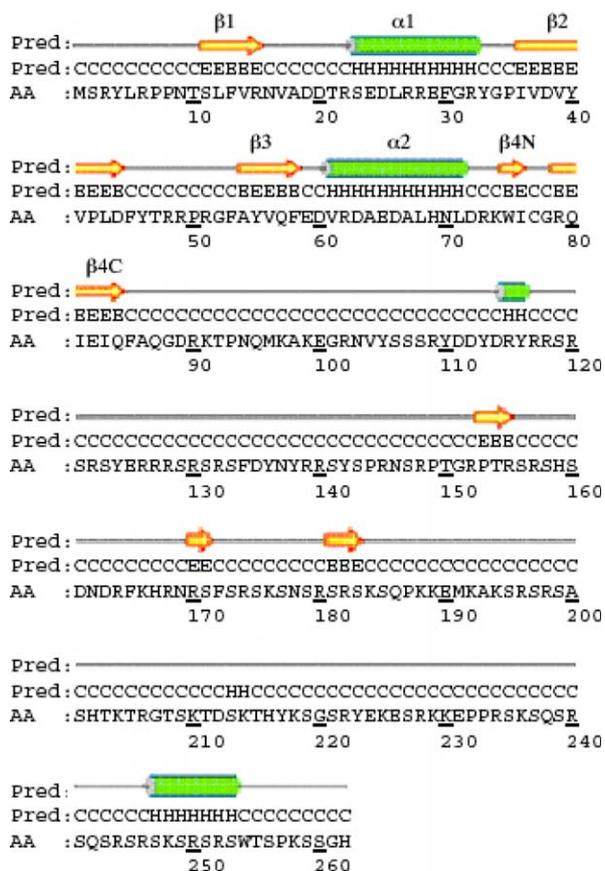


Fig. 1. Predicted secondary structure of NSSR. The predicted secondary structure of NSSR 1 was produced from its full primary sequence by the PSIPRED program using its default settings. E and yellow arrows indicate β -strands, and α -helices are shown by H and green columns. C indicates possibility of random coils. From S11 to Q84, $\beta/\alpha/\beta/\alpha/\beta$ alignment was observed for NSSR1 as well as typical RRM. In the C-terminal of NSSRs' RRM, two β -strands were predicted (β 4N and β 4C) (For interpretation of the reference to color in this legend, the reader is referred to the web version of this article.).

2.5. Microarray analysis

The procedure for the microarray experiment was followed as described in the products manual. Briefly, the purified RNA was reverse transcribed using the T7 tagged oligo-dT primer and then, labeled RNA with Cy3 or Cy5 (Low RNA Input Fluorescent Linear Amplification Kit, Agilent). These labeled cRNAs were hybridized onto Mouse Oligo Microarray Kit (Agilent). Then, the hybridized microarray was scanned using CRBIO Iie (Hitachi Software Engineering). The scanned image was analyzed using ScanAnalyze version 2.50 (<http://rana.lbl.gov/EisenSoftware.htm>).

3. Results and discussion

3.1. Secondary structure analyses

The secondary structure of NSSR 1 was predicted using PSIPRED. As a result, α -helix structures were suggested in positions from 23 to 32 and from 61 to 71, while residues next to both helices at the C- and N-terminals were likely to form β -strand structures (Fig. 1). Because an RRM forms a $\beta/\alpha/\beta/\alpha/\beta$ structure from its N-terminal and is about 100 amino acids long, the prediction by PSIPRED suggested that NSSRs have a similar structure to other RRM (Fig. 2). The sequence features of NSSRs were further explored by comparison with other proteins. The following proteins were selected for comparison: solution structures of U2AF65 (accession; 1U2FA), hnRNP D0 RBD 1 (accession; 1HD0) and 2 (accession; 1IQT), and mushashi1 RBD 1 (accession; 2MSS), all of which were analyzed by NMR [7–10]. Their structures include four-stranded antiparallel β -sheets backed by two α -helices, which are characteristics of RRM. The comparison among the NSSRs' RRM and

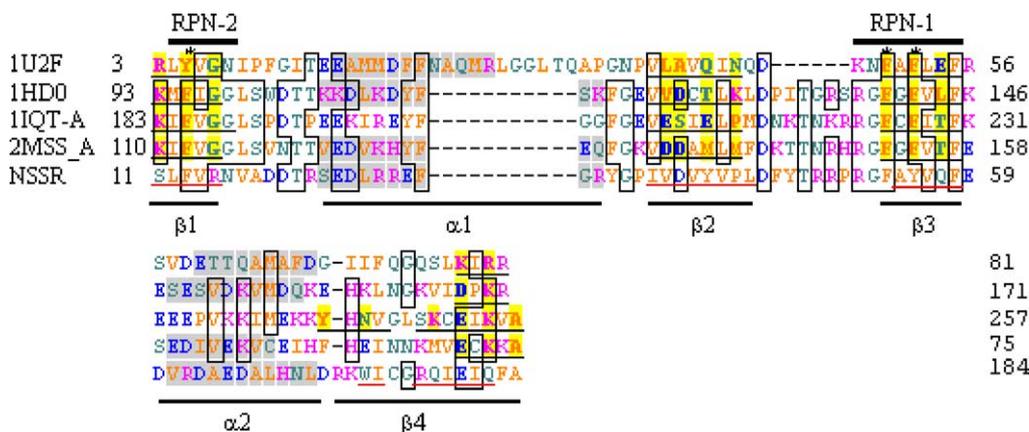


Fig. 2. Comparison of primary structure of NSSRs' RRM with other RRM characterized by NMR. The RRM, whose structures were determined by NMR, were aligned with NSSRs' RRM. Gray shaded regions and underlined regions form α -helices and β -strands, respectively. Yellow shaded parts indicate amino acids whose exposed solvent surrounds the RRM's β -sheets. RNP-1 and 2 motifs are conserved on third and first β -strands, respectively, and aromatic amino acids indicated by asterisks are exposed to the outside of the proteins (For interpretation of the reference to color in this legend, the reader is referred to the web version of this article.).



Fig. 3. Comparison of primary structure of NSSRs' RRM with PABP domains 1 and 2. To construct the tertiary structure of NSSRs' RRM, PABP domain 1 was chosen as a model. The primary sequences of NSSRs' RRM, PABP domains 1 and 2 were aligned to compare their amino acids that are contributing to the RNA binding. Asterisks indicate aromatic amino acids conserved in RNP-1 and 2. Residues of amino acids shaded by gray support RNAs. The yellow shaded amino acid backbone participates in holding RNA. Amino acids contributing to highly negative potential on the surface of NSSRs' RRM are shaded by green (For interpretation of the reference to color in this legend, the reader is referred to the web version of this article.).

above mentioned proteins revealed that the amino acid sequences S11, F13 and R15 of the first β -strand from N-terminal (β 1), V37, D38, Y40 and P42 of the second β -strand from N-terminal (β 2), and F53, Y55 and Q57 of the third β -strand from N-terminal (β 3) in the NSSRs' RRM are likely to be exposed to the solvent that surrounds them. β 1 and 3 were similar to the corresponding β -sheets of the NSSRs' RRM predicted by PSIPRED, which are reflecting RNP-1 and 2 motifs in β 3 and 1 of RRM, respectively. The third amino acids on β 1 and β 3 of RRM are usually aromatic amino acids (F or Y) that are on the surface of the proteins and interact with bases of the target nucleic acids by stacking. The conservation of aromatic amino acids at the 13 and 55th positions from the N-terminal of the NSSRs' RRM (F13 and Y55) suggested that these amino acids are also located on the surface of the protein and stack with the bases of the target nucleic acids. However, β 2 and β 4 (the fourth β -strand from N-terminal) were not well-conserved among the compared RRM, especially, the β 4 of hnRNP D0 RRM2 has an unusual structure [8]. Although only one β -strand follows after β 2 in a typical RRM, there are two possible regions to form β -strands in hnRNP D0 RRM2; in which, the N-terminal region (YHNV) forms an additional antiparallel strand (β 4) with β 4 consisting of the C-terminal region (SKCEIKVA). According to the prediction by PSIPRED, a highly possible region to form β -strands next to C-terminal of α 2 is divided into two by a random coil as well as hnRNP D0 RRM2 (Fig. 1). Therefore, the NSSRs' RRM may fold its structure in a similar manner as hnRNP D0 RRM2 consisting of five strands, β 1–4 and 4-does.

3.2. Analysis of the tertiary structure of the NSSRs' RRM

The tertiary structure of the NSSRs' RRM was built using SWISS-MODEL. The structure of the domain 1 in human poly A binding protein chain A (accession; 1CVJ_A) was used as a template [5]. As a result, the identity of the amino acid sequences between NSSRs' RRM and 1CVJ_A was 34.6% (Fig. 3), and the total final energy for NSSRs' RRM was -5046.952 KJ/mol. From these results, NSSRs' RRM may have a similar structure to a typical RRM (Fig. 4A). In Fig. 4C, F13, F53 and Y55 in NSSRs' RRM, which correspond to the solvent exposed aromatic amino

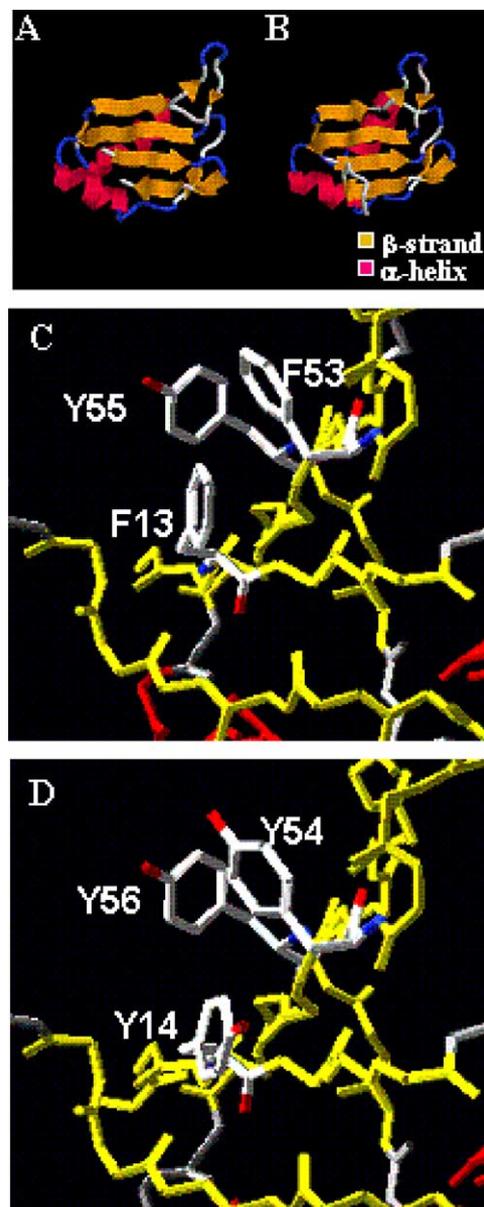


Fig. 4. The tertiary structure of NSSRs' RRM. The tertiary structure of NSSRs' RRM was constructed through SWISS MODEL. The cartoons of the structures for NSSRs' RRM and PABP domain 1 are shown in (A) and (B), respectively. The aromatic amino acids conserved in RNP-1 (F53 and Y55) and RNP-2 (F13) of NSSRs were exposed to the solvent surrounding them, (C), as well as those of PABP domain 1, (D)

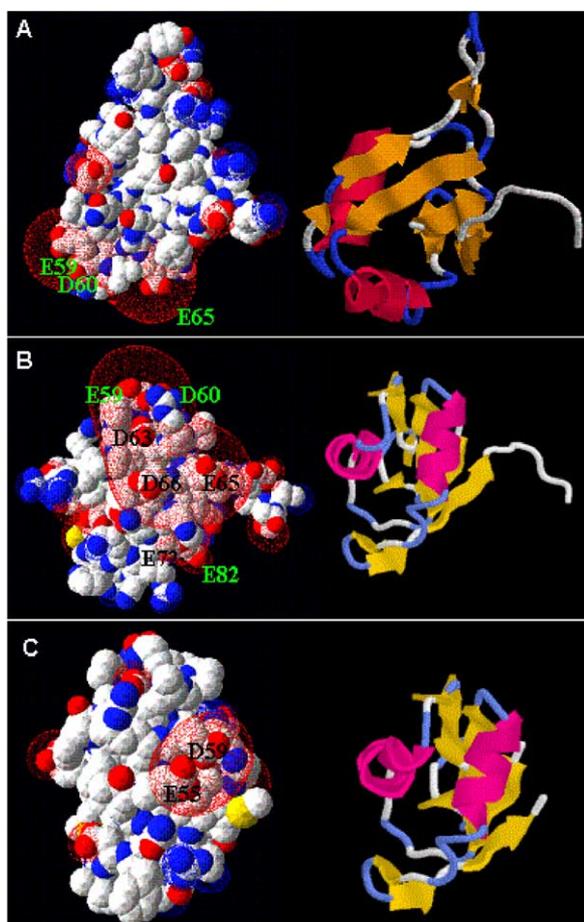


Fig. 5. Electrostatic potentials of NSSRs' RRM. Electrostatic potentials of NSSR's RRM and PABP domain 1 were computed for NSSR (A and B) and PABP (C). The second α -helix region of NSSRs' RRM was highly negatively charged (A and B), but only two negatively charged amino acids on the surface of same region of PABP were observed (C).

acids in the RNP-1 and 2 of PABP (Y14, Y54 and Y56 in Fig. 4D), were faced on the outside of the β -sheet. The secondary structure suggest that the RRM of NSSRs has the additional β -strand at the C-terminal as well as the hnRNP D0 domain 1 structure. The upstream short strand of the β 4 region (KWI) built a β -sheet with the next β -strand (RQIEIQFA) which also built the β -sheet structure with β 1. This structure was also observed in PABP domain 1. The tertiary structure analysis suggests that the NSSRs' RRM polypeptide is fold into a typical RRM structure in coincidence with the secondary structure prediction other than the additional β -strand next to β 4 (Fig. 4C).

The electrostatic potentials of NSSRs' RRM was calculated and shown in Fig. 5. The highly negative region was found around the second α -helix of NSSR's RRM (Fig. 5A and B), while PABP domain 1 did not have such a large negative area around its second helix region (Fig. 5C). Furthermore, the highly negative region around the second α -helix was absent in other RRM (Figs. 2 and 3). Most RRM recognize RNA on the surface of their β -sheets. Therefore, this unique feature of NSSRs' RRM suggested

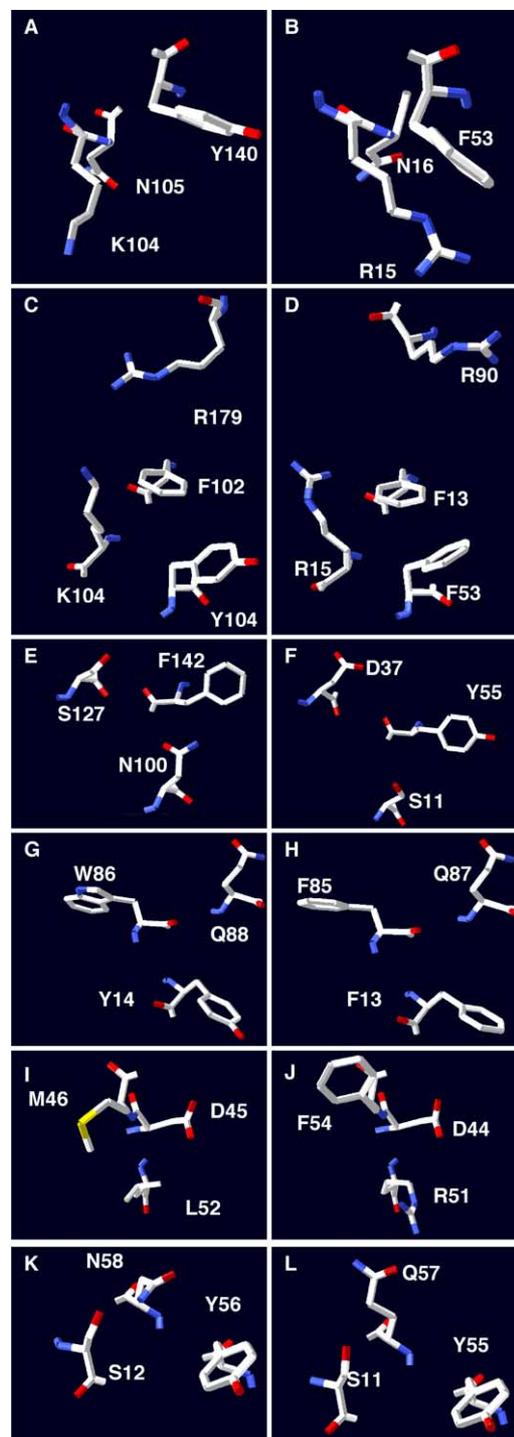


Fig. 6. The details of the tertiary structure of NSSRs' RRM. Amino acids contributing to holding an RNA in PABP (A, C, E, G, I and K) and their positions were well conserved in NSSRs' RRM (B, D, F, H, J and L). N105 and Y140, K104, N100 and S127, backbone of W86 and Q88, D45, and S12 and N58 in PABP are specified adenine, Ade2 (A), Ade3 (C), Ade4 (E), Ade6 (G), Ade7 (I) and Ade8 (K), respectively.

that there exists a specific interaction with certain positively charged proteins or repulsion of nucleic acids interacting with their β -sheets. Since the surface of the β -sheet in NSSRs' RRM was neutral as well as that of PABP

domain 1, the β -sheet in NSSRs' RRM may interact with bases of RNA by aliphatic and van der Waals interaction, base stacking, and hydrogen bond.

Next, the spatial positions of the amino acids which interact with RNA in PABP domains 1 and 2 were compared with those of equivalent NSSRs' amino acids. In the case of PABP domains 1 and 2, phosphates of RNA interact with K104, R89, Y14, Y54, and Y140 (Fig. 3). Among the equivalent positions of NSSRs' RRM, only R15, which corresponds to K104 in the domain 2 of PABP, potentially interacted with phosphate of RNAs. Therefore, the manner in which the residue supports the phosphate of an RNA backbone seemed to be different between NSSR and PABP. In respect of amino acids recognizing adenines, counterparts of NSSRs to the amino acids supporting Ade2 and Ade6 in PABP showed considerable similarity suggesting NSSRs hold adenines at those positions (Fig. 6 A, B, G, H, Table 1). Totally, the amino acids, which stack with adenines in PABP, were well conserved, but most of the counterparts for the amino acids specified adenines in NSSRs may not recognize adenines (Fig. 6, Table 1). The equivalent amino acids of NSSRs to amino acids interact with Ade5 were unclear, because the amino acids straddle both of RRM domains 1 and 2 in PABP.

Table 1
Comparison of amino acids responsible to recognize adenines in PABP with NSSR

Position	PABP	Interaction	Specifying	NSSR
Ade2	K104	van der Walls ^a		R15
	N105	O δ -N6	adenine	N16
	Y140	C δ 1-C2	Purine	F53
Ade3	Y140	OH-O1P		F53
	F102	Stacking		F13
	R179	Stacking		R90 ^b
	K104	N ζ -N6	Adenine	R15
Ade4	N100	N δ -N7	Purine	S11
	S127	O-N6 and O γ -N1 ^c	Adenine	D37
	F142	Stacking		Y55
Ade6	W86	O-N6 ^c	Adenine	F85
	Q88	N-N1 ³	Adenine	Q87
	Y14	Stacking		F13
Ade7	L52	Stacking		R51
	M46	C ϵ -C2	Purine	F45
	D45	O δ 1-N6	Adenine	D44
Ade8	S12	O γ -N6	Adenine	S11
	Y56	Stacking		Y55
	N58	N δ 2-N1, O- δ N6 ^c	Adenine	Q59

Candidates to interact with bases were predicted by the tertiary structures. The amino acids, which stack with adenines in PABP, were well conserved in NSSRs. The amino acids shown in bold potentially support adenines in a similar manner to PABP.

^a The aliphatic region of K104 interacts with the adenine face.

^b R90 is not the exact counterpart of R179 in PABP, but located in similar position.

^c The amino acid's backbone interacts with the adenine.

3.3. RNA Binding activity of NSSRs' RRM

The structural observation prompted us to explore the RNA binding activity of NSSR and the targets for NSSR. Generally, SR proteins recognize exons of pre-mRNA to include the selected exons into the spliced mRNA. By using this mechanism, it is possible to screen for the target of the specific splicing. Therefore, we tried to identify target RNAs of NSSRs' RRM. The DNA fragment coding RRM of NSSR was fused with CBD to produce the recombinant protein. The NSSR's RRM-CBD fusion protein was obtained from *E. coli* lysate, and absorbed on 5 μ g of cellulose resin. We confirmed that 2.6 μ g of the purified fusion protein were absorbed on the resin. As shown in Fig. 7A, a band at about 46 kDa was detected as a major protein on the resin. In addition, it was reactive to the anti-CBD antibody (Fig. 7B). These results suggested that the CBD fused NSSRs' RRM was successfully produced and trapped on the cellulose resin.

The mRNAs from murine brains were added onto the immobilized NSSRs' RRM-CBD fusion protein/cellulose column and washed four times with the binding buffer to eliminate unbound mRNAs. Next, the trapped mRNAs were extracted by denaturation. The recovered mRNAs were amplified and labeled with Cy5. The untreated murine brain mRNAs were labeled with Cy3 as a reference. Both labeled cRNAs were hybridized onto the mouse oligo microarray, and scanned with CRBIO IIe. The scanned image was analyzed by ScanAlyze.

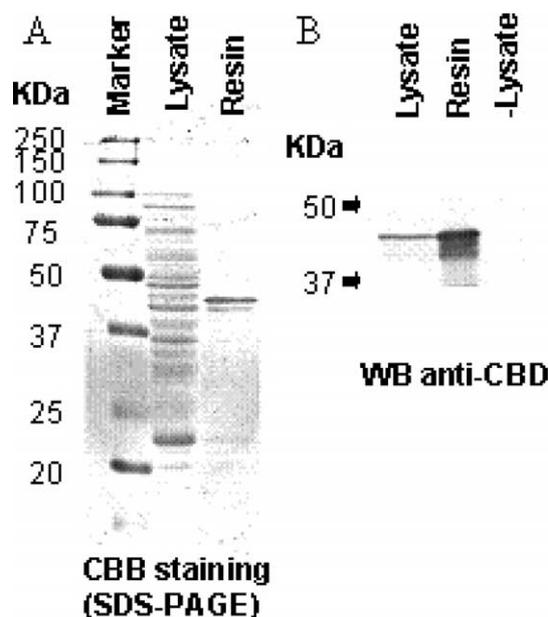


Fig. 7. Production of CBD tagged NSSRs' RRM and absorption onto cellulose. The CBD-tagged NSSRs' RRM was produced by *E. coli*, purified by using a cellulose column and analyzed with SDS-PAGE (A) and immunoblotting (B). The *E. coli* lysate contained the CBD anti-body reactive protein at 46 kDa (B, the left lane). It was confirmed that the recombinant proteins were successfully pulled-down with 5 μ g of CBInD resin (A, right lane and B, middle lane).

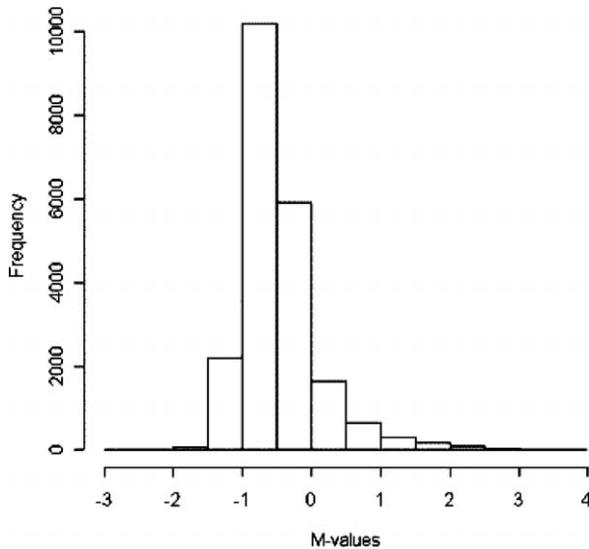


Fig. 8. The histogram for the M -values of the microarray data. The M -value for each spot was calculated using the following formula: $M = \log_2(\text{CH2}/\text{CH1})$, where both CH1 and CH2 are the background-subtracted intensity values for Channels 1 and 2, respectively. The histogram was drawn based on the M -values to observe the overall tendency of the microarray data. The x -axis indicates the M -values, and the y -axis shows the frequency distribution. As it can be observed from the figure, there is a tendency of the distribution of the M -values towards higher M -values (compare the frequencies around $M = -2$ and 1 and higher). This tendency corresponds to the fact that some gene transcripts are concentrated due to the attachment of the recombinant NSSRs' RRM.

Fig. 8 shows the distribution of the expression ratio for the microarray data. Channel 1 was the Cy5 fluorescence intensities for the mRNA from NSSRs' RRM trapped on the resin, and Channel 2 was the Cy3 fluorescence intensities for the reference murine brain mRNA. A single peak was observed, but the distribution was skewed towards the higher ratio than towards the lower, which might suggest that some mRNAs were concentrated for NSSRs' RRM trapped on the resin. The most concentrated RNA was the clone IMAGE: 3584181 (13-fold expression level increase compared to the reference), and 63 other RNAs were concentrated more than 4.9 times (Table 2). These RNAs coded many types of proteins, including structure proteins (e.g. neurofilament-L), kinases (e.g. Mapk12), secretory proteins (e.g. Chromogranin B) and RNA-binding proteins (e.g. mHuC-L). Many hypothetical genes and pseudogenes were also found among them. Chromogranin B was listed many times therefore the microarray analysis was likely to be reliable. The Database searches for these genes through Ensembl Genome Browser (<http://www.ensembl.org/>) revealed that 27.5% of the identified genes had alternative splicing isoforms. Although NSSRs were expected as an alternative splicing regulator, there were many genes, which were not suggested as splicing isoforms in the list. Therefore, the splicing isoforms of these genes may be unidentified so far. Another possibility is that NSSRs participate in not only alternative splicing but also constitutive splicing in the brain.

Table 2
Highly concentrated RNAs by NSSRs' RRM absorbed resin

Accession no.	Description	CH1	CH2	Ratio	Products
BC018347	Similar to translation initiation factor IF2, clone IMAGE: 3584181	2018	26180	13.0	1
BC004015	clone MGC: 7593 IMAGE: 3493893	2715	24,295	8.9	1
BC023025	clone IMAGE:5362343	1868	16,519	8.8	–
AJ278123	MRNA for putative synaptopodin	5852	49,917	8.5	2
AK046243	Hypothetical LIM domain, Villain headpiece domain containing protein	5157	43,490	8.4	–
M20480	Brain neurofilament-L	5602	46,434	8.3	1
NM_029842	Similar to THYRO1000124 PROTEIN	4340	35,554	8.2	1
AK009959	Ankyrin-like repeat protein	4811	39,116	8.1	1
NM_010123	Eukaryotic translation initiation factor 3 (Eif3)	1531	11,798	7.7	1
NM_008527	Killer cell lectin-like receptor subfamily B member 1C (K1rb1c)	3792	28,994	7.6	3
AK005564	Similar to 60S RIBOSOMAL PROTEIN L17 (L23)	6515	46,949	7.2	1
BC004722	clone IMAGE:3582796	2268	15,787	7.0	–
AK046455	Weakly similar to RIBOSOMAL PROTEIN L18A HOMOLOGUE	2155	14,759	6.8	–
BC047049	Similar to UPF3 regulator of nonsense transcripts homolog B	1608	10,922	6.8	2
NM_010880	Nucleolin (Ncl), mRNA	1796	12,188	6.8	1
BC010811	clone MGC:19,122 IMAGE: 4210,911	2074	14,073	6.8	–
NM_007694	Chromogranin B (Chgb)	7926	53,452	6.7	1
NM_027349	SIMILAR TO S164 PROTEIN homolog	1768	11,752	6.6	3
NM_007694	Chromogranin B (Chgb)	8047	52,766	6.6	–
NM_028071	Coactosin-like protein (Clp-pending)	7832	51,071	6.5	1
NM_013871	Mitogen-activated protein kinase 12 (Mapk 12)	3225	20,954	6.5	1
NM_023153	RIKEN cDNA 0610040D20 gene (0610040D20Rik)	6668	43,229	6.5	1
NM_007694	Chromogranin B (Chgb)	8401	54,294	6.5	–
NM_007786	Casein kappa (Csnk)	2327	14,940	6.4	1
NM_007684	Centrin 3 (Cetn 3)	2487	15,797	6.4	1
NG_001389	Hmgb1-rs13 pseudogene on chromosome X	1641	10,273	6.3	–

(continued on next page)

Table 2 (continued)

Accession no.	Description	CH1	CH2	Ratio	Products
AK088007	PEPTIDYLPROLYL ISOMERASE MATRIN CYP (EC 5.2.1.8) homolog	1360	8470	6.2	1
NM_007694	Chromogranin B (Chgb)	8548	52,940	6.2	
NM_007694	Chromogranin B (Chgb)	8877	54,896	6.2	
NM_007694	Chromogranin B (Chgb)	8313	51,087	6.1	
NM_007694	Chromogranin B (Chgb)	9177	55,450	6.0	
NM_025282	Myocyte enhancer factor 2C (Mef2c0)	3130	18,663	6.0	3
NM_007694	Chromogranin B (Chgb)	8645	51,415	5.9	
NM_007694	Chromogranin B (Chgb)	9052	53,790	5.9	
NM_007694	Chromogranin B (Chgb)	9026	53,630	5.9	
AK015949	CHROMODOMAIN HELICASE DNA BINDING PROTEIN CHD MI 2	1326	7850	5.9	1
BQ044583	cDNA clone IMAGE: 5685698 5	1401	8214	5.9	–
NM_010871	Baculoviral IAP repeat-containing if (Birc 1f)	2537	14,807	5.8	1
NG_001388	Hmgb1-rs 12 pseudogene on chromosome 18	1653	9633	5.8	
AK019393	Hypothetical Lysine-rich region containing protein	1428	8238	5.8	1
BC022717	clone IMAGE: 4501335	1963	1131	5.8	–
AK077741	ABC50 (FRAGMENT) homolog	2134	12,182	5.7	–
AK048208	Hypothetical RPEL repeat containing protein	4407	24,714	5.6	2
NM_025663	RIKEN cDNA 2610029K21 gene (2610029K21Rik)	2105	11,731	5.6	3
NM_013561	5-hydroxytryptamine (serotonin) receptor 3A	2269	12,458	5.5	1
NM_008989	Purine rich element binding protein A (pura)	1334	7292	5.5	1
NM_011856	Odd Oz/ten-m homolog 2 (Drosophila) (Odz2)	1892	10,326	5.5	1
NG_001965	Olfactory receptor MOR 175-6 pseudogene	1815	9901	5.5	
AK122230	mRNA for mKIAA0244 protein	1241	6709	5.4	3
U29148	RNA-binding protein mHuC-L	2876	15,333	5.3	1
BC027561	Similar to ubiquitin associated and SH3 domain containing	2368	12,584	5.3	–
NM_007926	Endothelial monocyte activating polypeptide 2 (Emap2)	5488	29,075	5.3	1
AK078225	Similar to HYPOTHETICAL 139.5 KDA PROTEIN	1156	6119	5.3	2
AK010338	Protein (peptidyl-prolyl cis/trans isomerase) NIMA-interaction, 4 (parvulin)	2024	10,651	5.3	1
NM_025387	RIKEN cDNA 1110021D01 gene (1110021D01Rik)	5632	29,518	5.2	1
AK046782	Hypothetical Acid phosphatase/Vanadium-dependent haloperoxidase structure	3421	17,912	5.2	1
NM_133801	Expressed sequence C76800 (C76800)	3339	17,190	5.1	2
NM_021519	Endothelial differentiation-related factor 1 (Edf1)	9181	46,928	5.1	1
AK044216	Nucleolar protein 5	2300	11,689	5.1	1
AK005378	Actin related protein 2/3 complex, subunit 1B (41 kDa)	1722	8734	5.1	–
NM_009115	S100 protein, beta polypeptide, neural (S100b)	8471	42,709	5.0	1
BC020382	clone IMAGE: 3582855	9360	46,666	5.0	–
NM_008687	Nuclear factor I/B (Nfib)	3837	18,979	4.9	4

CH1, fluorescent intensity by Cy3 from each spot; CH2, fluorescent intensity by Cy5. Ratio, CH2/CH1; products, number of suggested gene transcripts from the EMBL Genome Browser on the – indicates no entry in the database.

4. Conclusions

In this paper, computational methods were applied to predict the secondary and tertiary structures of NSSRs' RRM from its amino acid sequence. Through the homology searches, it was suggested that the RNA binding activity of NSSRs' RRM is similar to that of a typical RRM, PABP. Furthermore, the NSSR's targets were screen from the murine brain by a combination of recombinant technology and microarray analysis. These results suggested that NSSRs contribute to choices of exons for the alternative splicing by directly binding to its target pre-mRNAs.

Acknowledgements

This works was supported by the Center for Nano Materials and Technology, JAIST.

References

- [1] S.J. Melissa, J.M. Melissa, Pre-mRNA splicing: a wash in a sea of proteins, *Mol. Cell* 12 (2003) 5–14.
- [2] B.R. Graveley, Sorting out the complexity of SR protein functions, *RNA* 6 (2000) 1197–1211.
- [3] M. Komatsu, E. Kominami, K. Arahata, T. Tsukahara, Cloning and characterization of two neural-salient serine/arginine-rich (NSSR) proteins involved in the regulation of alternative splicing in neurones, *Genes Cells* 4 (1999) 593–606.
- [4] C. Oubridge, N. Ito, P.R. Evans, C. Teo, K. Ngai, Crystal structure at 1.92 Å resolution of the RNA-binding domain of the U1A spliceosomal protein complexed with an RNA hairpin, *Nature* 372 (1994) 432–438.
- [5] R.C. Deo, J.B. Bonanno, N. Sonenberg, S.K. Burley, Recognition of polyadenylate RNA by the poly(A)- binding protein, *Cell* 98 (1999) 835–845.
- [6] E. Birney, S. Kumar, A.R. Krainer, Analysis of the RNA-recognition motif and RS and RGG domains: conservation in metazoan pre-mRNA splicing factors, *Nucleic Acid. Res.* 21 (1993) 5803–5816.
- [7] T. Nagata, R. Kanno, Y. Kurihara, S. Uesugi, T. Imai, S. Sakakibara, H. Okano, M. Katahira, Structure, backbone dynamics and

- interactions with RNA of the C-terminal RNA binding domain of a mouse neural RNA-binding protein, *Musashi 1*, *J. Mol. Biol.* 287 (1999) 315–330.
- [8] M. Katahira, Y. Miyanori, Y. Enokizono, G. Matsuda, T. Nagata, F. Ishikawa, S. Uesugi, Structure of the C-terminal RNA-binding domain of hnRNP D0 (AUF1), its interactions with RNA and DNA, and change in backbone dynamics upon complex formation with DNA, *J. Mol. Biol.* 311 (2001) 973–988.
- [9] T. Nagata, Y. Kurihara, G. Matsuda, J. Saeki, T. Kohno, Y. Yanagida, F. Ishikawa, S. Uesugi, M. Katahira, Structure and interactions with RNA of the N-terminal UUAG-specific RNA-binding domain of hnRNP D0, *J. Mol. Biol.* 287 (1999) 221–237.
- [10] C.L. Kielkopf, N.A. Rodionova, M.R. Green, S.K. Burley, A novel peptide recognition mode revealed by the X-ray structure of a core U2AF35/U2AF65 heterodimer, *Cell* 106 (2001) 595–605.
- [11] A. Marchler-Bauer, J.B. Anderson, C. DeWeese-Scott, N.D. Fedorova, L.Y. Geer, S. He, D.I. Hurwitz, J.D. Jackson, A.R. Jacobs, C.J. Lanczycki, C.A. Liebert, C. Liu, T. Madej, G.H. Marchler, A.N. Mazumder, A.R. Nikolskaya, B.S. Panchenko, B.A. Rao, V. Shoemaker, J.S. Simonyan, P.A. Song, S. Thiessen, Y. Vasudevan, R.A. Wang, J.J. Yamashita, S.H. Yin, Bryant, CDD: a curated Entrez database of conserved domain alignments, *Nucleic Acid. Res.* 31 (2003) 383–387.
- [12] L.J. McGuffin, K. Bryson, D.T. Jones, The PSIPRED protein structure prediction server, *Bioinformatics* 16 (2000) 404–405.
- [13] T. Schwede, J. Kopp, N. Guex, M.C. Peitsch, SWISS-MODEL: an automated protein homology-modeling server, *Nucleic Acid. Res.* 31 (2003) 3381–3385.
- [14] J. Sambrook, D.W. Russell, *Molecular Cloning*, third ed., Cold Spring Harbor, New York, 2001.