



HAL
open science

Analysis of RNA binding properties of human Ku protein reveals its interactions with 7SK snRNA and protein components of 7SK snRNP complex

Olga Shadrina, Irina Garanina, Sergey Korolev, Timofei Zatsepin, Jeanne van Assche, Fadoua Daouad, Clémentine Wallet, Olivier Rohr, Marina Gottikh

► **To cite this version:**

Olga Shadrina, Irina Garanina, Sergey Korolev, Timofei Zatsepin, Jeanne van Assche, et al.. Analysis of RNA binding properties of human Ku protein reveals its interactions with 7SK snRNA and protein components of 7SK snRNP complex. *Biochimie*, 2020, 10.1016/j.biochi.2020.02.016 . hal-02497902

HAL Id: hal-02497902

<https://hal.science/hal-02497902v1>

Submitted on 3 Mar 2020

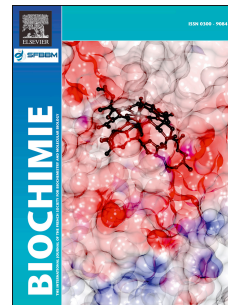
HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Journal Pre-proof

Analysis of RNA binding properties of human Ku protein reveals its interactions with 7SK snRNA and protein components of 7SK snRNP complex

Olga Shadrina, Irina Garanina, Sergey Korolev, Timofei Zatsepin, Jeanne Van Assche, Fadoua Daouad, Clementine Wallet, Olivier Rohr, Marina Gottikh



PII: S0300-9084(20)30044-4

DOI: <https://doi.org/10.1016/j.biochi.2020.02.016>

Reference: BIOCHI 5845

To appear in: *Biochimie*

Received Date: 20 November 2019

Accepted Date: 21 February 2020

Please cite this article as: O. Shadrina, I. Garanina, S. Korolev, T. Zatsepin, J. Van Assche, F. Daouad, C. Wallet, O. Rohr, M. Gottikh, Analysis of RNA binding properties of human Ku protein reveals its interactions with 7SK snRNA and protein components of 7SK snRNP complex, *Biochimie*, <https://doi.org/10.1016/j.biochi.2020.02.016>.

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

© 2020 Published by Elsevier B.V.

Abstract

Human Ku heterodimeric protein composed of Ku70 and Ku80 subunits plays an important role in the non-homologous end-joining DNA repair pathway as a sensor of double strand DNA breaks. Ku is also involved in numerous cellular processes, and in some of them it acts in an RNA-dependent manner. However, RNA binding properties of the human Ku have not been well studied. Here we have analyzed interactions of a recombinant Ku heterodimer with a set of RNAs of various structure as well as eCLIP (enhanced crosslinking and immunoprecipitation) data for human Ku70. As a result, we have proposed a consensus RNA structure preferable for the Ku binding that is a hairpin possessing a bulge just near GpG sequence-containing terminal loop. 7SK snRNA is a scaffold for a ribonucleoprotein complex (7SK snRNP), which is known to participate in transcription regulation. We have shown that the recombinant Ku specifically binds a G-rich loop of hairpin 1 within 7SK snRNA. Moreover, Ku protein has been co-precipitated from HEK 293T cells with endogenous 7SK snRNA and such proteins included in 7SK snRNP as HEXIM1, Cdk9 and CTIP2. Ku and Cdk9 binding is found to be RNA-independent, meanwhile HEXIM1 and Ku co-precipitation depended on the presence of intact 7SK snRNA. The latter result has been confirmed using recombinant HEXIM1 and Ku proteins. Colocalization of Ku and CTIP2 was additionally confirmed by confocal microscopy. These results allow us to propose human Ku as a new component of the 7SK snRNP complex.

Analysis of RNA binding properties of human Ku protein reveals its interactions with 7SK snRNA and protein components of 7SK snRNP complex

Olga Shadrina^{1,2*}, *Irina Garanina*³, *Sergey Korolev*^{2,4}, *Timofei Zatsepin*^{2,5}, *Jeanne Van Assche*⁶,
*Fadoua Daouad*⁶, *Clementine Wallet*⁶, *Olivier Rohr*⁶, *Marina Gottikh*^{2,4}

¹Faculty of Bioengineering and Bioinformatics, Lomonosov Moscow State University,
Moscow, 119234, Russia;

²Chemistry Department, Lomonosov Moscow State University, Moscow, 199991, Russia;

³Federal Research and Clinical Center of Physical-Chemical Medicine of Federal Medical
Biological Agency, Moscow, 119435, Russia;

⁴Belozersky Institute of Physico-Chemical Biology, Lomonosov Moscow State University,
119992, Moscow; Russia;

⁵Skolkovo Institute of Science and Technology, Skolkovo, 121205, Russia;

⁶Université de Strasbourg, EA7292, FMTS, IUT Louis Pasteur, Schiltigheim, France.

*Corresponding author: oashadrina92@gmail.com

Abstract

Human Ku heterodimeric protein composed of Ku70 and Ku80 subunits plays an important role in the non-homologous end-joining DNA repair pathway as a sensor of double-strand DNA breaks. Ku is also involved in numerous cellular processes, and in some of them, it acts in an RNA-dependent manner. However, RNA binding properties of the human Ku have not been well studied. Here we have analyzed interactions of a recombinant Ku heterodimer with a set of various RNA structures as well as eCLIP (enhanced crosslinking and immunoprecipitation) data for human Ku70. As a result, we have proposed a consensus RNA structure preferable for the Ku binding that is a hairpin possessing a bulge just near GpG sequence-containing terminal loop. 7SK snRNA is a scaffold for a ribonucleoprotein complex (7SK snRNP), which is known to participate in transcription regulation. We have shown that the recombinant Ku specifically binds a G-rich loop of hairpin 1 within 7SK snRNA. Moreover, Ku protein has been co-precipitated from HEK 293T cells with endogenous 7SK snRNA and such proteins included in 7SK snRNP as HEXIM1, Cdk9, and CTIP2. Ku and Cdk9 binding is found to be RNA-independent, meanwhile HEXIM1 and Ku co-precipitation depended on the presence of intact 7SK snRNA. The latter result has been confirmed using recombinant HEXIM1 and Ku proteins.

Colocalization of Ku and CTIP2 was additionally confirmed by confocal microscopy. These results allow us to propose human Ku as a new component of the 7SK snRNP complex.

Highlights

Human Ku interacts with bulge-containing stem-loop RNA structures

Ku binds 7SK snRNA at the terminal G-rich loop of its hairpin 1

Ku interacts with proteins of 7SK snRNP: HEXIM1, Cdk9 and CTIP2

Keywords

Human Ku protein, RNA binding, 7SK snRNA, 7SK ribonucleoprotein complex, HEXIM1

1. Introduction

The Ku protein is a highly abundant protein in mammalian cells made up by two subunits, Ku70 and Ku80 [1]. The most well-known Ku function is participation in the repair of DNA double-strand breaks (DSBs) through the non-homologous end joining (NHEJ) pathway where the Ku heterodimer acts as the main sensor of DSBs. Being bound to DNA ends, Ku recruits the catalytic subunit of DNA-dependent protein kinase (DNA-PKcs) and forms the trimeric complex DNA-PK, which triggers a cascade of phosphorylation events required for the subsequent DSB repair [2–4]. In addition to NHEJ, Ku is involved in various cellular processes such as V(D)J recombination, AP-site repair, telomere maintenance, apoptosis, transcription, and translation [1,5–9]. Its exact function in these processes remains the subject of intensive research. In particular, one cannot rule out a possibility of RNA-mediated participation of Ku in some of these processes. Indeed, even in the well-studied DSB repair, a direct role of RNA has been revealed in recent works [10–12]. In particular, the noncoding RNA LINP1 acts as a scaffold that links Ku and DNA-PKcs and enables efficient DNA repair through the NHEJ pathway in breast cancer cells [12].

Accumulating evidence demonstrates that Ku acts as an RNA binding protein in telomere-length maintenance in yeast [13–17] and in human cells [18]. There is no sequence similarity between the RNA components of yeast and human telomerases, but in both RNAs Ku associates with similar stem-loop structures [14,18]. The interaction between the Ku heterodimer and telomerase RNA (TLC1) from *Saccharomyces cerevisiae* is required for proper recruitment of yeast telomerase to the chromosome end and telomere synthesis [15,19].

Ku participation in translation can also be mediated by its interaction with RNA. For example, Ku70 is found to bind the 5'-UTR elements of mRNAs of some growth factors, and this finding provides data to propose a role for Ku in the modulation of IRES-mediated mRNA

translation [20]. Both Ku subunits are identified in a screen for proteins' binding to the 5'-UTR of the mRNA encoding the tumor suppressor protein p53, a master gene regulator of the DNA damage response [21]. The Ku binding with p53 mRNA results in repression of p53 protein synthesis and p53-mediated apoptosis and the translational repression is relieved after genotoxic stress due to Ku acetylation, which disrupts Ku-p53 mRNA interactions [6].

Another important cellular process where Ku participation might be mediated by its binding to RNA is alien DNA sensing and innate immune response. Several works have shown that Ku (alone or in its complex with DNA-PKcs) can activate an innate immune response when it detects foreign DNA [22–25]. Recently, it has been found that Ku is involved in DNA-mediated innate immune response being a component of a multi-subunit RNP complex, in which several proteins are associated with a long non-coding RNA NEAT1 [26]. This RNA is known as a paraspeckle marker [27], and in addition to Ku70, Ku80 and DNA-PKcs, this RNP complex also contains paraspeckle proteins (SFPQ, NONO, PSPC1, RBM14, and MATRIN3). Another interesting component of the RNP is hexamethylene-bis-acetamide-inducible protein 1 (HEXIM1), an RNA-binding protein involved in the control of transcription elongation by RNA polymerase II (RNAPII) [28,29]. Among all the proteins listed above, only SFPQ, HEXIM1, and Ku70 have been found to be co-localized with the NEAT1 accordingly to immuno-fluorescence *in situ* hybridization in HeLa cells. This data allows supposing a direct interaction of Ku70 with RNA NEAT1.

It should also be noted that Ku has been found to associate with RNA helicase A (RHA) [30], a multifunctional protein, which is involved in several steps of RNA metabolism, such as RNA processing, cellular transit of viral molecules, ribosome assembly, regulation of transcription, and translation of specific mRNAs [31]. Interestingly, Ku was immunoprecipitated with RHA in an RNA-dependent manner, and both proteins were identified in hnRNP complexes [30]. Moreover, in these complexes, DNA-PK consisting of Ku protein and a kinase subunit (DNA-PKcs) was able to perform RNA-dependent phosphorylation of hnRNP proteins and of RHA [30]. It was a surprising result since previously DNA-PK had been believed to exhibit its phosphorylating activity only after binding to DNA, but not RNA [32,33]. However, later it was shown that DNA-PK could phosphorylate heterogeneous nuclear ribonucleoprotein A1 (hnRNPA1), and its phosphorylation is stimulated by the presence of both DNA and the telomerase RNA component, hTR [34].

Participation of Ku protein in the transcription regulation of cellular and viral genes has been extensively studied [35–40], and no uniform mechanism has been revealed. Nothing is known about Ku interactions with any RNA during transcription regulation. However, Ku has been found to interact with TAR (trans-activation response) RNA [41] forming a hairpin

structure located at the 5'-end of the synthesized mRNA and known as a key player in active transcription from the HIV-1 promoter [42–44]. Therefore, one cannot rule out that at least the viral transcription regulation by Ku could be mediated by some Ku-RNA interplays.

The structural aspects of Ku-RNA binding have been well studied for the yeast Ku protein, and a minimum consensus structure for a Ku-binding site within RNAs has been proposed [16]. It represents a bulged stem-loop structure, and the presence of UA or UG bulge, and the size of the terminal loop, but not the loop sequence are the most important marks for Ku recognition. For *Saccharomyces cerevisiae*, it has been shown that Ku defective in DNA binding also loses its affinity to telomerase RNA (TLC1) [13]. This result allows considering that the yeast Ku protein has the same binding site for both types of nucleic acids. Interaction of Ku from mammalian cells with different RNA structures is less studied, but the existing data show that the human protein also prefers stem-loop RNAs containing a bulge [6,18]. Analysis of Ku binding with TAR RNA demonstrates that, in contrast to the yeast protein, the loop sequence may be important for recognition by human Ku [41]. Besides, a series of RNA aptamers for Ku heterodimer has been developed [33], however, the nature of their binding to Ku has not been clarified.

Here we have analyzed structural aspects of RNA interactions with human Ku heterodimer based on the results of eCLIP (enhanced crosslinking and immunoprecipitation) experiments available from ENCODE database, and our experimental study of recombinant Ku binding with different RNA molecules. As a result, we could propose a consensus RNA structure favorable for the Ku binding, which is a hairpin containing a bulge within a few nucleotides from the loop and possessing GpG sequence in the terminal loop. Also, we investigated Ku interplay with a non-coding small nuclear 7SK RNA (7SK snRNA), which is a scaffold for a ribonucleoprotein complex (7SK snRNP) participating in transcription regulation [45]. We have found that Ku binds mainly the hairpin 1 containing the most G-rich loop in the complex secondary structure of 7SK snRNA. Interestingly, the hairpin 1 is involved in 7SK snRNA interaction with HEXIM1 protein, an RNA-binding protein well-known as an inhibitor of the positive transcription elongation factor b (P-TEFb), which controls transcription elongation by RNAPII [29,46]. Of note, Ku does not compete with HEXIM1 for binding to 7SK snRNA, since we detected formation of a ternary complex composed of both proteins and 7SK snRNA. Ku has been co-precipitated with HEXIM1 from cells, however RNase treatment destroyed HEXIM1-Ku complex. Moreover, Ku interacts with some other proteins of the 7SK snRNP complex, namely cyclin-dependent kinase 9 (Cdk9), which is the part of P-TEFb, and the cellular repressor CTIP2. It is known that 7SK snRNA links HEXIM1 and CTIP2 to Cdk9 and CyclinT1 in an inactive

form of the P-TEFb complex [47]. Basing on the results obtained we can consider Ku protein as a new component of the 7SK snRNP complex.

2. Materials and methods

2.1. Plasmids

For the construction of heterodimer Ku expression, vector pET15b_His6_Ku70_(-1nt)_Ku80, PCR with primers Ku70-N-NdeI (5'-CATATGTCAGGGTGGGAGTCATATTACA-3') and Ku70-C-BamHI (5'-GGATCCGTCCTGGAAGTGCTTGGTGA-3') was used to amplify the human Ku70 coding sequence (corresponded to NP_001275905.1) from pGEX-6p-1-Ku70 without a stop codon [48], and the product predigested with NdeI and BamHI restriction endonucleases (Thermo Fisher Scientific, USA) was cloned between NdeI and BamHI sites into a pET15b vector (Merk Millipore, USA). A new XhoI restriction site was added following by the BamHI site in the resulted plasmid by site-directed mutagenesis (Quick Change II Site-Directed Mutagenesis Kit, Agilent Technologies, USA) with primers (5'-AACAAAGCCCGAAAGGAACTCGAGTTGGCTGCTGCCACC-3' and 5'-GGTGGCAGCAGCCAACTCGAGTTCCTTTCGGGCTTTGTT-3'). Then, human Ku80 coding sequence (corresponded to NP_066964.1) was amplified from a pCDNA3_Ku80_3xFLAG vector (described in [49]) with primers Ku80-N-BamHI (5'-GGATCCTTAAAAGAGGAATTATAATGGTGCGGTCGGGGAATAA-3') and Ku80-C-XhoI (5'-CTCGAGTTATATCATGTCCAATAAATCGTCCAC-3') and cloned between BamHI and XhoI restriction sites into the vector pET15b-Ku70, obtained at the previous step. The primer Ku80-N-BamHI introduced the Shine-Dalgarno sequence GAGG, the intercistronic sequence AATT, and a stop codon at the 3'-end of the Ku70 gene after the ligation of the PCR product into the vector for the optimal translation efficiency of the Ku80 gene located with -1nt shift relative to the stop codon of the Ku70 gene [50].

To obtain the sequence coding the human HEXIM1 protein, total RNA from HEK 293T cells was extracted by TRIzol Reagent (Thermo Fisher Scientific, USA) according to the manufacturer's protocol and cDNA was synthesized using the MMLV kit (Evrogen, Russia). Then, the HEXIM1 coding sequence (corresponded to NP_006451.1) was amplified from cDNA by PCR using primers 5'-ACTTCGAGGACTCTACTAGCC-3' and 5'-GTTTCAGTCTAGTCTCCAACCTTGG-3' and verified by Sanger sequencing. For the construction of the GST-HEXIM1 expression vector (pGEX-6p-1-hHEXIM1), HEXIM1 coding sequence was amplified using primers HEXIM1-N-BglII (5'-ATGCAGATCTATGGCCGAGCCATTCTTG-3') and HEXIM1-C-XhoI (5'-

Journal Pre-proof

ATGCCTCGAGTTACTAGTCTCCAAACTTGGAAAGCGG-3') and cloned into the pGEX-6p-1 vector between BglIII and XhoI restriction sites in frame with GST tag. The vector for mammalian expression of human HEXIM1 protein (pcDNA3_HEXIM1_3xFLAG) was obtained by cloning the HEXIM1 coding sequence into the AflIII restriction site after the Ku70 gene was cut out of pcDNA3_Ku70_3xFLAG (used in [49]) by the AflIII restriction enzyme.

pNTap-CTIP2 and RFP-CTIP2 plasmids were described previously [51].

2.2. Expression and purification of recombinant proteins

For recombinant human His6-Ku70/Ku80 production, *E.coli* BL21(DE3) Codon Plus/pLysS strain was transformed with plasmid pET15b_His6_Ku70_(-1nt)_Ku80, grown in 2 L of modified 2YT medium (16 g/L bactotryptone, 10 g/L yeast extract, 1 g/L K₂HPO₄, 1.5 g/L NaH₂PO₄, 1% glycerol, pH 7.5) in the presence of 100 mg/L ampicillin (Sigma, USA) at 37 °C until the culture reached OD₆₀₀ = 0.6. Recombinant protein expression was induced by adding IPTG water solution up to 0.1 mM, and bacteria were cultured at 18°C. After 6 h of cultivation, lactose in the concentration of up to 5 g/l was added. The total cultivation time was 20 h at 18°C. Cells were harvested by centrifugation (20 min, 4 °C, 4000 g), and 50 g of cellular debris was resuspended in 1 L of lysis buffer (50 mM Tris-HCl, pH 8.0, 3 mM 2-mercaptoethanol, 1 M NaCl, 5% sucrose, 0.05% TritonX-100, 2mM PMSF) with 1 g/L lysozyme and incubated on ice for 30 min. After sonication, the insoluble material was removed by centrifugation (40 min, 4 °C, 35000 g), the supernatant was purified from bacterial DNA by incubation with 0.05% polyethylenimine (pH 5.0) on ice for 30 min, and centrifuged 45 min, 4°C, 35000 g. The lysate was loaded on Ni-NTA-agarose beads (Thermo Fisher Scientific, USA), and incubated overnight at 4°C. After binding, the resin was washed with a buffer (50 mM Tris-HCl, pH 8.0, 2 mM 2-mercaptoethanol, 0.5 M NaCl, 5% glycerol, 0.05% TritonX-100, 1 mM PMSF), and a buffer (50 mM Tris-HCl, pH 8.0, 2 mM 2-mercaptoethanol, 0.1 M NaCl, 5% glycerol, 50 mM imidazole, 1 mM PMSF) and finally eluted with a buffer (50 mM Tris-HCl, pH 8.0, 2 mM 2-mercaptoethanol, 0.1 M NaCl, 10% glycerol, 500 mM imidazole, 1 mM PMSF). The fraction of the eluted protein was diluted by 50 mM Tris-HCl, pH 8.0 up to 50 mM NaCl concentration and then loaded on the heparin-sepharose (GE Healthcare, USA), washed with a buffer (50 mM Tris-HCl, pH 8.0, 2 mM 2-mercaptoethanol, 10% glycerol, 1 mM PMSF, 50 mM NaCl), and eluted with this buffer containing 400 mM NaCl.

For recombinant human HEXIM1 production, *E.coli* BL21(DE3) Codon Plus/pLysS strain was transformed with plasmid pGEX-6p-1-h-HEXIM1, grown in 2 L of 2% LB medium in the presence of 100 mg/L ampicillin (Sigma, USA) at 37 °C until the culture reached OD₆₀₀ =

0.6. Recombinant protein expression was induced by adding IPTG up to 0.1 mM and bacteria were cultured at 25°C for 3 hours. Cells were harvested by centrifugation (20 min, 4 °C, 4000 g) and resuspended in a lysis buffer (50 mM Tris-HCl, pH 7.8, 0.5 M NaCl, 0.1% TritonX-100, 0.1 mM PMSF) containing 1 g/L lysozyme, and incubated on ice for 30 min. After sonication, the insoluble material was removed by centrifugation (40 min, 4 °C, 35000 g), and the lysate was loaded on Pierce®Glutathione Agarose (Thermo Fisher Scientific, USA). After overnight incubation at 4°C, the resin was washed with 50 mL of lysis buffer and 50 mL volumes of a wash buffer (50 mM Tris-HCl, pH 7.8, 0.25 M NaCl, 2 mM 2-mercaptoethanol, 5% glycerol, 0.1 mM EDTA). The protein was eluted with 1 ml of the wash buffer containing 50 mM glutathione and was dialyzed against the wash buffer using SnakeSkin Dialysis Tubing (Thermo Fisher Scientific, USA) overnight.

Protein concentrations were measured by Bradford assay (Thermo Fisher Scientific, USA) using BSA (Thermo Fisher Scientific, USA) as a protein standard.

2.3. Western blot

Protein samples were separated using 10% SDS PAGE and analyzed by western blot with anti-His6 antibodies (produced in mouse, SAB1305538, Sigma, USA), anti-Ku70 rabbit polyclonal antibody (ab83502, Abcam, USA), anti-Ku80 rabbit polyclonal antibody (#2735, Cell Signaling, USA). HRP-conjugated anti-rabbit (produced in mouse, A6154, Sigma, USA) and anti-mouse antibodies (produced in goat, A9044, Sigma, USA) were used as secondary antibodies. Visualization of specific protein bands was performed with Clarity Western ECL substrate (Bio-Rad, USA) using ChemiDoc MP system (Biorad, USA).

2.4. Mass-spectrometry

The protein samples were prepared for MS identification and analyzed as described [52]. Mass-spectra were obtained on UltrafleXtreme Bruker Daltonics MALDI-TOF mass-spectrometer, analyzed by FlexAnalysis 3.3 (Bruker Daltonics, Germany), protein identification was performed using Mascot (www.matrixscience.com) (Table S1).

2.5. ³²P-labeled nucleic acid preparation

All oligonucleotides were synthesized using the phosphoramidite method on a Mermaid 12 synthesizer (Bioautomation, USA) by the standard operating procedures using commercially available reagents (Glen Research, USA). Oligonucleotides were double purified by denaturing PAGE, and RP-HPLC and their purity was confirmed by LC-MS. Purified DNA and RNA ligands were labeled with T4 polynucleotide kinase (Thermo Fisher Scientific, USA) and 50 μCi

Journal Pre-proof
[γ -³²P]-ATP (3000 Ci/mmol) for 30 min at 37 °C. T4 polynucleotide kinase was inactivated by the addition of EDTA followed by phenol-chloroform extraction. The resulting ligands were purified from unincorporated [γ -³²P]-ATP by 15% denaturing PAGE.

2.6. *In vitro* transcription

In vitro transcription was performed using MEGAscript™ T7 Transcription kit (Thermo Fisher Scientific, USA) as recommended by the manufacturer with minor modifications: reaction mixture contained additional 9 mM guanosine for subsequent ³²P labeling of 5' end of RNA. Primers 5'-TAATACGACTCACTATAGGATGTGAGGGCGATCTG-3', containing the sequence of T7 promoter, and 5'-AAAGAAAGGCAGACTGCC-3' were used for the amplification by PCR of 7SK template for *in vitro* transcription. The sequence of human 7SK snRNA (332 bases long, RefSeq NR_001445) was prepared from HEK 293T cells genome by PCR using primers 5'-CACCAATGGAGACTGCAGTATTTAGCA-3' and 5'-GCATCTAACTTAGATGGGTAATGGGTCAAAG-3'.

2.7. Nucleic acid binding assays

For electrophoretic mobility shift assays (EMSAs), ³²P-labeled nucleic acids (0.1 nM) were annealed and then incubated with increasing concentrations of Ku protein (0-600 nM) in 20 μ L of a buffer (50 mM Tris-HCl pH 8.0, 100 mM NaCl, 2 mM DTT, 1 mM EDTA, 3% glycerol, 5 μ g/ml yeast tRNA, 0.1 mg/ml BSA) for 20 min at room temperature (RT) and subsequently for 10 min on ice. Then, the resulting complexes were separated using 5% pre-run native PAGE at 4°C for 3 h at 60 V in 1xTBE buffer and visualized by autoradiography (³²P). For the super-shift assay, the reaction mixture containing 0.1 nM ³²P-labeled TAR RNA and 50 nM Ku protein was supplemented with 1 μ l of either anti-Ku70, anti-Ku80, anti-His6, and anti-actin (A2103, Sigma) antibody and incubated for 30 min at RT.

For competition experiments, a double-stranded 40-mer DNA duplex with the sense sequence 5'-GACTACGGTTCAAGTCAGCGTGTGGAAAATCTCTAGCAGT-3' was used. ³²P-labeled 0.1 nM hp1RNA was incubated with 100 nM Ku protein in the reaction buffer in the presence of increasing concentrations of the unlabeled DNA duplex (0, 0.1, 1, 10, 25, 50, 100, 1000 nM) for 20 min at RT. For the RNA displacement from its complex with Ku by the unlabeled DNA, ³²P-labeled 0.1 nM hp1RNA was pre-incubated with 100 nM Ku protein in the reaction buffer for 20 min at RT, followed by the addition of increasing amounts of the DNA and incubation for 20 min more.

To analyze binding of human full-length 7SK snRNA and recombinant Ku protein, 7SK snRNA was *in vitro* transcribed and 5'-end ³²P-labeled by [γ -³²P]-ATP and T4 polynucleotide

kinase. Then, 5 nM 7SK snRNA was incubated with increasing concentrations of recombinant Ku protein (0 – 400 nM) in the reaction buffer for 20 min at RT followed by separation in the native PAGE.

Typhoon FLA9500 Phosphorimager (GE Healthcare, USA) was used for gel visualization and quantification. The dissociation constants were determined through an approximation of experimental data using a standard equation for bimolecular ligand-receptor binding. Chart preparation and analysis of experimental data were performed using GraphPad Prism7.03 software.

2.8. RNase T1 footprinting assay

200 nM recombinant Ku protein was incubated with 10 nM T7-transcribed 5'-³²P-labeled full-length 7SK snRNA in 200 µl of a buffer (50 mM Tris-HCl pH 8.0, 100 mM NaCl, 2 mM DTT, 1 mM EDTA, 3% glycerol, 5 µg/ml yeast tRNA, 0.1 mg/ml BSA) for 20 min at room temperature. Then, samples containing either 7SK-Ku complex or only free 7SK snRNA were treated with 0.005 U RNase T1 (Thermo Fisher Scientific, USA) per 1 µl of the reaction mixture for 5 min at 37° C. RNase T1 digestion was stopped by adding phenol-chloroform and subsequent RNA-extraction and precipitation with ethanol in the presence of glycogen (5 µg). Reaction products were suspended in a loading buffer (80% formamide, 0.05% bromophenol blue, 0.05% xylene cyanol, 1xTBE) and separated on 6%-10% and 10%-15% polyacrylamide denaturing gels.

2.9. Cell culture

HEK 293T cells and microglial cells were cultured at 37°C and 5% CO₂ in DMEM medium supplemented with 10% FBS and 100 I.U./mL penicillin/100 µg/mL streptomycin solution (all from Thermo Fisher Scientific, USA).

2.10. Immunoprecipitation assay

For RNA-immunoprecipitation assay, HEK 293T cells in a 100-mm dish were transfected with 30 µg of the empty pCDNA3.1 vector (negative control) by calcium phosphate or co-transfected with 15 µg of pCDNA3_Ku70_3xFLAG and 15 µg of pCDNA3_Ku80_3xFLAG vectors described earlier [49]. 48 h after transfection, the cells were lysed for 30 min on ice in a lysis buffer (20 mM Tris-HCl pH 7.4, 150 mM NaCl, 10% glycerol, 1 mM EDTA, 5 mM MgCl₂, 0.05% Triton-X100, 0.1% Tween20, 1mM DTT, 0.5 mM PMSF), supplemented with Protease inhibitor cocktail (Thermo Fisher Scientific, USA) and RiboLock RNase Inhibitor (Thermo Fisher Scientific, USA). Extracts were cleared by centrifugation for 10 min at 14,000 g, and

protein concentration was measured on NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific). 0.1 mg of cell lysates were saved for input analysis. 1 mg of total protein was mixed with FLAG-antibody conjugated agarose (Sigma) and incubated overnight at 4°C. Beads were washed 5 times with lysis buffer for 10 min at 4°C, and bound protein complexes were eluted with 0.1 M glycine pH 2.5 for 20 min at room temperature. Elution fractions and inputs were divided for Western blot analysis and RNA extraction by TRIzol Reagent (Thermo Fisher Scientific, USA) according to the manufacturer's protocol. cDNA was synthesized using the MMLV kit (Evrogen, Russia), and qPCR was performed with primers for human 7SK snRNA (5'-GTCAAGGGTATACGAGTAGCTG-3' and 5'-TGACTACCCTACGTTCTCCTAC-3'), for human U6 RNA (5'-CGCTTCGGCAGCACATATAC-3' and 5'-AAAATATGGAACGCTTCACGA-3') using the qPCR Mix-HS SYBR (Evrogen, Russia) on a Biorad CFX96 amplifier (Biorad, USA).

For co-immunoprecipitation following harvest, cells were lysed for 30 min on ice in the lysis buffer (20 mM Tris-HCl pH 7.4, 150 mM NaCl, 10% glycerol, 1 mM EDTA, 5 mM MgCl₂, 0.05% Triton-X100, 0.1% Tween20, 1 mM DTT, 0.5 mM PMSF), supplemented with Protease inhibitor cocktail (Thermo Fisher Scientific, USA). After centrifugation 15 min 14000 rpm, the protein amount in the extract was quantified by Bradford assay. Antibodies anti-Cdk9 (sc13130, Santa Cruz Biotechnology, USA), anti-HEXIM1 (ab25388, Abcam, USA), anti-Ku80 (ab119935, Abcam, USA), and rabbit IgG isotype control (Thermo Fisher Scientific, USA) as a negative control were incubated with Dynabeads Protein G (Thermo Fisher Scientific, USA) for 3 h at 4°C, after that 1 mg of total extract per immunoprecipitation was added for overnight incubation. After several washes with IPLS buffer (50 mM Tris-HCl pH 7.4, 120 mM NaCl, 0.5 mM EDTA, 0.5% NP40, 10% glycerol) and IPHS buffer (50 mM Tris-HCl pH 7.4, 0.5 M NaCl, 0.5 mM EDTA, 0.5% NP40, 10% glycerol) proteins were eluted by Laemmli sample buffer (Biorad, USA) and analyzed by Western blot.

For HEXIM1-FLAG immunoprecipitation, HEK 293T cells in 100 mm dish were transfected with either 30 µg of the empty pcDNA3.1 vector (negative control) or pcDNA3.1_HEXIM1_3xFLAG by calcium phosphate. 48 h after transfection, the cells were lysed for 30 min on ice in the lysis buffer (20 mM Tris-HCl pH 7.4, 150 mM NaCl, 10% glycerol, 1 mM EDTA, 5 mM MgCl₂, 0.05% Triton-X100, 0.1% Tween20, 1mM DTT, 0.5 mM PMSF), supplemented with Protease inhibitor cocktail (Thermo Fisher Scientific, USA). Extracts were cleared by centrifugation for 10 min at 14000 rpm, and protein concentration was measured using NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific, USA). 0.1 mg of cell lysates were saved for input analysis. 1 mg of total protein was mixed with FLAG-antibody conjugated agarose (Sigma, USA) and incubated overnight at 4°C. Beads were washed 5 times

with the lysis buffer for 10 min at 4°C and then bound protein complexes were eluted by Laemmli sample buffer (Biorad, USA) and analyzed by Western blot.

For RNase treatment, prior to addition of antibody-bound Dynabeads Protein G or FLAG-antibody conjugated agarose to a cellular extract, it was incubated in the presence or absence of RNase A at 250 U/ml and RNase T1 at 1000 U/ml for 10 min at 37°C and subjected to immunoprecipitation as above.

For CTIP2 immunoprecipitation with 3FLAG-Ku70, cells cultured in 150-mm diameter dishes were transfected using a calcium phosphate co-precipitation method with the plasmids indicated. Lysates were prepared two days post-transfection as previously described [47]. Immunoprecipitations were performed using the standard technique, as in [47]. Finally, the immunoprecipitated complexes were processed for SDS-PAGE and immunoblot analysis. Proteins were visualized by chemiluminescence using the Super Signal Chemiluminescence Detection System (Thermo Fisher Scientific, USA).

2.11. Bioinformatics analysis

The eCLIP data were obtained from the ENCODE database. The data of experiments ENCSR571ROL and ENCSR258QKO, obtained for HepG2 and K562 lines, were used respectively. Detailed protocols of experiments and data processing are available on the pages of experiments (<https://www.encodeproject.org/experiments/ENCSR571ROL/> and <https://www.encodeproject.org/experiments/ENCSR258QKO/>). Uniquely mapped reads were used for reads coverage counting. Files were downloaded in “bed” format containing the coordinates of the RNA binding sites in the human genome (version GRCh38). For every site values of the normalized coverage (eCLIP signal) and p-value were given. Peaks were filtered by thresholds to a signal level of at least 6.3 and p-value no more than 0.01. For further analysis, only the peaks present in both biological replicates were used and the average coverage was calculated. If the partially overlapping sites, the sequences occurring in one replicate was selected, and the average coverage for the common fragment was counted.

For analysis we used the following software: RNAstructure for RNA secondary structure prediction [53], RNAforester from ViennaRNA Package 2.0 with parameter -r for RNA secondary structure similarity calculation [54], LocARNA for consensus RNA structure prediction [55]. Preferable for Ku binding consensus RNA structure was build using our tested RNAs (TAR, hp1RNA, let7a) and published structure from p53 mRNA [6] and human telomerase RNA [18]. Also, we included GpG sequence into consensus structure since UGGG > CAAA substitution in the terminal loop of TAR RNA decreased its binding with Ku.

2.12. Confocal microscopy

Microglial cells were cultivated on glass coverslips in a 24-well plate and transfected with Jetprime[®] (polyplus transfection) according to the manufacturer's protocol. 48-hours post-transfection, cells were fixed with PFA 4% and permeabilized with PFA 4%/ triton 1%. The cells were incubated in BSA blocking buffer, then incubated for 1 hour with primary mouse anti-Ku70 antibodies (Abcam 3114) followed by a 1-hour incubation with Alexa 647-labeled anti-mouse antibodies. Finally, the cells were incubated for 30 minutes with Hoechst to stain the cell nucleus. The cells were washed with PBS after each step. Fluorescence was recorded by confocal microscopy using a Zeiss laser scanning microscope (model 510 invert) equipped with a Planapo oil (63x) immersion lens (numerical aperture = 1.4).

3. Results

3.1. Expression and purification of recombinant Ku protein

The Ku protein usually used for *in vitro* studies is either an endogenous protein purified from mammalian cells or a protein produced in the baculovirus system [56,57]. Both approaches to Ku production are multi-step and result in relatively low yields and high costs. Purification from *E.coli* cells is mainly used only for truncated Ku proteins [58,59]. Earlier, we designed a method for a recombinant Ku70 subunit production, taking advantages of two-tags (His6- and GST-) protein purification [48,49], and here tried to apply this method to Ku80 production. However, the expression of Ku80 in *E. coli* resulted in its significant proteolysis and insolubility (data not shown). Therefore, we used an approach described earlier [60] for bi-cistronic expression of Ku70 and Ku80 subunits from the single T7 promoter with some modifications. We constructed a plasmid vector containing the Ku70 (XRCC6) coding sequence followed by the Ku80 (XRCC5) coding sequence. To increase the translation efficiency from the second cistron, we used an approach described in [50]: Ku80 sequence had an internal Shine-Dalgarno sequence in the open reading frame of Ku70 sequence and started with a -1nt shift related to Ku70 open reading frame (Fig. 1A).

Recombinant Ku protein bearing a His6-tag at the N-terminus of Ku70 subunit was expressed in *E.coli* BL21(DE3) Codon Plus/pLysS, and purified by two consecutive chromatography procedures. However, the protein prepared along with a full-size Ku70 (MW ~ 75 kDa) and Ku80 (MW ~ 90 kDa) subunits contained one major Ku80 proteolysis product (MW ~ 65 kDa) (40% according to Western blot) (Fig. 1B). Nonetheless, the final yield of the full-length Ku70/Ku80 was estimated at 0.45 mg/L, which is comparable to that described previously [60]. The first N-terminal peptide resulted from trypsin digestion of the Ku80 proteolysis product, according to mass-spectrometry analysis, was ²¹⁰MVMISLEGEDGLDEIYSFSESLR²³²

(Table S1). We supposed that it is C-terminal fragment of Ku80 with an approximate length of 523 amino acids. The truncated Ku80 lacked a major part of the N-terminal alpha-helix/beta-barrel von Willebrand A (vWA) domain, which participates in neither DNA nor Ku70 binding [61,62]. This explains the ability of the truncated Ku80 to form a stable complex with Ku70 during both purification steps.

The main function of Ku protein is recognition of DNA ends during double-stranded DNA break repair via NHEJ pathway [1,4,5]. The minimal length of DNA duplexes sufficient for Ku binding is about 14-20-mer, and this binding is not sequence-specific [61–63]. To examine the functional quality of the prepared recombinant Ku protein, we tested its DNA-binding capacity using an electrophoretic mobility shift assay (EMSA). Briefly, a ^{32}P -labeled 21-mer or 40-mer DNA duplex was incubated with increasing concentrations of the recombinant Ku protein, and then the mixture was resolved on a nondenaturing polyacrylamide gel. As shown in (Fig. 1C), Ku was able to form stable complexes with both DNA ligands; determined K_d values were 0.9 ± 0.3 nM and 0.5 ± 0.2 nM for 21-mer and 40-mer DNA duplexes, respectively, and correlated with published data [64]. As it was described earlier [60,62,64], we detected two lower migrating bands in the case of 40-mer DNA, probably, corresponding to complexes with DNA:Ku ratio equal to 1:1 and 1:2, since the latter appeared only at increased Ku concentrations. The 21-mer duplex formed with Ku only one complex, since the duplex length was evidently insufficient for binding two Ku heterodimers.

3.2. Ku protein binds RNA with a stem-loop structure *in vitro*

Having confirmed that the purified recombinant Ku can bind dsDNA, we next studied its RNA binding capacity by EMSA using a set of ^{32}P -labeled RNA ligands (Table 1). No interaction between the Ku heterodimer and a 21-mer double-stranded RNA (dsRNA) was observed (Fig. 2A, Table 1). Ku purified from HeLa cells is known to interact with HIV-1 TAR RNA [41], which has a hairpin structure with a 3-nt bulge just near the terminal loop. We analyzed the recombinant Ku binding with TAR RNA and found that incubation of the protein with the RNA resulted in a new band with reduced electrophoretic mobility (Fig. 2A). To verify the nature of the complex, the binding was performed in the presence of anti-Ku80, anti-Ku70, and anti-His6 antibodies (supershift assay). In all cases, the appearance of a new high-mass band was observed, thus confirming the RNA-protein complex is composed of Ku70 and Ku80 (Fig. 2B). Thus, we confirmed that our recombinant Ku formed a stable complex with TAR RNA (Table 1). At high protein concentrations, the second RNA-protein complex was detected suggesting binding of two Ku molecules to the ligand as it was demonstrated for Ku-DNA interactions. Earlier, we showed that the human Ku70 subunit purified from *E.coli* interacted

with RNA only in the absence of Mg^{2+} [48]. Due to an excess of the Ku70 subunit in the purified heterodimer, we performed Ku-TAR RNA binding in the presence and absence of Mg^{2+} to distinguish between Ku70-RNA and heterodimer Ku-RNA complex formation. No difference in the K_d values was determined in both experiments that confirmed the preferable formation of the heterodimer Ku-RNA complex. Further experiments were carried out in the absence of Mg^{2+} cations to minimize RNA degradation.

We also analyzed Ku interactions with two RNA hairpins derived from TAR RNA: TAR-M1 hairpin lacking a bulge, and TAR-M2 with UGGG > CAAA substitutions in the terminal loop. The latter decreased the RNA binding by Ku purified from HeLa cells [41]. Changing the loop sequence decreased the affinity of the recombinant Ku protein to TAR RNA as well; however, the decrease was not as strong as in the case of the bulge deletion, which led to an approximately 3-fold drop in K_d (Table 1).

Next, we examined Ku binding with several RNA hairpins of various structures (Fig. 2D, Table 1). Among them, hp1RNA, resembling TAR with the bulged stem-loop structure but without sequence similarity, was bound the most efficiently. The Ku affinity to RNAs without bulges, hp2RNA and microRNA mir-504, was comparable with that to TAR-M1. Interestingly, Ku binding to another microRNA, hsa-let-7a-2, containing a bulky bulge just near the loop was somewhat less efficient than to TAR and hp1RNA. This allowed assuming that the bulge in hsa-let-7a-2 was probably too bulky and created some steric tension when binding to Ku.

Under investigation of Ku interactions with human telomerase RNA [18] or p53 mRNA [6], it was found that endogenous Ku from HeLa cells efficiently bound a bulged hairpin structure in both RNAs. The p53 mRNA hairpin contained two consecutive bulges; the first one contained two nucleotides and was located just near the loop, and the second one was a large bulge composed of 8 nucleotides and located downstream the stem. Importantly, only the deletion of the first bulge led to the disruption of the Ku-RNA complex. Noteworthy, altogether our and literature data demonstrate the importance of the stem-loop RNA structure containing a bulge close to the loop for its recognition by the human Ku protein, and this is consistent with observations made with the yeast Ku protein [16].

To test whether the existence of a bulge is a structural motif sufficient for RNA recognition by human Ku, we constructed an RNA duplex (dsRNA-bulge) containing a 4-nucleotide bulge. A complex between dsRNA-bulge and recombinant Ku turned out to be detected only at the highest protein concentration (400 nM) (Fig. 2A). Thus, we could conclude that an unpaired region in the duplex is not enough for successful RNA recognition by Ku, and therefore, the RNA structure motif necessary for the efficient Ku binding is a hairpin with a bulge located near the loop.

To determine a region within heterodimeric Ku responsible for its interactions with a hairpin RNA, we used a competitor approach and carried out the binding assay in two variants. First, ^{32}P -labeled hp1RNA was incubated with Ku in the presence of increasing concentrations of an unlabeled double-stranded DNA competitor. Second, the competitor was added to a pre-formed complex of ^{32}P -labeled hp1RNA and Ku. In both cases, the dsDNA successfully interrupted the complex formation suggesting the same binding site for dsDNA and RNA hairpin (Fig. 2C). These data are consistent with observations made with Ku purified from mammalian cells [33].

3.3. In silico secondary structure analysis of RNAs interacting with Ku70 in cell culture

To determine RNA structures interacting with Ku protein in cells, we used eCLIP (enhanced crosslinking and immunoprecipitation) data available from the ENCODE database. This method allows the identification of protein binding sites within cellular RNAs and the estimation of enrichment values for the precipitated RNAs [65]. The ENCODE database contains the results of eCLIP experiments for HepG2 and K562 cell lines. Precipitation of cross-linked RNA-Ku complexes was carried out using antibodies to Ku70, in cells Ku70 is tightly associated with Ku80, and the identified RNAs were most likely bound to the Ku heterodimer. The data were obtained in a processed form and represented positions of the protein binding sites in the precipitated RNAs and peak intensity values. The intensity of the peaks was calculated as reads coverage of the protein binding site in RNAs in the eCLIP library, normalized to reads coverage of the site in the transcriptome library. Importantly, these data can be used not only to detect all RNAs associated with the protein but also to identify specific regions of RNA, which are recognized and bound to the protein.

The analysis of the eCLIP data revealed several transcripts that have at least one region bound by Ku70 and longer 10 bases (28 RNA regions listed in Table S2). The set of these transcripts differed between two cell lines, only 2 transcripts were identified in HepG2 line, and 14 transcripts in K562 line. To clarify whether there is a specific RNA motif recognized by Ku70, we analyzed structures of the identified RNA regions (N=28). For each region, a secondary structure was predicted using RNAstructure software [53]. At first sight, the predicted structures looked rather unlike each other and had nothing in common. Of note, the secondary structure of some of them resembled those of HIV-1 TAR RNA and hp1RNA as well as a consensus RNA hairpin preferable for the yeast Ku binding [16]. Also, we noticed that some identified RNAs had G-rich sequences, in particular, GpG dinucleotide, in the terminal loop. The presence of this sequence was found to be important for the Ku binding since mutations UGGG->CAAA in the TAR RNA loop slightly decreased its binding to Ku (Table 1).

Since RNAs had different signals in eCLIP, we assumed that the value of the eCLIP signal should depend on the RNA structure. To verify this assumption, we constructed a consensus RNA structure on the base of RNAs validated in our experiments as well as in the literature (Fig. 3A), and then we aligned structures of the identified RNAs with it. Using the alignment, the degree of similarity between the identified RNA fragments and the consensus structure was calculated. The similarity to the consensus structure was found to correlate positively with the magnitude of the eCLIP signal. The greatest correlation was observed taking into account the presence of the GpG sequence in the loop in the similarity counting (Spearman correlation coefficient $r = 0.48$) (Fig. 3B). Thus, we have shown that the structure of RNA determines the Ku-binding affinity not only *in vitro* but also in cells.

3.4. Ku interacts with 7SK snRNA

TAR RNA, which plays an important role in transcription regulation from the HIV-1 promoter, is known to have some structural similarities with 7SK snRNA [46,66,67]. The latter is a highly structured small nuclear RNA participating in the regulation of transcription elongation by RNAPII [68,69]. Taking into account the high affinity of Ku to TAR RNA, we decided to analyze its capacity to bind 7SK snRNA that forms four hairpins with a stem-loop structure.

The EMSA experiment with the recombinant Ku protein and ^{32}P -labeled T7-transcribed RNA showed that 7SK snRNA formed a complex with Ku, and K_d value for this complex was estimated as 30 nM (Fig. 4A). To elucidate the Ku-binding site within 7SK snRNA, we used a footprinting assay with RNase T1 digestion. We clearly demonstrated strong protection of G50-G52 and G55 residues located in the terminal loop of the conserved hairpin 1 (HP1) pointing out to the Ku binding with this bulged stem-loop region (Fig. 4B-C). We also observed the protection of several G-residues in the hairpin 3 (HP3), but it was much slighter than the protection of the G50-G55 region. We cannot rule out some weak interactions of Ku with HP3, but rather Ku being strung on HP1 by its bulk ring-like DNA-binding domain covers the structures lying next to it within the tertiary RNA structure.

Further, we performed the RNA-immunoprecipitation assay to find out interactions of Ku and 7SK snRNA in HEK 293T cells. Briefly, Ku70 and Ku80 tagged by 3xFLAG-peptide were overexpressed and pulled down by anti-FLAG resin, protein-RNAs complexes were washed and eluted, and the RNAs were extracted. Cells transfected with an empty pcDNA3 vector was used as a negative control. The level of the co-precipitated RNAs was estimated by RT-qPCR. We obtained two-fold enrichment of 7SK snRNA but not U6 snRNA, which we used as a control (Fig. 4, D). U6 snRNA is a non-coding small nuclear RNA with a stem-loop structure, however,

it contains neither bulge nor terminal G-rich loop [70,71] (Fig. S1). Thus we have first shown that Ku protein is able to bind 7SK snRNA *in vitro* or being over-expressed in HEK 293T cells.

3.5. Interactions of Ku with protein components of the 7SK ribonucleoprotein complex

7SK RNA is known to form a scaffold for a ribonucleoprotein complex [45,69]. We analyzed whether Ku can interact with such components of 7SK snRNP as HEXIM1 and Cdk9. For this purpose, we performed immunoprecipitation of these endogenous proteins using specific anti-HEXIM1 or anti-Cdk9 antibodies, and as a result, we observed the Ku80 subunit in the precipitated fraction (Fig. 5A). Moreover, endogenous HEXIM1 and Cdk9 were co-precipitated with endogenous Ku80 using anti-Ku80 antibodies (Fig. 5A). The co-precipitation of the proteins from 7SK snRNP with Ku could be mediated by interactions between Ku and 7SK snRNA. To clarify a possibility of a direct Ku binding to Cdk9, we carried out immunoprecipitation of endogenous Cdk9 in the presence and absence of RNase cocktail, and subsequently monitored the amounts of co-purified endogenous Ku80 (Fig. 5B). A significant amount of Ku80 was co-precipitated with Cdk9 in an RNase-resistant fashion suggesting direct protein-protein interactions between them.

In the case of HEXIM1, immunoprecipitation of Flag-tagged HEXIM1 was performed from cells transiently transfected with either pcDNA3 or pcDNA3-HEXIM1-3xFLAG vector (Fig 5C) also with or without RNase treatment. We could detect Ku80-HEXIM1 complex under native conditions, whereas the RNase treatment partially disrupted the complex. To verify if the association of these proteins is indeed RNA-dependent, a recombinant GST-HEXIM1 protein was purified, and interaction of recombinant proteins His6-Ku70/Ku80 and GST-HEXIM1 was studied in a GST-pull-down assay. We did not detect direct binding of Ku and HEXIM1 in the absence of 7SK snRNA (data not shown). To examine the formation of the ternary complex containing HEXIM1, Ku and 7SK snRNA, we performed an EMSA using labeled T7-transcribed 7SK snRNA and the recombinant proteins (Fig. 5D). In the absence of HEXIM1, we detected two complexes of Ku with full-length 7SK snRNA and a complex probably composed of Ku and degraded RNA. When HEXIM1 was added, bands corresponding to the complexes of Ku with full-length 7SK snRNA disappeared, and the formation of a higher-mass complex was observed. Notably, this complex possessed slightly lower mobility than a binary HEXIM1-RNA complex, suggesting it contained HEXIM1, Ku and RNA. Thus, we can conclude that the association of Ku with HEXIM1 protein occurs via their interactions with 7SK snRNA.

We have previously shown that the cellular repressor CTIP2 associates with 7SK snRNA and HEXIM1 to repress P-TEFb function [47]. Here we analyzed interactions of Ku with CTIP2 by co-immunoprecipitation experiments and confocal microscopy analysis. Co-

immunoprecipitation experiments were performed with nuclear extracts from HEK 293T cells expressing CTIP2 and Ku70-Flag. We found that CTIP2 does interact with Ku protein (Fig. 5E). To confirm Ku association with CTIP2 in a cellular context, we next analyzed Ku70 and CTIP2 localization patterns in microglial cells. These cells are considered as the main HIV-1 reservoirs of the central nervous system, and CTIP2 is shown to be a key regulator of HIV-1 gene transcription in microglial cells (reviewed in [72]). The confocal microscopy analysis of CTIP2 and Ku70 nuclear localization also pointed to the possibility of interactions of these proteins (Fig. 5F). Ku70 harbored a punctate pattern of localization in the nucleus of the cells in the absence of RFP-CTIP2 expression (lane A), whereas upon RFP-CTIP2 expression we observed a clear concentration of the endogenous Ku70 protein in CTIP2-induced structures. As shown by the yellow staining in Fig. 5F (lane C4), RFP-CTIP2 and Ku70 were co-localized in the nucleus. As a specificity control, no green staining was detected with the fluorescent secondary antibody alone (lane B). Altogether, these results suggest that the Ku protein physically interacts with the CTIP2-associated inactive pTEFb complex.

4. Discussion

Accumulating data point out to the importance of Ku-RNAs interactions in eukaryotic cells. There are studies elucidating the functional role of Ku-RNA binding in DSB repair, telomere maintenance, modulation of translation and innate immune response [10-14, 21-24, 29, 33]. However, the mechanism of RNA binding by human Ku heterodimer has not been yet clarified. To figure out the structural aspects of RNA interactions with the human Ku heterodimer, we developed a new system for recombinant Ku production in *E.coli*, and using the purified heterodimer and RNA molecules possessing different structures, we performed EMSA studies. Besides this, we analyzed the results of eCLIP (enhanced crosslinking and immunoprecipitation) experiments available for human Ku from the ENCODE database. The results obtained altogether allowed us to propose a consensus RNA structure favorable for the Ku binding, which is a hairpin containing a bulge within a few nucleotides from the loop and possessing the GpG sequence in the terminal loop (Fig. 3A). The structure proposed resembles consensus RNA preferable for binding with yeast Ku described earlier [16] except for the loop sequence importance for the recognition by the human protein, but not the yeast one.

We also investigated Ku interplay with a highly structured 7SK snRNA, which is a scaffold for a ribonucleoprotein complex participating in transcription regulation [45]. We found that the recombinant Ku protein and T7-transcribed RNA formed a stable complex with K_d value about 30 nM (Fig. 4A). Also, RNA-immunoprecipitation of transiently expressed FLAG-tagged Ku in HEK 293T cells confirmed the intracellular interaction of 7SK snRNA and Ku protein

(Fig. 4C). To better characterize the mode of Ku-7SK RNA interplay, we performed a footprinting assay with RNase T1 digestion, which allowed us to elucidate the Ku-binding site within 7SK snRNA. Despite a complex stem-loop structure of this RNA with a lot of hairpins and stem abnormalities preferable for Ku binding, we detected strong and reliable protection only for G50-G52 and G55 residues located in the terminal loop of a conservative hairpin 1 (HP1) (Fig. 4B). Upon our study of the Ku binding with different RNA structures, we detected the formation of a stable complex between Ku and TAR RNA, which is a bulged hairpin formed at the 5'-end of the nascent transcript of HIV-1. Interestingly, structural similarities of TAR and 7SK RNAs were thoroughly studied, and a structural mimicry was found for TAR and two hairpins of 7SK snRNA, HP1 [46] and HP4 [66]. However, we detected no protection in the HP4; the result that points out to a specific character of the Ku binding to 7SK snRNA.

Of note, HP1 is involved in interactions with the transcription repressor HEXIM1 that binds an inverted GAUC double-strand repeat located near the terminal loop [73,74]. Importantly, in 7SK snRNP, HEXIM1 also interacts with Cdk9, a component of P-TEFb bound with HP4, and inhibits its phosphorylating activity [75,76]. We verified the association of endogenous Ku with these proteins in HEK 293T cells. As for Cdk9, the interaction of Ku70 only with its minor 55 kDa-isoform was previously shown in HeLa cells [77]. In our study, Ku was co-precipitated with the main 42 kDa-isoform of Cdk9, and the binding of these proteins was RNA-independent. HEXIM1 was co-precipitated with Ku as well, however, RNase treatment destroyed HEXIM1-Ku complex (Fig. 4D-F). Moreover, recombinant Ku and HEXIM1 were not able to interact in the absence of 7SK snRNA. Importantly, we detected the formation of a ternary complex containing the recombinant HEXIM1, Ku, and full-length 7SK snRNA (Fig. 4G), although HEXIM1 and Ku bind the same HP1. We can propose two possible explanations of this fact: either HEXIM1 and Ku bind different regions within HP1 and therefore do not interfere with each other, or in the presence of Ku, HEXIM1 interacts with another 7SK snRNA region since it has been found to bind HP3 as well [74]. We have also elucidated the interaction of Ku with another component of 7SK snRNP, namely, the transcriptional repressor CTIP2 that is a major modulator of P-TEFb activity. CTIP2 is known to interact with an inactive P-TEFb complex containing the 7SK snRNA and HEXIM1[47].

There are numerous studies showing the participation of Ku protein in the transcription regulation of cellular and viral genes [40,78–84], and it is supposed that there is a specific mechanism of Ku-dependent regulation for each particular gene. Among the mechanisms proposed, there is no either involving the interaction of Ku with RNA, but we cannot completely rule out such a possibility. Also, it has been shown that depletion of the Ku protein level affects the efficiency of HIV-1 transcription in infected cells [84–86]. However, the nature of Ku

influence has not been elucidated, and no Ku-binding site within the HIV-1 LTR promoter is yet determined [87,88].

5. Conclusions

We have demonstrated that Ku may efficiently bind both TAR and 7SK snRNA that are known to be involved in HIV-1 transcription regulation [42,89,90]. Our data give reasons to hypothesize that Ku protein could be involved at least in the HIV-1 transcription regulation via its interactions with TAR and/or 7SK snRNA as well as protein components of 7SK snRNP. Indeed, we have previously reported that the cellular repressor CTIP2 binds both the 7SK snRNA HP2 and HEXIM1 to repress P-TEFb function and thereby cellular and HIV-1 genes transcription [47]. Here we further report that Ku binds and colocalizes with CTIP2 in the nucleus. These results confirm Ku association with the inactive P-TEFb complex and suggest the involvement of Ku protein in the regulation of P-TEFb-sensitive genes including HIV-1.

Funding

The work in the MSU laboratory was supported by RFBR (grants №18-34-00393 (Ku-7SK RNP interaction) and № 18-04-00542 (Ku-DNA interaction)) and the Russian Science Foundation grant 17-14-01107 (Ku interactions with TAR RNA, HEXIM1, and Cdk9) and 19-74-10021 (eCLIP data analysis). Work in the OR laboratory was supported by grants from the French Agency for Research on AIDS and Viral Hepatitis (ARNS), Sidaction, the European Union's Horizon 2020 research and innovation program under grant agreement No. 691119 - EU4HIVCURE - H2020-MSCA-RISE-2015, Alsace Contre le Cancer and University of Strasbourg.

Authors' Contributions

OS, OR, MG - conception of the work; OS, IG, SK, TZ, JA, FD, CW, - collection of data; OS, IG, OR, MG - analysis of data; OS, IG, SK, JA, OR, MG - writing of manuscript.
All the authors have approved the final article.

References

- [1] J.A. Downs, S.P. Jackson, A means to a DNA end: the many roles of Ku., *Nat. Rev. Mol. Cell Biol.* 5 (2004) 367–78. <https://doi.org/10.1038/nrm1367>.
- [2] A.J. Davis, B.P.C. Chen, D.J. Chen, DNA-PK: A dynamic enzyme in a versatile DSB repair pathway, *DNA Repair (Amst)*. 17 (2014) 21–9. <https://doi.org/10.1016/j.dnarep.2014.02.020>.
- [3] B.M. Sirbu, D. Cortez, DNA damage response: three levels of DNA repair regulation., *Cold Spring Harb. Perspect. Biol.* 5 (2013) a012724. <https://doi.org/10.1101/cshperspect.a012724>.
- [4] S.-H. Lee, C.-H. Kim, DNA-dependent Protein Kinase Complex: a Multifunctional Protein in DNA Repair and Damage Checkpoint, *Mol. Cells*. 13 (n.d.) 159–166. <http://www.molcells.org/journal/view.html?year=2002&volume=13&number=2&spage=159> (accessed July 8, 2019).
- [5] V.L. Fell, C. Schild-Poulter, The Ku heterodimer: Function in DNA repair and beyond, *Mutat. Res. - Rev. Mutat. Res.* 763 (2015) 15–29. <https://doi.org/10.1016/j.mrrev.2014.06.002>.
- [6] A. Lamaa, M. Le Bras, N. Skuli, S. Britton, P. Frit, P. Calsou, H. Prats, A. Cammas, S. Millevoi, A novel cytoprotective function for the DNA repair protein Ku in regulating p53 mRNA translation and function, 17 (2016) 1–11. <https://doi.org/10.15252/embr.201541181>.
- [7] T.S. Fisher, V.A. Zakian, Ku: A multifunctional protein involved in telomere maintenance, *DNA Repair (Amst)*. 4 (2005) 1215–1226. <https://doi.org/10.1016/j.dnarep.2005.04.021>.
- [8] E.J. Gapud, B.P. Sleckman, Unique and redundant functions of ATM and DNA-PKcs during V(D)J recombination., *Cell Cycle*. 10 (2011) 1928–35. <https://doi.org/10.4161/cc.10.12.16011>.
- [9] A.A. Kosova, O.I. Lavrik, S.N. Hodyreva, [The role of Ku antigen in the repair of apurinic/apyrimidinic sites in DNA]., *Mol. Biol. (Mosk)*. 49 (n.d.) 67–74. <http://www.ncbi.nlm.nih.gov/pubmed/25916111> (accessed July 8, 2019).
- [10] V. Sharma, T. Misteli, Non-coding RNAs in DNA damage and repair., *FEBS Lett.* 587 (2013) 1832–9. <https://doi.org/10.1016/j.febslet.2013.05.006>.
- [11] Y.-G. Yang, Y. Qi, RNA-directed repair of DNA double-strand breaks, *DNA Repair (Amst)*. 32 (2015) 82–85. <https://doi.org/10.1016/j.dnarep.2015.04.017>.
- [12] Y. Zhang, Q. He, Z. Hu, Y. Feng, L. Fan, Z. Tang, J. Yuan, W. Shan, C. Li, X. Hu, J.L. Tanyi, Y. Fan, Q. Huang, K. Montone, C. V Dang, L. Zhang, Long noncoding RNA

- LINP1 regulates repair of DNA double-strand breaks in triple-negative breast cancer., *Nat. Struct. Mol. Biol.* 23 (2016) 522–30. <https://doi.org/10.1038/nsmb.3211>.
- [13] J.S. Pflingsten, K.J. Goodrich, C. Taabazuing, F. Ouenzar, P. Chartrand, T.R. Cech, Mutually exclusive binding of telomerase RNA and DNA by Ku alters telomerase recruitment model., *Cell.* 148 (2012) 922–32. <https://doi.org/10.1016/j.cell.2012.01.033>.
- [14] S.E. Peterson, A.E. Stellwagen, S.J. Diede, M.S. Singer, Z.W. Haimberger, C.O. Johnson, M. Tzoneva, D.E. Gottschling, The function of a stem-loop in telomerase RNA is linked to the DNA repair protein Ku., *Nat. Genet.* 27 (2001) 64–7. <https://doi.org/10.1038/83778>.
- [15] A.E. Stellwagen, Z.W. Haimberger, J.R. Veatch, D.E. Gottschling, Ku interacts with telomerase RNA to promote telomere addition at native and broken chromosome ends., *Genes Dev.* 17 (2003) 2384–95. <https://doi.org/10.1101/gad.1125903>.
- [16] A.B. Dalby, K.J. Goodrich, J.S. Pflingsten, T.R. Cech, RNA recognition by the DNA end-binding Ku heterodimer., *RNA.* 19 (2013) 841–51. <https://doi.org/10.1261/rna.038703.113>.
- [17] E.P. Hass, D.C. Zappulla, The Ku subunit of telomerase binds Sir4 to recruit telomerase to lengthen telomeres in *S. cerevisiae*, *Elife.* 4 (2015). <https://doi.org/10.7554/eLife.07750>.
- [18] N.S.Y. Ting, Y. Yu, B. Pohorelic, S.P. Lees-Miller, T.L. Beattie, Human Ku70/80 interacts directly with hTR, the RNA component of human telomerase., *Nucleic Acids Res.* 33 (2005) 2090–8. <https://doi.org/10.1093/nar/gki342>.
- [19] T.S. Fisher, A.K.P. Taggart, V.A. Zakian, Cell cycle-dependent regulation of yeast telomerase by Ku, *Nat. Struct. Mol. Biol.* 11 (2004) 1198–1205. <https://doi.org/10.1038/nsmb854>.
- [20] D. Silvera, N. Koloteva-Levine, S. Burma, O. Elroy-Stein, Effect of Ku proteins on IRES-mediated translation, *Biol. Cell.* 98 (2006) 353–361. <https://doi.org/10.1042/BC20050060>.
- [21] M. Takagi, M.J. Absalon, K.G. McLure, M.B. Kastan, Regulation of p53 Translation and Induction after DNA Damage by Ribosomal Protein L26 and Nucleolin, *Cell.* 123 (2005) 49–63. <https://doi.org/10.1016/j.cell.2005.07.034>.
- [22] N.E. Peters, B.J. Ferguson, M. Mazzon, A.S. Fahy, E. Krysztofinska, R. Arribas-Bosacoma, L.H. Pearl, H. Ren, G.L. Smith, A mechanism for the inhibition of DNA-PK-mediated DNA sensing by a virus., *PLoS Pathog.* 9 (2013) e1003649. <https://doi.org/10.1371/journal.ppat.1003649>.
- [23] B.J. Ferguson, D.S. Mansur, N.E. Peters, H. Ren, G.L. Smith, DNA-PK is a DNA sensor for IRF-3-dependent innate immunity, *Elife.* 1 (2012) e00047. <https://doi.org/10.7554/eLife.00047>.
- [24] Y. Li, Y. Wu, X. Zheng, J. Cong, Y. Liu, J. Li, R. Sun, Z.G. Tian, H.M. Wei, Cytoplasm-

- Translocated Ku70/80 Complex Sensing of HBV DNA Induces Hepatitis-Associated Chemokine Secretion, *Front. Immunol.* 7 (2016).
<https://doi.org/10.3389/fimmu.2016.00569>.
- [25] H. Sui, M. Zhou, H. Imamichi, X. Jiao, B.T. Sherman, H.C. Lane, T. Imamichi, STING is an essential mediator of the Ku70-mediated production of IFN- λ 1 in response to exogenous DNA, *Sci. Signal.* 10 (2017) eaah5054.
<https://doi.org/10.1126/scisignal.aah5054>.
- [26] M. Morchikh, A. Cribier, R. Raffel, S. Amraoui, J. Cau, D. Severac, E. Dubois, O. Schwartz, Y. Bennasser, M. Benkirane, HEXIM1 and NEAT1 Long Non-coding RNA Form a Multi-subunit Complex that Regulates DNA-Mediated Innate Immune Response, *Mol. Cell.* 67 (2017) 387-399.e5. <https://doi.org/10.1016/j.molcel.2017.06.020>.
- [27] C.M. Clemson, J.N. Hutchinson, S.A. Sara, A.W. Ensminger, A.H. Fox, A. Chess, J.B. Lawrence, An architectural role for a nuclear noncoding RNA: NEAT1 RNA is essential for the structure of paraspeckles., *Mol. Cell.* 33 (2009) 717–26.
<https://doi.org/10.1016/j.molcel.2009.01.026>.
- [28] Q. Li, J.J. Cooper, G.H. Altwerger, M.D. Feldkamp, M.A. Shea, D.H. Price, HEXIM1 is a promiscuous double-stranded RNA-binding protein and interacts with RNAs in addition to 7SK in cultured cells., *Nucleic Acids Res.* 35 (2007) 2503–12.
<https://doi.org/10.1093/nar/gkm150>.
- [29] 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, O. Bensaude, Binding of the 7SK snRNA turns the HEXIM1 protein into a P-TEFb (CDK9/cyclin T) inhibitor, *EMBO J.* 23 (2004) 2608–2619.
<https://doi.org/10.1038/sj.emboj.7600275>.
- [30] S. Zhang, B. Schlott, M. Görlach, F. Grosse, DNA-dependent protein kinase (DNA-PK) phosphorylates nuclear DNA helicase II/RNA helicase A and hnRNP proteins in an RNA-dependent manner, *Nucleic Acids Res.* 32 (2004) 1–10.
<https://doi.org/10.1093/nar/gkg933>.
- [31] M. Fidaleo, E. De Paola, M.P. Paronetto, The RNA helicase A in malignant transformation., *Oncotarget.* 7 (2016) 28711–23.
<https://doi.org/10.18632/oncotarget.7377>.
- [32] A. Dvir, S.R. Peterson, M.W. Knuth, H. Lu, W.S. Dynan, Ku autoantigen is the regulatory component of a template-associated protein kinase that phosphorylates RNA polymerase II., *Proc. Natl. Acad. Sci. U. S. A.* 89 (1992) 11920–4.
<ftp://ftp.radiomaryja.pl.eu.org/vol/rzm1/cran/doc/contrib/Baggott-refcard-v2.pdf>.
- [33] S. Yoo, W.S. Dynan, Characterization of the RNA binding properties of Ku protein.,

- Biochemistry. 37 (1998) 1336–1343. <https://doi.org/10.1021/bi972100w>.
- [34] N.S.Y. Ting, B. Pohorelic, Y. Yu, S.P. Lees-Miller, T.L. Beattie, The human telomerase RNA component, hTR, activates the DNA-dependent protein kinase to phosphorylate heterogeneous nuclear ribonucleoprotein A1., *Nucleic Acids Res.* 37 (2009) 6105–15. <https://doi.org/10.1093/nar/gkp636>.
- [35] L.S. Treviño, M.J. Bolt, S.L. Grimm, D.P. Edwards, M.A. Mancini, N.L. Weigel, Differential Regulation of Progesterone Receptor-Mediated Transcription by CDK2 and DNA-PK., *Mol. Endocrinol.* 30 (2015) me20151144. <https://doi.org/10.1210/me.2015-1144>.
- [36] R.H.F.F. Wong, I. Chang, C.S.S.S. Hudak, S. Hyun, H.-Y.Y. Kwan, H.S. Sul, A Role of DNA-PK for the Metabolic Gene Regulation in Response to Insulin, *Cell.* 136 (2009) 1056–1072. <https://doi.org/10.1016/j.cell.2008.12.040>.
- [37] M.W. Knuth, S.I. Gunderson, N.E. Thompson, L.A. Strasheim, R.R. Burgess, Purification and characterization of proximal sequence element-binding protein 1, a transcription activating protein related to Ku and TREF that binds the proximal sequence element of the human U1 promoter., *J. Biol. Chem.* 265 (1990) 17911–20. <http://www.ncbi.nlm.nih.gov/pubmed/2211668> (accessed July 8, 2019).
- [38] W. Giffin, H. Torrance, D.J. Rodda, G.G. Préfontaine, L. Pope, R.J. Hache, Sequence-specific DNA binding by Ku autoantigen and its effects on transcription., *Nature.* 380 (1996) 265–8. <https://doi.org/10.1038/380265a0>.
- [39] J. Hazra, P. Mukherjee, A. Ali, S. Poddar, M. Pal, Engagement of Components of DNA-Break Repair Complex and NFκB in Hsp70A1A Transcription Upregulation by Heat Shock., *PLoS One.* 12 (2017) e0168165. <https://doi.org/10.1371/journal.pone.0168165>.
- [40] O.A. Shadrina, E.S. Knyazhanskaya, S.P. Korolev, M.B. Gottikh, Host Proteins Ku and HMGA1 As Participants of HIV-1 Transcription., *Acta Naturae.* 8 (n.d.) 34–47. <http://www.ncbi.nlm.nih.gov/pubmed/27099783> (accessed July 8, 2019).
- [41] W. Kaczmarek, S.A. Khan, Lupus Autoantigen Ku Protein Binds Hiv-1 TAR RNA in vitro.Pdf, *Biochem. Biophys. Res. Commun.* 196 (1993) 935–942.
- [42] S. Bannwarth, A. Gatignol, HIV-1 TAR RNA: the target of molecular interactions between the virus and its host., *Curr HIV Res.* 3 (2005) 61–71. <https://doi.org/15638724>.
- [43] A. Marcello, M. Zoppé, M. Giacca, Critical Review Multiple Modes of Transcriptional Regulation by the HIV-1 Tat Transactivator, *IUBMB Life.* 51 (2001) 175–181. <https://doi.org/10.1080/152165401753544241>.
- [44] A. Gatignol, Transcription of HIV: Tat and Cellular Chromatin, in: *Adv. Pharmacol.*, 2007: pp. 137–159. [https://doi.org/10.1016/S1054-3589\(07\)55004-0](https://doi.org/10.1016/S1054-3589(07)55004-0).

- [45] G. Diribarne, O. Bensaude, 7SK RNA, a non-coding RNA regulating P-TEFb, a general transcription factor, *RNA Biol.* 6 (2009) 122–128. <https://doi.org/10.4161/rna.6.2.8115>.
- [46] S.C. Sedore, S.A. Byers, S. Biglione, J.P. Price, W.J. Maury, D.H. Price, Manipulation of P-TEFb control machinery by HIV: Recruitment of P-TEFb from the large form by Tat and binding of HEXIM1 to TAR, *Nucleic Acids Res.* 35 (2007) 4347–4358. <https://doi.org/10.1093/nar/gkm443>.
- [47] T. Cherrier, V. Le Douce, S. Eilebrecht, R. Riclet, C. Marban, F. Dequiedt, Y. Goumon, J.C. Paillart, M. Mericskay, A. Parlakian, P. Bausero, W. Abbas, G. Herbein, S.K. Kurdistani, X. Grana, B. Van Driessche, C. Schwartz, E. Candolfi, A.G. Benecke, C. Van Lint, O. Rohr, CTIP2 is a negative regulator of P-TEFb, *Proc. Natl. Acad. Sci. U. S. A.* 110 (2013) 12655–12660. <https://doi.org/10.1073/pnas.1220136110>.
- [48] A.N. Anisenko, E.S. Knyazhanskaya, T.S. Zatsepin, M.B. Gottikh, Human Ku70 protein binds hairpin RNA and double stranded DNA through two different sites, *Biochimie.* 132 (2017) 85–93. <https://doi.org/10.1016/j.biochi.2016.11.001>.
- [49] A.N. Anisenko, E.S. Knyazhanskaya, A.O. Zalevsky, J.Y. Agapkina, A.I. Sizov, T.S. Zatsepin, M.B. Gottikh, Characterization of HIV-1 integrase interaction with human Ku70 protein and initial implications for drug targeting., *Sci. Rep.* 7 (2017) 5649. <https://doi.org/10.1038/s41598-017-05659-5>.
- [50] I.A. Osterman, S.A. Evfratov, P. V Sergiev, O.A. Dontsova, Comparison of mRNA features affecting translation initiation and reinitiation., *Nucleic Acids Res.* 41 (2013) 474–86. <https://doi.org/10.1093/nar/gks989>.
- [51] F. Forouzanfar, S. Ali, C. Wallet, M. De Rovere, C. Ducloy, H. El Mekdad, M. El Maassarani, A. Aït-Ammar, J. Van Assche, E. Boutant, F. Daouad, F. Margottin-Goguet, C. Moog, C. Van Lint, C. Schwartz, O. Rohr, HIV-1 Vpr mediates the depletion of the cellular repressor CTIP2 to counteract viral gene silencing, *Sci. Rep.* 9 (2019). <https://doi.org/10.1038/s41598-019-48689-x>.
- [52] Y. V Lyupina, O.G. Zatsepina, M. V Serebryakova, P.A. Erokhov, S.B. Abaturova, O.I. Kravchuk, O. V Orlova, S.N. Beljelarskaya, A.I. Lavrov, O.S. Sokolova, V.S. Mikhailov, Proteomics of the 26S proteasome in *Spodoptera frugiperda* cells infected with the nucleopolyhedrovirus, AcMNPV., *Biochim. Biophys. Acta.* 1864 (2016) 738–746. <https://doi.org/10.1016/j.bbapap.2016.02.021>.
- [53] J.S. Reuter, D.H. Mathews, RNAstructure: Software for RNA secondary structure prediction and analysis, *BMC Bioinformatics.* 11 (2010). <https://doi.org/10.1186/1471-2105-11-129>.
- [54] R. Lorenz, S.H. Bernhart, C. Höner zu Siederdissen, H. Tafer, C. Flamm, P.F. Stadler, I.L.

- Hofacker, ViennaRNA Package 2.0, Algorithms Mol. Biol. 6 (2011) 26.
<https://doi.org/10.1186/1748-7188-6-26>.
- [55] S. Will, T. Joshi, I.L. Hofacker, P.F. Stadler, R. Backofen, LocARNA-P: accurate boundary prediction and improved detection of structural RNAs., RNA. 18 (2012) 900–14. <https://doi.org/10.1261/rna.029041.111>.
- [56] A. Dvir, L.Y. Stein, B.L. Calore, W.S. Dynan, Purification and characterization of a template-associated protein kinase that phosphorylates RNA polymerase II, J Biol Chem. 268 (1993) 10440–10447. <http://www.ncbi.nlm.nih.gov/pubmed/8486698>.
- [57] M. Ono, P.W. Tucker, J.D.D. Capra, Production and characterization of recombinant human Ku antigen, Nucleic Acids Res. 22 (1994) 3918–3924.
<https://doi.org/10.1093/nar/22.19.3918>.
- [58] Z. Zhang, L. Zhu, D. Lin, F. Chen, D.J. Chen, Y. Chen, The three-dimensional structure of the C-terminal DNA-binding domain of human Ku70., J. Biol. Chem. 276 (2001) 38231–6. <https://doi.org/10.1074/jbc.M105238200>.
- [59] J. Wang, M. Satoh, C.H. Chou, W.H. Reeves, Similar DNA binding properties of free P70 (KU) subunit and P70/P80 heterodimer., FEBS Lett. 351 (1994) 219–24.
[https://doi.org/10.1016/0014-5793\(94\)00863-9](https://doi.org/10.1016/0014-5793(94)00863-9).
- [60] L.A. Hanakahi, 2-Step purification of the Ku DNA repair protein expressed in Escherichia coli, Protein Expr. Purif. 52 (2007) 139–145. <https://doi.org/10.1016/j.pep.2006.10.002>.
- [61] J.R. Walker, R.A. Corpina, J. Goldberg, Structure of the Ku heterodimer bound to DNA and its implications for double-strand break repair., Nature. 412 (2001) 607–14.
<https://doi.org/10.1038/35088000>.
- [62] S. Paillard, F. Strauss, Analysis of the mechanism of interaction of simian Ku protein with DNA., Nucleic Acids Res. 19 (1991) 5619–24. <https://doi.org/10.1093/nar/19.20.5619>.
- [63] M. Yaneva, T. Kowalewski, M.R. Lieber, Interaction of DNA-dependent protein kinase with DNA and with Ku: biochemical and atomic-force microscopy studies., EMBO J. 16 (1997) 5098–112. <https://doi.org/10.1093/emboj/16.16.5098>.
- [64] W.S. Dynan, S. Yoo, Interaction of Ku protein and DNA-dependent protein kinase catalytic subunit with nucleic acids., Nucleic Acids Res. 26 (1998) 1551–1559.
<https://doi.org/10.1093/nar/26.7.1551>.
- [65] E.L. Van Nostrand, G.A. Pratt, A.A. Shishkin, C. Gelboin-Burkhart, M.Y. Fang, B. Sundararaman, S.M. Blue, T.B. Nguyen, C. Surka, K. Elkins, R. Stanton, F. Rigo, M. Guttman, G.W. Yeo, Robust transcriptome-wide discovery of RNA-binding protein binding sites with enhanced CLIP (eCLIP)., Nat. Methods. 13 (2016) 508–14.
<https://doi.org/10.1038/nmeth.3810>.

- [66] M.A. Durney, V.M. D'Souza, Preformed protein-binding motifs in 7SK snRNA: structural and thermodynamic comparisons with retroviral TAR., *J. Mol. Biol.* 404 (2010) 555–67. <https://doi.org/10.1016/j.jmb.2010.08.042>.
- [67] V. V. Pham, C. Salguero, S.N. Khan, J.L. Meagher, W.C. Brown, N. Humbert, H. de Rocquigny, J.L. Smith, V.M. D'Souza, HIV-1 Tat interactions with cellular 7SK and viral TAR RNAs identifies dual structural mimicry, *Nat. Commun.* 9 (2018). <https://doi.org/10.1038/s41467-018-06591-6>.
- [68] T. Mizutani, A. Ishizaka, Y. Suzuki, H. Iba, 7SK small nuclear ribonucleoprotein complex is recruited to the HIV-1 promoter via short viral transcripts, *FEBS Lett.* 588 (2014) 1630–1636. <https://doi.org/10.1016/j.febslet.2014.01.067>.
- [69] 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) 317–22. <https://doi.org/10.1038/35104575>.
- [70] E.A. Dunn, S.D. Rader, Secondary structure of U6 small nuclear RNA: implications for spliceosome assembly., *Biochem. Soc. Trans.* 38 (2010) 1099–104. <https://doi.org/10.1042/BST0381099>.
- [71] S. Yamashita, Y. Takagi, T. Nagaike, K. Tomita, Crystal structures of U6 snRNA-specific terminal uridylyltransferase, *Nat. Commun.* 8 (2017). <https://doi.org/10.1038/ncomms15788>.
- [72] C. Wallet, M. De Rovere, J. Van Assche, F. Daouad, S. De Wit, V. Gautier, P.W.G. Mallon, A. Marcello, C. Van Lint, O. Rohr, C. Schwartz, Microglial Cells: The Main HIV-1 Reservoir in the Brain, *Front. Cell. Infect. Microbiol.* 9 (2019) 362. <https://doi.org/10.3389/fcimb.2019.00362>.
- [73] I. Lebars, D. Martinez-Zapien, A. Durand, J. Coutant, B. Kieffer, A.-C. Dock-Bregeon, HEXIM1 targets a repeated GAUC motif in the riboregulator of transcription 7SK and promotes base pair rearrangements., *Nucleic Acids Res.* 38 (2010) 7749–63. <https://doi.org/10.1093/nar/gkq660>.
- [74] D. Martinez-Zapien, J.M. Saliou, X. Han, C. Atmanene, F. Proux, S. Cianféroni, A.C. Dock-Bregeon, Intermolecular recognition of the non-coding RNA 7SK and HEXIM protein in perspective, *Biochimie.* (2015). <https://doi.org/10.1016/j.biochi.2015.03.020>.
- [75] 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) 630–42. <https://doi.org/10.1128/MCB.26.2.630-642.2006>.
- [76] 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. <https://doi.org/10.1073/pnas.1612331113>.
- [77] H. Liu, C.H. Herrmann, K. Chiang, T.-L. Sung, S.-H. Moon, L.A. Donehower, A.P. Rice, 55K isoform of CDK9 associates with Ku70 and is involved in DNA repair., *Biochem. Biophys. Res. Commun.* 397 (2010) 245–50. <https://doi.org/10.1016/j.bbrc.2010.05.092>.
- [78] L. Shi, D. Qiu, G. Zhao, B. Corthesy, S. Lees-Miller, W.H. Reeves, P.N. Kao, Dynamic binding of Ku80, Ku70 and NF90 to the IL-2 promoter in vivo in activated T-cells, *Nucleic Acids Res.* 35 (2007) 2302–2310. <https://doi.org/10.1093/nar/gkm117>.
- [79] G.L. Mayeur, W.J. Kung, A. Martinez, C. Izumiya, D.J. Chen, H.J. Kung, Ku is a novel transcriptional recycling coactivator of the androgen receptor in prostate cancer cells, *J. Biol. Chem.* 280 (2005) 10827–10833. <https://doi.org/10.1074/jbc.M413336200>.
- [80] H. Wang, R. Fang, J.Y. Cho, T.A. Libermann, P. Oettgen, Positive and negative modulation of the transcriptional activity of the ETS factor ESE-1 through interaction with p300, CREB-binding protein, and Ku 70/86, *J. Biol. Chem.* 279 (2004) 25241–25250. <https://doi.org/10.1074/jbc.M401356200>.
- [81] H. Mizuguchi, K. Miyagi, T. Terao, N. Sakamoto, Y. Yamawaki, T. Adachi, S. Ono, Y. Sasaki, Y. Yoshimura, Y. Kitamura, N. Takeda, H. Fukui, PMA-induced dissociation of Ku86 from the promoter causes transcriptional up-regulation of histamine H(1) receptor., *Sci. Rep.* 2 (2012) 916. <https://doi.org/10.1038/srep00916>.
- [82] G. Nolens, J.C. Pignon, B. Koopmansch, B. Elmoualij, W. Zorzi, E. De Pauw, R. Winkler, Ku proteins interact with activator protein-2 transcription factors and contribute to ERBB2 overexpression in breast cancer cell lines, *Breast Cancer Res.* 11 (2009) R83. <https://doi.org/10.1186/bcr2450>.
- [83] K.W. Trotter, H.A. King, T.K. Archer, Glucocorticoid Receptor Transcriptional Activation via the BRG1-Dependent Recruitment of TOP2 β and Ku70/86., *Mol. Cell. Biol.* 35 (2015) 2799–817. <https://doi.org/10.1128/MCB.00230-15>.
- [84] G. Manic, A. Maurin-Marlin, F. Laurent, I. Vitale, S. Thierry, O. Delelis, P. Dessen, M. Vincendeau, C. Leib-Mösch, U. Hazan, J.F. Mouscadet, S. Bury-Moné, Impact of the Ku Complex on HIV-1 Expression and Latency, *PLoS One.* 8 (2013). <https://doi.org/10.1371/journal.pone.0069691>.
- [85] S. Tyagi, A. Ochem, M. Tyagi, DNA-dependent protein kinase interacts functionally with the RNA polymerase II complex recruited at the human immunodeficiency virus (HIV) long terminal repeat and plays an important role in HIV gene expression, *J. Gen. Virol.* 92 (2011) 1710–1720. <https://doi.org/10.1099/vir.0.029587-0>.
- [86] S. Waninger, K. Kuhen, X. Hu, J.E. Chatterton, F. Wong-Staal, H. Tang, Identification of

- cellular cofactors for human immunodeficiency virus replication via a ribozyme-based genomics approach, *J Virol.* 78 (2004) 12829–12837.
<https://doi.org/10.1128/JVI.78.23.12829>.
- [87] L. Jeanson, F. Subra, S. Vaganay, M. Hervy, E. Marangoni, J. Bourhis, J.-F. Mouscadet, Effect of Ku80 depletion on the preintegrative steps of HIV-1 replication in human cells., *Virology.* 300 (2002) 100–108. <https://doi.org/10.1006/viro.2002.1515>.
- [88] C. Masson, S. Bury-Moné, E. Guiot, A. Saez-Cirion, D. Schoëvaërt-Brossault, C. Brachet-Ducos, O. Delelis, F. Subra, L. Jeanson-Leh, J.-F. Mouscadet, Ku80 participates in the targeting of retroviral transgenes to the chromatin of CHO cells., *J. Virol.* 81 (2007) 7924–7932. <https://doi.org/10.1128/JVI.02015-06>.
- [89] C.T. Sheline, L.H. Milocco, K.A. Jones, Two distinct nuclear transcription factors recognize loop and bulge residues of the HIV-1 TAR RNA hairpin, *Genes Dev.* 5 (1991) 2508–2520. <https://doi.org/10.1101/gad.5.12b.2508>.
- [90] P.D. Bieniasz, T.A. Grdina, H.P. Bogerd, B.R. Cullen, Recruitment of cyclin T1/P-TEFb to an HIV type 1 long terminal repeat promoter proximal RNA target is both necessary and sufficient for full activation of transcription, *Proc. Natl. Acad. Sci. U. S. A.* 96 (1999) 7791–7796. <https://doi.org/10.1073/pnas.96.14.7791>.

Figure legends

Fig. 1. Expression, purification, and DNA binding characteristics of the Ku heterodimer prepared in *E.coli*. **A.** Scheme of the expression vector fragment coding for the Ku protein subunits. **B.** Recombinant N-His6-Ku70/Ku80 protein after two-step purification on Ni-NTA-agarose and heparin-sepharose analyzed by SDS PAGE with subsequent Coomassie blue staining (lane 1), and Western blot with anti-Ku70 (lane 2) and anti-Ku80 (lane 3) antibodies. M - molecular mass markers in kDa. **C.** Gel-shift analysis of 0.1 nM DNA binding with increasing concentrations of recombinant N-His6-Ku70/Ku80. Protein concentrations are indicated above the gel.

Fig. 2. Binding of the recombinant Ku protein to ³²P-labeled RNA ligands tested by EMSA. **A.** Binding of 0.1 nM double-stranded RNA (dsRNA), bulged double-stranded RNA (dsRNA-bulge) and TAR RNA with increasing concentrations of the protein. Positions of free and Ku-bound RNAs are indicated by arrows. **B.** Complex formation between Ku and TAR RNA in the presence of anti-actin (negative control), anti-Ku80, anti-Ku70, or anti-His6 antibodies. Positions of free and Ku-bound RNAs and super-shift of the Ku-RNA complex by an antibody are indicated by arrows. **C.** Binding of 0.1 nM ³²P-labeled hp1RNA with Ku in the presence of increasing concentrations of unlabeled DNA competitor (left gel), and displacement of ³²P-labeled hp1RNA from its pre-formed complex with Ku by increasing concentrations of unlabeled DNA (right gel). **D.** Curves of RNA binding with increasing concentrations of Ku based on the gel-shift assay. Points represent the mean of at least three independent experiments.

Fig. 3. Analysis of RNA structures preferably bound by Ku in cell cultures. **A.** A consensus structure of RNA preferably bound by Ku. **B.** Dependence of the eCLIP signal of RNA bound by human Ku on its similarity of structure to the consensus structure. For black dots, the similarity was calculated relying on the RNA folding (Spearman correlation coefficient $r = 0.43$, p-value 0.021, N=28); for empty square, the similarity calculation considered the presence of GpG sequence in the loop in addition to the RNA folding (Spearman correlation coefficient $r = 0.48$, p-value = 0.009, N=28).

Fig. 4. Ku interacts with 7SK snRNA. **A.** Gel-shift analysis of binding of the ³²P-labeled T7-transcribed 7SK snRNA with increasing concentrations of the recombinant Ku protein. Positions of free and bound 7SK snRNA are indicated by arrows. **B.** Footprinting assay with RNase T1

Journal Pre-proof

digestion of the ^{32}P -labeled T7-transcribed 7SK snRNA in the presence and absence of the recombinant Ku protein. Products of the footprinting assay were resolved by denaturing PAGE (the left gel is 10-15%, and the right gel is 6-10%). 60-mer and 184-mer RNAs were used as mobility controls. Positions of the protected nucleotides are marked by arrows and lines. **C.** Conserved hairpins of 7SK snRNA with binding sites of HEXIM1 and P-TEFb proteins. Indicated G50, G51, G52, and G55 residues were protected by Ku protein in the RNase footprinting assay. **D.** Immunoprecipitation of Ku70-FLAG/Ku80-FLAG from transiently transfected HEK 293T cells using anti-Flag beads. Western blot of input and precipitated proteins with detection by anti-FLAG antibody. **E.** Level of co-precipitated with Ku70-Flag and Ku80-Flag endogenous 7SK and U6 snRNAs detected by RT-qPCR. Cells transfected with the pcDNA3 vector was used as a control. Data from three independent experiments are presented as mean of percentages of the bound input, error bars represent SD (** $p < 0.01$ as compared with the control according to Student's t test).

Fig. 5. Ku interacts with protein components of the 7SK snRNP complex. **A.** Immunoprecipitation of endogenous Ku80, HEXIM1, and Cdk9 from HEK 293T cells using specific anti-HEXIM1, anti-Cdk9, and anti-Ku80 antibodies. Input and precipitation steps were monitored for endogenous Ku80, HEXIM1, and Cdk9. Immunoprecipitation with rabbit IgG was used as a control. **B.** Immunoprecipitation of endogenous Ku80 and Cdk9 from HEK 293T cells in the presence (marked by a plus) or absence (marked by a minus) of RNase cocktail. Input and precipitation steps were monitored for endogenous Ku80 and Cdk9. Immunoprecipitation with rabbit IgG was used as a control. **C.** Anti-Flag immunoprecipitation of HEXIM1-FLAG from transiently transfected HEK 293T cells with either control pcDNA3 or pcDNA3-HEXIM1-3xFLAG vector. Prior immunoprecipitation the cellular lysate was incubated in the presence (+) or absence (-) of RNase cocktail. Western blot using anti-Ku80, anti-Ku70, and anti-Flag antibodies. **D.** Gel-shift analysis of binding of the ^{32}P -labeled T7-transcribed 7SK snRNA with Ku and HEXIM1. Positions of free and bound 7SK snRNA are indicated by arrows. **E.** Nuclear extracts from HEK 293T cells transfected with pNTap-CTIP2 combined to the pcDNA3-Flag or the pcDNA3-Ku70-3xFlag vector were subjected to immunoprecipitations targeting the Flag peptide and analyzed by Western blot for the presence of CTIP2 as described previously [51]. Controls of CTIP2 and Ku70 expressions in nuclear extracts of HEK 293T cells were also accessed by Western blot. **F.** Control microglial cells (lane A) and microglial cells expressing RFP-CTIP2 (Lanes B and C) were subjected to immunostainings with mouse anti-Ku70 / Alexa 647-labeled goat anti-mouse IgG (Lane A and C) or with Alexa 647-labeled goat anti-mouse IgG

alone (Lane B). Nuclei were stained by Hoechst as indicated. Patterns of CTIP2 and Ku70 expressions were analyzed by confocal microscopy.

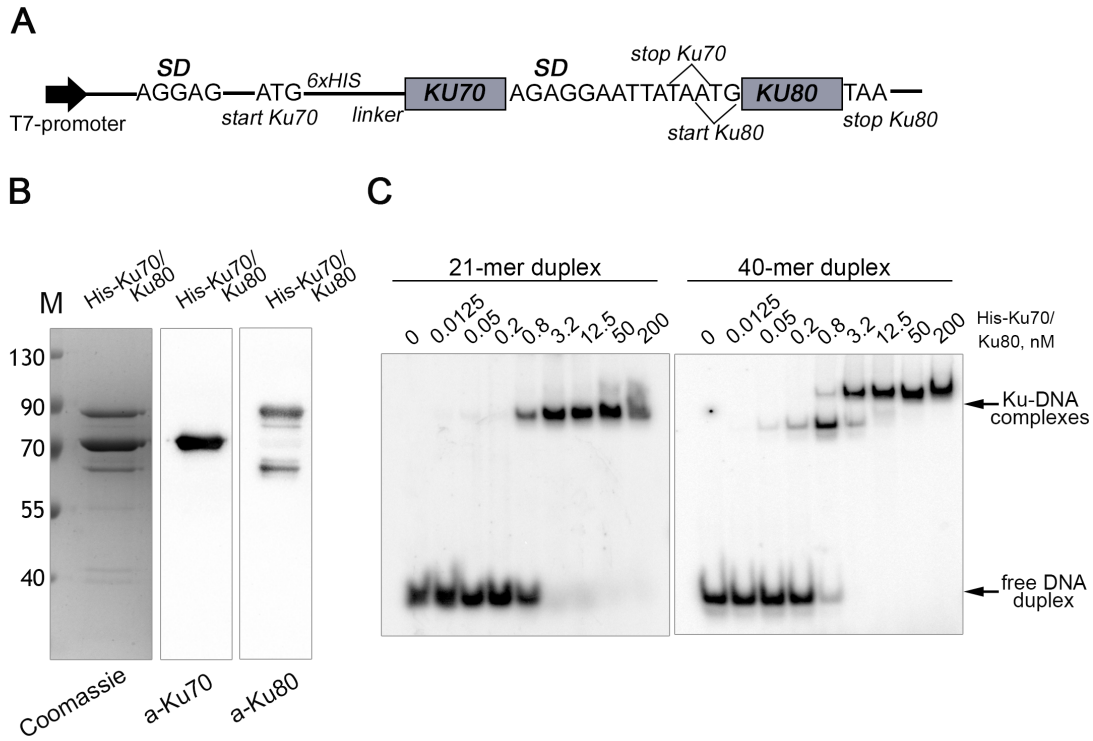
Journal Pre-proof

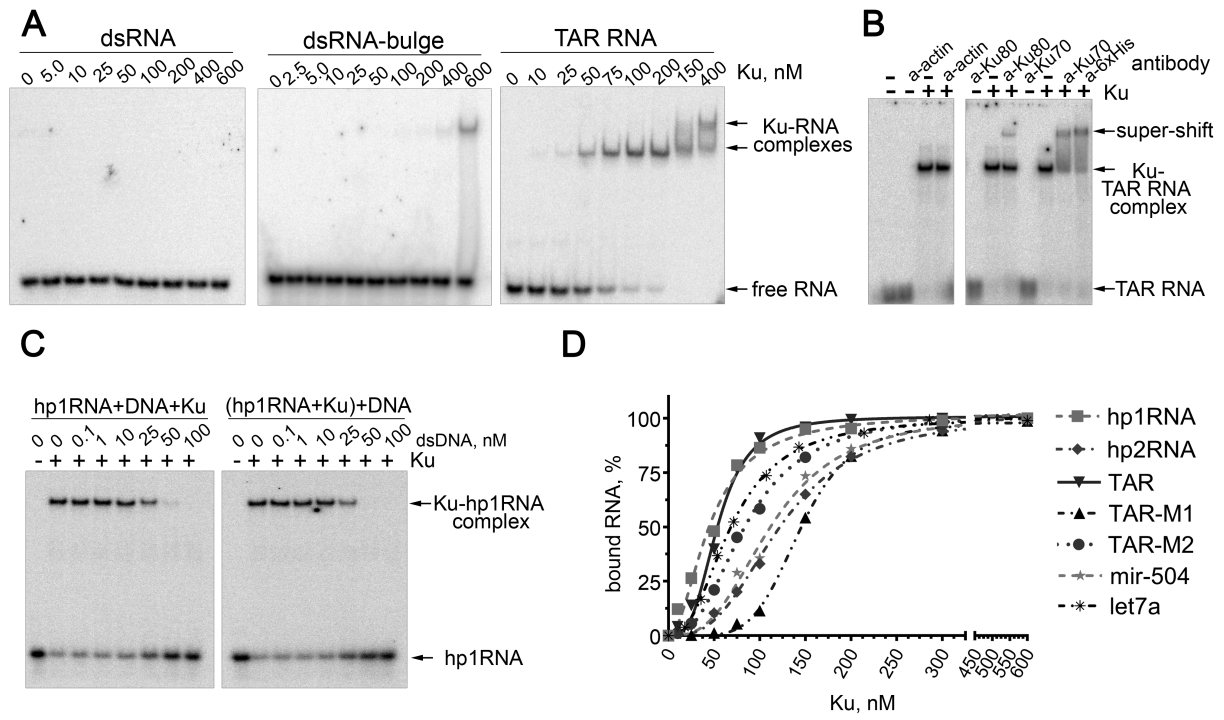
Table 1. Structures of DNA and RNA molecules used in the study and stability of their complexes with the recombinant Ku protein

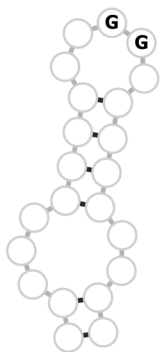
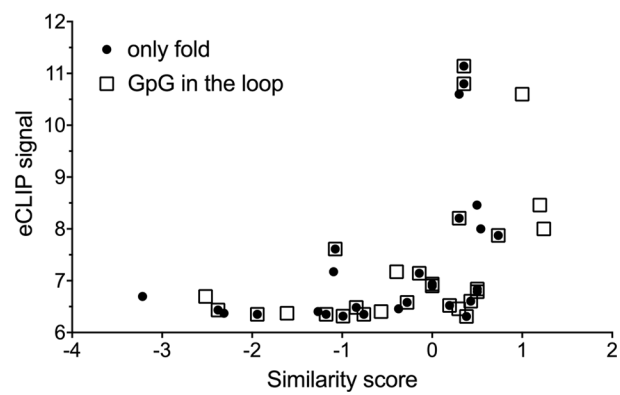
Name	Structure	K_d , nM*
21-mer DNA	5'-gtgtggaaaatctctagcagt-3' 3'-cacaccttttagagatcgtca-5'	0.9±0.3
40-mer DNA	5'-gactacggttcaagtcagcgtgtggaaaatctctagcagt-3' 3'-ctgatgccaaagttcagtcgcacaccttttagagatcgtca-5'	0.5±0.2
dsRNA	5'-aucaucaaacccguuuuagc-3' 3'-uaguaguuuuggcaaaaucg-5'	NB
dsRNA- bulge	5'-guugugcucc ucuacgcguc-3' 3'-caacacgagg _{acc} agaugcgag-5'	NB
TAR	5'-cugguuag ^a ccaga ^u c ^u gagc ^c u ^g 3'-gaucaauc ggucu cucg _a g ^g	54 ± 6
TAR-M1	5'-cugguuag ^a ccagagagc ^c u ^g 3'-gaucaauc ggucucucg _a g ^g	145 ± 17
TAR-M2	5'-cugguuag ^a ccaga ^u c ^u gagc ^c c ^a 3'-gaucaauc ggucu cucg _a a ^a	85 ± 16
hp1RNA	5'-gg cgucc ^c uc ccga ^a gc ^u g 3'-cc _a gcagg ac _a ggcu cg _c g ^c agc ^u	45 ± 10
hp2RNA	5'-gucccuc ^a cccacc ^c u ^c g ^g 3'-cagggag _c gggugg ^a acc ^c u ^a	125 ± 20
let7a	5'-cg _a uc ^c uccgacauguc ^a aa uag ^a g ^g 3'-gg _u ag _u agguuguauagu ^u auc ^a a ^g a _u ca _a g ^g g _a aa ^u	70 ± 7
mir-504	5'-gagaccug ^g ucugcacucu ^a uc ^u gua ^u u 3'-cuuugggac ^g ggacgugagg ^g ag ^u cau ^u c ^c g _a u	115 ± 11

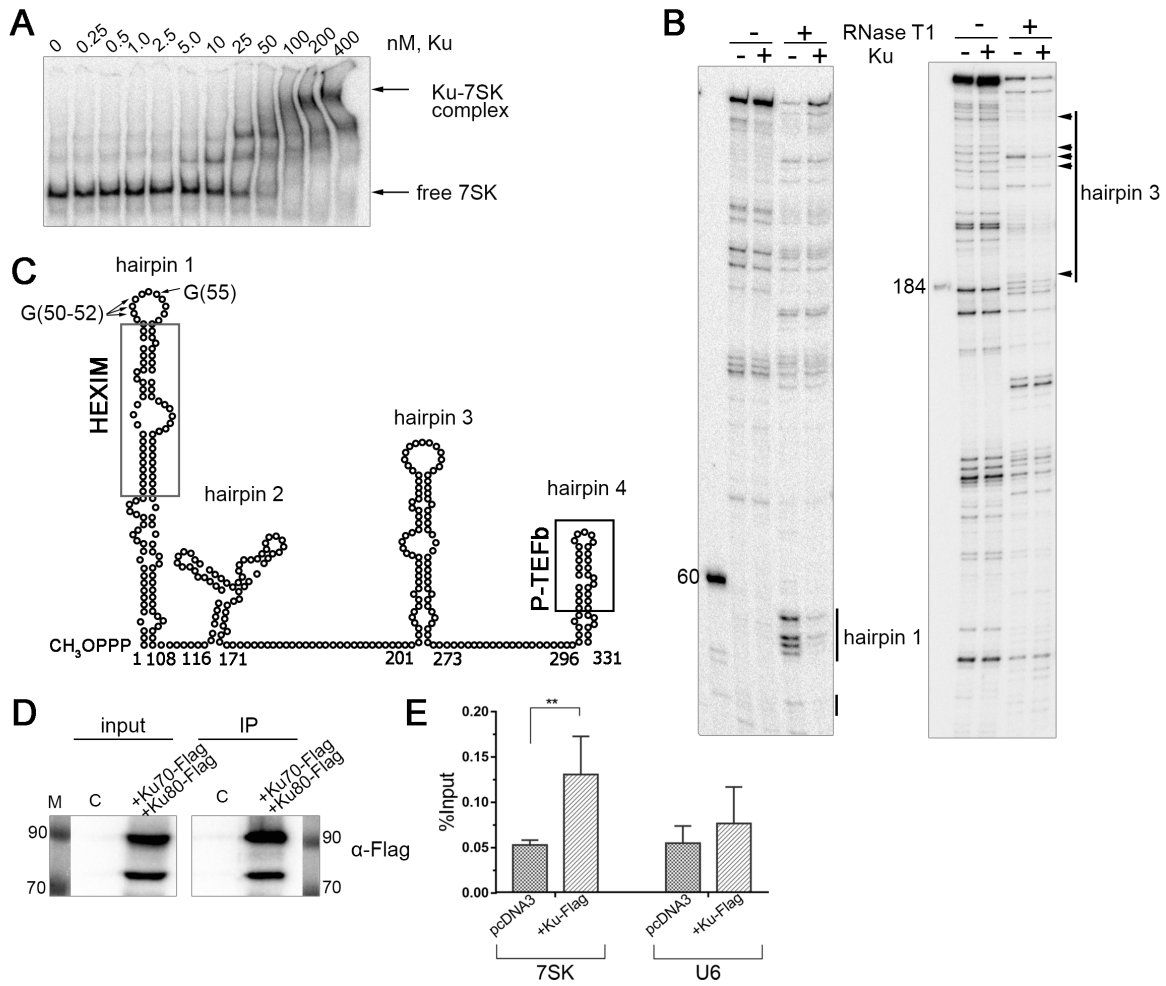
NB – No binding detected.

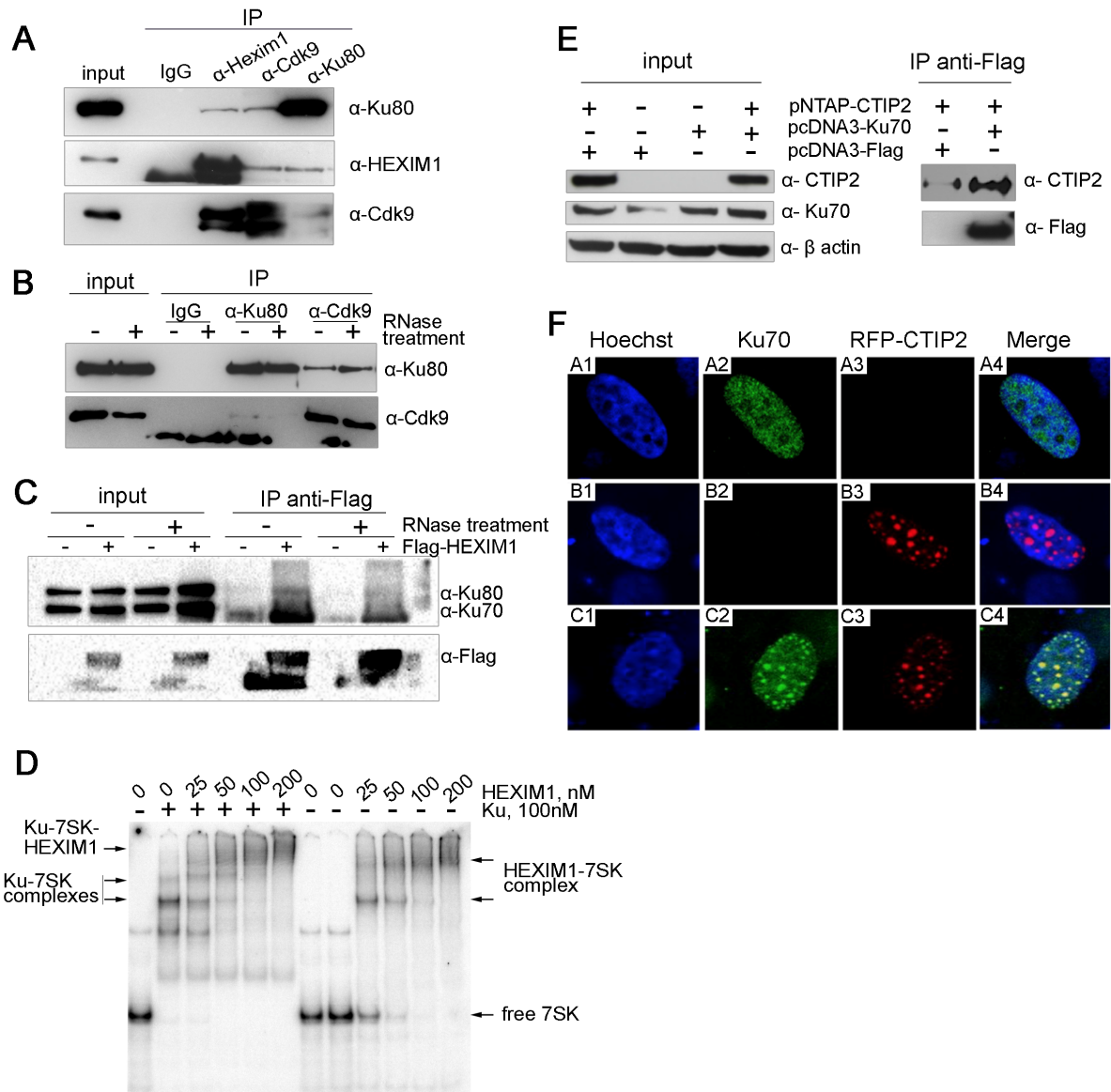
*The mean of at least three independent experiments with standard deviations





A**B**





Highlights

Human Ku interacts with bulge-containing stem-loop RNA structures

Ku binds 7SK snRNA at the terminal G-rich loop of its hairpin 1

Ku interacts with proteins of 7SK snRNP: HEXIM1, Cdk9 and CTIP2

Journal Pre-proof

Declaration of interests

✓ The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Journal Pre-proof