Supplementary data

Protein Hit1, a novel box C/D snoRNP assembly factor, controls cellular concentration of protein Rsa1p by direct interaction

Benjamin Rothé, Jean-Michel Saliou, Marc Quinternet, Régis Back, Decebal Tiotiu, Clémence Jacquemin, Christine Loegler, Florence Schlotter, Vlad Peña, Kelvin Eckert, Solange Moréra, Alain Van Dorsselaer, Christiane Branlant, Séverine Massenet, Sarah Sanglier-Cianférani, Xavier Manival, and Bruno Charpentier

Supplemental Materials and Methods

Mass spectrometry-based proteomic analysis

In gel digestion of SDS-PAGE gel slices was performed as previously described (Miguet et al. 2009). The resulting peptide extracts were analyzed by nanoLC-MS/MS on a nanoACQUITY Ultra-Performance-LC (UPLC, Waters, Milford, UK) coupled to a SYNAPT HDMS G1 hybrid quadrupole-time-of-flight mass spectrometer (Waters, Milford, UK) equipped with a Z-spray ion source and a lock mass system. Peptide mixtures were loaded on a Symmetry C18 trap precolumn (180 µm inner diameter × 20 mm, particle size 5 µm; Waters, Milford, UK) and peptides were separated on a ACQUITY UPLC BEH130 C18 column, 75 µm x 200 mm, 1.7 µm particle size (Waters, Milford, UK). The solvent system consisted of 0.1% formic acid in water (solvent A) and 0.1% formic acid in acetonitrile (solvent B). Trapping was performed during 3 min at 5 µL/min with 99% of solvent A and 1% of solvent B. Elution was performed at a flow rate of 400 nL/min, using a 1−50 % gradient of solvent B over 35 min at 45 °C followed by 65% solvent B over 5 min. The mass spectrometer parameters were set as follows: capillary voltage 3500 V and cone voltage 35 V. Mass calibration of the TOF was achieved using [Glu1]-fibrinopeptide B (GFP) over the *m/z* range 50−2000 in the positive ion mode. Online correction of this calibration was achieved using lock-mass on product ions derived from the [Glu1]-fibrinopeptide B (GFP). The ion $(M+2H)^{2+}$ at *m/z* 785.8426 was used

to calibrate MS data and the fragment ion $(M+H)^+$ at *m/z* 684.3469 was used to calibrate MS/MS data during the analysis. For tandem MS experiments, the system was operated with automatic switching between MS and MS/MS modes. The four most abundant peptides, preferably doubly and triply charged ions, were selected from each MS spectrum for further isolation and CID fragmentation. MS/MS fragmentation was performed in the collision cell using argon as collision gas. The complete system was fully controlled by MassLynx 4.1 (SCN 566, Waters, Milford, UK). Raw data collected during nanoLC-MS/MS analyses were processed and converted with ProteinLynx Browser 2.3 (23, Waters, Milford, UK) into *.pkl peak list format. The MS/MS data were submitted to the Mascot search engine (Matrix Science, London, UK) installed in its 2.2.04 version on a local server. Searches were performed against a composite target-decoy database including both forward (target) and reverse (decoy) SwissProt protein sequences (21 January 2010, 5 1028424 total entries). Searches were performed with a tolerance on mass measurements of 0.2 Da for precursor and 25 ppm for fragment ions. Carbamidomethylation of cysteine residues, oxidation of methionine residues and acetylation of protein N-terminal residues were searched as variable modifications. A single missed cleavage was allowed. Scaffold V3_00_01 (Proteome Science, Portland, USA) was used for identification validation and false positive rate estimation for protein identification. For each sample, peptides were filtered out according to the cutoff set for proteins hits with two or more peptides (Mascot Ion Score >20 and Mascot Ion Score - Mascot Identity Score >0) and no false positive proteins were identified. Further data treatment was performed with Microsoft Excel.

Supplemental references

Miguet L, Bechade G, Fornecker L, Zink E, Felden C, Gervais C, Herbrecht R, Van Dorsselaer A, Mauvieux L, Sanglier-Cianferani S. 2009. Proteomic analysis of malignant B-

cell derived microparticles reveals CD148 as a potentially useful antigenic biomarker for mantle cell lymphoma diagnosis. *J Proteome Res* **8**: 3346-3354.

Supplemental Figures & Tables

Table S1. Structural statistics obtained on $Rsa1p_{317–352}$–$Hit1p_{70–164}$. Constraints used for RECOORD scripts and statistics on the 20 NMR structures with the lowest final energies are listed.

	$Rsa1p_{317-352}$/$Hit1p_{70-164}$
NMR distances and dihedral constraints	
Distance constraints	
Total NOE	3378
Intra-residue	955
Inter-residue	2423
Sequential ($\|i - j\| = 1$)	810
Medium-range ($\|i - j\| < 4$)	868
Long-range ($\|i - j\| > 5$)	745
Total dihedral angle restraints	
ϕ	95
ψ	95
Structure statistics	
Violation occurences	
Distance constraints (> 0.5 Å)	0
Dihedral angle constraints (> 5º)	0
Deviations from idealized geometry	
Bond lengths (Å)	0.0032 + /- 0.0001
Bond angles (º)	0.465 +/- 0.013
Impropers (º)	1.129 + /- 0.041
R.M.S.D. on secondary structures (Å)	
Heavy	1.03 + /- 0.12
Backbone	0.53 + /- 0.09
Ramachandran statistics	
Residues in most favoured regions	87.7 %
Residues in additional allowed regions	10.5 %
Residues in generously allowed regions	0.8 %
Residues in disallowed regions	1.0 %
$Rsa1p_{317-352}$surface (Å^2)	4341 +/- 71
$Hit1p_{70-164}$ surface (Å^2)	11836 +/- 97
Complexe surface (Å^2)	8569 +/- 165
Complex buried surface area (Å^2)	3267 +/- 128

Table S2. DNA oligonucleotides used for rRNAs, snoRNAs and UsnRNAs detection.

Experiments	Oligonucleotides	Sequences	Target RNAs
pre-rRNA/rRNA Northern-blot	p1	CTTGCGACCGGCTATTC	5'ETS pre-rRNA
	p2	CATGGCTTAATCTTTGAGAC	18S rRNA
	p3	AAAGCTCTCATGCTCTTGCC	ITS1 pre-rRNA
	p4	CGCCTAGACGCTCTCTTCTTA	ITS2 pre-rRNA
	p5	CTCCGCTTATTGATATGC	25S rRNA
snoRNA/UsnRNA Northern-blot	NB-U3	CCGTCAGACTGTTCA	C/D snoRNA U3
	NB-U14	TCACTCAGACATCCTAG	C/D snoRNA U14
	NB-snR13	AGCTTGAGTTTTTCCACACC	C/D snoRNA snR13
	NB-snR45	GTATTTGCGCTCACGTTGC	C/D snoRNA snR45
	NB-snR63	TTATGTTGGCCACTCATCAC	C/D snoRNA snR63
	NB-snR70	TGACCAATCATCAATTCTCC	C/D snoRNA snR70
	NB-snR190	CGAGGAAAGAAGAGACACCATTATC	C/D snoRNA snR190
	NB-snR10	CCTTGCAACGGTCCTCATCCGGG	H/ACA snoRNA snR10
	NB-snR30	GTCCGAAGCGCCATCTAG	H/ACA snoRNA snR30
	NB-snR32	GGGTCAAGCTTTGCAATTCT	H/ACA snoRNA snR32
	NB-snR42	TCAAACAATAGGCTCCCTAAAGCATCACAA	H/ACA snoRNA snR42
	NB-snR44	GGTAATGGATATAGTAGCTA	H/ACA snoRNA snR44
	NB-U1	GACCAAGGAGTTTGCATCAATGA	UsnRNA U1
	NB-U4	TAAATTTCAACCAGGGGAAACACAATCTCGGACGAA	UsnRNA U4
	NB-U6	CAGGGGAACTGCTG	UsnRNA U6
RT-PCR	OG-Rev	CAGTGCAGGGTCCGAGGTATTCTGAAAACCAAACCTTTGG	pre-U3 (+exogenous anchor)
	OG-5'PCRa	TCTGTGTCGACGTACTTCA	pre-U3 (Exon 1)
	OG-5'PCRb	AGTCTTAGGTACTAGAGTTT	pre-U3 (Intron)
	OG-3'PCR	CAGTGCAGGGTCCGAGGTATT	pre-U3 (exogenous anchor)

Figure S1. Analysis of functional domains in the Hit1 protein. Various fragments of Hit1p were expressed in yeast strain Δ*HIT1* and tested for complementation of the growth defect.

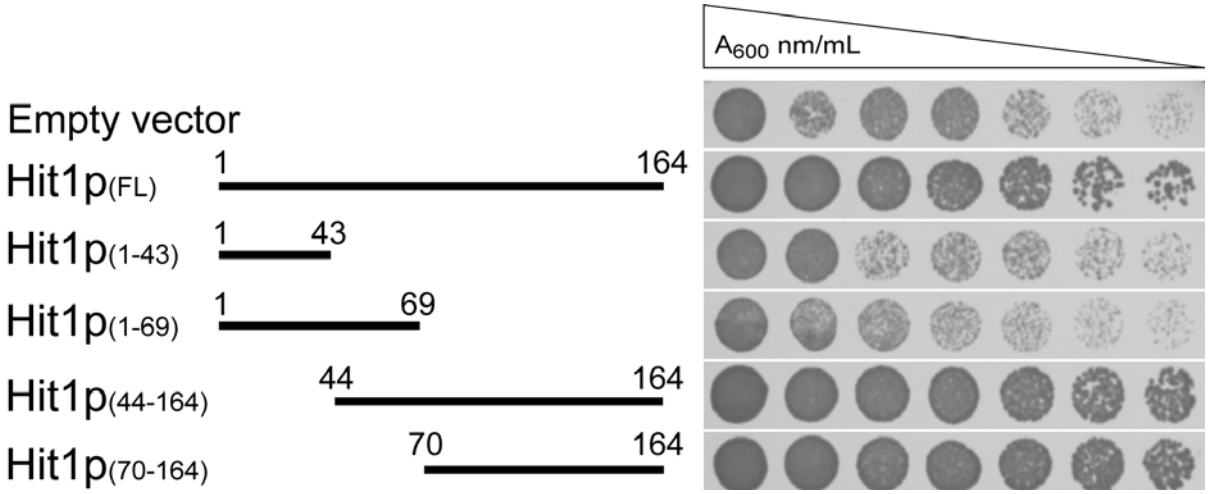

Figure S2. Comparison of HSQC ^{1}H-^{15}N spectra recorded at 293K on Rsa1p$_{317–375}$-Hit1p$_{44–164}$ (red), Rsa1p$_{317–375}$–Hit1p$_{70–164}$ (green) and Rsa1p$_{317–352}$–Hit1p$_{70–164}$. All complexes were obtained upon co-expression of the protein partners in *E. coli* and are ^{15}N-labelled. As the size of the complex decreases, the three spectra remain superimposable for the majority of resonances except those located in the middle portion of the HSQC map. Using 3D NMR spectra, the latter resonances have been assigned to the C-terminal tail of Rsa1p$_{317–375}$ (i.e. fragment 353−375) and to the N-terminal tail of Hit1p$_{44–164}$ (i.e. fragment 44−69). This observation indicates that these tails are not involved in stabilization of the complex.

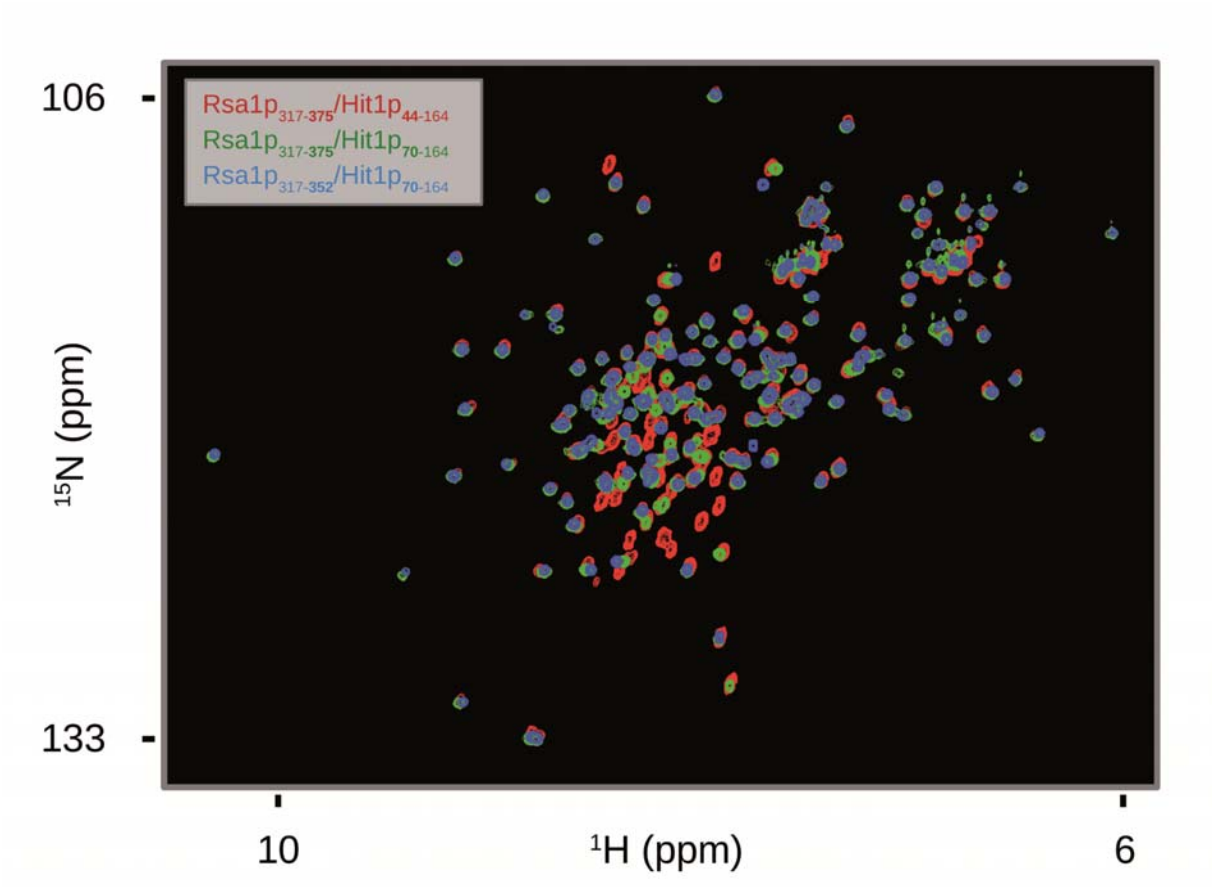

Figure S3. ^{1}H-^{15}N heteronuclear nOe (blue circles) and Het-SOFAST (red circles) ratios are plotted vs. the primary sequences of $Hit1p_{70–164}$ and $Rsa1p_{317–352}$. Secondary structures are indicated using grey squares.