AMINO ACID SEQUENCE INVOLVING THE REACTIVE THIOL GROUP OF FICIN[1]

Robert C. Wong and Irvin E. Liener

Department of Biochemistry

University of Minnesota

St. Paul, Minnesota

Received August 28, 1964

In contrast to the detailed information which we now have on the amino acid sequence in the vicinity of the active serine residue of the animal proteases (1), our knowledge concerning the active site of the plant proteases is very limited. Because of their sensitivity to thiol reagents, most proteolytic enzymes of plant origin, such as papain, ficin, and bromelain, have been generally regarded as enzymes in which a free SH group is required, either directly or not indirectly, for activity (2). Essential to an understanding of the mechanism whereby such SH groups participate in catalysis is specific information concerning the chemical nature of their immediate environment in the protein molecule. This paper reports a portion of the amino acid sequence which is adjacent to the SH group upon which the catalytic activity of ficin is dependent (3).

The preparation and properties of the ficin used in this study has already been described (3). When ficin was treated with a two-fold molar excess of iodoacetate-1-C^{14} (2 mc/mmole) at pH 5.5 in the presence of 0.001 N KCN, no enzymatic activity could be detected after 15 min. Amino acid analysis of the acid hydrolyzed protein (4) showed CM-cysteine[2] (0.6-0.7 mole/mole ficin) to be the only amino acid residue which had been

[1] Paper No. 5504, Scientific Journal Series, Minnesota Agricultural Experiment Station. This investigation was supported by grants from the National Institutes of Health, U. S. P. H. (GM 04616) and the National Science Foundation (G 13965).

[2] Abbreviations used: CM, carboxymethyl; $CySO_3H$, cysteic acid; $CMCySO_2$, CM-cysteine sulfone.

carboxymethylated. When the acid hydrolyzate was subjected to high voltage paper electrophoresis (1 hr, 1500 volts, pH 2) and scanned for radioactivity, only one radioactive component corresponding to CM-cysteine could be detected. The radioactivity per mole of CM-cysteine was based on the amount of CM-cysteine recovered from acid hydrolyzed $C^{14}M$-ficin.

$C^{14}M$-ficin was digested with crystalline pepsin (Worthington) at pH 2, 37°, for 10 hrs (enzyme:substrate ratio, 1:50). In order to avoid the complications produced by the oxidative decomposition of CM-cysteine that accompanies attempts to separate peptides containing this residue (5,6), the pepsin digest was oxidized with performic acid (7). This treatment oxidized CM-cysteine to CM-cysteine sulfone, which is stable during chromatography and electrophoresis, and converts half-cystine residues to cysteic acid. Since CM-cysteine sulfone was the only radioactive component present, a correction for its destruction during acid hydrolysis was made by counting the sample before and after hydrolysis. Although CM-cysteine sulfone and cysteic acid are both eluted in the hold-up volume during chromatography (4), the amount of CM-cysteine sulfone present in this peak can be calculated from the radioactivity of the total hydrolyzate, and the cysteic acid estimated by difference.

When the oxidized digest was chromatographed on Dowex-50X2, as described by Schroeder et al (8), 6 radioactive peaks were obtained. The major peak, containing 25% of the original activity, was selected for further study. Paper electrophoresis showed this particular fraction to be composed of 4 ninhydrin-positive, non-radioactive peptides and one ninhydrin-negative, radioactive peptide designated as P_3. The latter could be separated from the other peptides by chromatography on Dowex-1X2 (formate form) using the gradient produced by passing 1.0 M pyridine-formate buffer, pH 3.4, into a 1:20 dilution of the same buffer. The amino acid composition of peptide P_3 is shown in Table I.

Three successive Edman degradations (9,10) established Pro-Ileu-Arg as the N-terminal sequence of peptide P_3. Tryptic digestion of this pep-

tide produced an octapeptide, P_3-T, which was isolated by chromatography on Dowex-1X2 and was shown to have the composition given in Table I.

Table I

Amino Acid Composition of Peptide P_3 and Peptides Derived Therefrom by Enzymatic Degradation

Amino acid	Peptide			
	P_3	P_3-T	P_3-S_1	P_3-LAP
Arginine	0.95	-	-	-
Cysteic acid	0.89	1.02	1.02	1.00
CM-Cysteine sulfone*	1.00	1.00	1.00	1.00
Glutamic acid	3.08	3.03	0.98	-
Glycine	2.05	2.20	2.00	1.05
Isoleucine	1.01	-	-	-
Proline	1.05	-	-	-
Serine	1.05	1.20	0.90	0.95

* Based on the radioactivity of the purified peptide prior to hydrolysis. Molar ratios of amino acids calculated on the basis that this residue was 1.0.

After one stage of Edman degradation, glutamic acid (or glutamine) was lost from peptide P_3-T. When peptide P_3-T was digested with subtilisin (Nagarse), two peptides, P_3-S_1 and P_3-S_2, could be separated by chromatography on Dowex-1X2. Peptide P_3-S_1 was radioactive and its composition showed that it contained two fewer glutamic acid (or glutamine) residues than peptide P_3-T. Peptide P_3-S_2, which was non-radioactive, yielded glutamic acid after acid hydrolysis and only glutamine after digestion with leucine aminopeptidase. These results extend the N-terminal sequence of peptide P_3 to Pro-Ileu-Arg-GluNH_2-GluNH_2.

One Edman degradation of peptide P_3-S_1 established glycine as the N-terminal amino acid of this peptide. Leucine aminopeptidase released 1 mole of glycine and 0.8 mole of glutamine from peptide P_3-S_1, and a new radioactive tetrapeptide, P_3-LAP, could be isolated from the digest by high-voltage electrophoresis. Two successive Edman degradations of peptide P_3-LAP led to the loss of cysteic acid and glycine in that order.

Serine disappeared after the third stage of degradation, and a radioactive compound corresponding to CM-cysteine sulfone was obtained when the aqueous phase was examined by high voltage paper electrophoresis. Peptide P_3-S_1 thus has the sequence: Gly-GluNH_2-CySO_3H-Gly-Ser-CMCySO_2.

The evidence which led to the reconstruction of the complete sequence of peptide P_3 is summarized in Fig. 1. It is significant to note that the amino acid sequence shown here is strikingly similar to that recently reported as being adjacent to the active SH group of papain (11,12):

Pro-Val-Lys-Asp(NH_2)-Glu(NH)$_2$-Gly-Ser-CyS-Gly-Ser-CyS.

The amino acids which have been underlined are those which have been replaced by other amino acids in ficin. These replacements are of such a nature that there is no change in the charge distribution in this particular sequence of amino acids. This similarity in amino acid sequence in the vicinity of the reactive thiol group suggests that this region of the molecule may play a unique role in the catalytic function of the SH proteases.

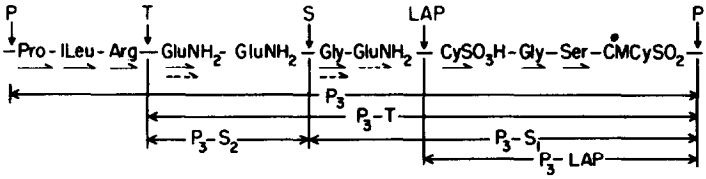

Fig. 1. Sequence of amino acids in peptide P_3 isolated from peptic digest of CM-ficin. P, T, S, and LAP show sites of cleavage by pepsin, trypsin, subtilisin, and leucine aminopeptidase respectively. Solid arrow shows sequence determined by Edman degradation and dashed arrow by LAP. * denotes radioactive residue.

References

1. Koshland, D. E., Jr., Science, 142, 1533 (1963).

2. Boyer, P. D., in The Enzymes, vol I, edited by P. D. Boyer, H. Lardy, and K. Myrback. Academic Press, Inc., New York, 1960, p. 511.

3. Liener, I. E., Biochim. Biophys. Acta, 53, 332 (1961).

4. Spackman, D. H., Stein, W. H., and Moore, S., Anal. Chem., 30 1190 (1958).

5. French, T. C., Dawid, I. B., and Buchanan, J. M., J. Biol. Chem., 238, 2186 (1963).

6. Li, T.-K. and Vallee, B. L., Biochemistry, 3, 869 (1964).

7. Hirs, C. H. W., J. Biol. Chem., 219, 611 (1956).

8. Schroeder, W. A., Jones, R. T., Cormik, J., and McCalla, K., Anal. Chem., 34, 1570 (1962).

9. Konigsberg, W. and Hill, R. J., J. Biol. Chem., 237, 2547 (1962).

10. Van Orden, H. O. and Carpenter, F. H., Biochem. Biophys. Res. Commun., 14, 399 (1964).

11. Light, A. and Smith, E. L. Abstracts VIth International Congress of Biochemistry, New York, 1964.

12. Light, A., Frater, R., Kimmel, J. R., and Smith, E. L., Proc. Natl. Acad. Sci., 52, 1964, in press.