Structure of Yeast Dom34

Marc Graille, Maxime Chaillet, Herman van Tilbeurgh

THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 283, NO. 11, pp. 7145–7154, March 14, 2008

Structure of Yeast Dom34

*A PROTEIN RELATED TO TRANSLATION TERMINATION FACTOR Erf1 AND INVOLVED IN No-Go DECAY**[S]

Received for publication, October 3, 2007, and in revised form, December 20, 2007 Published, JBC Papers in Press, January 7, 2008, DOI 10.1074/jbc.M708224200

Marc Graille[1], **Maxime Chaillet, and Herman van Tilbeurgh**

From the Institut de Biochimie et de Biophysique Moléculaire et Cellulaire, Université Paris-Sud, UMR8619-CNRS, IFR115, F-91405 Orsay, France

The yeast protein Dom34 has been described to play a critical role in a newly identified mRNA decay pathway called No-Go decay. This pathway clears cells from mRNAs inducing translational stalls through endonucleolytic cleavage. Dom34 is related to the translation termination factor eRF1 and physically interacts with Hbs1, which is itself related to eRF3. We have solved the 2.5-Å resolution crystal structure of *Saccharomyces cerevisiae* Dom34. This protein is organized in three domains with the central and C-terminal domains structurally homologous to those from eRF1. The N-terminal domain of Dom34 is different from eRF1. It adopts a Sm-fold that is often involved in the recognition of mRNA stem loops or in the recruitment of mRNA degradation machinery. The comparison of eRF1 and Dom34 domains proposed to interact directly with eRF3 and Hbs1, respectively, highlights striking structural similarities with eRF1 motifs identified to be crucial for the binding to eRF3. In addition, as observed for eRF1 that enhances eRF3 binding to GTP, the interaction of Dom34 with Hbs1 results in an increase in the affinity constant of Hbs1 for GTP but not GDP. Taken together, these results emphasize that eukaryotic cells have evolved two structurally related complexes able to interact with ribosomes either paused at a stop codon or stalled in translation by the presence of a stable stem loop and to trigger ribosome release by catalyzing chemical bond hydrolysis.

In living cells, elaborate quality control mechanisms are essential for ensuring the highest fidelity in the transfer and decoding of the genetic information, *i.e.* during DNA replication, transcription and translation of mRNAs into proteins (for reviews, see Refs. 1–5). During the last decade, RNA biogenesis has proven to be a niche for the identification of new surveillance pathways aimed at clearing nonfunctional RNAs from the cells (6). For instance, the various steps responsible for the maturation of a mRNA (splicing, polyadenylation, capping, . . .) are sources of errors that will greatly affect the overall accuracy of gene expression. To prevent the translation of faulty mRNAs by the ribosome, eukaryotic cells have evolved several quality control systems whose roles are to detect and degrade these pools of aberrant mRNAs. In addition, these mechanisms also contribute to the maintenance of cellular homeostasis and are implicated in certain diseases (7, 8).

Among these quality control systems, the best studied is nonsense-mediated mRNA decay (NMD),[2] a pathway targeting mRNAs harboring in-frame premature termination (nonsense) codons (9–13). NMD identifies and removes from the cytoplasm, aberrant mRNA that if translated could produce truncated proteins. This pathway functions through the concerted action between the translational apparatus (ribosome, release factors eRF1 and eRF3, Pab1), mRNA decay machineries (Dcp1-Dcp2, Xrn1, and exosome), and specific NMD factors whose number varies among species (Upf1, Upf2, Upf3, Smg1, and Smg5–7). Release factors eRF1 and eRF3 recognize stop codons in the ribosomal A-site and then interact with Upf1, a RNA helicase as well as with Pab1, which binds to the 3′-poly(A) tract. Growing evidence suggests that in yeast as well as in flies and worms, the discrimination between normal and aberrant stop codons is based on the aptitude of Pab1 and eRF3 to interact together. A premature stop codon would be too far from the 3′ poly(A) tract to allow the formation of the eRF3-Pab1 complex, leading to the assignment of this mRNA as aberrant and substrate of NMD (14–16). In mammals, NMD targets are recognized depending on a post-splicing exon junction complex of proteins that is deposited ~20–24 nucleotides upstream of exon-exon junctions (17, 18). Finally, the surveillance complex made by the Upf proteins (Upf1, Upf2, and Upf3) recruits the mRNA degradation machineries.

A second pathway, non-stop decay (NSD), is dedicated to the clearance of mRNAs devoid of in-frame stop codons or prematurely polyadenylated (19–21). NSD facilitates the release of the ribosomes stalled at the 3′ end of mRNAs and as a consequence, also protects cells from the production of aberrant proteins. In bacteria, proteins encoded by mRNAs lacking stop codons are marked for degradation by the addition of a transfer messenger RNA at their C-terminal extremities (22, 23). In the eukaryotic NSD pathway, these mRNAs are detected and

* This work was supported by the Agence Nationale pour la Recherche Grant ANR-06-BLAN-0075-02, the Association Française contre les Myopathies, and the EU "3D-Repertoire" program LSHG-CT-2005-512028. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

[S] The on-line version of this article (available at http://www.jbc.org) contains supplemental Figs. S1–S3.

The atomic coordinates and structure factors (code 2VGN and 2VGM) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

[1] To whom correspondence should be addressed. Tel.: 33-1-69-15-50-47; E-mail: marc.graille@u-psud.fr.

[2] The abbreviations used are: NMD, nonsense-mediated mRNA decay; NSD, non-stop decay; Mant, methylanthraniloyl; NGD, No-Go decay; SeMet, selenomethionine; r.m.s., root mean square; GMPPNP, guanyl-5′-yl imidodiphosphate.

degraded by a completely different mechanism, which involves Ski7, the exosome, and the Ski complex (19, 21). Ski7 connects the exosome to the Ski complex via its N-terminal domain (24). Its C-terminal domain is structurally related to the GTPase domain from translation elongation factor eEF-1A and the translation termination factor eRF3, two proteins that bind close to the ribosomal A-site.

More recently, a novel mRNA surveillance pathway called No-Go decay (NGD) has been unraveled in yeast (25–27). Engineering of a very stable stem loop into a mRNA forces the ribosome to pause during elongation, leading to the accelerated degradation of the mRNA via an endonuclease cleavage followed by rapid degradation by the exosome and the Xrn1 nuclease. NGD involves at least two proteins, Dom34 and Hbs1. In yeast, Dom34 functions in protein translation to promote G_1 progression and cellular differentiation (28). In *Drosophila melanogaster* and mouse, the Dom34 homolog named Pelota is required for correct mitotic and meiotic cell division (29–31). According to primary sequence analysis, Dom34 is related to the translation termination factor eRF1 (32). The eRF1 protein interacts with the translation termination factor eRF3 and its three-dimensional structure mimics that of a tRNA (33). Dom34 interacts with Hbs1, which has been identified as an Hsp70 subfamily B suppressor and is a member of the family of eEF-1A-like GTPases (34–36). The translation elongation factor eEF-1A that delivers aminoacyl-tRNAs to the ribosomal A-site (37) is the representative of this protein family that further includes eRF3 (38, 39) and Ski7, which is involved in exosome-mediated mRNA degradation and in the NSD pathway (19, 21). The Hsp70 Ssb1/2 chaperones interact with the nascent polypeptide chain at the exit of the ribosome channel during protein synthesis. In yeast, deletion of both *ssb1* and *ssb2* genes results in slower cell growth, in the decrease of the number of translating ribosomes and in slower translation (40). Overexpression of Hbs1 suppresses this phenotype in yeast *ssb1*/*ssb2* mutant strains. Based on these observations, Inagaki and co-workers (35) have suggested that during translation elongation, Hbs1 may help in stop codon-independent peptide release from ribosomes stalled by the absence of Hsp70-mediated nascent polypeptide channeling. This is further supported by the involvement of Hbs1 and Dom34 in the NGD pathway (26). Very recently, the crystal structure of the Dom34/Pelota protein from the archeon *Thermoplasma acidophilum* (hereafter called Ta Pelota) has been described (41). Lee and colleagues (41) showed *in vitro* that the N-terminal domains from both *T. acidophilum* and *Saccharomyces cerevisiae* Dom34/Pelota display endonuclease activity against mRNAs that contain a stem loop, supporting the hypothesis that Dom34 is the endonuclease involved in the NGD pathway.

Here, we present the 2.5-Å resolution crystal structure of *S. cerevisiae* Dom34 protein (hereafter called Dom34) and compare it to the structure of Ta Pelota. Yeast Dom34 is composed of three domains whose orientations are significantly different from those observed in the structure of Ta Pelota. Two of these domains are structurally homologous to eRF1. From the structure, we observe that all the eRF1 motifs known to be involved in the interaction with eRF3 have structural matches within Dom34. This resemblance is further supported by fluorescence experiments showing that Dom34 enhances the affinity of Hbs1 for GTP but not GDP as already observed for eRF1 with eRF3.

MATERIALS AND METHODS

Cloning, Expression, and Purification of Dom34 and Hbs1—The *DOM34 (YNL011w)* and *HBS1 (YKR084c)* genes were amplified by PCR using the genomic DNA of *S. cerevisiae* strain S288C as a template. An additional sequence coding for a His_6 tag was introduced at the 3′ end of each open reading frame during amplification. The PCR products derived from *DOM34* and *HBS1* were then cloned into a derivative of pET9 vector (Stratagene). Expression of these proteins was done at 15 °C using the transformed *Escherichia coli* Rosetta (DE3) pLysS (Dom34) or Gold (DE3) (Hbs1) strains (Novagen) and 2×YT medium (BIO101 Inc.) supplemented with kanamycin at 50 μg/ml. For both proteins, cells were harvested by centrifugation, resuspended in 30 ml of 20 mM Tris-HCl, pH 7.5, 200 mM NaCl, 5 mM β-mercaptoethanol, and stored at −20 °C. Cell lysis was achieved by sonication. Both His-tagged proteins were separately purified on a nickel-nitrilotriacetic acid column (Qiagen Inc.) followed by gel filtration on a Superdex™ 75 column (Amersham Biosciences) equilibrated in 20 mM Tris-HCl, pH 7.5, 200 mM NaCl, 10 mM β-mercaptoethanol. The Hbs1-Dom34 complex was purified on a Superdex™ 200 (16/60) column (Amersham Biosciences) after mixing of individual components.

Protein labeling with SeMet was conducted as described (42). Due to a decreased solubility compared with the native protein, the SeMet protein was purified as described above in the following buffer: 20 mM Tris-HCl, pH 7.5, 1 M NaCl, 10 mM β-mercaptoethanol. During protein concentration, the NaCl was decreased to 0.5 M. The homogeneity and SeMet labeling of Dom34 was verified by SDS-PAGE and mass spectrometry.

Fluorescence Measurements for Binding of Ligands—Methylanthraniloyl (Mant)-GTP and Mant-GDP were purchased from Jena Biosciences. Excitation was performed at 349 nm and emission scanned from 400 to 500 nm with a Cary Eclipse fluorospectrophotometer (Varian). Measurements were performed at 25 °C with 2 μM Hbs1 in 20 mM Tris-HCl, pH 7.5, 200 mM NaCl, 2.5 mM $MgCl_2$, 5 mM β-mercaptoethanol, 10% glycerol. Successive aliquots of ligands (from 0 to 20 mM) were added to the protein. Binding of ligands to Hbs1 or Hbs1-Dom34 (2 μM Dom34) was quantified by measuring fluorescence difference of Mant-guanine nucleotides at 448 nm against ligand concentration. Experiments were performed in triplicate.

Structure Determination—Crystallization trials were performed at 18 °C. Crystals for the native protein were grown from a mixture in a 1:1 ratio of 9 mg/ml protein solution in 20 mM Tris-HCl, pH 7.5, 200 mM NaCl, 10 mM β-mercaptoethanol and crystallization liquor containing 7.5% polyethylene glycol 4000, 50 mM NaH_2PO_4, pH 6.7, 3% xylitol. Crystals for the SeMet protein (in 20 mM Tris-HCl, pH 7.5, 500 mM NaCl, and 10 mM β-mercaptoethanol) were grown under the same conditions after microseeding from crystals of the native form. All crystals were cryo-protected by transfer into the crystallization solution with progressively higher glyc-

TABLE 1
Data collection statistics

Data collection	SeMet	Native
ESRF beamline	BM30A	ID29
Resolution (Å)	14.0-2.5 (2.6-2.5)	35.0-2.6 (2.74-2.6)
Space group	$P3_121$	$P3_221$
Unit cell parameters	$a = b = 75.6$ Å; $c = 324.5$ Å	$a = b = 75.5$ Å; $c = 164.0$ Å
Total number of reflections	157,243	48,718
Total number of unique reflections	69,296	16,966
R_{sym} (%)[a]	11.0 (49.2)	9.9 (49.2)
Completeness (%)	96.4 (91.1)	98.7 (99.9)
$I/\sigma(I)$	7.3 (1.75)	11.3 (1.4)
Redundancy	2.2 (2.1)	2.9 (2.9)
Refinement		
Resolution (Å)	14.0-2.5	35.0-2.6
R/R_{free} (%)[b]	21.0/25.1	22.6/28.8
R.m.s. deviation bonds (Å)	0.010	0.007
R.m.s. deviation angles (°)	1.330	1.079
Number of residue/water/ligand	716/228/3 (2 glycerol + 1 phosphate)	354/26/0
B factors (Å^2) (protein/water/ligand)	36/31/40	65/49/–
PDB code	2VGN	2VGM

[a] $R_{sym} = \Sigma_h\Sigma_i|I_{hi} - \langle I_h\rangle|/\Sigma_h\Sigma_i I_{hi}$, were I_{hi} is the ith observation of the reflection h, while $\langle I_h\rangle$ is the mean intensity of reflection h.
[b] $R_{factor} = \Sigma||F_o| - |F_c||/|F_o|$. R_{free} was calculated with a small fraction (5%) of randomly selected reflections.

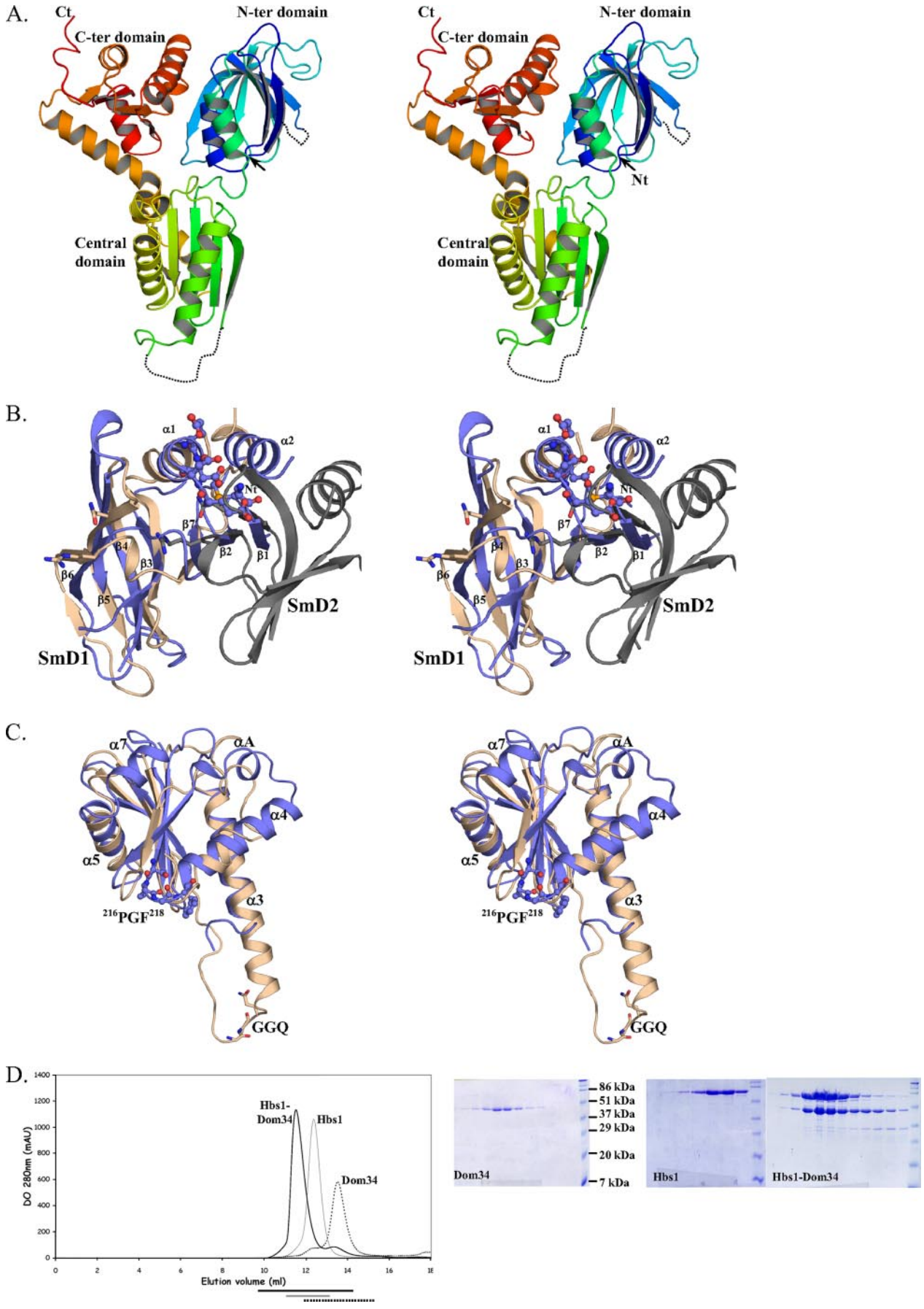
erol concentrations up to 30% (v/v) and then flash cooled in liquid nitrogen.

Native crystals diffracted to 2.6-Å resolution on beamline ID29 at the European Synchrotron Radiation Facility (ESRF, Grenoble, France). These crystals belong to space group $P3_221$ with one molecule in the asymmetric unit (cell dimensions $a = b = 75.5$ Å, $c = 164$ Å). Diffraction data (2.5-Å resolution) at the selenium edge were collected from a flash cooled SeMet crystal at 100 K on beamline BM30A (ESRF, Grenoble, France) (43). The crystal belongs to space group $P3_121$ with cell dimensions $a = b = 75.6$ Å, $c = 324.5$ Å (the c axis is twice longer than in native crystal), corresponding to two molecules per asymmetric unit and a solvent content of 60%. The structure was determined by the single wavelength anomalous dispersion method using the selenium anomalous signal. Data were processed using the XDS package (44). Eighteen of 20 (10 per Dom34 protomer) expected selenium atom sites were found with the program SHELXD in the 14-3.5-Å resolution range (45). These sites were then used by the SOLVE/RESOLVE package to calculate experimental electron density maps and to improve them by both NCS averaging and solvent flattening (46, 47). The quality of the experimental phases allowed automatic building of a partial model with the program RESOLVE (supplemental Fig. S1A). This model was completed by iterative cycles of manual rebuilding using the molecular modeling program "TURBO" followed by refinement with the program PHENIX in the 14-2.5-Å resolution range (48). In the final model, the following Dom34 regions Met^1-Phe^{47}, Thr^{60}-Tyr^{170}, and Asp^{180}-Asp^{382} could be modeled in the electron density as well as 228 water molecules, one phosphate ion from the crystallization solution and two glycerol molecules, which was used as cryoprotectant (supplemental Fig. S1B). In parallel, the structure of the native protein (one Dom34 protomer in the asymmetric unit) was refined using the program PHENIX and considering a twinning fraction of 6.5%, yielding R and R_{free} factors of 22.6 and 28.8%, respectively. The statistics for data collection and refinement are provided in Table 1. The atomic coordinates and structure factors for the SeMet and native structures have been deposited into the Brookhaven Protein Data Bank under accession numbers 2VGN and 2VGM, respectively.

RESULTS

Overall Structure—The structure of *S. cerevisiae* Dom34 was solved using the single wavelength anomalous diffraction method from the anomalous scattering from SeMet-substituted protein crystals and refined to 2.5-Å resolution (Table 1). The excellent quality of the electron density maps allowed the building of a large majority of the 386 amino acids of this protein (supplemental Fig. S1). 24 loop residues (Thr^{48}-Ser^{59}, Ser^{171}-Thr^{179}, and Asp^{383}-Glu^{386}) were absent in the final model due to the lack of electron density, reflecting intrinsic flexibility. Two copies of Dom34 related by a 2-fold axis are present in the asymmetric unit. Although the interface between both copies buries a large surface (2,100 Å^2), size exclusion chromatography indicated that the protein is monomeric in solution (data not shown). We have also refined the structure of the native protein to 2.6-Å resolution in a different space group with only one copy of Dom34 in the asymmetric unit.

Dom34 adopts a "T" shape with overall approximate dimensions $30 \times 55 \times 80$ Å^3. It is organized into three domains with the arms of the T formed by both N-terminal and C-terminal domains and the vertical one by the central domain (Fig. 1*A*). The central domain interacts loosely with N-terminal and C-terminal domains (buried accessible surface area of 520 and 640 Å^2, respectively), which are tightly packed together (buried accessible surface area of 1,030 Å^2). Superimposition of the three independent copies of Dom34 (two SeMet labeled and one native) onto their N-terminal domain (r.m.s. deviation ≈ 0.35–0.4 Å over 130 Cα positions) reveals flexibility of the central and C-terminal domains relative to the N-terminal domain (supplemental Fig. S2). The two conformations can be superposed by a rigid body movement of these domains whose amplitude is inversely correlated to their contact area with the N-terminal domain, *i.e.* the central domain exhibits larger shifts than the C-terminal domain.



The N-terminal domain (residues 1–135) consists of a highly bent 7-stranded antiparallel β-sheet (strand order β1, β2, β7, β3, β4, β5, and β6; Figs. 1*B* and 2) flanked on one side by two α-helices. This domain can be assigned to the Sm-fold family that is characterized by a N-terminal α-helix followed by a twisted five-stranded β-sheet. Compared with the canonical Sm-fold, two insertions (two N-terminal β-strands β1 and β2 and the C-terminal helix α2) are located on the same extremity of the domain (Fig. 1*B*). The closest structural homologs of this domain are *E. coli* Hfq (r.m.s. deviation of 2.5 Å over 55 Cα atoms; 7% sequence identity (49)) and SmD1 from human small nuclear ribonucleoprotein (r.m.s. deviation of 2.9 Å over 70 Cα atoms; 10% sequence identity (50)). Many Sm-fold containing proteins, including Hfq and SmD1, bind RNA and associate as homohexamers (bacterial Hfq) or heptamers (archaeal and eukaryotic proteins). The N-terminal domain of Dom34 has two important differences with members of the Sm family. First, the strands β1 and β2 from Dom34 occupy the location of strands β4 and β3 from the neighboring protomer in oligomeric Sm family members and therefore hamper the formation of an oligomeric structure as observed in Lsm proteins, for instance (Fig. 1*B*). Second, Dom34 lacks the Sm1 and Sm2 motifs implicated in RNA binding. Therefore, if the N-terminal domain of Dom34 is involved in RNA binding, the mechanism of recognition will be different to that of documented Sm family members.

Sequence analysis of Dom34 protein suggested that the central and C-terminal domains were structurally analogous to the eukaryotic class I release factor eRF1 (32). As illustrated in Fig. 1*C*, the central domains of eRF1 and Dom34 are very similar (r.m.s. deviation of 2.1 Å over 95 Cα atoms; 12% sequence identity (33)). In Dom34, this central domain (residues 136–277) has an α/β sandwich architecture with a five-stranded β-sheet (strand order β10, β9, β8, β11, and β12 with β9 antiparallel to the others) sandwiched by three helices (α3, α4, and αA) on one side and two (α5 and α7) on the other. The fold of the central domain is found in some endonucleases (*E. coli* RNase H1, RuvC resolvase, and the RNase H domain from human immunodeficiency virus reverse transcriptase (51–53)). The largest difference between the central domains of Dom34 and eRF1 resides in helix α3 that is 15 Å shorter in Dom34. This difference is very significant because the universally conserved GGQ motif of eRF1 is present in a loop at the C-terminal extremity of this helix. The central and C-terminal domains are connected by helix α7, which is kinked so that one-third belongs to the central domain and the remainder to the C-terminal domain. The latter adopts a mixed α/β-fold with a central four-stranded β-sheet (strand order β13, β16, β14, and β15 with β16 antiparallel to the others) flanked by two α-helices on each side. This domain is also structurally homologous to the eRF1 C-terminal domain (r.m.s. deviation of 2 Å over 96 Cα; 18% sequence identity; Fig. 3*A*).

Effect of Dom34 on GTP Binding by Hbs1—We further wanted to find out whether the structural analogies between Dom34 and eRF1 indicate deeper functional similarities. Dom34 is known to interact with Hbs1, a protein homologous to class II release factor eRF3 (34, 36). Eukaryotic release factors eRF1 and eRF3 form a heterodimer stabilized by GTP and govern translation termination (39, 54). The binding of eRF1 to eRF3 improves the affinity of eRF3 for GTP but not GDP by lowering the dissociation rate constant, hence eRF1 acts as a GTP dissociation inhibitor (55). The formation of the eRF1-eRF3-GTP complex is also necessary for coupling stop codon recognition and peptidyl-tRNA hydrolysis by eRF1 upon GTP hydrolysis by eRF3 (56) as well as for the post-translational modification of eRF1 by the Mtq2-Trm112 methyltransferase heterodimer (57).

We have investigated whether Dom34 influences guanine nucleotide binding to Hbs1. Yeast Dom34 and Hbs1 have been overexpressed and purified separately and stoichiometrically mixed with or without a nonhydrolysable analog of GTP (GMPPNP). Analytical gel filtration experiments showed that both proteins associate as a heterodimer independently of the presence of GTP (Fig. 1*D*). Next, the effect of Dom34 on the affinity of Hbs1 for guanine nucleotides has been investigated by fluorescence spectroscopy. Guanine dissociation constants to Hbs1 were measured by equilibrium fluorescence titration using the fluorescence signal at 448 nm of GTP/GDP derivatives carrying the Mant dye. The K_d values of Hbs1 alone for Mant-GDP and Mant-GTP were estimated to be 3.6 and 8.7 μM, respectively (measured in the presence of 2.5 mM $MgCl_2$, Table 2). The K_d of the Dom34-Hbs1 complex for Mant-GDP was comparable (4.2 μM; Table 2) but was lower for Mant-GTP (K_d = 1.6 μM; Table 2). As observed for the eRF1-eRF3 complex (55), the association of Dom34 with Hbs1 results in a 5–6-fold higher affinity for GTP.

DISCUSSION

A new mRNA degradation pathway (No-Go Decay or NGD) aimed at clearing cells from mRNAs that force ribosomes to pause during translation elongation has been described recently (26). NGD triggers endonucleolytic cleavage of these mRNAs and involves at least Dom34 and Hbs1, predicted to be structurally similar to translation termination factors eRF1 and eRF3, respectively (32, 36). However, the role of Dom34/Hbs1 in NGD remains largely unknown. The crystal structure of the *S. cerevisiae* Dom34 protein is an important step toward a better understanding of its function in the NGD pathway.

FIGURE 1. **Dom34 structure.** *A*, stereo view ribbon representation of the Dom34 structure. The chain is colored from *blue* (N-terminal) to *red* (C-terminal). Missing loops are represented by *dashed lines*. *B*, stereo view of the superposition of human SmD1 (*salmon*)-SmD2 (*gray*) and Dom34 N-terminal domains (*blue*). Only secondary structure elements of Dom34 are labeled. The side chains from the active site residues are shown as *ball and sticks*. The asparagine and arginine side chains from Sm1 and Sm2 motifs implicated in RNA binding are shown as *sticks*. *C*, stereo view of the superposition of human eRF1 (*salmon*) and Dom34 central domains (*blue*). Only secondary structure elements of Dom34 are labeled. The strictly conserved PGF motif from Dom34/Pelota and the GGQ motif from eRF1 are depicted as *ball and sticks* or *sticks*, respectively. *D*, gel filtration profiles of Hbs1, Dom34, and the Hsb1-Dom34 complex. Experiments were performed at 4 °C using a Superdex S200 10/30 HR from GE Healthcare with the following buffer: 50 mM Tris, pH 7.5, 500 mM NaCl, 10 mM β-mercaptoethanol. The content of each elution peak has been analyzed on a 14% SDS-PAGE (*right*). Fractions loaded on the gels are indicated by *bars under* the chromatograms.

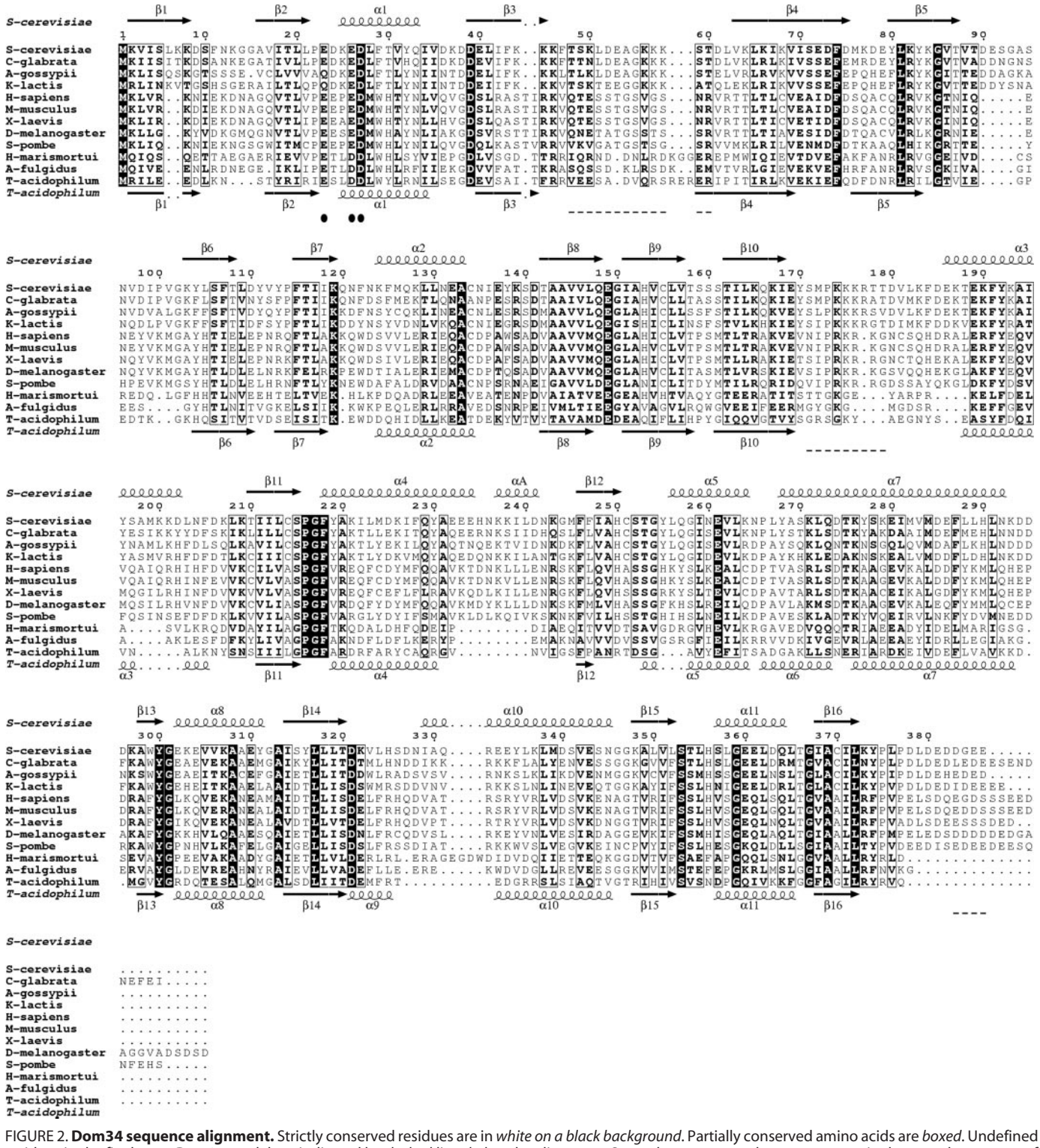


FIGURE 2. **Dom34 sequence alignment.** Strictly conserved residues are in *white on a black background*. Partially conserved amino acids are *boxed*. Undefined residues in the final yeast Dom34 model are indicated by *dashed lines below* the alignment. Secondary structure elements present in the crystal structures of yeast Dom34 (this study) and Ta Pelota (41) are shown *above* and *below* the alignment, respectively. Residues whose substitution by Ala in Ta Pelota resulted in disruption of the endonuclease activity are labeled by *filled circles below* the alignment.

Comparison with Ta Pelota—The yeast Dom34 structure confirms the organization in three domains as observed in the recently published structure of the archaeal ortholog *T. acidophilum* Pelota (the proteins display 21% sequence identity) (41). Despite this low sequence identity, the corresponding domains adopt very similar structures (each domain from Dom34 superposes onto the corresponding domain from Ta Pelota with r.m.s. deviation value ≈1.5–1.7 Å over 90–100 Cα atoms; 17% sequence identity for the N-terminal and central domains and 26% for the C-terminal domain). The main difference between the overall Ta Pelota and Dom34 structures resides in the relative positions of

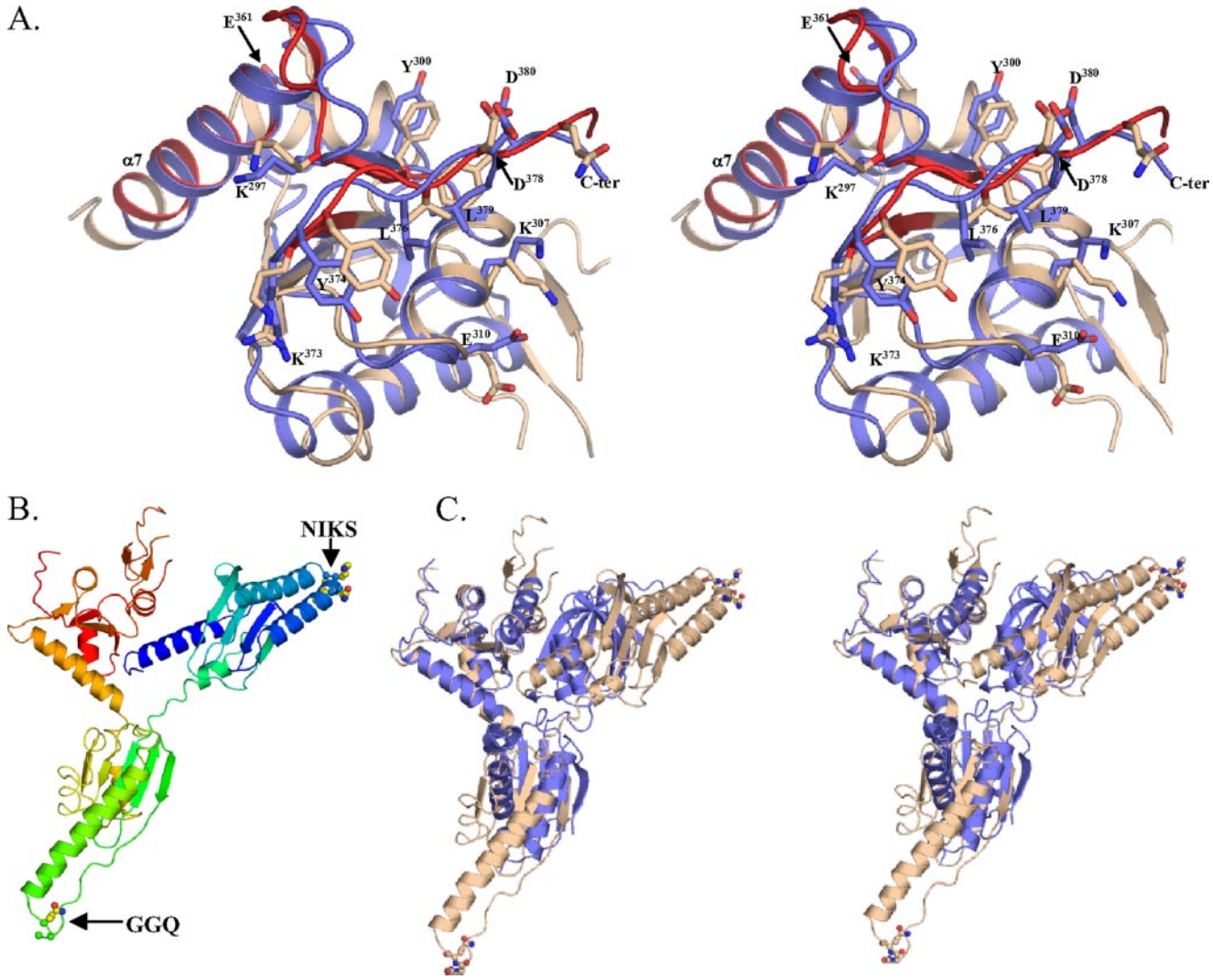


FIGURE 3. **Comparison between Dom34 and eRF1.** *A*, stereo view representation of the superimposition of the C-terminal domains of eRF1 (*salmon*) and Dom34 (*blue*). Side chains from residues strictly and highly conserved between eRF1 and Dom34 are shown as *sticks*. The backbone ribbon from the eRF1 amino acids described to be important for eRF3 binding are colored in *red*. For clarity, only Dom34 residues are labeled (see structure-based sequence alignment, supplemental Fig. S3). *B*, ribbon representation of the human eRF1 structure (PDB code 1DT9). The chain is colored from *blue* (N-terminal) to *red* (C-terminal). Amino acid side chains from the motifs involved in stop codon recognition (NIKS) and peptidyl-tRNA hydrolysis (GGQ) are shown as *ball and sticks*. *C*, stereo view ribbon representation of the superposition of eRF1 (*salmon*) and Dom34 (*blue*) structures onto their C-terminal domain. The orientation is the same as for *panel B*.

TABLE 2

Effect of Dom34 on the affinity of Hbs1 for nucleotides

Values in parentheses are those obtained in the presence of Mg^{2+} ions by Pisareva *et al.* (55) on eRF3 and eRF1-eRF3.

Complex	Nucleotide	K_d
		μM
Hbs1 (eRF3)	Mant-GDP	3.6 (0.8)
Hbs1-Dom34 (eRF3-eRF1)	Mant-GDP	4.2 (1.1)
Hbs1 (eRF3)	Mant-GTP	8.7 (23)
Hbs1-Dom34 (eRF3-eRF1)	Mant-GTP	1.6 (0.3)

their domains (Fig. 4). Superposition of the N-terminal domains of Ta Pelota and Dom34 reveals rotations of 30° and 90° along the horizontal and vertical axes for the C-terminal and central domains, respectively, resulting in a movement of 25 Å of the highly conserved PGF motif from the central domain (Fig. 4). Domain flexibility of lesser amplitude is also observed when the three copies of Dom34 (two SeMet labeled and one native) are superposed (Fig. S2).

As part of the NGD pathway, it was proposed that Dom34 is recruited to stalled ribosomes, where it directly or indirectly triggers endonucleolytic cleavage of mRNAs that cause a pause in translation elongation (26). Lee and colleagues (41) proved *in vitro* that the N-terminal domains from Ta Pelota and yeast Dom34 have divalent metal ion dependent endonuclease activity. These domains adopt a Sm-Lsm fold, which is found in proteins that interact with multiple factors implicated in RNA metabolism including splicing or mRNA decay (58). However, Ta Pelota and Dom34 are the first Sm-Lsm domains with ribonuclease activity. In bacteria, the homo-hexameric Sm protein Hfq interacts with the C-terminal domain of RNase E, a protein responsible for endonucleolytic cleavage of mRNAs. In eukaryotes, Lsm1–7 and U7 small nuclear ribonucleoprotein-specific Sm core complexes (both are heteroheptamers formed by the association of Lsm proteins) are involved in mRNA metabolism (59–61). In

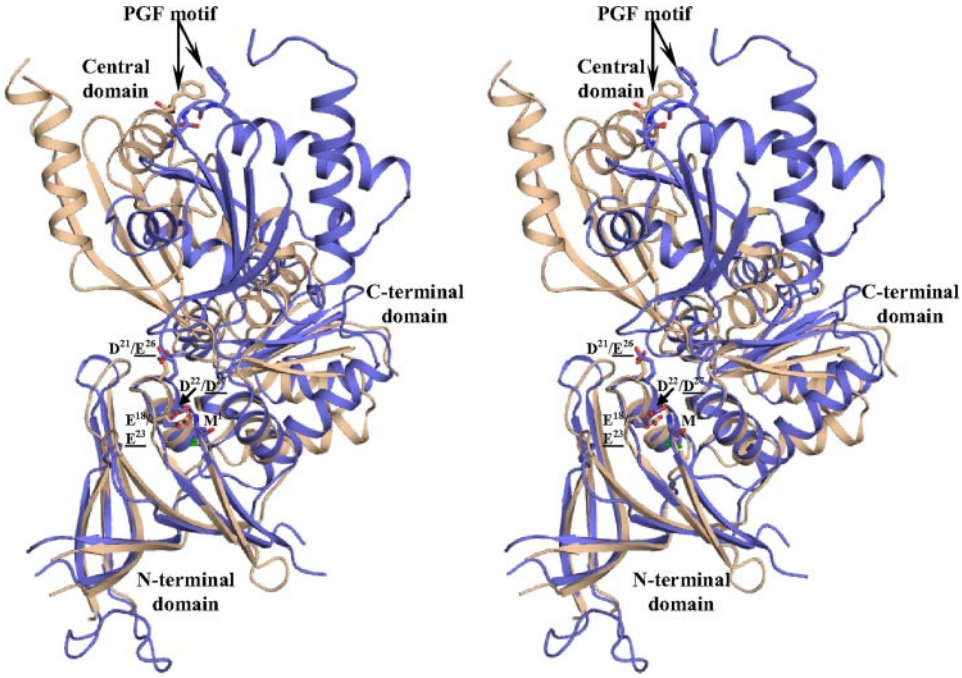


FIGURE 4. **Comparison with Ta Pelota crystal structure.** Stereo view ribbon representation of the superposition of Ta Pelota (*salmon*) and Dom34 (*blue*) structures onto their N-terminal domain. The active site residues and the strictly conserved PGF motifs are shown as *sticks*.

addition, Lsm proteins interact with various components of mRNA decay machinery such as the Xrn1 5′-3′ exonuclease, Pat1, and the exosome (59, 62, 63).

Single and double mutations of three well conserved acidic residues into alanines within the Ta Pelota N-terminal domain (E18A, E18A/D21A, E18A/D22A, and D21A/D22A) significantly diminished endonuclease activity (41). These three acidic residues can be perfectly superposed onto Glu^{23}, Glu^{26}, and Asp^{27} from Dom34 (Figs. 2 and 4). In the structure of Dom34, these residues are clearly less accessible than in Ta Pelota due to a difference in the orientation of the central domain relative to the N-terminal domain between the two proteins (see above). In Dom34, the central domain comes closer to the N-terminal domain than in Ta Pelota and the loop connecting strands β9 and β10 reduces the entry of the active site. Hence, a mRNA stem loop cannot access to the putative active site of Dom34 unless conformational changes bring the central and C-terminal domains in similar positions as those observed in Ta Pelota (Fig. 4). Endonuclease activity is often associated with acidic clusters as illustrated in RNase H1 and RuvC resolvase (51, 53), coordinating a divalent metal that activates the hydroxyl ion that attacks the pentacoordinate phosphorus intermediate (the higher activity is detected with Mg^{2+} and to a lesser extent with Mn^{2+}) (64). The loss of activity of the acidic mutants in Ta Pelota corroborates their involvement in catalysis. It should be noted, however, that the Glu^{18} and Asp^{22} side chains from Ta Pelota (Glu^{23} and Asp^{27} in Dom34) form direct charged hydrogen bonds with the N-terminal methionine amino group (Fig. 4 and supplemental Figs. S1). Therefore, mutations of these residues may affect protein stability and folding, causing indirectly loss of ribonuclease activity. This hypothesis is compatible with the observation that the Dom34 point mutants (E23A, E23K, E26A, E23K/E26A, E23K/E26K, E23K/D27A, E23K/D27K, E26A/D27A, and E26K/D27K) are produced as insoluble in *E. coli*, whereas wild type Dom34 N-terminal domain is soluble (41).

Dom34 Displays Structural and Functional Similarity to eRF1—It was previously noted that the global eRF1 envelope has a shape and charge distribution that mimics tRNAs (Fig. 3*B*) (33). Although functionally important sequence motifs of eRF1 were identified, little is known about its mechanism at the molecular level. The N-terminal domain carries the NIKS loop responsible for stop codon recognition and supposedly binds deeply into the ribosomal A-site (65, 66). The central domain contains the strictly conserved GGQ motif that interacts with the peptidyl transferase center and catalyzes peptide release (67). The C-terminal domain is implicated in the interaction with eRF3, a GTPase that facilitates eRF1 stop codon recognition and enhances polypeptide chain release (68–71). The overall organization as well as the structures of the central and C-terminal domains of Dom34 is similar to those of eRF1 (Figs. 1*C* and 3). The partners of eRF1 and Dom34, eRF3 and Hbs1, display 53% overall sequence similarity and hence possess similar three-dimensional structures. Considering the analogy between eRF1/Dom34 on one hand and eRF3/Hbs1 on the other, we hypothesize that Hbs1 interacts with Dom34 through the C-terminal domain of the latter. The structure of the eRF1-eRF3 complex is unknown but two-hybrid experiments mapped their interaction site to three distinct well conserved eRF1 regions: the 281–305 segment, the ^{411}GILRY^{415} motif (according to human eRF1 numbering), and the acidic C-terminal stretch (68, 70, 72). The superposition of the C-terminal domains from eRF1 and Dom34 reveals that the three eRF1 regions that interact with eRF3 have structural matches in Dom34 (Fig. 3*A*). First, the segment encompassing amino acids 281–305 from helix α8 in eRF1 (33) fits with residues 278–302 from helix α7 in Dom34. The sequences of these segments are well conserved among both eRF1 and Dom34 orthologs (Fig. S3). Second, the ^{411}GILRY^{415} eRF1 pentapeptide superposes well onto ^{370}CILKY^{374} from yeast Dom34. The side chains from Lys^{373} and Tyr^{374} are fully accessible to interact with Hbs1 (Fig. 3*A*) (70). Finally, the eRF1 C-terminal acidic stretch has been shown to be essential for the interaction with eRF3 in *S. cerevisiae* and *Schizosaccharomyces pombe* (68, 72). Dom34 also has an acidic C-terminal tail but shorter than for eRF1 (supplemental Fig. S3). Although most of the acidic residues from the C termini are disordered in both eRF1 and Dom34 structures, the few residues that are structured (situated at the N-terminal of the acidic peptide) superpose very well (Fig. 3*A*). Based on these

observations we propose that the mode of interaction between Hbs1 and Dom34 will be very similar to that of eRF1 with eRF3.

Despite the mentioned similarities between eRF1-eRF3 and Dom34-Hsb1 complexes we expect that they will interact differently with the ribosome. The eRF3 and Hbs1 ribosome binding sites probably overlap, but this may not be the case for eRF1-Dom34. It is known that the central domain of eRF1 contacts the peptidyl transferase center via its universally conserved GGQ motif, whereas the N-terminal domain binds deeply into the ribosomal A-site and contacts the stop codon through its NIKS signature. The structures of the N-terminal domain of Dom34 and eRF1 are unrelated and hence they likely interact differently with the ribosomal A-site. Several observations suggest that the N-terminal domain might even be excluded from the ribosomal A-site. First, the endonucleolytic cleavage performed *in vivo* by Dom34 seems to occur in the vicinity of the stem loop of the mRNA causing ribosomal stalls (26). Considering that stalled ribosomes abut against this stable stem loop, a ribosomal binding mode with the Dom34 N-terminal domain docked in the A-site would be incompatible with the observed cleavage sites. Second, because ribosomes in NGD become stalled by mRNA stem loops, there will be no stop codon present in the ribosomal A-site that therefore is likely be occupied by a cognate tRNA.

It was shown that eRF1, eRF3, and GTP act cooperatively to stimulate peptidyl-tRNA hydrolysis in the pre-termination complex (56). It remains an open question whether eRF3 increases the catalytic rate of peptide release and/or the affinity of eRF1 for the pre-termination complex. Considering the structural similarities between eRF1-eRF3 and Dom34-Hbs1 we suspected that GTP might also interfere with the function of Dom34. We therefore investigated whether the presence of GTP influences the interaction between Dom34 and Hbs1. We demonstrated that Hbs1 binds to guanine nucleotides *in vitro* using fluorescent GTP and GDP derivatives. We then showed that binding of Dom34 to Hbs1 enhances its affinity for Mant-GTP (5–6-fold effect) but not for Mant-GDP (Table 2). Similarly, a 75-fold increase in affinity for Mant-GTP and GTP is observed when eRF3 is bound to eRF1 (Ref. 55, see Table 2 for details).

Conclusion—The Hbs1-Dom34 complex from *S. cerevisiae* has been implicated in a newly described mRNA decay pathway called No-Go decay. This pathway releases ribosomes stalled in translation due to the presence of a stable stem loop within the mRNA and this complex triggers direct endonucleolytic cleavage of aberrant mRNA. Sequence analysis suggested that Dom34 and Hbs1 are related to translation termination factors eRF1 and eRF3, respectively. This is confirmed by the structure of yeast Dom34 and its archaeal ortholog Ta Pelota for the central and C-terminal domains, which are structurally similar to the corresponding domains in eRF1, but the N-terminal domain is unrelated. The domains and structural motifs in eRF1 that are important for the interaction with eRF3 are structurally conserved in Dom34. Dom34 enhances the affinity of Hbs1 for guanine nucleotides in the same way as eRF1 does with eRF3/GTP. Altogether, our data highlight both mechanistic similarities and differences between Dom34-Hbs1 and eRF1-eRF3 complexes. Eukaryotic cells have evolved two structurally related complexes able to interact with ribosomes either paused at a stop codon or stalled in translation by the presence of a stem loop. In both cases, they trigger ribosome release by catalyzing chemical bond hydrolysis.

Acknowledgments—We are grateful to Dr. B. Séraphin and M. A. Brooks for critical reading of the manuscript. We are indebted to Dr. K. Blondeau, B. Dray, and Dr. N. Leulliot for technical assistance. We thank Dr. P. Gouet for a special version of ESPRIPT used in supplemental Fig. S3.

REFERENCES

1. Echols, H., and Goodman, M. F. (1991) *Annu. Rev. Biochem.* **60,** 477–511
2. Ibba, M., and Soll, D. (1999) *Science* **286,** 1893–1897
3. Libby, R. T., and Gallant, J. A. (1991) *Mol. Microbiol.* **5,** 999–1004
4. Modrich, P., and Lahue, R. (1996) *Annu. Rev. Biochem.* **65,** 101–133
5. Ogle, J. M., and Ramakrishnan, V. (2005) *Annu. Rev. Biochem.* **74,** 129–177
6. Garneau, N. L., Wilusz, J., and Wilusz, C. J. (2007) *Nat. Rev. Mol. Cell. Biol.* **8,** 113–126
7. Holbrook, J. A., Neu-Yilik, G., Hentze, M. W., and Kulozik, A. E. (2004) *Nat. Genet.* **36,** 801–808
8. Kuzmiak, H. A., and Maquat, L. E. (2006) *Trends Mol. Med.* **12,** 306–316
9. Amrani, N., Sachs, M. S., and Jacobson, A. (2006) *Nat. Rev. Mol. Cell. Biol.* **7,** 415–425
10. Chang, Y. F., Imam, J. S., and Wilkinson, M. F. (2007) *Annu. Rev. Biochem.* **76,** 51–75
11. Conti, E., and Izaurralde, E. (2005) *Curr. Opin. Cell Biol.* **17,** 316–325
12. Maquat, L. E. (2004) *Nat. Rev. Mol. Cell. Biol.* **5,** 89–99
13. Rehwinkel, J., Raes, J., and Izaurralde, E. (2006) *Trends Biochem. Sci.* **31,** 639–646
14. Amrani, N., Ganesan, R., Kervestin, S., Mangus, D. A., Ghosh, S., and Jacobson, A. (2004) *Nature* **432,** 112–118
15. Behm-Ansmant, I., Gatfield, D., Rehwinkel, J., Hilgers, V., and Izaurralde, E. (2007) *EMBO J.* **26,** 1591–1601
16. Buhler, M., Steiner, S., Mohn, F., Paillusson, A., and Muhlemann, O. (2006) *Nat. Struct. Mol. Biol.* **13,** 462–464
17. Le Hir, H., Gatfield, D., Izaurralde, E., and Moore, M. J. (2001) *EMBO J.* **20,** 4987–4997
18. Le Hir, H., Moore, M. J., and Maquat, L. E. (2000) *Genes Dev.* **14,** 1098–1108
19. Frischmeyer, P. A., van Hoof, A., O'Donnell, K., Guerrerio, A. L., Parker, R., and Dietz, H. C. (2002) *Science* **295,** 2258–2261
20. Maquat, L. E. (2002) *Science* **295,** 2221–2222
21. van Hoof, A., Frischmeyer, P. A., Dietz, H. C., and Parker, R. (2002) *Science* **295,** 2262–2264
22. Karzai, A. W., Roche, E. D., and Sauer, R. T. (2000) *Nat. Struct. Biol.* **7,** 449–455
23. Keiler, K. C., Waller, P. R., and Sauer, R. T. (1996) *Science* **271,** 990–993
24. Araki, Y., Takahashi, S., Kobayashi, T., Kajiho, H., Hoshino, S., and Katada, T. (2001) *EMBO J.* **20,** 4684–4693
25. Clement, S. L., and Lykke-Andersen, J. (2006) *Nat. Struct. Mol. Biol.* **13,** 299–301
26. Doma, M. K., and Parker, R. (2006) *Nature* **440,** 561–564
27. Tollervey, D. (2006) *Nature* **440,** 425–426
28. Davis, L., and Engebrecht, J. (1998) *Genetics* **149,** 45–56
29. Adham, I. M., Sallam, M. A., Steding, G., Korabiowska, M., Brinck, U., Hoyer-Fender, S., Oh, C., and Engel, W. (2003) *Mol. Cell. Biol.* **23,** 1470–1476
30. Eberhart, C. G., and Wasserman, S. A. (1995) *Development* **121,** 3477–3486
31. Xi, R., Doan, C., Liu, D., and Xie, T. (2005) *Development* **132,** 5365–5374
32. Koonin, E. V., Bork, P., and Sander, C. (1994) *Nucleic Acids Res.* **22,** 2166–2167
33. Song, H., Mugnier, P., Das, A. K., Webb, H. M., Evans, D. R., Tuite, M. F.,

Hemmings, B. A., and Barford, D. (2000) *Cell* **100,** 311–321
34. Carr-Schmid, A., Pfund, C., Craig, E. A., and Kinzy, T. G. (2002) *Mol. Cell. Biol.* **22,** 2564–2574
35. Inagaki, Y., Blouin, C., Susko, E., and Roger, A. J. (2003) *Nucleic Acids Res.* **31,** 4227–4237
36. Wallrapp, C., Verrier, S. B., Zhouravleva, G., Philippe, H., Philippe, M., Gress, T. M., and Jean-Jean, O. (1998) *FEBS Lett.* **440,** 387–392
37. Negrutskii, B. S., and El'skaya, A. V. (1998) *Prog. Nucleic Acids Res. Mol. Biol.* **60,** 47–78
38. Kisselev, L. L., and Buckingham, R. H. (2000) *Trends Biochem. Sci.* **25,** 561–566
39. Zhouravleva, G., Frolova, L., Le Goff, X., Le Guellec, R., Inge-Vechtomov, S., Kisselev, L., and Philippe, M. (1995) *EMBO J.* **14,** 4065–4072
40. Nelson, R. J., Ziegelhoffer, T., Nicolet, C., Werner-Washburne, M., and Craig, E. A. (1992) *Cell* **71,** 97–105
41. Lee, H. H., Kim, Y. S., Kim, K. H., Heo, I., Kim, S. K., Kim, O., Kim, H. K., Yoon, J. Y., Kim, H. S., Kim do, J., Lee, S. J., Yoon, H. J., Kim, S. J., Lee, B. G., Song, H. K., Kim, V. N., Park, C. M., and Suh, S. W. (2007) *Mol. Cell* **27,** 938–950
42. Hendrickson, W. A., Horton, J. R., and LeMaster, D. M. (1990) *EMBO J.* **9,** 1665–1672
43. Roth, M., Carpentier, P., Kaikati, O., Joly, J., Charrault, P., Pirocchi, M., Kahn, R., Fanchon, E., Jacquamet, L., Borel, F., Bertoni, A., Israel-Gouy, P., and Ferrer, J. L. (2002) *Acta Crystallogr. D Biol. Crystallogr.* **58,** 805–814
44. Kabsch, W. (1993) *J. Appl. Cryststallogr.* **26,** 795–800
45. Schneider, T. R., and Sheldrick, G. M. (2002) *Acta Crystallogr. D Biol. Crystallogr.* **58,** 1772–1779
46. Terwilliger, T. C. (1999) *Acta Crystallogr. D Biol. Crystallogr.* **55,** 1863–1871
47. Terwilliger, T. C., and Berendzen, J. (1999) *Acta Crystallogr. D Biol. Crystallogr.* **55,** 849–861
48. Adams, P. D., Grosse-Kunstleve, R. W., Hung, L. W., Ioerger, T. R., McCoy, A. J., Moriarty, N. W., Read, R. J., Sacchettini, J. C., Sauter, N. K., and Terwilliger, T. C. (2002) *Acta Crystallogr. D Biol. Crystallogr.* **58,** 1948–1954
49. Schumacher, M. A., Pearson, R. F., Moller, T., Valentin-Hansen, P., and Brennan, R. G. (2002) *EMBO J.* **21,** 3546–3556
50. Kambach, C., Walke, S., Young, R., Avis, J. M., de la Fortelle, E., Raker, V. A., Luhrmann, R., Li, J., and Nagai, K. (1999) *Cell* **96,** 375–387
51. Ariyoshi, M., Vassylyev, D. G., Iwasaki, H., Nakamura, H., Shinagawa, H., and Morikawa, K. (1994) *Cell* **78,** 1063–1072
52. Davies, J. F., 2nd, Hostomska, Z., Hostomsky, Z., Jordan, S. R., and Matthews, D. A. (1991) *Science* **252,** 88–95
53. Yang, W., Hendrickson, W. A., Crouch, R. J., and Satow, Y. (1990) *Science* **249,** 1398–1405
54. Kobayashi, T., Funakoshi, Y., Hoshino, S., and Katada, T. (2004) *J. Biol. Chem.* **279,** 45693–45700
55. Pisareva, V. P., Pisarev, A. V., Hellen, C. U., Rodnina, M. V., and Pestova, T. V. (2006) *J. Biol. Chem.* **281,** 40224–40235
56. Alkalaeva, E. Z., Pisarev, A. V., Frolova, L. Y., Kisselev, L. L., and Pestova, T. V. (2006) *Cell* **125,** 1125–1136
57. Heurgue-Hamard, V., Graille, M., Scrima, N., Ulryck, N., Champ, S., van Tilbeurgh, H., and Buckingham, R. H. (2006) *J. Biol. Chem.* **281,** 36140–36148
58. Wilusz, C. J., and Wilusz, J. (2005) *Nat. Struct. Mol. Biol.* **12,** 1031–1036
59. Bouveret, E., Rigaut, G., Shevchenko, A., Wilm, M., and Seraphin, B. (2000) *EMBO J.* **19,** 1661–1671
60. Pillai, R. S., Will, C. L., Luhrmann, R., Schumperli, D., and Muller, B. (2001) *EMBO J.* **20,** 5470–5479
61. Tharun, S., He, W., Mayes, A. E., Lennertz, P., Beggs, J. D., and Parker, R. (2000) *Nature* **404,** 515–518
62. Fromont-Racine, M., Mayes, A. E., Brunet-Simon, A., Rain, J. C., Colley, A., Dix, I., Decourty, L., Joly, N., Ricard, F., Beggs, J. D., and Legrain, P. (2000) *Yeast* **17,** 95–110
63. Lehner, B., and Sanderson, C. M. (2004) *Genome Res.* **14,** 1315–1323
64. Yang, W., and Steitz, T. A. (1995) *Structure* **3,** 131–134
65. Bertram, G., Bell, H. A., Ritchie, D. W., Fullerton, G., and Stansfield, I. (2000) *RNA* (*Cold Spring Harbor*) **6,** 1236–1247
66. Chavatte, L., Seit-Nebi, A., Dubovaya, V., and Favre, A. (2002) *EMBO J.* **21,** 5302–5311
67. Frolova, L. Y., Tsivkovskii, R. Y., Sivolobova, G. F., Oparina, N. Y., Serpinsky, O. I., Blinov, V. M., Tatkov, S. I., and Kisselev, L. L. (1999) *RNA* (*Cold Spring Harbor*) **5,** 1014–1020
68. Eurwilaichitr, L., Graves, F. M., Stansfield, I., and Tuite, M. F. (1999) *Mol. Microbiol.* **32,** 485–496
69. Frolova, L., Le Goff, X., Zhouravleva, G., Davydova, E., Philippe, M., and Kisselev, L. (1996) *RNA* (*Cold Spring Harbor*) **2,** 334–341
70. Merkulova, T. I., Frolova, L. Y., Lazar, M., Camonis, J., and Kisselev, L. L. (1999) *FEBS Lett.* **443,** 41–47
71. Salas-Marco, J., and Bedwell, D. M. (2004) *Mol. Cell. Biol.* **24,** 7769–7778
72. Ito, K., Ebihara, K., and Nakamura, Y. (1998) *RNA* (*Cold Spring Harbor*) **4,** 958–972