Proc. Natl. Acad. Sci. USA
Vol. 74, No. 6, pp. 2256–2258, June 1977
Biochemistry

Amino acid sequence of human platelet factor 4

(thrombosis/heparin-binding protein/proteoglycan/automated Edman degradation)

THOMAS F. DEUEL*†‡, PAMELA S. KEIM*, MARTHA FARMER*†, AND ROBERT L. HEINRIKSON*

* Departments of Medicine and Biochemistry, The Pritzker School of Medicine, The Franklin McLean Memorial Research Institute, The University of Chicago, Chicago, Illinois 60637; and the † Departments of Medicine and Biological Chemistry, Washington University School of Medicine, and the Jewish Hospital of St. Louis, St. Louis, Missouri 63110

Communicated by Donald F. Steiner, March 14, 1977

ABSTRACT **Human platelet factor 4, a protein that binds heparin, has been purified to apparent homogeneity and the complete amino acid sequence of the protein has been determined. The 70-residue polypeptide chain contains no methionine, tryptophan, or phenylalanine, and contains only a single tyrosyl residue. The sequence analysis demonstrates a highly negatively charged amino-terminal region. The carboxyl-terminal region of the polypeptide is unusual in that it contains a repetitive clustering of positively charged and hydrophobic pairs of amino acids; preliminary evidence suggests that this domain may play a role in the binding of heparin.**

Heparin is a potent anticoagulant that is used extensively in clinical medicine for the treatment of thrombotic disease (1). Platelets release a small protein that neutralizes the anticoagulant effect of heparin (2). A plasma factor with this property was first predicted in 1948 by Conley (3) when he noticed that patients with thrombocytopenia had increased sensitivity to heparin. Partial purification of the specific, platelet-derived protein, referred to as platelet factor 4 (4), was reported by Deutsch *et al.* (5), and, more recently, various investigators have published methods for isolation of the protein in an apparently homogeneous form (6–13). Platelet factor 4 is a relatively small, heat-stable protein that is insoluble in solutions of low ionic strength, but that is soluble in high ionic strength solutions and in acidic buffers (12, 13). Platelet factor 4 is released with and bound to a carrier proteoglycan molecule from which it is displaced by heparin (8, 10).

Because of a possible relationship of platelet factor 4 to arterial and venous thrombosis (14–20), knowledge of the structure of the protein is of special importance. Such knowledge would be desirable also both from the standpoint of learning more about how platelet factor 4 combines with heparin and the proteoglycan carrier molecule, as well as from the standpoint of forming a basis for understanding what role, if any, platelet factor 4 may play in normal coagulation. The complete amino acid sequence analysis of platelet factor 4 described herein represents the initial efforts of this laboratory to provide detailed information upon which to base further inquiry into the complex structural and functional properties of the protein.

METHODS

Platelet packs were obtained from local blood banks after the expiration date for use in transfusions had passed. The platelets were stored frozen and thawed immediately before use in the purification of platelet factor 4. The purified protein was obtained by precipitation between ammonium sulfate concentrations of 50 and 100%, by affinity chromatography on columns of heparin– ϵ-aminocaproic acid–Sepharose prepared as described by Schmer (21), and by gel filtration through Sephadex G-100 in 1 M NaCl. These methods and the modified thrombin time used in the assay of platelet factor 4 were adapted from those reported by Levine and Wohl (12). Purified platelet factor 4 binds 27 units of heparin per mg of protein. The purified protein appears as a single band after polyacrylamide gel electrophoresis (15% acrylamide) in the presence of sodium dodecyl sulfate and is estimated to be at least 95% pure both by this criterion and by automated Edman degradation of the intact *S*-carboxamidomethylated protein. The purity of the preparation was also established by immunological techniques. The IgG fraction of antisera obtained from rabbits immunized with platelet factor 4 was isolated by ammonium sulfate precipitation and DEAE-cellulose chromatography (22). The purified IgG gave a single precipitin line against purified platelet factor 4 in immunodiffusion experiments, suggesting that the purified protein behaves as a single major antigenic species.

Platelet factor 4 (5–10 mg) was reduced and S-[^{14}C]carboxamidomethylated by reaction with iodo[2-^{14}C]acetamide (Amersham-Searle, specific radioactivity = 60,000 cpm/μmol) as described earlier (23), except the final reaction volume was 2 ml. This protein derivative (R[^{14}C]CAM-factor 4) was subjected to automated Edman degradation in order to establish the partial sequence of platelet factor 4 at the NH_2-terminus. The fragments required to complete the structural analysis were derived by tryptic hydrolysis at arginyl residues in citraconylated R[^{14}C]CAM-factor 4. All methods were as described in an earlier report from this laboratory (23).

Acid hydrolysates (6 M HCl, 110°, 24 hr under reduced pressure) of R[^{14}C]CAM-factor 4 and peptides derived therefrom by various cleavage procedures were analyzed for amino acid content by automated ion-exchange chromatography with a Durrum D-500 analyzer. This method also served in the identification and quantitation of residues liberated from the COOH-terminus of R[^{14}C]CAM-factor 4 by digestion with carboxypeptidase Y (24) and of the amino acids regenerated from selected phenylthiohydantoin derivatives by hydrolysis in HI (25).

Automated Edman degradation (26–28) of intact R[^{14}C]CAM-factor 4 and fragments obtained from it by tryptic cleavage at arginyl residues was performed in a Beckman Protein-Peptide Sequencer (model 890-C) programmed with the "new improved" fast peptide procedure (102974) supplied by the manufacturer. Phenylthiohydantoin derivatives were identified and quantitated by gas chromatography (29), thin-layer chromatography (30), and, in some cases, by amino acid analysis after HI hydrolysis (25).

Abbreviations: R[^{14}C]CAM-factor 4, S-[^{14}C]carboxamidomethylated protein derivative of factor 4; Tc, tryptic peptide.

‡ Present address: Division of Hematology-Oncology, Departments of Medicine and Biological Chemistry, Washington University School of Medicine, and The Jewish Hospital of St. Louis, St. Louis, MO 63110. To whom reprint requests should be addressed.

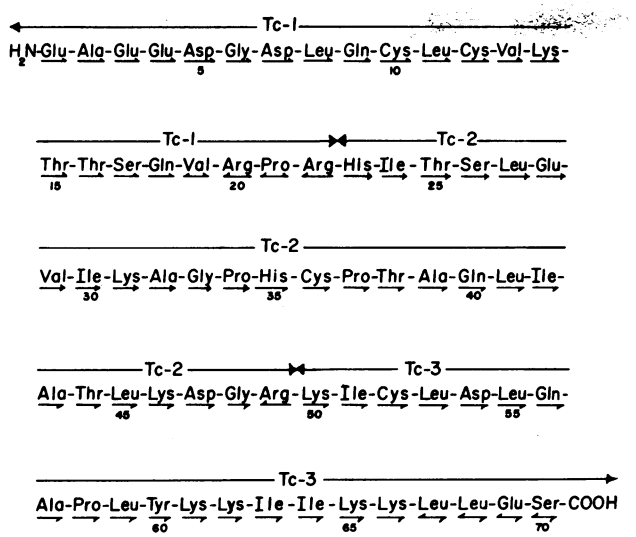

FIG. 1. Amino acid sequence of human platelet factor 4. Identification and quantitation of most residues was by gas chromatography (29); Arg, Lys, and His were determined by amino acid analysis after hydrolysis of the phenylthiohydantoin derivatives in HI (25). The latter procedure, as well as thin-layer chromatography (30), was used in verification of certain assignments. Tc, trypic peptides. Half arrows (→) indicate assignments based upon automated Edman degradation of R[^{14}C]CAM-factor 4. Half arrows (→) indicate assignments based upon similar analysis of Tc-2 and Tc-3. Half arrows pointing to the left indicate identifications confirmed by digestion with carboxypeptidase Y. Half-cystines were identified by radioactivity in the phenylthiohydantoin derivatives.

RESULTS

The amino acid sequence of platelet factor 4 is presented in Fig. 1, together with symbols designating the methods used to establish the structure. Experimental details will be published in a later report (R. L. Heinrikson, P. S. Keim, M. Farmer, and T. F. Deuel, unpublished data), but the overall strategy was to obtain as much information as possible by automated Edman degradation of the intact protein and the large derivative fragments.

The compositional analysis of R[^{14}C]CAM-factor 4 given in Table 1 shows the absence of methionine, tryptophan, and phenylalanine and the presence of three residues of arginine. Automated Edman degradation of 6 mg (0.77 μmol) of R[^{14}C]CAM-factor 4 provided identification of the 30 NH_2-terminal residues, including two of the three arginines in the polypeptide. The sequence analysis was completed essentially by automated Edman degradation of the tryptic peptides from citraconylated R[^{14}C]CAM-factor 4 (12 mg; 1.54 μmol). Three fragments were produced by tryptic cleavage at Arg-22 and Arg-49, and these were aligned immediately on the basis of the earlier sequence analysis. As expected, no cleavage was observed at Arg-20 since the next residue in sequence is a proline (Fig. 1). These peptides are designated Tc-1 (residues 1–22), Tc-2 (residues 23–49), and Tc-3 (residues 50–70) in Fig. 1. The COOH-terminal fragment, Tc-3, was separated from a mixture of the other two by gel filtration of the digest on a column (2.5 × 180 cm) of Sephadex G-50 developed in 0.5 M NH_4HCO_3. Elution with solutions of high ionic strength was essential in order to provide good yields (70%) of Tc-3 since this peptide tended to smear over the elution profile at lower salt concentrations. Tc-2 and Tc-1, the more retarded peptides, were separated by gel filtration on a column (1.5 × 95 cm) of Sephadex G-25 (superfine) in 0.1 M acetic acid. Under these conditions Tc-2 eluted slightly ahead of Tc-1. After appropriate corrections for contaminating fragments, the amino acid compositions of Tc-1, Tc-2, and Tc-3 (Table 1) account for all of the residues present in the parent polypeptide.

Table 1. Amino acid compositions of reduced and *S*-[^{14}C]-carboxamidomethylated platelet factor 4 and tryptic peptides isolated from the citraconylated protein derivative*

Amino acid	R[^{14}C]CAM factor 4	Tc-1†	Tc-2	Tc-3‡	Total
Lysine	7.9 (8)	1.0 (1)	2.0 (2)	4.8 (5)	8
Histidine	1.9 (2)	0.0 (0)	2.0 (2)	0.0(0)	2
Arginine	2.7 (3)	1.7 (2)	0.9 (1)	0.0 (0)	3
Carboxy methyl-cysteine	3.7 (4)	1.8 (2)	0.9 (1)	0.6 (1)	4
Aspartic acid	4.1 (4)	2.0 (2)	1.0 (1)	0.9 (1)	4
Threonine	4.3 (5)	1.9 (2)	2.9 (3)	0.0 (0)	5
Serine	2.5 (3)	0.8 (1)	0.8 (1)	0.6 (1)	3
Glutamic acid	8.5 (9)	4.6 (5)	2.0 (2)	1.9 (2)	9
Proline	4.2 (4)	1.1 (1)	2.0 (2)	1.1 (1)	4
Glycine	2.9 (3)	1.0 (1)	1.9 (2)	0.4 (0)	3
Alanine	5.0 (5)	1.0 (1)	2.8 (3)	1.2 (1)	5
Valine	2.9 (3)	1.6 (2)	0.6 (1)	0.0 (0)	3
Isoleucine	5.4 (6)	0.0 (0)	2.4 (3)	3.0 (3)	6
Leucine	9.5 (10)	1.7 (2)	2.8 (3)	4.8 (5)	10
Tyrosine	0.9 (1)	0.0 (0)	0.0 (0)	0.9 (1)	1
Total	70	22	27	21	70

* Protein and peptide samples of 10–20 nmol were hydrolyzed in 6 M HCl at 110° for 24 hr under reduced pressure and aliquots were analyzed for amino acid content on a Durrum D-500 analyzer. Relative molar quantities are given for each sample; the figures in parentheses indicate integral residue numbers. No corrections were made for decompositional losses or for incomplete liberation of residues in 24 hr.

† Values corrected for 15% contamination with Tc-3.

‡ Values corrected for 10% and 15% contamination with Tc-2 and Tc-1, respectively.

The sequence analysis of platelet factor 4 was completed by automated Edman degradation of Tc-2 (0.35 μmol) and Tc-3 (0.15 μmol). In both cases, the complete peptide sequences were established in a single run. The high efficiency of program 102974 in the analysis of relatively short peptides was noted in another report from this laboratory (23). Confirmation of the sequence at the COOH-terminus of Tc-3 was made by analysis of the release of amino acids by carboxypeptidase Y.

DISCUSSION

Venous and arterial thrombosis are major causes of death in humans. Heparin remains the most effective of the known antithrombotic drugs available for the prevention of venous thrombosis (1). Platelet factor 4, a product of the platelet release reaction, is a low-molecular-weight protein that binds heparin specifically. Because platelet factor 4 may be important in coagulation and may serve as a marker for platelet activation, detailed information on the structure of the protein is desirable. The determination of its primary structure is the subject of this communication.

Platelet factor 4 is a single polypeptide chain of 70 amino acids, including four half-cystines and a single tyrosine. The molecular weight calculated from the primary structure is 7767. The protein contains no methionine, tryptophan, or phenylalanine. A recent abstract by Chesterman *et al.* (31) reported the sequence of the 20 NH_2-terminal residues in platelet factor 4;

our results are in complete agreement with theirs. However, the authors' conclusion that the protein contains 63 residues and 3 half-cystines is incompatible with the sequence analysis reported herein. Especially noteworthy in the structure of platelet factor 4 are the sequences at the NH_2- and COOH-termini of the molecule. The NH_2-terminal region is highly acidic; indeed, five of the first seven residues in sequence are either glutamic or aspartic acids. A distinctive sequence also occurs at the COOH-terminal region (residues 61–68) in which pairs of lysyl residues are separated by pairs of leucines or isoleucines (-Lys-Lys-Ile-Ile-Lys-Lys-Leu-Leu-). This repetitive clustering of positively charged and hydrophobic domains is strongly suggestive of a role of the COOH-terminal segment in the function of the protein. Indeed, Handin and Cohen (13) have shown that treatment of platelet factor 4 with *O*-methylisourea results in reduced binding of heparin by the modified protein derivative. If, as the data suggest, lysine residues are involved in the binding of heparin, it seems quite possible that the binding site is located in this region of the COOH-terminal segment of the protein. Preliminary results from our laboratory indicate that the COOH-terminal peptide, Tc-3 (Fig. 1), reduces the heparin-dependent prolongation of the thrombin time assay used in our experiments (12) even though the peptide is *S*-carboxamidomethylated at Cys-52. However, the possibility that this effect may be due to the presence of low levels of Tc-1 and Tc-2 in the Tc-3 preparation (Table 1) cannot be ruled out at present. Knowledge of the amino acid sequence will facilitate interpretation of these and other chemical modification studies directed toward understanding the function of platelet factor 4. Furthermore, detailed analysis of the binding site of heparin should be possible through the study of peptides generated during the present investigation.

Note Added in Proof. Some of the results described here were reported recently in abstract form (32). Since submission of the manuscript, Gottfried Schmer (University of Washington, Seattle) called attention to a single discrepancy between our sequence and that determined independently in Seattle (Hermodson, M., Schmer, G. & Kurachi, K., *J. Biol. Chem.*, in press). This involved residue 56, which we initially identified as Glu. We also have shown independently that residue 56 is Gln, in agreement with Hermodson *et al.*

We gratefully acknowledge the kindness of Drs. Kenneth Robbins and Thiru Vaithianathan, and Mssrs. Roger Sullivan and Robert Gilbert for providing the platelets used in these studies. We are grateful also to Dr. Shirley Levine who made her manuscript available to us prior to its publication. Supported in part by Grants CA-13980 and HL-18577 from the National Institutes of Health, by Grant BMS-75-23506 from the National Science Foundation, and by a contract with the U.S. Energy Research and Development Administration. The Franklin McLean Research Institute is operated by the University of Chicago for the U.S. Energy Research and Development Administration under Contract EY-76-C-02-0069. T.F.D. holds Faculty Research Award 133 from the American Cancer Society.

The costs of publication of this article were defrayed in part by the payment of page charges from funds made available to support the research which is the subject of the article. This article must therefore be hereby marked "advertisement" in accordance with 18 U. S. C. §1734 solely to indicate this fact.

1. Wessler, S. (1977) *Fed. Proc.* **36**, 66–69.
2. Niewiarowski, S., Poplawski, A., Lipinski, B. & Farbiszewski, R. (1968) *Exp. Biol. Med.* **3**, 121–128.
3. Conley, C. L., Hartmann, R. C. & Lalley, J. S. (1948) *Proc. Soc. Exp. Biol. Med.* **69**, 284–287.
4. Deutsch, E. & Kain, W. (1961) in *Henry Ford Hospital International Symposium: Blood Platelets*, eds. Johnson, S. A., Monto, R. W., Rebuck, J. W. & Horn, R. C., Jr. (Little Brown and Co., Boston, MA), pp. 337–345.
5. Deutsch, E., Wawersich, E. & Franke, G. (1957) *Thromb. Diath. Haemorrh.* **1**, 397–412.
6. Kaser-Glanzmann, R., Jakabova, M. & Luscher, E. F. (1972) *Experientia* **28**, 1221–1223.
7. Nath, N., Lowery, C. T. & Niewiarowski, S. (1975) *Blood* **45**, 537–550.
8. Moore, S., Pepper, D. S. & Cash, J. D. (1975) *Biochim. Biophys. Acta* **379**, 370–384.
9. Nath, N., Niewiarowski, S. & Joist, J. H. (1973) *J. Lab. Clin. Med.* **82**, 754–768.
10. Barber, A. J., Kaser-Glanzmann, R., Jakabova, M. & Luscher, E. F. (1972) *Biochim. Biophys. Acta* **286**, 312–329.
11. Pepper, D. S. & Moore, S. (1973) *Br. J. Haematol.* **24**, 666.
12. Levine, S. P. & Wohl, H. (1976) *J. Biol. Chem.* **251**, 324–328.
13. Handin, R. I. & Cohen, H. J. (1976) *J. Biol. Chem.* **251**, 4273–4282.
14. Walsh, P. N. (1976) in *Heparin: Chemistry and Clinical Usage*, eds. Dakkar, V. V. & Thomas, D. P. (Academic Press, New York), pp. 125–143.
15. O'Brien, J. R. (1974) *Thromb. Diath. Haemorrh.* **32**, 116–123.
16. Cotton, R. C., Shaikh, M. S. & Dent, R. V. (1968) *J. Atheroscler. Res.* **8**, 959–966.
17. Farbiszewski, R., Niewiarowski, S., Worowski, K. & Lipinski, B. (1968) *Thromb. Diath. Haemorrh.* **19**, 578–583.
18. McDonald, L. & Edgill, M. (1961) *Lancet* **i**, 844–847.
19. Gormsen, J. & Haxholdt, B. F. (1960) *Acta Chir. Scand.* **120**, 121–136.
20. O'Brien, J. R., Tulevski, V. G., Etherington, M., Madgwick, T., Alkjaersig, N. & Fletcher, A. (1974) *J. Lab. Clin. Med.* **83**, 342–354.
21. Schmer, G. (1972) *Trans. Am. Soc. Artif. Intern. Organs* **18**, 321–323.
22. Fahey, J. L. (1967) in *Methods in Immunology and Immunochemistry* (Academic Press, New York), Vol. 1, pp. 321–332.
23. Heinrikson, R. L., Krueger, E. T. & Keim, P. S. (1977) *J. Biol. Chem., in press.*
24. Hayashi, R., Moore, S. & Stein, W. H. (1973) *J. Biol. Chem.* **248**, 2296–2302.
25. Smithies, O., Gibson, D., Fanning, E. M., Goodfliesch, R. M., Gilman, J. G. & Ballantyne, D. L. (1971) *Biochemistry* **10**, 4912–4921.
26. Edman, P. (1950) *Acta Chem. Scand.* **4**, 277–282.
27. Edman, P. & Begg, G. (1967) *Eur. J. Biochem.* **1**, 80–91.
28. Fietzek, P. & Kühn, K. (1972) *Fortschr. Chem. Forsch.* **29**, 1–28.
29. Pisano, J. J. & Bronzert, T. J. (1969) *J. Biol. Chem.* **244**, 5597–5607.
30. Jeppsson, J. O. Sjöquist, J. (1967) *Anal. Biochem.* **18**, 264–269.
31. Chesterman, C. N., Morgan, F. J., McGready, J. R. & Begg, G. S. (1976) *Sixteenth International Congress of Hematology, Kyoto, Japan. Abstract 8-34*, p. 317.
32. Deuel, T. F., Keim, P., Farmer, M. & Heinrikson, R. L. (1977) *Clin. Res.* **25**, 474A.