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.

THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 279, No. 20, Issue of May 14, pp. 20836–20849, 2004
Printed in U.S.A.

Structural Characterization of the SARS-Coronavirus Spike S Fusion Protein Core*

Received for publication, January 23, 2004, and in revised form, March 1, 2004
Published, JBC Papers in Press, March 2, 2004, DOI 10.1074/jbc.M400759200

Brian Tripet‡, Megan W. Howard§, Michael Jobling§, Randall K. Holmes§, Kathryn V. Holmes§, and Robert S. Hodges‡¶

From the ‡Department of Biochemistry and Molecular Genetics and the §Department of Microbiology, University of Colorado Health Sciences Center, Denver, Colorado 80262

The spike (S) glycoprotein of coronaviruses mediates viral entry into host cells. It is a type 1 viral fusion protein that characteristically contains two heptad repeat regions, denoted HR-N and HR-C, that form coiled-coil structures within the ectodomain of the protein. Previous studies have shown that the two heptad repeat regions can undergo a conformational change from their native state to a 6-helix bundle (trimer of dimers), which mediates fusion of viral and host cell membranes. Here we describe the biophysical analysis of the two predicted heptad repeat regions within the severe acute respiratory syndrome coronavirus S protein. Our results show that in isolation the HR-N region forms a stable α-helical coiled coil that associates in a tetrameric state. The HR-C region in isolation formed a weakly stable trimeric coiled coil. When mixed together, the two peptide regions (HR-N and HR-C) associated to form a very stable α-helical 6-stranded structure (trimer of heterodimers). Systematic peptide mapping showed that the site of interaction between the HR-N and HR-C regions is between residues 916–950 of HR-N and residues 1151–1185 of HR-C. Additionally, interchain disulfide bridge experiments showed that the relative orientation of the HR-N and HR-C helices in the complex was antiparallel. Overall, the structure of the heterostranded complex is consistent with the structures observed for other type 1 viral fusion proteins in their fusion-competent state.

In the Fall of 2002, the World Health Organization witnessed an outbreak of atypical pneumonia, termed severe acute respiratory syndrome (SARS),[1] that quickly spread to >25 countries world wide. The causative agent of SARS was quickly identified to be a previously unknown member of the family of coronaviruses (1–4). The coronaviruses are a diverse group of large enveloped, positive-stranded RNA viruses that cause respiratory and enteric diseases in humans and other animals. For example, previously known human coronaviruses, HCoV-229E and HCoV-OC43, cause up to 30% of common colds and rarely cause pneumonia in older adults or immunocompromised patients (5–7). Coronaviruses of animals (*e.g.* porcine transmissible gastroenteritis virus, murine hepatitis virus (MHV), and avian infectious bronchitis virus) cause respiratory, gastrointestinal, neurological, or hepatic disease in their respective hosts (8). In the case of the SARS-associated coronavirus (SARS-CoV), the nucleotide sequence shows closest identity to novel coronavirus strains isolated from Himalayan palm civets and a raccoon dog in the Guandong province of China (9, 10). The human isolate of SARS-CoV can also grow in monkeys, mice, cats, and ferrets.

Coronavirus infection is achieved through fusion of the lipid bilayer of the viral envelope with host cell membranes. Membrane fusion is mediated by the viral spike (S) glycoprotein on the viral envelope (11–14). The S glycoprotein is synthesized as an ~180-kDa precursor that oligomerizes in the endoplasmic reticulum and is incorporated into budding virions in a pre-Golgi compartment. In some strains, S is cleaved by trypsin or related enzymes to yield two non-covalently associated subunits, S1 and S2 (15, 16). S1 contains the receptor-binding site and thus defines the host range of the virus (17, 18). S2 is the transmembrane subunit that mediates fusion between viral and cellular membranes. S2 contains two 4,3-hydrophobic repeat domains (HR) that are predicted to form coiled-coil structures (14, 19). These regions are denoted HR-N and HR-C and are separated by an intervening stretch of ~140 amino acid residues called the interhelical domain. These coiled-coil regions are thought to play an important role in defining the oligomeric structure of the spike protein in its native state and its fusogenic ability (20). Furthermore, the presence of the HR regions in conjunction with recent studies by Bosch *et al.* (19) indicate that coronavirus spike proteins can be classified as type 1 viral fusion proteins.

Binding of S protein of murine coronavirus MHV to a specific soluble or cell surface glycoprotein receptor induces global changes in the conformation of S protein that converts S protein from a native pre-fusogenic state to a fusion active state (Fig. 1). In the case of the MHV-JHM strain, interaction with the host cell receptor CEACAM1 causes the S1 domain to readily dissociate from the S1/S2 complex, resulting in display of a new hydrophobic surface area that allows the virions to bind to liposomes and exposes previously hidden trypsin cleavage sites in S2 (21–23). This receptor-induced conformational change is believed to facilitate the release of a hydrophobic fusion peptide from the interior of S2 and position it to interact with the host cell membrane. Next, the fusion active state of type 1 viral fusion proteins resolves itself to a fusogenic or post-fusion state (24). This latter change is thought to arise from an association between the two coiled-coil regions within

* This work was supported by National Institutes of Health Grants R01AI48717 (to R. S. H.), 940 (to R. K. H.), and RO1AI25231 (to K. V. H.) and the John Stewart Chair in Peptide Chemistry (to R. S. H.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

¶ To whom correspondence should be addressed. Tel.: 303-315-8837; Fax: 303-315-1153; E-mail: robert.hodges@uchsc.edu.

[1] The abbreviations used are: SARS, severe acute respiratory syndrome; SARS-CoV, SARS-associated coronavirus; TFE, trifluoroethanol; HPLC, high performance liquid chromatography; MHV, murine hepatitis virus; S, spike; HR, 4,3-hydrophobic repeat; SEC, size-exclusion chromatography; HA, hemagglutinin.

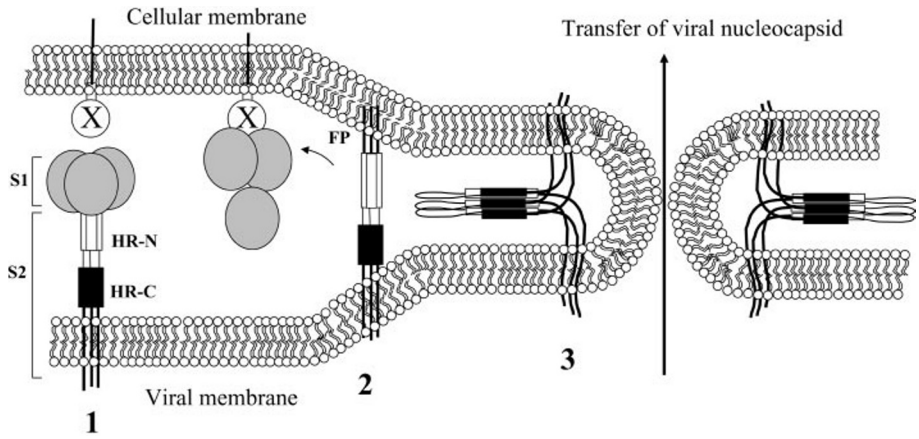

FIG. 1. **Schematic representation of the three different states of the coronavirus S fusion protein during viral entry.** *State 1* (native), the S protein is denoted S1 and S2 for its N- and C-terminal domains; *state 2* (intermediate state), the N-terminal (S1) domain is dissociated (shedding) to expose the fusion peptide (*FP*) region, and *state 3* (collapsed 6-helix bundle or fusion active state), the collapsed S2 domain draws the viral and cellular membranes together causing fusion and release of the viral nucleocapsid into the host cell. HR-N and HR-C denote coiled coils at the N terminus and C terminus of the S2 domain. *X* denotes the host cell surface receptor.

the S2 subunit, in which the N-terminal HR region from three S proteins forms a parallel triple-stranded coiled coil and the interhelical polypeptide chain loops around to reverse the polypeptide chain direction, positioning the HR-C region antiparallel to HR-N. The residues at the a and d positions (of the heptad repeat positions denoted $(abcdefg)_n$) of HR-C then pack into grooves formed by the residues at the a, d, e, and g positions of the HR-N core coiled coil to complete the 6-stranded α-helical bundle structure also termed "trimer of hairpins" or "trimer of dimers." This structure ultimately draws the viral and cellular membranes close together, destabilizing the lipid bilayers that surround the virus and target cell, causing fusion and release of the viral nucleocapsid into the host cell. The high resolution structures of the ectodomains of type 1 viral fusion proteins from many enveloped viruses show very similar structures in the fusogenic state, suggesting a general membrane fusion mechanism (25–36).

Because fusion mediated by several type 1 viral fusion proteins, *e.g.* human immunodeficiency virus, gp160, can be inhibited by peptide mimics of the HR-N or HR-C regions (19, 37–46), and there is a pressing need for effective anti-viral therapies to treat SARS, we have carried out a detailed study of the biophysical properties of the HR regions within the ectodomain of the SARS-CoV S glycoprotein. We demonstrate that the HR-N and HR-C regions of the SARS-CoV S glycoprotein can independently form α-helical coiled-coil structures. Furthermore, we show that a mixture of HR-N and HR-C can form a very stable trimer of dimers structure similar to other type 1 viral fusion proteins with the two helical HR-N and HR-C regions running anti-parallel to each other. Systematic peptide mapping has localized the site of interaction to residues 916–950 within the HR-N domain and 1151–1185 in the HR-C domain.

MATERIALS AND METHODS

Peptide Synthesis—The HR-N and HR-C peptides of SARS CoV S glycoprotein were prepared by solid-phase synthesis methodology using 4-benzylhydrylamine hydrochloride resin with conventional *N-t*-butyloxycarbonyl chemistry as described by Tripet *et al.* (47). Peptides were N-terminally acetylated, cleaved from the resin, and purified by reversed-phase high performance liquid chromatography (HPLC) to homogeneity and characterized by amino acid analysis and electrospray mass spectrometry.

Plasmid Construction—Production of the peptide corresponding to SARS S protein amino acids 882–973 (HR-N1) was done using directional sub-cloning and bacterial expression techniques. PCR fragments were prepared from the plasmid SARS S number 18 (containing the entire S protein from SARS, Urbani strain (GenBank™ AY278741)). Primers were designed to incorporate an NdeI restriction site upstream (GTACGTACGCATATGATGCAAATGGCATATAGGTTC) and an EcoRI site downstream (GCGAATTCCCTTTGTCGTCGTCGTCGCAGCCGCCTACCTCCGCCTCGACTTTATCAAG) of the amplified fragment. The NdeI site contains an ATG start codon just upstream of the SARS sequence. The reverse primer was engineered to incorporate the amino acid sequence Gly-Gly-Cys-Asp-Asp-Asp-Asp-Lys to insert an enterokinase site (DDDDK) just downstream of the synthesized sequence and just upstream of the EcoRI site. The nucleotide fragment corresponding to residues 882–973 was amplified by PCR and subcloned into the EcoRI/NdeI site of the pT7SH6 plasmid, in-frame with the His_6 tag directly downstream of the EcoRI site to create the plasmid pT7SH6 SARS-(882–973). pT7SH6 contains an oligonucleotide-encoded His_6 tag followed by two stop codons cloned into the EcoRI-BamHI sites of pT7.7 so that cloning into the EcoRI site creates a C-terminal protein fusion encoding NSHHHHHH*XX*. Plasmid pT7.7 (based on the plasmids of Tabor and Richardson (48)) enables genes to be expressed by T7 RNA polymerase using the highly efficient translation signals of T7 gene 10 by cloning into an NdeI (CATATG) site where the ATG is the initiation codon. Plasmid SARS S number 18 was created by cloning the full-length SARS S protein into pcDNA3.1/V5-His TA cloning vector (Invitrogen). The SARS S sequence was amplified using reverse transcription from SARS genomic RNA (kindly provided by W. Bellini under a material transfer agreement with the Centers of Disease Control, Atlanta, GA), followed by PCR using primers F21488 (CACCATGTTTATTTTCTTATTATTT) and R25259STOP (TTATGTGTAATGTAATTTGACACC). The reverse primer contains a stop codon to halt transcription before the V5-His_6 tag in the cloning vector.

Bacterial Protein Expression and Purification—BL21 AI *Escherichia coli* cells (Invitrogen) were transformed with pT7SH6 SARS-(882–973). These cells were grown overnight in Luria Broth (LB) containing ampicillin (Sigma). The overnight culture was used to inoculate a flask of LB, containing ampicillin, and the cells were grown to an A_{600} of 0.7. The cells were induced by adding 10% L-arabinose (Sigma) to a final concentration of 0.2%. Three hours later the cells were pelleted, resuspended in 1/10th volume of 20 mM Tris/HCl, pH 8.0, and protease inhibitors (Complete EDTA-free, Roche Applied Science). The sample was frozen at −80 °C and thawed at 50 °C three times. 1 mg of DNase (Promega, Madison, WI) was added to the sample and incubated at room temperature for 20 min. The sample was centrifuged at 14,000 × *g* for 10 min. The soluble supernatant medium was purified with Hi-Trap nickel affinity columns (Amersham Biosciences) as described previously (49). The column was washed with 20 mM Tris/HCl, pH 8.0, and bound proteins were eluted with a 0–500 mM gradient of imidazole (Sigma) in 20 mM Tris/HCl, pH 8.0. Fractions containing the SARS-(882–973) (HR-N1) peptide were identified using immunoblot analysis with an anti-His_6 polyclonal antibody (AbCam, Cambridge, MA). These fractions were pooled and placed over a reversed phase chromatography column for purification.

Circular Dichroism Spectroscopy—Circular dichroism spectra were recoded on a Jasco J-810 spectropolarimeter (Jasco Inc., Easton, MD). The CD wave scans were measured from 190 to 255 nm in benign buffer (0.1 M KCl, 0.05 M PO_4, pH 7). For samples containing trifluoroethanol (TFE), the above buffer was diluted 1:1 (v/v) with TFE. Temperature denaturation midpoints ($T_{½}$) for the peptides were determined by following the change in molar ellipticity at 222 nm from 4 to 95 °C in a 1-mm path length cell and a temperature increase rate of 1 °C/min. Ellipticity readings were normalized to the fraction of peptide folded (f_f) or unfolded (f_u) using the standard equations $f_f = ([\theta] - [\theta]_u)/([\theta]_n - [\theta]_u)$ and $f_u = (1 - f_f)$; where $[\theta]_n$ and $[\theta]_u$ represent the ellipticity values for the fully folded and fully unfolded species, respectively. $[\theta]$ is the observed ellipticity at 222 nm at any temperature.

Sedimentation Analysis—Sedimentation equilibrium experiments

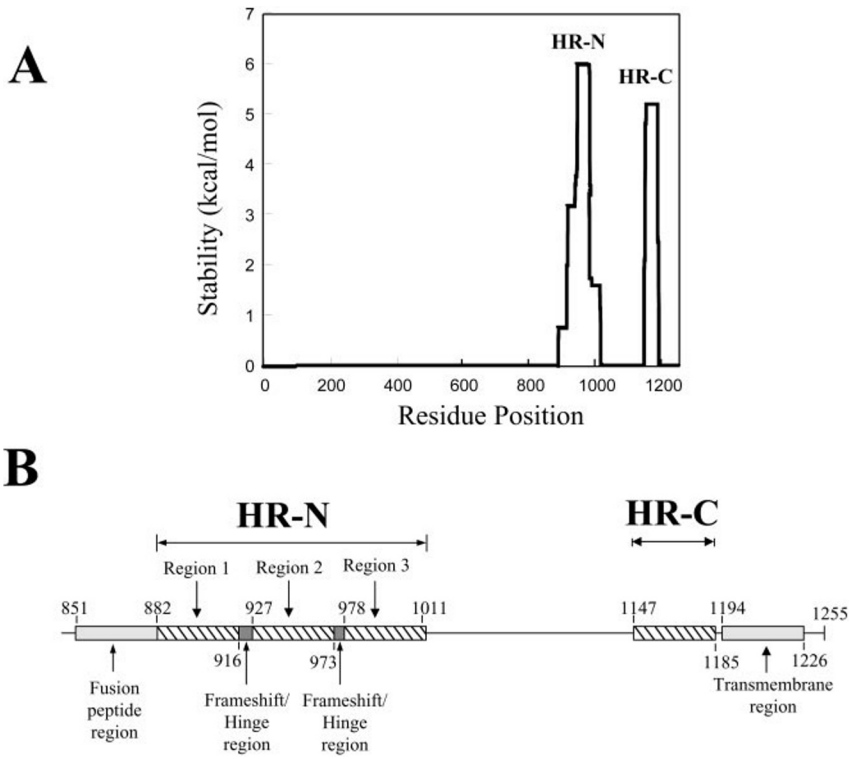

FIG. 2. **Coiled-coil prediction analysis of the SARS-CoV S glycoprotein.** *A*, plot of the predicted stability *versus* sequence position for the SARS-CoV S protein using the coiled-coil prediction algorithm STABLECOIL (53) with a 35-residue window width. The two regions of interest to this study are denoted HR-N (for N-terminal heptad repeat) and HR-C (for C-terminal heptad repeat). *B*, schematic representation of the relative locations of the predicted coiled-coils HR-N and HR-C within the S protein of the SARS-CoV. The location of the predicted transmembrane spanning region and fusion peptide region are also shown. HR-N can be divided into three sub-regions (*regions 1–3*) based on the assignment of the a and d positions. Changes in the a and d register are denoted as a frameshift/hinge region.

were performed on the analytical ultracentrifuge (XLA from Beckman-Coulter) as described by Tripet *et al.* (47). In general HR-N and HR-C peptides were dialyzed exhaustively against 50 mM K_2HPO_4, 100 mM KCl, pH 7 buffer at 4 °C. For full-length HR-N1 peptide, HR-N1 was dialyzed against the low salt containing buffer 50 mM K_2HPO_4, pH 6, for enhanced solubility. Three 100-μl aliquots of the sample were loaded into a 12-mm 6-sector, charcoal-filled Epon cell and centrifugation proceeded for 48 h with rotor speeds of 20,000, 26,000, and 30,000 rpm. The HR-N and HR-C peptides were each loaded at initial total peptide concentrations of 50, 100, and 250 μM, and the full-length HR-N1 peptide was loaded at 10, 20, and 50 μM. The partial specific volume of each peptide was calculated using the program SednTerp version 1.08 (50). The density of the solvent was calculated to be 1.009 g/ml. The data were evaluated using a non-linear least squares curve-fitting algorithm contained in the WinNonLin analysis program version 1.06 (51).

Gel Electrophoresis—HR-N and HR-C region peptides singly or as an equimolar mixture (200 μM of each peptide) were dissolved in 100 mM KCl, 50 mM PO_4, pH 7 buffer and incubated at room temperature. Samples were then diluted with 1 volume of 2× Laemmli sample buffer and analyzed by SDS-PAGE in a 15% Tris/glycine gel.

Size-exclusion and Reversed-phase Chromatography—HR-N and HR-C region peptides singly or as an equimolar mixture (200 μM of each peptide) were dissolved in 100 mM KCl, 50 mM PO_4, pH 7, and equilibrated at room temperature for 30 min. A 10-μl aliquot of the mixture was loaded onto a high performance size-exclusion column, Superdex 75™ (1 × 30 cm, Amersham Biosciences) equilibrated in a buffer consisting of 50 mM PO_4, 100 mM KCl, pH 7, at a flow rate of 0.75 ml/min and ambient temperature. For peptides that formed stable heterostranded complexes, the complex peak was collected and analyzed by reversed phase chromatography on an analytical C8 column (Zorbax 300SB-C8, 15 cm × 4.6 mm inner diameter, 6.5 μM particle size, 300-Å pore size; Agilent Technologies). The peptides were eluted from the column by employing a linear AB gradient of 2% B/min, where eluent A is 0.05% aqueous trifluoroacetic acid and eluent B is 0.05% trifluoroacetic acid in acetonitrile at a flow rate of 1.0 ml/min at room temperature. To calculate the peptide ratio in the complex, the peak areas of each component were compared with peak areas of known standard solutions of each peptide.

Formation of Parallel and Anti-parallel Disulfide-bridged Hetero Two-stranded Molecules—Preferential disulfide bridge formation between the HR-N9 and HR-C1 regions was performed similar to that described by Semchuck *et al.* (52). In brief, 10 mg of 2,2′-dithiopyridine was dissolved in 100 μl of dimethylformamide with sonication. A 10-μl (3.4 μmol) aliquot of this solution was added to a solution of the HR-N9 peptide (2 mg, 0.8 μmol) dissolved in a 3:1 (v/v) acetic acid/H_2O buffer and stirred for 6 h. The solution was then extracted three times with 500 μl of ether, and the aqueous layer was applied to a Sephadex G-25 desalting column conditioned with 50 mM NH_4 acetate, pH 5.5 running buffer. Fractions (1 ml) eluted from the column were collected, and those with 220 nm absorbance were pooled. To one-half of the pooled solution was added HR-C1 peptide with either an N-terminal CGG or a C-terminal GGC linker (1 mg of dissolved in 1 ml of 50 mM NH_4Ac, pH 5.5 buffer) in 100-μl aliquots over 30 min. The reactions were then stirred for 1 h, and the final complex was purified by reversed-phase HPLC and freeze-dried.

RESULTS

Sequence Analysis—To identify the heptad repeat (HR) regions within the SARS-CoV S protein, we utilized the coiled-coil prediction algorithm STABLECOIL (53). Fig. 2*A* shows the graphical output of the analysis using a 35-residue window width. The program predicts two coiled-coil regions: the HR-N coiled coil (residues 882–1011) and the HR-C coiled coil (residues 1147–1185) with an interhelical domain of ~140 amino acid residues between HR-N and HR-C. The location of these regions appears to be consistent with other type 1 viral fusion proteins, in that the HR-C region is located adjacent to the predicted transmembrane region (residues 1194–1226), and the HR-N region is located C-terminal to a region of high sequence similarity to the fusion peptide (residues 851–882) (Fig. 2*B*). The coiled-coil prediction analysis also indicates that the HR-N region can be further sub-divided into three regions (regions 1–3) based on the assignment of the heptad a and d positions (data not shown). The region where the heptad registers change is denoted as a frameshift/hinge region (Fig. 2*B*). The program also predicts a second (lower scoring) heptad repeat register within HR-N region 2 that can run continuous with the heptad register observed in HR-N region 3 (shown below the sequence in Fig. 3, *top*). The presence of this alternate heptad register indicates that hydrophobic residues also occur in the interfacial e and g positions in this location. In contrast, the HR-C coiled-coil region shows only a single continuous heptad register (Fig. 3, *bottom*).

Circular Dichroism Analysis of the HR-N and HR-C Regions—To examine the structural characteristics of the HR-N

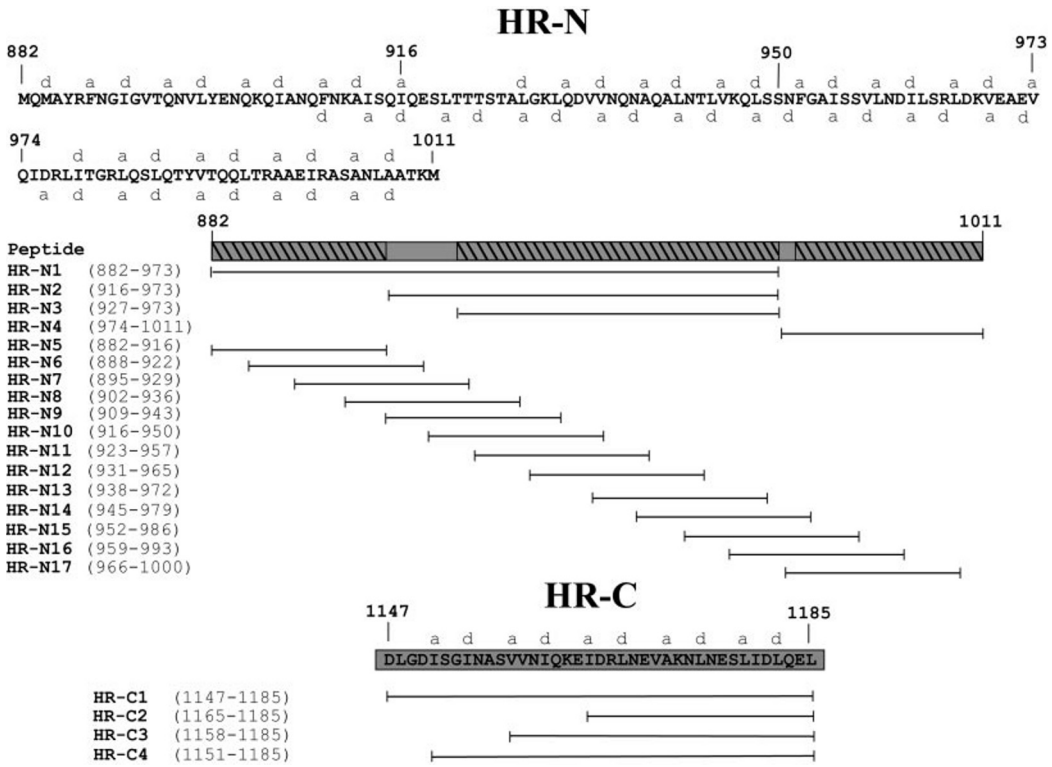

FIG. 3. *Top*, sequence of the predicted HR-N region from residues 882–1011 of the SARS-CoV S protein. The a and d positions of the strongest predicted coiled-coil heptad repeats $(abcdefg)_n$ are shown *above* the sequence. The a and d positions of the alternate weaker scoring heptad repeat is shown *below* the sequence. *Middle*, the names and sequence regions (denoted in *parentheses*) of the HR-N peptides used in this study. The *rectangle* denotes the HR-N sequence from residues 882 to 1011. The *hatched areas* denote regions 1–3 (Fig. 2). The *bars below* the HR-N domain indicate the locations of the peptides within the HR-N region used in this study. *Bottom*, the sequence of the predicted HR-C domain from residues 1147–1185. The a and d positions are shown *above* the sequence. The names, sequences, and locations (*bars*) of the HR-C peptides used in this study are shown *below* the HR-C domain.

and HR-C regions of the SARS CoV S protein, peptides (see Fig. 3) corresponding to various regions between residues 882 and 1185 were synthesized and analyzed by CD spectroscopy. Fig. 4 (*A* and *B*) shows representative far-UV CD spectra and thermal denaturation profiles of the HR-N1, HR-N3, HR-N4, HR-N5, and HR-C1 peptides. Analysis of the HR-N1 peptide, which represents almost the entire length of the predicted HR-N region, showed limited solubility in 0.1 M KCl, 0.05 M K_2HPO_4, pH 7 buffer; thus the CD study could only be carried out at 15 μM. The CD profile of HR-N1 showed the characteristic α-helical spectrum with double minima at 208 and 222 nm. The molar ellipticity at 222 nm (−15,100°) corresponds to 40 helical residues out of the maximum 62 helical residues induced in 50% TFE for the molecule under the conditions analyzed. Temperature denaturation of HR-N1 showed the α-helical structure had high stability with 50% of its structure unfolded at 77 °C (Table I).

Mapping of the HR-N region further with smaller peptides showed that HR-N5 and HR-N4 (regions 1 and 3) were also largely insoluble, and the limited amount that did dissolve (15 μM) exhibited random coil spectra (Fig. 4*A*). The peptides, however, could be induced into an α-helical conformation in the presence of 50% TFE, a helix-inducing solvent, indicating they do have an intrinsic ability to be helical (Table I). In contrast, the HR-N3 peptide corresponding to the central region of HR-N (*region 2*, Fig. 2) showed good solubility characteristics (>5 mg/ml) and exhibited a strong α-helical spectrum (Fig. 4*A*). The mean residue molar ellipticity value (−24,600°) indicated that HR-N3 was fully helical relative to the maximum inducible α-helix in TFE but only 68% helical relative to the theoretical value for a fully folded helical molecule of its 47-residue length (−36,000°). The thermal melting profile of this peptide showed that it had high stability similar to that of the longer length HR-N1 peptide (compare $T_{1/2}$ values of 63 *versus* 77 °C for HR-N3 and HR-N1, respectively). An extension of HR-N3 produced a peptide HR-N2 of 58 residues. This peptide had a similar number of helical residues (Table I) and a $T_{1/2}$ of 68 °C.

We analyzed the CD spectra of several overlapping 35-residue peptides shifted 1 heptad (7 residues) toward the C terminus of the HR-N region beginning at HR-N5 and ending with HR-N17 (Fig. 3). All of these peptides showed mainly random coil spectra with the exception of HR-N12, HR-N13, HR-N14, and HR-N15 that displayed helical spectra profiles. Most interesting, these peptides corresponded to the same 47 residues of the sequence of HR-N3 that was also observed to be helical (Fig. 4 and Table I). It is also notable that the region in HR-N that was predicted to have the highest likelihood of forming a coiled-coil structure (Fig. 2*A*) corresponds with the HR-N3 sequence region identified to form a stable α-helical structure. Most interesting, the stability of peptides (as measured by their $T_{1/2}$ values) for HR-N12 ($T_{1/2}$ 52 °C), HR-N3 ($T_{1/2}$ 63 °C), HR-N2 ($T_{1/2}$ 68 °C), and HR-N1 ($T_{1/2}$ 77 °C) increased with increasing chain length from 35, 47, 58, and 92 residues, respectively (Table I). The number of helical residues in benign medium for these peptides also increased with chain length from 17 to 40 residues.

The far-UV CD spectrum of HR-C1 exhibited double minima at 208 and 222 nm (Fig. 4*A*) indicating α-helical structure. The mean residue molar ellipticity was −12,600° at 100 μM and increased to −22,000° at high concentrations (>500 μM) similar to the value observed in the presence of TFE. Temperature denaturation of HR-C1 showed the peptide has low stability as 50% of the molecule could be unfolded at 33 °C (Fig. 4*B* and Table I).

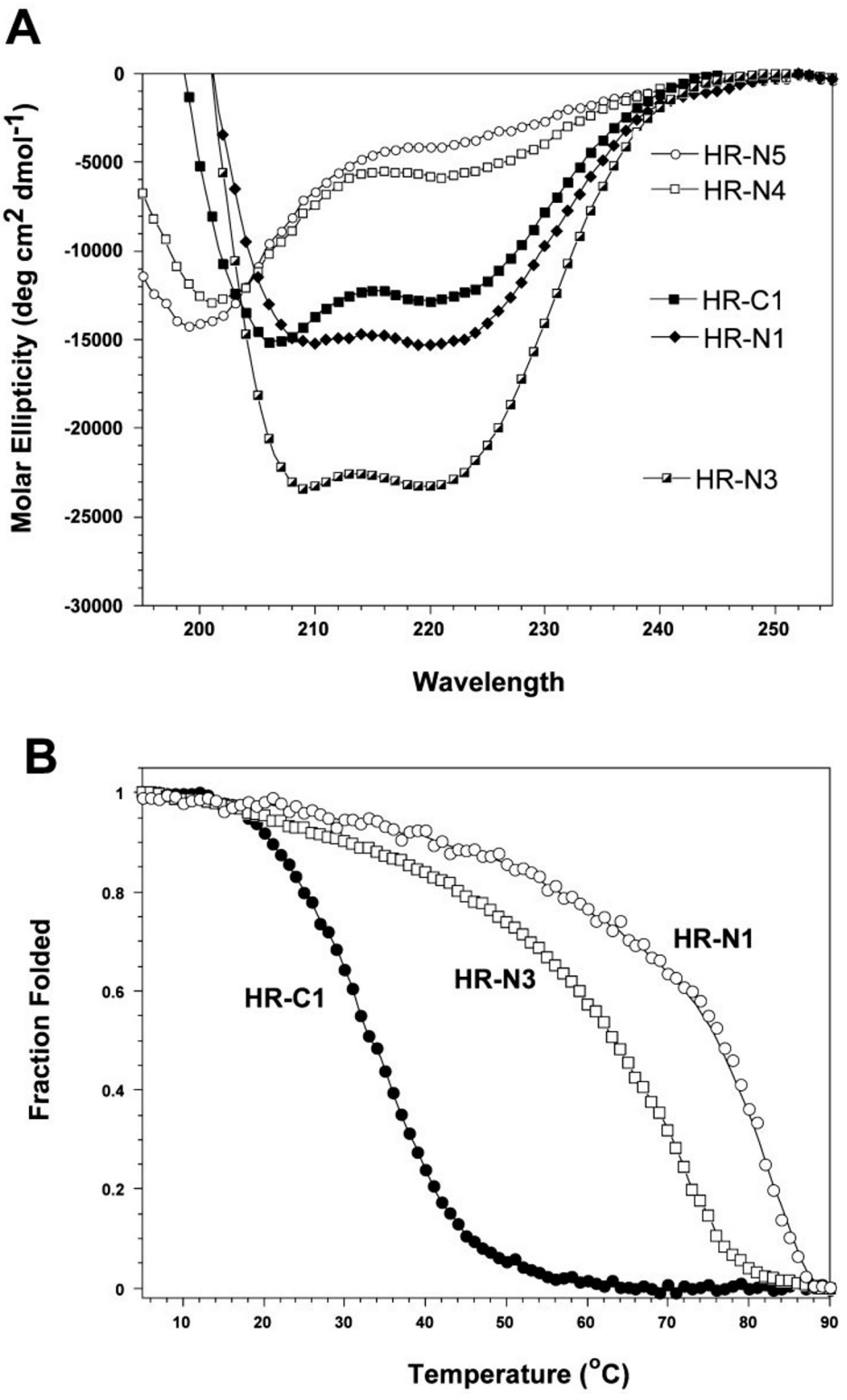

FIG. 4. *A*, CD spectra of peptides corresponding to the HR-N and HR-C of the SARS-CoV S protein. Spectra were recorded in a 0.1 M KCl, 0.05 M PO_4, pH 7 buffer. Peptide concentrations were 100 μM for HR-N3 and HR-C1, and 15 μM for HR-N1, HR-N4, and HR-N5. *B*, temperature denaturation profiles of the HR-N and HR-C peptides monitored by CD at 222 nm in a 0.1 M KCl, 0.05 M PO_4, pH 7 buffer. Concentrations were 15 μM for HR-N1 and 100 μM for HR-N3 and HR-C1 peptides. The fraction folded (f_f) of each peptide was calculated as $f_f = ([\theta] - [\theta]_u)/([\theta]_n - [\theta]_u)$, where $[\theta]$ is the observed mean residue ellipticity at 222 nm at any particular temperature, and $[\theta]_n$ and $[\theta]_u$ are the mean residue ellipticities at 222 nm of the native folded state at 4 °C and unfolded states, respectively.

To see if we could define further the α-helical region within HR-C (39 residues), we analyzed three truncated peptides of HR-C1. CD analysis of HR-C2 and HR-C3 (21 and 28 residue length, respectively) displayed random coil spectra in benign buffer, whereas HR-C4 (35-residue length) showed α-helical structure with a $[\theta]_{222}$ value of −17,500° (Table I). Additionally, the thermal melting profile of HR-C4 indicated the same temperature transition midpoint (33 °C) as full-length HR-C1 (Table I). Thus the minimal length required for folding of the HR-C region coiled coil appears to be 35 residues corresponding to the sequence 1151–1185.

Oligomeric State of the HR Regions—Sedimentation equilibrium experiments were carried out to determine the oligomeric states of the HR-N and HR-C peptides. The HR-N1 peptide was studied at three different concentrations and three different rotor speeds. The data obtained fit well to a single species model with a weight average molecular weight of 48,750 (Fig. 5, *left*). Surprisingly, this value corresponds well to that expected for a tetramer of HR-N1, based on the amino acid sequence ($4 \times 12{,}004 = 48{,}016$ Da). The high quality of the single species fit at the different concentrations analyzed suggests that the HR-N1 peptide was exclusively tetrameric over the entire con-

TABLE I
Ellipticity and stability of the synthetic peptides studied

Peptide name[a]	No. residues	$[\theta]_{222}$[b]		Helical residues[c]		% Helix (benign/TFE)[d]	$T_{½}$[e]
		Benign	50% TFE	Benign	50% TFE		
		degrees·cm²·dmol⁻¹					°C
HR-N1[g]	**92**	**−15,100**	**−25,800**	**40**	**62**	**65**	**77**
HR-N2	**58**	**−21,100**	**−23,300**	**33**	**37**	**89**	**68**
HR-N3	**47**	**−24,600**	**−21,500**	**32**	**28**	**100**	**63**
HR-N4[h]	38	−6,100	−31,300		34		
HR-N5[h]	35	−4,300	−24,200		24		
HR-N6	35	−3,300	−20,100		20		
HR-N7	35	−2,800	−24,200		24		
HR-N8	35	−2,500	−25,100		25		
HR-N9	35	−2,800	−26,800		27		
HR-N10	35	−3,600	−28,500		29		
HR-N11	35	−3,200	−25,200		25		
HR-N12	**35**	**−16,500**	**−21,000**	**17**	**21**	**81**	**52**
HR-N13	35	−8,700[f]	−20,500		21		
HR-N14	35	−7,600[f]	−12,000		12		
HR-N15	35	−6,200[f]	−14,900		15		
HR-N16	35	−4,800	−19,000		19		
HR-N17	35	−5,500	−23,300		24		
HR-C1	**39**	**−13,600**	**−23,800**	**15**	**26**	**58**	**33**
HR-C2	21	−3,600	−15,400		10		
HR-C3	28	−5,800	−19,200		16		
HR-C4	**35**	**−17,500**	**−19,800**	**18**	**20**	**90**	**33**

[a] Name of each peptide studied. The sequence position of each peptide is shown in Fig. 3.
[b] The mean residue molar ellipticities at 222 nm were measured at 22ºC in benign buffer (0.1 M KCl, 0.05 M K_2PO_4, pH 7). For samples containing TFE, the above buffer was diluted 1:1 (v/v) with TFE.
[c] The number of α-helical residues was calculated based on the predicted molar ellipticity. The predicted molar ellipticity for a completely α-helical peptide was calculated using $[\theta]_{\text{theoretical}} = \times 40{,}000\ (1 - 4.6/N)$, where N is the number of residues in the polypeptide chain. The number of α-helical residues in benign or TFE conditions was calculated as the observed $[\theta]$ divided by the theoretical $[\theta]$.
[d] The percent α-helix in benign medium at the concentration used was calculated as the number of helical residues in benign divided by the maximum inducible α-helical residues in TFE. Note that in oligomerizing α-helices the helical content is concentration-dependent. Higher concentrations shift the monomer to oligomer equilibrium increasing the α-helical content based on a two-state unfolding mechanism (random coil to coiled coil).
[e] $T_{½}$ is the transition midpoint temperature at which there is a 50% decrease in molar ellipticity $[\theta]_{222}$ compared with the fully folded peptide as determined by CD at 5 °C. All peptides in which $T_{½}$ was determined are in boldface type.
[f] These peptides all showed a characteristic α-helical spectrum with minima at 208 and 222 nm even though the $[\theta]_{222}$ was low.
[g] Results obtained at a peptide concentration of 15 μM instead of 100 μM used for the other determinations.
[h] Peptides were studied in 0.05 M K_2PO_4 buffer, pH 7, due to low solubility.

centration range of the experiment. This result was surprising considering earlier analyses of fusion proteins and ectodomain cores of other viruses have been observed to be trimeric (54–59). Our results, however, were consistent with the sedimentation equilibrium analysis of the smaller truncated HR-N peptide, HR-N3, which also displayed α-helical structure. Sedimentation equilibrium data obtained for this peptide best fit to a single species model with weight average molecular weight of 19,000, close to that expected for a tetramer (theoretical molecular weight 20,212, 4 × 5,053, data not shown).

In contrast, sedimentation equilibrium experiments of the HR-C1 peptide at three different concentrations and three speeds showed that the data were best fit globally to an associating monomer to trimer equilibrium model. Fig. 5 (*right*) shows a plot of the absorbance *versus* radial distance squared divided by 2 and the residuals of the fit for the HR-C1 peptide. At the highest concentration (250 μM) and highest speed (30,000 rpm), the HR-C1 peptide showed a weight average molecular weight of 13,900, slightly greater than the theoretical mass for a trimer (13,007 Da). The deviations from the expected trimeric species near the bottom of the cell are likely due to aggregation at this high concentration (~1 mM).

Interaction between the HR-N and HR-C—Previously, it has been shown that HR-N and HR-C of the murine coronavirus MHV S glycoprotein self-associate to form a 6-helix bundle structure typical of other type 1 viral fusion proteins (19). To identify the corresponding sequences within the SARS-CoV S glycoprotein, mixtures of the HR-N and HR-C peptides were screened by CD spectroscopy and SEC. Initially, mixtures between the three HR-N peptides, HR-N5, HR-N3, and HR-N4 (which correspond to regions 1–3 of HR-N, Fig. 2), with HR-C1 were screened. These peptide mixtures demonstrated no detectable interaction, *i.e.* the CD spectra were super-imposable with the theoretical spectra calculated for two non-interacting peptides. Additionally, the SEC profiles were similar to those of the individual peptides alone. We next analyzed the HR-N peptides corresponding to the intervening frameshift/hinge regions (Fig. 2). No interactions were observed for peptides HR-N6, HR-N7, and HR-N8. In contrast, HR-N9 showed the induction of 3,500° at 222 nm; HR-N10 showed the induction of 9,600°, and shifting 7 residues (1 heptad) further toward the C terminus (HR-N11) resulted in the absence of any observable interaction once again with HR-C1 (Table II and Fig. 6). Similar mapping results were also observed in the temperature denaturation profiles and SEC analysis of the mixtures. The mixture between HR-N9/HR-C1 showed a $T_{½}$ of 36 °C, a 3 °C rise relative to HR-C1 alone ($T_{½}$ 33 °C), and the SEC showed the formation of a higher molecular weight complex of ~30% of the observed area units. HR-N10/HR-C1 ($T_{½}$ 57 °C) showed a 24 °C increase in $T_{½}$ from HR-C1 (Fig. 6*C*), and the SEC showed largely one single peak of higher molecular weight (Fig. 7). The mixture between HR-N11/HR-C1 showed no complex peak in the SEC. It is worthy of mention that the CD signal at 222 nm for the strongest interacting peptides (HR-N10/HR-C1) only showed a molar ellipticity of −17,700, ~51% of a theoretical fully helical peptide of 35 residues in length despite the significant increase in stability observed (data not shown).

To confirm that the structure of the HR-N region is not a factor in binding to HR-C1, as HR-N10 was a random coil on its own (Fig. 6*A*), we extended the length of HR-N3 (residues 927–973) that exhibited an α-helical structure by 11 additional N-terminal residues (HR-N2) to contain the complete binding

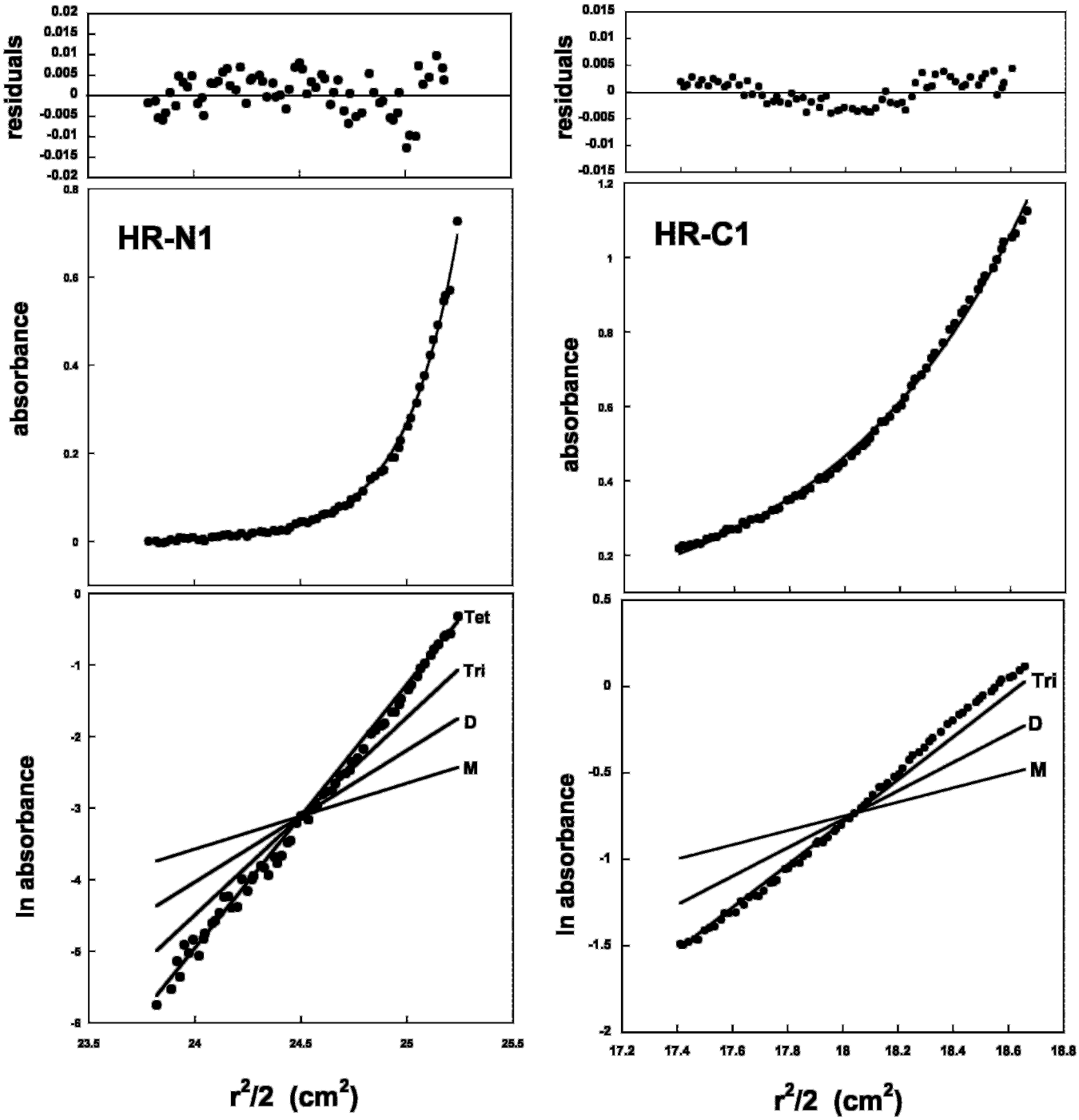

FIG. 5. **Equilibrium ultracentrifugation analysis of the HR-N and HR-C peptides.** *Left side*, *middle*, a plot absorbance *versus* radial distance squared divided by 2 for HR-N1 at 22 °C and 26,000 rpm in a 0.1 M KCl, 0.05 M PO_4, pH 6 buffer. Protein concentration was 15 μM. *Left side*, *bottom*, a plot of the natural logarithm of the absorbance *versus* radial distance squared divided by 2 for the same data. The theoretical lines for a single species monomer (*M*), dimer (*D*), trimer (*Tri*), and tetramer (*Tet*) are shown for comparison. The data best fit a single species model with tetrameric molecular weight. *Right side*, *middle*, a plot of absorbance *versus* distance squared divided by 2 of the HR-C1 peptide at 22 °C and 30,000 rpm in a 0.1 M KCl, 0.05 M PO_4, pH 7 buffer. Peptide concentration was 250 μM. *Right side*, *bottom*, a plot of the natural logarithm of the absorbance *versus* radial distance squared divided by 2 for the same data. The theoretical lines for monomer, dimer, and trimer are shown. The data were fit best to a monomer to trimer associating model in equilibrium. Residuals from both fits are shown above.

site of HR-N10. This peptide showed strong complex interactions with HR-C1 as assessed by SEC. Furthermore, the temperature denaturation profile indicated the midpoint rose 6 °C from 68 °C for HR-N2 alone to 74 °C for the complex HR-N2/HR-C1 (Table I and Table II). The denaturation profile was cooperative for this complex *versus* a biphasic transition profile observed for HR-N3 with HR-C1 which showed $T_{½}$ values corresponding to the individual domain $T_{½}$ values indicative of non-interacting helices. Thus HR-C1 forms a complex with the HR-N region whether the N-region is unfolded (HR-N10) or folded (HR-N2).

We also analyzed the truncated peptides of HR-C1 in an attempt to localize the binding site for HR-N10. HR-C2 and HR-C3 peptides of 21 and 28 residue length showed no ability to interact the HR-N10 region as assessed by CD spectroscopy and SEC (Table II). In contrast, the HR-C4 peptide of 35-residues which folded independently into an α-helical structure showed strong complex formation with the HR-N10 peptide

TABLE II
CD and SEC analysis of the HR-N/HR-C peptide mixtures

Peptide mixture[a]	$[\theta]_{222}$[b]		Change in molar ellipticity[c]	$T_{½}$[d]	SEC[e]
	Observed	Predicted			
	degrees·cm²·dmol⁻¹			*°C*	
HR-N2/HR-C1	−17,900	−17,000	900	74	++
HR-N3/HR-C1	−18,000	−18,000	0		−
HR-N4/HR-C1[f]	−9,400	−9,400	0		−
HR-N5/HR-C1[f]	−8,500	−8,500	0		−
HR-N6/HR-C1	−8,600	−8,600	0		−
HR-N7/HR-C1	−7,400	−7,400	0		−
HR-N8/HR-C1	−7,900	−7,900	0		−
HR-N9/HR-C1	−10,800	−7,300	3,500	36	+
HR-N10/HR-C1	−17,700	−8,100	9,600	57	++
HR-N11/HR-C1	−8,000	−8,000	0		−
HR-N12/HR-C1	−14,900	−14,900	0		−
HR-N13/HR-C1	−10,300	−10,300	0		−
HR-N14/HR-C1	−12,500	−12,500	0		−
HR-N15/HR-C1	−9,500	−9,500	0		−
HR-N16/HR-C1	−9,000	−9,000	0		−
HR-N17/HR-C1	−9,500	−9,500	0		−
HR-N10/HR-C2	−3,800	−3,800	0		−
HR-N10/HR-C3	−4,900	−4,900	0		−
HR-N10/HR-C4	−7,900	−10,500	7,400	56	++
HR-N10/HR-C1(antiparallel)ox	−9,200			76	
HR-N10/HR-C1(antiparallel)ox	−3,300[g]			74[g]	
HR-N10/HR-C1(parallel)ox	−12,200[g]			57[g]	

[a] Name of the peptides in the mixture. For sequence regions see Fig. 3. Names followed by "ox" indicate peptide mixtures that have been covalently linked by a disulfide bridge.
[b] The mean residue molar ellipticities at 222 nm were measured at 22 °C in benign buffer (0.1 M KCl, 0.05 M K_2PO_4, pH 7).
[c] The difference in molar ellipticity from the observed and predicted values (column 2 and column 3).
[d] $T_{½}$ is the transition midpoint temperature at which there is a 50% decrease in molar ellipticity $[\theta]_{222}$ compared with the fully folded peptide as determined by CD at 5 °C. When no change in molar ellipticity (column 3) was observed, the $t_{½}$ was not determined.
[e] Summary of the analysis of the peptide complexes by SEC. − denotes no complex formation; + denotes weak complex formation between the two peptides (~30% of total peak area or less); ++ denotes strong complex formation between the two peptides (>95% of total peak area).
[f] Results were obtained for peptide complexes scanned in a 0.05 M K_2PO_4 buffer.
[g] Peptide complex concentration was 100 and 15 μM. The predicted molar ellipticity for two non-interacting peptides was calculated by summing the two individual spectra.

observed by CD and SEC (Table II). For example, the temperature denaturation profile of HR-N10/HR-C4 displayed a similar temperature midpoint and molar ellipticity observed for HR-N10/HR-C1 (compare 57 °C with 56 °C and −17,700 with −17,900, respectively). Thus the binding interaction between the HR-N and HR-C regions appears to be localized to residues 916–950 in the center of HR-N and 1151–1185 within the HR-C region (based on the N-terminal truncations).

HR-N and HR-C Form a 3:3 Hexameric Complex—To determine the ratio of the peptides within the HR-N10/HR-C1 complex, mixtures of HR-N10 and HR-C1 peptides at molar ratios of 1.5:1, 1:1, and 1:1.5, respectively, were pre-incubated for 30 min and then applied to a size-exclusion column (Fig. 7). The earlier eluting complex was collected and applied to a reversed-phase C8 analytical HPLC column and eluted (Fig. 7, *inset*). Both HR-N10 and HR-C1 peptides were observed. Integration of the peak areas for each component and conversion of these values to mole amounts showed that the two peptides within the complex were present at a 1:1 molar ratio.

Next, we analyzed the molecular mass of the complex. Bosch *et al.* (19) had shown for MHV-A59 that complexes of HR-1 and HR-2 are stable during SDS-PAGE as long as the samples are not heated before loading. Therefore, HR-C1, HR-N10, or HR-N2 peptides separately or as pre-incubated equimolar mixtures were subject to Tris/SDS-15% PAGE (Fig. 8). The HR-N10 peptide, which CD spectroscopy had shown to be a random coil, migrated in the gel to a location corresponding to its molecular weight (~3800 Da). In contrast, the HR-C1 peptide and HR-N10/HR-C1 complex migrated with higher molecular masses. The mobility of the HR-C1 peptide ran according to a molecular mass of ~13 kDa (just below the 16.8-kDa marker) indicating a trimeric state (theoretical mass of 13,008 Da), and the HR-N10/HR-C1 complex migrated according to a molecular mass of ~24 kDa corresponding to a hexameric (6-stranded) complex (theoretical mass of 24,279 Da) (Fig. 8*A*).

Similar results were also seen for the HR-N2/HR-C1 peptides (Fig. 8*B*). The HR-N2 peptide migrated as a broad band between ~24 and ~17 kDa with the most pronounced band at ~17 kDa, indicating a molecular mass between tetramer and trimer (theoretical masses 24,512 and 18,384 Da, respectively), and the HR-C1 peptide migrated according to a mass of ~13 kDa (close to a trimeric mass), and the HR-N2/HR-C1 complex migrated according to a molecular mass of 29 kDa, close to the mass of a 3:3 hexamer (theoretical mass 31,692 Da).

Sedimentation equilibrium analysis was carried out on the more stable HR-N2/HR-C1 complex. As shown in Fig. 9, the data were fit best to a single species model with molecular mass of 31,500 Da, close to the theoretical mass of a 3:3 mole ratio (6-stranded) complex between the HR-N2 and HR-C1 peptides (theoretical mass 31,692 Da).

Antiparallel Helix Orientation—To investigate the orientation of the helices within the HR-N10/HR-C1 complex, we chose to form disulfide-bridged heterodimers between HR-N10 and HR-C1 in the parallel and antiparallel polypeptide chain orientation. The premise of this experiment was that the correct polypeptide chain orientation should form a tightly folded helical structure of high solubility and greater stability due to elimination of the concentration dependence between the two peptides, whereas the incorrect orientation would form head-to-tail complexes, which would aggregate and show little or no increase in stability (Fig. 10*A*). To do this, we added Cys-Gly-Gly residues to the N terminus of HR-N10 and either the N or C terminus of HR-C1. The peptides were then selectively covalently linked using the thiol coupling agent 2,2′-dithiopyridine (see "Materials and Methods").

Table II and Fig. 10 show helicity and stability of the two

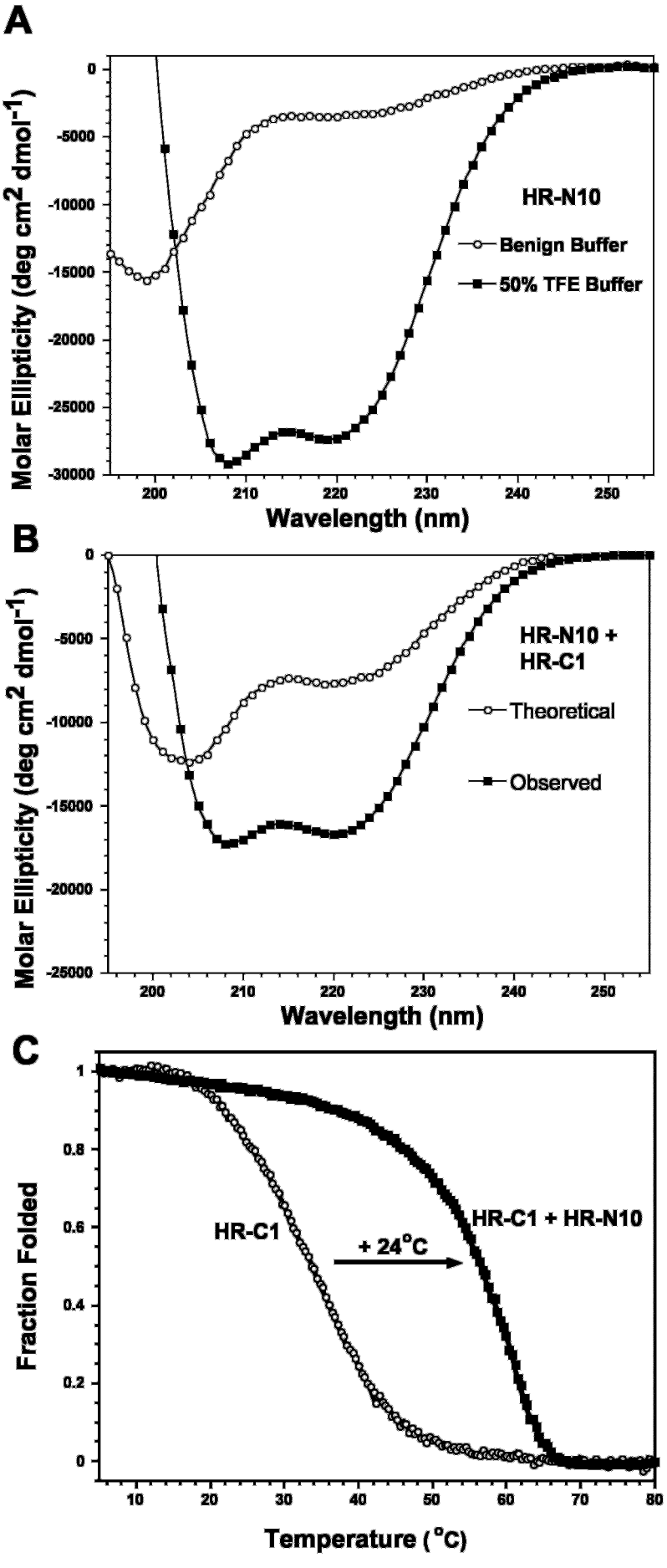

FIG. 6. **Interaction between HR-N and HR-C peptides.** *A*, circular dichroism spectra of HR-N10 peptide alone. Spectra were recorded at 25 °C in a 0.1 M KCl, 0.05 M PO_4, pH 7 buffer. For the spectrum containing 50% TFE, the above buffer was diluted 1:1 (v/v) with TFE. *B*, CD spectrum of a 1:1 molar complex between HR-N10 and HR-C1 peptides at 25 °C in a 0.1 M KCl, 0.05 M PO_4, pH 7 buffer. Peptide concentrations were 80 μM. The theoretical spectrum for two noninteracting peptides is shown for comparison. This spectrum is generated by adding the individual peptide spectra at the same concentrations. *C*, temperature denaturation profiles of HR-C1 alone and a 1:1 molar HR-C1 with HR-N10 complex monitored by CD at 222 nm in a 0.1 M KCl, 0.05 M K_2PO_4, pH 7 buffer. Concentrations were 100 μM. Fraction folded was calculated as described in Fig. 4 legend.

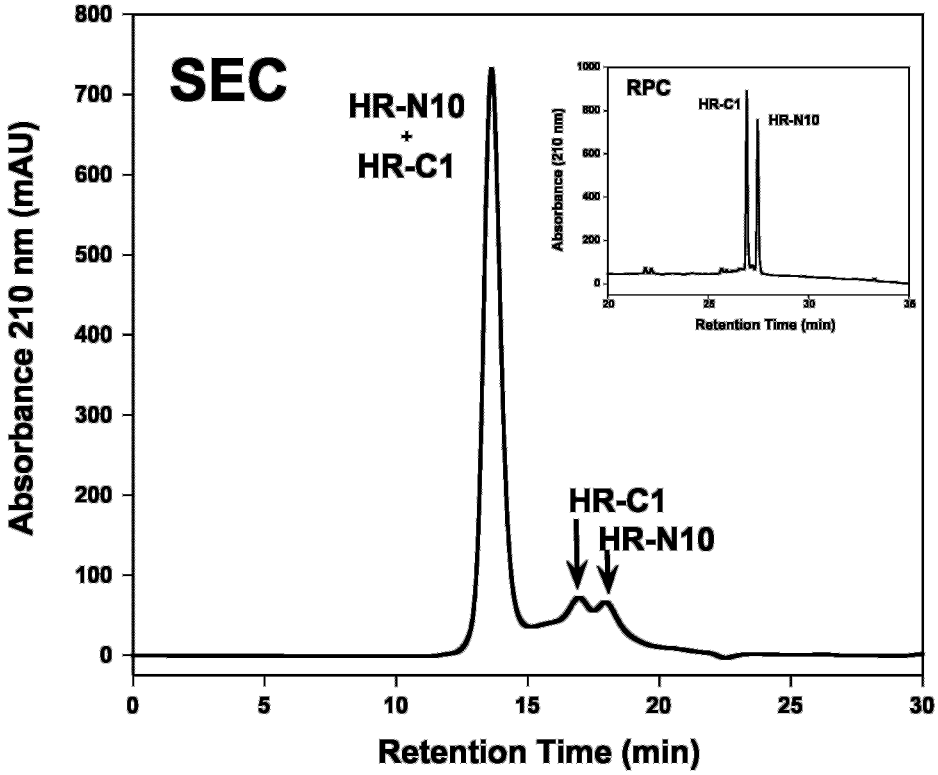

FIG. 7. **HPLC analysis of the HR-N10/HR-C1 complex.** HR-N10 (2 nmol) and HR-C1 (2 nmol) were pre-incubated together for 30 min in 10 μl of running buffer and then applied to a Superdex^{TM} 75 SEC column equilibrated in a buffer consisting of 0.1 M KCl, 0.05 M PO_4, pH 7, and a flow rate of 0.7 ml/min. The absorbance peak at 13 min corresponding to the HR-N10/HR-C1 complex was collected and subsequently analyzed by reversed phase chromatography (*inset*) on an analytical C8 Zorbax column employing a linear AB gradient of 2% B/min, where A = 0.05% aqueous trifluoroacetic acid and B is 0.05% trifluoroacetic acid/acetonitrile. Each absorbance peak is labeled accordingly.

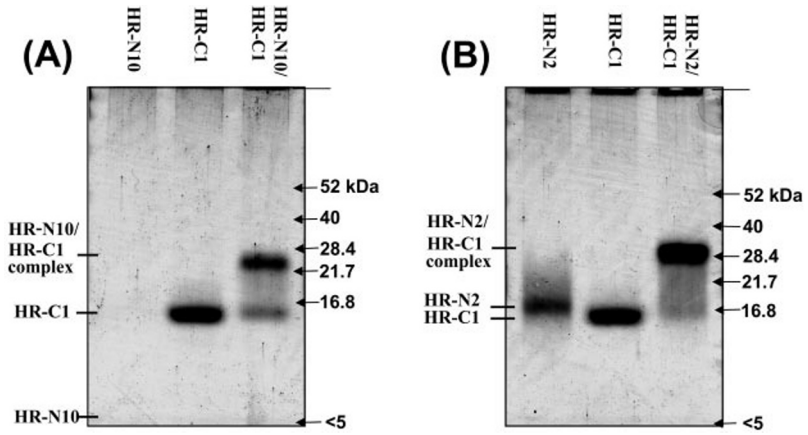

FIG. 8. **Molecular mass of the complex formed between the HR-N and HR-C regions as determined by gel electrophoresis.** *A*, HR-N10 and HR-C1 peptides on their own or as pre-incubated equimolar (200 μM of each peptide) mixture were subjected to Tris/SDS-15% PAGE. Samples were incubated for 30 min in 0.1 M KCl, 0.05 M PO_4, pH 7 buffer, and the diluted 1:1 (v/v) with 2× Laemmli sample buffer at room temperature and loaded into the gel. The positions of HR-N10, HR-C1, and HR-N10/HR-C1 complex are indicated on the *left side* of the gel, whereas the positions of the molecule mass markers are indicated on the *right side* of the gels with *arrows*. Note the absence of the HR-N10 band was a result of the peptide eluting from the gel during the staining/destaining steps. *B*, HR-N2 and HR-C1 peptides treated in a manner similar to that described in *A*. The lanes of the gel, peptide locations, and molecular mass markers are labeled accordingly.

disulfide-bridged complexes. The first complex in which the HR-N10 and HR-C1 peptides were linked in an anti-parallel orientation showed high solubility (>5 mg/ml or 1 mM) and a temperature transition midpoint of 76 °C, an increase of 19 °C relative to the non-disulfide-bridged complex analyzed at the same concentration (HR-N10/HR-C1 complex, $T_{½}$ of 57 °C, Table II). In contrast, the complex between HR-N10 and HR-C1 disulfide-bridged in the parallel orientation was largely insoluble in the same buffer conditions (solubility ~15 μM). The low solubility was completely dependent upon the disulfide bridge link between the peptides, as addition of dithiothreitol to reduce the disulfide bridge completely reversed the solubility characteristics and allowed the complex to be fully soluble at high concentrations once again. A temperature denaturation melting profile was obtained at 15 μM for the parallel complex and was compared with the anti-parallel complex at the same concentration (Fig. 10*B*). The temperature transition midpoint of the parallel complex was 57 °C, similar to that observed for the non-disulfide-bridged complex, indicating the interactions between HR-N10 and HR-C1 are likely inter-molecular, whereas the anti-parallel complex showed a temperature transition midpoint of 74 °C, indicating that it was 17 °C greater than the opposite orientation. It is noteworthy that a CD scan of the anti-parallel complex at 100 μM concentration only showed a molar ellipticity reading of −19,200° at 222 nm (a fully folded α-helix of 35 residues is predicted to have a molar ellipticity of −34,700°). Thus the highly stable complex either has unfolded residues at the termini of the helices or, when in a complex, one or the other helix exists partly in a non-helical form. Taken together, however, the results suggest the HR-N

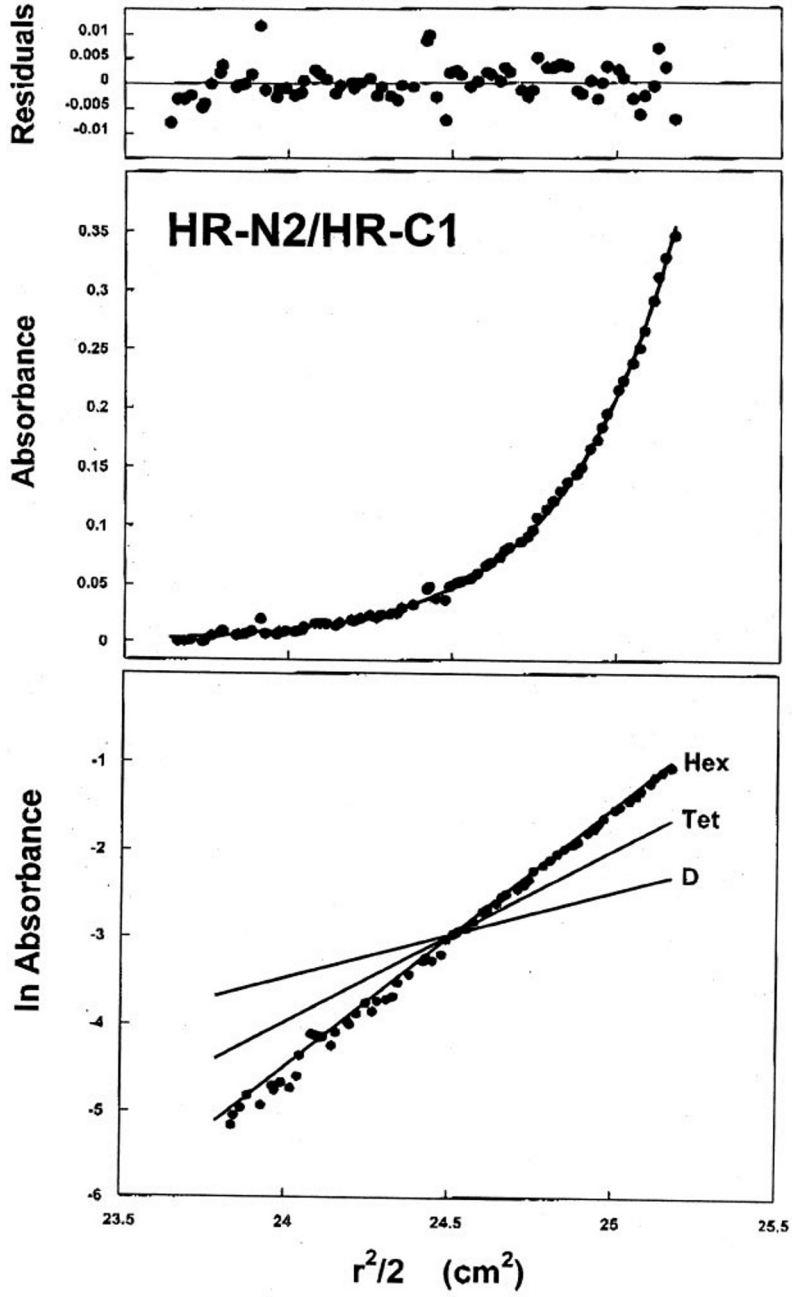

FIG. 9. **Molecular mass of the complex formed between HR-N2 and HR-C1 peptides as determined by sedimentation equilibrium analysis at 22 °C and 30,000 rpm.** The *middle panel* (absorbance *versus* radial distance squared divided by 2) shows the best fit curve for a single species model indicating a molecular mass of 30,500 Da. The theoretical mass for a 3:3 mole ratio HR-N2/HR-C1 hexamer is 31,692 Da. The *upper panel* shows the residuals from the curve fit, and the *lower panel* shows the ln absorbance *versus* $r^2/2$ of the data compared with theoretical 1:1 dimer (*D*), 2:2 tetramer (*Tet*), and 3:3 hexamer (*Hex*).

and HR-C regions associate in an anti-parallel manner in the hexameric complex state.

DISCUSSION

Here we used CD spectroscopy, SEC, and sedimentation equilibrium analysis to characterize the predicted HR-N and HR-C regions in the ectodomain of the spike S glycoprotein of the SARS-CoV. In isolation, the HR-N region folds as an independent α-helical coiled-coil structure of high stability, but with only ~40 helical residues of a possible 92 residues. Analysis of smaller peptides of HR-N showed that the helical content is largely localized to residues 927–973 in the central portion of HR-N and the N and C termini were unfolded. It is interesting that the N-terminal region closest to the predicted fusion peptide sequence (residues 851–882, Fig. 2, a region of high sequence similarity to the fusion peptides of other type 1 fusion proteins, which is thought to insert into the target membrane during the fusion process) did not show any α-helical coiled-coil structure, indicating this region does not spontaneously fold into a coiled coil and therefore requires other conditions or parts of the molecule. This is similar to observations in influenza HA, Ebola virus GP2, human T-cell leukemia virus type 1, gp21, and RSV F fusion proteins. In the case of influenza HA, the N terminus of the HA_2 coiled coil at neutral pH exists in part as an extended conformation but undergoes a conformational change to a more stable elongated coiled coil only when HA encounters a low pH acid environment in a series of steps termed the spring-loaded mechanism (60–63). It should be noted that we did analyze the HR-N1 peptide at lower pH (4.5) and did not observe any significant increase in helicity (data not shown). Moreover, CoV S proteins are generally stable at acid pH, and MHV S can be triggered to its fusogenic state at pH 8, 37 °C (23).

It is noteworthy that studies on the HR regions of the S protein of the mouse hepatitis virus strain A59 showed the α-helical content of the separate HR-N and HR-C peptides were calculated to be ~89%, and that the helical content of their

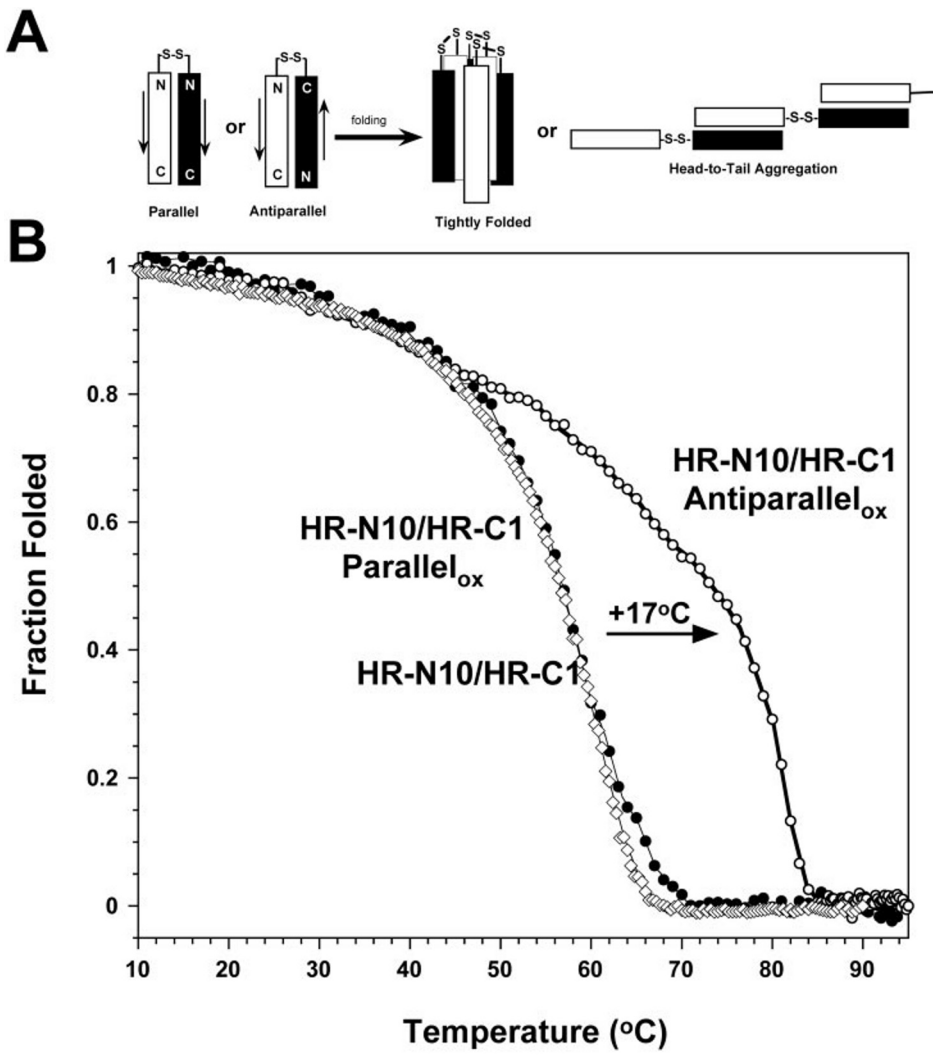

FIG. 10. **Helix orientation in the HR-N/HR-C complex.** *A*, schematic representation of the use of selective disulfide bridge formation to form parallel and anti-parallel hetero-stranded polypeptides, and the effect the relative helix orientations could have on the final folded state of the complex. The HR-N helix is represented as a *white rectangle* and corresponds to residues 916–950. The HR-C helix is represented as a *black rectangle* and corresponds to residues 1147–1185. The disulfide bridge between two cysteine residue side chains is denoted by *S-S*. *B*, temperature denaturation profiles of parallel and anti-parallel disulfide-bridged HR-N10/HR-C1 complexes in 0.1 M KCl, 0.05 M PO_4, pH 7 buffer monitored by CD at 222 nm. Peptide concentrations were 15 μM. Fraction folded was calculated as described in Fig. 4 legend. The temperature melting profile of the non-disulfide-bridged HR-N10/HR-C1 peptide is shown for comparison.

equimolar mixture was calculated to be ~82% (19). Furthermore, electron microscopy showed rod-like structures of ~14.5 nm correlating well with the length predicted for an α-helix the size of HR-N (96 residues). Because the sequences between coiled-coil domains of the SARS-CoV Urbani strain analyzed here and MHV-A59 have a high degree of sequence identity, this difference was quite surprising. One possibility for the difference may be due to the buffer conditions in which the samples were analyzed. The MHV-A59 peptide was analyzed in H_2O *versus* physiological conditions used in this study (100 mM KCl, 50 mM K_2HPO_4, pH 7 buffer). Alternatively, the difference could be due to the low concentration at which we analyzed SARS HR-N1 (15 μM). The α-helical contents of peptides that form coiled-coil structures are affected by concentration due to the monomer-oligomer equilibrium (two-state folding mechanism random coil to coiled coil). It should be noted, however, that we observed a strong α-helical coiled coil in HR-N between residues 927 and 973, and this region correlates well with the protease-resistant portion of the HR-N region determined by Bosch *et al.* (19), and MHV S residues 976–1040 correspond to residues 909–973 of the SARS-CoV S protein sequence.

Sedimentation equilibrium analysis showed that both the long and short folded HR-N region peptides associated as a tetramer; however, in SDS-PAGE the shorter HR-N region (HR-N2) displayed a smear between tetramer and trimer with the more dominant band displaying a trimeric molecular weight. This difference may indicate that the HR-N region naturally associates as a trimer, but in the absence of the HR-C region (or other parts of the native sequence) the hydrophobic residues occurring in the e and g positions are exposed. To sequester these residues from the aqueous environments, the protein switches from a trimer to a tetramer state. In agreement with this proposal is the observation that more interfacial surface areas, in particular the e and g positions, are buried in a tetrameric coiled-coil structure relative to a trimeric or dimeric coiled coil (64).

The HR-C region in isolation showed a very concentration-dependent α-helical coiled-coil structure that at a very high concentration (>500 μM) indicated only ~24 helical residues. The truncation studies showed that the N-terminal four residues could be deleted without adversely affecting the structure. Additionally both sedimentation equilibrium and gel electrophoresis showed that the entire HR-C region alone associates as a homotrimer. Consistent with the helicity data, stability analysis showed that HR-C1 was of low stability with 50% of its structure unfolded at 33 °C. This and the above characteristic are surprising, considering that 9 out of the 10 heptad a and d core positions contain large optimal hydrophobic residues (*e.g.* Ile, Ile, Val, Ile, Ile, Leu, Leu, Leu, and Leu), and no apparent intermolecular electrostatic repulsions can be predicted. The low stability, however, does fit with the general mechanism of type 1 viral fusion proteins, where the C-terminal HR must separate in order to form a complex with the N-terminal HR. In agreement with this is the observation that there are no possible inter-chain electrostatic attractions that can occur between peptide chains. Such electrostatic interactions have been

shown to add considerable stability to coiled coils and stabilize their ability to stay associated (65).

Upon mixing the HR-C1 peptide with several truncated HR-N region peptides, HR-C1 associated preferentially with HR-N10 to form a stable α-helical 6-stranded complex, consisting of a 3:3 mole ratio of HR-N10:HR-C1. Additionally, the selective disulfide bridge experiments indicated that the relative orientation of HR-N10 and HR-C1 peptides in the complex were anti-parallel relative to one another. This orientation is consistent with that observed for other type 1 fusion-active core complexes (25–36). Furthermore, the peptide mapping study also revealed that a 35-residue length is optimal for the binding interaction with HR-C1, as shifting either 7 residues toward the N or C terminus of HR-N (HR-N9 or HR-N11) caused significant loss in the ability to interact with HR-C1. In general, mapping of the HR-N10 region as the primary site of interaction with HR-C1 appears to be entirely consistent with the location recently identified in the MHV-A59 coronavirus strain. Bosch *et al.* (19) identified MHV S residues ~969–1014 as the site of interaction with HR-C. The SARS-CoV HR-10 peptide corresponds to residues 984–1013 of MHV by sequence alignment. Moreover, the strong conservation in location of the HR-C-binding site between the two coronaviruses, despite slight sequence differences in HR-C, likely reveals the importance of this site for the coronavirus fusion mechanism.

Most interesting, despite CD analysis of the HR-N10/HR-C1 complex showing a significant increase in thermal stability and α-helical induction, the complex did not adopt a fully helical structure ($-17{,}700°$ molar ellipticity observed). Some loss can always be attributed to end fraying and concentration dependence (66–68). However, when HR-N10 and HR-C1 were disulfide-bridged, removing the concentration dependence between these two peptides, the molecule only displayed $-19{,}200°$ molar ellipticity. One possible explanation may be that the interaction between HR-C and HR-N10 is not fully helical but more akin to the binding interactions observed for HR-C to HR-N of the F protein of the paramyoxovirus SV5 (25). The structure of the fusion-active core region of the SV5 F protein shows that the HR-C (HR2) domain contains both an α-helix and an extended conformation that stretches out along the HR-N (HR2) region. Whether this is also the case for S of SARS-CoV or not will be best resolved with subsequent high resolution structural studies.

It is generally accepted that viral fusion proteins can adopt at least two different conformations (*i.e.* native and fusogenic states) and that the change in conformation is required for fusion of the viral and cellular membranes during virus entry (24). For the HA protein of influenza virus, HA is most stable in its fusogenic state, whereas HA in its native state is metastable, and formation of the more stable helical hairpin complex provides the energy to overcome the activation energy needed to bring the two lipid bilayers close together at the fusion site (60, 62, 63). The crystal structure of a large number of other viral fusion proteins in the more stable fusogenic state (25–36) confirms many aspects of this proposed model. Because the α-helical heterotrimeric (6-helix) complex consisting of HR-N10 and HR-C1 peptides of SARS-CoV S protein displays greater stability than the separate domains, we speculate that this 6-helix bundle structure represents the core of the fusion-competent state of the SARS-CoV S protein.

Finally, the observation of the 6-helix bundle conformation suggests that peptides corresponding to the SARS-CoV S protein HR regions should efficiently inhibit viral fusion as observed with other type 1 viral fusion proteins (19, 37–46). In these cases, free HR-N or HR-C peptides compete for binding to the HR-N and HR-C coiled coils in the native state thereby blocking the conformational transition to the fusion-active form. Analysis of the inhibitory effects of the SARS-CoV S peptides is now in progress.

Taken together, these studies of the SARS-CoV HR regions support the emerging structural theme for type 1 viral fusion proteins. In the native state, the two HR regions of the fusion protein ectodomain oligomerize to form homotrimeric coiled-coil cores, which when stimulated by receptor interaction undergo a conformational change to an antiparallel 6-stranded α-helical structure (trimer of dimers) in the fusion active state to initiate the fusion of viral and cellular membranes.

Acknowledgments—We thank Mark Genest, Jennifer Labrecque, Larissa Thackray, and Mark Young for help with peptide synthesis, cloning, and purification. CD and analytical ultracentrifugation measurements were performed at the Biophysics Core Facility, University of Colorado Health Sciences Center.

REFERENCES

1. Drosten, C., Gunther, S., Preiser, W., van der Werf, S., Brodt, H. R., Becker, S., Rabenau, H., Panning, M., Kolesnikova, L., Fouchier, R. A., Berger, A., Burguiere, A. M., Cinatl, J., Eickmann, M., Escriou, N., Grywna, K., Kramme, S., Manuguerra, J. C., Muller, S., Rickerts, V., Sturmer, M., Vieth, S., Klenk, H. D., Osterhaus, A. D., Schmitz, H. & Doerr, H. W. (2003) *N. Engl. J. Med.* **348,** 1967–1976
2. Ksiazek, T. G., Erdman, D., Goldsmith, C. S., Zaki, S. R., Peret, T., Emery, S., Tong, S., Urbani, C., Comer, J. A., Lim, W., Rollin, P. E., Dowell, S. F., Ling, A. E., Humphrey, C. D., Shieh, W. J., Guarner, J., Paddock, C. D., Rota, P., Fields, B., DeRisi, J., Yang, J. Y., Cox, N., Hughes, J. M., LeDuc, J. W., Bellini, W. J. & Anderson, L. J. (2003) *N. Engl. J. Med.* **348,** 1953–1966
3. Peiris, J. S., Lai, S. T., Poon, L. L., Guan, Y., Yam, L. Y., Lim, W., Nicholls, J., Yee, W. K., Yan, W. W., Cheung, M. T., Cheng, V. C., Chan, K. H., Tsang, D. N., Yung, R. W., Ng, T. K. & Yuen, K. Y. (2003) *Lancet* **361,** 1319–1325
4. Poon, L. L., Wong, O. K., Luk, W., Yuen, K. Y., Peiris, J. S. & Guan, Y. (2003) *Clin. Chem.* **49,** 953–955
5. El-Sahly, H. M., Atmar, R. L., Glezen, W. P. & Greenberg, S. B. (2000) *Clin. Infect. Dis.* **31,** 96–100
6. Folz, R. J. & Elkordy, M. A. (1999) *Chest* **115,** 901–905
7. Holmes, K. V. (2001) in *Fields Virology* (Knipe, D. M. & Howley, P. M., eds) 4th Ed., pp. 1187–1203, Lippincott Williams & Wilkins, Philadelphia
8. McIntosh, K. (1974) *Curr. Top. Microbiol. Immunol.* **63,** 85–129
9. Martina, B. E., Haagmans, B. L., Kuiken, T., Fouchier, R. A., Rimmelzwaan, G. F., Van Amerongen, G., Peiris, J. S., Lim, W. & Osterhaus, A. D. (2003) *Nature* **425,** 915
10. Guan, Y., Zheng, B. J., He, Y. Q., Liu, X. L., Zhuang, Z. X., Cheung, C. L., Luo, S. W., Li, P. H., Zhang, L. J., Guan, Y. J., Butt, K. M., Wong, K. L., Chan, K. W., Lim, W., Shortridge, K. F., Yuen, K. Y., Peiris, J. S. & Poon, L. L. (2003) *Science* **302,** 276–278
11. Bos, E. C., Heijnen, L., Luytjes, W. & Spaan, W. J. (1995) *Virology* **214,** 453–463
12. De Groot, R. J., Van Leen, R. W., Dalderup, M. J., Vennema, H., Horzinek, M. C. & Spaan, W. J. (1989) *Virology* **171,** 493–502
13. Luo, Z. & Weiss, S. R. (1998) *Virology* **244,** 483–494
14. Spaan, W., Cavanagh, D. & Horzinek, M. C. (1988) *J. Gen. Virol.* **69,** 2939–2952
15. Frana, M. F., Behnke, J. N., Sturman, L. S. & Holmes, K. V. (1985) *J. Virol.* **56,** 912–920
16. Sturman, L. S., Ricard, C. S. & Holmes, K. V. (1985) *J. Virol.* **56,** 904–911
17. Cavanagh, D., Davis, P. J., Darbyshire, J. H. & Peters, R. W. (1986) *J. Gen. Virol.* **67,** 1435–1442
18. Taguchi, F. (1995) *J. Virol.* **69,** 7260–7263
19. Bosch, B. J., van der Zee, R., de Haan, C. A. & Rottier, P. J. (2003) *J. Virol.* **77,** 8801–8811
20. Luo, Z., Matthews, A. M. & Weiss, S. R. (1999) *J. Virol.* **73,** 8152–8159
21. Matsuyama, S. & Taguchi, F. (2002) *J. Virol.* **76,** 11819–11826
22. Taguchi, F. & Matsuyama, S. (2002) *J. Virol.* **76,** 950–958
23. Zelus, B. D., Schickli, J. H., Blau, D. M., Weiss, S. R. & Holmes, K. V. (2003) *J. Virol.* **77,** 830–840
24. Skehel, J. J. & Wiley, D. C. (1998) *Cell* **95,** 871–874
25. Baker, K. A., Dutch, R. E., Lamb, R. A. & Jardetzky, T. S. (1999) *Mol. Cell* **3,** 309–319
26. Chan, D. C., Fass, D., Berger, J. M. & Kim, P. S. (1997) *Cell* **89,** 263–273
27. Caffrey, M., Cai, M., Kaufman, J., Stahl, S. J., Wingfield, P. T., Covell, D. G., Gronenborn, A. M. & Clore, G. M. (1998) *EMBO J.* **17,** 4572–4584
28. Fass, D., Harrison, S. C. & Kim, P. S. (1996) *Nat. Struct. Biol.* **3,** 465–469
29. Kobe, B., Center, R. J., Kemp, B. E. & Poumbourios, P. (1999) *Proc. Natl. Acad. Sci. U. S. A.* **96,** 4319–4324
30. Malashkevich, V. N., Chan, D. C., Chutkowski, C. T. & Kim, P. S. (1998) *Proc. Natl. Acad. Sci. U. S. A.* **95,** 9134–9139
31. Malashkevich, V. N., Schneider, B. J., McNally, M. L., Milhollen, M. A., Pang, J. X. & Kim, P. S. (1999) *Proc. Natl. Acad. Sci. U. S. A.* **96,** 2662–2667
32. Weissenhorn, W., Dessen, A., Harrison, S. C., Skehel, J. J. & Wiley, D. C. (1997) *Nature* **387,** 426–430
33. Weissenhorn, W., Carfi, A., Lee, K. H., Skehel, J. J. & Wiley, D. C. (1998) *Mol. Cell* **2,** 605–616
34. Tan, K., Liu, J., Wang, J., Shen, S. & Lu, M. (1997) *Proc. Natl. Acad. Sci. U. S. A.* **94,** 12303–12308
35. Yang, Z. N., Mueser, T. C., Kaufman, J., Stahl, S. J., Wingfield, P. T. & Hyde,

C. C. (1999) *J. Struct. Biol.* **126,** 131–144
36. Zhao, X., Singh, M., Malashkevich, V. N. & Kim, P. S. (2000) *Proc. Natl. Acad. Sci. U. S. A.* **97,** 14172–14177
37. Bewley, C. A., Louis, J. M., Ghirlando, R. & Clore, G. M. (2002) *J. Biol. Chem.* **277,** 14238–14245
38. Ghosh, J. K., Peisajovich, S. G., Ovadia, M. & Shai, Y. (1998) *J. Biol. Chem.* **273,** 27182–27190
39. Medinas, R. J., Lambert, D. M. & Tompkins, W. A. (2002) *J. Virol.* **76,** 9079–9086
40. Pinon, J. D., Kelly, S. M., Price, N. C., Flanagan, J. U. & Brighty, D. W. (2003) *J. Virol.* **77,** 3281–3290
41. Pritsker, M., Jones, P., Blumenthal, R. & Shai, Y. (1998) *Proc. Natl. Acad. Sci. U. S. A.* **95,** 7287–7292
42. Rapaport, D., Ovadia, M. & Shai, Y. (1995) *EMBO J.* **14,** 5524–5531
43. Tomasi, M., Pasti, C., Manfrinato, C., Dallocchio, F. & Bellini, T. (2003) *FEBS Lett.* **536,** 56–60
44. Yang, C. & Compans, R. W. (1997) *J. Virol.* **71,** 8490–8496
45. Zhou, G., Ferrer, M., Chopra, R., Kapoor, T. M., Strassmaier, T., Weissenhorn, W., Skehel, J. J., Oprian, D., Schreiber, S. L., Harrison, S. C. & Wiley, D. C. (2000) *Bioorg. Med. Chem.* **8,** 2219–2227
46. Young, J. K., Hicks, R. P., Wright, G. E. & Morrison, T. G. (1997) *Virology* **238,** 291–304
47. Tripet, B., Wagschal, K., Lavigne, P., Mant, C. T. & Hodges, R. S. (2000) *J. Mol. Biol.* **300,** 377–402
48. Tabor, S. & Richardson, C. C. (1985) *Proc. Natl. Acad. Sci. U. S. A.* **82,** 1074–1078
49. Zelus, B. D., Wessner, D. R., Williams, R. K., Pensiero, M. N., Phibbs, F. T., deSouza, M., Dveksler, G. S. & Holmes, K. V. (1998) *J. Virol.* **72,** 7237–7244
50. Hayes, D. B., Laue, T. & Philo, J. (2003) *Sedimentation Interpretation Program,* Version 1.08, University of New Hampshire
51. Johnson, M. L., Correia, J. J., Yphantis, D. A. & Halvorson, H. R. (1981) *Biophys. J.* **36,** 575–588
52. Semchuk, P. D., Monera, O. D., Kondejewski, L. H., Gannon, C., Daniels, L., Wilson, I. & Hodges, R. S. (1996) in *Peptides: Chemistry, Structure and Biology* (Kaumaya, P. T. P. & Hodges, R. S., eds) pp. 73–74, Mayflower Scientific Ltd., UK
53. Tripet, B. & Hodges, R. S. (2001) in *Peptides: The Wave of the Future* (Lebl, M. & Houghten, R. A., eds) pp. 365–366, American Peptide Society, San Diego, CA
54. Chen, L., Gorman, J. J., McKimm-Breschkin, J., Lawrence, L. J., Tulloch, P. A., Smith, B. J., Colman, P. M. & Lawrence, M. C. (2001) *Structure* **9,** 255–266
55. Yang, X., Farzan, M., Wyatt, R. & Sodroski, J. (2000) *J. Virol.* **74,** 5716–5725
56. Yang, X., Florin, L., Farzan, M., Kolchinsky, P., Kwong, P. D., Sodroski, J. & Wyatt, R. (2000) *J. Virol.* **74,** 4746–4754
57. Matthews, J. M., Young, T. F., Tucker, S. P. & Mackay, J. P. (2000) *J. Virol.* **74,** 5911–5920
58. Weissenhorn, W., Calder, L. J., Wharton, S. A., Skehel, J. J. & Wiley, D. C. (1998) *Proc. Natl. Acad. Sci. U. S. A.* **95,** 6032–6036
59. Gibbons, D. L., Ahn, A., Chatterjee, P. K. & Kielian, M. (2000) *J. Virol.* **74,** 7772–7780
60. Bullough, P. A., Hughson, F. M., Skehel, J. J. & Wiley, D. C. (1994) *Nature* **371,** 37–43
61. Carr, C. M., Chaudhry, C. & Kim, P. S. (1997) *Proc. Natl. Acad. Sci. U. S. A.* **94,** 14306–14313
62. Watowich, S. J., Skehel, J. J. & Wiley, D. C. (1994) *Structure* **2,** 719–731
63. Wilson, I. A., Skehel, J. J. & Wiley, D. C. (1981) *Nature* **289,** 366–373
64. Harbury, P. B., Zhang, T., Kim, P. S. & Alber, T. (1993) *Science* **262,** 1401–1407
65. Kohn, W. D., Kay, C. M. & Hodges, R. S. (1998) *J. Mol. Biol.* **283,** 993–1012, and references therein
66. Holtzer, M. E., Lovett, E. G., d'Avignon, D. A. & Holtzer, A. (1997) *Biophys. J.* **73,** 1031–1041
67. Zhou, N. E., Kay, C. M. & Hodges, R. S. (1992) *J. Biol. Chem.* **267,** 2664–2670
68. Zhou, N. E., Zhu, B. Y., Kay, C. M. & Hodges, R. S. (1992) *Biopolymers* **32,** 419–426