Virology 253, 319–326 (1999)
Article ID viro.1998.9519, available online at http://www.idealibrary.com on IDEAL®

Rice Tungro Bacilliform Virus Open Reading Frame 3 Encodes a Single 37-kDa Coat Protein

Philippe Marmey,* Brian Bothner,† Emmanuel Jacquot,‡ Alexandre de Kochko,* Ching Ang Ong,§ Pierre Yot,‡ Gary Siuzdak,† Roger N. Beachy,* and Claude M. Fauquet*,[1]

*ILTAB (ORSTOM/TSRI), Plant Division, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, California 92037; †Department of Chemistry, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, California 92037; ‡IBMP, CNRS, Université Louis Pasteur, 67084 Strasbourg Cedex, France; and §Department of Virology, Malaysian Agricultural Research and Development Institute, P.O. Box 12301, Kuala Lumpur, Malaysia

Received September 16, 1998; returned to author for revision October 7, 1998; accepted November 11, 1998

Rice tungro bacilliform virus (RTBV) is a plant pararetrovirus and a member of the Caulimoviridae family and closely related to viruses in the *Badnavirus* genus. The coat protein of RTBV is part of the large polyprotein encoded by open reading frame 3 (ORF3). ORF3 of an RTBV isolate from Malaysia was sequenced (accession no. AF076470) and compared with published sequences for the region that encodes the coat protein or proteins. Molecular mass of virion proteins was determined by mass spectrometry (matrix-assisted laser desorption/ionization–TOF) performed on purified virus particles from three RTBV isolates from Malaysia. The N- and C-terminal amino acid sequences of the coat protein were deduced from the mass spectral analysis, leading to the conclusion that purified virions contain a single coat protein of 37 kDa. The location of the coat protein domain in ORF3 was reinforced as a result of immunodetection reactions using antibodies raised against six different segments of ORF3 using Western immunoblots after SDS–PAGE and isoelectrofocusing of proteins purified from RTBV particles. These studies demonstrate that RTBV coat protein is released from the polyprotein as a single coat protein of 37 kDa.

INTRODUCTION

Rice tungro is a major rice disease in southeast Asia, with annual crop losses estimated at ~$680 million (Herdt, 1991). The disease is caused by a complex of two viruses (Hibino *et al.*, 1978; Jones *et al.*, 1991): rice tungro spherical virus (RTSV), a single-stranded RNA virus and member of the *Waikavirus* genus (Shen *et al.*, 1993), and rice tungro bacilliform virus (RTBV), a plant pararetrovirus (Hay *et al.*, 1991; Qu *et al.*, 1991) and member of the Caulimoviridae family that is closely related to the *Badnavirus* genus (Pringle, 1997). RTSV is required for the transmission of the two viruses by the leafhopper vector (*Nephotettix virescens*) (Hibino, 1983a, 1983b), and RTBV is responsible for symptoms of the disease (Hibino *et al.*, 1978).

The nonenveloped bacilliform (130 × 30 nm) RTBV particles contain a circular and relaxed, double-stranded DNA of 8.0 kbp with two site-specific discontinuities (Bao and Hull, 1994) resulting from replication by reverse transcription (Bao and Hull, 1994; Laco and Beachy, 1994). The RTBV genome has four open reading frames (ORFs) (for a review, see Hull, 1996). ORF3 encodes a polyprotein (P3; 194 kDa) possessing the analogs of the gag, pro, and pol products of retroviruses (for a review, see Rothnie *et al.*, 1994). Comparative analysis of sequences of related badnaviruses (Bouhida *et al.*, 1993; Hagen *et al.*, 1993; Hay *et al.*, 1991; Qu *et al.*, 1991) revealed that P3 contains a putative movement protein (MP), the coat protein (CP), the aspartyl protease (PR), the reverse transcriptase (RT), and a ribonuclease H activity. The amino acid (aa) sequence DSGS located within the aspartyl protease is found in the RTBV P3 polyprotein between CP and RT domains. The change of D to A in this sequence affects the proteolytic processing of the RT (Laco *et al.*, 1995).

Previous reports indicated that antibodies raised against RTBV used in Western immunoblots detected peptides with molecular masses of 150, 62, and/or 37 and 33 kDa from purified virus (Hay *et al.*, 1991, 1994; Qu *et al.*, 1991). Qu *et al.* (1991) found the 37-kDa peptide in both infected tissue and purified virus, whereas the 33-kDa peptide was present only in purified virus. Qu *et al.* (1991) determined the N-terminus of the 33-kDa peptide; however, the 37-kDa peptide was blocked for sequencing. These authors suggested that 37- and 33-kDa peptides may share the same C-terminus, with the 33-kDa peptide being a degradation product of the 37-kDa protein. Hull (1996) suggested the possibility that there are two CPs of about the same molecular mass; this "tandem CP" model does not fit with the hypothesis presented by Qu *et al.* (1991) (Fig. 1). In the single CP model, it is likely that P3 would contain intervening regions of undeter-

[1] To whom reprint requests should be addressed. Fax: (619) 784-2994. E-mail: iltab@scripps.edu.

0042-6822/99 $30.00

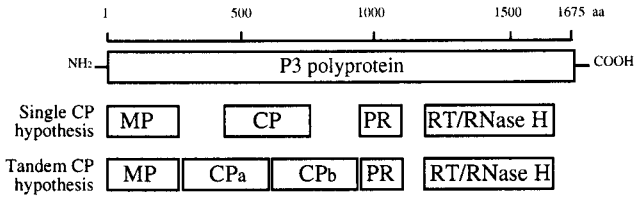

FIG. 1. Schematic representation of polyprotein P3 for the two hypotheses related to the CP of RTBV.

mined function between MP and CP and between CP and PR, as it does between PR and RT (Laco *et al.*, 1995).

The development of electrospray and matrix-assisted laser desorption/ionization (MALDI) techniques has made the accurate molecular mass determination of proteins routine (Siuzdak, 1996). More recently, MALDI mass spectrometry has been applied to the analysis of viral capsid proteins (Bothner *et al.*, 1998; Lewis *et al.*, 1998; Siuzdak *et al.*, 1996).

The RTBV genome sequenced by Qu *et al.* (1991) was an isolate from the Philippines [International Rice Research Institute (IRRI)]. In this study, we used three Malaysian isolates of RTBV: one named Serdang from MARDI (Malaysian Agriculture Research Development Institute, Malaysia) and the other two from the central region (Kuala) and the south of Malaysia. These isolates have identical characteristics with regard to the 37- and 33-kDa peptides as the isolate from IRRI (Qu *et al.*, 1991). We describe the analysis of RTBV particles by mass spectrometry and two-dimensional gel electrophoresis to investigate the hypothesis of a single versus two 37-kDa CP species in RTBV virions.

RESULTS

Comparative analysis of sequences in the P3 polyprotein between Serdang and IRRI isolates

Because the Serdang and IRRI isolates of RTBV are from two different countries, some differences were expected at the nucleotide (nt) and aa levels. We isolated and sequenced regions of the genome that contained the coat protein; the correct aa sequence then would be deduced and later used to interpret the mass spectrometry analysis.

DNA from tissues infected with the Serdang isolate was amplified using RTBV-specific primers to sequence its genome (accession no. AF076470). We then compared sequences for the CP, PR, and intervening region (IR) domains of the genome. Putative domains considered were the CP 37-kDa domain suggested by Qu *et al.* (1991) and the 13.5-kDa PR domain suggested by Hay *et al.* (1994). The percentages of nt identity between the Serdang and IRRI isolates were 93.6%, 90.3%, and 93.5% for the CP, IR, and PR domains, respectively. When comparing derived aa sequences, sequence homologies were 99.0%, 91.3%, and 99.1% for CP, IR, and PR regions, respectively. In the CP and PR regions of the Serdang isolate, only four substitutions were identified relative to the IRRI isolate: the predicted replacement of S519 to G, V668 to I, V669 to I, and N1007 to S. In the IR region of the Serdang isolate, 14 differences in aa were found.

Determination of molecular masses

MALDI mass spectrometry is accomplished by directing a pulsed laser beam onto a sample cocrystallized in a matrix (for a review, see Siuzdak, 1996). The matrix absorbs the laser energy and is vaporized; proteins released are then analyzed. The mass spectrometer determines the molecular masses of compounds by separating ions according to their mass-to-charge ratio (*m/z*). The primary advantages of the MALDI technique include high sensitivity, soft ionization, and tolerance to salt in most samples.

Analyses were conducted on three virus preparations from Malaysia. A volume of 0.5 μl of virus preparation (optical density of 3.0 at 260 nm) was mixed with 0.5 μl of the sinapinic acid matrix on the plate and left for 10 min for the matrix to crystallize before analysis. Three peaks were detected for each RTBV isolate tested. There was no signal at 37,000 *m/z* in any of the samples tested, and only one sample showed a signal in the 34000 *m/z* range. The Serdang isolate presented signals at 15,210, 22,117, and 34,105 *m/z*, whereas peaks obtained for the two other Malaysian isolates were, successively, 11,983, 15,197, and 22,099 and 11,922, 15,182, and 22,111 (Table 1 and Fig. 2). The sample that yielded a signal at 34,105 *m/z* did not produce a signal in the 12,000 *m/z* range, and we concluded that the signal expected at 37,000 *m/z* was

TABLE 1

Results of Mass Spectral Analyses Performed on Three Malaysian Isolates of RTBV

Virus sample	Mass (*m/z*)			
Serdang isolate		15,210 + 0.32%	22,117 − 0.19%	34,105 + 0.15%
Kuala isolate	11,983 + 0.60%	15,197 + 0.24%	22,099 − 0.27%	
Southern isolate	11,922 + 0.09%	15,182 + 0.14%	22,111 −0.22%	

Note. The mass spectrometer determines the molecular masses of compounds by separating ions according to their mass-to-charge ratio (*m/z*). The percentage difference between measured molecular mass and that predicted from the genome sequence is indicated.

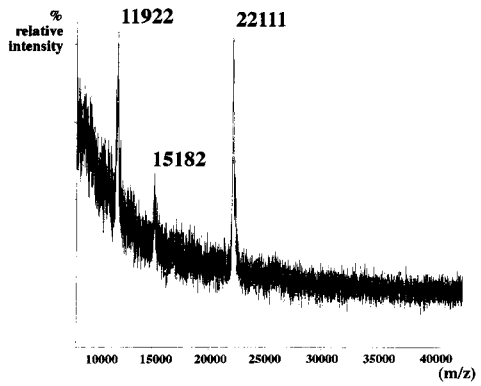

FIG. 2. Mass spectrum of RTBV particles from the Southern Malaysian isolate. Three signals are observed at *m/z* 11,922, 15,182, and 22,111.

split into signals at ~22,000 and ~15,000 *m/z*. For two isolates, the signal at 34,000 *m/z* was split into signals at 22,000 (similar to the one described above) and at 12,000 *m/z*. It is likely that the signal at 34,105 *m/z* represents the 33-kDa protein reported in the literature (molecular mass estimated by SDS–PAGE).

To test the possibility that RTBV virus particles were sensitive to the conditions used for mass spectrometry, we determined the stability of the CP in presence of the matrix used for the analysis. An aliquot of virus was mixed with the same amount of sinapinic acid (as above) in an open Eppendorf tube for 10 min before adding Laemmli's sample buffer and boiling before SDS–PAGE. A Western immunoblot was performed using the antibody raised against the virus (Ab-RTBV). When incubated in the presence of sinapinic acid, the 37- and 34-kDa peptides, which are found in mature virions, were unstable and yielded products with apparent molecular masses of 12, 15, and 22 kDa (Fig. 3). Two other matrix compounds that are often used to form matrices, α-cyano-4-hydroxy cinnamic acid and 2,5-dihydrobenzoic acid, cleaved the two peptides in the same manner (data not shown).

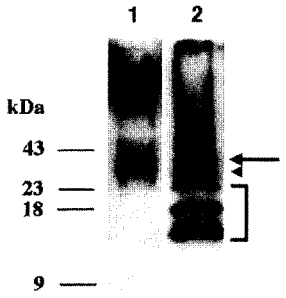

FIG. 3. Immunodetection of RTBV CP-related proteins. Immunogenic proteins were revealed with antibodies raised against RTBV particles. Lane 1 corresponds to purified virions alone; lane 2, virions treated with sinapinic acid. Arrow and arrowhead indicate 37- and 34-kDa peptides, and bracket denotes cleavage products as peptides of 22, 15, and 12 kDa. At the left, location of molecular weight standards.

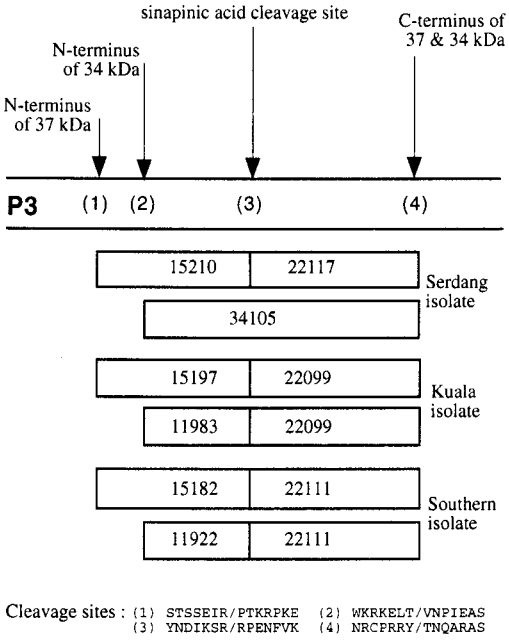

FIG. 4. Schematic representation of the cleavage sites according to the mass spectral analyses. Arrows indicate positions of cleavage, and numbers in brackets indicate the corresponding sequence. Squares represent peptides for each isolate; numbers in squares are *m/z* signals read on mass spectra.

The peptides produced by mass spectrometry were most likely derived from the CP. By combining the molecular masses of peptide fragments (i.e., 12 kDa plus 22 kDa, and 15 kDa plus 22 kDa), we obtained the expected molecular masses of the 34- and 37-kDa proteins, respectively (Fig. 4). We determined the sinapinic acid-induced cleavage site to be between aa 603 and 604 by combining the protein molecular mass of 11,922 or 11,983 Da with the N-terminus of the 34-kDa protein previously shown to be at aa 502 (Qu *et al.*, 1991). The C-terminus of the 34-kDa protein was determined by combining either the protein molecular mass of 22,099, 22,111, or 22,117 Da with the sinapinic acid cleavage site or the protein molecular mass of 34,105 Da with the known N-terminus sequence of the 34-kDa protein; these led us to conclude that the two peptides share the same C-terminus, which was determined to be at aa 791. The best fit for the 37-kDa protein indicates that it has the same C-terminal aa as the 34-kDa protein because of the presence of the single peak for each sample of ~22,000 *m/z*. The C-terminal aa sequence of the 15-kDa peptide was calculated to be the sinapinic acid-induced cleavage site (aa 603) (Fig. 4). The N-terminus of the 15-kDa peptide would then be the N-terminus of the 37-kDa peptide and is determined at aa 477.

Localization of the coat protein domain in the P3 polyprotein

To confirm the CP domain within P3, we used a combination of different antibodies and Western blot assays.

TABLE 2

Oligonucleotides Used as Primers to Generate DNA Fragments of ORF3

Designation	Orientation	Primer sequence[a]	Region of ORF3 corresponding to PCR product[b]
MP1	Forward	5′-GTAAGTGCCcatatgAGCCTTAGACCATTTACTGG-3′ *Nde*I	1–210
	Reverse	5′-AGGGCTGTGggatccTCAGGTGGTCAATCTTCCTA-3′ *Bam*HI	
MP2	Forward	5′-GTAAGTGCCcatatgGCAATCAAAACTATAGGAAG-3′ *Nde*I	201–445
	Reverse	5′-AGGGCTGTGggatccTCAGAGGGCCATATCATAGT-3′ *Bam*HI	
CP1	Forward	5′-GTAAGTGCCcatatgGAAGAAATAACTGAATGGGA-3′ *Nde*I	401–606
	Reverse	5′-AGGGCTGTGggatccTCATTCAGGTCTATCACCTC-3′ *Bam*HI	
CP2	Forward	5′-GTAAGTGCCcatatgAGAGATGATTGGAAAAGAAA-3′ *Nde*I	492–606
	Reverse	5′-AGGGCTGTGggatccTCATTCAGGTCTATCACCTC-3′ *Bam*HI	
MP1-CP	Forward	5′-GTAAGTGCCcatatgAGCCTTAGACCATTTACTGG-3′ *Nde*I	1–606
	Reverse	5′-AGGGCTGTGggatccTCATTCAGGTCTATCACCTC-3′ *Bam*HI	
MP2-CP	Forward	5′-GTAAGTGCCcatatgGCAATCAAAACTATAGGAAG-3′ *Nde*I	201–606
	Reverse	5′-AGGGCTGTGggatccTCATTCAGGTCTATCACCTC-3′ *Bam*HI	
CP3	Forward	5′-ATCCCATAggatccTCAAAAGAT-3′ *Bam*HI	506–775
	Reverse	5′-GGTGATTTTAAGCTTGTCAGATAT-3′ *Hin*dIII	
IR	Forward	5′-CTAATAGATggatccGACGAAGATATGGTATC-3′ *Bam*HI	806–961
	Reverse	5′-CATCCTGaagcttTTTAAATTGATG-3′ *Hin*dIII	

[a] Lowercase letters indicate restriction enzyme sites.
[b] Amino acid coordinates according to Qu *et al.* (1991).

We first expressed different parts of ORF3 in *Escherichia coli* and determined whether the polypeptides were recognized by the antiserum raised against virions (Ab-RTBV). We then determined whether virus particles were recognized by different antisera raised against specific regions of ORF3. Finally, we used the antisera to identify virion-related proteins within infected plants.

Specific primers were used to amplify by PCR different regions of ORF3 as described in Table 2. PCR fragments were cloned into either pET28 or pTrcHisA vectors, and gene expression was induced in *E. coli* to produce specific fragments of P3. After SDS–PAGE and electrotransfer onto nitrocellulose membranes, peptides were reacted with Ab-RTBV. Peptides that included aa 401–775 were detected by antibodies raised against RTBV, whereas fragments that compose aa 1–445 and 806–961 did not share common epitopes with the purified virions (Fig. 5).

The second approach involved using different antibodies raised against specific peptides derived from ORF3. Some of the peptides described in Table 2 were purified and injected to rabbits to induce antibodies. Antisera Ab-MP1, Ab-MP2, Ab-CP1, Ab-CP2, and Ab-CP3 were raised against peptides MP1, MP2, CP1, CP2, and CP3, respectively (Table 2). All antibodies were tested by reaction with the inducing peptide and shown to be positive. RTBV particles were disrupted and subjected to SDS–PAGE and transferred onto membranes and subsequently reacted with the different antibodies. Ab-MP1 and Ab-MP2 did not react with the CP, whereas bands at 37 and 34 kDa were detected using Ab-CP1, Ab-CP2, and Ab-CP3 (Fig. 6 and Table 3). Similar results were obtained with extracts from infected plants where Ab-CP1, Ab-CP2, and Ab-CP3 revealed the expected 37-kDa band (Table 3).

The results obtained in these analyses lead to the conclusion that sequences upstream of aa 401 and downstream of aa 806 do not contain CP sequences.

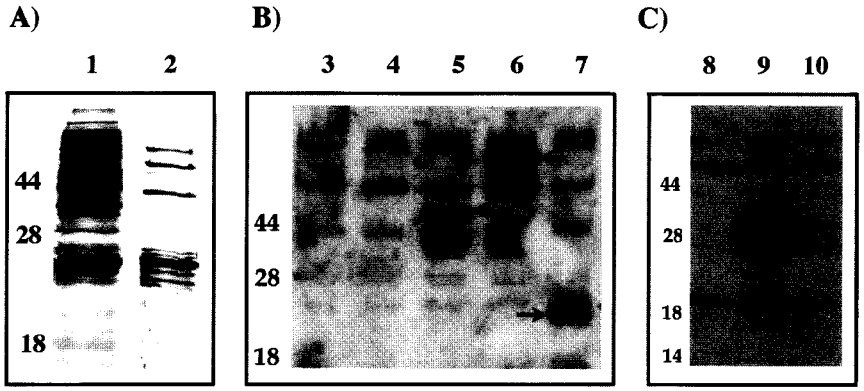

FIG. 5. Western blots analyses performed on induced peptides from *E. coli*. (A) Coomassie blue-stained gel and corresponding immunoblot using antibodies raised against peptide MP1. (B and C) Immunoblots using antibodies raised against RTBV particles of peptides expressed in pET- and pTrcHis-based vectors, respectively. Lanes 1–3 correspond to peptides expressed from plasmid pMP1; lane 4, peptides expressed from pMP2; lane 5, peptides expressed from pMP2CP; lane 6, peptides expressed from pMP1CP; lane 7, peptides expressed from pCP1; lane 8, peptides expressed from pTrcHis with no insert; lane 9, peptides expressed from pCP3; and lane 10, peptides expressed from pIR. (Left) Molecular weight markers. Arrows indicate induced peptides immunogenic to antibodies raised against RTBV particles. Only peptides MP2-CP, MP1-CP, CP1, and CP3 were detected by antibodies raised against RTBV.

This suggests that there is a single 37-kDa coat protein in P3, as proposed in Fig. 1.

Electrophoretic determination of the number of 37-kDa CP domains

We considered the possibility that virions contain one or two CP species of 37 kDa that were not separated by SDS–PAGE. The hypothesis of two CP molecules encoded in tandem in ORF3 (see Fig. 1) would lead to two peptides of very different isolectric points (pI); based on the predicted sequences of ORF3, the proteins would have pIs of ~4.6 and ~8.6. Sequence analysis of the putative single CP domain predicted a high percentage of basic aa and a pI of 9.4 for both 34- and 37-kDa domains. We used nonequilibrium pH gradient electrophoresis (NEPHGE) (O'Farrel *et al.*, 1977) for these analyses. Western blot assays using Ab-RTBV after NEPHGE clearly showed major spots at ~37 and ~34 kDa and a variety of other proteins near the basic region of the gel (Fig. 7). These results led to the conclusion that RTBV contains only CP molecules with pIs as predicted for a single CP (37 kDa) that is cleaved, producing a 34-kDa protein.

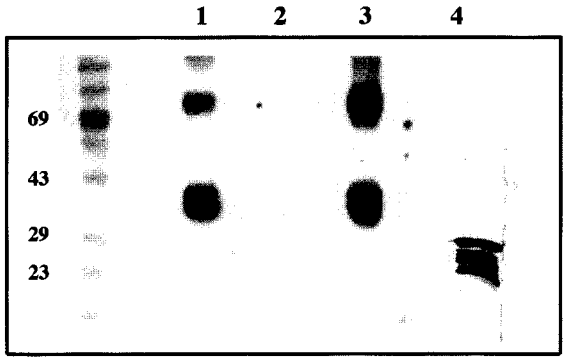

FIG. 6. Western blot analyses on purified virus using different antibodies. Lane 1 corresponds to the immunoreaction using antibodies raised against RTBV; lane 2, the immunoreaction using Ab-MP2; and lane 3, the immunoreaction using Ab-CP1. Lane 4 shows induced peptides from pMP2, detected using Ab-MP2. The 37-and 34-kDa peptides are immunogenic to Ab-RTBV and Ab-CP1 but not to Ab-MP2. Upper bands correspond to dimers. (Left) Molecular weight markers.

DISCUSSION

The nature of RTBV CP has been the subject of research in the past few years, with reports that suggested that the virus contains a variety of CP molecules of different size. However, no single work reported the full characterization of the CP. In this study, we identified the N- and C-terminal sequences of the CP using MALDI mass spectrometry on virus particles and used two different approaches to validate the hypothesis of a single 37-kDa CP.

Virus preparations used in our analysis were from Malaysia. Mass spectrometry analysis is highly accurate and can determine molecular masses with a good resolution when based on the correct aa sequence. The genome of the Serdang isolate was sequenced for the putative CP domain, and its sequence was compared with the IRRI isolate reported previously (Qu *et al.*, 1991).

Two RTBV isolates from IRRI have been described as having genomes of 8000 and 8002 nucleotides (Hay *et al.*, 1991; Qu *et al.*, 1991); the predicted aa sequences of ORF3 of these isolates were identical. Fan *et al.* (1996) looked at the variation in the genome of six isolates from Bangladesh, India, Indonesia, Malaysia, Thailand, and Philippines. Using restriction endonuclease maps and cross-hybridization, they concluded that the isolates fell into two groups: those from the Indian subcontinent and those from south-east Asian countries. Malaysian iso-

TABLE 3

Summary of the Western Blot Analysis Using Different Antibodies

Protein	Antibody					
	Ab-MP1	Ab-MP2	Ab-CP1	Ab-CP2	Ab-CP3	Ab-RTBV
Purified virus						
37 kDa	No	No	Yes	Yes	Yes	Yes
34 kDa	No	No	Yes	Yes	Yes	Yes
Infected plants						
37 kDa	No	No	Yes	Yes	Yes	Yes
34 kDa	No	No	No	No	No	No

Note. Indicated is the recognition (yes or no) of coat protein from purified Serdang isolate virus and from plants infected with Serdang isolate using antibodies raised against known sequences from RTBV-ORF3 (as described in Table 2) or against RTBV particles (Ab-RTBV).

lates so far studied show similar restriction site patterns; predicted protein sequences seem to be highly conserved between isolates from Malaysia.

RTBV particles are relatively unstable, and attempts to dialyze virions against water resulted in their disruption (unpublished observation). Similarly, treatment of virions with sinapinic acid before exposure to MALDI-TOF cleaved the 37- and 34-kDa proteins at a single site. The 37-kDa protein from all three isolates was split into two peptides of ~15 and ~22 kDa. The 34-kDa protein from the Serdang isolate was not cleaved, whereas the peptide from the other Malaysian isolates was cleaved to produce a 12-kDa peptide and a 22-kDa peptide. Because a single peak was observed for the 22-kDa peptide, we concluded that the sinapinic acid cleavage site was at the same location in the 34- and 37-kDa proteins and that these two proteins shared the same C-terminal aa sequence. From these data, we concluded that the N-terminus of the 37-kDa protein was at aa 477 and that the C-terminal was at aa 791.

The hypothesis of a single 37-kDa CP domain was verified by analysis using different antibodies. Antibodies directed against proteins between aa 1–210 and 201–445 did not recognize the RTBV CP. Furthermore, the IR domain that overlaps residues 806–961 was not immunoreactive with RTBV antibodies. The hypothesis of a single 37-kDa CP domain was further verified by NEPHGE analysis, with major spots at ~37 and ~34 kDa, each of which exhibited a high pI. In contrast, the two-CP hypothesis would have led to two peptides with a very different pI. Western immunoblots analyses indicate as expected that 34-kDa peptide is only present in purified virus; the 34-kDa peptide is likely a degradation product of the 37-kDa peptide that occurs during virus purification.

In this study, we demonstrate that the RTBV CP is released from the polyprotein as a single CP of 37 kDa. The CP identified has a molecular mass of 37,303 Da with an estimated pI of 9.43. It contains near its C-terminus (aa 776–789) the "zinc finger-like" motif CXCX2CX4HX4C (Hay *et al.*, 1991; Qu *et al.*, 1991). Other plant pararetroviruses share this consensus sequence as a characteristic of the RNA binding site of retroviruses (Covey, 1986; Fütterer and Hohn, 1987).

Processing of the polyprotein P3 (aa 1–1675) is believed to produce at least four proteins: MP, CP, PR, and RT. Laco *et al.* (1995) determined the N- and C-termini of

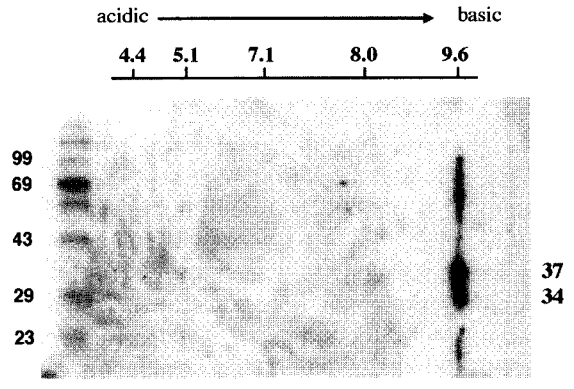

FIG. 7. Western blot analysis of viral particles after NEPHGE and immunodetection with RTBV antibody. Arrow indicates the direction of the isoelectric focusing; scale is measured pI gradient.

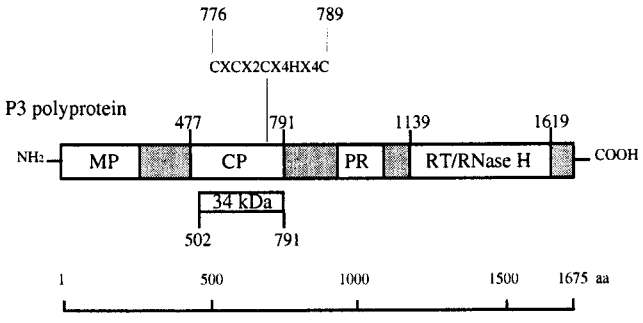

FIG. 8. Schematic organization of CP domain within P3 polyprotein. Positions of amino- and carboxyl-termini of the 37-kDa CP are 477 and 791, respectively. The 34-kDa peptide CP is positioned below the polyprotein with its amino- and carboxyl-termini corresponding to aa 502 and 791, respectively. The RT domain was characterized by Laco *et al.* (1995). The aa coordinates are according to Qu *et al.* (1991).

the RT to be aa 1139 and 1619, respectively. The MP was predicted, based on similarities among the badnaviruses, to lie within the first 300 aa of P3 (Bouhida *et al.,* 1993; Hagen *et al.,* 1993). Although the N- and C-termini of the PR are not yet determined, a protein of 13.5 kDa was identified in RTBV particles to contain the domain predicted to include the active site of the aspartyl protease (aa 987–990) (Hay *et al.,* 1994). As summarized in Fig. 8, the regions of P3 upstream and downstream of the CP (i.e., intervening regions) do not contain CP sequences and would be similar in molecular mass (~20 kDa) and charge (pI ≈ 4.5). To date, proteins of this size and charge have not been identified in RTBV particles or infected cells; a role, if any, of these sequences in virus replication is not known. One possibility is that these regions are in some way involved in the activity of the protease in releasing the 37-kDa CP.

MATERIALS AND METHODS

RTBV isolates used for the analysis

Three isolates of RTBV from Malaysia were used. They were purified at MARDI (Malaysian Agriculture Research Development Institute, Malaysia) and were named Serdang (from the MARDI area), Kuala (from the Kuala Lumpur area), and Southern (from south of Malaysia). Serdang isolate was sequenced (AF076470). The isolate used for sequences comparison is the RTBV isolate sequenced by Qu *et al.* (1991) (M65026) from the Philippines (IRRI).

Virus purification

Rice infected leaves (250 g) were frozen, ground, and thawed in 4 ml/g tissue of buffer A (0.1 M sodium citrate, pH 6.0, 10 mM EDTA). The mixture was digested with Cellulclast (Novo-Enzymes) at 5% (v/v) for 3 h at 40°C. The mixture was centrifuged at 15,000*g* for 15 min at 4°C. The supernatant was mixed for 1 h at room temperature with 7% (v/v) PEG 8000, 200 mM NaCl, and 1% (v/v) Triton X-100 before centrifugation at 22,000*g* for 20 min at 4°C. The pellet was resuspended overnight in 10 ml of buffer A. The sample then was centrifuged through a 10% sucrose cushion at 250,000*g* for 2.5 h at 4°C. The pellet was resuspended overnight in buffer A. After centrifugation at 8000*g* for 15 min at 4°C, the supernatant was centrifuged through a 10–50% sucrose density gradient at 250,000*g* for 2 h at 10°C. Opalescent bands were collected and diluted in buffer B (100 mM sodium citrate, pH 7.0), followed by a final centrifugation at 250,000*g* for 3 h at 4[degC. The final pellet was then resuspended in 1 mM phosphate buffer, pH 7.4.

MALDI mass spectrometry

Then, 0.5 μl of the virus sample, with an optical density of 3.0 at 260 nm, was loaded onto the sample plate and mixed with an equal volume of the matrix sinapinic acid. MALDI mass spectrometry experiments were performed on a PerSeptive Biosystems Voyager-Elite with delayed extraction. Samples were irradiated with a nitrogen laser (Laser Science Inc.) operated at 337 nm, and the laser beam was attenuated by a variable attenuator and focused on the sample target. Ions produced in the ion source were accelerated with a deflection voltage of 30,000 V. The ions were then differentiated according to their *m/z* using a time-of-flight mass analyzer.

Sequencing of viral DNA

Extraction of DNA from tissue infected with the Serdang isolate was performed according to Dellaporta *et al.* (1983). RTBV DNA was amplified by PCR using specific primers; amplified fragments were subjected to automatic sequencing (model 373A, Applied Biosystem).

Expression of proteins

Different primers were designed to amplify specific sequences from the RTBV genomes (Table 2). PCR fragments corresponding to domains MP1, MP2, CP1, CP2, MP1-CP, and MP2-CP were cloned into an expression vector pET28 (Novagen). PCR fragments corresponding to domains CP3 and IR were cloned into pTrcHisA (In-Vitrogen). The resulting plasmids were designated pMP1, pMP2, pCP1, pCP2, pCP3, pMP1CP, pMP2CP, and pIR and allowed synthesis of peptides corresponding to regions between residues 1–210, 201–445, 401–606, 492–606, 506–775, 1–606, 201–606, and 806–961, respectively. Peptides were expressed in *E. coli* strain BL21/DE3 (pLys S), except for pTrcHis-derived constructs, which were induced in DH5α. Expression was induced as previously described (Jacquot *et al.,* 1996).

Antibodies and Western blot analysis

Peptides expressed in bacteria were used to immunize rabbits and raise antisera as follows. Proteins from the bacterial lysate were separated by 18% SDS–PAGE and stained with Coomassie blue. Gel bands containing ~80 μg of peptides were excised and crushed in PBS buffer (v/v). The crushed gel was then emulsified with an equal volume of complete Freund's adjuvant and divided into two fractions before being subcutaneously injected into two rabbits. Injections of proteins emulsified with incomplete Freund's adjuvant was performed every 2 weeks, and rabbits were bled 1 week after each injection. Peptides MP1, MP2, CP1, CP2, and CP3 were used to produce Ab-MP1, Ab-MP2, Ab-CP1, Ab-CP2, and Ab-CP3, respectively. An antiserum (Ab-RTBV) was also raised against purified virions.

Proteins were subjected to electrophoresis in SDS–PAGE (14% polyacrylamide) (Laemmli, 1970) and transferred to a nitrocellulose membrane. Blots were incubated with the appropriate antiserum at a 1:5000 dilution

using the Biomax chemiluminescent detection system (Kodak).

NEPHGE

Viral particles were subjected to NEPHGE according to O'Farrel *et al.* (1977), using X-Cell II apparatus (Novex). Ampholines in the pH range of 3.0–10.0 were used, and running conditions were 100 V for 15 min, 200 V for 15 min, and 400 V for 35 min. The second dimension was run on a 14% SDS–polyacrylamide gel. Gels were then subjected to Western blotting.

ACKNOWLEDGMENTS

The authors thank Dr. P. Más for critical reading of the manuscript. P.M., A.K., and C.M.F. were supported by ORSTOM (French Scientific Research Institute for Development through Cooperation). R.N.B. was supported in part by The Scripps Family Chair. NATO is acknowledged for its support for a collaborative research grant between ILTAB and IBMP (Grant CRG 960760). Major funding of this work was provided by the Rockefeller Foundation.

REFERENCES

Bao, Y., and Hull, R. (1994). Replication intermediates of rice tungro bacilliform virus DNA support a replication mechanism involving reverse transcription. *Virology* **204**, 626–633.

Bothner, B., Dong, F., Bibbs, L., Johnson, J., and Siuzdak, G. (1998). Evidence of viral capsid dynamics using limited proteolysis and mass spectrometry. *J. Biol. Chem.* **273**, 673–676.

Bouhida, M., Lockhart, B. E. L., and Olszewski, N. E. (1993). An analysis of the complete sequence of a sugarcane bacilliform virus genome infectious to banana and rice. *J. Gen. Virol.* **74**, 15–22.

Covey, S. N. (1986). Amino acid sequence homology in gag region of reverse transcribing elements and the coat protein gene of cauliflower mosaic virus. *Nucleic Acids Res.* **14**, 623–633.

Dellaporta, S. L., Wood, J., and Hicks, J. B. (1983). A plant DNA preparation: Version II. *Plant Mol. Biol. Rep.* **4**, 19–21.

Fan, Z., Dahal, G., Dasgupta, I., Hay, J., and Hull, R. (1996). Variation in the genome of rice tungro bacilliform virus: Molecular characterization of six isolates. *J. Gen. Virol.* **77**, 847–854.

Fütterer, J., and Hohn, T. (1987). Involvement of nucleocapsids in reverse transcription: A general phenomenon? *Trends Biochem. Sci.* **12**, 92–95.

Hagen, L. S., Jacquemond, M., Lepingle, A., Lot, H., and Tepfer, M. (1993). Nucleotide sequence and genomic organization of cacao swollen shoot virus. *Virology* **196**, 619–628.

Hay, J., Grieco, F., Druka, A., Pinner, M., Lee, S.-C., and Hull, R. (1994). Detection of rice tungro bacilliform virus gene products *in vivo*. *Virology* **205**, 430–437.

Hay, J. M., Jones, M. C., Blakebrough, M. L., Dasgupta, I., Davies, J. W., and Hull, R. (1991). An analysis of the sequence of an infectious clone of rice tungro bacilliform virus, a plant pararetrovirus. *Nucleic Acids Res.* **19**, 2615–2621.

Herdt, R. W. (1991). Research priorities for rice biotechnology. *In* "Rice Biotechnology" (G. H. Toenniessen and G. S. Kush, eds.), pp. 19–54. CAB International, Wallingford, UK.

Hibino, H. (1983a). Relations of rice tungro bacilliform and rice tungro spherical viruses with their vector *Nephotettix virescens*. *Ann. Phytopathol. Soc. Jpn.* **49**, 545–553.

Hibino, H. (1983b). Transmission of two rice tungro-associated viruses and rice waika virus from doubly or singly infected source plants by leafhopper vectors. *Plant Dis.* **67**, 774–777.

Hibino, H., Roechan, M., and Sudarisman, S. (1978). Association of two types of virus particles with penyakit habang (Tungro disease) of rice in Indonesia. *Phytopathology* **68**, 1412–1416.

Hull, R. (1996). Molecular biology of rice tungro viruses. *Annu. Rev. Phytopathol.* **34**, 275–297.

Jacquot, E., Hagen, L. H., Jacquemond, M., and Yot, P. (1996). The open reading frame 2 product of cacao swollen shoot badnavirus is a nucleic acid-binding protein. *Virology* **225**, 191–195.

Jones, M. C., Gough, K., Dasgupta, I., Subba Rao, B. L., Cliffe, J., Qu, R., Shen, P., Kaniewska, M., Blakebrough, M., Davies, J. W., Beachy, R. N., and Hull, R. (1991). Rice tungro disease is caused by an RNA and a DNA virus. *J. Gen. Virol.* **72**, 757–761.

Laco, G. S., and Beachy, R. N. (1994). Rice tungro bacilliform virus encodes reverse transcriptase, DNA polymerase and ribonuclease H activities. *Proc. Natl. Acad. Sci. USA* **91**, 2654–2658.

Laco, G. S., Kent, S. H., and Beachy, R. N. (1995). Analysis of the proteolitic processing and activation of the rice tungro bacilliform virus reverse transcriptase. *Virology* **208**, 207–214.

Laemmli, U. K. (1970). Cleavage of the structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680–685.

Lewis, K., Bendahmane, M., Smith, T. J., Beachy, R. N., and Siuzdak, G. (1998). Identification of viral mutants by mass spectrometry. *Proc. Natl. Acad. Sci. USA* **95**, 8596–8601.

O'Farrell, P. H., Goodman, H. M., and O'Farrell, P. H. (1977). High resolution two-dimensional electrophoresis of basic as well as acidic proteins. *Cell* **12**, 1133–1142.

Pringle, C. R. (1997). Virus taxonomy 1997, Proceedings of the 26th meeting of the Executive Committee of the ICTV. *Arch. Virol.* **142**, 1727–1733.

Qu, R., Bhattacharyaa, M., Laco, G., Kochko de, A., Subba Rao, B. L., Kaniewska, M., Elmer, J. S., Rochester, D. E., Smith, C. E., and Beachy, R. N. (1991). Characterization of the genome of rice tungro bacilliform virus: Comparison with *commelina* yellow mottle virus and caulimoviruses. *Virology* **185**, 354–364.

Rothnie, H. M., Chapdelaine, Y., and Hohn, T. (1994). Pararetroviruses and retroviruses: A comparative review of viral structure and gene expression strategies. *Adv. Virus Res.* **44**, 1–67.

Shen, P., Kaniewska, M., Smith, C., and Beachy, R. N. (1993). Nucleotide sequence and genomic organization of rice tungro spherical virus. *Virology* **193**, 621–630.

Siuzdak, G. (1996). "Mass Spectrometry for Biotechnology." Academic Press, San Diego.

Siuzdak, G., Bothner, B., Yeager, M., Brugidou, C., Fauquet, C., Hoey, K., and Chang, C-M. (1996). Mass spectrometry and viral analysis. *Chem. Biol.* **3**, 45–48.