Hepatitis E virus (HEV) is a positive-sense RNA virus and member of the genus Orthohepevirus in the family Hepeviridae. Although HEV RNA-dependent RNA polymerase (HEV-RdRp) plays an important role in the HEV life cycle, its template specificities are not completely understood. We expressed HEV-RdRp protein with His-tag in a bacterial system and analysed template specificities using different putative cis-regulatory elements in the HEV genome. The enzyme showed highest affinity for the 3' non-coding region (NCR), then for the 5'NCR and least for the putative subgenomic promoter (SgP). The enzyme could co-bind to 3'NCR and putative SgP templates together, as evident from the supershift in binding assay, indicating presence of different binding sites for these elements. Proteomic analysis revealed that the RNA elements share two common peptides for binding, while a third peptide, which is highly conserved across different HEV genotypes, is specific for 3'NCR. We propose that, during the early phases of replication, as negative sense antigenome copies accumulate at the replication site, they probably initiate promoter swapping from 3'NCR to SgP, to favour synthesis of subgenomic RNA and to prevent synthesis of genomic RNA. The conserved site for 3'NCR binding could be potential antiviral target and needs further evaluation.

INTRODUCTION

Hepatitis E virus (HEV) is a major causative agent of acute viral hepatitis in developing countries, particularly in Asia, Africa and Central America (Arankalle et al., 1995; Purcell et al., 2001). HEV is primarily transmitted via the faecal–oral route through contaminated water. Mammalian HEV strains form four distinct phylogenetic clusters/genotypes (HEV-1–4). Of these, HEV-1 and -2 strains are restricted to humans, while HEV-3 and HEV-4 are zoonotic and can infect humans, pigs and several other animals (Pavio et al., 2010; Meng et al., 2011; Arankalle et al., 2002, 2006).

HEV belongs to the genus Orthohepevirus in the family Hepeviridae and is a member of an alphavirus-like superfamily. HEV genome is a single-stranded, capped, polyadenylated positive-sense (PS) RNA of ~7.2 kb which contains three overlapping ORFs - ORF1, ORF2 and ORF3 (Tam et al., 1991). The genome is flanked by short 5' and 3' non-coding regions (NCRs). The 5'NCR is ~25–27 nt long and 3'NCR is ~65 nt long. The capsid protein is known to bind to a 76 nt stretch at the 5'end of the HEV genome indicating the importance of this sequence in encapsidation (Surjit et al., 2004). A 51 nt sequence from this stretch is highly conserved amongst different alphaviruses (Niesters et al., 1990). Interaction of HEV RNA-dependent RNA polymerase (HEV-RdRp) with 3'NCR and adjacent upstream sequence harbouring two stem–loop structures (SL1 and SL2) is documented (Agrawal et al., 2001). Poly-A tail and SL2 are also known to be essential during HEV replication (Emerson et al., 2001; Graff et al., 2005a). The junction region (JR) between ORF1 and ORF2, known as the putative subgenomic promoter (SGP), contains a highly conserved stem–loop structure (Cao et al., 2010; Huang et al., 2007). Silent mutations in this region completely eliminate ORF2 and ORF3 protein synthesis, and thereby prove its role as the regulatory element in the HEV genome (Graff et al., 2005b). HEV synthesizes a capped 2.2 kb bicistronic subgenomic RNA (sgRNA), which initiates at 5122 nt in HEV-1 (Graff et al., 2006). ORF2 and ORF3 proteins are translated from the same bicistronic sgRNA.
Based on the sequence homologies with other PS RNA viruses, HEV ORF1 protein is proposed to contain putative domains for methyltransferase, papain-like cysteine protease, RNA helicase, RdRp and domains of unknown function X and Y (Koonin et al., 1992). It is presumed that upon entry into the host cell, ORF1 is translated to generate HEV non-structural polyprotein; however, it is not yet clear whether it functions as a multidomain protein or is cleaved and processed into different functional units. The first step in HEV replication is synthesis of replicative intermediate RNA [negative-sense (NS) antigenome] from the genomic RNA (gRNA). The antigenome then serves as a template for synthesis of two classes of PS RNA molecules, gRNA and sgRNA (Huang et al., 2007; Graff et al., 2006). The sgRNA requires RdRp to initiate synthesis from internal promoter (SgP) within the antigenome. The recognition elements that determine template specificities of HEV-RdRp to ensure amplification of appropriate viral RNA species are yet not well characterized.

We cloned the RdRp encoding region, expressed the protein in a bacterial system and tested its enzymatic activity. In vitro binding assays were carried out to know the binding affinities of the protein for the putative cis-acting regulatory elements in the HEV genome/antigenome. HEV-RdRp showed highest binding affinity for 3’NCR. The enzyme showed cooperative binding with putative SgP and 3’NCR, which could be a novel mechanism for initiating sgRNA synthesis.

**RESULTS**

*Cloning, expression and purification of HEV-RdRp*

Both wild-type (T4GDD-RdRp) and mutant (T4GAA-RdRp) RdRp constructs expressed proteins of desired mass of 48 kDa. Since the majority of the protein was in insoluble form, denaturing buffer conditions were required for the solubilization and purification of the protein. The pooled purified protein fractions from Ni-affinity chromatography showed enrichment of the expected 48 kDa recombinant HEV-RdRp (T4GDD-RdRp) protein (Fig. 1a, d). Western blot analysis with anti-His antibodies further confirmed the expression and purification of recombinant HEV-RdRp proteins (Fig. 1b, e). High performance liquid chromatography (HPLC) gel filtration column fractions showed a peak corresponding to 48 kDa HEV-RdRp (Fig. 1c, f). Overall purity of the protein after gel filtration chromatography was about 90 %, as seen from the SDS-PAGE analysis. Inclusion of MgCl₂ in buffers improved the yield as well as the stability of the protein. An attempt to concentrate purified protein above 1 mg ml⁻¹ resulted in aggregates, and hence concentration of the protein was always kept below 1 mg ml⁻¹ while storing purified protein. Further, it was also noted that fractions containing similar amounts of proteins exhibited differential enzymatic activities indicating that they contained different amounts of properly folded protein molecules. Presence of a highly hydrophobic palm domain of the protein (data not shown) probably interfered in the proper folding and solubility of the protein. To check the secondary structure of the refolded protein, we carried out circular dichroism (CD) spectroscopy analysis. The CD spectra showed presence of similar structural elements in the protein that were seen in the de novo three-dimensional (3-D) structure predicted using software on the LOOPP server (data not shown). For testing biological activity of the protein, each HPLC fraction was screened for both polymerase and RNA binding activities using 3’NCR (3PS) as template (Fig. 2b), and fractions having both activities were pooled together as a single batch stored in aliquots at -20 °C and used for further assays.
Fig. 2. Selection of putative regulatory elements from the HEV genome. (a) Schematic representation of HEV genome. (b) Different genomic regions of HEV used as templates in the current study, black colour indicates positive sense and grey colour indicates negative sense. The PCR products were used as templates to generate RNA templates using Riboprobe® in vitro transcription system. (Figure is made only for the representation and it is not to scale.) (c) In vitro polymerase activity of recombinant HEV-RdRp proteins: HEV 3’NCR positive-sense RNA (3PS) was used as the template and RdRp polymerase activity was checked using recombinant wild-type HEV-RdRp (T4GDD-RdRp) protein (WT); mutant HEV-RdRp (T4GAA-RdRp) protein (Mu); WT protein with actinomycin D (WT, Act-D). The 40 µl reaction contained 50 mM HEPES (pH 7.5), 5 mM MgCl₂, 250 mM NaCl, 5 mM DTT, 10 % glycerol, 500 µM each of CTP, GTP, UTP and 20 µCi of α³²P-ATP, 100 ng template RNA and
In vitro RdRp polymerase activity assay

Polymerase activity of purified HEV-RdRp was tested using the 3PS template. The recombinant protein could synthesize complementary strand as visualized by labelled (α-32P-ATP incorporated) RNA product having equal length to that of the template RNA. In addition, a high molecular weight band, equivalent to twice the length of template RNA was visualized (Fig. 2c). It was assumed that the larger product was due to loop back extension. Incubation with active site mutant RdRp protein resulted in synthesis of a very low amount of negative-strand RNA (Fig. 2c). Polymerase activity of HEV-RdRp was found to be resistant to actinomycin D (a potent inhibitor of DNA-dependent RNA polymerase) confirming that the synthesis was due to polymerase activity of purified HEV-RdRp (Fig. 2c).

To see whether HEV-RdRp has the ability to initiate primer-independent synthesis of RNA, we blocked 3′-OH ends of 3PS template RNA by NaIØ4 treatment. Blocking of 3′-OH ends was confirmed by 3′-end labelling of the template with α-32P-ATP. As expected, NaIØ4 treated template could not be end labelled (results not shown). HEV-RdRp could synthesize RNA product equivalent to template length from the 3′-OH blocked template (Fig. 2d), however, the enzyme was not able to synthesize larger RNA. These results confirmed de novo priming and complementary strand synthesis by HEV-RdRp. Absence of a larger fragment with blocked template also confirmed that the polymerase was also extending the template strand from the 3′-OH end by a loop back mechanism to generate a larger RNA product. Polymerase activity of the enzyme was found to be optimum at the concentration of 3 µg protein/40 µl reaction, and this concentration was used further for all polymerase assays. An attempt to develop a quantitative assay for RdRp polymerase failed due to very low activity of the purified protein.

Binding of HEV-RdRp to putative regulatory sequences in HEV genome

To analyse binding of recombinant RdRp protein to different RNA templates, we designed PS and negative-sense (NS) RNA templates containing putative cis-regulatory elements from the HEV-4 genome (Fig. 2a, b). An electrophoretic mobility shift assay (EMSA) was optimized using 3′NCR (3PS) RNA template and the assay was further used to analyse binding of different putative regulatory elements. Due to the presence of secondary structures, the template RNA elements, 5PS, 5NS, SG4347, 5NCR250PS and 5NCR250NS migrated as two distinct bands in polyacrylamide gels. However, gel retardation of these RNA fragments could be analysed due to the presence of higher molecular weight complexes. Of the PS RNA templates, enzyme could specifically bind to SG5096PS, 3PS and 5NCR250PS fragments (Fig. 3a, c), while there was no detectable binding with 5PS, SG4347PS and CPS (Fig. 3a). Of the NS templates, enzyme could bind to SG5096NS, 3NS and 5NCR250NS fragments (Fig. 3b and 3c), but there was no binding with 5NS, SG4347NS and CNS (Fig. 3b). The shorter 5′NCR templates, 5PS and 5NS, contained 60 nt 5′ end sequence of HEV genome. While extended 5′NCR templates 5NCR250PS and 5NCR250NS contained 250 nt 5′ end genome sequence, encompassing 51 nt highly conserved alphavirus homology region (Surjit et al., 2004; Huang et al., 1992). Though enzyme was not able to bind to 5PS or 5NS, it showed binding with both PS as well as NS 5NCR250 templates suggesting requirement of additional secondary structures in the ORF1 coding region for binding (Fig. 3c). Overall, the HEV-RdRp could specifically bind to genomic (positive sense) 3PS, SG5096PS, 5NCR250PS templates and antennogenic (negative sense) 3NS, SG5096NS, 5NCR250NS templates. It did not show any binding with templates from ORF2 encoding region (CPS and CNS) and the region homologous to alphavirus SgP (SG4347PS and SG4347NS), suggesting sequence/structure-specific binding of HEV-RdRp.

RdRp polymerase activity with putative regulatory elements in HEV genome

After evaluating binding of HEV-RdRp to putative regulatory elements in the HEV genome, we assessed whether the enzyme has the ability to initiate RNA synthesis from these templates by carrying out a polymerase assay. We included 100-fold excess of non-specific RNA, either tRNA or total Escherichia coli RNA, to mimic cellular milieu in these reactions. Enzyme could initiate RNA synthesis from 3PS (Figs 2c, d and 4c), 5NCR250NS and SG5096NS templates (Fig. 4a, b). Surprisingly, it could also initiate synthesis from 5NCR250PS, with comparable efficiency as that with 5NCR250NS template (Fig. 4a). There was also some synthesis from SG5096PS template, while synthesis from 3NS template was almost nil (Fig. 4b, c). These results showed that although HEV-RdRp binds to SG5096PS and 3NS templates, it is not able to initiate efficient synthesis from these templates. These templates probably lacked sequences/
secondary structures that are essential for initiation of RNA synthesis by HEV-RdRp.

**Relative binding affinities for different RNA templates**

After analysing binding specificities and abilities of different RNA templates as potential HEV-RdRp promoters to initiate complementary RNA strand synthesis, we next analysed relative binding affinities of these templates. Templates showing positive results in EMSA and polymerase assay, i.e. 5NCR250PS, 5NCR250NS, 3PS and SG5096NS, were tested further in competitive mobility shift assays. For that, the test template was labelled with $\alpha\text{-}^{32}\text{P}$-ATP/CTP and incubated with either equimolar or excess concentrations of competitor cold template along with the HEV-RdRp.

Competition assay between labelled 3PS and cold SG5096NS template at equimolar concentration (10 nM) showed partial inhibition of 3PS binding, indicating comparatively higher affinity of HEV-RdRp for 3PS. However, two additional high molecular bands were also seen above.

**Fig. 3.** Electrophoretic mobility shift assays. (a) Binding abilities of the recombinant HEV-RdRp protein for different $\alpha\text{-}^{32}\text{P}$-labelled positive sense (PS) putative regulatory elements in the HEV genome. The 20 µl reaction contained 50 mM HEPES (pH 7.5), 5 mM MgCl$_2$, 250 mM NaCl, 5 mM DTT, 100-fold excess of the carrier RNA (Escherichia coli total RNA), 20 nM template RNA and 1.5 µg (30 pmol) of purified HEV-RdRp. The reactions were incubated at 30 °C for 10 min, analysed on to 8 % native PAGE and visualized by autoradiography. (b) Binding abilities of the recombinant HEV-RdRp protein for different $\alpha\text{-}^{32}\text{P}$-labelled negative-sense (NS) putative regulatory elements in the HEV genome. (c) Binding of HEV-RdRp with $\alpha\text{-}^{32}\text{P}$-labelled positive and NS 5NCR250 templates. Experiments were repeated at least three times to check reproducibility of the results. Data from one of the reproducible experiments is shown.

**Fig. 4.** In vitro polymerase activity using different templates. (a) The polymerase assay with the extended 5’NCR templates (5NCR250PS and 5NCR250NS). (b) The polymerase assay with the putative SgP templates (SG5096PS and SG5096NS). The respective $\alpha\text{-}^{32}\text{P}$-labelled positive-sense RNA fragments were used as size markers along with an RNA ladder. (c) The polymerase assay with 3’NCR templates (3NS and 3PS). The 40 µl reaction contained 50 mM HEPES (pH 7.5), 5 mM MgCl$_2$, 250 mM NaCl, 5 mM DTT, 10 % glycerol, 500 µM each of CTP, GTP, UTP and 20 µCi of $\alpha\text{-}^{32}\text{P}$-ATP, 100 ng template RNA and 3 µg purified RdRp protein. The reaction mixture was incubated at 30 °C for 1 h. All products were analysed on 8 M urea-8 % PAGE and processed for the autoradiography for visualization. Experiments were repeated at least three times to check reproducibility of the results. Data from one of the reproducible experiments is shown.
the protein–3PS RNA complex indicating co-binding of HEV-RdRp with 3PS and SG5096NS templates. At 50-, 100- and 200-fold molar excess of cold SG5096NS, the supershift due to co-binding of two fragments was more clearly evident (Fig. 5a). In the reverse competition experiment, wherein labelled SG5096NS was competed out with cold 3PS, higher affinity of RdRp towards 3PS was further confirmed. At equimolar concentrations, SG5096NS binding was almost completely inhibited by cold 3PS; however, the higher molecular weight complexes due to co-binding were also visible, though at very low levels. The supershift could be clearly seen with only higher concentration of cold 3PS (50-, 100- and 200-fold excess) (Fig. 5b).

Competition of labelled 5NCR250PS with equimolar and 20-fold molar excess of cold 5NCR250NS completely inhibited binding of 5NCR250PS and also formed dsRNA molecules due to complementarity between these two molecules (Fig. 5c). The dsRNA molecule migrated faster than the single-stranded template, probably due to secondary structures in the single-stranded template (Fig. 5c). Competition of 5NCR250PS with cold 3PS showed complete inhibition of 5NCR250PS binding, at equimolar as well as at higher concentration of the 3PS RNA indicating higher binding affinity of 3PS. Cold SG5096NS could inhibit binding of 5NCR250PS only partially at 20-fold molar excess concentration, indicating very high affinity of the protein for 5NCR250PS than that for SG5096NS (Fig. 5c).

Competition of labelled 5NCR250NS with equimolar and 20-fold molar excess of cold 5NCR250PS showed only partial inhibition of 5NCR250PS binding. Again double-stranded, faster migrating RNA bands were visible in the competition reactions due to complementarity between two templates in the assay. However, it was clear from these results that as compared to 5NCR250PS, the enzyme has very high affinity for 5NCR250NS (Fig. 5d). Competition of 5NCR250PS with cold 3PS showed complete inhibition of 5NCR250PS binding, both at equimolar and 20-fold molar excess of 3PS, further confirming higher binding affinity for 3PS. While, competition with cold SG5096NS showed negligible inhibition of 5NCR250PS binding, at equimolar and

---

**Fig. 5.** Competition binding assays using 3PS and SG5096NS templates. (a) Competition between α-32P-labelled 3PS and cold SG5096NS templates. The 20 µl reaction contained 50 mM HEPES (pH 7.5), 5 mM MgCl₂, 250 mM NaCl, 5 mM DTT, 100-fold excess of the carrier RNA (E. coli total RNA), 10 nM labelled 3PS RNA, 500/1000/2000 nM cold SG5096NS RNA and 1.5 µg (30 pmol) of purified HEV-RdRp. The reactions were incubated at 30 °C for 10 min, analysed onto 8 % native PAGE and visualized by autoradiography. (b) Competition between α-32P-labelled SG5096NS and cold 3PS templates. (c) Competition between α-32P-labelled 5NCR250PS and cold 5NCR250NS, 3PS and SG5096NS templates. (d) Competition between α-32P-labelled 5NCR250NS and cold 5NCR250PS, 3PS and SG5096NS templates. Experiments were repeated at least three times to check reproducibility of the results. Data from one of the reproducible experiments is shown.

---

**Table 1.** Concentration of RNA templates used in competition binding assays. The 20 µl reaction contained 50 mM HEPES (pH 7.5), 5 mM MgCl₂, 250 mM NaCl, 5 mM DTT, 100-fold excess of the carrier RNA (E. coli total RNA), 10 nM labelled 3PS RNA, 500/1000/2000 nM cold SG5096NS RNA and 1.5 µg (30 pmol) of purified HEV-RdRp. The reactions were incubated at 30 °C for 10 min, analysed onto 8 % native PAGE and visualized by autoradiography.

<table>
<thead>
<tr>
<th>Template</th>
<th>Concentration (nM)</th>
<th>Cold RNA (nM)</th>
<th>5NCR250NS</th>
<th>3PS</th>
<th>SG5096NS</th>
</tr>
</thead>
<tbody>
<tr>
<td>α-32P-3PS (nM)</td>
<td>10</td>
<td>–</td>
<td>–</td>
<td>10</td>
<td>1000</td>
</tr>
<tr>
<td>5NCR250NS (nM)</td>
<td>–</td>
<td>–</td>
<td>10</td>
<td>10</td>
<td>1000</td>
</tr>
<tr>
<td>3PS (nM)</td>
<td>–</td>
<td>–</td>
<td>10</td>
<td>10</td>
<td>1000</td>
</tr>
<tr>
<td>SG5096NS (nM)</td>
<td>–</td>
<td>–</td>
<td>10</td>
<td>10</td>
<td>1000</td>
</tr>
</tbody>
</table>

---
combinations of different concentrations of these two templates. Surprisingly, HEV-RdRp showed more efficient synthesis from SG5096NS when both 3PS and SG5096NS templates were present in equimolar concentrations (100 nM) (Fig. 6a, b). On increasing 3PS template concentration (200, 300 and 400 nM) further though there was reduced synthesis from SG5096NS; it was not completely inhibited even by a fourfold higher concentration of 3PS. These results indicated that though HEV-RdRp has more binding affinity for 3PS, it can initiate RNA synthesis more efficiently from SG5096NS as compared with 3PS. Moreover, in the reverse experiment, on doubling SG5096NS template concentration, there was a complete inhibition of synthesis from the 3PS template, proving that SG5098NS was comparatively a more efficient promoter than the 3PS (Fig. 6a, b).

Identification of peptide motifs of HEV-RdRp that bind to 3PS and SG5096NS

A first step towards identifying possible mechanisms behind regulation of SG5096NS and 3PS promoters during HEV replication was to know whether these two RNA elements were interacting with each other or the enzyme has separate binding sites/motifs for these RNA elements. We carried out in silico sequence analysis to evaluate possibility of sequence-based interactions between 3PS and SG5096NS RNA templates. There were no sequence complementarities between these two elements. Further, incubation of the two labelled RNA templates together in EMSA reaction in absence of enzyme with BSA to analyse possible RNA–RNA interaction, did not reveal any positive results (data not shown).

To see the possibility of separate interacting motifs within RdRp, the 3PS and SG5096NS RNA templates were allowed to bind separately to HEV-RdRp and resolved by gel electrophoresis. After confirming mobility shifts (Fig. 7a, d), UV cross-linking was done and the protein–RNA complex was digested with trypsin and proteinase K. The resulting RNA–RdRp peptide complex/es were eluted from the gel and subjected to N-terminal sequencing and Quadrupole Time of Flight (Q-TOF) analysis. Mass spectrometry results showed that two peptides, I: QSRNAAA (aa 274–280 in HEV-4 RdRp protein) and II: YGRRTKLYEAAA-HADVRGSL (aa 25–43), interacted with SG5096NS (Fig. 7b, c). While the 3PS RNA template interacted with peptides, III: NAAALIAGCGLK (aa 277–288), IV: LYEAA–HADVR (aa 31–40) and V: DVSRTFFQKDCNK (aa 86–99) (Fig. 7e, f). To locate these peptides on the RdRp molecular structure, we carried out de novo 3-D structure modelling of the HEV-4 RdRp protein. The CD spectroscopy results of refolded RdRp protein were consistent with the predicted 3-D structure of the protein. It was observed that peptides I and II are closely located within the molecule (Fig. 7b, c). The 3PS and SG5096NS RNA templates shared closely localized motifs I and II (III was part of
peptide I and IV was part of peptide II). However, binding of 3PS involved an additional peptide V, that was comparatively away from I and II peptides on the RdRp molecule (Fig. 7e, f).

**DISCUSSION**

Due to paucity of efficient cell culture systems for HEV studies, molecular events during viral replication are yet not completely understood. One of the key steps in initiating viral RNA synthesis requires specific recognition of RNA elements in the viral genome by the viral polymerase. For characterization of viral RdRp and its regulatory elements in the HEV genome, the predicted RdRp domain of ORF1, encompassing all conserved I–VIII motifs (Koonin et al., 1992) was expressed in a bacterial system and tested for its enzymatic activity. Selection of putative regulatory elements in the HEV genome was based on previous reports. The importance of the 3'-end sequence containing two stem–loop structures (SL1 and SL2) and poly A tail in 3’NCR in

---

**Fig. 7.** Identification of SG5096NS and 3PS template interacting sites within the enzyme. (a) Mobility shift assay was performed using both, ³²P-labelled and cold SG5096NS templates, and the products were separated by loading them in adjacent wells on the gel. RNA–protein complexes were UV cross-linked, wells containing unlabelled template were separated and processed separately, and remaining gel was visualized on the phosphorimager. (b) Predicted 3-D ribbon structure of the RdRp protein with highlighted peptides which showed binding to the SG5096NS template. (c) Surface structure of the RdRp protein with highlighted peptides which were binding to the SG5096NS template. (d) Mobility shift assay with 3PS performed similarly as mentioned above to analyse binding sites in the enzyme. (e) Predicted 3-D ribbon structure of the RdRp protein with highlighted peptides which showed binding to the 3PS template. (f) Surface structure of the RdRp protein with highlighted peptides which were binding to the 3PS template. The experiment was repeated twice to check reproducibility of the results.
binding and initiating RNA synthesis has been previously reported (Agrawal et al., 2001; Emerson et al., 2007; Graff et al., 2005a). These observations prompted us to use 176 nt 3'-end sequence which contained 3'NCR and the end of the ORF2 encoding region of the HEV genome (3PS). The 3'NCR template included 64 nt NCR along with poly A tail and SL1 and SL2 from the ORF2 encoding region. The JR between ORF1 and ORF2, containing the putative promoter for the sgRNA synthesis has been shown to be important during HEV replication and transcription (Cao & Meng, 2012; Huang et al., 2007; Graff et al., 2005b). We selected 145 nt stretch from the ORF1-ORF2 JR which included a highly conserved stem–loop structure and start of the ORF3/ORF2 encoding region (SG5096). The 5'NCR of the HEV genome is only 26 nt long and has a methylated cap, however, its exact role as regulatory element or cis-reactive element during HEV replication is not yet established. We tested two RNA templates for 5'NCR, a 60 nt stretch including a 26 nt 5' non-coding sequence, plus downstream a 34 nt ORF1 encoding region (5PC) and an extended 5'-end fragment (5NCR250) containing 76 nt stretch from ORF1 that was found to be involved in binding to ORF2 protein (Surjit et al., 2004).

HEV-RdRp could initiate synthesis of the complementary RNA strand from the PS 3'NCR RNA template (3PS). Further, using an active site mutant protein, T4GAA-RdRp, and actinomycin D (a known inhibitor of DNA-dependent RNA polymerases); we confirmed that the polymerase activity was due to HEV-RdRp and not due to contamination with E. coli DNA-dependent RNA polymerases (Fig. 2c, d). The HEV-RdRp could synthesize RNA strand of the length equivalent to the template strand, and also extended the template from the 3'-OH end via a loopback mechanism and synthesized a twofold longer fragment. To confirm this, we blocked 3'-OH of the template and used it in the polymerase reactions. Presence of a single product equivalent to the length of the template in these reactions (Fig. 2d) indicated that HEV-RdRp carries out two types of initiations, it has ability to initiate primer-independent (de novo) RNA synthesis from the RNA template and can also utilize the template end as the primer to extend it from the 3′OH end. The significance of this second activity during virus replication is not understood! Synthesis of small amount of negative-strand RNA by the active site mutant protein (GDD changed to GAA) was surprising (Fig. 2c). It is our previous observation that similar mutations in HEV-RdRp could establish specific binding to the enzyme, suggesting requirement of the three stem–loop structures, SL1, SL2 and SL3, extending from 130 to 250 nt, for this interaction. The initial stretch of 5′NCR may not be essential for this interaction; however, in depth analysis is required to precisely locate HEV-RdRp interacting element/s at the 5′ end of the HEV genome. With these results, we propose presence of a third cis-regulatory element at the 5′-end of the HEV genome, in addition to previously proposed 3'NCR and putative SgP cis-regulatory elements.

Similar to 5'NCR, the enzyme could bind to both genomic and antigenomic sense of SgP templates, SG5096PS and SG5096NS (Fig. 3a, b), however, it could initiate RNA synthesis efficiently only from the antigenomic template (SG5096NS) (Fig. 4b), indicating that the template contains RNA elements that are essential for its function as the promoter. Our competitive binding assays also showed comparatively higher affinity of the enzyme for antigenomic sense SG5096NS template.

We used enzymatically active protein further to analyse binding affinities for the putative regulatory elements in the HEV genome/antigenome. For that, the 5′NCR, 3′NCR and putative SgP RNA templates were made in both genomic and antigenomic sense and used in EMSA. The protein could bind to 5NCR250PS, 3PS, SG5096PS genomic sense templates and 3NS, 5NCR250NS, SG5096NS antigenomic sense templates (Fig. 3a, b, c). This binding was highly specific since these RNA elements were present in a mixed pool of 100-fold excess of non-specific RNA molecules. HEV-RdRp was not able to bind to SG4347PS and SG4347NS RNA templates, a region homologous to SgP of Sindbis virus, which was previously thought to be important for HEV replication (Purdy et al., 1993).

We further analysed whether RdRp initiates de novo synthesis from the templates that showed positive binding in EMSA. HEV-RdRp could bind to both genomic sense 5NCR250PS and antigenomic sense 5NCR250NS 5′NCR templates and exhibited more affinity for the 5NCR250NS as compared with 5NCR250PS (Fig. 5c, d). Surprisingly, enzyme initiated synthesis from both 5NCR250PS and 5NCR250NS templates (Fig. 4a). It was anticipated that the HEV-RdRp will bind to 5NCR250NS (5′-end of the antigenome) to initiate synthesis of gRNA, however, synthesis from the PS template was unexpected. We speculate that once synthesized in abundance, HEV capsid protein would bind to the 76 nt region in the 5′NCR and inhibit further binding of HEV-RdRp. This would eventually prevent synthesis from this promoter. Further, it was also evident from our competition results that the enzyme had more affinity for 3PS (3'NCR of the HEV genome) as compared with 5′NCR. This would further restrict binding of the HEV-RdRp to 5′NCR. We tested 60 and 250 nt stretches from the 5′-end of the HEV genome and saw that only the 250 nt template, containing 76 nt ORF2 interacting sequence (130–206 nt from the HEV genome) (Surjit et al., 2004) could establish specific binding to the enzyme, suggesting requirement of the three stem–loop structures, SL1, SL2 and SL3, extending from 130 to 250 nt, for this interaction. The initial stretch of 5′NCR may not be essential for this interaction; however, in depth analysis is required to precisely locate HEV-RdRp interacting element/s at the 5′ end of the HEV genome. With these results, we propose presence of a third cis-regulatory element at the 5′-end of the HEV genome, in addition to previously proposed 3′NCR and putative SgP cis-regulatory elements.

For the 3′NCR, the enzyme could bind and initiate synthesis only from the genomic sense (3PS) template. Competition assays between 3′NCR, SgP and 5′NCR templates showed that the protein has highest affinity for the 3′NCR genomic sense template, then for the antigenomic sense 5′NCR and then antigenomic SgP templates. Interestingly, competition between genomic sense 3′NCR (3PS) and the antigenomic sense SgP (SG5096NS) showed a supershift, indicating
co-binding of these two fragments to the enzyme (Fig. 5a, b). This prompted us to identify binding sites for these templates. It has been suggested that viral RdRps may have different sites for the recognition of genomic promoters and SgPs in brome mosaic virus (Grdzelishvili et al., 2005). Sindbis virus RdRp has distinct binding sites for SgP and genomic promoters (Li & Stollar, 2004, 2007).

On analyzing binding sites in the HEV-RdRp, it was seen that SG5096NS interacted with peptides I (17aa) and II (19aa) (Fig. 7b, c), while 3PS interacted with peptides III, IV and V in the enzyme (Fig. 7e, f). These two RNA elements shared two peptides III (12 aa region within the pepti- de I) and IV (10 aa region within the motif II) during binding, however, the third peptide, V (17 aa) interacted specifically with the 3′NCR. Moreover, on analysing represen- tative HEV genomic sequences from the four mammalian genotypes (HEV-1–4), the 3′NCR specific peptide V was found to be highly conserved across all HEV genotypes, suggesting its critical role during HEV replication. It would be worthwhile to do in depth analysis of peptide V to understand its role during HEV replication.

From our results, it was evident that HEV-RdRp does not have separate binding sites for 3′NCR and SgP, but it has an additional site for selective interaction with 3′NCR. HEV- RdRp probably binds to SgP via motifs I and II, and to 3′NCR with III, IV and V via its unique interacting peptide V. The co-bound RdRp probably jumps to the adjacent antige- nomic molecule via 3′NCR of the HEV genome, while it is still binding to peptide V and then initiates sgRNA synthe- sis. This cooperative selective recruitment of the protein would prevent its binding to the 5NCR250NS in the anti- genome to favour synthesis of sgRNA. Further experiments are required to find out the exact mechanism involved in switching of promoters by the HEV-RdRp.

To see how synthesis of new RNA occurs from the co- bound templates, we tested 3′NCR and SgP templates together in a polymerase assay. Surprisingly, SgP was found to be a stronger promoter as compared with 3′NCR and initiation synthesis efficiently even in the presence of equimolar concentration of 3′NCR (Fig. 6a, b). It is previously reported that, interaction between promoters at the 3′ and 5′-ends of Sindbis virus genome is essential for negative- strand RNA synthesis (Frolov et al., 2001). By having interaction between promoters with different efficiencies, HEV probably manages to enhance synthesis of sgRNA for capsid protein accumulation and restricts synthesis of antigenomic RNA copies that would be generated from the 3′NCR prom- otter. It is previously documented that at any given time there is always a two- to threefold excess of the PS HEV RNA compared with the NS RNA in HEV replicating cells (Kumar et al., 2010). Preferred recruitment of HEV-RdRp to SgP via 3′NCR would also prevent binding to 5NCR250NS. There could be some additional, yet unknown mechanism, to switch promoters from SgP to 5NCR250NS, to initiate synthesis of gRNA. For these need to have a robust polymerase assay.

METHODS
Cloning, expression and purification of HEV-RdRp. The RdRp encoding region encompassing all I–VIII conserved motifs of the protein (Rehman et al., 2008) (1262b from 3886–5148 nt, 1286–1706aa of the ORF1) (HEV-4, GenBank accession no. AY723745) was PCR amplified with RDRPF and RDRPR primers (Table S1, available in the online Sup- plementary Material) using Platinum® Pfx DNA polymerase (Invitro- gen) and cloned into pET15b (Invitrogen) in frame with N-terminal polyhistidine tag (pET15b(41TRdRp). Protein expression was carried out in E. coli BL21 (RIL) codon plus cells (Stratagene) using 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG) and 10 mM MgCl₂ for 4 h at 37 °C. Protein was purified using ProBond Nickel chelating resin column (Invitrogen). Briefly, cell pellet was lysed using lysis/binding buffer (8 M urea, 20 mM sodium phosphate, pH 7.8, 0.5M NaCl, 10 mM MgCl₂) and was centrifuged at 10 000 g for 30 min. The supernatant was filtered through 0.45 µm syringe filter (Millipore) and loaded on to ProBond Nickel chelating resin column. Unbound proteins were removed with washing buffer (8 M urea, 20 mM sodium phosphate pH 6.0, 0.5M NaCl, 10 mM MgCl₂) and protein was eluted in elution buffer (8 M urea, 20 mM sodium phosphate pH 4.0, 0.5M NaCl, 10 mM MgCl₂). Fractions containing protein of expected size (48 kDa) were pooled, concentrated and renatured by exchanging denaturing buffer with native buffer (500 mM NaCl, 20 % glycerol, 50 mM HEPES, pH 7.5, 10 mM MgCl₂, 50–100 µM GTP) and protein was eluted in elution buffer (8 M urea, 20 mM sodium phosphate pH 4.0, 0.5M NaCl, 10 mM MgCl₂). Frac- tions containing desired protein were analysed by Western blot analysis using anti- His mAbs (Sigma). Protein concentration was determined by Lowry’s method and stored at −20 °C until use.

Generation of active site mutant of RdRp by site directed mutagenesis. RdRp mutant protein was generated by changing amino acids GDD to GAA using QuikChange® Site-Directed Mutagenesis Kit (Stratagene) for use as a negative control in polymerase assays. It is our observation that HEV replicons harbouring GDD to GAA mutations fail to replicate in cells (Devhare et al., 2016).

Circular dichroism measurements. Circular dichroism (CD) spec- tra of refolded recombinant RdRp protein was collected using a Jasco J-815 CD spectrophotometer (Tokyo) with the following instrument settings: measurement wavelength range: 250–190 nm and slit width: 100 µm. CD spectra of 33.4 µg ml⁻¹ of RdRp protein sample were recorded and baseline correction was done using buffer as blank. Dichroic software was used to analyse the CD data and curves were fitted with con- strained least square self-consistent method to obtain the value of alpha-helical, beta-sheet and random coil conformations of protein (Sreerama et al., 1999).

Generation of RNA templates. Selective regions of the HEV-4 full genome clone were PCR amplified using specific primers containing T7 promoter sequences (see Table S1). The PCR products were used as templates to generate RNA templates using Riboprobe® in vitro transcription system (Promega). The positive sense (PS) and negative sense (NS) RNA transcripts generated included: 5′NCR (first 60 nt of HEV genome) - 3PS and NS5; 5′NCR extended (first 250 nt of HEV genome) - 5NCR250PS and 5NCR250NS; 4347–4407 region (60 nt region homologous to SgP in alphavirus) - SG4347PS and SG4347NS; putative SgP for HEV (5096–5240, 145 nt) - SG5096PS and SG5096NS; 3′NCR (with SL1, SL2 comprising of nucleotides 7173–7260, 176 nt) - SG5096PS and SG5096NS; 3′NCR (with SL1, SL2 comprising of nucleotides 7173–7194 and 7089– 7163 and poly A tail 7084–7260, 176 nt) – 3PS and NS5 and a non-spe- cific RNA control (6421–6540, 120 nt region from ORF2) - CPS and CN5, (Fig. 2a, b). α²P-labelled RNA fragments were generated to use as markers by in vitro transcription.
**In vitro RdRp polymerase activity assay.** RdRp activity of the recombinant protein was evaluated by an *in vitro* assay using α[^32]P-ATP (BRIT) and 3PS as template. The 40 µl reaction contained 50 mM HEPES (pH 7.5), 5 mM MgCl₂, 250 mM NaCl, 5 mM DTT, 10% glycerol, 500 µM each of CTP, GTP, UTP and 20 µCi of α[^32]P-ATP, 100 ng template RNA and 3 µg purified RdRp protein. Reaction mixture was incubated at 30°C for 1 h. Unincorporated rNTPs were removed by using Centricip column (Princeton Separations) and eluted RNA was vacuum dried in speed vac, pellet was resuspended in 5.0 µl RNA loading dye (Ambion), denatured at 70°C for 10 min, chilled on ice and separated on 8% denaturing PAGE containing 8 M urea. The products were visualized by autoradiography using Carestream Kodac Biomex.

**Periodate (NaO₄) treatment and RdRp polymerase activity.** The periodate treatment was carried out as described previously (You *et al.*, 1999) to block the 5’-OH groups in the template RNA. For that template RNA was dissolved in reaction buffer (50 mM NaOAc, pH 5.0), 20 mM NaO₄ was added and incubated for 1 h at room temperature. Lysine (60 mM) was added to saturate excess periodate and further incubated for 3 h at room temperature. The treated RNA was purified using Centricip column and quantitated spectrophotometrically (NanoDrop, Thermo Fisher). Integrity of RNA was checked by agarose gel electrophoresis. Templates obtained after periodate treatment were used for RdRp assay.

**Electrophoretic mobility shift assay (EMSA).** 3’NCR (3PS) was used as template to develop EMSA to evaluate binding activity of purified HEV-RdRp protein. The 20 µl reaction contained 50 mM HEPES (pH 7.5), 5 mM MgCl₂, 250 mM NaCl, 5 mM DTT, 100-fold excess of the carrier RNA (E. coli total RNA), 20 nM template RNA and 1.5 µg (30 pmol) of purified HEV-RdRp. The template RNA was heated at 65°C for 5 min and allowed to come to room temperature slowly to let the formation of native conformational structure and then used in the assay. In addition, reaction mixture was incubated with RNasin (Promega) at 37°C for 30 min prior to addition of RNA. After adding template RNA and carrier RNA, reactions were incubated at 30°C for 10 min and immediately loaded on to 8% native PAGE to analyse electrophoretic mobility shift. Competition assays were performed using unlabelled (cold) competitor RNA in 20-, 50-, 100- and 200-fold molar excess as compared to the labelled template RNA and analysed similarly.

**Identification of RdRp interacting motifs.** *In silico* enzymatic digestion of HEV-RdRp was carried out with MS-Digest programme using Protein Prospector proteomic tool (http://prospector.ucsf.edu/prospector/mshome.htm) and based on the results, trypsin and proteinase K were used for gel digestion. Briefly, EMSA was carried out with the labelled as well as cold 3PS and SN5096NS RNA templates in separate binding reactions, products were separated by electrophoresis in parallel wells and in gel UV cross-linking was carried out for 30 min. The retarded labelled fragment was visualized on the Phosphorimager and both labelled and cold cross-linked products were excised from the gel, washed with 10 mM ammonium bicarbonate solution and processed overnight in gel digestion with trypsin (20 µg ml⁻¹ Trypsin Gold Mass spectroscopy grade; Promega) and proteinase K (20 µg ml⁻¹; Promega). Digested products were separated on 20% SDS-PAGE, labelled RNA was visualized on Phosphorimager and the parallel band, corresponding to cold RNA, was excised from the gel. The excised band was subjected to alklylation and reduction and eluted using 1% trifluoroacetic acid. The resultant peptide masses were analysed by N-terminal sequencing and Q-TOF (Bruker Daltonics). The mass spectra produced from Q-TOF were searched against protein databases NCBI, MSDB and Swiss-prot using the Mascot, MSfit and Profound search engine, and peptide sequences were determined.

**3-D structure modelling and visualization.** Structural prediction of HEV-RdRp protein was carried out using LOOPP server (http://biosapps.tc.cornell.edu/loopp.aspx). The output of the structural prediction algorithms was imported into 3-D viewer Chimera [Pettersen et al., 2004] and structures were visualized.

**REFERENCES**


