USE OF A HSPC117 MOLECULE AS RNA LIGASE

The present invention relates to the field of cell and molecular biology tools, in particular to RNA ligases, and methods for using and suppressing an RNA ligase for analysis and therapeutics.

Natural enzymes that ligate RNA or DNA generally join a nucleic acid molecule having a phosphoryl group at the 5' position to a second nucleic acid molecule having a hydroxyl group at the 3' position. The phosphate on the 5' end is usually provided by ATP in an energy transferring step.

Transfer RNAs (tRNAs) are essential adaptor molecules for the translation of messenger RNA (mRNA) into proteins. Similar to other RNA molecules, precursor tRNA transcripts (pre-tRNAs) are subjected to extensive posttranscriptional processing before they are to fulfil their biological functions. In addition to the removal of 5′-leader and 3′-trailer sequences, extensive base and sugar modifications and template-independent addition of nucleotides, some tRNAs have to undergo excision of an intervening sequence. Removal of tRNA introns is accomplished by a splicing process that differs from canonical spliceosome-dependent processing of mRNA. tRNA splicing rather requires a specialized endonuclease excising the intron and a ligase to join the exon halves (Fig. 1A).

The WO 2004/087884 A2 describes methods to screen for small organic molecules involved in tRNA processing. Pascal et al., Current Opinion in Structural Biology 18 (1) (2008): 96-105, relates to differences in PNA and RNA ligases. Kato et al., Journal of Molecular Biology 239 (5) (2003): 903-911, describe a crystal structure of a RNA ligase of Thermus thermophilus. Wang et al., RNA 11 (6) (2005): 966-975 performed a structure-function analysis of yeast tRNA ligase. Okada et al., PROTEINS 63 (4) (2006): 1119-1122, provide a crystal structure of a RtcB homologuous protein from Pyrococcus horikoshii, a RNA cyclase.

Although the presence of introns in tRNA genes seems to be common to all domains of life, evolution of the splicing machinery has apparently diverged at the ligation step. Two major ligation

pathways have been described which can be attributed to distinct kingdoms of life (Abelson et al., 1998). The fungal and plant clades use a common ligation mechanism catalyzed by single multifunctional polypeptides that are homologous to bacteriophage T4 RNA ligase 1. This pathway requires the action of cyclic phosphodiesterase and polynucleotide kinase activities to prepare the exon termini for subsequent ligation. As a consequence, an exogenous phosphate originating from a nucleoside triphosphate (NTP) donor is incorporated into the mature tRNA (Fig. 1A, upper branch). In contrast, the animal and archaebacteria clades may ligate tRNA exons by directly joining the 2′,3′-cyclic phosphate and 5′-hydroxyl (5′-OH) termini left after the endonuclease reaction (Laski et al., 1983). This ligase reaction (Fig. 1A, lower branch) is dependent on 2′,3′-cyclic-phosphate terminated RNA (RNA>p) and results in incorporation of the precursor-derived 2′,3′-cyclic phosphate into the splice junction of mature tRNA as a 3′,5′-phosphodiester (Filipowicz and Shatkin, 1983). Key biochemical aspects of this RNA>p ligase reaction have been elucidated. However, despite many biochemical, bioinformatics and genetic efforts, no suitable RNA ligase has been identified since the tRNA splicing pathway was initially postulated.

It is therefore a goal of the present invention to provide an RNA ligase that is at least also capable of using 2', 3'-cyclic phosphate terminated RNA as substrate ("RNA>p ligase"). This goal is achieved by the subject matter of the claims.

In particular, the present invention relates to the use of a HSPC117 molecule as RNA ligase as a molecular biology tool and in therapeutics. HSPC117 has been sequenced (e.g. Genbank ACC NO: NP_055121 or CAG33456), and located at chromosome 22 orf 28 ("C22ORF28"). HSPC117 is the human homolog of the bacterial/archaeal RtcB gene family characterised by a highly conserved domain of unknown function (UPF0027) and a unique protein fold. UPF0027 proteins form a cluster of orthologous genes (KOG3833) with no detectable representatives in the plant or fungal model organisms. This phyletic distribution is highly reminiscent of the exclusive occurrence of RNA>p ligase activity in animals and archaea. HSPC117 is also referred herein as HSPC117/C22ORF28 or

RtcB/HSPC117. As used herein the expression "HSPC117 molecule" refers to any homologous or orthologous molecule in this cluster which has now been identified to catalyze an RNA ligase reaction. Example sequences of such "HSPC117 molecules" are given in Fig. 6 as SEQ ID NO: 1 to SEQ ID NO: 11. All HSPC117 molecules have been found to contain the catalytic cysteine residue corresponding to C122 of SEQ ID NO: 1.

HSPC117 molecules and sequences have been further described, e.g. in the US 2007/0204352 A1 (especially SEQ ID NOs: 15, 16, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78), however without prior recognition of the inventive uses. The US 2007/0204352 A1 relates to a screening of genes potentially involved in the aggregation of alpha-synclein. At paragraph [50], the US 2007/0204352 A1 provides background on the remarkable degree of evolutionary conservation of HSPC117 genes, which proteins can be used in a method of the present invention, in the knowledge of a new function of these HSPC117 molecules.

The inventive RNA ligase can be used to catalyze the transfer of a first RNA to a second RNA. The ends of both RNA can be connected by the ligase. This connection is usually a covalent connection of a phosphor diester bond between both RNA. In particular, one RNA may comprise a 3' phosphate, in particular in form of a 2', 3'-cyclic phosphate, and the other may comprise a 5'-OH terminus. The ability to form a connection between such ends in particular by using a 5'-OH terminus is a unique feature of the inventive HSPC117 molecule.

In general, the RNA ligation can be an inter- or intra-strand ligation. Two separate RNA strands may be connected on the 3' and 5' end, respectively. Furthermore, in an intra-strand ligation, the 5' and 3' end of one RNA molecule is connected.

In a further embodiment of the present invention the RNA is double-stranded. In particular, the first and/or second RNA molecule connected by the inventive RNA ligase may comprise a double-stranded section or is fully double-stranded or alternatively single-stranded. In particular preferred, the 3' end mentioned above as well as the 5' terminus of the other RNA end,

which are connected by the RNA ligase reaction, may be double-stranded. Further portions of the RNA can also be single-stranded, in particular of RNA splicing there is usually a single-stranded 3' overhang of a pre-tRNA. Also, the 5' and/or 3' ends, which are connected by the RNA ligase reaction may be single-stranded - as is usually the case in pre-tRNA processing. Double-strandedness may be a base pairing between the first and the second RNA molecules, or alternatively may be base pairing to further RNA strands.

In particular preferred embodiments the present inventive HSPC117 molecule is used for RNA splicing. In an RNA splicing reaction an intron section is removed between two exons, which are connected by the inventive RNA ligase. A typical splicing reaction is the reaction of an exon1-intron-exon2 sequence to exon1-exon2. Other splicing reactions may remove several introns and, optionally also exons between these intron sections.

The RNA that is connected by the inventive use of the HSPC117 molecule as RNA ligase may be of any length. Example RNA lenghts are 2 to 3000 nucleotides or base pairs in length. In special embodiments, the first RNA or the second RNA may be more than 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 40, 45, 50, 60, 70, 80, 90 or more than 100 nucleotides or base pairs in length. Alternatively or in addition thereto, the RNA, either the first RNA or second RNA or both, may be up to 3000, 2000, 1500, 1200, 1000, 900, 800, 700, 600, 500, 400, 300, 200, 100, 80, 70, 60 or up to 50 nucleotides or base pairs in length.

The inventive HSPC117 molecule may be of anyone of SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11. The sequences provide example HSPC117 molecules of H.sapiens, M.musculus, X.leavis, S.purpuratus, D.melanogaster, C.elegans, C.reinhardtii, M.jannaschii, P.horikoshii, T.thermophilus, or E.coli, respectively. Further HSPC117 sequences or sequence variants are disclosed in SEQ ID NOs: 12 to 23, providing further nucleic acid and amino acid sequences from homo sapiens, C. elegans, Drosophila, Danio rerio, bovine, mouse and rat. The inventive HSPC117 molecule may be ob-

tained from any of these organisms. In preferred embodiments the inventive HSPC117 molecule is of an animal or archaea, in particular of a mammal, such as a primate, including human, or rodent, in particular mouse or rat.

The inventive HSPC117 molecule may be further modified by one or more amino acid substitution or deletion. Furthermore, the inventive HSPC117 molecule may be expressed as part of a fusion protein and may comprise further additional amino acids or polypeptide sequences. In particular preferred, the inventive HSPC117 molecule has a sequence identity of at least 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90% or at least 95% to any one of SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23. In preferred embodiments the sequence identity is related to SEQ ID NO: 1. Sequence identities are usually calculated over the whole length sequences of SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23. Of course such a HSPC117 molecule variant maintains the RNA ligase activity as mentioned above as can be easily determined by standard assays as shown in the example section herein. In particular of importance is that the HSPC117 molecule maintains catalytically important residues, such as cysteine 122 of SEQ ID NO: 1. Variants of the inventive HSPC117 molecules are e.g. described in US 2007/0204352 A1 (especially SEQ ID NOs: 15, 16, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78 thereof), incorporated herein by reference as SEQ ID NOs: 12 to 23, and can be used for the inventive purposes.

In the case of amino acid substitution, in preferred embodiments at least 50%, at least 60%, at least 70%, at least 80%, at least 90% or at least 95% of the substitutions are conserved amino acid substitutions. Conserved substitutions are mutations within an amino acid group. Amino acids are usually grouped according to their polarity, charge and/or size. The following groups are noteworthy: basic amino acids: arginine, histidine, lysine;

acidic amino acids: aspartic acid, glutamic acid; polar amino acids: asparagine, glutamine; small amino acids: alanine, serine, threonine, methionine, glycine; aromatic amino acids: phenylalanine, tryptophan, tyrosine, histidine; hydrophobic amino acids: leucine, isoleucine, valine. Cysteine is a special case, as it may usually be conservatively substituted with serine and any other polar uncharged sidechain and vice versa. Glycine may be used as substituent for any amino acid. Glycin can be substituted usually by a small sidechain such as by alanine, serine, threonine. Proline may be usually substituted, or used as substituent for glycin.

In a further aspect, the present invention relates to the method of ligating at least two RNA molecules - as e.g. described above - comprising using a HSPC117 molecule as described above. Herein, the expressions "use ... as RNA ligase" and "method of ligating RNA molecules" are used interchangeably.

In preferred embodiments the inventive use or method may comprise contacting at least two RNA molecules with the HSPC117 molecule in a cell. The invention also relates to the use of recombinant HSPC117. Recombinant HSPC117 (including any homologs or orthologs as mentioned above) can be readily obtained by expression of genetic constructs comprising one or more HSPC117 DNA sequences operable linked to regulatory DNA sequences (which may be heterologuous regulatory sequences), such as promoters or enhancers, in host cells. Example host cells are bacterial, archaea, fungal (including yeast), plant or animal (including insect or mammalian) cells. In such constructs, the design of which is described in common laboratory manuals and is routine to a skilled artician, the regulatory sequences may be operably linked to a polynucleotide encoding the HSPC117 molecule or an active variant thereof having RNA ligase activity.

The inventive HSPC117 molecule may be used *in vivo* such as in a cell, e.g. artificially provided therein or recombinantly expressed in the cell. Two RNA molecules may be ligated in said cell according to an embodiment of the present invention. The cell may be any cell as described above, preferably a non-human cell or an isolated human cell.

In a further embodiment the RNA molecules may be contacted with a HSPC117 molecule *in vitro* or *in situ* such as e.g. including outside a cell or in a cell free solution. With the inventive HSPC117 molecule it is possible to ligate RNA molecules in an isolated fashion, *ex vivo*.

According to the present invention it was found that HSPC117 is the catalytically active protein that may naturally also be contained in a complex *in vivo*. Therefore, according to a further embodiment of the present invention the inventive HSPC117 molecule is also provided in such a complex. The complex may be e.g. of spliceosomal particles such as SF3B particles that can be isolated from nuclear HeLa cell extracts. Other complex members, in particular with regard to other HSPC117 molecules of other organisms or variants thereof may be isolated from cellular extracts of the respective organisms.

In particular preferred embodiments the inventive complex may comprise DEAD box helicase DDX1, a FAM98B molecule, a CGI-99 molecule, ASW or any combination thereof, in particular preferred a combination of the DDX1 and FAM98B molecule. Further complex members that can be comprised in the complex of HSPC117 molecules may be any one of tables 1 and 2 shown in example 5.

According to the present invention the HSPC117 molecule may be provided per se. Alternatively, HSPC117 molecules may be used or provided as a component of a kit.

Thus, in a further aspect the invention relates to a kit that contains HSPC117 molecule. The kit may further comprise a reaction buffer for the RNA ligase comprising buffer components or one or more metal ions selected from Mg^{2+}, Mn^{2+}, Ni^{2+} or mixtures thereof. In preferred embodiments the metal ions are included in an amount for use in a final concentration range of ca 0.1-20 mM, preferably 1-10mM, in particular preferred 2-5mM.

Besides the above-mentioned metal ions, the buffer of the kit may contain the usual buffer components that are well known in the art. Such buffers may e.g. include phosphate, Hepes, Tris-

HCl components. Preferably the buffer is in the range of physiological pH values of e.g. pH 6 to pH 9, preferably pH 7-8, especially preferred about pH 7.4. The buffer may comprise tonic substances or salts ranging from about 10-200mM KCl or NaCl. Furthermore, the buffer may contain non-ionic tonicity substances such as glycerol.

In the form of a test kit, the kit may further comprise an RNA molecule that is a substrate of the inventive HSPC117 molecule, especially a RNA molecule with a 2', 3' cyclic phosphate. This RNA molecule may e.g. further comprise a label such as a radioactive label to detect the RNA molecule before or after the RNA ligase reaction. Such a kit is useful for all types of reactions and to monitor RNA processing or hybridisation. The inventive HSPC117 molecule or kit may be especially used for RNA ligation or splicing studies.

The present invention in a further aspect relates to a transgenic cell comprising an exogenously expressed HSPC117 molecule. The cell may be a cell line or comprised in an animal model, in particular a non-human animal model. A cell line may be also a human cell line that stably expresses HSPC117 molecules.

Stable expression of the exogenously expressed HSPC117 molecule is achieved by inserting a HSPC117 DNA, under the control of a promoter, preferably an inducible promoter, into the cell. In certain embodiments this DNA can be inserted in the genome of the cell, which can be achieved by conventional methods such as commercially available systems like the tetracycline-inducible system such as the t-REx system (invitrogen). Such cells are useful in combination with RNA that can be ligated, especially RNA with 2', 3' cyclic phosphate or 5'-OH to ligate the RNA molecules.

The present invention further relates to method of reducing RNA ligase activity, in particular RNA>p ligase activity, in a cell comprising inhibiting a HSPC117 molecule in a cell, preferably by knock-out or RNAi. RNA>p ligase activity, as mentioned above, relates to RNA ligase reactions using 2', 3'-cyclic phosphate terminated RNA as substrate. Such a method can be used to reduce

tRNA production or processing in said cell. A reduction of HSPC117 can be achieved by administering a ligand to HSPC117 that binds, segregates or generally inactivates HSPC117 in said cell or by inhibiting HSPC117 expression. Such a binding inhibitor is e.g. a HSPC117 antibody, which is e.g. commercially available. A "HSPC117-antibody" includes any functional equivalents and derivatives thereof, including antibody fragments such as Fab, $F(ab)_2$, Fv, or single chain antibodies (scAb) that binds HSPC117. In preferred embodiments the inhibition is achieved by reducing expression of an HSPC117 molecule, preferably an endogenous HSPC117 molecule, in said cell. A suitable inhibitor to reduce HSPC117 expression is a HSPC siRNA molecule to induce RNAi.

Preferred methods of inhibiting HSPC117 expression are knock-out or RNAi. For a knock-out, a genomic HSPC117 is modified to present expression, transcription or translation of a functional HSPC117 molecule. Such modifications may include large stretch deletion of e.g. up to 200 or more nucleotides or selective modifications (deletions or substitutions) in the catalytic centre. E.g. a modification in the catalytic C122 according to the human HSPC117 sequence of SEQ ID NO:1 is sufficient to prevent expression of a functional molecule. Of course the skilled man in the art can readily select alternative modifications, which are within the routine ability of a molecular cell biologist.

A further preferred method is RNAi (RNA interference). For antagonizing cellular HSPC117 expression preferably siRNA molecules are administered to reduce the expression and function. RNA interference is a mechanism to suppress gene expression in a sequence specific manner. RNA interference is highly effective methodology for suppression of specific gene function in eukaryotic cells. When applied to cells and organisms, RNAi entails the degradation of target mRNA upon transfection of short interfering RNA (siRNA) oligos or short-hairpin RNA (shRNA) encoding vectors. Various methods of RNAi have been described and are generally known for altering gene expression in plant cells, drosophila and human melanoma cells as is described for example in US 2002/0162126 and US 2002/0173478. The siRNA for use in the methods and compositions of the invention are selected to target

a HSPC117 molecule. In this manner they are targeted to various RNAs or portions thereof corresponding to the HSPC117 gene. It is understood by one of skill in the art that the siRNA as herein described may also include altered siRNA that is a hybrid DNA/RNA construct or any equivalent thereof, double-stranded RNA, microRNA (miRNA), as well as siRNA forms such as siRNA duplications, small hairpin RNA (shRNA) in viral and non-viral vectors and siRNA or shRNA in carriers.

In a further embodiment the invention relates to a HSPC117 knock-out cell or cell with reduced or inhibited endogenous HSPC117 expression.

Such cell lines can be further used in RNA ligation or splicing studies, i.e. to study the function of RNA ligation. Reduction of an endogenous HSPC117 expression has also the benefit that there is no background activity of the inventive HSPC117 ligase that mediates the conversion of 2', 3' cyclic phosphate onto 5'-OH RNA molecules. In combination with a transgenic cell wherein HSPC117 is under control of an inducible promoter this allows specific on/off studies of the RNA ligase and is a useful tool to control the ligase activity, be it only for splicing studies or as a cellular biochemical engineering tool. Therefore, in a preferred embodiment, the invention relates to a HSPC117 knock-out cell that does not express endogenous HSPC117 but is in addition exogenously transfected with an HSPC117 molecule under control of an inducible promoter. In preferred embodiments the cell is a mammal cell, especially preferred a cell of a primate, in particular of a human or of a rodent such as a mouse cell. These cells, including cells with increased or decreased HSPC117 expression as described above, can be used for RNA ligation or splicing studies.

In a further aspect the present invention relates to the treatment of diseases with abnormal tRNA processing or diseases dependent on (increased) tRNA processing. In particular embodiments the invention provides the use of a HSPC117 molecule as an RNA ligase or a method of inhibiting a HSPC117 molecule, with the proviso that methods for treatment of the human or animal body by therapy are excluded, or the use of a HSPC117 molecule

or HSPC117 inhibitor for use as medicament. A HSPC117 inhibitor is any molecule that reduces HSPC117 activity or expression as described above, preferably a HSPC117 antibody or HSPC117 siRNA.

HSPC117 molecule inhibition may have therapeutic effect in several diseases. Such diseases include proliferative diseases, in particular cancer. By reducing tRNA processing the proliferative activity can be greatly decreased resulting in reduced cell growth. Therefore the present invention provides a method of reducing tumor cell growth comprising administering an HSPC117 molecule inhibitor to said cell. It is known that tumor cells have abnormally high rates of Polymerase (Pol) III transcription(Marshall & White, 2008). Since Pol III synthesizes tRNAs, targeting the tRNA ligase will turn (high) tRNA production rate-limiting in cancer cells. The importance of tRNA splicing components for proliferation is e.g. disclosed in the WO2004/087884 A2 (incorporated herein by reference).

In a further aspect the present invention provides the treatment of or a disease or infection dependent on host polymerases, such as hepatitis delta virus infection comprising administering an HSPC117 molecule inhibitor to said cell. The human hepatitis delta virus is the only animal virus known to replicate its RNA genome using a host polymerase. Host factors involved in the replication of the virus are elusive. A ligase-host factor in circularizing the viral genome during replication is implicated (Reid & Lazinski, 2000). In a combined proteomic-RNAi screen identified more than 100 proteins associated to the hepatitis delta antigens. A portion of the identified proteins has roles in RNA metabolism, and one of those is HSPC117. Combined this shows that HSPC117 is a decisive target for treating a hepatitis delta virus infection.

Furthermore the present invention relates to a method of treating a disease in a subject associated with dysfunctional tRNA splicing, in particular being deficient in a tRNA ligation by RNA>p, preferably pontocerebellar hypoplasia, comprising administering a HSPC117 molecule to said subject. A link has been established between the tRNA splicing pathway and Pontocerebellar hypoplasia. This disease belongs to a group of degenerative au-

tosomal recessive disorders with prenatal onset, atrophy or hypoplasia of the cerebellum and other motor impairments. Mechanistically these diseases are associated with aberrant removal of introns and ligation of exons during tRNA splicing. Therefore administration of a functional HSPC117 molecule can restore normal splicing and intron removal and treat the disease, while it is also well known, but not understood at a molecular level, that tRNA metabolism has a special impact on brain function.

Cells respond to oxidative stress by secreting Angiogenin, a factor that displays ribonuclease activity besides its known role in angiogenesis. Angiogenin cleaves mature tRNAs at the anticodon loop, thereby generating tRNA pieces known as tiRNAs, for tRNA-derived stress-induced RNAs. tiRNA accumulation impairs protein synthesis and is therefore detrimental to cell health and function. Inactivating the human tRNA ligase HSPC117 leads to an increase in tiRNAs in culture cells. Increased HSPC117 reverts Angiogenin cleavage and reduces tiRNA levels. HSPC117 therefore may have a distinct role in re-ligating Angiogenin-cleaved tRNA. This angiogenin reaction cannot be reverted upon inhibition of HSPC117 molecule. Thus, the present invention also relates to a method of modulating tiRNA amounts in a cell, such as by increasing or decreasing HSPC117 activity in a cell.

In a further aspect, the present invention provides a pharmaceutical composition comprising a HSPC117 molecule expressing nucleic acid, preferably in form of a expression vector, or a HSPC117 molecule inhibitor, preferably an antibody or siRNA or variant thereof as described above. Such a composition can be a ready to use composition, e.g. for the treatment of any disease described above. Pharmaceutical compositions or formulations for therapeutic or prophylactic use may comprise a pharmaceutically acceptable diluent, carrier, solubilizer, emulsifier and/or preservative. The invention also provides for pharmaceutical compositions comprising a therapeutically effective amount of an HSPC117 inhibitor or expression nucleic acid. The term "therapeutically effective amount" means an amount which provides a therapeutic effect for a specified condition and route of administration. The composition may be in a liquid or lyophilized form and comprises a diluent (Tris, acetate or phosphate buff-

ers) having various pH values and ionic strengths, solubilizer such as Tween or Polysorbate, carriers such as human serum albumin or gelatin, preservatives such as thimerosal or benzyl alcohol, and antioxidants such as ascrobic acid or sodium metabisulfite. Selection of a particular composition will depend upon a number of factors, including the condition being treated, the route of administration and the pharmacokinetic parameters desired. Nucleic acids and siRNA formulations are preferably administered in liposome formulations. Compositions of the invention may be administered by injection, either subcutaneous, intravenous or intramuscular, or by oral, nasal, pulmonary or rectal administration. The route of administration eventually chosen will depend upon a number of factors and may be ascertained by one skilled in the art.

The present invention will be now explained more in detail with reference to the figures and examples, without being limited thereto.

Figures

Fig. 1. RNA ligation mechanisms and identification of the RNA>p ligase HSPC117. (A) Scheme illustrating mechanistic differences in described RNA ligation pathways (B) [5′-^{32}P]-pCp-radiolabeled 5′-OH, 3′-P RNA oligonucleotides (depicted in grey, asterisk marks position of radiolabel) were incubated with or without AP and annealed to RNA strands (depicted in black) incubated with or without T4 Pnk in presence of ATP. The obtained RNA duplexes were used as substrates for inter-strand ligation in HeLa cell extracts. Aliquots of ligation reactions were withdrawn at indicated time points and analyzed by denaturing gel electrophoresis. (C) Fractionation scheme established for partial purification of RNA>p ligase from HeLa extracts. Throughout the purification inter-strand ligation of [5′-^{32}P]-pCp radiolabeled 5′-OH, 3′-P dsRNA was used to monitor RNA>p ligase activity. (D) Dilution series of protein extracts prepared from HSPC117-siRNA-transfected cells and control-siRNA-transfected cells were assayed for inter-strand ligation. Numbers above lanes indicate extract dilution. (E) The same extracts were assayed for processing of [α-^{32}P]-GTP-radiolabeled pre-tRNA in time course experiments. (F) Reduction of HSPC117 levels in extracts was con-

firmed by Western Blot.

Fig. 2. Affinity purification of c-myc-HSPC117 from stably transfected HeLa cell lines yields an RNA>p ligase complex. (A) IPs of WT or C122A c-myc-HSPC117 were incubated with [α-^{32}P]-GTP-radiolabeled tRNA exon halves. An IP prepared from a non-expressing clone were used as negative control. (B) Specificity of the affinity purification and presence of equal amounts of WT and mutant c-myc-HSPC117 in compared IPs was confirmed by Western blot. (C) tRNA exon halves were incubated with Clp1 or T4 Pnk in presence or absence of ATP. RNA was recovered and used as a substrate in incubations with immunopurified c-myc-HSPC117. (D) Scheme illustrating the course of procedures used for nearest neighbor analysis of splice junctions in mature tRNA and circular intron (Np, nucleoside 3'-monophosphate; pN, nucleoside 5'-monophosphate). (E) RNAse T1 fragments derived from [α-^{32}P]-UTP-radiolabeled mature tRNA generated either by T4 Pnk/Rnl1 or affinity purified c-myc-HSPC117 were resolved by denaturing gel electrophoresis. 6-mers, 7-mers and 8-mers were isolated from the gel, digested by RNAse T2 and analyzed by TLC in solvent D. (F) Circular, [α-^{32}P]-ATP-radiolabeled intron generated either by T4 Pnk/Rnl1 or affinity purified c-myc-HSPC117 was isolated from gels, digested with RNAse P1 and analyzed by TLC in solvent C.

Fig. 3. Active RNA>p ligase complex co-purifies with SF3B complexes. (A) Affinity protocol established for purification of SF3B-complexes from HeLa nuclear extracts. After depletion of U2 by immobilized monoclonal antibody H20 recognizing the trimethylguanosine ($m_3^{2,2,7}$G) mRNA cap structure, HSPC117- and SF3B-complexes were bound by immobilized monoclonal SF3B155-directed antibody 13E12. (B) Upon elution from the anti-SF3B155 13E12 column, SF3B and RNA>p ligase complexes were separated by size exclusion on Superose 6. Fractions were analyzed by SDS-PAGE and proteins stained with Coomassie blue. (C) Fractions eluted from the Superose 6 column were assayed for RNA>p ligase activity using tRNA exon halves as substrate. RNA recovered from reactions was analyzed by denaturing gel electrophoresis. (D) The RNA>p ligase peak fraction was analyzed by SDS-PAGE, protein bands were visualized by staining with Coomassie blue. Marked bands were identified by mass spectrometry. (E) Loss of inter-strand

ligation and tRNA maturation activities upon silencing of HSPC117 but not the other components of RNA>p ligase complex. Cells were treated with siRNAs targeting C22ORF28/HSPC117, DDX1, C14ORF166/CGI-99, FAM98B, C2ORF49/ASW and EGFP as a control. Extracts were prepared from these cells and assayed for inter-strand ligation of [5′-^{32}P]-pCp-radiolabeled dsRNA or (F) processing of [α-^{32}P]-GTP-radiolabeled pre-tRNA. (G) Efficient depletion of C22ORF28/HSPC117, DDX1, C14ORF166/CGI-99, FAM98B and C2ORF49/ASW mRNAs was confirmed by quantitative PCR. Results represent mean and standard deviation of triplicate PCR reactions. (H) Western blot for HSPC117, DDX1, FAM98B and β-Actin as a loading control confirms efficient depletion of individual RNA>p complex members.

Fig. 4. Silencing of HSPC117 abolishes inter-strand ligation in living HeLa cells and partially impairs tRNA processing *in vivo*. (A) [5′-^{32}P]-pCp-radiolabeled dsRNA were transfected into HeLa cells pre-transfected with siRNAs targeting EGFP as a control gene, HSPC117 or RTCD1. Transfections were stopped at indicated time points. RNA was isolated and analyzed by denaturing gel electrophoresis. (B) Treatment scheme used for transfection of siRNAs and reporter constructs and induction of tagged tRNA transcription. (C) Schematic representation of elements present in tagged tRNA reporter constructs. (D) Tet-repressor expressing HeLa cells were cotransfected with siRNAs targeting the control gene EGFP, HSPC117 or TSEN2 and with the reporter construct pSTet-Ile expressing tagged pre-tRNA Ile. After induction, RNA was isolated at indicated time points and analyzed by Northern blot. (E) Quantification of multiple Northern blot experiments. Transfection of siRNAs and reporter construct were done in triplicates. RNA was recovered from cells, mature tRNA detected by Northern blot and quantitated by Phosphorimaging using ImageQuant. Error bars represent standard deviations.

Fig. 5. Genetic rescue of the *in vitro* RNA ligation defect in extracts depleted of HSPC117 by RNAi. (A) Stable mouse-BAC transgenic HeLa cell pools or wild type HeLa cells were transfected with siRNA targeting the non-conserved 3′-UTR of human HSPC117 or control siRNAs. Extracts prepared from these cells were assayed for inter-strand ligation with [5′-^{32}P]-pCp-

radiolabeled dsRNA. Numbers above lanes indicate relative amounts of ligation product. The signal in lane 1 was arbitrarily set to 1.0. (B) RNA was isolated from the same experiments in parallel and analyzed for levels of mouse and human HSPC117 by quantitative PCR. Expression levels of human HSPC117 in HeLa cells and human and murine HSPC117 in mouse-BAC transgenic HeLa cell pools were arbitrarily set to 100 %.

Fig. 6. Sequence alignment of HSPC117/RtcB proteins. HSPC117/RtcB proteins are widely spread in archaea, bacteria and animals, but not in plants and fungi. The asterisk indicates the position of the characterized C122A mutation in the presumed active site of human/murine HSPC117. The following sequence identifiers are used: H.sapiens: SEQ ID NO: 1, M.musculus: SEQ ID NO: 2, X.leavis: SEQ ID NO: 3, S.purpuratus: SEQ ID NO: 4, D.melanogaster: SEQ ID NO: 5, C.elegans: SEQ ID NO: 6, C.reinhardtii: SEQ ID NO: 7, M.jannaschii: SEQ ID NO: 8, P.horikoshii: SEQ ID NO: 9, T.thermophilus: SEQ ID NO: 10, E.coli: SEQ ID NO: 11.

Fig. 7. Detailed explanation of the protocol used for nearest neighbor analysis of the splice junction phosphate. (A) The depicted pre-tRNA is body-labeled with [α-^{32}P]-UTP. For simplicity only radiolabeled phosphates are shown. Cleavage of this pre-tRNA with recombinant splicing endonuclease of *Methanocaldococcus jannaschii* (MjTSEN) yields 5'-exon halves with a radiolabeled terminal 2',3'-cyclic phosphate. Ligation of the exon halves with c-myc-HSPC117 leads to formation of mature tRNA molecules containing the precursor-derived splice junction phosphate. During ligation of the same exon halves with a mixture of T4 Pnk and T4 Rnl1 the splice junction phosphate is replaced by exogenous phosphate provided by nucleoside triphosphate. As a consequence, the radioactive splice junction label is lost. Mature tRNA is isolated from both ligation reactions and digested with RNAse T1. The splice junction is now contained in a RNAse T1 7-mer which can be isolated by preparative denaturing gel electrophoresis. Complete digestion of the RNAse T1 7-mers with RNAse T2 releases the splice junction phosphate as a Guanosine 3'-monophosphate (Gp). Therefore, detection of radiolabeled Gp indicates incorporation of the precursor-derived, radiolabeled

terminal 2',3'-cyclic phosphate into the splice junction as a 3',5'-phosphodiester. (B) The depicted pre-tRNA is body-labeled with [α-^{32}P]-ATP. Cleavage of this precursor produces linear intron with a radiolabeled terminal 2',3'-cyclic phosphate. Ligation of this linear intron with c-myc-HSPC117 yields circularized intron containing the radiolabeled terminal phosphate. During ligation of linear intron with T4 Pnk and Rnl1 this phosphate is removed and an exongenous phosphate provided by nucleoside triphosphate is deposited at the 5'-OH of the first Uridine nucleotide of the linear intron. Circularized intron is isolated by preparative denaturing gel electrophoresis. During ligation with c-myc-HSPC117 the radiolabeled 2',3'-cyclic phosphate becomes the 5'-phosphate of the first Uridine nucleotide of the linear intron. Digestion of circularized intron with RNAse P1 therefore releases the junction phosphate as Uridine 5'-monophosphate (pU). Therefore, detection of radiolabeled pU indicates incorporation of the radiolabeled terminal 2',3'-cyclic phosphate of the linear intron into circular intron as a 3',5'-phosphodiester.

Fig. 8. Biochemical rescue of the *in vitro* RNA ligation defect in extracts depleted of HSPC117 by RNAi. Extracts of cells depleted of HSPC117 by RNAi were complemented by Superose 6 fractions of HSPC117-complex copurified with SF3B particles. Fractions were assayed for processing of [α-^{32}P]-GTP-radiolabeled pre-tRNA.

Fig. 9. Interfering with RNA>p ligase by specific inhibition with cruciform DNA. (A) [5'-^{32}P]-pCp-radiolabeled dsRNA was incubated with HeLa cell extracts in presence of a 2000-fold molar excess of cruciform (ScraPal) or control DNA (ScraPalΔ) duplexes over dsRNA substrate. Aliquots of reactions were withdrawn at indicated time points and analyzed by denaturing gel electrophoresis. (B) HeLa extracts were assayed for processing of [α-^{32}P]-GTP-radiolabeled pre-tRNA in presence of a 2000-fold molar excess of cruciform or control DNA duplexes over RNA substrate. Aliquots of reactions were withdrawn at indicated time points and analyzed by denaturing gel electrophoresis.

Fig. 10. Design of tagged tRNA Ile and validation of RNAi-

mediated depletion of HSPC117 and TSEN2 for *in vivo* tRNA processing experiments. Alignment of the pSTet insert encoding tagged pre-tRNA Ile^{TAT} with exon sequences derived from annotated human genomic tRNA Ile loci (Chan and Lowe, 2009). Capital letters in the bottom line indicate universally conserved residues, asterisks indicate mutated positions with respect to chr19/trna10-Ile^{TAT}.

Examples

Example 1: Identification of mammalian RNA>p ligase by a new biochemical strategy

A novel strategy to detect RNA>p ligase as applied potentially leading to the identification of a tRNA ligase. It was serendipitously discovered that 3′-phosphorylated (3′-P), 5′-OH double stranded RNA molecules (dsRNA) are covalently linked upon incubation with human cell extracts. In these extracts, 3′-P dsRNA is converted into 2′,3′-cyclic phosphate terminated dsRNA by human RNA terminal cyclase RTCD1. Therefore, we decided to use 3′-P dsRNA as a stable surrogate substrate for the elusive tRNA ligase. Inter-strand ligation requires a 3′-P single stranded RNA annealed to a 5′-OH complementary strand. Removal of the 3′-P by incubation with alkaline phosphatase (AP) or phosphorylation of the 5′-OH by incubation with bacteriophage T4 polynucleotide kinase (T4 Pnk) in presence of ATP or a combination of both inhibited inter-strand ligation (Fig. 1B). This result led to identify RNA>p ligase by classical activity guided protein chromatography. Monitoring inter-strand ligation, we were able to follow RNA>p ligase activity through four purification steps (Fig 1C). Unable to further fractionate inter-strand ligation activity, we identified proteins contained in the most enriched MonoQ fraction by in solution tryptic digest followed by tandem mass spectrometric analysis (MS). One out of 91 (Tab. S1) identified polypeptides, HSPC117/C22ORF28, appeared to be of particular interest for the following reasons. First, HSPC117 is the human homolog of the bacterial/archaeal RtcB gene family characterized by a highly conserved domain of unknown function (UPF0027) and a unique protein fold harboring a putative metal ion binding site. Interestingly, in *E. coli* RtcB together with RtcA, the RNA 3′-P terminal cyclase, resides within a σ54-regulated operon (Genschik et al., 1998). As a consequence,

RtcB/HSPC117 proteins have previously been predicted to have a function in RNA processing or modification (Galperin and Koonin, 2004). Second, UPF0027 proteins form a cluster of orthologous genes (KOG3833) with no detectable representatives in the plant and fungal model organisms *Arabidopsis thaliana*, *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe*. This phyletic distribution is highly reminiscent of the exclusive occurrence of RNA>p ligase activity in animals and archaea (Abelson et al., 1998).

Example 2: HSPC117 is required for inter-strand ligation and tRNA maturation

Therefore it was tested whether HSPC117 is indeed involved in inter-strand ligation and tRNA processing. HeLa cells were transfected with small interfering RNAs (siRNAs) targeting HSPC117 or EGFP as a control. Extracts were prepared from these cells and assayed for inter-strand ligation. Remarkably, depletion of HSPC117 by RNA interference (RNAi) affected inter-strand ligation (Fig. 1D, compare lanes 2-4 with lanes 5-7). 3'-P dsRNAs were merely used as surrogate substrates and RNA>p ligation is assumed to be required for splicing of tRNAs. Next, the impact of silencing HSPC117 on tRNA maturation was examined *in vitro*. Suitable tRNA precursor (pre-tRNA) transcripts are processed to mature tRNA in HeLa cytoplasmic extracts (Laski et al., 1983). Therefore, [α-^{32}P]-GTP-radiolabeled pre-tRNA transcripts were included with extracts depleted of HSPC117 by RNAi (Fig. 1E). Reduced levels of HSPC117 impaired formation of mature tRNA and led to a concomitant accumulation of exon halves, consistent with a biochemical function in ligation of tRNA exons. Efficient reduction of HSPC117 levels in extracts was verified by Western blot analysis (Fig. 1F). Specificity of siRNA-mediated depletion of HSPC117 was confirmed with a different set of siRNAs and by rescue of the RNAi phenotype (Fig. 5) by expression of an RNAi-resistant form of HSPC117 from a bacterial artificial chromosome (BAC) derived from mouse genomic DNA.

Example 3: HSPC117 is the catalytic component of mammalian RNA>p ligase

Next it was investigated whether HSPC117 is associated with RNA>p ligase activity. Therefore, stably transfected clonal HeLa

cell lines expressing c-myc-tagged murine HSPC117 were established. Based on multiple sequence alignments of UPF0027 proteins (Fig. 6) and guided by the published crystal structure of RtcB from *Pyrococcus horikoshii* (Okada et al., 2006) a stably transfected clonal cell line expressing the point mutant c-myc-HSPC117 C122A was generated. To detect tRNA ligase activity independent of RTCD1 and tRNA endonuclease, tRNA exon halves were prepared by cleaving a suitable [α-^{32}P]-GTP radiolabeled hybrid pre-tRNA (Englert, 2005) with recombinant splicing endonuclease from *Methanocaldococcus jannaschii* (MjTSEN). Affinity purification of c-myc-HSPC117 yielded an immunoprecipitate (IP) that was able to ligate tRNA exon halves. In contrast, the point mutant c-myc-HSPC117 C122A was inactive as an RNA ligase (Fig. 2A, compare lanes 6 and 7 with lane 8). Comparison of equal amounts of wild-type (WT) and C122A mutant c-myc-HSPC117 was confirmed by Western blot (Fig. 2B, compare lanes 2 and 3). It was therefore concluded that HSPC117 is the catalytic component of a tRNA ligase.

Example 4: HSPC117 catalyzes the direct ligation of 2′,3′-cyclic phosphate and 5′-OH RNA termini

RNA>p ligase is predicted to require a 5′-OH at the terminus of its substrates. Consistently, no ligase activity could be detected in c-myc-HSPC117 IPs when tRNA exon halves were preincubated with recombinant 5′-OH RNA kinase CLP1 in presence of ATP to convert the 5′-OH of the 3′-exon into 5′-P (Fig. 2C, compare lanes 2 and 3). In addition to its 5′-OH RNA kinase activity, T4 Pnk is known to harbor a 2′,3′-cyclic phosphodiesterase and a 3′-phosphatase activity. As expected for RNA>p ligase, we could not detect any ligase activity when we used tRNA exon halves preincubated with T4 Pnk in absence (to remove the 2′,3′-cyclic phosphate) or presence of ATP (to remove the 2′,3′-cyclic phosphate and to additionally convert 5′-OH into 5′-P) as substrates for ligation with c-myc-HSPC117 IP (Fig. 2C, lanes 4 and 5). Thus, ligation by the described RNA>p ligase is favourable on 5′-OH (see also Fig. 1B) and 2′,3′-cyclic phosphate.

The most characteristic feature of ligation of tRNA exon halves by RNA>p ligase is the generation of mature tRNAs that contain the precursor-derived splice junction phosphate in a canonical 5′,3′-phosphodiester bond (Fig. 1A, lower branch). To test

whether this is true for RNA ligation catalyzed by c-myc-HSPC117 IP a nearest neighbor analysis was performed of the splice junction phosphate (Fig. 2D). tRNA exon halves with radiolabeled 2′,3′-cyclic phosphate at the terminus of the 5′-exon half were prepared by cleaving [α-^{32}P]-UTP-radiolabeled pre-tRNA with recombinant splicing endonuclease (Fig. 7). These exon halves were ligated either by incubation with c-myc-HSPC117 IP or with a mixture of T4 Pnk and bacteriophage T4 RNA ligase 1 (T4 Rnl1) as a negative control. Mature tRNA was isolated from these reactions and digested by RNAse T1 which cleaves RNA 3′ of every guanosine nucleotide and processes its substrates into fragments terminated by guanosine 3′-monophosphate (Gp). Fragments ranging from 6-8 nucleotides in length were isolated, digested by RNAse T2 and resulting nucleotide 3′-monophosphates resolved by thin layer chromatography. Following this procedure, radiolabeled Gp was only detected in RNA 7-mers derived from mature tRNA generated by c-myc-HSPC117, indicating retention of the precursor-derived splice junction phosphate (Fig. 2E, lane 3). As expected, the splice junction phosphate was exchanged in control ligations performed with T4 Pnk and T4 Rnl1 (Fig. 2E, lane 4). Since no radiolabeled NTP was added to any ligation reactions, no radiolabeled Gp could be detected in RNAse T2 digests of RNAse T1-7-mers derived from mature, T4 Rnl1-ligated tRNA (Fig. 2E, lane 4). A similar nearest neighborhood analysis was also carried out with 2′,3′-cyclic phosphate terminally radiolabeled linear intron after its conversion to circularized intron by c-myc-HSPC117 IP. Detection of radiolabeled Uridine 5′-monophosphate (pU) in P1 nuclease digests of intron circularized by c-myc-HSPC117 but not by T4 Pnk/Rnl1 supported retention of the splice junction phosphate (Fig. 2F, compare lanes 1 and 2). It was concluded that the identified RNA>p ligase joins tRNA exon halves by incorporating the precursor-derived splice junction phosphate into the mature tRNA in a canonical 3′,5′-phosphodiester bond.

Example 5: HSPC117 resides in a stable heteromeric protein complex

Strikingly, in parallel to our efforts to purify RNA>p ligase, a stable and highly homogenous HSPC117-containing complex of so far unknown function was found to copurify with spliceosomal

SF3B particles prepared from nuclear HeLa cell extracts. Therefore, a SF3B-associated and highly purified HSPC117-complex was tested to exhibit RNA>p ligase activity. SF3B-associated HSPC117-complex was obtained by dissociation of SF3B and U2 spliceosomal complexes followed by selective immunodepletion of U2. Unbound SF3B- and HSPC117-complexes were recovered from the flow-through, captured by an anti-SF3B155-coupled resin and specifically eluted with antibody epitope peptide (Fig 3A). After elution from the affinity column, SF3B- (Fig. 3B, fractions 12-19) and HSPC117-complexes (Fig. 3B, fractions 23-26) were separated by size exclusion chromatography on Superose 6. In agreement with the initial data obtained with c-myc-HSPC117 IPs, fractions containing the HSPC117-complex but not the SF3B spliceosomal complex converted tRNA exon halves and linear intron into mature tRNA and circularized intron, respectively (Fig. 3C). In addition, the same fractions could biochemically rescue the tRNA splicing defect of extracts RNAi-depleted of HSPC117 (Fig. 8). Proteins identified in this HSPC117-containing complex were analyzed by tryptic digest of single bands isolated from Coomassie blue stained gels followed by MS analysis. In addition to C22ORF28/HSPC117 identified proteins comprised the DEAD box helicase DDX1 and the ninein interacting protein C14ORF166/CGI-99 in addition to FAM98B and substoichiometric amounts of ASW/C2ORF49 (Fig. 3D). This set of proteins overlaps with the results of the MS analysis of immunoprecipitates of c-myc-HSPC117 from stable cell clones (Tab. 2) and the proteins identified in the MonoQ RNA>p ligase fraction (Fig. 1C and Tab. 1).

Proteins Identified in Enriched RNA Ligase MonoQ Fraction.

Gene Symbol	Number of Unique Peptides	Gene Symbol	Number of Unique Peptides	Gene Symbol	Number of Unique Peptides
EEF2	34	IDH1	11	WDR4	5
MSN	34	BASP1	10	ATE1	4
ELAC2	28	FAM129B	10	CMPK1	4
CCT2	23	FKBP4	10	CPS1	4

CCT8	22	GMPS	10	EEF1A1	4
HSPA8	22	BLVRA	9	ERP29	4
PKM2	20	**C22orf28**	**9**	MAT2A	4
SYNCRIP	20	LRRC40	9	NME1-NME2	4
NARS	19	METTL1	9	PPWD1	4
PUS7L	17	PAPSS2	9	RNASEH2C	4
SND1	16	TSFM	9	UBE2N	4
CCT5	15	CCT7	8	CHMP2A	3
HSPA1B	15	SERBP1	8	DUS2L	3
RDX	15	SLC9A3R1	8	**FAM98B**	**3**
CACYBP	14	PRPS1	7	HNRNPK	3
CCT6A	14	SHMT2	7	HSPA14	3
DUS3L	14	CWF19L1	6	LRPAP1	3
G6PD	14	**DDX1**	**6**	PFKM	3
GLT25D1	14	NAMPT	6	PUS7	3
NMT1	14	PRPSAP2	6	RNASEH2A	3
CCT3	13	TWF1	6	TRMT112	3
FARSB	13	ANKZF1	5	WBSCR22	3
MTHFD1	13	CALD1	5	AKR7A2	2
ACLY	12	EIF4B	5	**C14orf166**	**2**
EIF4A3	12	EIF4H	5	PAWR	2
GARS	12	FARSA	5	RAN	2
LGTN	12	GART	5	RANBP1	2
SSB	12	PAPSS1	5	RDBP	2
TCP1	12	RNASEH2B	5	UBE2V1	2
ABCE1	11	RPRD1B	5		
HSPA9	11	RRM1	5		

Tab. 1. Proteins identified by in solution tryptic digest and MS analysis of MonoQ fraction enriched for RNA>p ligase.

Proteins Identified in c-myc-HSPC117 Immunopurifications sorted by

enrichment.

Gene Symbol	Number of Unique Peptides		Gene Symbol	Number of Unique Peptides		Gene Symbol	Number of Unique Peptides	
	Control	Sample		Control	Sample		Control	Sample
DDX1	**0**	**35**	HNRNPU	20	21	ARL6IP4	3	2
C22orf28	**0**	**27**	HNRPUL1	22	23	RPS23	5	3
FAM98B	**0**	**14**	PRKDC	14	14	RPS6	7	4
C14orf166	**0**	**12**	RPL4	5	5	RPL6	6	3
FAM98A	**0**	**10**	RPL13	5	5	RPL23	4	2
HSPA5	0	5	HNRNPH1	5	5	RPL10	4	2
RPS14	0	3	HIST1H4H	5	5	DSP	11	5
RPL11	0	3	RPS11	4	4	HNRNPAB	6	2
KPNA3	0	3	RPL18	4	4	JUP	10	3
YWHAE	0	2	RPL17	4	4	IGH@	7	2
RPL7	0	2	HIST1H2BD	4	4	ZNF192	4	0
RPL32P18	0	2	UBC	3	3	VCP	3	0
RPL10A	0	2	RPS26	3	3	TF	2	0
PRSS3	0	2	RBBP8	3	3	TAF15	3	0
LMNA	0	2	EEF1A1	3	3	SMC3	9	0
FAM178A	0	2	ZNF326	2	2	SMC1A	5	0
DNAH14	0	2	YBX1	2	2	SBSN	2	0
CSTF3	0	2	RPL37A	2	2	RPS9	2	0
CSTF1	0	2	RPL35A	2	2	RPS4X	8	0
ATP5B	0	2	RPL29	2	2	RPS13	3	0
AHCY	0	2	MYC	2	2	RPL36	2	0
DDX23	3	7	HIST1H2A	2	2	RPL19	2	0

			B					
DDX21	4	9	TRIM21	21	20	RPL18A	2	0
RPS3	2	4	LGALS3BP	16	14	RP9	4	0
RPL21	2	4	TFG	8	7	RIF1	7	0
MTDH	3	5	MATR3	8	7	RAD21	2	0
RPL28	5	8	ACTB	8	7	PCMT1	2	0
TUBA1B	4	6	RBM14	15	13	PABPC1	2	0
RPS18	2	3	RPL7A	7	6	MUCL1	2	0
RPL24	2	3	HNRNPH3	7	6	IgM	2	0
EHBP1L1	19	26	FUS	7	6	IGHA1	3	0
HSPA8	11	15	UTP14A	13	11	IgG light chain	2	0
KPNA2	6	8	VIM	6	5	IgA light chain	3	0
RPL34	3	4	TNRC6B	33	27	Ig kap-pa chain	3	0
RPL3	3	4	TNRC6A	27	22	HNRNPM	3	0
RPL15	3	4	ZNF280D	5	4	GAPDH	2	0
HIST1H3D	3	4	RPL27	5	4	FLG2	3	0
ILF3	20	26	RPL14	5	4	FLG	6	0
TUBB	7	9	APOBEC3B	8	6	FAM133B	2	0
TFRC	4	5	BBX	4	3	DSG1	5	0
ILF2	4	5	HNRNPA1	7	5	CIR1	2	0
HNRNPA3	5	6	BEND3	21	14	C5	2	0
EIF2C2	7	8	RPL8	3	2	C3	6	0
DDX41	9	10	RPL27	3	2	AZGP1	2	0
HNRN-PA2B1	12	13	NKAP	3	2	APOD	2	0

KRT1	26	28	HIST1H1D	3	2	A2M	3	0
HNRN-PUL2	20	21	CSTA	3	2			

Tab. 2. Proteins identified by in solution tryptic digest and MS analysis of affinity purified c-myc-HSPC117.

Silencing of the mentioned interactors of HSPC117 did not have an effect on interstrand ligation (Fig. 3E) or tRNA maturation (Fig. 3F), despite the efficient reduction of mRNA levels (Fig. 3G). Western blot analysis of protein extracts depleted either of HSPC117, DDX1, CGI-99, FAM98B and CGI-99 confirmed efficient depletion of HSPC117, DDX1 and FAM98B (Fig 3H). Remarkably, depletion of HSPC117 also led to decreased protein levels of the other complex members DDX1 and FAM98B. Similarly, depletion of both DDX1 and CGI-99 was accompanied by depletion of DDX1 and FAM98B. Interestingly, HSPC117 containing complexes were reported to exhibit affinity for cruciform DNA structures (Drewett et al., 2001). We were able to inhibit both inter-strand ligation (Fig. 9A) and tRNA maturation (Fig. 9B) with cruciform but not control double stranded DNA. Taken together, these data provide robust evidence for a stable complex composed of HSPC117/C22ORF28, DDX1, CGI-99/C14ORF166 and FAM98B exhibiting RNA>p ligase activity. The copurification of SF3B spliceosomal particles and RNA>p ligase reveals a link between canonical mRNA splicing and RNA>p ligase intriguing.

Example 6: HSPC117 is involved in tRNA ligation in living cells

After demonstrating that HSPC117 is a member of a complex capable of directly joining RNA 2',3'-cyclic phosphate and 5'-OH termini *in vitro*, we next investigated the role of HSPC117 in RNA ligation *in vivo*. We first tested whether inter-strand ligation can occur in living cells by transfection of [5'-^{32}P]-pCp-radiolabeled dsRNA into HeLa cells, RNAi-depleted of RNA terminal cyclase (RTCD1), HSPC117 or EGFP as a control gene (Fig. 4A). Isolation and analysis of total RNA by gel electrophoresis demonstrated that the radiolabeled 3'-P dsRNAs were taken up by the cells and subsequently ligated. *In vivo*, inter-strand ligation was dependent on HSPC117 (Fig. 4A, compare lanes 4 and 7 with lanes 3 and 6, respectively) and to a less pronounced ex-

tent on RTCD1 (Fig. 4A, compare lanes 5 and 8 with lanes 3 and 6, respectively). In order to test the effect of RNAi-mediated depletion of HSPC117 on processing of *de novo* synthesized pre-tRNA, a system that has previously been applied to study mitochondrial import of tRNA in *Trypanosoma brucei* (Bouzaidi-Tiali et al., 2007) was emploeyed. siRNAs targeting HSPC117, the human splicing endonuclease subunit TSEN2 or EGFP as controls together with a reporter construct, encoding tagged pre-tRNA Ile^{TAT} (Fig. 9) under control of a Tetracyclin (Tet) responsive promoter (Fig. 4B), were transfected into HeLa cells stably expressing Tet-repressor. After induction at selected time points (Fig. 4C), RNA was extracted and analyzed by Northern blot. Probing of the blot for the 5'-exon of the tagged tRNA Ile^{TAT} (probe I5E, Fig 4B) revealed a clear delay in the formation of mature tRNA upon RNAi-mediated depletion of HSPC117 and TSEN2 but not upon EGFP control siRNA treatment (Fig. 4D, compare lanes 1-3 with lanes 4-6 or 7-9, respectively and Fig. 4E). Taken together, these data establish a role for HSPC117 as an RNA ligase with broad substrate specificity and with a function in tRNA processing.

Example 7: Potential functions of HSPC117 as the first RNA ligase to be discovered in humans

HSPC117 is the catalytic subunit of a human RNA ligase complex. The known ATPase and unwinding activity of helicases suggest a potential role for DDX1 in ATP-dependent conformational changes to support tRNA splicing. In fact, DDX1 is the only polypeptide associated with the RNA>p ligase-dependent tRNA splicing pathway that has been shown to require ATP for its action as a DEAD-box helicase. This could explain the supportive but dispensable role of ATP in the endonuclease-RNA>p ligase reaction cascade which, mechanistically, does not require any additional energy. Little is known about the potential functions of CGI-99, FAM98B and ASW. The data herein, based on RNAi-mediated depletion, suggest a non-essential role for the ligation of tRNA halves (Fig. 7B and C).

Both RNA>p ligase and T4 Pnk/Rnl1-like ligation mechanisms can be detected in human cells (Filipowicz and Shatkin, 1983). Here, we have identified components of an RNA>p ligase pathway. Pro-

teins catalyzing the 5′-OH kinase and 2′,3′-cyclic phosphodiesterase healing reactions as well as the final phosphotransferase step (required for this pathway have been identified in humans. In addition, it has recently been demonstrated by heterocomplementation studies in yeast that these proteins can indeed function as tRNA splicing enzymes *in vivo*. Earlier studies show that the RNA>p ligase pathway is the prevalent tRNA ligation pathway in mammals (Filipowicz and Shatkin, 1983).

The RNA>p ligase complex described here is the first identified RNA ligase protein that is capable of directly joining the 2′,3′-cyclic phosphate, 5′-OH displaying products of the tRNA endonuclease reactions without the necessity of prior "healing" of RNA termini. This implies a more general role of RNA>p ligase in RNA repair pathways. Instructive examples of RNA repair systems are known from prokaryotes. Recently, the role of RNA repair in eukaryotes has received increasing attention. RNA>p ligase is involved in human RNA repair routes and non-canonical splicing events mediated by HSPC117. Stress-induced cleavage and subsequent spliceosome-independent ligation of mRNA during the unfolded protein response (UPR) exemplifies such a non-canonical splicing event requiring an RNA ligase. The pathway has been extensively investigated in *S. cerevisiae* and has implicated the yeast tRNA ligase Trl1p in the final ligation step. A similar UPR pathway that relies on stress-induced non-conventional splicing of the *XBP1* mRNA exists in human cells. However, the responsible ligase was previously unknown. Previous work suggests that the phosphotransferase TPT1, which is a potential component of the T4 Pnk-like RNA ligation pathway in humans, is dispensible for stress-induced splicing of the *XBP1* mRNA. HSPC117 acts as an RNA>p ligase in the human UPR pathway.

In addition, RNA ligases have been proposed to be involved in viral replication in humans. For example, a host-encoded ligase has previously been postulated to circularize the RNA genome of Hepatitis Delta virus (HDV) during rolling cycle replication in humans. It appears that the HSPC117-containing ligase complex described here participates in host-mediated viral genome circularization. Recently, both HSPC117 and DDX1 have been reported to be required for replication of HDV in human cells. Further-

more, all members of the described RNA>p ligase complex have been shown to interact with kinesin-associated RNA transport granules in mouse brain extracts. Intriguingly, RTCD1 was also found to be associated with these RNA transport granules, thereby establishing a potential functional link between RNA terminal phosphate cyclization and RNA ligation.

The high degree of conservation of HSPC117/RtcB proteins in organisms as distantly related as humans and *E. coli* shows the universal and important roles for this protein family. Initial work indeed established an essential function for HSPC117 in mammalian development. On the other hand, no biological function could so far be assigned to the operon harboring HSPC117/RtcB and RtcA in some bacteria. Herein the molecular function of HSPC117/RtcB proteins has been elucidated.

References

Abelson, J., Trotta, C.R., and Li, H. (1998). tRNA splicing. The Journal of biological chemistry *273*, 12685-12688.

Bouzaidi-Tiali, N., Aeby, E., Charriere, F., Pusnik, M., and Schneider, A. (2007). Elongation factor 1a mediates the specificity of mitochondrial tRNA import in T. brucei. The EMBO journal *26*, 4302-4312.

Chan, P.P., and Lowe, T.M. (2009). GtRNAdb: a database of transfer RNA genes detected in genomic sequence. Nucleic acids research *37*, D93-97.

Drewett, V., Molina, H., Millar, A., Muller, S., von Hesler, F., and Shaw, P.E. (2001). DNA-bound transcription factor complexes analysed by mass-spectrometry: binding of novel proteins to the human c-fos SRE and related sequences. Nucleic acids research *29*, 479-487.

Englert, M. (2005). Mechanismus des pre-tRNA Spleißens: Struktur und Funktion pflanzlicher und animaler RNA Ligasen. In PhD thesis, Faculty of Chemistry and Pharmaceutics (Würzburg, Bayerische Julius-Maximilians-Universität), pp. 139.

Filipowicz, W., and Shatkin, A.J. (1983). Origin of splice junction phosphate in tRNAs processed by HeLa cell extract. Cell *32*, 547-557.

Galperin, M.Y., and Koonin, E.V. (2004). 'Conserved hypothetical' proteins: prioritization of targets for experimental study. Nucleic acids research *32*, 5452-5463.

Genschik, P., Drabikowski, K., and Filipowicz, W. (1998). Characterization of the Escherichia coli RNA 3'-terminal phosphate cyclase and its sigma54-regulated operon. The Journal of biological chemistry *273*, 25516-25526.

Laski, F.A., Fire, A.Z., RajBhandary, U.L., and Sharp, P.A. (1983). Characterization of tRNA precursor splicing in mammalian extracts. The Journal of biological chemistry *258*, 11974-11980.

Okada, C., Maegawa, Y., Yao, M., and Tanaka, I. (2006). Crystal structure of an RtcB homolog protein (PH1602-extein protein) from Pyrococcus horikoshii reveals a novel fold. Proteins *63*, 1119-1122.

Reid CE, Lazinski DW (2000) A host-specific function is required for ligation of a wide variety of ribozyme-processed RNAs. *Proc Natl Acad Sci USA* 97(1): 424-429

Claims:

1. The use of a HSPC117 molecule as RNA ligase, preferably for the transfer of a first RNA end to a second RNA end, in particular preferred wherein the first end comprises a 3' phosphate, preferably a 2',3'-cyclic phosphate, and/or wherein the second end comprises a 5'-OH terminus.

2. The use of claim 1 for inter- or intra-strand ligation.

3. The use of claim 1 or 2 wherein the RNA is double-stranded.

4. The use of any one of claims 1 to 3 for RNA splicing.

5. The use of any one of claims 1 to 4 wherein the RNA is a mRNA or tRNA.

6. The use of any one of claims 1 to 5 wherein the RNA is 2 to 3000 nucleotides or basepairs in length.

7. The use of any one of claims 1 to 6 wherein the HSPC117 molecule is of any one of SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23.

8. The method of ligating at least two RNA molecules comprising using a HSPC117 molecule according to any one of claims 1 to 7, preferably comprising contacting the at least two RNA molecules with said HSPC117 molecule in a cell, or preferably comprising contacting the at least two RNA molecules with said HSPC117 molecule *in vitro*, in particular preferred further comprising providing the HSPC117 molecule in a complex, preferably with spliceosomal particles, preferably wherein the complex comprises a DEAD box helicase DDX1 molecule, a FAM98B molecule, a CGI-99 molecule, ASW or any combination thereof.

9. A kit suitable for performing RNA ligation, comprising a HSPC117 molecule,

a RNA ligase reaction buffer comprising buffer components and one or more metal ions selected from Mg^{2+}, Mn^{2+}, Ni^{2+} or mixtures thereof, for use in a final concentration range of ca 0.1-20 mM, preferably 1-10mM, in particular preferred 2-5 mM, optionally further comprising a RNA molecule with 2',3'-cyclic phosphate, preferably further with a label.

10. A transgenic cell comprising an exogenously expressed HSPC117 molecule, preferably with an inducible promoter.

11. A HSPC117 knock-out cell or cell with reduced inhibited endogenous HSPC117 expression.

12. Cell according to claim 10 or 11 being a mammal cell.

13. A method of reducing RNA ligase activity, in particular RNA>p ligase activity, in a cell comprising inhibiting a HSPC117 molecule in a cell, preferably by knock-out or RNAi.

14. The method of claim 13, characterized in that inhibition is by reducing expression of an HSPC117 molecule, preferably an endogenous HSPC117 molecule, in said cell.

15. The method of claim 13 or 14 for reducing tRNA production or processing in said cell.

16. The use of a cell according to any one of claims 10 to 12 for RNA ligation or splicing studies, preferably to modulate tiRNA amounts in a cell.

17. A method of reducing tumor cell growth or a disease or infection dependent on host polymerases, such as hepatitis delta virus infection, comprising administering an HSPC117 molecule inhibitor to said cell, or of treating a disease in a subject associated with dysfunctional tRNA splicing, preferably pontocerebellar hypoplasia, comprising administering a HSPC117 molecule to said subject.