ScienceAsia 29 (2003): **181-188**

Structure and Dynamics of SARS Coronavirus Proteinase: The Primary Key to the Designing and Screening for Anti-SARS Drugs

Vannajan Sanghiran Lee,[a,b] Kitiyaporn Wittayanarakul,[a] Tawun Remsungnen,[a] Vudhichai Parasuk,[a] Pornthep Sompornpisut,[a] Wasun Chantratita,[c] Chak Sangma,[d] Sornthep Vannarat,[f] Piyawut Srichaikul,[f] Supa Hannongbua,[d] Patchreenart Saparpakorn,[d] Witcha Treesuwan,[d] Ornjira Aruksakulwong,[a] Ekawat Pasomsub,[c] Siriporn Promsri,[a] Daungmanee Chuakheaw,[d] and Supot Hannongbua[a,*]

[a] Department of Chemistry, Faculty of Science, Chulalongkorn University, Bangkok 10330, Thailand.

[b] Department of Chemistry, Faculty of Science, Chiang Mai University, Chiang Mai 50200, Thailand.

[c] Virology and Molecular Microbiology, Department of Pathology, Faculty of Medicine, Ramathibodi Hospital, Mahidol University, Bangkok 10400, Thailand.

[d] Department of Chemistry, Faculty of Science, Kasetsart University, Bangkok 10900, Thailand.

[f] High Performance Computing Research and Development Division, National Electronics and Computer Technology Center, Thailand Science Park, 112 Pahonyothin Road, Klong 1, Klong Luang, Pathumthani 12120, Thailand.

* Corresponding author, E-mail: supot.h@chula.ac.th

Received 3 Jun 2003
Accepted 30 Jun 2003

Abstract Structure and dynamics of SARS coronavirus (SARS-CoV) proteinase have been investigated using molecular dynamics (MD) simulation technique. The simulations were carried out under the NPT ensemble at 298 K and 1 atm, for free enzyme in both monomer and dimer forms and the monomer-inhibitor complex. The systems were observed to reach equilibrium after 200 ps. Enzyme conformation and its structural changes were monitored in terms of root mean square displacement of all 306 amino acid residues. The results show, as expected, that the proteinase in complex form is, in average, less flexible than the free form. Interest is centered on the two regions, at N-termini (residues 1-2) and around the active site (residues 56-62), in which the flexibilities in complex form is lower than those in free form. This behavior is supposed to facilitate the binding between enzyme and substrate. With the obtained MD structure, molecular dockings have been carried out in order to search for potent SARS-CoV proteinase inhibitors. Preliminary results show that among sixteen antiviral drugs taken from the NCI database, four of them with trade-name Nevirapine, Glycovir, Virazole, and Calanolide A, are observed to fit well in the active site of the SARS-CoV proteinase.

KEYWORDS: proteinase, protease, molecular structure, molecular dynamics, SARS, Coronavirus, drug design.

Introduction

Severe acute respiratory syndrome (SARS) has recently emerged as a new human disease, resulting globally in 682 deaths from 8,046 probable cases (as of 23 May 2003).[1] The discovery that the virus can be readily isolated in a monkey-kidney cell line was the key to the rapid molecular characterization of this novel coronavirus and the development of diagnostic tests for SARS.[2] The nucleotide sequence of the SARS-associated coronavirus genome differs substantially from sequences of all known coronaviruses.[3, 4]

SARS-associated coronavirus was finally proven to be the cause of SARS. Inoculation of monkeys with SARS-associated coronavirus from cell cultures caused lower respiratory tract disease, fulfilling Koch's postulate.[5] The understanding of the aetiology of SARS will expedite the development of diagnostic tests, antiviral therapies and vaccines.

Vaccines are available for some animal coronaviruses. However, some vaccines against feline coronaviruses actually promoted the disease when vaccinated animals were exposed to wild-type virus, and antibody enhancement of disease is a potential risk of SARS vaccines in human.[6] As a matter of fact, it would take much longer time to develop a good preventive vaccine against viruses, than to screen or improve compound which exhibits high potential to be a good inhibitor from available drugs or from natural products. In order to do that, a three dimensional structure of the enzyme target is required. Unfortunately, it has not been available experimentally for the SARS coronavirus.

Anand *et al*[7] recently proposed a three dimensional structure of the SARS coronavirus proteinase, an enzyme involved in viral replication, based on the homology modeling technique. They also recommended to use a rhinovirus inhibitor (codename AG7088) which is already in clinical trials as anti-common cold drug, to be the model for the design of anti-SARS drugs. Nevertheless, it has been found that this compound is inactive against SARS *in vitro*.[8] Based on theoretical approach, another three dimensional structure of this enzyme was determined by homology modeling and then subjected to molecular dynamics (MD) method.[9] This structure should be the appropriate model for the structure based drug design as well as for drug screening, *in silico*.

In this study, structure and dynamics of SARS-CoV proteinases were investigated using molecular dynamics technique. Simulations were performed for both free enzyme and its complex with the SARS-CoV substrate-analog inhibitor.

Calculation Details

Preparation of the Initial Structures

The sequence alignment of coronavirus proteinases of porcine transmissible gastroenteritis virus (1LVO_A) and SARS-associated coronavirus (SARS_protein) using Bioedit software and CLUSTAL X (1.81) multiple sequence alignment software with mutation data matrix (BLOSUM) was compared in Fig 1.[3-4, 7]

SARS-CoV proteinase consists of 306 amino acids. This sequence has 35% identity and 65% similarity in comparison to amino acid residues of porcine coronavirus (transmissible gastroenteritis virus, TGEV).

Starting from X-ray structure of the proteinase of porcine coronavirus and using homology-modeling technique,[10] initial structure of the SARS-CoV proteinase in monomer form (Mono-Free) was prepared. The dimer form (Di-Free) was generated by superposition of Mono-Free with a dimer form (Chains A and B) taken from the PDB entry of the TGEV proteinase. For the structure of the monomer complex (Mono-Cpx), the SARS-CoV substrate-analog inhibitor (Thr-Ser-Ala-Val-Leu-Gln) was derived from the P6-P1 residues of the NH_2-terminal auto-processing site of SARS-CoV^{Mpro}.[7] It was, then, inserted into the SARS-CoV proteinase in an area close to H41 and C145 in the conformation analogue to that found for the porcine coronavirus proteinase-inhibitor complex. Note that H41 and C145 are known to form an active site based on the sequence comparison with the porcine coronavirus proteinase.

Molecular Dynamics Simulations

Three MD simulations were carried out for the systems consisting of the following 3 forms of the SARS-CoV proteinase: monomer in free form, dimer in free form and monomer complexed with SARS-CoV substrate-analog inhibitor.

The Mono-Free, Di-Free and Mono-Cpx were

```
SARS_protein    SGFRKMAFPSGKVEGCMVQVTCGTTTLNGLWLDDTVYCPRHVICTAEDMLNPNYEDLLIR 60
1LVO_A          SGLRKMAQPSGLVEPCIVRVSYGNNVLNGLWLGDEVICPRHVIASDTTRV-INYENEMSS 59
                **:**** *** ** *:*:*: *...******.* * ******.:    :  ***: :

SARS_protein    KSNHSFLVQAGNVQLRVIGHSMQNCLLRLKVDTSNPKTPKYKFVRIQPGQTFSVLACYNG 120
1LVO_A          VRLHNFSVSKNNVFLGVVSARYKGVNLVLKVNQVNPNTPEHKFKSIKAGESFNILACYEG 119
                   *.* *. .** * *:.   :.  * ***:  **:**::**  *:.*::*.:****:*

SARS_protein    SPSGVYQCAMRPNHTIKGSFLNGSCGSVGFNIDYDCVSFCYMHHMELPTGVHAGTDLEGK 180
1LVO_A          CPGSVYGVNMRSQGTIKGSFIAGTCGSVGYVLENGILYFVYMHHLELGNGSHVGSNFEGE 179
                .*..**   **.: ******: *:*****: :: . : * ****:** .* *.*:::**:

SARS_protein    FYGPFVDRQTAQAAGTDTTITLNVLAWLYAAVINGDRWFLNRFTTTLNDFNLVAMKYNYE 240
1LVO_A          MYGGYEDQPSMQLEGTNVMSSDNVVAFLYAALINGERWFVTNTSMSLESYNTWAKTNSFT 239
                :** : *: : *  **:.  : **:*:****:***:***:.. : :*:.:*  * . .:

SARS_protein    PLTQDHVDILGPLSAQTGIAVLDMCAALKELLQNGMNGRTILGSTILEDEFTPFDVVRQC 300
1LVO_A          ELSS--TDAFSMLAAKTGQSVEKLLDSIVRLNK-GFGGRTILSYGSLCDEFTPTEVIRQM 296
                 *:.  .* :. *:*:** :* .:  :: .* : *:.*****.   * ***** :*:**

SARS_protein    SGVTFQ 306
1LVO_A          YGVNLQ 302
                 **.:*
```

Fig 1. The sequence alignments using CLUSTAL X (1.81) multiple sequence alignment software with mutation data matrix (BLOSUM) of porcine transmissible gastroenteritis virus (1LVO_A) and SARS-associated coronavirus (SARS_protein). The (*), (:) and (.) denote the identity and similarity of amino acid residues, while (-) illustrates the position of amino acid insertion or deletion.

solvated by 9572, 10746 and 19370 TIP3P water molecules, respectively, in a cubical cell and treated in the simulation under periodic boundary conditions. Energy minimizations were carried out to relax the system prior to MD runs. A cutoff distance of 12 Å was applied for non-bonded pair interaction. The Particle Mesh Ewald (PME) method was employed for correcting electrostatic interaction. Sodium and Chloride ions were added to neutralize the system. The simulation time step was set at 2 femtosecond (fs). The temperature of the system was gradually raised to 298 K for the first 60 picosecond (ps), and then kept constant according to the Berendensen algorithm[11] with a coupling time of 0.2 ps. The trajectories were collected for 200 ps after equilibration in order to be used for the data evaluation.

The molecular mechanics potential energy minimizations and MD simulations were carried out using the program package AMBER7.[12-13] Calculations were performed using the parm99 force field. All MD runs reported here were done under an isobaric-isothermal ensemble (NPT) using constant pressure of 1 atm and constant temperature of 298 K.

Molecular Docking

With the three dimensional structure of the SARS-CoV proteinase yielded from molecular dynamics simulations, molecular docking was performed using the Autodock 3.05 program.[14-16] The 3D affinity grid fields were created using the auxiliary program AutoGird. The center of protein mass was chosen as the grid center. In this stage, the protein was embedded in the 3D grid and a probe atom was placed at each grid point. The affinity and electrostatic potential grid was calculated for each type of atom in the proteinase molecule. The number of grid points in x, y, z-axis was 60 x 60 x 60 with grid points separated by 0.375 Å. Kollman-United atomic charges were added to the enzyme and the investigated compounds. The interaction between enzyme and ligands were calculated using the force field scoring function.

Results and Discussion

Reliability of the Homology Model

The quality of the geometry and of the stereochemistry of the protein structure was validated using PROCHECK.[17] The calculation of main chain torsion angles of the protein showed no severe distortion of the backbone geometry. A total of 97% of the backbone dihedrals, f and y, fall within the structurally favorable regions in Ramachandran plot. The overall average of the PROCHECK's score was above the minimum requirement. This is a good indication for the quality of the starting homology model prior to the refinement with molecular dynamics.

Three Dimensional Structure of SARS Proteinase

During the MD simulation, the changes in molecular conformation were monitored in terms of a root mean square displacement (RMSD) of all atoms with respect to the starting geometry. The results, as a function of time steps, for Mono-Free, Di-Free and Mono-Cpx are given in Fig 2. To understand detailed properties of the system, RMSD of each chain of the dimer (Fig 2b) as well as that of enzyme and substrate in the complex form (Fig 2a) were plotted separately.

In the first 50 ps, the global conformation of the enzyme was slightly different from the starting structure. The RMSD increases exponentially during 50-100 ps and remains constant afterward. The plots describe the molecular structure relaxation, as for example when the molecule in the crystalline form was dissolved in the solution. It can be seen that the

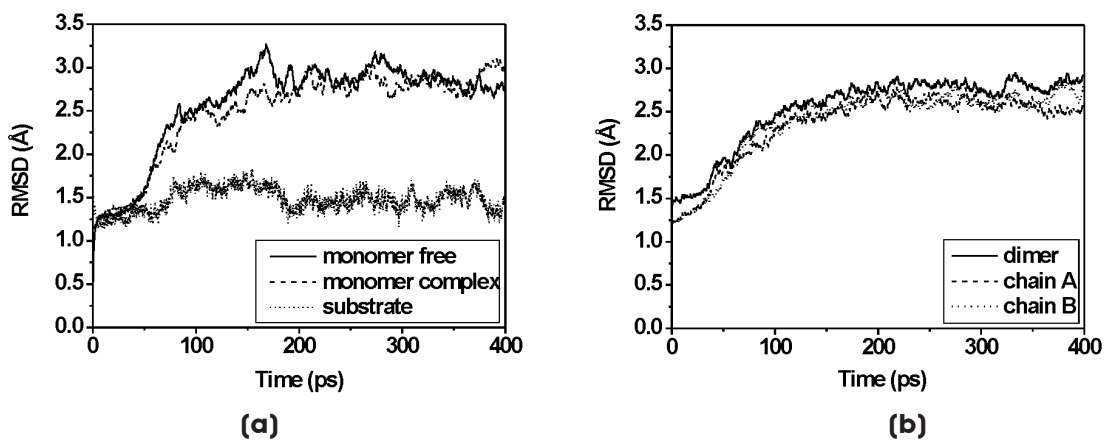


Fig 2. Root mean square displacement (RMSD) of all atoms extracted from the first 400 ps in the molecular dynamics simulations of the proteinase of the SARS coronavirus in (a) monomer form and its complex with the SARS coronavirus substrate-analog inhibitor and (b) dimer in free form.

magnitude of changes in the conformation is in the following order: Mono-Free > Mono-Cpx > Di-Free. It means, as expected, that the overall movement of the conformation in the monomeric free enzyme is greater than that in the complex form while no significant difference in RMSD is found between the two chains of the dimer, when they are in the solution. In addition, RMSD of the substrate is much lower than those of the enzymes because there are much fewer atoms in the substrate to contribute to the RMSD calculation in comparison to the enzyme.

From the plot, the fluctuation of RMSDs is small after 200 ps. Therefore, all information reported in this study was extracted from individual properties of each atom after 200 ps.

In Fig 3, three-dimensional structures of these three forms are displayed. Note that the X-ray structure of this enzyme is not yet available due to complexity of the enzyme crystallization. This is the first time that the structure of SARS-CoV proteinase was determined using homology modeling followed with MD simulations. The molecular structure and sizes of active sites differ significantly among the three investigated forms.[9] All of these will directly affect the binding affinity between enzyme and substrate as well as between enzyme and inhibitors which are governed by molecular geometry and electron distribution of the molecule. In addition to the steric and electronic effects, the enzyme-substrate binding depends strongly on flexibility of the enzyme.

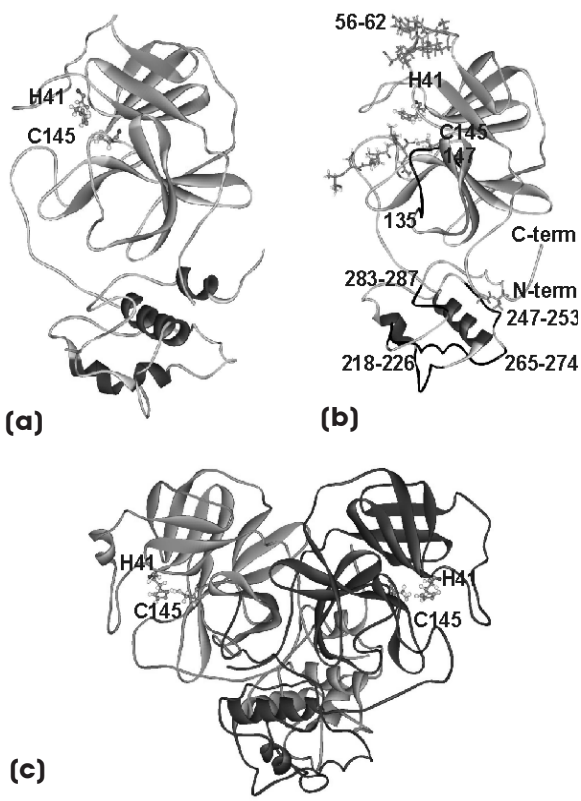


Fig 3. Schematic representation of the average MD structure calculated from those snapshots at 200 ps ≤ t ≤ 400 ps of the proteinase of the SARS coronavirus in (a) monomer form, (b) monomer complexed with the SARS-CoV substrate-analog inhibitor and (c) dimer form. Residues H41 and C145 are shown by ball-and-stick. The residue numbers given in (b) indicate significant change in RMSD (see text).

Flexibility of the SARS Coronavirus Proteinase

To ascertain in more details on the flexibility of the fragments in the enzyme structure, the atomic coordinates of all non-hydrogen atoms of the structures taken from the MD trajectory between 200 - 400 ps were used to calculate the RMSD with respect to their average structure.

Fig 4 shows the difference of the RMSDs, as a function of residue, between Mono-Free and Mono-Cpx, Mono-Cpx and Di-Free (A), and Mono-Cpx and Di-Free (B), where A and B denote chains A and B of the dimer. Therefore, a residue with a positive difference of (a) minus (b) is more flexible when it is in (a) than in (b) forms, and *vice versa*. For example, positive values in Figure 4a indicate higher flexibility of the residues of the Mono-Free over the Mono-Cpx etc.

Comparing the total positive and negative filled areas in Fig 4a, it can be clearly concluded that the monomer in free form is, in average, more flexible than the monomer in complex form. This observation can be explained by the interaction between substrate and enzyme. However, sum of the positive and negative filled regions in Figs 4b and 4c are almost equivalent. This seems to indicate that in average there is no significant difference between flexibility of monomer-substrate complex versus each chain of the dimer.

Interest is centered on the flexibility of each residue of the enzyme. Pronounced peaks, where the absolute value of the difference between the two RMSDs (filled bars) of monomer free and monomer is approximately higher than 0.5 Å, take place in the following seven residue ranges, 1-2, 56-62, 135-147, 218-226, 247-253, 265-274 and 283-287 (Fig 4a). The first two regions indicate that the N-terminus (residues 1-2) and areas around active site (residues 56-62) of the Mono-Cpx are more flexible than those of Mono-Free (see Fig 3b for the schematic representations of the two regions of the residues). This behavior is supposed to facilitate the binding between enzyme and substrate. However, the reason for such finding is not yet understood. In contrast to the first two regions, the other areas (residues 135-147, 218-226, 247-253, 265-274 and 283-287) are observed to be more rigid (the peaks are positive) when the complex is formed. Increasing of rigidity in

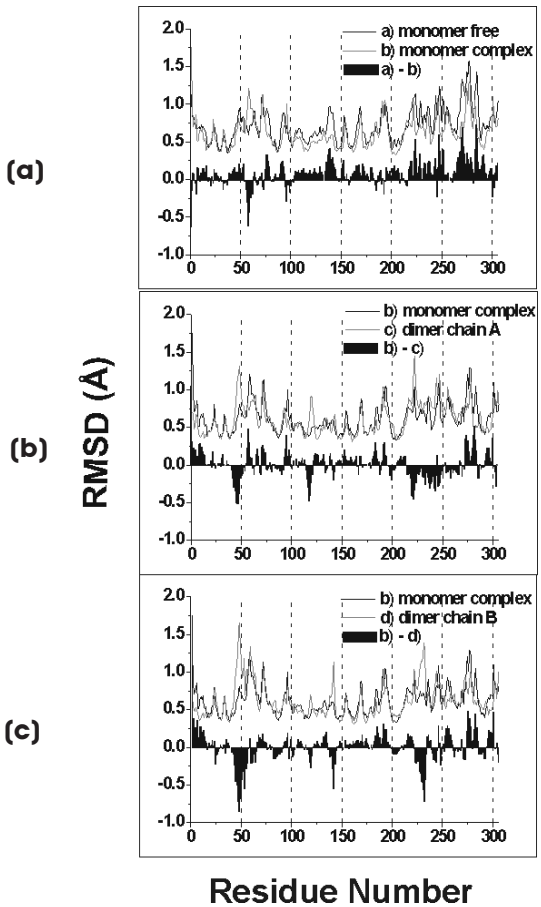


Fig 4. Comparison of RMSD (solid and dotted lines) with respect to their average structure for non-hydrogen atoms of the structures taken from the MD trajectory between 200-400 ps and their residue-wise subtraction (filled bars) for the indicated pairs of the SARS-CoV proteinase structures.

the active site regions of residues 135-147 in the complex form is due to the enzyme-substrate binding. However, description on the less flexibility of the other last four regions (218-226, 247-253, 265-274 and 283-287), which are far away from the active site (see numbers labeled in Fig 3b), cannot be made yet.

Detailed comparisons are made between the Mono-Cpx and each chain of Di-Free (Figs 4b-4c). Several pronounced peaks, both in positive and negative regions, are observed, indicating significant differences in the flexibility of each residue, especially in the active site of the SARS-CoV proteinase, between the two forms.

Flexibility of the residues can be visualized in terms of superposition of the enzyme coordinates via its mobility in the molecular dynamics simulation. Snapshots of molecular structure were taken, every 5 ps, from the MD trajectories of the free monomer form. The results are given in Fig 5 together with the RMSD plot.

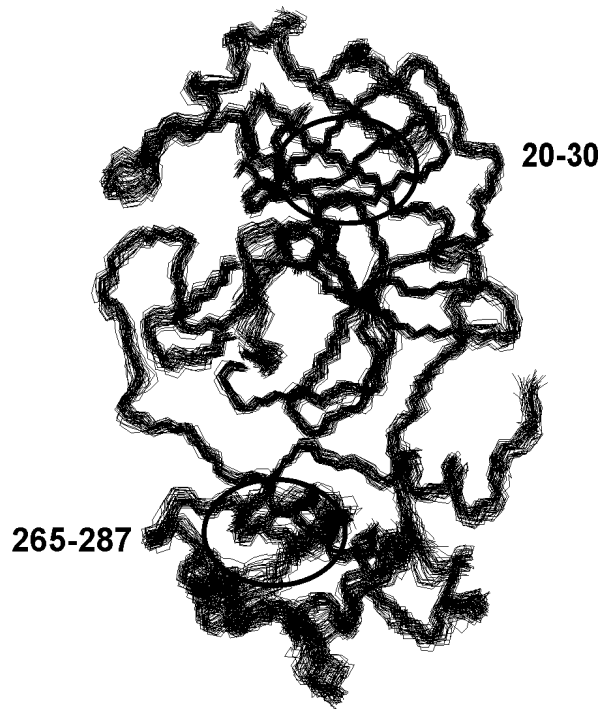


Fig 5. Superposition of the 50 structures of the proteinase of the SARS-CoV. The structures were the snapshots taken every 5 ps from the MD data. The flexible region which RMSD > 1 Å (residues 265-287) and the rigid region with RMSD < 1 Å (residues 20-30) were circled..

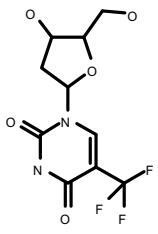
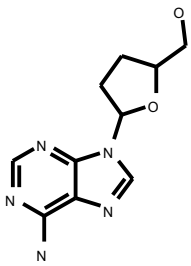
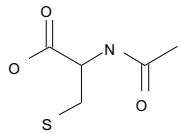
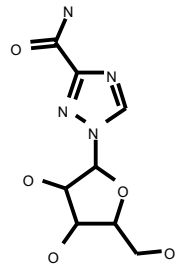
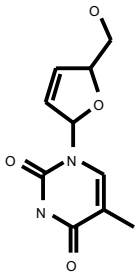
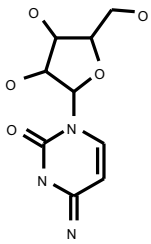
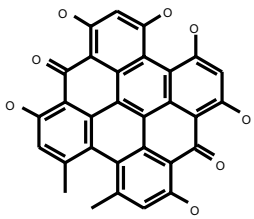
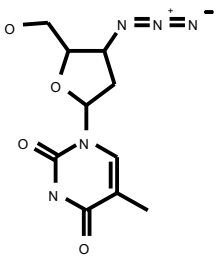
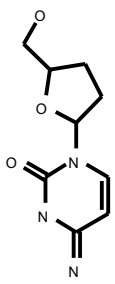
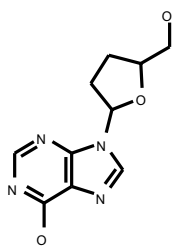
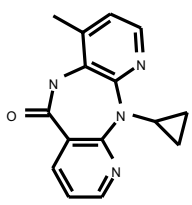
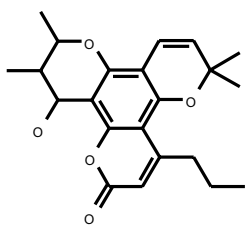
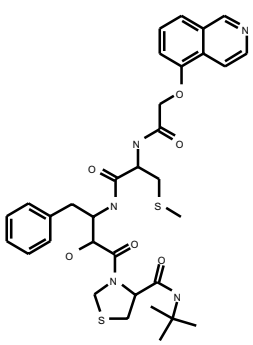
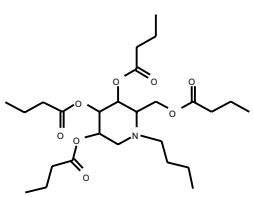
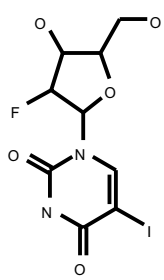
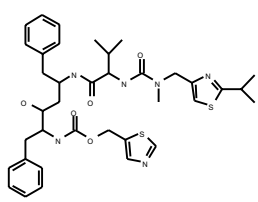
The superposition plot in Fig 5 clearly reflects the rigidity and flexibility on each part of the enzyme. It is fully consistent with the RMSD of the monomer in free form shown in Fig 4a (solid line). For example the residues 265-287 where the RMSD is higher than 1 Å and residues 20-30 where the RMSD is lower than 0.5 Å in which flexibility and rigidity of these residues were, respectively, represented by the broad and narrow distribution of the superposition plots.

Searching for Potent SARS Coronavirus Proteinase Inhibitor Using Molecular Docking Technique

By visual inspection of the docking results, 4 out of sixteen investigated antiviral drugs taken from the National Cancer Institute (NCI) database (Table 1),[18] can fit into the substrate-binding cleft of the proposed SARS-CoV proteinase while the rest are far away from the binding region. These substrate-binding schemes are in good agreement with the X-ray structure of TGEV.[7] Schematic representation of the complex is shown in Fig 6. The hydrophobic binding sites (marked as spheres) predicted by PASS[19] method are indicated for comparison.

The trade name of these four detected drugs,

Table 1. NSC numbers and structures of 16 compounds selected from NCI database.[18]

NSC number			
75520 	98700 	111180 	163039 
163661 	287459 	407313 	602670 
606170 	612049 	641530 	650886 
651714 	670880 	678514 	693184 

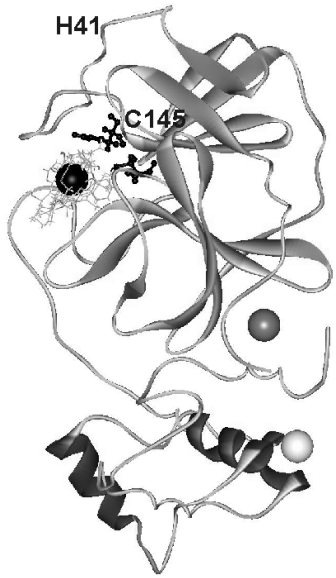


Fig 6. Schematic representation of the proteinase of the SARS-CoV complexed with the available antiviral drugs in which the configuration is predicted based on molecular docking technique. Black, gray and unfilled spheres stand for the predicted binding sites with high, medium, and low priorities, respectively

Calanolide A, Nevirapine, Virazole, and Glycovir, with the Number System Character (NSC) number and the corresponding free energy of binding are summarized in Table 2. The first three compounds are available on the market and in clinical use. The Nevirapine is known as anti HIV-1 reverse transcriptase (RT). For the Calanolide A, it has been proven to be active against HIV-1 RT and is under clinical testing. In addition, the binding site of the inhibitor observed by the molecular docking agrees well with those predicted by PASS method in which the black, gray and unfilled spheres stand for the predicted binding sites with high, medium and low priorities, respectively.

Note that the screening method used in this study is one of the most effective strategies, which is commonly used in the drug industries to discover antiviral drugs. It is a powerful tool to reduce the scope of random sampling. However, the predicted results are in a very preliminary stage. The predicted compounds show high potential to be potent drugs, *in silico*. To be in clinical use, they need to be tested *in vitro* and *in vivo*, respectively.

Conclusion

The initial 3D structure of the SARS-CoV proteinase was derived from the known X-ray structure of the porcine coronavirus proteinase using homology modeling method. It was used as a starting configuration for Molecular Dynamic Simulation of 3 forms of SARS-CoV proteinase: monomer in free form, dimer in free form and monomer complexed with SARS-CoV substrate-analog inhibitor. The simulations were carried out under the NPT ensemble at 298 K and 1 atm. At the time being, no known experimental X-ray structure of the true SARS-CoV proteinase has been reported.

Table 2. The preliminary screening for possible docking of 16 available antiviral drugs detects Calanolide A (650886), Nevirapine (641530), Virazole (163039), and Glycovir (670880) to fit well at the substrate binding cleft of our 3D structures.

NSC number	Estimated Free Energy of Binding (kcal/mol)
650886	-10.58
641530	-9.47
163039	-7.35
670880	-5.40

In the structural aspect, the monomer in free form is more flexible than the monomer in complex form while no significant difference was observed between flexibility of monomer in complex form versus each chain of the dimer in free form. Interest is centered on the two regions, at N-terminus (residues 1-2) and around the active site (residues 56-62), in which the flexibilities in complex form is lower than those in free form. This behavior is supposed to facilitate the binding between enzyme and substrate.

With the obtained MD structure, molecular dockings have been carried out in order to search for potent SARS-CoV proteinase inhibitors. The preliminary screening for possible docking of 16 available antiviral drugs detects Nevirapine, Glycovir, Virazole, and Calanolide A to fit well at the substrate binding cleft of our 3D structures. However, extensive work for each specific docking has to be done to be conclusive, even *in silico*.

Acknowledgements

This work was financially supported by the Thailand Research Fund. The structure-based sequence alignment of coronavirus proteinase was done at the Department of Pathology, Faculty of Medicine, Ramathibodi Hospital, Mahidol University. All simulations and calculations were performed on multiple Linux cluster systems where computing facilities were provided by the National Electronics and Computer Technology Center, the Computing Center for Advance Research at the Faculty of Science, Chulalongkorn University, the Computational

Chemistry Unit Cell at Chulalongkorn University, and the Chemistry Department, Faculty of Science, Kasetsart University.

REFERENCES

1. Cumulative number of reported probable cases of SARS (http://www.who.int/csr/sars/en/).
2. WHO collaborative multi-center research project on Severe Acute Respiratory Syndrome (SARS) diagnosis (http://www.who.int/csr/sars/ project/en/).
3. Marra MA, Jones SJ, Astell CR, Holt RA, Brooks-Wilson A, Butterfield YS, Khattra J and Asano JK, et al (2003) The genome sequence of the SARS-associated coronavirus. *Science* **300**, 1399-1404.
4. Rota PA, Oberste MS, Monroe SS, Nix WA, Campagnoli R, Icenogle JP, Penaranda S and Bankamp B, *et al*. (2003) Characterization of a novel coronavirus associated with Severe Acute Respiratory Syndrome. *Science* **300**, 1394-9.
5. Fouchier RA, Kuiken T, Schutten M, Van Amerongen G, Van Doornum GJ, Van Den Hoogen BG, Peiris M, Lim W, Stohr K, Osterhaus AD (2003) Aetiology: Koch's postulates fulfilled for SARS virus. *Nature* **423**, 240.
6. Holmes KV (2003) SARS-Associated Coronavirus. *N Engl J Med* **20**, 1948.
7. Anand K, Ziebuhr J, Wadhwani P, Mesters JR, Hilgenfeld R. (2003) Coronavirus main proteinase (3CLpro) structure: Basis for design of anti-SARS drugs. *Science* **300**, 1763-7.
8. Clarke T (2003) SARS' Achilles' heel revealed, *Science update* (http://www.nature.com/nsu/030505/030505-4.html)
9. Lee VS, Parasuk V, Sompornpisut P, Chantratita W, Sangma C, Vannarat S, Srichaikul P and Hannongbua S, *et al.* (2003) Three Dimensional structure of dimer of SARS proteinase: An appropriate model for design and screen for anti-SARS drugs, *Submitted*.
10. Combet C, Jambon M, Deléage G, Geourjon C (2002) Gen3D: Automatic comparative molecular modelling of protein. *Bioinformatics* **18**, 213.
11. Berendsen HJC, Postma JMP, van Gunsteren WF, DiNola A, Haak JR (1984) Molecular dynamics with coupling to an external bath. *J Comp Phys* **81**, 3684.
12. Case DA, Pearlman DA, Caldwell JW, Cheatham III TE, Wang J, Ross WS, Simmerling C and Darden T, et al (2002) AMBER7, University of California, San Francisco.
13. Pearlman, DA, Case DA, Caldwell J W, Ross WS, Cheatham III TE, DeBolt S, Ferguson D, Seibel G, Kollman P, (1995) AMBER, a package of computer programs for applying molecular mechanics, normal mode analysis, molecular dynamics and free energy calculations to simulate the structural and energetic properties of molecules. *Comp Phys Commun* **91**, 1.
14. Goodsell, DS, Olson AJ, (1990) Automated docking of substrates to proteins by simulated annealing. *Proteins* **8**, 195.
15. Morris GM, Goodsell DS, Halliday RS, Huey R, Hart WE, Belew RK, Olson AJ (1998) Automated docking using a Lamarckian Genetic Algorithm and Empirical Binding Free Energy Function. *J Comp Chem* **19**, 1639.
16. Morris GM, Goodsell DS, Huey R, Olson AJ, (1996) Distributed automated docking of flexible ligands to proteins: Parallel applications of AutoDock 2.4. *JCAMD* **10**, 293.
17. Laskowski RA, Mac Arthur MW, Moss DS, Thornton JM, (1993) Procheck-a program to check the stereochemical quality of protein structures. *J Appl Cryst* **26**, 283-291.
18. National Cancer Institute (NCI) Database (http://cactus.nci.nih.gov/ncidb2/).
19. Brady Jr. GP, Stouten PFW, (2000) Fast prediction and visualization of protein binding pockets with PASS. *JCAMD* **14**, 383. (http://www.delanet.com/~bradygp/pass/).