Molecular Cloning and Expression of a 26 S Protease Subunit Enriched in Dileucine Repeats*

Quinn Deveraux, Chris Jensen, and Martin Rechsteiner†

From the Department of Biochemistry, University of Utah School of Medicine, Salt Lake City, Utah 84132

The 26 S protease is a multisubunit enzyme required for ubiquitin-dependent proteolysis. Recently, we identified a 50-kDa subunit (S5) of this enzyme that binds ubiquitin polymers (Deveraux, Q., Ustrell, V., Pickart, C., and Rechsteiner, M. (1994) J. Biol. Chem. 269, 7059–7063). We have now isolated, sequenced, and expressed a cDNA encoding a novel 50-kDa subunit of the 26 S protease. The recombinant protein does not bind ubiquitin polymers. Two-dimensional electrophoresis reveals that two subunits of the 26 S protease have apparent molecular masses of 50 kDa. Antibodies specific for the recombinant protein recognize the more basic of the two subunits (S5b), whereas the more acidic subunit (S5a) binds ubiquitin chains. Thus, the 26 S protease contains at least two distinct subunits with apparent molecular masses of 50 kDa.

Ubiquitin has been implicated in diverse cellular phenomena including proteolysis of key regulatory molecules such as GCN4 and some cyclins (1–5). Although ubiquitin may have nonproteolytic roles, covalent attachment of ubiquitin to proteins clearly produces substrates for the 26 S ATP-dependent protease. This enzyme is composed of the multicatalytic protease or proteasome and a regulatory ATPase complex (6). Both multicatalytic protease and the regulatory complex are multisubunit structures that associate in the presence of ATP to form the 26 S enzyme (7). Although the multicatalytic protease can degrade a variety of peptides, few intact proteins have been shown to be substrates. Subunits of the regulatory complex are, therefore, thought to confer the ability to recognize and degrade intact proteins, especially those conjugated to ubiquitin chains. In support of this idea, we identified a 50-kDa subunit of the regulatory complex that binds ubiquitinated lysozyme as well as free polymers of ubiquitin (8). We called this protein subunit 5 (S5) based upon its relative mobility on SDS-polyacrylamide gels (9). S5 efficiently binds tetrameric ubiquitin and selects for longer ubiquitin polymers (8). This property is consistent with characteristics expected of a component that selects ubiquitin conjugates for proteolysis since it has been shown that β-galactosidase molecules attached to long ubiquitin chains are degraded more efficiently than those conjugated to shorter chains (10). As a further characterization of the ubiquitin-binding subunit, we have employed standard procedures to obtain a cDNA encoding the 50-kDa subunit of the 26 S protease. We have found that two distinct proteins comprise the band previously designated as S5 on SDS gels. Here we distinguish between the two 50-kDa proteins and describe the cloning, sequence, and expression of a cDNA encoding the more basic subunit, S5b.

EXPERIMENTAL PROCEDURES

Peptide Sequencing—Human red blood cell 26 S proteases (11) were resolved on SDS gels, and the subunits were transferred to polyvinylidenefluoride (PVDF)1 membranes (9). The PVDF membrane was stained with Ponceau S, and the S5 band was excised. Direct sequencing of S5 indicated that it has a blocked amino terminus. Therefore, S5 bands were excised from PVDF membranes and treated with CNBr. The resulting peptides were eluted from the membrane, electrophoresed on SDS gels, and transferred to PVDF. The S5 fragments were then excised and sequenced directly using an ABI gas phase sequenator. Amino acid sequences were obtained from three of the isolated peptides.

Isolation of a cDNA Encoding S5b—Degenerate oligonucleotides derived from two of the S5 peptides were used in the polymerase chain reaction to synthesize a non-overlapping probe. The probe was labeled by the random prime method and was used to isolate two clones from a λ ZAP II HeLa cDNA library (Stratagene). Both cDNAs encoded all three of the peptides obtained from S5 digestion. Expression of the these cDNAs, however, resulted in products smaller than the 50-kDa subunit 5 of the 26 S protease. To obtain the complete coding region for S5, we employed the 5′ rapid amplification of cDNA ends (5′-RACE) procedure (Life Technologies, Inc.). Numerous subclones containing the 5′-RACE products were sequenced to identify polymerase chain reaction errors prior to construction of the full-length cDNA. The 5′-RACE products were similar in length, indicating the reverse transcriptase terminated at a single position. The longest open reading frame predicted a 56-kDa protein. This full-length construct was subcloned into pAE4 (12), and both sense and antisense DNA strands were sequenced using the Sequenase protocol (U. S. Biochemical Corp.). The complete sequence was compared with a cDNA sequence submitted to GenBank (accession number D31889) and found to be identical except for the initiating methionine.

Expression and Purification of S5b—The full-length S5b construct was expressed in E. coli by T7 polymerase and was translated in reticulocyte lysate containing [35S]methionine (Stratagene). For in vivo expression, the same plasmid was transformed into BL21 cells and was induced with 1 mM isopropyl-1-thio-β-galactopyranoside (IPTG) under standard conditions. An insoluble fraction was prepared following sonication in lysis buffer (50 mM Tris, pH 8.0, 50 mM NaCl, 0.25% Triton X-100) by centrifugation for 30 min at 16,000 × g. The insoluble material was washed twice with lysis buffer and then suspended in sample buffer and boiled for 2 min prior to electrophoresis. Induction of authentic S5b was confirmed by NH2-terminal amino acid sequence analysis of protein resolved on SDS gels and transferred to PVDF. Purified S5b was obtained by subcloning the S5b sequence into the pET 16b vector (Novagen), which places a histidine-containing leader sequence at its amino terminus. The pET 16b vector containing the S5b sequence was transformed into BL21 cells, expressed, and purified as recommended by Novagen.

Antibody Preparation, Western Blotting, and Ubiquitin-Polymer Binding Assays—The purified histidine-S5b fusion protein was used for antibody production in rabbits using Titermax as described by the manufacturer (Vaxcel, Inc.). These antibodies react with both the histidine-

* This work was supported by National Institutes of Health Grant GM37009 and by a grant from the Lucile P. Markey Charitable Trust. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

† To whom correspondence and reprint requests should be addressed.

1 The abbreviations used are: PVDF, polyvinylidene difluoride; 5′-RACE, 5′ rapid amplification of cDNA ends.

2 The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.
A Leu/Asp-rich 26 S Protease Subunit

RESULTS AND DISCUSSION

Subunit 5 was identified as a ubiquitin recognition component of the 26 S protease by its ability to bind ubiquitin-labeled protein conjugates. Subsequent analyses using a mixture of ubiquitin polymers of different lengths revealed that S5 selects for longer ubiquitin polymers even after SDS denaturation of the subunit (8). In addition, our calculations of subunit stoichiometry suggest that only one S5 molecule is present per regulatory complex. In view of these findings, we hypothesized that repeated sequences in S5 might bind multiple sites present in ubiquitin chains. This idea can be tested by obtaining the subunit's sequence. Consequently, we used a combination of polymerase chain reaction, cDNA library screening, and 5'-RACE to obtain a composite full-length nucleotide sequence encoding the S5 subunit (see "Experimental Procedures").

The expressed protein, however, does not bind ubiquitin polymers (data not shown). This finding raised the possibility that more than one 26 S protease subunit comprises the band designated as S5 on SDS gels. This was tested by resolving...
components of the 26 S protease on two-dimensional gels. Proteins were transferred to nitrocellulose and incubated with either \(^{125}\text{I}\)-labeled ubiquitin polymers or antibodies specific for the recombinant S5 protein. As shown in Fig. 3, the labeled ubiquitin polymers bound to a 50-kDa subunit with a pl of 4.6 (panel B, white arrow), and the anti-S5 antibodies recognized a protein focusing at pH 5.3 (panel C, black arrow). These results demonstrate that there are two distinct subunits of the 26 S protease in the band previously called S5. We have termed the two 50 kDa proteins S5a (acidic) and S5b (basic). The cDNA sequence shown in Fig. 1 codes for S5b.

The three CNBr peptides initially sequenced from the S5 band are present in the cDNA encoding S5b. However, the data in Fig. 3 clearly reveal two distinct 50-kDa proteins at the S5 position. For this reason, we separated large amounts of the regulatory complex subunits on SDS-gels and asked whether peptides other than those present in S5b could be identified. By excising proteins that migrated slightly above the major S5 band, we obtained the sequences of eight endoproteinase Lys-C digested peptides, three of which are not encoded by S5b. These three peptides (RIAFVGSPVELDTK, VNVDIINGFEEVNE, and AGTGSVLVTYPVG) are homologous to a recently identified multibuquin-binding protein from Arabidopsis thaliana.\(^2\) The identification of peptides encoded by two distinct cDNAs confirms the presence of two 50-kDa subunits in the 26 S protease.

When 26 S protease and regulatory complexes were electrophoresed on native gels and transferred to nitrocellulose, antibodies raised against S5b recognize both native particles. Similarly, these antibodies identified a 50-kDa protein after subunits of each native complex were resolved by SDS-PAGE (data not shown). Therefore, S5b is an integral component of the 26 S protease. The function of subunit 5b, however, is not currently understood. Its sequence is enriched in leucine residues; 26 of the first 100 amino-terminal residues are leucines. Interestingly, the residues surrounding the dileucine pairs in S5b are also highly enriched in charged and polar side chains (Table I). Conceivably, the presence of dileucine repeats has been implicated in the negative regulation of the platelet-derived growth factor \(\beta\)-receptor, possibly by promoting the degradation of the ligand-activated receptor (16). In addition, the apparent role of the 26 S protease in antigen presentation (17) might require its direct association with the endoplasmic reticulum. Peptides generated by the protease would then be immediately available for transport.

\(^2\) S. van Nocker, Q. Deveraux, M. Rechsteiner, and R. Vierstra, submitted for publication.

### Table I

A comparison of dileucine repeats in S5b and various transmembrane proteins

| Human S5b dileucine pairs | Human S5b dileucine pairs |
|---------------------------|---------------------------|
| LIMPR (rat)               | RALLRE                   |
| CD4 (human)               | LALLRE                   |
| CD-M6PR/IF-1R (bovine)    | LALLRE                   |
| CD-M6PR (bovine)          | LALLRE                   |
| CD3 g-chain (mouse)       | LALLRE                   |
| CD3 d-chain (mouse)       | LALLRE                   |
| Human insulin receptor    | LALLRE                   |

\[\text{residues 476-479}\]
across the endoplasmic reticulum membrane.

In summary, we have isolated a cDNA that encodes a 50-kDa subunit of the 26 S protease. Although we cannot currently assign a function to S5b, the availability of both antibodies to S5b and a histidine-tagged version of the recombinant protein should facilitate future characterization of this 26 S protease component.

REFERENCES

1. Kornitzer, D., Raboy, B., Kulka, R. G., and Fink, G. R. (1994) EMBO J. 13, 6021–6030
2. Deshaies, R. J., Chau, V., and Kirschner, M. (1995) EMBO J. 14, 303–312
3. Glotzer, M., Murray, A. W., and Kirschner, M. W. (1991) Nature 349, 132–138
4. Ciechanover, A. (1994) Cell 79, 13–21
5. Hershko, A., Gandot, D., Pehrson, J., Palazzo, R. E., and Cohen, L. H. (1991) J. Biol. Chem. 266, 16376–16379
6. Rechsteiner, M., Hoffman, L., and Dubé, W. (1993) J. Biol. Chem. 268, 6065–6068
7. Hoffman, L., Pratt, G., and Rechsteiner, M. (1992) J. Biol. Chem. 267, 22362–22368
8. Deveraux, Q., Ustrell, V., Pickart, C., and Rechsteiner, M. (1994) J. Biol. Chem. 269, 7059–7061
9. Dubé, W., Ferrel, K., Pratt, G., and Rechsteiner, M. (1992) J. Biol. Chem. 267, 22699–22702
10. Chau, V., Tobias, J. W., Bachmair, A., Marriott, D., Ecker, D. J., Gonda, D. K., and Varshavsky, A. (1989) Science 243, 1576–1583
11. Hough, R., Pratt, G., and Rechsteiner, M. (1987) J. Biol. Chem. 262, 8303–8313
12. Realini, C., Dubé, W., Pratt, G., Ferrel, K., and Rechsteiner, M. (1994) J. Biol. Chem. 269, 20727–20732
13. Deveraux, Q., Wells, R., and Rechsteiner, M. (1990) J. Biol. Chem. 265, 6323–6329
14. Pickart, C. M., Haldeman, M. R., Kasperk, E. M., and Chen, Z. (1992) J. Biol. Chem. 267, 14418–14423
15. Haft, C. R., Klausner, R. D., and Taylor, S. I. (1994) J. Biol. Chem. 269, 26286–26294
16. Mori, S., Heldin, C.-H., and Claesson-Welsh, L. (1993) J. Biol. Chem. 268, 577–583
17. Germain, R. N. (1994) Cell 76, 287–299