Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Structure *14*, 889–899, May 2006 DOI 10.1016/j.str.2006.03.007

Structures and Polymorphic Interactions of Two Heptad-Repeat Regions of the SARS Virus S2 Protein

Yiqun Deng,[1,2] Jie Liu,[1,2] Qi Zheng,[1] Wei Yong,[1] and Min Lu[1,*]
[1] Department of Biochemistry
Weill Medical College of Cornell University
New York, New York 10021

Summary

Entry of SARS coronavirus into its target cell requires large-scale structural transitions in the viral spike (S) glycoprotein in order to induce fusion of the virus and cell membranes. Here we describe the identification and crystal structures of four distinct α-helical domains derived from the highly conserved heptad-repeat (HR) regions of the S2 fusion subunit. The four domains are an antiparallel four-stranded coiled coil, a parallel trimeric coiled coil, a four-helix bundle, and a six-helix bundle that is likely the final fusogenic form of the protein. When considered together, the structural and thermodynamic features of the four domains suggest a possible mechanism whereby the HR regions, initially sequestered in the native S glycoprotein spike, are released and refold sequentially to promote membrane fusion. Our results provide a structural framework for understanding the control of membrane fusion and should guide efforts to intervene in the SARS coronavirus entry process.

Introduction

The recent global outbreak of an atypical pneumonia—severe acute respiratory syndrome (SARS)—has been associated with a newly identified coronavirus termed SARS coronavirus (SARS-CoV) (Guo et al., 2003). Coronaviruses are enveloped, positive-stranded RNA viruses that infect a wide range of animals. Infection of cells by coronaviruses requires fusion of the target and viral membranes, a process mediated by the viral spike (S) glycoprotein (De Groot et al., 1989). Coronavirus S proteins are synthesized as single-chain precursors that oligomerize in the endoplasmic reticulum and are processed through the Golgi, eventually forming long, petal-shaped spikes that protrude from the virion surface (Lai and Holmes, 2001). In some coronaviruses, S is posttranslationally cleaved into two chains, known as S1 (the receptor binding protein) and S2 (the transmembrane fusion protein) (Frana et al., 1985; Sturman et al., 1985). Coronavirus S2 is functionally and structurally related to a large group of so-called class I viral fusion proteins, including those of orthomyxoviruses such as influenza virus, paramyxoviruses such as SV5, retroviruses such as HIV, and filoviruses such as Ebola virus (Earp et al., 2005; Eckert and Kim, 2001; Harrison, 2005; Weissenhorn et al., 1999). A general fusion mechanism for these viral fusion proteins posits, in simple terms, that binding of S1 to a receptor on the target cell membrane triggers a series of conformational changes in S2 that ultimately lead to the formation of a highly stable postfusion trimer-of-hairpins structure (Duquerroy et al., 2005; Supekar et al., 2004; Xu et al., 2004). A large body of evidence suggests that the fusion reaction proceeds via a regulated sequence of transitions involving one or more on-pathway intermediate(s) (Eckert and Kim, 2001).

S2 is a type I integral membrane protein with a large N-terminal ectodomain, a single transmembrane domain, and a C-terminal cytoplasmic tail. The ectodomain includes two highly conserved regions consisting of heptad-repeats (HR) of hydrophobic residues characteristic of coiled coils. The first HR (HR_N) is adjacent to the N terminus of S2, while the second HR (HR_C) immediately precedes the transmembrane domain (Figure 1A). By analogy with other class I viral fusion proteins, the N-terminal region of S2 contains a hydrophobic "fusion peptide" (Earp et al., 2005), although its exact location in the sequence is not known. The fusion peptide and HR regions must remain sequestered in the native S glycoprotein spike (Baker et al., 1999; Wang et al., 2002; Yin et al., 2006). During the fusion process, the fusion peptide is exposed and inserted into the target cell bilayer, providing attachment points for drawing together the viral and cellular membranes (Skehel and Wiley, 2000). Protein dissection studies from several groups demonstrated that the HR_N and HR_C regions of S2 associate to form a stable, α-helical trimer of antiparallel heterodimers (Bosch et al., 2003; Ingallinella et al., 2004; Liu et al., 2004; Tripet et al., 2004). X-ray crystallographic analysis revealed that three HR_N helices form an interior, parallel trimeric coiled coil, whereas three C-terminal HR_C helices pack in an antiparallel manner into three hydrophobic grooves on the surface of this coiled coil (Duquerroy et al., 2005; Supekar et al., 2004; Xu et al., 2004). This rod-shaped α-helical core of the S2 trimer-of-hairpins shares features with the other known class I fusion proteins (Earp et al., 2005; Eckert and Kim, 2001; Harrison, 2005; Weissenhorn et al., 1999). Formation of the stable six-helix bundle has been proposed to bring the viral and cellular membranes into close apposition, thereby facilitating membrane fusion and subsequent viral entry (Chan et al., 1997; Tan et al., 1997; Weissenhorn et al., 1997). Interestingly, isolated peptides derived from the HR_C region of S2 can have antiviral activity with IC_{50} values in the micromolar range (Bosch et al., 2004; Liu et al., 2004). Thus, in addition to providing insights into fusion-activating conformational changes in S2, determining the structural basis for interactions between the HR regions could assist anti-SARS drug development efforts.

In this report, we have applied a protein dissection approach to identify and study two isolated HR domains and two HR_N/HR_C complexes from SARS-CoV S2. The set of structures and their widely varying stabilities suggest a plausible mechanism for achieving the controlled activation of S2. We discuss the implications of

*Correspondence: mlu@med.cornell.edu
[2] These authors contributed equally to this work.

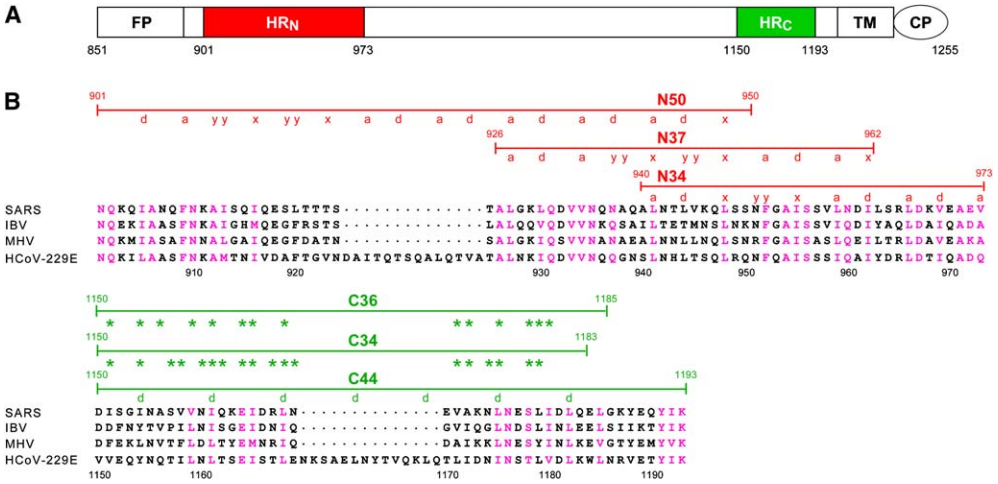

Figure 1. Amphipatic α-Helical Regions Identified in SARS-CoV S2

(A) Schematic diagram of S2. The positions of the transmembrane region (TM), the cytoplasmic domain (CP), and the two HR regions predicted to form amphipatic helices (HR_N and HR_C) are shown. The exact location of the fusion peptide (FP) is uncertain; for convenience, it is shown at the N terminus of the protein. Residues are numbered according to their position in the Urbani SARS-CoV S sequence.

(B) Sequence alignment of the HR_N (top) and HR_C (bottom) regions derived from SARS-CoV S2 with the corresponding residues of S2 proteins from other coronavirus genera. The sequences of SARS-CoV (Urbani strain used in this study), HCoV-229E (human coronavirus, strain 229E), MHV (mouse hepatitis virus, strain JHM), and IBV (avian infectious bronchitis virus, strain LX4) are shown. The locations of the two isolated HR domains (N34 and C44) and two HR_N/HR_C complexes (N37/C34 and N50/C36) from SARS-CoV S2 are indicated. Core residue positions (*a*, *d*, *x*, or *y*) for the N34, N37, N50, and C44 coiled coils are shown (see text and Figure 5). Residues of the C34 and C36 peptides in direct contact with their respective N37 and N50 coiled coils are indicated by asterisks. Chemically similar residues among coronaviruses are colored red.

polymorphic interactions of the HR regions for SARS-CoV membrane fusion and its inhibition.

Results and Discussion

The HR_C Tetramer

We began by studying a 44 residue peptide (called C44) corresponding to the predicted HR_C region of SARS-CoV S2 (Figure 1). On the basis of circular dichroism (CD) measurements at 50 μM peptide concentration in neutral phosphate-buffered saline (PBS), C44 is ~75% helical at 4ºC and displays a cooperative thermal unfolding transition with a melting temperature (T_m) of 40ºC (Table 1 and Figures 2A and 2B). Sedimentation equilibrium experiments indicate that C44 is a tetramer (Figure 2C). The X-ray crystal structure of the C44 peptide at 1.70 Å resolution (Table 2) reveals a unique left-handed superhelix consisting of four antiparallel α helices crossing at an angle of ~18º (Figure 3). These helices form a cylinder with an overall diameter of ~25 Å and a length of ~60 Å. An exact dyad is perpendicular to the superhelical axis. The distance between the axes of diagonally related helices (with the same relative orientation) is ~13 Å, whereas that between the axes of adjacent antiparallel helices is ~9 Å.

Unexpectedly, only the Leu and Ile side chains at the *d* positions interact between parallel helices (on diagonal) and stagger axially to form the hydrophobic core of the tetramer. Each cross-sectional *d* layer contains either two Leu or two Ile residues, which alternate from one pair of parallel helices to the other. In nine of ten layers, the dihedral angles χ_1 and χ_2 of the Leu and Ile side chains are approximately −60º,180º, corresponding to the most abundant rotamers for these amino acids (Ponder and Richards, 1987). Residues at positions *a* and *g* in the neighboring helices pack against the leucines and isoleucines at *d* positions to complete the hydrophobic core (see Figure 3D). Compared to the side chains of isolated helices, residues at the *a*, *d*, and *g* positions of the tetramer are substantially buried (>88%), those at the *c* and *e* positions are partly buried (~28%), while the *b* and *f* positions remain completely exposed. Approximately 6,000 $Å^2$ of solvent-accessible surface area (44% of the total accessible surface area of the four helical monomers) is buried in the tetramer.

The C44 tetramer interface shows nonclassical "knobs-to-knobs" packing of the Leu and Ile side chains at the *d* positions between parallel helices (Gottschalk, 2005; Lupas and Gruber, 2005). By contrast, adjacent

Table 1. Summary of Physicochemical Analysis

Peptide	$-[\theta]_{222}$ (deg. cm^2 $dmol^{-1}$)	T_m (ºC)	$M_{obs}/M_{monomer}$ [a]
C44	25,000	40	3.9
recSARS-1	26,500	>100	3.1
N50	31,000	63	2.9
C36	11,400	<0	– [b]
N50/C36	30,000	>100	3.0
recSARS-2	25,200	72	2.0
N37	10,900	<0	1.1
C34	10,100	<0	1.2
N37/C34	33,200	64	2.0
N37(L6)C34	32,000	72	2.0
N63	18,200	80	3.1
N34	33,000	70	3.0

[a] $M_{obs}/M_{monomer}$ is the apparent molecular mass determined from sedimentation equilibrium data divided by the expected mass of a monomer.

[b] The presence of a species that could not be assigned.

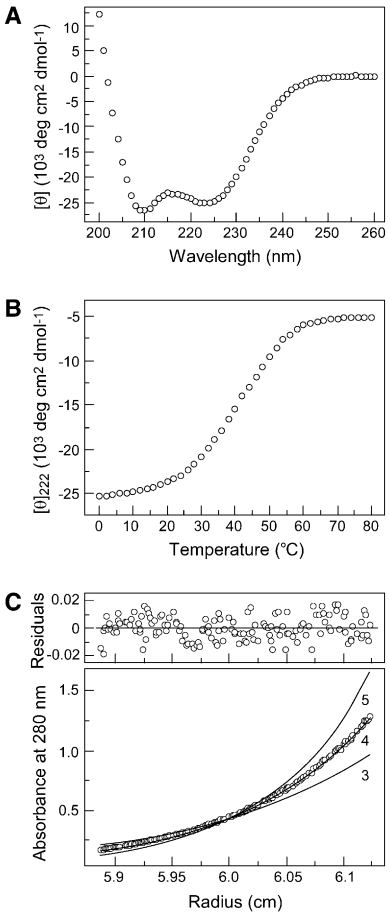

Figure 2. Folding of C44 as an α-Helical Tetramer

(A) Circular dichroism (CD) spectrum at 4ºC in PBS and 50 μM peptide concentration.

(B) Thermal melt monitored by CD at 222 nm.

(C) Sedimentation equilibrium data for a 50 μM sample at 20ºC and 25,000 rpm in PBS. The data fit closely to a tetrameric complex. Curves expected for trimeric and pentameric models are indicated for comparison. The deviation in the data from the linear fit for a tetrameric model is plotted (upper).

antiparallel helices exhibit a new type of "knobs-into-holes" interaction (Bryson et al., 1995; Crick, 1953; Harbury et al., 1993), whereby each Leu or Ile knob at a *d* position packs into a hole formed by the *a* and *g* residues of the neighboring helix and by two *d* residues in adjacent layers along the superhelical axis. This geometry results in a similar placement of atoms around the side chains at positions *a* and *g*. Knobs formed by *a* residues of one helix fit into holes formed by the spaces between the *d* and *e* residues on the neighboring helix and by two *a* residues in adjacent layers along the superhelical axis. Similarly, knobs at *g* positions pack into holes formed by the *c* and *d* residues of the neighboring helix and by two *g* residues in adjacent layers. Thus, the *a*, *d*, and *g* residues of the HR in the C44 tetramer segregate into four geometrically distinct helix-helix interfaces. By contrast, we note that residues at the *a*, *d*, and *e* positions of the lac repressor HR region participate in interhelical hydrophobic interactions (Deng et al., 2006; Friedman et al., 1995; Lewis et al., 1996). In summary, the conserved C-peptide region within the S2 ectodomain folds into a labile yet well-ordered, antiparallel, four-stranded coiled coil with unusual structural features.

The Six-Helix Bundle

To investigate the interaction between the two HR regions in S2, we constructed a recombinant protein, denoted recSARS-1, in which residues 901–973 and 1150–1193 are tethered by a glycine-rich linker (Figure 1A). CD and sedimentation equilibrium studies indicate that recSARS-1 forms an extremely stable helical trimer that does not unfold upon heating to 98ºC at 50 μM protein concentration in PBS (Table 1). Proteolysis of recSARS-1 by proteinase K yields two peptide fragments corresponding to residues 901–950 (N50) and 1150–1185 (C36) (Figure 1B). Isolated N50 folds into an α-helical, trimeric structure (the T_m of a 50 μM solution is 63ºC) (Table 1 and Figures 4A–4C), while C36 is predominantly unfolded (Figure 4A). An equimolar mixture of these two peptides is ~90% helical and sediments as a trimer of heterodimers (Table 1 and Figures 4A and 4D). The N50/C36 complex has a thermal stability that exceeds 100ºC in PBS and unfolds cooperatively with an apparent T_m of 70ºC in the presence of the denaturant GuHCl at 2 M concentration (Figure 4B). Thus, N50 and C36 associate to form an exceedingly stable six-helix bundle.

The crystal structure of the N50/C36 complex at 1.95 Å resolution (Table 2) reveals three C36 chains (partly α-helical) packed in an oblique, antiparallel manner into hydrophobic grooves on the surface of an interior, parallel coiled-coil trimer formed by three N50 helices (Figures 5A and 5B). The six-helix bundle forms an overall rod-shaped structure ~75 Å in length with a maximum diameter of 28 Å. With the exception of the C-terminal-most core residue Leu948 directed at the 3-fold axis ("*x*-like" packing) (Bullough et al., 1994), the next six *a* and *d* residues in the N50 coiled coil exhibit classical knobs-into-holes side chain packing (Figure 1B). By contrast, the six helical turns closest to the N terminus lack any regular 4-3 hydrophobic periodicity. Instead, two layers containing an *x*-like symmetric pattern (Figure 5C) alternate with two layers containing an atypical two-core residue structural motif (*da*-like layer [Gruber and Lupas, 2003]; we call it "*y*-like" packing for convenience; see Figure 5D). As a result, the N-terminal half of the superhelix is underwound to an average pitch of ~250 Å. Interestingly, the 19 residue α-helix of each C36 peptide (residues 1161–1179) intercalates into each of the grooves on the outside of the flattened coiled-coil segment in a "ridges-into-grooves" arrangement (Chothia et al., 1981). Beyond buttressing the core α helix, the extended N- and C-terminal peptide regions of C36 tuck into the adjoining grooves through hydrophobic contacts, forcing the extreme ends of the N50 and C36 chains into the rod-like structure. Fourteen amino acid residues from each C36 peptide and 38 residues from two adjacent N50 peptides contribute to an interfacial interaction that buries ~3,500 Å^2 of solvent-accessible surface area. Recently, independent results on three crystal structures of similar α-helical domains from SARS S2 have been reported (Duquerroy et al., 2005; Supekar et al., 2004; Xu et al., 2004). All four crystal structures can be superimposed on each other with

Table 2. Summary of Crystallographic Analysis

Diffraction Data							
Data Set	λ (Å)	Resolution (Å)	Number of Reflections (Total/Unique)	Completeness (%)	I/σ(I)	R_{sym} [a] (%)	Phasing Power (Ano/Iso)
C44							
Native	1.0055	45.2–1.70	66,857 (9,539)[b]	99.9 (100)	16.5 (5.7)	6.3 (38.2)	
NaBr λ1	0.9203	50–2.0	45,596 (1,0619)	97.9 (99.3)	18.8 (8.0)	6.4 (19.1)	0.4/0.5
NaBr λ2	0.9200	50–2.0	45,719 (10,700)	97.9 (99.3)	16.1 (6.8)	7.1 (24.9)	0.7/0.2
NaBr λ3	0.9070	50–2.0	44,184 (10,689)	97.6 (98.2)	15.6 (5.7)	7.8 (31.8)	0.5/–
N50/C36							
	0.9788	70.7–1.95	81,203 (32,684)	94.5 (94.8)	11.1 (3.1)	5.5 (38.7)	
N37(L6)C34							
	1.0358	72.6–1.50	69,620 (11,906)	98.0 (99.2)	15.6 (7.1)	5.8 (25.5)	
N34							
	0.9788	51.6–1.70	33,119 (11,231)	96.9 (94.7)	21.1 (5.1)	3.6 (21.0)	

Refinements	C44	N50/C36	N37(L6)C34	N34
Resolution (Å)	45.2–1.70	70.7–1.95	72.6–1.50	51.6–1.70
Number of reflections	9,076	31,031	11,341	10,695
Number of protein atoms	551	3,668	551	763
Number of water molecules	73	176	77	100
R_{cryst}/R_{free} (%)[c]	20.9/25.8	20.3/27.4	21.1/24.6	19.4/23.8
Rmsd bond lengths (Å)	0.015	0.033	0.013	0.014
Rmsd bond angles (°)	1.6	2.3	1.4	1.3
Average B factor (Å^2)	15.9	17.1	11.9	18.2
Rmsd B values (Å^2)	1.9	4.7	1.3	2.1

[a] $R_{sym} = \Sigma|I - \langle I \rangle|/\Sigma I$, where I is the integrated intensity of a given reflection.
[b] Numbers in parentheses represent the statistics for the shell comprising the outer 10% (theoretical) of the data.
[c] $R_{cryst} = \Sigma|F_o - F_c|/\Sigma F_o$, $R_{free} = R_{cryst}$ calculated by using 5% of the reflection data chosen randomly and omitted from the start of refinement.

a root mean square deviation (rmsd) for α carbon atoms of less than 1.0 Å. The high degree of sequence conservation among S2 proteins (see Figure 1B) suggests that each forms an α-helical bundle in a similar manner, but with interesting variations because some of the coronaviruses (e.g., human respiratory viruses 229E and NL63)

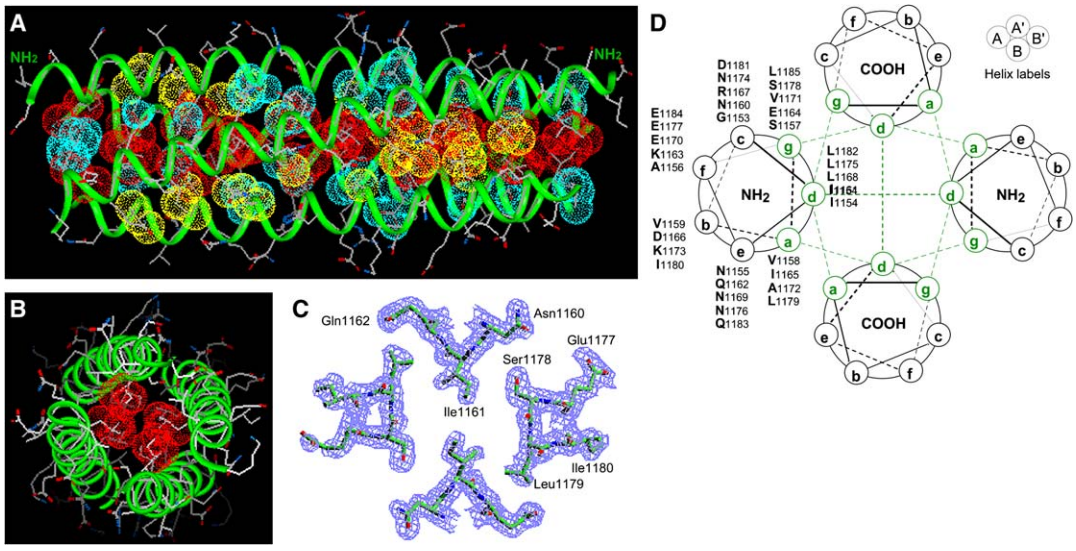

Figure 3. Crystal Structure of the C44 Tetramer

(A) Lateral view of the C44 tetramer. Yellow van der Waals surfaces identify residues at the *a* positions, red surfaces identify residues at the *d* positions, and light-blue surfaces identify residues at the *g* positions. The N termini of helices A and B are indicated.

(B) Axial view of the C44 tetramer. The red van der Waals surfaces of the Ile1154(*d*) and Leu1182(*d*) side chains are depicted.

(C) Cross-section of the tetramer in the Ile1161(*d*) layer. The 1.70 Å $2F_o - F_c$ electron density map (contoured at 1.2σ) is shown with the refined molecular model.

(D) Helical wheel representation of residues 1153–1185 of the C44 tetramer. Heptad-repeat positions are labeled *a*–*g*. The C44 helices interact through a previously uncharacterized type of packing interaction between the *a*, *d*, and *g* side chains (colored green).

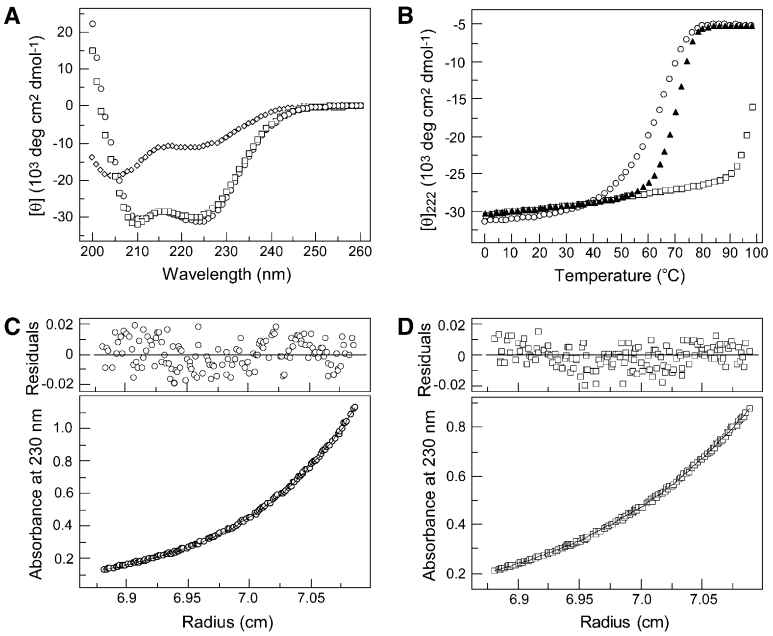

Figure 4. Solution Properties of the N50/C36 Complex

(A) CD spectra of N50 (open circles), C36 (open rhombs), and the N50/C36 complex (open squares) at 4ºC in PBS and 50 μM peptide concentration.

(B) Thermal melts of N50 and the N50/C36 complex monitored by CD at 222 nm. The filled triangles show the N50/C36 data collected in the presence of 2 M GuHCl, a denaturant.

(C) Sedimentation equilibrium data for a 50 μM sample of N50 at 20ºC and 28,000 rpm. The data fit closely to a trimeric complex. The deviation in the data from the linear fit for a trimeric model is plotted (upper).

(D) Sedimentation equilibrium data for a 50 μM sample of the N50/C36 complex at 20ºC and 17,000 rpm. The data fit closely to a trimeric complex. The deviation in the data from the linear fit for a trimeric model is plotted (upper).

contain 14 residue in-phase insertions in both HR sequences.

The Four-Helix Bundle

To test whether the C-terminal half of the predicted HR_N sequence can interact with the C-peptide region, we generated a variant of recSARS-1 that is truncated by 25 residues at the N terminus. Serendipitously, we found that this recSARS-2 molecule folds into a stable, helical, dimeric structure (T_m = 72ºC at 50 μM peptide concentration), as measured by CD and sedimentation equilibrium measurements (Table 1 and Figures 6A and 6B). Digestion of recSARS-2 by proteinase K produces two peptide fragments corresponding to residues 926–962 (N37) and 1150–1183 (C34) (Figure 1B). The isolated N37 and C34 peptides individually display little secondary structure (Table 1). Upon mixing, however, they form a dimer of heterodimers with 100% helix content (Table 1 and Figures 6A and 6C). This complex undergoes a cooperative thermal unfolding transition with a T_m of 64ºC (Figure 6B). Thus, N37 and C34 associate to form a highly stable four-helix structure. To facilitate crystallographic studies, we connected the N37 and C34 peptides via a glycine-rich linker. This subdomain, denoted N37(L6)C34, possesses the same dimeric structure and thermostability as the parent recSARS-2 molecule (Table 1).

The crystal structure of N37(L6)C34 at 1.50 Å resolution (Table 2) reveals a previously uncharacterized type of antiparallel four-α-helical bundle. Two C34 helices entwine obliquely around a parallel coiled-coil dimer formed by two N37 helices in an antiparallel orientation

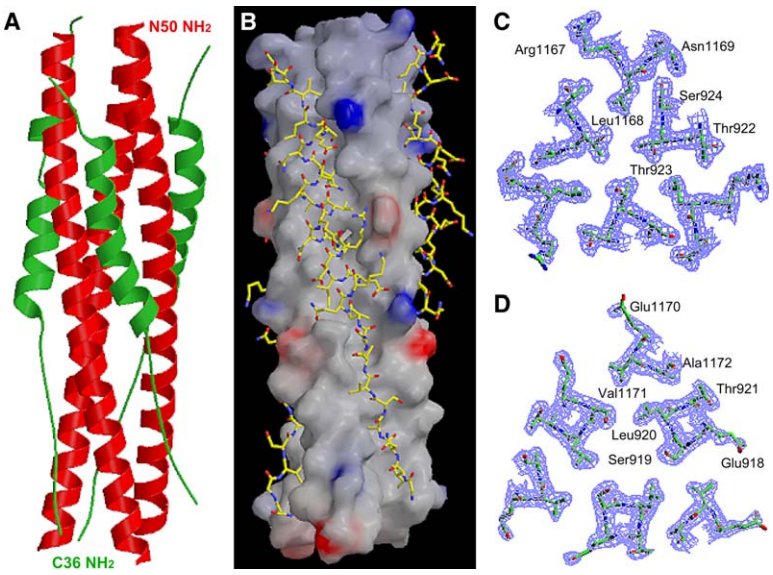

Figure 5. Crystal Structure of the N50/C36 Complex

(A) Ribbon diagram of the N50/C36 complex. The N termini of the N50 (red) and C36 (green) chains are indicated.

(B) Conserved grooves on the surface of the N50 coiled-coil trimer. The C36 peptides drawn as an atomic model are shown against a surface representation of the N50 trimer. The view is in approximately the same orientation as in (A). The solvent-accessible surface is colored according to the local electrostatic potential; colors range from dark blue, representing the most positive area, to deep red, representing the most negative area.

(C) Cross-section of the N50/C36 complex in the Thr923 layer showing "*x*-like" packing of side chains that project simultaneously toward the 3-fold axis. The $2F_o - F_c$ electron density map contoured at 1.5σ is shown with the refined molecular model.

(D) Cross-section of the N50/C36 complex in the Ser919–Leu920 layer showing "*y*-like" packing of alternating small and large side chains in a hexagonal arrangement.

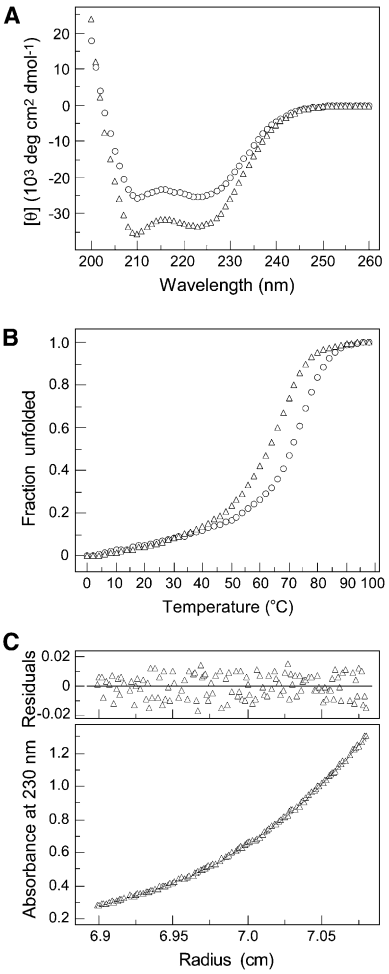

Figure 6. Solution Properties of the N37/C34 Complex

(A) CD spectra of recSARS-2 (circles) and the N37/C34 complex (triangles) at 4ºC in PBS and 50 μM peptide concentration.

(B) Thermal melts monitored by CD at 222 nm. The decrease in the fraction of a folded molecule is shown as a function of temperature.

(C) Sedimentation equilibrium data for a 50 μM sample of the N37/C34 complex at 20ºC and 25,000 rpm. The data fit closely to a dimeric complex. The deviation in the data from the linear fit for a dimeric model is plotted (upper).

(Figures 7A and 7B). This four-helix bundle comprises a cylinder ~56 Å in length and 27 Å in diameter. N37 folds into an 11-turn α helix stretching the entire length of the coiled coil, with an irregular 4-3 hydrophobic periodicity (Figure 1B). An alternating *y*- and *x*-like core packing in the middle four helical turns is flanked by the *a* and *d* layers, except for the C-terminal most Ile962 (*x*) layer. The flattening out of the superhelical twist in this central coiled-coil segment (pitch of 194 Å), in conjunction with its disrupted core, appears to be critical in guiding formation of the four-helix bundle (see below). An electrostatic potential map of the N37 dimer shows that its surface is largely uncharged (Figure 7C). The grooves that are the sites for C34 interaction are lined with predominantly hydrophobic residues. In general, residues at positions *a* and *d* of C34 pack against the *a*, *d*, *x*, and *y* residues of the N37 helices, although contacts at other positions are also observed (Figure 1B). The interaction of each C34 helix with the N37 dimer buries ~2,560 Å^2 of solvent-accessible surface area. It is noteworthy that the *y*-like packing geometry in the N37 dimer leads to the formation of two large holes in which the symmetry-related Asn937–Ala938 and Leu944–Val945 residues converge (Figures 7D–7F). However, when the two holes are occupied by the respective Leu1175 and Leu1168 residues from the abutting C34 helices, a wedges-into-grooves mode of packing (Chothia et al., 1981) results, with the interhelical Leu side chains in van der Waals contact. Such an elaborate N37 coiled-coil structure ensures a snug fit for C34. All of the interfacial residues at the buried core positions of this N37/C34 complex are highly conserved among coronavirus S2 proteins (Figure 1B), presumably reflecting selective pressure on the interactions between the HR regions in membrane fusion. So far as we know, this four-helix bundle structure has not seen before in viral fusion proteins.

The N34 Trimer

A 73-residue peptide corresponding to the predicted HR_N region of SARS-CoV S2 (residues 901–973) forms an insoluble aggregate under normal solution conditions. In contrast, a closely related 63 residue peptide (N63) beginning at residue 911 has ~50% helix content and forms a trimer, as determined by CD and sedimentation equilibrium measurements (Table 1). We localized the α-helical segment of N63 to its C-terminal half (residues 940–973) by truncation studies (Figure 1B). This 34-residue peptide (N34) remains trimeric and is fully helical (T_m = 70ºC at 50 μM peptide concentration) (Table 1 and Figures 8A and 8B), demonstrating the preference of this segment to form an autonomously folded subdomain. Its crystal structure at 1.70 Å resolution (Table 2) reveals a parallel triple-stranded α-helical coiled coil that is ~21 Å wide and ~52 Å long (Figures 8C–8E). Ten hydrophobic residues and one polar residue from each chain point into the center of the trimer (Figure 1B), and seven *a* and *d* positions show a classical acute knobs-into-holes packing characteristic of trimeric coiled coils (Harbury et al., 1994). However, three adjacent layers, starting at the third layer from the N terminus, show an unusual *x–y–x*-like core packing interaction (see Figure 1B). As a consequence, the backbones of the three helices in the N34 trimer are less sharply curved and wrap less tightly around the superhelical axis, resulting in a larger spacing between adjacent helices (12.5 Å). The presence of the *x–y–x*-like break in the classical 4-3 hydrophobic periodicity of N34 is thus mirrored in its overall coiled-coil assembly and geometry. Such recurring heptad-repeat anomalies (Gruber and Lupas, 2003) may serve a general role in facilitating polymorphic structural transitions in viral fusion proteins (Baker et al., 1999).

Potential Biological Implications

The stable six-helix bundle formed by the N50 and C36 peptides is a well-known feature of class I fusion proteins and is likely to represent the core of the postfusion trimer-of-hairpins structure of S2 (Duquerroy et al., 2005; Supekar et al., 2004; Xu et al., 2004). Current thinking postulates that a folding back of the S2 ectodomain drives the C-terminal transmembrane anchor toward the fusion peptide, allowing the two membrane attachment points to come together in the postfusion structure

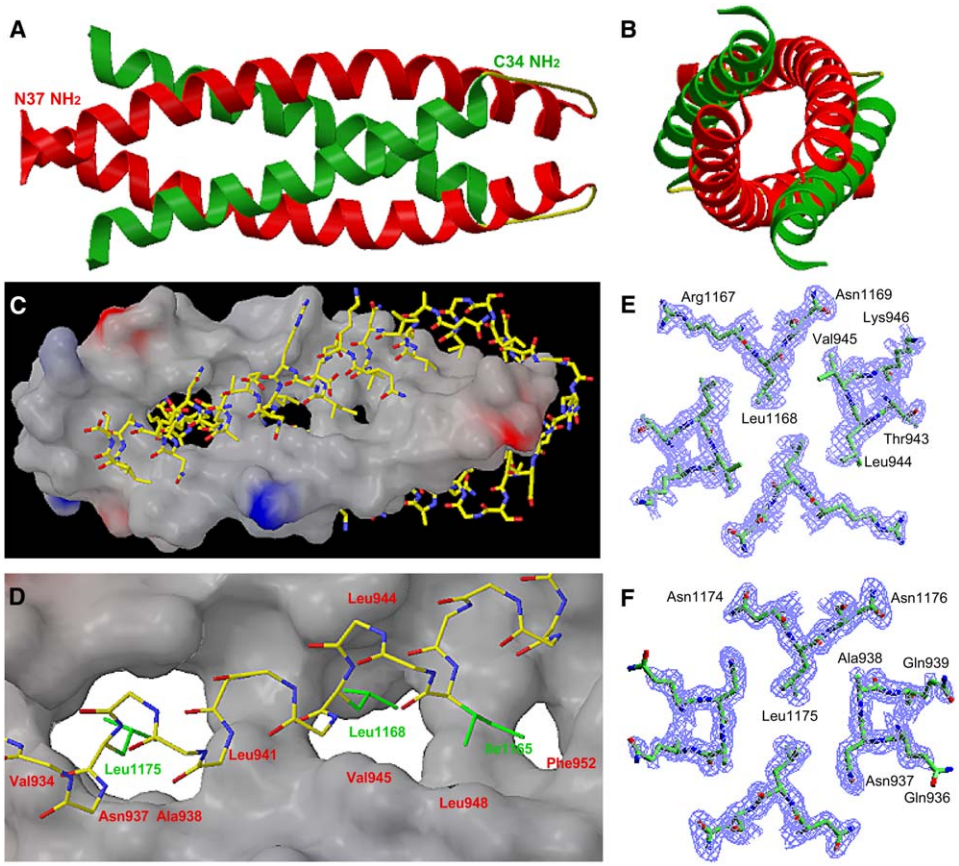

Figure 7. Crystal Structure of the N37(L6)C34 Dimer

(A) Lateral view of the N37(L6)C34 dimer. The N termini of the N37 (red) and C34 (green) helices are indicated.

(B) Axial view of the N37(L6)C34 dimer.

(C) An electrostatic surface potential representation of the N37 two-stranded coiled coil with the C34 helices drawn as an atomic model. The view represents a 45° counterclockwise rotation about the vertical axis of the lateral view in (A). The solvent-accessible surface is colored according to the local electrostatic potential; colors range from dark blue, representing the most positive area, to deep red, representing the most negative area.

(D) Close-up view of three holes on the N37 coiled-coil surface. To clarify the drawing, only side chain atoms (green) in C34 that dock into these holes are shown. Residues that form the holes in the N37 coiled-coil interface are labeled in red.

(E) Omit map showing side chain packing of Leu1168 of C34 against Leu944 and Val945 at the *y* positions of N37 in a $2F_o - F_c$ difference Fourier synthesis, calculated with the Thr943, Leu944, Val945, Lys946, Arg1167, Leu1168, and Asn1169 residues removed from the model and contoured at 1.0σ.

(F) Cross-section of the N37(L6)C34 dimer showing side chain packing of Leu1175 of C34 against Asn937 and Ala938 at the *y* positions of N37. The 1.50 Å $2F_o - F_c$ electron density map contoured at 1.5σ is shown with the refined molecular model.

(Chan et al., 1997; Tan et al., 1997; Weissenhorn et al., 1997). In the native S glycoprotein spike, the structure of the S1/S2 complex is trapped in a metastable state that precludes formation of the thermodynamically preferred postfusion conformation (Carr et al., 1997; Ruigrok et al., 1986). The structures and relative stabilities of the isolated C44 and N34 domains and the N50/C36 and N37/C34 complexes described here suggest a possible mechanism whereby the HR regions, initially sequestered in the native S spike, are sequentially refolded to promote membrane fusion, with the three distinct α-helical domains formed by the respective C44, N37(L6)C34, and N34 peptides representing different intermediate states of S2 in the fusion reaction. We should emphasize nonetheless that biological evidence for the role of these structures in the fusion process is currently lacking.

For example, the C44 structure suggests that the HR_C region, which is known to be involved in the stable postfusion state, may also regulate activation of the metastable native state. Indeed, the tetrameric symmetry of this structure is consistent with topological considerations that argue against a symmetric transition from the native oligomer to the postfusion trimer-of-hairpins structure (Root et al., 2001). In a similar vein, the α-helical rods of the N50/C36 and N37/C34 complexes share the attribute of having the helices oriented so as to direct the membrane proximal end of S2 toward the N-terminal fusion peptide. This leads us to speculate that the four-helix bundle plays a role in bringing the attached target and viral membranes into proximity, probably prior to formation of the still more stable six-helix bundle. This idea is appealing because the subsequent structural rearrangement to the trimer-of-hairpins conformation could then be coupled to the actual membrane fusion event. In this regard, it is noteworthy that cellular membrane fusion processes mediated by the SNARE proteins also

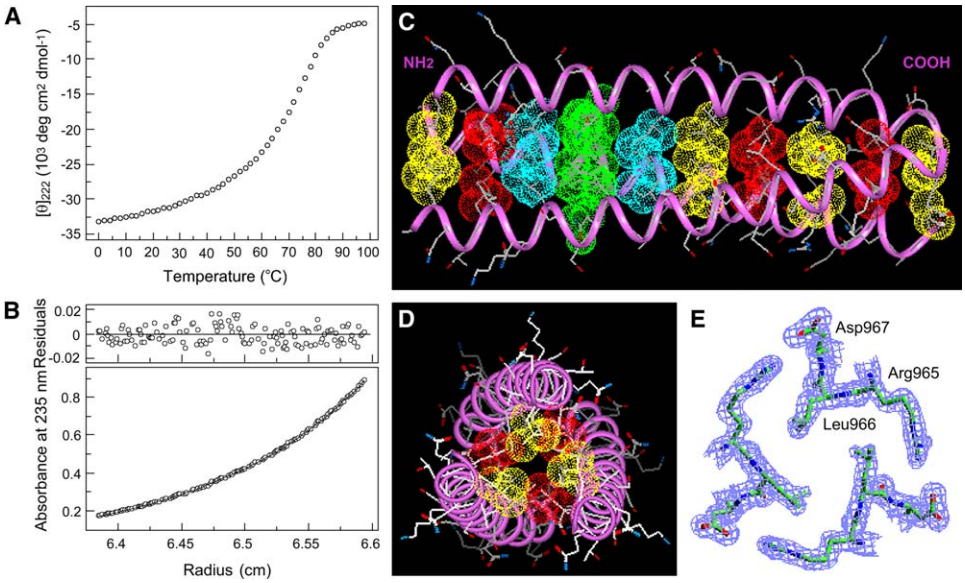

Figure 8. The N34 Peptide Forms a Parallel, Three-Stranded Coiled Coil

(A) Thermal melt monitored by CD at 222 nm in PBS and 50 μM peptide concentration.

(B) Sedimentation equilibrium data for a 200 μM sample at 20ºC and 32,000 rpm. The data fit closely to a trimeric complex. The deviation in the data from the linear fit for a trimeric model is plotted (upper).

(C) Lateral view of the N34 trimer. Yellow van der Waals surfaces identify residues at the *a* positions, red surfaces identify residues at the *d* positions, light-blue surfaces identify residues at the *x* positions, and green surfaces identify residues at the *y* positions. The N and C termini of one N34 helix are indicated.

(D) Axial view of the N34 trimer. The van der Waals surfaces of the Leu941(*a*) and Leu944(*d*) side chains are depicted. The view is from the N terminus toward the superhelical axis.

(E) Region around the Leu966(*a*) layer showing packing of core residue side chains in "holes" on the counterclockwise-related helices. The 1.70 Å $2F_o - F_c$ electron density map (contoured at 1.5σ) is shown with the refined molecular model.

rely on a four-helix bundle structure to achieve apposition of the vesicle and target membranes (Sutton et al., 1998).

Pursuing this hypothetical model further, we note that the overlapping segments of the N34 and N50 peptides (residues 941–949) adopt the same three-stranded coiled-coil conformation (with an rmsd for α carbon atoms of 0.87 Å) in the respective trimer and six-helix bundle structures, which differs from that seen for the N37 dimer in the four-helix bundle. This observation suggests that the N34 region may play a role as an adaptor to break the symmetry during the postulated S2 dimer-to-trimer transition by nucleating trimer formation during the conformational change. Cooperativity and irreversibility could be achieved when the HR_C region folds onto the newly formed HR_N trimer surface. Resolution of the four-helix bundle intermediate might thus be mechanistically and thermodynamically linked to formation of the trimer-of-hairpins structure. The highly stable N34 subdomain (with a T_m value of 70ºC, as compared to that of 63ºC for the longer N50 trimer at the same concentration) could then represent a late intermediate state of S2 in which the central HR_N trimer interactions in the postfusion state are present, but the zippering up of the outer layer has not yet begun. The structures and polymorphic interactions of the S2 HR regions presented here serve as the starting point for addressing this and other essential structural questions about the mechanism of SARS-CoV entry into cells.

Synthetic C-peptides (peptides corresponding to the C helix of S2) have been shown to inhibit SARS-CoV entry at micromolar concentrations (Bosch et al., 2004; Liu et al., 2004). C-peptides could interfere with formation of either or both of the four-helix bundle and six-helix bundle structures in a dominant-negative mechanism. Fusion inhibitors developed by using these two different strategies could even act in synergy. Speculatively, the highly conserved, large pockets in the N37 coiled-coil dimer that accommodate conserved C-peptide residues may be attractive targets for the development of new peptidomimetic or small-molecular drugs. Finally, the four-helix bundle intermediate considered here may be exposed in the S2 fusion process, in which case antibodies directed against N37(L6)C34 may inhibit viral entry.

Experimental Procedures

Protein Expression, Purification, and Proteolysis

Two distinct HR_N/HR_C proteins that differ only at their N termini were generated for this study: (1) recSARS-1, which is derived from Urbani SARS-CoV residues 901–973 and 1150–1193 connected by an -SGGRGG- linker, and (2) recSARS-2, which starts at residue 926. The HR_N/HR_C constructs were appended to the *TrpLE′* leader sequence (Kleid et al., 1981). All constructs were cloned into the pET24a vector (Novagen) by using standard molecular biology techniques. The HR_N/HR_C peptides were expressed in *E. coli* BL21 (DE3)/pLysS, purified from inclusion bodies, and cleaved from the *TrpLE′* leader sequence with cyanogen bromide as described (Shu et al., 1999). Final purification of all peptides was performed by reverse-phase HPLC on a C18 preparative column by using a water-acetonitrile gradient in the presence of 0.1% trifluoroacetic acid. Peptide identities were confirmed by electrospray mass spectrometry (PerSeptive Biosystems Voyager Elite, Cambridge, MA). Proteinase K

digestion was performed at protease:protein ratios of 1:200 (wt/wt) at room temperature in 50 mM Tris-HCl (pH 8.0). Proteolytic fragments were analyzed by reverse-phase HPLC and were assigned by N-terminal sequencing and mass spectrometry.

CD Spectropolarimetry

Circular dichroism (CD) spectra were measured on an AVIV 62A/DS spectropolarimeter (Lakewood, NJ) equipped with a thermoelectric temperature control at 4ºC in PBS (50 mM sodium phosphate [pH 7.0], 150 mM NaCl). A $[\theta]_{222}$ value of −33,000 deg. cm^2 $dmol^{-1}$ was taken to correspond to 100% helix (Chen et al., 1974). Thermal melts were performed in the same buffer and also, for the N50/C36 complex, with the addition of 2 M guanidine hydrochloride (GuHCl) to facilitate unfolding. The reversibility of thermal transitions was verified by repeated scans. In all cases, superimposable folding and unfolding curves were observed, and >90% of the signal was regained upon cooling. Values of T_m were estimated by evaluating the maximum of the first derivative of $[\theta]_{222}$ versus temperature data (in 2º steps) (Cantor and Schimmel, 1980).

Analytical Ultracentrifugation

All sedimentation equilibrium experiments were performed on a Beckman XL-A analytical ultracentrifuge (Fullerton, CA) equipped with an An-60 Ti rotor (Fullerton, CA) at 4ºC as described (Shu et al., 1999). Protein solutions were dialyzed overnight against PBS, loaded at initial concentrations of 50, 200, and 500 μM, and analyzed at two rotor speeds. Data were acquired at two wavelengths per rotor speed setting and were processed simultaneously with a nonlinear least squares fitting routine (Johnson et al., 1981). Solvent density and protein partial specific volume were calculated according to solvent and protein composition, respectively (Laue et al., 1992). Random residuals were observed in all cases. The apparent molecular masses derived from the complete data sets followed by expected molecular mass (shown in parentheses) and the rotor speeds in thousand revolutions per minute at which data were collected (shown in brackets) are as follows: C44, 19,400 (20,075) [22, 25]; recSARS-1, 41,500 (40,119) [12, 15]; N50, 15,700 (16,287) [25, 28]; N50/C36, 27,900 (28,313) [17, 20]; recSARS-2, 21,700 (21,194) [21, 24]; N37, 4,200 (3,884) [36, 39]; C34, 3,900 (3,766) [36, 39]; N37/C34, 15,600 (15,301) [25, 28]; N37(L6)C34, 15,900 (16,208) [25, 28]; N63, 21,100 (20,144) [21, 24]; N34, 11,100 (10,936) [29, 32].

Crystallization and Structure Determination

C44 was crystallized at room temperature by using the hanging drop vapor diffusion method by equilibrating a solution containing 1 μl 10 mg ml^{-1} peptide in water and 1 μl reservoir buffer against reservoir buffer (0.1 M Tris-HCl [pH 8.5], 10 mM $NiCl_2$, 22% PEG MME 2000). Crystals belong to space group C222 (a = 43.7 Å, b = 90.2 Å, c = 42.3 Å) and contain two monomers in the asymmetric unit. A useful heavy atom derivative was produced by soaking the protein crystal in the mother liquor containing 1 M sodium bromide for 3 min. The crystals were harvested in 0.1 M Tris-HCl (pH 8.0), 10 mM $NiCl_2$, 25% PEG MME 2000, 15% glycerol and were frozen in liquid nitrogen. Native and MAD data sets were collected on beamline X4A at the National Synchrotron Light Source (Table 2). All diffraction data were indexed and scaled by using DENZO and SCALEPACK (Otwinowski and Minor, 1997). Two bromide sites in the asymmetric unit were located with Solve (Terwilliger and Berendzen, 1999). Density modification and phase extension to 1.7 Å with Arp/Warp (Lamzin and Wilson, 1993) improved the quality of the electron density maps and allowed ~95% of the final structure to be interpreted. Before refinement, 5% of the reflections were set aside for calculation of R_{free} (Brunger, 1993). Iterative rounds of model building with O (Jones et al., 1991) and refinement with Refmac (Murshudov et al., 1997) as well as the addition of ordered solvent clarified the trace, except for the helix terminal regions, which are not visible in the electron density maps and therefore must be disordered. Refinement was concluded with overall anisotropic thermal factors by using TLS groups for each monomer (Schomaker and Trueblood, 1998). The final model (R_{cryst} = 20.9% and R_{free} = 25.8% for the resolution range 45.2–1.70 Å) contains residues 1150–1187 (monomer A), residues 1153–1185 (monomer B), 1 chloride ion, and 73 water molecules. The model has excellent stereochemistry and structural quality with rmsds from ideal geometry of 0.015 Å for bond lengths and 1.6º for bond angles. All main chain torsional angles fall within the helical regions of the Ramachandran plot. Some noninterpreted electron density, enough for 2 or 3 residues, can be observed at the 0.5σ contour level in σ_A-weighted electron density maps and may correspond to part of the disordered C terminus of the C44 molecule.

The N50/C36 complex was refolded by renaturation from urea and was purified by size exclusion on a Superdex 200 column. Purified complex was exchanged into 10 mM Tris-HCl (pH 8.0) and concentrated to 20 mg ml^{-1} by ultrafiltration. Crystals were obtained from 0.1 M sodium cacodylate (pH 6.4), 0.15 M zinc acetate, 18% PEG 8000 at room temperature and were cryoprotected with 15% glycerol. Crystals belong to space group P1 (a = 31.1 Å, b = 55.6 Å, c = 71.7 Å, α = 81.3º, β = 87.1º, γ = 84.1º) and contain six N50/C36 monomers in the asymmetric unit. Data were collected on beamline X4A at the National Synchrotron Light Source and were reduced and scaled with DENZO and SCALEPACK (Otwinowski and Minor, 1997) (Table 2). The structure of the N50/C36 complex was solved by molecular replacement with MOLREP (Vagin and Teplyakov, 1997) by using the MHV S2 core structure (Xu et al., 2004) as a search model. Electron density map interpretation and model building were carried out with O (Jones et al., 1991), and the structure was refined at 1.95 Å resolution by using Refmac (Murshudov et al., 1997) with TLS groups assigned for each N50 or C36 monomer (Schomaker and Trueblood, 1998). The final model (R_{cryst} = 20.3% and R_{free} = 27.4% for the resolution range 70.7–1.95 Å) consists of residues 901–949 (monomer A), residues 1154–1180 (monomer B), residues 902–947 (monomer C), residues 1154–1182 (monomer D), residues 901–949 (monomer E), residues 1153–1182 (monomer F), residues 901–950 (monomer G), residues 1150–1183 (monomer H), residues 901–949 (monomer I), residues 1153–1182 (monomer J), residues 901–950 (monomer K), and residues 1152–1185 (monomer L) in the asymmetric unit, 2 zinc ions, 6 sodium ions, 1 acetate ion, 1 cacodylate ion, and 176 water molecules. The model exhibits rmsds from ideal bond lengths and bond angles of 0.033 Å and 2.3º, respectively. 98.0% of residues occupy the most favored regions of the Ramachandran plot, with none in disallowed regions.

A 20 mg ml^{-1} stock of N37(L6)C34 was prepared in 10 mM Tris-HCl (pH 8.0) by following the procedures described above for the N50/C36 complex. Crystals of N37(L6)C34 were grown from 0.1 M sodium formate, 26%–30% PEG 1500 at room temperature. Crystals belong to space group $C222_1$ (a = 41.0 Å, b = 146.6 Å, c = 24.0 Å) and contain one N37(L6)C34 molecule in the asymmetric unit. The crystals were transferred into cryosolution containing the reservoir buffer and 15% glycerol, harvested, and frozen in liquid nitrogen. Diffraction data were collected on beamline X4A at the National Synchrotron Light Source. Reflection intensities were integrated and scaled with DENZO and SCALEPACK (Otwinowski and Minor, 1997) (Table 2). Initial phases were determined by molecular replacement with Phaser (Storoni et al., 2004) by using the structure of the C44 monomer (residues 1154–1186) as a search model. Two C44 molecules were oriented and placed in the asymmetric unit with a Z score of 6.8 and a final refined LLG of 111.9, corresponding to the N37 and C34 helical regions, respectively. This model and the data set for N37(L6)C34 were directly fed to Arp/Warp (Lamzin and Wilson, 1993), which provided a largely complete asymmetric unit of the N37(L6)C34 chain and allowed ~93% of the final model to be interpreted. The resulting experimental electron density map was of excellent quality and showed the location of most of the side chains. Crystallographic refinement of the N37(L6)C34 structure was carried out by using Refmac (Murshudov et al., 1997). Density interpretation and manual model building were done with O (Jones et al., 1991). An overall anisotropic thermal factor correlation was applied by using the N37 and C34 TLS groups (Schomaker and Trueblood, 1998). The final model (R_{cryst} = 21.1% and R_{free} = 24.6% for the resolution range 72.6–1.50 Å), which contains 75 residues (the two most C-terminal residues are disordered) and 77 water molecules, was verified by omit maps. The model fits the $2F_o - F_c$ map well, and the $F_o - F_c$ map contoured at ± 3σ has no interpretable features. Bond lengths and bond angles of the model have rmsds from ideality of 0.013 Å and 1.4º, respectively. All main chain dihedral angles but one (Ile1183) fall within the most preferred regions of the Ramanchandran space, and most side chains assume well-populated rotamer conformations. Ile1183 lies in an additionally allowed

region of the Ramachandran space and is the second residue from the C terminus of the structured C34 model.

N34 was crystallized from 7.5 mg ml^{-1} peptide in water, 0.1 M sodium citrate (pH 5.0), 16% PEG 4000, 20% isopropanol at room temperature. Crystals belong to space group C2 (a = 58.9 Å, b = 34.4 Å, c = 51.7 Å, β = 93.9º) and contain three monomers in the asymmetric unit. Cryoprotection was achieved by raising the concentration of PEG 4000 to 20% supplemented with 15% glycerol. The crystals were frozen in liquid nitrogen, and data were collected on beamline X4A at the National Synchrotron Light Source. The data were processed with DENZO and SCALEPACK (Otwinowski and Minor, 1997) (Table 2). The initial model of the N34 trimer was obtained by molecular replacement with Phaser (Storoni et al., 2004) by using the N50 trimer structure as a search model. The coordinates were refined against data up to 1.70 Å resolution by using Refmac (Murshudov et al., 1997), followed by manual rebuilding with O (Jones et al., 1991), in iterative cycles. Later cycles included overall anisotropic refinement with TLS groups for each monomer (Schomaker and Trueblood, 1998). The final model (R_{cryst} = 19.4% and R_{free} = 23.8% for the resolution range 51.6–1.70 Å) contains residues 940–973 (monomer A), residues 941–973 (monomer B), residues 940–973 (monomer C), and 100 water molecules. All protein residues are in the most favored regions of the Ramachandran plot.

Structure Analysis

Rmsds were calculated with LSQKAB in the CCP4i program suite (Potterton et al., 2003). Buried surface areas were calculated from the difference of the accessible side chain surface areas of the oligomer structure and of the individual helical monomers by using CNS 1.0 (Brunger et al., 1998). To calculate an omit map, target residues were removed, and the remaining atoms were shaken randomly by 0.3 Å to minimize model bias and then refined. Figures were generated with SETOR (Evans, 1993), Insight II (Accelrys, San Diego, CA), and GRASP (Nicholls et al., 1991).

Acknowledgments

We thank John Schwanof at the National Synchrotron Light Source for support at beamline X4A and Neville Kallenbach, Jack Nunberg, and David Eliezer for comments on the manuscript. This work was supported by National Institutes of Health grant AI511151 and by the Irma T. Hirschl Trust.

Received: January 16, 2006
Revised: March 1, 2006
Accepted: March 15, 2006
Published: May 16, 2006

References

Baker, K.A., Dutch, R.E., Lamb, R.A., and Jardetzky, T.S. (1999). Structural basis for paramyxovirus-mediated membrane fusion. Mol. Cell *3*, 309–319.

Bosch, B.J., van der Zee, R., de Haan, C.A., and Rottier, P.J. (2003). The coronavirus spike protein is a class I virus fusion protein: structural and functional characterization of the fusion core complex. J. Virol. *77*, 8801–8811.

Bosch, B.J., Martina, B.E., Van Der Zee, R., Lepault, J., Haijema, B.J., Versluis, C., Heck, A.J., De Groot, R., Osterhaus, A.D., and Rottier, P.J. (2004). Severe acute respiratory syndrome coronavirus (SARS-CoV) infection inhibition using spike protein heptad repeat-derived peptides. Proc. Natl. Acad. Sci. USA *101*, 8455–8460.

Brunger, A.T. (1993). Assessment of phase accuracy by cross validation: the free R value. Methods and applications. Acta Crystallogr. D Biol. Crystallogr. *49*, 24–36.

Brunger, A.T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.S., Kuszewski, J., Nilges, M., Pannu, N.S., et al. (1998). Crystallography & NMR system: a new software suite for macromolecular structure determination. Acta Crystallogr. D Biol. Crystallogr. *54*, 905–921.

Bryson, J.W., Betz, S.F., Lu, H.S., Suich, D.J., Zhou, H.X., O'Neil, K.T., and DeGrado, W.F. (1995). Protein design: a hierarchic approach. Science *270*, 935–941.

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

Cantor, C., and Schimmel, P. (1980). Biophysical Chemistry, *Volume III* (New York: W.H. Freeman and Co.).

Carr, C.M., Chaudhry, C., and Kim, P.S. (1997). Influenza hemagglutinin is spring-loaded by a metastable native conformation. Proc. Natl. Acad. Sci. USA *94*, 14306–14313.

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

Chen, Y.H., Yang, J.T., and Chau, K.H. (1974). Determination of the helix and β form of proteins in aqueous solution by circular dichroism. Biochemistry *13*, 3350–3359.

Chothia, C., Levitt, M., and Richardson, D. (1981). Helix to helix packing in proteins. J. Mol. Biol. *145*, 215–250.

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

De Groot, R.J., Van Leen, R.W., Dalderup, M.J., Vennema, H., Horzinek, M.C., and Spaan, W.J. (1989). Stably expressed FIPV peplomer protein induces cell fusion and elicits neutralizing antibodies in mice. Virology *171*, 493–502.

Deng, Y., Liu, J., Zheng, Q., Eliezer, D., Kallenbach, N.R., and Lu, M. (2006). Antiparallel four-stranded coiled coil specified by a 3-3-1 hydrophobic heptad repeat. Structure *14*, 247–255.

Duquerroy, S., Vigouroux, A., Rottier, P.J., Rey, F.A., and Bosch, B.J. (2005). Central ions and lateral asparagine/glutamine zippers stabilize the post-fusion hairpin conformation of the SARS coronavirus spike glycoprotein. Virology *335*, 276–285.

Earp, L.J., Delos, S.E., Park, H.E., and White, J.M. (2005). The many mechanisms of viral membrane fusion proteins. Curr. Top. Microbiol. Immunol. *285*, 25–66.

Eckert, D.M., and Kim, P.S. (2001). Mechanisms of viral membrane fusion and its inhibition. Annu. Rev. Biochem. *70*, 777–810.

Evans, S.V. (1993). SETOR: hardware-lighted three-dimensional solid model representations of macromolecules. J. Mol. Graph. *11*, 134–138, 127–138.

Frana, M.F., Behnke, J.N., Sturman, L.S., and Holmes, K.V. (1985). Proteolytic cleavage of the E2 glycoprotein of murine coronavirus: host-dependent differences in proteolytic cleavage and cell fusion. J. Virol. *56*, 912–920.

Friedman, A.M., Fischmann, T.O., and Steitz, T.A. (1995). Crystal structure of lac repressor core tetramer and its implications for DNA looping. Science *268*, 1721–1727.

Gottschalk, K.E. (2005). A coiled-coil structure of the alphaIIbbeta3 integrin transmembrane and cytoplasmic domains in its resting state. Structure *13*, 703–712.

Gruber, M., and Lupas, A.N. (2003). Historical review: another 50th anniversary—new periodicities in coiled coils. Trends Biochem. Sci. *28*, 679–685.

Guo, Q., Ho, H.T., Dicker, I., Fan, L., Zhou, N., Friborg, J., Wang, T., McAuliffe, B.V., Wang, H.G., Rose, R.E., et al. (2003). Biochemical and genetic characterizations of a novel human immunodeficiency virus type 1 inhibitor that blocks gp120–CD4 interactions. J. Virol. *77*, 10528–10536.

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

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

Harrison, S.C. (2005). Mechanism of membrane fusion by viral envelope proteins. Adv. Virus Res. *64*, 231–261.

Ingallinella, P., Bianchi, E., Finotto, M., Cantoni, G., Eckert, D.M., Supekar, V.M., Bruckmann, C., Carfi, A., and Pessi, A. (2004). Structural characterization of the fusion-active complex of severe acute respiratory syndrome (SARS) coronavirus. Proc. Natl. Acad. Sci. USA *101*, 8709–8714.

Johnson, M.L., Correia, J.J., Yphantis, D.A., and Halvorson, H.R. (1981). Analysis of data from the analytical ultracentrifuge by nonlinear least-squares techniques. Biophys. J. *36*, 575–588.

Jones, T.A., Zou, J.Y., Cowan, S.W., and Kjeldgaard. (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. Acta Crystallogr. A *47*(Pt 2), 110–119.

Kleid, D.G., Yansura, D., Small, B., Dowbenko, D., Moore, D.M., Grubman, M.J., McKercher, P.D., Morgan, D.O., Robertson, B.H., and Bachrach, H.L. (1981). Cloned viral protein vaccine for foot-and-mouth disease: responses in cattle and swine. Science *214*, 1125–1129.

Lai, M.M.C., and Holmes, K.V. (2001). Coronaviridae: the viruses and their replication. In Fields Virology, D.M. Knipe and P.M. Howley, eds. (Philadelphia: Lippincott Williams & Wilkins), pp. 1163–1185.

Lamzin, V.S., and Wilson, K.S. (1993). Automated refinement of protein models. Acta Crystallogr. D Biol. Crystallogr. *49*, 129–149.

Laue, T.M., Shah, B.D., Ridgeway, T.M., and Pelletier, S.L. (1992). Computer-aided interpretation of analytical sedimentation data for proteins. In Analytical Ultracentrifugation in Biochemistry and Polymer Science, J.C. Horton, ed. (Cambridge, UK: Royal Society of Chemistry), pp. 90–125.

Lewis, M., Chang, G., Horton, N.C., Kercher, M.A., Pace, H.C., Schumacher, M.A., Brennan, R.G., and Lu, P. (1996). Crystal structure of the lactose operon repressor and its complexes with DNA and inducer. Science *271*, 1247–1254.

Liu, S., Xiao, G., Chen, Y., He, Y., Niu, J., Escalante, C.R., Xiong, H., Farmar, J., Debnath, A.K., Tien, P., and Jiang, S. (2004). Interaction between heptad repeat 1 and 2 regions in spike protein of SARS-associated coronavirus: implications for virus fusogenic mechanism and identification of fusion inhibitors. Lancet *363*, 938–947.

Lupas, A.N., and Gruber, M. (2005). The structure of α-helical coiled coils. Adv. Protein Chem. *70*, 37–78.

Murshudov, G.N., Vagin, A.A., and Dodson, E.J. (1997). Refinement of macromolecular structures by the maximum-likehood method. Acta Crystallogr. D Biol. Crystallogr. *53*, 240–255.

Nicholls, A., Sharp, K.A., and Honig, B. (1991). Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. Proteins *11*, 281–296.

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

Ponder, J.W., and Richards, F.M. (1987). Tertiary templates for proteins. Use of packing criteria in the enumeration of allowed sequences for different structural classes. J. Mol. Biol. *193*, 775–791.

Potterton, E., Briggs, P., Turkenburg, M., and Dodson, E. (2003). A graphical user interface to the CCP4 program suite. Acta Crystallogr. D Biol. Crystallogr. *59*, 1131–1137.

Root, M.J., Kay, M.S., and Kim, P.S. (2001). Protein design of an HIV-1 entry inhibitor. Science *291*, 884–888.

Ruigrok, R.W., Martin, S.R., Wharton, S.A., Skehel, J.J., Bayley, P.M., and Wiley, D.C. (1986). Conformational changes in the hemagglutinin of influenza virus which accompany heat-induced fusion of virus with liposomes. Virology *155*, 484–497.

Schomaker, V., and Trueblood, K.N. (1998). Correlation of internal torsional motion with overall molecular motion in crystals. Acta Crystallogr. B *54*, 507–514.

Shu, W., Ji, H., and Lu, M. (1999). Trimerization specificity in HIV-1 gp41: analysis with a GCN4 leucine zipper model. Biochemistry *38*, 5378–5385.

Skehel, J.J., and Wiley, D.C. (2000). Receptor binding and membrane fusion in virus entry: the influenza hemagglutinin. Annu. Rev. Biochem. *69*, 531–569.

Storoni, L.C., McCoy, A.J., and Read, R.J. (2004). Likelihood-enhanced fast rotation functions. Acta Crystallogr. D Biol. Crystallogr. *60*, 432–438.

Sturman, L.S., Ricard, C.S., and Holmes, K.V. (1985). Proteolytic cleavage of the E2 glycoprotein of murine coronavirus: activation of cell-fusing activity of virions by trypsin and separation of two different 90K cleavage fragments. J. Virol. *56*, 904–911.

Supekar, V.M., Bruckmann, C., Ingallinella, P., Bianchi, E., Pessi, A., and Carfi, A. (2004). Structure of a proteolytically resistant core from the severe acute respiratory syndrome coronavirus S2 fusion protein. Proc. Natl. Acad. Sci. USA *101*, 17958–17963.

Sutton, R.B., Fasshauer, D., Jahn, R., and Brunger, A.T. (1998). Crystal structure of a SNARE complex involved in synaptic exocytosis at 2.4 Å resolution. Nature *395*, 347–353.

Tan, K., Liu, J., Wang, J., Shen, S., and Lu, M. (1997). Atomic structure of a thermostable subdomain of HIV-1 gp41. Proc. Natl. Acad. Sci. USA *94*, 12303–12308.

Terwilliger, T.C., and Berendzen, J. (1999). Automated MAD and MIR structure solution. Acta Crystallogr. D Biol. Crystallogr. *55*, 849–861.

Tripet, B., Howard, M.W., Jobling, M., Holmes, R.K., Holmes, K.V., and Hodges, R.S. (2004). Structural characterization of the SARS-coronavirus spike S fusion protein core. J. Biol. Chem. *279*, 20836–20849.

Vagin, A., and Teplyakov, A. (1997). MOLREP: an automated program for molecular replacement. J. Appl. Crystallogr. *30*, 1022–1025.

Wang, S., York, J., Shu, W., Stoller, M.O., Nunberg, J.H., and Lu, M. (2002). Interhelical interactions in the gp41 core: implications for activation of HIV-1 membrane fusion. Biochemistry *41*, 7283–7292.

Weissenhorn, W., Dessen, A., Harrison, S.C., Skehel, J.J., and Wiley, D.C. (1997). Atomic structure of the ectodomain from HIV-1 gp41. Nature *387*, 426–430.

Weissenhorn, W., Dessen, A., Calder, L.J., Harrison, S.C., Skehel, J.J., and Wiley, D.C. (1999). Structural basis for membrane fusion by enveloped viruses. Mol. Membr. Biol. *16*, 3–9.

Xu, Y., Lou, Z., Liu, Y., Pang, H., Tien, P., Gao, G.F., and Rao, Z. (2004). Crystal structure of severe acute respiratory syndrome coronavirus spike protein fusion core. J. Biol. Chem. *279*, 49414–49419.

Yin, H.S., Wen, X., Paterson, R.G., Lamb, R.A., and Jardetzky, T.S. (2006). Structure of the parainfluenza virus 5 F protein in its metastable, prefusion conformation. Nature *439*, 38–44.

Accession Numbers

Coordinates and structure factors have been deposited in the Protein Data Bank under ID codes 1ZV7 (C44), 1ZV8 (N50/C36), 1ZVA (N37/C34), and 1ZVB (N34).