Proc. Natl. Acad. Sci. USA
Vol. 81, pp. 7338–7342, December 1984
Biochemistry

Rat liver asialoglycoprotein receptor lacks a cleavable NH_2-terminal signal sequence

(cDNA cloning/oligonucleotide probe/transmembrane protein biosynthesis)

ERIC C. HOLLAND, JUDITH O. LEUNG, AND KURT DRICKAMER

Department of Biochemistry and Molecular Biology, The University of Chicago, 920 East 58th Street, Chicago, IL 60637

Communicated by Joseph L. Goldstein, August 6, 1984

ABSTRACT **Two cDNA clones encoding the predominant form of the asialoglycoprotein receptor from rat liver (the major rat hepatic lectin; RHL-1) were identified by screening a rat liver cDNA library with a mixed oligonucleotide probe 35 nucleotides long. One clone was a nearly full-length copy of the mRNA for RHL-1, while the other was shortened at both ends. The sequences of these clones demonstrate that this transmembrane receptor is not synthesized with an NH_2-terminal signal sequence. The only proteolytic processing occurring in the biosynthesis of RHL-1 is the removal of the NH_2-terminal initiator methionine residue. Insertion of RHL-1 into the membrane is postulated to occur by the recognition of the internal transmembrane region as a signal sequence.**

The hepatic lectins are receptors that mediate the endocytosis of glycoproteins. The rat liver receptors recognize terminal galactose, which is exposed after removal of sialic acid from the complex carbohydrate group of plasma glycoproteins. The chicken receptor binds glycoproteins only after both the sialic acid and galactose are removed to expose *N*-acetylglucosamine (1). Preparations of the rat hepatic lectin contain two different receptor polypeptide species; primary structure analysis reveals that the two species are 50% identical in the COOH-terminal 100 amino acid residues (2). The relative abundance of these species of receptor, designated RHL-1 and RHL-2/3, is approximately 4:1. The chicken hepatic lectin shows strong sequence homology with the rat receptors near the COOH terminus.

Several lines of evidence indicate that both the chicken hepatic lectin and the predominant species of rat hepatic lectin (RHL-1) are transmembrane proteins oriented with their NH_2 termini in the cytoplasm and their COOH termini outside the cell. Each protein can be divided into three domains: an NH_2-terminal hydrophilic sequence of 20–40 residues, a hydrophobic sequence of 20–25 residues, and a large COOH-terminal hydrophilic sequence. Since the NH_2-terminal domain of the chicken hepatic lectin contains a site of phosphorylation, it has been suggested that this part of the polypeptide is accessible to cytoplasmic kinases (3). Both the chicken and rat hepatic lectins are glycosylated on residues in their COOH-terminal domains, which indicates that this region is on the extracellular surface of the plasma membrane (2, 4). Vectorial labeling studies provide direct evidence that in whole cells the COOH termini of the hepatic lectins are exposed to the medium, while the NH_2 termini are exposed to the cytoplasm (33).

Many secretory proteins as well as some transmembrane proteins have been shown to be synthesized with an NH_2-terminal signal sequence that is removed cotranslationally (5). Transmembrane proteins synthesized by this mechanism invariably are arranged with their NH_2 termini on the outside of the cell (6). Since the glycoprotein receptors have the opposite orientation, they appear to be members of a different class of membrane proteins. Other proteins in this class include the invariant γ chain of the class II histocompatibility antigens (7) and the influenza virus neuraminidase (8). Therefore, insertion of this group of proteins into the membrane must proceed by a different pathway.

To study the biosynthesis of this latter class of membrane proteins we have obtained a nearly full-length cDNA copy of the mRNA for RHL-1. The structure of this cDNA indicates that the mRNA does not encode a hydrophobic NH_2-terminal extension of RHL-1 and that the only proteolytic processing occurring in the biosynthesis of this protein is the removal of the initiator methionine residue.

MATERIALS AND METHODS

Materials. $(dT)_{12-18}$ was purchased from Collaborative Research (Waltham, MA). Radioactively labeled nucleotides and most restriction enzymes were obtained from Amersham. *Dde* I was purchased from New England Biolabs. Guanidine thiocyanate was a product of Fluka, while oligo(dT)-cellulose, globin mRNA, and phage T4 polynucleotide kinase were obtained from Bethesda Research Laboratories. RNAsin (placental ribonuclease inhibitor) was purchased from Promega Biotec (Madison, WI). S1 endonuclease was from Boehringer Mannheim. Terminal transferase was purchased from P-L Biochemicals and reverse transcriptase was from Life Sciences (St. Petersburg, FL).

Synthesis of Oligonucleotide Probe. The oligonucleotide used as a probe for RHL-1 cDNA and mRNA was synthesized on a solid support by the phosphoramidite method, using an Applied Biosystems 380A DNA synthesizer (9). It was end-labeled by using [γ-^{39}P]ATP and T4 polynucleotide kinase (10) and characterized by sequence analysis (see below).

Preparation of mRNA. Extraction of total RNA from rat livers (frozen in liquid nitrogen immediately after sacrifice of the animal) with guanidine thiocyanate was followed by a single acid/ethanol precipitation and centrifugation through CsCl (11). Poly(A)$^+$ RNA was purified by two cycles of oligo(dT)-cellulose chromatography (10). The yield of poly(A)$^+$ RNA was approximately 1% of total RNA; typically, 50 μg of poly(A)$^+$ RNA was obtained from 1 g of liver.

Blot Hybridization of Poly(A)$^+$ RNA. RNA denatured with glyoxal was separated by electrophoresis on a 1% agarose gel (10). The RNA was transferred to nitrocellulose by capillary blotting in 20× NaCl/Cit (1× NaCl/Cit = 0.15 M NaCl/15 mM sodium citrate, pH 7.0), baked at 80°C for 2 hr, and boiled in 50 mM Tris·HCl, pH 8.0, for 10 min. The filter was incubated at 58°C for 2 hr in a solution containing 20% (vol/vol) deionized formamide, 0.75 M NaCl, 50 mM Na_2HPO_4, 0.5 mM EDTA, 5× Denhardt's solution (see p. 327 of ref. 10), denatured salmon sperm DNA at 150 μg/ml, and 0.1% $NaDodSO_4$. The filter-bound RNA was hybridized

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviation: RHL, rat hepatic lectin.

in the same solution with 4×10^4 cpm/ml of end-labeled oligonucleotide (4×10^6 cpm total) for 48 hr at 58°C. The filter was washed in two changes of 5× NaCl/Cit at 58°C for 30 min and exposed to Kodak XAR-5 x-ray film with a DuPont Lightning Plus intensifying screen for 7 days at −70°C.

Construction and Screening of Recombinant Plasmids. Double-stranded cDNA was prepared by the method of Wickens *et al.* (12) with the following modifications. For the first strand synthesis, the concentration of each of the four nucleoside triphosphates was 1 mM; RNAsin was included at 400 units/ml. After first strand synthesis, RNA was hydrolyzed in 20 μl of 300 mM NaOH and 1 mM EDTA for 1 hr at 42°C. The resulting solution was neutralized with 1 M Tris·HCl, pH 7.5, and passed through a 150-μl column of Sephadex G-100 equilibrated with second strand synthesis buffer. Second strand synthesis conditions were exactly as for the first strand except that KCl was omitted from the buffer. S1 endonuclease digestion of double-stranded cDNA was performed in a buffer containing 300 mM NaCl, 30 mM sodium acetate at pH 4.5, 3 mM $ZnSO_4$, and 30 units of S1 nuclease per 50 μl for 1 hr at 37°C. cDNA at each step in the synthesis was assessed by alkaline agarose gel electrophoresis (10). The S1 endonuclease-digested double-stranded cDNA was tailed with oligo(dC) by using terminal transferase (13) and was annealed with plasmid pUC 8 that had been digested with *Pst* I and tailed with oligo(dG) (14). Transformation of *Escherichia coli* strain HB101 was by the $CaCl_2$ shock procedure (ref. 10, p. 250). To achieve higher transformation efficiency, transformants were plated directly onto agar containing ampicillin (35 μg/ml). After colonies were grown up, they were transferred to nitrocellulose by placing a dry, sterile filter on top of the agar. When the filter was completely wet it was lifted and placed inverted on a fresh agar plate. If the filter is initially dry, this procedure provides complete and efficient transfer of colonies from agar to nitrocellulose. After replication and preparation for hybridization (ref. 10, pp. 314–316), filters were washed for 2 hr at 58°C in a solution containing 50 mM Tris·HCl at pH 8.0, 1 mM EDTA, 1 M NaCl, and 0.1% $NaDodSO_4$ to reduce background. Colonies bound to nitrocellulose were screened by the method of Grunstein and Hogness (15), using end-labeled oligonucleotide as probe. Prehybridization, hybridization, and washing procedures were the same as for the RNA blot hybridization analysis except that *E. coli* DNA was included in the prehybridization and hybridization buffers at 50 μg/ml.

Sequence Analysis. Plasmid DNA was isolated by the alkali lysis procedure followed by a single CsCl density centrifugation, ribonuclease treatment, and centrifugation through 1 M NaCl (ref. 10, pp. 90–95). After restriction endonuclease digestion (ref. 10, p. 104), DNA was prepared for sequence analysis either by labeling at the 5′ end with [γ-^{32}P]ATP and polynucleotide kinase or by labeling at the 3′ end with dideoxy-[α-^{32}P]ATP and terminal transferase, following the instructions provided by the manufacturer. Labeled fragments were sequenced by the chemical cleavage method of Maxam and Gilbert (16).

RESULTS AND DISCUSSION

Design and Synthesis of Oligonucleotide Probe. Although it has been reported that RHL-1 may make up as much as 0.1% of the total protein in rat liver (17), the value may actually be as low as 1/10th of this (unpublished observations). These results indicate that the mRNA for RHL-1 is likely to be a very low abundance species, constituting as little as 0.01% of the total messenger pool. The cloning of such low-abundance mRNA has been greatly simplified by the use of synthetic oligonucleotide probes (18). Since the complete amino acid sequence of RHL-1 has been determined, it is possible to predict a portion of the mRNA sequence and to design a very specific probe that will hybridize to the RHL mRNA.

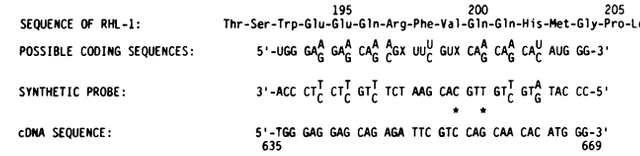

FIG. 1. Design of oligoncleotide probe. The top line indicates the amino acid sequence of a portion of RHL-1 (2). The mRNA sequences that could code for this segment are given on the second line (X refers to any of the four nucleosides). The third line gives the synthetic oligonucleotide used as a probe, while the bottom line is the actual sequence of the cDNA (sense strand) derived from sequence analysis of clones 1 and 22. Mismatches between the cDNA sequence and the synthetic probe are indicated with asterisks.

The strategy for the design of a suitable oligonucleotide probe is summarized in Fig. 1. It was decided to make a probe in excess of 30 nucleotides long so that hybridization and washing could be carried out under moderately stringent conditions. In spite of the degeneracy of the genetic code, it is possible to select a stretch of 12 amino acid residues in the known sequence of RHL-1 for which the mRNA can be predicted with few ambiguities. The amino acids from positions 193 to 204, with the exceptions of Arg-197 and Val-199, can be encoded by only one or two codons each (see Fig. 1). Even so, synthesis of all sequences that could encode this portion of RHL-1 would result in a mixture of 4096 different probes. While the mixture of probes actually synthesized contained all possible coding sequences in the terminal portions of the probe, ambiguity in the central region was eliminated in order to reduce the complexity of the probe mixture. Selection of bases in the central region was based on codon usage frequency (19) and the G·T base-pairing rule, since G·T base pairs lend some stability to double-stranded DNA (20). This resulted in a mixture of 32 different 35mers, all with a common (and, as it turns out, partially incorrect) central sequence. The protein sequence between residues 193 and 204 of RHL-1 is not highly homologous to RHL-2/3; therefore this probe mixture is specific for RHL-1.

RNA Blot. When rat liver mRNA is separated by electrophoresis through agarose, transferred to nitrocellulose, and probed with the synthetic oligonucleotide described above, a single band is detected (see Fig. 2). The RNA species detected is approximately 1375 bases long; an mRNA of this length

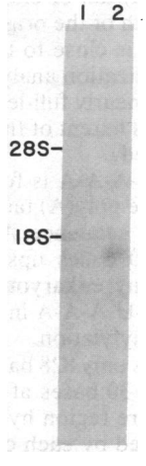

FIG. 2. Blot analysis of RNA from rat liver. Rat liver RNA was separated by electrophoresis through agarose and transferred to nitrocellulose. The filter was hybridized with the end-labeled oligonucleotide probe. Lane 1, 50 μg of poly(A)$^-$ RNA; lane 2, 30 μg of poly(A)$^+$ RNA. The mobilities of 18S and 28S rRNAs are indicated to the left.

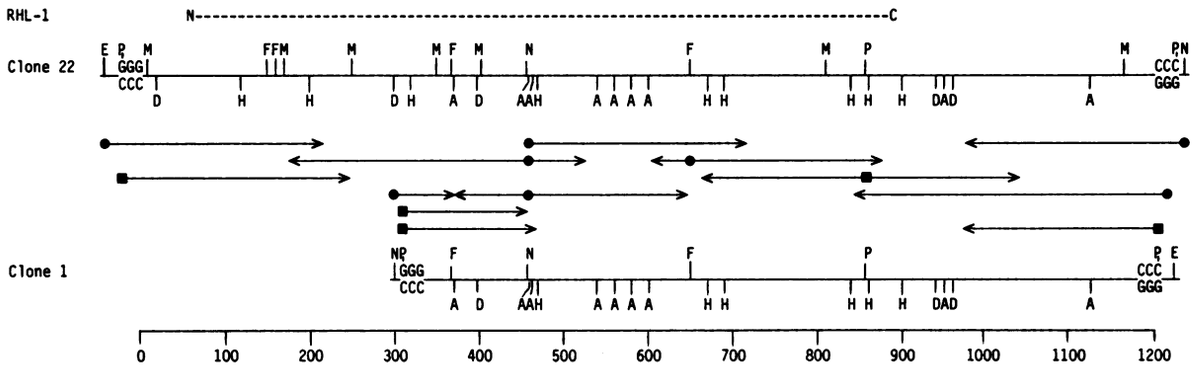

FIG. 3. Restriction map and sequencing strategy for clones 1 and 22. RHL-1 is aligned above the two clones for orientation. At the bottom, bases are numbered from the 5′ end of clone 22. Restriction enzymes are designated as follows: A, *Alu* I; D, *Dde* I; E, *Eco*RI; F, *Hin*fI; H, *Hae* III; M, *Mbo* II; N, *Hin*dIII; and P, *Pst* I. Digested DNA was labeled at its 5′ termini (●) or 3′ termini (■) and subjected to secondary restriction digestion to separate the labeled ends. Fragments in the top three lines were derived from clone 22, while those in the lower three lines derived from clone 1.

could encode a protein the size of RHL-1 (852 bases) plus 3′ and 5′ untranslated regions as well as a poly(A) tail (21). Inclusion of formamide in the hybridization solution and use of elevated temperatures for hybridization and washing results in a low level of nonspecific binding of the radioactive probe to the nitrocellulose filter and to unrelated RNA sequences. From these results it was concluded that the oligonucleotide can be employed under these hybridization conditions to identify colonies containing cloned copies of the RHL-1 mRNA.

Construction and Screening of cDNA Library. A cDNA library was constructed, using $(dT)_{12-18}$ to prime reverse transcription of total rat liver mRNA. Conditions for cDNA synthesis were optimized by testing on globin mRNA. When annealed with linearized pUC 8 vector, 0.2 μg of the double-stranded cDNA gave 17,000 transformants, of which 2 hybridized to the oligonucleotide probe. These 2, designated clones 1 and 22, were reproducibly positive upon rescreening, although clone 1 consistently gave a weaker signal than did clone 22.

Sequence Analysis. The strategy employed to sequence the inserts in clones 1 and 22 is summarized in Fig. 3. The results show that both clones are copies of the message for RHL-1 (see Fig. 4). The cDNA insert in clone 22 contains 1169 bases preceding a stretch of 17 A residues at the 3′ end. Eukaryotic mRNAs typically have poly(A) tails with a length of 200 bases (21). If the poly(A) tail of clone 22 were extended to 200 bases, the total length of the original mRNA would have been 1369 bases, which is close to the length of 1375 bases measured by blot hybridization analysis. These data suggest that clone 22 contains a nearly full-length copy of the RHL-1 mRNA, although the full extent of the 5′ untranslated region remains to be established.

The sequence A-G-U-A-A-A is found 21 bases upstream from the beginning of the poly(A) tail. This sequence closely matches the consensus sequence A-A-U-A-A-A, which is found approximately 20 bases upstream from the site of poly(A) addition in many eukaryotic mRNAs (22). Therefore, the sequence A-G-U-A-A-A in the RHL-1 mRNA may be a signal for polyadenylylation.

The insert in clone 1 is only 828 bases long, and it lacks 328 bases at the 5′ end and 30 bases at the 3′ end compared to clone 22. Since the entire region hybridizing to the oligonucleotide probe is encoded by each clone, the differential affinity of the probe for the two clones is not readily explained. It might reflect a different copy number for the insert-bearing plasmids in the two clones.

Comparison of cDNA and Protein Sequences. Examination of the base sequence at the 5′ end of clone 22 reveals that the coding sequence for RHL-1 immediately follows the first AUG codon. This indicates that a single methionine residue directly precedes the threonine residue found at the NH_2 terminus of mature RHL-1 (2). Removal of the initiator methionine is common in cytoplasmic proteins (23), a fact that is consistent with evidence that the NH_2 terminus of this receptor is located in the cytoplasm. The presence of a stop codon in frame with the coding sequence 54 bases upstream from this AUG codon eliminates the possibility that translation initiation could occur at another AUG codon upstream from the cloned portion of the mRNA.

The sequence of bases around the first AUG codon in the RHL-1 mRNA closely resembles the consensus sequence found near other eukaryotic translation initiation sites: specifically, there are purines at positions −3 and +4 as well as pyrimidines at positions −5, −4, −2, and −1 (24). The RHL-1 mRNA contains adenosines at −3 and +4, which are optimal for promoting translation (25).

When the amino acid sequence of RHL-1 obtained by Edman degradation is compared with that deduced from the DNA sequence, only four differences are found (see Fig. 4). All of these discrepancies are in the region directly following the hydrophobic sequence; since this region was sequenced by extended Edman degradation of large peptides, misidentification of residues in this region is not surprising. As a result of these corrections, the hydrophobic sequence is revealed to be 21 residues long rather than 27 as originally proposed (2). The presence of an arginine residue at the COOH terminus of the hydrophobic sequence means that the transmembrane segment of RHL-1 is bounded at both ends by arginine residues; in this respect, RHL-1 resembles the chicken hepatic lectin (4) and many other membrane-spanning proteins (26).

When the DNA sequences of clones 1 and 22 are compared, a discrepancy is observed. While clone 22 contains an adenosine residue at position 685, clone 1 contains a guanosine residue at the corresponding position. Since the presence of an adenosine residue at this position would change the codon for Trp-209 to a termination codon, it was concluded that the adenosine residue found in clone 22 was an error in reverse transcription (27) and that the guanosine residue found in clone 1 corresponds to the original mRNA sequence.

Analysis of purified RHL-1 shows that the COOH-terminal residue is Asn-283 (2). In the RHL-1 mRNA, the sequence encoding this residue is followed immediately by a termination codon (see Fig. 4). Therefore, the COOH terminus of the initial translation product is identical to that of the mature protein.

Implications for Receptor Biosynthesis. The "signal hypothesis" was developed to explain the discharge of secre-

```
                                                                                                        58
                                         T TAG CCT GGG CTC TTC GTG GTG CTG AGG GAC CTT CAG TCC TGC CCC AGT GCT ATC ATG
                                           ***                                                                    Met-

                                                                                                               148
ACA AAG GAT TAT CAA GAT TTC CAG CAC TTG GAC AAT GAG AAC GAC CAC CAT CAA CTC CAG AGA GGG CCA CCT CCC GCT CCA AGG CTC TTG
Thr-Lys-Asp-Tyr-Gln-Asp-Phe-Gln-His-Leu-Asp-Asn-Glu-Asn-Asp-His-His-Gln-Leu-Gln-Arg-Gly-Pro-Pro-Pro-Ala-Pro-Arg-Leu-Leu-

                                                                                                               238
CAG CGA CTC TGC TCT GGA TTC CGT CTC TTC CTG CTT TCC CTG GGC CTC AGC ATC CTG CTG CTG GTG GTT GTC TGT GTG ATC ACA TCC CGA
Gln-Arg-Leu-Cys-Ser-Gly-Phe-Arg-Leu-Phe-Leu-Leu-Ser-Leu-Gly-Leu-Ser-Ile-Leu-Leu-Leu-Val-Val-Val-Cys-Val-Ile-Thr-Ser-Arg-
                                    ------------------------------------------------------------Ser---------------- Gln

                                                                                                               328
AAT TCC CAA CTC CGG GAA GAT CTG CGG GTT CTA AGG CAG AAT TTC AGC AAC TTT ACC GTG AGC ACT GAG GAC CAA GTC AAG GCC CTG ACC
Asn-Ser-Gln-Leu-Arg-Glu-Asp-Leu-Arg-Val-Leu-Arg-Gln-Asn-Phe-Ser-Asn-Phe-Thr-Val-Ser-Thr-Glu-Asp-Gln-Val-Lys-Ala-Leu-Thr-
Val             Val

|<- Clone 1                                                                                                    418
ACC CAG GGA GAG AGA GTG GGA AGA AAG ATG AAG TTA GTC GAG TCA CAG CTG GAA AAA CAT CAG GAG GAT CTG AGG GAA GAC CAC TCT AGA
Thr-Gln-Gly-Glu-Arg-Val-Gly-Arg-Lys-Met-Lys-Leu-Val-Glu-Ser-Gln-Leu-Glu-Lys-His-Gln-Glu-Asp-Leu-Arg-Glu-Asp-His-Ser-Arg-

                                                                                                               508
TTG CTA CTG CAT GTA AAG CAG TTA GTG TCT GAC GTG CGA AGC TTG AGC TGC CAG ATG GCC GCA CTT CGG GGC AAT GGC TCT GAA AGG ATC
Leu-Leu-Leu-His-Val-Lys-Gln-Leu-Val-Ser-Asp-Val-Arg-Ser-Leu-Ser-Cys-Gln-Met-Ala-Ala-Leu-Arg-Gly-Asn-Gly-Ser-Glu-Arg-Ile-

                                                                                                               598
TGC TGC CCC ATC AAC TGG GTG GAG TAT GAA GGC AGC TGC TAC TGG TTC TCC AGC TCT GTG AAG CCT TGG ACG GAA GCT GAC AAG TAC TGC
Cys-Cys-Pro-Ile-Asn-Trp-Val-Glu-Tyr-Glu-Gly-Ser-Cys-Tyr-Trp-Phe-Ser-Ser-Ser-Val-Lys-Pro-Trp-Thr-Glu-Ala-Asp-Lys-Tyr-Cys-

                                                                                                          *    688
CAG CTG GAG AAC GCC CAC CTG GTG GTG GTG ACT TCC TGG GAG GAG CAG AGA TTC GTC CAG CAA CAC ATG GGC CCC TTA AAT ACT TGG ATT
Gln-Leu-Glu-Asn-Ala-His-Leu-Val-Val-Val-Thr-Ser-Trp-Glu-Glu-Gln-Arg-Phe-Val-Gln-Gln-His-Met-Gly-Pro-Leu-Asn-Thr-Trp-Ile-

                                                                                                               778
GGC CTA ACT GAC CAG AAC GGA CCC TGG AAA TGG GTG GAT GGG ACA GAC TAT GAG ACA GGC TTC AAG AAC TGG AGA CCA GGG CAG CCA GAT
Gly-Leu-Thr-Asp-Gln-Asn-Gly-Pro-Trp-Lys-Trp-Val-Asp-Gly-Thr-Asp-Tyr-Glu-Thr-Gly-Phe-Lys-Asn-Trp-Arg-Pro-Gly-Gln-Pro-Asp-

                                                                                                               868
GAC TGG TAC GGA CAT GGG CTT GGA GGG GGT GAA GAC TGT GCC CAC TTC ACC ACC GAT GGC CAC TGG AAT GAT GAC GTC TGC AGG AGG CCC
Asp-Trp-Tyr-Gly-His-Gly-Leu-Gly-Gly-Gly-Glu-Asp-Cys-Ala-His-Phe-Thr-Thr-Asp-Gly-His-Trp-Asn-Asp-Asp-Val-Cys-Arg-Arg-Pro-

                                                                                                               958
TAC CGC TGG GTC TGT GAG ACA GAG TTG GGC AAG GCC AAT TAG GAG CCT CCA TTC CCC TGA TTT ATT TCC TGA GTG CCT TGA GCT GTG AGG
Tyr-Arg-Trp-Val-Cys-Glu-Thr-Glu-Leu-Gly-Lys-Ala-Asn ***                     ***             ***         ***

                                                                                                              1048
CTT AGA GTT GGG AGG TCC TAC TCA AGG GTC TCC CCA TCT TTG GGA ACT TTC ATC TAG GAT TTT AAG GGG CTG GTA AAG AAT GGT GTT TGA
                                                                        ***                                         ***

                                                                                                              1138
GGA ATG TGG TAT GAT GCC TGG TGG TGG TGG TGG TGG TGG TTA TTG GAA CCC ATG ACA CTT TGT GCA GTG TGC AGC TTA CTA CTG CCA ACT

            Clone 1 ->|                         1186
TGT TTT TGA G[AG TAA A]AG AAG GAA GAT AAA CAA AAA AAA AAA AAA AAA
        ***       ***
```

FIG. 4. Nucleotide sequence of RHL-1 cDNA as derived from clones 1 and 22. Amino acids encoded in the open reading frame are shown below the DNA sequence. Termination codons are marked by *** and the putative signal for poly(A) addition is enclosed in a box. Differences between the amino acid sequence of RHL-1 deduced from the cDNA sequence and the amino acid sequence determined by Edman degradation are indicated by underlining; the original assignments from the protein sequence analysis are shown below the DNA-derived assignments. The membrane-spanning sequence of RHL-1 is indicated by a broken line. The single * at base 685 marks the discrepancy between clones 1 and 22. Clone 22 contains an A at this point, which would change the tryptophan codon to a termination codon.

tory proteins into the lumen of the endoplasmic reticulum (5). An NH_2-terminal signal sequence on a nascent polypeptide results in movement of the ribosome complex to the endoplasmic reticulum membrane. As translation continues, the nascent peptide is directed through the membrane. In the lumen of the endoplasmic reticulum the signal sequence is cleaved from the remainder of the protein by a signal peptidase. Translation and transfer through the membrane continue until the entire protein is in the lumen.

The signal hypothesis has been extended to explain the insertion of proteins into the endoplasmic reticulum membrane (6). As in the case of secretory proteins, an NH_2-terminal signal sequence directs the initial insertion of the protein into the membrane, but transfer through the membrane is believed to be halted when a hydrophobic, membrane-spanning sequence is encountered; the protein remains embedded in the membrane with its NH_2 terminus in the lumen and its COOH terminus in the cytosol. In fact, several membrane-spanning proteins that have this transmembrane orientation are actually synthesized with transient NH_2-terminal signal sequences (28, 29).

Insertion of RHL-1 into the membrane cannot be explained by this extension of the signal hypothesis for two reasons: first, the results reported here demonstrate that there is no NH_2-terminal signal sequence preceding the coding region for RHL-1, and second, the final orientation of RHL-1 in the plasma membrane is the opposite of that which would result from this insertion mechanism. Thus we must propose an alternative model for the insertion of this receptor into the membrane. Although the insertion process must be different from what is described above, it is still possible that a signal mechanism is involved. The hydrophobicity

profile of the transmembrane region of RHL-1 is very similar to that of known signal sequences (30). Therefore, the signal machinery [including the signal recognition particle (31)] could recognize the internal hydrophobic region of the nascent polypeptide as a signal sequence and promote insertion of this region as a hairpin into the membrane. Translation and transfer across the endoplasmic reticulum membrane would proceed as in the case of a secretory protein, except that the NH_2-terminal portion of the protein would remain in the cytoplasm and signal peptidase would not cleave the peptide as it emerges into the lumen of the endoplasmic reticulum. The COOH terminus would eventually traverse the membrane, leaving the NH_2 terminus still in the cytoplasm and the hydrophobic region embedded in the lipid bilayer. Thus the transmembrane segment would serve as an internal signal sequence during the biosynthesis of RHL-1.

Such a mechanism would explain the biosynthesis of several membrane proteins in addition to the hepatic lectins. Since the invariant γ chain of the class II histocompatibility antigens (8), the influenza virus neuraminidase (7), and membrane glycoprotein PE_2 of Sindbis virus (32) all have the same transmembrane orientation as RHL-1 and none have NH_2-terminal signal sequences, it is proposed that all of these proteins are inserted into membranes by a similar mechanism. Consistent with this proposal, recent studies indicate that the internal, membrane-spanning sequences of influenza virus neuraminidase and Sindbis virus glycoprotein PE_2 can, in fact, act as signal sequences (7, 32).

We thank Donald Steiner and Shu Chan for their help in synthesizing the oligonucleotide probe and Naoko Tanese and Stephen Goff for their helpful suggestions. This research was supported by National Institutes of Health Grants GM30823 and AM20595 and by Grant 82-A-127 from the Searle Scholars Program. K.D. is the recipient of a Career Development Award from the Schweppe Foundation. E.C.H. is supported by National Institutes of Health Training Grant 5T32 GM07183.

1. Ashwell, G. & Harford, J. (1982) *Annu. Rev. Biochem.* **51,** 531–554.
2. Drickamer, K., Mamon, J. F., Binns, G. & Leung, J. O. (1984) *J. Biol. Chem.* **259,** 770–778.
3. Drickamer, K. & Mamon, J. F. (1982) *J. Biol. Chem.* **257,** 15156–15161.
4. Drickamer, K. (1981) *J. Biol. Chem.* **256,** 5827–5839.
5. Dobberstein, B. & Blobel, G. (1975) *J. Cell Biol.* **67,** 835–851.
6. Rothman, J. E. & Lenard, J. (1977) *Science* **195,** 743–753.
7. Bos, T. J., Davis, A. R. & Nayak, D. P. (1984) *Proc. Natl. Acad. Sci. USA* **81,** 2327–2331.
8. Claesson, L., Larhammar, D., Rask, L. & Peterson, P. A. (1983) *Proc. Natl. Acad. Sci. USA* **80,** 7395–7399.
9. Beaucage, S. L. & Caruthers, M. H. (1981) *Tetrahedron Lett.* **22,** 1859–1862.
10. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY).
11. Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J. & Rutter, W. J. (1979) *Biochemistry* **18,** 5294–5299.
12. Wickens, M. P., Buell, G. N. & Schimke, R. T. (1978) *J. Biol. Chem.* **253,** 2483–2495.
13. Nelson, T. & Brutlag, D. (1979) *Methods Enzymol.* **68,** 41–50.
14. Vieira, J. & Messing, J. (1982) *Gene* **19,** 259–268.
15. Grunstein, M. & Hogness, D. S. (1975) *Proc. Natl. Acad. Sci. USA* **72,** 3961–3965.
16. Maxam, A. M. & Gilbert, W. (1980) *Methods Enzymol.* **65,** 499–560.
17. Hudgin, R. L., Pricer, W. E., Jr., Ashwell, G., Stockert, R. J. & Morell, A. G. (1974) *J. Biol. Chem.* **249,** 5536–5543.
18. Itakura, K., Rossi, J. J. & Wallace, R. B. (1984) *Annu. Rev. Biochem.* **53,** 323–356.
19. Jaye, M., de la Salle, H., Schamber, F., Balland, A., Kohli, V., Findeli, A., Tolstoshev, P. & Lecocq, J.-P. (1983) *Nucleic Acids Res.* **11,** 2325–2335.
20. Uhlenbeck, O. C., Martin, F. H. & Doty, P. (1971) *J. Mol. Biol.* **57,** 217–229.
21. Lewin, B. (1980) *Gene Expression* (Wiley Interscience, New York), Vol. 3, p. 658.
22. Proudfoot, N. J. & Brownlee, G. G. (1976) *Nature (London)* **263,** 211–214.
23. Wold, F. (1984) *Trends Biochem. Sci.* **9,** 256–257.
24. Kozak, M. (1981) *Nucleic Acids Res.* **9,** 5233–5252.
25. Kozak, M. (1984) *Nature (London)* **308,** 241–246.
26. Sabatini, D. D., Kreibich, G., Morimoto, T. & Adesnik, M. (1982) *J. Cell Biol.* **92,** 1–22.
27. Battula, N. & Loeb, L. A. (1974) *J. Biol. Chem.* **249,** 4086–4093.
28. Lingappa, V. R., Katz, F. N., Lodish, H. F. & Blobel, G. (1978) *J. Biol. Chem.* **253,** 8667–8670.
29. McCune, J. M., Lingappa, V. R., Fu, S. M., Blobel, G. & Kunkel, H. G. (1980) *J. Exp. Med.* **152,** 463–468.
30. von Heijne, G. (1982) *J. Mol. Biol.* **159,** 537–541.
31. Walter, P. & Blobel, G. (1980) *Proc. Natl. Acad. Sci. USA* **77,** 7112–7116.
32. Bonatti, S., Migliaccio, G., Blobel, G. & Walter, P. (1984) *Eur. J. Biochem.* **140,** 499–502.
33. Chiacchia, K. B. & Drickamer, K. (1984) *J. Biol. Chem.*, in press.