JOURNAL OF VIROLOGY, May 1983, p. 626–628
0022-538X/83/050626-03$02.00/0
Copyright © 1983, American Society for Microbiology
Vol. 46, No. 2

Structural Relationships Between Minor and Major Proteins of Hepatitis B Surface Antigen

WERNER STIBBE† AND WOLFRAM H. GERLICH*

Department of Medical Microbiology, University of Göttingen, Göttingen, Federal Republic of Germany

Received 2 November 1982/Accepted 31 January 1983

The minor glycoproteins from hepatitis B surface antigen, GP33 and GP36, contain at their carboxy-terminal part the sequence of the major protein P24. They have 55 additional amino acids at the amino-terminal part which are coded by the pre-S region of the viral DNA.

Hepatitis B surface antigen (HBsAg) consists of a major protein, P24, and its glycosylated form, GP27 (9). The gene for P24 starts at the third conserved AUG codon of a larger open reading frame in the DNA of hepatitis B virus (2, 8, 14). HBsAg 20-nm particles from viremic donors (13) also contain two minor glycoproteins, GP33 and GP36 (11, 12), which have not yet been well characterized. Here we present evidence that these proteins are coded by the same DNA region. Their sequence probably starts at the second conserved AUG of the HBsAg coding region and ends at the same stop codon as that of P24.

The close relationship between the major and the minor HBsAg glycoproteins was first shown by amino acid analysis. Acid hydrolysates of GP27, GP33, and GP36 had identical amino acid compositions within the technical error (data not shown). Further comparison was done by partial proteolysis mapping. ^{125}I-labeled P24 produced trypsin fragments of 21 kilodaltons (kd) and 13 kd; GP27 generated instead fragments of 24 and 16 kd (Fig. 1). The transitory formation of fragments which were 3 kd smaller than P24 or GP27 was due to cleavage at arginine 24 of their sequence (see Fig. 3). The formation of a stable 13-kd band from P24 suggested cleavage at arginine 122. GP27 has a complex glycan at asparagine 146 (7), and so the size of its carboxy-terminal fragment was 16 instead of 13 kd. The central 13-kd fragment from GP27 was undetectable in autoradiography. Surface iodination of 20-nm particles apparently labeled only tyrosine 134.

Tryptic digests of ^{125}I-GP33 had the 21-kd and 13-kd fragments in common with P24. Digests of ^{125}I-GP36 contained 24-kd and 16-kd fragments instead, and these fragments were in common with GP27 (Fig. 2A). Small amounts of the 13-kd fragment in digests from GP36 and of the 16-kd fragment in digests of GP33 were due to slight cross-contaminations. Removal of neuraminic acid from GP36 with neuraminidase reduced the size of the 24- and 16-kd fragments slightly. All other fragments from GP36 or GP33 were not changed by neuraminidase (data not shown). These results suggested that GP33 and GP36 were identical in their carboxy-terminal portions

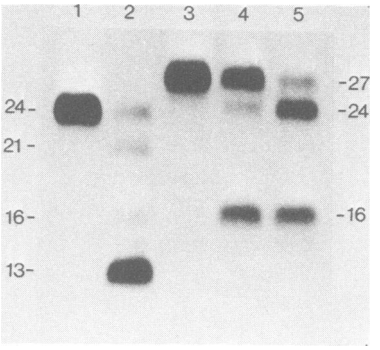

FIG. 1. Tryptic fragments of the major HBsAg proteins. HBsAg 20-nm particles were isolated from the blood plasma of one chronic HBsAg carrier by gel chromatography and isopycnic centrifugation (3, 4). Surface labeling with sodium [^{125}I]iodide and a solid-phase oxidant and isolation of the denatured ^{125}I-labeled proteins by polyacrylamide gel electrophoresis are described elsewhere (13). Reaction mixtures contained 10^5 to 10^6 cpm of ^{125}I-protein, 3 μg of bovine serum albumin, and trypsin as indicated in 15 μl. The reaction was stopped with 2× sample buffer (6), and the samples were boiled for 4 min. The mixtures were separated in 15% acrylamide gels (6), and the ^{125}I-labeled fragments were detected by autoradiography. Lane 1, P24; lane 2, P24 digested with 4 μg of trypsin (DPCC treated; Sigma) for 3 h at 37°C; lane 3, GP27; lanes 4 and 5, GP27 digested with 4 μg of trypsin at 20 or 37°C, respectively. The numbers given the apparent size in kd.

† Present address: Department of Clinical Chemistry, University of Göttingen, Göttingen, Federal Republic of Germany.

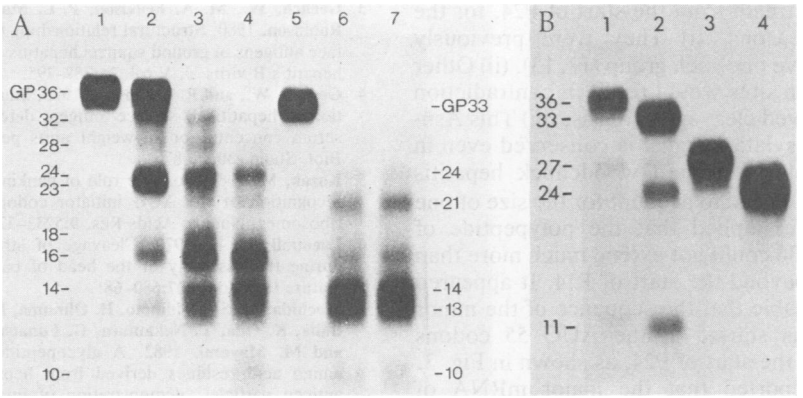

FIG. 2. Proteolytic fragments of the minor ^{125}I-labeled HBsAg proteins. (A). Digestion with trypsin for 2 h at 37°C in 0.2% dithiothreitol–0.1 M Tris-hydrochloride (pH 7.2). Lane 1, GP36; lanes 2 through 4, GP36 with 1, 5, or 10 μg of trypsin, respectively. Lane 5, GP33; lanes 6 and 7, GP33 with 1 or 5 μg of trypsin, respectively. (B) Digestion with 5 μg of *S. aureus* V8 protease (Miles) in 0.1% sodium dodecyl sulfate for 24 h at 37°C. Lane 1, GP36; lane 2, GP33; lane 3, GP27; lane 4, P24.

to P24 or GP27 and that they differed only by one complex glycan group which was bound to the carboxy-terminal 16-kd fragment at GP36.

The larger size of the minor glycoproteins suggested an amino-terminal extension beyond the start of P24 in a way predicted by the DNA sequence (Fig. 3). In agreement with the DNA sequence, the 32- and 28-kd fragments from GP36 (Fig. 2A, lane 3) suggested the presence of basic amino acids 35 and 70 positions upstream of arginine 24 in P24. The corresponding fragments of GP33 were, as expected, 4 kd smaller (only the 24-kd fragment is visible in Fig. 2A). The proposed arrangement of the amino-terminal region in GP33 and GP36 was further confirmed by digestion with *Staphylococcus aureus* V8 protease. The bands of P24 and GP27 were not altered by the protease. However, GP33 was slowly degraded to 24- and 11-kd fragments. GP36 also generated the 11-kd fragment and a 27-kd fragment (Fig. 2B). This suggested cleavage at the glutamic acid 2 in P24 (see Fig. 3).

Both minor proteins were assumed to have one mannose-rich glycan at the asparagine 51

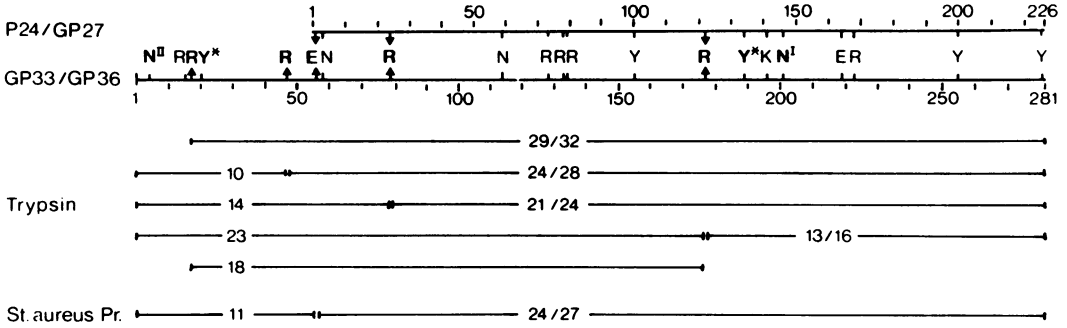

FIG. 3. Proposed structural relationship between the minor HBsAg proteins GP33 and GP36 and the major proteins P24 and GP27. The tentative numbering for GP33 and GP36 starts with the next AUG codon in the hepatitis B virus DNA upstream from the amino-terminal methionine of P24. The positions of the basic amino acids (R, K), glutamic acid (E), tyrosine (Y), and glycosylation signal sequences (N) are given in the one-letter code and were read from the DNA sequence. This sequence, from Pasek et al. (8), was selected for the model because the HBsAg used for analysis came from a virus carrier whose viral DNA had a closely related restriction pattern (U. Böttcher, personal communication). Sites which were found to be available for surface labeling, proteolysis, or glycosylation are shown in bold face. Proteolytic cleavage sites are marked by an arrow. Asterisks identify sites of surface labeling. The six lines at the bottom show the positions of all observed radiolabeled proteolytic fragments from GP33 and GP36 in the proposed sequence. All carboxy-terminal fragments in GP36 and GP27 are considered to be 3 to 4 kd larger than those of GP33 and P24 because of their additional type I glycan (N^I). The amino-terminal fragments of GP33 and GP36 are 3 to 5 kd larger than the expected size because of the type II glycan (N^{II}) in position 4.

position upstream from the start of P24, for the following reasons. (i) They were previously shown to have one such group (12, 13). (ii) Other glycosylation sites would result in contradiction to the observed cleavage patterns. (iii) This Asn-X-Ser glycosylation signal is conserved even in the DNA of the related woodchuck hepatitis virus (1). The need to account for the size of one carbohydrate implied that the polypeptide of GP33 or GP36 could not extend much more than 55 codons beyond the start of P24. It appeared most reasonable that the sequence of the minor glycoproteins started at the AUG 55 codons upstream of the start of P24, as shown in Fig. 3.

It was reported that the major mRNA of HBsAg-producing cells covers most of the open coding sequence of HBsAg (10). Thus, both minor and major HBsAg proteins could be translated from the same mRNA. The tentative initiation signal for translation of GP33 and GP36 is probably weak because it lacks the flanking bases C or AXXAUGG (5), and so initiation would occur more frequently at the start of P24, which has the appropriate structure. Thus, the different amounts of major and minor HBsAg proteins may be due to regulation at the translational level.

H. Kratzin kindly performed amino acid analysis.

This work was supported by Deutsche Forschungsgemeinschaft grant Ge 345/6.

LITERATURE CITED

1. **Galibert, F., T. N. Chen, and E. Mandart.** 1982. Nucleotide sequence of a cloned woodchuck hepatitis virus genome: comparison with the hepatitis B virus sequence. J. Virol. **41**:51–65.
2. **Galibert, F., E. Mandart, F. Fitoussi, P. Tiollais, and P. Charnay.** 1979. Nucleotide sequence of the hepatitis B virus genome (subtype ayw) cloned in E. coli. Nature (London) **281**:646–650.
3. **Gerlich, W., M. A. Feitelson, P. L. Marion, and W. S. Robinson.** 1980. Structural relationships between the surface antigens of ground squirrel hepatitis virus and human hepatitis B virus. J. Virol. **36**:787–795.
4. **Gerlich, W., and R. Thomssen.** 1975. Standardized detection of hepatitis B surface antigen: determination of its serum concentration in weight units per volume. Dev. Biol. Stand. **30**:78–87.
5. **Kozak, M.** 1981. Possible role of flanking nucleotides in recognition of the AUG initiator codon by eukaryotic ribosomes. Nucleic Acids Res. **9**:5233–5252.
6. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227**:680–685.
7. **Machida, A., S. Kishimoto, H. Ohnuma, H. Miyamoto, K. Baba, K. Oda, T. Nakamura, G. Funatsu, Y. Mijakawa, and M. Mayumi.** 1982. A glycopeptide containing 15 amino acid residues derived from hepatitis B surface antigen particles: demonstration of immunogenicity to raise anti-HBs in mice. Mol. Immunol. **19**:1087–1093.
8. **Pasek, M., T. Goto, W. Gilbert, B. Zink, H. Schaller, P. McKay, G. Leadbetter, and K. Murray.** 1979. Hepatitis B virus genes and their expression in E. coli. Nature (London) **282**:575–579.
9. **Peterson, D. L.** 1981. Isolation and characterization of the major protein and glycoprotein of hepatitis B surface antigen. J. Biol. Chem. **256**:6975–6983.
10. **Pourcel, C., A. Louise, M. Gervais, N. Chenciner, M. F. Dubois, and P. Tiollais.** 1982. Transcription of the hepatitis B surface antigen gene in mouse cells transformed with cloned viral DNA. J. Virol. **42**:100–105.
11. **Sanchez, Y., I. Ionescu-Matiu, G. R. Dreesman, F. B. Hollinger, and J. L. Melnick.** 1981. Evidence for the presence of repeating antigenic determinants in the major and minor polypeptides derived from hepatitis B surface antigen. Virology **114**:71–80.
12. **Skelly, J., C. R. Howard, and A. J. Zuckerman.** 1978. The labelling of galactose residues in hepatitis B surface antigen glycoprotein. J. Gen. Virol. **41**:447–457.
13. **Stibbe, W., and W. H. Gerlich.** 1982. Variable protein composition of hepatitis B surface antigen from different donors. Virology **123**:436–442.
14. **Valenzuela, P., M. Quiroga, J. Zaldivar, P. Gray, and W. J. Rutter.** 1981. The nucleotide sequence of the hepatitis B viral genome and the identification of the major viral genes, p. 57–70. *In* B. Fields, R. Jaenisch, and C. F. Fox (ed.), Animal virus genetics. Academic Press, Inc., New York.