Proc. Natl. Acad. Sci. USA
Vol. 93, pp. 12150–12154, October 1996
Biochemistry

Three-dimensional model of the potyviral genome-linked protein

DANUTA PŁOCHOCKA, MAREK WEŁNICKI, PIOTR ZIELENKIEWICZ*, AND WŁODZIMIERZ OSTOJA-ZAGÓRSKI

Institute of Biochemistry and Biophysics, Polish Academy of Sciences, 5A Pawińskiego Street, 02-106 Warsaw, Poland

Communicated by Michael Sela, The Weizmann Institute of Science, Rehovot, Israel, August 1, 1996 (received for review January 28, 1996)

ABSTRACT **The full sequence of the genome-linked viral protein (VPg) cistron located in the central part of potato virus Y (common strain) genome has been identified. The VPg gene codes for a protein of 188 amino acids, with significant homology to other known potyviral VPg polypeptides. A three-dimensional model structure of VPg is proposed on the basis of similarity of hydrophobic–hydrophilic residue distribution to the sequence of malate dehydrogenase of known crystal structure. The 5′ end of the viral RNA can be fitted to interact with the protein through the exposed hydroxyl group of Tyr-64, in agreement with experimental data. The complex favors stereochemically the formation of a phosphodiester bond [5′-(O^4-tyrosylphospho)adenylate] typical for representatives of picornavirus-like viruses. The chemical mechanisms of viral RNA binding to VPg are discussed on the basis of the model structure of protein–RNA complex.**

The genome-linked viral protein (VPg) covalently bound to a 5′-terminal nucleotide of the genome has been identified for several animal and plant viruses, with both DNA and RNA genomes (1). Among positive-sense RNA viruses five families—luteoviruses, picornaviruses, comoviruses, nepoviruses, and potyviruses—are known to contain 5′-terminal VPg (1). The latter four families also share other similarities in RNA structure [3′-terminal poly(A) tail], genome organization (the presence of a characteristic cluster of replication genes), and strategy of expression (expression of viral RNA as a single polyprotein undergoing self-processing), and they are therefore often arranged in a supergroup of picornavirus-like viruses (2). In the potyvirus family, the area of our interest, genomic RNAs carry at their 5′ end a VPg of molecular mass 22–24 kDa (3–5). As in polioviruses (6, 7), potyviral VPg binds to the 5′-phosphate through the hydroxyl group of a tyrosine (8). Sequencing the potato virus Y, common strain (PVY^0), we detected a part of the viral open reading frame covering the putative VPg cistron, which translates into a protein of molecular mass equal to 25 kDa. In a study of the sequences of phenotypically different PVY strains, the PVY variants were recently classified into three phylogenetic groups (9). Strikingly, the first 20 5′-proximal RNA bases, covering the fragment which could be implicated in interaction with VPg, seem to be fully conserved in representatives of all subgroups (9). The distal 5′ nontranslated region varies significantly between the strains. The conservation of the 5′ terminus and of tyrosine implicated in RNA binding (8) suggests that constraints are imposed on their interaction.

It can be inferred from the data on poliovirus (10, 11) that VPg binding to RNA is intimately related to the synthesis of progeny plus and minus RNA strands. It is suggested that a covalent [5′-(O^4-tyrosylphospho)adenylate] bond could be formed during potyviral RNA replication either by transesterification with nascent RNA chain or as a result of adenylylation of specific VPg tyrosine. Such VPg mono- (or oligo)adenylate could prime the progeny RNA synthesis. The two mechanisms are believed not to be mutually exclusive and could reflect the role of VPg in differentiation of viral plus and minus RNA synthesis by a strand-specific initiation mechanism (11). From the chemical point of view, the first mechanism assumes that the phenolic hydroxyl group of tyrosine may directly participate in transesterification with the phosphodiester bond of the RNA chain. The second mechanism points toward direct recognition of the 5′-nucleotide phosphate by appropriately exposed tyrosine. We therefore decided to visualize the interaction by modeling the polypeptide and the 5′ end of PVY RNA, allowing for the best fit of both complex-forming macromolecules. One could hope that stereochemical insight into the complex would be of interest in evaluating the mechanisms of unusual bond formation between protein and RNA.

Up to the present, there has been no general method for safely predicting the three-dimensional structure of a protein from sequence data (12, 13). The most promising method of structure prediction is homology-based modeling (14, 15). Nevertheless homology-based modeling requires that at least one three-dimensional structure among related proteins is known, to be used as a model template. Recently, some interesting attempts to eliminate this problem have been published (16, 17). It should be noted that protein structures are able to adapt to amino acid replacements (18), especially when the general pattern of hydrophobic–hydrophilic residue distribution is conserved (19). The tolerance to amino acid substitution has been pointed out in studies on mutant proteins (20, 21). Detailed analysis of λ repressor mutants (22) has suggested that the hydrophobic core residues play an essential role in determining protein conformation and stability, and it is possible that the pattern of hydrophobic and hydrophilic residues determines the protein topology (23, 24).

Taking this into account and selecting an appropriate template, we decided to model the structure of PVY VPg. This led to the first, to our knowledge, modeling of the three-dimensional structure based on similarity of distribution of hydrophobic and hydrophilic residues. Quite surprisingly, the VPg model agrees with several biochemical predictions on peptide interaction with the 5′ end of the viral genome (10).

MATERIALS AND METHODS

Sequencing of PVY cDNA. PVY-specific insert was excised from the PVY15 plasmid (25) with *Pst*I restriction endonuclease and was inserted into the *Pst*I site of the pBluescript SK (+) (Stratagene) vector. The new construct was then digested with *Xba*I/*Bst*XI or with *Xho*I/*Eco*RV, for the sequencing from M13 reverse or universal primer, respectively (26). To obtain a set of unidirectional nested deletions, the Erase-A-Base system was used, and the linearized plasmid was treated with exonuclease III, S1 nuclease, the Klenow fragment of DNA polymerase, and T4 DNA ligase according to the system supplier (Promega). The resulting constructs with shortened

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: VPg, the genome-linked viral protein; PVY, potato virus Y; PDB, Protein Data Bank.

Data deposition: The sequence reported in this paper has been deposited in the GenBank data base (accession no. Z29526).

*To whom reprint requests should be addressed.

cDNA inserts were used to transform competent *Escherichia coli* DHα5-F′ cells, purified (26), and subjected to sequencing by the dideoxynucleotide method using the A.L.F. (automated laser fluorescent) DNA sequencer (Pharmacia).

Alignment of VPg Protein Sequences. The programs for the sequence homology search and multiple alignment were those implemented in the Wisconsin GCG package, version 7.3. The sequence of PVY15 VPg protein was compared with all the sequences in the Swiss-Prot data base, release 29 (27), using FASTA. The PILEUP program was used for the alignment of sequences of potyviral proteins.

Homology Modeling. We analyzed sequences of 50-residue fragments of the studied protein and looked for homologous segments in the Protein Data Bank (PDB) data base (28). As only sequences with little evidence of homology were found, these fragments were considered inappropriate for modeling. Therefore, a simplified approach was adopted by considering a distribution of hydrophobic and hydrophilic residues along the polypeptide chain as the basis for homology modeling.

We compared hydrophobic–hydrophilic distribution (29) of the VPg protein's residues to all the sequences available from the PDB data base using PRO-EXPLORE version 1.21 (Oxford Molecular). The best similarity (52%) of that distribution for 188 residues of VPg was found to the fragment (from Gly-94 to Asp-281) of cytoplasmic malate dehydrogenase of known crystal structure (30, PDB entry 4MDH). The sequence comparison of the VPg protein and malate dehydrogenase fragment is shown in Fig. 1. As can be seen, there are no insertions or deletions, making the modeling task relatively easy. The malate dehydrogenase residues were mutated to the corresponding VPg sequence by using INSIGHT/HOMOLOGY version 2.3.0 (Biosym, San Diego, CA). This procedure did not cause any serious steric clashes that could not be removed by manual adjustment of side-chain torsions. The VPg structure obtained in such a manner was subjected to energy minimization using CVFF forcefield as implemented in DISCOVER version 2.3.4. After initial minimization for 300 steps using the steepest descents algorithm, conjugate gradients minimization *in vacuo* was carried out until the maximum derivative was less than 10^{-3}.

It is worth noting that random replacement of amino acid residues in the VPg produced a model differing significantly from the template. To test this, we adopted the following procedure:

The VPg sequence was randomized with the GCG version 8.1 program SHUFFLE. This gave a sequence of the same length and residue composition, but random. The resulting sequence was SSEFEKRAQG KDLMDVEDAI DKDDEQEFLI FHNRAKDITM QIFIGSRIHP IKQKLHKLIY VSFSNFYDWP EEIRFPQTKK GRTANGFPDD EIVMERAKAK SGTERNVFSI TNKKPLDERG RKELAPMEMT ANDGDFAGVV DEENTKAYNV KNDGKAKDAT FGLIRVPSRH YVGRVDKIER GLYCGTQQ.

The shuffled sequence was the subject of the same modeling procedure as the VPg sequence on the template structure, including energy minimization. The resultant structure is 4.38 Å (main-chain rms) apart from the initial malate dehydrogenase backbone. This value is significantly higher than that obtained for the VPg structure. Also, both malate dehydrogenase and VPg do have significant content of regular secondary structures, whereas the structure obtained from the shuffled VPg sequence does not. The only piece of secondary structure here is the short α-helix (residues 116–118) as calculated with the BIOPOLYMER module of SYBYL 6.2 (Tripos Associates, St. Louis). In the randomized polypeptide, neither the position of tyrosine nor the distribution of positively charged residues allows specific interaction with the RNA molecule.

Structure Modeling of VPg–RNA Complex. It was assumed that Tyr-64 is covalently bound through its hydroxyl oxygen to the 5′ end of the RNA. Therefore the 5′-end adenosine bound to the protein constituted the initial starting structure in our model. The subsequent nucleotides were added by systematic searching (when moving from the bound adenosine in 5′→3′ direction) of the most probable interactions of VPg surface positive charges with subsequent phosphate groups of the RNA. This procedure is justified by experimental evidence (9, 31) showing that the 5′ end of potyviral RNA is probably single-stranded. Therefore we excluded the possibility that RNA–protein interaction specificity results from the formation of the RNA loop structures. The resultant model structure was again subjected to energy minimization.

RESULTS AND DISCUSSION

Sequencing of the virus-specific insert of the PVY15 clone has revealed that the 1784-bp cDNA fragment corresponds to the central part of the PVY genome between nucleotides 4708 and 6491 without any deletion or insertion (numbering according to ref. 32). The insert covers the C-terminal fragment (44.3%) of cytoplasmic inclusion protein (CI), full sequences of $6K_2$ and putative VPg, and the N-terminal fragment (29.1%) of NI^{pro} protease (Fig. 2). The putative VPg protein is composed of 188 amino acid residues and is highly conserved among PVY strains: PVY^{N}, a French isolate (32); PVY^{NTN}, a Hungarian isolate (33); and PVY^{0}, a Japanese isolate (34) (identity 94.1–98.9%).

Homology with the VPg sequences of nine other potyviruses was found. Multiple alignment and descriptions of these proteins

```
VPg GKNKSKRIQALKFRHARDKRAGFEIDNNDDTIEEFFGSAYRKKGKGKGTT 50
    : :::: :::    :  :  :  :: ::: ::  :   : ::  : : :
MDH GMERKDLLKANVKIFKCQGAALDKYAKKSVKVIVVGNPANTNCLTASKSA 143

VPg VGMGKSSRRFINMYGFDPTEYSFIQFVDPLTGAQIEENVYADIRDIQERF 100
      ::::: :::: : ::  ::::::  :   :    :  :: :::: :
MDH PSIPKENFSCLTRLDHNRAKAQIALKLGVTSDDVKNVIIWGNHSSTQYPD 193

VPg SEVRKKMVENDDIEMQALGSNTTIHAYFRKDWSDKALKIDLMPHNPLKVC 150
     :: :  :: ::: :  : ::::::     :  ::   : :   :  :::
MDH VNHAKVKLQAKEVGVYEAVKDDSWLKGEFITTVQQRGAAVIKARKLSSAM 243

VPg DKTNGIAKFPERELELRQTGPAVEVDVKDIPAQEVEHE 188
    :  : :::: :: :  : :: :::: : : : :    :
MDH SAAKAICDHVRDIWFGTPEGEFVSMGIISDGNSYGVPD 281
```

FIG. 1. Comparison of VPg and the residues 94–281 fragment of malate dehydrogenase (MDH) sequences. Colons indicate similar hydrophobic–hydrophilic pattern.

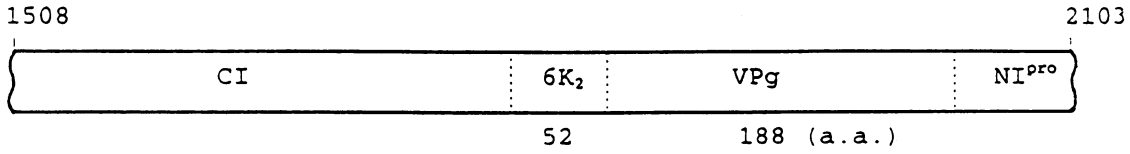

FIG. 2. Presumed structure of the central part of PVY^0 (British isolate) polyprotein deduced from the nucleotide sequence of the PVY15 clone. Vertical dotted lines indicate the cleavage sites of the polyprotein, which may be processed to cytoplasmic inclusion protein (CI), 6-kDa protein ($6K_2$), genome-linked protein (VPg), and small nuclear inclusion protease (NI^{pro}). The fragment corresponds to the region of the PVY^N polyprotein between amino acid residues 1508 and 2103 (numbering according to ref. 31). The number of amino acid residues of VPg and $6K_2$ cistron is shown below the box.

are shown in Fig. 3. There are 19% residues which are fully conserved among 10 compared proteins, and 38% are invariant according to their hydrophobic–hydrophilic character.

The proposed molecular structure of VPg obtained as discussed in the previous section is presented in Fig. 4. The rms difference between the main-chain atoms of VPg and malate

```
PVY15   GKNKSKRIQALKFRHARDKRAGFEIDNNDDTIEEFFGSAYRKKGKGKG..
PPVD    .GFNRRQRQKLKFRQARDNRMAREVYGDDSTMEDYFGSAYSKKGKSKG..
TUMV    ...GKKQRQKLKFRNARDNKMGREVYGDDDTIEHFFGDAYTKKGKSKG..
PRV     .GFSARQRQKLRFKSAANAKLGREVYGDDGTIEHYFGEAYTKKGNKKG..
TVMV    ....GKSRRRLQFRKARDDKMGYIMHGEGDTIEHFFGAAYTKKGKSKG..
SBMV    ..GKKRQIQKLKFRDAFDRKVGREVYADDYTMEHTFGEAYTKKGKQKGST
PEMV    GRSKTKRIQALKFRKARDKRAGFEIDNNEDTIEEYFGSAYTKKGKGKG..
TEV     ..GKKNQKHKLKMREARGARGQYEVAAEPEALEHYFGSAYNNKGKRKG..
OMV     ..................................................
PSBMV   ...GKSKAKTLRFRQARDNNAKYEVFADEDTKRHYFGEAYTKKGKKSG..

PVY15   TTVGMGKSSRRFINMYGFDPTEYSFIQFVDPLTGAQIEENVYADIRDIQE
PPVD    KTRGMGTKTRKFVNMYGYDPTDYNFVRFVDPLTGHTLDEDPLMDINLVQE
TUMV    RTRGIGHKNRKFINMYGFDPEDFSAVRFVDPLTGATLDDNPFTDITLVQK
PRV     KMHGMGVKTRKFVATYGFKPEDYSYVRYLDPLTGETLDESPQTDISMVQD
TVMV    KTHGAGTKAHKFVNMYGVSPDEYSYVRYLDPVTGATLDESPMTDLNIVQE
SBMV    RTKGMGRKSRNFIHLYGVEPENYSMIRVVDPLTGHTMDEHPRVDIRMVQQ
PEMV    TTVGMGRTNRRFINMYGFEPGQFSYIKFVDPLTGAQMEENVYADIVDVQE
TEV     TTRGMGAKSRKFINMYGFDPTDFSYIRFVDPLTGHTIDESTNAPIDLVQH
OMV     .....................EYTIVRYVDPLTGATQDENPLMAIDLVQE
PSBMV   KARGMGVKTKKFVNVYGFDPCEYSLVRFVDPLTGLTYDRHPMEHMMDVQE

PVY15   RFSEVRKKMVENDDIEMQALGSNTTIHAYFRKDWSDKALKIDLMPHNPLK
PPVD    HFSQIRNDYIGDDKITMQHIMSNPGIVAYYIKDATQKALKVDLTPHNPLR
TUMV    HFGDIRMDLLGEDELDSNEIRMNKTIQAYYMNNKTGKALKVDLTPHIPLK
PRV     HFSDIRRKYMDSDSFDRQALIANNTIKAYYVRNSAKAALEVDLTPHNPLK
TVMV    HFGEIRREAILADAMSPQQ..RNKGIQAYFVRNSTMPILKVDLTPHIPLK
SBMV    EFEEIRKDMIGEGELDRQRVYHNPGLQAYFIGKNTEEALKVDLTPHRPTL
PEMV    KFGDIRRQMILDDELDRRQTDVHNTIHAYLIKDWSNKALKVDLTPHNPLR
TEV     EFGKVRTRMLIDDEIEPQSLSTHTTIHAYLVNSGTKKVLKVDLTPHSSLR
OMV     YFAKIRSQLVSEEKLETQNIIANPGIQAYYMKNRGDAALKVDLTPHNPLL
PSBMV   TIGDDRREAMWNDELDKQLFVTRPTIEAYYIKDKTTPALKIDLNPHNPMR

PVY15   VCDKTNGIAKFPERELELRQTGPAVEVDVKDIPAQEVEHE.........
PPVD    VCDKTATIAGFPEREFELRQTGHPIFVEPNAIPKINEEGDEEVDHE...
TUMV    VCDLHATIAGFPERENELRQTGKAQPINIDEVPRANNE.LVPVDHESNS
PRV     VCDNKLTIAGFPDREAELRQTGPPRTIQVDQVPPPSK....SVHHE...
TVMV    VCESN.NIAGFPEREGELRRTGPTETLPFDALPPEKQEVAFE.......
SBMV    LCQNSNAIAGFPEREDDLRQTGLPQVVSKSDVPRAKERVEME.......
PEMV    VSDKASAIMKFPEREGELRQTGQAVEVDVCDIPKEVVKH..........
TEV     ASEKSTAIMGFPERENELRQTGMAVPVAYDQLPPKNEDLTFE.......
OMV     V.TKTGTIAGFPENEFILRQTGKAVNVKMSEVPVENELEEVEHEG....
PSBMV   VCDKAETIAGFPEREFELRQSGSATLVPYSEVPVQNEKQEFDEEHWRTE
```

FIG. 3. Sequence alignment of VPg proteins from potyviruses. Residues conservative through all the sequences are shown in boldface. PVY15, potato virus Y, British isolate; PPVD, plum poxvirus, strain D; TUMV, turnip mosaic virus; PRV, papaya ringspot virus, isolate HA; TVMV, tobacco vein mottling virus; SBMV, soybean mosaic virus, strain G-6, isolate K-1; PEMV, pepper mottle virus, California strain; TEV, tobacco etch virus; OMV, ornithogalum mosaic virus; PSBMV, pea seed-borne mosaic virus, isolate DP1. Note that OMV VPg cistron was partially sequenced.

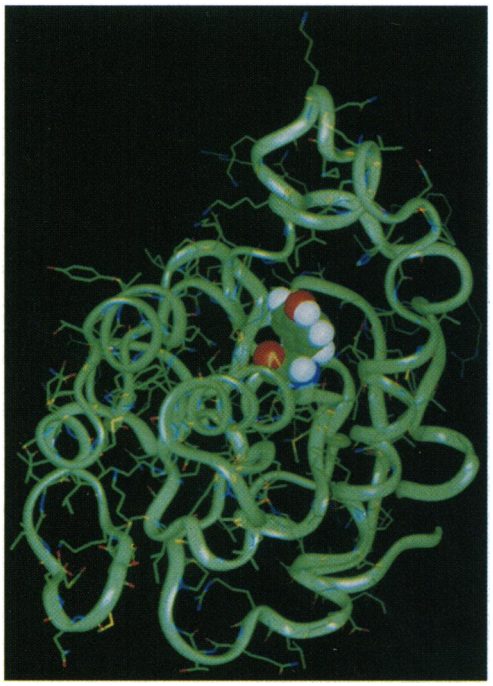

FIG. 4. Hypothetical model of VPg protein structure. The green ribbon traces the polypeptide backbone. A CPK (Corey–Pauling–Koltun) space-filling representation is used to show the Tyr-64 position.

dehydrogenase is 3.44 Å. It is noteworthy that mutating the malate dehydrogenase residues to the corresponding VPg sequence did not produce any serious steric clashes in the structure. This gives additional confidence that the cores of both proteins are very similar and that the procedure adopted by us is correct. The VPg model structure is compact, and its core was found by calculation of solvent accessibilities (data not shown). The model contains several helical fragments, namely: Tyr-40–Gly-47, Tyr-64–Gln-75, Thr-174–Ala-176, Thr-122–Asp-134, Leu-137–Phe-143, and Val-149–Glu-165. Some of the solvent-inaccessible residues are located on the buried sites of these helices.

The model structure of VPg was validated by using three different algorithms: (*i*) three-dimensional profiles (35) as implemented in the Profiles 3D module of BIOSYM software package version 95.0; (*ii*) the method of Godzik and Skolnick (36) as implemented in the MatchMaker module of SYBYL VERSION 62.0, and (*iii*) the BIOTECH Validation Suite for Protein Structures (which may be accessed at http://www.embl-heidelberg.de). All three methods examine the goodness of the fit of the sequence to the three-dimensional structure. The results from these methods are in agreement, finding three regions in our model in which the residue environment is different from the one usually found in protein structures: 105–107, 259–262, and 276–280. We note that the regions 105–107 and 276–280 are implicated in RNA binding (the model of RNA–VPg interaction is discussed below), as probably is the central Arg-260 from the third region. Similar disparities are present in other nucleic acid-binding proteins (data not shown).

Earlier studies (6) of the poliovirus genome have established that the VPg molecule is linked to the viral RNA by a covalent bond between the O4 of tyrosine (Tyr-1860 in the polyprotein) and the 5′-P of the terminal uridylic acid residue. The tyrosine residue was also found to be involved in binding potyviral (TVMV) VPg to the viral genomic RNA (8). This tyrosine (Tyr-64 in the VPg sequence) is located at the N terminus of the buried helix, with only the hydroxyl group exposed. The Tyr-64 hydroxyl is located in the crevice between Tyr-122–Asp-134, Leu-137–Phe-143, and Val-149–Glu-165 helices. The nearest neighbors of Tyr-64 are Trp-132, Ile-139, Lys-148, and Asp-151 at distances of 5.15 Å, 3.14 Å, 4.64 Å, and 5.30 Å, respectively. Moreover, Tyr-64 is conserved in related proteins (Fig. 3). Tyr-64 of the VPg substitutes for Leu-157 in the malate dehydrogenase sequence, which is involved in binding of the nicotinamide ring of NAD (30). Location of Tyr-64 in the place analogous to the binding site of malate dehydrogenase as well as solvent accessibility of its hydroxyl group in the protein model structure has tempted us to propose a structural model of the VPg–viral RNA complex.

Since the terminal residue of PVY RNA is adenosine, we postulate that Tyr-64 of VPg is bound to the terminal adenylic acid residue. The molecular structure of VPg linked to an 11-nucleotide fragment of RNA: VPg (Tyr-O)-pAAUUAAAACAA, is presented in Fig. 5. This RNA fragment participates in 32 hydrogen bonds with VPg protein. Among these 27 are formed by phosphate groups of the RNA chain, 3 are contacts between the bases of RNA and protein residues, and 2 are ribose–amino acid interactions. Some of the interactions may be of importance for RNA–protein recognition, because amino acid residues taking part in them are conservative or invariant according to their hydrophilic character. These are the Lys-6, Lys-12, Lys-45, Gly-48, Lys-138, His-144, and Lys-148 residues.

Evaluating the model, we would emphasize that VPg polypeptide folding results in a simple compact domain. This could be expected for a protein liberated from a polyprotein by specific proteolysis. The protein seems to be highly structured with no clear alternative conformation. The model structure creates a surface accepting the 5′ end of PVY RNA, exposing the tyrosine hydroxyl group in a position favorable for synthesis of the appropriate diester. Complementarity between VPg surface and the 5′ end of viral RNA agrees with the structure of the genomic PVY strand. Covalent binding of RNA to Tyr-64 exposes the 5′-proximal bases on the surface of the protein in a manner allowing for complementary hydrogen bond formation. Such an "open" complex would not create steric hindrance during termination of minus strand synthesis during replication of the 5′ terminus of genomic RNA. Calculated atomic distances would allow for adenylylation of Tyr-64. The adenylate donor can be either adenosine triphosphate or a phosphodiester bond in a continuous RNA chain (data not shown). It is worthwhile to point out that a

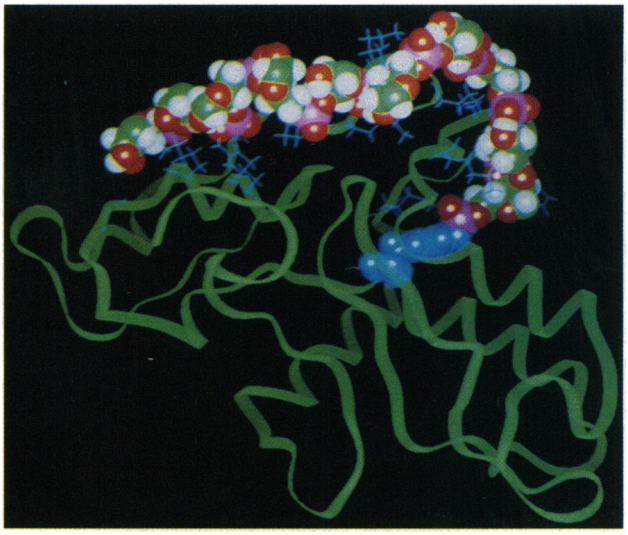

FIG. 5. Schematic ribbon drawing illustrating interactions of VPg protein with RNA fragment. The RNA backbone and Tyr-64 are shown in space-filling representation. The side-chain residues that are in close contact with RNA are shown in blue.

stretch of 11 sugar-phosphate backbone units can contact the protein surface. The phosphodiester bond preceding adenylate in such a continuous RNA chain would be sterically accessible to nucleophilic attack by the hydroxyl group of Tyr-64. Therefore the proposed structure conforms with the results of experiments suggesting the simultaneous existence of two different chemical mechanisms of VPg binding to RNA (10, 11). Confirmation of predictions requires *in vitro* experiments with purified PVY VPg and appropriate RNA precursors.

Judging the generality of the model, one should stress that despite the similarities in genome organization and expression within the picorna-like group (2), it is not clear to what extent VPg proteins have similar functions for all viruses of that group. There are differences in VPg molecular mass (22–24 kDa for potyviruses versus 2 kDa for picornaviruses and comoviruses), low similarity of amino acid sequence of potyviral VPg proteins to that of other families of viruses (identity 10–16%, data not shown), and the type of residue that links VPg to the viral RNA [tyrosine for potyviruses and picornaviruses (25), and serine for comoviruses–cowpea mosaic virus (37)]. However, certain observations: (*i*) comoviral VPg protein exists *in vivo* in the form of the membrane-bound 60-kDa precursor (38); (*ii*) replication of comovirus occurs in the membranous cytopathological structures; (*iii*) 5′ ends of both plus and minus strands of comovirus RNA are covalently linked to the VPg (39); and (*iv*) the sequence of potyviral VPg on the polyprotein is directly preceded by a short polypeptide, $6K_2$, that contains a stretch of hydrophobic amino acids—a likely candidate for a donor peptide to anchor the VPg (or its precursor) in the membranes (40)—suggest that similarity between picorna-like viruses might also concern the replication of viral RNAs, thus implicating an important role of the plant virus VPg in the initiation of RNA synthesis.

Note added in Proof. Recent *in vitro* experiments confirm the molecular mass of the VPg protein and the position of the N terminus assumed for the construction of the model presented in this paper.†

†Chiang, A. N., Hwang, D. J. & Tumer, N. E., Tenth International Congress of Virology, Aug. 11–16, 1996, Jerusalem, abstr. W 19-4).

We acknowledge Dr. D.C. Baulcombe (The Sainsbury Laboratory, Norwich, U.K.) for supplying the PVY15 plasmid. We thank Prof. D. Shugar for critical reading of the manuscript. This work was partly supported by Grant 6 6380 9203 from the Polish State Committee for Scientific Research.

1. Wimmer, E. (1982) *Cell* **28,** 199–201.
2. Strauss, J. H. & Strauss, E. G. (1988) *Annu. Rev. Microbiol.* **42,** 657–683.
3. Murphy, J. F., Rhoads, R. E., Hunt, A. G. & Shaw, J. G. (1990) *Virology* **178,** 285–288.
4. Siaw, M. F. E., Shahabuddin, M., Ballard, S., Shaw, J. G. & Rhoads, R. E. (1985) *Virology* **142,** 134–143.
5. Riechmann, J. L., Lain, S. & Garcia, J. A. (1989) *J. Gen. Virol.* **70,** 2785–2789.
6. Rothberg, P. G., Harris, T. J. R., Nomoto, A. & Wimmer, E. (1978) *Proc. Natl. Acad. Sci. USA* **75,** 4868–4872.
7. Ambros, V. & Baltimore, D. (1978) *J. Biol. Chem.* **253,** 5263–5266.
8. Murphy, J. F., Rychlik, W., Rhoads, R. E., Hunt, A. G. & Shaw, J. G. (1991) *J. Virol.* **65,** 511–513.
9. Marie-Jeanne Tordo, V., Chachulska, A. M., Fakhfakh, H., Le Romancer, M., Robaglia, C. & Astier-Manifacier, S. (1995) *J. Gen. Virol.* **76,** 939–949.
10. Tobin, G. J., Young, D. C. & Flanegan, J. B. (1989) *Cell* **59,** 511–519.
11. Takeda, N., Kuhn, R. J., Yang, C.-F., Takegami, T. & Wimmer, E. (1986) *J. Virol.* **60,** 43–53.
12. Hilbert, M., Böhm, G. & Jaenicke, R. (1993) *Proteins* **17,** 138–151.
13. Wilson, C., Gregoret, L. M. & Agard, D. A. (1993) *J. Mol. Biol.* **229,** 996-1006.
14. Swindells, M. B. & Thornton, J. M. (1991) *Curr. Opin. Struct. Biol.* **1,** 219–223.
15. Šali, A. & Blundell, T. L. (1993) *J. Mol. Biol.* **234,** 779–815.
16. Dandekar, T. & Argos, P. (1994) *J. Mol. Biol.* **236,** 844–861.
17. Sudarsanam, S., March, C. J. & Srinivasan, S. (1994) *J. Mol. Biol.* **241,** 143–149.
18. Chelvanayagam, G., Roy, G. & Argos, P. (1994) *Protein Eng.* **7,** 173–184.
19. Sweet, R. M. & Eisenberg, D. (1983) *J. Mol. Biol.* **171,** 479–488.
20. Rennell, D., Bouvier, S. E., Hardy, L. W. & Poteete, A. R. (1991) *J. Mol. Biol.* **222,** 67–87.
21. Bowie, J. U., Reidhaar-Olson, J. F., Lim, W. A. & Sauer, R. T. (1990) *Science* **247,** 1306–1310.
22. Lim, W. A. & Sauer, R. T. (1991) *J. Mol. Biol.* **219,** 359–376.
23. Kamtekar, S., Schiffer, J. M., Xiong, H., Babik, J. M. & Hecht, M. H. (1993) *Science* **262,** 1680–1685.
24. Yue, K. & Dill, K. A. (1995) *Proc. Natl. Acad. Sci. USA* **92,** 146–150.
25. Baulcombe, D. C. & Fernandez-Northcote, E. N. (1988) *Plant Disease* **72,** 307–309.
26. Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab. Press, Plainview, NY), 2nd Ed.
27. Bairoch, A. & Boeckmann, B. (1993) *Nucleic Acids Res.* **21,** 3093–3096.
28. Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyer, E. F., Jr., Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T. & Tasumi, M. (1977) *J. Mol. Biol.* **112,** 535–542.
29. Hopp, T. P. & Woods, K. R. (1981) *Proc. Natl. Acad. Sci. USA* **78,** 3824–3828.
30. Birktoft, J. J., Rhodes, G. & Banaszak, L. J. (1989) *Biochemistry* **28,** 6065–6081.
31. Nicolas, O. & Laliberté, J.-F. (1992) *J. Gen. Virol.* **73,** 2785–2793.
32. Robaglia, C., Durand-Tardif, M., Tronchet, M., Boudazin, G., Astier-Manifacier, S. & Casse-Delbart, F. (1989) *J. Gen. Virol.* **70,** 935–947.
33. Thole, V., Dalmay, T., Burgyan, J. & Balazs, E. (1993) *Gene* **123,** 149–156.
34. Hidaka, M., Yoshida, Y., Masaki, H., Namba, S., Yamashita, S., Tsuchizaki, T. & Uozumi, T. (1992) *Nucleic Acids Res.* **20,** 3515.
35. Lüthy, R., Bowie, J. U. & Eisenberg, D. (1992) *Nature (London)* **356,** 83–85.
36. Godzik, A. & Skolnick, J. (1992) *Proc. Natl. Acad. Sci. USA* **89,** 12098–12102.
37. Jaegle, M., Wellink, J. & Goldbach, R. (1987) *J. Gen. Virol.* **68,** 627–632.
38. Vos, P., Verser, J., Jaegle, M., Wellink, J., van Kammen, A. & Goldbach, R. (1988) *Nucleic Acids Res.* **16,** 1967–1985.
39. Lomonossoff, G. P. (1994) in *Encyclopedia of Virology*, eds. Webster, R. G. & Granoff, A. (Academic, London), Vol. 1, pp. 249–254.
40. Riechmann, J. L., Lain, S. & Garcia, J. A. (1992) *J. Gen. Virol.* **73,** 1–16.