Dissertation zur Erlangung des Doktorgrades
der Fakultät für Chemie und Pharmazie
der Ludwig-Maximilians-Universität München

Isolation and characterization of *in vivo* RNA ligands for RIG-I and MDA5 from measles virus-infected cells.

Simon Runge
aus
Hannover, Deutschland

2013

Erklärung

Diese Dissertation wurde im Sinne von § 7 der Promotionsordnung vom 28. November 2011 von Herrn Prof. Dr. Karl-Peter Hopfner betreut.

Eidesstattliche Versicherung

Diese Dissertation wurde eigenständig und ohne unerlaubte Hilfe erarbeitet.

München, 27.09.2013

_____

Simon Runge

Dissertation eingereicht am 01. Oktober 2013

1. Gutachter: Prof. Dr. Karl-Peter Hopfner
2. Gutachter: Prof. Dr. Karl-Klaus Conzelmann

Mündliche Prüfung am 05. Dezember 2013

This thesis has been prepared from January 2010 to September 2013 in the laboratory of Professor Dr. Karl-Peter Hopfner at the Gene Center of the Ludwig-Maximilians-Universität München (LMU).

Publications

Parts of this thesis are published in the Bachelor thesis *"Development of an analysis pipeline for high-throughput genome-wide PAR-CLIP measurements of virus-infected cells."* (Hembach, 2012).

Parts of this thesis are in the process of being published:

Runge S., Sparrer K.M.J., Lässig C., Hembach K., Baum A., Garcia-Sastre A., Söding J., Conzelmann K.K., Hopfner K.P. (2013). ***In vivo ligands of MDA5 and RIG-I in virus-infected cells.*** Manuscript submitted.

"Wie alles sich zum Ganzen webt,

Eins in dem andern wirkt und lebt!"

Goethe (1749-1832), Faust I

Meinen Eltern

Table of contents

1. Introduction

1.1. Basic principles in immunity

The mammalian immune system is a sophisticated network comprising a plethora of host factors and cells, which orchestrate a complex immune response upon pathogen invasion. Immune signaling can be generally divided into two distinct branches: the innate and the adaptive immune response. However, innate and adaptive immunity should not be considered as separate systems because the initial recognition of a pathogen by innate immune factors triggers a response that leads to the activation of adaptive immunity. Like gears in clockwork, innate and adaptive immunity intertwine to provide a fast and efficient protection against pathogen infection.

1.1.1. Innate immunity

The early phase of response against pathogens is controlled by the innate immune system, which is a universal and ancient form of host defense against infection. It includes a variety of defense strategies and is present at all times in the host. Innate immune responses are not specific to a particular pathogen but depend on a variety of molecules and phagocytic cells, including macrophages and neutrophiles, which recognize conserved molecular structures of microorganisms known as **p**athogen-**a**ssociated **m**olecular **p**atterns (PAMPs) (Janeway Jr. and Medzhitov, 2002). PAMPs are sensed by a large superfamily of germline-encoded receptor proteins generally called **p**attern **r**ecognition **r**eceptors (PRRs). These receptors include Toll-like receptors (TLRs), RIG-I-like receptors (RLRs), Nod-like

receptors (NLRs) and others (Melchjorsen, 2013). In addition, PRRs could recognize cell-derived **d**anger-**a**ssociated **m**olecular **p**atterns (DAMPs), which are expressed by an infected cell. Effective sensing of PAMPs and DAMPs rapidly induces host immune responses via complex intracellular signal transduction pathways that culminate in the expression of type I interferons (IFNs) and proinflammatory cytokines (Kawai and Akira, 2010; Medzhitov, 2007; Takeuchi and Akira, 2010). Cytokines are the key players in the immune response by modulating cell activation and differentiation and attracting immune cells to the sites of infection. Especially, type I IFNs and proinflammatory cytokines play a major role during infection by inducing the expression of numerous antiviral proteins (Takeuchi and Akira, 2010).

1.1.2. Regulation of adaptive immunity by the innate immune response

As mentioned above, the innate immune system is capable to mount an effective defense against infectious agents through the initiation of adaptive immunity. The adaptive immune response is mediated through the generation of antigen-specific B and T lymphocytes and provides long lasting protection against a pathogen by immunological memory (Hoebe et al., 2004; Iwasaki and Medzhitov, 2010).

PRRs that recognize microbial pathogens are either expressed as transmembrane receptors (e.g. TLRs) or as cytosolic factors (e.g. RLRs). Most PRRs that are capable to activate the transcription factors NF-κB and IRFs are in principle able to induce T and B cell responses (Iwasaki and Medzhitov, 2004). However, the relative contribution of different PRRs to the induction of an adaptive immune response is not fully understood.

1.2. Pattern recognition receptors

To date several families of PRRs (e.g. TLRs, RLRs, NLRs, and cytosolic DNA sensors) are known to play a crucial role in host defense mechanisms. This section gives a brief overview about recent advances in the field of PAMP recognition by PRRs and the signaling pathways activated thereby.

1.2.1. Toll-like receptor (TLR) proteins

TLRs are the best characterized PRRs and are considered to be the primary sensors of pathogens. In humans, 10 TLR family members have been identified since the discovery of toll proteins in flies in 1996 (Lemaitre et al., 1996). TLRs belong to the family of type I membrane glycoproteins that consist of an extracellular **l**eucine **r**ich **r**epeat (LRR) domain, which is required for PAMP recognition and a cytoplasmic **T**oll/**i**nterleukin-1 **r**eceptor (TIR) domain, responsible for downstream signaling (Kawai and Akira, 2011). TLRs are expressed on various immune cells including macrophages, dendritic cells, B and T lymphocytes, and even on non-immune cells like fibroblasts and epithelial cells. Expression of these receptors is modulated in response to pathogens and a variety of cytokines and may occur on extracellular or intracellular membranes.

TLR 1, 2, 4, 5, and 6 are predominately expressed on the cell surface and recognize PAMPs of bacterial, fungal, and protozoal origin. In contrast, TLR 3, 7, and 9 are exclusively expressed on endocytic compartments recognizing nucleic acids derived from various bacteria and viruses. TLR 10 is not well characterized and its function remains unclear (Takeuchi and Akira, 2007; Kumar et al., 2009; Kawai and Akira, 2010).

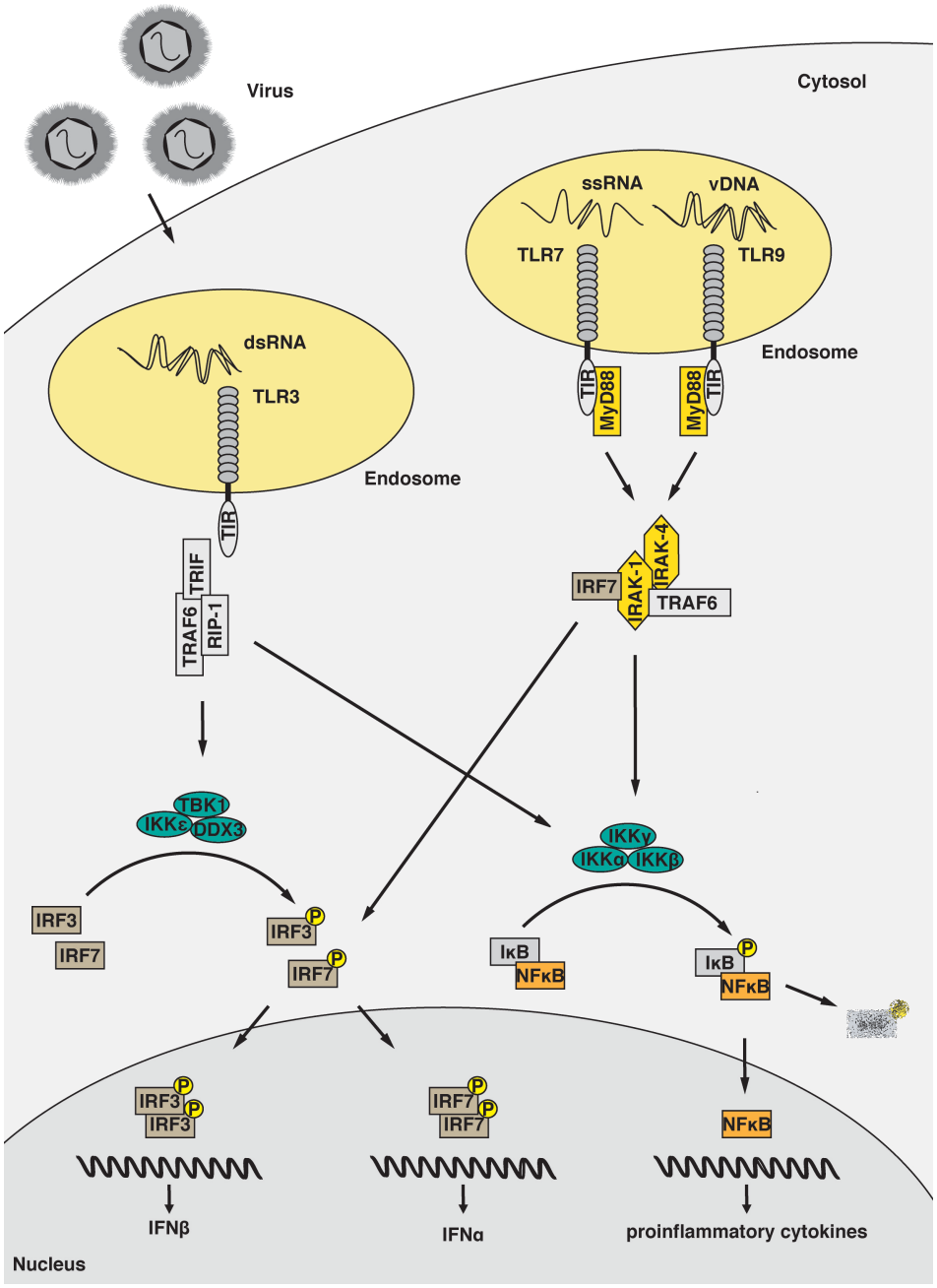


Figure 1.1. TLR signaling in viral infection. Viral nucleic acids are recognized by TLR3 (dsRNA), TLR7 (ssRNA, and TLR9 (vDNA) in endosomal compartments. Upon ligand binding, TLRs elicit a signaling cascade via TIR domain-containing adaptor molecules (TRIF and MyD88) that leads to the activation of type I interferons and proinflammatory cytokines and finally to the induction of an antiviral state.

Upon recognition of the respective PAMPs, TLRs dimerize and recruit TIR domain-containing adaptor molecules comprising MyD88, TIRAP, TRIF or TRAM. This recruitment initiates a downstream signaling

cascade that leads to the secretion of type-I IFNs, proinflammatory cytokines, and chemokines (Kawai and Akira, 2010).

For example, intracellular TLRs transmit the initial signal to MyD88 that recruits IRAK-1 and IRAK-4. These kinases phosphorylate the E3 ubiquitin ligase TRAF6, which then activates NF-κB to initiate the expression of proinflammatory cytokines. In addition, IRAK-1 triggers the activation of IRF7, which induces the transcription and translation of type I IFNs (**Figure 1.1.**) (Tseng et al., 2010).

Activated TLR3 and TLR4 transduce the signal via a TRIF-dependent pathway, which is initiated through the recruitment of TRAF6 and RIP1. The interaction of TRAF6 with TRIF results in a signaling mechanism similar to the MyD88-dependent pathway. The interaction with RIP1 leads to K63-linked polyubiquitination of TRIF and to the recruitment of TBK1 and IKKε kinases that activate the transcription factor IRF3 and type I IFN expression (**Figure 1.1.**).

1.2.2. Nucleotide-binding, oligomerization domain-like receptors (NLRs)

NLRs comprise a large group of multidomain proteins that contain a C-terminal LRR domain, a central **n**ucleotide binding, **o**ligomerization **d**omain (NOD), and an N-terminal effector domain (Inohara and Nunez, 2001; Harton et al., 2002). NLRs can be grouped into subfamilies based on their effector domains (Ting et al., 2008).

NOD1 and NOD2 are the best-studied members of the NLRC (**N**OD-**l**ike **r**eceptor containing a **C**ARD [**c**aspase **a**ctivation and **r**ecruitment **d**omain]) family. These proteins are cytosolic receptors that recognize peptidoglycans of gram-positive and gram-negative bacteria. More recent studies revealed that NOD2 also recognizes viral ssRNA (Sabbah et al., 2009), demonstrating that these receptors trigger

innate immune responses upon sensing of a large variety of microbial pathogens.

Upon binding of peptidoglycans to NOD1 and NOD2, the proteins oligomerize and transduce a signal via RIP2/NEMO(IKKγ) to activate NF-κB. In contrast, viral ssRNA recognition by NOD2 is RIP2-independent and involves CARD-CARD interactions between the NLR and the **m**itochondrial **a**nti**v**iral **s**ignaling protein MAVS. This leads to the recruitment of numerous downstream factors and finally induces NF-κB and IRF signaling and the expression of proinflammatory cytokines and type I IFNs (**Figure 1.2.)** (Elinav et al., 2011).

In contrast to NOD1 and NOD2, the NLR family members NLRP1 and NLRP3 (**N**OD-**l**ike **r**eceptor containing a **P**YD domain) assembles into large multiprotein complexes called inflammasomes, which are responsible for the activation of the inflammatory caspase-1. Currently, three different inflammasome complexes have been linked with antiviral signaling: the NLRP3-inflammasome (Kanneganti et al., 2006; Allen et al., 2009; Ichinohe et al., 2009), the RIG-I inflammasome (Poeck et al., 2009), and the AIM2 inflammasome (Fernandes-Alnemriet al., 2009; Hornung et al., 2009).

Upon recognition of viral RNA NLRP3 associates with the **a**poptosis-associated **s**peck-like protein containing a **C**ARD (ASC), which provides a link between the receptor and the pro-form of caspase-1. This interaction occurs through homotypic PYD and CARD interactions and leads to the activation of caspase-1. Activated caspase-1 cleaves pro-IL-1β and pro-IL-18 to release the biologically active forms, which become secreted to induce inflammation and antiviral responses (**Figure 1.2.**) (Dinarello et al., 1996 and 1998).

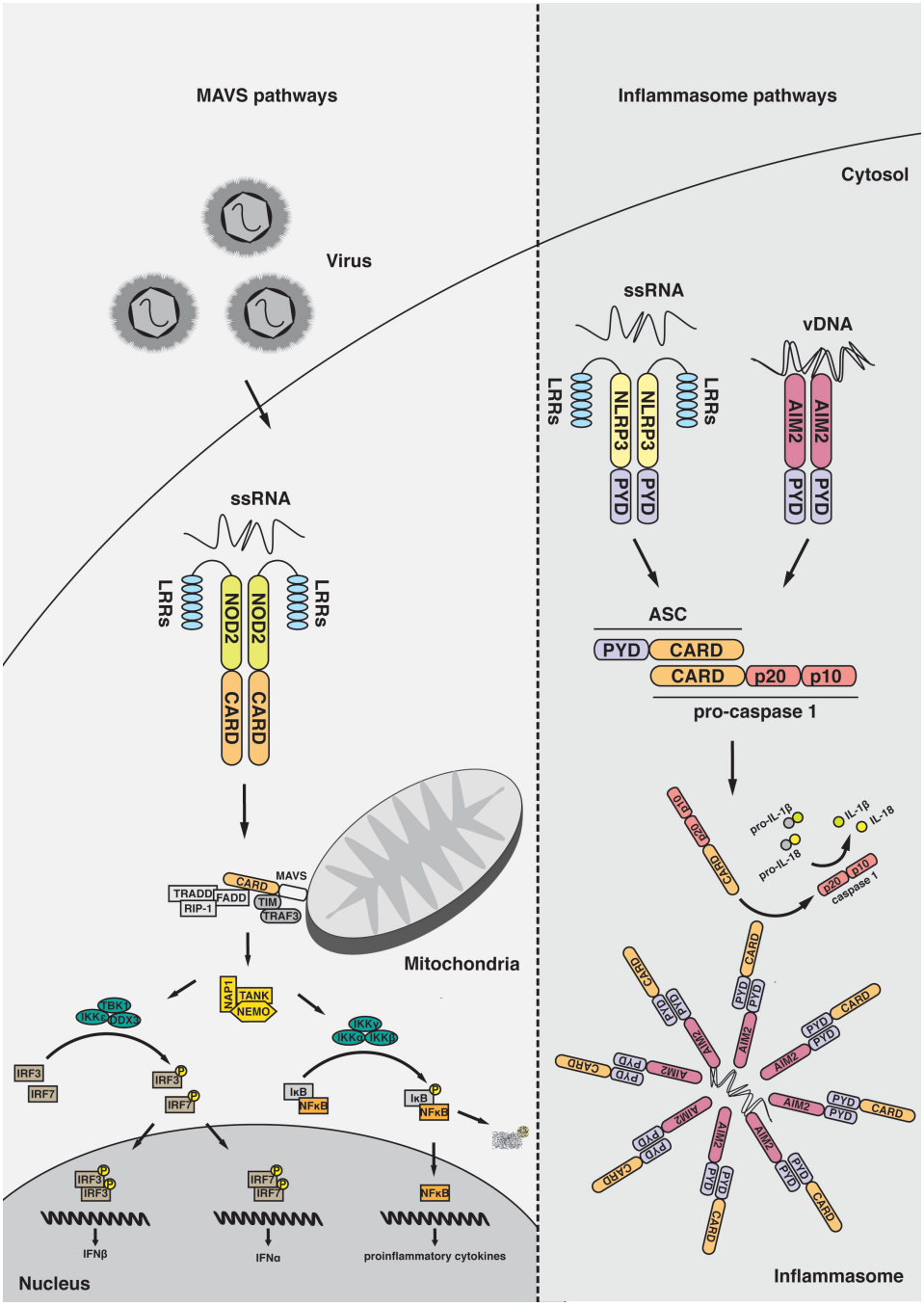


Figure 1.2. NLR signaling and inflammasome formation in the context of viral infection. *Left*: MAVS-dependent signaling. NOD2-mediated sensing of ssRNA leads to the activation of MAVS and stimulate the transcription of type I interferons and proinflammatory cytokines. *Right*: Inflammasome-dependent signaling. Activation of NLRP3 and AIM2 results in the assembly of inflammasomes that activate caspase-1. Upon binding to ssRNA (NLRP3) or vDNA (AIM2) the proteins assemble into inflammsome complexes via an ASC-dependent mechanism. This leads to the activation of caspase-1 and to the processing of pro-IL-1β and pro-L-18.

1.2.3. Recognition of cytoplasmic DNA: AIM2 and other DNA sensors

DNA that gains access to the cytosol is recognized by a variety of sensor molecules that initiate signaling responses to induce inflammatory molecules and immune factors. Endogenous DNA that is inappropriately cleared from the cytosol can induce chronic and pathological inflammation that result in autoimmune diseases like SLE (**s**ystemic **l**upus **e**rythematosus). In this section, the most common DNA sensors are outlined.

a) The AIM2 inflammasome

The PYHIN family protein AIM2 consists of an N-terminal PYD and a C-terminal HIN200 domain, which is responsible for DNA binding. Like NLRP1 and NLRP3, AIM2 activates caspases-1 via an ASC-dependent mechanism. The sensor protein is localized in the cytosol and recognizes cytosolic DNA of viral origin (Fernandes-Alnemriet al., 2009; Hornung et al., 2009). Upon DNA binding the AIM2 inflammasome assembles around the DNA, which provides the scaffold for the multiprotein complex formation (**Figure 1.2.**) (Jin et al., 2012). The AIM2 inflammasome is essential for caspase-1 activation, but completely dispensable for type I IFN production in response to cytosolic dsDNA. This emphasizes the different roles of AIM2-mediated signaling and type I IFN production, which is largely dependent on STING (Barber, 2011).

b) The cGAS-STING pathway

Although the AIM2 response is important in host defense, the predominant response to cytosolic DNA is the induction of type I IFNs. A central molecule in this signal transduction cascade is the adaptor molecule STING, which is localized at the endoplasmatic reticulum

(ER) (Ishikawa and Barber, 2008; Jin et al., 2008; Zhong et al., 2008). Recent discoveries indicate that the nucleotidyl transferase cGAS produces the second messenger molecule cGAMP upon binding to cytosolic dsDNA (Civril et al., 2013; Gao et al., 2013; Kranzusch et al., 2013; Wu et al., 2013). cGAMP is a cyclic molecule produced from ATP and GTP in the presence of DNA that binds to the adaptor molecule STING, facilitating its dimerization and activation and thereby triggering type I IFN expression upon TBK1 binding (Tanaka and Chen 2012). This recent finding provides new insight into the complex signaling pathways of DNA sensors.

c) TLR9 and the detection of CpG DNA

CpG DNA is a common feature of bacteria and viruses but it is absent in vertebrates (Hemmi et al., 2000). It is internalized via an endocytic pathway and traffics to lysosomal compartments were it associates with TLR9 (Sasai et al., 2010). Upon CpG DNA binding, TLR9 interacts with MyD88 triggering the activation of IRF7 and NF-κB transcription factors (Iwasaki and Medzhitov, 2010).

d) RNA Polymerase III detects cytosolic DNA and induces type I IFNs in a RIG-I-dependent pathway

Transfection studies have identified RNA polymerase III as a novel DNA sensor that triggers a type I IFN response through the RIG-I signaling pathway (Ablasser et al., 2009; Chiu et al., 2009). It was shown that transfected poly(dA/dT) dsDNA served as a template for RNA polymerase III, which was then transcribed into A/U-rich dsRNA containing 5'triphosphate moieties. Activation of RIG-I by this dsRNA induces a potent type I IFN response and activation of transcription factor NF-κB. The utilization of RNA polymerase III as a DNA sensor for bacteria and DNA viruses may therefore allow the host to take

advantage of the RIG-I signaling pathway to defend against a large variety of microbial pathogens. However, the physiological relevance of this pathway still needs to be clarified.

1.3. RIG-I-like receptor (RLR) proteins

1.3.1. RLRs are superfamily 2 DExD/H-box helicases

The **r**etinoic acid **i**nducible **g**ene I (RIG-I)-**l**ike **r**eceptor (RLR) proteins are key players in innate immunity and act by recognizing viral RNA (vRNA) in the cytosol. The RLR family consists of the members RIG-I (**r**etinoic acid **i**nducible **g**ene **I**), MDA5 (**m**elanoma **d**ifferentiation **a**ssociated protein **5**), and LGP2 (**l**aboratory of **g**enetics and **p**hysiology **2**) (Yoneyama et al., 2004; Kato et al., 2011; Loo and Gale Jr. 2011). RLRs are DExD/H-box helicase-like proteins sharing a central ATP-dependent helicase domain and a C-terminal regulatory domain (RD) that is responsible for initial RNA binding. In addition, RIG-I and MDA5 possess N-terminal tandem CARDs that are responsible for downstream signaling (**Figure 1.3.**) (Yoneyama et al., 2004; Fujita et al., 2007; Wang et al., 2010).

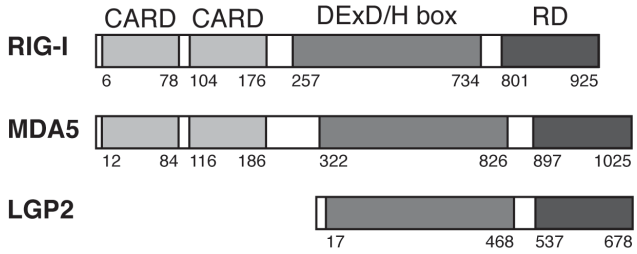


Figure 1.3. RLR structural architecture (Loo and Gale Jr., 2011).

Several crystal structures of RIG-I have shown that in the absence of virus, the protein exists in an auto-inhibited state, where the CARDs are sterically unavailable for signal transduction (Civril et al., 2011; Kowalinski et al., 2011; Ferrage et al., 2012). Upon viral infection and initial vRNA binding, the protein undergoes large conformational changes driven by ATP hydrolysis, leading to the exposure of CARDs to the cytosol. Lysine residues of the CARD2 domain are now available for polyubiquitination by the E3 ubiquitin ligase TRIM25 (Gack et al., 2007; Zeng et al., 2010), which is thought to trigger the interaction with the adaptor protein MAVS in the mitochondrial membrane.

In contrast to RIG-I, MDA5 does not exist in an auto-inhibited state but is rather thought to adopt an open conformation in the absence of vRNA (Berke and Modis, 2012). Upon dsRNA binding, the helicase domain of MDA5 forms a ring-like structure around the phosphate backbone of the ligand, which exhibits a more open conformation in comparison to RIG-I (Wu et al., 2013).

The RD domains between RLR proteins reveal differences that provide the molecular basis for recognition of distinct RNA structures. While the RIG-I RD comprises a pocket for the binding of 5' triphosphate containing blunt-ended dsRNA (Leung and Amarasinghe, 2012), the flat surface of the MDA5 RD rather allows stem loop recognition of dsRNA (Wu et al., 2013).

Structural and biophysical analyses support a model, in which MDA5 cooperatively forms ATP-sensitive filaments upon binding to long dsRNA molecules (Peisley et al., 2011; Berke et al., 2012; Berke and Modis, 2012; Wu et al., 2013). This is a remarkable difference in regard to RIG-I, which is thought to form monomeric or dimeric signaling complexes upon binding to blunt ended 5' triphosphate dsRNA (Cui et al., 2008; Kohlway et al., 2013). However, recent data suggest that RIG-I forms ATP hydrolysis-driven filaments that propagate from the

dsRNA blunt end to the interior of the molecule (Patel et al., 2013; Peisley et al., 2013). These recent findings require further analyses to investigate the exact mode of ATP hydrolysis by RLR proteins and its role in filament formation.

In contrast to RIG-I and MDA5, the role of LGP2 in cytosolic RNA sensing remains unclear. Some reports suggest that LGP2 is required for type I IFN production to some RIG-I- and MDA-dependent viruses, whereas others describe LGP2 as a negative regulator for RIG-I-dependent signaling (Bruns and Horvath, 2012).

1.3.2. RLR signal transduction via MAVS

Activated RIG-I and MDA5 transmit the initial signal to the downstream signaling pathway by binding to MAVS through homotypic CARD-CARD interactions (Takeuchi and Akira, 2010). In recent studies it has been suggested that the engagement of MAVS by RLRs, triggers the formation of large aggregates of the protein at the mitochondrial membrane (Hioscott et al., 2006; Ohmann et al., 2009; Dixit et al., 2010; Hou et al., 2011). The assembly of such "signalosomes" results in a large scale amplification of the signaling cascade and allows for highly sensitive detection of very small amounts of viral RNA that then induce the antiviral response (Zeng et al., 2010).

Formation of the MAVS "signalosome" triggers the recruitment of numerous downstream signaling factors (**Figure 1.4.**) including **t**umor **n**ecrosis **f**actor **r**eceptor (TNFR)-**a**ssociated **d**eath **d**omain (TRADD) protein, which exists in complex with **F**as-**a**ssociated **d**eath **d**omain-containig (FADD) protein and the death domain kinase RIP1. Signaling to this complex results in the recruitment and activation of TANK, NAP1 and NEMO to facilitate the activation of IRF3 and IRF7

by the kinases TBK1 and IKKε (Loo and Gale Jr., 2011). In addition, TANK and NAP1 recruit the kinases IKKα, β, and γ leading to the

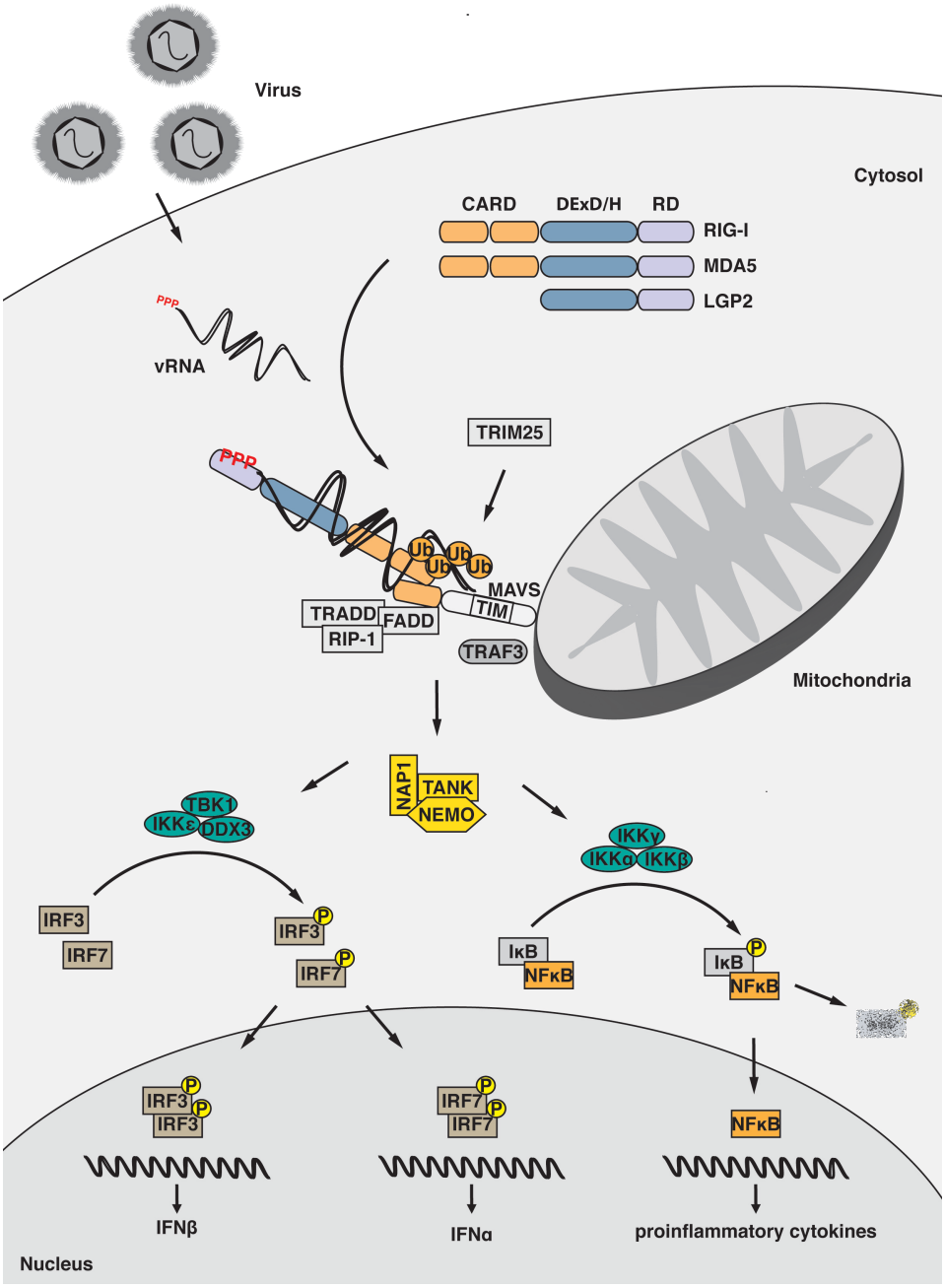


Figure 1.4. RLR signaling pathway. Upon binding to its respective RNA ligands, RIG-I and MDA5 become activated and transmit the signal via direct CARD-CARD interaction to MAVS, which is located in the mitochondrial membrane. Activation of MAVS triggers the formation of a "signalosome" complex by recruiting numerous downstream signaling factors including protein kinases and ubiquitin ligases. Finally, IRFs and NF-κB induces the expression of type I interferons and proinflammatory cytokines.

activation of NF-κB-dependent expression of proinflammatory cytokines (**Figure 1.4.**). Furthermore, the E3 ubiquitin ligase TRAF6, which binds to the **T**RAF-**i**nterracting **m**otif (TIM) in the proline-rich region of MAVS, facilitates the recruitment of IKKε kinase to the signalosome, thereby enhancing the IRF-3-dependent induction of type I IFN gene expression (Haecker et al., 2006; Oganesyan et al., 2006). Taken together, RLR signaling is regulated through multiple steps including posttranslational modifications like polyubiquitination and phosphorylation. This complexity in the signaling pathway may serve to fine-tune the RLR response to specific pathogenic stimuli.

1.3.3. Preferential roles of RLRs for virus recognition

In vitro studies revealed that most RNA viruses with a non-segmented negative-strand RNA genome, including measles virus (MeV), rabies virus (RABV), and Sendai virus (SeV) are preferentially recognized by RIG-I (Plumet et al., 2007; Loo et al., 2008). In contrast MDA5, preferentially senses picornaviruses with a positive-strand RNA genome (Loo et al., 2008). However, the observed preferences are unlikely to be exclusive. In case of MDA5, a minor contribution to recognition of MeV and SeV has been reported (Yount et al., 2008; Ikegame et al., 2010).

1.3.4. RNA structures recognized by RLRs

The viral specificity of RLRs emphasizes their remarkable capability to detect RNAs that are only present in infected cells. How RLRs actually discriminate between self and non-self RNA is not fully understood. Several *in vitro* studies report that RIG-I preferentially binds 5' triphosphate containing blunt-ended RNA with double-stranded

regions (Hornung et al., 2006; Pichlmair et al., 2006; Schlee et al., 2009; Schidt et al., 2009). Consistent with this, viral 5' triphosphate containing dsRNA was identified as the physiological RIG-I agonist in cells infected with influenza and SeV (Baum et al., 2010; Rehwinkel et al., 2010; Weber et al., 2013). Recently, it was shown that 5' triphosphate containing RNA is not the only ligand for RIG-I. Specific U/C-rich regions within certain viral genomes seem to contribute to efficient recognition by the protein (Saito et al., 2008; Schnell et al., 2012).

In contrast to RIG-I agonists, MDA5 ligands from virus-infected cells are not as well understood. Since picornaviruses are preferably sensed by MDA5 (Weber et al., 2006; Pichlmair et al., 2009), it was speculated that the protein could act as a simple dsRNA sensor in the cytosol. Consistent with this model it was shown that MDA5 becomes activated upon binding to poly (I/C) RNA, which is a synthetic analog of dsRNA (Takeuchi and Akira et al., 2010). Furthermore, it was suggested that MDA5 predominantly binds to long dsRNA stretches, while RIG-I is responsible for recognition of RNAs with short double-stranded regions (Kato et al., 2008). However, the precise nature of MDA5 agonists in infected cells might actually be more complex than dsRNA and could comprise structures or sequences that occur in the cytosol in the course of viral infection. For example, data on the enzyme RNase L suggest that small RNA cleavage products from the enzyme fold into RIG-I and MDA5 agonists containing 5' hydroxyl groups and 3' monophosphate moieties that efficiently activate the RLR signaling response (Malahti et al., 2007; Malathi et al., 2010; Luthra et al., 2011). In light of the divergent data, further work will be required to characterize naturally occurring RLR agonists.

1.4. Interplay between *Mononegavirales* and the innate immune system

Among viruses with RNA genomes, the order of negative-single-strand viruses (*Mononegavirales*) comprises many human pathogens causing severe diseases, like measles virus (MeV), rabies virus (RABV), Sendai virus (SeV), and Ebola virus (EBOV). These viruses replicate in the cytosol and use nucleoprotein complexes or nucleocapsids as replication templates instead of naked RNA (Gerlier and Lyles, 2011).

1.4.1. Virus replication cycle

Mononegavirales are enveloped viruses that share many common features and their genome organization is well conserved throughout this virus order. For example, the MeV genome consists of 15894 nucleotides encoding for 6 genes that are arranged in 3' to 5' direction (**Figure 1.5.**). The genes encode for a nucleoprotein (N), a phosphoprotein (P) and two non-structural proteins (V, C), a matrix protein (M), a fusion protein (F), a hemaglutinin protein (H), and a large RNA-dependent RNA polymerase (L). The P gene gives rise to 3 proteins (P, V, and C) through RNA editing (Lamb and Parks, 2007). The MeV genome is flanked by two small non-coding RNA sequences serving as promoter regions for replication and transcription: the leader (leRNA) and trailer (trRNA) regions.

Upon viral entry into the cell, the incoming P+L polymerase complex starts immediately to transcribe all genes from the genomic RNA of negative polarity (Lamb and Parks, 2007). The transcripts from every gene are 5' capped and polyadenylated by the polymerase. By arriving at an intergenic junction, the transcriptase either resumes RNA

synthesis at the next downstream gene or reinitiates upstream transcription, resulting in an mRNA gradient declining from 3' to 5' direction (**Figure 1.5. A**) (Cattaneo et al., 1987). New P+L transcriptases accumulate and enhance the transcription until enough N protein is synthesized. At this state, the virus switches from transcription to replication with a replicase complex consisting of N, P, and L proteins. The replication consists of an uninterrupted synthesis of the complementary antigenome, which is concurrently encapsidated by N proteins to form the so-called nucleocapsids (NCs). The new antigenomic NC serves as a template for the synthesis of genomic RNA, which is also immediately encapsidated into NCs. Consequently, the genome and antigenome do not exist naked in the infected cells and cannot anneal to each other or to viral transcripts (Gerlier and Lyles, 2011).

1.4.2. Viral RLR agonists from *Mononegavirales*: Towards a rational model.

Numerous *in vitro* studies focused on the nature of RLR agonists (Hornung et al., 2006; Pichlmair et al., 2006; Schlee et al., 2009; Schidt et al., 2009). It is now widely accepted in the field that the preferred RIG-I ligand is a 5' triphosphate containing dsRNA molecule. However it should be noted, that most of these *in vitro* data depend on transfection of RNAs extracted from virions or virally infected cells, which only provide limited or even misleading information about the physiological RIG-I agonist. In fact, these experiments were performed with extracted, denatured and purified RNA molecules comprising negative and positive RNA strands that could easily anneal to each other into dsRNA in the absence of N protein, which may never exist in an infected cell (Gerlier and Lyles, 2011). Therefore, the following

section will discuss the occurrence of viral RNA species in infected cells and their putative role as RLR agonists.

As already discussed in section 1.4.1., the genomic and antigenomic replication products from *Mononegavirales* do not exist as naked RNAs in the cytolsol, but are tightly packed into NC particles, in which the RNA is largely inaccessible. Biochemical experiments showed that RNA is resistant to nuclease degradation and RNA interference within these complexes (Blumberg et al., 1981; Bitko and Barik, 2001;Reuter et al., 2006; Mottet-Osmann et al., 2007). Therefore, the viral genome and antigenome should not be considered as putative RLR agonists.

However, the viral replication machinery is rather imprecise giving rise to abortive 5' genomes and antigenomes, which are not properly packed into NC particles (**Figure 1.5. B**). Furthermore, replication results in nucleocapsids with large internal deletions, called defective interfering (DI) particles, which may also fail to be correctly encapsidated. Upon self-hybridization into panhandle RNA structures, or upon genome-antigenome hybridization, these RNA species could make 5' triphosphate containg blunt-ended dsRNAs (Strähle et al., 2006). Indeed, a recent study on SeV-infected cells proposed a 546-nt-long copy-back DI particle encompassing the end of the L gene and trRNAs as a ligand for RIG-I (Baum et al., 2010). Furthermore, read-through L-trRNA mRNA could potentially hybridize to 5' triphosphate containing trRNA, which is then recognized by the receptor protein as an intermolecular dsRNA (**Figure 1.5. B**). It should be noted that in contrast to (+) RNA viruses (e.g. picornaviruses), (-) RNA viruses generate only short dsRNA species (Weber et al., 2006). This observation may provide a possible explanantion for the minor contribution of MDA5 to (-) RNA virus recognition (Yount et al., 2008; Ikegame et al., 2010).

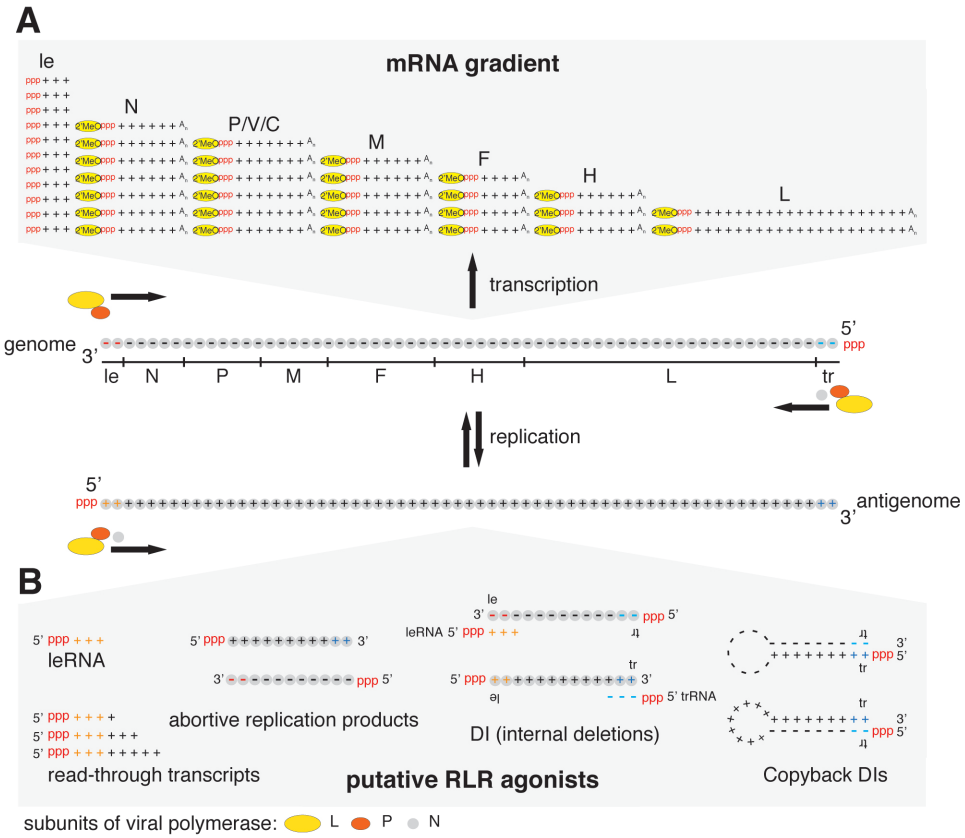


Figure 1.5. Viral RNA species and genomic organization of measles virus. **(A)** Viral mRNA gradient generated in the course of sequential transcription by the P+L transcriptase. In case of the antigenome only the trRNA is transcribed. With the exception of leRNA and trRNA, all transcripts are readily capped, ribose-O-methylated and polyadenylated. The replicase complex N+P+L carries out RNA replication. It consists of an uninterrupted synthesis of the genome and antigenome, which is readily packed into nucleocapsids by the N protein. **(B)** Potential RLR agonists due to transcriptase or replicase errors (adapted and modified from Gerlier and Lyles, 2011).

Beside trRNA, the leRNA transcripts are possible candidates for 5' triphosphate containing RIG-I ligands. However, it has been reported that these RNA molecules only exist in very limited amounts in infected cells, likely because of a lack in stability (Leppert et al., 1979; Vidal and Kolakofsky, 1989).

Finally, kinetic investigations on MeV-infected cells indicate that IFN-β expression strongly correlates with virus transcription (Plumet et al., 2007). However, fully transcribed mRNA molecules should not be considered as RLR agonists, since these species are subsequently capped, 5' O methylated, and polyadenylated by the viral polymerase

and exhibit exact the same characteristics as endogenous mRNA (**Figure 1.5. A**) (Li et al., 2009; Rahmeh et al., 2009). In this context it should be emphasized, that transcription also gives rise to non-functional byproducts. Read-through transcripts occur at frequencies of a few percent, resulting in bi- or tri-cristronic RNAs. Furthermore, 5' triphosphorylated leader-N and L-antitrailer RNAs are transcribed (Cattaneo et al., 1987) (**Figure 1.5. B**), that may explain the correlation of type I IFN production and virus replication (Plumet et al., 2007).

According, to the complex RNA composition in a virus-infected cell, the prediction of putative RNA ligands is challenging and further *in vivo* investigations are required to clarify, which RNAs are recognized by RLR proteins under physiological conditions.

1.5. Viral evasion of type I IFN signaling

The type I IFN system mediates a broad variety of antiviral effects. Consequently, successful viruses have been evolutionary selected to develop countermeasures to circumvent the IFN response. For example, the *Mononegavirales* can either mask the RNA structures that are recognized by RLRs or minimize their production (Vidal and Kolakofsky, 1983; Manuse and Parks, 2009). As discussed in section 1.4.2., the encapsidation of viral genome and antigenome into NCs efficiently prevent dsRNA formation and RLR recognition. Paramyxovirus C proteins and the RSV (Respiratory syncytial virus) NS1 protein negatively regulate the viral transcription rate, thereby limiting the production of putative RLR agonists (Cadd et al., 1996; Atreya et al., 1998; Baron and Barrett, 2000). The polymerase subunit VP35 of Ebola virus was shown to bind to dsRNA, which thereby

sequesters the molecule from the cytosol and from recognition by the immune receptors (Cardenas et al., 2006).

Furthermore, viruses encode for numerous proteins that directly target RNA sensors or their downstream signaling, thereby efficiently interfering with the antiviral immune response (Gerlier and Lyles, 2011; Taylor and Mossmann, 2012). Members of the paramyxovirus family (e.g. measles, parainfluenza, and Sendai virus) encode for V proteins that interfere with type I IFN expression by targeting MDA5, LGP2, and NF-κB (Childs et al., 2007; Schumann et al., 2011). Measles virus V protein was also shown to bind to JAK1 interfering with STAT1 phosphorylation (Caignard et al., 2007). By determining the structure of MDA5 in complex with parainfluenza virus V protein, the inhibitory mechanism has been previously elucidated. Upon binding to MDA5, the viral protein unfolds the ATPase domain via a β-hairpin motif that consequently leads to the disruption of the MDA5 ATP hydrolysis site and prevents RNA bound filament formation (Motz et al., 2013). The high efficiency of MDA5 inhibition by the V protein may explain the low contribution of MDA5 in mediating the activation of type I IFNs in response to paramyxoviruses (Gitlin et al., 2010; Ikegame et al., 2010).

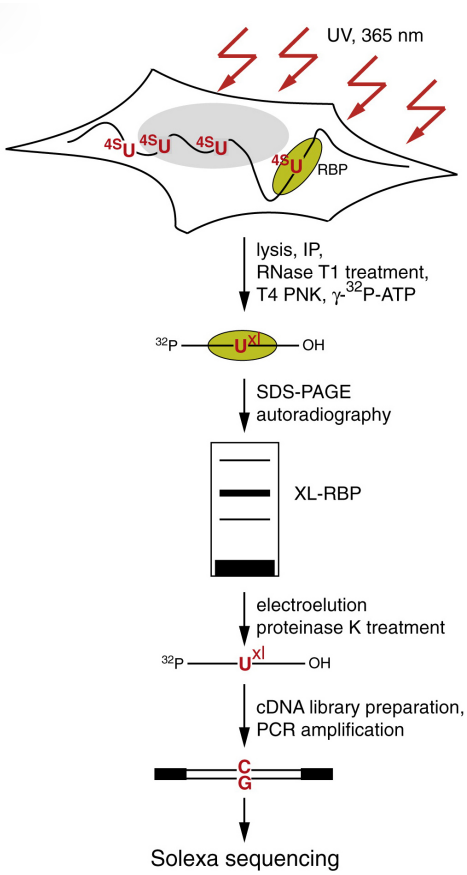
Besides the direct inhibition of RLRs, viral proteins interfere with the immune response by disrupting the downstream signaling cascade. C proteins from RPV (Rinder pest virus), MeV and SeV inhibit the activation of the IFN-β gene downstream of the phosphorylation and dimerization of IRF3, suggesting a nuclear mode of action (Shaffer et al., 2003; Strähle et al., 2007; Boxer et al., 2009; Sparrer et al., 2012). Rubaluvirus V protein, RABV P protein, and filovirus VP35 were also shown to inhibit IFN-β production by preventing IRF3 phosphorylation via TBK-1 and IKKε (Brzozka et al., 2005; Lu et al., 2008; Prins et al., 2009).

Efficient RIG-I inhibitors include the NS2 protein of RSV, which binds to the N-terminal CARD of RIG-I, thereby preventing the recruitment of MAVS by competition (Ling et al., 2009). The HIV (Human immunodeficiency virus) protease decreases cytoplasmic RIG-I levels by targeting the protein to the lysosome (Solis et al., 2011).

The list of additional viral evasion factors and their distinct mechanisms of interfering with the immune system is long. However, these few examples already give insights into the impressive diversity of tactics, viruses have developed to escape from the antiviral cellular system.

1.6. Objectives of the thesis

As discussed above, many studies focused on the identification of RNA agonists for RLR proteins (Hornung et al., 2006; Pichlmair et al., 2006; Weber et al., 2006; Saito et al., 2008; Schlee et al., 2009; Schmidt et al., 2009; Pichlmair et al., 2009; Baum et al., 2010; Rehwinkel et al., 2010; Schnell et al., 2012; Weber et al., 2013). However, these studies largely relied on *in vitro* experiments that may result in misleading or even false interpretation. To elucidate the exact nature and origin of viral RLR agonists it is therefore indispensable to isolate and investigate RLR-associated RNA from virus-infected cells. Recent studies on SeV-infected cells identified SeV DI particles as potent RIG-I inducers (Baum et al., 2010). However the physiological ligand for MDA5 and LGP2 still needs to be investigated. A major challenge in regard to examining MDA5 agonists is the apparently transient interaction between the protein and its RNA ligand, making common pull down approaches inadequate for studying MDA5-associated RNA.



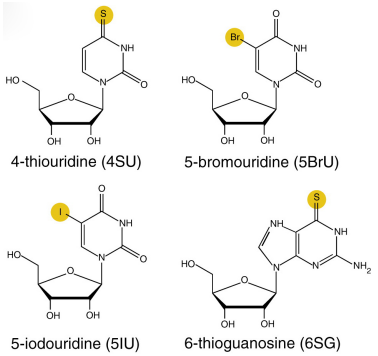


Figure 1.7. Examples of photoactivatable nucleoside analogs (Hafner et al., 2010)

Figure 1.6. PAR-CLIP methodology. Cells expressing the RNA-binding protein (RBP) of interest are treated with 4SU. 24 h later, cells are *in vivo* crosslinked with 365 nm UV. The cells are harvested and lyzed and the respective proteins are immunprecipitated. cDNA libraries can be generated from the isolated co-purified RNA and subjected to Next generation sequencing. (Hafner et al., 2010).

To improve the stability of interaction between MDA5 and its RNA ligand a crosslinking approach was established in the lab, adopted from the PAR-CLIP (**P**hoto**a**ctivatable-**R**ibonucleoside-Enhanced **C**ross**l**inking and **I**mmuno**p**recipitation) methodology (Hafner et al., 2010) (**Figure 1.6.**). In this approach, virus-infected cells are treated with the photoactivatable nucleoside analog 4-thiouridine (4SU) (**Figure 1.7.**), which is then incorporated into newly synthesized viral RNA. Upon 365 nm UV exposure of living cells, a covalent linkage between the 4SU-labeled RNA and the receptor protein is induced.

This leads to a higher RNA recovery from RLR immunoprecipitates and facilitates the analysis of putative RLR agonists.

To shed light on the nature of physiological RLR ligands in an unbiased manner, the crosslinking approach was combined with **N**ext **g**eneration **s**equencing (NGS) analysis. Additional biochemical experiments were performed to independently validate the obtained results.

2. Material & Methods

2.1. Material

2.1.1. Chemicals

All chemicals used in this thesis were purchased from Roth (Karlsruhe, Germany) and Sigma (Deisenhofen, Germany) unless otherwise stated. Mammalian cell culture media and reagents were purchased from Invitrogen, Life Technologies (Grand Island, NY, USA).

Reagents and siRNAs for RNA biology were obtained from Ambion, Life Technologies.

Enzymes and nucleotides for molecular biology were supplied by MBI Fermentas (St. Leon-Rot, Germany) or New England Biolabs (Ipswich, MA, USA).

For deep sequencing analysis reagents and kits were supplied by Illumina (San Diego, CA, USA) and Epicentre Biotechnologies (Madison, WI, USA).

Reagents for quantitative PCR analysis were purchased from Roche Diagnostics (Mannheim, Germany).

Primary antibodies to human MDA5 (AT113) and RIG-I (Alme-1) were purchased from Enzo Life Science (Loerrach, Germany). Antibody to GFP (ab1218) and monoclonal anti-β tubulin (TUB 2.1) were obtained from Abcam (Cambridge, UK) and Sigma, respectively. Antibodies to actin and measles P and N proteins were kindly provided by Prof. Dr. Karl-Klaus Conzelmann (LMU, Munich). Secondary antibodies were supplied by GE Healthcare (Buckinghamshire, UK).

2.1.2. Viral strains

Sendai virus Cantell (SeV-C) and Encephalomycarditis virus (EMCV) were generously provided by Prof. Dr. Adolfo Garcia-Sastrè (Mount Sinai Hospital, New York). Measles and rabies strains were obtained by Prof. Dr. Karl-Klaus Conzelmann. Mengo virus was provided by PD Dr. Anne Krug (TU Muenchen, Munich).

2.1.3. Mammamlian cell lines

In this thesis two human cell lines were used for the infection experiments. A549 human lung carcinoma cell lines were used for infection experiments and for the investigation of natural RNA ligands for RLR proteins. The cells were kindly provided by Prof. Dr. Karl-Klaus Conzelmann.

For immunostimulatory experiments human embryonic kidney (HEK) 293T ISRE FF/RFP cell lines were used. These cells stably integrated an interferon-stimulated response element (ISRE) in the genome, controlling the expression of the reporter genes firefly luciferase and RFP. The cells are a gift from Prof. Dr. Adolfo Garcia-Sastrè.

2.1.4. Mammalian cell culture

Cell lines were cultured in complete growth media and grown in a humidified, 37 ºC, 5% CO_2 incubator. When cells reached a confluency of approx. 80%, the medium was removed and cells were washed with PBS. Cells were detached by trypsin (EDTA) solution (Invitrogen) and diluted to the desired concentration in complete growth media.

Buffer	Composition
Complete growth medium	DMEM, 10% FBS, 2 mM L-glutamine, 1% Penicilin/Streptavidin solution
PBS	10 mM phosphate, 137 mM NaCl, 2.7 mM KCl, pH 7.5

Table 2.1. Media and solutions for mammalian cell culture.

2.1.5. Oligoribonucleotides

The following tables outline the primers and oligonucleotides used in this thesis. For quantitative PCR analysis primers were used as previously described (Plumet and Gerlier, 2005).

Genomic region	Primer	Sequence Sequence (5' → 3')
Le-N	forward	ACCAAACAAAGTTGAGTAAG
	reverse	GATGTAATGGGTGGTGGTTTGTC
N	forward	AGTGAGAATGAGCTACCG
	reverse	TGTCTAGGGGTGTGCC
P	forward	GAACTCAATCCCGACCTGAA
	reverse	GCTGTCCTCTGGAACTGGTC
M	forward	AACGCAAACCAAGTGT
	reverse	TGAAGGCCACTGCATT
F	forward	TGCACTTCACCAGTCC
	reverse	AGCTTCTGGCCGATTA
H	forward	ACATACCTACCTGCGG
	reverse	AGTAAGAAAATGAGCGGC
L	forward	GTGTGAAAGCGACGAG
	reverse	TGTTCCACGAAGATCCT
L-tr	forward	GAGAAACAGATTATTATGACGGG
	reverse	TGTTCCACGAAGATCCT

Table 2.2. Primers used for quantitative PCR.

For synthesis of *in vitro* transcripts following primers were used (**Table 2.3.**). The *in vitro* transcripts were generated by Konstantin Sparrer (Prof. Dr. Karl-Klaus Conzelmann).

In vitro transcript	Primer	Sequence (5' → 3')
1	forward	GCGTAATACGACTCACTATAGGG
		ACCAAACAAAGTTGGGTAA
	reverse	GATTCCTCTGATGGCTCCAC
2	forward	GCGTAATACGACTCACTATAGGG
		CAAACACATTATTATAGTACC
	reverse	TAACGTCAGGGTCATCGGTG A
3	forward	GCGTAATACGACTCACTATAGGG
		CTTAGGAACCAGGTCCACAC
	reverse	CTCCTGTCCTGGGTTGTCTGA
4	forward	GCGTAATACGACTCACTATAGGG
		CAGGTGCACCTGCGGGGAAT
	reverse	CTTCTGATTATCCTCGTGTAT
5	forward	GCGTAATACGACTCACTATAGGG
		AACAGCCCTGACACAAGGCC
	reverse	CTTGTGCGGT TCGGTTGTGG A
6	forward	GCGTAATACGACTCACTATAGGG
		TGTCCATCATGGGTCTCAAGG
	reverse	GAGAGTTATATTGGGCATTAA
7	forward	GCGTAATACGACTCACTATAGGG
		GAACCTAGCCTTAGGTGTAAT
	reverse	CCAAAACATATTGGAGATCTT
8	forward	GCGTAATACGACTCACTATAGGG
		GAGACACACACCTGTATTCTT
	reverse	CTAGAAGCTCTGTATACCTAG
9	forward	GCGTAATACGACTCACTATAGGG
		GAAGAGTCAGATACATGTGG
	reverse	AGAAAACCCGTTTTGGTCAAG

In vitro transcript	Primer	Sequence (5' → 3')
10	forward	GCGTAATACGACTCACTATAGGG
		GATGAAGGTACTTATCATGAG
	reverse	TGACCTTTCATCAGAGTCTCA
11	forward	GCGTAATACGACTCACTATAGGG
		GCCAAGTGATTGCTGAAAATC
	reverse	GGTACTTGTGTGGACTGGGCT
12	forward	GCGTAATACGACTCACTATAGGG
		ACAATTGTTTCATCACATTTTT
	reverse	TATGCAAGGTAACGGTCATAA
13	forward	GCGTAATACGACTCACTATAGGG
		AGAAATGTCCTCATTGACAAA
	reverse	CATCCGTAGTTGACTGATCCA
14	forward	GCGTAATACGACTCACTATAGGG
		GAATCGGGTTGAACTCATCTG
	reverse	CCTTTGACCAGATCTAGAATT
15	forward	GCGTAATACGACTCACTATAGGG
		AGTAGGTAATATTGTCAAAGT
	reverse	GCAGAGCCATCGATAAGATGG
16	forward	GCGTAATACGACTCACTATAGGG
		ATTTGGTTATGACAGATCTCA
	reverse	TGTAAGTTTTTTCAGAGTAGG
17	forward	GCGTAATACGACTCACTATAGGG
		TTTGAAACGTGAGTGGGTTTTT
	reverse	ACCAGACAAAGCTGGGAATAG

Table 2.3. Primers used for PCR of in vitro transcripts.

Based on the primers listed in table 2.3, the following *in vitro* transcripts were generated by PCR and T7 *in vitro* transcription.

<u>*In vitro* transcript #1</u> (le-N, 1-200)

ACCAAACAAAGTTGGGTAAGGATAGTTCAATCAATGATCATCTTCTAGTGCACTTAGGATTCAAG
ATCCTATTATCAGGGACAAGAGCAGGATTAGGGATATCCGAGATGGCCACACTTTTAAGGAGCTT

AGCATTGTTCAAAAGAAACAAGGACAAACCACCCATTACATCAGGATCCGGTGGAGCCATCAGAG
GAATC

In vitro transcript #2 (N, 200-400)

CAAACACATTATTATAGTACCAATCCCTGGAGATTCCTCAATTACCACTCGATCCAGACTTCTGGA
CCGGTTGGTGAGGTTAATTGGAAACCCGGATGTGAGCGGGCCCAAACTAACAGGGGCACTAATAG
GTATATTATCCTTATTTGTGGAGTCTCCAGGTCAATTGATTCAGAGGATCACCGATGACCCTGACG
TTA

In vitro transcript #3 (P, 1745-1945)

CTTAGGAACCAGGTCCACACAGCCGCCAGCCCATCAACCATCCACTCCCACGATTGGAGCCAATGG
CAGAAGAGCAGGCACGCCATGTCAAAAACGGACTGGAATGCATCCGGGCTCTCAAGGCCGAGCCCA
TCGGCTCACTGGCCATCGAGGAAGCTATGGCAGCATGGTCAGAAATATCAGACAACCCAGGACAGG
AG

In vitro transcript #4 (P, 2600-2800)

CAGGTGCACCTGCGGGGAATGTCCCCGAGTGTGTGAGCAATGCCGCACTGATACAGGAGTGGACAC
CCGAATCTGGTACCACAATCTCCCCGAGATCCCAGAATAATGAAGAAGGGGGAGACTATTATGAT
GATGAGCTGTTCTCTGATGTCCAAGATATTAAAACAGCCTTGGCCAAAATACACGAGGATAATCA
GAAG

In vitro transcript #5 (M, 4600-4800)

AACAGCCCTGACACAAGGCCACCACCAGCCACCCCAATCTGCATCCTCCTCGTGGGACCCCCGAGGA
CCAACCCCCAAGGCTGCCCCCGATCCAAACCACCAACCGCATCCCCACCACCCCCGGGAAAGAAACC
CCCAGCAATTGGAAGGCCCCTCCCCCTCTTCCTCAACACAAGAACTCCACAACCGAACCGCACAAG

In vitro transcript #6 (F, 5450-5650)

TGTCCATCATGGGTCTCAAGGTGAACGTCTCTGCCATATTCATGGCAGTACTGTTAACTCTCCAAA
CACCCACCGGTCAAATCCATTGGGGCAATCTCTCTAAGATAGGGGTGGTAGGAATAGGAAGTGCA
AGCTACAAAGTTATGACTCGTTCCAGCCATCAATCATTAGTCATAAAATTAATGCCCAATATAACT
CTC

In vitro transcript #7 (H, 8650-8850)

GAACCTAGCCTTAGGTGTAATCAACACATTGGAGTGGATACCGAGATTCAAGGTTAGTCCCTACCT
CTTCACTGTCCCAATTAAGGAAGCAGGCGAAGACTGCCATGCCCCAACATACCTACCTGCGGAGGT

GGATGGTGATGTCAAACTCAGTTCCAATCTGGTGATTCTACCTGGTCAAGATCTCCAATATGTTTT
GG

In vitro transcript #8 (L, 9800-10000)

GAGACACACACCTGTATTCTTCACTGGTAGTTCAGTTGAGTTGCTAATCTCTCGTGACCTTGTTGC
TATAATCAGTAAAGAGTCTCAACATGTATATTACCTGACATTTGAACTGGTTTTGATGTATTGTG
ATGTCATAGAGGGGAGGTTAATGACAGAGACCGCTATGACTATTGATGCTAGGTATACAGAGCTT
CTAG

In vitro transcript #9 (L, 10000-10200)

GAAGAGTCAGATACATGTGGAAACTGATAGATGGTTTCTTCCCTGCACTCGGGAATCCAACTTATC
AAATTGTAGCCATGCTGGAGCCTCTTTCACTTGCTTACCTGCAGCTGAGGGATATAACAGTAGAAC
TCAGAGGTGCTTTCCTTAACCACTGCTTTACTGAAATACATGATGTTCTTGACCAAAACGGGTTTT
CT

In vitro transcript #10 (L, 10200-10400)

GATGAAGGTACTTATCATGAGTTAACTGAAGCTCTAGATTACATTTTCATAACTGATGACATACA
TCTGACAGGGGAGATTTTCTCATTTTTCAGAAGTTTCGGCCACCCCAGACTTGAAGCAGTAACGGC
TGCTGAAAATGTTAGGAAATACATGAATCAGCCTAAAGTCATTGTGTATGAGACTCTGATGAAAG
GTCA

In vitro transcript #11 (L, 10900-11100)

GCCAAGTGATTGCTGAAAATCTAATCTCAAACGGGATTGGCAAATATTTTAAGGACAATGGGATG
GCCAAGGATGAGCACGATTTGACTAAGGCACTCCACACTCTAGCTGTCTCAGGAGTCCCCAAAGAT
CTCAAAGAAAGTCACAGGGGGGGGCCAGTCTTAAAAACCTACTCCCGAAGCCCAGTCCACACAAGT
ACC

In vitro transcript #12 (L, 11700-11900)

ACAATTGTTTCATCACATTTTTTTGTCTATTCAAAAGGAATATATTATGATGGGCTACTTGTGTCC
CAATCACTCAAGAGCATCGCAAGATGTGTATTCTGGTCAGAGACTATAGTTGATGAAACAAGGGC
AGCATGCAGTAATATTGCTACAACAATGGCTAAAAGCATCGAGAGAGGTTATGACCGTTACCTTG
CATA

In vitro transcript #13 (L, 12600-12800)

AGAAATGTCCTCATTGACAAAGAGTCATGTTCAGTGCAGCTGGCGAGAGCTCTAAGAAGCCATAT
GTGGGCGAGGCTAGCTCGAGGACGGCCTATTTACGGCCTTGAGGTCCCTGATGTACTAGAATCTAT
GCGAGGCCACCTTATTCGGCGTCATGAGACATGTGTCATCTGCGAGTGTGGATCAGTCAACTACGG
ATG

In vitro transcript #14 (L, 14500-14700)

GAATCGGGTTGAACTCATCTGCTTGCTACAAAGCTGTTGAGATATCAACATTAATTAGGAGATGC
CTTGAGCCAGGGGAGGACGGCTTGTTCTTGGGTGAGGGATCGGGTTCTATGTTGATCACTTATAA
GGAGATACTTAAACTAAACAAGTGCTTCTATAATAGTGGGGTTTCCGCCAATTCTAGATCTGGTC
AAAGG

In vitro transcript #15 (L, 14750-14950)

AGTAGGTAATATTGTCAAAGTGCTCTTTAACGGGAGGCCCGAAGTCACGTGGGTAGGCAGTGTAG
ATTGCTTCAATTTCATAGTTAGTAATATCCCTACCTCTAGTGTGGGGTTTATCCATTCAGATATAG
AGACCTTGCCTGACAAAGATACTATAGAGAAGCTAGAGGAATTGGCAGCCATCTTATCGATGGCT
CTGC

In vitro transcript #16 (L, 15100-15300)

ATTTGGTTATGACAGATCTCAAGGCTAACCGGCTAATGAATCCTGAAAAGATTAAGCAGCAGATA
ATTGAATCATCTGTGAGGACTTCACCTGGACTTATAGGTCACATCCTATCCATTAAGCAACTAAGC
TGCATACAAGCAATTGTGGGAGACGCAGTTAGTAGAGGTGATATCAATCCTACTCTGAAAAAACT
TACA

In vitro transcript #17 (L-tr, 15695-15894)

TTTGAAACGTGAGTGGGTTTTTAAGGTAACAGTCAAGGAGACCAAAGAATGGTATAAGTTAGTCG
GATACAGTGCCCTGATTAAGGACTAATTGGTTGAACTCCGGAACCCTAATCCTGCCCTAGGTGGTT
AGGCATTATTTGCAATATATTAAAGAAAACTTTGAAAATACGAAGTTTCTATTCCCAGCTTTGTC
TGGT

2.2. Microbiological methods

2.2.1. Polyacrylamide gel electrophresis

a) Denaturing polyacrylamide gel electrophresis for protein analysis

For separation of proteins by **s**odium **d**odecyl **s**ulfate **p**oly**a**crylamide **g**el **e**lectrophoresis (SDS-PAGE), a 15% separation and a 5% stacking gel was used. For denaturation, protein samples were diluted with SDS loading dye and heated up to 95 ºC for 5 min. The samples were separated on the gel by electrophoresis with SDS-PAGE running buffer. After separation the gel was either stained by Coomassie or separated proteins were transferred to a membrane by Western blot.

Buffer	Composition
4x Separation buffer	3 M Tris, pH 8.5, 0.4% SDS
4x Stacking buffer	0.5 M Tris, pH 6.8; 0.4% SDS
15% separation gel	2.5 mL H_2O; 2.5 mL separation buffer; 5 mL acrylamide; 50 µL 10% APS; 5 µL TEMED
5% stacking gel	2.5 mL H_2O; 1 ml stacking buffer; 500 µL acrylamide; 40 µL 10% APS; 4.0 µL TEMED
SDS loading dye	110 mM Tris, pH 6.8, 15% glycerin, 4% SDS, 5% β-mercaptoethanol

Table 2.4. Buffers and solutions used for SDS-PAGE.

b) Native polyacrylamide gel electrophoresis for RNA analysis

For separation of RNA a 15% polyacrylamide gel was used. The gel was pre-run at 200 V for 30 min in Tris-glycine buffer and afterwards RNA samples were separated by electrophoresis (200 V, 2h).

Buffer	Composition
Tris-glycine buffer (5x)	125 mM Tris; 1 M glycine
RNA loading dye (4x)	40% glycerol in Tris-glycine buffer; 0.025% bromophenol blue
Native polyacrylamide gel (15%)	3.75 mL acrylamide; 2 mL Tris-glycine buffer (1x); 4.25 mL H_2O; 37.5 µL 10% APS; 7.5 µL TEMED

Table 2.5. Buffers and solutions used for native PAGE.

2.2.2. Western blot assay

Proteins were transferred from the gel to a nitrocellulose membrane using a semi-dry blotting system (Bio-Rad) and Western blot transfer buffer. The blotting was performed for 1 h with 2 mA/cm^2 according to the membrane size. The membrane was incubated with TBS-T containing 5% milk powder (Roth) for 30 min at room temperature (RT), washed 3 times with TBS-T, and incubated overnight at 4ºC in TBS-T, containing 5% milk powder and the primary antibody (1 to 5 µg/mL). The membrane was again washed 3 times with TBS-T and subsequently incubated with 1 µg/mL of secondary antibody in TBS-T for 1 h at 4 ºC. The membrane was washed 3 times, incubated and developed with the ECL detection kit (Thermo Scientific) according to manufacturer's instructions. The signal was visualized using a chemiluminescence film (Amersham Bioscience).

Buffer	Composition
TBS-T	10 mM Tris, pH 8.0; 150 mM NaCl; 0.01% Tween 20
Wetsern blot transfer buffer	25 mM Tris, pH 8.5; 192 mM glycione; 0.05% SDS; 20% methanol

Table 2.6. Buffers and solutions used for Western blot.

2.2.3. Immunoprecipitation

Protein G Dynabeads (Invitrogen) were resuspended and washed 2 times with citrate-phosphate buffer (CPB). Washed beads were resuspended in 2 bead volumes CPB containing 1 µg/mL of the appropriate antibody. The beads were incubated for 1 h at 4 ºC to allow binding of the antibody. After incubation the beads were washed 2 times with CPB and resuspended in 2 bead volumes. Finally the beads were incubated with cell lysate for 1 to 4 h at 4 ºC constantly shaking. Afterwards, beads were washed 3 times with low-salt washing buffer. Depending on the following analysis, the beads were resuspended in SDS loading dye for SDS PAGE analysis, or protein-associated RNA was extracted.

Buffer	Composition
Citrate-phosphate buffer	50 mM Sodium hydrogenphosphate dibasic; 25 mM citric acid
Low salt washing buffer	50 mM HEPES; 300 mM KCl; 0.05% % NP-40; 0.5 mM DTT; protease inhibitor

Table 2.7. Buffers and Solutions used for immunoprecipitation.

2.2.4. RNA purification

a) RNA extraction

To isolate RNA from RNP complexes, the beads were resuspended in proteinase K reaction buffer containing 10 U/µL proteinase K (Fermentas). The reaction was incubated for 30 min at 55 ºC and 800 rpm. To extract RNA, the solution was mixed with 1 volume of acidic phenol (pH 4.2; Sigma) and centrifuged for 1 min at 20.000 x g. The aqueous phase was transferred into a new tube and mixed with 1

volume of a chloroform/isoamylalcohol solution (24:1 (v/v); Sigma). After centrifugation the aqueous phase was again transferred into a new tube and RNA was ethanol-precipitated as described below.

Buffer	Composition
Proteinase K reaction buffer (2x)	100 mM Tris, pH 7.5; 150 mM NaCl; 12.5 mM EDTA; 2% SDS

Table 2.8. Buffers and solutions used for RNA extraction

b) RNA precipitation

The aqueous phase containing the RNA was mixed with 4 volumes of 100% ethanol. The solution was mixed with lithium chloride (100 mM final concentration), sodium actetate (220 mM final concentration) and 2 µL glycogen (Ambion). RNA was precipitated overnight at -20 ºC and pelleted by centrifugation. The RNA was washed with 80% ethanol and resuspended in the desired volume of RNase-free water.

2.2.5. RNA analysis

a) 5' radiolabeling of RNA with T4 PNK

RNA was radiolabeled in the presence of 2 µCi/µL [γ-P^{32}] ATP (Hartmann Analytic) with 0.4 U/µL T4 polynucleotide kinase (PNK) (Fermentas). The reaction was incubated for 15 min at 37 ºC. For saturation the reaction was mixed with 100 µM non-radioactive ATP and incubation was continued for additional 5 min. The reaction was terminated by addition of an equal volume of stop mix and incubation at 95 ºC for 2 min. The samples were separated on a 15% denaturing acrylamide gel. The gel was exposed to a phosphorimaging screen for 1 h and radioactivity was quantified by a STORM 860 imaging system (Molecular Dynamics).

Buffer	Composition
PNK buffer	50 mM Tris, pH 7.5; 50 mM NaCl; 10 mM $MgCl_2$; 5mM DTT
Stop mix	15 mg bromophenol blue/xylene cyanol 150 µL 0.5 M EDTA, 600 µL H_2O, 14.25 mL deionized formamide, pH 8.0

Table 2.9. Buffers and solutions used for RNA 5' radiolabeling.

b) Luciferase transfection assay

Immunoactivity experiments were carried out in 24-well plates**.** 2.5 x 10^5 HEK 293T ISRE FF/RFP cells were transfected with up to 250 ng of recovered RNA or 500 ng *in vitro* transcript using Lipofectamine 2000 (Invitrogen) according to manufacturer's protocol. After 24 h incubation, cells were subjected to immunoactivity experiments using the *Dual-Glo luciferase* assay system (Promega) according to manufacturer's instructions. The luciferase activity was validated in a 96-well plate reader.

c) Separation of 4SU-labeled RNA from unlabeled RNA species

RNA (~1 µg) from 4SU-treated cells was diluted in biotinylation buffer containing 0.2 µg/µL biotin-HDPD (Thermo Scientific). The reaction mix was incubated for 2 h at RT and 800 rpm in a thermomixer. The RNA was purified as described in section 2.2.4. For separation of labeled and unlabeled RNA, the samples were incubated for 10 min at 65 ºC and subsequently cooled down to 4º C on ice. The RNA (~100 µg) was loaded to 100 µL of µMacs streptavidin beads (Miltenyi Biotec) and the reaction was incubated with slight shaking for 15 min at RT. The mixture was applied to µMacs columns (Miltenyi Biotec) and beads were separated from the supernatant by a magnetic separator.

The flow-through containing unlabeled RNA was collected in 600 µL isopropanol and centrifuged at 20.000 x g for 10 min at 4 ºC. The supernatant was removed and RNA was precipitated as described above. The RNA pellet was diluted in the desired volume of RNase-free water. The µMacs streptavidin beads were washed one time with 65 ºC washing buffer. The 1st wash, still containing unlabeled RNA, was treated as the flow-through. The bound RNA was washed 5 times with cold washing buffer and eluted with 100 µL elution buffer. The elution step was repeated one time. The labeled RNA was purified according to the RNeasy MinElute Cleanup protocol (Qiagen).

Buffer	Composition
Biotinylation buffer (10x)	100 mM Tris, pH 7.5; 10 mM EDTA
Washing buffer	100 mM Tris, pH 7.5; 10mM EDTA; 1M NaCl; 0.1% Tween 20
Elution buffer	100 mM DTT in RNase-free H_2O

Table 2.10. Buffers and solutions used for RNA labeling.

2.2.6. Fluorescence microscopy

Approximately 5×10^4 A549 or 293T ISRE FF/RFP cells were grown on glass slides in 24-well microtiter plates. Cells were infected with virus as indicated and 24 hours post infection (hpi), infection was visualized using a Zeiss LSM510 microscopy system. The images were edited with Zeiss LSM 5 Image Examiner.

2.2.7. T7 RNA transcription

Templates were generated for *in vitro* transcription in a PCR reaction adding the T7 promoter sequence (TAATACGACTCACTATA GGG) to

the 5' end of the desired MeV genome fragment. The used oligonucleotides in the reaction are outlined in table 2.3. The PCR products were subsequently purified on agarose gels. RNA was transcribed from these templates using the Ambion MEGAshortscript T7 Kit according to the manufacturer's protocols. The reaction was incubated over night at 37°C, followed by RNA precipitation using LiCl at -20°C for at least 30 minutes. The purified RNA was then subjected to triphosphate digestion and purified on denaturating polyacrylamide gels at 25 mA constant current. Gel slices containing RNA were incubated overnight with 450 µL probe elution buffer. Eluted RNA was isolated by Phenol/Chloroform/Isoamylalkohol extraction and precipitated with Ethanol (see section 2.2.4).

Buffer	Composition
Probe elution buffer	0.5 M NH_4; 1 mM EDTA; 0.2% SDS

Table 2.11. Buffers and solutions used for T7 RNA transcription.

2.2.8. ATPase hydrolysis assays

ATPase hydrolysis activity was determined using [γ-P^{32}] ATP. Mouse MDA5 was purified as described previously (Motz et al., 2013) and 2 mM of protein was preincubated with 80 pM *in vitro* transcribed RNA for 10 min at RT. The reaction was initiated by addition of ATPase hydrolysis buffer containing 2 mM ATP and 0.2 µCi [γ-P^{32}] ATP. The hydrolysis rate was monitored over 1h and analyzed by thin layer chromatography (TLC). The membrane was exposed to a phosphorimaging screen for 1 h and quantified as described in section 2.2.5.

Buffer	Composition
ATPase hydrolysis buffer	20 mM HEPES, pH 7.5; 150 mM NaCl; 1.5 mM $MgCl_2$, and 2 mM DTT

Table 2.12. Buffers and solutions used for ATP hydrolysis assays.

2.2.9. SYBR green quantitative PCR analysis

Quantitative PCR analysis was performed with primer pairs previously described (Plumet and Gerlier, 2004) using a LightCycler carousel based system (Roche). Total RNA from immunoprecipitations was subjected to DNase I treatment in order to remove DNA contaminations. To allow transcription of total RNA including mRNA species and small RNAs, a random hexanucleotide mix (Roche) was used for cDNA synthesis. The reaction was incubated with reverse transcriptase (Roche) according to manufacturer's protocol. PCR products were quantified by utilization of SYBR green in relation to a PCR standard of known concentration. The relative abundance of RNA was calculated with reference to the GFP control sample.

2.3. PAR-CLIP approach

2.3.1. Crosslinking and Immunoprecipitation of RLR-associated RNA from virus-infected cells

A549 human lung carcinoma cells were infected with virus with an MOI (multiplicity of infection) of 1.0 in the presence of 400 µM 4SU. Infection was allowed to occur for 24 h and living cells were washed with PBS and exposed to 1 J/cm^2 365 nm UV light using a photocrosslinker (Vilbert Lourmat). Cells were harvested and incubated in Nonidet P-40 lysis buffer for 10 min on ice. The lysate was cleared by centrifugation and endogenous proteins were immunoprecipitated for 4 h with the respective antibodies (1 µg/mL) bound to protein G Dynabeads. The beads were washed five times with high-salt wash buffer and incubated with Proteinase K for 30 min at 55 ºC. The RNA was isolated as described in section 2.2.4.

Buffer	Composition
NP-40 lysis buffer	50 mM HEPES, pH 7.5; 150 mM KCl; 1mM NaF; 10 µM $ZnCl_2$; 0.5% NP-40; 0.5 mM DTT; protease inhibitor
High-salt wash buffer	50 mM HEPES, pH 7.5; 500 mM KCl; 0.05% NP-40; 0.5 mM DTT; protease inhibitor

Table 2.13. Buffers and solutions used for the PAR-CLIP approach.

2.3.2. cDNA library preparation and deep sequencing analysis

Total RNA isolated from immunoprecipitates was prepared for Illumina sequencing using the mRNA-Seq library preparation kit

(Epicentre) according to manufacturer's protocol. To remove ribosomal RNA species from the sequencing libraries a Ribo-Zero rRNA removal kit (Epicentre) was used. Quality of RNA-Seq libraries was validated on an Agilent bioanalyzer DNA chip. Sequencing was performed on the Illumina Genome Analyzer in the Gene Center sequencing facility (LAFUGA). Obtained sequences were processed with the FASTX toolkit in order to remove adapter sequences and reads with PHRED scores below 30. Remaining sequences were mapped to human and viral genomes by utilization of the Bowtie algorithm, allowing maximal one mismatch per unique read. The Bowtie sequence alignments were converted with SAMtools to pileup format, which was subsequently used for further data analysis. The relative sequence abundance was analyzed between RLR pull down samples and the GFP control. The results were visualized in R.

2.3.4. Generation of Copy number variation data

Copy number variations were calculated by determining the relative sequence abundance at each position on the genomic segment and calculating the average of the RLR/GFP ratios over a dynamic window of 200 reads. Relative sequence abundances with Log2 ratios above +1 were defined as significantly enriched in the RLR library.

2.4. Bioinformatical analysis of NGS data

The bioinformatical analysis was done by Katharina Hembach (Prof. Dr. Johannes Soeding, LMU). All bioinformatical scripts, designed for the analyses, were published in the Bachelor thesis "Development of an analysis pipeline for high-throughput genome-wide PAR-CLIP measurements of virus-infected cells" (Hembach, 2012).

2.4.1. Transition analysis

The pileup format generated from NGS data by SAMtools were filtered to find prominent transitions from T to C. As reference the antigenome of the MeVvac2 genome was used. All possible transitions were extracted with the script *parsePileup.py* written by Katharina Hembach and T to C transitions were isolated using *filterSNP.R*. (Hembach, 2012).

2.4.2. Base content analysis

The MeVvac2 genome was analyzed to find a possible correlation between the base content and the read coverage of RLR libraries in a shifting window of 10 nucleotides. Based on the method *windowComp* written by Katharina Hembach the A, T, G, C, AT, and GC content of (+) and (-) RNA strands were determined in a shifting window of 1 nucleotide (Hembach, 2012). Windows started at the 5' end of the MeVvac2 genome and shifted one base to the 3' end after the base content was determined. This was repeated until the 3' end was reached. The analysis was performed for both RLR data sets. The obtained results were visualized as boxplots (Chambers et al., 1983).

2.4.3 5-mer analysis

5-mer analysis was performed on the MeVvac2 genome. All possible 1024 5-mers were generated and the appearance of the possible 5-mers on the MeVvac2 genome was validated. Each 5-mer, which appears in the MeV genomic sequence, was set into correlation with the median read coverage of the (+) and (-) RNA strands. This analysis was done for both RLR data sets. Finally, the median of the read coverage for each 5-mer was computed. For the analysis, the method *kmerAnalysis* was written by Katharina Hembach (Hembach, 2012). The parameters used in this approach were k = 5 to search for 5-mers and the median coverage threshold T = 0 to report all 5-mer appearances independent from the read coverage.

To identify, 5-mers with a significant median read coverage, the coverage expected by chance was calculated as a reference. Therefore, the MeVvac2 genome was divided into fragments of 20 nucleotides in length. The fragments were then rearranged in a randomized manner. This shuffled genome was used to repeat the shuffling and 5-mer analysis 100 times. After each search the method *repeatedAnalysis*, written by Katharina Hembach, computed the medians of the read coverage of each 5-mer and sorted them in descending order (Hembach, 2012).

Finally, 5-mers from the original genome were sorted according to their medians and compared to the medians of the shuffled genome.

2.4.4. *In silico* RNA secondary structure prediction

RNA secondary structure prediction from MeV genome or *in vitro* transcripts was performed by utilization of RNAfold from the ViennaRNA package, using standard parameter settings. Therefore,

the genome was divided into 201 nucleotide fragments with a shifting window of 5 nucleotides. The sequences were folded *in silico* and each predicted secondary structure, represented in dot-bracket notation, was then analyzed for specific structural features (Hembach, 2012). The linear relationship of structural parameters between different data sets was determined with the Pearson correlation coefficient and set into correlation with the mean read coverage of NGS libraries.

2.4.5. *In silico* prediction of RNase L cleavage sites

Based on the literature the 8 most prominent RNase L cleavage sites were investigated in this thesis (Han et al., 2004). For this, the MeVvac2 genome was divided into fragments of 500 nucleotides. These fragments were then *in silico* folded using RNAfold. Based on the dot-bracket notation of the predicted structures, RNase L cleavage sites in single-stranded regions were determined. The occurrence of each analyzed cleavage site in the 500 nucleotides window was then set into correlation with the read coverage of RLR libraries. The results were visualized with R.

3. Results

3.1. Establishing the PAR-CLIP method

3.1.1. Validation of 4SU incorporation efficiency into measles virus RNA

To establish the PAR-CLIP (**P**hoto**a**ctivatable-**R**ibonucleoside-Enhanced **C**ross**l**inking and **I**mmuno**p**recipitation) technique (**Figure 1.6.**) in the lab and to test its suitability for the identification of naturally occurring RNA ligands of RLR proteins, the incorporation level of photoactivatable 4SU into viral RNA was validated. A549 human lung carcinoma cells were infected with MeV at a MOI (multiplicity of infection) of 1.0 in the presence of 400 µM 4SU. Infection was allowed to occur for 24 hours. Cells were harvested and total RNA was isolated. The RNA was incubated with biotin to separate 4SU-labeled RNA by streptavidin purification (see section 2.2.5.). The fractions of 4SU-labeled and unlabeled RNA were analyzed by quantitative (q)PCR analysis (**Figure 3.1.**).

The data indicate that photoactivatable nucleosides are incorporated into viral RNA. The relative incorporation levels of 4SU into the N gene and into the 5' end of the MeV genome, representing the L gene-trailer region, is roughly 50%. However, according to the control, large amounts of 4SU are still incorporated into endogenous RNA, as it is shown for the relative RNA ratios of GAPDH (**Figure 3.1.**). Thus, the data suggest that 4SU is indeed incorporated into viral RNA, which is the first requirement for studying viral RNA ligands utilizing the PAR-CLIP approach. However, significant amounts of 4SU also incorporate

into endogenous RNA, suggesting a competition between viral and endogenous RNA for the photoactivatable nucleoside analog.

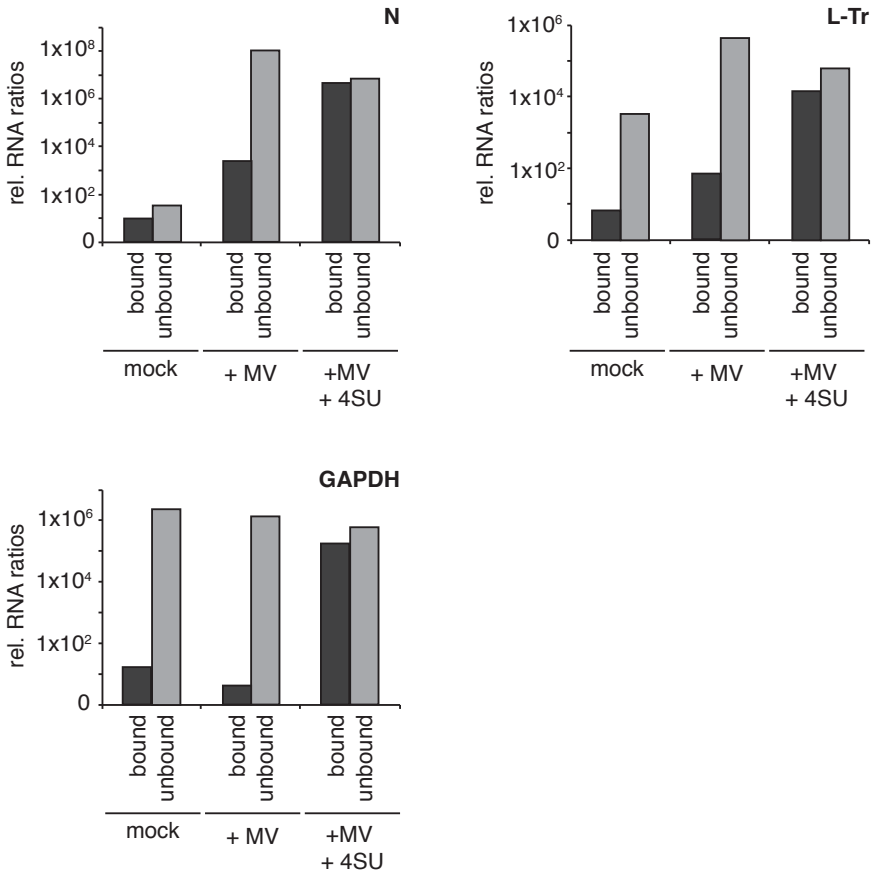


Figure 3.1. Incorporation efficiency of 4SU into viral RNA. 4SU incorporation efficiency was analyzed for the N gene and a region corresponding to the L-trailer region at the genomic 5' end. MeV-infected cells were treated with 4SU and 24 h later 4SU-labeled RNA was separated from unlabeled RNA by biotinylation and streptavidin purification. The fractions were analyzed by (q)PCR. As controls, incorporation levels from mock-infected samples and 4SU-free samples were analyzed. Incorporation into endogenous RNA was validated in the case of GAPDH.

3.1.2. Comparison of RNA recovery between PAR-CLIP and a common pull down approach

For further validation, PAR-CLIP was compared with a common pull down technique previously used for the identification of SeV DI particles as potent RIG-I inducers (Baum et al., 2010). A549 cells were

infected with SeV at high MOI in the presence of 400 µM 4SU and infection was allowed to occur for 24 h. Part of the cells was then exposed to 365 nm UV for crosslinking and endogenous RIG-I was immunopurified. The recovered RNA was isolated and subjected to (q)PCR analysis and immunoactivity experiments.

The analysis reveals that treatment of cells with 4SU and exposure to 365 nm UV lead to a reduction of immunostimulatory activity of RIG-I-associated RNA to 50% (**Figure 3.2. A**). However, (q)PCR data show that the crosslinking approach yields improved RNA recovery with an increase of 50% in SeV DI particles in comparison to the non-crosslinking method (**Figure 3.2. C**). Furthermore, fluorescence microscopy experiments confirm that treatment of cells with photoactivatable nucleoside does not affect cell viability or virus replication (**Figure 3.3.**). In conclusion, the data indicate that the crosslinking technique is a promising tool to study *in vivo* occurring RNA ligands for RLR proteins.

3.1.3. Determination of viral concentration for infection experiments

A549 cells were infected with different concentrations of a genetically modified MeVvac2 strain encoding GFP (provided by Prof. Dr. Karl-Klaus Conzelmann) in order to identify the optimal viral concentration for the PAR-CLIP experiment. 24 hpi (hours post infection) MeV replication was visualized by fluorescence microscopy (**Figure 3.4.**). The fluorescence images indicate that an increasing number of cells become infected in a concentration-dependent manner. According to the results an MOI of 1.0 was chosen for further experiments. At the given concentration A549 cells remain viable and a reasonable amount of cells become infected (**Figure 3.4. C**). Although, higher titers would

lead to a better infection efficiency, these concentrations are not feasible due to practical problems in MeV production.

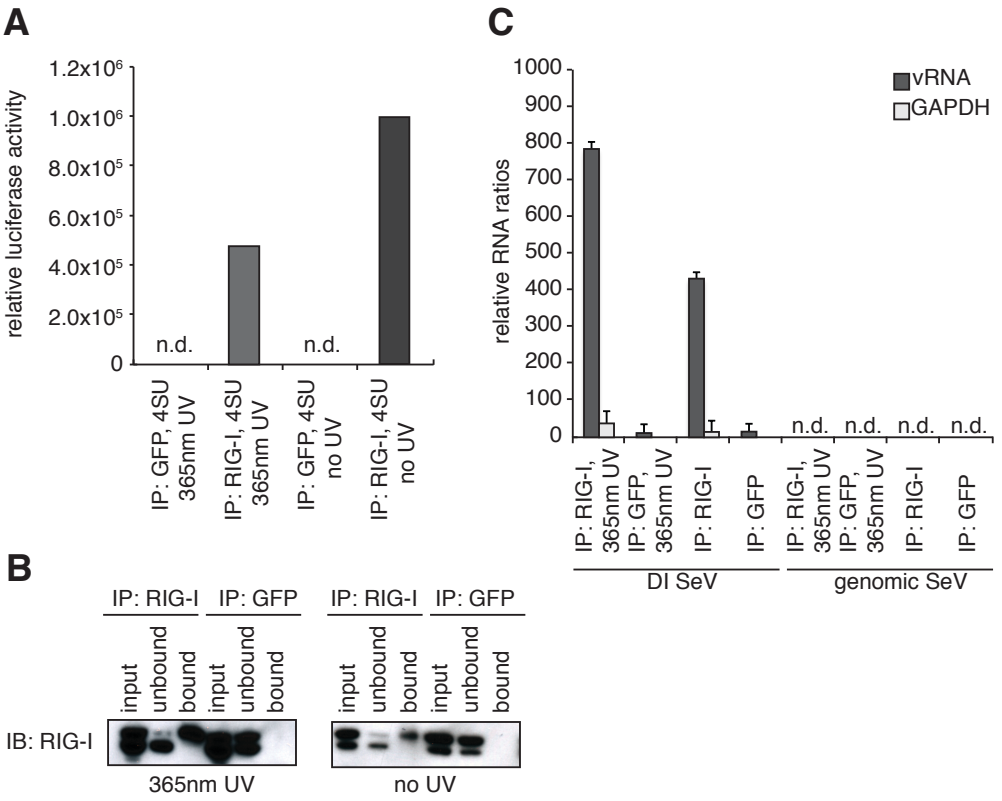


Figure 3.2. Validation of crosslinking and immunoprecipitation of RIG-I/RNA complexes from 24 h virus-infected cells. **(A)** Immunostimulatory activity of RNA from RIG-I and control (GFP) crosslinking samples (365 nm UV) in comparison to non-crosslinking immunoprecipitates (no UV) were validated by transfection of RNA into 293T ISRE FF/RFP reporter cells. **(B)** Immunoblot (IB) analysis of RIG-I pull down experiments shows the efficiency of RIG-I immunopurification in comparison to GFP. **(C)** Comparison of RNA recovery levels by (q)PCR analysis of RIG-I-associated RNA from SeV-infected cells (n=3). Q-PCR reactions incorporated human actin β controls and relative abundance of each RNA was calculated with reference to this control.

Next, MeV infection was investigated on a protein level. Therefore, A549 cells were infected with virus (MOI of 1.0) in the presence or absence of 4SU. Viral protein expression was validated by Western blot analysis 6 and 24 hpi, respectively (**Figure 3.4. B**). The obtained data show slight expression of P and V protein 6 hpi indicating that the transcription machinery is already activated shortly after the virus entered the cell. 24 hpi translation of viral proteins become strongly

upregulated as reflected by the robust expression of viral proteins in the cell lysate of A549 cells.

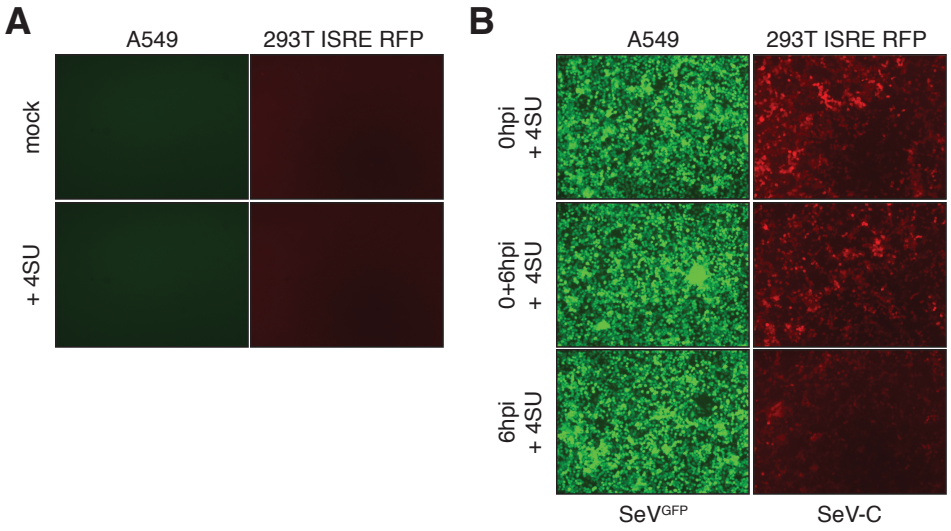


Figure 3.3. Effect of 4SU on viral replication. A549 cells and 293T ISRE RFP cells were infected with SeV^{GFP} and SeV-C, respectively. Viral replication was allowed to proceed over 24 h and then analyzed by fluorescence microscopy. **(A)** Cells were grown in the absence of virus as a negative control. **(B)** Cells were infected with the respective virus at a high MOI and 4SU was added as indicated (at 0 hpi and/or 6 hpi). Viral replication was validated by viral GFP expression or by RFP production from the reporter cells.

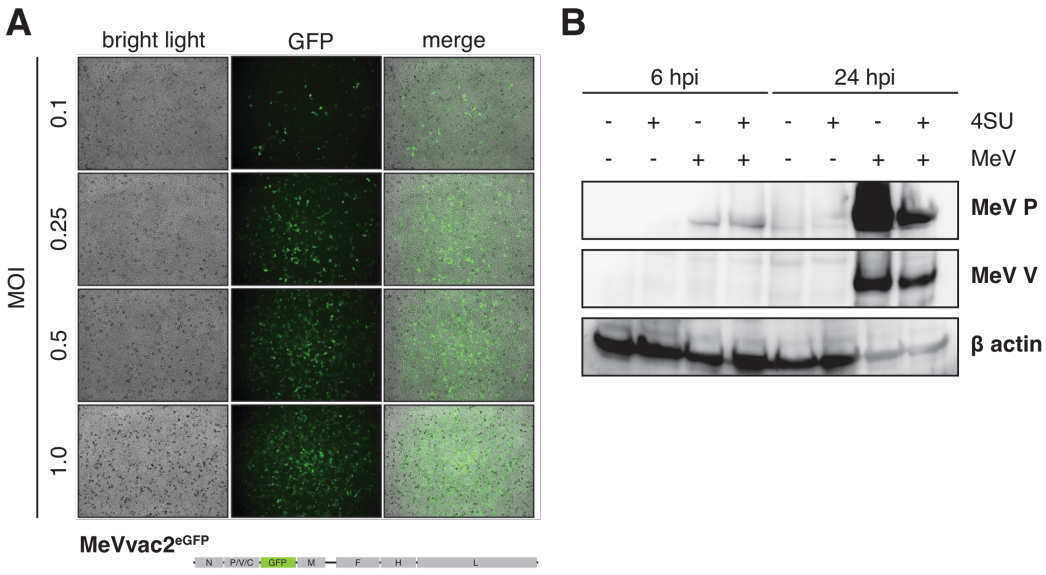


Figure 3.4. (A) MeVvac2^{GFP}-infected A549 cells. A549 cells were infected with MeVvac2^{GFP} at different MOIs and 24 hpi, infection was monitored by fluorescence microscopy. *Lower*: Schematic representation of the genome organization of the genetically modified MeVvac2^{GFP} strain. **(B)** Viral replication at different time points. By Western blot analysis production of viral P and V protein was monitored 6 hpi and 24 hpi in the presence or absence of 400 µM 4SU, respectively. β actin was used as a loading control .

Comparison between 4SU treated and untreated cells indicate that the nucleoside has no significant impact on viral gene expression, which is in good agreement to the above-mentioned results (**Figure 3.4. B**). However, it is noted that endogenous protein levels (β-actin) dramatically decrease 24 hpi, suggesting that at late time points of infection the virus already interferes with the cell cycle.

3.1.4. Validation of RNA recovery efficiency from MeV-infected cells

To finally show that the crosslinking strategy yields higher RNA recovery levels from RIG-I and MDA5 immunoprecipitates, A549 cells were infected with MeV in the presence of 4SU followed by 365 nm UV exposure 24 hpi. As a reference, control pull downs for RIG-I and MDA5 were performed. The recovered RNA from normalized cell lysates was radioactively labeled and separated by native PAGE electrophoresis (**Figure 3.5.**). The phosphorimages of separated RNAs clearly show an increased RNA recovery from virus-infected samples treated with 4SU and UV light in comparison to the non-crosslinking approach (**Figure 3.5. A**). Quantification of the RLR-associated RNA confirms an improved RNA recovery level up to 30% (**Figure 3.5. B**). Remarkably, RLR immunoprecipitates from non-infected cells do not yield improved RNA recovery even with PAR-CLIP, suggesting that RLRs require viral infection in order to sense RNA ligands.

3.1.5. 4SU treatment and 365 nm UV exposure yield immunoactive RNA in an RLR-dependent manner

After determining the initial parameters for the PAR-CLIP experiment, different viruses including negative-strand (-) RNA viruses (MeV and Rabies) and postive-strand (+) RNA viruses

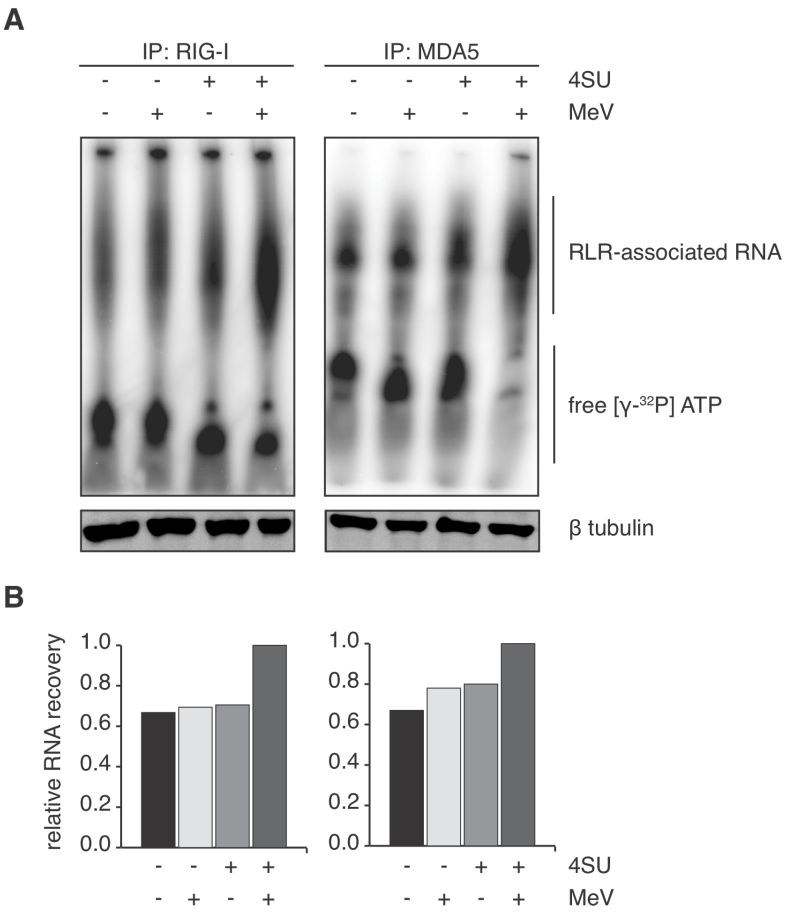


Figure 3.5. RNA recovery efficiency from MeV-infected cells. **(A)** *Upper*: Phosphorimage of ^{32}P-labeled RNA from different pull down conditions. RLR-associated RNA from MeV-infected or non-infected A549 cells in the presence or absence of 4SU, were radioactively labeled and separated by native PAGE. *Lower*: Western blot analysis of normalized cell lysates. **(B)** Quantification of relative RNA recovery levels from RLR immunoprecipitates. Phosphorimages of RLR-associated RNAs were quantified with ImageJ. The highest value was set to 1.

(Encephalomyocarditis virus (EMCV) and, Mengo virus) were tested. A549 cells were infected with the respective virus in the presence of the photoactivatable nucleoside and 365 nm UV crosslinked 24 hpi. RIG-I and MDA5 were immunoprecipitated and the recovered RNA was subjected to immunoactivity experiments. The data reveal that immunoactive RNA is co-purified in a RIG-I- and MDA5-dependent manner from MeV-infected cells. The induction is significant in comparison to the negative control (**Figure 3.6.**). In case of RIG-I-associated RNA an immunostimulatory effect is observed that is 2600-

fold higher in comparison to the control. For MDA5 a 800-fold induction is reported (**Figure 3.6.**). However, RNA co-purified from additional viruses does not show significant immunostimulatory activity. The reason for this may lie in the heterogeneity of viral replication cycles. Especially, replication of picornaviruses reveals remarkable differences in regard to *Mononegavirales*. But also closely related viruses like measles and rabies exhibit differences in their mode of replication that hamper the investigation of viral RNA ligands from different viruses with one established protocol. In respect to such discrepancies, the crosslinking approach has to be fine-tuned and adjusted for each viral background.

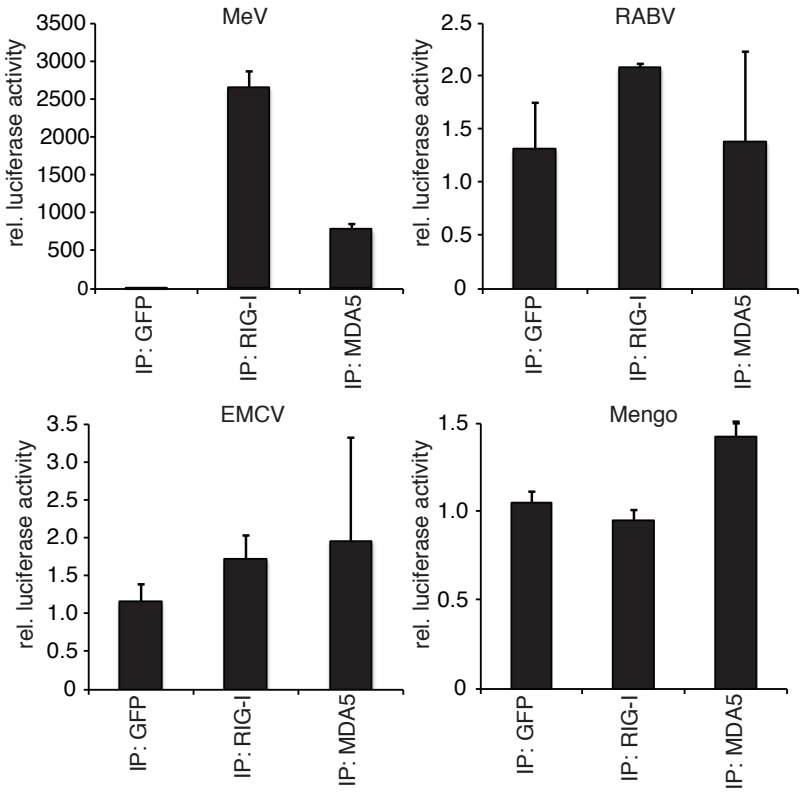


Figure 3.6. Immunoprecipitation of RLR-associated RNA from cells infected with different types of viruses. Validation of immunostimulatory activity of RNA from RIG-I, MDA5, and GFP immunoprecipitates upon transfection into 293T ISRE-FF/RFP cells (n=3).

3.2. Next generation sequencing (NGS): an unbiased method to study RLR-associated viral RNA

3.2.1. NGS reveals regions within the Measles virus genome recognized by RIG-I and MDA5

In light of the results mentioned in section 3.1., the focus of this thesis was on the investigation of endogenous RLR ligands from MeV-infected cells. Concerning this question, A549 cells were the preferred cell system, since these cells adhere as a monolayer to the cell culture dish, allowing for homogenous 365 nm UV crosslinking. Furthermore, A549 cells exhibit robust RLR expression, making them ideally suited for the analysis of naturally occurring RNA ligands of the endogenous sensor proteins.

Due to the complex RNA composition in a virus-infected cell, the analysis of specific RNA ligands for a receptor protein is challenging (section 1.4.). To shed light on the exact nature of RIG-I and MDA5-associated RNAs derived from MeV-infected cells, NGS (**N**ext **g**eneration **s**equencing) analysis on isolated RNA species was performed. The application of NGS holds the advantage to investigate RNA without any prior knowledge about the nature of the molecule. In more traditional methods, like Northern blot and quantitative PCR, the nature of probes and primers can greatly influence the experimental outcome and commonly introduce bias to the data set. These disadvantages can be eliminated by the NGS technology.

Co-purified RNA from MeV crosslinking experiments was prepared for NGS analysis according to the Epicentre RNA-Seq protocol and sequenced on the Illumina Genome Analyzer (see section 2.3.2.). The obtained sequences were mapped to the MeV antigenome and relative

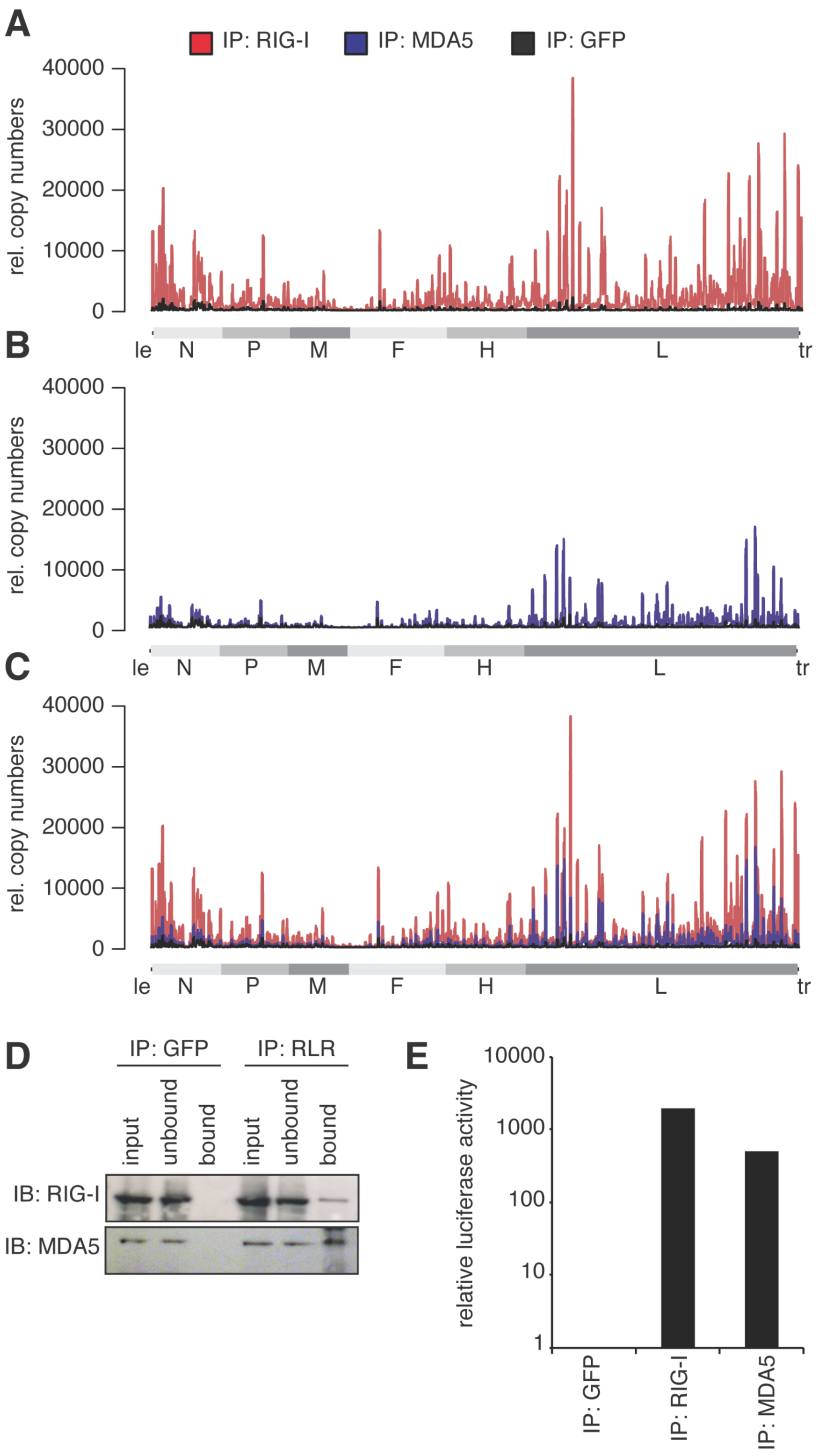


Figure 3.7. Next generation sequencing (NGS) of RLR-associated RNA from MeV-infected cells. **(A)** RNA from RIG-I pull down (red), **(B)** MDA5 pull down (blue), and **(C)** GFP pull down (black) from MeV-infected cells were subjected to Illumina NGS analysis. The obtained sequences were mapped to the MeV antigenome and relative abundances of these sequences between RIG-I, MDA5, and GFP pull down were compared. The lower panel shows an overlay of all three libraries **(D)** Western blot analysis shows the efficiency of RIG-I and MDA5 immunopurification in comparison to the GFP pull down. **(E)** Validation of immunostimulatory activity of RLR-associated RNA.

abundances of these sequences between RIG-I, MDA5, and GFP pull downs were compared. Individual peaks on the graph correspond to a sequencing read that starts at that particular position. The x-axis corresponds to all possible 15,894 positions in the MeV antigenome and the y-axis shows the number of reads that begin at that position. (**Figure 3.7.**).

Analysis of the reads reveals that RIG-I and MDA5 bind to same regions within the L genomic segment. In addition, RIG-I, but not MDA5, binds to the 3' and the 5' ends of the MeV genome (**Figure 3.7. A, B**). These regions correspond to leRNA and trRNA that are generated in the course of replication (Mottet et al., 1990). Furthermore, the data indicate that RIG-I binds to MeV-derived RNA with higher affinity in comparison to MDA5. This observation suggests a major role for RIG-I in RNA recognition from MeV-infected cells, which is also reflected by the increased immunostimulatory activity of isolated RNA from RIG-I immunoprecipitates. Accorrding to the immunostimulatory activity, RIG-I-associated RNA gives a 4-fold induction in regard to MDA5-associated RNA (**Figure 3.7. D, E**).

3.2.2. Analysis of NGS data reveals striking differences in the strand-specificity of RIG-I and MDA5

Based on the protocol used for cDNA library preparation, sequencing reads could be separated according to their strand orientation. During cDNA synthesis adaptors were specifically ligated to the 3' and 5' ends, thereby keeping the information of strand specificity during the sequencing run. Separation of sequences reveals remarkable differences between RIG-I and MDA5 libraries (**Figure 3.8.**). In regard to RIG-I, the sequence distribution indicate that RNA sequences of positive polarity, which represent either antigenome or mRNA

transcripts, are predominantly enriched in the genomic leader region and within distinct region of the L genomic segment (**Figure 3.8., red**). Sequences of negative polarity, which represent genomic RNA, originate from the 5' end corresponding to the trailer region and from distinct regions within the L gene (**Figure 3.8., magenta**).

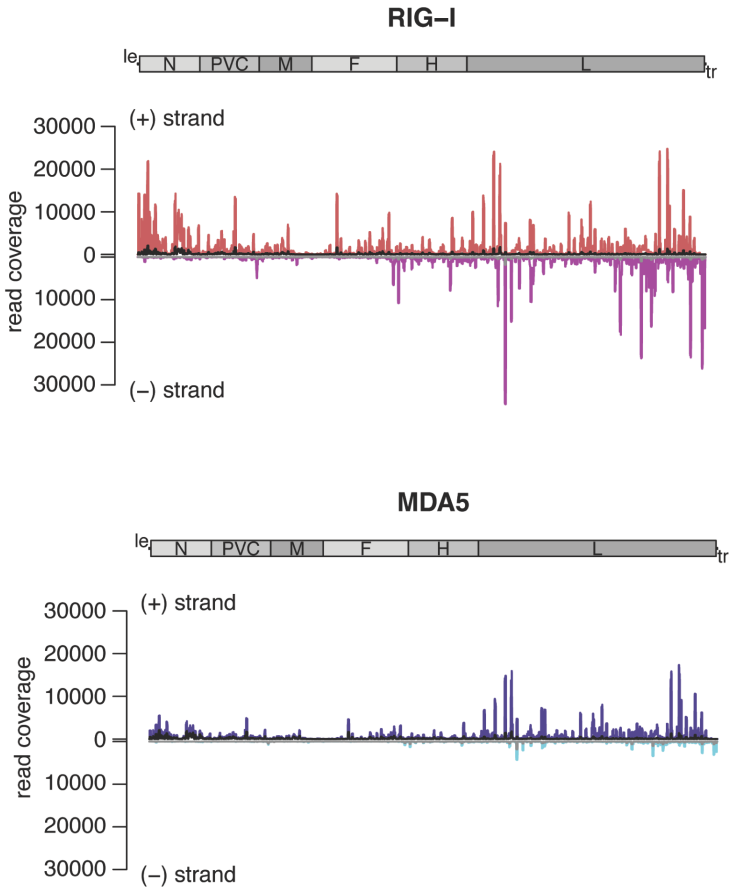


Figure 3.8. Analysis of NGS data revealed differences in the strand specificity of RNAs recognized by RIG-I and MDA5. RNA NGS libraries were generated based on the strand-specific mRNA sample preparation protocol from Epicentre (see section 2.3.2). The Epicentre protocol encompasses sequential ligation of 5' and 3' adapters to RNA molecules, thus preserving strandness information. RLR-associated sequences were separated into (+) and (-) RNA and mapped against the MeV antigenome. *Upper:* RIG-I-associated sequences in comparison to the control (black, grey) mapped to the antiviral genome. (+) RNA is shown in red; (-) RNA is shown in magenta. *Lower:* Sequences representing (+) RNA (blue) and (-) RNA (cyan) from MDA5 libraries were mapped in comparison to the control (black, grey).

Analysis of MDA5-associated RNA reveals that sequences of positive polarity are enriched within the L gene originating from similar

regions as (+) RNA from the RIG-I library (**Figure 3.8., blue**). However, MDA5 does not bind to sequences originating from the 3' and 5' ends of the MeV genome. Comparison of enriched (-) RNA from RIG-I and MDA5 libraries shows that in contrast to RIG-I, MDA5 does not bind to RNA of negative polarity, which represents the viral genome (**Figure 3.8., cyan**). Moreover, the data evidently rules out the possibility that MDA5 recognizes heterogeneous RNA duplexes of (+) and (-) RNA that might represent replication intermediates as previously described (Feng et al., 2012). In fact, the results suggest that MDA5 bind (+) RNA that could either represent MeV mRNA or antigenomic RNA species.

3.2.3. Copy number variation (CNV) analysis: a bioinformatical tool to validate significant sequence enrichment between NGS libraries

To validate the significant enrichment of RIG-I- and MDA5-associated RNA in comparison to the negative control, copy number variation (CNV) data were generated (Wood et al., 2010). This analysis allowed to compare enriched sequences in each library in regard to the control. CNVs were validated by determining the normalized sequence abundances at each position on the genomic segment and by calculating the average of the RLR/GFP ratios over a dynamic window of 200 reads. Sequence abundances with Log2 ratios above +1 were defined as significantly enriched in the RLR libraries (**Figure 3.9.**).

The CNV karyograms indicate that RIG-I-associated RNA of positive polarity is remarkably enriched across the whole genome with the majority of specific sequences in the N genomic and L genomic regions (**Figure 3.9., A and 3.10.**). In the case of (-) RNA, CNV data reveal

slight enrichment for the 5' end of the MeV genome but not for the 3' region (**Figure 3.9., A**).

In contrast to RIG-I-associated RNA, CNV analysis for MDA5-associated RNA clearly indicates that (+) RNA is significantly enriched from region corresponding to the L genomic segment, while no enrichment is observed for (-) RNA (**Figure 3.9. B and 3.10.**). The data nicely correlate with the NGS analysis demonstrating that distinct regions of the MeV genome are indeed specifically enriched in an RIG-I- and MDA5-dependent manner in comparison to the control.

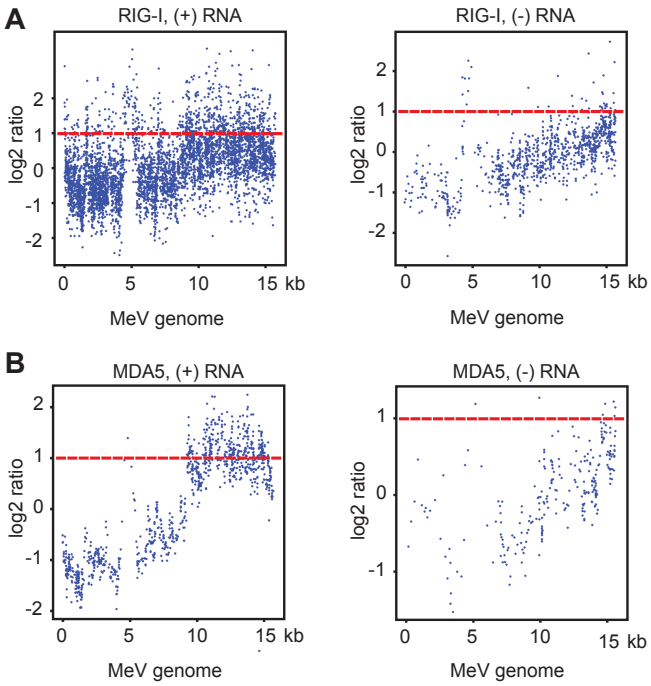


Figure 3.9. Generation of NGS-based copy number karyograms from RLR RNA sequencing libraries. Data points with log2 ratios above 1 represent sequencing reads that were significantly enriched in comparison to the control (GFP) library (red, dashed line). **(A)** The copy number profiles for RIG-I are illustrated. (*left*) copy number variation data representing (+) RNA; (*right*) copy number variation data representing (-) RNA. **(B)** Copy number profiles for MDA5 libraries. (*left*) copy number variation data representing (+) RNA; (*right*) copy number variation data representing (-) RNA.

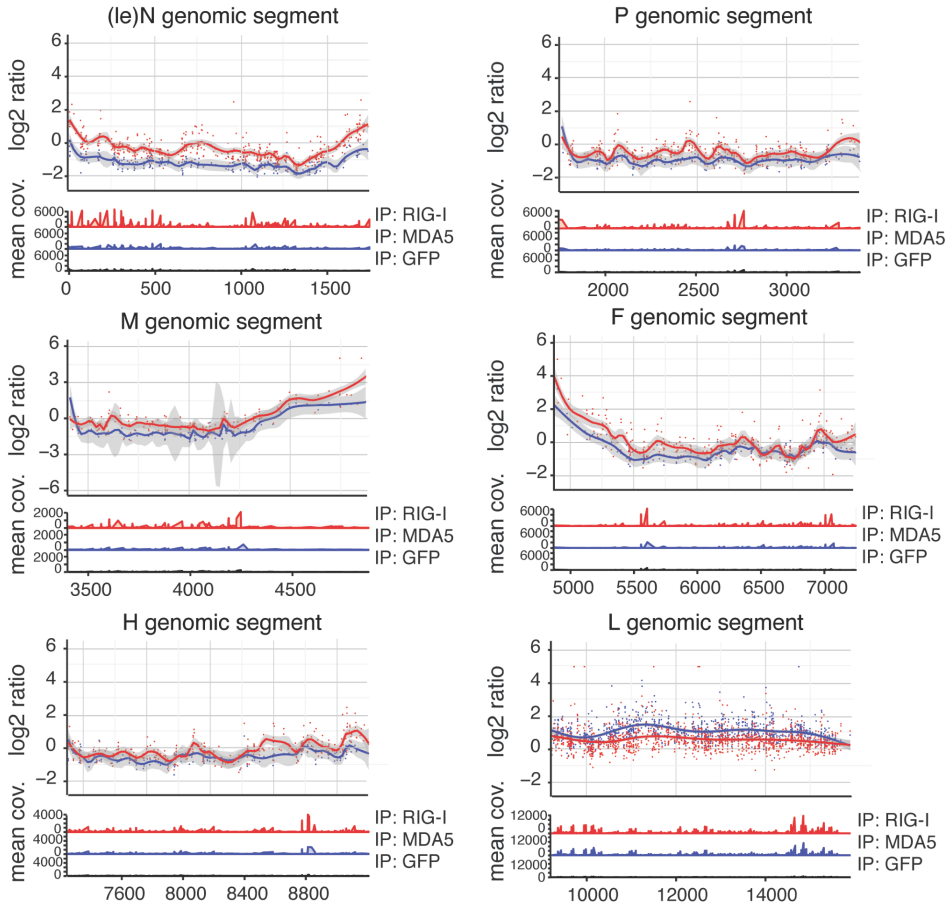


Figure 3.10. Karyograms of copy number variation analysis for MeV genomic segments. The calculation is based on a dynamically calculated window size. The algorithm determines the start and end points in an interval 200 reads which are used for the CNV calculation. CNV values with a log2 ratio above 1 are significantly enriched in comparison to the control library. (*Upper*) Copy number variation data for each genomic segment. Each data point corresponds to a CNV value for the calculated window. The mean values for RIG-I and MDA5 are shown in red and blue, respectively. Standard deviations are represented in grey. The high log2 ratios in the M and F genes, are caused by the low copy numbers in GFP in comparison to RLR libraries (*Lower*). Alignment of sequences against the MeV antigenome from RLR and GFP immunoprecipitates. The mean coverage of sequences are shown for RIG-I (red), MDA5 (blue), and GFP (black) for (+) RNA.

3.3. Quantitative PCR analysis: Confirmation of NGS experiments by an independent approach

To independently validate the relative amount of RLR-associated RNA, (q)PCR analysis was performed. The obtained copy read numbers were normalized to compare the genomic segments in the RIG-I and MDA5 samples (**Figure 3.11. A, B**). Analysis of relative abundances confirms that RIG-I specifically enriches sequences in the 3' and the 5' regions of the viral genome, representing antigenome or viral mRNA. Interestingly, the analysis shows that RIG-I-associated RNA from the 3' end most likely represent leRNA-N mRNA read-through transcripts or abortive replication products, but no mRNA from the N genomic segment (**Figure 3.11. A**). In regard to MDA5, RNA is enriched in the case of the L genomic segment and partly from the H gene, but not from upstream genomic regions (**Figure 3.11. A**).

The analyzed regions within the MeV genome do not exactly reflect the results of the NGS experiment. For each genomic segment, primer pairs were chosen that represent 200-nucleotide regions at the end of each gene, thereby covering only a very restricted area of the genome. Nevertheless, the obtained data are in good agreement to the results of the NGS analysis, indicating that RIG-I and MDA5 indeed recognize RNA originating from the L gene of the MeV genome.

Furthermore, the relative abundances in the RIG-I sample were up to 40-fold higher in comparison to MDA5 (**Figure 3.11. C**). This again indicates that RIG-I binds to MeV RNA with a much higher affinity in comparison to MDA5. This conclusion is further supported by immunoactivity experiments, where relative immunostimulatory activity of RIG-I-associated RNA is 20-fold higher in comparison to MDA5 (**Figure 3.11. D**).

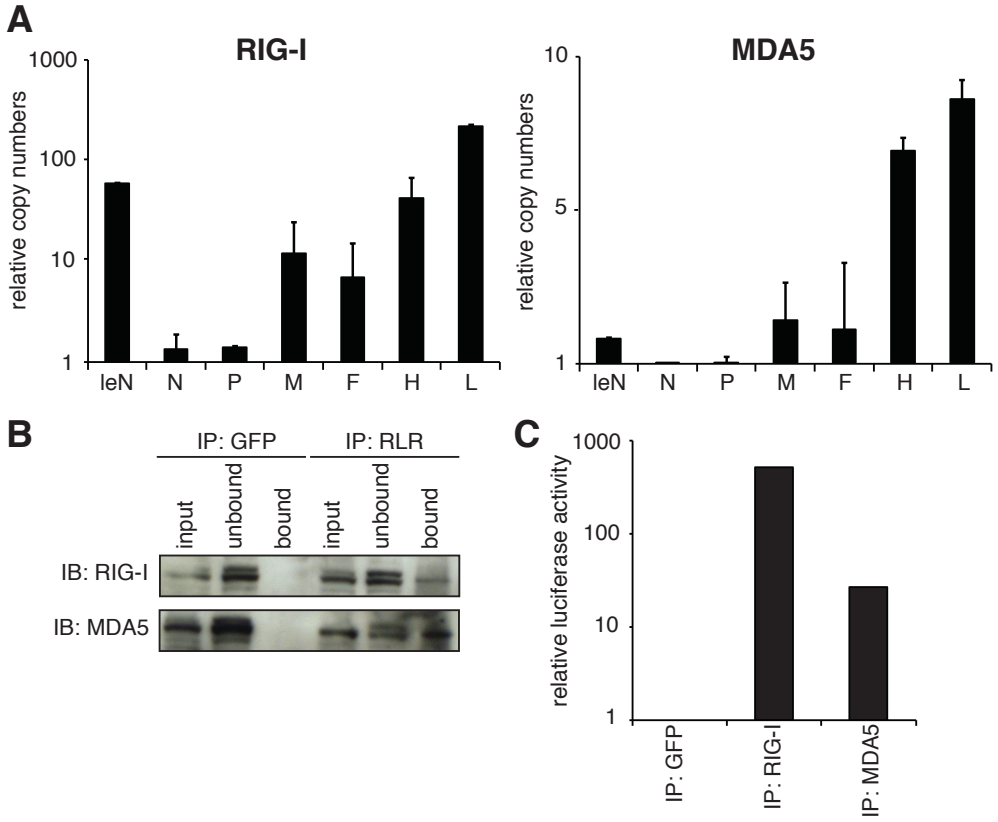


Figure 3.11. Quantitative PCR analysis of RLR-associated RNA from MeV-infected cells. **(A)** Comparison of (+) RNA levels between RIG-I and MDA5 immunoprecipitates for each genomic segment at 24 hpi. Relative RNA ratios were normalized against the control (GFP) library. **(B)** Western blot analysis of RLR pull down experiments show the efficiency of RIG-I and MDA5 immunopurification in comparison to the GFP pull down. **(C)** Immunstimulatory activity of RLR-associated RNA at 24 hpi. The qPCR reactions incorporated multiplexed human actin β internal controls and relative abundance of RNA was calculated with reference to this control.

3.4. Bioinformatical analysis: Characterization of the RLR-associated RNAs

The quality of NGS data was assessed by bioinformatical approaches. Based on the obtained results, the nature of endogenous RNA ligands for RLRs was further characterized. The bioinformatical analysis was performed in collaboration with the group of Dr. Johannes Soeding. The main contribution to this part of the thesis came from Katharina

Hembach, who wrote the bioinformatical scripts for the analysis of the NGS data and did the evaluation. Parts of the results were published in her Bachelor thesis "Development of an analysis pipeline for high-throughput genome-wide PAR-CLIP measurements of virus-infected cells" (Hembach, 2012).

3.4.1. Mutational analysis: Deciphering the exact binding site

According to the paper from Hafner et al. the most significant advantage of the PAR-CLIP technique to commonly used pull down strategies, is the introduction of specific T to C transitions at the crosslinking sites between 4SU-labeled RNA and the RNA-binding protein (Hafner et al., 2010). It is speculated that these transitions are caused by a chemical structure change of the 4SU nucleobase upon crosslinking to the amino acid side chain of the protein. This alteration could result in modified donor/acceptor properties of the hydrogen bond, directing the preferential incorporation of dG rather than dA during reverse transcription (Hafner et al., 2010). The resulting T to C transitions could be used to exactly pinpoint the binding sites of the respective protein in the NGS libraries.

To identify putative binding sites by mutational analysis, the NGS data was filtered for all possible transitions and transversions (see section 2.4.1.). The analysis of (+) RNA from RIG-I immunoprecipitates revealed that beside T to C transitions, other mutations (e.g. C to T and G to T) occur at frequencies that are comparable to those of T to C (**Figure 3.12. A**). Furthermore, the relative mismatch frequencies are comparable between the different libraries, with the exception for G to T, which is the most frequent mutation in (+) RNA of the RIG-I library (**Figure 3.12. A**).

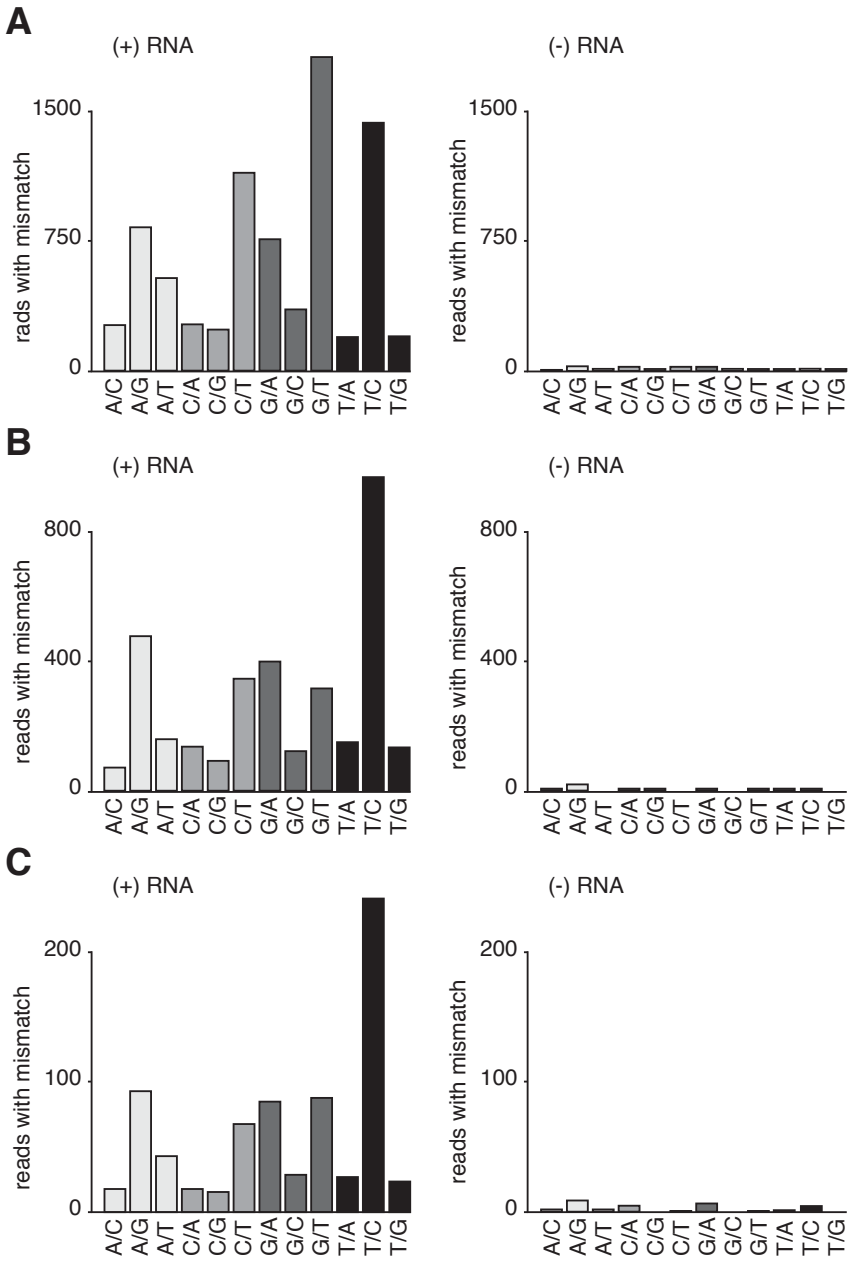


Figure 3.12. Analysis of transition/transversion frequencies in the NGS data set. For the mutational analysis, pileup files were filtered for all possible transition/transversions. Each barplot represents the conversion frequency of a particular nucleobase. The y-axis reflects the total number of reads that contain the specified mutation (x-axis). **(A)** Barplots for each possible transition in the RIG-I dataset. **(B)** MDA5 dataset. **(C)** GFP dataset. *Left*: (+) RNA; *right*: (-) RNA.

This transversion could result from crosstalk phenomena in the base-calling step. The nucleobase pairs A/C and G/T exhibit similar emission spectra that could result in wrong base calls in the Illumina

sequencing strategy (Li and Speed, 1999). However, if this were the case, (+) RNA from MDA5 and GFP would also show an increased frequency in G to T transversions. Interestingly, no significant mismatch numbers were found in (-) RNA of the RIG-I library despite the high copy numbers (**Figure 3.8. and 3.12. A**).

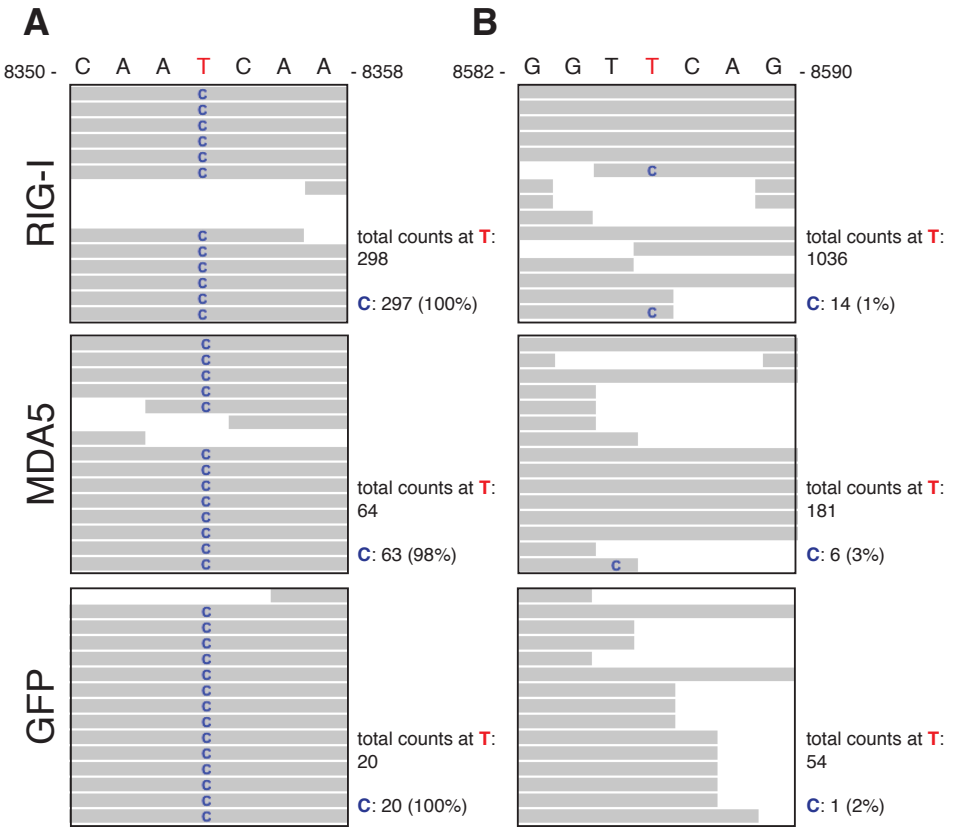


Figure 3.13. Characterization of T to C transitions in the NGS data. The observed T to C transitions can be divided into two groups. **(A)** T to C conversion with a high frequency originate from 100% of the read counts in the NGS libraries. These mismatches could represent wrongly annotated nucleobases but are unlikely to result from specific crosslinks at 4SU moieties. **(B)** The majority of T to C conversions occurs only at a few percent in the read counts of NGS libraries (below 5%), which indicates that these transitions originate from non-crosslinked 4SU in the RNA molecules and therefore do not represent interaction sites between the RNA and the receptor protein. T to C transitions were visualized using the Integrative Genomic Viewer (James et al., 2011).

In regard to MDA5 and GFP, T to C transitions are indeed the most frequent mutations observed in the NGS data (**Figure 3.12. B, C**). These conversions have a 2-fold higher frequency in comparison to the

second most common transition (A to G). At this stage of the analysis it is not clear, whether the observed T to C transitions result from crosslinks between MeV RNA and the receptor proteins, or whether unspecific conversions from non-crosslinked 4SU-labeled RNA contribute to the T to C transition rate. However, the slight increase in T to C rather suggests that the crosslinking did not induce covalent linkages at 4SU moieties of the RNA. This is further supported by the observation that some T to C conversions can be found in almost 100% of the total read counts of all libraries (**Figure 3.13. A**). This indicates that the T to C mutations result from wrongly annotated nucleobases in the reference genome, rather than from specific crosslinks at the interaction sites. Even if the crosslinking strategy efficiently induced T to C conversions, one would expect a maximum transition rate of 80% in comparison to a non-crosslinking method (Hafner et al., 2010). Furthermore, the majority of T to C transitions occurs only at a few percent in the total read counts (**Figure 3.13. B**). According to Hafner et al. the average T to C conversion rate derived from background non-crosslinked 4SU-labeled sites is near 5% (Hafner et al., 2010). This and the occurrence of T to C in the GFP library, indicate that these transitions do not result from specific crosslinking sites, but rather represent background conversions or mismatches introduced during the sequencing run.

It is possible that the improved RNA recovery in comparison to non-crosslinking techniques (**Figure 3.5.**) is based on the high doses of UV exposure, applied in the modified PAR-CLIP protocol. This could result in unspecific crosslinks between non-substituted nucleobases and the interacting amino acid side chains of the protein. Based on the mutational analysis, T to C transitions are preliminary considered to be not reliable to investigate RLR binding sites in the viral RNA.

3.4.2. Analysis of the base content in a shifting window of 10

Since the transition analysis did not allow a reliable interpretation for preferred binding sites, the base content of the MeV genome was analyzed and set into correlation to the mean read coverage of sequences. It is possible that a simple nucleotide bias in the viral genome leads to a preferred binding of RNA sequences towards other species. To address this, the genome was analyzed in windows of 10 nucleotides with a shifting number of 1 nucleotide. The numbers of bases in the analyzed windows were set into correlation with the mean read coverage of RIG-I and MDA5 libraries (**Figure 3.14. and 3.15.**). The results are represented as boxplots for the indicated occurrence of the base content (x-axis) and the corresponding mean read coverage in the window (y-axis). The black line in each box represents the median for each analyzed occurrence of the respective nucleobase. Each boxplot is drawn with a notch around the median. Non-overlapping notches between two boxes indicate that their medians differ statistically significant (Chambers et al., 1983).

Based on the analysis of the calculated medians, no significant nucleotide bias is observed that might explain the sequence enrichment in RLR libraries (**Figure 3.14. and 3.15.**). In regard to RIG-I, a slight correlation between the G content of (+) RNA and the mean read coverage of sequences is observed. However, this is in contrast to the G/C composition, which does not show any correlation with the mean read coverage of (+) RNA sequences (**Figure 3.14. A**). Indeed, the data rather suggest that A/U-rich sequences slightly correlate with the mean read coverage of (+) RNA (**Figure 3.14. A**), while sequences with a high G/C content, tend to correlate with the mean read coverage of (-) RNA (**Figure 3.14. B**).

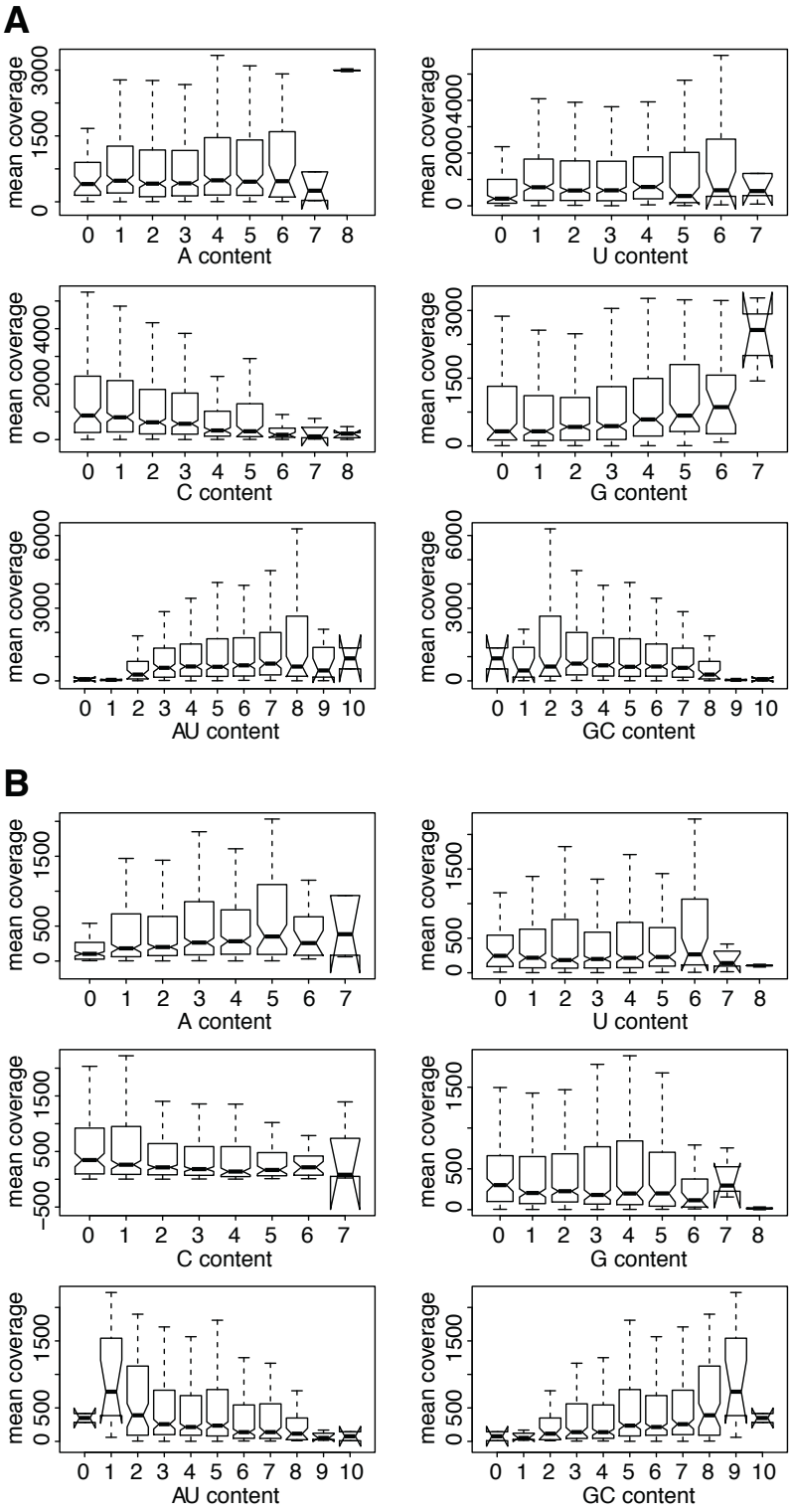


Figure 3.14. Analysis of the base content in a shifting window of 10. Six individual plots represent the base content in a shifted window of 10 nucleotides. Each plot contains boxplots of the indicated occurrence (x-axis) of the particular base content and the corresponding mean read coverage in the window (y-axis) **(A)** Boxplots of the base content and the mean read coverage of (+) RNA from RIG-I libraries. **(B)** Boxplots of the base content and the mean read coverage of (-) RNA from RIG-I libraries.

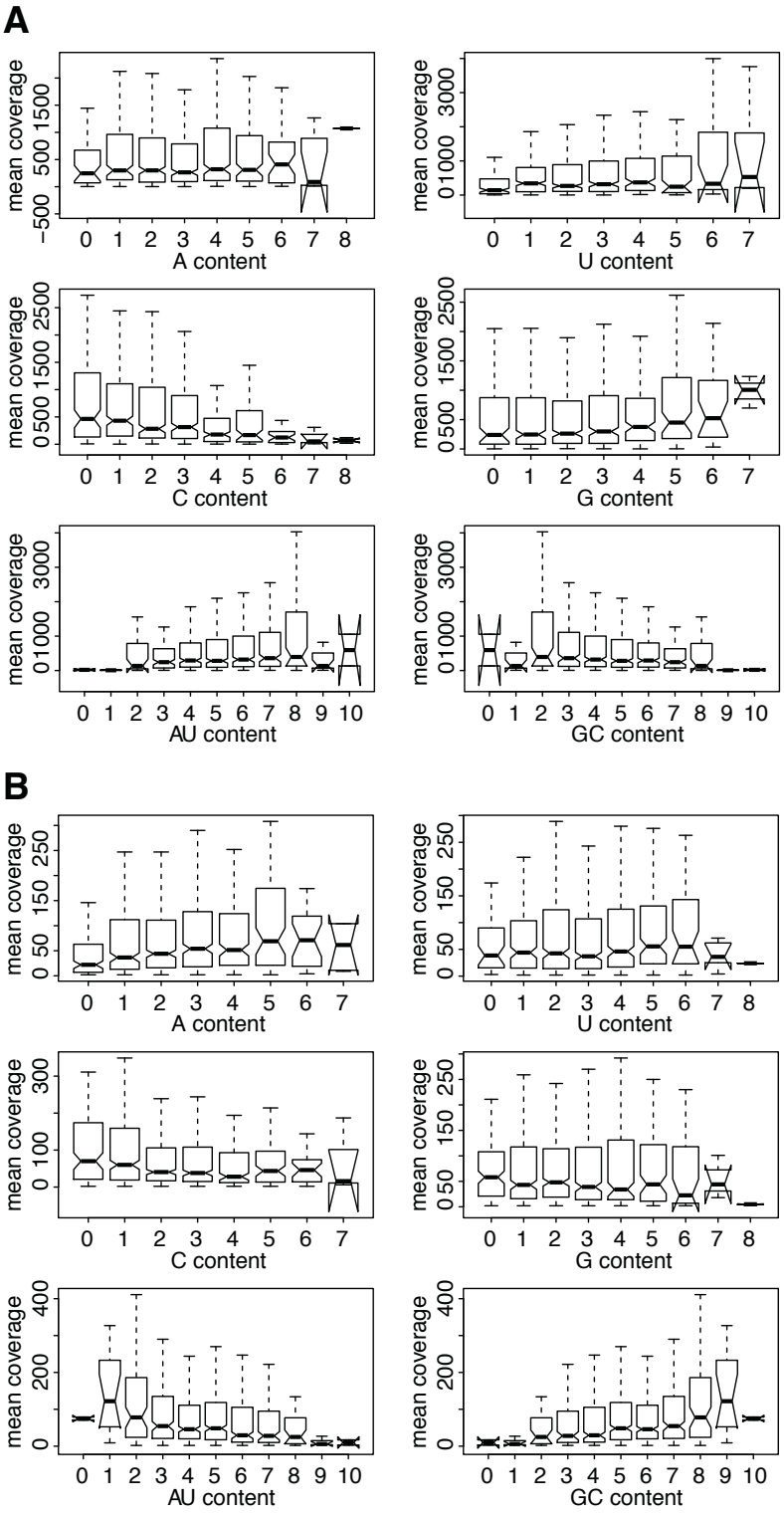


Figure 3.15. Analysis of the base content in a shifting window of 10. Six individual plots represent the base content in a shifted window of 10 nucleotides. Each plot contains boxplots of the indicated occurrence (x-axis) of the particular base content and the corresponding mean read coverage in the window (y-axis). **(A)** Boxplots of the base content and the mean read coverage of (+) RNA from MDA5 libraries. **(B)** Boxplots of the base content and the mean read coverage of (-) RNA from MDA5 libraries.

In the case of MDA5, similar results were obtained (**Figure 3.15. A, B**) indicating that enriched sequences from RIG-I and MDA5 immunoprecipitates have the same base composition. According to the high similarity between the different medians and the overlapping notches of the particular boxes, it is obvious that the sequence enrichment, observed in the RLR libraries, is not caused by a simple nucleotide bias.

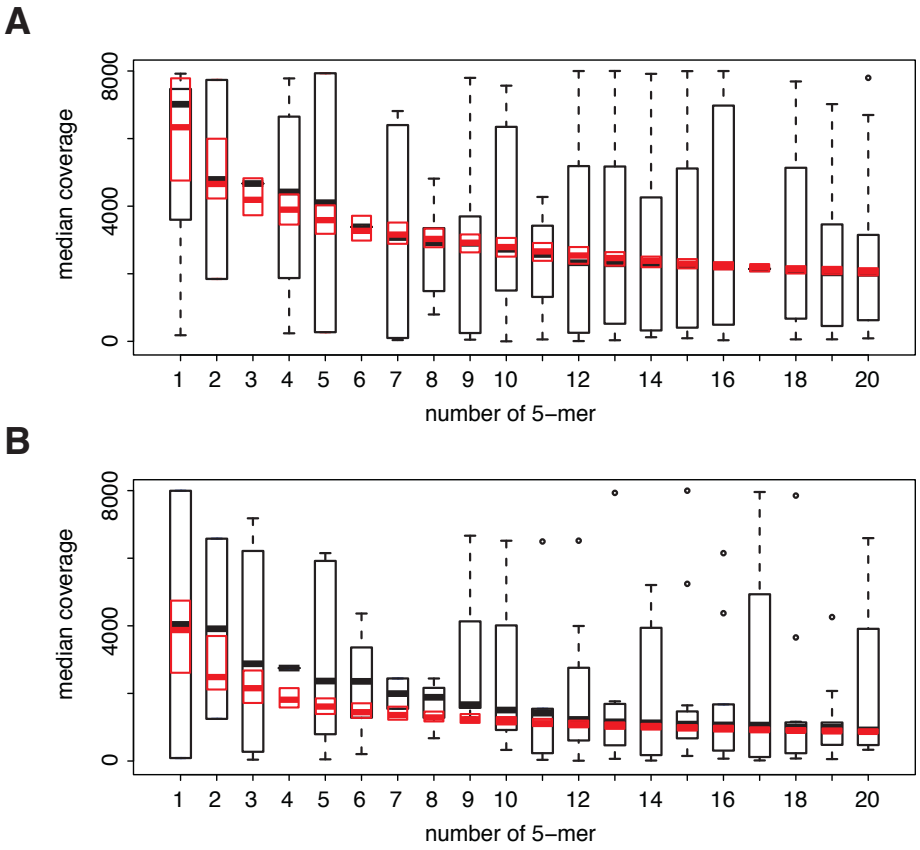
3.4.3. 5-mer analysis with the objective to study viral sequence motifs

Towards the analysis of possible sequence motifs in the MeV genome, 5-mers and their correlation with the median read coverage of RLR libraries were investigated. For this, all 1024 possible 5-mers were generated and for each 5-mer, which appears in the MeV genome, the median read coverage was determined. In order to identify 5-mers with a significant median read coverage; the coverage expected by chance was also calculated. Therefore, the MeV genome was divided into fragments of 20 nucleotides and rearranged in a randomized order. With this shuffled genome the same 5-mer analysis was performed. 5-mers with the 20 highest medians were sorted in descending order (**Figure 3.16. and 3.17., Table 3.1. and 3.2.**).

Comparison of 5-mers with an appearance above 5 and a higher median read coverage in regard to the expected values indicate that these 5-mers tend to be rich in A/U rather than G/C (**Table 3.1. and 3.2.**). Especially, the sequence UAGUA is slightly enriched in comparison to the expected medians in the (+) RNA of RIG-I and MDA5 libraries (**Figure 3.16. A (#4) and 3.17. A (#2)**).

In addition, the sequence CAUUC from MDA5 (+) RNA shows a slightly higher median in comparison to the control value (**Figure**

3.17. A (#7)), which was not the case for RIG-I. In regard to (-) RNA, the sequence motifs GCUAC and GCUUA from the RIG-I library slightly correlate with a higher median read coverage (**Figure 3.16. B (#5, #9)**), while (-) RNA from MDA5 samples does not contain any 5-mer, which is enriched in comparison to the expected value.

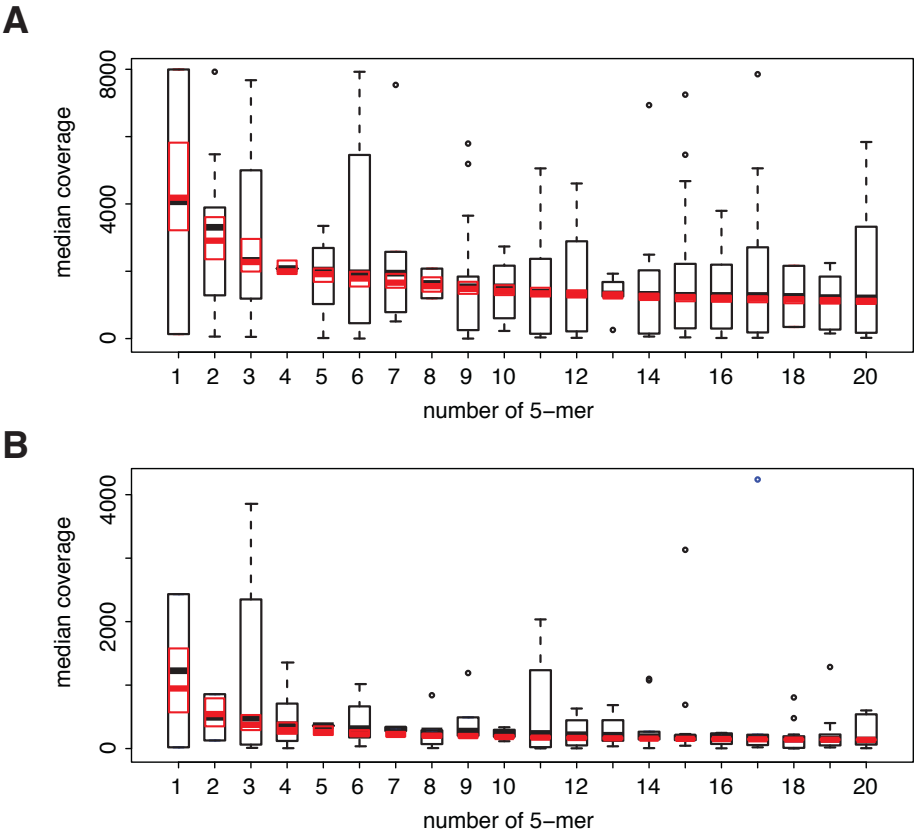


Figue 3.16. RIG-I 5-mer analysis. Boxplots represent the first 20 5-mers, which were sorted after their median read coverage in descending order. **(A)** 5-mers from the MeV (+) RNA strand are shown in black. The expected medians are in red. 5-mer number 4 (UAGUA), with an appearance of 9 (see Table 3.1), tend to be slightly enriched in comparison to the expected value. **(B)** In case of MeV (-) RNA, 5-mers number 5 (GCUAC) and number 9 (GCUUA) (see Table 3.1), which have appearances above 5 (see Table 3.1), show higher medians in comparison to the expected values (red).

Taken together, the analysis of RIG-I and MDA5 libraries indicates, that with the exception of a slight correlation with a few 5-mers rich in A/U, there is no motif, which is significantly enriched in regard to the expected values. The observed medians approach the expected ones

exactly towards the end of the plots, evidently ruling out the possibility of an RNA motif-dependent recognition by the receptor proteins.

In regard to viral evolution the lack of such a sequence motif in the viral genome is reasonable, since viruses are fast replicating species constantly changing their genome upon mutations. A conserved motif would have been lost over time and would no longer be available as a putative PAMP for RLR proteins.



Figue 3.17. MDA5 5-mer analysis. Boxplots represent the first 20 5-mers, which were sorted after their median read coverage in descending order. **(A)** 5-mers from the MeV (+) RNA strand are shown in black. The expected medians are in red. 5-mers number 2 (UAGUA) and 7 (CAUUC), which have appearances above 5 (see Table 3.2), tend to be slightly enriched in comparison to the expected values. **(B)** In case of MeV (-) RNA, no 5-mers (black) is enriched in comparison to the expected values (red).

RIG-I				
	(+) RNA		(-) RNA	
ranking of 5-mer	5-mer sequence	appearance	5-mer sequence	appearance
1	GGGCG	3	CGUAU	2
2	UUCGC	2	GACGA	2
3	AGCGU	1	CGAAA	4
4	**UAGUA**	9	CGUAC	1
5	CGCUA	2	**GCUAC**	6
6	UACGC	1	ACGCA	3
7	CGCCU	4	UUACG	2
8	CAUUC	6	UACGG	3
9	UACCA	15	**GCUUA**	6
10	AGUUA	12	GGCGA	3
11	CGUGG	3	ACGAA	5
12	GAAUA	10	GCGAG	7
13	AUAGU	18	GAGAA	7
14	CUAAG	11	ACUUA	10
15	GUAGG	11	AAUUA	11
16	AAUAG	8	CUACA	9
17	GUACG	1	CGUGC	4
18	GUAGA	11	GACAG	11
19	UGGAU	15	UAAGA	13
20	UGAAU	13	UGACG	4

Table 3.1. RIG-I 5-mer analysis. Sequences of the first 20 5-mers sorted after their medians in descending order. Bold sequences represent 5-mers with a higher median than the expected value and which appear more than 5 times in the MeV genome.

MDA5				
	(+) RNA		(-) RNA	
ranking of 5-mer	5-mer sequence	appearance	5-mer sequence	appearance
1	CGCUA	2	CGUAU	2
2	**UAGUA**	9	GACGA	2
3	GGGCG	3	CGAAA	4
4	AGCGU	1	GCUAC	6
5	CGUGG	3	CGUAC	1
6	AGUUA	12	AGCCA	3
7	**CAUUC**	6	UUACG	2
8	UUCGC	2	ACGAA	5
9	UUUGU	15	GCUUA	6
10	GCGUU	5	UACGG	3
11	GUAGG	11	CGUGC	4
12	AAUAG	8	ACUUA	10
13	UAACC	6	GGCGA	3
14	UUAAC	10	CUACA	9
15	GGAUG	22	AAUUA	11
16	GUACC	14	UACUA	9
17	UGUAG	9	GAGAA	7
18	UAGCG	2	UAUGU	11
19	UUACU	9	GUAGA	14
20	AUAGU	18	AGCUA	7

Table 3.2. MDA5 5-mer analysis. Sequences of the first 20 5-mers sorted after their medians in descending order. Bold sequences represent 5-mers with a higher median than the expected value and which appear more than 5 times in the MeV genome.

3.4.4. *In silico* analysis of RNA secondary structures

Due to its constitution, RNA forms extensive three-dimensional secondary structures, which contribute to the biological function of the molecule. RNA secondary structures are based on intrastrand base-pairing probabilities in the nucleotide sequence, which can give rise to a large variety of three-dimensional shapes (**Figure 3.18.**).

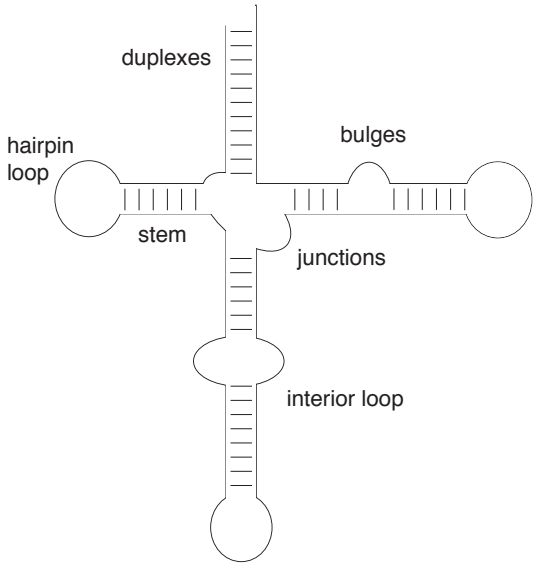


Figure 3.18. Possible RNA secondary structures.

As indicated by the analyses in sections 3.4.2. and 3.4.3., the nucleotide composition in the MeV genome by its own does not explain the mechanism of recognition by RLR proteins. Therefore, the focus was turned on the secondary structure of viral RNA. It is known that RLR proteins bind to higher-order structures of viral RNA (Pichlmair et al., 2009), which activate the antiviral signaling cascade. To study a possible correlation between the read coverage in the NGS libraries and the secondary structure of viral RNA, the MeV genome was divided into fragments of 201 nucleotides with a shifting window of 5 nucleotides. Each fragment was folded *in silico* by RNAfold (Vienna

RNA package) using the standard parameter settings. The predicted secondary structures were then analyzed for specific RNA structure features, which were set into correlation with the mean read coverage of the NGS libraries (**Figure 3.19. and 3.20.**).

The heatscatter plots represent the correlation of the corresponding structure feature with the mean coverage of the library. The color indicates how many data points overlap at one particular position. From grey to yellow there is an increase in data points. The calculated Pearson correlation coefficients reflect the relationship between the analyzed parameters and the mean coverage of RNA copy numbers.

The data indicate that the A/U content correlates with the mean coverage in the RIG-I ((+) RNA = 0.273 and (-) RNA = 0.334) (**Figure 3.19.**) and MDA5 ((+) RNA = 0.358 and (-) RNA = 0.348) (**Figure 3.20.**) libraries of positive and negative RNA strands. This moderate correlation suggests that sequences enriched by the receptor proteins tend to have a higher A/U content in comparison to sequences not recognized by RIG-I and MDA5. The result raises the possibility of a direct contribution of A/U-rich sequences to RLR activation. Furthermore, *in silico* analysis shows a correlation of 0.314 (0.262) in the case of RIG-I (MDA5) between the number of paired nucleotides and the mean coverage of (+) RNA (**Figure 3.19. A and 3.20. A**). Interestingly, this correlation negatively correlates with the G/C content of the enriched sequences (data not shown). Sequences rich in G/C generally form more stable RNA duplexes due to the formation of three hydrogen bonds. However, the data indicates that base pairs in enriched sequences are built up from A/U rather than G/C.

All other analyzed structure features have Pearson correlation coefficients below 0.2 (-0.2), which are considered as not significant (**Figure 3.19. and 3.20.**). Therefore, no further interpretations could be made on the secondary structure of putative RLR ligands.

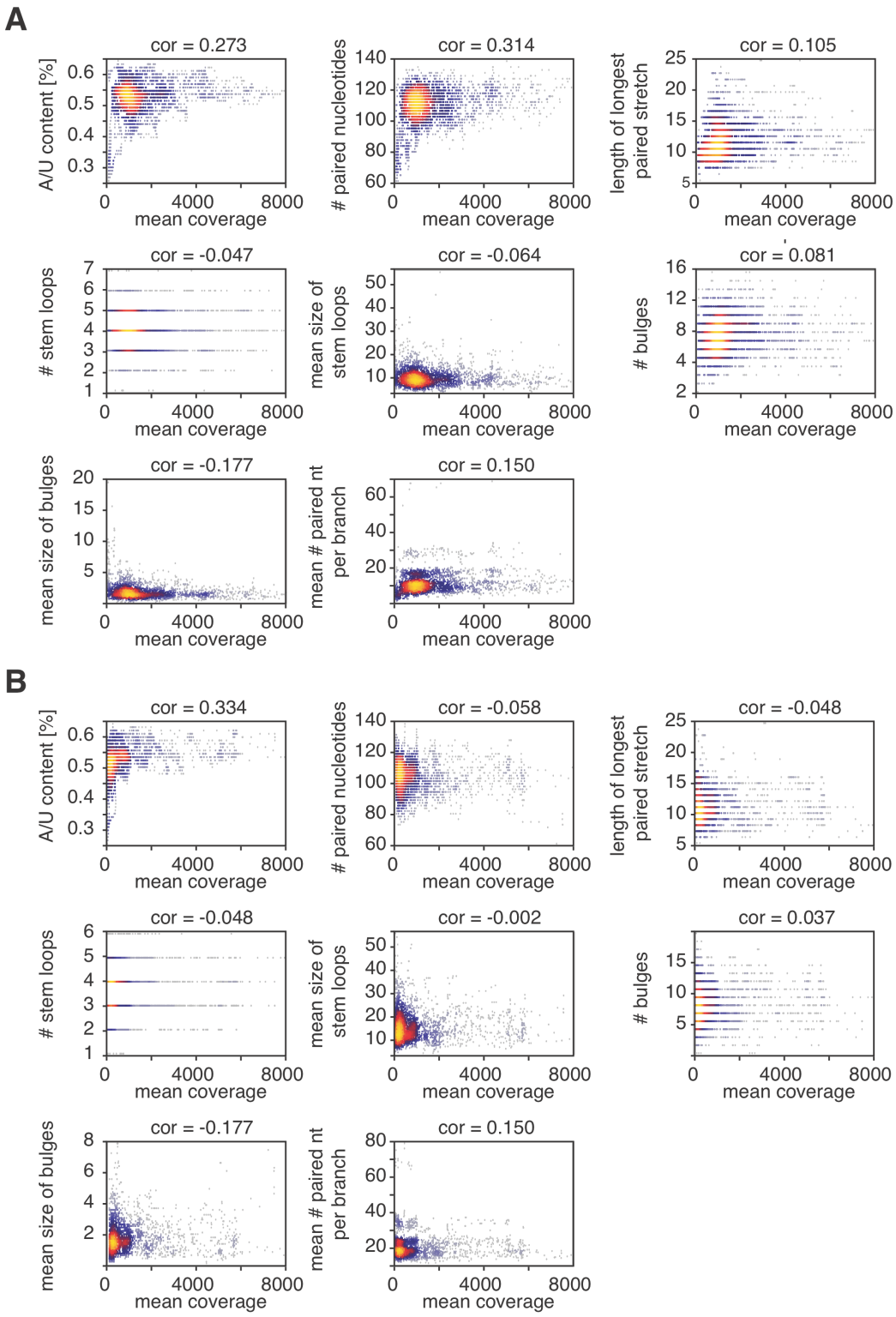


Figure 3.19. RNA secondary structure analysis of RIG-I-associated RNA. Heatscatter plots show the correlation between the analyzed parameters of RLR-associated sequences and the mean coverage of sequencing reads. **(A)** Mean read coverage of (+) RNA. **(B)** Mean read coverage of (-) RNA.

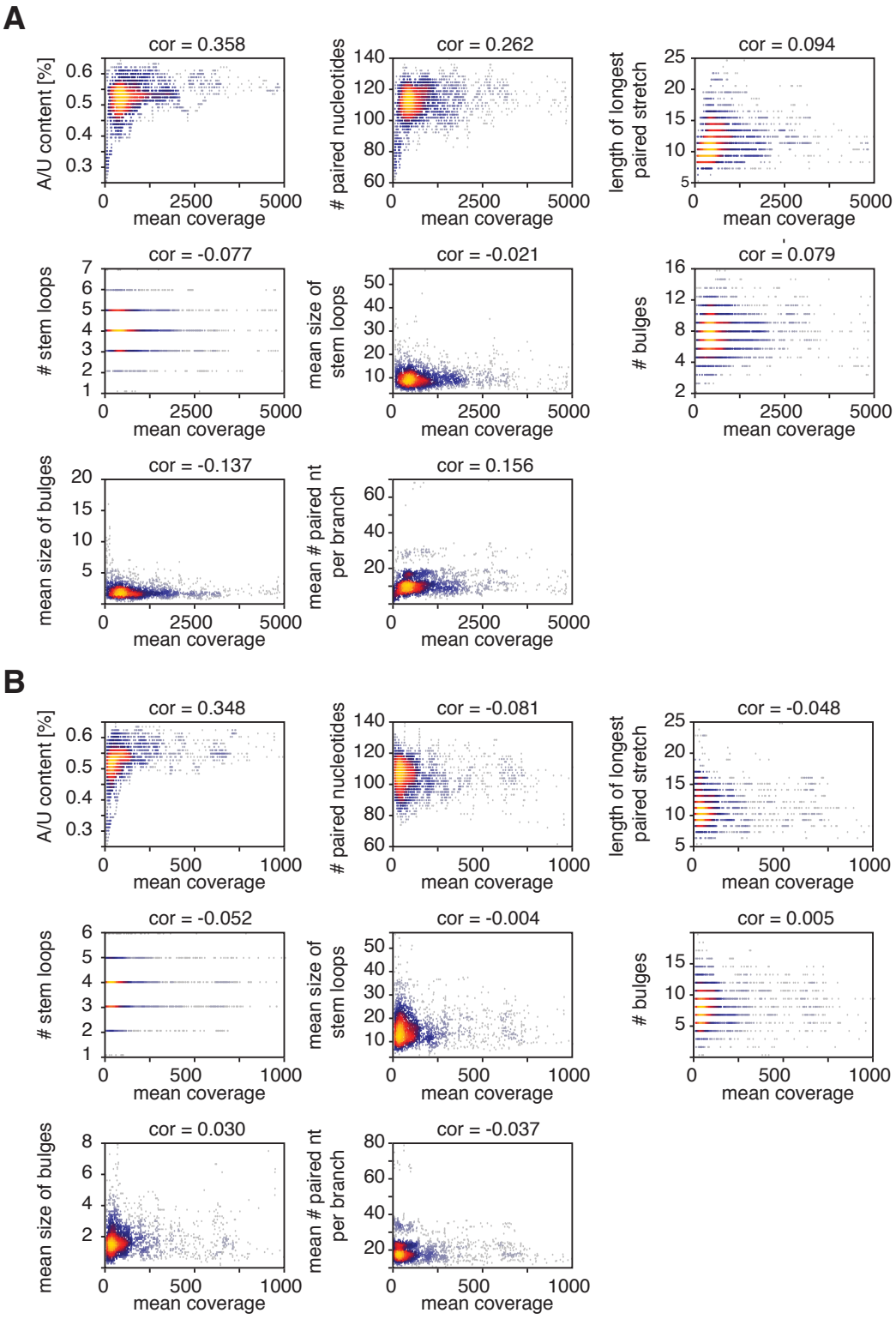


Figure 3.20. RNA secondary structure analysis of MDA5-associated RNA. Heatscatter plots show the correlation between the analyzed parameters of RLR-associated sequences and the mean coverage of sequencing reads. **(A)** Mean read coverage of (+) RNA. **(B)** Mean read coverage of (-) RNA.

Based on the literature, dsRNA like poly (I/C) is the only known agonist for MDA5 so far (Takeuchi and Akira, 2010). If this *in vitro* data would reflect the real situation in a virus-infected cell, a much higher correlation between the number of base pairings and the mean read coverage would be expected. Thus, it is speculated that double-strand regions in vRNA are not sufficient to characterize an RNA molecule as a putative PAMP for the receptor protein. All together, the data indicate that RIG-I and MDA5 recognizes A/U-rich sequences partially forming double-stranded regions in a higher-order RNA structure.

3.5. *In vitro* analysis of MeV RNA: confirmation of the NGS data

The bioinformatical analysis of NGS data suggests a possible role of the A/U composition in the activation of RLR signaling. To further proof this hypothesis, 17 *in vitro* transcripts (IVTs) were generated covering different regions of the MeV genome representing 200 nt long RNAs of positive polarity (section 2.1.5.).

As a first step, the immunostimulatory activity of RNA transcripts was validated. The results reveal that high copy numbers from NGS libraries correlate with the stimulatory activities of IVTs (**Figure 3.21. and 3.22.**). According to the immunostimulatory experiment, increased immune stimulation is observed for transcripts 9 and 12 (**Figure 3.21.**). Interestingly, transcript 9 exactly corresponds to those regions at the 5' end of the L gene, which contain the highest copy numbers of (+) RNA in the RLR libraries (**Figure 3.7.**). In general, it appears that IVTs representing regions within the L gene have higher

immunostimuatory activities in comparison to the upstream genomic segments. This is in good agreement to the NGS data.

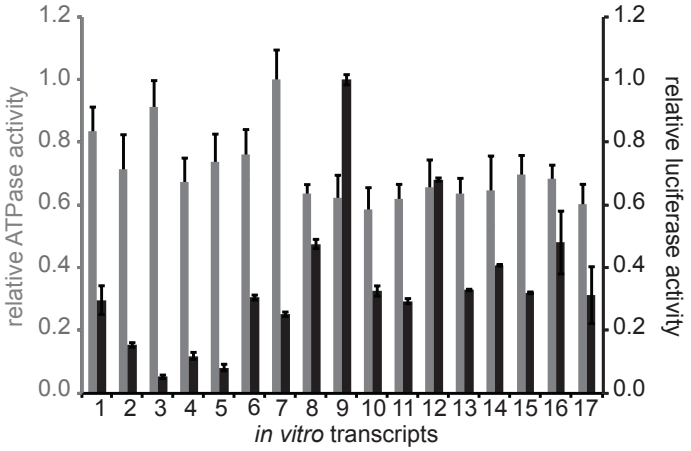


Figure 3.21. Analysis of *in vitro* transcribed RNA of the measles virus genome. Sequences were generated according to the NGS data. The transcripts were either transfected into 293T ISRE-FF/RFP cells in order to validate the immunostimulatory potential or ATPase hydrolysis experiments were performed in the presence of recombinant *m*MDA5. Comparison of luciferase activities (black) and relative ATPase activities (grey) of *in vitro* transcribed RNAs (n=3 and n=2) are shown here. Highest values were set to 1. In the case of immunoactivity experiments, highest stimulation is observed for transcripts number 9 and 12. Surprisingly, the ATP hydrolysis rate by *m*MDA5 is highest for transcript 1, 3, and 7, which show remarkably less immunostimulatory potential.

Furthermore, calculated Pearson correlations indicate that the best correlation between maximal numbers of sequencing reads and the immunostimulatory activity of RNA transcripts is observed for MDA5 sequencing data (cor = 0.526), while RIG-I and GFP samples show less correlation (cor = 0.369 and cor = 0.217) (**Figure 3.22. A**).

In order to find a possible explanation for the different immunostimulatory potentials of IVTs, the transcripts were analyzed *in silico* (RNAfold). The obtained data reveal that the immunostimulatory potential strongly correlates with the A/U content of IVTS (cor = 0.599) (**Figure 3.22. C**), which is consistent with bioinformatical data from the NGS experiment (see section 3.4.4.). Visualization of the transcripts on the Agilent bioanalyzer RNA chip

indicates that no higher-order structures due to the sequence composition are formed that might explain differences in immuostimulatory activity (data not shown). Taken together, the data allow speculating that the A/U content of sequences causes differences in immunostimulatory potentials of the tested transcripts.

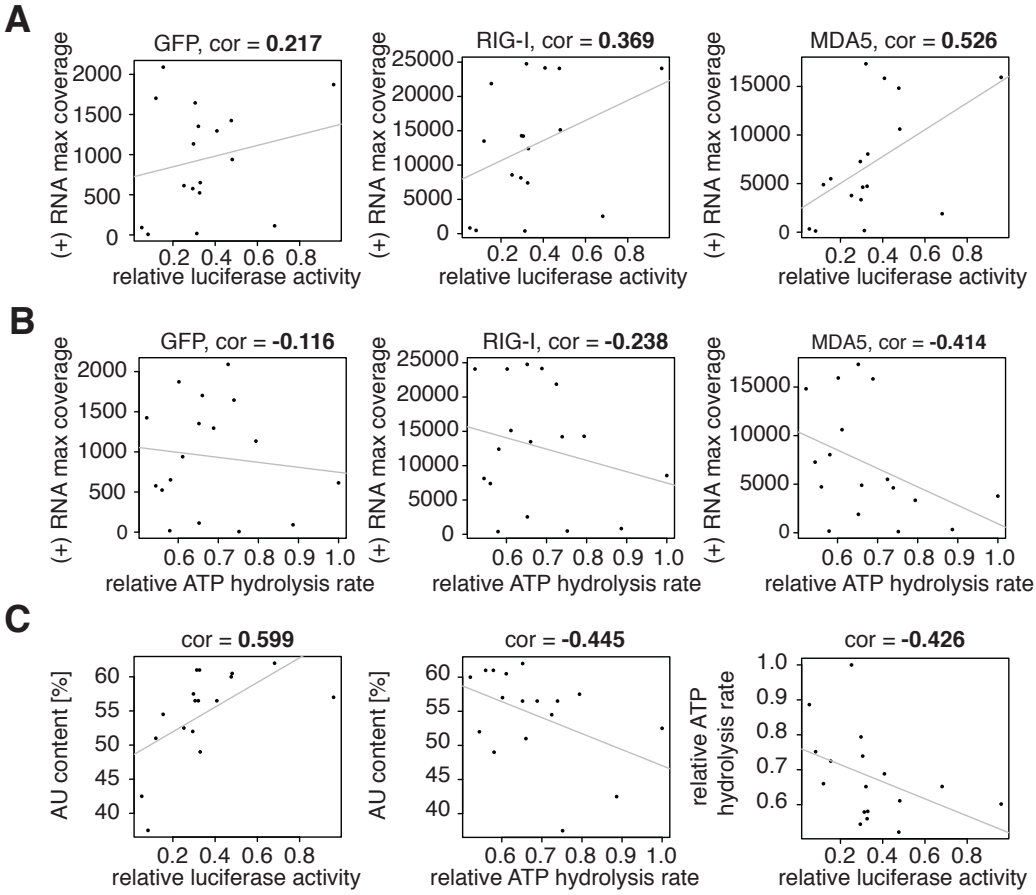


Figure 3.22. Analysis of *in vitro* transcribed RNA of the Measles virus genome. **(A)** Pearson correlation between (+) RNA maximal coverage and the relative luciferase activity. **(B)** Pearson correlation between (+) RNA maximal coverage and the relative ATPase activity. **(C)** Correlation analysis between the AU content and the luciferase or ATPase activity, and between the ATPase activity and the luciferase activity.

Finally, a possible correlation between the ATPase hydrolysis activity of MDA5 and the immunostimulatory potential of IVTs was investigated. Therefore, the ATP hydrolysis rate of recombinant mouse

MDA5 in the presence of RNA transcripts was tested (**Figure 3.21. and 3.23.**).

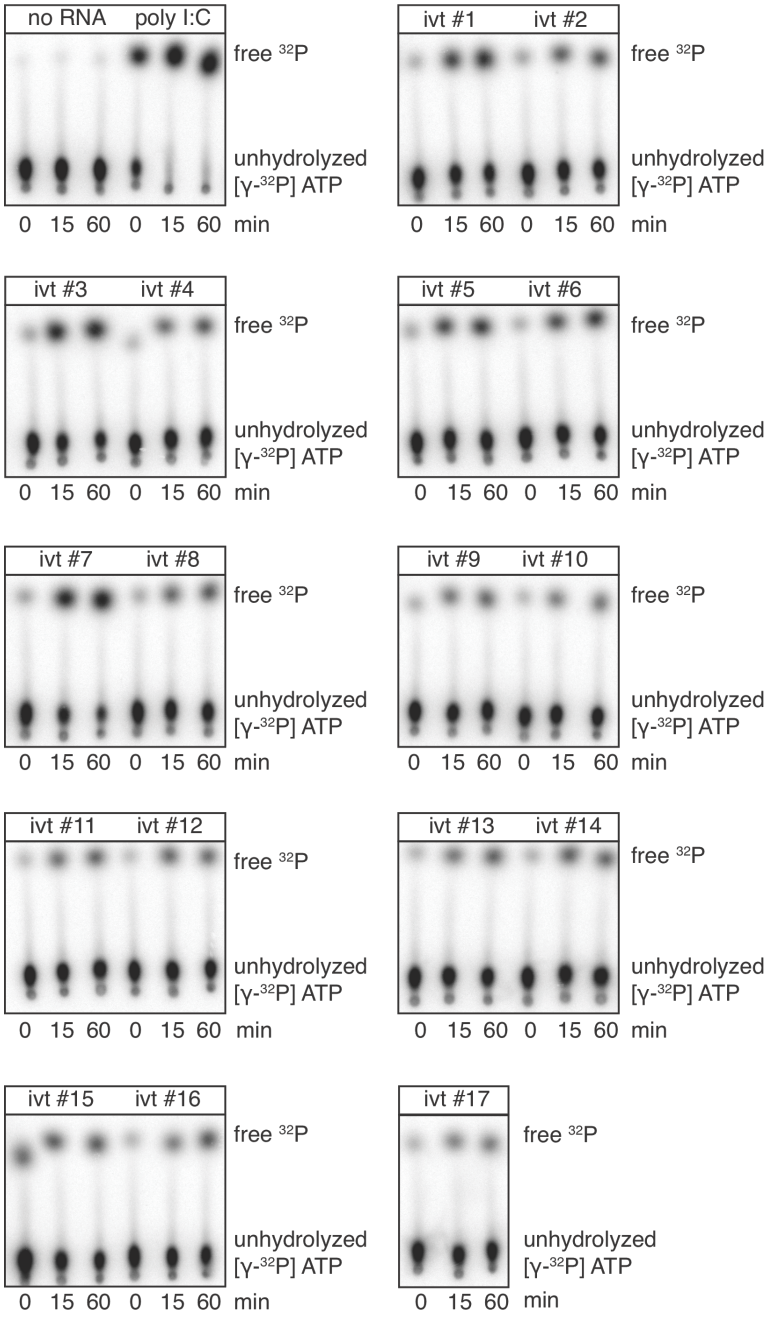


Figure 3.23. ATPase activity assay. *m*MDA5 was incubated with the respective IVT in the presence of ATP. The reaction was incubated over a time course of 1 h. Free phosphate was separated by thin layer chromatography (TLC) and visualized on a phosphorimage screen. The ATPase hydrolysis rate was determined by quantifying free phosphate at 15 min after adding ATP to the reaction mix. IVTs 1, 3, and 7 strongly stimulate ATP hydrolysis by *m*MDA5, as it is shown by the release of ^{32}P. As a control, *m*MDA5 was stimukated with poly (I/C).

According to the data, the highest ATP hydrolysis rate is observed for transcripts 1, 3, and 7 (**Figure 3.23.**), which show remarkably less immunostimulatory activity in comparison to transcript 9 and 12 (**Figure 3.21.**). The results further indicate a negative correlation between the maximum numbers of sequencing reads in the MDA5 library (cor = -0.414) and the ATP hydrolysis rate (**Figure 3.22. B**). Bioinformatical analysis of the *in vitro* data reveal that A/U-rich sequences lead to a decrease in ATP hydrolysis activity of MDA5 (cor = -0.445). Furthermore, the ATP hydrolysis rate negatively correlates with the immunostimulatory activity of RNA transcripts (cor = -0.426) (**Figure 3.22. C**). These results suggest that the ATPase hydrolysis activity of MDA5 interferes with the immunostimulatory potential of the RNA transcripts and could therefore provide a first model of RNA recognition by the receptor protein (discussed in section 4.8.).

3.6. Investigation of RNase L cleavage sites

The observed correlation between the mean coverage in RLR libraries and the A/U composition in RNA sequences, raises the possibility of a potential role for RNase L in the processing of RNA species during MeV infection. RNAse L is an endoribonuclease, which is widely expressed in a variety of mammalian tissues (Zhouh et al., 2005), existing as a monomer in the uninfected state. Several studies showed that the enzyme is activated upon viral infection by the production of 2'5'-oligoadenylate (2'5'A) molecules that bind to RNase L and lead to its dimerization (Chakrabati et al., 2011; Luthra et al., 2011). The activation of the enzyme is believed to take place in a second round of

antiviral signaling. Initial recognition of the virus leads to the activation of oligoadenylate synthetases (OAS), which produce 2'5'A molecules, required for RNase L activation (**Figure 3.24.**). Activated RNase L is suggested to cleave cytosolic RNA at A/U-rich sequences in single-stranded regions, thereby generating RNA species that could fold into potent PAMPs for RIG-I and MDA5 proteins (Malathi et al., 2010; Malathi et al., 2007; Washenber at al., 2007). This feedback loop could represent the secondary response against an invading virus.

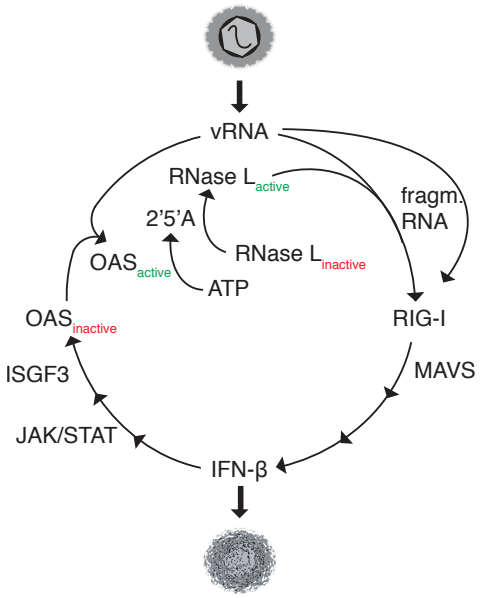


Figure 3.24. Potential model of the OAS/RNase L pathway in antiviral signaling. Initial virus sensing occurs via RIG-I, which transduces the signal to a MAVS-dependent signaling cascade leading to the expression of type I interferons. This in turn leads to the expression of interferon stimulated genes (ISG), including oligoadenylate synthetases (OASs). In the presence of dsRNA, OAS synthesizes 2'5'A molecules from ATP that bind to and activate the enzyme RNase L. Activated RNase L cleaves cytosolic RNA species at single-strand A/U-rich regions, generating RNA molecules that serve as ligands for RIG-I and MDA5. This second round of signaling enhances the antiviral response (adapted from Malathi et al., 2010).

To investigate the role of RNase L in RLR signaling, the MeV genome was divided into fragments of 500 nucleotides and folded *in silco* using RNAfold with the standard parameter settings. The folded fragments were then analyzed for potential RNase L cleavage sites in single-

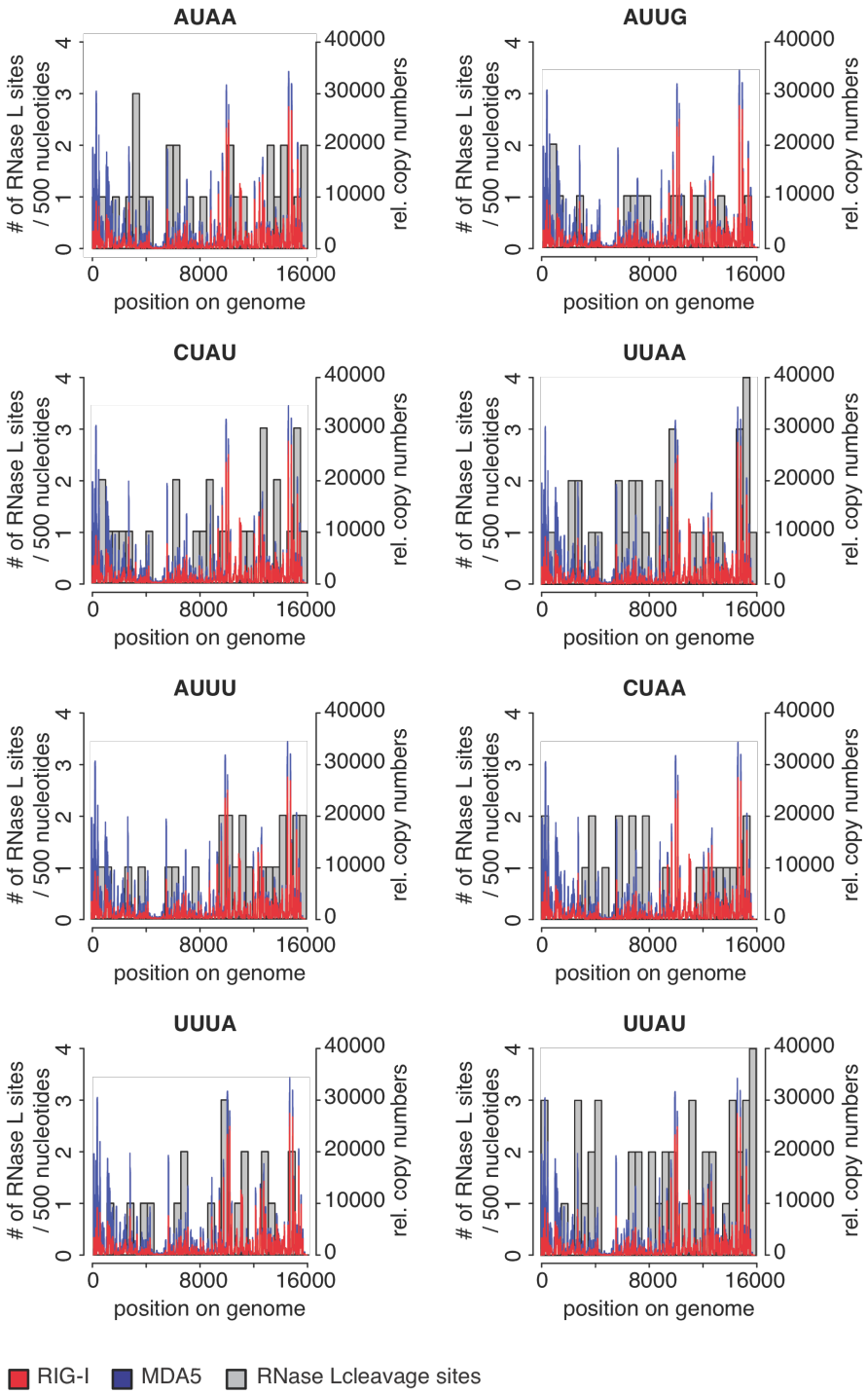


Figure 3.25. Analysis of putative RNase L cleavage sites. *Grey*: Number of respective RNA cleavage site in an *in silico* folded RNA fragment with the size of 500 nucleotide. *Red, blue*: Sequence copy numbers of RIG-I and MDA5 libraries.

stranded regions (Han et al., 2004). Comparison of cleavage sites with the NGS data (see section 3.2.1.) indicates that some of the analyzed sequences correlate with the copy numbers of RLR libraries (**Figure 3.25.**).

Especially, the sequences -UUAA-, -AUUU-, and -UUUA- nicely overlap with the NGS data. Furthermore, analysis of cleavage sites indicates that no particular cleavage site perfectly fits to the sequence distribution, suggesting that several RNase L recognition sites might contribute to the observed sequence distribution in the NGS libraries. In conclusion, the data principally allow raising the hypothesis of a potential role of RNase L in RLR ligand generation.

4. Discussion

4.1. PAR-CLIP: A promising tool to study RLR agonists?

Aim of this thesis was to characterize *in vivo* occurring RNA ligands for RLR proteins. Therefore, a method was established, which was previously described as the PAR-CLIP methodology (Hafner et al., 2010). In this approach photoactivatable nucleoside analogs are applied to live cells and a covalent linkage between the RNA ligand and the RNA-binding protein is induced upon 365 nm UV exposure. Due to the utilization of 4SU, Hafner and colleagues observed prominent transitions from T to C at the interaction sites of the RNP complexes, facilitating the identification of binding sites by mutational analysis. Compared to regular 254 nm UV crosslinking in the absence of photoactivatable nucleosides, PAR-CLIP yields higher RNA recovery levels and the usage of lower energy wavelength holds the advantage of a more specific crosslink and hence to less unspecific RNA, which is co-purified upon immunopurification of the respective protein (Hafner et al., 2010).

To validate the approach, a line of initial experiments were performed to analyze the 4SU incorporation efficiency into viral RNA and to compare the RNA recovery levels between crosslinked and non-crosslinked immunoprecipitates. An important requirement for using the method was to proof the incorporation of photoactivatable nucleoside analogs into viral RNA. Figure 3.1. displays the 4SU incorporation levels obtained by (q)PCR analysis. The data suggest that the nucleoside analog is indeed incorporated into viral RNA. However, the experimental set up does not allow a precise conclusion over the incorporation levels, since only two viral genomic segments

were investigated. Furthermore, no conclusion about the nature of analyzed RNA could be made. Although, the primers used for the experiment should give rise to (+) RNA species representing mRNA transcripts, it is possible that only one single antigenome was co-purified and analyzed by (q)PCR. Comparison of incorporation levels between viral RNA and endogenous RNA indicates that large amounts of 4SU are incorporated into host RNA, suggesting a competition between viral and endogenous RNA for the photoactivatable nucleoside analog (**Figure 3.1.**).

Next, PAR-CLIP was compared to a commonly used pull down technique (section 3.1.2.). The data show that treatment of cells with 4SU and 365 nm UV exposure leads to an improved RNA copurification in comparison to the non-crosslinking strategy. Independent experiments indicate that the recovery levels increase between 20 to 50% depending on the experimental set up (**Figure 3.2. C and 3.5.**). However, RLR-associated RNA from crosslinking samples shows a reduced immunostimulatory potential in comparison to non-crosslinked RNA (**Figure 3.2. A**). This observation raises the possibility that after protein digestion in the course of the experiment single amino acids or small polypeptide chains are still bound to nucleotides of the RNA that could interfere with the immunostimulatory effect of the molecule.

Time- and dose-dependent experiments revealed that 4SU treatment does not interfere with cell viability or viral replication (**Figure 3.3. and 3.4.**). Based on the initial findings, following parameters were chosen for the experimental set up. For infection experiments, A549 cells were treated with virus at a MOI of 1.0 in the presence of 400 μM 4SU. To ensure that viral replication and transcription is fully activated at the time of UV crosslinking, infection was allowed to occur for 24 hours. However, at this time point a decrease in endogenous

protein levels was observed indicating that cells were already seriously affected by viral infection (**Figure 3.4. B**).

4.2. Validation of immunostimulatory activity of RLR-associated RNA

To validate the immunostimulatory potential of RLR-associated RNA, 293T ISRE FF/RFP cells were transfected with co-purified RNA from RLR immunoprecipitates. The reporter gene read out allows speculating to what extend the enriched molecules are specifically recognized by RIG-I or MDA5. Upon transfection of RNA, potent RLR agonists lead to activation of RLR signaling, resulting in type I interferon production. This in turn activates the reporter gene expression via STAT/STAT signaling. However, a positive read out does not necessarily indicate a direct interaction between the receptor proteins and the transfected RNA. It is also possible that an indirect interaction, with additional steps involving unknown interaction partners, activates the reporter gene expression. Nevertheless, the reporter cell system is an established approach to investigate RLR-associated RNAs.

As shown in Figure 3.6., immunostimulatory RNA was obtained from RIG-I and MDA5 immunoprecipitates of MeV-infected cells. The induction was significant in comparison to the negative control. Consistent with the literature, RIG-I-associated RNA shows increased immunostimulatory activity in comparison to MDA5-associated RNA, suggesting a major role for RIG-I in the immune response against MeV infection (Plumet et al., 2010). Measles belongs to the family of *Paramyxoviridae*, which encode for a non-structural protein V. This

protein is known to inhibit RLR signaling at various levels including a direct interaction with the DExC/H box helicase domain of MDA5 (Motz et al., 2013). The inhibitory binding of V protein could interfere with RNA binding and hence explains the reduced immunostimulatory activity of RNA from MDA5 immunoprecipitates (**Figure 3.6.**).

In regard to other viruses, co-purified RNA from RLR proteins was not capable to induce an immune response, which highlights the differences in viral replication cycles (**Figure 3.6.**). In fact, picornaviruses like EMCV and Mengo would be the preferred viral species to study endogenous MDA5 agonists (Loo et al., 2008). However, EMCV has a higher pathogenicity in comparison to MeV, which under the given conditions lead to dramatic cell death after 24 hours. Reduction of the infection time is not recommended, since the protein levels for MDA5 would also decrease, directly affecting the levels of co-purified RNA. In the uninfected state, cells express basal levels of MDA5, which become strongly up-regulated in the course of an infection. To ensure suitable amounts of MDA5 protein at the time of *in vivo* crosslinking, viral infection was therefore allowed to occur for 24 hours. To study EMCV-derived RNA ligands for MDA5, it is possible to express recombinant MDA5 prior to virus infection. This would ensure high levels of protein already at early stages of infection and would allow a reduction in infection time. However, overexpressed MDA5 is capable to induce an antiviral state, even in the absence of viral infection (Crampton et al., 2012; unpublished data), which could interfere with viral replication.

In conclusion, the established protocol was suitable to study RLR-associated RNAs from MeV-infected cells. However, the PAR-CLIP approach has to be adapted and modified for each viral infection.

4.3. Identification of RLR-associated RNA from MeV-infected cells using Next generation sequencing (NGS)

As already mentioned in section 3.2.1., the NGS technology offers numerous advantages to study RNA-protein interaction networks in comparison to commonly used techniques. Studying RNA species by NGS does not require any prior assumptions about the molecule to be investigated. In addition, the redundancy of specific primers reduces the likelihood of introducing artifacts and allows an unbiased analysis of the molecules of interest.

In this work, the co-purified RNA from RLR and GFP samples was processed into cDNA libraries. Based on constraints of the sequencing technology, the RNA was subjected to random fragmentation prior to cDNA synthesis (see section 2.3.2.). This was necessary, since longer RNA would interfere with the efficiency of bridge amplification and cluster formation in the Illumina NGS approach. Consequently, the obtained NGS data do not allow any conclusions about the possible size of the putative RNA ligands, or about the 5' and 3' ends of co-purified RNA species.

A major problem for NGS analysis is the high background of unspecifically co-purified RNAs from immunoprecipitates. The most prominent RNA species, which are commonly found in immunoprecipitates are of ribosomal origin. High levels of background RNA would interfere with the depth of analysis and should be removed from the RNA library. Therefore, an rRNA removal kit was used to enrich RLR-associated RNA. About 16% (RIG-I) and 7% (MDA5) of the total number of sequencing reads were uniquely mapped against the MeV genome, while the rest corresponded to endogenous RNA species.

Prior to the NGS run, the cDNA libraries were analyzed on an Agilent Bioanalyzer Chip array. This quality control was done to ensure that the libraries had a sufficient concentration and were properly fragmented into shorter RNA species.

After running the libraries on the Illumina Genome Analyzer, the obtained sequences were processed in order to remove adapter sequences and reads with PHRED scores below 30. PHRED quality scores are defined as a property, which is logarithmically related to the base-calling error probabilities (Ewing and Green, 1998). A PHRED score of 30 represents a base call accuracy of 99.9%. The remaining sequences were mapped to human and viral genomes, allowing maximal one mismatch per unique read. These filter steps guaranteed high quality sequencing reads that could be further analyzed.

NGS data of RLR-associated RNA revealed that RIG-I and MDA5 bind to same regions within the L genomic segment. In addition, RIG-I, but not MDA5, binds to the 3' and the 5' ends of the MeV genome. These regions probably represent leRNA or trRNA generated in the course of replication (**Figure 3.7.**). Separation of (+) and (-) RNA furthermore indicated that MDA5 exclusively enriches MeV RNA of positive polarity. This is a remarkable difference in comparison to RIG-I, where high copy numbers are observed for both (+) and (-) RNA (**Figure 3.8.**). The data evidently rules out the possibility that MDA5 recognizes RNA duplexes of (+) and (-) RNA that might represent replication intermediates as previously described (Feng et al., 2012). The result rather suggests that MDA5 binds (+) RNA corresponding to mRNA or MeV antigenome.

Since MDA5 obviously does not bind (-) RNA, which corresponds to the MeV genome, it is conceivable that (+) RNA enriched by MDA5 most likely originate from mRNA species of the L genomic segment, rather than antigenomic sequences. The antigenomic RNA is generated

during replication in order to synthesize a template for the genomic strand. Like the genome, the antigenome is subsequently packed into nucleocapsids after replication, where they are inaccessible for cytoplasmic proteins (Blumberg et al., 1981; Bitko and Barik, 2001). Hence, genome and antigenome should not be considered as RNA ligands for RLR proteins.

In addition to recognizing (+) RNA of the L gene, RIG-I also binds to RNA of positive polarity originating from the 3' end of the MeV genome, starting at the leader region. Analysis of these sequences by qPCR indicates that the molecules could comprise leRNA-N read-through transcripts or abortive replication products that could include 5' triphosphate ends (**Figure 3.11 A**). Consequently, these RNA molecules could act as putative RNA ligands for RIG-I. In regard to RNA of negative polarity, RIG-I binds to regions comprising the trailer region and the L genomic segment. It might be possible that these RNA species represent abortive replication products of the antigenomic replication process that lead to short RNA molecules comprising 5' triphosphate ends (Cattaneo et al., 1987).

A closer look to the sequencing distribution between RLR and GFP samples revealed that the pattern of enriched sequences is similar between the samples. This raised the question whether the observed sequences in RLR libraries are indeed specific ligands for the receptor proteins. To address this question CNV data were generated that clearly show specific enrichment of RNA in the main peaks of the RLR libraries (**Figure 3.9. and 3.10.**). The similarity in the sequence enrichment pattern between control and RLR samples is a commonly observed phenomenon in NGS libraries. An explanation for this may be the introduction of a small bias due to the protocol used for cDNA synthesis. However, immunostimulatory experiments and CNV data clearly indicate that the enriched sequences specifically interact with

RLR proteins (**Figure 3.7. and 3.9. and 3.10**). Furthermore, if an unspecific enrichment of MeV RNA due to higher abundances of certain RNA species were expected, one would observe a decreasing (+) RNA gradient from the N genomic segment to the L gene (Cattaneo et al., 1987) (see section 1.3.). This was apparently not the case.

4.4. Quantitative (q)PCR analysis generally confirms NGS data

NGS data were generally confirmed by (q)PCR analysis (**Figure 3.11.**). However, the results do not reflect all RNA sequences analyzed by NGS. For (q)PCR, primer pairs were used corresponding to short areas of the 3' ends of MeV mRNA transcripts. Therefore, the analysis excludes any investigation of enriched (-) RNA. Furthermore, the different primer pairs gave rise to PCR products varying between 150 and 200 nucleotides in length, which could also influence the amplification of the respective genomic region. Differences in annealing temperatures of primers did not allow performing the (q)PCR with an optimized amplification protocol. Therefore, a direct comparison between the different genomic segments is rather unreliable. The obtained copy numbers for each genomic segment were normalized to GFP and copy numbers from RIG-I and MDA5 samples were compared. Although this superficial analysis does not allow precise conclusions, the NGS data was in general confirmed showing a sequence enrichment of (+) RNA towards the 3' end of the MeV genome (**Figure 3.11. A**).

4.5. Bioinformatical analysis of NGS data

4.5.1 Analysis of transition frequencies in the NGS data

One major advantage of PAR-CLIP is the introduction of specific transition from T to C in NGS libraries. The identification of these transitions by mutational analysis can be used to assign the direct binding sites of the RNP in RNA molecules.

To determine interaction sites between viral RNA and RLR proteins, the frequency of all possible transitions was investigated (**Figure 3.12.**). The mutational analysis indicates that T to C conversions do not occur with a significant higher frequency in comparison to other mutations. In fact, the highest frequency is observed for G to T transversions in (+) RNA of RIG-I libraries. Since crosstalk phenomena are unlikely to cause the observed transversions (see section 3.4.1.), it seems conceivable that the observed mismatches represent the correct nucleobases in the sequencing reads. It is possible that errors were introduced into the RIG-I library during cDNA preparation that could have led to the transversion from G to T. Comparison between the RLR libraries and the control shows that transitions occur at similar rates in relation to the copy numbers of sequencing reads (**Figure 3.12.**). In regard to MDA5 and GFP, conversions from T to C have a 2-fold higher frequency in comparison to the second most common transition (A to G) (**Figure 3.12. B, C**). However, this slight increase in T to C is considered to be not significant. The observed T to C transitions either occur in 100% percent of the sequencing reads (**Figure 3.13. A**) at a particular transition site, or have an occurrence below 5% (**Figure 3.13. B**). Both cases indicate that the observed T to C mutations do not result from specific crosslinks, but rather originate from wrongly annotated nucleobases or background transitions from

non-crosslinked 4SU residues. Furthermore, mismatches that occur to a few percent in the sequencing read counts could represent errors, which were introduced in the sequencing process.

Possible reasons for the lack of prominent transitions from T to C are numerous. In fact, it could be possible that due to the experimental set up the incorporation of 4SU into viral RNA is rather inefficient and endogenous RNA strongly competes for the photoactivatable nucleoside analog (**Figure 3.1.**). Endogenous RNA is synthesized in the nucleus, while MeV replication and transcription exclusively takes place in the cytosol (Gerlier and Lyles, 2011). If the nucleus were the preferred compartment for incorporation of 4SU into newly synthesized RNA, vRNA would have a significant disadvantage towards endogenous RNA, resulting in poor 4SU incoporation levels. According to Hafner and colleagues only one uridine out of 25 is substituted by 4SU, further emphasizing the inefficiency of 4SU incorporation (Hafner et al., 2010). Furthermore, it is also thinkable that 4SU is efficiently incorporated into vRNA, but the crosslinking rates are too low to result in prominent T to C transition levels. A major disadvantage of PAR-CLIP in regard to this project is that RLR proteins recognize vRNAs forming higher-ordered structures with partially dsRNA regions. The relative inefficient crosslinking of extended regions of dsRNA could lead to a further decrease in the T to C conversion rate to a level, where these mutations could not be distinguished from background mutations.

Therefore, the arising question is what caused the improved recovery levels of RLR-associated RNA from MeV-infected cells? In regard to this thesis, the applied doses of 365 nm UV light were up to 10-fold higher in comparison to the original PAR-CLIP protocol (Hafner et al., 2010). Consequently, this could result in an unspecific crosslink of unmodified nucleosides as observed for 254 nm UV crosslinking

approaches (Meisenheimer and Koch, 1997) and to an enrichment of RLR-associated RNA.

4.5.2. Does the MeV genome contains sequence motifs that are recognized by RLR proteins?

Since the mutation analysis did not allow to draw any conclusions about preferred binding sites of the receptor proteins, the nucleotide composition of enriched sequences was further investigated.

For this, the base content of the MeV genome was set into correlation with the mean read coverage of RIG-I and MDA5 libraries in a shifting window of 10 nucleotides (**Figure 3.14. and 3.15.**). In addition, all occurring 5-mers were analyzed in regard to a possible correlation with the median read coverage of the RLR libraries (**Figure 3.16. and 3.17**). Theses analyses were done in order to investigate whether the viral genome comprises any nucleotide bias or any sequence motifs that might explain the mechanism of recognition by RLR proteins.

The data led to the conclusion that according to the converging medians and notches in the represented boxplots, a possible nucleotide bias could not be the explanation for the read coverage in the RLR libraries (**Figure 3.14. and 3.15.**).

The 5-mer analyses revealed that the broad majority of investigated 5-mers do not show any enrichment in comparison to the expected values. However, based on the data, one could speculate that a slight correlation is observed for A/U-rich 5-mers, since these sequences show a slightly higher median in comparison to the expected ones (**Figure 3.16. and 3.17.**).

Taken together, the data rule out the possibility of a sequence motif in the viral genome that explains the recognition of vRNA by RLR proteins. In principle, it should be discussed to what extend a sequence

motif would make sense for the virus. A conserved motif, which is recognized by the innate immune system, would represent a crucial disadvantage for the virus to survive and successfully replicate in the host cell. Furthermore, viruses have rather imprecise replication machineries introducing frequently point mutations in their genome, thereby successfully adapting to new environmental challenges. A possible binding motif in the viral genome would have been lost over viral evolution by constantly introducing point mutation in every viral life cycle. Therefore, it seems more conceivable that recognition of viral RNA depends on other features, including secondary structures or possibly a combination of both.

4.5.3 RNA secondary structure analysis

As discussed above, a simple binding motif is not sufficient for efficient recognition by RLR proteins. To validate a possible contribution of secondary structures in vRNA to the activation of RLRs, the MeV genome was subjected to RNA secondary structure prediction.

The Pearson correlation coefficients were calculated for a variety of RNA secondary structure parameters and the mean read coverage of the RIG-I and MDA5 libraries (**Figure 3.19. and 3.20.**). The Pearson correlation coefficient is a measure for the linear relationship between two variables. A Pearson correlation coefficient of 0 means that there is no linear correlation and a correlation of ± 1 declares a perfect positive or negative linear relationship (Upton and Cook, 1996). In this work, a Pearson correlation coefficient above/below ± 0.2 was considered as reliable. The results indicate that the majority of analyzed parameters do not correlate with the mean read coverage in the RLR libraries. However, slight correlations are observed for the A/U content and the number of base pairs in the NGS samples (**Figure 3.19. and 3.20.**).

When discussing *in situ* data, one should note that the approach of RNA secondary structure prediction is highly artificial and most likely does not reflect the real conditions *in vivo*. For this analysis, the genome was divided into fragments of 201 nucleotides. These fragments were then folded using RNAfold (Vienna RNA package) and its standard parameter settings. Based on the folding, secondary structure parameters were set into correlation to the mean coverage of the NGS libraries. However, the length of 201 nucleotides for each analyzed fragment was chosen at random and the real ligand for the receptor proteins could either be shorter or longer. The length of RNA greatly influences the outcome of the *in silico* folding and therefore the data could differ depending on the size of the analyzed fragments.

Moreover, one should note that *in silico* folding of RNA only represents a limited reflection of the real condition in the cell. Folding with RNAfold relies on base pairing probabilities and the minimum free energy of the isolated RNA sequence. However, in a living cell the RNA molecule may not exist as an isolated species but constantly interacts with numerous cytosolic factors including RNPs, which could influence the stability and the shape of the RNA. Thus, the RNA secondary structure should not be considered as a rigid formation but rather as a dynamic structure that may coexist with different formations in equilibrium. Those conditions are very difficult to take into account when analyzing RNA secondary structures *in silico*. To get a more reliable conclusion about possible RNA secondary structures in RLR ligands, a combination of *in silico* data and *in vitro* experiments could be applied.

4.5.4. Does the crosslinking approach introduce an A/U bias into the NGS data?

Taken together, the bioinformatical analysis suggests that the A/U composition contributes to MeV RNA recognition by RLR proteins. The enrichment of sequences with a high A/U content by RLR immunoprecipitates raises the possibility of an experimental bias that is introduced due to the utilization of 4SU. The increased incorporation of 4SU into regions with a high A/U content would consequently lead to an increase in the crosslinking rate and to a preferred enrichment of those sequences. However, if this would be the case an enrichment of transitions from T to C should occur in A/U-rich sequences of NGS libraries. According to the mutational analysis (section 3.4.1.) increased transition frequencies from T to C are not observed, which therefore excludes the possibility of an A/U bias in the NGS data due to the crosslinking of 4SU.

Another reason for a possible A/U bias in the enriched sequences could be the increased crosslinking dose. Beside photoactivatable nucleoside analogs, wild-type nucleobases themselves have photoreactive characteristics that are exploited in commonly used 254 nm UV crosslinking approaches. The optimum excitation wavelength for uridine is about 275 nm, while the remaining nucleosides show the maximum excitation between 250 and 265 nm (Meisenheimer and Koch, 1997). The range of the UV spectra of uridine raises the possibility that due to the higher optimum of the crosslinking wavelength, uridine residues become more readily excited upon 365 nm UV exposure than other nucleobases. This would consequently lead to preferred crosslinking between uridine moieties and the amino acids of the interacting protein. Since crosslinked wild-type uridine does not result in specific transitions in the course of reverse transcription, it is

difficult to proof based on NGS data, whether this possibility should be taken into consideration.

4.6. Confirmation of *in vivo* data by *in vitro* experiments

To proof the hypothesis of a possible role for A/U-rich sequences in RLR activation and to rule out the possibility of an A/U bias in the NGS data due to the experimental set up, *in vitro* experiments were performed.

For this, 200 nucleotide long *in vitro* transcripts were generated corresponding to different regions in the MeV genome. To ensure that remaining 5' triphosphate moieties did not influence further analysis, transcripts were enzymatically double-dephosphorylated.

Immunostimulatory experiments clearly revealed a correlation between the maximum coverage of RLR libraries and the immunostimulatory potential of the tested transcript (**Figure 3.22. A**). In addition, analysis of the *in vitro* data further showed that the immunostimulatory activities of RNA molecules strongly correlate with the A/U content of transcripts (**Figure 3.22 C**). This finding is in agreement with the bioinformatical data and evidently rules out the possibility that an A/U bias due to the experimental set up results in the observed RNA enrichment by RIG-I and MDA5.

Prior to transfection into the reporter cells the transcripts were heated up to 90 °C and cooled down either rapidly or by a slowly decreasing temperature gradient. By this, the influence of possible secondary structures for the immunostimulatory potential was investigated. However, no differences in the induction of reporter gene expression were observed, indicating that the immunostimulatory effect is indeed based on the sequence composition. This was also confirmed by

bioanalyzer data showing that all transcripts exist in one conformation representing a 200 nt molecule (data not shown).

To elucidate the underlying mechanism of RLR activation by A/U-rich sequences, transcripts were incubated with recombinant MDA5 in the presence of ATP. Surprisingly, the observed ATPase activity by MDA5 negatively correlates with the immunostimulatory activities of the transcripts. Bioinformatical analysis reveals an inverse correlation between the A/U content of sequences and the ATP hydrolysis rate of MDA5 (**Figure 3.22. C**). This finding obviously indicates that the ATPase activity of the protein interferes with the immunostimulatory potential of the RNA.

4.7. A possible role for RNase L in the generation of RLR agonists

The question of the exact nature of RLR ligands from MeV- infected cells still needs to be elucidated. In fact, NGS data and *in vitro* experiments revealed that the A/U composition of viral RNA might contribute to RLR activation. However, nothing is known about the length or structure of the putative RLR agonists.

Several studies in the past suggested a role of RNase L in the generation of RNA ligands for RIG-I and MDA5 (Han et al., 2004; Malathi et al., 2007; Luthra et al., 2011). The generated agonists can either be of viral or endogenous origin and are characterized by A/U-rich cleavage sites for the enzyme. The finding of a possible correlation between A/U-rich sequences and the activation of RLR signaling, raised the question whether RNase L plays a role in the immune response against MeV infection. To address this, the genome was

searched for the most prominent RNase L cleavage sites (Han et al., 2004). Comparison with the sequencing coverage revealed that some of these recognition sites indeed nicely overlap with the enriched sequences of RLR immunoprecipitates (**Figure 3.25.**), suggesting a possible role of RNase L in the generation of RNA ligands for the receptor proteins. A recent work on parainfluenza virus 5 (PIV5) further supports this theory (Luthra et al., 2011). Like Measles, PIV5 is a nonsegmented, negative-strand RNA virus belonging to the family of *paramyxoviridae* (Lamb and Kolakofsky, 2001). Robert Silverman and colleagues provided evidence that RNase L is involved in the cleavage of distinct region of the PIV5 L gene, leading to RNase L cleavage products that are capable to activate IFN expression in an MDA5-dependent manner (Luthra et al., 2011). This data is in good agreement with findings in this thesis and supports a possible role of an RNaseL-MDA5-dependent pathway in the immune response of MeV infection.

However, the hypothesis is speculative and further experiments need to be done in order to draw reliable conclusion in regard to RNAse L and its role in RLR agonist formation.

4.8. RLR-associated RNA is rich in A/U: A current working model

The results obtained by NGS experiments, *in silico* and *in vitro* approaches finally allows to set up a model to explain the possible mechanism of MeV RNA recognition by RLR proteins.

Based on the sequence distribution in NGS libraries (**Figure 3.7.**), the RNA species recognized by RLR proteins could be divided into two

classes. The first class would comprise RNA molecules originating from the 5' triphosphate ends of the genome or antigenome. These molecules could be generated in the course of read-trough transcription and abortive replication (Cattaneo et al., 1987; Plumet et al., 2005) and could therefore represent preferred ligands for RIG-I, as already shown in previous work (Baum et al., 2010). The second class of RNA molecules could be recognized by both receptor proteins. The data suggest that recognition of these RNAs might occur through the A/U composition of sequences, which would define a new basis for recognition of non-self RNA (Schnell et al., 2012). In this context, a possible role of RNase L in RLR agonist generation still needs to be investigated.

Furthermore, the data indicate that the main contribution of MeV recognition is taken by RIG-I, while MDA5 seems to play a minor role. It is possible that RIG-I initially recognizes leader-N read-through transcripts or abortive replication products containing 5' tri-phosphate ends, which then lead to the activation of the signaling cascade. In a second round of recognition, up-regulated RIG-I and MDA5 recognize viral transcripts that are rich in A/U. Especially in regard to MDA5, the data provide evidence that the preferred ligand for the protein is an RNA comprising A/U-rich sequences. Furthermore, the inverse correlation between the immunostimulatory potential of RNAs and the ATP hydrolysis rate by MDA5 (cor = -0.445; **Figure 3.22 C**) raises the possibility that the ATPase activity might not be necessary for, or even interferes with the immunoactivity of RNA ligands. Although this observation disagrees with recent findings about the role of ATP hydrolysis in RIG-I oligomerization on 5' triphosphate dsRNA (Patel et al., 2013; Peisley et al., 2013), it seems conceivable that MDA5 and RIG-I markedly differ in their mechanical activation and the mode of ATP hydrolysis. The presented data is supported by results suggesting

that MDA5 filament formation is abrogated in an ATP-sensitive manner. By electron microscopy (EM) analysis it was shown that MDA5 filaments disassemble in the presence of ATP indicating that ATP hydrolysis triggers the translocation of the protein along the dsRNA molecule, thereby interfering with downstream signaling (Peisley et al., 2011; Berke et al., 2012).

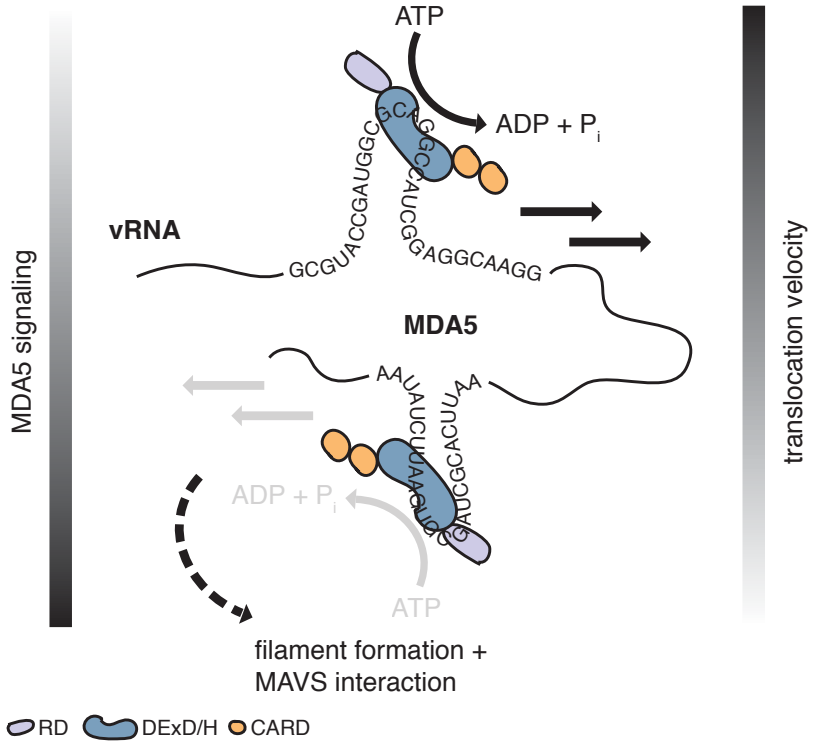


Figure 4.1. Proposed model of MDA5 activation by AU-rich RNA molecules. Upon binding to RNA, MDA5 translocates along the molecule in an ATP-dependent manner. As soon as the protein reaches A/U-rich double-stranded stretches, MDA5 binds with higher affinity to these regions. This results in a reduced ATP turn over and a slow down in translocation velocity. Additional MDA5 molecules are recruited that lead to MDA5 filament fomation and to downstream signaling via MAVS.

In light of the current data in the literature, one can hypothesize that binding of RNA to MDA5 activates the ATPase activity of the helicase domain, leading to a translocation of the protein along the RNA molecule. Upon reaching A/U-rich dsRNA stretches, MDA5 could bind

with a higher affinity to these regions, leading to a decrease in translocation velocity. This could allow a stronger interaction between the RNA and MDA5, which finally triggers a higher immunostimulatory signal via MDA5 filament formation (**Figure 4.1.**). Although, the exact mechanism of recognition and the role of RNase L in this complex arrangement of immune factors remains to be analyzed in more detail, the work provides a first insight into the molecular basis of RNA sensing by MDA5 and is a good starting point for future investigation.

4.9. Improvement of RNA-protein crosslinks using alternative methods

Although the PAR-CLIP technology yields improved RNA recovery levels from RLR immunoprecipitates, the lack of prominent T to C transitions highlights the disadvantages of the method in regard to the investigation of RLR-associated RNAs from virus-infected cells.

To improve the specific co-purification of putative RLR agonists 254 nm UV crosslinking could be tested. The higher energy wavelength could result in a more efficient crosslink between dsRNA regions and the receptor proteins in comparison to the 365 nm UV strategy. Methods involving such a crosslinking step include CLIP (Ule et al., 2003; Ule et al, 2005), HITS-CLIP (Licatalosi et al., 2008; Chi et al 2009), and iCLIP (Konig et al., 2010 and 2011).

In regard to the inefficiency of dsRNA crosslinking upon UV irradiation an alternative approach might be the utilization of methylene blue. Methylene blue is a dye, which intercalates into dsRNA and mediates RNA-protein crosslinking upon exposure to

visible light (Liu et al., 1996). However, the available data represent *in vitro* experiments and so far little is known about the uptake of methylene blue and its crosslinking efficiency in living cells.

Another possible procedure, is based on formaldehyde crosslinking, which could result in a much more efficient crosslinking between the RNA and the receptor protein. However, this strategy is rather unspecific, since crosslinking is not restricted to direct interaction partners but also occurs between unspecific molecules in close proximity. Furthermore, this *in vitro* method requires the co-purification of RNA in stable RNA-protein complexes. Especially in regard to MDA5 and its transient interaction with its putative RNA ligands, the approach could be limited and could at best serve as a reference experiment. Currently, UV crosslinking techniques are under way (Ule lab) that are specifically developed to study dsRNA ligands. These new approaches might facilitate the analysis of RLR-associated RNAs and circumvent the limitations of conventional UV-crosslinking protocols as discussed in this thesis.

5. Summary

The innate immune system provides the first line of defense against invading pathogens. Innate immunity relies on germline-encoded pathogen-recognition receptors (PRRs) that sense invariant molecular patterns of microbial origin, thereby triggering a fast but non-specific immune response. Cytosolic sensing of RNA viruses is mainly performed by the RIG-I-like receptors (RLRs) RIG-I, MDA5, and LGP2, which are helicase-like proteins sharing structural similarities. RLRs recognize viral infection to initiate and modulate an antiviral immune response. Upon binding of RNA molecules, which are absent from the uninfected host, the sensor proteins activate a MAVS-dependent signaling cascade, which leads to the expression of type I interferons (IFNs) and proinflammatory cytokines.

One remaining question in the field is, how RLR proteins are able to distinguish between RNA from the uninfected and infected cell. In the past decade extensive efforts were made to characterize the exact nature of RNA agonists of RLR proteins. *In vitro* and transfection studies identified 5' triphosphate containing double-strand (ds)RNAs as preferred ligands for RIG-I. In case of the RIG-I paralogs MDA5 and LGP2, the nature of RNA ligands is less well understood. Several studies suggest MDA5 as a simple dsRNA sensor in the cytosol. However, the experimental evidence for this is not conclusive, because the data mainly rely on *in vitro* transfection experiments that bring only limited and potentially misleading information about the physiological MDA5 agonist. To elucidate the exact nature and origin of RLR ligands, it is important to isolate and investigate RLR-associated RNA from virus-infected cells. A major challenge in regard to examining MDA5 agonists is the apparently transient interaction

between the protein and its RNA ligands. To improve the stability of interaction between RNA and MDA5, a UV crosslinking approach was established, in which virus-infected cells are treated with the photactivitable nucleoside analog 4 thiouridine (4SU). Upon 365 nm UV exposure of living cells, a covalent linkage between the 4SU-labeled RNA and the receptor protein is induced. This crosslink results in a higher RNA recovery from immunoprecipitates. Based on Next generation sequencing, bioinformatical analysis, and *in vitro* approaches, a correlation between the A/U-composition of viral RNA and its ability to induce an MDA5-dependent immune response was observed. The results suggest that RIG-I and MDA5 recognize A/U-rich RNA species, originating from the mRNA of the measles virus L gene. Interestingly, the immunostimulatory potential of A/U-rich RNA interferes with the ability to stimulate ATP hydrolysis by MDA5. Together with data in the literature, the results suggest that binding of RNA to MDA5 activates the ATPase activity of the helicase domain leading to translocation of the protein along the RNA molecule. The negative correlation between A/U-rich sequences and the ATP hydrolysis rate indicates that A/U-rich RNA could bind to MDA5 with a higher affinity than sequences rich in GC. This would lead to a decrease in translocation velocity and consequently to a stronger interaction between AU-rich RNA and MDA5, which then induces a higher immunostimulatory signal. At this stage, further analysis is required to confirm the model. Nevertheless, the work of the thesis provides a first insight into the molecular basis of RNA sensing by MDA5 and is a good starting point for future investigation of questions how recognition of viruses by RLRs is achieved.

6. References

Ablasser A., Bauernfeind F., Hartmann G., Latz E., Fitzgerald K.A., Hornung V. (2009). "RIG-I-dependent sensing of poly(dA:dT) through the induction of an RNA polymerase III-transcribed RNA intermediate." *Nat. Immunol.* **10**: 1065-1073.

Alberts B., Johnson, A., Walter, P., Raff, M., Roberts, K. and Lewis, J. (2002). "Molecular Biology of the Cell", Taylor & Francis.

Allen I.C., Scull M.A., Moore C.B., Holl E.K., McElvania-TeKippe E., Taxman D.J., Guthrie E.H., Pickles R.J., Ting J.P. (2009). "The NLRP3 inflammasome mediates in vivo innate immunity to influenza A virus through recognition of viral RNA." *Immunity* **30**: 556-565.

Atreya P. L., Peeples M. E., Collins P. L. (1998). "The NS1 protein of human respiratory syncytial virus is a potent inhibitor of minigenome transcription and RNA replication." *J. Virol.* **72**:1452-1461.

Barber G.N. (2011). "Cytoplasmic DNA innate immune pathways." *Immunol. Rev.* **243**: 99-108.

Baron M. D., Barrett T. (2000). "Rinderpest viruses lacking the C and V proteins show specific defects in growth and transcription of viral RNAs." *J. Virol.* **74**:2603-2611.

Baum A., Sachidanandam R., Garcìa-Sastre A. (2010). "Preference of RIG-I for short viral RNA molecule in infected cells revealed by next-generation sequencing." *Proc. Natl. Acad. Sci. USA* **107**: 16303-16308.

Berke I.C., Yu X., Modis Y., Egelmann E.H. (2012). "MDA5 assembles into a polar helical filament on dsRNA." *Proc. Natl. Acad. Sci. USA* **109**: 18437-18441.

Berke I.C., Modis Y. (2012). "MDA5 cooperatively forms dimers and ATP-sensitive filaments upon binding to double-stranded RNA." *EMBO J.* **31**: 1714-1726.

Bitko V., Barik S. (2001). "Phenotypic silencing of cytoplasmic genes using sequence-specific double-stranded short interfering RNA and its application in the reverse gentics of wild type negative-strand RNA viruses." *BMC Microbiol.* **1**: 34.

Blumberg B.M., Leppert M., Kolakofsky D. (1981). "Interaction of VSV leader RNA and nucleocapsid protein may control VSV genome replication." *Cell* **23**: 837-845.

Boxer E. L., Nanda S. K., Baron M. D. (2009). "The rinderpest virus non-structural C protein blocks the induction of type 1 interferon." *Virology* **385**: 134-142.

Brzozka K., Finke S., Conzelmann K. K. (2006). "Inhibition of interferon signaling by rabies virus phosphoprotein P: activation-dependent binding of STAT1 and STAT2." *J. Virol.* **80**: 2675-2683.

Bruns A.M., Horvath C.M. (2012). "Activation of RIG-I-like receptor signal transduction." *Crit Rev. Biochem. Mol. Biol.* **47**: 194-206.

Cadd T., Garcin D., Tapparel C., Itoh M., Homma M., Roux L., Curran J., Kolakofsky D. (1996). "The Sendai paramyxovirus accessory C proteins inhibit viral genome amplification in a promoter-specific fashion." *J. Virol.* **70**: 5067-5074.

Caignard G., Guerbois M., Labernardière J.L., Jacob Y., Jones L.M., Wild F., Tangy F., Vidalain P.O. (2007). "Measles virus V protein blocks Jak1-mediated phosphorylation of STAT1 to escape IFN-alpha/beta signaling." *Virology* **368**: 351-362**.**

Cárdenas W.B., Loo Y.M., Gale M. Jr., Hartman A.L., Kimberlin C.R., Martínez-Sobrido L., Saphire E.O., Basler C.F. (2006). "Ebola virus VP35 protein binds double-stranded RNA and inhibits alpha/beta interferon production induced by RIG-I signaling." *J. Virol.* **80**: 5168-5178.

Cattaneo R., Rebmann G., Schmid A., Baczko K., ter Meulen V., Billeter M.A. (1987). "Altered transcription of a defective measles virus genome derived from a diseased human brain." *EMBO J.* **6**: 681-688.

Chambers J.M., Cleveland W.S., Kleiner B., Tukey P.A. (1983). "Graphical Methods for Data Analysis." Wadsworth & Brooks/Cole.

Chi S.W., Zang J.B., Mele A., Darnell R.B. (2009). "Argonaute HITS-CLIP decodes microRNA-mRNA interaction maps." *Nature* **460**: 479-486.

Childs K., Stock N., Ross C., Andrejeva J., Hilton L., Skinner M., Randall R., Goodbourn S. (2007). "MDA-5, but not RIG-I, is a common target for paramyxovirus V proteins." **Virology** 359: 190-200.

Chui Y.H., MacMillan J.B., Chen Z.J. (2009). "RNA polymerase III detects cytosolic DNA and induces type I interferon's through RIG-I pathway." *Cell* **138**: 576-591.

Civril F., Bennett M., Moldt M., Deimling T., Witte G., Schiesser S., Carell T., Hopfner K.P. (2011). "The RIG-I ATPase domain structure reveals insights into ATP-dependent antiviral signalling." *EMBO Rep.* **12**: 1127-1134.

Civril F., Deimling T., de Oliveira Mann C.C., Ablasser A., Moldt M., Witte G., Hornung V. Hopfner K.P. (2013). "Structural mechanism of cytosolic DNA sensing by cGAS." *Nature* **498**: 332-337.

Crampton S.P., Deane J.A., Feigenbaum L., Bolland S. (2012). "Ifih1 gene dose effect reveals MDA5-mediated chronic type I IFN gene signature, viral resistance, and accelerated autoimmunity." *J. Immunol.* **188**: 1451-1459.

Cui S., S., Eisneaecher K., Kirchhofer A., Brzozka K., Lammens A., Lammens K., Fujita T., Conzelmann K.K., Krug A., Hopfner K.P. (2008). "The C-terminal regulatory domain is the RNA 5'triphosphate sensor of RIG-I." *Mol. Cell* **29**: 169-179.

Dinarello C.A. (1996). "Biological basis for interleukin-1 in desease." *Blood* **87**: 2095-2147.

Dinarello C.A. (1998). "Interleukin-1 beta, interleukin-18, and the interleukin-1 beta converting enzyme." *Ann. NY. Acad. Sci.* **856**: 1-11.

Dixit E., Boulant S., Zhang Y., Lee A.S., Odendall C., Shum B., Hacohen N., Chen Z.J., Whelan S.P., Fransen M., Nibert M.L., Superti-Furga G., Kagan J.C. (2010). "Peroxisomes are signaling platforms for antiviral innate immunity." *Cell* **141**: 668-681.

Elinav E., Strowing T., Kau A.L., Henao-Mejia J., Thaiss C.A., Booth C.J., Peaper D.R., Bertin J., Eisenbarth S.C., Gordon J.L., Flavell R.A. (2010). "NLRP6 inflammasome regulates colonic microbial ecology and risk for colitis." *Cell* **145**: 745-757.

Ewing B., Green P. (1998). "Base-calling of automated sequencer traces using phred. II. Error probabilities." *Genome Res.* **8**: 186-194.

Feng Q., Hato S.V., Langereis M.A., Zoll J., Virgen-Slane R., Peisley A., Hur S., Semler B.L., van Rij R.P., van Kuppeveld F.J. (2012). "MDA5 detectes the double-stranded RNA replicative form in picornavirus-infected cells." *Cell Rep.* **2**: 1187-1196.

Fernandes-Alnemri T., Yu J.W., Datta P., Wu J., Alnemri E.S. (2009). "AIM2 activates the inflammasome and cell death in response to cytoplasmic DNA." *Nature* **458**: 509-513.

Ferrage F., Duuta K., Nistal-Villan E., Patel J.R., Sànchez-Aparicio M.T. De Ioannes P., Buku A., Aseguinolaza G.G., Garcia-Sastre A., Aggarwal A.K. (2012). "Structure and dynamics of the second CARD of human RIG-I provide mechanistic insights into regulation of RIG-I activation." *Structure* **20**: 2048-2061.

Fujita T., Onoguchi K., Onomoto K., Hirai R., Yoneyama M. (2007). "Triggering antiviral response by RIG-I-related RNA helicases." *Biochimie* **89**: 754-760.

Gack M.U., Shin Y.C., Joo C.H., Urano T., Liang C., Sun L., Takeuchi O., Akira S., Chen Z., Inoue S., Jung J.U. (2007) "TRIM25 RING-finger E3 ubiquitin ligase is essential for RIG-I-mediated antiviral activity." *Nature* **446**: 916-920.

Gao, P., Ascano, M., Wu, Y., Barchet, W., Gaffney, B.L., Zillinger, T., Serganov, A.A., Liu, Y., Jones, R.A., and Hartmann, G., Tuschl T., Patel D.J. (2013). "Cyclic [G(2′,5′)pA(3′,5′)p] is the metazoan second messenger produced by DNA-activated cyclic GMP-AMP synthase." *Cell* **153**: 1094-1107.

Gerlier D., Lyles D.S. (2011). "Interplay between innate immunity and negative-strand RNA viruses: towards a rational model." *Microbiol. Mol. Biol. Rev.* **75**: 468-490.

Gitlin L., Benoit L., Song C., Cella M., Gilfillan S., Holtzman M.J., Colonna M. (2010). "Melanoma differentiation-associated gene 5 (MDA5) is involved in the innate immune response to Paramyxoviridae infection in vivo." *PLoS Pathog.* **6**: e1000734.

Haecker H., Redecke V., Blagoev B., Kratschmarova I., Hsu L.C. Wang G.G., Kamps M.P., Raz E., Wagner H., Haecker G., Mann M., Karin M. (2006). "Specificity in Toll-like receptor signaling through distinct effector functions of TRAF3 and TRAF6." *Nature* **439**: 204-207.

Hafner M., Landthaler M., Burger L., Khorshid M. Hausser J., Berninger P., Rothballer A., Ascano M. Jr., Jungkamp A.C., Munschauer M., Ulrich A., Wardle G.S., Dewell S., Zavolan M., Tuschl T. (2010). "Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP." *Cell* **141**: 129-149.

Han J.Q., Wroblewski G., Xu Z., Silverman R.H., Barton D.J. (2004) "Sensitivity of hepatitis C virus RNA to the antiviral enzyme ribonuclease L is determined by a subset of efficient cleavage sites." *J. Interferon Cytokine* **24**: 664-676.

Harton J.A., Linhoff M.W., Zhang J., Ting J.P. (2002). "Cutting edge: CATERPILLER: a large family of mammalian genes containing CARD, pyrin, nucleotide-binding, and leucine-rich repeat domains." *J. Immunol.* **169**: 4088-4093.

Hembach K. (2012). "Development of an analysis pipeline for high-throughput genome-wide PAR-CLIP measurements of virus-infected cells." Bachelor thesis in Protein Bioinformatics & Computational Biology.

Hemmi H., Takeuchi O., Kawai T., Kaisho T., Sato S., Sanjo H., Matsumoto M., Hoshino K., Wagner H., Takeda K., Akira S. (2000). "*A* Toll-like receptor recognizes bacterial DNA." *Nature* **408**: 740-745.

Hoebe K., Janssen E., Beutler B. (2004). "The interface between innate and adaptive immunity." *Nat. Immunol.* **5**: 971-974.

Hou F., Sun L., Zheng H., Skaug B., Jiang Q.X., Chen Z.J. (2011). "MAVS forms functional prion-like aggregates to activate and propagate antiviral innate immune responses." *Cell* **146**: 448-461.

Hiscott J., Lacoste J., Lin R. (2006). "Recruitment of an interferon molecular signaling complex to the mitochondrial membrane: disruption by hepatitis C virus NS3-4A protease." *Biochem. Pharmacol.* **72**: 1477-1484.

Hornung V., Ellegast J., Kim S., Brzózka K., Jung A., Kato H., Poeck H., Akira S., Conzelmann K.K., Schlee M., Endres S., Hartmann G. (2006). "5'-triphosphate RNA is the ligand for RIG-I." *Science* **314**: 994-997.

Hornung V., Ablasser A., Charrel-Dennis M., Bauernfeind F., Horvath G., Caffrey D.R., Latz E., Fitzgerald K.A. (2009). "AIM2 recognizes cytosolic dsDNA and forms a caspase-1-activating inflammasome with ASC." *Nature* **458**: 514-518.

Ichinohe T., Lee H.K., Ogura Y., Flavell R., Iwasaki A. (2009). "Inflammasome recognition of influenza virus is essential for adaptive immune response." *J. Exp. Med.* **206**: 79-87.

Ikegame S., Takeda M., Ohno S., Nakatsu Y., Nakanishi Y., Yanagi Y. (2010). "Both RIG-I and MDA5 RNA helicases contribute to the induction of alpha/beta interferon in measles virus-infected human cells." *J. Virol.* **84**: 372-379.

Inohara N., Nunez G. (2001). "The NOD: a signaling module that regulates apoptosis and host defense against pathogens." *Oncogene* **20**: 6473-6481.

Ishikawa H., Barber G.N. (2008). "STING is an endoplasmic reticulum adaptor that facilitates innate immune signalling." *Nature* **455**: 674-678.

Iwasaki A., Medzhitov R. (2004). "Toll-like receptor control of the adaptive immune responses." *Nat. Immunol.* **5**: 987-995.

Iwasaki A., Medzhitov R. (2010). "Regulation of adaptive immunity by the innante immune system." *Science* **327**: 291-295.

Janeway C.A. Jr., Medzhitov R. (2002). "Innate Immune Recognition." *Annu. Rev. Immunol.* **20**: 197-216.

Jin L., Waterman P.M., Jonscher K.R., Short C.M., Reisdorph N.A., Cambier J.C. (2008). "MPYS, a novel membrane tetraspanner, is associated with major histocompatibility complex class II and mediates transduction of apoptotic signals." *Mol. Cell. Biol.* **28**: 5014-5026.

Jin T., Perry A., Jiang J., Smith P., Curry J.A., Unterholzner L., Jiang Z., Horvath G., Rathinam V.A., Johnstone R.W., Hornung V., Latz E., Bowie A.G., Fitzgerald K.A., Xiao T.S. (2012). "Structures of the HIN domain:DNA complexes reveal ligand binding and activation mechanisms of the AIM2 inflammasome and IFI16 receptor." *Immunity* **36:** 561-571.

Kang D.C., Gopalkrishnan R.V., Lin L., Randolph A., Valerie K., Pestka S., Fisher P.B. (2004). "Expression analysis and genomic characterization of human melanoma differentiation associated gene-5, mda-5: a novel type I interferon-responsive apoptosis-inducing gene." *Oncogene* **23**: 1789-1800.

Kanneganti T.D., Ozören N., Body-Malapel M., Amer A., Park J.H., Franchi L., Whitfield J., Barchet W., Colonna M., Vandenabeele P., Bertin J., Coyle A., Grant E.P., Akira S., Núñez G. (2006). "Bacterial RNA and small antiviral compounds activate caspase-1 through cryopyrin/Nalp3." *Nature* **440**: 233-236.

Kato H., Takeuchi O., Mikamo-Satoh E., Hirai R., Kawai T., Matsushita K., Hiiragi A., Dermody T.S., Fujita T., Akira S. (2008). "Length-dependent recognition of double-stranded ribonucleic acids by retinoic acid-inducible gene-I and melanoma differentiation-associated gene 5." *J. Exp. Med.* **205**: 1601-1610.

Kato H., Takahasi K., Fujita T. (2011). "RIG-I-like receptors: Cytoplasmic sensors for non-self RNA." *Immunol. Rev.* 243: 91-98.

Kawai T., Akira S. (2010). "The role of pattern-recognition receptors in innate immunity: Update on Toll-like receptors." *Nat. Immunol.* **11**: 373-384.

Kawai T., Akira S. (2011). "Toll-like receptors and their crosstalk with other innate receptors in infection and immunity." *Immuntiy* **34**: 637-650.

Konig J., Zarnack K., Curk T., Gregor R., Kayikci M., Zupan B., Luscombe N.M., Ule J. (2010). "iCLIP reveals the function of hnRNP particles in splicing at individual nucleotide resolution." *Nat. Struct. Mol. Biol.* **17**: 909-915.

Konig J., Zarnack K., Rot G., Curk T., Kayikci M., Zupan B., Turner D.J., Luscombe N.M., Ule J. (2011). "iCLIP - Transcriptome-wide Mapping of Protein-RNA Interactions with Individual Nucleotide Resolution." *J. Vis. Exp.* **50**: 2638.

Kranzusch P.J., Lee A.S., BergerJ.M., Doudna J.A. (2013). "Structure of human cGAS reveals a conserved family of second-messenger enzymes in innate immunity." *Cell Rep.* **3**: 1362-1368.

Kumar H., Kawai T., Akira S. (2009). "Pathogen recognition in the innate immune response." *Nat. Immunol.* **420**: 1-16.

Lamb R.A., Kolakofsky D. (2001). "Paramyxoviridae: The viruses and their replication." Fields Virology, 4th Edition.

Lamb R.A., Parks G.D. (2007). "Paramyxoviridae: the viruses and their replication", p. 1449-1496. In B.N. Fields and P.M. Howley (ed.), Fields virology.

Lemaitre B., Nicolas E., Michaut L., Reichhart J.M., Hoffmann J.A. (1996). "The dorsoventral regulatory gene cassette spatzle/toll/cactus controls the potent antifungal response in drosophila adults." *Cell* **86**: 973-983.

Leppert M., Rittenhouse L., Perrault J., Summers D. F., Kolakofsky D. (1979). "Plus and minus strand leader RNAs in negative strand virus-infected cells." *Cell* **18**:735-747.

Leung D.W., Amarasinghe G.K. (2012). "Structural insights into RNA recognition and activation of RIG-I-like receptors." *Curr. Opin. Struct. Biol.* **22**: 297-303.

Li L., Speed T.P. (1999). "An estimate of the crosstalk matrix in four-dye fluorescence-based DNA sequencing." *Electrophoresis* **20**: 1433-1442.

Li X., Ranjith-Kumar C.T., Brooks M.T., Dharmaiah S., Herr A.B., Kao C., Li P. (2009). "Structural basis of double-stranded RNA recognition by the RIG-I like receptor MDA5." *Arch. Biochem. Biophys.* **488**: 23-33.

Licatalosi D.D., Mele A., Fak J.J., Ule J., Kayikci M., Chi S.W., Clark T.A., Schweitzer A.C., Blume J.E., Wang X.N., Darnell J.C., Darnell R.B. (2008). "HITS-CLIP yields genome-wide insights into brain alternative RNA processing." *Nature* **456**: 464-469.

Ling Z., Tran K. C., Teng M. N. (2009). "Human respiratory syncytial virus nonstructural protein NS2 antagonizes the activation of beta interferon transcription by interacting with RIG-I." *J. Virol.* **83**: 3734-3742.

Loo Y.M., Fornek J., Crochet N., Bajwa G., Perwitasari O., Martinez-Sobrido L., Akira S., Gill M.A., García-Sastre A., Katze M.G., Gale M. Jr. (2008). "Distinct RIG-I and MDA5 signaling by RNA viruses in innate immunity." *J. Virol.* **82**: 335-345.

Loo Y.M., Gale M. Jr. (2011). "Immune signaling by RIG-I-like receptors." *Immunity* **34**: 680-692.

Lu L.L., Puri M., Horvath C.M., Sen G.C. (2008). "Select paramyxoviral V proteins inhibit IRF3 activation by acting as alternative substrates for inhibitor of kappaB kinase epsilon (IKKe)/TBK1." *J. Biol. Chem.* **283**: 14269-14276.

Luthra P., Sun D., Silverman R.H., He B. (2011). "Activation of IFN-β expression by a viral mRNA through RNase L and MDA5." *Proc. Natl. Acad. Sci. USA* **108**: 2118-2123.

Malathi K., Dong B., Gale M., Jr., Silverman, R.H. (2007). "Small self-RNA generated by RNase L amplifies antiviral innate immunity." *Nature* **448**, 816-819.

Malahti K., Saito T., Crochet N., Barton D.J., Gale M. Jr., Silverman R.H. (2010). "RNase L a small RNA from HCV RNA that refolds into a potent PAMP." *RNA* **16**: 2108-2119.

Manuse M. J., Parks G. D. (2009). "Role for the paramyxovirus genomic promoter in limiting host cell antiviral responses and cell killing." *J. Virol.* **83**: 9057–9067.

Medzhitov R. (2007). "Recognition of moicroorganisms and activation of the immune response." *Nature* **449**: 819-826.

Meisenheimer K.M., Koch T.H. (1997). "Photocross-linking of nucleic acids to associated proteins." *Crit. Rev. Biochem. Mol. Biol.* **32**: 101-140.

Melchjorsen J. (2013). "Learning from the messengers: Innate sensing of viruses and cytokine regulation of immunity – clues for treatments and vaccines." *Viruses* **5**: 470-527.

Mottet G., Curran J., Roux L. (1990). "Intracellular stability of nonreplicating paramyxovirus nucleocapsids." *Virology* **176**: 1-7.

Mottet-Osman G., Iseni F., Pelet T., Wiznerowicz M., Garcin D., Roux L. (2007). " Suppression of the Sendai virus M protein through a novel short interfering RNA approach inhibits viral particle production but does not affect viral RNA synthesis." *J. Virol.* **81**: 2861-2868.

Motz C., Schuhmann K.M., Kirchhofer A., Moldt M., Witte G., Conzelmann K.K., Hopfner K.P. (2013). "Paramyxovirus V proteins disrupt the fold of the RNA sensor MDA5 to inhibit antiviral signaling." *Science* **339**: 690-693.

Oganesyan G., Saha S.K., Guo B., He J.Q., Shahangian A., Zarnegar B., Perry A., Cheng G. (2006). "Critical role of TRAF3 in the Toll-like receptor-dependent and –independent antiviral response." *Nature* **439**: 208-211.

Ohman T., Rintahaka J., Kalkkinen N., Matikainen S., Nyman T.A. (2009) "Actin and RIG-I/MAVS signaling components translocate to mitochondria upon influenza A virus infection of human primary macrophages." *J. Immunol.* **182**: 5682-5692.

Peisley A., Jo M.H., Lin C., Wu B., Orme-Johnson M., Walz T., Hohng S., Hur S. (2012). "Kinetic mechanism for viral dsRNA length discrimination by MDA5 filaments." *Proc. Natl. Acad. Sci. USA* **109**: E3340-E3349.

Peisley A., Wu B., Yao H., Walz T., Hur S. (2013). "RIG-I forms signaling-competent filaments in an ATP-dependent, ubiquitin-independent manner." *Mol. Cell* **51**: 573-583.

Pichlmair A., Schulz O., Tan C.P., Näslund T.I., Liljeström P., Weber F., Reis e Sousa C. (2006). "RIG-I-mediated antiviral responses to single-stranded RNA bearing 5'-phosphates." *Science* **314**: 997-1001.

Pichlmair A., Schulz O., Tan C.P., Rehwinkel J., Kato H., Takeuchi O., Akira S., Way M., Schiavo G., Reis e Sousa, C. (2009). "Activation of MDA5 requires higher-order RNA structures generated during virus infection." *J. Virol.* **83**: 10761-10769.

Poeck H., Bscheider M., Gross O., Finger K., Roth S., Rebsamen M., Hannesschläger N., Schlee M., Rothenfusser S., Barchet W., Kato H., Akira S., Inoue S., Endres S., Peschel C., Hartmann G., Hornung V., Ruland J. (2010). "Recognition of RNA virus by RIG-I results in activation of CARD9 and inflammasome signaling for interleukin 1 beta production." *Nat. Immunol.* **11**: 63-69.

Plumet S., Gerlier D. (2005). "Optimized SYBR green real-time PCR assay to quantify the absolute copy number of measles virus RNAs using gene specific primers." *J. Virol. Methods* **128**: 79-87.

Plumet S., Herschke F., Bourhis J.M., Valentin H., Longhi S., Gerlier D. (2007). "Cytosolic 5'-triphosphate ended viral leader transcript of measles virus as activator of the RIG I-mediated interferon response." *PLoS One* **2**: e279.

Prins K.C., Delpeut S., Leung D.W., Reynard O., Volchkova V.A. Reid S.P., Ramanan P., Cárdenas W.B., Amarasinghe G.K., Volchkov V.E., Basler C.F. (2010). "Mutations abrogating VP35 interaction with double-stranded RNA render Ebola virus avirulent in guinea pigs." *J. Virol.* **84**: 3004-3015.

Rahmeh A. A., Li J., Kranzusch P. J., Whelan S. P. (2009). "Ribose 2′-O methylation of the vesicular stomatitis virus mRNA cap precedes and facilitates subsequent guanine-N-7 methylation by the large polymerase protein." *J. Virol.* **83**: 11043-11050.

Rehwinkel J., Tan C.P., Goubau D., Schulz O., Pichlmair A., Bier K., Robb N., Vreede F., Barclay W., Fodor E., Reis e Sousa C. (2010). "RIG-I detects viral genomic RNA during negative-strand RNA virus infection." *Cell* **140**: 397-408.

Reuter T., Weissbrich B., Schneider-Schaulies S., Schneider-Schaulies J. (2006). "RNA interference with measles virus N, P, and L mRNs

efficiently prevents and with matrix protein mRNA enhances viral transcription." *J. Virol.* **80**: 5951-5957.

Robinson J.T., Thorvaldsdóttir H., Winckler W., Guttman M., Lander E.S., Getz G., Mesirov J.P. (2011). "Integrative Genomics Viewer." *Nat. Biotech.* 29: 24–26.

Rothenfusser S. (2009). "5'-triphosphate RNA requires base-paired structures to activate antiviral signaling via RIG-I." *Proc. Natl. Acad. Sci. USA* **106**: 12067-12072.

Sabbah A., Chang T.H., Harnack R., Frohlich V., Tominaga K., Dube P.H., Xiang Y., Bose S. (2009). "Activation of innate immune antiviral responses by NOD2." *Nat. Immunol.* 10: 1073-1080.

Saito T, Owen DM, Jiang F, Marcotrigiano J, Gale M Jr. (2008). "Innate immunity induced by composition-dependent RIG-I recognition of hepatitis C virus RNA." *Nature* **454**: 523-527.

Sasai M., Linehan M.M., Iwasaki A. (2010). "Bifurcation of Toll-like receptor 9 signaling by adaptor protein 3." *Science* **329**: 1530-1534.

Schlee M., Roth A., Hornung V., Hagmann C.A., Wimmenauer V., Barchet W., Coch C., Janke M., Mihailovic A., Wardle G., Juranek S., Kato H., Kawai T., Poeck H., Fitzgerald K.A., Takeuchi O., Akira S., Tuschl T., Latz E., Ludwig J., Hartmann G. (2009). "Recognition of 5' triphosphate by RIG-I helicase requires short blunt double-stranded RNA as contained in panhandle of negative-strand virus." *Immunity* **31**: 25-34.

Schmidt A., Schwerd T., Hamm W., Hellmuth J.C., Cui S., Wenzel M., Hoffmann F.S., Michallet M.C., Besch R., Hopfner K.P., Endres S.,

Schnell G, Loo YM, Marcotrigiano J, Gale M Jr. (2012). "Uridine composition of the poly U/UC tract of HCV RNA defines non-self recognition by RIG-I." *PLoS Pathog* **8**: e1002839.

Schuhmann K. M., Pfaller C. K., Conzelmann K. K. (2011). "*The measles virus V protein binds to p65 (RelA) to suppress NF-kappaB activity*." *J. Virol.* **85**: 3162-3171.

Shaffer J. A., Bellini W. J., Rota P. A. (2003). "The C protein of measles virus inhibits the type I interferon response." *Virology* **315**: 389-397.

Solis M., Nakhaei P., Jalalirad M., Lacoste J., Douville R., Arguello M., Zhao T., Laughrea M., Wainberg M.A., Hiscott J. (2011). "RIG-I-mediated antiviral signaling is inhibited in HIV-1 infection by a protease-mediated sequestration of RIG-I." *J. Virol.* **85**: 1224-1236.

Sparrer K.M., Pfaller C.K., Conzelmann K.K. (2012). "Measles virus C protein interferes with Beta interferon transcription in the nucleus." *J. Virol.* **81**: 12227-12237.

Strahle L., Garcin D., Kolakofsky D. (2006). "Sendai virus defective-interfering genomes and the activation of interferon-beta." *Virology* **351**: 101-111.

Strahle L., Marq J.B., Brini A., Hausmann S., Kolakofsky D., Garcin D. (2007). "Activation of the beta interferon promoter by unnatural Sendai virus infection requires RIG-I and is inhibited by viral C proteins." *J. Virol.* **81**: 12227-12237.

Tanaka Y., Chen Z.J. (2012). "STING specifies IRF3 phosphorylation by TBK1 in the cytosolic DNA signaling pathway." *Sci. Signal.* **5**, ra20.

Takeuchi O., Akira S. (2010). "Pattern recognition receptors and inflammation." *Cell* **140**: 805-820.

Taylor K.E., Mossman K.L. (2012). "Recent advances in understanding viral evasion of type I interferon." *Virology* **138**: 190-197.

Ting J.P., Lovering R.C., Alnemri E.S., Bertin J., Boss J.M., Davis B.K., Flavell R.A., Girardin S.E., Godzik A., Harton J.A., Hoffman H.M., Hugot J.P., Inohara N., Mackenzie A., Maltais L.J., Nunez G., Ogura Y., Otten L.A., Philpott D., Reed J.C., Reith W., Schreiber S., Steimle V., Ward P.A. (2008). "The NLR gene family: a standard nomenclature." *Immunity* **28**: 285-287.

Tseng P.H., Matsuzawa A., Zhang W., Mino T., Vignali D.A., Karin M. (2009). "Different modes of ubiquitination of the adaptor TRAF3 selectively activate the expression of type I interferon's and proinflammatory cytokines." *Nat. Immunol.* **11**: 70-75.

Ule J., Jensen K.B., Ruggiu M., Mele A., Ule A., Darnell R.B. (2003). "CLIP identifies Nova-regulated RNA networks in the brain." *Science* **302**: 1212-1215.

Ule J., Ule A., Spencer J., Williams A., Hu J.S., Cline M., Wang H., Clark T., Fraser C., Ruggiu M., Zeeberg B.R., Kane D., Weinstein J.N., Blume J., Darnell R.B. (2005). "Nova regulates brain-specific splicing to shape the synapse." *Nat. Genet.* **37**: 844-852.

Upton G., Cook I. (1996). "Understanding statistics." Oxford University Press.

Vidal S., Kolakofsky D. (1989). "Modified model for the switch from Sendai virus transcription to replication." *J. Virol.* **63**: 1951-1958.

Wang Y., Ludwig J., Schuberth C., Goldeck M., Schlee M., Li H., Juranek S., Sheng G., Micura R., Tuschl T., Hartmann G., Patel D.J. (2010). "Structural and functional insights into 5'-ppp RNA pattern recognition by the innate immune receptor RIG-I." *Nat. Struct. Mol. Biol.* **17**: 781-787.

Weber F.V., Wagner S., Rasmussen B., Hartmann R., PaludanS.R. (2006). "Double-stranded RNA is produced by positive-strand RNA viruses and DNA viruses but not detectable amounts by negative-strand RNA viruses." *J. Virol.* **80**: 5059-5064.

Weber M., Gawanbacht A., Habjan M., Rang A., Borner C., Schmidt A.M., Veitinger S., Jacob R., Devignot S., Kochs G.,Garcıa-Sastre A., Weber F. (2013). "Incoming RNA virus nucleocapsids containing a 5'-triphosphorylated genome activate RIG-I and antiviral signaling." *Cell Host Microbe* **13**: 336-346.

Wood H.M., Belvedere O., Conway C., Daly C., Chalkley R., Bickerdike M., McKinley C., Egan P., Ross L., Hayward B., Morgan J., Davidson L., MacLennan K. Ong T.K., Papagiannopoulos K., Cook I., Adams D.J., Taylor G.R., Rabbits P. (2010). "Using next-generation sequencing for high resolution multiplex analysis of copy number variation from nanogram quantities of DNA from formalin-fixed paraffin-embedded specimens." *Nucleic Acids Res.* **38**: e151.

Wu J., Sun L., Chen X., Du F., Shi H., Chen C., Chen Z.J. (2013). "Cyclic GMP-AMP is an endogenous second messenger in innate immune signaling by cytosolic DNA." *Science* **339**: 826-830.

Yoneyama M., Kikuchi M., Natsukawa T., Shinobu N., Imaizumi T., Miyagishi M., Taira K., Akira S., Fujita T. (2004). "The RNA helicase RIG-I has an essential function in double-stranded RNA-induced innate antiviral responses." *Nat. Immunol.* **5**: 730-737.

Yount J.S., Gitlin L., Moran T.M.., Lopez C.B. (2008). "MDA5 participates in the detection of paramyxovirus infection and is essential for the early activation of dendritic cells in response to Sendai virus defective interfering particles." *J. Virol.* **180**: 4910-4918.

Zeng W., Sun L., Jiang X., Chen X., Hou F., Adhikari A., Xu M., Chen Z.J. (2010) "Reconstitution of the RIG-I pathway reveals a signaling role of unanchored polyubiquitin chains in innate immunity." *Cell* **141**: 315-330.

Zhong B., Yang Y., Li S., Wang Y.Y., Li Y., Diao F., Lei C., He X., Zhang L., Tien P., Shu H.B. (2008). "The adaptor protein MITA links virus-sensing receptors to IRF3 transcription factor activation." *Immunity* **29**: 538-550.

Zhou A., Molinaro R.J., Malathi K., Silverman R.H. (2005) "Mapping of the human RNASEL promoter and expression in cancer and normal cells." *J. Interferon Cytokine* **25**: 595-603.

Abbreviations

4SU	4 thiouridine
AIM2	Absent in melanoma 2
APC	Antigen presenting cell
ASC	Apoptosis-associated speck-like protein containing a CARD
CARD	Caspase activation and recruitment domain
cDNA	complementary DNA
cGAMP	Cyclic-GMP-AMP
cGAS	cGAMP synthase
CNV	Copy number variation
DAMP	Danger-associated molecular pattern
DC	Dendritic cell
DExD/H	Asp-Glu-x-Asp/His
DI	Defective interfering particle
dsRNA	double-strand RNA
EBOV	Ebola virus
EMCV	Encephalomyocarditis virus
ER	Endoplasmatic reticulum
FADD	Fas-associated death domain
FF	Firefly luciferase
GFP	Green fluorescence protein
HIN200	Hematopoietic interferon-inducible nuclear antigens with 200 amino acid repeats
HITS-CLIP	High-throughput sequencing of RNA isolated by crosslinking immunoprecipitation
HIV	Human immunodeficiency virus

iCLIP	Individual-nucleotide resolution crosslinking and immunoprecipitation
IFN	Interferon
IL	Interleukin
IKK	Inhibitor of nuclear factor-κB kinase
IRAK	IL-1R-associated kinase
IRF	Interferon regulatory factor
ISRE	IFN stimulated respone element
IVT	*In vitro* transcript
leRNA	leader RNA
LGP2	Laboratory of genetics and physiology 2
LRR	Leucine rich repeat
MAVS	Mitochondrial antiviral signaling protein
MDA5	Melanoma differentiation-associated gene 5
MeV	Measles virus
MHC	Major histocompatibility complex
MOI	Multiplicicity of infection
mRNA	messenger RNA
MyD88	Myeloid differentiation protein 88
NAP1	Nucleosome-Associated Protein 1
NC	Nucleocapsid
NEMO	N-κB essential modulator
NF-κB	Nuclear factor-κB
NGS	Next generation sequencing
NLR	NOD-like receptor
NLRC	NOD-like receptors containing a CARD domain
NLRP	NOD-like receptors containing a PYD domain

NOD	Nucleotide-binding oligomerization domain
OAS	Oligoadenylate synthetase
PAMP	Pathogen-associated molecular pattern
PAR-CLIP	Photoactivatable-ribonucleoside-enhanced crosslinking and immunoprecipitation (CLIP)
PIV5	Parainfluenza virus 5
PRR	Patttern recognition receptor
PYD	Pyrin domain
PYHIN	Pyrin and HIN domain-containing protein
RABV	Rabies virus
RD	Regulatory domain
RFP	Red fluorescence protein
RIG-I	Retinoic acid inducible gene I
RIP-1	Receptor-Interacting Protein 1
RLR	RIG-I-like receptor
RNP	RNA-binding protein
RPV	Rinder pest virus
RSV	Respiratory syncytial virus
vRNA	viral RNA
SeV	Sendai virus
SLS	Systemic lupus erythematosus
STAT	Signal transducer and activator of transcription
STING	Stimulator of IFN genes
TANK	TRAF-associated NF-κB activator
TBK1	TANK-binding kinase 1
TIM	TRAF-interacting motif
TIR	Toll/interleukin-1 receptor

TIRAP	TIR domain-containing adaptor protein
TLR	Toll-like receptor
TRADD	Tumor necrosis factor receptor (TNFR)-associated death domain
TRAF	TNF receptor-associated factor
TRAM	TRIF-related adaptor molecule
TRIF	Toll/IL-1 receptor domain-containing adaptor inducing IFN
TRIM25	Tripartite motif 25
trRNA	trailer RNA

Acknowledgements

Zuerst möchte ich mich bei meinem Doktorvater Herrn Prof. Dr. Karl-Peter Hopfner für die Unterstützung und das Vertrauen, das er in mich gesetzt hat, ganz herzlich bedanken. Er hat mir alle nötigen Freiheiten gegeben, so dass ich mich in diesem spannenden Projekt verwirklichen konnte. Ich bin ihm zutiefst dankbar für sein Verständnis und seine Rücksicht in Bezug auf meine Rolle als Vater, die es mir ermöglicht haben, meine Arbeit im harmonischen Einklang mit meinen familiären Pflichten und Freuden zu verbinden.

Ebenfalls möchte ich mich aufrichtig bei Herrn Prof. Dr. Karl-Klaus Conzelmann für seine Bereitschaft bedanken, das Zweitgutachten zu übernehmen. Darüberhinaus danke ich ihm und Konstantin Sparrer für eine außerordentlich kollegiale Kollaboration. Durch zahlreiche anregende Diskussionen habe ich wichtige Grundkenntnisse der Virologie vermittelt bekommen, die mir in meinem Projekt sehr von Nutzen waren. Weiterhin möchte ich betonen, dass ohne diese Kollaboration das Projekt nicht realisierbar gewesen wäre, da uns Konstantin unter anderem die nötigen Viren zur Verfügung gestellt hat.

Aus meiner Arbeitsgruppe möchte ich vor allem Charlotte Lässig für die Unterstützung bei der bioinfomatischen Analyse, sowie bei der Einreichung des Papers danken. Im Allgemeinen danke ich der gesamten Gruppe für die angenehme Arbeitsatmosphäre und die netten, freundschaftlichen Gespräche während meiner Zeit im Labor.

Ich möchte mich bei Dr. Johannes Söding und Katharina Hembach für die Bioinformatik bedanken, ohne die ein großer Teil der Analysen nicht möglich gewesen wäre. Durch die Zusammenarbeit habe ich eine andere Sichtweise auf die in dieser Arbeit gegebenen Fragestellung kennen gelernt, die sehr hilfreich für mich war.

Weiterhin möchte ich mich ganz herzlich bei Prof. Dr. Simon Rothenfusser, Prof. Dr. Dietmar Martin und Prof. Dr. Klaus Förstemann bedanken, die sich bereit erklärt haben, Mitglieder meiner Prüfungskommission zu sein.

Zu guter Letzt danke ich von ganzem Herzen meiner Familie, die mir unendlich viel Liebe und Geborgenheit entgegenbringt. Ich habe von meinen Eltern und meiner Schwester jederzeit die volle Unterstützung und das volle Vertrauen erhalten. Worte reichen hier nicht aus, um meine Dankbarkeit auszudrücken und ihnen sei diese Arbeit gewidmet.
Ich danke meinen Kindern, Smilla und Emil, aus deren Lächeln und ihrer einzigartigen Art, ich viel Freude, Mut und Kraft schöpfe. Sie sind der Lohn aller Mühen.
Meiner Frau Petra danke ich nicht nur für das Korrekturlesen dieser Arbeit, sondern vielmehr für ihre liebe Art und das Gefühl, das sie mir gibt, mit meinen Sorgen niemals alleine zu sein.