Structural and functional characterization of a monoclonal antibody specific for the preS1 region of hepatitis B virus

Juan Carlos Pizarro[a], Brigitte Vulliez-le Normand[a], Marie-Madeleine Riottot[a], Agata Budkowska[b], Graham A. Bentley[a,*]

[a] *Unité d'Immunologie Structurale, Institut Pasteur, 25 rue du Dr. Roux, 75724 Paris Cedex 15, France*
[b] *Laboratoire d'Epidémologie Moléculaire des Entérovirus, Institut Pasteur, 25 rue du Dr. Roux, 75724 Paris Cedex 15, France*

Received 24 September 2001; revised 6 November 2001; accepted 6 November 2001

First published online 30 November 2001

Edited by Hans Eklund

Abstract The monoclonal antibody 5a19, raised against the *ay* serotype of hepatitis B virus, binds to the segment of the preS1 region comprising residues 37–43, which is implicated in attachment of the virus to hepatocytes. The dissociation constant, derived from kinetic studies using surface plasmon resonance techniques, is in the low nanomolar range. The nucleotide sequence of the variable domains has been determined and the corresponding germ-line genes have been identified. The three-dimensional structure of the Fab fragment has been determined by X-ray crystallography to 2.6 Å resolution.

Key words: preS1; Monoclonal antibody; Epitope mapping; Binding kinetics; Crystal structure

1. Introduction

Hepatitis B virus (HBV), a member of the *Hepadnaviridae* family, is a major cause of chronic liver disease and is one of the principle etiological factors associated with hepatocellular carcinoma [1]. The infectious viral particle contains a nucleocapsid with icosahedral symmetry enclosing a circular, partially double-stranded DNA genome (about 3.2 kb depending on serotype) and the P protein, which has DNA polymerase activity. The nucleocapsid itself is enveloped by a lipid bilayer into which are inserted three related transmembrane glycoproteins, S, M and L [2]. The genes encoding the three glycoproteins begin with distinct initiation codons, but share the same reading frame and stop codon. Thus, the protein S, the smallest envelope protein, is 226 amino acids in length. The protein M, in turn, carries and additional 55 amino acid segment, the preS2 region, in a position N-terminal to the S sequence. Similarly, the protein L is composed of the preS1 region, 108 or 119 amino acid residues in length depending on the serotype, positioned N-terminal to the M sequence. Two other forms of viral particles have been characterized: the spherical and filamentous particles. Being composed of the envelope material only (lipid and envelope glycoprotein), they are non-infectious.

Although HBV shows a strong tropism towards hepatocytes, it can invade other human cell lines. Nonetheless, only infection of the liver appears to lead to clinical pathology for hepatitis B [3]. Diverse candidates for the viral receptor have been proposed: the IgA receptor, glyceraldehyde-3-phosphate dehydrogenase, asialoglycoprotein receptor, annexin V, and several other less well characterized proteins [3]. The region of the virus involved in attachment to the hepatocyte, on the other hand, has been clearly identified as the preS region of the envelope proteins. In particular, the preS1 peptide fragment comprising residues 21–47 (preS1(21–47)), as well as antibodies induced by this peptide inhibit HBV binding to HepG2 cells [4,5], implicating this region of the L envelope protein in the virus–host cell interaction.

The murine monoclonal antibody (mAb) 5a19 recognizes a linear epitope preS1(36–43) [6] and inhibits invasion of cultured human hepatocytes [7]. Here, we present the sequence of the heavy and light chain variable domains (V_H and V_L, respectively), kinetic studies of antigen-binding, and fine epitope mapping of the preS1 region recognized by 5a19. We also describe the three-dimensional structure of the Fab fragment of 5a19, determined by X-ray crystallography to a resolution of 2.6 Å.

2. Materials and methods

2.1. Preparation and purification of IgG and Fab fragments

The mAb 5a19 (isotype IgG1,κ) was obtained by immunizing BALB/c mice with HBV particles of the *ay* serotype [6]. The immunoglobulin was purified from ascitic fluid by precipitation with 40% ammonium sulfate at pH 7.4, followed by ion-exchange chromatography on a DEAE-Sephacel (Pharmacia) column using an incrementing NaCl concentration gradient (0–500 mM) in Tris buffer at pH 8.0. Antibody fractions were pooled and dialyzed across a microsep 10K (Filtron) against 0.1 M potassium phosphate at pH 7.2.

After addition of mercaptoethanol (14 mM) and EDTA (6 mM), the IgG was treated with papain at an enzyme:substrate ratio of 1:300 at 37°C for 2 h. After incubating with 20 mM iodoacetamide for 30 min at room temperature, the buffer was changed to 0.02 M potassium phosphate at pH 7.8. The protein was then added to a DEAE-Sephacel (Pharmacia) column equilibrated with the same buffer. Fab 5a19 was eluted by the running phosphate buffer.

2.2. Epitope mapping

Peptides bound to cellulose were prepared by automated synthesis (SPOTs, Genosys Biotechnologies, Cambridge, UK). The cellulose membranes were treated overnight at 4°C with a blocking agent (supplied by the manufacturer) placed in Tris-buffered saline (50 mM Tris–HCl at pH 7.6, 150 mM NaCl) and 0.05% Tween 20, with 5% sucrose. After washing with Tween/Tris-buffered saline (T-TBS), the membranes were incubated for 1 h at room temperature with IgG

*Corresponding author. Fax: (33)-1-45 68 86 39.
E-mail address: bentley@pasteur.fr (G.A. Bentley).

PII: S0014-5793(01)03190-8

5a19 (1 μg/ml in blocking buffer). After extensive washing in T-TBS, the membranes were then incubated for 1 h with a 1:5000 dilution of peroxidase-conjugated sheep anti-mouse IgG, H+L (ICN). Following a final wash with T-TBS, ECL detection reagent was added for 30 s before exposing to Hyperfilm (Amersham).

2.3. Kinetic analysis of 5a19-binding

The association (k_{on}) and dissociation (k_{off}) rate constants were measured by surface plasmon resonance using a BIAcore 2000 instrument (BIAcore AB, Uppsala, Sweden). Two series of experiments were performed, one in which the antigenic peptide was fixed to the sensor chip, and the other where IgG 5a19 was fixed.

1. Sensor-fixed antigen: the biotinylated peptide preS1(21–48) was immobilized on a streptavidin-coated SA chip (BIAcore AB) and the Fab 5a19 was used as the analyte. Association and dissociation rate constants of the analyte were measured at 25°C using Fab concentrations ranging from 0.01 to 1 μM in HBS-EP buffer (10 mM HEPES pH 7.4, 150 mM NaCl, 3 mM EDTA and 0.005% Tween 20), and a flow rate of 50 μl/min.
2. Sensor-fixed antibody: IgG 5a19 was immobilized on the sensor chip and the biotinylated peptide preS1(21–48) was used as the analyte. The antibody was covalently fixed to a CM5 chip by its primary amine groups using an amine-coupling kit (BIAcore AB). IgG 5a19 was then injected at a concentration of 30 μg/ml in 10 mM sodium acetate at pH 5.0. The immobilization procedure was completed by a 35 μl injection of 1 M ethanolamine/HCl mixture at pH 8.5 to block remaining active groups present on the matrix. To avoid potential effects of mass transport, kinetic measurements were made with the biotinylated peptide at low concentrations, ranging from 1 pM to 100 μM, in a HBS-EP buffer and with a flow rate of 50 μl/min at 25°C.

In both series of measurements, the chip was regenerated by a 30 s injection pulse with 10 mM HCl, and one background surface on the chip was retained for control runs. Kinetic constants were derived with the BIAEVALUATION 3.01 software (BIAcore AB), fitting k_{on} and k_{off} separately, and using background-subtracted sensorgrams.

2.4. Sequencing of the variable domains of 5a19

Total RNA was extracted from about 10^7 5a19 hybridoma cells by guanidinium thiocyanate extraction using Tri reagent (Sigma). The RNA was used as substrate for cDNA synthesis using an oligo-dT 15-mer primer (Roche Molecular Biochemicals) and Moloney murine leukemia reverse transcriptase (Gibco BRL). cDNA for the Vκ domain was amplified by PCR using LC7 (5′-CCA GGT CCG AGC TCG TGA TGA CAC AGT CTC CA-3′) and Cκ_2 (5′-GAT GGA TAC AGT TGG TGC AGC ATC AGC CCG-3′) primers. The V_H domain was amplified with the primers V_HIIID (5′-GAA GTG CAG CTC GAG GAG TCT GGG GG-3′) and C_H (5′-GGC AGC GAT CCA GGG GCC AGT GG-3′). The PCR products were cloned and each sequence was analyzed from at least three different colonies.

2.5. Crystallization and diffraction data collection

Crystals of Fab 5a19 were grown by vapor diffusion using the hanging drop technique. A volume of 2 μl of protein was mixed with 0.5 μl of 0.1 M cysteine and 0.5 μl of buffer containing 15% (w/v) PEG 6000, 110 mM $MgCl_2$, 60 mM Tris–HCl, pH 8.5, and 15% glycerol. The drop, with a final protein concentration of 4.7 mg/ml, was sealed over a reservoir containing 1 ml of buffer. The crystals, which appeared after 1–2 weeks at 17°C, belong to the space group $P2_12_12_1$ with unit cell parameters $a = 47.01$ Å, $b = 88.46$ Å, $c = 111.67$ Å. The unit cell volume is consistent with one Fab molecule in the asymmetric unit.

Diffraction data were collected at the beam line D41 at LURE using a MarResearch imaging plate detector (18 cm diameter). A complete data set to 2.6 Å resolution was obtained from single crystal at ambient temperature using a monochromated X-ray wavelength of 1.375 Å. Diffraction intensities were integrated with the programme HKL [8], yielding a total of 14823 unique reflections ($R_{merge} = 7.7\%$, redundancy = 6.5).

2.6. Determination and refinement of the structure

A preliminary model of Fab 5a19 was obtained by molecular replacement using the programme AmoRe [9]. A variable domain dimer (V_L/V_H) and a constant domain dimer (C_L/C_H1), obtained from Protein Data Bank (PDB) entries 1hil and 1fdl, respectively, were used as independent search models in the rotation and one-body translation functions. A two-body translation search assembled the independently determined solutions for the (V_L/V_H) and (C_L/C_H1) dimers to give a sensible pairing of the light and heavy chains within the single Fab fragment present in the asymmetric unit, thus confirming the correctness of the rotation and translation function results. In addition, the molecular packing of the model in the unit cell was free of steric hindrance. The position and orientation of the variable and constant domain dimers were optimized by rigid-body refinement. After introducing the derived 5a19 sequence into the molecular replacement model, refinement of the atomic parameters was carried out with the programme X-PLOR [10] using data between the resolution limits 15.0 and 2.6 Å (5% of these data were randomly selected for calculating R_{free}). Bulk solvent corrections were applied using the procedure described by Jiang and Brünger [11].

3. Results and discussion

3.1. Epitope mapping

Peptide fragments contained within the segment preS1(31–47) from the *ay* serotype were synthesized to perform epitope mapping using the SPOTs system. The results of this analysis

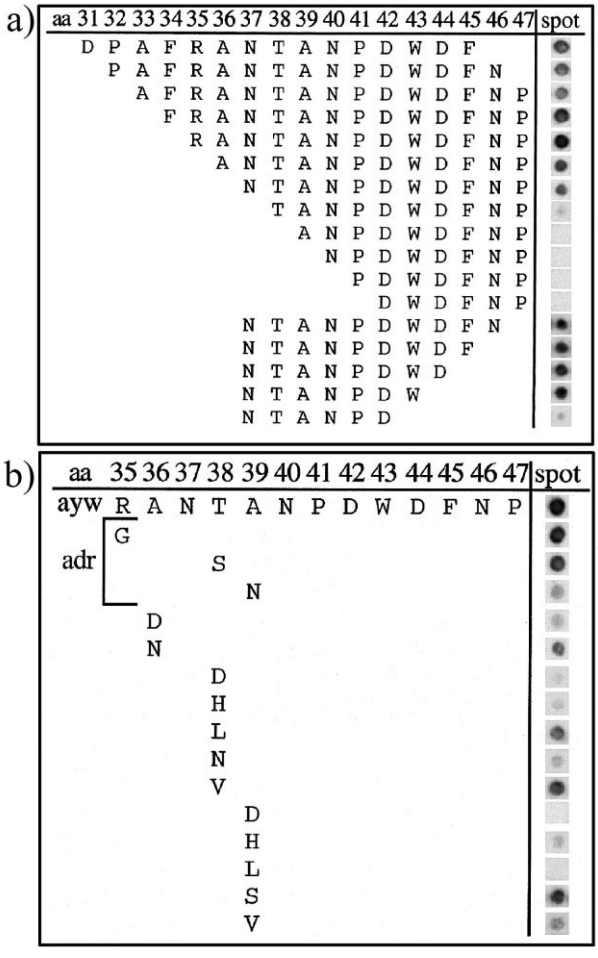

Fig. 1. Epitope mapping of 5a19 using the SPOTs test. The intensity of the spot indicates the avidity of the antibody. (a) Recognition of 6–15-mer peptides spanning the sequence 31–47 of the preS1 region (*ay* serotype). (b) The specificity of 5a19 was examined by substitution of each serotype-dependent residue by the one corresponding to the *ad* serotype (Gly35, Ser38, Asn39). See Section 3 for other mutations made at positions 36, 38 and 39.

are visualized in Fig. 1. The patterns of recognition by 5a19 of peptides truncated from the N- and C-termini indicate that the minimum epitope is the segment preS1(37–43). Alanine scanning across this seven-residue segment showed no significant effect on the ability of 5a19 to recognize the epitope (data not shown). The substitutions Arg35→Gly, Thr38→Ser and Ala39→Asn, corresponding to changes to the *ad* serotype, were also tried separately, but only the substitution at position 39 showed a reduction in recognition, consistent with the previously reported preference of 5a19 for the *ay* serotype [6].

The sensitivity to different residue types was further probed at positions 36, 38 and 39. The substitution Ala36→Asp, but not Ala36→Asn, reduces 5a19-binding. Substitution of Thr38 by the polar residues Asp, His and Asn showed a reduction in recognition. Substitution with Ser, Ala, Leu or Val, by contrast, showed no difference with respect to the native sequence, showing that hydrophobic residues of different size produced no observable effect on the binding of 5a19. Changes in reactivity to substitutions at Ala39 were sensitive to charge since either reduced binding, or none at all, was observed when this residue was replaced by His or Asp. Size of the side chain of residue 39 seems critical since reactivity is lost by substitution with Leu, but is observable if replaced by Val or Ser. Residue W43 was also tested by substitutions with Asp and His, but no difference with respect to wild-type *ay* antigen was observed (data not shown).

3.2. Kinetic analysis of 5a19-binding

1. Immobilized biotinylated preS1(21–48) using Fab 5a19 as analyte: the monovalent Fab fragment of 5a19 was used instead of the bivalent IgG molecule to avoid potential artefacts arising from avidity. The net surface resonance signal after addition of the biotinylated preS1(21–48) peptide was 25 RU (resonance units – 1 RU = 1 pg/mm^2) and the kinetic constants were calculated by fitting a simple bimolecular reaction model to the data. The averaged values for k_{on} and k_{off} were 1.2 ($\pm$0.6)$\times 10^5$ M^{-1} s^{-1} and 1.9 ($\pm$0.3)$\times 10^{-4}$ s^{-1}, respectively, giving an estimated affinity constant ($K_D = k_{off}/k_{on}$) of 1.6 ($\pm$0.8)$\times 10^{-9}$ M.

2. Immobilized antibody using biotinylated preS1(21–48) as analyte: mAb 5a19 was immobilized on the CM5 sensor chip surface at a level of 4600 RU. The analyte concentrations used were 5, 10, 50, 100 nM, corresponding to a response of 10–85 RU and the kinetic constants were also calculated from the data by assuming a simple bimolecular reaction. The averaged values for k_{on} and k_{off} were 5.8 ($\pm$1.2)$\times 10^4$ M^{-1} s^{-1} and 2.7 ($\pm$0.2)$\times 10^{-4}$ s^{-1}, respectively, giving an estimated dissociation constant of 4.6 ($\pm$1.0)$\times 10^{-9}$ M.

Although the estimates of K_D were obtained from two independent experimental procedures, they differ by less than three standard deviations, the major difference arising from the values of k_{on}. Reasons for this difference are difficult to identify, but may be related to the different kinetic behavior of the analyte in each case.

3.3. Sequence of the variable domains of 5a19

The variable domain sequences of 5a19 are given in Fig. 2. V_L is derived from the Vκ8-21 germ-line sequence [12], having a total of nine nucleotide changes due to somatic mutation, three of which are silent. This leads to the following changes in the amino acid sequence: Ala15→Val, Lys24→Arg, Ser27→Thr, Tyr36→Phe, Leu46→Leu and Asn93→Ser. The framework region FR4, encoded by the Jκ2 gene, has not been subjected to somatic mutations. Although no germ-line gene for V_H could be found, a 96.9% sequence identity occurs with the V_H domain of the IgM D2 [13], which belongs to the V_H7183 gene family. Of the seven nucleotide changes, two are silent and the remaining five give rise to the amino acid changes Ile51→Val, Gly53→Asp, Thr57→Ala, Val63→Leu and Arg94→Ser. These probably reflect real germ-line changes since IgM immunoglobulins are rarely subject to somatic hypermutation. The CDR-H3 and FR4 regions of V_H are encoded by the DQ52 (used in the second reading frame) and J_H3 minigenes, respectively, with some nucleotide deletions and additions occurring at the junctions.

The V_H and V_L nucleotide sequences have been deposited in the EMBL nucleotide sequence database with accession numbers AJ344367 and AJ344368, respectively.

3.4. Structure of 5a19

After completion of the refinement of the atomic parameters, the R and R_{free} values were 0.213 and 0.291, respectively, for a model comprising 3347 protein atoms and 132 solvent positions. The average protein temperature factor is 35.4 Å^2, and the root mean square deviation from ideal bond length and angles is 0.005 Å and 1.2°, respectively. Coordinates have been deposited in the PDB with entry code 1gm3.

The main chain of 5a19 could be traced without ambiguity for both the H and L chains, except for the region H128–H133, and the two C-terminal regions L212–L214 and H214–H215; these regions were not included in the final model. The side chain of Trp-H97, belonging to CDR-H3, had ill-defined electron density but was modelled as two discrete conformers. A total of 98.7% of the residues have main-chain conformation angles (φ–ψ) falling within the favored region of the Ramachandran plot (as defined by the programme PROCHECK [14]). Residue Ala-L51 forms part of a γ-turn characteristic of the unique canonical conformation of CDR-L2, and thus falls in an energetically unfavored region of the main-chain (φ–ψ) plot. Residues contributed by all six CDRs of Fab 5a19 make extensive contacts with a neighboring molecule, burying a total surface of 938 Å^2 from the solvent, thus precluding diffusion of the peptide ligand through the crystal lattice to the antigen-binding site.

The hypervariable regions CDR-L1, CDR-L2, CDR-H1 and CDR-H2 conform to the canonical conformations [15,16], belonging to the classifications 3 (17 residues), 1 (seven residues), 1 (five residues) and 3 (six residues), respectively (CDR lengths used are those defined in [15]). CDR-L3 has residue 95 deleted to give a hypervariable region of eight amino acids. This is due to a codon loss when the V_L and J gene segments were joined during somatic recombination. Although there is no accepted canonical classification to describe the CDR-L3 structure of 5a19, its conformation is the same as that observed in the anti-CD25 mAb CRIS [17] and the catalytic antibody 17E8 [18]. CDR-H3 is seven residues in

(a)

```
                                                                                                     ===========
             1   2   3   4   5   6   7   8   9   10  11  12  13  14  15  16  17  18  19  20  21  22  23  24  25  26
             Glu Leu Val Met Thr Gln Ser Pro Ser Ser Leu Ala Val Ser Val Gly Glu Lys Val Thr Met Ser Cys Arg Ser Ser
VL 5a19                                      TCC TCC CTG GCT GTG TCA GTA GGA GAG AAG GTC ACT ATG AGC TGC AGA TCC AGT
VL Y15982                                    --- --- --- --- --- --- -C- --- --- --- --- --- --- --- --- -A- --- ---
                                              *   *   *   *   *   *  Ala  *   *   *   *   *   *   *   *  Lys  *   *

             =================== CDR1 ==================================================================================
             27  A   B   C   D   E   F   28  29  30  31  32  33  34  35  36  37  38  39  40  41  42  43  44  45  46
             Gln Ser Leu Leu Asn Thr Arg Thr Arg Lys Ser Tyr Leu Ala Trp Phe Gln Gln Lys Pro Gly Gln Ser Pro Lys Met
VL 5a19      CAG AGT CTG CTC AAC ACT AGA ACC CGA AAG AGC TAC TTG GCT TGG TTC CAG CAG AAA CCA GGG CAG TCT CCT AAA ATG
VL Y15982    --- --- --- --- --- -G- --- --- --- --- --- --- --- --- --- -A- --- --- --- --- --- --- --- --- --- C--
              *   *   *   *   *  Ser  *   *   *   *   *   *   *   *   *  Tyr  *   *   *   *   *   *   *   *   *  Leu

                         =========== CDR2 ==========
             47  48  49  50  51  52  53  54  55  56  57  58  59  60  61  62  63  64  65  66  67  68  69  70  71  72
VL 5a19      Leu Ile Tyr Trp Ala Ser Thr Arg Glu Ser Gly Val Pro Asp Arg Phe Thr Gly Ser Gly Ser Gly Thr Asp Phe Thr
VL Y15982    TTG ATC TAC TGG GCA TCC ACT AGG GAA TCT GGG GTC CCT GAT CGC TTC ACA GGC AGT GGA TCT GGG ACA GAT TTC ACT
             C-- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
              *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *

                                                                             ================ CDR3 =================
             73  74  75  76  77  78  79  80  81  82  83  84  85  86  87  88  89  90  91  92  93  94  96  97  98  99
VL 5a19      Leu Thr Ile Ser Ser Val Gln Ala Glu Asp Leu Ala Val Tyr Tyr Cys Lys Gln Ser Tyr Ser Leu Tyr Thr Phe Gly
VL Y15982    CTC ACC ATC AGC AGT GTG CAG GCT GAA GAC CTG GCA GTT TAT TAC TGC AAA CAA TCT TAT AGT CTG TAC ACG TTC GGA
             --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --G --- --- --- -A- --T
              *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *  Asn  *
```

```
             96  97  98  99  100 101 102 103 104 105 106 107
             Tyr Thr Phe Gly Gly Gly Thr Lys Leu Glu Ile Lys
VL 5a19      TAC ACG TTC GGA GGG GGG ACC AAG CTG GAA ATA AAA
Jk2          --- --- --- --- --- --- --- --- --- --- --- ---
              *   *   *   *   *   *   *   *   *   *   *   *
```

(b)

```
             1   2   3   4   5   6   7   8   9   10  11  12  13  14  15  16  17  18  19  20  21  22  23  24  25  26
             Glu Val Gln Leu Glu Glu Ser Gly Gly Gly Leu Val Lys Pro Gly Gly Ser Leu Lys Leu Ser Cys Ala Ala Phe Gly
VH 5a19                                          GGC TTA GTG AAG CCT GGA GGG TCC CTG AAA CTC TCC TGT GCA GCC TTT GGA
VH M83724                                        --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --C ---
                                                  *   *   *   *   *   *   *   *   *   *   *   *   *   *   *  Ser  *

                             ======= CDR1 ======                                                     ===========
             27  28  29  30  31  32  33  34  35  36  37  38  39  40  41  42  43  44  45  46  47  48  49  50  51  52
             Phe Thr Phe Ser Ser Tyr Ala Met Ser Trp Val Arg Gln Ser Pro Glu Lys Arg Leu Glu Trp Val Ala Glu Val Ser
VH 5a19      TTC ACT TTC AGT AGT TAT GCC ATG TCT TGG GTT CGC CAG TCT CCA GAG AAG AGG CTG GAG TGG GTC GCA GAA GTT AGC
VH M83724    --- --- --- --- --C --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- A-- --T
              *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *  Ile  *

             =================== CDR2 ==============================
             A   53  54  55  56  57  58  59  60  61  62  63  64  65  66  67  68  69  70  71  72  73  74  75  76  77
             Ser Asp Gly Ser Tyr Ala Tyr Tyr Pro Asp Thr Leu Thr Gly Arg Phe Thr Ile Ser Arg Asp Asn Ala Lys Asn Thr
VH 5a19      AGT GAT GGT AGT TAC GCC TAC TAT CCA GAC ACT TTG ACG GGC CGA TTC ACC ATC TCC AGA GAC AAT GCC AAG AAC ACC
VH M83724    --- -G- --- --- --- A-- --- --- --- --- --- G-- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
              *  Gly  *   *   *  Thr  *   *   *   *   *  Val  *   *   *   *   *   *   *   *   *   *   *   *   *   *

                                                                                         =========== CDR3 ======
             78  79  80  81  82  A   B   C   83  84  85  86  87  88  89  90  91  92  93  94  96  97  98  99  100 101
             Leu Tyr Leu Glu Met Thr Ser Leu Arg Ser Glu Asp Thr Ala Met Tyr Tyr Cys Ala Ser Phe Asn Trp Asp Val Ala
VH 5a19      CTG TAC CTG GAA ATG ACC AGT CTG AGG TCT GAG GAC ACG GCC ATG TAT TAC TGT GCA AGT TTT AAC TGG GAC GTC GCT
VH M83724    --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --G
              *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *  Arg

             ===
             102 103 104 105 106 107 108 109 110 111 112 113 114
             Tyr Trp Gly Gln Gly Thr Leu Val Thr Val Ser Ala Ala
VH 5a19      TAC TGG GGC CAA GGG ACT CTG GTC ACT GTC TCT GCA GCC
```

```
             96  97  98  99  100 101 102 103 104 105 106 107 108 109 110 111 112 113
             Phe Asn Trp Asp Val Ala Tyr Trp Gly Gln Gly Thr Leu Val Thr Val Ser Ala
VH 5a19      TTT AAC TGG GAC GTC GCT TAC TGG GGC CAA GGG ACT CTG GTC ACT GTC TCT GCA
IGHD-Q52       C --- --- ---
JH3                      TGG T-T --- --- --- --- --- --- --- --- --- --- --- --- ---
                         Trp Phe  *   *   *   *   *   *   *   *   *   *   *   *   *
```

Fig. 2. Nucleotide and amino acid sequences for the variable domain of 5a19. The amino acid residue numbering and CDR limits are according to [20]. Differences with germ-line sequences and the corresponding amino acid sequences are shown. N-terminal residues falling in regions covered by the primers were identified by amino acid sequencing. (a) 5a19 V_L domain is compared with Vκ8-21 germ-line (GenBank entry Y15982). FR4 is compared with Jκ2. (b) 5a19 V_H domain is compared with IgM D2 [13] (GenBank entry M83724). CDR-H3 and FR4 are compared with DQ52 and J_H3 gene segments.

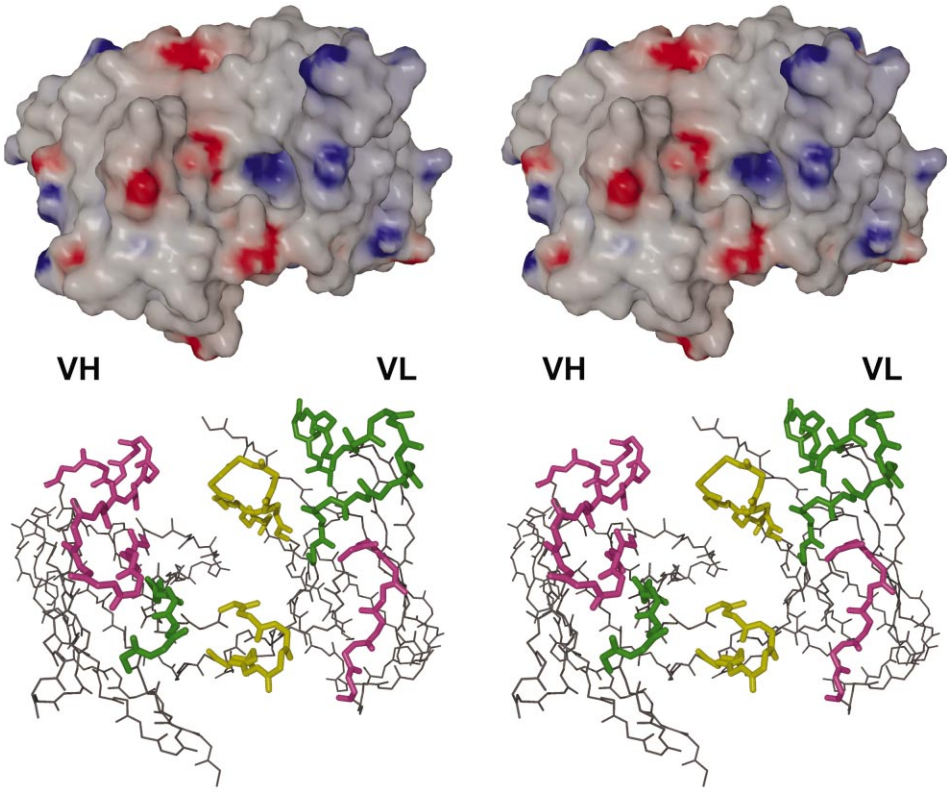

Fig. 3. A stereo view of the antigen-binding site of mAb 5a19, with V_H on the left and V_L on the right. Upper panel: surface representation with color coded for electrostatic potential, with red for negative charge and blue for positive charge. Lower panel: conformation of the antigen-binding loops, with CDR-1 in green, CDR-2 in mauve and CDR-3 in yellow. The figures were prepared with GRASP [21] and VMD [22], and both rendered with Raster3D [23].

length and is thus slightly shorter than the mean length of 8.7 residues noted by Wu et al. [19].

Hypermutation has introduced changes in the germ-line sequence that affect the antigen-binding site. The Vκ gene presents two changes in CDR-L1 (residues L24 and L27E) and one change in CDR-L3 (residue 93). V_H gene changes occur in CDR-H2 (residue H53) and CDR-H3 (residue H94), although the latter change could have arisen from the V-D recombination joining event rather than the latter somatic recombination event (see Fig. 2). Only Ser-L93 and Asp-H53, however, are favorably located in the antigen-binding site to form potential contacts to the antigen.

The antigen-binding site of 5a19 takes the form of a groove of length 27 Å, width 20 Å and depth 15 Å. The electrostatic potential (Fig. 3) is significantly negative in the groove, but with an important positive contribution on the side from CDR-L1. The groove-like topology, rare in anti-protein antibodies, is more reminiscent of that found in anti-peptide antibodies and would be consistent with the recognition by 5a19 of a linear epitope in the preS1 region. We are attempting to co-crystallize Fab 5a19 with peptide fragments containing the minimal epitope determined by our studies reported here to understand the nature of the interaction of 5a19 with the viral antigen.

Acknowledgements: We thank Dr. A. Lewit-Bentley, LURE, Orsay, France, for assistance on the beam-line D41. This research was supported by funds from the Institut Pasteur, the Centre National de la Recherche Scientifique, the Association pour la Recherche sur le Cancer and the Fondation pour la Recherche Médicale.

References

[1] Schafer, D.F. and Sorrel, M.F. (1999) Lancet 353, 1253–1257.
[2] Heermann, K.H., Goldman, U., Schwartz, W., Seyyfarth, T., Baugarten, H. and Gerlich, W.H. (1984) J. Virol. 52, 396–402.
[3] De Meyer, S., Gong, Z.J., Suwandhi, W., van Pelt, J., Soumillion, A. and Yap, S.H. (1997) J. Viral Hepat. 4, 145–153.
[4] Neurath, A.D., Kent, S.B., Strick, N. and Parker, K. (1986) Cell 46, 428–436.
[5] Neurath, A.D., Seto, B. and Strick, N. (1989) Vaccine 7, 234–236.
[6] Budkowska, A., Bedossa, P., Groh, F., Louise, A. and Pillot, J. (1995) J. Virol. 69, 840–848.
[7] Maeng, C.Y., Ryu, C.J., Gripon, P., Guguen-Guillouzo, C. and Hong, H.J. (2000) Virology 270, 9–16.
[8] Otwinowski, Z. and Minor, W. (1997) in: Methods in Enzymology: Macromolecular Crystallography (Carter, C.W. and Sweet, R.M., Eds.), Vol. 276, pp. 307–326, Academic Press.
[9] Navaza, J. (1994) Acta Crystallogr. A50, 157–163.
[10] Brünger, A.T., Kuriyan, J. and Karplus, M. (1987) Science 235, 458–460.
[11] Jiang, J.S. and Brünger, A.T. (1994) J. Mol. Biol. 243, 100–115.
[12] Kirschbaum, T., Pourrajabi, S., Zocher, I., Schwendinger, J., Heim, V., Roschenthaler, F., Kirschbaum, V. and Zachau, H.G. (1998) Eur. J. Immunol. 28, 1458–1466.
[13] Taub, R., Hsu, J.C., Garsky, V.M., Hill, B.L., Erlanger, B.F. and Kohn, L.D. (1992) J. Biol. Chem. 267, 5977–5984.
[14] Laskowski, R.A., Mc Arthur, M.W., Moss, D.S. and Thornton, J.M. (1993) J. Appl. Crystallogr. 26, 283–291.
[15] Chothia, C., Lesk, A.M., Tramontano, A., Lewit, M., Smith-Gill,

S.J., Air, G., Sheriff, S., Padlan, E.A., Davies, D., Tulip, W.R., Colman, P.M., Spinelli, S., Alzari, P.M. and Poljak, R.J. (1989) Nature 342, 877–883.
[16] Al-Lazikani, B., Lesk, A.M. and Chothia, C. (1997) J. Mol. Biol. 273, 927–948.
[17] Guarbé, A., Bravo, J., Calvo, J., Lorana, F., Vives, J. and Fita, I. (1996) Protein Sci. 5, 167–169.
[18] Zhou, G.W., Guo, J., Huang, W., Fletterick, R.J. and Scanlan, T.S. (1994) Science 265, 1059–1064.
[19] Wu, T.T., Johnson, G. and Kabat, E.A. (1993) Proteins Struct. Funct. Genet. 16, 1–7.
[20] Kabat, E.A., Wu, T.T., Perry, H.M., Gottesman, K.S. and Foeller, C. (1991) in: Sequences of Proteins of Immunological Interest, 5th edn., US Department of Health and Human Services, NIH, Bethesda, MD.
[21] Nicholls, A., Sharp, K.A. and Honig, B. (1991) Proteins Struct. Funct. Genet. 11, 281–296.
[22] Humprey, W., Dolke, A. and Schulten, K. (1996) J. Mol. Graph. 14, 33–38.
[23] Merrit, E.A. and Bacon, D.J. (1997) in: Methods in Enzymology: Macromolecular Crystallography Part B (Carter, C.W. and Sweet, R.M., Eds.), Vol. 277, pp. 505–524, Academic Press.