Structure and Assembly of the RNA Binding Domain of Bluetongue Virus Non-structural Protein 2*

Received for publication, January 16, 2004, and in revised form, May 19, 2004
Published, JBC Papers in Press, May 20, 2004, DOI 10.1074/jbc.M400502200

Carmen Butan, Hans van der Zandt, and Paul A. Tucker‡
From the European Molecular Biology Laboratory, c/o Deutsches Elektronen-Synchrotron, Notkestrasse 85, D-22603 Hamburg, Germany

Bluetongue virus non-structural protein 2 belongs to a class of highly conserved proteins found in orbiviruses of the Reoviridae family. Non-structural protein 2 forms large multimeric complexes and localizes to cytoplasmic inclusions in infected cells. It is able to bind single-stranded RNA non-specifically and has been suggested that the protein is involved in the selection and condensation of the Bluetongue virus RNA segments prior to genome encapsidation. We have determined the x-ray structure of the N-terminal domain (sufficient for the RNA binding ability of non-structural protein 2) to 2.4 Å resolution using anomalous scattering methods. Crystals of this apparently insoluble domain were obtained by in situ proteolysis of a soluble construct. The asymmetric unit shows two monomers related by non-crystallographic symmetry, with each monomer folded as a β sandwich with a unique topology. The crystal structure reveals extensive monomer-monomer interactions, which explain the ability of the protein to self-assemble into large homomultimeric complexes. The entirety of the monomer, one-third is used to create the interfaces of the curved multimeric assembly observed in the x-ray structure. The structure reported here shows how the N-terminal domain would be able to bind single-stranded RNA non-specifically protecting the bound regions in a heterogeneous multimeric but not polymeric complex.

Bluetongue virus (BTV) is a representative member of the Orbivirus genus within the Reoviridae family and has a ten-segment double-stranded RNA (dsRNA) genome enclosed within a double capsid. The segments code for seven capsid and viral core proteins (VP1–VP7). The remaining three segments encode non-structural proteins (NS1, NS2, and NS3/NS3A) that are produced in the host cell at different stages of the infectious cycle and are presumed to be involved in the various steps of virus morphogenesis.

Bluetongue virus RNA segment 8 encodes non-structural protein 2 (molecular mass ~41 kDa), which is phosphorylated (1) and able to self-associate to form multimeric assemblies in complex with single-stranded RNA (ssRNA) (2). The protein is synthesized in large amounts throughout the replication cycle of Bluetongue virus and is associated with large dense perinuclear structures called viral inclusion bodies in BTV-infected mammalian cells (3, 4). It is believed that BTV assembly occurs at the perimeter of these viral inclusion bodies (3). In insect cells, recombinant baculovirus-expressed NS2 has been shown to accumulate as multimers forming inclusion bodies, not unlike those observed in BTV-infected mammalian cells, even when synthesized independently of any other BTV-encoded proteins or viral RNA (5). Deletions at the C terminus of up to 130 amino acids do not affect the protein oligomerization (6).

Compared with the other BTV proteins, NS2 appears unique by virtue of its ability to bind ssRNA (but not dsRNA). The protein is of particular interest because of its possible involvement in genome recognition (preference for viral ssRNA versus cellular ssRNA), which generally is known to be an early crucial step in the assembly process of the viruses. Two deletion mutagenesis studies (6, 7) have indicated that the N-terminal half of the protein, which bears most of the homology within the NS2 family, is an RNA binding domain. Although the isolated N-terminal half of the protein can be functional, other RNA binding sites are required for the high affinity RNA binding (7). An analysis of the primary structure of NS2 does not reveal any sequence similarity to well known RNA binding motifs such as the ribonucleotide motif (8), the K homology (KH) domain originally described in the heterogeneous nuclear ribonucleoprotein (hnRNP) K protein (8), the arginine-rich motif (9), or the RGG and RS motifs (10).

NS2 shares a number of features with other non-structural proteins; for example, NSP2 of rotavirus (11) and αNS of reovirus (12), the functions of which have been implicated in genome packaging. It has been shown that cells infected with temperature-sensitive mutants of NSP2 contain few replication assembly factories (viroplasms) and produce virus particles that are mostly empty (13).

To obtain insight into how NS2 and ssRNA interact and function, we have determined the crystal structure of the RNA binding domain of NS2 and examined it for potential binding surfaces as well as defined flexible protein regions that might become ordered upon RNA binding.

**Experimental Procedures**

Protein Expression Studies of the N-terminal Domain Using Recombinant Baculovirus-infected Insect Cells and Escherichia coli—The S8 gene from BTV serotype 10 (encoding full-length NS2) contained in the recombinant plasmid pNS2 was used to generate the ns21–177 construct. DNA coding for residues 1–177 was produced by PCR amplification on the NS21–354 sequence with the primers A (5′-GGGATCCATATGAGCAAAAGCAACGTAG-3′) and B (5′-CTCGAGCGGCCGCTTACGGCCGCGCCACGCTATGAACTTGAAG-3′) inserting BamHI and EseFl sites before cloning the PCR product into a modified pVL1393 plasmid (Invitrogen) with a C-terminal His tag. The plasmid for the untagged NS21–177 protein was produced using the BamHI and Notl sites.

This paper is available on line at http://www.jbc.org
Crystal Structure of the RNA Binding Domain of NS2 from BTV

To produce C-terminal glutathione S-transferase-tagged protein, the primer 5'-ATATCCGACGCCGCAAAACCCGTTAATCCCGG- GCCATGGACATGTCACATTACGACCAC-3' and the reverse strand 3'-CCGCTCGAGATTTTATTTTCAGGGCCGCTGGATGGATGATGA- AAAATAGATTCTC-5' were used with a His<sub>6</sub> glutathione S-transferase tag and a tobacco etch virus (TEV) site (plasmid and sequence available upon request) as a template. The PCR product was cut with EagI and ligated with the C-terminal His<sub>6</sub>-tagged pVL1393NS2<sub>1-177</sub> vector, as well as cut with EagI. The constructs were checked by sequencing.

The NS2<sub>1-177</sub> constructs were expressed at 28 °C in High 5 insect cells (Invitrogen) by means of infection with recombinant baculovirus at inserted into the C-terminal His<sub>6</sub>-tagged pET-21a vector (Novagen). NdeI and EagI to produce the appropriate restriction sites and was pean Molecular Biology Laboratory, Heidelberg, Germany). The same

ments were generated using primers A and E (5'-H11032

and G. The PCR product was ligated into expression vector pET-22b( F( 5'-TGAG-3' /H11032

in Luria-Bertani medium containing ampicillin at a concentration of 50

ul of 354 with the TEV cleavage occurring between Gln and Gly. The

Construction of the TEV Insert Construct—The full-

length NS2 from Bluetongue virus serotype 10 was engineered with a

TEV protase cleavage site between residues 182 and 183. The insert

resulted in the sequence NS2<sub>1-177</sub>-Glu-Asn-Leu-Tyr-Pheny-Gly-NS2<sub>177-354</sub> with the TEV cleavage occurring between Gin and Gly. The

PCR amplification was performed in two steps. In the first step, frag-

ments were generated using primers A and E (5'-GGCTCTGAAATAA-

AGATCTTCGACATCTCCCTGGCCG-3') for the first fragment and F (5'-GAATTCACTTTTTCAGGGCCGCTGGATGGATGATGA-

GACATGTCACATTACGACCAC-3') for the second fragment. The second

PCR step used the two generated fragments together with primers A and G. The

PCR product was ligated into expression vector pET-22b( Novagen) at the NdeI and EagI sites. The plasmid for untagged NS2<sub>1-354</sub> was expressed using the NdeI and Xhol sites. The resulting C-terminal His<sub>6</sub>-tagged plasmid was transformed into E. coli BL21(DE3)Plys (Novagen), and 1-liter expression cultures were grown in Luria-Bertani medium containing ampicillin at a concentration of 50 μg/ml and chloramphenicol at 34 μg/ml. The cells were grown at 37 °C to an A<sub>595</sub> of ~0.6, and expression was induced at 25 °C with 0.4 mM isopropyl-β-D-galactopyranoside. After 6 h of induction, the cells were harvested by centrifugation and lysed by sonication in 100 mM sodium phosphate buffer, pH 8.0, 1 mM sodium chloride, 10 mM dithio-
threitol in the presence of protease inhibitor mixture (Roche Applied Science). After the sonication step, the supernatant was clarified by ultracentrifugation (Sorvall SS34 rotor, 20,000 rpm, 4 °C, 1 h), filtered through a membrane, and loaded on a Q-Sepharose column (Amerham Biosciences) equilibrated with 100 mM sodium chloride in 100 mM sodium phosphate buffer at pH 8.0. The protein eluted at a concentration of 400 mM sodium chloride. Purity of the protein was monitored by SDS-PAGE. The protein fractions at the absorption peak were pooled and diluted by a factor of two prior to loading onto a HiTrap Heparin column (5 ml) (Amersham Biosciences) and equilibrated with 200 mM sodium chloride, 10 mM sodium phosphate, pH 7.0 and dialyzed overnight. After a 4-column volume wash, a two step elution of 300 mM and of 600 mM sodium chloride was developed. The C-terminal His<sub>6</sub>-tagged fusion protein eluted at a concentration of 600 mM sodium chloride. To de-

crease the salt concentration, the eluted protein solution was dialyzed against the storage buffer consisting of 10 mM sodium phosphate buffer, pH 8.0, 500 mM sodium chloride, 2 mM dithiothreitol at 4 °C overnight. The selenomethionine-labeled protein was expressed in a methionine-

auxotrophic cell host, the E. coli B834(DE3)Plys strain (Novagen), and purified in the same manner as the native protein. To avoid oxidation of selenomethionine, all buffers were flushed with nitrogen and supple-

mented with 10 mM dithiothreitol.

Crystallography and Data Collection—Diffraction-quality crystals of the N-terminal domain of NS2 protein were obtained at 20 °C by the hanging drop vapor diffusion technique. The crystallization drops con-

sisted of 1 μl of mixed protein solution (8–15 mg/ml) with TEV protase (50 μg/ml) and 1 μl of reservoir solution. Each drop was equilibrated against 1 ml of the reservoir solution, and conditions of 10 mM sodium phosphate buffer plus 0.35–0.65 M sodium chloride at pH 7.5 and 10–25% (v/v) Jeflamine M-600 gave the best results. Selenomethionine derivative crystals could be grown by the hanging drop vapor diffusion method at 20 °C from 10 mM sodium phosphate buffer, pH 7.5, sodium chloride between 0.4 M and 1.2 M, 20–25% Jeflamine M-600 (v/v), plus 10 mM dithiothreitol. Needle-like crystals with dimensions of 0.7 × 0.01 × 0.01 mm<sup>2</sup> were observed for the native protein. The crys-
lals (thus diffusing to space group P6<sub>5</sub> from 2.4 Å resolution with cell

dimensions of a = b = 109.29 Å and c = 77.91 Å, and the value of the

Matthews coefficient (14) suggests that there are two molecules in the asymmetric unit (V<sub>M</sub> = 2.8 Å<sup>3</sup>/Da with a solvent content of 54%). The morphology and the size of selenomethionine derivative crystals were similar to those of the native crystals, and they diffracted to a maximum resolution of 2.9 Å. The native data set was collected on the European Molecular Biology Laboratory beamline X13 (Deutsches Elektronen-Synchrotron, Hamburg, Germany), and selenomethionine derivative data sets were collected at the European Molecular Biology Laboratory beamlines BW7A and X11 as well as at the European Synchrotron Radiation Facility beamline BM14 (Grenoble, France). A selenomethionine-containing crystal soaked with ssRNA, described under “ssRNA Soaking Experiments,” was used for data collection on the European Molecular Biology Laboratory X11 beamline. Data were reduced, merged, and scaled using the programs DENZO and SCALE-

PACK, respectively (15). Details of all crystallographic data collections are given in Table I.

ssRNA Soaking Experiments—A synthetic oligonucleotide having the rotavirus 5’ consensus sequence 5’-GCCUUUAAG-3’ was cleaved from a solid support, and all of the protecting groups were removed according to the supplier’s protocol (CRUACHEM, UK). The concentra-

tion of the RNA solution was 42 μM. Previous binding experiments showed that a suitable RNA binding buffer was 10 mM HEPES, pH 7.8, 40 mM potassium chloride, 1 mM EDTA, 1 mM dithiothreitol, and 5% glycerol. Crystals of selenomethionine-containing protein were grown from 10 mM HEPES buffer, pH 7–8, 0.7–0.9 M potassium chloride, 10 mM dithiothreitol, and 20–25% (v/v) Jeflamine M-600, and soaking experiments were performed by adding between 2 and 4 μl of ssRNA solution to the crystal-containing drops.

Structure Determination and Refinement—The crystal structure of the RNA binding domain of NS2 (Fig. 1) was solved by the SIRAS method using the 4.0 Å data set collected at the selenium absorption edge (0.973 Å). Fourteen of sixteen possible selenium sites were located. Eleven sites were identified using the program SnB (16), and three selenium atoms were refined after analyzing the residual maps produced by SHARP (17). Positons and occupancies were refin-

ed with the program MLPHARE (18) and further in the program

SHARP, which was also used to generate SIRAS phases. The final SIRAS phases had a figure of merit of 0.51 for the entire resolution range (18–4.0 Å). Density modification procedures, including solvent flattening and two-fold non-crystallographic symmetry averaging of the experimental map were used to extend the phases to 2.4 Å in the program RESOLVE (19). The non-crystallographic symmetry operator was determined with RESOLVE from the identified atom positions.

To obtain experimental phases to a higher resolution than 4.0 Å (thus including the initial basis of the space extension procedure and consequently obtaining a superior map) a three-wavelength anomalous dispersion data set was collected to a 2.9 Å resolution, and phases were determined using SOLVE (20). The initial phases had a figure of merit of 0.65 in the resolution range from 20.0 to 3.0 Å. Further phase improvement and extension to 2.4 Å using solvent flattening and two-

fold non-crystallographic symmetry averaging were carried out in the program RESOLVE. Two molecules in the asymmetric unit were built in the electron density map obtained from the three-wavelength anom-

alous dispersion data using the program O (21). Restrained refinement of coordinates and temperature factors was carried out using the program REFMAC5 (22). Bulk solvent correction, different types of non-

crystallographic symmetry and screw-rotation refinement were used during the refinement calculations with REFMAC5. Waters

C. Butan, H. van der Zant, and P. A. Tucker, unpublished results.
Crystal Structure of the RNA Binding Domain of NS2 from BTV

Table I
Summary of data collection and refinement statistics

|                | Native1 | SeMet1 | SeMet2 | SeMet2 | SeMet2 | SeMet3/RNA soak |
|----------------|---------|--------|--------|--------|--------|-----------------|
| Data statistics |         |        |        |        |        |                 |
| Synchrotron source | X13, EMBL | BM14, ESRF | BW7A, EMBL | BW7A, EMBL | BW7A, EMBL | X11, EMBL     |
| Wavelength (Å)    | 0.803   | 0.975  | 0.9785 | 0.9785 | 0.9897 | 0.841          |
| Space group       | P6      | P6     | P6     | P6     | P6     | P6             |
| Unit cell (Å)     | a = b = 102.29 c = 77.91 | a = b = 102.16 c = 77.67 | a = b = 102.16 c = 77.65 | a = b = 101.85 c = 77.40 | a = b = 101.85 c = 77.40 | a = b = 102.94 c = 72.27 |
| Resolution (Å)    | 20–2.4  | 18.0–4.0 | 20–2.9  | 20–2.9  | 20–2.9  | 20–3.0         |
| Resolution (Å)    | 2.46–2.40 | 4.07–4.00 | 2.97–2.90 | 2.97–2.90 | 2.97–2.90 | 2.97–3.00 |
| Rmerge (%)        | 5.9 (36.5) | 10.0 (15.5) | 13.3 (38.4) | 12.43 (4.98) | 16.67 (5.63) | 18.99 (6.43) |
| Complete (%)      | 99.9 (98.8) | 99.8 (99.5) | 98.6 (92.6) | 99.3 (99.6) | 99.8 (96.7) | 99.4 (97.3)  |
| No. of protein/solvent atoms | 2369 (103) | 24.73    | 0.012    | 1.265    | 91.8   | 8.2           |
| r.m.s.d. bond lengths (Å) | 2.97 | 2.90 | 2.97 | 2.97 | 2.97 | 2.97 |
| r.m.s.d. angles (%) | 91.8 | 91.8 | 91.8 | 91.8 | 91.8 | 91.8 |
| Multiplicity      | 9.4     | 7.0    | 6.7    | 7.0    | 9.7    | 5.5            |

Refinement statistics

|                |         |        |        |        |        |                 |
|----------------|---------|--------|--------|--------|--------|-----------------|
| Resolution range | 20–2.4  |        |        |        |        |                 |
| Number of reflections | 18113 |        |        |        |        |                 |
| Rmerge (%)        | 21.1%   |        |        |        |        |                 |
| Residues          | 306     |        |        |        |        |                 |
| Number of protein/solvent atoms | 2369 (103) | 24.73 | 0.012 | 1.265 | 91.8 | 8.2 |
| Average B-factor (Å²) | 2.97 | 2.90 | 2.97 | 2.97 | 2.97 | 2.97 |
| r.m.s.d. angles (%) | 91.8 | 91.8 | 91.8 | 91.8 | 91.8 | 91.8 |
| Multiplicity      | 9.4     | 7.0    | 6.7    | 7.0    | 9.7    | 5.5            |

RESULTS

Domain Identification—The full-length protein has been initially expressed using recombinant baculovirus according to the procedures described by Thomas et al. (5) and was purified to homogeneity as described under “Experimental Procedures.” In the presence of protease inhibitors, crystallization trials failed either because the domain structure of the protein introduced conformational heterogeneity or more probably because the protein was heterogeneously aggregated. When protease inhibitors were not used, needle-like or chunky crystals were obtained and diffractioned to a 3.7 Å maximum resolution. Data (not shown) were collected to low resolution and analyzed, and they revealed a hexagonal space group with unit cell dimensions a = b = 104 Å and c = 81.7 Å. SDS-PAGE showed that different crystals contained varying amounts of full-length protein and proteolytic products, but that despite this, the diffraction pattern was the same. This suggested that only the major proteolytic product was ordered in the crystals. We also observed that proteolysis could be accelerated by adding trypsin.

Mass spectroscopy analysis of the crystals indicated that the first 177 residues formed the major proteolytic product, which was also necessary and sufficient for RNA binding. Constructs of residues 1–177 were expressed with His6 or glutathione S-transferase tags at the C terminus using recombinant baculovirus-infected insect cells or expressed in E. coli with a His6 tag at the N and C termini for the purpose of crystallization. Although expression of the 178–354-residue domain resulted in soluble material, 2 we could not obtain soluble protein for the

---

**Fig. 1. Structure of the NS2<sub>α</sub> monomer.** A, ribbon representation of the NS2<sub>α</sub> monomer with the two layers of the β sandwich fold colored light gray and dark gray. The remaining structural elements are in black. B, topology diagram of the secondary structure elements of the NS2<sub>α</sub> monomer. The arrows represent β strands, and the cylinders represent α helices.
1–177-residue domain. NS2 is a phosphoprotein, and we have shown by ion exchange chromatography the presence of two species, possibly in different phosphorylation states when expressed using recombinant baculovirus-infected insect cells. Separation and crystallization of the individual species did not improve the crystal quality. As it was not possible to obtain soluble protein for the 1–177-residue construct, we decided that crystallization under proteo-
lytic conditions would be the method of choice. Crystals of the N-terminal domain (NS2_N) diffracting to higher resolution were obtained after engineering a TEV protease cleavage site between residues 182 and 183 of the full-length protein (Fig. 2A). Because the polypeptide region beyond residue 177 was predicted to be unstructured, we decided to use the next five residues of the sequence as a linker to the engineered TEV cleavage site. The insertion of this cleavage site resulted in a better control of the proteolytic cleavage process during crystallization, perhaps because TEV is a highly specific and only moderately active protease. The construct was expressed in *E. coli* to avoid previously identified heterogeneities due to variable phosphorylation. It was shown that TEV protease cleaves this construct and that small crystals could be obtained when this was done.

The Overall Structure—The structure of NS2_N was determined by anomalous scattering methods using crystals of selenomethionine-containing protein. The model presented here was refined to a resolution of 2.4 Å with an *R*-factor of 21.1% and a free *R* value of 26.7%. The domain crystallized with a dimer in the crystallographic asymmetric unit. The refined model contained residues 8–160 of both molecules and 103 water molecules. Residues 161–182, the six additional residues of the TEV cleavage site insertion, and residues 1–7 of the N terminus were not seen in the electron density map and were presumed to be disordered in the crystal. The refinement statistics as well as an evaluation of the quality of the current model are shown in Table I.

Each monomer exhibits a mixed β sandwich structure (Fig. 1A). One layer consists of β1, β2, β3, β4, β8, and β9. A short helical region (3_10 and α1) connects β1 with β2. Almost matching the length of one dimension of the monomer, the β8 strand intertwines with the β9 strand, and both are twisted. β8 forms a small antiparallel β sheet with β10 from the C terminus. The other layer consists of β5, β6, and β7. The β4 strand links the two β sheets forming a highly stable structural core (Fig. 1A). The C terminus of the construct (Met134-Arg160) extends out 22 Å from the body of the monomer making contacts with an adjacent subunit in the crystal. In this respect, NS2_N is similar to another RNA-binding protein, the ribosomal protein S15 from *Bacillus stearothermophilus* (28), which consists of a core and an N-terminal segment that protrudes 36 Å from the core of the protein and interacts with a neighboring monomer. No other proteins were found to exhibit a similar topology (Fig. 1B) when searching for structural similarity in the Protein Data Bank using the DALI server (29).

Structure of the Dimers and Dimer Interfaces—The two molecules in the asymmetric unit are related by a non-crystallographic two-fold axis and are superimposable with a root mean square deviation of 0.41 Å over all Ca atoms of residues 8–160. There are two possible descriptions for the homodimeric con-
figuration: monomers interact 1) through their C termini giving rise to an interface, which buries almost 20% of the total monomer surface (1831 Å²) (Fig. 3A), and 2) through a continuation of β sheet structure (Fig. 3B), which results in a buried surface area of 1075 Å² per monomer.

The formation of the largest buried surface (Fig. 3A) involves an extension of the C terminus (residues 134–160) folding over the adjacent molecule. The homodimer is maintained by hydrogen bonding interactions, hydrophobic packing, and a symmetrical salt bridge interaction. Asn134, which is absolutely conserved throughout the NS2 family, located on /H9252 interacts with Thr /H11032 located on /H9252 of the second monomer (Fig. 4A). Furthermore, the interaction between Thr /H102 and Tyr /H121 (β’9), which is mediated through a water molecule, places the C terminus of one monomer in interaction with the structural core of the adjacent monomer. The main chain of Pro /H145 makes a hydrogen bond with the hydroxyl group of conserved Tyr /H34 (β’2) from the neighboring monomer (Fig. 4A). Met /H148 lies between 3/10 and α2 and interacts through main chain hydrogen bonds with Arg /H107, which lies at the end of β’8 of the second monomer (Fig. 4B). In addition, the side chain of Ser /H150 is engaged in a hydrogen bonding interaction with the side chain of Asp /H116, located on β’9 (Fig. 4B). The orientation of the conserved Arg /H158 appears to be constrained by a salt bridge interaction with conserved Glu /H118 (β’9) (Fig. 4B). A number of hydrophobic interactions are involved in homodimerization. For example, conserved residue Leu /H144 and conservatively replaced residue Val /H147 are facing conserved residues Met /H104 (β’8), Val /H57 (β’4) and Ile /H119 (β’9). Of 27 residues involved in the interface formation, 11 are conserved, and 4 are conservatively replaced within the members of the NS2 family (Fig. 2B). A large pocket harboring water molecules is found at the 1831 Å² homodimeric interface. The network of hydrogen bonds between the water molecules and oxygens and nitrogens of residues pointing toward this interface strongly stabilizes the interaction between monomers.

**Fig. 4. Interactions at the homodimer interfaces.** A and B, diagram showing the interactions made by residues from the C terminus of one monomer (in violet) with side chains from the other monomer (in green). Hydrogen bonding interactions and a salt bridge interaction are indicated by thin lines. C and D, ribbon diagrams showing the hydrogen bonding interactions (indicated by dashed lines) at the β sheet interface viewed from opposite sides of the sheet.
The second homodimeric interface (Fig. 3B), the driving force for propagating the NS2N assembly, is mediated by formation of continuous β sheets. Upon dimerization, extended β sheets formed by eight (5 + 3) strands are observed on both sides of the homodimer. Although the monomer area buried at this interface is less than the buried area at the first interface, the strength of the interactions at the second homodimeric interface is substantial. The homodimerization determinants reside within the hydrogen bonding interactions between the main chains of residues located on β9252 and β9257 of one monomer with the main chains of residues located on β11032 and β9253 of the second subunit, resulting in a very stable structure (Fig. 4, C and D).

Oligomeric State—According to recent sedimentation experiments (2), the full-length protein in solution exists as an 8–10 S multimer with a molecular mass between 140 and 250 kDa and assembles from 6 ± 2 subunits. Our preliminary small angle-scattering experiments (data not shown) on the full-length protein suggest a molecular mass corresponding to about a decamer. In the crystal of the N-terminal domain, repetition of the two monomer-monomer interfaces, described in the previous section under “Structure of the Dimers and Dimer Interfaces,” gives rise to a helical structure in the crystal with a pitch equal to the length of the c axis (77.91 Å). The helical arrangement is generated through the application of the crystallographic 65 symmetry operator on two non-crystallographic symmetry-related monomers that form the asymmetric unit. Of the entire surface area of a single subunit, which is ∼9150 Å², 31% is used to create the interfaces that bring the subunits together. The helical structure in the crystal is infinite and has a central channel with a diameter of 76.7 Å (Fig. 3, C and D). Assuming that the volume occupied by the C-terminal domain is about equal to the volume of the present N-terminal construct, we find that the interior channel of the helical structure has insufficient volume to accommodate the C-terminal domains of each of the monomers, let alone the C-terminal domains plus some RNA. It could, however, accommodate some C-terminal domains, explaining why we observed the N-terminal domain with varying amounts of full-length protein on the SDS-PAGE of our original crystals. It seems likely that the N-terminal domain of NS2 alone drives the oligomerization, a conclusion further substantiated by the observation that the C-terminal domain is monomeric in solution.2

A notable feature of the helical structure is the presence on its surface of a high number of conserved solvent-accessible residues. Only 30% of the absolutely conserved residues are involved in the hydrophobic core of the protein. Large clusters
of proximally located conserved residues are observed on the exposed area of a single subunit and on the interior surface of the helical assembly. One area of conserved residues, located on the interior surface of the monomer and projecting into the solvent channel, consists of residues Lys158, Pro159, Lys160, Tyr161, Glu157, Trp91, and Arg158 as well as residues which have been shown to be implicated in RNA binding, namely Phe6, Thr7, and conservatively replaced Lys10. Positively charged RNA binding residues, Arg155, Arg158, and Lys160 from a neighboring monomer are located nearby (Fig. 5A). Another conserved patch includes acidic residues Glu25, Glu26, and Glu28 but also Thr7, Arg38, Pro159, Tyr161, Asn158, Ile159, Ile160, and Arg41 as well as two solvent-exposed tryptophan side chains, Trp31 and Trp34 (Fig. 5B). Asn158 bridges two water molecules infiltrated at the interface, interacting with Trp34 as well as with the conserved but not solvent-exposed Arg67. It is frequently reported that conserved and solvent-accessible residues are involved in interactions with ligands or other proteins. NS2 may interact with other BTV proteins, namely with the proteins from the transcriptase complex (VP1, VP4, VP6) (30) of the virus before genome encapsidation, and it is tempting to speculate that the solvent-exposed conserved residues will mediate those interactions.

**RNA Interaction Surface—**NS2 is reported to bind BTV ssRNAs as well as other ssRNA species in a sequence-independent manner indicating that the binding motif of the ssRNA is non-specific (7, 31), although not necessarily independent of the RNA structure (30). It has been demonstrated that multiple recombinant NS2 molecules bind to transcripts of rotavirus gene 8 ssRNA, suggesting that the RNA binding may require homomultimer formation (2).

In the crystal structure, both the N and C termini of the construct extend into the center of the helical structure, projecting into the central solvent channel (Fig. 3D). Although these termini are disordered, they contain residues that have been identified by mutational analysis as important for RNA binding. The absence of electron density for residues 1–7 and 161–182 implies that these residues have conformational flexibility. Such conformational flexibility would be in agreement with other data suggesting that 11 residues (2–11) (6) at the beginning of the construct and 14 residues (153–166) (7) at the C terminus of the construct are important for the RNA binding function of the protein. We therefore propose that RNA binds to the inner concave side of the helical assembly, as all the missing residues important for the RNA binding are exposed in this area (Fig. 3D). They could conceivably become more ordered upon RNA binding. The RNA soaking experiment of a crystal does not show increased order, but it does show a substantial change (Table I, 5 Å) in the pitch of the helix, suggesting that there is a degree of flexibility and consequent relative domain movement in the solid state. The region determining the relative positions of molecules within the helical structure is around residues 134–138 and is located before the largest homodimerization interface. These findings strengthen the previous observation that the largest homodimerization interface does not provide the same rigidity as the extended β sheet interface.

Guided by the distribution of the residues, which are likely to be involved in RNA binding, the putative RNA interaction surface on a homodimer can be delimited into two symmetrical regions made up of the N-terminal part of β1, the C-terminal part of β2, the loop connecting β2 with β3 (where completely conserved Arg41 is placed), and two residues from β9, α2 and the C-terminal extension (Arg158-Lys160) from the adjacent monomer (Fig. 6). It is conceivable that an extended region of ssRNA will non-specifically interact with the two sites symmetrically placed on the homodimer and will flexibly fit into the curved area between the two sites.

Mutagenesis of Glu2-Lys11 points to the importance of this region for RNA binding (6). Two arginines, Arg6 and Arg7, when mutated to leucines, result in a reduction in the affinity of the protein toward ssRNA, whereas when Lys6 is mutated to leucine, this results in a total abrogation of the ssRNA binding (6). The glutamic acid Glu2 was not required for the ssRNA interaction surface on a homodimer can be delimited into two symmetrical regions made up of the N-terminal part of β1, the C-terminal part of β2, the loop connecting β2 with β3 (where completely conserved Arg41 is placed), and two residues from β9, α2 and the C-terminal extension (Arg158-Lys160) from the adjacent monomer (Fig. 6). It is conceivable that an extended region of ssRNA will non-specifically interact with the two sites symmetrically placed on the homodimer and will flexibly fit into the curved area between the two sites.
Crystal Structure of the RNA Binding Domain of NS2 from BTV

This structural study increases the number of known ssRNA binding scaffolds and contributes to the characterization of the ssRNA-binding proteins. Known ssRNA binding folds are the ribonucleoprotein fold, observed in sex-lethal protein (40) and human polyanamine binding protein (41) proteins or the oligonucleotide and oligosaccharide binding fold, observed in the transcription factor Rho (27). All of these proteins have a common mode of ssRNA recognition, namely through β sheet binding, where one or multiple β sheets form binding pockets for the oligonucleotide bases.

Sequence comparison among the members of the NS2 family indicates that the first half of the protein (residues 1–182) is very well conserved. Conservation of NS2N within different orbiviruses (EHDV2, AHSV9, BTV) and different serotypes of BTV (BTV10, BTV17, BTV18, BRV1X) might be a reflection of the constraints on the protein to maintain the NS2N function. It is highly likely that the other members of the NS2 family have three-dimensional structures similar to that reported here.

Acknowledgments—We thank Prof. Polly Roy (Department of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine) for the initial gift of recombinant baculovirus. We also thank the BM14 support staff for assistance.

REFERENCES

1. Devaney, M. A., Kendall, J., and Grubman, M. J. (1988) Virus Res. 11, 151–164
2. Taraporewala, Z. F., Chen, D., and Patton, J. T. (2001) Virology 290, 221–231
3. Brooks, S. M., Hyatt, A. D., and Eaton, B. T. (1993) J. Gen. Virol. 74, 525–530
4. Epstein, B. T., Hyatt, A. D., and Brooks, S. M. (1990) Curr. Top Microbiol. Immunol. 162, 89–119
5. Thomas, C. P., Booth, T. F., and Roy, P. (1990) J. Gen. Virol. 71, 2073–2083
6. Zhao, Y., Thomas, C., Brenner, C., and Roy, P. (1994) J. Virol. 68, 2179–2185
7. Fillingmore, G. C., Lin, H., and Li, J. K. (2002) J. Virol. 76, 499–506
8. Nagai, K. (1996) Curr. Opin. Struct. Biol. 6, 53–61
9. Puglisi, J. D., Tan, R., Calnan, B. J., Franklin, A. D., and Williamson, J. R. (1992) Science 257, 76–80
10. Birney, E., Kumar, S., and Krainer, A. R. (1993) Nucleic Acids Res. 21, 5803–5816
11. Taraporewala, Z., Chen, D., and Patton, J. T. (1999) J. Virol. 73, 9934–9943
12. Gillian, A. L., Schmechel, S. C., Livny, J., Schiff, L. A., and Nibert, M. L. (2000) J. Virol. 74, 5939–5948
13. Taraporewala, Z. F., Schuck, P., Ramig, R. F., Silvestri, L., and Patton, J. T. (2002) J. Virol. 76, 7082–7093
14. Matthews, B. W. (1968) J. Mol. Biol. 33, 491–497
15. Otwinowski, Z., and Minor, W. (1997) Methods Enzymol. 276, 307–326
16. Weeks, C. M., and Miller, R. (1999) Acta Crystallogr. Sect. D 55, 482–500
17. de La Fortelle, E., and Briogno, G. (1997) Methods Enzymol. 276, 472–494
18. Bailey, S. (1994) Acta Crystallogr. Sect. D 50, 760–763
19. Perrakis, A., Morris, R., and Lamzin, V. S. (1999) Acta Crystallogr. Sect. D 55, 865–972
20. Terwilliger, T. C., and Berendzen, J. (1999) Acta Crystallogr. Sect. D 55, 849–861
21. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard, M. (1991) Acta Crystallogr. Sect. A 47, 110–119
22. Mutton, G. N., Vagin, A. A., and Dodson, E. J. (1997) Acta Crystallogr. Sect. D 53, 240–255
23. Ferrarri, A., Morris, R., and Lammam, Y. S. (1999) Nat. Struct. Biol. 6, 458–463
24. Laskowski, R. A., Macarthur, M. W., Moss, D. S., and Thornton, J. M. (1993) J. Appl. Crystallogr. 26, 858–950
25. Burton, G. J. (1991) Protein Eng. 4, 37–40
26. Bogden, C. E., Fass, D., Bergman, N., Nichols, M. D., and Berger, M. J. (1999) Mol. Cell 3, 487–493
27. Clemons, W. M., Jr., Davies, C., White, S. W., and Ramakrishnan, V. (1998) Structure 6, 429–438
28. Holm, L., and Sander, C. (1993) J. Mol. Biol. 233, 123–138
29. Lyneopheropulus, K., Wirlich, C., Brierley, I., and Roy, P. (2003) J. Biol. Chem. 278, 31722–31730
30. Huismans, H., van Dijk, A. A., and Bauskin, A. R. (1993) J. Gen. Virol. 74, 2589–2595
31. Grimes, J. M., Burroughs, J. N., Guet, P., Digre, J. M., Maley, R., Zentara, S., Mertens, P. F., and Stuart, D. I. (1998) Nature 395, 470–478
32. Grimes, J. M., Burroughs, J. N., Guet, P., Digre, J. M., Maley, R., Zentara, S., Mertens, P. F., and Stuart, D. I. (1999) Cell 97, 481–490
33. Handa, N., Nureki, O., Kurimoto, K., Kim, I., Sakamoto, H., Shimura, Y., Muto, Y., and Yokoyama, S. (1999) Nature 398, 579–585
34. Deo, R. C., Bonanni, J. B., Sonenberg, N., and Burley, S. K. (1999) Cell 98, 835–845