Short paper

Cornelia Estaller[○+], Wilhelm Schwaeble[△], Manfred Dierich[△] and Elisabeth H. Weiss[○]

Institut für Immunologie[○], München and Institut für Hygiene[△], Innsbruck

Human complement factor H: two factor H proteins are derived from alternatively spliced transcripts*

The human complement factor H is an important component in the control of the alternative pathway of complement activation. We have previously shown that at least three factor H homologous mRNA species of 4.3 kb, 1.8 kb and 1.4 kb in length are constitutively expressed in human liver. In addition, several factor H-related proteins have been detected in human sera using antibodies directed against the classical human factor H glycoprotein of 150 kDa. The structure of the additional polypeptides has not been shown so far. Circumstantial evidence suggests that the 1.8-kb mRNA might encode the 43-kDa factor H-like polypeptide. Here we report the isolation, characterization and eukaryotic expression of the first full-length cDNA representing the major 4.3-kb mRNA and the 1.8-kb mRNA of human factor H. We show that the 4.3-kb transcript encodes the 150-kDa-factor H glycoprotein and the 1.8-kb mRNA the 43-kDa factor H polypeptide. The identity of the two cDNA in a region of 1400 nucleotides suggests that the two factor H-related transcripts are derived from one gene by a process of alternative splicing.

1 Introduction

Factor H, a regulatory protein of the complement system, controls the alternative pathway of complement activation by competing for the interaction of C3b with factor B and C5 and by dislodging the Bb fragment of factor B from the active alternative pathway convertase (C3bBb), thus accelerating the decay of this enzyme [1–3].

Factor H also functions as a cofactor for factor I cleavage of C3b to inactivated C3b (iC3b) [4]. All these functions have been located on a 38-kDa aminoterminal fragment obtained by limited tryptic digestion of the major 150-kDa factor H protein derived from serum [5]. No biological activities have yet been assigned to the remaining 120-kDa portion of factor H [6, 7].

For human factor H, three different mRNA species of 4.3 kb, 1.8 kb and 1.4 kb are constitutively expressed in liver [8]. Recently shorter factor H-related polypeptides have been detected in human serum using polyclonal or monoclonal anti-factor H reagents. It is not clear whether these proteins are derived from the 150-kDa factor H glycoprotein by post-translational processing or whether they are encoded by the smaller factor H-related mRNA species present in liver [8, 9]. cDNA clones have been isolated and sequenced which represent the 1.8-kb mRNA species [10–13]. To date none of these cDNA encode functional 1.8-kb transcripts containing the correct 5′ sequence. For this reason it was not possible to demonstrate which, if any, of the short factor H-like proteins detected in serum are derived from the 1.8-kb transcript. We have previously identified and purified a novel factor H protein of 43 kDa from human plasma which is the most likely translational product for the 1.8-kb factor H mRNA *in vivo* [8, 14].

The amino acid sequence of the classical factor H (150 kDa) has been compiled from protein sequence data and overlapping cDNA sequences [13, 15]. Thus, it has been found that the translation products of the 4.3-kb and 1.8-kb factor H mRNA species share the amino terminal sequence, and it has been proposed that both transcripts are obtained from one factor H gene by a process of alternative splicing [11, 13]. Clear evidence is lacking, as no cDNA with the 5′ sequence unambiguously derived from the 4.3-kb mRNA has been isolated so far.

In order to study the complexity of factor H-like proteins and transcripts, we constructed a liver cDNA library in the eukaryotic expression vector CDM8 [16]. Full-length cDNA clones were isolated (pFH4.3 and pFH1.8) both for the 4.3-kb and the 1.8-kb mRNA species, respectively. By transient expression in COS cells, it can be shown that pFH4.3 codes for the 150-kDa factor H serum protein and pFH1.8 for the 43-kDa factor H-like polypeptide. DNA sequence data obtained from the 5′ and 3′ termini of the cDNA clones demonstrate that both transcripts are derived from one factor H gene.

2 Materials and methods

2.1 Construction and screening of a human cDNA library in the eukaryotic expression vector CDM8

RNA was prepared from human liver tissue by extraction with guanidinium thiocyanate and centrifugation through a cesium chloride gradient [17]. Poly(A)$^+$RNA was purified

[I 8636]

* This study was supported by a grant from the Genzentrum München and from the Austrian FWF (P6920).

+ This work is in partial fulfilment of C. Estaller's PhD thesis.

Correspondence: Elisabeth H. Weiss, Institut für Immunologie, Goethestr. 31, D-8000 München 2, FRG

0014-2980/91/0303-0799$3.50+.25/0

by oligo(dT) cellulose chromatography. Approximately 15 μg poly(A)$^+$RNA was employed for cDNA synthesis using the kit supplied by BRL, Karlsruhe, FRG. Size-selected cDNA fragments (> 1 kb) were ligated with Bst XI linker and cloned into purified Bst XI-linearized vector CDM8, following exactly the protocol provided by Dr. B. Seed, Boston, MA [16, 18]. About 300 000 colonies were differentially screened with the following probes: H19 cDNA insert, hybridizing to the 4.3-kb and 1.8-kb mRNA, the 3′ Bgl I/Pst I fragment of H19 detecting only the 1.8-kb transcript and H46 cDNA insert hybridizing to the 4.3-kb and 1.4-kb transcripts [8, 11]. More than 100 positive colonies were obtained with each probe. Only clones which might represent the full-length cDNA transcripts of the 4.3-kb and 1.8-kb mRNA were further characterized by restriction enzyme and DNA sequence analysis.

2.2 Restriction enzyme and DNA sequence analysis

The complete cDNA inserts could be excised with the restriction enzymes Sac I and Pst I present in the linker region of CDM8. DNA sequence analysis was performed with T7 polymerase and the reagent kit supplied by Pharmacia, Freiburg, FRG. Plasmid DNA was directly sequenced with two oligonucleotides, ocdm8-1 (CTGGCTAACTAGAGAAC) located 5′ and ocdm8-2 (GATCCTCTAGAGTCGC) located 3′ of the Bst XI cloning site.

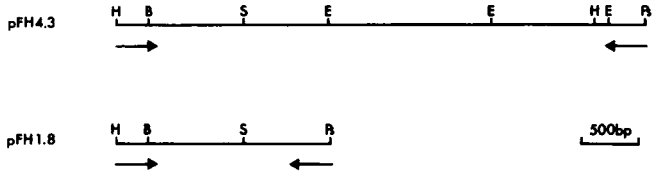

Figure 1. Full-length factor H cDNA clones in the vector CDM8. The restriction maps of the full-length cDNA clones pFH4.3 (factor H_{gp150}) and pFH1.8 (factor H_{p43}) is shown for the following restriction enzymes: B, Bgl II, E, Eco RV; H, Hind III; P, Pst I; S, Sma I. Short horizontal arrows show the direction of DNA sequence determination with oligonucleotides localized in the CDM8 linker region.

2.3 Northern blot analysis and hybridization

Approximately 1 μg of liver poly(A)$^+$RNA was separated on a formaldehyde agarose gel, blotted to nylon membrane (Hybond N, Amersham/Buchler, Braunschweig, FRG) following standard protocols [19] and hybridized with the 4.3-kb Sac I/Pst I fragment of pFH4.3, the 1.8-kb Hind III/Pst I insert of pFH1.8, and the 250-bp Eco RV/Pst I fragment of pFH4.3. The probes were labeled with α[^{32}P]dATP according to the random priming procedure [20]. Hybridization at a concentration of 5×10^6 cpm labeled fragment/ml hybridization solution and washing of colony filters and Northern blots was carried out by the method of Church and Gilbert [21].

2.4 COS cell transfection

The cDNA clones pFH4.3 and pFH1.8, and the vector CDM8 as control DNA were transfected into COS cells by electroporation using the Gene Pulser™ apparatus (Bio-Rad, Munich, FRG) 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 μF in the cold. The cells were then diluted with DMEM/25 mM Hepes medium containing 10% FCS and cultured in 9-cm petri dishes.

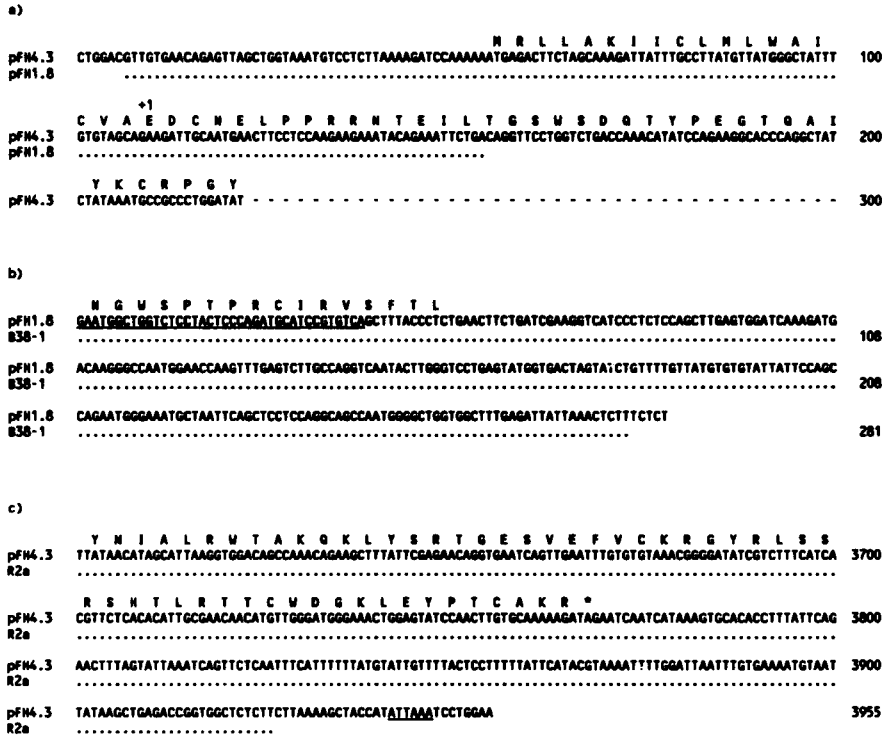

Figure 2. Partial nucleotide and derived amino acid sequences of factor H. (a) The 5′ sequence of the cDNA clones pFH4.3 and pFH1.8 is compared as far as determined. (b) The 3′ sequence of pFH1.8 is shown in comparison with the cDNA clone B38-1 also derived from the 1.8-kb mRNA. Here the nucleotide numbering corresponds to the B38-1 3′ sequence [13]. Sequence identity of the pFH4.3 sequence ending exactly with SCR 7 is indicated by dots. (c) The 3′ sequence of pFH4.3 is shown in comparison with the factor H_{gp150}-derived sequence R2a [7, 13]. The Hind III recognition sequence is present at position 3635. Here the nucleotide numbering corresponds to the R2a sequence. The putative polyadenylation signal in pFH4.3 is underlined.

After 24-h the medium was changed to DMEM containing 1 × Nutridoma HU (Boehringer Mannheim, Mannheim, FRG). SN were collected 72 h after the pulse, concentrated 20-fold, and assayed for the expression of factor H or factor H-related proteins by Western blot analysis.

2.5 Western blot analysis

SN of COS cell transfectants (80 µl) and human serum (3 µl) were analyzed by SDS-PAGE on a 10% gel under nonreducing conditions [22] and blotted to nitrocellulose according to the method of Towbin et al. [23]. Blots were stained with the mAb MAH-4 [11] and peroxidase-conjugated anti-mouse IgG using standard protocols.

3 Results

3.1 Isolation of the full-length cDNA clones pFH4.3 and pFH1.8

The orientation of the cDNA clones was determined by digestion with Hind III which cuts in the 3′ region of the 4.3-kb cDNA [13] and in the 5′ linker sequence of the vector CDM8 (see Figs. 1 and 2). The extent of the cloned 5′ sequence was resolved by digestion with Hind III and Bgl II.

Only one cDNA clone (pFH4.3) was found which contained a 4.3-kb insert in the correct 5′ to 3′ orientation. Determination of the DNA sequence of the 5′ and 3′ ends of the insert confirmed the orientation and that this cDNA was full length (see Fig. 2). All other clones hybridizing to both the H19 and H46 insert represented truncated cDNA of the 4.3-kb mRNA.

pFH4.3 is, in the region sequenced (222 bp of the 5′ sequence and 751 bp of the 3′ sequence), identical to the previously published factor H cDNA clones B38-1 and R2a. Thus, it contains 54 nucleotides of 5′ untranslated

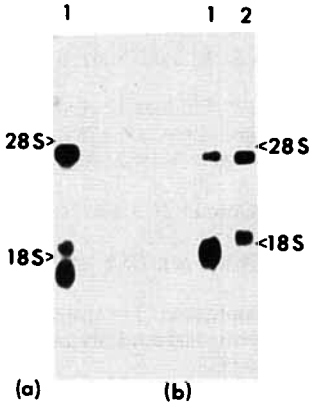

Figure 3. Northern blots of poly(A)⁺RNA from human liver. About 1 µg of poly(A)⁺ liver RNA was separated on an agarose/formaldehyde gel, transferred to nylon membrane and hybridized with the full-length pFH4.3 cDNA insert (a), the 250-bp Eco RV/Pst I fragment of pFH4.3, coding for the 3′ end of the 4.3-kb mRNA [(b) lane 1] and the full-length pFH1.8 cDNA insert [(b), lane 2]; Fig. 3a is derived from a different gel than Fig. 3b.

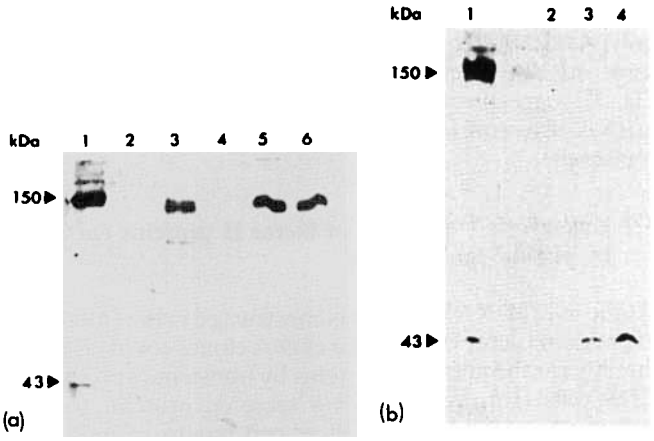

Figure 4. Western blot analysis for factor H expression in COS cell transfectants. Normal human serum, concentrated SN and cell lysates of the transfection assays were subjected to SDS-PAGE under nonreducing conditions and then analyzed by Western blotting with the anti-factor H mAb MAH-4. (a), Lane 1: 3 µl of human serum, lane 2: 80 µl of 20-fold concentrated SN of CDM8 COS cell transfectants, lane 3: lysate of 1 × 10^6 pFH4.3 transfected COS cells, lane 4: lysate of 1 × 10^6 CDM8 transfected COS cells, lane 5: 80 µl of 20-fold concentrated SN of COS cells transfected with pFH4.3 at 250 µF and 350 V, lane 6: same as 5 but electroporation was performed with a 380-V pulse. (b), Lane 1: 3 µl of human serum, lane 2: 80 µl of 20-fold concentrated SN of CDM8 COS cell transfectants, lane 3: 80 µl of 20-fold concentrated SN of COS cells transfected with pFH1.8 at 250 µF and 250 V, lane 4: same as 3 but electroporation was performed with a 380-V pulse. On the left, sizes of the detected factor H proteins are indicated in kDa.

sequence and 189 nucleotides of 3′ untranslated region which is 29 nucleotides longer than the R2a cDNA sequence ([13], see Fig. 2). Two putative polyadenylation signals are present in pFH4.3 at positions 3810 and 3942. Previous factor H cDNA clones ended prior to the second polyadenylation signal. The length of the 3′ untranslated region of pFH4.3 and of shorter factor H cDNA clones indicate that the second ATTAAA signal is recognized in human liver. This is confirmed by the same 3′ terminal sequence in H46 cDNA (Barbieri, Dierich, Schwaeble, unpublished results), isolated from a liver cDNA library prepared from RNA of another donor.

Several cDNA clones of 1.8 kb in length were characterized which hybridized to both the 5′ and 3′ fragment of H19, and thus should be derived from the 1.8-kb factor H mRNA. Only one full-length clone, pFH1.8, had the right orientation and could therefore directly be used for transfection. The completeness of this cDNA was confirmed by sequencing both ends (see Figs. 1 and 2). It contains 48 nucleotides of 5′ untranslated region and the 5′ nucleotide sequence is identical to pFH4.3. pFH1.8 corresponds to the previously published B38-1 cDNA, but does not contain any cloning artifact such as inversion within the 5′ sequence present in the previously described cDNA clones B38-1 [13] and H20 [8]. Thus, it has the same unique 3′ end as reported for the truncated cDNA B38-1, pH1050 [12] and H19 [11], characteristic for cDNA derived from the 1.8-kb factor H transcript (see Fig. 2).

Previous Northern blot results with different cDNA probes showed three factor H-related transcripts in human liver

[8]. We used the full length 4.3-kb cDNA insert to probe poly(A)$^+$RNA from human liver and detected the same three mRNA species as obtained with subfragments (see Fig. 3), suggesting that the smallest 1.4-kb factor H-related mRNA also represents a truncated form of the large 4.3-kb transcript.

3.2 Eukaryotic expression of factor H proteins encoded by pFH4.3 and pFH1.8

As the liver cDNA library was constructed in the eukaryotic expression vector CDM8, the cDNA clones could be tested directly for the encoded proteins by transient expression in COS cells [16]. Factor H is a secreted protein, for this reason we tested SN as well as cell lysates of individual transfection assays. As shown in Fig. 4a, pFH4.3 codes for a 150-kDa factor H protein which, in this assay, is indistinguishable from the major abundant factor H protein in human serum. No shorter factor H protein was detected in SN or cell lysates fo pFH4.3 transfectants. Therefore, it is unlikely that the smaller factor H-related proteins detected with anti-factor H reagents are derived from the 150-kDa protein by posttranslational processing events. Furthermore, the eukaryotic expression of pFH1.8, which represents the complete 1.8-kb factor H transcript, encodes a 43-kDa factor H protein demonstrating that this mRNA encodes a truncated factor H protein identical to that found in serum (see Fig. 4b).

4 Discussion

The results reported here demonstrate that the cDNA clones pFH4.3 and pFH1.8 represent full-length cDNA sequences derived from transcripts of the 4.3-kb and the 1.8-kb mRNA species of human factor H. The eukaryotic expression of both cDNA has proven that the 4.3-kb mRNA encodes the large classical 150-kDa factor H protein, while an additional truncated factor H-like protein of 43 kDa is encoded by the 1.8-kb mRNA. As both factor H proteins contain the apparent functional domains of cofactor H known so far, we propose the nomenclature factor H_{gp150} and factor H_{p43} to distinguish between the two molecules. Although the protein sequence of pFH1.8 contains one potential glycosylation site for Asn-linked carbohydrate, no evidence for such a modification of the 43-kDa factor H polypeptide in human serum has been found [13]. Several reports have to date described a truncated form of factor H and on SDS gels have shown a molecular weight of 41 000–45 000, much shorter than the calculated M_r of 49 000 [8, 9, 13]. This discrepancy in M_r is difficult to explain as a mobility artifact on SDS-PAGE, but may indicate post-transcriptional modification such that some amino acid residues, presumably of the carboxy terminus, have been removed. Our cDNA clones should be useful in future to produce recombinant native and mutant factor H proteins for the study of the complex functions of these factors.

The sequence identity of pFH4.3 and pFH1.8 in the 5′ part combined with the published sequences for the 4.3-kb and the 1.8-kb factor H cDNA clones covering in total 1400 nucleotides can only be explained by one gene coding for both transcripts. The arrangement of exactly the same 5′ untranslated region, in addition to an identical open reading frame in both mRNA (see this report, [11–13]), strongly suggests that they arise by alternative splicing from the transcription product of a single structural gene. This conclusion is supported by the partial gene structure for human factor H recently described by McAleer et al. [24], containing an extra exon encoding the specific 3′ end of the 1.8-kb transcript. In data reported here, the smallest transcript of 1.4-kb hybridized with only the 3′ factor H_{gp150} cDNA probes but not with 5′ fragments. The origin of this mRNA is unknown so far. Cloning of the 1.4-kb transcript coding for factor H-related molecules will show whether this mRNA is also derived from the common factor H gene and whether it only codes for a carboxyterminal fragment of factor H_{gp150}.

We thank Dr. Erik C. Böttger, Medizinische Hochschule, Hannover, for his helpful comments on preparation of the cDNA library and Dr. G. Arnold, Genzentrum, München, for synthesizing the two oligonucleotides ocdm8-1 and ocdm8-2.

Received June 6, 1990; in final revised form October 17, 1990.

5 References

1 Whaley, K. and Ruddy, S., *Science* 1976. *193:* 1011.
2 Whaley, K. and Ruddy, S., *J. Exp. Med.* 1976. *144:* 1147.
3 Conrad, D. H., Carlo, J. R. and Ruddy, S., *J. Exp. Med.* 1978. *147:* 1792.
4 Pangburn, M. K., Schreiber, R. D. and Müller-Eberhard, H. J., *J. Exp. Med.* 1977. *146:* 257.
5 Alsenz, J., Lambris, J. D., Schulz, T. F. and Dierich, M. P., *Biochem. J.* 1984. *224:* 389.
6 Alsenz, J., Lambris, J. D., Schulz, T. F. and Dierich, M. P., *Immunobiology. 168:* 97.
7 Day, A. J., Ripoche, J., Lyous, A., McIntosh, B., Harris, T. J. R. and Sim, R. B., *Biosci. Rep.* 1987. *7:* 201.
8 Schwaeble, W., Zwirner, J., Schulz, T. F., Linke, R. P., Dierich, M. P. and Weiss, E. H., *Eur. J. Immunol.* 1987. *17:* 1485.
9 Fontaine, M., Demares, M. J., Koistinen, V., Day, A. J., Davrinche, C., Sim, R. B. and Ripoche, J., *Biochem. J.* 1989. *258:* 927.
10 Schulz, T. F., Schwaeble, W., Stanley, K. K. and Dierich, M. P., *Mol. Immunol.* 1986. *23:* 1243.
11 Schulz, T. F., Schwaeble, W., Stanley, K. K., Weiss, E. H. and Dierich, M. P., *Eur. J. Immunol.* 1986. *16:* 1351.
12 Kristensen, T., Wetsel, R. A. and Tack, B. F., *J. Immunol.* 1986. *136:* 3407.
13 Ripoche, J., Day, A. J., Harris, T. J. R. and Sim, R. B., *Biochem. J.* 1988. *249:* 593.
14 Misasi, R., Huemer, H. P., Schwaeble, W., Söldner, E., Larcher, C. and Dierich, M. P., *Eur. J. Immunol.* 1989. *19:* 1765.
15 Sim, R. B. and DiScipio, R. G., *Biochem. J.* 1982. *205:* 285.
16 Seed, B., *Nature* 1987. *329:* 840.
17 Chirgwin, J. M., Przybyla, R. J., MacDonald, R. J. and Rutter, W. J., *Biochemistry* 1979. *18:* 5294.
18 Aruffo, A. and Seed, B., *Proc. Natl. Acad. Sci. USA* 1987. *84:* 8573.
19 Maniatis, T., Fritsch, E. F. and Sambrook, J., *Molecular cloning. A laboratory manual,* Cold Spring Harbor Laboratory, Cold Spring Harbor Press, New York 1982.
20 Feinberg, A. P. and Vogelstein, B., *Anal. Biochem.* 1983. *132:* 2.
21 Church, G. M. and Gilbert, W., *Proc. Natl. Acad. Sci. USA* 1984. *81:* 1991.
22 Laemmli, U. K. and Favre, M., *Nature* 1970. *227:* 680.
23 Towbin, H., Staehelin, T. and Gordon, J., *Proc. Natl. Acad. Sci. USA* 1979. *76:* 4350.
24 McAleer, M. A., Ripoche, J., Dominguez, O., Schwaeble, W. and Sim, R. B., *7th Intern. Congress of Immunology,* Abstract 24-17, S. Fisher, Stuttgart 1989, p. 140.