Partridge, S. M., Davis, H. F., and Adair, G. S. (1961), *Biochem. J. 79*, 15.
Rosenberg, L., Schubert, M., and Sandson, J. (1967), *J. Biol. Chem. 242*, 4691.
Schachman, H. K. (1957), *Methods Enzymol. 4*, 32.
Schachman, H. K. (1959), Ultracentrifugation in Biochemistry, New York, N. Y., Academic.
Schubert, M. (1958), *Federation Proc. 17*, 1099.
Spackman, D. H., Stein, W. H., and Moore, S. (1958), *Anal. Chem. 30*, 1190.
Stern, K. G., and Atlas, S. M. (1953), *J. Biol. Chem. 203*, 795.
Svedberg, T., and Pedersen, K. O. (1940), The Ultracentrifuge, London, Oxford University.
Svennerholm, L. (1958), *Acta Chem. Scand. 12*, 547.
Trevelyan, W. E., and Harrison, J. S. (1952), *Biochem. J. 50*, 298.
Warren, L. (1959), *J. Biol. Chem. 234*, 1971.
Woessner, J. F., Jr. (1961), *Arch. Biochem. Biophys. 93*, 440.

The Use of Peptide Synthesis to Establish the Amino Acid Sequence of Tobacco Mosaic Virus Protein Tryptic Peptide 2*

Janis Dillaha Young, Cherry Y. Leung, and Wilfred A. Rombauts†

ABSTRACT: Peptides corresponding to two previously proposed sequences of a tryptic peptide representing residues 42–46 of the tobacco mosaic virus protein (TMVP) having a sequence interchange at residues 44 and 45 were synthesized using the Merrifield solid-phase method. (This peptide was designated as TMVP tryptic peptide 2 according to the nomenclature proposed by Tsugita *et al.* (Tsugita, A., Gish, D. T., Young, J., Fraenkel-Conrat, H., Knight, C. A., and Stanley, W. M. (1960), *Proc. Natl. Acad. Sci. U. S. 46*, 1463).) Comparison of the properties of these synthetic peptides with the peptide isolated from the protein revealed that the sequence in the protein is: Thr-Val-Val-Gln-Arg. (1) On paper chromatography the synthetic Thr-Val-Val-Gln-Arg and TMVP tryptic peptide 2 cochromatographed and migrated slower than the synthetic Thr-Val-Gln-Val-Arg-(2) The two synthetic peptides gave markedly different paper chromatographic patterns after partial acid hydrolysis with the pattern of TMVP tryptic peptide 2 being identical with that of synthetic Thr-Val-Val-Gln-Arg. (3) Synthetic Thr-Val-Gln-Val-Arg was completely hydrolyzed by overnight acid hydrolysis, but both the synthetic Thr-Val-Val-Gln-Arg and the TMVP tryptic peptide 2 required a longer hydrolysis time for complete hydrolysis. (4) Though difficulties were encountered in performing N-terminal stepwise degradation on the peptides, these analyses confirmed that the sequence of TMVP tryptic peptide 2 is Thr-Val-Val-Gln-Arg.

Since peptide synthesis has been greatly simplified by the use of the Merrifield (1964) solid-phase method, synthesis may soon be a common tool for use in sequence analysis. Recently, Bornstein (1967) utilized synthetic peptides in ascertaining that a peptide of collagen was α-L-glutamyl-L-prolyl-glycine and not α-L-glutamyl-L-prolyl-glycine. We have used synthesis in ascertaining the sequence of TMVP tryptic peptide 2[1] containing residues 42–46. This peptide was reported by both the Berkeley group and the Tubingen group in 1960 as Thr-Val-Gln-Val-Arg (Anderer *et al.*, 1960; Tsugita *et al.*, 1960) but in 1965 the Tubingen laboratory reinvestigated this peptide and found that the sequence was Thr-Val-Val-Gln-Arg (Anderer *et al.*, 1965). This only remaining point of difference between the two laboratories regarding the TMVP amino acid sequence has been utilized in this paper to ascertain whether synthesis of the two proposed sequences and comparison of their properties to that of TMVP tryptic peptide 2 can be used as a method for determining the correct sequence in naturally occurring peptides.

Experimental Procedures

TMVP was obtained from TMV[2] by treatment with 67% acetic acid (Fraenkel-Conrat, 1957). Amino acid

* From the Laboratory of Medical Entomology, Kaiser Foundation Research Institute and Allergy Research Division, Kaiser Foundation Hospitals, San Francisco, California 94115, and the Virus Laboratory, University of California, Berkeley, California 94720. *Received December 20, 1967.* This investigation was supported in part by Public Health Service Research Grant No. AI 06040 from the National Institutes of Health and by National Science Foundation Grant No. GB 3107.

† Present address: Department of Physiological Chemistry, University of Louvain Medical School, Belgium.

[1] The tryptic peptide is numbered according to the nomenclature proposed by Tsugita *et al.* (1960).

[2] The tobacco mosaic virus was generously supplied both by Dr. C. A. Knight of the Virus Laboratory, University of California, Berkeley, and Dr. Stanley Mandeles of the Space Science Laboratory, University of California, Berkeley.

TABLE I: Amino Acid Ratios of Pentapeptides.

Peptide	Mole Ratio of Amino Acids and Amide Group				
	Arg	Thr	Glu	Val	Amide (NH_3)[d]
TMVP tryptic peptide 2[a]	1.09	0.86	1.06	2.00	
Synthetic Thr-Val-Gln-Val-Arg					
Peptide-resin[b]	1.00	1.06	0.92	2.04	
Column purified[c]	0.98	1.00	1.02	1.99	1.01
Synthetic Thr-Val-Val-Gln-Arg					
Peptide-resin[b]	1.00	1.00	0.96	1.52	
Column purified[a]	1.05	0.91	1.00	2.04	0.95

[a] Hydrolysis time, 72 hr. [b] Hydrolysis time, 15 hr, in 12 N HCl–dioxane (1:1) followed by a 15-hr hydrolysis in 6 N HCl. [c] Hydrolysis time, 15 hr. [d] Amide content was determined from the NH_3 released during hydrolysis of the peptide and is given as moles amide per mole of arginine (see text).

analyses and ammonia analyses were performed on a Spinco Model 120B amino acid analyzer (Spackman, 1963). The moles of amide group per mole of peptide (based on the arginine analysis) was determined by subtracting the NH_3 in the peptide prior to hydrolysis from the NH_3 found after overnight hydrolysis in 6 N HCl in a sealed tube at 100°. Duplicate analyses were determined for the NH_3 in the peptide both before and after hydrolysis. The same batch of ninhydrin was used for all of the ammonia analyses. Partial acid hydrolysis was performed using the conditions of Anderer *et al.* (1965).

Ion-exchange chromatography was performed using a 1 × 150 cm Dowex 1-X2 ion-exchange column equilibrated with a collidine–pyridine–acetic acid buffer (pH 8.8) and eluted with a pH 8.12 buffer as described by Funatsu (1964). The peptides eluted with the pH 8.12 buffer at approximately 50 ml. Acetic acid (50%), water, and pH 8.8 buffer were pumped successively through the column prior to the application of each sample.

Descending paper chromatography was performed on Whatman No. 1 using the solvent, 1-butanol–acetic acid–water–pyridine (15:3:12:10). Peptides isolated by paper chromatography were run on paper prewashed in the solvent overnight and were eluted with 0.1 N NH_4OH.

TMVP tryptic peptide 2[1] was obtained from a tryptic digest of TMVP by ion-exchange chromatography followed by paper chromatography of the first peak eluting from the ion-exchange column. The technique of double-solvent development as described by Ramachandran and Gish (1958) in isolating this peptide was used. The desired peptide was identified on side marker strips as the fastest moving spot staining yellow with ninhydrin (see the peptide map of Benjamini *et al.*, 1964).

Peptides (approximately 0.8 μmole) were digested for 2 hr with approximately 0.05 mg of leucine aminopeptidase (Worthington Biochemical Corp., Freehold, N. J., LAP-DFP-113) using the conditions for assay of the enzyme given in the 1967 Worthington Enzyme Manual. The digests were adjusted to pH 3 with acetic acid and dried *in vacuo*.

The Merrifield resin, chloromethylpolystyrene–2% divinylbenzene copolymer resin (lot no. H1502, 1.36 mequiv of Cl/g), was obtained from the Cyclo Chemical Corp., Los Angeles, Calif. The *t*-butyloxycarbonyl-L-amino acids (Boc), valine and nitroarginine (Schwarz BioResearch, Inc., Orangeburg, N. Y.), were of high purity as judged by thin-layer chromatography (Merrifield, 1964). Boc[3]-L-glutamine-*p*-nitrophenyl ester (Cyclo Chemical Corp.) showed several ninhydrin-positive spots on silica gel thin-layer chromatography plates developed in $CHCl_3$–MeOH–HOAc (85:10:5) and exposed to HCl vapor to remove the Boc group. This derivative was partially purified by chromatography on a 2 × 27 cm column of silicic acid (Bio-Sil A, Bio-Rad Laboratories, Richmond, Calif.) which was packed, equilibrated, and eluted with the above thin-layer chromatographic solvent. The glutamine derivative was checked for its identity by converting it into glutamine by a 30-min treatment with 1 N HCl in acetic acid followed by rotary evaporation. This material chromatographed identical with glutamine on the amino acid analyzer. Glutamine resolved adequately from threonine by using a pH 3.10 buffer instead of the usual pH 3.28 first buffer with glutamine eluting from the column 3 min after threonine and isoglutamine eluting 25 min after phenylalanine.

Solid-Phase Synthesis of Thr-Val-Gln-Val-Arg and of Thr-Val-Val-Gln-Arg. Boc-nitro-L-arginine-resin (0.35 mmole of Arg/g of resin) was prepared and analyzed as described by Stewart *et al.* (1966). The next four residues for each peptide were coupled to the nitroarginine essentially as described by Young *et al.* (1967) using 5-ml portions of reagents. In the cycle where glutamine was

[3] Abbreviations used that are not listed in *Biochemistry 5*, 1445 (1966), are: Boc, *t*-butyloxycarbonyl; DMF, dimethylformamide; PTH, phenylthiohydantoin.

coupled to the nitroarginine on the resin, the resin was washed with DMF (purified as described by Merrifield, 1964) instead of CH_2Cl_2 and Boc-L-glutamine-*p*-nitrophenyl ester dissolved in DMF was allowed to react with the peptide-resin for 4 hr. In each synthesis, 0.5 g of Boc-nitro-L-arginine-resin (0.17 mmole of arginine) was used. A 2.5 molar ratio of Boc-amino acid (0.43 mmole) to arginine, and except for glutamine, a 2.6 molar ratio of *N,N'*-dicyclohexylcarbodiimide (0.45 mmole) to arginine was used in each coupling reaction. Analyses of hydrolysates of the peptide-resins are given in Table I. The peptides were cleaved from the resin with anhydrous HBr in trifluoroacetic acid (Merrifield, 1964) and after evaporation was dissolved in 90% HOAc–10% H_2O and reduced by bubbling hydrogen overnight through the solution at atmospheric pressure using 5% palladium on barium sulfate as the catalyst. A portion of each of the peptides was subjected to column chromatography.

Edman Degradation. N-Terminal stepwise degradation of the peptides was performed using the paper strip method of Fraenkel-Conrat (1954) with the modifications of Schroeder *et al.* (1963). The time of exposure to aqueous acid vapor was increased from 7 to 11 hr in the hope that the valyl peptide bonds would be cleaved more quantitatively. The peptides (0.3–0.4 μmole) were degraded after applying equal amounts of the peptide on two paper strips. In each of the degradations a control was run consisting of two paper strips saturated with water. Thin-layer chromatography was performed using $^1/_{25}$ to $^1/_{10}$ of the total amount of PTH recovered. Identification of the phenylthiohydantoins was made on Eastman chromogram sheets type K 301R silica gel with fluorescent indicator (Doolittle, 1964; Edman and Begg, 1967) using solvents D or E of Sjöquist (1953) and solvents V or V and IV successively of the system of Jeppsson and Sjöquist (1967). For quantitation the dried PTH's were dissolved in 2 ml of absolute ethanol and their absorption spectra measured against an ethanol blank using a Cary 14 recording spectrophotometer (Sjöquist, 1959). Phenylthiohydantoins were converted into amino acids for the determination of valine and glutamic acid by oxygen-free alkaline hydrolysis (Africa and Carpenter, 1966). Residual peptides were eluted from the paper strips with 0.1 N HCl. Opening the pyrrolidonecarboxylyl rings was attempted using the method of Stark and Smyth (1963) which involves 3 N HCl hydrolysis at 100° in sealed tubes.

Results and Discussion

Amino acid analyses of the purified synthetic peptides and TMVP tryptic peptide 2 are given in Table I. Amide determinations were performed only on the synthetic peptides. The synthetic peptides contained the glutamine residue in α-peptide linkage since all the glutamic acid was in the amide form and leucine aminopeptidase digestion of both of the peptides released the expected molar ratios of threonine, glutamine, and valine as determined on the long column of the amino acid analyzer. No isoglutamine was released which would have been expected were any of the glutamic acid amide in α-peptide linkage. The native peptide was not digested with leucine aminopeptidase in this study but both Anderer *et al.* (1965) and Ramachandran and Gish (1958) obtained quantitative release of glutamine by leucine aminopeptidase digestion of TMVP peptide 2.[1] Based on the amino acid analyses of the peptide-resins, a 38% yield of Thr-Val-Gln-Val-Arg (31 μmoles) and a 15% yield of Thr-Val-Val-Gln-Arg (13 μmoles) was recovered after column purification. These quantities of peptides are quite easy to obtain by synthesis and provide generous amounts of material for testing various techniques for distinguishing the sequences without using the more difficult to obtain native peptide. Approximately 2.5 μmoles of the TMVP tryptic peptide 2[1] was obtained from 5 μmoles of TMVP. The paper chromatographic behavior of the two synthetic peptides was very similar and only by a 48-hr solvent development was the migration of the Thr-Val-Gln-Val-Arg clearly faster than that of the Thr-Val-Val-Gln-Arg and the native peptide as diagramed in Figure 1A. The cochromatography of the native peptide with each of the synthetic peptides established the proper sequence as Thr-Val-Val-Gln-Arg.

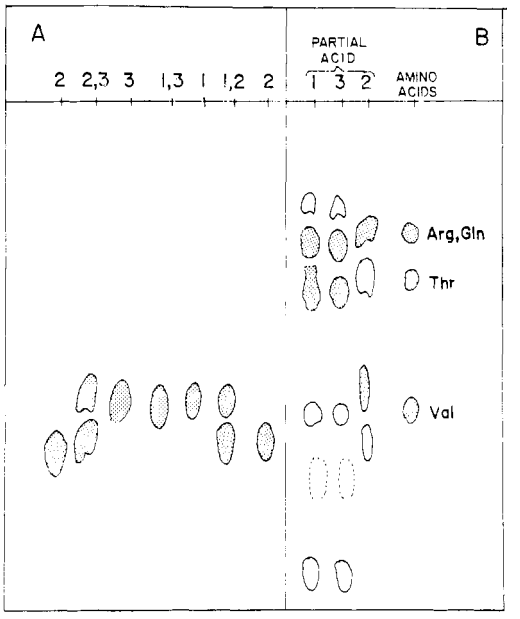

FIGURE 1: Paper chromatograms of 1, TMVP tryptic peptide 2; 2, synthetic Thr-Val-Gln-Val-Arg; and 3, synthetic Thr-Val-Val-Gln-Arg using 1-butanol–acetic acid–water–pyridine (15:3:12:10). (A) Untreated peptides; 48-hr solvent development; (B) parial acid hydrolysates of the peptides; 15-hr solvent development.

Partial acid hydrolysis of the peptides gave pronounced differences between the two sequences after overnight paper chromatography (Figure 1B). The partial acid hydrolysate of the native peptide had the pattern of the partial hydrolysate of Thr-Val-Val-Gln-Arg. A slow-moving spot and a fast-moving spot were present in the hydrolysate of synthetic Thr-Val-Val-Gln-Arg and TMVP peptide 2[1] both of which were absent in the hydrolysate of Thr-Val-Gln-Val-Arg. This difference in chromatographic behavior was probably due to the resistance of the Val–Val bond to acid hydorlysis since the Val–Val containing synthetic sequence and the na-

TABLE II: Results of Edman Degradation of Peptides.

Peptide	Edman Step	PTH Identified by	N-Terminal Amino Acid (% yield)
Synthetic Thr-Val-Gln-Val-Arg			
	1st	Thin-layer chromatography	Thr, 60%
	2nd	Thin-layer chromatography	Val, 30%; Thr (trace)
	3rd	Thin-layer chromatography	Nothing
	4th[a,b]	Thin-layer chromatography	Nothing
	5th	Thin-layer chromatography	Nothing
Synthetic Thr-Val-Val-Gln-Arg			
	1st	Thin-layer chromatography	Thr, 60%
	2nd	Thin-layer chromatography	Val, 30%
	3rd[b]	Thin-layer chromatography; amino acid analysis	Val, 18%; Gln, <1%
	4th	Thin-layer chromatography	Gln <5%; Val, <5%
TMVP peptide 2			
	1st	Thin-layer chromatography	Thr, 65%
	2nd	Thin-layer chromatography	Val, 30%
	3rd[b]	Thin-layer chromatography; amino acid analysis	Val, 18%; Gln, 3%
	4th	Thin-layer chromatography	Val ≤5%; Gln, ≤5%

[a] After this step the peptide was eluted and treated with acid in an attempt to open any pyrrolidonecarboxylyl residues prior to amino acid analysis and further degradation. [b] Half of the sample after this degradation was subjected to amino acid analysis and the remaining half used in the next degradation step.

tive peptide could not be completely hydrolyzed to the expected valine content by the usual overnight hydrolysis (Table I). Anderer *et al.* (1965) also used a 72-hr hydrolysis to obtain a complete hydrolysis of the native peptide.

The results of the stepwise degradation of each of the peptides are given in Table II. The low yields of phenylthiohydantoins were likely due to the paper strip method, as in general much higher yields can be obtained when carrying out the degradation in solution (Rombauts, 1966). We found, however, that we could not use the solution method since these peptides were repeatedly extracted into the benzene phase used to remove excess reagent after coupling with phenyl isothiocyanate. Synthetic Thr-Val-Gln-Val-Arg gave Thr-Val(Gln,Val,Arg) with no amino acids released after the second step even after acid treatment to open any pyrrolidonecarboxylyl ring. The amino acid composition of the peptide after the fourth degradation was: $Glu_{(0.65)}$,$Val_{(1.13)}$,$Arg_{(1)}$. Synthetic Thr-Val-Val-Gln-Arg gave Thr-Val-Val(Gln,-Arg) with very little released after the third step. The amino acid composition after the third step was: $Glu_{(0.77)}$,$Val_{(0.23)}$,$Arg_{(1)}$ with 60% recovery of the expected Arg. TMVP peptide 2 gave Thr-Val-Val(Gln,Arg). The amino acid composition after the third step was: $Glu_{(0.91)}$,$Val_{(0.59)}$,$Arg_{(1)}$ with a 78% recovery of the expected arginine giving analyses similar to that of the synthetic Thr-Val-Val-Gln-Arg.

A number of dipeptide or tripeptide sequences in proteins have been reported which have later been either inverted or reordered. Some residues or sequences appear to be more difficult to determine than others; for example, the asparagine or aspartic acid residue is very often involved, and Ile-Asn was inverted in both TMVP (residues 125–126) and lysozyme (residues 58–59). It is possible that similar "undiscovered" rearrangements exist in the published literature and may occur in the future especially where only one laboratory is involved in a protein determination and that synthesis may prove to be a definitive tool for ordering difficult short sequences.

Acknowledgments

We wish to thank Dr. Ben F. Feingold, the director of Kaiser Laboratory, for his interest and encouragement in this work. Other members of the Kaiser Laboratory which are acknowledged for their advice and discussions are Drs. E. Benjamini, D. Michaeli, and J. Kettman, and Mr. M. Shimizu. We also wish to thank Mr. Shimizu for the diagrams of the chromatograms. Dr. John Stewart of Rockefeller University is thanked for introducing us to solid-phase peptide chemistry and for his continuing advice on problems arising in peptide synthesis.

W. R. wishes to thank Dr. H. Fraenkel-Conrat, in whose laboratory the Edman degradation was done, for his hospitality and advice. He also thanks Mr. Hans

Kremer for technical assistance and Messrs. Vladimir Sturm and Ron Stickel for help with the amino acid analyzer.

References

Africa, B., and Carpenter, F. H. (1966), *Biochem. Biophys. Res. Commun. 24*, 113.
Anderer, F. A., Uhlig, H., Weber, E., and Schramm, G. (1960), *Nature 186*, 922.
Anderer, F. A., Wittmann-Leibold, B., and Wittmann, H. G. (1965), *Z. Naturforschg. 20b*, 1203.
Benjamini, E., Young, J. D., Shimizu, M., and Leung, C. Y. (1964), *Biochemistry 3*, 1115.
Bornstein, P. (1967), *Biochemistry 6*, 3082.
Doolittle, R. F. (1964), *Biochem. J. 94*, 742.
Edman, P., and Begg, G. (1967), *European J. Biochem. 1*, 80.
Fraenkel-Conrat, H. (1954), *J. Am. Chem. Soc. 76*, 3606.
Fraenkel-Conrat, H. (1957), *Virology 4*, 1.
Funatsu, G. (1964), *Biochemistry 3*, 1351.
Jeppsson, J. O., and Sjöquist, J. (1967), *Anal. Biochem. 18*, 264.
Merrifield, R. B. (1964), *Biochemistry 3*, 1385.
Ramachandran, L. K., and Gish, D. T. (1958), *J. Am. Chem. Soc. 81*, 884.
Rombauts, W. A. (1966), *Biochem. Biophys. Res. Commun. 23*, 549.
Schroeder, W. A., Shelton, J. R., Shelton, J. B., Cormick, J., and Jones, R. T. (1963), *Biochemistry 2*, 992.
Sjöquist, J. (1953), *Acta Chem. Scand. 7*, 447.
Sjöquist, J. (1959), *Arkiv Kemi 14*, 291.
Spackman, D. H. (1963), *Federation Proc. 22*, 244.
Stark, G. R., and Smyth, D. G. (1963), *J. Biol. Chem. 238*, 214.
Stewart, J. M., Young, J. D., Benjamini, E., Shimizu, M., and Leung, C. Y. (1966), *Biochemistry 5*, 3396.
Tsugita, A., Gish, D. T., Young, J., Fraenkel-Conrat, H., Knight, C. A., and Stanley, W. M. (1960), *Proc. Natl. Acad. Sci. U. S. 46*, 1463.
Young, J. D., Benjamini, E., Stewart, J. M., and Leung, C. Y. (1967), *Biochemistry 6*, 1455.