Folding of the SARS Coronavirus Spike Glycoprotein Immunological Fragment (SARS_S1B): Thermodynamic and Kinetic Investigation Correlating with Three-Dimensional Structural Modeling†

Changying Yu,‡,§ Chunshan Gui,‡,§ Haibin Luo,‡ Lili Chen,‡ Liang Zhang,‡ Hao Yu,|| Sheng Yang,|| Weihong Jiang,|| Jianhua Shen,‡ Xu Shen,*,‡ and Hualiang Jiang*,‡,⊥

Drug Discovery and Design Center, State Key Laboratory of Drug Research, Shanghai Institute of Materia Medica, Shanghai Institutes for Biological Sciences, Graduate School, Chinese Academy of Sciences, Shanghai 201203, China, Shanghai Institute of Plant Physiology and Ecology, Shanghai Institutes for Biological Sciences, Graduate School, Chinese Academy of Sciences, 300 Fenglin Road, Shanghai 200032, China, and School of Pharmacy, East-China University of Science and Technology, Shanghai 200237, China

Received August 14, 2004; Revised Manuscript Received November 13, 2004

ABSTRACT: Spike glycoprotein of SARS coronavirus (S protein) plays a pivotal role in SARS coronavirus (SARS_CoV) infection. The immunological fragment of the S protein (Ala251−His641, SARS_S1b) is believed to be essential for SARS_CoV entering the host cell through S protein−ACE-2 interaction. We have quantitatively characterized the thermally induced and GuHCl-induced unfolding features of SARS_S1b using circular dichroism (CD), tryptophan fluorescence, and stopped-flow spectral techniques. For the thermally induced unfolding at pH 7.4, the apparent activation energy (E_{app}) and transition midpoint temperature (T_m) were determined to be 16.3 ± 0.2 kcal/mol and 52.5 ± 0.4 °C, respectively. The CD spectra are not dependent on temperature, suggesting that the secondary structure of SARS_S1b has a relatively high thermal stability. GuHCl strongly affected SARS_S1b structure. Both the CD and fluorescent spectra resulted in consistent values of the transition middle concentration of the denaturant (C_m, ranging from 2.30 to 2.45 M) and the standard free energy change ($\Delta G°$, ranging from 2.1 to 2.5 kcal/mol) for the SARS_S1b unfolding reaction. Moreover, the kinetic features of the chemical unfolding and refolding of SARS_S1b were also characterized using a stopped-flow CD spectral technique. The obvious unfolding reaction rates and relaxation times were determined at various GuHCl concentrations, and the C_m value was obtained, which is very close to the data that resulted from CD and fluorescent spectral determinations. Secondary and three-dimensional structural predictions by homology modeling indicated that SARS_S1b folded as a globular-like structure by β-sheets and loops; two of the four tryptophans are located on the protein surface, which is in agreement with the tryptophan fluorescence result. The three-dimensional model was also used to explain the recently published experimental results of S1−ACE-2 binding and immunizations.

From the end of 2002 to June 2003, a severe epidemic disease called severe acute respiratory syndrome (SARS)[1] broke out in China and other countries in the world. SARS ever strictly menaced the worldwide population (*1*−*3*). SARS coronavirus (SARS_CoV) was identified as being responsible for SARS infection (*4*−*6*). Recently, remarkable achievements have been made in genome sequencing of SARS_CoV (*7*, *8*), SARS_CoV protein functional studies (*9*−*11*), three-dimensional (3D) structural determination and modeling of SARS_CoV proteins (*12*−*17*), clinical studies (*4*), and anti-SARS drug discovery (*18*, *19*). It has been demonstrated that the important proteins associated with SARS_CoV infection involve the RNA polymerase, the spike (S) glycoprotein, the envelope (E) protein, the membrane (M) protein, the nucleocapsid (N) protein, and the main protease (3C-like proteinase).

† This work was supported by the State Key Program of Basic Research of China (Grants 2003CB514125, 2002CB512801, 2002CB512802, and 2002CB512807), the National Natural Science Foundation of China (Grants 20372069 and 20472095), Shanghai Basic Research Project from the Shanghai Science and Technology Commission (Grants 02DJ14070, 03DZ19228, and 03DZ19212), the 863 Hi-Tech Program (Grants 2001AA235051, 2002CB512807, and 2002AA233011), the Sino-European Project on SARS Diagnostics and Antivirals (Proposal/Contract 003831), and the special programs of oppugning SARS from the Ministry of Science and Technology, Chinese Academy of Sciences, the National Natural Science Foundation of China, the Sino-German Science Center, and the Shanghai Science and Technology Commission.

* To whom correspondence should be addressed. E-mail: xshen@mail.shcnc.ac.cn (X.S.) or hljiang@mail.shcnc.ac.cn (H.J.). Telephone: 86-21-50807188. Fax: 86-21-50807088.

‡ Shanghai Institute of Materia Medica, Shanghai Institutes for Biological Sciences, Graduate School, Chinese Academy of Sciences.

§ These authors contributed equally to this work.

|| Shanghai Institute of Plant Physiology & Ecology, Shanghai Institutes for Biological Sciences, Graduate School, Chinese Academy of Sciences.

⊥ East-China University of Science and Technology.

[1] Abbreviations: SARS, severe acute respiratory syndrome; SARS_CoV, SARS coronavirus; SARS_S, spike glycoprotein of SARS_CoV; SARS_S1b, immunological fragment of SARS_S (Ala251−His641); CD, circular dichroism; GuHCl, guanidinium hydrochloride; T_m, transition midpoint temperature; C_m, transition midpoint concentration.

10.1021/bi0482396 CCC: $30.25
Published on Web 01/06/2005

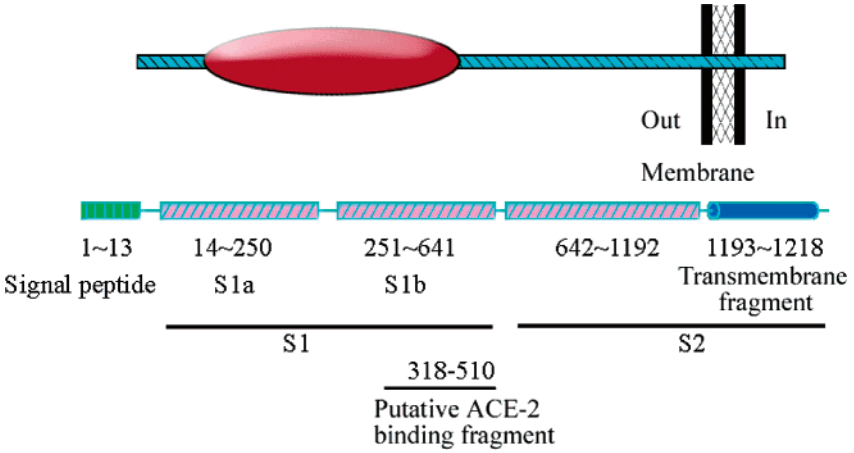


FIGURE 1: Diagrammatic representation of SARS_S.

The spike glycoprotein (S protein) plays an important role in virus entry, virus−receptor interactions, and virus tropism (*20*−*22*). Several succeeding studies have also revealed that S protein has other important functions, including binding of virus to susceptible cells, mediation of membrane fusion (both virus−cell and cell−cell fusions), and induction of neutralizing antibody responses in the host species (*23*−*26*). Structurally, S protein is a surface projection glycoprotein and may be cleaved by virus-encoded or host-encoded proteinases into two functional subunits, S1 and S2 (*27*). The N-terminal subunit (S1), forming the surface knoblike structure of the spike, seems to be more important in the recognition of the host membrane (*20*, *26*, *28*, *29*), while the C-terminal, membrane-anchored subunit (S2), which forms the stemlike structure beneath the knob, is involved in fusion activity.

The S protein of SARS_CoV consisting of 1255 amino acids (SARS_S) contains two hydrophobic regions; one is located at the N-terminus of the entire protein, including a short type I signal sequence, and the other is situated at the C-terminus with a transmembrane domain and a cytoplasmic tail rich in highly conserved cysteine residues (Figure 1) (*16*). Li et al (*26*) found that angiotensin-converting enzyme 2 (ACE-2) may efficiently bind to the S1 domain of the SARS_S protein, and this protein−protein binding possibly plays an essential role in SARS virus infecting the host cells. More accurately, Wong et al. found that the site of binding of the S1 protein to ACE-2 is located in the region of residues 318−510, which blocks S protein-mediated infection more efficiently than does the full-length S1 protein (*29*). This study further implicated that disrupting the ACE-2−S protein interaction might be a potential method of anti-SARS infection. For most of the viruses, such as the infectious bronchitis virus (IBV) (*28*) and human coronavirus 229E (*30*), the fragment located at the C-terminus of the S1 protein neighboring the S2 protein (Ala251−His641) is conservative (Figure 1). Recently, Sui et al. (*31*) reported that one single-chain variable region fragment 80R efficiently neutralized SARS_CoV and blocked the binding of S1 to ACE-2. Mapping of the 80R human monoclonal antibody epitope showed that it is located within N-terminal amino acids 261−672 of the S protein. More recently, He et al. demonstrated that the SRAR_S protein contains five linear immunodominant sites (I−V), and site IV (residues 528−635) is a major immunodominant epitope (*32*). Protein domain analysis indicated that SARS_S1b (Ala251−His641) is a domain of the S protein, mostly overlapping the ACE-2 binding site (*29*) and the immunodominant epitope (*31*, *32*). These data indicate that the S1b domain is obviously associated with the entrance mechanism of SARS infection. Accordingly, the structural and functional study of the SARS_S1b domain is significant.

In the following, we report the study of the unfolding and refolding processes of SARS_S1b induced by heating and a chemical denaturant using circular dichroism (CD) spectroscopy, tryptophan fluorescence, and stopped-flow spectral techniques. The three-dimensional (3D) structure of SARS_S1b was predicted using a homologue modeling method. The experimental and structural modeling results are consistent with each other.

MATERIALS AND METHODS

Enzymes and Chemicals. All chemicals were reagent grade or ultrapure quality and were purchased from Sigma (St. Louis, MO). Protease for tag cleavage and low-molecular weight markers for SDS−PAGE were from Amersham Pharmacia Biotech (Uppsala, Sweden).

Protein Preparation. The plasmid of pET32c-SARS_S1b was cloned according to the published method (*33*), and expression and purification of the recombinant SARS_S1b protein were performed on the basis of the literature result (*34*). The purity and identity of this protein were confirmed by SDS−PAGE and LC−MS spectral determination. Additionally, the SARS_S1b protein contained in the pellet after cell disruption was washed sequentially by 1% sucrose, 0.5% Triton X-100, 1% sucrose, and 2 M urea, and then dissolved in 6 M GuHCl or 8 M urea. The refolding procedure for the SARS_S1b protein was dialyzed against buffer A [20 mM Tris-HCl, 500 mM NaCl, 5 mM imidazole, and 1 mM 2-mercaptoethanol (pH 8.0)] supplemented with 5% glycerol. The concentration of protein was determined according to its molar extinction coefficient (ϵ_{280} = 63 520 M^{-1} cm^{-1}).

Fluorescence Spectroscopy. Fluorescence spectra were recorded on a Hitachi F2500 spectrometer. The samples (2.5 mL) were processed in a quartz cell with a path length of 1 cm. Fluorescence emission was monitored from 300 to 380 nm, with excitation at 280 nm, an excitation slit width of 10 nm, an emission slit width of 5 nm, a scan speed of 60 nm/

min, and a response of 1.0 s. A Neslab water bath was used to control the experimental temperature. The data were automatically collected when the temperature vibration in the cell was less than 0.1 °C. The pH value of all solutions was found to vary by less than 0.1 unit between 15 and 90 °C. All the other background effects were subtracted during the data analyses.

Circular Dichroism (CD) Spectroscopy. Far-UV CD spectra were recorded on a Jasco J-810 spectropolarimeter equipped with a Neslab water bath. All determinations were processed using a quartz cell with path length of 0.1 cm and a spectral bandwidth of 1.0 nm. For the mode of wavelength scans, the ellipticities from 250 to 195 nm were scanned at a rate of 100 nm/min and a time constant of 4 s. In the presence of a denaturant, a meaningful signal was restricted only above 210 nm due to the noise caused by the denaturant. Averages of six scans were recorded. For the mode of temperature scans, the movements of ellipticities at a wavelength of 215 nm were obtained at a scan rate of 2 °C/min from 10 to 90 °C. All solution blanks showed no changes in ellipticity with temperature and were thus neglected during data analysis. The data from three independent experiments were averaged.

Kinetic Measurements by Stopped-Flow Circular Dichroism. The rapid kinetics of protein folding and unfolding by a chemical denaturant was monitored using the stopped-flow circular dichroism (stopped-flow CD) system. The determinations were performed using a Jasco J-810 spectropolarimeter equipped with SFM-300 (Biologic Co.). The rectangular cuvette with a path length of 2.0 mm (FC-20) for loading samples was put in the path of light. In the unfolding process, the native protein and denaturant buffer B [0.1 mM sodium phosphate and 6 M GuHCl (pH 7.4)] were each loaded into a different syringe. Measurements were initiated by dilution of the native protein (SARS_S1b) from buffer B. The denaturant concentration was determined according to the injected volumes during each mixing of the protein with buffer B. On the other hand, the unfolded protein sample which was denatured in the presence of a chemical denaturant was then refolded by stopped-flow dilution with refolding buffer C [0.1 mM sodium phosphate and 5 mM 1,4-dithiothreitol (pH 7.4)]. The 3D structural model of SARS_S1b derived from the homology modeling suggests that it is impossible to form a disulfide bond between any two cysteine residues (see the result below), indicating that both the unfolded and refolded proteins are in the reductive state. Both unfolding and refolding procedures were monitored at 215 nm, and the dead time of the stopped-flow system was 4 ms. Kinetic traces are averages of at least eight acquisitions and were analyzed using the Bio-kine software (BioLogic). The CD values were fitted with the following equation

$$\theta = at + b + ce^{-k_{obs}t} \quad (1)$$

where θ is the observed CD value at any given time (t) and a, b, c, and k_{obs} represent the slope, offset, amplitude, and apparent rate constant during the denaturation procedure, respectively. The relaxation time of the transition was calculated from the reciprocal of the apparent rate constant.

Studying Equilibrium by Fitting to a Simple Two-State Transition. The denaturation experiments were performed by diluting the protein stock solution with 6 M GuHCl to attain the given protein concentration, 30 μM for CD and 10 μM for fluorescence measurements, at the desired denaturant concentration. The solutions were then incubated for 2 h to attain complete equilibrium. The typical two-state transition (N $\Leftrightarrow$ U) was presumed in the denaturalization. The thermodynamic stability of a macromolecular native state can be expressed in terms of the standard free energy of folding ΔG. Given the equilibrium constant (K) for the folding reaction, we have

$$\Delta G = -RT \ln K \quad K = \frac{[\text{native}]}{[\text{denatured}]} = \frac{A_1 - A}{A - A_2} \quad (2)$$

where [native] and [denatured] are the concentrations of the protein in the two states, A is the equilibrium value of the spectroscopic parameter of each sample in the presence of a denaturant at the given concentration, and A_1 and A_2 stand for the values of A characteristic of the unfolded and folded forms, respectively. ΔG values are then fitted to the linear regression by the concentration of denaturant (C) in eq 3

$$\Delta G = mC + \Delta G^\circ \quad (3)$$

where the slope m is the cooperative index. At the midpoint concentration of the denaturant (C_m), the folding equilibrium constant (K) reaches 1; therefore, $\Delta G = 0$, and $\Delta G^\circ = -mC_m$. Equation 3 is modified as

$$\Delta G = mC - mC_m \quad (4)$$

where m and C_m are thus associated with the experiments by the following equations. From eqs 2 and 4, we have

$$-RT \ln \frac{A_1 - A}{A - A_2} = m(C - C_m) \quad (5)$$

$$A = \frac{A_1 + A_2 e^{-m(C-C_m)}}{1 + e^{-m(C-C_m)}} \quad (6)$$

or

$$A = \frac{A_1 - A_2}{1 + e^{-m(C-C_m)}} + A_2 \text{ [sigmoidal (Boltzmann)]} \quad (7)$$

With nonlinear fitting to the curve of A versus C, all four parameters in eq 7 can be obtained.

3D Structure of SARS_S1b Constructed by Homology Modeling. Sequence similarity comparison of SARS_S1b against known crystal structure proteins was carried out using PSI-BLAST. The score and E value of the best PSI-BLAST record are 31 bits and 0.83, respectively. Obviously, there is no significant sequence homology between SARS_S1b and template proteins. Therefore, the routine method based on the sequence homology is not suitable for modeling the 3D structure of SARS_S1b. We adopted sequence alignment combined with protein fold recognition to construct the 3D model of SARS_S1b.

The SeqFold module of Insight II (Molecular modeling package, version 2000, Accelrys, San Diego, CA) was employed to identify the folding pattern of SARS_1b. An array of structures sharing significant fold similarities to SARS_S1b (with P value < 0.0001) were identified.

According to the scoring result, two proteins (PDB entries 1AOF and 1NIR) were selected as templates for 3D structural construction. Sequence alignments of SARS_1b with the two templates were performed using the Align123 module of Insight II, which is a sequence alignment method developed on the basis of the CLUSTAL_W program (*35*). On the basis of the sequence alignment, the 3D model of SARS_S1b was thus generated using MODELLER (*36*). The 3D model was refined by the following steps. (i) Loops were fixed by the Loop_refine program, and the side chain conformations of all residues were automatically rearranged by Auto_Rotamer. (ii) The Discover module of Insight II and Amber95 force field were employed to carry out a short time molecular dynamics simulation (MD) and energy minimization. During the structural refinement, a 100-step initial equilibration simulation was first carried out, and the system was then subjected to MD simulation with a time step of 1 fs for 1000 steps at a constant pressure of 1.0 bar and a constant temperature of 300 K. Energy minimization was performed on the structure resulting from the short MD simulation to obtain a low-energy structure. The system was subjected to a 100-step steepest descent energy minimization, followed by a 500-step conjugate gradient minimization.

Several structural analysis softwares were adopted to check the resulting structure. The Prostat module of Insight II was used to analyze the properties of bonds, angles, and torsions. The PROCHECK (*37*) suite was employed to assess the stereochemical quality and secondary structure properties of the SARS_S1b protein, while Profile-3D was used to check the structure and sequence compatibility.

RESULTS

Intrinsic Tryptophan Fluorescence. The SARS_S1b protein, excited at 280 nm, emits fluorescence at 20 °C, giving only one peak at ~340 nm ranging from 300 to 380 nm (Figure 2a). This fluorescence emission can be attributed to the tryptophan residues. The purified protein from the inclusion body has a similar fluorescence feature (data not shown), indicating that the refolded SARS_S1b has a tryptophan environment similar to the native one. As shown in Figure 2b, from 15 to 90 °C, the maximum fluorescence absorption moves gradually from 333 to 339 nm, while its emission intensity decreases from 430 to 285 nm.

The irreversible thermal denaturation of SARS_S1b protein was detected using the fluorescence spectrum method. The apparent activation energy (E_{app}) of the transition of SARS_S1b denaturation can be estimated by eq 8:

$$\ln\left(\ln\frac{1}{\alpha}\right) = \frac{E}{R}\left(\frac{1}{T_m} - \frac{1}{T}\right) \quad (8)$$

where α, the apparent fraction of the native protein, is defined as $(F - F_D)/(F_N - F_D)$, where F is the sample fluorescence intensity at a particular temperature and F_N and F_D are the corresponding values for the native and denatured state of the protein, respectively, and T_m is the temperature at which the maximum of the heat capacity curve occurs (Figure 2c). The plot of $\ln[\ln(1/\alpha)]$ versus $1/T$ of SARS_S1b at pH 7.4 is shown in Figure 2c, which indicates that the relationship is linear. Accordingly, the apparent activation energy of SARS_S1b denaturation can be estimated by eq 8. At pH 7.4, the apparent activation energy (E_{app}) of SARS_S1b denaturation and T_m were estimated to be ~16.3 ± 0.2 kcal/mol and ~52.5 ± 0.4 °C, respectively.

Temperature-Dependent CD Spectra. Figure 3a shows the CD spectra of SARS_S1b in buffer E [0.1 mM sodium phosphate and 5 mM 1,4-dithiothreitol (pH 7.4)] at 20 and 90 °C. Different from the effect of temperature on the fluorescence, few changes in the CD spectra were observed despite the increase in the temperature to 90 °C, as shown in Figure 3a. This implies that the polypeptide backbone of SARS_S1b protein is relatively thermostable. In addition, thermal melting of SARS_S1b at 215 nm gave similar results (Figure 3b). The thermal endurance of the unfolded protein in 6 M GuHCl is almost the same as that of the native one in the temperature scan process (Figure 3b).

Unfolding under Equilibrium Conditions. To inspect the chemical unfolding of SARS_S1b, the emission fluorescences of the native (in buffer C) and denatured (in buffer C containing a given concentration of GuHCl) proteins were determined. Figure 4a shows the unfolding equilibrium curves of the maximum emission wavelengths and the fluorescence intensities versus the denaturant concentrations. Similar to that for the thermally induced denaturation (Figure 2b), the maximum emission wavelength increases and the emission intensity decreases along with the concentration of the denaturant increasing. For example, in the presence of 4 M denaturant, the maximum emission wavelength shifts to 344.5 nm, and the emission intensity drops to 225. In addition, the far-UV CD spectra of SARS_S1b in both buffer C and buffer C containing different concentrations (0, 1, 2, 2.5, 3, and 4 M) of GuHCl were measured, which are shown in Figure 4b. As indicated from the CD spectra (Figure 4b inset), SARS_S1b becomes a typical random coil in the presence of a high concentration of the denaturant. The plots of the CD signals at 222 and 215 nm versus GuHCl concentration are also shown in Figure 4b, from which the thermodynamic parameters for SARS_S1b unfolding can be deduced using eq 7. The result is listed in Table 1.

Unfolding and Refolding Kinetics. CD kinetic traces obtained from different volume ratios of protein to denaturant contain the information about the reaction rate. The CD kinetic traces of native SARS_S1b unfolding in buffer D [0.1 mM sodium phosphate and 6 M GuHCl (pH 7.4)] with different protein:GuHCl ratios at 215 nm were determined using two BioLogic SFM-300 microvolume stopped-flow syringes with a JASCO CD detector. The results are shown in Figure 5a, where the unfolding rate constant (k_{obs}) and relaxation time ($\tau = 1/k_{obs}$) can be obtained. In a similar way, the kinetic traces of denatured SARS_S1b refolding were determined, as shown in Figure 5b, and the refolding rate constant and relaxation time were estimated accordingly. During the kinetic data fitting, we assumed that the unfolding and refolding processes make up a first-order reaction (U ⇔ N). Figure 6a shows the relaxation time of SARS_S1b unfolding and refolding in buffers with different volume ratios of protein to unfolding (or refolding) buffer. From Figure 6b, we can see that either unfolding or refolding of SARS_S1b is a rapid process, and refolding is faster than unfolding. Figure 6b represents the plot of relaxation time versus volume ratio (R) between protein and unfolding buffer, from which the volume ratio corresponding to the midpoint of relaxation time (R_m) was estimated to be 1.26 ± 0.08,

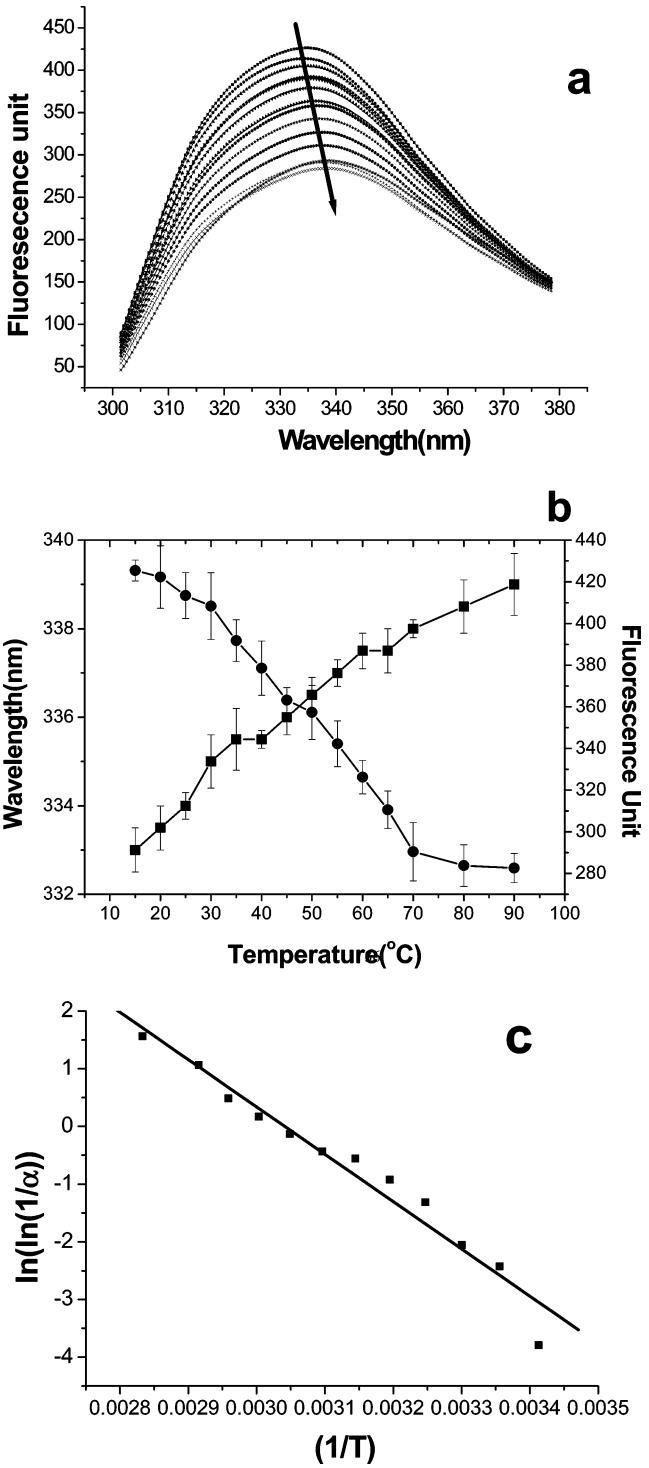


FIGURE 2: Fluorescence spectral analyses of SARS_S1b. (a) Emission fluorescence for a wavelength scan at temperatures ranging from 15 to 90 °C recorded by the temperature increment of 5 or 10 °C from the top down. The arrow shows the trend for the maximum values. (b) Red shift in the process of the thermally induced denaturation represented by plots of apex wavelength (■) and emission intensity peak (●) vs temperature. (c) Linear relationship between ln[ln(1/α)] and 1/*T* (eq 8). The correlation coefficient $R = 0.99$. From the slope and intercept of this line, the apparent average activation energy (E_{app}) was estimated to be 16.3 ± 0.2 kcal/mol, and the transition midpoint temperature (T_m) of SARS_S1b thermally induced unfolding can be calculated as 52.5 ± 0.4 °C.

and then the C_m of the GuHCl concentration was calculated to be 2.65 ± 0.12 M.

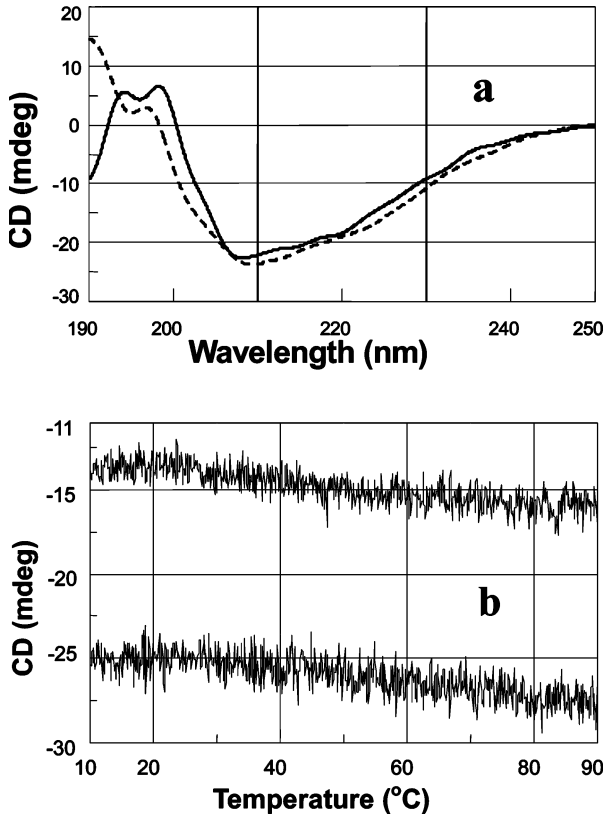


FIGURE 3: (a) Far-UV CD spectra of SARS_S1b protein in buffer E [0.1 mM sodium phosphate and 5 mM 1,4-dithiothreitol (pH 7.4)] at 20 (—) and 90 °C (- - -). (b) Raw data of thermal melting of the native protein (bottom trace) in buffer E and unfolding protein (top trace) in buffer D [0.1 mM sodium phosphate and 6 M GuHCl (pH 7.4)]. The CD signal was monitored at 208 nm; the scan rate was 2 °C/min, and the detection pitch was 0.1 °C.

2D and 3D Structural Predictions. Secondary structural analysis results from both 2D structural prediction and CD experiment for SARS_S1b are listed in Table 2 and shown in Figure 7a. It is found that the 2D structural prediction result with the percentages of 3, 35, and 62% for α-helix, *β*-sheet, and coil, respectively, is generally in agreement with the CD determination, indicating that *β*-sheet is the major component of the secondary structure of SARS_S1b. Sequence alignment showed that SARS_S1b is mostly homologous to cytochrome cd_1 nitrite reductase (PDB entries 1AOF and 1NIR) (*38*, *39*), and the final alignment of SARS_S1b with the sequences of the two proteins is shown in Figure 7b. On the basis of this alignment and using the crystal structures of these proteins as templates, a 3D model of SARS_S1b was generated, and then was refined by several methods. The 3D model is shown in Figure 8, which indicates that SARS_S1b folded as a globular-like structure in *β*-sheets. This is in agreement with the secondary structure prediction (Figure 7a). The stability of the structural model in aqueous solution was verified by a 5 ns molecular dynamics simulation (data not shown). As shown in Figure 8, Trp340 and Trp476 are fully buried inside the SARS_S1b structure, while Trp423 and Trp619 are located on the surface of the structure. Solvent accessible surface (SAS) analysis indicated that 85 residues are fully exposed to the solvent, and 62% of them are hydrophilic.

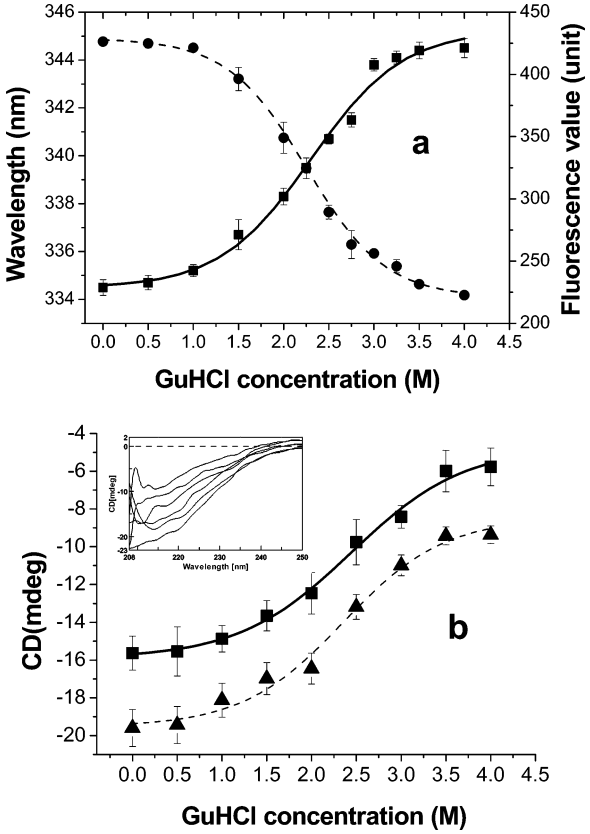


FIGURE 4: Unfolding equilibrium. (a) Result of fluorescence determination: plots of the maximum emission wavelength (■) and peak fluorescence value (●) vs GuHCl concentration. The curves were fitted by the sigmoidal method (eq 7). (b) Result of CD determination: the ellipticities at 222 (■) and 215 nm (▲) vs GuHCl concentration, which were also fitted by the sigmoidal fitting method. The inset shows the wavelength scans of SARS_S1b in the presence of 0, 1, 2, 2.5, 3, and 4 M GuHCl (from bottom to top).

Table 1: Thermodynamic Parameters for the Unfolding Transition of SARS_S1b Induced by GuHCl at 298 K

	F (intensity)[a]	F (λ_{max})[b]	CD (222 nm)	CD (215 nm)
C_m (M)	2.34 ± 0.05	2.36 ± 0.09	2.44 ± 0.11	2.35 ± 0.13
m (kcal mol^{-1} M^{-1})	0.89 ± 0.11	0.98 ± 0.13	0.95 ± 0.05	1.07 ± 0.07
$\Delta G°$ (kcal/mol)	2.08 ± 0.21	2.31 ± 0.31	2.32 ± 0.24	2.51 ± 0.29

[a] Data were fitted to the intensity of the fluorescence. [b] Data were fitted to the maximum emission wavelength.

DISCUSSION

Thermally Induced Unfolding. Since thermal denaturation of a protein is generally irreversible, the unfolding process is thereby characterized by the apparent activation energy (E_{app}) instead of the free energy change (ΔG) (*40*). The apparent activation energy for thermally induced unfolding of SARS_S1b was determined to be 16.3 ± 0.2 kcal/mol, which was relatively larger than the putative transition energies for other proteins, ranging from 5 to 15 kcal/mol (*41*, *42*). This suggests that the folded conformation of SARS_S1b is highly stable, which is also in agreement with the result for the denaturation transition analysis (see below).

The λ_{max} position of the fluorescence peak emission for tryptophan residues in a protein usually moves toward a longer wavelength during unfolding with the tryptophans exposed to a polar medium. Therefore, the λ_{max} position generally reflects the tryptophan environment. The 3D structural model of SARS_S1b indicates that two of the four tryptophans are exposed to the solvent (Figure 8a,b). In addition, CD spectra indicate that the secondary structures have not dramatically changed during thermal melting (Figure 3). All these are in agreement with the λ_{max} position change of tryptophans, which moves slightly, from 333 to 339 nm, during thermally induced unfolding. This suggests indirectly that our 3D model of SARS_S1b is reliable.

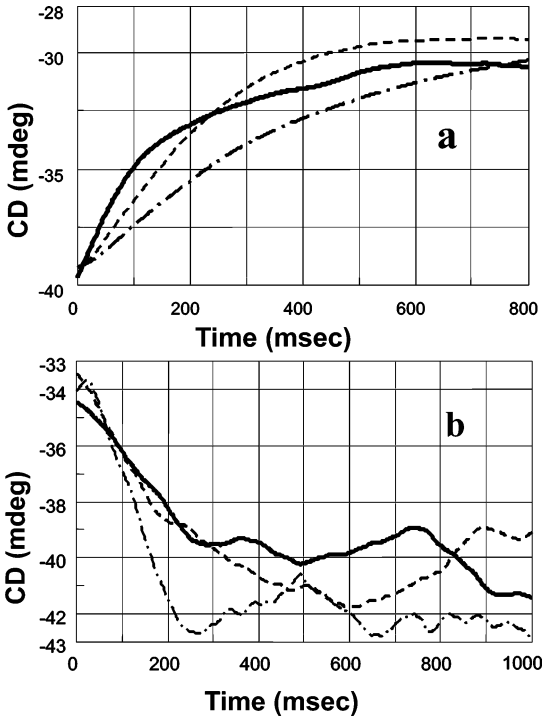


FIGURE 5: Stopped-flow trace of kinetic features of SARS_S1b unfolding and refolding. (a) The native SARS_S1b protein and denaturant in the unfolding process were mixed in volume ratios of 1:1 (—), 1:2 (- - -), and 1:3 (—•—). (b) The unfolded protein and refolding buffer were mixed in volume ratios of 1:1 (—), 1:2 (- - -), and 1:3 (—•—).

GuHCl-Induced Unfolding. Because the GuHCl-induced unfolding (refolding) process is reversible, a two-state transition model (N ⇔ U) is suitable for analyzing the chemical unfolding data. Four sets of parameters (C_m, m, and $\Delta G°$) have been obtained by fitting eq 5 on the basis of four different kinds of experimental data (Table 1). The consistency among these four sets of parameters indicates the reliability of our experimental methods in assessing SARS_S1b unfolding (refolding).

Deshpande et al. (*42*) pointed out that tryptophan in nonpolar solvents exhibits its maximum emission at 320 nm, whereas in a polar environment, the maximum is at 355 nm. In the work presented here, the λ_{max} position of SARS_S1b was shown to change from 333.5 to 344.5 nm during the chemical denaturation, which indicates that some of tryptophans might be located on the surface of the native structure of SARS_S1b. This in agreement with the 3D model of SARS_S1b, demonstrating again the reliability of our 3D model of SARS_S1b (Figure 8a,b). On the other hand, the movement of the λ_{max} position during chemical denaturation is larger than that in thermally induced unfolding, suggesting that the conformational change caused by

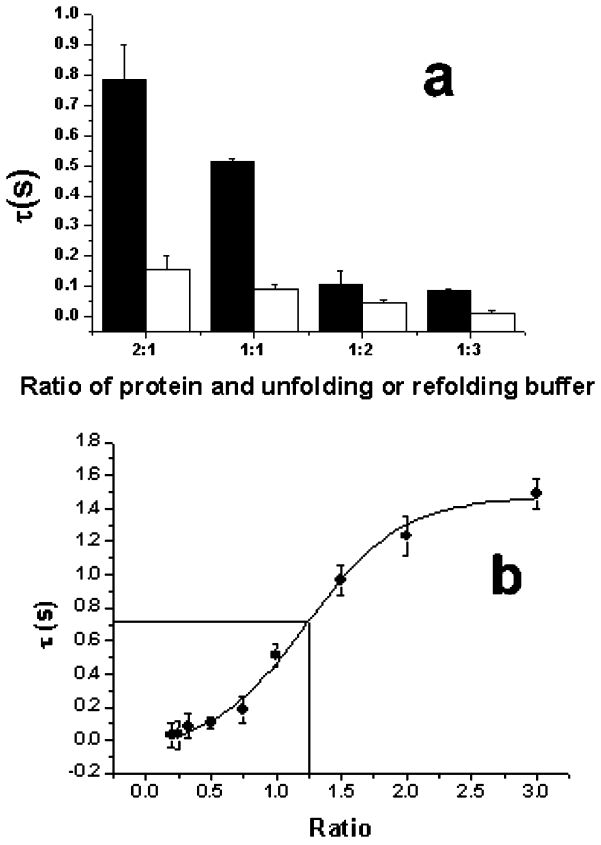


FIGURE 6: Relaxation time determined using the stopped-flow method. (a) Comparison of the relaxation times at the volume ratio of protein to unfolding buffer (black columns) in the unfolding phase with that of the fully unfolded protein to refolding buffer (white columns). (b) Plot of relaxation time vs the volume mixing ratio between the native protein in sodium phosphate buffer and 6 M GuHCl. The original data points were fitted by the sigmoidal method. The two beelines direct the middle point of the curve.

Table 2: Contributions of the Secondary Structure of SARS_S1b (%)

	estimations from CD data		secondary structure prediction	
	method 1[a]	method 2 [b]	this study	ref *14*
α-helix	16.9	15.6	3.0	2.6
β-sheet	28.4	35.1	35.0	33.0
other	54.7	49.3	62.0	64.4

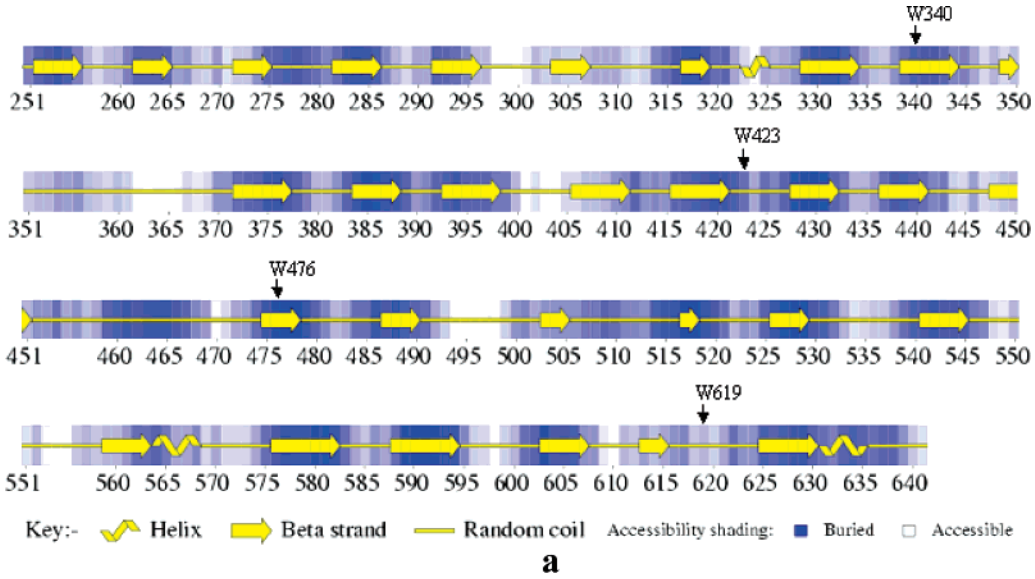
[a] Results estimated using the Yang method encoded in the software of the CD instrument (Jasco 810). [b] Results estimated using the Jfit method (http://www-structure.llnl.gov/cd).

chemical denaturation produced a different end state compared with those obtained with the heat treatment. This result was also reflected in the results of CD measurements, which showed that the chemical denaturation of SARS_S1b broke more secondary structures than the thermally induced unfolding (Figure 4).

Correlation of Denaturation between Equilibrium and Kinetic Conditions. The chemical denaturation of SARS_S1b was assessed at the steady state in the presence of a series of discontinued concentrations of denaturants. At the same time, the kinetic features of refolding and unfolding of SARS_S1b were monitored by a real-time stopped-flow instrument with a CD detector, on the basis of which was calculated the observed reaction rate (k_{obs}) using eq 1, and the reaction relaxation time (τ) of the kinetic reaction could be thus derived ($\tau = 1/k_{obs}$). Since the τ value is only dependent on the injected volume ratio of protein to buffer containing denaturant in the stopped flow (Figure 6a), a series of τ values corresponding to various volume ratios of protein to buffer were determined (Figure 6b), from which the midpoint concentration (C_m') of the denaturant was obtained (2.69 ± 0.12 M). To our knowledge, this is the first use of the real-time stopped-flow method in determining the midpoint concentration for protein unfolding (refolding). The value of C_m' is very close to those determined under equilibrium conditions (Table 1), indicating that this method can be extended to the folding study for other proteins. The consistency between C_m and C_m' values indicates that the steady state and kinetic analyses for protein folding can be bridged by reaction relaxation time (τ) determination.

2D and 3D Models Correcting the Experiments. For lack of the crystal structure of SARS_CoV spike protein, 2D and 3D structural predictions may disclose some intrinsic structural features for the protein. Recently, Spiga et al. (*14*) have modeled 3D structures for both S1 and S2 subunits of the SARS_CoV spike glycoprotein based on the crystal structure of *Clostridium botulinum* neurotoxin B (PDB entries 1Q4Z and 1Q4Y for S1 and S2, respectively). However, the 3D model of Spiga et al. for S1 may be unreasonable. (i) They modeled the structure taking the S1 protein as an entire domain, but S1 consists of two major domains, as has been indicated in Figure 1. (ii) The percentages of α-helix and *β*-sheet encoded in their 3D model (PDB entry 1Q4Z) are not in agreement with the secondary structure prediction (Table 2) because their 3D model does not show that *β*-sheet dominates the secondary structure, their model containing only a few percent of α-helix and *β*-sheet. (iii) They modeled the structures of S1 and S2 on the basis of the hypothesis that CD13 is the binding receptor of the SARS_S protein, but recently it was demonstrated experimentally that ACE-2 is its binding receptor (*26*, *29*). This indicates that the crystal structure of *C. botulinum* neurotoxin B is not a good template for modeling the 3D structure for the S1 protein in general and for the S1b domain in particular. Accordingly, we divided the S1 protein into two domains, S1a and S1b (Figure 1), and modeled their structures separately. In this paper, we report the predicted 2D and 3D structures of SARS_S1b (Table 2 and Figures 7 and 8).

As listed in Table 2, the percentages of α-helix (3%) and *β*-sheet (35%) for SARS_S1b obtained from the secondary structure prediction in this study are extremely close to the data reported by Spiga et al. (*14*), although a different prediction method was used in this study. Meanwhile, the secondary structure was also estimated from the CD spectra (*43*). In comparison with the percentages of secondary structure components for SARS_S1b listed in Table 2, the obvious difference between the prediction and CD estimation lies on the α-helix component. We intend to accept the predicted data rather than CD spectral curve fitting data, because the estimation program for CD spectra is often weak when tested on polypeptides made up of one dominant secondary structure (*44*). In addition, the 3D model has also revealed that SARS_S1b is most likely to be an all-*β*-sheet globular protein, which coincides with the reported result that the coronavirus spike protein S1 subunit is the globular part of S protein (*20*). Tryptophan fluorescence and CD spectra (Figures 2 and 3) revealed that some of the tryptophans are located on the surface of the globular protein,




```
1AOF PPEFGMKEMRESWKVHVAPEDRPTQQMNDWDLENLFSVTLRDAGQIALIDGSTYEIKTVLDTGYAVHISRLSASGRYLFV   80
1NIR --EWGMPEMRESWKVLVKPEDRPKKQLNDLDLPNLFSVTLRDAGQIALVDGDSKKIVKVIDTGYAVHISRMSASGRYLLV   78
S1B  ----------------------------------AYFVGYLKPTTFMLKYDENGTITDAVD-----------CSQNPLAE   35

1AOF IGRDGKVNMIDLWMKEPTTVAEIKIGSEARSIETSKMEGWEDKYAIAGAYWPPQYVIMDGETLEPKKIQSTRGMTYDEQE   160
1NIR IGRDARIDMIDLWAKEPTKVAEIKIGIEARSVESSKFKGYEDRYTIAGAYWPPQFAIMDGETLEPKQIVSTRGMTVDTQT   158
S1B  LKCSVKSFEIDKGIYQTSNFRVVPSGDVVRFPNITNLCPFGEVFNATKFPSVYAWERKKISNCVADYSVLYNSTFFSTFK   115

1AOF YHPEPRVAAILASHYRPEFIVNVKETGKILLVDYTDLNNLKTTEISAERFLHDGGLDGSHRYFITAANARNKLVVIDTKE   240
1NIR YHPEPRVAAIIASHEHPEFIVNVKETGKVLLVNYKDIDNLTVTSIGAAPFLHDGGWDSSHRYFMTAANNSNKVAVIDSKD   238
S1B  CYGVSATKLNDLCFSNVYADSFVVKGDDVRQIAPGQTGVIADYNYKLPDDFMGCVLAWNTRNIDATSTGNYNYKYRYLRH   195

1AOF GKLVAIEDTGGQTPHPGRGANFVHPTFGPVWATSHMGDDSVALIGTDPEGHPDNAWKILDSFPALGGGSLFIKTHPNSQY   320
1NIR RRLSALVDVG-KTPHPGRGANFVHPKYGPVWSTSHLGDGSISLIGTDPKNHPQYAWKKVAELQGQGGGSLFIKTHPKSSH   317
S1B  GKLRPFERDISNVPFSPDGKPCTPPALNCYWPLNDYGFYTTTGIGYQPYRVVVLSFELLNAPATVCGPKLSTDLIKNQCV   275

1AOF LYVDATLNPEAEISGSVAVFDIKAMTGDGSDPEFKTLPIAEWAGITEGQPRVVQGEFNKDGTEVWFSVWNGKDQESALVV   400
1NIR LYVDTTFNPDARISQSVAVFDLKNLDAK-----YQVLPIAEWADLGEGAKRVVQPEYNKRGDEVWFSVWNGKNDSSALVV   392
S1B  NFNFNGLTGTGVLTPSSKRFQPFQQFGRDVSDFTDSVRDPKTSEILDISPCSFGGVSVITPGTNASSEVAVLYQDVNCTD   355

1AOF VDDKTLELKHVIKDERLVTPTGKFNVYNTMTDTY--   434
1NIR VDDKTLKLKAVVKDPRLITPTGKFNVYNTQHDVY--   426
S1B  VSTAIHADQLTPAWRIYSTGNNVFQTQAGCLIGAEH   391
```

b

FIGURE 7: (a) Secondary structure prediction of SARS_S1b. (b) Sequence alignment between SARS_S1b and the template proteins (PDB entries 1AOF and 1NIR).

which is in agreement with our 3D model (Figure 8). The consistency between the experimental results and the structural predictions enhanced the reliability of our 3D model of SARS_S1b.

To further verify the reliability of our 3D model of SARS_S1b, we mapped the structure feature of the model with the current experimental results (*29*, *30*). Wong et al. determined that a 193-amino acid fragment (residues 318−510) of the SARS_S1b domain is essential for the association between the S1 protein and ACE-2 (*29*). This fragment occupies approximately half the volume of the structure of SARS_S1b (Figure 8a,b, left part colored purple). In addition, He et al. found that the fragment consisting of residues 528−635 is a major immunodominant epitope; this part fills ~25% of the structural space of SARS_S1b (Figure 8a,b, right part colored red). Sequence alignment for the S1b proteins isolated from different strains of SARS-CoV shows that there are 14 mutation sites, i.e., Gly311, Lys344, Phe360, Arg426, Asn437, Tyr442, Arg444, Leu472, Asp480, Thr487, Phe501, Ser577, Ala609, and Asp613 (Figure S1 of the Supporting Information). Among these 14 residues, the 10 residues from Lys344 to Phe501 are positioned in the ACE-2 binding fragment and the last three residues are located in the major immunodominant site. It is interesting that most of the residues at mutation sites are located on the surface of our 3D model. This is why these resides can be altered; they are more likely to interact with other proteins (receptor or antibody). Moreover, Glu452Ala and Asp454Ala mutations interfere with or abolish binding of the S1 protein to ACE-2 (*29*). These two acidic residues are located on the surface of our 3D model, and their side chains are exposed in the solvent (Figure 8a,b). A primary SARS_S1b−ACE-2 docking simulation (data not shown) indicates that ACE-2 may probably interact with the patch around Asp454

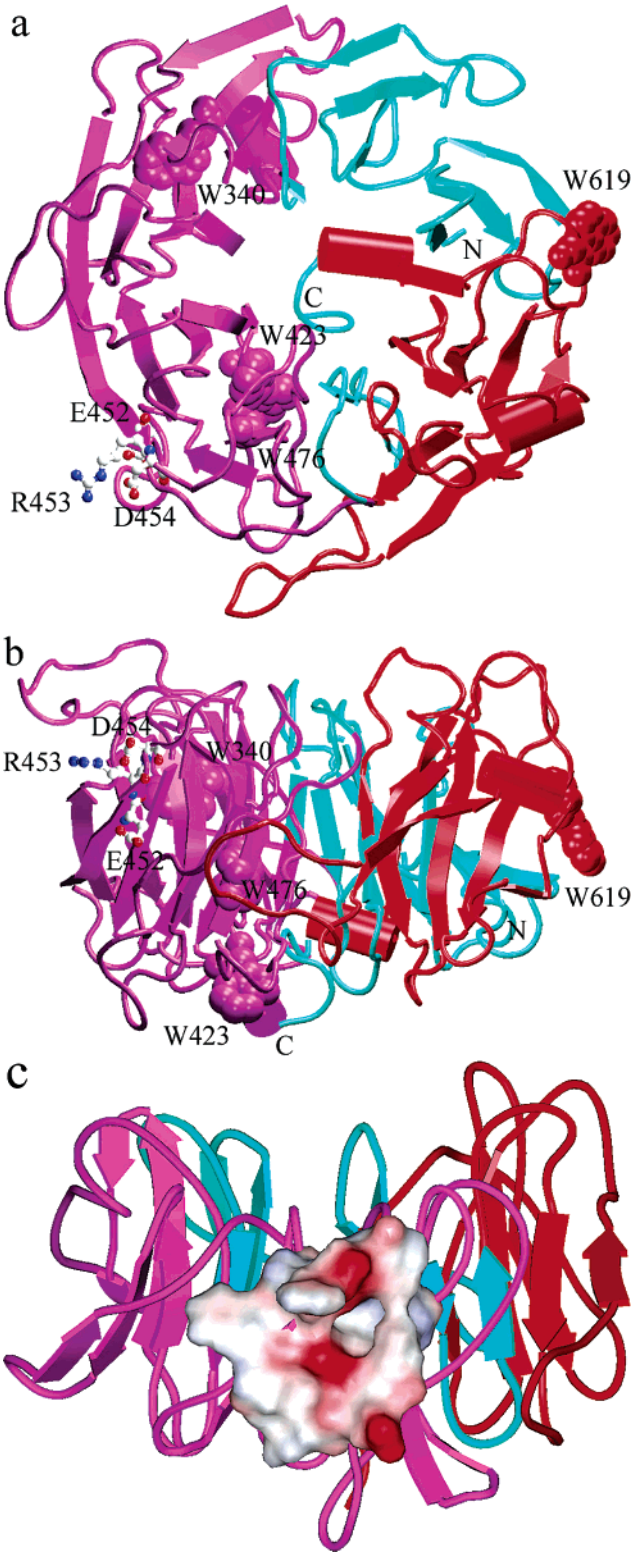


FIGURE 8: 3D model of the SARS_S1b domain that resulted from homology modeling. (a) Top view of the structure. (b) Side view of the structure. (c) Electrostatic surface for the possible ACE-2 binding site composed by Asn437, Tyr438, Pro450−Ile455, Thr487, Ile489, and Phe501−Leu503. The four tryptophans are represented as the CPK model, and the three residues (Glu452, Arg453, and Asp454) that are important for ACE-2 binding (*29*) are represented as a ball-and-stick model. The ACE-2 association part (residues 318−510) and the major immunodominant epitope (residues 528−635) are colored purple and red, respectively. N and C denote the N- and C-termini, respectively. Panels a and b were constructed and rendered using VMD (*45*) and Raster3D (*46*), respectively. Panel c was generated using WebLab ViewerPro (*47*).

composed by Asn437, Tyr438, Pro450−Ile455, Thr487, Ile489, and Phe501−Leu503 (Figure 8c). The Asp335−Ala342 loop of ACE-2 may fit into the binding patch, and the side chains of Glu452 and Asp454 of SARS_S1b form hydrogen bonds with Gln340 of ACE-2. This SARS_S1b−ACE-2 interaction model is in good agreement the mutagenesis result (*29*). These data demonstrate again the reliability of our 3D model, suggesting that it can be used for further biological study and drug discovery targeting the S protein.

SUPPORTING INFORMATION AVAILABLE

Sequence alignment for the variations of the SARS_S1b fragment and the locations of the mutation residues of the variations on the 3D model of SARS_S1b. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Poutanen, S. M., Low, D. E., Henry, B., Finkelstein, S., Rose, D., Green, K., Tellier, R., Draker, R., Adachi, D., Ayers, M., Chan, A. K., Skowronski, D. M., Salit, I., Simor, A. E., Slutsky, A. S., Doyle, P. W., Krajden, M., Brunham, R. C., and McGeer, A. J. (2003) Identification of Severe Acute Respiratory Syndrome in Canada, *N. Engl. J. Med. 348*, 1995−2005.
2. Lee, N., Hui, D., Wu, A., Chan, P., Cameron, P., Joynt, G. M., Ahuja, A., Yung, M. Y., Leung, C. B., To, K. F., Lui, S. F., Szeto, C. C., Chung, S., and Sung, J. J. (2003) A major outbreak of severe acute respiratory syndrome in Hong Kong, *N. Engl. J. Med. 348*, 1986−1994.
3. Tsang, K. W., Ho, P. L., Ooi, G. C., Yee, W. K., Wang, T., Chan-Yeung, M., Lam, W. K., Seto, W. H., Yam, L. Y., Cheung, T. M., Wong, P. C., Lam, B., Chan, J., Yuen, K. Y., and Lai, K. N. (2003) A cluster of cases of severe acute respiratory syndrome in Hong Kong, *N. Engl. J. Med. 348*, 1977−1985.
4. 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., and members of the SARS study group (2003) Coronavirus as a possible cause of severe acute respiratory syndrome, *Lancet 361*, 1319−1325.
5. Drosten, C., Preiser, W., van der Werf, S., Brodt, H. R., Becker, S., Rabenau, H., Panning, M., Kolesnikova, L., Fouchier, R. A., Berga, 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., and Doerr, H. W. (2003) Identification of a novel coronavirus in patients with severe acute respiratory syndrome, *N. Engl. J. Med. 348*, 1967−1976.
6. 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., Shilh, 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., and Anderson, L. J. (2003) A novel coronavirus associated with severe acute respiratory syndrome, *N. Engl. J. Med. 348*, 1953−1966.
7. Rota, P. A., Oberste, M. S., Monroe, S. S., Nix, W. A., Campagnoli, R., Icenogle, J. P., Penaranda, S., Bankamp, B., Maher, K., Chen, M., Tong, S., Tamin, A., Lowe, L., Frace, M., DeRisi, J. L., Chen, Q., Wang, D., Erdman, D. D., Peret, T. C., Burns, C., Ksiazek, T. G., Rollin, P. E., Sanchez, A., Liffick, S., Holloway, B., Limor, J., McCaustland, K., Olsen-Rasmussen, M., Fouchier, R., Günther, S., Osterhaus, D. M., Drosten, C., Pallansch, M. A., Anderson, L. J., and Bellini, W. J. (2003) Characterization of a novel coronavirus associated with severe acute respiratory syndrome, *Science 300*, 1394−1399.
8. Marra, M. A., Jones, S. J., Astell, C. R., Holt, R. A., Wilson, A. B., Butterfield, Y. S., Khattra, J., Asano, J. K., Barber, S. A., Chan, S. Y., Cloutier, A., Coughlin, S. M., Freeman, D., Girn, N., Griffith, O. L., Leach, S. R., Mayo, M., McDonald, H., Montgomery, S. B., Pandoh, P. K., Petrescu, A. S., G.Robertson, A., Schein, J. E., Siddiqui, A., Smailus, D. E., Stott, J. M., Yang, G. S., Plummer, F., Andonov, A., Artsob, H., Bastien, N., Bernard, K., Booth, T. F., Bowness, D., Czub, M., Drebot, M., Fernanado, L., Flick, R., Garbutt, M., Gray, M., Grolla, A., Jone, S.,

Feldmann, H., Meyers, A., Kabani, A., Li, Y., Normand, S., Stroher, U., Tipples, G. A., Tyler, S., Vogrig, R., Ward, D., Watson, B., Brunham, R. C., Krajden, M., Petric, M., Skowronski, D. M., Upton, C., and Roper, R. L. (2003) The Genome sequence of the SARS-associated coronavirus, *Science 300*, 1399−1404.

9. Shen, X., Xue, J. H., Yu, C. Y., Luo, H. B., Qin, L., Yu, X. J., Chen, J., Chen, L. L., Xiong, B., Yue, L. D., Cai, J. H., Shen, J. H., Luo, X. M., Chen, K. X., Shi, T. L., Li, Y. X., Hu, G. X., and Jiang, H. L. (2003) Small envelope protein E of SARS: Cloning, expression, purification, CD determination, and bioinformatics analysis, *Acta Pharmacol. Sin. 24*, 505−511.
10. Krokhin, O., Li, Y., Andonov, A., Feldmann, H., Flick, R., Jones, S., Stroeher, U., Bastien, N., Dasuri, K. V., Cheng, K., Simonsen, J. N., Perreault, H., Wilkines, J., Ens, W., Plummer, F., and Standing, K. G. (2003) Mass spectrometric characterization of proteins from the SARS virus: A preliminary report, *Mol. Cell. Proteomics 2*, 346−356.
11. Luo, C., Luo, H. B., Zheng, S. X., Gui, C. S., Yue, L. D., Yu, C. Y., Sun, T., He, P. L., Chen, J., Shen, J. H., Luo, X. M., Li, Y. X., Liu, H., Bai, D. L., Shen, J. K., Yang, Y. M., Li, F. Q., Zuo, J. P., Hilgenfeld, R., Pei, G., Chen, K. X., Shen, X., and Jiang, H. L. (2004) Nucleocapsid protein of SARS coronavirus tightly binds to human cyclophilin A, *Biochem. Biophys. Res. Commun. 321*, 557−565.
12. Yang, H., Yang, M., Ding, Y., Liu, Y., Lou, Z., Zhou, Z., Sun, L., Mo, L., Ye, S., Pang, H., Gao, G. F., Anand, K., Bartlam, M., Hilgenfeld, R., and Rao, Z. (2003) The crystal structures of severe acute respiratory syndrome virus main protease and its complex with an inhibitor, *Proc. Natl. Acad. Sci. U.S.A. 100*, 13190−13195.
13. Huang, Q., Yu, L., Petros, A. M., Gunasekera, A., Liu, Z., Xu, N., Hajduk, P., Mack, J., Fesik, S. W., and Olejniczak, E. T. (2004) Structure of the N-terminal RNA-binding domain of the SARS CoV nucleocapsid protein, *Biochemistry 43*, 6059−6063.
14. Spiga, O., Bernini, A., Ciutti, A., Chiellini, S., Menciassi, N., Finetti, F., Causarono, V., Anselmi, F., Proschi, F., and Niccolai, N. (2003) Molecular modelling of S1 and S2 subunits of SARS coronavirus spike glycoprotein, *Biochem. Biophys. Res. Commun. 310*, 78−83.
15. Anand, K., Ziebuhr, J., Wadhwani, P., Mesters, J. R., and Hilgenfeld, R. (2003) Coronavirus main proteinase ($3CL^{pro}$) structure: Basis for design of anti-SARS drugs, *Science 300*, 1763−1767.
16. Xiong, B., Gui, C. S., Xu, X. Y., Luo, C., Chen, J., Luo, H. B., Chen, L. L., Li, G. W., Sun, T., Yu, C. Y., Yue, L. D., Duan, W. H., Shen, J. K., Qin, L., Shi, T. L., Li, Y. X., Chen, K. X., Luo, X. M., Shen, X., Shen, J. H., and Jiang, H. L. (2003) A 3D model of SARS-CoV 3CL proteinase and its inhibitors design by virtual screening, *Acta Pharmacol. Sin. 24*, 497−504.
17. Von Grotthuss, M., Wyrwicz, L. S., and Rychlewski, L. (2003) mRNA cap-1 methyltransferase in the SARS genome, *Cell 113*, 701−702.
18. Cinatl, J., Morgenstern, B., Bauer, G., Chandra, P., Rabenau, H., and Doerr, H. W. (2003) Glycyrrhizin, an active component of liquorice roots, and replication of SARS-associated coronavirus, *Lancet 361*, 2045−2046.
19. Wu, C. Y., Jan, J. T., Ma, S. H., Kuo, C. J., Juan, H. F., Cheng, Y. S., Hsu, H. H., Huang, H. C., Wu, D., Brik, A., Liang, F. S., Liu, R. S., Fang, J. M., Chen, S. T., Liang, P. H., and Wong, C. H. (2004) Small molecules targeting severe acute respiratory syndrome human coronavirus, *Proc. Natl. Acad. Sci. U.S.A. 101*, 10012−10017.
20. Kubo, H., Takase-Yoden, S., and Taguchi, F. (1993) Neutralization and fusion inhibition activities of monoclonal antibodies specific for the S1 subunit of the spike protein of neurovirulent murine coronavirus JHMV c1-2 variant, *J. Gen. Virol. 74*, 1421−1425.
21. Suzuki, H., and Taguchi, F. (1996) Analysis of the receptor-binding site of murine coronavirus spike protein, *J. Virol. 70*, 2632−2636.
22. Collins, A. R., Knobler, R. L., Powell, H., and Buchmeier, M. J. (1982) Monoconal antibodies to murine hepatitis virus-4 (strain JHM) define the viral glycoprotein responsible for attachment and cell fusion, *Virology 119*, 358−371.
23. Hingley, S. T., Leparc-Goffart, I., and Weiss, S. R. (1998) The spike protein of murine coronavirus mouse hepatitis virus strain A59 is not cleaved in primary Glial cells and primary hepatocytes, *J. Virol. 72*, 1606−1609.
24. Arbour, N., Ekande, S., Lachance, C., and Chagnon, F. L. (1999) Persistent infenction of human oligodendrocytic and neuroglial cell lines by human coronavirus 229E, *Am. Soc. Microbiol. 73*, 3326−3327.
25. Bonavia, A., Zelus, B. D., Wentworth, D. E., Talbot, P. J., and Holmes, K. V. (2003) Identification of a receptor-binding domain of the spike glyocoprotein of human coronavirus HcoVE-229E, *J. Virol. 77*, 2530−2538.
26. Li, W., Moore, M. J., Vasilieva, N., Sui, J., Wong, S. K., Berne, M. A., Somasundaran, M., Sullivan, J. L., Luzuriaga, K., Greenough, T. C., Choe, H., and Farzan, M. (2003) Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus, *Nature 426*, 450−454.
27. John, Z., Snijder, E. J., and Gorbalenya, A. E. (2000) Virus-encoded proteinases and proteolytic processing in the Nidovirales, *J. Gen. Virol. 81*, 853−879.
28. Wang, C. H., Hong, C. C., and Seak, C. H. (2002) An ELISA for antibodies against infectious bronchitis virus using an S1 spike polypeptide, *Vet. Microbiol. 85*, 333−342.
29. Wong, S. K., Li, W., Moore, M. J., Choe, H., and Farzan, M. (2004) A 193-amino acid fragment of the SARS coronavirus S protein efficiently binds angiotensin-converting enzyme 2, *J. Biol. Chem. 279*, 3197−3201.
30. Pohl-Koppe, A., Raabe, T., Siddell, S. G., and Meulen, V. T. (1995) Detection of human coronavirus 229E-specific antibodies using recombinant fusion proteins, *J. Virol. Methods 55*, 175−183.
31. Sui, J., Li, W., Murakami, A., Tamin, A., Matthews, L. J., Wong, S. K., Moore, M. J., Tallarico, S. C., Olurinde, M., Choe, H., Anderson, L. J., Bellini, W. J., Farzan, M., and Marasco, W. A. (2004) Potent neutralization of severe acute repiratory syndrome (SARS) coronavirus by a human mAb to S1 protein that blocks receptor association, *Proc. Natl. Acad. Sci. U.S.A. 101*, 2536−2541.
32. He, Y., Zhou, Y., Wu, H., Luo, B., Chen, J., Li, W., and Jiang, S. (2004) Identification of immunodominant sites on the spike protein of severe acute respiratory syndrome (SARS) coronavirus: Implication for developing SARS diagnostics and vaccines, *J. Immunol. 173*, 4050−4057.
33. Yu, H., Yang, Y., Zhang, W., Xie, Y. H., Qin, J., Wang, Y., Zheng, H. B., Zhao, G. P., Yang, S., and Jiang, W. H. (2003) Expression and Purification of Recombinant SARS Coronavirus Spike Protein, *Acta Biochim. Biophys. Sin. 35*, 774−778.
34. Sun, H., Luo, H., Yu, C., Sun, T., Chen, J., Peng, S., Qin, J., Shen, J., Yang, Y., Xie, Y., Chen, K., Wang, Y., Shen, X., and Jiang, H. (2003) Molecular cloning, expression, purification, and mass spectrometric characterization of 3C-like protease of SARS coronavirus, *Protein Expression Purif. 32*, 302−308.
35. Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice, *Nucleic Acids Res. 22*, 4673−4680.
36. Sali, A., and Blundell, T. L. (1993) Comparative protein modelling by satisfaction of spatial restraints, *J. Mol. Biol. 234*, 779−815.
37. Laskowski, R. A., Macarthur, M. W., Moss, D. S., and Thornton, J. M. (1993) PROCHECK: A program to check the stereochemical quality of protein structures, *J. Appl. Crystallogr. 26*, 283−291.
38. Williams, P. A., Fulop, V., Garman, E. F., Saunders, N. F., Ferguson, S. J., and Hajdu, J. (1997) Haem-ligand switching during catalysis in crystals of a nitrogen-cycle enzyme, *Nature 389*, 406−412.
39. Nurizzo, D., Silvestrini, M. C., Mathieu, M., Cutruzzola, F., Bourgeois, D., Fulop, V., Hajdu, J., Brunori, M., Tegoni, M., and Cambillau, C. (1997) N-Terminal arm exchange is observed in the 2.15 Å crystal structure of oxidized nitrite reductase from *Pseudomonas aeruginosa*, *Structure 5*, 1157−1171.
40. Tell-slis, S. R., and Romero, G. B. (2001) Thermal denaturation of porcine pepsin: A study by circular dichrosim, *Int. J. Biol. Macromol. 28*, 129−133.
41. Pace, C. N. (1990) Conformational stability of globular proteins, *Trends Biochem. Sci. 15*, 14−17.
42. Deshpande, R. A., Khan, M. I., and Shankar, V. (2003) Equilibrium unfolding of RNase Rs from *Rhizopus stolonifer*. pH dependence of chemial and thermal denaturation, *Biochim. Biophys. Acta 1648*, 184−194.
43. Greenfield, N., and Fasman, G. D. (1969) Computed circular dichroism spectra for the evaluation of protein conformation, *Biochemistry 8*, 4108−4116.
44. Greenfield, N. J. (1996) Methods to estimate the conformation of proteins and polypetides from circular dichroism data, *Anal. Biochem. 235*, 1−10.

45. Humphrey, W., Dalke, A., and Schulten, K. (1996) VMD: Visual Molecular Dynamics, *J. Mol. Graphics 14*, 33−38.
46. Merritt, E. A., and Bacon, D. J. (1997) Raster3D: Photorealistic Molecular Graphics, *Methods Enzymol. 277*, 505−524.
47. *WebLab ViewerPro*, version 4.0 (2000) Molecular Simulations, Inc., San Diego.

BI0482396