Volume 262, number 2, 349–352 FEBS 08287 March 1990

A genetic polymorphism in a functional domain of human pregnancy zone protein: the bait region

Genomic structure of the bait domains of human pregnancy zone protein and α_2 macroglobulin

P. Marynen, K. Devriendt, H. Van den Berghe and J.J. Cassiman

Center for Human Genetics, University of Leuven, Campus Gasthuisberg, Herestraat 49, B-3000, Leuven, Belgium

Received 24 January 1990

Genomic clones containing the exons coding for the bait domain of human pregnancy zone protein and α_2 macroglobulin were isolated and fragments containing the bait exons were sequenced. It is shown that the bait domains of both α_2 macroglobulin and pregnancy zone protein are encoded by two exons, with conserved exon/intron boundaries. A genetic polymorphism showing either a Met or a Val residue as the sixth amino acid of the pregnancy zone protein bait domain was detected with the rare Met allele showing a gene frequency of 0.065.

Proteinase inhibitor; Macroglobulin, α-

1. INTRODUCTION

The α macroglobulin proteinase inhibitor family is represented in humans by α_2 macroglobulin (α_2M) and pregnancy zone protein (PZP). Unlike α_2M which is constitutively present at high levels in plasma, PZP levels, normally very low in both males and females (approx. 10 μg/ml), rise dramatically during pregnancy to reach 2 mg/ml at term [1]. Proteinase cleavage of a specific 'bait' domain of α macroglobulins, located near the middle of the polypeptide chains of the macroglobulin subunits, leads to a sequence of events resulting eventually in the physical entrapment of the proteinase, which is then no longer able to reach high molecular weight substrates such as proteins [2]. The resulting α macroglobulin-proteinase complexes are efficiently cleared from the circulation by receptor-mediated endocytosis [3]. The spectrum of proteinases inhibited by each particular α macroglobulin is therefore defined by the amino acid (aa) sequence of its bait domain, which must reflect the specificity of the inhibited proteinases. α Macroglobulins are highly conserved proteins, except for the bait domains which vary widely in length and aa sequence [4], and different α macroglobulins thus inhibit different (complementary) sets of proteinases. The suggestion was made that the occurrence of varying bait domains might be the result of exon shuffling and positive Darwinian selection [4].

The aa sequence of a cyanogen bromide fragment spanning the PZP bait domain and the sites of cleavage by several proteinases within the bait was published recently [4]. The isolation of the two exons coding for the PZP bait domain is reported here. Sequencing of this clone revealed a rare Val⟶Met polymorphism in this important functional domain of PZP. An assay system, relying on allele-specific oligonucleotides was developed allowing for the rapid detection of both PZP bait domain alleles. To evaluate the possible genetic origin of the diversity of different bait domains, the exon structure of the human α_2 macroglobulin bait domain was determined and compared to the genomic structure of the PZP bait.

2. MATERIALS AND METHODS

The PZP clone was isolated from a human placental DNA genomic library, cloned into the cosmid vector pWE15 and obtained from Stratagen (La Jolla, CA, USA). The α_2M clone was isolated using an α_2M cDNA probe [5] from a human genomic library cloned into Charon 4a (courtesy of Dr J. Darby, Stanford University, CA, USA). The screening, clone purification, restriction mapping and subcloning into pGEM.3Z (Promega, Madison, USA) were done according to established procedures. Sequencing was performed with T7 DNA polymerase (Pharmacia, Uppsala) with supercoiled plasmid as a template. Random shotgun cloning was used to obtain the sequence of both strands of the inserts.

Correspondence address: P. Marynen, Center for Human Genetics, University of Leuven, Campus Gasthuisberg, Herestraat 49, B-3000 Leuven, Belgium

The nucleotide sequence(s) presented here has (have) been submitted to the EMBL/GenBank database under the accession number no.Y07535

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/90/$3.50

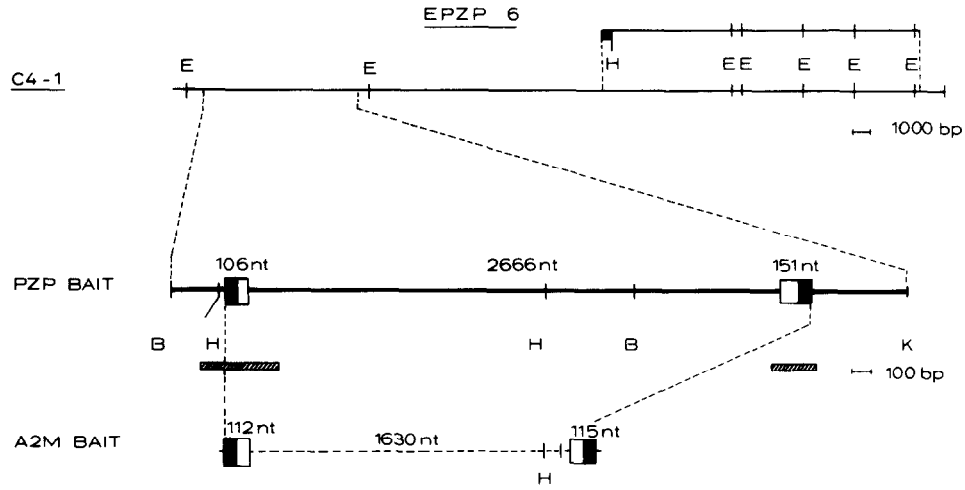

Fig.1. Schematic representation of the genomic PZP and α_2M clones. The newly isolated PZP clone (C4-1) is shown in relation to the previously described [5] PZP clone EPZP6. The fragment containing the two PZP bait exons is shown enlarged (PZP BAIT). E, *Eco*R; H, *Hind*; B, *Bgl*II; K, *Kpn*I site. The exons are illustrated as boxes, which are shaded for residues conserved between PZP and α_2M and open for bait domain residues. The crosshatched lines delimit the sequences shown in fig.2. Bottom, the structure of the α_2M clone containing the two bait exons is shown.

The forward primer (5′ TAC TGT CAA GTG TTT CGC CA) and reverse primer (5′ AGC TAT CAC CAG CTA AGT CA) for the polymerase chain reactions (PCR) were synthesized on a Cyclone (Biosearch, San Raphael, CA) DNA synthesizer and used without further purification. PCR was performed with 1 μg human genomic DNA in 50 mM KCl, 10 mM Tris-HCl, pH 8.3, 1.5 mM $MgCl_2$, 0.01% gelatinc, 100 pmol primers, 200 μM dNTPs and 2.5 units of Taq DNA polymerase (Perkin-Elmer Cetus, Norwalk, CA). Twenty-five amplification cycles (1 min at 94°C, 2 min at 60°C, 1 min at 72°C) were performed on a PCR thermal cycler (Perkin Elmer-Cetus); each experiment included negative controls. Two allele-specific oligonucleotides, ASO PZP BAIT-Val (5′ CCC TTC CGT

```
PZP-[tactgtcaagtgtttcgcca]taaaattaatataaggatactggaattacatagagaagtctctaaggataatatgttttatatattctaa     90
PZP-tcaagcttacaagtagattgaacaattatttttttccagGGGATGGGATTGAAGGTGTTCACTAACTCAAAAATCCGAAAACCAAAGTCG     180
                                  PZP -      G  M  G  L  K  V  F  T  N  S  K  I  R  K  P  K  S
                                             |  |  |  |     |  |  |  |  |  |  |  |  |  |  |
                                  A2M - 649- D  M  G  L  K  A  F  T  N  S  K  I  R  K  P  K  M
                                A2M-ttttgtttgcagGACATGGGCTTAAAGGCATTCACCAACTCAAAGATTCGTAAACCCAAAATG


PZP-TGTTCAGTCAT[CCCTTCCATGTCTGCA]GGAGCAGTAGGTCAAGGATACTATGGAGgtaaaacagcaatttgcttttaattttatttcatg     270
PZP-C  SerValIleProSerMetSerAlaGlyAlaValGlyGlnGlyTyrTyrGlyA/
   |
A2M-C  ProGlnLeuGlnGlnTyrGluMetHisGlyProGluGlyLeuArgValGlyPheTyrG/-686
A2M-TGTCCACAGCTTCAACAGTATGAAATGCATGGACCTGAAGGTCTACGTGTAGGTTTTTATGgtaaacaaaaa----------


PZP-tccaacaatattactatgcccaaaatgctaaaatgaggtaatgtagccctatttccactttatacacacaaagaaaccaagattcaaatt     360
PZP-aattgtttcttgctgagg<tgacttagctggtgatagct>----------2493 nt--------------------------------
PZP-cttactttagCAGGTCTAGGAGTAGTAGAGAGACCATATGTTCCTCAATTAGGCACATATAATGTGATACCCTTAAATAATGAACAAAGT     2951
       PZP - laGlyLeuGlyValValGluArgProTyrValProGlnLeuGlyThrTyrAsnValIleProLeuAsnAsnGluGlnSer

                                                A2M - 686-luSerAspValMetGlyArgGlyHisAlaArgLeuValHisVal
A2M----------------------1610 nt-------ctcaccatagAGTCAGATGTAATGGGAAGAGGCCATGCACGCCTGGTGCATGTT


PZP-TCAGGGCCAGTCCCTGAAACGGTGCGAAGCTATTTTCCTGAGACTTGGATCTGGGAGTTGGTGGCAGTGAAgtaagtaact     3032
PZP-SerGlyProValProE  T  V  R  S  Y  F  P  E  T  W  I  W  E  L  V  A  V  N/
                      |  |  |  |     |  |  |  |  |  |  |  |  |  |     |  |
A2M-GluGluProHisThrE  T  V  R  K  Y  F  P  E  T  W  I  W  D  L  V  V  V  N/-724
A2M-GAAGAGCCTCACACGGAGACCGTACGAAAGTACTTCCCTGAGACATGGATCTGGGATTTGGTGGTGGTAAAgtaagtaact
```

Fig.2. Sequence of PZP and α_2M bait exons. DNA sequences of both bait exons of PZP Met allele and α_2M are shown. Intron sequences are shown in lower case, exon sequences in upper case. The conserved protein sequences are given in single letter amino acid code, the bait domain sequences are given in three letter code. The numbering of α_2M amino acids is as in [4]. The sequences of the PCR primers are shown by arrows, the sequence of the ASO is boxed, the polymorphic Met residue is underlined. The PZP gene fragment was fully sequenced. The full sequence is submitted to Genbank (CA), access number Y07535. The intronic distance between the two α_2M bait domain exons was determined with restriction mapping.

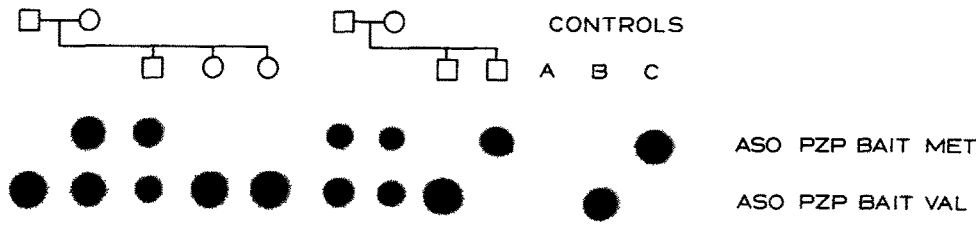

Fig.3. Mendelian inheritance of the PZP bait domain polymorphism. PCR products obtained from DNA of the different family members was denatured and dot spotted onto nylon membranes. The membranes were then hybridized to ^{32}P-labeled ASOs specific for the Met allele (upper row) or the Val allele (lower row). The controls are A: no genomic DNA added to the PCR reaction; B: 100 ng of a plasmid carrying the Val allele and C: 100 ng of a plasmid carrying the Met allele.

GTC TGC A) and ASO PZP BAIT-Met (5′ CCC TTC CAT GTC TGC A) were labeled with the polynucleotide kinase and hybridized (2000000 dpm/ml) to dot blots in 5 × SSPE, 0.2% SDS, 200 μg/ml heparin at 37°C. Allele-specific hybridization was observed on autoradiography after a final wash in 2 × SSPE, 0.2% SDS for 5 min at 50°C for ASO PZP BAIT-Met and 55°C for ASO PZP BAIT-Val.

3. RESULTS

We previously reported the characterization of a genomic PZP clone (EPZP6) spanning the 3′ end of the PZP gene [6]. With a probe derived from the 5′ end of EPZP6, a cosmid clone C4-1 was isolated reaching 21.5 kb upstream of EPZP6 (fig.1). Sequencing of a 3.7 kb fragment of C4-1, and comparison with the published partial aa sequence of PZP [7], revealed the presence of 2 exons coding for 35 and 51 aa, respectively (figs 1 and 2). Inspection of this sequence shows that the PZP bait domain, as defined by Sottrup-Jensen et al. [4], spans from aa 17 of the first exon to aa 33 of the second exon. Interestingly, the DNA sequence predicts a Met residue at position 6 of the PZP bait domain instead of a Val as published. In view of the possible significance of a mutation in this functional domain of PZP, this observation was further investigated. First, primers were designed to amplify the first exon of the PZP bait (fig.2) with the PCR. The bait exon 1 of the DNA of 8 unrelated individuals was amplified, cloned into *Sma*I-cut pGEM3Z and 10 independent clones for each individual were sequenced. Seven individuals showed only GTG (Val) as the 21th codon of bait exon 1. The 8th individual showed a GTG (Val) codon in 4 clones and an ATG (Met) codon in 6 clones, suggesting heterozygosity at this locus. To exclude PCR or cloning artifacts, the PCR, cloning and sequencing were repeated for individual 8 on another DNA sample, with identical results. Second, two allele-specific oligonucleotides (ASO PZP BAIT-Val and ASO PZP BAIT-Met) were designed and the hybridization and wash conditions for the specific detection of the Met and the Val allele were determined (section 2). The DNA samples of 32 unrelated individuals, parents of small nuclear families, were amplified and assayed. 28 samples were homozygous for the GTG codon while 4 heterozygotes, belonging to 3 families, showed both GTG and ATG alleles. Mendelian inheritance of the Met/Val polymorphism in two representative families is shown in fig.3.

Exon shuffling was suggested as a possible mechanism by which different bait domains could have arisen [4]. As shown here, the PZP bait domain is encoded by two exons separated by a type I intron. In view of the possible implications as to the mechanism of divergence of the bait domains, the corresponding region of the α_2M gene was cloned, exons coding for the α_2M bait were identified and the intron/exon boundaries determined by sequencing (fig.2). Comparison of these data with the α_2M cDNA sequence [5] shows an identical genomic organization with two exons (37 aa and 39 aa, respectively) coding for the α_2M bait domain, separated by a type I intron (figs 1 and 2).

4. DISCUSSION

The data presented here demonstrate a polymorphism in the bait domain of PZP, the sixth amino acid of this domain being either a Val or a Met. Assuming a Hardy-Weinberg equilibrium, the frequency of the Met allele can be estimated to be 0.065, with 0.4% of the individuals being homozygotes. This could explain why the polymorphism was not detected by protein sequencing. Furthermore, the presence of a new Met residue would generate fragments of different length upon cyanogen bromide cleavage, separating the Val bait peptides from the Met bait fragments upon further purification and sequencing.

The α macroglobulins represent an interesting case of molecular evolution containing a bait domain defining their specificity as proteinase inhibitors, differing in length and aa sequence in an otherwise well-conserved polypeptide [2]. Bait domains of α macroglobulins were shown to be exposed at the surface of the molecule [8,9], but comparison of bait domains of different α macroglobulins failed to detect common structural motives in this domain [4]. It is therefore conceivable that a mutation in the bait domain would not affect the molecular mechanism by which α macroglobulins en-

trap proteinases, although it would of course affect the specificity of the inhibitor. This, and positive Darwinian evolution as suggested by L. Sottrup-Jensen et al. [4], could then account for the rapid evolution of the bait domain of different α macroglobulins. Exon shuffling has also been mentioned as a mechanism by which different baits might have evolved [4]. The present data seem not to support this view. Indeed, both the α_2M and the PZP bait domains are encoded by two exons with boundaries which do not coincide with the boundaries of the respective exons, while the 5′ end of the first exon and the 3′ end of the second occur at conserved aa residues in both genes. The intron/exon boundaries at the rat α_2M bait domain are also identical to the one reported here for human PZP and α_2M, showing cross-species conservation of this genomic organization [10]. The PZP and the human α_2M bait exon, and the rat α_2M bait exons are all interrupted by a type I intron which would argue against, although not exclude, intron sliding [11] as a mechanism by which the lengths of different bait domains could vary.

The physiology of PZP is not yet well understood and the polymorphism of the bait domain described here does not occur at the primary site of cleavage by the proteinases reported recently [4]. This set of proteinases, however, might be different from the proteinases against which PZP protects the organism. While low PZP levels were reported to be correlated with spontaneous abortion early in pregnancy [12], no correlation was found between PZP levels and severe pre-eclampsia [13]. As shown here, overall PZP levels, however, do not necessarily represent the proteinase inhibitory potential of PZP, which could be influenced by polymorphisms like the bait domain polymorphism reported here. The probes for rapid detection of this polymorphism, developed here, will allow us to investigate proteinase inhibition by the different PZP allotypes and their possible relation to PZP physiology in the future.

Acknowledgements: We thank Hilde Braeken for expert technical assistance, Staf Doucet for the synthesis of oligonucleotides and Karel Rondou for the photography. This work was supported by a grant 'Geconcerteerde acties' from the Belgian government and by the Interuniversity Network for Fundamental Research sponsored by the Belgian government (1987–1991). This work was also supported by Grant 3.0027.88 from the National Fund for Scientific Research, Belgium.

REFERENCES

[1] Sottrup-Jensen, L. (1987) in: The Plasma Proteins: Structure, Function and Genetic Control, vol.V (Putnam, F.W. ed.) pp.189–291, Academic Press, New York.
[2] Sottrup-Jensen, L. (1989) J. Biol. Chem. 264, 11539–11542.
[3] Van Leuven, F., Cassiman, J.J. and Van den Berghe, H. (1979) J. Biol. Chem. 254, 5155–5160.
[4] Sottrup-Jensen, L., Sand, O., Kristensen, L. and Fey, G.H. (1989) J. Biol. Chem. 264, 15781–15789.
[5] Kan, C.-C., Solomon, E., Belt, K.T., Chain, A.C., Hiorns, L.R. and Fey, G. (1985) Proc. Natl. Acad. Sci. USA 82, 2282–2286.
[6] Devriendt, K., Zhang, J., Van Leuven, F., Van den Berghe, H., Cassiman, J.J. and Marynen, P. (1989) Gene 81, 325–334.
[7] Sottrup-Jensen, L., Folkersen, J., Kristensen, T. and Tack, B.F. (1984) Proc. Natl. Acad. Sci. USA 81, 7353–7357.
[8] Gettins, P. and Cunningham, L.W. (1986) Biochemistry 25, 5011–5017.
[9] Gettins, P., Beth, A.H. and Cunningham, L.W. (1988) Biochemistry 27, 2905–2911.
[10] Hattori, M., Kusakabe, S.-I., Ohgusu, H., Tsuchiya, Y., Ito, T. and Yoshiyuki, S. (1989) Gene 77, 333–340.
[11] Traut, T.W. (1988) Proc. Natl. Acad. Sci. USA 85, 2944–2948.
[12] Than, G.N., Csaba, I.F., Szabó, D.G., Karg, N.J. and Novák, P.F. (1976) Vox Sang. 30, 134–138.
[13] Armstrong, N.P.I., Teisner, B., Redman, W.G., Westergaard, J.G., Folkersen, J. and Grudzinskas, J.G. (1986) Br. J. Obstet. Gynaecol. 93, 811–814.