**Supplementary Information**

# Defining the functional determinants for RNA surveillance by RIG-I

Andrew Kohlway[1*], Dahai Luo [2,4*], David C. Rawling[1], Steve C. Ding[1], Anna Marie Pyle [2, 3, 4#]

[1] Department of Molecular Biophysics and Biochemistry, [2] Department of Molecular, Cellular, and Developmental Biology, [3] Department of Chemistry, Yale University, New Haven, Connecticut 06520, [4] Howard Hughes Medical Institute, Chevy Chase, Maryland 20815.

[#] To whom correspondence may be addressed.

E-mail: anna.pyle@yale.edu

*These authors contributed equally to this work.

**Inventory of Supplemental Information**

**Supplementary Methods**

**Supplementary Table 1. Crystallographic and structure refinement statistics.**

**Supplementary Table 2. Nucleic acids used in this study.**

**Supplementary Table 3. Poly I:C ng/µl to nanomolar calculations.**

**Supplementary Table 4. RNA stimulated ATP hydrolysis by RIG-I.**

**Supplementary Figure 1. $K_{m,ATP}$ and $K_{m,RNA}$ ATPase experiments on LMW poly I:C.**

**Supplementary Figure 2. $K_{m,ATP}$ and $K_{m,RNA}$ ATPase experiments on short duplex RNA.**

**Supplementary Figure 3. Mock control for HEK293T cell culture IFN production.**

**Supplementary Figure 4. Cell culture IFN production on different lengths of poly I:C.**

**Supplementary Movie 1. Scanning motions of RIG-I on duplex RNA. (morph5b.mpg / morph5c.mpg).** The morph is generated from the three structures presented in this study.

**Supplementary References**

## Supplementary Methods

### Cloning, Expression, and Purification

Briefly, the constructs were cloned into the pET-SUMO vector (Invitrogen) and transformed into Rosetta II(DE3) *E.coli* cells (Novagen). The proteins were expressed in LB media upon the addition of 0.5 mM isopropyl-β-D-thiogalactopyranoside (IPTG) and grown at 18 °C overnight for 20 hours. The cells were then lysed with a microfluidizer at 15,000 psi, clarified by centrifugation, and purified by batch binding with Ni-NTA beads (Qiagen). After collection and elution from Biorad polyprep columns, the RIG-I constructs were concentrated on a HiTrap Heparin HP column (GE Healthcare) and gel filtered over a HiPrep 16/60 Superdex 200 column (GE Healthcare) in buffer containing 25 mM Hepes, pH 7.4, 150 mM NaCl, 5 % glycerol, 5 mM β-ME. RIG-I preparations were concentrated to between 5-10 mg/mL with a 50k MW cutoff Amicon centrifugal concentrator (Millipore), and concentrations were determined spectrophotometrically, using the extinction coefficients of $\varepsilon = 99{,}700\ M^{-1}\ cm^{-1}$ for full length RIG-I and an $\varepsilon = 60{,}280\ M^{-1}\ cm^{-1}$ for the RIG-I (ΔCARDs:1-229) N-terminal deletion construct. The extinction coefficients were calculated theoretically from the RIG-I sequence and guanidinium chloride denaturation of protein preparations. The RIG-I preparations were flash frozen with liquid nitrogen and stored at -80 °C.

### RNA synthesis and transcription

The 2'O-methyl modification on the penultimate nucleotide of the template strand prevents T7 terminal transferase activity as previously described (Kao et al, 1999). Incorporation of 2'-OMe modifications within the DNA template prevents addition of +1 and +2 nucleotides by T7 RNA polymerase and results in transcription of RNA molecules with defined, uniform 3'-ends, which is obviously essential for studies of RIG-I binding and stimulation. All synthesized and transcribed RNA constructs were purified on 20% denaturing polyacrylamide gels. LMW

polyI:C (Invivogen) was dissolved in buffer containing Hepes (pH 7.4), 150 mM NaCl, 5% glycerol, 5 mM BME to a final concentration of 10 mg/ml. 500 µl of this solution was loaded on an analytical 10.300 Superdex 200 Column (GE) and eluted at 0.25 ml/min while collecting 1 ml fractions. Concentrations were determined spectrophotometrically.

All hairpin RNAs were purified by 8M urea PAGE. After gel extraction, the re-annealing step was performed at low RNA concentrations by heating the RNA at 96°C for 2 mins and rapidly cooling on ice. It is notable that these hairpins are stabilized by a terminal UUCG tetraloop, which is known to promote exclusive hairpin formation by short duplexes, including those as short as four base-pairs (Cheong et al, 1990; Nozinovic et al, 2010).

**Crystallization, Data Collection, Structure Determination, and Refinement**

Briefly, the RIG-I (ΔCARDs:1-229) complex with 5'OH-GC10 duplex was preassembled by incubating at a protein:RNA molar ratio of 1:1.5 on ice for 1 hour and then purified with a HiPrep 16/60 Superdex 200 column (GE Healthcare). The crystals of the binary complex of RIG-I (ΔCARDs:1-229):5'OH-GC10 were grown at 13 °C by mixing equal volumes of precipitating solution (0.1 M Bicine, pH 9.0, 22.5 % polyethylene glycol 6,000) and RIG-I (ΔCARDs:1-229) : 5'OH-GC10 complex (2-3 mg $ml^{-1}$) using the sitting drop method. The crystals grew into needle clusters within a week and were harvested within two weeks. The crystals were soaked in a cryoprotecting solution containing 0.1 M Bicine, pH 9.0, 30 % polyethylene glycol 6,000 for 12 hours before being flash frozen with liquid nitrogen. To grow the crystals of the ternary complex of RIG-I (ΔCARDs:1-229):5'OH-GC10:ADP-$Mg^{2+}$, the binary complex of RIG-I (ΔCARDs:1-229):5'OH-GC10 was first incubated with 2.5 mM ADP and 2.5 mM $MgCl_2$ at 2-3 mg $ml^{-1}$ for half an hour to one hour on ice, mixed with equal volumes of precipitating solution (0.1 M Bicine, pH 9.0, 26-28 % polyethylene glycol 6,000) and then

grown at 13 °C. Crystals also grew into needle clusters within three days and were harvested within two weeks. Crystals were soaked in a cryoprotecting solution containing 0.1 M Bicine, pH 9.0, 30 % polyethylene glycol 6,000 briefly before being flash frozen with liquid nitrogen. Diffraction intensities were recorded at NE-CAT beamline ID-24 at the Advanced Photon Source (Argonne National Laboratory, Argonne, Illinois). Integration, scaling and merging of the intensities were carried out by using the programs XDS (Kabsch, 2010) and SCALA (Evans, 2006).

Initial attempts to use the structure of RIG-I (ΔCARDs:1-229) : 5'OH-GC10 (PDB: 2ykg) as search model for molecular replacement were not successful. Rather, successful phasing was accomplished through molecular replacement by using the subgroups (HEL1: aa 236-455, HEL2-HEL2i: aa 456-793, CTD: aa 794-925, and dsRNA) of RIG-I (ΔCARDs:1-229) : 5'OH-GC10 (PDB: 2ykg) as search models in Phaser (McCoy, 2007). Refinement cycles were carried out using Phenix Refine (Adams et al, 2010) and REFMAC5 (Murshudov et al, 1997) with four TLS (translation, liberation, screw-rotation displacement) groups (HEL1: aa 236-455, HEL2-HEL2i: aa 456-793, CTD: aa 794-925, and dsRNA). Refinement cycles were interspersed with model rebuilding using Coot (Emsley & Cowtan, 2004). The quality of the structures was analyzed with PROCHECK (Laskowski et al, 1993). A summary of the data collection and structure refinement statistics is given in **Supplementary Table 2**. During the crystallographic studies, it was noticed that crystals with RIG-I:dsRNA captured in the conformation 1, the binary complex of RIG-I and 5'OH-GC10, is always associated with the longest *c* axis (225.1 Å) of the unit cells (conformation 2, 219.8 Å and conformation 3, 207.8 Å). Figures and movies were prepared by using the program Pymol (DeLano, 2002).

**Analytical Ultracentrifugation-Sedimentation Velocity (SV) Experiments**

Full length RIG-I protein was mixed with 5'ppp10L, 5'ppp20L, 5'ppp30L and 5'pppGC22 RNAs at a ratio of 4.5 μM RIG-I : 1.5 μM RNA (or RIG-I alone) in 450 μl aliquots in buffer containing 25 mM HEPES pH 7.4, 150 mM NaCl, 2 mM $MgCl_2$, and 5 mM BME. The SV experiments were run at 40,000 rpm in a Beckman Optima XL-I analytical ultracentrifuge. Partial specific volumes for RIG-I and RIG-I:RNA complexes and buffer density and viscosity parameters were calculated in SEDNTERP. Data analyses were performed in SEDFIT (Schuck, 2000; Schuck et al, 2002).

**NADH coupled ATPase Experiments**

Experiments were set up in 50 μl reaction volumes in 96 well format using Corning clear half-area flat bottom plates (#3695). Each 50 μl reaction contained 10 μl of 5x NADH enzyme buffer (1 mM NADH, 100 U of lactate dehydrogenase/mL, 500 U/mL of pyruvate kinase/mL, and 2.5 mM phosphoenolpyruvate), 5-10 nM of RIG-I, 5 μl of varying amounts of either RNA or ATP, and a remaining volume of 25 mM 3-(N-morpholino)propanesulfonic acid (MOPS) pH 7.4, 150 mM KCl, 2 mM DTT, and 0.1 % Triton X-100. The rate of ATP hydrolysis was indirectly determined by monitoring the loss of NADH by reading the absorbance at 340 nm using a Biotek Synergy H1 plate reader. For both the $K_{m,ATP}$ and the $K_{m,RNA}$ experiments, RIG-I and the RNA constructs were allowed to equilibrate for at least 2 hours before addition of ATP. Detergent was required to record reproducible ATPase rates in the 96-well Corning clear bottom plates, especially at low concentrations of RIG-I. The initial velocities ($v_0$) at various RNA concentrations were plotted and fit to the following quadratic solution to the Briggs-Haldane equation:

$$y = y_0 + (amp) * \frac{[M_t] + [S_t] + K_M - \sqrt{([M_t] + [S_t] + K_M)^2 - 4[M_t][S_t]}}{2[M_t]} \quad (1)$$

[Mt] is the total protein concentration, [St] is the total [RNA], $y_0$ is the basal activity, amp is the $k_{cat}$ (minus the basal activity), and $K_m$ is the apparent Michaelis constant for substrate activation. The $y_0$ was constrained to the average basal activity from the entire set of 0 nM RNA, 5 mM ATP wells. The initial velocities ($v_0$) at various ATP concentrations were plotted and fit to the hyperbolic form of the above equation:

$$y = \frac{(amp)*[\mathrm{ATP}]}{K_M + [\mathrm{ATP}]} \qquad (2)$$

For the ATPase experiments in which the ATP concentration was varied, the RNA concentration were held at 500 nM for the short RNA duplexes (**Fig. 4**, **Supp. Fig. 2**), 500 ng/µL for the poly I:C experiments (**Supp. Fig. 1**), and 15 ng/µL for the poly I:C fractions (**Fig. 3**). Although 15 ng/µL was suboptimal for the longer polyIC fractions, it was the highest we could manage for all of the fractions from a single gel filtration experiment. One row from a 96 well plate constituted a single experiment with the following 12 ATP solutions in µM (final concentrations listed) derived from a two-thirds dilution series: 0, 30.2, 50.4, 84.0, 140.0, 233.3, 388.8, 648, 1080, 1800, 3000, and 5000.

For the ATPase experiments in which the RNA concentration was varied, the ATP concentration was held at 5 mM, approximately 10-fold above the $K_{m,ATP}$ measured for each RNA fraction. One row from a 96-well plate constituted a single experiment with the following 12 RNA concentrations in nM (or ng/µl for poly I:C, final concentrations listed) from a two-fold dilution series: 0, 0.5, 1.0, 2.0, 3.9, 7.8, 15.6, 31.2, 62.5, 125, 250, and 500 (**Fig. 4, Supp. Fig. 1**, **Supp. Fig. 2**). For the gel filtered poly I:C fractions the following 12 RNA concentrations in ng/µL (final concentrations listed) were used from a two-fold dilution series: 0, 0.01, 0.03, 0.06, 0.12, 0.23, 0.47, 0.94, 1.88, 3.75, 7.50, 15 (**Fig. 3**). In order to calculate the nM amounts of each poly I:C fraction for **Figure 3D**, the following estimates were made for the duplex lengths of

fractions A1-A7 respectively based on the semi-denaturing polyacrylamide gel and molecular weight standards shown in Figure 6C: 500, 360, 180, 90, 60, 40, and 25. The nM concentration ranges used for the analysis of the gel filtered poly I:C fractions is shown in the **Supplementary Table 3**.

**Cell Culture Interferon-β Response**

Batches of 293T cells were grown to 70-80% confluency in 10 cm dishes in Dulbecco's Modified Eagle Medium (DMEM; Invitrogen) containing 10% heat-inactivated fetal calf serum (Hyclone) and non-essential amino acids (Invitrogen). For RIG-I transfections of 10 cm dishes of 293T cells, one 800 µl aliquot of Opti-MEM containing 4 µg of pUNO-RIG-I, 1 µg of pRL-TK, and 5 µg of an IFN-ß/Firefly luciferase reporter plasmid was mixed with a second 800 µl aliquot of Opti-MEM containing 50 µl of lipofectamine. After 45 minutes, the 1.6 mL aliquot was diluted four-fold with Opti-MEM and then added to a 10 cm dish of 293T cells. The transfection was allowed to proceed for 6-8 hours, and then 10 mL of fresh DMEM was added to the plate. The cells were split twice into 15 cm dishes over the course of three days in 3 µg/µL blasticidin, and then used for transfections in 96 well plates containing RNA hairpins or poly I:C fractions.

The 293T cells, in DMEM without blasticidin, and transfected with pUNO-RIG-I, pRL-TK, and IFN-ß/Firefly luciferase reporter, were seeded at 15,000 cells per well, with each well containing 5 µl of Lyovec (Invivogen), and the following final concentrations of 5'ppp RNA hairpin in nM: 39.1, 78.1, 156.3, 312.5, 625 or the following final concentrations of poly I:C in total ng per well: 15.6, 31.3, 62.5, 125, 250, 500. In each experiment, the RNA hairpin or poly I:C was tested three times at each concentration (or total RNA amount). Luminescence measurements were assayed between 16-24 hours after stimulation by the RNA.

For luciferase measurements, the Promega Dual Luciferase Reporter assay system was used to quantitate the cellular levels of firefly and Renilla luciferase. Briefly, media was aspirated from each 96 well plate and replaced with 60 µL of passive lysis buffer. After 15 minutes at room temperature, lysates were collected, clarified by centrifugation, and then 20 µl of lysate was assayed for firefly and Renilla luciferase using the luminometer from a Biotek Synergy H1 plate reader with a dual injector. The Renilla luciferase is an internal control for each experiment and set of transfections, and the ratio of firefly luciferase over Renilla luciferase is reported in this paper.

**Accession code**

The atomic coordinates and structure factors of the binary complex of RIG-I (ΔCARDs:1-229) : 5'OH-GC10 and the ternary complex of RIG-I (ΔCARDs:1-229) : 5'OH-GC10 : ADP-$Mg^{2+}$ have been deposited with the RCSB Protein Data Bank under the accession codes 3zd6 and 3zd7. The ternary complex of RIG-I (ΔCARDs:1-229) : 5'OH-GC10 : $SO_4^{2+}$ is already available under the accession code 2ykg and was published in **Luo et al., 2011**.

## Supplementary Tables

### Supplementary Table 1. Crystallographic and structure refinement statistics.

| **Data collection** | | | |
|---|---|---|---|
| Structure | RIG-I (ΔCARDs 1-229): GC10 (Conformation 1) | RIG-I (ΔCARDs 1-229): GC10:$SO_4$ (Conformation 2) [b] | RIG-I (ΔCARDs 1-229): GC10: ADP-Mg (Conformation 3) |
| Space group | $P2_12_12_1$ | $P2_12_12_1$ | $P2_12_12_1$ |
| Cell dimensions (Å) | 48.5, 78.0, 225.1 | 47.6, 76.2, 219.8 | 48.3, 76.1, 207.8 |
| Resolution (Å) | 48.5 - 2.8 (2.9 - 2.8) [a] | 47.6 - 2.5 (2.6 - 2.5) | 48.3 - 2.5 (2.6 - 2.5) |
| R merge (%) | 6.4 (70.6) | 7.5 (62.3) | 6.4 (57.3) |
| I / σ | 16.2 (1.8) | 14.5 (1.5) | 11.4 (1.9) |
| Completeness (%) | 99.3 (98.6) | 93.8 (58.8) | 98.1 (98.9) |
| Redundancy | 5.0 (4.8) | 5.0 (2.2) | 3.4 (3.1) |
| **Refinement** | | | |
| Resolution (Å) | 25.0 - 2.8 | 45.0 - 2.5 | 25.0 - 2.5 |
| R work / R free (%) | 22.2 / 27.9 | 22.4 / 27.5 | 22.9 / 28.8 |
| No. atoms | 5,380 | 5,517 | 5,369 |
| Protein | 4,947 | 4,985 | 4,857 |
| RNA/ADP-$Mg^{2+}$ | 424 | 424 | 424 / 28 |
| Water | 9 | 99 | 60 |
| B-factors ($Å^2$) | 74.7 | 57.9 | 66.2 |
| Protein | 75.1 | 58.1 | 66.4 |
| Ligand | 71.7 | 65.7 | 70.2 |
| Solvent | 74.6 | 48.2 | 51.9 |
| Ramachandran analysis | | | |
| Favored | 86.3 | 94.2 | 89.1 |
| Additionally allowed | 12.8 | 4.8 | 10.4 |
| Outliers | 0.9 | 1.0 | 0.5 |
| R.m.s. deviations | | | |
| Bond lengths (Å) | 0.099 | 0.086 | 0.097 |
| Bond angles (°) | 1.6 | 1.3 | 1.5 |
| **PDB ID** | 3zd6 | 2ykg | 3zd7 |

[a] Statistics for the highest resolution shell is shown in parenthesis.

[b] Reference data taken from http://www.rcsb.org/pdb/explore/explore.do?structureId=2ykg. (Luo et al., 2011.)

**Supplementary Table 2. Nucleic acids used in this study.**

| Name | Sequence and Chemical Composition |
|---|---|
| **GC8** | 5'OH-GCGCGCGC-3' |
| **GC10** | 5'OH-GCGCGCGCGC-3' |
| **GC12** | 5'OH-GCGCGCGCGCGC-3' |
| **GC14** | 5'OH-GCGCGCGCGCGCGC-3' |
| **GC18** | 5'OH-GCGCGCGCGCGCGCGCGC-3' |
| **GC22** | 5'OH-GCGCGCGCGCGCGCGCGCGCGC-3' |
| **5'pppGC10** | 5'ppp-GGCGCGCGCC-3' |
| **5'pppGC12** | 5'ppp-GGCGCGCGCGCC-3' |
| **5'pppCM12** | 5'ppp-GGACGUACGUCC-3' |
| **5'pppGC22** | 5'ppp-GGCGCGCGCGCGCGCGCGCGCC-3' |
| **5'ppp8L** | 5'ppp-GGCGCGGC UUCG GCCGCG CC-3' |
| **5'ppp10L** | 5'ppp-GGACGUACGU UUCG ACGUACGUCC-3' |
| **5'ppp20L** | 5'pppGGAUCGAUCGAUCGAUCGGCUUCGGCCGAUCGAUCG AUCGAUCC-3' |
| **5'ppp30L** | 5'pppGGAUCGAUCGAUCGAUCGGCAUCGAUCGGCUUCGGC CGAUCGAUGCCGAUCGAUCGAUCGAUCC-3' |
| **polyIC** | 5'OH-$I^n$:$C^n$-3'(0.02-1 kilo base pairs) |

**Supplementary Table 3. Poly I:C ng/µl to nanomolar calculations.** The estimates for the length of each fraction of poly I:C as well as the approximate molecular weights (**Fig. 3A**). The ng/µl concentrations used in the poly I:C $K_{m,RNA}$ experiment were converted to nanomolar of poly I:C strands based on the estimated length and molecular weights of each poly I:C strand.

| **Poly I:C (ng/µL)** | **A1 (nM)** | **A2 (nM)** | **A3 (nM)** | **A4 (nM)** | **A5 (nM)** | **A6 (nM)** | **A7 (nM)** |
|---|---|---|---|---|---|---|---|
| 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 0.01 | 0.04 | 0.06 | 0.12 | 0.24 | 0.36 | 0.55 | 0.87 |
| 0.03 | 0.09 | 0.12 | 0.24 | 0.48 | 0.73 | 1.09 | 1.75 |
| 0.06 | 0.17 | 0.24 | 0.48 | 0.97 | 1.45 | 2.18 | 3.49 |
| 0.12 | 0.35 | 0.48 | 0.97 | 1.94 | 2.91 | 4.36 | 6.98 |
| 0.23 | 0.70 | 0.97 | 1.94 | 3.88 | 5.82 | 8.73 | 13.96 |
| 0.47 | 1.40 | 1.94 | 3.88 | 7.76 | 11.64 | 17.45 | 27.93 |
| 0.94 | 2.79 | 3.88 | 7.76 | 15.51 | 23.27 | 34.91 | 55.85 |
| 1.88 | 5.59 | 7.76 | 15.51 | 31.03 | 46.54 | 69.82 | 111.71 |
| 3.75 | 11.17 | 15.51 | 31.03 | 62.06 | 93.09 | 139.63 | 223.41 |
| 7.50 | 22.34 | 31.03 | 62.06 | 124.12 | 186.18 | 279.26 | 446.82 |
| 15.00 | 44.68 | 62.06 | 124.12 | 248.24 | 372.35 | 558.53 | 893.65 |
| **Length** | 500 | 360 | 180 | 90 | 60 | 40 | 25 |
| **MW** | 335703 | 241706.16 | 120853.08 | 60426.54 | 40284.36 | 26856.24 | 16785.15 |

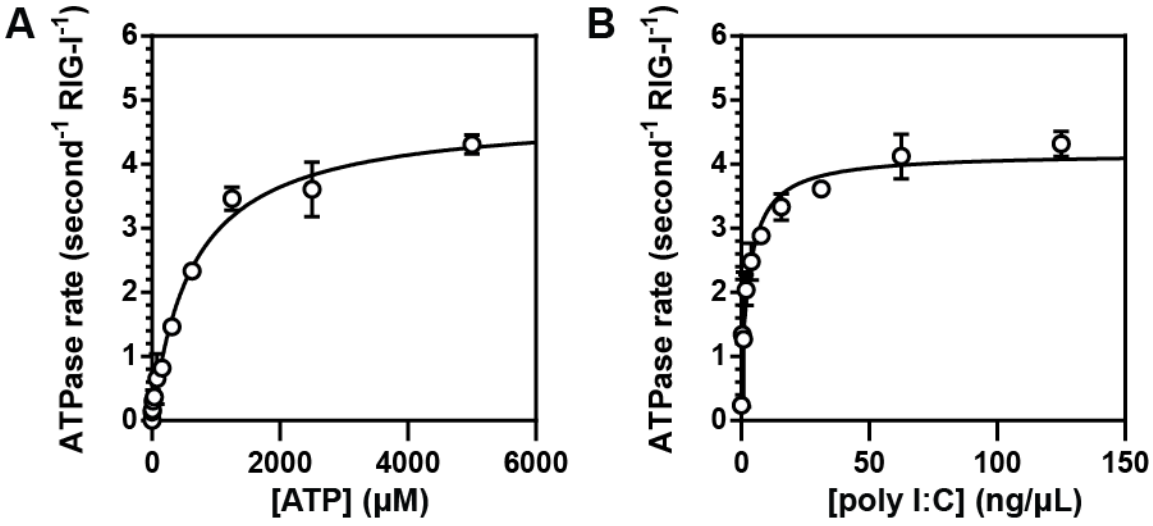
**Supplementary Table 4. RNA stimulated ATP hydrolysis by RIG-I.**

| RNA construct | $k_{cat}$ ± SD($s^{-1}$RIG-$I^{-1}$) | $K_{m,ATP}$ ± SD(μM) | $K_{m,RNA}$ ± SD(nM) |
|---|---|---|---|
| **5'ppp8L** | 7.45 ± 0.90 | 454 ± 18 | 5.16 ± 0.40 |
| **5'ppp10L** | 14.3 ± 2.5 | 556 ± 65 | 4.46 ± 0.50 |
| **5'ppp20L** | 12.6 ± 1.6 | 604 ± 40 | 10.1 ± 0.50 |
| **5'ppp30L** | 9.98 ± 1.3 | 622 ± 100 | 10.8 ± 1.0 |
| **GC8** | 7.34 ± 2.6 | 425 ± 66 | 91.4 ± 15 |
| **GC10** | 14.4 ± 3.0 | 511 ± 55 | 24.4 ± 1.5 |
| **GC12** | 15.9 ± 3.1 | 528 ± 38 | 13.1 ± 1.0 |
| **GC14** | 15.1 ± 1.9 | 537 ± 31 | 23.2 ± 1.3 |
| **GC18** | 12.5 ± 1.3 | 600 ± 37 | 26.7 ± 3.4 |
| **GC22** | 11.3 ± 1.1 | 570 ± 75 | 27.3 ± 1.1 |
| **5'pppGC10** | 18.8 ± 2.8 | 498 ± 35 | 1.16 ± 0.20 |
| **5'pppGC12** | 20.5 ± 3.3 | 535 ± 55 | 2.33 ± 0.20 |
| **5'pppCM12** | 15.9 ± 1.8 | 591 ± 57 | 2.62 ± 0.10 |
| **5'pppGC22** | 12.3 ± 1.9 | 536 ± 39 | 3.58 ± 1.0 |
| **LMW poly I:C** | 4.90 ± 0.50 | 690 ± 130 | 2.40 ± 1.1 (ng/μl) |

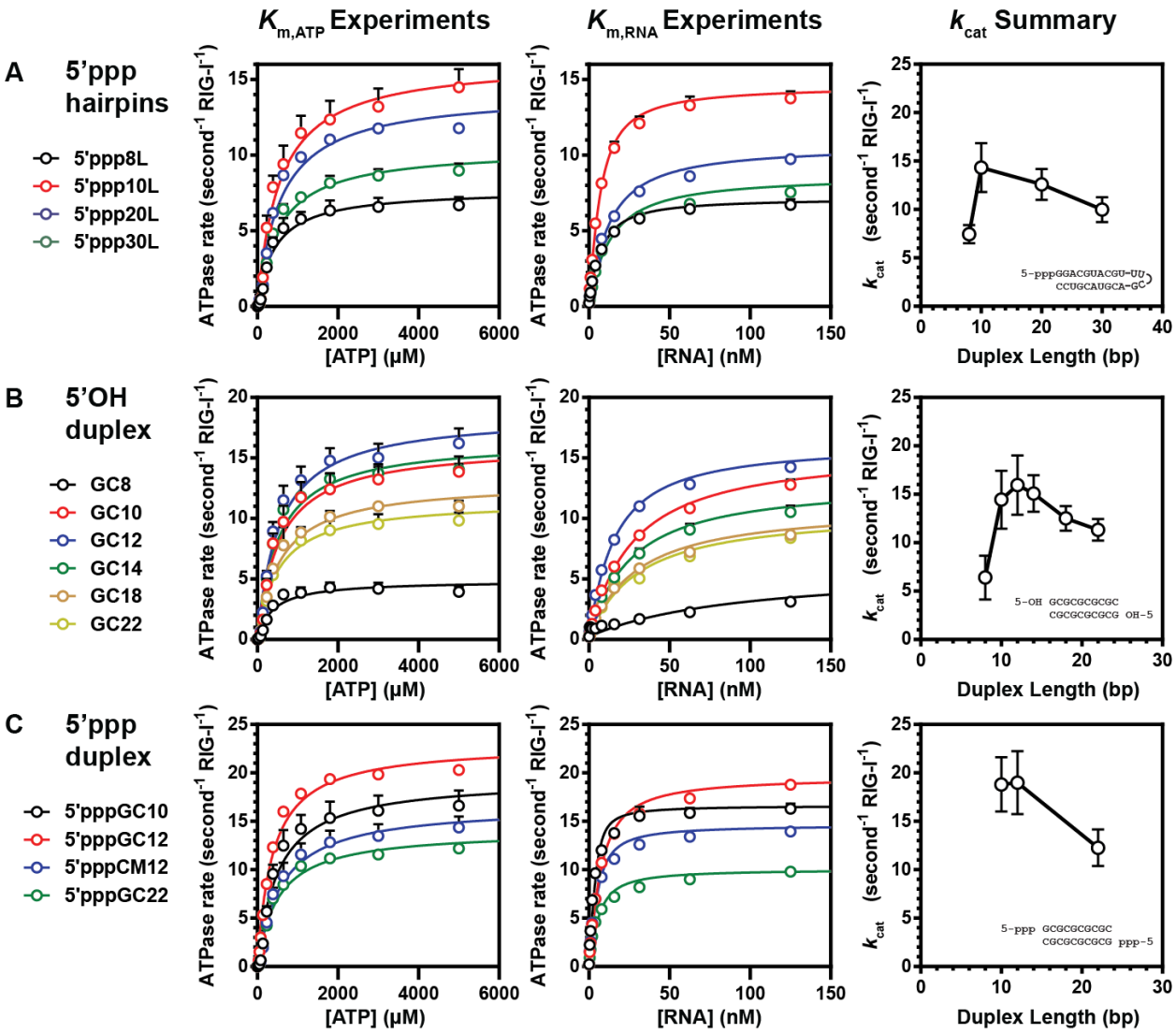
***Note that for the poly I:C $K_{m,ATP}$, the poly I:C concentration was kept at 500 ng/μl. And for the poly I:C $K_{m,RNA}$, the poly I:C concentration was varied up to 500 ng/μl.

**Supplementary Figure 1. $K_{m,ATP}$ and $K_{m,RNA}$ ATPase experiments on LMW poly I:C.**

**(A)** ATPase activity of RIG-I stimulated by 500 ng/μL LMW poly I:C while varying the ATP concentration from 0 to 5 mM ATP. Error bars report the standard deviation from 4 experiments. **(B)** ATPase activity of RIG-I at 5 mM ATP while varying LMW poly I:C from 0 to 500 ng/μL. Error bars report the standard deviation from 4 experiments. The average $k_{cat}$ from both experiments was 4.9 $s^{-1}$, and the $K_{m,ATP}$ was approximately 700 μM. The $K_{m,RNA}$ was 2.4 ng/μL, which is difficult to interpret because it cannot be expressed as a nanomolar value due to the heterogeneity of poly I:C samples

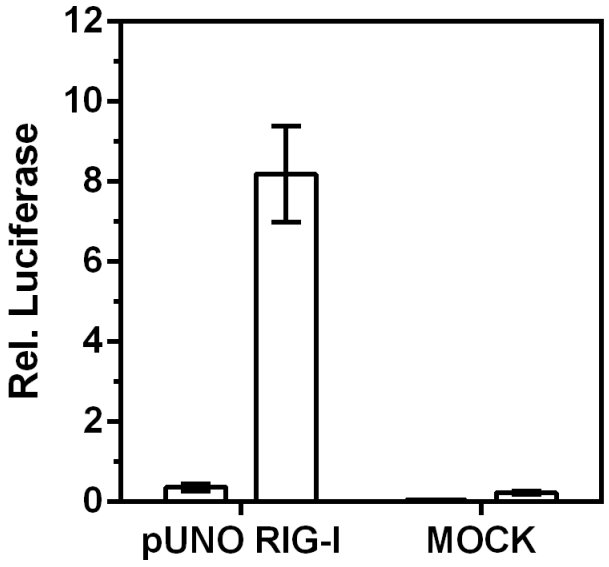


**Supplementary Figure 2. $K_{m,ATP}$ and $K_{m,RNA}$ ATPase experiments on short duplex RNA.** ATPase activity of RIG-I stimulated by various length RNAs including 5'triphosphorylated hairpins, 5'hydroxyl duplexes, and 5'triphosphorylated duplexes. The $K_{m,ATP}$ of RIG-I (10 nM enzyme) stimulated by each RNA was measured by varying the ATP concentrations ranging from 0 to 5 mM at 500 nM RNA. The $K_{m,RNA}$ of RIG- I (5 nM enzyme) stimulated by each RNA was measured by varing the RNA concentrations ranging from 0 to 500 nM at 5 mM ATP. A small basal activity (0 nM RNA) is measured for RIG-I of less than 1 per second. Error bars for the $K_{m,ATP}$ and $K_{m,RNA}$ experiments report the standard error of the mean from 4 experiments. The last column of graphs plots the average $k_{cat}$ values calculated from Briggs-Haldane fits from both the $K_{m,ATP}$ and $K_{m,RNA}$ experiments. Error bars for the $k_{cat}$ summary report the standard deviation measured across 6 experiments, in which each experiment was comprised of an averaged duplicate dataset for each RNA or ATP concentration. **(A)** ATPase measurements on 4 triphosphorylated hairpins with a duplex region of 8, 10, 20, and 30 nucleotides with a UUCG hairpin. **(B)** ATPase measurements on 6 double stranded RNA duplexes with 5'hydroxyl of length 8, 10, 12, 14, 18, and 22. **(C)** ATPase measurements on 4 double stranded RNA duplexes with a 5'triphosphate of length 10, 12, and 22. **Supplementary Table 2** lists the RNA sequences used in this study. We observed similar $k_{cat}$ values for RIG-I stimulated by the 5'ppp8L hairpin and GC8. However, in the case of the hairpin, we observed a 5.2 nM $K_{m,RNA}$, approximately 20-fold smaller than GC8, perhaps because 5'ppp8L contains a 'UUCG' tetraloop, which may accommodate the HEL2i flexibility seen in the crystal structures.

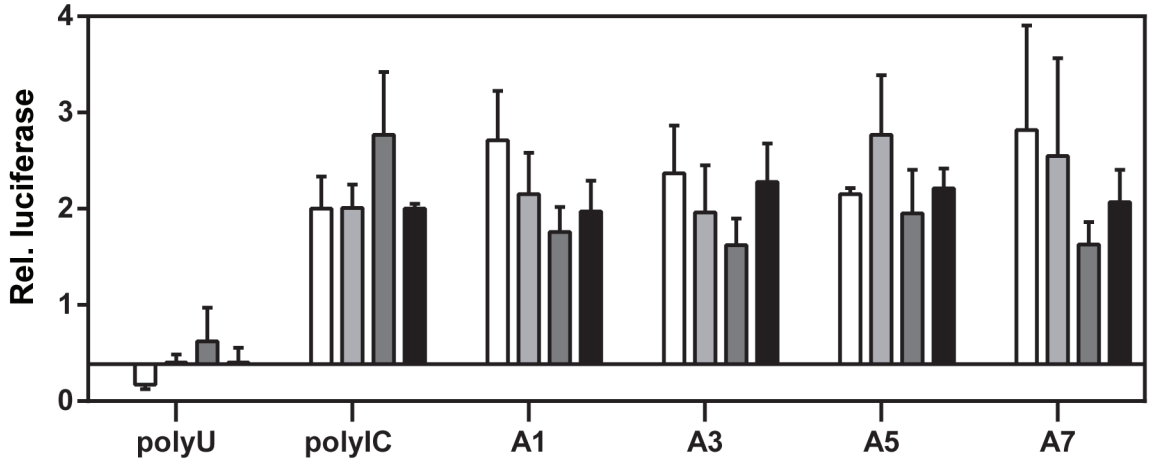

$K_{m,ATP}$ Experiments
$K_{m,RNA}$ Experiments
$k_{cat}$ Summary
A 5'ppp hairpins
5'ppp8L
5'ppp10L
5'ppp20L
5'ppp30L
B 5'OH duplex
GC8
GC10
GC12
GC14
GC18
GC22
C 5'ppp duplex
5'pppGC10
5'pppGC12
5'pppCM12
5'pppGC22
ATPase rate (second$^{-1}$ RIG-I$^{-1}$)
[ATP] (μM)
[RNA] (nM)
$k_{cat}$ (second$^{-1}$ RIG-I$^{-1}$)
Duplex Length (bp)
5-pppGGACGUACGU-UU
CCUGCAUGCA-GC
5-OH GCGCGCGCGC
CGCGCGCGCG OH-5
5-ppp GCGCGCGCGC
CGCGCGCGCG ppp-5

**Supplementary Figure 3. Mock control for HEK293T cell culture IFN production.**

The IFN-β production in 293T cells overexpressing RIG-I (and a mock control not overexpressing RIG-I) was measured in the absence (left) and presence (right) of poly I:C stimulation. The relative luciferase is the firefly luciferase (IFN-β reporter) divided by the Renilla luciferase. The following protocol was adapted from Luo et. al. (Luo et al, 2011). 293T cells were seeded at ~ 50,000 cells per well in 24 well plates. The next day, 293T cells were transfected with 30 ng of pRLTK, 178 ng of a firefly IFN-β reporter, and 3 ng (or none for mock) of pUNO-RIG-I per well using lipofectin (Invitrogen). After 24 hours, 293T cells were transfected with 1 μg of poly I:C (or none for negative control) using mRNA transfection reagent (MIRUS). After 16 hours, cells were harvested and assayed for firefly and *Renilla* luciferase using the Promega Dual Luciferase Reporter assay system. Error bars report the standard deviation from 6 experiments for unstimulated and 12 experiments for stimulated.



**Supplementary Figure 4. Cell culture IFN production on different lengths of poly I:C.** The IFN-ß responses to the fractions of poly I:C was measured in HEK 293T cells transfected with pUNO-RIG-I, an IFN-ß/Firefly luciferase reporter, and a pRL-TK reporter (note that the poly I:C data is the same as in Figure 4D and was done side by side with Fraction A1-A7 shown here). The charts display the measured relative luciferase ratio of Firefly luminescence over Renilla luminescence from 293T cells in which RIG-I was stimulated by the fractions of poly I:C, and also a single stranded poly U and no RNA (serum free media) control. The range of RNA concentrations spans between 31 to 250 ng per well displayed in the figure by a darkening color gradient from low to high RNA concentration. Error bars report the standard error of the mean from 3 measurements.



## Supplementary References


**Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung LW, Kapral GJ, Grosse-Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner R, Read RJ, Richardson DC, Richardson JS, Terwilliger TC, Zwart PH (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution.** ***Acta Crystallogr D Biol Crystallogr*** **66: 213-221**

**Cheong C, Varani G, Tinoco I, Jr. (1990) Solution structure of an unusually stable RNA hairpin, 5'GGAC(UUCG)GUCC.** ***Nature*** **346: 680-682**

**DeLano WL (2002)** ***The PyMOL User's Manual*****: DeLano Scientific, Palo Alto, CA, USA.**

**Emsley P, Cowtan K (2004) Coot: model-building tools for molecular graphics.** ***Acta Crystallogr D Biol Crystallogr*** **60: 2126-2132**

**Evans P (2006) Scaling and assessment of data quality.** ***Acta Crystallogr D Biol Crystallogr*** **62: 72-82**

**Kabsch W (2010) Xds.** ***Acta Crystallogr D Biol Crystallogr*** **66: 125-132**

**Kao C, Zheng M, Rudisser S (1999) A simple and efficient method to reduce nontemplated nucleotide addition at the 3 terminus of RNAs transcribed by T7 RNA polymerase.** ***Rna*** **5: 1268-1272**

**Laskowski RA, Macarthur M. W., Moss D. S., Thornton JM (1993) PROCHECK: a program to check the stereochemical quality of protein structures.** ***J Appl Cryst*** **26: 283-291**

**Luo D, Ding SC, Vela A, Kohlway A, Lindenbach BD, Pyle AM (2011) Structural insights into RNA recognition by RIG-I.** ***Cell*** **147: 409-422**

**McCoy AJ (2007) Solving structures of protein complexes by molecular replacement with Phaser.** ***Acta Crystallogr D Biol Crystallogr*** **63: 32-41**

**Murshudov GN, Vagin AA, Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method.** ***Acta Crystallogr D Biol Crystallogr*** **53: 240-255**

**Nozinovic S, Furtig B, Jonker HR, Richter C, Schwalbe H (2010) High-resolution NMR structure of an RNA model system: the 14-mer cUUCGg tetraloop hairpin RNA.** ***Nucleic Acids Res*** **38: 683-694**

**Schuck P (2000) Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling.** ***Biophysical journal*** **78: 1606-1619**

**Schuck P, Perugini MA, Gonzales NR, Howlett GJ, Schubert D (2002) Size-distribution analysis of proteins by analytical ultracentrifugation: strategies and application to model systems.** ***Biophysical journal*** **82: 1096-1111**