The structure of tobacco ringspot virus: a link in the evolution of icosahedral capsids in the picornavirus superfamily

Veda Chandrasekar[1,2] and John E Johnson[1]*

Background: Tobacco ringspot virus (TRSV) is a member of the nepovirus genus of icosahedral RNA plant viruses that cause disease in fruit crops. Nepoviruses, comoviruses and picornaviruses are classified in the picornavirus superfamily. Crystal structures of comoviruses and picornaviruses and the molecular mass of the TRSV subunit (sufficient to accommodate three β-barrel domains) suggested that nepoviruses may represent a link in the evolution of the picornavirus capsids from a T = 3 icosahedral virus. This evolutionary process is thought to involve triplication of the capsid protein gene, to encode a three-domain polyprotein, followed by development of cleavage sites in the interdomain linking regions. Structural studies on TRSV were initiated to determine if the TRSV subunit corresponds to the proposed uncleaved three-domain polyprotein.

Results: The 3.5 Å resolution structure of TRSV shows that the capsid protein consists of three β-barrel domains covalently linked by extended polypeptides. The order of connectivity of the domains in TRSV confirms the proposed connectivity for the precleaved comovirus and picornavirus capsid polyprotein. Structural differences between equivalent domains in TRSV and comoviruses are confined to the external surface loops, interdomain connecting polypeptides and N termini. The three different domains within TRSV and comoviruses are more closely related at the structural level than the three individual domains within picornaviruses.

Conclusions: The structural results confirm the notion of divergent evolution of the capsid polyproteins of nepoviruses, comoviruses and picornaviruses from a common ancestor. A number of residues were found to be conserved among various nepoviruses, some of which stabilize the quaternary structure of the three domains in the TRSV capsid protein subunit. Two conserved regions were identified on the external surface of TRSV, however, mutational studies will be needed to understand their functional significance. Nepoviruses transmitted by the same nematode species do not share regions with similar amino acid composition on the viral surface.

Addresses: [1]Department of Molecular Biology, MB-31, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, CA 92037, USA and [2]Department of Biological Sciences, Purdue University, 1392 Lilly Hall of Life Science, West Lafayette, IN 47907, USA.

*Corresponding author.
E-mail: jackj@noda.scripps.edu

Key words: comovirus, nematode transmission, nepovirus, picornavirus, tobacco ringspot virus, viral evolution

Received: **1 August 1997**
Revisions requested: **8 September 1997**
Revisions received: **19 December 1997**
Accepted: **19 December 1997**

Structure 15 February 1998, **6**:157–171
http://biomednet.com/elecref/0969212600600157

Introduction

High-resolution structures of several icosahedral RNA viruses belonging to eight different families have been determined (e.g. [1]). The capsids of these viruses are composed of 180 subunits and fall into two categories: the T = 3 viruses, in which all the subunits are formed from the same gene product, and the pseudo T = 3, picorna-like viruses, in which there are 60 copies of three different, but structurally similar, subunits. The three different subunits in the latter particle are synthesized as a polyprotein and occupy positions in the surface lattice comparable to those of the single gene product in a T = 3 virus.

Tobacco ringspot virus (TRSV), a small icosahedral RNA plant virus, is a member of the nepovirus genus [2,3]. The TRSV capsid is composed of 60 copies of a single capsid protein (56,000 Da; 513 amino acids [4,5]). The nepovirus and comovirus genera in the comoviridae family are classified in the picornavirus superfamily. This classification is based on their similarity to animal picornaviruses in the translation of the genomic RNA as a polyprotein, the cleavage of this polyprotein by a virally encoded proteinase, and the sequence similarity among non-structural proteins [6]. Nevertheless, there is no statistically significant sequence similarity between the capsid polyproteins of nepoviruses, comoviruses and picornaviruses. Previous structural studies of comoviruses and picornaviruses have suggested that picorna-like viruses have evolved from T = 3 viruses by triplication of the gene coding a β-barrel domain into a three-domain polyprotein, followed by independent evolution of the β barrels and the development of cleavage sites in the capsid polyprotein [7,8]. The

picornavirus capsid polyprotein is cleaved at two sites to yield three subunits, each of which is folded into a β-barrel domain. The comovirus polyprotein, on the other hand, is cleaved at only one site to yield two subunits: one composed of two β barrels and the other composed of one β barrel. Given the size of the nepovirus capsid protein, it was hypothesized that its capsid protein may contain three covalently linked β-barrel domains, and that nepoviruses may represent an example of an early stage in the development of picornaviruses [7]. The structure determination of a nepovirus and the structural comparison with comoviruses is, therefore, of evolutionary significance.

In addition to protecting the encapsidated nucleic acid from degradation, the capsid of a virus plays a crucial role in several aspects of the viral life cycle. In the case of nepoviruses, previous studies have indicated the role of the capsid in interactions with the soil-inhabiting nematode (eel worm) vector; these interactions dictate specificity during virus transmission [9,10]. The capsid has also been implicated in the movement of the virus across the plant cell wall through cytoplasmic connections formed by the viral movement protein (and other unknown host factors [11]). Hence, an examination of the surface topography of a nepovirus capsid structure may provide structural insights into these functions.

The TRSV genome is bipartite and consists of two positive-sense, single-stranded RNA molecules: RNA1 (2.6×10^6 Da) that encodes four non-structural proteins; and RNA2 (1.2×10^6 Da) that encodes the capsid protein and an upstream movement protein. The two RNA molecules are encapsidated separately in particles of 280 Å diameter [12,13]; both particles are required for infection. The virus particles purified from infected plants sediment as three separate components, depending on the RNA content, and include empty capsids containing no RNA [2]. The crystals used in the present study contain a mixture of all the three components. Here we present the 3.5 Å structure of the first nepovirus, TRSV, and a structural comparison of TRSV with comoviruses. Our structural results support the evolutionary link among the three groups in the picornavirus superfamily and confirm the connectivity of the domains in the nascent polyprotein proposed from the structures of picornaviruses [14,15] and comoviruses [7].

Results and discussion

Structure of the whole capsid

The capsid of TRSV consists of a protein shell with an average thickness of 40 Å (Figure 1a). The outer and inner surfaces of the capsid display considerable variation in the radial dimension resulting in a non-spherical density distribution. The most prominent features on the outer surface are the pronounced protrusions near the icosahedral fivefold axes. There is a minor protrusion at each icosahedral threefold axis and a depression near each icosahedral twofold axis, where the capsid has a minimum thickness of 14 Å. The inner and outer radial dimensions along the icosahedral symmetry axes (Figure 1a) are in agreement with the cryo-electron microscopy (cryo-EM) reconstruction of TRSV virus-like particles (VLPs) that result from expressing the capsid protein gene in a baculovirus system [16]. A comparison of the dimensions of TRSV with the comovirus cowpea mosaic virus (CPMV) [7,17] demonstrates that the overall capsid shape of the two viruses is similar.

Quality of the electron-density map

The overall quality of the electron-density map at 3.5 Å resolution is very good (Figure 1b) in spite of the lack of completeness in the diffraction data. The density corresponding to the polypeptide backbone is continuous at a level of 2σ above the background. Most of the amino acid sidechains have well-defined density, allowing the assignment of the known sequence for the TRSV capsid protein without ambiguity. All 513 residues present in the capsid protein were modeled as one continuous chain.

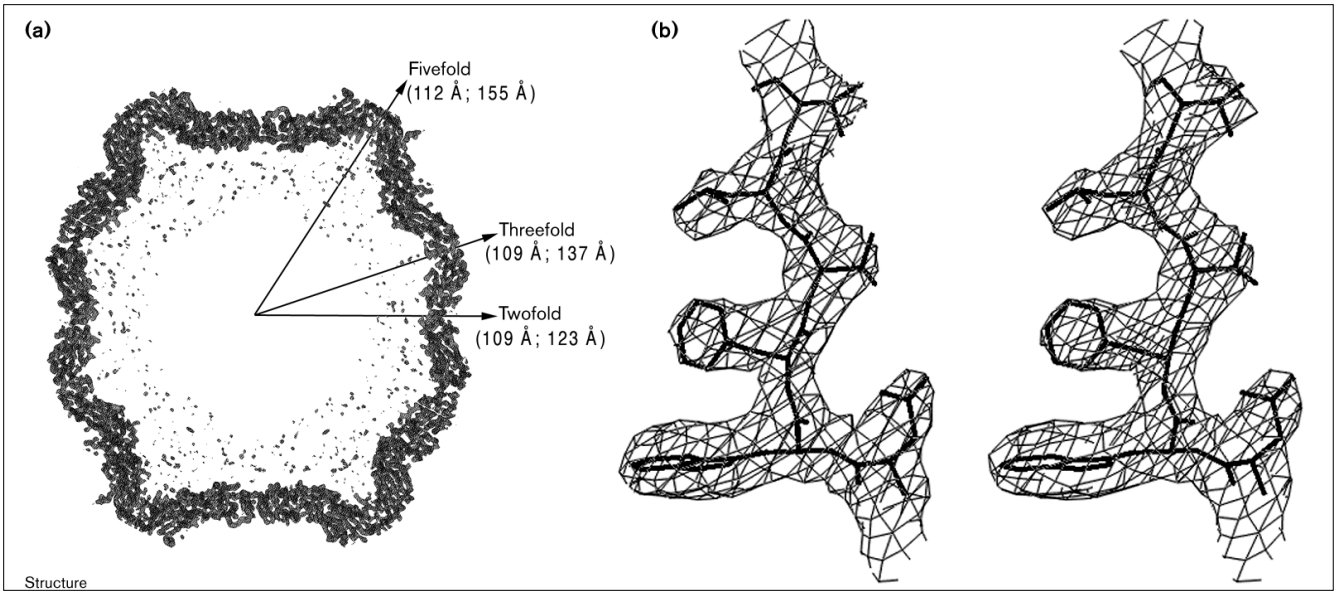
Structure of the capsid protein

The TRSV capsid is composed of 60 copies of a single capsid protein with domains arranged in a pseudo T = 3 icosahedral lattice [18] (Figure 2a). The capsid protein is folded into three trapezoid-shaped β-barrel domains (designated as C, B and A from the N to the C terminus) that are covalently linked together as 'beads on a string'. The B and C domains lie side-by-side with their narrow ends packed together around the virus threefold axes (Figure 2a). The β barrels in these domains lie with the strands of the β sheets tangential to the capsid surface. The prominent protrusion along the fivefold axes is formed by the tilted packing of the β barrels in the pentameric cluster of A domains, where the strands run roughly parallel to the nearby fivefold axis (Figure 2a).

The connected domains in the TRSV capsid protein (Figures 2b and 2c) agree with the connectivity proposed for the precleaved comovirus and picornavirus capsid polyprotein protomer [7,14,15]. The C domain (containing 179 amino acids) forms the N-terminal third of the capsid protein, the B domain contains the central third of 170 amino acids and the A domain (containing 164 amino acids) constitutes the C-terminal third of the protein. The junction point within the two extended polypeptides linking the domains was assigned based on the structure of the comovirus beanpod mottle virus (BPMV) [17].

The tertiary structure of the wedge-shaped C, B and A domains in TRSV (Figure 2c) is a variation of the canonical eight-stranded antiparallel β-sandwich fold, where the eight strands βB to βI form the BIDG and the CHEF β sheets [19,20]. The twist of the β strands constituting the

Figure 1



The quality of the 3.5 Å electron-density map at 2σ contour level. **(a)** A thin slab of the electron-density map viewed down the crystallographic twofold axis and containing the central section of the TRSV particle. The inner and outer radial dimensions along the icosahedral symmetry axes are indicated beneath the axes labels. **(b)** Density corresponding to the sidechains of the residues Asn356, Trp357, Phe358, Thr359, Leu360 and Thr361 forming the βB strand in the A domain.

barrels in the three domains vary. As in comoviruses, the strands in the B domain have the largest twist relative to the other domains. In general, the loops connecting the strands of the β barrel in the three domains are short, but some are of sufficient length to adopt a secondary structure. Inserts of this type in TRSV are also present in the corresponding positions in comoviruses.

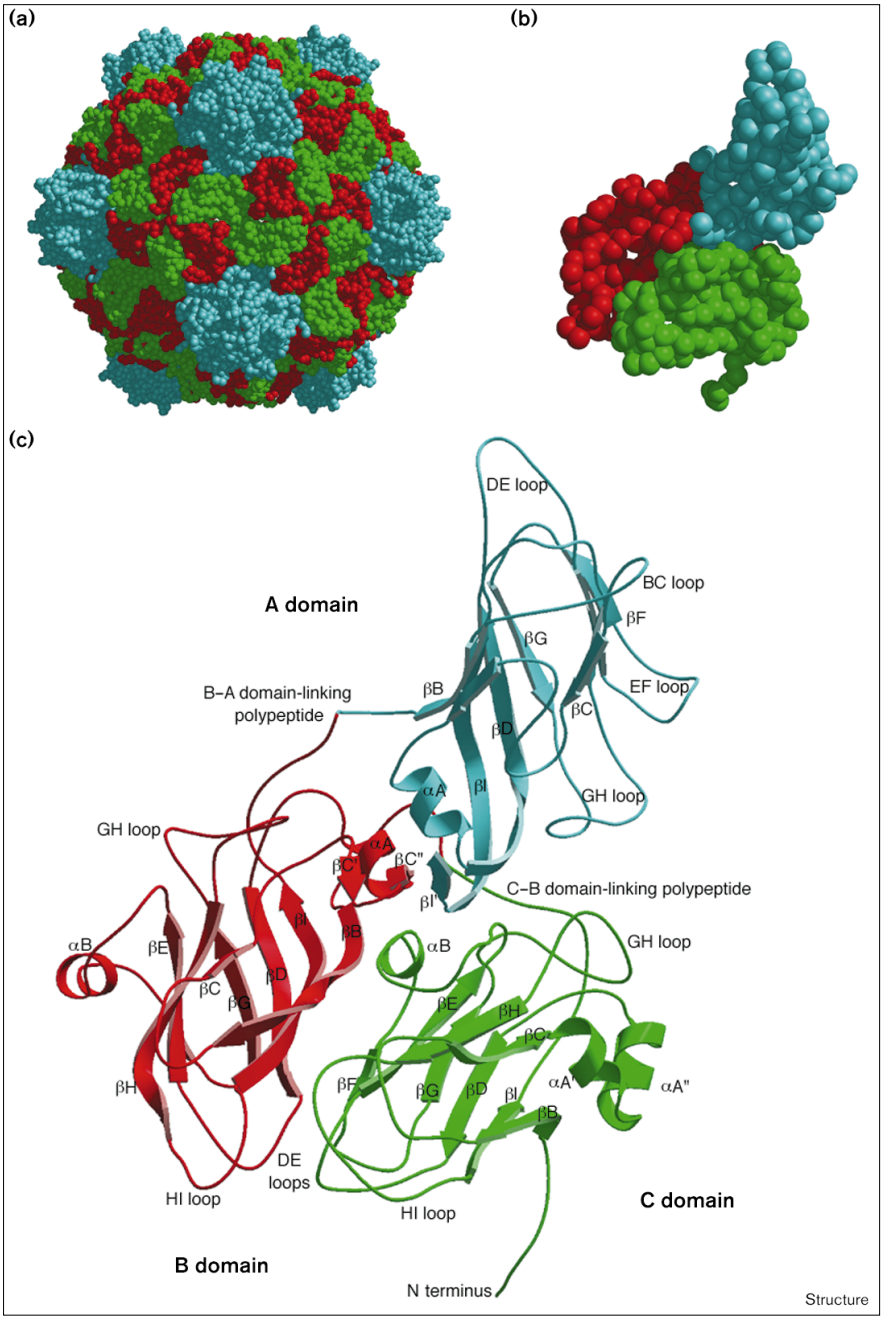
The C domain in TRSV has two helical insertions: the βC–βD insertion containing the αA′ and αA helices; and the βE–βF insertion containing the αB helix. The αC helix present in comoviruses between the βG and βH strands is absent in TRSV. The largest insertion in the B domain occurs between the βC and βD strands and includes two antiparallel β strands, βC′ and βC″, and the αA helix. The βE–βF insertion containing the αB helix is also present in this domain. The only insertion in the A domain is between the βC and βD strands; the insertion contains an extra β strand βC″ and the αA helix. Unlike comoviruses, where the strands βC′ and βC″ lie on top of the canonical β barrel leading to an unusual ten-stranded structure, in TRSV there is no βC′. The βC″ strand, however, does lie on top of the CHEF wall. The short N-terminal tail in the C domain, containing only non-basic sidechains, extends on the inner surface of the capsid to the B domain of the threefold-related capsid protein. The C terminus of the capsid protein present in the A domain forms an extra strand, βI′, and extends toward the outer surface of the virus. The similarity between the Cα traces of the three domains in TRSV (Figure 3) support the hypothesis that the three β barrels in the nepovirus subunit have evolved from a T = 3 virus as a result of gene triplication.

Interdomain connecting polypeptides

The extended polypeptide of 25 residues that connects the C and the B domains (Figure 4a), starts from the end of the βI strand in the C domain. The polypeptide runs roughly perpendicular to the β strands of the C domain on the inner surface of the capsid until it reaches the A domain, where it makes a sharp upwards turn and runs alongside the end of the βI strand in the A domain. Finally, the linking peptide makes another sharp downwards turn and reaches the βB strand in the B domain. The equivalent C–B domain-connecting polypeptide in comoviruses contains 14 residues and has a very different conformation (Figure 4b). This polypeptide in BPMV defines one side of a shallow pocket between the B and C domains that serves as the binding site for ordered RNA [7].

The B–A domain-connecting polypeptide of 15 residues, starts from the end of the βI strand in the B domain. This polypeptide runs perpendicular to the strands in the B domain on the inner surface of the capsid and then makes a turn at the residue Pro347 and ends at the βB strand in the A domain. The equivalent B–A domain-connecting polypeptide in comoviruses is formed by the C-terminal tail of the large subunit and the N-terminal tail of the small subunit, which arise as a result of

Figure 2



Structure of the TRSV capsid, capsid protein and the three domains in the capsid protein. **(a)** A CPK model of the pseudo T = 3 TRSV capsid showing the prominent surface features. The C and B domains (in green and red, respectively) are clustered around the threefold axes, while the A domains (cyan) are clustered around the fivefold axes. **(b)** An enlarged view of one of the 60 copies of the capsid protein in the TRSV capsid represented as a CPK model. **(c)** A cartoon representation of the tertiary structure of the C, B and A domains in the capsid protein. The secondary structure features in the three domains, the two domain-linking polypeptides and the N and C termini of the capsid protein are indicated. (Figure 2c was generated using MOLSCRIPT [40] modified by Robert Esnouf.)

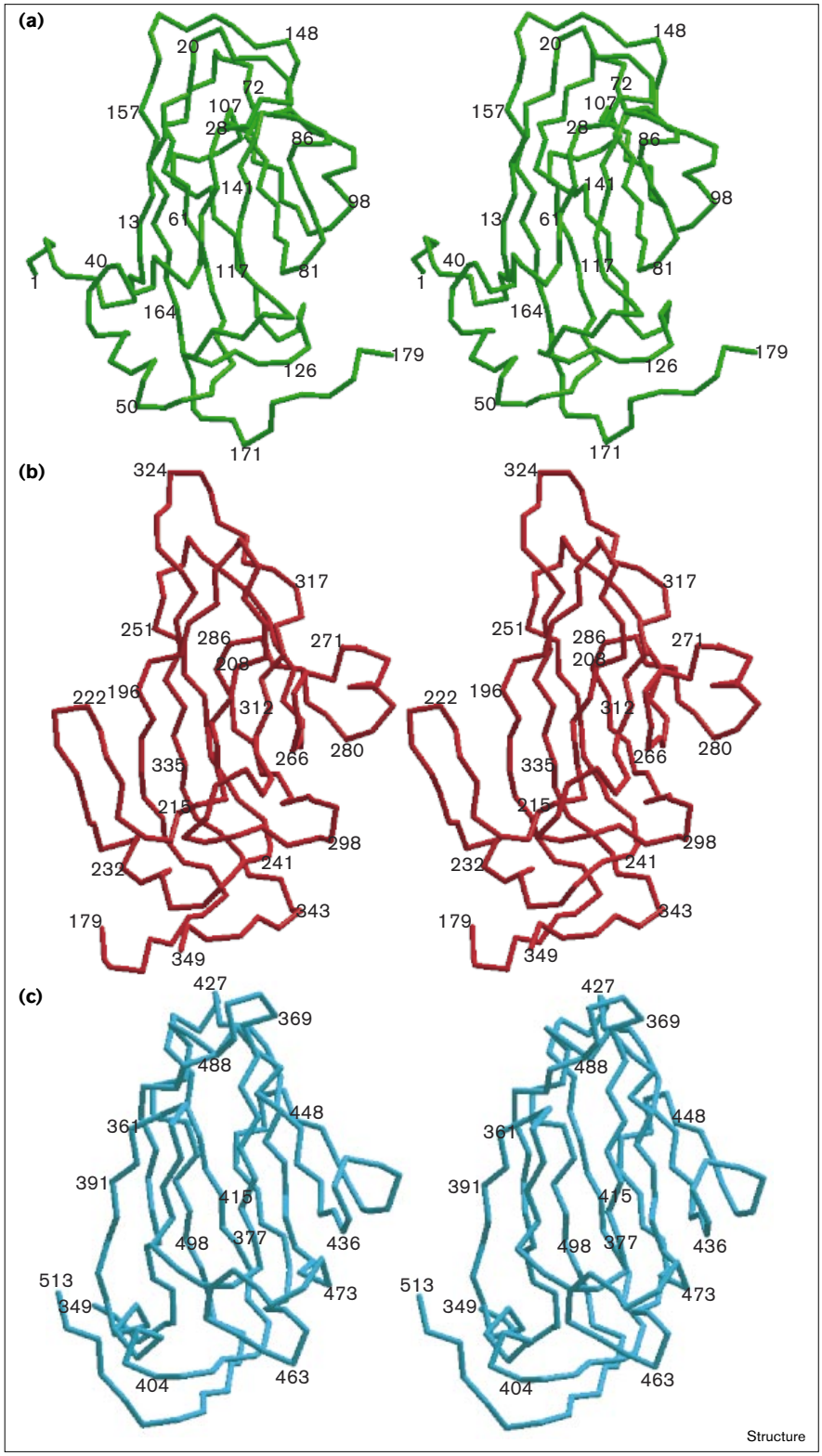
cleavage of the capsid polyprotein at a specific site in this loop by the virally encoded proteinase. The precleaved B–A domain-connecting polypeptide contains 30 to 36 residues in other comoviruses and the post cleaved N and C termini are separated from each other in BPMV and CPMV. While the new C termini of the B domains are in similar positions in BPMV and CPMV, the new N termini of the A domains have very different structures in these two comoviruses.

Quaternary interactions within the capsid protein

The quaternary interactions occurring within the TRSV subunit are defined by contacts at the A–B_5, C–B_5 and A–C interfaces (Figure 5a). The interactions at the C–B_5 and the A–B_5 interfaces are significantly more extensive than at the C–A interface. This suggests that the three domains C, B_5 and A in the TRSV capsid protein associate with each other successively as they are synthesized. Hydrophobic interactions play a dominant role in establishing the

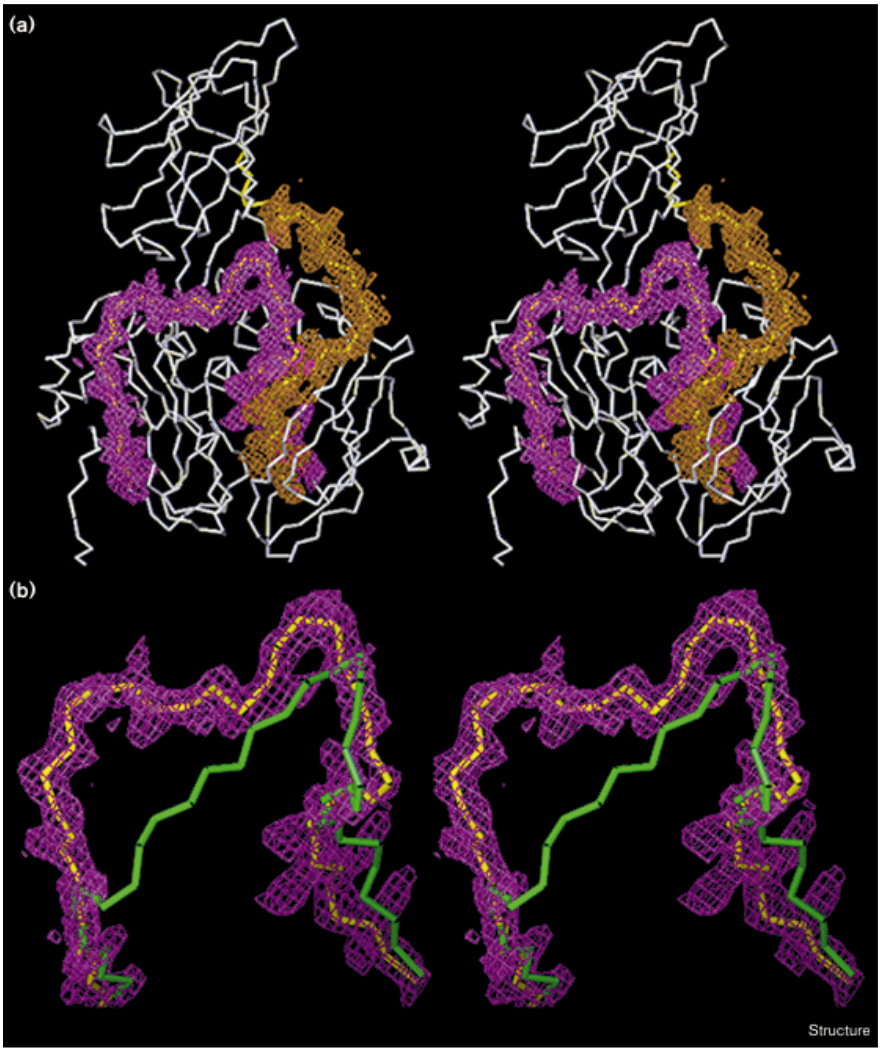
Figure 3

Stereo view Cα traces of the **(a)** C, **(b)** B and **(c)** A β-barrel domains in TRSV. The traces are all shown in the same orientation and every tenth residue is labeled.



quaternary organization of these domains in both TRSV and CPMV. There is no intertwining of the N and C termini in nepoviruses and comoviruses, while this is an important stabilizing interaction in picornaviruses [14,15,21,22].

Figure 4



Interdomain linking polypeptides in TRSV. **(a)** Stereo view of the Cα trace (white) of the capsid protein from inside the capsid with the electron density for the two domain-linking polypeptides. The pink density shows the link between the C and B domains and the brown density shows the link between the B and A domains. **(b)** An enlarged view of (a) showing the difference in the conformation of the C–B domain-connecting polypeptides in TRSV (yellow) and BPMV (green).

Quaternary interactions between the capsid proteins

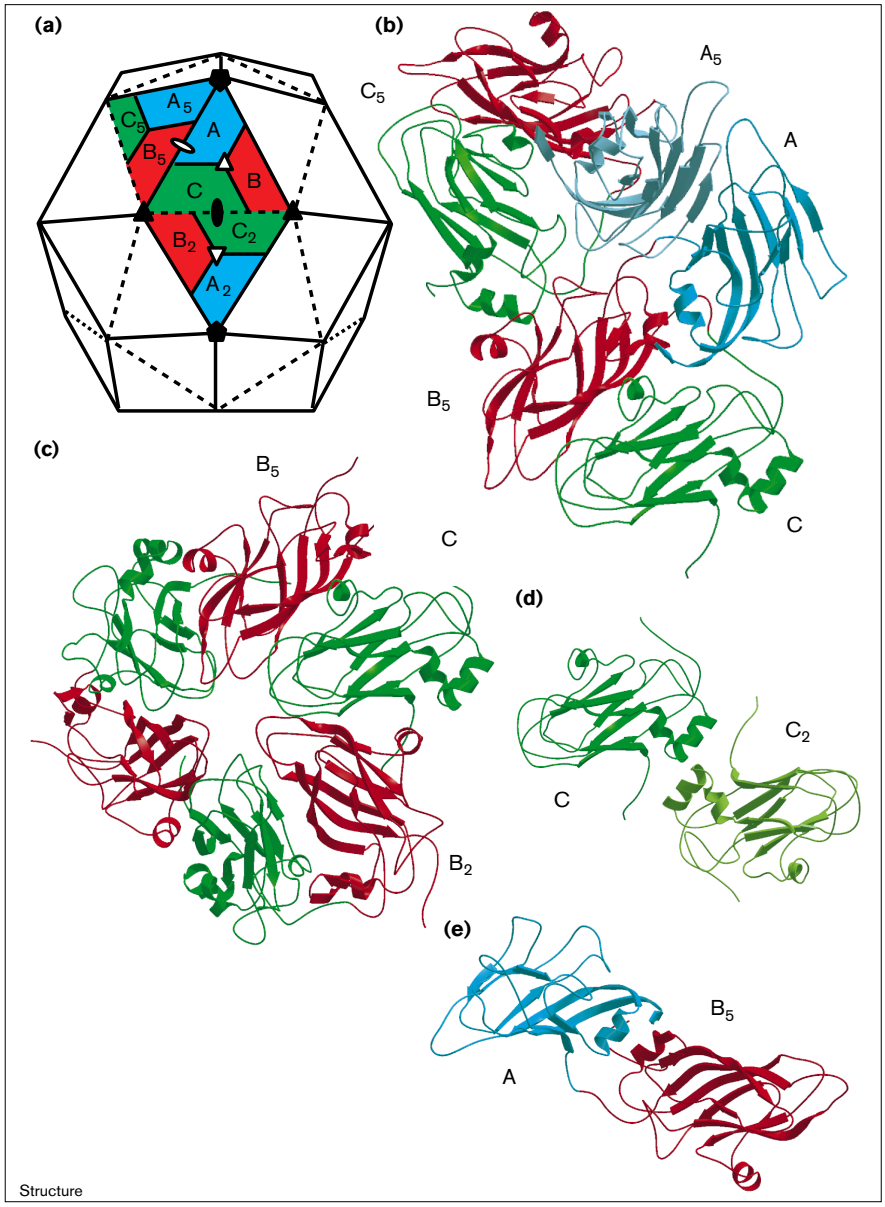
Quaternary interactions occur between pairs of capsid proteins related by icosahedral fivefold, threefold and twofold axes in the pseudo T = 3 capsid of TRSV.

Quaternary interactions at the fivefold contact consist of A–A_5, A–B and B–C domain interfaces (Figure 5b). The interactions at the A–A_5 and A–B interfaces are much more extensive than those at the B–C interface. Unlike comoviruses, where polar interactions play a dominant role in the association of the capsid proteins into a pentamer, hydrophobic interactions are more important in TRSV.

The only interdomain interface that defines the threefold contact between subunits in TRSV is C–B_2 (Figure 5c). Hydrophobic interactions and the N terminus play a prominent role in the association of the capsid proteins around the threefold axes. In contrast the N-terminal tail is not involved in this contact in CPMV. Protein–RNA interactions are involved in this contact in BPMV.

The twofold-related subunits contact each other only at the C–C_2 interface (Figure 5d). The interactions at this interface in TRSV are minimal. The N terminus of the C domain in comoviruses extends to the RNA-binding pocket in the twofold-related C_2 domain. The pseudo T = 3 nature of the TRSV capsid suggests comparing the A–B_5 contact with the C–C_2 contact as they are quasi-equivalent in the capsids formed by a single subunit type. The quasi or pseudo-equivalence breaks down dramatically in TRSV because the chemical interfaces are different at the two contacts even though similar structural elements are adjacent (Figure 5d,e). Interactions at the A–B_5 contact are much more extensive than at the C–C_2 contact.

Figure 5



Quaternary interactions between the TRSV capsid proteins. The C, B and A domains in the capsid are shown in green, red and cyan respectively. **(a)** A T = 3 icosahedral surface lattice represented as a rhombic triacontahedron, viewed approximately down an icosahedral twofold axis normal to the plane of the paper. The figure indicates the domains involved in fivefold, threefold and twofold contacts between capsid proteins. Closed pentagons, triangles and the closed oval symbolize the icosahedral fivefold, threefold and twofold axes, respectively. Open triangles and the open oval represent the quasi-threefold and twofold axes, respectively. The trapezoids labeled C, B_5 and A constitute three β-barrel domains that form one polypeptide chain going from the N to the C terminus. **(b)** Fivefold contact between capsid proteins viewed in the same orientation as (a). The three domain interfaces at this contact bury nearly 4600 Å^2 of the accessible surface area. **(c)** Hexamer of the B and C domains around the icosahedral threefold axes viewed in the same orientation as (a). Nearly 2570 Å^2 of the accessible surface area becomes buried at this threefold contact. **(d)** The icosahedral twofold contact at the C–C_2 domain interface (viewed down the icosahedral twofold axes) versus **(e)** the quasi-twofold contact at the A–B_5 domain interface (viewed approximately down the quasi-twofold axes). Only 606 Å^2 of the accessible surface area is buried at the C–C_2 interface compared to 2120 Å^2 at the A–B_5 interface.

The fivefold contacts between the capsid polyproteins are the strongest in the assembled capsid of TRSV and CPMV. This observation suggests that the TRSV capsid proteins assemble into pentamers rather than dimers or trimers, although experimental evidence for this has not been reported for nepoviruses or comoviruses. An assembly intermediate of picornaviruses has been identified which sediments at 14 S and is composed of a pentamer of three β sandwich subunits, VP1, VP2 and VP3, which correspond to A, C and B in TRSV [23,24].

Sequence alignment of nepoviruses

A multiple sequence alignment of the capsid proteins of 11 nepoviruses and five comoviruses [25] suggested that the two genera may share a common structure despite the low sequence identity within the nepovirus genus and between the two genera of viruses. A structural prediction for nepoviruses was attempted on the basis of the sequence alignment and the known comovirus structures. The predictions of the C and B domain structures were of reasonable quality when compared with the TRSV structure reported here. Significant differences between the predicted and actual structures were found in the A domain. A major readjustment of the structure prediction of the nepovirus cherry leaf roll virus (CLRV) was required. In Figure 6 a new sequence alignment of all the nepoviruses is shown with the reported TRSV structure as the reference.

Figure 6

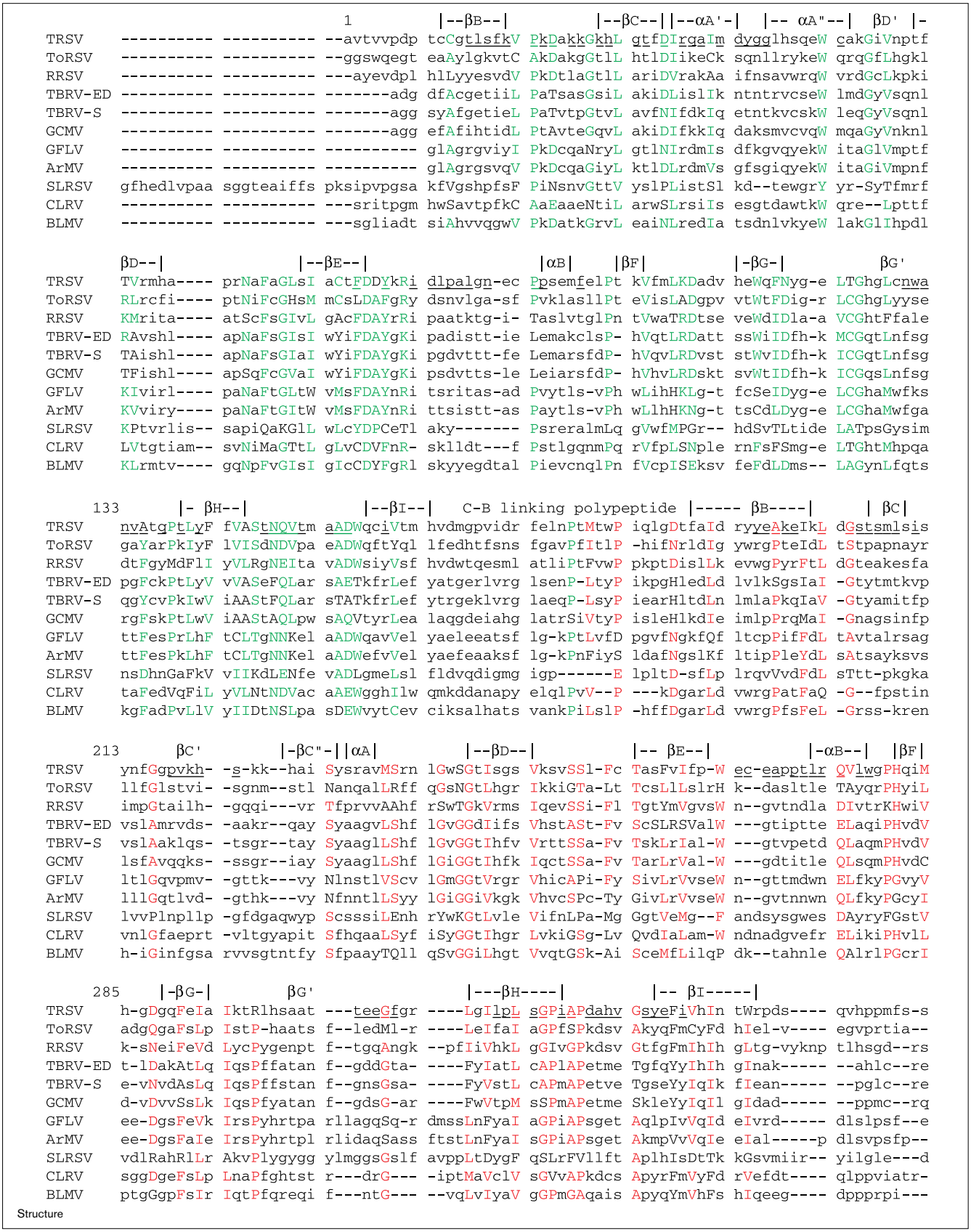
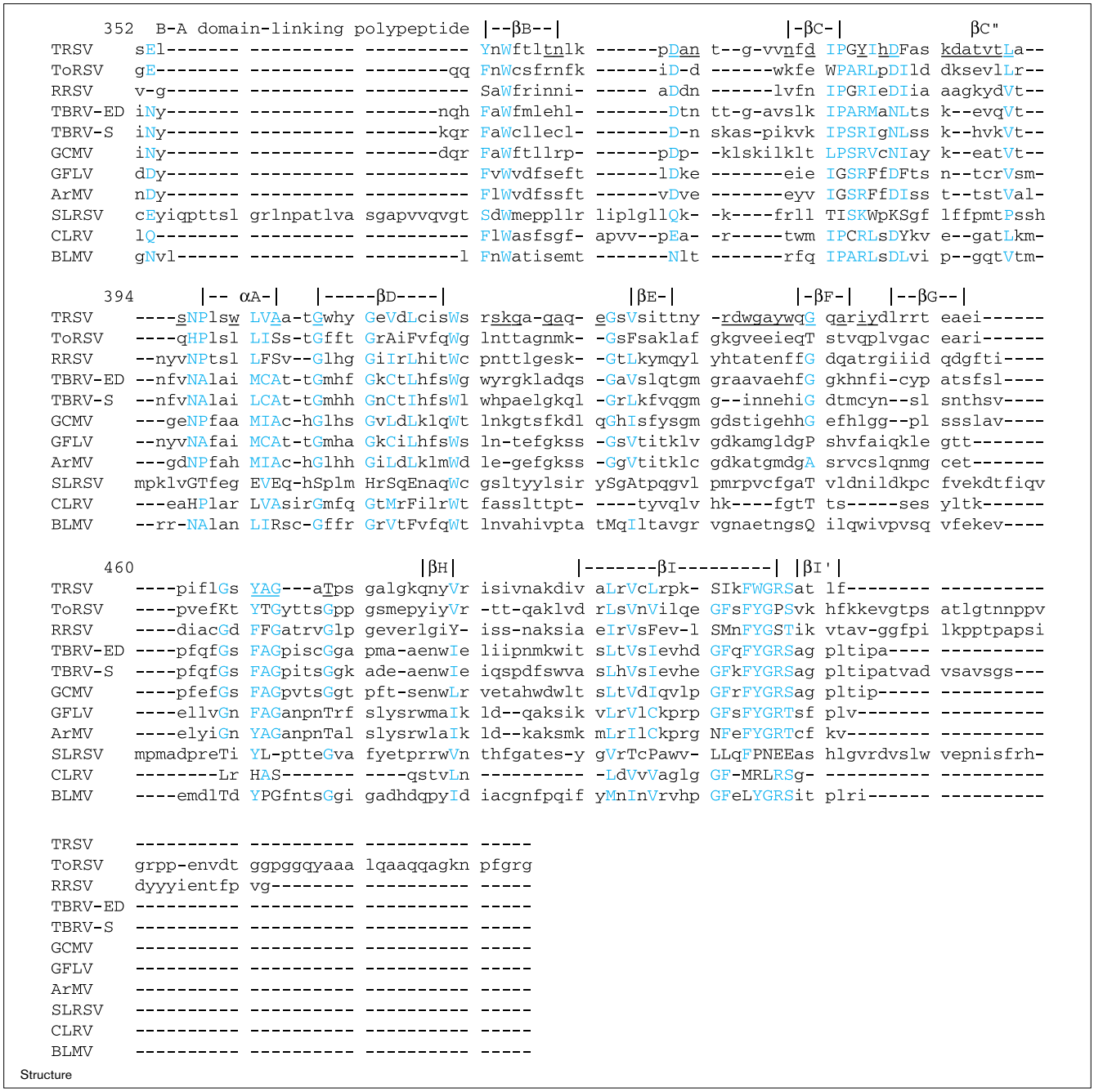

1 |--βB--| |--βC--||--αA'-| |-- αA"--| βD' |-
TRSV ---------- ---------- --avtvvpdp tcCgtlsfkV PkDakkGkhL gtfDIrgaIm dygglhsqeW cakGiVnptf
ToRSV ---------- ---------- --ggswqegt eaAylgkvtC AkDakgGtlL htlDIikeCk sqnllrykeW qrqGfLhgkl
RRSV ---------- ---------- ---ayevdpl hlLyyesvdV PkDtlaGtlL ariDVrakAa ifnsavwrqW vrdGcLkpki
TBRV-ED ---------- ---------- -------adg dfAcgetiiL PaTsasGsiL akiDLislIk ntntrvcseW lmdGyVsqnl
TBRV-S ---------- ---------- -------agg syAfgetieL PaTvtpGtvL avfNIfdkIq etntkvcskW leqGyVsqnl
GCMV ---------- ---------- -------agg efAfihtidL PtAvteGqvL akiDIfkkIq daksmvcvqW mqaGyVnknl
GFLV ---------- ---------- ---------- glAgrgviyI PkDcqaNryL gtlNIrdmIs dfkgvqyekW itaGlVmptf
ArMV ---------- ---------- ---------- glAgrgsvqV PkDcqaGiyL ktlDLrdmVs gfsgiqyekW itaGiVmpnf
SLRSV gfhedlvpaa sggteaiffs pksipvpgsa kfVgshpfsF PiNsnvGttV yslPListSl kd--tewgrY yr-SyTfmrf
CLRV ---------- ---------- ---sritpgm hwSavtpfkC AaEaaeNtiL arwSLrsiIs esgtdawtkW qre--Lpttf
BLMV ---------- ---------- ---sgliadt siAhvvqgwV PkDatkGrvL eaiNLredIa tsdnlvkyeW lakGlIhpdl
βD--| |--βE--| |αB| |βF| |-βG-| βG'
TRSV TVrmha---- prNaFaGLsI aCtFDDYkRi dlpalgn-ec Ppsemfelpt kVfmLKDadv heWqFNyg-e LTGhgLcnwa
ToRSV RLrcfi---- ptNiFcGHsM mCsLDAFgRy dsnvlga-sf Pvklasllpt eVisLADgpv vtWtFDig-r LCGhgLyyse
RRSV KMrita---- atScFsGIvL gAcFDAYrRi paatktg-i- TaslvtglPn tVwaTRDtse veWdIDla-a VCGhtFfale
TBRV-ED RAvshl---- apNaFsGIsI wYiFDAYgKi padistt-ie LemakclsP- hVqtLRDatt ssWiIDfh-k MCGqtLnfsg
TBRV-S TAishl---- apNaFsGIaI wYiFDAYgKi pgdvttt-fe LemarsfdP- hVqvLRDvst stWvIDfh-k ICGqtLnfsg
GCMV TFishl---- apSqFcGVaI wYiFDAYgKi psdvtts-le LeiarsfdP- hVhvLRDskt svWtIDfh-k ICGqsLnfsg
GFLV KIvirl---- paNaFtGLtW vMsFDAYnRi tsritas-ad Pvytls-vPh wLihHKLg-t fcSeIDyg-e LCGhaMwfks
ArMV KVviry---- paNaFtGItW vMsFDAYnRi ttsistt-as Paytls-vPh wLlhHKNg-t tsCdLDyg-e LCGhaMwfga
SLRSV KPtvrlis-- sapiQaKGlL wLcYDPCeTl aky------- PsreralmLq gVwfMPGr-- hdSvTLtide LATpsGysim
CLRV LVtgtiam-- svNiMaGTtL gLvCDVFnR- sklldt---f PstlgqnmPq rVfpLSNple rnFsFSmg-e LTGhtMhpqa
BLMV KLrmtv---- gqNpFvGIsI gIcCDYFgRl skyyegdtal PievcnqlPn fVcpISEksv feFdLDms-- LAGynLfqts
133 |- βH--| |--βI--| C-B linking polypeptide |----- βB----| | βC|
TRSV nvAtgPtLyF fVAStNQVtm aADWqciVtm hvdmgpvidr felnPtMtwP iqlgDtfaId ryyeAkeIkL dGstsmlsis
ToRSV gaYarPkIyF lVISdNDVpa eADWqftYql lfedhtfsns fgavPfItlP -hifNrldIg ywrgPteIdL tStpapnayr
RRSV dtFgyMdFlI yVLRgNEIta vADWsiyVsf hvdwtqesml atliPtFvwP pkptDislLk evwgPyrFtL dGteakesfa
TBRV-ED pgFckPtLyV vVASeFQLar sAETkfrLef yatgerlvrg lsenP-LtyP ikpgHledLd lvlkSgsIaI -Gtytmtkvp
TBRV-S qgYcvPkIwV iAAStFQLar sTATkfrLef ytrgeklvrg laeqP-LsyP iearHltdLn lmlaPkqIaV -Gtyamitfp
GCMV rgFskPtLwV iAAStAQLpw sAQVtyrLea laqgdeiahg latrSiVtyP isleHlkdIe imlpPrqMaI -Gnagsinfp
GFLV ttFesPrLhF tCLTgNNKel aADWqavVel yaeleeatsf lg-kPtLvfD pgvfNgkfQf ltcpPifFdL tAvtalrsag
ArMV ttFesPkLhF tCLTgNNKel aADWefvVel yaefeaaksf lg-kPnFiyS ldafNgslKf ltipPleYdL sAtsayksvs
SLRSV nsDhnGaFkV vIIKdLENfe vADLgmeLsl fldvqdigmg igp------E lpltD-sfLp lrqvVvdFdL sTtt-pkgka
CLRV taFedVqFiL yVLNtNDVac aAEWgghIlw qmkddanapy elqlPvV--P ---kDgarLd vwrgPatFaQ -G--fpstin
BLMV kgFadPvLlV yIIDtNSLpa sDEWvytCev ciksalhats vankPiLslP -hffDgarLd vwrgPfsFeL -Grss-kren
213 βC' |-βC"-||αA| |--βD--| |-- βE--| |-αB--| |βF|
TRSV ynfGgpvkh- -s-kk--hai SysravMSrn lGwSGtIsgs VksvSSl-Fc TasFvIfp-W ec-eapptlr QVlwgPHqiM
ToRSV llfGlstvi- -sgnm--stl NanqalLRff qGsNGtLhgr IkkiGTa-Lt TcsLlLslrH k--dasltle TAyqrPHyiL
RRSV impGtailh- -gqqi---vr TfprvvAAhf rSwTGkVrms IqevSSi-Fl TgtYmVgvsW n--gvtndla DIvtrKHwiV
TBRV-ED vslAmrvds- -aakr--qay SyaagvLShf lGvGGdIifs VhstASt-Fv ScSLRSValW ---gtiptte ELaqiPHvdV
TBRV-S vslAaklqs- -tsgr--tay SyaaglLShf lGvGGtIhfv VrttSSa-Fv TskLrIal-W ---gtvpetd QLaqmPHvdV
GCMV lsfAvqqks- -ssgr--iay SyaaglLShf lGiGGtIhfk IqctSSa-Fv TarLrVal-W ---gdtitle QLsqmPHvdC
GFLV ltlGqvpmv- -gttk---vy NlnstlVScv lGmGGtVrgr VhicAPi-Fy SivLrVvseW n--gttmdwn ELfkyPGvyV
ArMV lllGqtlvd- -gthk---vy NfnntlLSyy lGiGGiVkgk VhvcSPc-Ty GivLrVvseW n--gvtnnwn QLfkyPGcyI
SLRSV lvvPlnpllp -gfdgaqwyp ScsssiLEnh rYwKGtLvle VifnLPa-Mg GgtVeMg--F andsysgwes DAyryFGstV
CLRV vnlGfaeprt -vltgyapit SfhqaaLSyf iSyGGtIhgr LvkiGSg-Lv QvdIaLam-W ndnadgvefr ELikiPHvlL
BLMV h-iGinfgsa rvvsgtntfy SfpaayTQll qSvGGiLhgt VvqtGSk-Ai SceMfLilqP dk--tahnle QAlrlPGcrI
285 |-βG-| βG' |---βH----| |-- βI-----|
TRSV h-gDgqFeIa IktRlhsaat ---teeGfgr ----LgIlpL sGPiAPdahv GsyeFiVhIn tWrpds---- qvhppmfs-s
ToRSV adgQgaFsLp IstP-haats f--ledMl-r ----LeIfaI aGPfSPkdsv AkyqFmCyFd hIel-v---- egvprtia--
RRSV k-sNeiFeVd LycPygenpt f--tgqAngk --pfIiVhkL gGIvGPkdsv GtfgFmIhIh gLtg-vyknp tlhsgd--rs
TBRV-ED t-lDakAtLq IqsPffatan f--gddGta- ----FyIatL cAPlAPetme TgfqYyIhIh gInak----- --ahlc--re
TBRV-S e-vNvdAsLq IqsPffstan f--gnsGsa- ----FyVstL cAPmAPetve TgseYyIqIk fIean----- --pglc--re
GCMV d-vDvvSsLk IqsPfyatan f--gdsG-ar ----FwVtpM sSPmAPetme SkleYyIqIl gIdad----- --ppmc--rq
GFLV ee-DgsFeVk IrsPyhrtpa rllagqSq-r dmssLnFyaI aGPiAPsget AqlpIvVqId eIvrd----- dlslpsf--e
ArMV ee-DgsFaIe IrsPyhrtpl rlidaqSass ftstLnFyaI sGPiAPsget AkmpVvVqIe eIal-----p dlsvpsfp--
SLRSV vdlRahRlLr AkvPlygygg ylmggsGslf avppLtDygF qSLrFVllft AplhIsDtTk kGsvmiir-- yilgle---d
CLRV sggDgeFsLp LnaPfghtst r---drG--- -iptMaVclV sGVvAPkdcs ApyrFmVyFd rVefdt---- qlppviatr-
BLMV ptgGgpFsIr IqtPfqreqi f---ntG--- --vqLvIyaV gGPmGAqais ApyqYmVhFs hIqeeg---- dppprpi---

Figure 6 continued



Amino acid sequence alignment of eleven nepovirus capsid proteins. The secondary structure features and residue numbers for TRSV, based on the structure reported here, are indicated above its sequence. Residues that are conserved among at least seven nepoviruses in the C, B and A domains are shown in green, red and blue upper case letters, respectively. Underlined residues in the TRSV sequence occur on the outer surface of the capsid. Upper case black letters correspond to residues that are not conserved in columns where seven or more residues are conserved.

Most of the loops connecting the β strands in nepoviruses show some variations with at least one or two amino acid insertions or deletions. This is in contrast to comoviruses, where the insertions and deletions are confined to the BC and DE loops on the exterior surface of the capsid near the fivefold axes. The regions with the greatest variation in the three domains of TRSV include the loops present on the external surface of the capsid,

the two interdomain linking polypeptides and the N and C termini. A subset of nepoviruses comprising of strawberry latent ring spot virus (SLRSV), CLRV and blueberry leaf mottle virus (BLMV) vary the most relative to the other nepoviruses.

Most of the conserved hydrophobic residues are present in the strands forming the β-barrel with their sidechains contributing to the hydrophobic core in the three domains. Several externally exposed loops contain conserved residues interspersed with insertions and deletions.

Some of the sequence conservation may stabilize the quaternary structure of the three-domain subunit. The sequence motif Phe-Tyr-Gly-Arg-X-Ser (FYGRXS) present in the βI strand in the A domain interacts with a conserved stretch of residues (Pro177 to Pro182) in the C–B domain-connecting polypeptide. SLRSV is the only nepovirus that lacks both of these conserved regions. This suggests that the motif in the A domain in nepoviruses is important for maintaining the rigid structure of the C–B domain connector. Constraining the flexibility of this extended polypeptide may be important for the assembly of the nepovirus capsid proteins. As predicted previously [25], the sequence motifs FYGRXS in the A domain and Phe-Asp-Ala-Tyr-X-Arg (FDAYXR) present in the βE strand in the C domain occur adjacent to each other in the structure of TRSV. The motif present in the A domain is involved in stabilizing the quaternary structure of the domains within the capsid protein, while the motif in the C domain is accessible from outside the capsid and this region on the viral surface may have some functional significance in the nepoviruses.

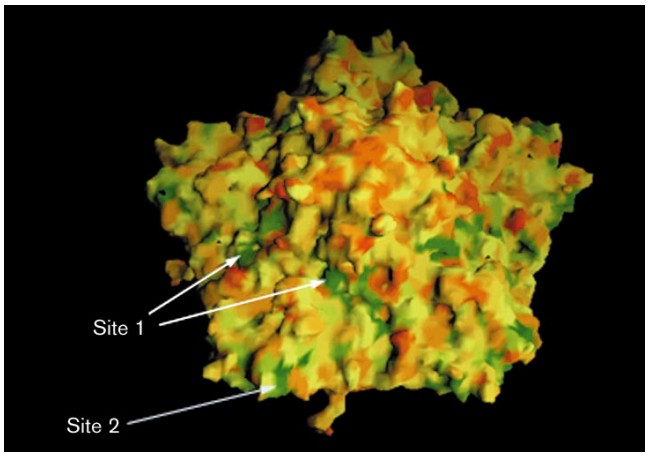
SLRSV is characterized as a nepovirus based on nematode transmission specificity. Unlike the other known nepoviruses, the processing of the capsid polyprotein of SLRSV is the same as for comoviruses (i.e. it is cleaved into two capsid protein subunits by a virally encoded protease at a specific site in the B–A domain-connecting polypeptide). When the length of the N terminus and the two domain-connecting polypeptides in TRSV are compared to other nepoviruses, only SLRSV has a longer N terminus, a shorter C–B domain connector and a much longer B–A domain connector. It is interesting that in these three trends SLRSV closely resembles comoviruses. The sequence Ser-Gly, recognized by the SLRSV protease, occurs at five different locations along the SLRSV capsid polyprotein. Three of these sites will occur on exposed loops based on the TRSV structure, but the protease cleaves only at the site in the B–A domain-linking polypeptide. This observation indicates that the presence of the cleavage site in an extended chain must be essential for cleavage by the SLRSV and comovirus proteases. A similar trend has been observed in picornaviruses based on their three-dimensional structures [26].

Surface amino acid conservation in nepoviruses

The variability of all the surface residues in the structure of TRSV was examined using the program GRASP [27] (Figure 7). A few residues are conserved among at least nine nepoviruses; these are confined to two regions, site 1 and site 2, on the TRSV surface. Site 1 is located in a minor depression at the junction of the A and C domains between the pseudo-twofold and the pseudo-threefold axes (Figures 5a and 7). Site 1 is comprised of residues from several regions: the EF loop in the C domain (including Asp79 present in the FDAYXR sequence motif); the βC–βD insertion in the A domain; and the GH loop in all the three domains. Site 2 occurs on a minor surface protrusion, close to the icosahedral threefold axes, where the conserved residues on the BC and HI loops in the B and C domains form a ring.

The nematode species that are required to transmit ten of the 11 nepoviruses shown in the sequence alignment in Figure 6 are known. The nepoviruses, with the exception of BLMV, fall into four groups based on the nematode species specificity: TRSV and tomato ringspot virus (ToRSV); raspberry ringspot virus (RRSV) and tomato blackring virus (TBRV); grapevine chrome mosaic virus (GCMV) and grapevine fourleaf virus (GFLV); and arabis mosaic virus (ArMV), SLRSV and CLRV. Regions

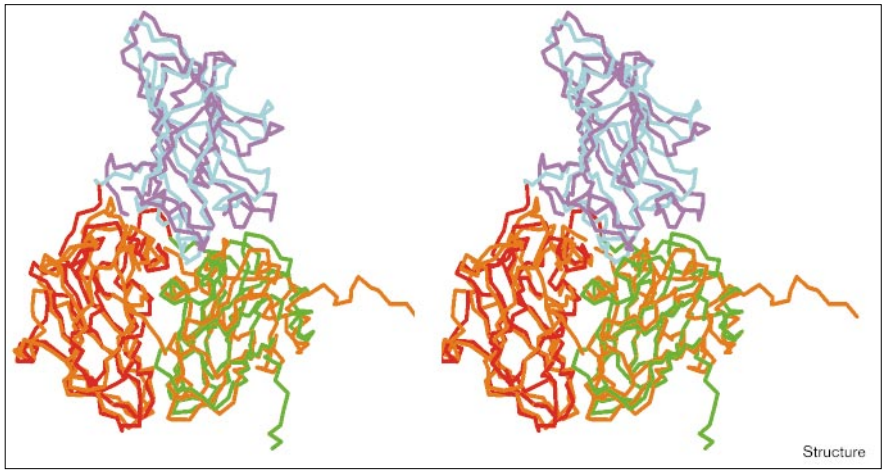
Figure 7



Depth-cued view of the external molecular surface of a pentamer of TRSV subunits. The figure shows the variability of the surface residues among nepoviruses. The surface residues in TRSV are color-coded according to their conservation number, which denotes the number of nepoviruses in Figure 6 containing a similar residue at that position. The color-code ranges from red (conservation number 1) through yellow (conservation number 6) to green (conservation number 11). Two regions, site 1 and site 2, containing some residues conserved among at least eight nepoviruses are indicated. (The figure was generated using GRASP [27].)

Figure 8

Stereo view Cα traces of the capsid proteins of TRSV superpositioned on the capsid proteins of BPMV. The superposition is viewed from outside the capsid. The C, B and A domains in TRSV are shown in green, red and cyan, respectively; the C and B domains of BPMV are shown in orange and the A domain is in purple.



on the viral surface that correlate with the nematode transmission specificity of the four groups of nepoviruses could not be identified on the basis of the surface amino acid variability. Nepoviruses transmitted by the same nematode species do not share regions of similar amino acid composition on the viral surface. This also suggests that the conserved site 1, containing Asp79, on the viral surface is not directly involved in determining vector specificity. This site may serve as a binding site for some unknown host factor or viral protein. Further mutation studies are required to identify the functional significance of site 1 in the nepovirus life cycle. The significance of the ring of conserved residues around the threefold axes (site 2) is also not clear. This region in CPMV constitutes a major antigenic site, as demonstrated by the cryo-EM structure of a complex formed between an Fab fragment of a monoclonal antibody raised to CPMV and the virus [28,29].

Structural similarity to comoviruses

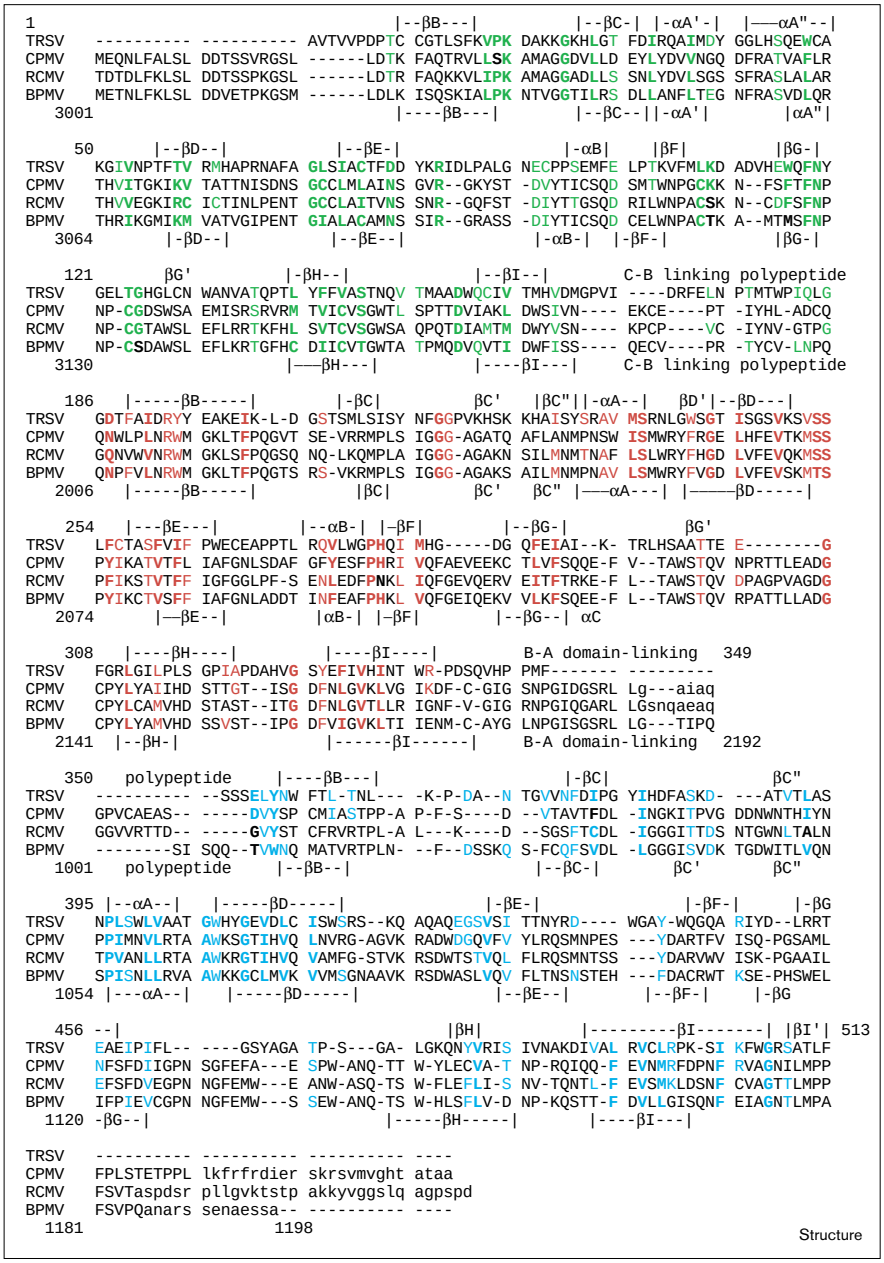
Structural superposition of the capsid proteins of TRSV and the comovirus BPMV was performed using the least squares superposition option in the graphics package O [30]. Based on the hydrogen-bonding pattern of the eight β strands forming the β barrel, the corresponding stretches of residues in the β strands of TRSV and BPMV were identified for superposition. The regions with the greatest deviation in distance between equivalenced Cα positions in TRSV and BPMV include the external surface loops, the two domain-connecting polypeptides and the N and C termini (Figure 8).

The sequence alignment of the C, B and A domains of TRSV and three comoviruses (BPMV, CPMV and RCMV), based on structure superposition, is shown in Figure 9. The three comoviruses are structurally very similar (W Wikoff, T Lin and JEJ, unpublished data). Most of the β strands forming the core β barrel in BPMV are longer than the corresponding strands in TRSV. About 58 of the 513 residues in TRSV are conserved among at least seven nepoviruses and the comoviruses. Most of these residues are hydrophobic and occur on the β strands and α helices with their sidechain buried in the hydrophobic core. Among the non-glycine and non-proline residues that are conserved, it is intriguing that the residues Asp79 and Arg83, present in the FDAYXR sequence motif in the C domain of TRSV, are also conserved as Asn3093 and Arg3097 in comoviruses. This indicates that the function of the conserved site 1 in nepoviruses may also be shared by comoviruses.

The statistics for the structural superpositions (Table 1) support the proposed evolutionary connections between the capsid polyproteins of the three groups in the picornavirus superfamily. Among the three equivalent domains in TRSV and BPMV, the B domains display the most structural similarity while the A domains exhibit the least (Table 1). The superposition of the three different domains within TRSV or BPMV onto each other indicates that the B and C domains are structurally the closest in both the viruses (Table 1). A similar trend has been observed among picornaviruses [21,31]. The equivalent domains in TRSV and BPMV are structurally closer than the three different domains within TRSV. Furthermore, the structures of the three β-barrel domains in comoviruses have diverged from each other slightly more than the domains in TRSV. The three β-barrel domains within a picornavirus [21], on the other hand, have clearly diverged much more than those in either comoviruses or TRSV. The comparisons suggest that the three domains in the nepovirus capsid protein diverged significantly before the comovirus polyprotein, with its cleavage site in the B–A domain-linking polypeptide, evolved.

Figure 9

Amino acid sequence alignment of the capsid proteins of TRSV and three comoviruses (CPMV, RCMV and BPMV) based on the structural superposition. The secondary structure elements and sequence numbers of TRSV are given on the top line. Residues conserved among at least seven nepoviruses (shown in Figure 6) and the three comoviruses in the C, B and A domains are shown in green, red and blue bold upper case letters, respectively. Residues that are conserved among TRSV and any of the three comoviruses in the C, B and A domains are shown in upper case and in green, red and blue, respectively. Residues in lower case letters were not observed in the structures of comoviruses.



Electron density in the RNA region

The structure of BPMV revealed the interaction of a trefoil-shaped cluster of RNA with the cleft between the covalently linked B and C domains in the capsid protein near the icosahedral threefold axes. Based on the observed protein–RNA interactions, it was proposed that the RNA in comoviruses may be recognizing a preformed icosahedral binding site in the assembling capsid [7]. Except for a piece of density that has been modeled as a nucleotide near the VP2 subunit, ordered RNA density has not been observed in the mature picornavirus structures [14,15,31]. The crystal structure of the P1 poliovirus empty capsid, however, revealed a trefoil-shaped depression on the inner surface of the capsid that was strikingly similar to the RNA-binding site in BPMV [32].

The inner surface architecture of the cleft between the B and C domains in TRSV is quite different from the RNA-binding pocket in BPMV due to differences in the conformations of the C–B domain-connecting polypeptide and the N terminus (Figure 8). There is no evidence for the presence of density corresponding to well-ordered RNA in this cleft in TRSV.

Table 1

Root mean square deviation (σ) in Cα displacements (Å) for residues in the core β strands and loops connecting the strands in the pairs of superimposed domains.

Structural comparison*	Superimposed pairs		Root mean square deviation	Number of equivalenced residues†	Percentage equivalence (%)†‡	
	Domain 1	Domain 2			Domain 1	Domain 2
TRSV versus BPMV	TRSV C	BPMV C	2.0	123 (129)	84 (72)	80 (71)
	TRSV B	BPMV B	1.4	127 (137)	83 (81)	78 (71)
	TRSV A	BPMV A	2.2	119 (122)	75 (74)	70 (66)
TRSV versus TRSV	TRSV A	TRSV B	2.3	81	58	57
	TRSV B	TRSV C	1.8	78	56	52
	TRSV C	TRSV A	2.2	78	50	55
BPMV versus BPMV	BPMV A	BPMV B	2.3	73	46	46
	BPMV B	BPMV C	2.0	82	54	57
	BPMV C	BPMV A	2.2	79	54	51

*TRSV = tobacco ring spot virus; BPMV = beanpod mottle virus. †The numbers given in parentheses refer to the superposition of all the residues (including the N and C termini) in the pair of domains. ‡The percentage equivalence shown in the last column is given by [number of equivalenced residues/total number of residues in the β-barrel domain] × 100.

Biological implications

Tobacco ringspot virus (TRSV) is a member of the nepovirus genus of icosahedral RNA plant viruses. Nepoviruses cause diseases in fruit crops, have a wide host range and are transmitted by adult and larval stages of soil-inhabiting eel worms (nematodes). Plant virus families are classified into superfamilies based on genetic similarities to animal viruses. Nepoviruses, comoviruses and picornaviruses constitute members of the picornavirus superfamily that have icosahedral capsids. The capsids of the icosahedral picorna-like viruses are composed of 60 copies of three capsid protein subunits (each with a similar β-barrel fold), which occupy positions in the surface lattice comparable to those of 180 copies of a single capsid protein subunit in a T = 3 virus capsid. The three different subunits are synthesized as a polyprotein, which is subsequently cleaved by a viral proteinase.

Previous structural studies of members of the picornavirus superfamily have suggested that picornaviruses evolved from a T = 3 virus by triplication of the gene encoding a β-barrel domain, followed by the independent evolution of the three β-barrel domains and the development of cleavage sites in the interdomain-linking regions in the capsid polyprotein. The picornavirus capsid polyprotein is cleaved at two sites to yield three subunits. The comovirus polyprotein, on the other hand, is cleaved at only one site to yield a large subunit with two β-barrel domains and a small subunit with one β-barrel domain. The capsids of nepoviruses are composed of 60 copies of a single capsid protein composed of three β-barrel domains and lack any cleavage sites. As such, the nepoviruses may represent an early stage in the evolution of picornavirus capsids. Structural studies on TRSV were initiated to investigate the proposed evolutionary connection among the three groups in the picornavirus superfamily.

The crystal structure of TRSV has revealed that the capsid protein subunit is folded into three β-barrel domains that are covalently linked together by extended polypeptides as 'beads on a string'. The observed order of connectivity of the three domains from the N to the C terminus in the capsid protein subunit is consistent with the proposed connectivity for the precleaved comovirus and picornavirus capsid polyproteins. The three different domains within TRSV and comoviruses are more closely related at the structural level than the three domains within picornaviruses. The results of structural comparison and a sequence alignment among nepoviruses and comoviruses support the notion that the capsid polyproteins of nepoviruses, comoviruses and picornavirus have evolved from a common ancestor via divergent evolution. The first structure of a nepovirus also provides a snapshot of how the development of cleavage sites in the capsid polyprotein yields new flexible N and C termini which can intertwine and stabilize the quaternary interactions between the subunits during capsid assembly.

Nepoviruses exhibit diversity in the capsid protein sequence and nematode species specificity during transmission. Previous studies have indicated that retention of the nepovirus capsid in the gut lining of the nematode vector dictates virus vector specificity. Nepoviruses transmitted by the same nematode species do not share regions on the viral surface with similar amino acid composition. A few residues on the TRSV capsid surface, clustered in two regions, are conserved among nepoviruses. Further studies are needed to understand the role of these regions as potential binding sites for host factors or viral proteins involved in either cell-to-cell movement or some other function shared by all nepoviruses.

Materials and methods

Crystals, data collection and processing

TRSV was propagated in tobacco plants, purified by a modified procedure reported by [33] and crystallized at room temperature using 2–3% polyethylene glycol (PEG) MW 3350 as precipitant. The virus sample used for crystallization contains empty capsids, particles with one molecule of RNA1 or RNA2 and particles containing two molecules of RNA2. X-ray diffraction data were collected at the Cornell High Energy Synchrotron Source (CHESS) on the F1 beam line ($\lambda = 0.91$ Å, F = 300mm) employing Fuji image plates. The TRSV particles crystallize in the monoclinic C2 space group (a = 407.1 Å, b = 399.7 Å, c = 285.9 Å and $\beta = 129.1°$) with half a virion in the crystallographic asymmetric unit giving 30-fold noncrystallographic redundancy. The crystals diffract X-rays to 3.3 Å and the diffraction data were only 22% complete to 3.5 Å resolution with an R_{merge} of 10.3%. The orientation of the virus particle in the unit cell was determined by rotation function analysis [34]. It was not possible to generate a more complete data set at this time because of problems associated with obtaining crystals of TRSV that diffract X-rays to high resolution consistently.

Structure determination and refinement

The structure of TRSV has been determined to 3.5 Å resolution using only the 22% complete data. A data set composed of 22% of the crystallographically unique reflections corresponds to nearly sevenfold noncrystallographic redundancy (100% of the data corresponds to 30 icosahedral asymmetric units). The power of real-space electron-density averaging allows this redundancy to be exploited explaining the success in solving the structure [35]. The polyalanine model for the three β-barrel domains in CPMV [7] were treated as three independent rigid bodies and subjected to XPLOR rigid-body minimization [36] using all the 4045 observed unique reflections for TRSV in the resolution range of 20 to 10 Å. After 40 cycles, the XPLOR R factor dropped from 56.5% to 49.5%. The XPLOR R factor $= [\Sigma|F_o - F_c|/\Sigma F_o] \times 100$, where F_o and F_c are the experimentally observed and atomic model based calculated structure-factor amplitudes, respectively. The phases calculated from this model were combined with the observed structure-factor amplitudes for TRSV in the 20 to 10 Å resolution range to generate an electron-density map.

Real-space electron-density map averaging and solvent flattening were performed using the programs in the RAVE and CCP4 suites [37]. An atomic mask, based on the coordinates of the initial phasing model using MAMA [38], was used for averaging. After 15 cycles of averaging at 10 Å, the overall CC and R_{MR} were 85.7% and 24.7%, respectively. $CC = \Sigma\ (\langle F_o \rangle - F_o)\ (\langle F_c \rangle - F_c)/[\Sigma(\langle F_o \rangle - F_o)^2 \Sigma\ (\langle F_o \rangle - F_o)^2]^{1/2}$ and $R_{MR} = [\Sigma|F_o - F_c|/\Sigma F_o] \times 100$, where F_o is as defined above and F_c is the structure-factor amplitude obtained by inverse Fourier transformation of the modified electron-density map. Phase extension from 10 Å to 3.5 Å resolution was carried out in 42 extension steps with a step size of one reciprocal lattice point and involved 630 cycles of phase refinement by averaging. Using the package O [30] a structural model containing all 513 residues present in the capsid protein of TRSV [5] was built to the 3.5 Å resolution electron-density map. Of the residues present in the solvent accessible surface loops on the capsid, 23 of these, including Lys17, Lys23, Arg83, Asn91, Glu92, Glu100, Lys109, Asp110, Asp112, Asn130, Arg172, Glu174, Lys201, Leu202, Asp203, Arg274, Glu275, Glu330, Arg339, Lys386, Lys421, Trp443 and Arg482 have ill-determined sidechain density due to higher sidechain mobility. A new atomic mask based on the 3.5 Å structural model of TRSV and an inner radial cut-off of 70 Å was generated. Ten cycles of averaging at 3.5 Å were carried out using the new mask and phases from the structural model, in order to locate any regions of ordered RNA that may be interacting with the protein shell.

The structural model was subjected to conjugate gradient energy minimization using XPLOR [36], followed by manual rebuilding. After 300 cycles of positional refinement using all the reflections between 8 Å to 3.5 Å, the R factor decreased from 31.4% to 26.9%. Further refinement of the structure has not been carried out due to the incompleteness of the diffraction data. The present structure of TRSV has an R factor of 28.5% for all the measured reflections between 20 Å and 3.5 Å resolution. The root mean square (rms) deviations in bond distances and bond angles are 0.011 Å and 2°, respectively.

Buried surface area calculation

The buried surface areas reported in Figures 5b–5d are based on the accessible contact surface areas calculated for the various domains and domain interfaces in the capsid, using a probe radius of 1.4 Å in the program ACCESS [39]. For instance, the difference between the accessible surface area for one intact subunit (containing three domains) and the sum of the accessible surface areas for the A, B_5 and C domains as separate modules will yield the buried surface area for the single subunit.

Accession numbers

The coordinates for the TRSV capsid protein will be deposited in the Brookhaven Protein Data Bank.

Acknowledgements

We wish to thank Abhinav Kumar for discussions on the methodology of structure solution, William Wikoff for useful discussions on the TRSV structure and Vijay Reddy for the use of his program to analyze the quaternary interactions in the capsid. We acknowledge CHESS for access to the F1 beamline and Tianwei Lin, Tim Schmidt, Bonnie Mckinney and the staff at CHESS and MacCHESS for their assistance in data collection. This work was supported by a grant from the National Institutes of Health R01GM 54076.

References

1. Stuart, D.I. (1993). Viruses. *Curr. Opin. Struct. Biol.* **3,** 167-174.
2. Stacie-Smith. (1985). Tobacco ringspot virus. In *CMI/AAB Descriptions of Plant Viruses.* **19**, No. 309, Association of Applied Biologists, National Vegetable Research Station, Wellesbourne, Warwick, UK.
3. Harrison, B.D. & Murant, A.F. (1977). Nepovirus group. In *CMI/AAB Descriptions of Plant Viruses*. **11**, No.185, Commonwealth Mycological Institute and the Association of Applied Biologists, Ferry Lane, Kew, Surrey, UK.
4. Mayo, M.A., Murant, A.F. & Harrison, B.D. (1971). New evidence on the structure of nepovirus. *J. Gen. Virol.* **12,** 175-178.
5. Buckley, B., Silva, S. & Singh, S. (1993). Nucleotide sequence and *in vitro* expression of the capsid protein gene of tobacco ringspot virus. *Virus Res.* **30**, 335-349.
6. Goldbach, R. & Wellink, J. (1988). Evolution of plus-strand RNA viruses. *Intervirology* **29,** 260-267.
7. Chen, Z.G., *et al*., Johnson, J.E. (1989). Protein–RNA interactions in an icosahedral virus at 3.0 Å resolution. *Science* **245**, 154-159.
8. Hosur, M.V., *et al.,* Rueckert, R.R. (1987). Structure of an insect virus at 3.0 Å resolution. *Proteins* **2**, 167-176.
9. Harrison, B.D., Murant, A.F., Mayo, M.A. & Roberts, I.M. (1974). Distribution of determinants for symptom production, host range and nematode transmissibility between the two RNA components of raspberry ringspot virus. *J. Gen. Virol.* **22**, 233-247.
10. Harrison, B.D., Robertson, W.M. & Taylor, C.E. (1974). Specificity of retention and transmission of viruses by nematodes. *J. Nematol.* **6**, 155-164.
11. Wieczorek, A. & Sanfacon, H. (1993). Characterization and subcellular localization of tomato ringspot nepovirus putative movement protein. *Virology* **194**, 734-742.
12. Diener, T.O. & Schneider, I.R. (1966). The two components of tobacco ringspot virus nucleic acid: origin and properties. *Virology* **29**, 100-105.
13. Murant, A.F., Taylor, M., Duncan, G.H. & Raschké, J.H. (1981). Improved estimates of molecular weight of plant virus RNA by gel electrophoresis and electron microscopy after denaturation with glyoxal. *J. Gen. Virol.* **53**, 321-332.
14. Rossmann, M.G., *et al*., Mosser, A.G. (1985). Structure of a human common cold virus and functional relationship to other picornaviruses. *Nature* **317**, 145-153.
15. Hogle, J.M., Chow, M. & Filman, D.J. (1985). Three-dimensional structure of poliovirus at 2.9 Å resolution. *Science* **229**, 1358-1365.

16. Singh, S., Rothnagel, R., Prasad, B.V. & Buckley, B. (1995). Expression of tobacco ringspot virus capsid protein and satellite RNA in insect cells and three-dimensional structure of tobacco ringspot virus-like particles. *Virology* **213**, 472-481.
17. Chen, Z., Stauffacher, C.V. & Johnson, J.E. (1990). Capsid structure and RNA packaging in comoviruses. *Sem. in Virol.* **1**, 453-466.
18. Caspar, D.L.D. & Klug, A. (1962). Physical principles in the construction of regular viruses. *Cold Spring Harbor Symp. Quant. Biol.* **27**, 1-24.
19. Chelvanayagam, G., Heringa, J. & Argos, P. (1992). Anatomy and evolution of proteins displaying the viral capsid jellyroll topology. *J. Mol. Biol.* **228**, 220-242.
20. Rossmann, M.G. & Johnson, J.E. (1989). Icosahedral RNA virus structure. *Annu. Rev. Biochem.* **58**, 533-573.
21. Luo, M., *et al.*, Palmenberg, A.C. (1989). The atomic structure of Mengo virus at 3.0 Å resolution. *Science* **235**, 182-191.
22. Acharya, R., Fry, E., Stuart, D., Fox, G., Rowlands, D. & Brown, F. (1989). The three-dimensional structure of foot-and-mouth disease virus at 2.9 Å resolution. *Nature* **337**, 709-716.
23. Rueckert, R.R., Dunker, A.K. & Stoltzfus, C.M. (1969). The structure of Maus-Elberfeld virus: a model. *Proc. Natl. Acad. Sci. USA* **62**, 912-919.
24. Rueckert, R.R. (1976). On the structure and morphogenesis of picornaviruses. In *Comprehensive Virology*. (Fraenkel-Conrat, H. & Wagner, R.R. eds.), pp. 131-213, **6**, Plenum Press, New York, NY.
25. Le Gall, O., Candresse, T. & Dunez, J. (1995). A multiple alignment of the capsid protein sequences of nepoviruses and comoviruses suggests a common structure. *Arch. Virol.* **140**, 2041-2053.
26. Arnold, E., *et al.*, Wimmer, E. (1987). Implications of the picornavirus capsid structure for polyprotein processing. *Proc. Natl. Acad. Sci. USA* **84**, 21-25.
27. Nicholls, A., Sharp, K. & Honig, B. (1991). Protein folding and association – insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins* **11**, 281-296.
28. Porta, C., Spall, V.E., Loveland, J., Johnson, J.E., Barker, P.J. & Lomonossoff, G.P. (1994). Development of cowpea mosaic virus as a high-yielding system for the presentation of foreign peptides. *Virology* **202**, 949-955.
29. Wang, G.J., Porta, C., Chen, Z.G., Baker, T.S. & Johnson, J.E. (1992). Identification of a Fab interaction footprint site on an icosahedral virus by cryoelectron microscopy and X-ray crystallography. *Nature* **355**, 275-278.
30. Jones, T.A., Cowan, S., Zou, J.Y. & Kjeldgaard, M. (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta. Cryst. A* **47**, 110-119.
31. Muckelbauer, J.K., *et al.*, Rossmann, M.G. (1995). The structure of coxsackievirus B3 at 3.5 Å resolution. *Structure* **3**, 653-667.
32. Basavappa, R., Syed, R., Flore, O., Icenogle, J.P., Filman, D.J. & Hogle, J.M. (1994). Role and mechanism of the maturation cleavage of VP0 in poliovirus assembly: structure of the empty capsid assembly intermediate at 2.9 Å resolution. *Protein Sci.* **3**, 1651-1669.
33. Rezaian, M.A. & Francki, R.I.B. (1973). Replication of tobacco ringspot virus. I. Detection of a low molecular weight double-stranded RNA from infected plants. *Virology* **56**, 238-249.
34. Chandrasekar, V., Munshi, S. & Johnson, J.E. (1997). Crystallization and preliminary X-ray analysis of tobacco ringspot virus. *Acta Cryst. D* **53**, 125-128.
35. Rossmann, M.G. (1990). The molecular replacement method. *Acta Cryst A* **46**, 73-82.
36. Brünger, A.T. (1993). X-PLOR Version 3.1 Manual. Yale University, New Haven, CT.
37. Collaborative Computational Project, Number 4. (1994). The CCP4 suite: programs for protein crystallography. *Acta Cryst. D* **50**, 760-763.
38. Kleywegt, G.J. & Jones, T.A. (1994). Halloween – masks and bones. In *From First Map to Final Model*. (Bailey, S., Hubbard, R. & Waller, D. eds.), pp. 59-66, SERC Daresbury Laboratory, Warrington, UK.
39. Lee, B. & Richards, F.M. (1971). The interpretation of protein structures: estimation of static accessibility. *J. Mol. Biol.* **55**, 379-400.
40. Kraulis, P.J. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Cryst.* **24**, 946-950.