Vol. 170, No. 2, 1990 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
July 31, 1990 Pages 915-922

ISOLATION AND CHARACTERIZATION OF A cDNA CLONE FOR A NOVEL SERINE-RICH NEUTROPHIL PROTEIN

P. Bellavite, F. Bazzoni, M.A. Cassatella, K.J. Hunter* and J.V. Bannister*,#

Institute of General Pathology, University of Verona, Italy

*Biotechnology Center, Cranfield Institute of Technology, Bedford, U.K.

#Department of Biomedical Sciences, University of Malta, Malta

Received June 11, 1990

A cDNA expression library from pig blood neutrophils was immunoscreened with a rabbit antiserum raised against a 32 kDa neutrophil membrane phosphoprotein. Previous work indicated this protein as a component of the superoxide-forming NADPH oxidase enzyme complex (1,2). Only one cDNA clone (B+) was highly positive. The B+ clone contained a 1109 bp insert, with an open reading frame encoding for 284 amino acids. The deduced B+ amino acid sequence contained a 72 amino acid domain with proline and glutamine repeats and two domains extremely enriched with serine residues. The isolated cDNA hybridizes with a 3.1 kb mRNA expressed in pig and human leukocytes. © 1990 Academic Press, Inc.

Stimulated neutrophils undergo to a series of functional and metabolic modifications such as enzyme secretion, oxygen radicals production, gene expression, etc. Cell activation is associated with phosphorylation of many proteins, including proteins that have been indicated as components of the free-radical generating enzyme system NADPH oxidase (2-7). One of these putative NADPH oxidase components, a membrane protein of 32 kDa, has received considerable attention in our laboratory in the past few years. From guinea pig and pig neutrophils, this protein was consistently copurified with the enzymatic activity and with cytochrome b_{558} (1-3). Moreover, it appeared to be markedly phosphorylated in membranes of activated neutrophils (2,3,8) and in cell-free system by protein kinase C (9), suggesting that phosphorylation of this component could be one of the mechanisms by which the enzyme activity is regulated.

Attempts to purify this polypeptide by conventional methods were unsuccessful because of the scarcity of material, its marked

0006-291X/90 $1.50
Copyright © 1990 by Academic Press, Inc.
All rights of reproduction in any form reserved.

tendency to aggregate and to denaturate after detergent treatment (1,2,10). For this reason, we took advantage of recombinant DNA technology in order to isolate the cDNA coding for the 32 kDa protein, to deduce its aminoacidic sequence, to compare it with other known sequences and to make available cDNA probes. In this report the initial results from this approach are presented, i.e. the isolation and sequencing of a cDNA identified by immunoscreening of a pig neutrophil library with antibodies raised against the pig 32 kDa protein.

MATERIALS AND METHODS

Purification of 32 kDa protein and preparation of antiserum. Pig neutrophils were isolated and activated by phorbol-myristate-acetate (PMA) as previously described (10). Extraction and subsequent purification of NADPH oxidase have been performed as already described in detail (2,8,10). The protein band localized at 32 kDa on polyacrylamide slab gel (1,2) was excised, electroeluted from the gel and precipitated with acetone. Polyclonal antibodies were raised in a New Zealand white rabbit by subcutaneous injections of approximately 50 µg of purified protein emulsified in Freund's adjuvant for the primary immunization and in incomplete adjuvant for three booster injections. The antiserum was characterized by assessing its ability to bind the 32 kDa protein in Western blot analysis of subcellular fractions separated from PMA-activated pig neutrophils (8). After electrophoresis according to Laemmli (11), samples were transferred onto a nitrocellulose membrane (Bio Rad) and Western blotting was performed essentially as described by Towbin et al. (12).

cDNA synthesis, cloning and immunoscreening of the library. A cDNA expression library was constructed as follows: poly(A)+ RNA isolated from pig neutrophils was used as a template for cDNA synthesis. cDNA was syntetized by the method of Gubler and Hoffman (13) and then a cDNA library was prepared in the Lambda gt11 expression vector (Amersham cDNA cloning system, Amersham International, England). The library was then screened with 1:200 anti-32 kDa polyclonal antibodies according to the procedure of Hyunh et al. (14).

Immunoblot analysis of fusion protein from cDNA positive clone. The B+ positive cDNA clone and an other random clone were used to lysogenize E. Coli Y1089 (15). Total protein extract from bacterial lysates were electrophoresed on a 12% SDS-polyacrylamide gel and transferred to nitrocellulose paper. Polyclonal antibodies raised against 32 kDa were used at dilution of 1:200 for immunoblot analysis of bacterial proteins as in the previous section.

Nucleotide sequence analysis. cDNA from Lambda gt11 B+ clone was cut with suboptimal concentrations of Eco RI and the fragments were subcloned into the Eco RI unique site of pUC18 plasmid by the "shot gun" cloning method (16). Nucleotide sequences were determinated by the dideoxynucleotide chain-termination method (17). Sequencing was performed with both Sequenase (United States Biochemical) and Taq polymerase (Promega), and by the use of the nucleotide analog dITP in order to avoid the problems associated with the G+C regions and secondary structures.

Northern blot analysis. Northern blot analysis was performed as previously reported (18). Total RNA was isolated from human neutrophils and pig neutrophils and monocytes by the guanidium isothiocyanate/cesium chloride procedure (19). Filters were hybridized with cDNA purified from the positive clone, which was labelled with [^{32}P]-dCTP by using Multiprime DNA labeling system (Amersham International, England).

RESULTS

A cDNA expression library made from pig neutrophils was screened with polyclonal antibodies raised against the 32 kDa protein. Immunoblotting was selected because two preliminary attempts made to obtain information about the amino acid sequence were unsuccessful due to the fact that the protein is blocked at the aminoterminal end and owing to the scarcity of pure protein other attempts based on limited proteolysis and sequencing were not

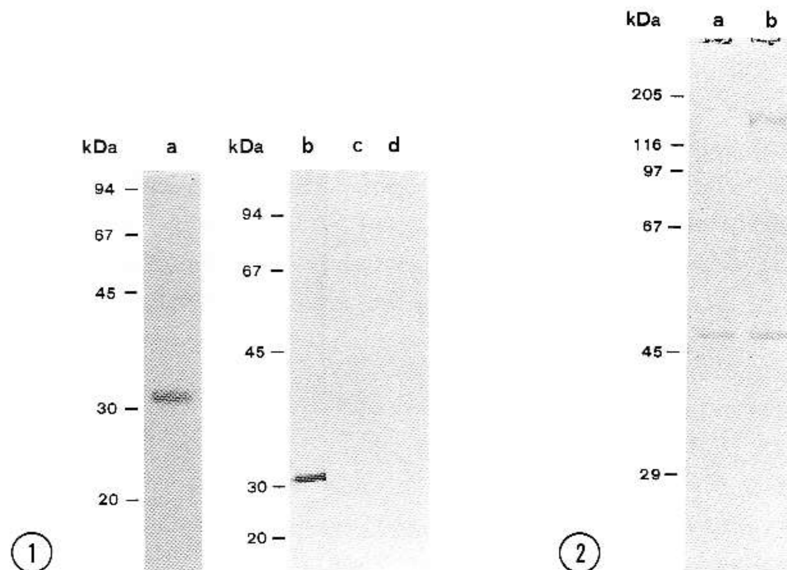

Fig. 1. SDS-polyacrylamide gel electrophoresis of the purified 32 kDa protein and immunoblot analysis. Lane a shows the Coomassie blue stained protein preparation that was utilized for raising antibodies, purified from PMA-activated pig neutrophils as described in the text. Lanes b-d show Western blot with rabbit anti-32 kDa protein antibodies of subcellular fractions from PMA-activated neutrophils. b, membranes; c, cytosol; d, granules (azurophil + specific granule fraction).

Fig. 2. Western blot analysis of hybrid recombinant protein. Proteins from E. Coli Y1089 lysogens induced with non-recombinant Lambda gt11 (a) and with B+ clone isolated on immunoscreening (b) were separated by SDS-polyacrylamide gel electrophoresis, transferred to nitrocellulose membranes and subjected to immunoblot analysis with rabbit anti-32 kDa protein antibodies as described in the methods.

possible. The antiserum prepared by immunizing a rabbit with a preparation of pure protein (Fig. 1a) allowed at 1:200 dilution the detection of approximately 0.2 µg of the 32 kDa protein both in western blots of neutrophil fractions (Fig. 1b-d) and in dot-blots of non denatured partially purified NADPH oxidase. The immunoreactive protein was recognized by the antiserum in plasmamembranes from pig neutrophils, where the reaction was highly specific (Fig. 1b). Only faint signals, probably derived from aspecific binding, were present in the cytosol (Fig. 1c) and the granular fraction (Fig. 1d).

This antiserum was used to screen a cDNA expression library from pig neutrophils containing a total $1.5x10^5$ recombinant p.f.u. The immunoscreening of the library resulted in the identification of only one plaque (named B+) that was highly positive in two successive cycles of plaque isolation and screening. This clone did not react with control serum from non-immunized rabbits. Immunoblot analysis of recombinant clones (Fig.2b) showed that anti-32 kDa polyclonal antibodies recognize the fusion protein (apparent m.w. of about 160 kDa) from a total extract of E. Coli Y1089 lysogenized with the B+ clone. The same protein band was also recognized by monoclonal antibodies against ß-galactosidase (not shown). The anti-32 kDa antibodies did not detect any fusion protein in other recombinant random clones (Fig. 2a). The band at 47 kDa was found in all clones tested and was probably due to aspecific reaction because it also reacted with serum from non-immunized rabbits.

Characterization of the B+ clone. Eco RI digestion and agarose gel electroporesis showed that the cDNA of B+ clone contained two fragments of 750 and 350 bp. The phage was then cut with suboptimal doses of Eco RI and the fragments were subcloned into pUC18 plasmid by "shot-gun" cloning. Of several randomly isolated plasmids that were analyzed by minipreps, most incorporated either the 750 bp or the 350 bp insert, and one of them contained a complete 1,100 bp insert that cross-hybridized with both smaller fragments. Plasmids containing only the 750 bp or 350 bp fragments were used for DNA sequencing, while the one containing the complete insert helped to choose the correct orientation of the sequences.

Fig. 3 shows the nucleotide and deduced amino acid sequence of the B+ cDNA. This cDNA contains a coding region of 853 bp, corresponding to 285 amino acid residues, and a 3' untranslated region of 256 bp. There are no initiation codons in this sequence. At positions 1086-1091 a putative polyadenylation signal, AATAAA

```
         10                        30                        50                        70
GAATTCTAGTGCTCTAGAAATGGAAAATGAGCCTGAGGAAAAAGAAGAAAAAGAATCCGAGCCCCAGCCT
 AsnSerSerAlaLeuGluMetGluAsnGluProGluGluLysGluGluLysGluSerGluProGlnPro

                   90                       110                       130
GAGCCTATGTCTCAGCCTCAGCCCCTGCCTCAGCCCCAGCCCCAGTCCCAGCCACAGCCACAGCCACAG
GluProMetSerGlnProGlnProLeuProGlnProGlnProGlnSerGlnProGlnProGlnProGln

        150                       170                       190
GCCCAGACCCAGGCCCAGTGCCAGTCCCAGTCCCAACCAGCACTCCAGCCCCAGCCTCTCTCTCAGCCT
AlaGlnThrGlnAlaGlnCysGlnSerGlnSerGlnProAlaLeuGlnProGlnProLeuSerGlnPro

210                   230                       250                       270
GAGACTATCCCCTTAGCTGTTTTACAGCCACCACCTCAGGCTATTCAGGAGCAAGGGCATTTACCACCT
GluThrIleProLeuAlaValLeuGlnProProProGlnAlaIleGlnGluGlnGlyHisLeuProPro

         290                       310                       330
GAGAGAAAGGAGTTTCCTGTAGAATCTGCAAAACTGACTGAGGTGACAGTAGAGCCAGTCTTGACAGTA
GluArgLysGluPheProValGluSerAlaLysLeuThrGluValThrValGluProValLeuThrVal

  350                       370                       390                       410
CATCCAGAGAGCAAGAGCAAAACCAAAACTAGGAGCAGAAGTAGAGGTCGAGCTAGAAATAAAACAAGC
HisProGluSerLysSerLysThrLysThrArgSerArgSerArgGlyArgAlaArgAsnLysThrSer

              430                       450                       470
AAGAGCAGAAGTCGAAGCAGTAGCAGTAGTAGTTCTAGTAGCAGTTCAACCAGTAGCAGCAGTGGAAGC
LysSerArgSerArgSerSerSerSerSerSerSerSerSerSerThrSerSerSerSerGlySer

   490                       510                       530                       550
AGTTCTAGCAGTGGCAGTAGTAGTAGTCGAAGTAGTTCCAGTAGCAGCTCCGATCAAGTGGCAGCAGCA
SerSerSerSerGlySerSerSerSerArgSerSerSerSerSerSerAspGlnValAlaAlaAla

              570                       590                       610
GCTAGAGACAGTAGCAGCAGCACTACCAGTAGTAGTGAGAGTAGAAGTCGAAGTAGGGGCCGGGCACAT
AlaArgAspSerSerSerSerThrThrSerSerSerGluSerArgSerArgSerArgGlyArgAlaHis

      630                       650                       670                       690
AACAGAGATAGAAAACACAGGAGGAGTGTGGATCGGAAGCGGAGGGATGCTTCTGGACTAGAAAGAAGT
AsnArgAspArgLysHisArgArgSerValAspArgLysArgArgAspAlaSerGlyLeuGluArgSer

                  710                       730                       750
CACAAATCTTCAAAAGGTGGTAGTAGTAGAGATACAAAAGGATCAAAGGATAAGAATTCCCGGTCTGAC
HisLysSerSerLysGlyGlySerSerArgAspThrLysGlySerLysAspLysAsnSerArgSerAsp

        770                       790                       810
AGAAAGAGGTCTATATCAGAGAGTAGTCGATCAGGCAAAAGATCCTCAAGAAGTGAAAGAGACCGAAAA
ArgLysArgSerIleSerGluSerSerArgSerGlyLysArgSerSerArgSerGluArgAspArgLys

830                    850                       870                       890
TCAGACAGGAAAGACAAAAGGCGTTAATGGAAGAAGCCAGGCTTTCTTAGCCTATTCTTGCAGCAGAAG
SerAspArgLysAspLysArgArgEnd

         910                       930                       950
ATTTCTTGATGAGTAAAAGGATTACCTTTCCTTGTAAGGAGAATGCTGCCTTAAGAATTCCATGTTGTA

 970                    990                    1010                     1030
AAAATTTTTTTGGAAAATACAGACTGTTTGTTTACCAGACTTCTTGTACTTTTGCATAATTTTGTAAGA

          1050                      1070                      1090
GTTATTTATCAAAATTATGTGAGGTTCCAAAATATGTAAAAATAATAATAATAAAAAAACGATTAACGGAA

TTC
```

Fig. 3. Nucleotide sequence and deduced amino acid sequence of clone B+ cDNA. A putative polyadenylation signal AATAAA is underlined.

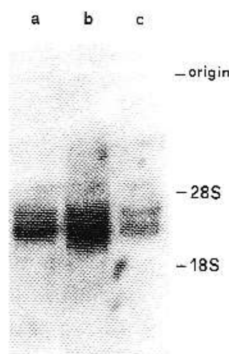

Fig. 4. Northern blot analysis. Total RNA (10 µg in each lane) from pig neutrophils (a), pig monocytes (b) and human neutrophils (c) was hybridized with ^{32}P-labelled clone B+ cDNA (750 bp insert) as described in the text. The size markers are 28S (5.5 kbp) and 18S (2.1 kbp) rRNA.

is present. Inspection of the predicted amino acid sequence shows distinct features such as repeated proline and glutamine residues located at positions 21-92 and two regions rich in serine at positions 138-202 and 229-271. Particularly interesting is a sequence of 35 amino acids (residues 144-178), which contains 31 serine residues. A search of the protein data bases (NBRF and New-EMBL) revealed the existence of some homology with other proteins bearing either proline/glutamine-rich or serine-rich domains. Among the proline/glutamine rich proteins, the highest scores have been found with bovine amelogenin (39% identity in an 84 residue overlap), basic proline-rich protein precursor (human) (29% identity in an 84 residue overlap), c-hordein (50% identity in an 59 residue overlap), α/ß gliadin precursor (prolamin) (52% identity in a 42 residue overlap), virion protein UL 36 (54% identity in 81 residue overlap). Among the serine-rich proteins, the following showed the best homologies: glucoamylase S1 (yeast) (34% identity in 103 residue overlap), vitellogenin II precursor (phosvitin) (53% identity in an 81 residue overlap), sericin precursor (28% identity in 159 residue overlap).

The 750 bp insert hybridized on Northern blot analysis with a mRNA species of approximately 3,100 bp, expressed in pig neutrophils, pig monocytes and also in human neutrophils (Fig. 4). Also the 350 bp insert gave the same 3,100 bp mRNA size on Northern blot analysis (not shown), indicating that both inserts belonging to B+ clone are representative of the same mRNA transcripts.

DISCUSSION

In this report we describe the cloning and characterization of a cDNA of a neutrophil protein isolated by immunoscreening of an expression library with polyclonal antiserum raised against the 32 kDa neutrophil phosphoprotein that we described in our laboratory as a component of the NADPH oxidase system in guinea pig (1) and pig (2,3,8) neutrophils.

The isolated cDNA is 1,109 bp long and does not represent the entire coding region of the protein. In fact: a) it contains an open reading frame starting at the 2nd nucleotide in the (+) strand and there is no evidence that one of the two ATG codons (at nucleotide positions 20 and 77) may be the starting codon, b) it hybridizes with mRNA from pig neutrophils, monocytes and human neutrophils of considerably larger size (3,100 bp). These features indicate that the B+ clone cDNA likely corresponds to the 3' terminus of a larger mRNA and that the deduced polypeptide represents the 285 COOH-terminus aminoacids of a larger protein.

Comparison of the predicted aminoacid sequence with available databases showed that this cDNA has not previously been described. Infact there is not significant homology with other proteins, except for the repeated sequences rich in proline/glutamine and those containing serine. Proteins with similar domains are not common in nature and very rare in mammals. The possible biological significance of the domains rich in proline/glutamine is not clear, whereas the regions that contain serine have been shown in other proteins to be the target of marked phosphorylation. For example, chicken vitellogenin contains a fragment - called phosvitin - that is composed of a series of 258 serine residues which are heavily phosphorylated before secretion from liver cells (20). The existence of a putative site of phosphorylation in the predicted aminoacid sequence in our cDNA is in keeping with the observation that the 32 kDa protein in pig neutrophil membranes incorporates the highest amount of phosphate during cell activation (2,8).

Although no definite proof that the B+ clone contains a cDNA coding for part of the 32 kDa protein that has been used for raising the antiserum can be presented (a cross-reaction with other neutrophil proteins cannot be excluded), such cross-reaction is very improbable since the antiserum was raised against the pure protein (Fig. 1a) and gave a very specific reaction with only the 32 kDa protein on Western blots of neutrophil membrane proteins (Fig. 1b).

However, even if the precise biochemical and functional identification of the protein coded by the B+ clone needs more detailed investigation, the present communication demonstrates that circulating neutrophils and monocytes transcribe the message for a serine-rich protein, that may be of relevance in the active and pleiomorphic posphorylation-related processes that take place in these cells during activation.

Acknowledgments: This work was done with contribution by Ministero P.I. (fondi 40%) and from Associazione Italiana Ricerca sul Cancro. P.B. thanks the European Molecular Biology Organization for a short-term fellowship. J.V.B. and P.B. acknowledge financial support from the European Economic Community.

REFERENCES

1. Serra, M.C., Bellavite, P., Davoli, A., Bannister, J.V., and Rossi, F. (1984) Biochim. Biophys. Acta 788, 138-146.
2. Bellavite, P., Papini, E., Zeni, L., Della Bianca, V. and Rossi, F. (1985) Free Rad. Res. Commun. 1, 11-29.
3. Bellavite, P. (1988) Free Rad. Biol. Med. 4, 225-261.
4. Garcia, R.C., and Segal, A.W. (1988) Biochem. J. 252, 901-905.
5. Kramer, I.M., Verhoeven, A.J., van der Bend, R., Weening, R.S., and Roos, D. (1988) J. Biol. Chem. 263, 2352-2357.
6. Okamura, N., Curnutte, J.T., Roberts, R.L., and Babior, B.M. (1988) J. Biol. Chem. 263, 6777-6782.
7. Teahan, C.G., Totti, N., Casimir, C.M., and Segal, A.W. (1990) Biochem. J. 267, 485-489.
8. Bellavite, P., Dusi, S., and Cassatella, M.A. (1987) Free Rad. Res. Commun. 4, 83-98.
9. Papini, E., Grzeskowiak, M., Bellavite, P., and Rossi, F. (1985) FEBS Lett. 190, 204-208.
10. Bellavite, P., Jones, O.T.G., Cross, A.R., Papini, E. and Rossi, F. (1984) Biochem. J. 223, 639-648.
11. Laemmli, U.K. (1970) Nature 227, 680-685.
12. Towbin, H., Stahelin, T., and Gordon, J. (1979) Proc. Natl. Acad. Sci. USA 76, 4350-4354.
13. Gubler, U., and Hoffman, B.J. (1983) Gene 25, 263-269.
14. Hyunh, T.V., Young, R.A., and Davis, R.W. (1985) in DNA Cloning Techniques, a Practical Approach (D. Glover, ed.), Vol. I, pp. 49-73. IRL Press, Oxford.
15. Young, R.A., and Davis, R.W. (1983) Proc. Natl. Acad. Sci. USA 80, 1194-1198.
16. Davis, L.G., Dibner, M.D., and Battey, J.F. (1986) Basic Methods in Molecular Biology. Elsevier, New York.
17. Sanger, F., Nicklen, S., and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463-5467.
18. Cassatella, M. A., Hartman, L., Perussia, B., and Trinchieri, G. (1989) J. Clin. Invest. 83, 1570-1580.
19. Maniatis, T., Fritsch, E.F., and Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Lab.
20. van het Ship, F.D., Samallo, J., Broos, J., Ophuis, J., Mojet, M., Gruber, M., and AB, G. (1987) J. Mol. Biol. 196, 245-260.

Reprinted from Biochemical and Biophysical Research Communications 170, 915-922 (1990)
Copyright © 1990 Academic Press, Inc. Printed in U.S.A.