Identification and Expression of Glycine Decarboxylase (p120) as a Duck Hepatitis B Virus Pre-S Envelope-binding Protein*

Jisu Li, Shuping Tong, and Jack R. Wands†

From the Molecular Hepatology Laboratory, Massachusetts General Hospital Cancer Center and Harvard Medical School, Charlestown, Massachusetts 02129

A 120-kilodalton protein (p120) was identified in the duck liver that binds to several truncated versions of duck hepatitis B virus (DHBV) pre-S envelope protein, suggesting p120 may serve as a DHBV co-receptor. The amino acid sequences of tryptic peptides from purified p120 were found to be the duck p protein of the glycine decarboxylase complex (DGD). DGD cDNA cloning revealed extensive protein conservation with the chicken homologue except for several insertions in the N-terminal leader sequence. The DGD cDNA contained no in-frame AUG codon at the predicted initiation site of the open reading frame, and site-directed mutagenesis experiments established an AUU codon as the translational initiator. The DGD protein expressed in rabbit reticulocyte lysates bound truncated DHBV pre-S protein identical to that of p120 derived from duck liver confirming DGD as p120. Moreover, transfection studies in liver- and kidney-derived cells revealed both cell surface and cytoplasmic expression of the protein. Cloning of the glycine decarboxylase cDNA will permit a direct test of whether it functions as a cell surface co-receptor or as a co-factor in the DHBV replication cycles.

The human hepatitis B virus and related animal viruses form hepatotropic DNA viruses or hepadnaviruses (1); because the early events of hepatocyte infection are unclear, studies were initiated via the duck hepatitis B virus (DHBV) model in a multifold approach to identify candidate cell surface receptor proteins. Interactive proteins for the pre-S domain of DHBV large envelope protein (the assumed ligand to viral receptor) were identified and cDNAs cloned to verify their potential role as DHBV receptor/co-receptor in nonsusceptible cell lines. Two pre-S interacting proteins, p170 (2) and p120 (3), were classified using DHBV pre-S domain fused to glutathione S-transferase (GST); p170 was analogous to the gp180 DHBV-envelope protein characterized by Kuroki et al. (4) and, based on its similarity to carboxypeptidases (5), was renamed duck carboxypeptidase D. Transfection of duck carboxypeptidase D cDNA into liver- and kidney-derived cell lines, conferred efficient DHBV binding and entry, indicating p170 (duck carboxypeptidase D) pre-S-binding protein served as a primary DHBV receptor.

Optimal p170 binding requires an entire pre-S domain (residues 1–161) (2). Although p120 interacts with this pre-S peptide, it also binds with high affinity to several cleaved pre-S polypeptides (92–161, 98–161, 1–102) (3); of these, 1–102 may be generated in vivo by cleavage of the large envelope protein present on virion particles by a di-basic endopeptidase. The sequence surrounding pre-S residue 102 (Arg-Glu-Ala-Phe) is underlined) fulfills the requirement for cleavage by furin, which processes many viral envelope protein precursors including human immunodeficiency virus (7). Thus, depending on the subcellular compartment of endopeptidase cleavage, p120 may serve as a cell surface co-receptor or an intracellular binding partner facilitating the disassembly of viral particles.

Further support of the involvement of p120 in the DHBV life cycle is proposed by 1) exclusive expression in the liver, kidney, and pancreas of known DHBV tissues susceptible to infection, contrasting sharply with widespread duck carboxypeptidase D tissue distribution (2–4); 2) the p120 binding site, mapped to pre-S residues 98–102, where a mouse monoclonal antibody blocks DHBV infection in primary duck hepatocytes (PDH) (8, 9); 3) a short pre-S peptide (residues 80–102) with affinity for p120 but not duck carboxypeptidase D substantially inhibited productive DHBV infection of PDH. Similarly, double-point mutations within the p120 binding site severely hampered viral infection in PDH (3). In the present investigation, the cDNA of p120 was cloned, and the protein was expressed to characterize its binding properties to pre-S peptides. More important, available protein on the surface of transfected cells was detected.

EXPERIMENTAL PROCEDURES

Purification and Microsequencing of the p120 Protein—Frozen duck liver (40 gm) was homogenized in 300 ml of lysis buffer (2). After sequential pre-clearing with empty glutathione–Sepharose beads and GST protein conjugated on beads, the liver lysates were incubated at 4 °C overnight with pre-S peptide 80–102 anchored on Sepharose beads via the GST tag. Bound proteins were separated by a 0.1% SDS-6% polyacrylamide gel (PAGE) and transferred to polyvinylidene difluoride membrane (Bio-Rad). After staining with 0.1% Ponceau red, the 120-kD band was excised, and 28 μg of protein was obtained for sequence analysis at the Harvard Microchemistry Facility. Briefly, p120 was digested with trypsin, and peptide fragments were separated by high pressure liquid chromatography (HPLC). Selected peptides were sequenced by the Edman degradation method.

For N-terminal sequencing, p120 was immunoprecipitated from 3 ml of lystate (1 g of duck liver) using 80 μl of rabbit antiserum raised against recombinant DGD protein expressed in baculovirus (α-Bac-

* This work was supported in part by National Institutes of Health Grants CA-35711 and AA-02666. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

† Current address and to whom correspondence should be addressed: Liver Research Center, 55 Claverick St., Fourth Floor, Providence, RI 02903. Tel.: 01-444-2795; Fax: 401-444-2939; E-mail: Jack Wands MD @Brown.edu.

‡ The abbreviations used are: DHBV, duck hepatitis B virus; GST, glutathione S-transferase; PDH, primary duck hepatocyte; DGD, duck glycine decarboxylase; CGD, chicken glycine decarboxylase; PAGE, polyacrylamide gel electrophoresis; HPLC, high pressure liquid chromatography; kb, kilobase(s); PCR, polymerase chain reaction; nt, nucleotide(s).
After electrophoresis, the p120 band (6 μg) was excised, sequenced through SDS-PAGE, and transferred to the polyvinylidene difluoride membrane.

**Construction of Duck Liver cDNA Libraries and Isolation of Duck Glycine Decarboxylase Clones**—Oligonucleotides used are listed in Table I. RNA extracted from frozen Pekin duckling liver with guanidinium thiocyanate was purified through oligo(dT) columns. Two types of cDNA expression libraries were constructed via Stratagene protocol: an oligo(dT)-primed library inserted into the λ ZAP II vector, and a randomly primed library prepared in the ZAP expression vector. Both were amplified once and contained approximately 6 × 10^6 independent recombinants. The DGD cDNA clones were initially isolated from the oligo(dT) library with a 2.3-kb cDNA fragment derived from chicken glycine decarboxylase (CGD) as a probe (10). The CGD fragment was PCR-amplified from first-strand chicken liver cDNA using primers CGD3 and CGD4 described in Fig. 1A. To enrich for clones containing extended 5′ ends, positive plaques were re-screened with a 0.6-kb PCR product derived from chicken liver with a 2.3-kb cDNA fragment derived from chicken glycine decarboxylase (CGD).3 After electrophoresis, the p120 band (6 μg) was excised, sequenced through SDS-PAGE, and transferred to the polyvinylidene difluoride membrane.

**Table I**

| Oligonucleotides used in this study |
|-----------------------------------|
| **CGD3** | S |
| **CGD4** | AS |
| **CGD8** | S |
| **CGD2** | AS |
| **DGD24a** | S |
| **DGD27** | S |
| **DGD30** | S |
| **AUC328AUG** | S |
| **AUC346AUG** | S |
| **AUA346AUG** | S |
| **AGC931AUG** | S |
| **CUG397CGC** | S |
| **Phe1** | S |
| **Phe2** | S |
| **Phe3** | S |
| **Phe4** | S |
| **Phe5** | S |
| **Asn2** | S |
| **tr1** | S |
| **tr2** | AS |
| **tr7** | AS |
| **tr9** | S |

**DGD**. After electrophoresis, the p120 band (6 μg) was excised, sequenced through SDS-PAGE, and transferred to the polyvinylidene difluoride membrane.

**Construction of Duck Liver cDNA Libraries and Isolation of Duck Glycine Decarboxylase Clones**—Oligonucleotides used are listed in Table I. RNA extracted from frozen Pekin duckling liver with guanidinium thiocyanate was purified through oligo(dT) columns. Two types of cDNA expression libraries were constructed via Stratagene protocol: an oligo(dT)-primed library inserted into the λ ZAP II vector, and a randomly primed library prepared in the ZAP expression vector. Both were amplified once and contained approximately 6 × 10^6 independent recombinants. The DGD cDNA clones were initially isolated from the oligo(dT) library with a 2.3-kb cDNA fragment derived from chicken glycine decarboxylase (CGD) as a probe (10). The CGD fragment was PCR-amplified from first-strand chicken liver cDNA using primers CGD3 and CGD4 described in Fig. 1A. To enrich for clones containing extended 5′ ends, positive plaques were re-screened with a 0.6-kb PCR product derived from chicken liver with a 2.3-kb cDNA fragment derived from chicken glycine decarboxylase (CGD).3 After electrophoresis, the p120 band (6 μg) was excised, sequenced through SDS-PAGE, and transferred to the polyvinylidene difluoride membrane.

**Table I**

| Oligonucleotides used in this study |
|-----------------------------------|
| **CGD3** | S |
| **CGD4** | AS |
| **CGD8** | S |
| **CGD2** | AS |
| **DGD24a** | S |
| **DGD27** | S |
| **DGD30** | S |
| **AUC328AUG** | S |
| **AUC346AUG** | S |
| **AUA346AUG** | S |
| **AGC931AUG** | S |
| **CUG397CGC** | S |
| **Phe1** | S |
| **Phe2** | S |
| **Phe3** | S |
| **Phe4** | S |
| **Phe5** | S |
| **Asn2** | S |
| **tr1** | S |
| **tr2** | AS |
| **tr7** | AS |
| **tr9** | S |

**DGD**. After electrophoresis, the p120 band (6 μg) was excised, sequenced through SDS-PAGE, and transferred to the polyvinylidene difluoride membrane.
placed at the 5' end of the insert. Translation of DGD from these constructs will initiate from the newly created AUG codon. For transfection experiments in mammalian cells, the entire 4.0-kb DGD insert was removed from the pBluescript vector by double digestion with BamHI and XhoI and cloned into the same sites of pcDNA3 vector (Invitrogen).

Construct tr1/2 was derived from clone 2.3.1 by attaching an artificial in-frame AUG codon into the 5' end. It expresses a DGD protein composed of residues 52–1024 but lacks 12 N-terminal residues of the mature protein. This construct was obtained by PCR amplification of clone 2.3.1 using primers tr1 and tr2 followed by cloning into the KpnI - XhoI sites of pBluescript vector. Constructs tr1/6, tr1/7, tr8/2, and tr9/2 were made in the same way using corresponding primer pairs. They expressed DGD residues 52–986, 52–922, 95–1024, 171–1024, respectively.

Cell-free Translation of DGD Protein and Pre-S binding Experiments—Plasmid DNA was purified by centrifugation through a CsCl gradient and translated in TNT-coupled transcription/translation system (Promega) using T3 or T7 RNA polymerase, rabbit reticulocyte lysates, and [35S]methionine (translational grade, NEN Life Science Products). An aliquot of the translational product (usually 1–5 μl) was applied to 0.1% SDS, 6% PAGE and after electrophoresis, radiolabeled proteins were detected by fluorography. For DHBV pre-S binding experiments, 5–10 μl of sample plus a similar amount of 35S-labeled, in

FIG. 1. Cloning and sequencing of duck glycine decarboxylase cDNAs. A, schematic representation of CGD fragments as probes for cross-hybridization and some DGD cDNA clones obtained. The CGD PCR fragments CGD3/4 and CGD8/2 match the 3' and 5' ends of the DGD cDNA, respectively. The DGD clones DRL6 and DRL5 were obtained from a randomly primed duck liver library and contained the extended 5' terminus. Clone 2.3.1 is the longest DGD clone obtained from an oligo(dT)-primed duck liver library. The complete DGD cDNA clone, DGD24a/2.3.1, was constructed by joining the 5' PstI fragment of DRL5 (1.0 kb) with the 3' PstI fragment (3.0 kb) of 2.3.1. The DGD-coding sequence spans nucleotides 346–3417. B, complete nucleotide sequence of DGD cDNA and predicted amino acid sequence of DGD. Please note that DGD protein expression is initiated from an AUU codon at position 346 (arrowhead). Underlines, peptide sequences of p120 purified from duck liver (see Table II). Dotted line, six amino acid residues missing in the chicken homologue. Boxed, poly(A) signal. Two amino acid residues as determined from peptide sequences are different from those deduced from the cDNA; an aspartic acid in the N-terminal sequence of p120 (Table II) is glutamic acid, specified by nucleotides 499–501, and an isoleucine in peak 30 (Table II) is valine, as specified by nucleotides 2359–2361.
vitro translated luciferase protein was diluted with 300 μl of lysis buffer and incubated with 2–4 μg of various GST-pre-S constructs immobilized on Sepharose beads. After 5 h of incubation at 4 °C and three washes with lysis buffer, retained 35S-labeled proteins were separated by SDS-PAGE and revealed by fluorography.

Expression of DGD cDNA Clones in Transfected Cell Lines—Plasmid DNA (20 μg) was transiently transfected into 60-mm dishes of cells by the calcium phosphate precipitation technique. At 36–42 h post-transfection, cells were metabolically labeled using the 35S express protein-labeling mix (NEN Life Science Products) at 100 μCi/ml medium for 5 h. Cells were lysed in 1 ml of lysis buffer, and an aliquot of lysate (500 μl) was precleared once with protein A Staphylococcus aureus. The DGD protein was immunoprecipitated from the precleared lysate by overnight incubation at 4 °C with 5 μl of rabbit antiserum (aBac-DGD). The immune complex was brought down by 5 μl of protein A-Sepharose beads, and radiolabeled DGD protein was detected by SDS-PAGE and fluorography. For the pre-S binding experiment, 500 μl of precleared lysate was incubated with pre-S peptide 80–102 as described above.

Immunofluorescent Staining of Cell Surface and Intracellular Expression of DGD—Cells grown on coverslips in 6-well plates were transfected with the DGD cDNA clone AUU346AUG. Cells were fixed with either paraformaldehyde (4%) or ethanol:acetic acid (95:5) 2 days post-transfection and incubated at 4 °C for 1 h in 3% bovine serum albumin/phosphate-buffered saline, 1:100 dilution of aBac-DGD in bovine serum albumin/phosphate-buffered saline for 1 h, and a 1:160 dilution of anti-rabbit Ig conjugated with fluorescein isothiocyanate for 1 h (Sigma). To prepare cell lines stably expressing the DGD protein, 293 cells were selected with G418 (200 μg/ml) 2 days after transfection. The expression of DGD was detected as described above. As a positive control for DGD expression, primary duck hepatocytes were grown on coverslips and stained as described above.

RESULTS

p120 Is the Duck p Protein Component of the Glycine Decarboxylase Complex—Purified p120 was digested with trypsin, and after HPLC separation, four peptide peaks were sequenced. Two peaks (pk30, pk68) yielded unique peptide sequences, whereas the remaining two (pk48 and pk51) produced additional minor peptide sequences (Table II). A protein data base search revealed complete homology of all the peptides with the p protein component of the multienzyme CGD (10), suggesting that the p120 pre-S-binding protein is the p protein component of DGD. The reported tissue-restricted distribution of glycine decarboxylase in the liver and kidney (13, 14) is also consistent with our previous p120 findings (liver, kidney, pancreas) (3). Subsequent Western blot analysis using rabbit antiserum prepared against DGD protein expressed in baculovirus (a Bac-DGD) confirmed its abundant expression in the liver, kidney, and pancreas (Fig. 2). Very little expression of DGD was found in the heart, lung, stomach, small intestine, gall bladder, spleen, and muscle (Fig. 2).

The cDNA Clone Encoding for the DGD Protein—DGD cDNA clones were obtained by screening an oligo(dT)-primed duck

| Peptide sequences of p120 | Location in CGD |
|-------------------------|-----------------|
| 30                      | IQPIEVDK        |
| 48                      | EVYRLAQTECHRRD  |
| 51                      | GRDQYGHP        |
| 68                      | DSQVVFDPPTEGK   |
| N-terminal              | VGGGGGAAAAA |
liver cDNA library with PCR fragments of CGD cDNA. Alignment with CGD showed the longest DGD clone, designated 2.3.1, still lacked coding sequences for the signal peptide and a few N-terminal residues of the mature protein. The missing 5’ end of DGD was subsequently obtained from the randomly primed duck liver cDNA library using the 5’ end of clone 2.3.1 as a probe. Of the two clones with the longest 5’ end, the 5’ terminus of clone DRL6 was identical to clone DRL5 (Fig. 1A). The complete nucleotide sequence of the DGD cDNA and deduced amino acid sequence of the DGD protein are shown in Fig. 1B. All p120 peptide sequences obtained were identified (Fig. 1B, underlined). The amino acid sequence of DGD demonstrated extensive homology with CGD (91% identity) except for the N-terminal leader sequence (Fig. 3) and the insertion of a VVQTRA hexapeptide (Fig. 1B, dashed line). The same hexapeptide was found in the human protein at the identical position (10). These observations would argue for a deleterional event in the chicken enzyme rather than an inserational event in the duck enzyme. The 3’-noncoding sequence has 496 nt excluding the poly(A) tail. The AATAAA polyadenylation signal was noted at nt 3889–3894 (Fig. 1B, boxed).

The N Terminus of the DGD Protein Contains Inserted Amino Acid Residues as Compared with CGD Protein—The N-terminal ~30 amino acid sequences of the p protein of glycine decarboxylase protein constitute the mitochondrial-targeting sequence, which is cleaved in the mature protein. Based on the nucleotide sequence of DRL5, the glycine decarboxylase reading frame becomes open after an in-frame TGA stop codon at nt 286–288. However, no in-frame AUG codon was found until position 589–591, which corresponds to codon 68 of CGD. In addition, the N-terminal sequence of DGD, as deduced from DRL5/DRL6, does not align with CGD. Evidence suggests that the DGD sequence derived from clone DRL5 and DRL6 is authentic.

1) 5’ rapid amplification of cDNA ends (RACE) (15) experiments using primers based on the 5’ end of clone 2.3.1 enabled us to obtain 70 nucleotides upstream of 2.3.1 identical to that of DRL5/DRL6 (data not shown).

2) N-terminal sequencing of p120 protein revealed a major peptide VGGGGGGGGG-GDAA and a minor dipeptide sequence GP (Table II). The coding sequence for the GP dipeptide can be found at nt 385–390. Careful inspection of the duck and chicken sequences revealed that by introducing gaps into the CGD sequence, extensive homology can be found between the two proteins even at the N terminus, as shown in Fig. 3. Therefore, the DGD protein contains several stretches of extra amino acid sequences as a result of insertions at the nucleotide level. In this regard, similar insertions have been reported in the N-terminal sequence of pea glycine decarboxylase relative to that of the oat enzyme (16). In addition, the N terminus of the human glycine decarboxylase is different in both length and primary sequence from that of the chicken enzyme despite an overall 84% homology when the mature proteins are compared (10).

Full-length DGD Protein Translation Is Initiated from a Non-AUG Codon—The 5’-coding sequence for DGD lacks an AUG initiation codon. The initiating AUG codon in CGD has been mutated to an AUC codon in DGD (position 328; Figs. 1B and 3), and no nearby in-frame AUG codon is found. The nearest methionine codon is found at position 589–591, corresponding to codon 68 of the CGD protein. In vitro translation of clone DGD24a/2.3.1, which contains the entire DGD-coding sequence preceded by 5’-nontranslated sequence, generated several protein species. The largest protein species is ~125 kDa (Fig. 4A). The fact that it migrated more slowly than DGD protein translated from tr1/2, which had an artificial AUG codon placed in front of nt 499 (Fig. 4A), suggests that it is translated not from AUG589 but rather from an non-AUG codon upstream of nt 499. Conversely, the second largest protein species expressed from the full-length cDNA clone migrated faster than the DGD protein generated from tr1/2 but co-migrated with the translational product of clone 2.3.1. Therefore, translation of this protein species is very likely initiated from a downstream AUG codon such as AUG589.

Determination of AUU346 as the Translational Initiation Site—Since the first amino acid residue of mature DGD protein is specified by nucleotides 463–465 (Table II and Fig. 1B), translation of the 125-kDa, full-length DGD protein should be initiated from a non-AUG codon upstream of nt 463. Besides the AUG codon at position 328 that aligns with the CGD initiation codon, several other amino acid codons upstream of position 463 differ from the AUG codon by a single nucleotide: AUU346, AUA364, AGG391, and CUG397 (Figs. 1B and 3). According to previous findings, such codons most likely act as noncanonical initiation sites (17–22). Considering that translational initiation from these different sites will generate proteins with slightly different lengths, we converted some of these codons into AUG and compared sizes of the proteins produced with that of DGD24a/2.3.1. As a result, the DGD protein produced from DGD24a/2.3.1 was smaller than that derived from the AUC328AUG construct but slightly larger than the one translated from AGG391AUG (Fig. 4A). This narrowed down the putative initiation codon to sequences between, such as AUU346 or AUA364.

To further define the initiation site, we introduced frameshift or nonsense mutations downstream of each of the five potential initiation codons. A frameshift mutation introduced between AUG328 and AUU346 (Phe1) did not effect the production of the 125-kDa protein (Fig. 4B), thus excluding AUG328 as the initiation site. In contrast, frameshift mutations placed downstream of AUU346 (Phe1), AUA364 (Phe2), AGG391 (Phe3), and CUG397 (Phe5) all abolished translation of this 125-kDa protein species (Fig. 4B). These results strongly implicate AUU346 as the initiation codon. Interestingly, the Phe2 mutation (a +1 frameshift) introduced downstream of AUU346 generated a novel protein product of about 130 kDa. This new protein species may well be a fusion product; translation initiated from an upstream non-AUG codon at −1 frame fused to the DGD open reading frame as a result of the +1 frameshift. To avoid such complications, an in-frame stop codon (Asn2) was introduced between AUU346 and AUA364. Indeed this mutation abolished the 125-kDa band without generating a new protein species (Fig. 4B). As a final proof that AUU346 is the initiation codon, we mutated the five candidate initiation codons in a codon that cannot serve as translational initiators:
Analysis of DGD Expression from Transfected Mammalian Cells; Initiation from AUU346 and Processing into a Mature Form—The DGD protein translated from the AUG346 initiation codon was slightly larger than DGD present in duck liver. Such a size difference may reflect a post-translational cleavage of the signal peptide in mammalian cells, as suggested by the N-terminal sequence of the mature DGD (Table II and Fig. 3) and the chicken sequence is shown at the bottom. Identical or related residues are linked by dots in the middle. Missing sequences in CGD are represented by dashed lines. The cleavage sites for DGD and CGD are based on N-terminal sequencing of mature proteins. For DGD, positions of the five potential translational initiation sites are shown. Horizontal arrowheads indicate the translational initiation sites for DGD and CGD, respectively.

AUG328CUC, AUU346AGU, AUAA34AGA, AGG391AGA, and CUG397CCG. Production of the 125-kDa protein was either immunoprecipitated with polyclonal antibody α Bac-DGD (A) or pulled down by GST-pre-S peptide 80–102 immobilized on Sepharose beads (B). Retained proteins were revealed by SDS-PAGE and fluorography. The small-sized protein generated from construct AUC328AUG, AUU346AUG, and DGD27/2.3.1 represents N-terminal-processed mature DGD protein. The protein band with slightly faster mobility may be the primary translational product initiated from a downstream AUG codon. This protein species is precipitated by the α Bac-DGD antibodies (A) but is not recognized by the GST-pre-S peptide (B).

CGD proteins (10). If AUU346 is the initiation site for DGD protein expression, transfection of AUU346AUG construct into mammalian cells should produce a processed DGD protein of the correct size. The AUU346AUG construct, together with AUC328AUG, AUAA34AGA, AGG391AGA, and CUG397AUG, were cloned into pcDNA3 vector and transfected into COS cells (Fig. 5). A doublet of DGD protein was produced from the AUU346AUG construct; one corresponded to the primary translational product (125 kDa), whereas the other migrated slightly faster (~120 kDa; Fig. 5A). The size of the shorter peptide is identical to that of DGD protein found in duck liver (data not shown). No such processed product was observed when AUAA34AGA, AGG391AGA, or CUG397AUG constructs were transfected, although a small amount of processed DGD protein was produced from the AUC328AUG construct (Fig. 5A). It is noteworthy that transfection of DGD27/2.3.1 into COS cells also produced a similar ratio of primary translational product and processed form (Fig. 5A). This finding indicates that AUU346 was also efficiently utilized for translational initiation in mammalian cells.

Recombinant DGD Protein Recapitulates Binding to Truncated DHBV Pre-S Protein—If DGD encodes for the pre-S-binding protein p120, then recombinant DGD protein should be capable of binding to truncated forms of pre-S protein as well. The DGD proteins translated from construct tr1/2 or expressed in COS cells from clones AUC328AUG, AUU346AUG, AUAA34AGA, AGG391AGA, and CUG397AUG were transfected, and all were shown to be competent for binding of DHBV pre-S protein (Fig. 5B and 6A). As illustrated in Fig. 6A for the tr1/2
construct, the radiolabeled DGD protein could be retained by pre-S constructs with N-terminal truncation to residue 92 or 98 (92–161 and 98–161) or C-terminal truncation to residue 102 (1–102) but not (or poorly) by the intact pre-S protein (1–161) or other types of deletion constructs such as 1–104. Moreover, several single amino acid substitutions at positions 100, 101, or 102, the critical p120 contact sites (3), abolished retention of DGD protein translated from construct tr1/2 but not (or poorly) by the intact pre-S protein (1–161) or other types of deletion constructs such as 1–104. Moreover, several single amino acid substitutions at positions 100, 101, or 102, the critical p120 contact sites (3), abolished retention of DGD protein translated from construct tr1/2 but not (or poorly) by the intact pre-S protein (1–161) or other types of deletion constructs such as 1–104, whereas the right 8 C-terminal residues (tr1/6) nearly abolished interaction.

The fact that DGD protein translated from construct tr1/2 (which misses the N-terminal 12 amino acid residues as compared with mature DGD protein) can associate with the truncated pre-S protein suggests that these residues are dispensable for the pre-S interaction. In an attempt to further define the pre-S binding site, we generated a series of N- or C-terminal deletion constructs of DGD protein and tested for their reactivity with the pre-S construct 80–102 (bottom). For each construct, three PCR clones were randomly picked for analysis. The pre-S binding capacity was abolished or severely reduced by truncation mutants with as little as 55 N-terminal residues (tr6/2) or 38 C-terminal residues (tr1/6).

**FIG. 6.** Recombinant DGD protein interacts specifically with truncated forms of DHBV pre-S protein. A, binding of DGD protein with various pre-S constructs. The 35S-labeled DGD protein translated in reticulocyte lysate from tr1/2 was mixed with lysate containing a similar amount of 35S-labeled luciferase protein and incubated at 4 °C with various pre-S constructs immobilized on Sepharose beads. After an extensive wash, the retained radiolabeled protein was separated on 10% SDS-PAGE and revealed by fluorography. The 9 lanes at the left represent retention of DGD protein by pre-S protein of different lengths, whereas the right 8 lanes show retention by wild-type (WT) pre-S peptide 80–102 or those constructs containing different point mutations. B, expression of various DGD truncation mutants in reticulocyte lysate (top) and their interaction with pre-S construct 80–102 (bottom). For each construct, three PCR clones were randomly picked for analysis. The pre-S binding capacity was abolished or severely reduced by truncation mutants with as little as 55 N-terminal residues (tr6/2) or 38 C-terminal residues (tr1/6).

**FIG. 7.** The DGD protein is available on the surface of transfected cells and primary duck hepatocytes. The PDH and DGD transiently transfected chicken hepatoma cells (LMH), Bosc, and stably transfected 293 cells were fixed with either paraformaldehyde (nonpermeabilizing (NP)) or ethanol:acetic acid (95:5) (permeabilizing (P)). The DGD cellular distribution was revealed by using a Bac-DGD polyclonal antibody followed by the fluorescein isothiocyanate-conjugated secondary antibody. Note the cell surface localization of DGD in NP cells (arrows) as compared with the diffuse cytoplasmic distribution in P cells.

Compared with the diffuse homogenous distribution of cytoplasmic DGD in permeabilized cells, the pattern observed in nonpermeabilized cells was characterized by granular or punctate distribution on the cell surface (Fig. 7). This distinct distribution was also observed with primary duck hepatocytes and 293 cells stably transfected with DGD cDNA (Fig. 7). Thus, the mitochondrial enzyme DGD may be found both in the cytoplasm and on the cell surface of reconstituted cells and PDH.

**DISCUSSION**

The p120 pre-S-binding protein (2) has now been established as the p protein of DGD following cDNA cloning. The partial amino acid sequences purified from p120 duck liver matched the translated cDNA sequences, and comparison from different species validated p120 as the p protein component of the glycine decarboxylase complex. The tissue distribution of DGD and, specifically, unique patterns of binding to truncated DHBV pre-S and mutants confirmed the identity of DGD as p120.

Although the duck p protein is highly homologous to that of the corresponding chicken molecule, there is significant divergence in the N terminus encoding the putative mitochondrial-targeting domain. Interestingly, the DGD protein not only is distributed in the cytoplasm but is also available on the cell
surface as described previously (3) and confirmed in the present study (Fig. 7); these DGD findings differ from CGD described as solely a mitochondrial protein. Whether the divergent 5’ sequence is responsible for subcellular and cell surface localization warrants further study.

The duck glycine decarboxylase p protein is translated from an AUU codon based on extensive cDNA mutational analysis followed by expression of the mutant constructs in a cell-free system. An in-frame nonsense mutation (Asn5) placed immediately downstream of AUU^346 abolished 125-kDa protein production and was similar to a point mutation that converted AUU^346 into an AGU codon (Fig. 4, B and C). Moreover, use of the non-AUG codon for initiation also occurred in transfected DGD27/2.3.1 construct in mammalian cells (Fig. 5), indicating that translation from cell lysates was not merely a result of relaxed specificity caused by high potassium concentration. Whether AUU codon selection (compared with other nearby AUG-like codons) requires structural motifs i.e. hairpin structure downstream to slacken the passage of scanning ribosomes, warrants further study.

Initiation from an internal AUU codon of the DGD cDNA was apparently not optimal in either reticulocyte lysates or transfected mammalian cells, since conversion of the AUU into the AUG codon and deletion of the 5’-nontranslated region greatly enhanced protein yield. The low efficiency of protein expression is due, in part, to the presence of the nontranslated sequence at the 5’ end, since the DGD27/2.3.1 construct produced a higher yield of protein than DGD24a/2.3.1. Whether such an inhibitory effect is caused by the translation of the upstream small open reading frames or by the presence of a secondary structure impeding the entry of scanning ribosomes remains unknown.

Victorin, the toxin produced by the fungus Cochliobolus victoriae, uses the p protein of the oat glycine decarboxylase as the binding protein (16); inhibition of the enzymatic function by victorin is believed to account for the blight of oats (6). The identification of glycine decarboxylase as the binding partner for DHBV pre-S protein and truncated species will allow us to directly test the role of glycine decarboxylase as a DHBV co-receptor or co-factor facilitating productive viral infection by cDNA transfection experiments.

REFERENCES

1. Ganem, D. (1996) in Fields Virology (Fields, B. N., Knipe, D. M., and Howley, P. M., eds) Lippincott-Raven Publishers, Philadelphia
2. Tong, S., Li, J., and Wands, J. R. (1995) J. Virol. 69, 7106–7112
3. Li, J., Tong, S., and Wands, J. R. (1996) J. Virol. 70, 6029–6035
4. Kuroki, K., Cheung, R., Marion, P. L., and Ganem, D. (1994) J. Virol. 68, 2091–2096
5. Kuroki, K., Eng, F., Ishikawa, T., Turck, C., Harada, F., and Ganem, D. (1995) J. Biol. Chem. 270, 15022–15028
6. Navarre, D., and Wolpert, T. J. (1995) Plant Cell 7, 463–471
7. Nakayama, K. (1997) Biochem. J. 327, 625–635
8. Lambert, V., Fernholz, D., Sprengel, R., Fourel, I., Deleage, G., Wildner, G., Peyret, C., Trepo, C., Cova, L., and Will, H. (1990) J. Virol. 64, 1290–1297
9. Chassot, S., Lambert, V., Kay, A., Godinot, C., Roux, B., Trepo, C., and Cova, L. (1993) J. Virol. 67, 217–223
10. Kume, A., Koyata, H., Sakakibara, T., Ishiguro, Y., Kure, S., and Hiraga, K. (1991) J. Biol. Chem. 266, 3323–3329
11. Ho, S. N., Hunt, H. D., Horton, R. M., Pullen, J. K., and Pease, R. (1989) Gene 77, 51–59
12. Tong, S., Li, J., Vitvitski, L., Leop, D., and Trepo, C. (1991) J. Med. Virol. 35, 228–231
13. Yoshida, T., and Kikuchi, G. (1973) J. Biochem. (Tokyo) 73, 1013–1022
14. Kure, S., Koyata, H., Kume, A., Ishiguro, Y., and Hiraga, K. (1991) J. Biol. Chem. 266, 3330–3334
15. Frohman, M. A., Dush, M. K., and Martin, G. R. (1988) Proc. Natl. Acad. Sci. U. S. A. 85, 8996–9002
16. Wolpert, T. J., Navarre, D. A., Moore, D. L., and Macko, V. (1994) Plant Cell 6, 1145–1155
17. Peabody, D. S. (1989) J. Biol. Chem. 264, 5031–5035
18. Mehdi, H., Ono, E., and Gupta, K. C. (1990) Gene 91, 173–178
19. Saric, C., Domen, J., and Berns, A. (1991) EMBO J. 10, 655–664
20. Chalut, C., and Egly, J. M. (1995) Gene 156, 43–45
21. Goldberger, G., Dalska, M., Grobner, P., and Piendl, W. (1995) J. Bacteriol. 177, 5994–5996
22. Schmitz, J., Pruler, D., Rohde, W., and Tacke, E. (1996) Nucleic Acids Res. 24, 257–263