3436–3445 Nucleic Acids Research, 2013, Vol. 41, No. 5 Published online 15 January 2013
doi:10.1093/nar/gks1477

Conformational rearrangements of RIG-I receptor on formation of a multiprotein:dsRNA assembly

Simone A. Beckham[1,2], Jason Brouwer[1], Anna Roth[1], Die Wang[2], Anthony J. Sadler[2], Matthias John[3], Kerstin Jahn-Hofmann[4], Bryan R. G. Williams[2], Jacqueline A. Wilce[1] and Matthew C. J. Wilce[1,*]

[1]Department of Biochemistry and Molecular Biology, Monash University, Clayton, Victoria 3800, Australia [2]Monash Institute of Medical Research, Monash University, Clayton, Victoria 3168, Australia, [3]Roche Kulmbach GmbH, 95326 Kulmbach, Germany and [4]Sanofi Deutschland GmbH, 65926 Frankfurt, Germany

Received October 9, 2012; Revised December 19, 2012; Accepted December 20, 2012

ABSTRACT

The retinoic acid inducible gene-I (RIG-I)-like family of receptors is positioned at the front line of our innate cellular defence system. RIG-I detects and binds to foreign duplex RNA in the cytoplasm of both immune and non-immune cells, and initiates the induction of type I interferons and pro-inflammatory cytokines. The mechanism of RIG-I activation by double-stranded RNA (dsRNA) involves a molecular rearrangement proposed to expose the N-terminal pair of caspase activation recruitment domains, enabling an interaction with interferon-beta promoter stimulator 1 (IPS-1) and thereby initiating downstream signalling. dsRNA is particularly stimulatory when longer than 20 bp, potentially through allowing binding of more than one RIG-I molecule. Here, we characterize full-length RIG-I and RIG-I subdomains combined with a stimulatory 29mer dsRNA using multi-angle light scattering and size-exclusion chromatography–coupled small-angle X-ray scattering, to build up a molecular model of RIG-I before and after the formation of a 2:1 protein:dsRNA assembly. We report the small-angle X-ray scattering–derived solution structure of the human apo-RIG-I and observe that on binding of RIG-I to dsRNA in a 2:1 ratio, the complex becomes highly extended and flexible. Hence, here we present the first model of the fully activated oligomeric RIG-I.

INTRODUCTION

Retinoic acid inducible gene-I (RIG-I) is a cytosolic RNA receptor that acts in the front line of defence against viral pathogens (1–5). While Toll-like receptors 3, 7 and 8 detect foreign RNA that enters the cell via endosomal compartments, RIG-I and related family members, MDA5 (melanoma differentiation–associated gene 5) and LGP2 (laboratory of genetics and physiology 2), recognize foreign RNA that enters the cytoplasm. Together, these proteins are responsible for triggering signalling cascades that ultimately up-regulate the expression of cytokine and chemokine anti-viral mediators, including type I interferon (6). Type I interferon, in turn, induces higher levels of RIG-I expression in surrounding cells, priming them against viral invasion (7).

RIG-I activation is initiated through recognition of double-stranded RNA (dsRNA) end structures that discriminate non-self- from self-RNA. Uncapped 5′-triphosphate moieties are a hallmark of viral genomic RNA and are thought to be fundamental for self-/non-self-RNA discrimination by RIG-I. RIG-I binds with highest affinity, and exhibits greatest activation to fully base-paired dsRNA, including that formed as part of panhandle structures, with a 5′ triphosphate (8,9). However, *in vitro* studies have shown that RIG-I is also activated by dsRNA ligands with end termini that lack a 5′-triphosphate group. Non–5′-triphosphate RNA ligands of RIG-I include dsRNA possessing 5′-monophosphate moieties (10,11), RNase L products exhibiting 3′ monophosphates (12) or synthetic dsRNA lacking both 5′ and 3′ phosphates (11,13). dsRNA termini with 3′-nucleotide overhangs fail to activate RIG-I, but dsRNA with short 5′-nucleotide overhangs or dsRNA displaying blunt ends activates RIG-I to trigger the innate immune response (11,13). In addition to the end structure, the length of the dsRNA impacts on its stimulatory potential. It has been shown that RIG-I requires dsRNA longer than 20 bp for optimal activation, and markedly increased activity is seen with increasing dsRNA length (4,13,14). Longer dsRNA are able to accommodate multiple

*To whom correspondence should be addressed. Tel: + 61 399 029 244; Fax: +61 399 029 500; Email: matthew.wilce@monash.edu
Present address:
Matthias John, Moderna Therapeutics, Cambridge, MA 02142, USA.

RIG-I molecules, and it may be that this is a requirement for effective immune stimulation (15).

The molecular basis for RIG-I recognition of and activation by RNA has been the focus of intense research effort since its recognized importance in innate cellular immunity (16). RIG-I is a 925 amino acid protein comprising a pair of N-terminal caspase activation recruitment domains (CARDs), a central DExH-box helicase domain (so named for the conserved helicase motif II sequence) and a C-terminal domain (CTD), also referred to as the repressor domain. In the absence of ligand, RIG-I is maintained in a repressed state (15). The activation of RIG-I by RNA is proposed to involve a conformational change that exposes the CARDs, enabling them to interact with the mitochondria-bound interferon-beta promoter stimulator 1 (IPS-1)-1 protein (also known as mitochondrial antiviral-signaling, Cardiff and virus-induced signaling adapter), most likely via a homotypic CARD–CARD interaction (17,18). After interaction with RIG-I, IPS-1 interacts with the tumour necrosis factor receptor–associated factors TRAF3/6 adaptor complex and activates inhibitor of κB kinases, inhibitor of nuclear factor κ-B kinase and the tumor necrosis factor receptor-associated factor (TRAF) family member-associated NF-kappa-B activator-binding kinase 1. These phosphorylate interferon regulatory factors 3 and 7 to drive the expression of type I interferons and interferon-stimulated genes, resulting in an anti-viral state (19).

Many aspects of the structural basis for the interaction of RIG-I with short dsRNA have recently been elucidated. It has been shown that the CTD is responsible for the recognition of dsRNA end structures (8,11,20–22). The CTD possesses a positively charged RNA-binding site that accommodates the end of blunt-ended dsRNA or single-stranded RNA (ssRNA) that possesses a 5′ triphosphate. It is this interaction that is suggested to underlie the discrimination between host and viral RNA, as most endogenous RNA molecules do not possess these end structures. On recognition of the dsRNA end structure by the CTD, the helicase domain is also able to interact with the dsRNA. The helicase domain comprises two RecA-like domains and a helical insertion domain that, together with the CTD, are able to completely encircle the dsRNA (23–25). The structural rearrangement of the helicase on binding to dsRNA forms an active adenosine triphosphatase (ATPase) site, helping to explain the increase in ATPase activity that occurs on dsRNA binding (23), although ATPase activity, in itself, is not required for RIG-I stimulatory activation (26). The crystal structure of duck RIG-I (53% identity to human) shows the way in which the CARDs are arranged before interaction with dsRNA (27). The CARDs possess the canonical CARD fold and are arranged head-to-tail against the helicase insertion domain. From this positioning of the duck RIG-I CTD on binding, it is suggested that the CARDs would be displaced to some extent on interaction of RIG-I with dsRNA.

Despite this detailed understanding of the RIG-I interaction with dsRNA, however, there are several aspects of its activation that remain unknown. First, the crystal structure of duck RIG-I in apo-form revealed the arrangement of the CARDs and helicase domain, but not the position of the CTD that has previously been reported to interact with the CARDs in its repressor domain capacity (15,27). Thus, the complete structure of apo–full-length RIG-I remains unknown. Second, the small-angle X-ray scattering (SAXS) studies of RIG-I do not reveal a dramatic repositioning of the CARD domains on binding to a dsRNA 10mer, (24). The CARD domains are modelled close against the helicase domain, and an overall compacting of RIG-I is reported on dsRNA interaction. Thus, structural data to date do not provide biophysical evidence for the release of the CARD domains as suggested in the current literature. Third, the structural information for RIG-I bound to dsRNA has only provided detailed information for 1:1 RIG-I:RNA complexes. dsRNAs <20 bp in length have been used, which are known to be significantly less stimulatory and are not able to accommodate more than one RIG-I molecule. An exception is a cryo-electron microscopy (EM) study of RIG-I bound to a dsRNA 24mer (28). The detail provided, however, is limited, and the positioning of the CTD domains is not consistent with X-ray crystal structures. There has also been an investigation of MDA5 binding to a 20-bp RNA in a 2:1 ratio using SAXS. However, their data were collected for MDA5 without CARDs and also suggest that MDA5 interacts with RNA differently from RIG-I (29). They thus do not provide detailed information for a multimeric complex of RIG-I bound to dsRNA.

Because a proposed property of RIG-I before and after being bound to oligonucleotide is a change in the accessibility and degree of mobility of the CARDs, we have used SAXS for determining the arrangement of RIG-I before and after binding to immunostimulatory dsRNA. Although several classes of dsRNA are substrates of RIG-I, here we selected the most simple case of a blunt-ended dsRNA 29mer without 5′ triphosphates that has been previously established as immunostimulatory (13). Analyses of size-exclusion chromatography (SEC) and use of multi-angle light scattering (MALS) for molecular weight determination of the full-length protein and subdomains in complex with RNA have allowed us to examine dsRNA binding and to assess the stoichiometry and monodispersity of the complexes formed. In-line SEC-coupled SAXS (SEC-SAXS) analysis has provided high-quality data, allowing us to model RIG-I proteins in both apo-form and in complex with 29mer dsRNA. These data describe the shape of RIG-I in apo-form and on formation of a 2:1 RIG-I:dsRNA complex, and reveal both the position of the CTD in human apo-RIG-I and the release of the CARD domains on the formation of a multi–RIG-I:dsRNA complex.

MATERIALS AND METHODS

A detailed description can be found in the Supplementary methods.

Sample preparation

RIG-I proteins representing the full-length protein (FL-RIG-I), RIG-I without the two N-terminal CARDs

(DC-RIG-I) and the CTD alone (CTD-RIG-I) were produced by overexpression in *Escherichia coli*. The proteins were purified by immobilized-metal affinity chromatography followed by gel-filtration and heparin affinity chromatography. The 29mer dsRNA was synthesized by Roche (Kulmbach, Germany), and the 8mer was synthesized by Dharmacon.

SEC-coupled MALS and SAXS

RIG-I proteins and RNA complexes were prepared at a final protein concentration of 8 mg/ml and mixed with RNA at a 1:1 molar ratio (effectively providing the RNA in excess in view of the formation of 2:1 protein:RNA complexes). All samples were duplicated; one of the duplicates was analysed using SEC-coupled MALS (SEC-MALS), and the other subjected to SEC-SAXS analysis. The same column and chromatographic parameters were maintained. SEC-SAXS data were collected at the SAXS/WAXS (small and wide angle x-ray scattering) beamline at the Australian Synchrotron.

SAXS data processing

SAXS data reduction was performed using beamline-specific software developed from SAXSID15d (Advanced Photon Source, Argonne National Laboratory). The ATSAS software package (http://www.embl-hamburg.de/biosaxs/software.html) was used for analysing the SAXS data. This includes *ab initio* structure calculation using DAMMIF and DAMAVER (30,31), multiphase modelling with MONSA (32), and rigid-body modelling using BUNCH (31) and CORAL (31). The ensemble optimization method (EOM) was also used (33).

Limited proteolysis of RIG-I:RNA complexes using trypsin

FL-RIG-I and DC-RIG-I were mixed with excess RNA (at a 1:1 ratio) with the 29mer and an 8mer, and were incubated with 0.2 µg of sequencing grade–modified trypsin (Promega) at 37°C. Over time, 10-µl aliquots were removed, and samples were subjected to sodium dodecyl sulphate–polyacrylamide gel electrophoresis (SDS–PAGE) and visualized with Coomassie blue stain. N-terminal sequencing of selected bands was performed using Edman degradation.

RESULTS

Full-length RIG-I and RIG-I deletion mutants ± dsRNA analysis by SEC-SAXS

To characterize the structure of RIG-I before and after binding to target dsRNA, we prepared FL-RIG-I, DC-RIG-I and CTD-RIG-I (Figure 1A and C). SEC-MALS confirmed that these proteins each behaves as monomeric proteins of the expected molecular weight (Figure 1B and Supplementary Table S1). For the formation of multi–RIG-I:RNA complexes, we used a target 29mer dsRNA based on an siRNA sequence previously determined to activate RIG-I in cellular assays (13). On the combination of the CTD-, DC- and FL-RIG-I with

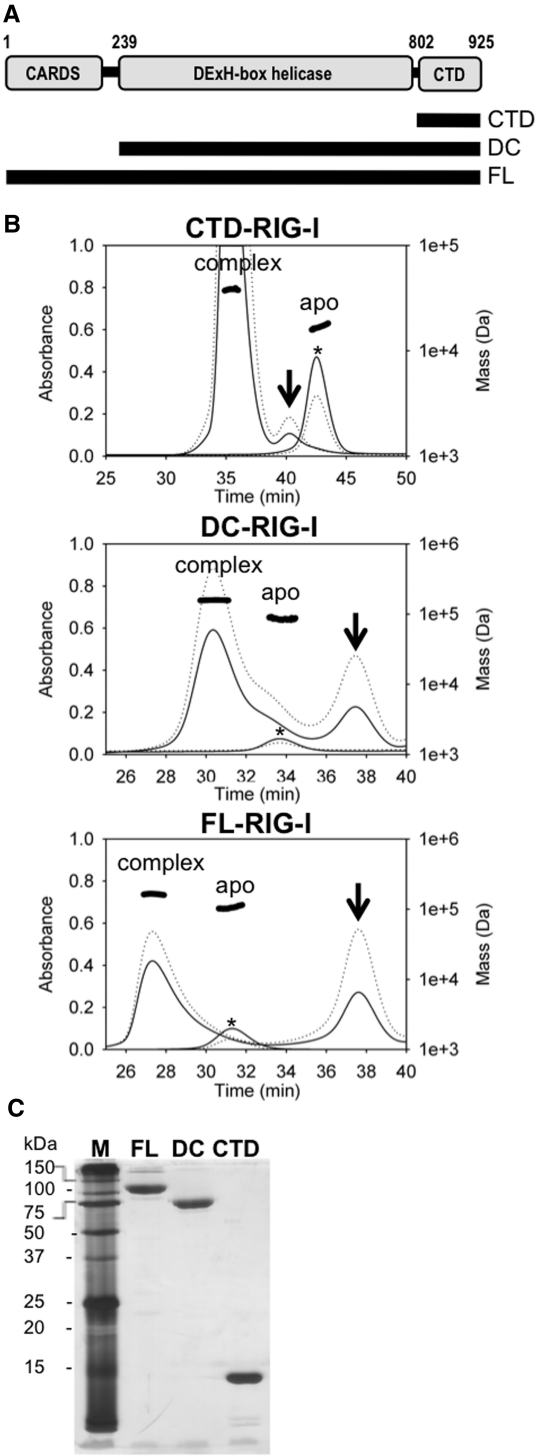

Figure 1. RIG-I constructs and RIG-I:RNA complexes. (**A**) Schematic representation showing RIG-I proteins used in this study and their abbreviations (CTD = C-terminal domain RIG-I; DC = deltaCARDs RIG-I; FL = full-length RIG-I). (**B**) Overlayed SEC profiles of apo-RIG-I proteins (peaks are indicated by *), and RIG-I proteins combined with 29mer RNA, detected at 260 nm (dotted lines) and 280 nm (solid lines). CTD-RIG-I:RNA (1:1) and excess RNA peaks are indicated with an arrow. MALS data collected simultaneously with peak elution and used to determine molecular mass are shown above the protein and protein:RNA complex peaks (thick dark line), confirming formation of 2:1 protein:RNA complexes. **(C)** FL-RIG-I (FL), DC-RIG-I (DC) and CTD-RIG-I (CTD) were subjected to SDS–PAGE and visualized by silver staining. M represents molecular weight markers.

the 29mer dsRNA, 2:1 protein:dsRNA complexes were formed (Figure 1B). The complex peaks eluted at earlier elution times, well separated from the apo-protein and the RNA. Their elution times and MALS-based molecular weight estimates were consistent with the expected molecular weights of 2:1 protein:RNA complexes (Supplementary Table S1). SAXS data were collected as apo-RIG-I or RIG-I:RNA complexes eluted from an SEC column, ensuring the analysis of uniform non-aggregated samples. The experimental and calculated scattering curves are shown in Figure 2A; the pairwise distribution function (P(r)) profiles are shown in Figure 2B; normalized Kratky plots are shown in Figure 2C; and structural parameters are reported in Supplementary Table S1.

RIG-I domain arrangement in the absence of dsRNA

The SAXS data for all three apo-RIG-I proteins were used to build up a model for the FL-RIG-I molecule. The CTD-RIG-I SAXS data were consistent with the size and shape of its crystallographically determined structure (PDB ID: 3OG8). The radius of gyration (R_g) for CTD-RIG-I was determined by Guinier analysis to be 14.5 Å, and the pairwise distribution function P(r) is unimodal with a narrow peak, reflecting a simple globular form with R_{max} = 43.1 Å. The DC-RIG-I SAXS data gave rise to an R_g of 37.0 Å and R_{max} = 129.7 Å, which is consistent with the dimensions of the apo-helicase domain (PDB ID: 3TBK) with the CTD (PDB ID: 3OG8) associated alongside. The P(r) profile is bimodal, indicating the presence of more than one lobe to the shape of the molecule. The normalized Kratky plot for the DC-RIG-I data shows a non-parabolic curve that does, however, decrease towards zero at higher sR_g and displays a maximum value at 1.73 sR_g (*i.e.* $\sqrt{3}$ sR_g), consistent with a predominantly rigid molecule (34). *Ab initio* shape reconstruction for DC-RIG-I was hence calculated using DAMMIF (30) and DAMAVER (35) and is displayed in Figure 3A. The DC-RIG-I helicase (PDB ID: 3TBK) and CTD (PDBID: 3OG8) domains were modelled as rigid bodies against the SAXS data using the program BUNCH (31). The model is shown superposed with the *ab initio* shape in Figure 3A, and the predicted scattering curve from this model fit the raw scattering data well (χ value = 0.41) (Figure 2A and Supplementary Table S1). The CTD fits within a well-defined lobe alongside the helicase domain, but is not associated with the helicase domain.

The FL-RIG-I SAXS data gave rise to an R_g of 39.6 Å and R_{max} = 138.6 Å, which is only slightly larger than that of DC-RIG-I. The P(r) profile is slightly less bimodal than that calculated for DC-RIG-I but is otherwise similar, consistent with there being no dramatic structural difference between these molecules, except for the presence of the CARDs. The normalized Kratky analysis (Figure 2C) indicates that FL-RIG-I is also predominantly rigid, with a maximum value at 1.73 sR_g and a decrease towards zero at higher sR_g, indicating that the CARDs do not add to the overall flexibility of FL-RIG-I. FL-RIG-I was modelled with the two CARD domains (from PDB ID: 4A2W), the helicase domain (from PDB ID: 4A2W) and the CTD (PDB ID: 3OG8) as rigid bodies against the DC- and FL-RIG-I SAXS data using BUNCH (31). This model is superposed within the *ab initio* shape reconstruction of FL-RIG-I calculated using DAMMIF and DAMAVER (30,31) in Figure 3B, and within the multi-phase *ab initio* reconstructed shape for FL-RIG-I

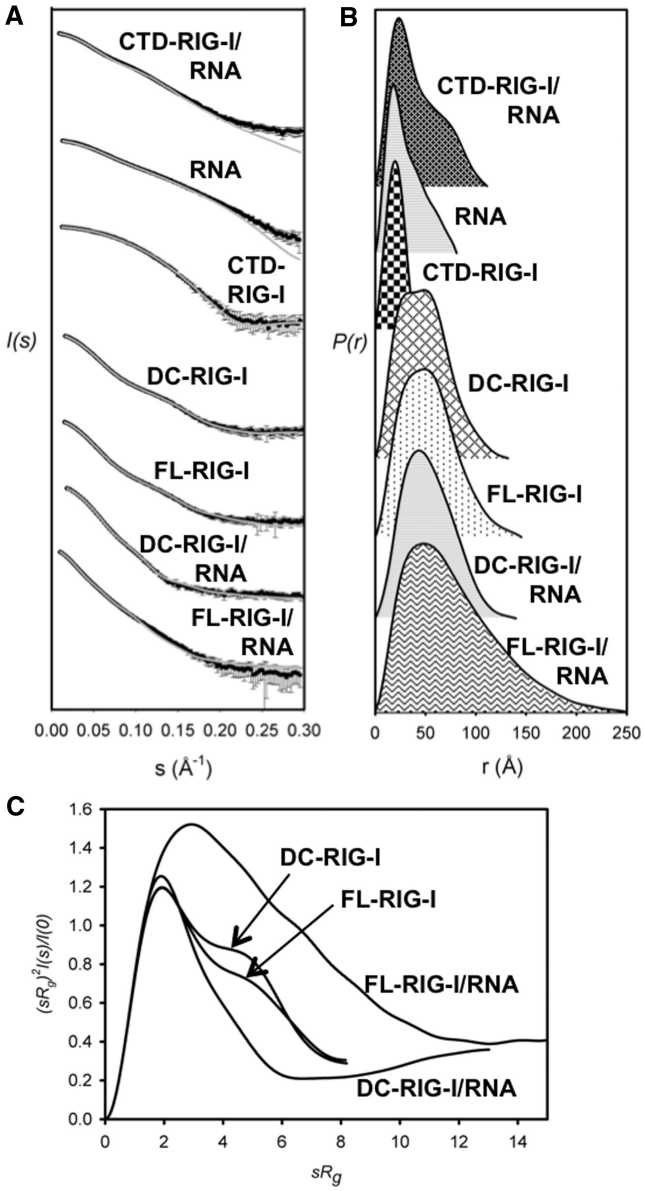

Figure 2. SAXS analysis of CTD-, DC- and FL-RIG-I before and after complex formation with 29mer dsRNA. (**A**) Scattering data for the three RIG-I proteins ± 29mer RNA and 29mer RNA alone (black, with error bars indicating the mean ± SD). The predicted scattering curves derived from models of these proteins are superposed (grey lines). The scattering intensity (I(s)) is plotted in arbitrary units versus the forward scattering vector (s) in units of inverse Å. (**B**) Pairwise distribution (P(r)) profiles for the three RIG-I proteins ± 29mer RNA and 29mer RNA alone, filled with different patterns for clarity, plotted against r in units of Å. (**C**) Normalized Kratky plots for DC- and FL-RIG-I before and after complex formation with dsRNA (29mer).

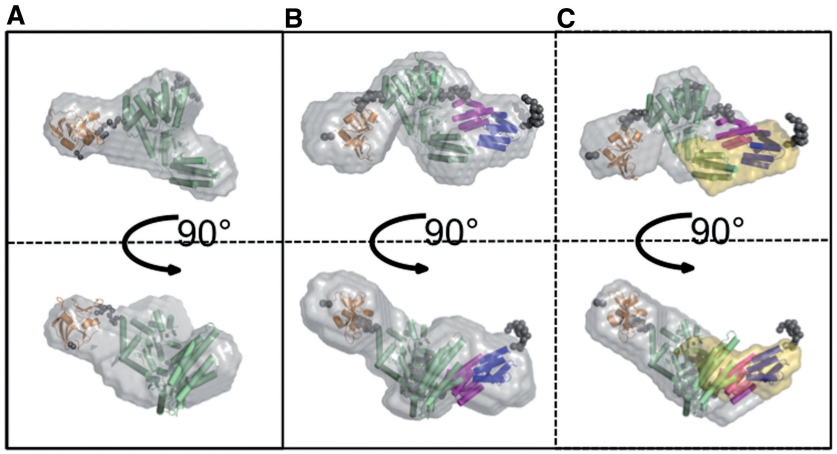

Figure 3. Models of DC- and FL-RIG-I derived from SAXS data. *Ab initio* reconstruction of DC-RIG-I (**A**) and FL-RIG-I (**B**) calculated from SAXS data using DAMMIF (30) and DAMAVER (35) shown as a grey molecular envelope. Superposed is a DC-RIG-I model generated using BUNCH (31). The helicase domain is coloured green, and the CTD orange. Linker regions are shown as grey spheres. (**C**) Multiphase *ab initio* reconstruction of FL-RIG-I calculated from SAXS data using MONSA (32) shown as a grey (helicase and CTD domains) and yellow (CARDs) molecular envelope indicating where DC-RIG-I and CARDs are positioned, respectively. Superposed is a model generated using BUNCH (31). The helicase domain is coloured green, the CTD orange and CARD domains 1 and 2 purple and pink, respectively.

calculated from DC- and FL-RIG-I SAXS data using MONSA (32) Figure 3C.

The resulting model is consistent with the positioning of the CARD domains against the helicase domain [in close agreement with the crystallographically derived structure (27)] and the CTD positioned near the C-terminal end of the helicase domain, as similarly modelled for DC-RIG-I. Good agreement between the predicted scattering curves from this model and the raw scattering data was observed (χ value = 0.41/0.51 DC-/FL-RIG-I) (Figure 2A and Supplementary Table S1). Together, these data show that apo-FL-RIG-I adopts an elongated arrangement of its helicase, CTD and CARD domains.

RIG-I domain arrangement on forming a 2:1 protein:dsRNA complex

The main goal of this study was to observe the change in conformation that occurs on the binding of multiple RIG-I molecules to an immunostimulatory dsRNA. A 29mer dsRNA was used that has been shown previously to be more stimulatory in p56 induction assays than shorter (<20 bp) dsRNA (13). While the crystal structures of RIG-I helicase domain with blunt-ended 10mer dsRNA (25), 14mer dsRNA (24) and 19mer dsRNA (27) formed 1:1 protein to oligonucleotide complexes, a 2:1 protein:dsRNA complex was formed with the 29mer, as this dsRNA can accommodate two RIG-I molecules.

The SAXS data for the DC-RIG-I:RNA complex gave rise to an R_g of 39.4 Å and R_{max} = 138, and the P(r) is almost unimodal with a slight shoulder. In addition, the normalized Kratky plot for the DC-RIG-I:29mer was bell shaped in appearance, returning towards zero at high sR_g and with a maxima at 1.73 sR_g. This is indicative that the complex is more ordered than the apo-DC-RIG-I, in agreement with observations by Jiang *et al.* (24) with a shorter dsRNA. Together, this is consistent with the formation of a compact interaction of DC-RIG-I with the dsRNA. A multiphase *ab initio* SAXS reconstruction for

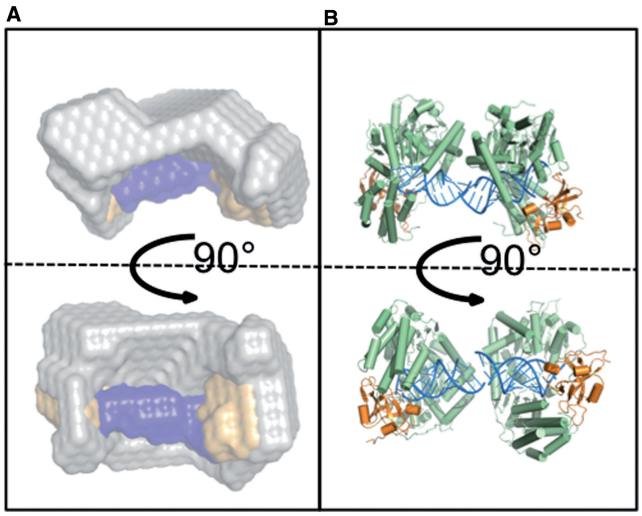

Figure 4. Models of DC-RIG-I with dsRNA 29mer derived from SAXS data. (**A**) Multiphase *ab initio* reconstruction of DC-RIG-I:29mer complex calculated from SAXS data using MONSA (32) shown as a molecular envelope (with grey indicating the helicase domains, yellow the CTDs and blue the 29mer RNA). (**B**) Model of DC-RIG-I:29mer RNA complex generated using SASREF (31). The helicase domain is coloured green, the CTD yellow and the RNA blue.

DC-RIG-I in complex with 29mer in a 2:1 ratio was calculated from four SAXS data sets (29mer RNA, CTD-RIG-I, CTD-RIG-I:29mer and DC-RIG-I:29mer) using MONSA (32) (Figure 4A). The molecular envelopes show the positioning of the CTD domains (yellow) at each end of the dsRNA (blue) and the helicase domains (grey) wrapped around this core, with some resemblance to a hotdog. A model of the DC-RIG-I:29mer complex was derived with the program SASREF (31) using two copies of the crystal structure of the helicase and CTD domain bound to 14mer dsRNA (PDB ID: 3TMI) as

rigid bodies connected via a 12-Å linker (to mimic a central extra base and to allow for flexibility) between the central phosphates of the dsRNA (Figure 4B). A good fit to the DC-RIG-I:29mer SAXS data was obtained when the two sections of dsRNA were not perfectly linear with respect to one another, but met at an angle of ~120 degrees (χ value = 0.85) [when the dsRNA was not allowed to bend the fit of the DC-RIG-I:29mer model to the SAXS, scattering was much poorer (χ value = 2.7)]. This resulted in a molecular model consistent with the arrangement of domains obtained by the multiphase *ab initio* SAXS reconstruction of the molecular envelope. These data thus reveal the way in which two DC-RIG-I molecules are arranged on the 29mer dsRNA and suggest that dsRNA may bend slightly on the binding of two RIG-I molecules.

In contrast, the SAXS data for FL-RIG-I in complex with dsRNA resulted in the formation of a greatly elongated complex structure, consistent with the release of CARDs. The R_g for FL-RIG-I:RNA was determined to be 62.1 Å, with R_{max} = 255 Å, which are both far larger than those of the DC-RIG-I:RNA complex and FL-RIG-I. The pairwise distribution P(r) is representative of an elongated molecule. The normalized Kratky analysis of the FL-RIG-I:RNA complex reveals that, compared with the DC RIG-I:RNA complex, there is an increase in the degree of flexibility exhibited by the complex (as indicated by the more gradual return of the curve at higher sR_g and a maxima at almost 3 sR_g). This is consistent with the CARD domains extending away from the rest of the complex on dsRNA binding by RIG-I. The EOM (33) was used to examine the position of the CARD domains in the 2:1 FL-RIG-I:29mer complex. Figure 5A shows the R_g and R_{max} distributions of the ensemble of randomly generated model structures (solid line) and the selection of structures from within the ensemble that together give rise to predicted SAXS data consistent with that obtained experimentally. The selected structures give rise to one major population that is larger, both in terms of the R_g and R_{max}, than the random pool. The scattering data for FL-RIG-I:29mer was best fit by a population of structures that have an average R_g of 64.4 Å and an R_{max} of 242.2 Å. These values are very close to those estimated from both Guinier and P(r) methods (Supplementary Table S1). The ensemble of structures selected by the EOM is shown in Figure 5B and is consistent with a highly flexible amino acid linker between the helicase and CARD domains, allowing the CARD domains to move freely with respect to the helicase domain.

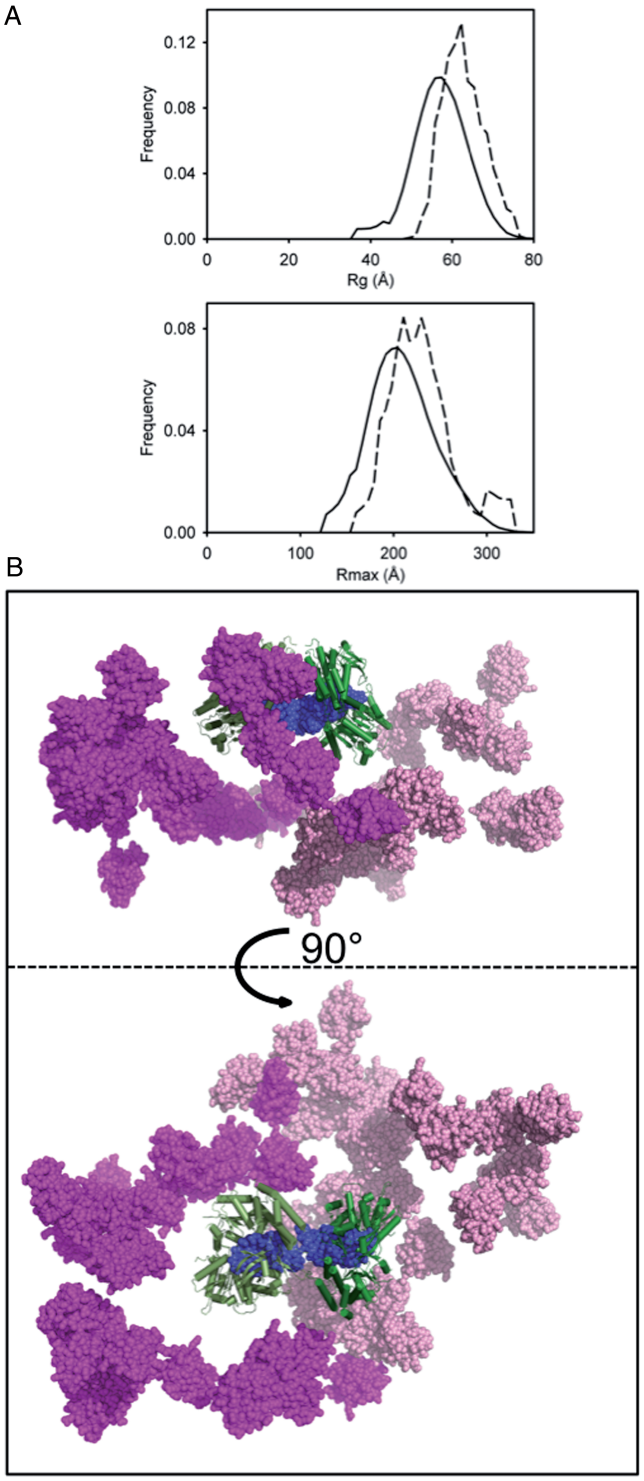

Figure 5. Ensemble optimization analysis of FL-RIG-I bound to dsRNA 29mer derived from SAXS data. (**A**) R_g (top graph) and R_{max} (bottom graph) distributions from the best-fitting ensembles calculated using EOM (33). The distribution for the pool of 10000 conformers (solid line) and the selected best-fitting ensemble (dashed line) are shown for the FL-RIG-I:29mer sample. (**B**) Superposition of the best-fitting ensemble from the main peaks shown as cartoon structures. The CTD and helicase domains (green) and RNA (blue) are directly superposed. The CARD domains (shown in two shades of pink) adopt a wide range of positions relative to the rest of the complex.

Tryptic digest of RIG-I in apo- and dsRNA-bound forms

To further verify that the CARDs become extended on the interaction of RIG-I with dsRNA 29mer, RIG-I was subjected to trypsin digestion in apo-form and bound to short (8mer) and long (29mer) dsRNA (Figure 6). The digestion pattern shows that in all cases, over the course of 90 min, the full-length RIG-1 (band 1; ~107 kDa) is digested, with the formation of an ~90-kDa band (band 2). On N-terminal sequencing, this was determined to represent

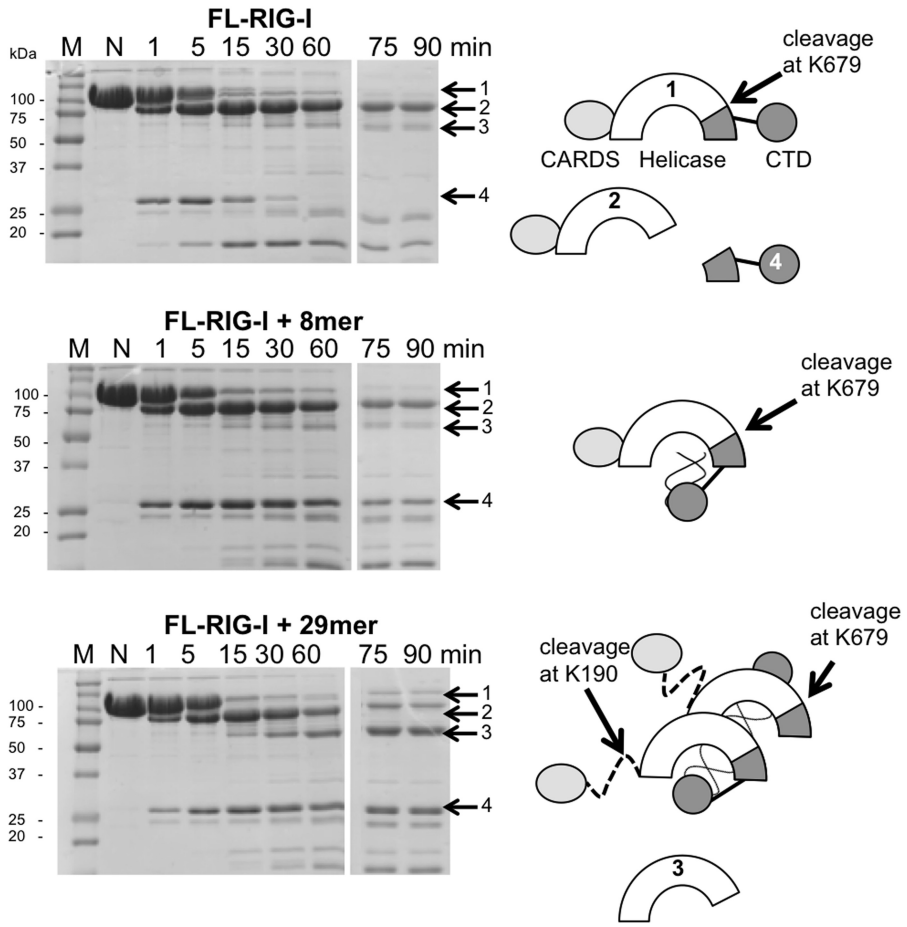

Figure 6. Tryptic digest pattern of FL-RIG-I with dsRNA of different lengths. Limited proteolysis was performed with RIG-I apo-form, RIG-I bound to 8mer or RIG-I bound to 29mer. Samples were taken at the indicated time points and subjected to SDS–PAGE. Specific protein fragments (indicated by arrows) were noted to persist for FL-RIG-I bound to 29mer, suggesting that the CARDs contribute to proteolytic resistance only when bound to longer dsRNA that allows multiple RIG-I binding. The N-terminal sequence of selected bands indicated the sequence of band 2 (arrow 2) to be the N-terminus, and band 3 (arrow 3) to represent cleavage site K190 and band 4 (arrow 4) K679.

the N-terminal portion of the molecule (with cleavage of the CTD), indicating that the CTD is vulnerable to trypsin cleavage in the presence or absence of dsRNA. Notably, a cleavage product of ~56 kDa (band 3) is produced in a higher amount when RIG-I is in complex with the 29mer RNA. The N-terminal sequence of this band revealed that the cleavage site is K190 (located in the linker region between CARD 2 and the helicase domain). Band 3, therefore, reflects the cleavage of the CARD domains (and CTD) from RIG-I and is consistent with greater accessibility of this cleavage site when RIG-I is bound to 29mer dsRNA. A fourth band (band 4, ~28 kDa) is seen to arise in all cases, but to persist significantly longer when RIG-I is in the presence of dsRNA, short or long. N-terminal sequencing determined that band 4 represents CTD from cleavage at K679, and its persistence in solution in the presence of oligonucleotide suggests that it is relatively protected from protease activity by either 8mer or 29mer dsRNA. Overall, these data support the observation that the CARD domains are completely released CARD domains on interaction with the long dsRNA, but not the short dsRNA, although both are bound by RIG-I.

DISCUSSION

RIG-I plays a critical role in detecting the presence of viral RNA and triggering anti-viral cellular responses. It does so through the direct recognition of the dsRNA end structure via its CTD domain and the formation of an oligomeric protein:RNA complex, and has been suggested to undergo a conformational change that allows its CARD domains to be ubiquitinated by tripartite motif 25 (TRIM25) and to interact with its downstream target IPS-1 (36). Recent crystallographic studies of RIG-I have revealed the way in which the helicase and CTD domains of RIG-I form a ring around the dsRNA (24,25), which implies that the CARDs are displaced from their positioning in the apo-RIG-I structure (27). The studies to date do not reveal where the CTD is positioned in the apo-RIG-I structure, nor how a second helicase domain may be positioned relative to the first, and most importantly, they do not reveal the position of the CARDs on the formation of an immunostimulatory multimeric RIG-I:RNA complex.

The current study describes the SAXS-derived models of human RIG-I before and after forming a 2:1 protein:RNA complex with a stimulatory 29mer dsRNA.

SAXS data were obtained for rigorously prepared samples by the collection of data directly as samples eluted from a size-exclusion column, ensuring sample homogeneity and avoiding aggregation. In addition, the collection of SAXS data for CTD-RIG-I, DC-RIG-I and FL-RIG-I, the dsRNA 29mer alone and each of the proteins in complex with the 29mer, facilitated the use of multiphase modelling of the RIG-I structures against the SAXS data.

The structure of apo-DC-RIG-I reveals that the CTD exists alongside the helicase domain. The FL-RIG-I forms a longer structure, with the CARDs positioned away from the CTD, juxtaposed against the helicase domain as also observed in the crystallographically derived structure of duck RIG-I (27). The distal positioning of the CARDs and the CTD has not previously been observed, as the crystal structure of duck RIG-I did not show electron density for the CTD, suggesting it was too mobile to be observed. It is possible that human and duck RIG-I structures differ in the mobility of the CTD, or that the SAXS method is more forgiving to limited molecular flexibility. Importantly, this model reveals that the CARDs and CTD are not intimately associated with each other, as has previously been hypothesized to underlie the repressed state of RIG-I (15). Our data are similar to previously reported SAXS data for DC-RIG-I and FL-RIG-I (24). Their data were not used for modelling but gave rise to similar R_g and R_{max} values, with the exception of the longer maximal dimension of FL-RIG-I (180 Å determined by Jiang *et al.*, compared with 138.6 Å). On this basis, they concluded that RIG-I becomes much more compact on binding to dsRNA. Our data for apo-FL-RIG-I indicate that it is not so extended.

The formation of a 2:1 RIG-I:dsRNA complex provides new insight into the way in which immune stimulation may be initiated by longer dsRNA species. Crystallographic efforts have been successful in delineating the structural features of single RIG-I proteins bound to shorter dsRNA (10mer–19mer), but no crystallographic details for multimeric RIG-I:dsRNA assemblies have been reported to date. The current study shows that CTD-RIG-I is able to bind to both ends of the dsRNA 29mer. Its positioning is consistent with the crystallographically derived structures of CTD:dsRNA complexes (8,21,22). Notably, however, on binding the dsRNA, DC-RIG-I does not form the dumbbell-like shape that would occur if DC-RIG-I:RNA crystal structures were modelled at each end of a dsRNA 29mer. Better fits to the SAXS data were obtained from models in which the helicase domains are slightly tilted towards each other and the dsRNA allowed to bend. This suggests that RIG-I binding affects the dsRNA structure and thereby influences the second RIG-I binding event, or that there is a predisposition for intermolecular RIG-I interactions that influence the structure of the multimeric complex. Aside from this, the data are consistent with the helicase and CTD forming two rings around the dsRNA in an arrangement similar to that seen for DC-RIG-I bound to shorter dsRNA.

This SAXS-derived 2:1 DC-RIG-I:dsRNA model is similar to the overall shape derived from reconstructed EM images of a DC-RIG-I bound to a dsRNA 24mer, but differs from the interpretation of the EM data (28). In the EM-derived model, two CTDs are placed at the same end of the complex, positioning the RIG-I molecules in a head-to-head orientation on the dsRNA. In contrast, our data position the CTDs at opposite ends of the dsRNA, forming a head-to-tail arrangement. In both cases, the helicase domains of the DC-RIG-I remain closely associated when bound to the dsRNA.

Based on the molecular arrangement of DC-RIG-I on the 29mer, SAXS data were then interpreted to model FL-RIG-I bound to the dsRNA. The SAXS data unambiguously indicate a greatly elongated structure, suggesting an extension of the CARD domains away from the rest of the molecule on dsRNA binding. The molecular model of FL-RIG-I:29mer dsRNA that best fits the scattering data is described as an ensemble of structures, with the CARDs moving freely at the end of an elongated stretch of 64 amino acids connecting them to the helicase domains. Protease digestion patterns also support the increased accessibility of the linker between the CARDs and helicase domain. This is in contrast to the SAXS-derived model for a single FL-RIG-I bound to a 10mer dsRNA (24). The model reported by Jiang *et al.* (24) shows the CARD domains protruding rigidly from the helicase, not unlike their positioning in apo-RIG-I (27). Here we demonstrate a complete release of the CARDs on the formation of a multimeric assembly of RIG-I on a longer dsRNA. It is possible that the complete release of the CARD domains is facilitated only when a second RIG-I molecule binds alongside the first, explaining the increased immunostimulatory effect of longer dsRNA. Thus, on binding of two or more RIG-I molecules to the dsRNA, the CARDs are no longer held next to the rest of RIG-I but are readily available for ubiquitination at K172 and subsequent signalling via IPS-1.

With respect to natural RIG-I ligands, RIG-I recognizes viral genomic RNA when 5′-triphosphate ssRNA is base-paired with a self-complementary sequence, to form blunt-ended dsRNA termini displaying a 5′ triphosphate with panhandle secondary structure, as typically found in ssRNA negative-strand viruses (9). The work of Schlee *et al.* (9) suggests that a stretch of at least 19 nucleotides must be base-paired at the termini of viral RNA for RIG-I to be activated. Given that the footprint of RIG-I is ~9 bp (37), it is therefore plausible that two RIG-I molecules can be accommodated on the panhandle structure, positioned along the 19-bp stretch of self-complementary viral RNA between the blunt-ended termini and a region of mismatch/hairpin loop. It is feasible that RIG-I associates with the termini of blunt-ended base-paired viral RNA in a manner equivalent to that detailed in this study. However, the conformation that is adopted by RIG-I, particularly the positions of the CTD and CARDs when RIG-I is associated with an adjacent hairpin/mismatched region of a panhandle structure, is not known and warrants further investigation.

In summary, our data provide insight into the conformational change that takes place during RIG-I multimerization on an immunostimulatory dsRNA. The binding of full-length RIG-I with an immunostimulatory

dsRNA results in a 2:1 multimeric complex in which the CARDs exhibit a large-scale movement and thus expose the region that has been shown to be ubiquitinated by TRIM25. It may be that the formation of such flexible multimeric structures underlies the greater immunostimulatory activity of longer dsRNAs that are able to accommodate more than one RIG-I molecule.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR online: Supplementary Table 1, Supplementary Methods and Supplementary References [38–43].

ACKNOWLEDGEMENTS

The authors acknowledge the Australian Synchrotron, where SAXS data were collected, and the assistance of beamline scientist Nigel Kirby. S.A.B., A.J.S., B.R.G.W., J.A.W. and M.C.J.W. designed research; S.A.B., J.B., A.R., D.W. and M.C.J.W. performed research; M.J. and K.J.-H. contributed new reagents; S.A.B., J.A.W. and M.C.J.W. analysed data; S.A.B., J.A.W. and M.C.J.W. wrote the paper.

FUNDING

National Health and Medical Research Council of Australia through a Project Grant [1006590 to B.R.G.W., M.C.J.W. and J.A.W.], and a Senior Research Fellowship to M.C.J.W., a Linkage Infrastructure, Equipment and Facilities Grant Australian Research Council [LE100100207 to M.C.J.W and B.R.G.W.], and the Victorian Government's Operational Infrastructure Support Program. Funding for open access charge: National Health and Medical Research Council of Australia.

Conflict of interest statement. K.J.-H. is employed by Sanofi Deutschland GmbH and M.J. was employed by Roche Kulmbach GmbH, 95326 Kulmbach, Germany, which may gain or lose financially as a result of the publication of this article.

REFERENCES

1. Kato,H., Sato,S., Yoneyama,M., Yamamoto,M., Uematsu,S., Matsui,K., Tsujimura,T., Takeda,K., Fujita,T., Takeuchi,O. *et al.* (2005) Cell type-specific involvement of RIG-I in antiviral response. *Immunity*, **23**, 19–28.
2. Loo,Y.M. and Gale,M. Jr (2011) Immune signaling by RIG-I-like receptors. *Immunity*, **34**, 680–692.
3. Onoguchi,K., Yoneyama,M. and Fujita,T. (2010) Retinoic acid-inducible gene-I-like receptors. *J. Interferon Cytokine Res.*, **31**, 27–31.
4. Schlee,M. and Hartmann,G. (2010) The chase for the RIG-I ligand–recent advances. *Mol. Ther.*, **18**, 1254–1262.
5. Yoneyama,M. and Fujita,T. (2010) Recognition of viral nucleic acids in innate immunity. *Rev. Med. Virol.*, **20**, 4–22.
6. Kato,H., Takahasi,K. and Fujita,T. (2011) RIG-I-like receptors: cytoplasmic sensors for non-self RNA. *Immunol. Rev.*, **243**, 91–98.
7. Zhao,M., Zhang,J., Phatnani,H., Scheu,S. and Maniatis,T. (2012) Stochastic expression of the interferon-beta gene. *PLoS Biol.*, **10**, e1001249.
8. Wang,Y., Ludwig,J., Schuberth,C., Goldeck,M., Schlee,M., Li,H., Juranek,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.
9. 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.
10. Rehwinkel,J., Tan,C.P., Goubau,D., Schulz,O., Pichlmair,A., Bier,K., Robb,N., Vreede,F., Barclay,W., Fodor,E. *et al.* (2010) RIG-I detects viral genomic RNA during negative-strand RNA virus infection. *Cell*, **140**, 397–408.
11. Takahasi,K., Yoneyama,M., Nishihori,T., Hirai,R., Kumeta,H., Narita,R., Gale,M. Jr, 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.
12. Malathi,K., Saito,T., Crochet,N., Barton,D.J., Gale,M. Jr and Silverman,R.H. (2010) RNase L releases a small RNA from HCV RNA that refolds into a potent PAMP. *RNA*, **16**, 2108–2119.
13. 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.
14. 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.
15. Saito,T., Hirai,R., Loo,Y.M., Owen,D., Johnson,C.L., Sinha,S.C., Akira,S., Fujita,T. and Gale,M. Jr (2007) Regulation of innate antiviral defenses through a shared repressor domain in RIG-I and LGP2. *Proc. Natl. Acad. Sci. USA*, **104**, 582–587.
16. Yoneyama,M., Kikuchi,M., Natsukawa,T., Shinobu,N., Imaizumi,T., Miyagishi,M., Taira,K., Akira,S. and Fujita,T. (2004) The RNA helicase RIG-I has an essential function in double-stranded RNA-induced innate antiviral responses. *Nat. Immunol.*, **5**, 730–737.
17. Potter,J.A., Randall,R.E. and Taylor,G.L. (2008) Crystal structure of human IPS-1/MAVS/VISA/Cardif caspase activation recruitment domain. *BMC Struct. Biol.*, **8**, 11.
18. Sumpter,R. Jr, Loo,Y.M., Foy,E., Li,K., Yoneyama,M., Fujita,T., Lemon,S.M. and Gale,M. Jr (2005) Regulating intracellular antiviral defense and permissiveness to hepatitis C virus RNA replication through a cellular RNA helicase, RIG-I. *J. Virol.*, **79**, 2689–2699.
19. Fujita,T., Onoguchi,K., Onomoto,K., Hirai,R. and Yoneyama,M. (2007) Triggering antiviral response by RIG-I-related RNA helicases. *Biochimie*, **89**, 754–760.
20. Cui,S., Eisenacher,K., Kirchhofer,A., Brzozka,K., Lammens,A., Lammens,K., Fujita,T., Conzelmann,K.K., Krug,A. and Hopfner,K.P. (2008) The C-terminal regulatory domain is the RNA 5'-triphosphate sensor of RIG-I. *Mol. Cell*, **29**, 169–179.
21. 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.
22. 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.
23. Civril,F., Bennett,M., Moldt,M., Deimling,T., Witte,G., Schiesser,S., Carell,T. and Hopfner,K.P. (2011) The RIG-I ATPase domain structure reveals insights into ATP-dependent antiviral signalling. *EMBO Rep.*, **12**, 1127–1134.
24. Jiang,F., Ramanathan,A., Miller,M.T., Tang,G.Q., Gale,M. Jr, Patel,S.S. and Marcotrigiano,J. (2011) Structural basis of RNA

recognition and activation by innate immune receptor RIG-I. *Nature*, **479**, 423–427.
25. Luo,D., Ding,S.C., Vela,A., Kohlway,A., Lindenbach,B.D. and Pyle,A.M. (2011) Structural insights into RNA recognition by RIG-I. *Cell*, **147**, 409–422.
26. Bamming,D. and Horvath,C.M. (2009) Regulation of signal transduction by enzymatically inactive antiviral RNA helicase proteins MDA5, RIG-I, and LGP2. *J. Biol. Chem.*, **284**, 9700–9712.
27. Kowalinski,E., Lunardi,T., McCarthy,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.
28. Ranjith-Kumar,C.T., Murali,A., Dong,W., Srisathiyanarayanan,D., Vaughan,R., Ortiz-Alacantara,J., Bhardwaj,K., Li,X., Li,P. and Kao,C.C. (2009) Agonist and antagonist recognition by RIG-I, a cytoplasmic innate immunity receptor. *J. Biol. Chem.*, **284**, 1155–1165.
29. Berke,I.C. and Modis,Y. (2012) MDA5 cooperatively forms dimers and ATP-sensitive filaments upon binding double-stranded RNA. *EMBO J.*, **31**, 1714–1726.
30. Franke,D. and Svergun,D.I. (2009) DAMMIF, a program for rapid ab initio shape determination in small-angle scattering. *J. Appl. Crystallogr.*, **42**, 342–346.
31. Petoukhov,M.V. and Svergun,D.I. (2005) Global rigid body modeling of macromolecular complexes against small-angle scattering data. *Biophys. J.*, **89**, 1237–1250.
32. Svergun,D.I. (1999) Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing. *Biophys. J.*, **76**, 2879–2886.
33. Bernado,P., Mylonas,E., Petoukhov,M.V., Blackledge,M. and Svergun,D.I. (2007) Structural Characterization of Flexible Proteins Using Small-Angle X-ray Scattering. *J. Am. Chem. Soc.*, **129**, 5656–5664.
34. Receveur-Brechot,V. and Durand,D. (2012) How random are intrinsically disordered proteins? A small angle scattering perspective. *Curr. Protein Pept. Sci.*, **13**, 55–75.
35. Volkov,V.V. and Svergun,D.I. (2003) Uniqueness of ab initio shape determination in small-angle scattering. *J. Appl. Crystallogr.*, **36**, 860–864.
36. Gack,M.U., Shin,Y.C., Joo,C.H., Urano,T., Liang,C., Sun,L., Takeuchi,O., Akira,S., Chen,Z., Inoue,S. *et al.* (2007) TRIM25 RING-finger E3 ubiquitin ligase is essential for RIG-I-mediated antiviral activity. *Nature*, **446**, 916–920.
37. Jiang,Q.X. and Chen,Z.J. (2012) Structural insights into the activation of RIG-I, a nanosensor for viral RNAs. *EMBO Rep.*, **13**, 7–8.
38. Gee,P., Chua,P.K., Gevorkyan,J., Klumpp,K., Najera,I., Swinney,D.C. and Deval,J. (2008) Essential role of the N-terminal domain in the regulation of RIG-I ATPase activity. *J. Biol. Chem.*, **283**, 9488–9496.
39. Kim,D.H., Behlke,M.A., Rose,S.D., Chang,M.S., Choi,S. and Rossi,J.J. (2005) Synthetic dsRNA Dicer substrates enhance RNAi potency and efficacy. *Nat. Biotechnol.*, **23**, 222–226.
40. Fischer,H., de Oliveira Neto,M., Napolitano,H.B., Polikarpov,I. and Craievich,A.F. (2010) Determination of the molecular weight of proteins in solution from a single small-angle x-ray scattering measurement on a relative scale. *J. Appl. Crystallogr.*, **43**, 101–109.
41. Svergun,D.I. (1992) Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *J. Appl. Crystallogr.*, **25**, 495–503.
42. Svergun,D.I., Barberato,C. and Koch,M.H.J. (1995) CRYSOL - a program to evaluate x-ray solution scattering of biological macromolecules from atomic coordinates. *J. Appl. Crystallogr.*, **28**, 768–773.
43. Whitten,A.E., Cai,S. and Trewhella,J. (2008) MULCh: ModULes for the analysis of small-angle neutron Contrast variation data from biomolecular assemblies. *J. Appl. Crystallogr.*, **41**, 222–226.