Volume 322, number 2, 135–138 FEBS 12403 May 1993
© 1993 Federation of European Biochemical Societies 00145793/93/$6.00

cDNA cloning of a new type of subunit of mammalian proteasomes

Stuart Thomson, Deborah F. Balson and A. Jennifer Rivett

Department of Biochemistry, University of Leicester, Leicester, UK

Received 15 February 1993; revised version received 16 March 1993

The primary structure of a new type of subunit (RN3) of rat proteasomes (multicatalytic proteinase complexes) has been determined from the nucleotide sequence of the cDNA. The cDNA encodes a protein of 232 amino acids but the directly determined N-terminal amino acid sequence suggests that the subunit is post-translationally processed to a M_r = 24k form. Sequence alignments reveal a similarity of RN3 to other proteasome subunits. It can be designated a B-type proteasomal subunit but is not closely related to the β subunit of the archaebacterial proteinase or to other members of the B group.

Proteasome; Multicatalytic proteinase; Subunit RN3; cDNA cloning; Polymerase chain reaction

1. INTRODUCTION

Proteasomes or multicatalytic proteinase complexes are high molecular mass cylindrical particles which are found in all types of eukaryotic cells (see [1–4] for reviews). They are localized in both the nucleus and in the cytoplasm [5] and seem to play a major role in non-lysosomal pathways of protein turnover. Purified mammalian proteasome preparations give rise to a complicated pattern of up to 25 spots on two-dimensional polyacrylamide gels and many of the polypeptides are antigenically distinct [6,7]. A related proteinase with only two different types of subunit, α and β, has been found in archaebacteria.

cDNAs for many subunits of proteasomes, including several rat subunits [8–13], have been cloned and their sequences found to be related to those of the archaebacterial subunits [15]. Proteasome subunits can be divided into two groups, A and B, depending upon whether they show the characteristics of the α or β archaebacterial subunit, respectively [1,15]. Most of the cloned cDNAs are of A-type subunits. A-type subunits share a highly conserved region at their N-terminus [1] and at least some of them are N-terminally blocked. B-type subunits, on the other hand, often have an unblocked N-terminus and have more variable sequences. Seven sequences have been obtained by direct N-terminal sequence analysis of rat proteasomes [14].

Using N-terminal amino acid sequence data for rat (sequence 3) [14] and human (β subunit) [16] proteasomes, we have cloned the cDNA for one of the B-type subunits of the rat proteasome which we have named RN3. The results provide new information about sequences of the mammalian B-type proteasome subunits.

Correspondence address: A.J. Rivett, Department of Biochemistry, University of Leicester, Leicester LE1 7RH, UK. Fax: (44) (533) 52 3369.

Abbreviations: PAGE, polyacrylamide gel electrophoresis; PCR, polymerase chain reaction; RSV, Rous sarcoma virus.

2. MATERIALS AND METHODS

Since cells transformed with Rous sarcoma virus have greatly enhanced levels of proteasomal mRNAs [17], a cDNA library of Rous sarcoma virus transformed rat-1 fibroblasts (ts LA29 cells) was constructed in the λZAP expression vector (Stratagene) using a Stratagene cDNA synthesis kit. The sequence of the RN3 subunit cDNA was then determined using the polymerase chain reaction method of Walker et al. [18]. Degenerate oligonucleotides corresponding to the protein sequences TQNPMV and VIAADM (obtained from the N-terminal protein sequence of the rat and human subunit [14,16]) were used as PCR primers. These oligonucleotides had the following sequences:

Forward: 5′-CCCGGGGAATTCAC(N)CA(A/G)AA(T/C)CC-(N)ATGGT-3′
EcoRI

Reverse: 5′-CCCGGGAAGCTTCAT(A/G)TC(N)GC(N)GC(A/T)AT(N)AC-3′
HindIII

The 102 nucleotide product of the PCR reaction was identified by Southern blotting using a degenerate probe based upon part of the intervening protein sequence GVKFEG [16] and having the following sequence:

5′-GG(N)GT(N)AA(A/G)TT(T/C)GA(A/G)GG-3′

The probe was radiolabelled using polynucleotide kinase and [γ-^{32}P]dATP. The PCR product was cut from the agarose gel, purified using the Stratagene Geneclean II kit and then cloned into M13 mp19 for sequence analysis using the dideoxy chain termination method [19]. The sequence was as follows:

5′ACCCAGAATCCGATGGTGACTGGGACATCCGTACTA-GGGGTGAAGTTCGACTGCGGAGTGGTGATAGCCGCGG-ACATG-3′

The underlined portion of the sequence was used to design a forward primer for a second PCR reaction:

5-'CCCGGGAATTC ACTGGGACATCCGTACTA-3'
EcoRI

in which the reverse primer was:

5'-CCCGGGAAGCTT$(T)_{18}$-3'.
HindIII

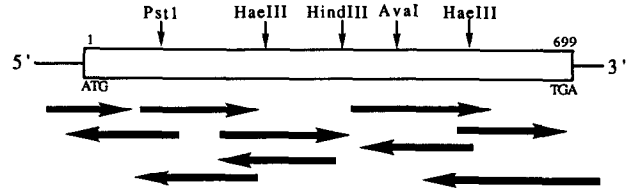
The PCR product (820 base pairs) was identified by Southern blotting using a nondegenerate probe (5'-TGAAGTTCGACTGCGGAGTG-3'), excised from the agarose gel and cloned into the pBluescript vector for DNA sequence analysis.

The 5' cDNA sequence was obtained by PCR using a forward primer based on a pBluescript sequence immediately 5' to the cDNA insert and a reverse primer based on internal RN3 cDNA sequence. The product was identified by Southern blot, excised from the agarose gel and cloned into pBluescript for sequencing. Radiolabelling of probes, Southern blotting and subcloning were all carried out by standard methods [20].

3. RESULTS AND DISCUSSION

The N-terminal PCR approach used here has been used very successfully by Walker et al. [18] to determine the primary structure of a large number of mitochondrial inner membrane proteins and of subunits of NADH:ubiquinone oxidoreductase. N-Terminal amino acid sequence was determined for only seven of the more than twenty rat proteasome polypeptides [14] and some of these correspond to subunits such as RC5 and RING12 which have already been cloned [10,21]. The success of the PCR approach using a rat cDNA library with degenerate primers designed partially from human protein sequence is not surprising in view of the high degree of identity between proteasome subunits from different animal species.

The nucleotide sequence of the RN3 cDNA and the deduced amino acid sequence are shown in Fig. 1. The lack of complete identity of the N-terminal region with the rat N-terminal sequence 3 [14] may be due to sequence errors since the subunit was not isolated for that study. The conclusion that the ATG located at nucleotide position 1–3 is the initiation codon is based upon the fact that it is the first ATG from the 5' end to encode




```
                                5'GGCACGAGTGACTAAGATGGAAGCGTTTTGGGAGTCACGGAC -61
TGGCCACTGGGCGGAGGTCCGGGGCAGTTTTACCGGCGTCCCGTCACCCCCAGCTGCCTC -1
 M  D  P  M  S  A  P  A  R  S  I  T  R  T  Q  N  P  M  L  T   20
ATGGATCCGATGTCTGCGCCTGCGAGGTCCATCACTCGGACCCAGAATCCGATGCTGACT 60
 G  T  S  V  L  G  V  K  F  D  C  G  V  V  I  A  A  D  M  L   40
GGGACATCCGTACTAGGGGTGAAGTTCGACTGCGGAGTGGTGATTGCTGCAGACATGCTG 120
 G  S  Y  G  S  L  A  R  F  R  I  I  S  R  I  M  R  V  N  D   60
GGCTCCTACGGCTCCCTGGCTCGTTTCCGCATTATCTCTCGTATTATGCGAGTCAACGAC 180
 S  T  M  L  G  A  S  G  D  Y  A  D  F  E  Y  L  K  E  V  L   80
AGCACTATGCTAGGTGCTTCAGGAGACTATGCTGATTTCCAGTATTTGAAGCAAGTTCTC 240
 G  E  M  V  I  D  E  E  L  F  G  D  G  H  S  Y  S  P  R  A  100
GGCCAGATGGTGATTGATGAGGAGCTGTTCGGAGATGGACATAGCTATAGTCCTAGAGCT 300
 I  H  S  W  L  T  R  A  M  Y  S  R  R  S  K  M  N  P  L  W  120
ATCCATTCATGGTTGACAAGGGCCATGTACAGTCGCCGCTCCAAGATGAACCCTCTGTGG 360
 N  T  K  V  I  G  G  Y  A  G  G  E  S  F  L  G  Y  V  D  M  140
AACACCAAGGTCATTGGAGGCTATGCTGGTGGAGAAAGCTTTCTTGGTTATGTGGACATG 420
 L  G  V  A  Y  E  A  P  S  L  A  T  G  Y  G  A  Y  L  A  E  160
CTTGGTGTAGCTTATGAAGCCCCTTCACTAGCCACTGGTTATGGTGCATACTTGGCTCAG 480
 P  L  L  R  E  V  L  E  K  Q  P  V  L  S  Q  T  E  A  R  E  180
CCTCTGCTTCGAGAAGTTCTAGAGAAGCAACCGGTGCTGAGTCAGACAGAGGCCCGAGAG 540
 L  V  E  R  C  M  R  V  L  Y  Y  R  D  A  R  S  Y  N  R  F  200
CTTGTGGAGCGCTGCATGCGAGTGCTGTACTACCGAGACGCTCGTTCGTATAACCGGTTT 600
 Q  V  A  T  V  T  E  K  G  V  E  I  E  G  P  L  S  A  Q  T  220
CAAGTTGCTACTGTAACTGAAAAAGGTGTGGAGATAGAAGGACCACTGTCAGCACAGACC 660
 N  W  D  I  A  H  M  I  S  G  F  E  *                      232
AACTGGGACATTGCTCACATGATCAGTGGCTTTGAATGAAATCCAGATGAAGTGTCCAGA 720
GTTAAGCTT 3'                                                729
```

Fig. 1. Nucleotide sequence and deduced amino acid sequence of subunit RN3. (Upper panel) Restriction endonuclease map of cloned cDNA for component RN3 and sequencing strategy. The box shows the coding region and the numbers indicate the position of the first nucleotide of the putative initiation codon and of the last nucleotide of the termination codon TGA. Arrows indicate the sequenced regions. (Lower panel) Nucleotide sequence of the cDNA and the predicted amino acid sequence. Nucleotides are numbered in the 5'-to-3' direction as above. The underlined sequence was that determined directly by protein sequencing.

protein sequence consistent with the directly determined N-terminal protein sequence of the RN3 subunit [14]. However, the possibility that the initiation codon is at position 10–13 cannot be ruled out at this stage. The cDNA encodes a protein of 232 amino acids (Fig. 1). The molecular mass of the protein coded from amino acid position 14 (24 kDa) is similar to that determined for subunit RN3 from SDS-PAGE gels [14].

Sequence alignment [22] with other proteasome subunits showed a clear relationship but, as for other pro-

```
          1                                                          60
RAT  RN3  MDPMSAPARSITRTQNPMLTGTSVLGVKFDCGVVIAADMLGSYGSLARFRIISRIMRVND
                                |:||||||||||||.||:|||||||||||||||||| |||||:||:
XEN BETA                        MVTGTSVLGVKFDGGVIIAADMLGSYGSLARFRNISRIMKVNE

         61                                                         120
          STMLGASGDYADFQYLKQVLGQMVIDEELFGDGHSYSPRAIHSWLTRAMYSRRSKMNPLW
          .|:||||||||||:|||||::||||||||.||||.|||:||||||||.||.||||||||||
          NTILGASGDYADYQYLKQVIDQMVIDEELVGDGHNYSPKAIHSWLTRVMYNRRSKMNPLW

        121                                                         180
          NTKVIGGYAGGESFLGYVDMLGVAYEAPSLATGYGAYLAQPLLREVLEKQPVLSQTEARE
          || ||||: .|||||||||.||||||||.:|||:|||||||||||||| |..:.||..|||:
          NTVVIGGFYNGESFLGYVDKLGVAYEAPTIATGFGAYLAQPLLREVTENKATLSKEEARQ

        181                                                 232
          LVERCMRVLYYRDARSYNRFQVATVTEKGVEIEGPLSAQTNWDIAHMISGFE
          ||:|||:||||||||||||||::.||||.|||:||||||.:|||:|||:|||||
          LVDRCMKVLYYRDARSYNRFEITTVTESGVEVEGPLSSETNWEIAHLISGFE
```

Fig. 2. Amino acid sequence alignment of *Xenopus* proteasome β subunit with rat subunit RN3. The data for putative proteasome subunit *Xenopus* β was taken from the EMBL database. Numbers refer to the RN3 sequence.

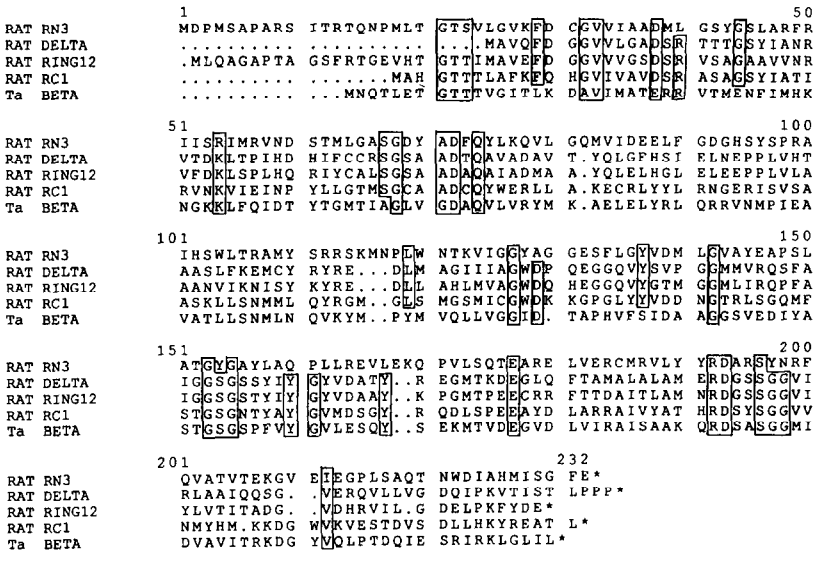


Fig. 3. Amino acid sequence alignment of rat B-type subunits with the archaebacterial β subunit sequence. Numbers refer to the RN3 sequence. Boxed regions indicate positions where residues are conserved in at least four of the five sequences.

teasome subunits, no overall homology was found to any other proteins currently in the SWISSPROT protein sequence database. The RN3 sequence is 80% identical to that of a putative *Xenopus* proteasome β subunit sequence determined from the nucleotide sequence of a cDNA clone (Fig. 2). These observations indicate a high degree of conservation of the RN3/β subunit between species as found for other proteasome subunits [1,4,24].

RN3 has been grouped with the B-type subunits [1] because it has an unblocked N-terminus, it does not contain the highly conserved N-terminal region of the A-type subunits, and its N-terminal region does show some obvious homology to the archaebacterial β subunit. However, like RC5, RN3 is not very closely related to the archaebacterial β subunit (Fig. 3) and in fact has a similar % identity with both the archaebacterial α and β subunits (approximately 22%). This level of identity is comparable to that between the α and β subunits themselves [15]. RN3 is also 22% identical to RC5 and shows similar levels of identity to other rat B-type subunits (Fig. 3). Comparison of the RN3 sequence with that of the other rat B-type subunits, which include the two MHC-encoded subunits RC1 (RING10, LMP7) [25] and RING12 (LMP2) [21], as well as the DELTA subunit [21], is shown in Fig. 3. Although RN3 does not share some of the highly conserved regions of DELTA, RC1 and RING12, there are a number of highly conserved amino acid residues between all the known rat B-type subunit sequences. The primary structure of RN3 is only 20% identical to that of the putative yeast proteasome subunit PUP1 [26] which was taken from the SWISSPROT database, but these two subunits do share a highly conserved region from amino acid position 20 to 38 in RN3.

The function of the N3 subunit of proteasomes has yet to be determined. It does not appear to contain the conserved putative nuclear localization signals, clusters of acidic amino acid residues or putative phosphorylation sites found in some of the A-type subunits [4].

Acknowledgements: We thank Dr. John Walker for advice on cloning strategies, Professor John Wyke for providing ts LA29 cells, Dr. Grant Mason for growing the cells, and Deborah Langton for the preparation of oligonucleotides. The work was supported by the Medical Research Council. A.J.R. is a Lister Institute–Jenner Research Fellow.

REFERENCES

[1] Rivett, A.J. (1993) Biochem. J. 291, 1 10.
[2] Orlowski, M. (1990) Biochemistry 29, 10289–10297.
[3] Goldberg, A.L. and Rock, K.L. (1992) Nature 357, 375–379.
[4] Tanaka, K., Tamura, T., Yoshimura, T. and Ichihara, A. (1992) New Biol. 4, 173–187.
[5] Rivett, A.J. and Knecht, E. (1993) Curr. Biol. 3, 127–129.
[6] Rivett, A.J. and Sweeney, S.T. (1991) Biochem. J. 278, 171–177.
[7] Kaltoft, M.B., Koch, C., Uerkvitz, W. and Hendil, K.B. (1992) Hybridoma 11, 507–517.
[8] Fujiwara, T., Tanaka, K., Kumatori, A., Shin, S., Yoshimura, T., Ichihara, A., Tokunaga, F., Aruga, R., Iwanaga, S., Kakizuka, A. and Nakanishi, S. (1989) Biochemistry 28, 7332–7340.
[9] Tanaka, K., Fujiwara, T., Kumatori, A., Shin, S., Yoshimura, T., Ichihara, A., Tokunaga, F., Aruga, R., Iwanaga, S., Kakizuka, A. and Nakanishi, S. (1990) Biochemistry 29, 3777–3785.
[10] Tamura, T., Tanaka, K., Kumatori, A., Yamada, F., Tsurumi, C., Fujiwara, T., Ichihara, A., Tokunaga, F., Aruga, R. and Iwanaga, S. (1990) FEBS Lett. 264, 91–94.
[11] Tanaka, K., Kanayana, H., Tamura, T., Lee, D.H., Kumatori, A., Fujiwara, T., Ichihara, A., Tokunaga, F., Aruga, R. and Iwanaga, S. (1990) Biochem. Biophys. Res. Commun. 171, 676–683.
[12] Kumatori, A., Tanaka, K., Tamura, T., Fujiwara, T., Ichihara, A., Tokunaga, F., Onikura, A. and Iwanaga, S. (1990) FEBS Lett. 264, 279–282.

[13] Sorimachi, H., Tsukahara, T., Kawasaki, H., Ishiura, S., Emori, Y., Sugita, H. and Suzuki, K. (1990) Eur. J. Biochem. 193, 775–781.
[14] Lilley, K.S., Davison, M.D. and Rivett, A.J. (1990) FEBS Lett. 262, 327–329.
[15] Zwickl, P., Grziwa, A., Pühler, G., Dahlmann, B., Lottspeich, F. and Baumeister, W. (1992) Biochemistry 31, 964–972.
[16] Lee, L.W., Moomaw, C.R., Orth, K., McGuire, M.J., DeMartino, G.N. and Slaughter, C.A. (1990) Biochim. Biophys. Acta 1037, 178–185.
[17] Balson, D.F., Skilton, H.E., Sweeney, S.T., Thomson, S. and Rivett, A.J. (1992) Biol. Chem. Hoppe-Seyler 373, 623–628.
[18] Walker, J.E., Arizmendi, J.M., Dupuis, A., Fearnley, I.M., Finel, M., Medd, S.M., Pilkington, S.M., Runswick, M.J. and Skehel, J.M. (1992) J. Mol. Biol. 226, 1051–1072.
[19] Sanger, F., Nicklen, S. and Coulsen, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463–5467.
[20] Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) Molecular Cloning: a Laboratory Manual, 2nd edn., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
[21] Tamura, T., Shimbara, N., Aki, M., Ishida, N., Bey, F., Scherrer, K., Tanaka, K. and Ichihara, A. (1992) J. Biochem. 112, 530–534.
[22] Devereux, J., Haeberli, P. and Smithies, O. (1984) Nucleic Acids Res. 12, 387–395.
[23] Van Riel, M.C.H.M. and Martens, G.J.M. (1991) FEBS Lett. 291, 37–40.
[24] Tamura, T., Lee, D.H., Osaka, F., Shin, S., Chung, C.H., Fujiwara, T., Tanaka, K. and Ichihara, A. (1991) Biochim. Biophys. Acta 1089, 95–102.
[25] Aki, M., Tamura, T., Tokunaga, F., Iwanaga, S., Kawamura, Y., Shimbara, N., Kagawa, S., Tanaka, K. and Ichihara, A. (1992) FEBS Lett. 301, 65–68.
[26] Haffter, P. and Fox, T.D. (1991) Nucleic Acids Res. 19, 5075.