

# Ddx42p—a human DEAD box protein with RNA chaperone activities

Heike Uhlmann-Schiffler\*, Carolin Jalal and Hans Stahl

FR 2.3 Medical Biochemistry and Molecular Biology, Theoretical Medicine, University of the Saarland, D-66421 Homburg, Germany

Received October 14, 2005; Revised and Accepted December 1, 2005

## ABSTRACT

**The human gene *ddx42* encodes a human DEAD box protein highly homologous to the p68 subfamily of RNA helicases. In HeLa cells, two *ddx42* poly(A)<sup>+</sup> RNA species were detected both encoding the nuclear localized 938 amino acid Ddx42p polypeptide. Ddx42p has been heterologously expressed and its biochemical properties characterized. It is an RNA binding protein, and ATP and ADP modulate its RNA binding affinity. Ddx42p is an NTPase with a preference for ATP, the hydrolysis of which is enhanced by various RNA substrates. It acts as a non-processive RNA helicase. Interestingly, RNA unwinding by Ddx42p is promoted in the presence of a single-strand (ss) binding protein (T4gp32). Ddx42p, particularly in the ADP-bound form (the state after ATP hydrolysis), also mediates efficient annealing of complementary RNA strands thereby displacing the ss binding protein. Ddx42p therefore represents the first example of a human DEAD box protein possessing RNA helicase, protein displacement and RNA annealing activities. The adenosine nucleotide cofactor bound to Ddx42p apparently acts as a switch that controls the two opposing activities: ATP triggers RNA strand separation, whereas ADP triggers annealing of complementary RNA strands.**

## INTRODUCTION

Nucleic acid metabolism in all organisms is made possible by the activity of a wealth of highly specific DNA and RNA helicases, with ~600 different proteins per eukaryotic cell [(1), reviewed in (2)]. Consistent with their central roles in nucleic acid metabolic processes, an increasing number of helicases are found to be involved in severe diseases, often associated with significant cancer susceptibility, and in viral

diseases (3–5). Helicases, by classical definition, are enzymes, which disrupt the hydrogen bonds of nucleic acid secondary structures to form single strands in an NTP-dependent manner. While several DNA helicases are also known to disrupt protein–DNA interactions [recently summarized in (6)], to date only two DExH/D proteins (the processive NPH-II helicase and the non-processive Ded1p) have been directly shown to couple ATP hydrolysis with the removal of proteins bound to RNA (7,8). Two spliceosomal DExH/D proteins from yeast have been inferred to rearrange ribonucleoprotein (RNP) complexes (9,10).

RNA helicases are of particular interest owing to the recently growing perception of RNA metabolism as a fundamental regulatory element in cellular biochemistry and metabolism. The most extensively studied are NPH-II helicase from vaccinia virus (11) and NS3 protein from hepatitis C virus (HCV) (12,13), which are critical for virus replication and capable of processive unwinding of RNA duplexes. Interestingly, the latter protein unwinds RNA as well as DNA duplexes (14), a property shared, e.g. by Dbp9p, a yeast DEAD box protein involved in ribosome biogenesis (15) and Simian Virus 40 large T antigen (SV40 T ag), a hexameric replicative helicase (16–18). However, the number of biochemically, let alone physiologically (with particular respect to their *in vivo* targets), well-characterized cellular RNA helicases is limited. Many of them seem to disrupt only short RNA duplexes, e.g. eukaryotic translation initiation factor 4A (eIF4A) [reviewed in (19,20)]. eIF4A appears to represent a ‘minimal’ RNA helicase, consisting of a core sequence common to most of these proteins, and is regarded as the prototype of DExH/D proteins. eIF4A, a non-processive helicase on RNA or DNA/RNA duplexes, unwinds 10–15 bp, depending on the stability of the respective duplex (21), and its activity is modulated by other translation initiation factors (22,23). Ded1p from yeast (24) appears closely related to eIF4A, although it has also been found in spliceosomes (25). Two other DEAD box proteins from higher eukaryotes studied in detail are p68 (26) and p72 (27). Apart from having low processive RNA helicase activity, they are capable of catalyzing RNA rearrangement events (28) and may thus participate in

\*To whom correspondence should be addressed. Tel: +49 6841 16 26294; Fax: +49 6841 16 26521; Email: heike.uhlmann.schiffler@uniklinikum-saarland.de

reorganization of RNA structures *in vivo*. While the exact biological role(s) of p68 and p72 is(are) as yet unknown, they seem to be related to cell growth and/or division and have appeared in various functional contexts: as transcriptional co-activators [e.g. (29,30)], as spliceosome components (31), in alternative splicing (32,33) and in rRNA metabolism (34). p68 has also been found in a complex with 5-methylcytosine DNA glycosylase (35). Very recently, it emerged as a cellular factor involved in HCV replication (36). Interestingly, p72 mRNA translation can start in-frame upstream of the AUG codon resulting in an N-terminally extended protein, p82, which is biochemically similar to p72 (37). Most notably, no DEAD box protein is as closely related to another as p68 and p72/p82 are. Their cores show >90% homology, and their biochemical profile is similar. This led to the proposal that they constitute a p68-type subfamily unique to 'higher' eukaryotic cells while the unicellular eukaryote *Saccharomyces cerevisiae* contains only one homologous protein, Dbp2p (38). Dbp2p (39) and p68 (40) are posttranscriptionally autoregulated via a conserved intron. A similar regulation mechanism for p72/82 is likely as a pre-mRNA retaining an unspliced intron is detected in the nucleus as well (37). These observations are consistent with a fundamental role for each of these proteins in the cell.

In this work, we identified, purified and extensively characterized a human DEAD box protein, Ddx42p, which was previously found to be associated with the SF3b subunit of human 17S U2 snRNP (41). According to its sequence, the protein is more closely related to the p68 subfamily than any other known DEAD box protein, but differs with respect to its biochemical characteristics. It acts as an NTPase and non-processive ATP-dependent RNA helicase that can be stimulated by a single-strand (ss) binding protein. Annealing of partially complementary ssRNA by Ddx42p, even in the presence of excess ss binding protein, is mediated most efficiently in the presence of ADP so that RNA helicase and RNA annealing activities appear to proceed in a regulated manner. These combined activities make Ddx42p well suited to restructuring of RNP complexes.

## MATERIALS AND METHODS

### Biochemicals, enzymes and routine procedures

Restriction endonucleases were purchased from Roche, plasmids pGEM-7Zf(-) and pGEM-3Z as well as M-MLV Reverse transcriptase were from Promega. HotStarTaq DNA polymerase was obtained from Qiagen. Cell culture media were from PAA Laboratories GmbH. All other reagents were of the highest purity grade commercially available. First strand cDNA synthesis and PCR were performed according to manufacturer's instructions. Cloning procedures, plasmid preparation, electrophoresis and western blotting were carried out according to standard protocols (42). Gels containing <sup>35</sup>S-labeled samples were incubated in enhancer solution (1 M salicylic acid, pH 8) for 30 min, dried and exposed at -70°C.

DNA sequencing was done by Dr R. Klein, University of Kaiserslautern, Germany. Database searches were done by BLAST (<http://www.ncbi.nlm.nih.gov>) and ExPASy (<http://www.expasy.ch>). Alignments were performed using

MacDNASIS with the implemented PAM250 homology scoring matrix (Hitachi Software Engineering Co., Ltd).

### Oligonucleotides, RNA substrates and antibodies

Sequences and use of oligonucleotides used in PCR (Supplementary data Table 1) as well as RNA substrates and their preparation (Supplementary data Table 2) are detailed in Supplementary Data. For preparation of Ddx42p-specific antibodies  $\alpha$ -D42N and  $\alpha$ -D42C, peptides corresponding to an N-terminal (H<sub>2</sub>N-AGKKEEPKLPQQSHS-COOH) and a C-terminal Ddx42p sequence (H<sub>2</sub>N-DSPRHGDGGRHGD-COOH), respectively, were synthesized on an Applied Biosystems 433A Peptide Synthesizer, coupled to hemocyanine, and used for multiple immunizations of rabbits. The respective antisera were used for detection of Ddx42p in western blots. Actin was detected with polyclonal antibody Actin (I-19)-R (Santa Cruz Biotechnology, Inc.).

### Northern blotting

Poly(A)<sup>+</sup> RNA was isolated from different mammalian cell lines (~10<sup>7</sup> cells each) using the Oligotex Direct mRNA kit (Qiagen) according to supplier's recommendations. A ddx42-specific antisense RNA probe (complementary to nt 1797–3074 of the ddx42 cDNA) was synthesized by NheI-digest of pGEMddx42 (see below) and *in vitro* transcription with T7 RNA polymerase in the presence of Dig-RNA labeling mix (Roche). Northern blotting of 100 ng poly(A)<sup>+</sup> RNA was carried out using the Dig-labeled ddx42-specific antisense RNA probe and detected with CSPD (Roche). A Dig-labeled  $\beta$ -actin RNA probe (Roche) was used to detect the ubiquitously abundant  $\beta$ -actin mRNA as a loading control.

### RT-PCR

DNase-treated HeLa poly(A)<sup>+</sup> RNA was used as a template in a two-step protocol including M-MLV reverse transcriptase, 1  $\mu$ g RNA and 0.5  $\mu$ g of oligo(dT)<sub>18</sub> primer at 42°C for 1 h. HotStarTaq polymerase was used in subsequent PCRs. In respective negative controls, the poly(A)<sup>+</sup> RNA was directly used in PCRs and reconfirmed the absence of traces of genomic DNA. The RT-PCR products were analyzed by DNA sequencing.

### *In vitro* translation of ddx42

PCR with primers ddx42\_N and ddx42\_C resulted in a full-length ddx42 coding sequence (cds), which was cloned into pGEM-7Zf(-) via SacI/BamHI, yielding pGEMddx42. Insert sequences were verified by DNA sequencing. The plasmid was linearized by BamHI and transcribed *in vitro* with SP6 RNA polymerase using the Riboprobe system (Promega) to yield ddx42 RNA. Concentrations of transcripts were determined photometrically, and their quality was checked by agarose gel electrophoresis. *In vitro* translation in a reticulocyte lysate was performed using <sup>35</sup>S-Met according to the Flexi Rabbit protocol (Promega). The translation mix was separated by SDS-PAGE and the product (Ddx42p-TL) was visualized by autoradiography.

### Cell culture and cellular extracts

Mammalian cell lines (HeLa, Hep3B, COS, TC7, HaCaT) were cultivated in DMEM plus 10% fetal calf serum to

apparent confluency. For transient expression in COS cells, the *ddx42* cds was recloned into pCMV5, a mammalian expression vector, via MluI/BamHI (pCddx42). A C-terminal deletion mutant was obtained by BmgBI/SmaI digest of pCddx42, removal of the small fragment and plasmid recircularization (pCddx42Δ). DNA constructs were verified by nucleotide sequencing. COS cells were transfected with the respective plasmid using Metafectene (Biontex Laboratories GmbH).

Cellular extracts were obtained by different methods. All steps were carried out at 4°C. About  $5 \times 10^7$  cells were harvested by centrifugation at 3000 *g* for 5 min. (i) The cell pellet was extracted in RIPA buffer (TE, pH 8.0, 0.1% SDS) and spun at 100 000 *g* for 30 min. (ii) Cells were resuspended and swollen in a hypotonic buffer (10 mM HEPES, pH 7.5, 5 mM MgCl<sub>2</sub> and 2 mM DTT) and dounced until lysis occurred. After centrifugation at 20 000 *g* for 10 min, the pellet contained nuclei, mitochondria and membranes. The supernatant represented the soluble cytosolic fraction and was supplemented with Complete Protease Inhibitor Cocktail (Roche). (iii) The cell pellet was resuspended in lysis buffer (10 mM MES, pH 6.2, 10 mM NaCl, 1.5 mM MgCl<sub>2</sub>, 5 mM DTT, 1 mM EGTA, 1 mM phenylmethylsulfonyl fluoride (PMSF), 10% glycerol and 1% NP-40, supplemented with Complete Protease Inhibitor Cocktail) and recentrifuged at 3000 *g* for 5 min. The supernatant was labeled 'ce' (soluble cellular extract). The pellet consisted mainly of nuclei, which were extracted at pH 10.5 as described in Ref. (26), neutralized, and the extract labeled 'ne' (nuclear extract).

#### Purification of Ddx42p from *Escherichia coli*

For expression of Ddx42p in *E.coli*, the cds was recloned via SacI/KpnI into the bacterial expression vector pRSET, yielding N-terminally hexa-His tagged Ddx42p. Optimum expression in BL21DE3 was observed 2 h after induction with isopropyl-β-D-thiogalactopyranoside. All subsequent steps were carried out at 4°C. Cells were harvested by centrifugation, washed once with phosphate-buffered saline, and resuspended in lysis/wash buffer (25 mM HEPES, pH 7.9, 500 mM NaCl, 0.5% Triton X-100, 5 mM β-mercaptoethanol and 5 mM imidazole) supplemented with 1× Complete™ protease inhibitor cocktail without EDTA (Roche) and 50 μg/ml PMSF. Cells were lysed by sonication (4 × 10 pulses for 1 s) and spun at 50 000 *g* for 25 min. The supernatant was subjected to immobilized metal affinity chromatography (IMAC) on a TALON metal affinity resin (BD Biosciences), washed with excess buffer, and eluted with buffer containing 150 mM imidazole. Fractions containing Ddx42p were further purified by gel filtration on a Superdex 200 HR column (Amersham Biosciences).

#### NTPase activity test

A modified malachite green method (43) was used to determine hydrolysis of NTPs by Ddx42p. Various amounts of Ddx42p in assay buffer (25 mM Tris, pH 7.5, 1.5 mM DTT, 50 mM NaCl, 2 mM MgCl<sub>2</sub> and 30 μg/ml BSA) were incubated with 250 μM of each NTP in separate reactions at 37°C for 15 min. To test for stimulation of the reaction, 500 ng poly(C), total RNA, tRNA, poly(A)<sup>+</sup> RNA, *in vitro* transcript, or chemically synthesized complete double-stranded (ds) RNA was added. Autohydrolysis of the NTPs

was checked in respective control samples lacking Ddx42p. Incubation was stopped by addition of 0.25 N H<sub>2</sub>SO<sub>4</sub> and incubation on ice for 10 min. The precipitate was spun and the supernatant mixed with the color reagent. Released P<sub>i</sub> was determined by measuring A<sub>630</sub> after 10 min and a standard curve which had been prepared under identical conditions. Kinetic analysis was done by fitting a hyperbolic function [ $f = a \cdot x / (b + x)$ ] to the data.

#### Preparation of RNA substrates for RNA helicase and annealing experiments

RNA single strands were obtained by *in vitro* transcription with T7 and SP6 RNA polymerases of modified pGEM-3Z or pGEM7Zf(-) plasmids with simultaneous [ $\alpha$ -<sup>32</sup>P]CTP labeling in the T7 reaction (Supplementary Data and Table 2). Unincorporated radiolabel was removed by gel filtration on Sephadex G-50. The amount of radiolabeled RNA was determined by Cerenkov counting and the quality by SDS-PAGE and autoradiography. Concentration and quality of the SP6-transcribed RNAs was determined photometrically and by agarose gel electrophoresis. Two RNA single strands with complementary regions were used as substrates in RNA annealing experiments. Partially dsRNA helicase substrates were obtained by hybridization at 40°C for 4 h of respective ssRNAs. Theoretical melting temperatures ( $T_m$ s) as a measure for RNA duplex stability were calculated by the program Oligonucleotide Properties Calculator (<http://www.basic.northwestern.edu/biotools/OligoCalc.html>) according to the nearest neighbor method (44) using RNA thermodynamic properties (45), and  $T_m$  was lowered by % non-complementary bases for loop-containing substrates.

#### RNA-dependent assays with Ddx42p

Binding of purified Ddx42p to RNA was assayed by incubation of 50 nM protein with 75 nM (in nucleotides) of a <sup>32</sup>P-labeled ssRNA (MT7) or partial duplex (M) in 25 mM Tris-HCl, pH 7.5 and 3.5 mM MgCl<sub>2</sub> at 37°C for 10 min. The reaction was stopped by adding 0.1% glutar aldehyde and subjected to 4% PAGE, followed by autoradiography. RNA binding affinity of purified Ddx42p and T4gp32 was studied in a double-filter RNA binding assay according to Ref. (46) in 25 mM Tris, pH 7.5, 50 mM NaCl and 35 nM (in nucleotides) of a <sup>32</sup>P-labeled ssRNA (MT7). To test for the effect of the adenosine nucleotide, 4 mM of either MgATP, MgATPγS or MgADP was added. Reaction mixtures containing various amounts of protein were incubated at room temperature for 10 min and subsequently filtered through nitrocellulose and nylon filters under vacuum. Radioactivity bound to the filters was quantified by Cerenkov counting, and the fraction of RNA bound to the protein determined.

RNA helicase experiments were performed in NTPase assay buffer containing 10 U/sample RNase inhibitor, 4 mM MgCl<sub>2</sub>, 4 mM ATP and 35 nM (in nucleotides) helicase substrate. The reaction was started by addition of variable amounts of Ddx42p. In respective control reactions, individual compounds were omitted. Samples with T4gp32 were pre-incubated (before addition of Ddx42p) for 5 min with a 100-fold molar excess of T4gp32 over RNA (in nucleotides); a 3-fold molar excess theoretically allows for complete ssRNA coating [calculated from (47,48)].

|        |   |
|--------|---|
| p68    | FEKNFYQEHPDLARRTAQEVETRYRRSKEITVVRGHNCPKPVLNIFYEANFPANVMDVIARQN |
| Ddx42p | FEKNFYNEHEEETNLTPQQLIDLRLKLNLRVSGAAPPRPGSSFAHFGFDEQLMHQIRKSE    |
| eIF4A  | -----MSASQDSRSRDNGPDGMEPEGVIESNWEIVDSFDDMNLSESLLRGIYAYG         |
| p68    | FTEPTAIQAQGWFPVALSGLDMVGVAQTGSGKTLSTLPAIVHINHQPFLERGDGPICLVL    |
| Ddx42p | YTQPTPIQCQGVFVALSGRDMIGIAKTGSGKTAAFIWPMLIHIMDQKELEPGDGPIAVIV    |
| eIF4A  | FEKPSAIQQRALLPCKIKGYDVIAQAQSGTGKTATF----AISILQQIELDL-KATQALVL   |
| p68    | APTRELAQQVQVAAEYCRACRLKSTCIYGGAPKGPQIRDLE-RGVEICTATPGRLLIDFL    |
| Ddx42p | CPTRELCCQIHAECKRFKAYNLRSVAVYGGGSMWEQAKALQ-EGAEIVVCTPGRLLIDHV    |
| eIF4A  | APTRELAQQIQKVVVALGDMGASCHACIGGTNVRAEVQKLQMEAPHIIVGTPGRVFDML     |
| p68    | ECGKTNLRRTTYLVLDEADRMLDMGFEPQIRKIVDQIRPDRQTLMWSATWPKEVRQLAED    |
| Ddx42p | KKKATNLQRVSYLVFDEADRMFDMGFQYQVRSIASHVRPDRQTLTLLFSATFRKKIEKLRD   |
| eIF4A  | NRRLSPKYIKMFVLDEADEMLSRGFKDQIYDIFQKLNSTQVVLSSATMPSDVLVETKK      |
| p68    | FLKDYIHINIGALELSANHNILQIVDVCH-DVEKDEKLIR-LMEEIMSEKENKTIVFVET    |
| Ddx42p | LLIDPIRVVQGDIG-EANEDVTQIVEILHSGSPKWNWLTR-RLVEF--TSSGSLVLLFVTK   |
| eIF4A  | FMRDPIRILVKKEE-LTLEGIRQF--YINVEREEWKLDLTLCDLYET--LTITQAVIFINT   |
| p68    | KRRCDDELTRKMRRDGPAMGIHGDKSQQERDQVWVLEFVKHGKAPILIATDVASRGLDVEDV  |
| Ddx42p | KANAELANLQEGHNLGLLHGDMQSERNKVISDFKKKDIIPVLVATDVAARGLDIPSI       |
| eIF4A  | RRKVDWLTEKMHARDFTVSAMHGDMQKERDVMIMREFRSGSSRVLITTDLLARGIDVQQV    |
| p68    | KFVINVDYPNSSDYIHRIGRTARSTKTGTAYTFFTPNNIKQVSDLISVLRANQAINPK      |
| Ddx42p | KTVINYDVARIDTHTHRIGRTGRAGEKGVAYTLLTPKDSNFAGDLVRNLEGANQHVSKE     |
| eIF4A  | SLVINVDLPTNRENYIHRIGRGRFGRKGVAINMVTEDDKRTLRLDIETFYNTSIEEMPLN    |
| p68    | LLQLV   |
| Ddx42p | LLDLA   |
| eIF4A  | VADLI   |

**Figure 1.** The core regions of Ddx42p and p68 are highly homologous. The sequences of p68 (amino acids 54–475), Ddx42p (amino acids 213–632) and eIF4A (amino acids 1–406) were aligned. Residues that are homologous (i.e. identical or similar) in Ddx42p and either p68 or eIF4A are highlighted in gray.

RNA annealing experiments were designed similarly except that two RNA single strands with complementary portions were the substrates of the reaction. The concentrations of individual compounds were varied as stated in the figure legends. For annealing assays in the presence of T4gp32, reaction conditions were identical except that RNA substrate concentration was reduced by one half in order to achieve 100% annealing in the control reaction. Variable amounts of T4gp32 (at least 20-fold molar excess over RNA) were again added 5 min before addition of Ddx42p. After incubation of RNA helicase or RNA annealing samples at 37°C for 30 min, SDS-PAGE loading buffer containing 4 mM EDTA was added and samples applied to a 10% gel. Radiolabeled RNA was detected by autoradiography, and quantification of dsRNA and ssRNA was done by ImageQuant (Molecular Dynamics).

## RESULTS

### Ddx42p sequence comparisons

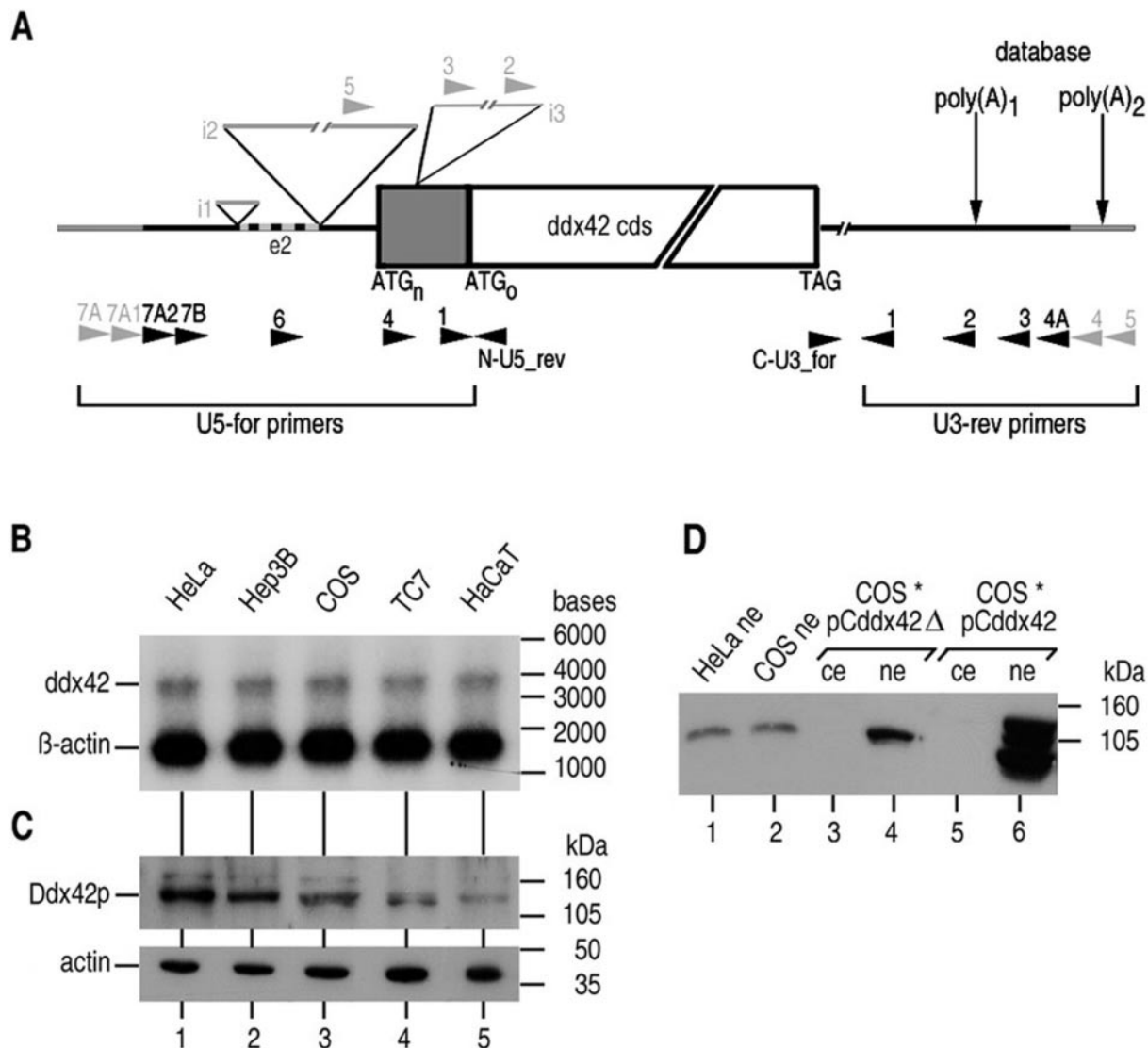
PCR with degenerate primers derived from p68 cDNA and a HeLa cDNA bank as template yielded a DNA fragment that was subsequently identified as part of the human gene *ddx42*. The cDNA encodes a 938 amino acid polypeptide, which is strictly conserved among vertebrates (more than 70% homology, i.e. identical residues plus conservative substitutions, of the full-length sequence), and invertebrates also contain a highly related hypothetical protein. Ddx42p contains all the

classical core motifs of a DEAD box protein. In addition, its homology to both p68 and p72/p82 is significantly above average. While the N- and C-terminal regions of the proteins differ, their core regions contain >80% homologous amino acid residues (43.6% identity, 37.4% similarity), in contrast to eIF4A, which has <70% homology (27.2% identity, 41.6% similarity) (Figure 1).

### Analysis of the *ddx42* mRNA

The mRNAs of the p68 protein subfamily contain unusual features (37,40). Therefore, owing to the apparent close relationship of Ddx42p to the p68 subfamily, we first characterized the *ddx42* RNA. Northern blotting with a *ddx42*-specific probe revealed an mRNA of ~4 kb, which occurs in similar concentration in HeLa, Hep3B, COS, TC7 and HaCaT cells (Figure 2B). The coding sequence (cds) of the *ddx42* mRNA was found by RT-PCR to be extended further 5'-ward than reported previously (41,49), resulting in a polypeptide consisting of 938 amino acids (compare Figure 2A).

The untranslated regions (UTRs) of the *ddx42* mRNA were characterized by RT-PCR analyses (Figure 2A) and comparison of the *ddx42* genomic sequence with *ddx42* partial sequence clones available in the GenBank database. This approach confirmed the existence of two *ddx42* poly(A)<sup>+</sup> RNA species (GenBank NM\_007372 and NM\_203499, respectively), which differ in their 5'-UTRs (overall length 255 and 184 bases, respectively) with the smaller one being the prominent product. Evidently, the two RNA species arise



**Figure 2.** Characterization of *ddx42* mRNA and protein. (A) Schematic representation of the structure of the *ddx42* mRNA. The *ddx42* cds is represented as a bar starting with ATG and ending with TAG. ATG<sub>n</sub> denotes the true translation initiation codon of Ddx42p (this study), while ATG<sub>o</sub> is the previously reported one, and gray parts are genomic sequences not found in *ddx42* poly(A)<sup>+</sup> RNA (i1, i2, i3 are introns 1, 2 and 3). Exon 2 (e2) is absent in the shorter *ddx42* poly(A)<sup>+</sup> RNA. PCR primers used for analysis are indicated as arrows, and those not leading to a PCR product are highlighted in gray. (B) *ddx42* mRNA is present in similar low levels in different cell lines. A northern blot of poly(A)<sup>+</sup> RNA (100 ng each) was probed with a *ddx42*-specific antisense RNA probe.  $\beta$ -Actin RNA served as a loading control. The size of RNA marker bands (Fermentas RNA ladder, high range) is indicated on the right. (C) The protein Ddx42p is expressed differentially in different cell lines. Ddx42p was detected by western blotting in whole cell extracts (15  $\mu$ g total protein each) with antiserum  $\alpha$ -D42N. Actin was detected as a loading control. The molecular weight of marker bands (RPN800, Amersham Biosciences) is indicated on the right. (D) Ddx42p is found exclusively in nuclear extracts. The western blot was performed as above using nuclear extracts (ne) from HeLa and COS cells (lanes 1 and 2). Soluble cellular extracts (ce; lanes 3 and 5) and ne (lanes 4 and 6) from COS cells transfected with full-length *ddx42* cDNA (pCddx42; lanes 5 and 6) or a C-terminally shortened form (pCddx42 $\Delta$ ; lanes 3 and 4), respectively, show that even overexpressed protein resides in the nuclei. The overexpressed full-length form is probably subject to proteolytic degradation.

from alternative splicing as the shorter variant lacks a sequence corresponding exactly to exon 2. Both variants, however, encode the full-length Ddx42p polypeptide. The GC content of both 5'-UTRs amounts to  $\sim$ 70%, which hints at throttled translation of *ddx42*.

The *ddx42* 3'-UTR was found to almost extend to the second poly(A) site (Figure 2A) given in the database (length 918 bases). The 3'-UTR corresponds to the continuous genomic sequence and its GC content is very low (35%). In addition, it shows 28 stretches of 3 or more U in a row, particularly towards the extreme 3' end, and an AU-element

(5'-AUUUA-3'). All these features are generally deemed RNA destabilizing.

#### Cellular context of Ddx42p

Cell extracts were prepared from HeLa, Hep3B, COS, TC7 and HaCaT cells and tested by western blotting with Ddx42p-specific antibodies raised against oligopeptides from the N-terminal ( $\alpha$ -D42N) and C-terminal regions of the protein ( $\alpha$ -D42C). HeLa, Hep3B and COS cells were found to contain at least 2- to 3-fold more Ddx42p than TC7 and HaCaT cells

(Figure 2C). Transient overexpression of Ddx42p from a mammalian expression vector (pCMV5) in COS cells did not cause any discernable phenotype.

Cellular localization of DEAD box proteins differs, e.g. p68 and p72/p82 are localized primarily to the nucleus, while eIF4A occurs in the cytoplasm. In order to obtain information on the cellular distribution of Ddx42p, western blot analysis of subcellular fractions was performed ( $\alpha$ -D42N and  $\alpha$ -D42C are not suited to immunofluorescence studies). The endogenous protein as well as Ddx42p overexpressed in COS cells were found exclusively in the nucleus. A Ddx42p deletion mutant lacking the C-terminal part of the protein, which contains a potential nuclear localization sequence (NLS; PPKRKKSR), could also be detected only in the nucleus (Figure 2D), which points to an internal NLS. Nuclear localization of Ddx42p corresponds to former results (41). Interestingly, p68 and p72 have also been found as spliceosome-associated proteins (31), and have been shown to form complexes with each other (50) and with a number of nuclear proteins, among them was p53 (51). As a starting point to identify protein–protein interactions of Ddx42p, we checked whether the protein binds to p68 or to p53. Preliminary immunoprecipitation studies with monoclonal antibodies C10 (against p68) and pab421 (against p53), however, do not indicate complex formation of Ddx42p with p68 or p53 (data not shown).

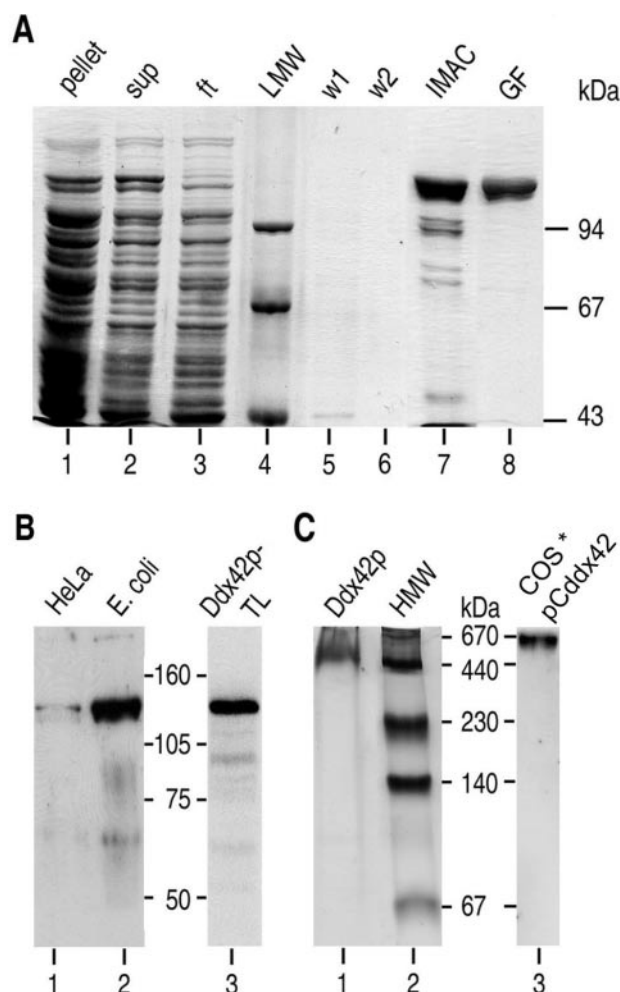
### Properties of Ddx42p

In order to obtain sufficient amounts of purified Ddx42p for further studies, we overexpressed the *ddx42* cds in *E. coli* with a hexa-His tag fused to its N-terminal end. Most of the Ddx42p protein was found in the soluble fraction after cell lysis and purified to apparent homogeneity by a two-step chromatography protocol (Figure 3A). Having a theoretical molecular mass of 104 kDa, recombinant and native Ddx42p appeared as a band of  $\sim$ 120 kDa in SDS–PAGE (Figure 3A and B), and no change in the electrophoretic behavior of Ddx42p was detected after the protein had been incubated at 37°C for several hours or stored at 4°C for up to 4 weeks. Native PAGE of the purified protein revealed high molecular weight complexes of Ddx42p of  $>$ 440 kDa (Figure 3C), which appear to be stable in the presence of ATP or ADP (not shown). This may suggest oligomer formation of the protein *in vivo*. Native PAGE of nuclear or cellular extracts followed by western blotting with  $\alpha$ -D42C revealed the incorporation of Ddx42p in complexes of even higher molecular weight, and no free monomer was detected even after overexpression (Figure 3C). This indicates that *in vivo*, the protein is indeed part of high molecular weight complexes, which may also contain RNA, however.

DEAD box proteins are RNA binding proteins, and therefore Ddx42p was examined by an electrophoretic mobility shift assay for its ability to bind to  $^{32}$ P-labeled RNA. Ddx42p was shown to bind to *in vitro* transcripts and partially dsRNAs (Figure 4A), but no band shift was observed for completely dsRNA (data not shown), indicating that Ddx42p binds to ss regions.

### RNA binding affinity of Ddx42p

Binding of Ddx42p to RNA was further characterized by a filter binding assay. The apparent  $K_D$  of Ddx42p for binding to an *in vitro* transcript varies only slightly in the absence

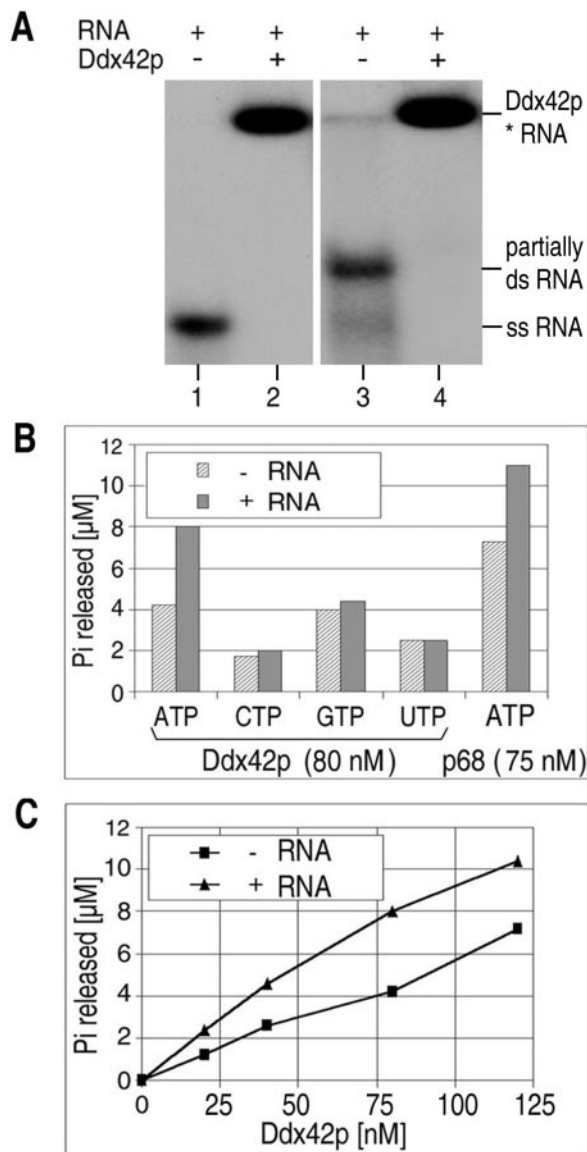


**Figure 3.** Purification and characterization of recombinant Ddx42p. (A) Ddx42p is purified to apparent homogeneity after expression in *E. coli*. A Coomassie-stained SDS–PAGE of the purification course is shown. sup, soluble fraction after sonication of *E. coli*; ft, IMAC flow-through; LMW, low molecular weight calibration kit (Amersham Biosciences); w1 (2), IMAC washes 1 (2); IMAC, Ddx42p-containing IMAC eluate; GF, Ddx42p eluate of gel filtration. Marker band sizes are indicated on the right. (B) Ddx42p is heterologously expressed and *in vitro* translated as a full-length protein. Lanes 1 and 2 show a western blot with antisera  $\alpha$ -D42C of a HeLa nuclear extract and Ddx42p expressed in *E. coli*. Lane 3 is an autoradiograph of  $^{35}$ S-labeled Ddx42p *in vitro* transcribed/translated *ddx42* cDNA. The molecular weight of the protein marker bands is indicated. (C) Ddx42p forms complexes of high molecular weight. Purified Ddx42p was subjected to native PAGE and silver staining (lane 1). The molecular weight of the high molecular weight calibration kit proteins (Amersham Biosciences, lane 2) is indicated. Lane 3 shows a nuclear extract of COS cells transfected with full-length *ddx42* cDNA (pCddx42) separated on a native PAGE and detected by western blotting with  $\alpha$ -D42C.

(100 nM) and in the presence of ATP or ATP $\gamma$ S (75 nM). These values are in the same range as those of T4gp32 [(52) and data not shown] and other DExH/D proteins like Ded1p (53) or Prp22 (54). However, in the presence of ADP, the apparent  $K_D$  drops to 10 nM, indicating an increased affinity of ADP-bound Ddx42p for ssRNA.

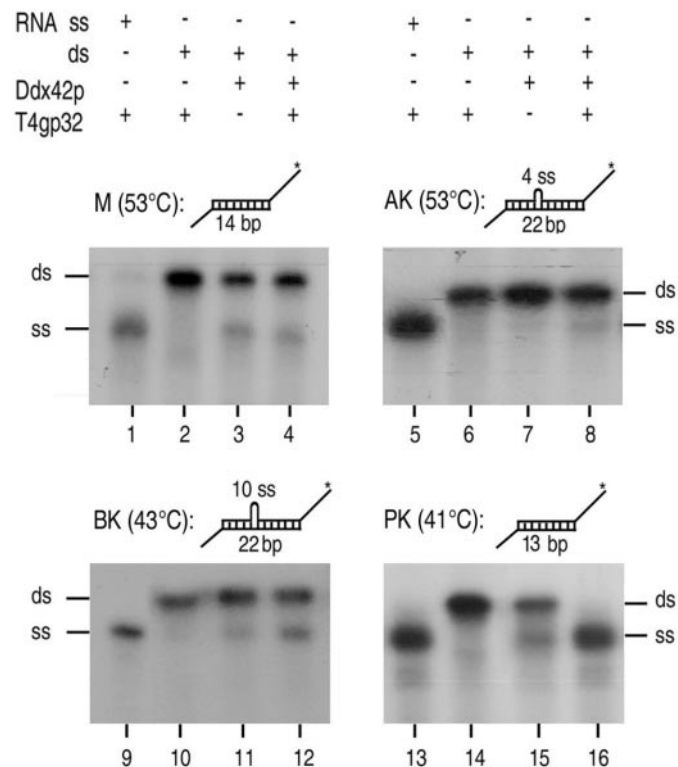
### Hydrolysis of NTPs by Ddx42p

While helicases depend on their NTPase activity to resolve the hydrogen bonds of ds nucleic acids and/or to move along a



**Figure 4.** Ddx42p is an RNA binding protein and RNA-stimulated ATPase. (A) RNA binding by Ddx42p. Purified Ddx42p was incubated with an *in vitro* transcript (lane 2) and a partial RNA duplex (lane 4), respectively. Lower electrophoretic mobility proves binding of Ddx42p to both RNAs. (B) NTPase activity of Ddx42p. Ddx42p shows a basal NTPase activity with all four NTPs, which is lower than that of p68. Only the Ddx42p ATPase is stimulated in the presence of heteropolymeric RNA. (C) Ddx42p ATPase as a function of Ddx42p concentration. Formation of orthophosphate owing to ATP hydrolysis was monitored at increasing Ddx42p concentrations in the absence and in the presence of poly(A)<sup>+</sup> RNA. Values depicted represent the average of at least three measurements per point.

single nucleic acid strand, their specificities for a particular NTP and the extent of NTPase stimulation by nucleic acids vary considerably. Therefore, we analyzed the hydrolysis of the four NTPs by Ddx42p in the absence and in the presence of nucleic acid. Basal hydrolysis activity of Ddx42p was found for all four NTPs. However, only the ATPase activity of Ddx42p could be stimulated by addition of RNA (Figure 4B and C). Interestingly, RNA homopolymers poly(C) and poly(A) as well as completely dsRNA did not stimulate the ATPase (data not shown), while poly(A)<sup>+</sup> RNA (Figure 4B



**Figure 5.** Ddx42p is a non-processive RNA helicase stimulated in the presence of a ss binding protein. RNA helicase assay samples were separated by SDS-PAGE and autoradiographed. M, AK, BK and PK denote the partially dsRNA substrates depicted above the respective autoradiograph. An asterisk marks the respective <sup>32</sup>P-labeled RNA strand. In lane 1 of each panel, the expected product of the helicase reaction, the released ss, is shown. The ss binding protein, T4gp32, does not unwind any of the dsRNA substrates used (lanes 2, 6, 10 and 14). Limited unwinding of substrates M and AK by Ddx42p was observed (lanes 3, 4, 7 and 8). For the less stable substrates BK and PK, respectively, the weak Ddx42p-mediated strand-separating activity (lanes 11 and 15, respectively) is promoted by T4gp32 (lanes 12 and 16, respectively). The positions of the respective dsRNA and ssRNA species are indicated.

and C), total RNA, tRNA or *in vitro* transcripts (not shown) all stimulate the Ddx42p ATPase by ~1.5- to 2-fold. Addition of RNase abolished this stimulation (data not shown). Kinetic analysis revealed an apparent  $K_M$  for ATP of 0.2  $\mu$ M and a turnover number of 2  $\text{min}^{-1}$  ( $R = 0.9998$ , standard error = 0.07). ATP hydrolysis by Ddx42p is comparable with that of p68 (Figure 4B).

### RNA helicase activity of Ddx42p

The high degree of sequence homology to p68, RNA binding and RNA-stimulated ATPase by Ddx42p suggested that the protein might function as an RNA helicase as well. The members of the p68 subfamily were shown to separate partially dsRNA substrates up to 25 bp (corresponding to a  $T_m$  of ~65°C as a measure of duplex stability) with 3' ss overhangs into the respective single strands (28,37). Accordingly, Ddx42p was offered p68 helicase substrates with 3' ss overhangs (M and K, respectively). Ddx42p unwound none of substrate K ( $T_m = 62^\circ\text{C}$ ; not shown) and about 10% of substrate M ( $T_m = 53^\circ\text{C}$ ; Figure 5, lane 3) in the presence of ATP under optimized reaction conditions (5 nM Ddx42p, final concentration of NaCl 50 mM, 30 min reaction time).

Increasing concentrations of Ddx42p gradually reduced the fraction of unwound substrate M, and addition of poly(C) as a competitor had no effect (data not shown). The assay, however, allows the detection of only completely unwound strands, as intermediates in which part of the duplex is still intact when the protein dissociates would snap back into the original ds substrate.

Indeed, some DEAD box proteins, like eIF4A, are non-processive RNA helicases, which can unwind 10–15 bp, depending on duplex stability, at most. In order to check if Ddx42p mediates strand separation of less stable duplexes, substrate PK ( $T_m = 41^\circ\text{C}$ ), and two substrates in which the duplex is interrupted by a loop of 4 (AK,  $T_m = 53^\circ\text{C}$ ) and 10 bases (BK,  $T_m = 43^\circ\text{C}$ ), respectively, were created. Substrate AK was not unwound, while about 15% of BK and about 20% of PK could be separated into their respective single strands by Ddx42p (Figure 5, lanes 7, 11 and 15), which therefore seems to represent a non-processive RNA helicase. A substrate with 5' ss overhangs (substrate 5,  $T_m = 45^\circ\text{C}$ ), on the other hand, was not unwound by Ddx42p (data not shown).

*In vivo*, RNA helicase activities can be modulated due to complex formation with other cellular factors. Such functional cooperations are known from spliceosomal complexes (2) or from the translation initiation complex, in which the RNA helicase activity of eIF4A is modulated. However, the eIF4A RNA helicase is also stimulated in an *in vitro* assay containing only isolated eIF4A and eIF4B (22), which do not seem to form a protein–protein complex (55). eIF4B binds not only to ss, but also to structured RNA, and the precise mechanism by which it stimulates the eIF4A helicase is unknown (22). However, classical *in vitro* RNA helicase assays as those described do not account for two important facts: (i) ssRNA occurs almost exclusively covered by proteins *in vivo* and (ii) ssRNA produced in an RNA helicase reaction is exposed to other RNA binding proteins, e.g. for subsequent reactions, or may be removed by protein-mediated transport processes. In order to mimic these effects in the *in vitro* experiments with Ddx42p, we wanted to check whether a ssRNA binding protein has an effect on the Ddx42p helicase. Thus we performed a helicase assay in the presence of T4gp32, the ss binding protein of bacteriophage T4 (56), as a model protein, which was unlikely to affect the intrinsic properties of Ddx42p. T4gp32 is capable of non-base-specific binding to ssDNA or ssRNA owing to its oligonucleotide/oligosaccharide binding domain (OB fold) (57). Excess T4gp32 was allowed to bind to RNA prior to addition of Ddx42p. Under the conditions of the helicase assay, T4gp32 binds RNA with an apparent  $K_D$  of about 100 nM [(52) and data not shown]. None of the helicase substrates used was separated into single strands by T4gp32 alone (control reactions lacking Ddx42p; Figure 5, lanes 2, 6, 10 and 14). Almost no effect on the Ddx42p unwinding activity was seen for substrates M and AK (Figure 5, lanes 4 and 8), indicating that T4gp32 did not affect the Ddx42p processivity. On the other hand, the fraction of unwound BK was increased, and PK was unwound completely under these conditions (Figure 5, lanes 12 and 16, respectively). This shows that T4gp32 assists in Ddx42p helicase action most likely by stabilizing newly generated single strands as neither Ddx42p ATPase activity nor Ddx42p RNA binding appeared to be affected by T4gp32 (not shown).

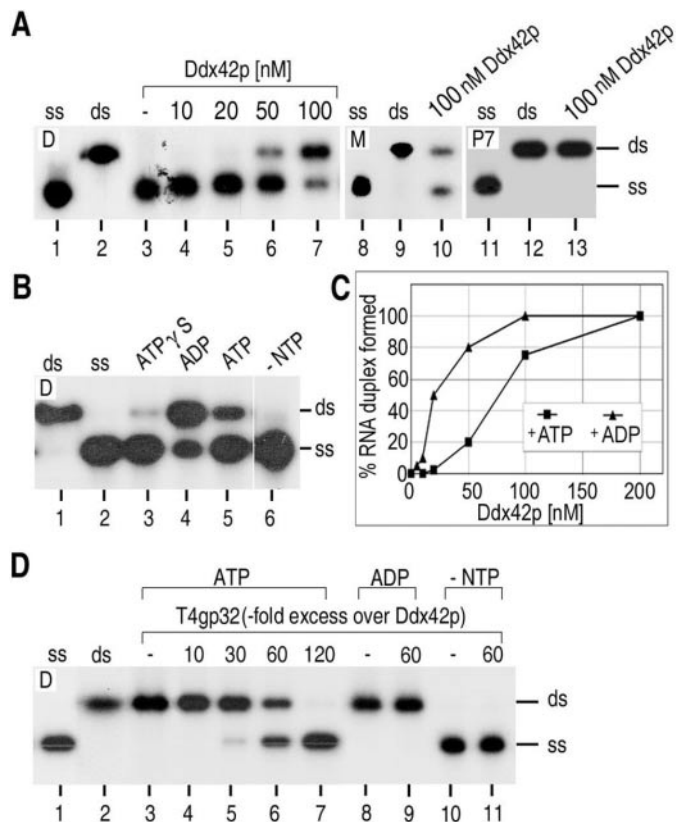
### Annealing of complementary RNA sequences by Ddx42p

We observed that higher concentrations of Ddx42p in the helicase assays did in fact decrease the fraction of unwound molecules. This led us to the presumption that the protein might also mediate annealing of complementary RNA regions as had been shown previously for p68 and p72 (28). Indeed, when partially complementary single strands (complementary regions of 14–123 bases) were offered to Ddx42p instead of a helicase substrate, the protein produced some ds molecules under helicase assay conditions, thus counteracting the helicase activity. While essentially no annealing was seen in the absence of Ddx42p, raising the protein concentration increased the yield of annealed RNA strands (Figure 6A, lanes 4–7). The reaction reached equilibrium after 10 min, and the efficiency of the annealing reaction correlates with the length, and thus with the stability, of the RNA duplex formed (Figure 6A, lanes 7, 10 and 13).

Protein-catalyzed RNA annealing activity proceeds independently of NTP for p68, p72 (28) and p53 (58), and DNA annealing by RecQ5 $\beta$  is more pronounced in the absence of NTP (59). Thus, the role for Ddx42p-mediated RNA annealing of the nucleotide cofactor was studied (Figure 6B). About 20% of the strands were annealed when ATP had been added to the reaction. Nearly no duplex was formed upon omission of a nucleotide cofactor or in the presence of ATP $\gamma$ S, a non-hydrolyzable analog of ATP. Addition of ADP as a model for the state of Ddx42p after ATP hydrolysis, on the other hand, led to significantly more (~80%) annealed product. Also, lower concentrations of Ddx42p were sufficient for duplex formation in the presence of ADP as compared with ATP (Figure 6C), indicating that Ddx42p after ATP hydrolysis favors local enrichment of complementary single strands.

In order to determine whether covering of ssRNA strands with a ss binding protein affects Ddx42p-mediated RNA duplex formation, we performed the RNA annealing assay with ssRNA preincubated with increasing amounts of T4gp32 (molar excess over RNA in nucleotides at least 20-fold). The binding site of T4gp32, taken as a model ss binding protein here, was determined to be 6–10 nt long (60), which is similar to that of, e.g. the RNP exon junction complex [8–10 nt; (8)]. No annealing was observed in control reactions lacking Ddx42p. T4gp32 started to significantly interfere with the annealing reaction from a >50-fold molar excess of T4gp32 over Ddx42p in the reaction containing ATP (Figure 6D, lanes 3–7). In contrast, essentially no inhibition of annealing by ADP-bound Ddx42p was observed under these conditions (lanes 8 and 9). As expected from the above experiments, RNA persisted in the ss state upon omission of the nucleotide (lanes 10 and 11). These results show that Ddx42p mediates RNA annealing even when the RNA has been covered by an RNA binding protein. Complex formation between T4gp32 and Ddx42p as a theoretically possible, though unlikely mechanism for removal of T4gp32 from the RNA was not detected (data not shown). Since binding of T4gp32 to RNA is ss specific, the Ddx42p-catalyzed annealing reaction shows that T4gp32 is efficiently displaced from the RNA.





**Figure 6.** RNA duplex formation by Ddx42p. (A) Ddx42p mediates annealing of complementary regions in RNA. RNA strands containing complementary regions of differing length were used as substrates for Ddx42p-catalyzed annealing reactions. The ss and ds control samples were incubated under identical conditions, but in the absence of Ddx42p. RNA annealing samples were subjected to SDS-PAGE and autoradiographed. D, M and P7 denote the partially ds product of the respective annealing reaction. The efficiency of RNA annealing by Ddx42p under helicase conditions increases with increasing length of the base pairing region: M (14 bp) < D (46 bp) < P7 (123 bp). The positions of the respective dsRNA and ssRNA species are indicated on the right. (B) Ddx42p-mediated RNA annealing is most efficient with ADP as the nucleotide cofactor. Annealing reactions with 50 nM Ddx42p were carried out with ssRNAs DT7 and DSP6 (complementary region of 46 bases) to exclude a possible counteracting of Ddx42p helicase activity. The reaction is most efficient with ADP, indicating that ATP must be hydrolyzed before RNA annealing. The weak RNA duplex band in lane 3 is attributed to 10% ADP content in the ATP $\gamma$ S preparation as declared by the manufacturer. (C) RNA annealing as a function of Ddx42p concentration. Formation of duplex D (46 bp) was monitored at increasing Ddx42p concentrations. Each point is the average of at least three measurements. In the presence of ADP (after ATP hydrolysis), significant annealing activity was observed even at a Ddx42p concentration used in helicase assays. (D) Ddx42p removes bound protein from RNA. Ddx42p-mediated RNA annealing reactions were carried out with the indicated nucleotide cofactors in the presence of RNA coated with saturating concentrations of the ss binding protein T4gp32. T4gp32 affects the Ddx42p-mediated RNA annealing reaction only when a 50-fold molar excess of T4gp32 over Ddx42p is used in the presence of ATP (lane 6). ADP-bound Ddx42p is even more effective as essentially no inhibition of annealing is seen under these conditions (lane 9).

## DISCUSSION

### ddx42 gene and mRNA structure imply regulated gene expression

The genomic sequence of *ddx42* extends over 45 kb. The occurrence of two introns in the *ddx42* 5'-UTR sets it apart from the intronless 5'-UTRs of the p68 and p72/p82 mRNAs

(37,40), but the UTRs of all three mRNAs rank with long eukaryotic ones. Additionally, sequence composition (high GC content of 5'-UTR, Us in 3'-UTR) and lack of a canonical poly(A) signal also point to translational regulation of *ddx42*. This implication is supported by the finding that in spite of a consistently low *ddx42* mRNA level in the different cell lines analyzed (Figure 2B), the amount of Ddx42p differs from cell line to cell line (Figure 2C). Remarkably, immortalized cells (HaCaT and TC7) show a lower Ddx42p content than transformed ones (HeLa, Hep3B, COS).

### Characteristics of Ddx42p

Ddx42p, according to its sequence, is highly related to the p68 subfamily of DEAD box proteins. The known members of this group of proteins have been found in the nucleus as is also shown here for Ddx42p. A predicted NLS in the C-terminal part of the protein (PPKRRKKS), which is similar to the functional NLSs of protein kinase A (PPKRRKV) and of SV40 T antigen (PKKRRKV) (61), does not or not exclusively determine nuclear localization of Ddx42p. Consequently, other sequence features at least contribute to its nuclear import. Sequences predicted (<http://cubic.bioc.columbia.edu/predictNLS>) to determine nuclear localization of p68 (KKFGNPGEKLVK) and of p72/p82 (RGGGGGGGG), respectively, are not present in Ddx42p, however. On the other hand, not all nuclear proteins are assigned a defined NLS and novel types of NLSs continue to be reported (62,63). Moreover, the sequence of Ddx42p contains stretches consisting of positively charged and apolar amino acids that might be recognized as a nuclear import signal.

Based on the high degree of sequence homology of Ddx42p to the p68 subfamily proteins and one common homolog in yeast, Dbp2p, one could assume similar functional contexts, and that Ddx42p may interact with the same protein partners as p68 and p72. Yet, we did not observe binding of Ddx42p to either p68 or p53, which form complexes with p68 and/or p72 (50,51). Characterization of Ddx42p interaction partners will be particularly interesting future work as Ddx42p probably has functions that differ from p68/p72. In addition, the biochemical similarity of Ddx42p and the p68 group is limited since some activities of Ddx42p resemble more those of non-processive helicases like eIF4A or Ded1p.

### RNA binding by Ddx42p is modulated by the adenosine nucleotide

Like most DExH/D proteins, Ddx42p requires a ss region for binding to RNA, as shown by its inability to bind to completely dsRNA. The ssRNA binding affinity of Ddx42p in the absence and in the presence of ATP or ATP $\gamma$ S is similar. However, ADP-bound Ddx42p shows an almost 10-fold increased affinity for ssRNA. The yeast protein Ded1p shows a similar behavior: (i) it does not bind to a complete RNA duplex, (ii) it shows an increased affinity to partial duplexes in the presence of ATP, but not ADP, and (iii) hydrolysis of ATP increases the affinity of Ded1p to ssRNA (53). This suggests that ATP hydrolysis modulates the relative affinities of Ddx42p for ssRNA versus dsRNA in a similar way, probably based on conformational changes of the protein as was shown for eIF4A (64).

### Ddx42p ATPase is stimulated by heteropolymeric RNA only

Though NTP binding and hydrolysis by DExH/D proteins is substantially based on Walker A and B motifs, usage of NTPs differs considerably. Even the members of the p68 subfamily differ in this respect: p68, with ATP being its best substrate, also harnesses CTP or UTP for RNA unwinding whereas p72 and p82 are strictly ATP-dependent (28,37). Ddx42p is an RNA binding protein and shows basal hydrolysis of all four NTPs, but only ATP hydrolysis by Ddx42p is enhanced by RNA binding. The 2-fold stimulation of the Ddx42p ATPase by RNA seen here is similar to that of other DExH/D proteins like p68 (Figure 4B) or spliceosomal helicases Prp2 (65) and UAP56 (66) under comparable conditions. The apparent  $K_M$  for ATP of 0.2  $\mu\text{M}$  indicates that Ddx42p binds ATP more tightly than other DEAD box proteins, and the turnover number for ATP hydrolysis ( $\sim 2 \text{ min}^{-1}$ ) is in the same range as that of eIF4A (67). ATP hydrolysis by Ddx42p is significantly stimulated in the presence of various heteropolymeric RNAs, but not with poly(A) and poly(C), indicating weak or no binding of Ddx42p to these RNAs. Correspondingly, poly(C) was no competitor for Ddx42p helicase. Little stimulation of ATPase activity by RNA homopolymers was shown previously, e.g. for Ded1p (53). It may be due to inherent low affinity of some DEAD box proteins to these non-natural RNAs, which adopt conformations different from RNA heteropolymers. Also, complete RNA duplexes do not stimulate Ddx42p ATPase, indicating that Ddx42p in ATP-bound state needs a ss region or ss/ds junction for binding to its RNA substrate. The lower ssRNA binding affinity of ATP-bound relative to ADP-bound Ddx42p (see above) is consistent with a preference of ATP-bound Ddx42p for ss/ds junctions.

### Ddx42p RNA helicase is stimulated in the presence of a single-strand binding protein

Testing of several RNA helicase substrates of different stability has revealed that Ddx42p is a non-processive RNA helicase in contrast to p68 and p72/p82 which can be regarded as low processive RNA helicases (28). This result corresponds well with the higher affinity of Ddx42p for ATP and lower turnover number. Single-stranded loops up to 10 bases in one strand of the helicase substrate are not accessible to Ddx42p, because it could be expected that substrate BK was unwound with higher efficiency then. The absence of processivity is characteristic for a number of RNA helicases, e.g. for eIF4A, Ded1p [reviewed in (68)], Prp16 and Prp22 [reviewed in (2)]. In most pathways that include RNA restructuring, processive separation of base pairs is not only unnecessary but could even lead to uncoordinated release of single strands which might then be available for a physiologically wrong interaction partner.

The non-processive Ddx42p helicase activity is promoted in the presence of a ss binding protein (T4gp32) with respect to the fraction of unwound substrate, but not the length of the ds region, i.e. processivity is not influenced. Absence of competition for the nucleic acid substrate and a stimulatory effect on DNA unwinding by SV40 T ag was found previously in the presence of excess SSB, leading to the conclusion that T antigen helicase preferentially binds to ss/ds junctions (69). This may also hold true for Ddx42p (see above).

Protein-mediated stabilization of nascent ss RNA during rearrangement of a protein–nucleic acid complex avoids wrong interactions and cooperation between helicases and ssRNA binding proteins is indeed found in spliceosomal rearrangement reactions (2). Thus, a scenario of Ddx42p helicase assisted by another ssRNA binding protein is not unlikely *in vivo*.

### Ddx42p efficiently anneals complementary strands of RNA

Ddx42p is also capable of mediating RNA annealing (which may counteract its helicase activity such that net unwinding is low), and the annealed region appears not to be limited to a certain length. In contrast to other helicases with annealing activity [p68, p72 (28) or RecQ5 $\beta$  (59)], Ddx42p RNA annealing depends on the presence of a nucleotide cofactor. Maximum Ddx42p annealing activity (even at Ddx42p concentrations where the protein is helicase active) is seen in the presence of ADP, i.e. the state of Ddx42p after ATP hydrolysis, so the ADP-bound conformation of Ddx42p apparently has the highest affinity to ssRNA. This is confirmed by RNA binding studies (see above). In the presence of ATP, annealing is observed starting at higher concentrations of Ddx42p. This suggests that ATPase activity is required to convert part of Ddx42p into the ADP-bound state, which then mediates annealing. Correspondingly, addition of a non-hydrolyzable ATP analog or omission of nucleotide abolishes RNA annealing. Taken together, the activities of Ddx42p appear accurately tuned, which makes sense for the coordination of helicase and annealing activities under physiological conditions.

Whereas the helicase mechanism is beginning to be understood, far less is known on the mechanism of protein-catalyzed intermolecular or intramolecular RNA annealing. Base pairing requires a local enrichment of complementary single strands, which is effected by Ddx42p possibly by the formation of oligomers (Figure 3C). Ddx42p apparently binds to RNA in a sequence-unspecific manner like most RNA helicases, i.e. it binds to the acidic sugar-phosphate backbone. A clue of how Ddx42p may promote strand annealing is provided by the RNA annealing protein gBP21 (70). (i) Clusters of basic amino acids, particularly RG-rich domains, assist in neutralization of RNA backbone charges. Clustering of basic residues is found in the C-terminal part of Ddx42p. (ii) ADP-bound Ddx42p has a high affinity to ssRNA and traps a ss until it encounters a complementary region. (iii) Base pairing leads to a conformational change in RNA, the affinity of Ddx42p to dsRNA is low, and the protein dissociates (possibly triggered by base pairing energy).

### Ddx42p disrupts protein–RNA interactions

Ddx42p displays RNA helicase as well as annealing activities not only on free but also on RNA covered by an excess of a ssRNA binding protein. At least the latter process requires removal of the pre-bound protein from RNA. Annealing of T4gp32-precoated ssRNA includes Ddx42p-mediated T4gp32 removal and duplex formation. ADP-bound Ddx42p, as in the assay with free RNA, was most efficient in this reaction. Competition by excess T4gp32 in the presence of ATP indicates impaired annealing due to shielding of ssRNA. This result

underscores the importance of ATP hydrolysis by Ddx42p for protein removal and strand annealing. The driving force for the reaction is most probably a conformational change in Ddx42p leading to increased affinity of the protein for ssRNA.

Taken together, these observations allow us to speculate on a possible mechanism by which a non-processive RNA helicase uses ATP hydrolysis to affect conformational changes in RNP assemblies: ATP hydrolysis leads to a conformational change in Ddx42p, thus enabling the protein by its increased affinity to ssRNA to remove a nearby bound obstacle (i.e. base pairs of a short ds duplex and/or protein). This process might be accompanied by conformational changes on the RNA strand, exposing it to a suitable binding partner, e.g. a protein or another complementary RNA. Ddx42p, perhaps as an oligomer, provides several RNA binding sites, thus raising the probability of accumulating complementary single strands in spatial proximity. Coordination of non-processive helicase, protein displacement and annealing activities residing in Ddx42p, without the need for the simultaneous presence of another protein, is an ideal prerequisite for its involvement in RNP restructuring, for instance in the remodeling of the spliceosomal complex SF3b as has been hypothesized recently (71). Upon U2 snRNP formation, SF3b interacts with the 12S U2 particle to constitute the intermediate 15S U2 complex, which in turn binds SF3a to build the mature 17S U2 snRNP. The complexes are highly flexible to allow for the necessary conformational changes. Ddx42p is apparently associated with splicing factor SF3b, but not with the 17S U2 snRNP itself (41). A possible role of Ddx42p (as a temporary component of SF3b) could thus include a remodeling of a protein–RNA interaction or a base pair contact (i) at the 5' end of the U2 snRNA where SF3b binds, or (ii) after 15S U2 formation, perhaps at stem–loops IIa/IIb, to facilitate SF3a joining. On the other hand, SF3b is also part of the U11/U12 di-snRNP of the minor spliceosome. It splices U12-type introns, which have not been found in yeast (72). The fact that no obvious homolog of Ddx42p exists in yeast raises the possibility of a specific involvement of Ddx42p in U11/U12 di-snRNP formation. Furthermore, Ddx42p may be involved in other spliceosomal rearrangements as it binds not only to U2, but also to U4 and U5 snRNAs (41), and last but not least, it might participate in the recycling of spliceosome components.

Finally, the results presented in this work provide evidence for the assumption that ATP hydrolysis-driven protein displacement from RNA is indeed a general feature of DExH/D proteins.

## SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

## ACKNOWLEDGEMENTS

We thank Anna-Maria Bohrer for excellent technical assistance and W. Nastainczyk for peptide synthesis and immunization of rabbits. HaCaT cells were kindly provided by M. Montenarh (Universität des Saarlandes, Homburg, Germany). We wish to thank N. Harraghy for valuable comments on the manuscript. This work was supported by a grant of the Saarland Ministry of Education, Culture and Science to H.U.-S. Funding to pay the

Open Access publication charges for this article was provided by the University of the Saarland.

*Conflict of interest statement.* None declared.

## REFERENCES

- Shiratori,A., Shibata,T., Arisawa,M., Hanaoka,F., Marakami,Y. and Eki,T. (1999) Systematic identification, classification, and characterization of the open reading frames which encode novel helicase-related proteins in *Saccharomyces cerevisiae* by gene disruption and northern analysis. *Yeast*, **15**, 219–253.
- Delagoutte,E. and von Hippel,P.H. (2003) Helicase mechanisms and the coupling of helicases within macromolecular machines. Part II: integration of helicases into cellular processes. *Q. Rev. Biophys.*, **36**, 1–69.
- Hickson,I.D. (2003) RecQ helicases: caretakers of the genome. *Nature Rev. Cancer*, **3**, 169–178.
- Abdelhaleem,M. (2004) Do human RNA helicases have a role in cancer? *Biochim. Biophys. Acta*, **1704**, 37–46.
- Kwong,A.D., Rao,B.G. and Jeang,K.T. (2005) Viral and cellular RNA helicases as antiviral targets. *Nature Rev. Drug Discov.*, **4**, 845–853.
- Byrd,A.K. and Raney,K.D. (2004) Protein displacement by an assembly of helicase molecules aligned along single-stranded DNA. *Nature Struct. Mol. Biol.*, **11**, 531–538.
- Jankowsky,E., Gross,C.H., Shuman,S. and Pyle,A.M. (2001) Active disruption of an RNA-protein interaction by a DExH/D RNA helicase. *Science*, **291**, 121–125.
- Fairman,M.E., Maroney,P.A., Wang,W., Bowers,H.A., Gollnick,P., Nilsen,T.W. and Jankowsky,E. (2004) Protein displacement by DExH/D 'RNA helicases' without duplex unwinding. *Science*, **304**, 730–734.
- Chen,J.Y.-F., Stands,L., Staley,J.P., Jackups,R.R., Latus,L.J. and Chang,T.-H. (2001) Specific alterations of U1-C protein or U1 small nuclear RNA can eliminate the requirement of Prp28p, an essential DEAD box splicing factor. *Mol. Cell*, **7**, 227–232.
- Kistler,A.L. and Guthrie,C. (2001) Deletion of MUD2, the yeast homolog of U2AF65, can bypass the requirement for sub2, an essential spliceosomal ATPase. *Genes Dev.*, **15**, 42–49.
- Gross,C.H. and Shuman,S. (1998) The nucleoside triphosphatase and helicase activities of vaccinia virus NPH-II are essential for virus replication. *J. Virol.*, **72**, 4729–4736.
- Yao,N., Hesson,T., Cable,M., Hong,Z., Kwong,A.D., Le,H.V. and Weber,P.C. (1997) Structure of the hepatitis C virus RNA helicase domain. *Nature Struct. Mol. Biol.*, **4**, 463–467.
- Bartenschlager,R. and Lohmann,V. (2000) Replication of hepatitis C virus. *J. Gen. Virol.*, **81**, 1631–1648.
- Pang,S.P., Jankowsky,E., Planet,P.J. and Pyle,A.M. (2002) The hepatitis C viral NS3 protein is a processive DNA helicase with cofactor enhanced RNA unwinding. *EMBO J.*, **21**, 1168–1176.
- Kikuma,T., Ohtsu,M., Utsugi,T., Koga,S., Okuhara,K., Eki,T., Fujimori,F. and Murakami,Y. (2004) Dbp9p, a member of the DEAD box protein family, exhibits DNA helicase activity. *J. Biol. Chem.*, **279**, 20692–20698.
- Gomez-Lorenzo,M.G., Valle,M., Frank,J., Gruss,C., Sorzano,C.O., Chen,X.S., Donate,L.E. and Carazo,J.M. (2003) Large T antigen on the simian virus 40 origin of replication: a 3D snapshot prior to DNA replication. *EMBO J.*, **22**, 6205–6213.
- Seinsoth,S., Uhlmann-Schiffler,H. and Stahl,H. (2003) Bidirectional DNA unwinding by a ternary complex of T antigen, nucleolin and topoisomerase I. *EMBO Rep.*, **4**, 263–268.
- Uhlmann-Schiffler,H., Seinsoth,S. and Stahl,H. (2002a) Preformed hexamers of SV40 T antigen are active in RNA and origin-DNA unwinding. *Nucleic Acids Res.*, **30**, 3192–3201.
- Rogers,G.W.Jr, Komar,A.A. and Merrick,W.C. (2002) eIF4A: the godfather of the DEAD box helicases. *Prog. Nucleic Acid Res. Mol. Biol.*, **72**, 307–331.
- Linder,P. (2004) The life of RNA with proteins. *Science*, **304**, 694–695.
- Rogers,G.W.Jr, Lima,W.F. and Merrick,W.C. (2001a) Further characterization of the helicase activity of eIF4A. Substrate specificity. *J. Biol. Chem.*, **276**, 12598–12608.
- Rogers,G.W.Jr, Richter,N.J., Lima,W.F. and Merrick,W.C. (2001b) Modulation of the helicase activity of eIF4A by eIF4B, eIF4H, and eIF4F. *J. Biol. Chem.*, **276**, 30914–30922.

23. Rocak, S. and Linder, P. (2004) DEAD-box proteins: the driving forces behind RNA metabolism. *Nature Rev. Mol. Cell Biol.*, **5**, 232–241.
24. de la Cruz, J., Iost, I., Kressler, D. and Linder, P. (1997) The p20 and Ded1 proteins have antagonistic roles in eIF4E-dependent translation in *Saccharomyces cerevisiae*. *Proc. Natl Acad. Sci. USA*, **94**, 5201–5206.
25. Stevens, S.W., Ryan, D.E., Ge, H.Y., Moore, R.E., Young, M.K., Lee, T.D. and Abelson, J. (2002) Composition and characterization of the yeast spliceosomal penta-snRNP. *Mol. Cell*, **9**, 31–44.
26. Hirling, H., Scheffner, M., Restle, T. and Stahl, H. (1989) RNA helicase activity associated with the human p68 protein. *Nature*, **339**, 562–564.
27. Lamm, G.M., Nicol, S.M., Fuller-Pace, F.V. and Lamond, A.I. (1996) p72: a human nuclear DEAD box protein highly related to p68. *Nucleic Acids Res.*, **24**, 3739–3747.
28. Rossler, O.G., Straka, A. and Stahl, H. (2001) Rearrangement of structured RNA via branch migration structures catalysed by the highly related DEAD-box proteins p68 and p72. *Nucleic Acids Res.*, **29**, 2088–2096.
29. Watanabe, M., Yanagisawa, J., Kitagawa, H., Takeyama, K., Ogawa, S., Arao, Y., Suzawa, M., Kobayashi, Y., Yano, T., Yoshikawa, H. *et al.* (2001) A subfamily of RNA-binding DEAD-box proteins acts as an estrogen receptor alpha coactivator through the N-terminal activation domain (AF-1) with an RNA coactivator, SRA. *EMBO J.*, **20**, 1341–1352.
30. Warner, D.R., Bhattacharjee, V., Yin, X., Singh, S., Mukhopadhyay, P., Pisano, M.M. and Greene, R.M. (2004) Functional interaction between Smad, CREB binding protein, and p68 RNA helicase. *Biochem. Biophys. Res. Commun.*, **324**, 70–76.
31. Zhou, Z., Licklider, L.J., Gygi, S.P. and Reed, R. (2002) Comprehensive proteomic analysis of the human spliceosome. *Nature*, **419**, 182–185.
32. Honig, A., Auboeuf, D., Parker, M.M., O'Malley, B.W. and Berget, S.M. (2002) Regulation of alternative splicing by the ATP-dependent DEAD-box RNA helicase p72. *Mol. Cell Biol.*, **22**, 5698–5707.
33. Guil, S., Gattoni, R., Carrascal, M., Abian, J., Stevenin, J. and Bach-Elias, M. (2003) Roles of hnRNP A1, SR proteins, and p68 helicase in c-H-ras alternative splicing regulation. *Mol. Cell Biol.*, **23**, 2927–2941.
34. Gendra, E., Moreno, A., Alba, M.M. and Pages, M. (2004) Interaction of the plant glycine-rich RNA-binding protein MA16 with a novel nuclear DEAD box RNA helicase protein from *Zea mays*. *Plant J.*, **38**, 875–886.
35. Jost, J.P., Schwarz, S., Hess, D., Angliker, H., Fuller-Pace, F.V., Stahl, H., Thiry, S. and Siegmund, M. (1999) A chicken embryo protein related to the mammalian DEAD box protein p68 is tightly associated with the highly purified protein–RNA complex of 5-MeC-DNA glycosylase. *Nucleic Acids Res.*, **27**, 3245–3252.
36. Goh, P.Y., Tan, Y.J., Lim, S.P., Tan, Y.H., Lim, S.G., Fuller-Pace, F. and Hong, W. (2004) Cellular RNA helicase p68 relocalization and interaction with the hepatitis C virus (HCV) NS5B protein and the potential role of p68 in HCV RNA replication. *J. Virol.*, **78**, 5288–5298.
37. Uhlmann-Schiffler, H., Rossler, O.G. and Stahl, H. (2002b) The mRNA of DEAD box protein p72 is alternatively translated into an 82-kDa RNA helicase. *J. Biol. Chem.*, **277**, 1066–1075.
38. Iggo, R.D., Jamieson, D.J., MacNeill, S.A., Southgate, J., McPheat, J. and Lane, D.P. (1991) p68 RNA helicase: identification of a nucleolar form and cloning of related genes containing a conserved intron in yeasts. *Mol. Cell Biol.*, **11**, 1326–1333.
39. Barta, I. and Iggo, R. (1995) Autoregulation of expression of the yeast Dbp2p 'DEAD-box' protein is mediated by sequences in the conserved DBP2 intron. *EMBO J.*, **14**, 3800–3808.
40. Rossler, O.G., Hloch, P., Schutz, N., Weitzenegger, T. and Stahl, H. (2000) Structure and expression of the human p68 RNA helicase gene. *Nucleic Acids Res.*, **28**, 932–939.
41. Will, C.L., Urlaub, H., Achsel, T., Gentzel, M., Wilms, M. and Luhrmann, R. (2002) Characterization of novel SF3b and 17S U2 snRNP proteins, including a human Prp5p homologue and an SF3b DEAD-box protein. *EMBO J.*, **21**, 4978–4988.
42. Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*. 2nd edn. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
43. Baykov, A.A., Evtushenko, O.A. and Avaeva, S.M. (1988) A malachite green procedure for orthophosphate determination and its use in alkaline phosphatase-based enzyme immunoassay. *Anal. Biochem.*, **171**, 266–270.
44. Breslauer, K.J., Frank, R., Blocker, H. and Marky, L.A. (1986) Predicting DNA duplex stability from the base sequence. *Proc. Natl Acad. Sci. USA*, **83**, 3746–3750.
45. Xia, T., SantaLucia, J., Burkard, M.E., Kierzek, R., Schroeder, S.J., Jiao, X., Cox, C. and Turner, D.H. (1998) Thermodynamic parameters for an expanded nearest-neighbor model for formation of RNA duplexes with Watson–Crick base pairs. *Biochemistry*, **37**, 14719–14735.
46. Wong, I. and Lohman, T.M. (1993) A double-filter method for nitrocellulose-filter binding: application to protein–nucleic acid interactions. *Proc. Natl Acad. Sci. USA*, **90**, 5428–5432.
47. Beckman, M.T.L. and Kirkegaard, K. (1998) Site size of cooperative single-stranded RNA binding by poliovirus RNA-dependent RNA polymerase. *J. Biol. Chem.*, **273**, 6724–6730.
48. Veaute, X., Jeusset, J., Soustelle, C., Kowalczykowski, S.C., Le Cam, E. and Fabre, F. (2002) The Srs2 helicase prevents recombination by disrupting Rad51 nucleoprotein filaments. *Nature*, **423**, 309–312.
49. Suk, K., Kim, S., Kim, Y.H., Oh, S.H., Lee, M.K., Kim, K.W., Kim, H.D. and Seo, Y.S. (2000) Identification of a novel human member of the DEAD box protein family. *Biochim. Biophys. Acta*, **1501**, 63–69.
50. Ogilvie, V.C., Wilson, B.J., Nicol, S.M., Morrice, N.A., Saunders, L.R., Barber, G.N. and Fuller-Pace, F.V. (2003) The highly related DEAD box RNA helicases p68 and p72 exist as heterodimers in cells. *Nucleic Acids Res.*, **31**, 1470–1480.
51. Bates, G.J., Nicol, S.M., Wilson, B.J., Jacobs, A.M., Bourdon, J.C., Wardrop, J., Gregory, D.J., Lane, D.P., Perkins, N.D. and Fuller-Pace, F.V. (2005) The DEAD box protein p68: a novel transcriptional coactivator of the p53 tumour suppressor. *EMBO J.*, **24**, 543–553.
52. Nadler, S.G., Roberts, W.J., Shamoo, Y. and Williams, K.R. (1990) A novel function for Zinc(II) in a nucleic acid-binding protein. *J. Biol. Chem.*, **265**, 10389–10394.
53. Iost, I., Dreyfus, M. and Linder, P. (1999) Ded1p, a DEAD-box protein required for translation initiation in *Saccharomyces cerevisiae*, is an RNA helicase. *J. Biol. Chem.*, **274**, 17677–17683.
54. Tanaka, N. and Schwer, B. (2005) Characterization of the NTPase, RNA-binding, and RNA helicase activities of the DEAD-box splicing factor Prp22. *Biochemistry*, **44**, 9795–9803.
55. Methot, N., Pause, A., Hershey, J.W.B. and Sonenberg, N. (1994) The translation initiation factor eIF-4B contains an RNA binding region that is distinct and independent from its ribonucleoprotein consensus sequence. *Mol. Cell Biol.*, **14**, 2307–2316.
56. Kowalczykowski, S.C., Lonberg, N., Newport, J.W. and von Hippel, P.H. (1981) Interactions of bacteriophage T4-coded gene 32 protein with nucleic acids. I. Characterization of the binding interactions. *J. Mol. Biol.*, **145**, 75–104.
57. Bycroft, M., Hubbard, T.J.P., Proctor, M., Freund, S.M.V. and Murzin, A.G. (1997) The solution structure of the S1 RNA binding domain: a member of an ancient nucleic acid-binding fold. *Cell*, **88**, 235–242.
58. Oberosler, P., Hloch, P., Ramsperger, U. and Stahl, H. (1993) p53-Catalyzed annealing of complementary single-stranded nucleic acids. *EMBO J.*, **12**, 2389–2396.
59. Garcia, P.L., Liu, Y., Jiricny, J., West, S.C. and Janscak, P. (2004) Human RECQ5beta, a protein with DNA helicase and strand-annealing activities in a single polypeptide. *EMBO J.*, **23**, 2882–2891.
60. Pant, K., Karpel, R.L., Rouzina, I. and Williams, M.C. (2004) Mechanical measurement of single-molecule binding rates: kinetics of DNA helix–destabilization by T4 gene 32 protein. *J. Mol. Biol.*, **336**, 851–870.
61. Nair, R., Carter, P. and Rost, B. (2003) NLSdb: database of nuclear localization signals. *Nucleic Acids Res.*, **31**, 397–399.
62. Yamaki, A., Kudoh, J., Shimizu, N. and Shimizu, Y. (2004) A novel nuclear localization signal in the human single-minded proteins SIM1 and SIM2. *Biochem. Biophys. Res. Comm.*, **313**, 482–488.
63. Streb, J.W. and Miano, J.M. (2005) Cross-species sequence analysis reveals multiple charged residue-rich domains that regulate nuclear/cytoplasmic partitioning and membrane localization of a kinase anchoring protein 12 (SseCKS/Gravin). *J. Biol. Chem.*, **280**, 28007–28014.
64. Lorsch, J.R. and Herschlag, D. (1998b) The DEAD box protein eIF4A. 2. A cycle of nucleotide and RNA-dependent conformational changes. *Biochemistry*, **37**, 2194–2206.
65. Kim, S.-H., Smith, J., Claude, A. and Lin, R.-J. (1992) The purified yeast pre-mRNA splicing factor PRP2 is an RNA-dependent NTPase. *EMBO J.*, **11**, 2319–2326.
66. Shi, H., Cordin, O., Minder, C.M., Linder, P. and Xu, R.-M. (2004) Crystal structure of the human ATP-dependent splicing and export factor UAP56. *Proc. Natl Acad. Sci. USA*, **101**, 17628–17633.
67. Lorsch, J.R. and Herschlag, D. (1998a) The DEAD box protein eIF4A. 1. A minimal kinetic and thermodynamic framework reveals coupled binding of RNA and nucleotide. *Biochemistry*, **37**, 2180–2193.
68. Linder, P. (2003) Yest RNA helicases of the DEAD-box family involved in translation initiation. *Biol. Cell*, **95**, 157–167.

69. Wiekowski, M., Schwarz, M.W. and Stahl, H. (1988) Simian virus 40 large T antigen DNA helicase. Characterization of the ATPase-dependent DNA unwinding activity and its substrate requirements. *J. Biol. Chem.*, **263**, 436–442.
70. Muller, U.F. and Goring, H.U. (2002) Mechanism of the gBP21-mediated RNA–RNA annealing reaction: matchmaking and charge reduction. *Nucleic Acids Res.*, **30**, 447–455.
71. Golas, M.M., Sander, B., Will, C.L., Luhrmann, R. and Stark, H. (2005) Major conformational change in the complex SF3b upon integration into the spliceosomal U11/U12 di-snRNP as revealed by electron cryomicroscopy. *Mol. Cell*, **17**, 869–883.
72. Will, C.L., Schneider, C., Reed, R. and Luhrmann, R. (1999) Identification of both shared and distinct proteins in the major and minor spliceosomes. *Science*, **284**, 2003–2005.