Cloning and Expression of a cDNA Coding for a Rat Liver Plasma Membrane Ecto-ATPase

THE PRIMARY STRUCTURE OF THE ECTO-ATPase IS SIMILAR TO THAT OF THE HUMAN BILIARY GLYCOPROTEIN

Sue-Hwa Lin* and Guido Guidotti

From the Department of Biochemistry and Molecular Biology, Harvard University, Cambridge, Massachusetts 02138

The amino acid sequence of the ecto-ATPase from rat liver was deduced from analysis of cDNA clones and a genomic clone. Immunoblots with antibodies raised against a peptide sequence deduced from the cDNA sequence indicated that the determined amino acid sequence is that of the ecto-ATPase. The deduced sequence predicts a 519-amino acid protein with a calculated molecular mass of 57,388 daltons. There are 16 potential asparagine-linked glycosylation sites in the protein. Hydropathy analysis of the deduced amino acid sequence indicates that the protein has two hydrophobic stretches. One is located at the N-terminal and the other is near the C-terminal end.

A full-length clone encoding the ecto-ATPase was expressed transiently in mouse L cells and human HeLa cells. The cell lysate from the transfected cells contained immunoreactive ecto-ATPase and Ca2+-stimulated ATPase activities. The expressed protein is glycosylated and has an apparent molecular weight (100,000) similar to that of the rat liver plasma membrane ecto-ATPase.

The presence of an ecto-ATPase activity has been observed on the surface of many cells. The general characteristics of this activity are its activation to a similar extent by either Ca2+ or Mg2+ and its ability to hydrolyze nonspecifically several different nucleotides. The function(s) for the enzyme is not clear. It has been suggested that the ecto-ATPase may be transported into the cell by nucleoside transporters for the recapture of nucleosides by the cell.

This work was supported in part by Grant HL08893 from the National Institutes of Health (to G. G.). 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.

EXPERIMENTAL PROCEDURES

Materials—The peptide was synthesized by Biosearch (San Rafael, CA). Trypsin was purchased from Calbiochem (San Diego, CA). Soybean trypsin inhibitor, KLH, and MBS were purchased from Sigma.

Protein Purification and Preparation of Tryptic Peptides—The ecto-ATPase was isolated from rat liver plasma membrane using the method described in Lin and Fain (1984) with modifications (Lin, 1989). The purified ecto-ATPase (300 μg) in a volume of 3 ml was reduced by the addition of 30 μl of β-mercaptoethanol and 800 μl of 10% SDS followed with heating in a boiling water bath for 20 min. The reduced and denatured ecto-ATPase was then carboxymethylated with iodoacetamide (159 mg) at room temperature overnight. The protein was dialyzed against 10 mM ammonium bicarbonate to remove SDS and unreacted iodoacetamide. An aliquot of the resulting protein solution containing 180 μg of modified ecto-ATPase was digested with 4 μg of trypsin (1:45 w/w ratio of trypsin/ecto-ATPase) at 37 °C. After 6 h, 4 μg of additional trypsin were added and the digestion was continued for an additional 12 h. The trypsin in the digestion mixture was inactivated by the addition of 4 μg of soybean trypsin inhibitor (1.5 μl/mg of trypsin inhibitor/trypsin). The digests were subjected to SDS-PAGE.
Sequence of Liver Ecto-ATPase

Construction of the Expression Vector (pExp) and Expression of Recombinant Ecto-ATPase in L Cells and HeLa Cells—Expression plasmid, cdm8 (Seed, 1987), was obtained from T. Maniatis (Harvard University). This expression vector has the human cytomegalovirus AD169 IE promoter and 3'-poly(A) addition signals. The two cDNA clones, Clone 1 and Clone 2, were joined at a common restriction enzyme site (Fig. 3A). The resulting cDNA sequence was further digested with restriction enzyme Nhel and the 1850-base pair Nhel/EcoRI fragment was isolated. A fragment obtained from the genomic clone which contains the 5' upstream sequence of the ecto-ATPase was digested with restriction enzymes XhoI and Nhel to generate a fragment of 249 base pairs (assigned as position −243 to +6 of the complete sequence). This genomic fragment was joined to the 1850-base pair Nhel/EcoRI fragment at the Nhel site and then further cut with restriction enzyme SnaI to generate a DNA fragment of 1837 base pairs containing the entire coding region for the ecto-ATPase (Fig. 3A). This 1807-base pair fragment was cloned into the expression vector cdm8 through XhoI, PstI restriction enzyme sites.

L cells and HeLa cells were transfected with 25 μg of plasmid DNA/100-mm plate by the calcium phosphate precipitation method. After 48 h, the culture medium was aspirated and the cells were rinsed sequentially with cell buffer (150 mM NaCl, 5 mM KCl, 10 mM Hepes, pH 7.4) plus 2 mM EDTA twice and cell buffer once. One ml of cell buffer was added to each dish, and the cells were scraped off dishes with a rubber policeman. Another 1-ml cell buffer was added to rinse the remaining cells out of the dish. Both fractions were combined and the cells were centrifuged in an Eppendorf microfuge for 30 s. The supernatants were removed and the cells were disrupted by mixing with 0.5 ml of distilled water followed by freezing and thawing to obtain a total cell lysate.

Other Procedures—To determine the position of the exon-intron junction in the genomic clone, a restriction fragment from the genomic clone was cloned into PBS-M13+ plasmid (Stratagen, CA) through XhoI. SacI restriction sites. Various lengths of ribonucleotide probes were made from the plasmid by T7 RNA polymerase. RNAse protection experiments were then performed as described by Melton et al. (1984) using 30 μg of total liver RNA per sample.

The Ca2+-stimulated ATPase activity of the cell lysate was measured by incubating 50 μl of the cell lysate in 2 ml at 37 °C in an assay medium which contained in a final volume of 250 μl, 20 mM sodium azide, 50 mM Hepes-Tris (pH 7.4), 2 mM EGTA, 2 mM ATP (disodium salt), with or without 2 mM Ca2+ (this gives 1.6 μM free [Ca2+]). The ATPase activity was determined by inorganic phosphate release as described (Lin, 1985b). The Ca2+-stimulated ATPase activity was determined by subtracting values obtained with EGTA alone from those with calcium plus chelator.

To test whether the expressed protein can bind to a ConA-Sepharose 4B column, the total lysate (150 μl) was solubilized by the addition of 100 μl of detergent C6E6 (100 μg/ml). The solubilized samples were then centrifuged in an Eppendorf microfuge, and the supernatant fractions were diluted through ConA-Sepharose 4B gel (75 μl) packed in 1-ml blue pipette tips. The flow-through fractions from the ConA-Sepharose 4B columns were lyophilized and used for SDS-PAGE analysis. The materials which bound to the ConA-Sepharose 4B gel were eluted by using 100 μl of 0.5 M α-methylmannoside in 1 mg/ml C6E6, and the eluates were lyophilized and used for SDS-PAGE analysis.

Protein was determined by the method of Peterson (1977) using bovine serum albumin as a standard.

RESULTS

Partial Amino Acid Sequence Analysis of the Ecto-ATPase and Confirmation of the Sequence by Using Antipeptide Antibodies—Attempts to obtain the N-terminal amino acid sequence of the ecto-ATPase were not successful indicating that the N-terminal amino acid of the ecto-ATPase is blocked.

Construction of the Expression Vector (pExp) and Expression of Recombinant Ecto-ATPase in L Cells and HeLa Cells—Expression plasmid, cdm8 (Seed, 1987), was obtained from T. Maniatis (Harvard University). This expression vector has the human cytomegalovirus AD169 IE promoter and 3'-poly(A) addition signals. The two cDNA clones, Clone 1 and Clone 2, were joined at a common restriction enzyme site (Fig. 3A). The resulting cDNA sequence was further digested with restriction enzyme Nhel and the 1850-base pair Nhel/EcoRI fragment was isolated. A fragment obtained from the genomic clone which contains the 5' upstream sequence of the ecto-ATPase was digested with restriction enzymes XhoI and Nhel to generate a fragment of 249 base pairs (assigned as position −243 to +6 of the complete sequence). This genomic fragment was joined to the 1850-base pair Nhel/EcoRI fragment at the Nhel site and then further cut with restriction enzyme SnaI to generate a DNA fragment of 1837 base pairs containing the entire coding region for the ecto-ATPase (Fig. 3A). This 1807-base pair fragment was cloned into the expression vector cdm8 through XhoI, PstI restriction enzyme sites.

L cells and HeLa cells were transfected with 25 μg of plasmid DNA/100-mm plate by the calcium phosphate precipitation method. After 48 h, the culture medium was aspirated and the cells were rinsed sequentially with cell buffer (150 mM NaCl, 5 mM KCl, 10 mM Hepes, pH 7.4) plus 2 mM EDTA twice and cell buffer once. One ml of cell buffer was added to each dish, and the cells were scraped off dishes with a rubber policeman. Another 1-ml cell buffer was added to rinse the remaining cells out of the dish. Both fractions were combined and the cells were centrifuged in an Eppendorf microfuge for 30 s. The supernatants were removed and the cells were disrupted by mixing with 0.5 ml of distilled water followed by freezing and thawing to obtain a total cell lysate.

Other Procedures—To determine the position of the exon-intron junction in the genomic clone, a restriction fragment from the genomic clone was cloned into PBS-M13+ plasmid (Stratagen, CA) through XhoI. SacI restriction sites. Various lengths of ribonucleotide probes were made from the plasmid by T7 RNA polymerase. RNAse protection experiments were then performed as described by Melton et al. (1984) using 30 μg of total liver RNA per sample.

The Ca2+-stimulated ATPase activity of the cell lysate was measured by incubating 50 μl of the cell total lysate for 2 h at 37 °C in an assay medium which contained in a final volume of 250 μl, 20 mM sodium azide, 50 mM Hepes-Tris (pH 7.4), 2 mM EGTA, 2 mM ATP (disodium salt), with or without 2 mM Ca2+ (this gives 1.6 μM free [Ca2+]). The ATPase activity was determined by inorganic phosphate release as described (Lin, 1985b). The Ca2+-stimulated ATPase activity was determined by subtracting values obtained with EGTA alone from those with calcium plus chelator.

To test whether the expressed protein can bind to a ConA-Sepharose 4B column, the total lysate (150 μl) was solubilized by the addition of 100 μl of detergent C6E6 (100 μg/ml). The solubilized samples were then centrifuged in an Eppendorf microfuge, and the supernatant fractions were diluted through ConA-Sepharose 4B gel (75 μl) packed in 1-ml blue pipette tips. The flow-through fractions from the ConA-Sepharose 4B columns were lyophilized and used for SDS-PAGE analysis. The materials which bound to the ConA-Sepharose 4B gel were eluted by using 100 μl of 0.5 M α-methylmannoside in 1 mg/ml C6E6, and the eluates were lyophilized and used for SDS-PAGE analysis.

Protein was determined by the method of Peterson (1977) using bovine serum albumin as a standard.

RESULTS
Peptide Sequence

SVLLAHNLQPQFPQV

Oligonucleotides

5' ACT TGI AAT TCT TGI GGI AII TTI TG 3'

Fig. 1. Amino acid sequence and the corresponding nucleotide sequence of a tryptic peptide obtained from the ecto-ATPase. The peptide sequence was obtained from the fraction eluting at 84 min. The isoaspartic containing oligonucleotides which were used as hybridization probes for the ecto-ATPase cDNA are shown and are complementary to the coding strand.

Pre-immune rabbit serum (Fig. 1B). The sequence HNLPQEFQV underlined in Fig. 1B was used as hybridization probes for the ecto-ATPase cDNA are shown.

Table 1. Pre-immune Anti- Anti-

A B C

PM ecto PM ecto PM ecto

Hydropathy analysis, according to the algorithm of Kyte and Doolittle (1982), of the deduced amino acid sequence indicates that the protein has two hydrophobic stretches. One of them, located at the N terminus of the protein (amino acids 17–34), has the character of a signal peptide (Van Heijne, 1983); the other, located near the C-terminal end (amino acids 424–448), is the most hydrophobic portion of the protein. This C-terminal hydrophobic segment is likely to be a membrane anchoring domain; the hydrophilic segment (amino acids 449–519) following the C-terminal hydrophobic segment is, therefore, the cytoplasmic domain of the protein. The deduced primary structure contains 16 potential N-linked glycosylation sites (Hubbard and Ivatt, 1981) located in a region between amino acids 87 and 374 suggesting that at least part of this region is extracellular. The C terminus of the postulated sequence for the ecto-ATPase contains N-terminal methionine residues (Fig. 1B) and is the true initiation codon. The combined sequence consists of a 1557-base pair open reading frame and a 300-base pair 3' untranslated sequence. There is no polyadenylation signal or poly(A) tract at the 3' end of the cDNA. The amino acid sequence determined for the ecto-ATPase clone obtained by overexpressing a genomic clone is shown in Fig. 3B.

To confirm that the determined amino acid sequence was that of the ecto-ATPase, antibodies were raised against a peptide with the sequence SVLLAHNLQPQFPQVC. The extra cysteine was added to the C-terminal end of the peptide in order to couple the peptide to keyhole limpet hemocyanin with the bifunctional coupling agent MBS. Rabbit antisera were reacted with both the plasma membrane ecto-ATPase and the purified ecto-ATPase (Fig. 2B). The pattern of the immunoblot by the antipeptide antibodies is similar to that of the immunoblot by the antisera 708 (Lin, 1989). The missing 5' sequence was therefore obtained by isolating a genomic ecto-ATPase clone and sequencing the relevant region. To determine the intron-exon junction at the 5' end of the clone, an RNAse protection experiment was performed. By using probes HindIII, SacI, and XbaI/SacI, bands of 58 and 127 nucleotides were observed, respectively (data not shown). This corresponds to an exon beginning position -64 in Fig. 3B. The entire sequence of the ecto-ATPase clone obtained by overlapping Clone 1, Clone 2, and the genomic clone is shown in Fig. 3B. The presence of an inframe termination codon in position -18 (Fig. 3B) and the results from the expression of the combined sequence, which will be discussed later, suggest that the ATG coding for the methionine labeled as the first amino acid of the ecto-ATPase is the true initiation codon. The combined sequence consists of a 1557-base pair open reading frame and a 300-base pair 3' untranslated sequence. There is no polyadenylation signal or poly(A) tract at the 3' end of the cDNA. The amino acid sequence determined for the tryptic peptide is located at amino acid residues 449–519.

Expression of the Ecto-ATPase in Mammalian Cell—To
**Fig. 3. Nucleotide sequence and deduced amino acid sequence of the ecto-ATPase.** A, alignment of rat liver ecto-ATPase cDNA clones and the genomic clone. Two cDNA clones obtained by screening two cDNA libraries and one genomic clone were aligned by comparison of nucleotide sequences. The combined sequence is obtained by combining the sequences of the three clones. The lines indicate introns. Several restriction endonuclease recognition sites are also indicated. B, nucleotide sequence and deduced amino acid sequence of the ecto-ATPase. The sequence which contains the entire coding region of the ecto-ATPase was obtained by using the entire sequence from Clone 2, 49 nucleotides (5' end) of Clone 1, and 68 nucleotides from the genomic clone. Stars mark the potential asparagine-linked glycosylation sites. The peptide whose structure was determined by amino acid sequencing is underlined. The sequence which contains the transmembrane region is indicated by a box.
methylmannoside (Fig. 4) while there was no detectable level of the expressed protein in the ConA-Sepharose 4B column flow through fraction (Fig. 4). These results indicate that the expressed protein has undergone post-translational glycosylation. When the cell lysate was separated into supernatant and pellet fractions by centrifugation, the expressed protein was associated with the pellet (data not shown), suggesting that the expressed protein is membrane-bound.

The Ca\(^{2+}\)-stimulated ATPase activity of the expressed protein was measured by using the cell lysates from cells transfected with a control plasmid and the plasmid pEXP. As shown in Table I, L cells have low basal plasma membrane Ca\(^{2+}\)-ATPase activity while HeLa cells had much higher basal Ca\(^{2+}\)-ATPase activity. Cells transfected with plasmid pEXP had higher Ca\(^{2+}\)-stimulated ATPase activity compared to that of controls indicating that the protein expressed from plasmid pEXP is the ecto-Ca\(^{2+}\)-ATPase.

**TABLE I**

<table>
<thead>
<tr>
<th>Cell type</th>
<th>Ca(^{2+})-ATPase activity</th>
<th>% control</th>
<th>nmol/mg/h</th>
</tr>
</thead>
<tbody>
<tr>
<td>L cells</td>
<td>Control</td>
<td>14.6</td>
<td>38.0</td>
</tr>
<tr>
<td>HeLa cells</td>
<td>pEXP</td>
<td>166.2</td>
<td>251.8</td>
</tr>
</tbody>
</table>

**FIG. 4. Expression of the rat liver ecto-ATPase protein in mouse L cell and human HeLa cells.** Preparation of the total cell lysate and the ConA-Sepharose 4B separation of proteins are described under "Experimental Procedures." The total lysate (50 µl) from each sample was used for SDS-PAGE analysis. All the samples were boiled in SDS sample buffer containing β-mercaptoethanol (Laemmli, 1970) for 5 min before loading onto a 10% SDS polyacrylamide gel. Blotting of proteins onto nitrocellulose membrane filters and immunoblot analysis of the expressed protein were performed as described under "Experimental Procedures." The antipeptide antiserum in 1-500-dilution was used.

**DISCUSSION**

This paper reports the primary sequence of the rat liver plasma membrane ecto-ATPase deduced by molecular cloning of the gene for the ecto-ATPase. Three lines of evidence show that the clone codes for the amino acid sequence of the ecto-ATPase. First, amino acids 50-69 are the same as those obtained by microsequencing of a tryptic peptide from purified ecto-ATPase; antibodies raised against the peptide recognize the ecto-ATPase protein. Second, mammalian cells transfected with a plasmid containing the clone produce a glyco-protein which is recognized by anti-ecto-ATPase antibodies and antipeptide antibodies in immunoblots. Third, the protein expressed in the transfected cells has Ca\(^{2+}\)-stimulated ATPase activity.

Due to the method used for the construction of the cDNA library, it is not uncommon, especially when there is low abundance of the specific mRNA, to have cDNA clones which are missing the 5′ end of the coding sequence. Several methods have been used to obtain information about the 5′ end. In a situation where the message abundance is high, sequencing poly(A) RNA directly using an oligonucleotide complementary to the specific RNA may be done (Geliebter, 1987). Another approach is to make a specifically primed cDNA library using, in the first strand reaction, an oligonucleotide sequence present in the 5′ end of the isolated clone; this method has been successfully used to obtain the complete cDNA sequence of the human vitamin D receptor (Baker et al., 1988) and for the bovine brain GABA receptor (Schofield et al., 1987). Attempts to use these two approaches were not successful. This may be due to the extremely low abundance of the message, as indicated by the fact that only one partial clone was obtained from each cDNA library screened, or because the mRNA encoding the ecto-ATPase has strong secondary structure. Finally, the missing 5′ end of the coding sequence may be obtained by cloning and sequencing the genomic clone for the protein. The problem with using the genomic sequence is to determine whether the deduced sequence is part of an intron or an exon. This problem can be addressed by RNase protection experiments using a ribonucleotide probe derived from the genomic sequence and tissue RNA. In this case, the complete coding sequence for the ecto-ATPase was obtained by a combination of cDNA and genomic DNA cloning. The following observations support the view that the five nucleotides (nucleotides 1-5) obtained from the genomic clone are the true 5′ end of the coding region for the ecto-ATPase. First, there is an inframe termination codon at position −18, indicating that the methionine codon is the first inframe initiation codon. Second, the N-terminal part of the deduced protein sequence (amino acids 17-34) has the characteristics of a hydrophobic signal peptide. Third, transfection of the combined sequence into mammalian cell lines results in the expression of a protein which is glycosylated and has the same size as the native protein indicating that the protein has a proper N-terminal signal sequence for translation and membrane insertion.

The binding of the ecto-ATPase to ConA-Sepharose 4B and periodic acid-Schiff staining demonstrate that the ecto-ATPase is a glycoprotein (data not shown). To determine the molecular weight of the protein without the carbohydrate, the effects of various glycosidases on the apparent molecular weight of the protein were examined (data not shown). After treatment of the purified ecto-ATPase with neuraminidase, the apparent molecular weight was reduced from approximately 100,000 to 85,000. This M, 85,000 protein still traveled as a diffused band in SDS-PAGE indicating that this protein was still glycosylated. However, no change in apparent molecular weight was observed when the purified ecto-ATPase was treated with either endo-β-N-acetylglucosaminidase H, O-glycanase, or N-glycanase. The combination of neuraminidase with endo-β-N-acetylglucosaminidase H, O-glycanase, or N-glycanase did not change the apparent molecular weight more than did neuraminidase alone. The possibility that the ecto-ATPase may be folded into a conformation not accessible to the glycosidases was tested. The purified ecto-ATPase was denatured with 5 M guanidine HCl followed by reduction and alkylation with N-ethylmaleimide. The modified protein was
then dialyzed and treated with various combinations of glycosidases. No change in apparent molecular weight by the glycosidases other than by neuraminidase was observed (data not shown). This result indicates that there are carbohydrates, which can not be removed by these glycosidases, on the protein. Comparison of the calculated molecular weight from the sequence ($M_w = 57,000$) with the apparent molecular weight by SDS-PAGE ($M_w = 100,000$) of the ecto-ATPase indicates that about 40% of the apparent mass is contributed by the carbohydrate moiety. This is consistent with the existence of 16 possible N-linked glycosylation sites in the primary sequence. It is also possible that some serine residues in the hydrophilic region may be glycosylated through O-linked sugar residues. The observation that transfection of the clone which could be translated into a protein with a molecular weight of 57,000 into HeLa and L cells results in expression of a glycoprotein with apparent molecular weight of 100,000 is consistent with the idea that the ecto-ATPase is heavily glycosylated. The precise determination of the glycosylation status of the protein will require expression of the clone in mutant mammalian cells defective in various glycosylation enzymes.

The hydrophathy plot shows that the deduced protein sequence for the ecto-ATPase has two hydrophobic stretches. One, which is located at the C-terminal end may serve as a membrane anchor while the other, located at the N terminus, may be a signal sequence or a transmembrane stretch. The structural arrangement predicts that most of the protein mass is on the outside of the cell, and the C terminal of the protein is intracellular. This structural information is consistent with the fact that the protein is an ecto-ATPase with its active site on the outside of the cell. Several ecto-enzymes, i.e. \( \gamma \)-glutamyltranspeptidase (Laperche et al., 1986) and alkaline phosphatase (Berg et al., 1987; Millan, 1986; Henthorn et al., 1986; Kam et al., 1985) have been found to have similar structural arrangements in the membrane. The N-terminal hydrophobic stretch may be the signal sequence for the ecto-ATPase. However, since the N-terminal amino acid of the purified ecto-ATPase is blocked, the true N terminus of the protein is unknown. By examining the deduced protein sequence, two possibilities may be raised to explain the blocked N-terminal amino acid. One is that the N-terminal methionine is acetylated \( \alpha \)-amino groups. Since the predicted N-terminal amino acid sequence for the ecto-ATPase has two hydrophobic stretches. Another possibility for the blocked N-terminal amino acid is that the ecto-ATPase has a signal sequence which is cleaved at residue 94 (alanine) leaving residue 35 (glutamine) as the N-terminal amino acid of the mature protein. In this case, the blocked N-terminal may be due to cyclization of \(^{34}\)Gln to pyroglutamate. These possibilities are presently under investigation.

The ecto-ATPase has broad substrate specificity. It can hydrolyze not only ATP, but also GTP, CTP, UTP to a similar extent (Lin, 1985b). It has been pointed out by Walker et al. (1982) that the consesus sequence G-(X)_3-GK(T) and R/K-(X)_3-G-(X)_3-L-(hydrophobic)_3-D is present in a number of ATPases and they suggested that this pair of sequences represent an ATP-binding domain in these proteins. In the deduced sequence for the ecto-ATPase, the sequence at amino acid positions 92–100, G-(PAYS)-GR(E)(T), and the sequence at amino acid position 335–348, K-(EL)-G-(SVT)-L-(TCFSK)-D, are similar to the consensus sequences. However, these putative nucleotide binding domains of the ecto-ATPase are much less homologous to the consensus sequences than those of ion transport ATPases (Shull and Greeb, 1988). It is possible that such a lower homology may contribute to the difference in the substrate specificity between the ion transport ATPase and the ecto-ATPase. The activities of ion transport ATPases, i.e. Na,K*-ATPase and Ca**-ATPase, are inhibited by vanadate. It was proposed that the inhibition of ATPase activity is due to the similarity in the structure of vanadate with that of the transition state for the phosphorylated intermediate. Molecular cloning of the sarcoplasmic reticulum Ca**-ATPases and Na,K*-ATPases and plasma membrane Ca**-ATPases suggests that a sequence (C-S-D-K-T-G-T-L-T) corresponding to the phosphorylation domain is conserved in these ion-pumping ATPases (Shull and Greeb, 1988). The ecto-ATPase activity is not inhibited by vanadate (Lin, 1985b) indicating that the mechanism of ATP hydrolysis by the ecto-ATPase may not involve an aspartyl phosphate intermediate. As a result, the phosphorylation domain of those ion-transporting ATPases is not found in the deduced ecto-ATPase sequence. However, it was also found that although P,Fo-type ATPases are not phosphorylated, the \( \beta \) chains have a sequence similar to the phosphorylation site sequence of ion-transport ATPases. In this corresponding sequence is T-T-T-K-K-S-I-T (Walker et al., 1982; Jorgensen and Andersen, 1988). A portion of the ecto-ATPase (T-V-K-E-L-G-S-V-T) located at residues 332–1161

**FIG. 5.** Comparison of the amino acid sequences of rat liver ecto-ATPase and human biliary glycoprotein 1. The comparison shows was performed using the program BESTFIT from the University of Wisconsin sequence analysis software. The upper sequence is in ecto-ATPase, and the lower sequence is in biliary glycoprotein 1.
Sequence of Liver Ecto-ATPase

341 is similar to that of the β chain of the ATPase of bovine mitochondria. Whether or not these regions of protein sequence are involved in nucleotide hydrolysis needs to be confirmed by other approaches.

Recently, Hinoda et al. (1988, 1989) reported the primary structure of a glycoprotein immunologically cross-reactive with carcinoembryonic antigen. The protein is believed to be a biliary glycoprotein 1 even though its cDNA was cloned from an adult human colon tissue cDNA library. A comparison of the primary structures of the rat liver ecto-ATPase with that of the human biliary glycoprotein 1 (BGPI) shows that they are highly homologous (Fig. 5), with 65% of the amino acids being identical. BGPI was known as a protein which has an antigenic site(s) that cross-react with carcinoembryonic antigen (Svenberg, 1976). According to Hinoda et al. (1988), the primary structure of the BGPI can be divided into several domains: the signal sequence, an N-terminal domain, an immunoglobulin-like domain consisting of two subdomains, a BGPI-specific domain, a membrane spanning domain, and a cytoplasmic domain. The N-terminal domain and the immunoglobulin-like domain are highly homologous to those of carcinoembryonic antigen. There is a BGPI-specific domain which consists of 108 amino acids and is located near the C-terminal part of the protein before the transmembrane domain.

The two putative nucleotide-binding sites of the ecto-ATPase can be identified in the BGPI sequence: one is located in the N-terminal domain and the other in the BGPI-specific domain. From the similarity in the primary sequences of the BGPI and the rat liver ecto-ATPase, it can be suggested that the BGPI is a human ecto-ATPase. It is possible that the sequence of the protein reported by Hinoda et al. (1988) is that of the colon enzyme and not that of the human liver enzyme. Thus, the homology between the rat liver and the human liver enzymes may be even greater than that observed here. In any event, the similarity in structure raises the question as to whether the BGPI is an ATPase.

In the previous studies on the localization of the ecto-ATPase in rat liver (Lin, 1989) we have suggested that depending on the location of the ecto-ATPase, the functions of the ecto-ATPase may be different. The sinuousid ecto-ATPase may function in regulating the concentration of ATP as a ligand for P2-purinergic receptors while the canalicular ecto-ATPase may function as part of the retrieval mechanism for the recovery of adenosine. It is possible, therefore, that human BGPI has similar functions. More evidence will be needed to confirm these speculations drawn from the sequence similarity.

Acknowledgments—We thank Drs. Richard Southgate, David Paul, and Eugene Rosenthal for help with molecular cloning; Dr. Mike Mueckler for providing the rat liver cDNA library; Chen-Ming Fan for help with expression and RNase protection experiments.

REFERENCES


Gellibter, J. (1987) Focus 9, 5-8


Helms, C., Graham, M. Y., Dutchtik, J. E., and Olson, M. V. (1985) DNA (N.Y.) 4, 39–49


