RAPID COMMUNICATION Jpn. J. Cancer Res. 81, 1191–1194, December 1990

X Protein of Hepatitis B Virus Resembles a Serine Protease Inhibitor

Shinako Takada and Katsuro Koike[1]

Department of Gene Research, Cancer Institute, Kami-Ikebukuro, Toshima-ku, Tokyo 170

The X protein of hepatitis B virus (HBV) has been shown to be a *trans*-activator for viral and cellular genes. Amino acid sequences in X protein were found to be highly homologous to functionally essential sequences in the "Kunitz domain," characteristic of Kunitz-type serine protease inhibitors. Mutations at these sequences completely abolished *trans*-activation. Consequently, HBV X protein resembles a serine protease inhibitor or its analogue, and may bring about *trans*-activation by activating certain transcriptional factors through proteolytic cleavage alteration.

Key words: Hepatitis B virus — X gene — *trans*-Activation — Serine protease inhibitor — Kunitz domain

The *trans*-activation of cellular oncogene(s) by hepatitis B virus (HBV) X gene product from integrated HBV DNA during chronic infection may be the predominant mechanism for HBV-directed hepatocarcinogenesis.[1] Although X protein of HBV has been shown to be a *trans*-activator for various viral and cellular genes,[2–5] the mechanism of *trans*-activation remains to be clarified. As the amino acid sequence (L-G-G-C-R-H-K) in the carboxy-terminal region of X protein was previously shown to be an essential region for *trans*-activation,[1] in this study, a search was made for homology and this amino acid sequence in X protein was found to be highly homologous to one essential sequence in the "Kunitz domain," characteristic of Kunitz-type serine protease inhibitors.[6–8] An amino acid sequence highly homologous to another functionally essential sequence (G-P-C) next to the reaction site in the Kunitz domain was also noted in X protein. Mutation at this sequence completely abolished *trans*-activation. Consequently, HBV X protein resembles a serine protease inhibitor and may bring about *trans*-activation by activating certain transcriptional factors through proteolytic cleavage alteration. This hypothesis provides a possible basis for understanding the mechanisms of hepatocarcinogenesis and the deregulation of cellular gene expression.

To determine the manner in which X protein brings about transcriptional *trans*-activation of various genes, the carboxy-terminal 30-amino-acid sequence was subjected to computer analysis. The No. 132–140 amino acid sequence in the carboxy-terminal region of HBV X protein, which is conserved among hepadna viruses, was found to be highly homologous to the consensus amino acid sequence (F-V/I-X-G-G-C-R/K) within the Kunitz domain essential for Kunitz-type serine protease inhibitor function[6–8] and incidentally to p18 protein in the *abl* gene of Abelson murine leukemia virus (A-MuLV), whose function is unknown (Fig. 1). Serine protease inhibitors are intracellular polypeptides present in many tissues and inhibit one or some serine proteases such as trypsin, kallikrein, chymotrypsin, and plasmin by virtue of their binding to the active center of the protease. Protease inhibitors regulate the biological activity of protease by becoming associated with or dissociated from them. Thus, the balance between protease and protease inhibitor affects various inter- and intracellular phenomena.

Serine protease inhibitors also possess the consensus amino acid sequence (G-P-C) within their Kunitz domain, in which the reaction site (R or K) is next to the (G-P-C) sequence.[6–8] This reaction site is situated close to the former consensus sequence (F-V/I-X-G-G-C-R/K) owing to the presence of an intramolecular disulfide bond.[6–8] The X proteins of hepadna viruses and p18 of A-MuLV also possess this (G-P-C) sequence, though the distance to the (F-V/I-X-G-G-C-R/K) sequence is different from that of known protease inhibitors with respect to the primary structure (Fig. 1). The C residue 6 or 7 amino acids upstream from the (G-P-C) sequence is conserved in X protein. The R residue corresponding to the cleavage site of bovine basic protease inhibitor precursor (BBPI) on releasing the inhibitor molecule is also conserved (Fig. 1). Thus, X protein possesses all the conserved amino acid residues within the Kunitz domain of Kunitz-type serine protease inhibitors except for the Y residue (No. 49 of BBPI) and the three C residues (Nos. 44, 65, and 69 of BBPI) (Fig. 1 and ref. 6–8).

If, in fact, X protein actually resembles serine protease inhibitors or related proteins, the mutation of the possible Kunitz domain would have a deleterious effect on its function. Deletion of the last 22 amino acids of HBV X protein was found in the previous study to result in the

[1] To whom correspondence should be addressed.

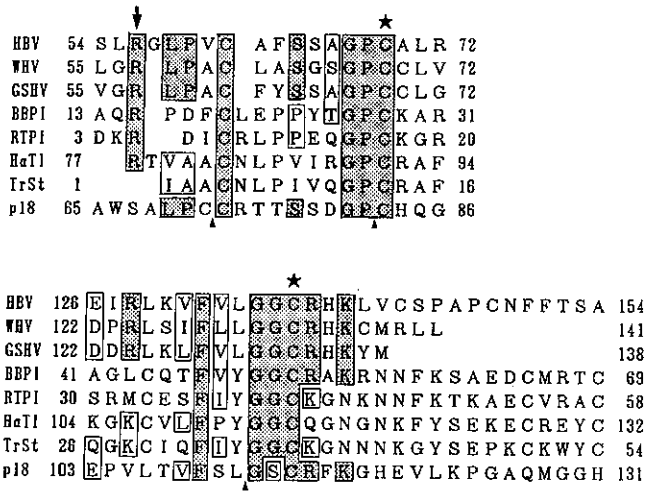

Fig. 1. Comparison of amino acid sequences of X proteins from hepadna viruses, the Kunitz domains of serine protease inhibitors and p18[abl]. Common amino acids between any one of the serine protease inhibitors and hepadna virus X proteins are boxed with shadows while conserved substitutions are boxed without shadows. Common amino acids of p18[abl] compared with those of serine protease inhibitor and X proteins are also boxed. Amino acid numbers of the first and last residues in each sequence are indicated on both sides. HBV, hepatitis B virus[22]; WHV, woodchuck hepatitis virus[23]; GSHV, ground squirrel hepatitis virus[23]; BBPI, bovine basic protease inhibitor precursor[24]; RTPI, red sea turtle basic protease inhibitor[25]; HαTI, human inter-α-trypsin inhibitor[26]; TrSt, trypstatin from rat mast cells.[26] ★ denotes two C residues possibly linked together by a disulfide bond in X protein. The arrow indicates No. 15 R residue of BBPI, the cleavage site for releasing an inhibitor molecule. Disulfide bonds in BBPI are as follows; C(19)-C(69), C(28)-C(52), C(44)-C(65). The carboxy-terminal 30-amino-acid sequence (Nos. 125–154) of HBV X protein was subjected to local homology search using the NBRF protein database with program SEQFP in IDEAS (Integrated Database and Extended Analysis System for nucleic acids and proteins).

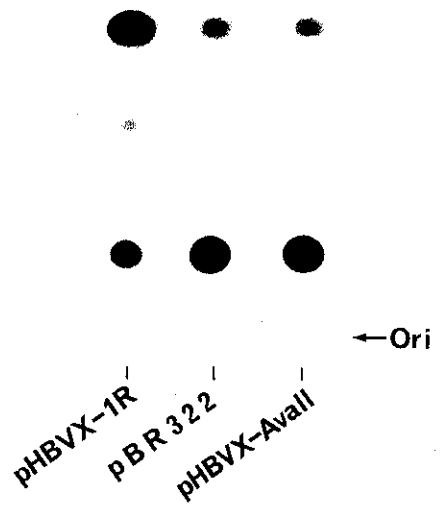

Fig. 2. CAT activity assay of X protein and its mutant. HepG2 cells were co-transfected with 2 μg of pSV2CAT and 18 μg of each indicated plasmid. pBR322 was used as the negative control. The plasmid pSV2CAT contains the chloramphenicol acetyltransferase (CAT) gene driven by the SV40 enhancer/early promoter. Transfection and the CAT assay were carried out as described.[1] The nucleotide sequence of the mutant plasmid, pHBVX-AvaII, was confirmed by the chemical modification method.[27]

complete abolition of *trans*-activation and the conserved amino acid sequence (L-G-G-C-R-H-K) to be essential for the *trans*-activation function of X protein.[1] Another mutant X gene was subsequently made and compared with a wild-type X gene. A wild-type X expression plasmid pHBVX-1R was constructed by inserting a 0.87 kb *Stu*I(n.t. 988)-*Bgl*II (n.t. 1860) HBV DNA fragment into the pBR322 *Bam*HI site through the *Bgl*II linker. The 0.87 kb HBV DNA fragment contains promoter/enhancer sequences, X ORF and a poly(A) addition signal. A mutant plasmid pHBVX-AvaII was constructed by S1 nuclease digestion of the *Ava*II site (n.t. 1446) followed by blunt-end ligation, and accordingly Nos. 67, 68 and 69 amino acids of X protein G-P-C were converted to R-C. Their *trans*-activation ability was examined using pSV2CAT as a reporter plasmid. Fig. 2 shows the mutant X to have completely lost its capacity to bring about *trans*-activation. Consistent with the above indication that the X protein may resemble a serine protease inhibitor in structure, the (G-P-C) sequence was thus essential for X protein to exert its *trans*-activation function. However, as shown in Fig. 1, owing to the absence of the K or R residue next to the (G-P-C) sequence possessing affinity for the serine residue in the reactive center of known proteases,[6] X protein may be incapable of inactivating the target protease. Even if this should be the case, the X protein might nevertheless compete with the substrate or its natural inhibitor by virtue of tertiary structural similarity. The X protein may thus deregulate the quantity and quality of transcription factors in cells by disordering protease action. Such metabolic deregulation of transcription factors would result in the abnormal transcriptional activation of some genes, leading to hepatocarcinogenesis. The exact role of the (F-V/I-X-G-G-C-R/K) sequence is still not known, but the sequence might contribute to the formation of the tertiary structure about the reaction site, this possibly

being essential to recognition between protease and protease inhibitor. The duck hepatitis B virus (DHBV) genome lacks these two sequences in the X+C region[9] and has no *trans*-activation activity.[10] This supports the view that *trans*-activation by the X gene occurs by way of a protease inhibitor-like structure.

Various serine protease inhibitors are present in cells, making up a superfamily, and the X gene may possibly derive from one of them. That is, the X gene might be of cellular origin and might have been transduced into the hepadna virus genome. This view is consistent with the previous indication by Miller and Robinson that X gene is of cellular origin based on the codon usage preference and the location of this gene in the viral genome.[11]

It is of interest that p18 protein, encoded not only in the v-*abl* region of A-MuLV[12] but also in the mouse c-*abl* locus,[13] is homologous to X protein. p18 is thought to have been transduced to the M-MuLV genome along with the so-called *abl* gene, whose product possesses tyrosine kinase activity. p18 is located 330 bp downstream from the end of $p120^{gag\text{-}abl}$ ORF in the A-MuLV genome.

The estradiol receptor was recently shown to possess proteolytic activity responsible for its own transformation which confers affinity for nuclei and DNA on the steroid-receptor complex, and serine protease inhibitors and substrates impair this transformation of the estradiol receptor.[14] Some protease inhibitors have been shown to be identical to growth stimulatory factor[15,16] or cell division cycle gene.[17] Although the mechanism for growth stimulation is unclear, these properties are similar to those of HBV X protein, which exhibits some growth-stimulatory effect.[18,19]

Since *trans*-activation function on the part of X protein depends on the particular cell type used, this protein may operate through interaction with specific cellular factor(s) in certain cells.[4,20] For detailed clarification of the mechanism, the target protein for X protein must be identified.

Chisari *et al.* recently found the accumulation of toxic quantities of HBsAg within hepatocytes to result in severe, prolonged hepatocellular injury which in turn leads to neoplasia, and that regardless of etiology, chronic cell injury and associated inflammatory and regenerative response constitute a preneoplastic condition which inevitably proceeds toward malignancy.[21] Imbalance of protease and protease inhibitor action may also cause the accumulation of abnormal protein. Thus, X protein may contribute to hepatocarcinogenesis in multiple modes as a serine protease inhibitor-like protein, that is, *trans*-activation of certain genes, growth stimulation, and the promotion of liver cell injury. Thus, hepadna virus X proteins may comprise a new oncogene family. The concept of protease-protease inhibitor balance provides a new insight into the mechanisms of hepatocarcinogenesis as well as deregulation of cellular gene expression.

This work was supported in part by a Grant-in-Aid from the Ministry of Health and Welfare, Japan for the Comprehensive 10-Year Strategy for Cancer Control and by a Grant-in-Aid for Cancer Research from the Ministry of Education, Science and Culture, Japan to K.K. Computer analysis was carried out on-line using the VAX system of the Institute of Medical Science, University of Tokyo.

(Received July 14, 1990/Accepted September 17, 1990)

REFERENCES

1) Takada, S. and Koike, K. Trans-activation function of a 3′ truncated X gene-cell fusion product from integrated hepatitis B virus DNA in chronic hepatitis tissues. *Proc. Natl. Acad. Sci. USA*, **87**, 5628–5632 (1990).
2) Spandau, D. F. and Lee, C. *trans*-Activation of viral enhancers by the hepatitis B virus X protein. *J. Virol.*, **62**, 427–434 (1988).
3) Seto, E., Yen, T. S. B., Peterlin, B. M. and Ou, J. H. Trans-activation of the human immunodeficiency virus long terminal repeat by the hepatitis B virus X protein. *Proc. Natl. Acad. Sci. USA*, **85**, 8286–8290 (1988).
4) Koike, K., Shirakata, Y., Yaginuma, K., Arii, M., Takada, S., Nakamura, I., Hayashi, Y., Kawada, M. and Kobayashi, M. Oncogenic potential of hepatitis B virus. *Mol. Biol. Med.*, **6**, 151–160 (1989).
5) Aufiero, B. and Schneider, R. J. The hepatitis B virus X-gene product *trans*-activates both RNA polymerase II and III promoters. *EMBO J.*, **9**, 497–504 (1990).
6) Laskowski, M., Jr. and Kato, I. Protein inhibitors of proteinases. *A. Rev. Biochem.*, **49**, 593–626 (1980).
7) Wunderer, G., Machleidt, W., and Frotz, H. The broad-specificity proteinase inhibitor 5 II from the sea anemone *Anemonia sulcata*. *Methods Enzymol.*, **80**, 816–820 (1981).
8) Strydom, D. J. and Joubert, F. J. Amino-acid sequence of a weak trypsin inhibitor-B from *Dendroaspis-polylepis* (black mamba). *Hoppe-Seyler's Z. Physiol. Chem.*, **362**, 1377–1384 (1981).
9) Mandart, E., Kay, A. and Galibert, F. Nucleotide sequence of a cloned duck hepatitis B virus genome: comparison with woodchuck and human hepatitis B virus sequences. *J. Virol.*, **49**, 782–792 (1984).
10) Colgrove, R., Simon, G. and Ganem, D. Transcriptional activation of homologous and heterologous genes by the hepatitis B virus X gene product in cells permissive for viral replication. *J. Virol.*, **63**, 4019–4026 (1989).

11) Miller, R. H. and Robinson, W. S. Common evolutionary origin of hepatitis B virus and retroviruses. *Proc. Natl. Acad. Sci. USA*, **83**, 2531–2535 (1986).
12) Reddy, E. P., Smith, M. J. and Srinivasan, A. Nucleotide sequence of Abelson murine leukemia virus genome: structural similarity of its transforming gene product to other *onc* gene products with tyrosine-specific kinase activity. *Proc. Natl. Acad. Sci. USA*, **80**, 3623–3627 (1983).
13) Wang, J. Y. J., Ledley, F., Goff, S., Lee, R., Groner, Y. and Baltimore, D. The mouse c-*abl* locus: molecular cloning and characterization. *Cell*, **36**, 349–356 (1984).
14) Puca, G. A., Abbondanza, C., Nigro, V., Armetta, I., Medici, N. and Morinari, A. M. Estradiol receptor has proteolytic activity that is responsible for its own transformation. *Proc. Natl. Acad. Sci. USA*, **83**, 5367–5371 (1986).
15) Mckeehan, W. L., Sakagami, Y., Hoshi, H. and Mckeehan, K. A. Two apparent human endothelial cell growth factors from human hepatoma cells are tumor-associated proteinase inhibitors. *J. Biol. Chem.*, **261**, 5378–5383 (1986).
16) Gasson, J. C., Golde, D. W., Kaufman, S. E., Westbrook, C. A., Hewick, R. M., Kaufman, R. J., Wong, G. G., Temple, P A., Leary, A. C., Brown, E. L., Orr, E. C. and Clark, S. C. Molecular characterization and expression of the gene encoding human erythroid-potentiating activity. *Nature*, **315**, 768–771 (1985).
17) Edwards, D. R., Waterhouse, P., Holman, M. L. and Denhardt, D. T. A growth-responsive gene (16C8) in normal mouse fibroblasts homologous to a human collagenase inhibitor with erythroid-potentiating activity: evidence for inducible and constitutive transcripts. *Nucleic Acids Res.*, **14**, 8863–8878 (1986).
18) Shirakata, Y., Kawada, M., Fujiki, Y., Sano, H., Oda, M., Yaginuma, K., Kobayashi, M. and Koike, K. The X gene of hepatitis B virus induced growth stimulation and tumorigenic transformation of mouse NIH3T3 cells. *Jpn. J. Cancer Res.*, **80**, 617–621 (1989).
19) Hohne, M., Schaefer, S., Seifer, M., Feitelson, M. A., Paul, D. and Gerlich, W. H. Malignant transformation of immortalized transgenic hepatocytes after transfection with hepatitis B virus DNA. *EMBO J.*, **9**, 1137–1145 (1990).
20) Seto, E., Zhou, D. X., Peterlin, B. M. and Yen, T. S. B. *trans*-Activation by the hepatitis B virus X protein shows cell-type specificity. *Virology*, **173**, 764–766 (1989).
21) Chisari, F. V., Klopchin, K., Moriyama, T., Pasquinelli, C., Dunsford, H. A., Sell, S., Pinkert, C. A., Brinster, R. L. and Palmiter, R. D. Molecular pathogenesis of hepatocellular carcinoma in hepatitis B virus transgenic mice. *Cell*, **59**, 1145–1156 (1989).
22) Kobayashi, M. and Koike, K. Complete nucleotide sequence of hepatitis B virus DNA of subtype adr and its conserved gene organization. *Gene*, **30**, 227–232 (1984).
23) Kodama, K., Ogasawara, N., Yoshikawa, H. and Murakami, S. Nucleotide sequence of a cloned woodchuck hepatitis virus genome. *J. Virol.*, **56**, 978–986 (1985).
24) Anderson, S. and Kingston, I. B. Isolation of a genomic clone for bovine pancreatic trypsin inhibitor by using a unique-sequence synthetic DNA probe. *Proc. Natl. Acad. Sci. USA*, **80**, 6838–6842 (1983).
25) Kato, I. and Tominaga, N. Trypsin-subtilisin inhibitor from red sea turtle eggwhite consists of two tandem domains -- one Kunitz -- one of a new family. *Fed. Proc.*, **38**, 832 (1979).
26) Kido, H., Yokogoshi, Y. and Katunuma, N. Kunitz-type protease inhibitor found in rat mast cells. *J. Biol. Chem.*, **263**, 18104–18107 (1988).
27) Maxam, A. M. and Gilbert, W. Sequencing end labeled DNA with base specific chemical cleavage. *Methods Enzymol.*, **65**, 499–560 (1980).