Z-RNA and the flipside of the SARS Nsp13 helicase

alan.herbert@insideoutbio.com

Alan Herbert[1,2*], Alexander Shein[2] and Maria Poptsova[2]

[1]InsideOutBio, 42 8th Street, Charlestown, MA

[2] Laboratory of Bioinformatics, Faculty of Computer Science, National Research University Higher School of Economics, 11 Pokrovsky Bulvar, Moscow, Russia 101000

*Corresponding author

Keywords: coronavirus;SARS;MERS;Z-RNA;ZBP1;Necoptosis

Abstract

We present evidence that the severe acute respiratory syndrome coronavirus (SARS) non-structural protein 13 (Nsp13) modulates the Z-RNA dependent regulated cell death pathway of necroptosis (*1*). We show that Z-prone sequences (called flipons (*2*)) exist in coronavirus and provide a signature that enables identification of the animal viruses which have become human pathogens. We also identify a potential RHIM in Nsp13. These two observations allow us to suggest a model in which Nsp13 may regulate Z-RNA-initiated RHIM-dependent cell death outcomes at two steps. The first step involves possible new ATP-independent Z-flipon helicase activity in Nsp13, which is distinct from the activity of the canonical A-RNA helicase. This activity unwinds/quenches nascent Z-RNAs, preventing their sensing by ZBP1. The second step involves RHIM-dependent inhibition of ZBP1, RIPK3 and/or RIPK1, preventing cell death downstream of Z-RNA sensing. Together the RHIM and Z-flipon helicase have the potential to alter the host response to the virus and the effectiveness of drugs targeting the NSP13 helicase.

Introduction

Z-RNA and Flipons. The left-handed Z-DNA and Z-RNA conformations form from their right-handed counterparts of B-DNA and A-RNA by flipping the bases over to produce the characteristic zig-zag backbone (*3*). The energy to produce the flip is produced by processive enzymes such as helicases and polymerases that generate negative supercoiling in their wake. In other situations, the flip reduces topological stress arising when different nucleic acids basepair and become entangled. Sequences that form these left-handed structures under physiological conditions are called flipons and often are composed of alternating pyrimidine and purine residues, such as $(CG)_n$, $(CA)_n$ and $(UG)_n$. The propensity to adopt the Z-conformation can be scored using the ZHUNT3 (ZH3) program (*4*). Much of the energy cost for forming Z-DNA and Z-RNA is in the creation of junctions between right-handed and left-handed helices. For Z-RNA, the flip occurs more easily when dsRNA contain basepair mismatches, non-canonical basepairs or unpaired residues at both ends of the Z-forming segment (*5*).

Z-RNA and innate immunity. Z-DNA and Z-RNA are recognized in a conformation specific manner by protein containing the Zα domain that was first discovered in the dsRNA editing enzyme ADAR1 (*6*), but is also found in many other proteins, including Z-DNA binding protein (ZBP1) and virally encoded proteins like E3 that is encoded by vaccinia virus (*7*). Interaction between Zα proteins regulates the regulated necroptotic cell death pathway (*8, 9*). The pathway is triggered when ZBP1 binds to Z-DNA or Z-RNA and activates receptor interacting protein kinases (RIPKs) 1 and 3, through the RIP homotypic interaction motif (RHIM) shared by all three proteins. RIPK3 then phosphorylates Mixed Lineage Kinase Domain Like (MLKL) pseudokinase, leading to necroptosis, while RIPK1 induces Caspase-8-dependent apoptosis, via the adaptor protein FADD. The RHIM is a ~40 aa motif, first identified in RIPK1 (*10*), which contains an (I/V)Q(I/L/V)G sequence at its core (*11*). The pathway plays an important role during infection by the negative RNA stranded influenza virus (*12*), as well as upon infection with the

24 herpesviruses murine cytomegalovirus (mCMV) and herpes simplex virus (HSV)-1/2 and the poxvirus
25 vaccinia virus (*13*).

26 V**iruses and ZBP1 dependent necroptosis.** Viruses that are prone to form Z-DNA and Z-RNA have
27 developed strategies to regulate ZBP1-dependent necroptosis. These include encoding ZBP1 homologs,
28 such as E3, that compete with ZBP1 for Z-RNA (*14*). Another strategy by which viruses regulate
29 ZBP1/RIPK signaling is by encoding proteins with RHIMs. So far, only RHIMs produced by DNA viruses
30 like the Herpesviridae mCMV, HSV-1 and HSV-2 are known to play an important role in virulence (*15*).
31 Notably, no RNA virus has yet been shown to encode a RHIM-containing protein.

32 **Results**

33 **A corona virus-specific Z-flipon signature.** We were interested in whether the SARS family
34 coronaviruses might also regulate Z-RNA dependent ZBP1 activation and cell death. We used a two-
35 pronged approach. First, we searched for Z-prone sequences in coronavirus using the program ZH3. We
36 were interested in those sequences that altered as the virus adapted to humans. We found that all
37 examined coronaviruses contained sequences that were Z-prone, and that the Z-signature for each
38 strain of virus is unique (Figure 1).Using the signature, it is easy to identify the host animal from which
39 the pathogenic human viruses arose. For example, we show that the signatures for Llama coronavirus
40 (CoV) and Middle East respiratory syndrome (MERS) are identical, as are those for civet CoV and Severe
41 Acute Respiratory Syndrome-CoV1 (SARS1) and those for bat CoV RaTG13 and SARS-CoV2 (SARS2)
42 (Figure 1 Panels A and B).

43 **Variations in Coronavirus Z-prone sequences** The ZH3 scoring is based on the energetics of Z-DNA
44 flipping, where a score of around 700 is sufficient to change flipon conformation under physiological
45 conditions. While not experimentally calibrated for the formation of Z-RNA, the scores observed for the
46 CoVs are far in excess of this value and it likely that the sequences in these viruses are flipons. With this

47 caveat in mind, we observed the loss in the 1ab transcript of a strong Z-forming sequence when
48 comparing SARS1 to SARS2. The reduced ZH3 score in SARS2 is due to three synonymous mutation in
49 adjacent codons that preserve a hairpin sRNA structure and the peptide sequence PARAR (residues 335-
50 339 of SARS2 Nsp13). This arginine rich sequence has the potential to bind nucleic acids (figure 2A and
51 B) (*16*). Two Z-prone sequences were present in the equivalent region of MERS Nsp13. The MERS
52 peptide, PAKAR, has a lysine in place of the first arginine, a conservative substitution.

53 **A putative RHIM associated with coronaviruses severely pathogenic in humans.** Second, we searched
54 for viral protein containing a RHIM and found that the prototypical VQIG sequence was present in the
55 Nsp13 protein of highly pathogenic human CoVs, but not in other CoVs thought to infect humans (Figure
56 2 D and E).

57 **The structure of NSP13 helicase**. Both strategies focused our attention of the Nsp13 helicase and its
58 potential novel role as a Z-flipon helicase. Nsp13 unwinds both DNA and RNA in the 5'->3' direction. It is
59 a member of the helicase superfamily 1B (SF1B) (*17*). Along with other CoV encoded proteins, it forms a
60 viral replication-transcription complex. High resolution crystallographic and electron-microscopy
61 structures of NSP13 have recently been published and suggests that a second Nsp13 cooperates with
62 the first to enhance translocation along the CoV genomic RNA (*18-20*). The Nsp13 helicase has two
63 canonical RecA domains between which ATP is bound and through which single-stranded RNA egresses
64 from the complex. Nsp1 also has three N-terminal domains that are unique to nidovirus helicases. The
65 stalk domain connects a zinc-binding (ZincBD) domain to the 1B domain (Figure 3).

66 The Nsp13 RHIM is present in the 1B domain, which bridges the RecA1 and RecA2 domains. Rather
67 surprisingly, the peptide PARA motif, which is in the RecA1 domain, interacts with the RHIM when the
68 Nsp13 is in the ATP-free closed conformation (Figure 3A). In the active conformation, the RHIM and
69 PARA separate to open a channel through which the single-stranded RNA passes (Figure 3B). Further

70 separation of the 1B domain from the RecA2 domain renders Nsp13 unable to translocate on RNA (*20*).
71 The cavity created by this separation appears large enough to accommodate a dsRNA structure (Figure
72 3C).

73 **Discussion**

74 The cavity is lined by the residues known to bind single-stranded RNA (ssRNA) (shown by white space fill
75 carbons in figure 4), both to the phosphate backbone and also in a base-specific manner. Into the newly
76 formed cavity projects the PARA peptide, with R337 and R339 forming a hook. Also, tyrosine Y205, along
77 with W178, faces the interior of the cavity. Potentially Y205 and W178 could engage each other in an
78 edge to face configuration similar to that of Y177 and W195 in the Zα domain and enable the
79 conformation-specific recognition of Z-RNA.

80 **Novel Nsp13 flipon helicase activity.** Docking Nsp13 to Z-RNA in silico leads us to propose a model
81 where Nsp13 acts as a flipon helicase, preventing Z-RNA formation by capturing single-stranded RNA
82 formed during the flip to and from A-RNA. The strand separation generated is powered by the free
83 energy stored in Z-RNA rather than through ATP hydrolysis. The process of strand capture would be
84 enhanced if Nsp13 had specificity for Z-prone sequences. Indeed, the recent Chen at al structures detail
85 the binding of Nsp13 to an alternating pyrimidine/purine sequence CAUGU substrate (*20*). Recognition
86 of Z-RNA by Nsp13leading to strand separation would then prevent binding by ZBP1 and activation of
87 RIPKs.

88 Binding to Z-RNA would also lead to the exposure of the Nsp13 RHIM, which is on the other side of the
89 1B domain (Figure 3C). Once the RHIM is sprung lose, it is free to contact ZBP1, RIPK1 and RIPK3 and
90 modulate their activities.

91

92 **Disease Implications .**At this stage, we only know of the possibilities and hope that the above
93 hypotheses will provide a framework for further investigation of Nsp13 effects on necroptosis in
94 properly qualified BSL3 laboratories. If it is found that pathogenic CoVs do indeed produce Z-RNA (as the
95 ZH3 algorithm predicts) and that the RHIM in Nsp13 regulates ZBP1, RIPK3 and/or RIPK1 signaling, then
96 we suggest that effects will depend on the stage of viral infection. At early stages, suppressing Z-RNA
97 formation and necroptosis would enable increased viral replication. Here flipons may act as a sink for
98 the topological stress arising within the dsRNA formed by template switching during viral transcription
99 and replication {Chen, 2020 #1552}.The recognition of Z-RNA by Nsp3 then blocks the viral polymerase
100 until the two RNA strands are fully separated by the classical Nsp13 helicase activity. At later stages of
101 infection, the dsRNA tangles formed as defective viral genomes accumulate may create Nsp13 clusters,
102 leading to the formation of RHIM filaments. The filaments are amyloid like structures capable of
103 activating the RIP necrosome (*21*) to cause cell death, as recently reported in SARS2 *in vitro* and *in vivo*
104 models.

105 **Coronavirus and Human Pathology.** From our analysis, it seems that the RHIM domain is a requirement
106 for CoV to cause severe pathology in humans. Acquiring a RHIM may be necessary for the successful
107 jump the virus made from its natural host to humans, although not a sufficient one as other mutations
108 that enable engagement of human cell surface receptors are required. The mutations that lead to loss of
109 the strong Z-prone sequence present in SARS1 may have favored spread of SARS2 (Figure 2A and 2B).
110 The potential combination of the RHIM domain with the Z-flipon helicase to modulate the cell death
111 pathways so far appears unique among human viruses and may account for its extreme virulence.

112 The high frequency in humans of protein variants in the necroptotic pathway that affect its function (*22*)
113 leave some individuals more vulnerable to SARS2 induced cell death. While Nsp13 inhibitors that target
114 the ATP-binding site are likely to be effective in early infection by inhibiting the classical helicase

115 function, they are unlikely to be effective against the ATP-independent Z-flipon helicase clusters that

116 potentially produce severe pathology at later stages of the disease.

117 **Methods**

118 The Z-forming potential of the Coronavirus sequences were assessed using ZHUNT3 using the following
119 command line "zhunt3nt 12 6 24" {Schroth, 1992 #1415}. The structures PDB files 7NIO (*18*), 6XEZ (*19*)
120 and 7RDX (*20*) wered visualize using the open access NGLViewer {Rose, 2015 #1541}. Amino acid
121 alignments and the Phylogram were prepared using MUSCLE {Edgar, 2004 #1557} using the following
122 NCBI sequences for MERS (NC_019843.3), SARS (NC_004718.3), SARS2 (NC_045512.2), Hum_CoV-229E
123 (KU291448.1), CoV-OC43 (MW587042.1), CoV_NL63 (MG428706.1), CoV-HKU1 (NC_006577), Bat
124 RaTG13 (MN996532), Civet_PC4-227 (AY613950.1) and Llama (MN507638).

125 **Author Contributions.** AH is an international supervisor for the HSE Bioinformatics Laboratory run by
126 MP where AS is a student who identified the strong Z-forming segment in SARS1. AH conceptualized and
127 wrote the paper with edits from MP.

128 **Acknowledgements**. We that Sid Balachandran for suggesting the possibility of RHIM domains in SARS2
129 and for helpful edits to the manuscript.

130 **Conflict of Interest**. None declared by the authors

References

1. L. Galluzzi *et al.*, Molecular mechanisms of cell death: recommendations of the Nomenclature Committee on Cell Death 2018. *Cell Death Differ* **25**, 486-541 (2018).
2. A. Herbert, A Genetic Instruction Code Based on DNA Conformation. *Trends Genet* **35**, 887–890 (2019).
3. A. Herbert, Z-DNA and Z-RNA in human disease. *Communications Biology* **2**, 7 (2019).
4. P. S. Ho, Thermogenomics: thermodynamic-based approaches to genomic analyses of DNA structure. *Methods* **47**, 159-167 (2009).
5. P. J. Nichols *et al.*, Recognition of non-CpG repeats in Alu and ribosomal RNAs by the Z-RNA binding domain of ADAR1 induces A-Z junctions. *Nat Commun* **12**, (2021).
6. A. Herbert *et al.*, A Z-DNA binding domain present in the human editing enzyme, double-stranded RNA adenosine deaminase. *Proc Natl Acad Sci U S A* **94**, 8421-8426 (1997).
7. A. Athanasiadis, Zα-domains: At the intersection between RNA editing and innate immunity. *Seminars in Cell & Developmental Biology* **23**, 275-280 (2012).
8. H. Guo *et al.*, Herpes simplex virus suppresses necroptosis in human cells. *Cell Host Microbe* **17**, 243-251 (2015).
9. Z. Huang *et al.*, RIP1/RIP3 binding to HSV-1 ICP6 initiates necroptosis to restrict virus propagation in mice. *Cell Host Microbe* **17**, 229-242 (2015).
10. X. Sun, J. Yin, M. A. Starovasnik, W. J. Fairbrother, V. M. Dixit, Identification of a novel homotypic interaction motif required for the phosphorylation of receptor-interacting protein (RIP) by RIP3. *J Biol Chem* **277**, 9505-9511 (2002).
11. F. K. Chan, N. F. Luz, K. Moriwaki, Programmed necrosis in the cross talk of cell death and inflammation. *Annu Rev Immunol* **33**, 79-106 (2015).
12. T. Zhang *et al.*, Influenza Virus Z-RNAs Induce ZBP1-Mediated Necroptosis. *Cell* **180**, 1115-1129 (2020).
13. S. Balachandran, E. S. Mocarski, Viral Z-RNA triggers ZBP1-dependent cell death. *Current Opinion in Virology* **51**, 134-140 (2021).
14. H. Koehler *et al.*, Vaccinia virus E3 prevents sensing of Z-RNA to block ZBP1-dependent necroptosis. *Cell Host Microbe*, (2021).
15. E. S. Mocarski, W. J. Kaiser, D. Livingston-Rosanoff, J. W. Upton, L. P. Daley-Bauer, True grit: programmed necrosis in antiviral host defense, inflammation, and immunogenicity. *J Immunol* **192**, 2019-2026 (2014).
16. T. S. Bayer, L. N. Booth, S. M. Knudsen, A. D. Ellington, Arginine-rich motifs present multiple interfaces for specific binding by RNA. *Rna* **11**, 1848-1857 (2005).
17. J.-L. E. P. H. Darlix *et al.*, Mechanism of Nucleic Acid Unwinding by SARS-CoV Helicase. *PLoS One* **7**, (2012).
18. J. A. Newman *et al.*, Structure, mechanism and crystallographic fragment screening of the SARS-CoV-2 NSP13 helicase. *Nat Commun* **12**, 4848 (2021).
19. J. Chen *et al.*, Structural Basis for Helicase-Polymerase Coupling in the SARS-CoV-2 Replication-Transcription Complex. *Cell* **182**, 1560-1573 e1513 (2020).
20. J. Chen *et al.*, Ensemble cryo-electron microscopy reveals conformational states of the nsp13 helicase in the SARS-CoV-2 helicase replication-transcription complex. *bioRxiv*, 2021.2011.2010.468168 (2021).
21. M. Mompean *et al.*, The Structure of the Necrosome RIPK1-RIPK3 Core, a Human Hetero-Amyloid Signaling Complex. *Cell* **173**, 1244-1253 e1210 (2018).

22. S. N. Palmer, S. Chappidi, C. Pinkham, D. C. Hancks, T. Leitner, Evolutionary Profile for (Host and Viral) MLKL Indicates Its Activities as a Battlefront for Extensive Counteradaptation. *Mol Biol Evol* **38**, 5405-5422 (2021).
23. R. Lorenz *et al.*, ViennaRNA Package 2.0. *Algorithms Mol Biol* **6**, 26 (2011).
24. R. C. Edgar, MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* **32**, 1792-1797 (2004).
25. A. S. Rose, P. W. Hildebrand, NGL Viewer: a web application for molecular visualization. *Nucleic Acids Res* **43**, W576-579 (2015).
26. T. Schwartz, M. A. Rould, K. Lowenhaupt, A. Herbert, A. Rich, Crystal structure of the Zα domain of the human editing enzyme ADAR1 bound to left-handed Z-DNA. *Science* **284**, 1841-1845 (1999).

Figure 1. Coronavirus Signatures based on Z-flipons **A.** The signatures for each human pathogenic virus are unique. **B.** The signature match those for the presumed natural hosts for these viruses. **C.** The signature differs from other known coronaviruses D. The viral signature also varies with the host animal infected. The signature is derived using the ZHUNT3 (ZH3) program (*4*) that scores the propensity of

sequences to flip to the left-handed helical Z-conformation. Sequences with a score about 700 are likely to adopt the Z-conformation under physiological conditions.

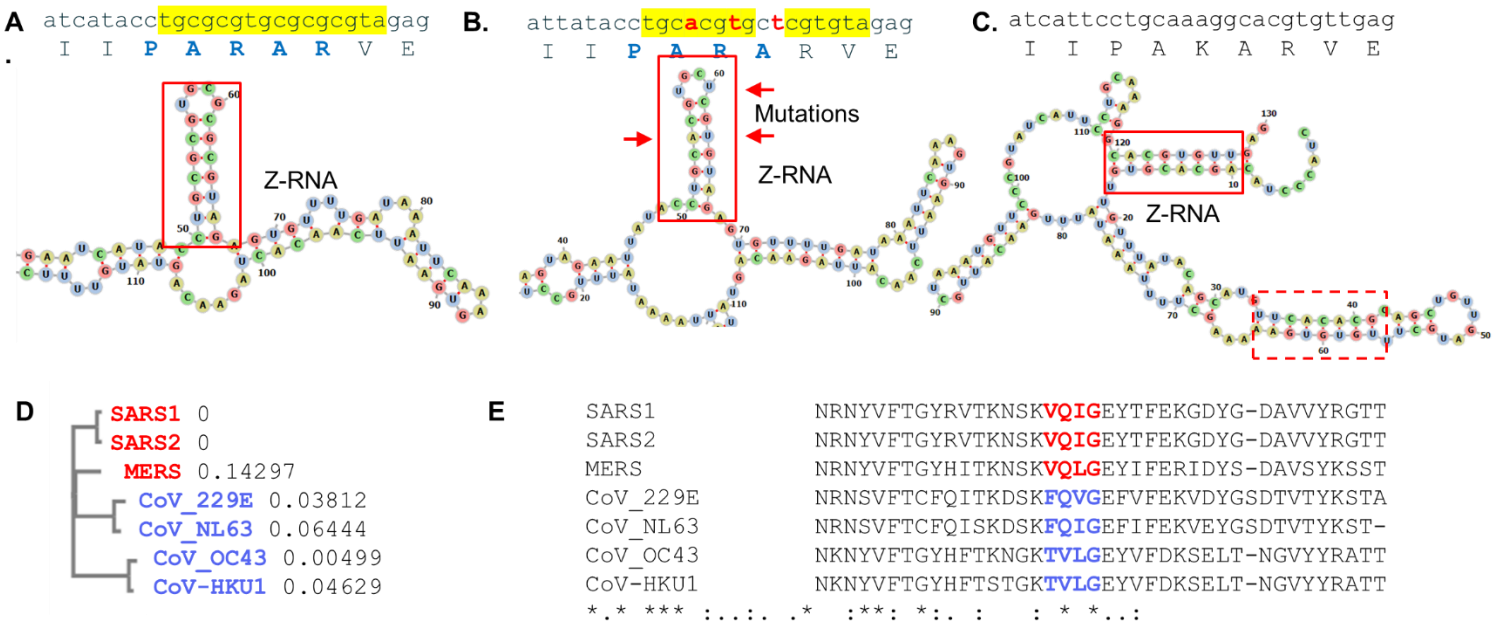

Figure 2 Nsp13 and Z-dependent necroptosis **A.** The high scoring ZH3 peak in SARS (NC_004718.3) maps to Nsp13 and encodes an arginine rich peptide with potential to bind to Z-RNA. Potential Z-RNA forming sequences are highlighted in yellow and form the Z-RNA stem within the red box identified using RNAfold (*23*). **B.** In SARS2 (NC_045512.2), three non-synonymous mutations conserve the Z-RNA stem and the peptide sequence, but diminish the Z-forming potential **C**. The equivalent Nsp13 region in MERS has a different peptide sequence. The Z-RNA forming element is encoded immediately after this sequence block, with another Z-RNA element close-by (in dashed box) **D.** Phylogram showing the evolutionary distance between coronavirus strains with high (in red) and low (in blue) human pathogenicity. **E**. The sequence aligner MUSCLE (*24*) reveals that highly pathogenic coronaviruses have a conserved RHIM domain in NSP13 (in red)

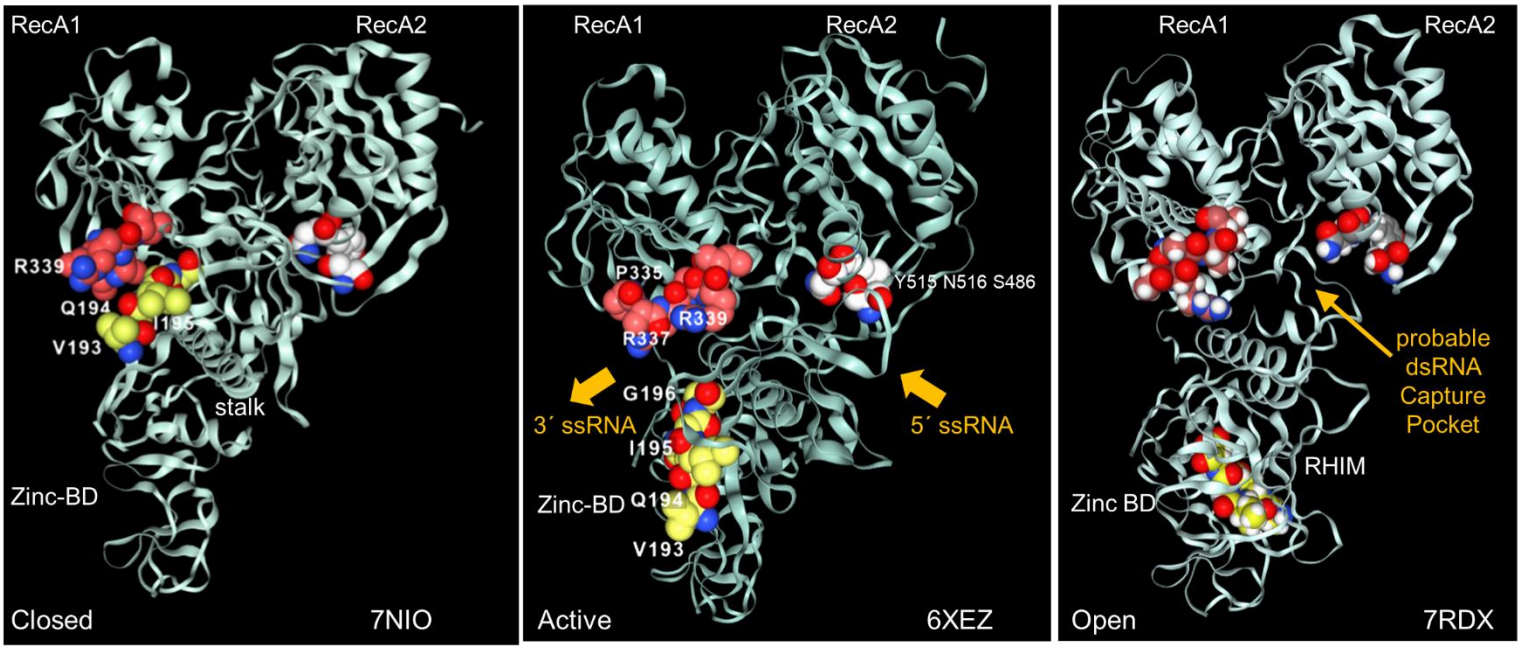

Figure 3 The interaction between the Nsp13 RHIM (residues 193-6) and tPARAR (residues 335-9) are conformation dependent. **A.** the ATP-free apo form of Nsp13, the RHIM (yellow space fill carbons) contacts PARAR339 (crimson space fill carbons) **B.** In the active state, RHIM and PARAR separate, opening the 5´ end of the single-strand RNA channel, that has the 3´end marked by N516 (white space fill carbons). **C**. The open complex in which the RHIM separates from the RecA2 domains to create a cavity that is large enough to accommodate dsRNA. This opening is associated with a rotation of the Zinc binding domain (ZincBD) relative to N516. PARAR also rotates, changing the position of R337 and R339. The structures are from PDB files 7NIO (*18*), 6XEZ (*19*) and 7RDX (*20*) as labeled, with images rendered using the NGLViewer (*25*).

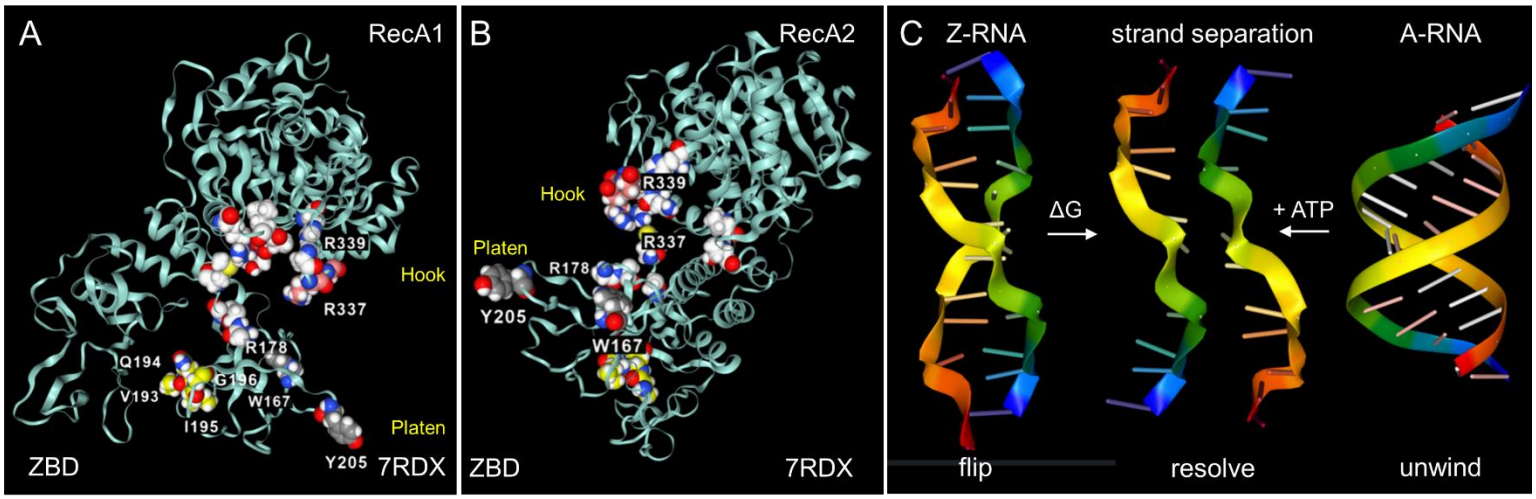

Figure 4 Nsp13 opens to expose a hook and platen. **A** The arginine rich PARAR hook (crimson carbons) and the tyrosine (Y205, grey carbons) platen create a surface for docking to dsRNA. Tryptophan (W168)(blue carbons) has the potential to orientate a to create a Z-RNA specific recognition element like that present in the Zα domain (*26*). The RHIM domain (yellow carbons) is free to engage other RHIM proteins. The space fill with white carbons show the residues identified as making base-specific contacts with single-stranded RNA in the active conformation shown in Figure 3B (residues 178, 179, 230, 233 in domain 1B, residues 311, 335,361, 362:E,363, 390, 408, 410:E in RecA1). **C.** In addition to the ATP-dependent helicase activity, Nsp13 has the potential to capture single stranded RNA when Z-RNA flips to and from A-RNA. The higher energy Z-RNA powers the strand-separation, providing the ΔG needed to fuel Z-flipon helicase activity.