“Role of the Cellular Decapping Activators LSm1-7 Complex in the Replication of Positive-Strand RNA Viruses”

Rui Pedro Galão

TESI DOCTORAL UPF/2010

DIRECTOR DE LA TESI

Dra Juana Díez Antón

UNIVERSITAT POMPEU FABRA

À minha Família

Acknowledgments

It was a long journey that started in Espinho, near the Atlantic, on the early hours of a "Día de Reys". A foggy dawn, I can still remember. The VW Polo, a long time faithful "friend", was packed not only with my belongings but also with hope and dreams. It was the beginning of a adventure that now reaches the end. I bear in mind all the motivation and high expectations with starting the PhD, the excitation of discovering new places, the curiosity to meet new people. Could not imagine, however, so many different people from so many different origins would come in touch with my work, with my life and somehow changed it… changing me. It is not possible to name everybody but, to everyone that along these years were side by side with me, sharing ups and downs, successes and problems… beers; to everyone that directly or indirectly contributed to this thesis; to everyone that simply made my life easier and better… to all of you thanks, muchas gracias, moltes gràcies, Vielen Dank, obrigado! Now that I used the normal artifact to have an excuse in case I forget someone (that for sure I will…), I would like to mention some of you in particular.

And I start obviously by Juana. Jefa, first of all thank you for the opportunity to work in your laboratory during these years. It was a long journey but it was pretty much worth it. There were good and difficult moments, both work and personal related. I always could count with your trust, advise, guidance, support, comprehension and patience… Muchíssimas gracias, por todo. Andreas, it was a pleasure to meet you and work in your laboratory (even considering that it was in the "exciting" Homburg-Saar). Thank you for the opportunity, support and for always try to make me feel as part of your group. I wish all the best life can bring to you both, Luca and Hanna.

And to finish the part that I acknowledge my supervisors… Prof Dr Fischer. Utz, it was a very nice time the period(s) I passed at Wuerzburg. Thank you for the opportunity, hospitability, advise and supervision during the last years of my PhD. It was quite a fundamental input not only for this thesis but also to our group. Please extend this acknowledgments to all at AG Fischer. All of you made me pass a wonderful time there. Obviously, a special word to Ashwin… you just made my life simpler. Thank you not only for all the help, discussions and friendly rescues in the lab, but also for the time we passed out of the lab. Wuerzburg night life is not the best, but we can not complain about the food. I wish to you and Elham all the best. By the way, Clemens and Ohmer (Simpson…), say something when you think you can beat me playing kicker…

Isabel. que posso dizer-te? Foste, em Barcelona, a minha primeira colega, a minha primeira amiga. Partilhamos projectos e escrita da tese, partilhamos "pisos" a cada regresso meu da Alemanha, partilhamos pensamentos, segredos, confidencias, e alguns telefonemas à tua irmã…, partilhamos séries, BGT (tu e eu sabemos o que isto é…), cervejas, nespressos, arroz de polvo e aquelas tuas sopas de tomate que ainda hoje não encontro adjectivos para descrever. Ainda assim soube a pouco. Ter-te conhecido não tem preço. Obrigado por tudo, linda!

Leonardo, boludo… por ser para ti voy a escribir en castellano que ya sé que te encanta mi portuñol. Dentitos, ensaladita, linda, tonta. Fue usando estas palabras, mientras compartíamos la campana de cultivos, qué empezamos una fuerte amistad en las frías tierras alemanas, perdón, de Saarland. Amistad que espero que se prolongue por mucho tiempo… al menos el tiempo que sea necesario para que aprendas a jugar dardos. A no ser que te encante verme bailar… Bueno, gracias por todo, Leo. Fue todo mucho más divertido y enriquecedor contigo cerca. Y gracias por la especial ayuda en la punta final. No hay cerveza suficiente en el Bitácora que pueda pagarte. Y por cierto, un petó a Mito… que sigan siempre tan felices como hoy. Por cierto, tranquilo que vos espero en 2012.

Meine liebe Freundin Nicoletta, Vielen Dank! It was very nice to have you nearby, first in Homburg, then in Barcelona and, at the end, as flatmate. You were since the day we met, someone I could always count with both at work, where you have been always present and helpful, and in the "normal" world where we built a comforting friendship in which you let me discover your peculiar sense of humour – missing it already - and drinking capacity! Not bad for a "Danish viking" girl… I wish you all the best and once again, thank you for everything.

I was several times abroad, several times in different groups, but this is the moment to thank all that form (or formed) part of what I always considered MY group. Mireia, if I appreciated so much giving the classes I have to thank you. Without your help (and notes), it would not be so nice. Thanks for all the help that you gave in the not so long time we shared in the lab. Daniela, Gemma Maité, and also Antonio and Encarna, thank you for all the nice moments we passed together, your good humour and friendship. I will never forget it. A special word to you Anna, to all your support and encouragement in this last stage of the thesis. It was very important to have someone like you nearby. Someone that I could trust and count with. Moltes gràcies per tot.

I want also to thank some other people that made the time I passed in Barcelona as one of the best periods in my life. The so-called Rabenessen Lab, "los portugueses

eses", ex-flatmates. In no special order, and trying to don't forget anyone in the way (forgive me if I do): Jordi, Bea, Katy, Luisa (escrever a tese a três… nunca me irei esquecer!), Annie, Vane, Fabien, Rosa, Olga, Rui Miguel, Tiago, Confras (aka Isabel Augusto), Joana, Aninhas, Pedro, Mito, Brian and Jeremy. Can not forget also my friends at Espinho and Lisboa. Thank you all so much.

I can not pass without acknowledge Anke. We didn't know each other before I started my PhD, we are no longer together now that is finished. However, for me, this thesis will be always linked to you. You were with me through most of it, you were always there when I needed it, you were a huge help and support. Thank you for all the nice moments we passed together, and they were so many. This thesis is also yours. A special acknowledgment to all your family. They simply made me feel part of it, be one of them. And this says everything. Danke Anke.

E para o fim o mais importante dos agradecimentos. A toda a minha família. Não é possível por em palavras tudo o que eu tenho para vos agradecer. Mãe e Pai, vocês desde cedo me deixaram tomar as decisões que eu achava ser melhor para mim, sempre me apoiaram, sempre me ajudaram. Desde sempre fizeram sacrifícios e deram-me tudo o que podiam dar para que eu pudesse atingir mais facilmente os meus objectivos, os meus sonhos. Nunca vos poderei retribuir uma pequena parte que seja de tudo isto. Este livrito, esta tese, tem o meu nome na capa, mas é antes demais vossa. Sem vocês simplesmente não existia. Obrigado.

Paula, Miguel… manitos consegui! Paula, o "bichinho" da ciência foste tu que o puseste cá dentro. Pois bem, parece que deu frutos. No próximo Natal eu faço uma apresentação só para ti! Miguel, não fazes a mínima ideia do que faço eu no laboratório mas sempre acreditaste em mim e sempre me motivaste a procurar mais e melhor. Obrigado por tudo, a vossa confiança e apoio. Por estarem sempre presentes mesmo não estando aqui. Tia, Zé, Márcia e Sofia. Também tenho muito que vos agradecer. Cada regresso a casa é especial e único. É óptimo saber que posso contar convosco e sentir-me tão querido por vocês. Obrigado a todos.

Rui Pedro, 2010

Summary

Members of the (+)RNA viruses group were shown to interact with the host decay machinery in order to express their genomes and regulate their life cycles. Previous works using the BMV/yeast system demonstrated that the proteins LSm1-7, Dhh1 and Pat1 - activators of mRNA decapping in the 5'-3'-deadenylation-dependent mRNA decay pathway - are essential cellular factors required for both translation and exit from translation to replication of the BMV genome. Here, we show that the corresponding human homologues of these proteins, LSm1-7, Rck/p54 and PatL1, are required for HCV RNA translation and replication. Furthermore, we also show that reconstituted LSm1-7 complexes specifically recognize important signals, either in BMV and HCV genomes, that regulate their translation and/or replication. These observations constitute the first evidence that the LSm1-7 complex is able to directly interact with viral genomes representing also novel LSm1-7 interaction sites. Given the common replication strategies of (+) RNA viruses and the conserved cellular functions of LSm1-7, Pat1 and Dhh1 from yeast to humans, our findings pinpoint a weak spot that may be exploited to generate broad-spectrum antiviral drugs.

Resumen

Se ha demostrado que los miembros del grupo de virus ARN de polaridad positiva interaccionan con mecanismos de degradación de ARNm de la célula hospedadora con el fin de expresar sus genomas y regular sus ciclos de vida. Trabajos anteriores utilizando el sistema BMV/levadura han demostrado que las proteínas LSm1-7, Dhh1 y Pat1 – activadores de *decapping* del ARNm en la vía de degradación de los ARNm celulares dependiente de la deadenilación - son requeridas tanto para la traducción como para la salida de la traducción a la replicación del genoma de BMV. Aquí, hemos demostrado que los correspondientes homólogos humanos de estas proteínas, LSm1-7, Rck/p54 y PatL1, son necesarios para la traducción y replicación del ARN del virus de la Hepatitis C. Por otra parte, anillos reconstituídos de LSm1-7 reconocen específicamente señales importantes, tanto en los genomas de BMV como del virus de la Hepatitis C que regulan su traducción y/o replicación. Estas observaciones constituyen la primera evidencia de que el complejo LSm1-7 es capaz de interactuar directamente con genomas virales y representan también novedosos patrones de interacción de este complejo con ARN. Teniendo en cuenta las estrategias de replicación en común de los virus de ARN de cadena positiva y las funciones celulares conservadas de LSm1-7, Pat1 Dhh1 de levadura a humanos, nuestros resultados señalan la posibilidad de explotar estas proteínas para la generación de medicamentos antivirales de amplio espectro.

Preface

The family of positive-strand RNA ((+)RNA) viruses constitutes an extensive genetic group of plant, animal, and human viruses that encompass over one third of all virus genera. Among them are important human pathogens such as Hepatitis C Virus, Poliovirus and SARS which are associated with syndromes such as colds, febrile illness, meningitis, paralysis, hepatitis and hemorrhagic fevers. The different families of (+)RNA viruses have evolved various genome organizations, virion morphologies and life cycles that exploit the biology and biochemistry of their hosts in diverse ways. Albeit these differences all (+)RNA viruses share fundamental similarities concerning their RNA replication. A deep understanding of the common replication strategies, namely the common cellular factors involved, is essential not only to increase our fundamental knowledge in cell and virus biology but also for the develpoment of broad-sprectrum antiviral strategies.

Taking advantage of three complementary experimental systems - the replication of the brome mosaic virus in yeast, the replication of HCV in human cells and the use of purified reconstituted protein complexes to address their interaction with viral RNAs - this thesis has been focused in the identification and characterization of a group of cellular factors that we demonstrate to have a very important role in the life cycle of (+)RNA viruses. The main results generated during this work have been included in three publications in *peer-reviewed* journals, which are presented in the "Results" section of this thesis. Complementary data is introduced in the form of appendices within the same section. Additionally, other review articles published during the thesis are included in a separated chapter.

Index

1. Introduction

1.1 Sm and Sm-like (LSm) Protein Family

The Sm and Sm-like (LSm) proteins constitute a conserved family whose members have been identified in all kingdoms of life and are known to be involved in RNA metabolism[1-3]. They are characterized by the presence of a Sm fold that includes two conserved domains, referred as Sm motifs 1 and 2, that are separated by a variable region[4]. The Sm fold itself consists of a closed barrel comprising five anti-parallel ß-strands and a N-terminal α-helix stacked on the top, enabling the association of these proteins into ring-shaped complexes. The canonical Sm proteins, named B/B', D1, D2, D3, E, F and G, form a heteroheptameric ring structure when binding to conserved single-stranded regions of U1, U2, U4 and U5 small nuclear RNAs (U-snRNAs), to which remains permanently associated following assembly[2]. In addition, two other heteroheptameric rings have been identified in eukaryotes: the LSm2-8 complex, composed of seven protein subunits named LSm2 to LSm8, and the LSm1-7 complex, in which the subunit LSm8 is exchanged by LSm1 (reviewed in[1-3]). Furthermore, the homolog of Sm and LSm proteins in prokaryotes, Hfq, shows a remarkably similarity to its eukaryotic counterparts as it forms a doughnut-shape complex, presenting however a homohexameric structure rather than a heteroheptameric one[3]. In contrast to the Sm complex, both LSm proteins and Hfq bind transiently to RNAs and are able to form ring-shaped complexes even in the absence of RNA.

These complexes are important players in multiple aspects of RNA metabolism. Sm proteins play a key role in the maturation of their target U-snRNAs, which are essential for nuclear pre-mRNA splicing[5]. The LSm2-8 complex is localized to the nucleus where is known to associate with U6 snRNAs and to be involved in processing and decay of various nuclear RNAs, as well as in pre-mRNA splicing. Conversely, the LSm1-7 complex localizes to the cytoplasm and functions as activator of decapping in the 5'-3' deadenylation-dependent mRNA decay pathway[6]. In mammalians, LSm1-7 complexes were shown to participate particularly in the decay of mRNAs containing AU-rich elements (ARE) and histone mRNAs[7-9].

Intriguingly, it was observed that the LSm1-7 complexes also serve to protect 3'-ends of mRNAs from trimming *in vivo,* as shown by the accumulation of 3'-end "trimmed" forms of several mRNAs in Δ*lsm1*-Δ*lsm7* loss of function yeast mutants[10,11]. The relationship between the functions of LSm1-7 complex in mRNA decay and 3'-end protection remains elusive. Finally, Hfq has been implicated in diverse cellular functions, like acting as translation regulator by interfering with ribosome binding, and as modulator of the activity from key bacterial mRNA decay enzymes[12-16].

Some complexes from the Sm family of proteins have been suggested to present RNA chaperone-like activities in connection to their cellular functions[2,3]. RNA chaperones are, by definition, non-specific RNA binding proteins that help RNA folding by facilitating RNA-RNA interactions, resolving misfolded structures or preventing their formation. Their function is of capital importance because some RNA molecules become easily trapped in inactive conformations given their structural and functional flexibility (reviewed in[17]). A hallmark of the interactions between chaperones and their target RNAs seems to be their transient nature: once the RNA has been folded, proteins can be removed without alteration of the RNA conformation[17,18]. The canonical Sm-complex was shown to present a chaperone-like function when associated to the U1 snRNA as it stabilizes the interaction between this snRNA and the 5' splice site[19]. Similarly, during pre-mRNA splicing, LSm2-8 facilitates multiple rearrangements of splicing complexes what was suggested to be a chaperone-type activity[20,21]. However, within the complexes from the Sm family, Hfq is the only one which chaperone activities has been proved. Hfq has the ability to resolve misfolded RNA structures[12], and enables translational control and mRNA stability by facilitating the interaction of small noncoding regulatory RNAs (ncRNAs) with their mRNA targets[22]. Furthermore, Hfq chaperone activities have been associated with other cellular functions such as stress responses and nitrogen fixation[23-25].

Interestingly, it has also been shown that LSm1-7 and Hfq complexes are involved in the replication of positive-strand RNA viruses. The LSm1-7 complex was shown to be required in different steps of Brome mosaic virus life cycle. However, the

molecular mechanisms underlying its functions are not well characterized. By its turn, the Hfq complex was first identified as a factor required for the replication of the bacteriophage Qß[26], which role was later associated with its RNA chaperone activity[27]. Similarly to the cellular RNAs, regions from viral RNA genomes can also adopt different structural conformations, some of them acting as misfolded intermediates lacking functional relevance. Thus, proteins with chaperone-like activities are likewise essential in the context of viral RNAs, where they are often required for an efficient replication. Accordingly, Hfq is capable of remodeling the structure of the Qß viral RNA in order to allow the access of the viral replicase to the 3'end. The growing list of cellular or viral encoded RNA chaperones involved in different steps of viral cycles is representative of their importance[18].

Complexes from the Sm family show a preferential binding to single-stranded RNA regions containing AU-rich sequences. In the case of the canonical Sm complex, its respective subunits present in the lumen of the ring exactly the same spacing as the nucleotides in a RNA molecule, what makes possible their binding to single-stranded RNA regions via a repeating motif[5]. This conserved motif is present in single-stranded regions of particular U-snRNAs and is denominated U-rich Sm-site ($AU_{(4-6)}G$), to which the Sm complexes are known to assemble around[2].

Knowledge of the binding specificities from LSm complexes is currently limited. Assuming the same rearrangement as in the Sm complex, a model was proposed to localize the respective LSm paralogs within the heptameric ring complexes. Further studies based in this model suggested that LSm rings have a more general RNA binding motif than the Sm complex[2,28,29]. Indeed, the nuclear LSm2-8 complex was shown to recognize both an internal and a 3'-terminal uridine-rich sequences in U6 snRNA[20,30], and to bind fully adenylated mRNAs, suggesting that this complex may interact with longer poly(A) tracts[31]. Additionally, studies with recombinant human LSm1-7 rings showed their direct interaction and subsequent stabilization of reporter RNAs containing 5'poly(A) tracts[32]. Further insights into the RNA-binding activity of the LSm1-7 complex come from the studies addressing its function as decapping activator in the 5'-3' mRNA decay pathway in yeast[6,33]. Binding assays with a purified yeast complex composed of LSm1-7 and other activator of decapping

named Pat1, showed a preferential binding of this complex to oligoadenylated versus polyadenylated RNAs and a direct interaction near or at 3´ends that contain oligo(U) stretches[34]. However, since the protein Pat1 was present in the complex, its contribution to the identified RNA binding properties remains unclear.

Finally, Hfq has been very well characterized and shown to bind both internal sequences and 3'terminal regions (reviewed in[2,3]). The distal face of the Hfq ring interacts with poly (A) sequences, while the proximal face is known to interact with internal AU-rich single-stranded regions situated near stem loops in ncRNAs targets[36,35], as well as to a large range of tRNAs either in their D- or T-sites[37]. The capacity of both faces of the ring to bind RNA implies that Hfq can potentially bind two RNA substrates simultaneously, or a single one at two sites, what is suitable to its role as an RNA chaperone that can facilitate RNA-RNA interactions[3,36]. The mechanism of binding to ncRNAs was proposed to involve the winding of AU-rich oligonucleotides around the central hole in the proximal face of the Hfq ring[27]. However, a structural effect rather than a sequence specificity should rule the binding of this complex to tRNAs given their lack of AU-rich and single-stranded RNA regions in their D- or T-sites[37].

1.2 Pathways of mRNA degradation in Eukaryotes

Eukaryotic mRNAs integrate two stability determinants that are incorporated co-transcriptionally — the 5′-7-methyl-guanosine cap (m^7G) and the 3′-poly(A) tail. These two structures interact, respectively, with the initiation factor eIF4E and the poly(A)-binding protein (PABP) in order to enhance translation initiation and protect the transcripts from exonucleases. mRNA decay initiation occurs when either one of these two structures is compromised or the transcript is internally cleaved by endonucleases (reviewed in[38]).

The process of mRNA degradation is of capital importance for the physiology of eukaryotic mRNAs. First, mRNA turnover is essential for the control of gene expression both by setting steady-state levels of gene expression, and by acting as a site of regulatory responses[39]. Second, mRNA decay pathways are also involved in

host responses to viral infections. This antiviral role includes not only the basal functions of the mRNA decay machinery, but also specialized systems such as the responses to dsRNA and the RNA interference, in which short interfering RNAs (siRNAs) initiate decay by endonucleolytic cleavage mediated by Argonaute protein-2 (AGO2)[40,41]. Finally, specialized surveillance machineries exist that identify and degrade aberrant mRNA molecules, thereby increasing the quality control of mRNA biogenesis and protecting the cell from potentially toxic protein production[42]. Among them there are the nonsense-mediated decay (NMD), that detects and degrades transcripts containing premature termination codons[43], the non-stop decay (NSD), that targets mRNAs lacking a stop codon[44] and, finally, the no-go decay (NGD) that detects and cleaves mRNAs with stalled ribosomes[45-47].

The majority of the cellular mRNAs are degraded by two general pathways of mRNA decay in eukaryotic cells – the 3'-5' and the 5'-3' deadenylation dependent mRNA decay (Figure 1). In both cases, the degradation of the transcript begins with the shortening of the poly(A) tail at the 3'end of the mRNA, in a process denominated deadenylation. Although there are some characterized eukaryotic deadenylases, such as PAN2-PAN3, CCR4-NOT and PARN (poly(A)-specific ribonuclease), it is still unknown exactly when and by which mechanism deadenylation is triggered. This is the only step in the mRNA degradation pathway that is potentially reversible as transcripts possessing the correct signals can be readenylated and return to polysomes. However, once the cell targets a mRNA for turnover, the deadenylated mRNAs can be degraded either by the 3'-5' or 5'-3' mRNA decay pathways[48,49].

The 3'-5' mRNA decay is carried out by the exosome, a 10-12 subunit complex that includes proteins with significant homology to 3'-5' phosphorolytic exo-ribonucleases, to hydrolytic 3'-5' exonucleases or to RNA helicases[38,39]. Furthermore, each of the core exosome subunits has an RNase PH domain, which might contribute to the catalytic activity. Following the 3'-5' decay, the remaining 5'cap structure is metabolized by the scavanger-decapping enzyme DcpS[50].

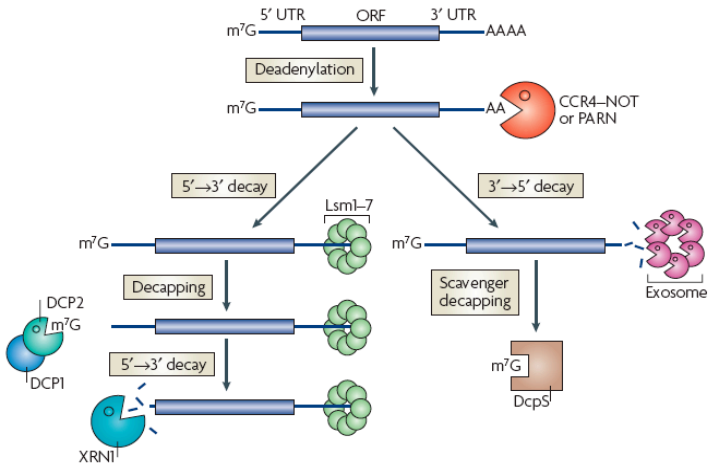

Figure 1. In eukaryotes, the bulk of mRNAs undergo decay by deadenylation-dependent pathways. The 3'-poly(A) tail is removed by deadenylases, such as the represented CCR4–NOT or PARN. After deadenylation, mRNAs can be degraded by two mechanisms, either the decapping-dependent 5'-3' decay pathway or the 3'-5' decay. In the decapping pathway, the LSm1–7 complex interacts with the 3′ end of the mRNA transcripts and activates decapping by the DCP1–DCP2 complex. The 5'end of the mRNA is then exposed to be degraded by the 5′-3′ exoribonuclease XRN1. Alternatively, the deadenylated mRNA can be degraded in the 3′-5′ direction by the exosome, with the remaining cap structure being hydrolysed by the scavenger-decapping enzyme DcpS. (extracted from[51]).

The 5'-3′ mRNA-decay pathway initiates with the removal of the m^7G cap, in a process denominated decapping. In *S.cerevisiae*, Dcp1 and Dcp2 proteins form a dimmer that acts as a decapping enzyme. Dcp2 supplies the catalytic activity through its MuT domain[39]. In higher eukaryotes there is a third protein, known as Hdles or Ge-1, which enhances decapping activity of the catalytic Dcp2 and promotes the formation of a complex between Dcp2 and Dcp1[52]. Following the decapping step, RNA is degraded by the 5'-3′ exoribonuclease Xrn1. LSm1-7 complex plus two other proteins, Dhh1 and Pat1, serve as accessory factors in this pathway by facilitating the transfer of the cellular mRNAs from a translating to non-translating state and acting as activators of decapping (reviewed in[53,54]).

1.3 The control of mRNA decapping

The regulation of eukaryotic gene expression occurs mainly through the control of mRNA translation and degradation in the cytoplasm. The presence of a m^7G cap in mRNAs is essential for initiation of translation and also for protecting them from the 5'-3' exonucleolytic decay[55-57]. Thus, the removal of the cap is a key step in the regulation of gene expression since leads to mRNA translation repression and mRNA turnover[58]. In this manner, mRNA decapping and mRNA translation are considered to be competing pathways (Figure 2).

Several proteins are known to function in the stimulation or inhibition of the decapping process. The eukaryotic initiation factor 4E (eIF4E), component of the eIF4F complex, associates with the m^7G cap and by that means protects it from decapping and activates translation initiation[59,60]. eIF4F forms a compact complex with the mRNA cap, which is further stabilized by an interaction between the eIF4G subunit of eIF4F complex and the cytoplasmic poly(A)- binding protein (Pab1p) at the mRNA poly(A) tail[61]. Thus, in order to the decapping process to occur it is required that, in a first step, the cap-binding translation factors are removed and that, in a second step, Dcp2 binds to the cap structure to catalyze decapping[61,62].

Proteins that promote this decapping process are know to act in either steps. Interestingly, LSm1-7 complex, plus Dhh1 and Pat1, are believed to activate decapping by stimulating the removal of the eIF4F complex what results in the repression of cellular mRNAs translation. This will lead to the second step in which the core decapping complex can access and remove the cap from the mRNA. This second step is known to be promoted by the enhancers of decapping Edc1 and Edc2, specific to yeast, that stimulate the Dcp2 cap-binding/catalysis step. Furthermore, the eukaryotic conserved Edc3 interacts directly with the decapping factors Dcp2 and Dcp1[59,60,62], what suggests that Edc3 can either recruit or activate the decapping complex on target mRNAs[63,64]. However, there are cases in which mRNAs are kept in a translational inactive state, meaning that not all translational repressed mRNAs are immediately targeted for decapping[56].

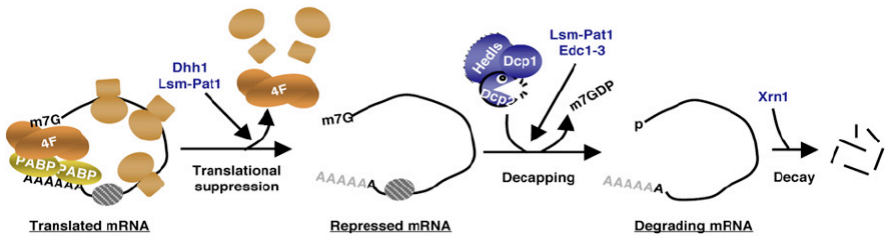

Figure 2. Multiple factors promote mRNA decapping. mRNA decapping activation is thought to occur in two distinct steps. First, the cap-binding complex (eIF4F) must be removed from the mRNA, a process that results in a mRNA translationally inactive. This process is stimulated, at least in yeast, by LSm1-7, Dhh1 and Pat1. The core decapping complex, including Dcp2, Dcp1, and possibly in metazoans Hedls, can then access and remove the mRNA m^{7}G cap through a process than can be stimulated by Edc1-3. The decapped mRNA is then degraded by Xrn1. (Adapted from[61,65])

Given that the proteins Dhh1 and Pat1, together with the already mentioned LSm1-7 complex, are in the core of this thesis, I will describe them in more detail. The *Saccharomyces cerevisiae* DEAD-box Dhh1 RNA helicase is a member of the helicase SF2/DDX6 subfamily, highly conserved across a wide range of organisms, with homologues in mammals (Rck/p54), in *Xenopus* (Xp54), *Drosophila* (Me31B), and *Caenorhabditis elegans* (CGH-1) (reviewed in[66]). DEAD-box helicases are involved in splicing, ribosome biogenesis, RNA transport, degradation, and translation repression, although their precise contribution to most of these processes is mainly unknown. Their catalytic core is composed of two RecA-like domains with nine conserved motifs, including the eponymous DEAD motif and the more recently identified Q motif, with roles in catalysis and substrate binding. Helicases are understood to use NTP (usually ATP) binding and hydrolysis to remodel RNA or RNA-protein complexes, resulting in double-strand RNA unwinding and/or in displacement of proteins from RNA (reviewed in[66-68]). Accordingly, the role of Dhh1 as translation repressor was suggested to be due to its ATPase activity, as it can release the eIF4E complex from the mRNP which simultaneously repress translation and stimulate decapping[69,70]. However, Dhh1 also represses translation of reporter mRNAs that do not depend of eIF4F for translation. Thus, Dhh1 could alternatively promote translational repression by a general destabilization of the eIF4F-mRNA cap complex in a more indirect manner, thus moving the equilibrium

toward decapping. Regarding the mammalian counterpart Rck/p54, it was shown to play a similar role as activator of decapping and repressor of translation, although this last mediated by microRNAs[71].

Pat1 was the only yeast protein involved in the conserved decapping-dependent 5'-3' mRNA degradation pathway, that did not have identified the corresponding human homologue. This question was addressed during the work here presented, where the protein PatL1 was identified as the human counterpart. The biochemical features of Pat1 protein are mostly unknown. However, recent studies described the existence of two primary functional domains: one promoting translation repression, assembly of processing bodies (discussed bellow) and mediating interactions with the mRNA decapping factors LSm1-7, Dcp1 and Edc3, while the second domain promotes mRNA decapping. In addition, there are evidences that Pat1 protein directly binds to RNA through two independent binding sites localized in both domains, showing a preference to poly(U) substrates[72].

Purification assays showed that in yeast, Pat1 and the LSm1-7 ring form a complex even in the absence of RNA. This complex promotes decapping of deadenylated yeast mRNAs to which, consistently, shows an inherent affinity for oligoadenylated mRNAs[73]. In addition to its role as translation repressor and activator of decapping, Pat1 is also required for translation initiation[10,34,74]. In fact, Pat1 is the only decapping activator protein that, besides interacting with deadenylated mRNAs, also binds to eIF4G-, eIF4E- and Pab1p-associated polyadenylated mRNAs and locates to polysomes[75]. The function of Pat1 in the antagonistic processes of translation initiation and translation repression/mRNA degradation suggest a key regulatory role for this protein.

1.4 Processing bodies (P-bodies)

P-bodies are highly dynamic cytoplasmic structures formed by the self-assembly and disassembly of translationally repressed mRNPs. In these bodies it is possible to find components of the 5'-3' mRNA decay machinery, such as LSm1-7, Dhh1 and Pat1, as well as protein components of the ARE-mRNA decay, miRNA and NMD pathways. In opposition, translation initiation factors and ribosomes are generally missing in such foci what suits the observation that a mRNA must be translationally repressed before it can be assembled into P-bodies[34,76].

P-bodies are proposed to be required for the regulation of the interplay between translation repression and mRNA decay[56]. In a generally accepted model for P-body assembly, translational repressing factors promote the formation of a P-body "monomer" by recruiting decapping factors and/or other P-bodies assembly components to the translationally repressed mRNA. Since Dhh1 and Pat1 act both as translational repressors and decapping activators - thus playing a key role in the transition between translation repression and mRNA turnover[39] - they are likely among the first P-body factors recruited to a mRNA and, consequently, playing an important function in promoting the formation of a repressed mRNP that is competent to assemble into a P-body[65].

In the cases in which mRNA decay enzymes are not a limiting factor, mRNAs may be degraded rapidly, prior to P-body assembly. However, if decay enzymes are limiting, the repressed mRNAs assemble into P-bodies where they can undergo mRNA decay or released back into the translational pool[56,55,57]. Accordingly, P-bodies substantially decrease in size and number when the amount of mRNA to be decayed is reduced by, e.g., blocking transcription or translation elongation, or when Dhh1 or Pat1 proteins are mutated or depleted. In the other hand, P-bodies increase in size and number when the 5'-3' decay pathway is overloaded with RNA substrates or when mRNA decay is impaired, for example, by mutation of decapping factors or the Xrn1 exonuclease.

Although the observations that many mRNA decay enzymes concentrate in P-bodies and that the depletion of mRNA decay enzymes leads to accumulation of

mRNAs in these foci, there is still no clear evidence that mRNAs are degraded while assembled in P-bodies. Furthermore, recent studies failed to link the loss of visible P-bodies with defects in the mRNA decay and translational repression pathways[56,55,57]. However, it was suggested that P-bodies might avoid indiscriminate mRNA decapping and decay in the cytoplasm by maintaining mRNA decay enzymes away from translating mRNAs. Moreover, it was also proposed the possibility that P-bodies, by concentrating mRNA decay enzymes, enhance only mRNA decay pathways for which mRNA decay enzymes are limiting[7,76]. In any case, it is in this moment unclear what role, if any, does the assembly of translationally repressed mRNAs into P-bodies have in promoting translational repression, decapping and/or mRNA decay.

1.5 Positive-strand RNA viruses

Viruses are self-replicant organisms defined as obligate intracellular parasites as they require the host sub-cellular machinery and cellular factors to bypass their lack of an independent metabolism. The family of positive-strand RNA ((+)RNA) viruses constitutes an extensive genetic group of plant, animal, and human viruses that encompass over one third of all virus genera. The most relevant representatives for mammals are members of the families *Coronaviridae* (SARS-coronavirus), *Flaviviridae* (West Nile Virus, Hepatitis C Virus), *Picornaviridae* (Poliovirus, Hepatitis A), and *Togaviridae* (Venezuelan Equine Encephalitis virus, Rubella virus). These (+)RNA viruses are associated with syndromes such as colds, febrile illness, meningitis, paralysis, hepatitis and hemorrhagic fevers.

The different families of (+)RNA viruses, in order to multiply and disseminate their genes, have evolved various genome organizations, virion morphologies and life cycles that exploit the biology and biochemistry of their hosts in diverse ways. Albeit these differences all (+)RNA viruses share fundamental similarities concerning their RNA replication[77] (Figure 3). In this group of viruses, viral particles contain the single-stranded, messenger-sense RNA genome that once introduced into cells is replicated in their cytoplasm through RNA intermediates, without natural DNA forms. Since the viral RNA-dependent-RNA polymerase (RdRp) is not encapsidated,

upon entering in the cell, the genomes from (+)RNA viruses must be first translated by the cellular translation machinery, in order to replicate their genome. The products of translation include proteins such as the RdRp, structural components, and other viral replication factors involved in processes such as membrane targeting, template recruitment, RNA capping and evasion from host antiviral responses.

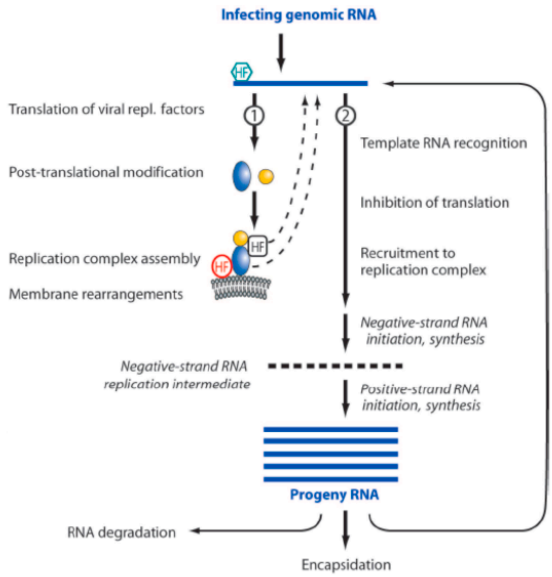

Figure 3. Schematic representation of the main steps in positive-strand RNA virus genome replication (discussed in the text). Arrows 1 and 2 indicate the sequential use of the infecting (+) genomic RNA as a template first for translation and then for RNA replication (Extracted from[78]).

Following translation and, when required, post-translational modifications, the viral genome is specifically recognized by viral factors and assembled together with the RdRp, further accessory non-structural proteins and host cell factors into a so-called replication complex. These replication complexes are always associated with virus-induced membrane structures derived from different cellular compartments, preferentially the endoplasmic reticulum, endosomes and mitochondrial membranes (reviewed in[77]). The exact function of such structures in viral RNA synthesis is yet not totally clear. However, they are thought to provide a physical structure where the replication complex can attach, confining viral RNA replication to a specific

cytoplasmic location where is possible to have simultaneously an increased concentration of essential components to this process and a way to divert the host defense mechanisms[79,80].

The genomic RNAs, once recruited to these membrane-associated replication complexes, serve as templates for the synthesis of negative-strand RNA replication intermediates that, by their turn, will be used as template for the production of progeny positive-strand genomic RNAs. Thus, for a successful infection to occur, the viral genomic RNAs have to be efficiently copied in order to serve as mRNA templates for additional rounds of replication and viral protein synthesis and, in addition, be specifically assembled into progeny virions.

One of the most interesting and puzzling aspects of (+)RNA viruses life cycle is the use of the same genomic RNA as a substrate for the competing processes of translation, replication and encapsidation. As these processes can not take place at the same time, mechanisms are required to regulate the sequential access of the respective machineries to a given RNA molecule, in order to avoid competition between them and allow the proper use of the genomic RNAs. Although the molecular features underlying these regulatory mechanisms are poorly understood, there is clear evidence that numerous interactions between viral *cis*-acting elements and host-encoded replication factors play an important role in these transitions[78,81]. The *cis*-acting signals that have been characterized in viral genomes mostly comprise not linear nucleotide sequences but RNA secondary structures such as bulged stem-loops, tRNA-like cloverleaves, and pseudoknots, which are thought to create distinctive three-dimensional molecular shapes that interact specifically with the viral and host factors[82]. These interactions permit to specifically target viral molecules from among thousands of cellular RNAs and selectively channel them to translation, replication and assembly complexes.

Along these lines it becomes clear the fundamental role that host factors play in several steps of viral life cycles. Accordingly to their importance, host factors are now being considered as promising targets in the development of new antiviral medicines. In fact, these novel strategies are becoming increasingly successful, and

include targeting factors such as cell surface receptors, protein kinases, nuclear receptors and proteosomes or exploiting the innate immune response system. These approaches are able to circumvent some of the main lacks that lead to the inefficiency some of the current therapies. The development and administration of live virus vaccines were able to keep under control important viral diseases, such as measles, small pox or polio. However, similar vaccines are ineffective aiming viruses with multiple serotypes, or with variable or complex antigenic structures. An alternative strategy is the development of chemicals able to target viral proteins, while sparing the cellular mechanisms. Although these antiviral drugs have notable effects controlling viral infections, they present very narrow spectra of activity and limited effect in eradicating chronic viral infections. Furthermore, their effectiveness is limited by the rapid appearance of antiviral resistances from viruses, such as HIV-1 and HCV, that rapidly mutate given their low-fidelity replication mode.

In contrast, the recent antiviral drugs targeting cellular pathways have the big advantage of being less susceptible to the emergence of drug resistances, since the human genes that encode the targeted cellular proteins are less likely to mutate in response to therapy (reviewed in[83]). Despite these approaches are presumed to be more toxic, it is reasonable to predict that there are cellular proteins unnecessary for cellular well-being, but critical to replication of viruses. In fact, it has been demonstrated that specific genomic changes can be well tolerated by mammals and yet offer strong protection against viral infections. For example, naturally occurring homozygous mutation in the human CCR5 (C-C-motif receptor 5) protein confers resistance to HIV infection[84], and Pfizer's (New York, NY) Selzentry/Celsentri (maraviroc) is an example of a recently approved drug targeting CCR5 for anti-HIV therapy[85].

In this manner, understanding the fundamental steps of virus life cycles, including the role of host factors at a molecular level, is essential for designing effective antiviral drugs. Furthermore, if host factors are identified that are required for multiple viruses, it will be possible to use to use them as targets to broad-spectrum antiviral agents. In this respect, the work performed during my thesis focused in the identification and characterization of specific host factors involved in the replication

of positive-strand RNA ((+)RNA) viruses. We centered our efforts in: (i) Brome mosaic virus, as it is a broadly used and established model system to study (+)RNA virus replication mechanisms and associated factors; and (ii) Hepatitis C Virus given its major clinical importance.

1.6 Brome Mosaic Virus

First isolated from brome grass (*Bromis inermis*), Brome Mosaic Virus (BMV) has the historically relevance of being the first RNA virus for which infectious genomic transcripts were generated *in vitro* from cDNA clones[86]. Furthermore, BMV is an established useful model system for studying different steps of (+)RNA viruses life cycles[87].

1.6.1 Organization of the viral genome

BMV is the type member of the icosahedral group of bromoviruses, belonging to the large alphavirus-like superfamily of plant and animal positive-strand RNA viruses. The genome of BMV consists of three RNAs with 5´-terminal m^7G-caps and a tRNA-like structure (TLS) located at the end of the 3´UTRs (Figure 4). The monocistronics BMV RNA1 (3,2Kb) and RNA2 (2,9Kb) encode, respectively, the essential viral replication factors 1a and 2a. The discistronic RNA3 (2,1Kb) encodes the cell-to-cell movement protein 3a and the coat protein, both required for systemic infection in BMV natural host but dispensable for RNA replication. The coat protein is translated not from the genomic RNA3 but from a subgenomic RNA4 generated during replication (reviewed in[81]).

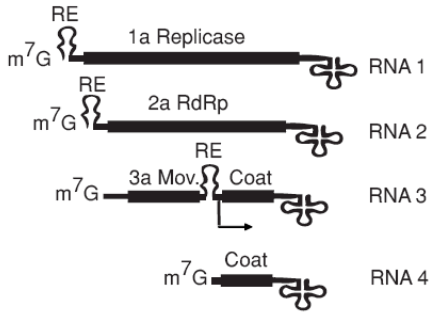

Figure 4. Schematic diagram of the BMV genome showing the ORFs (solid black boxes), the untranslated regions (UTRs) (single lines), the subgenomic mRNA start site (bent arrow) and the subgenomic RNA (RNA4). The BMV RNAs are capped (m^7G) and end in a tRNA-like structure (cloverleaf structure). The location of the recruitment element (RE) is shown for the genomic RNAs.

BMV 1a is a multifunctional protein (109kDa) with a NTPase/RNA helicase domain in the C-terminal half, and a RNA capping domain in its N-terminal half[88,89], while BMV 2a protein (94kDa) has a central domain with similarities to RNA-dependent RNA polymerases. BMV 1a protein localizes as a peripheral membrane protein in the endoplasmic reticulum (ER), inducing invaginations in which BMV RNA replication complexes are formed[90-92]. For the replication complexes to be active it is required the recruitment of BMV polymerase 2a and viral genomic RNAs to these sites at the ER[93,94]. The recruitment of BMV 2a results from the interaction of the 1a helicase-like domain with sequences in 2a protein N-terminus, while viral RNAs recruitment is also dependent on the 1a protein, as well as on *cis*-acting signals present in the viral genomes. During replication BMV 1a protein provides RNA capping, NTPase and possibly RNA helicase activities.

Following uncoating of the viral genomes within the cytoplasm, the natural replication cycle of BMV begins with the translation of the replication factors 1a and 2a. These proteins promote then the generation of viral replication complexes by inducing the formation of membrane patches in the ER. The transition between early translation and replication is regulated in a manner that allows sufficient synthesis of 2a polymerase proteins, thus securing efficient RNA replication. In association with host factors, the protein 1a recruits the BMV RNAs from the translation machinery and targets the 1a-2a-viral RNA complex to the membrane

replication sites. This targeting to the membranes is essential for BMV replication since the 2a polymerase is not functional in the cytosol before the assembly of the replication complex. In addition, the viral (+)RNA also plays a major role in the assembly and activation of the replication complex. The replication of BMV occurs in a asymmetric manner, with 100 positive BMV RNA strands being produced for every negative-strand. The transcription of subgenomic RNA4 to express the capsid protein (CP) occurs during the late stage of BMV infection. Subgenomic RNAs serve as translational templates, while genomic RNAs can be recruited for replication or encapsidation. The segmented BMV genomes are assembled within separate viral particles. The selectivity of the virion assembly was proposed to be assured by the specific interaction between RNAs and CP, where the highly conserved N-terminal arginine-rich motif of BMV CP is pointed as being responsible for both RNA binding and RNA packaging (reviewed in[81,87]).

In vitro and *in vivo* studies have identified and extensively characterized *cis*-acting signals present within BMV RNAs 1, 2 and 3 that are required for efficient RNA translation, replication, encapsidation and subgenomic RNA synthesis (reviewed in[81]). The 5'-untranslated regions from RNA1 and RNA2 share sequence similarity between them, with special note for a region denominated recruitment element (RE) (Figure 4). A striking feature of this element is a *cis*-acting motif matching the box B consensus sequence of RNA polymerase III promoter, which is subsequently transcribed into the TΨC stem-loop of tRNAs. Interestingly, this RE element exist also in BMV RNA3, however it is localized in the intergenic region of this RNA and not in the 5'UTR[95]. These RE elements are required for the proper 1a-mediated selection and recruitment of the viral RNAs from translation to the replication complex[95]. Moreover, there are partly overlapping sequences in 5'-UTR of all BMV genomic RNAs that control translation and initiation of positive-strand RNA synthesis[93,96].

The integenic region of RNA3 besides the RE element contains further *cis*-acting signals required for 1a-dependent recruitment of RNA3 into replication[97,98] and also the promoter for subgenomic RNA synthesis in the minus strand. This last consists of the 20 nucleotides upstream of the RNA4 initiation site. For full subgenomic

promoter activity there is an supplementary requirement for the oligo(A) tract located immediately upstream of the core promoter and additional AU-rich sequences[99,100].

The 3'UTRs of all BMV RNAs lack a poly(A) but have instead an highly conserved, highly structured tRNA-like structures (TLS) with key roles in BMV life cycle. These regions act as promoters for the synthesis of genomic-length negative-strand RNAs[101], and are also required for efficient translation, besides functioning as a nucleation site for coat protein assembly and encapsidation of the BMV RNAs into virions[102]. Interestingly the tRNA-like 3'ends terminate in 3'CCA_{OH} by action of (ATP, CTP):tRNA nucleotidyl transferase and are also specifically charged *in vivo* with tyrosine by cellular enzymes[103]. The biological meaning of this charging is not clear as it was shown that BMV proteins do not need tyrosylation for their initiation. However, the disruption of the TLSs strongly reduce the translation of the genomic RNA1 and RNA2, and less strongly RNA3[101].

The cumulative effect of the functional roles of these and, eventually, others yet to identify *cis*-acting elements, together with the interactions among them, as well as with cellular replication factors, coordinates the viral RNA functions for effective RNA replication.

1.6.2 BMV replication in yeast

The study of fundamental steps of virus life cycles, including the role played by host factors, has been deferred and challenging given the complexity of higher eukaryotic organisms. To circumvent the experimental difficulties associated with this complexity, systems were developed where higher eukaryotic viruses can direct RNA replication and gene expression in the yeast *Saccharomyces cerevisiae (S.cerevisiae).*

The yeast *S.cerevisiae,* given its facility to grow and be genetically manipulated, has been successfully used for many years as a model organism to unravel biological processes in higher eukaryotes. The complete yeast genome, known since 1996[104], comprises 6000 genes from which more than 60% have an assigned function. Noticeably, 40% of yeast genes share conserved amino acid sequences with at least

one known or predicted human protein. Moreover, 30% of human genes with a recognized involvement in human diseases have orthologs in yeast[105,106]. Due to this notable gene homology and the high conservation of fundamental biochemical pathways, studies in yeast have been used as model for study and understanding fundamental cellular processes such as mRNA translation and degradation, DNA repair mechanisms and the cell cycle[107,108].

The establishment of yeast/virus systems, allowing the replication of higher eukaryotic viruses in yeast, have made groundbreaking contributions for the dissection of the different steps from their life cycles such as the translation of viral proteins, the process of template selection, formation of the replication complex, viral RNA replication, encapsidation and recombination (reviewed in[87,109,110]). Furthermore, the use of traditional yeast mutational genetics and genome-wide screenings have been particularly successful in the identification of multiple host factors required for viral RNA replication and recombination of (+)RNA viruses. The list of higher eukaryotic viruses shown to replicate in yeast has been expanding and include viruses either with RNA or DNA genomes and that naturally infect plants or animals[111-118]. Importantly, these yeast systems reproduce the known features of those viruses replication cycles in their natural hosts. The first higher eukaryotic virus shown to replicate and encapsidate its genome in *S.cerevisiae* was the Brome mosaic virus[119,120].

BMV is able to recapitulate in the yeast *S.cerevisiae* all the steps of the life cycle as they occur in plants. The replication of BMV in yeast has been analyzed by following RNA3 or RNA3 derivatives. Yeast expressing the replicase proteins 1a and 2a, from DNA plasmids, are sufficient to support RNA3 or RNA3 derivatives replication and synthesize subgenomic mRNAs to express the coat gene[111,120,121]. Replicatable RNA3 derivatives can be introduced into yeast by transfection of *in vitro* transcripts, by *in vivo* transcription of RNA3 "launching cassettes" on yeast plasmids or integrated into yeast chromosomes. Most valuable for screening purposes has been the construction of RNA3 derivatives in which the coat gene was substituted with a reporter gene. Because its expression requires 1a-, 2a-directed negative-strand RNA3 and

subgenomic mRNA synthesis, the reporter gene provides a screenable marker for viral replication.

The segmented genome of BMV permits to study, independently, every single step of the replication process. For example, the study of recruitment of BMV RNA3 to the replication complex is performed in the absence of 2a polymerase and either expressing or not the 1a protein. Since the 1a protein is the only viral protein required for the template selection and for replication complex formation, when this protein is co-expressed with viral RNA3 it recruits this RNA to a high-stability membrane-bound state. This leads to an increase in the northern blot detected amounts of RNA3 when compared to the levels observed in the absence of 1a. The observed difference can be then correlated to effects on recruitment. By its turn, to study BMV RNA translation either BMV RNA1 or RNA2, with the original 5' and 3' ends, are expressed individually in yeast cells. Given that the protein 1a is not co-expressed the viral RNAs can not be recruited to the replication complex, meaning that the replication process stops at the translational step. Viral RNA translation efficiency is determined by normalizing viral protein levels detected by western-blot with the respective viral RNA levels detected via northern- blot.

The BMV/yeast system has also proven to be very fruitful in the identification of cellular factors involved in viral replication(reviewed in[81]). Using traditional yeast mutagenic analysis it has been possible to identify and characterize host genes that function in controlling different aspects of BMV life cycle. A more global and systematic approach to identify host factors affecting viral RNA replication was the use of a diploid yeast deletion library for a genome-wide screening. By this approach, approximately 4500 yeast deletion strains (~80% of all yeast genes) were screened and nearly 100 genes were identified whose absence either inhibited or stimulated BMV RNA replication and gene expression[122].

By using the BMV/yeast model system it was possible to identify the proteins LSm1, LSm6 and LSm7, which are subunits of the heptameric ring LSm1-7, and also Dhh1 and Pat1 as host factors playing an essential role in BMV genome replication[123-125]. The deletion of the respective genes diminishes the viral RNA recruitment from the

translation machinery to the replication complex. It was shown that the dependence of LSm1-7/Pat1/Dhh1 for BMV RNA recruitment was functionally linked to sequences in the 3'end, and proposed that this complex could mediate rearrangements in the BMV RNA that would facilitate the loss of ribosomes and translation factors, as well as the recognition of the 1a replicase. This would be consistent with the proposed function of LSm complexes as RNA chaperones facilitating multiple RNA-RNA and RNA-protein interactions and rearrangements during turnover and translation[1,3]. Additionally, it was also shown that this complex is required for the translation of the genomic RNA1, RNA2 and RNA3 but not of the subgenomic RNA4, the only BMV RNA that does not require to be recruited to the replication complex during the virus life cycle. Thus, it is apparent that the LSm1-7/Pat1/Dhh1 complex is able to function in apparently antagonistic processes, namely, virus RNA translation and recruitment to replication. Such role was suggested to allow rapid switching between processes in response to, for example, different host cell conditions or viral requirements[81]. However the molecular features underlying the roles of these proteins in BMV RNAs replication are still unknown.

1.7 Hepatitis C Virus

Estimated to infect 170 million people worldwide[126], the Hepatitis C Virus (HCV) is a major cause of chronic liver disease and liver transplantation in western Europe and United States[127]. Primary infections are often asymptomatic or initially associated with mild symptoms, but progression to chronic active hepatitis occurs in approximately 80% of infected individuals. The most severe manifestations associated with these chronic infections are liver cirrhosis and hepatocellular carcinoma[128].

The main obstacles to the eradication of HCV are the lack of an effective vaccine and the limitations associated with the actual treatments such as the combined therapy using ribavirin, a synthetic guanosine analogue, and a pegylated form of interferon[129]. This therapy does not achieve a sustained viral clearance in more than 50% of treated patients, and is associated with relevant side effects[130,131]. Thus, there is an urgent need to define approaches that lead to the development of more safe

and effective therapeutic strategies. In this regard, although new drugs designed to target certain viral proteins or their functional epitopes are currently being tested in clinical trials, the actual leading line of research in HCV field is the identification of cellular factors required for viral replication and the evaluation of their potential as antiviral targets[132]. Actually, various host-targeted compounds are already being tested in clinical trials as it is the case of *NA255,* that prevents de novo synthesis of sphingolipids, the major lipid rafts components, inhibiting HCV replication by disrupting anchoring of HCV nonstructural proteins on the lipid rafts[133]. As discussed, these host-targeted compounds have predictably the advantage of being less likely to promote the development of escaping mutants.

1.7.1 Organization of the Viral Genome

Hepatitis C virus (HCV) is an enveloped virus with a single positive-strand RNA genome, classified as a member of the genus *Hepacivirus* within the family Flaviviridae. The HCV genome is an uncapped linear molecule with a length of approximately 9.6kb, that carries a long open reading frame encoding a single polyprotein precursor, flanked by highly structured 5'- and 3'-untranslated regions (UTR's) essential for RNA translation and replication[134,135] (Figure 5).

Expression of the HCV viral proteins from the monocistronic genome, is primarily accomplished by the 5'UTR-mediated translation of the ~3000 aminoacid-long polyprotein precursor and its subsequent co- and post-translational processing. The polyprotein is proteolytically cleaved by both cellular and viral proteins into the structural proteins (core, envelope proteins E1 and E2), the hydrophobic peptide p7, and the non-structural (NS) proteins, NS2, NS3, NS4A, NS4B, NS5A and NS5B[134,135].

Core, E1 and E2 are the major components of the virus particles. The core is a RNA-binding protein that forms the viral nucleocapsid and functions in the regulation of HCV translation, particle assembly, cell signaling, apoptosis, carcinogenesis and lipid metabolism[136]. The ER-targeted type I transmembrane proteins E1 and E2 are the main components of the virion envelope, and are known to function in protein folding and virus entry. The protein p7 is a member of the "viroporin" family of proteins shown to form an ion-channel in cellular membranes,

and to be essential both for infection and a late step in virus assembly and release[137,138]. Regarding the non-structural proteins, NS2 is a membrane-associated cysteine protease, not required for replication but which protease domain is essential for viral assembly and infectious virus production[139,140]. The remaining NS proteins components of the HCV replication complex and are known to be sufficient for RNA replication in cell culture, as NS3-3'UTR subgenomic replicons efficiently replicate in a hepatic cell line[141,142]. The NS3 is a serine-type protease with nucleotide triphosphatase and helicase activity. NS4A is a cofactor for the NS3 serine protease and is also required for tethering the NS3 domain to intracellular membranes. NS4B induces the formation of membranous vesicles where the replication complexes are assembled, while NS5A is a phosphoprotein with RNA-binding properties, component of the viral replication machinery and also involved in assembly[133]. Finally, NS5B is the RNA-dependant RNA polymerase (RdRp) responsible for the synthesis of the HCV RNA genome (reviewed in[143]).

The polyprotein coding region is flanked by highly structured 5' and 3' untranslated regions that are among the most conserved segments of the viral genome and that include regulatory *cis*-acting elements that modulate different steps of the virus life cycle[134,135,143].

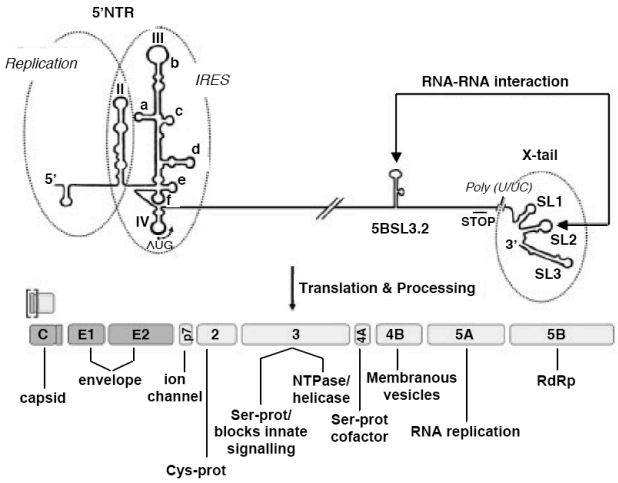

Figure 5. A schematic of the HCV genome with the secondary structure of *cis*-acting RNA elements is shown in the upper panel. The polyprotein with individual cleavage products is drawn below. (Adapted from[134,135,143])

The 5' UTR is 340nt-long and contains an internal ribosome entry site (IRES) that mediates cap-independent translation initiation of the HCV genome (Figure 5)[144]. This RNA element is composed by three stem-loops domains denominated SLII, SLIII and SLIV. The SLIII, by itself, consists of branching hairpin stem-loops (IIIa, -b, -c, -d, -e, -f) organized in 3- and 4-way junctions. The domains II, IIIcdef and IV bind the 40S ribosomal subunit in a way reminiscent of prokaryotic translation initiation, while the apical domains IIIab provide a platform for eIF3 (reviewed in[134]). Part of the IRES (domain II) overlaps with RNA signals essential for viral replication, which include also SLI, arguing for a possible role of domain SLII in regulating a translation-RNA replication switch[144]. Between SLI and SLII there is a spacer region (nt 21 to 43) that contains two target sequences for a liver-specific microRNA (miR122), which was shown to stimulate HCV translation by promoting the association of the small ribosomal unit with the HCV RNA[145]. Furthermore, already in the core protein-coding region of HCV, it was described the presence of highly conserved secondary RNA structures and a +1 alternative reading frame (ARF), in which was shown to exist a RNA element functionally important for replication[146,147].

The 3' UTR is required for initiation of negative synthesis and also acts as translation enhancer. It consists of three structural domains: an approximately 40-nt long variable region downstream the HCV coding-sequence, a internal poly(U/UC) tract of heterogeneous length, and a highly conserved 98-nt-long sequence designated X-tail that is composed of three stem-loops named 3'SL1-SL3 (Figure 5). The variable region is required for RNA replication and, together with the poly(U/UC) tract and part of the X-tail, enhances IRES-dependent translation (reviewed in[134,135]). An additional *cis*-acting RNA element has been identified in the 3' terminal coding region of NS5B. This element (5BSL3.2) forms a long distance RNA-RNA interaction with 3'SL2 in the X-tail, designated "kissing-loop", which is indispensable for RNA replication[148,149]. The poly(U/UC) tract beside functioning as a spacer between the kissing-loop interaction elements, it was also shown to be important for RNA accumulation in a length and sequence dependent manner. Finally, the 3' X-tail is essential for replication in cell culture and infectivity *in vivo*[150].

1.7.2 HCV replication cycle

Replication of Hepatitis C virus takes place in the cytoplasm of the host hepatocytes. HCV particles bind to the host cell via a specific interaction of the viral E2 glycoprotein with cellular receptors such as CD81, SR-B1 and Claudin-1 which were shown to be linked to HCV entry[151-153]. Other co-receptors candidates have been identified and include C-type lectins and the LDL receptor[154]. Bound particles enter the cell via receptor-mediated endocytosis in a clatherin-dependent manner, followed by a fusion step from within an acidic endosomal compartment[155]. Upon infection and release of the viral RNA into the cytoplasm, HCV proteins are generated by means of the IRES-mediated translation. Translation takes place at the rough endoplasmic reticulum where host cell signalases and signal peptide peptidases proteolitically cleave the generated polyprotein precursor between the core and p7 regions. The processing at all other sites within the polyprotein is performed by two viral enzymes, NS2 and NS3/4A proteases. Formation of membranous vesicles that accumulate to form a "membranous web" is induced by NS4B perhaps in collaboration with other viral and host cell factors. NS5B RdRp apparently assembles - in concert with other viral proteins, such as NS3/4A and NS5A, and host cell factors - at the 3'end of positive strand RNA to initiate de novo synthesis. Following, the newly generated negative strand RNA then serves as template for production of excess amounts of positive strands. These newly synthesized (+)-strand RNAs are either used for production of new polyprotein copies, a new round of RNA replication, or are incorporated into nucleocapsids what is assumed to occur in at the ER membrane (reviewed in[134,135]). The HCV core protein is associated with the surface of lipid droplets and the endoplasmic reticulum membranes closely surrounding these droplets, and its self-assembly drives virion budding[156]. Finally, virus particles are released out of infected cell via the constitutive secretory pathway[157].

Thus, similarly to all other (+)RNA viruses, HCV genomic RNA serves as substrate to essential and mutually exclusive processes of the virus life cycle such as acting as mRNA for translation of viral proteins, as template for minus-strand synthesis during replication and as progeny genomes for assembly of viral particles. Little is known about host and cellular factors that regulate the transit of HCV RNA from

translation to replication. However, strong evidence indicates that NS5A phosphorylation status serves as a molecular switch in the regulation of HCV RNA replication by affecting the association between NS5A and other components of the viral replication complex[158]. Additionally, the host NFAR proteins were shown to interact with regulatory regions present in both 5' and 3'UTRs of HCV genome and to promote its circularization. Curiously these proteins were also observed to stably interact with NS5A and a role in the switch between viral translation and viral replication was proposed[159]. Furthermore, recent studies showed that the protein SYNCRIP has dual functions, participating in both RNA replication and translation in HCV life cycle[160]. Given the involvement of several cellular proteins in the different steps of HCV life cycle, it is predictable that additional host factors play an key roles in the regulation of viral RNA fates.

1.7.3 Model systems for studying HCV life cycle *in vitro*

The lack of reliable cell culture systems supporting HCV RNA replication and particle formation were, for a long time, the main obstacle to the decipherment of the functional roles played by viral and cellular factors in HCV replication. Several efforts were made to propagate HCV in cell culture, but these attempts were entirely unsuccessful for long time[142]. However, we assisted in this decade to a impressive development in this area, starting with the first establishment of subgenomic replicons in 1999[141], up to the development, in 2005, of efficient systems capable of producing infectious virus particles in cell culture[161-163].

HCV subgenomic replicons are *in vitro* transcripts of a cloned DNA copy of the full-length viral genome in which the sequences coding the structural proteins, that are not required for RNA replication, were substituted by selectable markers or reporter genes. These replicons are able to efficiently replicate in the human hepatoma cell line Huh7[141]. However, although the introduction of cell culture-adaptive mutations in the HCV replicon allowed the generation of efficiently replicating full-length HCV genomes, virus production has not been observed[164-166]. This restraint has been circumvented with the JFH-1 isolate, a cloned genotype 2a consensus genome that was isolated from a Japanese patient with fulminant hepatitis[161-163,167]. Without requiring the introduction of cell culture-adaptive

mutations, the subgenomic JFH-1, when transfected into Huh-7 cells, can replicate at a much higher levels and lead to the production of HCV particles that are infectious both for naive cells and for chimpanzees.

The establishment of these HCV cell culture systems made possible to deep characterize HCV viral proteins, and to develop different assays to identify and study HCV-host proteins interactions influencing translation and RNA replication. The functional relevance of these proteins can be investigated by different methods, including RNA interference (RNAi). The sequence-specific gene silencing using RNA interference is a recent but already invaluable research tool. The highly effective targeting of either host or viral RNAs by specific siRNAs permits the characterization of gene functions, increases our understanding of virus-host interactions and potentially can reveal novel antiviral targets. The constant progress of RNAi techniques, together with genomics, allowed the development of genome-wide screening technologies, that systematically examine the human genome for the genes and pathways that are required by human pathogenic viruses such as *Human immunodeficiency virus* (HIV)[168,169], *Hepatitis C virus* (HCV)[171,170], *West Nile virus* (WNV)[172] and *Influenza A virus*[173].

These studies were able to identify host factors and associated pathways previously implicated in viral infection or pathogenesis, what supports the validity and accuracy of such screens. More importantly, they identified new genes and pathways that define virus-host interactions. In the concrete case of HCV, several host genes that modulate HCV replication and infectious virus production have been recently identified, however their specific role is mostly unknown[171,170].

Despite the obvious advantages and achievements associated with the use of RNAi screens, it should be taken into account the inherent obstacles of the technique. This includes off-target effects of siRNAs, incomplete knockdown of expression, and sensitivity limitations of the screening assay. Nevertheless, these limitations can be overcome with careful validation studies. When combined with biochemical and virological studies to validate and define gene function in supporting infection, and with improved bioinformatics analysis to map cellular

pathways exploited by viruses, the data extracted from RNAi screenings can provide a powerful approach to identify druggable factors for the treatment of viral infection.

2. Aims

By using the ability of the (+) RNA virus BMV to replicate in yeast it was previously shown that subunits of the LSm1-7 ring, as well as Pat1 and Dhh1 play an essential role in the transit of the BMV genome from translation to replication[123-125]. In non-infected cells, these proteins mediate the transition of cellular mRNAs from a translational to a non-translational state by activating decapping in the 5'-3' – deadenylation-dependent mRNA decay pathway[39]. Given the conservation of this pathway from yeast to humans and the common need of all (+)RNA viruses to regulate the transition of their genomes from active translation to a translationally inactive state to allow replication, an exciting possibility, and our work hypothesis, is that the function of LSm1-7, Dhh1 and Pat1 is used not only by BMV to replicate in yeast but also by other (+) RNA human viruses, such as HCV, to replicate in mammalian cells. Furthermore, given the key role of these proteins in a common step to all (+) RNA viruses, it is essential to characterize the not yet defined molecular mechanisms associated with such function. So, the **specific aims** of these thesis are:

1. Characterize the molecular features underlying the role of LSm1-7 ring in the translation and replication of Brome Mosaic Virus

2. Evaluate the role of P-bodies components LSm1-7, Rck/p54 and PatL1 in the replication of Hepatitis C Virus

3. Results

In this section scientific results are presented in two parts, both of them containing research papers and appendices. These appendices include additional data relevant for future studies and that complement the data presented in the papers.

Part I

Characterization of the molecular features underlying the role of LSm1-7 in the translation and replication of Brome Mosaic Virus

Publication I

LSm1-7 complexes bind to specific sites in viral RNA genomes and regulate their translation and replication. *RNA* ***16(4):*** *817-27 (2010).*

Appendices

Part II

Evaluation of the role of P-body components LSm1-7, Rck/p54 and PatL1 in the replication of Hepatitis C Virus

Publication II

Identification of PatL1, a human homolog to yeast P body component Pat1. *Biochim Biophys Acta* **1773:** 1786-1792. (2007)

Publication III

Translation and replication of hepatitis C virus genomic RNA depends on ancient cellular proteins that control mRNA fates. *PNAS* **106:** 13517-13522. (2009)

Appendices

3.1 Part I

Characterization of the molecular features underlying the role of LSm1-7 in the translation and replication of Brome Mosaic Virus

3.1.1 Publication I

Galão RP, Chari A, Alves-Rodrigues I, Lobão D, Mas A, Kambach C, Fischer U, Díez J. LSm1-7 complexes bind to specific sites in viral RNA genomes and regulate their translation and replication. *RNA* ***16(4):*** *817-27 (2010).*

The cellular LSm1-7 ring is required for both the translation and recruitment of the BMV genome. Yet, how this complex act mechanistically on the BMV RNAs remains elusive. My contribution to this manuscript, from which I have the first authorship, was to address this issue experimentally by establishing gel-shift assays employing recombinant LSm1-7 complexes and BMV viral transcripts. Using this *in vitro* binding approach I was able to provide evidence of a direct interaction between the LSm1-7 complexes with specific sites in BMV RNA genomes, as well as define previously unknown RNA-binding patterns for this complex (Figures 3-6). Further *in vivo* analysis showed that these LSm1-7 RNA target sequences act as regulators of translation and recruitment of the BMV genome suggesting that the intrinsic RNA-binding characteristics of the LSm1-7 complexes determine their function.

Galão RP, Chari A, Alves-Rodrigues I, Lobão D, Mas A, Kambach C, et al. LSm1-7 complexes bind to specific sites in viral RNA genomes and regulate their translation and replication. RNA. 2010; 16(4): 817-27.

3.1.2 Appendix I:

High amounts of recombinant LSm1-7 ring lead to formation of higher-order RNPs when interacting with BMV viral RNAs *in vitro*.

Previously, by using electromobility shift assays, we showed that reconstituted recombinant functional LSm1-7 rings are able to strongly interact with the 3'UTR of all BMV RNAs and with the intergenic region of BMV RNA3, reflected by a complete bandshift of the corresponding radiolabeled transcripts in the presence of this complex (Publication I). In order to determine the binding affinities of LSm1-7 complex with its viral RNA targets, we incubated BMV RNA3 3'UTR and BMV RNA3 IR ^{32}P-labeled transcripts with increasing amounts of LSm1-7 reconstituted complex. Curiously, the increasing amounts of LSm1-7 resulted in a gradual shift of the RNAs (Figure S1), rather than (a) distinct step(s) as observed in similar experiments performed by Bergman and colleagues using this same complex[32].

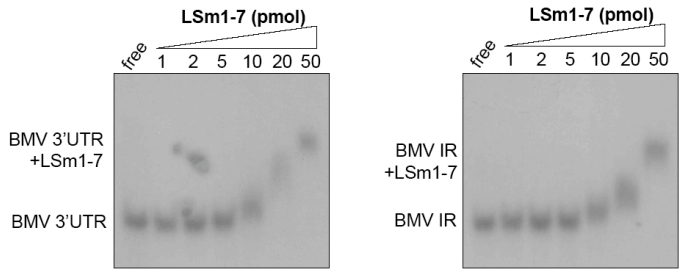

Figure S1. LSm1-7 complex forms higher-order structures when bounded to BMV RNA3 3'UTR and BMV RNA3 IR. Increasing amounts of LSm1-7 complex were incubated with radiolabeled BMV RNA3 3'UTR (left panel) or BMV RNA3 IR (right panel) and, after complex formation, products were separated on a non denaturing polyacrylamide gel. Gels were autoradiographed at -80°C in maximum sensitivity films.

These observations, together with the migration patterns presented by the same transcripts in competition experiments (Figure 6, publication I), suggest that binding of LSm1-7 complex to these viral RNAs leads, at least *in vitro*, to the formation of higher-order RNPs complexes involving an unknown number of LSm rings. This is supported by reports showing that LSm1-7 rings can rearrange into alternate higher-order structures in solution, principally dimers and occasionally trimers[174], and that

high-order complexes are frequently observed in experimental conditions in which RNA concentrations are limiting, as occurs in our electromobility shift assays[37]. Thus, the higher-order RNPs observed in our assays could be due to multiple rings interacting directly with the same transcript and/or the result of LSm1-7 rings being able to aggregate between them even when interacting with a target RNA. In any case, these findings precludes the simple determination of Kd values for the interaction of LSm1-7 with BMV transcripts. Similar results were obtained with a HCV 5'UTR transcript (data not shown).

In contrast, Bergman and colleagues were able to define Kd values for the binding of similar LSm1-7 recombinant preparations with RNA transcripts containing 5' poly (A) tracts. The different capacity to form higher-ordered RNPs can be due to i) the impossibility of more than one LSm1-7 ring to interact directly to the 5' poly(A) transcript, in opposition to what may happen in the binding to the tested BMV transcripts (e.g. BMV RNA3 IR binds to LSm1-7 through the loops L1 and L2), or ii) the probable existence of two binding sites with different characteristics in the LSm1-7 ring surface, eventually in opposing faces of the complex. This statement is supported by our additional data showing that HCV 5'- and 3'-UTRs bind to at least two different sites within the LSm1-7 complex (Part II – Appendix I). In such scenario, the transcripts used by Bergman could be interacting with LSm1-7 complex in a region different from the one used for the binding to BMV transcripts. If this region is simultaneously required for protein-protein interactions, it would interfere with the aggregation of additional LSm1-7 rings and consequently avoid the formation of higher-order complexes. A more detailed knowledge of the structure and mode of binding of the LSm-RNA complex is required to address these important points.

3.1.3 Appendix II:

LSm1-7 complexes bind to a ß-globin ARE mRNA through the same binding site used to interact with BMV 3'UTR and BMV IR

The cellular function of LSm1-7 ring is well characterized in yeast. This complex acts together with Dhh1 and Pat1 as decapping activators and they are believed to facilitate the exit of cellular mRNAs from active translation to degradation[6,39,69,175]. In human cells the function of the respective homologues has mainly been inferred from (i) the yeast work, (ii) their shared localization in P-bodies and (iii) their interactions with proteins of the 5´-3- deadenylation- dependent mRNA decay pathway. However, it has been described that human LSm1 promotes degradation of cellular mRNAs with special features such as mRNAs containing AU-rich elements (ARE) in their 3´UTR[7]. The mechanism by which LSm1 influences the fate of these specific mRNAs is unknown. To explore how LSm1 may exert its function we tested whether the LSm1-7 complex could specifically interact with the 3´UTR of the β-globin-ARE mRNA, the reporter mRNA used in the study of Stoecklin and colleagues. Indeed, LSm1-7 bound to the 3´UTR of β-globin only if it contained the ARE element. Thus, binding correlated with LSm1-7-dependent degradation (Figure S2a).

Following, we performed competition assays in which ^{32}P-labeled BMV RNA 3'UTR and BMV RNA3 IR transcripts were incubated with the LSm1-7 complex in the presence of increasing amounts of unlabeled 3'UTR from ß-globin ARE mRNA (Figure S2b, upper panel). LSm1-7 binding ß-globin ARE mRNA totally competed LSm1-7 binding to both BMV 3' UTR and BMV IR. In a reversed experimental set-up, the binding to both BMV 3'UTR and BMV IR completely competed LSm1-7 binding to the 3'UTR from ß-globin ARE mRNA (Figure S2b, lower panels). These observations suggest that BMV RNA3 3'UTR, BMV RNA3 IR and ß-globin ARE mRNA bind to the same site within the LSm1-7 complex.

Furthermore, the finding that a recombinant human LSm1-7 complex to is able to directly interact with a cellular mRNA and with (+)RNA viral genomes from HCV, a human virus (Publication III), and BMV, a plant virus capable of replicate also in

yeast, suggests a functional conservation of this complex across kingdoms, not only in their cellular roles but also in their function in the replication of (+)RNA viruses. To corroborate this statement one should not forget also that Hfq, the bacterial counterpart of LSm1-7 complex, binds its targets following similar binding patterns and play equal roles in the life cycles of a (+)RNA virus.

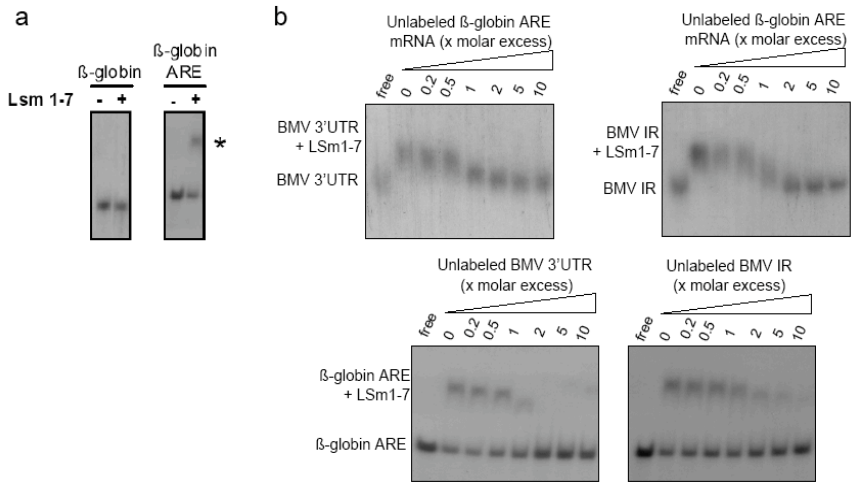

Figure S2. LSm1-7 complexes bind to a 3´UTR of β-globin mRNA that contains an ARE element through the same binding site used to bind to BMV 3'UTR and BMV IR. (a) Gel-shift analysis were performed using LSm1-7 complexes and radiolabeled transcripts of the 3'UTR of ß-globin mRNA, with or without a 51-nt long ARE-element derived from the murine GM-CSF gene (ß-globin ARE and ß-globin, respectively; (Stoecklin et al., EMBO Rep (7) 72 (2006)). The asterisk marks the position of the RNA transcripts bound to LSm1-7 complexes. (b) Labeled BMV 3'UTR and BMV IR transcripts were incubated with LSm1-7 complexes in the presence of increasing amounts of unlabeled 3'UTRs ß-globin-ARE RNA as competitor (upper panels) or radiolabeled 3'UTR of the ß-globin-ARE RNA transcripts were incubated with LSm1-7 complexes in the presence of unlabeled transcripts of either BMV 3'UTR or BMV IR (lower panels).

3.2 Part II

Evaluation of the role of P-body components LSm1-7, Rck/p54 and PatL1 in the replication of Hepatitis C Virus

3.2.1 Publication II

Scheller N, Resa-Infante P, de la Luna S, **Galão RP**, Albrecht M, Kaestner L, Lipp P, Lengauer T, Meyerhans A, Díez J. Identification of PatL1, a human homolog to yeast P body component Pat1. *Biochim Biophys Acta* **1773:** 1786-1792. (2007)

The 5'-3'-deadenylation-dependent decapping pathway of mRNA degradation is well conserved from yeast to humans. With the exception of Pat1, all the corresponding human homologues have been identified and shown to localize to P-bodies as do their yeast counterparts ([57] and references therein). In this research paper we described the identification of PatL1 as the human homolog of the yeast protein Pat1. PATL1 mRNA was shown to be ubiquitously expressed in all human tissues, as well as LSM1 and RCK/p54 mRNAs. Furthermore, as expected for the homolog of yeast Pat1, we observed that PatL1 protein accumulated in P-bodies and is required for its formation. My personal contribution for this manuscript was the setting up of the silencing conditions for the depletion of the proteins PatL1 and Rck/p54 within the cells and evaluate its effect in P-body formation (Figure 4).

Scheller N, Resa-Infante P, de la Luna S, Galão RP, Albrecht M, Kaestner L, et al. Identification of PatL1, a human homolog to yeast P body component Pat1. Biochim Biophys Acta. 2007; 1773(12): 1786-92.

3.2 Part II

Evaluation of the role of P-body components LSm1-7, Rck/p54 and PatL1 in the replication of Hepatitis C Virus

3.2.2 Publication III

Scheller N*, Mina LB*, **Galão RP***, Chari A, Giménez-Barcons M, Noueiry A**,** Fischer U, Meyerhans A, Díez J.Translation and replication of hepatitis C virus genomic RNA depends on ancient cellular proteins that control mRNA fates. *PNAS* **106:** 13517-13522 (2009). * *Equal contribution*

Given the conservation of the 5'-3'-deadenylation-dependent mRNA decay pathway from yeast to humans and the common need of all (+) RNA viruses to regulate the transition of their genomes from active translation to a translational inactive state to allow replication, an exciting possibility is that the function of Dhh1, LSm1-7 and Pat1 is utilized not only by BMV to replicate in yeast but also by human viruses to replicate in human cells. In this paper we showed that indeed the respective human homologues, namely Rck/p54, LSm1-7 and PatL1, are necessary for HCV replication. We also showed that the requirement of these proteins for efficient HCV RNA translation was linked exclusively to the 5'- and 3' untranslated regions (UTRs) of the viral genome. My main personal contribution to this manuscript was to give further insight into the model of action of LSm1-7 on HCV life cycle. By performing in vitro electromobility assays with reconstituted LSm1-7 rings, I was able to show that this complex specifically binds to defined sequences in the 5'- and 3' UTRs that are known to play key roles in the regulation of HCV translation and replication (Figures 3 and S5).

Scheller N, Mina, LB, Galão RP, Chari A, Giménez-Barcons M, Noueiry A, et al. Translation and replication of hepatitis C virus genomic RNA depends on ancient cellular proteins that control mRNA fates. Proc Natl Acad Sci U S A. 2009; 106(32): 13517-22.

3.2.3 Appendix I:

HCV 5'- and 3-'UTRs bind to at least two different sites within the LSm1-7 complex

As previously described in Part I of this "Results" section, LSm1-7 complex strongly binds to areas in Brome mosaic virus (BMV) RNAs that contain important regulation signals for translation and recruitment. In particular, we showed interactions with the 3' UTRs of all three genomic RNAs and with the intergenic region (IR) of BMV RNA3 (Publication I). Concerning cellular mRNAs, no specific binding target for the LSm1-7 complex has been published yet. However, as we described in Part I – Appendix II (Fig S2a), LSm1-7 complex specifically binds an ARE-containing β-globin mRNA, through the same binding site(s) used to bind BMV RNA3 3'UTR and BMV IR.

In order to address if the LSm1-7 complex directly interacts with HCV RNAs through the same binding sites used to bind the upper mentioned transcripts, we performed competition assays in which ^{32}P-labeled HCV 5'UTR or HCV 3'UTR transcripts were incubated with the LSm1-7 complex in the presence of increasing amounts of either unlabeled 3´ UTR from β-globin-ARE mRNA, BMV RNA3 3'UTR or BMV RNA3 IR. An unlabeled self transcript was used as a positive control for competition and two non-binding RNAs, the 3´ UTR from β-globin mRNA and the BMV RNA1 5´UTR, were used as negative controls (Figures S1b and S2a). Both BMV RNA transcripts tested, as well as the β-globin-ARE mRNA, competed LSm1-7 binding to the HCV regions. However, these competitions were not complete. In a reversed experimental set-up, the binding HCV RNA regions completely competed LSm1-7 binding to the 3´ UTR β-globin-ARE mRNA and the BMV RNA regions (Figure S1c and S2b). Together these data suggest that both HCV RNAs bind to at least two different sites within the LSm1-7 complex while only one of these is utilized by the BMV RNA3 3'UTR, the BMV RNA3 IR and the β-globin-ARE mRNA.

The existence of different binding sites in LSm1-7 complexes, together with the fact that LSm1-7 and the related Hfq complex present similar patterns of binding, may

indicate that LSm1-7 ring also has two independent RNA-binding surfaces, as occurs with Hfq[36]. RNA-binding sites at Hfq complex are present in both proximal and distal faces allowing the simultaneous binding of two RNA targets. This finding can have particular importance to understand not only how LSm1-7 ring interact with their different RNA targets, but also give insights into the molecular roles associated to their function in cellular mRNAs and viral life cycles.

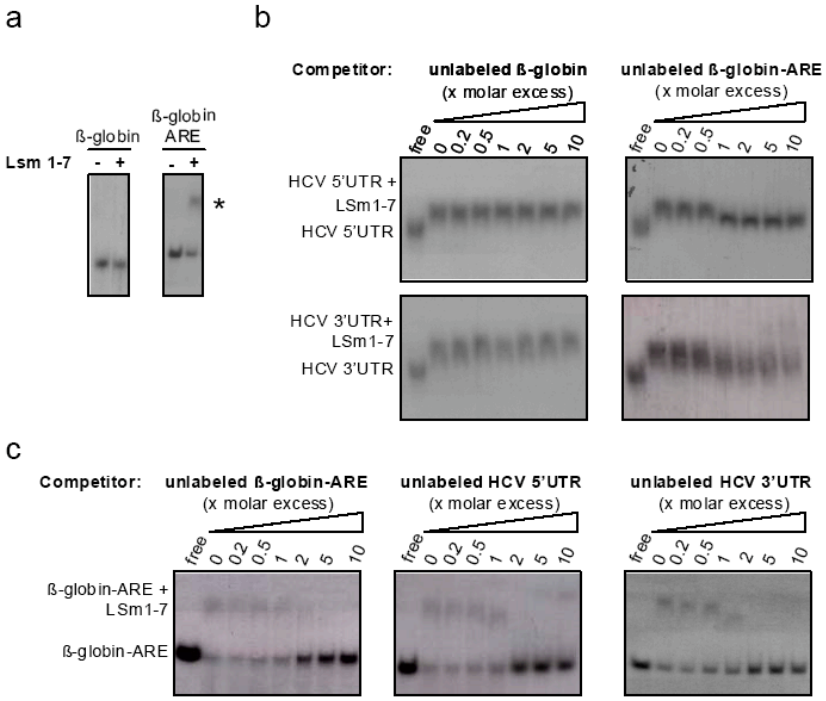

Figure S1. Interaction of LSm1-7 to a β-globin mRNA that contains an ARE element partially competes the binding of LSm1-7 to HCV regions.

(a) Gel-shift analysis were performed using LSm1-7 complexes and radiolabeled transcripts of the 3'UTR of ß-globin mRNA, with or without a 51-nt long ARE-element derived from the murine GM-CSF gene (ß-globin ARE and ß-globin, respectively; (Stoecklin *et al.*, EMBO Rep (7) 72 (2006)). The asterisk marks the position of the RNA transcripts bound to LSm1-7 complexes (b) Labeled HCV 5' (upper panels) and HCV 3'UTR (lower panels) transcripts were incubated with LSm1-7 complexes in the presence of increasing amounts of unlabeled 3'UTRs of ß-globin or ß-globin-ARE RNA as competitors. (c) Radiolabeled 3'UTR of the ß-globin-ARE RNA was incubated with LSm1-7 complexes in the presence of increasing amounts of unlabeled transcripts of either itself, HCV 5'UTR or HCV 3'UTR.

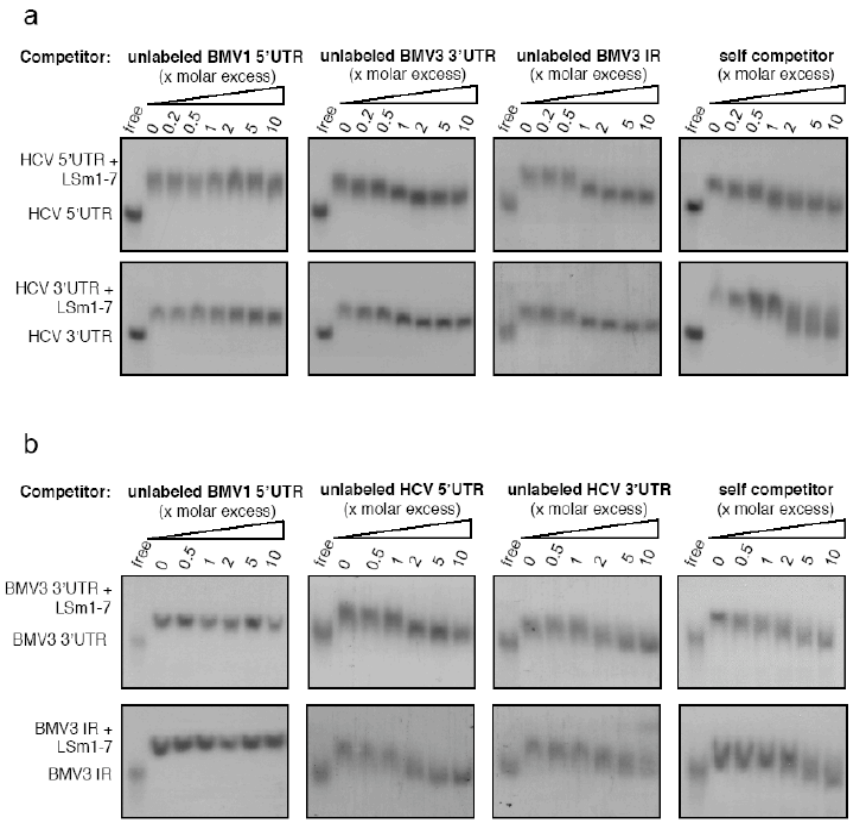

Figure S2. Binding of Lsm1-7 complexes to the 5'UTR and 3' UTR of HCV is competed by the 3'UTR and the Intergenic Region (IR) of BMV RNA3.

(a) Radiolabeled HCV 5'UTR (upper panels) and HCV 3'UTR (lower panels) were incubated with reconstituted LSm1-7 complexes in the presence of increasing amounts of either unlabeled 3'UTR or IR of BMV RNA3, or unlabeled transcripts of itself, as competitors. The non-binding 5´UTR of BMV RNA1 was used as a non-competing control. (b) Labeled 3'UTR (upper panels) or IR (lower panels) of BMV RNA3 were incubated with LSm1-7 complexes and increasing amounts of unlabeled 5'UTR of BMV RNA1, 5' UTR or 3'UTR of HCV, or unlabeled transcripts of itself.

3.2.4 Appendix II:

Deletion of single hairpin stem-loops at SLIII does not disrupt LSm1-7 binding to HCV IRES

The electro-mobility shift analysis presented in publication III, showed that SLIII was both necessary and sufficient for binding of the LSm1-7 ring to the 5'UTR region of HCV genome, as evidenced by binding to a RNA motif containing exclusively the corresponding positions 129-290 (highlighted in Figure S3, left panel) and by loss of binding upon removal of it. SLIII displays distinct 40S subunit and eIF3 interactions sites, thus playing key role in translation initiation. 40S subunit binds at the basal domain IIIdef and domain IIIc, while apical domains IIIab provide a platform for eIF3 binding (reviewed at [144]). Additionally, NFAR proteins were also shown to bind to this region indicating that SLIII functions in viral replication[159]. The potential overlapping of LSm1-7, NFAR proteins and eIF3 binding sites can have interesting molecular implications concerning translation regulation.

To better understand which are the exact binding site(s) of LSm1-7 complex within the SLIII, we decided to delete the different hairpin stem-loops present in this region and evaluate its effect in the binding of the complex to HCV 5'UTR. We focused in the hairpin stem-loops because of the existence of ssRNA stretches in these areas that, given the defined binding patterns for this complex (Publication I), may be important for LSm1-7 binding to HCV IRES. For these band-shift assays, new transcripts were generated using HCV 5'UTR transcript (pos 1-400) as backbone, from which stretches related to the different hairpin stem-loops were deleted: ΔSLIIIa (positions 154 to 171 deleted – (Δ154-171)), ΔSLIIIb (Δ191-206), ΔSLIIIc (Δ228-238) and ΔSLIIId (Δ249-281). Additionally, two more transcripts were generated: ΔSLIIIa/c, in which IIIa and IIIc were simultaneously deleted (Δ154-171 + Δ228-238), and Δknot, where a knot present in the apical part of the stem loop was removed (Δ181-187 + Δ210-218) (Figure S3, left panel). Curiously, the deletion of any of these segments affected the binding of LSm1-7 ring to HCV 5'UTR, as shown in Figure S3, right panel.

The data obtained in this set of band-shift assays do not define the position(s) where LSm1-7 ring binds to SLIII. Although, the findings presented here may indicate: i) the existence of two binding sites within the SLIII, where the removal of one of them would not provoke the disruption of the binding, similarly to what we observed in the binding of LSm1-7 to BMV IR that was only affected by the simultaneously removal of the binding sites L1 or L2 (Publication I); or ii) given that the deletions predictably did not caused major rearrangements in SLIII folding (data not shown), LSm1-7 binding to SLIII can be more dependent on the structure than in the sequence, as it occurs in its binding to BMV RNAs tRNA-like structures (Publication I). Further studies, such as additional band-shift assays with new transcripts containing simultaneous deletions of hairpin stem-loops, or assessing protection patterns for SLIII in the presence and absence of LSm1-7 by techniques like Tb(III) hydrolysis mapping[37,176], would be required to definitely define how LSm1-7 binds to HCV 5'UTR.

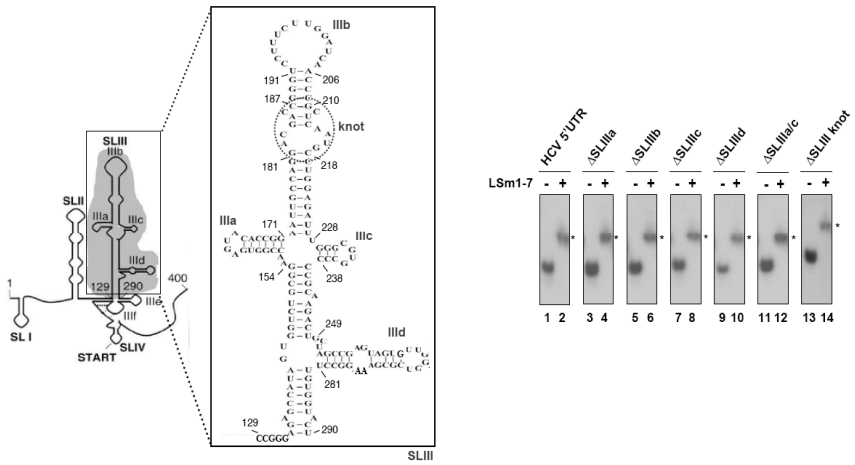

Figure S3. LSm1-7 binding to HCV 5'UTR is not affected by single removal of hairpin stem-loops in SLIII. (Left Panel) Detailed structural organization of the mapped region from SLIII shown to bind to LSm1-7 ring (Adapted from Publication III and [177]). Numeration refers to the nucleotide positions in the genome of the HCV Con1 strain. (Right Panel) Radiolabeled gel-purified RNA transcripts (described in the text) were incubated with reconstituted LSm1-7 rings. After complex formation, products were separated on a non-denaturating polyacrylamide gel and visualized by autoradiography.

3.2.5 Appendix III:

LSm1-7 complex strongly interacts with transcripts containing SLIIId-e, SLIV and core elements

Computer-assisted comparative analyses of diverse Hepatitis C Virus (HCV) sequences has revealed a remarkably high level of nucleotide sequence conservation in the core encoding region, in which synonymous mutations are suppressed[178,179]. These observations suggest that the nucleotide sequence can have a functional role beyond simply encoding the core protein[147]. Four conserved RNA stem-loop structures designated SL47, SL87, SL248, and SL443 have been predicted in the HCV core encoding region[146,179] (Figure S4A). Interestingly, the core structures SL47 and SL87, but not SL248 and SL443, were shown to be important *cis*-acting RNA elements required for HCV translation and replication. The integrity of these RNA elements was suggested to be important for the assembly of higher-order RNA structures and/or binding to (a) cellular protein(s), that would regulate HCV translation and replication[147,180]. Accordingly, since we showed that translation and replication of HCV depends on LSm1-7 proteins (Publication III), we considered the possibility of this complex to directly interact with these RNA elements present in the core region. Hence, to test this possibility we performed in vitro electro-mobility shift assays using recombinant LSm1-7 rings and a set of transcripts designed in order to map an eventual interaction with the core gene (Figure S4B).

Unpublished data from our laboratory, previous to the description of the existence of *cis*-acting signals in the stem-loops SL47 and SL87[180], showed that LSm1-7 complex strongly interacts with a transcript containing the nucleotides 250 to 650 (250-650, Figure S4B, lanes 7-8). This transcript comprises SLIIId, SLIIIe and SLIIIf from the IRES, plus the SLIV where is localized the starting codon of the core protein, and also, from the core gene, SL47, SL87 and almost the totality of SL248 (figure S4A). Taking into consideration the highly structured regions of this segment, we generated transcripts to address which of the sequence elements interacts with LSm1-7 complex. Firstly we tested transcripts designed to (i) exclude the SLs present in the core region (250-388); (ii) include SL47 (250-426; 290-426); include SL47 and SL87 (250-508; 290-508); or (iii) include SL47, SL87 and SL248 (250-650; 290-650).

The justification to design these transcripts starting in nt-250 or nt-290 is related to the fact that the region comprised between these nucleotides, and where SLIIId is included, are still part of the area mapped for the interaction of LSm1-7 to the HCV5'UTR (Publication III). Thus, this design allowed us to clearer interpret the outcome of these binding assays and evaluate the eventual binding of LSm1-7 to the core gene structures.

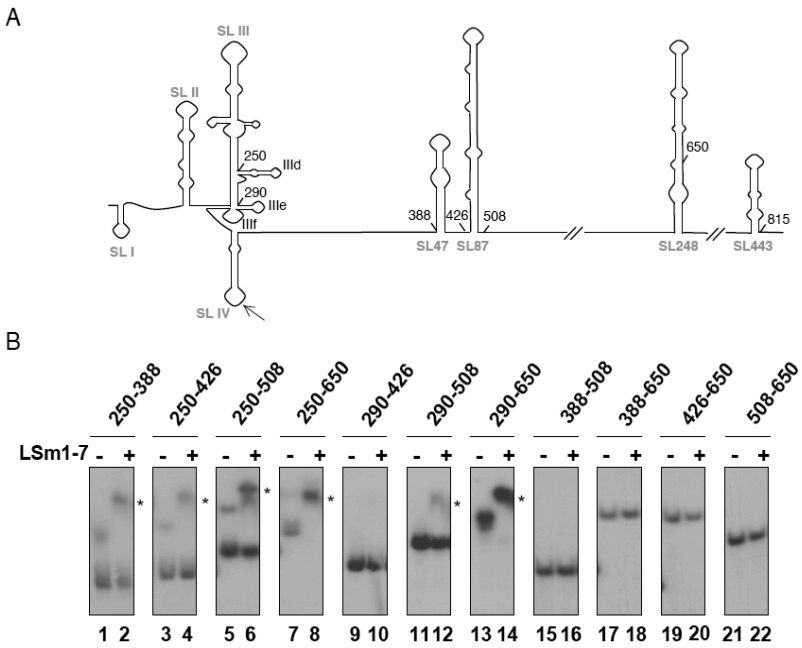

Figure S4. LSm1-7 complex strongly interacts with transcripts containing SLIIId-e, SLIV and core region elements.

(A) Predicted RNA structures SL47, SL87, SL248 and SL443 in the core enconding region. The polyprotein initiator AUG in SLIV is pointed out by a arrow, and hairpin stem loops IIId-e are indicated (Adapted from Publication III and [180]). (B) Gel-shift assays were performed using recombinant LSm1-7 rings and generated RNA transcripts used map an eventual interaction with the core region. After complex formation, samples were loaded on a nondenaturating polyacrylamide gel and visualized by autoradiography.

We observed that transcripts 250-388 and 250-246 weakly bound LSm1-7 complex (Figure S4B, lanes 1-4) while curiously, transcripts 290-400 (Publication III, Figure 3) and 290-426 showed no interaction with the ring (lanes 9-10). The difference in the binding of LSm1-7 to these constructs is eventually related with the presence or absence of SLIIId, what could indirectly suggest that this is the binding region of LSm1-7 to SLIII (Publication III). Although, the data presented previously excludes, at least, the possibility of SLIIId being the exclusive binding site since (i) the interaction of LSm1-7 with the SLIII results in a total shift of the corresponding transcript (Figure 3, publication III), while here we observed only a weak binding, and (ii) the removal of SLIIId do not disrupt the strong interaction between the complex and HCV 5'UTR (Part II – Appendix II). Furthermore, the finding that LSm1-7 do not bind to 290-426 transcript indicates that this complex do not interact with SL47. Following, when the sequence corresponding to SL87 was added (250-508 and 290-508) we observed a slight increase in LSm1-7 binding capabilities (Figure S4B, lanes5-6 and 11-12). Importantly, the inclusion of the sequent stretch corresponding to SL248 (250-650 and 290-650) led to a complete band-shift (Figure S4B, lanes 7-8 and 13-14). Taken together, these observations pointed out to an important role of the RNA elements present in the region between nt-508 and nt-650. To approach the importance of this segment for the binding to LSm1-7 complex, we tested transcripts without stretches of the IRES SLIII (388-508, 388-650, 426-650 and 508-650). We showed that the absence of these SLIII elements disrupted the previously observed bindings to LSm1-7 complex (compare lanes 7-8 and 11-12 with lanes 13-18) and that the SL248 alone (508-650) is not able to interact with this complex.

The data obtained in these experiments do not provide us enough evidence to conclude how LSm1-7 interacts with the core gene region. We could observe that sequences from the core region, nt508-650, are required for a total band shift of the transcripts 250-650 and 290-650. Although, that same elements by themselves do not bind to LSm1-7. Given that we showed in Publication I that not only the sequence but also the secondary structure of the RNA can be important for an efficient binding, it is possible that LSm1-7 complex can only interact with the regions of the core in very specific conformations, that only occur in the presence of

elements from the 5'UTR IRES. This would suggest that LSm1-7 may mediate an eventual RNA-RNA interaction between the IRES SLIII, or SLIV and RNA elements present in the core gene, such as the SL248. As suggested in the previous appendix, further mappings or protection assays in the presence or absence of LSm1-7 ring, are approaches that can guide to a better definition of the mechanism associated to the binding of LSm1-7 complex to these RNA elements.

3.2.6 Appendix IV:

LSm1-7 complex and NF/NFAR proteins interact simultaneously with HCV 5' and HCV 3'UTR

The cellular "NF/NFAR" proteins (Nuclear Factors associated with dsRNA) are a group of primarily nuclear proteins including NF90/NFAR-1, NF110/NFAR-2 (an isoform of NF90/NFAR-1), NF45 and RNA Helicase A (RHA). These proteins, with exception of NF45, belong to the widespread and functional diverse family of dsRNA binding proteins (DRBP) along with, for example, the dsRNA activated protein kinase PKR and the RNase III family including DICER[159,181]. NF/NFAR proteins are suggested to play crucial roles in mRNA post-transcriptional regulation, including mRNA stability, export and translation and may also have an important function in host defense (reviewed in [181]). Interestingly, the NF/NFAR proteins were shown to be involved in the HCV RNA replication process[159]. The work of Isken *et al.* revealed that in cells containing replicating HCV RNA, the predominantly nuclear NF/NFAR proteins accumulate in the cytoplasmic viral replication complexes. Curiously, NF/NFAR proteins were also shown to specifically bind to replication signals in the HCV genomic 5' and 3' termini that essentially overlap with the defined regions of interaction with the LSm1-7 complex ([159] and publication III). Furthermore, a purified [NF90/NFAR1, NF45, RHA] complex was able to trigger the formation of a loop structure of the HCV RNA, what was suggested to be linked with its binding to HCV 5'- and 3'UTRs and to have an important role for viral replication.

As mentioned previously, we have shown that LSm1-7 complex interacts with the same regions of HCV 5'- and 3'UTRs mapped to bind to NF/NFAR proteins. Thus, we considered the possibility that these complexes could compete for the binding to these regions of the HCV genome or, in alternative, if they were able to bind simultaneously. In contrast to our work, the mapping of NF/NFAR proteins binding sites in HCV UTRs was performed using UV-cross linking assays with cellular extracts, followed by Western blot to identify the proteins of interest[159]. In order to uniform procedures, we established a band-shift assay using a chromatographic fraction of the purified [NF90/NFAR1, NF45, RHA] complex,

kindly provided by Dr Behrens, and labeled transcripts of the HCV 5' and 3'UTRs. As expected, we showed a strong interaction between [NF90/NFAR1, NF45, RHA] complex and both 5' and 3'UTRs of HCV, in which 1-2μg of this complex was able to totally shift the RNA transcripts (Figure S5A). Curiously, we observed a stepwise pattern of migration dependent on the amount of [NF90/NFAR1, NF45, RHA] complex, what may indicate the formation of higher-orders complexes.

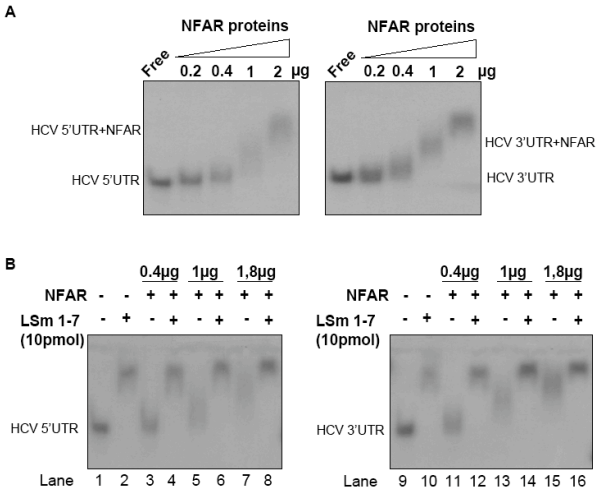

Figure S5. LSm1-7 and [NF90/NFAR1, NF45, RHA] complexes interact simultaneously to HCV 5' and 3'UTR. (A) Bandshifts assays were perfomed with labeled HCV 5' or HCV 3'UTR transcripts incubated with increasing amounts of a chromatographic fraction of the purified [NF90/NFAR1, NF45, RHA] complex (NFAR)), using essentially the same conditions described previously. (B) Similar assays were also performed in the presence or absence of 10pmol of reconstituted LSm1-7complex.

Following, to address if the presence of LSm1-7 complex affects the binding of NFAR proteins to HCV RNAs, we performed an experiment in which labeled HCV 5' or 3'UTR transcripts were incubated with increasing amounts of [NF90/NFAR1, NF45, RHA] complex, in the presence or absence of 10pmol of LSm1-7 complex (Figure S5B). In the absence of LSm1-7, as expected from the outcome of data presented in Figure S5A, we observed a stepwise pattern of migration from both transcripts dependent on the amount of [NF90/NFAR1, NF45, RHA] complex (Figure S5B, compare lanes 1, 3, 5 and 7 and lanes 9, 11, 13 and 15). Interestingly, the presence of LSm1-7 complex led to a super-shift of the transcripts (Figure S5B,

lanes 4, 6 and 8 and lanes 12, 14 and 16). Such observation could be due to i) a positive competition of the LSm1-7 complex that could provoke to the disruption of the binding of the [NF90/NFAR1, NF45, RHA] complex to the RNA or ii) a simultaneously binding of both complexes to the transcripts. In the first case, if the LSm1-7 complex totally competed the binding of NF/NFAR proteins to the target RNAs, one would expect that the transcripts would migrate always at the same level, given that the amount of LSm1-7 complex is constant. Although, we can observe that in the presence of LSm1-7 complex, the migration of HCV 5'UTR or HCV 3'UTR labeled transcripts has also a stepwise pattern (Figure S5B, compare lanes 2, 4, 6 and 8 and lanes 10, 12, 14 and 16). Given the previously described results we suggest that such pattern can be attributed to the increasing amount of [NF90/NFAR1, NF45, RHA] complex added. In resume, these experiments indicate that LSm1-7 and [NF90/NFAR1, NF45, RHA], complexes, previously described to bind to similar regions of HCV genome, can interact simultaneously with HCV 5' and 3'UTRs. Although, further experiments are required to clarify if both complexes are directly binding to the RNA, or if there is a protein-protein interaction involved, as well as if there is a biological meaning to these observations.

4. Discussion

Viruses life cycles are remarkably diverse, even within the same group, but have in common, however, the dependence on host factors to achieve their multiplication. Thus, the identification of common cellular factors requirements might provide not only clues about viruses origins and their evolutionary links, but also imply an important clinical impact, since such common factors can potentially be used as targets for broad-spectrum antiviral strategies.

Members of the group of (+)RNA viruses were shown to interact with the host decay machinery in order to express their genomes and regulate their life cycles. Previous works using the BMV/yeast system demonstrated that the proteins LSm1-7, Dhh1 and Pat1, activators of mRNA decapping in the 5'-3'-deadenylation-dependent mRNA decay pathway, are essential cellular factors required for both translation and exit from translation to replication of the BMV genome[123-125]. Given the conservation of the 5'-3' mRNA decay pathway from yeast to humans, together with the common requirement of all (+)RNA viruses to regulate the switch of their genomes from an active to an inactive translational state to allow replication, we hypothesized that LSm1-7, Dhh1, and Pat1 proteins function not only in BMV replication in yeast but can also be required for the replication human (+) RNA viruses in human cells.

The decapping activators LSm1-7, Dhh1 and Pat1 are conserved from yeast to humans

In yeast, the cellular proteins LSm1-7, Dhh1 and Pat1 interact with each other in an mRNA-independent way and accumulate in P-bodies with other components of the 5'-3' mRNA decay pathway[6,73,75,182]. In contrast to the other components of this mRNA decay pathway, the human homologue of the yeast Pat1 protein was still unknown. Pat1 is a particularly interesting protein since besides acting as activator of decapping, translation repressor and promoter of mRNPs assembly, it is also known to be involved in the antagonistic process of translation initiation, thus suggesting an important regulatory role[33,72,74,75]. We described the identification of an human gene, which we denominated PATL1, coding for an ORF that presents similar features to

the yeast PAT1. PATL1 mRNA was ubiquitously expressed in all tested human tissues, as well as the mRNAs from LSM1 and RCK/p54, the human homologues of the yeast proteins LSM1 and DHH1, respectively. This general expression can be considered expectable given the involvement of these proteins in a decay pathway that is conserved from yeast to humans, and that is one of the principal routes for mRNA turnover. Furthermore, the unequal levels of these mRNAs detected in different tissues - and assuming a direct correlation with the corresponding proteins levels - may indicate the existence of a tissue-specific regulation of mRNA turnover. Such tissue-specific regulation would not be an isolated case since mRNA translation and degradation can also be regulated by miRNAs, which are produced in a tissue-specific manner[183].

Regarding the localization of this protein within the cell, we showed that fluorescence-tagged PatL1 accumulates in distinct cytoplasmic foci which we recognized as P-bodies, since it co-localized with the P-body components LSm1, Rck/p54 and Dcp1, resembling the cellular localization of Pat1 in yeast. Furthermore, human PatL1 parallels yeast Pat1 which when suppressed leads to P-body loss[69,184]. Similar observations are described for the depletion of proteins acting in the early stages of the 5'-3' mRNA decay, such as Dhh1 in yeast, and Rck/p54 and LSm1 in humans, indicating their requirement for P-body formation[69,185]. This requirement was linked to their role in regulating the exit of the mRNAs from translation to a non-translating state that is essential for P-body formation. Together, these observations strongly indicate that PatL1 is the human homolog of the yeast Pat1, suggesting that this protein, together with LSm1-7 and Rck/p54, has a conserved role as regulator of mRNA fates both in yeast and humans.

LSm1-7, Dhh1 and Pat1 play a key role in the replication of (+)RNA viruses

Given the apparent functional conservation of LSm1-7, Dhh1 and Pat1 proteins in yeast and humans, we considered likely that their role in the replication of (+)RNA viruses could also be preserved in human cells. Remarkably, our subsequent studies demonstrated that the down-regulation of LSm1, Rck/p54 and PatL1 expression

affects translation and replication of HCV. This effect was observed in both genotype 1b replicons and genotype 2a viral derivatives. In the case of genotype 2a infectious HCV, although the observed inhibition of particle production and viral RNA accumulation is expected to be associated with an early step of the viral life cycles, such as translation or replication, one can not exclude an additional effect on RNA encapsidation. In addition to the roles of these proteins in BVM and HCV life cycles, others had previously observed that Hfq, the homolog of LSm1 in bacteria, is a factor required for the replication of the (+)RNA bacteriophage Qß and that efficient retrotransposition of the yeast retrovirus-like element Ty-1 and Ty-3 is also dependent on Dhh1, LSm1 and Pat1[26,186-188].

The life cycle of (+)RNA viruses and retroviruses have in common the need to separate the steps of replication and assembly from the translation of their genomes, in order to avoid competition between the respective replicative or packaging machineries and the elongating ribosomes for acceding the viral RNA[77]. Thus, the fact that some members from both virus groups are dependent on host factors which cellular function is linked to the regulation of the switch of the cellular mRNAs from a translational to a non-translational state, can indicate that the viruses hijacked the proteins cellular function to their own benefit. The implications regarding the involvement of proteins from this mRNA decay pathway in viral life cycles will be discussed further below.

LSm1-7 rings interact directly with important regulatory signals within BMV and HCV RNA genomes

To address the molecular mechanisms underlying the role of these proteins in the replication of BMV and HCV we decided to focus on the LSm1-7 ring because it interacts directly with cellular mRNAs and thus it was a good candidate to establish such kind of interactions with viral genomes. This complex belongs to the Sm-family of proteins known to interact directly with their target RNAs, in order to play their function in multiple aspects of RNA metabolism, as well as to facilitate RNA-RNA and RNA-protein interactions[3]. Accordingly, the function of LSm1–7 rings as activators of decapping of cellular mRNAs involves their direct binding to short

oligo (A) tracts at the 3'end of deadenylated cellular mRNAs[6,10,33,175]. To address the possibility that LSm1–7 rings interact directly with the BMV and HCV RNA genomes, we performed *in vitro* band-shift assays using the newly available functional recombinant reconstituted LSm1-7 complexes[174]. These assays provided important information that help us not only to understand better the role of this complex in viral life cycles, but also to newly define some biochemistry aspects associated with the LSm1-7/RNAs interactions, that I will be discussing next.

Previously to our work only two other studies addressed the binding specificity of LSm1-7 to RNA sequences other than homopolymers. Firstly, by performing band-shift assays with the same reconstituted LSm1-7 rings used in our study, it was shown that this complex is able to directly interact and stabilize reporter RNAs containing 5'poly(A) tracts[32]. In the other study, binding assays were performed using a purified yeast LSm1-7/Pat1 complex showing a preferential binding of this complex to oligoadenylated rather than to polyadenylated RNAs. Additionally it was also shown a direct interaction of this complex near or at 3'ends containing oligo(U) stretches[34]. However it is not clear how the presence of Pat1 protein in this complex contributes to its binding specificity.

Here, we were able to show that reconstituted LSm1-7 complexes specifically recognize important signals, either in BMV and HCV genomes, that regulate their translation and/or replication. In the BMV genome LSm1-7 was shown to directly interact with the tRNA-like structure (TLS) region present in the 3'end of all three genomic RNAs, and to two internal single-stranded A-rich loop sequences (L1 and L2) in the intergenic region of BMV RNA3. Concerning its association with the HCV genome, LSm1-7 complex recognizes the SLIII, that makes part of the highly structured IRES, present in the 5'UTR of the genome, and also the poly(U/UC) tail in the HCV 3'UTR. These data constitutes the first evidence that the LSm1-7 complex is able to directly interact with viral genomes representing also novel LSm1-7 interaction sites.

These newly defined RNA-target sites for LSm1-7 rings resemble the ones described for Hfq, their bacterial counterpart[3,37]. Hfq forms an homohexameric ring complex

that functions as regulator of mRNA decay and is also known to enable translational control and mRNA stability by facilitating RNA-RNA and RNA-protein interactions[3]. Furthermore, the upper mentioned role of Hfq as cellular factor required for the replication of the (+)RNA bacteriophage Qß[26] has been associated with its RNA chaperone activity, as this complex is capable of remodeling the structure of the viral RNA in order to allow the access of the viral replicase to the 3'end and subsequent initiation of replication[27]. Mutational analysis and competition assays have shown that Hfq complex binds RNA substrates on both proximal and distal faces of the ring, however with different binding specificities[35]. The distal face binds to poly(A) sequences while the proximal face binds to internal AU-rich regions located close to stem-loop structures in non-coding RNAs (ncRNAs) or to highly structured regions like tRNAs[36,37]. As for Hfq, LSm1-7 complexes bound to internal regions, in this case A-rich stretches present in L1 and L2 that flank a stem-loop RNA structure of the intergenic region of BMV and the poly(U/UC) present in the 3'UTR of HCV genome. Additionally, it also bound to the AU-rich element (ARE) present in a ß-globin mRNA reporter. Furthermore, we showed the LSm1-7 complex bound to two highly structured regions, the tRNA-like sequence (TLS) in the 3´UTRs of BMV RNAs and to the SLIII of HCV 5'UTR IRES. Interestingly, as described for the Hfq/tRNA interactions, is not the sequence motifs *per se* but the structural motifs of these highly structured regions what promote the binding of LSm1-7 to their RNA targets. By mapping the minimal BMV 3'UTR sequence necessary for optimal binding, we observed that LSm1-7 ring totally shifts the transcripts only in the presence of the complete 3'UTR sequence. Removal of stretches in the BMV 3'UTR either from the tRNA-like structure or from the non-tRNA-like structure led to a lower binding efficiency. Regarding the SLIII of HCV IRES, we showed that the single removal of the ssRNA hairpins existent – the predictable binding sites of LSm1-7 within SLIII - did not disrupt the binding this region, indicating that the interaction of LSm1-7 with SLIII is not dependent in the sequence or that exists at least two binding sites.

The fact that the described binding properties and functions of LSm1-7 complexes are similar to those observed for the bacterial Hfq, suggest a strong conservation

through evolution. However, it is still an unanswered question if the LSm1-7 complexes use both faces of the ring to interact with their target RNAs, as it occurs with Hfq. Our cross-competition binding analysis showed that both BMV sequences, as well as the ARE-element, bound through the same binding site(s) in the LSm1-7 complex. A similar observation occurred when HCV 5'- and 3'UTRs were competed with each other. However, the observation that the binding of LSm1-7 ring to HCV RNAs is only partially competed by BMV RNA3 3'UTR, BMV RNA3 IR or ß-globin-ARE mRNA, indicates that HCV 5' and 3'UTR bind to at least two different sites within LSm1-7 complex, from which only one of them is utilized by the other transcripts. Albeit our data point to the existence of multiple binding sites in the LSm1-7 complex, only the attainment of crystal structure from the ring will definitely clarify if this complex uses both faces to bind RNAs. In any case, the finding that this complex can bind two different transcripts simultaneously can have important impact to the clarification of its role in the regulation of either cellular mRNA and viral RNA fates. For example, Bergman and colleagues proposed recently that LSm1-7 regulates mRNAs generated by poxviridae[32]. Viruses of this family generate viral mRNAs with additional oligo(A) tract located at their 5'ends. They have shown that the binding of reconstituted LSm1-7 rings to such a tract in the 5'end of reporter RNAs does not result in mRNA decay but rather in RNA stabilization through inhibition of decapping and degradation. This effect was proposed to be the result of a 5'-3'-circularization of the viral RNA mediated by binding of the LSm1-7 complexes to the 5' and 3'ends.

The requirement of LSm1-7 complex for the efficient translation of Hepatitis C virus can be mediated by a similar process, as supported by our observation of a direct interaction of the LSm1-7 rings with sequences in both 5' and 3'UTRs. These interactions might facilitate rearrangements in the viral RNP structure and/or composition, by recruiting proteins such as Rck/p54 and PatL1, that instead of promoting decay, can possibly promote HCV RNA translation and subsequent transfer to replication. In this context, the potential overlap of the LSm1-7 and eIF3 binding sites in SLIII of HCV 5'UTR can have interesting implications, since at a later stage of the infection process and, eventually, in a different context of factors, LSm1-7 complex could also be part of the inhibition of translation by interfering

with the interaction of eIF3 with the IRES. The hypothesis that LSm1-7 is eventually involved in the circularization of the HCV genome is also supported by the observation that NFAR proteins interact essentially with the same regions in both HCV 5' and 3'UTRs[159] and that can do it simultaneously with LSm1-7 rings. The NFAR proteins are known to promote HCV RNA 5'-3' interactions although the mechanism has yet to be defined. However, the fact that both NFAR and LSm1-7 rings interact specifically with essential replication signals in the HCV UTRs suggests that genome circularization is a decisive precondition of the HCV life cycle, what supports the generally accepted concept that 5'-3' RNA-RNA interactions are required for the replication of (+)RNA viruses.

Regarding the molecular mechanisms associated with the role the LSm1-7 complex in the replication of BMV we showed that LSm1-7 rings interacted directly with tRNA-like structures present in the BMV RNA sequences, and with two A-rich loops present in the intergenic region of BMV RNA3. Importantly, *in vivo* analysis showed that these sites are required for the regulation of translation and recruitment of BMV RNA. For BMV RNA3, substitution of the TLS region by a poly(A) tail suppressed the requirement of LSm1-7 complexes for RNA3 translation and recruitment. Since poly(A) tails in cellular mRNAs mediate 5´-3´circularization via the binding of the poly(A)binding protein, the LSm1-7/TLS binding may promote such 5´- 3´interactions in BMV RNAs, what would mean that LSm1-7 might be playing a similar and conserved role in both HCV and BMV RNAs. Furthermore, the Sm complexes, as well as LSm2-8 and Hfq, were shown to present RNA chaperone-like activities, facilitating a variety of RNA-RNA and RNA-protein interactions[3]. Thus, we consider plausible that LSm1-7 functions similarly in their RNA targets. Accordingly, the binding of LSm1-7 to the TLS and the loops L1 and L2 of the internal region might facilitate rearrangements in the viral ribonucleoprotein (RNP) structure in order to promote the 5'-3' interaction. However, given that LSm1-7 rings do not interact with the 5'UTR of the BMV RNAs, this interaction might be mediated by other cellular proteins. A good candidate would be Pat1 since it interacts directly with the Lsm1-7 complex and has been shown to be important not only for activation of decapping but also for translation initiation (REF), The 5´-3´circularization, as occurs with the cellular

mRNAs, would then stabilize the RNA permitting an efficient translation and, additionally, the 1a-dependent recruitment via recognition of the RE element.

Function of the Lsm1-7 complex in the antagonistic processes of BMV RNA translation and recruitment

The observation that the absence of LSm1-7 complexes lead to the inhibition of both BMV RNA3 translation and recruitment indicates that the binding of LSm1-7 complex is required for both processes. Furthermore, as these processes are antagonistic, since they can not occur simultaneously in the same RNA molecule, alterations such as the presence of 1a protein expression or the absence of some of the LSm1-7 interaction sites in the viral genome would result in favoring one over the other. Accordingly, we observed that deletion of the L1 and L2 internal loops resulted in inhibition of translation and in enhancement of recruitment. Such deletion would prevent the binding of the LSm1-7 complex and hence inhibit translation by interfering with the 5'-3' circularization. This translation inhibition might then result in a LSm1-7-independent enhancement of recruiting due to the absence of the competing translation function. Alternatively, deletion of these internal sequences could affect the RNP reorganization in such a way that the RE signal might be more accessible to the 1a protein. However, we can not exclude that the observed effects may be also due to the binding to the loops L1 and L2 of additional host factors not identified yet.

Deletion of both L1 and L2 loops in the intergenic region of the RNA3 was required to abolish LSm1-7 binding, however deletion of either L1 or L2 inhibited translation and enhanced recruitment. It is yet unclear which are the basis of this particular disconnection between the *in vitro* binding assays and the *in vivo* functional data. However, it might be possible that in the *in vivo* setting, where the local concentration of viral RNA and LSm protein is low, the presence of both L1 and L2 sites are important to promote efficient LSm1-7 binding. Hence, under *in vivo* conditions, a suboptimal binding site would not suffice to allow LSm binding and function. In the *in vitro* setting, however, we may simply override this subtle effect. Such disconnection can also be linked to the eventual RNA chaperone-like activity

that LSm1-7 complex might present. Thus, the binding of LSm1-7 to L1 and L2 in the intergenic region can be promoting a RNA conformation changes that facilitates the recruitment of other cellular or viral components, or favoring a RNA-RNA interaction important for the regulation of the transition between translation and recruitment. If this is the case, the binding of LSm1-7 just to one of the loops could lead to a misfolded conformation of the viral RNA that would not be functional and able to correctly mediate such RNA-RNA or RNA-protein interactions, but somehow favoring viral RNA recruitment. However, this fact would not necessarily correspond to a loss in the efficiency of binding of LSm1-7 complex to the transcripts ΔL1 or ΔL2 when compared to IR transcript.

How Lsm1-7 complexes can function in both translation and replication of BMV and HCV genomes? Such an apparently antagonistic function, to promote both translation and exit from translation, is not without precedent as cellular proteins acting in two antagonistic processes such as translation initiation and translation repression have been described[189]. The regulation of this double function can be obtained by different mechanisms. CPEB is an example of a cellular translation regulator that can behave as a repressor or an activator depending on its phosphorylation status. By its turn, the regulator IRP acts as RNA-binding protein under low iron conditions repressing ferritin mRNA translation, while under high iron conditions shows cytoplasmic aconitase activity. These functions are mutually exclusive and, consistently, high iron conditions lead to a loss of affinity for RNA and the subsequent translation of ferritin mRNA. An advantage of using a single complex for opposing outcomes seems to be the possibility of responding rapidly to different cellular requirements. A similar advantage might apply for the regulation of the viral life cycle.

Role of P-bodies in the (+)RNA virus life cycle

P-bodies are cytoplasmic granules where translational repressed messenger ribonucleoproteins (mRNPs) accumulate in association with cytoplasmic decapping machinery components such as LSm1-7, Dhh1 and Pat1. Furthermore, mRNA and protein components of the miRNA, ARE-mRNA decay and NMD pathways can

also be observed in P-bodies (reviewed in[55]). However, the role, if any, that the assembly of individual mRNPs into P-bodies plays in promoting translational repression, decapping and/or mRNA decay is still an unresolved question. Recent studies have failed showing the loss of visible P-bodies as a cause for the inhibition of mRNA decay and translational repression pathway[7,63,76].

Some lines of evidence argue that the requirement of Lsm1-7, Pat1 and Dhh1 proteins for replication of BMV is due to a role of P-bodies in concentrating genomic RNAs and viral proteins and subsequently promote their interaction with cellular membranes (reviewed in[190]). First, it was observed that BMV RNA2 and RNA3 accumulate in P-bodies, being RNA3 accumulation dependent on *cis*-acting signals that contribute to replication efficiency. Second, the viral polymerase 2a was also shown to accumulate in P-bodies and to co-immunoprecipitate with LSm1. Finally, at least some P-bodies are localized and can be biochemically associated with membranes where BMV and other viruses replicate[191-193]. Thus, a proposed model suggests that P-bodies containing genomic RNAs not engaged in active translation and viral replication factors, would interact with the cellular membranes in order to facilitate the interaction between the membrane-bound replication factor 1a and the viral components in P-bodies, thereby leading to the assembly of a viral replication complex[190].

Our data demonstrate that the absence of LSm1-7 complex leads to an inhibition of translation and recruitment of the viral RNA3. As the number of P-bodies in considerably reduced in the absence of LSm1 one should expect that the non-translating RNA3 could not be targeted to these foci and, accordingly to the proposed model, this would affect the recruitment of the viral genomes by 1a protein and the subsequent formation of viral replication complexes in the membranes. In the other hand, the deletion of L1 and/or L2 lead also to translation repression but have opposite effect in the recruitment of BMV RNA3. In this case, the removal of L1 and/or L2, although inhibiting the translation, it should not affect the targeting to P-bodies as the required *cis*-acting signals are still present as well as the essential components for the assembly of these foci. Thus, once in P-bodies the viral RNAs could be recognized and recruited by 1a protein. As

discussed before, the lack of the interaction sites of LSm1-7 complex could facilitate the formation of a RNP that favors the 1a-recognition leading to the increased levels of recruited RNA3.

Finally, there are also observations for HCV, apart from the hereby described function of LSm1-7, Rck/p54 and PatL1 in its replication, that suggest a connection between HCV and P-bodies. First, the HCV core protein physically interacts with human helicase DDX3, which was recently shown to be essential for HCV-replication, and both co-localize in cytoplasmic foci reminiscent of P-bodies[194-197]. Secondly, the replication of HCV is enhanced by the interaction of miRNA122 with the 5'UTR of the HCV genome[198]. This observation is curious since one function of miRNAs appears to be the ability to target mRNAs to P-bodies in mammalian cells[199,200]. However still very speculative, it has been suggested that assembly of an RNP containing P-body components and the HCV RNA is important for HCV replication[190].

It is now evident that key protein components of P-bodies are required for the successful completion of the life cycles of some (+)RNA viruses. However, it is still not clear if P-bodies, as a structure *per se*, play any role in the replication of these viruses. Such hypothetical involvement is very difficult to prove since the deletion of most of the human components of P-bodies results in their disassembly. Subsequent studies combining biochemical, biophysical, molecular, and cell biological approaches are needed to resolve this interesting issue and well as other questions regarding the importance of P-bodies in cytoplasmic mRNA metabolism of decapping and the assembly of repressed mRNPs.

5. Supplemental Material & Methods

Additional Constructs used in gel-shift assays (used in Part II – Appendices II and III)

RNA sequences of interest were amplified by PCR using complete Con1 HCV genome (GenBank accession number AJ238799) as DNA template. Forward primers contain an *Eco*RI restriction site at the 5'end followed by the T7 promoter and the HCV sequence. Reverse primers contain a *Bam*HI site at the 3' end. The resulting *Eco*RI/*Bam*HI PCR-fragments were subcloned into pUC18 (New England Biolabs). The oligonucleotide sequences used for amplifying the different areas are given in the table bellow. DNA fragments used in Part II-AppendixIII, ΔSLIIIa, ΔSLIIIb, ΔSLIIIc, ΔSLIIId, ΔSLIII knot were obtained in a two-step PCR procedure. ΔSLIIIa/c was obtained also from a two-step PCR but using ΔSLIIIa as template.

***In vitro* transcriptions**

In brief, plasmid DNAs were linearized, extracted with phenol and chloroform, precipitated with ethanol, and dissolved in RNase-free water. α-32P-UTP labeled transcripts were *in vitro* synthesized from 1μg of the respective linerarized plasmid, at 37°C for 4h in 25μl reaction mixtures containing 40mM Tris-HCl (pH 7.9), 6mM MgCl2, 10mM DTT, 10mM NaCl, 2mM spermidine, 1mM each ATP, CTP, GTP and UTP, 10U RNasin, 20U of T7 or SP6 polymerase (Fermentas), and 10 μCi of [α-32P]UTP. After 2 hours of incubation, additional 20U the corresponding RNA polymerase was added. Similar reaction mixtures were used for synthesis of unlabeled transcripts, with the exception of using 2.5-5μg of linearized plasmid and a final volume of 50μl. The 32P-labeled transcription products were gel purified in a 5% acrylamide (19:1), 8M urea gel, and RNAs were then eluted from excised bands, over-night at 4°C, in an Elution buffer containing 300mM NaAc (pH8.2), 2mM EDTA and 0.1% SDS. Labeled transcripts were precipitated with ethanol, dried with speedvac and dissolved in RNase-free water. *In vitro* transcriptions reactions performed without 32-P UTP were terminated with 1U of RNAse-free DNAse per μg of plasmid DNA and 1 hour incubation at 37°C. RNA transcripts were extracted with phenol and chloroform, precipitated with ethanol, and dissolved in RNase-free

water. RNA concentration was determined by measurement of the optical density at 260nm, and RNA integrity was checked by denaturing agarose gel electrophoresis.

Expression and Purification of LSm1-8 Proteins and sub complexes

LSM1, LSM4 and LSM8 were expressed as individual LSm proteins from pQE30 vectors, while LSM2/3 and LSM5/6/7 sub-complexes were expressed from polycistronic cassettes inserted into pQE30 T5 expression vectors, as described in Zaric B, 2005. SG13009 *E.coli* cells were transformed with the respective plasmid DNA and plated on selective media. LB starter cultures were grown at 30°C overnight, and 2-12 liters of LB media were inoculated next day. Cultures were grown to an A600 of 0.8 at 37°C and induced with 1mM isopropyl 1-thio-D-galactopyranoside at temperatures ranging 25°C and 37°C. Cells were harvested after 4-48h of induction, depending on the construct. Cell pellets were resuspended in lysis buffer (20 mM HEPES-Na, pH 7.50, 0.5–1.0 M NaCl, 10 mM imidazole-Cl, pH 7.50, 5 mM _-mercaptoethanol), sonicated,and treated with DNase I. Insoluble material was removed by ultracentrifugation, and supernatants were purified by immobilized metal ion affinity chromatography (IMAC) on nickel-charged Hi-Trap chelating Sepharose columns (Amersham Biosciences). LSm proteins and sub-complexes were eluted with imidazole step gradients (60, 250, and 500 mM). If insufficiently pure, samples were subsequently dialyzed into 100 mM NaCl buffer without imidazole and subjected to ion exchange chromatography (100 mM to 1 M NaCl). Samples were frozen in liquid nitrogen in ion exchange buffer.

Reconstitution and Purification of LSm1-7 and LSm2-8 complexes

Individual LSm protein or sub-complex preparations were incubated in 4 M urea, 1 M NaCl, 20mM HEPES (pH7.5) and 5mM DTT dissociation buffer for 2 h at 37 °C without shaking. Following, each protein or dissociated protein subcomplex was mixed in equimolar amounts and incubated 2-5hr at 37°C and 120rpm for the assembly of the desired heptamer. The sample was dialyzed in a 4000-6000 MWCO 1.8ml/cm tube against buffer with progressively less salt (1 M and 0.5 M NaCl) overnight at 4 °C. Reconstituted LSm1–7 and LSm2–8 were then concentrated to 10k MWCO centrifugal concentrators and injected for purification into Superdex 200 10/300 GL gel filtration column, equilibrated in 20mM HEPES (pH7.5), 0.15M

NaCl, 5mM DTT. Pooled gel filtration fractions we then injected in a HiTrapQ HP 1ml anion exchange chromatography column.

Electro-mobility shift assays

In a typical binding experiment, 300cpm (≈50fmol) of labeled transcript is incubated with 20pmol of LSm1-7 complexes in a buffer containing 20mM HEPES-NaOH pH 7.5, 200mM NaCl, 2mM MgCl2, 0.1U/µl RNAsin and 0.1µg/µl yeast tRNA in a 5µl assay at 30°C for 1h. A 2x loading buffer containing 4M Urea and 20% Glycerol was added to each samples and then total volume was loaded on previously pre-run 5% native polyacrylamide, 4% Glycerol gels, and run at 4°C for 2h, 30mA. Gel were autoradiographed at -80°C on maximum sensivity Kodak Biomalx Films (Sigma-Aldrich, Munich, Germany). For gel shift assays that included RNA competition, we used 50-100fmol of labeled RNA and the indicated molar proportion of non-labeled transcripts, in a 20µl reaction. Buffer composition, LSm complex amount and remaining proceedings as described above.

For gel-shift assays performed with NFAR proteins, 300cpm of [α-32P]-labeled transcripts were incubated with increasing amounts (0 to 2µg), of chromatographically purified [NF90/NFAR1, NF45, RHA] complex, in a buffer containing 20mM HEPES-NaOH pH 7.5, 200mM NaCl, 2mM MgCl2, 0.1U/µl RNAsin and 0.1µg/µl yeast tRNA in a 20µl assay at 30°C for 1h. Remaining proceedings as described above.

Table 1: Primer sequences used to generate constructs for gel shift assays presented in Part II – Appendices II and III

Amplified UTR Region	Primer Name	Sequence from 5' to 3' end
250-388; 250-426; 250-508; 250-650	HCV250 fwd	GGG AAT TCG TAA TAC GAC TCA CTA TAG GCT AGC CGA G
250-388	HCV388 rev	GGG ATC CGT TTG GTT TTT C
250-426; 290-426	HCV426 rev	GGG ATC CCA CCG CCC G
250-508, 290-508; 388-508;	HCV508 rev	GGG ATC CGA CCG CTC GG
250-650; 388-650; 426-650; 508-650;	HCV 650 rev	GGG ATC CAG AGC CAC GGG G
290-426; 290-508;	HCV290 fwd	GGG AAT TCG TAA TAC GAC TCA CTA TAG GTG CCT GAT AGG
388-508; 388-650	HCV388 fwd	GGG AAT TCG TAA TAC GAC TCA CTA TAG GAC CGC CGC CCA C
426-650	HCV 426 fwd	GGG AAT TCG TAA TAC GAC TCA CTA TAG GCA GAT CGT CGG
508-650	HCV 508 fwd	GGG AAT TCG TAA TAC GAC TCA CTA TAG GCG CAA CCT CGT GG
ΔSLIIIa; ΔSLIIIb; ΔSLIIIc; ΔSLIIId; ΔSLIIIa/c; Δknot	HCVpos1 fwd	GGG AAT TCG TAA TAC GAC TCA CTA TAG GCC AGC CCC
ΔSLIIIa; ΔSLIIIb; ΔSLIIIc; ΔSLIIId; ΔSLIIIa/c; Δknot	HCV 400 rev	GGG ATC CTG TGG GCG GCG GTT GGT GTT AC
ΔSLIIIa	ΔSLIIIa fwd	ATA GTG GTC TGC GGA ATT GCC AGG ACG
ΔSLIIIa	ΔSLIIIa rev	CGT CCT GGC AAT TCC GCA GAC CAC TAT
ΔSLIIIb	ΔSLIIIb fwd	CCG GAA TTG CCA GGA CGA CCG GGC CCG CTC CAA TGC CTG G
ΔSLIIIb	ΔSLIIIb rev	CCA GGC ATT GGA GCG GGC CCG GTC GTC C TG GCA ATT CCG G
ΔSLIIIc; ΔSLIIIa/c	ΔSLIIIc fwd	CCC GCT CAA TGC CTG GAG ATT CCG CGA GAC TGC TAG CCG
ΔSLIIIc; ΔSLIIIa/c	ΔSLIIIc rev	CGG CTA GCA GTC TCG CG AAT CTC CAG GCA TTG AGC GGG
ΔSLIIId	ΔSLIIId fwd	GCG TGC CCC CGC GAG ACT TGT GGT ACT GCC TGA TAG GGT
ΔSLIIId	ΔSLIIId rev	ACC CTA TCA GGC AGT ACC ACA AGT CTC GCG GGG GCA CGC
Δknot	Δknot fwd	CCG GAA TTG CCA GGG GTC CTT TCT TGG ATC AAC CCC TGG AGA TTT GGG CGT GCC CCC
Δknot	Δknot rev	GGG GGC ACG CCC AAA TCT CCA GGG GTT GAT CCA AGA AAG GAC CCC TGG CAA TTC CGG

6. Other Publications during Thesis

Publication IV

Saccharomyces cerevisiae: A useful model host to study fundamental biology of viral replication. *Virus Research* **120:** 49-56. (2006)

Publication V

Saccharomyces cerevisiae: a versatile eukaryotic system in virology. *Microb Cell Fact.* **6:** 32. (2007)

6.1. Publication IV:

Alves-Rodrigues I, **Galão RP**, Meyerhans A, Díez J. *Saccharomyces cerevisiae:* A useful model host to study fundamental biology of viral replicaiton. *Virus Research* **120:** *49-56 (2006).*

Understanding the fundamental steps of virus life cycles including virus–host interactions is essential for the design of effective antiviral strategies. Such understanding has been deferred by the complexity of higher eukaryotic host organisms. To circumvent experimental difficulties associated with this, systems were developed to replicate viruses in the yeast *Saccharomyces cerevisiae*. The systems include viruses with RNA and DNA genomes that infect plants, animals and humans. By using the powerful methodologies available for yeast genetic analysis, fundamental processes occurring during virus replication have been brought to light. Here, we review the different viruses able to direct replication and gene expression in yeast and discuss their main contributions in the understanding of virus biology.

Alves-Rodrigues I, Galão RP, Meyerhans A, Díez J. Saccharomyces cerevisiae: a useful model host to study fundamental biology of viral replication. Virus Res. 2006; 120(1-2): 49-56.

6.2 Publication V:

Galão RP, Scheller N, Alves-Rodrigues I, Breinig T, Meyerhans A, Díez J. *Saccharomyces cerevisiae:* A versatile eukaryotic system in virology. *Microb Cell Fact.* ***6:*** *32 (2007).*

The yeast *Saccharomyces cerevisiae* is a well-established model system for understanding fundamental cellular processes relevant to higher eukaryotic organisms. Less known is its value for virus research, an area in which *Saccharomyces cerevisiae* has proven to be very fruitful as well. The present review will discuss the main achievements of yeast-based studies in basic and applied virus research. These include the analysis of the function of individual proteins from important pathogenic viruses, the elucidation of key processes in viral replication through the development of systems that allow the replication of higher eukayotic viruses in yeast, and the use of yeast in antiviral drug development and vaccine production.

Galão RP, Scheller N, Alves-Rodrigues I, Breinig T, Meyerhans A, Díez J. Saccharomyces cerevisiae: a versatile eukaryotic system in virology. Microb Cell Fact. 2007; 6:32.

7. Bibliography

1. Khusial, P., Plaag, R. & Zieve, G.W. LSm proteins form heptameric rings that bind to RNA via repeating motifs. *Trends Biochem. Sci* **30**, 522-528 (2005).
2. Beggs, J.D. Lsm proteins and RNA processing. *Biochem. Soc. Trans* **33**, 433-438 (2005).
3. Wilusz, C.J. & Wilusz, J. Eukaryotic Lsm proteins: lessons from bacteria. *Nat. Struct. Mol. Biol* **12**, 1031-1036 (2005).
4. Hermann, H. et al. snRNP Sm proteins share two evolutionarily conserved sequence motifs which are involved in Sm protein-protein interactions. *EMBO J* **14**, 2076-2088 (1995).
5. Kambach, C. et al. Crystal structures of two Sm protein complexes and their implications for the assembly of the spliceosomal snRNPs. *Cell* **96**, 375-387 (1999).
6. Tharun, S. et al. Yeast Sm-like proteins function in mRNA decapping and decay. *Nature* **404**, 515-518 (2000).
7. Stoecklin, G., Mayo, T. & Anderson, P. ARE-mRNA degradation requires the 5'-3' decay pathway. *EMBO Rep* **7**, 72-77 (2006).
8. Mukherjee, D. et al. The mammalian exosome mediates the efficient degradation of mRNAs that contain AU-rich elements. *EMBO J* **21**, 165-174 (2002).
9. Mullen, T.E. & Marzluff, W.F. Degradation of histone mRNA requires oligouridylation followed by decapping and simultaneous degradation of the mRNA both 5' to 3' and 3' to 5'. *Genes Dev* **22**, 50-65 (2008).
10. Tharun, S., Muhlrad, D., Chowdhury, A. & Parker, R. Mutations in the Saccharomyces cerevisiae LSM1 gene that affect mRNA decapping and 3' end protection. *Genetics* **170**, 33-46 (2005).
11. He, W. & Parker, R. The yeast cytoplasmic LsmI/Pat1p complex protects mRNA 3' termini from partial degradation. *Genetics* **158**, 1445-1455 (2001).
12. Moll, I., Leitsch, D., Steinhauser, T. & Bläsi, U. RNA chaperone activity of the Sm-like Hfq protein. *EMBO Rep* **4**, 284-289 (2003).
13. Vytvytska, O., Moll, I., Kaberdin, V.R., von Gabain, A. & Bläsi, U. Hfq (HF1) stimulates ompA mRNA decay by interfering with ribosome binding. *Genes Dev* **14**, 1109-1118 (2000).
14. Vecerek, B., Moll, I. & Bläsi, U. Translational autocontrol of the Escherichia coli hfq RNA chaperone gene. *RNA* **11**, 976-984 (2005).
15. Folichon, M., Allemand, F., Régnier, P. & Hajnsdorf, E. Stimulation of poly(A) synthesis by Escherichia coli poly(A)polymerase I is correlated with Hfq binding to poly(A) tails. *FEBS J* **272**, 454-463 (2005).
16. Hajnsdorf, E. & Régnier, P. Host factor Hfq of Escherichia coli stimulates elongation of poly(A) tails by poly(A) polymerase I. *Proc. Natl. Acad. Sci. U.S.A* **97**, 1501-1505 (2000).
17. Rajkowitsch, L. et al. RNA chaperones, RNA annealers and RNA helicases. *RNA Biol* **4**, 118-130 (2007).
18. Zúñiga, S., Sola, I., Cruz, J.L.G. & Enjuanes, L. Role of RNA chaperones in virus replication. *Virus Res* **139**, 253-266 (2009).
19. Zhang, D., Abovich, N. & Rosbash, M. A biochemical function for the Sm complex. *Mol. Cell* **7**, 319-329 (2001).
20. Achsel, T. et al. A doughnut-shaped heteromer of human Sm-like proteins

binds to the 3'-end of U6 snRNA, thereby facilitating U4/U6 duplex formation in vitro. *EMBO J* **18**, 5789-5802 (1999).
21. Verdone, L., Galardi, S., Page, D. & Beggs, J.D. Lsm proteins promote regeneration of pre-mRNA splicing activity. *Curr. Biol* **14**, 1487-1491 (2004).
22. Gottesman, S. The small RNA regulators of Escherichia coli: roles and mechanisms*. *Annu. Rev. Microbiol* **58**, 303-328 (2004).
23. Muffler, A., Fischer, D. & Hengge-Aronis, R. The RNA-binding protein HF-I, known as a host factor for phage Qbeta RNA replication, is essential for rpoS translation in Escherichia coli. *Genes Dev* **10**, 1143-1151 (1996).
24. Drepper, T. et al. The Hfq-like protein NrfA of the phototrophic purple bacterium Rhodobacter capsulatus controls nitrogen fixation via regulation of nifA and anfA expression. *FEMS Microbiol. Lett* **215**, 221-227 (2002).
25. Kaminski, P.A., Desnoues, N. & Elmerich, C. The expression of nifA in Azorhizobium caulinodans requires a gene product homologous to Escherichia coli HF-I, an RNA-binding protein involved in the replication of phage Q beta RNA. *Proc. Natl. Acad. Sci. U.S.A* **91**, 4663-4667 (1994).
26. Franze de Fernandez, M.T., Eoyang, L. & August, J.T. Factor fraction required for the synthesis of bacteriophage Qbeta-RNA. *Nature* **219**, 588-590 (1968).
27. Schuppli, D. et al. Altered 3'-terminal RNA structure in phage Qbeta adapted to host factor-less Escherichia coli. *Proc. Natl. Acad. Sci. U.S.A* **94**, 10239-10242 (1997).
28. Mayes, A.E., Verdone, L., Legrain, P. & Beggs, J.D. Characterization of Sm-like proteins in yeast and their association with U6 snRNA. *EMBO J* **18**, 4321-4331 (1999).
29. Stark, H., Dube, P., Lührmann, R. & Kastner, B. Arrangement of RNA and proteins in the spliceosomal U1 small nuclear ribonucleoprotein particle. *Nature* **409**, 539-542 (2001).
30. Vidal, V.P., Verdone, L., Mayes, A.E. & Beggs, J.D. Characterization of U6 snRNA-protein interactions. *RNA* **5**, 1470-1481 (1999).
31. Kufel, J., Bousquet-Antonelli, C., Beggs, J.D. & Tollervey, D. Nuclear pre-mRNA decapping and 5' degradation in yeast require the Lsm2-8p complex. *Mol. Cell. Biol* **24**, 9646-9657 (2004).
32. Bergman, N. et al. Lsm proteins bind and stabilize RNAs containing 5' poly(A) tracts. *Nat. Struct. Mol. Biol* **14**, 824-831 (2007).
33. Tharun, S. & Parker, R. Targeting an mRNA for decapping: displacement of translation factors and association of the Lsm1p-7p complex on deadenylated yeast mRNAs. *Mol. Cell* **8**, 1075-1083 (2001).
34. Chowdhury, A., Mukhopadhyay, J. & Tharun, S. The decapping activator Lsm1p-7p-Pat1p complex has the intrinsic ability to distinguish between oligoadenylated and polyadenylated RNAs. *RNA* **13**, 998-1016 (2007).
35. Schumacher, M.A., Pearson, R.F., Møller, T., Valentin-Hansen, P. & Brennan, R.G. Structures of the pleiotropic translational regulator Hfq and an Hfq-RNA complex: a bacterial Sm-like protein. *EMBO J* **21**, 3546-3556 (2002).
36. Mikulecky, P.J. et al. Escherichia coli Hfq has distinct interaction surfaces for DsrA, rpoS and poly(A) RNAs. *Nat. Struct. Mol. Biol* **11**, 1206-1214 (2004).
37. Lee, T. & Feig, A.L. The RNA binding protein Hfq interacts specifically with tRNAs. *RNA* **14**, 514-523 (2008).
38. Garneau, N.L., Wilusz, J. & Wilusz, C.J. The highways and byways of mRNA decay. *Nat. Rev. Mol. Cell Biol* **8**, 113-126 (2007).

39. Coller, J. & Parker, R. Eukaryotic mRNA decapping. *Annu. Rev. Biochem* **73**, 861-890 (2004).
40. Waterhouse, P.M., Wang, M.B. & Lough, T. Gene silencing as an adaptive defence against viruses. *Nature* **411**, 834-842 (2001).
41. van Hoof, A. & Parker, R. Messenger RNA degradation: beginning at the end. *Curr. Biol* **12**, R285-287 (2002).
42. Maquat, L.E. & Carmichael, G.G. Quality control of mRNA function. *Cell* **104**, 173-176 (2001).
43. Conti, E. & Izaurralde, E. Nonsense-mediated mRNA decay: molecular insights and mechanistic variations across species. *Curr. Opin. Cell Biol* **17**, 316-325 (2005).
44. Frischmeyer, P.A. et al. An mRNA surveillance mechanism that eliminates transcripts lacking termination codons. *Science* **295**, 2258-2261 (2002).
45. Doma, M.K. & Parker, R. Endonucleolytic cleavage of eukaryotic mRNAs with stalls in translation elongation. *Nature* **440**, 561-564 (2006).
46. van Hoof, A., Frischmeyer, P.A., Dietz, H.C. & Parker, R. Exosome-mediated recognition and degradation of mRNAs lacking a termination codon. *Science* **295**, 2262-2264 (2002).
47. Inada, T. & Aiba, H. Translation of aberrant mRNAs lacking a termination codon or with a shortened 3'-UTR is repressed after initiation in yeast. *EMBO J* **24**, 1584-1595 (2005).
48. Muhlrad, D. & Parker, R. Mutations affecting stability and deadenylation of the yeast MFA2 transcript. *Genes Dev* **6**, 2100-2111 (1992).
49. Shyu, A.B., Belasco, J.G. & Greenberg, M.E. Two distinct destabilizing elements in the c-fos message trigger deadenylation as a first step in rapid mRNA decay. *Genes Dev* **5**, 221-231 (1991).
50. Liu, H., Rodgers, N.D., Jiao, X. & Kiledjian, M. The scavenger mRNA decapping enzyme DcpS is a member of the HIT family of pyrophosphatases. *EMBO J* **21**, 4699-4708 (2002).
51. Houseley, J., LaCava, J. & Tollervey, D. RNA-quality control by the exosome. *Nat. Rev. Mol. Cell Biol* **7**, 529-539 (2006).
52. Steiger, M., Carr-Schmid, A., Schwartz, D.C., Kiledjian, M. & Parker, R. Analysis of recombinant yeast decapping enzyme. *RNA* **9**, 231-238 (2003).
53. Fenger-Grøn, M., Fillman, C., Norrild, B. & Lykke-Andersen, J. Multiple processing body factors and the ARE binding protein TTP activate mRNA decapping. *Mol. Cell* **20**, 905-915 (2005).
54. Simon, E., Camier, S. & Séraphin, B. New insights into the control of mRNA decapping. *Trends Biochem. Sci* **31**, 241-243 (2006).
55. Franks, T.M. & Lykke-Andersen, J. The control of mRNA decapping and P-body formation. *Mol. Cell* **32**, 605-615 (2008).
56. Parker, R. & Sheth, U. P bodies and the control of mRNA translation and degradation. *Mol. Cell* **25**, 635-646 (2007).
57. Eulalio, A., Behm-Ansmant, I. & Izaurralde, E. P bodies: at the crossroads of post-transcriptional pathways. *Nat. Rev. Mol. Cell Biol* **8**, 9-22 (2007).
58. Cougot, N., van Dijk, E., Babajko, S. & Séraphin, B. 'Cap-tabolism'. *Trends Biochem. Sci* **29**, 436-444 (2004).
59. Beelman, C.A. & Parker, R. Differential effects of translational inhibition in cis and in trans on the decay of the unstable yeast MFA2 mRNA. *J. Biol. Chem* **269**, 9687-9692 (1994).
60. Schwartz, D.C. & Parker, R. Mutations in translation initiation factors lead to

increased rates of deadenylation and decapping of mRNAs in Saccharomyces cerevisiae. *Mol. Cell. Biol* **19**, 5247-5256 (1999).
61. Brengues, M., Teixeira, D. & Parker, R. Movement of eukaryotic mRNAs between polysomes and cytoplasmic processing bodies. *Science* **310**, 486-489 (2005).
62. Schwartz, D.C. & Parker, R. mRNA decapping in yeast requires dissociation of the cap binding protein, eukaryotic translation initiation factor 4E. *Mol. Cell. Biol* **20**, 7933-7942 (2000).
63. Decker, C.J., Teixeira, D. & Parker, R. Edc3p and a glutamine/asparagine-rich domain of Lsm4p function in processing body assembly in Saccharomyces cerevisiae. *J. Cell Biol* **179**, 437-449 (2007).
64. Tritschler, F. et al. A divergent Sm fold in EDC3 proteins mediates DCP1 binding and P-body targeting. *Mol. Cell. Biol* **27**, 8600-8611 (2007).
65. Teixeira, D., Sheth, U., Valencia-Sanchez, M.A., Brengues, M. & Parker, R. Processing bodies require RNA for assembly and contain nontranslating mRNAs. *RNA* **11**, 371-382 (2005).
66. Weston, A. & Sommerville, J. Xp54 and related (DDX6-like) RNA helicases: roles in messenger RNP assembly, translation regulation and RNA degradation. *Nucleic Acids Res* **34**, 3082-3094 (2006).
67. Cordin, O., Banroques, J., Tanner, N.K. & Linder, P. The DEAD-box protein family of RNA helicases. *Gene* **367**, 17-37 (2006).
68. Bleichert, F. & Baserga, S.J. The long unwinding road of RNA helicases. *Mol. Cell* **27**, 339-352 (2007).
69. Coller, J. & Parker, R. General translational repression by activators of mRNA decapping. *Cell* **122**, 875-886 (2005).
70. Coller, J.M., Tucker, M., Sheth, U., Valencia-Sanchez, M.A. & Parker, R. The DEAD box helicase, Dhh1p, functions in mRNA decapping and interacts with both the decapping and deadenylase complexes. *RNA* **7**, 1717-1727 (2001).
71. Chu, C. & Rana, T.M. Translation repression in human cells by microRNA-induced gene silencing requires RCK/p54. *PLoS Biol* **4**, e210 (2006).
72. Pilkington, G.R. & Parker, R. Pat1 contains distinct functional domains that promote P-body assembly and activation of decapping. *Mol. Cell. Biol* **28**, 1298-1312 (2008).
73. Bouveret, E., Rigaut, G., Shevchenko, A., Wilm, M. & Séraphin, B. A Sm-like protein complex that participates in mRNA degradation. *EMBO J* **19**, 1661-1671 (2000).
74. Wyers, F., Minet, M., Dufour, M.E., Vo, L.T. & Lacroute, F. Deletion of the PAT1 gene affects translation initiation and suppresses a PAB1 gene deletion in yeast. *Mol. Cell. Biol* **20**, 3538-3549 (2000).
75. Bonnerot, C., Boeck, R. & Lapeyre, B. The two proteins Pat1p (Mrt1p) and Spb8p interact in vivo, are required for mRNA decay, and are functionally linked to Pab1p. *Mol. Cell. Biol* **20**, 5939-5946 (2000).
76. Eulalio, A., Behm-Ansmant, I., Schweizer, D. & Izaurralde, E. P-body formation is a consequence, not the cause, of RNA-mediated gene silencing. *Mol. Cell. Biol* **27**, 3970-3981 (2007).
77. Ahlquist, P. Parallels among positive-strand RNA viruses, reverse-transcribing viruses and double-stranded RNA viruses. *Nat. Rev. Microbiol* **4**, 371-382 (2006).
78. Ahlquist, P., Noueiry, A.O., Lee, W., Kushner, D.B. & Dye, B.T. Host factors

in positive-strand RNA virus genome replication. *J. Virol* **77**, 8181-8186 (2003).
79. Denison, M.R. Seeking membranes: positive-strand RNA virus replication complexes. *PLoS Biol* **6**, e270 (2008).
80. Miller, S. & Krijnse-Locker, J. Modification of intracellular membrane structures for virus replication. *Nat. Rev. Microbiol* **6**, 363-374 (2008).
81. Noueiry, A.O. & Ahlquist, P. Brome mosaic virus RNA replication: revealing the role of the host in RNA virus replication. *Annu Rev Phytopathol* **41**, 77-98 (2003).
82. Fields, B.N., Knipe, D.M. & Howley, P.M. *Fields' virology*. (Lippincott Williams & Wilkins: 2007).
83. Tan, S., Ganji, G., Paeper, B., Proll, S. & Katze, M.G. Systems biology and the host response to viral infection. *Nat. Biotechnol* **25**, 1383-1389 (2007).
84. de Angelis, D.S.A., Freire, W.S., Pannuti, C.S., Succi, R.C.D.M. & Machado, D.M. CCR5 genotypes and progression to HIV disease in perinatally infected children. *Braz J Infect Dis* **11**, 196-198 (2007).
85. Wheeler, J., McHale, M., Jackson, V. & Penny, M. Assessing theoretical risk and benefit suggested by genetic association studies of CCR5: experience in a drug development programme for maraviroc. *Antivir. Ther. (Lond.)* **12**, 233-245 (2007).
86. Ahlquist, P. & Janda, M. cDNA cloning and in vitro transcription of the complete brome mosaic virus genome. *Mol. Cell. Biol* **4**, 2876-2882 (1984).
87. Ahlquist, P. Bromovirus RNA replication and transcription. *Curr. Opin. Genet. Dev* **2**, 71-76 (1992).
88. Ahola, T. & Ahlquist, P. Putative RNA capping activities encoded by brome mosaic virus: methylation and covalent binding of guanylate by replicase protein 1a. *J. Virol* **73**, 10061-10069 (1999).
89. Kroner, P.A., Young, B.M. & Ahlquist, P. Analysis of the role of brome mosaic virus 1a protein domains in RNA replication, using linker insertion mutagenesis. *J. Virol* **64**, 6110-6120 (1990).
90. Schwartz, M. et al. A positive-strand RNA virus replication complex parallels form and function of retrovirus capsids. *Mol. Cell* **9**, 505-514 (2002).
91. Schwartz, M., Chen, J., Lee, W., Janda, M. & Ahlquist, P. Alternate, virus-induced membrane rearrangements support positive-strand RNA virus genome replication. *Proc. Natl. Acad. Sci. U.S.A* **101**, 11263-11268 (2004).
92. Ahola, T., den Boon, J.A. & Ahlquist, P. Helicase and capping enzyme active site mutations in brome mosaic virus protein 1a cause defects in template recruitment, negative-strand RNA synthesis, and viral RNA capping. *J. Virol* **74**, 8803-8811 (2000).
93. Chen, J. & Ahlquist, P. Brome mosaic virus polymerase-like protein 2a is directed to the endoplasmic reticulum by helicase-like viral protein 1a. *J. Virol* **74**, 4310-4318 (2000).
94. Chen, J., Noueiry, A. & Ahlquist, P. Brome mosaic virus Protein 1a recruits viral RNA2 to RNA replication through a 5' proximal RNA2 signal. *J. Virol* **75**, 3207-3219 (2001).
95. Chen, J., Noueiry, A. & Ahlquist, P. An alternate pathway for recruiting template RNA to the brome mosaic virus RNA replication complex. *J. Virol* **77**, 2568-2577 (2003).
96. Sullivan, M.L. & Ahlquist, P. cis-acting signals in Bromovirus RNA replication and gene expression: networking with viral proteins and host factors. *Seminars*

in Virology **8**, 221-230 (1997).
97. Baumstark, T. & Ahlquist, P. The brome mosaic virus RNA3 intergenic replication enhancer folds to mimic a tRNA TpsiC-stem loop and is modified in vivo. *RNA* **7**, 1652-1670 (2001).
98. Sullivan, M.L. & Ahlquist, P. A brome mosaic virus intergenic RNA3 replication signal functions with viral replication protein 1a to dramatically stabilize RNA in vivo. *J. Virol* **73**, 2622-2632 (1999).
99. French, R. & Ahlquist, P. Intercistronic as well as terminal sequences are required for efficient amplification of brome mosaic virus RNA3. *J. Virol* **61**, 1457-1465 (1987).
100. French, R. & Ahlquist, P. Characterization and engineering of sequences controlling in vivo synthesis of brome mosaic virus subgenomic RNA. *J. Virol* **62**, 2411-2420 (1988).
101. Barends, S. et al. tRNA-like structure regulates translation of Brome mosaic virus RNA. *J. Virol* **78**, 4003-4010 (2004).
102. Choi, Y.G., Dreher, T.W. & Rao, A.L.N. tRNA elements mediate the assembly of an icosahedral RNA virus. *Proc. Natl. Acad. Sci. U.S.A* **99**, 655-660 (2002).
103. Felden, B., Florentz, C., Westhof, E. & Giegé, R. Non-canonical substrates of aminoacyl-tRNA synthetases: the tRNA-like structure of brome mosaic virus genomic RNA. *Biochimie* **75**, 1143-1157 (1993).
104. Goffeau, A. et al. Life with 6000 genes. *Science* **274**, 546, 563-567 (1996).
105. Parsons, A.B., Geyer, R., Hughes, T.R. & Boone, C. Yeast genomics and proteomics in drug discovery and target validation. *Prog Cell Cycle Res* **5**, 159-166 (2003).
106. Foury, F. Human genetic diseases: a cross-talk between man and yeast. *Gene* **195**, 1-10 (1997).
107. Hartwell, L.H. & Kastan, M.B. Cell cycle control and cancer. *Science* **266**, 1821-1828 (1994).
108. Tsukuda, T., Fleming, A.B., Nickoloff, J.A. & Osley, M.A. Chromatin remodelling at a DNA double-strand break site in Saccharomyces cerevisiae. *Nature* **438**, 379-383 (2005).
109. Alves-Rodrigues, I., Galão, R.P., Meyerhans, A. & Díez, J. Saccharomyces cerevisiae: a useful model host to study fundamental biology of viral replication. *Virus Res* **120**, 49-56 (2006).
110. Galao, R.P. et al. Saccharomyces cerevisiae: a versatile eukaryotic system in virology. *Microb. Cell Fact* **6**, 32 (2007).
111. Janda, M. & Ahlquist, P. RNA-dependent replication, transcription, and persistence of brome mosaic virus RNA replicons in S. cerevisiae. *Cell* **72**, 961-970 (1993).
112. Angeletti, P.C., Kim, K., Fernandes, F.J. & Lambert, P.F. Stable replication of papillomavirus genomes in Saccharomyces cerevisiae. *J. Virol* **76**, 3350-3358 (2002).
113. Panavas, T. & Nagy, P.D. Yeast as a model host to study replication and recombination of defective interfering RNA of Tomato bushy stunt virus. *Virology* **314**, 315-325 (2003).
114. Pantaleo, V., Rubino, L. & Russo, M. Replication of Carnation Italian ringspot virus defective interfering RNA in Saccharomyces cerevisiae. *J. Virol* **77**, 2116-2123 (2003).
115. Price, B.D., Rueckert, R.R. & Ahlquist, P. Complete replication of an animal

virus and maintenance of expression vectors derived from it in Saccharomyces cerevisiae. *Proc. Natl. Acad. Sci. U.S.A* **93**, 9465-9470 (1996).
116. Raghavan, V., Malik, P.S., Choudhury, N.R. & Mukherjee, S.K. The DNA-A component of a plant geminivirus (Indian mung bean yellow mosaic virus) replicates in budding yeast cells. *J. Virol* **78**, 2405-2413 (2004).
117. Zhao, K. & Frazer, I.H. Replication of bovine papillomavirus type 1 (BPV-1) DNA in Saccharomyces cerevisiae following infection with BPV-1 virions. *J. Virol* **76**, 3359-3364 (2002).
118. Navarro, B., Russo, M., Pantaleo, V. & Rubino, L. Cytological analysis of Saccharomyces cerevisiae cells supporting cymbidium ringspot virus defective interfering RNA replication. *J. Gen. Virol* **87**, 705-714 (2006).
119. Krol, M.A. et al. RNA-controlled polymorphism in the in vivo assembly of 180-subunit and 120-subunit virions from a single capsid protein. *Proc. Natl. Acad. Sci. U.S.A* **96**, 13650-13655 (1999).
120. Ishikawa, M., Janda, M., Krol, M.A. & Ahlquist, P. In vivo DNA expression of functional brome mosaic virus RNA replicons in Saccharomyces cerevisiae. *J. Virol* **71**, 7781-7790 (1997).
121. Janda, M. & Ahlquist, P. Brome mosaic virus RNA replication protein 1a dramatically increases in vivo stability but not translation of viral genomic RNA3. *Proc. Natl. Acad. Sci. U.S.A* **95**, 2227-2232 (1998).
122. Kushner, D.B. et al. Systematic, genome-wide identification of host genes affecting replication of a positive-strand RNA virus. *Proc. Natl. Acad. Sci. U.S.A* **100**, 15764-15769 (2003).
123. Noueiry, A.O., Diez, J., Falk, S.P., Chen, J. & Ahlquist, P. Yeast Lsm1p-7p/Pat1p deadenylation-dependent mRNA-decapping factors are required for brome mosaic virus genomic RNA translation. *Mol. Cell. Biol* **23**, 4094-4106 (2003).
124. Díez, J., Ishikawa, M., Kaido, M. & Ahlquist, P. Identification and characterization of a host protein required for efficient template selection in viral RNA replication. *Proc. Natl. Acad. Sci. U.S.A* **97**, 3913-3918 (2000).
125. Mas, A., Alves-Rodrigues, I., Noueiry, A., Ahlquist, P. & Díez, J. Host deadenylation-dependent mRNA decapping factors are required for a key step in brome mosaic virus RNA replication. *J. Virol* **80**, 246-251 (2006).
126. Hepatitis C--global prevalence (update). *Wkly. Epidemiol. Rec* **75**, 18-19 (2000).
127. Poynard, T. et al. Natural history of HCV infection. *Baillieres Best Pract Res Clin Gastroenterol* **14**, 211-228 (2000).
128. Lauer, G.M. & Walker, B.D. Hepatitis C virus infection. *N. Engl. J. Med* **345**, 41-52 (2001).
129. Hoofnagle, J.H. & di Bisceglie, A.M. The treatment of chronic viral hepatitis. *N. Engl. J. Med* **336**, 347-356 (1997).
130. Chevaliez, S. & Pawlotsky, J. Interferons and their use in persistent viral infections. *Handb Exp Pharmacol* 203-241 (2009).doi:10.1007/978-3-540-79086-0_8
131. Dash, S. et al. HCV-hepatocellular carcinoma: new findings and hope for effective treatment. *Microsc. Res. Tech* **68**, 130-148 (2005).
132. Sakamoto, N. & Watanabe, M. New therapeutic approaches to hepatitis C virus. *J. Gastroenterol* (2009).doi:10.1007/s00535-009-0084-0
133. Sakamoto, H. et al. Host sphingolipid biosynthesis as a target for hepatitis C virus therapy. *Nat. Chem. Biol* **1**, 333-337 (2005).
134. Appel, N., Schaller, T., Penin, F. & Bartenschlager, R. From structure to

function: new insights into hepatitis C virus RNA replication. *J. Biol. Chem* **281**, 9833-9836 (2006).
135. Bartenschlager, R., Frese, M. & Pietschmann, T. Novel insights into hepatitis C virus replication and persistence. *Adv. Virus Res* **63**, 71-180 (2004).
136. Tellinghuisen, T.L. & Rice, C.M. Interaction between hepatitis C virus proteins and host cell factors. *Curr. Opin. Microbiol* **5**, 419-427 (2002).
137. Sakai, A. et al. The p7 polypeptide of hepatitis C virus is critical for infectivity and contains functionally important genotype-specific sequences. *Proc. Natl. Acad. Sci. U.S.A* **100**, 11646-11651 (2003).
138. Steinmann, E. et al. Hepatitis C virus p7 protein is crucial for assembly and release of infectious virions. *PLoS Pathog* **3**, e103 (2007).
139. Jones, C.T., Murray, C.L., Eastman, D.K., Tassello, J. & Rice, C.M. Hepatitis C Virus p7 and NS2 Proteins Are Essential for Production of Infectious Virus. *J. Virol.* **81**, 8374-8383 (2007).
140. Jirasko, V. et al. Structural and Functional Characterization of Nonstructural Protein 2 for Its Role in Hepatitis C Virus Assembly. *J. Biol. Chem.* **283**, 28546-28562 (2008).
141. Lohmann, V. et al. Replication of subgenomic hepatitis C virus RNAs in a hepatoma cell line. *Science* **285**, 110-113 (1999).
142. Bartenschlager, R. & Sparacio, S. Hepatitis C virus molecular clones and their replication capacity in vivo and in cell culture. *Virus Res* **127**, 195-207 (2007).
143. Tellinghuisen, T.L., Evans, M.J., von Hahn, T., You, S. & Rice, C.M. Studying hepatitis C virus: making the best of a bad virus. *J. Virol* **81**, 8853-8867 (2007).
144. Lukavsky, P.J. Structure and function of HCV IRES domains. *Virus Res* **139**, 166-171 (2009).
145. Henke, J.I. et al. microRNA-122 stimulates translation of hepatitis C virus RNA. *EMBO J* **27**, 3300-3310 (2008).
146. Tuplin, A., Wood, J., Evans, D.J., Patel, A.H. & Simmonds, P. Thermodynamic and phylogenetic prediction of RNA secondary structures in the coding region of hepatitis C virus. *RNA* **8**, 824-841 (2002).
147. McMullan, L.K. et al. Evidence for a functional RNA element in the hepatitis C virus core gene. *Proc. Natl. Acad. Sci. U.S.A* **104**, 2879-2884 (2007).
148. You, S., Stump, D.D., Branch, A.D. & Rice, C.M. A cis-acting replication element in the sequence encoding the NS5B RNA-dependent RNA polymerase is required for hepatitis C virus RNA replication. *J. Virol* **78**, 1352-1366 (2004).
149. Friebe, P., Boudet, J., Simorre, J. & Bartenschlager, R. Kissing-loop interaction in the 3' end of the hepatitis C virus genome essential for RNA replication. *J. Virol* **79**, 380-392 (2005).
150. You, S. & Rice, C.M. 3' RNA elements in hepatitis C virus replication: kissing partners and long poly(U). *J. Virol* **82**, 184-195 (2008).
151. Evans, M.J. et al. Claudin-1 is a hepatitis C virus co-receptor required for a late step in entry. *Nature* **446**, 801-805 (2007).
152. Pileri, P. et al. Binding of hepatitis C virus to CD81. *Science* **282**, 938-941 (1998).
153. Scarselli, E. et al. The human scavenger receptor class B type I is a novel candidate receptor for the hepatitis C virus. *EMBO J* **21**, 5017-5025 (2002).
154. Zeisel, M.B., Barth, H., Schuster, C. & Baumert, T.F. Hepatitis C virus entry: molecular mechanisms and targets for antiviral therapy. *Front. Biosci* **14**, 3274-3285 (2009).

155. Blanchard, E. et al. Hepatitis C virus entry depends on clathrin-mediated endocytosis. *J. Virol* **80**, 6964-6972 (2006).
156. Roingeard, P. & Hourioux, C. Hepatitis C virus core protein, lipid droplets and steatosis. *J. Viral Hepat* **15**, 157-164 (2008).
157. Gastaminza, P. et al. Cellular determinants of hepatitis C virus assembly, maturation, degradation, and secretion. *J. Virol* **82**, 2120-2129 (2008).
158. Evans, M.J., Rice, C.M. & Goff, S.P. Phosphorylation of hepatitis C virus nonstructural protein 5A modulates its protein interactions and viral RNA replication. *Proc. Natl. Acad. Sci. U.S.A* **101**, 13038-13043 (2004).
159. Isken, O. et al. Nuclear factors are involved in hepatitis C virus RNA replication. *RNA* **13**, 1675-1692 (2007).
160. Liu, H.M. et al. SYNCRIP (synaptotagmin-binding, cytoplasmic RNA-interacting protein) is a host factor involved in hepatitis C virus RNA replication. *Virology* **386**, 249-256 (2009).
161. Wakita, T. et al. Production of infectious hepatitis C virus in tissue culture from a cloned viral genome. *Nat. Med* **11**, 791-796 (2005).
162. Zhong, J. et al. Robust hepatitis C virus infection in vitro. *Proc. Natl. Acad. Sci. U.S.A* **102**, 9294-9299 (2005).
163. Lindenbach, B.D. et al. Complete replication of hepatitis C virus in cell culture. *Science* **309**, 623-626 (2005).
164. Pietschmann, T. et al. Persistent and transient replication of full-length hepatitis C virus genomes in cell culture. *J. Virol* **76**, 4008-4021 (2002).
165. Ikeda, M., Yi, M., Li, K. & Lemon, S.M. Selectable subgenomic and genome-length dicistronic RNAs derived from an infectious molecular clone of the HCV-N strain of hepatitis C virus replicate efficiently in cultured Huh7 cells. *J. Virol* **76**, 2997-3006 (2002).
166. Blight, K.J., McKeating, J.A., Marcotrigiano, J. & Rice, C.M. Efficient replication of hepatitis C virus genotype 1a RNAs in cell culture. *J. Virol* **77**, 3181-3190 (2003).
167. Kato, T. et al. Sequence analysis of hepatitis C virus isolated from a fulminant hepatitis patient. *J. Med. Virol* **64**, 334-339 (2001).
168. Brass, A.L. et al. Identification of host proteins required for HIV infection through a functional genomic screen. *Science* **319**, 921-926 (2008).
169. Nguyen, D.G., Wolff, K.C., Yin, H., Caldwell, J.S. & Kuhen, K.L. "UnPAKing" human immunodeficiency virus (HIV) replication: using small interfering RNA screening to identify novel cofactors and elucidate the role of group I PAKs in HIV infection. *J. Virol* **80**, 130-137 (2006).
170. Randall, G. et al. Cellular cofactors affecting hepatitis C virus infection and replication. *Proc. Natl. Acad. Sci. U.S.A* **104**, 12884-12889 (2007).
171. Ng, T.I. et al. Identification of host genes involved in hepatitis C virus replication by small interfering RNA technology. *Hepatology* **45**, 1413-1421 (2007).
172. Krishnan, M.N. et al. RNA interference screen for human genes associated with West Nile virus infection. *Nature* **455**, 242-245 (2008).
173. Hao, L. et al. Drosophila RNAi screen identifies host genes important for influenza virus replication. *Nature* **454**, 890-893 (2008).
174. Zaric, B. et al. Reconstitution of two recombinant LSm protein complexes reveals aspects of their architecture, assembly, and function. *J. Biol. Chem* **280**, 16066-16075 (2005).
175. Chowdhury, A. & Tharun, S. Activation of decapping involves binding of the

mRNA and facilitation of the post-binding steps by the Lsm1-7-Pat1 complex. *RNA* **15**, 1837-1848 (2009).
176. Walter, N.G., Yang, N. & Burke, J.M. Probing non-selective cation binding in the hairpin ribozyme with Tb(III). *J. Mol. Biol* **298**, 539-555 (2000).
177. Babaylova, E. et al. Positioning of subdomain IIId and apical loop of domain II of the hepatitis C IRES on the human 40S ribosome. *Nucleic Acids Res* **37**, 1141-1151 (2009).
178. Ina, Y., Mizokami, M., Ohba, K. & Gojobori, T. Reduction of synonymous substitutions in the core protein gene of hepatitis C virus. *J. Mol. Evol* **38**, 50-56 (1994).
179. Smith, D.B. & Simmonds, P. Characteristics of nucleotide substitution in the hepatitis C virus genome: constraints on sequence change in coding regions at both ends of the genome. *J. Mol. Evol* **45**, 238-246 (1997).
180. Vassilaki, N. et al. Role of the hepatitis C virus core+1 open reading frame and core cis-acting RNA elements in viral RNA translation and replication. *J. Virol* **82**, 11503-11515 (2008).
181. Barber, G.N. The NFAR's (nuclear factors associated with dsRNA): evolutionarily conserved members of the dsRNA binding protein family. *RNA Biol* **6**, 35-39 (2009).
182. Fischer, N. & Weis, K. The DEAD box protein Dhh1 stimulates the decapping enzyme Dcp1. *EMBO J* **21**, 2788-2797 (2002).
183. Bartels, C.L. & Tsongalis, G.J. MicroRNAs: novel biomarkers for human cancer. *Clin. Chem* **55**, 623-631 (2009).
184. Sheth, U. & Parker, R. Decapping and decay of messenger RNA occur in cytoplasmic processing bodies. *Science* **300**, 805-808 (2003).
185. Andrei, M.A. et al. A role for eIF4E and eIF4E-transporter in targeting mRNPs to mammalian processing bodies. *RNA* **11**, 717-727 (2005).
186. Griffith, J.L. et al. Functional genomics reveals relationships between the retrovirus-like Ty1 element and its host Saccharomyces cerevisiae. *Genetics* **164**, 867-879 (2003).
187. Aye, M. et al. Host factors that affect Ty3 retrotransposition in Saccharomyces cerevisiae. *Genetics* **168**, 1159-1176 (2004).
188. Irwin, B. et al. Retroviruses and yeast retrotransposons use overlapping sets of host genes. *Genome Res* **15**, 641-654 (2005).
189. Abaza, I. & Gebauer, F. Trading translation with RNA-binding proteins. *RNA* **14**, 404-409 (2008).
190. Beckham, C.J. & Parker, R. P bodies, stress granules, and viral life cycles. *Cell Host Microbe* **3**, 206-212 (2008).
191. Beckham, C.J. et al. Interactions between brome mosaic virus RNAs and cytoplasmic processing bodies. *J. Virol* **81**, 9759-9768 (2007).
192. Wang, X. et al. Brome mosaic virus 1a nucleoside triphosphatase/helicase domain plays crucial roles in recruiting RNA replication templates. *J. Virol* **79**, 13747-13758 (2005).
193. Wilhelm, J.E., Buszczak, M. & Sayles, S. Efficient protein trafficking requires trailer hitch, a component of a ribonucleoprotein complex localized to the ER in Drosophila. *Dev. Cell* **9**, 675-685 (2005).
194. Mamiya, N. & Worman, H.J. Hepatitis C virus core protein binds to a DEAD box RNA helicase. *J. Biol. Chem* **274**, 15751-15756 (1999).
195. You, L.R. et al. Hepatitis C virus core protein interacts with cellular putative RNA helicase. *J. Virol* **73**, 2841-2853 (1999).

196. Ariumi, Y. et al. DDX3 DEAD-box RNA helicase is required for hepatitis C virus RNA replication. *J. Virol* **81**, 13922-13926 (2007).
197. Randall, G. et al. Cellular cofactors affecting hepatitis C virus infection and replication. *Proc. Natl. Acad. Sci. U.S.A* **104**, 12884-12889 (2007).
198. Jopling, C.L., Yi, M., Lancaster, A.M., Lemon, S.M. & Sarnow, P. Modulation of hepatitis C virus RNA abundance by a liver-specific MicroRNA. *Science* **309**, 1577-1581 (2005).
199. Liu, J., Valencia-Sanchez, M.A., Hannon, G.J. & Parker, R. MicroRNA-dependent localization of targeted mRNAs to mammalian P-bodies. *Nat. Cell Biol* **7**, 719-723 (2005).
200. Pillai, R.S. MicroRNA function: multiple mechanisms for a tiny RNA? *RNA* **11**, 1753-1761 (2005).