



# C-terminal determinants for RNA binding motif 7 protein stability and RNA recognition

Amr M. Sobeh, Catherine D. Eichhorn\*

Department of Chemistry, University of Nebraska, 639 North 12th St, Lincoln, NE 68588, USA

## ARTICLE INFO

### Keywords:

Solution state NMR spectroscopy  
Nuclear exosome targeting complex  
DNA damage repair  
RNA-protein interactions

## ABSTRACT

The 7SK ribonucleoprotein (RNP) is a critical regulator of eukaryotic transcription. Recently, RNA binding motif 7 (RBM7) containing an RNA recognition motif (RRM) was reported to associate with 7SK RNA and core 7SK RNP protein components in response to DNA damage. However, little is known about the mode of RBM7-7SK RNA recognition. Here, we found that RRM constructs containing extended C-termini have increased solubility compared to a minimal RRM construct, although these constructs aggregate in a temperature and concentration-dependent manner. Using solution NMR dynamics experiments, we identified additional structural features observed previously in crystal but not in solution structures. To identify potential RBM7-7SK RNA binding sites, we analyzed deposited data from *in cellulo* crosslinking experiments and found that RBM7 primarily crosslinks to the distal region of 7SK stem-loop 3 (SL3). Electrophoretic mobility shift assays and NMR chemical shift perturbation experiments showed weak binding to 7SK SL3 constructs *in vitro*. Together, these results provide new insights into RBM7 RRM folding and recognition of 7SK RNA.

## 1. Introduction

Proper regulation of RNA polymerase II (Pol II) transcription is required to maintain cellular homeostasis and efficient gene expression [1]. An essential checkpoint in transcription regulation is promoter-proximal pausing of Pol II  $\sim$  20–60 nt downstream of the transcription start site [2–5], controlled in part by the kinase positive transcription elongation factor b (P-TEFb) [5–8]. In metazoa, the 7SK ribonucleoprotein (RNP) sequesters P-TEFb in a catalytically inactive pool and P-TEFb must be released from 7SK RNP to restore P-TEFb kinase activity [9–11].

The 7SK RNP assembles into a core complex comprised of the 7SK RNA, a  $\sim$  331 nucleotide (nt) noncoding RNA in humans [12]; the methylphosphate capping enzyme (MePCE), which caps and remains bound to the 7SK RNA 5' end [13,14]; and the La-related protein 7 (Larp7), which binds to the 7SK RNA 3' end [15–18]. MePCE acts cooperatively with Larp7 to assemble with 7SK RNA to form a stable ternary complex [13,19]. The accessory hexamethylene bisacetamide-inducible protein 1 or 2 (HEXIM1/2) is required to assemble P-TEFb onto the 7SK RNP and inhibit P-TEFb kinase activity [20–22]. Together, these components minimally constitute the P-TEFb-inactivated 7SK RNP [16,23,24]. Numerous accessory proteins have been reported to

promote release of P-TEFb from 7SK RNP [5,25–34]. On release P-TEFb, Bromodomain 4 (Brd4), and other transcription factors trigger the transition of Pol II from a promoter-proximal paused to a processive elongation state [5,35,36]. P-TEFb dysregulation or 7SK RNP malfunction is associated with several diseases including blood and solid tumor cancers, cardiac hypertrophy, and primordial dwarfism [5,37]. Moreover, several viruses manipulate host 7SK RNP for viral survival, notably HIV-1 [38] and more recently SARS-CoV-2 [39].

Spatial organization of 7SK RNP on chromatin may enable a rapid cellular response to environmental conditions. In support of this model, genome-wide analyses report the presence of 7SK RNP at the promoters of protein-coding genes [33,40,41]. UV exposure results in the global release of promoter-proximally paused Pol II [42], stimulates the rapid release of P-TEFb from 7SK RNP, and induces activation of UV-response genes [2,34,36]. 7SK RNA knockout in HAP1 cells showed reduced viability after UV irradiation, with impaired activation of UV-response genes requiring 7SK RNP-P-TEFb to rescue activation [43]. In HEK293 and HeLa cells, exposure to a chemical mimetic of UV-induced DNA damage (4-nitroquinoline 1-oxide, 4-NQO) similarly resulted in decreased viability [44,45]. 4-NQO exposure induced phosphorylation of the RNA binding protein 7 (RBM7) by the p38<sup>MAPK</sup>-MK2 pathway [34,46] and transcription activation of genes involved in DNA damage

\* Corresponding author.

E-mail address: [ceichhor@unl.edu](mailto:ceichhor@unl.edu) (C.D. Eichhorn).

<https://doi.org/10.1016/j.bpc.2022.106928>

Received 12 May 2022; Received in revised form 13 October 2022; Accepted 30 October 2022

Available online 8 November 2022

0301-4622/© 2022 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

response (DDR) [34,47,48]. RBM7 individual-nucleotide resolution UV crosslinking and immunoprecipitation (iCLIP) experiments found that RBM7 interacts with 7SK RNA, and co-immunoprecipitation experiments showed that RBM7 interacts with core 7SK RNP proteins MePCE and Larp7 as well [34]. A model has been proposed in which genotoxic stress stimulates phosphorylation of RBM7 to promote release of P-TEFb from 7SK RNP [34]. More recently, p38<sup>MAPK</sup>-MK2-induced phosphorylation of RBM7 has been shown *in cellulo* after SARS-CoV-2 viral infection [39,49], suggesting a similar pathway may be activated during immune response.

The primary identified function of RBM7 is as a component of the nuclear exosome targeting complex (NEXT), a trimeric complex containing RBM7, the zinc-knuckle protein ZCCHC8, and the RNA helicase hMTR4 [50,51]. NEXT interacts robustly with the RNA cap-binding complex to induce transcription termination; in addition, NEXT is involved in the degradation of aberrant Pol II transcripts [52–54]. The role of RBM7 in 7SK RNP-mediated DDR is an alternate regulatory mechanism independent of function in NEXT. When phosphorylated, RBM7 has globally reduced interaction to RNA with reduced NEXT function [55]. Importantly, 7SK RNA abundance is not reduced after 4-NQO exposure, indicating that RBM7 is functioning independently of its role in NEXT [34].

Together, these data suggest that RBM7 plays an essential role in the release of P-TEFb from the 7SK RNP complex to respond to genotoxic stress [34]. However, little is known regarding how RBM7 assembles with 7SK RNA and core 7SK RNP protein components. RBM7 contains an N-terminal canonical RNA recognition motif (RRM) with conserved RNP1 and RNP2 sequences [56,57] (Fig. 1A–B) and a C-terminal serine-rich domain that is phosphorylated by p38<sup>MAPK</sup> [34,46,56,57]. *In vitro*, the RBM7 RRM preferentially interacts with polypyrimidine repeat sequences with micromolar binding affinity [56]. Four high-resolution structures have been determined of the RBM7 RRM by solution NMR spectroscopy or X-ray crystallography, with varying N- and C-terminal ends and observed structural features [56,58,59]. Recently, two intermediate-resolution cryo-electron microscopy (cryoEM) structures were reported of the NEXT complex containing a recombinant RBM7 RRM construct [60,61].

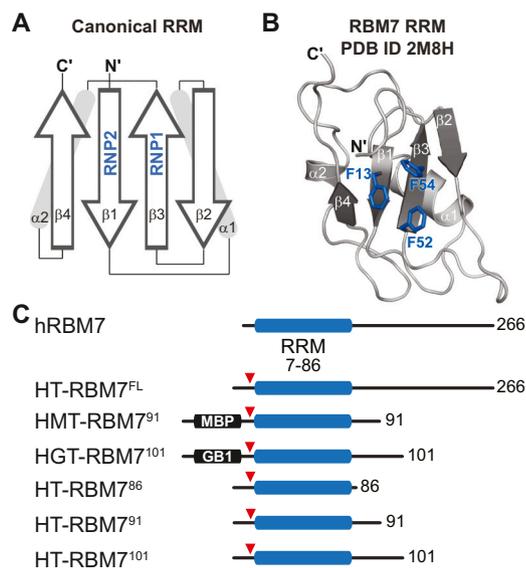
Here, we recombinantly expressed and purified RBM7 RRM constructs with various N-terminal fusion tags and C-terminal ends to investigate RRM folding and interaction with 7SK RNA. We found that, consistent with prior literature, constructs containing longer C-terminal ends have increased solubility compared to a minimal RRM construct. Using a combination of NMR spectroscopy, optical melting, and aggregation assays, we found that the isolated RRM is prone to concentration- and temperature-dependent aggregation. Consistent with solution NMR and X-ray structures, <sup>1</sup>H–<sup>15</sup>N heteronuclear NOE and chemical shift perturbation experiments reported disordered N- and C-termini that do not interact with the RRM. Unexpectedly, we observed an ordered  $\alpha$ 2– $\beta$ 4 region supporting the presence of an additional  $\beta$ -strand, which has been observed previously in X-ray structures but not the solution NMR structural ensemble. We compare available structures, find that this region is located at the protein-protein interface, and propose a model in which protein-protein interactions aid in ordering this region. To identify potential RBM7-7SK RNA binding sites, we analyzed deposited data from *in cellulo* crosslinking experiments and performed electrophoretic mobility shift assays and NMR chemical shift perturbation experiments. We found that RBM7 crosslinks specifically to the distal region of 7SK stem-loop 3 (SL3) and that the RRM binds weakly to single-stranded RNA sequences derived from 7SK SL3. We propose a mechanism in which RBM7 binds to unpaired residues in 7SK SL3. Together, these data provide new insights into RBM7 RRM folding and recognition of the 7SK RNA.

## 2. Methods

### 2.1. Recombinant protein construct design, expression, and purification

The DNA sequence encoding human RBM7 was codon optimized for *Escherichia coli* codon bias and purchased as a gblock (Integrated DNA Technologies, IDT). The RBM7 gene was cloned into a pET vector containing N-terminal His<sub>6</sub>, tobacco etch virus (TEV) protease recognition site, and maltose binding protein (MBP) domains with a kanamycin resistance gene (Addgene plasmid #29656) using *in vivo* assembly [62]. Briefly, Benchling was used to design overlapping primers complementary to the 5' and 3' ends of both the gene insert and vector insertion sites. PCR was performed to produce gene insert and vector with complementary 5' and 3' ends. PCR products were purified using a PCR cleanup kit (Zymo Research) and directly transformed into DH5 $\alpha$  competent cells. Individual colonies were selected and cultured in LB media with kanamycin, followed by plasmid purification using a plasmid miniprep kit (Qiagen). All constructs derived from this plasmid were generated using site-directed mutagenesis with the exception of the GB1 domain, which was inserted using *in vivo* assembly. Sequences were verified using Eurofins Genomics LLC and Plasmidsaurus.

For protein expression, plasmids were transformed into *E. coli* NiCo21(DE3) competent cells (New England Biolabs). Cells were cultured with shaking in M9 minimal media with kanamycin at 37 °C to an OD<sub>600</sub> of 0.6–0.8, transferred to 18 °C, expression induced with 0.5 mM isopropyl  $\beta$ -D-1-thiogalactopyranoside (IPTG), and grown for 18–20 h. For NMR experiments, <sup>13</sup>C-labelled glucose (Cambridge Isotope Laboratory) and/or <sup>15</sup>N-labelled ammonium chloride (Cambridge Isotope Laboratory) was used in the M9 media. Buffers used for protein purification were adapted from [56]. Cells were harvested by centrifugation and resuspended in buffer R (50 mM NaPO<sub>4</sub> pH 7.0, 1 M NaCl, 40 mM imidazole, 10 mM MgCl<sub>2</sub>, 10% glycerol, 1 mM tris(2-carboxyethyl) phosphine (TCEP)), 0.5 mM PMSF, and lysozyme. After sonication, soluble and insoluble fractions were separated by centrifugation at 35,000 xg for 45 min, and the supernatant was clarified by 0.45  $\mu$ m syringe filter. Affinity purification was performed using nickel nitrilotriacetic acid (Ni-NTA) affinity column (Qiagen) attached to an AKTA start system (GE Healthcare). After loading the supernatant, the column was washed with 10 column volumes (CV) of buffer R followed by elution in a linear gradient with buffer E (50 mM NaPO<sub>4</sub> pH 7.0, 250 mM NaCl,



**Fig. 1.** RBM7 domain topology and constructs used in study. A) Schematic of canonical RNA recognition motif (RRM) with conserved RNP1 and RNP2 shown in blue. B) Solution NMR structure of human RBM7 RRM (PDB ID 2M8H) with conserved Phe residues in RNP1 and RNP2 shown as blue sticks. C) Constructs used in this study. RRM shown in blue and TEV cleavage site indicated with red triangle. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

300 mM imidazole, 10% glycerol, 1 mM TCEP). Fractions were analyzed by 4–12% Bis-Tris NuPAGE™ gels (ThermoScientific). To remove the His<sub>6</sub>-MBP or His<sub>6</sub>-GB1 tag, protein was dialyzed against RBM7 storage buffer (20 mM NaPO<sub>4</sub> pH 5.5, 50 mM L-glutamate, 50 mM L-arginine) in the presence of TEV protease (Addgene #92414 [63] recombinantly expressed and purified in house) for 2–3 h at room temperature. After dialysis, cleaved RBM7 protein was purified from His<sub>6</sub>-MBP or His<sub>6</sub>-GB1 by a second Ni-NTA purification step (Qiagen).

As a final purification step, proteins were purified by size exclusion chromatography using a HiLoad16/600 Superdex 75 pg (GE Healthcare) column attached to a ÄKTA Prime M25 FPLC system (GE Healthcare) in RBM7 storage buffer. Fractions were analyzed by 4–12% Bis-Tris NuPAGE™ gel. Pure fractions were concentrated using a 10 kDa Amicon® concentrator. The absorbance at 280 nm was measured using a Nanodrop spectrophotometer to calculate sample concentration using the Beer-Lambert law [64]. Expsy ProtParam [65] was used to compute protein extinction coefficients (Supplementary Table S1), molecular weights, and theoretical pI values.

## 2.2. RNA sample preparation

RNA sequences used in this study are included in Supplementary Table S2. Single-stranded RNA (ssRNA) oligonucleotides were purchased from IDT, resuspended in ultrapure water, and diluted to 1 mM in ultrapure water. 7SK SL3 hairpin RNA constructs were produced using *in vitro* RNA transcription with T7 RNA polymerase (Addgene #124138 [66], prepared in-house) and chemically synthesized DNA templates (IDT) following established protocols [18]. Briefly, T7 RNA Polymerase, 500 nM DNA template, and transcription buffer (40 mM Tris pH 8, 1 mM spermidine, 0.01% Triton-X, 40 mM MgCl<sub>2</sub>, 2.5 mM DTT, 20% DMSO, and 2 mM each rATP, rCTP, rUTP, rGTP) were incubated at 37 °C for 6–8 h. RNA was purified by 15% denaturing polyacrylamide gel electrophoresis (PAGE) with 1× TBE running buffer and the RNA band was visualized by UV shadowing with a handheld UV lamp at 254 nm. After band excision, RNA was eluted from the gel using the ‘crush and soak’ method [67] by incubating gel pieces in crush and soak buffer (300 mM sodium acetate pH 5.2, 1 mM EDTA) for 24–48 h at room temperature. RNA was further purified to remove acrylamide contaminants by ion-exchange chromatography using a diethylaminoethanol (DEAE) column (GE Healthcare) and elution into buffer (10 mM sodium phosphate pH 7.6, 1 mM EDTA, 1.5 M KCl). RNA was diluted to <100 μM in ultrapure water and annealed by heating to 95 °C for 3 min, followed by snap cooling on ice for 1 h. RNA was then buffer exchanged into the appropriate buffer using a 3 kDa Amicon® concentrator.

## 2.3. NMR spectroscopy

Solution NMR spectroscopy experiments were performed at 293.15 K on a Bruker Neo 600 MHz NMR spectrometer equipped with a triple-resonance HCN cryoprobe. NMR samples were prepared in RBM7 storage buffer with added 5% D<sub>2</sub>O at 0.1–0.5 mM concentrations in 3 mm or 5 mm NMR tubes (Norell). Amide backbone resonance assignments were obtained from Biological Magnetic Resonance Data Bank (BMRB) [68] entry 19252 and validated with standard triple resonance assignment experiments [69,70]. Briefly, 3D HNCA, HNCACB, HNCOC, HNCACO, and <sup>15</sup>N-edited NOESY experiments from the Bruker experimental suite were acquired and analyzed using a 0.1 mM sample of HT-RBM7<sup>101</sup>. Weighted average chemical shift perturbations (CSP) were calculated using the equation  $\sqrt{\Delta H^2 + 0.1\Delta N^2}$  [70]. The TALOS+ webserver [71] was used to compute predicted secondary structure and S<sup>2</sup> order parameter values using the BMRB-deposited STAR file as input. TALOS+ output files are provided as Supplementary material. <sup>1</sup>H–<sup>15</sup>N heteronuclear nuclear Overhauser effect (NOE) experiments (hsqcnoef3gps) from the Bruker experimental suite were recorded in an interleaved manner with 32 scans and 2 s incremental delay for 0.3 mM protein

samples. The heteronuclear NOE is reported as the residue-specific ratio of peak intensity between the saturated and unsaturated experiments ( $I_{\text{sat}}/I_{\text{unsat}}$ ). Error was estimated as the standard deviation of noise in the saturated ( $\sigma_{I_{\text{sat}}}$ ) and unsaturated ( $\sigma_{I_{\text{unsat}}}$ ) experiments [72,73]:  $\frac{\sigma_{\text{NOE}}}{\text{NOE}} = \sqrt{\frac{\sigma_{I_{\text{sat}}}^2}{I_{\text{sat}}^2} + \frac{\sigma_{I_{\text{unsat}}}^2}{I_{\text{unsat}}^2}}$ .

Data were processed using NMRPipe [74] and analyzed using NMRFAM-Sparky 1.470 powered by Sparky 3.190 [75] in the NMRbox virtual machine [76]. RNA titration experiments were performed using 0.2 mM protein in RBM7 storage buffer/5% D<sub>2</sub>O, with a reference spectrum recorded prior to addition of final concentrations of 0.02 mM (0.1 equivalent) and 0.1 mM (0.5 equivalent) RNA sequentially added to the NMR sample. All plots were generated using Google Colaboratory notebooks.

## 2.4. Protein aggregation assay

Protein aggregation assays were performed in duplicate at ambient temperature (approximately 22 °C) using the PROTEOSTAT® Thermal shift stability assay kit (Enzo life sciences) and a Qubit Flex Fluorometer. Protein samples were prepared at 0.1 mM and 0.3 mM in RBM7 storage buffer with controls including 1× TS reagent dye in buffer and the kit-supplied β-lactoglobulin (4 mg/mL, 0.1 mM).

## 2.5. Circular Dichroism (CD) spectroscopy

CD experiments were performed on a Jasco815 spectrometer equipped with a Peltier temperature control device. Protein samples were prepared at 50 μM in CD buffer (20 mM sodium phosphate, 150 mM KCl, pH 5.5). CD spectra were collected at 20 °C before and after thermal unfolding experiments with the following parameters: 190–400 nm spectral range; 0.2 nm data interval, 1 s D.I.T., 1 nm band width, 100 nm/min scanning speed, and 5 accumulations. Thermal unfolding experiments were performed in duplicate with a temperature range of 5 °C – 100 °C (forward) or 100 °C – 5 °C (reverse) and a ramp rate of 2 °C/min, with molar ellipticity values measured at 222 nm and 211 nm every 5 °C, 1 s D.I.T., and 1 nm band width.

## 2.6. iCLIP data analysis

RBM7 iCLIP sequencing data [34] were obtained from accession code E-MTAB-6475 in the EMBL-EBI database. Data were processed using the University of Nebraska-Lincoln Holland Computing Center (HCC) following the workflow reported in previous studies [77,78]. First, FastQC [79] was run to analyze the per-base sequence quality of the data, per-base sequence content, barcode content and other quality content. Next, trim\_galore [80] was used to trim the adapters and enhance the quality of the reads according to the quality obtained from the raw data. Next, iCount [81] was used to demultiplex the barcode sequences. Next, STAR [82] was used to index the GRCh38.p13 genome v34 (Gencode), align the reads to the genome, and generate a BAM file. Bedtools 2.27 [83] was used to convert the file of aligned reads from BAM to BED format, shift reads by one nt to account for reverse transcriptase (RT) termination adjacent to the crosslinked site (hereafter referred to as RT stop), extract the residue position number of the RT stop, and concatenate reads to map read coverage across the genome. The BED file was used as input for an in-house python program (<https://github.com/ceichhorn2/7SKRNAextraction>) to extract reads from RN7SK (ENSG00000202198, chr6: 52995620–52995950), convert to 7SK RNA residue number (1–331), and produce a file with the RT stop position number, the associated chromosome number, and the number of reads with the RT stop position number. The read counts for RT stops for a given residue (positional RT stop counts) were normalized to the total number of 7SK reads using the equation  $\text{normalized RT stops} = \frac{\text{number of positional RT stop reads}}{\text{total number of 7SK reads}}$ . The average and standard deviation of the normalized RT stop values of the DMSO control (2 replicates) and 4-

NQO treated (4 replicates) experiments were calculated using Excel (Supplementary File 2) and plots were generated using matplotlib. The average normalized RT stop values in the 4-NQO treated experiments were plotted on a secondary structure model of the upper stem of 7SK SL3 RNA (residues 210–264) using VARNA [84].

## 2.7. Electrophoretic Mobility Shift Assays (EMSA)

RNA and protein samples were prepared as described above and diluted to 50  $\mu$ M and 170  $\mu$ M, respectively, in RBM7 storage buffer. To perform EMSA experiments, increasing amounts of protein samples were added to 50 pmol RNA (Supplementary Table S3). A protein-only sample was loaded in the first lane as a control to identify any nucleic acid contamination in the protein sample. The mixture was incubated on ice for 15–20 min, then 1  $\mu$ L of loading dye (30% glycerol, bromophenol blue) was added for a total volume of 7.9  $\mu$ L. Samples were loaded and run on 10% PAGE (37.5:1 bisacrylamide:acrylamide crosslinking ratio) at 50–60 V for 40–80 min depending on RNA construct length. RNA was visualized with toluidine blue stain for hairpin and SYBR Safe (ThermoScientific) for ssRNA constructs. Experiments were performed in duplicate or triplicate to ensure reproducibility.

## 3. Results

### 3.1. Characterization of RBM7 RRM expression and purification

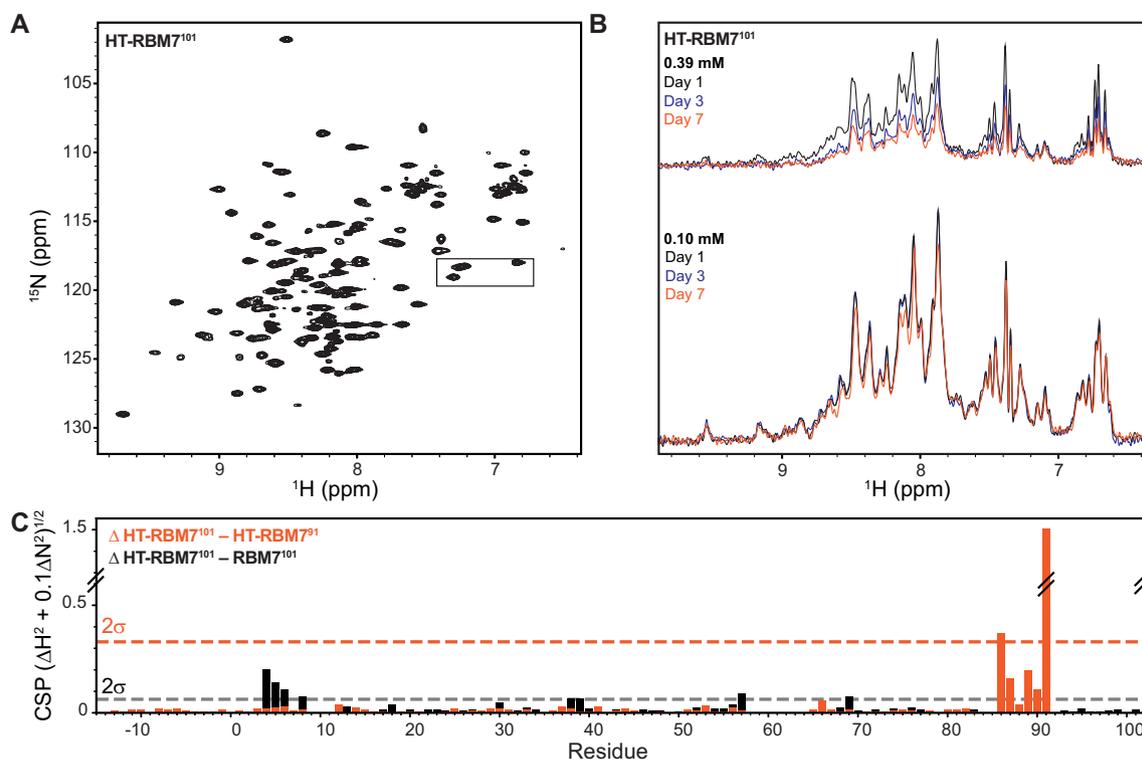
X-ray crystallographic, cryoEM, and solution state NMR structures of the human RBM7 RRM have been previously determined using constructs with varying C-terminal boundaries [58–61,85]. A survey of solved structures showed that the RRM contains a canonical  $\beta$ 1- $\alpha$ 1- $\beta$ 2- $\beta$ 3- $\alpha$ 2- $\beta$ 4 topology between amino acid (aa) residues T11-Q83; however, X-ray crystal structures report an additional  $\beta$ 4' strand in the  $\alpha$ 2- $\beta$ 4 loop and extended  $\beta$ 4 strand (Supplementary Fig. S6). Constructs used for X-ray crystallography structure determination had a C-terminal residue that was either aa S86 [58] or A91 [59], whereas the construct used for the solution NMR structure determination (named here RBM7<sup>2M8H</sup>) had a truncated N-terminus and slightly longer C-terminus (aa 6–94) [56]. We first attempted to recombinantly express the full-length human RBM7 protein with an N-terminal His<sub>6</sub> tag (HT-RBM7<sup>FL</sup>); however, this construct did not solubly express in *E. coli* (Supplementary Fig. S1). We next generated constructs of RBM7 RRM with varying C-terminal ends: 1–86, 1–91, and 1–101 (named HT-RBM7<sup>86</sup>, HT-RBM7<sup>91</sup>, HT-RBM7<sup>101</sup>). To assess the impact of solubility tags on protein expression and solubility, additional fusion constructs were generated including either MBP or GB1 domains between His<sub>6</sub> and TEV sites (HMT-RBM7<sup>91</sup> and HGT-RBM7<sup>101</sup>, respectively) (Supplementary Fig. S2). Protein expression was greatest for HMT-RBM7<sup>91</sup> and HGT-RBM7<sup>101</sup> constructs (approximately 50 mg purified protein per L cell culture); however, despite successful TEV protease cleavage neither His<sub>6</sub>-MBP nor His<sub>6</sub>-GB1 could be separated from the RBM7 domain after TEV cleavage (Supplementary Fig. S2A–B). Both MBP and GB1 are acidic (theoretical pI 5.1 and 4.5, respectively); it is likely that these domains have electrostatic interactions to the basic RBM7 (theoretical pI 9.5). The minimal HT-RBM7<sup>86</sup> construct had poor yield after purification (approximately 2 mg per L cell culture), with significant amounts of protein observed in the insoluble fraction after cell lysis (Supplementary Fig. S2C) suggesting misfolded or aggregated protein present in inclusion bodies in *E. coli* cells [86]. HT-RBM7<sup>86</sup> exhibited poor solubility and readily precipitated throughout purification and during concentration. Compared to HT-RBM7<sup>86</sup>, both HT-RBM7<sup>91</sup> and HT-RBM7<sup>101</sup> had higher yields (approximately 5–10 mg per L cell culture) with little protein observed in the insoluble fraction after cell lysis (Supplementary Fig. S2D–E). After purification, both HT-RBM7<sup>91</sup> and HT-RBM7<sup>101</sup> could be concentrated to 0.5 mM without visible precipitation in the buffer used in this study (20 mM sodium phosphate pH 5.5, 50 mM L-arginine, 50 mM L-glutamate).

### 3.2. Comparison of RBM7 RRM constructs reveals concentration and temperature dependent aggregation

Solution state NMR spectroscopy was used to evaluate and compare folding of HT-RBM7<sup>86</sup>, HT-RBM7<sup>91</sup>, and HT-RBM7<sup>101</sup>. The chemical shift assignments of RBM7<sup>2M8H</sup> (aa 6–94, BMRB entry 19252) [56] were transferred onto 2D amide <sup>1</sup>H–<sup>15</sup>N HSQC spectra of HT-RBM7<sup>101</sup> and verified with standard 3D triple resonance assignment experiments. The deposited chemical shifts lack assignments for the  $\beta$ 3- $\alpha$ 2 loop (aa K58-V63) and C-terminal residues (aa R85-S86, S89-H90). Good agreement between deposited and independently assigned chemical shifts was observed. Consistent with the prior study, we observed significant line-broadening for  $\beta$ 3- $\alpha$ 2 loop residue F57 and  $\alpha$ 2 residues Y65-N68, and were unable to assign intervening residues K58-V63 (Supplementary Fig. S3). Likewise, we observed line-broadening for  $\beta$ 4-adjacent residues Q83 and S86 and were unable to assign intervening residues F84-R85 (Supplementary Fig. S3) indicative of chemical exchange at the  $\alpha$ 2 helix and  $\beta$ 4 strand regions.

The 2D amide <sup>1</sup>H–<sup>15</sup>N HSQC spectrum of spectrum of HT-RBM7<sup>86</sup> showed reasonable chemical shift dispersion with residues S86-Q101 absent, as expected. Several resonances corresponding to  $\beta$ 1 (aa L12-G15),  $\alpha$ 1- $\beta$ 2 (aa H31-K38),  $\beta$ 3 (aa F52-V55, F57),  $\alpha$ 2 (aa A66-G72), and  $\beta$ 4 (aa I82-Q83) suffered from extreme line-broadening, were absent, or shifted relative to constructs with extended C-termini, indicative of global chemical exchange and conformational instability (Supplementary Fig. S3A). Within 24 h the sample began to precipitate, indicating poor solubility of the minimal HT-RBM7<sup>86</sup> construct. In contrast, the 2D amide <sup>1</sup>H–<sup>15</sup>N HSQC spectrum of both HT-RBM7<sup>91</sup> and HT-RBM7<sup>101</sup> showed excellent chemical shift dispersion and uniform peak intensity (Supplementary Fig. S3B–C and Fig. 2A). As expected, residues Q93-Q101 were absent from the HT-RBM7<sup>91</sup> spectrum with chemical shift perturbations observed for C-terminal residues S86-A91 (Supplementary Fig. S3B, Fig. 2C). No additional chemical shift changes were observed for RRM residues, indicating that C-terminal residues do not interact with the RRM. Due to the observed interaction of the RBM7 RRM with solubility tags, we next evaluated the potential impact of the N-terminal His<sub>6</sub> tag on the RRM. The construct RBM7<sup>101</sup> was generated by using TEV protease to cleave the His<sub>6</sub> tag (Fig. 1C). The 2D amide <sup>1</sup>H–<sup>15</sup>N HSQC spectra of HT-RBM7<sup>101</sup> and RBM7<sup>101</sup> are nearly identical (Supplementary Fig. S3D). As expected, resonances corresponding to the His<sub>6</sub> tag and TEV recognition site are absent in the RBM7<sup>101</sup> construct (Supplementary Fig. S3D) and chemical shift perturbations are observed for N-terminal residues A4-A8 (Fig. 2C).

For both HT-RBM7<sup>91</sup> and HT-RBM7<sup>101</sup> constructs, we observed that samples at standard NMR concentrations (0.3–0.5 mM) appeared to undergo signal loss over the course of several days although there was no visible precipitation or opacity in the NMR tube, suggesting that the RBM7 RRM may form soluble aggregates. To test this hypothesis, 1D <sup>1</sup>H amide proton spectra were compared over the course of several days for HT-RBM7<sup>101</sup> samples prepared at 0.10 and 0.39 mM (Fig. 2B). While the global signal decreased by approximately 50% for the 0.39 mM sample seven days after purification, there was little to no loss in signal for the 0.10 mM sample seven days after purification. To further probe RBM7 RRM aggregation, we performed an aggregation assay using the same buffer (20 mM sodium phosphate, pH 5.5, 50 mM L-arginine, 50 mM L-glutamate), temperature (20 °C), and protein concentrations (0.1 and 0.3 mM) as the NMR experiments (Supplementary Fig. S4). Compared to the  $\beta$ -lactoglobulin control protein, both HT-RBM7<sup>91</sup> and HT-RBM7<sup>101</sup> constructs aggregated within several hours. However, the apparent aggregation was more significant for the 0.30 mM samples compared to the 0.1 mM samples. Initially, both constructs at 0.10 mM had minimal apparent aggregation; however, after 24 h HT-RBM7<sup>101</sup> showed a substantial increase in aggregation compared to HT-RBM7<sup>91</sup>. At 0.3 mM, both HT-RBM7<sup>91</sup> and HT-RBM7<sup>101</sup> show nearly identical aggregation behavior. Together, these results indicate that at concentrations above 0.1 mM, the RBM7 RRM undergoes soluble aggregation.



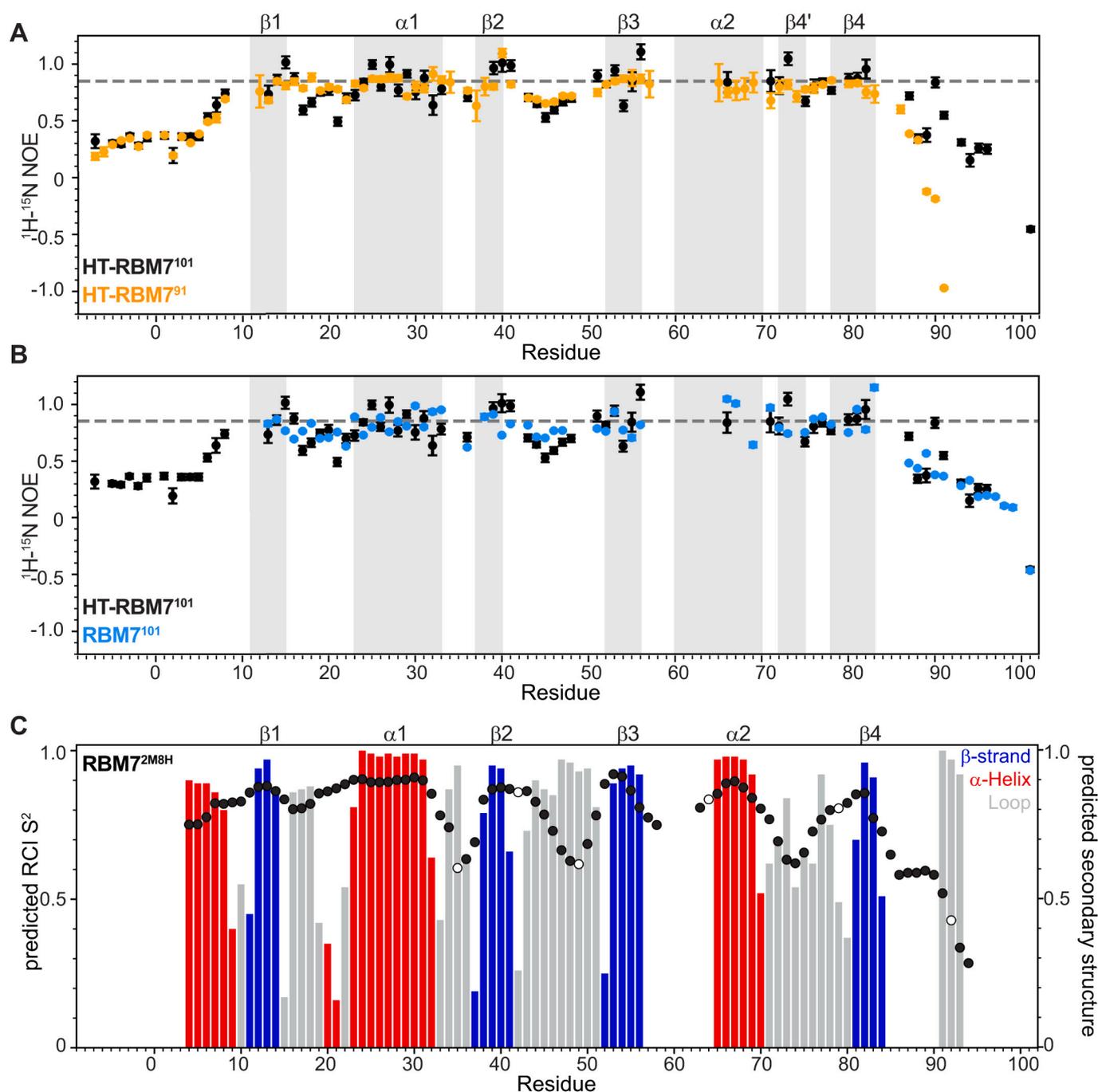
**Fig. 2.** Comparison of RBM7 constructs by solution state NMR. Backbone amide 2D  $^1\text{H}$ – $^{15}\text{N}$  HSQCs of A) HT-RBM7<sup>101</sup> (black). Resonances boxed in black are unassigned side-chain resonances folded into the spectrum. B) time-course 1D  $^1\text{H}$  amide proton experiment with different concentrations C) weighted average CSP for HT-RBM7<sup>101</sup> and either HT-RBM7<sup>91</sup> (orange) or RBM7<sup>101</sup> (black). Dashed lines indicate the  $2\sigma$  standard deviation of the weighted average CSP.

To investigate the thermodynamic stability of the RBM7 RRM, we used circular dichroism (CD) spectroscopy to perform thermal unfolding experiments for HT-RBM7<sup>91</sup> and HT-RBM7<sup>101</sup>. For these experiments, the buffer composition was adjusted to remove L-arginine and L-glutamate, which both have strong CD signals that could not be deconvoluted by baseline correction. Before the temperature unfolding experiment, both constructs had characteristic CD spectra indicative of a folded protein containing  $\alpha$ -helices and  $\beta$ -strands with a minimum at 210–220 nm (Supplementary Fig. S5A–B). In a typical biomolecular unfolding experiment, as the temperature increases the biomolecule will unfold, losing secondary structure, in a cooperative process in which the CD signal will reduce in value. Unexpectedly, the opposite result was observed for both HT-RBM7<sup>91</sup> and HT-RBM7<sup>101</sup> (Supplementary Fig. S5C–D), suggesting that these constructs increase in secondary structure with increasing temperature. For both constructs, this transition began after  $\sim 30^\circ\text{C}$ , with a midpoint at approximately  $60^\circ\text{C}$ . During the reverse experiment, in which the sample was cooled, the CD signal remained constant indicating an irreversible process. The CD spectrum of HT-RBM7<sup>91</sup> after melting showed a significant increase in structure, with a local minimum at 211 nm indicating increased  $\beta$ -strand character (Supplementary Fig. S5A–B). For HT-RBM7<sup>101</sup>, a similar transition to increased secondary structure was observed after  $\sim 30^\circ\text{C}$  that was irreversible. In addition, significant loss of CD signal was observed at  $\sim 85^\circ\text{C}$ , with visible precipitation in the cuvette (Supplementary Fig. S5D). From these experiments, we conclude that both HT-RBM7<sup>91</sup> and HT-RBM7<sup>101</sup> undergo temperature-induced aggregation that is irreversible. In addition, HT-RBM7<sup>101</sup> is more susceptible to aggregation compared to HT-RBM7<sup>91</sup>, consistent with our results from the aggregation assay.

### 3.3. NMR dynamics experiments support presence of additional secondary structure features observed previously in crystal but not solution structures

To assess the impact of residues at the N- and C-terminus on RBM7 RRM conformational dynamics we performed  $^1\text{H}$ – $^{15}\text{N}$  heteronuclear NOE experiments, which report on internal motions at ps-ns motional timescales, for HT-RBM7<sup>91</sup>, HT-RBM7<sup>101</sup>, and RBM7<sup>101</sup> (Fig. 3A–B).  $^1\text{H}$ – $^{15}\text{N}$  NOE values were unable to be measured for several  $\alpha 2$  residues due to missing assignments (K58–V63) and severe line broadening (F57, Y65, N68). The  $^1\text{H}$ – $^{15}\text{N}$  NOE values for HT-RBM7<sup>101</sup> show reduced values for N- and C-terminal residues, indicating gradually increasing disorder at the termini. Reduced values are also observed in  $\beta 1$ – $\alpha 2$  loop residues L17–T22 and  $\beta 2$ – $\beta 3$  loop residues D44–G47.  $^1\text{H}$ – $^{15}\text{N}$  NOE values for HT-RBM7<sup>101</sup>, HT-RBM7<sup>91</sup>, and RBM7<sup>101</sup> are overall similar (Fig. 3A–B), indicating that the disordered N- and C-termini do not alter RRM global folding and dynamics. Residues S89–A91 in HT-RBM7<sup>91</sup> have reduced values compared to HT-RBM7<sup>101</sup>, consistent with increased dynamics at the C-terminus. For all three constructs residues N71–I80, which comprise the  $\alpha 2$ – $\beta 4$  loop in the solution NMR structure, have values similar to those of structured regions of the RRM and are higher than expected for a ten-residue loop (Fig. 3A–B). In comparison, the ten-residue  $\beta 2$ – $\beta 3$  loop has reduced values compared to structured regions, indicating a flexible loop. In X-ray structures of the RBM7 RRM, these residues contain an additional  $\beta 4'$  strand (residues G72–L75) and an extended  $\beta 4$  strand (residues R78–R85) compared to the  $\beta 4$  strand in the solution NMR structure (residues K81–Q83) (Supplementary Fig. S6). These  $^1\text{H}$ – $^{15}\text{N}$  NOE values support the presence of a  $\beta 4'$  and extended  $\beta 4$  strand as observed in the X-ray structures. Together, the heteronuclear NOE data shows a disordered N- and C-termini, flexible  $\beta 2$ – $\beta 3$  loop, and presence of a  $\beta 4'$  strand and extended  $\beta 4$  strand.

The prior solution NMR study and this work used different RRM construct boundaries and different sample conditions. The RBM7<sup>2M8H</sup> construct from the prior study had shortened N- and C-termini (aa 6–94)



**Fig. 3.** Comparison of RBM7 RRM dynamics. A) Plot of  $^1\text{H}-^{15}\text{N}$  heteronuclear NOE values for HT-RBM7<sup>101</sup> (black) and HT-RBM7<sup>91</sup> (orange). B) Plot of  $^1\text{H}-^{15}\text{N}$  heteronuclear NOE values for HT-RBM7<sup>101</sup> (black) and RBM7<sup>101</sup> (blue). Dashed line at 0.85 indicates average value for structured residues. Secondary structure elements are indicated by gray boxes. C) TALOS+ predicted  $S^2$  values and secondary structure from BMRB entry 19252. Proline residues are shown as open circles. Secondary structure is plotted using the confidence score from TALOS+ output and is colored by  $\alpha$ -helix (red),  $\beta$ -strand (blue), and random coil (gray). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

compared to HT-RBM7<sup>101</sup> (aa 1–101). In addition, the buffer conditions in the prior study had increased pH and ionic strength compared to this study; for our constructs, we found that reducing the pH and including L-arginine and L-glutamate improved construct solubility. To further compare findings from the prior study on RBM7<sup>2M8H</sup> to this study, we used BMRB-deposited chemical shifts as input for the TALOS+ server [71] to compute predicted secondary structure and  $S^2$  order parameters for RBM7<sup>2M8H</sup> (Fig. 3C). The order parameter is a normalized degree of order that reports on the relative amplitude of internal motions and is scaled from 0 to 1, where 0 is maximum and 1 is minimum motions [87].

The overall pattern of the predicted  $S^2$  values for RBM7<sup>2M8H</sup> shows excellent agreement to our observed  $^1\text{H}-^{15}\text{N}$  NOE values, particularly for  $\beta 1$ ,  $\alpha 1$ ,  $\alpha 1$ - $\beta 2$  loop,  $\beta 2$ - $\beta 3$  loop, and C-terminal residues. For residues N71-I80, the predicted  $S^2$  values show decreased values and predicted random coil secondary structure, as observed in the solution NMR structure of RBM7<sup>2M8H</sup>. Interestingly, N-terminal residues A4-D9 have a predicted  $\alpha$ -helical secondary structure, which was not observed in the solution NMR structure.

### 3.4. RBM7 RRM binds 7SK RNA stem-loop 3

RBM7 was previously identified to interact with the 7SK RNP using iCLIP experiments that showed RBM7 crosslinking to the stem-loop 3 (SL3) region of 7SK RNA; however, the precise binding site was not defined in the prior study [34]. In addition, mutagenesis of the RRM domain showed that the RBM7 RRM RNP1 and RNP2 sequences were required for 7SK RNP association *in cellulo* [34]. To gain additional insights into specific 7SK RNA residues that crosslinked to RBM7, we processed and analyzed iCLIP data deposited to EMBL. The per-residue 7SK RT stops were normalized to the total number of 7SK reads for both DMSO control and 4-NQO treated samples to measure crosslinking enrichment across the 7SK RNA sequence (Fig. 4A). Consistent with the previous study [34], RBM7 crosslinking to 7SK RNA was observed in both control and treated samples (Fig. 4A) suggesting an interaction between RBM7 and the 7SK RNP even in the absence of DNA damage. Although the 7SK SL3 secondary structure and boundaries are not well characterized, independent chemical mapping studies report SL3 boundaries between residues 200–274 [88–90]. Crosslinking was enriched in the upper stem of 7SK SL3, particularly G232-U236 and C243-G249, with maximum crosslinking for residues A245-U246 (Fig. 4B).

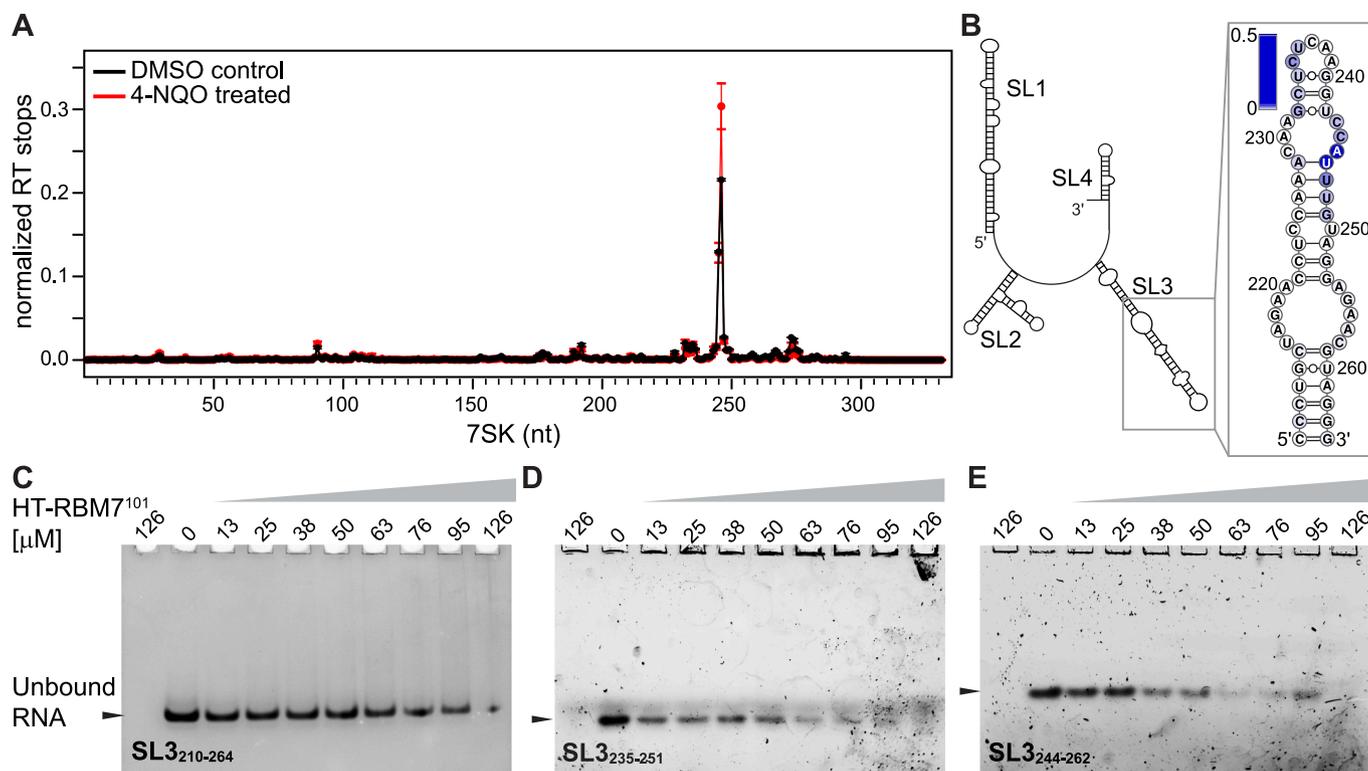
To investigate RBM7 RRM binding to 7SK SL3 *in vitro*, we performed qualitative EMSAs using HT-RBM7<sup>101</sup> and various constructs of 7SK SL3 comprising either hairpin or ssRNA sequences (Fig. 4C-F). Unfortunately, sample precipitation was observed shortly after addition of protein to RNA, indicating that although an RNA-protein complex was forming, the complex was not soluble. As a result, EMSAs showed only unbound RNA and not the complex. Binding was inferred from the decreasing intensity of the unbound band with increasing protein concentration. We first used a hairpin construct comprising the upper stem of 7SK RNA SL3 (SL3<sub>210-264</sub>) (Fig. 4B) and observed extremely weak binding requiring significant HT-RBM7<sup>101</sup> stoichiometric excess

(Fig. 4C). Given that canonical RRM generally bind ssRNA [91], and the RBM7 RRM preferentially binds to pyrimidine-rich ssRNA [56] we then assessed the binding of ssRNA constructs corresponding to 7SK SL3 regions with enriched crosslinking to RBM7: SL3<sub>235-251</sub>, which includes the SL3 apical loop, and SL3<sub>244-262</sub> (Fig. 4D-E). Both RNA constructs show relatively improved binding compared to the hairpin construct, although excess protein is still required to saturate binding.

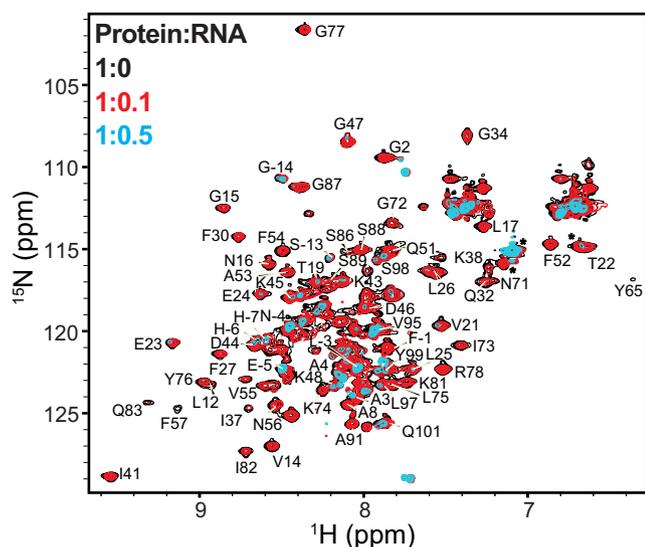
To further probe binding of HT-RBM7<sup>101</sup> to RNA we performed RNA titration experiments, in which SL3<sub>244-262</sub> was titrated into HT-RBM7<sup>101</sup>, and monitored binding with <sup>1</sup>H-<sup>15</sup>N HSQC experiments. Similar to EMSA experiments, immediate precipitation of the protein was observed upon addition of RNA. However, we were able to collect <sup>1</sup>H-<sup>15</sup>N HSQC spectra of HT-RBM7<sup>101</sup> with 0.1 and 0.5 equivalents of SL3<sub>244-262</sub> (Fig. 5). Addition of 0.1 eq RNA globally reduced the resonance signal intensity compared to protein in the absence of RNA, indicating that the loss in signal is due to RNA-protein complex formation and subsequent precipitation. The <sup>1</sup>H-<sup>15</sup>N HSQC spectra of HT-RBM7<sup>101</sup> in the absence and presence of 0.1 equivalents of RNA is identical with no apparent chemical shift perturbations, suggesting that the observed resonances are unbound protein. At 0.5 equivalents, the majority of backbone amide resonances of the RRM disappear (Fig. 5).

## 4. Discussion

As a component of NEXT, RBM7 targets RNA for processing and turnover [53,54]. RBM7 assembly with proteins is essential for its function in NEXT [50,58]. An alternate function for RBM7 has recently been identified, independent of NEXT, in which RBM7 assembles with both 7SK RNA and core 7SK RNP protein components to promote P-TEFb release and subsequent transcription activation of genes in response to genotoxic stress [34]. In this study, we used solution NMR spectroscopy in combination with biochemical and biophysical assays to investigate RBM7 RRM folding and probe RRM interactions with 7SK RNA. To



**Fig. 4.** RBM7 interaction with 7SK RNA. A) Plot of average normalized RT stop values from DMSO control (black) and 4-NQO-treated (red) iCLIP experiments. B) Average normalized RT stop values of 4-NQO-treated sample, shown in panel A, plotted as a heat map onto a secondary structure model of the upper region of 7SK SL3. C-E) Representative gel images of EMSAs for HT-RBM7<sup>101</sup> with 7SK SL3 RNA constructs for C) hairpin SL3<sub>210-264</sub>, D) ssRNA SL3<sub>235-251</sub>, E) ssRNA SL3<sub>244-262</sub>. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)



**Fig. 5.**  $^1\text{H}$ - $^{15}\text{N}$  HSQC titration of HT-RBM7<sup>101</sup> with SL3<sub>244-262</sub> shows disappearance of free RBM7 with increasing RNA. HT-RBM7<sup>101</sup> with no added RNA shown in black, with 0.1 equivalents of SL3<sub>244-262</sub> shown in red, and with 0.5 equivalents of SL3<sub>244-262</sub> shown in blue. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

evaluate the C-terminal boundary requirements for RBM7 RRM folding and stability, we generated RRM constructs corresponding to the X-ray structure of the individual RRM (aa 1–86, PDB ID 5IQQ), RRM-ZCCHC8 complex (aa 1–91, PDB IDs 5LXR and 5LXY), and a C-terminal extended construct ending at residue 101. We found that while all constructs could be recombinantly expressed and purified, the minimal HT-RBM7<sup>86</sup> construct was the least stable with the lowest yields, rapid precipitation, and deterioration of  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectral quality within one day. HT-RBM7<sup>91</sup> and HT-RBM7<sup>101</sup> constructs had improved yields, little to no observable precipitation, and NMR  $^1\text{H}$ - $^{15}\text{N}$  HSQC as well as CD spectra indicating a folded protein.  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra of RRM constructs showed line-broadening, indicative of chemical exchange, for residues in the  $\beta$ 3- $\alpha$ 2- $\beta$ 4 region on the underside of the RRM and global loss of signal for samples at concentrations exceeding 0.3 mM. Thermal melting and aggregation assay experiments showed that the RRM forms soluble aggregates in a concentration- and temperature-dependent manner. These findings are consistent with prior reports that RBM7 is prone to aggregation in the absence of protein binding partners *in vitro* [85].

CryoEM structures of RBM7 in the NEXT complex [60,61], as well as X-ray structures of the RBM7 RRM bound to NEXT component ZCCHC8 [58], show that the RBM7 RRM interacts with an acidic domain of ZCCHC8 (residues 277–328, theoretical pI 4.5) on the underside of the RRM ( $\alpha$ 1- $\alpha$ 2- $\beta$ 4') (Supplemental Fig. S6) and opposite to the RNA-binding interface on the  $\beta$ -sheet surface. Interestingly, we found that acidic solubility tags interact with the basic RRM, suggesting a similar mode of binding to the underside of the RRM. A protein-protein interface on the underside of the RRM ( $\alpha$ 2- $\beta$ 4'- $\beta$ 4, residues E60, P64, N68, K74, Y76, F84, R85) is also observed in the X-ray structure of the individual RBM7 RRM (PDB ID 5IQQ), in which the construct crystallized as a pentamer (Supplemental Fig. S7). These data support a model in which the RBM7 RRM interacts with acidic protein domains on the underside of the RRM, and forms soluble aggregates in the absence of protein cofactors when the underside of the RRM is solvent exposed.

Comparison of NMR  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra and  $^1\text{H}$ - $^{15}\text{N}$  heteronuclear NOE experiments for HT-RBM7<sup>91</sup>, HT-RBM7<sup>101</sup>, and RBM7<sup>101</sup> showed that the N- and C-termini are flexible and do not interact with the RRM, in agreement with previously solved structures [56,59]. Order parameters and secondary structure predicted from RBM7<sup>2M8H</sup> chemical shifts agree with the solution NMR structure that residues at the  $\alpha$ 2- $\beta$ 4

junction adopt an unstructured loop. However, our  $^1\text{H}$ - $^{15}\text{N}$  heteronuclear NOE data shows that residues N71-I80 are ordered, supporting the presence of an additional  $\beta$ 4' strand and extended  $\beta$ 4 strand in agreement with X-ray structures of the RBM7 RRM. The previous NMR study and this study use different RRM constructs as well as different buffer conditions, which may explain these observed differences. We find that the RRM adopts structured  $\beta$ 4' and  $\beta$ 4 strands in solution with the conditions used in this study, resolving the apparent discrepancy between X-ray and NMR structures of the RBM7 RRM. However, we observe line-broadening for residues in this region indicating that while these residues are structured, chemical exchange is present. Unusually, a proline residue (P79) is present at the beginning of the extended  $\beta$ 4 strand near the  $\beta$ 4'- $\beta$ 4 loop. Although prolines are generally disfavored in  $\alpha$ -helices and  $\beta$ -strands due to lack of an amide proton to participate in hydrogen bonding [92], proline can stabilize  $\beta$ -turns by reducing conformational entropy [93]. Interestingly, a human patient with spinal motor neuropathy was found to have an RBM7 P79R mutation that resulted in reduced RBM7 protein abundance, suggesting an important role of this residue in RBM7 function [94,95]. Part of the protein-protein binding interface (Supplemental Figs. S6–7), the  $\alpha$ 2- $\beta$ 4'- $\beta$ 4 region may be further stabilized by protein recognition.

A previous iCLIP study showed enriched crosslinking of RBM7 to 7SK SL3 [34]. In this study, we localized RBM7-7SK RNA crosslinking to the distal region of 7SK SL3, which is enriched in paired and unpaired pyrimidines. Previously, the RBM7 RRM was shown to bind heptamer ssRNA constructs, with affinities ranging from 60 to 150  $\mu\text{M}$  [56]. Highest affinities were observed to polyU and alternating CU-repeat ssRNA sequences, with weak to no binding detected to polyC and polyA ssRNA sequences. A recent cryoEM structure of NEXT in complex with RNA showed that RBM7 makes contacts with both U and A residues [60]. We show that *in vitro*, the RBM7 RRM has little apparent binding to a hairpin RNA construct but binds to ssRNA constructs derived from the 7SK SL3 apical loop and upper stem. Although base paired, the SL3 upper stem is comprised primarily of G•U and A-U base pairs, which may partially unfold *in vivo* or in the presence of protein chaperones to allow RBM7 RRM to bind 7SK RNA. We propose a mode of recognition in which RBM7 binds to the distal region of 7SK SL3 when residues 243–249 in the upper stem are unpaired.

In addition to binding to 7SK RNA, RBM7 interacts with core 7SK RNP proteins MePCE and Larp7 [34]. Given the apparent weak binding of RBM7 to 7SK RNA, RBM7 assembly onto 7SK RNP likely requires both protein-protein and protein-RNA interactions for stable association with 7SK RNP and for P-TEFb release. 7SK SL3 was recently shown to interact with another RNA binding protein, hnRNPA1 [89], which is also involved in P-TEFb release from 7SK RNP [15,96]. Both RBM7 and hnRNPA1 have increased association with 7SK RNP after 4-NQO treatment [34], suggesting a shared mechanism of P-TEFb release. RBM7, hnRNPA1, and additional accessory proteins may work cooperatively, or competitively, with core 7SK RNP proteins to coordinate P-TEFb release for targeted regulation of specific genes in response to genotoxic stress.

## Declaration of Competing Interest

Catherine D. Eichhorn reports financial support was provided by National Institutes of Health. Catherine D. Eichhorn reports financial support was provided by University of Nebraska-Lincoln Nebraska Center for Integrated Biomolecular Communication. Catherine D. Eichhorn reports a relationship with National Institutes of Health that includes: funding grants. Catherine D. Eichhorn reports a relationship with University of Nebraska-Lincoln Nebraska Center for Integrated Biomolecular Communication that includes: funding grants.

## Data availability

Samples, reagents, and raw data files generated in this work are available upon reasonable request. In-house programs generated in this

work are deposited on the GitHub repository <https://github.com/ceichhorn2/7SKRNAextraction>.

## Acknowledgements

We acknowledge funding support from the National Institutes of Health (1R35GM143030), the Nebraska Center for Integrated Biomolecular Communication (P20 GM113126), and UNL startup funds to CDE and UNL UCARE funding to AMS. We thank Ernest Atsrim for critical review of the manuscript, Momodou Camara for assistance with RNA sample preparation, Eng. Youssef Fathy for assistance with programming, Dr. Martha Morton for NMR assistance, and Drs Chia-sin Liew and John-Jack Riethoven at the Holland Computing Center for assistance with processing iCLIP data. We gratefully acknowledge Addgene for plasmids used for plasmid generation and to prepare enzymes involved in preparation of RBM7 and *in vitro* transcribed RNA samples. pDZ2087 was a gift from David Waugh (Addgene plasmid #92414); pQE30-His-T7RNAP was a gift from Sebastian Maerkl & Takuya Ued (Addgene plasmid #124138); pET His6 MBP TEV LIC cloning vector (1M) was a gift from Scott Gradia (Addgene plasmid #29656).

## Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.bpc.2022.106928>.

## References

- Adelman, J.T. Lis, Promoter-proximal pausing of RNA polymerase II: emerging roles in metazoans, *Nat Rev Genet* 13 (2012) 720–731.
- Zhou, T. Li, D.H. Price, RNA polymerase II elongation control, *Annu. Rev. Biochem.* 81 (2012) 119–143.
- Jonkers, J.T. Lis, Getting up to speed with transcription elongation by RNA polymerase II, *Nat Rev Mol Cell Biol* 16 (2015) 167–177.
- Core, K. Adelman, Promoter-proximal pausing of RNA polymerase II: a nexus of gene regulation, *Genes Dev.* 33 (2019) 960–982.
- Alexandre J.C. Quaresma, A. Bugai, M. Barboric, Cracking the control of RNA polymerase II elongation by 7SK snRNP and P-TEFb, *Nucleic Acids Res.* 44 (2016) 7527–7539.
- R.P. McNamara, C.W. Bacon, I. D'Orso, Transcription elongation control by the 7SK snRNP complex: releasing the pause, *Cell Cycle* 15 (2016) 2115–2123.
- Ivanov, Y.T. Kwak, J. Guo, R.B. Gaynor, Domains in the SPT5 protein that modulate its transcriptional regulatory properties, *Mol. Cell. Biol.* 20 (2000) 2970–2983.
- J. Kohoutek, P-TEFb- the final frontier, *Cell Div* 4 (2009) 19.
- N.F. Marshall, J. Peng, Z. Xie, D.H. Price, Control of RNA polymerase II elongation potential by a novel carboxyl-terminal domain kinase, *J. Biol. Chem.* 271 (1996) 27176–27183.
- Z. Yang, Q. Zhu, K. Luo, Q. Zhou, The 7SK small nuclear RNA inhibits the CDK9/cyclin T1 kinase to control transcription, *Nature* 414 (2001) 317–322.
- V.T. Nguyen, T. Kiss, A.A. Michels, O. Bensaude, 7SK small nuclear RNA binds to and inhibits the activity of CDK9/cyclin T complexes, *Nature* 414 (2001) 322–325.
- E. Ullu, V. Esposito, M. Melli, Evolutionary conservation of the human 7 S RNA sequences, *J. Mol. Biol.* 161 (1982) 195–201.
- Y. Xue, Z. Yang, R. Chen, Q. Zhou, A capping-independent function of MePCE in stabilizing 7SK snRNA and facilitating the assembly of 7SK snRNP, *Nucleic Acids Res.* 38 (2010) 360–369.
- Y. Yang, C.D. Eichhorn, Y. Wang, D. Cascio, J. Feigon, Structural basis of 7SK RNA 5'-gamma-phosphate methylation and retention by MePCE, *Nat. Chem. Biol.* 15 (2019) 132–140.
- B.J. Krueger, C. Jeronimo, B.B. Roy, A. Bouchard, C. Barrandon, S.A. Byers, C. E. Searcey, J.J. Cooper, O. Bensaude, E.A. Cohen, et al., LARP7 is a stable component of the 7SK snRNP while P-TEFb, HEXIM1 and hnRNP A1 are reversibly associated, *Nucleic Acids Res.* 36 (2008) 2219–2229.
- A. Markert, M. Grimm, J. Martinez, J. Wiesner, A. Meyerhans, O. Meyuhas, A. Sickmann, U. Fischer, The La-related protein LARP7 is a component of the 7SK ribonucleoprotein and affects transcription of cellular and viral polymerase II genes, *EMBO Rep.* 9 (2008) 569–575.
- E. Uchikawa, K.S. Natchiar, X. Han, F. Proux, P. Roblin, E. Zhang, A. Durand, B. P. Klaholz, A.-C. Dock-Bregeon, Structural insight into the mechanism of stabilization of the 7SK small nuclear RNA by LARP7, *Nucleic Acids Res.* 43 (2015) 3373–3388.
- C.D. Eichhorn, Y. Yang, L. Repeta, J. Feigon, Structural basis for recognition of human 7SK long noncoding RNA by the La-related protein Larp7, *Proc. Natl. Acad. Sci. U. S. A.* 115 (2018) E6457–E6466.
- Y. Yang, S. Liu, S. Egloff, C.D. Eichhorn, T. Hadjian, J. Zhen, T. Kiss, Z.H. Zhou, J. Feigon, Structural basis of RNA conformational switching in the transcriptional regulator 7SK RNP, *Mol. Cell* 82 (1724–1736) (2022), e1727.
- J.H. Yik, R. Chen, R. Nishimura, J.L. Jennings, A.J. Link, Q. Zhou, Inhibition of P-TEFb (CDK9/cyclin T) kinase and RNA polymerase II transcription by the coordinated actions of HEXIM1 and 7SK snRNA, *Mol. Cell* 12 (2003) 971–982.
- A.A. Michels, A. Fraldi, Q. Li, T.E. Adamson, F. Bonnet, V.T. Nguyen, S.C. Sedore, J. P. Price, D.H. Price, L. Lania, et al., Binding of the 7SK snRNA turns the HEXIM1 protein into a P-TEFb (CDK9/cyclin T) inhibitor, *EMBO J.* 23 (2004) 2608–2619.
- L. Kobbi, E. Demey-Thomas, F. Braye, F. Proux, O. Kolesnikova, J. Vinh, A. Poterszman, O. Bensaude, An evolutionary conserved Hexim1 peptide binds to the Cdk9 catalytic site to inhibit P-TEFb, *Proc. Natl. Acad. Sci. U. S. A.* 113 (2016) 12721–12726.
- S. Egloff, E. Van Herreweghe, T. Kiss, Regulation of polymerase II transcription by 7SK snRNA: two distinct RNA elements direct P-TEFb and HEXIM1 binding, *Mol. Cell. Biol.* 26 (2006) 630–642.
- N. He, N.S. Jahchan, E. Hong, Q. Li, M.A. Bayfield, R.J. Maraia, K. Luo, Q. Zhou, A La-related protein modulates 7SK snRNP integrity to suppress P-TEFb-dependent transcriptional elongation and tumorigenesis, *Mol. Cell* 29 (2008) 588–599.
- M. Briese, M. Sendtner, Keeping the balance: the noncoding RNA 7SK as a master regulator for neuron development and function, *Bioessays* 43 (2021), e2100092.
- P. Liu, Y. Xiang, K. Fujinaga, K. Bartholomeeusen, K.A. Nilson, D.H. Price, B. M. Peterlin, Release of positive transcription elongation factor b (P-TEFb) from 7SK small nuclear ribonucleoprotein (snRNP) activates hexamethylene bisacetamide-inducible protein (HEXIM1) transcription, *J. Biol. Chem.* 289 (2014) 9918–9925.
- J.R. Hogg, K. Collins, RNA-based affinity purification reveals 7SK RNPs with distinct composition and regulation, *RNA* 13 (2007) 868–880.
- T. Cherrier, V. Le Douce, S. Eilebrecht, R. Riclet, C. Marban, F. Dequiedt, Y. Goumon, J.C. Paillart, M. Mericskay, A. Parlakian, et al., CTIP2 is a negative regulator of P-TEFb, *Proc. Natl. Acad. Sci. U. S. A.* 110 (2013) 12655–12660.
- S. Eilebrecht, B.J. Benecke, A. Benecke, 7SK snRNA-mediated, gene-specific cooperativity of HMG1 and P-TEFb, *RNA Biol.* 8 (2011) 1084–1093.
- S. Eilebrecht, E. Wilhelm, B.J. Benecke, B. Bell, A.G. Benecke, HMG1 directly interacts with TAR to modulate basal and tat-dependent HIV transcription, *RNA Biol.* 10 (2013) 436–444.
- R.P. McNamara, J.E. Reeder, E.A. McMillan, C.W. Bacon, J.L. McCann, I. D'Orso, KAP1 recruitment of the 7SK snRNP complex to promoters enables transcription elongation by RNA polymerase II, *Mol. Cell* 61 (2016) 39–53.
- S.A. Gudiapaty, R.P. McNamara, E.L. Morton, I. D'Orso, PPM1G binds 7SK RNA and Hexim1 to block P-TEFb assembly into the 7SK snRNP and sustains transcription elongation, *Mol. Cell. Biol.* 35 (2015) 3810–3828.
- E. Calo, R.A. Flynn, L. Martin, R.C. Spitale, H.Y. Chang, J. Wysocka, RNA helicase DDX21 coordinates transcription and ribosomal RNA processing, *Nature* 518 (2015) 249–253.
- A. Bugai, A.J.C. Quaresma, C.C. Friedel, T. Lenasi, R. Duster, C.R. Sibley, K. Fujinaga, P. Kukanja, T. Hennig, M. Blasius, et al., P-TEFb activation by RBM7 shapes a pro-survival transcriptional response to genotoxic stress, *Mol. Cell* 74 (254–267) (2019), e210.
- M.K. Jang, K. Mochizuki, M. Zhou, H.S. Jeong, J.N. Brady, K. Ozato, The bromodomain protein Brd4 is a positive regulatory component of P-TEFb and stimulates RNA polymerase II-dependent transcription, *Mol. Cell* 19 (2005) 523–534.
- Q. Zhou, J.H. Yik, The yin and Yang of P-TEFb regulation: implications for human immunodeficiency virus gene expression and global control of cell growth and differentiation, *Microbiol. Mol. Biol. Rev.* 70 (2006) 646–659.
- C.W. Bacon, I. D'Orso, CDK9: a signaling hub for transcriptional control, *Transcription* 10 (2019) 57–75.
- M. Barboric, T. Lenasi, Kick-starting HIV-1 transcription elongation by 7SK snRNP depporTation, *Nat. Struct. Mol. Biol.* 17 (2010) 928–930.
- M. Boudhoudou, D. Memon, B. Meyer, K.M. White, V.V. Rezelj, M. Correa Marrero, B.J. Polacco, J.E. Melnyk, S. Ulferts, R.M. Kaake, et al., The global phosphorylation landscape of SARS-CoV-2 infection, *Cell* 182 (685–712) (2020), e619.
- X. Ji, Y. Zhou, S. Pandit, J. Huang, H. Li, C.Y. Lin, R. Xiao, C.B. Burge, X.D. Fu, SR proteins collaborate with 7SK and promoter-associated nascent RNA to release paused polymerase, *Cell* 153 (2013) 855–868.
- R.P. McNamara, C. Guzman, J.E. Reeder, I. D'Orso, Genome-wide analysis of KAP1, the 7SK snRNP complex, and RNA polymerase II, *Genom Data* 7 (2016) 250–255.
- M.D. Lavigne, D. Konstantopoulos, K.Z. Ntakou-Zamplara, A. Liakos, M. Fousteri, Global unleashing of transcription elongation waves in response to genotoxic stress restricts somatic mutation rate, *Nat. Commun.* 8 (2017) 2076.
- C. Studniarek, M. Tellier, P.G.P. Martin, S. Murphy, T. Kiss, S. Egloff, The 7SK/P-TEFb snRNP controls ultraviolet radiation-induced transcriptional reprogramming, *Cell Rep.* 35 (2021), 108965.
- M. Blasius, S.A. Wagner, C. Choudhary, J. Bartek, S.P. Jackson, A quantitative 14-3-3 interaction screen connects the nuclear exosome targeting complex to the DNA damage response, *Genes Dev.* 28 (2014) 1977–1982.
- M.E. Borisova, A. Voigt, M.A.X. Tollenaere, S.K. Sahu, T. Juretschke, N. Kreim, N. Mailand, C. Choudhary, S. Bekker-Jensen, M. Akutsu, et al., p38-MK2 signaling axis regulates RNA metabolism after UV-light-induced DNA damage, *Nat. Commun.* 9 (2018) 1017.
- H.C. Reinhardt, A.S. Aslanian, J.A. Lees, M.B. Yaffe, p53-deficient cells rely on ATM- and ATR-mediated checkpoint signaling through the p38MAPK/MK2 pathway for survival after DNA damage, *Cancer Cell* 11 (2007) 175–189.

- [47] H. Sobhy, R. Kumar, J. Lewerentz, L. Lizana, P. Stenberg, Highly interacting regions of the human genome are enriched with enhancers and bound by DNA repair proteins, *Sci. Rep.* 9 (2019) 4577.
- [48] S.T. Younger, J.L. Rinn, p53 regulates enhancer accessibility and activity in response to DNA damage, *Nucleic Acids Res.* 45 (2017) 9889–9900.
- [49] C.A. Higgins, B.E. Nilsson-Payant, A.P. Kurland, P. Adhikary, I. Golynger, O. Danziger, M. Panis, B.R. Rosenberg, B. Oever, J.R. Johnson, SARS-CoV-2 hijacks p38<sup>MAPK1</sup> to promote viral protein translation, *bioRxiv* (2021) (2021.2008.2020.457146).
- [50] M.P. Gustafson, M. Welcker, H.C. Hwang, B.E. Clurman, Zcchc8 is a glycogen synthase kinase-3 substrate that interacts with RNA-binding proteins, *Biochem. Biophys. Res. Commun.* 338 (2005) 1359–1367.
- [51] M. Lubas, M.S. Christensen, M.S. Kristiansen, M. Domanski, L.G. Falkenby, S. Lykke-Andersen, J.S. Andersen, A. Dziembowski, T.H. Jensen, Interaction profiling identifies the human nuclear exosome targeting complex, *Mol. Cell* 43 (2011) 624–637.
- [52] M. Hallais, F. Pontvianne, P.R. Andersen, M. Clerici, D. Lener, H. Benbahouche Nel, T. Gostan, F. Vandermoere, M.C. Robert, S. Cusack, et al., CBC-ARS2 stimulates 3'-end maturation of multiple RNA families and favors cap-proximal processing, *Nat. Struct. Mol. Biol.* 20 (2013) 1358–1366.
- [53] M. Lubas, P.R. Andersen, A. Schein, A. Dziembowski, G. Kudla, T.H. Jensen, The human nuclear exosome targeting complex is loaded onto newly synthesized RNA to direct early ribonucleolysis, *Cell Rep.* 10 (2015) 178–192.
- [54] C. Kilchert, S. Wittmann, L. Vasiljeva, The regulation and functions of the nuclear RNA exosome complex, *Nat Rev Mol Cell Biol* 17 (2016) 227–239.
- [55] C. Tiedje, M. Lubas, M. Tehrani, M.B. Menon, N. Ronkina, S. Rousseau, P. Cohen, A. Kotlyarov, M. Gaestel, p38MAPK/MK2-mediated phosphorylation of RBM7 regulates the human nuclear exosome targeting complex, *RNA* 21 (2015) 262–278.
- [56] D. Hrossova, T. Sikorsky, D. Potesil, M. Bartosovic, J. Pasulka, Z. Zdrahal, R. Stefl, S. Vanacova, RBM7 subunit of the NEXT complex binds U-rich sequences and targets 3'-end extended forms of snRNAs, *Nucleic Acids Res.* 43 (2015) 4236–4248.
- [57] C. Maris, C. Dominguez, F.H. Allain, The RNA recognition motif, a plastic RNA-binding platform to regulate post-transcriptional gene expression, *FEBS J.* 272 (2005) 2118–2131.
- [58] S. Falk, K. Finogenova, M. Melko, C. Benda, S. Lykke-Andersen, T.H. Jensen, E. Conti, Structure of the RBM7-ZCCHC8 core of the NEXT complex reveals connections to splicing factors, *Nat. Commun.* 7 (2016) 13573.
- [59] N. Sofos, M.B. Winkler, D.E. Brodersen, RRM domain of human RBM7: purification, crystallization and structure determination, *Acta Crystallogr F Struct Biol Commun* 72 (2016) 397–402.
- [60] M.R. Puno, C.D. Lima, Structural basis for RNA surveillance by the human nuclear exosome targeting (NEXT) complex, *Cell* 185 (2132–2147) (2022), e2126.
- [61] P. Gerlach, W. Garland, M. Lingaraju, A. Salerno-Kochan, F. Bonneau, J. Basquin, T.H. Jensen, E. Conti, Structure and regulation of the nuclear exosome targeting complex guides RNA substrates to the exosome, *Mol. Cell* 82 (2505–2518) (2022), e2507.
- [62] J.F. Watson, J. Garcia-Nafria, In vivo DNA assembly using common laboratory bacteria: A re-emerging tool to simplify molecular cloning, *J. Biol. Chem.* 294 (2019) 15271–15281.
- [63] S. Raran-Kurussi, S. Cherry, D. Zhang, D.S. Waugh, Removal of affinity tags with TEV protease, *Methods Mol. Biol.* 1586 (2017) 221–230.
- [64] S. Duvaud, C. Gabella, F. Lisacek, H. Stockinger, V. Ioannidis, C. Durinx, Expassy, the Swiss bioinformatics resource portal, as designed by its users, *Nucleic Acids Res.* 49 (2021) W216–W227.
- [65] E. Gasteiger, A. Gattiker, C. Hoogland, I. Ivanyi, R.D. Appel, A. Bairoch, ExPASy: the proteomics server for in-depth protein knowledge and analysis, *Nucleic Acids Res.* 31 (2003) 3784–3788.
- [66] Y. Shimizu, A. Inoue, Y. Tomari, T. Suzuki, T. Yokogawa, K. Nishikawa, T. Ueda, Cell-free translation reconstituted with purified components, *Nat. Biotechnol.* 19 (2001) 751–755.
- [67] M.R. Green, J. Sambrook, Isolation of DNA fragments from polyacrylamide gels by the crush and soak method, *Cold Spring Harbor Protocols* 2 (2019) (pdb. prot100479).
- [68] P.R. Romero, N. Kobayashi, J.R. Wedell, K. Baskaran, T. Iwata, M. Yokochi, D. Mazziuk, H. Yao, T. Fujiwara, G. Kuruusu, et al., BioMagResBank (BMRB) as a resource for structural biology, *Methods Mol. Biol.* 2112 (2020) 187–218.
- [69] D.G. Reid, L.K. MacLachlan, A.J. Edwards, J.A. Hubbard, P.J. Sweeney, Introduction to the NMR of proteins, *Methods Mol. Biol.* 60 (1997) 1–28.
- [70] J. Cavanagh, *Protein NMR Spectroscopy: Principles and Practice*, 2nd ed., Academic Press, Amsterdam, Boston, 2007.
- [71] Y. Shen, F. Delaglio, G. Cornilescu, A. Bax, TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts, *J. Biomol. NMR* 44 (2009) 213–223.
- [72] N.A. Farrow, O. Zhang, J.D. Forman-Kay, L.E. Kay, A heteronuclear correlation experiment for simultaneous determination of <sup>15</sup>N longitudinal decay and chemical exchange rates of systems in slow equilibrium 5th, 4, *J. Biomol. NMR*, 1994, pp. 727–734.
- [73] E.E. Metcalfe, J. Zamoan, D.D. Thomas, G. Veglia, <sup>1</sup>H/<sup>15</sup>N heteronuclear NMR spectroscopy shows four dynamic domains for Phospholamban reconstituted in Dodecylphosphocholine micelles, *Biophys. J.* 87 (2004) 1205–1214.
- [74] F. Delaglio, S. Grzesiek, G.W. Vuister, G. Zhu, J. Pfeifer, A. Bax, NMRPipe: a multidimensional spectral processing system based on UNIX pipes, *J. Biomol. NMR* 6 (1995) 277–293.
- [75] W. Lee, M. Tonelli, J.L. Markley, NMRFAM-SPARKY: enhanced software for biomolecular NMR spectroscopy, *Bioinformatics* 31 (2015) 1325–1327.
- [76] M.W. Maciejewski, A.D. Schuyler, M.R. Gryk, I.I. Moraru, P.R. Romero, E.L. Ulrich, H.R. Eghbalnia, M. Livny, F. Delaglio, J.C. Hoch, NMRbox: A resource for biomolecular NMR computation, *Biophys. J.* 112 (2017) 1529–1534.
- [77] A. Busch, M. Bruggemann, S. Ebersberger, K. Zarnack, iCLIP data analysis: A complete pipeline from sequencing reads to RBP binding sites, *Methods* 178 (2020) 49–62.
- [78] I. Huppertz, J. Attig, A. D'Ambrogio, L.E. Easton, C.R. Sibley, Y. Sugimoto, M. Tajnik, J. Konig, J. Ule, iCLIP: protein-RNA interactions at nucleotide resolution, *Methods* 65 (2014) 274–287.
- [79] <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>.
- [80] [https://www.bioinformatics.babraham.ac.uk/projects/trim\\_galore/](https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/).
- [81] M. Briese, N. Haberman, C.R. Sibley, R. Faraway, A.S. Elser, A.M. Chakrabarti, Z. Wang, J. Konig, D. Perera, V.O. Wickramasinghe, et al., A systems view of spliceosomal assembly and branchpoints with iCLIP, *Nat. Struct. Mol. Biol.* 26 (2019) 930–940.
- [82] <https://github.com/alexdobin/STAR>.
- [83] A.R. Quinlan, I.M. Hall, BEDTools: a flexible suite of utilities for comparing genomic features, *Bioinformatics* 26 (2010) 841–842.
- [84] K. Darty, A. Denise, Y. Ponty, VARNA: interactive drawing and editing of the RNA secondary structure, *Bioinformatics* 25 (2009) 1974–1975.
- [85] M.R. Puno, C.D. Lima, Structural basis for MTR4-ZCCHC8 interactions that stimulate the MTR4 helicase in the nuclear exosome-targeting complex, *Proc. Natl. Acad. Sci. U. S. A.* 115 (2018) E5506–E5515.
- [86] F. Baneyx, M. Mujacic, Recombinant protein folding and misfolding in *Escherichia coli*, *Nat. Biotechnol.* 22 (2004) 1399–1408.
- [87] G. Lipari, A. Szabo, Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 1. Theory and range of validity, *J. Am. Chem. Soc.* 104 (1982) 4546–4559.
- [88] S.W. Olson, A.W. Turner, J.W. Arney, I. Saleem, C.A. Weidmann, D.M. Margolis, K. M. Weeks, A.M. Mustoe, Discovery of a large-scale, cell-state-responsive allosteric switch in the 7SK RNA using DANCE-MaP, *Mol. Cell* 82 (1708–1723) (2022), e1710.
- [89] L. Luo, L.Y. Chiu, A. Sugarman, P. Gupta, S. Rouskin, B.S. Tolbert, HnRNP A1/A2 proteins assemble onto 7SK snRNA via context dependent interactions, *J. Mol. Biol.* 433 (2021), 166885.
- [90] R. Van Damme, K. Li, M. Zhang, J. Bai, W.H. Lee, J.D. Yesselman, Z. Lu, W. A. Velema, Chemical reversible crosslinking enables measurement of RNA 3D distances and alternative conformations in cells, *Nat. Commun.* 13 (2022) 911.
- [91] G.M. Daubner, A. Clery, F.H. Allain, RRM-RNA recognition: NMR or crystallography...and new findings, *Curr. Opin. Struct. Biol.* 23 (2013) 100–108.
- [92] N. Bhattacharjee, P. Biswas, Position-specific propensities of amino acids in the beta-strand, *BMC Struct. Biol.* 10 (2010) 29.
- [93] S.R. Trevino, S. Schaefer, J.M. Scholtz, C.N. Pace, Increasing protein conformational stability by optimizing beta-turn sequence, *J. Mol. Biol.* 373 (2007) 211–218.
- [94] M. Giunta, S. Edvardson, Y. Xu, M. Schuelke, A. Gomez-Duran, V. Boczonadi, O. Elpeleg, J.S. Muller, R. Horvath, Altered RNA metabolism due to a homozygous RBM7 mutation in a patient with spinal motor neuropathy, *Hum. Mol. Genet.* 25 (2016) 2985–2996.
- [95] D.J. Morton, E.G. Kuiper, S.K. Jones, S.W. Leung, A.H. Corbett, M.B. Fasken, The RNA exosome and RNA exosome-linked disease, *RNA* 24 (2018) 127–142.
- [96] C. Barrandon, F. Bonnet, V.T. Nguyen, V. Labas, O. Bensaude, The transcription-dependent dissociation of P-TEFb-HEXIM1-7SK RNA relies upon formation of hnRNP-7SK RNA complexes, *Mol. Cell. Biol.* 27 (2007) 6996–7006.