nature structural & molecular biology

Structure and RNA-binding properties of the Not1–Not2–Not5 module of the yeast Ccr4–Not complex

Varun Bhaskar[1], Vladimir Roudko[2,3], Jérôme Basquin[1], Kundan Sharma[4], Henning Urlaub[4], Bertrand Séraphin[2,3] & Elena Conti[1]

The Ccr4–Not complex is involved in several aspects of gene expression, including mRNA decay, translational repression and transcription. We determined the 2.8-Å-resolution crystal structure of a 120-kDa core complex of the *Saccharomyces cerevisiae* Not module comprising the C-terminal arm of Not1, Not2 and Not5. Not1 is a HEAT-repeat scaffold. Not2 and Not5 have extended regions that wrap around Not1 and around their globular domains, the Not boxes. The Not boxes resemble Sm folds and interact with each other with a noncanonical dimerization surface. Disruption of the interactions within the ternary complex has severe effects on growth *in vivo*. The ternary complex forms a composite surface that binds poly(U) RNA *in vitro*, with a site at the Not5 Not box. The results suggest that the Not module forms a versatile platform for macromolecular interactions.

The Ccr4–Not complex is a large assembly that regulates eukaryotic gene expression at multiple levels. The best-studied function of Ccr4–Not relates to its action as the major deadenylase involved in shortening the poly(A) tail of cellular mRNAs in the cytoplasm (reviewed in ref. 1). Deadenylation by Ccr4–Not is a key step in the constitutive and regulated turnover of mRNAs[2,3]. Ccr4–Not can also be targeted to *cis*-acting elements in the 3′ untranslated region (UTR) of specific transcripts to accelerate their decay (for example, in the case of ARE-containing mRNAs)[4,5] or to mediate microRNA-dependent repression[6–8] or translational repression (examples in refs. 9,10). In addition, Ccr4–Not has been implicated in transcription initiation and elongation in the nucleus as well as in ubiquitylation (reviewed in refs. 11,12). The nuclear and cytoplasmic functions of Ccr4–Not have long been thought of as disconnected. However, recent evidence is converging on the functional coupling between mRNA synthesis and degradation[13].

Ccr4–Not contains several evolutionarily conserved proteins (Not1, Caf1 (also called Pop2), Not2, Not3 or Not5 and Caf40) that are constitutive components of the complex in all species examined to date (yeast[14,15], humans[16,17], flies[18,19] and trypanosoma[20]). Other bona fide subunits of Ccr4–Not are peripheral (for example, Ccr4 and Not4)[16,18,21] and/or species specific[15,19,22,23]. Variants of the core complex are likely to exist, as homologs are present both in yeast (Not3 and Not5)[24] and humans (for Caf1 and for Ccr4)[16,17]. The core complex is built around Not1, a large scaffold protein of ~240 kDa (refs. 21,25). The N-terminal half of Not1 associates with the Caf1 and Ccr4 RNases and is involved in the formation of the deadenylase module of the complex[21,26]. The C-terminal half of Not1 binds Not2, Not3, Not4 and Not5 to form the so-called Not module[16,21,27]. Synthetic lethality between the yeast deadenylase subunits and Not subunits suggests that they have separate or only partially overlapping functions[21]. The deadenylase module is connected to the cytoplasmic activities of Ccr4–Not (reviewed in refs. 1,11) and has been studied at the structural and functional level[28,29]. How the Not module is structured and how it functions are far less clear (reviewed in ref. 30).

Genetic and biochemical studies have shown that Not2, Not3 and Not5 are closely associated[19,21]. *S. cerevisiae* Not3 and Not5 are currently thought of as paralogous proteins[30]. Yeast Not5 is reported to be crucial for vegetative growth, whereas Not3 deletion has milder phenotypes[24]. The only Not3 and Not5 homolog in metazoans (known as Not3) is essential in mice for embryonic development and for control of heart function[31] and metabolism[32] in adults. In metazoans, Not2 is believed to recruit Not3 into the complex[17,19], to be important for the integrity of Ccr4–Not[33,34] and to act as a repressor of promoter activity in the nucleus[35]. In yeast, Not2 and Not5 have been reported to interact with components of the transcription machinery, specifically with subunits of TFIID[36–38] and SAGA[33]. In addition to data pointing to connections with transcription (reviewed in refs. 11,12), the Not module has also been implicated in mRNA-decay pathways in the cytoplasm[18,39]. To shed light on how the Not module can mediate these different functions, we have determined the structure and biochemical properties of a core complex from *S. cerevisiae*.

RESULTS

Structure determination of a yeast Not1–Not2–Not5 complex

Yeast two-hybrid and coimmunoprecipitation assays have shown that Not1 (2,108 residues in *S. cerevisiae*) binds Not2, Not3 and Not5 in a region that spans approximately the last 700 residues[19,21,25] (**Fig. 1a**).

[1]Department of Structural Cell Biology, Max Planck Institute of Biochemistry (MPI Biochemistry), Martinsried, Germany. [2]Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC), Centre National de Recherche Scientifique (CNRS) UMR 7104, Institut National de Santé et de Recherche Médicale (INSERM), U964, Illkirch, France. [3]Université de Strasbourg, Illkirch, France. [4]Max Planck Institute of Biophysical Chemistry, Göttingen, Germany. Correspondence should be addressed to E.C. (conti@biochem.mpg.de).

Received 21 May; accepted 5 September; published online 13 October 2013; doi:10.1038/nsmb.2686

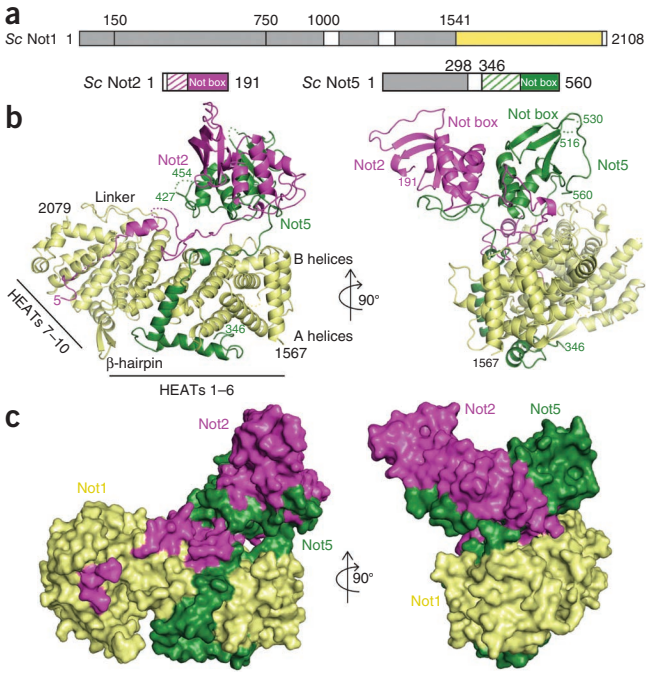

Figure 1 Structure of a yeast Not1–Not2–Not5 core complex. (**a**) Schematic representation of the domain organization of *S. cerevisiae* (*Sc*) Not1, Not2 and Not5. Color-filled rectangles indicate globular domains present in the crystal structure (yellow, Not1; magenta, Not2; green, Not5). Dashed rectangles indicate low-complexity regions of the molecules with ordered electron density. Gray rectangles indicate globular domains either from previous structures[28] or predicted from sequence analysis. (**b**) Structure of the complex shown in cartoon representation in two orientations (right, front view of the Not boxes; left, side view). Not1 features are labeled in black. Disordered regions are shown as dotted lines. The N- and C-terminal residues are indicated. The labeled linker and β-hairpin refer to the HEAT 6–7 and the HEAT 7–8 inter-repeat loops. This and all other cartoon drawings were generated with PyMOL (http://www.pymol.org/). (**c**) Surface representations of the complex in the same orientations and colors as in **b**.

Not2 (191 residues in *S. cerevisiae*) is predicted to contain a poorly structured N-terminal region followed by a conserved domain known as the Not-box domain[35] (**Fig. 1a**). Not5 (560 residues) contains an N-terminal coiled-coil region followed by a low-complexity linker and a C-terminal Not-box domain[35] (**Fig. 1a**). *S. cerevisiae* Not3 has a similar domain architecture as does Not5, but it could not be expressed as full length in a soluble form (V.B. and E.C., unpublished observations). We purified and reconstituted a complex containing the last ~750 residues of Not1, full-length (FL) Not2 and FL Not5, subjected it to limited proteolysis and identified the core complex composed of Not1 residues 1541–2093, Not2 FL and Not5 residues 298–560 (hereafter defined as Not1c, Not2 and Not5c, respectively) (**Supplementary Fig. 1**).

The Not1c–Not2–Not5c complex yielded crystals diffracting to 2.8-Å resolution. We determined the structure by SAD, using crystals derivatized with mercury, and refined it to an R_{free} of 22.6% and R_{work} of 18.1% with good stereochemistry (**Table 1**). The two independent copies of the Not1c–Not2–Not5c complex present in the crystal asymmetric unit are virtually identical (superposing with an r.m.s. deviation of 0.85 Å over all Cα atoms). The final model includes Not1 residues 1567–2079, Not2 residues 5–191 (with the exception of two disordered segments at residues 14–29 and 44–48) and Not5 residues 346–560 (with the exception of two loops at residues 428–453 and 517–529) (**Fig. 1b** and **Supplementary Fig. 2**).

The C-terminal region of Not1 is a scaffold of HEAT repeats

The Not1c–Not2–Not5c structure is organized around Not1 (**Fig. 1b**). Not1c is built almost entirely of antiparallel α-helices, forming an elongated molecule of the approximate dimensions 85 Å × 35 Å × 40 Å. The topology of the secondary-structure elements in Not1c is typical of that observed in HEAT-repeat proteins. Canonical HEAT repeats are characterized by a helix A–turn–helix B motif and are arranged in tandem in an almost-parallel fashion, with a 15° rotation between consecutive repeats[40]. Multiple repeats typically give rise to superhelical structures with a convex layer formed by the A helices and a concave layer formed by the B helices. Not1c contains ten HEAT repeats, which can be grouped into two units.

The first unit, comprising HEATs 1–6 (residues 1567–1849), has a regular architecture, albeit one less curved than for canonical HEAT-repeat proteins (**Fig. 1b**). The second unit, comprising HEATs 7–10 (residues 1888–2058), also adopts a regular architecture, with the exception of a long β-hairpin connecting HEATs 7 and 8 and of an additional C-terminal helix (residues 2059–2079). This unit contains four of the five HEAT repeats characteristic of the middle domain of eukaryotic initiation factor 4G (MIF4G)[41] and can therefore be described as an MIF4G-like domain. The 40-residue linker connecting HEATs 6 and 7 wraps around both units and contributes to formation of the extensive hydrophobic core of Not1c. The two HEAT-repeat units pack against each other in a perpendicular fashion resulting in a T-shaped architecture (**Fig. 1b** and **Supplementary Fig. 3a**). Interestingly, the domain formed by residues 193–745 in the N-terminal

Table 1 Data collection and refinement statistics

	Native	Hg derivative
Data collection[a]		
Space group	$P2_1$	$P2_1$
Cell dimensions		
a, *b*, *c* (Å)	110.45, 109.17, 133.62	109.67, 106.19, 136.02
α, β, γ (°)	90, 94.7, 90	90, 94.0, 90
		Peak
Wavelength	1.00004	1.00606
Resolution (Å)	49.15–2.80 (2.95–2.80)	47.77–3.20 (3.37–3.20)
R_{merge}	6.50 (42.90)	16.70 (80.60)
$I / \sigma I$	17.30 (2.90)	10.60 (2.40)
Completeness (%)	99.50 (97)	100 (100)
Redundancy	4.80 (4.50)	6.90 (7.20)
Refinement		
Resolution (Å)	49.15–2.80 (2.83–2.80)	
No. reflections	77,882	51,653
R_{work} / R_{free}	0.1812 / 0.2258	
No. atoms	14,019	
Protein	13,978	
Ligand/ion	36	
Water	5	
B factors	67.00	
Protein	67.10	
Ligand/ion	62.20	
Water	39.90	
r.m.s. deviations		
Bond lengths (Å)	0.009	
Bond angles (°)	1.12	

[a]One native and one Hg-derivative crystal were used for data collection. Values in parentheses are for highest-resolution shell.

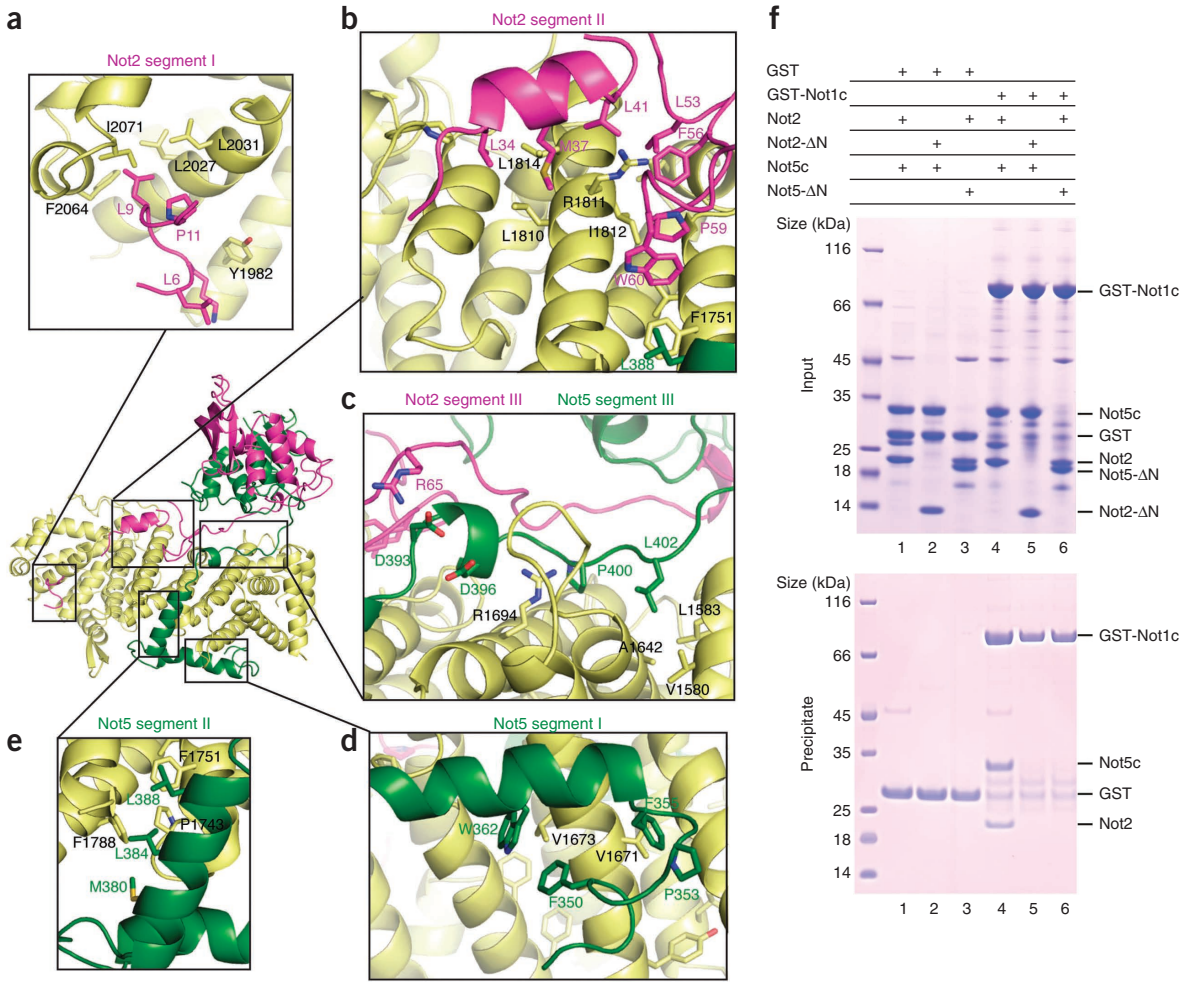

Figure 2 Not1 interacts with extended regions of Not2 and Not5. (**a**–**e**) Close-up views of the interactions of Not1 with Not2 and Not5 showing the three segments (I, II and III) of the N-terminal extensions of Not2 and Not5 that form the Not1-binding domains. The positions of the individual close-up views within the complex are indicated at center left. Interacting residues involved in evolutionarily conserved interactions are indicated and labeled (conservation shown in **Supplementary Fig. 2**). (**f**) Pulldown experiments of GST-tagged Not1c with untagged Not2, Not5c, Not2-ΔN and Not5-ΔN (ΔN refers to the removal of the N-terminal extension involved in Not1 binding). GST is shown as a control. Input samples (top) and samples precipitated on glutathione-agarose beads (bottom), analyzed on 4–12% Bis-Tris NuPAGE gel with MES running buffer, are shown. The proteins corresponding to the bands are indicated on the right.

arm of yeast Not1 is also formed by a MIF4G-like unit and a longer HEAT-repeat unit arranged perpendicularly to each other[28]. Although the relative orientations of the individual units differ in detail, the N-terminal and C-terminal arms of Not1 are built with remarkably similar principles (**Supplementary Fig. 3a**).

Extended regions of Not2 and Not5 wrap around Not1

Not2 and Not5 both contain a globular domain preceded by N-terminal extensions (**Fig. 1b**,**c**). In the N-terminal extensions, Not2 residues 5–75 and Not5 residues 346–404 mediate binding to Not1c, covering a distance of >100 Å each and burying a total surface of ~3,700 Å^2. The Not1-binding domain of Not2 can be described as composed of three segments (**Fig. 2**). The first segment (Not2 residues 5–13) binds the MIF4G-like unit of Not1c, mainly at the A helices of HEATs 9 and 10. Here, a conserved hydrophobic pocket of Not1 recognizes Not2 Leu9 (**Fig. 2a**), a conserved residue that has been shown to be functionally important in *in vivo* studies[33]. The second segment (Not2 residues 31–64) binds Not1c at the adjacent HEAT-repeat unit, zigzagging over the B helices of HEATs 4–6 (**Fig. 2b**). This segment of Not2 forms a short helix and a hairpin. The helix docks with hydrophobic residues on the conserved surface of HEAT 5 centered at Arg1811 and Leu1814. The hairpin wedges into another set of hydrophobic residues in a conserved groove at HEATs 4–5 (from Phe1751 to Ile1812). The third segment (Not2 residues 65–75) extends over the B helices of Not1c toward HEAT 3 (**Fig. 2c**).

The Not1-binding domain of Not5 wraps around HEATs 1–5 of Not1 (**Fig. 1b**) and can also be subdivided into three segments. The first segment (Not5 residues 346–373) contains an α-helix and binds the A helices of Not1c with apolar interactions (**Fig. 2d**). The second segment (Not5 residues 374–391) contains another α-helix and binds the edge of Not1c formed at the intrarepeat connections of HEATs 3–5 through hydrophobic interactions (**Fig. 2e**). The third segment (Not5 residues 392–404) stretches over the B helices of Not1c between HEATs 1–3, making both polar and apolar contacts (**Fig. 2c**). The third segment of Not5 flanks the third segment of Not2 and directly interacts with it through a salt bridge (between Asp393 and Arg65) (**Fig. 2c**). The structure suggests that the Not1-binding domains of Not2 and Not5 bind Not1c synergistically. We tested the effect of

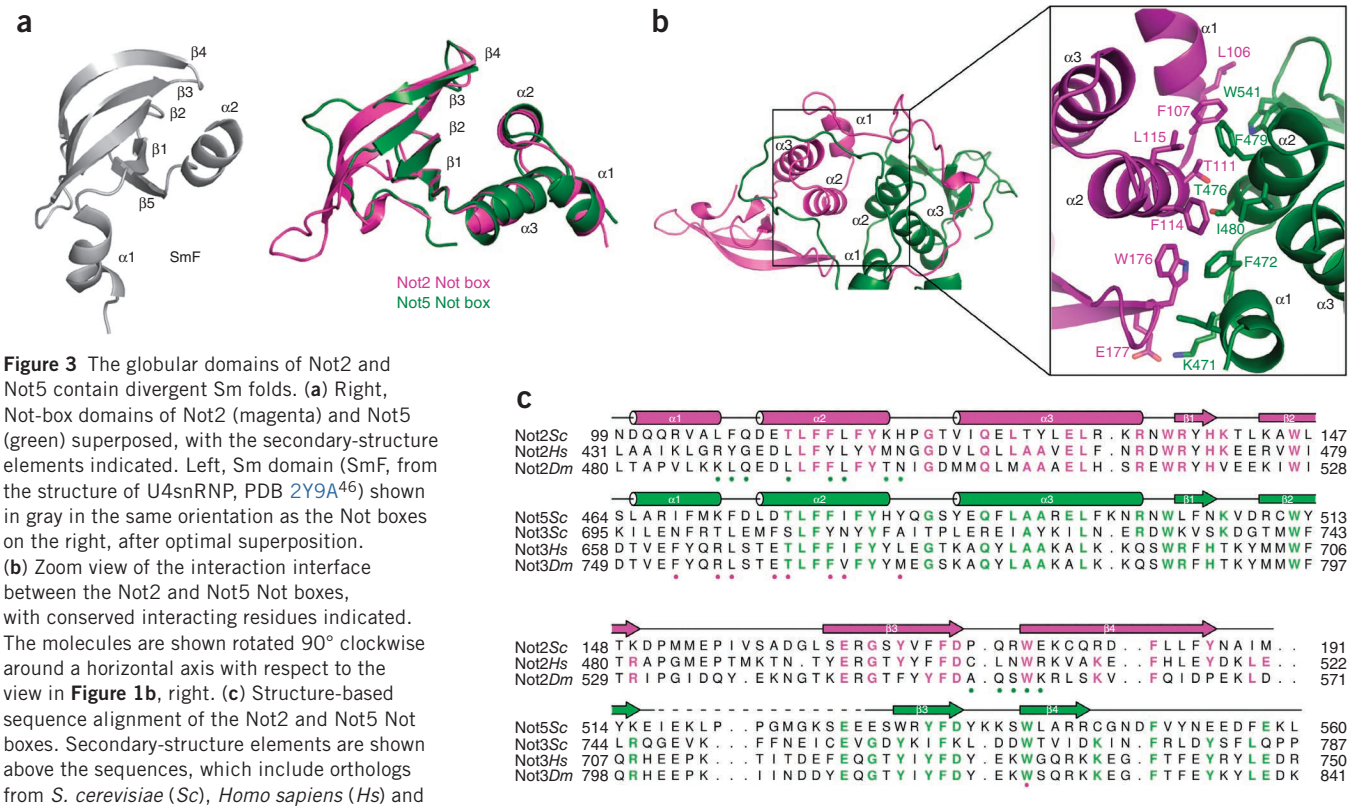

Figure 3 The globular domains of Not2 and Not5 contain divergent Sm folds. (**a**) Right, Not-box domains of Not2 (magenta) and Not5 (green) superposed, with the secondary-structure elements indicated. Left, Sm domain (SmF, from the structure of U4snRNP, PDB 2Y9A[46]) shown in gray in the same orientation as the Not boxes on the right, after optimal superposition. (**b**) Zoom view of the interaction interface between the Not2 and Not5 Not boxes, with conserved interacting residues indicated. The molecules are shown rotated 90° clockwise around a horizontal axis with respect to the view in **Figure 1b**, right. (**c**) Structure-based sequence alignment of the Not2 and Not5 Not boxes. Secondary-structure elements are shown above the sequences, which include orthologs from *S. cerevisiae* (*Sc*), *Homo sapiens* (*Hs*) and *D. melanogaster* (*Dm*). *S. cerevisiae* Not5 is similar to Not3. Sequence conservation is highlighted as magenta text for Not2 and green text for Not5. Below the sequences, the residues of Not2 that interact with Not5 are indicated with green circles, and residues of Not5 that interact with Not2 are indicated with magenta circles.

deleting either domain on Not1 binding in pulldown assays with purified proteins. As a control, Not2 and Not5c coprecipitated with glutathione S-transferase (GST)-tagged Not1c (**Fig. 2f**, lane 4). In this assay, Not5c was not coprecipitated with Not1c when Not2 was truncated (to Not2-ΔN, residues 76–191) (**Fig. 2f**, lane 5). Analogously, Not2 did not coprecipitate with Not1c when Not5c was truncated (to Not5-ΔN, residues 405–560) (**Fig. 2f**, lane 6). We concluded that formation of the core of the Not module requires the cooperative binding of Not2 and Not5.

The Not boxes of Not2 and Not5 have divergent Sm-like folds

The globular domains of Not2 and Not5 are positioned on top of the B helices of Not1 HEATs 1–4, sandwiching in between parts of the Not2 and Not5 N-terminal extensions (**Fig. 1b**,**c**). The globular domains contain the so-called Not boxes. The Not box of Not2 (residues 99–191) consists of three N-terminal helices (α1, α2 and α3) and a β-sheet formed by four antiparallel β-strands adjacent to each other (**Fig. 3a**, **Supplementary Fig. 3b** and **Supplementary Fig. 3c**). The β-sheet is highly bent: strands β3 and β4 are long and curved, with a conserved glycine residue (Gly166) at the bending point of β3. A short C-terminal extension packs against β1, creating a small β-barrel. The Not box of Not5 (residues 464–560) is similar in structure to that of Not2, superposing with an r.m.s. deviation of <1.3 Å over all Cα atoms (**Fig. 3a** and **Supplementary Fig. 3b**). The main difference is that in Not5 all the β-strands are short, thus resulting in a rather flat β-sheet.

Database searches on the DALI server[42] for structural similarities to the Not-box domains identified Sm domains as the most similar in fold (r.m.s. deviation of 2.2 Å and 2.7 Å with SmD3 and SmF, respectively[43]) (**Fig. 3a**). The Not boxes, however, differ from canonical Sm folds in several aspects. First, they lack the characteristic Sm1 and Sm2 signature motifs in the amino acid sequence. At the structural level, the Not boxes lack the fifth β-strand that in Sm proteins mediates the interaction forming dimeric Sm–Sm subcomplexes[43] and ring-like structures[44–46] (**Supplementary Fig. 3b**). The Not boxes of Not2 and Not5 also interact with each other, but in the absence of a fifth β-strand they do so with a different dimerization mechanism that involves the N-terminal α-helices (**Fig. 3b** and **Supplementary Fig. 3b**). Helix α1 of the Not2 Not box packs against the base of the β-sheet of the Not5 Not box and vice versa. Between them, the α2 helices of Not2 and Not5 pack against each other. The dimerization interface is mediated by extensive interactions centered at the conserved Phe114 and Leu115 of Not2 and the corresponding Phe479 and Ile480 of Not5 (**Fig. 3b**,**c**). Finally, the globular domains are also formed by parts of the N-terminal extensions. Residues 67–93 of Not2 wrap around the Not box of Not5, and residues 408–427 of Not5 wrap around the Not box of Not2 (**Fig. 3b**). The interactions of the N-terminal extensions effectively clamp the Not boxes on the Not1 scaffold (**Fig. 1b**,**c**).

Not1–Not2–Not5 mutations lead to growth defects *in vivo*

It has previously been shown that deletion of ~400 residues from the C terminus of Not1 is lethal in yeast[25,28]. In hindsight, these deletions generated Not1 proteins that lacked the last eight HEAT repeats (HEATs 3–10 in the Not1c structure). To test the functional importance of the Not module, we used the structural information to design point mutations that would disrupt specific interactions in the context of tagged full-length proteins.

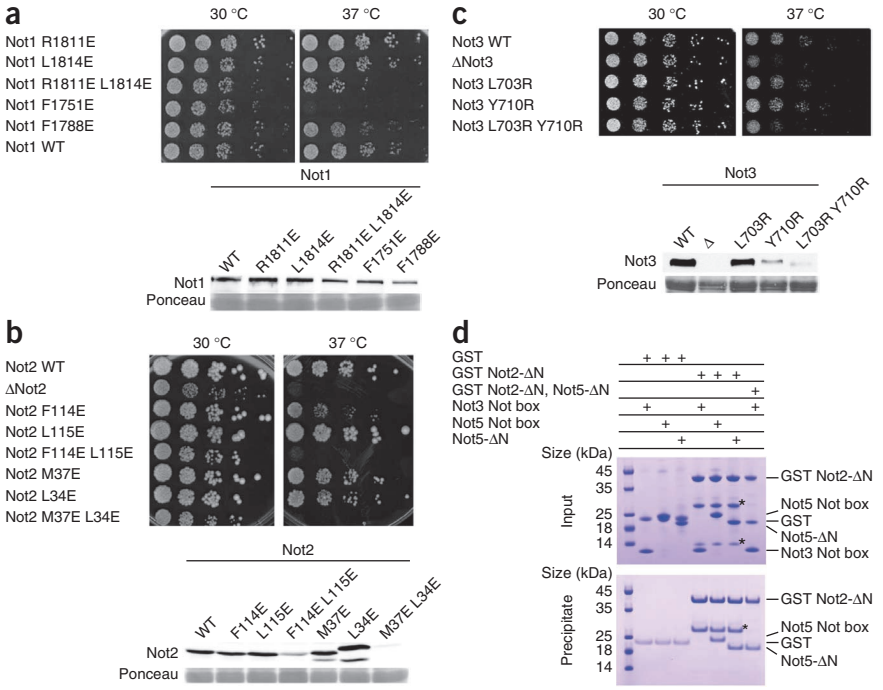

Figure 4 Analysis of mutants targeting interaction surfaces of the Not module. (**a**) Top, growth assays of Not1 mutants. Serial dilutions of cultures of strains expressing the indicated mutants incubated on YPDA medium at the indicated temperature are shown. Bottom, western blot analysis of Not1-TAP protein from cells expressing the wild-type (WT) or mutant proteins at 30 °C detected by peroxidase–antiperoxidase complex (PAP) prepared in rabbit. Ponceau staining of the membrane, used to assess equal loading in the different lanes, is shown. (**b**) Top, growth assays of Not2 mutants. Mutant strains were analyzed as for Not1 in **a**. Bottom, western blot analysis of Not2 mutant protein levels. Mutant strains were analyzed as for Not1 in **a**, with anti-VSV antibodies. (**c**) Top, growth assays of Not3 mutants. Serial dilutions of cultures of strains expressing the indicated mutant on YPDA medium containing 1 M NaCl at the indicated temperature is shown. Bottom, western blot analysis of Not3 mutant protein levels. Mutant strains were analyzed as for Not2 in **b**. (**d**) Pulldown experiment of GST Not2-ΔN with Not3 Not box, Not5 Not box and Not5-ΔN. The experiment was carried out as in **Figure 2f**. The Not3 Not box, Not5-ΔN and Not5 Not box include residues 685–800, 405–560 and 460–560 (with an N-terminal His-Z tag). The asterisks indicate a degradation product of GST Not2-ΔN.

We constructed four substitutions of Not1 residues contributing to the interaction with Not2 and Not5 (R1811E, L1814E, F1751E or F1788E) and a double mutant (R1811E L1814E) in a tandem affinity purification (TAP)-tagged plasmidic copy of the gene. R1811E and L1814E target the conserved binding site for the second segment of Not2 (**Fig. 2b**). Phe1751 is sandwiched between Not2 Trp60 and Not5 Leu388 and thus is expected to affect the binding of both proteins (**Fig. 2b**,**e**). F1788E targets the binding to the second segment of Not5 (**Fig. 2e**). Mutants were introduced in a *not1*Δ strain rescued by a *NOT1* gene (official symbol CDC39) on a *URA3*-marked plasmid. We recovered strains expressing only the mutant protein after counter selection for the *URA3* plasmid and scored the growth phenotypes at different temperatures. This revealed that Not1 R1811E or L1814E had little effect on cell growth at 30 °C and 37 °C, whereas strains expressing Not1 R1811E L1814E, F1751E or F1788E had a slow-growth phenotype at 30 °C that was exacerbated at 37 °C, particularly for the F1751E mutant (**Fig. 4a**). Western blot analyses demonstrated that the Not1 mutant proteins were expressed at comparable levels to those of the wild type (**Fig. 4a**). Coimmunoprecipitation experiments showed that the Not1 R1811E L1814E and F1751E mutants indeed blocked the interaction of Not1 with Not2 but maintained a normal interaction with Pop2 (**Supplementary Fig. 4a**).

Next, we engineered substitutions in Not2. The Not2 mutants L34E, M37E and the double mutant combining these substitutions target the Not1-binding site (**Fig. 2b**). The Not2 mutants F114E, L115E and combination of these substitutions target the binding to the Not5 Not box (**Fig. 3b**). We did not test the corresponding mutations in Not5 because of the lack of easily scorable phenotypes of Not5 mutants in our strain background (V.R. and B.S., unpublished observations). We introduced the Not2 mutants in yeast cells with the corresponding wild-type gene deleted and assayed the growth phenotypes of the resulting strains at various temperatures on appropriate medium. At 30 °C, the different mutations had no detectable effect, whereas growth of the double-mutant strains was severely impaired at 37 °C (**Fig. 4b**). Western blot analyses revealed that the double-mutant proteins were barely detectable (**Fig. 4b**). This observation that interfering with the dimerization of the Not2 Not box destabilizes Not2 is consistent with previous analyses of Not2 mutants showing protein instability with a concomitant reduction of Not5 protein level[33].

Differences between yeast and human Not3

In coimmunoprecipitation experiments, mutation on the surface of Not1 that interacts with Not2–Not5 also prevented the association with Not3 (**Supplementary Fig. 4a**), a result reinforcing the parallel between Not3 and Not5. The Not box of Not3 is predicted to have a similar fold and dimerization interface as those of Not5 or Not2. The central residues at the putative dimerization interface of the Not3 Not box are conserved, including Leu703 and Tyr710 (which are equivalent to *S. cerevisiae* Not5 Phe472 and PheF479; **Fig. 3b**,**c**). Using strategies described above, we constructed and evaluated yeast strains with the Not3 L703R and Y710R substitutions and a combination of both. As in the case of Not2, the single mutants had no growth phenotypes, whereas the growth of the double-mutant strain was severely impaired at 37 °C (**Fig. 4c**). Similarly as for the Not2 double mutants, low levels of the Not3 protein were present for the Not3 L703R Y710R double mutant (**Fig. 4c**). These results are consistent with the notion that Not3 is also destabilized if the Not-box domain is mutated at the putative dimerization interface.

Yeast Not3 and Not2 have been shown to associate *in vivo*[21,47,48] (**Supplementary Fig. 4b**), although with the caveat that the interaction might be indirect. The interaction between human Not2 and Not3 Not boxes has been shown to be direct by *in vitro* assays[19]. To test for direct interactions of the yeast proteins *in vitro*, we engineered a fragment of yeast Not3 encompassing a minimal Not-box region (residues 685–800). In contrast to that of Not5, the Not3 Not box failed to interact with GST Not2-ΔN in pulldown assays with purified proteins (**Fig. 4d**, comparison of lane 4 with lanes 5 and 6). The Not3 Not box also did not bind on top of the GST Not2-ΔN–Not5-ΔN complex (**Fig. 4d**, lane 7). A close inspection of the amino acid sequences revealed that a subset of residues at the putative dimerization interface

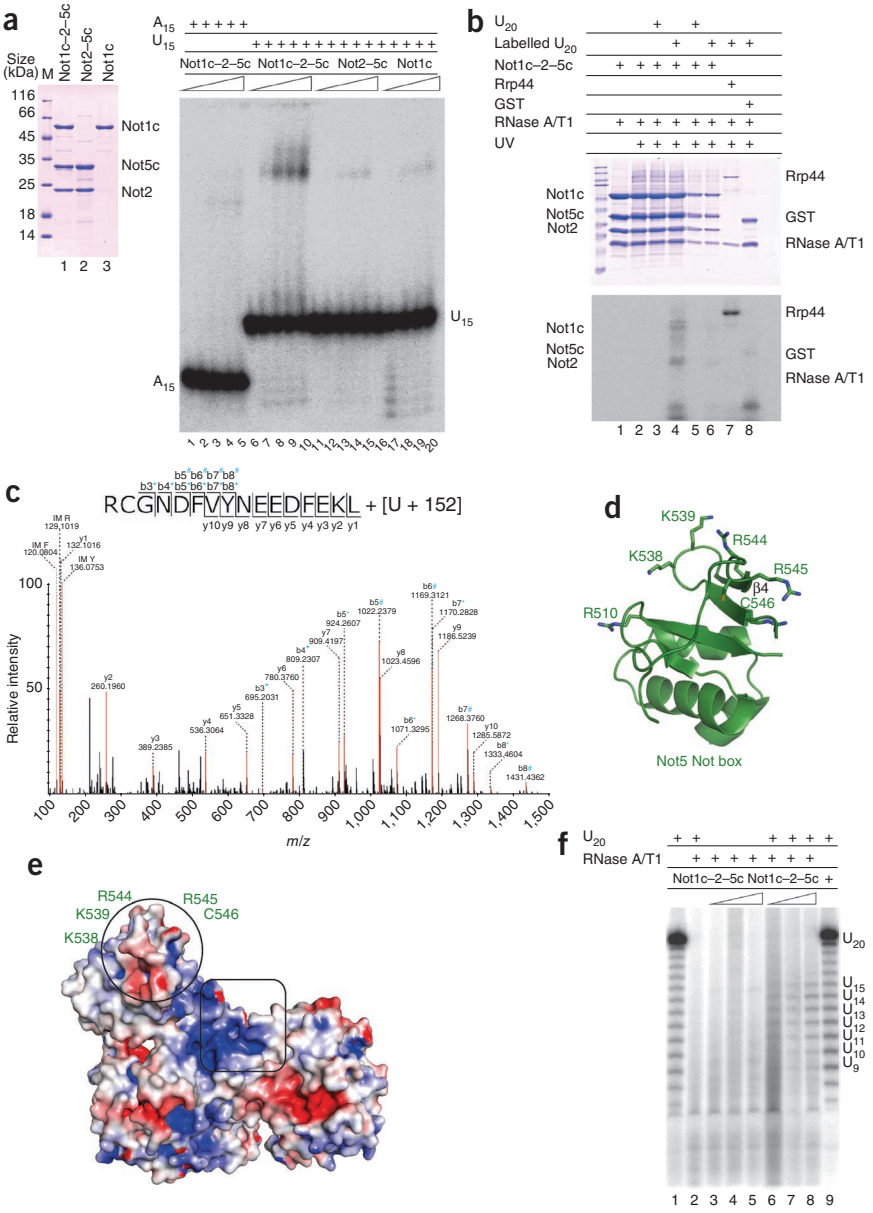

Figure 5 Not1c–Not2–Not5c binds poly(U) RNA. (**a**) Left, Coomassie-stained 13.5% SDS-PAGE gel with the protein samples used for the biochemical assays. Right, EMSA with A_{15} or U_{15} RNA (50 nM) labeled at the 5′ end with [γ-^{32}P]phosphate and incubated with increasing amounts of proteins. M, molecular weight marker. (**b**) Protein-RNA cross-linking. Proteins and γ-^{32}P body-labeled poly(U) 20-mer RNAs cross-linked under UV light and separated on 13.5% SDS PAGE are shown. The gel was stained with Coomassie blue (top) and analyzed by phosphorimaging (bottom). A/T1, mixture of RNases A and T1. (**c**) Tandem mass spectrum of Not5 residues 545–560, identifying an additional mass of 476.0338 Da corresponding to a U nucleoside with an adduct of 152 Da (associated with a cysteine). Peptide sequence and fragment ions are indicated at top. b ions with a mass shift corresponding to U-H_3PO_4 + 152 and to U + 152 are shown with an asterisk and hash mark, respectively. IM, immonium ions. (**d**) Structure of the Not5 Not box showing the position of the U–cross-linked Cys546 surrounded by a patch of positively charged residues. (**e**) Surface representation of the ternary complex colored by electrostatic potential (positive in blue and negative in red), calculated with PyMOL APBS tools. The structure is shown after a 180° rotation around a vertical axis with respect to **Figure 1b**, left. Circle, RNA-binding site; square, positively charged surface patch at the intersection of Not1, Not2 and Not5. (**f**) RNase protection assay. The protected RNA fragments obtained after treatment with a mixture of RNase A and T1 (A/T1), labeled at the 5′ end with γ-^{32}P, resolved by denaturing PAGE and analyzed by phosphorimaging are shown.

is conserved between Not5 and metazoan Not3 but diverges in *S. cerevisiae* Not3 (for example, *S. cerevisiae* Not3 Thr702, Phe706, Asn711 and Ala715 in **Fig. 3c**), thus rationalizing the different behavior of the yeast Not3 protein.

The Not1–Not2–Not5 complex is a binding platform for proteins

The interaction of the Not2 and Not5 Not boxes creates a V-shaped surface (**Fig. 1b**,**c**). In one molecule of the asymmetric unit, the β-sheet of Not2 is extended by a loop that mediates a crystal contact with the β4 strand of Not5 from a symmetry-related molecule (**Supplementary Fig. 3c**). Another interaction is present between the β4 strand of Not2 and the β-hairpin of Not1 from a symmetry copy (**Supplementary Fig. 3c**). These crystal-packing contacts are somewhat reminiscent of canonical Sm-Sm interactions and point to the Not boxes as possible interaction surfaces. Genetic evidence suggests that the Not box of Not5 interacts with Not4 (ref. 24), a conserved subunit of the complex with ubiquitin-ligase activity[49]. The Not box of Not2 interacts with ADA2, a component of the transcription-regulatory histone-acetylation complex SAGA[33]. Mutation of Not2 Arg165 has been shown to abrogate the interaction with ADA2 without affecting the integrity of the Ccr4–Not complex in yeast[33]. In the structure, Arg165 protrudes on the surface of the β-barrel and is indeed accessible to solvent (**Supplementary Fig. 5a**).

Not1–Not2–Not5 is a binding platform for poly(U) RNA

Mapping of the electrostatic potential on the molecular surface of the Not1c–Not2–Not5c complex showed patches of positively charged residues. We therefore tested whether the Not module can mediate RNA binding. In electrophoretic mobility shift assays (EMSAs), a single-stranded poly(U) 15-mer RNA (U_{15}) bound the Not1c–Not2–Not5c complex, whereas we detected no binding with a poly(A) 15-mer RNA (A_{15}) (**Fig. 5a**). The Not module recognized poly(U) RNA specifically, albeit with low affinity (in the low-micromolar range; **Supplementary Fig. 5b**). The EMSAs showed no binding of U_{15} RNA to either Not1c or the Not2–Not5c subcomplex in isolation (**Fig. 5a**), suggesting that the different portions of the Not module contribute together to RNA recognition. Indeed, after incubation of the Not1c–Not2–Not5c complex with a body-labeled U_{20} RNA and exposure to UV irradiation at 254 nm, all bands showed RNA cross-linking, which was strong in the case of Not1c and Not2 and less pronounced in the case of Not5c (**Fig. 5b**). In this experiment, Rrp44 and GST were positive and negative controls, respectively (**Fig. 5b**).

Next, we used MS to identify residues of the complex cross-linked to the U_{20} RNA. This approach is based on the detection and sequencing by LC-MS/MS of peptides conjugated to the mass of an RNA nucleotide (reviewed in ref. 50). The advantage of this approach is that RNA-contact sites are determined in an unbiased manner. The caveat is that the identification is limited to sites where the ribonucleotide is in proximity to amino acids with reactive groups (for example, thiol groups in cysteine residues) and is limited by the amounts of the cross-linked species and the complexity of the spectra. The MS analysis identified a Not5 peptide corresponding to residues 545–560, with a single U nucleoside cross-linked to Cys546 (**Fig. 5c**). In the structure, Cys546 is positioned at the top of the Not-box β-sheet and is part of a surface patch with positively charged residues (Lys515, Arg533, Arg544 and Arg545; **Fig. 5d**). This RNA-binding site (circle in **Fig. 5e**) differs from the U nucleoside–binding site of canonical Sm folds (**Supplementary Fig. 5a**) and is contiguous to a positively charged surface patch at the intersection of Not1, Not2 and Not5 (square in **Fig. 5e**). To estimate the length of the RNA-binding path on the complex, we carried out RNase protection assays. We found that fragments of 11–15 nucleotides accumulated in the presence of Not1c–Not2–Not5c (**Fig. 5f**, lanes 6–8). Fragments of this size could easily span a distance of 40–60 Å.

DISCUSSION

The core of the Not module that we investigated in this work is built around the C-terminal arm of Not1 by the cooperative binding of Not2 and Not5. The C-terminal arm of Not1 has a HEAT-repeat architecture similar to that found in the N-terminal arm[28]. It is thus possible to imagine that the two arms of Not1 might have originated from a duplication event. Not2 and Not5 interact through their C-terminal Not-box domains. At the structural level, the Not boxes resemble Sm folds. The similarity extends to their biochemical properties in terms of the ability of the Not boxes to dimerize and to bind poly(U) RNA stretches, although the interaction mechanisms have diverged from those of canonical Sm folds. The heterodimerization of the two Not boxes in the Not1–Not2–Not5 complex serves multiple purposes.

First, heterodimerization of the Not boxes tethers the N-terminal regions of Not2 and Not5, promoting their synergistic binding to Not1. Previous studies have shown that the N-terminal region of Not2 is essential for the structural integrity of the Not module because it recruits Not5 into the complex[33]. We found that the N-terminal region of Not5 is equally important in recruiting Not2. The two Not boxes thus contribute indirectly to Not1 binding by bringing the two separate N-terminal regions into spatial proximity, thus probably increasing their effective local concentration. Not-box heterodimerization is also important for the stability of the individual proteins *in vivo*, as shown by mutational analysis of Not2 as well as Not3. Yeast Not3 and Not5 are currently considered homologs with partially redundant functions[24,30]. Unexpectedly, we found that the yeast Not3 Not box has diverged from that of Not5 and does not interact with the Not2 Not box *in vitro*. The Not2-Not5 dimerization interface is instead conserved in Not2-Not3 from higher eukaryotes, thus suggesting that the protein that we currently refer to as metazoan Not3 is an ortholog of *S. cerevisiae* Not5. The identity of the direct interactions that mediate the recruitment of *S. cerevisiae* Not3 Not box in the complex is currently unclear and is an important question for future studies.

Second, the Not boxes together with Not1 form a composite platform for macromolecular interactions. Extensive data indicate that the Not module is closely connected to the transcriptional machinery and physically recruits transcription factors, such as ADA2 (ref. 33; reviewed in refs. 11,12). Evidence is also accumulating on the ability of the Not modules to mediate protein-protein interactions important for cytoplasmic mRNA metabolism. For example, in *Drosophila melanogaster* Not3 binds the translational repressor BicC[51], and in mice the C-terminal arm of Not1 binds the mRNA developmental regulator NANOS2 (ref. 52). We found that the Not module creates a composite RNA-binding surface for U nucleosides, with a specific site in the Not box of Not5. Although this RNA-binding activity of the Not module was unexpected, it rationalizes previous observations. In yeast, the decay of the Edc1 mRNA has been shown to proceed through a deadenylation-independent decapping pathway that depends on the Not proteins and on a poly-U tract in its 3′ UTR[39]. A model is conceivable in which binding of the Not module to this 3′ UTR might bring the mRNA into proximity of Dhh1, a decapping activator (known as DDX6 or RCK in metazoans) that is recruited to Ccr4–Not[27,53,54]. Interestingly, mouse Not3 has been shown to regulate the deadenylation of specific mRNAs by recruiting their 3′ UTRs[32], which also contain U-rich stretches. The emerging picture is that the Not module of the Ccr4–Not complex creates a platform for protein and nucleic acid interactions that is able to contribute to the many functions of the Ccr4–Not complex, including the degradation of specific mRNAs.

METHODS

Methods and any associated references are available in the online version of the paper.

Accession codes. Coordinates and structural factors have been deposited in the Protein Data Bank under accession code 4BY6.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS

We thank the MPI Biochemistry Crystallization Facility and Core Facility. We thank F. Bonneau and C. Basquin (MPI Biochemistry) for help with biochemical assays and for **Supplementary Figure 5b**; F. Lacroute, F. Gabriel and M.C. Daugeron (Centre de Génétique Moléculaire) for yeast strains; the staff of the Swiss Light Source synchrotron for assistance during data collection; and members of our laboratories for discussions and for critical reading of the manuscript. E.C. acknowledges support from the Max Planck Gesellschaft, the European Research Council (ERC Advanced Investigator Grant 294371, Marie Curie Initial Training Network RNPnet 289007) and the Deutsche Forschungsgemeinschaft (DFG SFB646, SFB1035, GRK1721, FOR1680, CIPSM). B.S. acknowledges support from the Centre Européen de Recherche en Biologie et en Médecine (CERBM)-IGBMC, the CNRS and the Ligue Contre le Cancer (Equipe Labellisée 2011).

AUTHOR CONTRIBUTIONS

V.B. and J.B. carried out the structure determination and the *in vitro* experiments; V.R. carried out the *in vivo* experiments; K.S. and H.U. carried out the MS analysis; E.C. and B.S. supervised the project; and E.C., V.B. and B.S. wrote the manuscript.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

1. Wahle, E. & Winkler, G.S. RNA decay machines: deadenylation by the Ccr4-Not and Pan2-Pan3 complexes. *Biochim. Biophys. Acta* **1829**, 561–570 (2013).
2. Tucker, M. *et al.* The transcription factor associated Ccr4 and Caf1 proteins are components of the major cytoplasmic mRNA deadenylase in *Saccharomyces cerevisiae*. *Cell* **104**, 377–386 (2001).
3. Daugeron, M.C., Mauxion, F. & Séraphin, B. The yeast POP2 gene encodes a nuclease involved in mRNA deadenylation. *Nucleic Acids Res.* **29**, 2448–2455 (2001).

4. Sandler, H., Kreth, J., Timmers, H.T.M. & Stoecklin, G. Not1 mediates recruitment of the deadenylase Caf1 to mRNAs targeted for degradation by tristetraprolin. *Nucleic Acids Res.* **39**, 4373–4386 (2011).
5. Fabian, M.R. *et al.* Structural basis for the recruitment of the human CCR4–NOT deadenylase complex by tristetraprolin. *Nat. Struct. Mol. Biol.* **20**, 735–739 (2013).
6. Braun, J.E., Huntzinger, E., Fauser, M. & Izaurralde, E. GW182 proteins directly recruit cytoplasmic deadenylase complexes to miRNA targets. *Mol. Cell* **44**, 120–133 (2011).
7. Chekulaeva, M. *et al.* miRNA repression involves GW182-mediated recruitment of CCR4–NOT through conserved W-containing motifs. *Nat. Struct. Mol. Biol.* **18**, 1218–1226 (2011).
8. Fabian, M.R. *et al.* miRNA-mediated deadenylation is orchestrated by GW182 through two conserved motifs that interact with CCR4–NOT. *Nat. Struct. Mol. Biol.* **18**, 1211–1217 (2011).
9. Goldstrohm, A.C., Hook, B.A., Seay, D.J. & Wickens, M. PUF proteins bind Pop2p to regulate messenger RNAs. *Nat. Struct. Mol. Biol.* **13**, 533–539 (2006).
10. Suzuki, A., Igarashi, K., Aisaki, K.-I., Kanno, J. & Saga, Y. NANOS2 interacts with the CCR4-NOT deadenylation complex and leads to suppression of specific RNAs. *Proc. Natl. Acad. Sci. USA* **107**, 3594–3599 (2010).
11. Collart, M.A. & Panasenko, O.O. The Ccr4-Not complex. *Gene* **492**, 42–53 (2012).
12. Miller, J.E. & Reese, J.C. Ccr4-Not complex: the control freak of eukaryotic cells. *Crit. Rev. Biochem. Mol. Biol.* **47**, 315–333 (2012).
13. Sun, M. *et al.* Comparative dynamic transcriptome analysis (cDTA) reveals mutual feedback between mRNA synthesis and degradation. *Genome Res.* **22**, 1350–1359 (2012).
14. Liu, H.Y. *et al.* The NOT proteins are part of the CCR4 transcriptional complex and affect gene expression both positively and negatively. *EMBO J.* **17**, 1096–1106 (1998).
15. Chen, J. *et al.* Purification and characterization of the 1.0 MDa CCR4-NOT complex identifies two novel components of the complex. *J. Mol. Biol.* **314**, 683–694 (2001).
16. Albert, T.K. *et al.* Isolation and characterization of human orthologs of yeast CCR4-NOT complex subunits. *Nucleic Acids Res.* **28**, 809–817 (2000).
17. Lau, N.C. *et al.* Human Ccr4-Not complexes contain variable deadenylase subunits. *Biochem. J.* **422**, 443–453 (2009).
18. Temme, C. *et al.* Subunits of the *Drosophila* CCR4-NOT complex and their roles in mRNA deadenylation. *RNA* **16**, 1356–1370 (2010).
19. Bawankar, P., Loh, B., Wohlbold, L., Schmidt, S. & Izaurralde, E. NOT10 and C2orf29/NOT11 form a conserved module of the CCR4-NOT complex that docks onto the NOT1 N-terminal domain. *RNA Biol.* **10**, 228–244 (2013).
20. Schwede, A. *et al.* A role for Caf1 in mRNA deadenylation and decay in trypanosomes and human cells. *Nucleic Acids Res.* **36**, 3374–3388 (2008).
21. Bai, Y. *et al.* The CCR4 and CAF1 proteins of the CCR4-NOT complex are physically and functionally separated from NOT2, NOT4, and NOT5. *Mol. Cell Biol.* **19**, 6642–6651 (1999).
22. Mauxion, F., Prève, B. & Seraphin, B. C2ORF29/CNOT11 and CNOT10 form a new module of the CCR4-NOT complex. *RNA Biol.* **10**, 267–276 (2013).
23. Färber, V., Erben, E., Sharma, S., Stoecklin, G. & Clayton, C. Trypanosome CNOT10 is essential for the integrity of the NOT deadenylase complex and for degradation of many mRNAs. *Nucleic Acids Res.* **41**, 1211–1222 (2013).
24. Oberholzer, U. & Collart, M.A. Characterization of NOT5 that encodes a new component of the Not protein complex. *Gene* **207**, 61–69 (1998).
25. Maillet, L., Tu, C., Hong, Y.K., Shuster, E.O. & Collart, M.A. The essential function of Not1 lies within the Ccr4-Not complex. *J. Mol. Biol.* **303**, 131–143 (2000).
26. Draper, M.P., Liu, H.Y., Nelsbach, A.H., Mosley, S.P. & Denis, C.L. CCR4 is a glucose-regulated transcription factor whose leucine-rich repeat binds several proteins important for placing CCR4 in its proper promoter context. *Mol. Cell Biol.* **14**, 4522–4531 (1994).
27. Maillet, L. & Collart, M.A. Interaction between Not1p, a component of the Ccr4-Not complex, a global regulator of transcription, and Dhh1p, a putative RNA helicase. *J. Biol. Chem.* **277**, 2835–2842 (2002).
28. Basquin, J. *et al.* Architecture of the nuclease module of the yeast Ccr4-Not complex: the Not1-Caf1-Ccr4 interaction. *Mol. Cell* **48**, 207–218 (2012).
29. Petit, A.-P. *et al.* The structural basis for the interaction between the CAF1 nuclease and the NOT1 scaffold of the human CCR4-NOT deadenylase complex. *Nucleic Acids Res.* **40**, 11058–11072 (2012).
30. Collart, M.A., Panasenko, O.O. & Nikolaev, S.I. The Not3/5 subunit of the Ccr4-Not complex: a central regulator of gene expression that integrates signals between the cytoplasm and the nucleus in eukaryotic cells. *Cell Signal.* **25**, 743–751 (2013).
31. Neely, G.G. *et al.* A global *in vivo Drosophila* RNAi screen identifies NOT3 as a conserved regulator of heart function. *Cell* **141**, 142–153 (2010).
32. Morita, M. *et al.* Obesity resistance and increased hepatic expression of catabolism-related mRNAs in *Cnot3*$^{+/-}$ mice. *EMBO J.* **30**, 4678–4691 (2011).
33. Russell, P., Benson, J.D. & Denis, C.L. Characterization of mutations in NOT2 indicates that it plays an important role in maintaining the integrity of the CCR4-NOT complex. *J. Mol. Biol.* **322**, 27–39 (2002).
34. Ito, K. *et al.* CNOT2 depletion disrupts and inhibits the CCR4-NOT deadenylase complex and induces apoptotic cell death. *Genes Cells* **16**, 368–379 (2011).
35. Zwartjes, C.G.M., Jayne, S., van den Berg, D.L.C. & Timmers, H.T.M. Repression of promoter activity by CNOT2, a subunit of the transcription regulatory Ccr4-Not complex. *J. Biol. Chem.* **279**, 10848–10854 (2004).
36. Badarinarayana, V., Chiang, Y.C. & Denis, C.L. Functional interaction of CCR4-NOT proteins with TATAA-binding protein (TBP) and its associated factors in yeast. *Genetics* **155**, 1045–1054 (2000).
37. Lemaire, M. & Collart, M.A. The TATA-binding protein-associated factor yTafII19p functionally interacts with components of the global transcriptional regulator Ccr4-Not complex and physically interacts with the Not5 subunit. *J. Biol. Chem.* **275**, 26925–26934 (2000).
38. Deluen, C. *et al.* The Ccr4-not complex and yTAF1 (yTaf(II)130p/yTaf(II)145p) show physical and functional interactions. *Mol. Cell Biol.* **22**, 6735–6749 (2002).
39. Muhlrad, D. & Parker, R. The yeast EDC1 mRNA undergoes deadenylation-independent decapping stimulated by Not2p, Not4p, and Not5p. *EMBO J.* **24**, 1033–1045 (2005).
40. Andrade, M.A., Petosa, C., O'Donoghue, S.I., Müller, C.W. & Bork, P. Comparison of ARM and HEAT protein repeats. *J. Mol. Biol.* **309**, 1–18 (2001).
41. Marcotrigiano, J. *et al.* A conserved HEAT domain within eIF4G directs assembly of the translation initiation machinery. *Mol. Cell* **7**, 193–203 (2001).
42. Holm, L. & Rosenström, P. Dali server: conservation mapping in 3D. *Nucleic Acids Res.* **38**, W545–W549 (2010).
43. Kambach, C. *et al.* Crystal structures of two Sm protein complexes and their implications for the assembly of the spliceosomal snRNPs. *Cell* **96**, 375–387 (1999).
44. Törö, I. *et al.* RNA binding in an Sm core domain: X-ray structure and functional analysis of an archaeal Sm protein complex. *EMBO J.* **20**, 2293–2303 (2001).
45. Khusial, P., Plaag, R. & Zieve, G.W. LSm proteins form heptameric rings that bind to RNA via repeating motifs. *Trends Biochem. Sci.* **30**, 522–528 (2005).
46. Leung, A.K.W., Nagai, K. & Li, J. Structure of the spliceosomal U4 snRNP core domain and its implication for snRNP biogenesis. *Nature* **473**, 536–539 (2011).
47. Azzouz, N. *et al.* Specific roles for the Ccr4-Not complex subunits in expression of the genome. *RNA* **15**, 377–383 (2009).
48. Tarassov, K. *et al.* An *in vivo* map of the yeast protein interactome. *Science* **320**, 1465–1470 (2008).
49. Albert, T.K. *et al.* Identification of a ubiquitin–protein ligase subunit within the CCR4–NOT transcription repressor complex. *EMBO J.* **21**, 355–364 (2002).
50. Schmidt, C., Kramer, K. & Urlaub, H. Investigation of protein-RNA interactions by mass spectrometry: techniques and applications. *J. Proteomics* **75**, 3478–3494 (2012).
51. Chicoine, J. *et al.* Bicaudal-C recruits CCR4-NOT deadenylase to target mRNAs and regulates oogenesis, cytoskeletal organization, and its own expression. *Dev. Cell* **13**, 691–704 (2007).
52. Suzuki, A., Saba, R., Miyoshi, K., Morita, Y. & Saga, Y. Interaction between NANOS2 and the CCR4-NOT deadenylation complex is essential for male germ cell development in mouse. *PLoS ONE* **7**, e33558 (2012).
53. Hata, H. *et al.* Dhh1p, a putative RNA helicase, associates with the general transcription factors Pop2p and Ccr4p from *Saccharomyces cerevisiae*. *Genetics* **148**, 571–579 (1998).
54. Coller, J.M., Tucker, M., Sheth, U., Valencia-Sanchez, M.A. & Parker, R. The DEAD box helicase, Dhh1p, functions in mRNA decapping and interacts with both the decapping and deadenylase complexes. *RNA* **7**, 1717–1727 (2001).

ONLINE METHODS

Protein purification. All the proteins were cloned and expressed individually in *E. coli* BL21(DE3) pLysS cells (Stratagene) in TB medium with IPTG induction overnight at 18 °C. $Not1_{1541–2093}$ (Not1c), Not5 FL and $Not5_{298–560}$ (Not5c) were expressed with an N-terminal His-SUMO tag (cleavable with the Senp2 protease). The Not1 C-terminal arm (starting at 1348), Not2 FL and $Not5_{460–560}$ (Not5 Not box) were expressed with an N-terminal His-Z tag (cleavable with TEV protease). $Not3_{685–800}$ (Not3 Not box) was expressed with an N-terminal His tag (cleavable with TEV protease). The cells were lysed in buffer A (50 mM sodium phosphate buffer, pH 7.5, 250 mM NaCl and 20 mM imidazole) by sonication. The lysates were cleared by centrifugation and were loaded on a 5-ml His-trap column (GE Healthcare). The column was washed with buffer B (50 mM phosphate buffer, pH 6.5, 1 M NaCl, 20 mM imidazole, 50 mM KCl, 10 mM $MgSO_4$ and 2 mM ATP) and with buffer A. The proteins were eluted by a gradient of buffer A and buffer C (buffer A supplemented with 500 mM imidazole). Except for the Not5 Not box and Not2, all other proteins were dialyzed overnight in gel-filtration buffer (without DTT) in the presence of TEV or Senp2 proteases and were then applied to the His-trap column to remove the cleaved tag (second affinity step). The proteins were further purified by size-exclusion chromatography in the gel-filtration buffer (20 mM Tris-Cl buffer, pH 7.5, 250 mM NaCl and 2 mM DTT). The complex of Not1c–Not2–Not5c was formed by mixing the purified protein in a 1:1.25:1 molar ratio and was incubated with TEV protease overnight at 4 °C to cleave the N-terminal His-Z tag of Not2. The complex was applied onto the 5-ml His-trap column (GE Healthcare) to remove the cleaved tag and was purified by gel filtration (Superdex 200 16/60, GE Healthcare) in the gel-filtration buffer.

For the pulldown assays, Not5c, Not5-ΔN and Not2-ΔN were expressed as N-terminal His-GST fusion proteins, whose tags were cleavable with 3C protease. Not2, Not5c and Not5-ΔN were affinity purified with a 5-ml His-trap column (GE Healthcare) as described above. Not2-ΔN was affinity purified at a pH of 8.5 (with buffer A and C at pH 8.5) instead of pH 7.5. Not2–Not5c, Not2-ΔN–Not5c and Not2–Not5-ΔN complexes were formed by mixing a 1:1.5 molar ratio of the larger to smaller protein and dialyzed in gel-filtration buffer (without 2 mM DTT) in the presence of 3C protease and TEV protease. The dialyzed proteins were subjected to a second His-affinity purification with a 5-ml His-trap column (GE Healthcare) and subsequent incubation with glutathione-agarose beads (Protino) for 2 h at 4 °C to remove the GST-tag contamination. The proteins were then purified by gel filtration (Superdex 75 10/30, GE Healthcare) in gel-filtration buffer. Not1c-GST was affinity purified as described above. The protein was dialyzed against heparin buffer A (20 mM Tris, pH 7.5, and 100 mM NaCl), applied onto the 5-ml heparin column (GE Healthcare) and purified with a gradient elution with heparin buffer A and heparin buffer A supplemented with 1 M NaCl. Not1c-GST was further purified by gel filtration (Superdex 200 10/30, GE Healthcare). The Not3 Not box was expressed as TEV protease–cleavable His_6 fusion protein and purified in a similar way as mentioned above.

For the RNA-binding experiments, Not2 and Not5c were expressed and purified as individual proteins as described above. The Not2–Not5c complex was formed by mixing the proteins in a 1.25:1 molar ratio and subsequent overnight incubation with TEV protease. The cleaved protein was subjected to His-affinity purification to remove the cleaved tag and a subsequent heparin-column purification. The complex was further purified by gel filtration (Superdex 200 16/60, GE Healthcare). Not1c was expressed and purified as above, with an additional step of heparin purification included after the second His-affinity step. Not1c–Not2–Not5c was purified by mixing Not1c and Not2–Not5c in a 1:1.25 molar ratio and subsequent gel filtration (Superdex 200 16/60, GE Healthcare).

Limited proteolysis experiment. 0.6 mg/ml of the Not1Δ1347–Not2–Not5 complex was incubated with elastase (Roche) at a 1:10 (w/w) enzyme/protein ratio for 30 min on ice. The products of the proteolysis were identified by N-terminal sequencing and MS analysis. The interacting core complex was identified by size-exclusion chromatography of the proteolyzed sample.

Crystallization and structure solution. The Not1c–Not2–Not5c complex was concentrated to 16 mg/ml and crystallized at room temperature in 8.5% (w/v) PEG 8000, 100 mM MES, pH 6.5, and 200 mM calcium acetate. The mercury derivative was prepared by cocrystallization of a solution of Not1c–Not2–Not5c with ethyl mercury phosphate (EMP) at 0.55 mM final concentration. The crystals were frozen in the presence of 20% glycerol as cryoprotectant. X-ray data were collected at 100 K at the SLS synchrotron (PXII and PXIII beamlines), with tuning of the wavelength at the Hg edge in the case of the EMP-containing crystals for SAD data collection. The data were processed with XDS[55]. The crystals belong to a monoclinic space group ($P2_1$) with two molecules per asymmetric unit. We used PHENIX.autosol[56] for phasing and Buccaneer[57] for the initial automatic model building. We completed the model with iterative rounds of manual model building with Coot[58] and restrained refinement with PHENIX[56]. The final model has 97.3% residues in the most-favored regions of the Ramachandran plot, as calculated with MolProbity[59].

Pulldown assays. For the experiments in **Figure 2f**, 50 pmol of bait (GST or GST-Not1c) were incubated with 100 pmol of prey (Not2–Not5c, Not2-ΔN–Not5c and Not2–Not5-ΔN) for 1 h at 4 °C in 40 mM Tris, pH 7.5, 150 mM NaCl, 2 mM DTT, 12.5% (v/v) glycerol and 0.1% (w/v) NP-40 (binding buffer). The protein mix was incubated with 20 µl of GSH-agarose beads (Protino) for 1 h with gentle rocking at 4 °C. The resin was washed three times with the binding buffer, and the proteins were eluted in 15 µl of binding buffer containing 100 mM glutathione. Input and precipitates were mixed with SDS loading dye, resolved on 4–12% Bis-Tris NuPAGE gel (Invitrogen) with MES as running buffer, and visualized by Coomassie-blue staining. A similar protocol was used for the GST pulldown assays in **Figure 4d** with 40 mM Tris, pH 8.5, 100 mM NaCl, 2 mM DTT and 12.5% (v/v) glycerol as binding buffer.

Sequence alignments and superpositions. All the sequence alignments were done with ClustalW[60] and ALINE[61], and the structural superpositions were done with SSM in Coot[58]. The r.m.s. deviations reported are from the output of Coot. Structure-based sequence alignment was done in STRAP[62] with the Aligner3D method and manually edited in ALINE.

Yeast strains. Yeast strains used in this study are all derivatives of W303 (*ade2-1*, *can1-100*, *leu2-3,112*, *his3-11,15*, *ura3-1*, *trp1-1*). Genes differing from W303 are as follows: T26N28 (*MATa*, *Δtrp1*, *ΔNOT1::HIS3 pFL 38* (*NOT1*)), BSY1110 (*MATa*, *Δtrp1*, *not2::HISMX6*), BSY1111 (*MATa*, *Δtrp1*, *not3::HISMX6*), BSY1230 (*MATa*, *Δtrp1*, *NOT3-VSV*, *NOT2-3HA::hisMX6*), BSY1231 (*MATa*, *Δtrp1*, *POP2-VSV*, *NOT2-3HA::hisMX6*), BSY1240 (*MATa*, *Δtrp1*, *NOT3-TAP::$TRP1_{Kl}$*, *POP2-VSV*, *NOT4-HA::hisMX6*) and BSY1242 (*MATa*, *Δtrp1*, *NOT3-TAP::$TRP1_{Kl}$*, *POP2-VSV*, *NOT2-HA::hisMX6*).

Coprecipitation assays. Protein extract preparation and coimmunoprecipitation were performed as described previously[28].

Mutant analyses. Mutations in Not1, Not2 and Not3 were constructed in plasmids pBS4806 (ref. 28) (Not1-TAP), pBS4968 (Not2-VSV) and pBS4975 (Not3-VSV) by one or multiple rounds of site-directed mutagenesis. The presence of the desired mutation was ascertained by sequencing. The resulting plasmids were introduced into yeast strains with the lithium acetate transformation procedure. Plasmid shuffling, growth assays, protein extraction and western blot analyses were performed with standard procedures as previously described[28].

Electrophoretic mobility shift assays. Proteins at 3, 10, 13 and 20 µM concentration (30, 100, 130 and 200 pmol, respectively) were incubated with 50 nM (0.5 pmol) of 5′-labeled RNA (A_{15} or U_{15}) at 4 °C overnight in 20 mM HEPES, pH 7.5, 5 mM EDTA, 0.1% NP-40 and 2 mM DTT (EMSA buffer). The reaction mixtures were complemented with gel-filtration buffer to a final NaCl concentration of 54 mM, resolved on a 6% (w/v) native PAGE and visualized by phosphorimaging.

RNA cross-linking. 200 pmol (20 µM) of Not1–Not2–Not5c complex were incubated with 2.5 pmol (250 nM) of body-labeled U_{20} RNA overnight in EMSA buffer at 4 °C. The cross-linking was performed by irradiation of the mix at a wavelength of 254 nm for 5 min on ice. The mixture was then treated with 1% SDS and 0.5 µl of RNase A/T1 mixture at 37 °C for 5 min. The samples were heated with SDS loading dye at 70 °C for 2 min, separated on 13.5% (w/v) SDS-PAGE gel and visualized by phosphorimaging and Coomassie-blue staining.

Mass spectrometry. *UV-induced protein-RNA cross-linking and enrichment of cross-linked peptides.* UV cross-linking and enrichment of cross-linked peptides was performed according to the established protocols described in ref. 63. Briefly, 1 nmol of the single-stranded U_{15} RNA oligonucleotide and 1 nmol of

Not1c–Not2–Not5c complex were mixed in a 1:1 molar ratio, and the total reaction volume was brought to 100 µl in 20 mM HEPES, pH 7.5, 50 mM NaCl, 2 mM DTT and 5 mM EDTA. The mixture was incubated on ice overnight for complex formation. The samples were then transferred to black polypropylene microplates (Greiner Bio-One) and irradiated at 254 nm for 10 min. After ethanol precipitation, the samples were denatured in 4 M urea and 50 mM Tris-HCl, pH 7.9, and digested for 2 h at 52 °C with 1 µg RNase A (Ambion, Applied Biosystems). After RNA digestion, proteolysis with trypsin (Promega) was performed overnight at 37 °C. The sample was desalted on an in-house–prepared C18 (Dr. Maisch GmbH) column, and the cross-linked peptides were enriched on an in-house–prepared TiO_2 (GL Sciences) column with the protocol described in ref. 63. The samples were dried and then resuspended in 10 µl sample solvent (5% v/v ACN and 1% v/v FA) for MS analysis.

Nano–liquid chromatography and MS analysis. 5 µL of the above sample was injected onto a nano–liquid chromatography system (Agilent 1100 series, Agilent Technologies) including a C18 trapping column of length ~2 cm and inner diameter 150 µm, in line with a C18 analytical column of length ~15 cm and inner diameter 75 µm (both packed in house; C18 AQ 120 Å 5 µm, Dr. Maisch GmbH). Analytes were loaded on the trapping column at a flow rate of 10 µL/min in buffer A (0.1% v/v FA) and subsequently eluted and separated on the analytical column with a gradient of 7–38% buffer B (95% v/v acetonitrile and 0.1% v/v FA) with an elution time of 33 min (0.87%/min) and a flow rate of 300 nL/min. Online ESI-MS was performed with an LTQ-Orbitrap Velos instrument (Thermo Scientific), operated in data-dependent mode with a TOP10 method. MS scans were recorded in the *m*/*z* range of 350–1,600 and for subsequent MS/MS the top ten most-intense ions were selected. Both precursor ions as well as fragment ions were scanned in the Orbitrap. Fragment ions were generated by higher-energy collision dissociation (HCD) activation (normalized collision energy = 40) and recorded from *m*/*z* = 100. As precursor ions as well as fragment ions were scanned in the Orbitrap, the resulting spectra were measured with high accuracy (<5 p.p.m.), both in the MS and MS/MS level.

Data analysis. The MS .raw files were converted into the .mzML format with msconvert[64]. Protein-RNA cross-links were analyzed with OpenMS[65,66] and OMSSA[67] as a search engine. Data-analysis workflows were assembled as described[11]. The high-scoring cross-linked peptides were manually annotated for confirmation. Protein-RNA interactions between the complex and poly(U) RNA were analyzed with UV-induced protein-RNA cross-linking followed by MS. Peptide RCGNDFVYNEEDFEKL in Not5 (position 545–560) was observed carrying an additional mass of 476.0338 Da corresponding to U nucleoside with an adduct of 152. The y-ion series could be observed from 1 to 10, unshifted. In contrast, b ions from b3 until b8 were observed with a mass shift corresponding to U-H_3PO_4 and 152 adduct (**Fig. 5c**). Also, the b ions from b5 until b8 were observed with a mass shift corresponding to U and 152. We have always observed that the 152 adduct is observed as a shift associated with cysteine, which could be the amino acid that is cross-linked. In the corresponding figure (**Fig. 5c**), the b ions that were observed with a mass shift corresponding to U-H_3PO_4 + 152 and to U + 152 are shown with an asterisk (*) and hash (#), respectively, and the immonium ions with IM.

RNase protection assays. 100, 150 and 200 pmol (10, 15 and 20 µM) of Not1c–Not2–Not5c complex were incubated with 0.5 pmol (50 nM) of U_{20} RNA in the EMSA buffer overnight at 4 °C. The reaction mixtures were treated with 0.5 µl of RNase A/T1 mix for 30 min at 4 °C. RNA was purified with phenol/chloroform/isoamyl alcohol extraction and ethanol precipitation. The purified RNA was 5′ labeled with [γ-^{32}P]ATP with T4 polynucleotide kinase, repurified by phenol/chloroform/isoamyl alcohol extraction and ethanol precipitation, separated on 22% (w/v) denaturing PAGE with 5 M urea and visualized by phosphorimaging.

Fluorescence anisotropy. 5′-6-carboxy-fluorescein (6-FAM)-labeled U_{15} RNA was used in fluorescence anisotropy measurements at 20 °C with Genios Pro (Tecan). RNA was 9.1 nM at final concentration and was incubated with varying concentrations of Not1c–Not2–Not5c complex in the gel-filtration buffer supplemented with 10 mM EDTA. We used 250 mM NaCl in the buffer for the measurement because the protein was not stable at 100 or 150 mM salt in such high concentration at 20 °C. The excitation and emission wavelengths were 485 nm and 535 nm, respectively. Each titration point was measured three times with ten reads with an integration time of 40 µsec. The data were analyzed by nonlinear regression fitting with Origin (OriginLab; http://www.originlab.com/).

55. Kabsch, W. Integration, scaling, space-group assignment and post-refinement. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 133–144 (2010).
56. Adams, P.D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 213–221 (2010).
57. Cowtan, K. The Buccaneer software for automated model building. 1. Tracing protein chains. *Acta Crystallogr. D Biol. Crystallogr.* **62**, 1002–1011 (2006).
58. Emsley, P., Lohkamp, B., Scott, W.G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501 (2010).
59. Davis, I.W., Murray, L.W., Richardson, J.S. & Richardson, D.C. MOLPROBITY: structure validation and all-atom contact analysis for nucleic acids and their complexes. *Nucleic Acids Res.* **32**, W615–W619 (2004).
60. Larkin, M.A. *et al.* Clustal W and Clustal X version 2.0. *Bioinformatics* **23**, 2947–2948 (2007).
61. Bond, C.S. & Schüttelkopf, A.W. ALINE: a WYSIWYG protein-sequence alignment editor for publication-quality alignments. *Acta Crystallogr. D Biol. Crystallogr.* **65**, 510–512 (2009).
62. Gille, C. & Frömmel, C. STRAP: editor for structural alignments of proteins. *Bioinformatics* **17**, 377–378 (2001).
63. Kramer, K. *et al.* Mass-spectrometric analysis of proteins cross-linked to 4-thio-uracil- and 5-bromo-uracil-substituted RNA. *Int. J. Mass Spectrom.* **304**, 184–194 (2011).
64. Kessner, D., Chambers, M., Burke, R., Agus, D. & Mallick, P. ProteoWizard: open source software for rapid proteomics tools development. *Bioinformatics* **24**, 2534–2536 (2008).
65. Sturm, M. *et al.* OpenMS: an open-source software framework for mass spectrometry. *BMC Bioinformatics* **9**, 163 (2008).
66. Bertsch, A., Gröpl, C., Reinert, K. & Kohlbacher, O. OpenMS and TOPP: open source software for LC-MS data analysis. *Methods Mol. Biol.* **696**, 353–367 (2011).
67. Geer, L.Y. *et al.* Open mass spectrometry search algorithm. *J. Proteome Res.* **3**, 958–964 (2004).

doi:10.1038/nsmb.2686