Published online 10 December 2011 *Nucleic Acids Research, 2012, Vol. 40, No. 7* ***3259–3274***
doi:10.1093/nar/gkr1186

Structural and functional analysis of the archaeal endonuclease Nob1

Thomas Veith[1,2], Roman Martin[1,3], Jan P. Wurm[1,2], Benjamin L. Weis[1], Elke Duchardt-Ferner[1,2], Charlotta Safferthal[1], Raoul Hennig[1], Oliver Mirus[1], Markus T. Bohnsack[1,3], Jens Wöhnert[1,2,3,*] and Enrico Schleiff[1,3,4,*]

[1]Institute for Molecular Biosciences, [2]Center of Biomolecular Magnetic Resonance (BMRZ), [3]Cluster of Excellence Frankfurt: Macromolecular Complexes and [4]Centre of Membrane Proteomics, Johann-Wolfgang-Goethe University, Max-von-Laue Str. 9, 60438 Frankfurt, Germany

Received July 5, 2011; Revised November 11, 2011; Accepted November 14, 2011

ABSTRACT

Eukaryotic ribosome biogenesis requires the concerted action of numerous ribosome assembly factors, for most of which structural and functional information is currently lacking. Nob1, which can be identified in eukaryotes and archaea, is required for the final maturation of the small subunit ribosomal RNA in yeast by catalyzing cleavage at site D after export of the preribosomal subunit into the cytoplasm. Here, we show that this also holds true for Nob1 from the archaeon *Pyrococcus horikoshii*, which efficiently cleaves RNA-substrates containing the D-site of the preribosomal RNA in a manganese-dependent manner. The structure of PhNob1 solved by nuclear magnetic resonance spectroscopy revealed a PIN domain common with many nucleases and a zinc ribbon domain, which are structurally connected by a flexible linker. We show that amino acid residues required for substrate binding reside in the PIN domain whereas the zinc ribbon domain alone is sufficient to bind helix 40 of the small subunit rRNA. This suggests that the zinc ribbon domain acts as an anchor point for the protein on the nascent subunit positioning it in the proximity of the cleavage site.

INTRODUCTION

Ribosomes are important macromolecular complexes that synthesize proteins in all organisms. Bacterial ribosomes have been studied extensively, both structurally and functionally (1,2), and their subunits can be assembled *in vitro* using purified components (3). Besides ribosomal proteins and RNAs, ribosome assembly in bacteria is facilitated by a small number of non-ribosomal assembly factors *in vivo* (4,5). In contrast, ribosome biogenesis in eukaryotes requires a plethora of cofactors. The process has been mainly studied in the yeast *Saccharomyces cerevisiae*, where at least 75 small nucleolar RNAs and more than 200 proteins are involved (6–8). These non-ribosomal proteins include many enzymes and RNA-binding cofactors such as endonucleases and exonucleases required for the processing of preribosomal RNA (pre-rRNA).

Ribosome biogenesis in yeast is initiated in the nucleolus by the transcription of the 35S pre-rRNA, which contains the sequences of the 18S, 5.8S and 25S rRNAs (9). The mature rRNAs are generated in a sequence of endonucleolytic cleavage and exonucleolytic trimming reactions mediated by different nucleases located in the nucleolus, nucleus and cytoplasm. After initial cleavages at the positions A_0 and A_1, the cleavage at A_2 results in the separation of the biogenesis pathways of the small ribosomal subunits (SSU) and large ribosomal subunits (LSU) and produces the 20S (SSU) and $27SA_2$ (LSU) pre-rRNAs. Several biogenesis steps of the LSU occur before nuclear export, such as processing steps to generate the 6S intermediate and the mature 25S and the incorporation of the separately transcribed 5S rRNA, while the final trimming of 6S to generate the mature 5.8S is cytoplasmic (10).

The cleavage of the 20S pre-rRNA to the mature 18S rRNA also occurs after nuclear export to the cytoplasm (11). Several protein cofactors have been suggested to be

*To whom correspondence should be addressed. Tel: +49 69 79829287; Fax: +49 69 79829286; Email: schleiff@bio.uni-frankfurt.de
Correspondence may also be addressed to Jens Wöhnert, Tel: +49 69 79829276; Fax: +49 69 79829527; Email: woehnert@bio.uni-frankfurt.de

The authors wish it to be known that, in their opinion, the first three authors should be regarded as joint First Authors.

involved in this step in the yeast *S. cerevisiae* including the endonuclease Nob1, the putative kinase Fap7, the RNA methyltransferase Dim1, the GTPase Ltv1 and the RNA helicase Prp43 with its cofactor Sqs1. Nob1 catalyzes the cleavage of the 20S pre-rRNA at site D, while the role of Fap7 has remained unclear so far (12,13). The RNA helicase Prp43 and its G-patch protein cofactor Sqs1 have been shown to genetically interact with Nob1 and Ltv1 and are thought to mediate structural rearrangements that allow efficient cleavage by the endonuclease Nob1 (13–15). The methyltransferase Dim1 is required for dimethylation of two adenines close to the 3′-end of the 18S rRNA (16), while Ltv1 is thought to act in cofactor release after export of the pre-40S subunits (17). Assembly of the head domain of the SSU and integration of the ribosomal protein rpS5 has further been shown to influence pre-rRNA cleavage at site D (18).

Nob1 was originally identified in yeast as a protein associated with the proteasome (19), before its requirement for pre-rRNA processing was discovered (12). It was suggested to interact with the pre-SSU RNA in the vicinity of site D in yeast (20) and cross-links to helix 40 in the 20S pre-rRNA *in vivo* (21). Based on the analysis of sequence similarities, it has been proposed that Nob1 contains a PilT N-terminus (PIN) domain common to many other exonucleases or endonucleases (22,23) and a zinc ribbon domain (24). Yeast Nob1 has been shown to cleave pre-rRNA sequences at site D *in vitro* (13) and its enzymatic function has been suggested to be regulated by a conformational switch of the RNA and by Dim2 *in vivo* (15,25). Mutagenesis of residues in the catalytic core of the endonuclease, the proposed PIN domain, lead to accumulation of the 20S pre-rRNA and decrease in levels of mature 18S in yeast (26). In general, PIN domain proteins have been shown to possess endonucleolytic activity, e.g. the exosome component Rrp44 (27,28) and the NMD factor Smg6 (29).

Human Nob1 was also found to be involved in ribosome biogenesis (30,31). Furthermore, *nob1* was recently identified as one of the six marker genes to distinguish the chronic phase from the blast crisis of chronic myeloid leukemia (32) and as an oncogenic factor in ovarian cancer (33). It was also found to be upregulated in noise-injured cochlea and it is suggested to be involved in adaptation to an acoustic trauma (34).

In archaea, many components of the ribosome biogenesis pathway are thought to be conserved and to play similar roles in the process as in yeast. Several proteins required for 40S biogenesis are conserved in archaea, including Nep1, Fap7, Rio2 as well as the methyltransferase Dim1 and its associated protein Dim2. Their structural analysis unraveled fundamental principles of the molecular action of these factors (35–40), which were instrumental for understanding the function of the related factors in eukaryotes. Here, we characterize the Nob1 homolog from the thermophilic archaeon *Pyrococcus horikoshii* (PhNob1). The nuclear magnetic resonance (NMR)-solution structure of PhNob1 reveals the presence of an N-terminal PIN domain and a C-terminal zinc ribbon domain connected by a flexible linker. In contrast to eukaryotic Nob1 proteins, PhNob1 lacks a large insertion in the PIN domain and a long C-terminal tail. *In vitro,* the full-length protein endonucleolytically cleaves RNA substrates that resemble the D-site of the *P. horikoshii* pre-SSU RNA in a Mn^{2+}-dependent manner specifically at the site that corresponds to the mature 3′-end of the 16S rRNA. The zinc ribbon domain of PhNob1 is sufficient to bind helix 40 of *P. horikoshii* 16S rRNA. Based on the structure of PhNob1, we have mutated residues required for efficient and selective cleavage of the substrate RNA. These results show that both the domain structure and function of Nob1 are conserved in archaea and eukaryotes.

MATERIALS AND METHODS

Sequence search, alignment and phylogenetic analyses

Sequences were collected from PFAM (PF08772) (41) and by PSIBLAST from GenBank (42,43). Multiple sequence alignment was calculated with MAFFT v6.846b and a maximum likelihood phylogeny was reconstructed with RAxML v7 2.6 using gamma-distributed rate heterogeneity and the WAG substitution matrix (44–46). Branch support values were calculated with the rapid bootstrap method of RAxML (45) from 1000 bootstrap trees.

Protein and RNA constructs used in this study

The PhNob1-coding sequence was codon optimized for expression in *Escherichia coli* and the synthetic gene (Entelechon) was cloned (NdeI and BamHI) into pET11a. Mutants were generated by site directed mutagenesis. The PIN domain (amino acids 1–120) and Zinc ribbon motif (amino acids 121–161) were cloned into a pQE80 derivative generating a His10-ZZ-TEV-tag (47). RNA constructs for *in vitro* transcription were generated in a pGEM4Z vector (HindIII and EcoRI) and hairpin constructs were designed as described (13,28). RNA constructs used for *in vitro* cleavage assays contained the pre-rRNA sequence GGG AGA CAA GCU UAA GUC GUA ACA AGG UAG CCG UAG GGG AAC CUA CGG CUC GAU CAC CUC CUA UCG CCG GAA ACC CCG UCC GGG GGA AUU (PhNlong90), while the long and short hairpin loops included the sequences AUC ACC UCC UAU CGC C and CGA UCA CCU CCU AUC GCC, respectively. RNAs for gel shift assays, analytical gel filtration, fluorescence anisotropy and NMR-titration experiments were synthesized commercially (Dharmacon) and deprotected as described by the manufacturer. They were folded into monomeric species by denaturing them at 95°C for 5 min and subsequent dilution with ice-cold water.

Production and purification of recombinant proteins

Sample preparation for biochemical analysis. Proteins were expressed in 2YT medium for 3 h at 37°C. For purification, cells expressing PhNob1 were lysed in a buffer containing 50 mM Tris–HCl pH7.5, 100 mM NaCl, 0.1 mM $ZnCl_2$ and 0.1 mM PMSF, and nucleic acids were precipitated with 0.1% polyethylenimine pH 7.0, the supernatant was loaded onto SP-Sepharose and PhNob1 eluted with a linear salt gradient from 100 to 1000 mM

NaCl. Peak fractions were collected, pooled and buffer exchanged to 50 mM Bis–Tris pH 6.2, 50 mM KCl for NMR or biochemical assays. $PhNob1_{PIN}$ or $PhNob1_{Zn}$ were purified using NiNTA (Qiagen) followed by TEV cleavage using GST-tagged TEV-protease and removal of tags and TEV protease using Glutathion-Sepharose and NiNTA or size exclusion on a sephacryl S-100 column in 50 mM Tris pH 7.5, 120 mM NaCl for $PhNob1_{PIN}$ and $PhNob1_{Zn}$, respectively.

NMR sample preparation. For the expression of ^{15}N- and/or ^{15}N- and ^{13}C-labeled PhNob1, $PhNob1_{PIN}$ or $PhNob1_{Zn}$ M9 medium supplemented with 1 g/l $^{15}NH_4Cl$ and 2.5 g/l $^{13}C_6$-D-glucose as sole nitrogen and carbon source, respectively, was used. ^{2}H-, ^{13}C- and ^{15}N-labeling of $PhNob1_{PIN}$ was achieved using *E. coli*-OD2 CDN medium (Silantes GmbH) and expression for 6 h. Proteins were purified as above with subsequent gel filtration on a HighPrep Sephacryl S100 column on an Äkta purifier system (GE-Healthcare). NMR samples were prepared in 50 mM Bis–Tris, 50 mM KCl, pH 6.2, 10 % (v/v) heavy water (D_2O) at protein concentrations of 300–400 μM (^{15}N- and ^{13}C-labeled PhNob1 and $PhNob1_{Zn}$ and ^{2}H-, ^{13}C- and ^{15}N-labeled $PhNob1_{PIN}$) and 100–150 μM (^{15}N-labeled proteins for titration experiments), respectively. For titration, PhNlong23: GAU CAC CUC CUA UCG CCG GAA AC and PhNshort10: CCU CCU AUC G were used as substrates.

Endonuclease assays

Pre-rRNA fragments were transcribed *in vitro* using T7 polymerase (Fermentas) and labeled by incorporation of radiolabeled nucleotides or subsequent 5′-labeling. *In vitro* nuclease assays were performed as described (13). *Escherichia coli* lysates for RNA degradation controls were prepared as described for PhNob1 purification and used in different dilutions as indicated. Cleavage products were separated on a 8% PAA-8M urea sequencing gel and analyzed by autoradiography. Sequencing ladders were produced using a T7 sequencing kit (USB).

Structure determination

NMR-spectroscopy and resonance assignments. NMR-spectra were recorded at 37°C on Bruker AVANCE 600, 700, 800 and 900 MHz spectrometers equipped with cryogenic triple resonance probes. Chemical shifts were referenced to an external trimethylsilylpropionate standard dissolved in NMR-buffer. The standard set of triple resonance experiments was used for the assignment of PhNob1 backbone and side chain resonances as described elsewhere (48). In addition, the assignments for the backbone amide groups of lysines were verified using ^{15}N-heteronuclear single quantum correlation (HSQC) spectra of an amino acid selectively labeled PhNob1 sample prepared as described (49). A {^{1}H}- and ^{15}N-heteronuclear nuclear Overhauser effect (HetNOE)-experiment (50) for ^{1}H- and ^{15}N-labeled PhNob1 was recorded using a Bruker standard pulse sequence at 37°C with a recycle delay of 5 s and 32 scans per increment. All spectra were processed using Bruker TopSpin 2.1 and analyzed using the program CARA (51).

Structural constraints. Torsion angle constraints for the backbone torsion angles ϕ and ψ were generated by using TALOS+ for the analysis of backbone H, N, Cα, Cβ and CO chemical shifts (52) and used for all residues with {^{1}H}- and ^{15}N-HetNOE values >0.5. A long range 2D-H(N)CO-experiment with the NC-transfer delay set to 133 ms and with 128 scans per increment and 96 complex points in the indirect CO-dimension was recorded for the detection of hydrogen bonds in $PhNob1_{Zn}$ using ^{13}C- and ^{15}N-labeled protein. Similarly, a standard TROSY-3D-HNCO-experiment and a long range TROSY-3D-HNCO (53) with the NC-transfer delay set to 133 ms (32 scans per increment, 32 and 48 complex points in the indirect ^{15}N and ^{13}C dimensions, respectively) were recorded for identifying hydrogen bonds in $PhNob1_{PIN}$ using ^{2}H-,^{13}C- and ^{15}N-labeled protein. In both the experiments, appropriate deuterium decoupling was used. Hydrogen bonds detected in these long range HNCO-experiments on the isolated domains were converted into upper limit distance restraints with a H(N)–CO distance of 2.3 Å and an N–CO distance of 3.2 Å when the chemical shifts of amide groups involved were identical in the full-length protein and the isolated domain.

Nuclear Overhauser effect (NOE)-based distance restraints were derived from a 3D-^{15}N-edited nuclear Overhauser effect spectroscopy (NOESY)-HSQC experiment, 3D-^{13}C-edited 3D-NOESY-HSQC experiments with the ^{13}C offset and the transfer delays optimized for either the aliphatic or the aromatic carbons recorded in H_2O and a 3D-^{13}C-edited HSQC-NOESY-experiment in D_2O.

Structure calculations. The TALOS+-derived torsion angle constraints, the hydrogen bond constraints and the 3D-NOESY peak lists served as input for seven rounds of iterative NOE assignments and structure calculations using the ATNOS/CANDID module in UNIO (54,55) in combination with CYANA (56). This already resulted in a well converged bundle of structures with low target functions. The automatically obtained NOE-derived distance constraints were visually inspected and manually corrected in case of obvious artifacts or misassignments. Additional hydrogen bond constraints were introduced for amide protons which did not exchange against D_2O after 24 h in an H/D-exchange experiment and which formed hydrogen bonds in more than 15 out of the 20 initial UNIO structures with the lowest target function. The resulting edited NOE-derived distance constraints were then used together with the hydrogen bond and torsion angle constraints described above as input for a final round of structure calculation using constrained torsion angle molecular dynamics with CYANA (56).

The zinc atom was introduced in the final structure calculations via a covalent bond to the sulfur atom of C150 with a bond length of 2.3 Å and a bond angle (Zn-S-Cβ) of 96°. Distance constraints with upper and lower limits of 4.1 and 3.6 Å between the sulfur atoms of the four cysteine residues (C131, C134, C147 and C150) and 2.6 and 2.1 Å between the zinc atom and the sulfur atoms of the cysteine residues were used to enforce the tetrahedral coordination of the zinc atom (57). The 20 structures with the lowest target function out of the 200 structures calculated in this

final CYANA run were chosen to represent the solution structure of PhNob1 and minimized with the AMBER force field as implemented in OPAL (58). All figures of structures were prepared using MOLMOL (59) or UCSF-Chimera (60).

Structure comparison

Structures of PIN (scop:88723) and zinc ribbon domains (scop:57783 and scop:144206) were identified by their SCOP (61) classification and downloaded from the protein data bank (PDB) (www.pdb.org) (62). Structures were loaded in YASARA (www.yasara.org) and superimposed with its MUSTANG plugin (63) onto the domains of PhNob1 to calculate the root mean square deviation and sequence (RMSD) similarity.

Electrophoretic mobility shift assay

Wild type PhNob1 was incubated with radiolabeled helix 40, PhNlong23 or PhNlong90 RNA for 30 min in 50 mM Tris/HCl pH 7.5, 100 mM NaCl and 1 mM β-mercaptoethanol at 25°C. Complexes were separated on 6% native PAGE in 1× TBE (89 mM Tris, 89 mM boric acid and 2 mM ethylenediaminetetraacetic acid) and analyzed by phosphoimager.

Analytical gel filtration

Analytical gel filtration experiments for the detection of RNA-protein interactions were carried out using a Superdex S75 10/30 GL (GE Healthcare) gel filtration column with a flow rate of 1 ml/min on an Äkta purifier system (GE-Healthcare) in NMR buffer. As control a 50 μM sample of the pre-16S rRNA helix 40 mimic of *P. horikoshii* (H40) was applied to the column. Before loading the RNA was heated for 10 min and immediately cooled down on ice. For RNA-protein interactions, protein samples of PhNob1 and $PhNob1_{Zn}$ were added after the heating step. In all cases, the total applied volume was 100 μl. Elution was followed at $\lambda_1 = 280$ nm and $\lambda_2 = 260$ nm. Protein standards and PhNob1 and $PhNob1_{Zn}$ samples without RNA were applied to the system in a similar way without any heating step.

Anisotropy measurements

Fluorescence anisotropy measurements with 5′-fluorescein labeled helix 40 (5′-Fl-GACUGCCGGCGAUAAGCCG GAGGAAG-3′; Dharmacon), the shortened 5′-fluorescein labeled helix 40 (H40S 5′-Fl-CCGGCGAUAAGCCG G-3′; Dharmacon) or a stabilized H16 from *S. cerevisiae* 18S rRNA, where all mismatches were changed into Watson–Crick base pairs (ScH16 5′-Fl-UACAGGG CCCAUUGGGGCCCUGUA-3′; Dharmacon) were carried out at 25°C using a Fluorolog 3 (Horiba Jobin Yvon, Germany) spectrometer equipped with polarizers. Full-length PhNob1, $PhNob1_{PIN}$ and $PhNob1_{Zn}$ were titrated to 100 nM fluorescein labeled RNA in 100 mM NaCl, 25 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) and pH 7.0. Excitation and emission wavelength were set to 490 and 518 nm, respectively. Binding curves for full-length PhNob1 and $PhNob1_{Zn}$ were analyzed in Sigma Plot (Systat Software Inc.) by non-linear least square fitting to the equation: $R = R_0 + R_i c/(c+K_D)$ (R: anisotropy, R_0: anisotropy of free RNA, R_i: increase in anisotropy due to binding, c: protein concentration, K_D: dissociation constant) considering the large excess of protein.

RESULTS

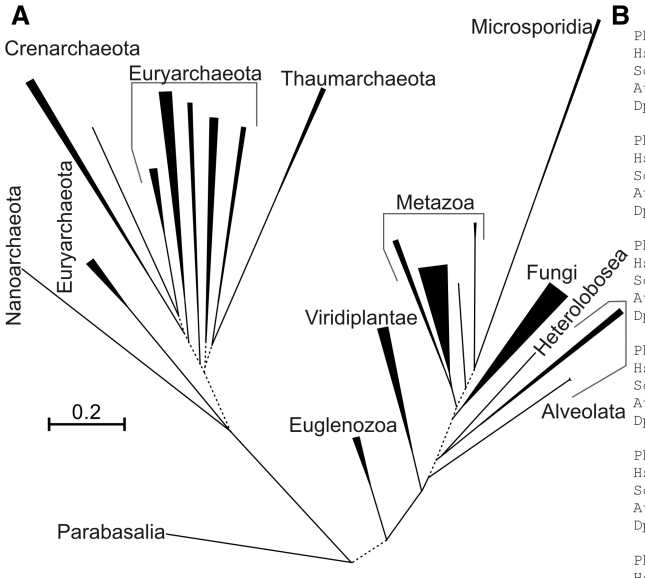
Identification of Nob1 homologs in archaea

Saccharomyces cerevisiae Nob1 has previously been shown to mediate processing of the 20S pre-rRNA at site D to generate the mature 18S rRNA (13). Remarkably, orthologous sequences of Nob1 can be identified in mammals, fungi, plants, protists and in many archaea (Figure 1A; Supplementary Figure S1 and Supplementary Table S1) suggesting that the RNA cleavage step catalyzed by this enzyme is largely conserved. We found Nob1 homologs in all of the major branches of archaea with the exception of the Korarchaeota branch; and even in *Nanoarchaeum equitans*, which is known to contain a highly compressed genome due to its parasitic lifestyle (67,68). In total, 82% of the sequenced archaeal genomes available in GenBank contained a Nob1-like protein as identified by a significant hit in a hidden Markov model search. The absence of Nob1 homologs in Korarchaeota, however, might be a result of the small number of genomes available for this phylum.

All identified Nob1 homologues apparently contain a putative PIN domain with highly conserved aspartate residues that are thought to be required for the endonucleolytic cleavage activity and a putative zinc ribbon domain with four highly conserved cysteine residues (Figure 1B). However, the archaeal and eukaryotic sequences vary significantly with respect to their lengths. All eukaryotic sequences contain a large insertion in the putative PIN domain and an additional long C-terminal tail of unknown structure and function, which are lacking in the archaeal proteins. To confirm the functional conservation, we analyzed an archaeal homolog. After testing a number of archaeal Nob1 homologs for expression yields, we chose Nob1 from the thermophilic archaeon *P. horikoshii* (PhNob1) for further studies. PhNob1 exhibits a sequence similarity of 68 and 66% to Nob1 from *S. cerevisiae* and *Homo sapiens* in the aligned regions, respectively. Thus, the *P. horikoshii* protein was expressed and purified without any additional tag for *in vitro* characterization and structure determination (Figure 1C and D).

PhNob1 binds Zn^{2+} via its zinc ribbon domain

Based on the presence of four conserved cysteine residues in their putative C-terminal Zn-ribbon domain, it was proposed that all Nob1 proteins bind a Zn^{2+}-ion, but Zn^{2+} binding has not been demonstrated experimentally so far. Therefore, the zinc content in the different purified proteins (Figure 1C and D) was determined to explore whether the Zn-ribbon domain coordinates this metal. Full-length PhNob1, the Zn-ribbon domain ($PhNob1_{Zn}$) and the PIN domain ($PhNob1_{PIN}$) were



B

```
PhNob1 M-------------------------------------LRNLKKTLVLDSSVFIQGI----
HsNob1 M-------------------------------------APVEHVVADAGAFLRHA--AL
ScNob1 M-----------------------------------TENQT-AHVRALILDATPLITQSYTHY
AtNob1 MDPKPTSMWSSIVKKDPPSKPPVNDGAPAAILGMVGNCKST-KGISIAVVDANAIIEGR-QSL
DpNob1 M-----------------------------GTLALTNEK-TTANFLIIDTNALISGV--RL

PhNob1 ---DIEGYTTPSVVEEIKDRESKIFLESLISAGKVKIAEPSKESIDRIIQVAKETGEVNELSK
HsNob1 QDIGKNIYTIREVVTEIRDKATRRRLAVL--PYELRFKEPLPEYVRLVTEFSKKTGDYPSLSA
ScNob1 QNYAQSFYTTPTVFQEIKDAQARKNLEIWQSLGTLKLVHPSENSIAKVSTFAKLTGDYSVLSA
AtNob1 TNFADKFVTVPEVLSEIRDPASRRRL-AFI-PFTIDTMEPSPESLSKVIKFARATGDLQSLSD
DpNob1 ETLGKELWTIPEVLEEVKDKKTSDFVTAF--PFEIKTRQPTPESVKAVIDFSMKTGDYPSLSA

PhNob1 ADIEVLALAYEL----------------------------------------------------
HsNob1 TDIQVLALTYQLEAEFVGVS-HLKQEPQ-KVKVSSSIQH--PETPLHISGFHLPYKPKPPQET
ScNob1 NDLHILALTYELEIKLNNGDWRLRKKPGDALDASKADVGTDGKQKLTEDNKKEEDSESVPKKK
AtNob1 VDLKLIALSYTLEAQVYGTK-NLRDVPPPIQTVRVKRLPEKDLPGWGS---------------
DpNob1 PDIKLIALAYTFEAEVNGTE-HIKTEPEKIQVTTSKPQQTSINSKSFNPSENNNN--------

PhNob1 ----------------------------------------------------------------
HsNob1 --EKGHSACEPENLEFSSFMFWRNPLPNIDHELQELLIDRGEDVPSE---------------E
ScNob1 NKRRGGK-------------------------------KQKAKREAREAREAENANLELESKAE
AtNob1 -----------NVANLEEWEALENETEEKSNANSKILPLKDLNMNIIASDNVSEVGSVVSHTE
DpNob1 ------------------------------------------NNNNTNKKEKKPKQPK

PhNob1 ----------------------------------------------------------------
HsNob1 EEEEE-----------------------------------NGFEDRKD-DS-----------
ScNob1 EHVEEAGSKEQ-----------------ICNDEN----IKESSDLNEVFED-----------
AtNob1 NHE---EDVQE-----------------GGKKHRRYPPKKTEIKLEG-KMVVEGVDASQGQY
DpNob1 EPKEPKEPKEPKEKNSKKNKLVVNLDKIEEEVSKEKVVDPNEVLRQLKQKQAEEAEERKKNRT

PhNob1 ----------------------------------------------------------------
HsNob1 --DDDGGGW--------------------------------------------------------
ScNob1 --ADDDGDW--------------------------------------------------------
AtNob1 DDDDDASDWRPAVSRSTHSKYLRRKARWEHYNALAEQEIQKDQEADKARHTKEANETHAKDSG
DpNob1 ISKEDEGEW--------------------------------------------------------

PHSD   ----------------------------------------------------------------
AtNob1 KNGEDISSILKDMRLEEESLRALQEETEETNAEATLINGEDDIDHDIEVEAEGIDVANQALEN

PhNob1 ----------------------------------------------------KGEIFSD
HsNob1 ----------------ITPSNIKQIQQEL--------EQCD-VPED---------VRVGCLTT
ScNob1 ----------------ITPENLTEAIIKDSGEDTTGSLGVEASEEDRHVALNRPENQVALATG
AtNob1 LEIASEAEDTFEASSIGDDG----SSEQS------WSLRALS---E---------SSVACITG
DpNob1 ----------------ITPDNISKVSTTYVQEKVH--------------------YDVGCITK

PhNob1 DYNVQNIASLLGLRFRTLKRG--IKKVIKWRYVCIGCGRK---FSTLPPGGVCPDCGSK-VKL
HsNob1 DFAMQNVLLQMGLHVLA-VNGMLIREARSYILRCHGCFKT---TSD-MSRVFCSHCGNK-TLK
ScNob1 DFAVQNVALQMNLNLMNFMSGLKIKRIRNYMLRCHACFKIFPLPKDGKPKHFCASCGQGQTLL
AtNob1 DYAMQNVILQMGLRLLA-PGGMQIRQLHRWILKCHACYTV---TPE-IGRIFCPKCGNGGTLR
DpNob1 DFSMQNVILQMGLHLIS-VDGVVIKQVKQFVLKCVACLNI---TTD-MEKIFCSHCGNK-SLY

PhNob1 IPRKR-------------------------------------------------------------
HsNob1 KVSVTVSD-DGTLHMHFSRNPKVLNPRGLRYSLPTPKGGKY-------------AINPH--LT
ScNob1 RCAVSVDSRTGNVTPHLKSN-FQWNNRGNRYSVASPLSKNSQKRYGKKGHVHSKPQENV-ILR
AtNob1 KVAVTIGA-NGAIIA--ACK-PRITLRGTQYSIPMPKGGRE-------------AITKNLILR
DpNob1 KATTYVDR-NGNQRVSVGSA-KQFNLRGTIFSIPKPKGGKK--------------HSDMIVT

PhNob1 ----------------------------------------------------------------
HsNob1 EDQRFPQLRLSQKAR---QKTNVFA--------PDYIAGVSPFVENDISSRSATLQVRDSTLG
ScNob1 EDQKEYEKVIKQEEWTRRHNEKILNNWIGGGSADNYI---SPFAITGLKQHNVRIGK------
AtNob1 EDQLPQKL----LHPRTKKKASKPG--------DEYFVS-DDVFLNHHSDR----KAPLQPPV
DpNob1 EEQYIHRLKVTGQYYKKKVSKEINL-------------------------DDLEMGFGHTGP

PhNob1 ------------------------------
HsNob1 AGRR-----RLNPN-ASRKKFVKKR-----
ScNob1 ---------GRYVN------SSKRRS----
AtNob1 RKAMSVFSQKRNPN-----DNHYSRSMH--
DpNob1 SDNIVIGYGNKNPNIARKRIGKKNKSISIF
```

Figure 1. PhNob1 sequence analysis and recombinant protein production. (**A**) The phylogenetic relation of Nob1 was calculated with RAxML v7.2.6 (64) as described (Supplementary Figure S1 and Supplementary Table S1). Dashed lines indicate low bootstrap values (<50%). The tree was prepared with MEGA (65) and filled triangles represent compressed subtrees. The triangle length and width are defined by the longest branch in a subtree and by the number of sequences, respectively. (**B**) The alignment of the amino acid sequence of the Nob1 protein from *P. horikoshii* (PhNob1, amino acids 1–161), *H. sapiens* (HsNob1, amino acids 1–412), *Arabidopsis thaliana* (AtNob1, amino acids 1–602), *Dictyostelium purpureum* (DpNob1, amino acids 1–429) and *S. cerevisiae* (ScNob1, amino acids 1–459) is shown. (**C**) The nomenclature, Nob1 domain annotation and PhNob1 constructs used in the manuscript are shown as bar diagrams. (**D**) The purified proteins used for biochemical characterization and structure determination were subjected to Tricin–PAGE (66) followed by Comassie blue staining.

purified and the zinc content analyzed by inductively coupled plasma mass spectrometry. A molar ratio of about 1 (protein to zinc) for PhNob1 and for $PhNob1_{Zn}$ (Table 1) was observed. In contrast, the ratio for $PhNob1_{PIN}$ was in the range of >60-fold protein excess over zinc (Table 1). These results clearly indicate that zinc is copurified with PhNob1 and the zinc-binding domain, as the enzyme was not recharged with zinc before the analysis. The high affinity of PhNob1 for zinc is consistent with our observation that bound zinc cannot be extracted *in vitro* by either EDTA or – N,N,N′, N′-*tetrakis*-(2-Pyridylmethyl)ethylenediamine (TPEN), which have reported affinities for Zn^{2+} of 10^{-14}–10^{-16} M, respectively (69).

Endonucleolytic activity of PhNob1

Purified PhNob1 (Figure 1) was tested in pre-rRNA cleavage assays using an *in vitro* transcript containing a pre-rRNA sequence covering the 3′-end of *P. horikoshii* SSU rRNA (16S rRNA), thus resembling site D of

Table 1. Zinc determination

Protein	MW (kDa)	Concentration		Protein:Zinc molar ratio
		Protein (μg/μl)	Zinc (ng/μl)	
PhNob1	17.9	3.5	12.4	1.02
$PhNob1_{Zn}$	4.9	0.05	0.74	0.90
$PhNob1_{PIN}$	13.8	1.4	0.11	59.9

pre-16S-rRNA. Cleavage of the internally labeled transcript resulted in two products, which occurred in a PhNob1 concentration-dependent manner (Figure 2A, internal). To determine which of them resembles the 5′-terminal fragment, 5′-end labeled RNA was used, resulting in the detection of only the longer cleavage product, again in a PhNob1 concentration-dependent manner (Figure 2A, 5′-labeled). This represents the 5′-fragment expected to result from cleavage at the 3′-end of the mature 16S rRNA. This was further confirmed by sequencing of PhNob1 cleavage products, the majority of which were processed at the position previously reported to represent the 16S rRNA 3′-end (12,13,20,26). The cleavage of the pre-rRNA is strictly dependent on the divalent cation manganese which cannot be substituted by Mg^{2+}-ions (Figure 2B). This finding parallels the observations previously made with yeast Nob1 (13). The observed nuclease activity is not the result of RNAse contamination by *E. coli* RNAses, because it is observed even after an extended heat shock (10 min, 75°C), which does not affect the activity of the thermophile protein (Supplementary Figure S3). In addition, incubation of substrate RNAs with *E. coli* lysates resulted in degradation patterns that did not resemble the specific cleavage products observed after incubation with purified PhNob1 (Supplementary Figure S3).

Specific hairpin constructs were designed to analyze whether the fragments obtained are generated by endonuclease activity or exonucleolytic digestion (Figure 3A) (13). These constructs carry sequences of the cleavage site in their loop structure and the ends are protected by a highly stable secondary structure. Incubation of the *in vitro* transcribed hairpin RNAs with recombinant PhNob1 resulted in distinct bands for both RNAs of sizes expected to result from endonucleolytic cleavage at the predicted site (Figure 3B). Thus, we conclude that *P. horikoshii* Nob1 is an endonuclease that can process pre-rRNA sequences by cleavage at the position corresponding to the 3′-end of the 16S rRNA.

PhNob1 consists of two structurally independent domains

To obtain structural insights, we performed NMR experiments. At first, we assigned the NMR signals of full length PhNob1. Isotopically labeled PhNob1 was expressed and purified and yielded high-quality ^{1}H, ^{15}N-HSQC spectra containing the expected number of backbone amide signals (Figure 4A). Essentially complete NMR assignments were obtained for the backbone and side chain

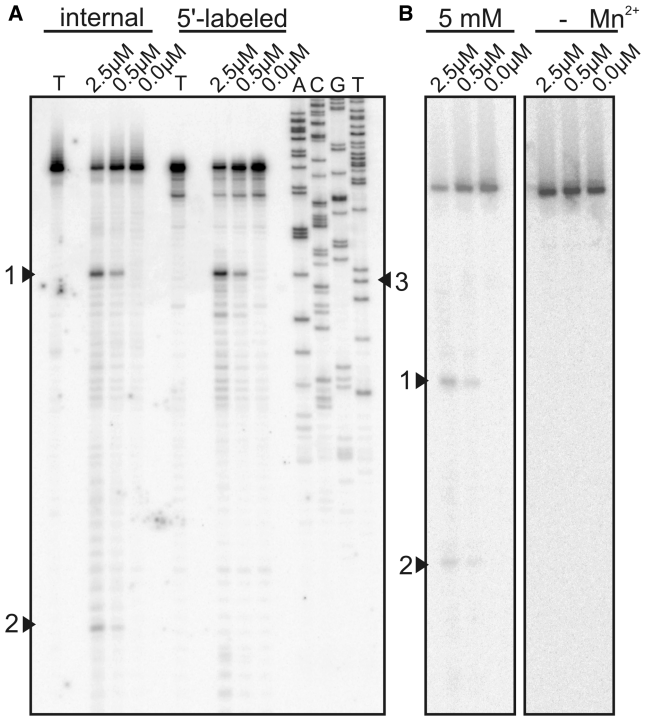


Figure 2. Processing of the 3′-end of *P. horikoshii* 16S rRNA is mediated by PhNob1. (**A**) *In vitro* transcripts of sequences from the 3′-end of *P. horikoshii* 16S rRNA and part of the internal transcribed spacer 1 (ITS1) were radiolabeled internally (left) or 5′-labeled (middle) and subjected to different concentrations of recombinant PhNob1. Cleavage of RNAs was then analyzed by denaturing SDS–PAGE followed by phosphorimaging. A sequencing ladder was prepared for size comparison (right). Arrows indicate the long (arrow 1) and short (arrow 2) cleavage products. Arrow 3 indicates the 3′-end of the 16S rRNA. Cleavage of internally labeled substrate results in two labeled products (arrows 1 and 2), while the 5′-phosphorylated substrate only results in labeling of the long cleavage product (arrow 1). T represents transcript subjected to electrophoresis without any previous treatment. (**B**) Recombinant PhNob1 was incubated with *in vitro* transcribed cleavage substrate as described for (A). Reactions were performed in the presence (5 mM) or absence (- mM) of manganese ions.

resonances of full-length PhNob1 (deposited in BioMagResBank (BMRB)—accession number 17595) (48). Chemical shift-based prediction of secondary structure elements using the programs PECAN (70) and TALOS+ (52) revealed an alternating pattern of five β-strands and six α-helices in the first 120 residues of PhNob1 and the presence of 2–3 short β-strands in the C-terminal part of the protein (48). The alternating pattern of five β-strands and 6 α-helices in the N-terminal portion of PhNob1 is in agreement with the presence of a PIN domain in PhNob1 as predicted based on sequence homologies. The Cα and Cβ chemical shifts of the four conserved cysteine residues in the C-terminal part of PhNob1 indicate that these cysteines are in their reduced form but are compatible with the complexation of a Zn^{2+}-ion (71) by these residues. Assignments for the backbone amide groups of $PhNob1_{Zn}$ and $PhNob1_{PIN}$ were obtained by transferring the assignment from PhNob1 due to the close similarity of their ^{1}H, ^{15}N-HSQC-spectra.

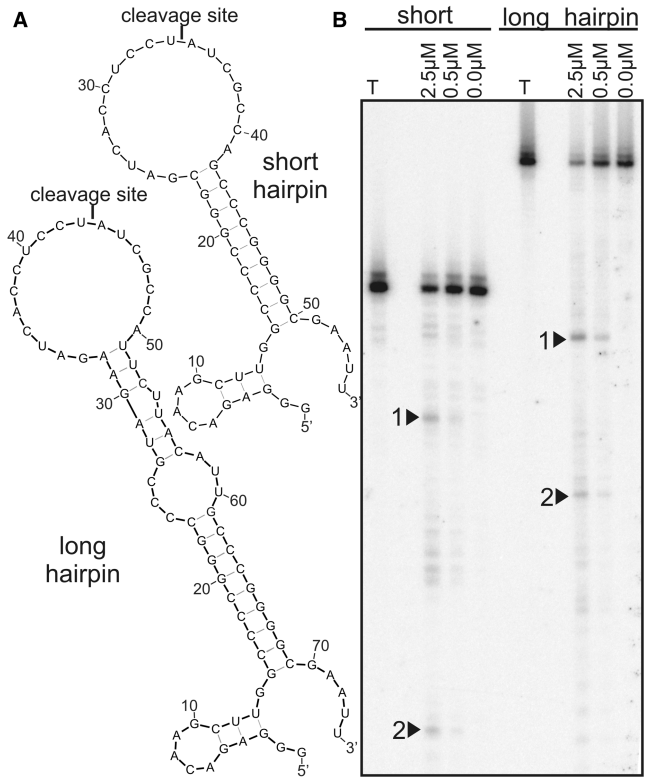


Figure 3. PhNob1 exhibits endonucleolytic activity. (**A**) Secondary structure prediction of the hairpin constructs used. (**B**) Radiolabeled *in vitro* transcripts of long (HP/L) or short (HP/S) hairpin constructs with loops containing sequences of the 3′-end of 16S rRNA were incubated with PhNob1 to analyze endonuclease activity. Analysis was performed as described for Figure 2.

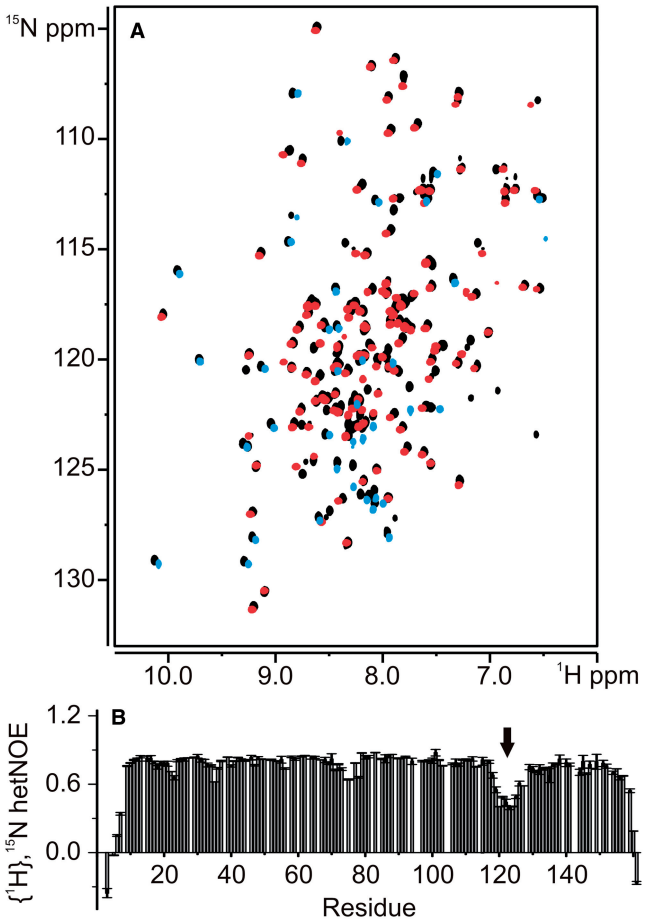


Figure 4. PhNob1 consists of two structurally independent domains connected by a flexible linker. (**A**) Overlay of ^{15}N-HSQC spectra of PhNob1 (black), PhNob1$_{PIN}$ (red) and PhNob1$_{Zn}$ (blue) with the assignments for backbone amid groups of full-length PhNob1 indicated. The two isolated domains exhibit virtually identical chemical shifts for backbone amide resonances in comparison with the full-length protein. (**B**) Plot of the measured {^{1}H}- and ^{15}N-hetNOE values for full-length PhNob1 against residue number. HetNOE values <0.5 point to an increased flexibility of the respective residues. Besides residues at N-termini and C-termini the amino acids 118–126 corresponding to the predicted domain boundary between PIN and zinc ribbon domain show strongly reduced hetNOE values. The position of the putative flexible interdomain linker is indicated by the black arrow.

A {^{1}H}- and ^{15}N-heteronuclear single quantum correlation (HetNOE)-experiment demonstrated increased flexibility of amino acids 120–125 (Figure 4B) in agreement with the presence of a flexible linker between the N-terminal and the C-terminal domain. The backbone chemical shifts of the linker amino acids have random coil values. The isolated N-terminal (amino acids 1–120) and C-terminal fragments (amino acids 121–165) yielded ^{1}H, ^{15}N-HSQC-spectra typical for well folded, globular proteins. The overlay of the HSQC-spectra for these two domains and for the full-length PhNob1 showed that the HSQC spectrum of PhNob1 is essentially the sum of the spectra of the two isolated domains (Figure 4A). This indicates that the structures of the PIN and the zinc ribbon domains are highly similar in isolation and in the context of the full-length protein. Thus, PhNob1 consists of two structurally independent domains connected by a flexible linker (amino acids 119–126).

Solution structure of PhNob1

The final solution structure of PhNob1 was calculated using 228 torsion angle constraints, 18 distance constraints to ensure the presence of a tetrahedrally coordinated zinc ion (57) and 3542 nuclear Overhauser effect spectroscopy (NOESY)-based distance constraints. The direct analysis of the hydrogen bonding patterns in 2D and 3D-long range HNCO experiments on the two isolated domains where signal overlap is reduced yielded a total of 34 hydrogen bonds (Figure 5A). The presence of the same hydrogen bonds in the full-length protein was verified by analyzing 3D-NOESY-HSQC and H/D-exchange experiments (Figure 5B). The structure calculation resulted in a set of converged structures with low target functions and reasonable stereochemistry even in the absence of the hydrogen bonding constraints derived from the HNCO-experiments. However, hydrogen bonding constraints were included in the final rounds of structure calculations (Table 2; the 20 best structures are deposited in PDB—2lcq).

In agreement with the presence of a flexible interdomain linker and the structural independence of the two domains, no NOEs were found between

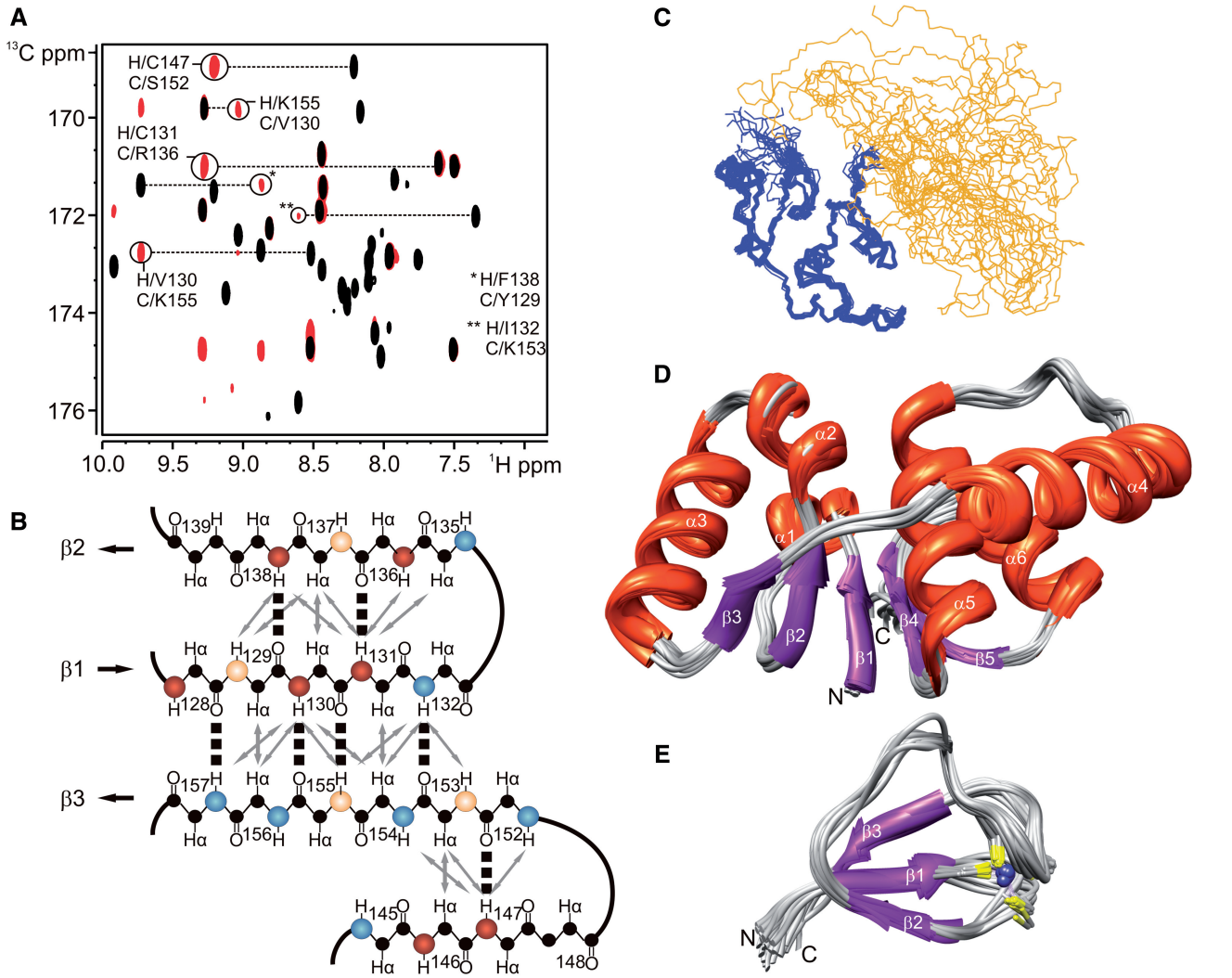


Figure 5. Solution structure of PhNob1. (**A**) Overlay of a long range 2D-H(N)CO-spectrum (red) and a standard 2D-H(N)CO spectrum (black) recorded for the isolated zinc ribbon domain for direct identification of hydrogen bond donor and acceptor groups. Identified hydrogen bonds are indicated. (**B**) Secondary structure and hydrogen bonding scheme for the zinc ribbon domain as derived by the long range H(N)CO-experiment. Directly identified hydrogen bonds are indicated by bold-dashed lines between donor hydrogen and acceptor CO group. Hydrogen bonds identified in the isolated domain are present in full-length PhNob1 as indicated by the corresponding NH–NH and NH–Hα–NOEs (double-edged grey arrows) and the observed protection patterns in an H/D-exchange experiment. Positions of slowly exchanging NH-protons in the full-length protein are marked by orange (NH present after 6 h) or red (NH present after 24 h) dots. Fast exchanging NH-protons are colored in light-blue. (**C**) Bundle of the 20 structures with the lowest target function aligned using the backbone heavy atoms of residues 11–117 of the PIN domain. The PIN domain is colored dark blue, the zinc ribbon domain is colored gold. (**D**) Overlay of the PIN domain (residues 11–117) taken from the 20 structures with the lowest target function in cartoon representation. β-strands are colored magenta whereas α-helices are colored orange. (**E**) Overlay of the zinc ribbon domain (residues 129–159) taken from 20 structures with lowest target function in cartoon representation. The secondary structure elements are annotated. The position of the bound Zn^{2+}-ion assuming tetrahedral coordination is indicated (blue sphere). The four metal-coordinating cysteine residues are shown as sticks with their sulfur atoms colored in yellow.

amino acid residues of the two domains despite careful manual checking. Accordingly, the orientation of the two domains with respect to each other is not defined by our NMR data (Figure 5C). However, the structures of the two domains themselves are well defined as judged by the backbone heavy atom root mean square deviations (RMSDs) of 0.51 Å and 0.52 Å for the PIN domain (residues 9–118) and the zinc ribbon domain (residues 127–159), respectively (Figure 5D and E).

The structure of the PIN domain (Figure 5D) reveals a central five-stranded β-sheet with all five strands in a parallel orientation. The order of strands in this β-sheet is β5↑–β4↑–β1↑–β2↑–β3↑. The loops connecting the β-strands with each other include either one or two α-helices of different lengths. Helix α1 is located between strand β1 and β2 and helix α6 between β4 and β5, whereas two α-helices are inserted between β2 and β3 and β3 and β4, respectively. Two sets of three α-helices each (α1–α3 and α4–α6, respectively) form a helical bundle on either side of the 5-stranded β-sheet. In each of these bundles, one helix (α1 and α4, respectively) is packed perpendicular against the other two, which are either oriented parallel (α5 and α6) or antiparallel with respect to each other.

Table 2. NMR and refinement statistics for the 20 lowest energy structures of PhNob1

NMR constraints	
Total NOE distance constraints	3542
Intraresidue	651
Sequential	931
Medium range	639
Long range	1321
Hydrogen bond	47
Zinc coordination	18
Dihedral restraints (Talos+)	228
Structural statistics	
Target function value	3.96 ± 0.44
Distance constraints (Å)	0.013 ± 0.001
Maximum distance constraint violation (Å)	0.34 ± 0.09
Dihedral angle constraints (°)	0.61 ± 0.08
Maximum dihedral angle violation (°)	3.90 ± 0.96
Amber energies total [kcal/mol]	−4750 ± 239
RMSD from idealized geometry	
Bond length (Å)	0.0147 ± 0.0001
Bond angles (°)	1.82 ± 0.03
Residues in most favored regions	84.9%
Residues in additionally allowed regions	13.6%
Residues in generously allowed regions	1.1%
Residues in disallowed regions	0.5%
RMSD values (Å)	
Backbone (residues 9–118)	0.51
Backbone (residues 127–159)	0.52
Heavy atoms (residues 9–118)	1.19
Heavy atoms (residues 127–159)	1.18

The overall twist of the 5-stranded parallel β-sheet is approximately 160°. The 5-stranded parallel β-sheet, the strand order of the sheet and its decoration with α-helices concur with the expected consensus structure of PIN domains (Figure 5D) (22,72,73).

The second domain contains a three-stranded antiparallel β-sheet (Figure 5E). The topology of this β-sheet can be described as: β3↑–β1↓–β2↑. The four zinc-coordinating cysteine residues are located at the C-terminus of β1, in the loop connecting β1 and β2 and in the long irregular loop connecting β2 and β3 and brought into close spatial proximity even when the structure is calculated without constraints between the sulfur atoms of the four cysteine side chains and the putative zinc ion.

Mapping of the catalytic center and the substrate binding site of PhNob1

The endonucleolytic cleavage activity of PhNob1 is dependent on the presence of divalent manganese ions (Figure 2). Manganese ions are paramagnetic and induce line broadening of NMR signals leading to disappearance of amide signals of amino acids in the vicinity of Mn^{2+}-binding sites. To locate the active site of PhNob1 for endonucleolytic cleavage, ^{1}H, ^{15}N-HSQC-spectra of ^{15}N-labeled PhNob1 were recorded without and with addition of up to 50 μM $MnCl_2$. At a concentration of 10 μM, Mn^{2+} signals for the amino acids D12, S13, S14, V15, L78, K80, A81, I83, V85, L86, D99, D100, Y101 and N102 could no longer be detected (an example is given in Figure 6A). When mapped on the surface of the protein 3D structure, these residues are clustered on a defined surface of the PIN domain of PhNob1, which is in agreement with the presence of a single, well-defined Mn^{2+}-binding site (Figure 6B and C). Three of the affected aspartates (D12, D82 and D100; Figure 6B) are conserved in all archaeal and eukaryotic sequences of Nob1 (Figure 1). The 3D structure of the domain brings the aspartate side chains of D12 at the end of β1, D82 at the N-terminus of α5 and D99 and D100, which follow β5 in close spatial proximity. Two of these amino acids, namely D12 and D82 were already implicated as necessary for D-site cleavage activity by mutational studies (13,26) in ScNob1 (scD15 and scD92). Notably, no line broadening was observed for signals of amino acid residues in the C-terminal zinc ribbon domain, which is apparently not involved in Mn^{2+}-binding.

To map amino acids possibly involved in substrate binding we titrated ^{15}N-labeled PhNob1 with RNAs of different lengths (10 and 23 nt, PhNlong23 and PhNshort10, see 'Materials and Methods' section) containing the D-site sequence as well as with single stranded oligo-U-RNA (U_9). We monitored the resulting chemical shift changes in ^{1}H, ^{15}N-HSQC-spectra (Figure 6B). Only limited and gradual chemical shift changes were observed upon addition of increasing amounts of all RNAs either in the absence or presence of Mg^{2+} as divalent cation. This is typical for the formation of low-affinity RNA–protein complexes in fast exchange on the NMR time scale. Notably, all affected signals (S13, S14, K70, S79, D99, D100, V103, L109 and R115-K118) correspond to amino acids in the N-terminal PIN domain. No chemical shift changes were observed for the C-terminal zinc ribbon domain despite their strongly basic amino acid composition. Mapping of the affected residues on the 3D structure of the PIN domain reveals that they line the manganese binding pocket as expected for a genuine substrate binding site (Figure 6C).

The fact that U_9 induces similar albeit smaller chemical shift changes as the genuine substrates for endonucleolytic cleavage indicates that substrate recognition by the PIN domain itself is only marginally specific for short RNA fragments. The low affinity for the short substrates observed in the NMR-experiments is in agreement with the K_D estimated by electrophoretic mobility shift assays for PhNlong23, which is larger than 30 μM (Supplementary Figure S4). However, the longer substrate used in the cleavage assays (PhNlong90) is bound with a higher affinity reflected by a dissociation constant of about 0.7 μM estimated from electrophoretic mobility shift assays (Supplementary Figure S4). The same holds true for the isolated PIN domain, which binds with a higher affinity reflected by a lower dissociation constant ($K_D = 5\,\mu M$) to PhNlong90 than to PhNlong23 ($K_D > 60\,\mu M$; Supplementary Figure S4). However, the increase of the dissociation constant by a factor of 10 for PhNlong90 and by a factor of 2 for PhNlong23 by omitting the zinc ribbon domain suggests that the latter contributes to the interaction *in vitro,* but to a smaller extent for the shorter RNA. The RNA Phlong90, however, is unfortunately unsuitable for NMR-titration experiments, due to its size.

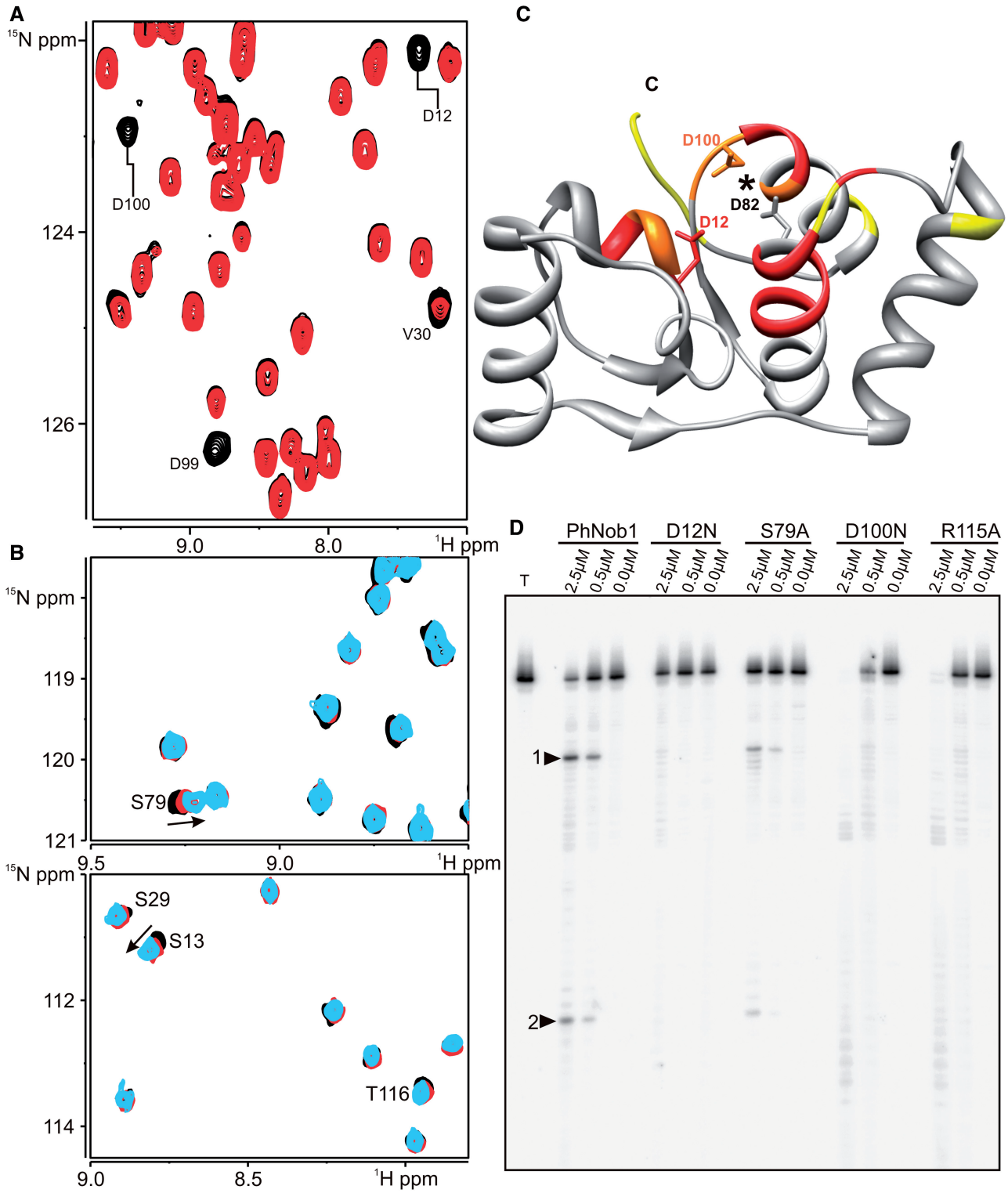


Figure 6. Identification of the active center and the substrate binding site of PhNob1. (**A**) Overlay of a selected section of the ^{1}H, ^{15}N-HSQC spectra of PhNob1 in the absence (black) or presence (red) of 10 μM Mn^{2+}-ions. The signal assignments for affected amino acids are given. (**B**) Overlays of selected regions of ^{1}H, ^{15}N-HSQC spectra of PhNob1 in the absence (black) and in the presence of 1.2 (red) or 2.4 (light blue) equivalents of substrate RNA. Affected signals are indicated. (**C**) Mapping of the effects of Mn^{2+} or substrate RNA addition on the structure of the PIN domain. Amino acids affected by Mn^{2+} are colored red, those affected by RNA are colored yellow and those affected by both are shown in orange. The three indicated aspartates are shown in stick representation. D82 (asterisks) cannot be analyzed due to signal overlap in the HSQC-spectrum but both adjacent amino acids are influenced by Mn^{2+}. (**D**) Wild-type or mutant PhNob1 were recombinantly produced, purified and incubated with radiolabeled *in vitro* transcripts as described for Figure 2. The expected migration of the cleavage products is indicated as in Figure 2.

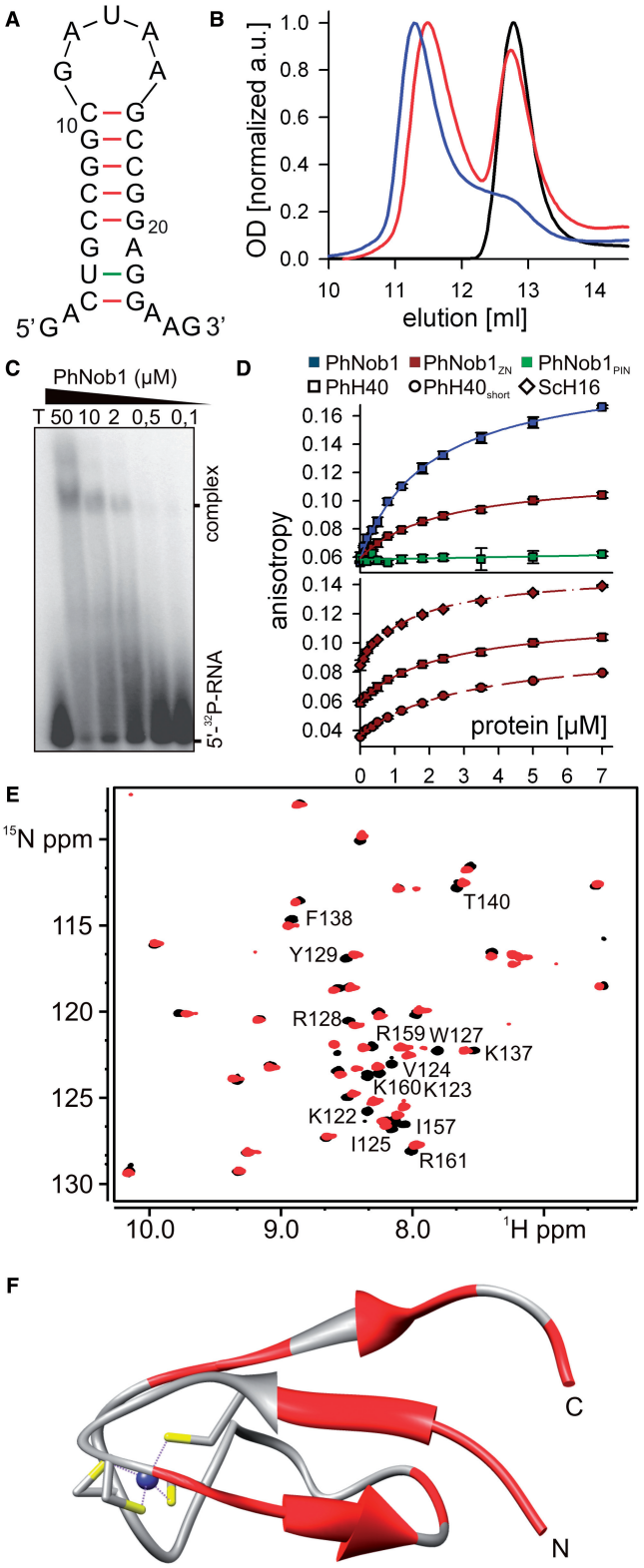


Figure 7. The interaction of PhNob1 with helix 40. (**A**) Secondary structure prediction of the 26-mer helix 40 RNA used for the binding studies. (**B**) Binding of recombinant PhNob1 to helix 40 RNA was probed by gel filtration analysis. The gel filtration profile determined at 260 nm of helix 40 is shown in black, the corresponding profiles of a approximately 1:1 and an approximately 1:4 mixture of helix 40 with PhNob1 are shown in red and blue, respectively. (**C**) PhNob1 was incubated at indicated concentrations with 5′-labeled synthetic *P. horikoshii* helix 40 rRNA and binding was analyzed by detection of the electrophoretic mobility shift on a native polyacrylamide gel. (**D**) The results of the fluorescence anisotropy measurements with full-length PhNob1 (blue), the PIN domain ($PhNob1_{PIN}$, green) and the Zn ribbon domain ($PhNob1_{Zn}$, red) to the H40 from *P. horikoshii* (square), a 5-bp stem–loop containing variant of PhH40 (circle) and the stabilized H16 from *S. cerevisiae* 18S RNA (diamond) are shown. (**E**) The chemical shift perturbations caused by 1.2 equivalents of helix 40 on the ^{15}N-HSQC-spectrum of the isolated zinc ribbon domain were recorded and the spectra of the protein in the absence (black) or presence of RNA are shown (red). The amino acid residues with the largest chemical shift changes upon RNA binding are labeled. (**F**) Residues labeled in (**E**) are colored red on the structural cartoon of the isolated zinc ribbon domain.

The PIN domain is required for RNA cleavage

To further test the functional relevance of the amino acids identified as part of the catalytic center and the substrate binding site (Figure 6C), some of these residues in the PIN domain were mutagenized. It was previously shown that the D15N mutation in yeast abolished Nob1 cleavage activity (13). The corresponding residue in *P. horikoshii* Nob1 is the aspartate at position 12. Mutagenesis of this residue (D12N) resulted in a complete loss of the catalytic activity of PhNob1 (Figure 6D). The mutations of serine 79 (S79A), which showed the largest chemical shift change upon substrate addition, did result in a significant reduction of cleavage which is in-line with its proposed role for substrate binding. When analyzing the activity of the PhNob1 carrying a mutation of D100 to N or R115 to A ($PhNob1_{D100N}$ and $PhNob1_{R115A}$), we observed unspecific degradation of the substrate. This is in agreement with the observation that these residues are involved in substrate binding (Figure 6), and the data suggest that these residues are playing a role in the correct positioning of the substrate with respect to the catalytic center. This indeed explains the loss in PhNob1 specificity for the conventional cleavage site. In summary, the results obtained for the different mutants support the substrate binding site mapped by NMR spectroscopy.

The C-terminal zinc ribbon domain is sufficient for binding to helix 40 of the SSU RNA

It was recently described that yeast Nob1 cross-links to helix 40 of 20S pre-rRNA (21). However, the addition of short substrate RNA did not induce chemical shift changes in the C-terminal zinc ribbon domain although this domain harbors the majority of basic amino acid residues of PhNob1. Thus, we speculated that the C-terminal domain might be important for binding to other rRNA elements apart from the D-site and might be involved in anchoring PhNob1 in the nascent ribosomal subunit. To substantiate this notion, we used a chemically synthesized helix 40 RNA from *P. horikoshii* (Figure 7A) to analyze the ability of the heterologously produced protein to interact with RNA *in vitro*. Using analytical gel filtration, we observed a decrease in the retention time of helix 40 RNA (Figure 7B; black line) in the

Table 3. Dissociation constants determined by anisotropy measurements

Protein	RNA	K_D [μM]
PhNob1	PhH40	2.0 ± 0.1
$PhNob1_{Zn}$	PhH40	2.3 ± 0.1
	$PhH40_{short}$	3.1 ± 0.2
	ScH16	1.5 ± 0.2
$PhNob1_{PIN}$	PhH40	>1000

presence of one (red) or four (blue) equivalents of full-length PhNob1. In contrast, oligoU (U_9) shows no shift in its retention time in similar experiments. This is consistent with the concentration-dependent formation of a helix 40 RNA/full-length PhNob1 complex observed by gel-shift analysis (Figure 7C).

To analyze whether the zinc ribbon domain alone is sufficient to mediate the interaction with helix 40 which was observed in the gel shift and gel filtration experiments with the full-length protein, the interaction of the helix 40 RNA (PhH40) with the isolated domains and the full-length protein was analyzed by fluorescence anisotropy measurements using 5′-fluoresceine-labeled RNA (Figure 7D). For full-length PhNob1, we observed a K_D value of 2.0 ± 0.1 μM (Table 3). While essentially no interaction was observed between PhH40 and $PhNob1_{PIN}$, the dissociation constant for the isolated zinc ribbon domain is similar to that of the full-length protein (Figure 7D top, Table 3). The dissociation constant observed for the zinc ribbon domain is sensitive to stem length as the dissociation constant increased slightly from 2.3 to 3.1 μM, when the 9 bp stem of PhH40 (Figure 7A) was shortened to a 5 bp stem of $PhH40_{short}$ (Figure 7D, bottom, square versus circle). In turn, the dissociation constant observed for the interaction of the zinc ribbon domain and the stabilized H16 from *S. cerevisiae* containing 10 bp is 1.5 μM, which is moderately lower than for PhH40 (Table 3 and Figure 7D, bottom, diamond).

The interaction between the isolated zinc ribbon domain and PhH40 was additionally analyzed by titration of the RNA and monitoring the chemical shift changes in the ^{1}H, ^{15}N-HSQC-spectra. Changes in the NMR spectra were only observed until an RNA–protein ratio of 1:1 was reached (Figure 7E). Furthermore, there were no gradual chemical shift changes but new signals appeared in the spectra upon addition of the RNA, whereas signals belonging to the free protein disappeared. Thus, the zinc ribbon domain forms a specific complex with PhH40 in slow exchange on the NMR-time scale indicative of a K_D for the RNA–protein interaction is approximately $\leq$1μM, which is in agreement with the dissociation constant measured by fluorescence anisotropy (Figure 7D). The amino acids affected by by PhH40 addition include many of the basic residues of the zinc ribbon domain, parts of the interdomain linker and the C-terminal portion of the domain; they form a continuous surface on the 3D structure of the domain located at the flat side of the β-sheet (Figure 7F).

DISCUSSION

D-cleavage of the SSU rRNA is conserved

Yeast Nob1 has been shown to mediate the final processing of the SSU rRNA in the nascent small ribosomal subunit by endonucleolytic cleavage of the 20S pre-rRNA. Remarkably, this enzyme is phylogenetically conserved in all eukaryotes and most archaea (Figure 1). All identified Nob1 homologs contain a PIN and a zinc ribbon domain. However, sequences of the eukaryotic kingdom contain an extended PIN domain–intrinsic loop and a C-terminal extension following the zinc ribbon domain (Figure 1). The length, sequence and the predicted secondary structure of the PIN domain–intrinsic loop vary significantly between the three different eukaryotic lineages (Supplementary Figure S2). In all cases, the N-terminus and C-terminus of the intrinsic loop contain a number of conserved acidic and hydrophobic residues suggesting that it might play a role in additional interactions with basic ribosomal or pre-ribosomal proteins not occurring or not required in archaea. The C-terminal extension of the eukaryotic Nob1 proteins contains a number of conserved charged and predominantly basic residues as well as conserved hydrophobic amino acids consistent with a role of the extension in RNA binding. The general lack of conservation in the eukaryotic extension segments among different eukaryotic lineages suggests that these extensions are required either for very specialized functions of Nob1 or a fine tuning of Nob1′s functionality to the specific environment in different organisms. The archaeal Nob1 seems to represent the minimal functional core required for its function in ribosome biogenesis.

Despite its generally smaller size compared to its eukaryotic counterparts and to the biochemically characterized yeast protein in particular, the *P. horikoshii* Nob1 homolog specifically cleaves RNA substrates containing a sequence surrounding the D-site of *P. horikoshii* SSU rRNA in an endonucleolytic manner specifically at the predicted site (Figures 2 and 3). As reported for the yeast protein, the reaction depends on the presence of Mn^{2+} ions, but the *in vitro* cleavage reaction appears to be much more efficient in comparison with the yeast protein (Figure 2) (13,15) even when carried out at 37°C, which is not the optimal temperature for a thermophilic enzyme. The presence of Nob1 homologs in archaea and the observed cleavage specificity *in vitro* strongly suggest that the 3′ terminal maturation of the pre-SSU-rRNA by processing at the D-cleavage site is conserved.

The structural basis for Nob1 function

The three-dimensional structure of PhNob1 reveals a PIN domain with a canonical structure and a zinc ribbon domain with a three-stranded β-sheet (Figure 5). The two domains are structurally independent from one another and connected by a flexible linker suggesting that the orientation of the two domains with respect to each other is flexible and can be influenced by interactions with ribosomal RNA, ribosomal proteins or other ribosome assembly factors.

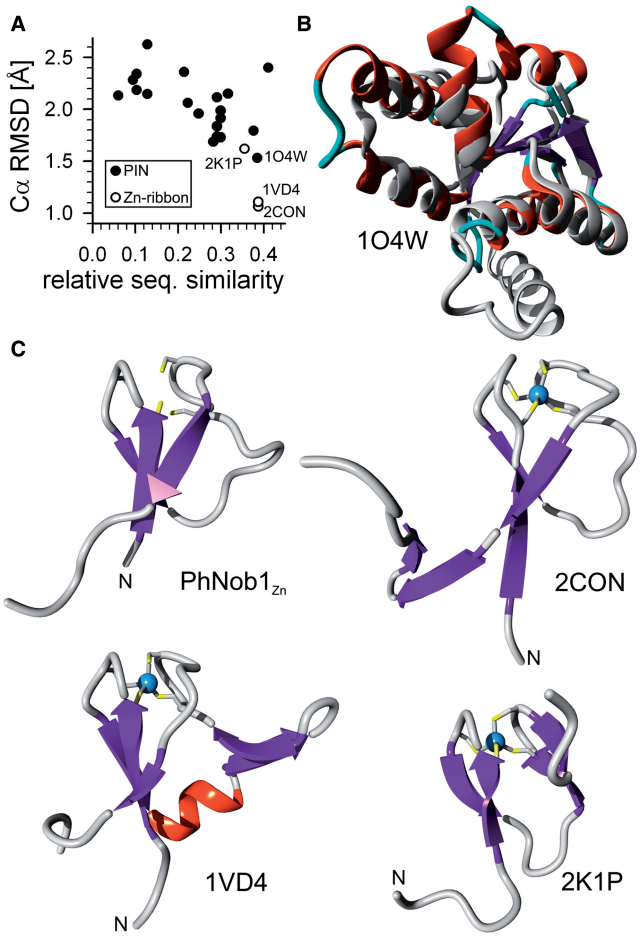


Figure 8. Comparison of the determined structure with known PIN and Zn-binding domains. (**A**) PIN and Zn ribbon domains from the PDB were structurally aligned with the corresponding domains of PhNob1 using the MUSTANG plugin in YASARA (www.yasara.org). The output consists of the alignment from which the sequence similarity and the Cα RMSD [Å] were calculated, which are presented for the PIN domains (full circle) or the Zn ribbon domains (open circle). (**B**) The overlay of the structures of $PhNob1_{PIN}$ (gray) and 1o4w (secondary structures are colored) is shown as a cartoon representation. (**C**) Representative structure of the deposited NMR-ensembles of $PhNob1_{ZN}$, 2con, 2k1p (structurally alike 3g9y) and 1vd4 are shown with the same color code as in B. For all structures the same orientation according to the structural alignment is shown.

The PIN domain is structurally most similar to the canonical PIN domain of AF0591 annotated as succinyl-CoA synthetase β-chain and not containing a Zn-ribbon (Figure 8A, PDB:1O4W) (74) from *Archaeoglobus fulgidus* with an Cα RMSD of 1.53 Å and a sequence similarity of 38.5%. The β-sheet of both the proteins is very similar although β3 is somewhat shorter in AF0591. With respect to the overall fold AF0591 contains an additional helix between α1 and β2, but does not contain α4 present in $PhNob1_{PIN}$ (Figure 8B). Unfortunately, the function of this factor is currently unknown and thus, no functional comparison of the two proteins is possible.

The insertion point of the eukaryote-specific loop in the PIN domain would be located between α-helix α5 and β-sheet β4. Remarkably, the loop does not share significant sequence similarities between fungi, mammals or plants, although the region is generally enriched in negatively charged amino acids. Even within these groups only a moderate similarity between the sequences restricted to the N- and C-terminal portion in fungi and mammals is observed (Supplementary Figure S2). In plants, the loop is significantly enlarged and in contrast to the other sequences two clusters of positive charges can be identified (Supplementary Figure S2). Up to now, no other PIN domains that carry an extension with respect to the canonical PIN domain fold at the same position as the eukaryotic Nob1s have been described at structural level.

The topology of the zinc ribbon domain is unusual (Figures 5 and 7). Only two other zinc ribbon domains with a similar topology were identified in the databases: the human general transcription factor transcription factor name (TFIIE) (Figure 8C, PDB:1VD4) (75), and the yet undescribed NMR structure of the isolated zinc ribbon domain of Nob1 from *Mus musculus* (PDB:2CON). Both exhibit a Cα RMSD of approximately 1.1 Å to the zinc ribbon domain of PhNob1. However, the loop of $PhNob1_{ZN}$ between β2 and β3 has a different structural arrangement in TFIIE with one N-terminal helix and a C-terminal additional β-hairpin (Figure 8C). Remarkably, the isolated mouse Nob1 domain contains two additional short β-strands (β4, β5) at its C-terminus forming a β-hairpin unconnected to the central three-stranded β-sheet and a long disordered C-terminal tail (Figure 8C). This additional β-hairpin represents the N-terminal part of the C-terminal extension observed only in eukaryotic Nob1 proteins in comparison with the archaeal proteins (Figure 1 and Supplementary Figure S2). The entire C-terminal extension is in general more positively charged and shows on its N-terminus a moderate conservation among all eukaryotic species, which according to the structure (2con) forms the β-hairpin structure (Figure 8C and Supplementary Figure S2).

In addition, the two zinc ribbon domains of the serine/arginine (SR)-rich Ran-binding protein ZRanB2 (Figure 8C, PDB:3G9Y) (76,77) are similar to the PhNob1 zinc ribbon domain with respect to the structure of the first two β-strands and the spatial orientation of the zinc binding residues. However, they contain a second β-hairpin not connected to β2 and β3 in contrast to β3 of the PhNob1 zinc ribbon domain. The RMSD between the 24 equivalent Cα positions is approximately 1.6 Å. Interestingly, for this domain an RNA–protein complex has been crystallized that shows how single stranded RNA is recognized in a sequence-specific manner. The RNA is bound to the outer edge of β-sheet β2 and by hydrogen bonds and π-stacking interactions to basic and aromatic residues, respectively, in the loop connecting β2 and β3. However, the RNA-binding residues are not conserved between the ZRanB2 zinc fingers and PhNob1. In particular, the loop connecting β2 and β3 in PhNob1 does not contain many basic and hydrophilic amino acids and lacks the tryptophan observed to stack between two guanines in ZRanB2. Therefore, it is an unlikely candidate for RNA interactions in the PhNob1 zinc ribbon domain. This already suggests that the RNA-recognition mode of

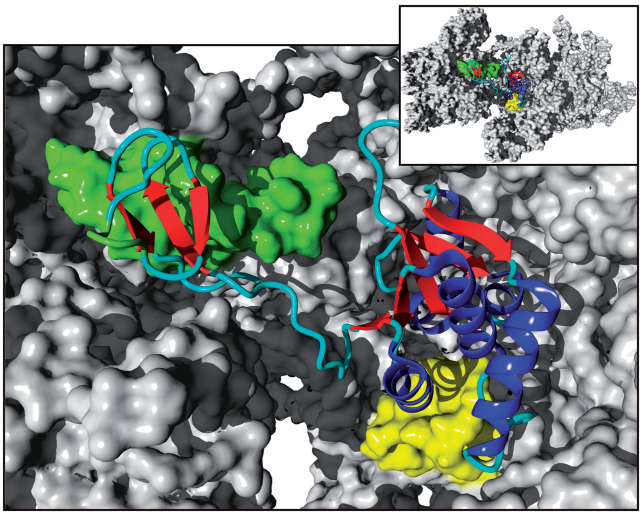


Figure 9. Model visualizing the possible positioning of Nob1 on the small subunit of the ribosome. Shown is the ribosomal surface (gray, pdb: 3o2z) (82) with H40 highlighted in green and site D in yellow. One representative structure of the PhNob1 NMR ensemble is shown as overlay in a cartoon representation. The PIN domain is oriented such that the conserved aspartates of the active site (D12, D82 and D100) point to the last nucleotide of the 3′-end of the rRNA. The Zn ribbon domain is packed with its β-sheet surface against H40. The linker between PIN and Zn ribbon domain is in a relaxed and not fully extended conformation. The inset on the upper right side shows the positioning in the context of the small ribosomal subunit.

the PhNob1 zinc ribbon domain differs from that of the ZRanB2 zinc finger despite their structural similarities.

The two domains of PhNob1 are not only structurally independent but also show a clear division of function. The PIN domain contains the catalytically active center required for endonucleolytic substrate cleavage as mapped in Mn^{2+}-titration experiments and it binds single stranded RNAs in the vicinity of the active center. The binding to single-stranded RNA is rather weak, but we observe a binding preference for larger substrate RNAs. In line, mutations in amino acid residues involved in substrate binding either diminish the cleavage activity or result in the loss of sequence specificity of the cleavage reaction (Figure 6). Thus, the experimentally observed cleavage specificity is probably not only the result of the interactions of the substrate with the PIN domain alone.

The isolated basic zinc ribbon domain is sufficient to bind a RNA helix such as helix 40 of the SSU RNA with a micromolar affinity (Figure 7), but it has a significantly lower affinity for single-stranded substrate RNAs. These observations are consistent with the *in vivo* cross-link results observed for yeast Nob1 (21) and show that this interaction is conserved even in archaea. This suggests an anchoring function for the zinc ribbon domain in Nob1. However, our results also suggest that additional factors influence the positioning of Nob1 as the zinc ribbon domain itself does not exhibit a high specificity for helix 40 *in vitro* (Figure 7).

A combination of a high-affinity RNA-anchoring domain connected to a catalytic domain with a low affinity for its target RNA by a flexible linker has also been suggested for other RNA remodeling enzymes, such as the RNA helicases YxiN (78,79), Hera (80) and members of the RNase III family (see, for example, Ref. 81). It is thought that in some cases, the binding of additional protein factors can modulate the interdomain orientation (78). The specificity for the substrate is in these cases defined by the RNA-binding module which recognizes a primary RNA-binding site with high affinity. This then restricts the catalytic domain to targets in the spatial vicinity, which are bound with much lower affinities. Indeed, the structural dimensions of PhNob1 perfectly match the distance between helix 40 and site D in the mature small ribosomal subunit of *S. cerevisiae* (Figure 9) (82). As Nob1 is recruited to the preribosomal complex within the nucleolus but site D cleavage occurs only in the cytosol (11), three alternative explanations for the regulatory mechanism can be proposed: (i) the distance between Helix 40 and site D might be significantly larger in earlier (nuclear) preribosomal complexes and structural changes in the complex, e.g. triggered by release of export factors or ribosomal protein binding, might deliver helix 40-bound Nob1 into close proximity of the cleavage site; (ii) additional interactions with proteins or RNA might protect the D-cleavage site and restrict access of Nob1 before their release after export; or (iii) other cofactors might keep Nob1 in an deactivated form in the nucleus or activate its catalytic activity in the cytoplasm for D-cleavage. In future, it will be important to analyze if and how RNA and ribosomal protein and ribosome assembly factor binding influences the relative domain orientation in Nob1 and the activation of its endonucleolytic activity.

ACCESSION NUMBER

21cq.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online: Supplementary Table 1 and Supplementary Figures 1–4.

ACKNOWLEDGEMENTS

We thank Maike Ruprecht for her technical support and Drs Oliver Ohlenschläger and Torsten Herrmann for helpful discussions about UNIO.

FUNDING

This project was funded by the Cluster of Excellence Frankfurt 'Macromolecular Complexes', the Goethe University in the Förderfond program and the SFB 902 'Molecular mechanisms of RNA-based regulation' (to M.T.B., J.W. and E.S.), the Center of Biomolecular Magnetic Resonance (to J.W.) as well as the Aventis Foundation (to J.W.). Funding for open access charge: SFB 902 'Molecular mechanisms of RNA-based regulation'.

Conflict of interest statement. None declared.

REFERENCES

1. Schmeing,T.M. and Ramakrishnan,V. (2009) What recent ribosome structures have revealed about the mechanism of translation. *Nature*, **461**, 1234–1242.
2. Dunkle,J.A. and Cate,J.H. (2010) Ribosome structure and dynamics during translocation and termination. *Annu. Rev. Biophys.*, **39**, 227–244.
3. Nierhaus,K.H. (1991) The assembly of prokaryotic ribosomes. *Biochimie*, **73**, 739–755.
4. Kaczanowska,M. and Rydén-Aulin,M. (2007) Ribosome biogenesis and the translation process in *Escherichia coli*. *Microbiol. Mol. Biol. Rev.*, **71**, 477–494.
5. Connolly,K. and Culver,G. (2009) Deconstructing ribosome construction. *Trends Biochem. Sci.*, **34**, 256–263.
6. Fromont-Racine,M., Senger,B., Saveanu,C. and Fasiolo,F. (2003) Ribosome assembly in eukaryotes. *Gene*, **313**, 17–42.
7. Tschochner,H. and Hurt,E. (2003) Pre-ribosomes on the road from the nucleolus to the cytoplasm. *Trends Cell Biol.*, **13**, 255–263.
8. Henras,A.K., Soudet,J., Gerus,M., Lebaron,S., Caizergues-Ferrer,M., Mougin,A. and Henry,Y. (2008) The post-transcriptional steps of eukaryotic ribosome biogenesis. *Cell Mol. Life Sci.*, **65**, 2334–2359.
9. Venema,J. and Tollervey,D. (1999) Ribosome synthesis in *Saccharomyces cerevisiae*. *Annu. Rev. Genet.*, **33**, 261–311.
10. Thomson,E. and Tollervey,D. (2010) The final step in 5.8S rRNA processing is cytoplasmic in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.*, **30**, 976–984.
11. Zemp,I. and Kutay,U. (2007) Nuclear export and cytoplasmic maturation of ribosomal subunits. *FEBS Lett.*, **581**, 2783–2793.
12. Fatica,A., Oeffinger,M., Dlakić,M. and Tollervey,D. (2003) Nob1p is required for cleavage of the 3′ end of 18S rRNA. *Mol. Cell Biol.*, **23**, 1798–1807.
13. Pertschy,B., Schneider,C., Gnädig,M., Schäfer,T., Tollervey,D. and Hurt,E. (2009) RNA helicase Prp43 and its co-factor Pfa1 promote 20S to 18S rRNA processing catalyzed by the endonuclease Nob1. *J. Biol. Chem.*, **284**, 35079–35091.
14. Bohnsack,M.T., Martin,R., Granneman,S., Ruprecht,M., Schleiff,E. and Tollervey,D. (2009) Prp43 bound at different sites on the pre-rRNA performs distinct functions in ribosome synthesis. *Mol. Cell*, **36**, 583–592.
15. Lamanna,A.C. and Karbstein,K. (2011) An RNA conformational switch regulates pre-18S rRNA cleavage. *J. Mol. Biol.*, **405**, 3–17.
16. Lafontaine,D., Delcour,J., Glasser,A.L., Desgres,J. and Vandenhaute,J. (1994) The DIM1 gene responsible for the conserved m62Am62A dimethylation in the 3′-terminal loop of 18 S rRNA is essential in yeast. *J. Mol. Biol.*, **241**, 492–497.
17. Fassio,C.A., Schofield,B.J., Seiser,R.M., Johnson,A.W. and Lycan,D.E. (2010) Dominant mutations in the late 40S biogenesis factor Ltv1 affect cytoplasmic maturation of the small ribosomal subunit in *Saccharomyces cerevisiae*. *Genetics*, **185**, 199–209.
18. Neueder,A., Jakob,S., Pöll,G., Linnemann,J., Deutzmann,R., Tschochner,H. and Milkereit,P. (2010) A local role for the small ribosomal subunit primary binder rpS5 in final 18S rRNA processing in yeast. *PLoS One*, **5**, e10194.
19. Tone,Y., Tanahashi,N., Tanaka,K., Fujimuro,M., Yokosawa,H. and Toh-e,A. (2000) Nob1p, a new essential protein, associates with the 26S proteasome of growing *Saccharomyces cerevisiae* cells. *Gene*, **243**, 37–45.
20. Lamanna,A.C. and Karbstein,K. (2009) Nob1 binds the single-stranded cleavage site D at the 3′-end of 18S rRNA with its PIN domain. *Proc. Natl Acad. Sci. USA*, **106**, 14259–14264.
21. Granneman,S., Petfalski,E., Swiatkowska,A. and Tollervey,D. (2010) Cracking pre-40S ribosomal subunit structure by systematic analyses of RNA-protein cross-linking. *EMBO J.*, **29**, 2026–2036.
22. Arcus,V.L., Bäckbro,K., Roos,A., Daniel,E.L. and Baker,E.N. (2004) Distant structural homology leads to the functional characterization of an archaeal PIN domain as an exonuclease. *J. Biol. Chem.*, **279**, 16471–16478.
23. Lebreton,A., Tomecki,R., Dziembowski,A. and Séraphin,B. (2008) Endonucleolytic RNA cleavage by a eukaryotic exosome. *Nature*, **456**, 993–996.
24. Qian,X., Gozani,S.N., Yoon,H., Jeon,C.J., Agarwal,K. and Weiss,M.A. (1993) Novel zinc finger motif in the basal transcriptional machinery: three-dimensional NMR studies of the nucleic acid binding domain of transcriptional elongation factor TFIIS. *Biochem.*, **32**, 9944–9959.
25. Woolls,H.A., Lamanna,A.C. and Karbstein,K. (2011) Roles of Dim2 in ribosome assembly. *J. Biol. Chem.*, **286**, 2578–2586.
26. Fatica,A., Tollervey,D. and Dlakić,M. (2004) PIN domain of Nob1p is required for D-site cleavage in 20S pre-rRNA. *RNA*, **10**, 1698–1701.
27. Schaeffer,D., Tsanova,B., Barbas,A., Reis,F.P., Dastidar,E.G., Sanchez-Rotunno,M., Arraiano,C.M. and van Hoof,A. (2009) The exosome contains domains with specific endoribonuclease, exoribonuclease and cytoplasmic mRNA decay activities. *Nat. Struct. Mol. Biol.*, **16**, 56–62.
28. Schneider,C., Leung,E., Brown,J. and Tollervey,D. (2009) The N-terminal PIN domain of the exosome subunit Rrp44 harbors endonuclease activity and tethers Rrp44 to the yeast core exosome. *Nucleic Acids Res.*, **37**, 1127–1140.
29. Huntzinger,E., Kashima,I., Fauser,M., Saulière,J. and Izaurralde,E. (2008) SMG6 is the catalytic endonuclease that cleaves mRNAs containing nonsense codons in metazoan. *RNA*, **14**, 2609–2617.
30. Wild,T., Horvath,P., Wyler,E., Widmann,B., Badertscher,L., Zemp,I., Kozak,K., Csucs,G., Lund,E. and Kutay,U. (2010) A protein inventory of human ribosome biogenesis reveals an essential function of exportin 5 in 60S subunit export. *PLoS Biol.*, **8**, e1000522.
31. Wyler,E., Zimmermann,M., Widmann,B., Gstaiger,M., Pfannstiel,J., Kutay,U. and Zemp,I. (2011) Tandem affinity purification combined with inducible shRNA expression as a tool to study the maturation of macromolecular assemblies. *RNA*, **17**, 189–200.
32. Oehler,V.G., Yeung,K.Y., Choi,Y.E., Bumgarner,R.E., Raftery,A.E. and Radich,J.P. (2009) The derivation of diagnostic markers of chronic myeloid leukemia progression from microarray data. *Blood*, **114**, 3292–3298.
33. Lin,Y., Peng,S., Yu,H., Teng,H. and Cui,M. (2011) RNAi-mediated downregulation of NOB1 suppresses the growth and colony-formation ability of human ovarian cancer cells. *Med. Oncol.*, February 2 (doi:10.1007/s12032-010-9808-5; epub ahead of print).
34. Han,Y., Hong,L., Chen,Y., Zhong,C., Wang,Y., Zhao,D., Xue,T., Qiao,L. and Qiu,J. (2011) Up-regulation of Nob1 in the rat auditory system with noise-induced hearing loss. *Neurosci. Lett.*, **491**, 79–82.
35. LaRonde-LeBlanc,N. and Wlodawer,A. (2004) Crystal structure of A. fulgidus Rio2 defines a new family of serine protein kinases. *Structure*, **12**, 1585–1594.
36. Coltri,P.P., Guimarães,B.G., Granato,D.C., Luz,J.S., Teixeira,E.C., Oliveira,C.C. and Zanchin,N.I. (2007) Structural insights into the interaction of the Nip7 PUA domain with polyuridine RNA. *Biochemistry*, **46**, 14177–14787.
37. Taylor,A.B., Meyer,B., Leal,B.Z., Kötter,P., Schirf,V., Demeler,B., Hart,P.J., Entian,K.D. and Wöhnert,J. (2008) The crystal structure of Nep1 reveals an extended SPOUT-class methyltransferase fold and a pre-organized SAM-binding site. *Nucleic Acids Res.*, **36**, 1542–1554.
38. Ng,C.L., Waterman,D.G., Koonin,E.V., Walters,A.D., Chong,J.P., Isupov,M.N., Lebedev,A.A., Bunka,D.H., Stockley,P.G. and Ortiz-Lombardía,M. (2009) Conformational flexibility and molecular interactions of an archaeal homologue of the Shwachman-Bodian-Diamond syndrome protein. *BMC Struct. Biol.*, **9**, 32.
39. Pulicherla,N., Pogorzala,L.A., Xu,Z.O., Farrell,H.C., Musayev,F.N., Scarsdale,J.N., Sia,E.A., Culver,G.M. and Rife,J.P. (2009) Structural and functional divergence within the Dim1/KsgA family of rRNA methyltransferases. *J. Mol. Biol.*, **391**, 884–893.
40. Jia,M.Z., Horita,S., Nagata,K. and Tanokura,M. (2010) An archaeal Dim2-like protein, aDim2p, fo,rms a ternary complex with a/eIF2 alpha and the 3′ end fragment of 16S rRNA. *J. Mol. Biol.*, **398**, 774–785.
41. Finn,R.D., Mistry,J., Tate,J., Coggill,P., Heger,A., Pollington,J.E., Gavin,O.L., Gunasekaran,P., Ceric,G., Forslund,K. *et al.* (2010) The Pfam protein families database. *Nucleic Acids Res.*, **38**, D211–D222.
42. Altschul,S.F., Madden,T.L., Schäffer,A.A., Zhang,J., Zhang,Z., Miller,W. and Lipman,D.J. (1997) Gapped BLAST and

PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.*, **25**, 3389–3402.
43. Benson,D.A., Karsch-Mizrachi,I., Lipman,D.J., Ostell,J. and Sayers,E.W. (2009) GenBank. *Nucleic Acids Res.*, **37**, D26–D31.
44. Whelan,S. and Goldman,N. (2001) A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. *Mol. Biol. Evol.*, **18**, 691–699.
45. Stamatakis,A., Hoover,P. and Rougemont,J. (2008) A rapid bootstrap algorithm for the RAxML Web servers. *Syst. Biol.*, **57**, 758–771.
46. Katoh,K. and Toh,H. (2010) Parallelization of the MAFFT multiple sequence alignment program. *Bioinformatics*, **26**, 1899–1900.
47. Leulliot,N., Bohnsack,M.T., Graille,M., Tollervey,D. and Van Tilbeurgh,H. (2008) The yeast ribosome synthesis factor Emg1 is a novel member of the superfamily of alpha/beta knot fold methyltransferases. *Nucleic Acids Res.*, **36**, 629–639.
48. Veith,T., Wurm,J.P., Duchardt-Ferner,E., Weis,B., Hennig,R., Martin,R., Safferthal,C., Bohnsack,M.T., Schleiff,E. and Wöhnert,J. (2011) NMR resonance assignments for an archaeal homolog of the endonuclease Nob1 involved in ribosome biogenesis. *Biomol.*, NMR Assign, July 6 (doi:10.1007/s12104-011-9323-4; epub ahead of print).
49. Wurm,J.P., Duchardt,E., Meyer,B., Leal,B.Z., Kötter,P., Entian,K.D. and Wöhnert,J. (2009) Backbone resonance assignments of the 48 kDa dimeric putative 18S rRNA-methyltransferase Nep1 from *Methanocaldococcus jannaschii*. *Biomol. NMR Assign.*, **3**, 251–254.
50. Farrow,N.A., Muhandiram,R., Singer,A.U., Pascal,S.M., Kay,C.M., Gish,G., Shoelson,S.E., Pawson,T., Forman-Kay,J.D. and Kay,L.E. (1994) Backbone dynamics of a free and phosphopeptide-complexed Src homology 2 domain studied by 15N NMR relaxation. *Biochem.*, **33**, 5984–6003.
51. Keller,R. (2004) *The Computer Aided Resonance Tutorial*. CANTINA Verlag, Goldau, Switzerland.
52. Shen,Y., Delaglio,F., Cornilescu,G. and Bax,A. (2009) TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *J. Biomol. NMR*, **44**, 213–223.
53. Cordier,F. and Grzesiek,S. (1999) Direct observation of hydrogen bonds in proteins by interresidue $^{3h}J_{NC}$ scalar couplings. *J. Am. Chem Soc.*, **121**, 1601–1602.
54. Herrmann,T., Güntert,P. and Wüthrich,K. (2002) Protein NMR structure determination with automated NOE-identification in the NOESY spectra using the new software ATNOS. *J. Biomol. NMR*, **24**, 171–189.
55. Fiorito,F., Damberger,F., Herrmann,T. and Wüthrich,K. (2008) Automated amino acid side-chain NMR assignment of proteins using (13)C- and (15)N-resolved 3D [(1)H, (1)H]-NOESY. *J. Biomol. NMR*, **42**, 23–33.
56. Güntert,P., Mumenthaler,C. and Wüthrich,K. (1997) Torsion angle dynamics for NMR structure calculation with the new program DYANA. *J. Mol. Biol.*, **273**, 283–293.
57. Ohlenschläger,O., Seiboth,T., Zengerling,H., Briese,L., Marchanka,A., Ramachandran,R., Baum,M., Korbas,M. and Meyer-Klaucke,W. (2006) Solution structure of the partially folded high-risk human papilloma virus 45 oncoprotein E7. *Oncogene*, **25**, 5953–5959.
58. Luginbühl,P., Güntert,P., Billeter,M. and Wüthrich,K. (1996) The new program OPAL for molecular dynamics simulations and energy refinements of biological macromolecules. *J. Biomol. NMR*, **8**, 136–146.
59. Koradi,R., Billeter,M. and Wüthrich,K. (1996) MOLMOL: a program for display and analysis of macromolecular structures. *J. Mol. Graph.*, **14**, 51–54.
60. Pettersen,E.F., Goddard,T.D., Huang,C.C., Couch,G.S., Greenblatt,D.M., Meng,E.C. and Ferrin,T.E. (2004) UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.*, **25**, 1605–1612.
61. Murzin,A.G., Brenner,S.E., Hubbard,T. and Chothia,C. (1995) SCOP: a structural classification of proteins database for the investigation of sequences and structures. *J. Mol. Biol.*, **247**, 536–540.
62. Berman,H.M., Westbrook,J., Feng,Z., Gilliland,G., Bhat,T.N., Weissig,H., Shindyalov,I.N. and Bourne,P.E. (2000) The protein data bank. *Nucleic Acids Res.*, **28**, 235–242.
63. Konagurthu,A.S., Whisstock,J.C., Stuckey,P.J. and Lesk,A.M. (2006) MUSTANG: a multiple structural alignment algorithm. *Proteins*, **64**, 559–574.
64. Stamatakis,A. (2006) "RAxML-VI-HPC: Maximum Likelihood-based Phylogenetic analyses with Thousands of Taxa and Mixed Models". *Bioinformatics*, **22**, 2688–2690.
65. Tamura,K., Peterson,D., Peterson,N., Stecher,G., Nei,M. and Kumar,S. (2011) MEGA5: Molecular Evolutionary Genetics Analysis using Maximum Likelihood, Evolutionary Distance and Maximum Parsimony Methods. *Mol. Biol. Evol.*, **28**, 2731–2739.
66. Schägger,H. (2006) Tricin-SDS-PAGE. *Nat. Prot.*, **1**, 16–22.
67. Huber,H., Hohn,M.J., Stetter,K.O. and Rachel,R. (2003) The phylum Nanoarchaeota: present knowledge and future perspectives of a unique form of life. *Res. Microbiol.*, **154**, 165–171.
68. Waters,E., Hohn,M.J., Ahel,I., Graham,D.E., Adams,M.D., Barnstead,M., Beeson,K.Y., Bibbs,L., Bolanos,R., Keller,M. *et al.* (2003) The genome of *Nanoarchaeum equitans*: insights into early archaeal evolution and derived parasitism. *Proc. Natl Acad. Sci. USA*, **100**, 12984–12988.
69. Fahrni,C.J. and O'Halloran,T.V. (1999) Aqueous coordination chemistry of quinoline-based fluorescence probes for the biological chemistry of zinc. *J. Am. Chem. Soc.*, **121**, 11448–11458.
70. Eghbalnia,H.R., Wang,L., Bahrami,A., Assadi,A. and Markley,J.L. (2005) Protein energetic conformational analysis from NMR chemical shifts (PECAN) and its use in determining secondary structural elements. *J. Biomol. NMR*, **32**, 71–81.
71. Kornhaber,G.J., Snyder,D., Moseley,H.N. and Montelione,G.T. (2006) Identification of zinc-ligated cysteine residues based on 13Calpha and 13Cbeta chemical shift data. *J. Biomol. NMR*, **34**, 259–269.
72. Glavan,F., Behm-Ansmant,I., Izaurralde,E. and Conti,E. (2006) Structures of the PIN domains of SMG6 and SMG5 reveal a nuclease within the mRNA surveillance complex. *EMBO J.*, **25**, 5117–5125.
73. Takeshita,D., Zenno,S., Lee,W.C., Saigo,K. and Tanokura,M. (2007) Crystal structure of the PIN domain of human telomerase-associated protein EST1A. *Proteins*, **68**, 980–989.
74. Levin,I., Schwarzenbacher,R., Page,R., Abdubek,P., Ambing,E., Biorac,T., Brinen,L.S., Campbell,J., Canaves,J.M., Chiu,H.J. *et al.* (2004) Crystal structure of a PIN (PilT N-terminus) domain (AF0591) from *Archaeoglobus fulgidus* at 1.90 A resolution. *Proteins*, **56**, 404–408.
75. Okuda,M., Tanaka,A., Arai,Y., Satoh,M., Okamura,H., Nagadoi,A., Hanaoka,F., Ohkuma,Y. and Nishimura,Y. (2004) A novel zinc finger structure in the large subunit of human general transcription factor TFIIE. *J. Biol. Chem.*, **279**, 51395–51403.
76. Nguyen,C.D., Mansfield,R.E., Leung,W., Vaz,P.M., Loughlin,F.E., Grant,R.P. and Mackay,J.P. (2011) Characterization of a family of RanBP2-type zinc fingers that can recognize single-stranded RNA. *J. Mol. Biol.*, **407**, 273–283.
77. Loughlin,F.E., Mansfield,R.E., Vaz,P.M., McGrath,A.P., Setiyaputra,S., Gamsjaeger,R., Chen,E.S., Morris,B.J., Guss,J.M. and Mackay,J.P. (2009) The zinc fingers of the SR-like protein ZRANB2 are single-stranded RNA-binding domains that recognize 5′ splice site-like sequences. *Proc. Natl Acad. Sci. USA*, **106**, 5581–5586.
78. Karow,A.R. and Klostermeier,D. (2010) A structural model for the DEAD box helicase YxiN in solution: localization of the RNA binding domain. *J. Mol. Biol.*, **402**, 629–642.
79. Hardin,J.W., Hu,Y.X. and McKay,D.B. (2010) Structure of the RNA binding domain of a DEAD-box helicase bound to its ribosomal RNA target reveals a novel mode of recognition by an RNA recognition motif. *J. Mol. Biol.*, **402**, 412–427.
80. Rudolph,M.G. and Klostermeier,D. (2009) The *Thermus thermophilus* DEAD box helicase Hera contains a modified RNA recognition motif domain loosely connected to the helicase core. *RNA*, **15**, 1993–2001.
81. Macrae,I.J., Zhou,K., Li,F., Repic,A., Brooks,A.N., Cande,W.Z., Adams,P.D. and Doudna,J.A. (2006) Structural basis for double-stranded RNA processing by Dicer. *Science*, **311**, 195–198.
82. Ben-Shem,A., Jenner,L., Yusupova,G. and Yusupov,M. (2010) Crystal structure of the eukaryotic ribosome. *Science*, **330**, 1203–1209.