Volume 236, number 1, 123–126 FEB 06187 August 1988

Complete amino acid sequence of a human platelet proteoglycan

Patrick M. Alliel, Jean-Pierre Périn, Philippe Maillet, François Bonnet, Jean-Philippe Rosa[+] and Pierre Jollès

Laboratoire des Protéines, UA CNRS no. 1188, Université de Paris V, 45 rue des Saints-Pères, F 75270 Paris Cedex 06 and [+] Unité INSERM no. 150, UA CNRS no. 334, Hôpital Lariboisière, 6 rue Guy Patin, F 75475 Paris Cedex 10, France

Received 27 June 1988

The primary structure of a human platelet proteoglycan (P.PG) core was established by a combination of amino acid sequence analysis and cDNA cloning. The deduced 131 amino acid long protein contains eight Ser-Gly repeats. The significance of homologies observed between P.PG and promyelocytic leukemia cell line proteoglycans is discussed.

Proteoglycan; (Human platelet, HEL cell line)

1. INTRODUCTION

A human platelet proteoglycan (P.PG) has been characterized as polydisperse molecules, the polydispersity being due to variations in the glycosylation and presumably sulphation levels. However the protein core appeared homogeneous for at least 90%; its N-terminal amino acid sequence has been reported previously [1]. The present work deals with the elucidation of the complete primary structure of the P.PG established by two complementary techniques: peptide sequencing of various P.PG fragments and sequencing of a cDNA coding for a proteoglycan in a human leukemic cell line (HEL). The latter presents megakaryocytic features [2–4]: megakaryocytes have been shown to be implicated in platelet proteoglycan biosynthesis [5].

2. MATERIALS AND METHODS

2.1. *Sequence determination of P.PG peptides*

The human P.PG core was obtained as previously described [1]. Trypsin and chymotrypsin treatments were carried out in a 50 mM Tris-HCl, pH 7.6, buffer at 37°C for 6 h at an enzyme: substrate ratio of 1:50. Cyanogen bromide (CNBr) treatment was performed in 70% formic acid for 18 h at room temperature. Reverse-phase HPLC on an aquapore RP-300 column (Brownlee) was performed in 0.1% trifluoroacetic acid (TFA) using a 0–40% acetonitrile gradient (Gilson chromatograph). Sephadex G-50 gel filtration was performed on a 90 × 1 cm column in a 50 mM sodium bicarbonate, pH 8, buffer. Superose 6 gel filtration was performed on a HR 10/30 column (Pharmacia) using a 50 mM sodium acetate, 4 M guanidinium hydrochloride (Gdn-HCl), pH 5.8, buffer. Automated Edman degradation was carried out on a 470 A sequencer equipped with a 120 PTH analyser (Applied Biosystems).

Correspondence address: P. Jollès, Laboratoire des Protéines, Université de Paris V, 45, rue des Saints-Pères, F 75270 Paris Cedex 06, France

2.2. *Screening of a HEL cell-derived cDNA library*

We took advantage of the homologies and differences between human P.PG [1] and rat yolk sac tumor proteoglycan core structures (PG19) [6] to synthesize a 26-mer oligonucleotide probe.

The probe corresponding to residues 8–16 in P.PG:

3′-AT$^{A}_{G}$GT$^{T}_{C}$ACCCAGGCGAC$^{A}_{G}$TT$^{A}_{G}$GGTCT-5′ OH

was 5′-end labelled to screen a HEL cDNA λgt10 library [7]. Filter hybridization was carried out in 6 × SSC, 5 × Denhardt, 0.5% SDS, 250 μg/ml salmon sperm DNA for 18 h at 37°C. The filters were then washed five times at 42°C for 10 min in 4 × SSC and subjected to autoradiography. The HEL derived clone was subcloned in M13 mp18 and mp19 and sequenced by the dideoxy-chain termination method [8].

3. RESULTS

3.1. *Protein core sequencing of the human P.PG*

Fig.1 presents the elution profile of the HPLC separation of the P.PG core tryptic peptides; the amino acid sequences of ten of them (T1–T10)

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/88/$3.50 © 1988 Federation of European Biochemical Societies

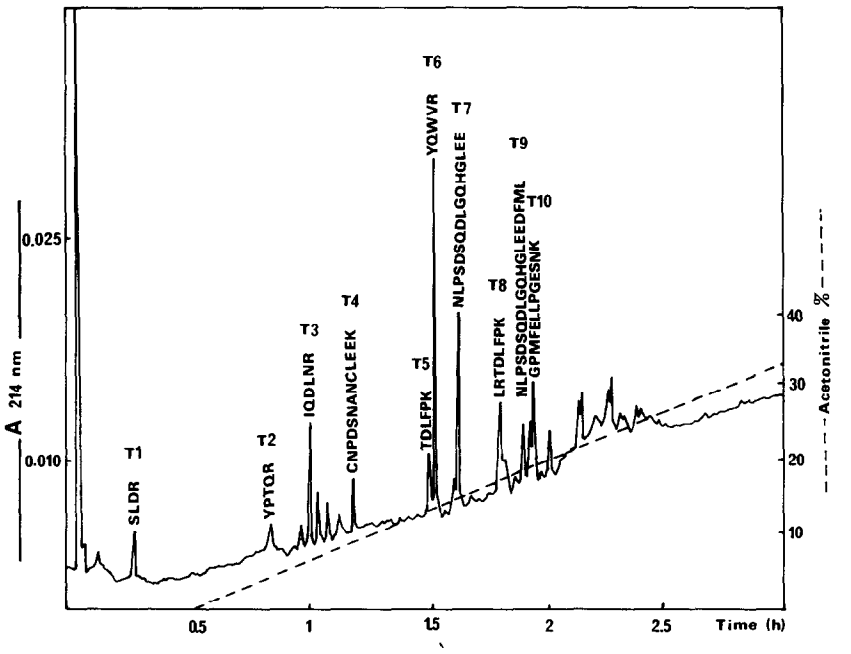


Fig.1. HPLC elution profile of a tryptic digest of the P.PG core. The elution was achieved with an acetonitrile linear gradient (0–40%) in 0.1% TFA, flow rate 1 ml/min. The sequences of the peptides (T1–T10) are indicated using the one letter amino acid abbreviation system.

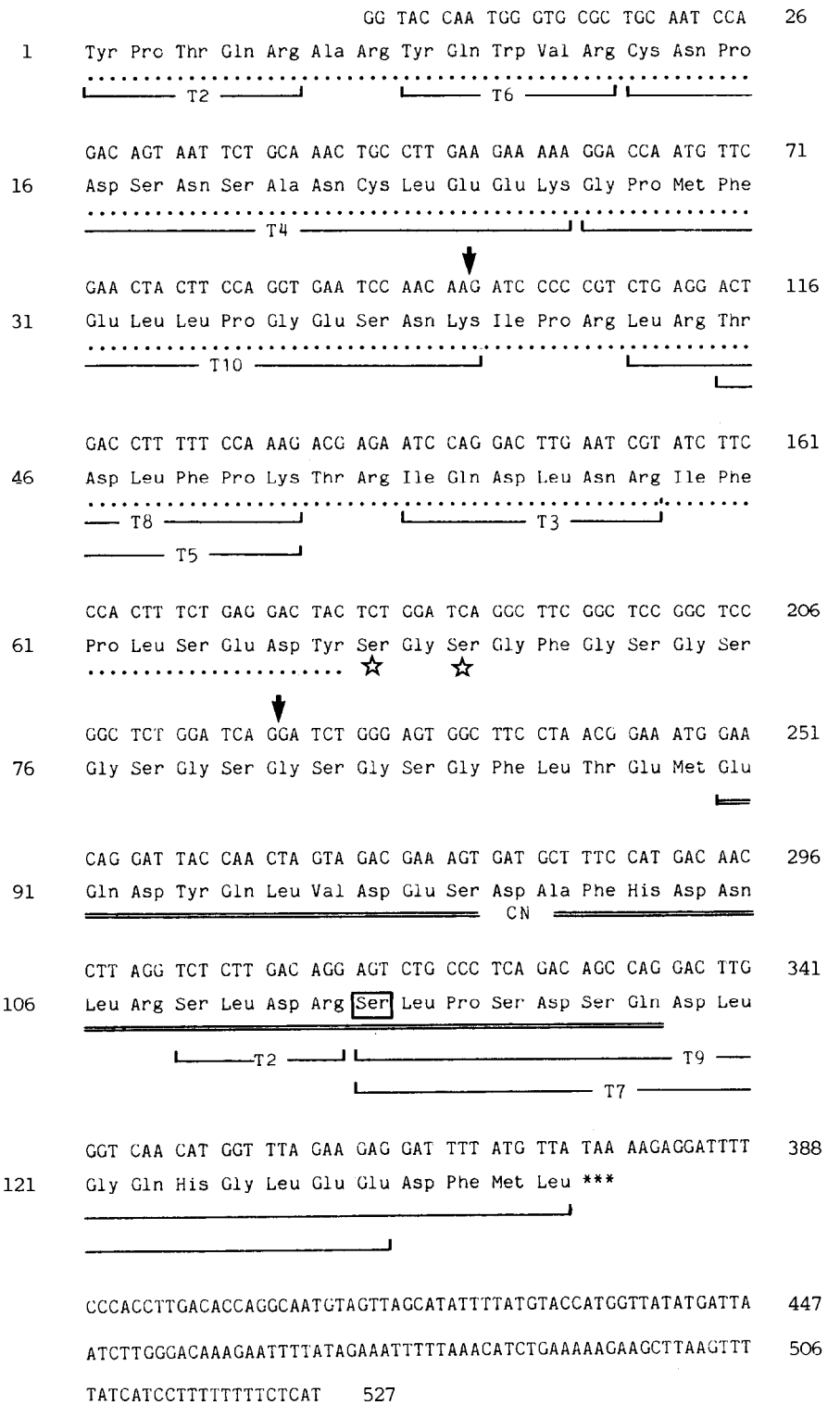
were determined. A CNBr peptide (CN) from the P.PG protein core was isolated after Sephadex G-50 gel filtration (not shown); its N-terminal amino acid sequence was determined up to residue 29: the precise localization of these various fragments is presented below (fig.2).

3.2. *Cloning and sequencing of the HEL cell-derived cDNA: primary structure of the human P.PG*

Three positive clones were selected after screening of 10^5 plaques. One of them (P13) was P.PG specific. Its nucleotide and deduced amino acid sequences are presented in fig.2. It consists of a 527 nucleotides sequence coding for 94% of the P.PG protein core and a tail of 153 nucleotides in the 3′-non coding region. Overlapping of the deduced protein sequence from the HEL cell-derived cDNA with the various P.PG peptides presented here and elsewhere [1] allowed us to establish the complete primary structure of the human P.PG core: 82% of the protein sequence was covered by protein sequencing data. The P.PG protein core consists of a 131 amino acid long protein (M_r 14641) with eight Ser-Gly repeats.

The only detected difference between the deduced sequence from the cDNA clone and the various sequenced peptides resides at position 112 in the P.PG core. A Ser residue was deduced at this position from the cDNA clone sequence (sequenced in both directions) while an Asn residue was identified in peptides T7, T9 and CN. The presence of Ser in the cDNA clone might be due either to a cloning artifact or to polymorphism.

Fig.2. Human P.PG core protein structure established from peptide sequencing experiments of the P.PG and the deduced amino acid sequence from the HEL cDNA clone nucleotide sequence. The tryptic peptides (T1–T10) characterized in the P.PG core are underlined, the double line indicates the CNBr fragment (CN), the dotted line represents the previously reported amino acid sequence (residues 1–66) [1]. The boxed (deduced) Ser residue was identified as an Asn residue during protein sequencing experiments of the P.PG core; (☆) the two glycosylation sites located in this work; (↓) the *Xho*II restriction sites. Stop codon is indicated by three asterisks. The numbers on the right and left refer to the nucleotides and amino acids in the respective sequence.


GG TAC CAA TGG GTG CGC TGC AAT CCA 26
1 Tyr Pro Thr Gln Arg Ala Arg Tyr Gln Trp Val Arg Cys Asn Pro
T2
T6
GAC AGT AAT TCT GCA AAC TGC CTT GAA GAA AAA GGA CCA ATG TTC 71
16 Asp Ser Asn Ser Ala Asn Cys Leu Glu Glu Lys Gly Pro Met Phe
T4
GAA CTA CTT CCA GGT GAA TCC AAC AAG ATC CCC CGT CTG AGG ACT 116
31 Glu Leu Leu Pro Gly Glu Ser Asn Lys Ile Pro Arg Leu Arg Thr
T10
GAC CTT TTT CCA AAG ACG AGA ATC CAG GAC TTG AAT CGT ATC TTC 161
46 Asp Leu Phe Pro Lys Thr Arg Ile Gln Asp Leu Asn Arg Ile Phe
T8
T3
T5
CCA CTT TCT GAG GAC TAC TCT GGA TCA GGC TTC GGC TCC GGC TCC 206
61 Pro Leu Ser Glu Asp Tyr Ser Gly Ser Gly Phe Gly Ser Gly Ser
GGC TCT GGA TCA GGA TCT GGG AGT GGC TTC CTA ACG GAA ATG GAA 251
76 Gly Ser Gly Ser Gly Ser Gly Ser Gly Phe Leu Thr Glu Met Glu
CAG GAT TAC CAA CTA GTA GAC GAA AGT GAT GCT TTC CAT GAC AAC 296
91 Gln Asp Tyr Gln Leu Val Asp Glu Ser Asp Ala Phe His Asp Asn
CN
CTT AGG TCT CTT GAC AGG AGT CTG CCC TCA GAC AGC CAG GAC TTG 341
106 Leu Arg Ser Leu Asp Arg Ser Leu Pro Ser Asp Ser Gln Asp Leu
T2
T9
T7
GGT CAA CAT GGT TTA GAA GAG GAT TTT ATG TTA TAA AAGAGGATTTT 388
121 Gly Gln His Gly Leu Glu Glu Asp Phe Met Leu ***
CCCACCTTGACACCAGGCAATGTAGTTAGCATATTTTATGTACCATGGTTATATGATTA 447
ATCTTGGGACAAAGAATTTTATAGAAATTTTTAAACATCTGAAAAAGAAGCTTAAGTTT 506
TATCATCCTTTTTTTTCTCAT 527

3.3. *Localization of glycosylated serine residues in P.PG*

After trypsin or chymotrypsin treatment of the native P.PG, two fractions, Tpg and Cpg respectively, were isolated by Superose 6 gel filtration; they contained the total hexuronate amount of the digested P.PG. The Tpg sequence (IFPLSEDYX-GXGF...) began at position 59 whereas Cpg (SEDYXGXGF...) began at position 63 in the P.PG (see fig.2). No amino acid (X) could be identified at identical positions for both of them while serine residues were deduced from the cDNA sequence. This result suggests that at least the two first serine residues in the Ser-Gly repeats are glycosylated in the P.PG. Tpg and Cpg behaved during Superose 6 gel filtration as untreated P.PG; this observation shows that P.PG is protease sensitive even in the vicinity of its glycosaminoglycan side chains, the latter being the main contributors of P.PG, Cpg and Tpg K_{av} values. This statement contrasts with the proposed protease resistance of secretory granules proteoglycans [9].

4. DISCUSSION

Our data are consistent with the localization of a Ser-Gly repeat containing proteoglycan in human platelets. Such proteoglycan protein cores exist in various tumoral or transformed cell lines in rats [6,10] as well as in humans ([11] and present data).

The cDNA sequence of a human promyelocytic leukemia cell line (HL-60) [11] appeared identical to that of our HEL clone. The HL-60 cell line is able to be induced to differentiate under certain in vitro conditions into cells that resemble neutrophils, monocytes-macrophages, eosinophils and basophils [11], while the HEL cell line presents megakaryocytic features [2–4]. The human P.PG described in this report might be the representative of a ubiquitous proteoglycan core in the hemopoietic lineage. Similarly a structural identity between Ser-Gly repeat containing proteoglycans has been shown in two different rat cell lines [6,10].

The comparison between such human and rat proteoglycans reveals a high level of strict homology in their N-terminal part (55%/66 amino acids of the human P.PG core). However the length of the Ser-Gly repeat structure and the C-terminal region are much more different (15 additional Ser-Gly repeats in the rat proteoglycans and only 35% of strict homology following the Ser-Gly repeats). The extent of Ser-Gly repeats and the C-terminal region appear therefore to be species specific.

Acknowledgements: The skillful technical assistance of Miss M. Rougeot and Mr Ly Quan-Le is gratefully acknowledged.

REFERENCES

[1] Périn, J.-P., Bonnet, F., Maillet, P. and Jollès, P. (1988) Biochem. J., in press.
[2] Tabilio, A., Rosa, J.-P., Testa, U., Kieffer, N., Nurden, A.T., Del Canizo, M.C., Breton-Gorius, J. and Vainchenker, W. (1984) EMBO J. 3, 453–459.
[3] Bray, P.F., Rosa, J.-P., Lingappa, V.R., Kan, Y.W., McEver, R.P. and Shuman, M.A. (1986) Proc. Natl. Acad. Sci. USA 83, 1480–1484.
[4] Kieffer, N., Debili, N., Wicki, A., Titeux, M., Henri, A., Mishal, Z., Breton-Gorius, J., Vainchenker, W. and Clemetson, K.J. (1986) J. Biol. Chem. 261, 15854–15862.
[5] Schick, B.P. and Schick, P.K. (1986) Semin. Hematol. 23, 68–87.
[6] Bourdon, M.A., Shiga, M. and Ruoslahti, E. (1986) J. Biol. Chem. 261, 12534–12537.
[7] Bray, P.F., Rosa, J.-P., Johnston, G.I., Shiu, D.T., Cook, P., Lau, C., Kan, Y.W., McEver, R.P. and Schuman, M.A. (1987) J. Clin. Invest. 80, 1812–1817.
[8] Sanger, F., Nicklen, S. and Coulsen, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463–5467.
[9] Tantravahi, R.L., Stevens, R.L., Austen, K.F. and Weis, J.H. (1986) Proc. Natl. Acad. Sci. USA 83, 9207–9210.
[10] Avraham, S., Stevens, R.L., Gartner, M.C., Austen, K.F., Lalley, P.A. and Weis, J.H. (1988) J. Biol. Chem. 263, 7292–7296.
[11] Stevens, R.L., Avraham, S., Gartner, M.C., Bruns, G.A.P., Austen, K.F. and Weis, J.H. (1988) J. Biol. Chem. 263, 7287–7291.