THE JOURNAL OF BIOLOGICAL CHEMISTRY
© 1989 by The American Society for Biochemistry and Molecular Biology, Inc.

Vol. 264, No. 11, Issue of April 15, pp. 6273-6279, 1989
Printed in U.S.A.

Primary Structure of Belladonna Mottle Virus Coat Protein*

(Received for publication, August 15, 1988)

S. Suryanarayana‡, N. Appaji Rao, Mathur R. N. Murthy§, and Handanahal S. Savithri¶

From the Department of Biochemistry and the §Molecular Biophysics Unit, Indian Institute of Science, Bangalore 560 012, India

The coat protein of belladonna mottle virus (a tymovirus) was cleaved by trypsin and chymotrypsin, and the peptides were separated by high performance liquid chromatography using a combination of gel permeation, reverse phase, and ion pair chromatography. The peptides were sequenced manually using the 4-*N*, *N*-dimethylaminoazobenzene-4′-isothiocyanate/phenyl isothiocyanate double-coupling method. The chymotryptic peptides were aligned by overlapping sequences of tryptic peptides and by homology with another tymovirus, eggplant mosaic virus. The belladonna mottle virus is more closely related to eggplant mosaic virus than to turnip yellow mosaic virus, the type member of this group, as evident from the sequence homologies of 57 and 32%, respectively. The accumulation of basic residues at the amino terminus implicated in RNA-protein interactions in many spherical plant viruses was absent in all the three sequences. Interestingly, the amino-terminal region is the least conserved among the tymoviruses. The longest stretch of conserved sequence between belladonna mottle virus and eggplant mosaic virus was residues 34–44, whereas it was residues 96–102 in the case of belladonna mottle virus and turnip yellow mosaic virus. A tetrapeptide in the region (residues 154–157) was found to be common for all the three sequences. It is possible that these conserved regions (residues 34–44, 96–102, 154–157) are involved in either intersubunit or RNA-protein interactions.

BDMV[1] belongs to the tymogroup of plant viruses whose particles are stabilized predominantly by protein-protein interactions (1). It consists of a single-stranded RNA genome encapsidated in an icosahedral shell of 180 identical coat protein subunits of molecular weight ~20,000. The virus particles purified by sucrose density gradient centrifugation contain calcium ions and polyamines (2). The exchange of polyamines by monovalent cations like potassium or cesium results in the loss of stability of virions upon increasing the pH to 8.0 or above (2). The determination of the amino acid sequence of BDMV coat protein was envisaged with a view to locating probable cation binding sites, sites of protein-protein and protein-RNA interactions, and to obtain information on the evolution of tymoviruses. Furthermore, the sequence would aid the interpretation of electron density maps of BDMV being obtained through x-ray diffraction studies (3). In an earlier investigation, it was shown that the amino terminus of BDMV coat protein is blocked by an acetyl group, and the amino acid sequence of a 43-residue-long CNBr fragment, arising from the carboxyl-terminal end, was determined (4). In this paper, we report the primary structure of BDMV coat protein obtained by sequencing chymotryptic and tryptic peptides purified on high performance liquid chromatography. The amino acid sequence of BDMV coat protein is compared with that of TYMV (5) and EMV (6).

MATERIALS AND METHODS AND RESULTS[2]

DISCUSSION

Reliability of the Sequence Determination—The primary structure of BDMV coat protein is presented in Fig. 1. The amino acid sequence was obtained essentially from an analysis of purified chymotryptic peptides. Most of the overlapping sequences were obtained from tryptic peptides. The peptides CH4–CH7 were aligned solely by comparison with the sequence of EMV coat protein (6). The sequence presented in Fig. 1 agreed very well with the composition derived by amino acid analysis except for an extra Val and Ile in the sequence and 1 Ala in the composition (Table 1). The low yields of Val and Ile were probably due to the presence of bonds like Val-Val, Ile-Leu. Molecular weight of the protein derived from the amino acid sequence (20,219) agreed with that determined by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (20,000).

Amino Acid Composition of BDMV—Like other tymoviruses, BDMV coat protein contained an unusually large number of Pro residues. This could be attributed to a large excess of cytosine in the coat protein messenger RNA (7). However, this argument does not account for the abundance of Thr in the composition. On the other hand, the Gly content was lower than the average value for soluble proteins (8). The average residue volume was 130.9, 130.8, and 128.3 Å^3 for BDMV, TYMV, and EMV, respectively, which agreed well with the average value of 131.6 Å^3 calculated for 207 unrelated proteins (8, 9). The total number of acidic amino acids in BDMV was 12 compared with 16 Lys and Arg. This slight excess of positive charge is insufficient to neutralize to any appreciable extent the charges on the RNA. Hence, additionally bound cations may be required for the stability of the

* 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.

‡ Recipient of University Grants Commission Research Fellowship.

¶ To whom correspondence should be addressed.

[1] The abbreviations used are: BDMV, belladonna mottle virus; DABITC, 4-*N*,*N*-dimethylaminoazobenzene-4′-isothiocyanate; DABTH, 4-*N*,*N*-dimethylaminoazobenzene-4′-thiohydantoin; PITC, phenylisothiocyanate; EMV, eggplant mosaic virus; TYMV, turnip yellow mosaic virus; SBMV, southern bean mosaic virus; HPLC, high performance liquid chromatography; RP, reverse phase; Ac, acetyl; Cm, carboxymethyl; TFA, trifluoroacetic acid; NH_4OAC, ammonium acetate; Gm-HCl, guanidine HCl.

[2] Portions of this paper (including "Materials and Methods," part of "Results," Figs. 4–9, and Tables 1–3) are presented in miniprint at the end of this paper. Miniprint is easily read with the aid of a standard magnifying glass. Full size photocopies are included in the microfilm edition of the Journal that is available from Waverly Press.

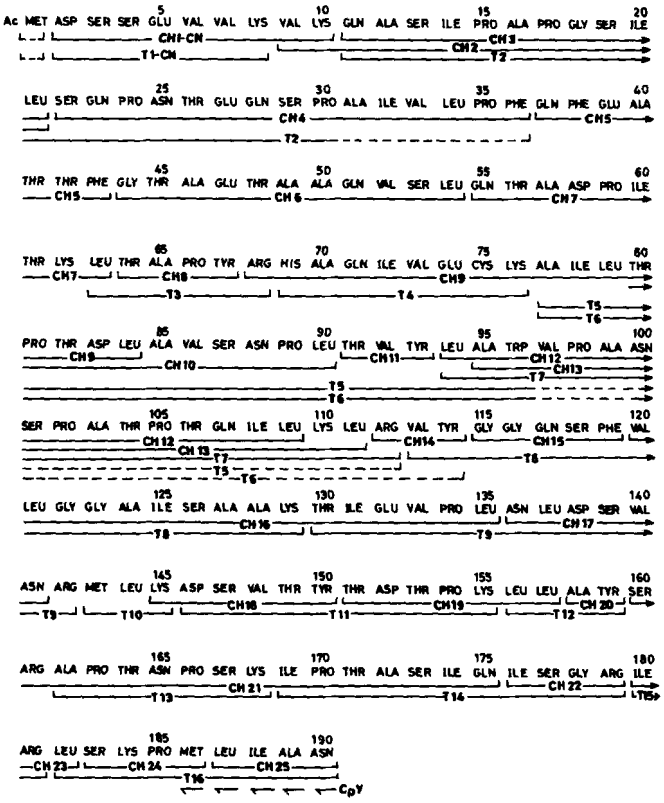


FIG. 1. **The complete amino acid sequence of BDMV coat protein.** Peptides obtained by chymotrypsin and trypsin digestion of carboxymethylated protein are designated as *CH* and *T*, respectively. *CH1-CN* and *T1-CN* refer to peptides obtained by subcleavage of *CH1* and *T1* with CNBr. *CpY* indicates the amino acids released by carboxypeptidase Y digestion. *Dotted line* indicates the regions not assigned by sequence analysis but compatible with amino acid composition of peptides.

virus particles. The sequence contained 1 His residue in contrast to 3 each in TYMV and EMV. The thermal stability of BDMV was not altered between pH 5.5 and 8.5, whereas that of TYMV underwent a transition near the pK_a of His. This difference was attributed to the absence of histidines in BDMV when compared with TYMV (10). However, the present analysis revealed that the BDMV coat also contains a His, raising some doubts on the validity of the explanation for instability of BDMV suggested by earlier workers (10).

Comparison of Tymoviral Coat Protein Sequences—Fig. 2 shows the alignment of the coat protein sequences of BDMV, EMV, and TYMV. The alignment was obtained by matrix methods using the Needleman and Wunsch procedure (11). The weighting scheme was based on Dayhoff's log-odds matrix (12) with a gap penalty of 6. The matrix elements were made positive by addition of 6 to all the elements. The most striking difference between the tymoviral sequences was the presence of two additional residues Leu-Arg after residue 110 in BDMV. This insertion was flanked by regions that are well conserved.

Dupin *et al.* (13) introduced a gap after the second residue in their comparison of primary sequences of EMV with TYMV, whereas the procedure used in this study suggests a gap after residue 19. With the gap introduced in this position, the score between TYMV and EMV was marginally higher than the score obtained with a gap after the 2nd residue in TYMV. Such ambiguity in alignment was not surprising in the light of the low degree of sequence homology at the amino termini. Introduction of a gap after residue 17 in TYMV sequence resulted in optimal alignment with the sequence of BDMV in this region. The number of residues conserved between BDMV, EMV; BDMV, TYMV; EMV, TYMV; are 109, 62, and 60, respectively. The percentage homology between these three pairs including the gaps was 56.8, 32.3, and 31.2. It is clear that TYMV is evolutionarily nearly equidistant from BDMV and EMV, whereas the latter two are more closely related. This conclusion is in agreement with the observation that antisera to BDMV or EMV cross-react with EMV or BDMV but not with TYMV. Antiserum to TYMV cross-reacts weakly with BDMV and not with EMV (14).

The distribution of the lengths of peptides conserved between BDMV and EMV or TYMV did not agree with that expected for this degree of homology. An 11-residue peptide (residues 34–44) was conserved between BDMV and EMV, whereas the theoretical distribution did not predict conserved sequences longer than 6 residues. A tripeptide and a total of 5 residues were conserved in all the three sequences in this region. Similarly, between BDMV and TYMV, a heptapeptide (residues 96–102) was conserved whereas the maximum length expected was only 4 residues. Two dipeptides were conserved in all the three sequences in this heptapeptide. Therefore, these regions are likely to have structural or functional importance. In contrast, the distribution of the sizes of conserved peptides between EMV and TYMV corresponded well with the expected distribution (13). One tetrapeptide, four tripeptides, five dipeptides, and 19 single amino acids were conserved in all the three sequences, representing 23.8% of the sequences. Dupin *et al.* (13) pointed out that the conserved tetrapeptide and the flanking regions with good homology are likely sites of protein-nucleic acid interactions in tymoviruses.

The secondary structure of the coat proteins of BDMV, EMV, and TYMV, predicted by the method of Chou and Fasman (33), did not reflect the similarities anticipated from their sequence homology. The structure predictions suggested 20, 21, and 13 helix and 20, 16, and 41% β-structure for BDMV, EMV, and TYMV, respectively. However, cross-rotation functions[3] computed using x-ray diffraction data obtained from crystals of BDMV (3) and SBMV (16) revealed that their three-dimensional structures were similar. An examination of the structure of SBMV showed 16% helix and 33% β-structure (16). This discrepancy in the predicted β-structure of BDMV and EMV was probably due to inherent limitations of the procedure employed (33).

Amino Acid Exchanges—Comparison of the degree of changes of different amino acids in these sequences to the relative mutabilities of amino acids showed that Gly was strongly conserved. Six glycines were found at common positions in all the sequences, of which 4 occurred as Gly-Gly dipeptides. Thus, these glycines are likely to be important for the tertiary structure of the coat proteins. In general, the small polar residues were better conserved than larger hydrophobic residues. These results suggested that the amino acid exchanges in evolutionarily related viral coat proteins might depart appreciably from the values observed for families of homologous proteins. However, an analysis of the amino acid exchanges in SBMV and tomato bushy stunt virus coat proteins suggested that the exchanges were similar to those of globular proteins (15). Hence, the pattern of exchanges in tymoviruses probably reflects the unique structural constraints in these viral capsids.

Hydropathy Plots and Antigenic Determinants—Fig. 3 depicts the hydropathic character of BDMV, EMV, and TYMV coat proteins based on hydropathy indices devised by Kyte

[3] M. R. N. Murthy, S. K. Munshi, and C. N. Hiremath, unpublished results.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
BDMV	MET	ASP	SER	SER	GLU	VAL	VAL	LYS	VAL	LYS	GLN	ALA	SER	ILE	PRO	ALA	PRO	GLY	SER	ILE
EMV	MET	GLU	ASP	THR	ALA	ILE	ILE	ARG	SER	PRO	GLN	PRO	SER	ILE	ASN	ALA	PRO	GLY	PHE	HIS
TYMV	MET	GLU	ILE	ASP	LYS	GLU	LEU	ALA	PRO	GLN	ASP	ARG	THR	VAL	THR	VAL	ALA	•••	THR	VAL
	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
BDMV	LEU	SER	GLN	PRO	ASN	THR	GLU	GLN	SER	PRO	ALA	ILE	VAL	LEU	PRO	PHE	GLN	PHE	GLU	ALA
EMV	LEU	PRO	PRO	THR	ASP	SER	GLN	GLN	SER	SER	ALA	ILE	GLU	LEU	PRO	PHE	GLN	PHE	GLU	ALA
TYMV	LEU	PRO	ALA	VAL	PRO	GLY	PRO	SER	PRO	LEU	THR	ILE	LYS	GLN	PRO	PHE	GLN	SER	GLU	VAL
	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60
BDMV	THR	THR	PHE	GLY	THR	ALA	GLU	THR	ALA	ALA	GLN	VAL	SER	LEU	GLN	THR	ALA	ASP	PRO	ILE
EMV	THR	THR	PHE	GLY	ALA	THR	GLU	THR	ALA	ALA	GLN	ILE	SER	LEU	ALA	SER	ALA	ASN	ALA	ILE
TYMV	LEU	PHE	ALA	GLY	THR	LYS	ASP	ALA	GLU	ALA	SER	LEU	THR	ILE	ALA	ASN	ILE	ASP	SER	VAL
	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80
BDMV	THR	LYS	LEU	THR	ALA	PRO	TYR	ARG	HIS	ALA	GLN	ILE	VAL	GLU	CYS	LYS	ALA	ILE	LEU	THR
EMV	THR	LYS	LEU	ALA	SER	LEU	TYR	ARG	HIS	VAL	ARG	LEU	THR	GLN	CYS	ALA	ALA	THR	ILE	THR
TYMV	SER	THR	LEU	THR	THR	PHE	TYR	ARG	HIS	ALA	SER	LEU	GLU	SER	LEU	TRP	VAL	THR	ILE	HIS
	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
BDMV	PRO	THR	ASP	LEU	ALA	VAL	SER	ASN	PRO	LEU	THR	VAL	TYR	LEU	ALA	TRP	VAL	PRO	ALA	ASN
EMV	PRO	THR	ALA	ALA	ALA	ILE	ALA	ASN	PRO	LEU	THR	VAL	ASN	ILE	VAL	TRP	VAL	SER	ASP	ASN
TYMV	PRO	THR	LEU	GLN	ALA	PRO	THR	PHE	PRO	THR	THR	VAL	GLY	VAL	CYS	TRP	VAL	PRO	ALA	ASN
	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120
BDMV	SER	PRO	ALA	THR	PRO	THR	GLN	ILE	LEU	LYS	LEU	ARG	VAL	TYR	GLY	GLY	GLN	SER	PHE	VAL
EMV	SER	THR	ALA	LYS	PRO	THR	GLU	ILE	LEU	ASN	•••	•••	VAL	PHE	GLY	GLY	SER	SER	TYR	THR
TYMV	SER	PRO	VAL	THR	PRO	ALA	GLN	ILE	THR	LYS	•••	•••	THR	TYR	GLY	GLY	GLN	ILE	PHE	CYS
	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140
BDMV	LEU	GLY	GLY	ALA	ILE	SER	ALA	ALA	LYS	THR	ILE	GLU	VAL	PRO	LEU	ASN	LEU	ASP	SER	VAL
EMV	PHE	GLY	GLY	ALA	LEU	ASN	ALA	THR	LYS	PRO	LEU	THR	ILE	PRO	LEU	PRO	MET	ASN	SER	VAL
TYMV	ILE	GLY	GLY	ALA	ILE	ASN	THR	LEU	SER	PRO	LEU	ILE	VAL	LYS	CYS	PRO	LEU	GLU	MET	MET
	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160
BDMV	ASN	ARG	MET	LEU	LYS	ASP	SER	VAL	THR	TYR	THR	ASP	THR	PRO	LYS	LEU	LEU	ALA	TYR	SER
EMV	ASN	CYS	MET	LEU	LYS	ASP	SER	VAL	LEU	TYR	THR	ASP	CYS	PRO	LYS	LEU	LEU	ALA	TYR	SER
TYMV	ASN	PRO	ARG	VAL	LYS	ASP	SER	ILE	GLN	TYR	LEU	ASP	SER	PRO	LYS	LEU	LEU	ILE	SER	ILE
	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180
BDMV	ARG	ALA	PRO	THR	ASN	PRO	SER	LYS	ILE	PRO	THR	ALA	SER	ILE	GLN	ILE	SER	GLY	ARG	ILE
EMV	ALA	ALA	PRO	SER	SER	PRO	SER	LYS	THR	PRO	THR	ALA	THR	ILE	GLN	ILE	HIS	GLY	LYS	LEU
TYMV	THR	ALA	GLN	PRO	THR	ALA	PRO	PRO	ALA	SER	THR	CYS	ILE	ILE	THR	VAL	SER	GLY	THR	LEU
	181	182	183	184	185	186	187	188	189	190	191									
BDMV	ARG	LEU	SER	LYS	PRO	MET	LEU	ILE	•••	•••	ALA	ASN								
EMV	ARG	LEU	SER	SER	PRO	LEU	LEU	GLN	•••	•••	ALA	ASN								
TYMV	SER	MET	HIS	SER	PRO	LEU	ILE	THR	ASP	THR	SER	THR								

FIG. 2. **Comparison of the primary structures of BDMV, EMV, and TYMV coat proteins.** The sequences were aligned using matrix methods. The numbering corresponds to BDMV coat protein sequence. The regions proposed to be involved in intersubunit or RNA-protein interactions are *underlined*.

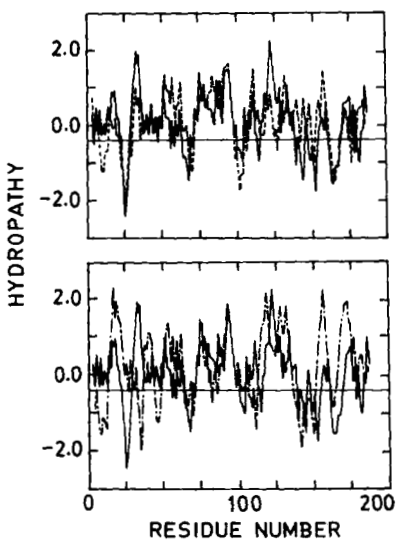


FIG. 3. **Hydropathy profiles of the coat protein sequences of BDMV (——), EMV (– – – –) (*top*) and BDMV (——), TYMV (— – —) (*bottom*).** Hydropathy profiles were obtained according to Kyte and Doolittle (17), choosing 11-residue segments for averaging the values.

and Doolittle (17). In the case of soluble, globular proteins, this scale predicts confidently buried and exposed regions as maxima and minima of the plots. As anticipated, the plots were very similar for BDMV and EMV, whereas larger differences were found between TYMV and the other two viruses. In particular an amino-terminal segment (residues 15–23) and a carboxyl-terminal segment (residues 168–175) were more hydrophobic in TYMV. In general, the three plots were very similar in the central region of the sequences when compared with their termini. Similar observations were made when the sequences of SBMV and tomato bushy stunt virus were compared (15). The average hydropathy index over the entire length of the proteins was 0.1 for BDMV and EMV, whereas it was 0.25 for TYMV. This index had an average value −0.4 for globular proteins and for membrane-bound proteins the range was 0.28–0.79 (17). This observation is of interest in the light of the conclusion that the stability of these viruses is based on hydrophobic interactions between their coat protein subunits. Among the five experimentally mapped antigenic determinants of TYMV (segments 1–12, 33–45, 57–64, 143–152, and 183–189), four fall in the hydrophilic regions of the plots (segments 1–12, 33–45, 143–152, and 183–189). In contrast, the hydrophilicity values of Hopp and Woods (18) were not well correlated with antigenicity of TYMV (13). The antigenic determinants in BDMV and EMV have not yet been mapped.

The amino acid sequence of the coat proteins of several plant viruses exhibits a nonrandom distribution of charged residues with positive charges appearing frequently at the amino terminus. This basic terminal was shown to interact with the RNA (19, 20) and direct the mode of assembly in some viruses (20). Also, in the three-dimensional structure of SBMV (16) and tomato bushy stunt virus (21), this region was disordered presumably because of its interactions with the disordered RNA. In contrast, tymoviral sequences exhibit a random distribution of charged residues all along the sequence. Furthermore, these charged residues are not particularly well conserved. This suggests that the mode of assembly in tymoviruses is likely to be different.

Partial charge neutralization and stability in tymoviruses was derived from bound polyamines. These polyamines are likely to be bound to the nucleic acid as they are co-isolated with the viral RNA (Ref. 22 and Footnote 4). The exchangeability of polyamines with other cations and its effect on particle stability have been studied in BDMV (2) and TYMV (23). The polyamines in BDMV were easily exchanged for other cations, whereas drastic conditions were required in the case of TYMV. The difference in exchangeability probably reflects differential permeability of the capsids. Permeability in turn would depend on the hydrophobicity of the ion access channels likely to be present in the intersubunit region. The regions of higher hydrophobicity in TYMV could account for the observed differences.

[4] H. S. Savithri and A. N. K. Jacob, unpublished results.

Acknowledgments—We thank T. U. Sahana for her technical assistance in the purification of the virus. Several of our former colleagues generously donated chemicals which made this work possible. Thanks are due to the Department of Science and Technology for financial support.

REFERENCES

1. Kaper, J. M. (1975) in *Frontiers of Biology* (Neuberger, A., and Tatum, E. L., eds) pp. 273–320, American Elsevier, New York
2. Savithri, H. S., Munshi, S. K., Suryanarayana, S., Divakar, S., and Murthy, M. R. N. (1987) *J. Gen. Virol.* **68,** 1533–1542
3. Munshi, S. K., Hiremath, C. N., Murthy, M. R. N., and Savithri, H. S. (1987) *Acta Crystallogr.* **B43,** 376–382
4. Suryanarayana, S., Rajgopal Rao, D., Appaji Rao, N., and Savithri, H. S. (1987) *Biochem. Int.* **15,** 65–72
5. Peter, R., Stehelin, D., Reinbolt, J., Collot, D., and Duranton, H. (1972) *Virology* **49,** 615–617
6. Dupin, A., Peter, R., Collot, D., Das, B. C., Peter, C., Bouillon, P., and Duranton, H. (1984) *C. R. HEBD. Seances Acad. Sci. Ser. C Sci. Chim.* **298,** 219–221
7. Koenig, R., and Lesemann, D.-E. (1979) *Tymovirus Group,* Commonwealth Mycological Institute/Association of Applied Biologists Descriptions of Plant Viruses No. 214, Commonwealth Mycological Institute, Surrey, Great Britain
8. Klapper, M. H. (1977) *Biochem. Biophys. Res. Commun.* **78,** 1018–1024
9. Zamyatnin, A. A. (1972) *Prog. Biophys. Mol. Biol.* **24,** 107–123
10. Virudachalam, R., Harrington, M., and Markley, J. L. (1985) *Virology* **146,** 138–140
11. Needleman, S. B., and Wunsch, C. D. (1970) *J. Mol. Biol.* **48,** 443–453
12. Dayhoff, M. O., Schwartz, R. M., and Orcutt, B. C. (1979) in *Atlas of Protein Sequence and Structure* (Dayhoff, M. O., ed), Vol. 5, Suppl. 3, pp. 345–352, National Biomedical Research Foundation, Washington, D. C.
13. Dupin, A., Collot, D., Peter, R., and Witz, J. (1985) *J. Gen. Virol.* **66,** 2571–2579
14. Koenig, R. (1976) *Virology* **72,** 1–5
15. Hopper, P., Harrison, S. C., and Sauer, R. T. (1984) *J. Mol. Biol.* **177,** 701–713
16. Abad-Zapatero, C., Abdel-Meguid, S. S., Johnson, J. E., Leslie, A. G. W., Rayment, I., Rossmann, M. G., Suck, D., and Tsukihara, T. (1980) *Nature* **286,** 33–39
17. Kyte, J., and Doolittle, R. F. (1982) *J. Mol. Biol.* **157,** 105–132
18. Hopp, T. P., and Woods, K. R. (1981) *Proc. Natl. Acad. Sci. U. S. A.* **78,** 3824–3828
19. Vriend, G., Verduin, B. J. M., and Hemminga, M. A. (1986) *J. Mol. Biol.* **191,** 453–460
20. Rossmann, M. G., Abad-Zapatero, C., Erickson, J. W., and Savithri, H. S. (1983) *J. Biomol. Struct. Dyn.* **1,** 565–579
21. Harrison, S. C., Olson, A. J., Schutt, C. E., Winkler, F. K., and Bricogne, G. (1978) *Nature* **276,** 368–373
22. Katouzian-Safadi, M., Charlier, M., and Maurizot, J.-C. (1985) *Biochimie* **67,** 1007–1013
23. Noort, A., van Den Dries, C. L. A. M., Pleij, C. W. A., Jaspars, E. M. J., and Bosch, L. (1982) *Virology* **120,** 412–421
24. Chang, J.-Y., Brauer, D., and Wittmann-Liebold, B. (1978) *FEBS Lett.* **93,** 205–214
25. Yang, C.-Y., Pauly, E., Kratzin, H., and Hilschmann, N. (1981) *Hoppe-Seyler's Z. Physiol. Chem.* **362,** 1131–1146
26. Liu, T.-Y. (1972) *Methods Enzymol.* **25,** 44–55
27. Liu, T.-Y., and Chang, Y. H. (1971) *J. Biol. Chem.* **246,** 2842–2848
28. Hayashi, R., Moore, S., and Stein, W. H. (1973) *J. Biol. Chem.* **248,** 2296–2302
29. Lee, H.-M., and Riordan, J. F. (1978) *Biochem. Biophys. Res. Commun* **85,** 1135–1142
30. Ambler, R. P. (1972) *Methods Enzymol.* **25,** 143–154
31. Runswick, M. J., and Walker, J. E. (1983) *J. Biol. Chem.* **258,** 3081–3089
32. Yang, C.-Y. (1979) *Hoppe-Seyler's Z. Physiol. Chem.* **360,** 1673–1675
33. Chou, P. Y., and Fasman, G. D. (1978) *Annu. Rev. Biochem.* **47,** 251–276
34. Irvine, G. B., and Shaw, C. (1986) *Anal. Biochem.* **155,** 141–148
35. Irvine, G. B. (1987) *J. Chromatogr.* **404,** 215–222

Supplementary Material to
Primary Structure of Belladonna Mottle Virus Coat Protein
by
S. Suryanarayana, N. Appaji Rao, Mathur R.N. Murthy and
Handanahal S. Savithri

EXPERIMENTAL PROCEDURES

Materials. Trypsin (N-tosyl-L-phenylalanyl chloromethyl ketone-treated), chymotrypsin (α -N-tosyl-L-lysyl chloromethyl ketone-treated), tryptamine, HCl, pepstatin A, iodoacetic acid, cyanogen bromide, DABITC were obtained from Sigma Chemical Company (St. Louis, MO). Carboxypeptidase Y, p-toluenesulfonic acid, PITC were obtained from Pierce Chemical Company (Rockford, IL). Gm-HCl and Tris were obtained from Schwarz/Mann (Organgeburg, N.Y.). For manual sequencing, pyridine, trifluoroacetic acid, DABITC were purified as described by Chang (24). Polyamide thin layer sheets were obtained from Schleicher & Schuell (Dassel). All other reagents were of the highest grade available from commercial sources.

HPLC columns. Gel filtration column TSK-G 2000 SW (7.5 x 600 mm, 10 ± 2 μm) and Lichrosorb RP-18 (4 x 250 mm, 10 μm) were obtained from LKB (Sweden). Spherisorb 5 ODS (4.6 x 250 mm, 5 μm) and Aquapore RP-300 (4.6 X 250 mm, 7 μm) were obtained from Pye Unicam (U.K.).

Purification and Carboxymethylation of BDMV Coat Protein. BDMV was propagated on *Nicotiana glutinosa* plants. Purification of BDMV and isolation of coat protein was performed as described earlier (4). The lyophilized BDMV coat protein (25 mg) was dissolved in 1.25 ml of 0.3 M Tris-HCl, pH 8.3 containing 6M Gm-HCl and 2 mM EDTA. The protein was reduced with dithiothreitol and carboxymethylated using iodoacetic acid. Gm-HCl was removed slowly by dialyzing against 3 x 100 ml of 0.1M N-ethylmorpholine-acetate, pH 8.2 for 24h. This protein suspension was used for enzymic cleavages.

Enzymic Cleavages:

Chymotrypsin digestion: BDMV coat protein (20 mg) in about 2 ml of 0.1M N-ethylmorpholine-acetate, pH 8.2 was hydrolyzed with chymotrypsin (enzyme to substrate ratio was 1:50 (w/w)). The digestion was carried out at 37°C for 24h with continuous stirring and stopped by lyophilization.

Trypsin digestion: Five mg of BDMV coat protein in 0.5 ml of 0.1M N-ethylmorpholine-acetate, pH 8.2 was incubated with trypsin at 37°C for 4 h with continuous stirring. Enzyme to substrate ratio was 1:50 (w/w). At the end of the digestion, the peptide mixture was centrifuged and the supernatant was collected. The insoluble fraction (4% w/w) was washed with distilled water (3 x 200 μl) and the washings were pooled with the soluble fraction. Both soluble and insoluble fractions were lyophilized seperately.

CNBr Cleavage: Peptides CH1 and T1 (approx. 5 nmol each) were seperately digested in 0.5 ml of 70% formic acid with 100 times molar excess of CNBr over Met residues. The reaction mixture was incubated at 37°C for 24h; diluted with water and lyophilized. The digest was directly sequenced without further purification.

Peptide Purifications: All peptides were purified by HPLC (LKB, Sweden).

Chymotryptic peptides: The lyophilized chymotryptic digest of the protein was dissolved in 0.1% TFA containing 35% methanol and 6M Gm-HCl. Aliquots (200 μl,2 mg) of the digest were loaded on to a TSK-G 2000 SW gel permeation column equilibrated with 0.1% TFA containing 35% methanol. The column was developed with the same solvent. Peak fractions (see Fig.5) were lyophilized. Each of these fractions were redissolved in 0.1% TFA containing 6M Gm-HCl. Aliquots of 50 μl (approx. 400 μg) were loaded on to Lichrosorb RP-18 column. The peptides were eluted using solvent system I which consisted of 0.1% TFA in water (A) and 0.1% TFA containing 70% acetonitrile (B). Most of the peak fractions were further purified by ion-pair chromatography using a Spherisorb 5 ODS column. The column was developed using solvent system II which consisted of 25 mM ammonium acetate, pH 6.0 (A) and 40% of 50 mM ammonium acetate, pH 6.0 plus 60% acetonitrile (B) (25).

Tryptic Peptides: Soluble fraction of the tryptic digest was dissolved in 0.1% TFA containing 6M Gm-HCl. Aliquots of 50 μl (~400 μg) of tryptic digest were loaded on to Spherisorb 5 ODS column and developed using solvent system I described above. All the peak fractions were concentrated by lyophilization and further purified by chromatography on Aquapore RP-300 using solvent system II.

Purity of the peptides at each step of purification was checked by end-group analysis adopting DABITC/PITC double coupling method (24).

Amino Acid Analysis. Carboxymethylated BDMV coat protein (10 nmol) was hydrolysed in 200 μl of 3 N p-tolunesulfonic acid solution containing 0.2% tryptamine (tryptamine was prepared from its hydrochloride as described by Liu (26)) and 10 nmol of norleucine as an internal standard (27) in evacuated tubes. Hydrolysis was performed at 110°C for 22h, 48h and 72h. At the end of the hydrolysis, the pH was adjusted to around 2.0 by adding 5N LiOH (80 μl) and directly loaded on to an LKB Alpha analyzer using lithium buffers and ninhydrin detection system.

The peptides (2-5 nmol) were hydrolyzed in 200 μl of constant boiling HCl (5.7N) containing 0.1% (w/v) phenol and 0.05% (v/v) 2-mercaptoethanol. Hydrolysis was done in evacuated tubes at 110°C for 24h. After drying, the hydrolysates were analyzed on an LKB Alpha analyzer.

Tryptophan content of peptides CH12, CH13 and T7 were determined from the characteristic absorbance spectrum of tryptophan ($\epsilon_{280\ nm}$ = 5690).

Carboxypeptidase Y Digestion. Carboxymethylated BDMV coat protein (100 nmol) was digested with carboxypeptidase Y in 1 ml of 0.1M pyridine-acetate, pH 5.5 containing 6M urea at room temperature (28). Norleucine was used as an internal standard. Reaction was started by adding carboxypeptidase Y which had been previously treated with pepstatin A (29). Enzyme to substrate ratio was 1:100 (w/w). Aliquots were taken at different intervals of time and the digestion was stopped by adding 100 mg of Dowex 50 X 8 resin (H^+ form, mesh size 50 - 100 μm) (30), centrifuged and the settled resin was washed with distilled water (3 x 500 μl). Resin-bound amino acids were extracted into 5N ammonia (3 x 300 μl); the extract was dried, redissolved in loading buffer and applied on to the analyzer.

Sequence Determination: Peptides were sequenced by a scaled down version of DABITC/PITC double coupling method (24) as described by Runswick and Walker (31). DABTH-amino acid derivatives were identified by thin layer chromatography on polyamide sheets (24). DABTH-Ile and DABTH-Leu were distinguished by TLC on polyamide sheets using 10% formic acid-ethanol (10:9, v/v) as the solvent system (32). With this procedure 2-5 nmol of peptide could be sequenced upto the twentieth residue.

Nomenclature of Peptides: Chymotryptic and tryptic peptides are designated as CH and T respectively, followed by a number corresponding with their order in the sequence.

Sequence Analysis: The sequence of BDMV coat protein was compared and aligned with those of TYMV and EMV following the Needleman and Wunsch algorithm (11). The variation of the hydropathic character of peptides along the protein chain was displayed following the procedure of Kyte and Doolittle (17). The segment length for evaluating averages was 11 residues. Secondary structures were predicted according to Chou and Fasman (33).

RESULTS

Purification and Sequencing of Chymotryptic Peptides: The chymotryptic peptides were initially size-fractionated on TSK-G 2000 SW column (Fig. 5). Peaks I-IV were refractionated on Lichrosorb RP-18 column (Fig. 6a, b and c). The purity of the peptides was checked by N-terminal analysis at this stage. Only peptides CH2, 5, 15, 16 and 25 were found to be pure. Rest of the peptides were purified further by chromatography on Spherisorb 5 ODS column using ammonium acetate as an ion-pair reagent. The use of Spherisorb 5 ODS column with a smaller particle size (5 µm) and ammonium acetate resulted in the complete purification of the rest of the peptides. Fig.7a, b, c and d shows the purification of some of the peptides on Spherisorb 5 ODS column. Amino acid composition of the purified chymotryptic fragments are presented in Table 2. The sequence of all the peptides were determined completely from the N to the C terminus (Fig. 1). Peptides purified from peak IV did not yield any unique sequence.

The reduced and carboxymethylated BDMV coat protein was sparingly soluble in the buffer used for the chymotryptic digestion. This necessitated longer period of digestion (24h) to obtain a complete set of fragments. However, digestion for longer period of time resulted in the cleavage at minor sites (e.g. Arg$^{\downarrow}$Ile-Arg-Leu) and the generation of partially digested fragments. The digest consisted of a complex mixture of peptides which were separated by using TSK-G 2000 SW gel filtration, reverse phase chromatography on Lichrosorb RP-18 and Spherisorb 5 ODS HPLC columns. The chymotryptic cleavage also resulted in short fragments which were amenable for manual sequencing. Only the minimum peptide sequences required for establishing the complete primary structure are reported here, although a large number of subfragments was sequenced.

Purification of Tryptic Peptides: Tryptic peptides were initally fractionated on Spherisorb 5 ODS using TFA/acetonitrile solvent system (Fig.8). Each peak fraction was further purified on Aquapore RP-300 using ammonium acetate/acetonitrile solvent system (e.g. Fig.9). The amino acid composition of tryptic peptides is presented in Table 3. The sequence of all these peptides were determined manually (Fig.1). Peptides T2, T5 and T6 being bigger in size could be sequenced only partially as indicated by dotted lines in Fig.1. Apart from two small fragments (residues 9-10, 160-161) one longer tryptic fragment could not be traced. However, this longer fragment spanning the region 11-62 was probably part of the insoluble fraction of the tryptic digest which was not analyzed.

Alignment of Peptide Sequences: Most of the chymotryptic peptides could be aligned by homology with EMV coat protein sequence. The complete sequence, except for a tripeptide (142-144, Fig.1), was obtained from the purified chymotryptic peptides. The tryptic fragments provided many of the overlaps to chymotryptic fragments. The positioning of the chymotryptic peptide Arg-Val-Tyr, which resulted in the introduction of a gap in the EMV sequence, was confirmed from the sequence of peptides T7 and T8.

N-Terminal and C-Terminal Sequence of BDMV Coat Protein: Peptides CH1 and T1 were shown to be blocked to sequencing when 5 nm of each failed to yield any DABTH derivative when subjected to mannual sequence analysis. The amino acid compositions of CH1 and T1 was Asx 1, Ser 2, Glx 1, Val 3, Met 1, Lys 2 and Asx 1, Ser 2, Glx 1, Val 2, Met 1, Lys 1, respectively. Since the amino terminus of BDMV coat protein was blocked by an acetyl group (4), these peptides were presumed to come from the N-terminus of the protein. These peptides were subjected to CNBr cleavage and sequenced to yield: Asp-Ser-Ser-Glu-Val-Val-Lys-Val-Lys (CH1-CN) and Asp-Ser-Ser-Glu-Val-Val-Lys (T1-CN). These two sequences agreed with their respective amino acid compositions, except for the absence of one Met. The specificity of CNBr cleavage suggested that this Met was at the N-terminus of the peptides (Fig.1). Sequence of CH1-CN also overlapped with peptide CH2, confirming its placement at the N-terminus.

Carboxypeptidase Y digestion of BDMV coat protein was carried out in 6 M urea due to the tendency of the protein to precipitate in the digestion buffer. It was necessary to treat carboxypeptidase with pepstatin A as some endopeptidase activity was associated with carboxypeptidase Y. C-terminal analysis yielded the sequence Met-Leu-Ile-Ala-Asn (Fig.1 and 4).

Special Comments: The separation of peptides on TSK-G 2000 SW did not appear to be strictly determined by the size of the peptides at acidic pH (pH of 0.1% TFA is ~2.0). Among the peptides with similar molecular weights, those with the highest proportion of basic amino acids were eluted earlier. For example peptides CH2 (residues 9-21), CH9 (68-84) and CH21 (160-175) each of which contained two or three basic residues eluted (pool I of Fig.5) before those containing single or no basic residues such as CH4 (residus 22-36), CH12 (94-109) and CH13 (95-111) (pool II of Fig.5). The dependance of elution on charge has also been observed in the chromatography of proteins on TSK-G SW columns (34, 35).

Spherisorb 5 ODS column was best suited for the separation of peptides of varying sizes using TFA-CH_3CN solvent system and for the rechromatography of smaller fragments from chymotryptic digest using NH_4OAC-CH_3CN solvent system. However, this column was found to be unsuitable for rechromatography of large tryptic fragments using NH_4OAC-CH_3CN solvent system. Because of its strong hydrophobic bonded phase, higher concentrations of organic solvent were required to elute the larger fragments and the yields of peptides were poor. Therefore, tryptic fragments were rechromatographed on Aquapore RP-300 column using NH_4OAC-CH_3CN solvent system. This column gave excellent resolution of all the fragments in good yields and required low concentrations of organic solvent. The yields of peptides containing basic residues were in general poor from all the HPLC columns used.

Cleavage at some unusual sites were observed with chymotrypsin as well as with trypsin. Chymotrypsin cleaved at sites Val-Lys$^{\downarrow}$Val-Lys-Gln, Val-Lys-Val-Lys$^{\downarrow}$Gln and Gly-Arg$^{\downarrow}$Ile-Arg-Leu. Cleavage by chymotrypsin was not observed at Phe in the sequence Gly-Phe-Glu-Ala. $^{\downarrow}$Trypsin cleaved at Pro-Phe$^{\downarrow}$Gln-Phe, Val-Tyr$^{\downarrow}$Leu-Ala, Val-Tyr$^{\downarrow}$Gly and Ala-Tyr$^{\downarrow}$Ser-Arg sequences.

In the partial sequence reported earlier (4), Met at 186 was misidentified as Gln because DABTH-homoserine overlapped with DABTH-Gln in two dimensional chromatography (24). This particular tryptic peptide was obtained in very low amounts and its amino acid composition was not determined earlier. Due to its low amounts Lys at 184 had also been misidentified as Ser, a carry-over from the previous cycle.

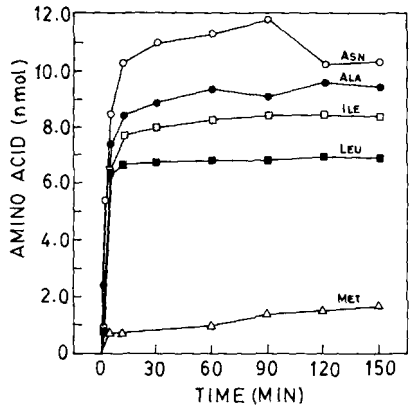


Figure 4. Carboxypeptidase Y digestion of reduced and carboxymethylated BDMV coat protein. Protein was digested with CpY as described in "Materials and Methods".

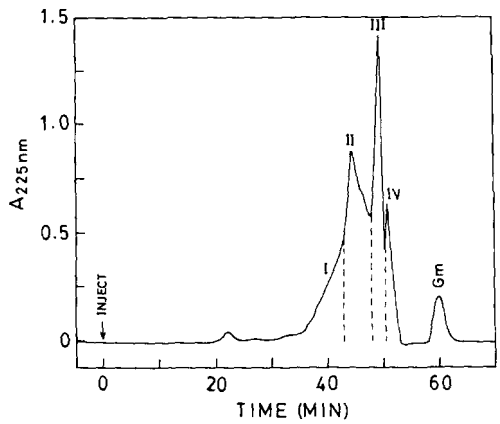
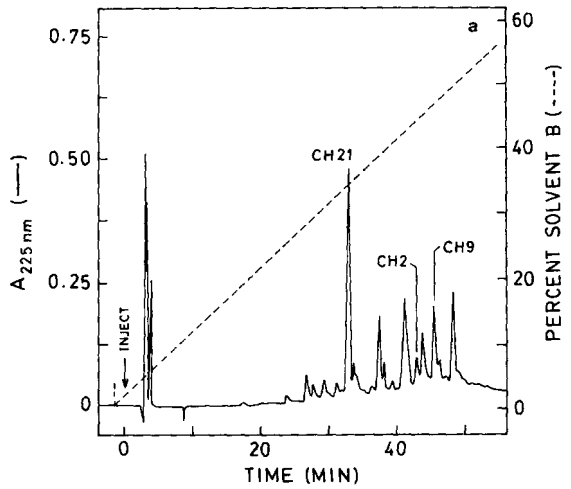
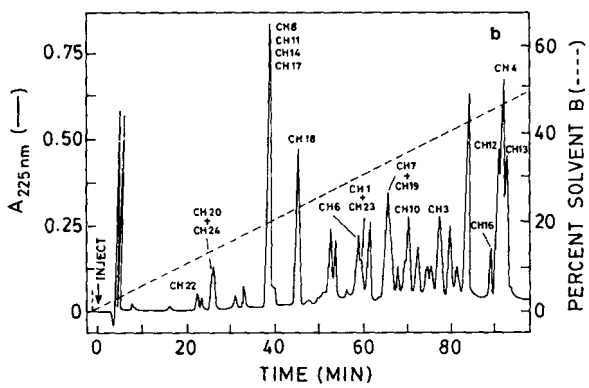


Figure 5. Initial separation of chymotryptic digest of BDMV coat protein by size exclusion HPLC. The fully reduced and carboxymethylated protein was digested with chymotrypsin as described in "Materials and Methods" and chromatographed on a TSK-G 2000 SW column using 0.1 % TFA containing 35% methanol as the mobile phase at a flow rate of 0.5 ml/min. Peptides were pooled as indicated by vertical dotted lines.





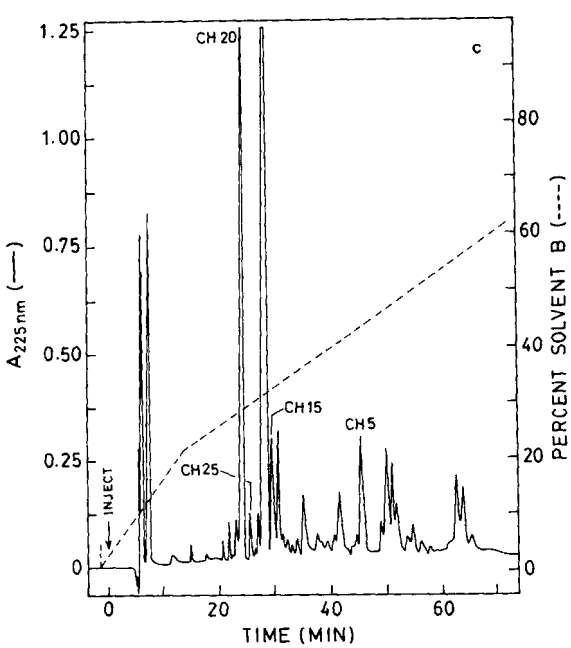


Figure 6. RP-HPLC of chymotryptic peptides in pools I-III from size exclusion HPLC. The pools indicated in Figure 5 were dried and chromatographed on a Lichrosorb RP-18 column using the TFA-CH_3CN solvent system described in "Materials and Methods" at a flow rate of 0.5 ml/min.

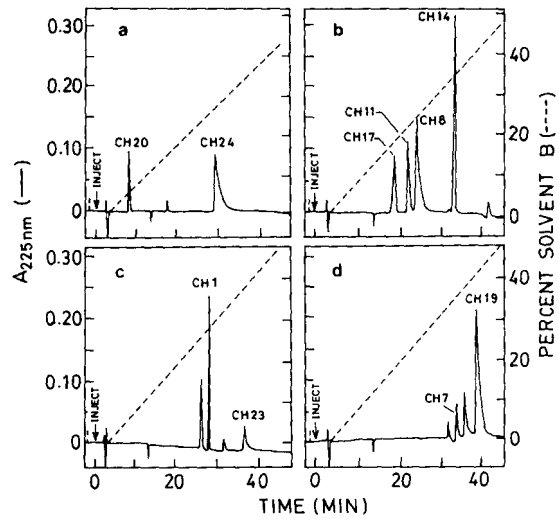


Figure 7. Repurification of chymotryptic peptides by RP-HPLC. Peaks from Figure 6 were rechromatographed on a Spherisorb 5 ODS column using NH_4OAc-CH_3CN solvent system described in "Materials and Methods" at a flow rate of 1 ml/min.

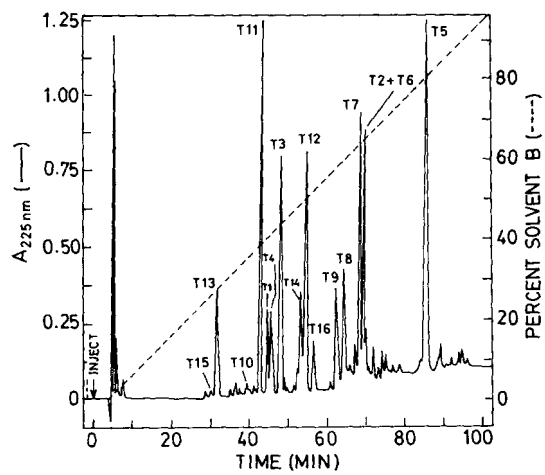


Figure 8. Initial separation of tryptic digest of BDMV coat protein by RP-HPLC. The digest was fractionated on Spherisorb 5 ODS column using TFA-CH_3CN solvent system at a flow rate of 0.7 ml/min.

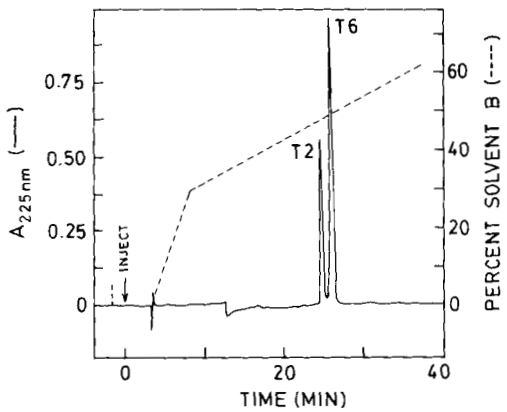


Figure 9. Rechromatography of a peptide mixture from Figure 8. Peptide mixture eluting at 69.06 min was fractionated on a Aquapore RP-300 column using NH_4OAc-CH_3CN solvent system at a flow rate of 1 ml/min.

Table 1. Amino Acid Composition of BDMV Coat Protein

Amino acid	Residues derived from Analysis	Sequence
Cys	1.0[a]	1
Asx	13.3	13
Thr	19.2	19
Ser	17.7	18
Glx	16.4	16
Pro	18.1	18
Gly	6.9	7
Ala	22.8	22
Val	13.3[b]	14
Met	3.1	3
Ile	13.0[b]	14
Leu	18.1[b]	18
Tyr	5.1	5
Phe	4.4	4
Lys	10.0	10
His	1.1	1
Arg	6.1	6
Trp	0.9[c]	1
Total		190

[a]Determined as Cm-cysteine
[b]Values at 72h hydrolysis
[c]Extrapolated to time 0

Table 2. Amino Acid Composition of Chymotryptic Peptides of BDMV Coat Protein[a]

Amino Acid	CH1	CH2	CH3	CH4	CH5	CH6	CH7	CH8
Cys								
Asx	1.0(1)			1.0(1)			0.9(1)	
Thr				1.0(1)	2.1(2)	2.0(2)	1.8(2)	0.9(1)
Ser	1.8(2)	1.8(2)	1.9(2)	1.8(2)		1.0(1)		
Glx	1.0(1)	1.2(1)	1.0(1)	3.0(3)	2.0(2)	2.0(2)	1.0(1)	
Pro		1.8(2)	1.9(2)	3.2(3)			0.9(1)	1.0(1)
Gly		1.0(1)	1.0(1)			1.0(1)		
Ala		2.0(2)	2.0(2)	1.0(1)	1.1(1)	3.0(3)	1.0(1)	1.0(1)
Val	3.1(3)	1.1(1)		0.7(1)		1.2(1)		
Met	0.9(1)							
Ile		1.9(2)	2.1(2)	0.7(1)			1.1(1)	
Leu		1.2(1)	1.2(1)	1.1(1)		1.2(1)	1.2(1)	
Tyr								1.0(1)
Phe				1.1(1)	2.1(2)			
Lys	1.8(2)	1.0(1)					0.9(1)	
His								
Arg								
Trp								
Total	10	13	11	15	7	11	9	4
Residues	1-10	9-21	11-21	22-36	37-43	44-54	55-63	64-67

Table 2. Contd...

Amino Acid	CH9	CH10	CH11	CH12	CH13	CH14	CH15	CH16
Cys	1.0[b](1)							
Asx	1.0(1)	2.0(2)		1.1(1)	1.1(1)			
Thr	1.9(2)	1.8(2)	0.9(1)	1.8(2)	1.7(2)			0.8(1)
Ser		0.9(1)		1.0(1)	0.9(1)		0.9(1)	1.4(1)
Glx	2.0(2)			1.0(1)	1.0(1)		1.0(1)	1.2(1)
Pro	0.9(1)	2.1(2)		3.1(3)	2.8(3)			1.1(1)
Gly							1.7(2)	1.6(2)
Ala	2.0(2)	0.9(1)		3.0(3)	3.0(3)			2.7(3)
Val	0.8(1)	1.1(1)	1.1(1)	1.0(1)	1.1(1)	1.0(1)		2.0(2)
Met								
Ile	1.8(2)			0.8(1)	0.9(1)			2.0(2)
Leu	2.1(2)	2.3(2)		2.0(2)	2.1(2)			2.1(2)
Tyr			1.0(1)			1.0(1)		
Phe							1.0(1)	
Lys	0.9(1)				1.1(1)			0.9(1)
His	0.9(1)							
Arg	0.9(1)					0.8(1)		
Trp				0.9[c](1)	1.1[c](1)			
Total	17	11	3	16	17	3	5	16
Residues	68-84	80-90	91-93	94-109	95-111	112-114	115-119	120-135

Table 2 contd...

Amino Acid	CH17	CH18	CH19	CH20	CH21	CH22	CH23	CH24	CH25
Cys									
Asx	3.0(3)	1.0(1)	1.0(1)		1.2(1)				1.0(1)
Thr		0.9(1)	1.8(2)	1.0(1)	1.8(2)				
Ser	0.8(1)	0.9(1)			2.6(3)	0.9(1)		1.0(1)	
Glx					1.0(1)				
Pro			1.0(1)		2.9(3)			0.9(1)	
Gly						1.0(1)			
Ala				1.0(1)	2.0(2)				1.0(1)
Val	1.2(1)	1.1(1)							
Met								1.0(1)	
Ile					1.9(2)	1.0(1)	1.0(1)		0.8(1)
Leu	1.2(1)		2.4(2)				1.2(1)		1.1(1)
Tyr		1.0(1)							
Phe									
Lys		0.9(1)	0.9(1)		0.9(1)			1.0(1)	
His									
Arg					0.9(1)	0.9(1)	0.9(1)		
Trp									
Total	6	6	7	2	16	4	3	4	4
Residues	136-41	145-50	151-57	158-59	160-75	176-79	180-82	183-86	187-90

[a]The numbers in parentheses indicate the composition predicted from the amino acid sequence
[b]Determined as Cm-cysteine
[c]Determined from the characteristic absorbance spectrum of tryptophan

Table 3. Amino Acid Composition of Tryptic Peptides of BDMV Coat Protein[a]

Amino Acid	T1	T2	T3	T4	T5	T6	T7	T8
Cys				1.0[b](1)				
Asx	1.1(1)	1.2(1)			2.8(3)	3.1(3)	1.0(1)	
Thr		1.0(1)	0.9(1)		5.0(5)	4.5(5)	1.6(2)	
Ser	2.1(2)	3.9(4)			2.1(2)	2.1(2)	1.1(1)	1.9(2)
Glx	1.1(1)	3.6(4)		2.1(2)	1.1(1)	0.7(1)	1.0(1)	1.1(1)
Pro		5.1(5)	0.9(1)		5.0(5)	5.2(5)	2.9(3)	
Gly		1.0(1)						3.9(4)
Ala		2.8(3)	1.1(1)	1.0(1)	5.0(5)	4.9(5)	3.0(3)	3.2(3)
Val	1.6(2)	0.8(1)		0.9(1)	3.2(3)	4.3(4)	1.0(1)	2.0(2)
Met	1.0(1)							
Ile		2.5(3)		1.1(1)	2.2(2)	2.0(2)	1.3(1)	1.2(1)
Leu		2.2(2)	1.3(1)		5.8(6)	5.7(6)	2.7(3)	1.3(1)
Tyr			1.0(1)		1.1(1)	1.6(2)		1.0(1)
Phe		1.0(1)						1.0(1)
Lys	1.0(1)			1.1(1)	0.6(1)	0.8(1)	1.0(1)	1.0(1)
His				1.0(1)				
Arg			1.3(1)		1.0(1)	1.1(1)	1.0(1)	
Trp					N.D.(1)	N.D.(1)	0.9[c](1)	
Total	8	26	6	8	36	38	19	17
Residues	1-8	11-36	63-68	69-76	77-112	77-114	94-112	113-29

Table 3 contd...

Amino Acid	T9	T10	T11	T12	T13	T14	T15	T16
Cys								
Asx	3.1(3)		1.8(2)		1.0(1)			1.0(1)
Thr	0.8(1)		2.8(3)		1.0(1)	1.0(1)		
Ser	0.7(1)		1.1(1)		1.0(1)	1.8(2)		1.1(1)
Glx	0.8(1)					1.1(1)		
Pro	1.1(1)		1.1(1)		2.0(2)	0.9(1)		0.9(1)
Gly						1.0(1)		
Ala				1.0(1)	1.0(1)	1.0(1)		1.0(1)
Val	1.8(2)		1.3(1)					
Met		1.0(1)						1.0(1)
Ile	0.8(1)					3.0(3)	1.0(1)	1.0(1)
Leu	2.0(2)	1.0(1)		2.2(2)				1.9(2)
Tyr			1.0(1)	1.0(1)				
Phe								
Lys		1.1(1)			1.0(1)			1.0(1)
His			0.9(1)					
Arg	0.9(1)					1.1(1)	0.9(1)	
Trp								
Total	13	3	10	4	7	11	2	9
Residues	130-42	143-45	146-55	156-59	162-68	169-79	180-81	182-90

[a]The numbers in parentheses indicate the composition predicted from the amino acid sequence
[b]Determined as Cm-cysteine
[c]Determined from the characteristic absorbance spectrum of tryptophan
N.D., Not determined