Vol. 214, No. 1, 1995 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
September 5, 1995 Pages 195-199

CHARACTERIZATION OF THE GENOMIC STRUCTURE AND THE PROMOTER REGION OF THE HUMAN INTESTINAL TREFOIL FACTOR

Thomas Seib, Steven Dooley and Cornelius Welter*

Department of Human Genetics, University of Saarland, 66421
Homburg, Germany

Received July 24, 1995

Trefoil proteins form a specific group of stable secretory polypeptides. They are expressed in a lot of human cancers and inflammatory processes of the gastrointestinal tract. Recently a new human trefoil protein, ITF/hP1.B was isolated. This gene is expressed mainly in globet cells of intestinum and colon. The genomic structure of the murine and rat homologues genes were described previously and a high nucleotide sequence conservation (78-95%) was found in the coding as well as the regulatory regions. We have isolated the human genomic ITF-gene region from a P1 library and a detailed analysis yielded in a similar gene structure with a three basepair displacement of the second exon intron boundary. The promoter region of the human gene interestingly is markedly different from rat and mouse.

Trefoil proteins form a specific group of stable secretory polypeptides. Members of this group are characterized by the presence of at least one copy of a 40-amino acid protein domain, termed the trefoil motif, that contains three conserved disulfide bonds (1). Three different trefoil proteins are known from human; the pS2/BCEI (breast cancer estrogen-induced), the spasmolytic polypeptide hSP and the intestinal trefoil factor (ITF/hP1.B). Physiologically pS2 and hSP are transcribed in human stomach mucosa (2-4) and ITF/hP1.B is mainly expressed in goblet cells of intestinum and colon (5,6). The function of these proteins is yet unknown. An activation of these proteins is induced by inflammatory processes of the gastrointestinal tract (6-8). It has been shown that trefoil peptides promote reestablishment of mucosal integry after injury (9). Furthermore is the expression of pS2 and hSP in different human tumors of interest (10-19). We have previously shown that the ITF/hP1.B gene in addition to pS2 is activated in human breast cancer (20). In the present study we have characterized the genomic locus and the promoter region of the human ITF gene and compaired its structure with the previously described rat and murine genes.

*Corresponding author. 49-6841-16-6600 (Fax); Email: hgsdoo@med-rz.uni-sb.de.

0006-291X/95 $12.00

Materials and Methods

Library screening

We screened a commercial P1 human genomic library using PCR technology (Genome Systems, St.Louis, MO, USA). The employed primers for PCR were derived from the 3_ untranslated region of the human intestinal trefoil factor c-DNA (5,6).
The sequences are: 5'-CCAGGCACTGTTCATCTCAG-3' and 5'-GGAGCATGGG ACCTTTATTC-3'. Using the following conditions a 108 bp fragment was amplified and used as screening probe: 93°C, 3 min, 30 x 92°C for 1 min, 60°C for 30 sec and 72°C for 30 sec.

DNA preparation

The recombinant DNA of the obtained P1 clone was purified according to the protocol of Kimmerly (21).

DNA sequencing

Sequencing was performed using an ABI 373A DNA sequencer and the Taq DyeDeoxy Terminator Cycle Sequencing kit (Applied Biosystems). The reaction mixture contained 6-8 ug of purified P1-DNA and 3.2 pmol of the appropriate primer. The sequencing reactions were carried out in a Perkin Elmer Cetus Model 9600 Thermocycler according to the ABI sequencing protocol.
After identification of exon/intron boundaries we selected new sequencing primers according to the nucleotide sequences of the next exon. These primers were used in subsequent PCR experiments to amplify the entire introns and thus determine their size. The resulting PCR products underwent direct sequencing after purification with the Wizard DNA clean-up system (Promega).
In addition we applied a primer walking strategy to reveal the sequence of the 5'-regulatory region. The resulting sequence was analysed in regard to protein binding sites and enhancer motifs by application of MacVector software (Kodak). The human and rat promoter sequences were compared using the MACAW program. All primers used for sequencing and PCR amplification were presented in Table 1.

Results and Discussion

Four different P1 clones were obtained containing the complete human ITF genomic region. The length of the gene and the exon-intron boundaries were determined. Three exons and two introns span a distance of 3.2 kb DNA

Table1

Tre1	5'-ATGCTGGGGCTGGTCCTG-3'
Tre2	5'-GGAGCATGGGACCTTTATTC-3'
ITF2	5'-CCAGGCACTGTTCATCTCAG-3'
ITF3_	5'-AAACCAGTGTGCCGTGCC-3'
ITF4	5'-CTGGAGGTGCCTCAGAAG-3'
ITF6	5'-CTTGAAACACCAAGGCACTC-3'
ITF-Pr	5'-AAGGCCAGGACCAGCCCC-3'
ITF-Pr2	5'-TGTCAATTGCCCTCCTAG-3'
ITF-Pr3	5'-GGATGAAATGAATTGATTG-3'
ITF-Pr/rev	5'-ACTTAGGAGGCTGAGGTGGG-3'
ITF-Pr/rev2	5'-CAATAGATGCCTCTGGATAG-3'

sequence. In addition we have characterized 930 bp of the ITF 5′-flanking sequences (Fig 1). Compaired to the published rat and murine ITF genes (20,22) the first exon intron boundary is conserved and the second one is displaced of three basepairs (Fig 2). The length of both introns also is different in human and rat. Moreover the promoter regions of these genes are markedly different. In contrast to the rat sequence no TATA-box exists within the human ITF gene and the rat AT rich sequence, a cannonical Pit-1 homeodomain the 14 bp homeodomain-like element also could not be identified. In the human gene we identified potential binding sites for several transcription factors (see Fig 1) whose involvement in the regulation of ITF transcription remains to be determined. A comparison of the 5′-flanking regions of rat and human ITF genes revealed no striking sequence homology except for a region of 60 bp with 85 % conservation.

```
                    SP1
-930  ACTTAGGAGG CTGAGGTGGG AGGATCACTT GAGCCCCGCG AGGGCGGGGG
-880  TCTCAGTGAG TTGAGATTGT GCCACTGCAC TCCAGCCTGG GTAACAGAGC
-830  AAGACCCTGT CTCCAAAAAA AAAAAAACAA AAAAACAAAA CAAAAAACAA
-780  ACAAACAAAC AAACAAACAA AAACAAACCA AAAAAACGTG CTTAGATCCA
-730  GAGAGAAAAA TTCTGAAAGG ATACCTAAAA AAGTGTTCAC AATAGATCAA
-680  TAGATGCCTC TGGATAGGAA AGGCTCTTTT GTATTATTTG ATTTTAAAAA
-630  TCAATTATGT GTACCATGTT TTTACTAACA TATTTTCAAT CAATTCATTT
-580  CATCCACTTA TTAAAATGTA GTTACTAATT TTTAGGGGAG AAAGCAAAAA
-530  GGAAAGACAA GGAATCTCTG TGTTTCAGGA GTTGTGAGAG AGCCGCAGGG
          AP-1
-480  TCCTGACTCA CTCAGAGCTG CCTGTCTCCG AGGCCGATCT GGGATGAAGC
-430  AGCCTGGGGC TCTCTTGTCA TGGGACCAGG GGTGTTCTGA GGGCTTCTGG
               SP1      CF1
-380  CTGGGAGGCT GAGATGGAAC GGACACCACA CCCTGGTCCT GCCACCCCAC
     CF1
-330  ATGGCTCCTG CACACTACAC CAGGCCAGGC TAGGAGGGCA ATTGACACAC
-280  ATCCGCTCCC CAGTAGAGGA CCCGGAACCA GAACTGGAAT CCGCCCTTAC
-230  CGCTTGCTGC CAAAACAGTG GGGGCTGAAC TGACCTCTCC CCTTTGGGAG
-180  AGAAAAACTG TCTGGGAGCT TGACAAAGGC ATGCAGGAGA GAACAGGAGC
-130  AGCCACAGCC AGGAGGGAGA GCCTTCCCCA AGCAAACAAT CCAGAGCAGC
                          GHF-1GH.1
-80   TGTGCAAACA ACGGTGCATA AATGAGGCCT CCTGGACCAT GAAGCGAGTC
                                      +1
-30   CTGAGCTGCG TCCCGGAGCC CACGGTGGTC ATG GCT GCC AGA GCG
                                       M   A   A   R   A

+16   CTC TGC ATG CTG GGG CTG GTC CTG GCC TTG CTG TCC TCC AGC
       L   C   M   L   G   L   V   L   A   L   L   S   S   S

+58   TCT GCT GAG GAG TAC GTG GGC CTG T  gtgagtactg-1200BP-
       S   A   E   E   Y   V   G   L  S  -------Intron1----

+83   tctgggacag  CT GCA AAC CAG TGT GCC GTG CCA GCC AAG GAC
      ----------      A   N   Q   C   A   V   P   A   K   D

+115  AGG GTG GAC TGC GGC TAC CCC CAT GTC ACC CCC AAG GAG TGC
       R   V   D   C   G   Y   P   H   V   T   P   K   E   C

+157  AAC AAC CGG GGC TGC TGC TTT GAC TCC AGG ATC CCT GGA GTG
       N   N   R   G   C   C   F   D   S   R   I   P   G   V

+199  CCT TGG TGT TTC AAG CCC CTG CAG GAA GCA G  gtaaggcccc-
       P   W   C   F   K   P   L   Q   E   A  E  --Intron2--

+230  1600Bp-ctgtttgcag  AA TGC ACC TTC tgaggcacct ccagctgccc
      -----------------      C   T   F

      ccggccgggg gatgcgaggc tcggagcacc cttgcccggc tgtgattgct
      gccaggcact gttcatctca gcttttctgt ccctttgctc ccggcaagcg
      cttctgctga aagttcatat ctggagcctg atgtcttaac gaataaaggt
      cccatgctcc
```

Figure1. Nucleotide sequence of the 5′-region of the HITF gene. Putative binding sites for transcription factors are indicated, the region with significant homology with the promoter of the RITF-gene is double underlined and the start codon is marked by bold letters.The positive numbering refers to the coding sequence. The lowercase letters indicate intron sequences and the 3′ untranslated region.

```
Exon1
     1                                          82
hITF ATGGCTGCCA ..............GTGGGCCTGT
mITF ATGGAGACCA...............GTTGGCCTGT
rITF ATGGAGACCA...............GTTGGCCTGT

Exon2
     83                                         199
hITF CTGCAAACCA...............CAGGAAGCAG
mITF CTCCAAGCCA............CTGCAGGAGA
rITF CTCCAAGCCA............CTGCAAGAGA

Exon3
        200
hITF    AATGCACCTT
mITF CAGAATGCACATT
rITF CAGAATGTACATT
```

Figure 2. Exon-intron boundaries of human intestinal trefoil factor (hITF) compared to the homologue genes of mouse (mITF) and rat (rITF); The numbers refer to the human sequence shown in Fig1.

Because of the discrepancy between these two conserved genes we have performed several control experiments to verify our results. We have sequenced three independently isolated P1 clones and always yielded the same sequence and we analyzed a PCR fragment amplified from human genomic DNA containing sequences from a part of 5'-flanking region, first exon and intron which also gave identical results. It has been shown that the ITF gene specifically is expressed in globet cells (5,6). Based on these observations it is likely that the proximal 5_-flanking region of the ITF gene contains elements which are capable of directing globet cell specific expression. In this context the 60 bp conserved region could be of importance and gene regulatory studies are presently under investigation to confirm this assumption.

Acknowledgment

This study was supported by a grant of the Deutsche Krebshilfe.

References

1.Hoffmann, W. and Hauser, F. (1993) Trends. Biochem. Sci. 18, 239-243.
2.Rio, M.C., Bellocq, J.P., Daniel, J.Y., Tomasetto, C., Lathe, R., Chenard, M.P., Batzenschlager, A., and Chambon, P. (1988) Science 241, 705-708.
3.Tomasetto, C., Rio, M.C., Gautier, C., Wolf, C., Hareuveni, M., Chambon, P., and Lathe, R. (1990) EMBO J. 9, 407-414.
4.Hanby, A.M., Poulsom, R., Singh, S., Elia, G., Jeffery, R.E., and Wright, N.A. (1993) Gastroenterology 105, 1110-1116.
5.Podolsky, D.K., Lynch-Devaney, K., Stow, J.L., Oates, P., Murgue, B., DeBeaumont, M., Sands, B.E., and Mahida, Y.R. (1993) J. Biol. Chem. 268, 6694-6702.
6.Hauser, F., Poulsom, R., Chinery, R., Rogers, L.A., Hanby, A.M., Wright, N.A., and Hoffmann, W. (1993) Proc. Natl. Acad. Sci. U. S. A. 90, 6961-6965.
7.Rio, M.C., Chenard, M.P., Wolf, C., Marcellin, L., Tomasetto, C., Lathe, R., Bellocq, J.P., and Chambon, P. (1991) Gastroenterology 100, 375-379.

8.Wright, N.A., Poulsom, R., Stamp, G., Van-Noorden, S., Sarraf, C., Elia, G., Ahnen, D., Jeffery, R., Longcroft, J., Pike, C., and et-al, (1993) Gastroenterology 104, 12-20.
9.Dignass, A., Lynch-Devaney, K., Kindon, H., Thim, L., and Podolsky, D.K. (1994) J. Clin. Invest. 94, 376-383.
10.Pichon, M.F. and Milgrom, E. (1993) Crit. Rev. Oncol. Hematol. 15, 13-21.
11.Henry, J.A., Bennett, M.K., Piggott, N.H., Levett, D.L., May, F.E., and Westley, B.R. (1991) Br. J. Cancer 64, 677-682.
12.Luqmani, Y.A., Ryall, G., Shousha, S., and Coombes, R.C. (1992) Int. J. Cancer 50, 302-304.
13.Welter, C., Seitz, G., and Blin, N. (1993) Acta Oncol. 32, 315-317.
14.Theisinger, B., Welter, C., Seitz, G., Rio, M.C., Lathe, R., Chambon, P., and Blin, N. (1991) Eur. J. Cancer 27, 770-773.
15.Möller, W. and Borchard, F. (1993) J. Pathol. 171, 263-269.
16.Welter, C., Theisinger, B., Seitz, G., Tomasetto, C., Rio, M.C., Chambon, P., and Blin, N. (1992) Lab. Invest. 66, 187-192.
17.Seitz, G., Thelsinger, B., Tomasetto, G., Rio, M.C., Chambon, P., Blin, N., and Welter, G. (1991) Am. J. Gastroenterol. 86, 1491-1494.
18.Welter, C., Theisinger, B., Rio, M.C., Seitz, G., Schuder, G., and Blin, N. (1994) Int. J. Cancer 56, 52-55.
19.Wysocki, S.J., Hahnel, E., Masters, A., Smith, V., McCartney, A.J., and Hahnel, R. (1990) Cancer Res.50, 1800-1802.
20.Theisinger, B., Seitz, G., Dooley, S., and Welter, C. (1995) Breast Cancer Res. Treat., in press.
21.Kimmerly, W.J., Kyle, A.L., Lustre, V.M., Martin, C.H., and Palazzolo, M.J. (1994) Gen Anal Tech App 11(5-6), 117-128.
22.Sands, B.E., Okata, H., Lynch-Devaney, K., DeBeaumont, M., Ezzell, R.M., and Podolsky, D.K. (1995) J. Biol. Chem. 270, 9353-9361.