npg

LETTER TO THE EDITOR

Crystal structure and nucleotide selectivity of human IFIT5/ISG58

Dear editor,

Interferon-stimulated gene 56 (ISG56) family members, ISG56, ISG54, ISG58 and ISG60, also known as interferon-induced protein with tetratricopeptide repeats (TPR) 1 (IFIT1), IFIT2, IFIT5 and IFIT3, respectively, play crucial roles in human antiviral defense systems. Recent studies have shown that ISG56 specifically recognizes viral 5′-pppRNA [1]. Using structural biology methods, we found that ISG54 possesses a right-handed helical channel in its C-terminal region, which recognizes AU-rich RNAs [2]. ISG54 mutants whose nucleotide-binding sites were destroyed showed a significant loss of their antiviral activities [2]. As the expression patterns of ISG56 family proteins have been reported to be cell- and inducer-specific [3], we hypothesize that different ISG56 family members likely recognize distinct substrates.

Recently, ISG58 expression was shown to be induced upon infection with influenza virus [4]. However, the exact roles of ISG58 in host defense have yet to be elucidated. From an evolutionary point of view, human ISG58 is different from all the other ISG56 family members: it does not have a mouse homologue, and is the only ISG56 family protein so far identified in birds, marsupials, frogs and fish [5]. ISG58 was found to copurify with a diverse mixture of small RNAs consisting mainly of a subset of tRNAs, and a number of noncoding RNAs [6].

To further understand the functional mechanism of this protein, we determined the 2.1 Å crystal structure of ISG58 (Figure 1A). The crystallographic data and refinement statistics are shown in Supplementary information, Table S1. The crystal structure shows that ISG58 consists of 24 α-helices that pack into a special "V"-like structure. All the helices (helices 3-24) in ISG58, except the first two helices at the N-terminus, can be separated into 11 tandem TPR motifs (Supplementary information, Figure S1A). Like ISG54, the ISG58 structure can be divided into three regions: the N-terminal (helices 1-6), middle (helices 7-20) and C-terminal (helices 21-24) regions (Figure 1A). The N-terminal sequence of ISG58 is conserved among ISG56 family members [2]. Loop L1

that connects helix 1 and helix 2 covers the inner surface of the N-terminus of ISG58 (Supplementary information, Figure S1B). Helices 7-20 are located at the bottom of the "V"-like structure, and form a spiral-like channel just like that observed in the ISG54 structure (Supplementary information, Figure S1C). Helices 21-24 form the C-terminus of ISG58, which is poorly conserved among ISG56 family proteins [2]. It is therefore possible that this region might play an important role in determining the different functions of different ISG56 family members.

It was previously reported that ISG58 associates with small RNAs in pull-down assays [6], indicating that it can interact directly or indirectly with nucleotides. Our structural study shows that positively-charged residues including lysines and arginines are enriched at the inner surface of the "V"-like structure (Figure 1B). In addition, several aromatic residues are also present in this region, such as phenylalanine and tyrosine. This observation led us to hypothesize that ISG58 can bind directly to negatively charged nucleotides, and that the large spiral-like channel within the ISG58 structure serves as a nucleotide-binding site.

To test this hypothesis, the nucleotide-binding ability of ISG58 was analyzed by electrophoretic mobility shift assays (EMSA), using different oligos including 14-nucleotide (nt) polyA, polyU and polyC single-stranded (ss) RNAs and the corresponding ssDNAs (Supplementary information, Table S2). Our results indicate that ISG58 binds to polyA and polyU ssRNAs with high affinity (Figure 1C and 1D). In contrast, it only weakly binds to ssDNA such as polyA ssDNA in high concentration (Figure1C and 1D). The binding affinity of ISG58 to ssRNA, as measured in a semiquantitative manner by EMSA using 100 nM polyA ssRNA and ISG58 protein in various concentrations, is within the range of 10-100 nM (Figure 1D). After testing 12-nt AU-rich and GC-rich double-stranded (ds) RNAs and dsDNAs, we unexpectedly found that, unlike ISG54 that recognizes AU-rich dsRNA, ISG58 selectively binds to AT-rich dsDNA (Figure 1C), and only weakly binds to AU-rich dsRNA

(Figure 1D). The AT-rich dsDNA-binding affinity of ISG58 is comparable to its polyA ssRNA-binding ability (Figure 1D). We also generated an RNA/DNA duplex by incubating a well-recognized 16-nt ssRNA with a ssDNA of complementary sequence. However, ISG58 failed to recognize this RNA/DNA hybrid substrate (Figure 1D). In conclusion, these results indicate that the nucleotide-binding ability of ISG58 differs from that of ISG54: ISG54 displays a highly selective interaction with dsR-NAs, whereas ISG58 selectively binds to certain types of ssRNAs (polyA and ployU ssRNAs) and dsDNA (AT-rich dsDNA). This DNA-binding ability distinguishes ISG58 from all the other ISG56 family proteins.

The putative nucleotide-binding sites of ISG58 were
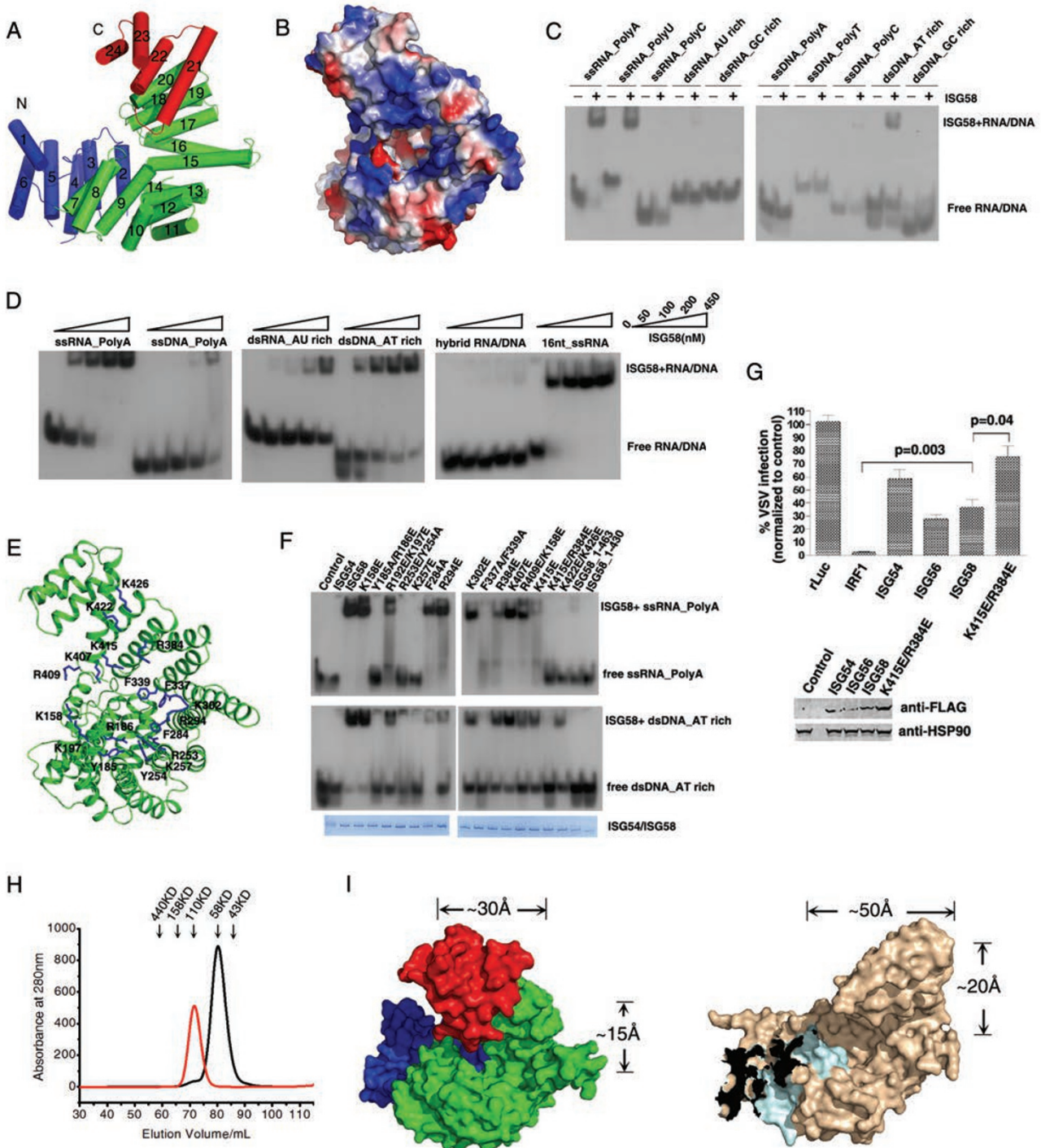
investigated using mutagenesis experiments (Figure 1E). We mutated positively-charged arginine and lysine residues in the channel of ISG58 to glutamate, and aromatic residues to alanine (Y185A/R186E, R192E/K197E, R253E/Y254A, K257E, F284A, R294E, K302E, F337A/F339A, R384E, K415E, K415E/R384E and K422E/K426E). Several mutations on the rim of the channel, such as K158E, K407E and R409E/K158E, were also studied. As the C-terminal regions of ISG56 proteins are poorly conserved [2], we constructed two C-terminal truncated forms of ISG58, ISG58_1-430 and ISG58_1-463, in order to investigate whether the C-terminus itself is important for nucleotide binding of ISG58. The nucleotide-binding ability of each ISG58 mutant was analyzed using both polyA ssRNA and AT-rich dsDNA as substrates. The EMSA results indicate that the C-terminal truncations and mutations in the putative nucleotide-binding channel generally abolished the nucleotide-binding ability of ISG58, while mutations on the rim of the channel had little effect (Figure 1F). The mutations Y185A/R186E, R253E/Y254A, K257E, F337A/F339A, K415E/R384E and K422E/K426E resulted in reduced binding to both RNA and DNA. Interestingly, two mutations within the channel, F284A and R294E, significantly altered the nucleotide-binding selectivity of ISG58. Both mutants still selectively bound to ssRNA, but the binding to dsDNA was greatly reduced (Figure 1F). Based on these results, we suggest that ISG58 binds to both types of nucleotides by using its C-terminal channel. The inner radius of the channel is 15 Å, which is suitable for ssRNA binding (Figure 1I). However, for dsDNA binding, the C-terminal channel of ISG58 likely needs to undergo a conformational change, in which residues such as F284 and R294 would be exposed and contribute to the binding surface.

ISG58 expression can be induced strongly upon viral infection [4], therefore, we sought to examine the potential inhibitory effect of ISG58 on viral replication by using a model of vesicular stomatitis virus expressing GFP (VSV-GFP). Earlier work reported that both ISG54 and ISG56 inhibited VSV replication when they were overexpressed in HEK293T cells [1, 7]. Similarly, ISG58 reduced VSV replication by more than 60% (Figure 1G). Importantly, the ISG58 mutant K415E/R384E exhibited impaired antiviral activity, consistent with its impaired nucleotide-binding ability (Figure 1G). These results suggest that the antiviral activity of ISG58 is dependent on its nucleotide-binding ability.

Although the sequence similarity of ISG58 and ISG54 is greater than 50% (Supplementary information, Figure S2), they tend to bind to different types of nucleotides. To explore why ISG58 binds to both ssRNA and dsDNA, we next compared its structure with that of ISG54 that recognizes dsRNA. First, in contrast to the dimer-forming ISG54 [1, 2], ISG58 is stably monomeric in the solution state (Figure 1H). We found that helices 1-6 and helices 10-16 of ISG58 superimpose well with the corresponding regions of one monomer (monomer A) of the ISG54 dimer, while helices 7-9 of ISG58 overlap with helices 7-9 from the other ISG54 monomer (monomer B) (Supplementary information, Figure S3A). Superimposing the structure of ISG58 on an ISG54 chimera composed of helices 1-6 and 10-22 from monomer A and helices 7-9 from monomer B revealed that, unlike ISG54, the unconserved C-terminus of ISG58 bends towards its N-terminus (Supplementary information, Figure S3B). Thus, we propose that the differences in the C-terminal structures of ISG54 and ISG58 provide a possible structural basis for the differences in their nucleotide-binding patterns. Second, we observed that the length of the ISG54 channel is 50 Å, while that of ISG58 is 30 Å. The inner radius of the bottom of the helical channel in

Figure 1 Crystal structure, nucleotide-binding activity and antiviral activity of ISG58. **(A)** Cartoon diagram of ISG58 structure. The N-terminal (helices 1-6), middle (helices 7-20) and C-terminal (helices 21-24) regions are shown in blue, green and red, respectively. **(B)** Electrostatic surface representation of ISG58 generated by Pymol. The positive, negative and neutral residues are shown in blue, red and white, respectively. **(C)** Binding of ISG58 to different nucleotides. The concentrations of the nucleotides and ISG58 protein used were 30 nM and 150 nM, respectively. **(D)** The nucleotide-binding affinities of ISG58. Binding of ISG58 to polyA ssRNA, polyA ssDNA, AU-rich dsRNA, AT-rich dsDNA and hybrid RNA/DNA were tested by EMSA. 16-nt ssRNA was used as positive control. The final concentration of the nucleotides used was 100 nM. **(E)** The putative nucleotide-binding sites of ISG58. Residues speculated to be involved in nucleotide binding are labeled and shown as blue sticks. **(F)** Binding of WT and mutant ISG58 to polyA ssRNA and AT-rich dsDNA. The concentrations of the nucleotides and the proteins used were 100 nM and 500 nM, respectively. Lane 1 was loaded with nucleotides alone. Equal amounts of proteins were loaded in each lane, as shown in the bottom panels. **(G)** HEK293T cells were transfected with ISG or control constructs (Renilla luciferase as negative control and IRF1 as positive control) as indicated and infected with VSV-GFP 40 h post-transfection. Cells were collected 8 h later, and GFP expression was quantified using flow cytometry. The expression levels of FLAG-tagged ISG54, ISG56, ISG58 and ISG58 K415E/R384E were assessed by immunoblotting with an anti-FLAG antibody. The data shown represent three independent experiments. **(H)** Purified ISG58 (black line) and ISG54 (red line) proteins were applied to a Hi-load Superdex-200 column. The elution volume of samples and the sizes of marker proteins are indicated. **(I)** Differences of the nucleotide-binding channel between ISG54 and ISG58. The dimensions of the nucleotide binding channel of ISG58 (left panel) and ISG54 (right panel) are indicated next to the surface diagrams.

ISG54 and ISG58 is about 20 Å and 15 Å, respectively (Figure 1I). The difference in channel size provides a plausible explanation as to why ISG58 binds to ssRNAs. Also, we hypothesize that the C-terminus of ISG58 likely undergoes a conformational change in order to accommodate the larger dsDNA. In addition, the C-terminus of ISG54 contains a long, positively-charged region, whereas that of ISG58 contains multiple hydrophobic residues. The different shapes and electrostatic features of ISG54 and ISG58 suggest that they likely bind to different nucleotides *in vivo*. We thus propose that ISG58 and ISG54 may have different antiviral spectra.

While this paper was under revision, two other groups [8, 9] also published ISG58 crystal structure, which is consistent with the results presented here. Collectively, these studies, together with the present study, show that ISG58 binds to certain types of RNA, and that its RNA-binding ability is important to its antiviral activity. Furthermore, our current study suggests that ISG58 also binds to DNA, though the exact function of this DNA-binding activity is currently unknown. It is conceivable that the DNA-binding ability may enable ISG58 to recognize DNA viruses, thus expanding the virus-recognition spectrum of ISG58.

Acknowledgments

Feng Feng[1, *], Lingmin Yuan[2, *], Yao E Wang[3, *], Christopher Crowley[3], Zongyang Lv[1], Jingjing Li[1], Yingfang Liu[1], Genhong Cheng[3], Su Zeng[2], Huanhuan Liang[1]

[1]*State Key Laboratory of Biomacromolecules, Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101, China; [2]Department of Pharmaceutical Analysis and Drug Metabolism, Zhejiang University, Hangzhou, Zhejiang 310058, China; [3]Department of Microbiology, Immunology and Molecular Genetics, University of California, Los Angeles, CA 90095, USA*
*These three authors contributed equally to this work.
Correspondence: Huanhuan Liang
E-mail: hhliang@moon.ibp.ac.cn

References

1 Pichlmair A, Lassnig C, Eberle CA, *et al. Nat Immunol* 2011; **12**:624-630.
2 Yang Z, Liang H, Zhou Q, *et al. Cell Res* 2012; **22**:1328-1338.
3 Fensterl V, Sen GC. *J Interferon Cytokine Res* 2011; **31**:71-78.
4 Vanderven HA, Petkau K, Ryan-Jean KE, *et al. Mol Immunol* 2012; **51**:316-324.
5 Zhou X, Michal JJ, Zhang L, *et al. Int J Biol Sci* 2013; **9**:200-208.
6 Hogg JR, Collins K. *Genes Dev* 2007; **21**:3067-3072.
7 Fensterl V, Wetzel JL, Ramachandran S, *et al. PLoS Pathog* 2012; **8**:e1002712.
8 Abbas YM, Pichlmair A, Gorna MW, *et al. Nature* 2013; **494**:60-64.
9 Katibah GE, Lee HJ, Huizar JP, *et al. Mol Cell* 2013; **49**:743-750.

(**Supplementary information** is linked to the online version of the paper on the *Cell Research* website.)