Nucleotide sequence of the gene coding for the major protein of hepatitis B virus surface antigen

Pablo Valenzuela, Patrick Gray, Margarita Quiroga, Josefina Zaldivar, Howard M. Goodman & William J. Rutter

Department of Biochemistry and Biophysics, University of California, San Francisco, San Francisco, California 94143

DNA extracted from hepatitis B virus Dane particles has been cloned in bacteria using a plasmid vector. A full-length clone has been examined by restriction endonuclease analysis, and the nucleotide sequence of an 892-base pair fragment from cloned hepatitis B viral DNA encoding the surface antigen gene is reported. The amino acid sequence deduced from the DNA indicates that the surface antigen is a protein consisting of 226 amino acids and with a molecular weight of 25,398. The portion of the gene coding for this protein apparently contains no intervening sequences.

VIRAL HEPATITIS is a major worldwide public health problem. Its aetiology is associated with at least three distinct viral families: hepatitis A, hepatitis B and non-hepatitis A/hepatitis B (ref. 1). The characterisation of the viruses, the pathology associated with the infection and the development of effective means of control have been severely hampered by the very narrow host range specificity of the viruses, and the inability to replicate the virus and obtain cytopathology in tissue culture systems. Nevertheless, progress has been made in identification and preliminary characterisation of the viruses from the sera of infected individuals and from livers at autopsy. The hepatitis A virus is apparently a picornavirus detected as a 27-nm particle^{2,3}. In addition to being infective in humans and chimpanzees, it has recently been transferred to marmosets and has been replicated in tissue culture⁴⁻⁶.

The plasma of individuals infected with heptatitis B show three major particulate structures containing the antigenic determinants apparently specified by the hepatitis viral genome. These include the predominant 22-nm particles, 22-nm filaments of various lengths, and the 42-nm spherical form known as the Dane particle^{7,8}. The Dane particle is probably the hepatitis B virion. Recently, a human hepatoma cell grown in tissue culture was shown to produce small quantities of hepatitis B antigens⁹.

In instances where productive infections and high titres of viruses cannot be obtained in alternative hosts or tissue culture cells, recombinant DNA methods can be used to obtain large quantities of the virus genome for characterisation and biological activity studies. Structural analysis of the various antigens may allow identification and sequence determination of the various genes of the virus. This information, plus the availability of the viral genome, will greatly aid the study of the pathology and eventually the development of antiviral therapy. In particular, it may be possible, using segments of the viral genome, to produce appropriate viral antigens in alternative hosts such as bacteria or yeast for the development of a vaccine.

Hepatitis B is a particularly attractive paradigm for the development of this approach. The hepatitis B viral genome is relatively simple, perhaps consisting of only 3,200 bases^{10,11}. The surface coat, which can be removed by treatment with detergents, contains two main polypeptides, P1 (22–24,000 MW) and P2 (25–29,500 MW)¹². These molecules are antigenically indistinguishable and protein components are probably identical. Their amino acid compositions are the same

within the limits of the analyses used, and the first 19 residues from the N-terminus and the three amino acid residues at the carboxy-terminus seem to be identical¹³. The higher molecular weight molecule (P2) is a glycoprotein, and it has been suggested that the major virus-specific surface antigens are contained within a single gene. Five to seven other polypeptides ranging upward in size to 97,000 MW have been reported to be present in surface antigen preparations¹⁴. It is not known whether any or all of these molecules are coded for by the viral genome. Viral core preparations contain a major polypeptide of MW 17,000 to

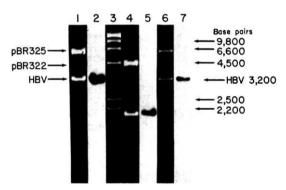


Fig. 1 Cloning and hybridisation of hepatitis B viral DNA. Doublestranded DNA was synthesised in Dane particles (isolated from human sera by Merck Laboratories) by the endogenous DNA polymerase reaction 10 Incubation was for 3 h at 37 °C in a reaction mixture containing Dane particles (250 ng of DNA), 50 mM Tris-HCl (pH 7.5), 0.1 M NaCl, 20 mM MgCl₂, 40 mM NH₄Cl, 0.4% NP40, 10 mM 2-mercaptoethanol and 0.2 mM each dATP, dGTP, dCTP and dTTP. DNA was isolated by phenol extraction and ethanol precipitation after digesting the Dane particles for 60 min at $56\,^{\circ}\text{C}$ with proteinase K $(0.8\,\text{mg ml}^{-1})$. Dane DNA was digested to completion with endonuclease EcoRI (producing a single 3,200-base pair fragment) and with BamHI (yielding two fragments of 2,100 and 1,100 base pairs). Digested Dane DNA (250 ng) was ligated to the appropriate plasmid (50 ng of EcoRI digested pBR325 (ref. 18), or 680 ng of phosphatasetreated, BamHI-digested pBR322 (ref. 30)) for 15 h at 14 °C in a reaction mixture containing 50 mM Tris-HCl (pH 8), 1 mM ATP, 10 mM MgCl₂, 20 mM dithiothreitol and 1 unit of T4 DNA ligase (New England Biolabs). The reaction mixture was used to transform E. coli HB101 (in P3/HV1 containment conditions) or E. coli x1776 (in P2/HV2 containment conditions). Recombinant colonies derived from the DNA ligated to the BamHI site of pBR322 were selected by their ampicillin resistance and tetracycline sensitivity¹⁷ and screened for plasmid size¹⁹. A clone (pHBV-2100) of approximately 6,500 base pairs was identified by a modified toothpick assay19. Recombinant colonies derived from the DNA ligated to the EcoRI site of pBR325 (ref. 18) were selected by their sensitivity to chloramphenicol. A clone (pHBV-3200) containing plasmid DNA of approximately 8,600 base pairs was identified by the toothpick assay. Supercoiled plasmid DNA was isolated from either clone by a cleared lysate procedure³¹, purified by CsCl-ethidium bromide gradient centrifugation31, digested with EcoRI or BamHI and analysed in 1% agarose gels. DNA fragments were hybridised according to Southern³² to Dane particle DNA labelled by nick translation³³. Lane 1, ethidium bromide staining of fragments derived from digestion of pHBV-3200 with EcoRI. Lane 2, hybridisation of the same fragments to ³²P-Dane particle DNA and autoradiography. Lane 3, HindIIIdigested & DNA as MW standards. Lane 4, ethidium bromide staining of fragments derived from digestion of pHBV-2100 with BamHI. Lane 5, hybridisation of the same fragments to ³²P-Dane particle DNA and autoradiography. Lane 6, ethidium bromide staining of fragments derived from digestion of pHBV-64 (a clone of pHBV-3200 viral DNA inserted in the Psil site of pBR322). Lane 7, hybridisation of the same fragments to ³²P-Dane particle DNA and autoradiography.

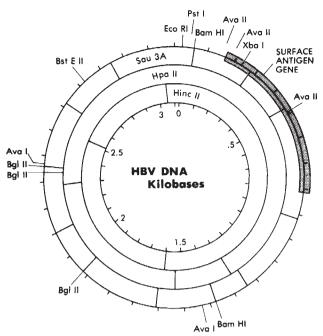


Fig. 2 Restriction endonuclease cleavage map of cloned HBV DNA in plasmid pHBV-3200. Cleavage with restriction endonucleases was carried out in the conditions described by the suppliers with excess amounts of each enzyme. Analysis of the resulting fragments was by electrophoresis in 7% or 10% acrylamide slab gels in 50 mM Tris-borate (pH 8.3) and 1 mM EDTA or in 1-2% agarose slab gels in 50 mM Tris-acetate (pH 8.3) and 1 mM

19,000 and a variety of other larger polypeptides of MWs 25,000 to 200,000 (ref. 12). Whether these polypeptides are virus-specific determinants is unknown. A third class of antigens associated with hepatitis infection are the e antigens¹⁵, which recent studies suggest are a dissociated form of the core antigen¹⁶.

Here we report the cloning in bacteria and structural analysis of DNA from hepatitis B virus (HBV). We have defined the location of the viral surface antigen gene in the restriction endonuclease cleavage map of the cloned HBV genome. Nucleotide sequence of an 892-base pair region encoding the surface antigen gene allows us to deduce the complete amino acid sequence of this protein.

Cloning of DNA from hepatitis B virus Dane particles

DNA from Dane particles (prepared from pooled sera received from Merck, Sharp and Dohme) was labelled with 32P-dATP and 32P-dCTP using the endogenous DNA polymerase reaction¹⁰, and purified as described by Landers et al.¹¹. Several different DNA preparations were cleaved by the restriction endonuclease EcoRI into a single fragment of approximately 3,200 base pairs and by BamHI into two fragments of approximately 2,100 and 1,100 base pairs (results not shown). For cloning, the Dane particle DNA was digested with BamHI endonuclease and separately with EcoRI, ligated to the respective sites of pBR322 (ref. 17) and pBR325 (ref. 18) and used to transform Escherichia coli x1776. Ampicillin-resistant colonies derived from the DNA ligated into the BamHI site were further screened for their sensitivity to tetracycline 17 and the size of the plasmids they contained 19. Similarly, recombinant colonies derived from the DNA ligated into the EcoRI site of pBR325 were screened for their sensitivity to chloramphenicol 18 and by analysis of their plasmids¹⁹.

Putative recombinant plasmids were isolated and examined by agarose gel electrophoresis after digestion by *EcoRI* and *BamHI*. One clone from the *EcoRI* experiment (named pHBV-3200) was found to contain a hybrid plasmid with an inserted DNA fragment of approximately 3,200 base pairs, the reported size of linearised viral DNA²⁰. A clone from the *BamHI*

experiment (named pHBV-2100) was found to contain a plasmid with a 2,100-base pair insert. The identities of the inserts within pHBV-3200 and pHBV-2100 were verified by cleavage of the plasmids with EcoRI and BamHI and hybridisation of the fragments with ³²P-labelled Dane particle DNA (Fig. 1). Physical mapping of both inserted fragments with restriction endonucleases showed that pHBV-2100 contains a BamHiderived insert also present in pHBV-3200.

Location and sequence analysis of the fragment containing the HBV surface antigen gene

A physical map containing the cleavage sites for several restriction endonucleases was prepared from pHBV-3200. This map has been obtained by a combination of enzyme digestions and DNA sequence analysis (Fig. 2). No restriction sites were found for XmaI, SstI, KpnI, HpaI, XhoI and SacI endonucleases. The map is similar, but not identical, to those recently described by others for cloned hepatitis viral DNA²¹⁻²³. For example, the map reported by Charnay et al.²³ shows an extra BamHI site in the region containing the surface antigen gene. Examination of the nucleotide sequence in this region (Figs 3, 4) shows that the sequence GGATCA coding for the amino acids Gly-Ser may be converted to a BamHI site GGATCC by a change of one base in the third position of the codon, which would therefore not result in a change in the amino acid sequence.

The HBV surface antigen gene was located by extensive nucleotide sequence analysis of the entire cloned viral genome to identify the sequence coding for the 19 amino acids present at the amino-terminus of the protein¹³. The location of the gene within the physical map of the cloned viral DNA is shown in Fig. 2.

The nucleotide sequence of an 892-base pair region encoding this gene was determined as summarised in Fig. 5 legend. The sequence, including the restriction enzyme cleavage sites, is presented in Fig. 3. The translation of this sequence in one of the phases (Fig. 4) predicts precisely the sequence of the 19 Nterminal amino acids of the proteins P1 and P2, except for residue 15, in which the DNA sequence predicts a leucine instead of the serine. Re-examination of the original data of Peterson and Vyas has shown that residue 15 is indeed leucine, and that the original report was incorrect. Furthermore, the amino acid sequencing data up to residue 31 is completely compatible with the amino acid sequence predicted by the DNA, with the exception of residue 24 (arginine), in which insufficient amounts of the phenylthiohydantoin amino acid were formed for analysis (D. Peterson, personal communication). The reported C-terminal sequence, Val-Tyr-Ile¹³, is in phase from 204 amino acid residues towards the 3' RNA terminus just before the ochre termination codon, UAA. The polypeptide encoded by these sequences is 226 amino acids long and has a MW of 25,398, in satisfactory agreement with the mass for P1 (22-24,000) determined by SDS gel electrophoresis¹². The amino acid composition also agrees very closely with that reported for this protein¹³. Note also the relatively high content of proline (10.2%), tryptophan (5.8%), aromatic amino acids (Trp, Tyr, Phe) (15.0%) and hydrophobic amino acids (Val, Ile, Leu) (26.6%).

Because of the prevalence of intervening sequences in eukaryotic genes (see ref. 24 and refs therein), it is not possible to presume the colinearity of a gene with the amino acid sequence of the protein product. However, there is no evidence for an intervening sequence in the surface antigen gene, as the molecule predicted by the DNA sequence closely approximates the characteristics of the isolated surface antigen P1. Any intervening sequence(s) would have to be small (<150 bases); most intervening sequences in structural genes are longer. The N-terminal and C-terminal ends of the molecule are in phase, thus any intervening sequence must also maintain the phase. Furthermore, preliminary studies suggest that the number of major peptides resolved from trypsin-treated aminoethylated

P1 agrees quite well with those predicted from the amino acid sequence derived from the gene (D. Peterson, personal communication). Further studies on P1 or its mRNA should resolve this issue decisively. In the meantime, it seems justified to assume colinearity of the gene with the mRNA for purposes of examination of the primary structure of P1. The high concentration of proline residues and their dispersion throughout the molecule preclude the possibility of a high content of α -helix. Indeed, according to the Fasman rules²⁵, there is only approximately 4% α -helix. However, there is a significant amount of β form ($\sim 30\%$) and also more than 10% β -turns. Thus, the molecule probably has a globular configuration. Perhaps the most striking primary structural feature of the molecule is a 19-amino acid long hydrophobic region. The sequences flanking this region are particularly rich in proline and cysteine residues. This hydrophobic region may be the site for intermolecular interaction of P1 monomers, and the abundant cysteine residues could make intra- and intermolecular linkages to form the surface coat of the virus. Because of the likelihood that P1 is glycosylated, we searched for possible sites of glycosylation at asparagine residues (Asn-X-Ser/Thr)²⁶. There are three such

Directions for future study

Although P1 and P2 are the major peptides, five or six other proteins of larger size (up to ~100,000 MW) have been observed in surface antigen preparations¹². One of these, P6 (~72,000 MW), is sometimes present as a major component¹². This family of molecules cross-reacts antigenically and hence must be related structurally. Whether they represent aggregates of P1/P2 or are distinct molecules is not known. We have not detected nucleotide sequences encoding similar amino acid sequences in the other regions of the viral genome. The possibility that they are formed by processing from a single peptide precursor should also not be overlooked. However, the DNA sequence suggests that it is unlikely that a precursor could extend from the C-terminus because several other termination codons exist in the 3' RNA direction. Furthermore, the nucleotide sequence which could code for a protein of 190 amino acids begins with the methionine codon AUG, 41 bases after the termination codon of the surface antigen gene (Fig. 5). Screening of the sequence of the entire virus suggested that this may be the gene coding for the major core protein. Further structural

sites at amino acid positions 3, 59 and 146. analysis of the core protein must be carried out before a BamHI Hinfl MboII MnIIGGATCCCAGAGTCAGGGGTCTGTATCTTCCTGCTGGTGGCTCCAGTTCAGGAACAGTAAACCCTGCTCCGAATATTGCCTCTCACATCTC CCTAGGGTCTCAGTCCCCAGACATAGAAGGACGACCACCGAGGTCAAGTCCTTGTCATTTGGGACGAGGCTTATAACGGAGAGTGTAGAG MboI Tacl Hinf1 U1 U1 GTCAATCTCCGCGAGGACTGGGGACCCTGTGACGAACATGGAGAACATCACATCAGGATTCCTAGGACCCCTGCTCGTGTTACAGGCGGG ${\tt CAGTTAGAGGGGGCTCCTGACCCCTGGGACACTGCTTGTACCTCTTGTAGTGTAGTCCTAAGGATCCTGGGGACGACAATGTCCGCCC}$ $Mn\Pi$ Ava2 HindII Hinf Hinf1 HinfI GTTTTCTTGTTGACAAGAATCCTCACAATACCGCAGAGTCTAGACTCGTGGTGGACTTCTCTCAATTTTCTAGGGGGGATCTCCCGTGTG CAAAAAGAACAACTGTTCTTAGGAGTGTTATGGCGTCTCAGATCTGAGCACCACCTGAAGAGTTAAAAGATCCCCCTTAGAGGGCACAC XhaI Ball $Mn\Pi$ Hph Mn/I MnIIApyI Mnll MboII MboI MnIIMnlIGCGTTTTATCATATCCTCTTCATCCTGCTATGCCTCATCTTCTTATTGGTTCTTCTGGATTATCAAGGTATGTTGCCCGTTTGTCC MboII EcoRII MboII HinfT MnIIUl TCTAATTCCAGGATCAACAACAACCAGTACGGACCATGCAAAACCTGCACGACTCCTGCTCAAGGCAACTCTATGTTTCCCTCATGTTG ĀĞATTAAĞĞTCCTAĞTTĞTTĞTTĞĞTCATĞCCCTĞĞTACĞTTTTĞĞACĞTĞCTĞAĞĞACĞAĞTTCCĞTTĞAĞATACAAAĞ<mark>ĞĞAĞ</mark>TACAAC Apyl **EcoRII** CTGTACAAAACCTACGGATGGAAATTGCACCTGTATTCCCATCCCATCGTCTTGGGCTTTCGCAAAATACCTATGGGAGTGGGCCTCAGT GACATGTTTTGGATGCCTACCTTTAACGTGGACATAAGGGTAGGGTAGCAGGCCCGAAAGCGTTTTATGGATACCCTCACCCGGAGTCA HaeIII AluI CCGTTTCTCTTGGCTCAGTTTACTAGTGCCATTTGTTCAGTGGTTCGTAGGGCTTTCCCCCACTGTTTGGCTTTCAGCTATATGGATGAT

AsuI HinfI

GGCAAAGAGAACCGAGTCAAATGATCACGGTAAACAAGTCACCAAGCATCCCGAAAGGGGGGTGACAAACCGAAAGTCGATATACCTACTA

Mbol

Fig. 3 Nucleotide sequence of the HBV surface antigen gene and adjacent regions. Sequence analysis was carried out by the method of Maxam and Gilbert³⁴ as outlined in Fig. 4. The enzyme *Mbo*I is an isoschizomer of *Sau3A*. U1 indicates *Sau₉₆₁* sites.

G GAU CCC AGA GUC AGG GGU CUG UAU CUU CCU GGU GGU GGC UCC AGU UCA GGA ACA GUA AAC CCU GCU CCG AAU AUU GCC UCU CAC AUC met glu asn ile thr ser gly phe leu gly pro leu leu val leu gln ala UCG UCA AUC UCC GCG AGG ACU GGG GAC CCU GUG ACG AAC AUG GAG AAC AUC ACA UCA GGA UUC CUA GGA CCC CUG CUC GUG UUA CAG GCG 30 gly phe phe leu leu thr arg ile leu thr ile pro gln ser leu asp ser trp trp thr ser leu asp phe leu gly gly ser pro val GGG UUU UUC UUG UUG ACA AGA AUC CUC ACA AUA CCG CAG AGU CUA GAC UCG UGG UGG ACU UCU CUC AAU UUU CUA GGG GGA UCU CCC GUG cys leu gly gln asn ser gln ser pro thr ser asn his ser pro thr ser cys pro pro ile cys pro gly tyr arg trp met cys leu ugu cuu ggc caa aau ucg cag ucc cca acc ucc aau cac uca acc ucc ugu ccu ugu ccu cca auu ugu ccu ggu uau cgc ugg aug ugu cuu arg arg phe ile ile phe leu phe ile leu leu leu cys leu ile phe leu leu val leu leu asp tyr gln gly met leu pro val cys cgg cgu uuu auc aua uuc cuc uuc auc cug cug cua ugc cuc auc uuc uua uug guu cuu cug gau uau caa ggu aug uug ccc guu ugu 130 120 150 cys cys thr lys pro thr asp gly asn cys thr cys ile pro ile pro ser ser trp ala phe ala lys tyr leu trp glu trp ala ser UGC UGU ACA AAA CCU ACG GAU GGA AAU UGC ACC UGU AUU CCC AUC CCA UCG UCC UGG GCU UUC GCA AAA UAC CUA UGG GAG UGG GCC UCA 180 val arg phe ser trp leu ser leu leu val pro phe val gln trp phe val gly leu ser pro thr val trp leu ser ala ile trp met GUC CGU UUC UCU UGG CUC AGU UUA CUA GUG CCA UUU GUU CAG UGG UUC GUA GGG CUU UCC CCC ACU GUU UGG CUU UCA GCU AUA UGG AUG 220 210 met trp trp gly pro ser leu tyr ser ile val ser pro phe ile pro leu leu pro ile phe phe cys leu trp val tyr ile OC AUG UGG UAU UGG GGG CCA AGU CUG UAC AGC AUC GUG AGU CCC UUU AUA CCG CUG UUA CCA AUU UUC UUU UGU CUC UGG GUA UAC AUU UAA

ACC CUA ACA AAA CAA AAA GAU GGG GUU AUU CCC UAA ACU UCA UGG GCU ACA UAA UUG GAA GUU GGG GAA CUU UGC CAC AGG AUC

Fig. 4 The amino acid sequence of the HBV surface antigen gene. The nucleotide sequence presented in Fig. 3 was translated in one reading frame so as to correspond to the known first 19 amino acids of the N-terminus of the HBV surface antigen¹³.

definitive assignment can be made.

It is possible that a putative precursor could extend from the N-terminus. In certain instances, promoter sites in eukaryotic genes may be identified from the DNA sequence. Hogness and Goldberg (personal communication) have postulated that the canonical sequence TATAAATA may interact with RNA polymerase, as E. coli polymerase interacts with the Pribnow box²⁷. Initiation of transcription usually occurs 23 ± 1 nucleotides from this site. We have been unable to find this sequence in the 200 nucleotides immediately preceding the N-terminal methionine. However, more distant variations on the Hogness-Goldberg sequence occur (D. S. Hogness, personal communication; E. Ziff, personal communication and ref. 28). For example, the sequence TATATT is found 184 bases in the 5' RNA direction from the initiator methionine codon of the surface antigen gene and also occurs in the IgG2 light chain²⁹. However, there are no specific data and the site of initiation of transcription and the structure of the initial gene product of the HBV surface antigen gene therefore remain in doubt.

Tiollais and coworkers have defined the single-stranded regions in the DNA of the virus²⁰. From their map, it is clear that

single-stranded regions of DNA extend into the surface antigen gene, but probably not as far as the initiation of transcription. These observations localise the coding strand of the gene on the full-length DNA strand.

The availability of abundant quantities of hepatitis B viral DNA through molecular cloning procedures will allow a comprehensive study of its gene products and the pathology of the virus. The present work enables specific studies on the expression of the gene coding for the major surface antigen to be carried out. The availability of this gene may also provide an alternative means for production of the antigen which in turn can be used as a vaccine against hepatitis B infection.

We thank Drs D. Peterson and G. Vyas for helpful discussions and for making their unpublished data available to us, Drs J. Birnbaum, R. Hirschman, M. Hilleman and A. Tytell for the Dane particle preparation and for helpful discussions, and Drs J. T. Yang and G. R. I. Chang for consultations on the putative structure of the surface antigen protein. This research was supported by a grant from Merck Sharpe and Dohme. H.M.G. is an Investigator of the Howard Hughes Medical Institute.

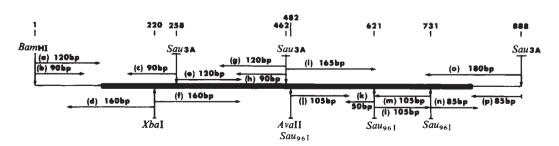


Fig. 5 Sequence strategy for the 889-base pair fragment containing the HBV surface antigen gene. Nucleotide residues are numbered in the direction from 5' to 3' in the message strand, beginning at the BamHI site. The solid bar represents the 678-base pair region coding for the protein. Only the restriction sites used as starting points for sequencing are shown. For labelling, DNA was treated with 5 μ g of bacterial alkaline phosphatase (Worthington) at 37 °C for 60 min in 20 mM Tris-HCl (pH 8.0), phenol extracted and treated with 10 units of T4 polynucleotide kinase (Boehringer) at 37 °C for 30 min in a reaction mixture containing 50 mM glycine (pH 9.5), 10 mM MgCl₂, 10 mM dithiothreitol, 0.1 mM spermidine and 0.001 mM [γ -32P]ATP (\sim 4,000 Ci mmol⁻¹). Fragments labelled only at one end were isolated by gel electrophoresis after either digestion with a restriction enzyme or strand separation. DNA sequence was carried out by the method of Maxam and Gilbert³⁴. Arrows indicate the direction and number of base pairs sequenced in each experiment. Sequences were determined in both strands or in duplicate for all but 50 residues. In this 50-base pair region the data were very reliable. (Sequence data are available from the authors on request.)

Received 28 June: accepted 30 July 1979.

- 1. Krugman, S. in Viral Hepatitis. A Contemporary Assessment of Etiology, Epidemiology, Pathogenesis and Prevention (eds Vyas, G. N., Cohen, S. N. & Schmid, R.) 3-10 (Franklin Institute Press, Philadelphia, 1978).
- Feinstone, S. M., Kapikian, A. Z. & Purcell, R. H. Science 182, 1026-1028 (1973).
- Provost, P. J. et al. Proc. Soc. exp. Biol. Med. 148, 532-539 (1975). Deinhardt, F., Holmes, A. W., Capps, R. B. & Popper, H. J. exp. Med. 125, 673-688 (1967).
- Mascoli, C. C. et al. Proc. Soc. exp. Biol. Med. 142, 276–282 (1973).

 Provost, P. J. & Hilleman, M. R. Proc. Soc. exp. Biol. Med. 160, 213–221 (1979).

 Dane, D. S., Cameron, C. H. & Briggs, M. Lancet i, 695–698 (1970).

- Blumberg, B. S. Science 197, 17-25 (1977).

 Destmyter, J., Ray, M. G., Bradburne, A. F. & Alexander, J. J. in Viral Hepatitis. A Contemporary Assessment of Etiology, Epidemiology, Pathogenesis and Prevention (eds Vyas, G. N., Cohen, S. N. & Schmid, R.) 459-463 (Franklin Institute Press, Philadelphia, 1978).
- 10. Hruska, J. F., Clayton, D. A., Rubenstein, J. L. & Robinson, W. S. J. Virol. 21, 666-672 (1977).
- . Landers, T. A., Greenberg, H. B. & Robinson, W. S. J. Virol. 23, 368-376 (1977)
- 12. Gerin, J. L. & Shi, J. W. K. in Viral Henatitis. A Contemporary Assessment of Etiology Epidemiology, Pathogenesis and Prevention (eds Vyas, G. N., Cohen, S. N. & Schmid, R.)
- Epidemiology, Pathogenesis and Prevention (eds Vyas, G. N., Cohen, S. N. & Schmid, R.)
 147-153 (Franklin Institute Press, Philadelphia, 1978).
 Peterson, D. L., Chien, D. Y., Vyas, G. N., Nitecki, D. & Bond, H. E. in Viral Hepatitis. A Contemporary Assessment of Etiology, Epidemiology, Pathogenesis and Prevention (eds Vyas, G. N., Cohen, S. N. & Schmid, R.) 569-573 (Franklin Institute Press, Philadelphia, 1978).
- 14. Shi, J. W. K. & Gerin, J. L. J. Virol. 21, 347 (1977).

- Magnius, L. O. in Viral Hepatitis. A Contemporary Assessment of Etiology, Epidemiology, Pathogenesis and Prevention (eds Vyas, G. N., Cohen, S. N. & Schmid, R.) 177–182 (Granklin Institute Press, Philadelphia, 1978).

 16. Takahashi, K. et al. J. Immun. 122, 275-279 (1979).

 17. Bolivar, F. et al. Gene 2, 95-113 (1977).

- Bolivar, F. Gene 4, 121-136 (1978).
- 19. Barnes, W. M. Science 195, 393-394 (1977).
- 20. Summers, J. A., O'Connel, A. & Millman, I. Proc. natn. Acad. Sci. U.S.A. 72, 4597-4601
- (1975).
 21. Burrell, C. J., Mackay, P., Greenaway, P. J., Hofschneider, P. H. & Murray, K. Nature 279,
- 22. Sninsky, J. J., Siddiqui, A., Robinson, W. J. & Cohen, S. N. Nature 279, 346-348 (1979). 23. Charnay, P., Pourcel, L. A., Fritsch, A. & Tiollais, P. Proc. natn. Acad. Sci. U.S.A. 76,
- 2222-2226 (1979).
- Robertson, M. S. et al. Nature 278, 370-372 (1979).
 Chou, P. Y. & Fasman, G. Biochemistry 1-222, 222-245 (1974).
- 26. Winzler, R. J. in Hormonal Proteins and Peptides (ed. Li, C. H.) 1-15 (Academic, New York,
- 27. Pribnow, D. Proc. nam. Acad. Sci. U.S.A. 72, 784-788 (1975).
- Ziff, E. B. & Evans, R. M. Cell 15, 1463-1475 (1978)
- Tonegawa, S., Maxam, A. M., Tizard, R., Bernard, O. & Gilbert, W. Proc. natn. Acad. Sci. U.S.A. 75, 1485-1489 (1978).
- Ullrich, A. et al. Science 196, 1313-1319 (1977).
- Clewell, D. B. & Helinski, D. R. Proc. natn. Acad. Sci. U.S.A. 62, 1159-1166 (1969).
 Southern, E. M. J. molec. Biol. 98, 503-517 (1975).
- Rigby, P. W. J., Dieckman, M., Rhodes, C. & Berg, P. J. molec. Biol. 113, 237-251 (1977).
- Maxam, A. & Gilbert, W. Proc. natn. Acad. Sci. U.S.A. 74, 560-564 (1977).

letters

Discovery of IR bursts from Liller I/MXB1730 – 333

THE first detection of IR bursts from the object known as Liller I which has been identified² with the rapid X-ray burster MXB1730-333 is reported here. The rapid burster is the only one of its kind that has been extensively studied in the X-ray region³⁻⁵ and some of its known characteristics can be summarised as follows. When it is active it gives several thousand X-ray bursts per day. It operates in two modes: mode I (usually in March and September every year), mode II (usually in April and October every year). In mode I it gives large X-ray bursts with energy in the range 10^{39} – 10^{40} erg and also small bursts with energy in the range of 10^{38} – 10^{39} erg; while in mode II it primarily emits bursts with energy typically close to 10³⁹ erg. These bursts are called type II on the basis of the constant character of their spectrum during the decay phase. The rapid burster occasionally also gives bursts known as type I which are characterised by the softening of the X-ray spectrum during the decay phase. The energy in these bursts is usually in excess of 10³⁹ erg and their frequency is about one every few hours.

Liller¹ searched for an optical counterpart of the rapid burster and found a highly reddened compact cluster with a red magnitude $m_r = 21$ in the error box of the rapid burster⁶. This has been identified as a globular cluster by Kleinmann et al.7, with the help of IR observations. The identification of this cluster with the rapid burster was supported by Doxsey et al.2 who obtained a more precise position for the X-ray burster. Kleinmann et al.⁷ observed the globular cluster continuously for about an hour on 29 May 1976 and did not detect any variation greater than 10% of the total IR flux. However, the rapid burster may have ceased activity in April 1976.

Two of us (K.M.V.A. and S.M.C.)⁸ have proposed a model for the X-ray emission from the rapid burster based on accretion of matter onto the poles of a rotating magnetic neutron star. More recently (unpublished data), the cyclotron emission of electrons in the accreting column above the poles of the magnetised

neutron star has been considered and emission in the IR and optical regions suggested; these calculations have prompted us to undertake the present experiment. The optical emission from the rapid burster may, however, be difficult to detect because of the large interstellar extinction^{1,7} of $A_v = 11 \pm 1$ mag. On the other hand, the interstellar extinction in the IR is very small which makes it possible to observe the radiation.

The observations were made on the 1-m telescope of the Indian Institute of Astrophysics at Kavalur (lat. 12° N), India on the night of 4-5 April 1979 between 21.29 and 00.10 UT. At the Cassegrain focus a liquid-nitrogen cooled photometer was used with a set of IR filters, three apertures and a Fabry mirror imaging the primary of the telescope on a 0.5 mm × 0.5 mm InSb detector. The incoming beam was f/20 giving the scale of 10 arc s mm⁻¹ at the focal plane. For the present measurement we used a 2 mm aperture corresponding to 20 arc s field of view. The incoming beam was chopped alternately on the source and a neighbouring part of the sky with a throw of 20 arc s by a tertiary mirror inclined at 45° to the incident beam with a frequency of 16 Hz. The servo-controlled square wave gave a duty cycle of 75%. The Liller I signal was detected with the phase sensitive detector followed by the low pass filter with a time constant of 0.6 s. In the case of standard stars, the phase sensitive detector

Table 1 IR bursts from Liller I/MXB1730-333

Burs No.	st o	Time of occurrence (UT)			Dura- tion (a)	FWHM (s)	Peak luminosity (10 ³⁷ erg s ⁻¹ in 0.3 μm interval at 1.6 μm)	Burst energy (10 ³⁸ erg in 0.3 µm interval at 1.6 µm)
1	21 h	33 min	00 s	2	37	12	2.2	3.0
2	22	41	24	2	36	12	1.9	2.0
3	23	45	06	3	36	12	2.0	2.6
4	23	47	30	3	24	9	1.7	1.5
5	00	03	00	2	36	10	2.2	2.5
6a	00	04	54	2	20	6	2.3	1.7
6b	00	05	12	2	19	6	1.6	0.9