The amino acid sequence of a pentadecapeptide obtained from a tryptic digest of the protein of tobacco mosaic virus

In recent communications[1,2] the countercurrent distribution of a tryptic digest of the protein of tobacco mosaic virus and the elucidation of the structures of six of these components were reported. The present communication presents the amino acid sequence of a peptide, designated K-0.7-A, with the amino acid composition of Ser_2, Glu_3, Val_3, Pro_2, Phe, Try, Thr, Lys, Arg, isolated by electrophoresis on cellulose at pH 5.5 from the fraction with distribution coefficient of about 0.7. The peptide gave a positive ninhydrin test (due to the ε-amino group of lysine) but failed to yield an N-terminal amino acid with a free amino group by either the dinitrophenyl or the phenylthiohydantoin method.

When the peptide was hydrolyzed with chymotrypsin the peptides, C-1 through C-7, shown in Table I were obtained by a combination of paper chromatography in *n*-butanol–acetic acid–water–pyridine (30 : 6 : 24 : 20) and paper electrophoresis at pH 5.5. The peptides C-1 and C-7 were resistant to leucine aminopeptidase and carboxypeptidase and the sequences were determined by partial acid hydrolysis and analysis of the fragments by the dinitrophenyl method. The sequence of C-4 could be deduced from the N-terminal amino acids and compositions of C-3 and C-4. This was confirmed by two enzymic procedures using leucine aminopeptidase and a mixture of carboxypeptidase-A and carboxypeptidase-B[3]. The peptide C-4 was shown to be the C-terminal peptide from K-0.7-A by treating the latter with a mixture of the two carboxypeptidases.

TABLE I

SEQUENCES OF PEPTIDES FROM A CHYMOTRYPTIC DIGEST OF K-0.7-A

Peptide	*Amino acid sequence*
C–1	NH_2 Lys.Pro.Ser.Pro.Glu.
C-2	NH_2 Lys.Pro.Ser.Pro.Glu.Val.Thr.
C-3	Val.Thr.
C-4	Val.Thr.Val.Arg.
C-5	NH_2 Ser.Glu.Val.Try.
C-6	NH_2 Pyroglu.Phe.Ser.Glu.Val.Try.
C-7	Pyroglu.Phe.
K-0.7-A	NH_2 NH_2 Pyroglu.Phe.Ser.Glu.Val.Try.Lys.Pro.Ser.Pro.Glu.Val.Thr.Val.Arg.

The tryptophanyl-lysine bond was established by treating K-0.7-A with N-bromosuccinimide or N-bromoacetamide as described by PATCHORNIK, LAWSON AND WITKOP[4] and isolation of the two fragments. Both of the peptides C-6 and C-7 were ninhydrin-negative and failed to react with fluorodinitrobenzene. It was apparent from the above data that these peptides were derived from the N-terminal end of K-0.7-A. The identity of pyrrolidone carboxylic acid (pyroglutamic acid) as the N-terminal residue of peptides C-6 and C-7, and therefore of K-0.7-A, was indicated by the above data and the fact that the peptide Glu.Phe was isolated from the partial acid hydrolyzate of peptide C-6. It is believed that the N-terminal amino acid of K-0.7-A was converted to pyrrolidone carboxylic acid during or subsequent to tryptic digestion of the protein. The conversion of glutamic acid and glutamine to pyrrolidone carboxylic acid under acidic conditions is well-known[5–10] and the solvent used in the countercurrent distribution for the initial separation of the components of the tryptic digest had a pH of about 2. An analysis of the tryptic digest indicated that one N-terminal glutamic acid peptide was present in the digest[1].

From the electrophoretic mobilities of K-0.7-A, C-1 and C-5 it appeared that both of the internal residues of glutamic acid were present as glutamine residues. This was also indicated by the fact that three moles of ammonia were found in an acid hydrolyzate by the chromatographic method of SPACKMAN, STEIN AND MOORE[11] and by the method of STONE[12]. The third mole of ammonia may have been derived from the destruction of amino acids during hydrolysis. Tryptophan was determined by the spectrophotometric method of GOODWIN AND MORTON[13] and the colorimetric method of SPIES AND CHAMBERS[14] prior to hydrolysis. The complete structure of K-0.7-A is shown in Table I.

Of considerable interest is the fact that this peptide has a lysine bond that is resistant to trypsin. The ε-amino group is free since ε-DNP-lysine was found after dinitrophenylation and hydrolysis of K-0.7-A. It had been previously reported, that in denatured tobacco-mosaic-virus protein only one of the two ε-amino groups of lysine is reactive towards the fluorodinitrobenzene reagent[12] or O-methylisourea[16]. Even in the tryptic digest of this protein only one ε-amino group reactive towards fluorodinitrobenzene was found[1]. The peptide Phe.Pro.Asp.Phe.Ser.Lys, bearing an ε-amino group reactive towards fluorodinitrobenzene, had already been isolated from the tryptic digest of tobacco-mosaic-virus protein[2]. The fact that the lysine bond of K-0.7-A was intact after treatment of the protein with trypsin and that apparently only one of the two ε-amino groups of lysine in tobacco-mosaic-virus protein is reactive towards fluorodinotrobenzene and O-methylisourea indicates that the ε-amino group of K-0.7-A may be masked in tobacco-mosaic-virus protein. However, it may be pointed out that SHEPHARD *et al.*[17], and HIRS *et al.*[18] have reported peptides containing lysylproline bonds resistant to trypsin and tyrosylproline and phenylalanylproline bonds resistant to chymotrypsin. Further work is now in progress on these unique aspects of the structure of this peptide.

Appreciation is expressed to Mrs. GLADYS PEREZ-MENDEZ and Miss MARTHA LANGLEY for their capable technical assistance.

This paper has been aided by the U.S. Public Health Service Grant.

Virus Laboratory, University of California, Berkeley, Calif. (U.S.A.) DUANE T. GISH

[1] D. T. GISH, L. K. RAMACHANDRAN AND W. M. STANLEY, *Arch. Biochem. Biophys.*, 78 (1958) 433.
[2] L. K. RAMACHANDRAN AND D. T. GISH, *J. Amer. Chem. Soc.*, 81 (1959) 884.
[3] J. E. FOLK AND J. A. GLADNER, *J. Biol. Chem.*, 231 (1958) 398.
[4] A. PATCHORNIK, W. B. LAWSON AND B. WITKOP, *J. Amer. Chem. Soc.*, 80 (1958) 4747.
[5] A. C. CHIBNALL AND R. G. WESTALL, *Biochem. J.*, 26 (1932) 122.
[6] H. WILSON AND R. CANNAN, *J. Biol. Chem.*, 119 (1937) 309.
[7] P. B. HAMILTON, *J. Biol. Chem.* 158 (1945) 375.
[8] K. NARITA *Biochim. Biophys. Acta*, 30 (1958) 352.
[9] H. WILSON AND R. CANNAN, *J. Biol. Chem.*, 119 (1937) 309.
[10] P. B. HAMILTON, *J. Biol. Chem.*, 158 (1945) 375.
[11] D. H. SPACKMAN, W. H. STEIN AND S. MOORE, *Anal. Chem.*, 30 (1958) 1190.
[12] W. E. STONE, *Proc. Soc. Exptl. Biol. Med.*, 93 (1956) 589.
[13] T. W. GOODWIN AND R. A. MORTON, *Biochem. J.*, 40 (1946) 628.
[14] J. R. SPIES AND D. C. CHAMBERS, *Anal. Chem.*, 21 (1949) 1249.
[15] H. FRAENKEL-CONRAT AND B. SINGER, *J. Amer. Chem. Soc.*, 76 (1954) 180.
[16] L. K. RAMACHANDRAN, *Biochim. Biophys. Acta*, 32 (1959) 557.
[17] R. G. SHEPHERD, K. S. HOWARD, D. S. DAVIES, E. A. EIGNER AND N. E. SHAKESPEARE, *J. Am. Chem. Soc.*, 78 (1956) 5067.
[18] C. H. W. HIRS, S. MOORE AND W. H. STEIN, *J. Biol. Chem.*, 219 (1956) 623.

Received July 20th, 1959

Enzymic formation of adenylserine and an unknown carboxyl-activated compound

The first step in protein synthesis is thought to be a carboxyl activation of amino acids. This reaction[1], which is analogous to that for activation of fatty acids, is as follows:

$$\text{Enzyme} + \text{ATP} + \text{amino acid} \overset{Mg^{++}}{\rightleftharpoons} \text{enzyme (adenyl amino acid)} + \text{pyrophosphate}$$

This communication deals with the identification of adenylserine and the observation of an unknown carboxyl-activated compound as products of a reaction mixture containing large amounts of a purified serine-activating enzyme[2].

20–50 mg enzyme, 0.2 μmole of DL-[3-^{14}C]serine containing 120,000 counts/min, 1–2 μmoles ATP, $MgCl_2$, Tris(hydroxymethyl)aminomethane buffer (pH 7.4), and crystalline pyrophosphatase were incubated in 2.7 ml at 37° for 5 min. After the enzyme was precipitated by cold trichloroacetic acid, the supernatant was adjusted to pH 3.0 and chromatographed at 4° on Dowex-1.

Two radioactive compounds were separated from serine (Fig. 1). Compound I (Peak I) chromatographed identically with synthetic adenylserine and contained 990–1200 counts/min. When incubated in neutral NH_2OH, it formed another compound which chromatographed on Dowex-50 identically with synthetic serine hydroxamate. Furthermore, Compound I incorporated label simultaneously from [8-^{14}C]ATP and [3-^{14}C]serine and the stoichiometry of this incorporation was 1:1. Radioactivity from either DL-[1-^{14}C]serine or [α-^{32}P]ATP was also incorporated into Compound I. When Compound I, derived from [α-^{32}P]ATP, was subjected to paper electrophoresis in 0.05 *M* citrate buffer, pH 3.9, the radioactivity migrated toward the cathode and coincided with the ultraviolet absorption of synthetic adenylserine. The foregoing evidence strongly suggests that Compound I was adenylserine. This represents the second instance in which an amino acid acyl adenylate has been isolated from an enzymic reaction mixture and identified[3,4].

Abbreviations: ATP, adenosine triphosphate; RNA, ribonucleic acid.