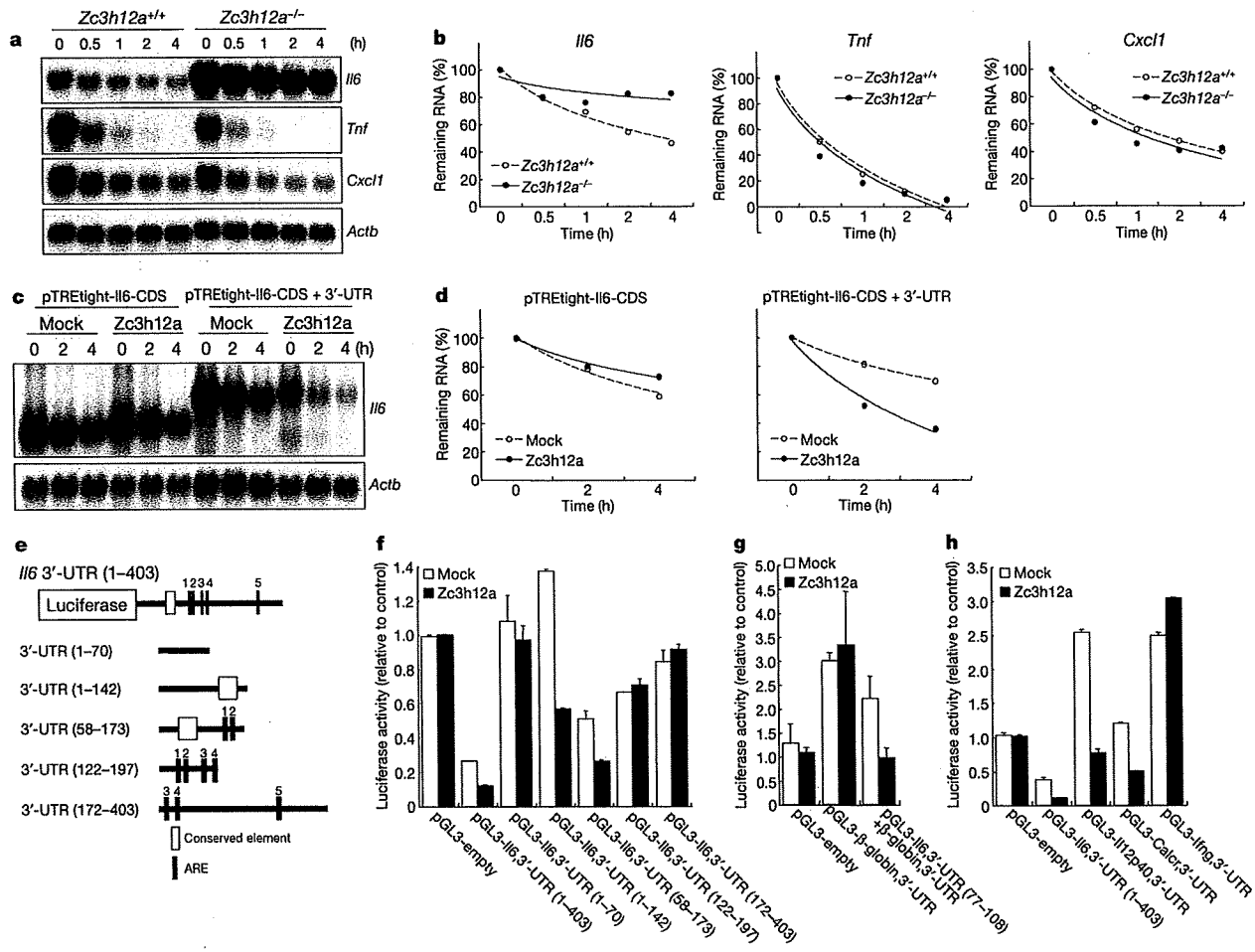


Figure 3 | Zc3h12a destabilizes mRNA from a set of genes through their 3′-UTRs. a, b, Peritoneal macrophages were treated with LPS (100 ng ml^{-1}) for 120 min and then treated with actinomycin D (2 μg ml^{-1}) for the indicated times. Total RNA (10 μg) was extracted and subjected to RNA blot analysis for the expression of *Il6*, *Tnf*, *Cxcl1* and β-actin (*Actb*) probes (**a**). Similar results were obtained in three independent experiments. The autoradiograph was quantified and the ratio of *Il6*, *Tnf* and *Cxcl1* to *Actb* was used to determine the remaining mRNA levels (**b**). **c, d,** HEK293 Tet-off cells were cotransfected with pTREtight-Il6-CDS or pTREtight-Il6-CDS + 3′-UTR, together with the Zc3h12a expression plasmid or control (empty) plasmid. Cells were divided 3 h after transfection and incubated overnight. Total RNA was prepared after dox (1 μg ml^{-1}) treatment, and *Il6* and *Actb* levels were determined by northern blot analysis (**c**). The autoradiograph was quantified and the ratio of *Il6* to *Actb* was used to determine remaining mRNA levels (**d**). **e–g,** Determination of Zc3h12a responsive regions in the *Il6* 3′-UTR. Schematic presentation of *Il6* 3′-UTR and its deletion constructs (**e**). HEK293 cells were transfected with the indicated pGL3 plasmids containing various sequences of *Il6* 3′-UTR (**f**) and β-globin 3′-UTR (**g**), together with the Zc3h12a expression plasmid or control (empty) plasmid. The luciferase activity was determined after 48 h. **h,** HEK293 cells were transfected with pGL3 harbouring the 3′-UTR for *Il6*, *Il12p40*, *Calcr* or *Ifng*, together with the Zc3h12a expression plasmid or control (empty) plasmid and the luciferase activity was determined after 48 h. Error bars indicate the s.d. of duplicates. Similar results were obtained in three independent experiments.

degrade RNA, indicating that the conserved pocket indeed functions as an RNase active site (Fig. 4f, h). The Zc3h12a(D141N) mutant failed to destabilize RNA containing the *Il6* 3′-UTR, indicating that the RNase activity is essential for the function of Zc3h12a (Fig. 4c, i).

This study clearly demonstrates that Zc3h12a is essential for the inhibition of the development of severe autoimmune responses culminating in the lethality of mice. Production of IL-6 and IL-12p40, but not TNF, was increased in *Zc3h12a*$^{-/-}$ macrophages due to mRNA decay failure. CCCH-type zinc-finger proteins have been shown to control mRNA decay by associating with the 3′-UTR. For example, tristetraprolin (Ttp, also known as Zfp36) and its homologues Zfp36l1, Zfp36l2 and Zfp36l3, are critical for the decay of the mRNAs for TNF, GM-CSF, CXCL1 and so on[4,9]. Aged *Ttp*$^{-/-}$ mice develop autoimmune arthritis owing to TNF production[10]. However, to our knowledge, there is no report showing that *Ttp*$^{-/-}$ cells produce increased amounts of IL-6 in response to TLR stimulation. Interestingly, the loss of Zc3h12a did not affect the expression of *Tnf* mRNA in macrophages, indicating that Ttp and Zc3h12a control mRNA decay for different cytokines. Zc3h12a targeted RNA sequences other than AREs, and the IL-6 AREs seem to be regulated by an unknown Zc3h12a-independent mechanism. Considering the profound pathological findings observed in *Zc3h12a*$^{-/-}$ mice, genes other than *Il6* and *Il12p40* are probably critically involved in the pathogenesis too. Identification of Zc3h12a target genes in response to other stimuli or in other cell types will improve our knowledge of the whole mechanism of abnormalities observed in *Zc3h12a*$^{-/-}$ mice. Zc3h12a was recently reported to be a monocyte chemotactic protein-1 (MCP-1)-induced protein[11], and overexpression of Zc3h12a protein was shown to suppress cytokine production in macrophages through inhibition of NF-κB activation[12]. However, the present study shows that Zc3h12a is involved in mRNA decay, but not in TNF regulation, inconsistent with this report.

The Zc3h12a protein has intrinsic RNase activity responsible for the decay of *Il6* mRNA. The mechanism is unique compared to the regulation of other ARE-mediated mRNA decay pathways. For instance, Ttp has been shown to recruit deadenylases for removing polyA tails, facilitating the subsequent degradation of target mRNAs

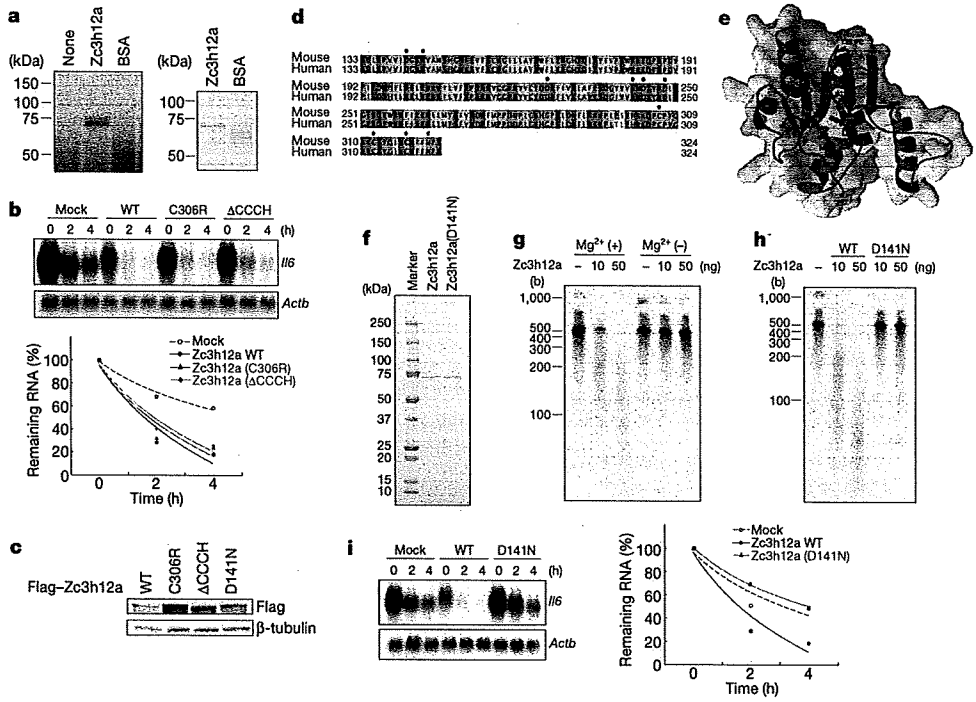


Figure 4 | Zc3h12a contains RNase activity essential for destabilizing *Il6* mRNA. a, Binding of Zc3h12a, but not BSA, to *Il6* 3′-UTR (1–403) mRNA by an ultraviolet cross-linking assay. **b**, Role of CCCH zinc-finger motif in destabilizing *Il6* mRNA. HEK293 Tet-off cells were cotransfected with pTREtight-Il6-CDS + 3′-UTR, together with expression plasmids encoding Flag–Zc3h12a and its mutants (C306R and ΔCCCH). Then cells were treated with dox for the indicated periods and *Il6* expression was determined by northern blot analysis (top). The autoradiograph was quantified and the ratio of *Il6* to *Actb* was used to determine remaining mRNA levels (bottom). WT, wild type. **c**, Expression levels of Zc3h12a mutant proteins determined by immunoblot. **d**, Alignment of N-terminal and CCCH domains in mouse and human Zc3h12a. **e**, The structural model of the Zc3h12a N-terminal domain. **f–h**, Zc3h12a possesses RNase activity. The expression levels of synthesized Zc3h12a and Zc3h12a(D141N) are shown (**f**). The RNase activity of Zc3h12a in degrading *Il6* 3′-UTR mRNA (1–403) in the presence or absence of 5 mM Mg^{2+} (**g**). The RNase activity of Zc3h12a and Zc3h12a(D141N) proteins (**h**). Synthesized RNA was incubated with increasing amounts of indicated proteins. RNA size marker is indicated in the left; 'b' denotes bases. **i**, HEK293 Tet-off cells were cotransfected with pTREtight-Il6 full and Zc3h12a(D141N). The cells were then treated with dox for the indicated periods and *Il6* expression was determined by northern blot analysis.

by exonucleases[4]. Thus, it is intriguing that Zc3h12a has endonuclease activity, which, at least *in vitro*, does not show sequence specificity. The target specificity may be determined by binding partner(s) of Zc3h12a, or Zc3h12a may have a preferential sequence for degradation under certain conditions. The mechanism of how Zc3h12a induces decay of mRNAs is an intriguing topic for further exploration. The RNase domain is conserved in four Zc3h12 family members, and the homologues of this protein family are found in metazoans such as *Drosophila melanogaster* (NCBI accession number CG10889) and *Caenorhabditis elegans* (NCBI accession number C30F12.1). Thus, regulation of mRNA by the RNase domain and CCCH zinc-finger domain seems to be evolutionally conserved.

Another RING-type ubiquitin ligase protein containing a CCCH zinc-finger motif called roquin (also known as Rc3h1) is essential for suppressing autoimmunity by controlling expression of the ICOS costimulatory molecule[13]. Roquin and several microRNAs seem to share an *Icos* 3′-UTR RNA segment for suppressing its degradation[14]. Given that each CCCH zinc-finger protein seems to have target mRNA specificity and 60 CCCH-type zinc-finger proteins have been identified in the mammalian genome[15], control of mRNA decay might be as important as the control of transcription in terms of regulation of innate immune responses. Future studies on CCCH zinc-finger proteins will be important for understanding the mechanisms of immune regulation by the control of mRNAs.

METHODS SUMMARY

Mice, reagents, cells and plasmids. Details are given in Methods.

ELISA. IL-4, IL-6, IL-12p40, IL-17, IFN-γ and TNF in culture supernatants, and mouse ANA in serum, were measured by ELISAs following manufacturer's protocols. ELISAs for mouse IgM, IgG1, IgG2a, IgG2b, IgG3 and anti-double-stranded DNA antibodies in serum were preformed as previously described[16,17].

Northern blotting, immunoblotting and EMSA. Northern blotting, immunoblotting and EMSA were performed as previously described[16].

Determination of haematological values. Haematological analysis of bloods prepared from wild-type and *Zc3h12a*$^{-/-}$ mice were performed at SRL Inc.

Flow cytometry. Details of flow cytometry are given in Methods. Cells were stained with indicated antibodies. The data were then acquired on a FACS Calibur or FACS Canto II flow cytometer (BD Biosciences), and analysed using FlowJo.

Measurement of RNA stability. mRNA stability was determined using three different methods. First, northern blot analysis was performed using total RNAs prepared from *Zc3h12a*$^{-/-}$ macrophages stimulated with LPS for 2 h followed by treating with actinomycin D. Second, the northern blot Tet-off system was used with HEK293 Tet-off cells transfected by pTRE-Il6 plasmids together with pFlag-Zc3h12a plasmids. Third, the luciferase assay of cell lysates prepared from HEK293 cells transfected with pGL3 plasmids together with pFlag-Zc3h12a plasmid was used.

***In vitro* RNA cleavage assay.** Cleavage activities of wild-type and mutant forms of Zc3h12a were analysed as described previously[18]. After incubation of recombinant Zc3h12a proteins with *in vitro* transcribed [^{32}P]-labelled RNAs, the cleaved RNAs were purified and analysed by denaturing PAGE and autoradiography. Details of the cleavage assay, expression of recombinant proteins and synthesis of [^{32}P]-labelled RNAs are given in Methods.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 23 January; accepted 25 February 2009.
Published online 25 March 2009.

1. Akira, S., Uematsu, S. & Takeuchi, O. Pathogen recognition and innate immunity. *Cell* 124, 783–801 (2006).

2. Beutler, B. *et al.* Genetic analysis of host resistance: Toll-like receptor signaling and immunity at large. *Annu. Rev. Immunol.* **24**, 353–389 (2006).
3. Medzhitov, R. Recognition of microorganisms and activation of the immune response. *Nature* **449**, 819–826 (2007).
4. Anderson, P. Post-transcriptional control of cytokine production. *Nature Immunol.* **9**, 353–359 (2008).
5. Barabino, S. M., Hubner, W., Jenny, A., Minvielle-Sebastia, L. & Keller, W. The 30-kD subunit of mammalian cleavage and polyadenylation specificity factor and its yeast homolog are RNA-binding zinc finger proteins. *Genes Dev.* **11**, 1703–1716 (1997).
6. Kanadia, R. N. *et al.* A muscleblind knockout model for myotonic dystrophy. *Science* **302**, 1978–1980 (2003).
7. Zhao, W., Liu, M. & Kirkwood, K. L. p38α stabilizes interleukin-6 mRNA via multiple AU-rich elements. *J. Biol. Chem.* **283**, 1778–1785 (2008).
8. Paschoud, S. *et al.* Destabilization of interleukin-6 mRNA requires a putative RNA stem-loop structure, an AU-rich element, and the RNA-binding protein AUF1. *Mol. Cell. Biol.* **26**, 8228–8241 (2006).
9. Datta, S. *et al.* Tristetraprolin regulates CXCL1 (KC) mRNA stability. *J. Immunol.* **180**, 2545–2552 (2008).
10. Taylor, G. A. *et al.* A pathogenetic role for TNF-α in the syndrome of cachexia, arthritis, and autoimmunity resulting from tristetraprolin (TTP) deficiency. *Immunity* **4**, 445–454 (1996).
11. Zhou, L. *et al.* Monocyte chemoattractant protein-1 induces a novel transcription factor that causes cardiac myocyte apoptosis and ventricular dysfunction. *Circ. Res.* **98**, 1177–1185 (2006).
12. Liang, J. *et al.* A novel CCCH-zinc finger protein family regulates proinflammatory activation of macrophages. *J. Biol. Chem.* **283**, 6337–6346 (2008).
13. Vinuesa, C. G. *et al.* A RING-type ubiquitin ligase family member required to repress follicular helper T cells and autoimmunity. *Nature* **435**, 452–458 (2005).
14. Yu, D. *et al.* Roquin represses autoimmunity by limiting inducible T-cell co-stimulator messenger RNA. *Nature* **450**, 299–303 (2007).
15. Liang, J., Song, W., Tromp, G., Kolattukudy, P. E. & Fu, M. Genome-wide survey and expression profiling of CCCH-zinc finger family reveals a functional module in macrophage activation. *PLoS One* **3**, e2880 (2008).
16. Sato, S. *et al.* Essential function for the kinase TAK1 in innate and adaptive immune responses. *Nature Immunol.* **6**, 1087–1095 (2005).
17. Fukuyama, H., Nimmerjahn, F. & Ravetch, J. V. The inhibitory Fcγ receptor modulates autoimmunity by limiting the accumulation of immunoglobulin G^+ anti-DNA plasma cells. *Nature Immunol.* **6**, 99–106 (2005).
18. Miyoshi, K., Uejima, H., Nagami-Okada, T., Siomi, H. & Siomi, M. C. *In vitro* RNA cleavage assay for Argonaute-family proteins. *Methods Mol. Biol.* **442**, 29–43 (2008).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank all colleagues in our laboratory, E. Kamada for secretarial assistance, and Y. Fujiwara, M. Kumagai and R. Abe for technical assistance. We thank S. Sato for discussions and W. Zhao and K. Kirkwood for plasmids. This work was supported by the Special Coordination Funds of the Japanese Ministry of Education, Culture, Sports, Science and Technology, grants from the Ministry of Health, Labour and Welfare in Japan, the Global Center of Excellence Program of Japan, and the NIH (P01 AI070167).

Author Contributions K.M. generated *Zc3h12a*$^{-/-}$ mice and performed most experiments. O.T. identified Zc3h12a, designed the research and wrote the paper. D.M.S. and H.N. carried out structural modelling. Y.K. analysed microarray data, and T.T. was responsible for histological analysis. T.K., T.M., T.S. and H.K. helped with experiments. S.A. designed the research and supervised the project.

Author Information Microarray data are deposited in the Gene Expression Omnibus (accession number GSE14890 for series of *Myd88*$^{-/-}$ and *Trif*$^{-/-}$ macrophages, and GSE14891 for series of *Zc3h12a*$^{-/-}$ macrophages). The structure model of Zc3h12a nuclease domain has been deposited in the Protein Model DataBase (PMDB) under accession number PM0075640. Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to S.A. (sakira@biken.osaka-u.ac.jp).

METHODS

Generation of *Zc3h12a*$^{-/-}$ mice. Genomic DNA containing *Zc3h12a* was isolated from GSI-I embryonic stem cells and characterized by restriction enzyme mapping and sequencing analysis. A targeting vector was designed to replace exon 3 to exon 5 containing the CCCH type zinc-finger domain, with a neomycin-resistance gene. A 1.1-kb ClaI–BamI fragment was used as the 3′ homology, and a 5.9-kb NotI–SalI fragment was used as the 5′ homology region. A total of 30 μg of NotI-linearized vector was electroporated into GSI-I embryonic stem cells. After selection with G418, drug-resistant clones were picked up and screened by PCR and Southern blot analysis. These clones were individually microinjected into blastocysts derived from C57BL/6 mice and transferred to pseudopregnant females. Matings of chimaeric male mice to C57BL/6 female mice resulted in the transmission of the mutant allele to the germ line. Resulting *Zc3h12a*$^{+/-}$ mice were intercrossed to generate *Zc3h12a*$^{-/-}$ mice. All animal experiments were done with the approval of the Animal Research Committee of the Research Institute for Microbial Diseases (Osaka University).

Reagents and cells. ELISA kits for mouse IL-4, IL-6, IL-12p40, IL-17, IFN-γ and TNF were purchased from R&D systems. The mouse ANA antibody ELISA kit was purchased from Alpha Diagnostic. Monoclonal anti-YY1 (H-10) and HRP-conjugated monoclonal anti-β-tubulin (D-10) antibodies were obtained from SantaCruz. HRP-conjugated anti-Flag antibody was purchased from Sigma. TLR ligands, including MALP-2, poly(I:C), LPS from *Salmonella Minnesota* Re595, R-848 and CpG oligonucleotide (ODN1668) were obtained as described previously[19].

Peritoneal exudate cells were isolated from the peritoneal cavities of mice 3 days after injection with 2 ml of 4.0% thioglycollate medium (Sigma) by washing with ice-cold Hank's buffered salt solution (Invitrogen). The HEK293 Tet-off cell line was purchased from Clontech.

Plasmids. *Zc3h12a* cDNA was inserted into a pFlag-CMV2 vector (Invitrogen). Point mutations (C306R and D141N) and deletion of the CCCH domain were carried out using the above-mentioned plasmid using QuickChangeII Site-Directed Mutagenesis Kit (Stratagene). pGL3 vector containing full-length (1–403) or parts (1–70, 58–173, 172–403) of *Il6* 3′-UTR sequences were supplied by W. Zhao and K. Kirkwood[7]. Parts (1–92, 1–102, 1–112, 1–132, 1–142 and 122–197) of *Il6* 3′-UTR cDNA were inserted in the pGL3 vector. 3′-UTR cDNA of β-globin (1–130) with or without *Il6* 3′-UTR (77–108) sequence, and the 3′-UTR cDNAs of *Il12p40* (1–781), *Calcr* (1–1601) and *Ifng* (1–631) were inserted in the pGL3 vector. *Il6* CDS and *Il6* CDS + 3′-UTR were inserted in pTREtight vector (Clontech). Wild-type and mutant (D141N) *Zc3h12a* cDNA were inserted in the pGEX-6P1 vector (GE Healthcare). *Il6* 3′-UTR cDNA was inserted downstream of the T7 promoter of pBluescript.

Flow cytometry. Antibodies for flow cytometry were purchased from BD Biosciences. Cell suspensions of spleen were prepared by sieving and gentle pipetting. For surface staining, cells were maintained in the dark at 4 °C throughout. Cells were washed in ice-cold FACS buffer (2% FCS, 0.02% NaN_3 in PBS), then incubated with each antibody for 15 min and washed twice with FACS buffer. Foxp3$^+$ regulatory T cell was stained using Mouse Regulatory T Cell Staining Kit (eBioscience) following the manufacturer's instructions. Intracellular cytokines were stained using Cytofix/Cytoparm Plus Fixation/Permeabilization Kit (BD Biosciences) following the manufacturer's instructions. Data were acquired on a FACS Calibur or FACS Canto II flow cytometer (BD Biosciences), and analysed using FlowJo.

Stability of mRNA in macrophages. Peritoneal macrophages (1×10^6) from wild-type and *Zc3h12a*$^{-/-}$ mice were stimulated with LPS (100 ng ml^{-1}) for 2 h. Actinomycin D (2 μg ml^{-1}) was then added to the culture medium to stop transcription, and total RNAs were prepared at the indicated time periods. The RNAs were subjected to northern blot analysis to determine *Il6*, *Tnf*, *Cxcl1* and *Actb* mRNA levels.

The Tet-off system. HEK293 Tet-off cells (3×10^6) were transfected with pTREtight-Il6-CDS (which have an *Il6* coding sequence) or pTREtight-Il6-CDS + 3′-UTR (which have a *Il6* coding and non-coding 3′-UTR sequence), together with wild-type or mutant forms of Zc3h12a expression plasmids or control (empty) plasmid. After 3 h, the cells were subdivided into three 60-mm dishes and cultured overnight. mRNA transcription from pTREtight vectors were terminated by the addition of dox (1 μg ml^{-1}), and total RNA was prepared at the indicated time periods. The RNA was subjected to northern blot analysis to determine *Il6* and *Actb* mRNA levels.

Luciferase assay. HEK293 cells were transfected with pGL3-Il6 3′-UTR plasmids or pGL3-empty plasmid together with Zc3h12a expression plasmid or empty control plasmid. After 48 h of cultivation, cells were lysed and luciferase activities in the lysates were determined using the Dual-luciferase reporter assay system (Promega). The *Renilla* luciferase gene was simultaneously transfected as an internal control.

Bone marrow transfer. Bone marrow cells were prepared from wild-type and *Zc3h12a*$^{-/-}$ mice. The cells were intravenously injected into lethally irradiated CD45.1 C57BL/6 mice. The chimaeric mice were given neomycin and ampicilin in their drinking water for 4 weeks. The mice were analysed at least 8 weeks after reconstitution. More than 90% of splenocytes from chimaeric mice were CD45.2-positive.

Expression of Zc3h12a protein in bacteria. The proteins were expressed in *Escherichia coli* BL21-Gold(DE3)pLysS (Stratagene) transformed with pGEX-6P1-Zc3h12a or Zc3h12a(D141N) mutant. After expression of the proteins, the cells were collected and resuspended in PBS. The cells were lysed by sonication followed by addition of Triton X-100 at a final concentration of 1% and incubation for 30 min at 4 °C with gentle shaking. The debris were then removed by centrifugation and supernatants were incubated with Glutathione Sepharose 4B (GE Healthcare) for 30 min at 4 °C with gentle shaking. The resins were collected and washed five times with PBS and resuspended in PreScission Protease cleavage buffer (50 mM Tris, 150 mM NaCl, 1 mM EDTA and 1 mM dithiothreitol (DTT)). PreScission Protease (GE Healthcare) (80 U) was added and incubated for 4 h at 4 °C with gentle shaking. Supernatants were collected and stored at −80 °C as purified recombinant protein solutions.

Synthesis of [^{32}P]-labelled RNAs. The pBluescript-Il6 3′-UTR (1–430) plasmid was used as a template for the synthesis of RNA having an *Il6* 3′-UTR sequence. *In vitro* RNA synthesis and [^{32}P] labelling was performed by using Riboprobe in vitro Transcription system (Promega) following manufacturer's instructions. The 5′-end labelling was performed using non-labelled RNA and Kinase Max 5′-end labelling Kit (Ambion) following manufacturer's instructions. The 3′-end labelling was performed by incubation of non-labelled RNA with T4 RNA Ligase (Takara) and [^{32}P]pCp (GE Healthcare).

RNA binding assay. [^{32}P]-labelled RNA (1×10^6 c.p.m.) were mixed with recombinant protein or BSA (Pierce) in a buffer (25 mM HEPES, 50 mM potassium acetate, 5 mM DTT) and incubated for 20 min at room temperature. Heparin was then added at a final concentration of 0.5 μg ml^{-1} and incubated further 10 min. The samples were cross-linked by irradiation with 254-nm ultraviolet light using FUNA-UV-LINKER FS-800 (Funakoshi) at a distance of 5 cm from the light source for 20 min on ice. The cross-linked samples were treated with RNaseT (100 U) for 20 min at room temperature, followed by treatment with RNaseA (1 μg) for 15 min at 37 °C. After the digestion, the proteins bound with [^{32}P]-labelled RNA were analysed by SDS–PAGE and autoradiography.

***In vitro* RNA cleavage assay.** Recombinant proteins and *in vitro* transcribed [^{32}P]-labelled RNAs (5,000 c.p.m.) were mixed in cleavage buffer (25 mM HEPES, 50 mM potassium acetate, 5 mM DTT) with or without 5 mM magnesium acetate[2] in the presence of Rnasin plus (40 U) (Promega). The cleaved RNA was purified with Trizol (Invitrogen) and analysed by denaturing PAGE using 6% TBE–urea gel (Invitrogen) and autoradiography.

Microarray analysis. Peritoneal macrophages from wild-type, *MyD88*$^{-/-}$ and *Trif*$^{-/-}$ mice were stimulated with 100 ng ml^{-1} LPS for 0, 1 and 4 h. Total RNA was extracted with an RNeasy kit (Qiagen), double-stranded cDNA was synthesized from 10 μg of total RNA with the SuperScript Choice System (Invitrogen) primed with T7-(dT) 24 primer. These cDNAs were used to prepare biotin-labelled complementary RNA by an *in vitro* transcription reaction performed using T7 RNA polymerase in the presence of biotinylated ribonucleotides, according to the manufacturer's protocol (Enzo Diagnostics). The cRNA product was purified using an RNeasy kit (Qiagen), fragmented, and hybridized to Affymetrix mouse expression array 430A microarray chips, according to the manufacturer's protocol. For determination of LPS-inducible genes in *Zc3h12a*$^{-/-}$ macrophages, peritoneal macrophages were stimulated with 100 ng ml^{-1} LPS. Total RNA was then extracted with Trizol (Invitrogen Life Technologies) and further purified using an RNeasy kit. Biotin-labelled cDNA was synthesized from 100 ng of the purified RNA using Ovation Biotin RNA Amplification and Labelling System (Nugen) according to the manufacturer's protocol. Hybridization, staining, washing and scanning of Affymetrix mouse Genome 430 2.0 microarray chip was done following the manufacturer's instructions. Robust multichip average (RMA) expression values were calculated using R and Bioconductor affy package. For hierarchical clustering, probes having a more than two- or fivefold increase compared to 0 h after stimulation were selected. The RMA expression values were transformed to fit averages and standard deviations to zero and one, respectively, by each probe. For analysis of LPS-inducible genes in *MyD88*$^{-/-}$ and *Trif*$^{-/-}$ macrophages, distances between probes were calculated using Pearson's correlation coefficient as a distance function. For analysis of LPS inducible genes in *Zc3h12a*$^{-/-}$ macrophages, principle component analysis for RMA values was performed and Euclidean distances between probes were computed using first to fifth principle components. Hierarchical clustering was carried out using these distances with Ward's method. These calculations and generation of heat map representation were carried out using R and Bioconductor.

Immunohistochemistry. The tissues was fixed with 10% formalin neutral buffer solution, embedded in paraffin, and cut into 5-μm thick sections. Sections were heated in Target Retrieval Solution (Dako) at 98 °C for 40 min to facilitate antigen retrieval. The sections were incubated with peroxidase-conjugated goat IgG fraction to mouse IgA (alpha chain) (MP Biomedicals) diluted 1:50 by antibody diluent (ChemMate, Dako), or peroxidase-conjugated goat affinity purified F(ab')2 fragment to mouse IgG (whole molecule) (MP Biomedicals) diluted 1:25 by antibody diluent, for 30 min at room temperature. Immunoreacted cells for mouse IgA and IgG were visualized with diaminobenzidine (Dako). The sections were lightly counterstained with haematoxylin.

Structure modelling. A model of the Zc3h12c N-terminal domain was constructed as follows: first, the sequence was submitted to the BioInfoBank Meta Server (http://bioinfo.pl), and the top ten models were built using default settings. The best model was then chosen by submitting each to the SeSAW functional annotation server (http://pdbjs6.pdbj.org/SeSAW/), and selecting the model with the highest score. The model chosen was built form the structural genomics template 2qip, using the FFAS03 server (http://ffas.ljcrf.edu/ffas-cgi/cgi/ffas.pl) and Modeller[20]. This model, which also had the highest three-dimensional Jury score, contained a cluster of conserved aspartic acids (D141, D226, S242, D244 and D248) that are also conserved in the active sites of Flap endonucleases (for example, Protein Data Bank ID 1UT5)[21]. Electrostatic surfaces were prepared using the *e*F-surf server (http://ef-site.hgc.jp/eF-surf/) and *e*F-site[22].

19. Kawagoe, T. *et al.* Sequential control of Toll-like receptor-dependent responses by IRAK1 and IRAK2. *Nature Immunol.* 9, 684–691 (2008).
20. Eswar, N. *et al.* Comparative protein structure modeling using MODELLER. *Curr. Protoc. Bioinformatics* Chapter 5, Unit 5.6 (2006).
21. Feng, M. *et al.* Roles of divalent metal ions in flap endonuclease-substrate interactions. *Nature Struct. Mol. Biol.* 11, 450–456 (2004).
22. Kinoshita, K. & Nakamura, H. eF-site and PDBjViewer: database and viewer for protein functional sites. *Bioinformatics* 20, 1329–1330 (2004).

Author's Choice

THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 284, NO. 20, pp. 13348–13354, May 15, 2009

Identification of Loss of Function Mutations in Human Genes Encoding RIG-I and MDA5

*IMPLICATIONS FOR RESISTANCE TO TYPE I DIABETES**

Received for publication, December 16, 2008, and in revised form, March 25, 2009 Published, JBC Papers in Press, March 26, 2009, DOI 10.1074/jbc.M809449200

Taeko Shigemoto[‡§], Maiko Kageyama[‡§], Reiko Hirai[‡], JiPing Zheng[‡], Mitsutoshi Yoneyama[‡§¶], and Takashi Fujita[‡§1]

From the [‡]Laboratory of Molecular Genetics, Institute for Virus Research, and [§]Laboratory of Molecular Cell Biology, Graduate School of Biostudies, Kyoto University, Kyoto 606-8507 and [¶]PRESTO, Japan Science and Technology Agency, 4-1-8 Honcho Kawaguchi, Saitama 332-0012, Japan

Retinoic acid-inducible gene I (RIG-I) and melanoma differentiation-associated gene 5 (MDA5) are essential for detecting viral RNA and triggering antiviral responses, including production of type I interferon. We analyzed the phenotype of non-synonymous mutants of human RIG-I and MDA5 reported in databases by functional complementation in cell cultures. Of seven missense mutations of RIG-I, S183I, which occurs within the second caspase recruitment domain repeat, inactivated this domain and conferred a dominant inhibitory function. Of 10 mutants of MDA5, two exhibited loss of function. A nonsense mutation, E627*, resulted in deletion of the C-terminal region and double-stranded RNA (dsRNA) binding activity. Another loss of function mutation, I923V, which occurs within the C-terminal domain, did not affect dsRNA binding activity, suggesting a novel and essential role for this residue in the signaling. Remarkably, these mutations are implicated in resistance to type I diabetes. However, the A946T mutation of MDA5, which has been implicated in type I diabetes by previous genetic analyses, affected neither dsRNA binding nor *IFN* gene activation. These results provide new insights into the structure-function relationship of RIG-I-like receptors as well as into human RIG-I-like receptor polymorphisms, antiviral innate immunity, and autoimmune diseases.

Innate and adaptive immune systems constitute the defense against infections by pathogens. Immediately after an infection occurs, various cells in the body sense the virus and initiate antiviral responses in which type I IFN[2] plays a critical role, both in viral inhibition and in the subsequent adaptive immune response (1). The production of IFN is initiated when sensor molecules such as Toll-like receptors (TLRs) and RLRs detect virus-associated molecules. TLRs detect pathogen-associated molecular patterns (PAMPs) at the cell surface or in the endosome in immune cells such as dendritic cells and macrophages (2). RLRs sense viral RNA in the cytoplasm of most cell types and induce antiviral responses, including the activation of *IFN* genes (3). RLRs include RIG-I, MDA5, and laboratory of genetics and physiology 2 (LGP2).

It is proposed that RLRs sense and activate antiviral signals through the coordination of their functional domains (4). The N-terminal region of RIG-I and MDA5 is characterized by two repeats of CARD and functions as an activation domain (3). This domain is responsible for the transduction of signals downstream to IFN-β promoter stimulator 1 (IPS-1) (also known as MAVS, VISA, and Cardif). The primary sequence of the CTD, consisting of ~140 amino acids, is conserved among RLRs. The CTD of RIG-I functions as a viral RNA-sensing domain as revealed by biochemical and structural analyses (5, 6). Both dsRNA and 5′-ppp-ssRNA, which are generated in the cytoplasm of virus-infected cells, are recognized by a basic cleft structure of RIG-I CTD. In addition to its RNA recognition function, the CTD of RIG-I and LGP2 functions as a repression domain through interaction with the activation domain. The repression domain is responsible for keeping RIG-I inactive in non-stimulated cells (3, 7). The helicase domain, with DEXD/H box-containing RNA helicase motifs, is the largest domain found in RLRs. Once dsRNA or 5′-ppp-ssRNA is recognized by the CTD, the helicase domain causes structural changes to release the activation domain. ATP binding and/or its hydrolysis is essential for the conformational change because Walker's ATP-binding site within the helicase domain is essential for signaling by RIG-I and MDA5.

Analyses of knock-out mice have revealed that RIG-I and MDA5 recognize distinct RNA viruses (8, 9). Picornaviruses are detected by MDA5, but many other viruses such as influenza A, Sendai, vesicular stomatitis, and Japanese encephalitis are detected by RIG-I. The difference is based on the distinct non-self RNA patterns generated by viruses, as demonstrated by the finding that RIG-I is selectively activated by dsRNA or 5′-ppp ssRNA, whereas MDA5 is activated by long dsRNA (10–12).

Single nucleotide polymorphisms (SNPs) of the human *RIG-I* and *MDA5* genes including several non-synonymous SNPs (nsSNPs), which potentially alter the function of the proteins encoded, are reported in databases. In this report, we investigated the functions of nsSNPs of RIG-I and MDA5 by functional complementation using respective knock-out cells. We identified

* This work was supported by grants from the Ministry of Education, Culture, Sports, Science and Technology of Japan, Japan Science and Technology Agency, The Mochida Memorial Foundation for Medical Pharmaceutical Research, and Nippon Boehringer Ingelheim.

Author's Choice—Final version full access.

[1] To whom correspondence should be addressed: Laboratory of Molecular Genetics, Institute for Virus Research, Kyoto University, Kyoto 606-8507, Japan. Tel.: 81-75-751-4031; Fax: 81-75-751-4031; E-mail: tfujita@virus.kyoto-u.ac.jp.

[2] The abbreviations used are: IFN, interferon; IPS-1, IFN-β promoter stimulator 1; RIG-I, retinoic acid-inducible gene I; MDA5, melanoma differentiation-associated gene 5; LGP2, laboratory of genetics and physiology 2; TLR, Toll-like receptor; RLR, RIG-I-like receptor; PAMP, pathogen-associated molecular pattern; CARD, caspase recruitment domain; CTD, C-terminal domain; SNP, single nucleotide polymorphisms; nsSNP, non-synonymous SNP; MEF, mouse embryonic fibroblast; SeV, Sendai virus; T1D, type I diabetes; dsRNA, double-stranded RNA; ssRNA, single-stranded RNA; EMSA, electrophoretic mobility shift assay; WT, wild type.

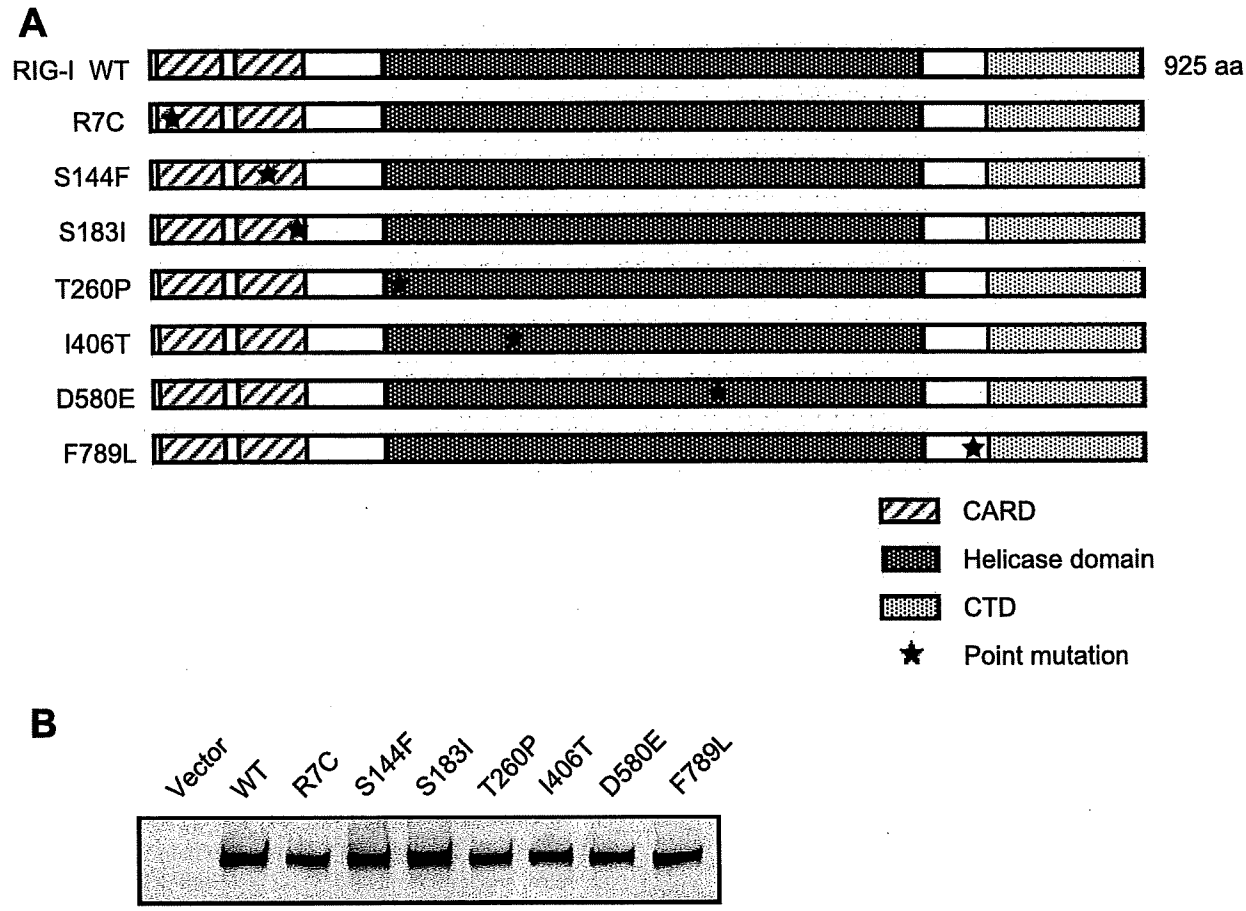


FIGURE 1. **RIG-I nsSNP mutants and their expression in MEFs.** *A*, schematic representation of the RIG-I wild type and nsSNP-containing mutants. RIG-I has a tandem CARD, RNA helicase domain, and CTD. Positions of the mutations are indicated by *asterisks*. *aa*, amino acids. *B*, FLAG-tagged WT RIG-I and SNPs were produced in RIG-I$^{-/-}$ MEFs and detected by immunoblotting using an anti-FLAG antibody.

loss of function mutations of RIG-I and MDA5. Notably, two MDA5 mutations, E627* and I923V, recently reported to have a strong association with resistance to T1D (13), were severely inactive. The results suggest a novel molecular mechanism for the activation of RLRs and will contribute to our understanding of the functional effects of RLR polymorphisms and the critical relationship between RLR nsSNPs and diseases.

EXPERIMENTAL PROCEDURES

Cells, DNA Transfection, and Preparation of Cell Extracts—L929 cells were maintained in minimum essential medium Eagle (Sigma) with 5% fetal bovine serum and penicillin/streptomycin. Mouse embryonic fibroblasts (MEFs) were obtained from Dr. S. Akira (Osaka University). MEFs and 293T cells were maintained in Dulbecco's modified Eagle's medium with 10% fetal bovine serum and penicillin/streptomycin. L929 cells were transiently transfected with the DEAE-dextran method. MEFs and 293T cells were transiently transfected by FuGENE 6 (Roche Applied Science). For the preparation of cell extracts, cells were lysed with lysis buffer (50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 1 mM EDTA, 1% Nonidet P-40, 0.1 mg/ml leupeptin, 1 mM phenylmethylsulfonyl fluoride, and 1 mM sodium orthovanadate) and centrifuged at 245,000 × *g* for 10 min. The supernatant was used for SDS-PAGE and electrophoretic mobility shift assays (EMSAs).

Oligonucleotides—Oligonucleotides (800 bp) corresponding to the coding sequence of green fluorescent protein were amplified by PCR using T7-primer (5′-CGTAATACGACTCACTATAGGGGATATCAGCAAAGGAGAAGAACTTTT-3′) and T3-primer (5′-GCAATTAACCCTCACTAAAGGGAGGCCTAGGGAGAAGACAGTGAGCTC-3′). Long dsRNA (ds800) was prepared by annealing complementary strands separately synthesized by *in vitro* transcription using the AmpliScribe T7 flash transcription kit (EPICENTRE Biotechnologies) and the AmpliScribe SP6 high yield transcription kit (EPICENTRE Biotechnologies). The annealed dsRNA was treated with S1 nuclease (Takara Bio) to generate a blunt end and alkaline phosphatase (Takara Bio) to remove 5′-phosphate. ^{32}P-ds800 was prepared by labeling the 800-bp dsRNA using T4 polynucleotide kinase and [γ-^{32}P]ATP.

Reporter Assay—MEFs were transfected with p55-C1Bluc, pRLtK, and expression plasmids for RIG-I mutants or MDA5 mutants. Cells were split into two aliquots, stimulated with RNA (5′-pppGG25) (6) or poly(I-C) transfection or Sendai virus (SeV) infection for 12 h, and harvested at 48 h after DNA transfection. L929 cells were transfected similar to MEFs but were stimulated by Newcastle disease virus infection. The infections were performed as described previously (14). The RNA transfection and poly(I-C) transfection were performed using Lipofectamine™ RNAiMAX

(Invitrogen). The luciferase assay was performed with a Dual-Luciferase reporter assay system (Promega). Luciferase activity was normalized using *Renilla* luciferase activity (pRLtk) as a reference.

Quantitative PCR Assay—Quantitative PCR was performed as described previously (3).

Plasmid Constructs—p-55C1Bluc and pRLtk, pEF-FLAG-RIG-I, pEF-FLAG-RIG-ICARD, pEF-FLAG-MDA5, and pEF-FLAG-MDA5CARD were described previously (14). Data on SNPs for *RIG-I* and *MDA5* were obtained from the NCBI (www.ncbi.nlm.nih.gov) and HapMap databases. The expression plasmids for RIG-I mutants (pEF-FLAG-RIG-IR7C, pEF-FLAG-RIG-IS144F, pEF-FLAG-RIG-IS183I, pEF-FLAG-RIG-IT260P, pEF-FLAG-RIG-II406T, pEF-FLAG-RIG-ID580E, pEF-FLAG-RIG-IF789L, and pEF-FLAG-RIG-ICARDS183I) and MDA5 mutants (pEF-FLAG-MDA5T260S, pEF-FLAG-MDA5L274I, pEF-FLAG-MDA5K351E, pEF-FLAG-MDA5I442V, pEF-FLAG-MDA5H460R, pEF-FLAG-MDA5E627*, pEF-FLAG-MDA5H843R, pEF-FLAG-MDA5I923V, pEF-FLAG-MDA5A946T, and pEF-FLAG-MDA5D1014E) were generated using a GeneEditor *in vitro* site-directed mutagenesis system (Promega). The mutations were confirmed by sequencing.

Antibodies and Immunoblotting—The anti-FLAG (M2; Sigma) antibody is a commercial product. SDS-PAGE and immunoblotting were performed as described previously (14).

EMSA—293T cells (1×10^6/6-cm dish) were transfected with 1 μg of expression vector. At 24 h after transfection, cell extract was prepared and mixed with anti-FLAG beads (Sigma) to adsorb FLAG-tagged proteins. The beads were washed, and bound protein was eluted with FLAG peptide (Sigma). The method of EMSA was described previously (6).

RESULTS

Construction of RIG-I Mutants and Their Biological Activities in MEFs Derived from RIG-I Knock-out Mice—nsSNPs were selected from nucleotide sequence polymorphisms of human RIG-I reported in databases and introduced into the RIG-I expression vector by site-directed mutagenesis. Domain structure and locations of the mutations are indicated in Fig. 1*A*. In RIG-I$^{-/-}$ MEFs transiently transfected with these vectors, the wild type and seven human RIG-I mutants were expressed at comparable levels (Fig. 1*B*). These results indicate that none of the amino acid substitutions significantly affect the synthesis and/or stability of RIG-I. The signaling function of the mutants was analyzed by temporarily complementing the function of RIG-I in RIG-I$^{-/-}$ MEFs using a virus-responsive luciferase reporter gene (Fig. 2*A*). Although the transfection of 5′-pppRNA did not activate the reporter gene significantly in RIG-I$^{-/-}$ MEFs (Fig. 2*A*, *Vector*), the ectopic expression of WT RIG-I restored the responsiveness to the ligand. All the mutant constructs exhibited functional complementation except the R7C, S144F, and S183I mutants, which exhibited a reduced response to 5′-pppRNA, particularly the S183I mutant, which exhibited a severe defect. Because the activity of S144F was not reproducible, we did not investigate this mutant further. It has been shown that RIG-I senses SeV and activates the IFN promoter. Therefore, the ability of the mutants to respond to this viral inducer was tested (Fig. 2*B*). SeV efficiently activated the reporter

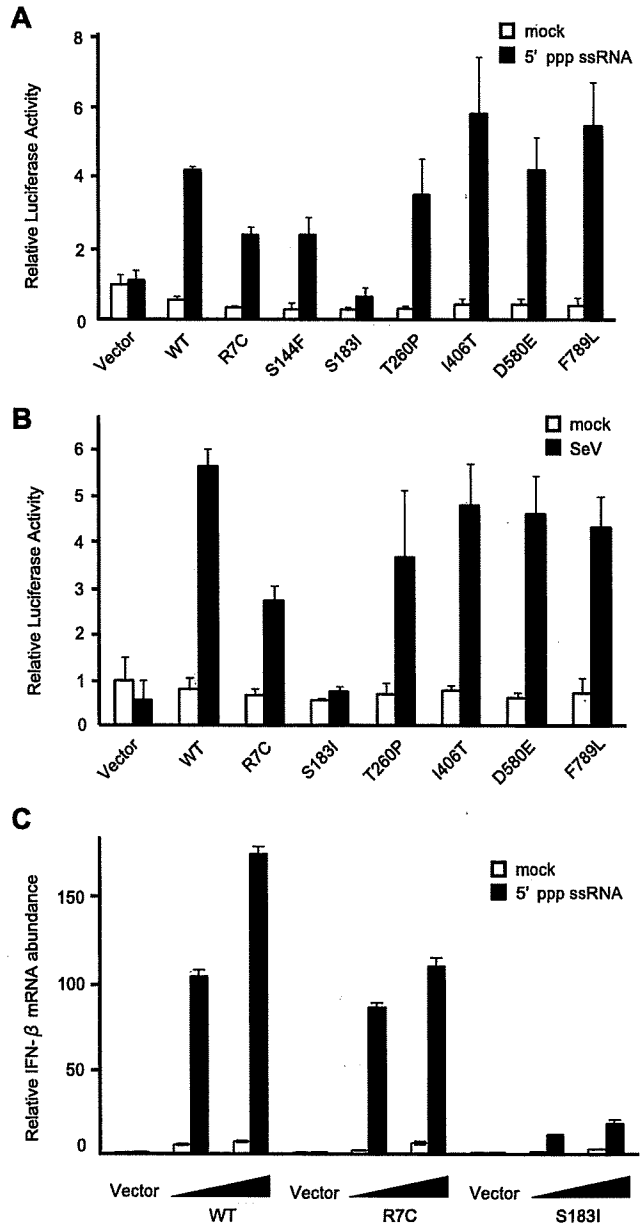


FIGURE 2. **Functional analysis of RIG-I nsSNP mutants.** *A* and *B*, RIG-I$^{-/-}$ MEF cells were transiently transfected with p-55CIBluc together with empty vector (*Vector*) or the indicated constructs. The cells were subjected to a Dual-Luciferase assay after stimulation with 5′-ppp-ssRNA (12 h) (*A*) or SeV (12 h) (*B*). The relative firefly luciferase activity, normalized to the *Renilla* luciferase activity, is shown. *Error bars* show the SDs for triplicate transfections. *mock*, mock-treated. *C*, RIG-I$^{-/-}$ MEFs were transfected with empty vector (*Vector*) or expression vectors for WT RIG-I or mutants as indicated (the total amount of plasmid was kept at 6 μg by adding empty vector). To observe the dose response, cells were transfected with 3 or 6 μg of the expression plasmid. Cells were mock-treated or transfected with 5′-ppp-ssRNA for 12 h, and *IFN-β* mRNA was quantified by quantitative PCR by using the Applied Biosystems primer set for mouse interferon-β1: Mm00439546_S1.

gene when WT RIG-I was expressed; however, S183I was virtually inactive and R7C exhibited partial activity. Other mutants sensed the virus as efficiently as the WT. We further confirmed the effect of the R7C and S183I mutations by monitoring the expression of endogenous mouse *IFN-β* mRNA (Fig. 2*C*). The results clearly demonstrate that S183I is barely active and that R7C is partially

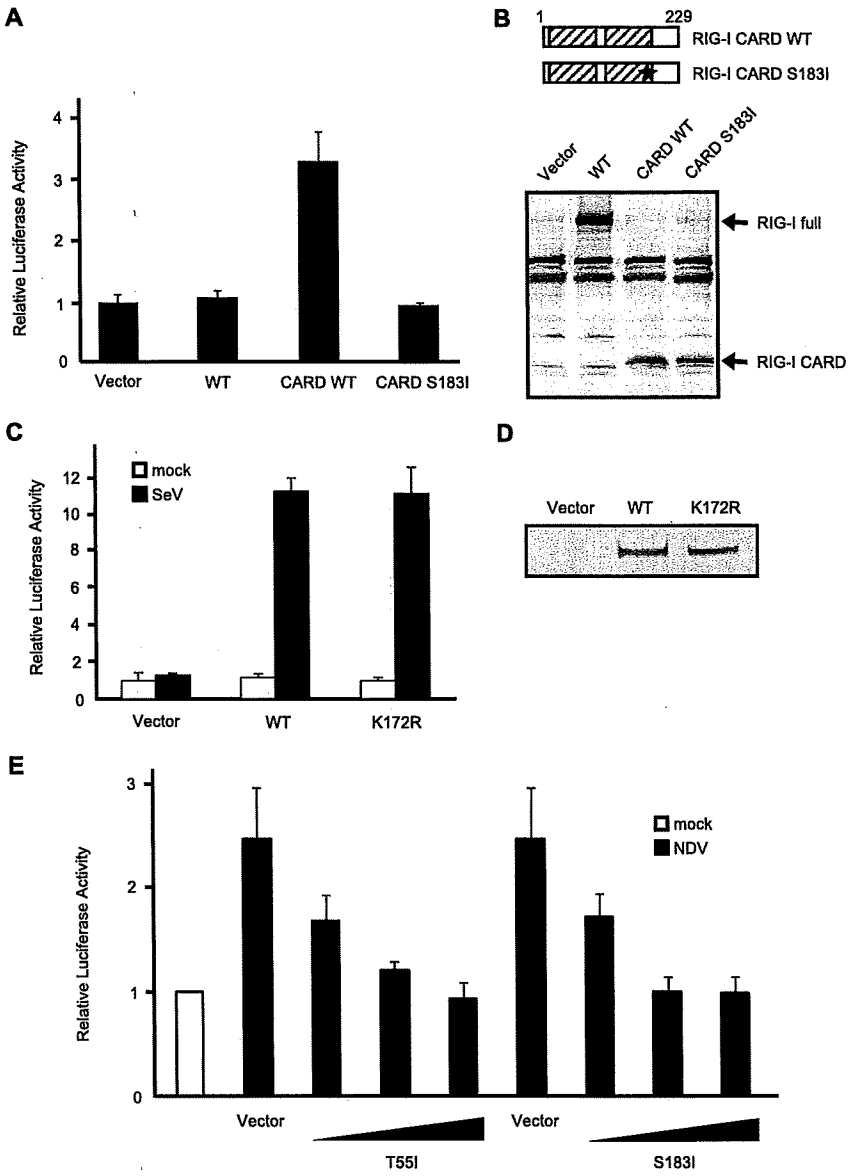


FIGURE 3. **Characterization of RIG-I S183I.** *A*, RIG-I$^{-/-}$ MEF cells were transfected with reporter genes together with empty vector (*Vector*) or plasmid expressing FLAG-tagged WT RIG-I, RIG-I CARD (the N-terminal region, amino acid 1–229), or RIG-I CARD S183I. After transfection (24 h), the cells were subjected to a Dual-Luciferase assay. *Error bars* show the S.D. values for triplicate transfections. *B*, each protein expressed in RIG-I$^{-/-}$ MEF cells was detected by immunoblotting using an anti-FLAG antibody. *RIG-I full*, full-length RIG-I. *C*, reporter assay of the K172R mutant was performed as in Fig. 2*B*. *D*, protein levels were determined by immunoblotting. RIG-I K172R and WT RIG-I were expressed at comparable levels. *E*, empty vector or expression vectors for full-length RIG-I with the T55I or S183I mutation were introduced into L929 cells (the total amount of plasmid was kept at 9 μg by adding empty vector) and infected with Newcastle disease virus, and reporter activity was analyzed as in *panel A*. To observe the dose response, cells received 1, 5, or 9 μg of the expression plasmid for T55I and S183I as indicated. *NDV*, Newcastle disease virus.

active. The above results strongly suggest that Ser-183 is critical for RIG-I to sense transfected 5′-pppRNA as well as SeV-derived PAMPs. Ser-183 resides within the second CARD, and this prompted us to explore the impact of the S183I substitution on the signaling function of the isolated RIG-I CARD. Unlike that of the full-length RIG-I, overexpression of the truncated RIG-I (1–229), which encompasses the two repeats of CARD, constitutively activated the reporter p-55C1B without a viral stimulus (Fig. 3*A*). However, RIG-I (1–229) with S183I failed to activate the reporter gene. In these cells, levels of RIG-I (1–229) with or without the S183I mutation were comparable, suggesting that Ser-183 is critical for the signaling function but does not affect protein levels of RIG-I CARD (Fig. 3*B*). It was reported that human RIG-I undergoes ubiquitination at Lys-172 in the second CARD, and this process is essential for RIG-I signaling (15). To compare the mutation at Ser-183, we generated a K172R mutant and tested its activity (Fig. 3, *C* and *D*). Surprisingly, the mutant exhibited virus-induced signaling activity comparable with the WT, suggesting that the ubiquitination of Lys-172 plays a minor role in the regulation of RIG-I and that the phenotype of the S183I mutant is unlikely due to a failure of ubiquitination.

It is known that human RIG-I with the amino acid substitution T55I acts as a dominant inhibitor (7). This mutation within the first CARD inactivates the signaling function of the isolated tandem CARDs. We next tested whether RIG-I S183I exhibits a dominant negative phenotype. L929 cells were transfected with control vector or RIG-I mutants and then activated by infection of Newcastle disease virus. Cells transfected with the vector exhibited *IFN* promoter activity due to endogenous RIG-I (Fig. 3*E*). Expression of T55I as well as S183I significantly reduced the promoter activity in a dose-dependent manner, suggesting that S183I causes a dominant negative phenotype similar to T55I.

Construction of MDA5 Mutants and Their Biological Activities in MEFs Derived from MDA5 Knockout Mouse—Next, the biological activity of human MDA5 mutants was analyzed similarly using *MDA5*-deficient MEFs. Ten nsSNPs including A946T identified in familial T1D (16) were introduced into the MDA5 expression vector (Fig. 4*A*). The wild type and mutants were expressed in MDA5$^{-/-}$ MEFs at comparable levels (Fig. 4*B*), showing that the mutations, including the E627* truncation, did not affect MDA5 protein levels. The biological activity of the mutants was assayed similarly to that of the RIG-I mutants. As a MDA5 agonist, a commercial poly(I-C) with an average length of 2 kbp was used, which selectively activates MDA5 (6, 11). Wild-type MDA5 clearly conferred responsiveness to the poly(I-C) (Fig. 5). Eight MDA5 mutants, including A946T, which was implicated in human T1D (16), exhibited complementing activity comparable

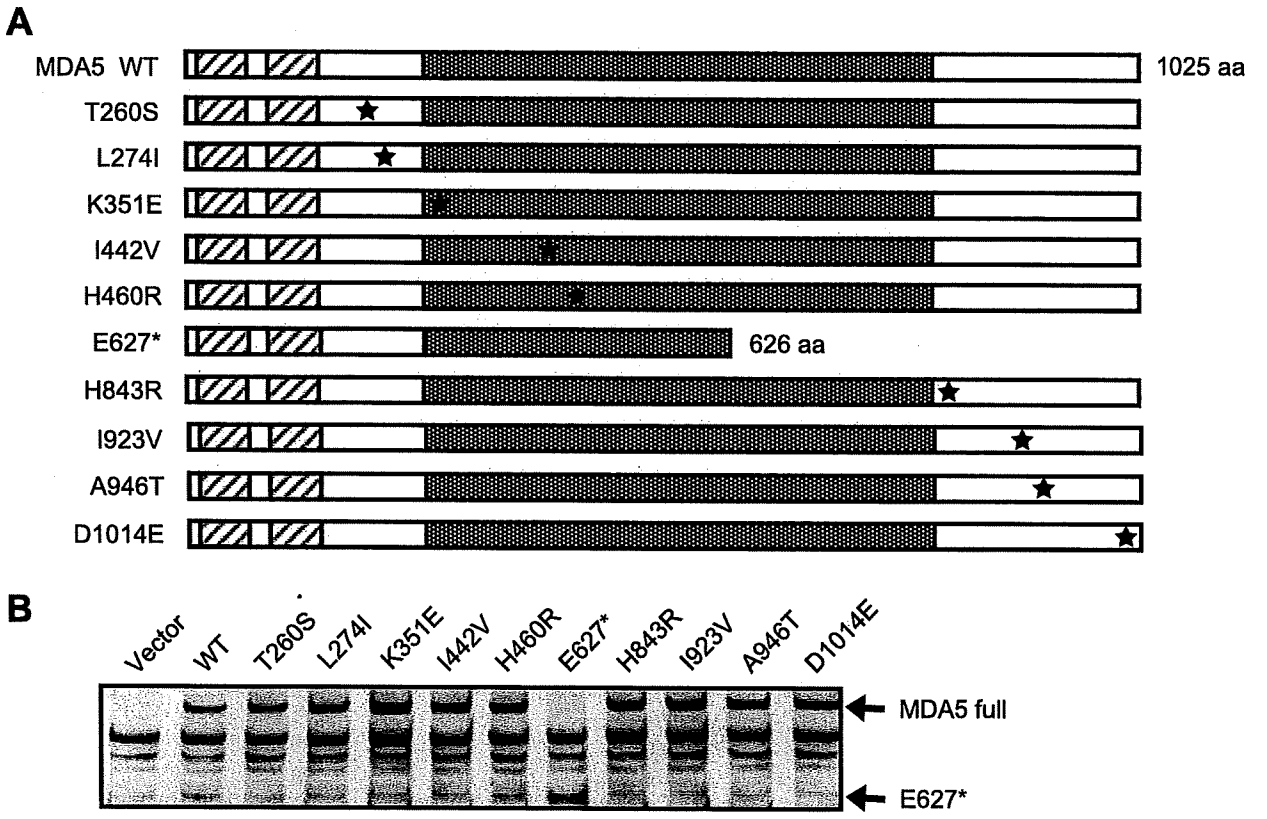


FIGURE 4. **MDA5 nsSNP mutants and their expression in MEFs.** *A*, schematic representation of WT MDA5 and its nsSNPs. Point mutations are indicated by *asterisks*. E627* is the nonsense mutant. *aa*, amino acids. *B*, FLAG-tagged MDA5 SNPs were produced in MDA5$^{-/-}$ MEFs and detected by immunoblotting using an anti-FLAG antibody. *Vector*, empty vector; *MDA5 full*, full-length MDA5.

with the wild type. However, E627* and I923V showed significantly low levels of activity. The phenotypes of these mutations were further tested by transient expression in MDA5$^{-/-}$ MEFs and monitoring endogenous *IFN-β* mRNA (Fig. 5*B*). Although MDA5 is absent, intracellular poly(I-C) induced *IFN-β* gene expression in the control cells (*Vector*). This is presumably due to the activation of RIG-I by a short poly(I-C) present in the preparation we used. Unlike the reporter assay, which monitors only transfected cells, this quantitative PCR assay detects *IFN-β* transcripts from both transfected and non-transfected cells, increasing the background signal. Irrespective of the background, overexpression of WT MDA5 resulted in an enhancement of *IFN-β* expression by poly(I-C), and this induction was not observed with E627* and I923V. Further, the effect was observed at different levels of MDA5 expression. The E627* mutant lacks a part of the helicase domain and the entire CTD in which Ile-923 resides.

Because in the case of RIG-I, the CTD determines RNA recognition specificity, we investigated the RNA binding activity of these mutants by EMSA using ^{32}P-labeled poly(I-C). Wild-type and recombinant MDA5 proteins were expressed in 293T cells and purified. The recombinant proteins were virtually free of other cellular proteins as analyzed by Coomassie Brilliant Blue staining (Fig. 5*C*). Wild-type MDA5 clearly formed a complex with poly(I-C) (Fig. 5*D*), but E627* did not exhibit detectable binding activity. I923V and A946T exhibited activity to bind poly(I-C) as strongly as the wild type under these conditions. We further compared the RNA binding at different protein concentrations (Fig. 5*E*) and confirmed that MDA5 WT and the I923V mutant bind to dsRNA in a comparable fashion. These results suggest that the E627* mutant is biologically inactive due to its failure to recognize its agonist.

DISCUSSION

Human RIG-I Polymorphism—We identified S183I as a loss of function mutation of RIG-I. This serine residue is conserved in human, monkey, cow, and pig RIG-I. The mutation apparently inactivates the tandem CARD, which relays signals downstream. Furthermore, S183I exhibits a dominant inhibitory phenotype, suggesting that individuals retaining this mutation as a heterozygote would exhibit hypersensitivity to viral infections. It has been shown that RIG-I mutants with a CARD deletion (RIG-IC), CARD point mutation (T55I), and ATP-binding site mutation (RIG-I K270A) all function as a dominant inhibitor (3, 7). It has been reported that the repression domain present in the C-terminal region of RIG-I and LGP2 dominantly suppresses the activation of RIG-I in *trans* through interaction with CARD and the helicase loop region (7). Because the repression domain encompasses the RNA-binding domain of RIG-I (6), one mechanism is likely competition of RNA binding with WT RNA. However, the RNA binding-deficient RIG-I mutants, K888A/K907A and K858A/K861A (6), functioned as a dominant negative inhibitor.[3] Interestingly, LGP2 mutants, K643E and K651E, which correspond to Lys-888

[3] J.-P. Zheng and T. Fujita, unpublished observation.

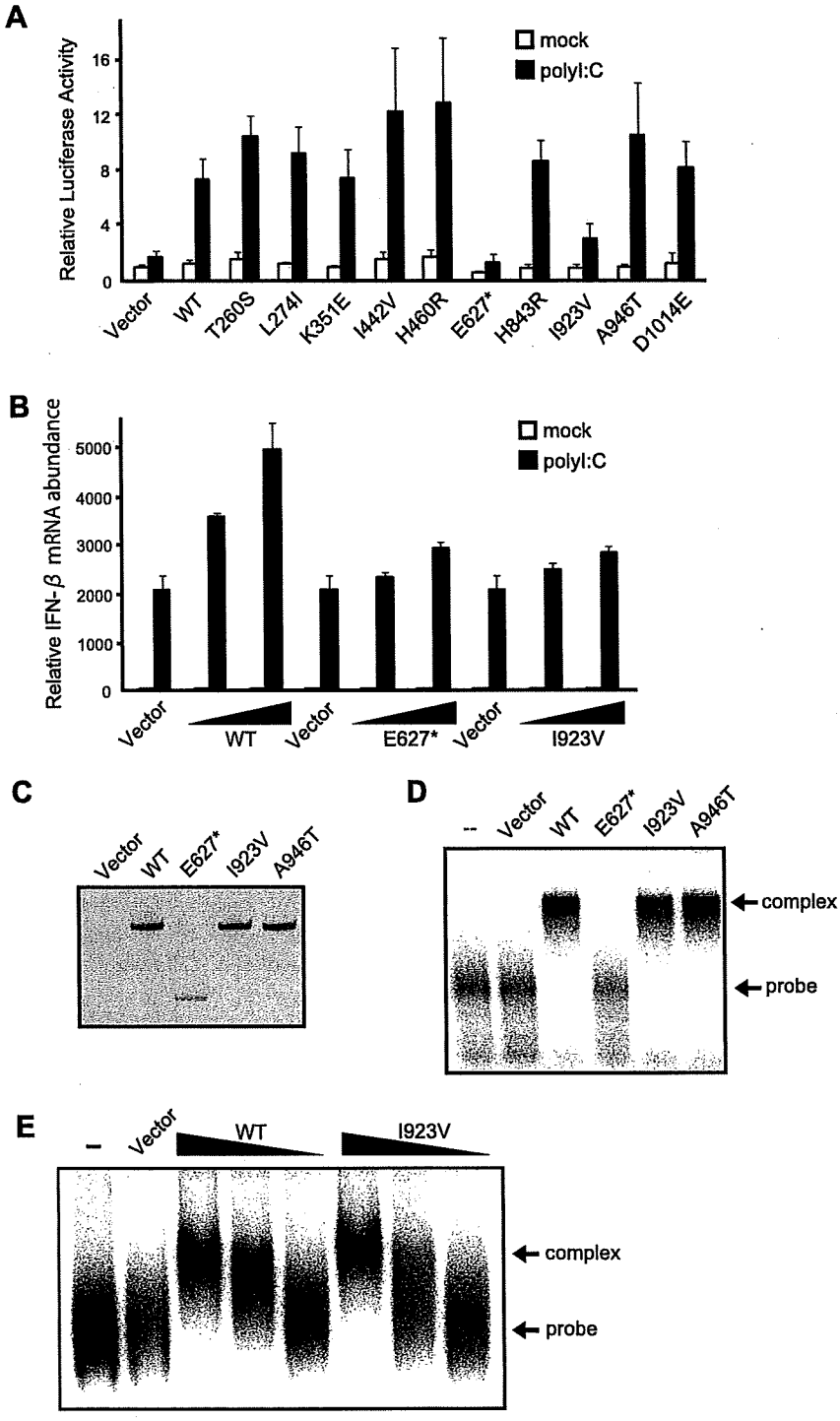


FIGURE 5. **Functional analysis of MDA5 nsSNP mutants.** *A*, MDA5$^{-/-}$ MEFs were transfected with reporter genes together with the indicated constructs as in Fig. 2A. After stimulation with poly(I-C) (12 h), the cells were subjected to a Dual-Luciferase assay. *Error bars* show the S.D. values for triplicate transfections. *mock*, mock-treated. *Vector*, empty vector. *B*, MDA5$^{-/-}$ MEFs were transfected with expression vectors for WT MDA5 or E627* or Ile-923 mutants (the total amount of plasmid was kept at 6 μg by adding empty vector). To observe the dose response, cells were transfected with 3 or 5.7 μg of the expression plasmid. Cells were mock-treated or transfected with poly(I-C) for 12 h, and *IFN-β* mRNA was quantified by quantitative PCR as in Fig. 2C. *C*, 293T cells were transfected with empty vector, WT MDA5, E627*, I923V, or A946T, and the produced proteins were purified using anti-FLAG ("Experimental Procedures"). The purified proteins were separated by SDS-PAGE and stained by Coomassie Brilliant Blue. *D*, EMSA of the purified MDA5 proteins (500 ng) using ^{32}P-labeled poly(I-C) as a probe. *complex*, probe protein complex; *probe*, free probe. *E*, dose response of RNA binding by WT MDA5 and the I923V mutant. EMSA was performed using 500, 300, and 100 ng of MDA5 and Ile-923 protein.

and Lys-907 of RIG-I, lost RNA binding activity but retain repression function (17). These results strongly suggest an RNA-independent mechanism.

One proposed function of the second CARD is to conjugate to ubiquitin (at Lys-172), which may be essential for signaling activity. However, human RIG-I with K172R, which is resistant to ubiquitination, did not affect phenotype (Fig. 3, *C* and *D*). Moreover, the corresponding amino acid in mouse RIG-I is glutamine, and the corresponding position of MDA5 is glutamic acid (human, monkey, mouse, cow, and pig), suggesting that the ubiquitination of Lys-172 has a minor impact on signaling activity. The precise function of the second CARD is not clear, but a comparative study of Ser-183 and WT RIG-I will elucidate the molecular function of this residue.

Human MDA5 Polymorphism—We identified two loss of function mutations, E627* and I923V, in human MDA5. Each of these mutations is actually present in the human population (13). It is worth investigating the phenotype of the homozygote because *MDA5* knockout mice clearly exhibit hypersusceptibility to the family *Picornaviridae*, genus Cardiovirus. E627*, which lacks CTD and a part of the helicase domain, lost its dsRNA binding activity; hence there was no signaling activity. Although the I923V mutation occurs in the CTD, this mutant exhibited intact dsRNA binding, suggesting a novel function of Ile-923 other than the recognition of RNA. It is worth noting that this isoleucine is conserved in human, monkey, mouse, cow, and pig MDA5. It is tempting to speculate that Ile-923 participates in an interaction with some other domain of MDA5 or other unknown regulatory protein(s).

A new report by Nejentsev *et al.* (13) describing the relationship between susceptibility to T1D and *MDA5* polymorphism was published. The report describes that four rare mutations (two mutations

in the exon and two mutations in the intron) in the human *MDA5* gene are associated with protection against T1D. Our analysis included the two exon mutations (E627* and I923V), which exhibited a loss of function phenotype. The genetic analysis included another rare nsSNP, H460R, shown to be independent of T1D resistance. The phenotype of this mutant was normal in our analysis. On the other hand, A946T, which was suggested to associate with T1D in previous reports (13, 16), did not exhibit a loss of function phenotype. However, because Nejentsev *et al.* (13) suggested that the association of A946R with T1D was due to the effect of another nsSNP, R843H, a combination of these mutations might confer loss of function on MDA5. In summary, our analysis strongly suggests that loss of function mutations of MDA5 have a causative role in resistance to T1D. Although T1D has a complex pathology, these findings may provide a new strategy for establishing an animal model for T1D.

The human genome encodes multiple sensors, including TLRs and RLRs for viral PAMPs. Some of these may act redundantly to secure defense against viral infections. In the case of RLRs, RIG-I and MDA5 detect a distinct spectrum of viruses, as suggested from the phenotype of respective knock-out mice. There is variation within human populations in susceptibility to a particular viral infection. We argue for a possible contribution of the genetic diversity of RLRs, including those identified in the current investigation, to susceptibility. In addition to the impact of the infection itself, secondary effects such as autoimmunity, which is remotely triggered by certain viral infections, may be influenced at least in part through the functional diversity of RLRs. In summary, our results suggest a critical relationship between RLR polymorphisms and diseases including viral infections and autoimmunity.

Acknowledgments—We greatly appreciate the gift of MEFs from Dr. S. Akira (Research Institute for Microbial Diseases, Osaka University, Osaka). We give our thanks to Dr. F. Matsuda (Center for Genomic Medicine, Kyoto University, Kyoto) for discussion.

REFERENCES

1. Samuel, C. E. (2001) *Clin. Microbiol. Rev.* **14**, 778–809
2. Akira, S., Uematsu, S., and Takeuchi, O. (2006) *Cell* **124**, 783–801
3. Yoneyama, M., Kikuchi, M., Natsukawa, T., Shinobu, N., Imaizumi, T., Miyagishi, M., Taira, K., Akira, S., and Fujita, T. (2004) *Nat. Immun.* **5**, 730–737
4. Yoneyama, M., and Fujita, T. (2008) *Immunity* **29**, 178–181
5. Cui, S., Eisenacher, K., Kirchhofer, A., Brzozka, K., Lammens, A., Lammens, K., Fujita, T., Conzelmann, K. K., Krug, A., and Hopfner, K. P. (2008) *Mol. Cell* **29**, 169–179
6. Takahasi, K., Yoneyama, M., Nishihori, T., Hirai, R., Kumeta, H., Narita, R., Gale, M., Jr., Inagaki, F., and Fujita, T. (2008) *Mol. Cell* **29**, 428–440
7. Saito, T., Hirai, R., Loo, Y. M., Owen, D., Johnson, C. L., Sinha, S. C., Akira, S., Fujita, T., and Gale, M., Jr. (2007) *Proc. Natl. Acad. Sci. U. S. A.* **104**, 582–587
8. Gitlin, L., Barchet, W., Gilfillan, S., Cella, M., Beutler, B., Flavell, R. A., Diamond, M. S., and Colonna, M. (2006) *Proc. Natl. Acad. Sci. U. S. A.* **103**, 8459–8464
9. Kato, H., Takeuchi, O., Sato, S., Yoneyama, M., Yamamoto, M., Matsui, K., Uematsu, S., Jung, A., Kawai, T., Ishii, K. J., Yamaguchi, O., Otsu, K., Tsujimura, T., Koh, C. S., Reis e Sousa, C., Matsuura, Y., Fujita, T., and Akira, S. (2006) *Nature* **441**, 101–105
10. Hornung, V., Ellegast, J., Kim, S., Brzozka, K., Jung, A., Kato, H., Poeck, H., Akira, S., Conzelmann, K. K., Schlee, M., Endres, S., and Hartmann, G. (2006) *Science* **314**, 994–997
11. Kato, H., Takeuchi, O., Mikamo-Satoh, E., Hirai, R., Kawai, T., Matsushita, K., Hiiragi, A., Dermody, T. S., Fujita, T., and Akira, S. (2008) *J. Exp. Med.* **205**, 1601–1610
12. Pichlmair, A., Schulz, O., Tan, C. P., Naslund, T. I., Liljestrom, P., Weber, F., and Reis e Sousa, C. (2006) *Science* **314**, 997–1001
13. Nejentsev, S., Walker, N., Riches, D., Egholm, M., and Todd, J. A. (March 5, 2009) *Science* 10.1126/science.1167728
14. Yoneyama, M., Kikuchi, M., Matsumoto, K., Imaizumi, T., Miyagishi, M., Taira, K., Foy, E., Loo, Y. M., Gale, M., Jr., Akira, S., Yonehara, S., Kato, A., and Fujita, T. (2005) *J. Immunol.* **175**, 2851–2858
15. Gack, M. U., Shin, Y. C., Joo, C. H., Urano, T., Liang, C., Sun, L., Takeuchi, O., Akira, S., Chen, Z., Inoue, S., and Jung, J. U. (2007) *Nature* **446**, 916–920
16. Smyth, D. J., Cooper, J. D., Bailey, R., Field, S., Burren, O., Smink, L. J., Guja, C., Ionescu-Tirgoviste, C., Widmer, B., Dunger, D. B., Savage, D. A., Walker, N. M., Clayton, D. G., and Todd, J. A. (2006) *Nat. Genet.* **38**, 617–619
17. Li, X., Ranjith-Kumar, C. T., Brooks, M. T., Dharmaiah, S., Herr, A. B., Kao, C., and Li, P. (March 28, 2009) *J. Biol. Chem.* 10.1074/jbc.M900818200

THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 284, NO. 26, pp. 17465–17474, June 26, 2009

Solution Structures of Cytosolic RNA Sensor MDA5 and LGP2 C-terminal Domains

IDENTIFICATION OF THE RNA RECOGNITION LOOP IN RIG-I-LIKE RECEPTORS[*S]

Received for publication, January 12, 2009, and in revised form, April 11, 2009 Published, JBC Papers in Press, April 20, 2009, DOI 10.1074/jbc.M109.007179

Kiyohiro Takahasi[‡], Hiroyuki Kumeta[‡], Natsuko Tsuduki[‡], Ryo Narita[§¶], Taeko Shigemoto[§¶], Reiko Hirai[§¶], Mitsutoshi Yoneyama[§¶,‖], Masataka Horiuchi[‡], Kenji Ogura[‡], Takashi Fujita[§¶], and Fuyuhiko Inagaki[‡1]

From the [‡]Department of Structural Biology, Graduate School of Pharmaceutical Sciences, Hokkaido University, N-21, W-11, Kita-ku, Sapporo 001-0021, the [§]Laboratory of Molecular Genetics, Institute for Virus Research, Kyoto University, Kyoto 606-8507, the [¶]Laboratory of Molecular Cell Biology, Graduate School of Biostudies, Kyoto University, Kyoto 606-8507, and [‖]PRESTO, Japan Science and Technology Agency, 4-1-8 Honcho Kawaguchi, Saitama 332-0012, Japan

The RIG-I like receptor (RLR) comprises three homologues: RIG-I (retinoic acid-inducible gene *I*), MDA5 (melanoma differentiation-associated gene *5*), and LGP2 (laboratory of genetics and physiology *2*). Each RLR senses different viral infections by recognizing replicating viral RNA in the cytoplasm. The RLR contains a conserved C-terminal domain (CTD), which is responsible for the binding specificity to the viral RNAs, including double-stranded RNA (dsRNA) and 5′-triphosphated single-stranded RNA (5′ppp-ssRNA). Here, the solution structures of the MDA5 and LGP2 CTD domains were solved by NMR and compared with those of RIG-I CTD. The CTD domains each have a similar fold and a similar basic surface but there is the distinct structural feature of a RNA binding loop; The LGP2 and RIG-I CTD domains have a large basic surface, one bank of which is formed by the RNA binding loop. MDA5 also has a large basic surface that is extensively flat due to open conformation of the RNA binding loop. The NMR chemical shift perturbation study showed that dsRNA and 5′ppp-ssRNA are bound to the basic surface of LGP2 CTD, whereas dsRNA is bound to the basic surface of MDA5 CTD but much more weakly, indicating that the conformation of the RNA binding loop is responsible for the sensitivity to dsRNA and 5′ppp-ssRNA. Mutation study of the basic surface and the RNA binding loop supports the conclusion from the structure studies. Thus, the CTD is responsible for the binding affinity to the viral RNAs.

A variety of pathogen-associated molecular patterns, including microbial peptidoglycan, lipopolysaccharide, β-1,3-glucan, and viral DNA or RNA are recognized by pattern recognition receptors that evoke the innate immune responses of host cells. In viral infections, double-stranded RNA (dsRNA)[2] is recognized by Toll-like receptor-3 in the early endosome and by RIG-I like receptors (RLRs) in the cytoplasm. These two receptors initiate the innate immune responses including the production of cytokines and type-I interferon, which are critical for the subsequent adaptive immune response (1).

The RLR comprises three homologs: RIG-I (retinoic acid-inducible gene *I*), MDA5 (melanoma differentiation-associated gene *5*), and LGP2 (laboratory of genetics and physiology *2*) (see Fig. 1*A*) (2), and they sense a viral infection by recognizing replicating viral RNA in the cytoplasm. The RIG-I and MDA5 consist of three functional domains: tandem-CARDs (caspase activation and recruitment domain), a DEAD box helicase-like domain, and a well conserved C-terminal domain (CTD), whereas LGP2 has only the DEAD box helicase like domain and well conserved CTD. The three RLRs are considered to play different roles in the recognition of pathogen-associated molecular patterns and to be activated by different viruses and different viral RNAs. RIG-I is activated by a variety of viruses, including paramyxovirus, rhabdovirus, and orthomyxovirus, recognizing not only dsRNA but also 5′-triphosphated single-stranded RNA (5′ppp-ssRNA) (3, 4), and MDA5 is mainly activated by picornavirus (5, 6), whereas LGP2 lacking the tandem CARDs was originally identified as a negative regulator, but gene disruption study showed that it may function as a positive regulator (2, 7).

Our previous study has shown that the three domains of RIG-I cooperatively take part in pathogen-associated molecular pattern recognition and signal transduction (8). Further, the tandem CARDs are essential to transduce the signal via CARD-CARD interaction with the downstream CARD containing signal element, IPS-1/MAVS/VISA/Cardif (9–12). The CTD plays a critical role in the specific recognition of dsRNA and

* This work was supported by grants-in-aid (Tokutei-ryoiki (Matrix of Infection Phenomena), Kiban (S)) and National Projects on Targeted Proteins Research Program from the Japanese Ministry of Education, Culture, Sports, Science, and Technology, by a Research Grant of the Uehara Memorial Foundation, and by Nippon Boehringer Ingelheim Co., Ltd.

The atomic coordinates and structure factors (codes 2RQB and 2RQA) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

[S] The on-line version of this article (available at http://www.jbc.org) contains supplemental Figs. 1–4 and supplemental Table 1.

[1] To whom correspondence should be addressed: Dept. of Structural Biology, Graduate School of Pharmaceutical Sciences, Hokkaido University, N-21, W-11, Kita-ku, Sapporo 001-0021, Japan. Tel.: 81-11-706-9011; Fax: 81-11-706-9012; E-mail: finagaki@pharm.hokudai.ac.jp.

[2] The abbreviations used are: dsRNA, double-stranded RNA; ssRNA, single-stranded RNA; RLR, RIG-I-like receptor; CARD, caspase recruitment domain; IPS-1, interferon promoter stimulator-1; RIG-I, retinoic acid-inducible gene *I*; MDA5, melanoma differentiation-associated gene *5*; LGP2, laboratory of genetics and physiology 2; 5′ppp-ssRNA, 5′-triphosphated single-stranded RNA; GST, glutathione *S*-transferase; CTD, C-terminal domain; PBS, phosphate-buffered saline; Bis-Tris, 2-[bis(2-hydroxyethyl)amino]-2-(hydroxymethyl)propane-1,3-diol; DTT, dithiothreitol; NOESY, nuclear Overhauser effect spectroscopy; wt, wild type; EMSA, electrophoretic mobility shift assay; SPR, surface plasmon resonance.

5′ppp-ssRNA, while the DEAD box helicase cooperatively enhances the affinity to dsRNA through a conformational change. In the resting state, RIG-I is believed to be in closed conformation via interaction with these three domains (13). Upon viral infection, CTD specifically recognizes viral RNA and induces a domain rearrangement that allows RIG-I to form a stable complex with viral RNA and simultaneously exposes the tandem CARDs so that RIG-I is assumed to become an open conformation to interact with IPS-1/MAVS/VISA/Cardif (8, 14).

In the previous report, we have solved the solution structure of RIG-I CTD and shown the presence of a large positively charged surface that is responsible for viral RNA binding. In this report, we compare RNA-binding activity of the three CTDs (RIG-I, MDA5, and LGP2) and determine the solution structures of MDA5 and LGP2 CTDs. The RNA-binding activity is correlated to a conserved RNA binding loop, including a critical phenylalanine residue of RIG-I and LGP2. This residue is not conserved in MDA5, reflecting the distinct specificity of this sensor. Our results provide new insights into the recognition of viral RNA in cytoplasm.

EXPERIMENTAL PROCEDURES

Recombinant Proteins—The C-terminal domains of human MDA5 (896–1025) and LGP2 (546–678) were amplified by PCR and inserted into a pGEX-6P-1 vector to produce GST fusion protein (Amersham Biosciences). Expression vectors were introduced into *Escherichia coli* BL21(DE3) cells and cultured in M9 medium. The MDA5 CTD expression was induced by addition of 0.01 mM isopropyl 1-thio-β-D-galactopyranoside when the absorbance at 600 nm was ~0.4. The cells were then grown at 16 °C for 20 h; LGP2 CTD expression was induced by 1 mM isopropyl 1-thio-β-D-galactopyranoside then incubated at 25 °C for 20 h. The cells were suspended in PBS containing 1 mM 4-(2-aminoethyl)benzenesulfonylfluoride hydrochloride, lysed via sonication, and centrifuged. The supernatant was mixed with glutathione-Sepharose 4B (Amersham Biosciences) for 12 h, and the protein bound to glutathione-Sepharose 4B was washed first with PBS then with PBS containing 500 mM NaCl to remove the nucleic acid bound to the protein. The proteins were eluted with 50 mM Tris, pH 8, 150 mM NaCl, 20 mM reduced glutathione. The eluted protein was digested by PreScission protease (Amersham Biosciences). The proteins were further purified with size-exclusion Superdex 75 columns (Amersham Biosciences) with PBS containing 500 mM NaCl.

Site-directed Mutagenesis and Preparation of Mutant Proteins Using 293T Cells—Mutations of the RLR were introduced by the overlap extension PCR (15). The mutagenized cDNA, which was designed to possess an N-terminal FLAG tag, was subcloned into pEFBOS. The expression vectors were transiently transfected to 293T cells and purified by anti FLAG beads (8). The mutations did not alter the stability of the RLR proteins drastically as the production levels, and the recoveries from the transfected mammalian cells are comparable (Figs. 5–7).

Preparation of Uniformly Labeled Proteins—Uniformly Labeled MDA5 and LGP2 CTDs were expressed as above except here using medium containing ^{15}N-ammonium chloride and D-glucose (or [D-^{13}C]glucose), and purified similarly. The NMR samples for the structure determination were prepared at 0.94 mM MDA5 CTD in 20 mM Bis-Tris, pH 7.0, 250 mM NaCl, 10 mM DTT, and 1.00 mM LGP2 CTD in 50 mM Tris, pH 7, 250 mM NaCl, and 1 mM DTT.

TABLE 1
Structural statistics for 20 MDA5 CTD

NOE distance constraints	
Total	2970
Short range |i-j|< = 1	1662
Medium range 1<|i-j|<5	372
Long range |i-j|> = 5	982
Dihedral angle constraints	
ϕ	76
ψ	73
Residual NOE violations	
Number > 0.3 Å	0
Residual angle violations	
Number > 5.0°	0
Ramachandran statistics (%)	
Residues in most favored region	74.0
Residues in additionally allowed region	24.1
Residues in generally allowed region	1.9
Residues in disallowed region	0
Structural coordinates root mean square deviation (Å) (residues 897–1017)	
Backbone atoms for the final ensemble	0.40
All heavy atoms for the final ensemble	0.76

TABLE 2
Structural statistics for 20 LGP2 CTD

NOE distance constraints	
Total	2593
Short range |i-j|< = 1	1517
Medium range 1<|i-j|<5	322
Long range |i-j|> = 5	754
Dihedral angle constraints	
ϕ	58
ψ	62
Residual NOE violations	
Number > 0.3 Å	2
Residual angle violations	
Number > 5.0°	0
Ramachandran statistics (%)	
Residues in most favored region	71.0
Residues in additionally allowed region	27.4
Residues in generally allowed region	1.7
Residues in disallowed region	0
Structural coordinates root mean square deviation (Å) (residues 549–657)	
Backbone atoms for the final ensemble	0.49
All heavy atoms for the final ensemble	0.84

NMR Measurements—The NMR spectra were acquired at 25 °C on Varian Unity Inova 500, 600, and 800 spectrometers. The data were processed using NMRPipe (16) and analyzed using Sparky (T. D. Goddard and D. G. Kneller, SPARKY 3, University of California, San Francisco). For assignments of the protein backbone and side-chain ^{1}H, ^{15}N, and ^{13}C resonances, two- and three-dimensional (2D and 3D) spectra were obtained. Backbone resonances were assigned using the 2D ^{1}H-^{15}N HSQC, 3D HNCA, HN(CO)CA, HNCACB, CBCA(CO)NH, HNCO, (HCA)CO(CA)NH, HNCAHA, and HBHA(CO)NH spectra. The aliphatic atoms of the side chains were assigned using 2D ^{1}H-^{13}C HSQC, 3D C(CO)NH, H(CCO)NH, CCH-TOCSY, and HCCH-TOCSY spectra, while aromatic side-chain atoms were assigned using the 2D ^{1}H-^{13}C HSQC, (Hb)Cb(CgCd)Hd, (Hb)Cb(CgCdCe)He, and 3D HCCH-TOCSY spectra.

Structural Determination—The 3D ^{15}N-edited NOESY and ^{13}C-edited NOESY spectra (t_{mix} = 80 ms) were measured to

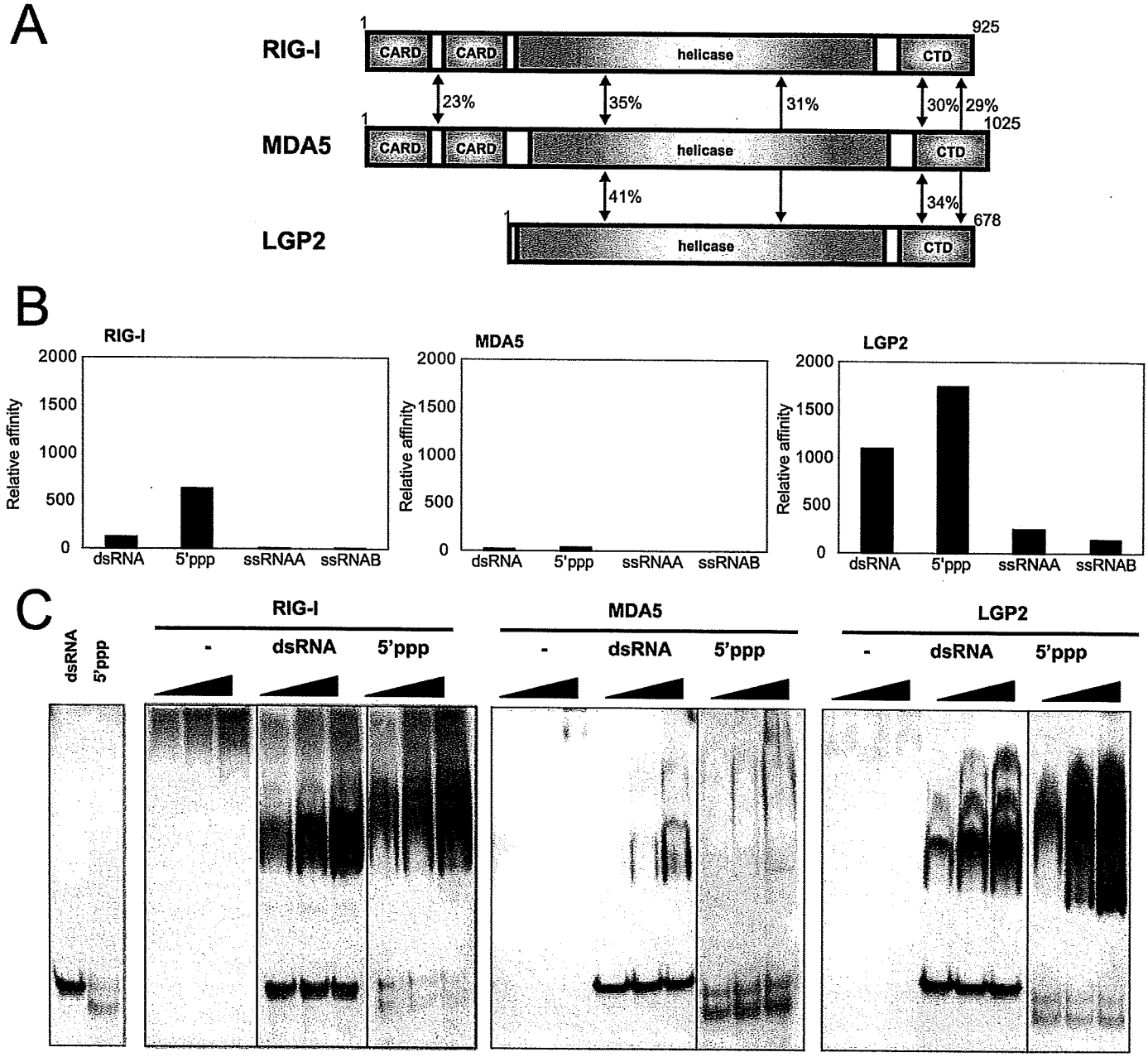


FIGURE 1. **Functional analysis of RLR CTDs.** *A*, domain structures of RIG-I, MDA5, and LGP2. Sequence identities of each domain among the RLRs are indicated. *B*, RNA-binding activity of RLR CTDs determined by SPR. GST-RIG-I CTD, GST-MDA5 CTD, and GST-LGP2 CTD were captured by the anti-GST anti-body immobilized onto the sensor chip, then dsRNA, 5′ppp-ssRNA, and two ssRNAs (*ssRNAA* and *ssRNAB*) were injected. Each resonance unit of RNA bound to a GST fused protein are standardized by molecular weight of the RNAs, then normalized by the resonance unit and molecular weight of the captured GST-fused protein. Normalized data are summarized and shown as *bar graphs*. RNAs bound to RIG-I, MDA5, and LGP2 are indicated from *left* to *right* in the *panels*. *C*, RNA-binding activity of RLR CTDs determined by EMSA. CTDs without the GST tag were prepared and subjected to EMSA. Increasing amounts of CTD (10, 20, and 40 pm) were reacted with the indicated probe and analyzed by native PAGE. The gels were silver-stained to visualize protein and RNA probe. −, no RNA; *dsRNA*, 25/25c probe; *5′ppp*: 5′pppGG25 probe.

obtain NOE distance constraints. Backbone ϕ and ψ dihedral angle constraints were generated using the TALOS program (17). The structures of MDA5 CTD and LGP2 CTD were determined using CANDID/CYANA 2.1 (18, 19). Structural statistics for the best 20 structures of MDA5 CTD and LGP2 CTD are shown in Tables 1 and 2. The 20 lowest energy structures of both MDA5 CTD and LGP2 CTD were deposited at the Protein Data Bank (PDB codes 2RQB and 2RQA, respectively).

NMR Titration Study of LGP2—The chemical shift perturbation study of the amide nitrogen and proton signals of LGP2 CTD in ^{1}H-^{15}N HSQC spectra was performed upon addition of the following RNAs: GG25/2 + 25c (dsRNA), 5′-triphosphated GG25 (5′ppp-ssRNA), two single-stranded RNAs GG25 (ssRNAA), and 2 + 25C (ssRNAB) were also titrated as controls. The GG25/2 + 25c was prepared by mixing equal molar GG25 and 2 + 25c, and then annealed in 1×M buffer (10 mM Tris, pH 7.5, 10 mM $MgCl_2$, 1 mM DTT). The final concentration of dsRNA was 0.71 mM. To equalize the condition of the titration experiments, GG25, 2 + 25C, and 5′-triphosphated GG25 was prepared at 0.71 mM in 1×M

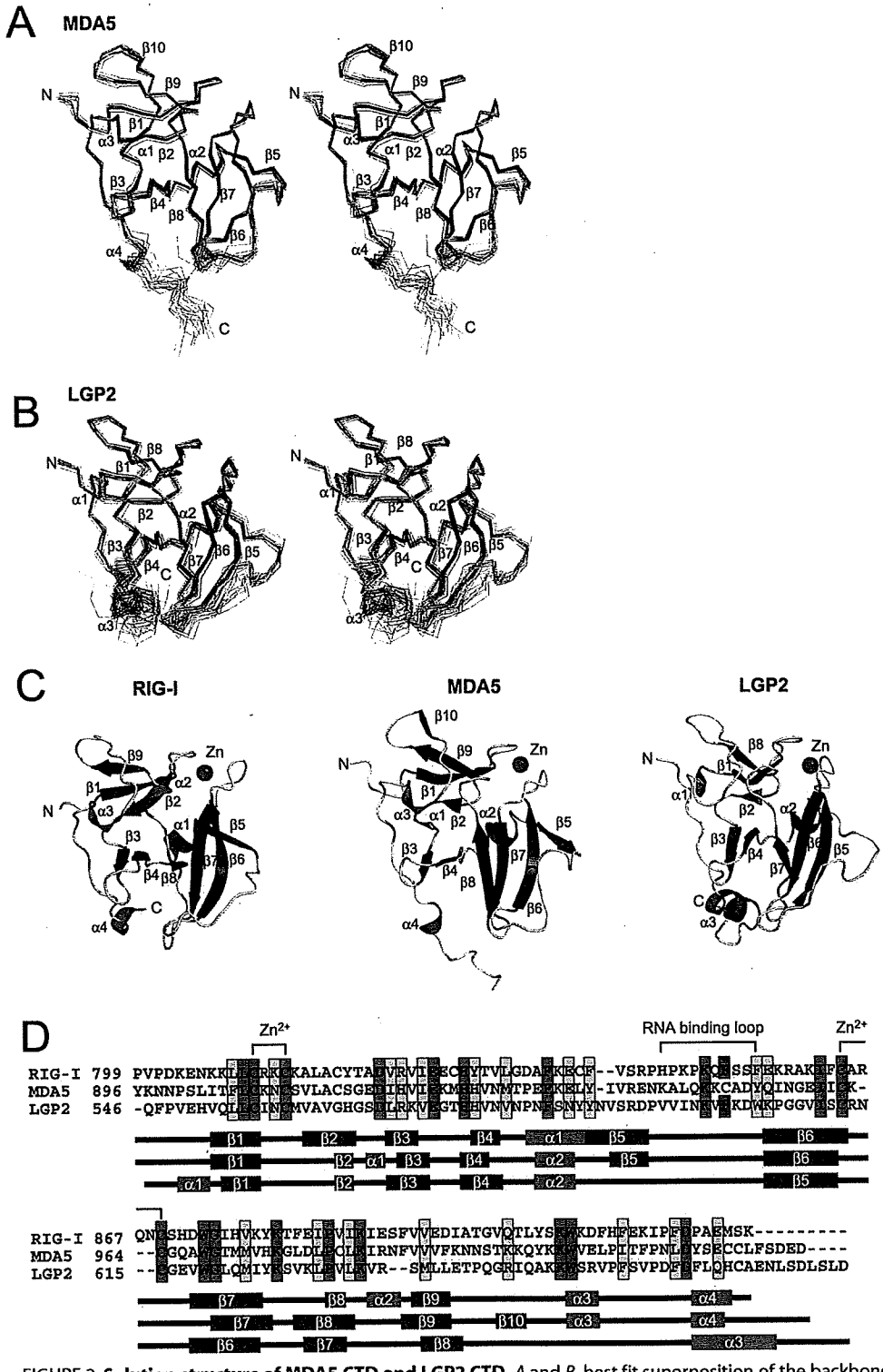


FIGURE 2. **Solution structure of MDA5 CTD and LGP2 CTD.** *A* and *B*, best fit superposition of the backbone atoms of 20 NMR-derived MDA5 CTD (*A*) and LGP2 CTD (*B*). Structures are shown in stereo. β-Strands and α-helices are shown in *blue* and *red*, respectively. *C*, *ribbon diagrams* of the structure of RIG-I CTD, MDA5 CTD, and LGP2 CTD (*left* to *right*). Secondary structure elements are labeled. The figure was prepared using PyMOL. *D*, sequence alignment of human RIG-I, MDA5, and LGP2 CTDs. ClustalX was used to align the sequences. The secondary structure elements of each CTD are indicated below the alignment. The amino acids in *red* and *yellow* indicate conserved (*red*) and type-conserved (*yellow*) residues with the Zn^{2+} binding Cys-*X*-*X*-Cys motifs and RNA binding loop. The Phe residues conserved in RIG-I and LGP2 in the RNA binding loop are colored *green*.

buffer. The ^{15}N-labeled LGP2 CTD was prepared at 92.4 μM in 250 μl of 50 mM Tris, pH 7, 250 mM NaCl, 1 mM DTT, and 10% D_2O. The volume of RNA at the addition of 1.0 equivalent molar ratio was 32.5 μl. 1×M buffer alone was also titrated, and it was confirmed that there was no effect on the NMR spectrum by the buffer. GG25, 5′-GGAAACUAAAAGGGAGAAGUGAAAGUG-3′; 2 + 25C, 5′-AUCACUUUCACUUCUCCCUUUCAGUUU-3′.

NMR Titration Study of MDA5—The experiment was performed in the same manner as in the study of LGP2 except for slightly different conditions of RNA and protein concentrations. Final concentration of RNAs were 0.644 mM, and the MDA5 CTD was prepared at 105 μM in 250 μl of 50 mM Tris, pH 7, 250 mM NaCl, 1.5 mM DTT, and 10% D_2O. The volume of RNA at the addition of 1.0 equivalent molar ratio of RNA was 40.8 μl.

Surface Plasmon Resonance Analysis—A Biacore X (Amersham Biosciences) was used for the SPR study. First, GST-fused RIG-I, MDA5, and LGP2 CTD were trapped by anti-GST antibody immobilized on the surface of a sensor CM4 chip, then RNAs were injected to detect the interaction. The standard running buffer used in the analysis was 50 mM Tris, pH 7.5, 150 mM NaCl, and the proteins and RNAs were prepared in the same buffer. All experiments were performed at 25 °C. Immobilization of anti-GST antibody and regeneration of the surface was carried out using a GST Capture Kit and an Amine Coupling Kit (Amersham Biosciences) according to the instruction manuals. The GST-RLR CTDs were prepared at 20 μg/ml and injected for 300 s at the flow rate of 20 μl/min confirming the amount of CTDs captured to the anti-GST antibody with the Resonance Unit (RU) to be ~800. As a control, GST was captured in an anti-GST antibody immobilized reference cell. The RNAs (dsRNA, 5′ppp-ssRNA, ssRNAA, and ssRNAB) were prepared at 1 μM and injected for 120 s at the flow rate of 20 μl/min and data were acquired as the difference in RU between GST-RLR CTDs captured cells and reference cells.

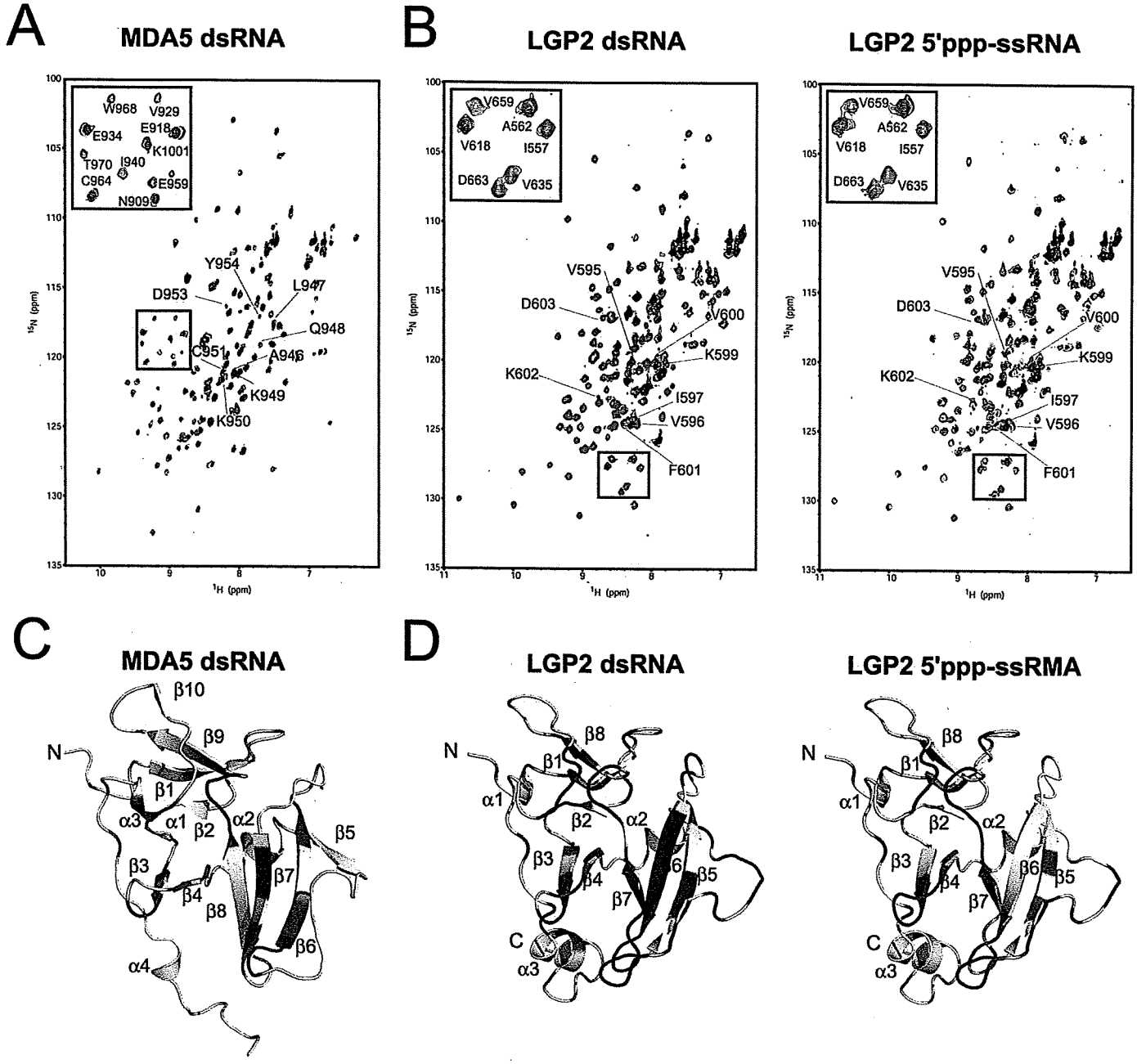
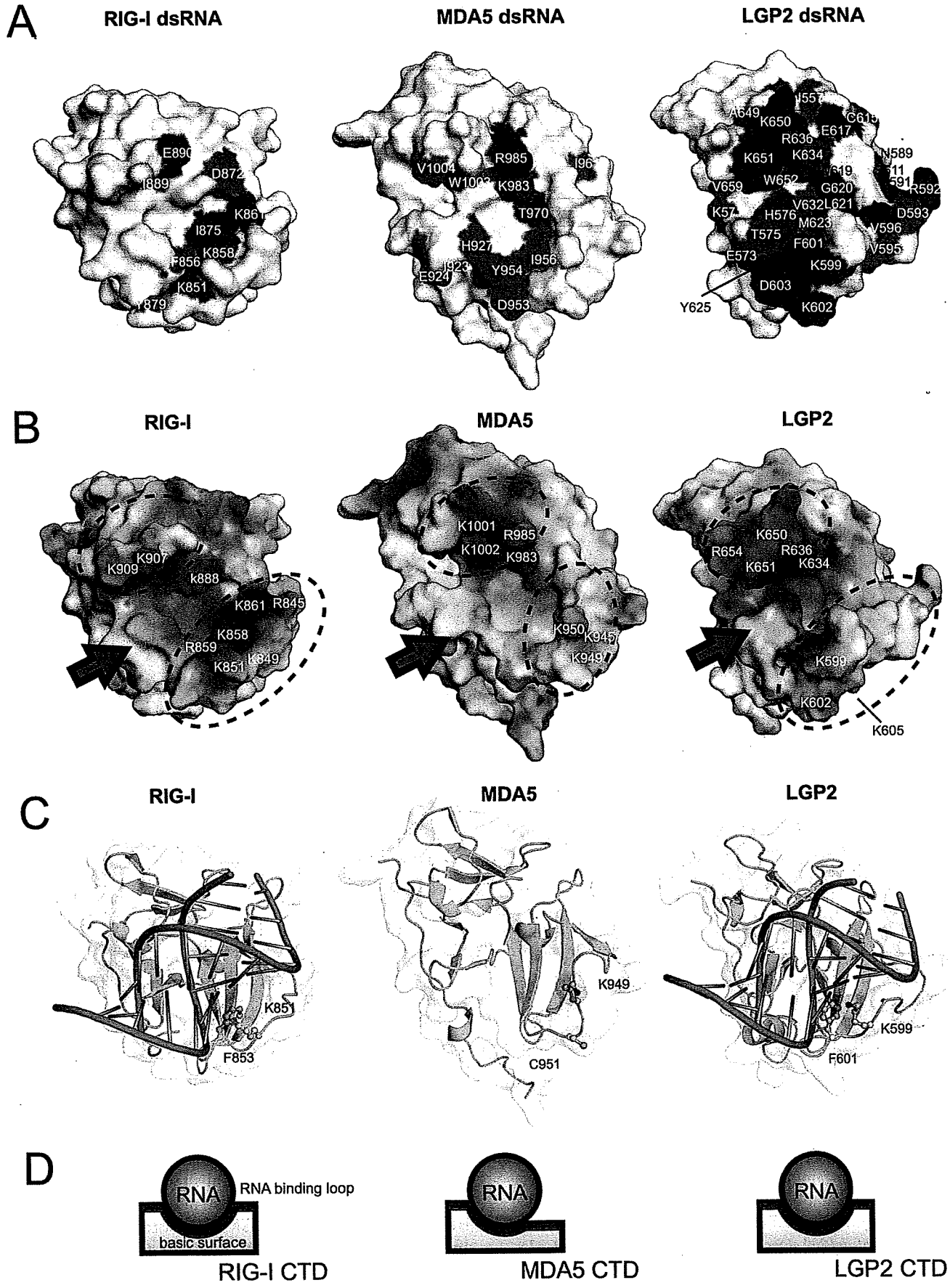


FIGURE 3. **NMR titration results for MDA5 CTD and LGP2 CTD.** *A*, NMR titration of MDA5 CTD with dsRNA. An overlay of ^{1}H-^{15}N HSQC spectra are shown in *red*, *yellow*, *green*, and *blue* at 0, 0.25, 0.5, and 1.0 equivalents (molar ratio of dsRNA to proteins), respectively, where the residues in the RNA binding loop are labeled. The *inset* shows the excerpt of the enclosed region of the spectrum with assignment. *B*, NMR titration of LGP2 CTD with dsRNA (*left panel*) and 5′ppp-ssRNA (*right panel*). The figures are prepared in the same manner as in *A*. *C*, mapping of the residues of MDA5 CTD affected by the addition of dsRNA on the *ribbon diagram*. The residues whose peaks disappeared on addition of 0.25 and 0.5 equivalent molar ratios of dsRNA to CTD are colored *blue* and *green*, respectively. *D*, mapping of the residues of LGP2 CTD affected by the addition of dsRNA and 5′ppp-ssRNA (*left* and *right panels*) on the *ribbon diagram*, colors are the same as in *C*. The orientations are the same as in Fig. 2, *A* and *B*.

EMSA—EMSA was performed essentially as described previously (8). The native gel was silver-stained (for unlabeled probe and protein), or subjected to radioactivity detection (for a ^{32}P-labeled probe). The 5′ppp-ssRNA was labeled by using 5′-^{32}P-pCp (PerkinElmer Life Sciences) and T4 RNA ligase (Takara, Ohtsu, Siga, Japan) following the manufacturer's instructions.

Preparation of dsRNA Docked with RIG-I and LGP2 CTDs—The crystal structure of RIG-I CTD (20) and short dsRNA derived from the crystal structure (PDB ID: 1YYW) were docked with the Molecular Docking Algorithm Patchdock (21). Ten residues from the basic surface of RIG-I CTD were applied as potential binding sites for RIG-I CTD, whereas no sites were applied for dsRNA. Then ten highest scoring structures were checked to establish whether they satisfy the NMR titration data. 6 of the 10 structures satisfied the NMR data, and orientations of dsRNA in those structures were similar. A dsRNA-docked model with LGP2 CTD was similarly calculated and 5 of 10 structures showed similar orientations of the dsRNA.

Solution Structures of MDA5 CTD and LGP2 CTD



RESULTS

Characterization of RLR CTDs for Specific Binding to RNAs by SPR and EMSA—To investigate the RNA binding specificity of LGP2 CTD (546–678) and MDA5 CTD (896–1025), both CTDs were expressed as GST fusion proteins and were applied to the SPR analysis together with RIG-I CTD (792–925) (Fig. 1*B* and supplemental Fig. 1). The GST fusion proteins were captured on the sensor chip, and their affinities to 5′ppp-ssRNA, dsRNA, and the two complementary ssRNAs were tested. In the SPR experiment, RIG-I CTD was bound to dsRNA and 5′ppp-ssRNA specifically, but was not bound to the two ssRNAs as has also previously been observed by EMSA (8). The LGP2 CTD was strongly bound to dsRNA and 5′ppp-ssRNA, but it exhibited lower binding affinity to ssRNA. The MDA5 CTD did not bind to any of the RNAs tested under these conditions.

The RIG-I CTD has been reported to specifically bind to both dsRNA and 5′ppp-ssRNA by EMSA (8). We produced CTDs without the GST tag and compared the RNA-binding activity by EMSA (Fig. 1*C*). The CTD of RIG-I and LGP2 specifically bound to dsRNA and 5′ppp-ssRNA, but not to ssRNAs (data not shown). This result is consistent with that of the SPR analysis. The EMSA data shows the appearance of slower mobility complexes, most clearly observed with LGP2 CTD and dsRNA. These complexes are likely corresponding to dsRNA bound to multiple CTD molecules (see "Discussion"). The MDA5 CTD exhibited low binding affinity to dsRNA, and its biding affinity to 5′ppp-ssRNA was very low. This is reminiscent of the previous observation that full-length MDA5 bound much more weakly to poly(I:C)-Sepharose than to RIG-I (22) (supplemental Fig. 3). Summarizing, these results suggest that RLR CTDs play a critical role in the RNA recognition by RLR.

Solution Structures of MDA5 and LGP2 CTDs—The solution structures of MDA5 and LGP2 CTDs were determined by NMR (Fig. 2, *A* and *B*), and both MDA5 and LGP2 CTDs have a fold similar to that of RIG-I CTD and contain a single conserved Zn^{2+} binding site (Fig. 2, *C* and *D*). We confirmed that both CTDs contain a single Zn^{2+} ion, by atomic absorption spectroscopy, similar to RIG-I CTD (supplemental Table 1). The core of the structure is composed of a central anti-parallel β sheet (Fig. 2, *C* and *D*) (β3–β8 in RIG-I and MDA5, β3–β7 in LGP2). A further anti-parallel β sheet (β1, -2, and -9 in RIG-I; β1, -2, -9, and -10 in MDA5; β1, -2, and -8 in LGP2) is located on top of the central β sheet, and there are several short helices (α1–3 in RIG-I and MDA5, α1 and -2 in LGP2) attached to the two β sheets that help stabilize the structure of the CTDs. The C-terminal region of RIG-I, MDA5, and LGP2 CTDs has a long loop that is surrounding one edge of the central β sheet and the C-terminal helix (α4 in RIG-I and MDA5, α3 in LGP2) is located at the bottom of the central β sheet. The long loop between β5 and β6 that was implicated in RNA recognition in RIG-I, is also present in all RLR CTDs (termed "RNA Binding Loop" below). Recently, the crystal structure of LGP2 CTD was reported. Although the RNA binding loop was not observed in the crystal structure, the author implied that the loop is important for RNA specificity (23).

Although the RLR CTDs have a similar global fold with a large basic surface on the central β-sheet (see Fig. 4*B*), a close inspection of the RLR CTD structures shows small but appreciable differences (Fig. 2*C*). Similar to RIG-I, the C-terminal helices of LGP2 and MDA5 are located at the bottom of the central β-sheet and interacts with the loop between β6 and β7 (β7 and β8 in MDA5), whereas the C-terminal helix of MDA5 is projected outward. However, the regions following the C-terminal helix are flexible, because these regions exhibit small steady-state NOEs (<0.3 in supplemental Fig. 2). There is also an appreciable difference in the RNA binding loop. The N-terminal part of the RNA binding loop in MDA5 CTD is clipped onto the surface of β6 so that the conformation of the rest of the RNA binding loop is restricted and gives rise to an open and flat conformation (Fig. 2, *A* and *C*, *middle*). Thus, the basic surface becomes a more open structure (Figs. 2*C*, 4*B*, and 4*D* (*middle*)). Although the C-terminal RNA binding loop in LGP2 is restrained but stands upright on β6 and encloses the central β-sheet together with the central part of the loop between β8 and α3, a loop N terminus to β8, forming a basic groove (Figs. 2*C*, *right*, and 4*B*, *right*). A similar basic groove was identified as the RNA binding surface in RIG-I (Figs. 2*C* and 4*B*, *left*). Both RIG-I and LGP2 but not MDA5 share the common structural feature for the RNA binding loop. It is to be noted that the RNA binding loops in MDA5 and LGP2 are relatively restricted, based on the steady-state NOE measurements (>0.5 in supplemental Fig. 2).

NMR Titration of MDA5 and LGP2 CTDs—To establish whether the basic surfaces of MDA5 and LGP2 CTDs are responsible for the binding to viral RNAs, we performed NMR titration with the RNAs (dsRNA, 5′ppp-ssRNA, and ssRNAs), used in the SPR and EMSA assays reported above, against ^{15}N-labeled MDA5 and LGP2 CTDs (Fig. 3, *A* and *B*).

In MDA5 CTD, there were negligibly small chemical shift changes in the ^{1}H-^{15}N HSQC spectra even with addition of 2.0 equivalent molar ratios of 5′ppp-ssRNA and two ssRNAs, indicating an absence of interaction between MDA5 CTD and these RNAs (data not shown). However, upon addition of 0.5 equivalent molar ratio of dsRNA, an appreciable number of peaks disappeared possibly due to intermediate exchange processes (Fig. 3, *A* and *C*). The disappeared peaks were assigned to the residues on β3, β4 and their connecting loop, β6, β7, the loop C terminus to β8, and α3. However, most of the residues on the RNA binding loop other than Asp-953 and Tyr-954 did not

FIGURE 4. **Structural comparisons of RLR CTDs.** *A*, surface representation of the dsRNA affected surface of RLR CTDs (*left* to *right*, RIG-I CTD, MDA5 CTD, and LGP2 CTD). The residues that disappeared in NMR titration experiments upon addition of 0.25 and 0.5 equivalents of dsRNA are colored in *blue* and *green*, respectively. The residue numbers are also shown. The data of the dsRNA affected surface of RIG-I CTD was derived from our previous study (8). *B*, electrostatic surface potentials of the RLR CTDs (*left* to *right*, RIG-I CTD, MDA5 CTD, and LGP2 CTD). *Dotted circles* indicate the banks surrounding the basic surface. *C*, *left*, dsRNA-bound model of RIG-I CTD. *Right*, dsRNA-bound model of LGP2 CTD. The Lys and Phe residues in the RNA binding loop conserved in RIG-I and LGP2 are shown in *ball-and-stick models* and labeled. *Middle*, structure of MDA5 CTD, the residues corresponding to the Lys-851 and Phe-853 in RIG-I are shown. All figures are shown in the same structural orientation as in Fig. 2*A*. *D*, schematic diagrams of dsRNA bound to the basic surface of RIG-I CTD (*left*), MDA5 CTD (*mid*), and LGP2 CTD (*right*) viewed from the *arrows* indicated in Fig. 4*B*.

disappear (Fig. 3*C*). This strongly suggests that the MDA5 CTD uses the basic surface to interact with dsRNA similar to RIG-I, but the RNA binding loop does not appear to be involved in RNA binding. This result is consistent with EMSA (Fig. 1*C*) but not with the SPR analysis (Fig. 1*B*), where there was no interaction between MDA5 and dsRNA. This inconsistency may be due to the higher concentration of MDA5 CTD and dsRNA used in the EMSA and NMR titration studies than in the SPR studies.

For LGP2, even with addition of 2.0 equivalent molar ratios of ssRNAs, ^{1}H-^{15}N HSQC of LGP2 CTD spectra showed neither chemical shift changes nor any loss of signal intensities (data not shown), indicating that LGP2 CTD does not recognize ssRNAs. However, the addition of a 0.25 equivalent molar ratio of dsRNA and 0.5 equivalent molar ratio of 5′ppp-ssRNA caused the disappearance of a large number of peaks (Fig. 3*B*). When the residues corresponding to the peaks, which disappeared by dsRNA titration, were mapped on the three-dimensional structure of LGP2 CTD, these were located on β3, β4 and the connecting loop, the RNA binding loop, β5, β6 and the loop C terminus to β6, β7 and the loop between β7 and β8, and the C terminus to β8 (α3 in RIG-I) (Fig. 3*D*, *left*). The binding surface for 5′ppp-ssRNA is similar to that of dsRNA but is less extensive (Fig. 3*D*, *right*). It is to be noted that unlike MDA5, a large number of the peaks assigned on the RNA binding loop (Fig. 3*B*) disappeared in both cases, indicating that the RNA binding loop plays a critical role in the recognition of RNA ligands in LPG2. Residues of CTD involved in the interaction with dsRNA along with surface charges are mapped for comparison among RLRs (8) (Fig. 4, see "Discussion").

Analysis of the Structure-Function Relationship of RLR by Mutagenesis—We generated MDA5 mutants considering the results of the chemical shift perturbation study and the electrostatic surface potential (Fig. 4, *A* and *B*, *middle*). The KR983/985AA and KK1001/1002AA mutants were expressed in MDA5$^{-/-}$ mouse embryo fibroblasts, and next the ability to mediate the virus-responsive reporter gene activation was examined. To stimulate MDA5, the cells were transfected with poly(I:C), a known chemical ligand for MDA5 (Fig. 5*A*). Irrespective of the absence of MDA5, vector-transfected cells exhibited reporter activation upon stimulus by poly(I:C), presumably due to the activation of RIG-I by short poly(I:C) present in this batch of poly(I:C) (300–3000 bp, data not shown). Expression of wt MDA5 strongly enhanced this response, suggesting this increase is mediated by ectopically expressed MDA5. However both of the mutants exhibited only slight activation. Comparable levels of wt and mutant MDA5 were expressed in these cells (Fig. 5*B*), suggesting that mutated basic amino acid residues perform critical functions.

The solution structures and chemical shift perturbation studies of RIG-I and LGP2 may highlight the presence of the RNA binding loop. Because the specificity of RIG-I and LGP2 CTDs to viral RNAs are similar, we focused on a Phe residue that are conserved in the RNA binding loop of RIG-I and LGP2 but is not conserved in MDA5. First, we produced and purified full-length wt RIG-I and mutants with F853A and F853C that mimic MDA5 (Fig. 2*D*) and KK888/907AA (8) in 293T cells (Fig. 6*A*). Comparable levels of wt and mutant proteins were

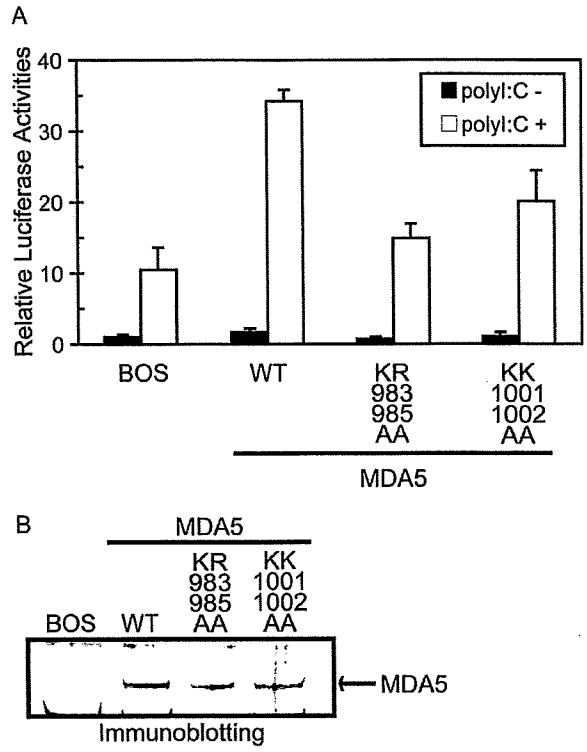


FIGURE 5. **Functional analysis of basic residue mutations of MDA5 on the basic surface.** *A*, MDA5$^{-/-}$ mouse embryo fibroblasts were transfected with the reporter gene, p-125Luc, and pRL-tk, together with the expression vector for MDA5 and mutants. Cells were stimulated by transfection with poly(I:C) and subjected to a dual-luciferase assay. The values are the means ± S.D. from triplicate experiments. The relative luciferase activity was calculated by considering the luciferase activity from cells transfected with empty vector (*BOS*) as 1.0. The cell lysates were analyzed for expression of MDA5 and mutants by Immunoblotting (*B*).

produced and recovered, suggesting that these mutations did not alter the stability of RIG-I (Fig. 6*A*). The purified proteins were subjected to EMSA using dsRNA and 5′ppp-RNA as probes (Fig. 6, *B* and *C*). The results clearly show that, along with the previously identified basic residues on the basic surface, Phe-853 in the RNA binding loop is critical for binding with dsRNA and 5′ppp-ssRNA. Further, the mutation of Phe-853 reduced the signaling activity of RIG-I when stimulated with these RNA species, particularly with dsRNA (Fig. 6, *D* and *E*). Next the involvement of the RNA binding loop in RNA recognition by LGP2 was examined. Three basic residues in the RNA binding loop or the conserved Phe were substituted (KKK599/602/605AAA, F601A, and F601C). These LGP2 mutant proteins along with the wt protein were produced and recovered at similar levels suggesting that these mutations did not affect protein stability dramatically (Fig. 7*A*). The purified LGP2 proteins were examined for RNA binding. The results in Fig. 7 (*B* and *C*) clearly show that the residues on the RNA binding loop are critical for LGP2 to recognize dsRNA and 5′ppp-ssRNA.

To directly compare RNA-binding activity of full-length RIG-I, MDA5, and LGP2 and their mutants, EMSA was performed under the same conditions (supplemental Fig. 3). The result shows that the binding properties of full-length RLRs are