Proc. Natl. Acad. Sci. USA
Vol. 79, pp. 3260–3264, May 1982
Genetics

γ heavy chain disease in man: cDNA sequence supports partial gene deletion model

(immunoglobulin genes/recombinant DNA/deletion mutant/*in vitro* translation/amino acid sequence)

ALICE ALEXANDER*, MICHAEL STEINMETZ†, DENIS BARRITAULT‡§, BLAS FRANGIONE‡, EDWARD C. FRANKLIN‡¶, LEROY HOOD†, AND JOEL N. BUXBAUM*‡‖

*General Medical Research, New York Veterans Administration Medical Center, 408 First Avenue, New York, New York 10010; †Division of Biology, California Institute of Technology, Pasadena, California 91125; and ‡Irvington House Institute, Departments of Medicine and Pathology, New York University School of Medicine, 550 First Avenue, New York, New York 10016

Contributed by Edward C. Franklin, February 25, 1982

ABSTRACT **Human γ heavy chain disease (HCD) is characterized by the presence in serum of a short monoclonal Ig γ chain unattached to light chains. Although most HCD proteins have internal deletions, in some the defect is NH_2-terminal. The OMM γ3 HCD serum protein is of the latter type, having undergone an extensive NH_2-terminal deletion with a sequence starting within the hinge. A cell line synthesizing the OMM protein has enabled us to study the biogenesis of the abnormal molecule. *In vitro* translation of isolated mRNA yields a protein containing a hydrophobic NH_2-terminal leader sequence. In the intact cell, the precursor molecule is processed normally to yield a protein with an NH_2-terminal sequence homologous to the beginning of the variable (V) region. The nucleotide sequence of cDNA prepared from the OMM mRNA encodes a 19-amino acid leader followed by the first 15 residues of the V region. An extensive internal deletion encompasses the remainder of the V and the entire C_H1 domain. Immediately following the short V region, there is information in the cDNA for the entire normal hinge. The primary synthetic product is thus an internally deleted molecule that undergoes postsynthetic degradation to yield the NH_2-terminally deleted serum protein. The structure of the OMM mRNA suggests that the protein abnormality results from a partial gene deletion rather than defective splicing.**

In heavy chain disease (HCD), a naturally occurring human lymphoproliferative disorder, variant monoclonal Ig heavy (H) chain fragments are found in the patients' serum or urine. HCD proteins of the γ class have been studied in detail, although those of the α and μ classes have also been described (1, 2). The abnormal serum proteins of the γ class all show complete deletion of the C_H1 domain of the constant (C) region. Additional features separate the proteins into three types. Most of the variant molecules studied to date have a normal or aberrant NH_2-terminal variable (V)-region sequence of variable size, followed by an extensive internal deletion encompassing most of the V region and the C_H1 domain. Normal structure usually resumes at the beginning of the hinge. In the second group, the deletion continues through the hinge, with normal sequence starting in the C_H2 domain. Least frequently, γ HCD proteins are found that lack the entire V and C_H1 domains. In such cases, the serum protein sequence starts within the hinge, and it is therefore unclear whether the molecules are of degradative or synthetic origin.

A γ3 HCD protein was isolated from the serum of patient OMM and shown to have a monomeric molecular weight of 40,000 and an unblocked NH_2 terminus (3). It had undergone an extensive NH_2-terminal deletion with a homogeneous sequence starting within the hinge. In addition to the HCD protein, the patient's serum also contained a homogeneous $(\gamma3)_2\lambda_2$ molecule with apparently normal H and light (L) chains (4). Because the HCD protein was NH_2-terminally deleted, it was not possible to tell whether it was the product of partial proteolytic degradation.

A permanent cell line, established from the peripheral blood of the patient, gave us the opportunity to study the HCD protein at several stages in its biosynthetic history (5).

Analysis of the radiolabeled molecules synthesized by the cells and precipitated with specific anti-γ and anti-L chain antisera showed that the line produced only the HCD protein and that the intracellular and secreted forms were identical in mobility on $NaDodSO_4$/polyacrylamide gels. Short-term labeling studies did not reveal degradation of a normal-sized H chain to yield the short polypeptide.

mRNA was extracted from the cells and shown, by sucrose gradient centrifugation, to be smaller than the message coding for a normal γ3 chain (15.5S vs. 17S, respectively). When translated *in vitro* by using a protein-synthesizing extract derived from wheat germ (6) or reticulocytes (7), the mRNA directed the synthesis of a protein 2,000 daltons larger than the unglycosylated protein synthesized by the cells in the presence of the glycosylation inhibitor 2-deoxyglucose. Such an extension corresponds to approximately 20 additional amino acids.

MATERIALS AND METHODS

mRNA Isolation. The OMM cells were maintained in culture, in RPMI-1640 medium with 5% fetal calf serum. For preparation of mRNA (5), cells were harvested from spinner cultures at a density of 5×10^5 per ml and frozen at −80°C. In early preparations, cytoplasmic RNA was extracted by using the $NaDodSO_4$/phenol/chloroform method. In later work, total cellular RNA was prepared by the guanidine thiocyanate procedure (8). The poly(A)-containing mRNA fraction was isolated by chromatography on oligo(dT)-cellulose (P-L Biochemicals) (5).

***In Vitro* Translation.** For preparation of the *in vitro* translated OMM protein (OMM_{Tr}) (5), mRNA was added to wheat germ lysate (6) in the presence of [^{35}S]methionine, [^{3}H]leucine, or [^{3}H]phenylalanine, and the radiolabeled OMM_{Tr} was precipitated from the translation mixtures by the addition of anti-

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.

Abbreviations: HCD, heavy chain disease; H and L, heavy and light chains of immunoglobulins; V and C, variable and constant regions of immunoglobulin chains; kb, kilobase(s).

§ Present address: Institut National de la Sante et de la Recherche Medicale, Unite de Recherches Gerontologiques, 29, rue Wilhem, 75016 Paris, France.

¶ Deceased, Feb. 20, 1982.

‖ To whom reprint requests should be addressed, at the New York Veterans Administration Medical Center.

human γ antiserum plus carrier human IgG. After extensive washing, the OMM_{Tr} specific precipitates were dissolved in 5 M guanidine/0.1 M Tris·HCl, pH 8.3. The disulfide bonds were reduced by incubation with 5 mM dithiothreitol for 1 hr at 37°C and alkylated with 10 mM iodoacetamide for an additional hour. In preparation for amino acid sequence analysis, the completely reduced protein was dialyzed extensively against 1 M HOAc.

Amino Acid Sequence Analysis. Amino acid sequences were determined automatically with a Beckman model 890C sequencer by the method of Edman and Begg (9). The positions of the radiolabeled amino acids were determined by measuring radioactivity at each degradation step.

Pronase Digestion. Digestion by Pronase (Sigma) (10) of the completely reduced and alkylated HCD protein secreted by the cells was performed in 0.1 M NH_4HCO_3, pH 8.2, for 3 hr at 37°C at an enzyme-to-substrate ratio of 1:50 (wt/wt). The digest was applied to a Dowex AG50W-X2 (Bio-Rad) column in water to separate the peptides containing pyrrolidone carboxylic acid.

Nucleic Acid Sequence Analysis. Sequencing procedures were those of Maxam and Gilbert (11), except that a modified G+A cleavage reaction was used (12). Restriction endonuclease-digested DNA was labeled either at protruding 3′ ends by using [α-^{32}P]3′-dATP (cordycepin; New England Nuclear) in the presence of terminal deoxynucleotidyltransferase (P-L Biochemicals) or at recessed 3′ ends by using the four deoxynucleoside [α-^{32}P]triphosphates (New England Nuclear) in the presence of *Escherichia coli* DNA polymerase I (Klenow fragment; Boehringer–Mannheim). After labeling, fragments labeled at one end were produced either by redigestion of double-stranded DNA with a different restriction endonuclease or by strand separation on 5% polyacrylamide gels (11). After degradation reactions, the DNA fragments were electrophoresed on either 32 × 40 × 0.04 cm 25% polyacrylamide sequencing gels for the separation of oligonucleotides 1 to 35 bases in length or 32 × 80 × 0.04 cm 5% polyacrylamide gels for separation of oligonucleotides containing 30 or more bases.

Preparation of cDNA Clones. To prepare a fraction enriched in the OMM message, the total poly(A)-containing RNA was size-fractionated by centrifugation through 10–40% sucrose gradients. RNA molecules were selected that sedimented in the 15.5–16S region, because previous work had shown that cell-free translation of these fractions yielded *bona fide* HCD protein (5). Procedures for the preparation of recombinant plasmids from the enriched mRNA molecules were as described by Steinmetz *et al.* (13) and are summarized as follows: cDNA copies of the messages in the enriched fractions were made by extending oligo(dT) primers with avian myeloblastosis virus reverse transcriptase (obtained from J. Beard, Life Sciences, St. Petersburg, FL). After alkaline hydrolysis of the RNA, double-stranded cDNA was prepared by incubation of the single strands with DNA polymerase I and the four deoxynucleoside triphosphates. The resulting hairpin loop was cleaved with S1 nuclease. The double-stranded molecules, tailed at the 3′ ends with poly(dC) in the presence of terminal deoxynucleotidyltransferase, were inserted in the *Pst* I endonuclease site of plasmid pBR322 similarly tailed with poly(dG). *E. coli* strain MC1061 was transformed with the recombinant plasmids by using the procedure of Kushner (14). Successful transformations were identified using the high density screening method of Hanahan (15). Preliminary studies in our laboratory established conditions under which human immunoglobulin genomic sequences would hybridize with either genomic or cDNA clones containing homologous murine Ig DNA sequences (unpublished results). A murine (BALB/c) genomic clone (γ_{42}) containing a 6.7-kilobase (kb) *Eco*RI fragment which included the γ2b constant region sequences was obtained from M. Davis and S. Kim of California Institute of Technology. DNA from the clone was digested with *Eco*RI and the 6.7-kb fragment was isolated by agarose gel electrophoresis and benzoylated DEAE-cellulose chromatography. Two micrograms of the isolated 6.7-kb fragment were nick-translated to yield a ^{32}P-labeled molecule with a specific activity of 6×10^7 cpm/μg. The labeled probe was used for colony hybridization (15).

RESULTS AND DISCUSSION

Amino Acid Sequence of the OMM Translation Product and Cellular Protein. When the OMM mRNA was added to wheat germ lysate in the presence of [^{35}S]methionine, [^{3}H]leucine, or [^{3}H]phenylalanine, a radiolabeled translation product (OMM_{Tr}), specifically precipitable with anti-γ antiserum, was obtained, and the amino acid sequence of the NH_2-terminal region was determined as described in *Materials and Methods*. OMM_{Tr} contains methionine at the NH_2 terminus followed by a hydrophobic region containing phenylalanine at positions 6 and 7 and leucine at positions 4, 8, 9, 10, and 18. Such a structure corresponds in length, position, and hydrophobicity to the leader sequences found at the NH_2 termini of newly synthesized secretory proteins (16).

In contrast to the *in vitro* translated material, the cytoplasmic protein contained a blocked NH_2 terminus. To determine the NH_2-terminal residues, unlabeled HCD protein secreted by the cultured cells was isolated by specific adherence to an insoluble immunoabsorbent column of rabbit anti-human IgG. The HCD protein eluted from the column was digested with Pronase as described in *Materials and Methods*. A dipeptide was obtained that contained pyrrolidone carboxylic acid (a cyclic form of glutamine or glutamic acid) followed by valine. Human V regions of subgroups I and II commonly start with such a sequence. Unlike the cellular protein, the serum protein began with a sequence starting at the glycine, 7 residues past the normal beginning of the γ3 hinge (3). The amino acid sequences of the three biosynthetic forms of the OMM protein are shown in Fig. 1, with the normal γ3 hinge sequence aligned for comparison.

We had previously determined that the message coding for the OMM protein was smaller in size than that coding for a normal γ chain (5). We therefore interpret the data in Fig. 1 to indicate that the short cytoplasmic mRNA codes for a precursor HCD protein having a deletion and from which the leader piece is cleaved in a normal way by the intact cell to produce a molecule whose NH_2 terminus is Glx-Val. Such a sequence corresponds to the normal NH_2 terminus of the V region. The similarity in size between the cellular and serum proteins, as determined by electrophoretic mobility on Na-$DodSO_4$/polyacrylamide gels, indicated that only a small portion of normal V region could be present and that the deletion would, therefore, include most of the V and the C_H1 domain, with normal sequence resuming at the hinge. Thus, when secreted, the OMM HCD protein resembles the majority of internally deleted HCD molecules. Cleavage of the cellular protein between the 6th and 7th residues of the hinge (Leu-Gly) would produce the NH_2 terminus of the protein found in the patient's serum. It was previously determined that the hinge region of intact γ3 H chains is particularly sensitive to the action of common proteases (17). The replacement of the normal V and C_H1 domains by a short V-region sequence immediately preceding the hinge may expose potential proteolytic cleavage sites in that region that would not otherwise be available for enzymatic attack.

Preparation of cDNA Clones from the γ3 HCD mRNA. pBR322 plasmids containing DNA inserts complementary to OMM cellular messages were prepared and used to transform *E. coli* strain MC1061 as described in *Materials and Methods*.

	Leader Sequence	Short V Region	Hinge
In Vitro Translation Product (OMM_{Tr})	Met--Leu-PhePheLeuLeuLeu-------Leu..	....	....
Cell Protein		<GluVal....................	....
Serum Protein			GlyAspThrThr...
Normal γ3 Hinge			GluLeuLysThrProLeuGlyAspThrThr...

FIG. 1. The NH_2-terminal sequences of the three different biosynthetic forms of the OMM HCD protein (-, undetermined amino acids; <Glu, pyrrolidone carboxylic acid). The sequence at the beginning of the normal γ3 hinge region (17) is shown for comparison with the NH_2-terminal sequence of the serum protein.

Four of approximately 6,000 recombinant clones clearly contained DNA sequences complementary to the mouse γ2b DNA used as a hybridization probe. Plasmid DNA from the four clones was bound to nitrocellulose filters and used to select mRNA molecules from total poly(A)-containing RNA extracted from the OMM cells (18). The duplexes were melted and the mRNA was translated in cell-free translation systems extracted from both wheat germ and reticulocytes (6, 7). Immunologic precipitation of the labeled translation products followed by $NaDodSO_4$/polyacrylamide gel electrophoresis revealed that the mRNA hybridizing with the plasmid DNA coded for the γ3 HCD protein.

DNA from each of the clones (designated POMMA, -B, -C, and -D) was digested with *Pst* I. The sizes of the cDNA inserts, as determined by agarose gel electrophoresis, were 1.8, 1.0, 0.7, and 0.5 kb, respectively. The clones with the two largest inserts, A and B, were chosen for further study.

Nucleotide Sequence of OMM cDNA. The amino acid sequence deduced from the nucleotide sequence of the OMM cDNA (Fig. 2) confirms and extends the results obtained by

```
                                  |Leader
......CCUGGACCUCCUGUGCAAGAAC|AUG AAA CAN CUG UGG UUC UUC CUU CUC CUG GUG GCA GCU CCC AGA UGG GUC CUG UCC
                            |Met Lys His Leu Trp Phe Phe Leu Leu Leu Val Ala Ala Pro Arg Trp Val Leu Ser
                            |-19                                                                 -3  -2  -1

|V Region                                                   |Hinge
|CAG GUG CAC CUG CAG GAG UCG GGC CCA GGA CUG GGG AAG CCU CCA|GAG CUC AAA ACC CCA CUU GGU GAC ACA ACU CAC ACA UGC CCA CGG
|Gln Val His Leu Gln Glu Ser Gly Pro Gly Leu Gly Lys Pro Pro|Glu Leu Lys Thr Pro Leu Gly Asp Thr Thr His Thr Cys Pro Arg
| 1   2   3                                              15|216 217 218 219         220 221 222 223 224 225 226 227 228

UGC CCA GAG CCC AAA UCU UGU GAC ACA CCU CCC CCG UGC CCA CGG UGC CCA GAG CCC AAA UCU UGU GAC ACA CCU CCC CCA UGC CCA CGG
Cys Pro Glu Pro Lys Ser Cys Asp Thr Pro Pro Pro Cys Pro Arg Cys Pro Glu Pro Lys Ser Cys Asp Thr Pro Pro Pro Cys Pro Arg
229 230

                                                            |CH2
UGC CCA GAG CCC AAA UCU UGU GAC ACA CCU CCC CCG UGC CCN NNG UGC CCA|GCA CCU GAA CUC UUG GGA GGA CCG UCA GUC UUC CUC UUC
Cys Pro Glu Pro Lys Ser Cys Asp Thr Pro Pro Pro Cys Pro     Cys Pro|Ala Pro Glu Leu Leu Gly Gly Pro Ser Val Phe Leu Phe
                                                                   |231                                                243

CCC CCA AAA CCC AAG GAU ACC CUU AUG AUU UCC CGG ACC CCU GAG GUC ACG UGC GUG GUG GUG GAC GUG AGC CAC GAA GAC CCN NNN GUC
Pro Pro Lys Pro Lys Asp Thr Leu Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val Val Asp Val Ser His Glu Asp Pro     Val
244                                                                                                                   273

CAG UUC AAG UGG UAC GUG GAC GGC GUG GAG GUG CAU AAU GCC AAG ACA AAG CUG CGG GAG GAG CAG UAC AAC AGC ACG UUC CGU GUG GUC
Gln Phe Lys Trp Tyr Val Asp Gly Val Glu Val His Asn Ala Lys Thr Lys Leu Arg Glu Glu Gln Tyr Asn Ser Thr Phe Arg Val Val
274                                                                                                                   303

AGC GUC CUC ACC GUC CUG CAC CAG GAC UGG CUG AAC GGC AAG GAG UAC AAG UGC AAG GUC UCC AAC AAA GCC CUC CCA GCC CCC AUC GAG
Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile Glu
304                                                                                                                   333

                           |CH3
AAA ACC AUC UCC AAA GCC AAA GGA|CAG CCC <-- 36 N --> GAG GAG AUG ACC AAG AAC CAA GUC AGC CUG ACC UGC CUG GUC AAA GGC UUC
Lys Thr Ile Ser Lys Ala Lys Gly|Gln Pro              Glu Glu Met Thr Lys Asn Gln Val Ser Leu Thr Cys Leu Val Lys Gly Phe
334                         341|342                  356                                                               372

UAC CCC AGC GAC AUC GCC GUG GAG UGG GAG AGC AAU GGG CAG CCG GAG AAC AAC UAC AAC ACC ACG CCU CCC AUG CUG GAC UCC GAC GGC
Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Asn Thr Thr Pro Pro Met Leu Asp Ser Asp Gly
373                                                                                                                   402

UCC UUC UUC CUC UAC AGC AAG CUC ACC GUG GAC AAG AGC AGG UGG CAG CAG GGG AAC AUC UUC UCA UGC UCC GUG AUG CAU GAG GCU CUG
Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser Arg Trp Gln Gln Gly Asn Ile Phe Ser Cys Ser Val Met His Glu Ala Leu
403                                                                                                                   432

                                          End of CH3|
CAC AAC CGC UAC ACG CAG AAG AGC CUC UCC CUG UCU CCG GGU|AAA UGA GUGCCAUGGCCGGCAAGCCCCCGCUCCCCGGGCUCUCGGGGUCGCGCGAGGAUGC
His Asn Arg Tyr Thr Gln Lys Ser Leu Ser Leu Ser Pro Gly|Lys End
433                                                 446|

UUGGCACGUACCCCGUGUACAUACUUCCCAGGCACCCAGCAUGGAAAUAAAGCACCCAGCGCUGCCCUGGAAAAAAAAAAAA - GC linker
```

FIG. 2. Nucleotide sequence of the OMM HCD mRNA (N, nucleotide not determined). The deduced amino acid sequence is shown beneath the nucleotide sequence. The numbers correspond to amino acids of the γ1 protein Eu (19). The nucleotide sequence was derived primarily from the POMMB cDNA clone, which contained the sequence starting with the last 9 nucleotides of the leader through the 3′ untranslated region. The entire sequence of the leader was determined from the larger POMMA clone, which also contained the POMMB sequence, as determined by restriction enzyme mapping and partial nucleotide sequencing.

direct amino acid sequence analysis of the OMM protein. The leader encompasses 19 amino acids and ends with a serine residue, as do many of the published mouse H- and L-chain leaders. The highly hydrophobic character is apparent from the complete sequence. The 15 amino acids immediately following the leader are clearly homologous to the first 15 residues of human H-chain V regions. A comparison with the subgroup II protein NEWM (20) shows identity at 11 of the first 15 residues. The differences at the remaining four positions can be accounted for by single-base changes.

The direct amino acid sequence data did not define the specific point where normal structure resumes after the deletion. Because the serum protein appeared to be the product of limited NH_2-terminal proteolysis, the NH_2-terminal glycine was not likely to be the true COOH-terminal boundary of the deletion. It is clear from the OMM cDNA nucleotide sequence that the coding information for the entire normal hinge region is present in the cytoplasmic message, hence the deletion extends from amino acid 16 in the V region through residue 215 [Eu numbering (19)], which marks the end of the C_H1 domain.

The amino acid sequence deduced from the OMM cDNA agrees with published γ3 protein sequences (21, 22) with the exception of seven residues. Positions 283 and 386 involve Glu/Gln interchanges which probably reflect uncertainties in the protein sequence. At position 291, a residue not involved in subclass determination, a single base change leads to the replacement of proline with leucine. The remaining four variant amino acids, 296, 339, 384 and 436, are identical to residues found in the normal human γ1 sequence.

The heavy chain of a recently described intact γ3 myeloma protein, GOE (23), also contains γ1 residues at three of the above four positions, the fourth residue being undetermined. These data suggest that the GOE and OMM proteins may be allelic forms of γ3 or may represent another γ subclass, the product of a separate γ3-like gene. The similarity between OMM and the H chain of the GOE protein, an apparently normal Ig molecule, suggests that most of the discrepancies between the published γ3 protein sequence and that deduced from the OMM cDNA can be attributed to a naturally occurring sequence polymorphism.

The normal γ3 hinge region consists of the quadruplication of a basic unit 15 amino acids in length. The first of the four units includes an addition of two amino acids and is homologous but not identical to the others, whereas the latter three units are exact repetitions. Because it is thought likely that the γ3 hinge evolved by quadruplication of the ancestral 15-residue unit (17), we expected to see a greatly conserved codon usage when comparing a particular position in one unit with the same position in the other three. This proved to be correct. Except for one silent base change in a proline codon, the nucleotide sequences of the last three repeat units are identical. In the first unit, with the exception of one position, the amino acid sequence divergence reflects only single base changes in the codons.

Immediately before the stop codon in the cDNA sequence, there is information for an additional lysine that is not present in any of the published γ3 amino acid sequences. Because the cDNA is a direct copy of the mature cytoplasmic message, the extra codon is presumably translated. The nucleotide sequences of the human γ4 (24) and mouse γ1 (25), γ2a (26, 27), and γ2b (28, 29) genes also encode information for a COOH-terminal lysine not present in the mature proteins. Apparently, the lysine is removed from Ig H chains of both species postsynthetically.

The sequence of the OMM cDNA 3′ untranslated region is shown in Fig. 3, compared to the corresponding region in the human γ4 gene (24). The two sequences differ in 5 out of 133 nucleotides, with no additions or deletions being necessary to maintain homology. The sequence A-A-T-A-A-A, thought to be a signal for poly(A) addition (30), is identically positioned in both sequences. In the POMMB cDNA clone, poly(A) was added 19 nucleotides 3′ to the A-A-T-A-A-A region. The POMMA clone contained a different poly(A) addition site eight nucleotides 3′ from the POMMB site. Because both cDNA clones have the same extensive internal deletion in the coding region, it is probable that the mRNAs that served as templates for the cDNA copies were transcripts of the same gene. Published nucleotide sequences of the cDNAs coding for mouse Ig H chains indicate that in the γ2a (31) and γ2b (32) subclasses poly(A) is added at a position homologous to that seen in the POMMA clone. The site for poly(A) addition in the cDNA coding for an intact γ1 chain (33) is located one base 3′ to the POMMB site, although the deleted γ1 H chain IF-2 (34) has poly(A) added at the POMMA site. These data suggest that the finding of two different functional poly(A) addition sites in the OMM RNA transcripts is not related to the structural abnormality in the coding region. Multiple poly(A) addition sites have also been described for the murine dihydrofolate reductase (35), α-amylase (36), and Ig μ-chain (37) genes, although in the latter case the two sites are associated with the membrane and secretory forms of the message. It thus appears that poly(A) addition is not fixed to one site per gene.

Molecular Events Underlying Synthesis of the Protein Having the Deletion. Because immunochemical and amino acid sequence analyses of the HCD proteins revealed that the deletions encompassed whole domains with resumption of normal structure at the beginning of the hinge, it was hypothesized that the hinge represents a separate genetic element (38). Structural studies of germ-line and active mouse (25–29, 39, 40) and germ-line human (24) γ-chain genes revealed that the DNA coding for the V region and the domains of the constant region were separated by noncoding intervening sequences (IVS). Hence, the hinge region is coded by a separate gene segment. The entire H-chain unit is transcribed into a single precursor RNA from which the noncoding sequences are removed, and the coding regions are spliced together to yield the cytoplasmic mRNA. In the HCD cell line, a generalized abnormality of the splicing mechanism is very unlikely, because the mRNAs coding for

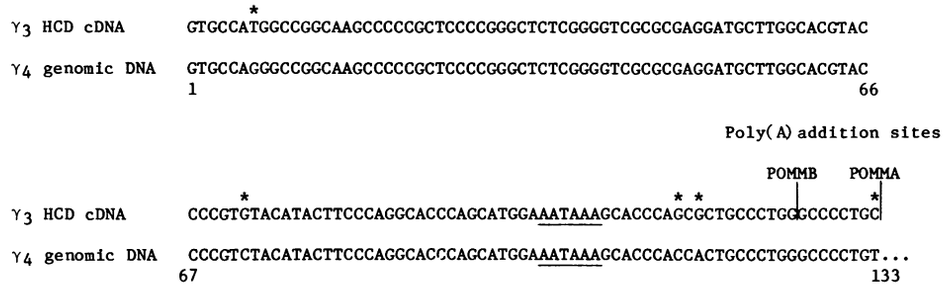

FIG. 3. The 3′ untranslated region of the OMM HCD cDNA compared to that of human γ4 genomic DNA (24) (* indicates difference between the two sequences).

other cell proteins are correctly spliced. A mutation or deletion resulting in an altered or missing splice site could force a normal splicing system to select another available position and thus join two coding regions that would not normally be contiguous. It has been demonstrated that this is the case in both the murine Ig H-chain variant IF-2 (41) and a clone of the murine plasmacytoma MPC-11 that synthesizes an L-chain fragment (42–44).

Examination of available human H-chain V-region amino acid sequences (45) reveals no case in which there is a proline at residue 15. The residues commonly found at position 15 are: for V_H1, Gly (8 out of 8); for V_H2, Thr (3 out of 5), Lys (1 out of 5), Ser (1 out of 5); for V_H3, Gly (48 out of 48). Hence, it is possible that the proline in position 15 of the OMM protein is abnormal and may reflect an event that created the codon CCA followed by a sequence capable of serving as the donor site of a splice junction. The creation of a new donor site would allow the joining of the short V-gene fragment to the hinge exon. To yield the amino acid sequence observed, at least two normal donor sites, those at the ends of the joining (J) and the C_H1 exons, had to be bypassed. The most likely explanation to account for these observations is a substantial DNA deletion. The definition of the actual event awaits the determination of the nucleotide sequence of genomic clones isolated from the HCD cells.

The present data now make it appear likely that NH_2-terminally deleted HCD serum proteins are synthesized initially as internally deleted molecules and have biosynthetic histories similar to the history of the OMM serum protein. A possible scenario for the generation of the aberrant proteins could begin with a deletion or mutation at the DNA level that results in altered or missing splice sites in the primary nuclear RNA transcript. This would cause the splicing system to align coding regions that are not normally contiguous. Thus, a short message would be formed in which a large part of the normal coding region is deleted. Nevertheless, the molecule would be entirely capable of directing translation of an internally deleted protein that would be processed normally to yield an internally deleted secretory product. In some cases, the abnormal NH_2 terminus may become the substrate for proteolytic enzymes in the serum, resulting in proteins with apparent NH_2-terminal deletions.

J.N.B. gratefully acknowledges the social and intellectual hospitality of L.H.'s colleagues in the Divisions of Biology and Chemistry at the California Institute of Technology during his sabbatical stay, without which the construction and cloning of the recombinant plasmids would not have been possible. The authors also acknowledge the technical expertise of Donald Hauser and the secretarial proficiency of Randi Klein. The studies were supported by Veterans Administration research funds and National Institutes of Health Grant AM01431. D.B. was a Fogarty International Fellow, M.S. was a Fellow of the Deutsche Forschungsgemeinschaft, and J.N.B. was a Scholar in Cancer Research of the American Cancer Society.

1. Franklin, E. C. & Frangione, B. (1975) in *Contemporary Topics in Molecular Immunology*, eds. Inman, F. P. & Mandy, W. (Plenum, New York), pp. 89–126.
2. Buxbaum, J. N. (1976) *La Ricerca* **6**, 301–317.
3. Adlersberg, J., Frangione, B. & Franklin, E. C. (1975) *Proc. Natl. Acad. Sci. USA* **72**, 723–727.
4. Adlersberg, J. B., Grann, V., Zucker-Franklin, D., Frangione, B. & Franklin, E. C. (1978) *Blood* **51**, 85–96.
5. Alexander, A., Barritault, D. & Buxbaum, J. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 4774–4778.
6. Roberts, B. E. & Paterson, B. M. (1973) *Proc. Natl. Acad. Sci. USA* **70**, 2330–2334.
7. Pelham, H. R. B. & Jackson, R. J. (1976) *Eur. J. Biochem.* **67**, 247–256.
8. Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J. & Rutter, W. J. (1979) *Biochemistry* **18**, 5294–5299.
9. Edman, P. & Begg, G. (1967) *Eur. J. Biochem.* **1**, 80–91.
10. Press, E. M., Piggot, P. J. & Porter, R. R. (1966) *Biochem. J.* **99**, 356–366.
11. Maxam, A. M. & Gilbert, W. (1980) *Methods Enzymol.* **65**, 499–560.
12. Gray, C. P., Sommer, R., Polke, C., Beck, E. & Schaller, H. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 50–53.
13. Steinmetz, M., Frelinger, J. G., Fisher, D., Hunkapiller, T., Pereira, D., Weissman, S. M., Uehara, H., Nathenson, S. & Hood, L. (1981) *Cell* **24**, 125–134.
14. Kushner, S. (1978) in *Genetic Engineering*, eds. Boyer, H. W. & Nicosia, S. (Elsevier/North-Holland Biomedical, Amsterdam), pp. 17–23.
15. Hanahan, D. & Meselson, M. (1980) *Gene* **10**, 63–67.
16. Sabatini, D. D., Kreibich, G., Morimoto, T. & Adesnik, M. (1982) *J. Cell Biol.* **92**, 1–22.
17. Michaelsen, T. E., Frangione, B. & Franklin, E. C. (1977) *J. Biol. Chem.* **252**, 883–889.
18. Ricciardi, R. P., Miller, J. S. & Roberts, B. E. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 4927–4931.
19. Edelman, G. M., Cunningham, B. A., Gall, W. E., Gottlieb, P. D., Rutishauser, U. & Waxdal, M. J. (1969) *Proc. Natl. Acad. Sci. USA* **63**, 78–85.
20. Poljak, R. J., Nakashima, Y., Chen, B. L. & Konigsberg, W. (1977) *Biochemistry* **16**, 3412–3420.
21. Frangione, B., Rossenwasser, E., Prelli, F. & Franklin, E. C. (1980) *Biochemistry* **19**, 4304–4308.
22. Wolfenstein-Todel, C., Frangione, B., Prelli, F. & Franklin, E. C. (1976) *Biochem. Biophys. Res. Commun.* **71**, 907–914.
23. Recht, B., Frangione, B., Franklin, E. C. & van Loghem, E. (1981) *J. Immunol.* **127**, 917–923.
24. Ellison, J., Buxbaum, J. & Hood, L. (1981) *DNA* **1**, 11–18.
25. Honjo, T., Obata, M., Yamawaki-Kataoka, Y., Kataoka, T., Kawakami, T., Takahashi, N. & Mano, Y. (1979) *Cell* **18**, 559–568.
26. Ollo, R., Auffray, C., Morchamps, C. & Rougeon, F. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 2442–2446.
27. Yamawaki-Kataoka, Y., Mityata, T. & Honjo, T. (1981) *Nucleic Acids Res.* **9**, 1365–1381.
28. Tucker, P. W., Marcu, K. B., Newell, N., Richards, J. & Blattner, F. R. (1979) *Science* **206**, 1303–1306.
29. Yamawaki-Kataoka, Y., Kataoka, T., Takahashi, N., Obata, M. & Honjo, T. (1980) *Nature (London)* **283**, 786–789.
30. Proudfoot, N. & Brownlee, G. G. (1976) *Nature (London)* **263**, 211–214.
31. Sikovav, J.-L., Auffray, C. & Rougeon, F. (1980) *Nucleic Acids Res.* **8**, 3143–3155.
32. Tucker, P. W., Marcu, K. B., Slightman, J. L. & Blattner, F. R. (1979) *Science* **206**, 1299–1303.
33. Obata, M., Yamawaki-Kataoka, Y., Takahashi, N., Kataoka, T., Shimizu, A., Mano, Y., Seidman, J. C., Peterlin, B. M., Leder, P. & Honjo, T. (1980) *Gene* **9**, 87–97.
34. Dunnick, W., Rabbitts, T. H. & Milstein, C. (1980) *Nucleic Acids Res.* **8**, 1475–1485.
35. Setzer, D. R., McGrogan, M., Nunberg, J. H. & Schimke, R. T. (1980) *Cell* **22**, 361–370.
36. Tosi, M., Young, R., Hagenbüchle, D. & Shibler, U. (1981) *Nucleic Acids Res.* **9**, 2313–2323.
37. Early, P. W., Rogers, J., Davis, M., Calame, K., Bond, M., Wall, R. & Hood, L. (1980) *Cell* **22**, 313–319.
38. Frangione, B. & Wolfenstein-Todel, C. (1972) *Proc. Natl. Acad. Sci. USA* **69**, 3673–3676.
39. Maki, R., Traunecker, A., Sakano, H., Roeder, W. & Tonegawa, S. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 2138–2142.
40. Sakano, H., Rogers, J. H., Huppi, K., Brack, C., Traunecker, A., Maki, R., Wall, R. & Tonegawa, S. (1979) *Nature (London)* **277**, 627–633.
41. Dunnick, W., Rabbitts, T. H. & Milstein, C. (1980) *Nature (London)* **286**, 669–675.
42. Schnell, H., Steinmetz, M. & Zachau, H. G. (1980) *Nature (London)* **286**, 170–173.
43. Choi, E., Kuehl, M. & Wall, R. (1980) *Nature (London)* **286**, 776–779.
44. Seidman, J. G. & Leder, P. (1980) *Nature (London)* **286**, 779–783.
45. Kabat, E. A., Wu, T. T. & Bilofsky, H. (1979) *Sequences of Immunoglobulin Chains* (U.S. Dept. of Health, Education & Welfare, Washington, DC).