RLR-mediated IFN signaling and aberrant activation

Han Yang1,2; Yunran Feng1,2; Chunjing Feng2,3*; Xin Mu1,2
1 School of Pharmaceutical Science and Technology, Tianjin University, Tianjin 300072, China.
2 Tianjin University and Health-Biotech United Group, Joint Laboratory of Innovative Drug Development and Translational Medicine, Tianjin University, Tianjin 300072, China.
3 Health-Biotech Group Stem Cell Research Institute, Tianjin 301799, China.

*Corresponding Author: Chunjing Feng
Tianjin University and Health-Biotech United Group Joint Laboratory of Innovative Drug Development and Translational Medicine, Tianjin University, Tianjin 300072, China.
Email: fengchunjing@health-biotech.com

Equal contribution: Feng C & Mu X

Received: Jan 16, 2022
Accepted: Feb 15, 2022
Published: Feb 22, 2022
Archived: www.jcimcr.org
Copyright: © Feng C (2022).
DOI: www.doi.org/10.52768/2766-7820/1682

Abstract

Interferon (IFN) signaling is important for host cells in defending microorganisms. “Self” and “non-self” recognition is the key step for proper IFN functioning. Pathogen-derived double-stranded RNAs (dsRNAs) are immunogenic to the cytosolic sensor, RIG-I and MDA5. Together known as RIG-I-like receptors (RLRs), they form filament along the length of dsRNA substrate which in turn activate downstream cascades. Despite RLRs themselves and host encoded regulators managing to avoid self-intolerance, extensive studies in recent years uncovered that the boundary of “self” and “non-self” can be vague. Retrotransposon elements embedded in the genome of humans can turn into RLR ligands under certain circumstances, like incorrect RNA metabolism or epigenetic drug treatment, which in turn, lead to IFN signaling activation. The consequences of such activation can be different according to different circumstances. Here, we summarized the biological features of RLR-mediated IFN signaling and discussed aberrant IFN signaling through RLRs, including situations of gene mutations, irradiation, and anti-cancer drug treatment. We speculated that combining current cancer therapy and RLR-mediated IFN signaling activation would bring beneficial effects to cancer treatment.

Keywords: RLR; IFN; MDA5; RIG-I; ADAR1; DNA methylation.

**Introduction**

Interferon (IFN) signaling pathway is an important defense mechanism for host cells to resist the invasion of microorganisms. It leads to the expression of IFN-stimulated genes (ISGs), including effectors that directly fight against pathogens and regulators modulating this stressed response as well as communicating with innate and adaptive immunity [1,2]. ISGs can be classified into three groups according to sequences, functions, and their receptors [3]. The type-I IFNs (IFN-I, e.g., IFN-a, IFN-b) and type-III IFNs (IFN-III, e.g., IFN-γ) engage in anti-microbialism defense while type-II IFN (IFN-II, e.g., IFN-γ) participates in regulating the host immunity [4,5].

All nucleated cells express receptors for IFN-I [6]. The proper functioning of this pathway is firstly dependent on the correct recognition of “self” and “non-self” molecules by the receptors. These proteins are called pattern recognition receptors (PRRs) as they recognize pathogen-associated molecular patterns (PAMPs) or damage-associated molecular patterns (DAMPs), instead of binding to specific sequences [7]. The “non-self” molecules can be DNA, RNA, lipoproteins, etc. Immunogenic nucleic acids can be derived exogenously from pathogenic infection in the cytoplasm [8,9]. From here, two branches of IFN signaling can be considered: One is mediated by the interaction of dsRNA with RIG-I and MDA5 activating downstream cascades for the stimulation of ISGs. The other branch is mediated through CARDMAVS motifs [17,33]. Such aggregation creates a filament and triggers the release of MAVS for binding through the CARDMAVS interaction. Multiple MAVS are induced to further impair the integrity of the host cell [12]. In contrast, virus-derived nucleic acids can be DNA, RNA, lipoproteins, etc. Immunogenic nucleic acids can be derived exogenously from pathogenic infection in the cytoplasm [10]. Endogenous RNAs are mainly single-strand RNA (ssRNA) or dsRNA [7,11]. Because cellular DNAs are confined to the nucleus and mitochondria, pathogen-derived RNAs in different viruses may have little sequence commonality. RIG-I and MDA5 share sequence and domain architecture similarities and use the same adaptor Mitochondrial antiviral-signaling protein (MAVS) for activating downstream cascades [10,16]. The N-terminus of RLR is a tandem caspase activation signaling protein (MAVS) for activating downstream cascades, while the C-terminal domain (CTD) are together responsible for dsRNA binding [17,18] (Figure 1A). Structural studies on RIG-I demonstrated that in the absence of dsRNA ligand, it exerts a conformation of auto-repressed manner, in which case the 2CARDs folds back to associate with helicase domain so that 2CARDs are not exposed to eliciting signaling [19] (Figure 1B). A recent study suggested that endogenously encoded long non-coding RNA (IncRNA), Inc-Lsm3b, may contribute to the silence of RIG-I by sequestering it in a homeostasis state as well [20].

Although both receptors recognize and bind to dsRNAs, the binding patterns and preferred substrates are different. RIG-I recognizes the end of dsRNA bearing the 5’-ppp or 5’-pp’ structures through CTD then binds to it [21,22]. Next, this RIG-I translocates to the interior of dsRNA upon hydrolyzing ATP through its helicase domain. The unoccupied 5’-end on dsRNA can recruit another molecule of RIG-I [23]. Cycles of ATP hydrolysis and translocation produce a RIG-I filament along the length of dsRNA [23] (Figure 1C). Such mode of action confers RIG-I is best stimulated by dsRNAs at a short length, e.g., 40-150 bp [24,25]. MDA5, on the other hand, binds to the interior of dsRNA as the speed-limiting step. New MDA5s bind quickly to this nucleation site and a filament is growing towards both sides [26]. Biochemical assays showed that MDA5 hydrolyses ATP for disassembly, but only the ones at the ends of filament are able to leave [26,27] (Figure 1D). This assembly-and-disassembly kinetics of binding determines that MDA5 prefers longer duplex RNAs, e.g., ~0.5-7 kb long dsRNA [25,26]. For both receptors, the ATPase activity was shown to serve as a proof-reading mean for correct substrates [26,28,29].

The differences in preference suggest that they are non-redundant dsRNA sensors. RIG-I is stimulated by negative-sense RNA viruses, like Sendai virus (SeV) and Vesicular Stomatitis Virus (VSV) [13,30], while MDA5 is intolerant to positive-sense RNA virus infection, such as picornavirus EMCV (Encephalomyocarditis virus) [13]. Besides, mice assays clearly showed that RIG-I is the responsible receptor for Japanese encephalitis virus (JEV) despite it being a positive-sense RNA virus [13]. RNA substrate determination has always been a challenging project for dsRNA-binding proteins. In the case of MDA5, it also binds to ssRNA randomly and with equivalent affinity in biochemical systems [31], which in turn, interferes with the analysis of assay results. One solution is to utilize the formation of filament, which encloses dsRNA substrate and thus protects it from RNase digestion in mild digesting conditions [12]. Yet, given that RLRs bind to dsRNA in a sequence-independent manner, bound RNAs in different viruses may have little sequence commonalities in general and thus is less informative. As noted, synthetic dsRNA polynosinic–polycytidylic acid (poly I:C) is composed of two nucleotides only, where the lower molecular weight (LMW) version with a shorter length is better in activating RIG-I and the longer version is best in stimulating MDA5 [25].

**Downstream cascades**

Albeit the filaments are formed by different mechanisms, they similarly induce the intermolecular 2CARDs forming oligomers [17,32]. The 2CARDs oligomers, in turn, recruit the mitochondria-anchoring protein MAVS for binding through the 2CARDsRLR-CARDMAVS interaction. Multiple MAVS are induced to form a prion-like aggregate via the MAVS-MAVS interaction through CARDMAVS motifs [17,33]. Such aggregation creates
a new platform to recruit downstream factors, such as tumor necrosis factor receptor-associated factors (TRAFs) and the subsequently associated TANK Binding Kinase 1 (TBK1)/inhibitor of NF-κB kinase e (IKKe), NF-κB essential modulator (NEMO)/IKKa/IKKb, etc. TBK1/IKKe and NEMO/IKKa/IKKb then activate the type-I IFN and NF-κB signaling respectively [34,35]. A schematic view of RLRs-mediated IFN signaling is illustrated in Figure 2.

A feature of RLR-mediated IFN signaling is the formation of filament, which arises from monomeric protein to large complex. Recent studies uncovered a post-filament event by host-encoded E3 ubiquitin ligases to regulate both RIG-I and MDA5 activity. RIPLET (also known as RNF135, Ring Finger Protein 135) and Tripartite Motif Containing 65 (TRIM65) were identified as RIG-I filament- and MDA5 filament-binding proteins, respectively [24,36,37]. Intriguingly, neither of the E3 ligases could bind to the corresponding RLR at monomeric form. Both ligases dictate the K63-linked ubiquitination of target RLRs, and such modification is required for RLR-mediated IFN activation.

Homeostasis and aberrant activation

dsRNA editing by ADAR1: As mentioned earlier, PAMPs or DAMPs are not sequence-dependent but pattern-dependent, which means that the silencing of IFN-I signaling in the absence of infection is very largely dependent on the control and homeostasis of cellular endogenous molecules [38]. For RLRs, such control comes from the modulation of the endogenous dsRNA pool. The adenosine deaminase acting on RNA 1 (ADAR1) is a dsRNA-specific adenosine deaminase that catalyzes adenosine (A) converting into inosine (I) in RNA duplex, such reaction melts the integrity of dsRNA structures by introducing mismatches (A: U becomes I_U) [39,40]. To be precise, ADAR1 controls the homeostasis of MDA5 [12,41]. Studies showed that the loss-of-expression of ADAR1 results in the embryonic lethal phenotype in mice and is completely rescued by the additional loss-of-expression of MDA5 [41]. ADAR1 in humans is suggested to mainly edit endogenous inverted-repeat Alu (IR-Alu) elements [42,43], implicating that IR-Alu duplex is targeted by human MDA5 if without A-to-I editing. In line with this, loss-of-function (LOF) mutations of ADAR1 and GOF mutations of IFIH1 (encoding MDA5) identified in Aicardi–Goutières syndrome (AGS) patients were shown to aberrantly activate IFN signaling via the same endogenous dsRNA species, IR-Alu elements [12]. That is being said, the GOF mutants of MDA5 form signaling-transducing filament on IR-Alu duplexes even these RNAs are edited by ADAR1, or the wildtype (WT) MDA5 binds to unedited IR-Alu duplexes to activate IFN-I signaling, in the absence of viral infection.

AGS and SMS: AGS is an immune disorder affecting mainly skin and brain and is the prototype of type-I interferonopathies conceptualized firstly by Yanick J. Crow in 2011 [44]. The type-I interferonopathies are featured by the aberrant activation of IFN-I signaling in the absence of pathogen invasion, due to monogenic mutations of somatic genes and such mutations obey Mendel’s law [9,44]. Another example of such a disease is Singleton-Merten syndrome (SMS) [14,15]. Both RIG-I and MDA5 mutations were identified in SMS [14,15]. It is interesting that up to now, RIG-I mutations were not yet identified in AGS patients. Manifestations of SMS include but are not limited to, aortic calcification, dental dysplasia, and abnormalities in the skin [45,46]. Given that the GOF mutations of both receptors aberrantly activate IFN-I signaling to rely on the same downstream cascades, AGS and SMS share some manifestations clinically, for example, glaucoma [47]. In fact, as the aberrant activation of IFN-I signaling is a feature in type-I interferonopathies, their symptoms were shown to have some overlapped manifestations. Phenotypic overlaps were also seen in AGS and systemic lupus erythematosus (SLE) [48], AGS and spondyloenchondromatosis (SPENCD) [49], etc. An excellent review on this topic can be found in the following reference [50]. We summarized RLR gene mutations in Table 1 [15,51-62].

From the table, it is shown that mutations in RIG-I and MDA5 occurred in motifs responsible for dsRNA-recognition and binding, the helicase domain and CTD, implicating these GOF mutants break the immune balance by losing tolerance to the endogenous self RNAs. The prime-specific retrotransposon, IR-Alu, was identified in human cells as the ligand for GOF MDA5 in 2018 [12], yet the endogenous ligands for GOF RIG-I are not identified so far. It is noted that the embryonic lethal effect in mice by deficiency of ADAR1 was rescued by MDA5 or MAVS knockout, but not RIG-I [41,63,64], implicating that the dsRNAs edited by ADAR1 are not targeted by RIG-I. Moreover, take IR-Alu as an example, it normally locates inside the intron in nucleus or 3’-UTR of mRNAs [12,42] where no 5’-ppp or 5’-pp end is exposed for RIG-I binding. Thus, the endogenous RNA substrates for GOF MDA5 and RIG-I are different. Evidence suggested that under certain circumstances such as ionizing radiation (IR) therapy, small nuclear RNAs U1 and U2 can translocate to the cytoplasm and be targeted by RIG-I to activate IFN sig-

Figure 1: Schematic views of domain architecture and filament formation of RLRs.

(A) Domains of RIG-I and MDA5. The helicase domain is composed of Hel1, Hel2i, and Hel2. P, pincer domain. The number below each protein is the amino acid number.

(B) Auto-repressive structure of RIG-I in the absence of dsRNA ligand. (C) RIG-I forms filament through the end of dsRNA. Upon ATP hydrolysis, RIG-I translocates to the interior of dsRNA and new RIG-I binds to the end for the next cycle. (D) MDA5 forms filament in the interior of dsRNA. ATP hydrolysis leads to the dissociation of proteins at the filament end where new proteins are ready to associate.
naling [65]. This observation suggested that the nuclear non-coding RNAs may have the ability to activate RIG-I, however, under normal circumstances, they are confined to the nucleus. Whether other non-coding RNAs present in cytoplasm become the ligand of GOF RIG-I is of interest to know.

Impacts of DNA methylation: Homeostasis breakdown not only happens in the case of gene mutations but also occurs in the treatment of pathological conditions, especially cancer chemotherapies and irradiation (as mentioned previously). DNA methylation plays important role in regulating DNA transcription epigenetically. Methyltion and transcription are generally negatively correlated with each other in the case of transposon elements [66,67]. An epigenetic anti-cancer drug, 5-aza-2' deoxycytidine (5-AZA-CdR), inhibits DNA methylation and is used in the treatment of hematological malignancies [68]. Evidence demonstrated that MDA5 is the key player in the effectiveness of 5-AZA-CdR in killing cancer cells [69,70]. Furthermore, the interplay between ADAR1 and MDA5 determines the efficacy of 5-AZACdR in killing cancer cells [71]. All these observations implicated that endogenous transposon elements can induce IFN-I signaling when activated in transcription. In line with this, inhibition of LSD1 activity in breast tumors was reported to induce IFN-I signaling which in turn, beneficial to antitumor treatment [72]. LSD1 stands for lysine-specific demethylase 1 and is a histone H3K4 demethylase. Trimethylated H3K4 histone is found in actively transcribed regions [73] where LSD1 suppresses gene expression by converting trimethylated H3K4 to mono- or unmethylated one [74]. Data showed that both sense-strand and antisense-strand endogenous retroviruses (ERV) (belong to retrotransposon) transcription were increased in LSD1-KO cells, thus producing endogenous dsRNA species that activate MDA5 in the cytoplasm [75].

In a word, we now are fully aware that the concept “self” and “non-self” are not based on the origins of molecules. Changed conditions such as chemotherapy or gene mutations can transfer the “self” genome into a reservoir of stimulant ligands, and to lead undesired consequences to the cell and the host.

Outlooks

A combining effort of biochemistry, biophysics, cellular approaches, and in vivo assays for almost two decades has unmasked the biology of the two cytosolic dsRNA sensors. Given that RLR-mediated IFN signaling is a stress response, it is regulated at multiple steps by a variety of host-encoded regulators. The recent findings that two E3 ligases (RIPTLET and TRIM65) specifically bind to the filamentous form of the corresponding RLRs but not the monomeric form to regulate substrate protein’s activity unleashed an open question of whether such mode of action represents a new kind of mechanism by which the cells apply to regulate PRRs. Biochemical data showed that at least several members of the TRIM family are capable of binding to different RNA helicases in a conformation-selecting manner (LGP2-TRIM14, Dicer-TRIM25, etc.) [37].

Mounting evidence from studies on autoimmune diseases and anti-cancer therapies demonstrated that our genome not only contains genes necessary for cellular survival and proliferation but also is embedded with repeated elements that are potentially immunogenic when transcribed in certain conditions. These repeated elements are retrotransposons integrated into the human genome as a remanence from an ancient age. Evolutions have shaped RLRs with balanced capacities between sensivity and selectivity. Under this logic, as others had suggested, the IFN pathway is more like a pathway hosts utilize to monitor cellular status, rather than inhibiting pathogen invasion only. PAMPs and DAMPs then may also be divided into the following two groups: first, wrong molecules (such as perfect, unedited dsRNA) derived endogenously or exogenously; second, right molecules in the wrong place (such as DNA in cytoplasm or ssRNA in endosome).

In the case of AGS raised from mutations of MDA5, treatment, for now, focuses on restricting the IFN signaling pathway, for example, the use of JAK inhibitors (ruxolitinib, baricitinib, etc.) [9,76,77], or antibodies against IFN-a [78]. Blocking the recognition of the IR-Alu duplex by GOF MDA5 is definitely a very important direction to investigate, but it is technically hard. RNAs are normally knocked down by treatment of siRNA or miRNA [79]. However, this is difficult for IR-Alu RNAs because these elements are pre-formed duplexes. Furthermore, the enormous number of Alu elements in the genome [80,81] makes it impossible to knockout by the CRISPR-Cas9 system. Yet it is worth noting that IR-Alu duplexes are not perfectly matched. The edition by ADAR1 further melts the integrity of these duplexes [12,82]. Thus, bulges and mismatches are present in every single IR-Alu and can be potential sites for unwinding.
For the repeated elements in the genome, on the one hand, they have the potential to elicit unwanted stress signaling; on the other hand, they can be manipulated to serve as good tools in killing cancer cells. Indeed, in the study of LSD1, it was suggested that deficiency of LSD1 overcomes tumor resistance to PD-1 antibody treatment in mice [75]. Thus, the combining of anti-tumor antibodies, immune cells, and epigenetic drugs would become a promising direction of cancer therapy in the future.

Table 1: RLR mutations identified in AGS and/or SMS. Gene and mutation sites are described, as well as the belonging domains.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Mutation</th>
<th>Domain</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>DDX58</td>
<td>c.803G&gt;T (p. Cys268Phe)</td>
<td>Helicase</td>
<td>51</td>
</tr>
<tr>
<td></td>
<td>c.1118A&gt;C (p. Gly373Ala)</td>
<td>Helicase</td>
<td>51</td>
</tr>
<tr>
<td></td>
<td>c.1529A&gt;T (p. Gly510Val)</td>
<td>Helicase</td>
<td>52</td>
</tr>
<tr>
<td></td>
<td>c.1551G&gt;C (p. Gln517His)</td>
<td>Helicase</td>
<td>53</td>
</tr>
<tr>
<td>IFIH1</td>
<td>c.992C&gt;G (p. Thr331Arg)</td>
<td>Helicase</td>
<td>54</td>
</tr>
<tr>
<td></td>
<td>c.1009A&gt;G (p. Arg337Gly)</td>
<td>Helicase</td>
<td>55</td>
</tr>
<tr>
<td></td>
<td>c.1114C&gt;T (p. Leu372Phe)</td>
<td>Helicase</td>
<td>56</td>
</tr>
<tr>
<td></td>
<td>c.1165G&gt;A (p. Gly389Arg)</td>
<td>Helicase</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>c.1178A&gt;T (p. Asp393Val)</td>
<td>Helicase</td>
<td>55</td>
</tr>
<tr>
<td></td>
<td>c.1178A&gt;C (p. Asp393Ala)</td>
<td>Helicase</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>c.1331A&gt;G (p. Glu444Gly)</td>
<td>Helicase</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>c.1347C&gt;G (p. Asn449Lys)</td>
<td>Helicase</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>c.1354G&gt;A (p. Ala452Thr)</td>
<td>Helicase</td>
<td>56</td>
</tr>
<tr>
<td></td>
<td>c.1465G&gt;A (p. Ala489Thr)</td>
<td>Helicase</td>
<td>57</td>
</tr>
<tr>
<td></td>
<td>c.1465G&gt;T (p. Ala489Ser)</td>
<td>Helicase</td>
<td>58</td>
</tr>
<tr>
<td></td>
<td>c.1483G&gt;A (p. Gly495Arg)</td>
<td>Helicase</td>
<td>55</td>
</tr>
<tr>
<td></td>
<td>c.1747A&gt;G (p. Ile583Val)</td>
<td>Helicase</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>c.2156C&gt;T (p. Ala719Val)</td>
<td>Helicase</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>c.2159G&gt;A (p. Arg720Gln)</td>
<td>Helicase</td>
<td>55</td>
</tr>
<tr>
<td></td>
<td>c.2335C&gt;T (p. Arg779Cys)</td>
<td>Helicase</td>
<td>55</td>
</tr>
<tr>
<td></td>
<td>c.2336G&gt;A (p. Arg779His)</td>
<td>Helicase</td>
<td>55</td>
</tr>
<tr>
<td></td>
<td>c.2336G&gt;T (p. Arg779Leu)</td>
<td>Helicase</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>c.2390A&gt;T (p. Asp797Val)</td>
<td>Helicase</td>
<td>59</td>
</tr>
<tr>
<td></td>
<td>c.2342G&gt;A (p. Gly781Glu)</td>
<td>Helicase</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>c.2407A&gt;T (p. Ile803Phe)</td>
<td>Helicase</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>c.2439A&gt;T (p. Gln813Asp)</td>
<td>Helicase</td>
<td>60</td>
</tr>
<tr>
<td></td>
<td>c.2465G&gt;A (p. Arg822Gln)</td>
<td>Helicase</td>
<td>61</td>
</tr>
<tr>
<td></td>
<td>c.2471G&gt;A (p. Arg824Lys)</td>
<td>Helicase</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>c.2486C&gt;G (p. Thr829Ser)</td>
<td>Helicase</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>c.2544T&gt;G (p. Asp848Glu)</td>
<td>Helicase</td>
<td>62</td>
</tr>
<tr>
<td></td>
<td>c.2561T&gt;A (p. Met854Lys)</td>
<td>Helicase</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>c.2866A&gt;G (p. Ile956Val)</td>
<td>CTD</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>c.2936T&gt;G (p. Leu979Trp)</td>
<td>CTD</td>
<td>14</td>
</tr>
</tbody>
</table>

For the repeated elements in the genome, on the one hand, they have the potential to elicit unwanted stress signaling; on the other hand, they can be manipulated to serve as good tools in killing cancer cells. Indeed, in the study of LSD1, it was suggested that deficiency of LSD1 overcomes tumor resistance to PD-1 antibody treatment in mice [75]. Thus, the combining of anti-tumor antibodies, immune cells, and epigenetic drugs would become a promising direction of cancer therapy in the future.

Declarations

**Funding:** This work was supported by the State Key Laboratory of Veterinary Etiological Biology, CAAS, and the Guangdong Provincial Key Laboratory of Precision Medicine and Clinical Translation Research of Hakka Population (Grants SKLVEB2020KFKT001 and 2018B030322003KF03 to X.M.)

**Conflicts of interest:** The authors declare no conflicts of interest.

**Author contributions:** H.Y., Y.F., C.F., and X.M. wrote and revised the manuscript, H.Y. draws the figures.

**References**


41. BJ Liddicoat, R Piskol, AM Chalk, G Ramaswami, M Higuchi, et al. RNA editing by ADAR1 prevents MDAS sensing of endogenous dsRNA as nonsense, Science. 2015; 349: 1115-1120.


