doi:10.1006/jmbi.2000.4250 available online at http://www.idealibrary.com on IDEAL J. Mol. Biol. (2000) 304, 861–871

JMB

Crystal Structure of the Oligomerization Domain of NSP4 from Rotavirus Reveals a Core Metal-binding Site

Gregory D. Bowman[1,2]†, Ilana M. Nodelman[1,2]†, Orlie Levy[3], Shuo L. Lin[3] Peng Tian[3], Timothy J. Zamb[3], Stephen A. Udem[3] Babu Venkataraghavan[3] and Clarence E. Schutt[2]*

[1]*Department of Molecular Biology, Lewis Thomas Laboratories, Princeton University, Princeton NJ 08544, USA*

[2]*Department of Chemistry Henry H. Hoyt Laboratory Princeton University Princeton, NJ 08544, USA*

[3]*Viral Vaccine Research Wyeth-Lederle Vaccines 401 North Middletown Road Pearl River, NY 10965, USA*

**Corresponding author*

During the maturation of rotaviral particles, non-structural protein 4 (NSP4) plays a critical role in the translocation of the immature capsid into the lumen of the endoplasmic reticulum. Full-length NSP4 and a 22 amino acid peptide ($NSP4_{114-135}$) derived from this protein have been shown to induce diarrhea in young mice in an age-dependent manner, and may therefore be the agent responsible for rotavirally-induced symptoms. We have determined the crystal structure of the oligomerization domain of NSP4 which spans residues 95 to 137 ($NSP4_{95-137}$). $NSP4_{95-137}$ self-associates into a parallel, tetrameric coiled-coil, with the hydrophobic core interrupted by three polar layers occupying *a* and *d*-heptad positions. Side-chains from two consecutive polar layers, consisting of four Gln123 and two of the four Glu120 residues, coordinate a divalent cation. Two independent structures built from MAD-phased data indicated the presence of a strontium and calcium ion bound at this site, respectively. This metal-binding site appears to play an important role in stabilizing the homo-tetramer, which has implications for the engagement of NSP4 as an enterotoxin.

Keywords: rotavirus; nsp4; ns28; fusion; parallel coiled-coil

Introduction

Rotavirus is the leading cause of severe gastroenteritis during the first few years of life, resulting in more than 870,000 deaths annually in the third world (Glass *et al*., 1994). A member of the *Reoviridae* family, rotavirus possesses a non-enveloped, triple-layered protein coat encapsulating a double-stranded RNA genome. During infection, the immature capsid particle (ICP) buds into the lumen of the endoplasmic reticulum (ER) (for review see, Estes, 1991). During this process a lipid coat is acquired which is shed prior to the exit of the mature virion.

NSP4, a virally-encoded receptor embedded in the membrane of the ER, is required for the budding of the ICP into the ER (Au *et al*., 1989). Most of the NSP4 protein is cytoplasmic, with the exception of a 25-residue transmembrane region, H2 (Bergmann *et al*., 1989), which joins to H1, a small, doubly-glycosylated, lumenal domain (Figure 1(a); Both *et al*., 1983; Kabcenell & Atkinson, 1985). Crosslinking experiments have shown that NSP4 oligomerizes into dimers and tetramers (Maass & Atkinson, 1990; Taylor *et al*., 1996). Heptad repeats spanning residues 95-137 suggest that oligomerization occurs *via* an α-helical coiled-coil interaction (Taylor *et al*., 1996). This coiled-coil oligomerization domain is resistant to protease degradation, unlike residues C-terminal to this putative stalk domain (Taylor *et al*., 1996; O'Brien *et al*., 2000). In addition, the extreme carboxyl-terminal 20 amino acid residues of NSP4 have been shown to be the minimal binding domain for interacting with VP6, the major structural protein of the ICP (Meyer *et al*., 1989; Taylor *et al*., 1996; O'Brien *et al*., 2000).

†These authors contributed equally to this work.

Abbreviations used: ASU, asymmetric units; NSP4, non-structural protein 4; ICP, immature capsid particle; MAD, multi wavelength anomalous dispersion; PEG, polyethylene glycol.

E-mail address of the corresponding author: schutt@chemvax.princeton.edu

0022-2836/00/050861–11 $35.00/0

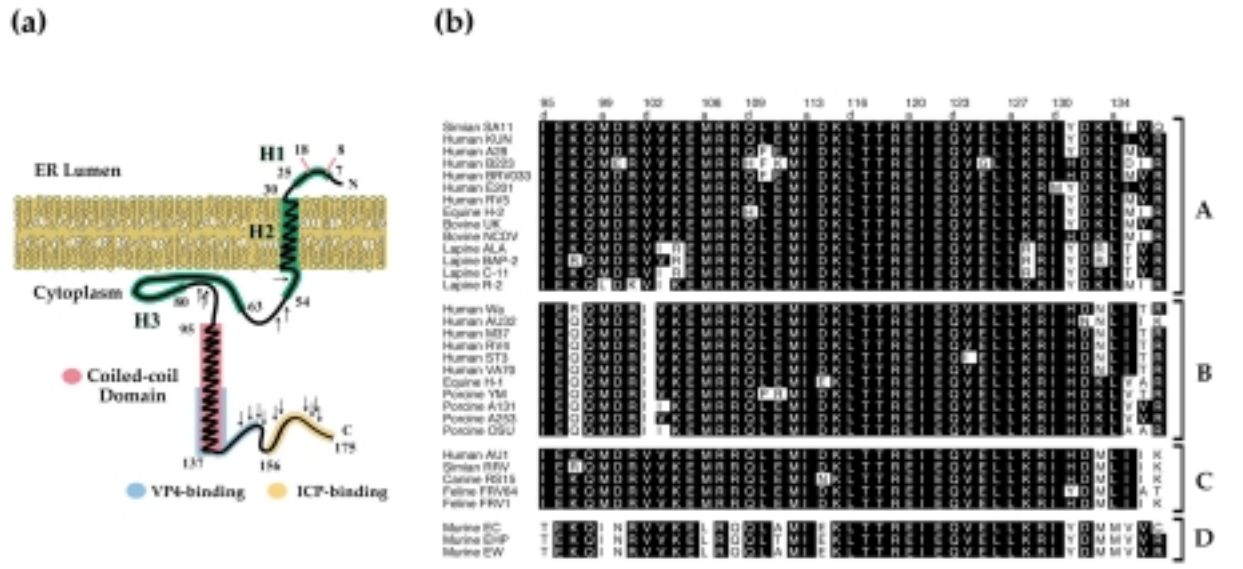

Figure 1. (a) Schematic of NSP4. Three hydrophobic regions (H1, H2, and H3; Bergmann *et al*., 1989) are highlighted in green. H2 anchors NSP4 in the lipid bilayer of the ER, and H1 resides in the ER lumen with glycosylation sites at amino acid positions 8 and 18 (Both *et al*., 1983; Kabcenell & Atkinson, 1985). The H3 domain (residues 63-80; green), the coiled-coil oligomerization domain (residues 95-137; pink; Taylor *et al*., 1996), and the ICP-binding domain (residues 156-175; yellow; O'Brien *et al*., 2000; Taylor *et al*., 1996) reside in the cytoplasm. Residues 112-148 (light blue) have been shown to be critical for VP4 binding (Au *et al*., 1993), although a minimal binding site has not been defined. Arrows indicate proteolytically-sensitive sites (Bergmann *et al*., 1989; O'Brien *et al*., 2000). (b) A sequence alignment of $NSP4_{95\text{-}137}$ from a sampling of various rotaviral strains. The sequence used for crystallization was that of SA11, a strain widely employed in functional studies. The strains are categorized according to their genotype (A-D; Ciarlet *et al*., 2000; Horie *et al*., 1999). The *a* and the *d*-positions of the heptad repeat are denoted above the first sequence. The alignment was generated using ALSCRIPT (Barton, 1993).

In addition to acting as a receptor for the ICP, NSP4 plays several important roles during viral maturation. Upon budding of the ICP into the ER, NSP4 hetero-oligomerizes with the two outer-coat proteins, VP4 and VP7 (Maass & Atkinson, 1990). Deletion mutagenesis of NSP4 demonstrates that some or all of residues 112-148 are critical for VP4 binding (Au *et al*., 1993). Moreover, deletion of 27 amino acid residues encompassing the VP6 binding site at the C terminus of NSP4 noticeably increases the affinity for VP4, suggesting that the interaction of NSP4 with VP4 may be coordinated with ICP binding (Au *et al*., 1993). After budding is completed, the subsequent shedding of the lipid envelope is dependent on the glycosylation of NSP4, as treatment with the glycosylation-inhibitor tunicamycin blocks this uncoating (Poruchynsky *et al*., 1991).

The gastrointestinal symptoms resulting from rotavirus infection may be directly attributable to NSP4. Injection of full-length NSP4 or a peptide corresponding to residues 114-135 ($NSP4_{114\text{-}135}$) into young mice results in diarrhea as a consequence of an efflux of chloride ions from intestinal cells *via* a calcium-dependent signaling pathway (Ball *et al*., 1996). Baculovirus expression of NSP4 in insect cells or the exogenous treatment of uninfected insect cells with $NSP4_{114\text{-}135}$ triggers the release of calcium from the ER (Tian *et al*., 1995, 1994). The difficulty in establishing stable cell lines expressing NSP4 might be due to cytotoxic levels of calcium (Tian *et al*., 1994).

NSP4 is a model system for studying how non-enveloped virus particles translocate intracellularly between compartments. We have determined the structure of the coiled-coil domain of NSP4 in order to gain insight into the enterotoxic and oligomerization properties of this unusual virally encoded receptor molecule.

Results

Structure determination

A peptide corresponding to residues 95-137 ($NSP4_{95\text{-}137}$; Figure 1(b)) with a single selenomethionine residue at position 112 was synthesized for structural studies. Diffraction-quality crystals were obtained at pH 8.5 using 20-28 % PEG 400 and a variety of salts between 0.32-0.40 M (see Methods). Multiwavelength data sets were collected for crystals grown in magnesium sulfate and strontium chloride (Table 1) and independently-phased electron density maps were obtained for each. In both of these MAD-phased (Hendrickson, 1991) maps, helical density allowed for unambiguous placement of all main-chain and the majority of side-chain atoms (Figure 2).

Table 1. Crystallographic and refinement statistics

Crystal	λ(Å)	*a*(Å)	*b*(Å)	*c*(Å)	d_{min}(Å)	Completeness(%)[b]	$\langle I \rangle / \langle \sigma \rangle$	R_{sym} (%)[d]
A. *Data collection*								
Calcium[a]		35.47	35.40	60.94				
λ_1 (peak)	0.9785				2.0	99.9	11.7	4.5
λ_2 (edge)	0.9788				2.0	99.9	11.0	4.8
λ_3 (remote)	0.9500				2.4	100	8.9	5.7
Strontium		34.92	35.91	60.94				
λ_1 (peak)	0.9787				1.86	100	19.3	5.1
λ_2 (edge)	0.9789				1.86	99.9	18.8	4.9
λ_3 (remote)	0.9500				2.4	100	25.5	4.7

B. *Phasing statistics (15-2.3 Å)*	Isomorphous		Anomalous	
	R_{cullis} [e]	Phasing power[f]	R_{cullis} [e]	Phasing power[f]
Calcium[a]				
λ_1 (peak)	0.98	0.34	0.65	1.85
λ_2 (edge)	NA[c]	NA[c]	0.48	2.88
λ_3 (remote)	0.79	1.09	0.62	1.98
Strontium				
λ_1 (peak)	0.99	0.21	0.76	1.25
λ_2 (edge)	NA[c]	NA[c]	0.44	2.80
λ_3 (remote)	0.76	1.22	0.62	1.86

C. *Refinement statistics*	Calcium[a]	Strontium
Resolution range (Å)	14-2.00	20-1.86
No. of reflections	9863	12,450
R_{work} (%)[g]	23.5	21.6
R_{free} (%)[h]	24.5	23.9
Test set size (%)	8.5	10.1
No. of atoms		
Protein	694	737
Solvent	67	89
r.m.s. deviation from ideality		
Bond lengths (Å)	0.010	0.006
Bond angles (deg.)	1.0	0.9
Temperature factor average		
Main-chain (Å^2)	24.3	21.0
Side-chain (Å^2)	29.0	26.4
Water (Å^2)	36.9	37.5
Metal (Å^2)	37.1	19.3

[a] These crystals were grown in magnesium sulfate as described in the Results.
[b] These values include the entire resolution range as defined in Refinement statistics.
[c] Not applicable as this was the reference wavelength.
[d] $R_{sym} = \Sigma|I_i - \langle I \rangle| / \Sigma\, I_i$ where I_i is the intensity of an individual reflection and $\langle I \rangle$ is the mean intensity of that reflection for unmerged Friedel pairs.
[e] $R_{cullis} = \varepsilon / \langle |F_{PH} - F_P| \rangle$, where ε is the phase-integrated lack of closure.
[f] Phasing power $= \langle |F_H(calc)| \rangle / \varepsilon$.
[g] $R_{cryst} = \Sigma||F_o| - |F_c|| / \Sigma|F_o|$, where F_o and F_c are the observed and calculated structure factors, respectively.
[h] R_{free} was calculated as R_{work} and these reflections were omitted from refinement.

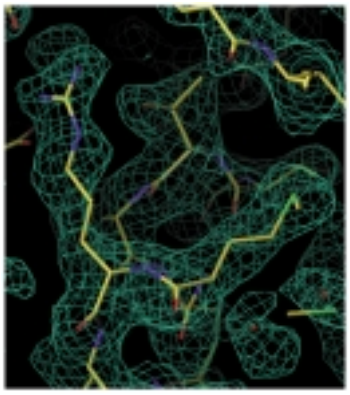

Figure 2. An initial 2.3 Å resolution electron density map of $NSP4_{95\text{-}137}$, obtained by multi-wavelength anomalous dispersion (MAD) phasing and solvent flipping, is contoured at 1σ. The final model of the $NSP4_{95\text{-}137}$ is shown as rods. This Figure was produced using the program O (Jones *et al*., 1991).

Structural overview

$NSP4_{95\text{-}137}$ packs as a parallel four-stranded coiled-coil (Figure 3(a)). The asymmetric unit of the crystal contains two molecules (A and B) which are related to two apposing molecules (A′ and B′) by a crystallographic 2-fold axis. These four molecules define a pseudo-4-fold symmetry axis which superimposes on the crystallographic 2-fold axis. The helical region of each chain spans residues 95 to 134 with the final three residues in an extended conformation. Superposition of main-chain C^{α} atoms between residues 95 to 134 of chain A onto B results in an r.m.s. deviation of 0.37 Å, showing that the helical characteristics of each monomer are similar. The coiled-coil region of the tetramer has a supercoil pitch of ~178 Å, and can be considered a cylinder with a length of ~58 Å and a diameter of ~26 Å. The tetrameric core of the $NSP4_{95\text{-}137}$ oligomer contains three polar layers. A layer of Gln109 residues occupies the *d*-position of the heptad repeat, and a more extensive polar region at the *a* and *d*-heptad repeat positions is formed by the Glu120 and Gln123 layers.

The exterior of the stalk domain exhibits a highly charged surface in which the positive and negative potentials follow the curvature of the coiled-coil (Figure 3(b)). There are several interhelical salt bridges and hydrogen bonds, including Arg107: Glu105, Arg119:Glu120, the backbone carbonyl of Thr117:Arg119, as well as an amino-aromatic interaction between Arg129 and Tyr131.

Metal-binding site

In the experimentally-phased electron density maps, a 4σ sphere of density was seen in the tetrameric core between the Glu120 and Gln123 layers. Based on the relatively high concentration of magnesium sulfate in the crystallization mixture, this density was initially modeled as a Mg^{2+}. After several rounds of building and refinement, a decrease in the *B*-factor value for the Mg^{2+} relative to the surrounding side-chain atoms (17.1 $Å^2$ *versus* 27.3 $Å^2$), and the relatively long distances to nearby oxygen atoms (>2.3 Å), indicated that the more likely candidate for this density was a calcium ion (see Methods). Unlike magnesium, which is more restricted in liganding distance (2.0-2.1 Å) and coordination, calcium allows for distances of 2.4 Å and greater and typically accommodates six to eight oxygen atoms (Kyte, 1995; McPhalen *et al*., 1991).

Although the backbone atoms of NSP4 at the metal-binding site approximate 4-fold symmetry, the metal binding ligands do not. Two apposing glutamic acid side-chains at position 120 point towards the metal in a monodentate arrangement with oxygen atoms at a distance of 2.82 Å from the cation; whereas the other two glutamic acid side-chains are oriented away from the core (Figure 4), forming salt bridges with Arg119 occupying the *g*-heptad position of a neighboring chain. Two water molecules, 2.41 Å away from the cation, replace the positions of carboxylate groups of the displaced glutamate groups. In contrast to the asymmetry of the Glu120 layer, all four Gln123 amino acid residues are oriented with each side-chain oxygen atom an average of 2.47 Å from the central Ca^{2+} ion. In addition, the side-chain nitrogen atom from each Gln123 hydrogen bonds with a carboxylate oxygen atom of Glu120 from a neighboring chain regardless of the Glu120 rotomer conformation. Two of the four Gln123 side-chain oxygen atoms also hydrogen bond with the nitrogen atom of a neighboring Gln123 side-chain, which in turn hydrogen bonds to a water molecule positioned on the 2-fold axis beneath the Gln123 layer (Figure 4).

Given that strontium is similar to calcium in its chemical properties, we expected that if this site were selective for calcium, it should show an affinity for strontium as well. Diffraction data were collected on crystals grown in strontium chloride to 1.86 Å resolution. A MAD-phased map contoured at 10σ clearly shows a spherical density at the metal-binding site, which is modeled as a Sr^{2+}. The refined *B*-factor of the Sr^{2+} is 19.3 $Å^2$, similar to that of neighboring side-chain atoms. As corroboration that the density is correctly modeled, refinement of calcium at this position drops the *B*-factor value to ~3 $Å^2$. The Glu120 and Gln123 rotomer conformations are nearly identical between the calcium and strontium structures (Figure 4). Water molecules surrounding the Sr^{2+} occupy analogous positions to those observed around the putative Ca^{2+}. Overall, the primary difference between the calcium and strontium structures is a lengthening of the metal to oxygen distances for the two closest water molecules (2.77 Å for Sr^{+2} compared with 2.41 Å for Ca^{2+}).

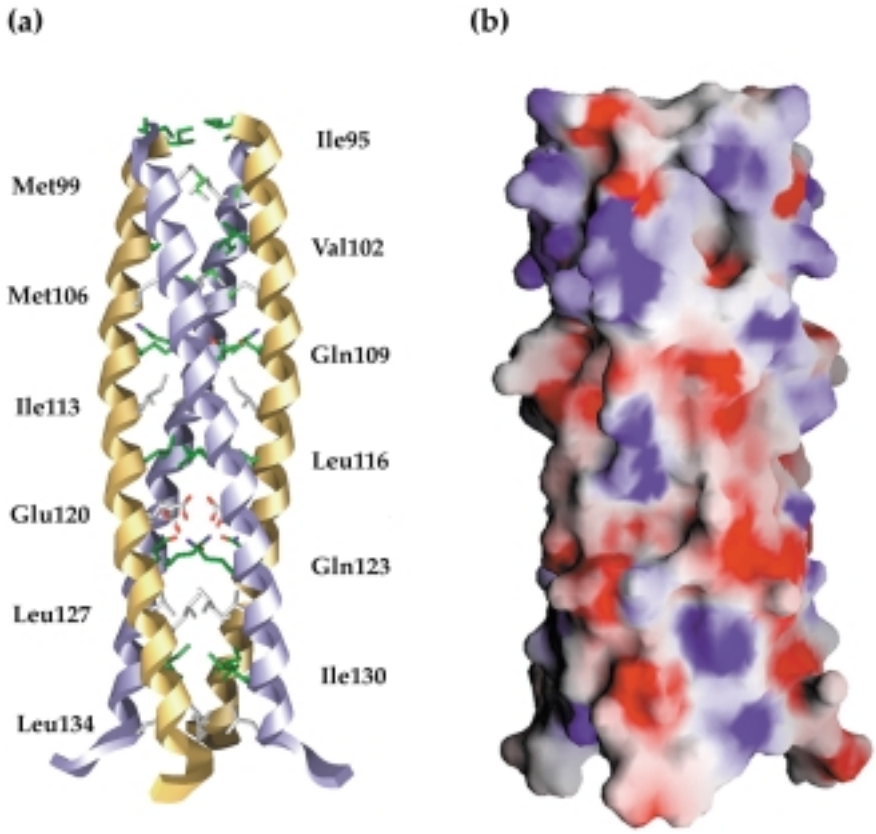

Figure 3. (a) A ribbon representation of $NSP4_{95\text{-}137}$ homo-tetramer with the *a*- (gray) and *d*- (green) heptad positions rendered as sticks (nitrogen, blue; oxygen, red; sulfur, bright green) and labeled. Similarly colored ribbons represent identical chains related by the crystallographic 2-fold. The helical nature of all four chains extends from residue 95 (top) to 134, with the final few residues splaying out from the body of the tetrameric stalk. (b) Electrostatic surface potential of the $NSP4_{95\text{-}137}$ tetramer rendered at ±10kT/e. Globally, the surface of the stalk is highly charged with a symmetrical distribution. Angular strips of positive and negative charge potential alternatively cover the C-terminal half of the stalk surface. As in (a) the amino-terminus is at the top of the Figure. This image was generated using GRASP (Nicholls, 1993).

Core packing of $NSP4_{95\text{-}137}$

Although the parallel organization of the four helices about a crystallographic 2-fold axis gives the appearance of 4-fold symmetry, a closer inspection of both side-chain and main-chain atoms reveals significant deviations from strict symmetry. The joint C^{α} superposition of chains A and B onto chains B and A′ (an approximate rotation of 90° about the pseudo-4-fold axis) of residues 95-134 results in an r.m.s. deviation of 0.78 Å. Visual inspection of this superposition reveals that the region amino-terminal to the metal-binding site displays differences up to 30° in orientation between the A and B chains with respect to the central axis of the coiled-coil. In contrast, the backbone positioning of the helical stretch C-terminal to and including the metal binding site is more 4-fold in character. Joint C^{α} superpositions of AB onto BA′ for residues 95-120 and 120-134 yield r.m.s. deviations of 0.93 Å and 0.29 Å, respectively.

The classical packing expected for a parallel, tetrameric coiled-coil is that of the *a* and *d*-heptad positions packing as knobs into holes (Crick, 1953) in ''perpendicular'' and ''parallel'' arrangements, respectively (Harbury *et al*., 1994, 1993). As viewed from the C terminus, the clockwise twist of chain A with respect to the orientation of chain B indicates a deviation from idealized helix packing (Figure 5(a) and (b)). As shown in Figure 5(b), the Gln109 (*d*-position) and Ile113 (*a*-position) layers clearly deviate from the expected tetrameric coiled-coil packing that is seen in the Gln123 (*d*-position) and Leu127 (*a*-position) layers. The extent of distortion of the idealized tetrameric packing is such that the C^{α}-C^{β} vectors of layers 109 and 113 more closely resemble the acute angles of trimeric coiled-coils. C^{α} superpositions of five-residue stretches from chain A onto helices of trimeric viral fusion proteins indicate that the Gln109 and Ile113 layers of the NSP4 tetramer display helical orientations similar to those of trimeric coiled-coils (Figure 5(c)).

In the absence of polar layers, the identities of the hydrophobic residues at *a* and *d*-heptad positions have been shown to dictate the oligomerization state of coiled-coils, with tetramers showing a preference for β-branched (primarily isoleucine) and non-β-branched (primarily leucine) amino acid

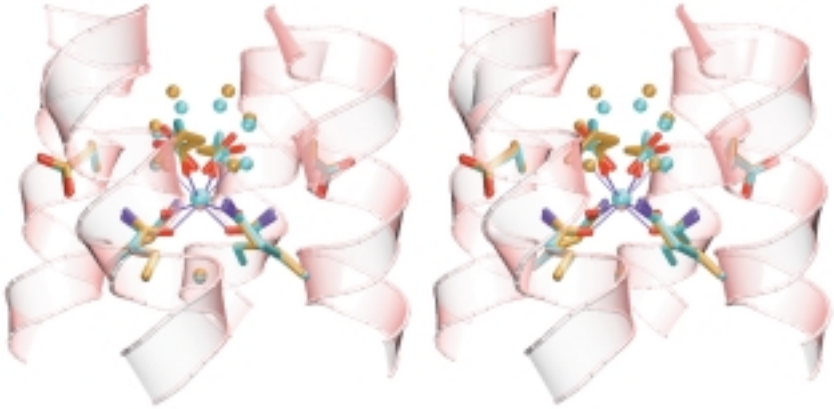

Figure 4. Superposition of the calcium and strontium-bound $NSP4_{95\text{-}137}$ core metal-binding sites, illustrated in tan and turquoise, respectively. Metal coordination is highlighted by broken lines in the calcium-bound structure (oxygen, red; nitrogen, blue). The rotomer conformations of the glutamic acid and glutamine side-chains at positions 120 and 123, respectively, are nearly identical in the two structures. The Ca^{2+} and Sr^{2+} and the seven surrounding water molecules are shown as small spheres, with the Ca^{2+} partially obscured by the larger Sr^{2+}. This image was generated using VMD (Humphrey *et al.*, 1996) and Raster3D (Merritt & Bacon, 1997).

side-chains at the *d* and *a*-positions, respectively (for a review see, Kohn & Hodges, 1998). Immediately amino-terminal to the metal-binding site, the $NSP4_{95\text{-}137}$ tetramer contains Ile113 at an *a*-position and Leu116 at a *d*-position, which would not be expected to be preferred in a tetramer. In contrast to the *a* and *d*-positions preceding the metal-binding site, the three core residues C-terminal to Gln123 (Leu127, Ile130, and Leu134) display the preferred amino acid identities for tetrameric *a* and *d*-heptad positions. The presence of the divalent cation between the Glu120 and Gln123 layers may play a role in the establishment of a more idealized packing arrangement.

A global rendering of the $NSP4_{95\text{-}137}$ tetramer interior illustrates the loose packing of side-chains. The $NSP4_{95\text{-}137}$ homo-tetramer has seven interior pockets that range in volume from 20.6 $Å^3$ to an upper limit of 70.0 $Å^3$ (Figure 6). The largest of these pockets is occupied by a water network proximal to the metal site and bordered by the Glu120 and the Leu116 layers. A second, smaller pocket is filled by a water molecule which lies on the crystallographic 2-fold axis and interacts with the Gln123 layer. In contrast, five of the other cavities do not contain ordered water molecules, resulting in a total volume of 186 $Å^3$.

Solvent-inaccessible pockets and water molecules are also seen in the homo-trimeric, parallel, coiled-coil domains of the classical viral fusion proteins. Crystal structures of HIV gp41 (Chan *et al.*, 1997) and Moloney Murine Leukemia virus (MoMuLV) transmembrane domain (Fass *et al.*, 1996) exhibit three to four internal cavities totaling 88.7 $Å^3$ and 152.0 $Å^3$, respectively. In addition to displaying between two to four core pockets, influenza hemagglutinin (Bullough *et al.*, 1994) and Ebola GP2 ectodomain (Malashkevich *et al.*, 1999; Weissenhorn *et al.*, 1998) have three and five buried water molecules, respectively. The significance of these pockets is unclear but may be due to limitations of homomeric, parallel coiled-coil packing.

Discussion

The structure of the oligomerization domain of NSP4 supports previous biochemical characterization and shows a detailed view of the interactions of this homo-tetramer. The $NSP4_{95\text{-}137}$ homo-tetramer shares features in common with other coiled-coil proteins, yet the presence of a negatively charged layer of residues in a core position neutralized by a bound metal distinguishes this from other coiled-coils. This structure is a first step towards understanding the atomic details of NSP4 as an intracellular viral receptor as well as an enterotoxin.

In a general sense, the role of NSP4 in ICP budding into the ER may be compared to that of the fusion proteins of enveloped viruses. The classical viral fusion proteins, embedded in the lipid envelopes of their virions, are trimeric coiled-coils which bring the target cell membrane to the lipid envelope of the virion (for a review see, Skehel & Wiley, 1998). Crystallographic analysis of influenza hemagglutinin indicates that fusion is induced by conformational changes in two distinct regions of the protein: extension of an amino-terminal fusion-peptide from the viral surface towards the target membrane; and a hinging of the C-terminal domain about the base of the coiled-coil, such that anti-parallel packing against the stalk is achieved (Bullough *et al.*, 1994). This domain orientation appears to be the same for NSP4 with respect to the target membrane and the virus particle. The N-terminal transmembrane domain is stationed in the target ER membrane, and the cytoplasmic C-terminal stretch, following the coiled-coil domain, presents the ICP-binding residues. Deletion of this C-terminal stretch increases the

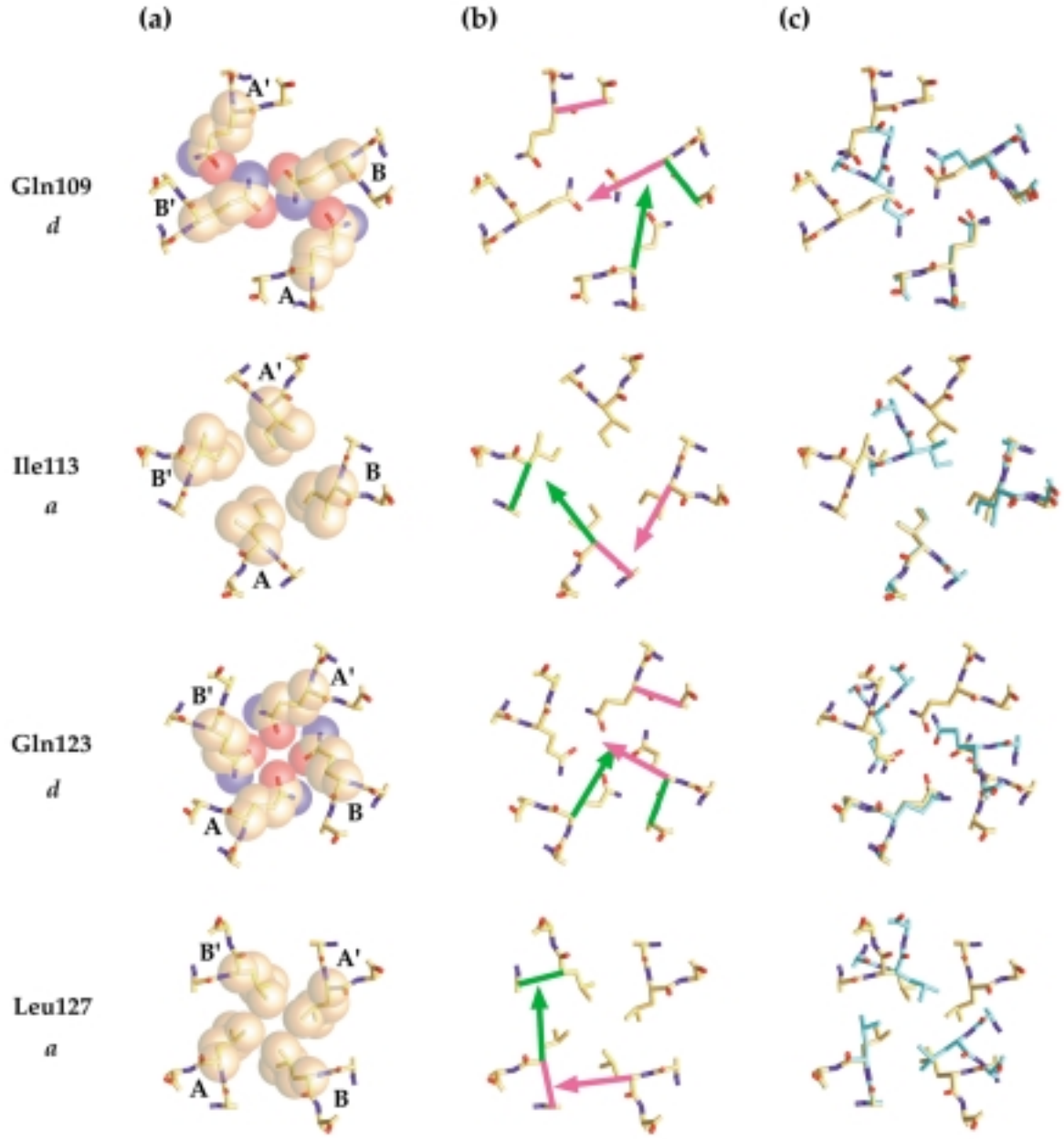

Figure 5. Deviations in packing geometry of the layers amino-terminal, but not carboxyl-terminal to the cation-binding site. As viewed from the C terminus down the central coiled-coil axis, cross-sectional slices of the $NSP4_{95\text{-}137}$ tetramer at the Gln109, Ile113, Gln123 and Leu127 layers highlight the internal packing of the heptad *a* and *d*-positions (carbon, tan; nitrogen, blue; oxygen, red). (a) CPK representation of the side-chain atoms reveals an approximate 4-fold arrangement and closer packing in the layers C-terminal (Leu127) and including (Gln123) the metal-binding site relative to those N-terminal (Gln109 and Ile113). A and B indicate the two chains in the asymmetric unit. (b) A coiled-coil vector diagram indicates that layers amino-terminal to the metal-binding site deviate from the ''parallel'' and ''perpendicular'' packing expected for a parallel tetrameric coiled-coil. In an idealized packing arrangement of a parallel four-stranded coiled-coil, the *a* and *d*-heptad positions pack in ''parallel'' and ''perpendicular'' orientations, respectively, as defined by the angle of the C^{α}-C^{β} vector with respect to a C^{α}-C^{α} of the neighboring strand (Crick, 1953; Harbury *et al.*, 1994, 1993). The C^{α}-C^{β} vectors (green for chain A; magenta for chain B) are represented as arrows, and the C^{α}-C^{α} vectors as thick lines. (c) The relative orientation of the A and B helices in the Gln109 and Ile113 layers approximate those found in trimeric coiled-coils. For each NSP4 layer, only the indicated layer residue position of chain A and flanking C^{α} atoms were superimposed onto a single helix from a trimeric viral fusion protein. The resulting orientation of chain B in several of the layers amino-terminal to the metal-binding site coincide with the position of a second helix of the trimer. Trimers (carbon, light blue; nitrogen, blue; oxygen, red) used for comparison were HIV gp41 (PDB accession code:1AIK) and Moloney Murine Leukemia virus TM domain (PDB accession code:1MOF).

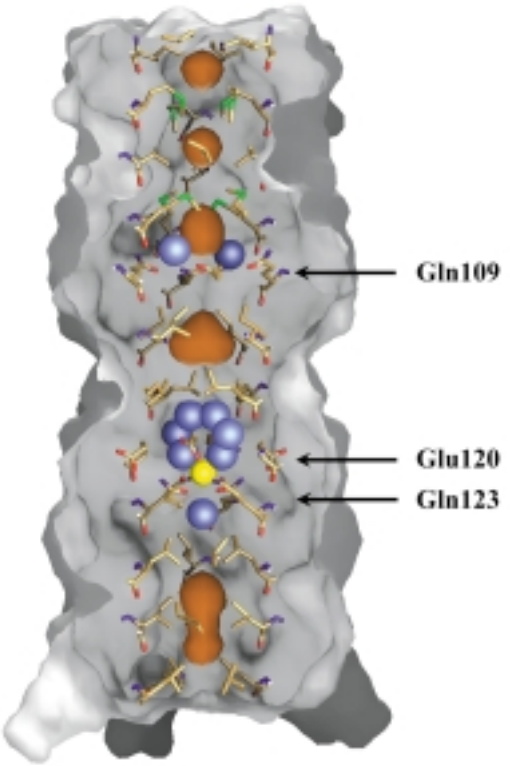

Figure 6. A longitudinal slice revealing cavities on the crystallographic 2-fold axis within the tetrameric core. Five of the seven cavities (orange) are present in the final model of $NSP4_{95\text{-}137}$, with two pockets (not rendered) filled with ordered water molecules. Overall, there are nine ordered water molecules (blue) buried in the interior, two at the Gln109 layer and seven proximal to the Gln123/Glu120 layers. The calcium ion is shown in yellow. The side-chains for the *a* and *d*-heptad positions are rendered as sticks and the polar layers are labeled. Surfaces were generated in GRASP (Nicholls, 1993) using a sphere radius of 1.4 Å.

affinity of NSP4 for VP4 (Au *et al.*, 1993), suggesting that the C-terminal domain might be partially obstructing the VP4-binding site. Upon ICP binding, the C-terminal region may be displaced allowing for more efficient binding of VP4.

Anatomical features observed in the crystal structure of NSP4 can be compared and contrasted to other coiled-coils domains, including those of the classical fusion proteins. $NSP4_{95\text{-}137}$ has three interior polar layers, two of which are located at adjacent *a* and *d*-heptad positions and coordinate a divalent cation and water network approximately four helical turns from the base of the coiled-coil stalk. Core polar layers consisting of either asparagine or glutamine residues are present in all structures of viral fusion proteins solved to date. In addition, buried water molecules and chloride ions have been observed in a number of homomeric coiled-coil structures. In the MoMuLV TM domain and Ebola GP2, a chloride ion is positioned on the central axis of the coiled-coil by three asparagine side-chains (Fass *et al.*, 1996; Malashkevich *et al.*, 1999; Weissenhorn *et al.*, 1998) approximately three helical turns from the C-terminal base of the stalk. It appears that positioning a chloride binding site in the interior of a trimeric coiled-coil domain may be sufficiently accomplished by placement of a single layer of either glutamine or asparagine residues as was exemplified by an engineered variant of GCN4 (Eckert *et al.*, 1998). In contrast, a core charged layer (Glu120) neighboring a layer of glutamine residues arranged in a tetrameric context in NSP4, forms a cation-binding site, and thus differs from that previously seen in other naturally occurring coiled-coil proteins.

Recently, structures of two homomeric four-stranded coiled-coils were reported: the 114 amino acid oligomerization domain of the P protein of Sendai virus (Tarbouriech *et al.*, 2000); and the 52-residue protease binding region of the tetrabrachion complex from *Staphylothermus marinus* (Stetefeld *et al.*, 2000). These structures display deviations from the classical 3,4-hydrophobic repeat, which affect the coiled-coil pitch and the orientation of *a* and *d*-positions with respect to the central coiled-coil axis. Despite deviations from idealized coiled-coil packing, these proteins differ from $NSP4_{95\text{-}137}$ in that they maintain 4-fold symmetry along the central axis of the tetramer. Like the case of NSP4, the P protein of Sendai virus binds to a calcium ion in the coiled-coil core. However, this calcium-binding site is composed of a single layer of four uncharged ligands situated at a distance of ~3.5 Å (Tarbouriech *et al.*, 2000). Together with $NSP4_{95\text{-}137}$, these tetrameric structures expand our understanding of homo-oligomeric, parallel coiled-coils.

The packing arrangement for core positions of coiled-coils is typically described as the orientation of the C^{α}-C^{β} vector of the *a* or *d*-heptad positions with respect to a C^{α}-C^{α} vector of a neighboring helix. In the case of dimers, these vectors display a ''parallel'' arrangement at *a*-positions and a ''perpendicular'' arrangement at *d*-positions, whereas the reverse, perpendicular packing at *a*-positions and parallel packing at *d*-positions, is observed for tetramers (Harbury *et al.*, 1993). In the structure of $NSP4_{95\text{-}137}$, deviations from these idealized packing arrangements are observed amino-terminal to the metal-binding site, with some of the layers displaying the acute angles seen in trimeric coiled-coils (Figure 5(b),(c)). These deviations coincide with a relative twisting of the A and A′ helices with respect to the B and B′ helices, breaking the 4-fold symmetry that might be expected for a homo-tetramer. Both the departure from 4-fold symmetry and the deviations from idealized tetrameric packing may be due to the identity of the *a* and *d*-position residues preceding the metal-binding site. Mutational studies of coiled-coil *a* and *d*-positions have demonstrated that isoleucine displays the strongest preference for a particular geometric packing environment (Harbury *et al.*,

1993; Tripet *et al.*, 2000; Wagschal *et al.*, 1999). As has been shown with variants of GCN4, the constraint for isoleucine to be located at a non-perpendicular packing position appears sufficient to discriminate between dimers, trimers and tetramers (Harbury *et al.*, 1993). In NSP4, the presence of Ile113 at an *a*-position may be inducing a twist in the backbone to accommodate better side-chain packing. With the exception of Ile95 and Ile130 at *d*-positions, the remaining hydrophobic residues would not be expected to be sufficient for formation of a tetrameric coiled-coil. Similarly, the Gln109 layer shows a marked asymmetry in the side-chain packing, yet it is unclear if glutamine residues are unpreferred in *d*-positions of tetrameric coiled-coils or if this is due to the unpreferred layers (Ile113 and Leu116) which follow this position. The packing of the Gln109 layer contrasts the symmetrical positioning observed in the Gln123 layer, a symmetry which is likely explained by the bound cation. Although several layers display an organization characteristic of trimeric coiled-coils, $NSP4_{95\text{-}137}$ may be maintained as a tetramer due to a combination of a bound divalent cation and tetramerically preferred residues.

The crystal structure of $NSP4_{95\text{-}137}$ provides an understanding as to how a parallel coiled-coil can self-associate with a charged residue at a core heptad position. In the case of a parallel four-stranded coiled-coil, the residues at the *a* and *d*-heptad positions must be able to form self-interacting layers; therefore, the presence of a charged residue at these positions would be expected to be unfavorable due to electrostatic repulsion. In this structure the charged glutamic acid layer is accommodated by a conformational asymmetry in which two side-chains are pointing away from the core and interacting with arginine residues on adjacent helices within the tetramer, and two are buried in the core, interacting with a bound divalent cation and water network. The absence of a divalent cation would appear to be destabilizing for the tetramer, as glutamate:glutamate repulsion might orient all of the Glu120 side-chains towards the Arg119 residues, thus creating a large core cavity. Given the helical twisting observed immediately preceding the metal-binding site, one might imagine this site to be the target for modulation by other protein components, and thus act as a ''switch'' in determining the oligomerization state of NSP4.

The X-ray structure of $NSP4_{95\text{-}137}$ solved from crystals grown in magnesium sulfate raises the question of the identity of the divalent cation bound at the metal-binding site. Although initially modeled as a magnesium ion, the metal-to-oxygen distances, ligation number, and coordination geometry were more consistent with that of a calcium ion. Since magnesium was the only divalent cation added to the crystallization mix, the ion at this site must be a contaminant. The structure determination from a crystal grown in strontium chloride gives further evidence this site not only binds divalent cations, but should show an affinity for calcium.

The binding of a calcium ion between the polar layers in the oligomerization domain of NSP4 has implications for understanding the enterotoxic effect of this viral receptor. Ectopic expression of NSP4 in baculovirus-infected insect cells results in the release of Ca^{2+} from the ER, effectively increasing the intracellular calcium concentrations fivefold (Tian *et al.*, 1995, 1994). These local calcium increases would be expected to play a role in the stabilization of the tetrameric state of ER-bound NSP4. However, this cation-binding site may be even more critical for stability of the homotetramer in an extracellular environment. Following the lysis of infected cells, the exposure of the NSP4 cytoplasmic domain to the extracellular milieu presumably results in the digestion of proteolytically sensitive regions (O'Brien *et al.*, 2000; Taylor *et al.*, 1996) leaving the resistant coiled-coil stalk domain intact. As Ball and coworkers demonstrated, the enterotoxic effects of rotavirus infection can be elicited through injection of $NSP4_{114\text{-}135}$ peptide into mice (Ball *et al.*, 1996). $NSP4_{114\text{-}135}$ constitutes the carboxyl-terminal half of the coiled-coil domain and encompasses a putative calcium-binding site, suggesting the causative agent may be extracellular homo-tetramer populations of NSP4.

Materials and Methods

Crystallization and data collection

A synthetic 43 amino acid peptide corresponding to residues 95-137 of NSP4 was resuspended in 100 mM Hepes, pH 7.0. Crystals grew at room temperature in hanging drops equilibrated over well solutions containing 20-28 % PEG 400, 100 mM Tris-HCl (pH 8.5), and a variety of salts ($MgSO_4$, $CaCl_2$, $MnCl_2$, $SrCl_2$, $BaCl_2$, Li_2SO_4, and NaCl), several of which were not suitable for diffraction studies without further optimization. Peptides derivatized with seleno-methionine at position 112 yielded crystals under the same conditions, and were maintained in the reduced state by the addition of 10 mM dithiothreitol to all solutions. Both native and selenium-derivatized peptides crystallized in space group $P2_12_12$ with two molecules per ASU (see Table 1 for statistics). These two non-crystallographic symmetry mates are related by an approximate 90° rotation about the crystallographic 2-fold axis, yielding a pseudo-4-fold symmetry axis aligned in the center of the coiled-coil tetramer. In combination with the similar lengths of the *a* and *b* unit cell axes, the pseudo-4-fold axis allowed for pseudo-merohedral twinning, resulting in an apparent spacegroup of $P42_12$ with one molecule in the asymmetric unit, similar to that seen in crystals of rubredoxin oxygen reductase from *Desulfovibrio gigas* (Frãzao *et al.*, 1999). Diffraction data from selenium-derivatized peptide crystals displayed a lower twinning fraction than those from the underivatized peptide (R_{merge} of $P42_12$-scaled data from 30.0 to 5.4 Å was 11-13 % for derivatized compared to 6 % for underivatized). Therefore, data collected from the selenium-derivatized crystals were used for all model building and refinement.

Two multiwavelength anomalous dispersion (MAD) datasets were collected at 100 K on crystals grown in $MgSO_4$ and $SrCl_2$, respectively, using beamline X12C at the National Synchrotron Light Source, Brookhaven National Laboratory. Data were processed, integrated and scaled using HKL2000 and SCALEPACK (Otwinowski & Minor, 1997).

Structure determination

For each MAD dataset, the determination of the two selenium atom positions and calculation of experimentally phased, solvent-flipped electron density maps were performed using the CNS program suite (Brünger, 1998). Polyalanine helical segments were docked into these initial maps, which displayed unbroken backbone density for the entire length of the peptides. Discovery of a metal-binding site in the tetramer core prompted independent building and refinement of the two structures to allow for an unbiased comparison of all side-chain residues. During building and refinement of the data collected from crystals grown in $MgSO_4$, a magnesium ion was initially positioned in the model, but was later replaced with a calcium ion. Calcium may have been introduced as a contaminant in the commercial preparation of the magnesium sulfate stock used for crystallization (contaminating levels of up to 0.001 %) as well as other reagents used in crystallization and peptide preparation. Cycles of building and refinement were performed with O (Jones *et al*., 1991) and CNS (Brünger, 1998), respectively.

Coordinates

The coordinates and experimental data have been deposited in the RCSB Protein Data Bank (PDB accession codes:1G1I and 1G1J).

Acknowledgments

We wish to thank Zbigniew Dauter for insightful recognition of the twinning problem. We gratefully acknowledge Ron Shigeta for crystallographic advice and Charles Dismukes for scientific discussion. We appreciate the NSLS X12-C beamline assistance of Jon Skinner, Michael Becker, Anand Saxena and Robert Sweet. Diffraction data for this study were collected at Brookhaven National Laboratory in the Biology Department single-crystal diffraction facility at beamline X12-C in the National Synchrotron Light Source. NIH grants supported I.M.N. (GM08309), G.D.B. (GM44038), and C.E.S. (GM44038).

References

Au, K.-S., Chan, W.-K., Burns, J. W. & Estes, M. K. (1989). Receptor activity of rotavirus non-structural glycoprotein NS28. *J. Virol.* **63**, 4553-4562.

Au, K.-S., Mattion, N. M. & Estes, M. K. (1993). A subviral particle binding domain on the rotavirus nonstructural glycoprotein NS28. *Virology*, **194**, 665-673.

Ball, J. M., Tian, P., Zeng, C. Q.-Y., Morris, A. P. & Estes, M. K. (1996). Age-dependent diarrhea induced by a rotaviral nonstructural glycoprotein. *Science,* **272**, 101-104.

Barton, G. J. (1993). ALSCRIPT - A tool to format multiple sequence alignments. *Protein Eng.* **6**, 37-40.

Bergmann, C. C., Maass, D., Poruchynshy, M. S., Atkinson, P. H. & Bellamy, A. R. (1989). Topology of the non-structural rotavirus receptor glycoprotein NS28 in the rough endoplasmic reticulum. *EMBO J.* **8**, 1695-1703.

Both, G. W., Siegman, L. J., Bellamy, A. R. & Atkinson, P. H. (1983). Coding assignment and nucleotide sequence of simian rotavirus SA11 gene segment 10: location of glycosylation sites suggests that the signal peptide is not cleaved. *J. Virol.* **48**, 335-339.

Brünger, A. (1998). Crystallography and NMR system (CNS): a new software system for macromolecular structure determination. *Acta Crystallog. sect. D,* **54**, 905-921.

Bullough, P. A., Hughson, F. M., Skehel, J. J. & Wiley, D. C. (1994). Structure of influenza haemagglutinin at the pH of membrane fusion. *Nature,* **371**, 37-43.

Chan, D. C., Fass, D., Berger, J. M. & Kim, P. S. (1997). Core structure of gp41 from the HIV envelope glycoprotein. *Cell,* **89**, 263-273.

Ciarlet, M., Liprandi, F., Conner, M. E. & Estes, M. K. (2000). Species specificity and interspecies relatedness of NSP4 genetic groups by comparative NSP4 sequence analyses of animal rotaviruses. *Arch. Virol.* **145**, 371-383.

Crick, F. H. C. (1953). The packing of α-helices: simple coiled-coils. *Acta Crystallog.* **6**, 689-697.

Eckert, D. M., Malashkevich, V. N. & Kim, P. S. (1998). Crystal structure of GCN4-pI_QI, a trimeric coiled coil with buried polar residues. *J. Mol. Biol.* **284**, 859-865.

Estes, M. K. (1991). Rotaviruses and their replication. In *Fundamental Virology* (Fields, B. N. & Knipe, D. M., eds), 2nd edit., pp. 619-642, Raven Press, New York.

Fass, D., Harrison, S. C. & Kim, P. S. (1996). Retrovirus envelope domain at 1.7 Å resolution. *Nature Struct. Biol.* **3**, 465-469.

Frãzao, C., Sieker, L., Coelho, R., Morais, J., Pacheco, I., Chen, L., LeGall, J., Dauter, Z., Wilson, K. & Carrondo, M. A. (1999). Crystallization and preliminary diffraction data analysis of both single and pseudo-merohedrally twinned crystals of rubredoxin oxygen oxidoreductase from *Desulfovibrio gigas*. *Acta Crystallog. sect. D,* **55**, 1465-1467.

Glass, R. I., Gentsch, J. & Smith, J. C. (1994). Rotavirus vaccines: success by reassortment? *Science,* **265**, 1389-1391.

Harbury, P. B., Kim, P. S. & Alber, T. (1994). Crystal structure of an isoleucine-zipper trimer. *Nature,* **371**, 80-83.

Harbury, P. B., Zhang, T., Kim, P. S. & Alber, T. (1993). A switch between two, three, and four-stranded coiled coils in GCN4 leucine zipper mutants. *Science,* **262**, 1401-1407.

Hendrickson, W. A. (1991). Determination of macromolecular structures from anomalous diffraction of synchrotron radiation. *Science,* **254**, 51-58.

Horie, Y., Nakagomi, O., Koshimura, Y., Nakagomi, T., Suzuki, Y., Oka, T., Sasaki, S., Matsuda, Y. & Watanabe, S. (1999). Diarrhea induction by rotavirus NSP4 in the homologous mouse model system. *Virology,* **262**, 398-407.

Humphrey, W., Dalke, A. & Schulten, K. (1996). VMD - Visual Molecular Dynamics. *J. Mol. Graphics,* **14**, 33-38.

Jones, T. A., Zou, J.-Y., Cowtan, S. W. & Kjeldgaard, M. (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallog. sect. A,* **47**, 110-119.

Kabcenell, A. K. & Atkinson, P. H. (1985). Processing of the rough endoplasmic reticulum membrane glycoproteins of rotavirus SA11. *J. Cell Biol.* **101**, 1270-1280.

Kohn, W. D. & Hodges, R. S. (1998). *De novo* design of α-helical coiled-coils and bundles: models for the development of protein-design principles. *Trends Biotechnol.* **16**, 379-389.

Kyte, J. (1995). *Mechanism in Protein Chemistry*, Garland Publishing, Inc., New York.

Maass, D. R. & Atkinson, P. H. (1990). Rotavirus proteins VP7, NS28, and VP4 form oligomeric structures. *J. Virol.* **64**, 2632-2641.

Malashkevich, V. N., Schneider, B. J., McNally, M. L., Milhollen, M. A., Pang, J. X. & Kim, P. S. (1999). Core structure of the envelope glycoprotein GP2 from Ebola virus at 1.9 Å resolution. *Proc. Natl Acad. Sci. USA,* **96**, 2662-2667.

McPhalen, C. A., Strynadka, N. C. & James, M. N. G. (1991). Calcium binding sites in proteins: a structural perspective. In *Metalloproteins: Structural Aspects* (Anfinsen, C. B., Edsall, J. T., Richards, F. M. & Eisenberg, D. S., eds), vol. 42, pp. 77-144, Harcourt Brace Jovanovich.

Merritt, E. A. & Bacon, D. J. (1997). Raster3D - photorealistic molecular graphics. *Methods Enzymol.* **277**, 505-524.

Meyer, J. C., Bergmann, C. C. & Bellamy, A. R. (1989). Interaction of rotavirus cores with the nonstructural glycoprotein NS28. *Virology,* **171**, 98-107.

Nicholls, A. (1993). *GRASP: Graphical Representation and Analysis of Surface Properties*, Columbia University, New York, USA.

O'Brien, J. A., Taylor, J. A. & Bellamy, A. R. (2000). Probing the structure of rotavirus NSP4: a short sequence at the extreme C terminus mediates binding to the inner capsid particle. *J. Virol.* **74**, 5388-5394.

Otwinowski, Z. & Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307-326.

Poruchynsky, M. S., Maass, D. R. & Atkinson, P. H. (1991). Calcium depletion blocks the maturation of rotavirus by altering the oligomerization of virus-encoded proteins in the ER. *J. Cell Biol.* **114**, 651-661.

Skehel, J. J. & Wiley, D. C. (1998). Coiled-coil in both intracellular vesicle and viral membrane fusion. *Cell,* **95**, 871-874.

Stetefeld, J., Jenny, M., Schulthess, T., Landwehr, R., Engel, J. & Kammerer, R. A. (2000). Crystal structure of a naturally occurring parallel right-handed coiled-coil tetramer. *Nature Struct. Biol.* **7**, 772-776.

Tarbouriech, N., Curran, J., Ruigrok, R. W. H. & Burmeister, W. P. (2000). Tetrameric coiled-coil domain of Sendai virus phosphoprotein. *Nature Struct. Biol.* **7**, 777-781.

Taylor, J. A., O'Brien, J. A. & Yeager, M. (1996). The cytoplasmic tail of NSP4, the endoplasmic reticulum-localized non-structural glycoprotein of rotavirus, contains distinct virus binding and coiled coil domains. *EMBO J.* **15**, 4469-4476.

Tian, P., Hu, Y., Schilling, W. P., Lindsay, D. A., Eiden, J. & Estes, M. K. (1994). The non-structural glycoprotein of rotavirus affects intracellular calcium levels. *J. Virol.* **68**, 251-257.

Tian, P., Estes, M. K., Hu, Y., Ball, J. M., Zeng, C. Q.-Y. & Schilling, W. P. (1995). The rotavirus non-structural glycoprotein NSP4 mobilizes Ca^{2+} from the endoplasmic reticulum. *J. Virol.* **69**, 5763-5772.

Tripet, B., Wagschal, K., Lavigne, P., Mant, C. T. & Hodges, R. S. (2000). Effects of side-chain characteristics on stability and oligomerization state of a *de novo* designed model coiled-coil: 20 amino acid substitutions in position "d". *J. Mol. Biol.* **300**, 377-402.

Wagschal, K., Tripet, B. & Hodges, R. S. (1999). *De novo* design of a model peptide sequence to examine the effects of single amino acid substitutions in the hydrophobic core on both stability and oligomerization state of coiled-coils. *J. Mol. Biol.* **285**, 785-803.

Weissenhorn, W., Carfi, A., Lee, K.-H., Skehel, J. J. & Wiley, D. C. (1998). Crystal structure of the Ebola virus membrane fusion subunit, GP2, from the envelope glycoprotein ectodomain. *Mol. Cell,* **2**, 605-616.

Edited by D. Rees

(Received 7 August 2000; received in revised form 16 October 2000; accepted 17 October 2000)