

# Molecular Cloning and Characterization of a Putative Nuclear DEAD Box RNA Helicase in the Spruce Budworm, *Choristoneura fumiferana*

D-Y. Zhang,<sup>1,2</sup> D.R. Ampasala,<sup>1,2</sup> S-C. Zheng,<sup>1,2</sup> M. Cusson,<sup>3</sup> X-W. Cheng,<sup>1,2,4</sup> P.J. Krell,<sup>2</sup> and Q-L. Feng<sup>1\*</sup>

RNA helicases play important roles in cellular processes such as pre-mRNA splicing, rRNA processing, ribosomal biogenesis, and translation. A full-length DEAD box RNA helicase cDNA (*CfHlc113*) was isolated from the spruce budworm, *Choristoneura fumiferana*. *CfHlc113* contained the eight functional motifs, which are highly conserved in the DEAD box RNA helicase family, and an arginine-serine-aspartate (RSD) domain at its *N*-terminal end. *CfHlc113* was highly homologous to *Rattus norvegicus* HEL117 and human prp5 genes, both of which are suggested to be involved in RNA splicing. The results of Northern and Western blotting showed that expression of the *CfHlc113* gene was low or undetectable in eggs, larvae, pupae, and adults. High levels of expression were, however, detected in the three in vitro cultured cell lines, CF-203, CF-124T, and CF-70, which were developed from the midgut, ovaries, and neonate larvae, respectively. Immunocytochemistry revealed that *CfHlc113* protein was present exclusively in the nuclei of these cell lines. Arch. Insect Biochem. Physiol. 61:209–219, 2006. © 2006 Wiley-Liss, Inc.

KEYWORDS: RNA unwinding; RNA splicing; ribosome biogenesis; translation

## INTRODUCTION

ATP-dependent RNA helicases are a family of proteins that are capable of unwinding double-stranded RNA and DNA/RNA hybrids (Lüking et al., 1998). Some members of this family have been experimentally demonstrated to play important roles in transcription, translation, mRNA splicing, and DNA replication (Lüking et al., 1998; de la Cruz et al., 1999). The most extensively studied samples include: eukaryotic initiation factor-4A (eIF-4A, Abramson et al., 1987; Dorn et al., 1993); pre-RNA

processing proteins (prp proteins, Dalbadie-McFarland and Abelson, 1990; Dayyeh et al., 2002); *Drosophila vasa* (Lasko and Ashburner, 1988); rat HEL117 (Sukegawa and Blobel, 1995); and human p68 (Ford et al., 1988; Hloch et al., 1990). This helicase family is characterized by a common core region that consists of eight highly conserved signature motifs (Lüking et al., 1998; Aubourg et al., 1999). Many members of this family have a DEAD (Asp-Glu-Ala-Asp) box in Motif IV. Therefore, they are called DEAD box RNA helicases (Schmid and Linder, 1992; Lüking et al., 1998).

<sup>1</sup>Great Lakes Forestry Centre, Canadian Forest Service, Sault Ste. Marie, Ontario, Canada

<sup>2</sup>Department of Microbiology, University of Guelph, Guelph, Canada

<sup>3</sup>Laurentian Forestry Centre, Canadian Forest Service, Ste-Foy, Canada

<sup>4</sup>Department of Microbiology, Miami University, Oxford, Ohio, USA

Grant sponsor: Canadian Biotechnology Strategy Fund; Grant sponsor: Genome Canada.

\*Correspondence to: Qili Feng, E-mail: qfeng@nrcc.gc.ca; Current address: College of Life Science, South China Normal University, Guangzhou, China, 510631.

E-mail: qifeng@scnu.edu.cn

Received 12 June 2005; Accepted 10 August 2005

RNA helicases have been identified from almost all of organisms. In insects, several *Drosophila* RNA helicases have been well studied. *Drosophila* maleless protein (MLE) is an RNA helicase required for X chromosome dosage compensation and essential for male viability (Kuroda et al., 1991). A *Drosophila* homologue of the eukaryotic initiation factor-4A (eIF-4A) gene is maternally expressed in the oocyte; mutation of the gene is recessive lethal (Dorn et al., 1993). *Drosophila* ME31B is also a maternally expressed DEAD box RNA helicase and is strongly expressed during oogenesis (de Valoir et al., 1991). *Drosophila hel* protein is an enhancer of *white* variegation and is associated with chromosomes in cell nuclei of embryos and ovaries (Eberl et al., 1997). Two *vasa* genes were isolated from insects. *Drosophila vasa* gene is homologous to eIF-4A and its product is required in only the female germ line (Lasko and Ashburner, 1988) and is involved in the formation of polar granules and germ cells (Hay et al., 1988). The *Bombyx mori vasa* gene is the first reported lepidopteran RNA helicase and is expressed only in the germline (Nakao, 1999). The biological functions of most insect RNA helicases are not clear and need to be further investigated.

We have previously identified a 64-kDa DEAD box RNA helicase from the spruce budworm, *Choristoneura fumiferana* (Zhang et al., 2004), which is a homologue of human nucleolar RNA helicase NOH61 (Zirwes et al., 2000) and yeast RNA helicase Dbp9p (Daugeron et al., 2001). We report here molecular cloning and characterization of another DEAD box RNA helicase cDNA (*CfrHlc113*) from *C. fumiferana*. *CfrHlc113* was a homologue of rat HEL117 (Sukegawa and Blobel, 1995) and human prp5 (Will et al., 2002) genes, which are involved in RNA splicing.

## MATERIALS AND METHODS

### Experimental Insects

East spruce budworm (*Choristoneura fumiferana* Clem., Lepidoptera: Tortricidae) eggs were maintained at 22°C and 70% relative humidity (RH) and allowed to hatch into first instar larvae on the balsam fir needles. The first instar larvae molted

into second instar in 6 days. The second instar larvae were maintained at 16°C for one week and then the diapausing second instar were stored at 2°C for 27 weeks to satisfy the obligatory requirement for cold treatment. At the end of this period, the larvae were moved from 2° to 16°C for 1 week and then placed on artificial diet (McMorran, 1965) at 22°C, 70% RH and a photoperiod of 12-h light and 12-h darkness and reared until they reached the adult stage.

### Cell Lines

Three spruce budworm cell lines, FPMI-CF-203, IPRI-CF-124T, and IPRI-CF-70, were used in this study. FPMI-CF-203 was developed from the midgut tissues (CF-203, Sohi et al., 1993); IPRI-CF-124T was derived from neonate larvae (CF-124T, Billmorria and Sohi, 1977), and IPRI-CF-70 was from ovary (CF-70, Caputo Guido, personal communication). CF203 cells were grown in 25-ml flasks with SF900 medium supplemented with 5% FBS (Life Technologies, Inc., Gaithersburg, MD). Cells of CF-124T and CF-70 were grown in modified Grace's media (Grace, 1962).

### Construction and Screening of cDNA Library

Messenger RNA was isolated from the midgut cell line CF-203, which is responsive to juvenile hormone (Feng et al., 1999). A cDNA library was constructed in Uni-ZAP XR vector (Stratagene, La Jolla, CA) according to the manufacturer's instructions. Polyclonal antibodies raised against *Rst(1)JH* protein (formally called MET, Ashok et al., 1998; Pursley et al., 2000) of *D. melanogaster* were used to screen the cDNA library for *Rst(1)JH* homolog in the spruce budworm. Positive plaques were selected after three rounds of screenings and plasmid cDNA was isolated.

### Sequence Analysis

Sequencing was performed using ALFexpress™ AutoRead™ sequencing Kit and ALFexpress™ DNA

sequencer (Amersham Pharmacia Biotech, Piscataway, NJ). Annotation, comparison, and alignment of sequences were performed using the National Center for Biotechnology Information BLAST search services (Altschul et al., 1990) and Clustal Alignment Program (Higgins and Sharp, 1988) of DNASTAR (DNASTAR, Inc., Madison, WI). Prediction of the subcellular localization of protein was conducted using PSORT program (Nakai and Kanehisa, 1992; <http://psort.nibb.ac.jp/>).

### Production of Recombinant Protein in Baculovirus Expression System

The open reading frame of *CfrHlc113* was cloned into the pFastBac1 donor plasmid (Invitrogen, Burlington, Ontario, Canada). The generated recombinant plasmid was used to transform DH10Bac cells, in which the *CfrHlc113* cDNA was transposited to Bacmid DNA (modified *Autographa californica* multicapsid nucleopolyhedrovirus DNA, AcMNPV). Insertion of *CfrHlc113* cDNA into the AcMNPV genome was confirmed by using *CfrHlc113*-specific primers and PCR according to the manufacturer's instruction. Sf21 cells were transfected with the recombinant AcMNPV. Expression of *CfrHlc113* was examined using SDS-PAGE and Western blotting.

### SDS-PAGE

Proteins were denatured at 100°C for 5 min in an equal volume of 2× protein loading buffer (0.1 M Tris buffer, pH 6.8, 4% SDS, 0.2% β-mercaptoethanol, 40% glycerol and 0.002% bromophenol blue) and were then separated in 8% acrylamide gels in Tris-glycine-SDS buffer (10 mM Tris, 50 mM glycine, 0.1% SDS, pH 8.0) in a mini vertical electrophoresis system (Bio-Rad Laboratories, Hercules, CA). Thirty micrograms of protein was used per lane. The gels were stained with Coomassie Blue R-250.

### Antibody Production

The recombinant protein expressed in an AcMNPV recombinant was excised from SDS-

PAGE gels. Polyclonal antiserum was made in a rabbit by Cedarlane Inc. (Hornby, Ontario, Canada). Antiserum was collected after three boost injections, each with 200 ng protein in the Freund's adjuvant. Pre-immune serum collected from the same rabbit prior to immunization was used as a control.

### Western Blotting

Proteins were transferred from SDS-PAGE gels to nitrocellulose membranes after electrophoresis. Thirty micrograms of protein was used per lane. The anti-*CfrHlc113* antibodies were used as primary antibodies at 1:1,000 dilution. The goat anti-rabbit IgG alkaline phosphatase conjugate (Sigma Chemical Company, St. Louis, MO) was used as secondary antibody at 1:2,000 dilutions. Color development of the alkaline phosphatase reactions was performed using 5-bromo-4-chloro-3-indolyl phosphate and nitroblue tetrazolium as substrates.

### Immunofluorescence Localization

Immunofluorescence localization of *CfrHlc113* was performed according to the general procedures described by Watkins (1996) for insect cells. Cells were grown on glass slides and fixed with 2% paraformaldehyde fixative plus 0.1% Triton X-100 on ice for 30 min. Cells were then permeabilized in cold methanol for 5 min. Antibodies at 1:1,000 dilution were applied to the cells on the slide and incubated for 1 h. The cells were blocked for 1 h in 1.5% bovine serum albumin and were then washed four times in phosphate buffered saline (PBS), followed by treatment with fluorescein conjugated anti-rabbit IgG (Sigma-Aldrich Canada Ltd. Oakville, ON, Canada) (1:500 dilution) for 1 h. The cells were washed three times in PBS and counter-stained with 4',6-diamidino-2'-phenylindole dihydrochloride (DAPI) for 30 min and examined by fluorescence microscopy. The photographs were taken as double exposures using fluorescence and DAPI filters.

## Northern Blotting

Ten micrograms of total RNA per lane were separated on 1% formaldehyde-agarose gels. The RNA was visualized by staining with ethidium bromide and photographed under UV light. The RNA was then transferred to nylon membranes. The blots were pre-hybridized in a pre-hybridization solution (Rapid-hyb buffer, Amersham Pharmacia Biotech) for 4 h and then hybridized in a hybridization solution (pre-hybridization solution plus  $^{32}\text{P}$ -dCTP labeled *CfrHlc113* DNA probe) at 65°C for at least 12 h. After hybridization, the membranes were washed twice in 2× SSC plus 0.1% SDS at 42°C for 15 min and twice with 0.5× SSC plus 0.1% SDS at 55°C for 15 min and once with 0.1× SSC plus 0.1% SDS at 65°C for 15 min.

## RESULTS

### Cloning and Sequence of *CfrHlc113*

In an attempt to clone *Drosophila Rst(1)JH* protein homologue, which was formally called MET and has been suggested to be a JH receptor (Ashok et al., 1998; Pursley et al., 2000), from the spruce budworm, we used the anti-*Rst(1)JH* protein antibodies to screen an expression cDNA library of the spruce budworm CF-203 cell line, which was responsive to JH I (Feng et al., 1999). Ten immunologically positive clones were isolated from a total of approximately  $10^6$  plaques. Sequencing and comparison with the sequences in the GenBank database revealed that three of the ten positive clones were identical and encoded an already identified *C. fumiferana* member (64 kDa) of the DEAD box RNA helicases family (*CfrHlc64*, Zhang et al., 2004). Four other clones were identical and encoded a different 113-kDa RNA helicase that also belongs to the DEAD box family of RNA helicases (this report). Another clone encoded a nuclear exportin (unpublished data) and the remaining two clones encoded two unknown proteins (unpublished data). The present report describes the *C. fumiferana* 113-kDa RNA helicase, *CfrHlc113*.

The *CfrHlc113* cDNA was 3,196 base pairs in length. The longest open reading frame was 3,039 nucleotides in length, potentially encoding a protein of 1,012 amino acids with a calculated molecular mass of 113 kDa and a pI of 9.58 (GenBank accession number for the sequence is AY559246). All of the eight conserved functional motifs that define the RNA helicase family were found in *CfrHlc113*. A tetrapeptide Asp-Glu-Ala-Asp (DEAD) was present in the sequence, indicating that it was a member of the DEAD box RNA helicase family.

Comparison of the deduced amino acid sequence of *CfrHlc113* with 43 RNA helicases from different species revealed that *CfrHlc113* showed high identities to homologues of *Mus musculus* (70%, NP\_666087.1), *Anopheles gambiae* (70%, XP\_311375.1), *D. melanogaster* (64%, NP\_573020.2), *Rattus norvegicus* HEL117 (57%, Sukegawa and Blobel, 1995. NP\_620798.1), and *Homo sapiens* prp5 (57%, Will et al., 2002. AAH12304.1). *CfrHlc113* also showed 18–46% identity to homologues from other species ranging from fungus to plants. In general, these RNA helicases can be clustered into three broad phylogenetic groups (Fig. 1A). Group I includes *CfrHlc113*, the well-characterized *R. norvegicus* HEL117 (Sukegawa and Blobel, 1995; NP\_620798.1), and human prp5 (Will et al., 2002; AAH12304.1). It has been suggested that these two proteins are involved in mRNA splicing. *CfrHlc113* also showed 30% identity to human p68 (Ford et al., 1988; NP\_004387), and 27% identity to *D. melanogaster vasa* (Lasko and Ashburner, 1988; CAA31405) and *Bombyx mori vasa*-like gene (Nakao, 1999; BAA19572), which are believed to be involved in early embryogenesis and oogenesis. Functions of most other members in this group have not been reported. Group II includes bacterial RNA helicases (NP\_755783, Q8XA87 and NP\_457662) from *E. coli* and *Buchnera aphidicola*, which have been suggested to be cold-shock helicases (Jones et al., 1996). The well-characterized transcription factor *Dm*elF4A (Dorn et al., 1993; CAA48790), *Dm*ME31B (De Valoir et al., 1991; AAA28603), and *Dm*hel (Eberl et al., 1997; AAB65835) are also clustered into this group. Group III is represented by *CfrHlc64* (Zhang et al.,

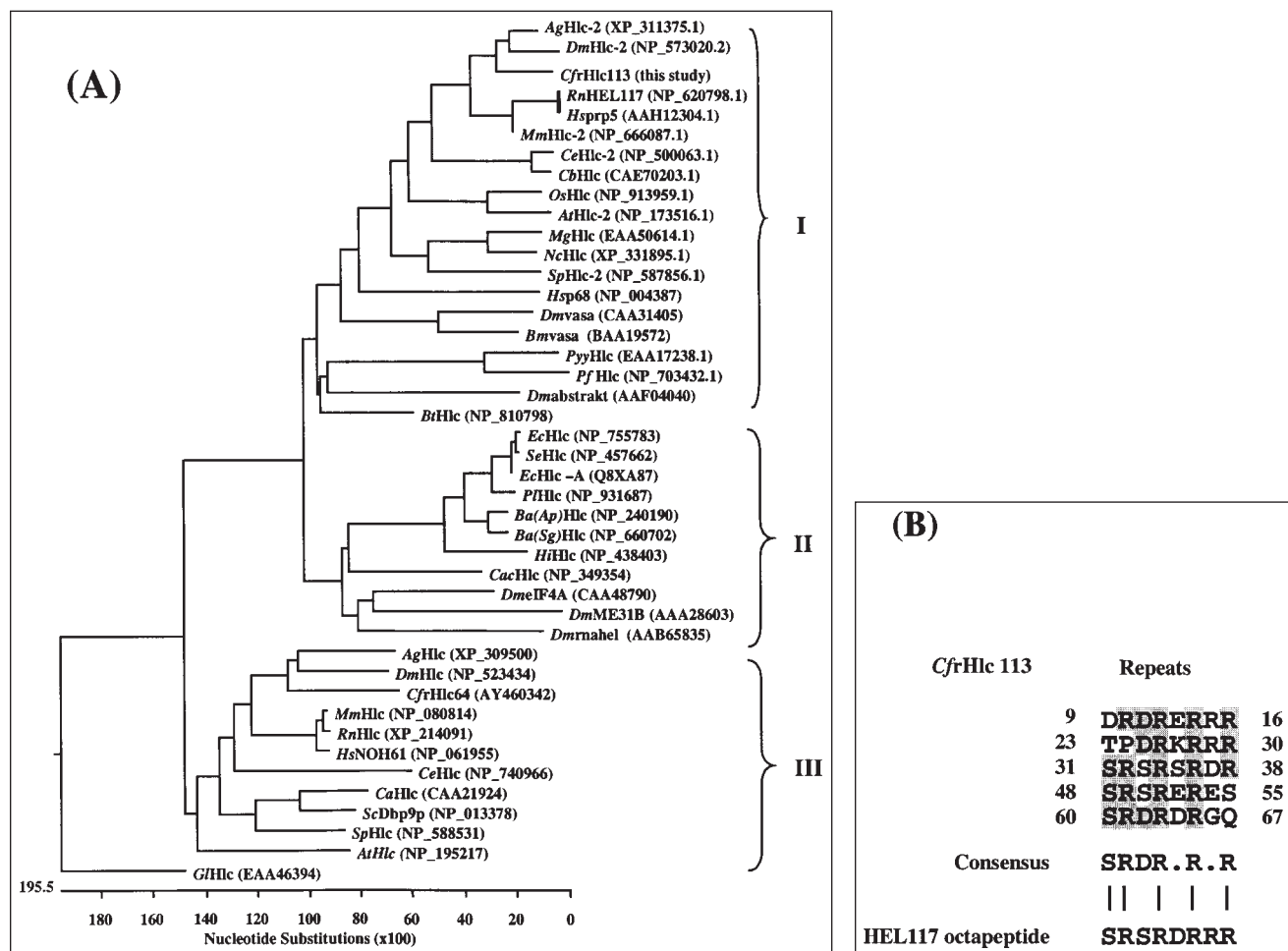


Fig. 1. A: Phylogenetic analysis of 44 RNA helicases from different species: CfrHlc113 (*C. fumiferana*, this study, AY559246), CfrHlc64 (*C. fumiferana*, Zhang et al., 2004; AY460342), AgHlc-2 (*A. gambiae*, XP\_311375.1), DmHlc-2 (*D. melanogaster*, Lasko and Ashburner, 1988; NP\_573020.2), RnHEL117 (*R. norvegicus*, Sukegawa and Blobel, 1995; NP\_620798.1), Hsprp5 (*H. sapiens*, Will et al., 2002; NP\_12304.1), MmHlc-2 (*M. musculus*; NP\_666087.1), CeHlc-2 (*Caenorhabditis elegans*, NP\_500063.1), CbHlc (*C. briggsae*, CAE70203.1), OsHlc (*Oryza sativa*, NP\_913959.1), AtHlc-2 (*Arabidopsis thaliana*, NP\_173516.1), MgHlc (*Magnaporthe grisea*, EAA50614.1), NcHlc (*Neurospora crassa*, XP\_331895.1), SpHlc-2 (*Schizosaccharomyces pombe*, NP\_587856.1), Hsp68 (*H. sapiens*, Ford et al., 1988; NP\_004387), Dmvasa (*D. melanogaster*, Lasko and Ashburner, 1988; CAA31405), Bmvasa (*Bombyx mori*, Nakao, 1999; BAA19572), PyyHlc (*Plasmodium yoelli yoelli*, EAA17238.1), PfHlc (*P. folci*, NP\_703432.1), Dmabstrakt (*D. melanogaster*, AAF04040), BtHlc (*Bacteroides thetaiotaomicron*, NP\_810798), EcHlc (*Escherichia coli* cold-shock helicase, NP\_755783), SeHlc (*Salmonella enterica*, NP\_457662),

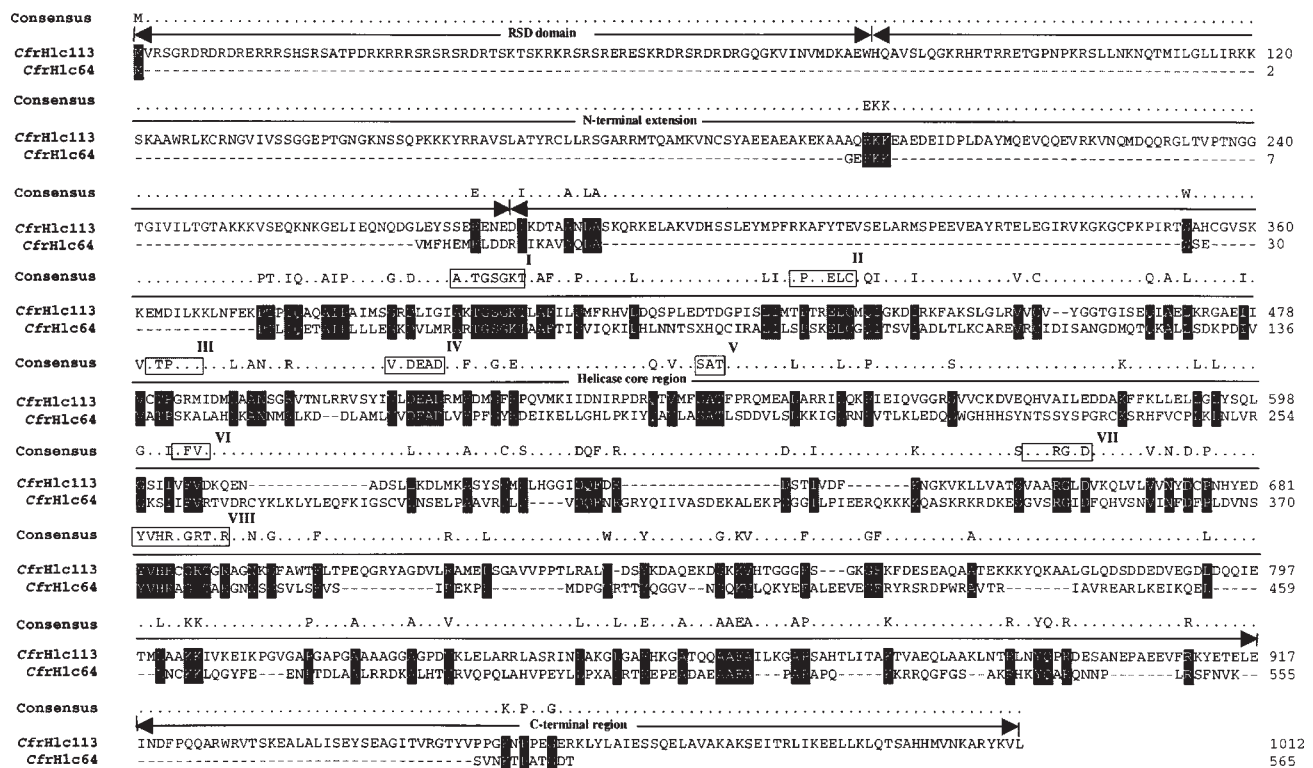
EccsdA (*E. coli* cold-shock helicase A, Q8XA87), PIHlc (*Photobacterium luminescens*, NP\_931687), Ba(Ap)Hlc [*Buchnera aphidicola* str. *Aps* (*Acyrtosiphon pisum*), NP\_240190], Ba(Sg)Hlc [*B. aphidicola* str. *Sg* (*Schizaphis graminum*), NP\_660702], HiHlc (*Haemophilus influenzae*, NP\_438403), CacHlc (*Clostridium acetobutylicum*, NP\_349354), DmeIF4A (*D. melanogaster*, CAA48790), DmME31B (De Valoir et al., 1991; *D. melanogaster*, AAA28603) Dmhel (*D. melanogaster*, Eberl et al., 1997; AAB65835), AgHlc (*A. gambiae*, XP\_309500), DmHlc (*D. melanogaster*, NP\_523434), MmHlc (*M. musculus*; NP\_080814), RnHlc (*R. norvegicus*, XP\_214091), HsNOH61 (*H. sapiens*, NP\_061955), CeHlc (*C. elegans*, NP\_740966), CaHlc (*Candida albicans*, CAA21924), ScDbp9p (*Saccharomyces cerevisiae*, NP\_013378), SpHlc (*S. pombe*, NP\_588531); AtHlc (*A. thaliana*, NP\_195217), GfHlc (*Giaria lamblia*, EAA46394). B: Octapeptide repeats of the RSD domain in CfrHlc113. The residues that are identical to the HEL117 octapeptide are shaded. The consensus residues are given when at least three residues match in the five repeats.

2004; AY460342), the human nucleolar RNA helicase, NOH61 (Zirwes et al., 2000; NP\_061955), and the yeast RNA helicase, Dbp9p (Daugeron et al., 2001; NP\_013378). NOH61 and Dbp9p are suggested to be essential for ribosomal biogenesis.

The eight highly conserved helicase core motifs were found in all these sequences, but their C- and N-termini were variable. The members in Group I have an RSD domain, EK domain, and/or Q domain at either the C- or N-terminal ends (Aubourg

et al., 1999). The members of Group II and III do not usually have extending domains at either end. There was an arginine-serine-aspartate-rich (RSD) domain at the N-terminus of *CfrHlc113* (Fig. 2A). In the RSD domain, 48 out of 79 amino acids residues (61%) were arginine/serine/aspartate and present as Arg-Ser (RS) or Ser-Arg (SR) or Arg-Asp (RD) dipeptides. Five repeats of an octapeptide motif, SRDRxRxR, were found in this domain (Fig. 1B). The repeats were conserved with that in *R.*

(A)



(B)

Motifs	I	II	III	IV	V	VI	VII	VIII
DEAD box RNA helicases	NH2--AQSGSKT--	APTRELA--	ATPGRL--	VLDEAD--	SAT--	FVNT--	VAARGLD--	YIHRIGRTGR--COOH
<i>CfrHlc113</i>	NH2--AKTGSKT--	TPTRELC--	CTPGRM--	VLDEAD--	SAT--	FVDK--	VAARGLD--	YVHRCGRTGR--COOH
<i>CfrHlc64</i>	NH2--ARTGSKT--	SPSKELC--	ATPSKA--	VVDEAD--	SAT--	FVRT--	V-SRGID--	YVHRAGRTAR--COOH
Consensus	NH2--A--GSKT--	--P--EL--	--TP----	V-DEAD--	SAT--	FV--	--V--RG-D--	Y--HR-GRT-R--COOH

Fig. 2. A: Alignment of amino acid sequences of *CfrHlc113* and *CfrHlc64*. The eight conserved motifs are boxed and numbered I to VIII. The RSD-domain and the core region are based on Aubourg et al. (1999). B: Alignment of the functional motifs of DEAD box RNA helicases

with *CfrHlc113* and *CfrHlc64*. The conserved amino acid sequences of the eight functional motifs are cited from Aubourg et al. (1999). The residues that match the consensus residues are bolded.

*norvegicus* HEL117 and other RSD domain proteins (Sukegawa and Blobel, 1995; Aubourg et al., 1999). *CfrHlc113* did not have Asp-Glu-Arg-Lys (DERK) and Arg-Gly-Gly (RGG) domains or a consecutive Gly hinge, which are found in some RNA-binding proteins between the RSD domain and the RNA helicase core region (Birney et al., 1993; Gibson and Thompson, 1994).

Prediction of subcellular localization using Reinhardt's method (Reinhardt and Hubbard, 1998) revealed that *CfrHlc113* was a nuclear protein (77% reliability). By using the *k*-NN Prediction (Horton and Nakai, 1997), the possibility for nuclear localization was 87%, much higher than the possibilities for cytoplasmic (8.7%) and cytoskeletal (4.3%) localizations, indicating the possible nuclear localization of *CfrHlc113*.

### Comparison of *CfrHlc113* and *CfrHlc64*

Two DEAD box helicases, *CfrHlc113* (this report) and *CfrHlc64* (Zhang et al., 2004; AY 460342), were isolated using antibodies against *Rst(1)JH* protein. They differ in size and belong to different groups in the structural and phylogenetic tree analysis (Fig. 1). Comparison of these two sequences (Table 1) revealed that overall identities were only 20 and 26% at the amino acid and nucleotide levels, respectively (Fig. 2A). All eight of the functional motifs that are common in RNA helicases were found in these two sequences. They were highly conserved with 68% identity in overall residues within these motifs (Fig. 2B). The major difference between these two sequences was that

TABLE 1. Comparison of *CfrHlc113* and *CfrHlc64*

Properties	<i>CfrHlc113</i>	<i>CfrHlc64</i>
No. of nucleotides	3,196	1,999
GC%	53	49
No. of amino acid residues	1,012	555
Molecular mass (kDa)	112.9	63.5
pI	9.58	9.55
RSD domain	Yes	No
RNA helicase motifs	8	8
Predicted localization	Nuclear	Nuclear

*CfrHlc113* contained a longer *N*-terminus with an RSD domain and a longer C-terminal region, whereas *CfrHlc64* did not have an RSD domain in its short *N*-terminus (Fig. 2A).

### Expression of Recombinant *CfrHlc113* Protein in Baculovirus System

In vitro expression of *CfrHlc113* cDNA in a baculovirus system resulted in a recombinant protein with an apparent molecular mass of approximately 132 kDa in SDS-PAGE gel (Fig. 3), confirming that the open reading frame of the cDNA did encode a protein. The difference between apparent (132 kDa) and calculated (113 kDa) molecular mass values may be due to the high number of basic amino acids in the sequence and possible post-translational modification.

### Expression of *CfrHlc113* in Tissues and Cell Lines

Northern blotting analysis detected a 3.2-kb transcript of *CfrHlc113* in six-day-old eggs, 1st instar larvae, 6th instar larvae, pupae, and adults, each at 2 days old (Fig. 4A). The embryos and 1st instar larvae had low levels of transcripts, while pupae appeared to have the highest levels. The fat body, midgut, and epidermis of 6th instar larvae had similar though low levels of the *CfrHlc113* transcript. Interestingly, the three cell lines CF-203, CF-124T, and CF-70, which were derived from the midgut, neonate larvae, and ovary, respectively, con-

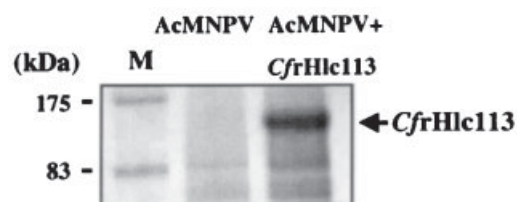


Fig. 3. Recombinant protein of *CfrHlc113* expressed in recombinant AcMNPV infected SF21 cells. The protein was isolated from SF21 cells 24 h post infection with AcMNPV+*CfrHlc113* virus or the control virus AcMNPV. The protein was separated in a 8% SDS-PAGE gel and stained with Coomassie Blue R-250.



tained much higher levels of the *CfrHlc113* transcript than the in vivo tissues (Fig. 4A).

Western blotting analysis did not reveal detectable *CfrHlc113* protein in six-day-old eggs, 1st instar larvae, 6th instar larvae, pupae, and adults, each at 2 days old (data not shown). However, high levels of *CfrHlc113* protein were detected in all three cell lines tested in this study (Fig. 4B), which is in concurrence with the results of the Northern blotting analysis (Fig. 4A).

### Cellular Localization of *CfrHlc113* in CF203 Cells

Immunocytochemistry analyses showed that the *CfrHlc113* protein was present exclusively in the nuclei of these three cell lines (Fig. 5). This observation confirmed its nuclear localization as predicted in the sequence analysis mentioned above.

### DISCUSSION

*CfrHlc113* was isolated by screening an expression cDNA library of the CF-203 cell line using the antibodies raised against *Rst(1)JH* protein (Pursley et al., 2000). It is not known if the *Rst(1)JH* antibodies contain non-specific antibodies to the helicase gene product or the *Rst(1)JH* antibody recognizes the helicase. However, only 11.6% overall amino acid identity was found between *CfrHlc113* and the *Rst(1)JH* proteins. There is no evidence to indicate a structural or functional similarity between *CfrHlc113* and the *Rst(1)JH* protein. Therefore, identification of this RNA helicase using the *Rst(1)JH* antibody may be spurious though fortuitous.

Interestingly, the same antibody identified the two RNA helicases, *CfrHlc113* and *CfrHlc64*. Sequence comparison revealed that *CfrHlc113* and

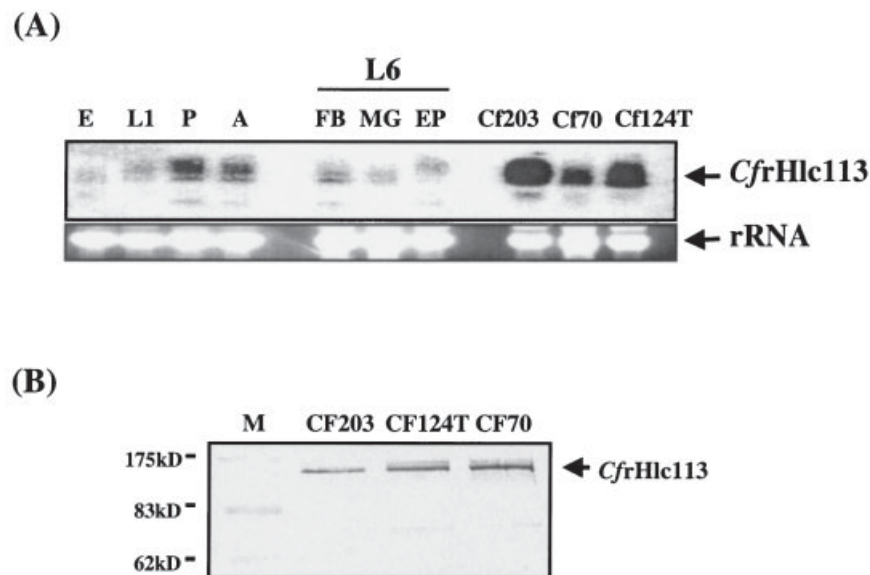


Fig. 4. **A:** Northern blotting analysis of expression of *CfrHlc113* mRNA in tissues and cell lines of *C. fumiferana*. Ten micrograms of total RNA hybridized with a  $^{32}\text{P}$ -labeled *CfrHlc113* cDNA as a probe. The estimated size of *CfrHlc113* mRNA was 3.2 kb. Ribosomal RNA stained with ethidium bromide indicates equivalent loading of total RNA. CF-203, CF-124T, and CF-70 were used for protein analysis after 4 days in culture. E: Six-day-old eggs; L1: two-day-old 1st instar larvae; P: two-day-old pupae; A: two-

day-old adults; L6: two-day-old 6th instar larvae; FB: fat body; MG: midgut; EP: epidermis. **B:** Western blotting analysis of expression of *CfrHlc113* in cell lines of *C. fumiferana*. CF203 cells were grown in 25-ml flasks with SF900 medium supplemented with 5% FBS. Cells of CF-124T and CF-70 were grown in modified Grace's media. Proteins were extracted from the cells for protein analysis after 4 days in culture. M: molecular marker. Thirty micrograms of protein was loaded in each lane.



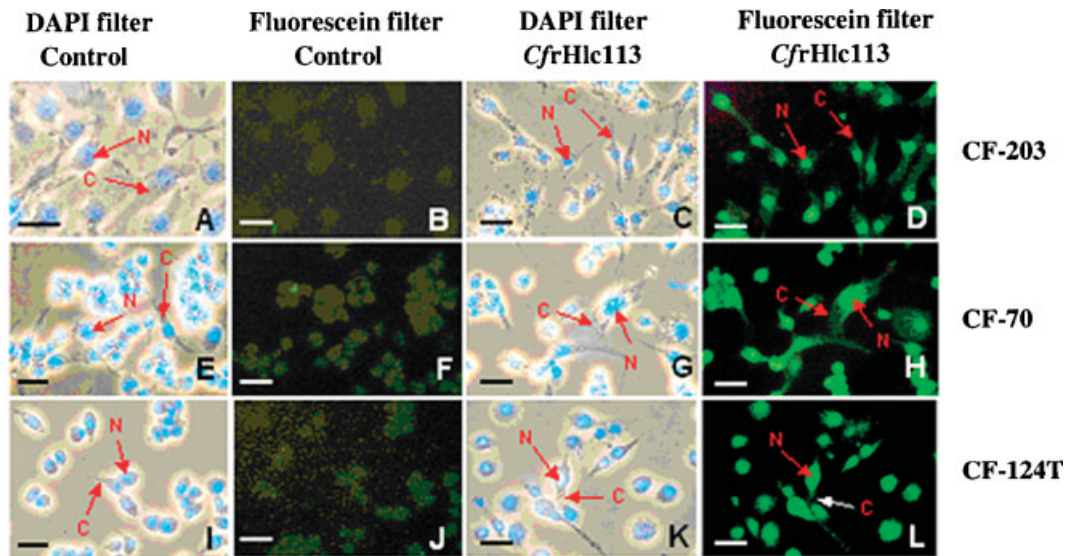


Fig. 5. Immunocytochemical localization of *CfrHlc113* in CF-203, CF-124T, and CF-70 cells by immunofluorescence and DAPI staining. CF203 cells were grown in SF900 medium supplemented with 5% FBS; cells of CF-124T and CF-70 were grown in modified Grace's media. A–D: CF-203; E–H: CF-70T; I–L: CF-124T. A, E, I: Controls stained with DAPI; B, F, J: controls stained with fluorescein; C, G, K: treated with anti-*CfrHlc113* antibody and stained with DAPI; D, H, L: treated with anti-*CfrHlc113* antibody and stained with fluorescein. The green fluorescence indicates

the presence of *CfrHlc113* protein and the blue indicates the presence of nuclear DNA. The primary antibody was rabbit anti-*CfrHlc113* at a dilution of 1:1,000, and the secondary antibody was fluorescein-labeled sheep anti-rabbit IgG F(ab)<sub>2</sub> fragment at a dilution of 1:500. The staining was examined using fluorescein and DAPI filters under an Olympus fluorescence microscope (Olympus BX50) and photographically documented. N: nuclei; C: cytoplasm. Scale bars = 60  $\mu$ m.

*CfrHlc64* displayed high similarity only in their helicase core regions, with the *N*- and *C*-terminal ends being dissimilar. For example, *CfrHlc113* had an RSD domain in its *N*-terminal end, whereas *CfrHlc64* did not. We have found that expression of *CfrHlc64* was up-regulated by the ecdysone agonist tebufenozide (RH5992) (Zhang et al., 2004). However, expression of *CfrHlc113* was not inducible by either 20E or RH5992 (data not shown). These features suggest that these two proteins may play different cellular functions.

In an attempt to predict its function, *CfrHlc113* was aligned with other known RNA helicases. *CfrHlc113* showed high identities to homologues in *M. musculus* (70%), *A. gambiae* (70%), and *D. melanogaster* (64%). However, the functions of these homologues have not yet been reported. *CfrHlc113* was also 56% identical to rat HEL117 (Sukegawa and Blobel, 1995) and human prp5p

(Will et al., 2002). Rat HEL117 also contains an RSD domain and a DERK region. It is co-located in the nucleus with splicing factor SC35 and may be involved in pre-mRNA splicing (Fu and Maniatis, 1990). Human prp5 is involved in pre-mRNA splicing and pre-spliceosome assembly (Will et al., 2002) by enhancing an ATP-dependent structural change in the U2 small nuclear ribonucleoprotein (snRNP), facilitating the interaction of the snRNP with the complementary region of the pre-mRNA (O'Day et al., 1996; Dayyeh et al., 2002).

From the above structural and sequence analyses, we hypothesize that *CfrHlc113* is a pre-mRNA processing protein or splicing factor. The subcellular localization prediction based on sequence analysis and the observation that *CfrHlc113* was localized in nuclei of the three cell lines also support this hypothesis. This finding is consistent with the reports on intranuclear localization of other

RS-domain RNA helicases such as HEL117 (Fu and Maniatis, 1990; Sukegawa and Blobel, 1995).

One of the interesting findings in this study is that *CfrHlc113* protein was undetectable in the stages and tissues of larvae tested using Western blotting analysis, whereas high levels of *CfrHlc113* protein were detected in the three cell lines. Failure to detect the protein in larval tissues may be due to too low levels of expression to be detected by Western blotting. Another possibility is that the protein expression might have been restricted to very narrow developmental stages. High expression in the cell cultures suggests that *CfrHlc113* may be necessary for in vitro growth and cell division. In addition, high expression in the cells was not dependent on the original tissues from which the cell lines were developed because all three cell lines yielded high expression levels. *CfrHlc113* was also similar (30%) to the nuclear helicase, p68, that is found in dividing cells of many mammals and amphibians, but absent from quiescent cells (Ford et al., 1988). Understanding the mechanism behind the difference in *CfrHlc113* expression between the in vivo tissues and the in vitro cells will help us to understand the biological roles of this RNA helicase.

## ACKNOWLEDGMENTS

Anti-*Rst(1)JH* antibody was a gift from Dr. Thomas Wilson's lab. We thank Mrs. Karen Jamieson for her editorial assistance. This work was supported in part by Canadian Biotechnology Strategy Fund and Genome Canada through the Ontario Genomics Institute.



## LITERATURE CITED

- Abramson RD, Dever TE, Lawson TG, Ray BK, Thach RE, Merrick WC. 1987. The ATP-dependent interaction of eukaryotic initiation factors with mRNA. *J Biol Chem* 262:3826–3832.
- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. *J Mol Biol* 215:403–410.
- Ashok M, Turner C, Wilson TG. 1998. Insect juvenile hormone resistance gene homology with the bHLH-PAS family of transcriptional regulators. *Proc Natl Acad Sci USA* 95:2761–2766.
- Aubourg S, Kreis M, Lecharny A. 1999. The DEAD box RNA helicase family in *Arabidopsis thaliana*. *Nucleic Acids Res* 27:628–636.
- Billmoria SL, Sohi SS. 1977. Development of an attached strain from a continuous insect cell line. *In Vitro* 13:461–466.
- Birney E, Kumar S, Krainer AR. 1993. Analysis of the RNA-recognition motif and RS and RGG domains: conservation in metazoan pre-mRNA splicing factors. *Nucleic Acids Res* 21:5803–5816.
- Dalbadie-McFarland G, Abelson J. 1990. PRP5: a helicase-like protein required for mRNA splicing in yeast. *Proc Natl Acad Sci USA* 87:4236–4240.
- Daugeron MC, Kressler D, Linder P. 2001. Dbp9p, a putative ATP-dependent RNA helicase involved in 60S-ribosomal-subunit biogenesis, functionally interacts with Dbp6p. *RNA* 7:1317–1334.
- Dayyeh BKA, Quan TK, Castro M, Ruby SW. 2002. Probing interactions between the U2 small nuclear ribonucleoprotein and the DEAD-box protein, Prp5. *J Biol Chem* 277:20221–20233.
- de la Cruz J, Kressler D, Linder P. 1999. Unwinding RNA in *Saccharomyces cerevisiae*: DEAD-box proteins and related families. *Trends Biochem Sci* 24:192–198.
- de Valoir T, Tucher MA, Belikoff EJ, Camp LA, Bolduc C, Beckingham K. 1991. A second maternally expressed *Drosophila* gene encodes a putative RNA helicase of the “DEAD box” family. *Proc Natl Acad Sci USA* 88:2113–2117.
- Dorn R, Morawietz H, Reuter G, Saumweber H. 1993. Identification of an essential *Drosophila* gene that is homologous to the translation initiation factor eIF-4A of yeast and mouse. *Mol Gen Genet* 237:233–240.
- Eberl DF, Lorenz LJ, Melnick MB, Sood V, Lasko P, Perrimon N. 1997. A new enhancer of position-effect variegation in *Drosophila melanogaster* encodes a putative RNA helicase that binds chromosomes and is regulated by the cell cycle. *Genetics* 146:951–963.
- Feng QL, Ladd TR, Tomkins BL, Sundaram M, Sohi SS, Retnakaran, A, Davey KG, Palli SR. 1999. Spruce budworm (*Choristoneura fumiferana*) juvenile hormone esterase: hormonal regulation, developmental expression and cDNA cloning. *Mol Cell Endocrinol* 148:95–108.
- Ford MJ, Anton IA, Lane DP. 1988. Nuclear protein with se-

- quence homology to translation initiation factor eIF-4A. *Nature* 332:736–738.
- Fu XD, Maniatis T. 1990. Factor required for mammalian spliceosome assembly is localized to discrete regions in the nucleus. *Nature* 343:437–441.
- Gibson TJ, Thompson JD. 1994. Detection of dsRNA-binding domains in RNA helicase A and *Drosophila* maleless: implications for monomeric RNA helicases. *Nucleic Acids Res* 22:2552–2556.
- Grace TDC. 1962. Establishment of four strains of cells from insect tissues grown in vitro. *Nature* 195:788–789.
- Hay B, Jan LY, Jan YN. 1988. A protein component of *Drosophila* polar granules is encoded by *vasa* and has extensive sequence similarity to ATP-dependent helicases. *Cell* 55:577–587.
- Higgins DG, Sharp PM. 1988. CLUSTAL: a package for performing multiple sequence alignment on a microcomputer. *Gene* 73:237–244.
- Hloch P, Schiedner G, Stahl H. 1990. Complete cDNA sequence of the human p68 protein. *Nucleic Acids Res* 18:3045.
- Horton P, Nakai K. 1997. Better prediction of protein cellular localization sites with the k nearest neighbors classifier. *Proc Int Conf Intell Syst Mol Biol* 5:147–152.
- Jones PG, Mitta M, Kim Y, Jiang W, Inouye M. 1996. Cold shock induces a major ribosomal-associated protein that unwinds double-stranded RNA in *Escherichia coli*. *Proc Natl Acad Sci USA* 93:76–80.
- Kuroda MI, Kernan MJ, Kreber R, Ganetzky B, Baker BS. 1991. The maleless protein associates with the X chromosome to regulate dosage compensation in *Drosophila*. *Cell* 66:935–947.
- Lasko PF, Ashburner M. 1988. The product of the *Drosophila* gene *vasa* is very similar to eukaryotic initiation factor-4A. *Nature* 335:611–617.
- Lüking A, Stahl U, Schmidt U. 1998. The protein family of RNA helicases. *Crit Rev Biochem Mol Biol* 33:259–296.
- McMorran A. 1965. A synthetic diet for the spruce budworm, *Choristoneura fumiferana* (Lepidoptera: Tortricidae). *Can Entomol* 97:58–62.
- Nakai K, Kanehisa M. 1992. A knowledge base for predicting protein localization sites in eukaryotic cells. *Genomics* 14:897–911.
- Nakao H. 1999. Isolation and characterization of a *Bombyx* vasa-like gene. *Dev Genes Evol* 209:312–316.
- O'Day CL, Dalbadie-McFarland G, Abelson J. 1996. The *Saccharomyces cerevisiae* Prp5 protein has RNA-dependent ATPase activity with specificity for U2 small nuclear RNA. *J Biol Chem* 271:33261–22367.
- Pursley S, Ashok M, Wilson TG. 2000. Intracellular localization and tissue specificity of the Methoprene-tolerant (*Met*) gene product in *Drosophila melanogaster*. *Insect Biochem Mol Biol* 30:839–845.
- Reinhardt A, Hubbard T. 1998. Using neural networks for prediction of the subcellular location of proteins. *Nucleic Acids Res* 26: 2230–2236.
- Schmid SR, Linder P. 1992. D-E-A-D protein family of putative RNA helicases. *Mol Microbiol* 6:283–292.
- Sohi SS, Lalouette W, MacDonld JA, Gringorten JL, Budau CB. 1993. Establishment of continuous midgut cell lines of spruce budworm (Lepidoptera: Tortricidae). *In vitro Cell Develop Biol* 29A (part II), 56A.
- Sukegawa J, Blobel G. 1995. A putative mammalian RNA helicase with an Arginine-Serine-rich domain colocalizes with a splicing factor. *J Biol Chem* 270:15702–15706.
- Watkins S. 1996. Immunohistochemistry. In: Ausubel FM, Brent R, Kingston RE, Moore DD, Seidman JG, Smith JA, Strunl K, editors. *Current protocols in molecular biology*. New York: John Wiley & Sons, Inc. p 14.0.1–14.9.17.
- Will CL, Urlaub H, Achsel T, Gentzel M, Wilm M, 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.
- Zhang DY, Zheng SC, Zheng YP, Ladd TR, Pang ASD, Davey KG, Krell PJ, Arif BM, Retnakaran A, Feng QL. 2004. An ecdysone-inducible putative “DEAD box” RNA helicase in the spruce budworm (*Choristoneura fumiferana*). *Insect Biochem Mol Biol* 34:273–281.
- Zirwes RF, Eilbracht J, Kneissel S, Schmidt-Zachmann MS. 2000. A novel helicase-type protein in the nucleolus: protein NOH61. *Mol Biol Cell* 11:1153–1167.