Mutation of Gly-11 on the Dimer Interface Results in the Complete Crystallographic Dimer Dissociation of Severe Acute Respiratory Syndrome Coronavirus 3C-like Protease
CRYSTAL STRUCTURE WITH MOLECULAR DYNAMICS SIMULATIONS*⒮

Shuai Chen[‡1], **Tiancen Hu**[‡1], **Jian Zhang**[‡1], **Jing Chen**[‡], **Kaixian Chen**[‡], **Jianping Ding**[§], **Hualiang Jiang**[‡2], and **Xu Shen**[‡3]

From the [‡]*Drug Discovery and Design Center, State Key Laboratory of Drug Research, Shanghai Institute of Materia Medica, Chinese Academy of Sciences, 555 Zuchongzhi Road, Pudong, Shanghai 201203 and the* [§]*State Key Laboratory of Molecular Biology, Institute of Biochemistry and Cell Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai 200031, China*

SARS-CoV 3C-like protease (3CL^pro) is an attractive target for anti- severe acute respiratory syndrome (SARS) drug discovery, and its dimerization has been extensively proved to be indispensable for enzymatic activity. However, the reason why the dissociated monomer is inactive still remains unclear due to the absence of the monomer structure. In this study, we showed that mutation of the dimer-interface residue Gly-11 to alanine entirely abolished the activity of SARS-CoV 3CL^pro. Subsequently, we determined the crystal structure of this mutant and discovered a complete crystallographic dimer dissociation of SARS-CoV 3CL^pro. The mutation might shorten the α-helix A′ of domain I and cause a mis-oriented N-terminal finger that could not correctly squeeze into the pocket of another monomer during dimerization, thus destabilizing the dimer structure. Several structural features essential for catalysis and substrate recognition are severely impaired in the G11A monomer. Moreover, domain III rotates dramatically against the chymotrypsin fold compared with the dimer, from which we proposed a putative dimerization model for SARS-CoV 3CL^pro. As the first reported monomer structure for SARS-CoV 3CL^pro, the crystal structure of G11A mutant might provide insight into the dimerization mechanism of the protease and supply direct structural evidence for the incompetence of the dissociated monomer.

Severe acute respiratory syndrome (SARS)[4] is a highly infectious disease and has indeed ever been a severe threat to the worldwide population from the end of 2002 to June of 2003. As a positive-sense single strand RNA virus, SARS coronavirus (SARS-CoV) has been identified as the etiological agent responsible for SARS infection (1, 2). The genome of SARS-CoV contains 14 functional open reading frames, and two large 5′-terminal open reading frames, 1a and 1b, encode two overlapping polyproteins, pp1a and pp1ab, necessary for viral RNA replication and transcription. pp1a and pp1ab can be cleaved extensively by 3C-like protease (3CL^pro) and papain-like cysteine protease (PL2^pro) to yield a multisubunit protein complex called "viral replicase-transcriptase" (3). The functional indispensability of 3CL^pro in the SARS-CoV life-cycle has made it an attractive target in discovering new anti-SARS agents (4).

SARS-CoV 3CL^pro forms a dimer in the crystal with two monomers oriented perpendicularly to one another (5). Each monomer contains three domains: domains I and II form a chymotrypsin fold, and domain III is a globular cluster of five α-helices connecting to domain II by a long loop region. The 16-residue loop region has been implicated to involve in substrate binding (6). Current experimental results all indicated that only the dimer is the biological functional form of SARS-CoV 3CL^pro, and the dimerization-activity relationship of the protease has been extensively characterized (7–12). Because dimerization is convincingly proven to be indispensable for enzymatic activity, the dimer interface of SARS-CoV 3CL^pro has thus been suggested as another potential target for rational inhibitors design. As revealed by the crystal structure and molecular dynamics simulations (5, 13), the dimer interface of SARS-CoV 3CL^pro mainly involves: 1) interactions between the two helical domain III of each monomer and 2) the hydrogen bonding and electrostatic interactions between the N-terminal finger (residues 1–7) of one monomer and residues near the S1 substrate-binding subsite of the other monomer.

To date, the contributions of the residues on the dimer interface to SARS-CoV 3CL^{pro} dimerization and enzymatic activity have been identified by several groups (10, 14, 15). The N-terminal finger of the protease is considered to play an important role in both dimerization and activity. Hsu *et al.* (10) reported that the fourth residue on the N-terminal finger (Arg-4) is vital for stabilizing the dimer structure to give a correct conformation of the active site. An octapeptide interface inhibitor, designed according to the sequence of the N-terminal finger, was also found to bind with the protease specifically and prevent its dimerization (16), further supporting the importance of the N-terminal finger on maintaining the dimer state of SARS-CoV 3CL^{pro}. Furthermore, the residues of domain III were revealed to extensively mediate monomer-monomer interactions and be responsible for positioning the N-terminal finger of one monomer to interact with the active site of the other monomer (6, 14). In addition, we observed that the α-helix A′ of domain I (residues 10–15) is also a critical part of the dimer interface and contributes well to the dimer stability and catalytic activity of SARS-CoV 3CL^{pro}.[5]

Coronavirus 3CL^{pro} exists as a dimeric form in all solved crystal structures (4, 5, 9, 13, 17), and dimerization is also observed in solution with a concentration-dependent manner (8, 12, 18). In this study, we performed site-directed mutation of the conserved residue Gly-11 on the α-helix A′ and then evaluated the effect of the mutation on the folding manner and catalytic activity of SARS-CoV 3CL^{pro}. Subsequently, the crystal structure of the Gly-11 $\rightarrow$ Ala mutant was successfully solved and analyzed in conjunction with molecular dynamics simulations, which reveals that the mutant protease exists only as a monomer in crystal and mutation of Gly-11 on the dimer interface could result in the complete crystallographic dimer dissociation. Although the conclusion that the dimer is the functional form of SARS-CoV 3CL^{pro} has been confirmed by various experiments, why the dissociated monomer is inactive still remains unknown due to the absence of the structural information of the monomer. As the first reported structure of a dissociated monomer, the G11A mutant is expected to shed more light on understanding the dimerization and catalytic mechanism of SARS-CoV 3CL^{pro}.

EXPERIMENTAL PROCEDURES

Cloning, Expression, and Purification—The coding sequence of the wild-type 3CL^{pro} was cloned from SARS-CoV Tor2 strain (19) and inserted into the BamHI and XhoI sites of the plasmid pGEX4T-1 (Amersham Biosciences). Mutation of Gly-11 to Ala was performed with the QuikChange site-directed mutagenesis kit (Stratagene) using pGEX4T-1-SARS-CoV 3CL^{pro} as a template. The nucleotide sequences of the primers used for single point mutation were: 5′-GGCATTCCCGT-CAGCTAAAGTTGAAGGGTGC-3′ (forward) and 5′-GCAC-CCTTCAACTTTAGCTGACGGGAATGCC-3′ (reverse).

The resulting plasmids were verified by sequencing and then transformed into *Escherichia coli* BL21(DE3) cells. The cells were cultured in 500 ml of LB medium containing 100 μg/ml

ampicillin. When the absorbance of the LB medium at 600 nm (A_{600}) reached 0.8, isopropyl 1-thio-β-D-galactopyranoside was added to 0.5 mM, and the cell cultures were incubated at 25 °C for 6 h. After harvesting by centrifugation at 8,000 rpm, the pellet was resuspended in phosphate-buffered saline buffer (140 mM NaCl, 2.7 mM KCl, 10 mM Na$_2$HPO$_4$, 1.8 mM KH$_2$PO$_4$, pH 7.4) and sonicated on ice. The lysate was centrifuged at 15,000 rpm for 30 min, and the supernatant was loaded onto 2 ml of a glutathione-Sepharose 4B affinity column (Amersham Biosciences) and washed with 20 column volumes of phosphate-buffered saline. After that, 10 ml of the reduced glutathione (50 mM) was added onto the column to elute the glutathione *S*-transferase fusion proteins. The glutathione *S*-transferase fusion proteins were subsequently cleaved by thrombin (50 units) at 25 °C for 6 h and dialyzed against phosphate-buffered saline to remove the glutathione. Afterward, the recombinant protein was reloaded onto the glutathione-Sepharose column to remove the free glutathione *S*-transferase tag, and the flow-through fractions were dialyzed against buffer A (20 mM Tris-HCl, pH 7.5, 1 mM EDTA, 5 mM DTT). The dialyzed samples were then loaded onto an 8-ml Mono Q 10/100 GL column (Amersham Biosciences) pre-equilibrated with buffer A. The column was washed, and then eluted with a NaCl gradient (0–1.0 M) in buffer A. Peak fractions were analyzed by SDS-PAGE and those containing SARS-CoV 3CL^{pro} Gly-11 $\rightarrow$ Ala mutant were pooled, and then concentrated with a Centricon concentrator (Millipore). The protein concentration was determined by the absorbance at 280 nm (A_{280}) using a molar extinction coefficient (ϵ_{280}) for the monomer of 34390 M^{-1} cm^{-1}. Finally, the purified and concentrated proteins (10 mg/ml) were dialyzed against 10 mM Tris-HCl, pH 7.5, 1 mM EDTA, 5 mM DTT and stored at −20 °C.

CD Spectroscopy—Far-UV CD spectra were recorded on a JASCO-810 spectropolarimeter, and the protein samples were prepared in 10 mM sodium phosphate, pH 7.5, 100 mM NaCl at 25 °C with concentration of 10 μM. CD spectra from 190 to 250 nm were collected in 1-nm bandwidth using 0.1-cm path length cuvette and normalized by subtracting the baseline of the buffer. Each measurement was repeated for three times, and the final result was the average of three independent scans.

Fluorescence Spectroscopy—The fluorescence experiments were performed on a Hitachi F-2500 fluorescence spectrophotometer, and the protease samples were prepared in 10 mM Tris-HCl, pH 7.5, 100 mM NaCl with a concentration of 3–5 μM. Upon excitation at 280 nm at 25 °C, the fluorescence emission spectra of the samples were collected from 300 to 380 nm in a 1-ml quartz cuvette with 1-cm path length, and the spectral slit width was 5 nm for excitation and emission. The final spectra were corrected for the buffer contribution and averaged from three parallel measurements.

Enzymatic Activity Assay—The catalytic activities of the Gly-11 $\rightarrow$ Ala mutant and wild-type SARS-CoV 3CL^{pro} were measured according to our published method (20) by using a 12-amino acid fluorogenic substrate, EDANS-VNSTLQSGL-RK(Dabcyl)-M. Enzymatic activity was the average of three parallel assays, and the activity of the wild-type SARS-CoV 3CL^{pro} was taken as 100%.

Crystallization and Data Collection—SARS-CoV 3CL^pro Gly-11 → Ala mutant was crystallized by the hanging drop, vapor diffusion method at 4 °C, and the initial crystals grew in the published conditions (5). After further optimization of the crystallization conditions, the best plate-like crystals were obtained from 0.1 M Mes, pH 6.2, 10% polyethylene glycol 6000, 1 mM DTT, 5% Me$_2$SO, with a protein concentration of 10 mg/ml.

Diffraction data were collected in-house on a Rigaku rotating-anode x-ray generator operated at 100 kV and 100 mA (λ = 1.5418 Å). Diffraction images were recorded by a Rigaku R-AXIS IV^{++} imaging-plate detector with an oscillation step of 1°. The crystal was harvested with a nylon loop and flash-cooled in liquid nitrogen. Data collection was performed at 100 K in the cryoprotectant solution containing 30% of glycerol and 70% of the original reservoir solution. The data were indexed, integrated, and scaled using the program suite CrystalClear. Analysis of the diffraction data indicated that the crystal belongs to space group P2$_1$2$_1$2$_1$. Matthews coefficient suggested the presence of only one monomer in an asymmetric unit, corresponding to a V_{max} of 2.03 Å^3/Da and a solvent content of 40% (21).

Structure Determination, Refinement, and Model Building—Molecular replacement was carried out with Molrep (22) of the CCP4 program suite (23), using the chain A of wild-type SARS-CoV 3CL^pro (PDB code 1UJ1) as the search model. At first, we failed to place the entire protomer correctly in the unit cell. After some trials, we found that the chymotrypsin fold (domain I and II) and domain III must be located separately to obtain the successful solution. This result was in agreement with our later finding that the positional relationship of these two parts in the Gly-11 → Ala mutant has changed dramatically comparing with the wild-type protease.

Model refinement was initially performed using CNS (24), including rigid body, simulated annealing, minimization, and *B*-factor refinement. Coot and refmac5 (25) were subsequently employed for iterative cycles of model building and refinement. The water molecules were picked by inspecting the $>3\sigma F_o - F_c$ difference map. In the later stage, O (26) and CNS were used to produce the final model. The determined structure of the Gly-11 → Ala mutant consists of residues 4–299, except residues 155, 222–224, and 277–279 missing in the electron density. The geometry of the model was validated by Procheck (27). Refinement statistics are summarized in Table 1.

The superposition of various structural models of SARS-CoV 3CL^pro mentioned under "Results" was performed by LSQKAB (28) of CCP4, and the r.m.s.d. values were calculated by CNS. The interface between domain III and the rest part of a specific model were determined by the EBI PISA web server. The structural figures were all produced by Pymol.[6]

Data Bank with Accession Number—Coordinates and structure factors for SARS-CoV 3CL^pro Gly-11 → Ala mutant have been deposited in the Protein Data Bank with accession number of 2PWX.

Molecular Dynamics Simulations—The crystal structures of Gly-11 → Ala monomer and a series of reported dimers (PDB codes: 1UK3 and 1UJ1) were taken as the starting points for molecular dynamics (MD) simulations, using the AMBER suite of programs (version 8.0) with the parm99 force field (30). Each structure was prepared by using the xLeap module in AMBER, which involves adding protons to the structure, aligning the principle axes of the protein with the Cartesian axes of the box. All ionizable side chains were maintained in their standard protonation states at pH 7.0. The proteins were solvated in a truncated octahedron box of TIP3P water molecules, with the water thickness extending at least 10 Å apart from the protein surface. To avoid the instability that might occur during the MD simulations, the solvated system was subjected to minimization for 5000 cycles with protein restrained and followed by another 5000 cycles with the whole system relaxed. Then, the system was gradually heated from 0 to 300 K during the first 60 ps by three intervals, followed by equilibrium for 80 ps under constant volume and temperature condition. Afterward, the system was switched to constant pressure and temperature condition and equilibrated for 100 ps to adjust the system to a correct density. During the heating and equilibrating process, harmonic positional restraints were imposed on the protein atoms to allow the solvent to equilibrate around the protein, and the force constants of the positional restraints were gradually reduced in magnitude. Finally, the production simulations were carried out in the absence of any restraint under constant pressure and temperature, and data were collected. The protocol was applied to all simulation systems.

All the MD simulations were performed using the parallel version of PMEMD in AMBER suit. The particle mesh Ewald method was employed to calculate the long range electrostatic interactions, whereas the lengths of the bonds involving hydrogen atoms were fixed with the SHAKE algorithm (31, 32). During the simulations, the integration time step of 2 fs was adopted, and structural snapshots were flushed every 500 steps (1 ps). The non-bonded cutoff was set to 10.0 Å, and the non-bonded pair list was updated every 25 steps. Each production simulation was coupled to a 300 K thermal bath at 1.0 atm pressure by applying the Berendsen algorithm (33). The temperature and pressure coupling constants were set to 2.0 and 1.0 ps, respectively. All the MD simulations were run on Origin3800.

RESULTS

Preparation of SARS-CoV 3CL^pro Gly-11 → Ala Mutant—As has been reported (5, 13), the dimer interface of SARS-CoV 3CL^pro mainly consists of extensive interactions between two helical domain III of each monomer, as well as the hydrogen bonding and electrostatic interactions between the N-terminal residues of one monomer and residues near the substrate-binding subsite S1 of the other monomer. Besides the well characterized N-terminal finger and Domain III, the α-helix A' of domain I (residues 10–15) can also mediate the monomer-monomer interactions (supplemental Fig. S1) and play an important role in both dimerization and activity of SARS-CoV 3CL^pro.[5] Guided by this valuable information, we selected the residue Gly-11 on the α-helix A' for site-directed mutagenesis and then evaluated the catalytic activity and structural features of the Gly-11 → Ala mutant.
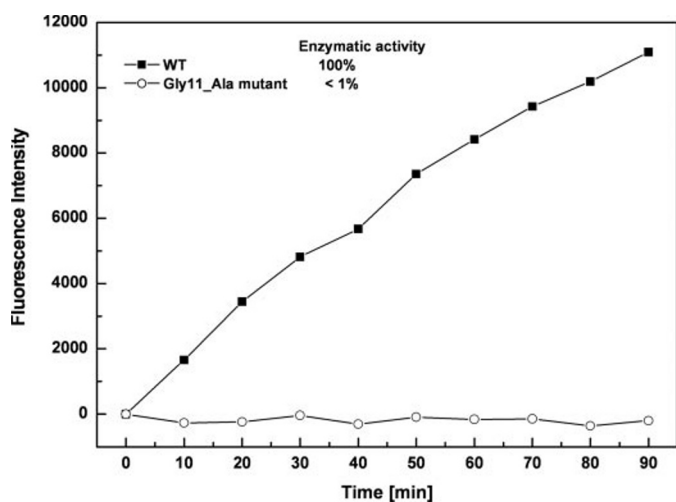

FIGURE 1. **Enzymatic activities of the Gly-11 → Ala mutant and wild-type SARS-CoV 3CL^{pro}.** The fluorogenic substrate at a concentration of 10 μM was incubated with 1 μM Gly-11 → Ala mutant or wild-type SARS-CoV 3CL^{pro} in 10 mM Tris-HCl, pH 7.5, 100 mM NaCl, 5 mM DTT, 1 mM EDTA, at 25 °C. Increase of emission fluorescence intensity at 488 nm wavelength was recorded at 10-min intervals, λ_{EX} = 340 nm. The emission spectrum was recorded for 90 min, and the activity of wild-type SARS 3CL^{pro} was taken as 100%.

TABLE 1

Crystallographic data collection and refinement statistics

Data collection statistics	
Space group	P2$_1$2$_1$2$_1$
Unit cell dimensions	
a, b, and c (Å)	34.147, 66.052, 129.030
a, b, and r (deg)	90.000, 90.000, 90.000
Resolution range (Å)	19.66–2.50 (2.59–2.50)[a]
No. of total reflections	48,386 (4,890)
No. of unique reflections	10,537 (1,038)
Redundancy	4.59 (4.71)
R_{sym}[b]	0.144 (0.306)
$\langle I/\sigma(I) \rangle$	4.2 (1.9)
Completeness	98.9%
Refinement statistics	
R-factor[c]	0.243
Free R-factor[c]	0.295
Number of residues	289
Number of protein atoms	2,240
Number of water molecules	83
Average B-factor of all atoms (Å^2)	
Protein main chain atoms	32.869
Protein side chain atoms	34.130
Water molecules	24.890
r.m.s.d. bond lengths (Å)[d]	0.007
r.m.s.d. bond angles (°)[d]	1.32
Ramachandran plot (%)	
Most favored regions	85.0
Allowed regions	13.8
Generously allowed regions	1.2
Disallowed regions	0

[a] Numbers in the parentheses represent statistics in the highest resolution shell.

[b] $R_{sym} = \Sigma_h \Sigma_i |I_{hi} - \langle I_h \rangle| / \Sigma_h \Sigma_i I_{hi}$, where I_{hi} and $\langle I_h \rangle$ are the ith and mean measurement of the intensity of reflection h, respectively.

[c] $R_{work} = \Sigma_h |F_{oh} - F_{ch}| / \Sigma_h F_{oh}$, where F_{oh} and F_{ch} are the observed and calculated structure factor amplitudes, respectively, and the summation is over 95% of the reflections in the specified resolution range. The remaining 5% of the reflections are randomly selected before the structure refinement and not included in the structure refinement. R_{free} is calculated over these reflections using the same equation as for R_{work}.

[d] r.m.s.d. bond lengths are the root-mean-square deviation from the parameter set for ideal stereochemistry.

Supplemental Fig. S2a shows the Far-UV CD spectra of the Gly-11 → Ala mutant and wild-type SARS-CoV 3CL^{pro}. The spectrum of the Gly-11 → Ala mutant is similar to that of the wild-type protease, indicating that the Gly-11 → Ala mutant also has well defined secondary structures and excluding the possibility of structural misfolding caused by mutation of Gly-11. In addition, we also measured the fluorescence emission spectra of the Gly-11 → Ala mutant and wild-type protease with results in supplemental Fig. S2b. Both the emission λ_{max} values of the wild-type protease and the Gly-11 → Ala mutant are ~325 nm, and such similar emission λ_{max} further demonstrates that replacement of Gly-11 on the dimer interface by alanine has not changed the folding manner of SARS-CoV 3CL^{pro}.

Enzymatic Activity of SARS-CoV 3CL^{pro} Gly-11 → Ala Mutant—SARS-CoV 3CL^{pro} has been reported to form a dimer in the crystal structure and exist as a mixture of monomer and dimer in solution (5, 8). It has been considered that only the dimer is the biological function form of SARS-CoV 3CL^{pro}, and the dissociated monomer might be enzymatic inactive (5, 34). Because the α-helix A′ of domain I (residues 10–15) is also revealed to involve in formation of the dimer interface (supplemental Fig. S1), mutation of Gly-11 might possibly disrupt the catalytic activity of SARS-CoV 3CL^{pro}. To verify this prediction, we determined the enzymatic activity of the Gly-11 → Ala mutant by a 12-amino acid fluorogenic substrate (20). As shown in Fig. 1, the fluorescence significantly increase with the hydrolysis of the substrate by the wild-type protease in a time-dependent manner, whereas the fluorescence profile of the Gly-11 → Ala mutant is obviously different from that of the wild-type protease, indicating that mutation of Gly-11 on the α-helix A′ could result in a complete loss of the catalytic activity. Such a result further supports the possibility that the monomer-monomer interactions regulated by the residue Gly-11 might stabilize the dimeric structure of SARS-CoV 3CL^{pro}, which is vital for maintaining the full enzymatic activity.

Overall Structure of SARS-CoV 3CL^{pro} Gly-11 → Ala Mutant—To accurately evaluate the contribution of the residue Gly-11 to 3CL^{pro} dimerization, the crystal structure of the Gly-11 → Ala mutant was subsequently analyzed in conjunction with molecular dynamics simulations. To date, the reported crystal structures of SARS-CoV 3CL^{pro} are all in dimeric form (5, 13, 17, 35, 36). However, the solved crystal structure of Gly-11 → Ala mutant in this work clearly revealed a monomer state. The crystallographic statistics of the structure are summarized in Table 1. There is only one molecule in the asymmetric unit of the crystal cell, and the P2$_1$2$_1$2$_1$ space group of the crystal also allows no 2-fold axis related dimer to be reconstituted by crystallographic symmetry. In addition, we observed that the crystal contact sites mainly involve the regions away from the dimer interface (residues 120–170) (37). These results thus indicated that the crystallographic dissociation of the dimer should not be an artifact of crystal packing but caused by Gly-11 mutation-induced conformational changes. To further test whether the monomer structure of the Gly-11 → Ala mutant is stable in solution, we conducted a 4-ns general MD simulation on the crystal structure for probing its behavior in the solvent environment. Fig. 2a showed the root-mean-square deviation (r.m.s.d.) from the crystal structure of all atoms *versus* simulation time. After ~0.5 ns, the r.m.s.d. of the simulation system tends to be convergent, indicating that the crystal structure might be stable and the system has been equilibrated well.
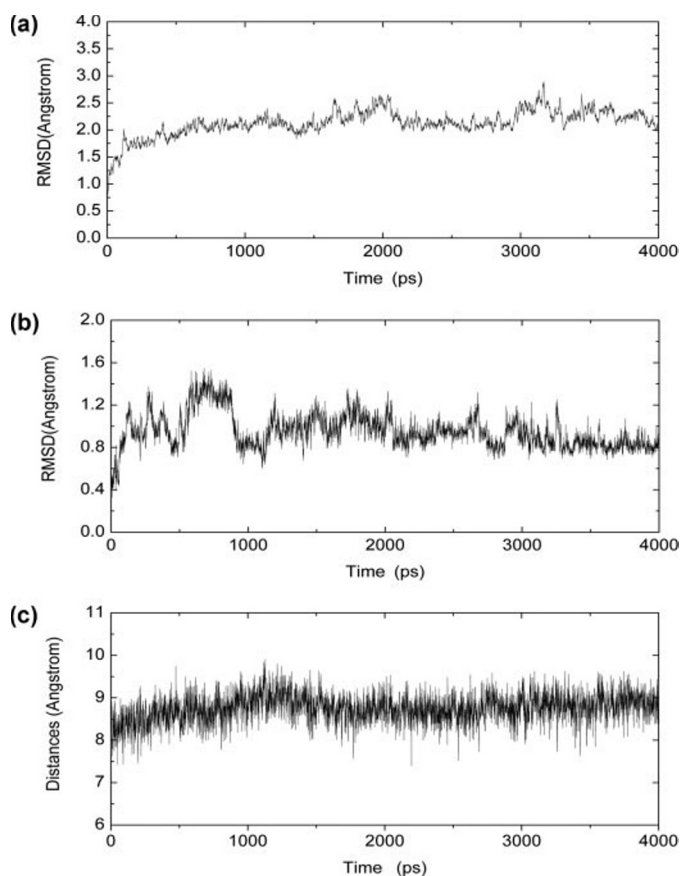
FIGURE 2. *a*, the r.m.s.d. of G11A monomer relative to the initial crystal structure during 4-ns MD simulation process. *b*, the r.m.s.d. of residues Phe-140 to Cys-145 in the S1 substrate-binding subsite of G11A monomer relative to the initial crystal structure during 4-ns MD simulation. *c*, time dependence of the centroid distance between residues Phe-140 and His-163 in G11A monomer.
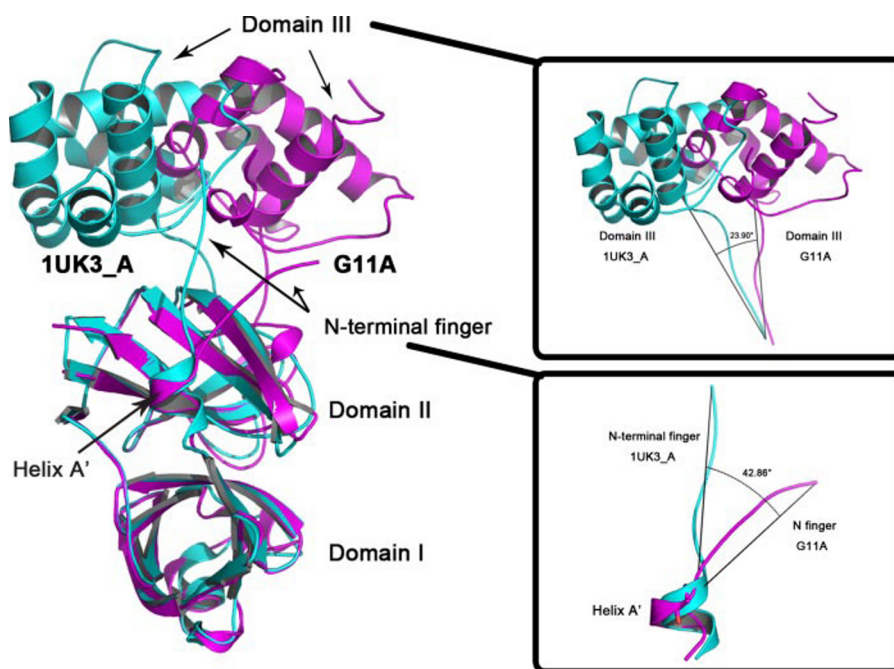

FIGURE 3. **Overall structure comparison between G11A monomer and the active protomer 1UK3_A in SARS-CoV 3CL^pro^ dimer.** The proteins are shown as schematics, and the structural elements are labeled. G11A and 1UK3_A are colored in *magenta* and *cyan*, respectively. The two *boxed figures* specially illustrate the structural comparison of domain III and the N-terminal finger, and the relative angles are also indicated.

*Structural Comparison of the Gly-11 → Ala Mutant with Wild-type SARS-CoV 3CL^pro^*—To better elucidate the conformational changes induced by Gly-11 → Ala mutation, we performed structural comparison of the Gly-11 → Ala mutant (G11A for short) with wild-type SARS-CoV 3CL^pro^ (5), including the protomer A in dimer at pH 8.0 (PDB codes: 1UK3 and 1UK3_A, for short, representative of the active form) and the protomer B in dimer at pH 6.0 (PDB codes: 1UJ1 and 1UJ1_B, for short, representative of the inactive form).

Similar to the protomers in wild-type SARS-CoV 3CL^pro^ dimer, the crystal structure of G11A monomer is still composed of three domains. However, the interdomain arrangement of G11A monomer has changed dramatically. To differentiate the slight but important changes of the active site conformation, we compared these monomeric molecules mainly based on the superposition of domains I and II. As shown in Fig. 3, compared with 1UK3_A, the most obvious conformational changes of G11A monomer are the rotation of domain III (tuned by about 24°, measured by the angle among Cα atoms of Thr201_1UK3_A, Thr190_1UK3_A, and Thr201_G11A) and the different orientation of the N-terminal finger (turned by about 43°, measured by the angle among Cα atoms of Phe3_1UK3_A, Gly11_1UK3_A, and Arg4_G11A), which might explain the difficulties we met in molecular replacement (see "Experimental Procedures"). The global conformation of the catalytic domains I and II has not changed significantly but the substrate-binding pocket between these two domains has altered considerably in G11A monomer, especially at the S1 subsite. After superposing domains I and II, the r.m.s.d. values between the G11A monomer and the two different wild-type protomers are shown in Table 2. The results further supported the existence of large conformational differences of domain III and the N-terminal finger among these monomeric molecules. In addition, the long loop region connecting domains II and III, particularly residues 193–200, also orients differently in G11A monomer.

Catalytic Dyad—Lai *et al.* (38) demonstrated that SARS-CoV 3CL^pro^ undergoes a general serine protease catalysis mechanism, and the residues His-41 and Cys-145 have been identified as the catalytic dyad. It has been suggested that hydrogen bond formation between His-41 NE2/ND1 and Cys-145 SG could indicate the right conformation of the catalytic dyad (39). Therefore, we monitored the distance between His-41 and Cys-145 and the time occupancy of the hydrogen bond His-41 to Cys-145 in G11A monomer during the whole 4-ns MD simulation process (supplemental Fig. S3 and Table 3). The results revealed that the distance between His-41 NE2 and Cys-145

TABLE 2

The r.m.s.d. between G11A monomer and two different protomers in wild type SARS-CoV 3CL^pro dimers

Wild-type protomer	r.m.s.d.[a]				
	Overall structure	N-terminal finger (residues 4–10)	Domains I and II (residues 12–184)	Long loop region (residues 185–200)	Domain III (residues 201–299)
	Å				
1UK3_A	10.94	8.84	1.88	5.67	18.73
1UJ1_B	9.93	7.96	1.71	5.50	16.99

[a] The r.m.s.d. values were obtained by superposing domains I and II of G11A monomer with two different protomers in wild-type dimers, respectively.

TABLE 3

Hydrogen bonds formation among the key residues in the substrate-binding pocket of G11A monomer monitored by 4-ns MD simulation

Residues	Atoms	Hydrogen-bonded residues	Atoms	Distance	Occupancy[a]
				Å	%
His-41	NH	Pro-39	O	3.043 ± 0.162	99.88
	NH	Arg-40	O	2.238 ± 0.041	100.0
	O	Val-42	NH	2.242 ± 0.041	100.0
	O	Cys-44	NH	3.055 ± 0.179	99.38
	HIE	Cys-145	HG	3.514 ± 0.182	0.124
	HE2	Asp-187	OD2	2.981 ± 0.245	91.57
Cys-145	O	Asn-28	HD22	3.074 ± 0.253	56.57
	HG	His-41	HIE	3.514 ± 0.182	0.124
	NH	Ser-144	O	2.240 ± 0.041	100.0
	SG	Gly-146	NH	3.413 ± 0.147	41.63
	O	Gly-146	NH	2.252 ± 0.042	100.0
	SG	His-163	HE2	3.510 ± 0.142	17.06
	SG	His-163	HIE	3.445 ± 0.163	26.48
	HG	His-164	O	3.446 ± 0.167	50.63
Tyr-161	OH	His-134	O	3.110 ± 0.261	87.88
	NH	Gly-149	O	3.278 ± 0.202	94.35
	O	Gly-149	NH	2.861 ± 0.115	100.0
	NH	Cys-160	O	2.244 ± 0.041	100.0
	O	Met-162	NH	2.241 ± 0.041	100.0
	O	His-163	NH	None	0.099
	O	Gly-174	NH	3.324 ± 0.193	89.14
His-163	HE2	Leu-138	O	2.902 ± 0.161	99.87
	HE2	Cys-145	SG	3.510 ± 0.142	17.06
	HIE	Cys-145	SG	3.445 ± 0.163	26.48
	NH	Ser-147	O	2.889 ± 0.113	100.0
	NH	Met-162	O	2.250 ± 0.040	100.0
	O	Gly-146	NH	2.920 ± 0.152	99.85
	NH	Tyr-161	O	None	0.099
	O	His-164	NH	2.249 ± 0.041	100.0
	ND1	Met-165	NH	3.540 ± 0.121	42.18
	HE2	Glu-166	OE2	3.242 ± 0.268	2.93
Glu-166	OE1	Asn-142	HD22	2.908 ± 0.220	20.43
	OE2	Gly-143	NH	3.005 ± 0.235	66.77
	OE2	Cys-145	NH	3.222 ± 0.250	1.64
	OE2	His-163	HE2	3.242 ± 0.268	2.93
	O	Leu-167	NH	2.349 ± 0.040	100.0
	OE1	His-172	HE2	3.055 ± 0.248	16.56
	OE2	His-172	HE2	3.010 ± 0.264	15.30

[a] Occupancy represents the ratio of hydrogen bonds existence during the whole 4-ns MD simulation process.

SG is 3.48 Å, whereas few hydrogen bonds could be formed between these two residues, indicating that the catalytic dyad in G11A monomer might fail to exhibit the right conformation and thus be incapable of performing catalysis.

Substrate Binding Pocket—The substrate-binding pocket of SARS-CoV 3CL^pro is composed of six subsites, namely S1–S6, corresponding to the P1–P6 residues on the peptide substrate. Among them, S1 subsite is the most important, because it recognizes the Gln-P1 residue of the peptide substrate and its correct conformation confers the enzyme absolute specificity for Gln at P1 position (5, 13). In the structure of wild-type protomer complexed with a substrate (1UK4_A) (Fig. 4a), the interactions between Gln-P1 and S1 subsite residues mainly involve the following hydrogen bonds: the side-chain OE1 of Gln-P1 with His-163 NE2; the side-chain NE2 of Gln-P1 with Glu-166 OE2; the main-chain oxygen of Gln-P1 with the main-chain nitrogen atoms of Gly-143, Ser-144, and Cys-145. Nota-

bly these three nitrogen atoms form an "oxyanion hole" structure, which is believed not only to bind the main-chain oxygen of Gln-P1 but also to stabilize the tetrahedral intermediate during the catalytic process. Conclusively the binding of Gln-P1 to the S1 subsite comprises three key structural elements: the oxyanion hole, His-163 and Glu-166. Therefore, we will describe the conformational changes of these three elements in G11A monomer separately.

Oxyanion Hole—As shown in Fig. 4a, the oxyanion hole of the active protomer 1UK4_A is large enough to accommodate the main-chain oxygen of Gln-P1 as well as the tetrahedral intermediate during catalysis. The residue Phe-140, as the major stabilizing force of the hole, can form hydrogen bonds with the N terminus Ser-1 from the counterpart protomer of the dimer, thus is held in place to stack against the imidazole ring of His-163 and supports the oxyanion hole. Besides, the hydrogen bond between Asn-28 ND2 and main-chain oxygen
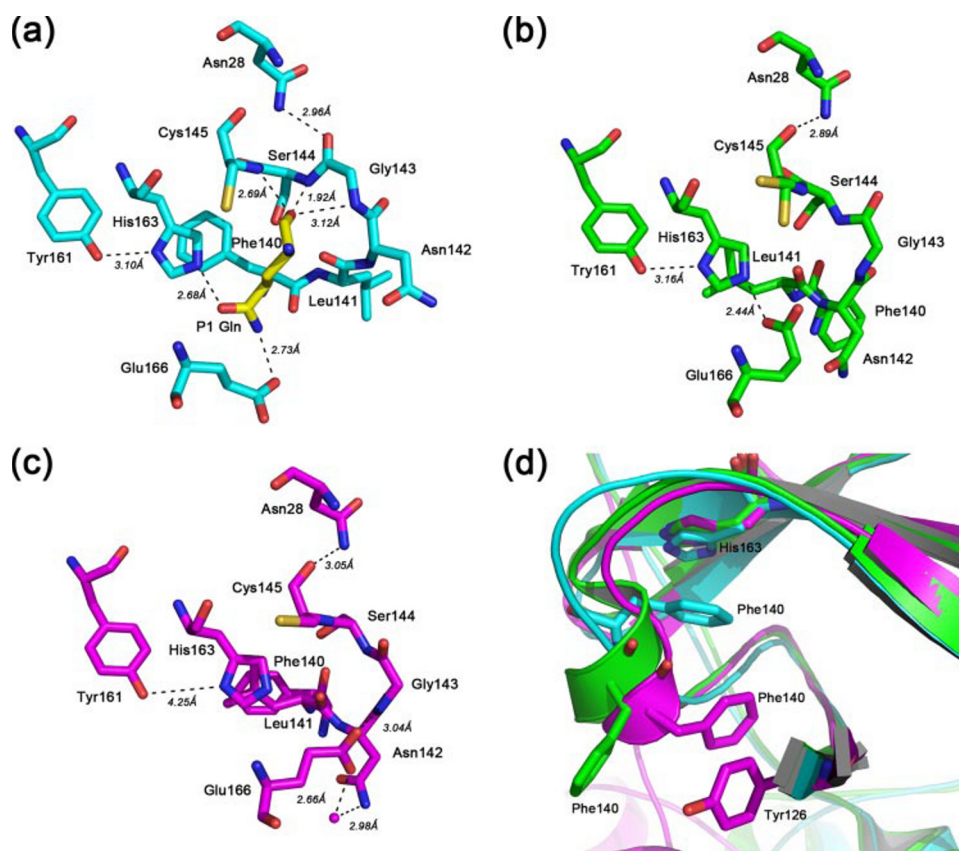
FIGURE 4. **Conformational variations in the S1 subsite of the substrate-binding pocket.** *a*, the substrate-complexed active protomer 1UK4_A; *b*, the inactive protomer 1UJ1_B; *c*, G11A monomer. All the residues are shown as *sticks*. The carbon atoms are colored in *cyan* (1UK4_A), *green* (1UJ1_B), and *magenta* (G11A). The nitrogen, oxygen, and sulfur atoms are colored in *blue*, *red*, and *yellow*, respectively. *Dashes* represent the key hydrogen bonds involved in substrate binding, and the residue-residue distances are also indicated. *d*, superposition of the three S1 subsites. The color scheme is the same as in *a*–*c*. The labeled residues are shown as *sticks*, and the rest of the proteins are shown as *schematics*.

His-163—In the substrate-complexed protomer 1UK4_A (Fig. 4*a*), His-163 specifically recognizes the side-chain OE1 of Gln-P1. This residue is held in the correct place mainly by two stabilizing forces in the protomer. One is the hydrogen bond between the side-chain hydroxyl group of Tyr-161 and His-163 ND1, the other is the packing between the side-chain rings of Phe-140 and His-163. However, these two stabilizing forces are both absent in G11A monomer (Fig. 4*c*) as indicated by the following facts. First, the distance between Tyr-161 OH and His-163 ND1 increases from 3.10 Å in 1UK4_A and 3.16 Å in 1UJ1_B to 4.25 Å in G11A monomer, implying the impossibility of the hydrogen bond formation, which is in agreement with the MD simulation result (Table 3, the occupancy of the hydrogen bond is less than 0.1%). Second, as mentioned above, Phe-140 in the G11A monomer escapes from the initial position to stack against Tyr-126 and no longer packs with His-163. This result has also been supported by the MD simulation in which the centroid distance between the imidazole ring of His-163 and the phenyl ring of Phe-140 remains larger than 8 Å during the whole simulation process (Fig. 2*c*), suggesting that no hydrophobic interaction might exist between these two residues in G11A monomer. In addition, similar to 1UJ1_B, Leu-141 in G11A monomer seems to replace the position originally occupied by Phe-140 and then interacts with His-163.

Glu-166—In the active protomer 1UK4_A (Fig. 4*a*), Glu-166 specifically recognizes the side-chain NE2 of Gln-P1 and locates at the entrance of the substrate binding pocket. The side-chain of Glu-166 points outside the binding pocket and is in an "open" state. In the inactive protomer 1UJ1_B (Fig. 4*b*), due to the formation of a hydrogen bond to His-163 NE2 by its OE2 atom, the side-chain of Glu-166 orients inward and presumably blocks the binding of the substrate (39). In G11A monomer (Fig. 4*c*), however, Glu-166 is in a unique conformation that has never been reported before. The OE2 atom of the side chain of Glu-166 forms a hydrogen bond with the main-chain nitrogen of Gly-143, and its OE1 atom also interacts with Asn-142 ND2 via a water molecule. Thus, Glu-166 might form a "closed door" at the entrance of the substrate binding pocket to inhibit substrate binding. In addition, the distance between Glu-166 OE2 and His-163 NE2 in G11A monomer becomes 4.79 Å, implying the disappearance of the hydrogen bond between these two residues, which is also in agreement with the

of Gly-143 might also contribute to the maintenance of the oxyanion hole. In G11A monomer, however, the oxyanion hole almost completely collapses (Fig. 4*c*), the main-chain nitrogen atoms of Gly-143, Ser-144, and Cys-145 move inward considerably, leaving no space for the main-chain oxygen of Gln-P1 and the tetrahedral intermediate. The similar structural collapse has also been observed in the oxyanion hole of the inactive protomer 1UJ1_B at pH 6.0 (Fig. 4*b*). But different from 1UJ1_B, in which the stabilizing element Phe-140 flips outward and induces the formation of an extra helix, Phe-140 of G11A monomer induces a similar helix but flips inward and packs against the side-chain ring of Tyr-126 in β10 strand via π-π interaction (Fig. 4*d*). This new position of Phe-140 might be one of the unique structural properties of the dissociated monomeric SARS-CoV 3CL^pro. In addition, similar to 1UJ1_B, Asn-28 ND2 of G11A monomer also forms a hydrogen bond to the main-chain oxygen of Cys-145 instead of Gly-143 (Table 3).

To further investigate the dynamic behavior of the oxyanion hole in G11A monomer, we performed a 4-ns MD simulation to calculate the r.m.s.d. of residues from Phe-140 to Cys-145 relative to the initial crystal structure (Fig. 2*b*). The result showed that the r.m.s.d. is relatively stable and fluctuates slightly around 1.0 Å after ~1 ns simulation, indicating that the collapsed oxyanion hole might be stable in G11A monomer structure.
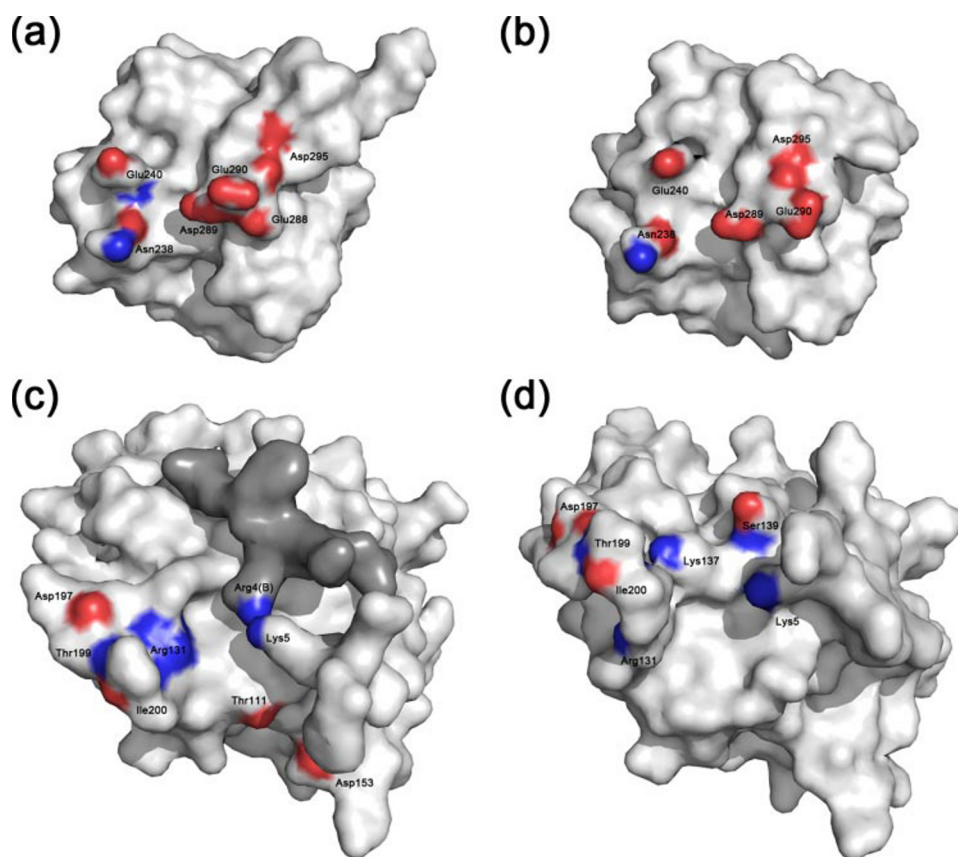
FIGURE 5. **The interface between domain III and the chymotrypsin fold of SARS-CoV 3CL^{pro}.** The domain III surface of the active protomer 1UK3_A (*a*) and G11A monomer (*b*) are represented, respectively. *c* and *d*, the surface of the rest of 1UK3_A and G11A. The *gray surface* in *c* represents the N-terminal finger of the other protomer 1UK3_B. The contact residues on the interface are labeled, and the atoms involved in forming hydrogen bonds or salt bridges are colored with oxygen in *red* and nitrogen in *blue*.

dimer and dissociated monomer, we explored the interface between domain III and the rest part of the molecule. In wild-type dimer of SARS-CoV 3CL^{pro}, we identified that the regions around Glu-290 (Glu-288, Asp-289, Glu-290, and Asp-295) and Asn-238 (Asn-238 and Glu-240) form extensive interactions with other structural elements of the molecule (Fig. 5, *a* and *c*, and supplemental Table S1). Among them, the salt bridge between Glu-290 of one protomer and Arg-4 from the N-terminal finger of another protomer is believed to be the major contribution in stabilizing domain III and monomer-monomer associations (7). Collaboratively these interactions might fix domain III in its correct position as observed in wild-type dimer. In G11A monomer, we were surprised to find that despite the large spatial movement of domain III relative to domains I and II, the two regions mentioned above still dominate the interactions of domain III with the rest part of the protease, but the residues that they make contact with have changed comparing with those in wild-type dimer (Fig. 5, *b* and *d*, and supplemental Table S1). Noticeably the side chain of Glu-290 rotates about 60° relative to its position in the dimer and forms extensive hydrogen bonds with Ser-139. Furthermore, the main-chain oxygen of Glu-290 forms a hydrogen bond to Lys-5 NZ. These interactions, which are not observed in wild-type dimer, are believed to have well stabilized Glu-290 and the spatial position of domain III in G11A monomer.

MD simulation result (Table 3, the occupancy of the hydrogen bond is less than 3%).

Dimerization-related Structures (the N-terminal Finger and Domain III)—Several reports have proven that the N-terminal finger plays a key role in mediating dimerization of SARS-CoV 3CL^{pro} (10, 34). As has been indicated, the residue Gly-11 is located on the first helix of domain I in SARS-CoV 3CL^{pro} (α-helix A'). Usually the α-helix A' of wild-type protease comprises the residues from Ser-10 to Gly-15, however in G11A monomer, mutation of residue Gly-11 to Ala has shortened the helix to Ala-11 to Glu-14, probably due to the different dihedral angle restraints between glycine and alanine. As a result, the N-terminal finger (residues 1–7) connected to the N terminus of the helix points out in a distinct direction that is about 43° away from the normal direction found in wild-type dimer (Fig. 3). This mis-orientation of N-terminal finger might have prevented itself from squeezing into the partner monomer and trigger the dimer dissociation of the G11A mutant in the crystal structure.

Besides the N-terminal finger, it has been reported that domain III also contributes well to dimerization of SARS-CoV 3CL^{pro} (6, 14). In agreement with this result, domain III undergoes a dramatic conformational switch relative to the chymotrypsin fold (domain I and II) in G11A monomer, rotating by 24° comparing with its normal position in wild-type dimer. To investigate how and why domain III is positioned differently in

DISCUSSION

As a critical target for anti-SARS drug design, SARS-CoV 3CL^{pro} has been extensively characterized for its structural property and enzymatic activity (5, 7–9, 18, 38, 40). Much progress has also been made for understanding the correlation between dimerization and catalytic activity of the protease (9–12). In the present study, we performed single point mutation of SARS-CoV 3CL^{pro} targeting the residue Gly-11 on the α-helix A' of domain I, which mediates extensive monomer-monomer interactions as revealed by the crystal structure of wild-type dimer (supplemental Fig. S1). Mutation of Gly-11 to Ala does not change the folding manner of the protease (supplemental Fig. S2), but the catalytic activity of G11A mutant is completely abolished (Fig. 1), indicating that the residue Gly-11 might play a important role in maintaining the dimer structure of SARS-CoV 3CL^{pro}. To better elucidate the mutation-induced influence on SARS-CoV 3CL^{pro} structure, the crystal structure of G11A mutant was subsequently analyzed, which

reveals that it exists only as a monomer in crystal (Table 1 and Fig. 3). To our knowledge, the crystal structure of G11A mutant is the first reported monomer structure of SARS-CoV 3CL^pro, which might hopefully help to explore the structural difference between the dimeric and monomeric forms of the protease at atomic level. Together with the results of MD simulations, this structure will also provide useful information for illustrating why only the dimer can perform the catalytic function and the dissociated monomer is inactive, which is the most intriguing aspect of SARS-CoV 3CL^pro.

Why Does Mutation of Gly-11 Cause the Complete Dissociation of the Dimer in Crystal?—To date, the N-terminal finger (residues 1–7) and domain III have been identified to extensively mediate monomer-monomer interactions of SARS-CoV 3CL^pro (6, 10, 14, 15, 18). Now we have brought forward a new structural element that is also vital for dimerization of the protease, namely α-helix A′ (residues 10–15). Several single point mutations of the residues on α-helix A′, including Gly-11, have almost completely abolished the activity of the protease (Fig. 1 and unpublished data). Considering the structural details of G11A monomer (Fig. 3), we speculate that α-helix A′ might determine the correct spatial orientation of the N-terminal finger. In the dimer structure, the N-terminal finger of one protomer can squeeze into the space between domain III of its parent protomer and domain II of the neighboring protomer (5), however, the mis-oriented N-terminal finger caused by damage of α-helix A′ (like that of G11A mutant) might be unable to insert correctly into the pocket of another monomer and thus destabilize the dimer structure.

Why Does the Relative Position of Domain III Change Dramatically in G11A Monomer?—Besides the mis-oriented N-terminal finger, the considerable movement of domain III is another striking observation in G11A monomer (Fig. 3 and Table 2). We believe that this positional relationship between domain III and the chymotrypsin fold represents a real and stable conformation unique to the dissociated monomer of SARS-CoV 3CL^pro, which is supported by two evidences. First, the MD result indicated that the structure of G11A monomer might be relatively stable in solution environment (Fig. 2a). Second, domain III forms extensive interactions with the rest part of G11A monomer (Fig. 5, b and d, and supplemental Table S1). It seems that upon dimer dissociation, domain III has "glided" from its original position and properly "anchored" in its new position in G11A monomer, forming new interactions with the regions on the surface of the chymotrypsin fold distinct from those in the dimer. Noticeably, Glu-290, which is reported to be important for monomer-monomer associations (7), forms extensive hydrogen bonds with Ser-139 and Lys-5 in G11A monomer. These interactions, which are absent in the dimer, are believed to well stabilize Glu-290 and thus fix the position of domain III in G11A monomer. In addition, the long loop region connecting domain II and III might be flexible (especially residues 193–200, Table 2) in order to tolerate the large conformational change of domain III. Therefore, we concluded that the unique interdomain arrangement of domain III and the chymotrypsin fold in G11A monomer might be an intrinsic property of monomeric SARS-CoV 3CL^pro.
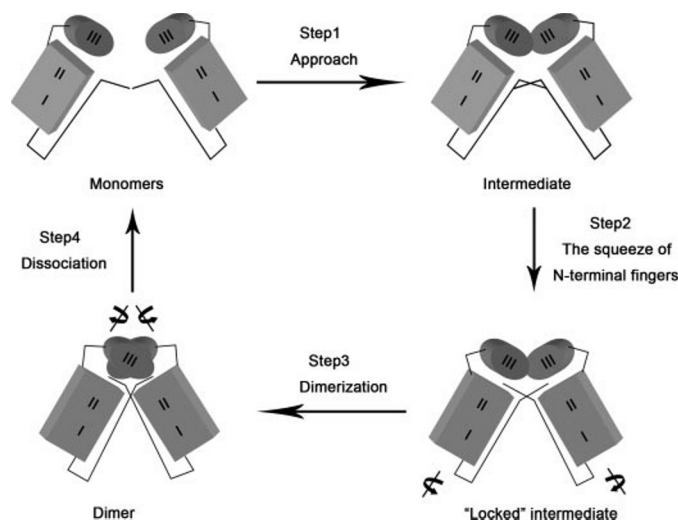

FIGURE 6. **A putative dimerization model of SARS-CoV 3CL^pro.** Domain I and II of SARS-CoV 3CL^pro are shown as *boxes*, domain III is shown as a *cylinder*. The dimerization model of SARS-CoV 3CL involves four steps. *Step 1*: initially two monomers approach each other and their domains III form an "intermediate" dimer, which induce the rotations of domains I and II. *Step 2*: subsequently the N-terminal fingers mutually squeeze in the space between domains I and II of one protomer and domain III of the other protomer, thus locking the dimer in a stable state. *Step 3*: meanwhile, domain III switches to the "final" conformations to produce the active dimer. *Step 4*: because the dimerization process is in equilibrium, the dimer can also dissociate into monomers and enter the cycle again.

In addition, Shi *et al.* (6) reported that domain III of SARS-CoV 3CL^pro itself possesses an intrinsic intention to form dimer even at a very low concentration. According to this finding and the structure of G11A monomer, we proposed a dimerization model of SARS-CoV 3CL^pro (Fig. 6). When two monomers approach each other, their domains III might initially form an "intermediate" dimer and then induce the relative rotations of their chymotrypsin folds (domain I and II). After these conformational changes, two N-terminal fingers mutually squeeze into the pockets formed by domain III of one monomer and domain II of the other monomer, just like a "key" to "lock" the dimer at a stable state. In the meantime, domains III switch to their "final" conformations by the Arg-4 to Glu-290 salt bridge as seen in the dimer crystal structure. Based on this model, the mis-orientated N-terminal finger, which is caused by mutation of Gly-11 on α-helix A′, is unable to properly lock the dimer structure and pushes the dimer-monomer equilibrium toward the monomer state, thus eventually induces the complete dimer dissociation of SARS-CoV 3CL^pro in the crystal structure.

Why Does the Dissociated G11A Monomer Completely Lose Enzymatic Activity?—Although numerous reports have pointed out that the monomer is the inactive form of SARS-CoV 3CL^pro, the related evidences were only obtained from computational simulations due to the absence of the crystal structure of the monomeric protease (37, 39). Our crystal structure of G11A monomer has provided the direct proof and detailed description of the incompetent state of the dissociated monomer of SARS-CoV 3CL^pro.

Lai *et al.* (39) have suggested that the right conformation of the catalytic dyad could be indicated by the hydrogen bond formed between SG atoms in Cys-145 and the ND1/NE2 atom in His-41. We monitored this parameter in G11A monomer by

4-ns MD simulation and found that few hydrogen bonds could be formed between these atoms (Table 3), suggesting that the catalytic dyad is unlikely to be competent in the monomer form of the protease.

In addition, SARS-CoV 3CL^{pro} confers absolute specificity for Gln at the P1 position of the peptide substrate (29, 40), which requires that the S1 subsite of the substrate-binding pocket should retain a correct conformation to accommodate Gln-P1. The damage of the S1 subsite is likely to entirely abolish the enzymatic activity, as observed in the crystal structure of the inactive protomer obtained at pH 6.0 (PDB code: 1UJ1_B) (5). In G11A monomer, several key structural elements of the S1 subsite, including the oxyanion hole, His-163 and Glu-166, are considerably impaired. Furthermore, the conformation of the S1 subsite in G11A monomer is also different from that in the inactive protomer 1UJ1_B, and is regarded as a unique property of the dissociated monomer.

The oxyanion hole, composed of the main-chain nitrogens of Gly-143, Ser-144, and Cys-145, is believed to accommodate the main-chain oxygen of Gln-P1 as well as the tetrahedral intermediate during catalysis. This hole is fully open in the active protomer (Fig. 4*a*), probably supported by Phe-140 that is held in place by forming hydrogen bond with Ser-1 from the N-terminal finger of another protomer and stacking against His-163. When losing the counteracting force from the neighboring protomer, as in G11A monomer, Phe-140 escapes from the original position and turns to pack with Tyr-126, which is a unique property of the monomer form (Fig. 4, *c* and *d*). Thus the oxyanion hole collapses and leaves no room for either Gln-P1 or the tetrahedral intermediate. Furthermore, the hydrogen bond between Asn-28 ND2 and main-chain oxygen of Cys-145 might have stabilized the collapsed hole, which is in agreement with the MD simulation result that the conformation of this incompetent oxyanion hole might also be stable (Fig. 2*b*). Besides the oxyanion hole, His-163 is also essential for the substrate binding and catalysis of the protease. The correct position of His-163 is mainly held by two forces: 1) the hydrogen bond to Tyr-161 and 2) the stacking of its imidazole ring against Phe-140. These two forces are both absent in G11A monomer. In detail, the distance between His-163 ND1 and Tyr-161 OH increases to 4.25 Å (Fig. 4*c*), and Phe-140 no longer packs against His-163 (Fig. 2*c*). Therefore, although the position of His-163 in G11A monomer does not change much comparing with the protomer of the dimer, the stability of this residue has been severely undermined, which probably also results in the incompetence of the dissociated monomer. In addition, Glu-166 recognizes the side-chain NE2 of Gln-P1 and locates at the entrance of the S1 subsite in the dimer, probably acting as a door to regulate the substrate binding. In the active protomer 1UK4_A, Glu-166 is held in place by interacting with Ser-1 of the N-terminal finger from another protomer and stays in an open state to interact with Gln-P1 NE2 (Fig. 4*a*). In the inactive protomer 1UJ1_B, Glu-166 forms a hydrogen bond to His-163 and points inside the S1 subsite, thus hindering the binding of Gln-P1 (Fig. 4*b*). In G11A monomer, however, Glu-166 forms hydrogen bond to Gly-143 and interacts with Asn-142 via a water molecule (Fig. 4*c* and Table 3). With this unique "door-closed" conformation, Glu-166 completely blocks the entrance of the S1 subsite,

which is probably another major cause for the inactivation of the dissociated monomer. Furthermore, the different spatial position of Glu-166 between G11A monomer and the inactive protomer 1UJ1_B is likely to be a result of the instability of His-163. The unsteady His-163 in G11A monomer no longer forms hydrogen bond with Glu-166 as it does in the inactive protomer 1UJ1_B (Table 3). Therefore, Glu-166 takes this unique and stable conformation in the dissociated monomer, but still blocks the binding of Gln-P1.

CONCLUSION

In summary, we reported the first crystal structure of the monomeric SARS-CoV 3CL^{pro} induced by mutation of Gly-11 to Ala. The mutation might shorten the α-helix A′ of the protease and cause a mis-oriented N-terminal finger that could no longer exactly squeeze into the pocket of another monomer, thus destabilize the dimer structure. The dimer dissociation could disrupt some key structural features and finally completely inactivate the protease. In addition, the positional relationship between domain III and the chymotrypsin fold in the G11A monomer has changed dramatically comparing with that in the dimer. Based on these findings, we proposed a dimerization model of SARS-CoV 3CL^{pro}, in which domains III of two monomers might initially form an "intermediate" dimer, then induce the rotation of the chymotrypsin folds whose N-terminal fingers subsequently squeeze into the right positions and fix the dimer in a stable state. Our current work provided valuable insight into the dimerization mechanism of SARS-CoV 3CL^{pro}, and supplied the direct structural evidence for the incompetence of the dissociated monomer. Furthermore, the new structural features that we found to be important for maintaining the dimer-monomer equilibrium and enzymatic activity of SARS-CoV 3CL^{pro}, *e.g.* the α-helix A′, the orientation of the N-terminal finger, the interface of domain III contacting with the chymotrypsin fold, and the flexibility of the long loop region, could probably be used as new potential targets for developing 3CL^{pro} inhibitors.

REFERENCES

1. Fouchier, R. A., Hartwig, N. G., Bestebroer, T. M., Niemeyer, B., de Jong, J. C., Simon, J. H., and Osterhaus, A. D. (2004) *Proc. Natl. Acad. Sci. U. S. A.* **101,** 6212–6216
2. 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., and Yuen, K. Y. (2003) *Lancet* **361,** 1319–1325
3. Thiel, V., Ivanov, K. A., Putics, A., Hertzig, T., Schelle, B., Bayer, S., Weissbrich, B., Snijder, E. J., Rabenau, H., Doerr, H. W., Gorbalenya, A. E., and Ziebuhr, J. (2003) *J. Gen. Virol.* **84,** 2305–2315
4. Anand, K., Ziebuhr, J., Wadhwani, P., Mesters, J. R., and Hilgenfeld, R. (2003) *Science* **300,** 1763–1767
5. 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) *Proc. Natl. Acad. Sci. U. S. A.* **100,** 13190–13195
6. Shi, J., Wei, Z., and Song, J. (2004) *J. Biol. Chem.* **279,** 24765–24773
7. Chou, C. Y., Chang, H. C., Hsu, W. C., Lin, T. Z., Lin, C. H., and Chang, G. G. (2004) *Biochemistry* **43,** 14958–14970
8. Fan, K., Wei, P., Feng, Q., Chen, S., Huang, C., Ma, L., Lai, B., Pei, J., Liu, Y., Chen, J., and Lai, L. (2004) *J. Biol. Chem.* **279,** 1637–1642
9. Hsu, M. F., Kuo, C. J., Chang, K. T., Chang, H. C., Chou, C. C., Ko, T. P., Shr, H. L., Chang, G. G., Wang, A. H., and Liang, P. H. (2005) *J. Biol. Chem.*

280, 31257–31266
10. Hsu, W. C., Chang, H. C., Chou, C. Y., Tsai, P. J., Lin, P. I., and Chang, G. G. (2005) *J. Biol. Chem.* **280,** 22741–22748
11. Graziano, V., McGrath, W. J., DeGruccio, A. M., Dunn, J. J., and Mangel, W. F. (2006) *FEBS Lett.* **580,** 2577–2583
12. Graziano, V., McGrath, W. J., Yang, L., and Mangel, W. F. (2006) *Biochemistry* **45,** 14632–14641
13. Tan, J. Z., Verschueren, K. H. G., Anand, K., Shen, J. H., Yang, M. J., Xu, Y. C., Rao, Z. H., Bigalke, J., Heisen, B., Mesters, J. R., Chen, K. X., Shen, X., Jiang, H. L., and Hilgenfeld, R. (2005) *J. Mol. Biol.* **354,** 25–40
14. Shi, J., and Song, J. (2006) *FEBS J.* **273,** 1035–1045
15. Barrila, J., Bacha, U., and Freire, E. (2006) *Biochemistry* **45,** 14908–14916
16. Ding, L., Zhang, X. X., Wei, P., Fan, K., and Lai, L. (2005) *Anal. Biochem.* **343,** 159–165
17. Lee, T. W., Cherney, M. M., Huitema, C., Liu, J., James, K. E., Powers, J. C., Eltis, L. D., and James, M. N. (2005) *J. Mol. Biol.* **353,** 1137–1151
18. Chen, S., Chen, L., Tan, J., Chen, J., Du, L., Sun, T., Shen, J., Chen, K., Jiang, H., and Shen, X. (2005) *J. Biol. Chem.* **280,** 164–173
19. 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) *Protein Expr. Purif.* **32,** 302–308
20. Chen, S., Chen, L. L., Luo, H. B., Sun, T., Chen, J., Ye, F., Cai, J. H., Shen, J. K., Shen, X., and Jiang, H. L. (2005) *Acta Pharmacol. Sin.* **26,** 99–106
21. Matthews, B. W. (1968) *J. Mol. Biol.* **33,** 491–497
22. Vagin, A., and Teplyakov, A. (1997) *J. Appl. Crystallogr.* **30,** 1022–1025
23. Collaborative Computational Project, Number 4. (1994) *Acta Crystallogr. D. Biol. Crystallogr.* **50,** 760–763
24. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr. D. Biol. Crystallogr.* **54,** 905–921
25. Emsley, P., and Cowtan, K. (2004) *Acta Crystallogr. D. Biol. Crystallogr.* **60,** 2126–2132
26. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard, M. (1991) *Acta Crystallogr. Sect. A* **47,** 110–119
27. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *J. Appl. Crystallogr.* **26,** 283–291
28. Kabsch, W. (1976) *Acta Crystallogr. Sect. A* **32,** 922–923
29. Gao, F., Ou, H. Y., Chen, L. L., Zheng, W. X., and Zhang, C. T. (2003) *FEBS Lett.* **553,** 451–456
30. Wang, J., Cieplak, P., and Kollman, P. A. (2000) *J. Comput. Chem.* **21,** 1049–1074
31. Ryckaert, J. P., Ciccott, G., and Berendsen, H. J. C. (1977) *J. Comput. Phys.* **23,** 327–341
32. Miyamoto, S., and Kollman, P. A. (1992) *J. Comput. Chem.* **13,** 952–962
33. Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F., DiNola, A., and Haak, J. R. (1984) *J. Comput. Phys.* **81,** 3684–3690
34. Wei, P., Fan, K., Chen, H., Ma, L., Huang, C., Tan, L., Xi, D., Li, C., Liu, Y., Cao, A., and Lai, L. (2006) *Biochem. Biophys. Res. Commun.* **339,** 865–872
35. Xu, T., Ooi, A., Lee, H. C., Wilmouth, R., Liu, D. X., and Lescar, J. (2005) *Acta Crystallograph. Sect. F. Struct. Biol. Cryst. Commun.* **61,** 964–966
36. Xue, X., Yang, H., Shen, W., Zhao, Q., Li, J., Yang, K., Chen, C., Jin, Y., Bartlam, M., and Rao, Z. (2007) *J. Mol. Biol.* **366,** 965–975
37. Zheng, K., Ma, G., Zhou, J., Zen, M., Zhao, W., Jiang, Y., Yu, Q., and Feng, J. (2007) *Proteins* **66,** 467–479
38. Huang, C., Wei, P., Fan, K., Liu, Y., and Lai, L. (2004) *Biochemistry* **43,** 4568–4574
39. Chen, H., Wei, P., Huang, C., Tan, L., Liu, Y., and Lai, L. (2006) *J. Biol. Chem.* **281,** 13894–13898
40. Fan, K., Ma, L., Han, X., Liang, H., Wei, P., Liu, Y., and Lai, L. (2005) *Biochem. Biophys. Res. Commun.* **329,** 934–940