โครงการ การศึกษาโครงสร้างและไดนามิกส์ของสารยับยั้งเชื้อ SARS coronavirus
Structure and Dynamic Studies of SARS Coronavirus Inhibitors

โดย ดร. วีณา นุกูลการ และคณะ

รายงานวิจัยฉบับสมบูรณ์

โครงการ การศึกษาโครงสร้างและไดนามิกส์ของสารยับยั้งเชื้อ SARS coronavirus
Structure and Dynamic Studies of SARS Coronavirus Inhibitors

คณะผู้วิจัย	สังกัด
1. ดร.วีณา นุกูลการ	คณะเภสัชศาสตร์ มหาวิทยาลัยมหิดล
2. ศ.ดร. สุพจน์ หารหนองบัว	คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
3. ดร. วรรณจันทร์ แสงหิรัญ ลี	คณะวิทยาศาสตร์ มหาวิทยาลัยเชียงใหม่
4. นางสาวมธุรส มาลัยศรี	คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

โครงการ การศึกษาโครงสร้างและไดนามิกส์ของสารยับยั้งเชื้อ **SARS coronavirus**
Structure and Dynamic Studies of SARS Coronavirus Inhibitors

Abstract

SARS-CoV main protease is one of the important targets for drug development. To date, many inhibitors are investigating by many researchers but no effective drugs are established. However, it was found that a mixture of HIV-I proteinase inhibitors, lopinavir and ritonavir, exhibits signs of effectiveness against the SARS virus. To understand the dynamics behaviors in the hope that comparative analysis of structural details would provide information to future design of new potent SARS-CoV selective inhibitors, the molecular dynamic (MD) simulations of enzyme complexed with inhibitors; ritonavir and lopinavir, were carried out. The flexibility of the inhibitors in the binding region was discussed. The results show that flap closing was clearly observed when inhibitors bind to the active site of SARS-CoV. Six hydrogen bonds were detected in the SARS-LPV system while seven hydrogen bonds were found in SARS-RTV system. In addition, ritonavir was observed to fit better to the SARS-CoV 3CLpro cavity than the lopinavir.

บทคัดย่อ

เอนไซม์ซาร์สโคโรนาไวรัส โปรติเอส เป็นหนึ่งในเอนไซม์เป้าหมายสำคัญสำหรับการพัฒนายาต้านโรคซาร์ส ถึงปัจจุบันมีการศึกษาหาตัวยับยั้งที่ออกฤทธิ์กับเอนไซม์โปรติเอสนี้หลายชนิดโดยนักวิจัยหลากหลายกลุ่ม แต่ถึงขณะนี้ ยังไม่มียาที่ใช้ในการรักษาโดยตรง อย่างไรก็ตาม มีรายงานการวิจัยว่า การใช้ยาผสมริโทนาเวียและโลปินาเวีย แสดงผลที่ดีในทางคลินิก และเพื่อความเข้าใจข้อมูลพื้นฐานทางโครงสร้างของยาและเอนไซม์ งานวิจัยนี้จึงทำการศึกษาสารประกอบเชิงซ้อนของซาร์สโปรติเอสกับยาริโทนาเวียและโลปินาเวีย ด้วยวิธีจำลองกลศาสตร์เชิงพลวัต โดยหวังว่าข้อมูลและผลวิเคราะห์ทางโครงสร้างที่ได้ จะเป็นประโยชน์ทางหนึ่งสำหรับการคัดกรองและออกแบบยาเพื่อใช้ในการรักษาโรคซาร์สต่อไปในอนาคต สำหรับผลการศึกษาในส่วนของความยืดหยุ่นของโครงสร้างยาเมื่อจับกับเอนไซม์ พบว่าในส่วนของยาโลปินาเวียเกิดพันธะไฮโครเจนกับเอนไซม์จำนวนหกพันธะ ส่วนยาริโทนาเวียเกิดเจ็ดพันธะและจับกับเอนไซม์ซาร์สโปรติเอสได้ดีกว่า

Executive Summary

In this research, the evidence for the potential benefits of HIV-1 protease inhibitors against the SARS coronavirus (SARS-CoV) is provided by the flexible docking studies of five HIV-1 protease inhibitors (lopinavir, ritonavir, saquinavir, nelfinavir, indinavir) to the active sites of the main SARS-CoV protease using BioMedCache 2.0. The theoretical docking result indicates ritonavir as the compound with the lowest PMF dock scoring = -43.295 kcal/mol. Consequently, molecular dynamic simulations were performed to investigate the dynamics behaviors of three systems (SARS-Free, SARS-LPV and SARS-RTV) in the hope that comparative analysis of structural details would provide information to future design of new potent SARS-CoV selective inhibitors. Free binding affinities, structure properties and structure change in complexes, the effect of residues to the binding of their inhibitor with catalytic residues, the characteristics of the inhibitor and residues in complexes were discussed. Results indicated that ritonavir is higher fit with SARS-CoV 3CLpro cavity than lopinavir due to the flexibility of its side chain in the binding region.

1. Introduction

Severe acute respiratory syndrome (SARS) is an acute respiratory illness, which has seriously threatened global public health and socioeconomic stability. SARS was first reported in Asia in February 2003. Over the next few months, the illness spread to more than two dozen countries in North America, South America, Europe, and Asia before the SARS global outbreak. World Health Organization (WHO) reported that a total of 8,098 people worldwide became sick with SARS and of these 774 patients died. Since the spread of the new disease increases rapidly and so far no efficacious therapy is available, the useful knowledge and technology will lead to effective drugs against SARS.

The causative agent of SARS was identified to be a novel coronavirus, SARS-associated coronavirus (SARS-CoV). The genome of SARS-CoV contains 11-14 major open-reading frames and encodes several proteins, including the replicase polyproteins, S (spike protein), polymerase, M (membrane protein), N (nucleocapsid protein) and E (small envelope protein). It has been shown that all the protein functions required for SARS coronavirus replication is encoded two overlapping polyproteins (pp1a and pp1b), from which the functional proteins are released by the extensive proteolytic processing. This is primarily achieved by the 33 kDa main proteinase (Mpro), which is frequently also called 3C-like proteinase (3CLpro) to indicate a similarity in substrate specificity with the 3C proteinase of picornaviruses. The 3C proteinase of rhinovirus has been selected as a target for the development of drugs against the common cold. Similarly, the functional importance of the SARS-CoV 3CLpro in the viral life-cycle makes it an attractive target for discovering anti-SARS drugs. As a matter of fact, it would take much longer time to develop a good preventive vaccine against viruses, than to screen or improve compounds which display high potential to be a good inhibitor from available drugs or from natural products. In order to do that, computational simulation would be desirable. Recently, the crystal structure of SARS-CoV 3CLpro (pdb 1D: 1Q2W and 1UK4) was independently solved by two groups, Rao *et al*. reported the structure of SARS-CoV 3CLpro in complex with a covalently attached

substrate-analogue inhibitor, thus providing insights into the substrate binding site (Yang, H. *et al*., 2003).

Sequence alignment of the SARS-CoV 3CLpro with orthologues from other coronaviruses indicated that the enzyme is highly conserved, with 40% and 44% sequence identity, respectively, to human CoV (HCoV) 229E M^{pro} and procine transmissible gastroenteritis virus (TGEV) M^{pro} (Anand, K., *et al*., 2003). The crystal structure of the SARS-CoV 3CLpro has been determined (Yang, H., 2003). This structure, like those of other CoV M^{pro}s, comprises three domains (Figure 1). Domains I (residues 8-101) and II (residues 102-184) are β-barrels and together resemble the structure of chymotrypsin, where as domain III (residues 201-306) consists mainly of α-helices. Domain II and III are connected by a long loop (residues 185-200). The active site of the Mpro comprises a catalytic dyad consisting of the conserved residues Cys145 and His41 (Figure 2) (Huang, C., *et al*., 2004). The SARS CoV main proteinase is conserved among all the SARS coronavirus genome sequences and is highly homologous to other coronavirus 3C-like proteinase.
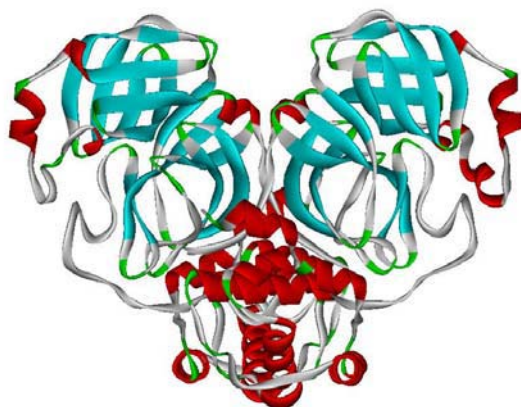

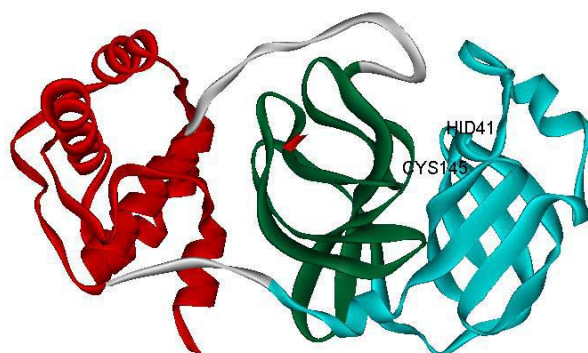**Figure1.** Structure of the dimer of SARS-CoV 3CLpro

Figure 2. Structure of the monomer of SARS-CoV 3CLpro (a) Domain I (light blue) and II (green), each contains β-barrels and Domain III (red) is composed mainly of α-helices (Tan, J., *et al*., 2005)

Availability of protein structures and the biological characteristics of 3CL^{pro}, providing it a key target for searching drug directly against SARS. Currently, neither effective antiviral drug nor vaccine is available. As its structure is similar to other proteinase, study of known proteinase inhibitors as anti-SARS drug may be useful for

starting point of treatment of SARS. Several reports indicated that HIV inhibitors have potential for designing SARS-CoV proteinase inhibitors (Zhang, X. W., *et al*., 2004, Yamamoto, N., *et al*., 2004, Jenwitheesuk, E., 2003). In particular, the proteinase inhibitor kaletra, a mixture of proteinase inhibitors-lopinavir and ritonavir, exhibits signs of effectiveness against the SARS virus (Vastag, B., 2003).

In the present research, the evidence for the potential benefits of HIV-1 protease inhibitors against the SARS coronavirus (SARS-CoV) is provided by the flexible docking studies of two HIV-1 Protease Inhibitors (lopinavir and ritonavir) to the active sites of the main SARS-CoV protease. Following, molecular dynamic simulations were performed to investigate the dynamics behaviors of three systems (SARS-Free, SARS-LPV nad SARS-RTV) (Figure 3).
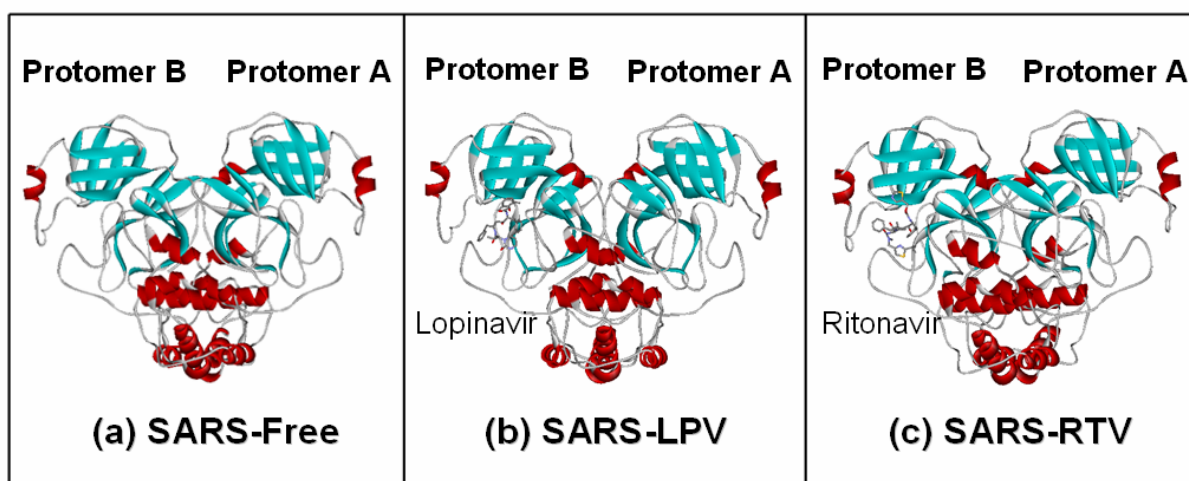

Figure 3. Structures of (a) SARS-free enzyme (b) SARS-Lopinavir complex and (c) SARS-Ritonavir complex.

2. Research objectives

The purpose of this project aims to analyze whether the existing drugs, especially HIV-1 protease inhibitors, could be the target of the SARS 3CLpro.

This project pursuer the following steps;

- Computational simulation approaches using molecular dynamics (MD) and flexible docking techniques; will be applied to analyze free binding affinities, structure properties and structure change in complexes, the effect of residues to the binding of their inhibitor with catalytic residues, the characteristics of the inhibitor and residues in complex, and the behavior of enzyme and inhibitor in aqueous solution.

- MM and PBSA module in AMBER 7.0 program and Delphi program will be investigated the relative stability and structure of inhibitor in SARS 3CLpro complexes and calculated free energy for each complex.

3. Literature review

Beginning in Guangdon province of China in late of 2002, severe acute respiratory syndrome (SARS) subsequently spread to over 25 countries and more than 700 people died (Donnelly et al., 2003). In March 2003, 4 months after the first case was reported, a new coronavirus (CoV) was identified as the causative agent of SARS (Kuiken, T., *et al*, 2003, Peiris,J. S., *et al*., 2003). The complete genome of the virus was sequence in one month later (Marra, M. A., *et al*., 2003), providing the opportunity

for more in depth studies and the identification of potential targets for drug and vaccine development. SARS-CoV particles contain a single positive stranded RNA genome that is approximately 30 kb in length and has a 5′ cap structure and 3′ poly (A) tract. The SARS-CoV genome encodes for replicase, spike, envelope, membrane, and nucleocapsid. The replicase gene encodes two large overlapping polypeptides (replicase 1a and 1ab), including 3C-like protease (3CLpro, also called main protease, M^{pro}), RNA-dependent RNA polymerase, and RNA helicase for viral replication and transcription (Ziebuhr, J., et al., 2000). The SARS-CoV 3CLpro mediates the proteolytic processing of replicase polypeptides 1a and 1ab into functional proteins, playing an important role in viral replication. Therefore, the SARS-CoV 3CLpro becomes an attractive target for developing effective drugs against SARS. The computational technique has been known to be one of the rapid and effective methods to find an attractive molecule to serve as anti-SARS drug.

There are attempts to propose three dimensional structure of SARS-CoV proteinase before the crystallographic structure released. Anand, K. et al., who solved the structure of the TGEV proteinase (PDB code: 1lvo), have produced a model for the SARS-CoV proteinase (PDB code: 1p9t and 1pa5) using the former as a template (Anan, K., et al., 2003). They also recommended using 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, AG7088 by itself has failed to inhibit the SARS-CoV in vitro (Clarke, 2003). Structure and dynamics of SARS-CoV proteinase have been investigated using molecular dynamics simulation technique by Lee et al. (Lee et al., 2003, 2004). The initial 3D structure was derived from the known X-ray structure of the porcine coronavirus proteinase using homology modeling method. However, in July 2003, first crystallographic structure of 3CLpro was released (PDB entry 1q2W by Bonnano, J. B., et al.). And a second structure was published in November in the same year (Yang, H., et al., 2003). Eleven cleavage sites of the 3CLpro on the viral polyprotein have been mapped using the computer prediction based on the substrate conservation among CoV main protease (Gao, F., et al., 2003), being confirmed by the in vitro trans-cleavage of 11 substrate peptides (Fan, K., et al., 2004). The active site of the SARS-Co V 3CLpro contains a catalytic dyad defined by His41 and Cys145, which is similar to the arrangement found in other coronavirus protease. The catalytic dyad is located in a predominantly β-structure (residues 1-184) reminiscent of the two domain fold found in the chymotrypsin family. Residues 201-303 form a third compact α-helical domain connected to the catalytic site by a long loop region (residues 185-200).

Structural conclusions from active site similarity within the coronavirus family and virtual screening on homology models have provided some clues regarding the class of compounds that could interact with SARS protease. Virtual screening on a 3D model of the SARS 3CLpro against a database of 73 protease inhibitors shows that available protease inhibitors could provide clues toward anti-SARS drug design (Xiong, B., et al., 2003). Molecular dynamics and docking studies using a database of 29 FDA approved compounds suggested that L-700417, a pseudo C2 symmetric HIV protease inhibitor, fits well into the active site compared to AG7088 (Jenwitheesuk, E. and Samudrala, R., 2003). Bifunctional aryl boronic acid compounds targeting the cluster of serine residues (Ser139, Ser144, and Ser147) near the active site cavity were found to be effective protease inhibitors (Bacha, U., et al., 2004). Researchers revisited the old adage "old drugs for new bugs" and they have shown that a few old drugs could be used as templates for designing SARS 3CLpro inhibitors (Zhang and Yap, 2004)

References:
Anand, K., Ziebuhr, J., Wadhwani, P., Mesters, J.R., and Hilgenfeld, R.. Coronavirus main protease (3CLpro) structure: Basis for design of anti-SARS drugs. *Science* 2003; 300: 1763-1767.

Bacha, U., Barrila, J., Velazquez-Campoy, A., Leavitt, SA., Freire, E. Identification of novel inhibitors of the SARS coronavirus protease 3CL(pro). *Biochemistry* 2004; 43: 4906-4912.

Case, D. A., Pearlman, J. W., Caldwell, T. E., Cheatham, J. III., Wang, W. S., Ross, C. L., Simmerling, T. A., Darden, K. M., Merz, R. V., Stanton, A. L., Cheng, J. J., Vincent, M., Crowley, V., Tsui, H., Gohlke, R. J., Radmer, Y., Duan, J., Pitera, I., Massova, G. L., Seibel, U. C., Singh, P. K., and Kollman, P. A. AMBER. Version 7.0 ed, University of California, San Francisco, CA. (2002).

Clarke, T. SARS' Achilles' heel revealed. *Science update* 2003; http://www.nature.com/nsu /030505/030505-4.html)

Donnelly, C.A., Ghani, A.C., Leung, G.M., Hedley, A.J., Fraser, C., Rieley, S., Abu-Raddad, L.J., Ho, L.M., Thach, T.Q., Chau, P. Epidermiological determinants of spread of causal agent of sevre acute respiratory syndrome in Hong Kong. *Lancet* 2003; 361: 1761-1766.

Fan, K., Wei, P., Feng, Q., Chen, S., Huang, C., Ma., L., Lai, B., Pei, J., Liu, Y., Chen, J., and Lai, L. *J. Bio.Chem.* 2004; 279: 1637-1642.

Gao, F., Ou, H.Y., Chen, L.L., Zheng, W.X., and Zhang, C.T. Prediction of proteinase cleavage sites in polyprotein of coronaviruses and its applications in analyzing SARS-CoV genomes. *FEBS Lett.* 2003; 553: 451-456.
Huang, C., Wei, P., Fan, K., Liu, Y., and Lai, L. 3C-like proteinase from SARS coronavirus catalyzes substrate hydrolysis by a general base mechanism. *Biochemistry*. 43, 4568-4574 (2004).

Jenwitheesuk, E. and Samudrala, R. Identifying inhibitors of the SARS coronavirus proteinase. *Bioorg. Med. Chem. Lett.* 2003; 13: 3989-3992.

Kuiken, T., Fouchier, R.A., Schutten, M., Rimmelzwaan, G.F., van Amerongen, G., van Riel, D., Laman, J.D., de Jong, T., van Doornum, G., Lim, W. Newly discovered coronavirus as the primary cause of severe acute respiratory syndrome. *Lancet* 2003; 362: 263-270.

Lee, V. S., Wittayanarakul, K., Remsungnen, T., Parasuk, V., Sompornpisut, P., Chantratita, W., Sangma, J., Vannarat, S., Srichaikul, P., Hannongbua, S., Saparpakorn, P., Treesuwan, W., Aruksakulwong, O., Pasomsub, E., Promsri, S., Chuakheaw, D., Hannongbua, S. Structure and Dynamics of SARS Coronavirus Proteinase: The Primary Key to the Designing and Screening for Anti-SARS Drugs *ScienceAsia* 2003; 29: 181-188.

Lee, V. S., Wittayanarakul, K., Remsungnen, T., Parasuk, V., Sompornpisut, P., Chantratita, W., Sangma, J., Vannarat, S., Srichaikul, P., Hannongbua, S., Saparpakorn, P., Treesuwan, W., Aruksakulwong, O., Pasomsub, E., Promsri, S., Chuakheaw, D.,

Hannongbua, S. "Three Dimensional Structure of SARS Protease of Coronavirus and Its Complex with SARS Substrate: Molecular Dynamics Simulations", 2004 *submitted*.

Marra, M.A., Jones, S.J., Astell, C.R., Holt, R.A., Brooks-Wilson, A., Butterfield, Y.S., Khattra, J., Asano, J.K., Barber, S.A., Chan, S.Y., et al. The genome sequence of the SARS-associated coronavirus. *Science* 2003; 300: 1399-1404.

Muegge, Y. C. Martin. A General and Fast Scoring Function for Protein-Ligand Interactions: A Simplified Potential Approach, *J. Med. Chem.*, **42**, 791-804 (1999).

Peiris, J.S., Chu, C.M., Cheng, V.C., Chan, K.S., Hung, I.F., Poon, L.L., Law, K.I., Tang, B.S., Hon, T.Y., Chan, C.S., Chan, K.H., Ng, J.S., Zheng, B.J., Ng, W.L., Lai, R.W., Guan, Y., and Yeun, K.Y. et al. Clinical progression and viral load in a community outbreak of coronavirus-associated SARS pneumonia: a prospective study. *Lancet* 2003; 361: 1767-1772.

Ryckaert, J. P., Ciccotti, G., and Berendsen, H. J. C. Numerical integration of the Caresian equations of motion of a system with constaints: molecular dynamics of n-alkanes. *J. Comput. Phys*. **23**, 327-341 (1977).

Tan, J., Verschueren, K. H.G., Anand, K., Shen, J., Yang, M., Xu, Y., Rao, Z., Bigalke, J., Heisen, B., Mesters, J. R., Chen, K., Shen, X., Jiang, H., Hilgenfeld, R. pH-Dependent conformational flexibility of the SARS-CoV main proteinase (Mpro) dimmer: molecular dynamics simulations and multiple X-ray structure analyses. *J. Mol. Biol*. **354**, 25-40 (2005).

Vastag, B. Old drugs for a new bug. *JAMA*. 2003; 290: 1695-1696.

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., Shin, T.L., Li, Y.X., Chen, K.X., Luo, X.M., Shen, X., Shen, J.H., and Jiang, H.L. A 3D model of SARS-CoV 3CL proteinase and its inhibitors design by virtual screening. *Acta Pharmacol*. *Sin*. 2003; 24: 497-504.

Yamamoto, N., Yang, R., Yoshinaka, Y., Amari, S., Nakano, T., Cinatl, J., Rabenay, H., Doerr, H. W., Hunsmann, G., Otaka, A., Tamamura, H., Fujii, N., Yamamoto, N. HIV protease inhibitor nelfinavir inhibits replication of SARS-associated coronavirus. *Biochem. Biophys. Res. Comm*. 2004; 318: 719-725.

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. The crystal structures of severe acute respiratory syndrome virus main protease and its complex with an inhibitor. *Proc. Natl. Acad., Sci. U.S.A.* 2003; 100: 13190-13195.

Zhang, X.W., and Yap, YL. Old drugs as lead compounds for a new disease? Binding analysis of SARS coronavirus main protease with HIV, psychotic and parasite drugs. *Bioorg. Med. Chem.* 2004; 12: 2517-2521.

Ziebuhr, J., Snijder, E.J., and Gorbalenya, A.E. Virus-encoded proteinases and proteolytic processing in the Nidovirales. *J. Gen. Virol.* 2000; 81: 853-879.
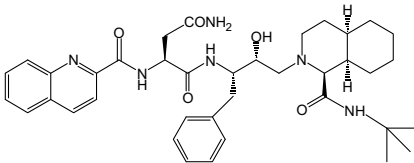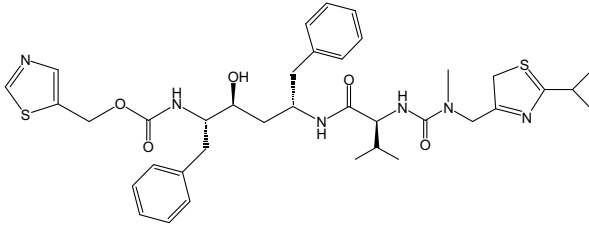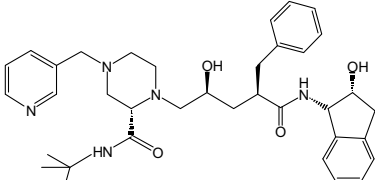
4. Methodology

4.1 Starting Structure and Protein Preparation

In order to understand ligand-enzyme interactions as well as effects of inhibitor on the structure and dynamics properties of free enzyme, three MD simulations were performed, SARS-CoV 3CL^{pro} free enzyme (SARS) and its complexes with lopinavir (SARS-LPV) and ritonavir (SARS-RTV). X-ray structures of the SARS-CoV 3CL^{pro} (PDB code 1UK3; 2.4 Å resolution) was used as initial model to construct the SARS-RTV and SARS-LPV complexes (Figure 1). Prior to the simulations, residues SerA1 and GlyA2, which had not been visible in the electron density maps, were modeled using the LEaP module in the AMBER 7 software package. The modeled conformation at the beginning of each simulation was similar to that of residues SerB1 and GlyB2 of the other polypeptide chain in the dimer, which exhibit well defined electron density. To incorporate the solvent and counter ions into the models, each system was neutralized with the counter ions and solvated with the TIP3P water model. The total atoms were 77866, 77966 and 86788 for the SARS, SARS-LPV and SARS-RTV, respectively. The crystallographic waters were also kept in the simulation box.

4.2 Flexible docking study

The flexible molecular docking with Genetic Algorithm was performed using BioMedCache 2.0 Software to find the most favorable binding interaction. Five drugs, saquinavir, ritonavir, indinavir, nelfinavir, and lopinavir were docked to SARS-CoV 3CLpro binding site defined at HID41 and CYS145 of chain B. The structure of SARS-CoV 3CLpro was obtained from X-ray crystallographic structure (PDB ID: 1UK3) [Anand, K. *et al*., 2003]. The structures of the inhibitors were generated from X-ray structured and optimized by semiempirical AM1 method. The PMF scores of the drugs were evaluated by genetic algorithm with population size of 50, crossover rate of 0.80, elitism of 5, mutation rate of 0.2 and the maximum cycles generation is set to be 40000. The size of the grid box is 25Å x 25Å x 25Å. Finally, the complex structures were analyzed and the interaction energy between the ligand and protein was calculated.

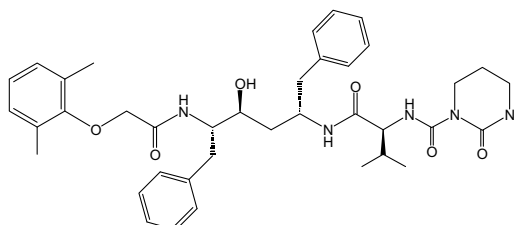| Generic name	Brand name	Chemical structures
Saquinavir	Invirase® Fortovase®	
Ritonavir	Norvir®	
Indinavir	Crixivan®	

Nelfinavir   Viracept®
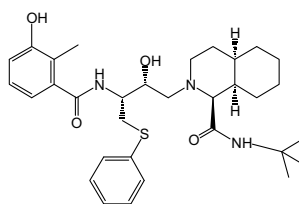
Lopinavir   Keletra®
                   (LPV-r)

4.3 Molecular dynamics simulations

Three MD simulations (SARS, SARS-RTV and SARS-LPV) were carried out using AMBER 7 simulation package. Energy minimization and MD simulations were achieved using the SANDER module of AMBER 7 with the Cornell force field. The parm99 force field which describes the structures and conformations of compounds, were applied. All MD run reported here was achieved under an isobaric-isothermal ensemble (NPT) using a constant pressure of 1 atm and a constant temperature of 298 K. The SHAKE algorithm was used to constrain all bonds involving hydrogens. Energy minimizations were carried out to relax the system prior to MD runs. Sodium and chloride ions were added to neutralize the system. The simulation time step was set at 2 femtosecond (fs). The simulation consists of thermalization, equilibration, and production phases. Initially, the temperature of the system was gradually raised to 298 K for the first 60 picosecond (ps) and then kept constant according to the Berendsen algorithm with a coupling time of 0.2 ps. The systems were maintained at 298 K until MD reached 600 ps of the simulation. Finally, the production phase started from 600 ps to 2 nanosecond (ns). The MD trajectories were collected every 0.2 ps. The 1400 ps trajectories of the production phase were used to calculate the average structure. All MD simulation was carried for 2 ns. The quality of the geometry and the stereochemistry of the protein structure were validated using PROCHECK. MD trajectories were evaluated in term of root mean square displacement (RMSD), changes of the distances, torsion angles and H-bond using the CARNAL and Ptraj modules of AMBER7.

5. Results and Discussion

5.1 Flexible docking study

The PMF score using united atom (without hydrogen atoms) and with hydrogens were shown in Fig. 4.
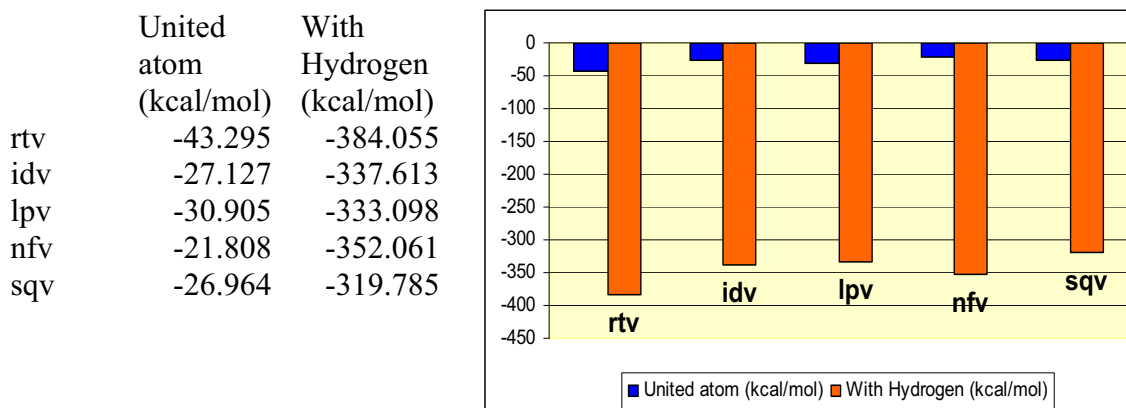
	United atom (kcal/mol)	With Hydrogen (kcal/mol)
rtv	-43.295	-384.055
idv	-27.127	-337.613
lpv	-30.905	-333.098
nfv	-21.808	-352.061
sqv	-26.964	-319.785

Figure 4. PMF score of HIV-1 protease inhibitors complex with SARS-CoV 3CL^{pro}

Because the lower PMF score is, the greater the binding affinity is, hence HIV protease drug ritonavir is the compound that bind to the substrate binding site of SARS-CoV proteinase with the highest binding affinity. Interestingly, nelfinavir which has strongly inhibited SARS-CoV replication in vitro (Yamamoto, N., *et al*., 2004), showed the lowest PMF score.

The close views of the interactions between SARS-CoV main protease and ritonavir and lopinavir are exhibited in Figure 3. The results show that the thiazole and benzene group of ritonavir are inserted into S1 and S2 specificity pocket, while the half of lopnavir is left outside the catalytic site. However, the combination of ritonavir/lopinavir decreases significantly overall death in clinical use. Taken together, our study illustrates that existing HIV-1 protease drugs may be used as starting points for the discovery of rationally designed anti-SARS proteinase drugs.
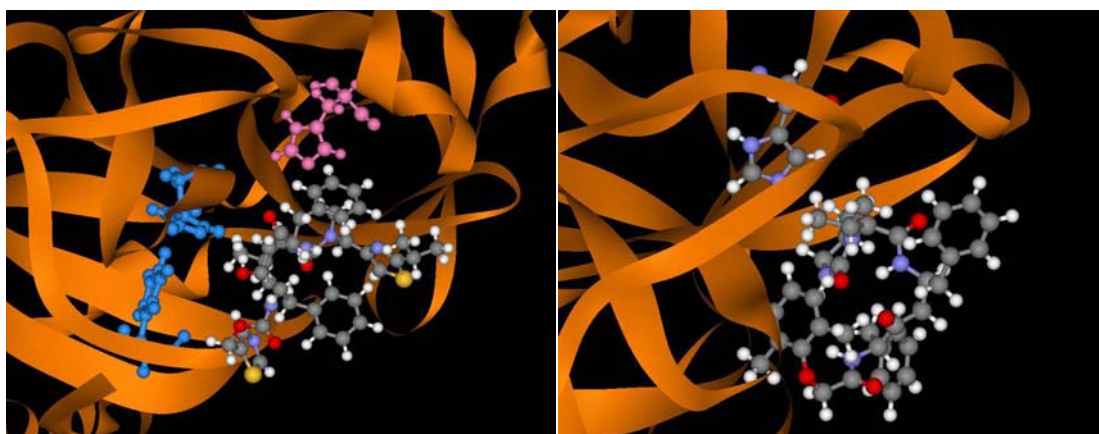

Figure 5. The close views of the interactions between SARS-CoV main protease and ritonavir (*left*) and lopinavir (*right*).

5.2 Overall enzyme and inhibitor structure

The 2 ns MD trajectories of the three simulated systems were generated and the RMSDs of the overall structure and two inhibitors with respect to the initial configuration were analyzed and plotted in Figure 6. The overall RMSD for the three

systems appears to reach equilibrium after 600 ps. While no significant different was found for the overall (black) and the back bone (dark grey) RMSDs of all plots, that of the RTV in the SARS-RTV complex (light grey) was found to change substantially. This is due to the rotation of the RTV in the SARS-CoV 3CLpro pocket, *i.e.,* initial (Docking) conformation of the RTV is totally different from that after equilibration.
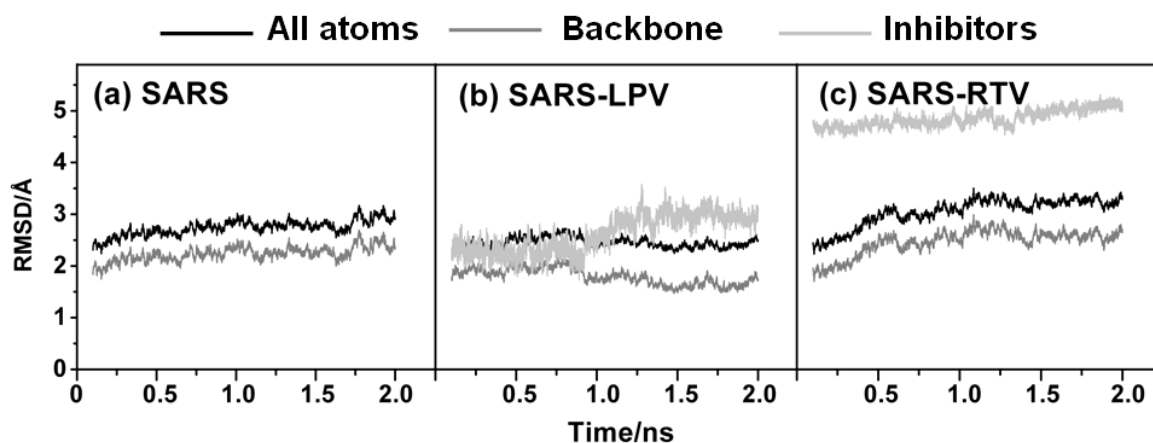

Figure 6. RMSDs of SARS-CoV 3CL^{pro} free enzyme (SARS) and its complexes with lopinavir (SARS-LPV) and ritonavir (SARS-RTV).

5.3 Flexibility of the inhibitor-binding loop regions

Consideration of protein flexibility is essential to the critical evaluation of ligand-binding affinity. Conformational changes of the catalytic binding domain play a critical role in ligand binding. Here, the ligand binding site of the SARS-CoV was evaluated focusing on the amino acid sequence around the active sites (H41 and C145) of the coronavirus proteinase. The inhibitor binding is likely to affect the relative motion between the apex of the loop with respect to D48 and Q189 loops upon the insertion of inhibitor. Figure 7, probability distributions of the associated interatomic distances from the center of mass of the following residues for the three systems were evaluated and plotted, *d1*: H41-C145 and *d2:* D48-Q189. It was shown that SARS-CoV has different dynamic flexibilities in the binding region when the SARS-selective inhibitors are bound to their active sites. In SARS-CoV we found that the most probable distance between H41 and C145 decreases from a sharp peak around 10.0 Å in the free form to 7.1–7.3 Å in the enzyme–inhibitor complexes. No significant change in the distance between D48 and Q189 loop region between free and complex forms was observed since the SARS-CoV free from were initially generated from the complexed SARS-CoV structure (PDB ID: 1UK3 which identified as the close loop conformation. This kind of flap (loop) - closing due to inhibitor binding was also observed when inhibitors bind in the active site of HIV-1 protease. Higher *d1* and *d2* distances for SARS-RTV than SARS-LPV were observed due to the more flexibility of the longer side chain region of RTV than that of LPV. Our results can therefore implied that a potent SARS-CoV selective inhibitor, when bound to the active site in SARS-CoV, should bind tightly to the catalytic site involving the closed contact between His and Cys and keep the disordered loop more motionally restricted.
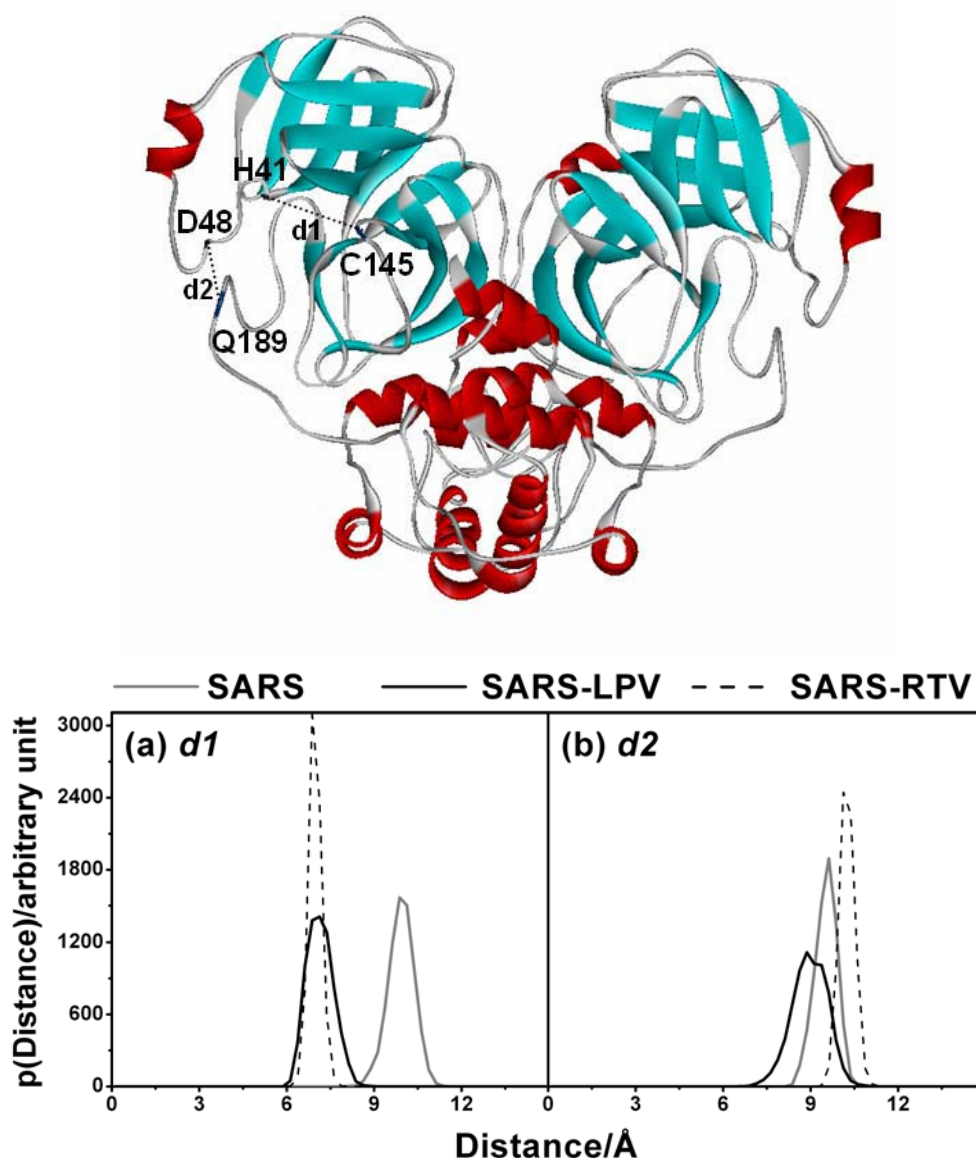
Figure 7. The probability distributions of the associated interatomic distances between the distances *d1* and *d2* from the center of mass of the following residues, H41 to C145 and D48 to Q189, respectively.

5.4 Conformations of the catalytic residues

Another structural feature common to the enzyme–inhibitor complexes is the conformations of enzyme binding site. The parameter analyzed was the changes in the dihedral angle of the key interactions of the enzyme involving the amino acids at the the catalytic residues and the tip of the loop upon inhibitor insertion. In this study we present conformational analysis of the two torsion angles of catalytic residues (TOR1: CA-CB-CG-CD of H41and TOR2: N-CA-CB-S of C145) and the two amino acids at the tip of the loop (TOR3: CA-CB-CG-OD of D48 and TOR4: CB-CG-CD-CA of Q189 as illustrated in Figure 8 during 600-2000 ps of the dynamics simulations. Table 1 summarizes the four dihedral angles. For SARS-Free, TOR1:TOR2:TOR3:TOR4 are about -90º: 60º (sharp): -60º: -100º and -150º (broad), for SARS-RTV are 150º:60º (sharp); -150º and 30º: 150º (broad), and for SARS-LPV are 60º (sharp): -60º (sharp): 150º: 90º. From this information, the torsion angle for inhibitor in binding with catalytic

residue is quite specific and more tightly bound with Cys than His observed from a single sharp peak (TOR2) and more broad peak (TOR1) indicated more loosely bound with His. It has been shown that the specific torsion angle of both inhibitors in binding with Cys were at -60º where as in the free form was at 60º. RTV and LTV bound slightly different to His catalytic site with the torsion angle (TOR1) of 90º and 150º, respectively. The torsion angle at the side chain at the tip of the loop is quite varies with broad peaks and/or with two probably distribution torsions.

Table 1. The four dihedral angles of TOR1: CA-CB-CG-CD of H41, TOR2: N-CA-CB-S of C145, TOR3: CA-CB-CG-OD of D48, and TOR4: CB-CG-CD-CA of Q189

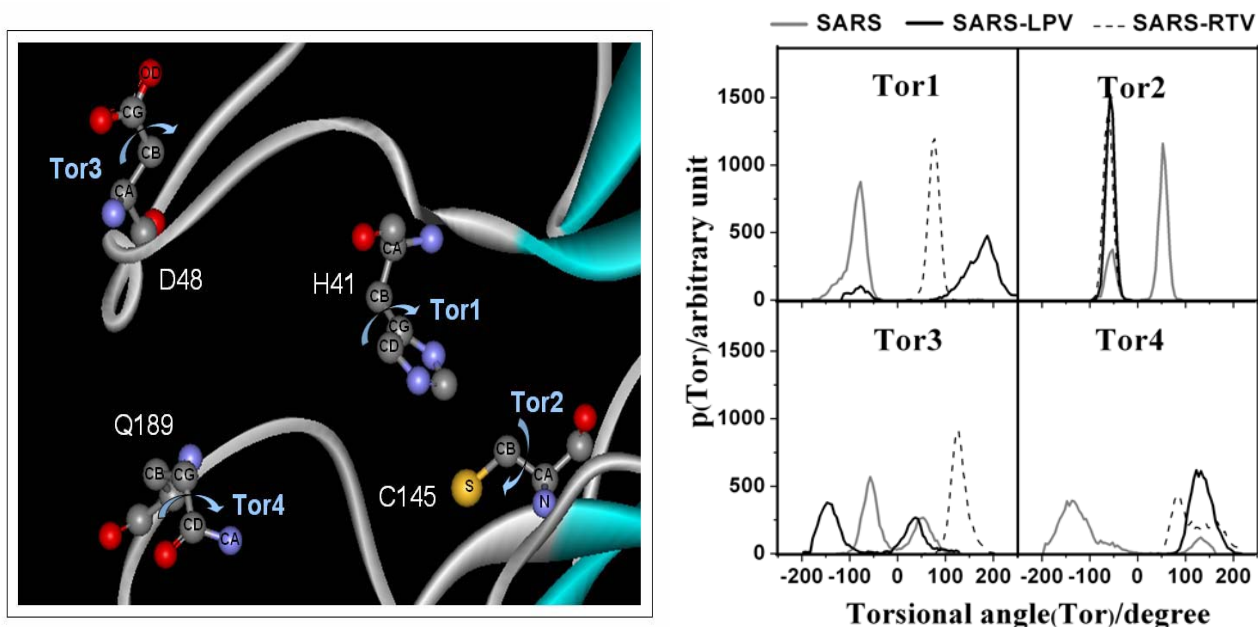|	TOR1(degree)	TOR2(degree)	TOR3(degree)	TOR4 (degree)
SARS	-77.5	-52.5, 52.5	-57.5, 52.5	-137.5, 132.5 (broad)
SARS-LPV	-77.5, 187.5	-57.5 (sharp)	-147.5, 37.5	132.5 (broad)
SARS-RTV	77.5 (sharp)	-62.5 (sharp)	127.5	82.5

Figure 8. The probability distribution of the two torsion angles of catalytic residues (TOR1: CA-CB-CG-CD of H41and TOR2: N-CA-CB-S of C145) and the two amino acids at the tip of the loop (TOR3: CA-CB-CG-OD of D48 and TOR4: CB-CG-CD-CA of Q189) during 600-1700 ps of the dynamics simulations.

5.5 Inhibitor-Enzyme Interaction

The SARS-LPV and SARS-RTV interactions were analyzed by means of hydrogen bond interaction and the molecular mechanical energy obtained from the electrostatic and the van der Waals interactions within the systems. The results were collected in Table 2.

To analyze hydrogen bond interaction, percentage and number of hydrogen bonding between the inhibitor and binding pocket residues were determined based on the following criteria: *(i)* proton donor-acceptor distance ≤ 3.5 Å and *(ii)* donor-H-acceptor bond angle ≥ 120°. In the MD simulations, the hydrogen bonds for the two

systems are almost comparable. Six hydrogen bonds were detected in the SARS-LPV system while seven hydrogen bonds were found in SARS-RTV system.

However, the differences of inhibitors's structure between lopinavir and ritonavir can directly affect via the inhibitor-enzyme interaction. The number and %occupation at Cys-His catalytic dyad, Cys145 and His41, were detected for the SARS-RTV whereas in the SARS-LPV system was lost (Table 2 and Figure 9). Referred to the inhibitor structure, the longer side chains of ritonavir in SARS-RTV can interact with the neighbor residue in active site more than lopinavir in SARS-LPV system. Consequently, the SARS-RTV system was observed to interact more tightly to the important residues in the catalytic site of SARS-CoV protease than the SARS-LPV system.

It was reconfirmed in term of the gas phase MM interaction energy (ΔE_{MM}) shown in Table 3 where molecular mechanical interaction energy for the SARS-RTV complex of -51.61 kcal/mol is lower than that of -46.63 kcal/mol for the SARS-LPV. This result indicates that ritonavir structure is highly fitted with amino acid in the SARS-CoV cavity in comparison with lopinavir.

Table 2. Percentage occupation of hydrogen bonding between inhibitor and specific residues of SAR-CoV enzyme.

Residue	Type	%Occupation	
		SARS-LPV	SARS-RTV
T26	OG1_HG1---O3_LPV	18.7	-
T26	LPV_N4_H47---OG1	22.8	-
T26	LPV_N4_H47---O	36.9	-
H41	RTV_N2_H5---NE2	-	1.7
Y118	OH_HH---O5_LPV	2.8	-
N119	LPV_N4_H47---OD1	3.3	-
C145	RTV_N2_H5---SG	-	56.3
C145	SG_HG---O3_RTV	-	4.4
E166	N_H---N1_RTV	-	29.9
Q189	NE2_HE21---O1_LPV	9.9	-
Q189	NE2_HE21---N1_RTV	-	11.6
Q189	NE2_HE21---S1_RTV	-	7.6
Q189	NE2_HE21---O2_RTV	-	1.3

Table 3. Comparison of the molecular mechanical interactionenergies (kcal/mol) of the two simulated systems, SARS-LPV and SARS-RTV

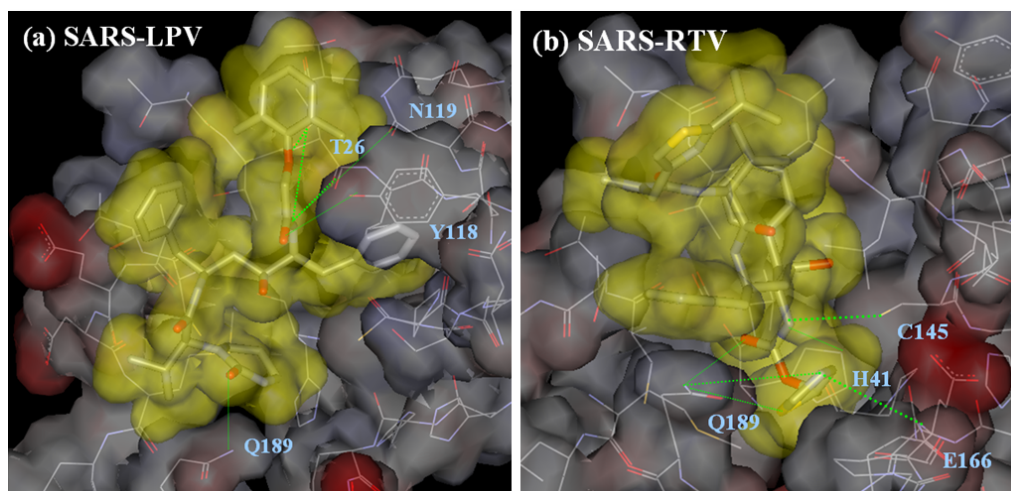|	SARS-LPV	SARS-RTV
ΔE_{ele}	-2.91 ± 4.21	-11.39 ± 3.28
ΔE_{vdw}	-43.72 ± 3.31	-40.23 ± 2.92
ΔE_{MM}	-46.63 ± 5.42	-51.61 ± 4.85

Figure 9. The Electrostatic potential and hydrogen bond in the binding pocket of the two simulated systems, SARS-LPV and SARS-RTV, where Van der Waal cavity of the inhibitor are in yellow, negative regions are in red and positive regions are in blue.

6. Conclusions

MD simulations and analysis of SARS, SARS-RTV and SARS-LPV systems give an insight into dynamic characteristics in term of the flexibility, conformation, and inhibitor-enzyme interactions. These data indicated a potent SARS-CoV selective inhibitor, when bound to the active site in SARS-CoV, should bind tightly to the catalytic site involving the closed contact between His and Cys and keep the disordered loop more motionally restricted

7. Output

- The potential compounds, HIV-1 proteinase inhibitors, inhibiting the SARS 3CLpro, were investigated in the hope that comparative analysis of structural details would provide information to future design of new potent SARS-CoV selective inhibitors

- A manuscript of 'Molecular Dynamic Simulations Analysis of Ritronavir and Lopinavir as SARS-CoV 3CL pro Inhibitor'

8. Acknowledgments

Molecular Dynamic Simulations Analysis of Ritronavir and Lopinavir as SARS-CoV 3CL ^{pro} Inhibitor

Veena Nukoolkarn[a], Vannajan Sanghiran Lee[b], Maturos Malaisree[c], Ornjira Aruksakulwong[d], Supot Hannongbua[c*]

[a]*Department of Pharmaognosy, Faculty of Pharmacy, Mahidol University, Bangkok 10400 Thailand*

[b]*Computational Simulation and Modeling Laboratory (CSML), Department of Chemistry, Faculty of Science, Chiang Mai University, Chiang Mai 50200 Thailand*

[c]*Computer Chemistry Unit Cell (CCUC) Department of Chemistry, Faculty of Science, Chulalongkorn University, Bangkok 10330 Thailand*

[d]*Department of Chemistry, Faculty of Science, Rangsit University, Pathumtani 12000 Thailand*

Abstract

Since the emerging of the severe acute respiratory syndrome (SARS), neither effective antiviral drug nor vaccine is available. However, it was found that a mixture of HIV-I proteinase inhibitors, lopinavir and ritonavir, exhibits signs of effectiveness against the SARS virus. To understand detailed interactions via complexation, molecular dynamics simulations were carried out for the SARS-CoV 3CL^{pro} free enzyme (SARS) and its complexes with lopinavir (SARS-LPV) and ritonavir (SARS-RTV). The results show that flap closing was clearly observed when inhibitors bind to the active site of SARS-CoV. Six hydrogen bonds were detected in the SARS-LPV system while seven hydrogen bonds were found in SARS-RTV system. In addition, ritonavir was observed to fit better to the SARS-CoV 3CLpro cavity than the lopinavir.

Keywords: SARS; proteinase, MD simulations, ritonavir, lopinavir

1. Introduction

Beginning in Guangdon province of China in the late of 2002, severe acute respiratory syndrome (SARS) subsequently spread to over 25 countries and more than 700 people died[1]. In March 2003, 4 months after the first case was reported, a new coronavirus (CoV) was identified as the causative agent of SARS[2,3]. The complete genome of the virus was sequenced in a month later[4], providing the opportunity for more in depth studies and the identification of potential targets for drug and vaccine development. The genome of SARS-CoV contains 11-14 major open-reading frames and encodes several proteins, including the replicase polyproteins, S (spike protein), polymerase, M (membrane protein), N (nucleocapsid protein) and E (small envelope protein). Viral proteinase and replicase are preferred targets for searching and designing of antiviral compounds. The SARS-CoV main proteinase, 3CL[pro] (also called M[pro]), exhibits a key role in proteolytic processing of the replicase polyproteins.

Sequence alignment of the SARS-CoV 3CL[pro] with orthologues from other coronaviruses indicated that the enzyme is highly conserved, with 40% and 44% sequence identity, respectively, to human CoV (HCoV) 229E proteinase and procine transmissible gastroenteritis virus (TGEV) proteinase[5]. The crystal structures of the 3CL[pro] has been determined[6-8], revealing that it is similar to those of other coronavirus proteinases. SARS-CoV 3CL[pro] forms a homodimer, comprising of three domains. Domains I (residues 8-101) and II (residues 102-184) are β-barrels and together resemble the structure of chymotrypsin, respectively, whereas domain III (residues 201-306) consists mainly of α-helices. Domains II and III are connected by a long loop (residues 185-200). The active site of the Mpro comprises a catalytic dyad consisting of the conserved residues His41 ansd Cys145, which locate at the cleft between domains I and II[9]. Availability of protein structures and the biological characteristics of 3CL[pro] encourage it to be a key target for searching drug directly against SARS.

Currently, neither effective antiviral drug nor vaccine is available. Several reports indicated that HIV-1 protease inhibitors have potential for designing SARS-CoV proteinase inhibitors[10-12]. In particular, the proteinase inhibitor kaletra, a mixture of proteinase inhibitors, lopinavir and ritonavir, exhibits signs of effectiveness against the SARS virus[13]. Therefore, understanding of known proteinase inhibitors as anti-SARS drug can be starting point to design and discover inhibitors for treatment of SARS.

In the present research, the evidence for the potential benefits of HIV-1 protease inhibitors against the SARS coronavirus (SARS-CoV) is provided by the flexible docking studies of the two HIV-1 Protease Inhibitors (lopinavir and ritonavir) in to the active sites of the main SARS-CoV protease. Following, molecular dynamic simulations were performed to

investigate the structure and dynamics behaviors of free enzyme and its complexes with lopinavir and ritonavir inhibitors (Figure 1).
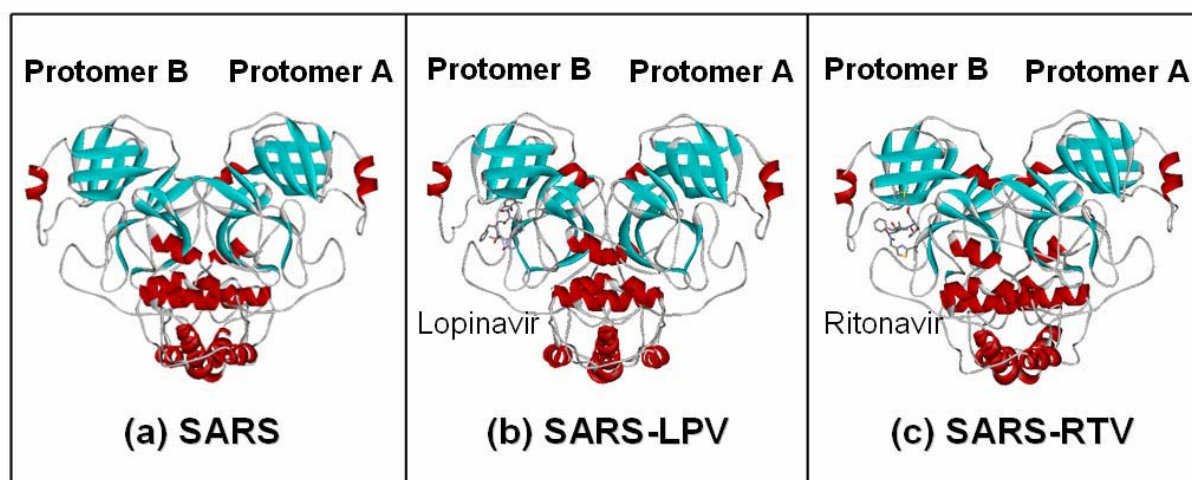

Figure 1. Structures of (a) SARS-free enzyme, (b) SARS-lopinavir complex and (c) SARS-ritronavir complex.
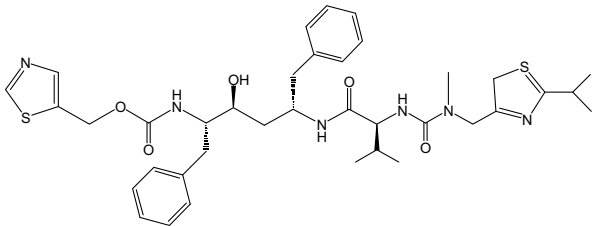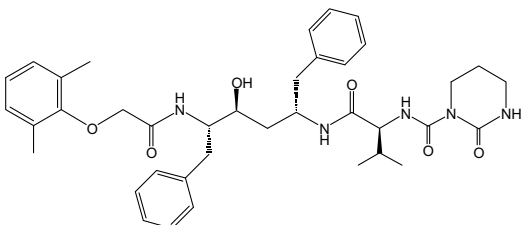
2. Methodology

2.1 Starting Structure and Protein Preparation

In order to understand ligand-enzyme interactions as well as effects of inhibitor on the structure and dynamics properties of free enzyme, three MD simulations were performed, SARS-CoV 3CL[pro] free enzyme (SARS) and its complexes with lopinavir (SARS-LPV) and ritonavir (SARS-RTV). X-ray structures of the SARS-CoV 3CL[pro] (PDB code 1UK3; 2.4 Å resolution) was used as initial model to construct the SARS-RTV and SARS-LPV complexes (Figure 1). Prior to the simulations, residues SerA1 and GlyA2, which had not been visible in the electron density maps, were modeled using the LEaP module in the AMBER 7 software package[14]. The modeled conformation at the beginning of each simulation was similar to that of residues SerB1 and GlyB2 of the other polypeptide chain in the dimer, which exhibit well defined electron density. To incorporate the solvent and counter ions into the models, each system was neutralized with the counter ions and solvated with the TIP3P water model[15]. The total atoms were 77866, 77966 and 86788 for the SARS, SARS-LPV and SARS-RTV, respectively. The crystallographic waters were also kept in the simulation box.

2.2 Flexible Docking Study

The flexible molecular docking[16] with Genetic Algorithm was performed using BioMedCache 2.0 Software to find the most favorable binding interaction. Two drugs, ritonavir and lopinavir were docked into the SARS-CoV 3CLpro binding site defined at

HID41 and CYS145 of chain B. The structure of SARS-CoV 3CLpro was obtained from X-ray crystallographic structure (PDB ID: 1UK3)[5]. The structures of the inhibitors were generated from X-ray structured and optimized by semiempirical AM1 method. The PMF scores of the drugs were evaluated by genetic algorithm with population size of 50, crossover rate of 0.80, elitism of 5, mutation rate of 0.2 and the maximum cycle generation is set to be 40000. The size of the grid box is 25Å x 25Å x 25Å. Finally, the complex structures were analyzed and the interaction energy between the ligand and protein was calculated.

Generic name	Brand name	Chemical structures
Ritonavir	**Norvir®**	
Lopinavir	**Keletra®** **(LPV-r)**	

2.3 Molecular Dynamics Simulations

Three MD simulations (SARS, SARS-RTV and SARS-LPV) were carried out using AMBER 7 simulation package. Energy minimization and MD simulations were achieved using the SANDER module of AMBER 7 (Case *et al*., 2002) with the Cornell force field[17]. The parm99 force field which describes the structures and conformations of compounds, were applied. All MD run reported here was achieved under an isobaric-isothermal ensemble (NPT) using a constant pressure of 1 atm and a constant temperature of 298 K. The SHAKE algorithm[18] was used to constrain all bonds involving hydrogens. Energy minimizations were carried out to relax the system prior to MD runs. Sodium and chloride ions were added to neutralize the system. The simulation time step was set at 2 femtosecond (fs). The simulation consists of thermalization, equilibration, and production phases. Initially, the temperature of the system was gradually raised to 298 K for the first 60 picosecond (ps) and then kept constant according to the Berendsen algorithm with a coupling time of 0.2 ps. The systems were maintained at 298 K until MD reached 600 ps of the simulation. Finally, the production

phase started from 600 ps to 2 nanosecond (ns). The MD trajectories were collected every 0.2 ps. The 1400 ps trajectories of the production phase was used to calculate the average structure. All MD simulation was carried for 2 ns. The quality of the geometry and the stereochemistry of the protein structure were validated using PROCHECK. MD trajectories were evaluated in term of root mean square displacement (RMSD), changes of the distances, torsion angles and H-bond using the CARNAL and Ptraj modules of AMBER7.

3. Results and Discussion

3.1 Overall enzyme and inhibitor structure

The 2 ns MD trajectories of the three simulated systems were generated and the RMSDs of the overall structure and two inhibitors with respect to the initial configuration were analyzed and plotted in Figure 2. The overall RMSD for the three systems appears to reach equilibrium after 600 ps. While no significant different was found for the overall (black) and the back bone (dark grey) RMSDs of all plots, that of the RTV in the SARS-RTV complex (light grey) was found to change substantially. This is due to the rotation of the RTV in the SARS-CoV 3CLpro pocket, *i.e.,* initial (Docking) conformation of the RTV is totally different from that after equilibration.
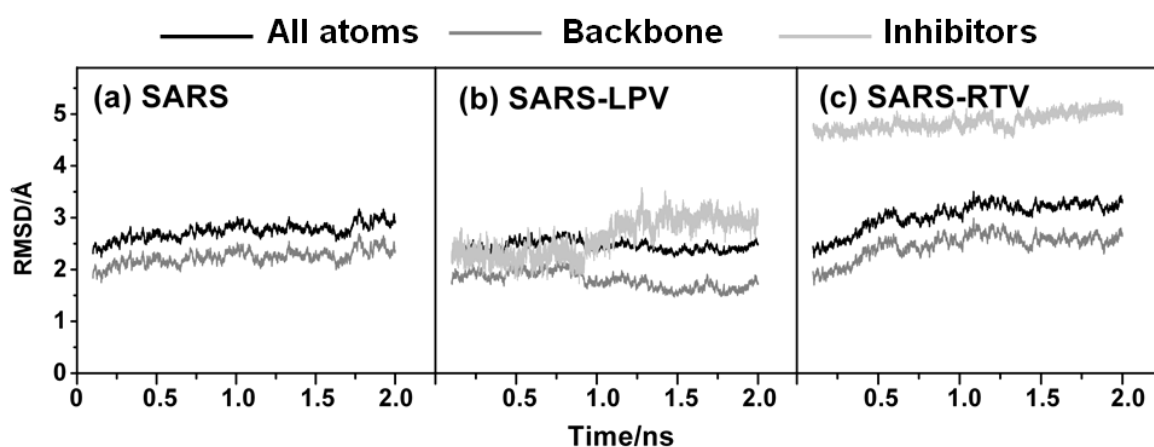

Figure 2. RMSDs of SARS-CoV 3CL^{pro} free enzyme (SARS) and its complexes with lopinavir (SARS-LPV) and ritonavir (SARS-RTV).

3.2 Flexibility of the inhibitor-binding loop regions

Consideration of protein flexibility is essential to the critical evaluation of ligand-binding affinity[19-22]. Conformational changes of the catalytic binding domain play a critical role in ligand binding. Here, the ligand binding site of the SARS-CoV was evaluated focusing on the amino acid sequence around the active sites (H41 and C145) of the coronavirus

proteinase[5,6,23]. The inhibitor binding is likely to affect the relative motion between the apex of the loop with respect to D48 and Q189 loops upon the insertion of inhibitor. Figure 3, probability distributions of the associated interatomic distances from the center of mass of the following residues for the three systems were evaluated and plotted, *d1*: H41-C145 and *d2:* D48-Q189. It was shown that SARS-CoV has different dynamic flexibilities in the binding region when the SARS-selective inhibitors are bound to their active sites. In SARS-CoV we found that the most probable distance between H41 and C145 decreases from a sharp peak around 10.0 Å in the free form to 7.1–7.3 Å in the enzyme–inhibitor complexes. No significant change in the distance between D48 and Q189 loop region between free and complex forms was observed since the SARS-CoV free from were initially generated from the complexed SARS-CoV structure (PDB ID: 1UK3 which identified as the close loop conformation. This kind of flap (loop) - closing was clearly observed when inhibitors bind in the active site of HIV-1 protease[24,25]. Higher *d1* and *d2* distances for SARS-RTV than SARS-LPV were observed due to the more flexibility of the longer side chain region of RTV than that of LPV. Our results can therefore implied that a potent SARS-CoV selective inhibitor, when bound to the active site in SARS-CoV, should bind tightly to the catalytic site involving the closed contact between His and Cys and keep the disordered loop more motionally restricted.
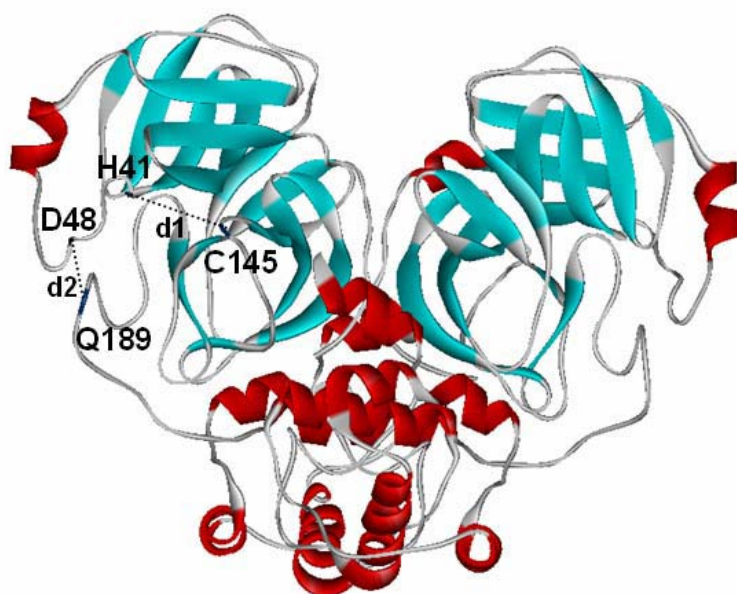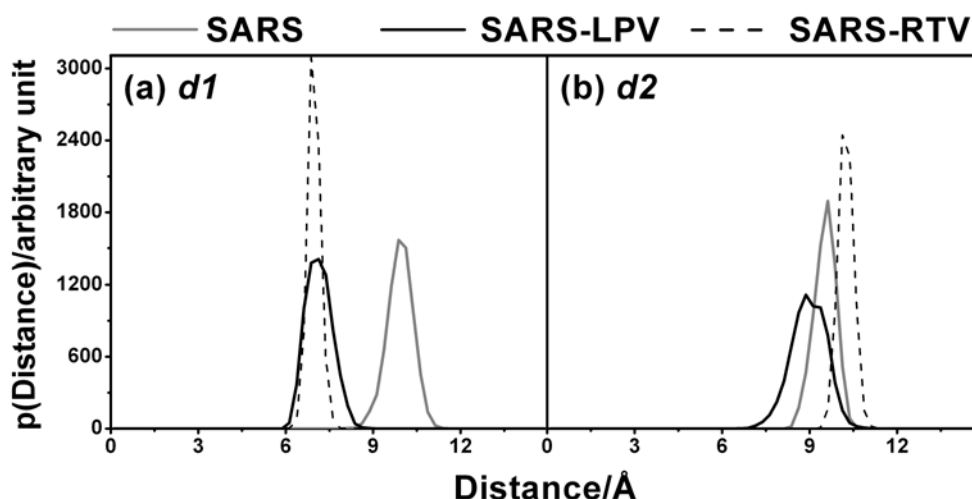
Figure 3. The probability distributions of the associated interatomic distances between the distances *d1* and *d2* from the center of mass of the following residues, H41 to C145 and D48 to Q189, respectively.

3.3 Conformations of the catalytic residues

Another structural feature common to the enzyme–inhibitor complexes is the conformations of enzyme binding site. The parameter analyzed was the changes in the dihedral angle of the key interactions of the enzyme involving the amino acids at the the catalytic residues and the tip of the loop upon inhibitor insertion. In this study we present conformational analysis of the two torsion angles of catalytic residues (TOR1: CA-CB-CG-CD of H41and TOR2: N-CA-CB-S of C145) and the two amino acids at the tip of the loop (TOR3: CA-CB-CG-OD of D48 and TOR4: CB-CG-CD-CA of Q189 as illustrated in Figure 4 during 600-2000 ps of the dynamics simulations. Table 1 summarizes the four dihedral angles. For SARS-Free, TOR1:TOR2:TOR3:TOR4 are about -90º: 60º (sharp): -60º: -100º and -150º (broad), for SARS-RTV are 150º:60º (sharp); -150º and 30º: 150º (broad), and for SARS-LPV are 60º (sharp): -60º (sharp): 150º: 90º. From this information, the torsion angle for inhibitor in binding with catalytic residue is quite specific and more tightly bound with Cys than His observed from a single sharp peak (TOR2) and more broad peak (TOR1) indicated more loosely bound with His. It has been shown that the specific torsion angle of both inhibitors in binding with Cys were at -60º where as in the free form was at 60º. RTV and LTV bound slightly different to His catalytic site with the torsion angle (TOR1) of 90º and 150º, respectively. The torsion angle at the side chain at the tip of the loop is quite varies with broad peaks and/or with two probably distribution torsions.

Table 1. The four dihedral angles of TOR1: CA-CB-CG-CD of H41, TOR2: N-CA-CB-S of C145, TOR3: CA-CB-CG-OD of D48, and TOR4: CB-CG-CD-CA of Q189

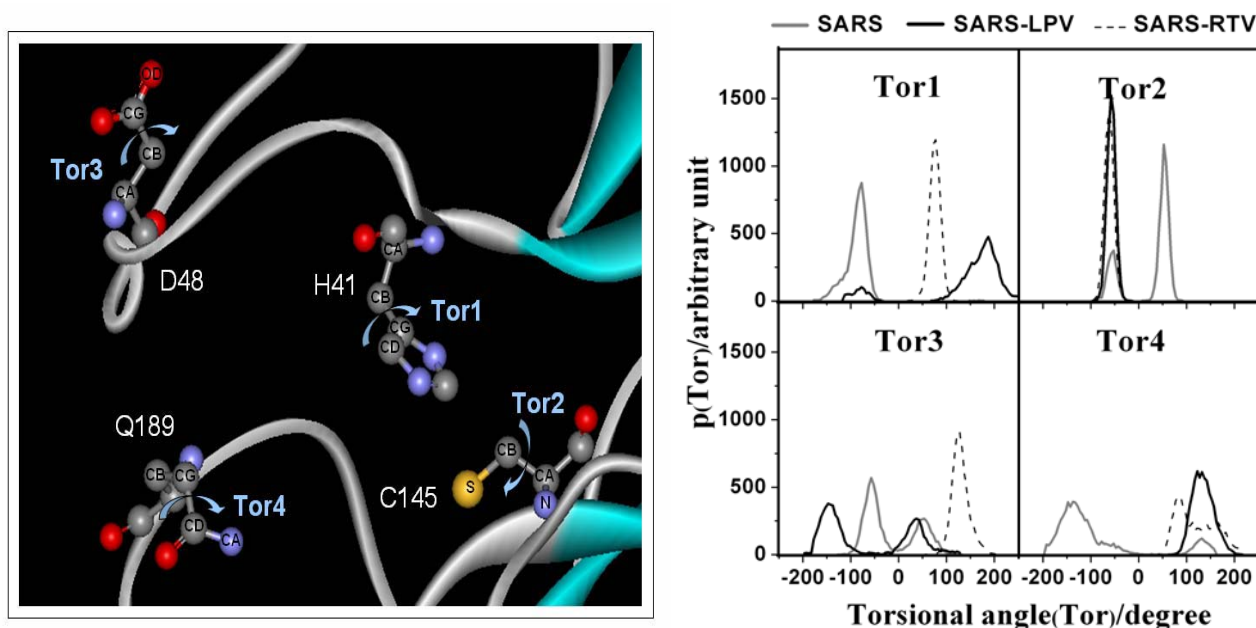|	TOR1(degree)	TOR2(degree)	TOR3(degree)	TOR4 (degree)
SARS	-77.5	-52.5, 52.5	-57.5, 52.5	-137.5, 132.5 (broad)
SARS-LPV	-77.5, 187.5	-57.5 (sharp)	-147.5, 37.5	132.5 (broad)
SARS-RTV	77.5 (sharp)	-62.5 (sharp)	127.5	82.5

Figure 4. The probability distribution of the two torsion angles of catalytic residues (TOR1: CA-CB-CG-CD of H41and TOR2: N-CA-CB-S of C145) and the two amino acids at the tip of the loop (TOR3: CA-CB-CG-OD of D48 and TOR4: CB-CG-CD-CA of Q189) during 600-1700 ps of the dynamics simulations.

3.4 Inhibitor-Enzyme Interaction

The SARS-LPV and SARS-RTV interactions were analyzed by means of hydrogen bond interaction and the molecular mechanical energy obtained from the electrostatic and the van der Waals interactions within the systems. The results were collected in Table 2.

To analyze hydrogen bond interaction, percentage and number of hydrogen bonding between the inhibitor and binding pocket residues were determined based on the following criteria: *(i)* proton donor-acceptor distance ≤ 3.5 Å and *(ii)* donor-H-acceptor bond angle ≥ 120°. In the MD simulations, the hydrogen bonds for the two systems are almost comparable. Six hydrogen bonds were detected in the SARS-LPV system while seven hydrogen bonds were found in SARS-RTV system.

However, the differences of inhibitors's structure between lopinavir and ritonavir can directly affect via the inhibitor-enzyme interaction. The number and %occupation at Cys-His catalytic dyad, Cys145 and His41, were detected for the SARS-RTV whereas in the SARS-LPV system was lost (Table 2 and Figure 6). Referred to the inhibitor structure, the longer side chains of ritonavir in SARS-RTV can interact with the neighbor residue in active site more than lopinavir in SARS-LPV system. Consequently, the SARS-RTV system was observed to interact more tightly to the important residues in the catalytic site of SARS-CoV protease than the SARS-LPV system.

It was reconfirmed in term of the gas phase MM interaction energy (ΔE_{MM}) shown in Table 3 where molecular mechanical interaction energy for the SARS-RTV complex of -51.61 kcal/mol is lower than that of -46.63 kcal/mol for the SARS-LPV. This result indicates that ritonavir structure is highly fitted with amino acid in the SARS-CoV cavity in comparison with lopinavir.

Table 2. Percentage occupation of hydrogen bonding between inhibitor and specific residues of SAR-CoV enzyme.

Residue	Type	%Occupation	
		SARS-LPV	SARS-RTV
T26	OG1_HG1---O3_LPV	18.7	-
T26	LPV_N4_H47---OG1	22.8	-
T26	LPV_N4_H47---O	36.9	-
H41	RTV_N2_H5---NE2	-	1.7
Y118	OH_HH---O5_LPV	2.8	-
N119	LPV_N4_H47---OD1	3.3	-
C145	RTV_N2_H5---SG	-	56.3
C145	SG_HG---O3_RTV	-	4.4
E166	N_H---N1_RTV	-	29.9
Q189	NE2_HE21---O1_LPV	9.9	-
Q189	NE2_HE21---N1_RTV	-	11.6
Q189	NE2_HE21---S1_RTV	-	7.6
Q189	NE2_HE21---O2_RTV	-	1.3

Table 3. Comparison of the molecular mechanical interactionenergies (kcal/mol) of the two simulated systems, SARS-LPV and SARS-RTV

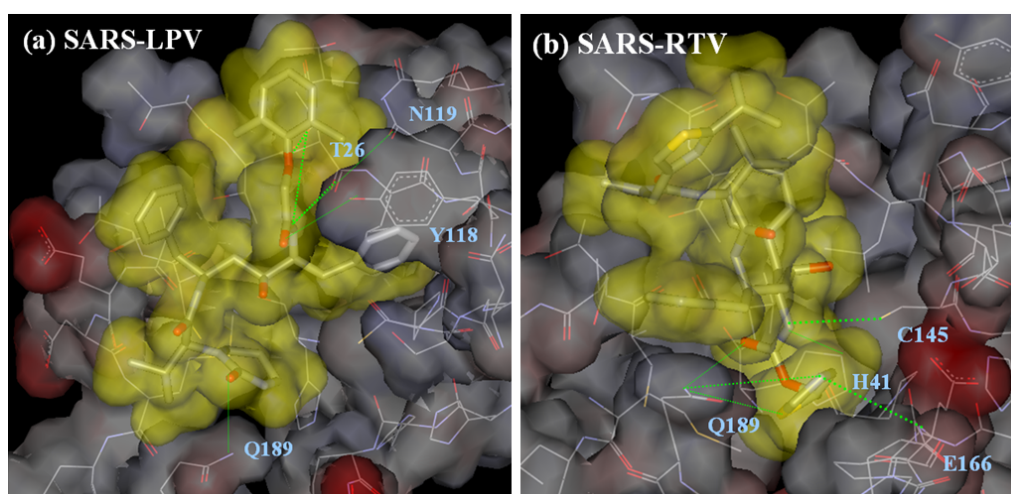|	SARS-LPV	SARS-RTV
ΔE_{ele}	-2.91 ± 4.21	-11.39 ± 3.28
ΔE_{vdw}	-43.72 ± 3.31	-40.23 ± 2.92
ΔE_{MM}	-46.63 ± 5.42	-51.61 ± 4.85

Figure 6. The Electrostatic potential and hydrogen bond in the binding pocket of the two simulated systems, SARS-LPV and SARS-RTV, where Van der Waal cavity of the inhibitor are in yellow, negative regions are in red and positive regions are in blue.

Conclusions

MD simulations and analysis of SARS, SARS-RTV and SARS-LPV systems give an insight into dynamic characteristics in term of the flexibility, conformation, and inhibitor-enzyme interactions. These data indicated a potent SARS-CoV selective inhibitor, when bound to the active site in SARS-CoV, should bind tightly to the catalytic site involving the closed contact between His and Cys and keep the disordered loop more motionally restricted

Acknowledgments

References

1. Donnelly, C. A., Ghani, A. C., Leung, G. M., Hedley, A. J., Fraser, C., Rieley, S., Abu-Raddad, L. J., Ho, L.M., Thach, T.Q., Chau, P. Epidermiological determinants of spread of causal agent of sevre acute respiratory syndrome in Hong Kong. *Lancet* 2003; 361: 1761-1766.

2. Kuiken, T., Fouchier, R. A., Schutten, M., Rimmelzwaan, G. F., van Amerongen, G., van Riel, D., Laman, J. D., de Jong, T., van Doornum, G., Lim, W. Newly discovered coronavirus as the primary cause of severe acute respiratory syndrome. *Lancet* 2003; 362: 263-270.

3. Peiris, J. S., Chu, C. M., Cheng, V. C., Chan, K. S., Hung, I. F., Poon, L. L., Law, K. I., Tang, B. S., Hon, T. Y., Chan, C. S., Chan, K. H., Ng, J. S., Zheng, B. J., Ng, W. L., Lai, R.W., Guan, Y., and Yeun, K.Y. *et al*. Clinical progression and viral load in a community outbreak of coronavirus-associated SARS pneumonia: a prospective study. *Lancet* 2003; 361: 1767-1772.

4. Marra, M. A., Jones, S. J., Astell, C. R., Holt, R. A., Brooks-Wilson, A., Butterfield, Y. S., Khattra, J., Asano, J. K., Barber, S. A., Chan, S.Y., *et a*l. The genome sequence of the SARS-associated coronavirus. *Science* 2003; 300: 1399-1404.

5. Anand, K., Ziebuhr, J., Wadhwani, P., Mesters, J. R., and Hilgenfeld, R. Coronavirus main protease (3CLpro) structure: Basis for design of anti-SARS drugs. *Science* 2003; 300: 1763-1767.

6. 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. The crystal structures of severe acute respiratory syndrome virus main protease and its complex with an inhibitor. *Proc. Natl. Acad., Sci. U.S.A.* 2003; 100: 13190-13195.

7. 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., Liang, P. H. Mechanism of the maturation process of SARS-CoV 3CL protease. *J. Biol. Chem.* 2005; 280: 31257-31266.

8. Lee, T. W., Cherney, M. M., Huitema, C., Liu, J., James, K E., Powers, J. C., Eltis, L. D., James, M. N. Crystal structure of the main peptidase from the SARS coronavirus inhibited by a substrate-like aza-peptide epoxide. *J. Mol. Biol.* 2005; 353: 1137-1151.

9. Huang, C., Wei, P., Fan, K., Liu, Y., and Lai, L. 3C-like proteinase from SARS coronavirus catalyzes substrate hydrolysis by a general base mechanism. *Biochemistry*. 2004; 43: 4568-4574.

10. Zhang, X. W., and Yap, Y. L. Old drugs as lead compounds for a new disease? Binding analysis of SARS coronavirus main protease with HIV, psychotic and parasite drugs. *Bioorg. Med. Chem.* 2004; 12: 2517-2521.

11. Yamamoto, N., Yang, R., Yoshinaka, Y., Amari, S., Nakano, T., Cinatl, J., Rabenay, H., Doerr, H. W., Hunsmann, G., Otaka, A., Tamamura, H., Fujii, N., Yamamoto, N. HIV protease inhibitor nelfinavir inhibits replication of SARS-associated coronavirus. *Biochem. Biophys. Res. Comm*. 2004; 318: 719-725.

12. Jenwitheesuk, E. and Samudrala, R. Identifying inhibitors of the SARS coronavirus proteinase. *Bioorg. Med. Chem. Lett.* 2003; 13: 3989-3992.

13. Vastag, B. Old drugs for a new bug. *JAMA*. 2003; 290: 1695-1696.

14. Case, D. A., Pearlman, J. W., Caldwell, T. E., Cheatham, J. III., Wang, W. S., Ross, C. L., Simmerling, T. A., Darden, K. M., Merz, R. V., Stanton, A. L., Cheng, J. J., Vincent, M., Crowley, V., Tsui, H., Gohlke, R. J., Radmer, Y., Duan, J., Pitera, I., Massova, G. L., Seibel, U. C., Singh, P. K., and Kollman, P. A. AMBER. Version 7.0 ed., University of California, San Francisco, CA. (2002).

15. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys*. 1983; 79: 926-935.

16. Muegge, Y. C. Martin. A General and Fast Scoring Function for Protein-Ligand Interactions: A Simplified Potential Approach, *J. Med. Chem.* 1999; 42: 791-804.

17. Cornell, W. D., Cieplak, P., Bayly, C. I., Gould, I. R., Merz, K. M., Ferguson, D. M., Spellmeyer, D. C., Fox, T., Caldwell, J. W., Kollman, P. A. A second generation force-field for the simulation of proteins, nucleic acids, and organic-molecules. *J. Am. Chem. Soc*. 1995; 117: 5179-5197.

18. Ryckaert, J. P., Ciccotti, G., Berendsen, H. J. C. Numerical integration of the Caresian equations of motion of a system with constaints: molecular dynamics of n-alkanes. *J. Comput. Phys*. 1977; 23: 327-341.

19. Teague, S. J. Implications of protein flexibility for drug discovery. *Nature Rev. Drug Discov*. 2003; 2: 527– 541.

20. Carlson, H. A., McCammon, J. A. Accommodating protein flexibility in computational drug design. *Mol. Pharmacol.* 2000; 57: 213 –218.

21. Carlson, C. A. Protein flexibility and drug design: how to hit a moving target. *Curr. Opin. Chem. Biol*. 2002; 6: 447–452.

22. Lin, J. H., Perryman,A. L., Schames, J. R., McCammon, J. A. Computational Drug Design Accommodating Receptor Flexibility: The Relaxed Complex Scheme**.** *J. Am. Chem. Soc.* 2002; 124: 5632 – 5633.

23. Anand, K., Palm, G. J., Mesters, J. R., Siddell, S. G., Ziebuhr, J., Hilgenfeld, R. Structure of coronavirus main proteinase reveals combination of a chymotrypsin fold with an a-extra helical domain. *EMBO J.* 2002; 21: 3213–3224.

24. Hornak, V., Okur, A., Rizzo, R. C., Simmerling, C. HIV-1 protease flaps spontaneously open and reclose in molecular dynamics simulations. *Proc. Natl. Acad. Sci. U S A*. 2006**;** 103: 915-920.

25. Hornak, V., Okur, A., Rizzo, R. C., Simmerling, C. HIV-1 protease flaps spontaneously close to the correct structure in simulations following manual placement of an inhibitor into the open state. *J. Am. Chem. Soc*. 2006**;** 128**:** 2812-2813.