# Identification of hemicatenane-specific binding proteins by fractionation of HeLa nuclei extracts

Oumayma Rouis [1], Cédric Broussard [2], François Guillonneau [2], Jean-Baptiste Boulé [3] and Emmanuelle Delagoutte [3,*]

[1] Current address: YposKesi, 26 rue Henri Auguste-Desbrueres F-91100 Corbeil-Essonnes, France

[2] Plateforme de Protéomique 3P5, Université de Paris, Institut Cochin, INSERM U1016, CNRS UMR8104, F-75014 Paris, France

[3] Genome Structure and Instability laboratory, Life Adaptations department, CNRS UMR 7196, INSERM U1154, National Museum of Natural History, F-75005 Paris, France

* Corresponding author. Tel: + 33 1 4079 4832; E-mail: emmanuelle.delagoutte@mnhn.fr

Running title: Isolation of HC-binding proteins

## Abstract

DNA hemicatenanes (HCs) are four-way junctions in which one strand of a double stranded helix is catenated with one strand of another double stranded DNA. Frequently mentioned as DNA replication, recombination and repair intermediates, they have been proposed to participate in the spatial organization of chromosomes and in the regulation of gene expression. To explore potential roles of HCs in genome metabolism, we sought to purify proteins capable of binding specifically HCs by fractionating nuclear extracts from HeLa cells. This approach identified three RNA-binding proteins: the Tudor-Staphylococcal Nuclease Domain 1 (SND1) protein and two proteins from the Drosophila behavior human splicing family, the ParaSpeckle Protein Component 1 and the Splicing Factor Proline- and Glutamine- rich protein. Since these proteins were partially pure after fractionation, truncated forms of these proteins were expressed in *E. coli* and purified to near homogeneity. The specificity of their interaction with HCs was re-examined *in vitro*. The two truncated purified SND1 proteins exhibited specificity for HCs, opening the interesting possibility of a link between the basic transcription machinery and HC structures *via* SND1.

**Key words**: DNA hemicatenanes/protein nuclear extract/fractionation/binding specificity/Tudor-Staphylococcal Nuclease Domain 1

## Introduction

DNA hemicatenane (HC) consists of a junction of four DNA strands in which one strand of one double helix passes in between the two strands of another double helix (Fig 1A). They can implicate one or two different double stranded DNA molecules and, in this later case, the two DNA helices may or may not be homologous. DNA HCs might appear at several times during the cell cycle. For instance, the convergent migration of two Holliday junctions formed during DNA recombination may lead to a HC depending on the relative topology of the two junctions. HC may be created during the initiation [1] and termination [2,3] phases of DNA replication or when replication is inhibited [4]. Their presence close to a replication fork in progress may help repair and restart a stalled DNA replication fork, due to the proximity of the newly synthesized daughter DNA molecules [1]. Finally, the ultra fine bridges that can be observed during the anaphase step of mitosis after a perturbed or incomplete DNA replication may consist of two homologous chromosomes topologically linked by HCs (for recent reviews see [5–7]).

Faithful and accurate processing of HCs is essential for genome stability. For HCs resulting from the convergent migration of two Holliday junctions, the process of unlinking the two DNA molecules of the HC is called dissolution (for reviews see [8,9]). In humans, the dissolvasome complex (BTRR) is responsible for the branch migration and dissolution reactions. In this complex, the RecQ-like Bloom helicase (B in BTRR) uses its ATPase activity to migrate the Holliday junctions [10,11] and its oligomeric structure may be responsible for the convergent migration of two Holliday junctions [8]. The Topoisomerase IIIα (T in BTRR), a type IA topoisomerase, through a transesterification and strand passage reaction, decatenates the HC. Its activity is stimulated by the RecQ-mediated instability, RMI1 and RMI2 (RR in BTRR), factors [12–14]. Devoid of catalytic activity, the RMI1 might stabilize the open form of the Topoisomerase IIIα to favor strand passage. Finally, one family of ultra fine bridges has been shown to be decorated by the BTRR complex and it is possible that the decatenase activity of Topoisomerase IIIα be used to resolve their inter-strand links [15,16].

*In vitro* experiments have shown that re-association of DNA fragments containing the polyCA/polyTG repeat sequences leads to the formation of an alternative structure that has been proposed to be a HC [17]. The High Mobility Group Box 1 and 2 (HMGB1 and HMGB2) proteins stimulate this process [18,19]. Following this observation, it has been proposed that HCs may participate in the genome organization and in the regulation of its transcription [20]. In this model, genome is organized in large loops maintained at their base by topological knots, such as HCs. Unless stabilized, knots can migrate and their location on the chromosome can change. Through the possibility of knots migration and stabilization, chromosomal loops may expose a variety of different DNA sequences, making thus possible the expression of different genes. Inherent to this model is the reversible and transient stabilization of the knots, either by a specific sequence or by some HC-specific proteins.

Our goal in this work was to explore further these hypotheses by seeking nuclear proteins capable of stabilizing HC structures by specifically binding to them. We adapted

a previously published procedure to build HCs from two distinct double stranded DNA mini-circles (dsMCs) [21] and used the purified HCs to fractionate nuclear extracts from HeLa cells over several chromatographic media. This approach identified three RNA-binding proteins capable of binding to HC with significant specificity in our binding assay: the Tudor-Staphylococcal Nuclease Domain 1 (SND1) protein and two proteins from the Drosophila behavior human splicing (DBHS) family, the ParaSpeckle Protein Component 1 (PSPC1) and the Splicing Factor Proline- and Glutamine- rich protein (SFPQ). Since these proteins were partially pure after fractionation of the HeLa nuclear extracts, and because of the difficulty to purify them in a full length form, truncated forms of these proteins were expressed in *E. coli* and purified to near homogeneity. Specificity of the interaction between the purified proteins and HCs was re-examined *in vitro*. Our results show that among the three recombinant proteins, SND1 exhibits the most specific binding to HC. These results suggest that SND1 could bridge the basic transcription machinery to HC structures.

## Materials and Methods

### Materials

The HeLa nuclei were from Ipracell and were resuspended in 10 mM Hepes pH 7.9, 1.5 mM $MgCl_2$, 10 mM KCl, 0.5 mM DTT. Oligonucleotides were from Eurogentec. T4 polynucleotide kinase, T4 DNA ligase, Nt.BbvcI and Nb.BbvcI nicking enzymes were from New England Biolabs. Wheat germ topoisomerase was from Inspiralis. *E. coli* SSB was from Ubs. $\gamma^{32}$P-ATP was from Perkin Elmer. Trypsin was from Promega (USA). Cellulose phosphate was from Sigma. The HiTrap ButylS-FF column of 1 mL and the gel filtration column Superdex 200 10/300 GL were from GE Healthcare. MacroPrep HighS resin was from Biorad. All columns were prepared as recommended by the manufacturer and equilibrated in the indicated loading buffer before use. λ DNA was from Sigma. Acrylamide and bisacrylamide stock solutions were from Euromedex. Proteinase inhibitor cocktail was from Roche. Anti-SND1 and anti-PTPB1 antibodies were from St John's Laboratory. Anti-SFPQ antibody was from Novusbio. Anti-PSPC1 was from Abcam. The sequence of the oligonucleotide OL21 (from Eurogenetec) was:

5'-CCCTAACCCTAACCCTAACCC-3'.

### Preparation of HCs

HCs were prepared as described in [21] with three modifications in the procedure. Firstly, the production of dsMCs was performed in the presence of ethidium bromide (0.35 μg/mL) to favor circularization of DNA fragments into monomeric circles. Secondly, dsMCs were purified on a 4% acrylamide (29:1 = acrylamide:bisacrylamide mass ratio) gel made in TBE 0.5X before performing the nicking reaction. Thirdly, after the nicking reaction, the circular and linear single stranded fragments were purified on a 4% acrylamide (29:1 = acrylamide:bisacrylamide mass ratio) gel made in TBE 0.5X + 7M Urea. The sequences of dsMCs, linker and helper oligonucleotides can be found in [21].

## Preparation of protein extract from HeLa nuclei

6 mL of lysis buffer containing 10 mM Hepes pH 7.8, 0.6 M NaCl, 1.5 mM $MgCl_2$, 5 mM DTT, 0.8 mg/mL PMSF and a cocktail of proteinase inhibitors were added to the 12 mL containing $15x10^9$ HeLa nuclei and left on ice until thawed. While gently vortexing the nuclei, EDTA was added to a final concentration of 2.5 mM and 410 µL of NaCl 5 M were added four times at intervals of five minutes. The 0.6 M NaCl final concentration that was reached at the end permitted to the chromatin to gently swell. The sample was left on ice for 45 minutes and centrifuged 60 minutes at 4°C and at 184,000 g. The supernatant of the centrifugation that constituted the nuclei protein extract was next fractionated by different means (see "Fractionation of nuclei proteins" section). Each step of fractionation generated fractions that were tested for containing HC-specific binding proteins (see "Test of chromatographic fractions for containing HC-specific binding activity"). Interesting fractions were next combined and applied to the next step of the fractionation procedure.

## Test of chromatographic fractions for containing a HC-specific binding activity

Unless indicated otherwise, the test consisted of performing an interaction between DNA (HC or dsMC) and chromatography fractions by mixing 8 µL of each fraction with 1 µL of DNA (0.1 femtomole) in a 1.5 mL Eppendorf tube. Competitor λ DNA (48502 base pairs) was added at the indicated concentration (from 0 to 30 ng (1 femtomole) of DNA molecule). The mixtures were incubated for 30 minutes at 22°C. At the end of the incubation glycerol was added to a final concentration of 5%. The samples were loaded on a 4% acrylamide (29:1 = acrylamide:bisacrylamide mass ratio) gel made in Tris-Borate 89 mM, boric acid 89 mM EDTA 2 mM (TBE) 0.5X. Electrophoresis was performed at 4°C for 4 hours and at 150 V. The gel was dried and exposed on a $^{32}P$ sensitive screen. The fractions that permitted the HC (and not the dsMC) to shift by decreasing its mobility in the gel were considered as containing HC-specific binding proteins.

## Fractionation of nuclei protein extract

The first step of protein fractionation was a protein precipitation that was performed with 5% of ammonium sulfate. Ammonium sulfate was slowly added to the protein extract under slow stirring and at 4°C. Proteins were let precipitated overnight. Precipitated proteins were recovered by centrifugation (20 minutes, 4°C, 1200 g). The pellet of proteins was resolubilized in 27 mL of a low salt (Low KP) buffer containing 25 mM potassium phosphate pH 7.5, 2 mM EDTA, 10% glycerol, 12.5 mM DTT. The resolubilized proteins were next loaded at a 0.5 mL/min flow rate on a column of cellulose phosphate of 4 mL equilibrated in the Low KP buffer. After the load of proteins, the column was washed with 20 mL of Low KP buffer. The bound proteins were eluted with a high salt buffer containing 500 mM potassium phosphate pH 7.5, 2 mM EDTA, 10% glycerol, 12.5 mM DTT and contained HC-specific binding proteins. They were combined and dialyzed against a high salt buffer (High KP) containing 500 mM potassium phosphate pH 7.5, 2 mM EDTA, 10% glycerol, 2 mM DTT. After dialysis, the protein sample was loaded at a flow rate of 0.3 mL/min on the 1 mL HighTrap ButylS-FF column equilibrated with the High KP buffer. The flow through of the column was collected and dialyzed overnight against a low salt buffer containing 25 mM potassium phosphate pH 6.6, 25 mM KCl, 2 mM EDTA, 10% glycerol, 2 mM DTT. After dialysis, the protein sample was loaded on a 1 mL MacroPrep High S at a flow rate of 0.7 mL/min.

After loading, the column was washed with 15 mL of the low salt buffer (25 mM potassium phosphate pH 6.6, 25 mM KCl, 2 mM EDTA, 10% glycerol, 2 mM DTT). The proteins were eluted with a 15 mL salt gradient (KCl raising from 25 mM to 500 mM). Fractions containing HC-specific binding proteins were pooled, concentrated and loaded on a Superdex 200 (10/300 GL) equilibrated in a buffer containing 25 mM potassium phosphate pH 7.5, 150 mM KCl, 2 mM EDTA, 10% glycerol, 2 mM DTT. Fractions of interest were kept at -80°C in small aliquots.

**Sample preparation for mass spectrometry analysis**

A large scale interaction between 16 femtomoles of HC and the designated chromatography fraction was performed in a final volume of 150 µL of buffer containing 7.5 mM potassium phosphate buffer pH 7.5, 34 mM Tris HCl pH 7.5, 45 mM KCl, 34 mM NaCl, 1 mM EDTA, 0.03 % Triton X100, 3% glycerol, 2 mM DTT. When present competitor λ DNA was added at the indicated concentration. The control sample that was included in the analysis consisted of the protein fraction alone (no HC) made in the same interaction buffer. The mixtures were next incubated for 30 minutes at 22°C. At the end of the incubation glycerol was added to a final concentration of 5%. The samples were loaded on a 4% acrylamide (29:1 = acrylamide:bisacrylamide mass ratio) gel made in TBE 0.5X. Electrophoresis was performed at 4°C for 4 hours and at 150 V. After electrophoresis the gel was exposed on a $^{32}P$ sensitive screen. The material (as gel cubes manually excised from the gel) that migrated above the HC was analyzed by mass spectrometry for its protein content.

**Protein identification by LC-MS/MS analysis**

In-gel digestion was carried out with trypsin: sample were washed twice with a mixture of 100 mM ammonium bicarbonate (ABC) and 50% (vol/vol) acetonitrile (ACN) for 20 min at room temperature and then dehydrated using 100% ACN for 20 min, before being reduced with 25 mM ABC containing 10 mM DTT for 1 h at 56 °C and alkylated with 55 mM iodoacetamide in 25 mM ABC for 30 min in the dark at room temperature. Gel pieces were washed twice with 25 mM ABC and dehydrated (twice, 20 min) with 100% ACN. Gel cubes were incubated with sequencing grade-modified trypsin (12.5 ng/µl in 40 mM ABC with 10% ACN, pH 8.0) overnight at 37 °C. After digestion, peptides were extracted twice from gel pieces with a mixture of 50% ACN – 5% formic acid (FA) and then with 100% ACN. Extracts were dried using a vacuum centrifuge concentrator plus (Eppendorf).

Mass spectrometry (MS) analyses were performed on a Dionex U3000 RSLC nano-LC system coupled to an Orbitrap Fusion Tribrid mass spectrometer (Thermo Fisher Scientific). After drying, peptides were solubilized in 7 µL of 0.1 % trifluoroacetic acid (TFA) containing 10 % acetonitrile (ACN). One µL was loaded, concentrated and washed for 3 min on a $C_{18}$ reverse phase precolumn (3 µm particle size, 100 Å pore size, 75 µm inner diameter, 2 cm length, Thermo Fisher Scientific). Peptides were separated on a $C_{18}$ reverse phase resin (2 µm particle size, 100 Å pore size, 75 µm inner diameter, 25 cm length from Thermo Fisher Scientific) with a 30 minutes gradient starting from 99 % of solvent A containing 0.1 % FA in $H_2O$ and ending in 90 % of solvent B containing 80 % ACN, 0.085 % FA in $H_2O$. The mass spectrometer acquired data throughout the elution process and operated in a data-dependent scheme with full MS scans acquired with the Orbitrap, followed by MS/MS HCD fragmentations acquired with the Ion Trap on the most abundant ions detected in top speed mode for 3 seconds. Resolution was set to

60,000 for full scans at AGC target 2.0e5 within 60 ms maximum injection ion time (MIIT). The MS scans spanned from 350 to 1500 m/z. Precursor selection window was set at 1.6 m/z, and MS/MS scan resolution was set with AGC target 2.0e4 within 100 ms MIIT. HCD Collision Energy was set at 30 %. Dynamic exclusion was set to 30 s duration. For the spectral processing, the software used to generate .mgf files was Proteome Discoverer 1.4 (ThermoFisher Scientific). The mass spectrometry data were analyzed using Mascot v2.5 (Matrix science) on *Homo sapiens* (20,243 sequences) from the SwissProt databank containing 553,655 sequences; 198,177,566 residues (April 2017). The enzyme specificity was Trypsin's and up to 1 missed cleavage was tolerated. The precursor mass tolerance was set to 4 ppm and the fragment mass tolerance to 0.55 Da for Fusion data. Carbamidomethylation of cysteins and oxidation of methionines were set as variable modifications. Among positive identifications based on a Mascot score above the significance threshold p<5%, we selected proteins identified with at least 2 peptides with an ion score > 25 for each of them.

**Sucrose gradient**

The solution used to prepare the sucrose gradient contained 50 mM Tris HCl p7.5, 50 mM NaCl, 0.5 mM EDTA, 2 mM DTT and either 5 % or 40 % sucrose. A 5 mL gradient of 5 to 40 % sucrose was prepared with a gradient maker and a peristaltic pump and equilibrated at 8°C for two hours before use. The interaction between HC and the designated fraction from the gel filtration was performed in a 90 µL volume and its product was carefully loaded on top of the sucrose gradient. Two control samples were also run at the same time: a sample containing only the HC and a sample containing only the designated fraction from the gel filtration. The centrifugation was performed at 8°C, 64,090g in SW 55Ti rotor for 12 hours. At the end of the run, fractions of 200 µL were recovered from the top to the bottom of the gradient and tested for the presence of (i) HC by measuring their reactivity level and (ii) the SND1 protein by western blot.

**Western Blot**

Western blot was performed with nitrocellulose membrane and in TBS buffer supplemented with 0.1 % Tween20 and 5 % low fat milk. Anti-PSPC1 and anti-SND1 antibodies were used at a final concentration of 1 µg/mL, anti-SFPQ at 0.4 µg/mL (final concentration) and anti-PTBP1 at 0.1 µg/L (final concentration).

**Purification of SND1 proteins**

Plasmids overexpressing the human Tudor-Staphylococcal Nuclease-like protein from amino acids 33 to 888 (SND1-110) and from amino acids 315 to 863 (SND1-64) were a gift from Pr. Hanna S. Yuan. The purification of the two proteins was performed as described in [22].

**Purification of PSPC1 homo-dimer and PSPC1-SFPQ hetero-dimer**

Plasmids overexpressing PSPC1 protein from residues 53 to 320 and SFPQ from residues 276 to 535 were a gift from Pr. Mihwa Lee. The purification of PSPC1 was performed as described in the thesis manuscript of Daniel Michael Passon [23]. The purification of the PSPC1-SFPQ heterodimer was performed as described in [24].

### Interaction between DNA and purified proteins

Unless mentioned otherwise, the interaction between 0.1 femtomole of DNA (radiolabeled HC, radiolabeled dsMC or radiolabeled circular single stranded DNA C10ss) and the purified proteins was performed in 8 µL of TENT buffer (10 mM Tris HCl pH pH 7.5, 50 mM NaCl, 0.5 mM EDTA, 0.05 % Triton X100) supplemented with 2 mM DTT. The concentrations of proteins were as indicated in the figure legend. The mixtures were incubated for 30 minutes at 22°C. At the end of the incubation glycerol was added to a final concentration of 5 %. The samples were loaded on a 4 % acrylamide (29:1 = acrylamide:bisacrylamide mass ratio) gel made in TBE 0.5X. Electrophoresis was performed at 4°C for 3 hours and at 150 V. The gel was dried and exposed on a $^{32}$P sensitive screen. After exposure, the screen was scanned and quantification of the gel was performed using the ImageQuant TL Software.

## Results and discussion

Our search for HC-specific binding proteins required radiolabeled DNA HCs. These DNA structures were built as described in [21] and the procedure of their preparation is summarized in Fig 1B. The starting material was two distinct radiolabeled dsMCs, one of 215 base pairs (dsMC10) and one of 235 base pairs (dsMC09), produced after circularization of $^{32}$P-end labeled fragments. The nicking of the dsMCs at specific sites provided linear single stranded fragments of 215 (L10ss) and 235 (L09ss) nucleotides, and circular single stranded DNAs of 215 (C10ss) and 235 (C09ss) nucleotides (Fig 1B). The linear single stranded L09ss and L10ss fragments were annealed to the linker oligonucleotide (Fig 1B). This reaction permitted to bring L09ss and L10ss in close proximity, due to the complementarity of the linker oligonucleotide to 25 nucleotides of L09ss and 25 nucleotides of L10ss. Under these conditions and in the presence of helper oligonucleotides, the circularization of the linear single stranded L09ss and L10ss fragments into single stranded catenanes was favored. C09ss and C10ss were re-annealed on the single stranded catenanes with the wheat germ topoisomerase I to finally give DNA HCs.

### Fractionation of HeLa nuclei protein extracts

To isolate proteins that bind specifically HCs, we first prepared a protein nuclear extract from HeLa nuclei cells at 0.6 M NaCl. At this salt concentration, the nuclei chromatin swells and the proteins that fall off the DNA can be recovered by a high-speed centrifugation. We next fractionated this nuclear protein extract following a procedure (Fig 2A) that consisted in five steps: a 5 % ammonium sulfate precipitation followed by four chromatographies: 1) phosphocellulose chromatography, 2) hydrophobic chromatography, cationic exchange chromatography and finally 4) a size exclusion chromatography. Along the fractionation procedure, the fractions were selected for containing HC-specific binding proteins based on a simple and robust binding assay (Fig 2B): for example, proteins contained in the fraction A10 of the size exclusion chromatography did bind HC (lanes 5-8 of Fig 2B) but not dsMC09 (lanes 1-4 of Fig 2B). Furthermore, the proteins assembled on HC efficiently resisted the competition with λ DNA since excess of λ DNA did not completely abolish the protein-DNA interaction

(lanes 5-8 of Fig 2B). At the end of this 5-step procedure the fractions were not homogeneous (Fig 2C) but greatly enriched in proteins potentially exhibiting binding specificity to HC.

**Identification of HC-specific binding proteins by mass spectrometry**

To identify the proteins assembled on the HC, a large scale (150 μL) interaction between HC and the fraction A10 of the size exclusion chromatography was performed, free DNA and protein-DNA complexes were resolved by electrophoresis on a polyacrylamide gel (lanes 3 and 5of Fig 3A) and the material that migrated slower than the naked HC was analyzed for its protein content by mass spectrometry (Table 1). Our mass spectrometry analysis only focused on the protein-HC complexes that had the same mobility (relative to HC) in the small (lane 5 of Fig 2B) and large scale experiments (Figs 3A and 3B). In particular, the protein-HC complexes of high molecular weight (lanes 3 and 5 of Fig 3A and lane 3 of Fig 3B) were not included as these complexes were not observed in the small scale experiment (lane 5 of Fig 2B). Note in addition that (i) two concentrations of competitor λ DNA were tested: 12 (lanes 2 and 3 of Fig 3A) or 23 ng (lanes 4 and 5 of Fig 3A) and (ii) the material above the naked HC and used for mass spectrometry analysis was divided into two gel pieces (green rectangles on Fig 3A). Control experiments included the fraction A10 of the size exclusion chromatography without HC but supplemented with the indicated amount of competitor λ DNA (lanes 2 and 4 of Fig 3A). The mass spectrometry analysis identified a number of proteins (listed in Supplementary Table 1). Proteins identified in the samples containing the protein fraction with HC (green rectangles, lanes 3 and 5 of Figure 3A) were compared with the proteins identified in the samples containing the protein fraction without HC (green rectangles, lanes 2 and 4 of Figure 3A). Only the proteins identified specifically in the samples containing the protein fraction with HC with a good score (Mascot score > 25 and number of peptides ≥ 2) are listed in Table 1. Among them were the ParaSpeckle Protein Component 1 (PSPC1), the Splicing Factor Proline- and Glutamine- rich protein (SFPQ), the Tudor-Staphylococcal Nuclease-like protein (SND1) and the Annexin A2 protein score (Table 1, lanes 3 and 5 of Fig 3A). We analyzed the fraction A11 of the size exclusion chromatography in the same manner (Fig 3B) and found with a good score the PSPC1 and the Polypyrimidine Tract Binding Protein (PTBP1) proteins specifically in the sample containing HC and proteins (Table 1, lane 3 of Fig 3B). The High Mobility Group Box 1 and 2 (HMGB1 and HMGB2) proteins were previously found to stimulate the re-association of DNA fragments containing the polyCA/polyTG repeat sequences into an alternative structure that was proposed to be a HC [17]. Interestingly, they were not retrieved by our analysis. The protein nuclear extract from HeLa nuclei cells was prepared at 0.6 M NaCl. Therefore it is possible that HMGB1 and HMGB2 proteins were still bound to the chromatin at this salt concentration and therefore not recovered by the high speed centrifugation. Alternatively, HMGB1 and HMGB2 could exhibit sequence specificity. Our results call for testing the binding of purified HMGB1 and HMGB2 to our HC substrate.

We focused our interest on three proteins, SND1, PSPC1 and SFPQ for two reasons. Firstly, HCs may participate in the regulation of the genome transcription [20], and secondly, SND1, PSPC1 and SFPQ are key players of the RNA metabolism, including DNA

transcription [25–31]. We checked by western blot for their presence in the size exclusion fractions A10 and A11 (Fig 3C). Results indicated that the three proteins could indeed be detected into the protein fraction A10. Together with the PSPC1 protein, PTPB1 identified in the 11p-β gel piece (Table 1, Fig 3B) could also convincingly be detected in the fraction A11 (Fig 3C). We note that the anti-PSPC1 antibody highlighted two bands in the A10 and A11 fractions, the band of ≈ 65 kDa (indicated by an arrow on lanes 2 and 3 of Fig 3C) corresponding to the full length protein (predicted MW = 59 kDa).

## Characterization of the PSPC1-HC and (PSPC1-SFPQ)-HC nucleoprotein complexes with purified proteins

PSPC1 and SFPQ are two members of the Drosophila behavior human splicing (DBHS) family that includes nuclear proteins implicated in various nuclear functions, such as RNA biogenesis and transport, paraspeckle formation or DNA repair (for a review on DBHS proteins, see [31]). Devoid of catalytic activities but capable of binding a variety of proteins and nucleic acids, members of the DBHS family have been proposed to serve as a "multipurpose molecular scaffold". DBHS proteins carry a highly conserved DBHS region that consists of two tandem RNA recognition motifs (RRM1 and RRM2 on Fig 4A), a NonA/paraspeckle domain (NOPS on Fig 4A) and a coiled-coil domain. These four domains are responsible for homo- and/or hetero-dimerization of the proteins. They are flanked at their carboxy terminal side by an intrinsically disordered region, G-rich in case of SFPQ protein and GP-rich in case of PSPC1 and a nuclear localization signal (NLS on Fig 4A). The sequences that flank the DBHS core at the amino terminal side of the core are also of low complexity and unstructured. The additional amino terminal domain of PSPC1 is AP-rich whereas that of SFPQ is longer, carries a GPQ-rich domain and a DNA binding region (DBD on Fig 4A) that might function in binding to gene promoters [32] and to DNA damages [33].

To characterize the PSPC1-HC and (PSCP1-SFPQ)-HC nucleoprotein complexes, we purified truncated forms of PSPC1 and SPFQ proteins known to assemble into dimers [23,24]: the truncated PSCP1 homo-dimer was from amino acids 53 to 320 and the truncated SFPQ protein in the PSCP1-SFPQ hetero-dimer was from residue 276 to 535 (indicated as double arrows underneath each schematic representation of the protein on Fig 4A). The choice to purify truncated forms of the proteins stems from the difficulty to purify the full length proteins. Both forms of truncated dimers contained the two tandem RNA recognition motifs, the NonA/paraspeckle domain but had a truncated carboxy terminal coiled-coiled domain. The purified proteins were tested for their efficiency of binding to HC and to the circular single stranded DNA coming from the nicking of dsMC10 mini-circle, C10ss (Fig 1B). High concentrations of PSPC1 and PSCP1-SFPQ (> 1 μM) were required to make possible the binding of the proteins to the HC (Figs 4B and 4D) whereas within these ranges of concentration, binding to C10ss was highly efficient (Figs 4C and 4E). For instance, at 9 μM PSPC1, 37% +/-1.5% of the HC was engaged in a complex with PSPC1 (Lane 5 of Fig 4B) whereas the extent of assembly of PSPC1 on C10ss reached 95% +/- 1.5 % (Lane 5 of Fig 4C). These results indicated that the binding of the truncated forms of PSPC1 homo-dimer and PSCP1-SFPQ hetero-dimer to HC was not highly specific. It is possible that the full length forms of the homo- and hetero-dimers and/or their post-translational modifications mainly carried by the

intrinsically disordered amino- and carboxy-terminal extremities make them acquire specificity for HC structures. Alternatively, proteins present in the A10 or A11 gel filtration fractions may be required to confer HC specificity to PSPC1 homo-dimer and PSCP1-SFPQ hetero-dimer.

## Detection of the SND1-HC complex on a sucrose gradient

SND1 belongs to the group of proteins that was partially purified from the HeLa cell nuclear extracts as a HC-specific binding protein (Table 1). To confirm this interaction that was detected by electrophoresis on a polyacrylamide under native conditions, we investigated whether the SND1-HC complex could also be isolated on a sucrose gradient. A 5 to 40% sucrose gradient of 5 mL was poured and the mixture of the interaction between radiolabeled HC and the fraction A10 of the size exclusion chromatography was loaded on the gradient. After centrifugation, fractions of 200 μL were collected from the top to the bottom of the gradient and analyzed for their DNA and SND1 content. DNA content was estimated by radioactivity measurement and SND1 content by western blot. Results indicate that SND1 peaked at fraction 8 when the fraction A10 of the size exclusion chromatography was sedimented alone (white bars on the graph of Fig 3D) where as its peak was shifted to fraction 5 when the fraction A10 was sedimented with HC (hatched bars on the graph of Fig 3D). Analysis of the radioactivity in fractions 5 to 10 of the sucrose gradient indicated that sedimentation profile of HC was also modified by the proteins in the fraction A10 and that fraction 5 did contain radiolabeled HC (Fig 3E). Altogether, the sucrose gradient sedimentation supported the formation of a complex between SND1 and HC.

## Characterization of the SND1-HC nucleoprotein complex with purified proteins

SND1 has multiple protein- and nucleic acid- binding partners (for reviews see [34–36] and is involved in many aspects of gene expression (*e.g.* transcription [25–30,37], RNA splicing [38,39], RNA interference [22,40], microRNA processing and decay [41,42], RNA protection in stress granules [43,44], protein modification and degradation [45,46]. While fulfilling these functions, it may have either a scaffolding role or a catalytic (nuclease) activity. Its domain composition comprises at the amino terminus of the protein, four staphylococcal nuclease (SN)-like domains assembled in tandem (SN1 to SN4 on Fig 5A) and at the carboxy terminal end of the protein a fusion of a Tudor domain with a partial SN domain (SN5 on Fig 5A).

We purified two truncated versions of this protein, SND1-64 and SND1-110. SND1-64, from amino acids 315 to 863, lacks the first two SN domains and the 20 last amino acids and SND1-110, from amino acids 33 to 888, is truncated in the first SN domain since it lacks the 32 first amino acids (Fig 5A). We first characterized the interaction between SND1-64 and three DNA constructs (Fig 5B): the HC and the two double stranded DNA mini-circles dsMCs that were used to prepare the HC, dsMC09 and dsMC10. Results indicated that over the range of concentration tested (from 0 to 2.7 μM) SND1-64 did interact with HC (lanes 11 to 15 of Fig 5B) but not with the dsMCs (lanes 1 to 10 of Fig 5B). The same result was obtained with SDN1-110 although a single concentration of protein (70 nM) was tested owing to the low concentration of our purified SDN1-110

(Fig 5C). Quantification of the gels indicated that SND1-110 exhibited a stronger affinity to HC than SND1-64 (compare Figs 5E and 5G), possibly due to the presence of two additional SN domains at the amino terminal side of SND1-110.

Single stranded DNA might be exposed at the junction of the HC strands and be the substrate recognized by SND1 proteins. To check this hypothesis, we first characterized the interaction between HC and *E. coli* SSB, the single stranded DNA binding protein from *E. coli* (Fig 6). Results showed that *E. coli* SSB interacted with the HC (Fig 6A) but not with dsMC09 (Fig 6B), indicating that the HC molecule carried regions of single stranded DNA. Finally, to verify that SDN1 did not recognize primarily the single stranded DNA of the HC and but the HC junction itself, we performed two types of additional experiments. We first compared the binding of SND1 to HC and the binding of SND1 to the circular single stranded DNA coming from the nicking of dsMC10 (C10ss, Fig 1B). Results showed that SND1-64 did bind single stranded DNA but the binding was weaker than that measured with HC (Figs 5D-E). In the second type of experiment, we investigated whether a short (21 nucleotides) single stranded oligonucleotide devoid of secondary structure (OL21, see Materials and Methods for sequence) could compete with HC for the binding of SND1. To that end, we mixed the HC substrate (10 pM) with increasing amount of OL21 (from 0 to 7 nM), added SND1-64 to the mixture and, after 25 minutes of incubation at 22°C, resolved the species by electrophoresis on a polyacrylamide gel under native conditions. Results indicated that the (SND1-64)-HC complex was highly resistant to the presence of the oligonucleotide competitor, since all the HC was still bound by SND1-64 at a 700-fold molar excess of oligonucleotide OL21 (Fig 5H). This experiment confirmed that the single stranded DNA was not the only binding substrate recognized by SND1-64 on the HC. We performed the same kind of experiments with the SND1-110 protein. Results indicated that SND1-110 bound almost equally well HC and C10ss DNAs (Figs 5F-G) but that OL21 oligonucleotide still poorly competed with HC for binding to SND1-110, confirming a strong affinity of SND1-110 to HC (Fig 5I).

## Conclusion

In this work we identified SND1 as a protein exhibiting strong affinity and definite specificity for hemicatenated structures *in vitro*. It is known that SND1 protein interacts with various proteins of the transcription machinery, including the RNA polymerase II [30] and the transcription initiation factor TFIIE [37]. Considering the proposed role of HCs as potential structures affecting dynamically genome organization and transcription [20], SND1 could for example act by bridging hemicatenated structures and the basal transcription machinery. The dynamics of HC structures, whose location may change with time, cell types or upon external stresses, could therefore be modulated by specific binding of SND1. Given the essential role of the SND1 in transcription regulation and in oncogenic progression, our results invite to explore further the interaction between SND1 and HCs.

## Acknowledgments

The research was supported by core funding of the laboratory by INSERM, CNRS, and the National Museum of Natural History. The authors thank members of the Genome Structure and Instability laboratory for helpful discussions. ED is particularly grateful to Dr François Strauss for introducing her to the field of hemicatenanes and for the support and constructive discussions along the course of the project.

## Author contributions

ED conceived and performed experiments except the sucrose gradient and western blotting experiments that were performed by OR. CB and FG performed the identification of the HC-binding proteins by LC-MS/MS analysis at the proteomic platform 3P5 (Université de Paris, Institut Cochin, INSERM, U1016, CNRS, UMR8104, F-75014 Paris, France). ED, OR and JBB analyzed the data. ED and JBB wrote the manuscript. All authors provided detailed and fruitful comments.

## Conflict of interest

The authors declare that they have no conflict of interest.

## References

1 Lopes, M., Cotta-Ramusino, C., Liberi, G. and Foiani, M. (2003) Branch migrating sister chromatid junctions form at replication origins through Rad51/Rad52-independent mechanisms. Mol Cell **12**, 1499–1510.

2 Sogo, J. M., Stahl, H., Koller, T. and Knippers, R. (1986) Structure of replicating simian virus 40 minichromosomes. The replication fork, core histone segregation and terminal structures. J. Mol. Biol. **189**, 189–204.

3 Laurie, B., Katritch, V., Sogo, J., Koller, T., Dubochet, J. and Stasiak, A. (1998) Geometry and Physics of Catenanes Applied to the Study of DNA Replication. Biophys. J., Cell Press **74**, 2815–2822.

4 Lucas, I. and Hyrien, O. (2000) Hemicatenanes form upon inhibition of DNA replication. Nucleic Acids Res. **28**, 2187–93.

5 Fernández-Casañas, M. and Chan, K.-L. (2018) The Unresolved Problem of DNA Bridging. Genes (Basel). **9**, 623.

6 Bizard, A. H. and Hickson, I. D. (2018) Anaphase: a fortune-teller of genomic instability. Curr. Opin. Cell Biol. **52**, 112–119.

7 Liu, Y., Nielsen, C. F., Yao, Q. and Hickson, I. D. (2014) The origins and processing of ultra fine anaphase DNA bridges. Curr. Opin. Genet. Dev. **26**, 1–5.

8 Swuec, P. and Costa, A. (2014) Molecular mechanism of double Holliday junction dissolution. Cell Biosci. **4**, 36.

9 Bizard, A. H. and Hickson, I. D. (2014) The dissolution of double Holliday junctions. Cold Spring Harb. Perspect. Biol. **6**, a016477.

10 Cejka, P., Plank, J. L., Bachrati, C. Z., Hickson, I. D. and Kowalczykowski, S. C. (2010)

Rmi1 stimulates decatenation of double Holliday junctions during dissolution by Sgs1-Top3. Nat. Struct. Mol. Biol. **17**, 1377–82.
11 Plank, J. L., Wu, J. and Hsieh, T. S. (2006) Topoisomerase IIIalpha and Bloom's helicase can resolve a mobile double Holliday junction substrate through convergent branch migration. Proc. Natl. Acad. Sci. U. S. A. **103**, 11118–11123.
12 Raynard, S., Zhao, W., Bussen, W., Lu, L., Ding, Y.-Y., Busygina, V., Meetei, A. R. and Sung, P. (2008) Functional role of BLAP75 in BLM-topoisomerase IIIalpha-dependent holliday junction processing. J. Biol. Chem. **283**, 15701–8.
13 Wu, L., Bachrati, C. Z., Ou, J., Xu, C., Yin, J., Chang, M., Wang, W. and Li, L. (2006) BLAP/RMI1 promotes the BLM-dependent dissolution of homologous recombination intermediates. Proc Natl Acad Sci USA **103**, 4068–4073.
14 Kennedy, J. A., Daughdrill, G. W. and Schmidt, K. H. (2013) A transient α-helical molecular recognition element in the disordered N-terminus of the Sgs1 helicase is critical for chromosome stability and binding of Top3/Rmi1. Nucleic Acids Res. **41**, 10215–27.
15 Chan, K.-L., North, P. S. and Hickson, I. D. (2007) BLM is required for faithful chromosome segregation and its localization defines a class of ultrafine anaphase bridges. EMBO J. **26**, 3397–409.
16 Sarlós, K., Biebricher, A. S., Bizard, A. H., Bakx, J. A. M., Ferreté-Bonastre, A. G., Modesti, M., Paramasivam, M., Yao, Q., Peterman, E. J. G., Wuite, G. J. L., et al. (2018) Reconstitution of anaphase DNA bridge recognition and disjunction. Nat. Struct. Mol. Biol. **25**, 868–876.
17 Gaillard, C. and Strauss, F. (2000) DNA loops and semicatenated DNA junctions. BMC Biol. **1**, 1.
18 Gaillard, C. and Strauss, F. (2000) High affinity binding of proteins HMG1 and HMG2 to semicatenated DNA loops. BMC Mol. Biol. **1**, 1.
19 Jaouen, S., de Koning, L., Gaillard, C., Muselíková-Polanská, E., Stros, M. and Strauss, F. (2005) Determinants of specific binding of HMGB1 protein to hemicatenated DNA loops. J. Mol. Biol. **353**, 822–37.
20 Gaillard, C. and Strauss, F. (2006) DNA topology and genome organization in higher eukaryotes: a model. J. Theor. Biol. **243**, 604–7.
21 Gaillard, C. and Strauss, F. (2015) Construction of DNA Hemicatenanes from Two Small Circular DNA Molecules. PLoS One.
22 Li, C.-L., Yang, W.-Z., Chen, Y.-P. and Yuan, H. S. (2008) Structural and functional insights into human Tudor-SN, a key component linking RNA interference and editing. Nucleic Acids Res. **36**, 3579–89.
23 Passon, D. M. (2012) Dimerization and RNA binding of paraspeckle proteins.
24 Huang, J., Casas Garcia, G. P., Perugini, M. A., Fox, A. H., Bond, C. S. and Lee, M. (2018) Crystal structure of a SFPQ/PSPC1 heterodimer provides insights into preferential heterodimerization of human DBHS family proteins. J. Biol. Chem. **293**, 6593–6602.
25 Callebaut, I. and Mornon, J. P. (1997) The human EBNA-2 coactivator p100: multidomain organization and relationship to the staphylococcal nuclease fold and to the tudor protein involved in *Drosophila melanogaster* development. Biochem. J. **321**, 125–132.
26 Duan, Z., Zhao, X., Fu, X., Su, C., Xin, L., Saarikettu, J., Yang, X., Yao, Z., Silvennoinen, O., Wei, M., et al. (2014) Tudor-SN, a Novel Coactivator of Peroxisome Proliferator-activated Receptor γ Protein, Is Essential for Adipogenesis. J. Biol. Chem. **289**, 8364–8374.

27 Leverson, J. D., Koskinen, P. J., Orrico, F. C., Rainio, E. M., Jalkanen, K. J., Dash, A. B., Eisenman, R. N. and Ness, S. A. (1998) Pim-1 kinase and p100 cooperate to enhance c-Myb activity. Mol. Cell **2**, 417–25.
28 Paukku, K., Yang, J. and Silvennoinen, O. (2003) Tudor and nuclease-like domains containing protein p100 function as coactivators for signal transducer and activator of transcription 5. Mol. Endocrinol. **17**, 1805–14.
29 Su, C., Zhang, C., Tecle, A., Fu, X., He, J., Song, J., Zhang, W., Sun, X., Ren, Y., Silvennoinen, O., et al. (2015) Tudor Staphylococcal Nuclease (Tudor-SN), a Novel Regulator Facilitating $G_1$/S Phase Transition, Acting as a Co-activator of E2F-1 in Cell Cycle Regulation. J. Biol. Chem. **290**, 7208–7220.
30 Yang, J., Aittomäki, S., Pesu, M., Carter, K., Saarinen, J., Kalkkinen, N., Kieff, E. and Silvennoinen, O. (2002) Identification of p100 as a coactivator for STAT6 that bridges STAT6 with RNA polymerase II. EMBO J. **21**, 4950–8.
31 Knott, G. J., Bond, C. S. and Fox, A. H. (2016) The DBHS proteins SFPQ, NONO and PSPC1: a multipurpose molecular scaffold. Nucleic Acids Res. **44**, 3989–4004.
32 Guillaumond, F., Boyer, B., Becquet, D., Guillen, S., Kuhn, L., Garin, J., Belghazi, M., Bosler, O., Franc, J.-L. and François-Bellan, A.-M. (2011) Chromatin remodeling as a mechanism for circadian prolactin transcription: rhythmic NONO and SFPQ recruitment to HLTF. FASEB J. **25**, 2740–2756.
33 Salton, M., Lerenthal, Y., Wang, S.-Y., Chen, D. J. and Shiloh, Y. (2010) Involvement of Matrin 3 and SFPQ/NONO in the DNA damage response. Cell Cycle **9**, 1568–76.
34 Jariwala, N., Rajasekaran, D., Srivastava, J., Gredler, R., Akiel, M. A., Roberston, C. L., Emdad, L., Fisher, P. B. and Sarkar, D. (2015) Role of the staphylococcal nuclease and tudor domain containing 1 in oncogenesis (Review). Int. J. Oncol. **46**, 465–473.
35 Chidambaranathan-Reghupaty, S., Mendoza, R., Fisher, P. B. and Sarkar, D. (2018) The multifaceted oncogene SND1 in cancer: focus on hepatocellular carcinoma. Hepatoma Res. **4**, 32.
36 Gutierrez-Beltran, E., Denisenko, T. V, Zhivotovsky, B. and Bozhkov, P. V. (2016) Tudor staphylococcal nuclease: biochemistry and functions. Cell Death Differ. **23**, 1739–1748.
37 Tong, X., Drapkin, R., Yalamanchili, R., Mosialos, G. and Kieff, E. (1995) The Epstein-Barr virus nuclear protein 2 acidic domain forms a complex with a novel cellular coactivator that can interact with TFIIE. Mol. Cell. Biol. **15**, 4735–44.
38 Yang, J., Välineva, T., Hong, J., Bu, T., Yao, Z., Jensen, O. N., Frilander, M. J. and Silvennoinen, O. (2007) Transcriptional co-activator protein p100 interacts with snRNP proteins and facilitates the assembly of the spliceosome. Nucleic Acids Res. **35**, 4485–94.
39 Gao, X., Zhao, X., Zhu, Y., He, J., Shao, J., Su, C., Zhang, Y., Zhang, W., Saarikettu, J., Silvennoinen, O., et al. (2012) Tudor staphylococcal nuclease (Tudor-SN) participates in small ribonucleoprotein (snRNP) assembly via interacting with symmetrically dimethylated Sm proteins. J. Biol. Chem. **287**, 18130–41.
40 Caudy, A. A., Ketting, R. F., Hammond, S. M., Denli, A. M., Bathoorn, A. M. P., Tops, B. B. J., Silva, J. M., Myers, M. M., Hannon, G. J. and Plasterk, R. H. A. (2003) A micrococcal nuclease homologue in RNAi effector complexes. Nature **425**, 411–4.
41 Li, C.-L., Yang, W.-Z., Shi, Z. and Yuan, H. S. (2018) Tudor staphylococcal nuclease is a structure-specific ribonuclease that degrades RNA at unstructured regions during microRNA decay. RNA **24**, 739–748.
42 Elbarbary, R. A., Miyoshi, K., Hedaya, O., Myers, J. R. and Maquat, L. E. (2017) UPF1 helicase promotes TSN-mediated miRNA decay. Genes Dev. **31**, 1483–1493.

43 Gao, X., Ge, L., Shao, J., Su, C., Zhao, H., Saarikettu, J., Yao, X., Yao, Z., Silvennoinen, O. and Yang, J. (2010) Tudor-SN interacts with and co-localizes with G3BP in stress granules under stress conditions. FEBS Lett. **584**, 3525–3532.
44 Weissbach, R. and Scadden, A. D. J. (2012) Tudor-SN and ADAR1 are components of cytoplasmic stress granules. RNA **18**, 462–71.
45 Rajasekaran, D., Jariwala, N., Mendoza, R. G., Robertson, C. L., Akiel, M. A., Dozmorov, M., Fisher, P. B. and Sarkar, D. (2016) Staphylococcal Nuclease and Tudor Domain Containing 1 (SND1 Protein) Promotes Hepatocarcinogenesis by Inhibiting Monoglyceride Lipase (MGLL). J. Biol. Chem., American Society for Biochemistry and Molecular Biology **291**, 10736–46.
46 Yu, L., Liu, X., Cui, K., Di, Y., Xin, L., Sun, X., Zhang, W., Yang, X., Wei, M., Yao, Z., et al. (2015) SND1 Acts Downstream of TGFβ1 and Upstream of Smurf1 to Promote Breast Cancer Metastasis. Cancer Res., American Association for Cancer Research **75**, 1275–1286.

## Figure legends

### Figure 1.

A Schematic representation of a DNA hemicatenane. In this scheme, the two double stranded DNAs that are topologically linked by a hemicatenane are not homologous and therefore are represented by two different colors: blue and red. The light and dark blue strands are complementary, as the light and dark red strands.

B Scheme of construction of DNA HCs from two dsMCs. Radiolabeled dsMCs09 and dsMCs10 were prepared from $^{32}$P -end labeled linear double stranded fragments of 215 (colored in red) and 235 (colored in blue) base pairs. Their nicking by specific enzymes led to two linear single stranded fragments (L09ss in blue and L10ss in red) and two single stranded circles (C09ss in blue and C10ss in red). Single stranded catenanes were assembled from the circularization of L09ss and L10ss under specific conditions: the linker oligonucleotide complementary to 25 nucleotides of L09ss and 25 nucleotides of L10ss brings together the two strands and favors the catenation. The single stranded catenanes were gel-purified and re-annealed with the C09ss and C10ss using the wheat germ Topoisomerase I to finally get the DNA HCs. Adapted from [21].

### Figure 2. Fractionation of HeLa nuclear extracts.

A Scheme of the procedure of fractionation of HeLa nuclear extracts. The procedure starts with an ammonium sulfate precipitation that is followed by four chromatographies on different media, as indicated.

B Proteins contained in 1 μL of fraction A10 of the size exclusion chromatography were incubated with radiolabeled dsMC09 (lanes 1-4, 9) or HC (lanes 5-8, 10) in the presence of increasing amount of λ DNA (0 (lanes 4 and 8), 5 (lanes 3 and 7), 10 (lanes 2 and 6), 30 (lanes 1 and 5) ng). After incubation, species were resolved by electrophoresis on a native polyacrylamide gel. Free DNA (dsMC09 or HC) and Protein-HC complexes are indicated.

C Analysis of the protein fractions obtained during the fractionation of HeLa nuclear extracts. Fractions were analyzed by electrophoresis on polyacrylamide gels under denaturing (SDS) conditions. Lanes 1, 3 and 6: molecular weight (MW); the size of the proteins is given in kDa; Lane 2: crude extract (CE); Lane 4: aliquot of the sample loaded on the phosphocellulose column (PC-load). Lane 5: aliquot of the sample loaded on the ButylS-FF column (BFF-load). Lane 7: aliquot of the sample loaded on the size exclusion S200 column (GF-load). Lanes 8 and 10: fraction A10 and 14 of the size exclusion chromatography. 10 μL of CE, PC-load, BFF-load and GF-load were loaded on the gel whereas 30 μL of fractions A10 and A14 were loaded.

**Figure 3. Identification of HC-specific binding proteins in the fractions A10 and A11 of the size exclusion chromatography.**

A Products of the interaction between radiolabeled HC (16 femtomoles) and proteins of the fraction A10 of the size exclusion chromatography (45 μL) were resolved by electrophoresis on a polyacrylamide gel (lanes 3, 5, 8, 10). 12 (lanes 2 and 3) or 23 (lanes 4 and 5) ng of λ DNA were added to reduce non specific binding. Sample without protein (lane 1) or without HC (lanes 2 and 4) were also loaded to serve as control for mass spectrometry analysis. Free DNA (dsMC or HC) and Protein-HC complexes are indicated. The gel pieces that were analyzed by mass spectrometry are shown as green rectangles and the name of the gel pieces of interest is indicated on the right side of the figure.

B Same as in (A) but with fraction A11 of the size exclusion chromatography. No λ DNA was added in the reaction mixes. Lane 1: HC without fraction A11; lane 2: fraction A11 without HC; lane 3: HC + fraction A11. On the right panel, the green rectangles represent the two gel pieces that were analyzed by mass spectrometry. The name of the gel piece of interest is indicated on the right side of the figure.

C Fractions A10 and A11 were tested for their content of PSPC1, SND1, SFPQ and PTBP1 proteins by western blot. Lanes 1, 4, 7, 10: MW; the size of the proteins is given in kDa; lanes 2, 5, 8, 11: fraction A10; lanes 3, 6, 9, 12: fraction A11. The identity of the antibodies used to probe the fraction is indicated above each panel. The arrow points to the protein of interest.

D and E Three samples were prepared and centrifuged on a sucrose gradient: sample with HC, sample with fraction A10 and sample with fraction A10 mixed with HC. After centrifugation, 200 μL fractions were collected from the top to the bottom of the gradient. Number of the fraction increases from top to bottom. In (D), fractions 5 to 10 of the sucrose gradients were tested for their SND1 content by western blot (band indicated by an arrow on the right side of the panel). Lanes 1, 3, 5, 7, 9, 11: the fraction A10 was loaded on the sucrose gradient. Lanes 2, 4, 6, 8, 10, 12: the mixture (fraction A10 + HC) was loaded on the sucrose gradient. Bands on the membrane were quantified using Image J software and the relative intensity of each band is plotted as a function of fraction. The white bars correspond to the fraction A10 sample and the hatched bars to the (fraction A10 + HC) sample. In (E), fractions 5 to 10 of the sucrose gradients were tested for their radioactivity content. An aliquot of each fraction (0.5 μL) was spotted on a nitrocellulose membrane. When dried, the membrane was exposed on a $^{32}$P-sensitive screen. After exposure, the screen was scanned. The radioactivity profile of the HC sample is compared with that of the (fraction A10 + HC) sample.

**Figure 4. Interaction between purified PSPC1 or PSPC1-SFPQ with DNA.**

A The organization of SFPQ and PSPC1 in domains is shown. The double arrow underneath each schematic representation corresponds to the truncated protein expressed in *E. coli* and purified.

B to E Molecular interactions were performed in a final volume of 7.75 μL with 0.1 femtomole of radiolabeled DNA and the indicated amount of purified protein. Species were resolved by electrophoresis on a polyacrylamide native gel. Free DNAs and bound DNAs are indicated.

B The DNA used is HC and the tested protein is PSPC1 homo-dimer. Lane 1: no PSPC1; lane 2: 300 nM PSPC1; lane 3: 1 μM PSPC1; lane 4: 3 μM PSPC1; lane 5: 9 μM PSPC1.

C The DNA used is C10ss and the tested protein is PSPC1. Lane 1: no PSPC1; lane 2: 300 nM PSPC1; lane 3: 1 μM PSPC1; lane 4: 3 μM PSPC1; lane 5: 9 μM PSPC1.

D The DNA used is HC and the tested protein is PSPC1-SFPQ hetero-dimer. Lane 1: no PSPC1-SFPQ; lane 2: 500 nM PSPC1-SFPQ; lane 3: 1.5 μM PSPC1-SFPQ; lane 4: 4.5 μM PSPC1-SFPQ; lane 5: 13 μM PSPC1-SFPQ.

E The DNA used is C10ss and the tested protein is PSPC1-SFPQ hetero-dimer. Lane 1: no PSPC1-SFPQ; lane 2: 500 nM PSPC1-SFPQ; lane 3: 1.5 μM PSPC1-SFPQ; lane 4: 4.5 μM PSPC1-SFPQ; lane 5: 13 μM PSPC1-SFPQ.

**Figure 5. Interaction between purified SND1 proteins and various DNA constructs.**

A The organization of SND1 in domains is shown. The double arrow underneath the schematic representation corresponds to the proteins that were expressed in *E. coli* and purified. The name of the purified protein is indicated on the left side of the double arrow.

B Interactions were performed in a final volume of 7.5 μL with 0.1 femtomole of radiolabeled DNA and the indicated amount of purified protein. Species were resolved by electrophoresis under native conditions. Three DNAs were tested for their binding to SND1-64: dsMC09 (lanes 1-5); dsMC10 (lane 6-10); HC (lanes 11-15). Concentrations of protein were as indicated (lanes 1, 6, 11: 0; lanes 2, 7, 12: 0.1 μM; lanes 3, 8, 13: 0.3 μM; lanes 4, 9, 14: 0.9 μM; lanes 5, 10, 15: 2.7 μM). Free DNAs and bound DNAs are indicated.

C Interactions were performed in a final volume of 13.25 μL with 0.1 femtomole of radiolabeled DNA and the SDN1-110 at 70 nM. Species were resolved by electrophoresis under native conditions. Three DNAs were tested for their binding to SND1-110: dsMC09 (lanes 1 and 2); dsMC10 (lanes 3 and 4); HC (lanes 5 and 6). Free DNAs and bound DNAs are indicated.

D Interactions were as described in (A). The DNA was C10ss, the single stranded circle obtained after nicking of the dsMC10 with Nt.BbvcI. 0.1 femtomole of C10ss was included in the reaction mixture and the concentrations of protein were as indicated

(lane 1: 0; lane 2: 0.1 μM; lane 3: 0.3 μM; lane 4: 0.9 μM; lane 5: 2.7 μM). Free DNAs and bound DNAs are indicated.

E The two curves (one for HC and one for C10ss) show the percentage of (SDN1-64)-DNA complexes as a function of protein concentration. Error bars correspond to the standard deviation. Percentages are the mean of three independent experiments.

F Interactions were as described in (C). C10ss is the single stranded circle obtained after nicking of the dsMC10 with Nt.BbvcI. 0.1 femtomole of C10ss or HC was included in the reaction mixture and the concentration of protein was 70 nM. Free DNAs and bound DNAs are indicated.

G The plot shows the percentage of (SDN1-110)-DNA complexes assembled at 70 nM SND1-110. Error bars correspond to the standard deviation. Percentages are the mean of three independent experiments.

H The interaction between radiolabeled HC and SND1-64 was tested in the presence of OL21, an oligonucleotide long of 21 nucleotides. SDN1-64 was at 2.7 μM and HC at 14 pM. HC was premixed with increasing amount of OL21 (lane 1: no OL21, no SND1-64; lane 2: no OL21; lane 3: 0.2 nM OL21; lane 4: 0.7 nM OL21; lane 5: 2 nM OL21; lane 6: 7 nM OL21) before adding SND1-64. Free DNAs and bound DNAs are indicated.

I The interaction between radiolabeled HC and SND1-110 was tested in the presence of OL21, an oligonucleotide of length of 21 nucleotides. SDN1-110 was at 70 nM and HC at 7.5 pM. HC was premixed with increasing amount of OL21 (lane 1: no OL21, no SND1-110; lane 2: no OL21; lane 3: 0.2 nM OL21; lane 4: 0.7 nM OL21; lane 5: 2 nM OL21; lane 6: 7 nM OL21) before adding SND1-110. Species are separated by electrophoresis on a polyacrylamide gel. Free DNAs and bound DNAs are indicated. The plot shows the percentage of (SND1-110)-HC complexes as function of concentration of OL21. The standard deviation is calculated from two independent experiments.

**Figure 6. Interaction between *E. coli SSB* and DNA.**

Interactions were performed in a final volume of 7.75 μL with 0.1 femtomole of radiolabeled DNA and the indicated amount of purified protein. Species were resolved by electrophoresis under native conditions. Free DNAs and bound DNAs are indicated.

A The DNA used was HC. Lane 1: no *E. coli* SSB; lane 2: 20 nM *E. coli* SSB; lane 3: 200 nM *E. coli* SSB; lane 4: 2 μM *E. coli* SSB; lane 5: 20 μM *E. coli* SSB.

B The DNA used is dsMC09. Lane 1: 2 μM *E. coli* SSB; lane 2: 0.4 μM *E. coli* SSB; Lane 3: no *E. coli* SSB

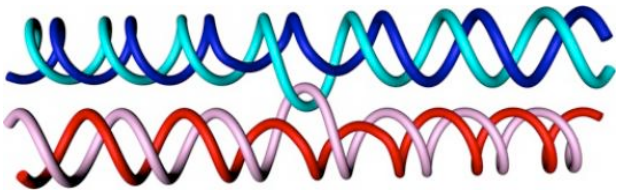
## Table

**Table 1. Names and parameters of identification of HC-binding proteins by mass spectrometry.**

| Name of the protein | Name of the gel piece | Mascot score | Number of identified peptide | Sequence coverage |
|---|---|---|---|---|
| Annexin A2 | 10p2-β | 64 | 2 | 5 |
| PSPC1 (ParaSpeckle Protein Component 1) | 10p2-γ | 78 | 3 | 5.9 |
| | 10p1-β | 74 | 3 | 5 |
| | 11p-β | 91 | 2 | 3.6 |
| PTBP1 (Polypyrimidine Tract-Binding Protein 1) | 11p-β | 77 | 3 | 5.1 |
| SFPQ (Splicing Factor Proline Glutamine-rich) | 10p1-β | 97 | 3 | 4.8 |
| SND1 (Staphylococcal Nuclease Domain-like protein 1) | 10p1-β | 98 | 2 | 2.6 |

Only proteins specifically identified by mass spectrometry in the indicated gel pieces are listed. For the all the proteins identified by mass spectrometry in each gel piece designated in the Figs 3A and 3B, see supplementary Table 1.

**A**



**B**

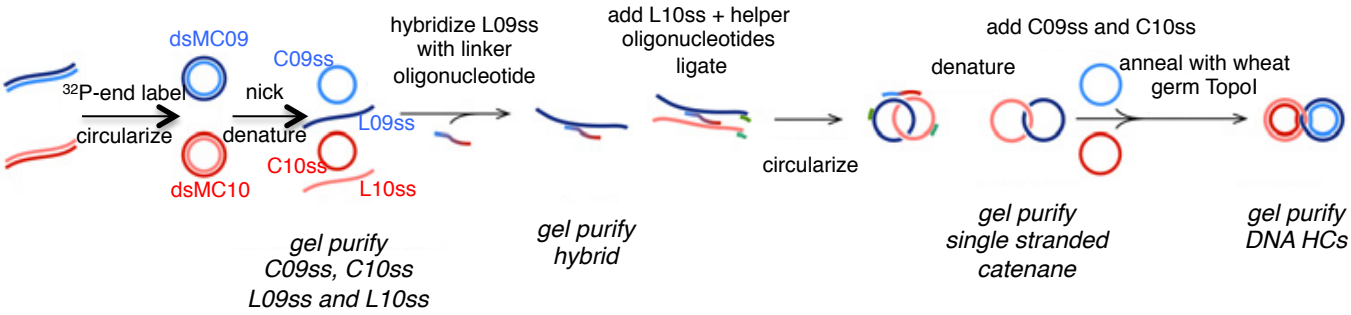

dsMC09
C09ss
hybridize L09ss
with linker
oligonucleotide
add L10ss + helper
oligonucleotides
ligate
add C09ss and C10ss
denature
anneal with wheat
germ TopoI
32P-end label
circularize
nick
denature
L09ss
C10ss
dsMC10
L10ss
circularize
gel purify
C09ss, C10ss
L09ss and L10ss
gel purify
hybrid
gel purify
single stranded
catenane
gel purify
DNA HCs

**Figure 1**

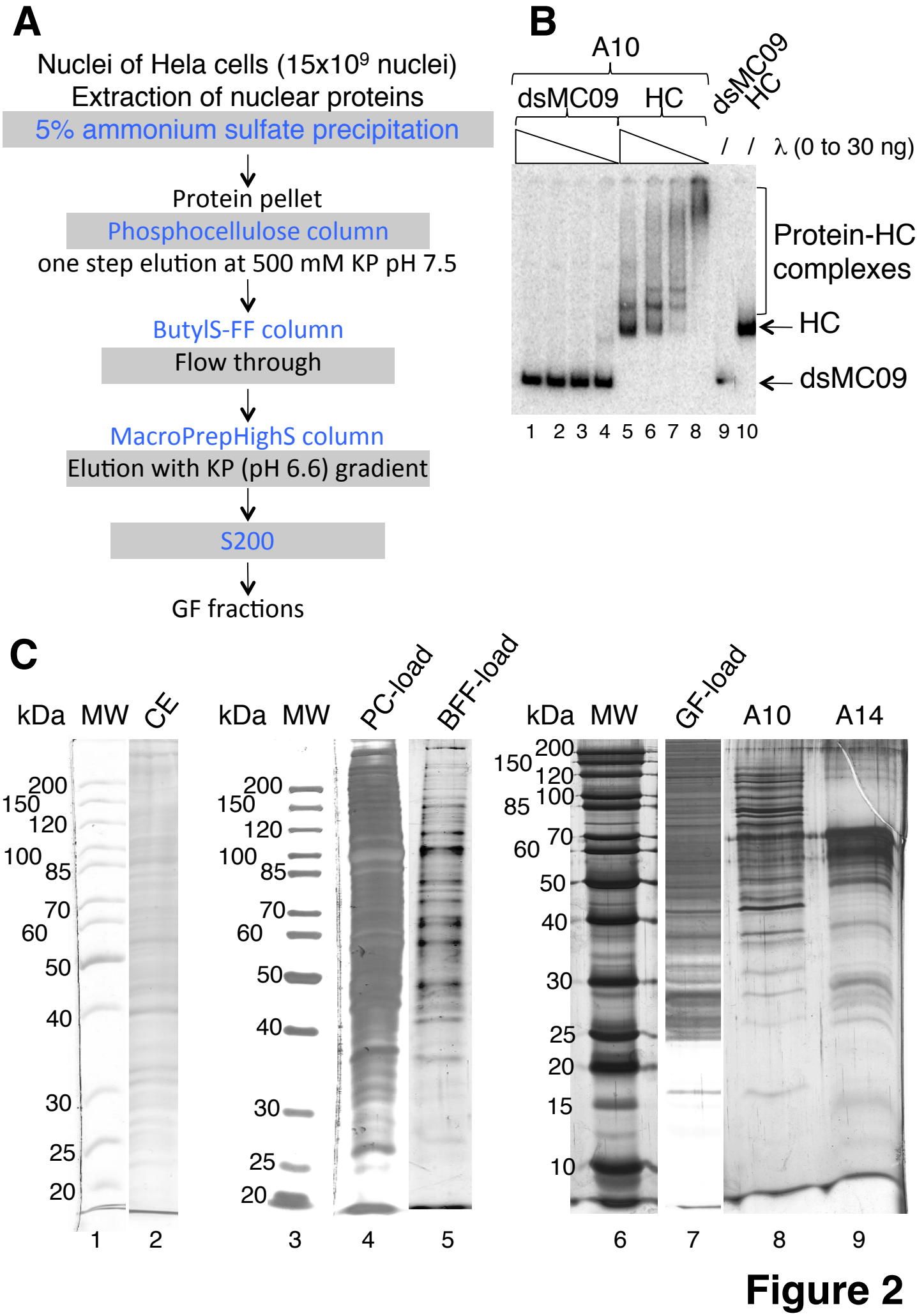
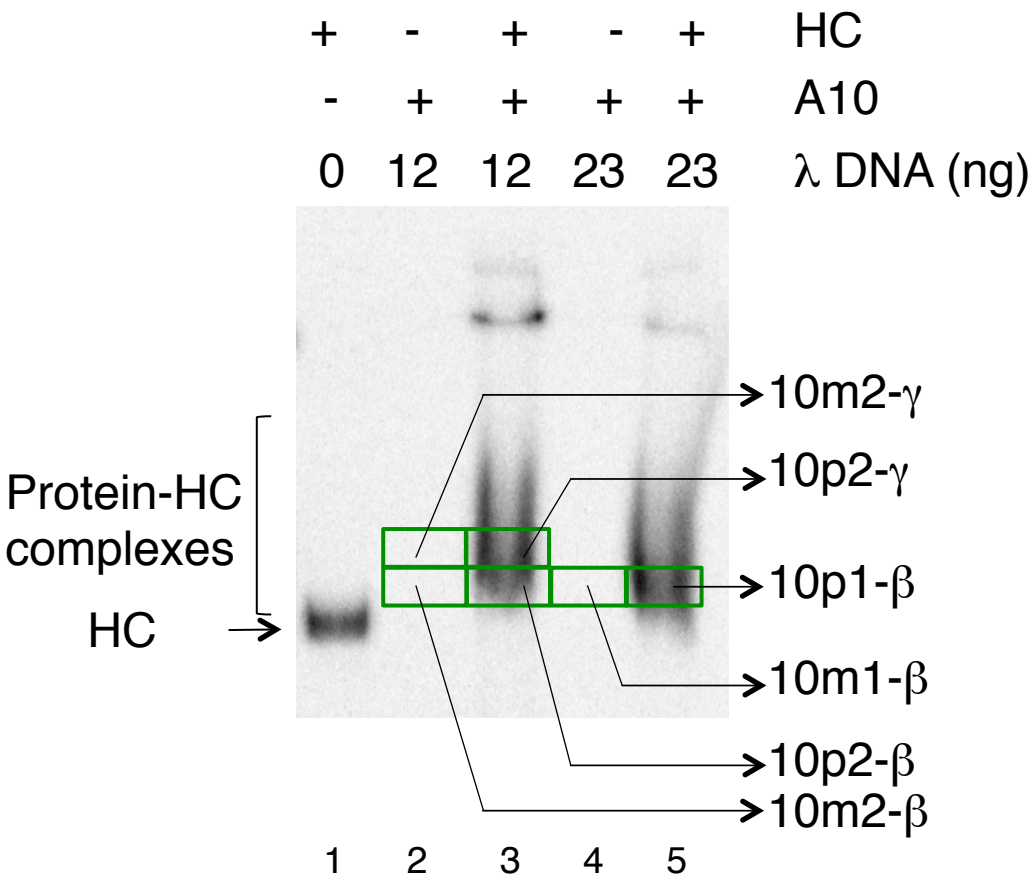

A
Nuclei of Hela cells (15x10⁹ nuclei)
Extraction of nuclear proteins
5% ammonium sulfate precipitation
Protein pellet
Phosphocellulose column
one step elution at 500 mM KP pH 7.5
ButylS-FF column
Flow through
MacroPrepHighS column
Elution with KP (pH 6.6) gradient
S200
GF fractions
B
A10
dsMC09
HC
dsMC09 HC
/ / λ (0 to 30 ng)
Protein-HC complexes
HC
dsMC09
1 2 3 4 5 6 7 8 9 10
C
kDa MW CE
kDa MW PC-load BFF-load
kDa MW GF-load A10 A14
200
150
120
100
85
70
60
50
40
30
25
20
15
10
1 2 3 4 5 6 7 8 9

**Figure 2**

**A**


+ - + - + HC
- + + + + A10
0 12 12 23 23 λ DNA (ng)
10m2-γ
10p2-γ
10p1-β
10m1-β
10p2-β
10m2-β
Protein-HC complexes
HC
1 2 3 4 5

**B**

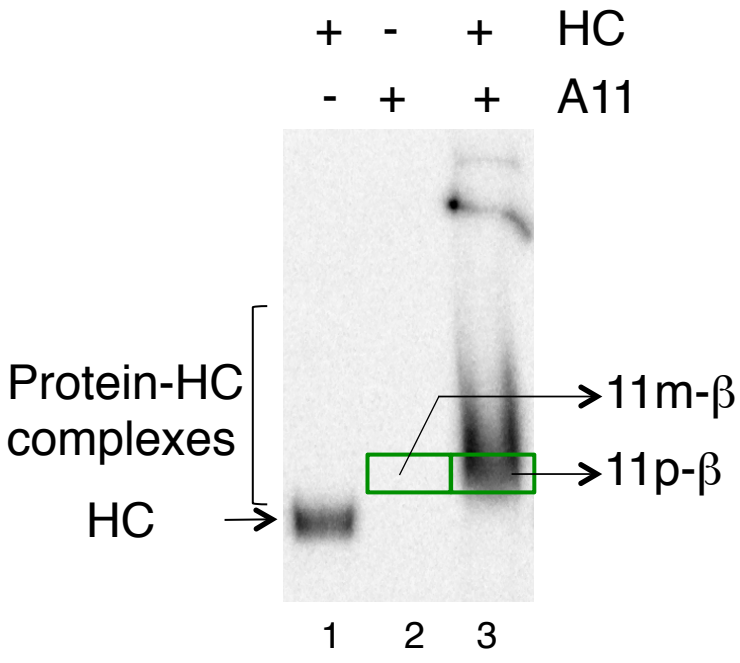
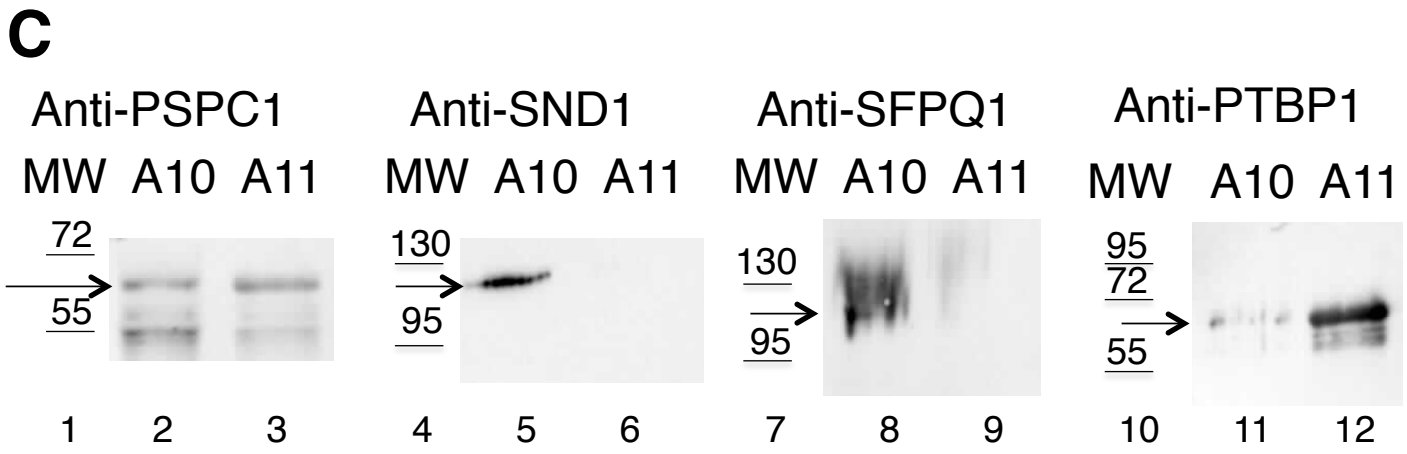
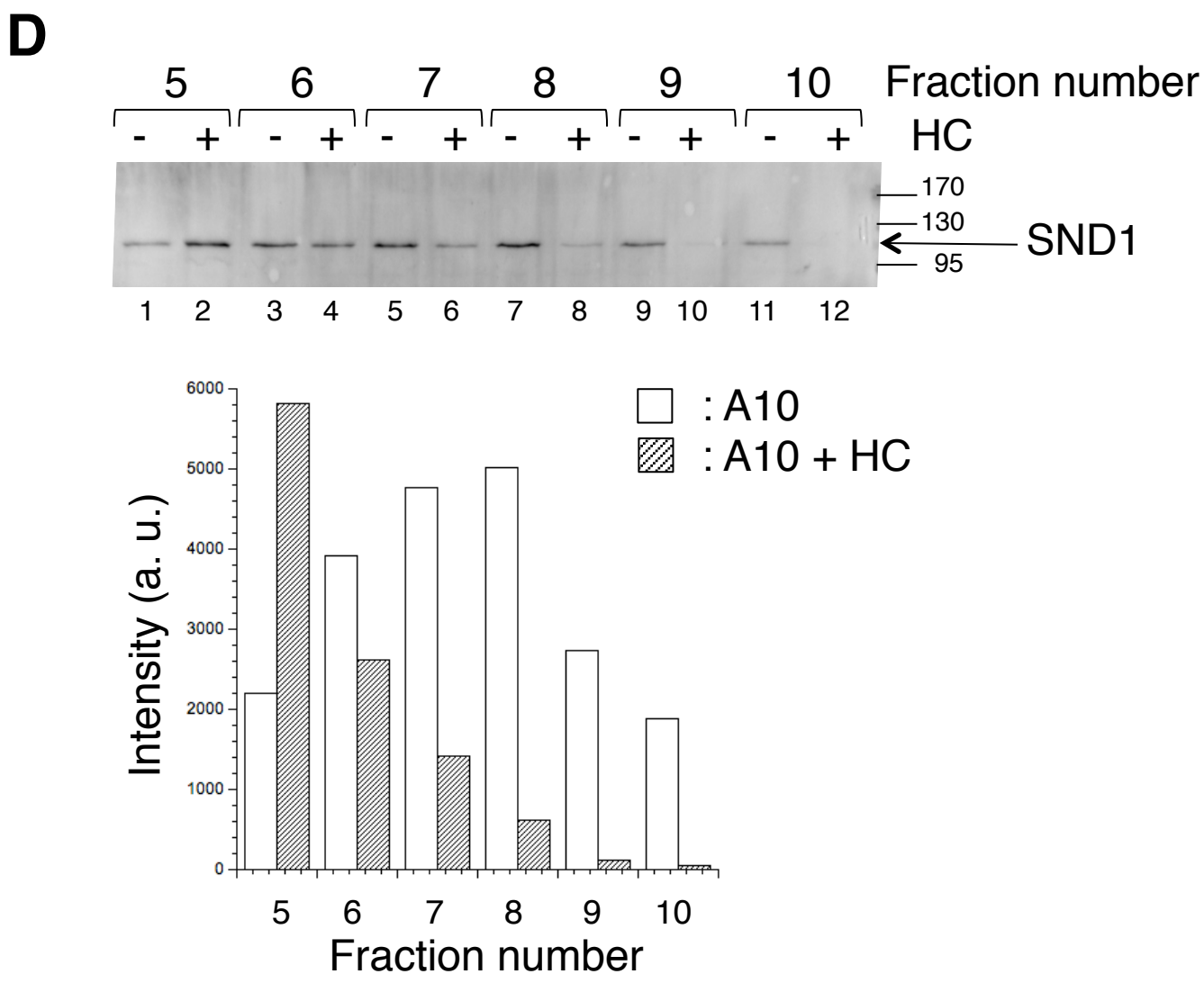

+ - + HC
- + + A11
11m-β
11p-β
Protein-HC complexes
HC
1 2 3

**Figure 3**


C
Anti-PSPC1
MW A10 A11
72
55
Anti-SND1
MW A10 A11
130
95
Anti-SFPQ1
MW A10 A11
130
95
Anti-PTBP1
MW A10 A11
95
72
55
1 2 3 4 5 6 7 8 9 10 11 12


D
5 6 7 8 9 10 Fraction number
- + - + - + - + - + - +
HC
170
130
95
SND1
1 2 3 4 5 6 7 8 9 10 11 12
: A10
: A10 + HC
Intensity (a. u.)
6000
5000
4000
3000
2000
1000
0
5 6 7 8 9 10
Fraction number

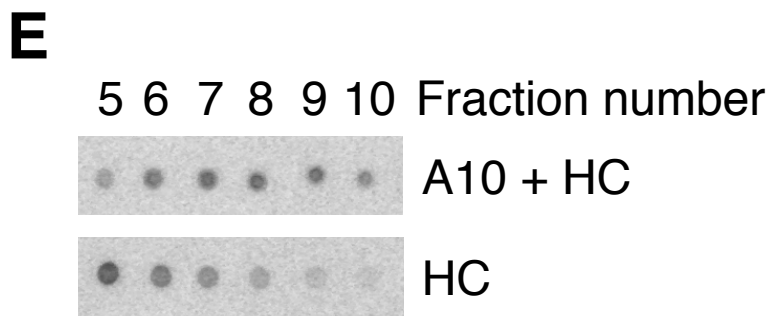

E
5 6 7 8 9 10 Fraction number
A10 + HC
HC

**Figure 3**

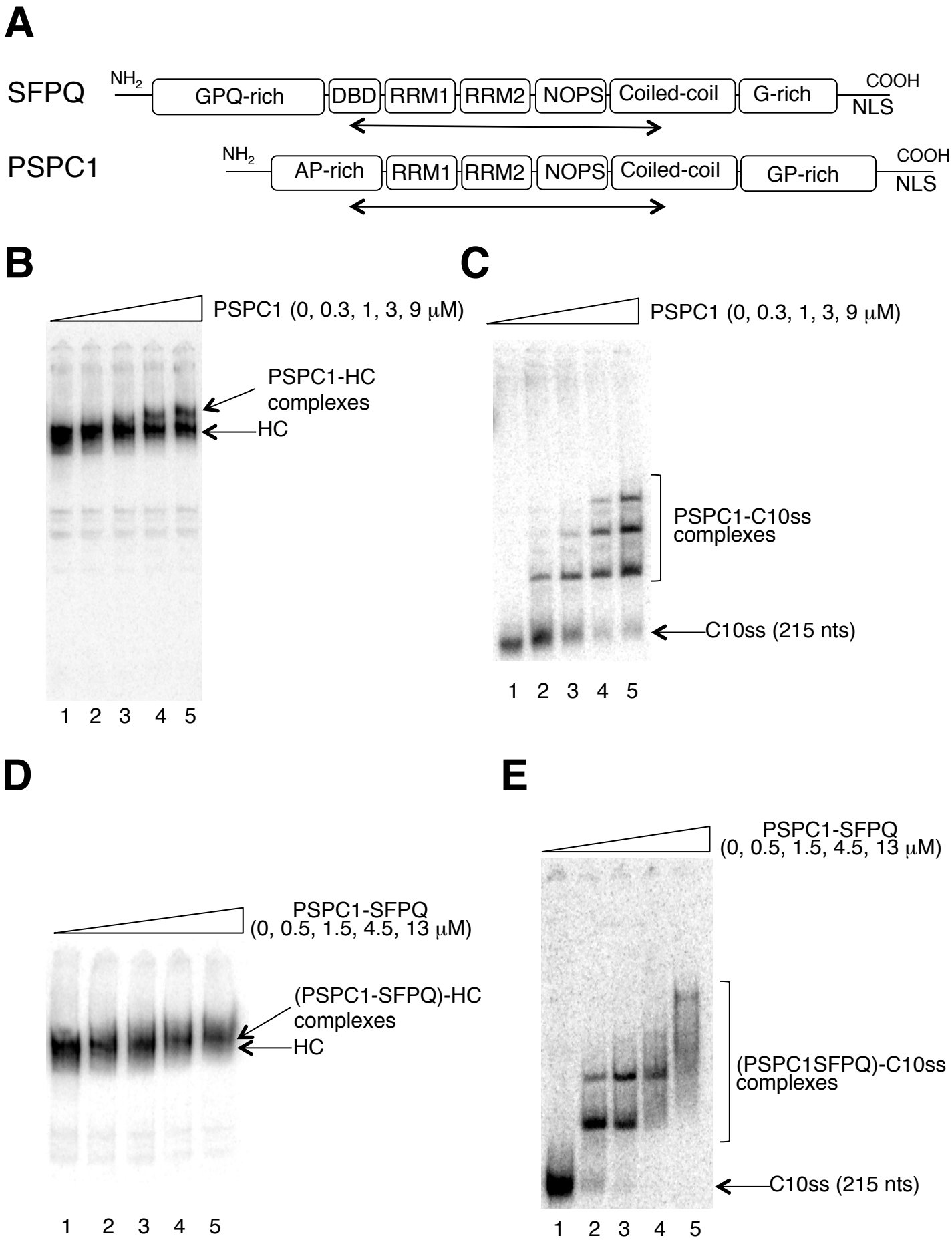

A
SFPQ
NH2
GPQ-rich
DBD
RRM1
RRM2
NOPS
Coiled-coil
G-rich
COOH
NLS
PSPC1
NH2
AP-rich
RRM1
RRM2
NOPS
Coiled-coil
GP-rich
COOH
NLS
B
PSPC1 (0, 0.3, 1, 3, 9 μM)
PSPC1-HC complexes
HC
1 2 3 4 5
C
PSPC1 (0, 0.3, 1, 3, 9 μM)
PSPC1-C10ss complexes
C10ss (215 nts)
1 2 3 4 5
D
PSPC1-SFPQ (0, 0.5, 1.5, 4.5, 13 μM)
(PSPC1-SFPQ)-HC complexes
HC
1 2 3 4 5
E
PSPC1-SFPQ (0, 0.5, 1.5, 4.5, 13 μM)
(PSPC1SFPQ)-C10ss complexes
C10ss (215 nts)
1 2 3 4 5

**Figure 4**

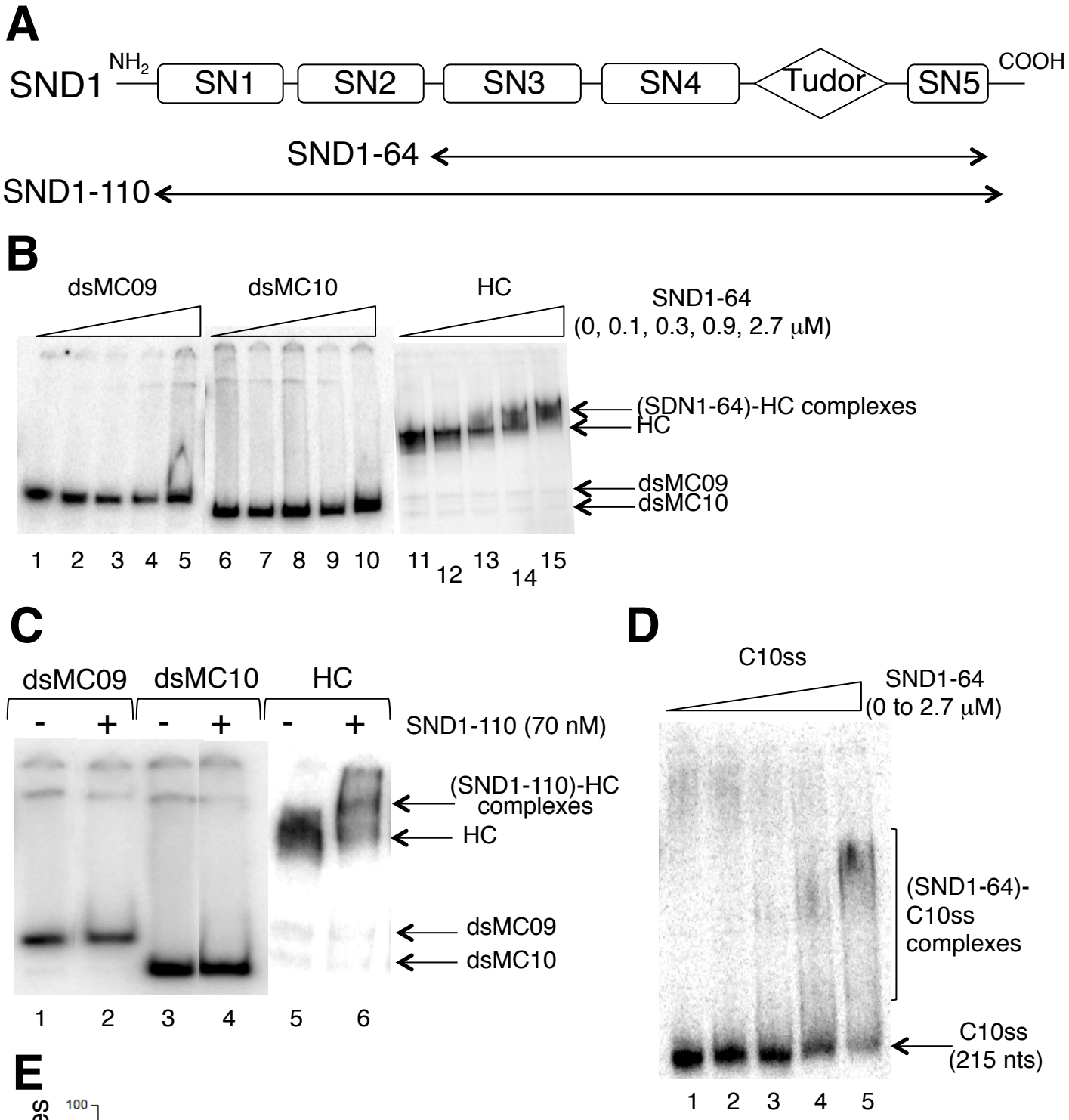

A
SND1
NH2
SN1
SN2
SN3
SN4
Tudor
SN5
COOH
SND1-64
SND1-110
B
dsMC09
dsMC10
HC
SND1-64
(0, 0.1, 0.3, 0.9, 2.7 μM)
(SDN1-64)-HC complexes
HC
dsMC09
dsMC10
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
C
dsMC09
dsMC10
HC
- + - + - +
SND1-110 (70 nM)
(SND1-110)-HC complexes
HC
dsMC09
dsMC10
1 2 3 4 5 6
D
C10ss
SND1-64
(0 to 2.7 μM)
(SND1-64)-C10ss complexes
C10ss
(215 nts)
1 2 3 4 5
E
% (SDN1-64)-DNA complexes
100
80
60
40
20
0
0 0,5 1 1,5 2 2,5 3
[SND1-64] in μM
HC
C10ss

**Figure 5**

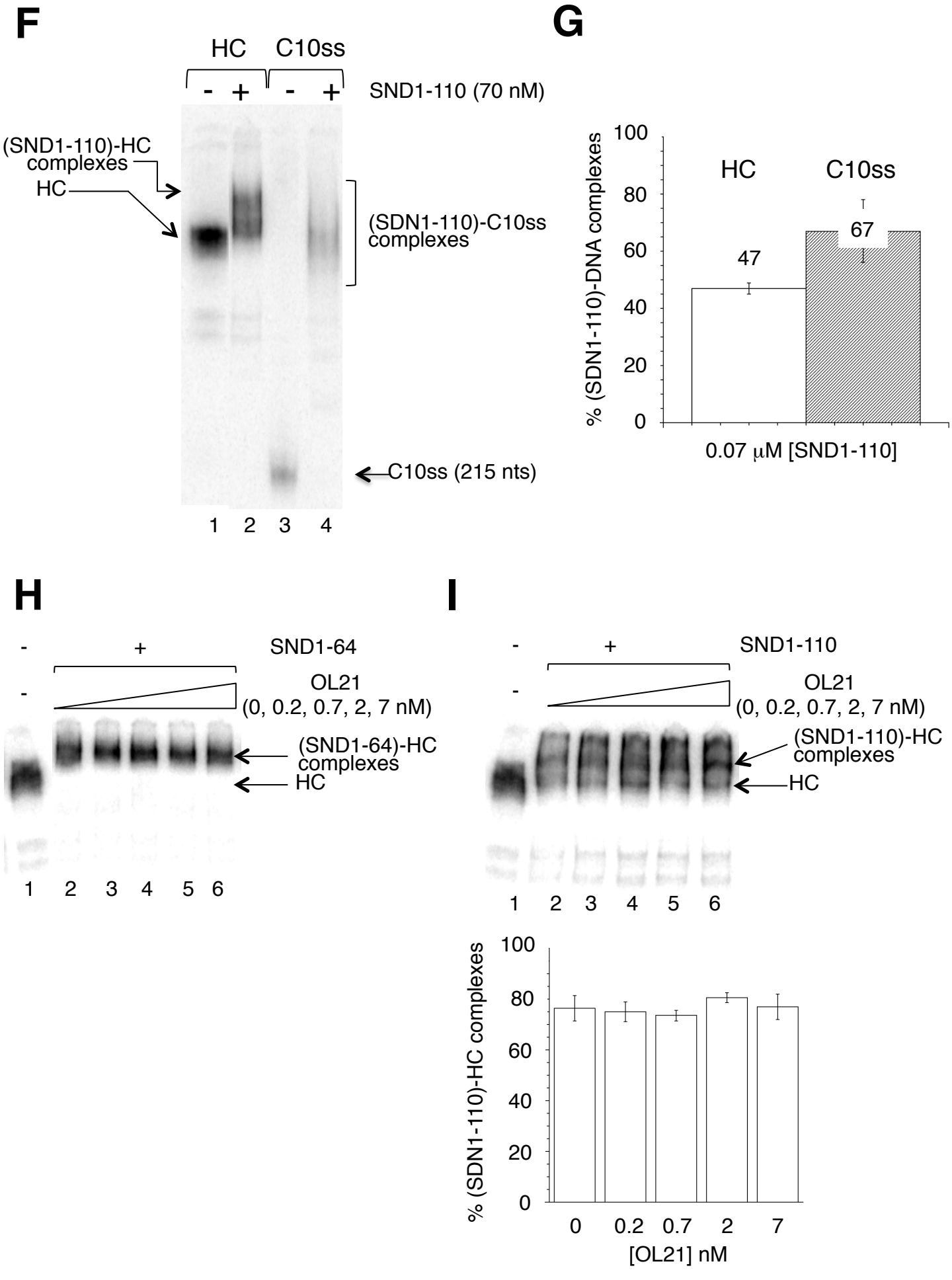

F
HC
C10ss
- + - +
SND1-110 (70 nM)
(SND1-110)-HC complexes
HC
(SDN1-110)-C10ss complexes
C10ss (215 nts)
1 2 3 4
G
HC
C10ss
47
67
% (SDN1-110)-DNA complexes
100
80
60
40
20
0
0.07 μM [SND1-110]
H
-
+
SND1-64
-
OL21
(0, 0.2, 0.7, 2, 7 nM)
(SND1-64)-HC complexes
HC
1 2 3 4 5 6
I
-
+
SND1-110
-
OL21
(0, 0.2, 0.7, 2, 7 nM)
(SND1-110)-HC complexes
HC
1 2 3 4 5 6
% (SDN1-110)-HC complexes
100
80
60
40
20
0
0 0.2 0.7 2 7
[OL21] nM

**Figure 5**

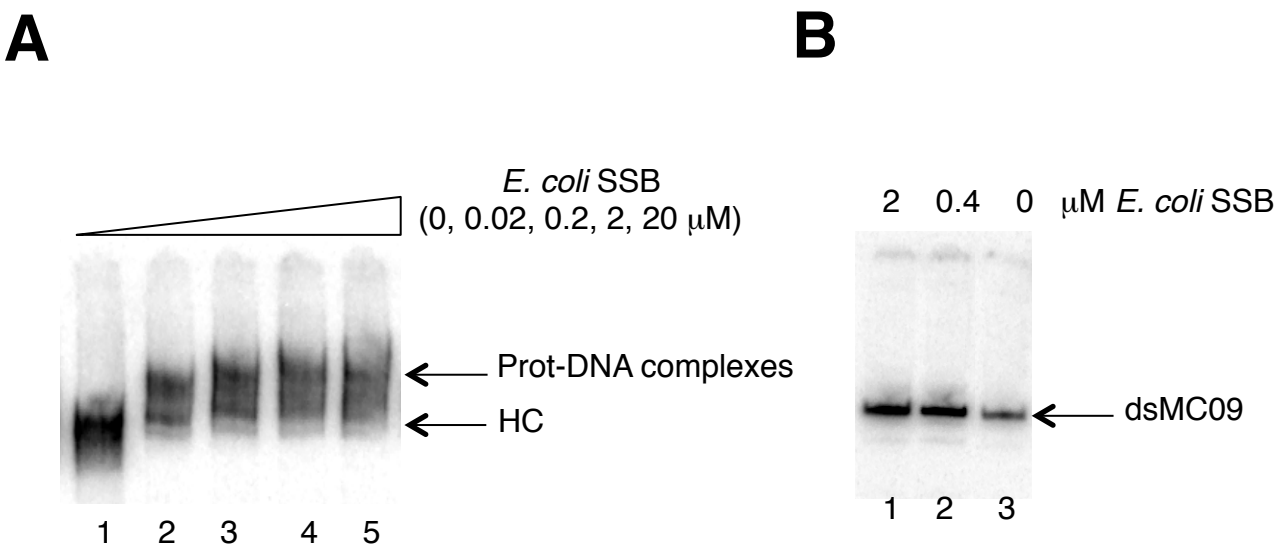

A
E. coli SSB
(0, 0.02, 0.2, 2, 20 μM)
Prot-DNA complexes
HC
1 2 3 4 5
B
2 0.4 0 μM E. coli SSB
dsMC09
1 2 3

**Figure 6**

## Supplementary Table 1: List of the proteins identified by mass spectrometry in the indicated gel piece

### Gel piece: 11p-beta

| | UnirProtKB ID | Name of the protein | Mascot score | Number of identified peptides | TheoricalMW (Da) | Theorical pI | Sequence coverage |
|---|---|---|---|---|---|---|---|
| 1 | MOES_HUMAN | Moesin OS=Homo sapiens GN=MSN PE=1 SV=3 | 250 | 9 | 67778 | 6.08 | 15.8 |
| 2 | EZRI_HUMAN | Ezrin OS=Homo sapiens GN=EZR PE=1 SV=4 | 110 | 5 | 69370 | 5.94 | 7.5 |
| 3 | PSPC1_HUMAN | Paraspeckle component 1 OS=Homo sapiens GN=PSPC1 PE=1 SV=1 | 91 | 2 | 58706 | 6.26 | 3.6 |
| 4 | RADI_HUMAN | Radixin OS=Homo sapiens GN=RDX PE=1 SV=1 | 90 | 4 | 68521 | 6.03 | 6.3 |
| 5 | PTBP1_HUMAN | Polypyrimidine tract-binding protein 1 OS=Homo sapiens GN=PTBP1 PE=1 SV=1 | 77 | 3 | 57186 | 9.22 | 5.1 |
| 6 | FILA2_HUMAN | Filaggrin-2 OS=Homo sapiens GN=FLG2 PE=1 SV=1 | 74 | 1 | 247928 | 8.45 | 0.5 |
| 7 | NONO_HUMAN | Non-POU domain-containing octamer-binding protein OS=Homo sapiens GN=NONO PE=1 SV=4 | 65 | 3 | 54197 | 9.01 | 6.4 |
| 8 | DCD_HUMAN | Dermcidin OS=Homo sapiens GN=DCD PE=1 SV=2 | 64 | 2 | 11277 | 6.08 | 17.3 |
| 9 | SARNP_HUMAN | SAP domain-containing ribonucleoprotein OS=Homo sapiens GN=SARNP PE=1 SV=3 | 50 | 2 | 23656 | 6.1 | 8.6 |
| 10 | FUBP1_HUMAN | Far upstream element-binding protein 1 OS=Homo sapiens GN=FUBP1 PE=1 SV=3 | 44 | 1 | 67518 | 7.18 | 1.6 |
| 11 | G3P_HUMAN | Glyceraldehyde-3-phosphate dehydrogenase OS=Homo sapiens GN=GAPDH PE=1 SV=3 | 39 | 1 | 36030 | 8.57 | 2.4 |
| 12 | SFPQ_HUMAN | Splicing factor. proline- and glutamine-rich OS=Homo sapiens GN=SFPQ PE=1 SV=2 | 38 | 1 | 76102 | 9.45 | 1.1 |
| 13 | HDGF_HUMAN | Hepatoma-derived growth factor OS=Homo sapiens GN=HDGF PE=1 SV=1 | 36 | 1 | 26772 | 4.7 | 4.6 |
| 14 | SHRM3_HUMAN | Protein Shroom3 OS=Homo sapiens GN=SHROOM3 PE=1 SV=2 | 33 | 1 | 216724 | 7.87 | 0.5 |
| 15 | CASPE_HUMAN | Caspase-14 OS=Homo sapiens GN=CASP14 PE=1 SV=2 | 29 | 1 | 27662 | 5.44 | 3.3 |
| 16 | TMPSD_HUMAN | Transmembrane protease serine 13 OS=Homo sapiens GN=TMPRSS13 PE=2 SV=4 | 29 | 1 | 63113 | 8.96 | 1.4 |

OS: organism name; GN: Gene Name; PE: Protein Existence; SV: Sequence Version.

**Gel piece: 11m-beta**

| | UnirProtKB ID | Name of the protein | Mascot score | Number of identified peptides | Theorical MW (Da) | Theorical pI | Sequence coverage |
|---|---|---|---|---|---|---|---|
| 1 | FILA2_HUMAN | Filaggrin-2 OS=Homo sapiens GN=FLG2 PE=1 SV=1 | 351 | 7 | 247928 | 8.45 | 4.3 |
| 2 | SARNP_HUMAN | SAP domain-containing ribonucleoprotein OS=Homo sapiens GN=SARNP PE=1 SV=3 | 167 | 3 | 23656 | 6.1 | 17.1 |
| 3 | SBSN_HUMAN | Suprabasin OS=Homo sapiens GN=SBSN PE=1 SV=2 | 161 | 5 | 60505 | 6.5 | 11.5 |
| 4 | DCD_HUMAN | Dermcidin OS=Homo sapiens GN=DCD PE=1 SV=2 | 97 | 2 | 11277 | 6.08 | 17.3 |
| 5 | DMKN_HUMAN | Dermokine OS=Homo sapiens GN=DMKN PE=1 SV=3 | 86 | 2 | 47054 | 6.8 | 4.2 |
| 6 | G3P_HUMAN | Glyceraldehyde-3-phosphate dehydrogenase OS=Homo sapiens GN=GAPDH PE=1 SV=3 | 75 | 2 | 36030 | 8.57 | 4.8 |
| 7 | FABP5_HUMAN | Fatty acid-binding protein. epidermal OS=Homo sapiens GN=FABP5 PE=1 SV=3 | 73 | 2 | 15155 | 6.6 | 7.4 |
| 8 | FILA_HUMAN | Filaggrin OS=Homo sapiens GN=FLG PE=1 SV=3 | 71 | 2 | 434922 | 9.24 | 0.4 |
| 9 | MOES_HUMAN | Moesin OS=Homo sapiens GN=MSN PE=1 SV=3 | 68 | 2 | 67778 | 6.08 | 3.1 |
| 10 | ACTA_HUMAN | Actin. aortic smooth muscle OS=Homo sapiens GN=ACTA2 PE=1 SV=1 | 53 | 2 | 41982 | 5.23 | 4.5 |
| 11 | SPR1A_HUMAN | Cornifin-A OS=Homo sapiens GN=SPRR1A PE=1 SV=2 | 46 | 2 | 9871 | 8.85 | 18 |
| 12 | DSG1_HUMAN | Desmoglein-1 OS=Homo sapiens GN=DSG1 PE=1 SV=2 | 45 | 2 | 113676 | 4.9 | 1.9 |
| 13 | SPR2D_HUMAN | Small proline-rich protein 2D OS=Homo sapiens GN=SPRR2D PE=2 SV=2 | 45 | 1 | 7900 | 8.77 | 12.5 |
| 14 | EZRI_HUMAN | Ezrin OS=Homo sapiens GN=EZR PE=1 SV=4 | 44 | 2 | 69370 | 5.94 | 2.4 |
| 15 | ALBU_HUMAN | Serum albumin OS=Homo sapiens GN=ALB PE=1 SV=2 | 41 | 1 | 69321 | 5.92 | 1.1 |
| 16 | PLAK_HUMAN | Junction plakoglobin OS=Homo sapiens GN=JUP PE=1 SV=3 | 40 | 1 | 81693 | 5.75 | 1.1 |
| 17 | SHRM3_HUMAN | Protein Shroom3 OS=Homo sapiens GN=SHROOM3 PE=1 SV=2 | 39 | 1 | 216724 | 7.87 | 0.5 |
| 18 | TMPSD_HUMAN | Transmembrane protease serine 13 OS=Homo sapiens GN=TMPRSS13 PE=2 SV=4 | 36 | 1 | 63113 | 8.96 | 1.4 |
| 19 | CYTA_HUMAN | Cystatin-A OS=Homo sapiens GN=CSTA PE=1 SV=1 | 34 | 1 | 11000 | 5.38 | 8.2 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 20 | FUBP1_HUMAN | Far upstream element-binding protein 1 OS=Homo sapiens GN=FUBP1 PE=1 SV=3 | 31 | 1 | 67518 | 7.18 | 2 |
| 21 | IF2A_HUMAN | Eukaryotic translation initiation factor 2 subunit 1 OS=Homo sapiens GN=EIF2S1 PE=1 SV=3 | 29 | 1 | 36089 | 5.02 | 3.8 |
| 22 | NONO_HUMAN | Non-POU domain-containing octamer-binding protein OS=Homo sapiens GN=NONO PE=1 SV=4 | 27 | 1 | 54197 | 9.01 | 1.5 |
| 23 | RS27A_HUMAN | Ubiquitin-40S ribosomal protein S27a OS=Homo sapiens GN=RPS27A PE=1 SV=2 | 27 | 1 | 17953 | 9.68 | 5.8 |
| 24 | KLK11_HUMAN | Kallikrein-11 OS=Homo sapiens GN=KLK11 PE=1 SV=2 | 26 | 1 | 31039 | 9.23 | 2.5 |

OS: organism name; GN: Gene Name; PE: Protein Existence; SV: Sequence Version.

## Gel piece : 10m1-beta

| | UnirProtKB ID | Name of the protein | Mascot score | Number of identified peptides | Theorical MW (Da) | Theorical pI | Sequence coverage |
|---|---|---|---|---|---|---|---|
| 1 | THIL_HUMAN | Acetyl-CoA acetyltransferase. mitochondrial OS=Homo sapiens GN=ACAT1 PE=1 SV=1 | 234 | 4 | 45171 | 8.98 | 11 |
| 2 | RADI_HUMAN | Radixin OS=Homo sapiens GN=RDX PE=1 SV=1 | 190 | 8 | 68521 | 6.03 | 12.7 |
| 3 | MOES_HUMAN | Moesin OS=Homo sapiens GN=MSN PE=1 SV=3 | 153 | 6 | 67778 | 6.08 | 9.2 |
| 4 | EZRI_HUMAN | Ezrin OS=Homo sapiens GN=EZR PE=1 SV=4 | 126 | 5 | 69370 | 5.94 | 7 |
| 5 | FABP5_HUMAN | Fatty acid-binding protein. epidermal OS=Homo sapiens GN=FABP5 PE=1 SV=3 | 96 | 2 | 15155 | 6.6 | 7.4 |
| 6 | THIM_HUMAN | 3-ketoacyl-CoA thiolase. mitochondrial OS=Homo sapiens GN=ACAA2 PE=1 SV=2 | 89 | 3 | 41898 | 8.32 | 7.8 |
| 7 | DCD_HUMAN | Dermcidin OS=Homo sapiens GN=DCD PE=1 SV=2 | 87 | 2 | 11277 | 6.08 | 17.3 |
| 8 | ALBU_HUMAN | Serum albumin OS=Homo sapiens GN=ALB PE=1 SV=2 | 73 | 2 | 69321 | 5.92 | 3.3 |
| 9 | DSG1_HUMAN | Desmoglein-1 OS=Homo sapiens GN=DSG1 PE=1 SV=2 | 70 | 2 | 113676 | 4.9 | 2.9 |
| 10 | HMGN5_HUMAN | High mobility group nucleosome-binding domain-containing protein 5 OS=Homo sapiens GN=HMGN5 PE=1 SV=1 | 66 | 2 | 31506 | 4.5 | 6 |
| 11 | PSIP1_HUMAN | PC4 and SFRS1-interacting protein OS=Homo sapiens GN=PSIP1 PE=1 SV=1 | 51 | 1 | 60067 | 9.15 | 1.7 |
| 12 | ZA2G_HUMAN | Zinc-alpha-2-glycoprotein OS=Homo sapiens GN=AZGP1 PE=1 SV=2 | 51 | 1 | 34237 | 5.71 | 3.4 |
| 13 | FILA2_HUMAN | Filaggrin-2 OS=Homo sapiens GN=FLG2 PE=1 SV=1 | 50 | 1 | 247928 | 8.45 | 0.5 |
| 14 | SBSN_HUMAN | Suprabasin OS=Homo sapiens GN=SBSN PE=1 SV=2 | 47 | 1 | 60505 | 6.5 | 3.1 |
| 15 | ACTA_HUMAN | Actin. aortic smooth muscle OS=Homo sapiens GN=ACTA2 PE=1 SV=1 | 46 | 2 | 41982 | 5.23 | 5.6 |
| 16 | SPR2D_HUMAN | Small proline-rich protein 2D OS=Homo sapiens GN=SPRR2D PE=2 SV=2 | 43 | 1 | 7900 | 8.77 | 12.5 |
| 17 | PLAK_HUMAN | Junction plakoglobin OS=Homo sapiens GN=JUP PE=1 SV=3 | 42 | 1 | 81693 | 5.75 | 1.1 |
| 18 | LAC2_HUMAN | Ig lambda-2 chain C regions OS=Homo sapiens GN=IGLC2 PE=1 SV=1 | 41 | 1 | 11287 | n.c. | 9.7 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 19 | G3P_HUMAN | Glyceraldehyde-3-phosphate dehydrogenase OS=Homo sapiens GN=GAPDH PE=1 SV=3 | 38 | 1 | 36030 | 8.57 | 2.4 |
| 20 | SHRM3_HUMAN | Protein Shroom3 OS=Homo sapiens GN=SHROOM3 PE=1 SV=2 | 38 | 1 | 216724 | 7.87 | 0.5 |
| 21 | THIO_HUMAN | Thioredoxin OS=Homo sapiens GN=TXN PE=1 SV=3 | 38 | 1 | 11730 | 4.82 | 10.5 |
| 22 | TMPSD_HUMAN | Transmembrane protease serine 13 OS=Homo sapiens GN=TMPRSS13 PE=2 SV=4 | 38 | 1 | 63113 | 8.96 | 1.4 |
| 23 | IREB2_HUMAN | Iron-responsive element-binding protein 2 OS=Homo sapiens GN=IREB2 PE=1 SV=3 | 37 | 1 | 104978 | 6.62 | 0.7 |
| 24 | CYTA_HUMAN | Cystatin-A OS=Homo sapiens GN=CSTA PE=1 SV=1 | 36 | 1 | 11000 | 5.38 | 8.2 |
| 25 | SF01_HUMAN | Splicing factor 1 OS=Homo sapiens GN=SF1 PE=1 SV=4 | 34 | 1 | 68286 | 9.07 | 2 |
| 26 | SPR1B_HUMAN | Cornifin-B OS=Homo sapiens GN=SPRR1B PE=1 SV=2 | 31 | 1 | 9881 | 8.85 | 9 |
| 27 | HDGF_HUMAN | Hepatoma-derived growth factor OS=Homo sapiens GN=HDGF PE=1 SV=1 | 31 | 1 | 26772 | 4.7 | 4.6 |
| 28 | NONO_HUMAN | Non-POU domain-containing octamer-binding protein OS=Homo sapiens GN=NONO PE=1 SV=4 | 30 | 1 | 54197 | 9.01 | 1.5 |
| 29 | CDSN_HUMAN | Corneodesmosin OS=Homo sapiens GN=CDSN PE=1 SV=3 | 29 | 1 | 51490 | 8.69 | 1.7 |
| 30 | HNRPU_HUMAN | Heterogeneous nuclear ribonucleoprotein U OS=Homo sapiens GN=HNRNPU PE=1 SV=6 | 29 | 1 | 90528 | 5.76 | 1.1 |
| 31 | SPG21_HUMAN | Maspardin OS=Homo sapiens GN=SPG21 PE=1 SV=1 | 26 | 1 | 34938 | 5.85 | 3.2 |
| 32 | H3C_HUMAN | Histone H3.3C OS=Homo sapiens GN=H3F3C PE=1 SV=3 | 26 | 1 | 15204 | 11.1 | 5.2 |
| 33 | TRY3_HUMAN | Trypsin-3 OS=Homo sapiens GN=PRSS3 PE=1 SV=2 | 26 | 1 | 32508 | 7.46 | 4.3 |
| 34 | FHDC1_HUMAN | FH2 domain-containing protein 1 OS=Homo sapiens GN=FHDC1 PE=1 SV=2 | 25 | 1 | 124684 | 9.17 | 0.6 |
| 35 | HSP7C_HUMAN | Heat shock cognate 71 kDa protein OS=Homo sapiens GN=HSPA8 PE=1 SV=1 | 25 | 1 | 70854 | 5.37 | 1.1 |

OS: organism name; GN: Gene Name; PE: Protein Existence; SV: Sequence Version.

## Gel piece: 10p1-beta

| | UnirProtKB ID | Name of the protein | Mascot score | Number of identified peptides | Theorical MW (Da) | Theorical pI | Sequence coverage |
|---|---|---|---|---|---|---|---|
| 1 | NONO_HUMAN | Non-POU domain-containing octamer-binding protein OS=Homo sapiens GN=NONO PE=1 SV=4 | 363 | 8 | 54197 | 9.01 | 19.3 |
| 2 | RADI_HUMAN | Radixin OS=Homo sapiens GN=RDX PE=1 SV=1 | 256 | 13 | 68521 | 6.03 | 20.9 |
| 3 | EZRI_HUMAN | Ezrin OS=Homo sapiens GN=EZR PE=1 SV=4 | 214 | 7 | 69370 | 5.94 | 10.9 |
| 4 | ALBU_HUMAN | Serum albumin OS=Homo sapiens GN=ALB PE=1 SV=2 | 193 | 6 | 69321 | 5.92 | 11.7 |
| 5 | THIL_HUMAN | Acetyl-CoA acetyltransferase. mitochondrial OS=Homo sapiens GN=ACAT1 PE=1 SV=1 | 193 | 6 | 45171 | 8.98 | 15 |
| 6 | MOES_HUMAN | Moesin OS=Homo sapiens GN=MSN PE=1 SV=3 | 143 | 7 | 67778 | 6.08 | 11.1 |
| 7 | FABP5_HUMAN | Fatty acid-binding protein. epidermal OS=Homo sapiens GN=FABP5 PE=1 SV=3 | 142 | 3 | 15155 | 6.6 | 17.8 |
| 8 | FILA2_HUMAN | Filaggrin-2 OS=Homo sapiens GN=FLG2 PE=1 SV=1 | 120 | 2 | 247928 | 8.45 | 0.9 |
| 9 | DCD_HUMAN | Dermcidin OS=Homo sapiens GN=DCD PE=1 SV=2 | 106 | 2 | 11277 | 6.08 | 17.3 |
| 10 | SND1_HUMAN | Staphylococcal nuclease domain-containing protein 1 OS=Homo sapiens GN=SND1 PE=1 SV=1 | 98 | 2 | 101934 | 6.74 | 2.6 |
| 11 | SFPQ_HUMAN | Splicing factor. proline- and glutamine-rich OS=Homo sapiens GN=SFPQ PE=1 SV=2 | 97 | 3 | 76102 | 9.45 | 4.8 |
| 12 | HMGN5_HUMAN | High mobility group nucleosome-binding domain-containing protein 5 OS=Homo sapiens GN=HMGN5 PE=1 SV=1 | 94 | 3 | 31506 | 4.5 | 11 |
| 13 | PSPC1_HUMAN | Paraspeckle component 1 OS=Homo sapiens GN=PSPC1 PE=1 SV=1 | 74 | 3 | 58706 | 6.26 | 5 |
| 14 | IF2A_HUMAN | Eukaryotic translation initiation factor 2 subunit 1 OS=Homo sapiens GN=EIF2S1 PE=1 SV=3 | 65 | 1 | 36089 | 5.02 | 3.8 |
| 15 | G3P_HUMAN | Glyceraldehyde-3-phosphate dehydrogenase OS=Homo sapiens GN=GAPDH PE=1 SV=3 | 62 | 2 | 36030 | 8.57 | 4.8 |
| 16 | ZA2G_HUMAN | Zinc-alpha-2-glycoprotein OS=Homo sapiens GN=AZGP1 PE=1 SV=2 | 51 | 1 | 34237 | 5.71 | 3.4 |
| 17 | LAC2_HUMAN | Ig lambda-2 chain C regions OS=Homo sapiens GN=IGLC2 PE=1 SV=1 | 48 | 1 | 11287 | n.c. | 9.7 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 18 | EF1A1_HUMAN | Elongation factor 1-alpha 1 OS=Homo sapiens GN=EEF1A1 PE=1 SV=1 | 47 | 1 | 50109 | 9.1 | 2.4 |
| 19 | THIM_HUMAN | 3-ketoacyl-CoA thiolase. mitochondrial OS=Homo sapiens GN=ACAA2 PE=1 SV=2 | 46 | 2 | 41898 | 8.32 | 5.8 |
| 20 | CYTA_HUMAN | Cystatin-A OS=Homo sapiens GN=CSTA PE=1 SV=1 | 45 | 1 | 11000 | 5.38 | 8.2 |
| 21 | IGJ_HUMAN | Immunoglobulin J chain OS=Homo sapiens GN=JCHAIN PE=1 SV=4 | 44 | 1 | 18087 | 5.12 | 7.5 |
| 22 | SYYM_HUMAN | Tyrosine--tRNA ligase. mitochondrial OS=Homo sapiens GN=YARS2 PE=1 SV=2 | 43 | 1 | 53166 | 9.07 | 1.5 |
| 23 | DEK_HUMAN | Protein DEK OS=Homo sapiens GN=DEK PE=1 SV=1 | 43 | 2 | 42648 | 8.69 | 2.7 |
| 24 | S10A9_HUMAN | Protein S100-A9 OS=Homo sapiens GN=S100A9 PE=1 SV=1 | 42 | 1 | 13234 | 5.71 | 11.4 |
| 25 | DSC1_HUMAN | Desmocollin-1 OS=Homo sapiens GN=DSC1 PE=1 SV=2 | 41 | 1 | 99924 | 5.25 | 1.5 |
| 26 | CALL5_HUMAN | Calmodulin-like protein 5 OS=Homo sapiens GN=CALML5 PE=1 SV=2 | 41 | 1 | 15883 | 4.34 | 6.2 |
| 27 | VIGLN_HUMAN | Vigilin OS=Homo sapiens GN=HDLBP PE=1 SV=2 | 41 | 1 | 141368 | 6.43 | 0.9 |
| 28 | IGHA1_HUMAN | Ig alpha-1 chain C region OS=Homo sapiens GN=IGHA1 PE=1 SV=2 | 39 | 1 | 37631 | 6.08 | 2.5 |
| 29 | ARGI1_HUMAN | Arginase-1 OS=Homo sapiens GN=ARG1 PE=1 SV=2 | 38 | 1 | 34713 | 6.72 | 2.2 |
| 30 | SBSN_HUMAN | Suprabasin OS=Homo sapiens GN=SBSN PE=1 SV=2 | 36 | 1 | 60505 | 6.5 | 3.1 |
| 31 | KLK11_HUMAN | Kallikrein-11 OS=Homo sapiens GN=KLK11 PE=1 SV=2 | 36 | 1 | 31039 | 9.23 | 2.5 |
| 32 | SHRM3_HUMAN | Protein Shroom3 OS=Homo sapiens GN=SHROOM3 PE=1 SV=2 | 35 | 1 | 216724 | 7.87 | 0.5 |
| 33 | BUB3_HUMAN | Mitotic checkpoint protein BUB3 OS=Homo sapiens GN=BUB3 PE=1 SV=1 | 35 | 1 | 37131 | 6.36 | 4 |
| 34 | SF01_HUMAN | Splicing factor 1 OS=Homo sapiens GN=SF1 PE=1 SV=4 | 35 | 1 | 68286 | 9.07 | 2 |
| 35 | FUBP1_HUMAN | Far upstream element-binding protein 1 OS=Homo sapiens GN=FUBP1 PE=1 SV=3 | 34 | 1 | 67518 | 7.18 | 1.2 |
| 36 | SPR1B_HUMAN | Cornifin-B OS=Homo sapiens GN=SPRR1B PE=1 SV=2 | 31 | 1 | 9881 | 8.85 | 9 |
| 37 | BASP1_HUMAN | Brain acid soluble protein 1 OS=Homo sapiens GN=BASP1 PE=1 SV=2 | 30 | 1 | 22680 | 4.64 | 6.2 |
| 38 | TMPSD_HUMAN | Transmembrane protease serine 13 OS=Homo | 28 | 1 | 63113 | 8.96 | 1.4 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | sapiens GN=TMPRSS13 PE=2 SV=4 | | | | | |
| 39 | HNRPU_HUMAN | Heterogeneous nuclear ribonucleoprotein U OS=Homo sapiens GN=HNRNPU PE=1 SV=6 | 28 | 1 | 90528 | 5.76 | 1.1 |
| 40 | FCSD2_HUMAN | F-BAR and double SH3 domains protein 2 OS=Homo sapiens GN=FCHSD2 PE=1 SV=3 | 27 | 1 | 84224 | 5.55 | 1.5 |
| 41 | FILA_HUMAN | Filaggrin OS=Homo sapiens GN=FLG PE=1 SV=3 | 27 | 1 | 434922 | 9.24 | 0.2 |
| 42 | TRY3_HUMAN | Trypsin-3 OS=Homo sapiens GN=PRSS3 PE=1 SV=2 | 27 | 1 | 32508 | 7.46 | 4.3 |
| 43 | C2C2L_HUMAN | C2 domain-containing protein 2-like OS=Homo sapiens GN=C2CD2L PE=1 SV=3 | 26 | 1 | 76134 | 7.62 | 1.4 |
| 44 | GLSL_HUMAN | Glutaminase liver isoform. mitochondrial OS=Homo sapiens GN=GLS2 PE=1 SV=2 | 26 | 1 | 66280 | 6.9 | 1.2 |

OS: organism name; GN: Gene Name; PE: Protein Existence; SV: Sequence Version.

## Gel piece: 10m2-beta

| | UnirProtKB ID | Name of the protein | Mascot score | Number of identified peptides | Theorical MW (Da) | Theorical pI | Sequence coverage |
|---|---|---|---|---|---|---|---|
| 1 | RADI_HUMAN | Radixin OS=Homo sapiens GN=RDX PE=1 SV=1 | 193 | 7 | 68521 | 6.03 | 10.3 |
| 2 | EZRI_HUMAN | Ezrin OS=Homo sapiens GN=EZR PE=1 SV=4 | 170 | 7 | 69370 | 5.94 | 10.2 |
| 3 | MOES_HUMAN | Moesin OS=Homo sapiens GN=MSN PE=1 SV=3 | 170 | 7 | 67778 | 6.08 | 11.3 |
| 4 | FILA2_HUMAN | Filaggrin-2 OS=Homo sapiens GN=FLG2 PE=1 SV=1 | 152 | 4 | 247928 | 8.45 | 2.3 |
| 5 | THIL_HUMAN | Acetyl-CoA acetyltransferase. mitochondrial OS=Homo sapiens GN=ACAT1 PE=1 SV=1 | 132 | 4 | 45171 | 8.98 | 10.8 |
| 6 | THIM_HUMAN | 3-ketoacyl-CoA thiolase. mitochondrial OS=Homo sapiens GN=ACAA2 PE=1 SV=2 | 93 | 3 | 41898 | 8.32 | 7.6 |
| 7 | DCD_HUMAN | Dermcidin OS=Homo sapiens GN=DCD PE=1 SV=2 | 89 | 2 | 11277 | 6.08 | 17.3 |
| 8 | SBSN_HUMAN | Suprabasin OS=Homo sapiens GN=SBSN PE=1 SV=2 | 74 | 2 | 60505 | 6.5 | 6.1 |
| 9 | IGHA1_HUMAN | Ig alpha-1 chain C region OS=Homo sapiens GN=IGHA1 PE=1 SV=2 | 49 | 1 | 37631 | 6.08 | 2.5 |
| 10 | SF3B1_HUMAN | Splicing factor 3B subunit 1 OS=Homo sapiens GN=SF3B1 PE=1 SV=3 | 47 | 1 | 145738 | 6.65 | 0.9 |
| 11 | FABP5_HUMAN | Fatty acid-binding protein. epidermal OS=Homo sapiens GN=FABP5 PE=1 SV=3 | 44 | 1 | 15155 | 6.6 | 6.7 |
| 12 | SPR2D_HUMAN | Small proline-rich protein 2D OS=Homo sapiens GN=SPRR2D PE=2 SV=2 | 42 | 1 | 7900 | 8.77 | 12.5 |
| 13 | TRY3_HUMAN | Trypsin-3 OS=Homo sapiens GN=PRSS3 PE=1 SV=2 | 42 | 1 | 32508 | 7.46 | 4.3 |
| 14 | TR150_HUMAN | Thyroid hormone receptor-associated protein 3 OS=Homo sapiens GN=THRAP3 PE=1 SV=2 | 40 | 1 | 108601 | 10.16 | 0.9 |
| 15 | SF01_HUMAN | Splicing factor 1 OS=Homo sapiens GN=SF1 PE=1 SV=4 | 40 | 1 | 68286 | 9.07 | 2 |
| 16 | G3P_HUMAN | Glyceraldehyde-3-phosphate dehydrogenase OS=Homo sapiens GN=GAPDH PE=1 SV=3 | 40 | 1 | 36030 | 8.57 | 2.4 |
| 17 | DSC1_HUMAN | Desmocollin-1 OS=Homo sapiens GN=DSC1 PE=1 SV=2 | 39 | 1 | 99924 | 5.25 | 1.5 |
| 18 | TMPSD_HUMAN | Transmembrane protease serine 13 OS=Homo sapiens GN=TMPRSS13 PE=2 SV=4 | 36 | 1 | 63113 | 8.96 | 1.4 |
| 19 | FCSD2_HUMAN | F-BAR and double SH3 domains protein 2 OS=Homo sapiens GN=FCHSD2 PE=1 SV=3 | 36 | 1 | 84224 | 5.55 | 1.5 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 20 | SHRM3_HUMAN | Protein Shroom3 OS=Homo sapiens GN=SHROOM3 PE=1 SV=2 | 35 | 1 | 216724 | 7.87 | 0.5 |
| 21 | ALBU_HUMAN | Serum albumin OS=Homo sapiens GN=ALB PE=1 SV=2 | 32 | 1 | 69321 | 5.92 | 1.1 |
| 22 | NONO_HUMAN | Non-POU domain-containing octamer-binding protein OS=Homo sapiens GN=NONO PE=1 SV=4 | 31 | 1 | 54197 | 9.01 | 1.5 |
| 23 | HDGF_HUMAN | Hepatoma-derived growth factor OS=Homo sapiens GN=HDGF PE=1 SV=1 | 31 | 1 | 26772 | 4.7 | 4.6 |
| 24 | ZA2G_HUMAN | Zinc-alpha-2-glycoprotein OS=Homo sapiens GN=AZGP1 PE=1 SV=2 | 31 | 1 | 34237 | 5.71 | 3.4 |
| 25 | ARGI1_HUMAN | Arginase-1 OS=Homo sapiens GN=ARG1 PE=1 SV=2 | 30 | 1 | 34713 | 6.72 | 2.2 |
| 26 | HMGN5_HUMAN | High mobility group nucleosome-binding domain-containing protein 5 OS=Homo sapiens GN=HMGN5 PE=1 SV=1 | 28 | 1 | 31506 | 4.5 | 2.8 |
| 27 | ITPR2_HUMAN | Inositol 1.4.5-trisphosphate receptor type 2 OS=Homo sapiens GN=ITPR2 PE=1 SV=2 | 27 | 1 | 307867 | 6.01 | 0.3 |
| 28 | HSP7C_HUMAN | Heat shock cognate 71 kDa protein OS=Homo sapiens GN=HSPA8 PE=1 SV=1 | 25 | 1 | 70854 | 5.37 | 1.1 |

OS: organism name; GN: Gene Name; PE: Protein Existence; SV: Sequence Version.

## Gel piece: 10m2-gamma

| | UnirProtKB ID | Name of the protein | Mascot score | Number of identified peptides | Theorical MW (Da) | Theorical pI | Sequence coverage |
|---|---|---|---|---|---|---|---|
| 1 | THIL_HUMAN | Acetyl-CoA acetyltransferase. mitochondrial OS=Homo sapiens GN=ACAT1 PE=1 SV=1 | 242 | 5 | 45171 | 8.98 | 13.8 |
| 2 | EZRI_HUMAN | Ezrin OS=Homo sapiens GN=EZR PE=1 SV=4 | 223 | 10 | 69370 | 5.94 | 14.2 |
| 3 | MOES_HUMAN | Moesin OS=Homo sapiens GN=MSN PE=1 SV=3 | 186 | 7 | 67778 | 6.08 | 11.3 |
| 4 | RADI_HUMAN | Radixin OS=Homo sapiens GN=RDX PE=1 SV=1 | 165 | 8 | 68521 | 6.03 | 13.2 |
| 5 | FILA2_HUMAN | Filaggrin-2 OS=Homo sapiens GN=FLG2 PE=1 SV=1 | 152 | 3 | 247928 | 8.45 | 1.9 |
| 6 | THIM_HUMAN | 3-ketoacyl-CoA thiolase. mitochondrial OS=Homo sapiens GN=ACAA2 PE=1 SV=2 | 107 | 4 | 41898 | 8.32 | 8.1 |
| 7 | PSIP1_HUMAN | PC4 and SFRS1-interacting protein OS=Homo sapiens GN=PSIP1 PE=1 SV=1 | 96 | 2 | 60067 | 9.15 | 4.2 |
| 8 | DCD_HUMAN | Dermcidin OS=Homo sapiens GN=DCD PE=1 SV=2 | 90 | 2 | 11277 | 6.08 | 17.3 |
| 9 | NONO_HUMAN | Non-POU domain-containing octamer-binding protein OS=Homo sapiens GN=NONO PE=1 SV=4 | 85 | 4 | 54197 | 9.01 | 8.1 |
| 10 | TRY3_HUMAN | Trypsin-3 OS=Homo sapiens GN=PRSS3 PE=1 SV=2 | 73 | 1 | 32508 | 7.46 | 4.3 |
| 11 | HMGN5_HUMAN | High mobility group nucleosome-binding domain-containing protein 5 OS=Homo sapiens GN=HMGN5 PE=1 SV=1 | 73 | 3 | 31506 | 4.5 | 9.6 |
| 12 | ACTB_HUMAN | Actin. cytoplasmic 1 OS=Homo sapiens GN=ACTB PE=1 SV=1 | 58 | 2 | 41710 | 5.29 | 5.3 |
| 13 | SBSN_HUMAN | Suprabasin OS=Homo sapiens GN=SBSN PE=1 SV=2 | 57 | 2 | 60505 | 6.5 | 6.1 |
| 14 | IF2A_HUMAN | Eukaryotic translation initiation factor 2 subunit 1 OS=Homo sapiens GN=EIF2S1 PE=1 SV=3 | 51 | 1 | 36089 | 5.02 | 3.8 |
| 15 | G3P_HUMAN | Glyceraldehyde-3-phosphate dehydrogenase OS=Homo sapiens GN=GAPDH PE=1 SV=3 | 51 | 2 | 36030 | 8.57 | 6.9 |
| 16 | HDGF_HUMAN | Hepatoma-derived growth factor OS=Homo sapiens GN=HDGF PE=1 SV=1 | 46 | 1 | 26772 | 4.7 | 4.6 |
| 17 | BASP1_HUMAN | Brain acid soluble protein 1 OS=Homo sapiens GN=BASP1 PE=1 SV=2 | 45 | 2 | 22680 | 4.64 | 12.3 |
| 18 | FABP5_HUMAN | Fatty acid-binding protein. epidermal OS=Homo sapiens GN=FABP5 PE=1 SV=3 | 44 | 1 | 15155 | 6.6 | 6.7 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 19 | TR150_HUMAN | Thyroid hormone receptor-associated protein 3 OS=Homo sapiens GN=THRAP3 PE=1 SV=2 | 42 | 1 | 108601 | 10.16 | 0.9 |
| 20 | ZA2G_HUMAN | Zinc-alpha-2-glycoprotein OS=Homo sapiens GN=AZGP1 PE=1 SV=2 | 42 | 1 | 34237 | 5.71 | 3.4 |
| 21 | SPR2D_HUMAN | Small proline-rich protein 2D OS=Homo sapiens GN=SPRR2D PE=2 SV=2 | 39 | 1 | 7900 | 8.77 | 12.5 |
| 22 | SHRM3_HUMAN | Protein Shroom3 OS=Homo sapiens GN=SHROOM3 PE=1 SV=2 | 39 | 1 | 216724 | 7.87 | 0.5 |
| 23 | ALBU_HUMAN | Serum albumin OS=Homo sapiens GN=ALB PE=1 SV=2 | 38 | 1 | 69321 | 5.92 | 1.1 |
| 24 | NACAM_HUMAN | Nascent polypeptide-associated complex subunit alpha. muscle-specific form OS=Homo sapiens GN=NACA PE=1 SV=1 | 36 | 1 | 205295 | 9.6 | 0.7 |
| 25 | IGHA1_HUMAN | Ig alpha-1 chain C region OS=Homo sapiens GN=IGHA1 PE=1 SV=2 | 34 | 1 | 37631 | 6.08 | 2.5 |
| 26 | DSG1_HUMAN | Desmoglein-1 OS=Homo sapiens GN=DSG1 PE=1 SV=2 | 32 | 1 | 113676 | 4.9 | 1 |
| 27 | SPR1B_HUMAN | Cornifin-B OS=Homo sapiens GN=SPRR1B PE=1 SV=2 | 31 | 1 | 9881 | 8.85 | 9 |
| 28 | HSP7C_HUMAN | Heat shock cognate 71 kDa protein OS=Homo sapiens GN=HSPA8 PE=1 SV=1 | 31 | 1 | 70854 | 5.37 | 1.1 |
| 29 | S10A9_HUMAN | Protein S100-A9 OS=Homo sapiens GN=S100A9 PE=1 SV=1 | 30 | 1 | 13234 | 5.71 | 11.4 |
| 30 | TMPSD_HUMAN | Transmembrane protease serine 13 OS=Homo sapiens GN=TMPRSS13 PE=2 SV=4 | 30 | 1 | 63113 | 8.96 | 1.4 |
| 31 | CYTA_HUMAN | Cystatin-A OS=Homo sapiens GN=CSTA PE=1 SV=1 | 27 | 1 | 11000 | 5.38 | 8.2 |
| 32 | KLK11_HUMAN | Kallikrein-11 OS=Homo sapiens GN=KLK11 PE=1 SV=2 | 27 | 1 | 31039 | 9.23 | 2.5 |
| 33 | SPG21_HUMAN | Maspardin OS=Homo sapiens GN=SPG21 PE=1 SV=1 | 26 | 1 | 34938 | 5.85 | 3.2 |
| 34 | C2C2L_HUMAN | C2 domain-containing protein 2-like OS=Homo sapiens GN=C2CD2L PE=1 SV=3 | 26 | 1 | 76134 | 7.62 | 1.4 |
| 35 | H2B1A_HUMAN | Histone H2B type 1-A OS=Homo sapiens GN=HIST1H2BA PE=1 SV=3 | 25 | 1 | 14159 | 10.31 | 5.5 |

OS: organism name; GN: Gene Name; PE: Protein Existence; SV: Sequence Version.

**Gel piece : 10p2-beta**

| | UnirProtKB ID | Name of the protein | Mascot score | Number of identified peptides | Theorical MW (Da) | Theorical pI | Sequence coverage |
|---|---|---|---|---|---|---|---|
| 1 | PLAK_HUMAN | Junction plakoglobin OS=Homo sapiens GN=JUP PE=1 SV=3 | 332 | 9 | 81693 | 5.75 | 14.1 |
| 2 | DSG1_HUMAN | Desmoglein-1 OS=Homo sapiens GN=DSG1 PE=1 SV=2 | 166 | 6 | 113676 | 4.9 | 9.4 |
| 3 | G3P_HUMAN | Glyceraldehyde-3-phosphate dehydrogenase OS=Homo sapiens GN=GAPDH PE=1 SV=3 | 150 | 5 | 36030 | 8.57 | 16.1 |
| 4 | CYTA_HUMAN | Cystatin-A OS=Homo sapiens GN=CSTA PE=1 SV=1 | 115 | 4 | 11000 | 5.38 | 49 |
| 5 | FABP5_HUMAN | Fatty acid-binding protein. epidermal OS=Homo sapiens GN=FABP5 PE=1 SV=3 | 112 | 3 | 15155 | 6.6 | 14.1 |
| 6 | TRY3_HUMAN | Trypsin-3 OS=Homo sapiens GN=PRSS3 PE=1 SV=2 | 79 | 1 | 32508 | 7.46 | 4.3 |
| 7 | ANXA2_HUMAN | Annexin A2 OS=Homo sapiens GN=ANXA2 PE=1 SV=2 | 64 | 2 | 38580 | 7.57 | 5 |
| 8 | SBSN_HUMAN | Suprabasin OS=Homo sapiens GN=SBSN PE=1 SV=2 | 61 | 2 | 60505 | 6.5 | 6.1 |
| 9 | THIL_HUMAN | Acetyl-CoA acetyltransferase. mitochondrial OS=Homo sapiens GN=ACAT1 PE=1 SV=1 | 49 | 1 | 45171 | 8.98 | 4 |
| 10 | SPB3_HUMAN | Serpin B3 OS=Homo sapiens GN=SERPINB3 PE=1 SV=2 | 43 | 1 | 44537 | 6.35 | 2.6 |
| 11 | RS27A_HUMAN | Ubiquitin-40S ribosomal protein S27a OS=Homo sapiens GN=RPS27A PE=1 SV=2 | 40 | 1 | 17953 | 9.68 | 5.8 |
| 12 | TMPSD_HUMAN | Transmembrane protease serine 13 OS=Homo sapiens GN=TMPRSS13 PE=2 SV=4 | 35 | 1 | 63113 | 8.96 | 1.4 |
| 13 | SHRM3_HUMAN | Protein Shroom3 OS=Homo sapiens GN=SHROOM3 PE=1 SV=2 | 34 | 1 | 216724 | 7.87 | 0.5 |
| 14 | FILA2_HUMAN | Filaggrin-2 OS=Homo sapiens GN=FLG2 PE=1 SV=1 | 34 | 1 | 247928 | 8.45 | 0.5 |
| 15 | HSP7C_HUMAN | Heat shock cognate 71 kDa protein OS=Homo sapiens GN=HSPA8 PE=1 SV=1 | 33 | 1 | 70854 | 5.37 | 1.1 |
| 16 | DCD_HUMAN | Dermcidin OS=Homo sapiens GN=DCD PE=1 SV=2 | 32 | 1 | 11277 | 6.08 | 10 |
| 17 | LEG7_HUMAN | Galectin-7 OS=Homo sapiens GN=LGALS7 PE=1 SV=2 | 29 | 1 | 15066 | 7.03 | 8.8 |
| 18 | KLK11_HUMAN | Kallikrein-11 OS=Homo sapiens GN=KLK11 PE=1 SV=2 | 28 | 1 | 31039 | 9.23 | 2.5 |
| 19 | ACTA_HUMAN | Actin. aortic smooth muscle OS=Homo sapiens GN=ACTA2 PE=1 SV=1 | 26 | 1 | 41982 | 5.23 | 1.9 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 20 | DSC1_HUMAN | Desmocollin-1 OS=Homo sapiens GN=DSC1 PE=1 SV=2 | 26 | 1 | 99924 | 5.25 | 1.5 |
| 21 | ARGI1_HUMAN | Arginase-1 OS=Homo sapiens GN=ARG1 PE=1 SV=2 | 26 | 1 | 34713 | 6.72 | 2.2 |
| 22 | PCMD1_HUMAN | Protein-L-isoaspartate O-methyltransferase domain-containing protein 1 OS=Homo sapiens GN=PCMTD1 PE=1 SV=2 | 26 | 1 | 40650 | 5.46 | 2 |

OS: organism name; GN: Gene Name; PE: Protein Existence; SV: Sequence Version.

## Gel piece: 10p2-gamma

| | UnirProtKB ID | Name of the protein | Mascot score | Number of identified peptides | Theorical MW (Da) | Theorical pI | Sequence coverage |
|---|---|---|---|---|---|---|---|
| 1 | EZRI_HUMAN | Ezrin OS=Homo sapiens GN=EZR PE=1 SV=4 | 211 | 9 | 69370 | 5.94 | 14 |
| 2 | NONO_HUMAN | Non-POU domain-containing octamer-binding protein OS=Homo sapiens GN=NONO PE=1 SV=4 | 195 | 7 | 54197 | 9.01 | 17.8 |
| 3 | RADI_HUMAN | Radixin OS=Homo sapiens GN=RDX PE=1 SV=1 | 156 | 7 | 68521 | 6.03 | 9.9 |
| 4 | THIL_HUMAN | Acetyl-CoA acetyltransferase. mitochondrial OS=Homo sapiens GN=ACAT1 PE=1 SV=1 | 139 | 5 | 45171 | 8.98 | 11.2 |
| 5 | MOES_HUMAN | Moesin OS=Homo sapiens GN=MSN PE=1 SV=3 | 119 | 5 | 67778 | 6.08 | 8 |
| 6 | ALBU_HUMAN | Serum albumin OS=Homo sapiens GN=ALB PE=1 SV=2 | 114 | 3 | 69321 | 5.92 | 5.6 |
| 7 | FILA2_HUMAN | Filaggrin-2 OS=Homo sapiens GN=FLG2 PE=1 SV=1 | 101 | 2 | 247928 | 8.45 | 0.8 |
| 8 | DCD_HUMAN | Dermcidin OS=Homo sapiens GN=DCD PE=1 SV=2 | 94 | 2 | 11277 | 6.08 | 17.3 |
| 9 | SF01_HUMAN | Splicing factor 1 OS=Homo sapiens GN=SF1 PE=1 SV=4 | 87 | 3 | 68286 | 9.07 | 5.8 |
| 10 | PSPC1_HUMAN | Paraspeckle component 1 OS=Homo sapiens GN=PSPC1 PE=1 SV=1 | 78 | 3 | 58706 | 6.26 | 5.9 |
| 11 | BASP1_HUMAN | Brain acid soluble protein 1 OS=Homo sapiens GN=BASP1 PE=1 SV=2 | 56 | 2 | 22680 | 4.64 | 12.3 |
| 12 | G3P_HUMAN | Glyceraldehyde-3-phosphate dehydrogenase OS=Homo sapiens GN=GAPDH PE=1 SV=3 | 56 | 2 | 36030 | 8.57 | 4.8 |
| 13 | IF2A_HUMAN | Eukaryotic translation initiation factor 2 subunit 1 OS=Homo sapiens GN=EIF2S1 PE=1 SV=3 | 49 | 1 | 36089 | 5.02 | 3.8 |
| 14 | SYYM_HUMAN | Tyrosine--tRNA ligase. mitochondrial OS=Homo sapiens GN=YARS2 PE=1 SV=2 | 43 | 1 | 53166 | 9.07 | 1.5 |
| 15 | BUB3_HUMAN | Mitotic checkpoint protein BUB3 OS=Homo sapiens GN=BUB3 PE=1 SV=1 | 43 | 1 | 37131 | 6.36 | 3 |
| 16 | KLK11_HUMAN | Kallikrein-11 OS=Homo sapiens GN=KLK11 PE=1 SV=2 | 43 | 1 | 31039 | 9.23 | 2.5 |
| 17 | SFPQ_HUMAN | Splicing factor. proline- and glutamine-rich OS=Homo sapiens GN=SFPQ PE=1 SV=2 | 37 | 1 | 76102 | 9.45 | 1.1 |
| 18 | SND1_HUMAN | Staphylococcal nuclease domain-containing protein | 36 | 1 | 101934 | 6.74 | 1.8 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | 1 OS=Homo sapiens GN=SND1 PE=1 SV=1 | | | | | |
| 19 | SHRM3_HUMAN | Protein Shroom3 OS=Homo sapiens GN=SHROOM3 PE=1 SV=2 | 35 | 1 | 216724 | 7.87 | 0.5 |
| 20 | SBSN_HUMAN | Suprabasin OS=Homo sapiens GN=SBSN PE=1 SV=2 | 34 | 1 | 60505 | 6.5 | 3.1 |
| 21 | IREB2_HUMAN | Iron-responsive element-binding protein 2 OS=Homo sapiens GN=IREB2 PE=1 SV=3 | 33 | 1 | 104978 | 6.62 | 0.7 |
| 22 | ARGI1_HUMAN | Arginase-1 OS=Homo sapiens GN=ARG1 PE=1 SV=2 | 32 | 1 | 34713 | 6.72 | 2.2 |
| 23 | FABP5_HUMAN | Fatty acid-binding protein. epidermal OS=Homo sapiens GN=FABP5 PE=1 SV=3 | 31 | 1 | 15155 | 6.6 | 7.4 |
| 24 | TMPSD_HUMAN | Transmembrane protease serine 13 OS=Homo sapiens GN=TMPRSS13 PE=2 SV=4 | 30 | 1 | 63113 | 8.96 | 1.4 |
| 25 | IGHA1_HUMAN | Ig alpha-1 chain C region OS=Homo sapiens GN=IGHA1 PE=1 SV=2 | 30 | 1 | 37631 | 6.08 | 2.5 |
| 26 | HMGN5_HUMAN | High mobility group nucleosome-binding domain-containing protein 5 OS=Homo sapiens GN=HMGN5 PE=1 SV=1 | 29 | 1 | 31506 | 4.5 | 3.2 |
| 27 | CYTA_HUMAN | Cystatin-A OS=Homo sapiens GN=CSTA PE=1 SV=1 | 28 | 1 | 11000 | 5.38 | 8.2 |
| 28 | GLSL_HUMAN | Glutaminase liver isoform. mitochondrial OS=Homo sapiens GN=GLS2 PE=1 SV=2 | 27 | 1 | 66280 | 6.9 | 1.2 |
| 29 | THIM_HUMAN | 3-ketoacyl-CoA thiolase. mitochondrial OS=Homo sapiens GN=ACAA2 PE=1 SV=2 | 26 | 1 | 41898 | 8.32 | 3.3 |
| 30 | SF3B1_HUMAN | Splicing factor 3B subunit 1 OS=Homo sapiens GN=SF3B1 PE=1 SV=3 | 26 | 1 | 145738 | 6.65 | 0.9 |
| 31 | PARP1_HUMAN | Poly [ADP-ribose] polymerase 1 OS=Homo sapiens GN=PARP1 PE=1 SV=4 | 26 | 1 | 113012 | 8.99 | 1 |

OS: organism name; GN: Gene Name; PE: Protein Existence; SV: Sequence Version.