Protein information sheet

Gene name	**HNRNPA1**
Uniprot ID	P09651
Region	M1 – F320

Description	heterogeneous nuclear ribonucleoprotein A1
Synonyms	ALS19, ALS20, HNRPA1, HNRPA1L3, IBMPFD3, UP 1, hnRNP A1, hnRNP-A1
Construct ID	HNRNPA1A-c000
Protein sequence	*SM*SKSESPKEPEQLRKLFIGGLSFETTDESLRSHFEQWGTLTDCVVMRDPNTK RSRGFGFVTYATVEEVDAAMNARPHKVDGRVVEPKRAVSREDSQRPGAHLTVK KIFVGGIKEDTEEHHLRDYFEQYGKIEVIEIMTDRGSGKKRGFAFVTFDDHDS VDKIVIQKYHTVNGHNCEVRKALSKQEMASASSSQRGRSGSGNFGGGRGGGFG GNDNFGRGGNFSGRGGFGGSRGGGGYGGSGDGYNGFGNDGSNFGGGGSYNDFG NYNNQSSNFGPMKGGNFGGRSSGPYGGGGQYFAKPRNQGGYGGSSSSSSYGSG RRF*SSKGGYGLNDIFEAQKIEWHE*

Protein mass	**with biotin, + 226.1**	**Protein pI**	**Extinction (M^{-1} cm^{-1})**
36673.9	38899.9	9.00	30370

Expression host	E. coli BL21(DE3)-R3-BirA
Expression medium	Terrific broth (TB) with 50 µg/ml kanamycin, 50 µg/ml Streptomycin and Chloramphenicol (34 µg/ml) included in starter cultures; chloramphenicol omitted in final growth. Biotin (10 mM stock): dissolve 24mg of d-biotin in 10 ml of warm (microwaved) 10mM bicine buffer (pH 8.3) and sterilize of the solution with a syringe and a 0.2 micron filter.
Transformation and storage	Cells were transformed with the expression plasmid and selected on LB-agar plates containing kanamycin, streptomycin and chloramphenicol. Several colonies were picked from the overnight transformation and used to inoculate cultures in LB supplemented with KAN/Cam/Str which were grown overnight and stored as glycerol (15% v/v) stock at -80°C.
	*The following protocol has been optimised for HNRNPA1. Previous purification attempts have found this protein to be highly susceptible to proteolysis during expression and purification resulting in protein degradation. Mass spectrometry characterisation of HNRNPA1 has identified the intact expected protein mass along with three cleavage sites resulting in sequence truncations. Additional the cleaved polypeptide sequences were identified using PAWS (Protein Analysis Worksheet). To minimise proteolysis the standard protocol was optimised; Expression conditions, induction time was reduced to 3 hours at 37°C. Additional Trypsin/Serine protease inhibitors were added to all lysis buffers,

Opher Gileadi, Structural Genomics Consortium
opher.gileadi@sgc.ox.ac.uk

Protein information sheet

Expression and *in vivo* biotinylation	Benzamidine HCL and Phenylmethanesulfonylfluoride (PMSF). Other steps favourable to reducing proteolysis during purification, keep cells and buffers on ice particularly during cell lysis, work as quickly as possible with the protein to avoid prolonged storage periods during purification steps. 10ml starter cultures in LB supplemented with KAN/Cam/Str were grown overnight at 37°C with shaking at 250rpm. Next day, 10 ml was used to inoculate 1L of TB in a 2LUltraYield flask supplemented with Streptomycin 50µg/ml and Kanamycin 50µg/ml. **Note:** the medium need not contain chloramphenicol at this step; the pRARE2 plasmid is not lost during the growth process and also ensures a more predictable growth. Seal flask with an adhesive air porous film during cell growth. Culture was grown in a 37°C incubator with shaking at 200rpm. Cell growth was monitored until mid-log phase was reached and cell density OD_{600nm} was approximately 2.00 – 3.00 in TB. Incubator temperature was maintained at 37ºC. Protein expression was induced by the addition of IPTG at 0.1mM final concentration and Biotin solution (10 ml/ 1L culture). Incubation of the culture was continued for 3 hours at 37°C. The same day, the cells were collected by centrifugation (15 min, 5,000 RPM in JLA-8.100 rotors). The supernatants were discarded and the cell pellets transferred into 50ml conical polypropylene tubes. Cell pellets were weighed, frozen and stored at -80°C.
Purification buffers	2x lysis buffer: 100 mM HEPES, pH 7.5, 1.0 M NaCl, 10% v/v glycerol, 1mM TCEP, 10mM imidazole. Protease inhibitors (1:200 Merck cocktail IV) were added before use. 1x lysis buffer: 50 mM HEPES, pH 7.5, 0.5 M NaCl, 5% v/v glycerol, 1 mM TCEP, 10 mM imidazole Wash buffer: 50 mM HEPES, pH 7.5, 0.5 M NaCl, 5% v/v glycerol, 1 mM TCEP, 40 mM imidazole ***To all *Ni^{2+}-IDA* buffers 1mM PMSF (100mM Stock) and 2mM Benzamidine (200mM Stock) was also added for this purification in addition to 1:200 Merck cocktail IV. Add to lysis buffer immediately prior to use** Elution buffers: 50 mM Hepes pH 7.5, 500mM NaCl, 5% Glycerol, 1mM TCEP, 100mM, 150mM, 250 mM, 500 mM Imidazole, respectively GF buffer: 50 mM HEPES, pH 7.5, 0.5M NaCl, 5% v/v glycerol, 1 mM TCEP IEX Low Salt Buffer: 50 mM HEPES, pH 7.5, 50 mM NaCl, 5% v/v glycerol, 1 mM TCEP IEX High Salt Buffer: 50 mM HEPES, pH 7.5, 1M NaCl, 5% v/v glycerol, 1 mM TCEP PEI 5%, pH 7.5: 5% polyethyleneimine (diluted from 50% stock, Sigma) and **titrated to pH 7.5 with HCl.**
Purification step 1: IMAC	**Cell Lysis:** Frozen cell pellets from -80°C were thawed by intermittently placing in a 37°C water bath, then transferred directly to ice. Cells were re-

Opher Gileadi, Structural Genomics Consortium
opher.gileadi@sgc.ox.ac.uk

	suspended in pre-chilled lysis buffer by vortex or pipetting gentle until cells were completely in suspension. 1 volume (v/w) of 2x lysis buffer was added, and the cells were thoroughly suspended. After further addition of 3 volumes (3 ml per 1g of pellet) of 1x lysis buffer the cells were broken by sonication. The lysate was then transferred to a 50ml Beckman polypropylene centrifuge tube. Cell debris and DNA from the lysate was precipitated by adding Polyethylenimine (PEI) to the tube from a 5% stock solution to a final concentration of 0.15% and incubated on ice for 15 – 30 mins for precipitation. Cell lysates were clarified by centrifugation using a JA25.50 rotor at 21,500rpm for 60 min at 4°C. The cleared supernatant was collected and filtered using a 0.45um acrodisc syringe filter and transferred to a fresh 50ml-conical polypropylene tube. **Ni^{2+}-IDA IMAC (column 1):** A 25ml Econo-Pac polypropylene column (BioRad) was packed with 5ml of Ni-Sepharose resin and equilibrated in 1x lysis buffer. The cell lysate was loaded onto the column and passed through resin at gravity flow. The unbound flow-through was collected, a wash step of 10 column volumes of wash buffer was included and collected as the wash fraction. Protein was eluted in five intermediate fractions with a step gradient of 5ml imidazole elution buffers; 50mM, 100mM, 150mM, 250mM and 500mM imidazole, respectively, each gradient eluted fraction was collected. IMAC fractions were analysed by SDS-PAGE gel for expected molecular weight and purity, the selected eluted fractions were pooled together. **Mass spectrometry characterisation:** Identity and biotinylation efficiency was verified by mass spectrometry, using a Quadrupole-Time of flight (Q-TOF) analyser, Biotin quantitation is measured by: MW of biotinylated protein = MW of sequenced Protein + 226 Da (biotinylation modification). An intact mass spectrometry was performed directly after the first IMAC, column 1. Proteolysis at this stage was minimal with ≥ 80% of intact biotinylated HNRNPA1. A second observed mass of 35411 was also identified, results from Tandem MS/MS were shown to determine the mass to be BIRA_ECOLI (UniProt: P06709). BirA ligase is known to sometimes co-purify with the target biotinylated protein (co-purified BirA may be separated from the target protein with an additional IEX chromatography step if separation cannot be achieved through SEC). Selected eluted fractions were pooled and prepared for removal of the histidine tag.
Step 2: Removal of His_6 tag and Ni^{2+}-IDA reverse binding	**Enzymatic removal of histidine tag:** Separate cleaved protein from any un-cleaved protein by passing over a Ni^{2+}-IDA column. Cleaved protein may be recovered in the flow-through fractions. Non-specific contaminating proteins with any un-cleaved protein bound to the column can be eluted with imidazole. Selected, pooled eluted fractions from Ni^{2+}-IDA IMAC were concentrated using a 30,000 Amicon-Ultra MWCO centrifugal concentrator to approximately 10mls in volume. Cleavage of the histidine tag was performed by the addition of TEV protease in a 1:20 w/w ratio to the pooled eluted protein fractions. The protein sample plus TEV was transferred to Snakeskin dialysis tubing (3,500 MWCO) and dialysed in 1L Gel Filtration

Opher Gileadi, Structural Genomics Consortium
opher.gileadi@sgc.ox.ac.uk

	buffer overnight at 4°C on a magnetic stirrer plate for digestion. Next day, the protein was applied to a Ni^{2+}-IDA reverse-binding column.
	Ni^{2+}-IDA Rebinding column (Column 2): To a 25ml Econo-Pac polypropylene column (BioRad), 2ml of packed Ni-Sepharose resin was added, resin was washed in 3 column volumes of distilled deionized water then equilibrated in 5 column volumes of gel filtration buffer. The overnight protein plus TEV sample was loaded and passed through the column by gravity flow. The flow through was collected, to wash any remaining unbound protein from the column a 2x wash step of 2 column volumes of gel filtration buffer was included. Any bound protein was eluted with 2 column volume washes of 40mM and 500mM imidazole elution buffer. Collected fractions were analysed for verification of cleaved protein and purity by SDS-PAGE gel and mass spectrometry characterisation.
	Observation of SDS-PAGE gel analysis identified two protein bands after column 2 in the sample flow-through, GF buffer washes 1 and 2 and 40 mM Imidazole wash, plus some additional lower molecular weight protein bands. Mass spectrometry results verified the expected mass with biotinylation plus cleavage of the histidine tag for HNRNPA1 with proteolysis remaining ≤ 80%, a second observed mass for BirA ligase was also observed in the protein sample.
Step 3: SEC: Gel filtration	**Size Exclusion Chromatography (SEC) (Column 3):** Verified protein from rebinding column 2 was concentrated to required volume for gel filtration (<5 ml), using an Amicon-Ultra 30,000 MWCO centrifugal concentrator. The GF column (Superdex S200 16/60 HiLoad, 120ml on an ÄKTA Express chromatography system) was equilibrated in gel filtration buffer. The concentrated protein sample was loaded onto the GF column for further purification; protein was fractionated on the column in GF buffer at 1ml/min and 1ml fractions were collected at the A280 peaks.
	Four protein peaks were observed after evaluation of the GF chromatogram. All peak fractions were analysed by SDS-PAGE gel and mass spectrometry. Results were shown as; peak 1, biotinylated HNRNPA1, dimer or possible tetramer only; peak 2, biotinylated HNRNPA1 and BirA ligase protein; peak 3 and 4, protein proteolysis fragments. GF fractions of peak 1 and 2 were pooled and an additional chromatography IEX column was performed for further separation of BirA ligase and proteolysis fragments.
Step 4: IEX	**Cation Ion Exchange Chromatography (IEX) (Column 4):** The IEX column (MonoS 15/50GL, 1ml on an ÄKTA Pure chromatography system) was equilibrated in low salt buffer. The selected peak fractions from SEC (GF) were exchanged by dilution into a low salt buffer and loaded onto the column using a super loop The sample load was washed through the column, collecting any unbound protein material. Protein was eluted using a 100% increasing gradient of 50 mM to 1M NaCl at 0.6ml/min over 30CV. Peak fractions were collected and analysed by SDS-PAGE gel and mass spectrometry.
	Protein purity was improved after additional IEX step with separation of degradation products and elimination of bound contaminating non-specific proteins and BirA ligase with protein purity increasing to >90%. Analysis of

Opher Gileadi, Structural Genomics Consortium
opher.gileadi@sgc.ox.ac.uk

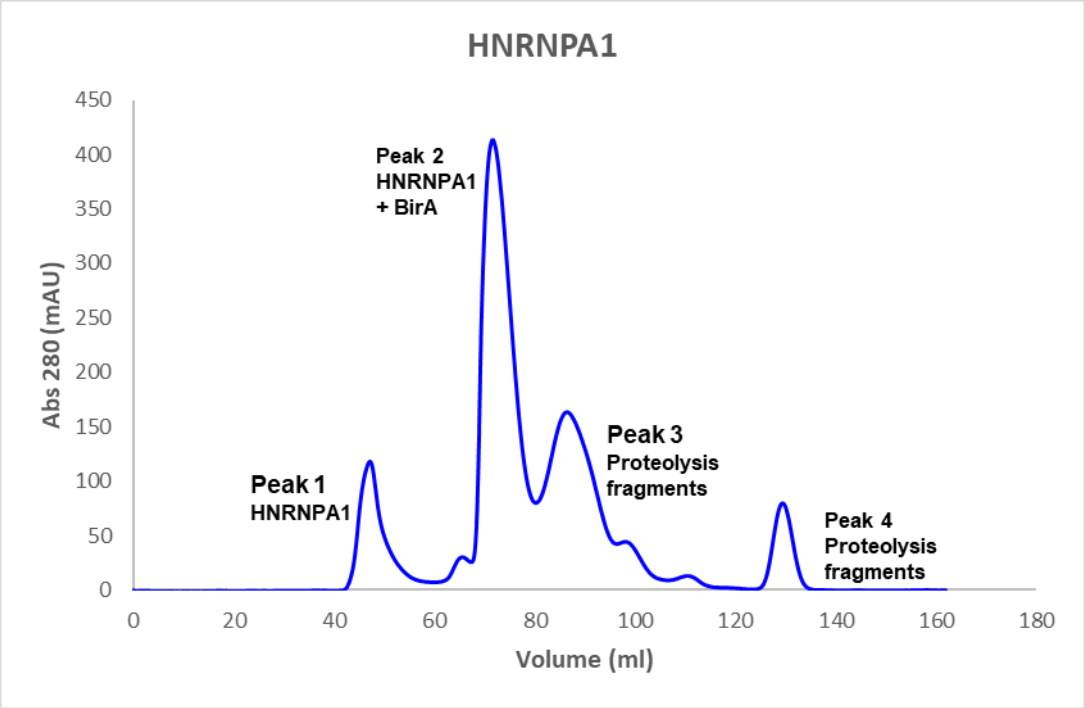
Protein information sheet

Step 5: Protein storage	IEX peaks found peak 3 protein fractions to contain intact biotinylated HNRNPA1 of the correct mass without any post-translational modifications. **Protein concentration mg/ml and storage:** Protein was concentrated using a 30,000 Amicon-Ultra MWCO centrifugal concentrator and final protein concentration was determined from the absorbance at 280nm, measured using a nanodrop spectrometer (Nanodrop ND-1000, NanoDrop Technologies). Final protein concentration was measured at 1mg/ml with a final volume of 2ml. Protein was transferred into 200 µl aliquots, flash-frozen in liquid nitrogen and stored at -80°C.

Opher Gileadi, Structural Genomics Consortium
opher.gileadi@sgc.ox.ac.uk

Protein information sheet

Gel Filtration Chromatogram (Superdex S200)

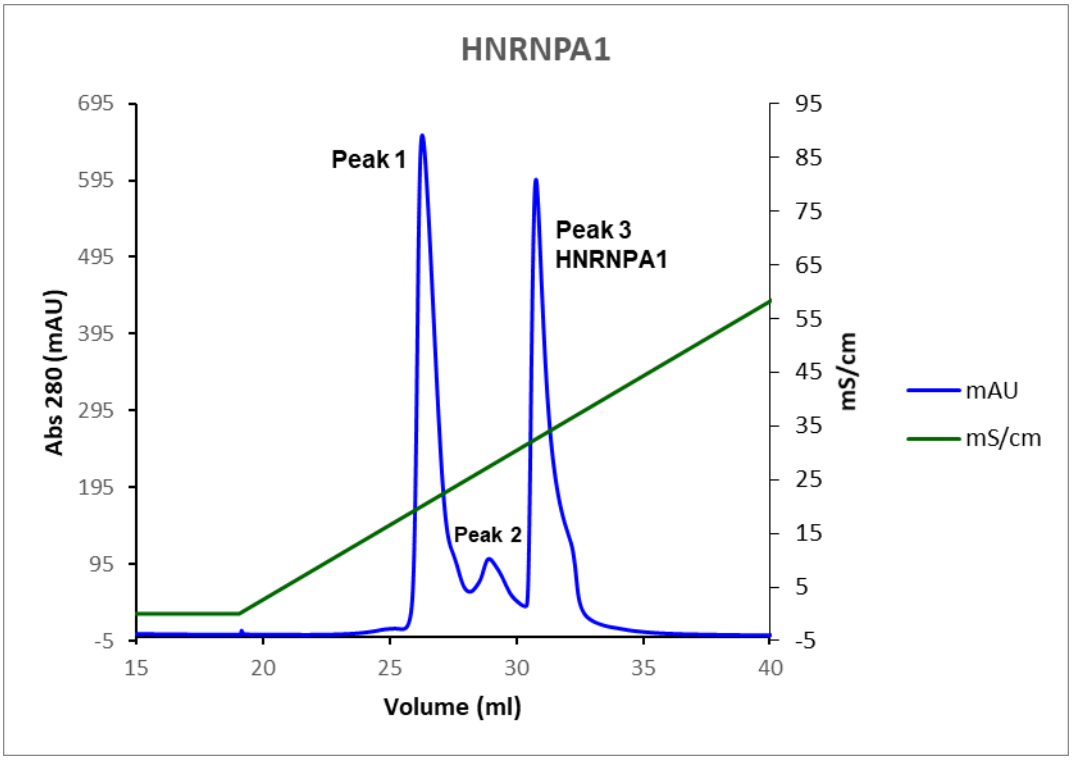


HNRNPA1A-c000, Maximum peak elution time (Peak 2) was 74.05ml with absorbance at 280nm of 412.90 mAU. SEC analysis validates protein is a Dimer / tetramer.

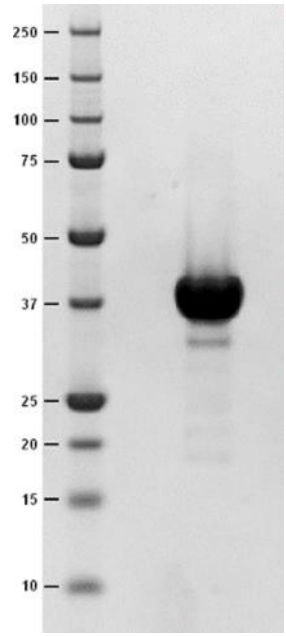
Opher Gileadi, Structural Genomics Consortium
opher.gileadi@sgc.ox.ac.uk

Protein information sheet

Cation Ion Exchange Chromatogram (MonoS)



Final Sample SDS-PAGE



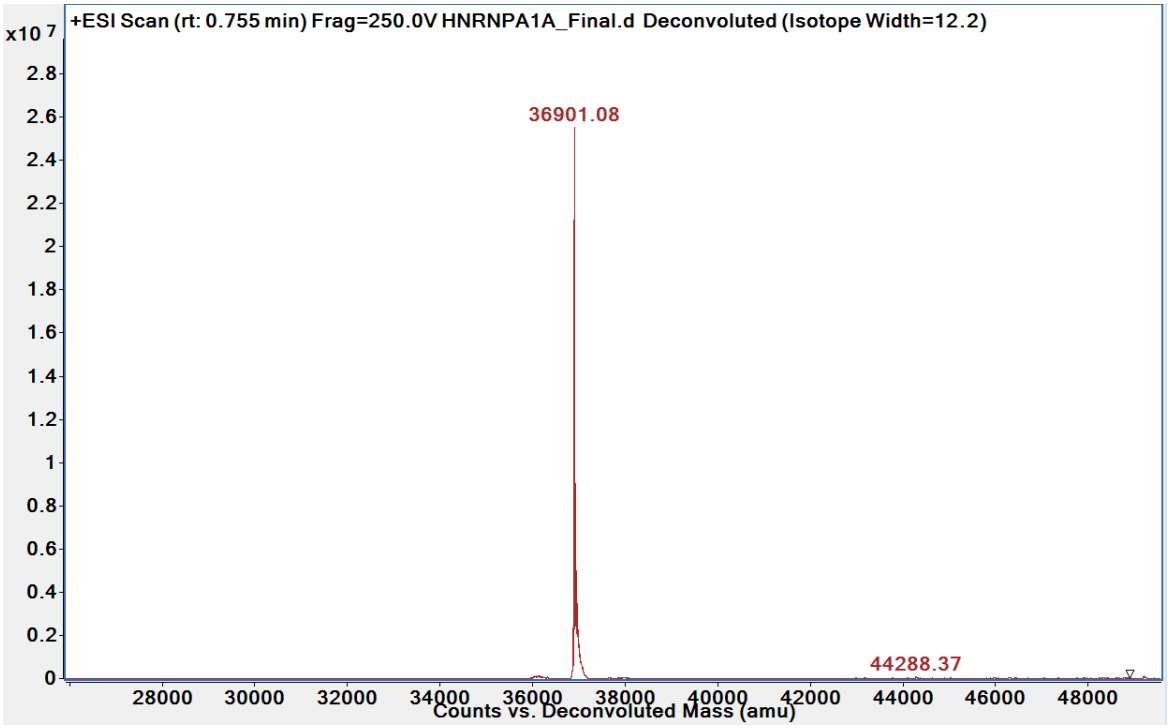
Protein	Expected mass + Biotinylation (Da)	Final concentration (mg/ml)
HNRNPA1A-c000	38899.9	1

Opher Gileadi, Structural Genomics Consortium
opher.gileadi@sgc.ox.ac.uk

Protein information sheet

Mass Spectrometry characterisation

Protein	Vector	Expected mass cleaved + Biotinylation (Da)	Observed mass	Comments
HNRNPA1A-c000	pNIC-Bio3	38899.9	39901.08	100% Biotinylated



Opher Gileadi, Structural Genomics Consortium
opher.gileadi@sgc.ox.ac.uk