ELSEVIER

Biochimica et Biophysica Acta 1493 (2000) 356–362

BIOCHIMICA ET BIOPHYSICA ACTA
BBA

www.elsevier.com/locate/bba

Short sequence-paper

Pregnancy-associated plasma protein-E (PAPP-E)[1]

Martin Farr *, Jörg Strübe, Harald-Gerhard Geppert, Andreas Kocourek,
Martina Mahne, Harald Tschesche

Department of Biochemistry I, Faculty of Chemistry and Biochemistry, University of Bielefeld, D-33501 Bielefeld, Germany

Received 17 May 2000; accepted 27 July 2000

Abstract

A full-length cDNA encoding a novel human protein was cloned from placenta cDNA. The corresponding 1542 amino acid protein sequence was termed 'pregnancy-associated plasma protein-E' (PAPP-E) as it shows a 62% homology to the human pregnancy-associated plasma protein-A (PAPP-A) that is a diagnostic marker for trisomies, especially Down syndrome. The conserved domain structure contains five motifs related to the short consensus repeats of complement proteins and selectins, three motifs related to the lin-notch motifs of proteins regulating early tissue differentiation, and a putative zinc-binding motif and active site of the metzincin-superfamily of metalloproteases. The PAPP-E gene was localized to chromosome 1q23–25. Northern blot analysis showed that PAPP-E is predominantly expressed in placenta. © 2000 Elsevier Science B.V. All rights reserved.

Keywords: Novel human cDNA; Pregnancy-associated plasma protein; Chromosome 1q23–25; Placenta

Pregnancy-associated plasma protein-A (PAPP-A) is a maternal serum protein of placental origin that is disulfide-bonded to the proform of eosinophil major basic protein [1]. The complex is detectable from the fourth to sixth week of pregnancy, and reaches concentrations up to 50 mg/l in the last trimester [2]. As the serum PAPP-A concentration is reduced by about 50–77% in the sixth to 14th week of pregnancies with fetal trisomies, PAPP-A in combination with other markers is used for non-invasive, early detection of trisomies [3]. The biological role of PAPP-A is not known, but a recent study has demonstrated that PAPP-A cleaves the insulin-like growth factor binding protein-4 (IGFBP-4) and could be involved in IGF-signaling [4]. Proteolysis might result from the putative zinc-coordination site followed by a putative Met-turn that are conserved active site motifs of the metzincin-superfamily of metalloproteinases [5,6]. Besides this, PAPP-A comprises five motifs related to short consensus repeats (SCRs) of complement proteins and selectins, and three motifs related to lin-notch repeats (LNRs) of proteins regulating early tissue differentiation [5,7,8].

Using the tblastn program [9] we have screened the GenBank human EST entries with the cDNA sequence of the catalytic domain of the human metzincin MMP-8 (matrix metalloproteinase-8, neutrophil collagenase). Besides known sequences of MMPs we have found a 315 bp EST (EST79446, gi: 2020543) that codes for the conserved zinc-binding motif of metzincins and is similar to the PAPP-A cDNA. We have cloned this EST by preparing RNA from placenta using the RNeasy Total RNA System (Qiagen) and performing reverse transcription with SuperScript II Rnase H$^-$ reverse transcriptase (Life Technologies), the polymerase chain reaction (PCR) with *Taq* DNA Polymerase (Roche) and the oligonucleotides 5′-tgaaggagctgaaggaggccctgc-3′ and 5′-cggtgtcggcacagaggtctcccg-3′, and TOPO TA Cloning (Invitrogen). Subsequent 5′- and 3′-RACE (rapid amplification of cDNA ends) with Marathon-Ready Placenta cDNA and Advantage 2 Polymerase Mix (Clontech) and TOPO XL PCR Cloning kit (Invitrogen) resulted in three cDNA fragments comprising a 4872 bp open reading frame with 5′- and 3′-flanking sequences (Fig. 1). Re-screening of the GenBank entries shows that the genomic sequences AL031734.9 (HS652L8) and AL031290.1 (HS774I24) partially cover the 3′ and 5′ PAPP-E cDNA. The deduced PAPP-E amino acid sequence shows global similarity only with PAPP-

---

* Corresponding author. Present address: Heart Center North Rhine-Westphalia, University Hospital of the Ruhr – University of Bochum, D-32545 Bad Oeynhausen, Germany. Fax: +49-5731-972194; E-mail: mfarr@hdz-nrw.de

[1] The cDNA and protein data reported here have been deposited in the EMBL Nucleotide Sequence Database under accession number AJ278348, HSA278348.

0167-4781/00/$ – see front matter © 2000 Elsevier Science B.V. All rights reserved.
PII: S0167-4781(00)00195-0

```
                      tccgagggcgcccgggcaggtgttctgccaactctgagctgggctggacacgcaagaaatccttggttgagagggaacacctgaatcaggtgctgttggaaggagaacgt  -601

 -600  tgttggctgggggccaaggttcgaagacccagagcttctccacagcatcacctctttggagtctaccccagcagggctgggaactacctaaggccctaccccgtgggggagcaagaaatc  -481

 -480  catcatacaggacgcagcaaaccagacactgaaggaaatgctgtgagccttgttcccccagacctgactgaaaatccagcaggactgaggggtgcagttgaagagccggctgccccatgg  -361

 -360  gtaggggatagtcctattgggcaatctgagctgctgggagatgatgacgcttatctcggcaatcaaagatccaaggagtctctaggtgaggccgggattcagaaaggctcagccatggct  -241
                                                                                                                        M  A    -81

 -240  gccactactaccaccgccattttcacaaccctgaacgaacccaaaccagagacccaaaggaggggctgggccaagtccaggcagcgtcgccaagtgtggaagaggcggcggaagatggg  -121
  -80   A  T  T  T  T  A  I  F  T  T  L  N  E  P  K  P  E  T  Q  R  R  G  W  A  K  S  R  Q  R  R  Q  V  W  K  R  R  A  E  D  G    -41

 -120  cagggagactccggtatctcttcacatttccaaccttggcccaagcattcccttaaacacagggtcaaaaagagtccaccggaggaaagcaaccaaaatggtggagagggctcctaccga    -1
  -40   Q  G  D  S  G  I  S  S  H  F  Q  P  W  P  K  H  S  L  K  H  R  V  K  K  S  P  P  E  E  S  N  Q  N  G  G  E  G  S  Y  R     -1

    1  gaagcagagacctttaactcccaagtaggactgcccatcttatacttctctgggaggcgggagcggctgctgctgcgtccagaagtgctggctgagattccccgggaggcgttcacagtg   120
    1   E  A  E  T  F  N  S  Q  V  G  L  P  I  L  Y  F  S  G  R  R  E  R  L  L  L  R  P  E  V  L  A  E  I  P  R  E  A  F  T  V     40

  121  gaagcctgggttaaaccggagggaggacagaacaacccagccatcatcgcaggtgtgtttgataactgctcccacactgtcagtgacaaaggctgggccctggggatccgctcagggaag   240
   41   E  A  W  V  K  P  E  G  G  Q  N  N  P  A  I  I  A  G  V  F  D  N  C  S  H  T  V  S  D  K  G  W  A  L  G  I  R  S  G  K     80

  241  gacaagggaaagcgggatgctcgcttcttcttctccctctgcaccgaccgcgtgaagaaagccaccatcttgattagccacagtcgctaccaaccaggcacatggacccatgtggcagcc   360
   81   D  K  G  K  R  D  A  R  F  F  F  S  L  C  T  D  R  V  K  K  A  T  I  L  I  S  H  S  R  Y  Q  P  G  T  W  T  H  V  A  A    120

  361  acttacgatggacggcacatggccctgtatgtggatggcactcaggtggctagcagtctagaccagtctggtcccctgaacagccccttcatggcatcttgccgctctttgctcctgggg   480
  121   T  Y  D  G  R  H  M  A  L  Y  V  D  G  T  Q  V  A  S  S  L  D  Q  S  G  P  L  N  S  P  F  M  A  S  C  R  S  L  L  L  G    160

  481  ggagacagctctgaggatgggcactatttccgtggacacctgggcacactggttttctggtcgaccgccctgccacaaagccatttcagcacagttctcagcattcaagtgaggaggag   600
  161   G  D  S  S  E  D  G  H  Y  F  R  G  H  L  G  T  L  V  F  W  S  T  A  L  P  Q  S  H  F  Q  H  S  S  Q  H  S  S  E  E  E    200

  601  gaagcgactgacttggtcctgacagcgagctttgagcctgtgaacacagagtgggttccctttagagatgagaagtacccacgacttgaggttctccagggctttgagccagagcctgag   720
  201   E  A  T  D  L  V  L  T  A  S  F  E  P  V  N  T  E  W  V  P  F  R  D  E  K  Y  P  R  L  E  V  L  Q  G  F  E  P  E  P  E    240

  721  attctgtcgcctttgcagcccccactctgtgggcaaacagtctgtgacaatgtggaattgatctcccagtacaatggatactggccccttcggggagagaaggtgatacgctaccaggtg   840
  241   I  L  S  P  L  Q  P  P  L  C  G  Q  T  V  C  D  N  V  E  L  I  S  Q  Y  N  G  Y  W  P  L  R  G  E  K  V  I  R  Y  Q  V    280

  841  gtgaacatctgtgatgatgagggcctaaaccccattgtgagtgaggagcagattcgtctgcagcacgaggcactgaatgaggccttcagccgctacaacatcagctggcagctgagcgtc   960
  281   V  N  I  C  D  D  E  G  L  N  P  I  V  S  E  E  Q  I  R  L  Q  H  E  A  L  N  E  A  F  S  R  Y  N  I  S  W  Q  L  S  V    320

  961  caccaggtccacaattccaccctgcgacaccgggttgtgcttgtgaactgtgagcccagcaagattggcaatgaccattgtgaccccgagtgtgagcacccactcacaggctatgatggg  1080
  321   H  Q  V  H  N  S  T  L  R  H  R  V  V  L  V  N  C  E  P  S  K  I  G  N  D  H  C  D  P  E  C  E  H  P  L  T  G  Y  D  G    360

 1081  ggtgactgccgcctgcagggccgctgctactcctggaaccgcagggatgggctctgtcacgtggagtgtaacaacatgctgaacgactttgacgacggagactgctgcgacccccaggtg  1200
  361   G  D  C  R  L  Q  G  R  C  Y  S  W  N  R  R  D  G  L  C  H  V  E  C  N  N  M  L  N  D  F  D  D  G  D  C  C  D  P  Q  V    400

 1201  gctgatgtgcgcaagacctgctttgaccctgactcacccaagagggcatacatgagtgtgaaggagctgaaggaggccctgcagctgaacagtactcacttcctcaacatctactttgcc  1320
  401   A  D  V  R  K  T  C  F  D  P  D  S  P  K  R  A  Y  M  S  V  K  E  L  K  E  A  L  Q  L  N  S  T  H  F  L  N  I  Y  F  A    440

 1321  agctcagtgcgggaagaccttgcaggtgctgccacctggccttgggacaaggacgctgtcactcacctgggtggcattgtcctcagcccagcatattatgggatgcctggccacaccgac  1440
  441   S  S  V  R  E  D  L  A  G  A  A  T  W  P  W  D  K  D  A  V  T  H  L  G  G  I  V  L  S  P  A  Y  Y  G  M  P  G  H  T  D    480

 1441  accatgatccatgaagtgggacatgttctgggactctaccatgtctttaaaggagtcagtgaaagagaatcctgcaatgacccctgcaaggagacagtgccatccatggaaacgggagac  1560
  481   T  M  I  H  E  V  G  H  V  L  G  L  Y  H  V  F  K  G  V  S  E  R  E  S  C  N  D  P  C  K  E  T  V  P  S  M  E  T  G  D    520

 1561  ctctgtgccgacaccgcccccactcccaagagtgagctgtgccgggaaccagagcccactagtgacacctgtggcttcactcgcttcccaggggctccgttcaccaactacatgagctac  1680
  521   L  C  A  D  T  A  P  T  P  K  S  E  L  C  R  E  P  E  P  T  S  D  T  C  G  F  T  R  F  P  G  A  P  F  T  N  Y  M  S  Y    560

 1681  acggatgataactgcactgacaacttcactcctaaccaagtggcccgaatgcattgctatttggacctagtctatcagcagtggactgaaagcagaaagcccacccccacccccattcca  1800
  561   T  D  D  N  C  T  D  N  F  T  P  N  Q  V  A  R  M  H  C  Y  L  D  L  V  Y  Q  Q  W  T  E  S  R  K  P  T  P  T  P  I  P    600

 1801  cctatggtcatcggacagaccaacaagtccctcactatccactggctgcctcctattagtggagttgtatatgacagggcctcaggcagcttgtgtggcgcttgcactgaagatgggacc  1920
  601   P  M  V  I  G  Q  T  N  K  S  L  T  I  H  W  L  P  P  I  S  G  V  V  Y  D  R  A  S  G  S  L  C  G  A  C  T  E  D  G  T    640

 1921  tttcgtcagtatgtgcacacagcttcctcccggcgggtgtgtgactcctcaggttattggaccccagaggaggctgtggggcctcctgatgtggatcagccctgcgagccaagcttacag  2040
  641   F  R  Q  Y  V  H  T  A  S  S  R  R  V  C  D  S  S  G  Y  W  T  P  E  E  A  V  G  P  P  D  V  D  Q  P  C  E  P  S  L  Q    680

 2041  gcctggagccctgaggtccacctgtaccacatgaacatgacggtcccctgccccacagaaggctgtagcttggagctgctcttccaacacccggtccaagccgacaccctcaccctgtgg  2160
  681   A  W  S  P  E  V  H  L  Y  H  M  N  M  T  V  P  C  P  T  E  G  C  S  L  E  L  L  F  Q  H  P  V  Q  A  D  T  L  T  L  W    720

 2161  gtcacttccttcttcatggagtcctcgcaggtcctctttgacacagagatcttgctggaaaacaaggagtcagtgcacctgggccccttagacactttctgtgacatcccactcaccatc  2280
  721   V  T  S  F  F  M  E  S  S  Q  V  L  F  D  T  E  I  L  L  E  N  K  E  S  V  H  L  G  P  L  D  T  F  C  D  I  P  L  T  I    760

 2281  aaactgcacgtggatgggaaggtgtcgggggtgaaagtctacacctttgatgagaggatagagattgatgcagcactcctgacttctcagccccacagtcccttgtgctctggctgcagg  2400
  761   K  L  H  V  D  G  K  V  S  G  V  K  V  Y  T  F  D  E  R  I  E  I  D  A  A  L  L  T  S  Q  P  H  S  P  L  C  S  G  C  R    800

 2401  cctgtgaggtaccaggttctccgcgatcccccatttgccagtggtttgcccgtggtggtgacacattctcacaggaagttcacggacgtggaggtcacacctggacagatgtatcagtac  2520
  801   P  V  R  Y  Q  V  L  R  D  P  P  F  A  S  G  L  P  V  V  V  T  H  S  H  R  K  F  T  D  V  E  V  T  P  G  Q  M  Y  Q  Y    840

 2521  caagttctagctgaagctggaggagaactgggagaagcttcgcctcctctgaaccacattcatggagctccttattgtggagatgggaaggtgtcagagagactgggagaagagtgtgat  2640
  841   Q  V  L  A  E  A  G  G  E  L  G  E  A  S  P  P  L  N  H  I  H  G  A  P  Y  C  G  D  G  K  V  S  E  R  L  G  E  E  C  D    880

 2641  gatggagaccttgtgagcggagatggctgctccaaggtgtgtgagctggaggaaggtttcaactgtgtaggagagccaagcctttgctacatgtatgagggagatggcatatgtgaacct  2760
  881   D  G  D  L  V  S  G  D  G  C  S  K  V  C  E  L  E  E  G  F  N  C  V  G  E  P  S  L  C  Y  M  Y  E  G  D  G  I  C  E  P    920

 2761  tttgagagaaaaaccagcattgtagactgtggcatctacactcccaaaggatacttggatcaatgggctacccgggcttactcctctcatgaagacaagaagaagtgtcctgtttccttg  2880
  921   F  E  R  K  T  S  I  V  D  C  G  I  Y  T  P  K  G  Y  L  D  Q  W  A  T  R  A  Y  S  S  H  E  D  K  K  K  C  P  V  S  L    960
```

Fig. 1. cDNA and deduced protein sequence of PAPP-E. The putative prepropeptide is in italics and according to the PAPP-A sequence [5] negatively numbered; start and stop codons are boldface.

```
2881  gtaactggagaacctcattccctaatttgcacatcataccatccagatttacccaaccaccgtcccctaactggctggtttccctgtgttgccagtgaaaatgaaactcaggatgacagg  3000
 961   V  T  G  E  P  H  S  L  I  C  T  S  Y  H  P  D  L  P  N  H  R  P  L  T  G  W  F  P  C  V  A  S  E  N  E  T  Q  D  D  R   1000

3001  agtgaacagccagaaggtagcctgaagaaagaggatgaggtttggctcaaagtgtgtttcaatagaccaggagaggccagagcaatttttatttttttgacaactgatggcctagttccc  3120
1001   S  E  Q  P  E  G  S  L  K  K  E  D  E  V  W  L  K  V  C  F  N  R  P  G  E  A  R  A  I  F  I  F  L  T  T  D  G  L  V  P   1040

3121  ggagagcatcagcagccgacagtgactctctacctgaccgatgtccgtggaagcaaccactctcttggaacctatggactgtcatgccagcataatccactgattatcaatgtgacccat  3240
1041   G  E  H  Q  Q  P  T  V  T  L  Y  L  T  D  V  R  G  S  N  H  S  L  G  T  Y  G  L  S  C  Q  H  N  P  L  I  I  N  V  T  H   1080

3241  caccagaatgtccttttccaccataccacctcagtgctgccgaatttctcatccccacgggtcggcatctcagctgtggctctaaggacatcctcccgcattggtctttcggctcccagt  3360
1081   H  Q  N  V  L  F  H  H  T  T  S  V  L  P  N  F  S  S  P  R  V  G  I  S  A  V  A  L  R  T  S  S  R  I  G  L  S  A  P  S   1120

3361  aactgcatctcagaggacgaggggcagaatcatcagggacagagctgtatccatcggccctgtgggaagcaggacagctgtccgtcattgctgcttgatcatgctgatgtggtgaactgt  3480
1121   N  C  I  S  E  D  E  G  Q  N  H  Q  G  Q  S  C  I  H  R  P  C  G  K  Q  D  S  C  P  S  L  L  L  D  H  A  D  V  V  N  C   1160

3481  acctctataggcccaggtctcatgaagtgtgctatcacttgtcaaaggggatttgcccttcaggccagcagtgggcagtacatcaggcccatgcagaaggaaattctgctcacatgttct  3600
1161   T  S  I  G  P  G  L  M  K  C  A  I  T  C  Q  R  G  F  A  L  Q  A  S  S  G  Q  Y  I  R  P  M  Q  K  E  I  L  L  T  C  S   1200

3601  tctgggcactgggaccagaatgtgagctgccttcccgtggactgcggtgttcccgacccgtctttggtgaactatgcaaacttctcctgctcagagggaaccaaatttctgaaacgctgc  3720
1201   S  G  H  W  D  Q  N  V  S  C  L  P  V  D  C  G  V  P  D  P  S  L  V  N  Y  A  N  F  S  C  S  E  G  T  K  F  L  K  R  C   1240

3721  tcaatctcttgtgtcccaccagccaagctgcaaggactgagcccatggctgacatgtcttgaagatggtctctggtctctccctgaagtctactgcaagttggagtgtgatgctccccct  3840
1241   S  I  S  C  V  P  P  A  K  L  Q  G  L  S  P  W  L  T  C  L  E  D  G  L  W  S  L  P  E  V  Y  C  K  L  E  C  D  A  P  P   1280

3841  attattctgaatgccaacttgctcctgcctcactgcctccaggacaaccacgacgtgggcaccatctgcaaatatgaatgcaaaccagggtactatgtggcagaaagtgcagagggtaaa  3960
1281   I  I  L  N  A  N  L  L  L  P  H  C  L  Q  D  N  H  D  V  G  T  I  C  K  Y  E  C  K  P  G  Y  Y  V  A  E  S  A  E  G  K   1320

3961  gtcaggaacaagctcctgaagatacaatgcctggaaggtggaatctgggagcaaggcagctgcattcctgtggtgtgtgagccaccccctcctgtgtttgaaggcatgtatgaatgtacc  4080
1321   V  R  N  K  L  L  K  I  Q  C  L  E  G  G  I  W  E  Q  G  S  C  I  P  V  V  C  E  P  P  P  P  V  F  E  G  M  Y  E  C  T   1360

4081  aatggcttcagcctggacagccagtgtgtgctcaactgtaaccaggaacgtgaaaagcttcccatcctctgcactaaagagggcctgtggacccaggagtttaagttgtgtgagaatctg  4200
1361   N  G  F  S  L  D  S  Q  C  V  L  N  C  N  Q  E  R  E  K  L  P  I  L  C  T  K  E  G  L  W  T  Q  E  F  K  L  C  E  N  L   1400

4201  caaggagaatgcccaccacccccctcagagctgaattctgtggagtacaaatgtgaacaaggatatgggattggtgcagtgtgttccccattgtgtgtaatcccccccagtgaccccgtg  4320
1401   Q  G  E  C  P  P  P  P  S  E  L  N  S  V  E  Y  K  C  E  Q  G  Y  G  I  G  A  V  C  S  P  L  C  V  I  P  P  S  D  P  V   1440

4321  atgctacctgagaatatcactgctgacactctggagcactggatggaacctgtcaaagtccagagcattgtgtgcactggccggcgtcaatggcacccagaccccgtcttagtccactgc  4440
1441   M  L  P  E  N  I  T  A  D  T  L  E  H  W  M  E  P  V  K  V  Q  S  I  V  C  T  G  R  R  Q  W  H  P  D  P  V  L  V  H  C   1480

4441  atccagtcatgtgagcccttccaagcaaatggttggtgtgacactatcaacaaccgagcctactgccactatgacgggggagactgctgctcttccacactctcctccaagaaggtcatt  4560
1481   I  Q  S  C  E  P  F  Q  A  N  G  W  C  D  T  I  N  N  R  A  Y  C  H  Y  D  G  G  D  C  C  S  S  T  L  S  S  K  K  V  I   1520

4561  ccatttgctgctgactgtgacctggatgagtgcacctgccgggaccccaaggcagaagaaaatcagtaactgtgggaacaagcccctccctccactgcctcagaggcagtaagaaagaga  4680
1521   P  F  A  A  D  C  D  L  D  E  C  T  C  R  D  P  K  A  E  E  N  Q Stop

4681  ggccgacccaggaggaaacaaagggtgaatgaagaagaacaatcatgaaatggaagaaggaggaagagcatgaaggatcttataagaaatgcaagaggatattgataggtgtgaactagt  4800

4801  tcatcaagtagcccaagtaggagagaatcataggcaaaagtttctttaaagtggcagttgattaacatggaaggggaaatatgatagatatataaggaccctccacctgcccgggcggcc  4920

4921  gctc
```

Fig. 1 (*continued*).

A. Both proteins have about 44% identical and 62% similar amino acids [10], a corresponding size of 1547 and 1542 (PAPP-E) amino acids, a conserved pattern of cysteine residues, and a homologous domain structure (Fig. 2). Therefore, we named the novel sequence 'pregnancy-associated plasma protein-E' (PAPP-E). (In the 1970s, the terms PAPP-B, -C, and -D were assigned to a 1300 kDa octadecamer protein, the pregnancy-specific β_1-glycoprotein and the placental lactogen, which are not related with PAPP-A and PAPP-E [11–13].)

According to the PAPP-A motifs [5], the PAPP-E sequence comprises a putative prepropeptide, a putative zinc-binding site and Met-turn, five motifs that are related to short consensus repeats of complement proteins and selectins, and three motifs that are related to the lin-notch motifs of proteins regulating early tissue differentiation (Fig. 3). For example, proteins showing significant partial homologies with PAPP-E are P-, E-, and L-selectins (CD62), complement decay-accelerating factor (CD55), complement receptor types 1 and 2, complement factor H, neurogenic locus notch (homolog) proteins, *Caenorhabditis elegans* lin-12 protein, and the catalytic domains of

Table 1
Comparison of PAPP-A and PAPP-E protein and genetic parameter

Parameter	PAPP-A	PAPP-E
Amino acids	1547/80	1542/82
Cysteines	82/2	86/0
M_r (kDa) without glycosylation	172.3/8.7	171.1/9.4
Putative glycosylation sites	14×Asn; 7×Ser	15×Asn; 12×Ser
Theoretical p*I*	5.4/12.3	5.0/10.7
Predominant expression	Placenta	Placenta
Chromosomal localization	9q33.1	1q23–25

Values are mature protein/preproprotein. Asn, putative *N*-glycosylation site; Ser, putative sites for the attachment of glycosaminoglycans. PAPP-A data are according to [5].

```
                             ▨                                            ◆               ▨
E: 1       EAETFNSQVGLP--ILYFSGRRERLLLRPEVLAEIPREAFTVEAWVKPEGGQNNPAIIAGVFDNCSHTVSDKGWALGIRSGKDKGKRDARFFFSLCTDRVKKATILISHSRYQPGTWTHV
           EA         P   LYFSGR E+L +    L E+PR+AFT++ W++ EGGQ +PA+I G++D CS+   D+GW +GI +  D+  +D R+FFSL TDR ++ T + +H  Y PG W ++
A: 1       EARGATEEPSPPSRALYFSGRGEQLRVLRADL-ELPRDAFTLQVWLRAEGGQRSPAVITGLYDKCSYISRDRGWVVGIHTISDQDNKDPRYFFSLKTDRARQVTTINAHRSYLPGQWVYL

                                   ▨
E: 119     AATYDGRHMALYVDGTQVASSLDQSGPLNSPFMASCRSLLLGGDSSEDGHYFRGHLGTLVFWSTALPQSHFQHSSQHSSEEEEATDLVLTASFEPVNTEWVPFRDEKYPRLEV--LQGFE
           AATYDG+ M LYV+G QVA+S +Q G + SP    C+ L+LGG +    H +RG++     W  A  Q                  L+L  +++ V   W P +D   P++E     GF
A: 120     AATYDGQFMKLYVNGAQVATSGEQVGGIFSPLTQKCKVLMLGGSALN--HNYRGYIEHFSLWKVARTQREILSDMETHGAHTALPQLLLQENWDNVKHAWSPMKDGSSPKVEFSNAHGFL

                                                                                       ◆           ◆           LNR-1
E: 237     PEPEILSPLQPPLCGQTVCDNVELISQYNGYWPLRGEKVIRYQVVNICDDEGLNPIVSEEQIRLQHEALNEAFSRYNISWQLSVHQVHNSTLRHRVVLVNCEPSKIGNDHCDPECEHPLT
            +    + L+PPLCGQT+CDN E+I+ YN     R  KV+RY+VVN+ +D+  NP V+ EQ+  QH  L EAF +YNISW+L V +V NS+LR R++L NC+ SKIG+++CDPEC H LT
A: 238     LD----TSLEPPLCGQTLCDNTEVIASYNQLSSFRQPKVVRYRVVNLYEDDHKNPTVTREQVDFQHHQLAEAFKQYNISWELDVLEVSNSSLRRRLILANCDISKIGDENCDPECNHTLT

                 LNR-2                                                         ◆
E: 357     GYDGGDCR-LQGRCYSWNRRDGLCHVECNNMLNDFDDGDCCDPQVADVRKTCFDPDSPKRAYMSVKELKEALQLNSTHFLNIYFASSVREDLAGAATWPWDKDAVTHLGGIVLSPAYYGM
           G+DGGDCR L+   +   + +G+C ++CN    +FD G+CCDP++ +V +TCFDPDSP RAY+ V ELK  L+L+ +  LNI+FA S  E+LAG ATWPWDK+A+ HLGGIVL+P++YGM
A: 354     GHDGGDCRHLRHPAFVKKQHNGVCDMDCNYERFNFDGGECCDPEITNVTQTCFDPDSPHRAYLDVNELKNILKLDGSTHLNIFFAKSSEEELAGVATWPWDKEALMHLGGIVLNPSFYGM

                   Zn2+                                                                            ◆
E: 476     PGHTDTMIHEVGHVLGLYHVFKGVSERESCNDPCKETVPSMETGDLCADTAPTPKSELCREPEPTSDTCGFTRFPGAPFTNYMSYTDDNCTDNFTPNQVARMHCYLDLVYQQWTESRKPT
           PGHT TMIHE+GH LGLYHVF+G+SE +SC+DPC ET PS ETGDLC DT P PK + C +P P +DTCGF  F   P+ N+MSY DD+CTD+FTPNQVARMHCYLDLVYQ W  SRKP
A: 474     PGHTHTMIHEIGHSLGLYHVFRGISEIQSCSDPCMETEPSFETGDLCNDTNPAPKHKSCGDPGPGNDTCGFHSFFNTPYNNFMSYADDDCTDSFTPNQVARMHCYLDLVYQGWQPSRKPA

                       ◆           ▨       ▨                            ▨                                  ◆
E: 596     PTPIPPMVIGQTNKSLTIHWLPPISGVVYDRASGSLCGACTEDGTFRQYVHTASSRRVCDSSGYWTPEEAVGPPDVDQPCEPSLQAWSPEVHLYHMNMTVPCP-TEGCSLELLFQHPVQA
           P  + P V+G T  S+T+ W PPI G  ++R  GS C  C E     QY   ASS   C  SG+W+P EA G PDV+QPC+ S++ WSP   +    +   CP  +GC LEL F +P+
A: 594     PVALAPQVLGHTTDSVTLEWFPPIDGHFFERELGSACHLCLEGRILVQYASNASSPMPCSPSGHWSPREAEGHPDVEQPCKSSVRTWSPNSAVNPHTVPPACPEPQGCYLELEFLYPLVP

                                                                    ▨                           ▨                ▨
E: 715     DTLTLWVT--SFFMESSQVLFDTEILLENKESVHLGPLDTFCDIPLTIKL-HVDGKVSGVKVYTFDERIEIDAALLTSQPHSPLCSGCRPVRYQVLRDPPFASGLPVVVTHSHRKFTDVE
           ++LT+WVT  S   +SS  + D ++L  + +++ LGP + FCD+PLTI+L  V  +V G+++YT DE +EIDAA+LTS   +PLC  C+P++Y+V+RDPP    +  ++ H +RKF D++
A: 714     ESLTIWVTFVSTDWDSSGAVNDIKLLAVSGKNISLGPQNVFCDVPLTIRLWDVGEEVYGIQIYTLDEHLEIDAAMLTSTADTPLCLQCKPLKYKVVRDPPLQMDVASIL-HLNRKFVDMD

                                                                 ▨
E: 832     VTPGQMYQYQVLAEAGGELGEASPPLNHIHGAPYCGDGKVSERLGEECDDGDLVSGDGCSKVCELEEGFNCVGEPSLCYMYEGDGICEPFERKTSIVDCGIYTPKGYLDQWATRAYSSHE
           +  G +YQY V+  +G E  E SP + +IHG  YCGDG + +  GE+CDD + ++GDGCS  C  E  FNC+ EPS CY ++GDG+CE FE+KTSI DCG+YTP+G+LDQWA+ A  SH+
A: 833     LNLGSVYQYWVITISGTEESEPSPAVTYIHGRGYCGDGIIQKDQGEQCDDMNKINGDGCSLFCRQEVSFNCIDEPSRCYFHDGDGVCEEFEQKTSIKDCGVYTPQGFLDQWASNASVSHQ

                                                      ◆                                                                ◆
E: 952     DKKKCPVSLVTGEP-HSLICTSYHPDLPNHRPLTGWFPCVASENETQDDRSEQPEGSLKKEDEVWLKVCFNRPGEARAIFIFLTTDGLVPGEHQQPTVTLYLTDVRGSNHSLGTYGLSCQ
           D ++CP  ++ G+P  S +C +   DL        W+PC  S   +Q             +   WL+  F++P  A A+ + L TDG   G+ +Q T+++ L D +  +H LG + LSC+
A: 953     D-QQCPGWVIIGQPAASQVCRTKVIDLSEGISQHAWYPCTISYPYSQ-----------LAQTTFWLRAYFSQPMVAAAVIVHLVTDGTYYGDQKQETISVQLLDTKDQSHDLGLHVLSCR

                       ◆           ◆                                                   SCR-1       ◆                         ▨
E: 1071    HNPLIINVTHHQNVLFHHTTSVLPNFSSPRVGISAVALRTSSRIGLSAPSNCISEDEGQNHQGQSCIHRPCGKQDSCPSLLLDHADVVNCTSIGP-GLMKCAITCQRGFALQASSGQYIR
           +NPLII V H  +  F+H+ +V  +FSSP V IS VALR+         S+C    E  +   QSC+H  C K D CP L +++A  +NC+S       +C ++C+ G+ LQ       +
A: 1061    NNPLIIPVVHDLSQPFYHSQAVRVSFSSPLVAISGVALRSFDNFDPVTLSSC-QRGETYSPAEQSCVHFACEKTD-CPELAVENAS-LNCSSSDRYHGAQCTVSCRTGYVLQIRRDDELI

                        ▨     ◆       SCR-2       ◆                                                SCR-3
E: 1190    PMQ--KEILLTCSSGHWDQNVSCLPVDCGVPDPSLVNYANFSCSEGTKFLKRCSISCVPPAKLQGLSPWLTCLEDGLWSLPEVYCKLECDAPPIILNANLLLPHCLQDNHDVGTICKYEC
             Q    + +TC+ G W++ V+C PVDC +PD   V  A+FSC EGT F  +CS  C  PA+L+G +  LTC+EDGLWS PE  C+L C APP + NA+L    C ++ H VG+ CKY+C
A: 1178    KSQTGPSVTVTCTEGKWNKQVACEPVDCSIPDHHQVYAASFSCPEGTTFGSQCSFQCRHPAQLKGNNSLLTCMEDGLWSFPEALCELMCLAPPPVPNADLQTARCRENKHKVGSFCKYKC

                                                 SCR-4                                                           SCR-5
E: 1308    KPGYYVAESAEGKVRNKLLKIQCLEGGIWEQGSCIPVVCEPPPPVFEGMYECTNGFSLDSQCVLNC-----NQEREKLPILCTKEGLWTQEFKLCENLQGECPPPPSELNS-VEYKCEQG
           KPGY+V  S+  K + +  K QC + G W++G+C+PV C+PPPP F G+Y+CTNGF  +S+C + C     +Q      I C K+G W   F +C+ +QG+C   P+ELNS ++ +C  G
A: 1298    KPGYHVPGSSR-KSKKRAFKTQCTQDGSWQEGACVPVTCDPPPPKFHGLYQCTNGFQFNSECRIKCEDSDASQGLGSNVIHCRKDGTWNGSFHVCQEMQGQC-SVPNELNSNLKLQCPDG

                                  ◆                                      LNR-3
E: 1422    YGIGAVCSPLCVIPPSDPVMLPENITADTLEHWMEPVKVQSIVCTGRRQWHPDPVLVHCIQSCEPFQANGWCDTINNRAYCHYDGGDCCSSTLSSKKVIPFAADCDLD-ECTCRDPKAEE
           Y IG+ C+  C+   S+ ++LP N+T   + HW+ P +V+ +VCT   +W+P P L+HC++ CEPF  + +CD INNRA+C+YDGGDCC+ST+ +KKV PF   CDL  +C CRDP+A+E
A: 1416    YAIGSECATSCLDHNSESIILPMNVTVRDIPHWLNPTRVERVVCTAGLKWYPHPALIHCVKGCEPFMGDNYCDAINNRAFCNYDGGDCCTSTVKTKKVTPFPMSCDLQGDCACRDPQAQE


PAPP-E:  1541 N
              +
PAPP-A:  1536 H
```

Fig. 2. Alignment of PAPP-E and PAPP-A sequences and putative motifs. LNR, lin-notch repeat; SCR, short consensus repeat; Zn^{2+}, zinc-binding site; ■, *N*-glycosylation site Asn–Xaa–Ser/Thr; ◆, site for attachment of glycosaminoglycans.

```
Short consensus repeats (SCRs):

SCR-1 (PAPP-A):  C--PELAVENAS-LN-CSSSDRYHGAQCTVSCRTGYVLQIRRDDELIKSQTGPSVTVT------CT-EGK-WNKQVA---C
SCR-1 (PAPP-E):  C--PSLLLDHADVVN-CTSIGP-GLMKCAITCQRGFALQASSGQYIRPMQ--KEILLT------CSS-GH-WDQNVS---C
SCR-2 (PAPP-A):  CSIPDHHQVYAASFS-CPEGTT-FGSQCSFQCRHPAQLKGNNSLLT------------------CMEDGL-WSFPEAL--C
SCR-2 (PAPP-E):  CGVPDPSLVNYANFS-CSEGTK-FLKRCSISCVPPAKLQGLSPWLT------------------CLEDGL-WSLPEVY--C
SCR-3 (PAPP-A):  CLAPPPVPNADLQTARCRENKHKVGSFCKYKCKPGYHVPGSSR-KSKKRAFKTQ----------CTQDGS-W-QEGA---C
SCR-3 (PAPP-E):  CDAPPIILNANLLLPHCLQDNHDVGTICKYECKPGYYVAESAEGKVRNKLLKIQ----------CLEGGI-WEQGS----C
SCR-4 (PAPP-A):  CDPPPPKFHG-LYQ--CTNGFQFN-SECRIKCEDSDASQGLGSNVIH-----------------CRKDGT-WNGSFHV--C
SCR-4 (PAPP-E):  CEPPPPVFEG-MYE--CTNGFSLD-SQCVLNC-----NQEREKLPIL-----------------CTKEGL-WTQEFKL--C
SCR-5 (PAPP-A):  CSVP-NELNSNLKLQ-CPDGYAI-GSECATSCLDHNSESIILPMNVTVRDIPHWLNPTRVERVVCTA-GLKWYPHPALIHC
SCR-5 (PAPP-E):  CPPPPSELNS-VEYK-CEQGYGI-GAVCSPLCVIPPSDPVMLPENITADTLEHWMEPVKVQSIVCTGRRQ-WHPDPVLVHC

                                                                                    D
PAPP-A SCRs   :  C--P------------C-------GS-C---C--------------------------------C--EG--W----A---C


Lin-notch repeats (LNRs):

LNR-1 (PAPP-A):  CDISKIGDEN-------CDPE-----CNHTLTGHDGGD-
LNR-1 (PAPP-E):  CEPSKIGNDH-------CDPE-----CEHPLTGYDGGD-
LNR-2 (PAPP-A):  CRHLRHPAFVKKQHNGVCDMD-----CNYERFNFDGGEC
LNR-2 (PAPP-E):  CR-LQGRCYSWNRRDGLCHVE-----CNNMLNDFDDGDC
LNR-3 (PAPP-A):  CEPFMGDNY--------CDAINNRAFCNY-----DGGDC
LNR-3 (PAPP-E):  CEPFQANGW--------CDTINNRAYCHY-----DGGDC

                                             H        E
PAPP-A LNRs   :  C----------------CD-------CNY-----DGGDC


Zinc-binding motifs:

PAPP-A:         HEIGHSLGLYH
PAPP-E:         HEVGHVLGLYH

Metzincins:     HE-GH--G--H
```

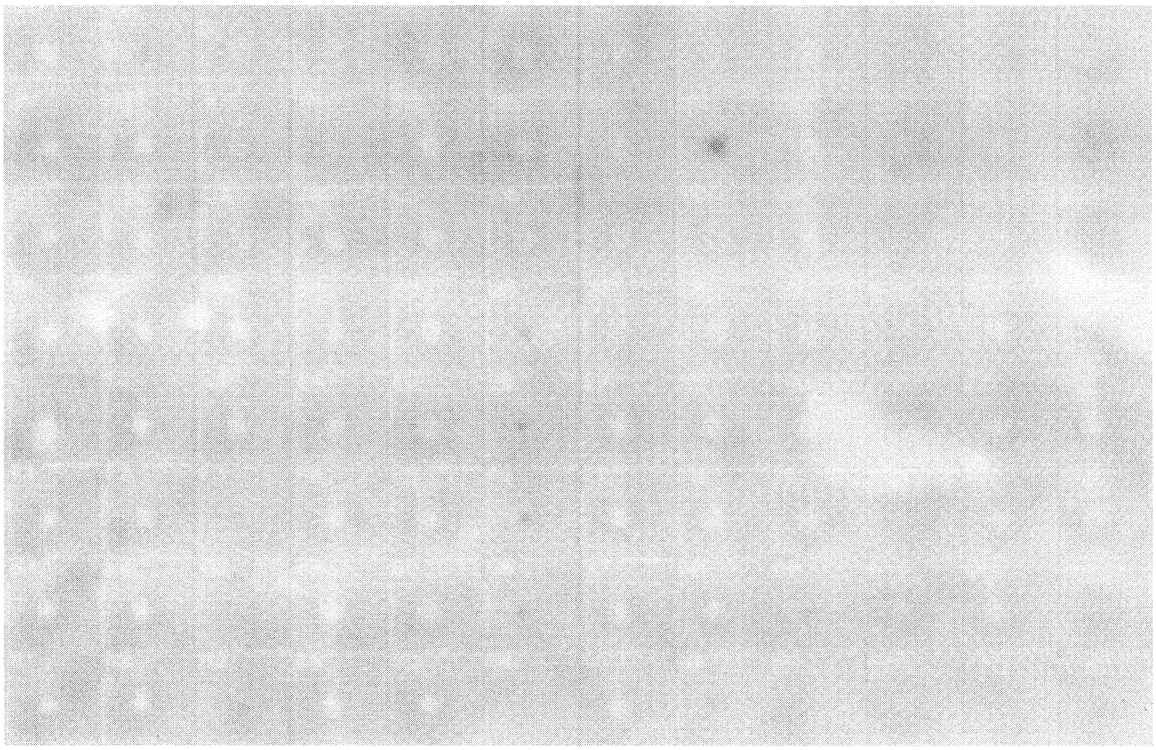
Fig. 3. Alignment of PAPP-A and PAPP-E consensus sequences. PAPP-A motifs are according to [5].

metzincins [9]. Primary structure analysis with the ExPASy proteomics tools shows that corresponding to PAPP-A, PAPP-E has a molecular mass (without glycosylation) of about 170 kDa, several putative sites for *N*-glycosylation and the attachment of glycosaminoglycan groups (Fig. 2 and Table 1), and an acidic theoretical isoelectric point (p*I*) of mature protein and a basic p*I* of prepropeptide (Table 1).

For Northern blot analysis, the 315 bp PAPP-E cDNA fragment coding for the putative active site was biotinylated with the North2South Biotin Random Prime kit (Pierce) and used for hybridization with the MTE Array 2 (Clontech), which represents expression of 75 human tissues. Chemiluminescent detection was performed with a streptavidin–horseradish peroxidase conjugate and luminol peroxidase reaction using North2South Chemiluminescent Hybridization and Detection (Pierce) and CL-XPosure X-ray film (Pierce) with Eukobrom paper developer and Superfix fixing solution (Tetenal). The only significant signal was produced by placenta mRNA (Fig. 4), so that placenta seems to be the main source of PAPP-E expression. This correlates with the predominant expression of PAPP-A by the trophoblastic tissue of the placenta [14] and defines both proteins as pregnancy-associated proteins.

To clarify the chromosomal localization of the PAPP-E gene, radiation hybrid mapping [15] with the human/hamster Genebridge 4 RH Panel (Research Genetics) and the oligonucleotides 5′-agcccagcatattatgggatgcctggcc-3′ and 5′-ggtgtcggcacagaggtctcccgtttcc-3′ for PCR with *Taq* DNA polymerase (Qiagen) was performed. Analysis of the resulting data vector shows that the PAPP-E gene is localized between the markers D1S242 and AFM210WC11 which are in the q23–25 region of chromosome 1 (Fig. 5).

In conclusion, we have cloned a novel human cDNA encoding a PAPP-A homologous protein that has been named pregnancy-associated plasma protein-E (PAPP-E). Both proteins share the same domain structure and probably form a new subfamily of metzincins. The PAPP-E gene has been assigned to chromosome 1 and is predominantly expressed in placenta. The biological role of PAPP-E regarding its putative proteolytic and regulatory functions and involvement in pregnancy and trisomy re-



whole brain		substantia nigra	heart	esophagus	colon, transverse	kidney	lung	liver	leukemia, HL-60	fetal brain	yeast total RNA
cerbral cortex	cerebellum, right	nucleus accumbens	aorta	stomach	colon, descending	skeletal muscle	placenta	pancreas	HeLa S3	fetal heart	yeast tRNA
frontal lobe	corpus callosum	thalamus	atrium, left	duodenum	rectum	spleen	bladder	adrenal gland	leukemia, K-562	fetal kidney	E. coli tRNA
parietal lobe	amygdala	pituitary gland	atrium, right	jejunum		thymus	uterus	thyroid gland	leukemia, MOLT-4	fetal liver	E. coli DNA
occipital lobe	caudate nucleus	spinal cord	ventricle, left	ileum		peripheral blood leukocyte	prostate	salivary gland	Burkitt's lymphoma, Raji	fetal spleen	poly r(A)
temporal lobe	hippo-campus		ventricle, right	ileocecum		lymph node	testis	mammary gland	Burkitt's lymphoma, Daudi	fetal thymus	human C_0t-1 RNA
paracentral gyrus of cerebr. cort.	medulla oblongata		inter-ventricular septum	appendix		bone marrow	ovary		colorectal adeno-carcinoma, SW480	fetal lung	human DNA 100 ng
pons	putamen		apex of the heart	colon, ascending		trachea			lung carcinoma, A549		human DNA 500 ng

Fig. 4. Northern blot analysis of PAPP-E mRNA. Chemiluminescent detection of a 75-tissue expression array (Clontech) hybridized with a biotinylated PAPP-E cDNA fragment.

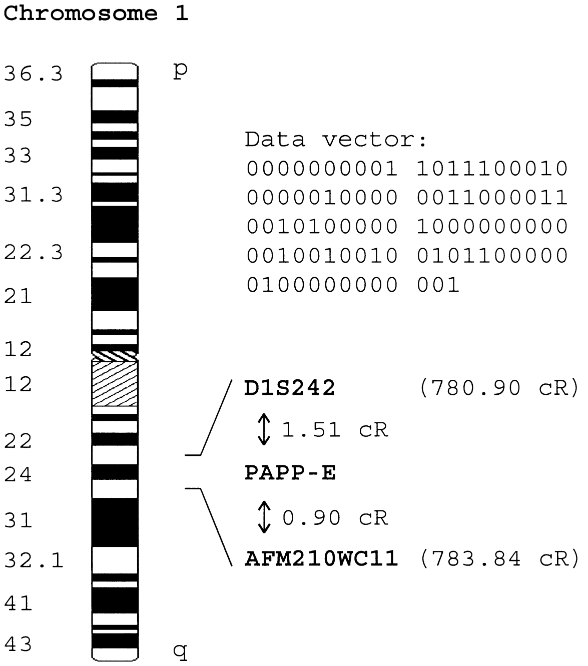


Fig. 5. Chromosomal assignment of the PAPP-E gene to chromosome 1. The data vector for the radiation hybrid mapping (Genebridge 4 panel), linkage to neighboring markers in centirad (cR), and the corresponding cytogenetic location are shown.

mains to be analyzed. The recombinant expression for biochemical characterization of the new human protein is in progress.

We thank Herbert Wenzel and Rainer Beckmann for scientific and technical assistance. Part of this work was supported by the Special Collaborative Programme SFB 549 and its service project Z3 (Thomas Schmitt-John and Harald Jockusch).

References

[1] C. Oxvig, O. Sand, T. Kristensen, G.J. Gleich, L. Sottrup-Jensen, Circulating human pregnancy-associated plasma protein-A is disulfide-bridged to the proform of eosinophil major basic protein, J. Biol. Chem. 268 (1993) 12243–12246.

[2] J.G. Westergaard, B. Teisner, J.G. Grudzinskas, Serum PAPP-A in normal pregnancy: relationship to fetal and maternal characteristics, Arch. Gynecol. 233 (1983) 211–215.

[3] N.J. Wald, L. George, D. Smith, J.W. Densem, K. Petterson, Serum screening for Down's syndrome between 8 and 14 weeks of pregnancy, International Prenatal Screening Research Group, Br. J. Obstet. Gynaecol. 103 (1996) 407–412.

[4] J.B. Lawrence, C. Oxvig, M.T. Overgaard, L. Sottrup-Jensen, G.J. Gleich, L.G. Hays, J.R. Yates, C.A. Conover, The insulin-like growth factor (IGF)-dependent IGF binding protein-4 protease secreted by human fibroblasts is pregnancy-associated plasma protein-A, Proc. Natl. Acad. Sci. USA 96 (1999) 3149–3153.

[5] T. Kristensen, C. Oxvig, O. Sand, N.P. Møller, L. Sottrup-Jensen, Amino acid sequence of human pregnancy-associated plasma protein-A derived from cloned cDNA, Biochemistry 33 (1994) 1592–1598.

[6] W. Bode, F. Grams, P. Reinemer, F.X. Gomis-Rüth, U. Baumann, D.B. McKay, W. Stöcker, The metzincin-superfamily of zinc-peptidases, Adv. Exp. Med. Biol. 389 (1996) 1–11.

[7] M.K. Pangburn, K.L. Pangburn, V. Koistinen, S. Meri, A.K. Sharma, Molecular mechanisms of target recognition in an innate immune system: interactions among factor H, c3b, and target in the alternative pathway of human complement, J. Immunol. 164 (2000) 4742–4751.

[8] S. Artavanis-Tsakonas, M.D. Rand, R.J. Lake, Notch signaling: cell fate control and signal integration in development, Science 284 (1999) 770–776.

[9] S.F. Altschul, T.L. Madden, A.A. Schäffer, J. Zhang, Z. Zhang, W. Miller, D.J. Lipman, Gapped BLAST and PSI-BLAST: a new generation of protein database search programs, Nucleic Acids Res. 25 (1997) 3389–3402.

[10] T.A. Tatusova, T.L. Madden, BLAST 2 Sequences, a new tool for comparing protein and nucleotide sequences, FEMS Microbiol. Lett. 174 (1999) 247–250.

[11] H. Bohn, Detection and characterization of pregnancy proteins in the human placenta and their quantitative immunochemical determination in sera from pregnant women, Arch. Gynakol. 210 (1971) 440–457.

[12] T.M. Lin, S.P. Halbert, W.N. Spellacy, S. Gall, Human pregnancy-associated plasma proteins during the postpartum period, Am. J. Obstet. Gynecol. 124 (1976) 382–387.

[13] M. Bossi, K.H. Winterhalter, N.A. Bersinger, Pregnancy-associated plasma protein B (PAPP-B): Isolation from human late pregnancy serum and biochemical characterization, Biochem. Mol. Biol. Int. 29 (1993) 93–102.

[14] D. Tornehave, J. Chemnitz, B. Teisner, J. Folkersen, J.G. Westergaard, Immunohistochemical demonstration of pregnancy-associated plasma protein A (PAPP-A) in the syncytiotrophoblast of the normal placenta at different gestational ages, Placenta 5 (1994) 427–431.

[15] D.R. Cox, M. Burmeister, E.R. Price, S. Kim, R.M. Myers, Radiation hybrid mapping: a somatic cell genetic method for constructing high-resolution maps of mammalian chromosomes, Science 250 (1990) 245–250.