Interaction of hnRNP-C1/C2 proteins with RNA: analysis using the yeast three-hybrid system

Nadejda Koloteva-Levine, Maya Amichay, Orna Elroy-Stein*

Department of Cell Research and Immunology, George S. Wise Faculty of Life Sciences, Tel Aviv University, Tel Aviv 69978, Israel

Received 16 April 2002; revised 13 May 2002; accepted 5 June 2002

First published online 19 June 2002

Edited by Lev Kisselev

Abstract Three-hybrid assays for the analysis of RNA–protein interactions in vivo are usually used, due to technical limitations, only for RNA baits that do not contain runs of four or more consecutive uridines. The present study provides the first example of a three-hybrid analysis of synthetic and natural uridine-rich RNA sequences. The use of the three-hybrid assay enabled us to demonstrate a functional difference between two closely related proteins, heterogeneous nuclear ribonucleoprotein C1 (hnRNP-C1) and hnRNP-C2. The hnRNP-C2 protein, an alternatively spliced variant of hnRNP-C1, contains an additional 13 amino acids between an RNA binding domain (RBD) and a basic leucine zipper-like motif (bZLM), also implied in RNA binding. This study shows that (i) for efficient binding of hnRNP-C1/C2 to RNA, the context of the U-stretch is more important than the stretch itself; (ii) both the RBD and the bZLM bind RNA independently; and (iii) the C2-related 13-amino acid insert enhances the specificity of either the RBD, the bZLM, or the full-length protein towards its ligand, allowing it to bind only the most high-affinity sequences while discriminating against those that do not perfectly match this category. The three-hybrid system is a powerful tool to work out the functional significance of peptide 'modules' within RNA binding proteins generated by alternative splicing.

Key words: Three-hybrid system; RNA–protein interaction; Heterogeneous nuclear ribonucleoprotein; hnRNPC

1. Introduction

Heterogeneous nuclear ribonucleoproteins (hnRNPs) are found in the eukaryotic nucleus, associated with heterogeneous nuclear RNA (hnRNA) to form higher order structures. hnRNP-C, together with hnRNP-A and B proteins, compose the 'core' of the 40S hnRNP particle. hnRNP-C proteins are believed to be involved in pre-mRNA packaging, spliceosome assembly and in nuclear retention of unspliced hnRNA (reviewed in [1]). hnRNP-C1 and hnRNP-C2, 41 kDa and 43 kDa, respectively, form a stable heterotetramer, $(C1)_3(C2)_1$. hnRNP-C2, an alternatively spliced variant of hnRNP-C1, contains an additional 13 amino acids and is expressed at one-third the level of hnRNP-C1 [2]. The protein sequence of the hnRNP-C1/C2 proteins, highly conserved among vertebrates, contains an RNA binding domain (RBD) and a basic leucine zipper (bZIP)-like motif (bZLM), both implied in RNA binding [3,4]. The additional 13 amino acids of the hnRNP-C2 isoform are located between these two domains.

The RNA binding specificities of hnRNP-C1/C2 proteins have been studied using various experimental approaches. The tetramer was characterized as a sequence-independent RNA chaperonin that is distributed along nascent transcripts by binding 230 nt of RNA in a high cooperative manner [4,5]. UV-cross-linking experiments characterized hnRNP-C1/C2 as poly(U) binding proteins, which preferably bind to the pyrimidine-rich sequences that precede 3′ intron–exon junctions [6,7], the 5′ stem-loop of U2 snRNA [8], the poly(U) stretches at the 3′ end of U4/U6 snRNAs [9], U-rich sequences downstream of polyadenylation cleavage sites [10], and to the AUUUA repeats in the 3′ untranslated region (UTR) of several labile mRNAs [11,12]. hnRNP-C1/C2 were also found to be associated with U-rich sequences within 3′ UTRs of mRNAs in correlation to their stabilization [13–15], and were implicated in enhanced translation of some mRNAs upon binding to U-stretches within their 5′ UTRs [16,17]. In vitro selection/amplification of RNA oligonucleotides from a pool of random sequences (the SELEX procedure) selected oligonucleotides with five or more contiguous uridines as the high affinity substrate for recombinant hnRNP-C1 [3]. Competition binding assays under equilibrium conditions using fluorescence spectroscopy and hnRNP-C tetramer isolated from HeLa cells confirmed the high-affinity binding to one of the SELEX-defined oligonucleotides, termed the 'winner' sequence. The latter experiment also showed that contiguous uridines are not obligatory for generation of high affinity ligand, suggesting that high affinity may depend on a unique context or structure of the RNA [18].

All previous studies relied on biochemical in vitro approaches to analyze the binding properties of hnRNP-C proteins with RNA. Moreover, the possible differences in the RNA binding properties of hnRNP-C1 and hnRNP-C2 have never been addressed. In the current study, we analyzed the binding characteristics of both hnRNP-C1 and hnRNP-C2 to synthetic and natural RNA sequences using a three-hybrid system. The three-hybrid assay provides an easy way to study RNA-protein interactions in the nucleus of a living yeast cell. Although it does not yield quantitative information that is comparable to any in vitro biochemical approach, it provides important information regarding RNA-protein interactions in vivo. The system is based on binding of a bifunctional hybrid RNA to each of two hybrid proteins in order to activate the transcription of a reporter gene, leading to the selection of live

*Corresponding author. Fax: (972)-3-642 2046.
E-mail address: ornaes@post.tau.ac.il (O. Elroy-Stein).

0014-5793/02/$22.00
PII: S0014-5793(02)02938-1

colonies. Two of such systems were developed independently: one utilizes RNA polymerase III, whereas the other utilizes RNA polymerase II for the synthesis of the hybrid RNA [19,20]. Since RNA polymerase III often terminates at a run of four or more consecutive uridine residues [21], it imposes a major problem for the analysis of U-rich RNA baits. Therefore, we have used in this study the RNA polymerase II-based system for the analysis of the hnRNP-C1/C2 ligands. We analyzed the binding of hnRNP-C1 and hnRNP-C2 with a synthetic 'winner' sequence [18] and with two natural sequences representing two U-rich domains within the 5′ UTR of PDGF2/c-*sis* mRNA [16]. This study presents the first example of a successful analysis of RNA-protein interaction to U-rich sequences using the three-hybrid system. It provides in vivo evidence for the binding preferences of hnRNP-C proteins towards natural RNA sequences representing different regions within a natural mRNA, and it demonstrates for the first time the functional differences between the two closely related proteins, hnRNP-C1 and hnRNP-C2.

2. Materials and methods

2.1. Plasmid construction

2.1.1. Hybrid RNA vectors. The sequence of nucleotides 1–227 ('A') or 475–797 ('B') of the human PDGF2/c-*sis* 5′ UTR was PCR amplified using psis4.0 [22] as template and oligonucleotide primers 5′-CCCCACTAGTGGCAACTTCTCCTCC-3′ (JB7) and 5′-GCGAGCCATGGCTGCTCCGG-3′ (OS70) or 5′-GGGTACTAGTGCTGCCGTTGC-3′ (OS66) and 5′-CAGGCCATGGGTCCGAGGCCGGCTA-3′ (OS62), respectively. The PCR products were cloned into blunted *Mlu*I site of pPGKRRE [20] between the Rev responsive element (RRE) and the PGK transcription terminator, to create the intermediate plasmids pGKRRE-'A' and pGKRRE-'B', respectively. The entire transcription unit containing the hybrid RNA between the PGK promoter and terminator, was excised from the above plasmids as an *Xho*I-*Sal*I fragment and cloned into the *Sal*I site of pDBRevM10 [20] to create plasmids pDBRevM10-'A' and pDBRevM10-'B'. The plasmid pDBRevM10-'winner' was constructed by inserting a double-stranded oligonucleotide (prepared by annealing the synthetic oligonucleotides 5′-CGCGTAGTATTTTTGTGGAGC-3′ and 5′-GGCCGCTCCACAAAAATACTA-3′) into the *Mlu*I–*Not*I sites of pRevRX between RRE and PGK terminator. Plasmid pRevRX expresses both 'hybrid protein 1' and the transcription unit that drives the transcription of the hybrid RNA (kindly provided by U. Putz and D. Kuhl). To generate pRevRX-IRE, a 200 bp *Eco*RI fragment containing IRE-MS2 was isolated from pIIIA/IRE-MS2–2 [19], blunted, and cloned into the *Sma*I site of the pRevRX vector. The resulting plasmid was digested with *Sma*I and *Not*I to remove the MS2 sequence, blunted, and self ligated.

2.1.2. 'Hybrid protein 2' vectors. Fragments containing the complete coding sequence of hnRNP-C1 (residues 1–290) or hnRNP-C2 (residues 1–303) were generated using pHC12 or pHCC2 templates, respectively [2], and oligonucleotide primers 5′-CCCCGAGCTCATGGCCAGCAACGTTACC-3′ (OS72) and 5′-GATAAGAATTCTAAACCCCACTA-3′ (OS73). The PCR products were digested with *Sac*I, filled in, digested with *Eco*RI and cloned into the *Sma*I–*Eco*RI sites of pACTII (Clontech) to generate plasmids pACTII-C1 and pACTII-C2, respectively. hnRNP-C1 truncations were generated by PCR using pHC12 template and the following oligonucleotide primers: OS72 and 5′-GGGTCGACCATCTCCGCTGCAG-3′ (M1-M104) for C1Δ1 (residues 1–104); 5′-CCCATGGACTATTATGATAGGATG-3′ (Y119-G290) and OS73 for C1Δ2 (residues 119–290); OS72 and 5′-GGGTCGACAGACTCCATCTTCAC-3′ (M1-S240) for C1Δ3 (residues 1–240); Y119-G290 and M1-S240 for C1Δ4 (residues 119–240). hnRNP-C2 truncations were generated by PCR using pHCC2 template and the following oligonucleotide primers: OS72 and 5′-CCCTCGAGCTGAGTAGAG-3′ (M1-S119) for C2Δ1 (residues 1–119); 5′-CCCCATGGGGTCAGTAAC-3′ (G106-G290) and OS73 for C2Δ2 (residues 106–303); OS72 and M1-S240 for C2Δ3 (residues 1–253); G106-G290 and M1-S240 for C2Δ4 (residues 106–253). PCR fragments C1Δ1, C1Δ3, and C2Δ1 were digested with *Sac*I, filled-in, digested with *Sal*I and cloned into the *Sma*I–*Xho*I sites of pACTII to generate pACTII-C1Δ1, pACTII-C1Δ3, and pACTII-C2Δ1, respectively. Fragment C2Δ1 was digested with *Sac*I, filled-in, digested with *Xho*I and cloned into *Sma*I–*Xho*I sites of pACTII to generate plasmid pACTII-C2Δ1. Fragments C1Δ2 and C2Δ2 were digested with *Nco*I and *Eco*RI and cloned into the *Nco*I–*Eco*RI sites of pACTII to generate pACTII-C1Δ2 and pACTII-C2Δ2, respectively. Fragments C1Δ4 and C2Δ4 were digested with *Nco*I and *Sal*I and cloned into *Nco*I–*Xho*I sites of pACTII to generate plasmids pACTII-C1Δ4 and pACTII-C2Δ4, respectively. pACTII-IRP [19] expressing the GAL4 activating domain fused to the iron responsive element binding protein (IRP) was kindly provided by M. Wickens.

2.2. Yeast

Yeast strains CG-1945 and Y190 (Clontech) were used. Transformants were grown on YNB plates lacking tryptophan and leucine for selection of the TRP1 and LEU2 marker genes. Double transformants forming a stable three-hybrid complex were selected for HIS3 expression by dropping 10 μl of a liquid culture grown to $OD_{600} = 0.1$ on YNB plates lacking tryptophan, leucine and histidine and containing 0.5 mM 3-aminotriazole (3-AT) for the CG-1945 strain or 100 mM 3-AT for the Y190 strain. Cells were allowed to grow for 7–10 days at 30°C.

2.3. Colony color assay (LacZ expression)

Logarithmically growing cells (2 μl, OD_{600} 0.3) were dropped onto a nitrocellulose filter (diameter 82 mm, BA85; Schleicher and Schuell) placed on a YNB plate lacking tryptophan and leucine. Following 48 h at 30°C, filters were removed and incubated overnight in −80°C. The filters were then thawed on 3 mm paper at room temperature for 1 min and then placed on No. 576 paper (Schleicher and Schuell) that had been soaked with Z buffer (Clontech) containing 0.03% Xgal (5-bromo-4-chloro-3-indolyl-β-D-galactoside). Blue color was developed during 2–5 h incubation at 30°C.

3. Results

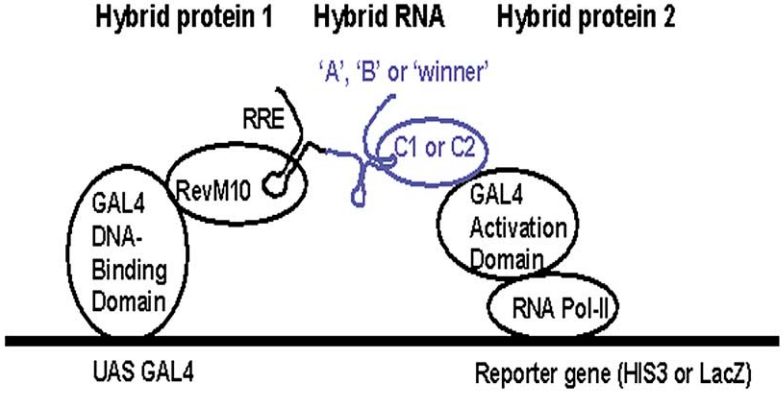
3.1. hnRNP-C1 interacts with the central part of PDGF2/c-sis mRNA leader, but not with another U-rich sequence within the mRNA

Using UV-cross-linking experiments, we previously identified the central region of the PDGF2/c-*sis* 5′ UTR as a high affinity site for hnRNP-C1/C2 binding. This region, spanning nucleotides 475–797 of the PDGF2/c-*sis* mRNA (termed sequence 'B' throughout the current study), contains a stretch of uridines in a context similar to that of the 'winner' sequence, a synthetic sequence previously selected as a high-affinity ligand for hnRNP-C proteins [3,18]. In contrast, another U-rich region within the PDGF2/c-*sis* 5′ UTR, spanning nucleotides 1–227 (termed sequence 'A' throughout the current study), did not show a significant binding activity towards hnRNP-C1/C2 proteins as judged by the UV-cross-linking assay [16].

To confirm the specific binding of hnRNP-C1 to sequence 'B' in vivo, we employed the yeast three-hybrid system in which 'hybrid protein 1' (harboring the DNA binding domain of GAL4) is brought to a close proximity with 'hybrid protein 2' (harboring the GAL4 activation domain) via an interaction of both hybrid proteins with a hybrid RNA. Upon the formation of a stable three-hybrid complex, the transcription of the *HIS3* and *lacZ* genes is activated [20]. In this system, the hybrid RNA is composed of the Rev responsive element (RRE) fused to the RNA sequence of interest, e.g. the bait. Its synthesis is carried out by RNA polymerase II, which can efficiently transcribe an RNA with uridine stretches. Since hnRNP-C1 over-expression in yeast was found to be lethal [23], we have used a single copy plasmid for the expression of 'hybrid protein 2' to lower its toxicity to a level that permitted selection of colonies that express *HIS3* due to the for-

mation of a stable three-hybrid complex. *LacZ* expression was demonstrated by a qualitative colony color assay since it was below the detection limits of a quantitative liquid assay.

As illustrated in Fig. 1, sequence 'A' or 'B' within the 5′ UTR of PDGF2/c-*sis* or the 'winner' sequence [18] were fused to RRE to create the hybrid RNAs. 'Hybrid protein 1' harbors a mutated Rev protein (RevM10) that binds RRE without mediating its nuclear export. Due to the RevM10 component, the three-hybrid complex is retained in the nucleus and thus can activate the transcription of the reporter genes. 'Hybrid protein 2' was designed to carry the RNA binding protein of interest, e.g. hnRNP-C1 or hnRNP-C2 (see Fig. 1) being tested for its ability to bind to the bait. As a positive control for the three-hybrid interaction, the IRP and its cognate RNA, the iron responsive element (IRE), were used as a 'hybrid protein 2'/hybrid RNA pair. Combinations of the three-hybrid components were introduced into the yeast Y190 strain followed by selection for growth on histidine-lacking media supplemented with 3-AT, which further increases the selectivity for the *HIS3* marker. Colonies formed due to sufficient activation of *HIS3* transcription were further tested for activation of *LACZ* expression. Effective growth on selective media and formation of a blue color is indicative of an RNA-protein interaction. Fig. 2A demonstrates the efficient interaction of IRP with the 'IRE' sequence but not with sequence 'A' nor with sequence 'B'. Fig. 2A also demonstrates the efficient interaction of hnRNP-C1 with the 'winner' sequence and with sequence 'B' which contains a stretch similar to the 'winner', but not with sequence 'A' which also contains a stretch of uridines (see Fig. 1). RT-PCR analysis confirmed the efficient expression of 'A' and 'B' sequences. In fact, as demonstrated in Fig. 2B, sequence 'A' was better expressed than sequence 'B' and yet a more stable three-hybrid complex was formed with sequence 'B'. This data clearly shows that for efficient binding to hnRNP-C1, the context of the U-stretch is more important than the stretch itself, demonstrating for the first time the sequence specificity of hnRNP-C1 binding in vivo.



Nucleotides 1-227 of c-sis 5'UTR (sequence 'A')

5'-GGUGGCAACUUCUCCUCCUGCAGCCGGGAGCGGCCUGCCUGCCUCCCUGCGCACCCGCA
GCCUCCCCCGCUGCCUCCCUAGGGCUCCCCUCCGGCCGCCAGCGCCCAUUUUUCAUUCCCU
AGAUAGAGAUACUUUGCGCGCACACACAUACAUACGCGCGCAAAAAGGAAAAAAAAAAAAAA
AGCCCACCCUCCAGCCUCGCUGCAAAGAGAAAACCGGAGCAGCC-3'

Nucleotides 475-797 of c-sis 5'UTR (sequence 'B')

5'-CUCGAGCUGCCGUUGCAAAGCCAACUUUGGAAAAAGUUUUUUGGGGGAGACUUGGGCCU
UGAGGUGCCCAGCUCCGCGCUUUCCGAUUUUGGGGGCCUUUCCAGAAAAUGUUGCAAAAAA
CUAAGCCGGCGGGCAGAGGAAAACGCCUGUAGCCGGCGAGUGAAGACGAACCAUCGACUGG
CCGUGUUCCUUUUCCUCUUGGAGGUUGGAGUCCCCUGGGCGCCCCCACACGGCUAGACGC
CUCGGCUGGUUCGCGACGCAGCCCCCCGGCCGUGGAUGCUCACUCGGGCUCGGGAUCCGC
CCAGGUAGCGGCCUCGGACCCA-3'

'winner' sequence

5'-AGUAUUUUUGUGGA-3'

Fig. 1. Schematic presentation of the three-hybrid system. The system is composed of three hybrid components as detailed by Putz et al. [20]. Hybrid protein 2 is composed of complete or truncated hnRNP-C1 or hnRNP-C2 coding sequences (see Fig. 4) fused to the GAL4 activation domain. The hybrid RNA is composed of the RRE fused to nucleotides 1–227 (sequence 'A') or 475–797 (sequence 'B') of the human PDGF2/c-*sis* 5′ UTR, respectively, or to the synthetic oligomer selected as the 'winner' sequence for hnRNP-C binding [3,18]. U-rich regions are underlined. Sequence similarities are highlighted by red bold letters.

3.2. hnRNP-C1 and hnRNP-C2 differ in their RNA binding characteristics

hnRNP-C2, an alternatively spliced variant of hnRNP-C1, contains an additional 13 amino acids between the RNA binding domains RBD and bZLM [3]. Since the 13-amino acid insertion within hnRNP-C2 does not interfere with the primary sequence of either of the RNA binding motifs, the RNA binding specificities of hnRNP-C1 and hnRNP-C2 were considered to be identical and have never been compared. Here, we used the three-hybrid system to test the binding of both hnRNP-C1 and hnRNP-C2 proteins to the different RNA baits. Fig. 3A demonstrates that hnRNP-C2 was able to form a stable complex with the 'winner' sequence, as judged by the growth on histidine-lacking plates and by *LACZ* expression in strain Y190. However, in contrast to hnRNP-C1, hnRNP-C2 was not able to efficiently bind sequence 'B' (compare Fig. 2A to Fig. 3A). Western analysis confirmed the efficient expression of both hnRNP-C1 and hnRNP-C2 (Fig. 3B). Note that although hnRNP-C proteins are toxic for yeast cells, they are expressed under selective pressure that selects for three-hybrid complex formation. The negative effect of the 13-amino acid insert in hnRNP-C2 on its binding to sequence 'B' was established using an additional yeast strain, CG-1945 (Fig. 3A). *LacZ* expression in this strain was below detection limits due to the weak *CYC1* minimal promoter. However, by assaying for *HIS3* expression we were able to confirm the observation that the interaction of hnRNP-C2 with sequence 'B' was significantly less efficient than the interaction of hnRNP-C1 with sequence 'B'. This data suggests that the hnRNP-C2-related small peptide insertion enhances the specificity of the protein towards its ligand, allowing it to bind only high-affinity sequences (such as the 'winner') and to discriminate more stringently against sequences that do not perfectly match this category (such as sequence 'B' or 'A').

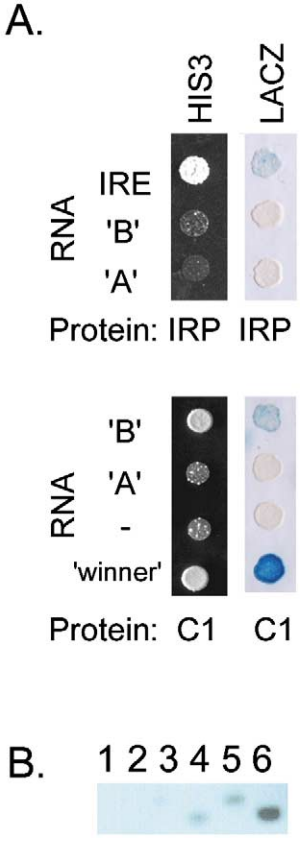


Fig. 2. hnRNP-C1 forms a stable complex with sequence 'B' but not with sequence 'A'. A: pACTII-IRP or pACTII-C1 expressing IRP or hnRNP-C1 as part of 'hybrid protein 2' were co-transfected into Y190 yeast strain together with pRevRX-IRE, pDBRevM10-'A', pDBRevM10-'B', or pDBRevM10-'winner' expressing the hybrid RNA bearing the different baits as indicated, in addition to 'hybrid protein 1'. Colonies were selected for *HIS3* expression and assayed for *LACZ* expression. B: Total RNA was extracted from colonies expressing pACTII-C1 and pDBRevM10-'A'/pDBRevM10-'B' subjected to RT-PCR using oligonucleotide primers that are described in Section 2 and separated on 2% agarose gel electrophoresis. 1,2: 25 PCR cycles without RT; 3,4: RT followed by 20 PCR cycles; 5,6: RT followed by 25 PCR cycles. 1,3,5: amplification of sequence 'B'; 2,4,6: amplification of sequence 'A'.

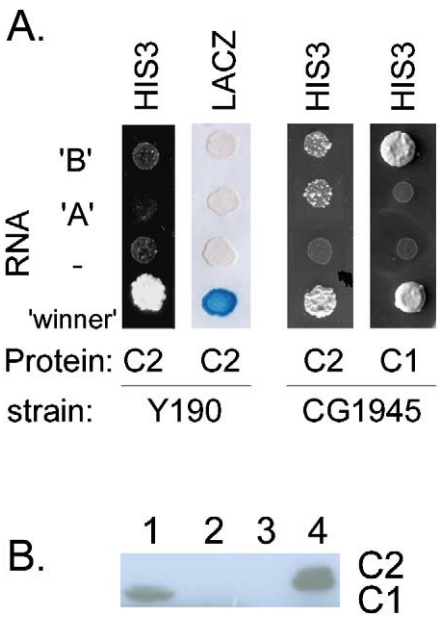


Fig. 3. hnRNP-C1 and hnRNP-C2 differ in their RNA binding characteristics. A: pACTII-C1 or pACTII-C2 expressing hnRNP-C1 or hnRNP-C2 as part of 'hybrid protein 2', were co-transfected into Y190 or CG-1945 yeast strains together with pDBRevM10-'A', pDBRevM10-'B', or pDBRevM10-'winner' expressing the different baits as indicated, in addition to 'hybrid protein 1'. Colonies were selected for *HIS3* expression and assayed for *LACZ* expression. B: 50 μg protein extracted from Y190 strain co-transfected with pDBRevM10-'winner' and pACTII-C1 (lanes 1,2) or pDBRevM10-'winner' and pACTII-C2 (lanes 3,4), not selected (lanes 2,3) or selected (lanes 1,4) for survival in 100 mM 3AT, were analyzed by 10% SDS–PAGE followed by Western analysis using antibodies specific for HA epitope.

3.3. The 13-amino acid insertion in hnRNP-C2 affects the RNA binding characteristics of both the RBD and the bZLM domains

The three-hybrid system was further used to test the ability of hnRNP-C1 or hnRNP-C2 truncated forms to interact with the different RNA baits. The RNA binding capacities of the RBD alone, the bZLM alone or both of them together were checked. Two versions of each truncation, e.g. with or without the hnRNP-C2-related 13-amino acid insertion, were analyzed in order to determine whether the effect of the small peptide on RNA binding was specific to one or both of the RNA binding domains. The expression of all the truncated proteins was confirmed by Western analysis (not shown). Fig. 4, showing the three-hybrid assay, demonstrates that both the RBD

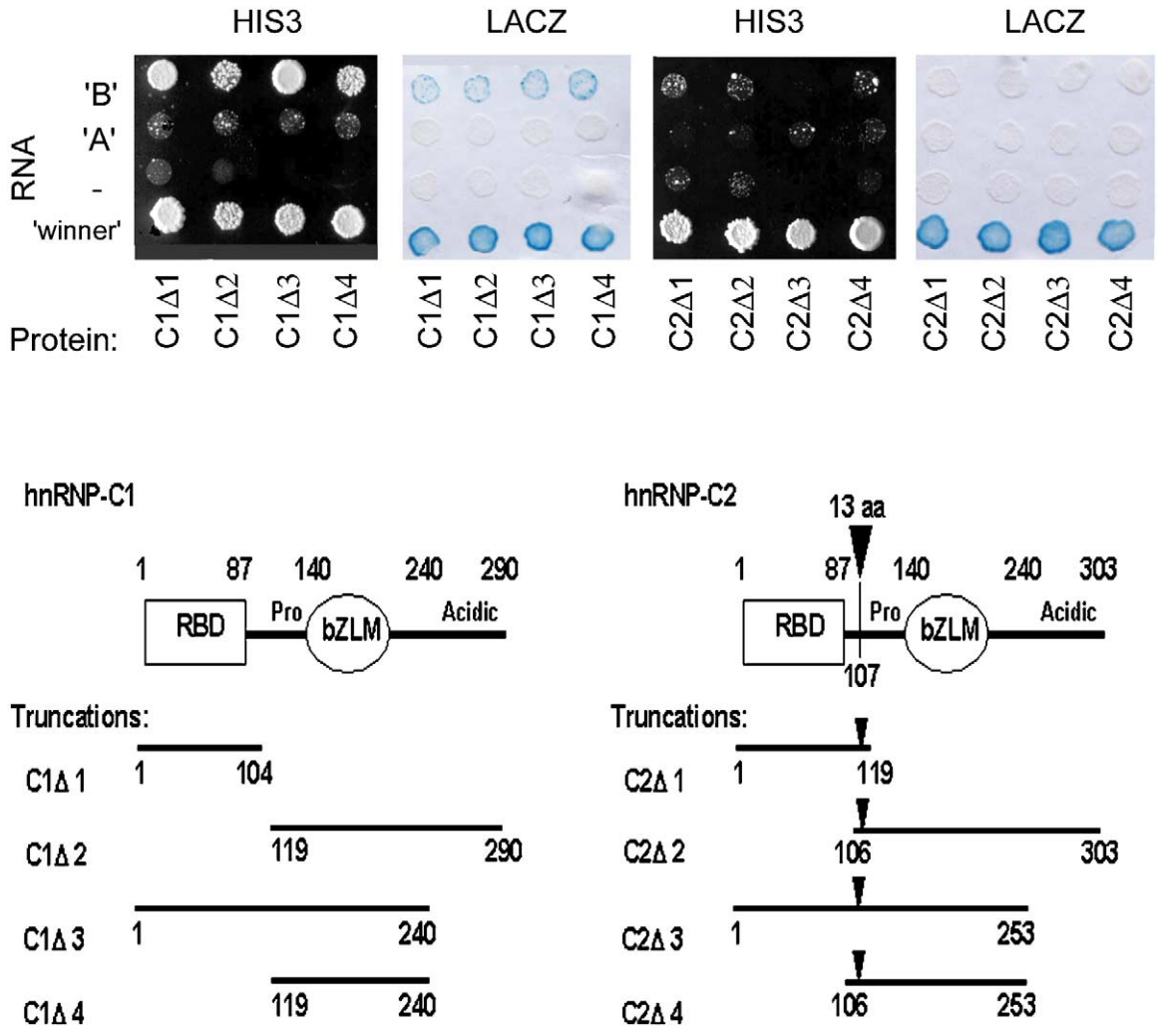


Fig. 4. The effect of the hnRNP-C2-related small peptide on RNA binding by the RBD and the bZLM. A: Schematic presentation of the hnRNP-C1/C2 truncations detailed in Section 2. B: Each pACTII plasmid expressing the indicated truncated protein as part of 'hybrid protein 2' was co-transfected into Y190 yeast strain together with pDBRevM10 plasmid expressing the different baits as indicated, in addition to 'hybrid protein 1'. Colonies were selected for *HIS3* expression and assayed for *LACZ* expression.

and the bZLM were able to bind RNA independently, with efficiencies similar to the full-length form of the protein. Interestingly, the presence of the hnRNP-C2-related 13-amino acid stretch impaired the ability of either the RBD or the bZLM to bind the PDGF2/c-*sis* mRNA-derived sequences, but did not affect their ability to bind to the 'winner' sequence. This data demonstrates that the hnRNP-C2-related small peptide enhances the RNA binding specificity of either the RBD or the bZLM domains.

4. Discussion

An important question concerning hnRNP-C proteins, centers on their RNA binding characteristics, e.g. their binding preferences and the elements conferring RNA binding specificity. High-affinity binding sites for purified recombinant hnRNP-C1 have been identified by Gorlach et al. [3] using the SELEX procedure that revealed 28 oligonucleotides, each harboring one or more groups of five or more contiguous uridines. Accordingly, preference for U-rich sequences has been documented by numerous UV-cross-linking studies [6–17]. However, Soltaninassab et al. [18] have shown that although native hnRNP-C proteins purified from HeLa cells possess enhanced affinity towards one of the SELEX-identified oligonucleotide, termed the 'winner' sequence, they possess no enhanced affinity to other U-rich oligonucleotides. The present study is consistent with the latter observation. It clearly demonstrates that under physiological conditions, e.g. in the nucleus of a living yeast cell, hnRNP-C1 efficiently binds the synthetic 'winner' sequence but possesses a clear preference towards one of the two natural U-stretch-containing sequences. Since sequence 'A', which contains a stretch of five contiguous uridines, did not mediate the formation of a three-hybrid complex, it seems that contiguous uridines do not alone define a high affinity substrate. The high similarity between sequence 'B' and the 'winner' and the ability of sequence 'B' to mediate the formation of a three-hybrid complex suggest that unique structures with steric nucleotide positioning may affect hnRNP-C protein interactions with natural RNA molecules in vivo, in agreement with a previous study [18]. How does this observation fit with the description of hnRNP-C1/C2 tetramer as an RNA chaperonin that binds 230 nt along the RNA in a sequence-independent, highly cooperative manner [4]? One possibility may be that high-affin-

ity sequences along the long hnRNA molecule serve as nucleation sites for initial hnRNP-C binding followed by further binding along the length of the nascent transcripts.

This study demonstrates that the 13-amino acid insert that distinguishes hnRNP-C2 from hnRNP-C1 directly influences RNA binding. The hnRNP-C2-associated insertion, beginning after Gly106, negatively affected the binding affinity to sequence 'B', whereas no significant influence on binding to the 'winner' sequence was observed. Interestingly, Gorlach et al. [3] previously reported that the 10 amino acids immediately preceding the COOH-terminal to the RBD motif (amino acids 95–104) prevented the minimal RBD from binding nonspecific RNA sequences. Here we show that the additional 13 residues further affect RNA binding by enhancing the discrimination ability against sequences that do not belong to the group of the highest-affinity ligands. What is the physiological significance of hnRNP-C2? It is an open question whether the RNA binding performance by $C1_3C2$ tetramer differs from that of $C1_4$ tetramer. It is possible that the $C1_3C2$ formula constitutes the best design to fit a physiological function such as binding to certain high-affinity sequences followed by a more uniform, less discriminative, high-cooperative binding along nascent transcripts. This concept awaits experimental verification. Interestingly, a remarkably similar example relates to two additional hnRNP core proteins, hnRNP-A2 and hnRNP-B1, that differ by only a 12-amino acid insert in hnRNP-B1 [2]. It is an open question whether the B1-related small peptide has a direct effect on RNA binding.

The RNA binding activity of hnRNP-C proteins is conferred by its canonical RBD comprising the amino-terminal 94 amino acids [3,24,25], and by a bZLM motif comprising residues 140–207 of hnRNP-C1 or 153–220 of hnRNP-C2 [4,26,23]. The data regarding RNA binding by both domains has been accumulated via a series of in vitro studies, some of which were performed under physiological conditions. The in vivo data obtained here from the yeast three-hybrid system demonstrate the ability of each of the two domains to mediate the formation of a stable three-hybrid complex, confirming the ability of the RBD and of the bZLM to independently bind RNA. The current study provides interesting new information regarding the ability of the C2-related 13 amino acids to affect the RNA binding characteristics of both the RBD and the bZLM. In the presence of the small peptide, both domains were more competent in distinguishing between the various RNA sequences and picked only high-affinity ligands. The three-hybrid system is a powerful tool to work out the functional significance of peptide 'modules' within RNA binding proteins generated by alternative splicing.

Acknowledgements: We thank U. Putz and D. Kuhl for the three-hybrid system that includes the pPGKRRE, pDBRevM10, pACTII, pRevRX plasmids and the CG1945 yeast strain, M. Wickens for pIIIA/IRE-MS2-2 and pACTII-IRP plasmids, and G. Dreyfuss for pHC12 and pHCC2 plasmids. We thank Debbie Silvera for comments on the manuscript. This work was supported by Grant 4367 from the Chief Scientist's Office of Ministry of Health, Israel, and by a grant from the Israel Cancer Association. N. Koloteva-Levine's fellowship was supported by the Israel Cancer Research Fund.

References

[1] Dreyfuss, G., Matunis, M.J., Pinol-Roma, S. and Burd, C.G. (1993) Annu. Rev. Biochem. 62, 289–321.
[2] Burd, C.G., Swanson, M.S., Gorlach, M. and Dreyfuss, G. (1989) Proc. Natl. Acad. Sci. USA 86, 9788–9792.
[3] Gorlach, M., Burd, C. and Dreyfuss, G. (1994) J. Biol. Chem. 269, 23074–23078.
[4] McAfee, J., Shahied-Milam, L., Soltaninassab, S. and LeStourgeon, W. (1996) RNA 2, 1139–1152.
[5] Huang, M., Rech, J.E., Northington, S.J., Flicker, P.F., Mayeda, A., Krainer, A.R. and LeStourgeon, W.M. (1994) Mol. Cell. Biol. 14, 518–533.
[6] Swanson, A. and Dreyfuss, G. (1988) Mol. Cell. Biol. 8, 2237–2241.
[7] Swanson, M.S. and Dreyfuss, G. (1988) EMBO J. 7, 3519–3529.
[8] Tetsamani, J. and Pederson, T. (1996) J. Biol. Chem. 271, 24922–24926.
[9] Forne, T., Rossi, F., Labourier, E., Antoine, E., Cathala, G., Brunel, C. and Tazi, J. (1995) J. Biol. Chem. 270, 16476–16481.
[10] Wilusz, J., Feig, D.I. and Shenk, T. (1998) Mol. Cell. Biol. 8, 4477–4483.
[11] Hamilton, B.J., Nagy, E., Malter, J.S., Arrick, B.A. and Rigby, W.F. (1993) J. Biol. Chem. 268, 8881–8887.
[12] Sokolowski, M. and Schwartz, S. (2001) Virus Res. 73, 163–175.
[13] Zaidi, S.H.E. and Malter, J.S. (1995) J. Biol. Chem. 270, 17292–17298.
[14] Rajagopalan, L.E., Westmark, C.J., Jarzembowski, J.A. and Malter, J.S. (1998) Nucleic Acids Res. 26, 3418–3423.
[15] Nanbu, R., Montero, L., d'Oraazio, D. and Nagamine, Y. (1997) Eur. J. Biochem. 247, 169–174.
[16] Sella, O., Gerlitz, G., Le, S.Y. and Elroy-Stein, O. (1999) Mol. Cell. Biol. 19, 5429–5440.
[17] Millard, S.S., Vidal, A., Markus, M. and Koff, A. (2000) Mol. Cell. Biol. 20, 5947–5959.
[18] Soltaninassab, S.R., McAfee, J.G., Shahied-Milam, L. and LeStourgeon, W.M. (1998) Nucleic Acids Res. 26, 3410–3417.
[19] SenGupta, D.J., Zhang, B., Kraemer, B., Pochart, P., Fields, S. and Wickens, M. (1996) Proc. Natl. Acad. Sci. USA 93, 8496–8501.
[20] Putz, U., Skehel, P. and Kuhl, D. (1996) Nucleic Acids Res. 24, 4838–4840.
[21] Geiduschek, E.P. and Tocchini-Valentini, G.P. (1988) Annu. Rev. Biochem. 57, 873–914.
[22] Rao, C.D., Pech, M., Robbins, K.C. and Aaronson, S.A. (1988) Mol. Cell. Biol. 8, 284–292.
[23] Tan, J., Kajiwara, Y., Shahied, L., Li, J., McAfee, J. and LeStourgeon, W.M. (2001) J. Mol. Biol. 305, 829–838.
[24] Gorlach, M., Wittekind, M., Beckman, R., Mueller, L. and Dreyfuss, G. (1992) EMBO J. 11, 3289–3295.
[25] Wan, L., Kim, J.K., Pollard, V. and Greyfuss, G. (2001) J. Biol. Chem. 276, 7681–7688.
[26] Shahied-Milam, L., Soltaninassab, S., Iyer, G. and LeStourgeon, W.M. (1998) J. Biol. Chem. 273, 21359–21367.