Proc Indian Natn Sci Acad, **70, A,** No. 4, July 2004, pp. 597–604

NMR STUDIES TOWARDS UNDERSTANDING THE ACTIVATION MECHANISM OF dsRNA DEPENDENT PROTEIN KINASE, PKR

RAVI P BARNWAL, SWATI GIRDHANI AND SAMBASIVARAO NANDURI*

Department of Chemical Sciences, Tata Institute of Fundamental Research, Homibhabha Road, Colaba, Mumbai-400 005 (India)

(Received 14 February 2003 ; Accepted 18 March 2003)

The double stranded RNA activated protein kinase, PKR, is an interferon (IFN) regulated host defense essential for higher vertebrates to suppress the spread of different viruses. PKR is a key regulator of protein synthesis and gene transcription. In response to specific signals, PKR also mediates phosphorylation of important transcriptional regulators, including the inhibitor of NF-kB, IkB, and IRF-1. A key step in the cellular processes mediated by PKR is its activation by double-stranded viral or cellular RNA. PKR has been suggested to undergo conformational change, dimerize and auto-phosphorylate before it activates its down-stream regulators. This review focuses on the recent developments in understanding the dsRNA mediated activation mechanism of PKR.

Key Words: PKR; Protein Kinase; dsRBD; dsRNA Binding Domain; dsRNA; NMR Structure; Interferon; Anti-Viral Protein

Introduction

During the viral infections, cellular defense system produces secreted proteins, called interferons. The interferons in turn produce a number of other antiviral proteins[1]. Protein Kinase R, PKR, is one of the key proteins induced by the interferons[2,3]. *In vitro*, PKR is activated by binding to RNA duplexes and *in vivo* the enzyme is believed to be activated by viral double-stranded RNA or viral replicative intermediates comprising dsRNA (Fig. 1). The activation of PKR by viral dsRNA results in the phosphorylation of the eukaryotic initiation factor (eIF2) on Ser51 of its α-subunit. The phosphorylated eIF-2α is an inhibitor of the protein factor, eIF-2B, whose role is to recycle the inactive eIF-2-GDP complex into its activated form eIF-2-GTP, a complex necessary to begin a new round of initiation[4]. The net effect of the phosphorylation of eIF-2α is the inhibition of protein synthesis and consequently inhibition of virus production. In addition to its role as a regulator of translation, PKR also acts as a signaling molecule in dsRNA- and cytokine-mediated signal transduction pathways[5]. Transcription factors, including nuclear factor-kB (NF-kB), Interferon Regulatory Factor-1 (IRF-1), Signal transducer and activator of transcription 1 (STAT1) and STAT3 have been shown to be regulated by PKR-mediated signaling[6-10]. PKR signals in the stress-activated protein kinase (p38) and c-Jun NH2-terminal kinase (JNK) pathways in response to extracellular stimuli[11]. PKR has also been implicated in the control of splicing of the tumor necrosis factor-α (TNF-α) mRNA in a kinase-dependent pathway[12] .

Protein Kinase, PKR

Human PKR is a 551 amino acid protein containing mainly two domains: the N-terminal dsRNA binding regulatory domain and C-terminal catalytic or kinase domain (Fig. 2). The kinase domain shares high homology with serine/threonine kinases in 11 of the 12 kinase subdomains[13]. It has a motif in subdomain VI (aa 414-419) characteristic of serine/threonine kinases, and has been predicted to play functional role on the basis of several crystal structures of different kinases. Also the Lys residue in subdomain II is found to be essential for enzymatic activity. Mutation of this lysine in PKR to either arginine or proline results in kinase inactivity[14, 15]. Activation of PKR involves the phosphorylation of several serine and threonine residues between the dsRNA binding and catalytic domains. Mutations of Thr_{258}, Thr_{255} and Ser_{242} attenuate the kinase activity[16], indicating an autoregulatory function for this region. Cytoplasmic PKR is multiple-phosphorylated, while the nuclear PKR is homogeneous[17], suggesting differences in the functional roles in the different forms of PKR. The N-terminal

*Author for Correspondence :
E-mail : nanduri@tifr.res.in;
Phone : (022) 2280-4545 Ext. 2251; *Fax :* (022) 2280-4610

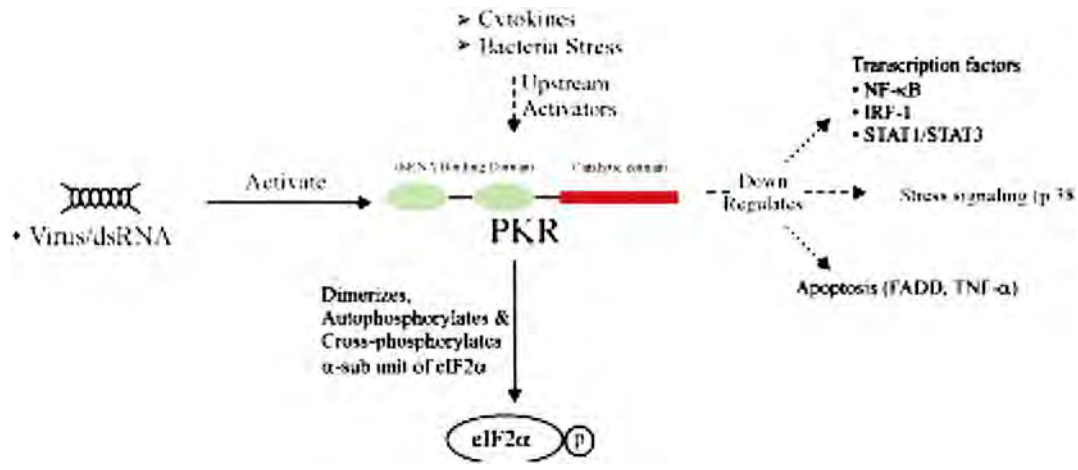

Fig. 1 Role of PKR in different signaling pathways. Binding of viral RNA activates PKR. Cytokines or the Stress signaling protein kinases are the upstream PKR activators. The downstream regulators are phosphorylated by the activated PKR.

dsRNA binding domain has two dsRNA binding motifs, each consisting of about 65 residues. Currently there are about 100 such motifs found in nine protein families from diverse sources, including viruses, bacteria and vertebrates[18]. Most of the dsRNA binding protein families contain multiple copies of dsRBMs. Four families (RHA, ADAR2, NF90, PKR) contain two dsRBMs, while ADAR1 has three, staufen has either five (Drosophila) or four (human) copies of the motif. TRBP family has either two or three copies. Rnase III and E3L families contain only one copy of this motif. These motifs either bind to ssRNA (eg. RNP, KH binding domains) or dsRNA (eg. PKR, Drosophila staufen dsRBDs).

Structure of dsRNA Binding Domain of PKR

The NMR structures of E. Coli. Rnase dsRBM III[19] and Drosophila Staufen dsRBM III[20] have a common tertiary fold of α–β–β–β–α. The structure of dsRBD of PKR exhibits a dumb bell shape, with two dsRBMs joined by a 22 residue linker (Fig. 3). Each of the two dsRBMs has folding topology, α–β–β–β–α, similar to the earlier reported Rnase and staufen dsRBMs. The two α-helices lying on one side of the three-stranded anti-parallel β-sheet. The two α-helices in dsRBMs are amphipathic and the residues from their hydrophobic faces pack onto the side of the β-sheet to form a hydrophobic core. This central hydrophobic core formed by the residues like F10, F43, V45, I47, A71 and V72 for dsRBM1 and Y101, Y133, C135, M137 and A161 for dsRBM2, stabilizes the conserved fold of dsRBM. In spite of the differences in the functionality, the two motifs of dsRBD of PKR have the same fold. This could be because of their 50% sequence homology. The long and highly flexible linker region well separates the two motifs, so as they do not have any contact NOEs. Point mutations of the conserved RNA binding residues, such as R39, F41, S59, K60, K61 and K64 in dsRBM1 have severely decreased or completely abolished the dsRNA binding[21-23]. The residues in dsRBM1 form a positively charged binding surface, residing in the secondary structure elements such as the beginning of β3, the loop between β3 and α2 and N-terminal part of α2, while F41 is partially involved in the edge of the hydrophobic core. Also the mutations of the conserved residues F131, K150 and K154 in dsRBM2 also abolishes the dsRNA binding of PKR[24], suggesting the existence of a dsRNA binding site similar to dsRBM1. The structure of dsRBD explains other deletional and mutational studies. The deletions of the regions 1-24, 39-50 and 58-69 all remove important structural elements like α-helix or β-strand. The point mutations of the α1-helix end-residues R18 and Q19 destabilizes the helical structure. G57 makes tight turn between β3 and α2, and hence its mutation into Ala introduces a steric clash. A68D, L75A and A158D mutations are defective because they all destabilize the hydrophobic cores of dsRBM1 and dsRBM2 respectively.

DsRBM – dsRNA Complex Structures

Binding of dsRBD to a sequence independent dsRNA is a first step in the activation mechanism of PKR. Subsequently PKR undergoes conformational rearrangement and auto-phosphorylation, which leads to the cross-phosphorylation of down stream substrates

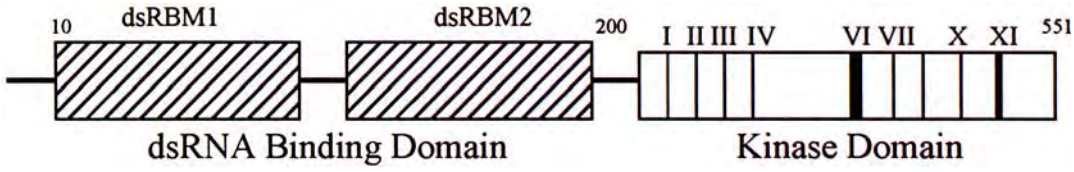

Fig. 2 Schematic representation of structure and functional motifs of human PKR protein.

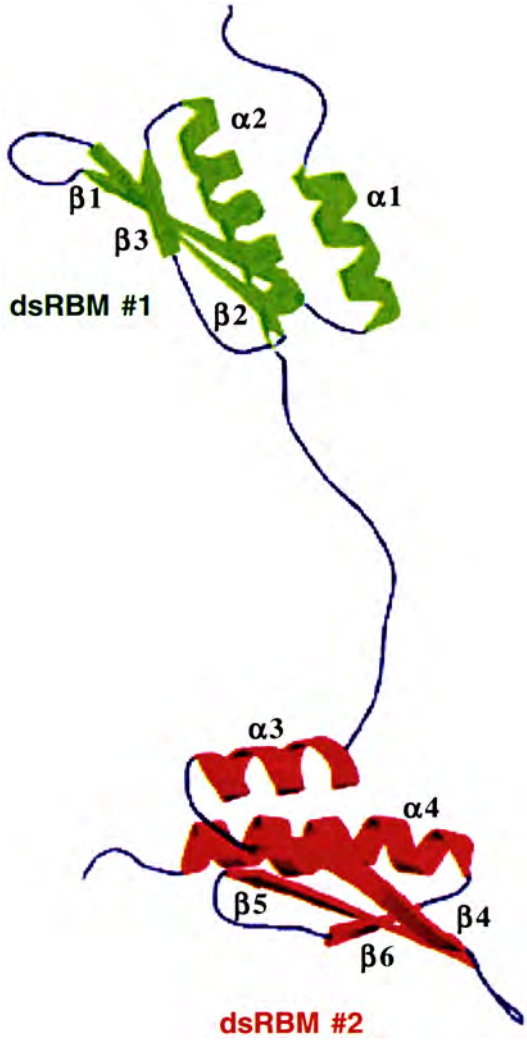

Fig. 3 Ribbon diagram of dsRBD of PKR containing dsRBDM1 (green) and dsRBM2 (red) linked by a 22 amino acid loop[44]

in PKR signaling pathway. Thus, it is critical to understand the dsRNA recognition by dsRNA binding domain of PKR. Even though the details of dsRBD structure and the identification of its RNA binding sites could provide preliminary understanding of dsRNA regulated PKR activation, the complete details of the dsRNA recognition mechanism will be known if a complex structure involving this multi-motif dsRBD is available. There are couple of single motif dsRBM-dsRNA complex structures that have been reported recently. The crystal structure of the second dsRBM of Xenopus laevis RNA binding protein A (Xlrbpa-2) in complex with a 10mer palindrome (self-complementary) dsRNA[25] and the NMR structure of Drosophilae staufen dsRBM3-dsRNA complex[26], could throw some initial light on the interaction pattern in dsRBD – dsRNA complexes. In the crystal structure of Xlrbpa-2 and dsRNA complex, the dsRBM interacts with one and half molecules of the dsRNA. Also, the dsRBM interacts with two successive minor grooves and across the intervening major groove on one face of the dsRNA helix, covering a total of 16bp. This is similar to the observations made by Bevilacqua and Cech[27], from their extensive biochemical experiments in the case of dsRBD of PKR which they found would require a minimum of 16bp of dsRNA for the stable binding of dsRBD and dsRNA. There are three dsRNA molecules in the asymmetric unit of the cell, which form nearly ideal A-form helices. Of the three regions in the protein, where the interaction with dsRNA was observed (Fig. 4), the two regions (the N-terminal α-helix and the β1-β2 loop regions of Xlrbpa-2) involve the direct and/or water-mediated interactions with the 2'-OH groups of the minor groove of dsRNA. In the third region, the C-terminal α-helix and the preceding loop have contacts with the non-bridging oxygens of the phosphodiester backbone across the major groove of the RNA. The total surface area of interaction is 1680 Å^2. Ramos *et al.*[26] reported an NMR solution structure of Drosophilae staufen dsRBD3 bound to uninterrupted 12bp (unlike in case of Xlrbpa-2) of A-form dsRNA stem-loop, in which the conserved residues within the β1-β2 loop and the β3-α2 loop interact with the minor groove and the phosphodiester backbone across the adjacent major groove of dsRNA. In addition, α1 helix of dsRBD3 has interactions with the single stranded loop that caps the RNA helix. The area buried in this complex is 1450 Å^2, 12% of the total surface area. ^{15}N NMR relaxation studies show that the loops L2 and L4 of dsRBD3 of staufen retain significant conformational flexibility in its complex with

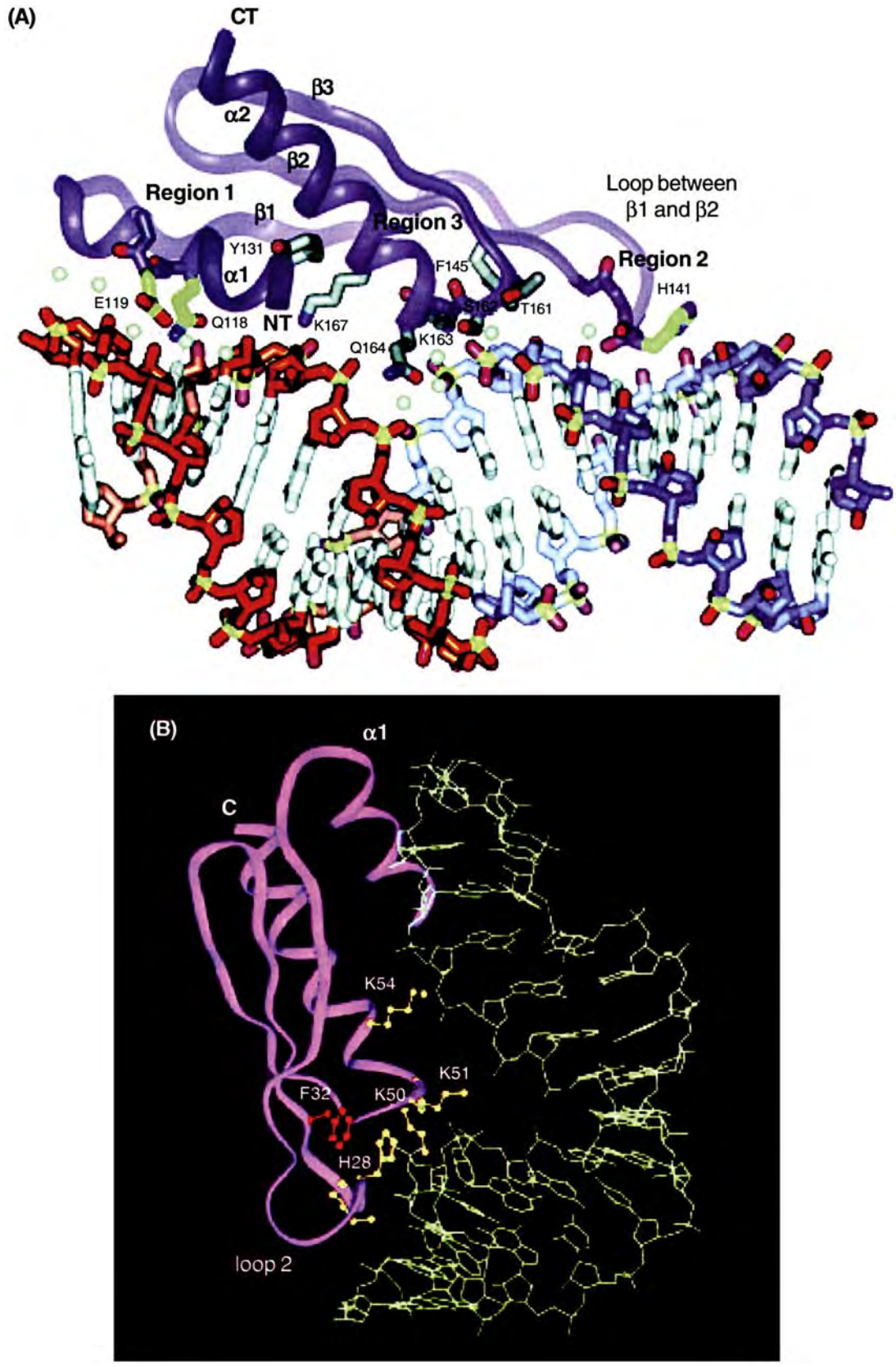

Fig. 4 Over view of the interactions between (A) Xenopus laevis RNA dsRBM2 and dsRNA[25] and (B) Drosophila Staufen protein dsRBM3 and dsRNA[26]. The α1-helix, Loop2-Loop3 and Loop4-α2helix regions in the protein interact with the two minor grooves and the intervening major groove of dsRNA.

dsRNA and the C-terminal tailor residues in the α2 region show complete disorder in the complex. The high flexibility of the interface region could be due to the indirect and non-specific interactions. The complex structure does not show any significant differences, when compared to free-form dsRBD and free-form dsRNA. The only minor changes observed are (a) the rotation of loop2 of dsRBD towards RNA and (b) a kink in dsRNA at the stem-loop junction resulting a bend for a simultaneous interaction of α1 helix and the tetraloop and the contact between loop 2 and RNA minor groove. Thus both Xlrbpa and staufen dsRBD-dsRNA complex structures have the same three regions of contact. In both the structures loop2 regions interact with the dsRNA minor groove and loop4 interacts with phosphodiester backbone across the major groove, while the α1-helix interacts with the minor groove of second RNA duplex in Xlrbpa-2 crystal structure, and it interacts with the tetraloop of dsRNA in NMR structure of staufen dsRBD-dsRNA complex.

Role of Multiple Motifs in dsRNA Recognition in PKR

The common interactions involving α1-helix, loop2 and loop4 in both NMR and X-ray structures of the staufen dsRBM3-dsRNA and Xlrbpa2-dsRNA complexes (Fig. 4) suggest that these three interactions are the basis for the dsRBD-dsRNA recognition mechanism. However, from such single motif complex structures, it is not clearly understood the necessity of the multi-motif dsRBDs in recognizing dsRNAs. In the case of dsRBD of PKR, it has been shown that the two motifs have different binding affinities to dsRNA, the first one has about 100 fold higher affinity than the second motif[28-31]. Mutant PKR proteins which contain two tandem copies of dsRBM1 bind dsRNA with the same efficiency as wild type protein, but the tandem repeats of dsRBM2 are severely deficient in binding[28], indicating that the two regions, while nearly identical in sequence, are not equivalent. Also swapping dsRBM1 and dsRBM2 or altering the distance between them had no effect on binding[28]. A recent reminiscent investigation on the RNA-specific adenosine deaminase (ADAR1), an RNA editing enzyme that converts adenosine to inosine, reveals functional selectivity of double stranded RNA binding domain by showing dramatic reduction in editing activity with natural substrates, when two of its motifs are replaced by the dsRBMs of PKR[32]. Spanggord and Beal[33] findings from their EDTA-Fe affinity cleavage experiments indicate that the activating dsRNAs bind to both motifs of dsRBD of PKR while the inhibiting RNAs bind to dsRBM1. They found that dsRBM1 binds near the loop region of the activating HIV1 TAR RNA, just like in NMR and crystal structures of the complexes, while the dsRBM2 has binding site near the 3'and 5' end regions of RNA (Fig. 5). Similarly the activating RNA ligand discovered during SELEX experiments (aptamer)[34] has two dsRBM1 binding sites while only one of them simultaneously binds with the dsRBM2 motif. The viral inhibiting VA1 RNA from Adenovirus has single major binding site for dsRBM1 or dsRBM2. The EBER1 RNA of Epstein-Barr virus (EBV), which prevent inhibition of translation by PKR *in vitro* has two binding sites with dsRBM1 while dsRBM2 does not have any binding site[35]. In order to investigate if the two motifs has a preferable conformation relative to each other in solution, either while interacting with an activating RNA or when forming a dimer, we having been using residual dipolar couplings in oriented media. We found some large NH dipolar couplings for some of the loop residues near the two domains, indicating that the two motifs may be having a preferable orientation with respect to the neighbourhood loop residues. The details of our results will be discussed elsewhere.

NMR Relaxation Studies on dsRBD of PKR

NMR relaxation experiments provide valuable information about the dynamics of proteins and their complexes. There are numerous instances in the literature where the dynamics pattern of the individual or a group of amino acids have been correlated to the function of the proteins. For example, (i) Drosophila calmodulin has been found from NMR relaxation

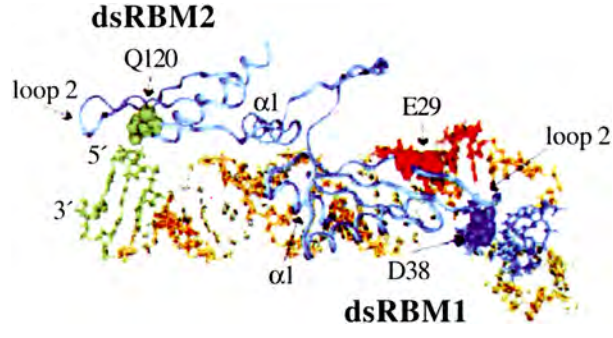

Fig. 5 Model of the dsRBD of PKR bound to TAR RNA. E29, D38 and Q120 are the EDTA.Fe attached positions in dsRBM1 and dsRBM2 of PKR and the positions of the two motifs are based on the cleavage pattern observed in the affinity cleavage experiments[33].

studies to have highly flexible central helix and this has been interpreted to facilitate the binding to target proteins[36], (ii) Feher and Cavanagh[37] have found that the amino acid regions in the B.subtilis response regulator Spo0F, which are critical for protein-protein interactions have been found to exhibit motions on the millisecond timescale, (iii) from the dynamics of the methyl groups at the interfaces of the Cy1 and Syp phosphatase SH2 domain complexes with phosphopeptides, Kay *et al.*[38] established a correlation between the binding energy and restriction of motions for the specific binding and (iv) Hosur and coworkers[39] have interpreted the variations in the motion in different locations of barnase and barnase-barstar complex binding interface with the regulation of the enzyme action. The NMR relaxation parameters such as T1, T2 and NOE are dependent on shape and the local and/or global motion of the systems under investigation. The backbone dynamics studies using ^{15}N NMR relaxation parameters T1, T2 and NOE, revealed interesting differences in the dynamics pattern of the two motifs of dsRBD of PKR[40]. Several residues in dsRNA binding regions of dsRBM1, as compared to dsRBM2, have been shown to exhibit millisecond-microsecond time scale motions (Fig. 6), indicating a strong correlation between the protein dynamics and the dsRNA binding to dsRBM. All residues in the three dsRNA binding regions of dsRBM1 of PKR show conformational exchange either with R_{ex} (>1) or line-broadening. Also the other residues that are having network of interactions in α1, β1 and α2, which pack against each other, have also exhibited R_{ex} values more than 1, indicating a concerted motion of the motif at millisecond-microsecond timescale. Such concerted motion is important since multiple conformational substrates of dsRBM may be required for binding to various dsRNAs, including the structured RNAs, that contain only short segments of uninterrupted Watson-Crick base pairing, such as HIV-Tar, adenovirus VAI RNA, EBV RNAs, EBER-2 and EBER-2 etc.[41]or RNAs with secondary structure defects such as a GA:AG mismatch and non-contiguous helix[34]. In contrast, majority of the residues in dsRBM2 undergo restricted picosecond-nanosecond motions with only a few residues (interestingly all of them clustered in dsRNA binding region) having R_{ex}. We have been further investigating to see if the packing of hydrophobic cores of the two motifs is different, using the ^{13}C methyl group spin-relaxation dynamics studies. The

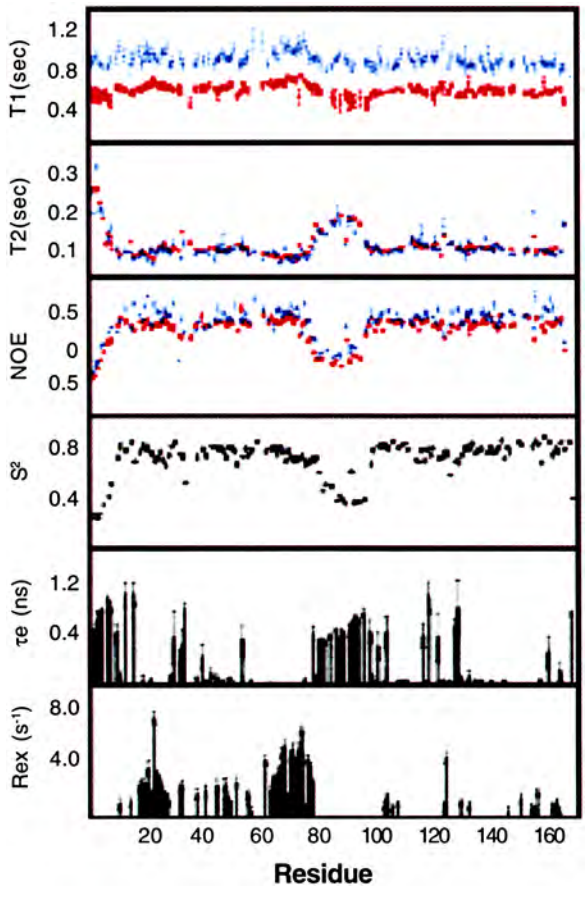

Fig. 6 ^{15}N relaxation parameters of dsRBD. The panel F exposes the differential distribution of R_{ex} values between dsRBM1 and dsRBM2 motifs. The Red and Blue points in panels A, B and C are the data from 500 and 600 MHz spectrometers respectively[40]

preliminary analysis of the data provided some interesting differences in the Methyl dynamics of the two motifs and the details of our data will be presented elsewhere.

DsRBM2-Kinase Domain Interaction and the Model for PKR Activation

Qin and co-workers[40] showed from their NMR titration experiments that the dsRBM2, but not dsRBM1, makes specific contacts with C-terminal kinase domain. The ^{1}H-^{15}N NMR heteronuclear correlation (HSQC) spectrum of the 1:1 mixture of the unlabeled MBP-kinase domain and the ^{15}N labeled dsRBD of PKR contain only the peaks from dsRBM1 motif and all the remaining peaks from the dsRBM2 residues have been found to disappear. This observation has been interpreted to some kind of engagement of dsRBM2 motif with the C-terminal kinase domain. The Surface plasmone resonance

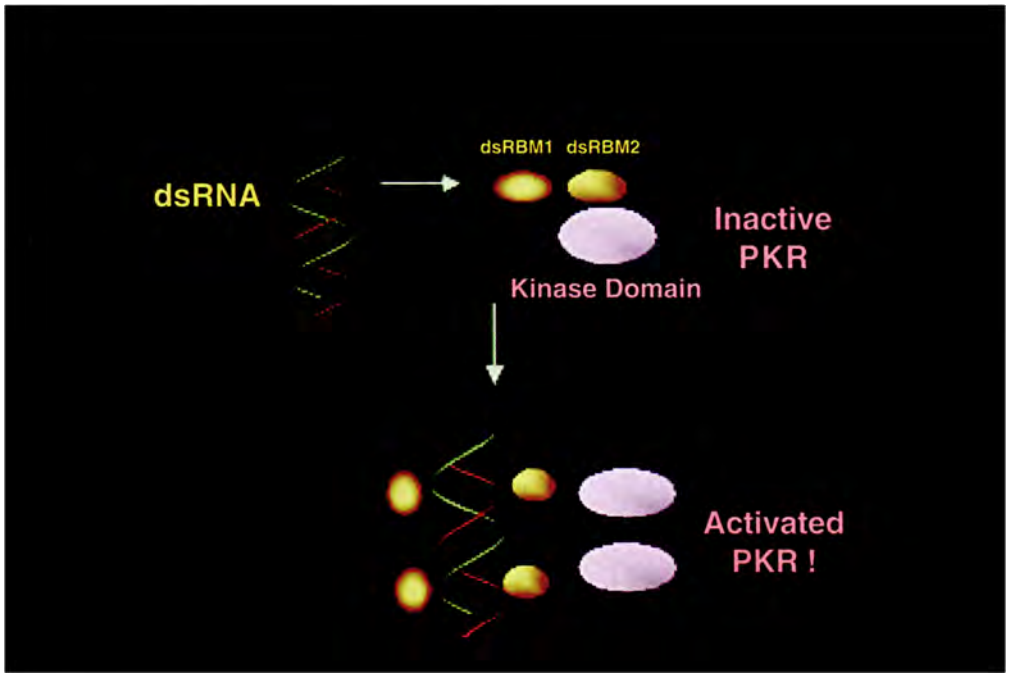

Fig. 7 A model for dsRNA-mediated PKR activation. DsRBM2 forms a closed conformation with the C-terminal kinase domain of in-active PKR. Any activating viral RNA first anchored to the free-dsRBM1, inducing a cooperative binding to dsRBM2, which then exposes the kinase domain. This releases the dimerization region (244-296) and the subsequent steps for activation such as dimerization, autophosphorylation will follow[40]

experiments have supported this notion, by showing a binding affinity in the range of ~500 nM, between the N- and C-terminal domains of PKR. On the other hand, titration of full-length PKR or the putative dimerization domain (244-296aa) into dsRBD did not result in any spectral change, indicating that dsRBM2 keeps the full-length PKR in closed or inactive conformation. There have been other instances in the literature to indicate the existence of such a possibility. They include: (i) deletion of entire dsRBD constitutively activates PKR[42], (ii) swapping of dsRBM1/dsRBM2 substantially increases the kinase activity[30], (iii) ATP becomes associated with PKR only after exposure to dsRNA[43] and (iv) the activating dsRNAs (like for example HIV1 TAR RNA) bind to both dsRBM1 and dsRBM2, unlike the inhibiting dsRNAs (VAI RNA from adenovirus) which bind only to dsRBM1[33]. Thus cooperative binding of dsRNA to dsRBM1 and dsRBM2 releases the restraint of dsRBM2 on the kinase domain, leading to an open active conformation (Fig. 7). The dsRBM1 while fluctuating in millisecond-microsecond motions serves as an anchor for dsRNA, thereby inducing a cooperative RNA binding for dsRBM2 to expose the kinase domain.

References

1 C E Samuel *Clin Microbiol Rev* **14** (2001) 778

2 E Meurs, K Chong, J Galabru, N S Thomas, I M Kerr, B R G Williams and A G Hovanessian *Cell* **62** (1990) 379

3 A E Koromilas, S Roy, G N Barber, M G Katze and N Sonenberg *Science* **257** (1992) 1685

4 P J Farrel, K Balkow, T Hunt, R J Jackson and H Trachsel *Cell* **11** (1977) 187

5 B R G Williams *Seminars Virology* **6** (1995) 191

6 A Kumar, J Haque, J Lacoste, J Hiscott and B R G Williams *Proc Natl Acad Sci USA* **91** (1994) 6288

7 A Kumar, Y L Yang, V Flati, S Der, S Kadereit, A Deb, J Haque, L Reis, C Weissmann and B R G Williams *EMBO Jour* **16** (1997) 406

8 A Maran, R K Maitra, A Kumar, B Dong, W Xiao, G Li, B R G Williams, P F Torrence and R H Silverman *Science* **265** (1994) 789

9 L Beretta, M Gabbay, R Berger, S M Hanash and N Sonenberg *Oncogene* **12** (1996) 1593

10 A H Wong, J E Durbin, S Li, T E Dever, T Decker and A E Koromilas *Jour Biol Chem* **276** (2001)13727

11 M S Iordanov, J M Paranjape, A Zhou, J Wong, B R Williams, E F Meurs, R H Silverman and B E Magun *Mol Cell Biol* **20** (2000) 617

12 S D Der, Y L Yang, C Weissmann and B R G Williams *Proc Natl Acad Sci USA* **94** (1997) 3279

13 S K Hanks, A M Quinn and T Hunter Science **241**(1988) 42

14 K L Chang, L Feng, K Schappert, E Meurs, T F Donahue, A G Hovanessian and B R G Williams *EMBO Jour* **11** (1992) 1553

15 M G Katze, M Wambach, M L Wong, M Garfinkel, E Meurs, K Chong, B R Williams, A G Hovanessian and G N Barber *Mol Cell Biol* **11** (1991) 5497

16 D R Taylor, S B Lee, P R Romano, D R Marshak, A G Himmebusch, M Esteban and M B Mathews *Mol Cell Biol* **16** (1996) 6295

17 I Jeffrey, S Kadereit, T Metzget, M Bachmann, M Schwemmle, A G Hovanessian M Clemens *Exptl Cell Res* **218** (1995) 17

18 I Fierro-Monti and M B Mathews *Trends in Biol Sci (TIBS)* **25** (2000) 241

19 A Kharrat, M J Macias, T J Gibson, M Nilges and A Pastore *EMBO Jour* **14** (1995) 3572

20 M Bycroft, S Grunert, A G Murzin, M Proctor and D Johnson *EMBO Jour* **14** (1995) 3563

21 P A Clarke and M B Mathews *RNA* **1** (1995) 7

22 N A McMillan, B W Carpick, B Hollis, W M Toone, M Zamanian-Daryoush and B R G Williams *Jour Biol Chem* **270** (1995) 2601

23 R C Patel, P Stanton, N M McMillan, B R G Williams and G C Sen *Proc Natl Acad Sci USA* **92** (1995) 8283

24 R C Patel, P Stanton and G C Sen *J Biol Chem* **271** (1996) 25657

25 J M Ryter and S C Schultz *EMBO Jour* **17** (1998) 7505

26 A Ramos, S Grunert, J Adams, D R Micklem, M R Proctor, S Freund, M Bycroft, D Johnston and G Varani, *EMBO Jour* **19** (2000) 997

27 P C Bevilacqua and T R Cech *Biochemistry* **35**(1996) 9983

28 S R Green and M B Mathews *Genes Dev* **6** (1992) 2478

29 S J McCormack, L G Ortega, J P Doohan and C E Samuel *Virology* **198** (1994) 92

30 P R Romano, S R Green, G N Barber, M B Mathews and A G Hinnebusch *Mol Cell Biol* **15** (1995) 365

31 C Schmedt, S R Green, L Manche, D R Taylor, Y Ma and M B Mathews *Jour Mol Biol* **249** (1995) 29

32 Y Liu, M Lei and C E Samuel *Proc Natl Acad Sci USA* **97** (2000) 12541

33 R J Spanggord, M Vuyisich and P A Beal *Biochemistry* **41** (2002) 4511

34 P C Bevilacqua, C X George, C E Samuel and T R Cech *Biochemistry* **37** (1998) 6303

35 M Vuyisich, R J Spanggord and P A Beal *EMBO Reports* **3** (2002) 622

36 G Barbato, M Ikura, L E Kay, R W Pastor and A Bax *Biochemisty* **31** (1992) 5269

37 V A Feher and J Cavanagh *Nature* **400** (1999) 289

38 L E Kay, D R Muhandiram, G Wolf, S E Shoelson and J D Forman-Kay *Nat Struct Biol* **5** (1998) 156

39 S C Sahu, A K Bhuyan, J B Udgaonkar and R V Hosur *Jour Biomol NMR* **18**(2) 2000 107

40 S Nanduri, F Rahman, B R Williams and J Qin *EMBO Jour* **19** (2000) 5567

41 M J Clemens, K G Laing, I W Jeffrey, A Schofield, T V Shartp, A Elia, V Matys, M C James and V J Tilleray *Biochimie* **76** (1994) 770

42 S Wu and R J Kaufman *J Biol Chem* **272** (1997) 1291

43 J Galabru and A Hovanessian *J Biol Chem* **262** (1987) 15538

44 S Nanduri, B W Carpick, Y Yang and B R Williams J Qin *EMBO Jour* (1998) 175458