Arch Virol (2004) 149: 903–914
DOI 10.1007/s00705-003-0268-5

Archives of Virology
Printed in Austria

Molecular characterisation of potexviruses isolated from three different genera in the family *Cactaceae**

R. Koenig[1], **C. W. A. Pleij**[2], **S. Loss**[3], **W. Burgermeister**[1], **H. Aust**[3], and **J. Schiemann**[1]

[1]Biologische Bundesanstalt für Land- und Forstwirtschaft, Institut für Pflanzenvirologie, Mikrobiologie und biologische Sicherheit, Braunschweig, Germany
[2]Leiden Institute of Chemistry, Leiden University, Leiden, The Netherlands
[3]Technical University Braunschweig, Institut für Mikrobiologie, Braunschweig, Germany

Received August 18, 2003; accepted November 3, 2003
Published online January 29, 2004

Summary. The genome properties of three potexviruses which previously had been isolated from different genera in the family *Cactaceae* and had been found to be only distantly related serologically have been studied. The sequence of the 3040 3′ terminal nucleotides of the genomic RNA of isolate K11 from *Schlumbergera bridgesii* and the complete RNA sequences of isolates B1 and CC10 from *Zygocactus* sp. and *Opuntia* sp., respectively, were determined. Starting sequences were obtained by means of immunocapture reverse transcription PCR using primers derived from highly conserved sequences in other potexviral RNAs. The known parts of the sequences were extended by means of random-primed cDNAs and specific primers derived from the known parts of the sequences. The genome structure of the three viruses resembles that of other potexviruses. The conserved motifs typical for replication-associated proteins, triple gene block (TGB) proteins and coat proteins of potexviruses were readily identified in the translation products of the five open reading frames. The 3′ untranslated regions of the three RNAs are folded into secondary structures containing three characteristic hairpins. Rather low percentages of amino acid sequence identities ranging from 62% to 76% for the coat proteins and 41% to 49% for TGB proteins 3 suggest that these viruses

*The GenBank accession numbers for the sequences reported in this paper are AY366208, AY366207 and AY366209 for the genomic RNAs of the B1 isolate from *Zygocactus* sp., the CC10 isolate from *Opuntia* sp. and the K11 isolate from *Schlumbergera bridgesii* (incomplete), respectively.

should be regarded as distinct virus species for which the names Zygocactus virus X, Schlumbergera virus X and Opuntia virus X are proposed. It is also suggested that the name *Cactus virus X* which originally was coined for all three virus isolates should no longer be used.

Introduction

Viruses with particles similar to those of *Potato virus X* (PVX) are widely spread in different genera of the family *Cactaceae*. More than 40 different species in the *Cactaceae* have been reported to be infected [1, 11]. Brandes and Bercks [1] have coined the name *Cactus virus X* (CVX) for an isolate from *Zygocactus* sp. which serologically was only distantly related to PVX, *Hydrangea ringspot virus* and *White clover mosaic virus*. The host reactions produced by this *Zygocactus* virus isolate (named B1) have been compared with those caused by an isolate from *Schlumbergera bridgesii* Lem. (named K11) and four isolates from *Opuntia monoacantha HAW.*, *O. microdasys* Pfeiff., *O. microdasys* var. *albispina* Fobe and *O. tomentosa* S.D., respectively, all from a botanical garden in Yugoslavia [16]. The K11 isolate produced much more severe symptoms on various species in the *Amaranthaceae*, *Caryophyllaceae* and *Chenopodiaceae* than did the B1 isolate and the four isolates from different *Opuntia* species. Milicic et al. [12] found that the *Zygocactus* isolate B1 and the four isolates from *Opuntia* were serologically closely related, whereas the K11 isolate was only distantly related to these viruses. They also studied another isolate named CC10 from an *Opuntia* plant (probably *Opuntia brasiliensis*) in the U.S.A. which was found to be only distantly related serologically to the Yugoslavian *Opuntia* isolates and also to the B1 and K11 isolates. Serological differentiation indices up to 7 and more can be calculated from the data given by [6] and [12]. Milicic et al. [12] have therefore discussed the question whether the three isolates B1, CC10 and K11 should be regarded as separate viruses. Because there was the possibility that the pronounced serological differences between these viruses might have been due to sequence differences confined to the coding regions for the antigenically important parts of the coat proteins, it was decided that they should be regarded as strains of the originally described CVX as long as further information on their proteins and genome properties was lacking. We have now sequenced the entire genomic RNAs of the B1 and CC10 isolates and the 3′ half of the K11 RNA. It was found that these viruses which were isolated from different genera in the *Cactaceae* differ not only in their coat proteins, but at least as much or even more so in most of their non-structural proteins.

Material and methods

The three isolates B1, CC10 and K11 were purified from infected leaves of *C. quinoa* as described by [6]. Viral RNA to be used as a template for PCR was obtained either by means

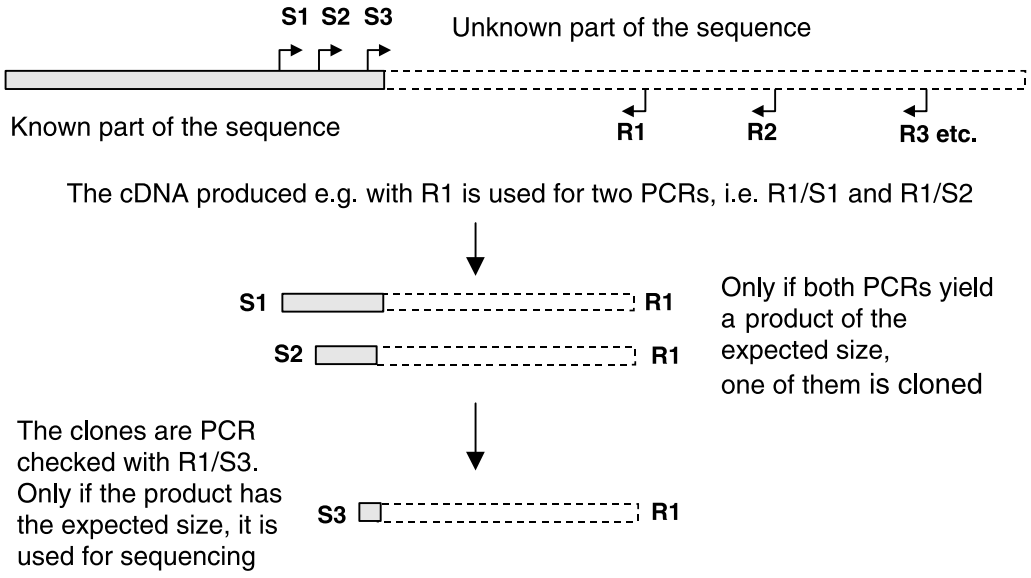


Fig. 1. Strategy used for cloning unkown sequences adjacent to known ones. S1, S2 and S3 are specific primers designed in the known part of the sequence. R1, R2, R3 etc. are c. 20 mer primers produced for other purposes which by chance hybridize at various sites in the unknown part of the sequence. For further details see text

of the Qiagen RNeasy Plant Mini Kit (Cat. No. 74904) or by the immunocapture procedure [8] using γ-globulin specific for each virus. The first portions of the genome of each virus were amplified by means of PCRs using primers derived from sequences which are highly conserved in potexviral RNAs. Four of these primers yielding PCR products were derived from nucleotides (nt) 451–424 (antisense), nt 3244–3274 (sense), nt 3960–3989 (sense) and nt 4316–4284 (antisense) of RNA of an isolate *Narcissus mosaic virus* (Acc. No. NC_001441) and one from nt 1–24 (sense) of PVX RNA (Acc. No. NC_001455). The starter sequences thus obtained were extended by means of an improved version (Fig. 1) of the random primed cDNA approach described earlier [7]. About 15 different 20-mer oligonucleotide primers, originally obtained for other purposes, were annealed separately as random primers to plus-strand RNA from immunocaptured virus particles (to obtain cDNAs for plus-strand RNA downstream of the known part of the sequences) or to RNA from denatured preparations of dsRNA (to obtain cDNA for minus-strand sequences upstream of the known sequence). These random primers are referred to as R1, R2, R3 etc. in Fig. 1 or as Rn below. The annealing was done under low stringency conditions and the annealing temperature lowered from 80 °C to room temperature over 1 hr. Under these conditions, cDNA synthesis starts at several sites, even when only five or six 3′-terminal nts of a given primer are complementary to a certain part of the RNA sequence. Three specific primers, designated as S1, S2 and S3 in Fig. 1, were designed at the end of the known part of the sequence. Two PCRs were done with each of the random-primed cDNAs using the respective random primer which had been employed for cDNA synthesis in combination with S1 and S2, respectively. Only when both primer combinations, i.e. Rn/S1 and Rn/S2, yielded products which showed the size difference expected from the location of S1 and S2 on the known part of the sequence, one of the two products was cloned. The correctness of the cloned sequences

was checked by means of PCR using the primer pair Rn/S3. Only those clones which gave a PCR-product of the expected size were sequenced. By means of this approach all clones sequenced extended the known part of the sequence correctly. For the newly obtained extended sequences again three specific primers were designed and the same set of random-primed cDNAs was used for obtaining further unknown parts of the sequence adjacent to the known parts.

The 3′ ends of the RNAs were obtained by means of oligo(dT)-primed cDNA and a specific primer and the 5′ends by means of the 5′ RACE System for Rapid Amplification of cDNA Ends (Invitrogen Cat. No. 18374-058). For cloning into the pGEM-T vector (Promega), PCR products were purified using the Jetsorb Gel Extraction Kit (Genomed). The sequencing was done by a commercial company (MWG-Biotech, D85560 Ebersberg/Germany). The sequences were analysed by means of the LINEUP/PILEUP programs of the UWGCG software version 8 [4] and trees were generated by the program DNAMAN (Lynnon Bio/Soft). Transmembrane segments were predicted by the DAS program [3]. Secondary structure analysis was done using the STAR program, which is able to predict RNA pseudoknots [5].

Results and discussion

The complete genomic RNA of the B1 isolate consists of 6624 nts nucleotides (nts) and that of the CC10 isolate of 6653 nts excluding the 3′ terminal poly(A) tails. A partial sequence of 3030 nts was determined for the K11 isolate which comprises the entire 3′UTR, the coat protein gene, a triple gene block (TGB) and part of the gene for the replication-associated protein including the entire region containing the RNA polymerase motifs. The genome organisation of these isolates proved to be identical to that of other potexviruses (see Fig. 2A for the B1 isolate). ORF1 codes for a large protein containing the three methyltransferase, the six helicase and the eight RNA-dependent RNA polymerase motifs (Fig. 2A, Fig. 3) which are typical for replication-associated proteins of ssRNA viruses [9]. ORF1 is followed by the TGB (Fig. 2A). The first TGB proteins of all three virus isolates contain in the region of amino acid (aa) 23–216 the six helicase motifs which are typically found in the corresponding proteins of other potexviruses (results not shown). These motifs are preceeded in position 15 by a conserved arginine which is supposed to be involved in RNA binding [13]. As in other potexviruses the TGB proteins 2 contain two transmembrane segments (shown for the B1 isolate in Fig. 2B), a conserved GD7xGGxYxG motif between them and a positively charged C-terminus [13]. The TGB proteins 3 of the B1, K11 and CC10 isolates – like those of other potex- and also carla-, allexi- and foveaviruses – contain only one transmembrane segment. It extends from aa 6 to 20 in the TGB protein 3 of the B1 isolate and is followed by the conserved C5xG7xC motif [13] in aa positions 28 to 42.

The location of ORF 5 near the 3′ end of the B1, CC10 and K11 RNAs and the presence in its translation products of several conserved sequence stretches which are typically present in the coat proteins of other potexviruses suggest that this ORF is the coat protein gene. The conserved sequences ‘RQFCRYFAK’ and ‘FAAFDFFD’ are present in the central parts of the B1 and K11 ORF 5 gene

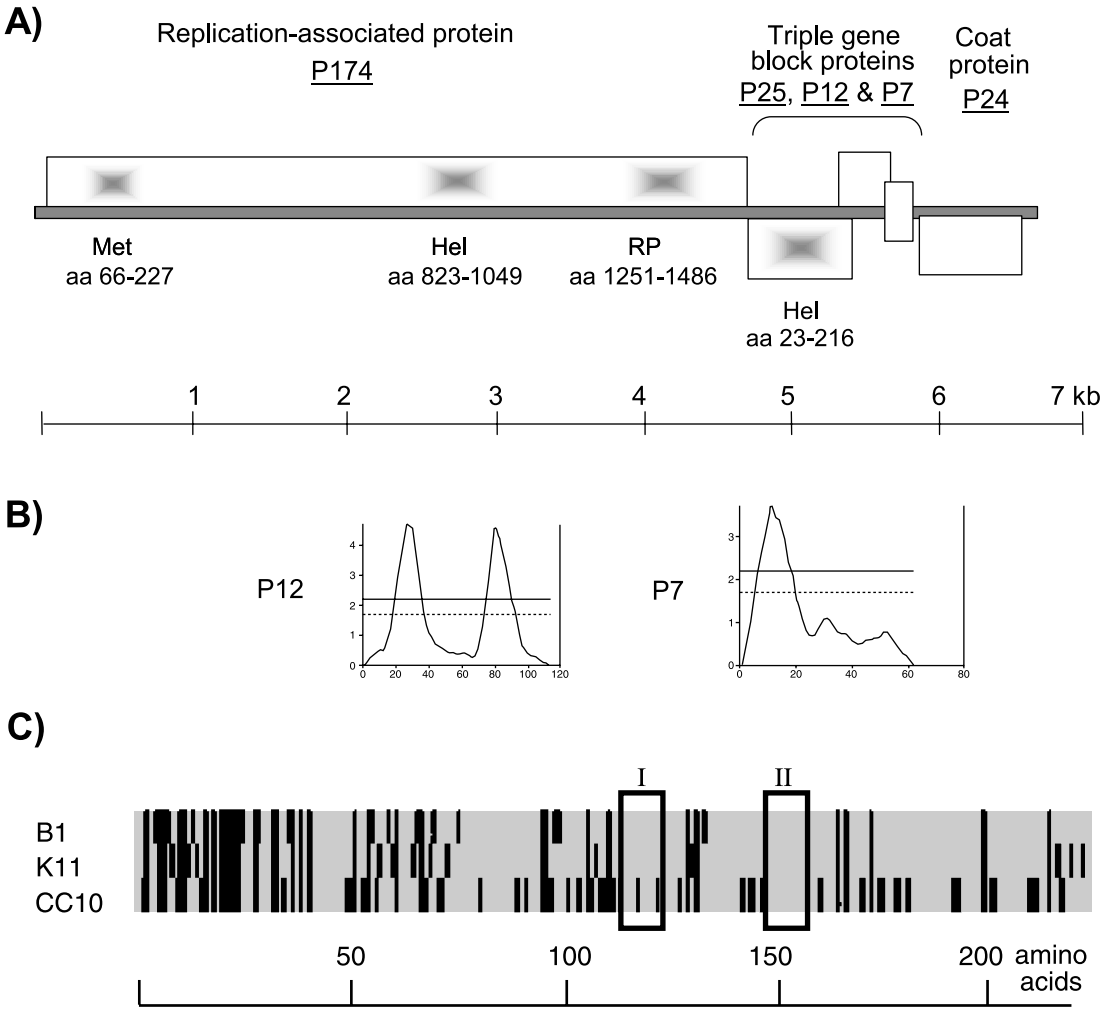


Fig. 2. Molecular properties of the potexvirus isolates B1, K11 and CC10 from various genera in the *Cactaceae*. **A**) Genome organization of the B1 isolate from *Zygocactus* sp. with coding regions for proteins of 174, 25, 12, 7 and 24 kD. The regions where the three methyltransferase (*Met*), the six helicase (*Hel*) and the eight RNA-dependent RNA polymerase (*RP*) motifs typical for replication-associated proteins are found on P174 and the region where the six helicase motifs typical for the first triple gene block proteins of potexviruses is found on P25 are indicated. **B**) Two or one transmembrane segment were readily detected on the second (*P12*) and the third (*P7*) triple gene block proteins, respectively, by means of the Dense Alignment Surface method [3]. **C**) Alignment of the coat protein amino acid (*aa*) sequences of the three virus isolates starting at the N terminus. A black bar indicates that the amino acid in the corresponding position differs from those in the same position of the coat proteins of the other viruses. One amino acid is represented by a single bar. Conserved regions in motifs I and II [9] are boxed

products and, with two exchanges in the first sequence, also in that of the CC10 ORF isolate (Fig. 2C). They form part of motifs I and II which contain a conserved R (positively charged) and a conserved D (negatively charged) (highlighted above

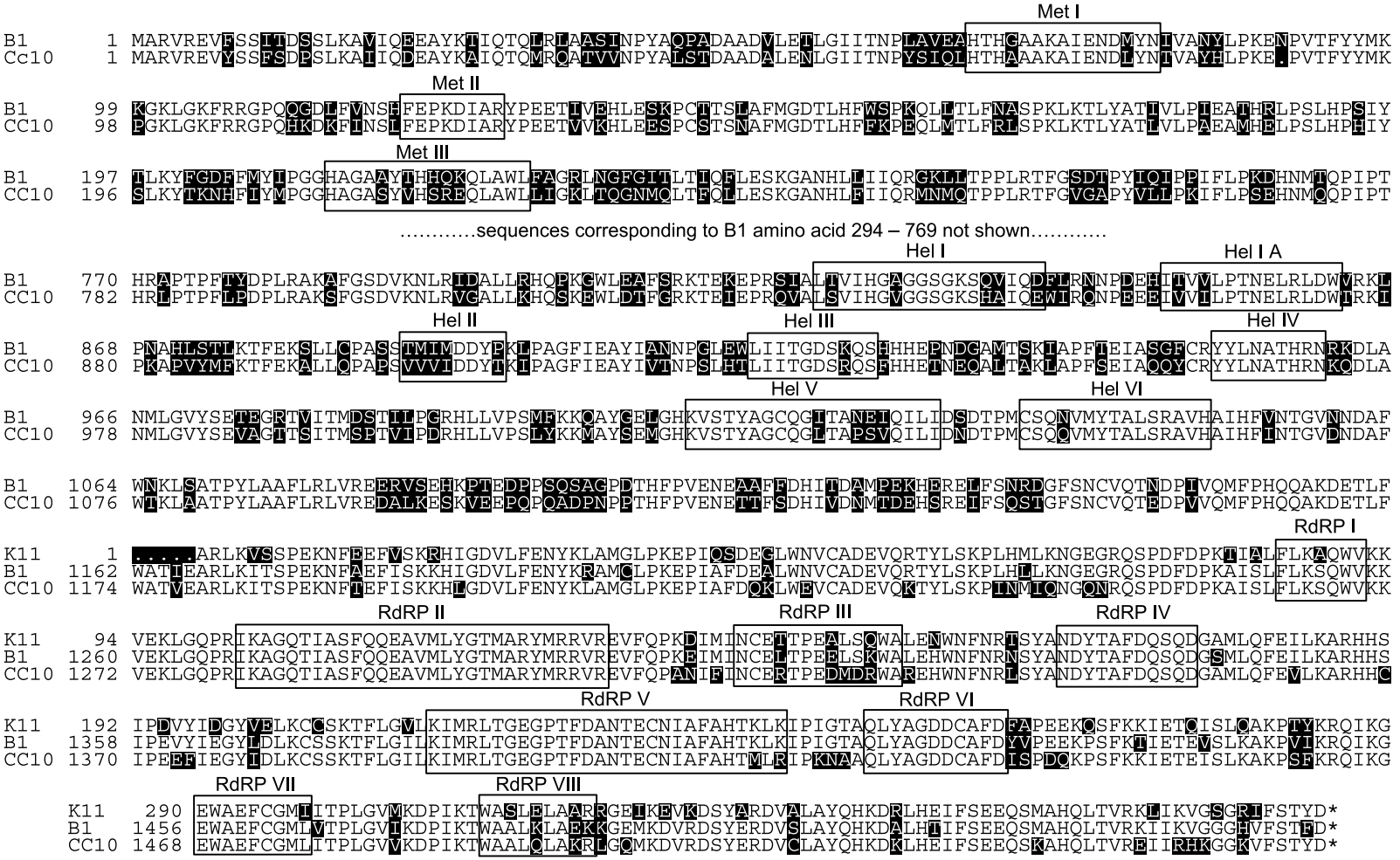
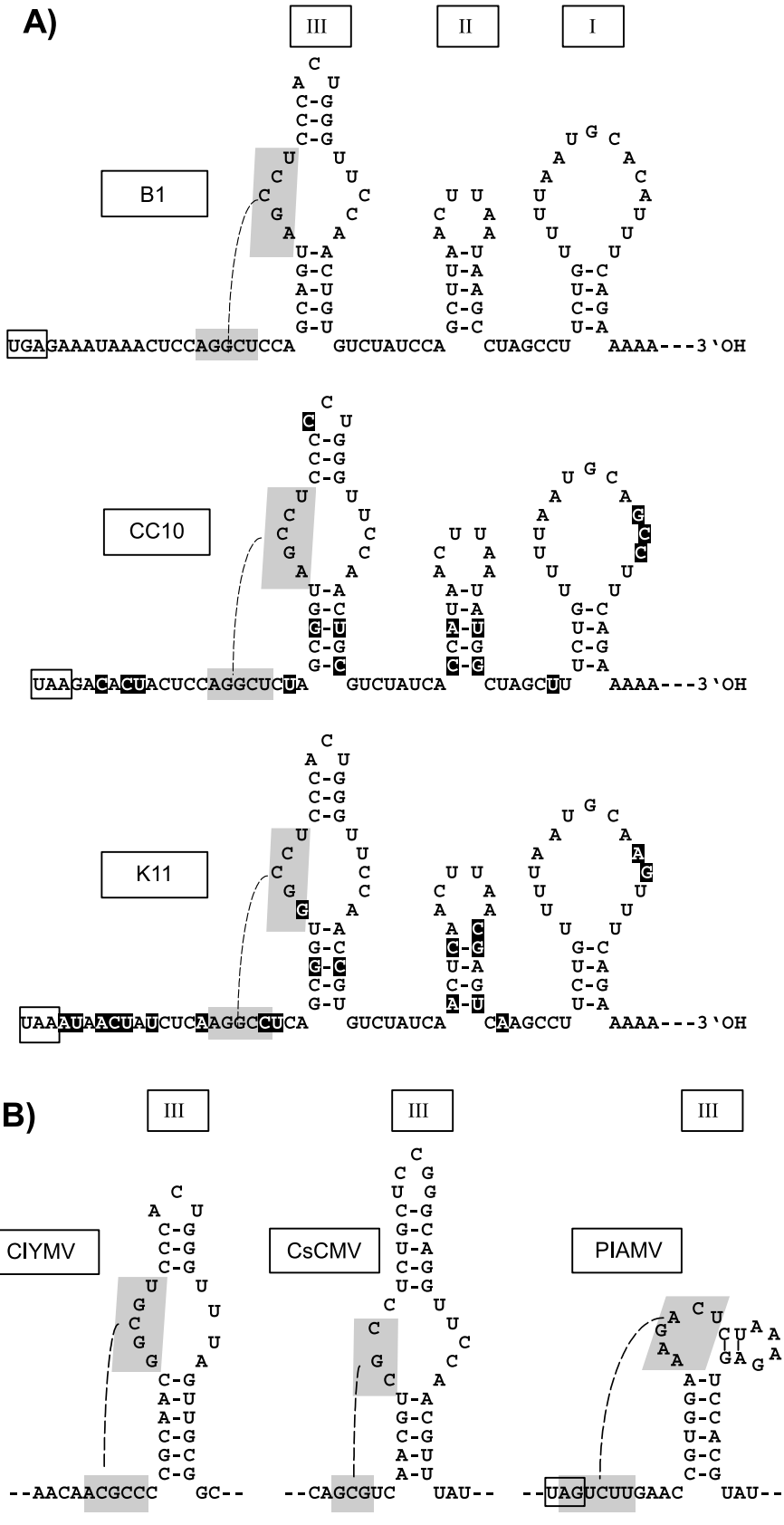


Fig. 3. Alignment of the replication-associated proteins of B1, CC10 (complete sequences) and K11 (partial sequence). The locations of the three methyltransferase (*Met I – III*), the six helicase (*Hel I – VI*) and the eight RNA-dependent RNA polymerase motifs (*RdRP I* to *VIII*) [9] are indicated. Amino acids differing in the three sequences are highlighted by white letters on a black background. The portion of the sequence of amino acids 294 to 769 in the B1 protein and the corresponding portion in the CC10 protein which are less conserved are not shown

by white letters on a black background), respectively. These two residues have been suggested to form a functionally important salt bridge [9]. An alignment of the coat proteins of the B1, CC10 and K11 virus isolates reveals considerable differences especially in their N-terminal parts (Fig. 2C) which may be the reason for the low degree of serological cross-reactivity found between these viruses [6, 12].

The 3′ UTRs of all three isolates are folded into an identical secondary structure, consisting of three hairpins (Fig. 4A). The predicted 3′ terminal hairpin I has a relatively large hairpin loop. The 4 bp stem is identical for all three isolates and is therefore not supported by covariations. This hairpin strongly resembles hairpin SL2 in the 3′UTR of PVX [14]. Covariations were found, however, for the stem of hairpin II and the bottom part of hairpin III. Hairpin II shows two covariations each in CC10 and K11 as compared to B1, which supports the existence of this stem, which in the case of K11 may consist only of four base pairs. Interestingly, in B1 and CC10 the loop consists exclusively of the conserved hexanucleotide motif (ACUUAA) which has been found to play an essential role in the viral RNA multiplication [14, 19]. In many other potexvirus 3′UTRs, e.g. those of PVX and *Bamboo mosaic virus* (BaMV), this motif shows up as part of a hairpin loop (results not shown). Hairpin III contains a symmetrical internal loop of which the 5′ side can basepair with a sequence outside the hairpin, thereby forming a pseudoknot structure. Although this basepairing possibility is only supported by a single covariation among the three isolates B1, CC10 and K11, strong support comes from the folding of a few other potexvirus 3′ UTRs, i.e. those of *Clover yellow mosaic virus*, *Cassava common mosaic virus* and *Plantago asiatica mosaic virus* (Fig. 4B). This type of pseudoknot, involving an internal loop is very rare and can be considered to be derived from a classic pseudoknot by inserting a stem-loop in loop L2 of the classic or H-type pseudoknot [10, 15]. In the case of PlAMV the upper stem was not predicted by the program STAR and may therefore be questionable, despite the resemblance it produces with the other proposed pseudoknot structures. If the two basepairs are not present indeed, it is reduced to just a classic pseudoknot. Analysis of the folding of other potexviruses 3′UTRs did not give indications for the existence of this pseudoknot structure nor was it reported before for BaMV and PVX [2, 14].

The relationships between the three isolates B1, CC10 and K11 among each other and with the other potexviruses for which sequence data are available were analysed by determining the percentages of sequence identities for their structural and non-structural proteins (Fig. 5). Similar trees were obtained by means of the neighbour joining method [17] (results not shown). The percentages of sequence identities among the coat proteins and among each of the three TGB proteins suggest somewhat closer relationships between the B1 and K11 isolates than between these two isolates and the CC10 isolate from the U.S.A. The percentages of sequence identities for the TGB proteins 1 of these isolates resemble those which were determined for their coat proteins. The relationships


A)
III
II
I
B1
UGAGAAAUAAACUCCAGGCUCCA GUCUAUCCA CUAGCCU AAAA---3'OH
CC10
UAAGACACUACUCCAGGCUCUA GUCUAUCA CUAGCUU AAAA---3'OH
K11
UAAAUAACUAUCUCAAGGCUCA GUCUAUCA CAAGCCU AAAA---3'OH
B)
III
III
III
ClYMV
CsCMV
PlAMV
--AACAACGCCC GC--
--CAGCGUC UAU--
--UAGUCUUGAAC UAU--

between their TBG proteins 2 and especially their TGB proteins 3 are even more distant (Fig. 5).

The relationships among the various proteins of the B1, CC10 and K11 isolates are somewhat closer than those to and among the other potexviruses (Fig. 5), but they are far more distant than those between strains in other virus genera, e.g. poty-, tobamo-, tombus- or tymoviruses which usually share sequence identities of at least 80% and often more than 90% [see also 18]. As mentioned above, Milicic et al. [12] had already raised the question whether these isolates should be considered as distinct viruses, but since at that time the possibility could not be excluded that the differences between these viruses might have been confined to the immunogenic surface structures of their coat proteins, they were provisionally regarded as strains of the same virus. The sequence data presented here revealed that these isolates differ not only in their coat proteins, but at least as much or even more so in most of their other gene products. This observation and also the fact that the virus isolates B1, K11 and CC10 were obtained from different genera in the *Cactaceae* and that the two of them which have been checked, i.e. B1 and K11, produce different symptoms on host plants [16] suggest that they should be regarded as separate viruses for which we propose the names Zygocactus virus X, Schlumbergera virus X and Opuntia virus X, respectively. The name *Cactus virus X* which was originally proposed for all three virus isolates [1, 12] should no longer be used.

In the various trees in Fig. 5 the names of those viruses have been underlined which form a similar pseudoknot structure in their 3′UTRs as the isolates B1, CC10 and K11 do (Fig. 4). In the trees for the replication-associated proteins and for the coat proteins these viruses are all found in the same cluster, but not so in the tree for TGB proteins 3 and not consistently in the trees for TGB proteins 1 and 2.

Because the isolates CC10 and K11 and especially B1 cause only relatively mild symptoms on *Chenopodium quinoa*, infectious cDNA clones of these viruses might serve as interesting tools for the expression of foreign genes in this plant which is a host for many plant viruses. We have, therefore, cloned the entire sequence of the B1 isolate into the 35S-promoter-containing vector p35Stu_pa which was kindly provided by Prof. Dr. E. Maiss, University of Hannover, Germany. Preliminary tests indicated that these clones were infectious. Detailed studies are now underway to check whether they can be modified in a such a manner that they can be used as expression vectors in plants.

Fig. 4. **A**) Folding of the 3′UTRs of the virus isolates B1, CC10 and K11 and **B**) Detection of stemloop III in the 3′UTRs of three other potexviruses, i.e. *Clover yellow mosaic virus* (ClYMV; nt 6914–6949), *Cassava common mosaic virus* (CsCMV; nt 6283–6321) and *Plantago asiatica mosaic virus* (PlAMV; nt 5998–6033). For sequence accession numbers see legend to Fig. 5. Nucleotides in the CC10 and K11 sequences which differ from those in the corresponding positions of the B1 sequence are highlighted by white letters on a black background

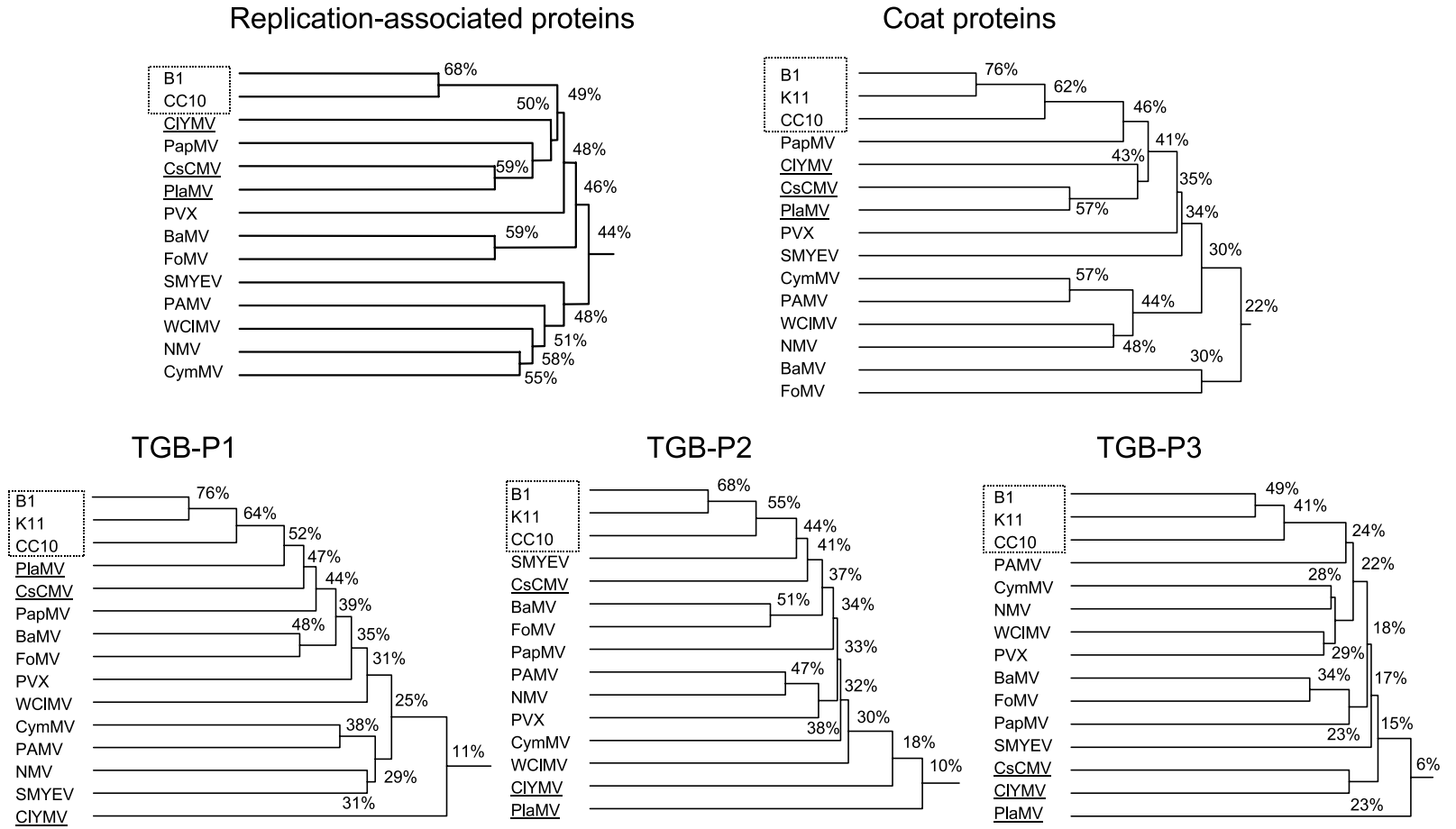


Fig. 5. Trees showing the percentages of amino acid sequence identities which the various proteins of the virus isolates B1, CC10 and K11 share with the corresponding proteins of other potexviruses, i.e. *Bamboo mosaic virus* (BaMV, Acc. No. D26017), *Cassava common mosaic virus* (CsCMV, Acc. No. U23414), *Clover yellow mosaic virus* (ClYMV, Acc. No. NC_001753), *Foxtail mosaic virus* (FoMV, Acc. No. M62730), *Narcissus mosaic virus* (NMV, Acc. No. NC_001441), *Papaya mosaic virus* (PapMV, Acc. No. D13957), *Plantago asiatica mosaic virus* (PlaMV, Acc. No. Z21647), *Potato aucuba mosaic virus* (PAMV, Acc. No. S73580), *Potato virus X* (PVX, Acc. No. M95516), *Strawberry mild yellow edge virus* (SMYEV, Acc. No. NC_003794) and *White clover mosaic virus* (WClMV, Acc. No. NC_003820). The names of the viruses which show similar stem loops III in their 3′UTRs as the three isolates B1, CC10 and K11 do (Fig. 3) are underlined

Acknowledgements

We are greatly indebted to the Deutsche Forschungsgemeinschaft (grant Ko 518/14) and the Niedersächsisches Ministerium für Wissenschaft und Kultur, Forschungsschwerpunkt Agrarbiotechnologie (grant ZN1401) for financially supporting different aspects of this research.

References

1. Brandes J, Bercks R (1962) Untersuchungen zur Identifizierung und Klassifizierung des Kakteen-X-Virus (Cactus virus X). Phytopath Z 46: 291–300
2. Cheng C-P, Tsai C-H (1999) Structural and functional analysis of the 3′ untranslated region of bamboo mosaic Potexvirus genomic RNA. J Mol Biol 288: 555–565
3. Cserzo M, Wallin E, Istcvan S, von Heijne G, Elofsson A (1997) Prediction of transmembrane alpha-helices in procaryotic membrane proteins: the Dense Alignment Surface method. Prot Eng 10: 673–676; http://www.sbc.su.se/∼miklos/DAS/maindas html
4. Devereux J, Haeberli P, Smithies O (1984) A comprehensive set of sequence analyses for the VAX. Nucleic Acids Res 12: 387–395
5. Gultyaev AP, van Batenburg FHD, Pleij CWA (1995) The computer simulation of RNA folding pathways using a genetic algorithm. J Mol Biol 250: 37–51
6. Koenig R, Bercks R (1968) Änderungen im heterologen Reaktionsvermögen von Antiseren gegen Vertreter der *Potato virus X*-Gruppe im Laufe des Immunisierungsprozesses. Phytopath Z 61: 382–398
7. Koenig R, Commandeur U, Loss S, Beier C, Kaufmann A, Lesemann D-E (1997) Beet soil-borne virus RNA 2: similarities and dissimilarities to the coat protein gene-carrying RNAs of other furoviruses. J Gen Virol 78: 469–477
8. Koenig R, Lüddecke P, Haeberlé AM (1995) Detection of beet necrotic yellow vein virus strains, variants and mixed infections by examining single-strand conformation polymorphisms of immunocapture RT-PCR products. J Gen Virol 76: 2051–2055
9. Koonin EV, Dolja VV (1993) Evolution and taxonomy of positive-strand RNA viruses: Implications of comparative analysis of amino acid sequences. Crit Rev Biochem Mol Biol 28: 375–430
10. Mans RMW, Pleij CWA (1993) RNA pseudoknots. In: Eckstein F, Lilley DMJ (eds) Nucleic acids and molecular biology, vol. 7. Springer, Berlin Heidelberg New York Tokyo, pp 250–270
11. Milicic D, Udjbinac Z (1961) Virus-Eiweissspindeln der Kakteen in Lokalläsionen von *Chenopodium*. Protoplasma 53: 584–596
12. Milicic D, Plese N, Bercks R, Brandes J, Casper R, Chessin M (1966) Vergleichende serologische und elektronenmikroskopische Untersuchungen an Isolaten des Kakteen-X-Virus. Phytopath Z 55: 211–217
13. Morozov SY, Solovyev AG (2003) Triple gene block: molecular design of a multifunctional machine for plant virus movement. J Gen Virol 84: 1351–1366
14. Pillai-Nair N, Kim K-H, Hemenway C (2003) Cis-acting regulatory elements in the Potato virus X 3′non-translated region differentially affect minus-strand and plus-strand RNA accumulation. J Mol Biol 326: 701–720
15. Pleij CWA, Rietveld K, Bosch L (1985) A new principle of RNA folding based on pseudoknotting. Nucleic Acids Res 13: 1713–1717
16. Plese N, Milicic D (1966) Vergleichende Untersuchungen an Isolaten des Kakteen-X-Virus mit Testpflanzen. Phytopath Z 55: 197–210
17. Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4: 406–425

18. van Regenmortel MHV, Bishop DHL, Fauquet MC, Mayo MA, Maniloff J, Calisher CH (1997) Guidelines to the demarcation of virus species. Arch Virol 142: 1505–1518
19. White KA, Bancroft JB, Mackie GA (1992) Mutagenesis of a hexanucleotide sequence conserved in potexvirus RNAs. Virology 189: 817–820

Author's address: Dr. R. Koenig, Biologische Bundesanstalt für Land- und Forstwirtschaft, Institut für Pflanzenvirologie, Mikrobiologie und biologische Sicherheit, Messeweg 11, 38104 Braunschweig, Germany; e-mail: r.koenig@bba.de