[CONTRIBUTION FROM THE VIRUS LABORATORY, UNIVERSITY OF CALIFORNIA, BERKELEY]

Studies on the Amino Acid Sequence of Tobacco Mosaic Virus (TMV) Protein. II. The Amino Acid Sequences of Six Peptides Obtained from a Tryptic Digest[1]

BY L. K. RAMACHANDRAN AND DUANE T. GISH

RECEIVED JUNE 25, 1958

The amino acid sequences of six peptides isolated from a tryptic digest of tobacco mosaic virus protein have been determined. These were found to be aspNH_2·arg, val·tyr·arg, thr·val·gluNH·val·arg, phe·pro·asp·phe·ser·lys, gly·thr·ser·gly·tyr·aspNH_2·arg and arg·val·asp·asp·ala·thr·val·ala·ileu·arg. These peptides contained 33 of the 164 amino acid residues in the protein sub-unit of TMV and included five with C terminal arginine and one with C-terminal lysine.

In a recent communication[2] the partial separation by countercurrent distribution of the components of a tryptic digest of tobacco mosaic virus (TMV) protein was described. The further purification and the elucidation of the structures of these components is now in progress and we wish to report the amino acid sequences of six of these peptides. Among the peptides were found a dipeptide, aspNH_2·arg,[3] a tripeptide, val·tyr·arg; a pentapeptide, thr·val·gluNH$_2$·val·arg; a hexapeptide, phe·pro·asp·phe·ser·lys, a heptapeptide, gly·thr·ser·gly·tyr·aspNH_2·arg and a decapeptide, arg·val·asp·asp·ala·thr·val·ala·ileu·arg. Thus 33 of the approximately 164 amino acid residues in the TMV protein sub-unit[4] have been accounted for in these six peptides. These peptides also account for six of the approximately eleven arginine residues and one of the two lysine residues believed to be present in the protein sub-unit.[4-6] It may also be noted that one of these peptides, the decapeptide, contains two arginine residues, one of which is N-terminal, the other C-terminal. This peptide (designated K-0.13) was recovered from the tryptic digest in an amount that would indicate that very few, if any, of the arginylvaline bonds in the portion of TMV protein from which this decapeptide was derived were split by trypsin. This peptide of molecular weight 1,115 would comprise 6.1% of the TMV protein sub-unit (mol. wt. 18,300).[4] However, since approximately 30% of the tryptic digest was in the form of sparingly soluble peptide material which was separated from the soluble portion of the digest,[2] this decapeptide would represent 8.9% of the soluble portion of the digest fractionated by countercurrent distribution. About 9.2% (68 mg.) of the approximately 740 mg. (780 mg. less 5% used for analytical procedures) recoverable from the distribution was recovered from the peak designated K-0.13, and this material was shown by paper chromatography and analysis to contain only small amounts of other peptide material. This fact reduces the number of arginine-containing peptides to be expected from a tryptic digest of TMV protein by one.

The ϵ-amino group of one of the two lysine residues of TMV protein appears to be masked and does not react with dinitrofluorobenzene.[7] It is possible, therefore, that only one peptide with C-terminal lysine will be found in a tryptic digest of the protein. A tryptic digest of TMV protein would thus be expected to contain twelve major peptide components, including the C-terminal peptide, ten peptides with C-terminal arginine and one peptide with C-terminal lysine. Minor amounts of other peptides can be expected due to some chymotryptic activity in the trypsin preparation and to non-specific splitting of peptide bonds during the countercurrent distribution and other techniques employed to separate and purify the various fractions.

The six peptides reported in this paper were recovered in an amount (see Table I) sufficient to justify the assumption that each is a product of tryptic hydrolysis of TMV protein and did not arise from transpeptidation or from an impurity. Furthermore, these six peptides occupied positions which coincided with the positions of six peptides in the pattern found for a tryptic digest of TMV protein when the digest was subjected to combined paper electrophoresis and paper chromatography as described by Knight.[8]

TABLE I

WEIGHT RECOVERIES OF VARIOUS COMPONENTS ISOLATED FROM 740 MG. OF THE SOLUBLE PORTION OF A TRYPTIC DIGEST OF TMV PROTEIN

Compound	Recovery of purified peptide, mg.	Cor. recovery, mg.	Calcd. recovery, mg.	% recovery, approximate
K-0.065-2	13	18	18	100
K-0.065-7	23	32	46	70
K-0.086-Ma	21	37	37	100
K-0.13	55	61	69	90
K-0.22-D	28	28	27	100
K-0.33	17	34	46	77

Experimental

Enzyme Preparations.—The leucine aminopeptidase employed was prepared essentially according to the procedure of Hill and Smith[9] with the exception that the final purification step was performed by electrophoresis on a starch block[10] instead of on a starch column. The purified product, which was dialyzed against *p*H 8.55 Veronal buffer (0.075 *M* and containing 0.005 *M* $MgCl_2$), had a specific activity of 60 C_1

(1) Aided by a grant from the United States Public Health Service and USPH Training Grant CRTY-5028.

(2) D. T. Gish, L. K. Ramachandran and W. M. Stanley, *Arch. Biochem. Biophys.*, in press.

(3) Abbreviations for the amino acid residues are those suggested by E. Brand and J. T. Edsall, *Ann. Rev. Biochem.*, **16**, 223 (1947).

(4) L. K. Ramachandran, *Virology*, **5**, 244 (1958).

(5) F. L. Black and C. A. Knight, *J. Biol. Chem.*, **202**, 51 (1953).

(6) H. Fraenkel-Conrat and B. Singer, *Biochim. et Biophys. Acta*, **24**, 540 (1957).

(7) H. Fraenkel-Conrat and B. Singer, THIS JOURNAL, **76**, 180 (1954).

(8) C. A. Knight, "The Nature of Viruses," Ciba Foundation Symposium, Ed. G. E. Wolstenholme and E. A. P. Millar, J. and A. Churchill, Inc., London, 1956, p. 69.

(9) R. L. Hill and E. L. Smith, *J. Biol. Chem.*, **228**, 577 (1957).

(10) H. G. Kunkel and R. J. Slater, *Proc. Soc. Exptl. Biol. Med.*, **80**, 42 (1952).

units[11] per mg. N. This enzyme preparation was found to have no proteolytic activity not attributable to the known specificity of leucine aminopeptidase.[9]

The following enzyme preparations were purchased from Worthington Biochemicals of Freehold, New Jersey: carboxypeptidase (lot no. 0570); chymotrypsin, crystalline, salt free (lot no. CD445-8); pepsin, twice recrystallized from ethanol (lot no. PM610).

N-Terminal and Amino Acid Analyses.—Fluorodinitrobenzene (FDNB)[12] was used for the determination of N-terminal amino acids and for many of the amino acid analyses. The general procedure described by Fraenkel-Conrat, Harris and Levy[13] was followed. Unless otherwise noted, the peptides or amino acids were converted to their dinitrophenyl (DNP) derivatives in the presence of a several fold excess of bicarbonate buffer and FDNB (added as a 2.5% solution in ethanol). The DNP-peptide was usually hydrolyzed in constant boiling HCl at 108° for 16 hr. The two-dimensional paper chromatographic system described by Levy[14] was used for the separation of the ether-soluble DNP-amino acids and the *t*-amyl alcohol-*p*H 6 phthalate system of Blackburn and Lowther[15] was used for paper chromatography of the aqueous soluble DNP-amino acids (ε-DNP-lysine and DNP-arginine). A useful variation of Levy's procedure was sometimes employed for the unequivocal identification and complete separation of DNP-glutamic acid and DNP-aspartic acid. After chromatography in the first dimension, a strip was cut from the dried chromatogram at right angles to the first dimension containing the DNP-glutamic acid and /or DNP-aspartic acid spot. The DNP-spot was covered and the strip was sprayed with buffer, dried and developed with the *t*-amyl alcohol-*p*H 5 phthalate system of Blackburn and Lowther.[15] DNP-glutamic acid and DNP-aspartic acid are easily separated and identified with this system in contrast to the results obtained with the *p*H 6 phosphate buffer of Levy's system where overlapping of these two DNP-derivatives usually is observed.

The DNP method for analysis of amino acids in a peptide hydrolysate usually served quite well. In those cases, however, where the peptide contained two or more residues of the same amino acid, chromatography of a hydrolysate on an ion-exchange column of Dowex 50-X4 was used.[16]

Purification and Elucidation of the Structures of Peptides. K-0.065.—The peptides described in this paper had been partially purified by countercurrent distribution as reported in the first paper of this series[2] and that paper should be consulted for a full description of each fraction. The material which had been recovered from tubes 160-210 after 3130 transfers (see Fig. 1) had been designated K-0.065 and weighed 48.3 mg. Paper chromatography of a sample of this material on Whatman No. 1 paper in the solvent system *n*-butanol-acetic acid-water-pyridine (30:6:24:20)[17] (this solvent system will henceforth be referred to as B-A-W-P) revealed that it contained two major ninhydrin-positive components with R_f values of 0.16 and 0.26. Both gave a positive Sakaguchi test for arginine[18] and the one of R_f 0.26 gave a positive test for tyrosine.[18] A trace of at least one other component was noted.

It was found that these components could be separated by electrophoresis at *p*H 5.5 in a pyridine-acetic acid buffer (24.3 ml. of pyridine and 5.7 ml. of glacial acetic acid diluted to one liter). For this purpose zone electrophoresis[19] in the apparatus described by Paigen[19] was used. The supporting medium was Whatman standard grade cellulose powder which had been washed several times with water and finally with buffer. The cellulose block was 1 cm. deep, 3.5 cm. wide and 45 cm. long. The electrophoresis was carried out at 200 v. for 21 hr. in a refrigerated box maintained at about 5°. The block was cut into 0.5 cm. segments and each segment was eluted with 2 ml. of water and centrifuged. Aliquots (0.1 ml.) of the eluates were taken for analysis by the Folin reaction as modified by Lowry, *et al.*,[20] which is given, to variant degrees, by all peptides, and by the Sakaguchi reaction as modified by Rosenberg, *et al.*,[21] which is given by arginine-containing peptides. When 10 mg. of K-0.065 was subjected to electrophoresis as described above, the results shown in Fig. 2 were obtained. The peptide designated K-0.065-7 was found to be a tyrosine-containing heptapeptide and the peptide designated K-0.065-2 was found to be a dipeptide. These differences would account for the great disparity between the color yield given by these two peptides in the Folin-Lowry reaction. The segments were pooled as shown and extracted three more times with water, the eluate being separated each time by filtration. After several extractions it was found that the cellulose tended to become finely dispersed, some passing through the filter. The addition of a drop or two of acetic acid to the eluate, however, caused the cellulose to coagulate. The eluates of each component were pooled and lyophilized. The recovery of K-0.065-7 was 4.6 mg. and that of K-0.065-2 was 2.6 mg.

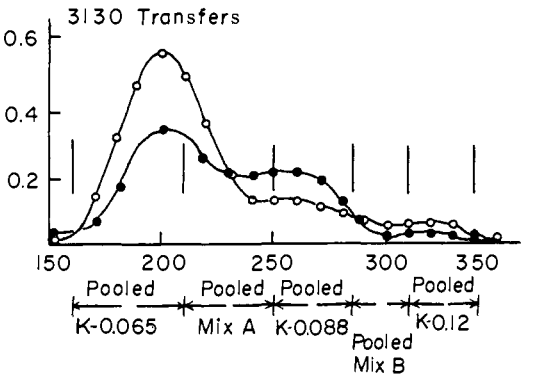

Fig. 1.—Countercurrent distribution curves of K-0.065 and K-0.086. The open circles are the o.d. values for the Folin-Lowry color and the closed circles the o.d. values for the modified Sakaguchi color.

The peptide K-0.065-2 was analyzed by the DNP method and the molar ratios found were aspartic acid 0.85, arginine 1.00, with aspartic acid being N-terminal. The peptide (1.23 mg.) was treated with about 1.4 C_1 units of leucine aminopeptidase in 0.2 ml. of 0.025 *M* *p*H 8.55 Veronal buffer containing 0.0017 *M* $MgCl_2$ at 37°. Aliquots were removed at various time intervals and analyzed by the DNP method. DNP-asparagine and DNP-arginine were found in equimolar amounts, the hydrolysis being complete in about 12 hr. The peptide K-0.065-2 was therefore aspNH_2.arg.

The amino acid composition of K-0.065-7 was determined by the DNP method. The molar ratios found were arg 1.00, asp 1.14, gly 0.98, ser 1.00, thr 1.00, tyr 0.86, plus the N-terminal amino acid, glycine, the DNP-derivative of which was recovered in 20% yield after acid hydrolysis. This peptide (0.51 mg.) was treated with about 0.23 C_1 unit of leucine aminopeptidase in 0.23 ml. of *p*H 8.55 Veronal buffer (0.015 *M*, containing 0.001 *M* $MgCl_2$) at a temperature of 37°. Aliquots were removed at various time intervals and analyzed by the DNP method. The results obtained are shown in Fig. 3. The N-terminal glycine and the next amino acid, threonine, were split off at almost the same rate and the release of these two amino acids was complete after 12 hr. Serine was apparently split off after threonine, the amount of serine always exceeding the amount of tyrosine, asparagine and the second mole of glycine. No quantitative data were obtained on the release of arginine because DNP-arginine failed to separate from other DNP-material, apparently a DNP-peptide. These results indicated the partial sequence of gly·thr·ser (gly, tyr, aspNH_2, arg) for this peptide.

The peptide K-0.065-7 was hydrolyzed with chymotrypsin. For this purpose 1 mg. of the peptide was dissolved in

(11) D. H. Spackman, E. L. Smith and D. M. Brown, *J. Biol. Chem.*, **212**, 255 (1955).

(12) F. Sanger, *Biochem. J.*, **39**, 507 (1945).

(13) H. Fraenkel-Conrat, J. I. Harris and A. L. Levy, in "Methods of Biochemical Analysis," Vol. II, Ed. D. Glick, Interscience Publishers, Inc., New York, N. Y., 1954, p. 359.

(14) A. L. Levy, *Nature*, **174**, 126 (1954).

(15) S. Blackburn and A. G. Lowther, *Biochem. J.*, **48**, 126 (1951).

(16) S. Moore and W. H. Stein, *J. Biol. Chem.*, **211**, 893 (1954).

(17) S. G. Waley and J. Watson, *Biochem. J.*, **55**, 328 (1953).

(18) R. Acher and C. Crocker, *Biochim. et Biophys. Acta*, **9**, 704 (1952).

(19) K. Paigen, *Anal. Chem.*, **28**, 284 (1956).

(20) O. H. Lowry, N. J. Rosebrough, A. L. Farr and R. J. Randall, *J. Biol. Chem.*, **193**, 265 (1951).

(21) H. Rosenberg, A. H. Ennor and J. F. Morrison, *Biochem. J.*, **63**, 153 (1956).

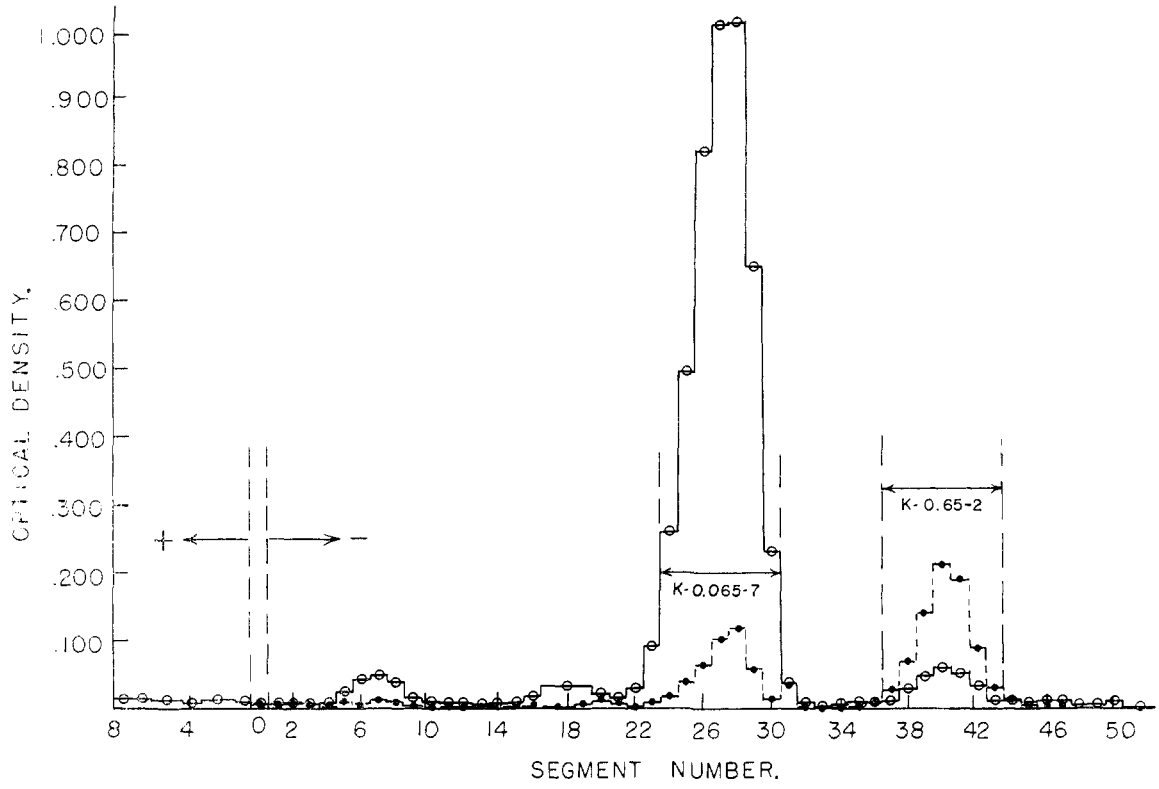

Fig. 2.—Electrophoresis of K-0.065. The open and closed circles have the same representation as in Fig. 1.

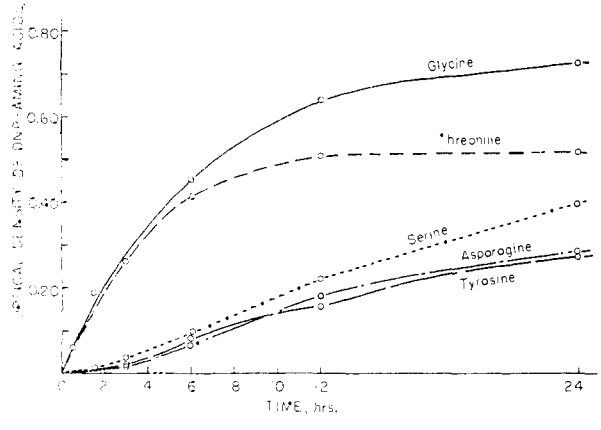

Fig. 3.—Hydrolysis of K-0.065-7 with leucine aminopeptidase.

0.5 ml. of water, the *p*H was adjusted to 7.8 and 0.025 mg. of chymotrypsin was added. A test reaction had shown that at 37° under these conditions the peptide was split rapidly into two peptides, the reaction being practically complete in 1 hr. After 2 hr. the digest was frozen and lyophilized. The residue was dissolved in water and chromatographed on Whatman No. 1 paper in the B–A–W–P solvent system. A guide strip was cut from one edge of the paper and stained successively with ninhydrin, the tyrosine reagent[18] and the arginine reagent.[18] Two zones were found, one with an R_f of 0.18 (designated K–0.065–7–C–1) and giving a positive test for arginine and the second with an R_f of 0.46 (designated K–0.065–7–C–2) and giving a positive test for tyrosine. The zones were eluted with water and the solutions were evaporated to dryness over concentrated H_2SO_4.

The peptide K–0.065–7–C–1 was analyzed by the DNP method. It contained equimolar amounts of aspartic acid and arginine (molar ratios found: arg 1.00, asp 1.08), with the former being N-terminal. Half of the peptide isolated from the chymotryptic digest was dissolved in water and the *p*H was adjusted to 8.5 with dilute NaOH. To this solution (0.33 ml.) was added 0.05 ml. of *p*H 8.55 Veronal buffer (0.075 *M*, containing 0.005 *M* $MgCl_2$) and about 0.17 C_1 unit of leucine aminopeptidase, and the hydrolysis was allowed to proceed at 37°. At various time intervals aliquots were dinitrophenylated and chromatographed in the *p*H 6 Blackburn and Lowther system. It was found that equimolar amounts of asparagine and arginine were released. The peptide K–0.065–7–C–1 was therefore aspNH_2·arg.

About one fourth of the peptide K–0.065–7–C–2 isolated from the chymotryptic digest of K–0.065–7 was treated with carboxypeptidase. The peptide was dissolved in 0.12 ml. of water and 0.25 ml. of *p*H 7.5 Veronal buffer (0.02 *M*, containing 0.1 *M* $NaCl_2$) and 0.03 mg. of carboxypeptidase was added. The mixture was kept at 37° and after various time intervals aliquots were dinitrophenylated. Both the ether-soluble and the water-soluble DNP-derivatives were chromatographed in the *p*H 6 Blackburn and Lowther system. It was found that tyrosine, and only tyrosine, was released from this peptide by carboxypeptidase. The reaction was rapid, no additional tyrosine being released after 15 min. The DNP-derivative of the residual peptide was water-soluble and insoluble in ether and was recovered after chromatography. This was accomplished by extraction of the paper spots with 1% $NaHCO_3$, evaporation of the extracts to dryness in an air stream at 45°, extraction of the residue with acetone containing 1% concentrated HCl and evaporation of the acid–acetone extract. Analysis of this DNP-peptide by the DNP method showed that, in addition to the N-terminal glycine which was recovered in 30% yield after hydrolysis of the DNP-peptide in concentrated HCl for 4 hr. at 108°, the peptide contained one mole each of threonine, serine and glycine (molar ratios found: gly 1.00, ser 1.06, thr 0.98). On the basis of the results of these experiments and of the treatment of K–0.065–7 with leucine aminopeptidase, the peptide K–0.065–7–C–2 could be assigned the sequence gly·thr·ser·gly·tyr. The peptide K–0.065–7 therefore was assigned the structure gly·thr·ser·gly·tyr·aspNH_2·arg.

K-0.086.—The material designated K-0.086 was isolated from tubes 250-285 after 3130 transfers (see Fig. 1) and weighed 62.4 mg. Chromatography on Whatman No. 1 paper in the B–A–W–P system revealed that the material was a mixture of two compounds along with a trace of a third component. The major component (based on nin-

hydrin color) had an R_f of 0 37 and contained arginine and the minor component had an R_f of 0.24.

A portion (20 mg.) of the material was dissolved in 0.2 ml. of water and streaked on a line 7 in. long and 3.5 in. from the top of a sheet of Whatman No. 3 paper, 9 × 22.5 in. The chromatogram was developed with the solvent system described above until the solvent had reached the bottom of the paper. The chromatogram was air-dried and developed a second time in the same direction. A guide strip was cut from one edge of the chromatogram and stained with ninhydrin. The two major zones thus located were cut from the paper and eluted with water. These eluates were lyophilized to yield 3.4 mg. of the minor component of *R*. 0.24 (henceforth designated K–0.086-Mi) and 6.8 mg. of the major component of R_f 0.37 (designated K–0.086-Ma).

Analysis of K-0.086-Ma by the DNP method revealed that it contained N-terminal threonine and one mole each of glutamic acid and arginine and two moles of valine (molar ratios found: arg 1.0, glu 1.15, thr 1.08, val 2.13). Hydrazinolysis[22] (performed by Dr. Kozo Narita) of a sample of this material confirmed that arginine was the C-terminal amino acid.

The peptide (0.80 mg.) was incubated at 37° with 1.33 C_1 units of leucine aminopeptidase in 0.4 ml. of *p*H 8.55 Veronal buffer (0.025 *M*, containing 0.0017 *M* $MgCl_2$). Aliquots were removed from time to time and analyzed for amino acids by the DNP method. The results, which are shown in Fig. 4, indicated that following the splitting of the

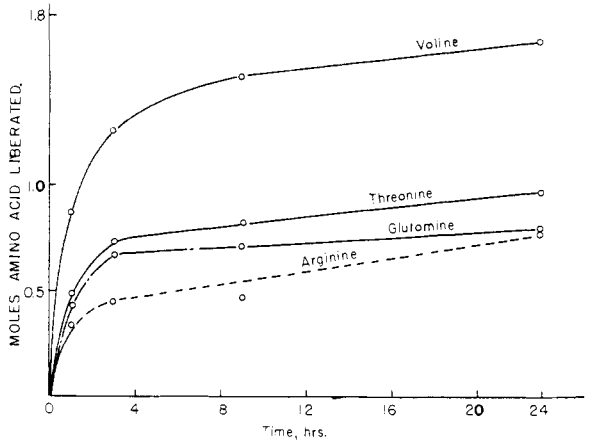

Fig. 4.—Hydrolysis of K-0.086-Ma with leucine aminopeptidase.

N-terminal threonine bond a valine residue was released by the enzyme. A mole of glutamine probably was released next, followed by the second valine residue and the arginine. The sequence known with certainty from the above experiments was therefore thr· val (gluNH$_2$, val) arg.

The peptide then was degraded by the Edman procedure.[23] The sample (0.40 mg.) was dissolved in 2 ml. of 50% pyridine, the *p*H was adjusted to 8.0 with dilute NaOH and 0.1 ml. of phenyl isothiocyanate was added. The mixture was stirred at room temperature for 2 hr. and then extracted 6 times with 10 ml. of benzene. The aqueous solution was evaporated to dryness with a stream of nitrogen. The residue was dissolved in 0.5 ml. of acetic acid and HCl gas was passed in at 80° for 15 min.[24] The mixture was evaporated to dryness in a nitrogen stream and the residue was dissolved in 0.5 ml. of water and extracted twice with 2-ml. portions of ethyl acetate. The extracts were combined and evaporated to a small volume and aliquots were chromatographed on starch-impregnated paper using Sjöquist's Solvent C,[25] and the phenylthiohydantoins (PTH) of the amino acids were revealed using Feigl's reagent. The PTH's of the amino acids were identified through the first four steps as follows: 1st, threonine; 2nd, valine, with a trace of threonine; 3rd, glutamine and a small amount of glutamic acid; 4th, valine. The residual material was chromatographed in the system *n*-butanol–formic acid–water (75:10:15) and arginine was the only major ninhydrin-positive spot found. The peptide K–0.086-Ma was therefore thr·val·gluNH$_2$·val·arg.

K-0.13.—The material designated K–0.13 had been recovered from tubes 410–500 after 4106 transfers of Fraction F (see Fig. 5) and weighed 68.4 mg. Paper chromatography

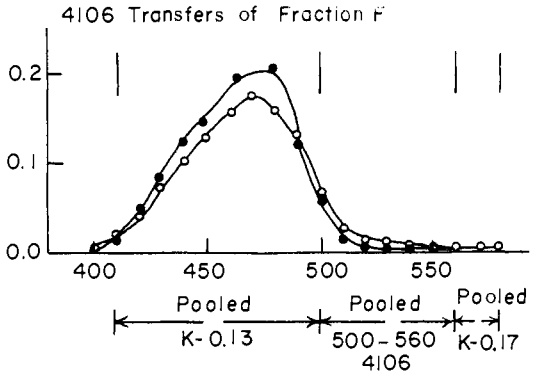

Fig. 5.—Distribution curve of K-0.13. The open and closed circles have the same representation as in Fig. 1.

of a sample of this material on Whatman No. 1 paper in the system B–A–W–P showed that it contained a component of R_f 0.36 and a small amount of material of lower R_f. When 7.5 mg. of K–0.13 was chromatographed on Whatman No. 3 paper in the B–A–W–P system and the main zone was eluted and the eluate was lyophilized, 6.0 mg. of purified material was obtained.

Chromatography of the amino acids from an acid hydrolysate of the peptide on Whatman No. 1 paper in the system 2-butanol–3% NH_4OH (3:1)[26] revealed the presence in the peptide of arginine, aspartic acid, alanine, isoleucine, valine and threonine. Analysis of the peptide by the DNP method yielded the following molar ratios: arg 2.00, asp 2.22, val 2.25, ala 2.12, thr 0.91 and ileu 0.87. Arginine was N-terminal. The peptide (1.2 mg.) was hydrolyzed for 48 hr. at 108° with 5.7 *N* HCl in a sealed tube and the hydrolysate was chromatographed on the ion-exchange resin Dowex 50-X4. The molar ratios found (based on isoleucine as 1.0) were: asp 1.89, thr 0.91, ala 1.93, val 1.84, ileu 1.00, NH_3 0.38 and arg 1.86. The recovery of only 0.38 mole of ammonia indicated that neither of the aspartic acid residues was present as asparagine. In a separate determination of amide ammonia,[27] 0.39 mole of ammonia per mole of peptide was found.

The peptide (0.8 μM) was dissolved in 0.3 ml. of *p*H 8.55 Veronal buffer (0.075 *M* containing 0.005 *M* $MgCl_2$), 1.4 C_1 units of leucine aminopeptidase was added and the mixture was incubated at 37°. Aliquots were taken from time to time and analyzed for amino acids by the DNP method. The results are shown in Fig. 6. On the basis of these results the partial sequence arg· val·val·asp. . . . . was postulated.

The peptide was digested with pepsin. For this purpose about 5 μM of the peptide was dissolved in water and the *p*H was adjusted to 1.8. To this solution (total volume 2.2 ml.) was added 0.12 mg. of pepsin and the digestion was allowed to proceed for 10 hr. at 37°. The mixture then was kept frozen until chromatographed. The peptic hydrolysis products were separated by chromatography on Whatman No. 3 paper in the B–A–W–P system. Four ninhydrin-positive zones were separated, all of which gave a positive Sakaguchi test. Two of the zones, with R_f values of 0.45 and 0.51, were present in small amount. The two main zones had R_f values of 0.35 (designated K–0.13–P–1) and 0.57 (designated K–0.13–P–4). The zones were eluted with water and the eluates were evaporated to dryness *in vacuo* over concentrated H_2SO_4. Paper chromatography of these two fractions in the solvent systems *n*-butanol–acetic acid–water (3:1:1) and *n*-butanol–formic acid–water (75:10:15)

(22) S. Akabori, K. Ohno and K. Narita, *Bull. Chem. Soc. Japan*, **25**, 214 (1952).

(23) P. Edman, *Acta Chem. Scand.*, **4**, 277 (1950).

(24) R. G. Shepherd, S. D. Willson, K. S. Howard, P. H. Bell, D. S. Davies, S. B. Davis, E. A. Eigner and N. E. Shakespeare, THIS JOURNAL, **78**, 5067 (1956).

(25) J. Sjöquist, *Acta Chem. Scand.*, **7**, 447 (1953).

(26) J. F. Roland and A. M. Gross, *Anal. Chem.*, **26**, 502 (1954).

(27) H. Fraenkel-Conrat and H. S. Olcott, THIS JOURNAL, **70**, 2675 (1948).

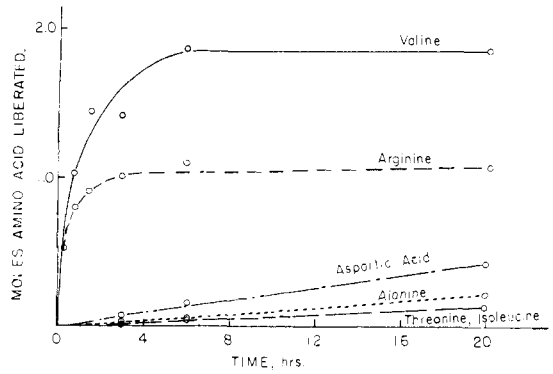

Fig. 6.—Hydrolysis of K-0.13 with leucine aminopeptidase.

showed that K–0.13–P–4 was an apparently homogeneous peptide but that K–0.13–P–1 could be resolved into a major and a minor component. The peptide K–0.13–P–1 was therefore chromatographed on Whatman No. 1 paper in the *n*-butanol–formic acid–water system and was separated into two zones, a major zone with R_f of 0.09 and a minor zone of R_f 0.04. The major zone was eluted with water and the eluate evaporated to dryness. This material was designated K–0.13–P–1a.

Analysis of K–0.13–P-1a by the DNP method showed that it contained one mole of N-terminal arginine and one mole each of valine, alanine and threonine and two moles of aspartic acid. Similar analysis of K–0.13–P–4 revealed that it was composed of one mole of N-terminal valine and one mole each of alanine, isoleucine and arginine.

The peptide K–0.13–P–1a was treated with carboxypeptidase (substrate to enzyme ratio of 4:1) at *p*H 7.8 and 37°. One mole of threonine, and only threonine, was liberated by this enzyme indicating that threonine is the C-terminal amino acid of K–0.13–P–1a. When the dethreoninated peptide was treated with leucine aminopeptidase, again the liberated arginine and valine had the ratio 1:2.

The peptide K–0.13–P–4 (2 μM) was dissolved in 0.32 ml. of the *p*H 8.55 Veronal buffer, and 0.7 C_1 unit of leucine aminopeptidase was added and the mixture was incubated at 37°. Aliquots were taken from time to time and analyzed by the DNP method. The results are shown in Fig. 7 and clearly indicated the sequence val·ala·ileu·arg for this peptide.

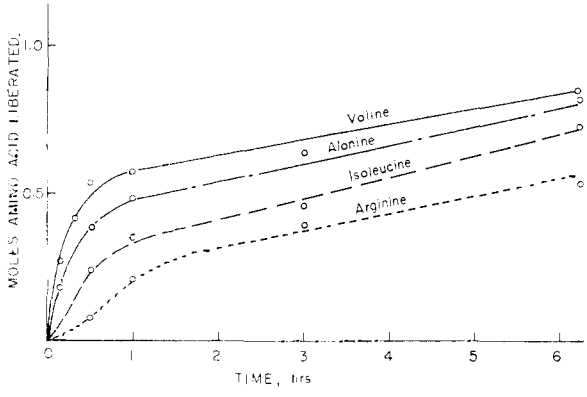

Fig. 7.—Hydrolysis of K-0.13-P-4 with leucine aminopeptidase.

Since the peptide K–0.13–P–4 was the C-terminal peptide derived from K–0.13 and contained one of the two valine residues found in K–0.13, it was obvious that the N-terminal sequence of K–0.13 could not be arg·val·val. The results of the treatment of K–0.13 and of the dethreoninated peptide derived from K–0.13–P–1a with leucine aminopeptidase which indicated the release of one arginine residue followed by the release of two valine residues must have been erroneous due either to release of valine from some other source besides the peptide or due to a poor recovery of arginine from this reaction mixture by the DNP method. The N-terminal sequence of K–0.13 was therefore investigated by the Edman degradation procedure.

The peptide K-0.13 (0.8 μM) was treated with 0.1 ml. of phenyl isothiocyanate in 2 ml. of 50% pyridine at *p*H 8 and 37° for 3 hr. The procedure was the same as that described above for the degradation of K-0.086-Ma except that the cyclization step was carried out in 2 *N* HCl at 37°.[28] The yield of the PTH of the amino acids was about 0.5 μM through the first five steps. Approximately one tenth of the PTH amino acid from each step was used for paper chromatography using Solvents B and C of Sjöquist. For the identification of the PTH of arginine released at step 1, an aliquot of the cyclization reaction mixture was used for chromatography since the PTH of arginine cannot be extracted into organic solvents. Unambiguous identification of the released PTH amino acids through the first five steps were as follows: 1, arg; 2, val; 3, asp; 4, asp; and step 5, ala. The identification of the PTH amino acids was confirmed by hydrolyzing a portion of each in 6 *N* HCl at 150° for 24 hr. and chromatographing the free amino acids thus obtained on Whatman No. 1 paper in the 2-butanol–3% NH_4OH (3:1) system.

From the results of the above experiments the structure of the decapeptide K-0.13 may be postulated as arg·val·asp·asp·ala·thr·val·ala·ileu·arg.

K-0.22.—This material had been recovered from tubes 410 to 480 after 2624 transfers of Fraction C (see Fig. 8). The fraction adjacent to this fraction had been designated Mix 2624 and was recovered from tubes 481 to 590. These two fractions would contain components common to both, due to overlapping of distribution curves. Chromatography on Whatman No. 1 paper in the B–A–W–P system of K-0.22 showed that it contained at least six ninhydrin-positive components, one of which (R_f 0.47) appeared to be the major component. Chromatography of the Mix 2624 material in the same system showed that it contained seven ninhydrin positive components, including a small amount of the material with R_f value of 0.47.

It was found that a considerable resolution of the components of K-0.22 could be achieved by electrophoresis at *p*H 7.0. A cellulose block 1 × 3.7 × 45 cm. provided the supporting medium and a *p*H 7 collidine–acetic acid buffer (9 ml. of collidine and 45 ml. of *N* acetic acid per liter) was used. When 10.0 mg. of K-0.22 was subjected to electrophoresis at an applied voltage of 250 v. for 16.5 hr. at 5°, the results shown in Fig. 9 were obtained. The main component, designated K-0.22-D, weighed 5.0 mg. When the fraction which had been designated Mix 2624 was similarly subjected to electrophoresis, the recovery of this same component amounted to only about 8% of the crude material.

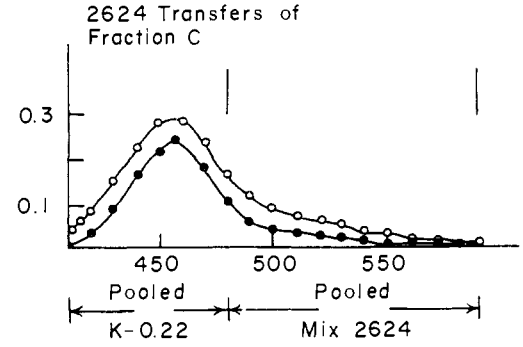

Fig. 8.—Distribution curve of K-0.22. The open and closed circles have the same representation as in Fig. 1.

Chromatography of K-0.22-D on Whatman No. 1 paper in the B–A–W–P system revealed that the peptide was still contaminated with a small amount of other ninhydrin-positive material. The peptide (3.5 mg.) was therefore further purified by chromatography on Whatman No. 1 paper in the above system. The aqueous eluate of the main band was concentrated to a small volume over concentrated H_2SO_4 and aliquots of this solution were used for structural studies.

When the purified peptide was analyzed by the DNP method, it was found to be composed of valine, tyrosine and

(28) H. Fraenkel-Conrat and J. I. Harris, THIS JOURNAL, **76**, 6058 (1954).

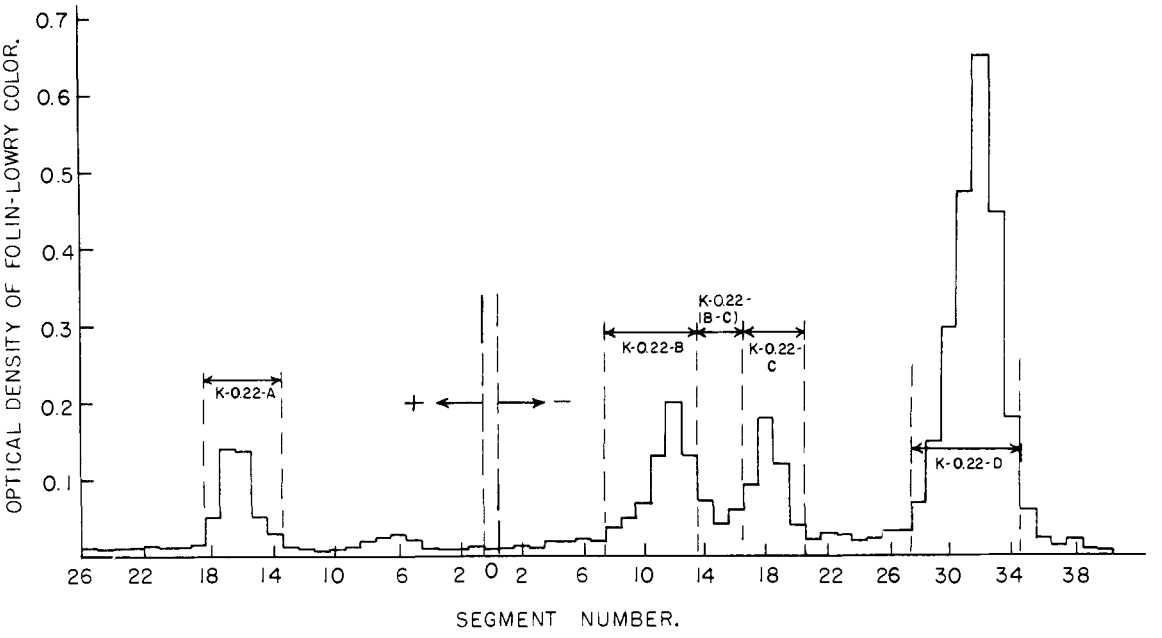

Fig. 9.—Electrophoresis of K-0.22.

arginine, with valine being N-terminal. Based on arginine as 1.0, the recovery of valine was 0.61 mole and that of tyrosine was only 0.35 mole. In addition an unknown DNP-compound of 0.21 mole was found. It appeared possible that the valyltyrosine bond was resistant to acid hydrolysis and that tyrosine in this peptide was subject to considerable destruction during hydrolysis. The resistance of this bond to acid hydrolysis could account for the fact that only about one-half mole of N-terminal valine had been found in the tryptic digest of TMV protein.[2]

The peptide (about 10% of the material isolated above) was treated with leucine aminopeptidase (0.18 C_1 unit) and the hydrolysate was analyzed by the DNP method. It was found that the peptide was split rapidly, valine, tyrosine and arginine being released in equimolar amounts (val 1.00, tyr 0.91, arg 1.09) within about 1 hr. Based upon these experiments and the specificity of trypsin, the peptide K–0.22-D may be assigned the sequence val·tyr·arg.

K-0.33.—This material was the fraction recovered from tubes 485–600 after 2155 transfers of Fraction E (see Fig. 10). The major fraction from tubes 485–560, designated

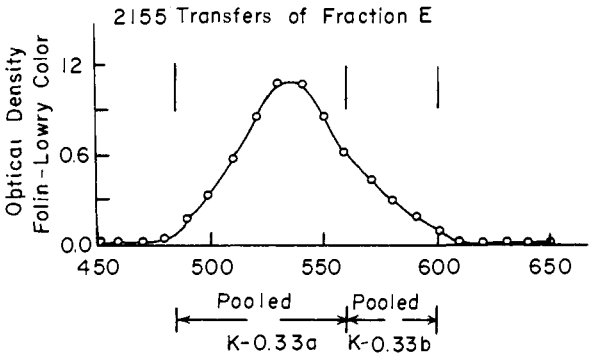

Fig. 10.—Distribution curve of K-0.33.

K-0.33a, amounted to 35 mg., while the minor fraction from tubes 561–600, designated K-0.33b, amounted to 6.5 mg. Chromatography of these two fractions on Whatman No. 1 paper in the B–A–W–P system revealed a major component of R_f 0.48 and a second component of R_f 0.24 amounting to only a few per cent., of the total based on ninhydrin color. Traces of other ninhydrin-positive substances also were present.

About 10 mg. of K-0.33a was subjected to zone electrophoresis on a 1 × 3.5 × 30 cm. cellulose block in a *p*H 7 collidine–acetic acid buffer (9.0 ml. of collidine and 45 ml. of *N* acetic acid per liter). The electrophoresis was conducted with an applied voltage of 300 v. for 16 hr. at 5°. Each 0.5 cm. segment was eluted with 2 ml. of water and the Folin–Lowry color was developed on 0.05-ml. aliquots. The results obtained are shown in Fig. 11. The segments were pooled in

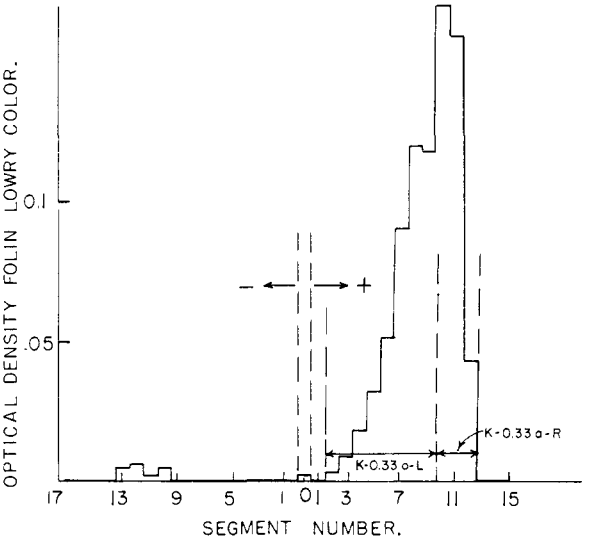

Fig. 11.—Electrophoresis of K-0.33a.

two fractions as shown, the cellulose was extracted twice more with water, the eluates were pooled, filtered and lyophilized. The fraction recovered from segments +2 to +9 amounted to 1.7 mg. (designated K-0.33a-L) and that recovered from segments +10 to +12 amounted to 2.6 mg. (designated K-0.33a-R). Chromatography of these two fractions on Whatman No. 1 paper in the B–A–W–P system revealed that K–0.33a–R contained only one ninhydrin-positive substance of R_f 0.45 and that K-0.33a-L contained a major component of R_f 0.45 and a small amount of material of lower R_f. The sequence studies described below were therefore performed on purified material.

A sample of the peptide was hydrolyzed in 6 *N* HCl and analyzed by the DNP method. This showed that the peptide contained one mole each of aspartic acid, serine, proline and lysine and two moles of phenylalanine (molar ratios found: phe 2.1, asp 0.88, ser 0.70, pro 1.00, lys 0.87). Phenylalanine was found to be N-terminal by this method.

Neither leucine aminopeptidase nor chymotrypsin showed any hydrolytic action on this peptide, the peptide bond between the N-terminal phenylalanine and the penultimate amino acid being resistant to the action of both enzymes and the internal phenylalanine also being resistant to chymotrypsin.

About 0.72 mg. of the peptide was degraded by the Edman procedure through five steps, the cyclization being effected in acetic acid–HCl. The identification of the PTH amino acid obtained at each step was made by hydrolyzing a portion of the PTH amino acid in constant boiling HCl in a sealed tube at 150° for 16 hr. and chromatographing the amino acid so obtained on Whatman No. 1 paper in the system methanol–water–pyridine (80:20:4) which separates all of the amino acids found in this peptide. The amino acids identified at each step were as follows: 1, phe; 2, pro with a trace of phe; 3, asp with a trace of phe; 4, phe with a trace of ser; 5, ser with a trace of phe and lys. Lysine was identified as the residue left after the fifth step. These results indicated the sequence phe·pro·(asp or aspNH_2)·phe·ser·lys. The yield of the PTH amino acid at each step was low, however, or about one fourth of that expected.

Because of the low yield obtained above, the Edman degradation was repeated using different conditions for the cyclization step. About 1.43 mg. (1.9 μM) was degraded. The cyclization in the first step was brought about in 0.5 N HCl at 37° and through the next five steps in 2 N HCl at room temperature. The yield of the PTH in steps 1 through 6 (in μM) was 1.75, 1.5, 1.2, 1.4, 1.4 and 0.7, respectively, or from about 40–90%. The low yield obtained in the first method apparently was due to the cyclization procedure employed. The identification of the amino acid at each step was the same as that first obtained.

The average of two determinations of amide ammonia content of this peptide was 0.16 mole per mole of peptide which indicated that aspartic acid, not asparagine, was present in the peptide. This fact also was indicated by the fact that the peptide is negatively charged at pH 7, moving toward the anode when subjected to electrophoresis at that pH. Were asparagine present in the molecule, the peptide would be a strongly basic peptide and would be positively charged at pH 7. The amino acid sequence of K-0.33 may therefore be written phe·pro·asp·phe·ser·lys.

The results are summarized in Table I and Table II. The calculated recoveries reported in Table I were obtained by the formula

$$\frac{\text{mol. wt. of peptide}}{18{,}000 - \text{mol. wt. insoluble peptide}} \times 740 \text{ mg.}$$

The total weight of peptide material which had been distributed was approximately 780 mg., about 40 mg. of which had been consumed for preparations of the various distribution curves.[3] The molecular weight of the protein sub-unit is about 18,000 and the "molecular weight" of the insoluble peptide material found in the digest is about 6,000.[2] The various fractions removed from the countercurrent distribution machine usually contained only part of each component of a particular distribution coefficient. The recovery of each purified peptide must therefore be multiplied by an appropriate factor to obtain the weight of the peptide that would have been recovered had all of this component been included in the fraction taken from the machine. Thus, it was estimated that the fraction designated K-0.065 (see Fig. 1) contained about 70% of the two components (K-0.065–2 and K-0.065-7) of this distribution coefficient while about half of the material with a distribution coefficient of 0.33 was estimated to be present in the fractions designated K-0.33a and K-0.33b (see Figs. 3, 16 and 20 of ref. 2).

TABLE II

PEPTIDES ISOLATED FROM A TRYPTIC DIGEST OF TMV PROTEIN

Component	Mol. wt.	Amino acid sequence
K-0.065-2	288.3	aspNH$_2$·arg
K-0.065-7	753.7	gly·thr·ser·gly·tyr·aspNH$_2$·arg
K-0.086-Ma	601.7	thr·val·gluNH$_2$·val·arg
K-0.13	1115.2	arg·val·asp·asp·ala·thr·val·ala·ileu·arg
K-0.22-D	436.5	val·tyr·arg
K-0.33	739.8	phe·pro·asp·phe·ser·lys

It can be seen from an examination of Table I that the peptides reported in this paper were recovered in an amount that would justify the assumption that all were major products of the action of trypsin on the protein sub-unit of TMV and that none were derived from transpeptidation or from impurities.

Acknowledgments.—The authors are indebted to Mrs. Gladys Perez-Mendez and Mr. Yusei Kato for valuable technical assistance.

BERKELEY, CALIF.

[CONTRIBUTION FROM THE RESEARCH DIVISION OF THE CLEVELAND CLINIC FOUNDATION AND THE FRANK E. BUNTS EDUCATIONAL INSTITUTE]

Syntheses of an Optically Pure Tetrapeptide Contained in Angiotensin[1,2]

BY HANS SCHWARZ AND F. MERLIN BUMPUS

RECEIVED AUGUST 4, 1958

The tetrapeptide derivative cbzo-nitro-L-arginyl-L-valyl-L-tyrosyl-L-isoleucine methyl ester was synthesized by the three routes possible, and the products were compared for yield, optical purity and by-products. A similar comparison was made with the two intermediate tripeptides obtained each by two different routes. Mixed anhydride and carbodiimide methods were used exclusively. A mixture of diastereoisomers was obtained with either method whenever an acyl di- or higher peptide was used for condensation. Both diastereoisomers, cbzo-L-valyl-L-tyrosyl-L-isoleucine methyl ester and cbzo-L-valyl-D-tyrosyl-L-isoleucine methyl ester, were isolated in crystalline form from the reaction of cbzo-L-valyl-L-tyrosine and L-isoleucine methyl ester. The best route to synthesize a polypeptide when using either mixed anhydride or carbodiimide method seems to be the one starting with the C-terminal amino acid and adding one amino acid at a time. The products obtained in good yield are optically pure.

In our recent synthesis of the isoleucine angiotensin octapeptide[3] we were handicapped by the fact that all derivatives of peptides containing four or more amino acids were amorphous, gelatinous products. Intermediates of this nature are extremely hard to purify and, as was pointed out recently,[4,5] may contain large amounts of diastereoisomers or other by-products. They, therefore, endanger the outcome of any synthesis in which optical purity is demanded, especially of those in which the end product cannot be compared by biological activity to a naturally occurring poly-

(1) This work was supported in part by the National Heart Institute, U. S. Public Health Service, Grant No. H-96 (C7).

(2) Angiotensin is the term suggested to replace the dual name angiotonin-hypertensin; E. Braun-Menendez and I. H. Page, *Science*, **127**, 242 (1958).

(3) H. Schwarz, F. M. Bumpus and I. H. Page, THIS JOURNAL, **79**, 5697 (1957).

(4) B. F. Erlanger, W. V. Curran and N. Kokowsky, *ibid.*, **80**, 1128 (1958).

(5) K. Hoffmann, M. E. Woolner, G. Spuhler and E. T. Schwartz, *ibid.*, **80**, 1486 (1958).