Volume 323, number 3, 276–278 FEBS 12531 June 1993

Peptide sequencing identifies MSS1, a modulator of HIV Tat-mediated transactivation, as subunit 7 of the 26 S protease

Wolfgang Dubiel[b], Katherine Ferrell[a] and Martin Rechsteiner[a]

[a]*Department of Biochemistry, University of Utah, School of Medicine, Salt Lake City, UT 84132, USA* and [b]*Institut für Biochemie, Medizinische Fakultät (Charité), Humboldt-Universität zu Berlin, Hessische Str 3–4, O-1040 Berlin, Germany*

Received 7 April 1993

Subunit 7 is an integral component of the human erythrocyte 26 S protease Peptide sequence analysis reveals that 22 amino acids from the N-terminus of subunit 7 correspond exactly to the N-terminus of MSS1, a modulator of HIV gene expression. Additional internal peptides from subunit 7 obtained by CNBr cleavage also match 100% with the deduced amino acid sequence of MSS1 Based on the fact that directly sequenced peptides from subunit 7 are identical to more than 12% of the hypothetical translation product of MSS1, and the fact that the molecular weight of subunit 7 (49 kDa) corresponds to the predicted molecular weight of MSS1 (48,633 Da), we conclude that subunit 7 is MSS1.

MSS1; HIV; Human 26 S protease; Putative ATPase; Tat

1. INTRODUCTION

The 26 S protease, the degradative component of the ubiquitin pathway, was isolated and characterized by Hough et al. [1–3]. The multimeric complex consists of a proteolytic core of small subunits (M_r's 20–32 kDa), which is identical with the multicatalytic protease (MCP) [4–7]. These associate with a set of larger polypeptides (M_r's 42–110 kDa), an ATPase complex previously termed the 'ball' [7], which appears to correspond to conjugate-degrading factors 1 and 2 [8]. Since the multicatalytic protease does not degrade ubiquitin conjugates and is not ATP-dependent, the 10 or more larger polypeptides must affect the activity and specificity of the 26 S protease.

In contrast to the multicatalytic protease, there is little known about the identity and function of proteins comprising the 26 S ATPase complex. Recently we reported the primary structure of subunit 4 (S4), an integral component of the 26 S ATPase complex [9]. It is a member of an ATPase family [10,11], and together with TBP1, SUG1, and MSS1, it forms a highly conserved subfamily [9].

TBP1 (Tat binding protein 1) was discovered by probing a λgt11 fusion cDNA library from human jurkat T-cells with the HIV protein, Tat, a transcriptional activator of HIV gene expression [12]. A human homolog of TBP1, called TBP7, was identified by screening a library with TBP1 cDNA [13]. SUG1, a yeast homolog of TBP1, was identified in two ways: its absence allows a C-terminally truncated yeast transcriptional activator, GAL4, to function [14]; SUG1 was also discovered upon sequencing the upstream region of a yeast gene encoding the large subunit of the initiation factor eIF-4F [15]. MSS1 was first identified as a mammalian suppressor of sgv1, a cold sensitive defect in the yeast gene (SGV) that encodes a CDC28/cdc2-like kinase. It was later found to enhance Tat-mediated transactivation [16,17]. The high degree of similarity among members of the subfamily led us to propose that all might be subunits of the ATP-dependent 26 S protease. This proposal is supported here by the identification of 26 S protease subunit 7 as MSS1.

Correspondence address W. Dubiel, Institut für Biochemie, Medizinische Fakultät (Charité), Humboldt-Universität zu Berlin, Hessische Str. 3–4, O-1040 Berlin, Germany

Abbreviations. SDS, sodium dodecyl sulfate; PAGE, polyacrylamide gel electrophoresis, HIV, human immunodeficiency virus; sLLVY-MCA, succinyl-Leu-Leu-Val-Tyr-7-amido-4-methyl coumarin; CNBr, cyanogen bromide, PVDF, polyvinylidene difluoride.

2. MATERIALS AND METHODS

2.1. *Purification of 26 S protease*

The 26 S protease was prepared from 20 units of outdated human blood following the protocol of Hough et al. [2] except that proportionally larger columns were employed. After gel filtration on TSK HW55-Sepharose, fractions displaying the highest ATP-dependent cleavage of the fluorogenic peptide, sLLVY-MCA, were pooled and used for SDS-PAGE separation of 26 S subunits

2.2. *Peptide sequencing*

For N-terminal sequence analysis, approximately 20 bands of subunit 7, each containing about 0.5 μg protein, were excised from a 10% acrylamide SDS gel (see Fig. 1). The gel slices were cut in small pieces (~ 1 × 1 mm) and the protein was eluted overnight [18]. Aliquots of 200–400 μl of the eluant were centrifuged in ProSpin Tubes (ABI) to immobilize the protein on a PVDF membrane. The protein bound to the PVDF membrane was sequenced directly on an ABI gas-phase sequenator. Internal peptide sequences were obtained by treating the

Published by Elsevier Science Publishers B V

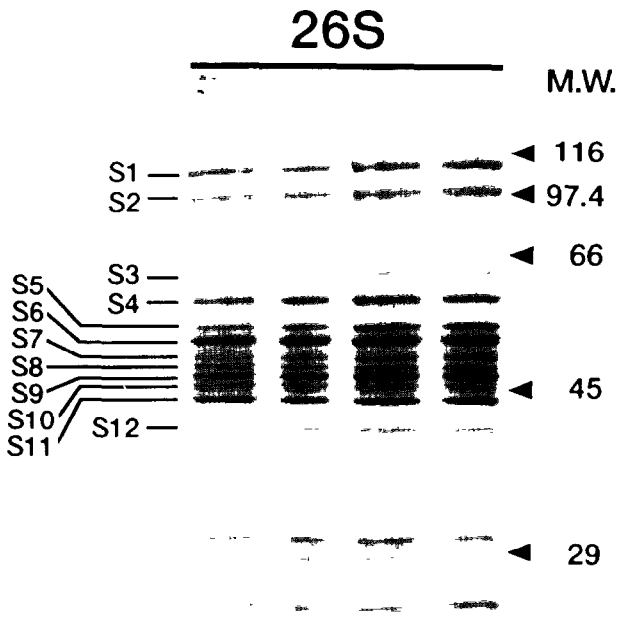

Fig. 1. SDS-PAGE of the 26 S protease. S1 to S12 denote subunits of the 26 S regulatory complex. Protein eluted from multiple bands of subunit 7 was either sequenced directly or digested with CNBr and processed as described in section 2.

gel slices directly with a solution of CNBr in 70% formic acid [18]. After microfiltration, resultant peptides were separated by reverse-phase HPLC on a C_{18} column (Vydac). Fractions of approximately 150 μl were collected at 1 min intervals (see Fig. 2). The indicated peaks were rechromatographed on a C_8 column, and selected peaks were sequenced as above. Peptide sequences were compared and aligned with the aid of the FASTA program [19].

3. RESULTS AND DISCUSSION

Fig. 1 shows a typical SDS-PAGE pattern of the human erythrocyte 26 S protease. Peptide mapping and sequence analysis (data not shown) revealed at least 10 different polypeptides in the molecular mass region from 40 kDa to 110 kDa, forming the regulatory complex of the 26 S protease. Direct sequencing of the protein band designated S7 resulted in the 22 amino acids labeled 1 in Fig. 3. A comparison of this sequence to known proteins in Genbank revealed a 100% match with the N-terminus of MSS1 except for the N-terminal methionine. Encouraged by these initial results we obtained internal peptide sequences from S7. Two internal peptides obtained after CNBr cleavage and reverse-phase HPLC were analyzed (see Fig. 2). The resulting sequences from peptides 2 and 3 also match exactly with MSS1. From the alignments in Fig. 3, it can be seen that a methionine precedes peptides 2 and 3 in the MSS1 sequence, as expected from CNBr cleavage. Peptides 1 and 2 match the hypothetical translation product of MSS1 in regions of low homology among the S4-like proteins [9], the best area for identification of a single member of this highly conserved subfamily. Peptide 1 ranges from 0% homology with SUG1, the N-terminus of which does not extend to this area, to 17% identity with subunit 4 of the human 26 S protease. Peptide 2, ranges from 15% identity with S4 to 40% identity with TBP1. Peptide 3 is a 12 residue peptide in the highly conserved ATPase module [9], but it still does not match exactly with any member of the subfamily except MSS1. Considering the fact that MSS1 is 433 residues long [17], the peptide sequences shown in Fig. 3 cover more than 12% of the MSS1 protein. In addition, the estimated molecular mass for MSS1 of 48,633 Da [17] is consistent with the SDS-PAGE molecular mass for S7 (49 kDa).

The results presented here provide strong evidence that S7 of the regulatory complex is identical to MSS1. The fact that S7 is missing its N-terminal methionine contrasts with most of the other 26 S protease subunits, which have blocked N-termini (data not shown). This may be caused by aminopeptidase activity during the preparation of the enzyme. As seen in Fig. 1, S7 appears as a doublet in SDS-PAGE. It is possible that only a portion of S7 subunits lack the N-terminal methionine. Our studies on the primary structure of subunit 4 of the 26 S protease and on S4-like proteins [9] led us to conclude that MSS1, TBP1 and SUG1 are, with high probability, all subunits of the 26 S ATP-dependent protease. Here we have presented strong evidence that MSS1 is a subunit of the human erythrocyte 26 S enzyme. Attempts to identify other members of the S4-like proteins in the 26S complex from human erythrocytes by peptide sequence analysis, however, have not yet been successful. This can be explained by assuming that 26 S complexes from different tissues and cell compartments (cytosol or nucleus) have modified subunit compositions that impart specificity.

The notion that the 26 S protease is not only localized in the cytosol but also in the nucleus is supported by studies on S4, which has a nuclear targeting sequence [9], on MSS1 (= S7), which has been found predomi-

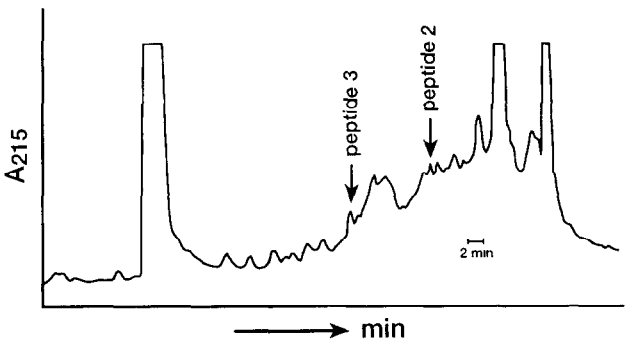

Fig. 2. Reverse-phase HPLC separation on Vydac C_{18} of peptide fragments resulting from CNBr digestion of subunit 7. Peaks 2 and 3 were further purified on a C_8 column and analyzed on an ABI sequenator, and resulted in the corresponding amino acid sequences in Fig. 3.

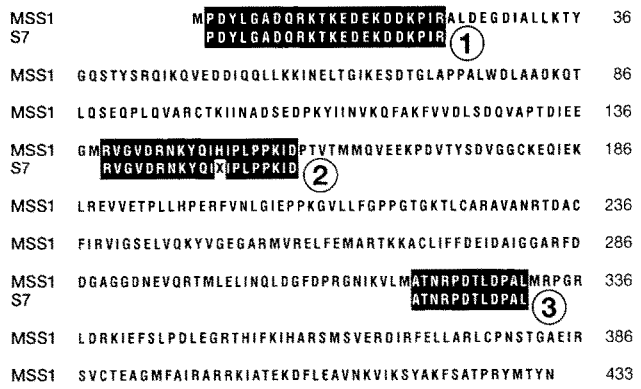

Fig. 3. Amino acid identity between MSS1 sequence and peptides from subunit 7 of the 26 S protease. Peptide 1 was sequenced directly from S7 transferred to the PVDF membrane. Peptides 2 and 3 were obtained by CNBr digestion

nantly in the nucleus [17], and by the fact that MCP, the proteolytic core of the 26 S protease, has also been localized to both nucleus and cytoplasm [20,21]. It has been suggested that MSS1, TBP1, TBP7 and SUG1 are members of a new family of transcription factors [13,14]. We cannot exclude the possibility that MSS1 has more than one function. Conceivably, in the 26 S complex the protein is involved in protease reactions, and in other complexes it might act as a transcription factor. Data supporting a transcriptional role for MSS1 and also for TBP1 and SUG1, however, can be interpreted as their functioning as proteolytic components. In fact, MSS1 presumably could not act as a transcription factor in erythrocytes since they lack a nucleus.

The finding that MSS1 is a 26 S protease subunit suggests that the enzyme itself may regulate transcription. In the case of HIV gene expression, the 26 S protease might be a modulator of HIV Tat-mediated transactivation by specific degradation of proteins that regulate transcription. It is reasonable to assume that a major part of the 26 S regulatory complex is an ATPase complex composed of proteins like S4 and MSS1 (= S7). There is a clear analogy to the *E. coli* Clp protease where the ATPase subunit, Clp A, forms homohexameric ATPase complexes in the presence of MgATP [22]. The 26 S ATPase appears to be a heteromeric complex that may function as a protein pump by moving substrates into the proteolytic core (MCP) of the 26 S protease.

Acknowledgements. These studies were supported by National Institutes of Health Grant GM 37009. We thank Laura Hoffman for helpful comments, and Linda Van Orden for excellent word processing.

REFERENCES

[1] Hough, R., Pratt, G and Rechsteiner, M. (1986) J. Biol. Chem. 261, 2400–2408.
[2] Hough, R., Pratt, G. and Rechsteiner, M (1987) J Biol. Chem. 262, 8303–8313
[3] Hough, R., Pratt, G and Rechsteiner, M. (1988) in Ubiquitin (Rechsteiner, M Ed.) pp. 207–238, Plenum Press, New York.
[4] Eytan, E., Ganoth, D., Armon, T. and Hershko, A. (1989) Proc. Natl. Acad Sci. USA 86, 7751–7755.
[5] Driscoll, J. and Goldberg, A.L (1990) J Biol. Chem. 265, 4789–4792.
[6] Peters, J.M., Harris, J.R. and Kleinschmidt, J.A. (1991) Eur. J. Cell Biol 56, 422–432.
[7] Hoffman, L., Pratt, G. and Rechsteiner, M. (1992) J. Biol. Chem. 267, 22362–22368.
[8] Ganoth, D., Leshinsky, E., Eytan, E. and Hershko, A. (1988) J. Biol Chem. 263, 12412–12419.
[9] Dubiel, W., Ferrell, K., Pratt, G. and Rechsteiner, M (1992) J. Biol. Chem. 267, 22699–22702.
[10] Erdmann, R., Wiebel, F F., Flessau, A., Rytka, J , Beyer, A., Fröhlich, K.U. and Kunau, W.H. (1991) Cell 64, 499–510.
[11] Fröhlich, K.U., Fries, H.W., Rüdiger, M , Erdmann, R., Botstein, D. and Mecke, D (1991) J. Cell Biol. 114, 443–453.
[12] Nelbock, P., Dillon, P.J , Perkins, A and Rosen, C.A. (1990) Science 248, 1650–1653
[13] Ohana, B., Moore, P.A., Ruben, S.M., Southgate, C.D., Green, M.R. and Rosen, C.A. (1993) Proc Natl. Acad. Sci. USA 90, 138–142
[14] Swaffield, J.C., Bromberg, J.F. and Johnston, S.A (1992) Nature 357, 698–700
[15] Goyer, C., Lee, H S., Malo, D. and Sonenberg, N. (1992) DNA Cell Biol. 11, 579–585.
[16] Irie, K., Nomoto, S., Miyajma, I and Matsumoto, K. (1991) Cell 65, 785–795.
[17] Shibuya, H., Irie, K., Ninomiya-Tsuji, J., Goebl, M , Taniguchi, T. and Matsumoto, K. (1992) Nature 357, 700–702.
[18] Kurth, J. and Stoffel, W (1990) Biol. Chem. Hoppe-Seyler 371, 675–685.
[19] Pearson, W R. and Lipman, D.J. (1988) Proc Natl. Acad. Sci USA 85, 2444–2448.
[20] Tanaka, K , Kumatori, A., Ii, K. and Ichihara, A. (1989) J Cell. Physiol. 139, 34–41.
[21] Amsterdam, A., Pitzer, F. and Baumeister, W (1993) Proc. Natl. Acad. Sci. USA 90, 99–103.
[22] Maurizi, M.R (1992) Experientia 48, 178–201.