การคัดกรองสารที่มีฤทธิ์ต้านเชื้อ Enterovirus71 และ Coxsackievirus A16

ที่ก่อให้เกิดโรคมือ เท้า และปาก โดยวิธี Steered Molecular Dynamic Simulation

SCREENING OF THE POTENT COMPOUNDS AGAINST ENTEROVIRUS 71

AND COXSACKIEVIRUS A16 OF HAND, FOOT AND MOUSE DISEASE

USING STEERED MOLECULAR DYNAMICS SIMULATION

โดย

นางสาววารินทร์ เจษฎาวิสุทธิ์

รายงานนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตร

ปริญญาวิทยาศาสตรบัณฑิต

ภาควิชาเคมี คณะวิทยาศาสตร์

จุฬาลงกรณ์มหาวิทยาลัย

ปีการศึกษา 2556

เรื่อง   การคัดกรองสารที่มีฤทธิ์ต้านเชื้อ Enterovirus71 และ Coxsackievirus A16 ที่
ก่อให้เกิดโรคมือ เท้า และปาก โดยวิธี Steered Molecular Dynamic Simulation

โดย   นางสาววารินทร์ เจษฎาวิสุทธิ์

ได้รับอนุมัติให้เป็นส่วนหนึ่งของการศึกษา

ตามหลักสูตรวิทยาศาสตรบัณฑิต ภาควิชาเคมี

คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

คณะกรรมการสอบโครงการ

.... ประธานกรรมการ

(รศ.ดร.วุฒิชัย พาราสุข)

.... อาจารย์ที่ปรึกษา

(ศ.ดร.สุพจน์ หารหนองบัว)

.... อาจารย์ที่ปรึกษาร่วม

(อ.ดร.ธัญญดา รุ่งโรจน์มงคล)

.... กรรมการ

(รศ.ดร.พลกฤษณ์ แสงวณิช)

รายงานฉบับนี้ได้รับความเห็นชอบและอนุมัติโดย หัวหน้าภาควิชาเคมี

....

(รศ.ดร.วุฒิชัย พาราสุข)

หัวหน้าภาควิชาเคมี

วันที่.......เดือน.................พ.ศ............

ชื่อโครงการ การคัดกรองสารที่มีฤทธิ์ต้านเชื้อ Enterovirus 71 และ Coxsackievirus A16

ที่ก่อให้เกิดโรคมือ เท้า และปาก โดยวิธี Steered Molecular Dynamics

Simulation

ชื่อนิสิตในโครงการ นางสาววารินทร์ เจษฎาวิสุทธิ์ **เลขประจำตัว** 533 31192 23

อาจารย์ที่ปรึกษาโครงการ ศ.ดร.สุพจน์ หารหนองบัว

อาจารย์ที่ปรึกษาร่วม อ.ดร.ธัญญดา รุ่งโรจน์มงคล

ภาควิชาเคมี คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย ปีการศึกษา 2556

บทคัดย่อ

โรคมือ เท้า และปากที่เกิดแพร่ระบาดในหลายประเทศ อาทิ มาเลเซีย จีน สิงคโปร์ เวียดนาม บรูไน กัมพูชา และอังกฤษ เป็นต้น ในประเทศไทยพบระบาดรุนแรงใน เด็ก ทำให้หลาย โรงเรียนต้องปิดการเรียนการสอน โดยสาเหตุของโรคเกิดจาก Enterovirus 71 (EV-71) และ Coxsackievirus A16 (CV-A16) ซึ่งทุกวันนี้มีหลายกลุ่มวิจัยที่มุ่งคิดค้นและพัฒนายา ต้านการทำงาน ของ EV-71 อย่างไรก็ตามสารประกอบที่ได้จากการศึกษาที่ผ่าน ยังมีประสิทธิภาพการออกฤทธิ์ต่ำ ใน การศึกษานี้มีเป้าหมายมุ่งไปที่การ คัดกรองสารประกอบจากฐานข้อมูล ZINC และ DrugBank โดย เริ่มต้นจากการคัดเลือกโครงสร้างที่คล้ายกับ rupintrivir ซึ่งเป็นสารประกอบที่กำลังอยู่ในขั้น in vitro รวมถึงที่ได้จากการดัดแปลงโครงสร้างของ rupintrivir จากนั้นใช้วิธี Molecular Docking ในการ สร้างสารประกอบเชิงซ้อนระหว่างลิแกนด์และ CV-A16 protease และคัดเลือกลิแกนด์ที่สามารถยึด จับกับเอนไซม์เป้าหมายอย่างแข็งแรง และนำมาศึกษาด้วยวิธี สเตียร์โมเลกุลาไดนามิกซิมูเลชัน . ซึ่งมี ความแม่นยำในการคัดกรองสารที่ออกฤทธิ์ทางชีวภาพได้ดีกว่า วิธี Molecular Docking จากการวิจัย พบว่าแรงฉีกที่ใช้ในการดึง ZINC11783883 (583 kJ/mol•nm) มีค่าสูงกว่า rupintrivir (514 kJ/mol•nm) จึงน่าจะเป็น lead compound ที่สามารถใช้พัฒนาสารเคมีตัวใหม่ที่เหมาะสำหรับการ รักษาโรคมือ เท้า และปากต่อไป

คำสำคัญ : การคัดกรองสาร, เอนไซม์โปรติเอส, โรคมือ เท้า และปาก , สเตียร์โมเลกุลาไดนามิกซิมูเล ชัน

Title SCREENING THE POTENT COMPOUNDS AGAINST ENTEROVIRUS 71

AND COXSACKIEVIRUS A16 OF HAND, FOOT AND MOUSE DISEASE

USING STEERED MOLECULAR DYNAMICS SIMULATION

Student name   Warin   Jetsadawisut   **ID** 533 31192 23

Advisor   Prof. Dr. Supot   Hannongbua

Co-advisor   Dr. Thanyada   Rungrotmongkol

Department of Chemistry, Faculty of Science, Chulalongkorn University, Academic year 2013

Abstract

Hand, foot and mouth disease (HFMD) outbreaks have occurred in many countries including Malaysia, China, Singapore, Vietnam, Brunei, Cambodia and England etc. In Thailand, increase in the prevalence of HFMD in children has forced many school to be closed across the outbreak. Enterovirus 71 (EV-71) and Coxsackievirus A16 (CV-A16) are two viruses that most commonly cause the pathogenic HFMD. Nowaday, many research groups have focused on drug design and development toward EV-71. However, compounds from study have relatively low efficiency. In this study, we aimed to screen the potent compounds from avialable databases and the rupintrivir analogs. Molecular docking was used to build the complex between ligand and CV-A16 protease. Then each complex was studied by steered molecular dynamics simulation (SMD) which shows better accuracy in screening the bioactivities compounds than molecular docking. Based on pulling force, the ZINC11783883 (583 kJ/mol•nm) was predicted to be a lead compound for further anti-HFMD drug design and development.

Keywords : Compound screening, protease, Hand Foot Mouth Disease, Steered Molecular Dynamics simulation

Acknowledgement

I would like to express my deepest appreciation to my advisors, Prof. Supot Hannongbua and Dr. Thanyada Rungrotmongkol, who have shown the attitude and the substance towards this project. They continually and persuasively conveyed a spirit of adventure of research, and an excitement of teaching. Without their supervision and constant help this dissertation would not been possible.

My gratitude is to Assoc. Prof. Dr. Vudhichai Parasuk, Assoc. Prof. Dr. Pornthep Sompornpisut and Assist. Prof. Dr. Somsak Pianwanit for their advices to make this project possible.

My gratitude is also to Ms. Arthitaya Meeprasert, who taught steered molecular dynamic simulation and imparting basic knowledge on this project, and Mr. Somphob Thompho, who taught optimization on ligand structures. I thank all members in ATC lab for their encouragement, advices and great ideas that you have shared with me. It was one of the most enjoyable moments of my life being with you.

The Scholarship Program from the Department of Chemistry, Faculty of Science, Chulalongkorn University is acknowledged.

I would like to thank all friends who motivated and supported me in all ways to reach this goal of my life. Finally, any words cannot explain how grateful I am to my parents for their undying love, understanding, encouragement and cheerfulness throughout my Bs.C study.

Content

Contents of figure

Contents of Table

Contents of diagram

1.1 Hand, foot and mouth disease

1.1.1 Introduction of hand, foot and mouth disease

Hand, foot and mouth disease (HFMD) is caused by viruses the Picornaviridae family, which coxsackievirus A16 (CV-A16) is the most common virus. CV-A16 presents no severe symptom and patients recover swiftly. In contrast, enterovirus 71 (EV-71) is rarer but can be fatal in young children and causes severe neurological symptoms[1, 2].

1.1.2 Signs and symptoms

HFMD can outbreak from excretion, saliva or feces of the illness. Sometimes viral transmission occurs from sneezing. HFMD commonly affects to children in summer and autumn periods.

The patients are likely to experience fevers and tiredness after 3-7 days of infection and lose their appetite. In the next step, they have a stuffed nose, cold sores, sore throat, and babies stop suckling. They have painful sores in the mouth, throat or both. Rashes form on the hands, feet, mouth, tongue, and inner cheeks[1, 3, 4]. In severe cases, enteroviruses can damage the central nervous systems (CNS) to viral meningitis, encephalitis and severe myocarditis, similar to fatal pulmonary edema. The children are affected more than adult because of the low efficiency of their immune systems[1, 3].

1.1.3 Outbreaks

EV-71 and CV-A16 associated HFMD are prominent in the southeast pacific region of the world, with millions of people infected every year and many fatalities amongst children. In 2013, 265 children died in China alone[5]. Although the disease is

more common in less developed countries, in 2013 there were relatively small outbreaks in Australia, and Hong Kong experienced a three-fold increase in infection. This increase in prevalence in areas of wealth likely leads to an increase in funding for combating the disease.

Table 1. Number of infected individuals in the South East Pacific region by country with incidences in 2012 and 2013. Note that EV-71 is not the primary cause of HFMD in Australia[5].

Country	Total HFMD infections 2012	Total HFMD infections 2013	2013/2012 incidence ratio
China	2,071,237	1,757,078	0.8
Hong Kong	1,625	495	3.3
Japan	68,691	298,879	4.3
Singapore	36,786	30,875	0.8
Thailand	*Jan 1 to July 22* 16,860	*Jan 1 to June 2* 11,678	0.7
Australia	-	32	-

1.1.4 Life cycle of enterovirus 71 and coxsackievirus A16
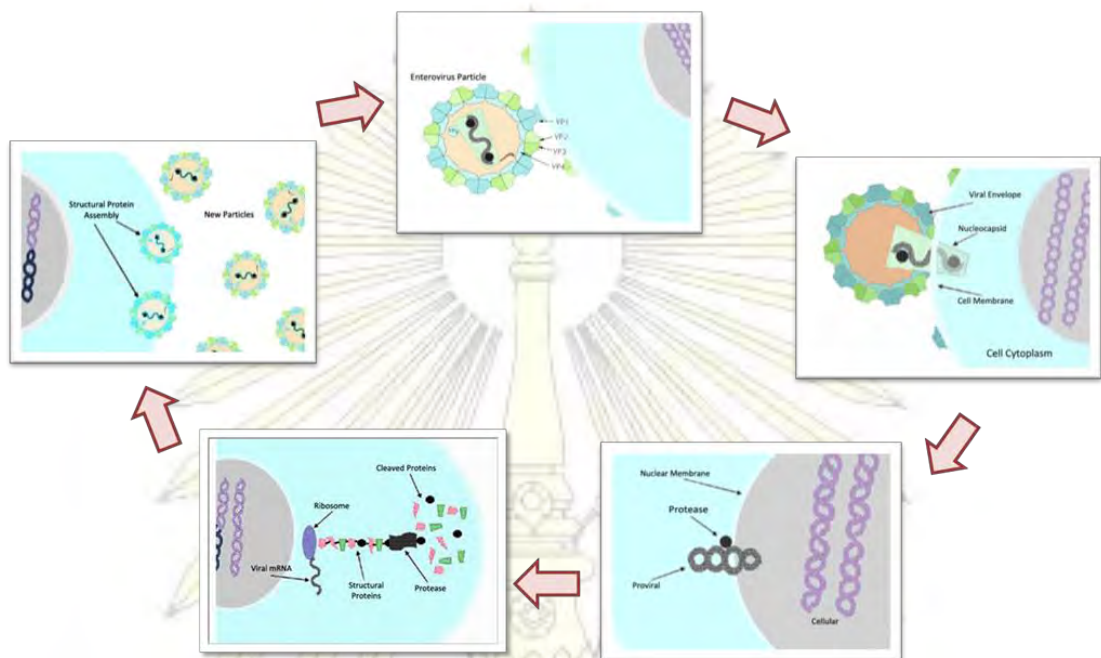
Figure 1: The life cycle of positive single-standed RNA of EV-71 and CV-A16[6]

The both viruses are in the family Picornaviridae[7] and contain a genome of single-standed positive-sense RNA with a single open reading frame (ORF) encoding a large polyprotein precursor. In infected cells, this polyprotein is further cleaved into structural and nonstructural proteins via the virus-encoded 2A and 3C proteases. Upon translation of the polyprotein, the 2A protease automatically cleaves the joining sequence between Vp1 and 2A. The 3C, the main protease, is responsible for the cleavage of the eight junction sites within the remainder of the polyprotein[7]. The 3C protease is a constituence of the replication complex of the viral genomic RNA[6]. Therefore, the protease cleavage step is focused for bulided inhibitor extract virus production[8].

1.2 Cysteine protease 3C

1.2.1 Nature of cysteine protease 3C
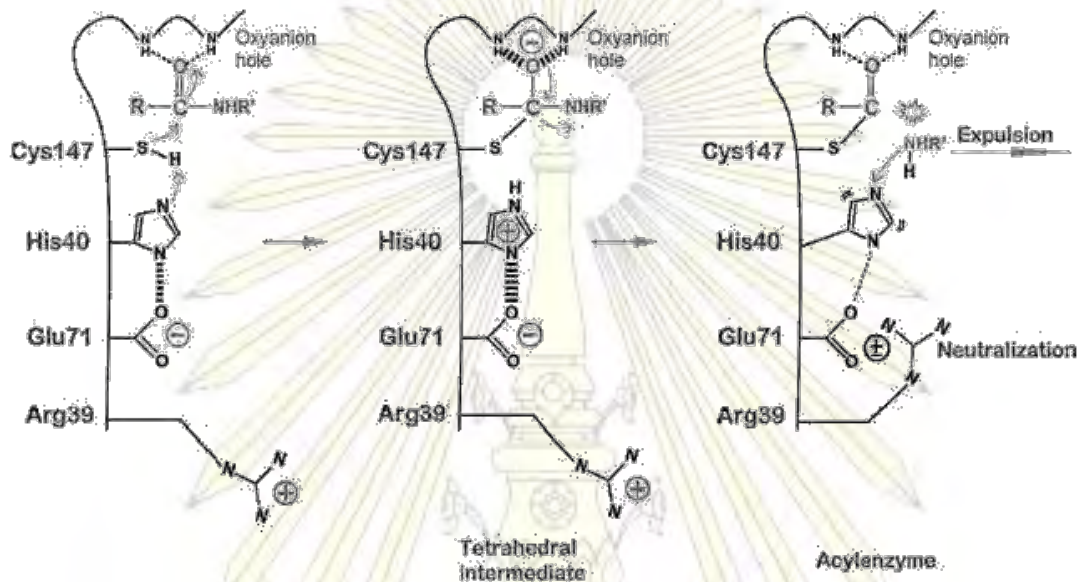
Figure 2: Proteolysis mechanism of EV-71 and substrate 3C protease involving the neutralizer Arg39. The catalytically important residues are shown in black and green, respectively. The electron relays are indicated by red arrows. Thicker and thinner dashed lines indicate stronger and weaker interactions[9].

3C is a cysteine protease essential for the production of EV-71 virion particles[10]. These particles are vital for the replication of the virus. It was also found that a novel mechanism of 3C protease impairs host immune function by interfering with polyadenylation of the host cell RNA by digesting CstF-64, a critical host factor for pre-mRNA processing[11]. The vital role of 3C protease in both viral replication and counter-immunity makes 3C protease a very attractive drug target.

1.3 Rupintrivir

1.3.1 Nature of rupintrivir
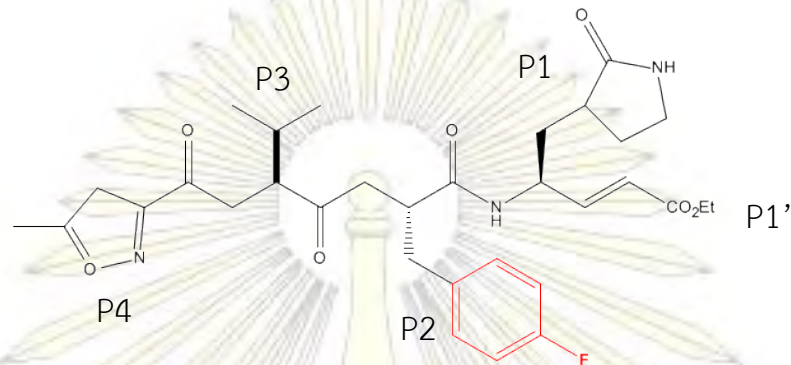
Figure 3: The 2D structure of rupintrivir. a lactam in P1, a fluorophenylalanine in P2, amino acid Val in P3, a 5-methyl-3-isoxazole in P4 and an α,β-unsaturated ester in P1'.

Rupintrivir is a specific peptide - mimic inhibitor against the 3C protease of human rhinovirus (HRV). It is a develop potent compound to inhibit against EV-71 contains rupintrivir[12].

1.3.2 Rupintrivir efficiency

Rupintrivir, also known as AG7088[12], is a compound designed originally to inhibit rhinovirus (HRV) 3C[pro], however has been found to exhibit broad-spectrum antiviral activity in other members of Picornaviridae. Rupintrivir has reached in phase II trials and is currently the most effective peptidomimetic 3C protease inhibitor with an IC_{50} of 2.1 μM[4, 13, 14]. Another desirable reason for targeting 3C protease is that rupintrivir is an effective inhibitor of the 4 other enterovirus strains, despite minor catalytic triad mechanism differences; CVB2, CVB5, EV6 and EV9[15].

1.3.3 Why we cannot use rupintrivir to treat hand, foot and mouth disease
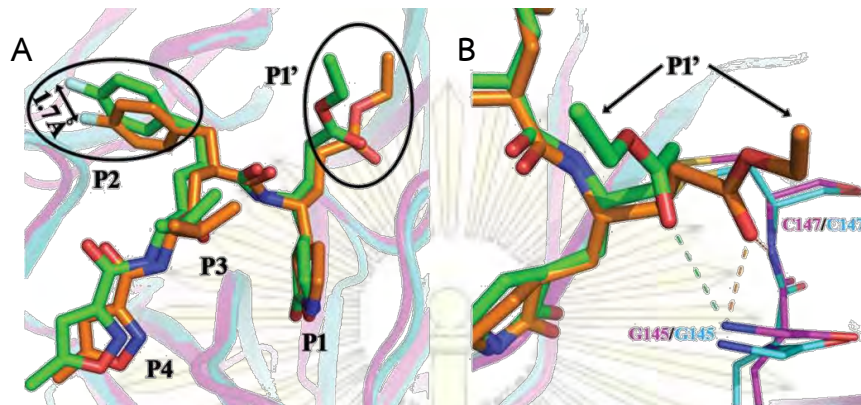
Figure 4: Rupintrivir in the 3C protease active site of HRV (cyan-orange) and CV-A16 (magenta-green). (A) Rupintrivir differenctly uplifts by 1.7 Å in P2 chain and P1' chain. (B) Hydrogen bonds[8] between P1' chain and protease residues.

Unfortunately, although rupintrivir is a potent inhibitor of HRV 3C protease, it is less than 100 times as effective against CV-A16 and EV-71 and thus not suitable for drug repurposing. Now that higher resolution structural models of the proteins are available, we are able to start investigation into structural modifications of rupintrivir so that it more effectively inhibits the 3C protease of CV-A16 and EV-71.

1.4 Objective of this study

To screen the potent compounds aginst CV-A16 3C protease of hand, foot and mouth disease.

Chapter II

Theory

2.1 Similarity search

The similarity search is a technique to screen the compounds using chemical structure from database of template. In this study, base on rupintrivir structure and set filter ex molecular weight, similarity threshold of 0.6 were applied. A higher similarity threshold results in less hits that are more similar to the query structure. Threshould value must be between 0.3 to 1

2.2 Molecular Docking

Molecular docking is used to predict conformation or nature of binding between molecules based on a lock and key process. A goal of molecular docking method is to find lead compound or approximately activity of ligand and to understand a binding mechanism between enzyme and ligand. In the current drug development, the docking method is mostly used for screening preference inhibitor.

There are two docking procedures, rigid docking and flexible docking. Docking algorithms including CDOCKER, LibDock, LigandsFit and MCSS are available techniques.

2.2.1 CDOCKER algorithm

CDOCKER algorithm uses a random initial ligand placement and CHARMm forcefield for ligand docking into the rigid receptor. In this method, the ligand is subjected to high-temperature molecular dynamics and random orientations of the conformations, which ligand with low-energy orientation could specific all located within the receptor active site[16]. Interaction energy between docked ligand and receptor (ΔE_{int}) is calculated in according to Equation.

$$\Delta E_{int} = E(A,B) - (E(A) + E(B))$$

The E(A) and E(B) are the energies of the isolate objects (free form) and E(A,B) is the energy of their interacting assembly (complex). In docking result, the ligand with the most negative value of ΔF_{int} is ranked in the top score.

2.3 Steered molecular dynamics simulation (SMD)

Steered molecular dynamics simulation (SMD) is an advanced molecular dynamics (MD) simulation. Which induces unbinding of ligand and conformational changes in biomolecules on time scales accessible to molecular dynamics simulation (MD). The system is applied external forces to pulling ligand out of enzyme[17].

2.3.1 Pulling path

Caver is a software tool for analysis and visualization of tunnels and channels in protein structures. A tunnel connects a protein cavity with bulk solvent. The shape of protein tunnel can be approximated as a pipeline with a varying width of cross section. This approximation is useful for estimation of the largest probe accessing the deepest place in the pocket. The size of a probe able to access internal cavity is limited by radius of tunnel gorge, *i.e.*, the most narrow place in the tunnel. The tunnel profile is a graph of cross section radius (radius of maximally inscribed ball) versus tunnel length measured from its deepest place to the surface opening[18].
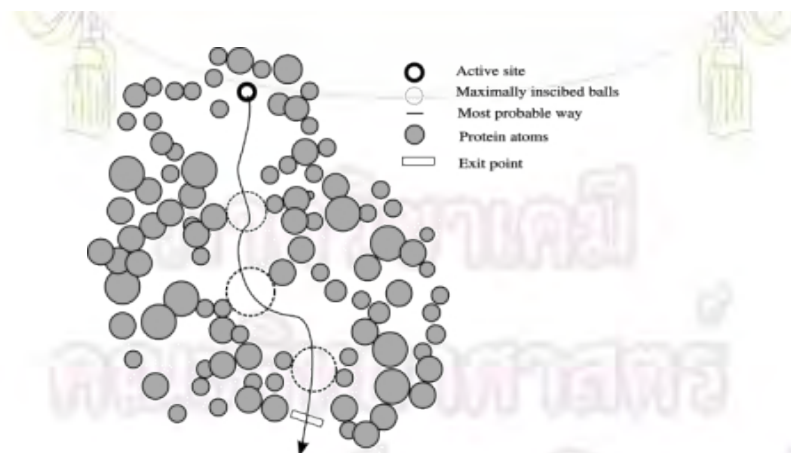
Figure 5: The shape of protein tunnel[4]

2.3.2 GROMACS-4.5.5

   The protein preparation, minimization, MD simulation and pulling ligand out of protein by SMD were performed by GROMACS-4.5.5 program. It is free and saves computing times more than other programs.

Chapter III

Methodology

3.1 Ligand selection and preparation
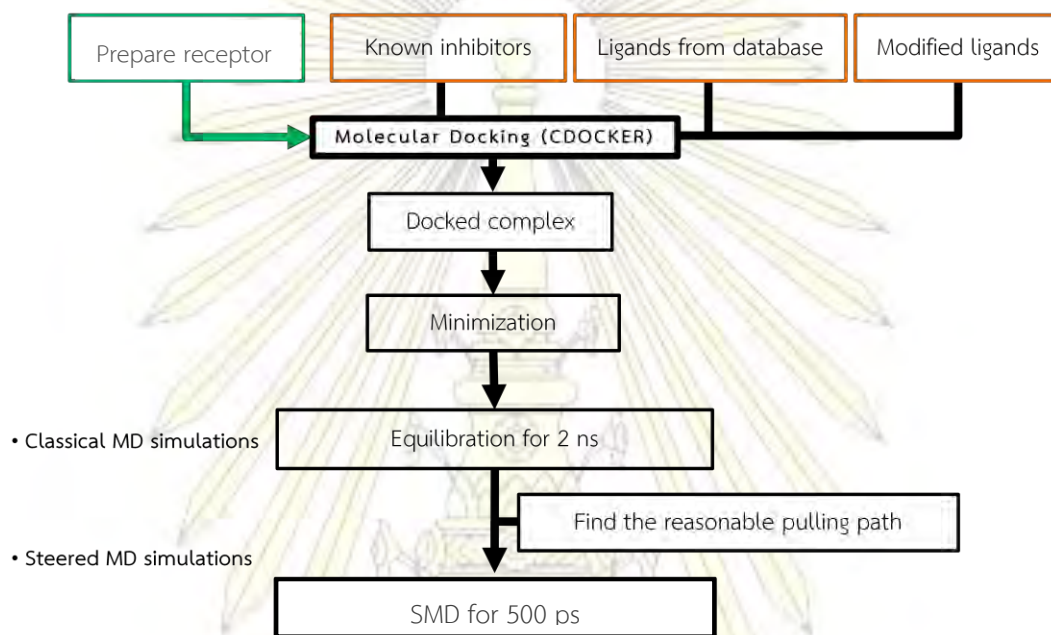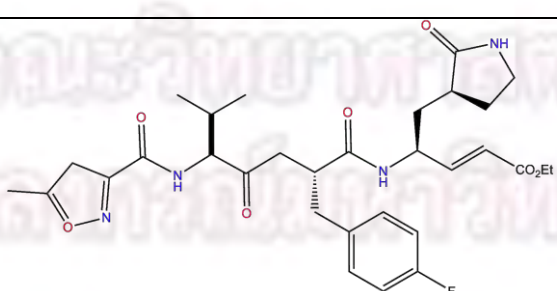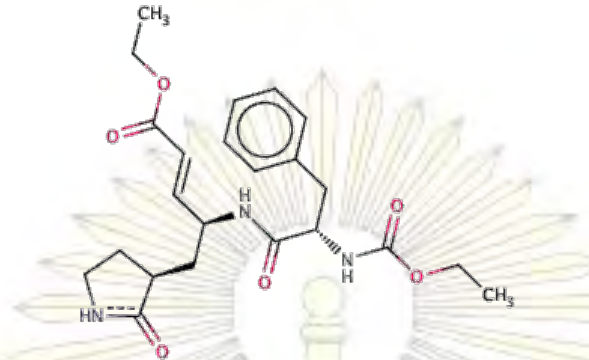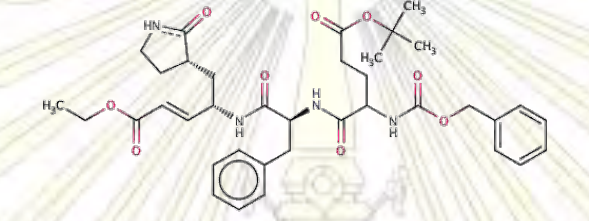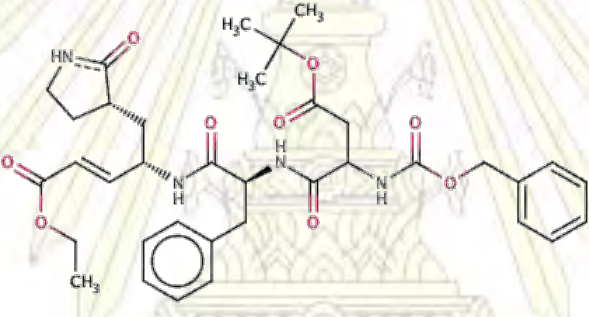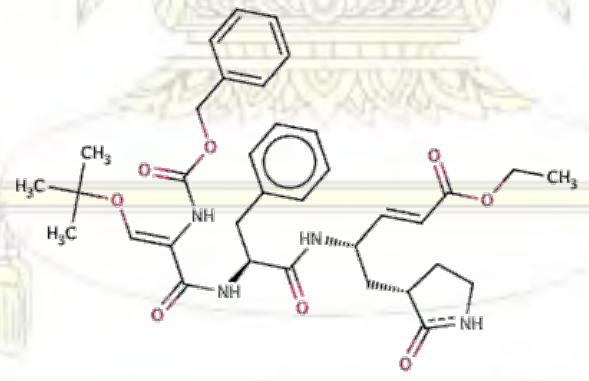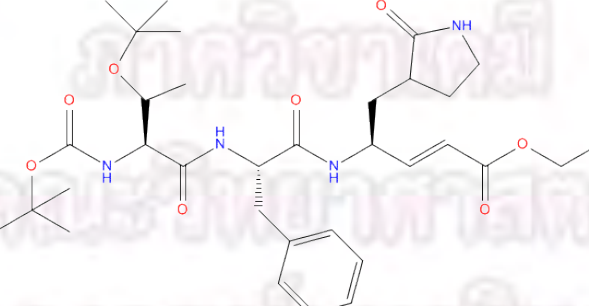
Diagram 1: Outline of overall process

3.1.1 Known inhibitors

Rupintrivir[8] and inhibitors[12] with 50% effective concentration (EC$_{50}$) to 3C protease of enterovirus 71 from the previous experimental studies are summarized in Table 2.

Table 2: Chemical structures and EC$_{50}$ values of rupintrivir[8] and 5 inhibitors[12]

Ligand	Chemical structure	EC$_{50}$ (μM)
Rupintrivir		0.3

SG81		> 40
SG82		1.2 ± 0.2
SG83		1.3 ± 0.5
SG85		1.0 ± 0.2
SG98		2.5 ± 0.5

3.1.2 Ligands selected from avialable databases

In the present study, the three available databases ZINC, DrugBank and CheMBL, were focused with 60% structure similarity toward rupintrivir. The 200 ZINC compounds, 28 DrugBank compounds and 16 ChEMBL compounds were selected for further docking study.

3.1.3 Modified rupintrivir compounds
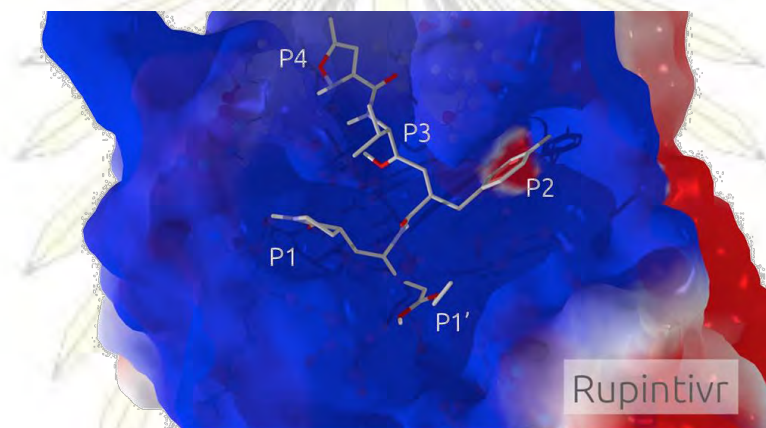
Figure 6: Rupintrivir in positively charged pocket of CV-A16.
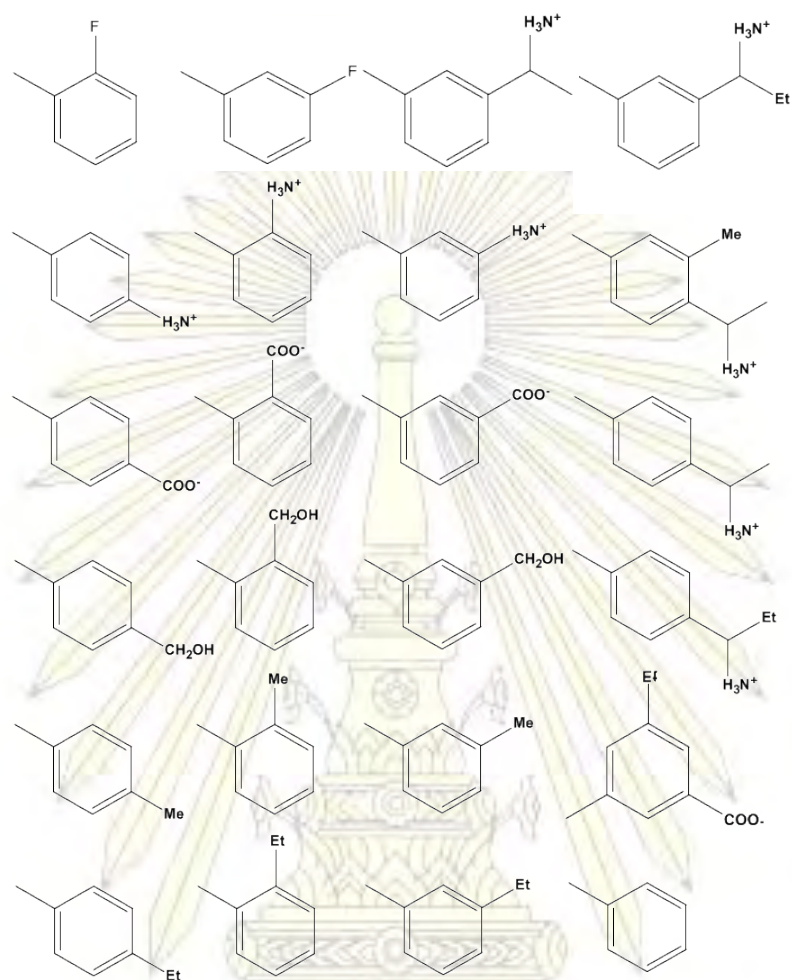
Figure 7: The Chemical structure of 24 P2-modified functional groups on rupintrivir.

To improve the ligand fitness at the P2 binding site, the P2 group of rupintrivir is modified with 24 different functional groups as shown in Figure 7.

3.2 Protein selection

The X-ray structures of cysteine protease of EV-71 and CV-A16 with rupintrivir bound are extracted from PDB databank (3R0F and 3SJI, respectively). From superimposition between the two complexes (Figure 8), both enzymes have sequence identity of 90% and root mean sequare deviation of 0.6 Å suggesting that there are relatively similar in structure and conformation. Since CV-A16 has the most outbreaks in many countries and thus we selected the CV-A16 as the target enzyme

in this study. We postulate that an efficient inhibitor of CV-A16 should be effective against EV-71 too.
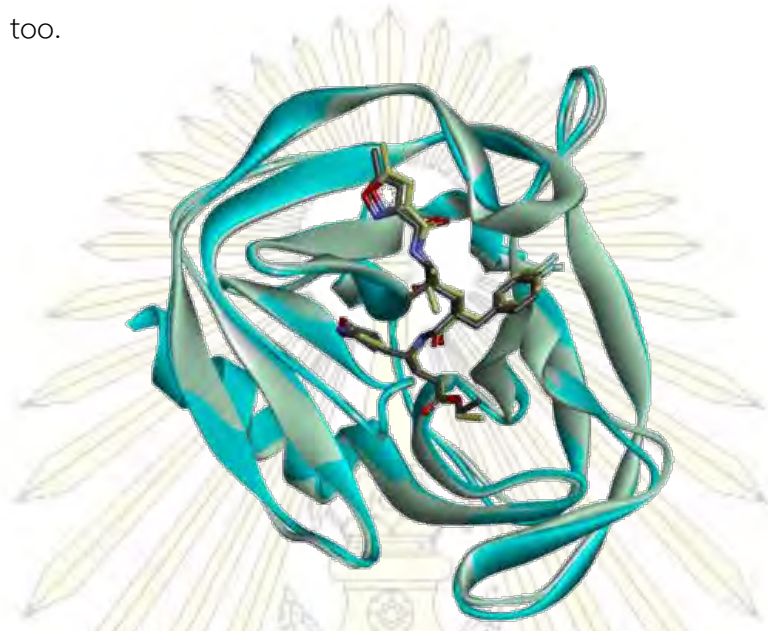
Figure 8: Crystal structures of EV-71 (grey) and CV-A16 (blue) in complex with rupintrivir are superimposed.
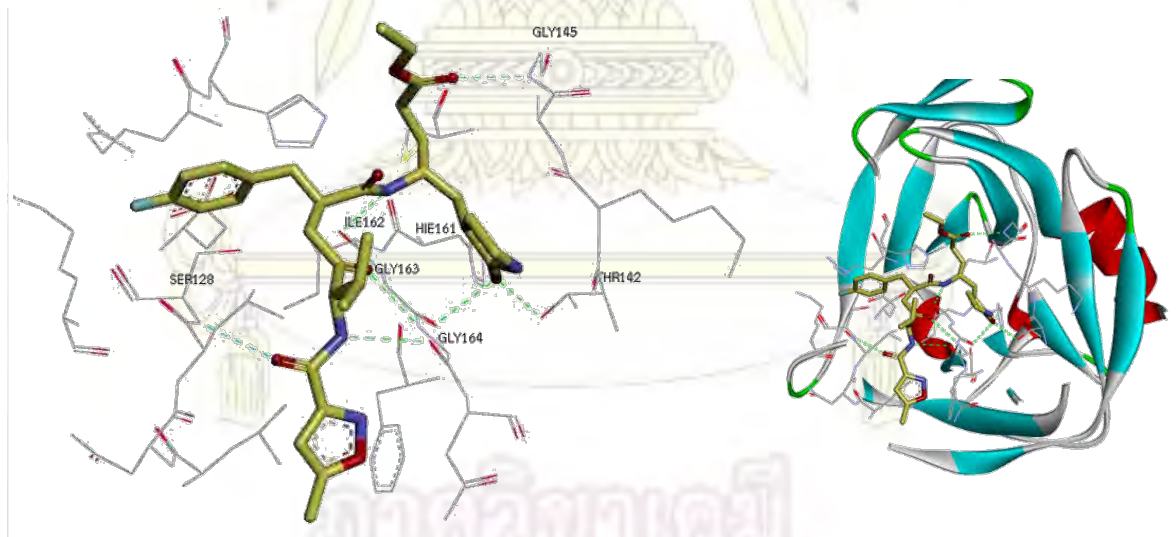
Figure 9: Hydrogen bonds between rupintrivir and 3C protease of CV-A16

3.3 Molecular docking method

3.3.1 Preparation of ligand structure (diagram 2)

Download structure from databases (ZINC, DrugBank and CheMBL)	Create structures of the known EC_{50} values and modified rupintrivir

Geometry optimization by MM method with MMFF94 partial charge and CHARMm force field

optimized strucutre
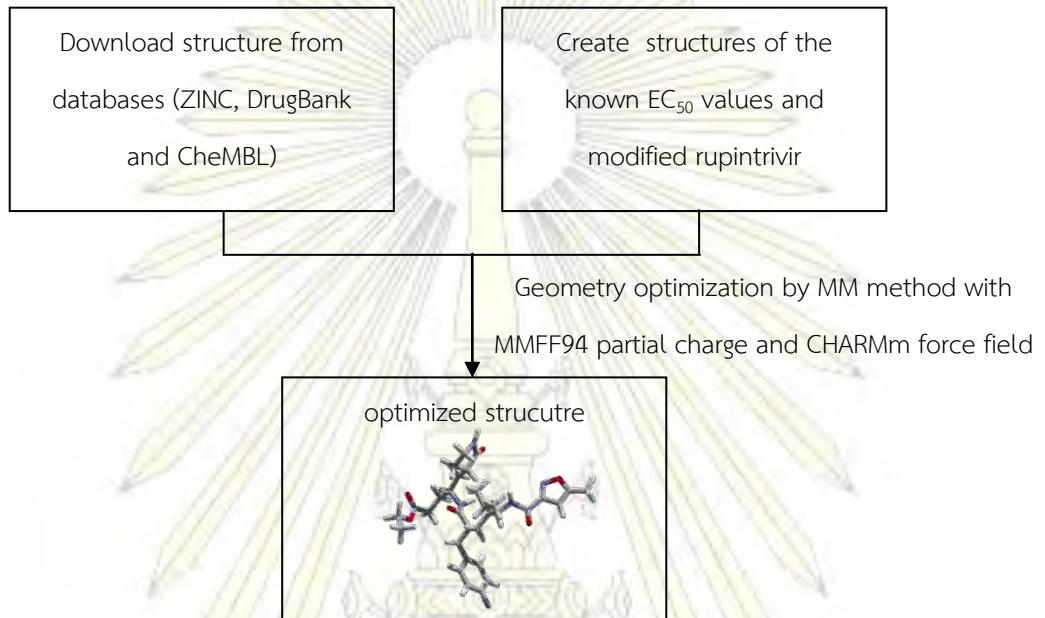
Diagram 2: Ligand optimization in molacular docking

3.3.2 Preparation of protein structure (diagram 3)

CV-A16 (3SJI)

Define and enlarge binding site using rupintrivir structure

Forcefield : CHARMm

Partial charge : MMFF94
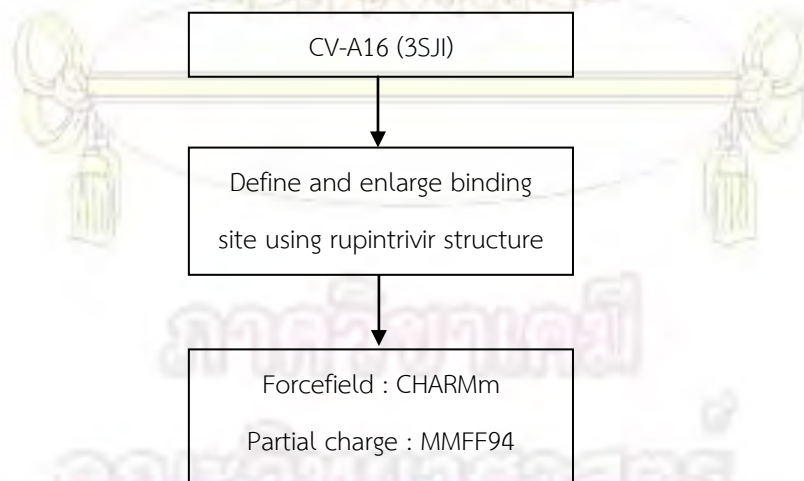
Diagram 3: Preparation of receptor in molacular docking
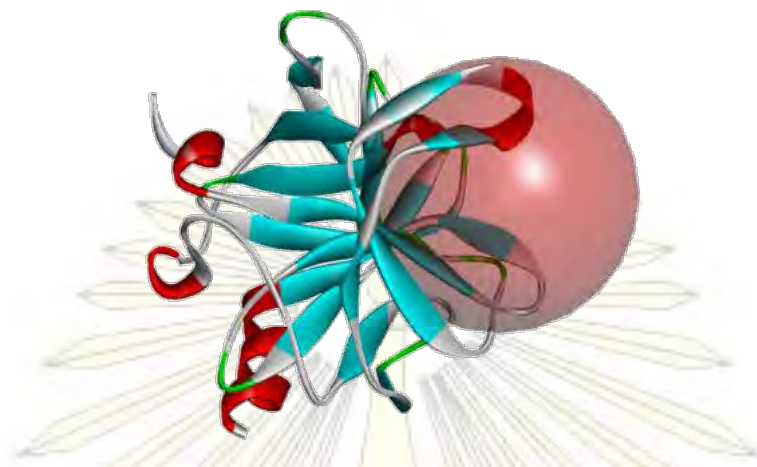
Figure 10: Defined binding site for ligand docking into the CV-A16 protaese

3.3.3 Docking procedure

Docking of each ligand into the active site of CV-A16 protease was performed by CDOCKER using CHARMm forcefield (Momany-rone partial charge method). The docking sphere is of 11 $\mathring{A}^3$ around rupintrivir structure. In annealing, the system was heated to 700 K and consequently with the cooling down to 300 K.
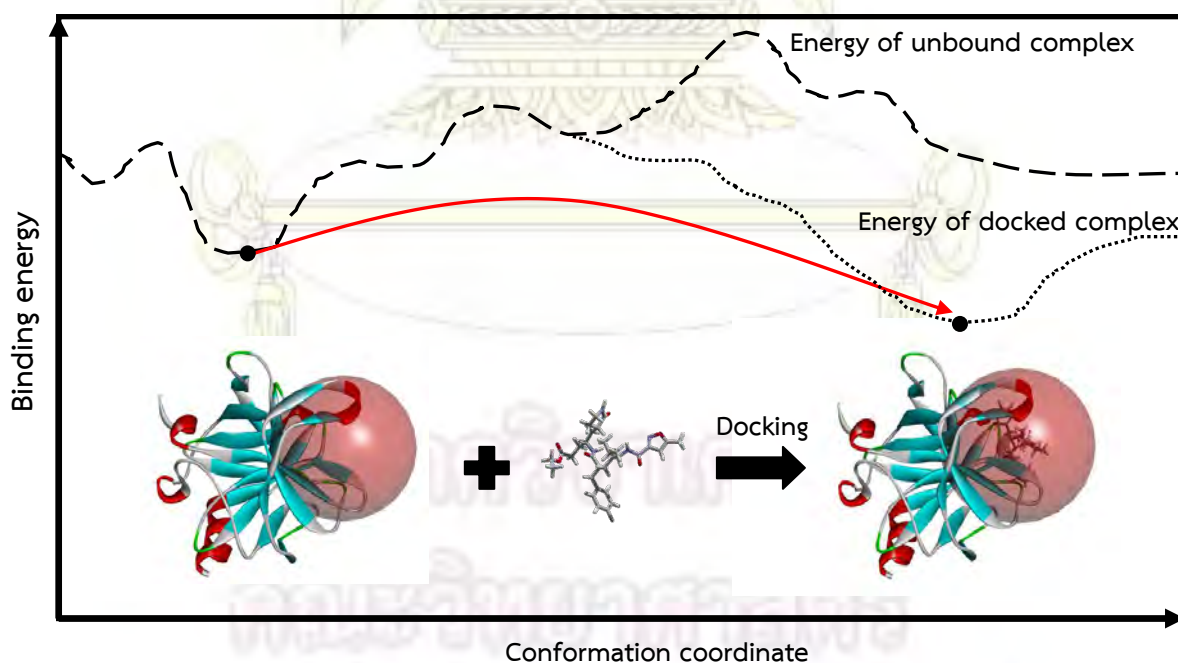
Figure 11: Binding energy profile along the conformation coordinates from unbound complex to docked complex.

3.4 Molecular dynamics simulation

We, defined the unit cell by having space left for pulling direction to allow the continuous pulling without interecting with the periodic images of the system. The system was solvented and neutralized by one Cl⁻ ion.
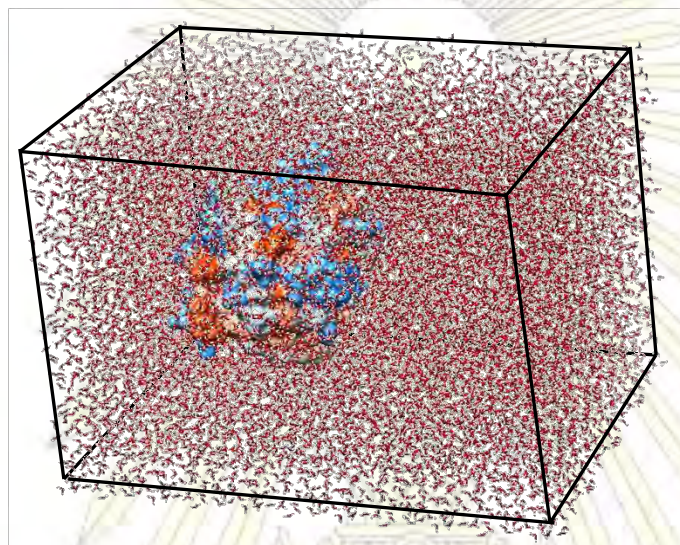
Figure 12: Unit cell for pulling simulation with size of 3.50 × 3.75 × 3.50 nm^3

The ligand optimization and a subsequent ESP calculation with basic set HF/6-31G* were performed by gaussian09 program. The RESP charge was calculated by antechamber module in AMBER program. Afterwards, charge format from AMBER was converted to gromacs format. In MD simulation, the AMBER ff03 forcefield was applied on each system. MD simulations were carried out under the periodic boundary condition with the NPT ensemble at 1 atm and 298 K using the GROMACS-4.5.5 program package in according to the standard procedures. The complex with added solvent and ion were energetically minimized using the steepest descent (SD) of 50,000 steps while the other molecules were constrained. Time step of 2 fs, a cutoff function at 1.0 nm for Van der Waal interactions and 1.0 Å for electrostatic interaction were used. The long-range electrostatic interactions were calculated via the particle mesh Ewald Sum method. Subsequently, each system was treated by heating up to 298 K for 100 ps with the *NVT* ensemble using the Berendsen

procedure. After that the simulation with *NPT* ensemble was performed at 298 K and a pressure of 1 atm for 200 ps using the Parrinello Rahman pressure couling approach to maintain a constant pressure. The simulation was fully equilibrated for 500 ps.

3.5 Steered molecular dynamics simulation

3.5.1 Pulling ligand pathway

Based on the hypothesis that the easiest pathway for escaping of ligand from active site would have the largest width and shortest depth, among the five paths in Figure 14 the path 1 was chosen for SMD study.
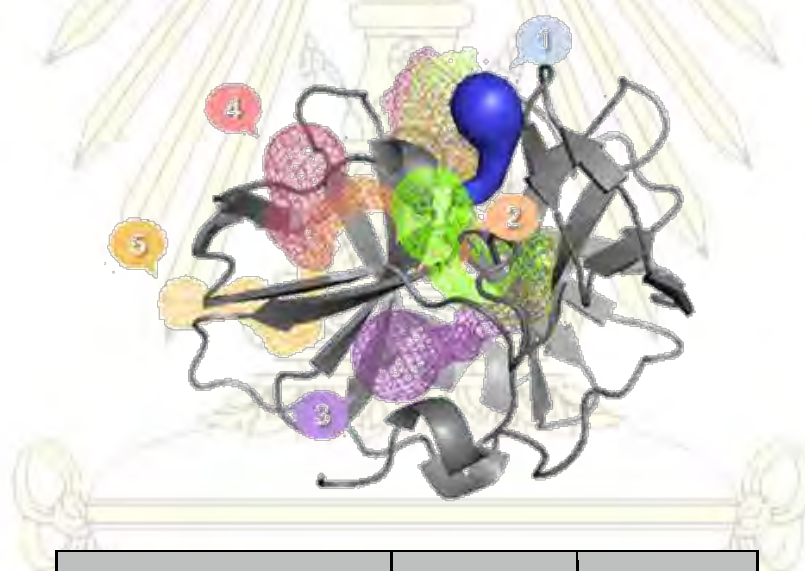
CV-A16 tunnelling path	Length (nm)	Radius (nm)
1	1.40	1.51
2	2.70	1.17
3	1.81	1.18
4	3.18	1.13
5	1.79	1.05

Figure 13: Possible paths for ligand pulling from protease

3.5.2 The pulling ligand from protein

After equilibration, the bound ligand was pulled out from the binding pocket of the protease with constant velocity (V) of 0.005 nm/ps along the Z-direction using harmonic potential. A time step of 1 fs with a van der Waals interaction cut off at 1.4 nm and the particle mesh Ewald (PME) summation method applied for the long-range electrostatic interactions were used. In equilibration for 500 ps, the C-alpha atoms of all amino acids were restrained, while the ligand was pulled from 3C protease in Z dimension, by a spring constant (k) of 600 kJ/mol•nm^2 (~996 pN/nm). The total force can be measured via equation of $F = k(vt-x)$, where x is the displacement of the pulled atom from the starting position.

Chapter IV

Results and discussion

4.1 Molecular docking results
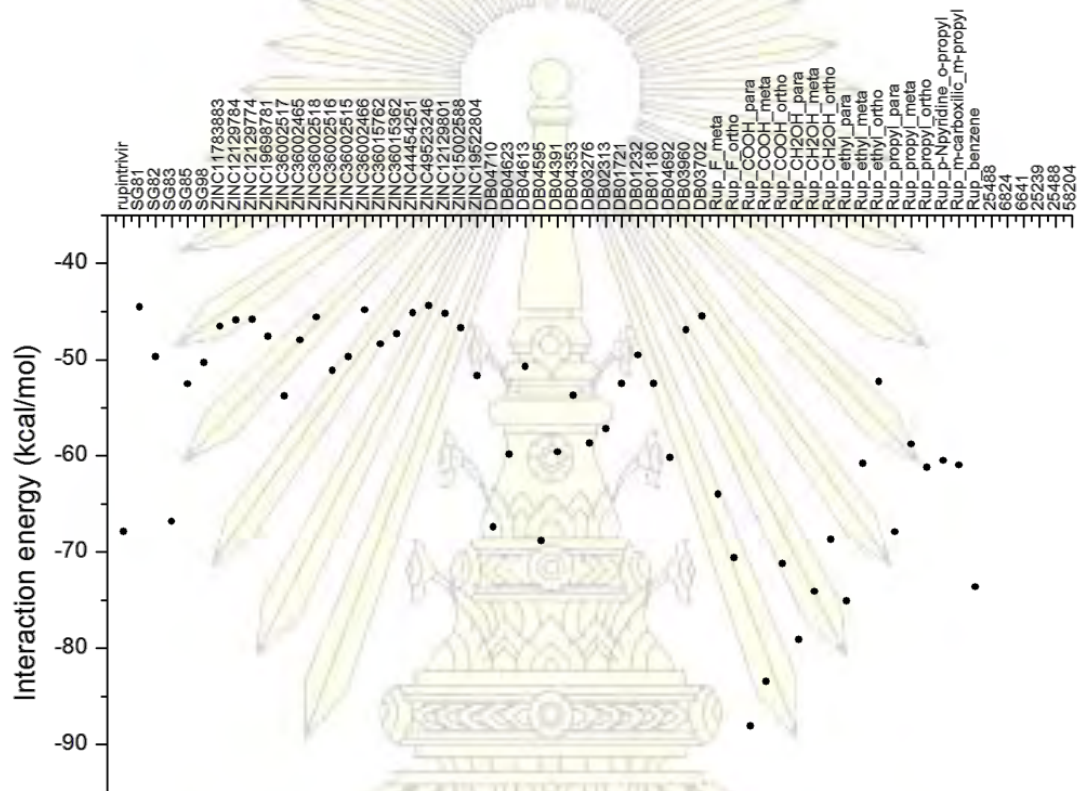
Figure 14: Binding interaction between the docked ligands and 3C protease of CV-A16

The low interaction energy between ligand and protease suggested the strong binding. From docking results in Figure 15, the top ranked ligands were chosen under the criterion of binding interaction energy less than -45 kcal/mol. And thus, the 46 selected ligands in complex with CV-A16 protease was further simulated using MD and SMD simulations.

4.2 SMD Validation
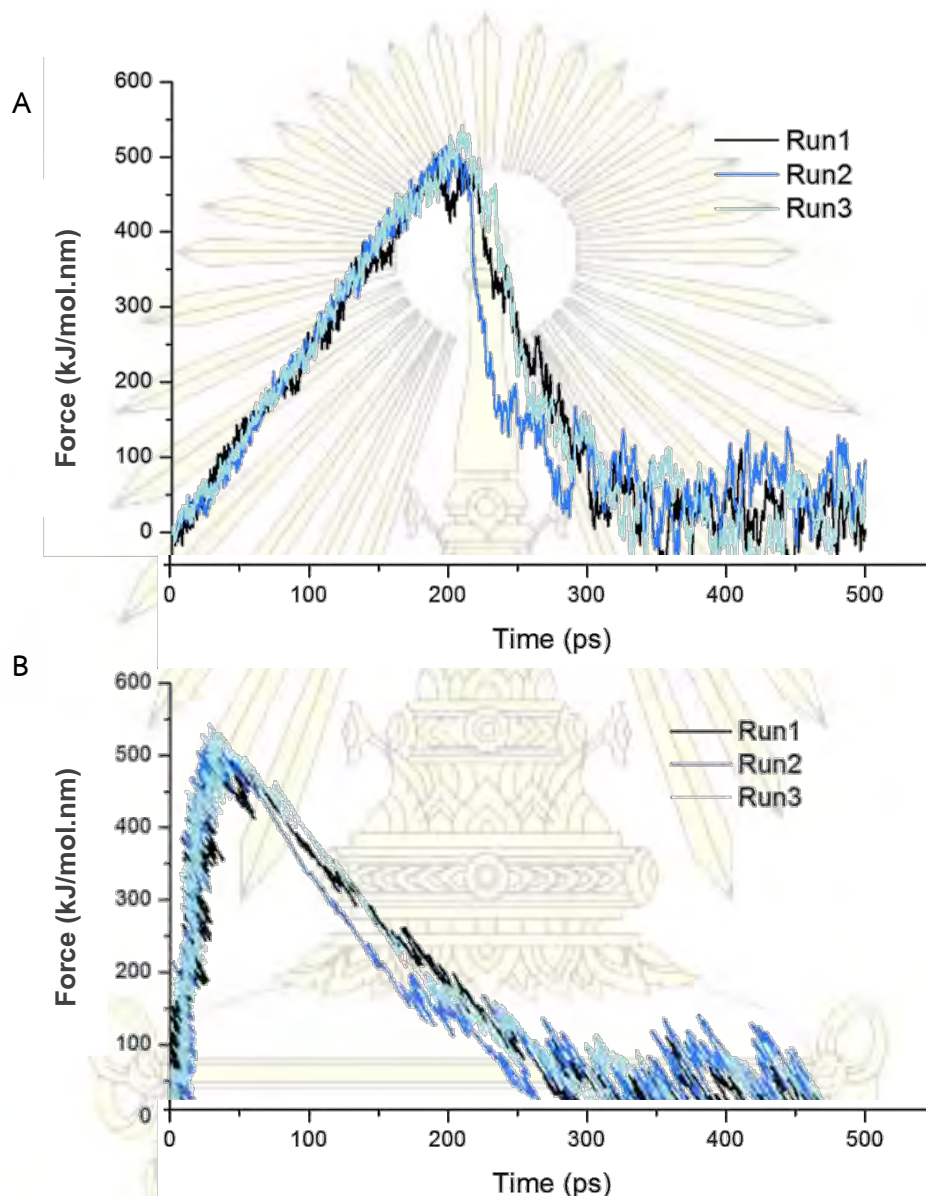
Figure 15 : Three times of the first pathway of rupintrivir and CV-A16 protease, (A) the rupture force about 514 kJ/mol•nm at 198 ps and (B) extension of ligand with protein at 0.13 nm

The derived force-time have the rupture force (F_{max}) obtain from the first tunnel, same the rupture force in three time and the extension refers to the distance between the ligand positions at time $t=t_0$ and $t=t_x$. So, in another ligands will once pulling
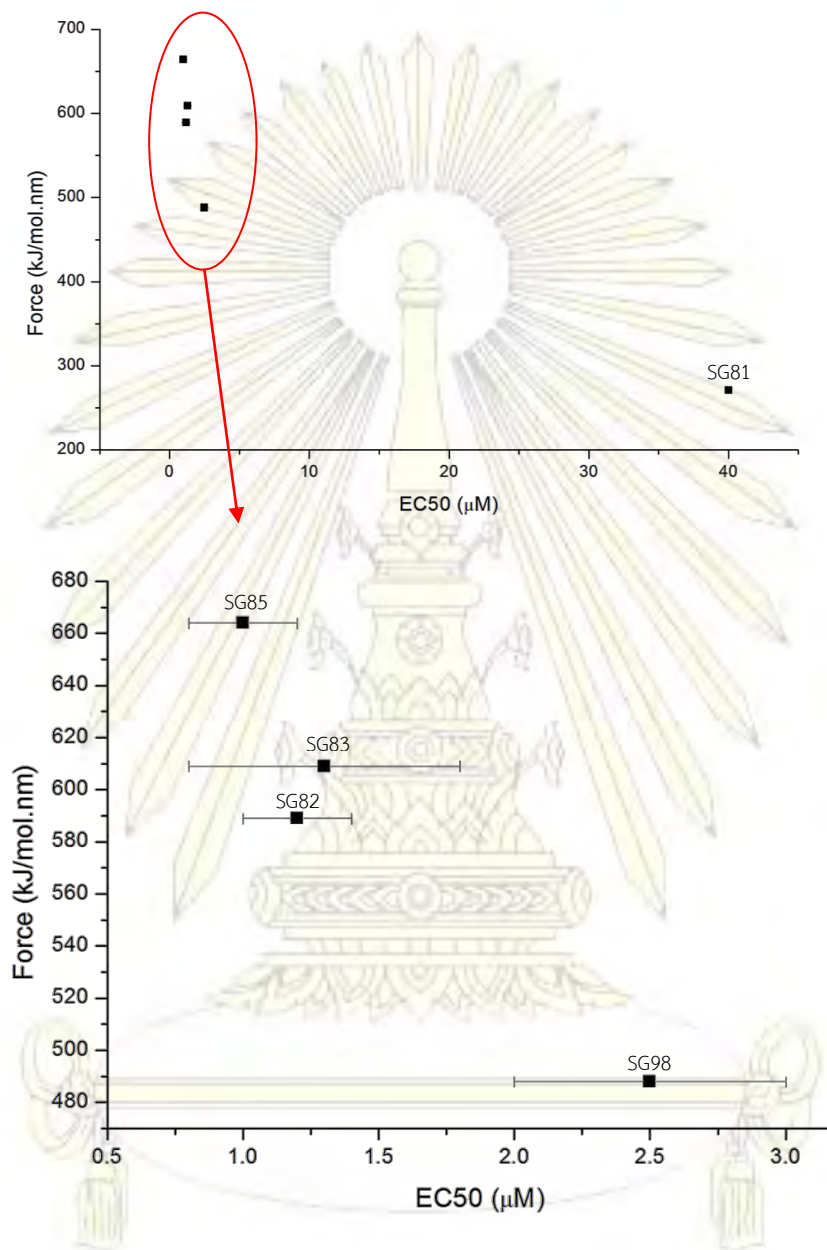
Figure 16: The correlation between the rupture force (F_{max}) and the calculated binding free energies obtained from the experimental values
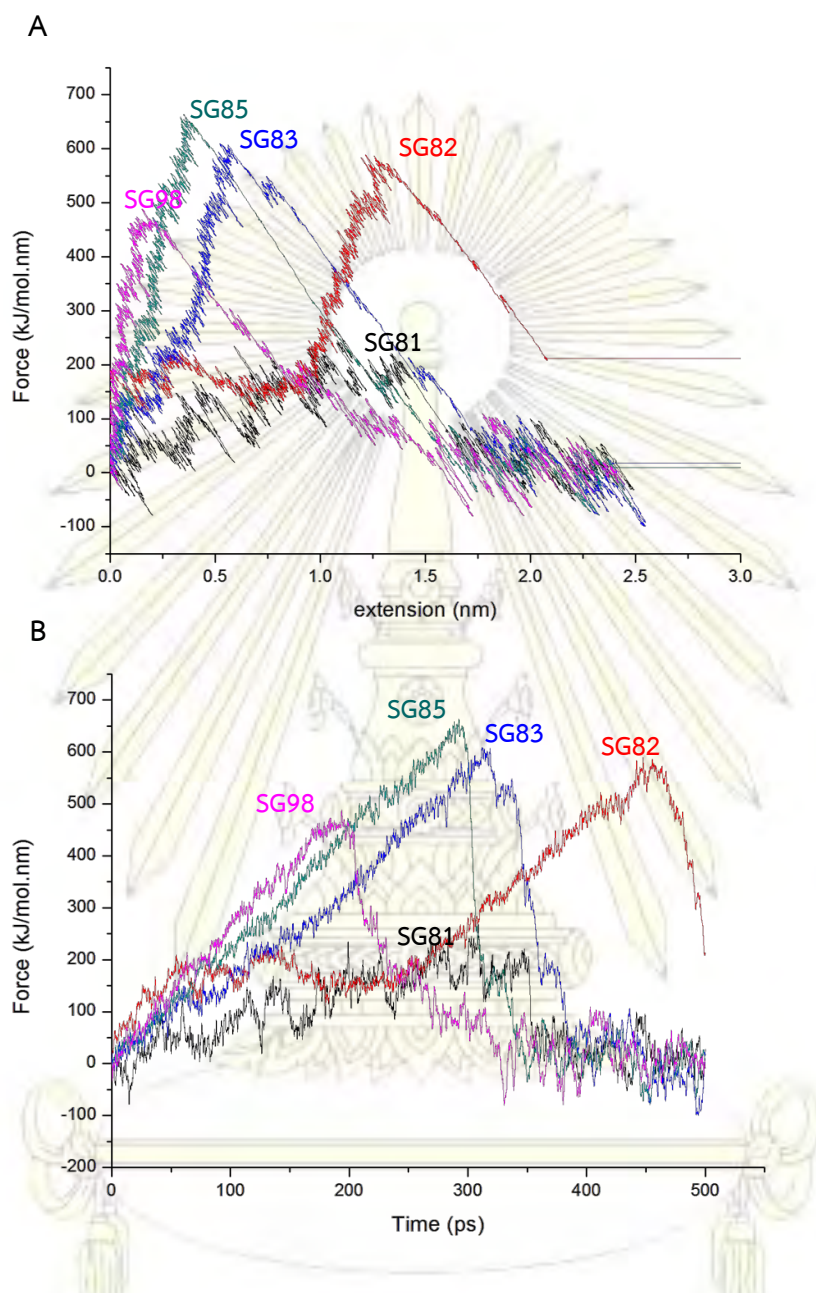
Figure 17: (A) The force-time and (B) the force-displacement profiles of the five known inhibitors pulled from the CV-A16 protease binding site through the selected tunnel (path 1 in Figure 12), where in (A) the X-axis refers to the extended distance from the reference position of the ligand (at t=0)

To test the potential reliability of the SMD simulations, the known inhibitors were pulled out from the binding pocket along path 1 and the resulted rupture forces were plotted versus the experimental EC_{50} values[15] in Figure 18. The force-

time and force-displacement profiles of each of the five known inhibitors (SG81, SG82, SG83, SG85 and SG98) are plotted in Figure 18. The force increased linearly to the maximum force, is defined as the rupture force (F_{max}).

With respect to the force-extension plot in Figure 16, the SG83 and SG85 ligands successfully escaped from the CV-A16 protease at nearly the same extension as each other at ~0.5 nm and after ~300 ps, the SG82 was freed later (~450 ps) at a extension of 1.4 nm, the SG98 was freed later (~200 ps) at a extension of 0.1 nm, the SG81 was freed later (~350 ps) at a extension of 1.5 nm which use low force pulling ligand from complex (~200 kJ/mol•nm) and the lowest activity. In comparison with the experimentally data[15], the theoretically (SMD) derived rupture force (F_{max} obtained from Figure 18) was plotted with the experimental EC_{50} values. The F_{max} values obtained from the SMD simulations corresponded relatively well to the experimental data, with ranked rupture force as SG85 > SG82 ≈ SG83 > SG98 >> SG81. Therefore, the SMD approach is a potential method to obtain qualitatively reliable binding affinities and so could be used for screening the hit-lead conpounds of the CV-A16 protease of HFMD.

4.3 Ranking of binding affinity of ligands by SMD results
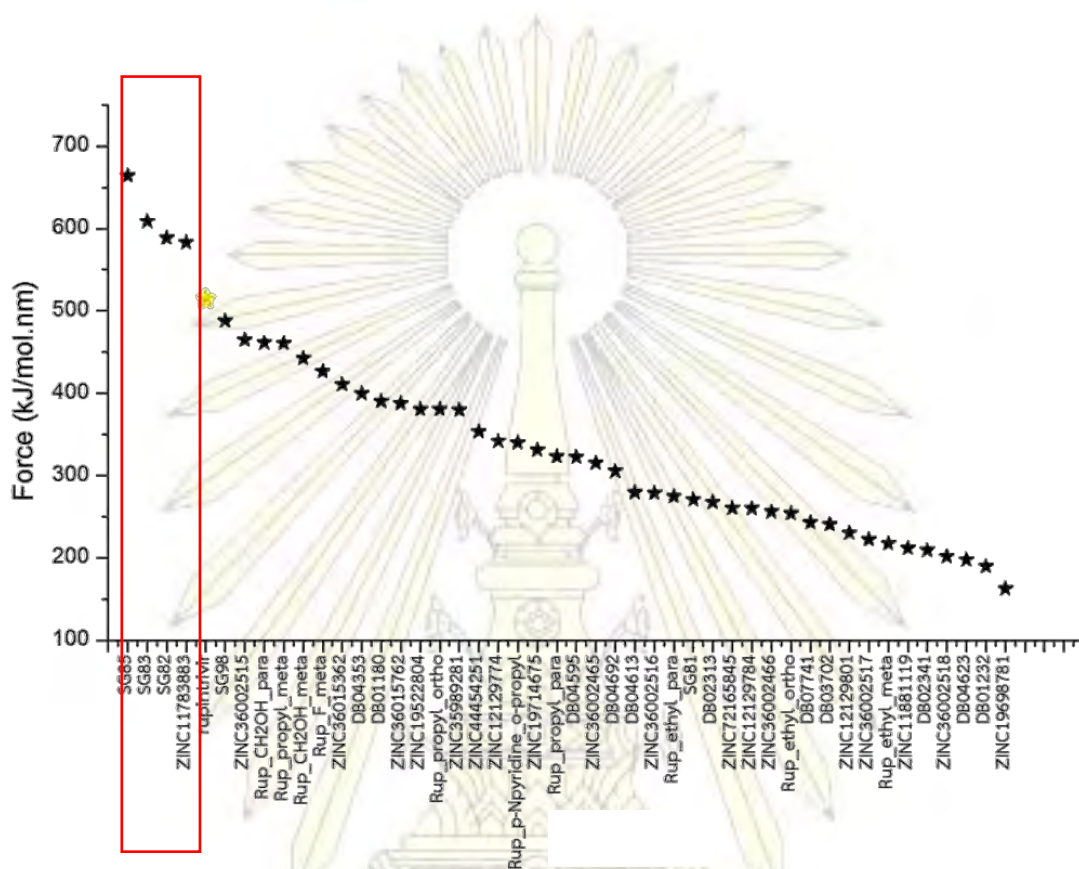
Figure 18: Ranking of the binding affinity towards the CV-A16 protease of the top 40 ligands from ZINC and DrugBank databases and rupintrivir modified analogs compared with rupintrivir (yellow) and known inhibitors

The other ligands were separately pulled out from protease binding pocket using the same procedure applied for the know inhibitor-protease complex described above. The F_{max} values for all 40 compounds were ranked with those of the six known inhibitors (Table 2) and plotted in Figure 19.

From the F_{max} values, rupintrivir, which is currently in phase II clinical trials, was predicted to have the high efficiency of ligand binding against the CV-A16 protease of HFMD with an F_{max} value of approximately 514 kJ/mol•nm. The best compound from ZINC database is ZINC11783883 (F_{max} value of 583 kJ/mol•nm). In contrast, the top ranked compounds from DrugBank database and modified analog of rupintrivir are DB04353, rup_CH$_2$OH_para, rup_propyl_meta, rup_CH$_2$OH_meta

have F_{max} value of 399, 461, 461, 442 kJ/mol•nm, respectively but not better than rupintrivir
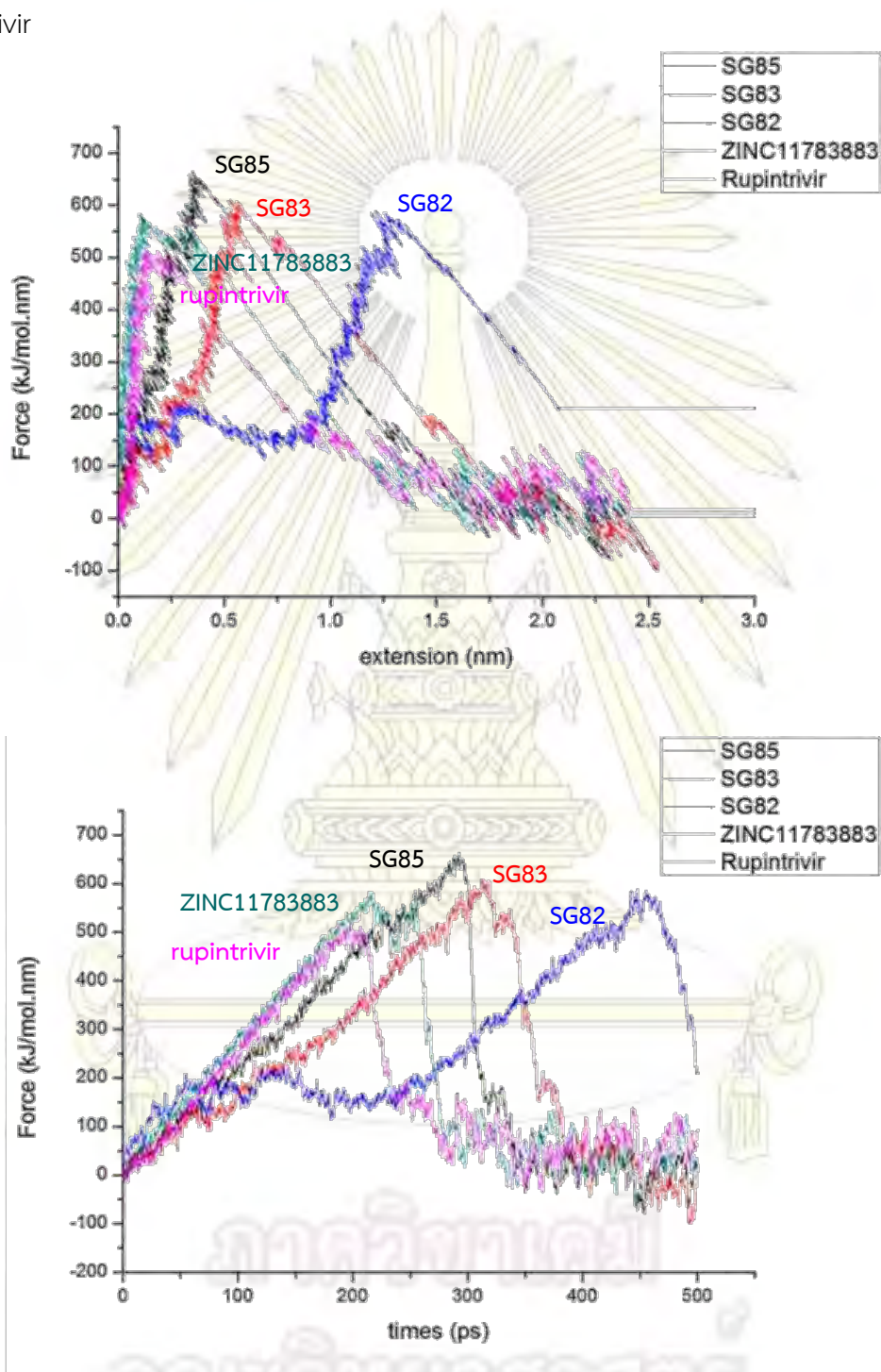
Figure 19: Comparison of the top ligands in terms of force-extension and force-time
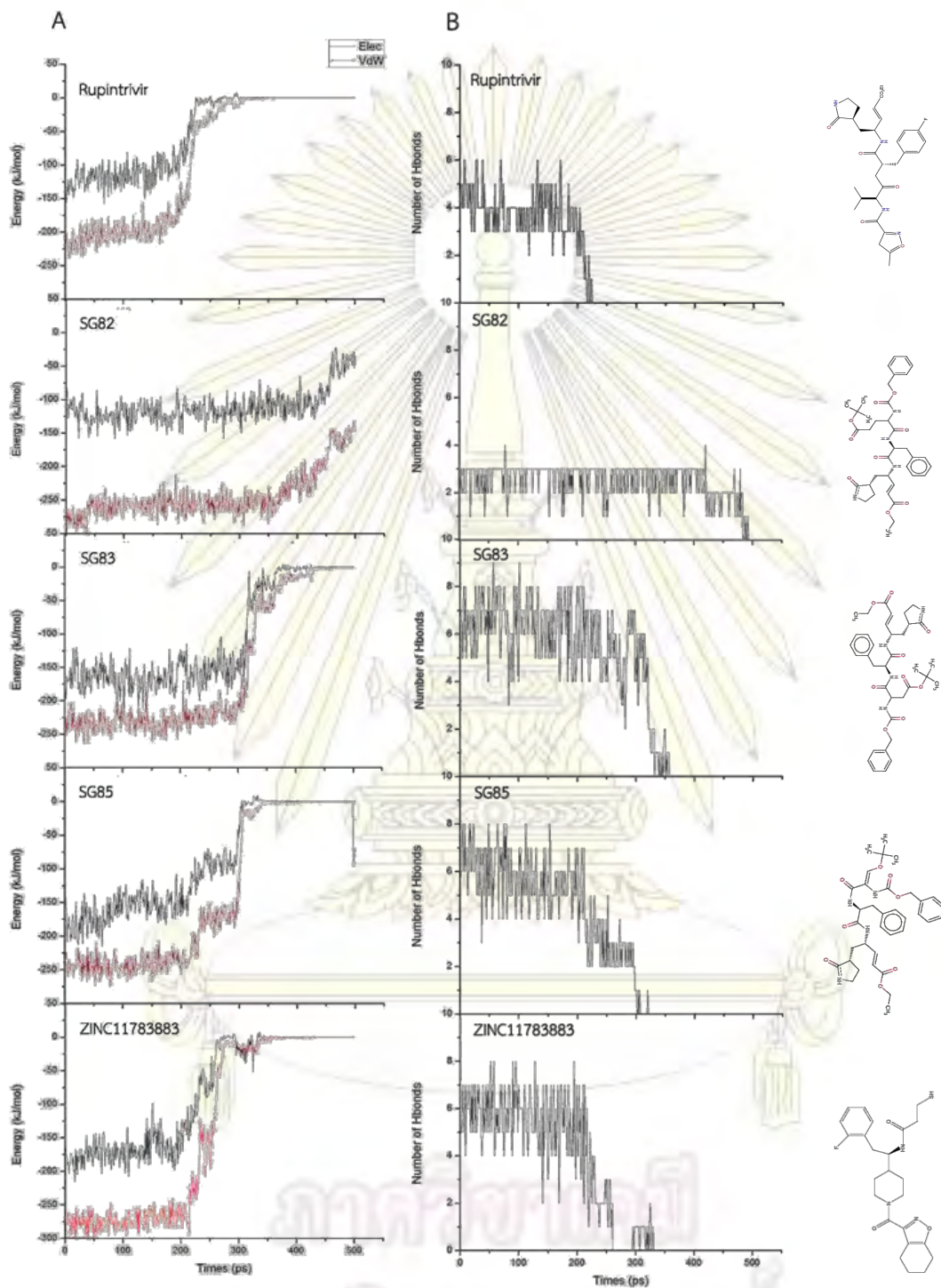
Figure 20: A is show energy of ligand with protein and B is Number of Hydrogen bonds, Both ralate with a time.

4.4 Conclusions

The SMD technique was applied to screen for potentially potent HFMD agent from the ZINC database, DrugBank database and modified analog of rupintrivir based on hypothesis that a high required rupture force (F_{max}) for pulling the ligand out of the CV-A16 binding site equates to a high predicted binding affinity and so inhibition efficiency. To validate the method, the six known HFMD inhibitor (rupintrivir, SG81, SG82, SG83, SG85 and SG98) were subjected to the same SMD analyses, where the derived F_{max} values were found to be in good agreement with the experimental EC_{50} data. According to the ranked F_{max} from the 40 top hit compounds from the databases and structural modified compounds, the ZINC11783883 was suggested to have a better inhibitory affinity than rupintrivir (but less than SG82, SG83 or SG85). Therefore, ZINC11783883 may serve as potential CV-A16 protease inhibitor or be further used as a template for lead optimization. In addition to the predicted binding affinities derived from the theoretically obtained F_{max} values, the VdW interaction was found to show a higher contribution towards stabilizing the ligand in the active site of CV-A16 protease interaction than electrostatics interaction.

Reference

1. Sarma, N. (2013). Hand, foot, and mouth disease: current scenario and Indian perspective. *Indian J Dermatol Venereol Leprol, 79*(2), 165-175. doi: 10.4103/0378-6323.107631

2. Wu Y, Lou Z, Miao Y, Yu Y, Dong H, Peng W, et al. Structures of EV71 RNA-dependent RNA polymerase in complex with substrate and analogue provide a drug target against the hand-foot-and-mouth disease pandemic in China. Protein & cell. 2010;1(5):491-500. doi: 10.1007/s13238-010-0061-7. PubMed PMID: 21203964.

3. Kaminska, K., Martinetti, G., Lucchini, R., Kaya, G., & Mainetti, C. (2013). Coxsackievirus A6 and Hand, Foot and Mouth Disease: Three Case Reports of Familial Child-to-Immunocompetent Adult Transmission and a Literature Review. *Case Rep Dermatol, 5*(2), 203-209. doi: 10.1159/000354533

4. Dragovich PS, Webber SE, Prins TJ, Zhou R, Marakovits JT, Tikhe JG, et al. Structure-based design of irreversible, tripeptidyl human rhinovirus 3C protease inhibitors containing N-methyl amino acids. Bioorg Med Chem Lett. 1999;9(15):2189-94. doi: Doi 10.1016/S0960-894x(99)00368-6. PubMed PMID: WOS:000081944000015

5. region WHOWP. Emerging disease surveillance and response2014 Jan 15,2014. Available from: http://www.wpro.who.int/emerging_diseases/HFMD/en/.

6. Shih SR, Chiang CY, Chen TC, Wu CN, Hsu JTA, Lee JC, et al. Mutations at KFRDI and VGK domains of enterovirus 71 3C protease affect its RNA binding and proteolytic activities. J Biomed Sci. 2004;11(2):239-48. doi: Doi 10.1159/000076036. PubMed PMID: WOS:000188957300012.

7. Knipe DM HP, editor. Picornaviridae: The Virus ans Their Repliation. 5th ed2007.

8. Lu G, Qi J, Chen Z, Xu X, Gao F, Lin D, et al. Enterovirus 71 and coxsackievirus A16 3C proteases: binding to rupintrivir and their substrates and anti-hand, foot, and mouth disease virus drug design. Journal of virology.

2011;85(19):10319-31. doi: 10.1128/JVI.00787-11. PubMed PMID: 21795339; PubMed Central PMCID: PMC3196414

9. Wang J, Fan TT, Yao X, Wu ZQ, Guo L, Lei XB, et al. Crystal Structures of Enterovirus 71 3C Protease Complexed with Rupintrivir Reveal the Roles of Catalytically Important Residues. Journal of virology. 2011;85(19):10021-30. doi: Doi 10.1128/Jvi.05107-11. PubMed PMID: WOS:000296253900034.

10. Marcotte LL, Wass AB, Gohara DW, Pathak HB, Arnold JJ, Filman DJ, et al. Crystal structure of poliovirus 3CD protein: virally encoded protease and precursor to the RNA-dependent RNA polymerase. Journal of virology. 2007;81(7):3583-96. doi: 10.1128/JVI.02306-06. PubMed PMID: 17251299; PubMed Central PMCID: PMC1866080.

11. Weng KF, Li ML, Hung CT, Shih SR. Enterovirus 71 3C protease cleaves a novel target CstF-64 and inhibits cellular polyadenylation. PLoS pathogens. 2009;5(9):e1000593. doi: 10.1371/journal.ppat.1000593. PubMed PMID: 19779565; PubMed Central PMCID: PMC2742901

12. Kuo C-J, Shie J-J, Fang J-M, Yen G-R, Hsu JTA, Liu H-G, et al. Design, synthesis, and evaluation of 3C protease inhibitors as anti-enterovirus 71 agents. Bioorganic & Medicinal Chemistry. 2008;16(15):7388-98. doi: http://dx.doi.org/10.1016/j.bmc.2008.06.015

13. Cui S, Wang J, Fan TT, Qin B, Guo L, Lei XB, et al. Crystal Structure of Human Enterovirus 71 3C Protease. J Mol Biol. 2011;408(3):449-61. doi: DOI 10.1016/j.jmb.2011.03.007. PubMed PMID: WOS:000290067400006.

14. Zhang XN, Song ZG, Jiang T, Shi BS, Hu YW, Yuan ZH. Rupintrivir is a promising candidate for treating severe cases of Enterovirus-71 infection. World J Gastroentero. 2010;16(2):201-9. doi: DOI 10.3748/wjg.v16.i2.201. PubMed PMID: WOS:000273713600008.

15. Binford SL, Maldonado F, Brothers MA, Weady PT, Zalman LS, Meador JW, et al. Conservation of amino acids in human rhinovirus 3C protease correlates with broad-spectrum antiviral activity of rupintrivir, a novel human rhinovirus 3C protease inhibitor. Antimicrob Agents Ch. 2005;49(2):619-26. doi: Doi 10.1128/Aac.49.2.619-626.2005. PubMed PMID: WOS:000226709900022.

16. accelrys. tutorial. Epub 3.0

17. S. Izrailev, S. S., B. Isralewitz, D. Kosztin, H. Lu, F. Molnar, W. Wriggers, and K. Schulten, *Steered Molecular Dynamics*. Springer-Verlag: Berlin, 1998; Vol. 4.

18. Barbora Kozlíková, Ondřej Strnad, Vilém Šustr, Antonín Pavelka, Martin Bezděka, Jan Byška, et al. Analysis of Tunnels in Static and Dynamic Protein Structures 2013. Available from: http://caver.cz/index.php?sid=120.

19. Tan CW, Chan YF, Sim KM, Tan EL, Poh CL. Inhibition of Enterovirus 71 (EV-71) Infections by a Novel Antiviral Peptide Derived from EV-71 Capsid Protein VP1. Plos One. 2012;7(5). doi: ARTN e34589

20. Lemkul J. GROMACS tutorial Umbrella Sampling 2012. Available from: http://www.bevanlab.biochem.vt.edu/Pages/Personal/justin/gmx-tutorials/umbrella/.

21. Ooi MH, Wong SC, Lewthwaite P, Cardosa MJ, Solomon T. Clinical features, diagnosis, and management of enterovirus 71. Lancet neurology. 2010;9(11):1097-105. doi: 10.1016/S1474-4422(10)70209-X. PubMed PMID: 20965438.

22. Medek P. BP, Sochor J. Computation of tunnels in protein molecules using Delaunay triangulation. Journal of WSCG. 2007;15(1-3):7.

23. Binford SL, Weady PT, Maldonado F, Brothers MA, Matthews DA, Patick AK. In vitro resistance study of rupintrivir, a novel inhibitor of human rhinovirus 3C protease. Antimicrob Agents Chemother. 2007;51(12):4366-73. doi: 10.1128/AAC.00905-07. PubMed PMID: 17908951; PubMed Central PMCID: PMC2167992

24. 2003. Available from: http://www.thebody.com/content/art14193.html

25. Clinic TM. Hand, Foot and Mouth Disease : signs & Symptoms May 5,2008]. Available from: http://www.mayoclinic.org/.

Appendix I : Show the interaction energy from Docking method and pulling force

from Steered Molecular dynamics simulation

No.	ID	EC50	Docking (-)	Force
		mM	kJ/mol	kJ/mol.nm
Reference				
1	rupintrivir	0.3	67.87	514.87
2	SG75	7.0 ± 1.1	51.64	-
3	SG81	> 40	44.51	271.21
4	SG82	1.2 ± 0.2	49.67	589.37
5	SG83	1.3 ± 0.5	66.81	609.54
6	SG84	2.5 ± 0.4	-	-
7	SG85	1.0 ± 0.2	52.53	664.11
8	SG98	2.5 ± 0.5	50.3	488.23
ZINC DATABASE				
9	11783883		46.5435	583.36
10	11881119		39.8666	212.58
11	12129784		45.8992	260.19
12	12129774		45.8417	342.22
13	19698781		47.5947	163.43
14	19714675		40.6643	331.16
15	35989281		40.7422	380.46
16	36002517		53.8002	222.52
17	36002465		47.9598	315.06
18	36002518		45.5841	201.59
19	36002516		51.117	278.51
20	36002515		49.6876	464.67
21	36002466		44.843	256.16
22	36015762		48.3752	387.70
23	36015362		47.3166	410.62
24	44454251		45.1163	353.50

25	44454241	42.4438	-
26	49523246	44.3956	-
27	72165845	40.4928	260.28
28	12129801	45.2011	230.53
29	15002588	46.6752	-
30	12752936	37.3045	-
31	19522804	51.7005	380.46
DRUGBANK			
32	DB04710	67.3999	-
33	DB04623	59.8576	197.68
34	DB04613	50.7078	279.44
35	DB04595	68.8372	322.78
36	DB04391	59.5914	-
37	DB04353	53.7196	399.43
38	DB03276	58.6744	-
39	DB02341	41.1102	209.31
40	DB02313	57.1961	267.36
41	DB01721	52.4953	-
42	DB01232	49.5377	189.48
43	DB01180	52.4852	390.42
44	DB07741	40.3768	242.81
45	DB04692	60.1905	305.45
46	DB03960	46.8977	-
47	DB03702	45.4784	240.69
MODIFIED			
48	Rup_F_meta	63.9869	426.75
49	Rup_F_ortho	70.5741	-
50	Rup_COOH_para	88.0961	-
51	Rup_COOH_meta	83.4482	-
52	Rup_COOH_ortho	71.1928	-
53	Rup_CH2OH_para	79.1045	461.16

54	Rup_CH2OH_meta	74.0876	442.10
55	Rup_CH2OH_ortho	68.6568	-
56	Rup_ethyl_para	75.0805	274.91
57	Rup_ethyl_meta	60.761	217.75
58	Rup_ethyl_ortho	52.2702	253.91
59	Rup_propyl_para	67.8957	323.14
60	Rup_propyl_meta	58.7837	461.07
61	Rup_propyl_ortho	61.1781	380.40
62	Rup_p-Npyridine_o-propyl	60.4803	339.69
63	Rup_m-carboxilic_m-propyl	60.9515	-
64	Rup_benzene	73.5996	-

Appendix II : Show example input file for minimization

NES STARTING WITH ';' ARE COMMENTS

title   = Minimization (fixed zn liand water)   ; Title of run

; Parameters describing what to do, when to stop and what to save

cpp   = /lib/cpp   ; location of cpp on SGI

define   = -DFLEXIBLE

constraints   = none   ; No constraints except for those defined explicitly in the topology

integrator   = steep   ; Algorithm (cg = conjugated gradient, steep = steepest descent
                                minimization)

nsteps   = 50000   ; Maximum number of (minimization) steps to perform

emtol   = 100.0   ; Stop minimization when the maximum force [kJ mol-1 nm-1]

emstep   = 0.01   ; Energy step size, used with steep

nstcomm   = 10   ; frequency for center of mass motion removal

nstxout   = 100   ; frequency to write coordinates to output trajectory file

nstvout   = 100   ; frequency to write velocities to output trajectory

nstfout   = 0   ; frequency to write forces to output trajectory

nstlog   = 100   ; frequency to write energies to log file

nstxtcout   = 100   ; frequency to write coordinates to xtc trajectory

energygrps   = system   ; Which energy group(s) to write to disk

; Parameters describing how to find the neighbors of each atom and how to calculate the
interactions

coulombtype   = PME   ; Treatment of long range electrostatic interactions

ns_type   = grid   ; Method to determine neighbor list (simple, grid)

nstlist   = 10   ; Frequency to update the neighbor list and long range forces

rlist   = 1.0   ; Cut-off for making neighbor list (short range forces)

rcoulomb   = 1.0   ; long range electrostatic cut-off (nm)

vdwtype   = cut-off

rvdw   = 1.0   ; long range Van der Waals cut-off (nm)

fourier spacing = 0.12

fourier_nx   = 0

fourier_ny   = 0

fourier_nz   = 0

pme_order   = 6

```
ewald_rtol   = 1e-5
optimize_fft = yes          ; Calculate the optimal FFT plan for the grid at startup
pbc          = xyz          ; Periodic Boundary Conditions (no=No pbc, xy=Use pbc in x and y
                              directions only, xyz=Use pbc all direction
```

Appendix III : Show example input file for Molecular Dynamics simulations

```
title                = NVT Equilibration

cpp                  = /lib/cpp          ; location of cpp on SGI

define               = -DPOSRES -DPOSRES_MOL


; Bond parameter

constraints          = all-bonds        ; Convert all bonds to constraints

constraint_algorithm = LINCS

integrator           = md               ; A leap-frog algorithm

emtol                = 1

emstep               = 0.005

dt                   = 0.002            ; unit is ps

nsteps               = 50000            ; total 100 ps.

nstcomm              = 10

comm_mode            = Linear

comm_grps            = Protein_MOL Water_and_ions

nstxout              = 1000

nstvout              = 1000

nstfout              = 0

nstlog               = 1000

nstenergy            = 1000

nstxtcout            = 1000

xtc_precision        = 5000


nstlist              = 10

ns_type              = grid

rlist                = 1.0

rcoulomb             = 1.0

rvdw                 = 1.4

coulombtype          = PME              ; Reaction-Field

vdwtype              = cut-off

fourierspacing       = 0.12

fourier_nx           = 0

fourier_ny           = 0

fourier_nz           = 0
```

```
pme_order          = 6
ewald_rtol         = 1e-5
optimize_fft       = yes
pbc                = xyz
DispCorr           = EnerPres        ; account for cut-off vdW scheme


; Berendsen temperature coupling is on in two groups
Tcoupl             = v-rescale
tc_grps            = Protein_MOL  Water_and_ions
tau_t              =  0.1     0.1
ref_t              = 310     310


; Energy monitoring
energygrps         =  Protein MOL Water_and_ions


; Pressure coupling is on
Pcoupl             = no               ; No pressure coupling, fixed box size
;Pcoupltype        = isotropic        ; Isotropic pressure coupling with time constant tau_p
;tau_p             = 0.1              ; time constant for coupling (ps)
;compressibility   = 4.5e-5           ; compressibility (bar), For water at 1 atm and 300 K the
                                      compressibility is 4.5e-5
;ref_p             = 1.0              ; reference pressure for coupling


; Generate velocites is off at 310 K.
gen_vel            = yes
gen_temp           = 310.0
gen_seed           = -1
```

Appendix IV : Show example input file for Steered Molecular Dynamics simulations

title	= smd	
cpp	= /usr/bin/cpp	
define	= -DPOSRES_CA	
constraint_algorithm	= lincs	
constraints	= none	
integrator	= md	
emtol	= 1	
emstep	= 0.005	
dt	= 0.001	
nsteps	= 500000	; 500ps
nstxout	= 1000	; write coordinate every 5 ps to output trajectory
nstvout	= 1000	; write velocity every 5 ps to output trajectory
nstfout	= 0	; write force to output trajectory
nstlog	= 1000	; write energies every 5 ps to log file
nstenergy	= 1000	; write energies to energy file every 0.5 ps
nstxtcout	= 1000	; write coordinate to xtc every 0.5 ps
xtc_precision	= 5000	
nstcomm	= 10	
comm_mode	= Linear	
comm_grps	= Protein_MOL water_and_ions	
nstlist	= 10	
ns_type	= grid	
rlist	= 1.0	
rcoulomb	= 1.0	
rvdw	= 1.4	
coulombtype	= PME	
vdwtype	= cut-off	
fourierspacing	= 0.12	
fourier_nx	= 0	
fourier_ny	= 0	
fourier_nz	= 0	
pme_order	= 6	
ewald_rtol	= 1e-5	
optimize_fft	= yes	

```
pbc                    = xyz
DispCorr               = EnerPres


; Berendsen temperature coupling is on in two groups
Tcoupl                 = v-rescale
tc-grps                = protein_MOL  Water_and_ions
tau_t                  = 0.1     0.1
ref_t                  = 310     310


; Energy monitoring
energygrps             = Protein MOL SOL CL


; Isotropic pressure coupling is now on
Pcoupl                 = Parrinello-Rahman
Pcoupltype             = isotropic
tau_p                  = 0.5
compressibility        = 4.5e-5
ref_p                  = 1.0
refcoord_scaling       = com


; Generate velocites is off at 310 K.
gen_vel                = no
;gen_temp              = 310.0
;gen_seed              = -1


; Steered Molecular Dynamics
pull                   = umbrella
pull_geometry          = direction
pull_vec1              = 0.0 0.0 1.0
pull_dim               = n n y
pull_k1                = 600
pull_group0            = Protein          ; reference group for pulling
pull_group1            = MOL              ; group to which pulling force is applied
pull_nstxout           = 10               ; write COMs of all pulled grp. every 0.01 ps
pull_nstfout           = 10               ; write force of all pulled grp. every 0.01 ps
```

pull_rate1   = 0.005
pull_start   = yes

Vitae

Miss Warin Jetsadawisut born on 11 December 1991 at Bangkok. I graduated high school in secondary school from Bodindecha (Sing Singhaseni) 2 school in Mathematics - science program at Bangkok in 2009. In 2010, I continously studied at bachelor of science, department of chemistry, chulalongkorn university. The present address after graduate : 56/151 soi Nawamin70 Nawamin road, Klongkum, Buengkum district, Bangkok, 10240