Immunobiology
Zeitschrift für Immunitätsforschung

Vol. 182

SEMPER BONIS ARTIBUS

Gustav Fischer Verlag · Stuttgart · New York · 1990/1991

ISSN Immunobiology · Zeitschrift für Immunitätsforschung · 0171-2985

Printed by Druckerei Ungeheuer+Ulmer KG GmbH+Co, Ludwigsburg
Printed in Germany

Contents Volume 182 · 1990/1991

Original Papers

Letters to the Editor

Special Report

Author's Index

Subject Index

Immunobiol., vol. 182, pp. 307–322 (1991)

[1] Institut für Hygiene, Innsbruck, Austria, [2] Institut für Immunologie, München, Germany, and [3] Red Cross Blood Transfusion Center, Helsinki, Finland

Human Complement Factor H: Molecular Cloning and cDNA Expression Reveals Variability in the Factor H-Related mRNA Species of 1.4 kb

WILHELM SCHWAEBLE[1*], ELISABETH FEIFEL[1], CORNELIA ESTALLER[2], ANTONIA BARBIERI[1], MARTIN MÖLGG[1], VESA KOISTINEN[3], ELISABETH H. WEISS[2], and MANFRED P. DIERICH[1]

Received December 21, 1990 · Accepted in Revised Form February 18, 1991

Abstract

Previously, we have shown that three different mRNA species of 4.3 kb, 1.8 kb and 1.4 kb, related to human complement factor H, are constitutively expressed in the human liver. Probing with our cDNA clone H-46 which represents 920 bp of the 3' end of the 4.3 kb mRNA of factor H on human liver RNA, we always detected the 4.3 kb and the additional, abundantly expressed mRNA species of 1.4 kb in length, indicating that the 1.4 kb transcript is highly homologous to the 3' end of the classical factor H mRNA of 4.3 kb. Using H-46 as a probe, several cDNA clones were isolated from a liver cDNA library and sequenced. The open reading frame of the novel mRNA species encodes a peptide consisting of five internal short consensus repeat motifs (SCR), identifying the translational product to be a member of the SCR family. Sequence comparison with cDNA clones derived from liver RNA of a different donor provided evidence for variability in the factor H related proteins encoded by the 1.4 kb mRNA species.

Interestingly, this variability was found to be restricted to the three carboxyterminal SCR domains. Expression data indicate that our variant is not recognized by the monoclonal antibody 3D11.

Introduction

Factor H is a regulatory protein of the complement system which controls the alternative pathway of complement activation by competing for the interaction of C3b with factor B and dislodging the Bb fragment of factor B from the active alternative pathway convertase (C3bBb) (1–3).

Factor H also serves as a cofactor for factor I in the conversion of C3b to inactivated C3b, iC3b (4). The binding site for C3b has been located on a

---

* Recipient of a fellowship from «Deutscher Akademischer Austauschdienst» Sonderprogramm Gentechnologie.

** This study was supported by a grant from the Austrian FWF (P6920), is a part of E. FEIFEL's, C. ESTALLER's and M. MÖLGG's Ph.D. thesis and of A. BARBIERI's M.D. thesis.

Mr 38,000 aminoterminal fragment obtained by limited tryptic digestion of the major Mr 150,000 factor H serum protein (5).

In addition, this Mr 38,000 fragment shows a cofactor activity in the factor I-mediated conversion of C3b to iC3b, while the remainder of the molecule is inactive in this respect (6).

For human factor H three different mRNA species are abundantly expressed in liver: 4.3 kb, 1.8 kb and 1.4 kb (7). The 4.3 mRNA codes for the major factor H serum protein of Mr 150,000 – termed H gp150 – and has been completely sequenced (8, 9). The sequence analysis of the 1.8 kb mRNA showed that this species is largely identical to the 5' portion of the 4.3 kb mRNA (8–11). A partial analysis of the factor H gene strongly suggests that the 1.8 kb and the 4.3 kb mRNA arise by alternative splicing from a single structural gene (12).

Factor H has been shown to be highly homologous to other regulatory proteins of the complement system such as the C4 binding protein (C4bp), the decay accelerating factor (DAF), the membrane cofactor protein (MCP) and the complement receptors type 1 (CR1) and type 2 (CR2). Characteristically, the amino acid sequence of these proteins has a structure of numerous internal domains, termed «short consensus repeats» (SCR). Each SCR covers a stretch of approximately 60 amino acids with a framework of four completely conserved cysteine residues along with highly conserved residues for glycine, tryptophan, tyrosine and proline (13–15).

The translational product of the 4.3 kb mRNA – Hgp150 – is composed of twenty such SCRs, while the 1.8 kb mRNA encodes the first seven repeating units of the aminoterminal portion of Hgp150, plus four unique amino acids at the COOH-terminal sequence. The genes encoding factor H and the other above-mentioned complement regulatory proteins have been localized to chromosome 1 forming a tightly linked gene cluster named the RCA (regulation of complement activation) cluster (16).

Recently, we identified a novel truncated factor H protein as the likely translational product of the 1.8 kb mRNA in fresh human plasma by Western blot analysis with monoclonal and polyclonal antibodies directed against the aminoterminal portion of H (7).

After purification by combined affinity chromatography and HPLC, we were able to show that this novel truncated factor H serum protein also has cofactor activity for the cleavage of C3b (17). The identity of this novel, truncated factor H protein with the translational product of the 1.8 kb mRNA was confirmed by eucaryotic expression of a full length cDNA clone of the 1.8 kb factor H mRNA species (9).

We have previously shown that in human liver RNA, a third factor H related mRNA species of 1.4 kb strongly hybridizes to our cDNA clone H-46 (H-46 represents 920 basepairs of the 3' end of the 4.3 kb mRNA). This mRNA species is only detectable in liver tissue, while the 4.3 kb and the 1.8 kb mRNA species of factor H are also expressed in fibroblasts, monocytes, endothelial cells and lung tissue (18).

In this report, we present a novel nucleotide sequence (H-69) for this factor H-related mRNA species and compare the sequence to other cDNA transcripts which we isolated in parallel from a different cDNA library (19). Procaryotic expression indicates that the significant variation observed for the factor H-related 1.4 kb mRNA leads (in the case of H-69) to a loss of the epitope recognized by monoclonal antibody 3D11 which was shown to react with the classical factor H molecule of Mr 150,000 and the two polypeptides encoded by the other variants of this factor H-related 1.4 kb mRNA species.

Materials and Methods

cDNA cloning and sequencing of H-69 and H-69/2

The cDNA clone H-69 was isolated from a commercially available human liver cDNA library in λ-gt10 (Clontech, Palo Alto, CA, USA) which was screened with our factor H-specific cDNA clone H-46. cDNA clone H-69/2 was isolated from a human liver primer extension cDNA library in which the oligo dT primer was replaced by a 21 bp long oligonucleotide, termed oligo 1.4/1,(3' GCC CCG ACC AGG TGG GGA GGG 5') derived from H-69 (see Fig. 2) to initiate the synthesis of the first strand. For the cDNA synthesis, poly (A^+) RNA was selected from total RNA preparations of human liver tissue using the «mRNA Purification Kit» (Pharmacia, Freiburg, Germany). First and second strand synthesis was carried out using the «You-Prime cDNA Synthesis Kit» (Pharmacia). Double-stranded cDNA was subcloned in λ-gt10 using the «cDNA cloning system λ-gt10» (Amersham, Aylesbury, UK). The library was screened with a second 18 bp long oligonucleotide, termed oligo 1.4/2, (3' CTA CAC AAA CAG GAC TTC 5') derived from H-69 5 to oligo 1.4/1. From five hybridizing plaques, H-69/2 showing the largest cDNA insertion was analyzed further. The cDNA insertions contained in H-69 and H-69/2 were amplified in a polymerase chain reaction (PCR) using «λ-gt10 Screening Amplimer sets» (Clontech). They were then digested with either EcoR I, or EcoR I and Rsa I, or EcoR I and Hind III and subcloned into M13 for sequencing by the Sanger dideoxy chain termination method (20). The sequencing strategy chosen for H-69 and H-69/2, is shown in Figure 1. Sequencing of H-69 and H-69/2 was performed with T7 polymerase and the reagent kit (Pharmacia).

Northern blot analysis

From human liver tissue, RNA was prepared according to the method of Chirgwin et al. (21). Twenty μg of total RNA was separated on a formaldehyde-containing agarose gel and blotted to Hybond N nylon filters (Amersham). Agarose gel electrophoresis, RNA transfer and hybridization of the blots were done by standard techniques (20). The DNA probes were labeled with P^{32}-dATP using the random priming method (22).

The probes utilized in these studies are H-69/2 (coding for the novel 5' sequence of the factor H related 1.4 kb mRNA), a 292 bp Hind III subfragment of H-69 (representing the 3' end of H-69), and a 718 bp long Alu I subfragment of the factor H specific cDNA clone R2a (8) which codes for SCR domains 12–16.

Hybridization was performed at 65 °C in the absence of formamide. The final washing step was done in 0.1 × SSC for 1 h at 65 °C.

Procaryotic and eucaryotic expression and Western blot analysis

For procaryotic expression, the cDNA contained in H-69 was digested with EcoR I, yielding a 1060 bp-long subfragment of H-69 which encodes, with the exception of 13 aminoterminal amino acids, the mature translational product of the 1.4 kb mRNA species.

This fragment was subcloned into the EcoR I site of the procaryotic expression vector pAX (23). The expression of clone pAX-H-69, containing the EcoR I fragment of H-69 in the correct reading frame and orientation, was induced by adding 1 ml of a 100 mM isopropyl-β-D-thiogalactoside solution (IPTG) to a 100 ml bulk culture for at least 3 h before harvesting.

The expression of the fusion protein encoded by H-46 (which was isolated from a human liver library cloned in the procaryotic expression vector pEX 2 (11)) was induced according to the method by STANLEY and LUZIO (24) by shifting a bulk culture of H-46 to 42 °C for 90 min. The bacteria were pelleted and lysed in 1/10 of the original volume of 50 mM TRIS/HCl pH 6.8 containing 5 % SDS, 1 % mercaptoethanol, 10 % glycerol.

The eucaryotic expression of the cDNA clone pFH1.4a was achieved by COS-cell transfection. Therefore, pFH1.4a and the vector CDM8 as control DNA were transfected into COS cells by electroporation using the «Gene PulserTM apparatus» (Bio-Rad, Munich, Germany) following the recommendations of the supplier. Briefly, 0.8 ml of COS cell suspension (1×10^7 cells/ml) were mixed with 60 µg DNA. Electroporation was performed in 272 mM sucrose, 7 mM Na_2HPO_4 pH 7.4 and 1 mM $MgCl_2$ with 380 V and 250 F in the cold. The cells were then diluted with DMEM/25 mM HEPES medium containing 10 % FCS and cultured in 9 cm petri dishes. After 24 h, the medium was changed to DMEM containing 1 × Nutridoma HU (Boehringer Mannheim). Supernatants were collected 72 h after the pulse, concentrated 20-fold, and assayed for the expression the translational product of 1.4 kb mRNA.

For Western blot analysis, 80 µl of concentrated supernatant of COS cell transfectants, 2 µl of serum and 20 µl of the bacterial lysates were electrophoresed on a 7.5 % SDS-polyacrylamide gel and blotted to nitrocellulose according to the method of TOWBIN et al. (25).

Individual lanes were stained with mAb 3D11 and a polyclonal goat anti-factor-H-antibody using peroxidase-conjugated anti-mouse IgG and peroxidase-conjugated anti-goat IgG, respectively, according to standard procedures.

Results

cDNA cloning and sequencing of H-69 and H-69/2

To isolate cDNA clones representing the unknown factor H related mRNA species of 1.4 kb, H-46 was used to screen a human liver cDNA library in λ-gt10. From 60,000 plaques, twenty hybridizing clones were isolated. From these, cDNA insertions were amplified in a polymerase chain reaction (PCR) using oligonucleotides corresponding to the polylinker region of λ-gt10. Amplified inserts were digested with Eco RI, subcloned in the bacteriophage M13 and sequenced from both ends (see Fig. 1). From those, nineteen were identified to be incomplete transcripts of the known 4.3 kb mRNA, while one – designated H-69 – was shown to have a 3' portion almost identical to that of the 4.3 kb mRNA, preceded by a unique 5' portion with only a slight degree of homology to the sequence of factor H (see Fig. 2).

Since neither a 5' untranslated region nor a translational start codon was contained at the 5' end of H-69, we used an oligonucleotide (3' GCC CCG ACC AGG TGG GGA GGG 5') corresponding to the 5' portion of the H-69 sequence for a primer extension cDNA synthesis in order to obtain a cDNA covering the full length of the 1.4 kb mRNA species. One primer extension clone, termed H-69/2, was sequenced extending the 5' end of H-69 for a length of 166 bp (see Fig. 1).

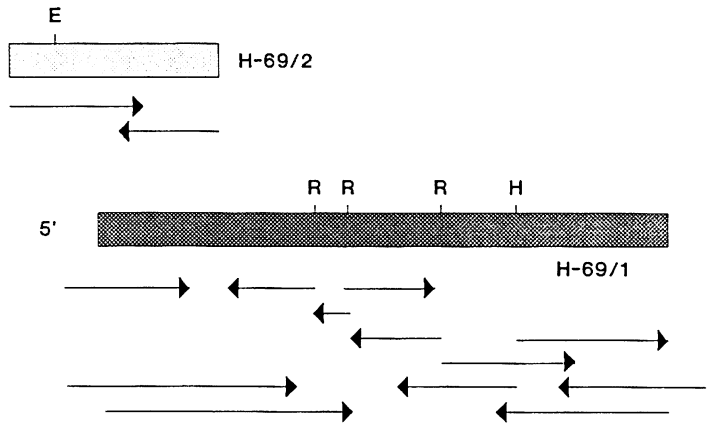

Figure 1. Subcloning and sequencing strategy used for the sequence analysis of the cDNA clones H-69/1 and H-69/2. The arrows mark the extent and direction of the DNA sequence determination below the restriction map with the following enzymes: E = EcoR I; R = Rsa I; H = Hind III. H-69/1 and H-69/2 overlap for a stretch of 270 bp.

The complete nucleotide sequence and the derived amino acid sequence, encoded by the cDNA clones H-69 and H-69/2 are shown in Figure 2. The combined nucleotide sequence of 1173 bp revealed an open reading frame of 993 bp starting with ATG at position 22 and ending with a stop codon (TAG) at position 1012. The coding sequence of the 1.4 kb mRNA covered by H-69 and H-69/2 is flanked by a 5' and a 3' untranslated region of 21 and 162 nucleotides, respectively. The 3' untranslated region includes a poly-adenylation signal (ATTAAA) located 40 nucleotides downstream from the translational termination codon. With exception of four nucleotide exchanges and one insertion (see Fig. 2), the sequence of the 3' untranslated region of H-69 is identical to the 3' untranslated region of the 4.3 kb mRNA species of factor H, represented by H-46.

Moreover, the sequence homology extends into the protein coding region which is nearly identical for a stretch of 557 bp encoding the three SCR domains SCR 18, 19, 20 of classical factor H (see Fig. 2). Regarding the criteria given by VON HEIJNE (26), the peptide sequence, encoded by the open reading frame of the 1.4 kb mRNA species, indicates an 18 amino acid-long leader peptide prior to the 312 amino acids of the mature protein.

Including the leader sequence, the molecular weight for the primary translational product was calculated to be 37,663 and 35,750 for the mature protein, respectively. As shown in Figure 3, the mature protein is composed of five approximately 60-amino acid-long short consensus repeats (SCR) which are typically observed in complement regulatory proteins (13–16). The three carboxyterminal SCRs of the protein encoded by H-69 and H-69/2 show five amino acid exchanges in comparison to the carboxyterminal SCR's of the major Mr 150,000 factor H protein, namely SCR-18-20. In

comparison, only a low degree of homology was observed between the SCRs 6 and 7 of the classical factor H protein of Mr 150,000 and the two novel SCR domains 1 and 2 of the aminoterminal portion of the protein encoded by the 1.4 kb mRNA.

Interestingly, the derived amino acid sequence of H-69 contains only one potential glycosylation site (NIS) in position 108, while the two other variants of the 1.4 kb mRNA (pFH1.4a and b) and H-46 (representing the 4.3 kb mRNA species) encode an additional, potential glycosylation site (NWT) in position 176.

```
                                  M  W  L  L  V  S  V  I  L  I  S  R  I
H-69      ATAAGATTGGAACTACCAAGCATGTGGCTCCTGGTCAGTGTAATTCTAATCTCACGGATA   60
pFH1.4a   ...........................................................   96

                          1                                  10
           S  S  V  G  G  E  A  T  F  C  D  F  P  K  I  N  H  G  I  L
H-69      TCCTCTGTTGGGGGAGAAGCAACATTTTGTGATTTTCCAAAAATAAACCATGGAATTCTA  120
pFH1.4a   ...........................................................  156

                    20                                 30
           Y  D  E  E  K  Y  K  P  F  S  Q  V  P  T  G  E  V  F  Y  Y
H-69      TATGATGAAGAAAAATATAAGCCATTTTCCCAGGTTCCTACAGGGGAAGTTTTCTATTAC  180
pFH1.4a   ...........................................................  216

                    40                                 50
           S  C  E  Y  N  F  V  S  P  S  K  S  F  W  T  R  I  T  C  T
H-69      TCCTGTGAATATAATTTTGTGTCTCCTTCAAAATCATTTTGGACTCGCATAACATGCACA  240
pFH1.4a   ...........................................................  276

                    60                                 70
           E  E  G  W  S  P  T  P  K  C  L  R  L  C  F  F  P  F  V  E
H-69      GAAGAAGGATGGTCACCAACACCAAAGTGTCTCAGACTGTGTTTCTTTCCTTTTGTGGAA  300
pFH1.4a   ...........................................................  336

                    80                                 90
           N  G  H  S  E  S  S  G  Q  T  H  L  E  G  D  T  V  Q  I  I
H-69      AATGGTCATTCTGAATCTTCAGGACAAACACATCTGGAAGGTGATACTGTGCAAATTATT  360
pFH1.4a   ...........................................................  396
oligo                ................
1.4/2

                   100                                110
           C  N  T  G  Y  R  L  Q  N  N  E  N  N  I  S  C  V  E  R  G
H-69      TGCAACACAGGATACAGACTTCAAAACAATGAGAACAACATTTCATGTGTAGAACGGGGC  420
pFH1.4a   ...........................................................  456
oligo                                                        GCCCCG
1.4/1

                   120                                130
           W  S  T  P  P  K  C  R  S  T  D  T  S  C  V  N  P  P  T  V
H-69      TGGTCCACCCCTCCCAAATGCAGGTCCACTGACACTTCCTGTGTGAATCCGCCCACAGTA  480
pFH1.4a   ...........................................................  516
H-46                                .....C.....................          216
oligo     ACCAGGTGGGGAGGG
1.4/1

                   140                                150
           Q  N  A  H  I  L  S  R  Q  M  S  K  Y  P  S  G  E  R  V  R
H-69      CAAAATGCTCATATACTGTCGAGACAGATGAGTAAATATCCATCTGGTGAGAGAGTACGT  540
pFH1.4a   .........T.....G.....................................          576
H-46      .........T.....G.....................................          276
                   Y     V
                   160                                170
           Y  Q  C  R  S  P  Y  E  M  F  G  D  E  E  V  M  C  L  N  G
H-69      TATCAATGTAGGAGCCCTTATGAAATGTTTGGGGATGAAGAAGTGATGTGTTTAAATGGA  600
pFH1.4a   ...........................................................  636
H-46      ...........................................................  336
```

Figure 2. Nucleotide and deduced amino acid sequence of cDNA clones H-69 and H-69/2 in comparison with another variant of the human factor related 1.4 kb mRNA species (pFH1.4a) and the 4.3 kb mRNA species represented by our cDNA clone H-46.

Northern blot analysis with H-69 and H-69/2 and fragments specific for the related 1.4 and the 4.3 kb mRNA species of factor H

To examine whether H-69 is a faithful transcript of the 1.4 mRNA species, we established 3' and 5' specific subfragments of H-69 to probe human liver RNA.

A 3' specific subfragment was isolated by using an internal unique Hind III-restriction site in position 882 resulting in a 292 bp 3' fragment.

```
                    180                                 190
             N  W  Q  E  P  P  Q  C  K  D  S  Q  G  K  C  G  P  P  P  P
H-69         AACTGGCAGGAACCACCTCAATGCAAAGATTCTCAGGGAAAATGTGGGCCCCCTCCACCT  660
pFH1.4a      ......AC.......................AC.......................  696
H-46         ......AC.......................ACA......................  396
                   T                           T

                    200                                 210
             I  D  N  G  D  I  T  S  F  P  L  S  V  Y  A  P  A  S  S  V
H-69         ATTGACAATGGGGACATTACTTCATTCCCGTTGTCAGTATATGCTCCAGCTTCATCAGTT  720
pFH1.4a      ............................................................  756
H-46         ............................................................  456

                    220                                 230
             E  Y  Q  C  Q  N  L  Y  Q  L  E  G  N  K  R  I  T  C  R  N
H-69         GAGTACCAATGCCAGAACTTGTATCAACTTGAGGGTAACAAGCGAATAACATGTAGAAAT  780
pFH1.4a      ............................................................  816
H-46         ............................................................  516

                    240                                 250
             G  Q  W  L  E  P  P  K  C  L  H  P  C  V  I  S  R  E  I  M
H-69         GGACAATGGCTAGAACCACCAAAATGCTTACATCCGTGTGTAATATCCCGAGAAATTATG  840
pFH1.4a      .........TC.................................................  876
H-46         .........TC.................................................  576
                   S
                    260                                 270
             E  N  Y  N  I  A  L  R  W  T  A  K  Q  K  L  Y  L  R  T  G
H-69         GAAAATTATAACATAGCATTAAGGTGGACAGCCAAACAGAAGCTTTATTTGAGAACAGGT  900
pFH1.4a      ............................................................  936
H-46         ............................................C...............  636

                    280                                 290
             E  S  A  E  F  V  C  K  R  G  Y  R  L  S  S  R  S  H  T  L
H-69         GAATCAGCTGAATTTGTGTGTAAACGGGGATATCGACTTTCATCACGTTCTCACACATTG  960
pFH1.4a      ..........................T.................................  996
H-46         .......T..........................T.........................  696
                   V
                    300                                 310
             R  T  T  C  W  D  G  K  L  E  Y  P  T  C  A  K  R  *
H-69         CGAACAACATGTTGGGATGGGAAACTGGAGTATCCAACTTGTGCAAAAAGATAGAATCAA 1020
pFH1.4a      ............................................................ 1056
H-46         ............................................................  756

H-69         TCATAAAATGCACACCTTTATTCAGAACTTTAGTATTAAATCAGTTCTTAATTTCATTTT 1080
pFH1.4a      ...........................................................A..... 1116
H-46         .......G...................................C..........  816

H-69         TAA-GTATTGTTTTACTCCTTTTTATTCATACGTAAAATTTTGGATTAATTGGTGAAAATG 1140
pFH1.4a      ...-.......................................T.......... 1176
H-46         .T.T.......................................T..........  876

H-69         TAATTATAAGCTGAGACCGGTGGCTCTCTTCTT                     1173
pFH1.4a      ............................                          1209
H-46         ............................                           936
```

The sequences corresponding to the oligonucleotides oligo 1.4/1 (used for the primer extension cDNA synthesis of H-69/2) and oligo 1.4/2 (used for screening the primer extension cDNA library) are marked.

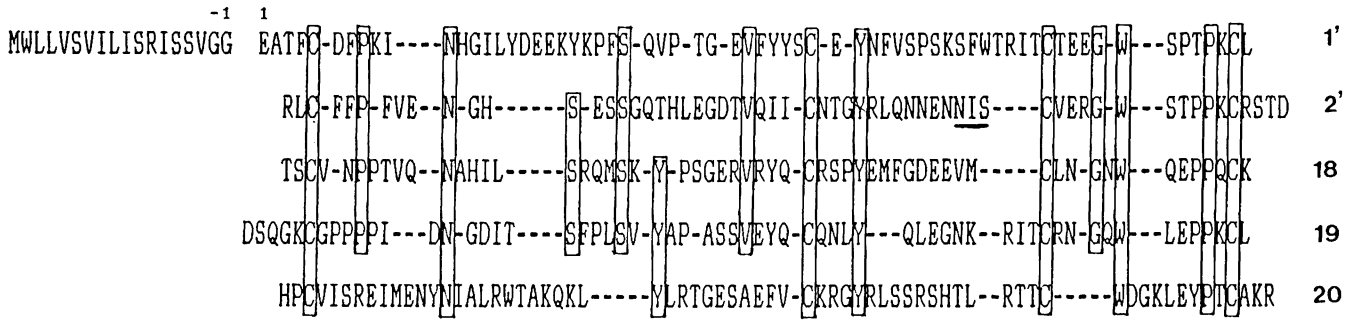

Figure 3. Assignment of the protein sequence derived from H-69 and H-69/2 to five homology subunits of about 60 amino acids each. Identical amino acids present in several of the five homologous regions are boxed. Gaps denoted by «–» have been introduced to maximize homology and to allow space for the regions with short inserted sequences. The potential glycosylation site (NIS) in position 108 is underlined.

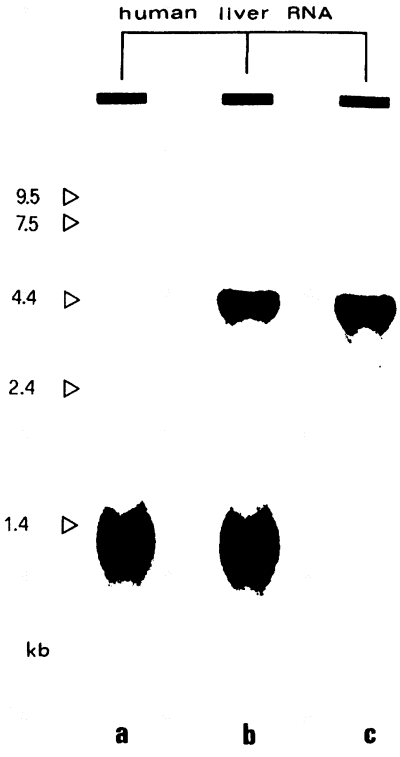

Figure 4. Northern blots of human liver RNA hybridized with different cDNA fragments. Twenty μg of total RNA each was separated on an agarose gel. Lane a was hybridized with H-69/2 coding for the novel 5' sequence of the 1.4 kb mRNA species. Lane b was hybridized with a 292 bp long Hind III restriction fragment representing the 3' coding and the 3' untranslated region of H-69. Lane c was hybridized with a 718 bp Alu I fragment representing the coding region for SCR 12–16 of the 4.3 kb mRNA of factor H.

This subfragment completely represents the 3' untranslated region in addition to 130 bp of the SCR 20 coding region of H-69. This 3' fragment of H-69 hybridizes with both, the 4.3 and the 1.4 kb mRNA species (see Fig. 4, lane b). This result confirms the cDNA sequencing data showing that both mRNA species share the same 3' region.

Probing primer extension clone H-69/2 (coding for the novel 5' sequence of the 1.4 kb derived cDNA) on total liver RNA, we exclusively detected the 1.4 kb mRNA species (see Fig. 4, lane a). This result shows that the 1.4 kb mRNA must have given rise to the unique 5' sequence linked to a segment which is shared by both, the 4.3 kb factor H mRNA and related 1.4 kb mRNA species.

Further confirmation for the direct linkage of SCR 1 and 2 of the 1.4 kb mRNA to the conserved sequence of SCR 18, 19 and 20 was obtained by probing with a subfragment of the 4.3 kb cDNA (718 bp Alu I fragment)

representing the coding region for SCR domains 12 to 16 (see Fig. 6). This probe exclusively hybridizes with the 4.3 kb and not with the 1.4 kb mRNA species (see Fig. 4, lane c).

Thus, the Northern blot results confirm that H-69 and H-69/2 are faithful reverse transcripts of the 1.4 kb mRNA species sharing a nearly identical 3' portion with the 4.3 kb long factor H mRNA, followed by a novel 5' portion of 456 nucleotides.

Procaryotic expression and Western blot analysis of the recombinant proteins encoded by H-69 and H-46

For procaryotic expression, H-69/1 was subcloned into the EcoR I site of the procaryotic expression vector pAX4C.

Clone pAX-H-69, which was shown to contain the insertion in the correct reading frame and orientation, was used for expression of the β-galactosidase/H-69 fusion protein. In parallel, the expression of the fusion protein encoded by H-46 (cloned in pEX2 (11)) was induced according the method by STANLEY and LUZIO (24).

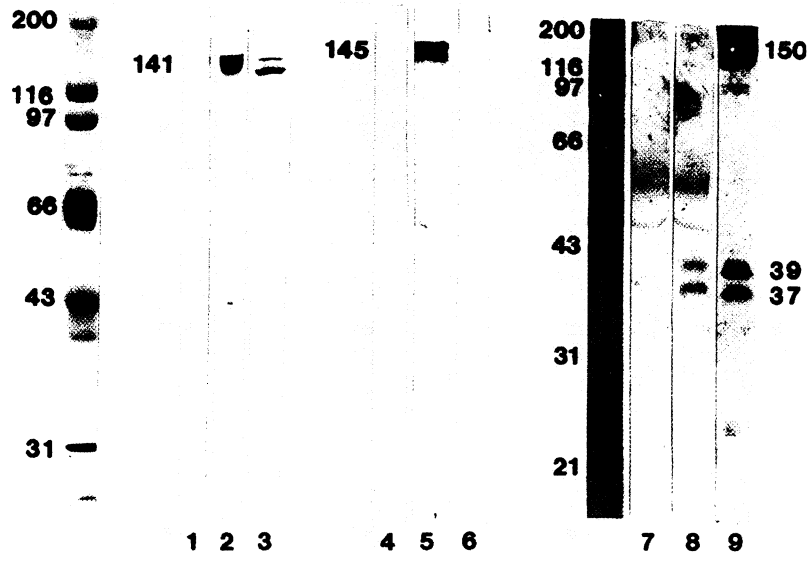

Figure 5. Western blot analysis of the translational products encoded by different variants of the 1.4 kb mRNA species, the 3' coding region of the 4.3 kb mRNA of factor H, and of the translational products represented in human plasma. In lanes 1 and 4, bacterial lysate and the galactosidase encoded by pEX2 (lane 1) and pAXC4 (lane 4) were exposed to a polyclonal anti-factor H-antibody as a negative control. The fusion protein encoded by H-46 was stained with the polyclonal (lane 2) and the monoclonal anti-factor-H-antibody mAb3D11 (lane 3) while the fusion protein encoded by pAX-H-69 was only stained with the polyclonal antibody (lane 5), but not with mAb3D11 (lane 6). In lane 8, concentrated supernatant from COS cells transfected with pFH1.4a was stained with mAb3D11, while supernatant from COS cells transfected with the vector CDM8 alone remained blank (lane 7). In lane 9, 2 µl of human plasma were stained with mAb3D11.

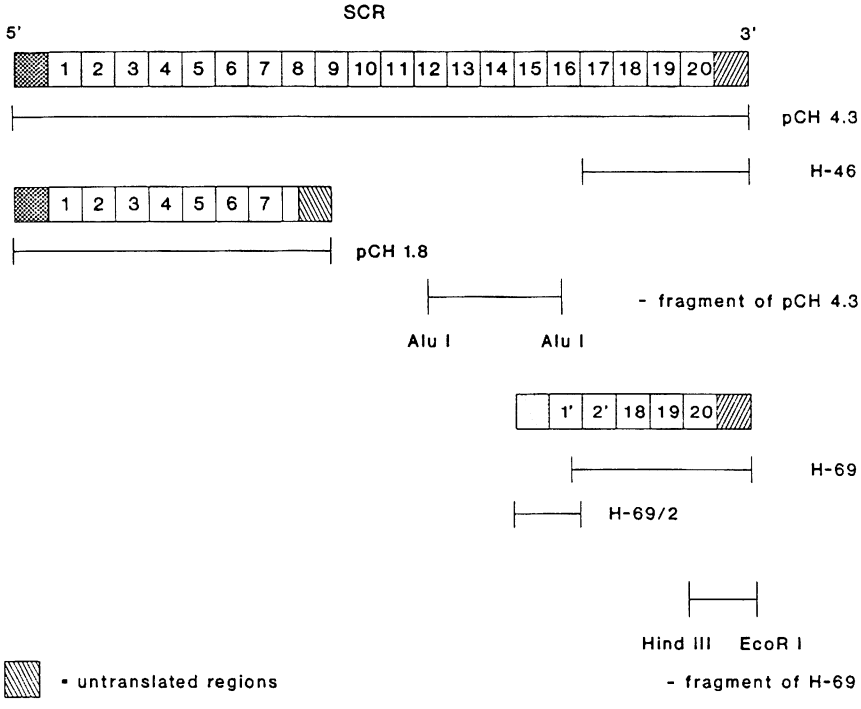

Figure 6. Diagram of the three mRNA species related to human factor H. The specific untranslated regions and the SCR repeat units encoded by the different mRNA species are boxed. The cDNA clones (pCH4.3 and H-46 for the 4.3 kb mRNA, pCH1.8 for the 1.8 kb mRNA, H-69 and H-69/2 for the 1.4 kb mRNA) and subfragments representing the different mRNA species are marked as lines.

After induction, the bacteria were lysed and the total lysate was separated on a 7.5 % SDS-polyacrylamide gel. To exclude any possible crossreactivity of the antibodies used with β-galactosidase or bacterial polypeptides in the Western blot analysis, proteins were also prepared from a pAX C4 clone and a pEX 2 clone without any cDNA insertion. As a positive control, human plasma was collected from four different donors in presence of 0.5 mM PMSF, 5 mM EDTA and 50 mM 6-aminohexanoic acid (EACA) to prevent any possible protein degradation. As shown in Figure 5, the polyclonal antibody stained both fusion proteins encoded by H-46 and pAX-H-69 while the extracts of the vector recombinant bacteria remained blank.

In contrast, the monoclonal antibody 3D11 only stained the fusion protein generated by H-46, but not the fusion protein of pAX-H-69 (Fig. 5, lanes 3 and 6). Again, the negative controls remained blank whereas, mAb 3D11 stained the classical Mr 150,000 factor H protein, the two polypeptides of Mr 37,000 and 39,000 in human plasma, and the eucaryotic expression products of pFH1.4a, which is a full length cDNA clone of the

1.4 kb factor H mRNA, but differs in five amino acid residues from the H-69 sequence (see Fig. 5).

Discussion

The Northern blot analysis and the cDNA sequencing have confirmed that the overlapping cDNA clones H-69 and H-69/2 represent faithful reverse transcripts of a novel factor H-related mRNA species of 1.4 kb, which is abundantly expressed in human liver. The open reading frame of H-69 codes for a peptide with one potential N-glycosylation site (NIS) in position 108. For the nonglycosylated, mature protein, the molecular weight was calculated to be 36,000. The primary structure of this protein is comprised of five SCR domains. The three carboxyterminal repeats are nearly identical with the carboxyterminal sequence of the major factor H protein of Mr 150,000. The sequence coding for the two novel aminoterminal SCR domains was shown to be unique for the 1.4 kb mRNA species in human liver by probing with primer extension cDNA clone H-69/2.

In parallel to this work, we describe full length cDNA clones (pFH1.4 a–f) derived from the 1.4 kb mRNA species which code for two distinct, presumably allelic transcripts (19).

The clone described here was isolated from a different cDNA library and differs from the pFH1.4 clones and the 4.3 kb factor H transcript in the three carboxyterminal SCR domains and the 3' untranslated region (Table 1). Thus, this cDNA represents another variant of the factor H-like protein encoded by the 1.4 kb mRNA species.

Interestingly, no differences were observed in the 5' sequence of the 1.4 kb mRNA coding for the two unique N-terminal domains, and no variation was found in the sequence of the different cDNA isolates representing the 4.3 kb mRNA (H-46, this paper; R2a (8); pFH4.3 (9)). The derived amino acid sequence of H-69 differs in three positions to the peptide sequence encoded by pFH1.4b, and in five amino acid residues to the peptide encoded by pFH1.4a (see Table 1). Although the pFH sequences share two potential aminoglycosylation sites in position 108 and 176 with classical factor H, the variant encoded by H-69 has only the recognition sequence

Table 1. Amino acid exchanges caused by nucleotide substitutions within three different variants of the factor H related 1.4 kb mRNA species and the 3' portion of the 4.3 kb mRNA of factor H (H-46)

	Pos. 139	Pos. 141	Pos. 178	Pos. 187	Pos. 239	Pos. 278
pFH1.4a	TAT=Tyr	GTG=Val	ACG=Thr	ACG=Thr	TCA=Ser	GCT=Ala
pFH1.4b	CAT=His	CTG=Leu	ACA=Thr	ACG=Thr	TCA=Ser	GCT=Ala
H-69	CAT=His	CTG=Leu	CAG=Gln	CAG=Gln	CTA=Leu	GCT=Ala
H-46	TAT=Tyr	GTG=Val	ACG=Thr	ACA=Thr	TCA=Ser	GTT=Val

(NWQ instead of NWT) in position 176. It is likely that the function of the three carboxyterminal repeats in the polypeptides encoded by the 1.4 kb mRNA species differs from the one carried by the carboxyterminus of classical factor H.

From a mouse liver cDNA library, four different classes of factor H related transcripts have recently been isolated and sequenced (27). Like the factor H related 1.4 kb mRNA species in man (18), these transcripts were found to be exclusively expressed in RNA preparations from liver tissue while the 4.4 kb factor H mRNA species of the mouse is also expressed in kidney, spleen, thymus and L cells (28).

With respect to the functional importance of the protein encoded by the human 1.4 kb mRNA species, it was of interest to learn that one of these factor H-related transcripts of the mouse, represented by the cDNA clone 13G1, shows remarkable similarities to the human 1.4 kb mRNA species:

(i) Similar to the reported human 1.4 kb mRNA species, the mRNA which has given rise to 13G1, encodes a peptide of similar size consisting of five SCR motifs. The two aminoterminal domains were shown to be unique while the three following SCRs have stringent homology to the carboxyterminal domains of the mouse classical factor H. In contrast to the human factor H related transcript, the protein encoded by 13G1 contains two SCR repeats homologous to SCR 19 of the classical factor H followed by one SCR 20 homologous domain.

(ii) As observed for the 1.4 kb mRNA in man, cDNA 13G1 shares a nearly identical 3' untranslated region with the 4.4 kb factor H mRNA. However, both use a different 5' untranslated region and leader sequence. Nucleotide and amino acid sequence comparison of the human and the mouse mRNA species shows that the polypeptide sequence is more conserved than the nucleotide sequence.

The polypeptide encoded by H-69 was investigated by the expression of a fusion protein (pAX-H-69) in *E. coli* in comparison with the homologous fragment of similar size derived from the carboxyterminus of the classical factor H cDNA (H-46 representing SCR 17–20). Whereas a polyclonal anti-factor-H-antibody detected both fusion proteins in a Western blot analysis, the monoclonal antibody 3D11 only stained the fusion protein of H-46, but not the fusion protein encoded by pAX-H-69. When mAb 3D11 was used to look for a translational product of the 1.4 kb mRNA species in human plasma, classical factor H of Mr 150,000 and in addition, two bands with a Mr of 39,000 and 37,000 were stained (see Fig. 5). This doublet was also seen with polyclonal antibodies directed against H and was detected regardless of whether the samples were run under reducing conditions or not. Moreover, the eucaryotic expression of both pFH1.4 variants (pFH1.4a and b) resulted in the synthesis of the two polypeptides of Mr 39,000 and 37,000 which were stained with mAb 3D11 (19). We therefore assume that the two smaller polypeptides detected in human plasma by mAb 3D11 and polyclonal antibodies are encoded by the factor H related 1.4 kb mRNA species and that the epitope recognized by mAb 3D11 must

be localized on SCR 18–20. The mAb 3D11 failed to bind to the fusion protein encoded by pAX-H-69 whereas a similar fragment of classical factor H expressed with the same fusion partner and the eucaryotic expression products of pFH1.4a/b were stained, indicates that this reagent does not recognize the variant represented by the H-69 cDNA.

The missing N-glycosylation site can not be responsible since procaryotic expression products do not have this modification. Thus, the three amino acid residues at position 178, 188 and 239 with unique substitutions in the H-69-encoded polypeptide, are part of the epitope recognized by 3D11. Screening human sera with polyclonal factor H reagents and the mAb 3D11 showed variations in the expression in the small factor H related polypeptides (unpublished results).

Since the donor of the liver used for cDNA construction is not known, we could not directly test the plasma proteins. In conclusion, our results indicate variations in the factor H-related polypeptides encoded by the 1.4 kb mRNA species detected by cDNA sequencing and by staining with the mAb 3D11. The posttranslational modification or processing event leading to the two polypeptides encoded by one 1.4 kb mRNA commonly detected in human sera, is not clear. It still remains to be clarified whether this novel factor H related serum protein has any of the functions of classical factor H known so far.

We are most grateful to Dr. R. B. Sim and Dr. A. J. Day, MRC Immunochemistry Unit, Oxford, UK, for helpful discussion.

References

1. Whaley, K., and S. Ruddy. 1976. Modulation of C3b hemolytic activity by a plasma protein distinct from C3b inactivator. Science **193:** 1011.
2. Whaley, K., and S. Ruddy. 1976. Modulation of the alternative complement pathway by β1Hglobulin. J. Exp. Med. **144:** 1147.
3. Conrad, D. H., J. R. Carlo, and S. Ruddy. 1978. Interaction of β1H globulin with cell bound C3b: Quantitative analysis of binding and influence of alternative pathway components on binding. J. Exp. Med. **147:** 1792.
4. Pangburn, M. K., R. D. Schreiber, and H. J. Müller-Eberhard. 1977. Human complement C3b inactivator: isolation, characterisation, and demonstration of an absolute requirement for the serum protein 1H for cleavage of C3b and C4b in solution. J. Exp. Med. **146:** 257.
5. Alsenz, J., J. D. Lambris, T. F. Schulz, and M. P. Dierich. 1984. Localisation of the complement component C3b binding site and the cofactor activity for factor I in the 38 kDa tryptic fragment of factor H. Biochem. J. **224:** 389.
6. Alsenz, J., T. F. Schulz, J. D. Lambris, R. B. Sim, and M. P. Dierich. 1985. Structural and functional studies on complement factor H with different enzymes and monoclonal anti-factor H antibodies. Biochem. J. **232:** 841.
7. Schwaeble, W., J. Zwirner, T. F. Schulz, R. P. Linke, M. P. Dierich, and E. H. Weiss. 1987. Human complement factor H: Expression of an additional truncated gene product of 43 kDa in human liver. Eur. J. Immunol. **17:** 1485.
8. Ripoche, J., A. J. Day, T. J. R. Harris, and R. B. Sim. 1988. The complete aminoacid sequence of human complement factor H. Biochem. J. **249:** 593.

9. ESTALLER, C., W. SCHWAEBLE, M. P. DIERICH, and E. H. WEISS. 1991. Human complement factor H: Two factor H serum proteins are derived from alternatively spliced transcripts. Eur. J. Immunol., in press.
10. KIRSTENSEN, T., R. A. WETSEL, and B. F. TACK. 1986. Structural analysis of human complement factor H: homology with C4b binding protein, -glycoprotein I and the Ba fragment of factor B. J. Immunol. **136:** 3407.
11. SCHULZ, T. F., W. SCHWAEBLE, K. K. STANLEY, E. H. WEISS, and M. P. DIERICH. 1986. Human complement factor H: isolation of cDNA clones and partial cDNA sequence of the 38 kDa tryptic fragment containing the binding site for C3b. Eur. J. Immunol. **16:** 1351.
12. MCALEER, M. A., J. RIPOCHE, O. DOMINGUEZ, W. SCHWAEBLE, and R. B. SIM. 1989. Partial gene structure for human factor H includes an alternative splice region. Abstract 24-17, International Congress of Immunology, p. 140. Gustav Fischer Verlag, Stuttgart.
13. HOLERS, V. M., J. L. COLE, D. M. LUBLIN, T. SEYA, and J. P. ATKINSON. 1985. Human C3b-regulatory proteins. A new multi-gene family. Immunol. Today **6:** 188.
14. REID, K. B. M., D. R. BENTLEY, R. D. CAMPBELL, L. P. CHENG, R. B. SIM, T. KIRSTENSEN, and B. F. TACK. 1986. Complement system proteins which interact with C3b or C4b. Immunol. Today **7:** 230.
15. REID, K. B. M., and A. J. DAY. 1989. Structure-function relationship of the complement components. Immunol. Today **10:** 177.
16. KLICKSTEIN, L. B., W. W. WONG, J. A. SMITH, J. H. WEIS, J. G. WILSON, and D. T. FEARON. 1987. Demonstration of long homologous repeating domains that are composed of the short consensus repeats characteristic of C3/C4 binding proteins. J. Exp. Med. **165:** 1095.
17. MISASI, R., H. P. HUEMER, W. SCHWAEBLE, E. SÖLDER, C. LARCHER, and M. P. DIERICH. 1989. Human complement factor H: an additional gene product of 43 kDa isolated from human plasma shows cofactor activity for the cleavage of the third component of complement. Eur. J. Immunol. **19:** 1765.
18. SCHWAEBLE, W., H. SCHWAIGER, R. A. BROOIMANS, A, BARBIERI, J. MÖST, M. TIEFENTHALER, M. HIRSCH-KAUFFMANN, D. F. LAPPIN, M. R. DAHA, K. WHALEY, and M. P. DIERICH. 1991. Human complement factor H: tissue specificity in the expression of three different mRNA species. Eur. J. Biochem., in press.
19. ESTALLER, C., W. SCHWAEBLE, V. KOISTINEN, M. P. DIERICH, and E. H. WEISS. 1991. Cloning of the 1.4 kb mRNA species of human factor H reveals a novel member of the SCR-family, related to the carboxyterminus of the classical 150 kDa molecule. J. Immunol., in press.
20. MANIATIS, T., E. F. FRISCH, and J. SAMBROOK. 1982. Molecular cloning. A laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor Press, New York.
21. CHIRGWIN, J. M., A. E. PRZYBYLA, R. J. MACDONALD, and W. J. RUTTER. 1979. Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. Biochemistry **18:** 5294.
22. FEINBERG, A. P., and B. VOGELSTEIN. 1983. A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Anal. Biochem. **137:** 266.
23. NAGAI, K., and H. C. THOGERSEN. 1984. Generation of β-globin by sequence-specific proteolysis of a hybrid protein produced in *Escherichia coli*. Nature **309:** 810.
24. STANLEY, K. K. and J. P. LUZIO. 1984. Construction of a new family of high efficiency bacterial expression vectors: identification of cDNA clones coding for human liver proteins. EMBO J. **3:** 1429.
25. TOWBIN, H., T. STAEHLIN, and J. GORDON. 1979. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedures and some applications. Proc. Natl. Acad. Sci. USA **76:** 4350.
26. VON HEIJNE, G. 1987. Sequence analysis in molecular biology. Academic Press, London.
27. VIK, D. P., J. B. KEENEY, P. MUNOZ-CANOVES, D. D. CHAPLIN, and B. F. TACK. 1990. Identification and sequence analysis of four complement factor H related transcripts in mouse liver. J. Biol. Chem. **265:** 3193.

28. MUNOZ-CANOVES, P., B. F. TACK, and D. P. VIK. 1989. Dexamethasone and retinoic acid increase the level of murine H mRNA. Complement Inflamm. **6:** 372.

Dr. WILHELM SCHWAEBLE, Department of Internal Medicine, Verfügungsgebäude für Forschung, Obere Zahlbacher Straße 63, 6500 Mainz, Germany