

# A conserved isoleucine in the binding pocket of RIG-I controls immune tolerance to mitochondrial RNA

Ann Kristin de Regt<sup>1</sup>, Kanchan Anand<sup>2</sup>, Katrin Ciupka<sup>1</sup>, Felix Bender<sup>1</sup>, Karl Gatterdam<sup>2</sup>, Bastian Putschli<sup>1</sup>, David Fusshöller<sup>2</sup>, Daniel Hilbig<sup>3</sup>, Alexander Kirchhoff<sup>1</sup>, Charlotte Hunkler<sup>1</sup>, Steven Wolter<sup>1</sup>, Agathe Grünewald<sup>1</sup>, Christina Wallerath<sup>1</sup>, Christine Schuberth-Wagner<sup>1</sup>, Janos Ludwig<sup>1</sup>, Katrin Paeschke<sup>1,3</sup>, Eva Bartok<sup>1,4,5</sup>, Gregor Hagelueken<sup>2</sup>, Gunther Hartmann<sup>1</sup>, Thomas Zillinger<sup>1,6</sup>, Matthias Geyer<sup>2</sup> and Martin Schlee<sup>1,\*</sup>

<sup>1</sup>Department of Clinical Chemistry and Clinical Pharmacology, University Hospital Bonn, Bonn, Germany

<sup>2</sup>Institute of Structural Biology, University Hospital Bonn, Bonn, Germany

<sup>3</sup>Department of Oncology, Hematology and Rheumatology, University Hospital Bonn, Bonn, Germany

<sup>4</sup>Institute of Experimental Haematology and Transfusion Medicine, University Hospital Bonn, Bonn, Germany

<sup>5</sup>Unit of Experimental Immunology, Department of Biomedical Sciences, Institute of Tropical Medicine, Antwerp, Belgium

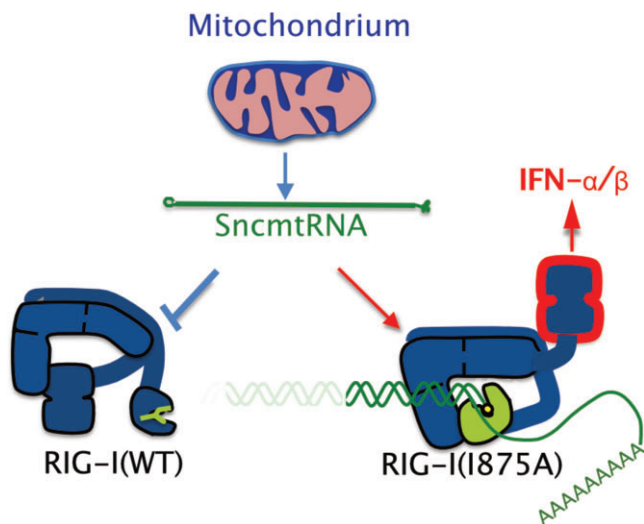
<sup>6</sup>Department of Biomedicine, Aarhus University, Aarhus, Denmark

\*To whom correspondence should be addressed. Tel: +49 228 287 16081; Fax: +49 228 287 51160; Email: martin.schlee@uni-bonn.de

## Abstract

RIG-I is a cytosolic receptor of viral RNA essential for the immune response to numerous RNA viruses. Accordingly, RIG-I must sensitively detect viral RNA yet tolerate abundant self-RNA species. The basic binding cleft and an aromatic amino acid of the RIG-I C-terminal domain (CTD) mediate high-affinity recognition of 5'-triphosphorylated and 5'-base-paired RNA (dsRNA). Here, we found that, while 5'-unmodified hydroxyl (OH)-dsRNA demonstrated residual activation potential, 5'-monophosphate (5'p)-termini, present on most cellular RNAs, prevented RIG-I activation. Determination of CTD/dsRNA co-crystal structures and mutant activation studies revealed that the evolutionarily conserved I875 within the CTD sterically inhibits 5'p-dsRNA binding. RIG-I(I875A) was activated by both synthetic 5'p-dsRNA and endogenous long dsRNA within the polyA-rich fraction of total cellular RNA. RIG-I(I875A) specifically interacted with long, polyA-bearing, mitochondrial (mt) RNA, and depletion of mtRNA from total RNA abolished its activation. Altogether, our study demonstrates that avoidance of 5'p-RNA recognition is crucial to prevent mtRNA-triggered RIG-I-mediated autoinflammation.

## Graphical abstract



Received: January 25, 2023. Revised: September 1, 2023. Editorial Decision: September 19, 2023. Accepted: September 21, 2023

© The Author(s) 2023. Published by Oxford University Press on behalf of Nucleic Acids Research.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License

(<http://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact [journals.permissions@oup.com](mailto:journals.permissions@oup.com)

## Introduction

The cytosolic nucleic acid receptors of the innate immune system are crucial for the initiation and control of the antiviral immune response in infected organisms. Receptor activation induces antiviral signaling pathways leading to the expression of type-I IFN, cytokines, chemokines and antiviral effector proteins. Since endogenous nucleic acids are also present in the host, pathogen-derived nucleic acids are generally sensed via their aberrant localization, e.g. cytosolic DNA, and/or the presence of uncommon/non-self structures or modifications, known as recognition motifs (1). Effective cytoplasmic RNA recognition is particularly challenging due to the physiological presence of a variety of endogenous RNA species. Here, the need to sensitively detect pathogenic RNA has to be balanced with the potential danger of ‘mistakenly’ sensing endogenous RNA (2). Whereas excessive tolerance favors the spread of infection, non-specific immune recognition of endogenous nucleic acids can lead to devastating autoinflammatory disease, such as Aicardi–Goutières Syndrome (AGS), Singleton–Merten Syndrome (SMS) and other type-I interferonopathies (3,4).

The members of the RIG-I like receptor family (RLR) – including RIG-I itself, MDA5 and LGP2—are capable of recognizing only a few copies of pathogenic RNA in the cytosol among an excess of endogenous RNA species, such as rRNA, tRNA, and mRNA (5). rRNA and tRNA are highly structured and have 5′-monophosphate (5′p) termini. mRNA from higher eukaryotes bears a 5′-cap1 structure, consisting of a 5′ to 5′ triphosphate(ppp) linkage between N7-methylguanosine (m7Gppp), which has a 2′O-methylated penultimate nucleotide(N1) (m7GpppN(Ome), cap1). Although long RNAs with secondary structures activate MDA5 in a 5′independent manner (5) and are also reported to activate RIG-I (6), activation of RIG-I by short dsRNA critically depends on the structure and modifications of the 5′-terminus. Blunt-ended dsRNA bearing a 5′-triphosphate (5′ppp, including cap0) or 5′-diphosphate (5′pp) is sensitively detected by RIG-I, while a cap1 structure abrogates activation (7–11). Moreover, 5′OH-dsRNA also activates RIG-I, albeit less potently and only at high concentrations and with slower activation kinetics (6,12). Since free 5′ppp, 5′pp, cap0 and 5′OH RNA ends are normally absent from the cytosol of non-infected vertebrate cells, they represent logical danger signals for RIG-I to detect (1). Conversely, active avoidance of RIG-I activation by self-structures (cap1 and 5′p) is also a plausible immune discriminatory strategy. Indeed, we previously identified such a mechanism for cap1 tolerance located within the CTD of RIG-I (10). However, the effect of 5′p on RIG-I activation has been controversial in the literature: While one group reported an enhancing effect for 5′p on RIG-I activation due to RNA stabilization (13), another, more recent, study observed that 5′p inhibited RIG-I binding and ATPase activity (14).

In the currently accepted model, prior to RIG-I activation, an inhibitory helicase subdomain (Hel-2i) binds the caspase activation and recruitment domains (CARDs), thus suppressing CARD signaling (Supplementary Figure S1A) (15,16). However, when pathogenic dsRNA binds to the flexibly connected CTD of RIG-I, the RNA ligand is oriented to form interactions with the helicase domains Hel-1 and Hel-2i, resulting in CARD displacement, homotypic CARD–CARD interaction with the adaptor protein MAVS and downstream signaling to type I IFN induction (Supplementary Figure S1A). Based

on this model, effective interaction of dsRNA with the CTD is a prerequisite of RIG-I activation. Indeed, several studies from our group and others have characterized how structural features of the CTD critically determine which RNA molecules can bind and activate RIG-I (17–19): 5′ppp/pp is recognized by a basic binding cleft (H847, K858, K861, K888) within the CTD, and the presence of a 5′-basepair enables a stacking interaction with F853 (Supplementary Figure S1B). Interestingly, ppp-dsRNA with 5′overhang that lacks a 5′-basepair considerably binds RIG-I but is not able to activate RIG-I (20). Thus, binding is required but not sufficient to activate RIG-I. Such ligands with non-productive binding fail to enable the release of the CARD domain. More recent studies described additional levels of RIG-I ligand recognition (‘proofreading’) involving CARD and helicase domains (21,22)

In the current study, we investigate the mechanism by which the 5′pRNA interaction with RIG-I is inhibited, thus preventing recognition of self RNA. We demonstrate that, while the RIG-I CTD can bind 5′ppp-dsRNA with high affinity and 5′OH-dsRNA with moderate affinity, 5′p-dsRNA-CTD binding was barely detectable. The molecular basis for this discrimination is an isoleucine residue (I875) in the RIG-I CTD that precludes 5′p-dsRNA binding. Co-crystal structures of 5′OH-dsRNA and 5′p-dsRNA with wt and I875A-mutated RIG-I CTD, respectively, show the steric conflict of 5′p-dsRNA with I875 compared to 5′OH-dsRNA. Correspondingly, I875A mutation completely restored RIG-I activation by 5′p-dsRNA to the level of 5′OH-dsRNA, and long-term expression of RIG-I(I875A) resulted in robust RIG-I activation by endogenous RNA. Moreover, within total RNA, we could identify mitochondrial RNA as the most active RNA species stimulating RIG-I(I875A). Our findings indicate that this active avoidance of recognition of endogenous 5′p-RNA is crucial to preventing RIG-I mediated autoinflammation, thus characterizing a new mechanism by which RIG-I discriminates between self and non-self RNA.

## Materials and methods

### Experimental model and subject details

#### Primary cells

Human PBMCs were isolated from buffy coats of healthy human donors and cultivated in RPMI 1640 medium (10% FCS, 100 U/ml penicillin, 100 mg/ml streptomycin) at 37°C.

#### Cell lines

A549 cells were cultivated in DMEM (10% FCS, 100 U/ml penicillin, 100 mg/ml streptomycin) at 37 °C. HEK Flp-In™ T-REx™ 293 were cultivated in DMEM (10% FCS, 100 U/ml penicillin, 100 mg/ml streptomycin) at 37°C.

### Methods details

#### Cell culture

Human PBMCs were isolated from whole human blood of voluntary, healthy donors by Ficoll-Hypaque density gradient centrifugation (Merck, Darmstadt, Germany). For stimulation experiments,  $4 \times 10^5$  PBMCs were cultured in 96-well plates. The PBMC studies were approved by the local ethics committee (Ethikkommission der Medizinischen Fakultät Bonn) according to the ICH-GCP guidelines. To inhibit TLR7/8 activity, cells were pre-incubated with 5 µg/ml chloroquine for

30 min. Cells were kept in RPMI 1640 medium (10% fetal calf serum (FCS), 100 U/ml penicillin, 100 mg/ml streptomycin).

All transfections were performed using Lipofectamine 2000 (LF2000, Thermo Fisher Scientific, Darmstadt, Germany). The transfection reagent was mixed with OptiMEM and was incubated for 5 min at room temperature (RT). RNA or DNA was mixed with an equal amount of OptiMEM and combined with the OptiMEM-transfection reagent mix. After 20 min of incubation at RT, the transfection mixture was added to the cells.

HEK Flp-In™ T-REx™ 293 cells (Thermo Fisher Scientific) were kept in DMEM (10% FCS, 100 U/ml penicillin, 100 mg/ml streptomycin). For stimulation experiments,  $3.5 \times 10^6$  cells were cultured in 96-well plates. Protein and RNA samples were harvested from 24-well plates that had been seeded with  $2 \times 10^5$  cells. Doxycycline (100 µg/ml or as indicated) was added to the cells in the plate 20 min before the experiments to induce the expression of RIG-I.

#### Generation of stable HEK Flp-In™ T-REx™ 293 expression cell lines

The HEK Flp-In™ T-REx™ 293 host cells were seeded in 6-well culture plates ( $1 \times 10^6$  cells in 3 ml/well) and allowed to adhere overnight. The following day 700 ng pOG44 plasmid, expressing the Flp recombinase, was transfected simultaneously with 70 ng of pcDNA5/FRT plasmid containing the gene of interest using LF2000. After 48 h of incubation, the cells were trypsinized, resuspended in selection media containing 150 µg/ml hygromycin and 10 µg/ml blasticidin and seeded in new 6-well plates. After the selection process, expression of the protein of interest was confirmed by functional assays and western blot.

#### Generation of gene-edited cell lines

HEK Flp-In™ T-REx™ 293 cells were transfected with LF2000 with 200 ng of a Cas9-gRNA expression plasmid targeting RIG-I (5-GGGTCTTCCGGATATAATCC(TGG)-3) or MDA5 (5-CTTCTAGTTAGAGACGTCT TGG (TGG)-3). One day after transfection cells were cultivated in limited dilution to obtain single-cell clones. After expansion knock-out (KO) was confirmed by immunoblot, Sanger sequencing and functional testing. To generate A549-RIG-I(I875A) point mutant cell lines, A549 cells were co-electroporated with Cas9-2A-EGFP-gRNA expression plasmid targeting 5'-CAAATGTCTTGTACTTCACA(TGG)-3' (100 µg/ml) a nd ssDNA repair oligo 5'-AAGATATTCTGTGCCCGACAG AACTGCAGCCATGACTGGGGAGCTCATGTGAAGTA CAAGACATTTGAGATTCCAGTTATAAAAAATTGAA-3' (5µM, I875A mutation bold, generates *SacI* site, underlined), using the Neon Transfection System (Thermo Fisher Scientific). Settings were as recommended by the manufacturer (1230 V, 30 ms, 2 pulses). 18h after electroporation, EGFP-positive cells were enriched by FACS-sort (BD FACS Aria III Sorter, Flow Cytometry Core Facility Bonn). Single-cell clones were obtained by limiting dilution and gene-editing was confirmed by *SacI*-digestion and Sanger sequencing of the amplified genomic site (fw 5'-TGGTTCTAACGTTCTCTTTTGTGTG-3', rev 5'-AGCGATCCATGATTATACCCACT-3') and functional testing. Clone genotypes are listed in Figure 2D.

#### Generation of RIG-I mutants

Point-mutations in full-length RIG-I were introduced using PCR amplification of plasmids containing the RIG-I WT gene with mutated primers (see table S3) and Gibson assembly. The mutant constructs were confirmed by sequencing. Protein expression of RIG-I mutants in HEK Flp-In™ T-REx™ 293 cells was compared by immunoblot with wild-type RIG-I using an antibody against RIG-I (Cell signaling, Frankfurt am Main, Germany).

#### Generation of mtDNA - and mtRNA-depleted cells

A549 cells were cultivated in DMEM containing ethidium bromide (50 ng/ml), sodium pyruvate (1 mM), uridine (50 µg/ml) and non-essential amino acid solution (NEAA; 1x) for at least 7 days. Successful mtRNA depletion was measured via gene expression analysis of mt-mRNA (ND1, ND5) using qPCR. Control cells were incubated in medium containing sodium pyruvate (1 mM), uridine (50 µg/ml) and NEAA (1x), but without EtBr.

#### Detection of cytokines

IFN-α levels were analyzed with commercial ELISA assay kits (Thermo Fisher Scientific) according to the manufacturer's protocol. The concentration of IFN-α present in the sample was determined by a standard curve, which was determined using known amounts of recombinant cytokines.

#### Dual luciferase assay

HEK Flp-In™ T-REx™ 293 were seeded in 96-well plates and were allowed to adhere overnight. The cells were then co-transfected with 10 ng *Gaussia* luciferase reporter plasmid (pGL3-IFNB1-GLuc) and 5 ng control Firefly Luciferase expression plasmid for normalization (pLenti-Ef1α-FLuc) (23). After at least 6 h, the cells were stimulated with the RNA ligands indicated in the figures. The luciferase assay was performed 20 h after stimulation. To measure *Gaussia* luciferase production, as a surrogate for IFN-β promoter activation, 25 µl coelenterazine-solution (1 µg/ml in H<sub>2</sub>O) was added to 25 µl supernatant in a white microplate. Chemiluminescence was immediately measured with the EnVision 2104 Multilabel plate reader (Perkin Elmer, Waltham, USA). To measure Firefly luciferase as a control, the supernatant was aspirated and 40 µl 1x SAP-solution (0.02% Saponin, 50 mM Tris HCl pH 7.8, 15 mM MgSO<sub>4</sub>, 4 mM EGTA, 10% glycerol) was added to the cells and then they were lysed for 20 min at RT on a shaker. Afterwards, 25 µl of lysed cells in SAP were added to 25 µl Firefly luciferase substrate (Luciferin) and luminescence was measured with the EnVision 2104 Multilabel plate reader.

#### Preparation of RNA

Triphosphorylated RNA oligonucleotides were chemically synthesized as described (24).

5'OH-RNA and 5'p-RNA were ordered from Biomers (RNA sequences are listed in Table S1). 60mer, 100mer, and 200mer RNA were transcribed using *in vitro* transcription (IVT, TranscriptAid™ T7 High Yield Transcription Kit, Thermo Fisher Scientific).

Whole-cell RNA (total RNA) was extracted using TRIzol™ reagent from A549 cells (Thermo Fisher Scientific) according to the manufacturer's instruction. polyA-RNA extraction:

Purification of polyA-containing RNA from the whole-cell RNA was performed with the NucleoTrap® mRNA Kit



(Macherey Nagel, Düren, Germany) according to the manufacturer's instructions.

### rRNA extraction

rRNA was extracted from whole cell RNA extracts (total RNA) using the RiboMinus™ Kit or by excision and extraction of the 18s and 28s rRNA bands from an agarose gel.

Extraction with the Ribominus™ Human/Mouse Transcriptome Isolation Kit (Thermo Fisher Scientific) was performed according to the manufacturer's instructions. The rRNA<sup>-</sup> fraction was isolated using columns while the rRNA<sup>+</sup> fraction was isolated from the beads using TRIzol extraction.

To extract rRNA using gel electrophoresis, whole cell RNA extracts were run on a 1% agarose gel in RNase free conditions. RNA bands according to the size of the 18s and 28s rRNA were cut from the agarose gel and extraction was performed using the NucleoSpin® Gel and PCR Clean-Up Kit (Macherey-Nagel). The extraction was performed according to the manufacturer's protocol for RNA extractions from agarose gels using the buffer NTC (Macherey Nagel).

### Small RNA extraction

Small RNA (<200 nucleotides) was extracted from whole cell RNA extracts (total RNA) using the miRNeasy Mini Kit and the RNeasy MiniElute Cleanup Kit (both Qiagen, Hilden, Germany). Extraction was performed according to the manufacturer's instructions.

### RIP (RNA immunoprecipitation)

HEK293 FT cells were cultured in DMEM (supplemented with 10% (v/v) FCS and 1% (v/v) Penicillin/Streptomycin) until they reached 70% confluency in a 10 cm cell culture dish. Transfection with the corresponding plasmids (RIG-I(WT) and RIG-I(I875A)) was conducted by using Lipofectamine2000. In short, 20 µg plasmid were mixed with 480 µl OptiMEM and 40 µl Lipofectamine2000 were mixed in a second tube with 460 µl OptiMEM. The plasmid mix was slowly added to the Lipofectamine2000 solution and briefly resuspended. After 10 minutes incubation at room temperature the solution was added dropwise to the cells in the 10 cm cell culture dish with 9 ml culture medium. The cells were incubated for 24 h at 37 °C with 5% CO<sub>2</sub> in an incubator. Subsequently, the cells were two times carefully washed with cold PBS and collected with a cell scraper. After centrifugation at 4 °C for 5 min at 200 × g the pellet was resuspended in 500 µl NP40 cell lysis buffer (50mM HEPES, pH 7.5, 150 mM KCl, 2 mM EDTA, 0.5 mM DTT, 0.5% (v/v) NP-40, protease inhibitors) and incubated on ice for 15 minutes. The solution was cleared by centrifugation at 4 °C for 5 min at 20 000 × g and the supernatant was transferred to a 2 ml reaction tube. To check for IP quality 50 µl were mixed with 500 µl TRIzol (Thermo Fisher Scientific) for input RNA and 50 µl were mixed with 10 µl 6xSDS loading dye for western blot analysis. The sample was filled up to 900 µl with NP40 cell lysis buffer for better binding during rotation. An amount of 30 µl anti-FLAG magnetic beads (Sigma M8823) per reaction were washed 2 times with 1 ml NP40 cell lysis buffer and resuspended in 100 µl NP40 cell lysis buffer. Afterwards, 100 µl anti-FLAG magnetic beads were added to each sample and incubated at 4 °C on a rotating wheel for 16 h. The anti-FLAG magnetic beads were washed 5 times with 1 ml NP40 cell lysis buffer and collected by a magnetic rack. 100 µl were saved for western blot IP analysis. The remaining anti-FLAG magnetic beads were col-

lected, the supernatant was removed and 500 µl TRIzol was added. The sample was stored at -20 °C. Subsequent RNA isolation and cDNA synthesis were conducted following the manufacturer's instructions.

The data analysis was performed following the ΔCt method. In short, the IP values were normalized to GAPDH levels to obtain the ΔCt value. Afterwards, the fold enrichment was calculated by 2<sup>-(ΔCt)</sup>. Finally, the fold enrichments were normalized to the non-specific control (anti-magnetic FLAG beads in non-transfected WT cells) to determine the fold changes of all analyzed targets. Significances were calculated by Student's *t* test. The recovered yield of the RIP was calculated to check for IP quality. Western blot analysis using anti FLAG antibody (Sigma F1804) was performed to check for RIG-I recovery after IP.

### Enzymatic treatment of RNA

FastAP Alkaline Phosphatase (Thermo Fisher Scientific) was used to remove 5'-phosphates from RNA samples to generate 5'-OH-RNA. 0.25 µg RNA was incubated in 1× FastAP reaction buffer containing 1 U FastAP for 30 min at 37°C.

RNA 5'Polyphosphatase (Lucigen, Middleton, USA) was used to generate 5'-p-RNA from IVT 5'ppp-RNA. 5 µg RNA were incubated in 1× polyphosphatase reaction buffer containing 20 U RNA 5'Polyphosphatase for 30 min at 37°C according to the manufacturer's protocol.

1 µg of whole-cell RNA was treated with 1 µl RNase A (Thermo Fisher Scientific) in low salt (50 mM NaCl) or high salt (200 mM NaCl) for 30 min at 37°C.

### qPCR

RNA from HEK Flp-In™ T-REx™ 293 and A549 cells was extracted using Zymo III columns (Zymo, Freiburg, Germany). Reverse transcription was performed using random hexamer primers (IDT) and RevertAid Reverse Transcriptase (Thermo Fisher Scientific). qPCRs (with the exception of SncmtRNA) was performed using 5× EvaGreen qPCR-Mix II (ROX) (Bio-Budget, Krefeld, Germany). Primers (Table S2) were validated with melting curve analysis and tested for efficiency using cDNA dilution. SncmtRNA qPCR was performed using Luna® Universal qPCR Master Mix (New England Biolabs). Primetime® Mini qPCR Assay FAM/Iowa Black FQ labeled (Integrated DNA Technologies, Leuven, Belgium) (Table S2) was tested for efficiency using cDNA dilution. The SncmtRNA qPCR product, which spans a part of the loop connecting sense and antisense mt-rRNA was validated by sanger sequencing (Seqlab) to be 100% in sequence with DQ386868. qPCR was performed on a QuantStudio™5 Real-Time PCR System (Thermo Fisher Scientific) using standard Taqman Relative Quantification protocol.

### Immunoblot

2 × 10<sup>5</sup> HEK Flp-In™ T-REx™ 293 cells were seeded in 24-well plates and RIG-I expression was induced by addition of doxycycline. 72 h later cells were lysed in 60 µl 1 × Laemmli-buffer (120 mM Tris pH 6.8, 4% (w/v) SDS, 20% (v/v) glycerol, 20 mM DTT and Orange G). Samples were sonicated and heated (95°C, 5 min) prior to loading. Primary and secondary antibodies are listed in the KeyResources table. Blots were recorded on an Odyssey FC Dual-Mode Imaging system (LI-COR Biosciences, Bad Homburg, Germany).

### Preparation of recombinant proteins

The coding sequence of RIG-I C-terminal domain (CTD) (802–925) and RIG-I (802–925, I875A) was cloned into pGEX-4T1 using *Nco*I and *Eco*R1 restriction sites at the 5' and 3' end, respectively. Proteins were expressed in *E. coli* BL21(DE3) cells overnight at 22 °C by induction with 0.4 mM IPTG. Cells were lysed by sonication using 20 mM Tris/HCl, pH 7.5, 200 mM NaCl, 2 mM MgCl<sub>2</sub>, 1 mM DTT, 12.5 µg/ml RNase A, spatula tip of lysozyme, and DNase. GST-RIG-I protein variants were affinity purified by GSTrap column (GE Healthcare). After sample loading, the column was washed with 20 mM Tris/HCl, pH 7.5, 1 M NaCl, 2 mM MgCl<sub>2</sub>, 1 mM DTT. Protein was eluted using 20 mM Tris/HCl, pH 7.5, 200 mM NaCl, 2 mM MgCl<sub>2</sub>, 1 mM DTT, and 10 mM GSH. For protein crystallization, pooled fractions were incubated with TEV protease (5 µg/ml) overnight at 4°C to cleave off the GST-tag. Sample was purified by S75 size exclusion column (GE Healthcare) and fractions containing RIG-I protein were concentrated to 10 mg/ml (RIG-I CTD(WT)) or 27 mg/ml (RIG-I CTD(I875A)), aliquoted in storage buffer, snap frozen in liquid nitrogen, and stored at –80°C.

### Preparation of SncmtRNA *in vitro* transcripts

We hybridized two *in vitro* transcripts (IVT) to obtain a SncmtRNA like structure (see also Figure 5A): Templates for *in vitro* transcription were generated by using 5'primers binding at position + 10 (#1) or adding 3 nucleotides (GCC, #2) upstream of the SncmtRNA sequence (DQ386868) to obtain templates that are transcribable by T7 polymerases and that harbor a basepaired 5'end (like the original transcript) (see Figure 5, Table S4). The reverse primer binds at 809. For the second IVT template (#3) covering the loop and antisense part of SncmtRNA, a forward primer binding at 810 and reverse primer starting at 2374 were used. Purified PCR products were used for T7 *in vitro* transcription according to the manufacturer's instructions (Thermoscientific). IVTs were treated with 5'polyphosphatase, producing 5'p-RNA from 5'ppp-RNA. IVT#1 or IVT#2 were hybridized with IVT#3.

### Surface plasmon resonance spectroscopy

Surface plasmon resonance experiments were performed using a Biacore 8K instrument (GE Healthcare). The flow system was cleaned using the maintenance 'Desorb' function (Desorb Kit, GE Healthcare). The system was flushed with running buffer (20 mM Tris pH 7.4, 150 mM NaCl, 5 mM KCl, 1 mM CaCl<sub>2</sub>, 1 mM MgCl<sub>2</sub>, 0.05% Tween20) and all steps were performed at 25°C chip temperature. For GST-RIG-I CTD (wt or I875A) capturing, a CM5 sensor chip was covalently modified by an anti-GST polyclonal goat antibody (GST capture Kit, GE Healthcare). The surface of flow cell 1 and 2 was activated for 60 s with 50 mM NaOH (10 µl/min), followed by activation of both flow cells for 7 min with a 1:1 mixture of 0.1 M NHS (N-hydroxysuccinimide) and 0.1 M EDC (3-(N,N-dimethylamino) propyl-N-ethylcarbodiimide) (10 µl/min). Before amine-coupling, the flow system was washed with 1 M ethanolamine pH 8.0. The anti-GST antibody (30 µg/ml in 10 mM sodium acetate pH 5.0, 5 µl/min) was immobilized on flow cell 1 and 2. Surfaces were blocked with a 7 min injection of 1 M ethanolamine pH 8.0 (10 µl/min). For blocking of high affinity sites, recombinant GST

was injected for three consecutive 180 s cycles (30 µl/min, 200 mM) followed by 120 s regeneration with 10 mM glycine pH 2.0 (30 µl/min). GST-RIG-I (wt or I875A) was immobilized for 240 s on flow cell 2 (4 µl/min, 2 µM) and recombinant GST for 90 s on reference flow cell 1 (10 µl/min, 200 nM). Before interaction measurements, the sensor chip was stabilized for 10 min with running buffer (30 µl/min).

For kinetic binding measurements the analyte was injected (30 µl/min, association: 60 s, dissociation: 240 s) over both flow cells at concentrations of 128, 256, 512, 1024, 2048 and 4096 nM (p/OH-dsRNA) or 0.25, 0.5, 1, 2, 4, 8 nM (3p-dsRNA). After each cycle, the surfaces were regenerated with a 120 s injection step of 10 mM glycine pH 2.1. Data were collected at a rate of 10 Hz. The data were double referenced by blank cycle and reference flow cell subtraction. Processed data were fitted by 1:1 interaction model using Biacore Insight Evaluation Software (version 2.0. 15.12933).

### Crystallization, structure solution and refinement

The purified and concentrated RIG-I (wild type, 10.3 mg/ml) was stored in 20 mM Tris–HCl pH 7.5, 50 mM NaCl, 1 mM dithiothreitol (DTT). Initial crystals of wild-type RIG-I in complex with RNA (OH-GFP2 Palindrome, 12mer) were obtained using the hanging drop vapor diffusion technique at 20°C by mixing 1 µl protein-complex (1:1 ratio) solution with 1 µl of the reservoir solution containing 0.1 M sodium cacodylate pH 6.5, 0.2 M NaCl, 22% PEG 3350 at 10°C. After optimization, the best hexagonal-shaped crystals (dimensions of ~0.3 × 0.3 × 0.4 mm) were obtained by using 100 mM Tris–HCl pH 8.0, 0.2M sodium chloride, 18% PEG 3350 as the reservoir and grew in ~9–11 days.

The crystals of mutated RIG-I (I875A, 20mg/ml) in complex with RNA (p-GFP2 Palindrome, 12 mer) were obtained using 1:1 ratio of protein–RNA. The optimized crystals were grown using 0.2 M thiocyanate pH 8.4, 20% PEG 3350, 10% ethylene glycol and 10% MPD in the reservoir solution.

Diffraction data of wild type RIG-I in complex with RNA (OH-GFP2, 12mer) were collected at the Synchrotron SOLEIL Proxima-I, Saint-Aubin, France. Data were collected at 0.97857 Å wavelength and 100 K temperature. Diffraction data of mutant RIG-I(I875A) in complex with RNA (P-GFP2, 12mer) were collected at PX-II beamline at Swiss Light Source Villigen at 0.9999097 Å wavelength and 100 K temperature using the EIGER detector.

The XDS package (25) was used to process, integrate and scale all the data. The structure of the 5'-p-dsRNA CTD I875A complex was solved by molecular replacement with PHASER (McCoy) using PDB-ID 3NCU as search model. The three monomers were found in three separate runs with a final TFZ-score of 29.3 and LLG of 1378.08. The structure was refined using phenix.refine (26) leading to R/Rfree values of 21.8/27.3. The 5'-OH-dsRNA CTD structure was also solved with PHASER (TFZ 28.4, LLG 926) and refined using phenix.refine with twin law (h,-h-k,-l; twin fraction 0.23) leading to R/Rfree values of 19.3/21.8. The geometry of the structures was checked and optimized with COOT and MOLPROBITY (27,28) (Supplementary Figure S3A).

### RNA-FISH

DNA oligo probes were labelled by digoxigenin-11-dUTP (DIG-11-dUTP, Jena Bioscience, Jena, Germany) tailing. For

that, 10 pmol of an equimolar mixture of four probes were incubated along with 1 nmol of DIG-11-dUTP and 10 U of terminal transferase according to the manufacturer's instructions.

A549 cells stably expressing mito-mCherry were seeded at 5000 cells per well in an optical black clear-bottom 96-well plate and cultured in medium supplemented with 50 µg/ml uridine with or without 10 µM of IMT1 (MedChem Express, Monmouth Junction, NJ, USA). After 72 h, medium was removed, cells were washed with PBS and fixed for 10 min at RT with 4% (w/v) PFA in PBS. Cells were permeabilised by incubation for 5 min in 0.5% (v/v) TritonX-100 in PBS at RT, afterwards blocked with Roti ImmunoBlock (Carl Roth, Karlsruhe, Germany), for 30 min at RT and then equilibrated for 10 min in FISH wash buffer (2× SSC (300 mM NaCl, 30 mM trisodium citrate) with 10% (v/v) formamide). Hybridisation with probes was carried out overnight at 37°C in a humidified chamber with 5 nM DIG-11-dUTP labelled probes in hybridisation buffer (FISH wash buffer supplemented with 10% (w/v) dextran sulfate and 150 µg/ml yeast tRNA). For visualisation of polyA RNA, 5 nM of a Cy5-labelled oligo(dT) probe was added in addition. After hybridisation, wells were washed two times with FISH wash buffer for 30 min at 37°C and once with 2× SSC. Subsequently, cells were incubated with primary rabbit anti-DIG antibody (1:500 in Roti ImmunoBlock) for 1 h at RT. Wells were washed 4× with PBS and secondary Alexa Fluor 488-conjugated anti-rabbit antibody was added (1:500 in Roti ImmunoBlock) together with 5 µg/ml DAPI for nuclear stain and incubated for 1 h at RT. Wells were washed 6× with PBS and stored in PBS supplemented with 2.5 mg/ml DABCO for imaging. Images were acquired via a Leica DMi8 epifluorescence microscope equipped with a HC PL APO CS2 63×/1.20 water-immersion objective and a Leica DFC9000 GT camera. The same microscope instrument settings were used for all samples. Fiji was used for image processing and profile plot generation (29).

#### Stable expression of mito-mCherry in A549 cells

A549 cells with stable expression of the transitpeptide of the *Neurospora crassa* ATP synthase subunit 9 (SU9) N-terminally fused to mCherry (referred to as mito-mCherry) under control of an EF1α promoter were generated via lentiviral transduction. Pseudotyped HIV-1-based lentiviral vector was produced by transfection of 293FT cells with lentiviral transfer vector (pLenti6-EF1α-SU9-mCherry), psPAX2 second generation lentiviral packaging plasmid as well as pMD2.G VSV-G envelope expressing plasmid using calcium phosphate. Virus was harvested from the culture supernatant after 40 h. Cleared, sterile-filtered supernatant was used to transduce cells by incubating A549 cells seeded at low density for three days in lentivirus-containing supernatant mixed 3:2 with fresh culture medium. After 72 h, cells were selected with 10 µg/ml Blasticidin S (InvivoGen) for four days after which single-cells were plated by limiting dilution. Single-cell clones showing strong and homogeneous mitochondrial mCherry signal across the whole cell population were stepwise expanded.

#### Quantification and statistical analysis

Statistical analysis was performed using GraphPad Prism 6 and 8 for Mac OS X. Statistical parameters are reported in the figure legends.

## Results

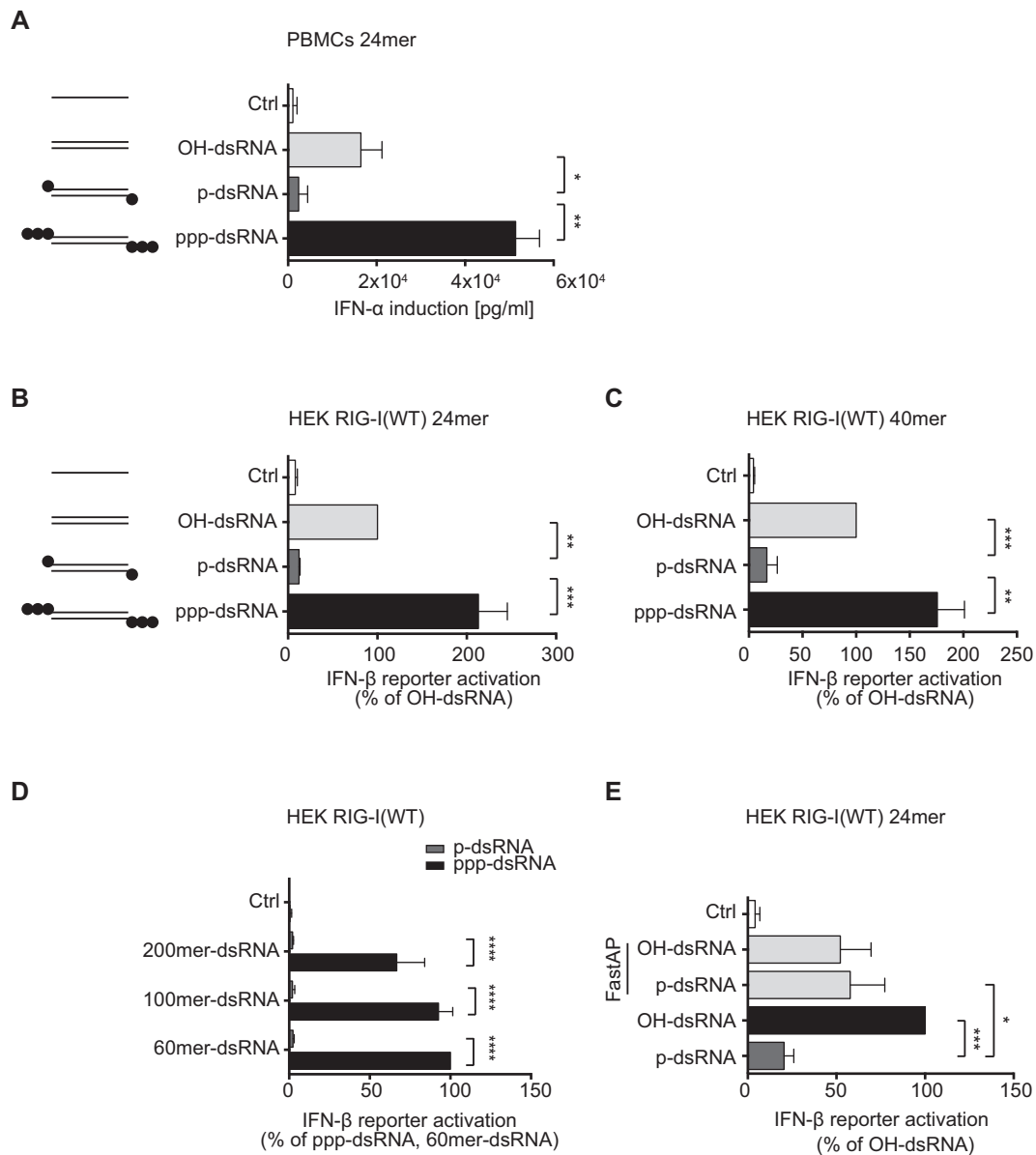
### Monophosphates at the 5' end of short and long dsRNA sterically inhibit recognition by RIG-I

Whether 5'p-dsRNA has the ability to stimulate RIG-I activity has remained controversial (13,14). Although Ren et al. detected weaker binding by 5'p-dsRNA than by 5'OH-dsRNA, the study did not demonstrate a significant difference in type-I-IFN induction by 5'p-dsRNA and 5'OH-dsRNA. We compared the activity of synthetic, blunt-ended 24mer and 40mer dsRNA with 5'ppp, 5'p, or 5'OH modifications at both ends. Here, using RIG-I/MDA5 deficient (RM<sup>-/-</sup>) T-REx-293 HEK cells expressing RIG-I(WT) and peripheral blood mononuclear cells (PBMCs) which were treated with chloroquine to block TLR7/8 activity, we found a substantial reduction of RIG-I stimulation by 5'p modification in comparison to 5'OH (Figure 1A–C). We also tested the sequences used by Takahashi et al., (13), which possessed 5'p on one strand of the dsRNA, and could not observe a significant stimulatory effect of 5'p over 5'OH-dsRNA in our cells (Supplementary Figure S1C, “6 + 4”). Of note, the dsRNA molecule used in the Takahashi et al. study has both a 5'p and a 5'OH terminus. In contrast, the presence of a 5'p modification at *both* ends of the dsRNA molecule almost completely abolished RIG-I activation (Supplementary Figure S1C, “6 + 7”). Taken together, this indicates that the presence of one non-phosphorylated end is sufficient to activate RIG-I, while the other 5'p end might have a different, auxiliary effect, such as increasing RNA stability (as discussed by (13)), leading to an enhancement of RIG-I stimulation of this sequence in some cell types or species. To investigate the impact of 5'p modification on longer dsRNA, we compared 60mer, 100mer and 200mer dsRNA (*in vitro* transcribed (IVT)RNA) with 5'ppp, and 5'p modifications (Figure 1D). 5'p-dsRNA was generated by treatment of IVT RNA with polyphosphatase, a pyrophosphatase that leaves 5'p ends after dephosphorylation of 5'ppp- or 5'ppRNA. Polyphosphatase treatment completely abolished RIG-I activation for all lengths of dsRNA. However, enzymatic removal of 5'p with alkaline phosphatase (FastAP) restored immune activation, demonstrating that the loss of RIG-I activation with 5'p-dsRNA is indeed due to the 5'p terminus (Figure 1E).

### Isoleucine 875 in the RIG-I CTD interferes with 5'p-dsRNA binding

Compared to 5'OH dsRNA, 5'p-dsRNA has two additional negative charges that could theoretically facilitate its interaction with the positively charged binding cleft. Thus, it was surprising that 5'p-dsRNA was less stimulatory than dsRNA without 5'phosphate. To identify the underlying structural mechanisms of this phenomenon, amino acids in the RIG-I CTD closer than 4 Å to the α-phosphate of the RNA were identified using the crystal structure of the RIG-I CTD with 5'pp-dsRNA (3NCU;(19)). With the exception of I875, all of the other amino acids identified (namely, K858, K861, K888 and H847) have been previously shown to be involved in binding to the triphosphate, and mutation of any of these four amino acids decreases or completely abrogates immune stimulation with 5'ppp-dsRNA (19). Although, I875 is highly conserved among species (Supplementary Figure S2A), it has not been implicated in 5'ppp binding so far, and the hydrophobic side chain of isoleucine could instead have repulsive effects on charged molecules like 5'phosphates. To test this





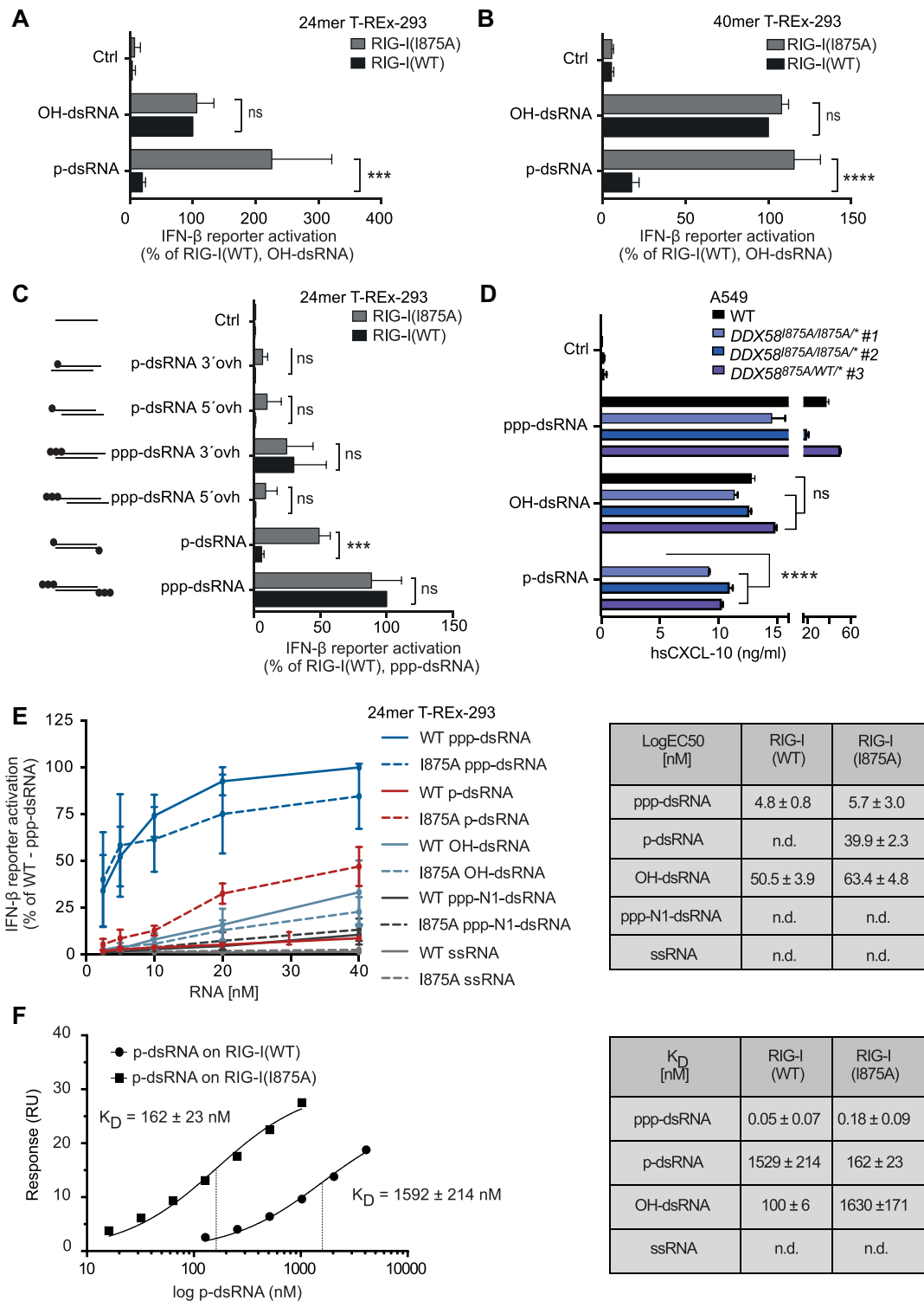
**Figure 1.** 5'p blocks RIG-I-dependent immune stimulation by short and long dsRNA. **(A)** Chloroquine pre-treated PBMCs were stimulated with the indicated synthetic RNA ligands (0.1  $\mu\text{g/ml}$ ). Concentrations of IFN- $\alpha$  in the supernatant were determined 20 h after transfection by ELISA. **(B, C)** HEK T-REx *RIG-I/MDA5(R/M)*<sup>-/-</sup> cells expressing RIG-I(WT) were stimulated with the indicated synthetic RNA ligands (0.1  $\mu\text{g/ml}$ ). IFN- $\beta$  reporter activation was determined 20 h after RNA transfection and is depicted as percentage of the signal of RIG-I(WT) stimulated by OH-dsRNA. **(A, B)** 24mer dsRNAs and **(C)** 40mer dsRNAs. **(D)** HEK T-REx *RM*<sup>-/-</sup> cells expressing RIG-I(WT) were stimulated with either untreated (ppp-dsRNA) or polyphosphatase-treated (p-dsRNA) IVT-transcribed dsRNA ligands (1 ng/ml) of the indicated lengths (60, 100 or 200mer). **(E)** HEK T-REx *RM*<sup>-/-</sup> cells expressing RIG-I(WT) were stimulated with the 24mer synthetic RNA ligands indicated (0.1  $\mu\text{g/ml}$ ). IFN- $\beta$  reporter activation was determined 20 h after RNA transfection and is depicted as percentage of the signal of RIG-I(WT) stimulated with D) 60mer ppp-dsRNA and (E) untreated 24mer OH-dsRNA. 'Fast AP' indicates alkaline phosphatase treatment of the RNA prior to transfection. (A–C, E)  $n = 4 + \text{SD}$ , one-way ANOVA, Dunnett's post hoc test, (D)  $n = 4 + \text{SD}$ , two-way ANOVA, Sidak's post hoc test. Statistical significance is indicated as follows: \* ( $P < 0.05$ ), \*\* ( $P < 0.01$ ), \*\*\* ( $P < 0.001$ ) \*\*\*\* ( $P < 0.0001$ ). 5'-modifications: 5'-hydroxyl (OH-), 5'-monophosphate (p-, ●), 5'-triphosphate (ppp-, ●●●). RNA sequences: Table S1.

hypothesis, RIG-I/MDA5 deficient (*RM*<sup>-/-</sup>) HEK cells expressing either wild-type RIG-I (WT) or RIG-I with I875 mutated to alanine (I875A) from a doxycycline-inducible locus (T-REx) were stimulated with 5'p-dsRNA and 5'OH-dsRNA. Indeed, the RIG-I(I875A) mutant could be activated by 5'p-dsRNA at a level similar to 5'OH-dsRNA for both sequences tested (24mer and 40mer, Figure 2A and B).

To test if mutation of I875 caused a general increase in immune activation, cells were stimulated with RNA with other inhibitory features, such as 5' or 3' overhangs (Figure 2C).

Moreover, to exclude artefacts from overexpressed RIG-I, we then generated A549 cells with the I875A mutation in the endogenous DDX58 locus (Figure 2D), demonstrating that RIG-I(I875) mutant protein expressed from the endogenous locus was also stimulated by p-dsRNA.

All other RNAs tested showed a similar immunostimulatory potential for wild-type RIG-I(WT) and RIG-I(I875A) cells, demonstrating the specific effect of RIG-I(I875A) 5'p-dsRNA recognition. Moreover, while the EC50 value for 5'p-dsRNA stimulation of RIG-I(WT) could not be determined



**Figure 2.** I875 in the RIG-I CTD prevents p-dsRNA binding by steric exclusion. **(A–C)** HEK T-REx RM<sup>-/-</sup> cells expressing either RIG-I(WT) or RIG-I(I875A) were stimulated with indicated synthetic RNA ligands (0.1 μg/ml). **(A, C)** 24mer dsRNA **(B)** 40mer dsRNA with indicated end-modifications and overhangs. IFN-β reporter activation was determined 20 h after RNA transfection and is depicted as percentage of the signal of RIG-I(WT) stimulated by **(A, B)** OH-dsRNA or **(C)** ppp-dsRNA. *n* = 3 + SD, two-way ANOVA, Sidak's post hoc test. **(D)** A549 DDX58(WT) and DDX58(I875A) gene-edited cell lines were stimulated with the RNA ligands indicated. CXCL10 from supernatants was measured after 24 h. \*\*\*\* *P* < 0.0001, two-way ANOVA, Dunnett's multiple comparisons test, brackets depict three comparisons versus WT condition. Shown are means of duplicates + SEM of one experiment, representative of two independent experiments. **(E)** RNA stimulation titration curves of HEK T-REx RM<sup>-/-</sup> cells expressing RIG-I(WT) or RIG-I(I875A) stimulated with indicated RNAs (c[RNA] = 0.0625–1 μg/ml). IFN-β reporter activation was determined 20 h after RNA transfection and is depicted as percentage of the signal of RIG-I(WT) stimulated by ppp-dsRNA. EC<sub>50</sub> values were calculated by nonlinear regression. n.d. = not determined (*n* = 3 + SE). **(F)** Dissociation constants of indicated RIG-I CTD/RNA complexes determined by surface plasmon resonance (SPR). Statistical significance is indicated as follows: \*\*\* (*P* < 0.001), \*\*\*\* (*P* < 0.0001), ns: not significant. 5'-modifications: 5'-hydroxyl (OH-), 5'-monophosphate (p-), 5'-triphosphate (3p-). N1-2'-O-methylated 5'-terminal nucleotide (N1). RNA sequences: Table S1.



owing to a lack of immune activation, the EC50 value for RIG-I(I875A) with 5'p-dsRNA was comparable to that of 5'OH-dsRNA in RIG-I(WT) and RIG-I(I875A) cells (Figure 2E). Compared with I875, mutation of another amino acid (N668) that has previously been associated with p-dsRNA discrimination (14) showed little effect on 5'p-dsRNA recognition (Supplementary Figure S2B). With the ligands used in our study, the RIG-I(N668A) mutant responded to p-dsRNA, OH-dsRNA and ppp-dsRNA in a manner similar to RIG-I(WT).

To confirm that the bulky side chain of I875 prevents CTD-5'p-dsRNA binding, we performed binding studies using surface plasmon resonance (SPR). WT CTD and I875A CTD had  $K_D$  values in the nanomolar range for 5'ppp-dsRNA while no binding was detected for ssRNA (Figure 2F). WT CTD binding to 5'OH-dsRNA was 15 times stronger than its binding to 5'p-dsRNA. Binding of 5'p-dsRNA to I875A CTD was 10 times stronger compared to WT CTD (Figure 2F). Unexpectedly, I875A mutation slightly reduced 5'ppp-dsRNA binding and considerably impaired binding of 5'OH to the RIG-I(I875A) CTD without changing the EC50 of OH-dsRNA stimulated full-length RIG-I(I875A) compared to RIG-I(WT).

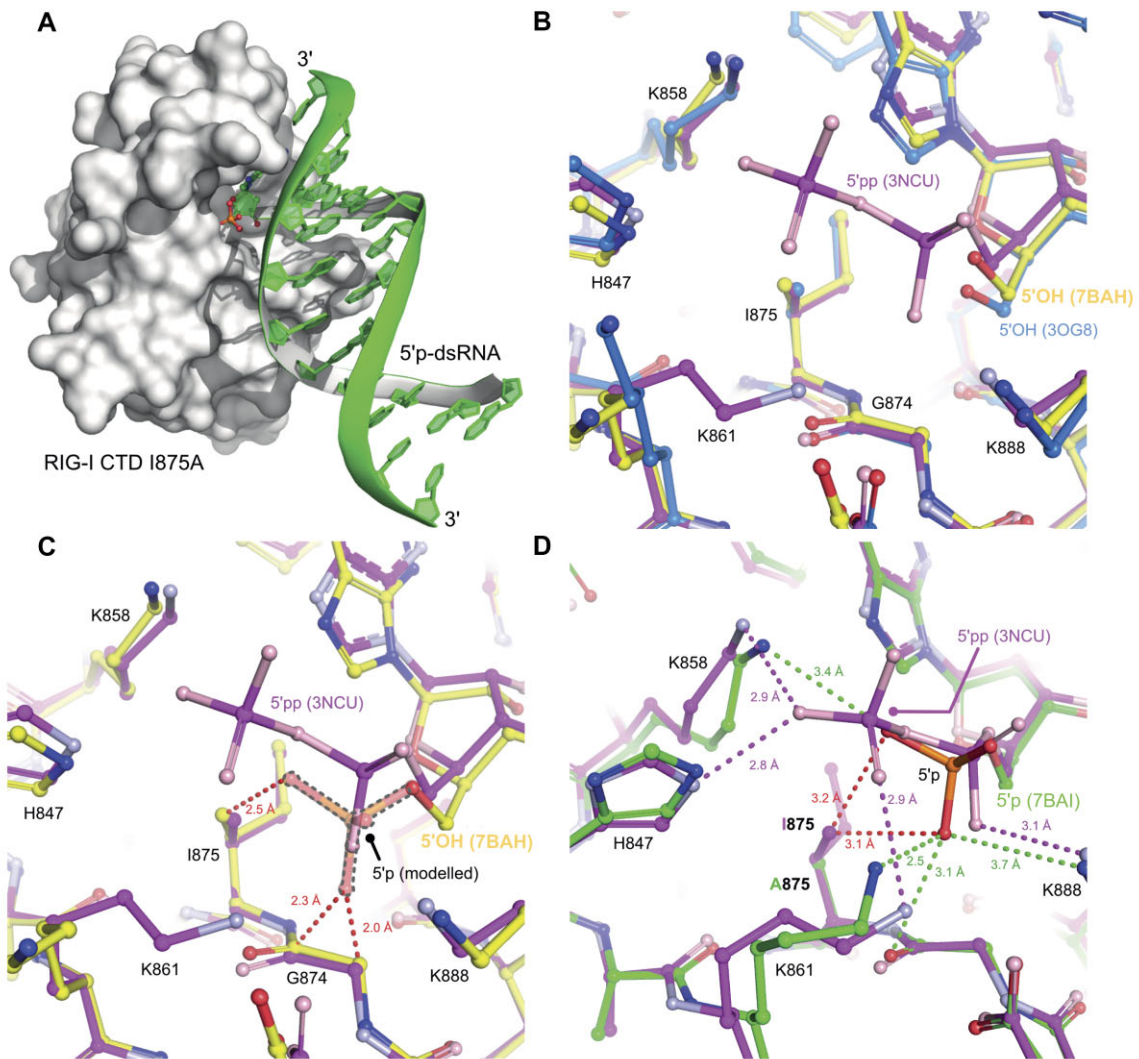
To corroborate the binding results, we determined two crystal structures: A 1.89 Å co-crystal structure of the human RIG-I(WT) CTD with 5'OH-dsRNA (7BAH) and a 3.4 Å structure of the RIG-I(I875A) CTD mutant in complex with 5'p-dsRNA (7BAI; Supplementary Figure S3A). Of note, we were not able to generate crystals of the RIG-I(WT) CTD with 5'p-dsRNA. Overall, the two 7BAH and 7BAI structures are very similar (Figure 3A). While the resolution of the 5'p-dsRNA structure was low (3.4 Å), a simulated annealing omit map unequivocally confirmed the presence of the 5'p group in the binding pocket of the RIG-I(I875A) CTD (Supplementary Figure S3B). Figure 3B shows a comparison of the RIG-I(WT) CTD/5'pp-dsRNA co-crystal structure (3NCU; (19)) that of RIG-I(WT) CTD/5'OH-dsRNA (7BAH, this work) and of a previously determined RIG-I(WT) CTD/5'OH-dsRNA structure with a different RNA sequence and length (3OG8 (17,18)). The 5'pp-dsRNA is held in position by numerous polar and ionic interactions between both the  $\alpha$ - (K861, K888) and  $\beta$ -phosphate (H847, K858, K861). Due to the absence of phosphate groups, 5'OH-dsRNA cannot interact with these residues but instead sinks deeper into the binding site. By modeling a 5'p onto the 5'OH-dsRNA structure, we determined that the 5'p would sterically clash with both I875 and G874 in this configuration (Figure 3C). The lack of interactions with the  $\beta$ -phosphate and the steric repulsion from I875 and G874 thus explain why 5'p does not productively bind to the pocket of the WT CTD. This was corroborated by our RIG-I(I875A) CTD/5'p-dsRNA structure (Figure 3D). Here, the less bulky side chain (of A875) together with numerous polar and ionic interactions with residues K888, G874, K861 and K858 allow the 5'p to be accommodated in the dsRNA binding site without steric clashes. Additionally, as revealed by the structure (7BAI), removal of the I875 sidechain enabled interaction of the amino group of K858 with the alpha phosphate. Thus, I875 seems to ensure selectivity for 5'diphosphates over 5'monophosphates by preventing the interaction of K858 with the alpha phosphate. Compared to the 5'pp complex structure (3NCU), some sidechain rearrangements were observed as shown in Figure 3D. Here, it should be noted that the 5'p has an additional negative charge at the position of the neutral phospho-anhydride bond of the 5'pp, which further strengthens the ionic interaction with the amino group of K858 (Figure 3, Supplemen-

tary Figure S3C). It should be noted that, due to the low resolution of the 5'p I875A structure, it is not possible to study the influence of the solvent molecules that are involved in the interaction.

### Tolerance of 5'p prevents recognition of endogenous mtRNA

Since the most abundantly expressed endogenous RNAs (rRNA and tRNA) are processed into molecules with 5'p termini, we speculated that the exclusion of structured 5'p-RNA by WT RIG-I is a mechanism to differentiate between self and non-self RNA. To test the degree of stimulation of RIG-I by endogenous RNA, expression of RIG-I(WT), RIG-I(I875A), and, as a control, RIG-I(H830A), a mutant which senses N1-2'O-methylated 5'pppRNA (10), were induced in RIG-I/MDA5 deficient (RM<sup>-/-</sup>) T-REx-293 cells using an integrated doxycycline inducible locus. Cell-intrinsic IFN signaling was measured via IFIT1 mRNA transcription and IRF3 phosphorylation levels 72 h after doxycycline treatment and without the addition of exogenous ligands (Figure 4A). Of note, after doxycycline treatment, the H830A and I875A mutants were expressed at the same levels as WT RIG-I (Supplementary Figure S4A, B). Moreover, as expected, long-term expression of RIG-I(H830A) caused an increase in immune activation compared to RIG-I(WT) (Figure 4A, Supplementary Figure S4D). However, this effect was even stronger in cells expressing RIG-I(I875A), which was strongly indicative of the presence of endogenous, monophosphorylated, base-paired RNA structures capable of stimulating RIG-I in the absence of a mechanism to actively exclude 5'p from the CTD. A549 cells expressing the RIG-I(I875A) mutant from the endogenous loci spontaneously secreted CXCL10 and induced an ISG signature including IFNB1, demonstrating the biological relevance of I875 to prevent RIG-I activation at baseline, homeostatic conditions (Figure 4B, C).

To identify the endogenous RNA species causing the observed RIG-I(I875A) activation, whole-cell RNA was isolated and transfected into the HEK T-REx cell lines. As expected, total RNA induced stronger IFNB promoter reporter activation in RIG-I(I875A) cells compared to RIG-I(WT) (Supplementary Figure S4E). To better characterize these endogenous RNA ligands, selective RNase A digestion was then employed: RNase A digests both ssRNA and dsRNA at low-salt concentrations (up to 100 mM NaCl), while at high-salt concentrations (>300 mM NaCl), it selectively digests ssRNA but not dsRNA. Whole-cell RNA lysate digested with RNase A in a high-salt buffer triggered the same level of IFNB promoter reporter activation as untreated RNA in HEK T-REx cells. However, total cellular RNA treated with RNase A under low-salt conditions reduced the immune stimulatory potential of the RNA significantly compared to untreated RNA, indicating that the stimulatory RNA consisted of base-paired structures (Figure 4D). Moreover, we detected RIG-I(I875A)-stimulating RNA within the fraction of long (>200nt) but not short RNAs (18–200 nt) (Figure 4E). rRNAs and tRNAs both form complex secondary structures with significant regions of base-pairing. Thus, to identify the specific stimulatory species within the total cellular RNA pool, RNA was fractionated and purified. tRNA isolated from yeast did not activate an immune response in either RIG-I(WT) or RIG-I(I875A) cells (Supplementary Figure S4F). To specifically test the stimulatory potential of 18S and 28S rRNA, total RNA was



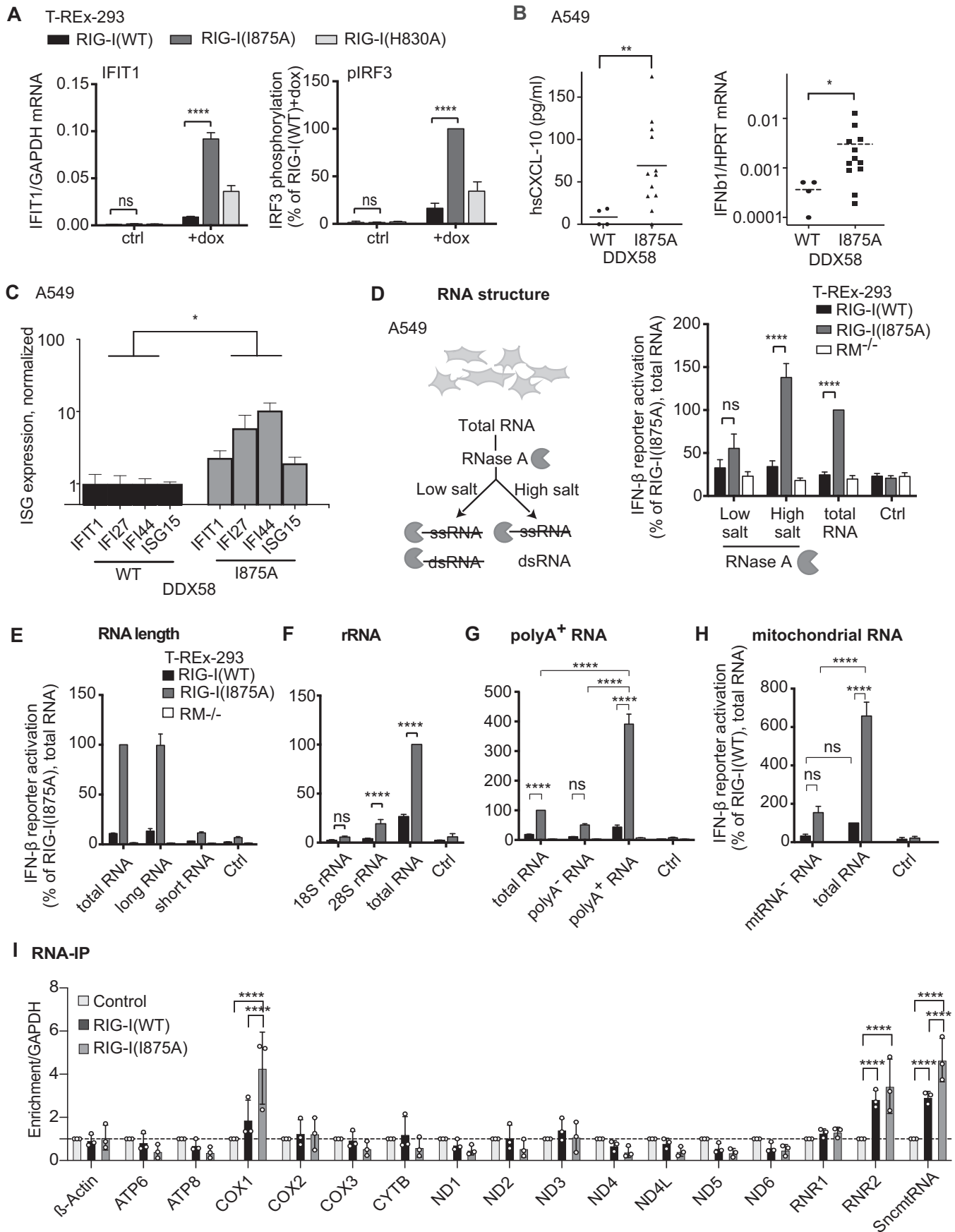
5'OH-dsRNA+RIG-I CTD(WT) : 7BAH

5'p-dsRNA+RIG-I CTD(I875A) : 7BAI

5'OH-dsRNA+RIG-I CTD(WT) : 3OG8

5'pp-dsRNA + RIG-I CTD(WT) : 3NCU

**Figure 3.** 5'p-dsRNA exclusion by the RIG-I CTD. **(A)** Overall structure of the RIG-I-CTD I875A/5'p-dsRNA complex. The CTD is shown as a white surface. The RNA is depicted in green with the terminal monophosphorylated base as a ball-and-stick model. The phosphate atom is colored orange and the phosphate oxygens are colored red. **(B)** The structural alignment of three ball-and-stick models of RIG-I CTD crystal structures: the 5'-OH-dsRNA/CTD WT complex structure (7BAH, this work, in yellow), the 5'pp-dsRNA/CTD WT structure (3NCU, in magenta), and a previously solved 5'-OH-dsRNA/CTD WT structure (3OG8, in blue) with a different RNA sequence. **(C)** Modelling a monophosphate onto our 5'OH-structure (7BAH, translucent) reveals that I875 and G874 together prevent binding of a monophosphorylated (5'p)-dsRNA. The red dashed lines indicate clashes between the modelled phosphate group and the CTD. **(D)** Crystal structure of the CTD I875A 5'p-dsRNA structure (7BAI, green ball-and-stick model) in comparison to the 5'pp-dsRNA/CTD WT structure (3NCU, magenta ball-and-sticks). Dashed lines indicate distances measured within the two structures. The red dashed lines indicate the very close contacts that would result from a monophosphate bound to the wild-type structure in the same configuration as found in the I875A mutant. RNA sequences: Table S1.



**Figure 4.** The 5'p tolerance mechanism prevents recognition of endogenous mtRNA by RIG-I. (A) RIG-I(WT), RIG-I(I875A) or RIG-I(H830A) expression was induced for 72 h with 10 pg/ml doxycycline (+dox) in HEK T-REx RM<sup>-/-</sup> cells in the absence of exogenous RIG-I stimuli. Left panel: RNA was



separated on an agarose gel and the 18S and 28S bands were excised and extracted. While stimulation with 18S rRNA had no significant effect on RIG-I(WT) or RIG-I(I875A) cells, 28S rRNA stimulation led to significant IFN- $\beta$  induction in RIG-I(I875A) cells (Figure 4F). Nonetheless, the observed IFN- $\beta$  induction was five times lower than stimulation with total RNA (Figure 4F). To exclude that this discrepancy resulted from the fractionation method used, a second rRNA extraction method was applied. Specific locked-nucleic acid (LNA) probes against 18S and 28S rRNA were used to pull-down rRNA from the total cellular RNA pool. Both rRNA isolated by this method (rRNA<sup>+</sup>) and the resultant fraction without rRNA (rRNA<sup>-</sup>) activated RIG-I(I875A) cells at a similar level, which was 6-fold higher than in RIG-I(WT) cells (Supplementary Figure S4G). While these data demonstrate that rRNA can contribute to the immune activation potential of total cellular RNA extracts, it also indicates that they are not the only immune-activating species within this RNA. Thus, we then examined the 3'-polyadenylated fraction of total RNA using polydT beads for RNA purification. mRNAs are the main polyadenylated RNA species, but rRNAs containing polyA tails have also been described (Slomovic et al., 2006). Like tRNAs and rRNAs, mRNAs are single-stranded but can form secondary structures with stretches of base-pairing. Of note, mRNAs have a cap1 structure (not a 5'p) and thus would not be expected to stimulate RIG-I(WT) or I875A. Surprisingly, the fraction containing the polyA RNA robustly induced IFN- $\beta$  reporter activation in RIG-I(I875A) cells, as compared to total RNA (Figure 4G). Accordingly, the fraction that had been depleted of polyA RNA demonstrated less IFN- $\beta$  reporter activation in comparison to total RNA (Figure 4G). Therefore, polyA RNA seems to be the main activating species of RIG-I(I875A).

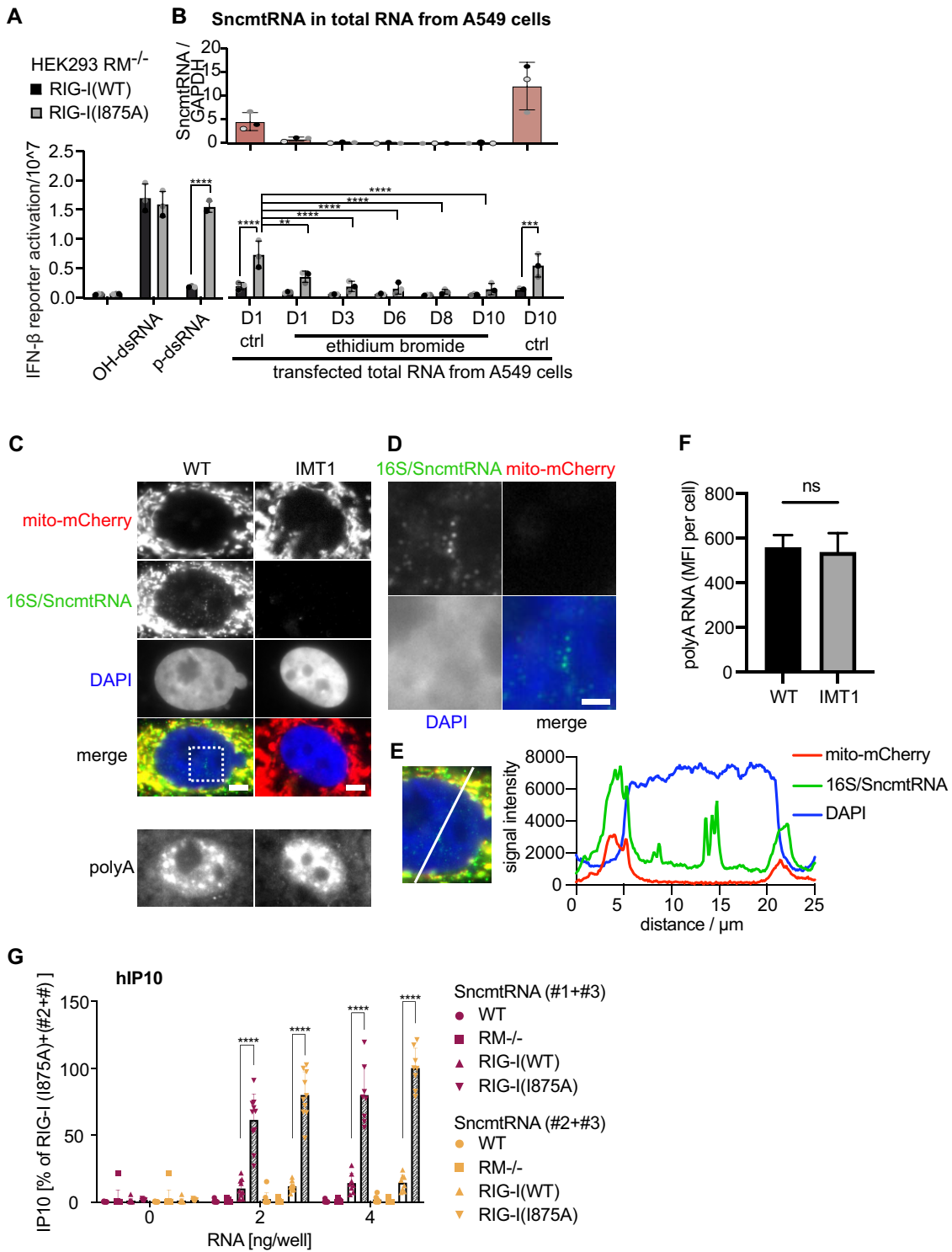
Although mRNA has a cap1 structure, thereby preventing activation of wild-type RIG-I and RIG-I(I875A), mitochondrial mRNAs (mt-mRNAs) are both uncapped and reported to possess a stable 3'-polyA tail (30,31). Indeed, we

could verify this observation via qPCR on the polyA RNA fraction, which revealed the presence of the mt-mRNAs ND1 and ND5 (Supplementary Figure S4H). Human mitochondria possess multiple copies of their 16.6-kb circular dsDNA genome that is bidirectionally transcribed and encodes 11 mt-mRNAs, two mt-rRNAs and 22 mt-tRNAs (32,33). The mt-tRNAs are excised from a long polycistronic precursor RNA by RNase P and Z, leading to the release of the mt-mRNAs, which are located between the tRNAs on the mitochondrial genomic DNA strands (34). RNase Z cleavage results in 3'OH and 5'P termini, thus generating 5'P mt-mRNAs, which are polyadenylated by a human nuclear-encoded mitochondrial poly(A) polymerase (33). Villegas *et al.* (35) identified mitochondria derived polyadenylated long non-coding RNAs (e.g. *SncmtRNA*; DQ386868.1) forming a hairpin with a long (200–800 bp) stem structure which was localized in the cytoplasm and nucleus of proliferating cells. Furthermore, the occurrence of double-stranded mtRNA that activates MDA5 and PKR has already been demonstrated (36,37).

To determine the contribution of mtRNA to the activation of RIG-I(I875A), we used ethidium bromide (EtBr) to deplete mtRNA levels. Long-term incubation of cells with low doses of EtBr causes a loss of mtDNA (and thus mtRNA) due to a decrease in its replication rate (38). Since HEK T-REx cells did not survive EtBr treatment (data not shown), A549 cells were used instead for mtRNA depletion. After incubation for at least seven days in medium containing EtBr, whole-cell RNA was extracted from these cells (mtRNA<sup>-</sup>) and from cells that were cultivated in a control medium (mtRNA<sup>+</sup>). The expression of rRNAs and nuclear-encoded mRNAs (GAPDH) did not change between RNA from EtBr-treated and mtRNA<sup>+</sup> A549 cells (Supplementary Figure S4I). However, expression of mt-mRNAs (ND1, ND5 and *SncmtRNA*) was greatly reduced in cells treated with EtBr, confirming the EtBr-mediated depletion of mt-DNA/RNA (Supplementary Figure S4I, J). RIG-I(WT) and RIG-I(I875A) cells were then transfected with the RNA from non-treated (mtRNA<sup>+</sup>) and EtBr-treated cells

extracted and mRNA expression of IFIT1 was analyzed by qPCR and normalized to GAPDH ( $n = 3-4$ , mean + SEM, two-way ANOVA, Tukey's post hoc test). Right panel: Phosphorylated IRF3 protein was analyzed by western blot and normalized to  $\beta$ -Actin ( $n = 4$ , mean + SEM, two-way ANOVA, Tukey's post hoc test). (B) Left panel: CXCL10 levels in WT and DDX58<sup>I875A/I875A</sup> mutant A549 cells (3 clones, 4 datapoints per clone), 24 h after medium exchange. Shown are cumulated results from two independent experiments. \*\*  $P < 0.01$ , Welch's  $t$ -test. Right panel: Spontaneous IFN- $\beta$ 1 mRNA expression in WT and DDX58<sup>I875A/I875A</sup> mutant A549 cells (3 clones) measured by qPCR. RNA was isolated from growing cultures twice per week. Shown are cumulated data from 4 timepoints, individual values and mean (dotted line). \*  $P < 0.05$ , Welch's  $t$ -test. (C) Spontaneous ISG signature of WT and DDX58<sup>I875A/I875A</sup> mutant A549 cells (3 clones) measured by qPCR. RNA was isolated from growing cultures twice per week. Shown are means + SEM of cumulated data from 4 timepoints per culture, normalized for the average of expression in WT. \*  $P < 0.05$ , Mann-Whitney test. (D) Total RNA from A549 cells was treated with RNase in low-salt (50 mM NaCl) or high-salt (300 mM NaCl) conditions or left untreated. RNA (1  $\mu$ g/ml) was transfected into HEK T-REx RM<sup>-/-</sup> cells expressing RIG-I(WT), RIG-I(I875A) or no RIG-I (RM<sup>-/-</sup>). IFN- $\beta$  reporter activation was determined 20 h after RNA transfection and is depicted as percentage of the signal of RIG-I(I875A) stimulated by non-treated total RNA ( $n = 5$ , mean + SEM, two-way ANOVA, Tukey's post hoc test). (E) Total RNA was separated into a short RNA(18–200nt) and a long RNA(>200nt) fraction. HEK T-REx RM<sup>-/-</sup> cells expressing either RIG-I(WT), RIG-I(I875A) or no RIG-I (RM<sup>-/-</sup>) were stimulated with the indicated RNA (1  $\mu$ g/ml). IFN- $\beta$  reporter activation was determined 20 h after RNA transfection and is depicted as percentage of the signal of RIG-I(I875A), stimulated with total RNA ( $n = 4$ , mean + SD). (F) Total RNA was separated on an agarose gel and 18s rRNA and 28s rRNA were extracted from the gel. HEK T-REx RM<sup>-/-</sup> cells expressing RIG-I(WT) or RIG-I(I875A) were transfected with total RNA and the extracted 18S rRNA and 28S rRNA (1  $\mu$ g/ml). IFN- $\beta$  reporter activation was determined 20 h after RNA transfection and is depicted as percentage of the signal of RIG-I(I875A), stimulated with total RNA ( $n = 4$ , mean + SD, two-way ANOVA, Sidak's post hoc test). (G) Total RNA was fractionated into a polyA<sup>+</sup> and polyA<sup>-</sup> fraction using polyA Beads. Fractionated RNA and total RNA (1  $\mu$ g/ml) was transfected into HEK T-REx RM<sup>-/-</sup> cells expressing RIG-I(WT), RIG-I(I875A) or no RIG-I (RM<sup>-/-</sup>). IFN- $\beta$  reporter activation was determined 20 h after RNA transfection and is depicted as percentage of the signal of RIG-I(I875A) stimulated by total RNA ( $n = 5$ , mean + SEM, two-way ANOVA, Tukey's post hoc test). (H) Total RNA was extracted using TRIzol from A549 cells that were cultivated in ctrl medium (total RNA) or ethidium bromide containing medium (mtRNA<sup>-</sup>) for at least 7 days. HEK T-REx RM<sup>-/-</sup> cells expressing RIG-I(WT) or RIG-I(I875A) were transfected with the RNA (1  $\mu$ g/ml). IFN- $\beta$  reporter activation was determined 20 h after RNA transfection and is depicted as percentage of the signal of RIG-I(WT) stimulated by total RNA ( $n = 3$  (RIG-I(WT) and RIG-I(I875A)), mean + SEM, two-way ANOVA, Tukey's post hoc test). (I) FLAG-RIG-I(WT) or FLAG-RIG-I(I875A) were overexpressed in HEK293FT cells. 24 h after transfection cells were lysed, FLAG-RIG-I proteins immunoprecipitated using anti-FLAG magnetic beads and bound RNA was extracted and used for qPCR ( $n = 3$  (RIG-I(WT) and RIG-I(I875A)), mean + SEM, two-way ANOVA, Tukey's post hoc test). Statistical significance is indicated as follows: \*\* ( $P < 0.01$ ), \*\*\* ( $P < 0.001$ ), \*\*\*\* ( $P < 0.0001$ ), ns: not significant. Dox: Doxycycline.





**Figure 5.** Mitochondrial 16S/SncmtRNA is localized to the nucleus and RIG-I(1875A) recognizes 5' monophosphorylated SncmtRNA *in vitro* transcripts. **(A)** 24mer OH-dsRNA, 24mer p-dsRNA or cellular RNA isolated from 1, 3, 6, 8 or 10 days EtBr- or untreated (ctrl) A549 cells were transfected into RIG-I(WT) or RIG-I(1875A) expressing HEK293 cells and RIG-I stimulation monitored by IFN-β reporter activation ( $n = 3$  (RIG-I(WT) and RIG-I(1875A)), mean + SEM, two-way ANOVA, Tukey's post-hoc test). **(B)** qPCR of SncmtRNA in total RNA from A549 cells used for stimulation in (A). **(C)** RNA-FISH of 16S/SncmtRNA in A549 cells. Stably expressed mito-mCherry served as mitochondrial marker and nuclei were stained with DAPI. PolyA RNA was visualised via a Cy5-labeled oligo-dT probe. Cells were cultured in absence (WT) or presence of the POLRMT inhibitor IMT1 before subjecting to RNA-FISH. Scale bar: 8 μm. **(D)** Magnification of nuclear region in (C), as indicated on the left. Scale bar: 2 μm. **(E)** Profile plot of DAPI (nucleus), mito-mCherry and 16S/SncmtRNA signals in the nuclear region, as indicated on the left. **(F)** Quantification of polyA signal of WT and IMT1-treated A549 cells. Mean fluorescence intensity (MFI) of  $n = 12$  cells per condition. Data were analysed via an unpaired *t* test. **(G)** *In vitro* transcripts corresponding to SncmtRNA as described in Supplementary Figure S5B were transfected into WT, RIG-I/MDA5 deficient (RM<sup>-/-</sup>) or RM<sup>-/-</sup> Flp-In™ T-REx™ 293 cells expressing RIG-I(WT) or RIG-I(1875A) upon addition of 0.1 μg/ml doxycycline. IP10 in supernatants was determined by ELISA 20 h after stimulation ( $n = 10$ , mean + SD, two-way ANOVA, Tukey's multiple comparisons test). Statistical significance is indicated as follows: \*\* ( $P < 0.01$ ), \*\*\*\* ( $P < 0.0001$ ), ns: not significant.

(mtRNA<sup>-</sup>). As expected, RNA from the mtRNA<sup>+</sup> cells led to an elevated immune response in RIG-I(I875A) expressing cells. However, this effect was largely abrogated by previous EtBr-treatment (Figure 4H). To identify mtRNA interacting with RIG-I, we expressed FLAG-RIG-I(WT) or FLAG-RIG-I(I875A) in HEK293T cells, immunoprecipitated FLAG-tagged proteins including bound RNAs from transfected cells or mock and analyzed RNAs from nuclear encoded house keeping genes and screened for all established mitochondrial genes by qPCR. While nuclear encoded RNAs (Actin, GAPDH and 28S rRNA) did not show enrichment (Figure 4I and Supplementary Figure S5A), three mitochondrial RNA transcripts (SncmtRNA, mt-RNR2 and COX1) were found to be enriched in pull-downs of FLAG-RIG-I(WT) or FLAG-RIG-I(I875A) expressing cells in comparison to mock (Figure 4I). Of note, qPCR of SncmtRNA and COX1 detected a significantly higher binding of COX1 and SncmtRNA to RIG-I(I875A). To further correlate mtRNA with RIG-I(I875A) stimulating activity in total RNA, we performed a time course of EtBr treatment of A549 cells and isolated total RNA from 1,3,6,8 or 10 days EtBr-treated or -untreated cells. The isolated total RNA was transfected into RIG-I(WT) or RIG-I(I875A) expressing HEK293 cells, and RIG-I stimulation was monitored by IFN- $\beta$  reporter activation (Figure 5A). RIG-I activation correlated well with SncmtRNA levels over different incubation times (Figure 5A,B), further supporting the notion that mtRNA is crucial for the observed RIG-I activation.

To evaluate if the detected mtRNAs are accessible by RIG-I in the intact cell we performed FISH with probes recognizing COX1 and RNR2/SncmtRNA. (Of note, SncmtRNA includes also the sequence of RNR2, which codes for the mt 16S rRNA subunit.) While COX1 did not show a signal outside of the mitochondria in the cytosol or in the nucleus (data not shown), we detected a robust FISH signal of RNR2/SncmtRNA out of mitochondria in the nucleus (Figure 5C–E). The specificity of the FISH signal was confirmed in cells treated with IMT1, an inhibitor of the mitochondrial RNA polymerase. IMT1 treatment abolished the RNR2/SncmtRNA signal (Figure 5C) without affecting the overall level of polyadenylated RNA (Figure 5C, F).

To further validate that monophosphorylated SncmtRNA is a potential RIG-I stimulating ligand, we synthesized sense and antisense strands by *in vitro* transcription which, after hybridization, simulate the structure of SncmtRNA. Since an IVT reaction cannot start with an uridine, we designed two different sense strands (1 or 2) that either start 10 nt behind the original start or have three additional nt upstream (5') of the uridine complementary to the antisense strand, yielding a 5'base-paired structure with 3'overhang, as found in the original sequence (Supplementary Figure S5B). 5'ppp was digested with polyphosphatase leaving a 5'monophorylated end. Here, only RIG-I(I875A) expressing cells showed a substantial immune response to monophosphorylated IVT SncmtRNA (Figure 5G).

Altogether, these results demonstrate that, within total RNA, polyadenylated mtRNA is the main species stimulating RIG-I(I875A) and that 28S rRNA has only a minimal to negligible role in the endogenous stimulation of RIG-I. Thus, we conclude that if RIG-I WT was not able to exclude 5'p-RNA from its binding site, mtRNA and particularly SncmtRNA would induce constitutive activation of RIG-I as observed for RIG-I(I875A) expressing cells, leading to chronic autoinflammation.

## Discussion

Most cytosolic antiviral innate immune receptors (PKR, OAS, RLR) recognize dsRNA (reviewed in (5)). Originally, it was presumed that dsRNA is a marker for non-self, as it was thought to only occur during viral replication and be otherwise absent from vertebrate cells (39). However, some recent studies, including our own, have exhibited the necessity for additional levels of control to prevent recognition of self-dsRNA by these receptors: ADAR1 modifies dsRNA stretches within polymerase II transcripts to abrogate base pairing that would otherwise activate PKR and MDA5 signaling (40–42); the RNA-degrading enzyme, SKIV2L, part of the RNA exosome, limits activation of RLR by endogenous RNA (43); the RNA helicase SUV3 and the RNA degrading enzyme polynucleotide phosphorylase (PNPase) restrict the levels of mitochondrial dsRNA that can stimulate MDA5 (36). Constitutive activation of RIG-I or MDA5 by endogenous RNA would likely have a phenotype similar to RIG-I or MDA5 gain-of-function mutations, which have been shown to cause autoimmune diseases like Singleton-Merten syndrome or Aicardi-Goutières syndrome (3,44). While MDA5 recognize very long dsRNA (>1000 bp) (45), RIG-I can detect short stretches of base-paired RNA (>20 bp) (8,12), the incidental occurrence of which is virtually unavoidable. Therefore, strict tolerance mechanisms for self-RNA are especially important for RIG-I activation. Our group previously discovered that a single amino acid, H830, in the RIG-I CTD is crucial to avoid activation by cap N1-2'O-methylated RNA (10). This mechanism mediates tolerance to endogenous cap1-RNA. In the present study, we find that the presence of a 5'-monophosphate terminus is sufficient to nearly completely prevent activation of RIG-I, even by long dsRNA. This finding clearly demonstrates that an entirely 5'terminus-independent RIG-I activation by long (>200 bp) dsRNA as suggested in a study comparing 5'ppp-dsRNA and 5'OH-dsRNA (6) does not exist. By systematically testing different endogenous RNA species, we could show that 5'p-discrimination has a strong impact on the recognition of mitochondrial RNA. Although tRNA and rRNA represent the largest fraction of 5'p-bearing RNA in the cytosol (up to 95%), our data, using depletion and enrichment of rRNA and monitoring of 28S rRNA specifically bound to FLAG-RIG-I pull-downs, indicate a rather subordinate role for (28S) rRNA in RIG-I(I875A) activation and no immune activation at all by tRNA. Here, it is plausible that the extensive internal post-transcriptional modifications of rRNAs and tRNAs prevent RIG-I binding and activation (46). Moreover, endogenous vault RNA, a short, structured 5'p-RNA, was described to stimulate RIG-I during viral infections if enriched as non-processed 5'ppp-RNA (47). However, we did not observe RIG-I(I875A) activation for the short RNA fraction (18–200 nt) from total RNA. In contrast, depletion of either polyadenylated RNA or mitochondrial RNA from isolated total RNA resulted in an almost complete loss of the RIG-I(I875A) signal. Of note, all mitochondrial RNA species with the exception of mt-tRNA are 3'polyadenylated (48). While nuclear-transcribed mRNA is capped at the 5'-end and therefore excluded from RIG-I binding, mt-mRNA lacks a 5'-cap structure. mtRNA is polycistronic and cleaved by RNase P and an RNase Z after transcription to form 5'p-mt-tRNA and 5'p-mt-mRNA. Because transcription of the circular mitochondrial genome is bidirectional, mtRNAs from opposite directions of transcription can form 5'p-dsRNA (37). The

5'p termini protect these mt-dsRNA species from wild-type RIG-I but not RIG-I(I875A) recognition. Additionally, long non-coding (lnc)-mtRNA, such as 'SncmtRNA' or 'ASncmtRNA', are generated by post-transcriptional events leading to the fusion of 16S mt-rRNA with complementary mtRNA and harbor a long stem (>800 bp) with 3'overhang. Such structures would be amenable targets for RIG-I if 5'p did not prevent RIG-I activation but not for MDA5, which requires more than 1000 bp (45). SncmtRNA is also polyadenylated and was found in the nucleus and cytosol of proliferating cells (35,49), indicating an active cytosolic transport and function in cell maintenance. Strikingly, we observed specific interaction of RIG-I(WT) and RIG-I(I875A) with the mitochondrial long non-coding (lnc)-RNA *SncmtRNA*, mt-rRNA (RNR2 = 16S) and COX1 mt-mRNA but not with 28S rRNA, housekeeping genes (GAPDH, Actin) or the other known mtRNAs (ATP6, ATP8, COX1, COX3, CYTB, ND1-6, RNR1).

In line with our findings, a previous study has demonstrated that most double-stranded cellular RNA originates from mtRNA (36). Of note, Kim et al. demonstrated that the cytosolic dsRNA binding receptor PKR is constantly activated by endogenous mtRNA (37), while Dhir et al. showed that MDA5 detects mtRNA if the enzymes involved in mtRNA degradation have been depleted (36). Since depletion of mtRNA from total RNA completely removed its RIG-I(I875A) stimulatory activity, we conclude that mt-dsRNA is the main RNA species activating RIG-I(I875A). While it is unclear if mt-mRNA is released at sufficient levels to activate RIG-I during cell maintenance, lnc-mtRNAs, which were described to reside in the cytosol and nucleus, are the more plausible RIG-I targets. Our data suggest that RIG-I would be constantly activated by mtRNA -at least in proliferating cells - if it lacked the ability to evade activation by 5'p-dsRNA. Of note, there are no RIG-I variants with SNP in amino acid 875 listed in the Genome Aggregation Database (gnomAD) of the Broad Institute. This may indicate that missense I875 mutations would have fatal autoinflammatory consequences for the individual and is in line with the exceptionally high evolutionary conservation of this amino acid. Interestingly, viral RNA can also bear a 5'p 'tolerance motif'. The (+)ssRNA genome of HCV is cleaved by RNase P, the same nuclease that also cleaves tRNA precursors and produces 5'p ends (50). Some (-)ssRNA viruses, including Crimean-Congo hemorrhagic fever virus (CCHTV), Hantaan virus (HTNV), and Bornavirus (BDV), cleave their 5'ppp-end to leave a 5'p-end (51). Thus, not only the absence of 5'ppp but also the presence of a 5'p contributes to the repression of the RIG-I mediated detection of these viruses, which is in line with the exploitation of other self tolerance motifs by other RNA viruses to avoid innate immune activation, e.g. N1-2'O-methylation of yellow fever RNA or cap-snatching by influenza.

Binding and activation assays revealed that 5'p-dsRNA exclusion results from the bulky methyl groups of I875 within the triphosphate binding pocket of RIG-I. Consequently, mutation of I875 to alanine facilitated binding and enabled activation of RIG-I by 5'p-dsRNA. RNAs with more than one phosphate, namely 5'pp-dsRNA and 5'ppp-dsRNA, are still able to bind RIG-I in the presence of I875. However, in this case, the additional, strong interactions between the amino groups of K861, H849, and K858 with the  $\beta$ -phosphate stabilize the ligand in a conformation that does not clash with I875. If 5'p-dsRNA were accommodated into the CTD in the same binding mode as 5'OH-dsRNA, as observed in 7BAI, the

5'p would clash with both I875 and the peptide backbone at G874. The proximity of I875 to the  $\alpha$ -phosphate or the clash of a modeled 5'p with G874 can also be observed in structures of human RIG-I(WT) CTD with 5'ppp-dsRNA (3LRN) or 5'OH-dsRNA (3OG8) with a different core sequence and length (16mer) (17,18). Thus far, we have no explanation why I875A mutation impairs binding of 5'OH-dsRNA to the RIG-I(I875A) CTD and why this does not have consequences for full-length RIG-I(I875A) activation by the same ligand.

The group of Fujita (13) described RIG-I stimulation by monophosphorylated dsRNA (5'p at one end, 5'OH at the other) in L929 cells. However, our data demonstrated that the presence of a 5'p at both ends of the same dsRNA sequence abolished the activation of RIG-I, thus strongly indicating that it was the 5'OH-end rather than the 5'p-end of that sequence that induced RIG-I activation. Additionally, the RNA-stabilizing effect that led to RIG-I stimulation in L929 cells, a murine fibroblast cell line, might not be present or relevant in the cells that we used in our study.

Ren *et al.* also described inhibition of RIG-I binding to 5'p-dsRNA when compared to 5'OH-dsRNA (14) but claimed a critical 'gatekeeper' role for N668 in the RIG-I helicase domain. However, considering that the first high affinity contact with the incoming RNA is made with the free CTD (before it has formed a complex with the helicase domain), it is unclear how amino acids outside of the CTD could plausibly take on the role of a 'gatekeeper'. Using 24mer dsRNA we could not detect a considerable difference between the RIG-I(WT) and RIG-I(N668A) response to p-dsRNA, OH-dsRNA or ppp-hpRNA. In contrast, we showed that I875A mutation in the CTD selectively allows 5'p-dsRNA binding and RIG-I activation, demonstrating the role of I875 in preventing 5'p-dsRNA binding. Furthermore, activation of other RNAs (5'OH-dsRNA, 5'ppp-dsRNA or RNA with overhangs) was unchanged between RIG-I(WT) and the RIG-I(I875A) mutant, highlighting the specificity for the exclusion of 5'p-dsRNA.

We can only speculate that the observed effect of the Ren group is due to the use of very short hairpin RNA and that the fourfold enhancement of *both* p-dsRNA and OH-dsRNA recognition observed for RIG-I(N668A) is due to a changed helicase unwinding or translocation activity of RIG-I that influences RIG-I activation or the stability of the RNA/RIG-I complex since N668 is close to T667, a residue critically involved in RIG-I translocation on the RNA (52). Furthermore, it cannot be excluded that N668 plays a role for discrimination of other ligand structures not analyzed in our work. In a follow up paper Wang & Pyle resolved the structure of full-length RIG-I(WT) with ppp-, pp-, p- and OH-dsRNA by cryo-EM(21). Here, they concluded that both, p-dsRNA as well as OH-dsRNA form a RIG-I-complex 'H<sub>F</sub>'(host RNA, Hel2 well folded) with high ATPase activity that favors transition from structure 3 to structure 2 (Supplementary Figure S1A) and therefore early release of the ligand p-dsRNA or OH-dsRNA while highly active RIG-I ligands (ppp-, pp-dsRNA) form a 'V<sub>U</sub>'(viral RNA, Hel2 unfolded) complex with low ATPase activity that favors signaling via MAVS. This mechanism would provide a second layer of discrimination after CTD mediated selection of high affinity binding ligands. The authors suggested that an autoinhibitory motif around N668 (Hel<sub>TNT</sub>), plays a critical role in H<sub>F</sub> versus V<sub>U</sub> formation (21). However, N668 is not a highly conserved amino acid in RIG-I (in vertebrates N668 can be replaced by D, S or T)



suggesting that asparagine specific side chain interactions may be not required for the Hel<sup>INT</sup>-function.

Comparison of our p-dsRNA/CTD structure (7BAI) with the corresponding region in the p-dsRNA/RIG(WT)-I-FL structure (7TNZ) reveals that the protein chains of the two structures superpose well with an r.m.s.d of 0.78 Å between 791 corresponding atoms while the relative orientation of the RNA chains is clearly different. Since our crystal contains only the CTD, we suspect that our crystal structure corresponds to the first contact of the free CTD with the 5' end of the dsRNA (see Supplementary Figure S1A structure 2). By contrast, since besides the CTD also other parts of RIG-I (helicase) are involved in p-dsRNA binding we conclude that the Wang&Pyle structure (7TNZ) corresponds to the situation after the CARDS are displaced by the CTD/dsRNA complex (see Supplementary Figure S1A structure 3). Interestingly, the 7TNZ structure shows another rotamer of I875 which in this arrangement provides more space for the 5'-monophosphate to be accommodated in the binding pocket. Here, the additional helicase interaction may support the assembly of the alternative I875 rotamer. Interestingly, also the OH-dsRNA/RIG-I(WT) structure (7TO0) but not the structures with the highly active ligands ppp/pp-dsRNA/RIG-I(WT) (7TNX, 7TNY) show the alternative I875 rotamer assembly in the binding pocket, indicating that I875 may also play a role in this second layer of ligand recognition.

Altogether, our study identifies a novel tolerance mechanism within the binding pocket of RIG-I that helps ensure that endogenous mitochondrial RNA does not lead to self-activation of the innate immune system. As, to our knowledge, RIG-I is expressed in almost all nucleated cells, inappropriate activation by mitochondrial RNA would have potentially devastating effects on the host. In line with previous studies (36,37), our study identifies endosymbiotic mitochondria as a source of unwanted, potentially stimulatory nucleic acids, underscoring the effort that the host cell has had to undertake to accommodate its ancient guest.

## Data availability

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Martin Schlee (martin.schlee@uni-bonn.de). Materials will be made available via material transfer agreement (MTA).

## Data and code availability

The crystal structures generated during this study are available at Protein Data Bank (7BAH, 7BAI).

## Supplementary data

Supplementary Data are available at NAR Online.

## Acknowledgements

We thank Meghan Campbell for her critical reading of the manuscript.

**Author contributions:** A.k.d.R., K.A., D.F., T.Z., B.P., K.G., A.G., S.W., D.H., K.C., F.B., A.K., C.H., C.W. and C.S.W. performed the experiments. A.k.d.R., K.A., K.G., D.H., K.C., F.B., A.K., S.W., E.B., K.P., J.L., T.Z., M.G. and M.S. con-

ceived experiments. A.k.d.R., K.G., E.B., K.P., G.H., Gr.H., T.Z., M.G. and M.S. analyzed and interpreted data and wrote the manuscript. All authors revised the manuscript.

## Funding

Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC2151 – 390873048 of which E.B., K.P., G.H., M.G. and M.S. are members; DFG, including TRR237 (E.B., T.Z., K.P., G.H., M.S.), SFB670 (E.B., G.H., M.S.), SFB704 (G.H.), GRK2168 (E.B., M.S.); DFG SCHL1930/1–2; E.B. and M.S. received financial support from BONFOR (University of Bonn); this work is part of the PhD thesis of AKdR at the University of Bonn and of the Bachelor thesis of Bastian Putschli at the Johannes Gutenberg University Mainz. Funding for open access charge: University Hospital Bonn.

## Conflict of interest statement

None declared.

## References

- Bartok,E. and Hartmann,G. (2020) Immune sensing mechanisms that discriminate self from altered self and foreign nucleic acids. *Immunity*, **53**, 54–77.
- Ahmad,S., Mu,X., Yang,F., Greenwald,E., Park,J.W., Jacob,E., Zhang,C.-Z. and Hur,S. (2018) Breaching self-tolerance to Alu duplex RNA underlies MDA5-mediated inflammation. *Cell*, **172**, 797–810.
- Rehwinkel,J. and Gack,M.U. (2020) RIG-I-like receptors: their regulation and roles in RNA sensing. *Nat. Rev. Immunol.*, **20**, 537–551.
- Rodero,M.P. and Crow,Y.J. (2016) Type I interferon-mediated monogenic autoinflammation: the type I interferonopathies, a conceptual overview. *J. Exp. Med.*, **213**, 2527–2538.
- Schlee,M. and Hartmann,G. (2016) Discriminating self from non-self in nucleic acid sensing. *Nat. Rev. Immunol.*, **16**, 566–580.
- Binder,M., Eberle,F., Seitz,S., Mucke,N., Huber,C.M., Kiani,N., Kaderali,L., Lohmann,V., Dalpke,A. and Bartenschlager,R. (2011) Molecular mechanism of signal perception and integration by the innate immune sensor retinoic acid-inducible gene-I (RIG-I). *J. Biol. Chem.*, **286**, 27278–27287.
- Hornung,V., Guenther-Biller,M., Bourquin,C., Ablasser,A., Schlee,M., Uematsu,S., Noronha,A., Manoharan,M., Akira,S., De Fougerolles,A., *et al.* (2005) Sequence-specific potent induction of IFN- $\alpha$  by short interfering RNA in plasmacytoid dendritic cells through TLR7. *Nat. Med.*, **11**, 263–270.
- Schlee,M., Roth,A., Hornung,V., Hagmann,C.A., Wimmenauer,V., Barchet,W., Coch,C., Janke,M., Mihailovic,A., Wardle,G., *et al.* (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.
- Goubau,D., Schlee,M., Deddouche,S., Pruijssers,A.J., Zillinger,T., Goldeck,M., Schuberth,C., Van Der Veen,A.G., Fujimura,T., Rehwinkel,J., *et al.* (2014) Antiviral immunity via RIG-I-mediated recognition of RNA bearing 59-diphosphates. *Nature*, **514**, 372–375.
- Schuberth-Wagner,C., Ludwig,J., Bruder,A.K., Herzner,A.-M., Zillinger,T., Goldeck,M., Schmidt,T., Schmid-Burgk,J.L., Kerber,R., Wolter,S., *et al.* (2015) A Conserved Histidine in the RNA Sensor RIG-I Controls Immune Tolerance to N1-2'O-Methylated Self RNA. *Immunity*, **43**, 41–51.
- Devarkar,S.C., Wang,C., Miller,M.T., Ramanathan,A., Jiang,F., Khan,A.G., Patel,S.S. and Marcotrigiano,J. (2016) Structural basis



- for m<sup>7</sup>G recognition and 2'-O-methyl discrimination in capped RNAs by the innate immune receptor RIG-I. *Proc. Natl. Acad. Sci. U.S.A.*, **113**, 596–601.
12. Marques, J.T., Devosse, T., Wang, D., Zamanian-Daryoush, M., Serbinowski, P., Hartmann, R., Fujita, T., Behlke, M.A. and Williams, B.R. (2006) A structural basis for discriminating between self and nonself double-stranded RNAs in mammalian cells. *Nat. Biotechnol.*, **24**, 559–565.
  13. Takahashi, K., Yoneyama, M., Nishihori, T., Hirai, R., Kumeta, H., Narita, R., Gale Jr., M., Inagaki, F. and Fujita, T. (2008) Nonself RNA-sensing mechanism of RIG-I helicase and activation of antiviral immune responses. *Mol. Cell*, **29**, 428–440.
  14. Ren, X., Linehan, M.M., Iwasaki, A. and Pyle, A.M. (2019) RIG-I selectively discriminates against 5'-monophosphate RNA. *Cell Rep.*, **26**, 2019–2027.
  15. Kolakofsky, D., Kowalinski, E. and Cusack, S. (2012) A structure-based model of RIG-I activation. *RNA*, **18**, 2118–2127.
  16. Kowalinski, E., Lunardi, T., McCarthy, A.A., Louber, J., Brunel, J., Grigorov, B., Gerlier, D. and Cusack, S. (2011) Structural basis for the activation of innate immune pattern-recognition receptor RIG-I by viral RNA. *Cell*, **147**, 423–435.
  17. Lu, C., Ranjith-Kumar, C.T., Hao, L., Kao, C.C. and Li, P. (2011) Crystal structure of RIG-I C-terminal domain bound to blunt-ended double-strand RNA without 5' triphosphate. *Nucleic Acids Res.*, **39**, 1565–1575.
  18. Lu, C., Xu, H., Ranjith-Kumar, C.T., Brooks, M.T., Hou, T.Y., Hu, F., Herr, A.B., Strong, R.K., Kao, C.C. and Li, P. (2010) The structural basis of 5' triphosphate double-stranded RNA recognition by RIG-I C-terminal domain. *Structure*, **18**, 1032–1043.
  19. Wang, Y., Ludwig, J., Schuberth, C., Goldeck, M., Schlee, M., Li, H., Juraneck, S., Sheng, G., Micura, R., Tuschl, T., et al. (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.
  20. Marq, J.B., Hausmann, S., Veillard, N., Kolakofsky, D. and Garcin, D. (2011) Short double-stranded RNAs with an overhanging 5' ppp-nucleotide, as found in arenavirus genomes, act as RIG-I decoys. *J. Biol. Chem.*, **286**, 6108–6116.
  21. Wang, W. and Pyle, A.M. (2022) The RIG-I receptor adopts two different conformations for distinguishing host from viral RNA ligands. *Mol. Cell*, **82**, 4131–4144.
  22. Schweibenz, B.D., Devarkar, S.C., Soltochi, M., Craig, C., Zheng, J., Pascal, B.D., Gokhale, S., Xie, P., Griffin, P.R. and Patel, S.S. (2022) The intrinsically disordered CARDs-Helicase linker in RIG-I is a molecular gate for RNA proofreading. *EMBO J.*, **41**, e109782.
  23. Gao, P., Ascano, M., Zillinger, T., Wang, W., Dai, P., Serganov, A.A., Gaffney, B.L., Shuman, S., Jones, R.A., Deng, L., et al. (2013) Structure-function analysis of STING activation by c[G(2',5')pA(3',5')p] and targeting by antiviral DMXAA. *Cell*, **154**, 748–762.
  24. Goldeck, M., Schlee, M., Hartmann, G. and Hornung, V. (2014) Enzymatic synthesis and purification of a defined RIG-I Ligand. *15-25*
  25. Kabsch, W. (2010) XDS. *Acta Crystallogr. D Biol. Crystallogr.*, **66**, 125–132.
  26. Liebschner, D., Afonine, P.V., Baker, M.L., Bunkóczi, G., Chen, V.B., Croll, T.I., Hintze, B., Hung, L.W., Jain, S., McCoy, A.J., et al. (2019) Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr. D Struct. Biol.*, **75**, 861–877.
  27. Emsley, P. and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta. Crystallogr. D Biol. Crystallogr.*, **60**, 2126–2132.
  28. Williams, C.J., Headd, J.J., Moriarty, N.W., Prisant, M.G., Videau, L.L., Deis, L.N., Verma, V., Keedy, D.A., Hintze, B.J., Chen, V.B., et al. (2018) MolProbity: more and better reference data for improved all-atom structure validation. *Protein Sci.*, **27**, 293–315.
  29. Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., et al. (2012) Fiji: an open-source platform for biological-image analysis. *Nat. Methods*, **9**, 676–682.
  30. Fernandez-Silva, P., Enriquez, J.A. and Montoya, J. (2003) Replication and transcription of mammalian mitochondrial DNA. *Exp. Physiol.*, **88**, 41–56.
  31. Ojala, D., Montoya, J. and Attardi, G. (1981) tRNA punctuation model of RNA processing in human mitochondria. *Nature*, **290**, 470–474.
  32. Anderson, S., Bankier, A.T., Barrell, B.G., de Bruijn, M.H., Coulson, A.R., Drouin, J., Eperon, I.C., Nierlich, D.P., Roe, B.A., Sanger, F., et al. (1981) Sequence and organization of the human mitochondrial genome. *Nature*, **290**, 457–465.
  33. Mai, N., Chrzanowska-Lightowlers, Z.M. and Lightowlers, R.N. (2017) The process of mammalian mitochondrial protein synthesis. *Cell Tissue Res.*, **367**, 5–20.
  34. Rossmann, W. (2012) Of P and Z: mitochondrial tRNA processing enzymes. *Biochim. Biophys. Acta*, **1819**, 1017–1026.
  35. Villegas, J., Burzio, V., Villota, C., Landerer, E., Martinez, R., Santander, M., Martinez, R., Pinto, R., Vera, M.I., Boccardo, E., et al. (2007) Expression of a novel non-coding mitochondrial RNA in human proliferating cells. *Nucleic Acids Res.*, **35**, 7336–7347.
  36. Dhir, A., Dhir, S., Borowski, L.S., Jimenez, L., Teitell, M., Rotig, A., Crow, Y.J., Rice, G.I., Duffy, D., Tamby, C., et al. (2018) Mitochondrial double-stranded RNA triggers antiviral signalling in humans. *Nature*, **560**, 238–242.
  37. Kim, Y., Park, J., Kim, S., Kim, M., Kang, M.G., Kwak, C., Kang, M., Kim, B., Rhee, H.W. and Kim, V.N. (2018) PKR senses nuclear and mitochondrial signals by interacting with endogenous double-stranded RNAs. *Mol. Cell*, **71**, 1051–1063.
  38. King, M.P. and Attardi, G. (1996) Isolation of human cell lines lacking mitochondrial DNA. *Methods Enzymol.*, **264**, 304–313.
  39. Weber, F., Wagner, V., Rasmussen, S.B., Hartmann, R. and Paludan, S.R. (2006) Double-stranded RNA is produced by positive-strand RNA viruses and DNA viruses but not in detectable amounts by negative-strand RNA viruses. *J. Virol.*, **80**, 5059–5064.
  40. Pfaller, C.K., Donohue, R.C., Nersisyan, S., Brodsky, L. and Cattaneo, R. (2018) Extensive editing of cellular and viral double-stranded RNA structures accounts for innate immunity suppression and the proviral activity of ADAR1p150. *PLoS Biol.*, **16**, e2006577.
  41. Tang, Q., Rigby, R.E., Young, G.R., Hvidt, A.K., Davis, T., Tan, T.K., Bridgeman, A., Townsend, A.R., Kassiotis, G. and Rehwinkel, J. (2021) Adenosine-to-inosine editing of endogenous Z-form RNA by the deaminase ADAR1 prevents spontaneous MAVS-dependent type I interferon responses. *Immunity*, **54**, 1961–1975.
  42. Pestal, K., Funk, C.C., Snyder, J.M., Price, N.D., Treuting, P.M. and Stetson, D.B. (2015) Isoforms of RNA-editing enzyme ADAR1 independently control nucleic acid sensor MDA5-driven autoimmunity and multi-organ development. *Immunity*, **43**, 933–944.
  43. Eckard, S.C., Rice, G.I., Fabre, A., Badens, C., Gray, E.E., Hartley, J.L., Crow, Y.J. and Stetson, D.B. (2014) The SKIV2L RNA exosome limits activation of the RIG-I-like receptors. *Nat. Immunol.*, **15**, 839–845.
  44. Lassig, C., Lammens, K., Gorenflos Lopez, J.L., Michalski, S., Fetscher, O. and Hopfner, K.P. (2018) Unified mechanisms for self-RNA recognition by RIG-I Singleton-Merten syndrome variants. *eLife*, **7**, e38958.
  45. Kato, H., Takeuchi, O., Mikamo-Satoh, E., Hirai, R., Kawai, T., Matsushita, K., Hiiragi, A., Dermody, T.S., Fujita, T. and 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.
  46. Hornung, V., Ellegast, J., Kim, S., Brzózka, K., Jung, A., Kato, H., Poeck, H., Akira, S., Conzelmann, K.-K., Schlee, M., et al. (2006)

- 5'-Triphosphate RNA is the ligand for RIG-I. *Science*, **314**, 994–997.
47. Zhao, Y., Ye, X., Dunker, W., Song, Y. and Karjane, J. (2018) RIG-I like receptor sensing of host RNAs facilitates the cell-intrinsic immune response to KSHV infection. *Nat. Commun.*, **9**, 4841.
48. Slomovic, S., Laufer, D., Geiger, D. and Schuster, G. (2005) Polyadenylation and degradation of human mitochondrial RNA: the prokaryotic past leaves its mark. *Mol. Cell. Biol.*, **25**, 6427–6435.
49. Landerer, E., Villegas, J., Burzio, V.A., Oliveira, L., Villota, C., Lopez, C., Restovic, F., Martinez, R., Castillo, O. and Burzio, L.O. (2011) Nuclear localization of the mitochondrial ncRNAs in normal and cancer cells. *Cell. Oncol.*, **34**, 297–305.
50. Nadal, A., Martell, M., Lytle, J.R., Lyons, A.J., Robertson, H.D., Cabot, B., Esteban, J.I., Esteban, R., Guardia, J. and Gomez, J. (2002) Specific cleavage of hepatitis C virus RNA genome by human RNase P. *J. Biol. Chem.*, **277**, 30606–30613.
51. Habjan, M., Andersson, J., Klingstrom, J., Schumann, M., Martin, A., Zimmermann, P., Wagner, V., Pichlmair, A., Schneider, U., Muhlberger, E., *et al.* (2008) Processing of genome 5' termini as a strategy of negative-strand RNA viruses to avoid RIG-I-dependent interferon induction. *PLoS One*, **3**, e2032.
52. Devarkar, S.C., Schweibenz, B., Wang, C., Marcotrigiano, J. and Patel, S.S. (2018) RIG-I uses an ATPase-powered translocation-throttling mechanism for kinetic proofreading of RNAs and oligomerization. *Mol. Cell*, **72**, 355–368.