

國立臺灣大學生命科學院生化科學研究所

碩士論文

Graduate Institute of Biochemical Sciences

College of Life Science

National Taiwan University

Master Thesis

探索對抗致命病毒的潛力藥物：病毒 3C 及 3CL 蛋白酶與流感  
病毒之神經氨酸酶的抑制

Exploring the potent drugs for life-threatening viruses: Inhibition  
of 3C & 3C-Like Viral Proteases and Influenza Neuraminidase



陳建儉

Kian-Pin Tan

指導教授：梁博煌博士

Advisor: Po-Huang, Liang Ph.D.

中華民國 100 年 7 月

July, 2011

## 誌謝

一瞬間我已經在台灣六年了，順利地大學畢業，碩班也在一番奮鬥後，終於也將畢業了。一路走下來，要感謝的人太多，但是有三個人卻是在我心中有著絕對重要的份量。第一個是我的媽媽，她一直都是我的最愛也一直是我做任何事情的動力來源。這六年來，家裡發生了太多事情，媽媽卻總是默默地繼續為這個家奮鬥，任勞任怨。我希望我的畢業會讓她開心，會覺得自豪，也希望自己能快點回到她的身邊，彌補這幾年來不在她身邊所錯過的一切美好時光。另外一個人是我的女朋友，是她在背後的默默鼓勵和照顧，陪我走過了許許多多的人生低潮，我希望在未來的日子，我們能繼續一起走下去直到坐看雲起時。其實我真的很幸運，大學在許多人的協助下半工半讀順利地畢業；研究所也在蕭老師的協助下完成了學業。我想我從她的身上學會了知福、惜福再造福，我會跟隨她的腳步，在未來的日子裡秉著助人為樂的心態，積極過每一天。除此之外，我還要感謝我的家人，是他們才有今天的我。謝謝大哥的犧牲奉獻，謝謝二哥的忠言逆耳，謝謝陳姐、阿輝和阿妹的默默關懷。

當然，我還要感謝梁老師在這兩年來給了我許多指導，更不吝嗇分享他的人生經驗和知識，讓我在這兩年學習的東西足以終身受用。我還要感謝蔡陰和老師和陳昭岑老師在百忙之中抽空當我的口試委員，更感謝他們在口試的過程中給予許多寶貴的指導和指正。還有要特別感謝鴻機學長，是他幫我的實驗室生活啟蒙，也教會了我許多東西。實驗室的學長姐們源峰、至剛、國勳、世勳、楚鈞、威至、曉苓、亞菲還有學弟妹們碩甫、偉廷、哲寬、爾婷，也讓我在非常歡樂的氣氛下完成了我的研究和論文。Thanks to Vathan and Rama also for our good cooperation, and I hope we could cooperate again in the future. 當然，還要感謝董至中和周憶安讓我擁有這份得來不易的友誼，未來我們都一直是好朋友！

謝謝這六年來所有幫助過我的人，謝謝你們，有你才有今天的我！

建儉

August 2011

## 中文摘要

嚴重急性呼吸系統綜合症是一種具高度傳染性和致命性的呼吸系統疾病。它是由一種在嚴重急性呼吸系統綜合症爆發之前都沒有被發現過的新型冠狀病毒所造成的，即 SARS 冠狀病毒。這種疾病於 2003 年爆發，且有近 10% 的疾病患者因此死亡。雖然至今已經七年沒有確診病例，但是有研究顯示蝙蝠是 SARS 冠狀病毒的自然寄主，所以對於此疾病的再度爆發仍有疑慮。在開發對抗這種病毒的藥物的過程當中，科學家發現 SARS 3CL 蛋白酶在病毒的複製過程當中扮演了非常重要的角色，也因此 SARS 3CL 蛋白酶的抑制也成為了一個非常重要的藥物開發方向。根據先前的研究指出，吡啶化合物可以作為 SARS 3CL 蛋白酶可逆抑制劑，一系列的嘧啶化合物也基於此基礎上被合成。此外，我們實驗室之前也應用高通量藥物篩選發現了一些吡唑啉酮化合物可以同時有效的抑制 3C 和 3CL 蛋白酶。作為實驗的延續，一系列的吡唑啉酮化合物也因此被合成。實驗結果表明，其中一個嘧啶化合物能夠有效的抑制 SARS 3CL 蛋白酶，且半抑制率為  $6.1\mu\text{M}$ 。此外，其中一個吡唑啉酮化合物可以同時分別以  $8.4\mu\text{M}$ ，及  $9.6\mu\text{M}$  的半抑制率有效地抑制 SARS 3CL 蛋白酶及柯薩奇 B3 3C 蛋白酶。

流感病毒是一種單股負鏈的 RNA 病毒，同時它也是造成的季節性流感和流感大流行的罪魁禍首。血凝素和神經氨酸酶在流感病毒的生活史中扮演了非常重要的角色。其中，神經氨酸酶催化唾液酸的水解協助成熟流感病毒脫離宿主細胞感染新的細胞，因此抑制流感神經氨酸酶一個非常重要的藥物開發方向。我們的實驗室在之前的高通量藥物篩選中發現，VK84 可以以  $16.6\mu\text{M}$  的半抑制率，有效的抑制 H5N1 的神經氨酸酶。之後我們也根據電腦模擬的結果為 VK84 做了改進，合成了一系列的衍生物。VK84 是一種吡唑啉酮化合物，它的衍生物之一 VK94 經實驗證明能以  $1.7\mu\text{M}$  證明能的半抑制率，有效的抑制 H1N1 神經氨酸酶。根據對接結果，我們發現 VK94 是一個非常有吸引力的和潛力可以進行改良的抑制劑，並且通過改良從而更有效的抑制神經氨酸酶。

此外，為了提高高通量藥物篩選準確度，並且從中節省資金和精力，我們應用了 Catalyst 4.10 這套軟體建構了一個 3D-QSAR 藥效團。經過一系列的驗證，Hypo1 似乎是一個具預測能力並且可用來進行虛擬高通量藥物篩選的可靠模型。總體而言，一個從虛擬高通量藥物篩選到抑制劑檢定的完整實驗流程已經被完整建立，並且肯定將在未來的藥物開發過程中扮演一定的角色。



## ABSTRACT

Severe acute respiratory syndrome (SARS) is a highly contagious and fatal respiratory disease. It is caused by a novel coronavirus which never been discovered before the SARS, namely SARS-CoV. The outbreak of this deadly disease in 2003 had killed almost 10% of the infected patients. Although it has been 7 years since the last reported case of SARS, the evidences that bats are natural reservoirs of SARS-CoV have led to the concerns about the reemergence of the deadly disease. In search of the potential drugs against the deadly disease, SARS 3CL<sup>pro</sup> a protease which playing essential roles during the viral replication has been targeted for the drugs development. A series of pyrimidine compounds have been synthesized based on the findings that pyridine compounds could act as SARS 3CL<sup>pro</sup> reversible inhibitors. Furthermore, a series of pyrazolone compounds have also been synthesized as part of our laboratory's continuation work of high throughput screenings. Some of the pyrazolone compounds have also been identified as common inhibitors for both 3C and 3CL protease, so that the Coxsackievirus B3 3C<sup>pro</sup> have also been purified for further inhibition assay. The results show that one of the pyrimidine analogs could inhibit the SARS 3CL<sup>pro</sup> with an IC<sub>50</sub> of 6.1μM. Meanwhile, a pyrazolone analog could inhibits SARS 3CL<sup>pro</sup> as well as inhibits CVB3 3C<sup>pro</sup> with an IC<sub>50</sub> of 8.4μM and 9.6μM respectively.

Influenza viruses are negative sense, single-stranded, segmented RNA viruses. It is the viruses which cause the seasonal flu and even flu pandemics. Two glycoproteins, hemagglutinin and neuraminidase play very crucial roles during the influenza virus infection and replication. Neuraminidase facilitates the release of the replicate viruses from the infected cells by enzymatically cleavage of the sialic acid groups from

host glycoproteins and hence it is a very good target for drugs development. VK84, an original hit that inhibits the H5N1 neuraminidase with an  $IC_{50}$  of 16.6  $\mu\text{M}$  has been modified based on the computer modeling and a series of analogs were then been synthesized. VK84 is also a pyrazolone compound and one of its analogs, VK94, inhibits the H1N1 neuraminidase with an  $IC_{50}$  of 1.7 $\mu\text{M}$ . According to the docking results, VK94 shows some unexpected H-bonds and this made VK94 as an attractive and potential inhibitor for further modifications as pyrazolone compounds that have never been reported as neuraminidase inhibitors.

Furthermore, in order to increase the accuracies and to save money and efforts for high throughput screening, Catalyst 4.10 was utilized to generate a reliable and predictive 3D-QSAR. The generated pharmacophore, Hypo1 has been validated and it seems that it is a reliable and predictive pharmacophore for virtual high throughput screening. In general, a complete procedure from the virtual high throughput screenings then to the inhibition assays has been developed.

## **Abbreviations**

Severe acute respiratory syndrome

3C protease

3CL protease

Influenza virus

Neuraminidase

Pyrimidine

Pyrazolone

3D-QSAR pharmacophore





## 中文摘要

嚴重急性呼吸系統綜合症

3C 蛋白酶

3CL 蛋白酶

流感病毒

神經胺酸酶

嘧啶化合物

吡唑啉酮化合物

3D-構效關係藥效團





## CONTENTS

口試委員會審定書	i
誌謝	ii
中文摘要	iii
Abstract	iv
<b>Chapter 1 INTRODUCTION</b>	
1.1 SARS-Coronavirus and SARS 3C-like Protease	1
1.2 Coxsackie Virus B3 and 3C Protease	3
1.3 Influenza Virus and Neuraminidase	5
<b>Chapter 2 MATERIALS AND METHODS</b>	
2.1 Inhibitors	8
2.2 Enzyme Expression and Purification	
2.2.1 SARS-3CL <sup>pro</sup>	8
2.2.2 CVB3 3C <sup>pro</sup>	9
2.2.3 H5N1 and H1N1 Viral Neuramidases	9
2.3 Enzyme Inhibition Assay and IC <sub>50</sub> Determination	
2.3.1 SARS 3CL <sup>pro</sup> and CVB3 3C <sup>pro</sup> Inhibition Assay	11
2.3.2 Influenza Viral Neuraminidase Inhibition Assay	11
2.3.3 IC <sub>50</sub> Determination	12
2.4 Computer Modeling	

2.4.1	Docking Studies of SARS 3CL <sup>pro</sup> and CVB3 3C <sup>pro</sup> -----	12
2.4.2	Docking Studies of Viral Neuraminidase-----	13
2.5	3D-QSAR Pharmacophore Generation	
2.5.1	Training Set-----	13
2.5.2	Generation of Hypo1-----	14
2.5.3	Cost Analysis-----	14
2.5.4	CatScramble-----	15
2.5.5	Test Set-----	16
<b>Chapter 3</b>	<b>RESULTS</b>	
3.1	Expression and Purification of SARS 3CL <sup>pro</sup> -----	17
3.2	Inhibition of Pyrimidine Compounds Against SARS 3CL <sup>pro</sup> and Their Binding Modes-----	17
3.3	Inhibition of Pyrazolone Compounds Against SARS 3CL <sup>pro</sup> and Their binding modes-----	18
3.4	Expression and Purification of CVB3 3C <sup>pro</sup> -----	20
3.5	Inhibition of Pyrazolone Compounds Against CVB3 3C <sup>pro</sup> and Their Binding Modes-----	20
3.6	Expression and Purification of H1N1 and H5N1 Neuraminidase-----	21
3.7	Inhibition of Pyrazolone Compounds Against the H1N1 and H5N1 NA-----	22
3.8	Generation of the Catalyst Hypotheses-----	24
3.9	Validation of the Hypo -----	24

**Chapter 4 DISCUSSIONS**

4.1 SARS-CoV 3CL<sup>pro</sup> and its Inhibitors-----26

4.2 Neuraminidase and its Inhibitors -----26

4.3 Hypo1 and its Modification -----27

**FIGURES**

Figure 1-----29

Figure 2a & 2b-----30

Figure 3a & 3b-----31

Figure 4a & 4b-----32

Figure 5a & 5b-----33

Figure 6a & 6b-----34

Figure 7a & 7b-----35

Figure 7c-----36

Figure 8a-----37

Figure 8b-----38

Figure 9 -----39

**TABLES**

Table 1-----40

Table 2-----41

Table 3-----42

Table 4-----43

Table 5-----44

Table 6-----45

**GRAPHS**

Graph 1-----46

Graph 2-----47

**SUPPLEMENTARY**

Supplementary figure 1 -----48

Supplementary figure 2 -----49

Supplementary table 1-----55

Supplementary table 2-----56

Supplementary table 3-----59

**APPENDIXES**

Appendix 1-----60

Appendix 2-----61

Appendix 3-----62

**REFERENCES** -----63



## Chapter 1 INTRODUCTION

### 1.1 SARS-Coronavirus and SARS 3C-like Protease

In November 2002, it was reported an emergence of severe acute respiratory syndrome (SARS) as a highly contagious and fatal respiratory disease. This infectious and deadly disease was first identified in Guangdong Province and then rapidly spread to over 25 countries (Lee, Hui et al. 2003). More than 8000 individuals were confirmed to be infected of which 9.6% patients died within a few months (Consortium. 2004). The most common symptoms of this atypical disease are cough, high fever (over 38 °C), malaise rigor, headache, chills, and progressive radiographic changes of the chest and lymphopenia (Peiris, Lai et al. 2003). A novel human coronavirus was identified as the causative viruses of SARS and was named SARS coronavirus (SARS-CoV) (Drosten, Gunther et al. 2003). SARS-CoV is mostly closed to group 2 coronaviruses, but it does not belongs to any one of three groups of coronaviruses as (Appendix 1) (Ksiazek, Erdman et al. 2003). Since the last infected human case was reported in June 2003, the spread of SARS is regarded fully contained. However, recent finding reports the isolation of SARS-CoV from animal including bats leads to a concern that would the SARS outbreak happen again (Guan, Zheng et al. 2003; Lau, Woo et al. 2005; Li, Shi et al. 2005). Due to the great impact and the high mortality rate it brings to the human, it is critical to explore the therapy for the SARS. There are no any effective antiviral drugs and vaccines have been marketed to date.

SARS-CoV is an enveloped, positive stranded RNA virus belonging to Coronaviridae. The SARS-CoV genome comprises about 29,700 nucleotides, which is one of the largest among RNA viruses. SARS-CoV is similar to other coronaviruses in that its

genome expresses two overlapping polyproteins namely PP1a (486 KDa) and PP2a (790 KDa). These overlapping polyproteins mediate all the functions required for viral replication and transcription (Marra, Jones et al. 2003). The polyproteins are then processed by two proteases that encoded by ORF1a and this process eventually leads to the maturation of the SARS-CoV. The proteases catalyze their own release from the polyproteins and other non-structural proteins (nsps) from the polyproteins and initiate virus mediated RNA replication (Ahlquist, Noueir et al. 2003). One of the proteases, which plays the very important roles in cleavage of the polyproteins, is Coronavirus main protease or 3C-like protease ( $3CL^{pro}$ ) with chymotrypsin fold (Snijder, Bredenbeek et al. 2003).

$3CL^{pro}$  cleaves the large replicase polyprotein at 11 cleavage sites, releasing the non-structural proteins including  $3CL^{pro}$  itself (nsp5), an RNA-binding protein (nsp9), an RNAdependent RNA polymerase (nsp12), an NTPase/RNA 5'-triphosphatase/helicase (nsp13), a nidoviral uridylate-specific endoribonuclease (nsp15), and two proteins containing a predicted exonuclease activity (nsp14) and an S-adenosylmethionine-dependent ribose 2'-*O*-methyltrans-ferase activity (nsp16) (Chen, Gui et al. 2005). This protease is hence absolutely essential for the viruses to make copies of itself and spread the infection, eventually the  $3CL^{pro}$  is considered as an attractive target for anti-SARS and other coronavirus infections. Furthermore, the availability of multiple crystal structures of the enzyme with co-crystallized ligands makes the target amenable to structure-based drug design.

The SARS-CoV  $3CL^{pro}$  is a dimeric protein with the molecular weight of 34.6 KDa of each protomer. Its active site contains a catalytic dyad where Cys145 acts as a

nucleophile and His41 acts as the general acid base (Anand, Ziebuhr et al. 2003). The sequences cleaved by the protease are well conserved, i.e there are a glutamine at P1, a hydrophobic residues at P2, and a small amino acid at P1' (Fan, Wei et al. 2004). The well conserved sequences have also led to the innovation of a new enzyme inhibition assay using a fluorogenic substrate, that is Dabcyl-KTSAVLQ↓SGFRKME-Edans, where ↓ indicated the cleavage site (Kuo, Chi et al. 2004).

A series of aromatic esters with the 3-chloropyridine moiety have been reported as potent SARS 3CL<sup>pro</sup> inhibitors with IC<sub>50</sub> in the nanomolar range (Zhang, Pettersson et al. 2007; Niu, Yin et al. 2008). It is also found that the binding of 3-chloropyridine moiety into the S1 pocket of the active site is preferable. LC-MS results showing that those inhibitors exert covalent bond binding to inhibit the protease and the result has led to the concern that those molecule are not suitable for drugs development (Zhang, Pettersson et al. 2007). A series of aryl methylene ketones and fluorinated methylene ketones with 3-chloropyridine moieties were then been synthesized and the evaluation of those molecules giving the IC<sub>50</sub> as low as 13 μM (Zhang, Huitema et al. 2008). In search of the reversible inhibitors for the SARS 3CL<sup>pro</sup>, a series of pyrimidine compounds have been synthesized and their activities against the SARS 3CL<sup>pro</sup> have been evaluated. Furthermore, a series of pyrazole compounds have been identified as 3C and 3CL protease common inhibitors, thus a series of pyrazolone analogs have been synthesized based on the computer modeling.

## 1.2 Coxsackie Virus B3 and 3C Protease

Coxsackieviruses (CVs) are positive sense, single-stranded RNA viruses, which belong to the Picornaviridae family (Fields 1985). Symptoms of infection with viruses in



the Coxsackie B group include fever, headache, sore throat, gastrointestinal distress, as well as chest and muscle pain. It is also known as pleurodynia or Bornholm disease in many areas. The infection may cause the myocarditis or pericarditis in some cases, which can eventually result in permanent heart damage or death. Aseptic meningitis is one of the potential symptoms that might be induced by Coxsackie B virus infection. Up to 50% of the unexpected sudden deaths of Coxsackievirus infections are related to this group (Maze and Adolph 1990).

Even though the patients suffered by the CVs infection would normally recover within few days, but drugs medication is needed in some severe infections. Unfortunately, there are no any drugs or vaccines available for CVs infection so far. A recent finding also shows that Coxsackie B3 viruses (CVB3) infection can reproduce human clinical heart disease in rodents (Rose and Hill 1996). Furthermore, CVB3 have been identified from patients with cardiovascular disease (Baboonian and Treasure 1997). It is more likely to develop effective drugs against CVs infection.

The 3C<sup>pro</sup> essential for CVB3 replication had served as a drug target (Binford, Maldonado et al. 2005; Lee, Lee et al. 2007; Kuo, Shie et al. 2008). Both 3C<sup>pro</sup> and 3CL<sup>pro</sup> have similar 3-D structures (Appendix 2), but unlike the dimeric 3CL<sup>pro</sup>, 3C<sup>pro</sup> is monomeric and utilizes Glu71-His41-Cys147 triad for catalysis (Lee, Kuo et al. 2009). Fluorogenic substrate, Dabcyl-KTSAVLQ↓SGFRKME-Edans could be used in the CVB3 3C<sup>pro</sup> inhibition assay as well due to the similarity in substrate specificity between SARS-CoV 3CL<sup>pro</sup> and CVB3 3C<sup>pro</sup> (Lee, Kuo et al. 2009). Furthermore, it is shown that some of the SARS 3CL<sup>pro</sup> inhibitors could also act as the CVB3 3C<sup>pro</sup> inhibitors as well (Kuo, Liu et al. 2009). In search of the potent inhibitors for both SARS 3CL<sup>pro</sup> and CVB3 3C<sup>pro</sup>, a

series of pyrimidines and pyrazolones have been synthesized, and their activity against the proteases have been evaluated.

### **1.3 Influenza Virus and Neuraminidase**

The 1918 flu pandemic (the Spanish Flu) was an influenza pandemic that emerged in 1918. The Spanish Flu pandemic was an extremely severe and deadly pandemic that spread across the world during the World War I. It is believed that 3% of the world population at that time, about 50 million people died of the disease. A few decades later, the emergence of the aggressive avian H5N1 influenza virus in 2003, particularly in Asia, has made the likelihood of a human influenza pandemic and the possible socio-economic impact a major worldwide concern. Luckily, the H5N1 did not cause the outbreak of another pandemic due to the poor infection rate though it is highly pathogenic. In April 2009, a novel flu strain combined genes from human, pig, and bird flu has emerged in Mexico. This strain of influenza virus, which was given name as influenza A/H1N1 later, spread all over the world within few months. The mortality rate of influenza A/H1N1 is relatively lower than the other flu pandemics, but the emergence of drug-resistant strain during the pandemic has led to the concern that the drug-resistant strain would cause the outbreaks of the next pandemic. On the other hand, influenza spreads around the world in seasonal epidemics, resulting in the deaths of between 250,000 and 500,000 people every year, up to millions in some pandemic years.

Vaccination against the influenza with influenza vaccines is highly recommended to the high risk groups such as children and elderly. Drug treatments are also available since the drugs amantadine and rimantadine targeting on the M2 ion channel of the influenza

viruses were discovered in 1963 and 1966 (Pinto and Lamb 2006). However, drug-resistant strains emerged rapidly and hence these drugs are no more recommended for the infection treatment (Hayden, Belshe et al. 1989; Cheng, Leung et al. 2009).

Influenza viruses are negative sense, single-stranded, segmented RNA viruses. Two glycoproteins, hemagglutinin and neuraminidase play very crucial roles during the influenza virus infection and replication. Hemagglutinin is responsible for the recognition of the targeted vertebrate cells and hence accomplishes the binding of the influenza viruses onto the target cells by the interaction between the glycoproteins and the sialic acid containing receptors. It also facilitates the entry of the viral genome by causing the targeted cells membrane and the viral membrane fusion (Skehel and Wiley 2000). Neuraminidase on the other hand facilitates the release of the replicate viruses from the infected cells by enzymatically cleave the sialic acid groups from host glycoproteins (Huang, Li et al. 2008). The importance of these glycoproteins to the influenza virus infection and replication has made them very attractive to be the drug targets.

Zanamivir, a drug using Relenza as a tradename under GlaxoSmithKline has been marketed in 1999. It is a neuraminidase inhibitor which was designed by getting aid of the enzyme X-ray structure and computer modeling (von Itzstein, Wu et al. 1993). A few months later, another drug known as Oseltamivir or GS4104, which also acts as a neuraminidase inhibitor, has also been marketed (Lew, Chen et al. 2000). Both of the inhibitors could inhibit the neuraminidase with the  $IC_{50}$  around 1 nM (Zhang, Yu et al. 2006). Zanamivir is a transition analog inhibitor which inhibits the neuraminidase by forming various H-bonds between the inhibitor and the amino acids of the neuraminidase. Oseltamivir inhibits the neuraminidase by a mechanism similar to the zanamivir except of

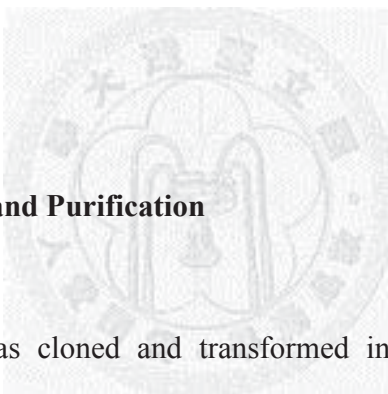
the pentyloxy substitution of the oseltamivir induces the reorientation of the Glu276 and hence the hydrophobic pocket is exposed. The exposure of the hydrophobic pocket provides an extra hydrophobic interaction for the inhibitor and leads to a better interaction (Russell, Haire et al. 2006) (Appendix 3). A few years later, peramivir and laninamivir, the neuraminidase inhibitors which exert the inhibition mechanism similar to the oseltamivir have also been developed (Babu, Chand et al. 2000; Makoto Yamashita 2009).

The rapid mutations of influenza viruses have led to the emergence of the oseltamivir and the peramivir-resistant influenza viruses (Russell, Haire et al. 2006; Memoli, Hrabal et al. 2010). There is no any zanamivir-resistant strain reported so far and the researchers have also found that both of the zanamivir and laninamivir work against the oseltamivir-resistant strains (Watanabe, Chang et al. 2010) but the poor oral bioavailability of the drugs has brought some limitations. In search of potent neuraminidase inhibitors with good oral bioavailability, high-throughput screenings have been performed. A series of the analogs of the original hit (VK84) have also been synthesized in order to improve the inhibition against influenza neuraminidase.

## Chapter 2 MATERIALS AND METHODS

### 2.1 Inhibitors

A series of 2-(benzylthio)-6-oxo-4-phenyl-1, 6-dihydropyrimidine (Ramajayam, Tan et al. 2010) and 4-benzylidene-1, 3-diphenyl-5-pyrazolone (Ramajayam, Tan et al. 2010) analogs have been synthesized by Dr. Ramajayam Ramachandran in our laboratory to test against the inhibition to the SARS-CoV 3CL protease and CVB3 3C protease. A series of 1-phenyl-5-pyrazolone analogs have been synthesized by Mr. Vathan Kumar to test against the inhibition to the SARS-CoV 3CL protease and influenza virus neuraminidase.



### 2.2 Enzyme Expression and Purification

#### 2.2.1 SARS-3CL<sup>pro</sup>

The SARS-3CL<sup>pro</sup> was cloned and transformed into *E. Coli BL21* for protein expression as reported (Kuo 2009). The 5-mL overnight culture of a single transformation was used to inoculate 500mL of fresh LB medium containing 100µg/mL ampicillin. The cells were grown to A600=0.6 and induced with 1mM isopropyl-β-thiogalactopyranoside. After 4–5 h, the cells were harvested by centrifugation at 6000g for 20 min. The enzyme purification was conducted at 4 °C. The SARS 3CL<sup>pro</sup> cell paste obtained from 2-L cell culture was suspended in 80mL lysis buffer containing 12mM Tris–HCl, pH 7.5, 120mM NaCl, and 0.1mM EDTA. Constant cell disruption system was used to disrupt the cells at 20,000 psi. The lysis solution was centrifuged and the debris was discarded. The cell free extract was loaded onto a 20mL Ni–NTA (Qiagen) column which was equilibrated with

12mM Tris-HCl, pH 7.5, 120mM NaCl, 0.1mM EDTA, and 5mM imidazole. The column was washed with 5mM imidazole followed by 30mM imidazole-containing buffer. His-tagged protease was eluted with 300mM imidazole-containing buffer. The protein solution was dialyzed against 2X2 L lysis buffer in the presence of different combinations of reducing agents (7.5mM b-ME, 7.5mM b-ME plus 1mM DTT, 17.5mM b-ME, or 2mM DTT). SARS His-tagged 3CL<sup>pro</sup> was then digested with FXa (Novagen) protease to remove the tag and the mixture was loaded onto Ni-NTA. The untagged protease in flowthrough (12mM Tris-HCl, pH 7.5, 120mM NaCl, 0.1mM EDTA, and 5mM imidazole containing the reducing agents) was dialyzed to buffer (12mM Tris-HCl, pH 7.5, 120mM NaCl, and 0.1mM EDTA with reducing agents) for storage.

### **2.2.2 CVB3 3C<sup>pro</sup>**

The CVB3 3C<sup>pro</sup> was cloned and transformed into E. Coli BL21 for protein expression as reported (Kuo 2009). The expression procedure of CVB3 3C<sup>pro</sup> is almost same with the SARS 3CL<sup>pro</sup> except of the lysis buffer (25 mM Tris-HCl, pH 7.5, 120 mM NaCl, 0.1mM EDTA). His-tag need not to be removed as it did not affect the protease catalytic activity (Kuo 2009).

### **2.2.3 H5N1 and H1N1 Viral Neuramidases**

H5N1 and H1N1 viral neuramidases were cloned into the baculovirus transfer vector by Dr. Yuan-Feng Lin. H5N1 and H1N1 full-length gene of the NA from A/Hanoi/30408/2005(H5N1) and A/WSN/1933 TS61(H1N1) were synthesized by Genearth Company (BioPark, Germany). In order to amplify H5N1 gene, DNA encoding N-terminal His-tagged ectodomain (63-449) of NA protein was constructed using sticky-end polymerase chain reaction (PCR) with *EcoRI/PstI* restriction site and cloned into the

baculovirus transfer vector, pAcGP67-A (BD Biosciences Pharmingen, San Jose, CA). The primers used in the sticky-end PCR were designed as follows: forward primers 1 (5'-AATTCCACCATCACCATCACCATGTGAAGCTGGCTGGTAACTCCTCC) and 2 (5'-CCACCATCACCATCACCATGTGAAGCTGGCTGGT AACTCCTCC); reverse primers 1 (5'-TCGAGTTACTTGTTCGATGGTGAAGGGCAGCTC) and 2 (5'-GTTACTTGTTCGATGGTGAAGGGCAGCTC). In the case of H1N1, the sticky-end polymerase chain reaction (PCR) with *EcoRI/PstI* restriction site was used to clone into the baculovirus transfer vector, pAcGP67-A. The primers used in the sticky-end PCR were designed as follows: forward primers 1 (5'-AATTC CACCATCACCATCACCATGTGATATTAACCGGCAATTCATCTC) and 2 (5'- C CACCATCACCATCACCATGTGATATTAACCGGCAATTCATCTC); reverse primers 1 (5'- TCGAGTTACTTGTCAATGGTGAACGGCAA) and 2 (5'-GTTACTTGTTCGATGGTGAAGGGCAGCTC). Both strain of NA using the same methods as below to transfer into insect cells. The recombinant NA-pAcGP67-A plasmid was amplified in *E. coli* strain, JM109, under ampicillin selection. The co-transfection of the NA-pAcGP67-A and BaculoGold linearized baculovirus DNA (BD Biosciences Pharmingen) was carried out in insect *Spodopterafrugiperda* (Sf9) cells (Invitrogen, Carlsbad, CA).

Insect Sf9 cells ( $2 \times 10^7$ /ml) were cultured in Sf-900 II medium (Invitrogen) supplemented with 5% fetal bovine serum (Sigma, St Louis, MO). After the cells had attached, the medium was removed and the cells were infected with the recombinant baculovirus at a multiplicity of infection (MOI) of 10 for 3 days at 27 °C. On day 3, culture media were collected, concentrated using 100K-limited Amicon Ultra-15 centrifugal filter



device (Millipore, Cork, Ireland), and dialyzed against buffer A (20 mM Tris-HCl, pH 7.5, 150 mM NaCl, 10 mM imidazole) containing 5 mM MgCl<sub>2</sub> and CaCl<sub>2</sub>. Subsequently, the dialyzed media were subjected to metal affinity chromatography using Ni-NTA resin (Amersham Biosciences, Tokyo, Japan). The column was washed with 10 volumes of buffer A and Ni<sup>2+</sup>-bound proteins were eluted with a linear gradient of buffer A to buffer B (20 mM Tris-HCl pH 7.5, 150 mM NaCl, and 500 mM imidazole). The fractions containing NA tetramer were pooled and dialyzed against phosphate-buffered saline (PBS), pH 7.4, to remove imidazole.

## **2.3 Enzyme Inhibition Assay and IC<sub>50</sub> Determination**

### **2.3.1 SARS 3CL<sup>pro</sup> and CVB3 3C<sup>pro</sup> Inhibition Assay**

A fluorogenic peptide substrate (Dabcyl-KTSAVLQSGFRKME-Edans) was used for assays of 3CL<sup>pro</sup> and 3C<sup>pro</sup> activities. The anti-SARS 3CL<sup>pro</sup> activity of the test compounds were performed in the solution containing 0.05 μM SARS 3CL<sup>pro</sup>, 10 μM fluorogenic substrate, and 50 μM of test compounds at 25 °C and the anti-CVB3 3C<sup>pro</sup> activity was assayed using 0.05 μM CVB3 3C<sup>pro</sup>. Enhanced fluorescence of the reactions in the buffer of 20 mM Bis-Tris at pH 7.0 was monitored at 538 nm with excitation at 355 nm using a fluorescence plate reader (Fluoroskan Ascent; ThermoLabsystems, Helsinki, Finland). The compounds which inhibited more than 50% of the protease activity at 50 μM were selected for IC<sub>50</sub> determination.

### **2.3.2 Influenza Viral Neuraminidase Inhibition Assay**

0.15 μM of neuraminidase was incubated with the 50 μM of test compounds in the 96-well plate with the reaction buffer (50 mM Tris-HCl, pH 7.2 and 1 mM CaCl<sub>2</sub>). After

incubating with the enzyme for 25 mins in room temperature, the fluorogenic substrate 2'-(4-methylumbelliferyl)- $\alpha$ -D-N-acetylneuraminic acid (MU-NANA) (Sigma, St. Louis, MO) was added to initiate the reaction (final concentration of 100 $\mu$ M). Enhanced fluorescence at 460 nm with an excitation wavelength at 365 nm was monitored by fluorescence plate reader. The compounds which inhibited more than 50% of the neuraminidase activity at 50 $\mu$ M were selected for IC<sub>50</sub> determination.

### 2.3.3 IC<sub>50</sub> Determination

The initial velocities of the inhibited reactions were plotted against the different inhibitor concentrations to obtain the IC<sub>50</sub> by fitting with the following equation:

$$A(I) = A(0) \times \{1 - [I/(I + IC_{50})]\}$$

In this equation, A(I) is the enzyme activity with inhibitor; A(0) is the enzyme activity without inhibitor; and I is the inhibitor concentration.

## 2.4 Computer Modeling

### 2.4.1 Docking Studies of SARS 3CL<sup>pro</sup> and CVB3 3C<sup>pro</sup>

The crystal structure of SARS 3CL<sup>pro</sup> in complex with a peptide inhibitor (PDB code 1UK4) was used (Yang, Yang et al. 2003). The crystal structures of CVB3 3C<sup>pro</sup> in complex with a peptide inhibitor (PDB code 2ZU3) was used (Lee, Kuo et al. 2009). Docking process was performed using an automated ligand-docking subprogram of the Discovery Studio Modeling 1.2 SBD (Accelrys Inc., San Diego, CA), with a set of parameters chosen to control the precise operation of the genetic algorithm. Docking runs were carried out using standard default settings 'grid resolution' of 5 Å, 'site opening' of 12

Å, and 'binding site' selected for defining the active site cavity. The computer modeling was performed by Mr. Hun-Ge Liu.

#### **2.4.2 Docking Studies of Viral Neuraminidase**

The crystal structure of H5N1 neuraminidase in complex with oseltamivir (PDB code 2HU0) was used. Docking process was performed using Genetic Optimisation for Ligand Docking 3.0.1 (GOLD 3.0.1, Cambridge Crystallographic Data Centre). GOLD is considered one of the most accurate docking programs (Perola, Walters et al. 2004). The protein and the ligands should be hydrogen added, computed charges and minimized energy before docking. The residues within 8 Å of the oseltamivir binding site were defined as the active site for docking. All the parameters were set to default except of the early termination was not allowed.

### **2.5 3D-QSAR Pharmacophore Generation**

#### **2.5.1 Training Set**

Pharmacophore modeling is a powerful method to build a model of the characteristics of the drugs. Catalyst is the most cited and successful collection of pharmacophore modeling tools. 18 neuraminidase inhibitors with various structures were collected from different articles (Chand, Babu et al. 1997; Matthew A. Williams 1997; Kim, Lew et al. 1998; Lew, Wu et al. 1998; Atigadda, Brouillette et al. 1999; Babu, Chand et al. 2000; Wang, Chen et al. 2001; Chand, Babu et al. 2004; Makoto Yamashita 2009). The inhibitors were selected by the following rules: (1) At least 16 compounds to assure statistical significance of the pharmacophore model; (2) activity range of the compounds should span at least 4 orders of magnitude; (3) each order of magnitude should be

represented by at least three compounds. All the structures were drawn in ChemDraw Ultra 11.0 (CambridgeSoft, PerkinElmer Inc.) and been proceed by MM2 energy minimization.

### **2.5.2 Generation of Hypogen**

Catalyst 4.10 software package was employed to construct the possible models, using Hypo as given name. Hypo I is normally the best among the generated hypothesis. Poling algorithm and CHARMM forcefield parameter were applied to the training set to generate the best conformation of each compound. A maximum number of 250 conformations of each compound were selected using ‘best conformer generation’. An energy threshold of 20 kcal/mol which is above the global energy minimum was used in order to ensure the maximum coverage of the conformational space. The values for minimum points and minimum subset points were kept as the default values of Catalyst, which are both 4. The uncertainty value, which represents the ratio of the uncertainty range of actual biological activity against the measured biological activity, was set to 3. Five features, that are hydrogen bond acceptor (HBA), hydrogen bond donor (HBD), hydrophobic (HD), positive ionization (PI) and negative ionization (NI), were selected for the hypotheses generation as they could characterize all the functional groups of all compounds based on the characters of each compound.

### **2.5.3 Cost Analysis**

Once the hypotheses (Hypo) were generated, an output logs file with several useful parameters information to help assessing the validity of the hypotheses was generated too (Debnath 2002). The parameters are listed as below:

1. Fixed cost: This value represents the simplest model that still fits all the data perfectly. It could also be regarded as a cost of a perfect hypothesis with no deviation between predicted and actual experimental activities.

$$\text{Fixed cost} = \text{Error cost (x=0)} + \text{Weigh cost (x=0)} + \text{Configuration cost}$$

2. Null cost: This value presumes that there is no statically significant structure in the data and which estimates activity to be the average of the activity data of the training set molecules.

$$\text{Null cost} = \text{Error cost (x=x)}$$

3. Total cost consists of weight cost, error cost and configuration cost.

$$\text{Total cost} = \text{Weight cost} + \text{Error cost} + \text{Configuration cost}$$

4. Errors cost is a value correspondent to the root-mean-square (RMS) deviation between the estimated and actual biological activity of the training set.
5. Configuration cost is equal to the entropy which depends on the complexity of the hypothesis. A good hypothesis should have configuration cost with value of smaller than 17 bits.

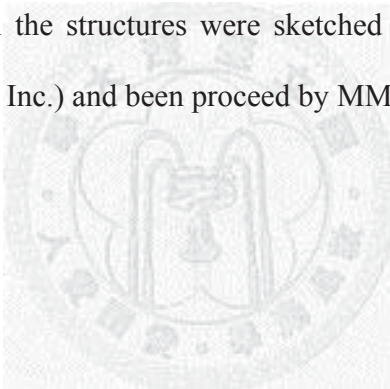
A generated Hypo with a fixed cost that is substantially below that of the null cost is likely to be statistically significant and bears visual inspection. The greater the difference between the fixed cost and null cost, the likely the generated hypothesis reflects true correlation rather than by chance (Zhang, Yu et al. 2006). The difference between fixed and null cost should be  $\geq 70$  bits to show over 90% statistical significance of the generated Hypo. The cost difference should be greater than 60 bits to represent a true correlation data.

#### **2.5.4 CatScramble**

A program, CatScramble which is embedded in the Discovery Studio was applied to further evaluate the statistical relevance of the model. This program which is based on Fischer's randomization was used to validate the relations between structures and biological activities of training set using retraining the structures of the training set with random activities (Du, Li et al. 2005; Bharatham, Bharatham et al. 2007). To achieve 95%, 98% and 99% confidence level, 19, 49 and 99 random Hypo have to be generated respectively.

### **2.5.5 Test Set**

93 compounds with different biological activity were selected to validate the accuracy of the generated Hypo I. All the structures were sketched in the ChemDraw Ultra 11.0 (CambridgeSoft, PerkinElmer Inc.) and been proceed by MM2 energy minimization (Zhang, Yu et al. 2006).



## Chapter 3 RESULTS

### 3.1 Expression and Purification of SARS 3CL<sup>pro</sup>

The SARS 3CL<sup>pro</sup> was expressed by transformation of the engineered plasmid into *E.coli* BL21. Ni-NTA column was employed for the His-tagged SARS 3CL<sup>pro</sup> purification. After the His-tag was cleaved by incubation with FXA for 3 days, the untagged SARS 3CL<sup>pro</sup> was loaded onto NI-NTA column again to remove the unwanted His-tag. As shown in the reducing SDS-PAGE (Figure 1), the flowthrough of the second Ni-NTA column contained highly pure untagged protease. The SARS 3CL<sup>pro</sup> was then rapid freeze by liquid nitrogen and kept in -80°C refrigerator for long-term storage.

### 3.2 Inhibition of Pyrimidine Compounds Against SARS 3CL<sup>pro</sup> and Their Binding Modes.

A series of pyrimidine analogs (6a-n) have been synthesized by Dr. Ramajayam Ramachandran (Figure 2a). The compounds which showed more than 50% inhibition at 50 $\mu$ M were chosen for further IC<sub>50</sub> determination. Compound 6m with R<sup>2</sup> group of nitro functionality at C-4 position is the most potent inhibitor with an enzyme inhibitory activity against SARS 3CL<sup>pro</sup> with an IC<sub>50</sub> of 6.1 $\mu$ M. The structures and IC<sub>50</sub> values are given in Table 1. The cytotoxicity of the test compounds was tested by performing the MTT assay and found that all compounds are devoid of cytotoxicity. Computer modeling was performed by Mr. Hun-Ge Liu in order to find out the possible binding modes of the inhibitors. Compound 6m has the lowest IC<sub>50</sub> among all and hence was used as model for docking in search of possible binding modes. The distance between the NH of the



pyrimidine ring and oxygen atom of Glu-166 was constrained in the distance of 2.0 Å. The orientation of the ligand has the nitro phenyl group situated in the S1 pocket, with the nitro group pointing towards the surface of the protein. One of the oxygen of the nitro group is in close proximity 2.3 Å to the Gly-143 and the other oxygen atom is forming hydrogen bond with Cys-145 at the distance of 2.1 Å (Figure 2b). The chlorophenyl ring fits into the S2 pocket and having hydrophobic interactions with Met-49 and Gln-189. The results of the docking studies presented here suggest that the nitro phenyl group of 6m can potentially occupy the active site cysteine residue in the enzyme. The oxygen of the nitro group forms a hydrogen bond with the side chain of Gly-143 and Cys-145 that was important for inhibition activity. The compounds lacking nitro functionality in the aryl ring lost the activity. Moderate electron withdrawing substituent R<sup>1</sup> like chloro in the compounds 6l and 6m favors the inhibitory activity when compared to the electron donating groups like methyl and methoxy in the compound (Table 1). This result suggests that the substituent R<sup>1</sup> can be electron withdrawing group to increase the inhibitory action.

### **3.3 Inhibition of Pyrazolone Compounds Against SARS 3CL<sup>Pro</sup> and Their binding modes.**

A series of pyrimidine analogs (2a-u) have been synthesized by Dr. Ramajayam Ramachandran (Figure 3a). The compounds which showed more than 50% inhibition at 50µM were chosen for further IC<sub>50</sub> determination. Compound 2p is the most potent inhibitor showing an IC<sub>50</sub> of 5.5 µM and 2t is the second with IC<sub>50</sub> of 6.8 µM against SARS 3CL<sup>Pro</sup>. The cytotoxicity of the compounds was tested by performing the MTT assay and found that all compounds are devoid of cytotoxicity at 200 µM. The structures and IC<sub>50</sub>

values are given in Table 2. In search of a computer model of the associated complex between the compound 2u and the proteases to rationalize its inhibitory activity, the orientation of the ligand has the N1-phenyl group situated in the S1' pocket of the 3CL<sup>pro</sup> (Figure 3b). One of the oxygen of the nitro group is in close proximity 2.7 Å and forms H-bond to the Gly-143. The C=O in the central pyrazolone ring is close to Glu-166 with the distance of 3.0 Å to form a H-bond. C-3 phenyl ring fits into the S2 pocket, having hydrophobic interactions with Met-49, Arg-188, and Gln-189 (hiding behind these residues in Figure 3b). The carboxyl benzylidene group is situated in the S3 pocket of the 3CL<sup>pro</sup>. The oxygen of the carboxyl group forms a hydrogen bond with the side chain of Gln-192 at a distance of 3.2 Å. It is important for inhibition activity since the compounds lacking carboxy functionality in the benzylidene lost the activity. Electron withdrawing R<sup>1</sup> substituents like cyano (2p), fluoro (2t), and nitro (2u), accompanied with R<sup>2</sup> carboxyl group favors the inhibitory activity.

Another series of pyrazolone analogs have been synthesized by Mr. Vathan Kumar (Figure 4a). Those of the compounds are initially synthesized for the purpose to be the neuraminidase inhibitors, but surprisingly, some of them could even inhibit the SARS 3CL<sup>pro</sup> with the IC<sub>50</sub> values around 3 μM. Four of the compounds show inhibition to the SARS 3CL<sup>pro</sup>. VK93 was the best against SARS 3CL<sup>pro</sup> with an IC<sub>50</sub> of 3.85 μM and VK99 is the second lowest with an IC<sub>50</sub> of 6.04 μM (Table 3). VK93 has a different binding mode comparing to 2u (Figure 4b). Replacing β-carboxybenzoic acid substitution by VK93 at the C4 position of pyrazolone resulted in re-orientation of the pyrazolone molecule, then making it unable to fit to S3 pocket of the enzyme. This could be attributed to bulkiness introduced by the molecule which makes it unable to fit to the S3 pocket. The 4-

chlorobenzoic acid at the C4 fits into the S1 pocket. The C=O in the central pyrazolone ring is close to Glu-166 with the distance of 1.9 Å to form a H-bond. The oxygen on the furan ring is in close proximity 2.8 Å and forms H-bond to the Glu-166 too. The N1 and C3 diphenyl ring fit into the S3 and S2 pocket respectively. The N1-phenyl of the VK93 occupies the S3 rather than the S1' like the N1-phenyl group of 2u does.

### 3.4 Expression and Purification of CVB3 3C<sup>pro</sup>

The CVB3 3C<sup>pro</sup> was expressed by transformation of the engineered plasmid into *E. coli* BL21. The cell lysates have been centrifuged and the cell debris was discarded. The supernatant was loaded onto the Ni-NTA column, and followed by the washing of the 5mM imidazole and 30mM imidazole respectively. 300 mM of imidazole was used to elute the pure CVB3 3C<sup>pro</sup> (Figure 5a) and then the enzyme was dialysis against 2x2L of the dialysis buffer in order to remove the imidazole. Further removal of the His-tag is not required as we have proven that the His-tag at the N-terminal does not affect the activity of the CVB3 3C<sup>pro</sup> (Lee, Kuo et al. 2009). The CVB3 3C<sup>pro</sup> was then rapid frozen by liquid nitrogen and kept in -80°C refrigerator for long-term storage.

### 3.5 Inhibition of Pyrazolone Compounds Against CVB3 3C<sup>pro</sup> and Their Binding Modes

In further evaluating the inhibitors against CVB3 3C<sup>pro</sup>, as shown in Table 2, we found 2p and 2t were moderate inhibitors against CVB3 3C<sup>pro</sup> (IC<sub>50</sub> = 20.8 and 22.4µM, respectively), but 2u was more active against CVB3 3C<sup>pro</sup> (IC<sub>50</sub> = 9.6 µM). According to

the modeling (Figure 5b) , the R1 nitro group of 2u forms H-bond with Gly-145 (2p and 2t without nitro group fail to form such a H-bond) and benzylidene carboxylate of 2u is H-bonded to Glu-71 in the active site of 3C<sup>pro</sup>. It was predicted that the C-3 phenyl ring of 2u is pointed to S1 site and the carboxyl benzylidene group is relocated to S2 in order to form the H-bond in 3C<sup>pro</sup> due to the subtle differences between the structures of 3CL<sup>pro</sup> and 3C<sup>pro</sup> (Lee, Kuo et al. 2009). However, it should be noted that computer modeling is speculation based on energy minimization to fit the SAR data. In conclusion, 2p and 2t are selective against 3CL<sup>pro</sup>, but 2u is a common inhibitor of 3CL<sup>pro</sup> and 3C<sup>pro</sup>, which may be potentially developed into anti-coronaviral and anti-picornaviral drugs.

Furthermore, it is really surprising that some of the pyrazolone compounds which initially synthesized to test against the inhibition to the neuraminidase also showing good inhibition to CVB3 3C<sup>pro</sup> with the IC<sub>50</sub> as low as 8.51  $\mu$ M (Supplementary table 1)

### **3.6 Expression and Purification of H1N1 and H5N1 Neuraminidase**

The ectodomain (63-449) of H1N1 and H5N1 NA protein that contains the enzyme activity were expressed as an N-terminal His-tagged protein and purified from the culture media using Ni-NTA column. The purified NA protein appeared on SDS-PAGE with molecular weight of ~43 kDa, and only the H5N1 NA was specifically recognized by the H5N1 NA antibody (Figure 6a). This enzyme was active in using the fluorogenic MU-NANA as substrate with expected kinetics (data not shown). Titration of the H1N1 and H5N1 NA with the oseltamivir giving the IC<sub>50</sub> values of 21.5 and 8.9 nM respectively (Figure 6b). These results gave us the confidence about the purity and the activity of the purified neuraminidase.

### 3.7 Inhibition of Pyrazolone Compounds Against the H1N1 and H5N1 NA

In order to explore the potent inhibitors with good oral bioavailability, high throughput screenings have been performed by Mr. Hun-Gen Liu. One of the compounds which was given name as VK84 showing an  $IC_{50}$  of 16.6  $\mu$ M against the H5N1 neuraminidase (Table 4). According to the docking result, the carboxylate of the 2-furanyl-4-(5-chloro-2-benzoic acid) on the  $R^3$  substitution plays a very crucial role in inhibiting the enzyme (Figure 7a). It acts as a H-bond acceptor to form three H-bonds with the residues nearby in the active site. Arg371, Arg292 and Tyr347 form H-bonds with the carboxylate and the distances are 3.5 Å, 2.6 Å, and 2.9 Å respectively. Furthermore, Arg152 also forms one H-bond with the oxygen on the furan ring of the  $R^3$  substituent with the distance of 2.7 Å. The 3-carboxylate of the  $R^1$  substituent also forms two H-bonds with the Glu119 and Arg156 in the distances of 2.8 Å and 2.6 Å respectively. The docking results also show the methyl group of the  $R^2$  substituent is orientating toward the hydrophobic pocket of the neuraminidase like oseltamivir does. It is really surprised that the superimposition of the oseltamivir and VK84 on the active site gives a similar binding mode.

In order to increase the hydrophobic interaction between the enzyme-inhibitor complex and also mimic the ionic interaction provided by the C4-amino of the oseltamivir, substituents like phenyl ring and amino group have been introduced into the original hit. A series of analogs were then synthesized in order to improve the inhibition against the neuraminidase (Table 4). In search of the possible binding mode, docking had also been performed. One of the compounds, VK94 inhibits the H1N1 neuraminidase with an  $IC_{50}$  of 1.7  $\mu$ M. The docking results show that VK94 basically inhibits the neuraminidase with the

binding mode that we had predicted (Figure 7b). The carboxylate of 2-furanyl-4-(5-chloro-2-benzoic acid) of the R<sup>3</sup> substituent acts as an H-bond acceptor to form five H-bonds with Tyr406, Arg371, Arg292 and Tyr347. Two H-bonds between the carboxylate and Arg292 are in distances of 3.4 Å and 2.8 Å. The H-bonds between the carboxylate and Tyr406, Arg371 and Tyr347 are 2.2 Å, 2.2 Å, and 3.0 Å respectively. The 2-amino of the R<sup>1</sup> substituent acts as a H-bond donor to form two H-bonds with Ser179 and Glu227. This amino group is in close proximity 2.4 Å and 1.3 Å respectively. These H-bonds are believed greatly contribute to the inhibition of the enzyme as another compound VK92, which is only different in the R<sup>1</sup> substitution (2-COOH instead of 2-NH<sub>2</sub>), has a dramatically increased IC<sub>50</sub> of 13.3 μM. It is proven by the docking that the carboxylate of the R<sup>1</sup> substituent of VK92 orients toward the Arg156 and act as a H-bond acceptor to form two H-bonds in the distances of 2.4 Å and 3.4 Å (Figure 7c). The reorientation of the R<sup>1</sup> substituent of VK92 also causes the loss of a H-bond between N1 nitrogen of the pyrazolone ring and the Arg152. R<sup>2</sup> substituent, a phenyl ring of both VK92 and VK94 fit into hydrophobic pocket of the enzyme. VK93, another inhibitor inhibits the H1N1 neuraminidase with an IC<sub>50</sub> of 3.8 μM. The binding mode according to the docking shows that the 2-furanyl-4-(5-chloro-2-benzoic acid) of the R<sup>3</sup> substituent binds to the residues similar to the VK92 and VK94, but the phenyl ring at the C3 position of the pyrazolone ring provides extra hydrophobic interaction thus the some of the H-bond are missing. The differences of the IC<sub>50</sub> values between the H1N1 and H5N1 NA are might be caused by the differences of these NA structure though they are conserved within the active sites. But it is interesting that the order of the inhibition activity of molecules is basically consistent between H1N1 and H5N1 NA.



### 3.8 Generation of the Catalyst Hypotheses

The features for the generation of the catalyst hypothesis are selected by the analysis of the functional groups of the training set (Figure 8a) (Chand, Babu et al. 1997; Matthew A. Williams 1997; Kim, Lew et al. 1998; Lew, Wu et al. 1998; Atigadda, Brouillette et al. 1999; Babu, Chand et al. 2000; Wang, Chen et al. 2001; Chand, Babu et al. 2004; Makoto Yamashita 2009). The selected features for the generation of the hypotheses are H-bond donor, H-bond acceptor, hydrophobic, positive ionization and negative ionization. 10 hypotheses have been generated and the cost analysis shows that all of the hypotheses are basically reliable (Table 5). The Null cost is 182.92 and the total fixed cost is 75.32, so that the difference between these costs is 107.60 which shows the 95% confidence level. Ten hypotheses have been generated and the Hypo1 is considered the most predictive model. Hypo1 is characterized by one hydrogen-bond donor feature (HD), one hydrogen-bond acceptor feature (HA), one hydrophobic feature (HY), one negative ionization feature (NI) and one positive ionization feature (PI) as shown in Figure 8b. The Hypo1 gives a RMS as low as 1.45. Then, the cost range between Hypo1 and the fixed cost is 7.22, while that between the Null hypothesis and Hypo1 is 82.54. Noticeably, the total cost of Hypo1 is much closer to the fixed cost than to the Null cost. Furthermore, the Hypo1 is also highly correlated to the training set with a correlation value of 0.92. The configuration cost of the Hypo1 is smaller than 17, that is 13.65. Cost analysis demonstrates that a reliable and highly predictive pharmacophore has been generated.

### 3.9 Validation of the Hypo1



All the molecules of the training set are divided into three groups by their inhibition activity. The inhibitors with the  $IC_{50}$  between lower than 100 nM are considered highly active, and the inhibitors with the  $IC_{50}$  values in the range from 100 nM to 10000 nM are considered moderately active, and the remaining inhibitors are considered inactive. As shown in the Table 6, 14 out of 18 inhibitors are classified correctly. Interestingly, the other 4 inhibitors are both underestimated by only one order of magnitude. Two marketed drugs, zanamivir and oseltamivir and one drugs which currently under phase III clinical trials, peramivir and the first neuraminidase inhibitor, DANA have also been predicted precisely (Figure 9). The superimpositions of the Hypo1 onto the enzyme active sites also show that the residues nearby have a possibility to interact with each feature (Supplementary figure 1).

With the aid of the CatScramble program, the experimental activities of molecules in the training set were scrambled randomly with the aid of the CatScramble program and the resulting training set was used for a hypotheses generation. This procedure was reiterated 49 times and none of the hypotheses show a better hypothesis fixed cost than the Hypo1 (Graph 1). 93 compounds with experimental activities have been used as the test set to validate the Hypo1 (Supplementary figure 2 and Supplementary table 2) (Zhang, Yu et al. 2006). The graph plotted by the actual  $IC_{50}$  versus estimated  $IC_{50}$  gives a correlation of 0.76 (Graph 2) and it is only slightly lower than the previously reported pharmacophore (Zhang, Yu et al. 2006). All the information above gives us the confidence about the reliability to the Hypo 1.

## 4. DISCUSSIONS

### 4.1 SARS-CoV 3CL<sup>pro</sup> and its Inhibitors

It is been 8 years since the last reported SARS, the impacts and the threats of this disease is almost been forgotten by the public. Some even think that the threats and the severities of SARS might even be overblown. The loss of the vigilance to the SARS of the public would somehow makes the situation even worse if SARS outbreak really happen again. The recent finding also reports the isolation of SARS-CoV from animals including bats leads to a concern that would the SARS outbreak happen again (Guan, Zheng et al. 2003; Lau, Woo et al. 2005; Li, Shi et al. 2005). No any vaccines and drugs are currently available so far are more likely reflects the severity of SARS reemergence.

One of the nonpeptidyl inhibitors our lab discovered, VK93 is a very potent inhibitor as it inhibits the SARS 3CL<sup>pro</sup> with an IC<sub>50</sub> of 3.85 $\mu$ M. It is possible that the inhibitor could inhibit the protease even better after some minor modifications. An H-bond donor or H-bond acceptor could be introduced to the para position of the phenyl ring at the N1 as it might provide extra H-bonds with the Glu192 as shown in the docking. Furthermore, the furan ring could also be modified like introduction of hydroxyl group at the C2 or C3 position to form interactions with the C145, an important catalytic amino acid of the SARS 3CL<sup>pro</sup>. Furthermore, pyrazolone compounds synthesized by Mr. Vathan Kumar could also be the inhibitors the of CVB3 3C<sup>pro</sup> because pyrazolone compounds have been proven potentially act as common inhibitors for both 3C and 3CL<sup>pro</sup> (Kuo, Liu et al. 2009).

### 4.2 Neuraminidase and its Inhibitors.

There is no doubt for the urgency of developments of the preventions and the treatments for influenza virus infections (flu) as it could cause the epidemics and it reemerged as seasonal flu every year. As part of our lab's effort to develop the drugs against the influenza viruses, VK94, a neuraminidase inhibitor with an  $IC_{50}$  of  $1.7\mu M$  has shown some unexpected H-bonds interaction with the neuraminidase. The initial purposes of the introduction of the amine substituent to the  $R^1$  of the inhibitor was forming an ionic interaction with the Glu119, but the docking results coming out with the ionic interaction with the Glu227 instead of Glu119 (von Itzstein, Wu et al. 1993). It is possible that the amine substituent could be replaced by the guanidine group as the ionic interaction could be even strong after the modification (Kim, Lew et al. 1997). Furthermore, the modification of the phenyl ring substituent located at the  $R^2$  position to the other hydrophobic group might also be preferable as it might greatly increase the hydrophobic interactions (Kim, Lew et al. 1997). The docking score between the VK94 and the modified VK94 did also confirm the hypothesis (data not shown). As a conclusion, the potency of VK94 to be modified has made it remains a very attractive target to be explored as the pyrazolone compounds have never been reported as neuraminidase inhibitors.

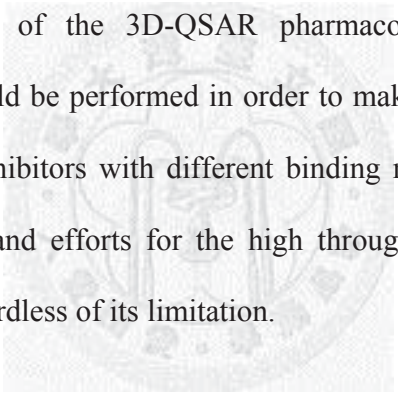
It is very interesting that some of the inhibitors could even act as common inhibitors for both SARS 3CL<sup>pro</sup> and Influenza virus neuraminidase like VK93 and VK99. These findings have also pointed out the potency of the pyrazolone compounds.

#### **4.3 Hypo1 and its Modification**

With the aid of the pharmacophore, the accuracies of the high throughput screenings could be increased but one of the limitations is that it could also limit the exploration of

the new inhibitors with the different functional groups, charges and the sizes. It is because 3D-QSAR pharmacophore was built based on the SAR of a series of known inhibitors.

The inhibitors of our lab has discovered so far were also been evaluated by the Hypo1 (Supplementary table 3). Almost all of the compounds were predicted accurately except of the VK94 and VK99. As shown in supplementary figure 1, most of the most potent inhibitors used for the generation of Hypo1 do not interact with the Glu227 and Ser179. Meanwhile, the amine substituent and the nitrate substituent located at the R<sup>1</sup> interact with the do interact with Glu227 and Ser119 according to the docking. This might explain why the Hypo1 has failed to predict the IC<sub>50</sub> of both VK94 and VK99 and also pointed out the limitations of the 3D-QSAR pharmacophore. So that the further modifications of Hypo1 should be performed in order to make it potentially predictive for all kind of neuraminidase inhibitors with different binding modes. In conclusion, Hypo1 could save a lot of money and efforts for the high throughput screening of the potent neuraminidase inhibitors regardless of its limitation.



## FIGURES

Figure 1

The SDS-PAGE of the purification of the SARS-CoV 3CL<sup>pro</sup> during purification. Lane 1 is shows the markers. Lane 6 is the eluted SARS 3CL<sup>pro</sup> by 300 mM imidazole. Lane 8 is the purified proteases after purification of the second Ni-NTA column.

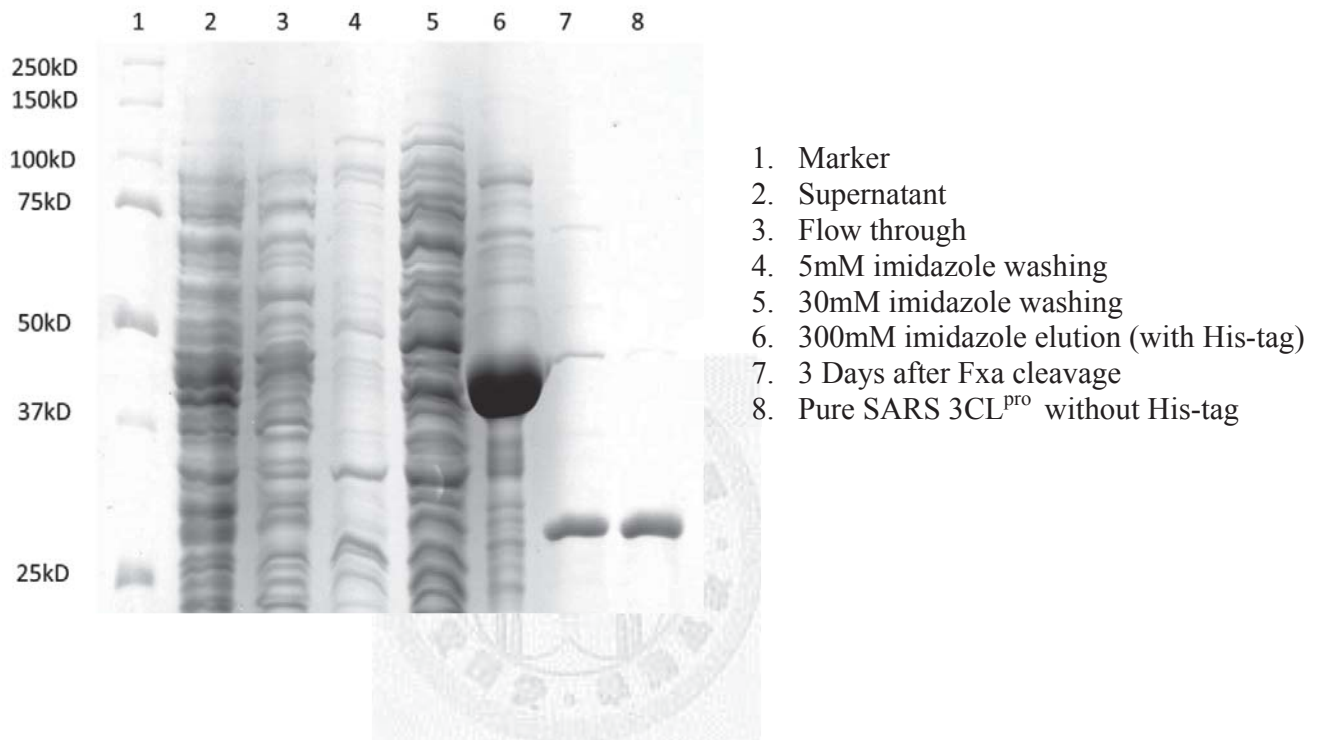


Figure 2a

14 pyrimidine analogs have been synthesized using the synthetic schemes below. The synthesis of the compounds were was performed by the Dr. Ramajayam Ramachandran in our laboratory.

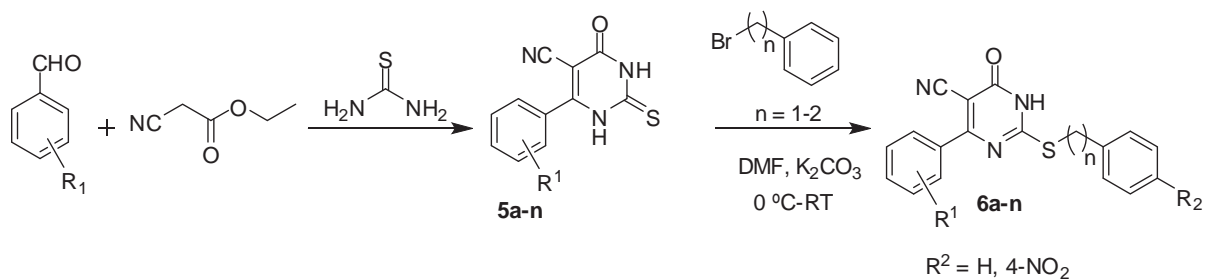


Figure 2b

The compound 6n which inhibits the protease with an IC<sub>50</sub> of 6.1 μM was docked into the protease (PDB code 1UK4) active site by Mr. Hun-Ge Liu.

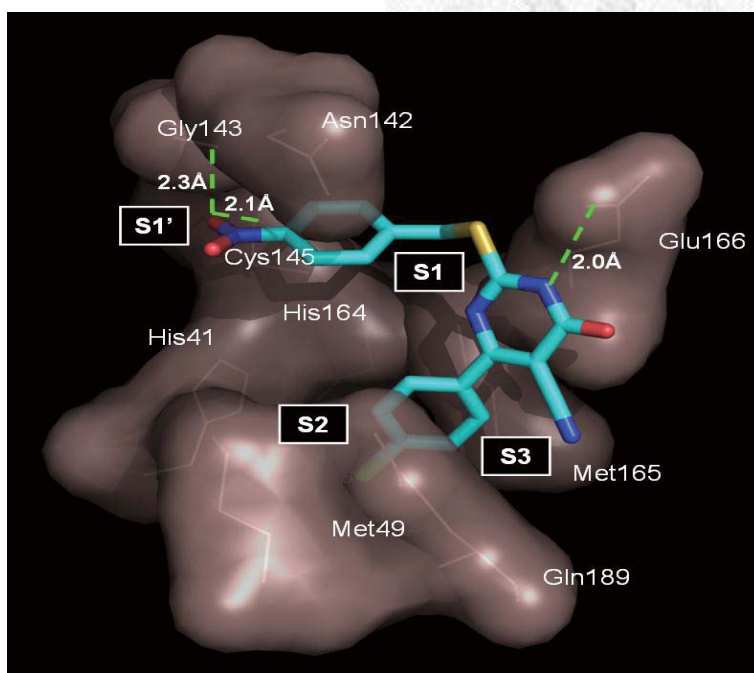


Figure 3a

21 pyrazolone analogs have been synthesized using the synthetic schemes below. The synthesis of the compounds were was performed by the Dr. Ramajayam Ramachandran.

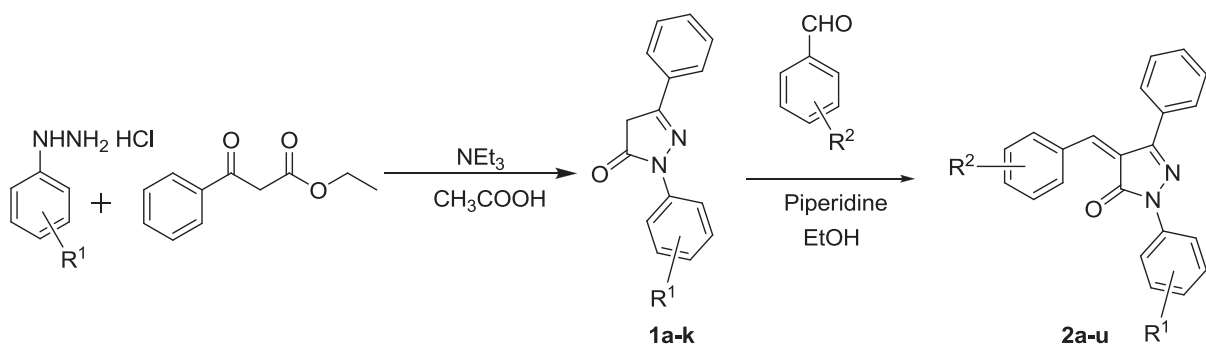


Figure 3b

The compound 2u which inhibits the protease with an IC<sub>50</sub> of 8.4μM was docked into the protease (PDB code 1UK4) active site by Mr. Hun-Ge Liu.

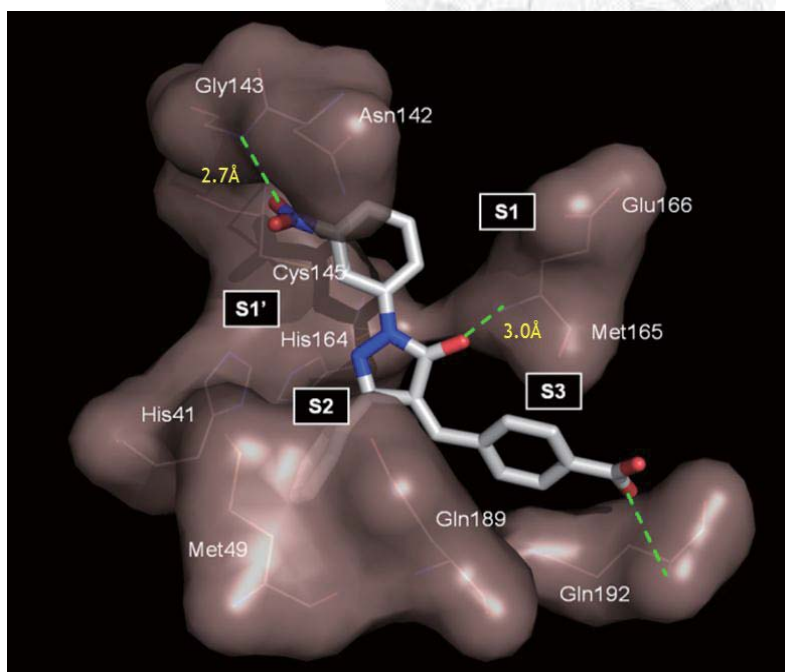




Figure 4a

14 pyrazolone analogs have been synthesized by Mr. Vathan Kumar using the synthetic scheme as shown below.

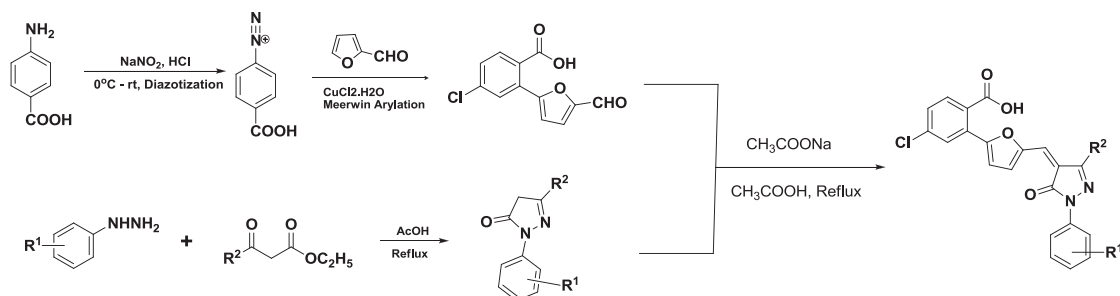


Figure 4b

VK93 which inhibits the SARS 3CL<sup>pro</sup> with an  $\text{IC}_{50}$  of  $3.85\mu\text{M}$  was docked into the SARS 3CL<sup>pro</sup> (PDB code 2AMD). The docking result shows that VK93 inhibits the protease using a mechanism different from the 2u.

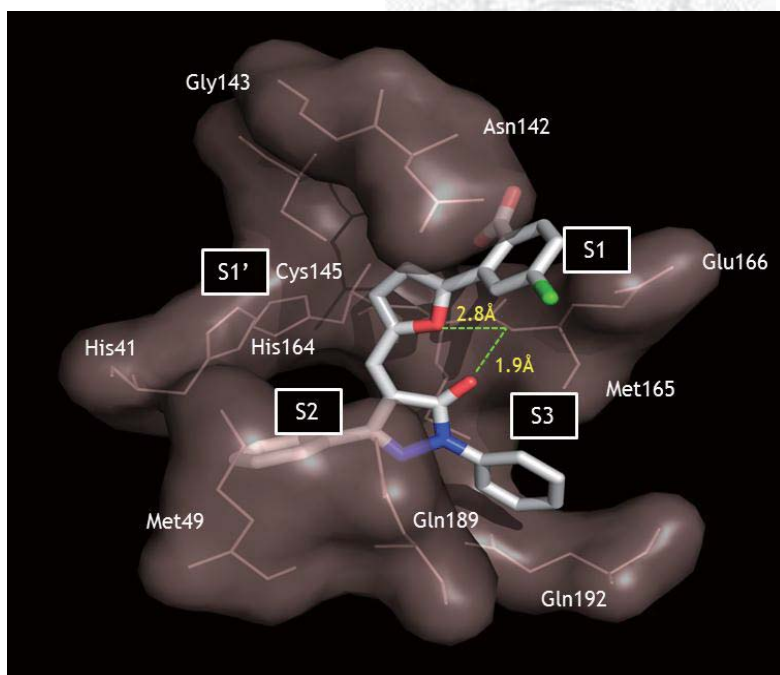


Figure 5a

The SDS-PAGE of the purification of the CVB3 3C<sup>pro</sup> during purification. Lane 1 is shows the markers and the lane 4 is the purified protease. The results show that the protease is not highly pure but it is good enough for the inhibition assay.

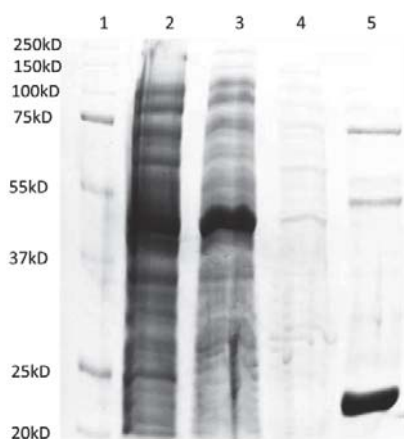


Figure 5b

2u which may act as a common inhibitor for both SARS 3CL<sup>pro</sup> and CVB3 3C<sup>pro</sup> has been docked into the CVB3 3C<sup>pro</sup> (PDB code 2ZU3). 2u inhibits the 3C<sup>pro</sup> with an IC<sub>50</sub> value of 9.6μM.

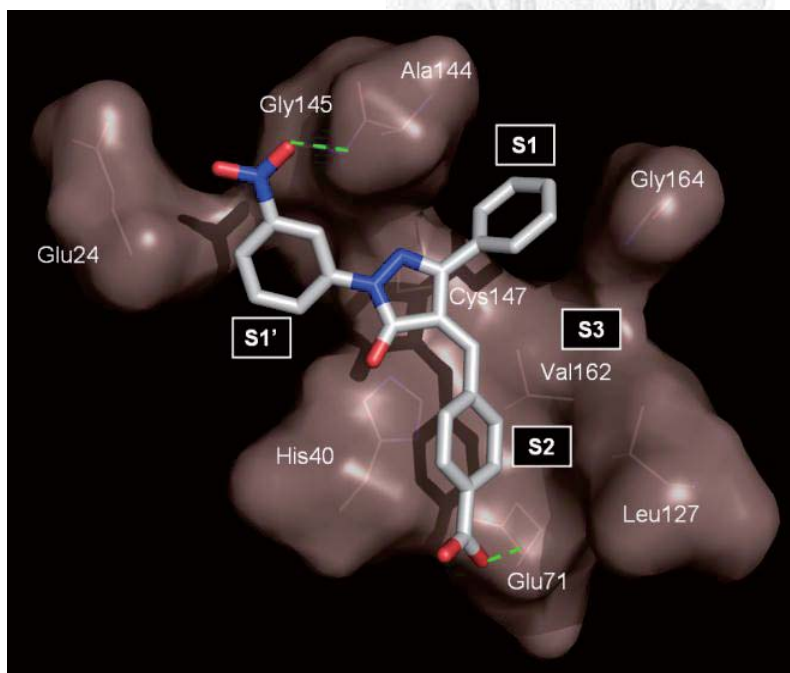


Figure 6a

The SDS-PAGE of the H1N1 NA and the H5N1 NA is on the left panel and the right panel showing shows that only the H5N1 NA can be recognized by the H5N1 antibody

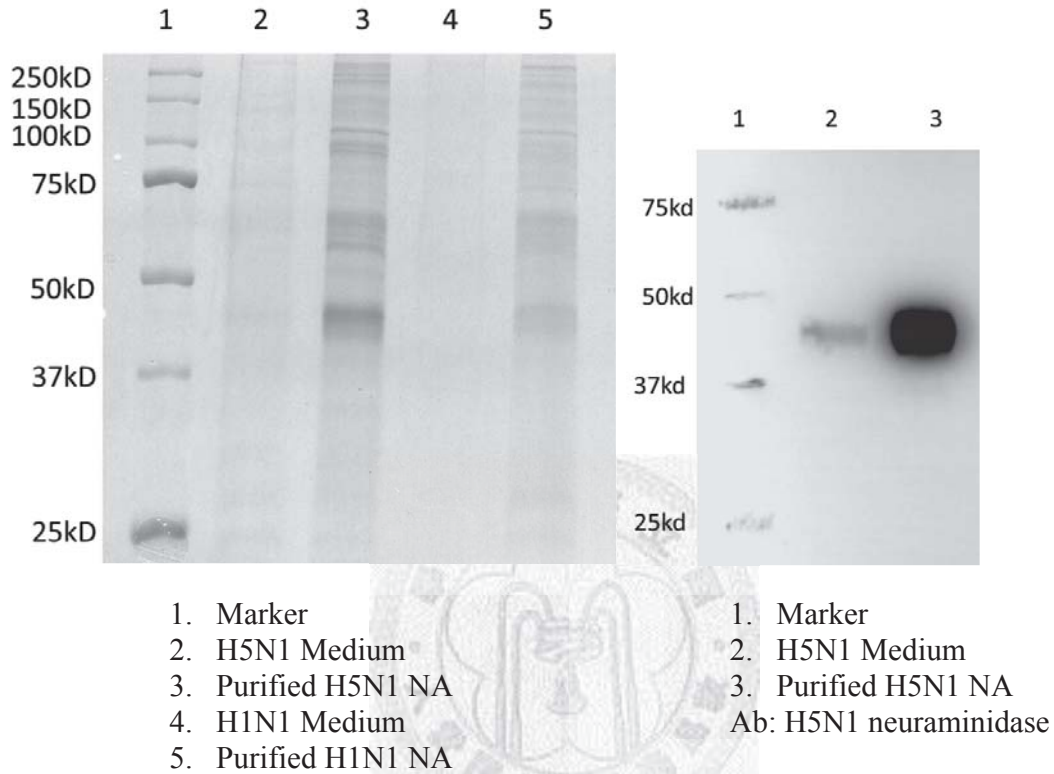


Figure 6b

The titration of the H1N1 NA and H5N1 NA by the oseltamivir.

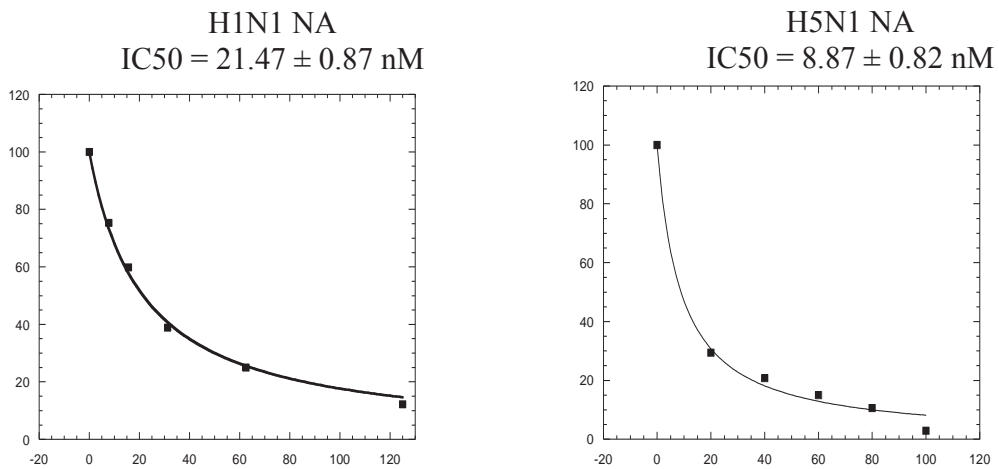


Figure 7a  
The superimposition of the docked VK84 and oseltamivir in the NA active site (PDB code 2HU0).

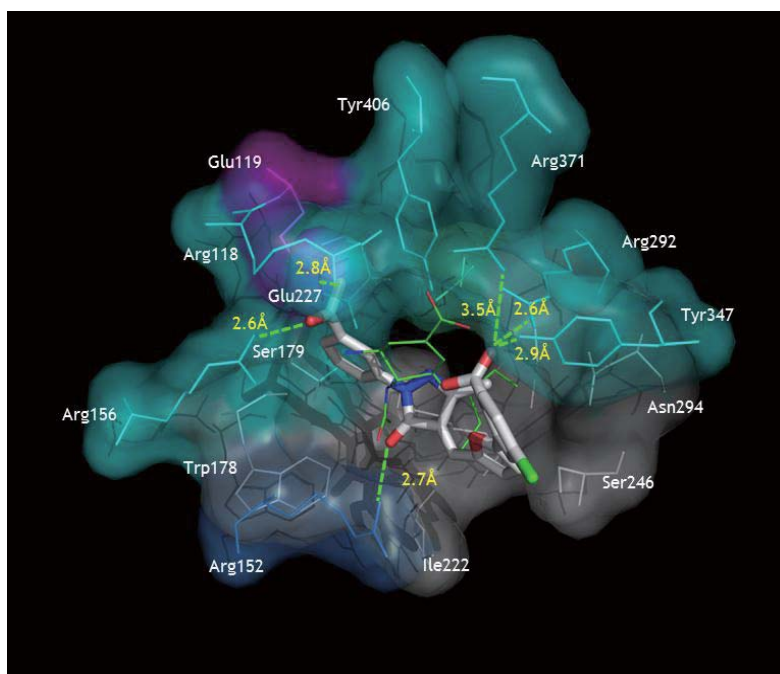


Figure 7b  
The docked poses of the VK94, an inhibitor with an  $IC_{50}$  of 1.7 $\mu$ M.

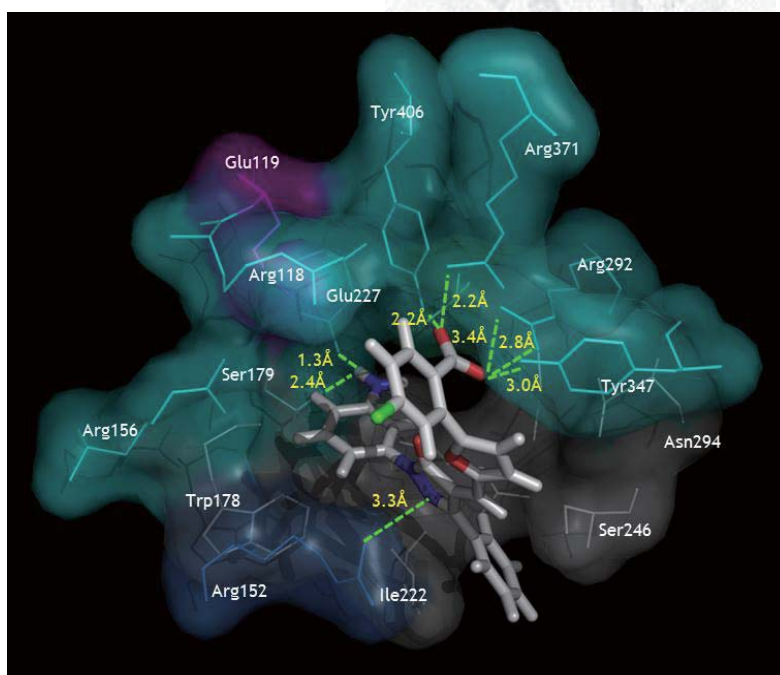


Figure 7c

The docked poses of the VK92, an inhibitor with an  $IC_{50}$  of  $13.3\mu M$

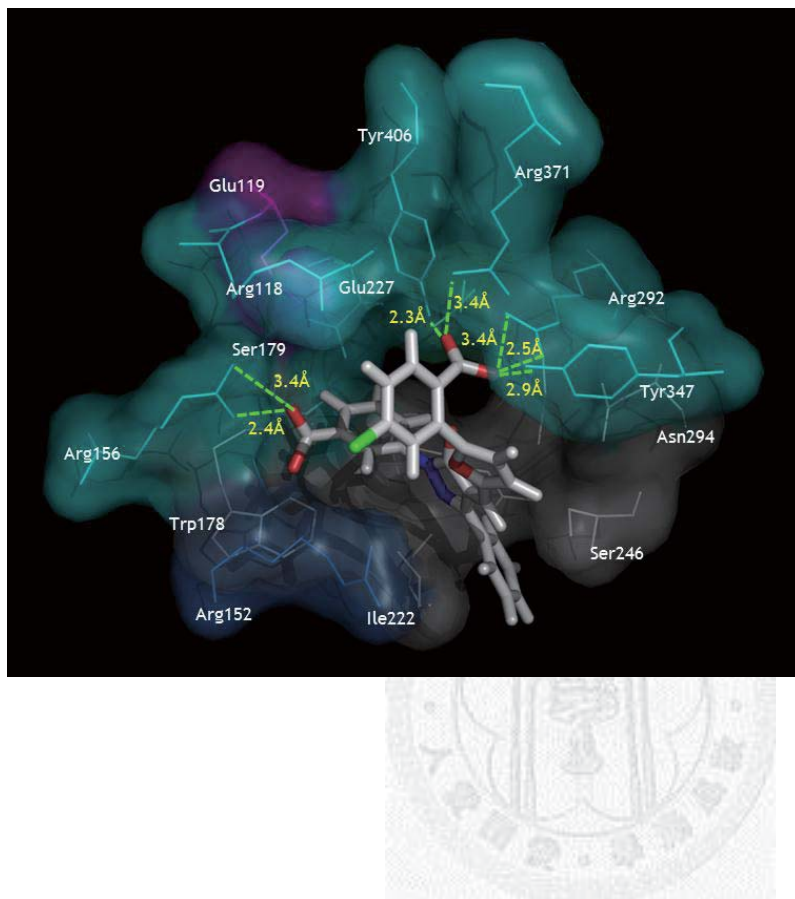


Figure 8a

The molecules that have been used as the training set for the HypoGen. Two marketed neuraminidase inhibitors, zanamivir and oseltamivir and another neuraminidase inhibitors under phase 3 clinical trials, laninamivir and peramivir are also included.

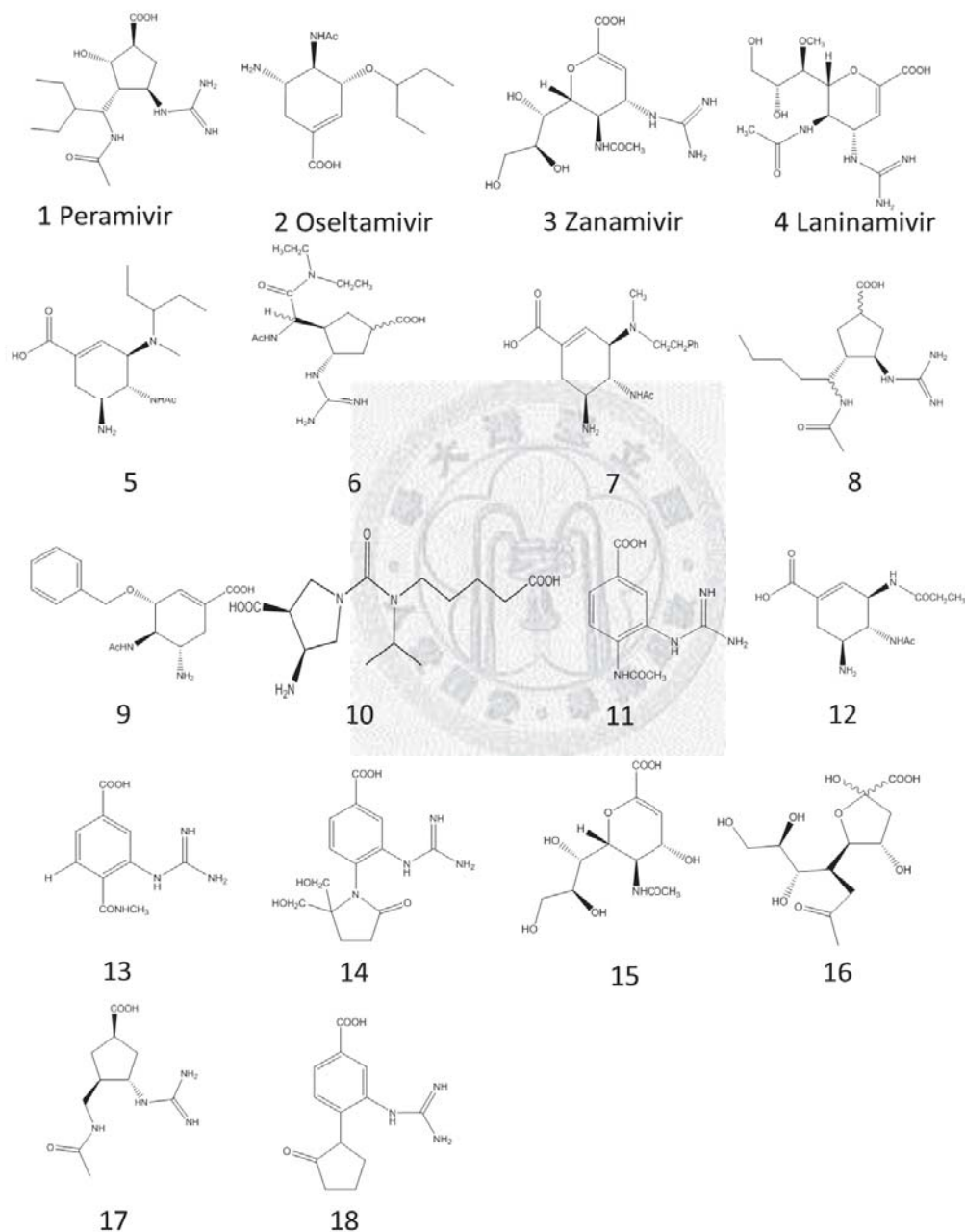


Figure 8b

Hypo1 is characterized by five features, there are HA (green), HD (violet), HY (cyan), NI (blue) and PI (red).

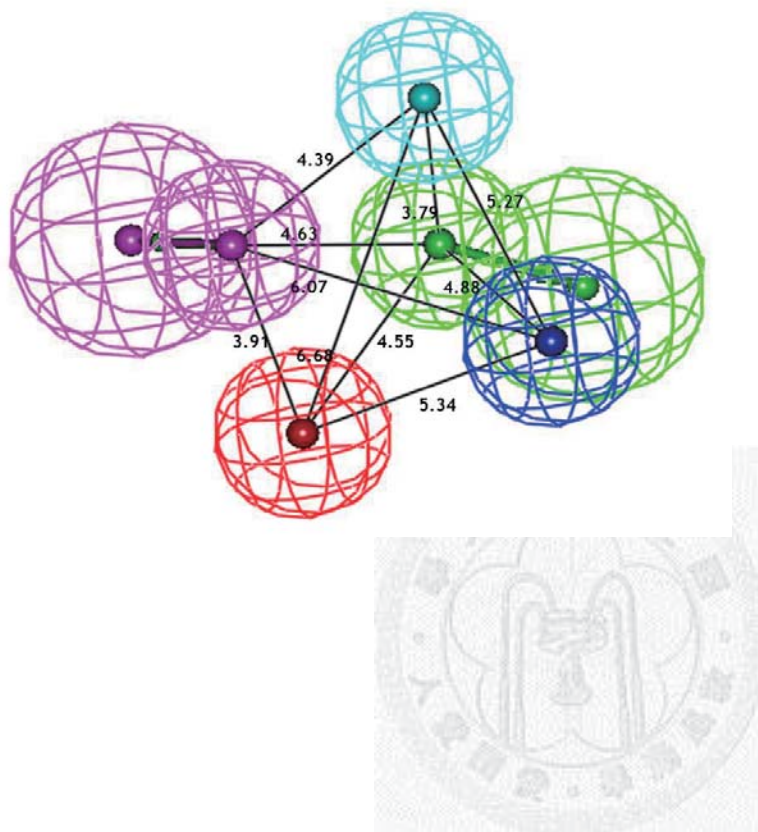
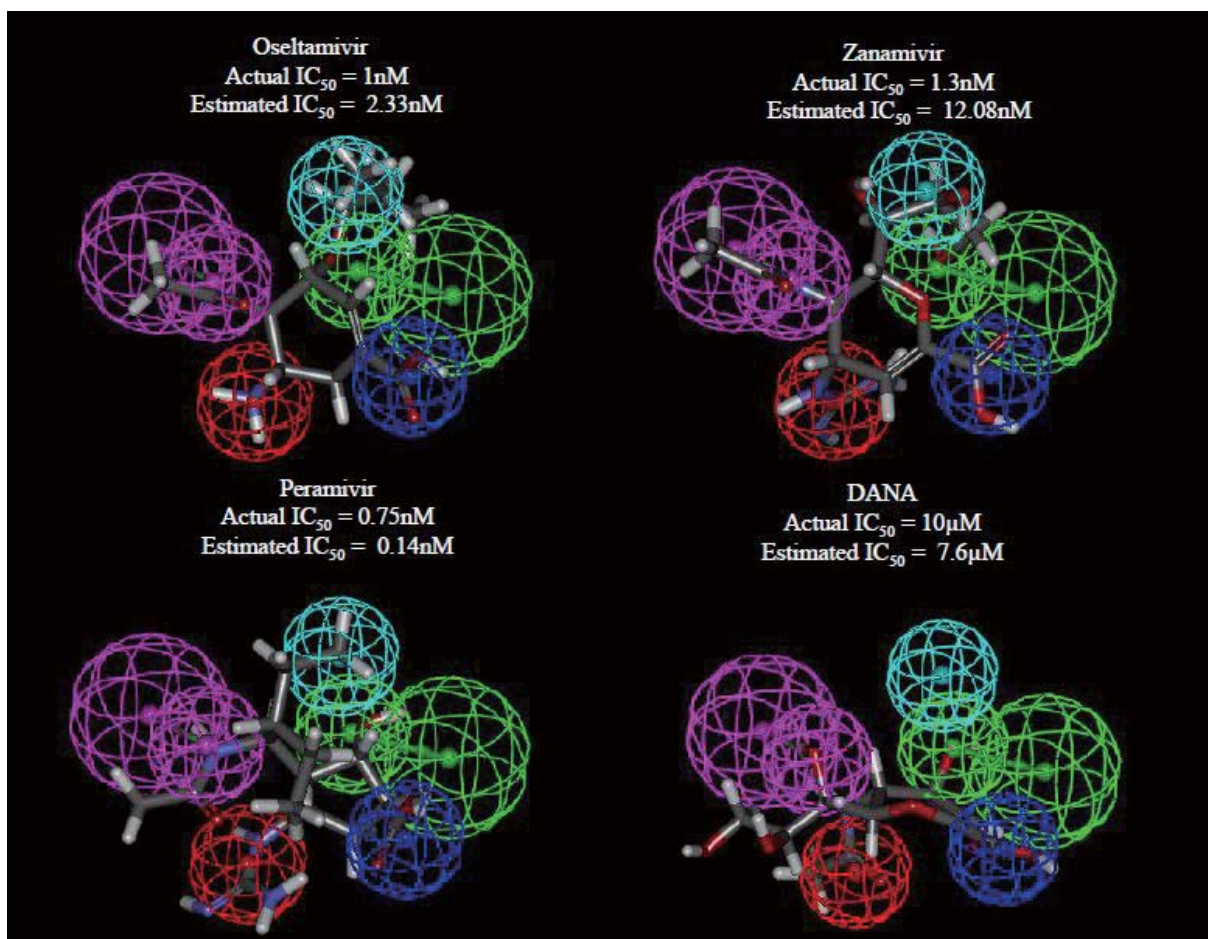




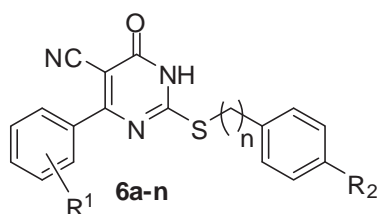
Figure 9

Inhibitors from the training set are mapped onto the Hypo1, and the slight differences between the actual and estimated  $IC_{50}$  shows pharmacophore is highly predictive.



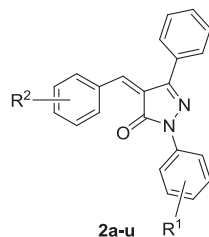
## TABLES

Table 1. Pyrimidine analogs synthesized by Dr. Ramajayam Ramachandran and their IC<sub>50</sub> against SARS 3CL<sup>pro</sup>



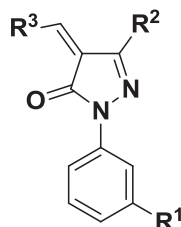
Compound	R <sup>1</sup>	R <sup>2</sup>	n	IC <sub>50</sub> (μM)
6a	H	H	1	N.D
6b	H	H	2	N.D
6c	H	4-NO <sub>2</sub>	1	35.3
6d	4-OCH <sub>3</sub>	H	1	N.D
6e	4-OCH <sub>3</sub>	H	2	20.3
6f	4-OCH <sub>3</sub>	4-NO <sub>2</sub>	1	26.3
6g	4-CH <sub>3</sub>	H	1	N.D
6h	4-CH <sub>3</sub>	H	2	N.D
6i	4-CH <sub>3</sub>	4-NO <sub>2</sub>	1	N.D
6j	3-NO <sub>2</sub>	H	2	N.D
6k	3-NO <sub>2</sub>	4-NO <sub>2</sub>	1	10.6±1.2
6l	4-Cl	H	2	16.9±1.3
6m	4-Cl	4-NO <sub>2</sub>	1	6.1±1.1
6n	3-Cl	H	1	N.D

Table 2. Pyrazolone analogs synthesized by Dr. Ramajayam Ramachandran and their IC<sub>50</sub> against SARS 3CL<sup>pro</sup> and CVB3 3C<sup>pro</sup>.



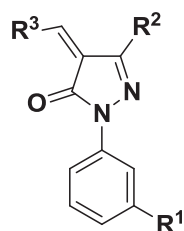
Compound	R <sup>1</sup>	R <sup>2</sup>	SARS 3CL <sup>pro</sup> IC <sub>50</sub> (μM)	CVB3 3C <sup>pro</sup> IC <sub>50</sub> (μM)
2a	H	H	N.D	N.D
2b	H	3-OCH <sub>3</sub>	N.D	N.D
2c	H	4-NHCOCH <sub>3</sub>	N.D	N.D
2d	H	4-COOH	18.0	51.1
2e	H	4-N(CH <sub>3</sub> ) <sub>2</sub>	N.D	N.D
2f	H	3-NO <sub>2</sub>	N.D	N.D
2g	4-Cl	H	N.D	N.D
2h	4-Cl	4-Cl	N.D	N.D
2i	4-Cl	4-COOH	13.9	25.0
2j	4-Cl	4-NHCOCH <sub>3</sub>	N.D	N.D
2k	4-Cl	4-OCH <sub>3</sub>	N.D	N.D
2l	4-Cl	4-OH	N.D	N.D
2m	4-OCH <sub>3</sub>	4-COOH	12.0	16.6
2n	4-CH(CH <sub>3</sub> ) <sub>2</sub>	4-COOH	N.D	N.D
2o	4-C(CH <sub>3</sub> ) <sub>3</sub>	4-COOH	N.D	N.D
2p	4-CN	4-COOH	5.5	20.8
2q	4-OCF <sub>3</sub>	4-COOH	42.0	98.8
2r	3-Cl	4-COOH	10.8	17.3
2s	3,4-Cl <sub>2</sub>	4-COOH	24.3	125.5
2t	4-F	4-COOH	6.8	22.4
2u	3-NO <sub>2</sub>	4-COOH	8.4	9.6

Table 3. Pyrazolone analogs synthesized by Mr. Vathan Kumar and their IC<sub>50</sub> against SARS 3CL<sup>pro</sup>.



Compound	R <sup>1</sup>	R <sup>2</sup>	R <sup>3</sup>	SARS 3CL <sup>pro</sup> IC <sub>50</sub> (μM)
Vk84	-COOH	-CH <sub>3</sub>		N.D
Vk85	-COOH	-CH <sub>3</sub>		N.D
Vk86	-COOH	-CF <sub>3</sub>	H	N.D
Vk87	-COOH	-Phenyl	H	N.D
Vk89	-COOH	-Phenyl		30.7
Vk90	-COOH	-CF <sub>3</sub>		N.D
Vk91	-COOH	-CF <sub>3</sub>		25.5
Vk92	-COOH	-Phenyl		N.D
Vk93	-H	-Phenyl		3.85± 0.73
Vk94	-NH <sub>2</sub>	-Phenyl		N.D
Vk99	-COOH	-Phenyl		6.04±0.45
Vk100	-COOH	-Phenyl		N.D
Vk101	-COOH	-Phenyl		N.D
vk102	-COOH	-Phenyl		37.3

Table 4. Pyrazolone analogs synthesized by Mr. Vathan Kumar and their IC<sub>50</sub> against H1N1 NA and H5N1 NA.



Compound	R <sup>1</sup>	R <sup>2</sup>	R <sup>3</sup>	H1N1 NA IC <sub>50</sub> (μM)	H5N1 NA IC <sub>50</sub> (μM)
Vk84	-COOH	-CH <sub>3</sub>		39.6	16.6±0.94
Vk85	-COOH	-CH <sub>3</sub>		20.1	26.0
Vk86	-COOH	-CF <sub>3</sub>	H	N.D	N.D
Vk87	-COOH	-Phenyl	H	N.D	N.D
Vk89	-COOH	-Phenyl		13.98±3.18	19.84±1.35
Vk90	-COOH	-CF <sub>3</sub>		36.5	39
Vk91	-COOH	-CF <sub>3</sub>		20.6	15.09±1.62
Vk92	-COOH	-Phenyl		13.31±0.83	49
Vk93	-H	-Phenyl		3.77±0.55	34.3
Vk94	-NH <sub>2</sub>	-Phenyl		1.70±0.03	11.23±0.97
Vk99	-COOH	-Phenyl		8.15±0.65	25
Vk100	-COOH	-Phenyl		N.D	N.D
Vk101	-COOH	-Phenyl		N.D	N.D
vk102	-COOH	-Phenyl		24.8	51.1

Table 5. Cost analysis of the 10 hypotheses generated by the HypoGen

	<b>Hypothesis Cost</b>	<b>Cost Difference</b>	<b>RMS</b>	<b>Correlation</b>	<b>Features</b>	<b>Max fit</b>
Hypo1	100.378	82.543	1.4498	0.923542	HA HD HY PI NI	15.2922
Hypo2	104.367	78.554	1.6168	0.902514	HA HD HY PI NI	15.0180
Hypo3	104.428	78.493	1.70263	0.888699	HD HD HY PI NI	13.7100
Hypo4	105.600	77.321	1.64347	0.899466	HD HD HY PI NI	15.2183
Hypo5	108.721	74.200	1.86574	0.863669	HA HA HY PI NI	13.0746
Hypo6	109.238	73.683	1.77269	0.881083	HA HD PI NI	12.0558
Hypo7	113.960	68.961	2.01169	0.839419	HA HD HY PI NI	13.1805
Hypo8	114.053	68.868	2.04134	0.833434	HD HD HY PI NI	12.3682
Hypo9	115.850	67.071	2.08359	0.855879	HA HA HY PI NI	12.5784
Hypo10	116.432	66.489	2.11579	0.819425	HA HD HY PI NI	11.9378

Null Cost: 182.921

Fixed Cost: 75.3195

$\Delta\text{Cost} = \text{Null Cost} - \text{Total Fixed Cost} = 107.6015$

Table 6. Output of the score hypothesis process on the training set

Name <sup>d</sup>	Activ	Estimate	Error <sup>a</sup>	Fit Value <sup>b</sup>	Activity Scale <sup>c</sup>		Mapped Features				
					Actual	Estimate	HBA	HBD	HY	NI	PI
1	0.75	0.14	-5.2	13.74	+++	+++	+	+	+	+	+
2	1	2.33	2.3	12.52	+++	+++	+	+	+	+	+
3	1.3	12.08	9.3	11.81	+++	+++	+	+	-	+	+
4	1.8	11.32	6.3	11.84	+++	+++	+	+	-	+	+
5	6	122.35	20.4	10.80	+++	++	+	+	-	+	+
6	15	22.98	1.5	11.53	+++	+++	+	+	+	+	+
7	100	83.74	-1.2	10.97	+++	+++	+	+	-	+	+
8	100	166.85	1.7	10.67	+++	++	+	+	+	+	+
9	620	278.87	-2.2	10.45	++	++	+	+	-	+	+
10	1,300	414.40	-3.1	10.27	++	++	+	-	+	+	+
11	2,500	14223.90	5.7	8.74	++	+	-	+	-	+	+
12	2,700	533.47	-5.1	10.16	++	++	-	+	+	+	+
13	5,000	2576.85	-1.9	9.48	++	++	+	+	-	+	+
14	5,000	26146.20	5.2	8.47	++	+	+	+	+	-	-
15	10,000	7594.19	-1.3	9.01	++	++	+	+	-	+	-
16	40,000	13133.20	-3.0	8.77	+	+	+	+	-	+	-
17	115,000	9979.76	-11.5	8.89	+	+	+	-	-	+	+
18	250,000	20642.00	-12.1	8.58	+	+	+	-	+	-	+

a. The error factor is computed as the ratio of the measured activity to the activity estimated by the hypothesis or the inverse if estimated is greater than measured.

b. Fit value indicates how well the features in the pharmacophore overlap the chemical features in the molecule.

c. Activity scale: +++, IC<sub>50</sub> < 100 nM (highly active); ++, 10,000 nM > IC<sub>50</sub> > 100 nM (moderately active); +, IC<sub>50</sub> > 10,000 nM (inactive).

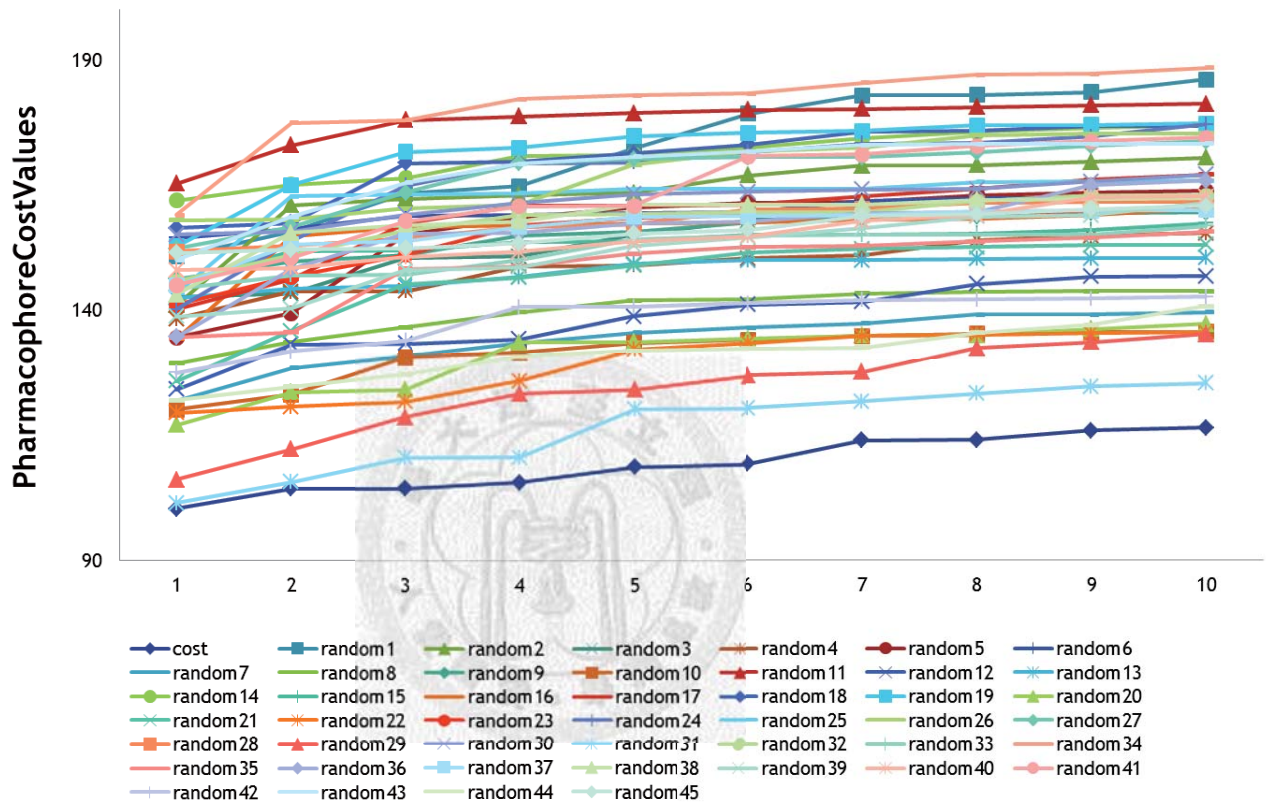
d. References: (Chand, Babu et al. 1997; Matthew A. Williams 1997; Kim, Lew et al. 1998; Lew, Wu et al. 1998; Atigadda, Brouillette et al. 1999; Babu, Chand et al. 2000; Wang, Chen et al. 2001; Chand, Babu et al. 2004; Makoto Yamashita 2009)



## GRAPHS

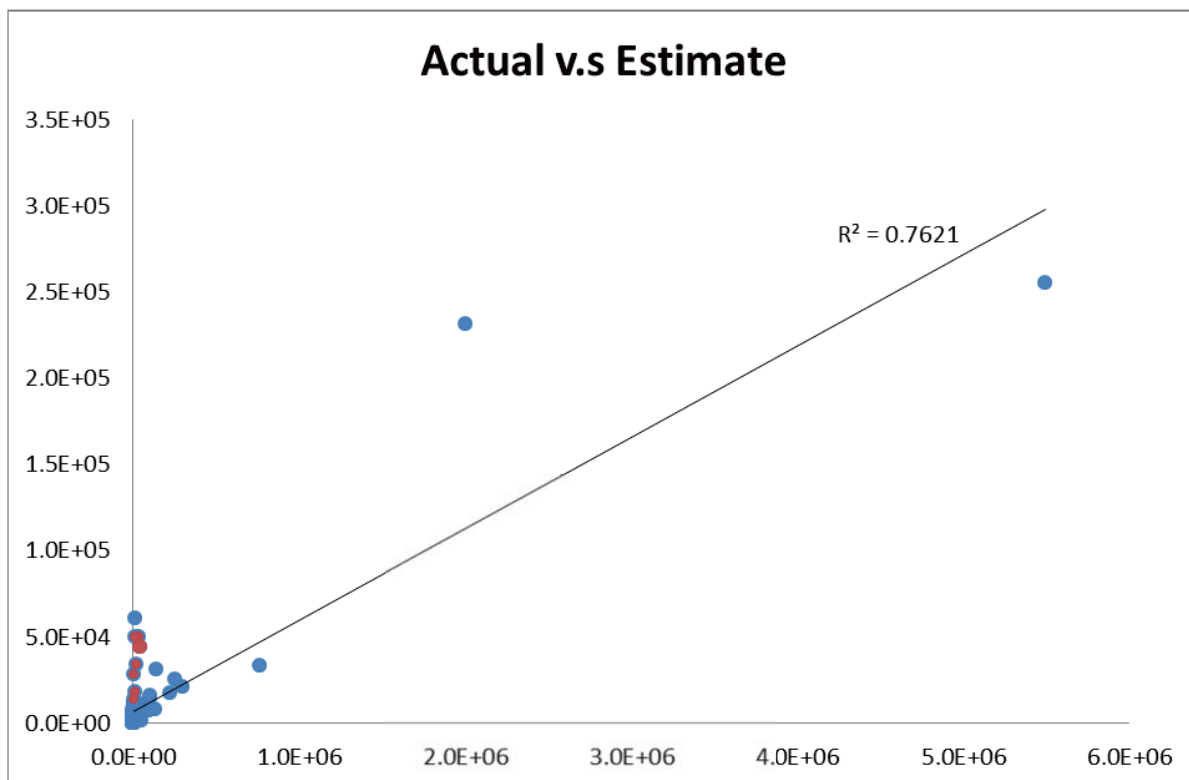
Graph 1

CatScramble result shows that none of the randomly generated hypotheses is better than the HypoGen. 98% confidence level was selected. Note that 4 of the scrambled run failed to generate the hypothesis.



Graph 2

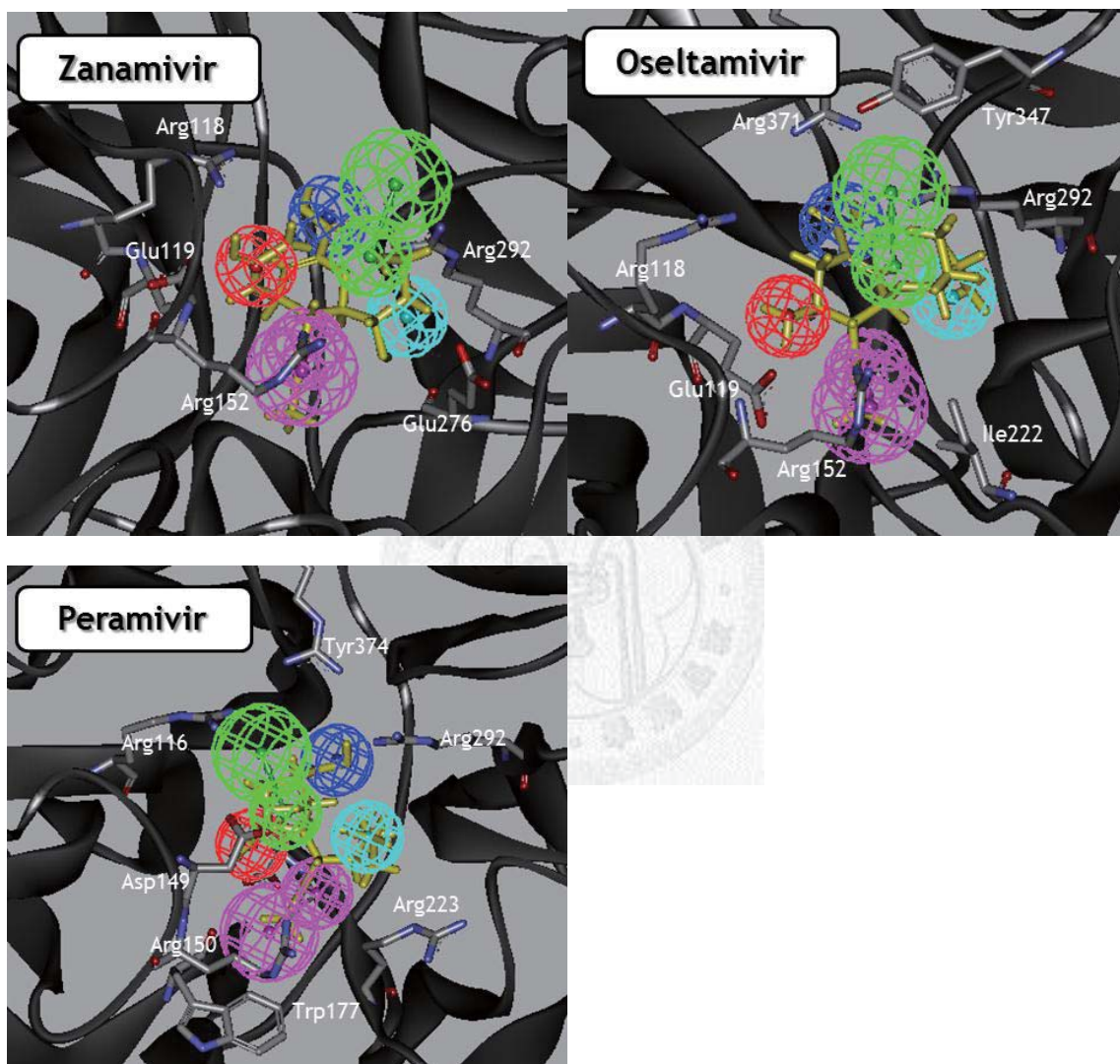
The plot of actual versus estimated  $IC_{50}$  values gives a correlation of 0.76. The compounds that our lab have synthesized are also included in the test set and they are marked in red.



## SUPPLEMENTARY

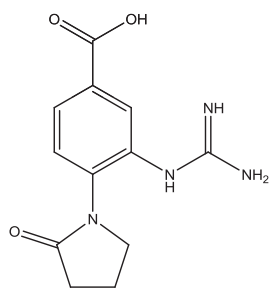
### Supplementary figure 1

The Hypo1 is mapped onto the active site of the NA. The crystal structure were is downloaded from PDB databank (PDB code 2HTY, 2HU0 and 2HTU)

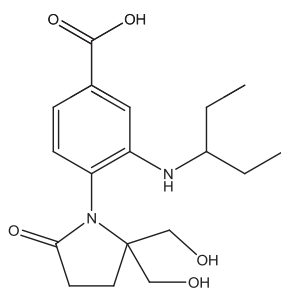


## Supplementary figure 2

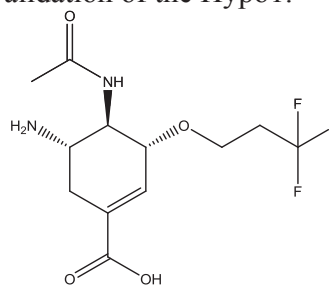
The structures of the test set being used for the validation of the Hypo1.



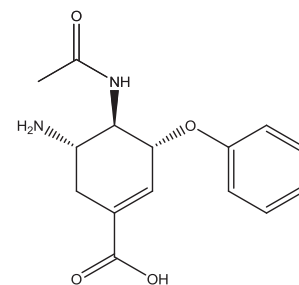
19



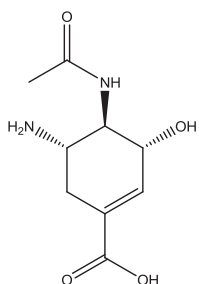
20



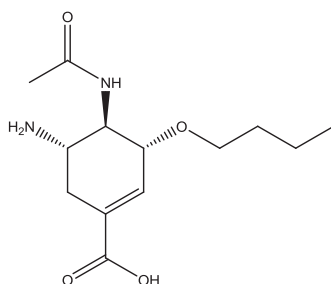
21



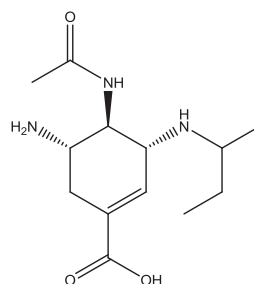
22



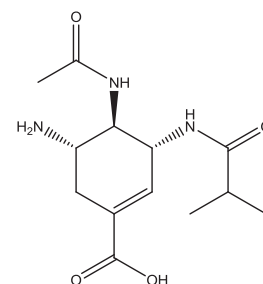
23



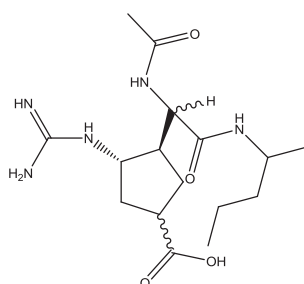
24



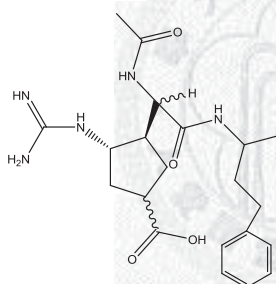
25



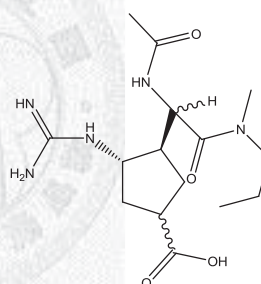
26



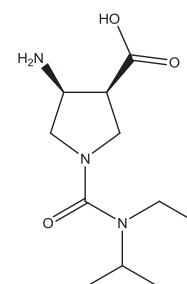
27



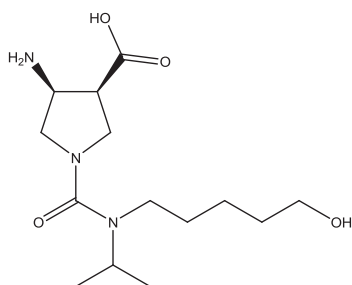
28



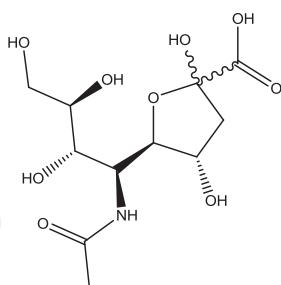
29



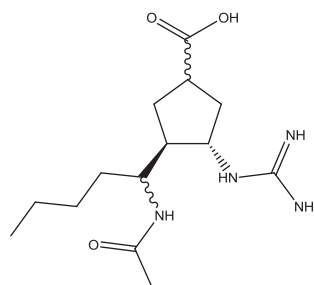
30



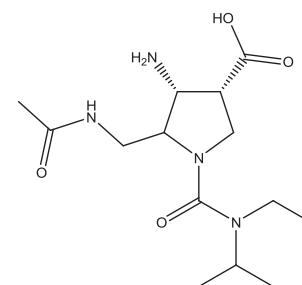
31



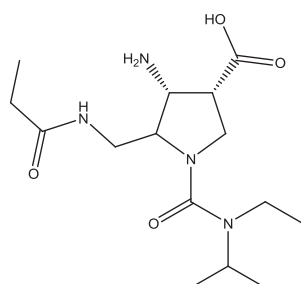
32



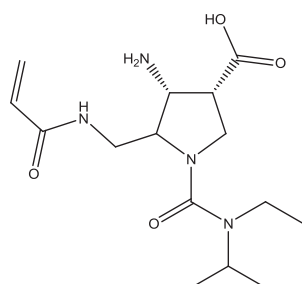
33



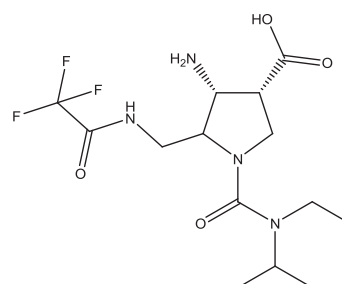
34



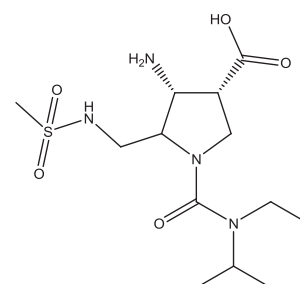
35



36

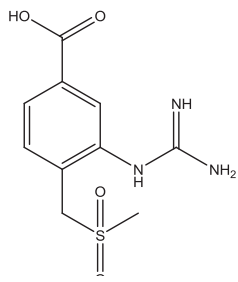


37

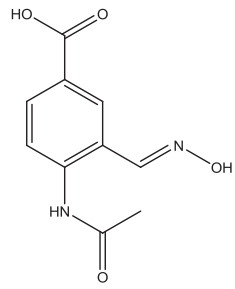


38

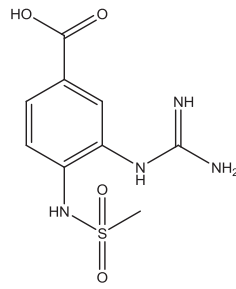
49



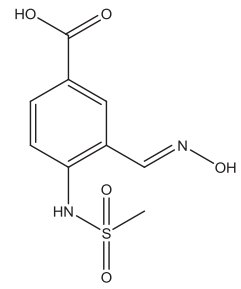
39



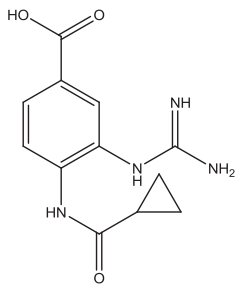
40



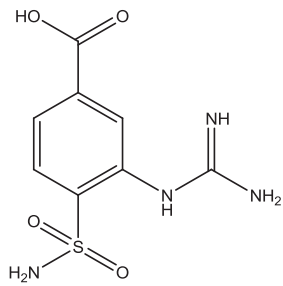
41



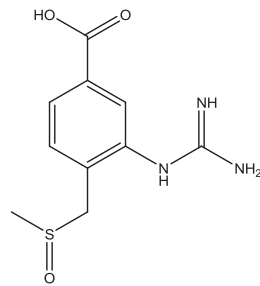
42



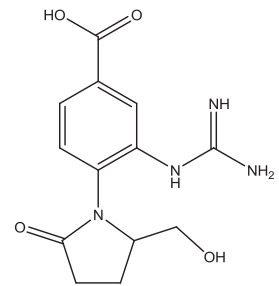
43



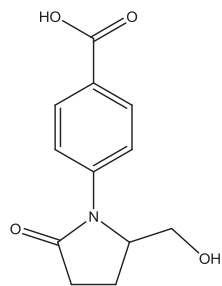
44



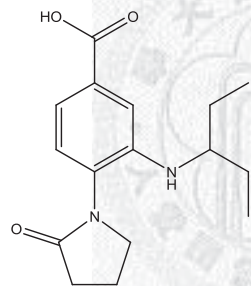
45



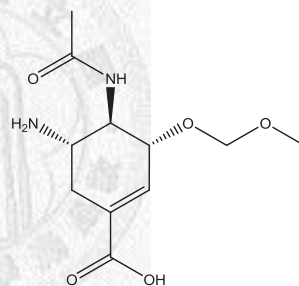
46



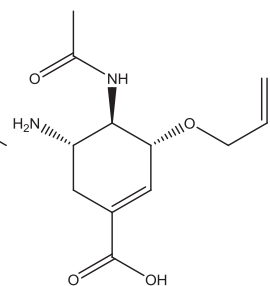
47



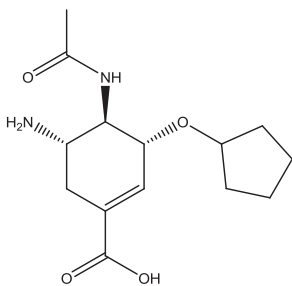
48



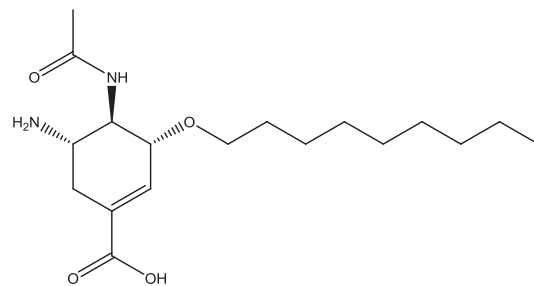
49



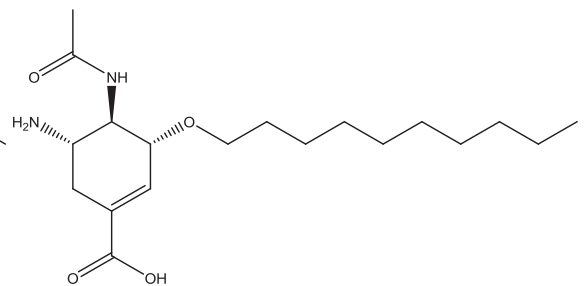
50



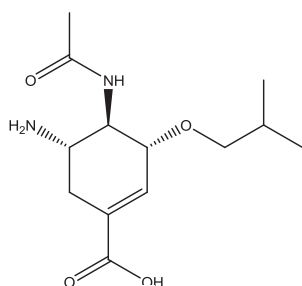
51



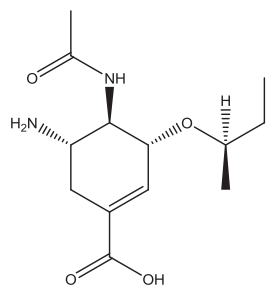
52



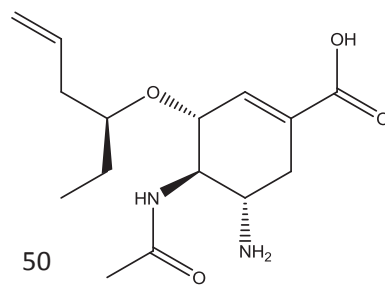
53



54

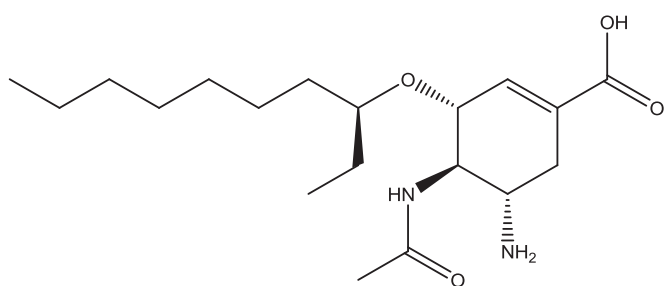


55

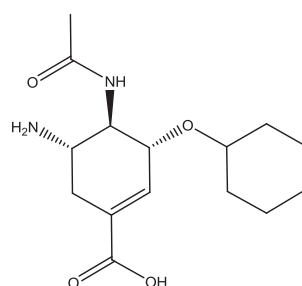


50

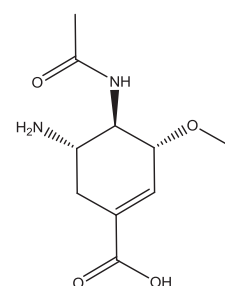
56



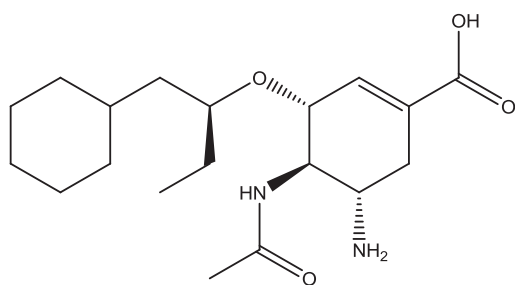
57



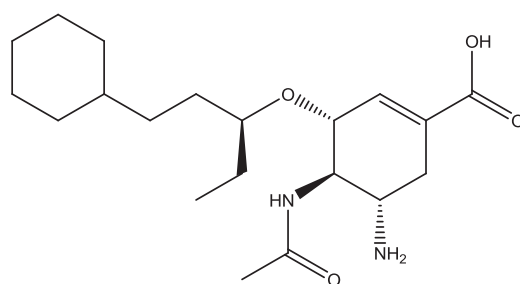
58



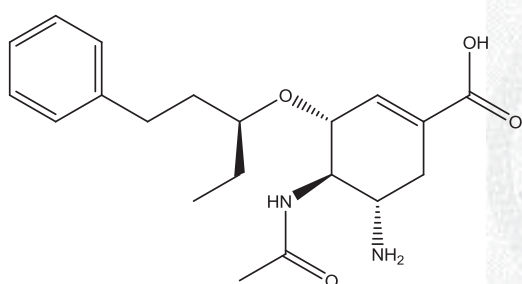
59



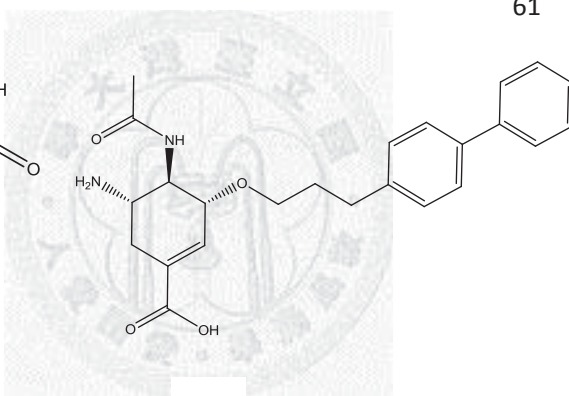
60



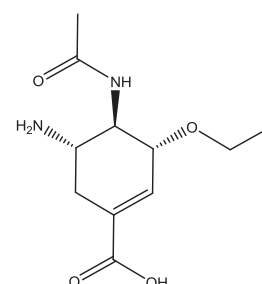
61



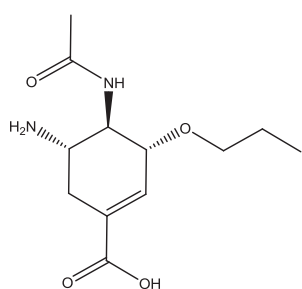
62



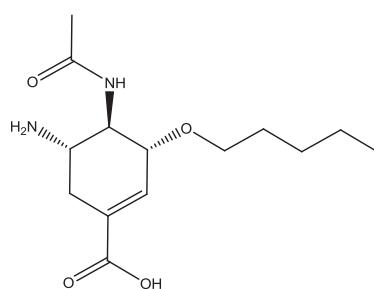
63



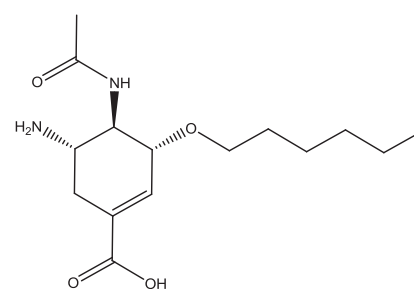
64



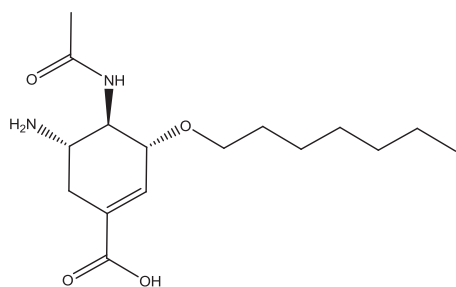
65



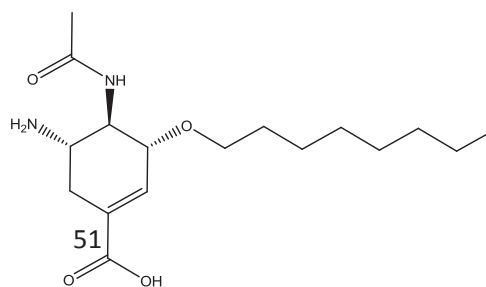
66



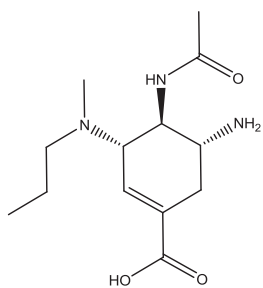
67



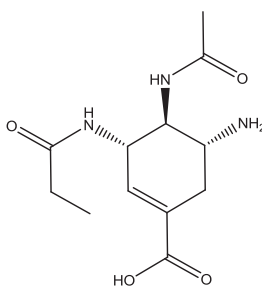
68



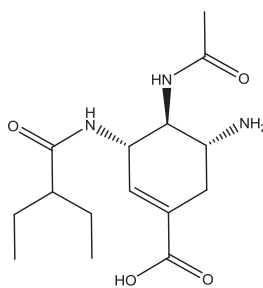
69



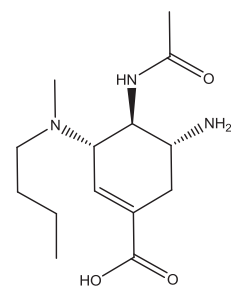
70



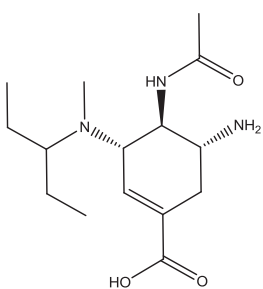
71



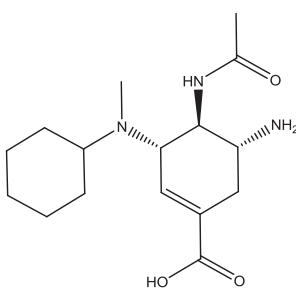
72



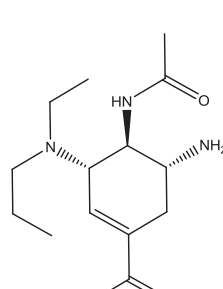
73



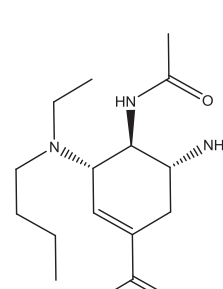
74



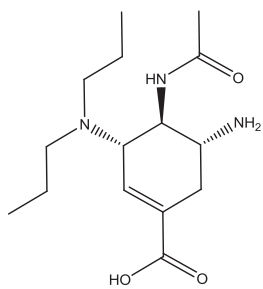
75



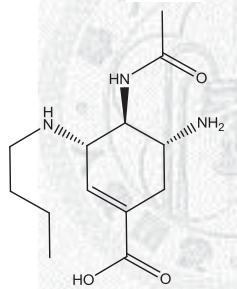
76



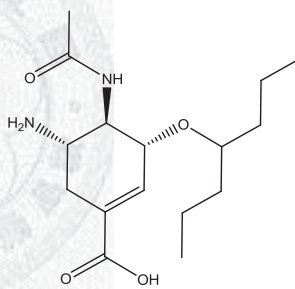
77



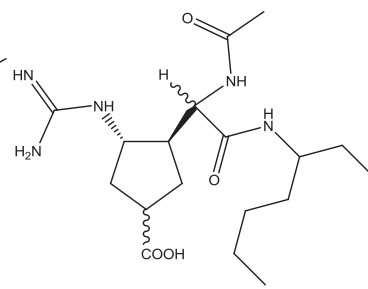
78



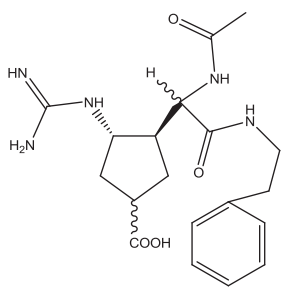
79



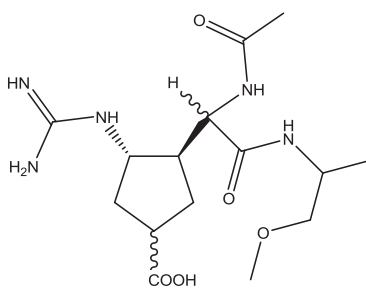
80



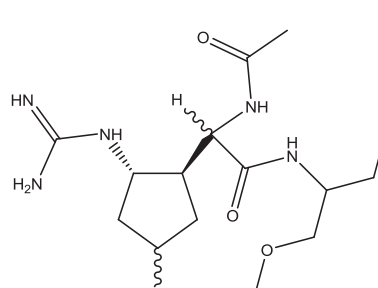
81



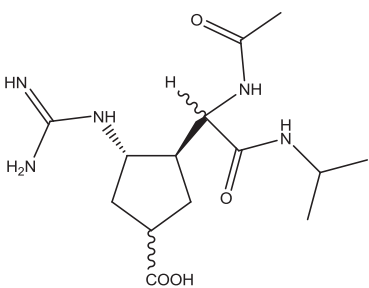
82



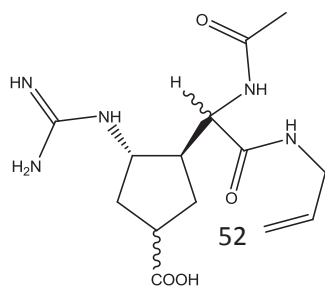
83



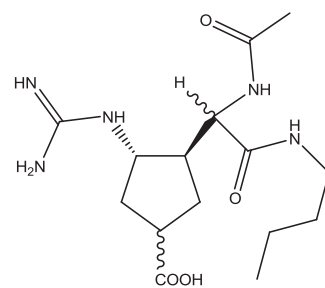
84



85

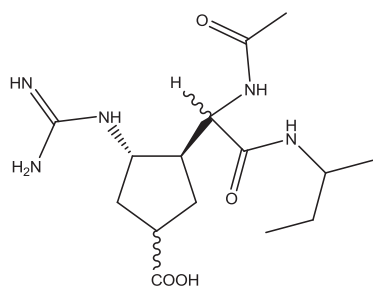


86

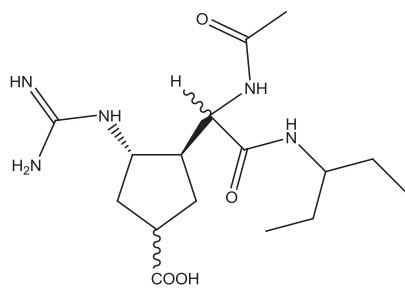


87

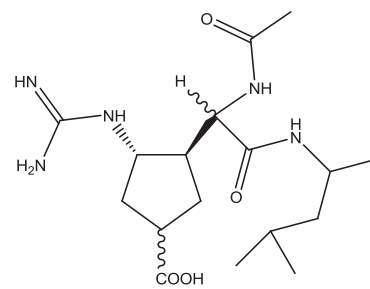




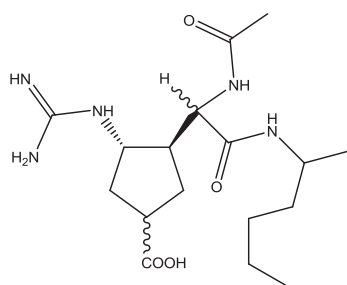
88



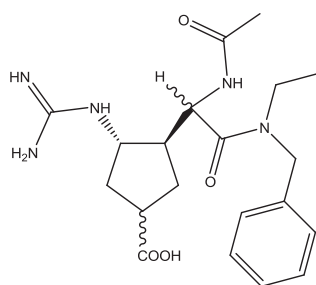
89



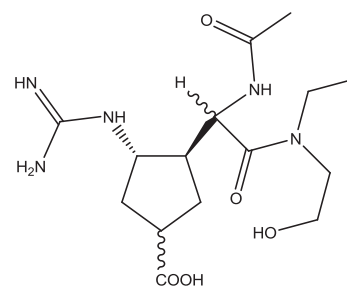
90



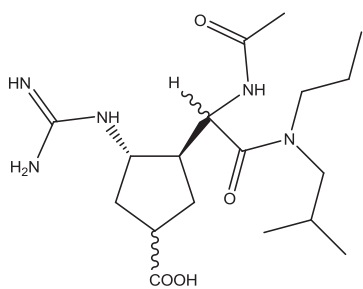
91



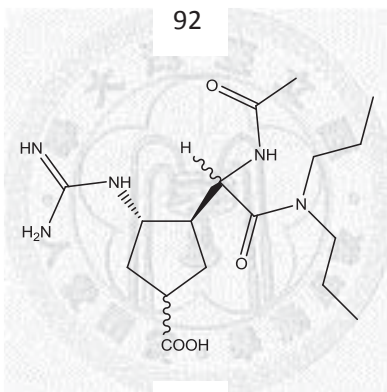
92



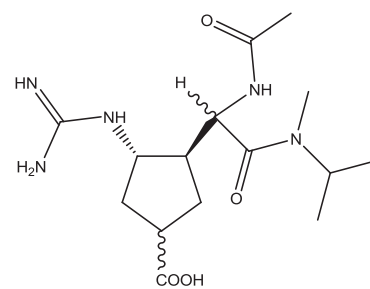
93



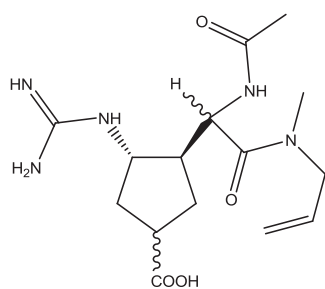
94



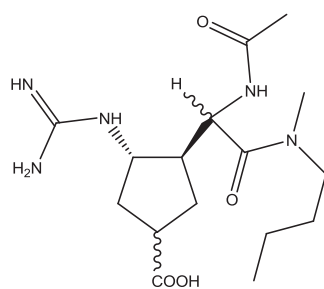
95



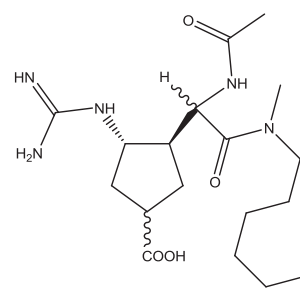
96



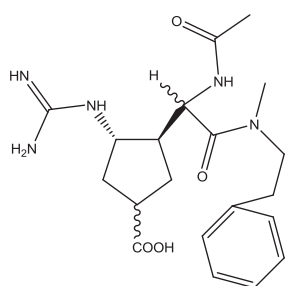
97



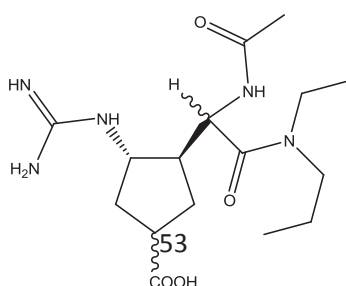
98



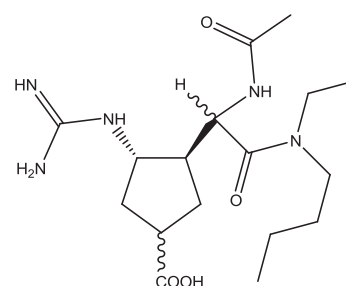
99



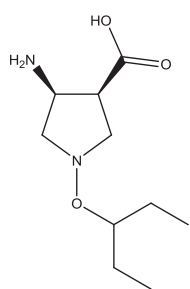
100



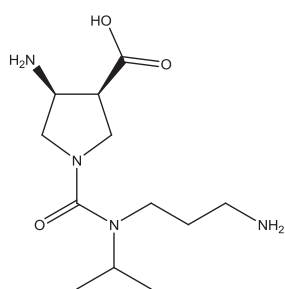
101



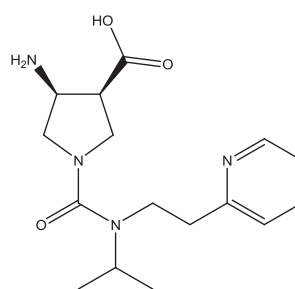
102



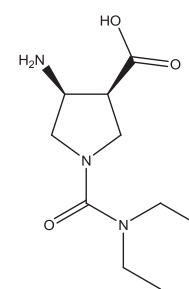
103



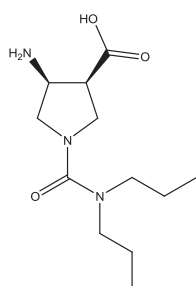
104



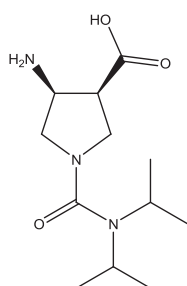
105



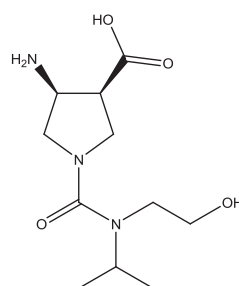
106



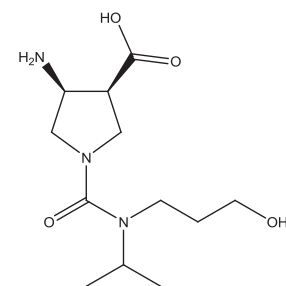
107



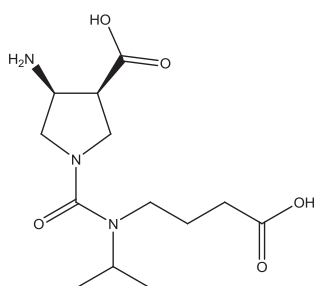
108



109



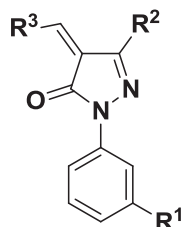
110



111



Supplementary table 1. Pyrazolone analogs synthesized by Mr. Vathan Kumar and their IC<sub>50</sub> against CVB3 3C<sup>pro</sup>.



Compound	R <sup>1</sup>	R <sup>2</sup>	R <sup>3</sup>	CVB3 3C <sup>pro</sup> IC <sub>50</sub> (μM)
Vk84	-COOH	-CH <sub>3</sub>		N.D
Vk85	-COOH	-CH <sub>3</sub>		N.D
Vk86	-COOH	-CF <sub>3</sub>	H	N.D
Vk87	-COOH	-Phenyl	H	N.D
Vk89	-COOH	-Phenyl		43.4
Vk90	-COOH	-CF <sub>3</sub>		25.3
Vk91	-COOH	-CF <sub>3</sub>		35.9
Vk92	-COOH	-Phenyl		8.51±1.73
Vk93	-H	-Phenyl		15.85± 2.43
Vk94	-NH <sub>2</sub>	-Phenyl		14.61±1.18
Vk99	-COOH	-Phenyl		N/A
Vk100	-COOH	-Phenyl		N.D
Vk101	-COOH	-Phenyl		N.D
vk102	-COOH	-Phenyl		N/A

Supplementary table 2. The actual and estimated IC<sub>50</sub> of the test set with t, the errors are listed in the table too.

Test set	IC <sub>50</sub> (nM)		Error	Fit Values	Scale		Mapped Features				
	Actual	Estimate			Actual	Estimate	HA	HD	HY	NI	PI
19	250,000	24,852.60	-10.06	8.50	+	+	+	-	+	-	+
20	48	351.573	7.32	10.35	+++	++	+	+	+	+	-
21	225	37.8149	-5.95	11.31	++	+++	+	-	+	+	+
22	530	101.472	-5.22	10.89	++	++	+	+	-	+	+
23	6,300	38.9609	-161.70	11.30	++	+++	+	+	-	+	+
24	300	135.77	-2.21	10.76	++	++	+	-	+	+	+
25	11	73.9947	6.73	11.02	+++	+++	-	+	+	+	+
26	6,400	820.93	-7.80	9.98	++	++	-	+	+	+	+
27	3,200	139.56	-22.93	10.75	++	++	-	+	+	+	+
28	230	104.29	-2.21	10.87	++	++	+	+	-	+	+
29	940	574.79	-1.64	10.13	++	++	+	+	+	+	+
30	1,600	9,991.36	6.24	8.89	++	+	+	-	+	-	+
31	2,000	7,504.09	3.75	9.02	++	++	+	-	+	-	+
32	40,000	6,919.18	-5.78	9.05	+	++	+	+	-	+	-
33	100	281.848	2.82	10.44	+++	++	-	+	+	+	+
34	7,500	6,655.94	-1.13	9.07	++	++	+	-	+	-	+
35	16,000	6,632.47	-2.41	9.07	+	++	+	-	+	-	+
36	96,000	6,901.24	-13.91	9.05	+	++	+	-	+	-	+
37	280	6,673.49	23.83	9.07	++	++	+	-	+	-	+
38	130,000	7,604.60	-17.09	9.01	+	++	+	-	+	-	+
39	300,000	20,641.30	-14.53	8.58	+	+	+	+	+	-	-
40	5,500,000	255,044	-21.56	7.49	+	+	+	+	-	+	-
41	100,000	14,257.80	-7.01	8.74	+	+	-	+	-	+	+
42	2,000,000	231,145	-8.65	7.53	+	+	+	+	+	-	-
43	100,000	15,741.20	-6.35	8.69	+	+	-	+	-	+	+
44	9,000	60,706.20	6.75	8.11	++	+	-	+	-	+	+
45	140,000	30,968.10	-4.52	8.40	+	+	+	+	+	-	-
46	20,000	3,245.79	-6.16	9.38	+	++	+	+	+	+	-
47	760,000	33,127.80	-22.94	8.37	+	+	+	+	+	-	-
48	222,000	17,519.60	-12.67	8.65	+	+	-	+	+	+	-
49	2,000	316.28	-6.32	10.39	++	++	+	+	-	+	+
50	2,200	683.77	-3.22	10.06	++	++	+	+	+	+	+
51	22	205.297	9.33	10.58	+++	++	-	+	+	+	+
52	210	15.7997	-13.29	11.69	++	+++	+	+	+	+	+

53	600	13.7271	-43.71	11.75	++	+++	+	-	+	+	+
54	200	2.05	-97.42	12.58	++	+++	+	+	+	+	+
55	10	1.28455	-7.78	12.78	+++	+++	+	+	+	+	+
56	1	1.67	1.67	12.67	+++	+++	+	+	+	+	+
57	1	33.0233	33.02	11.37	+++	+++	+	-	+	+	+
58	60	387.671	6.46	10.30	+++	++	+	-	+	+	+
59	3,700	357.889	-10.34	10.34	++	++	+	+	-	+	+
60	16	76.0945	4.76	11.01	+++	+++	+	+	-	+	+
61	1	47.6894	47.69	11.21	+++	+++	+	+	+	+	+
62	0.3	6.79097	22.64	12.06	+++	+++	+	+	+	+	+
63	90	22.61	-3.98	11.54	+++	+++	+	-	+	+	+
64	2,000	484.4	-4.13	10.21	++	++	+	+	-	+	+
65	180	163.26	-1.10	10.68	++	++	+	-	+	+	+
66	200	0.93706	-213.43	12.92	++	+++	+	+	+	+	+
67	150	104.19	-1.44	10.87	++	++	+	+	+	+	+
68	270	95.2521	-2.83	10.91	++	+++	+	+	+	+	+
69	180	56.1261	-3.21	11.14	++	+++	+	-	+	+	+
70	65	1,906.26	29.33	9.61	+++	++	+	-	+	+	+
71	2,700	1,309.40	-2.06	9.77	++	++	+	+	-	+	+
72	4,000	58.3924	-68.50	11.13	++	+++	+	+	+	+	-
73	180	756.385	4.20	10.01	++	++	+	-	+	+	+
74	6	31.881	5.31	11.39	+++	+++	+	-	+	+	+
75	200	8,039.41	40.20	8.99	++	++	-	+	-	+	+
76	90	949.41	10.55	9.91	+++	++	+	-	+	+	+
77	85	425.232	5.00	10.26	+++	++	+	-	+	+	+
78	12	3,420.58	285.05	9.36	+++	++	+	-	+	+	+
79	200	20.45	-9.78	11.58	++	+++	-	+	+	+	+
80	16	17.27	1.08	11.65	+++	+++	+	+	+	+	+
81	1,800	643.06	-2.80	10.08	++	++	+	+	-	+	+
82	6,400	220.62	-29.01	10.55	++	++	+	-	+	+	+
83	4,400	247.98	-17.74	10.50	++	++	+	+	-	+	+
84	2,200	163.74	-13.44	10.68	++	++	+	+	-	+	+
85	1,340	115.199	-11.63	10.83	++	++	+	+	+	+	+
86	720	3,390.20	4.71	9.36	++	++	+	+	-	+	+
87	9,900	904.764	-10.94	9.94	++	++	+	+	-	+	+
88	410	263.377	-1.56	10.47	++	++	+	-	+	+	+
89	80	858.299	10.73	9.96	+++	++	+	+	-	+	+
90	8,400	1,306.04	-6.43	9.78	++	++	+	+	+	+	-

91	8,700	316.884	-27.45	10.39	++	++	+	-	+	+	+
92	670	216.153	-3.10	10.56	++	++	+	-	+	+	+
93	720	33.6263	-21.41	11.37	++	+++	+	-	+	+	+
94	200	383.25	1.92	10.31	++	++	+	-	+	+	+
95	60	5,177.98	86.30	9.18	+++	++	+	-	+	+	+
96	3,210	113.867	-28.19	10.84	++	++	+	+	+	+	-
97	650	27.1795	-23.92	11.46	++	+++	+	+	+	+	+
98	2,400	411.228	-5.84	10.28	++	++	-	+	+	+	+
99	12,400	431.713	-28.72	10.26	+	++	-	+	+	+	+
100	8,000	1,737.01	-4.61	9.65	++	++	+	-	+	+	+
101	130	75.6277	-1.72	11.01	++	+++	+	+	+	+	+
102	4,300	463.85	-9.27	10.23	++	++	-	+	+	+	+
103	22,000	6,982.65	-3.15	9.05	+	++	+	-	+	-	+
104	46,000	1,518.42	-30.29	9.71	+	++	+	-	+	+	+
105	1,300	12,325.30	9.48	8.80	++	+	+	-	+	-	+
106	25,000	11,045	-2.26	8.85	+	+	+	-	+	-	+
107	32,000	10,814.60	-2.96	8.86	+	+	+	-	+	-	+
108	4,000	11,187	2.80	8.84	++	+	+	-	+	-	+
109	21,000	221.07	-94.99	10.55	+	++	+	+	+	-	+
110	2,100	3,349.55	1.60	9.37	++	++	+	+	+	-	+
111	19,000	612.57	-31.02	10.10	+	++	+	-	+	+	+

Supplementary table 3. The inhibitors to the NA that our lab have discovered so far are evaluated by the Hypo1

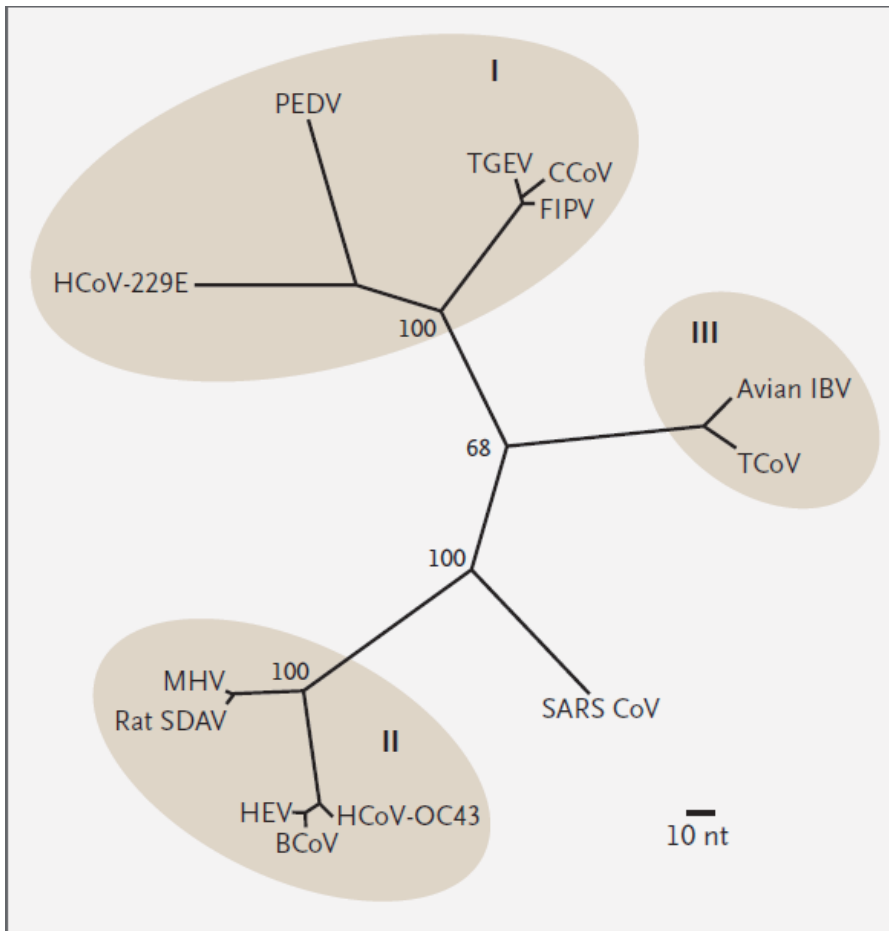
Compounds	Actual IC <sub>50</sub> (nM)	Estimated IC <sub>50</sub> (nM)	Error	Features				
				HA	HD	HY	NI	PI
VK84	39600	44285.5	1.12	1	0	1	1	0
VK89	13980	50085.3	3.58	1	0	1	1	0
VK90	36500	49782.6	1.36	1	0	1	1	0
VK91	20600	33714.7	1.64	1	0	1	1	0
VK92	13310	17959.6	1.35	1	0	1	1	0
VK93	3770	13297.7	3.53	1	0	1	1	0
VK94	1770	28143.2	16.66	1	0	1	1	0
VK99	8150	81844.3	6.36	1	0	1	1	0
VK102	24800	49717.5	2.00	1	0	1	1	0



## APPENDIXES

### Appendix 1

The phylogenetic tree of SARS-CoV



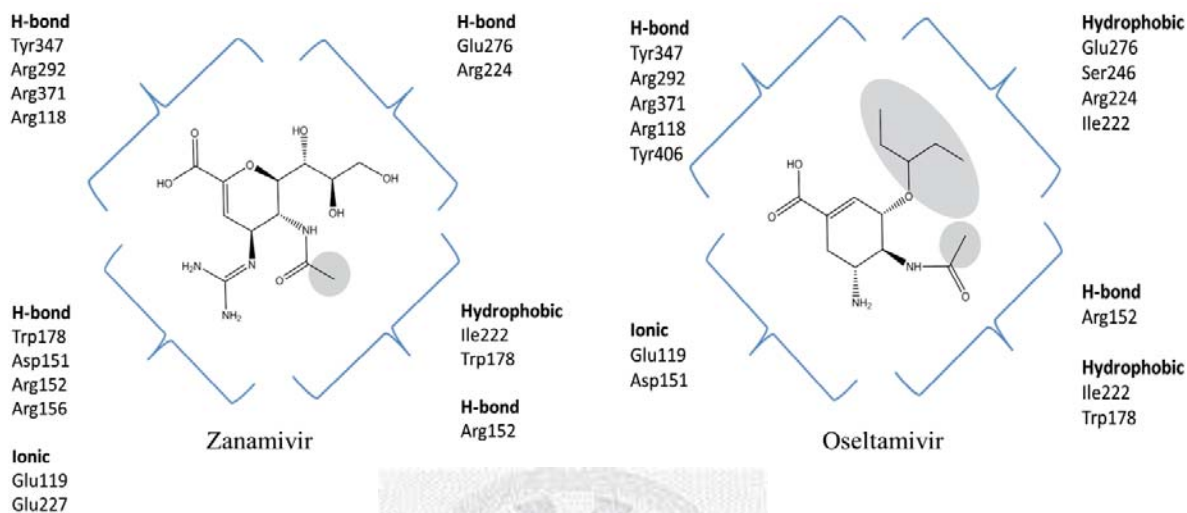
## Appendix 2

Superimposition of SARS-CoV 3CL<sup>pro</sup> (PDB code 2ZU4) and CVB3 3C<sup>pro</sup> (2ZU3).



## Appendix 3

The binding modes of the zanamivir and the oseltamivir with the NA. The pictures were drawn based on inhibitor-enzyme complex Xx-ray structure (PDB code 2HTY and 2HU0).



## REFERENCES

- "Influenza (Seasonal)." from <http://www.who.int/mediacentre/factsheets/fs211/en/>.
- "International Programs." from <http://www.census.gov/population/international/>.
- Ahlquist, P., A. O. Noueir, et al. (2003). "Host factors in positive-strand RNA virus genome replication." *J Virol* **77**(15): 8181-8186.
- Anand, K., J. Ziebuhr, et al. (2003). "Coronavirus main proteinase (3CLpro) structure: basis for design of anti-SARS drugs." *Science* **300**(5626): 1763-1767.
- Atigadda, V. R., W. J. Brouillette, et al. (1999). "Potent inhibition of influenza sialidase by a benzoic acid containing a 2-pyrrolidinone substituent." *J Med Chem* **42**(13): 2332-2343.
- Baboonian, C. and T. Treasure (1997). "Meta-analysis of the association of enteroviruses with human heart disease." *Heart* **78**(6): 539-543.
- Babu, Y. S., P. Chand, et al. (2000). "BCX-1812 (RWJ-270201): discovery of a novel, highly potent, orally active, and selective influenza neuraminidase inhibitor through structure-based drug design." *J Med Chem* **43**(19): 3482-3486.
- Bharatham, N., K. Bharatham, et al. (2007). "Pharmacophore identification and virtual screening for methionyl-tRNA synthetase inhibitors." *J Mol Graph Model* **25**(6): 813-823.
- Binford, S. L., F. Maldonado, et al. (2005). "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 Chemother* **49**(2): 619-626.

- Chand, P., Y. S. Babu, et al. (1997). "Design and synthesis of benzoic acid derivatives as influenza neuraminidase inhibitors using structure-based drug design." J Med Chem **40**(25): 4030-4052.
- Chand, P., Y. S. Babu, et al. (2004). "Syntheses and neuraminidase inhibitory activity of multisubstituted cyclopentane amide derivatives." J Med Chem **47**(8): 1919-1929.
- Chen, L., C. Gui, et al. (2005). "Cinanserin is an inhibitor of the 3C-like proteinase of severe acute respiratory syndrome coronavirus and strongly reduces virus replication in vitro." J Virol **79**(11): 7095-7103.
- Cheng, P. K., T. W. Leung, et al. (2009). "Oseltamivir- and amantadine-resistant influenza viruses A (H1N1)." Emerg Infect Dis **15**(6): 966-968.
- Consortium., C. S. M. E. (2004). "Molecular evolution of the SARS coronavirus during the course of the SARS epidemic in China." Science **303**(5664): 1666-1669.
- Debnath, A. K. (2002). "Pharmacophore mapping of a series of 2,4-diamino-5-deazapteridine inhibitors of Mycobacterium avium complex dihydrofolate reductase." J Med Chem **45**(1): 41-53.
- Drosten, C., S. Gunther, et al. (2003). "Identification of a novel coronavirus in patients with severe acute respiratory syndrome." N Engl J Med **348**(20): 1967-1976.
- Du, L. P., M. Y. Li, et al. (2005). "Characterization of binding site of closed-state KCNQ1 potassium channel by homology modeling, molecular docking, and pharmacophore identification." Biochem Biophys Res Commun **332**(3): 677-687.
- Fan, K., P. Wei, et al. (2004). "Biosynthesis, purification, and substrate specificity of severe acute respiratory syndrome coronavirus 3C-like proteinase." J Biol Chem **279**(3): 1637-1642.

- Fields, B. N. D. M. K., Robert M. Chanock, Joseph L. Melnick, Bernard Roizman, Robert E. Shope (1985). "Fields Virology." New York: Raven Press: 739-794.
- Guan, Y., B. J. Zheng, et al. (2003). "Isolation and characterization of viruses related to the SARS coronavirus from animals in southern China." Science **302**(5643): 276-278.
- Hayden, F. G., R. B. Belshe, et al. (1989). "Emergence and apparent transmission of rimantadine-resistant influenza A virus in families." N Engl J Med **321**(25): 1696-1702.
- Huang, I. C., W. Li, et al. (2008). "Influenza A virus neuraminidase limits viral superinfection." J Virol **82**(10): 4834-4843.
- Kim, C. U., W. Lew, et al. (1997). "Influenza neuraminidase inhibitors possessing a novel hydrophobic interaction in the enzyme active site: design, synthesis, and structural analysis of carbocyclic sialic acid analogues with potent anti-influenza activity." J Am Chem Soc **119**(4): 681-690.
- Kim, C. U., W. Lew, et al. (1998). "Structure-activity relationship studies of novel carbocyclic influenza neuraminidase inhibitors." J Med Chem **41**(14): 2451-2460.
- Ksiazek, T. G., D. Erdman, et al. (2003). "A novel coronavirus associated with severe acute respiratory syndrome." N Engl J Med **348**(20): 1953-1966.
- Kuo, C.-J. (2009). Characterization, Inhibition, and Engineering of 3C and 3C-Like Viral Proteases. Institute of Biochemical Science, National Taiwan University.
- Kuo, C. J., Y. H. Chi, et al. (2004). "Characterization of SARS main protease and inhibitor assay using a fluorogenic substrate." Biochem Biophys Res Commun **318**(4): 862-867.

- Kuo, C. J., H. G. Liu, et al. (2009). "Individual and common inhibitors of coronavirus and picornavirus main proteases." FEBS Lett **583**(3): 549-555.
- Kuo, C. J., J. J. Shie, et al. (2008). "Design, synthesis, and evaluation of 3C protease inhibitors as anti-enterovirus 71 agents." Bioorg Med Chem **16**(15): 7388-7398.
- Lau, S. K., P. C. Woo, et al. (2005). "Severe acute respiratory syndrome coronavirus-like virus in Chinese horseshoe bats." Proc Natl Acad Sci U S A **102**(39): 14040-14045.
- Lee, C. C., C. J. Kuo, et al. (2009). "Structural basis of inhibition specificities of 3C and 3C-like proteases by zinc-coordinating and peptidomimetic compounds." J Biol Chem **284**(12): 7646-7655.
- Lee, E. S., W. G. Lee, et al. (2007). "Development of potent inhibitors of the coxsackievirus 3C protease." Biochem Biophys Res Commun **358**(1): 7-11.
- Lee, N., D. Hui, et al. (2003). "A major outbreak of severe acute respiratory syndrome in Hong Kong." N Engl J Med **348**(20): 1986-1994.
- Lew, W., X. Chen, et al. (2000). "Discovery and development of GS 4104 (oseltamivir): an orally active influenza neuraminidase inhibitor." Curr Med Chem **7**(6): 663-672.
- Lew, W., H. Wu, et al. (1998). "A new series of C3-aza carbocyclic influenza neuraminidase inhibitors: synthesis and inhibitory activity." Bioorg Med Chem Lett **8**(23): 3321-3324.
- Li, W., Z. Shi, et al. (2005). "Bats are natural reservoirs of SARS-like coronaviruses." Science **310**(5748): 676-679.
- Makoto Yamashita, T. T., Masayo Kakuta, Akane Tokumitsu, Hatsumi Nasu, and Shuku Kubo (2009). "CS-8958, a Prodrug of the New Neuraminidase Inhibitor R-125489,



- Shows Long-Acting Anti-Influenza Virus Activity." Antimicrobial Agents and Chemotherapy **53**: 7.
- Marra, M. A., S. J. Jones, et al. (2003). "The Genome sequence of the SARS-associated coronavirus." Science **300**(5624): 1399-1404.
- Matthew A. Williams, W. L., Dirk B. Mendel, Chun Y. Tai, Paul A. Escarpe, W. Graeme Laver, Raymond C. Stevens and Choung U. Kim (1997). "Structure-activity relationships of carbocyclic influenza neuraminidase inhibitors." Bioorganic & Medicinal Chemistry Letters **7**(14): 6.
- Maze, S. S. and R. J. Adolph (1990). "Myocarditis: unresolved issues in diagnosis and treatment." Clin Cardiol **13**(2): 69-79.
- Memoli, M. J., R. J. Hrabal, et al. (2010). "Rapid selection of oseltamivir- and peramivir-resistant pandemic H1N1 virus during therapy in 2 immunocompromised hosts." Clin Infect Dis **50**(9): 1252-1255.
- Niu, C., J. Yin, et al. (2008). "Molecular docking identifies the binding of 3-chloropyridine moieties specifically to the S1 pocket of SARS-CoV Mpro." Bioorg Med Chem **16**(1): 293-302.
- Peiris, J. S., S. T. Lai, et al. (2003). "Coronavirus as a possible cause of severe acute respiratory syndrome." Lancet **361**(9366): 1319-1325.
- Perola, E., W. P. Walters, et al. (2004). "A detailed comparison of current docking and scoring methods on systems of pharmaceutical relevance." Proteins **56**(2): 235-249.
- Pinto, L. H. and R. A. Lamb (2006). "The M2 proton channels of influenza A and B viruses." J Biol Chem **281**(14): 8997-9000.

- Ramajayam, R., K. P. Tan, et al. (2010). "Synthesis and evaluation of pyrazolone compounds as SARS-coronavirus 3C-like protease inhibitors." Bioorg Med Chem **18**(22): 7849-7854.
- Ramajayam, R., K. P. Tan, et al. (2010). "Synthesis, docking studies, and evaluation of pyrimidines as inhibitors of SARS-CoV 3CL protease." Bioorg Med Chem Lett **20**(12): 3569-3572.
- Rose, N. R. and S. L. Hill (1996). "The pathogenesis of postinfectious myocarditis." Clin Immunol Immunopathol **80**(3 Pt 2): S92-99.
- Russell, R. J., L. F. Haire, et al. (2006). "The structure of H5N1 avian influenza neuraminidase suggests new opportunities for drug design." Nature **443**(7107): 45-49.
- Skehel, J. J. and D. C. Wiley (2000). "Receptor binding and membrane fusion in virus entry: the influenza hemagglutinin." Annu Rev Biochem **69**: 531-569.
- Snijder, E. J., P. J. Bredenbeek, et al. (2003). "Unique and conserved features of genome and proteome of SARS-coronavirus, an early split-off from the coronavirus group 2 lineage." J Mol Biol **331**(5): 991-1004.
- von Itzstein, M., W. Y. Wu, et al. (1993). "Rational design of potent sialidase-based inhibitors of influenza virus replication." Nature **363**(6428): 418-423.
- Wang, G. T., Y. Chen, et al. (2001). "Design, synthesis, and structural analysis of influenza neuraminidase inhibitors containing pyrrolidine cores." J Med Chem **44**(8): 1192-1201.

- Watanabe, A., S. C. Chang, et al. (2010). "Long-acting neuraminidase inhibitor laninamivir octanoate versus oseltamivir for treatment of influenza: A double-blind, randomized, noninferiority clinical trial." Clin Infect Dis **51**(10): 1167-1175.
- Yang, H., M. Yang, et al. (2003). "The crystal structures of severe acute respiratory syndrome virus main protease and its complex with an inhibitor." Proc Natl Acad Sci U S A **100**(23): 13190-13195.
- Zhang, J., C. Huitema, et al. (2008). "Aryl methylene ketones and fluorinated methylene ketones as reversible inhibitors for severe acute respiratory syndrome (SARS) 3C-like proteinase." Bioorg Chem **36**(5): 229-240.
- Zhang, J., H. I. Pettersson, et al. (2007). "Design, synthesis, and evaluation of inhibitors for severe acute respiratory syndrome 3C-like protease based on phthalhydrazide ketones or heteroaromatic esters." J Med Chem **50**(8): 1850-1864.
- Zhang, J., K. Yu, et al. (2006). "Neuraminidase pharmacophore model derived from diverse classes of inhibitors." Bioorg Med Chem Lett **16**(11): 3009-3014.



Contents lists available at ScienceDirect

## Bioorganic &amp; Medicinal Chemistry Letters

journal homepage: [www.elsevier.com/locate/bmcl](http://www.elsevier.com/locate/bmcl)

## Synthesis, docking studies, and evaluation of pyrimidines as inhibitors of SARS-CoV 3CL protease

R. Ramajayam<sup>a</sup>, Kian-Pin Tan<sup>b</sup>, Hun-Ge Liu<sup>b</sup>, Po-Huang Liang<sup>a,b,\*</sup><sup>a</sup> Institute of Biological Chemistry, Academia Sinica, 128 Academia Road, Taipei 11529, Taiwan<sup>b</sup> Institute of Biochemical Sciences, National Taiwan University, Taipei 106, Taiwan

## ARTICLE INFO

## Article history:

Received 6 March 2010

Revised 20 April 2010

Accepted 27 April 2010

Available online 20 May 2010

## Keywords:

SARS-CoV

Pyrimidines

## ABSTRACT

A series of 2-(benzylthio)-6-oxo-4-phenyl-1,6-dihydropyrimidine as SARS-CoV 3CL protease inhibitors were developed and their potency was evaluated by in vitro protease inhibitory assays. Two candidates had encouraging results for the development of new anti-SARS compounds.

© 2010 Elsevier Ltd. All rights reserved.

Severe acute respiratory syndrome (SARS) has been recognized as a global threat. SARS is characterized by high fever, malaise rigor, headache, chills, cough, and progressive radiographic changes of the chest and lymphopenia.<sup>1–3</sup> The initial outbreak of SARS was first identified in Guangdong Province, China in November 2002. This outbreak spread to several countries and has had significant health and economic impact. The mortality rate is nearly 10%.<sup>4</sup> With rigorous effort by the world health organization (WHO), researchers found that SARS is caused by a novel coronavirus, SARS-CoV.<sup>1,2,5</sup> The SARS-CoV is a positive-strand RNA virus and the genome is ~30 kb (Tor2 strain). The genome is constituted of five major open reading frames namely replicase polyproteins, nucleocapsid proteins, spike (S), envelope (E), and membrane (M) glycoproteins.

Resulting of structural and functional studies of coronaviral life-cycle has provided a number of significant targets for ceasing the viral replication. During the viral replication, the replicase polyprotein undergoes extensive processing by two viral proteases namely, chymotrypsin-like protease (3CL<sup>pro</sup>) and papain-like protease (PL<sup>pro</sup>), reside within the polyprotein. They catalyze their own release from the polyprotein and other non-structural proteins (nsps) from the polyproteins and initiate virus mediated RNA replication.<sup>6</sup> Because of their essential roles in viral replication, both proteases are recognized as attractive targets for development of anti-SARS agents.

To date various SARS-CoV protease inhibitors have been reported from both screened compound libraries and designed com-

pounds based on the substrate structure or active site properties. Their scaffolds are diverse, including C<sub>2</sub>-symmetric diols,<sup>7</sup> 3-quinolinecarboxylic acid derivatives,<sup>8</sup> thiophene-2-carboxylate derivatives,<sup>9</sup> cinanserin,<sup>10</sup> calmodulin,<sup>11</sup> keto-glutamine analogues,<sup>12</sup> anilide,<sup>13</sup> bifunctional boronic acid compounds,<sup>14</sup> isatin derivatives,<sup>15</sup> benzotriazole<sup>16</sup> as well as glutamic acid and glutamine peptides possessing a trifluoromethyl ketone group,<sup>17</sup> α,β-unsaturated esters,<sup>18</sup> and etacrynic acid derivatives.<sup>19</sup> With metal-conjugated structures, some molecules make a coordinate bond with Cys-145 at the active site of SARS-CoV 3CL<sup>pro</sup>.<sup>20</sup> However, no effective therapy has been developed so far and recent isolation of strains of SARS-CoV emphasizes the possibility of a reemergence. Therefore, it is still a great challenge to explore new chemical classes of SARS-CoV 3CL<sup>pro</sup> inhibitors that can be used in anti-SARS therapy in case the disease re-emerges.

In our previous study, from high throughput screening we have identified various heterocycles as novel anti-SARS agents with selective inhibition ranging from IC<sub>50</sub> 2–10 μM against SARS-CoV 3CL protease (compounds **1–4**, Fig. 1).<sup>21,22</sup> In this paper, as a part of our ongoing efforts to delineate a complete pharmacophore model, we designed several 2-(benzylthio)-6-oxo-4-phenyl-1,6-dihydropyrimidine derivatives as anti-SARS agents (compounds **6a–n**, Fig. 1).

From a synthetic point view, the preparation of the target compounds was envisioned following the synthetic routes illustrated in Scheme 1. The synthesis of 6-aryl-5-cyano-2-thiouracils **5a–n** was prepared by reaction between substituted benzaldehyde, ethyl cyanoacetate, and thiourea using the literature procedure.<sup>23</sup> The regioselective S-alkylation of 2-thiouracil **5a–n** achieved by slow addition of the respective halides to a solution of **5a–n** in DMF

\* Corresponding author. Tel.: +886 2 2362 0261x3091; fax: +886 2 2363 5038.

E-mail address: [phliang@gate.sinica.edu.tw](mailto:phliang@gate.sinica.edu.tw) (P.-H. Liang).

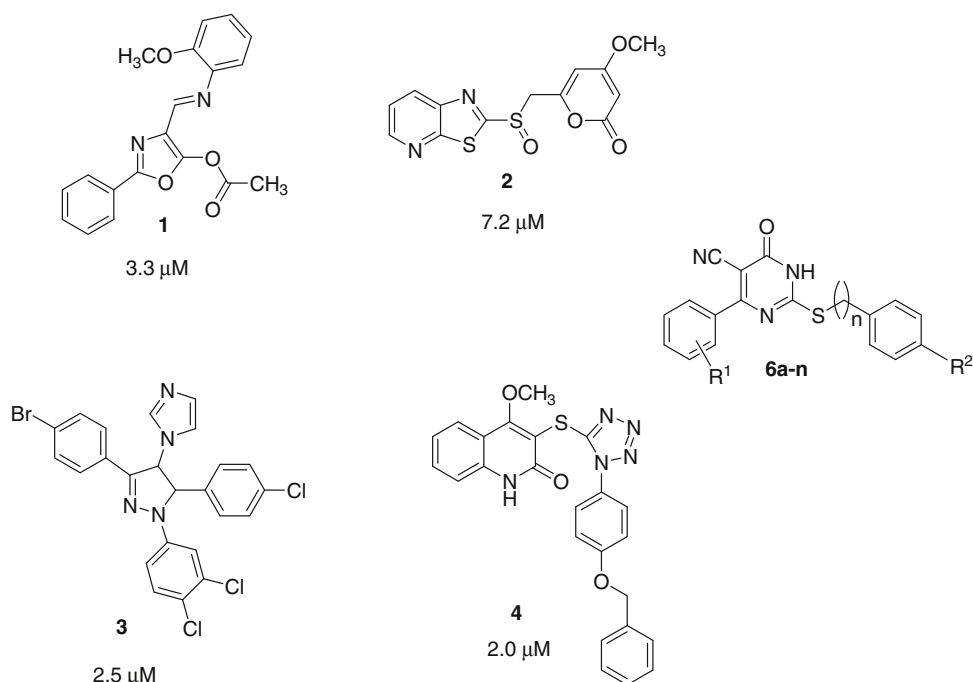
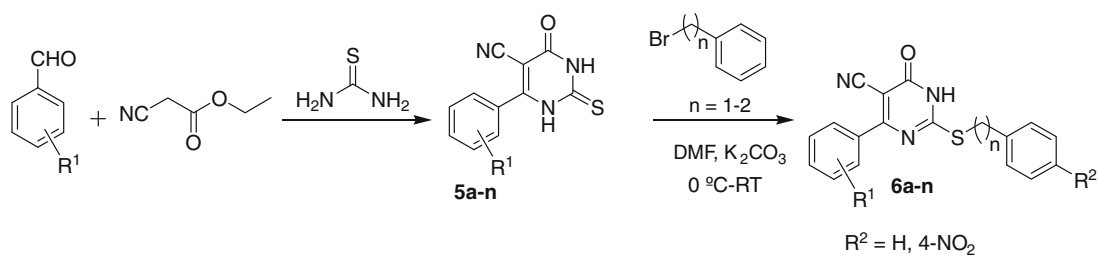


Figure 1. Hits from this study (6) and previous studies<sup>21,22</sup> in our laboratory.



Scheme 1. Synthesis of Inhibitors 6a–n.

using  $K_2CO_3$  as a base at 0–5 °C to yield the corresponding compounds 6a–n.<sup>24</sup> Both analytical and spectral data of all target compounds are accordant with the structures.

The target compounds were tested for anti-SARS activity against SARS-CoV 3CL<sup>pro</sup>, using previously developed assay method<sup>25</sup> containing 0.05  $\mu M$  SARS 3CL<sup>pro</sup>, 6  $\mu M$  fluorogenic substrate Dabcyl-KTSAVLQSGFRKME-Edans, and 50  $\mu M$  of test compounds. Enhanced fluorescence of the reactions in the buffer of 20 mM Bis-Tris at pH 7.0 was monitored at 538 nm with excitation at 355 nm using a fluorescence plate reader. The compounds which inhibited more than 50% of the protease activity at 50  $\mu M$  were selected for the next assay run at 10  $\mu M$  for  $IC_{50}$  calculation. Compound 6m with  $R^2$  group of nitro functionality at C-4 position is the most potent inhibitor with an enzyme inhibitory activity against SARS-CoV 3CL<sup>pro</sup> with an  $IC_{50}$  of 6.1  $\mu M$ . The structure and  $IC_{50}$  values are given in Table 1. The cytotoxicity of the test compounds was tested by performing the MTT assay and found that all compounds are devoid of cytotoxicity.<sup>24</sup>

To obtain molecular insight into the binding properties of these active compounds, we conducted docking studies in the 3CL<sup>pro</sup> active site. For modeling analysis, the crystal structure of SARS 3CL<sup>pro</sup> in complex with a peptide inhibitor (PDB code 1UK4) was used.<sup>26</sup> Docking process was performed using an automated ligand-docking subprogram of the Discovery Studio Modeling 1.2 SBD (Accelrys Inc., San Diego, CA), with a set of parameters chosen to control the precise operation of the genetic algorithm. Docking runs were carried out

Table 1

Structure and activity of compounds 6a–n.

Compound	R <sup>1</sup>	R <sup>2</sup>	n	IC <sub>50</sub> ( $\mu M$ )
6a	H	H	1	>50
6b	H	H	2	>50
6c	H	4-NO <sub>2</sub>	1	35.2
6d	4-OCH <sub>3</sub>	H	1	>50
6e	4-OCH <sub>3</sub>	H	2	20.3
6f	4-OCH <sub>3</sub>	4-NO <sub>2</sub>	1	26.3
6g	4-CH <sub>3</sub>	H	1	>50
6h	4-CH <sub>3</sub>	H	2	>50
6i	4-CH <sub>3</sub>	4-NO <sub>2</sub>	1	>50
6j	3-NO <sub>2</sub>	H	2	>50
6k	3-NO <sub>2</sub>	4-NO <sub>2</sub>	1	10.6±1.2
6l	4-Cl	H	2	16.9±1.3
6m	4-Cl	4-NO <sub>2</sub>	1	6.1±1.1
6n	3-Cl	H	1	>50

using standard default settings 'grid resolution' of 5 Å, 'site opening' of 12 Å, and 'binding site' selected for defining the active site cavity.



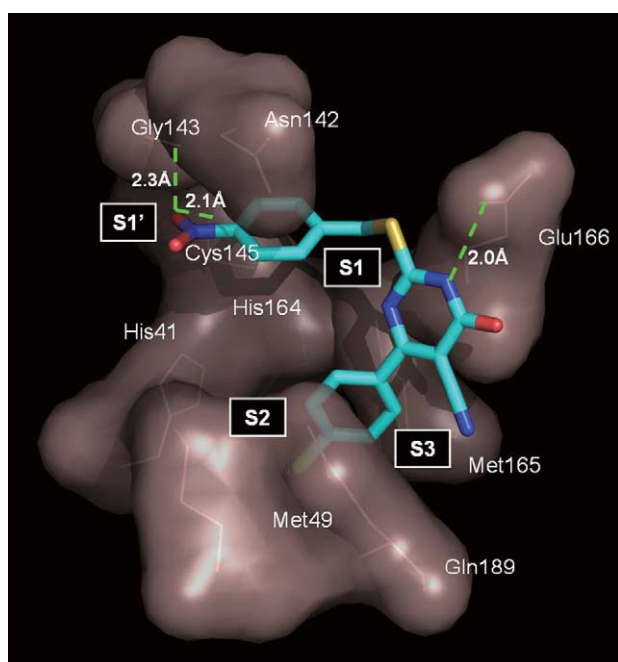


Figure 2. Computer modeling of **6m** binding in the active site of SARS 3CL<sup>pro</sup>.

In search of obtaining a model of the associated complex between the compound **6m** and protein, the distance between the NH of the pyrimidine ring and oxygen atom of Glu-166 was constrained in the distance of 2.0 Å. The orientation of the ligand has the nitro phenyl group situated in the S1 pocket, with the nitro group pointing towards the surface of the protein. One of the oxygen of the nitro group is in close proximity 2.3 Å to the Gly-143 and the other oxygen atom is forming hydrogen bond with Cys-145 at the distance of 2.1 Å (Fig. 2). The chlorophenyl ring fits into the S2 pocket and having hydrophobic interactions with Met-49 and Gln-189. The results of the docking studies presented here suggest that the nitro phenyl group of **6m** can potentially occupy the active site cysteine residue in the enzyme. The oxygen of the nitro group formed a hydrogen bond with the side chain of Gly-143 and Cys-145 that was important for inhibition activity. The compounds lacking nitro functionality in the aryl ring lost the activity. Moderate electron withdrawing substituent R<sup>1</sup> like chloro in the compounds **6l** and **6m** favors the inhibitory activity when compared to the electron donating groups like methyl and methoxy in the compound (see Table 1). This result suggests that the substituent R<sup>1</sup> can be electron withdrawing group to increase the inhibitory action.

In conclusion, we disclosed the inhibitory potency of compound **6m**, containing pyrimidine unit, as a SARS-CoV 3CL<sup>pro</sup> inhibitor. The measured inhibitory activity coupled with possible structure modifications revealed by 3D docking give us new directions for a fast development of much more potent inhibitors. Further investigations on this new family of compounds are currently in progress in our laboratory.

## References and notes

- Ksiazek, T. G.; Erdman, D.; Goldsmith, C. S.; Zaki, S. R.; Peret, T.; Emery, S.; Tong, S.; Urbani, C.; Comer, J. A.; Lim, W.; Rollin, P. E.; Dowell, S. F.; Ling, A.-E.; Humphrey, C. D.; Shieh, W.-J.; Guarner, J.; Paddock, C. D.; Rota, P.; Fields, B.; DeRisi, J.; Yang, J.-Y.; Cox, N.; Hughes, J. M.; LeDuc, J. W.; Bellini, W. J.; Anderson, L. J. *N. Engl. J. Med.* **2003**, *348*, 1953.
- Peiris, J. S. M.; Lai, S.-T.; Poon, L. L.-M.; Guan, Y.; Yam, L. Y.-C.; Lim, W.; Nicholls, J.; Yee, W. K.-S.; Yan, W. W.; Cheung, M.-T.; Cheng, V. C.-C.; Chan, K.-H.; Tsang, D. N.-C.; Yung, R. W.-H.; Ng, T. K.; Yuen, K.-Y. *Lancet* **2003**, *361*, 1319.
- Rota, P. A.; Oberste, M. S.; Monroe, S. S.; Nix, W. A.; Campagnoli, R.; Icenogle, J. P.; Penaranda, S.; Bankamp, B.; Maher, K.; Chen, M.-H.; Tong, S.; Tamin, A.; Lowe, L.; Frace, M.; DeRisi, J. L.; Chen, Q.; Wang, D.; Erdman, D. D.; Peret, T. C. T.; Burns, C.; Ksiazek, T. G.; Rollin, P. E.; Sanchez, A.; Liffick, S.; Holloway, B.; Limor, J.; McCaustland, K.; Olsen-Rasmussen, M.; Fouchier, R.; Gunther, S.; Osterhaus, A. D. M. E.; Drosten, C.; Pallansch, M. A.; Anderson, L. J.; Bellini, W. J. *Science* **2003**, *300*, 1394.
- He, J.-F.; Peng, G.-W.; Min, J.; Yu, D.-W.; Liang, W.-L.; Zhang, S.-Y.; Xu, R.-H.; Zheng, H.-Y.; Wu, X.-W.; Xu, J.; Wang, Z.-H.; Fang, L.; Zhang, X.; Li, H.; Yan, X.-G.; Lu, J.-H.; Hu, Z.-H.; Huang, J.-C.; Wan, Z.-Y.; Hou, J.-L.; Lin, J.-Y.; Song, H.-D.; Wang, S.-Y.; Zhou, X.-J.; Zhang, G.-W.; Gu, B.-W.; Zheng, H.-J.; Zhang, X.-L.; He, M.; Zheng, K.; Wang, B.-F.; Fu, G.; Wang, X.-N.; Chen, S.-J.; Chen, Z.; Hao, P.; Tang, H.; Ren, S.-X.; Zhong, Y.; Guo, Z.-M.; Liu, Q.; Miao, Y.-G.; Kong, X.-Y.; He, W.-Z.; Li, Y.-X.; Wu, C.-I.; Zhao, G.-P.; Chiu, R. W. K.; Chim, S. S. C.; Tong, Y.-K.; Chan, P. K. S.; Tam, J. S.; Lo, Y. M. D. *Science* **2004**, *303*, 1666.
- Drosten, C.; Gunther, S.; Preiser, W.; van der Werf, S.; Brodt, H. R.; Becker, S.; Rabenau, H.; Panning, M.; Kolesnikova, L.; Fouchier, R. A.; Berger, A.; Burguiere, A. M.; Cinatl, J.; Eickmann, M.; Escriviou, N.; Grywna, K.; Kramme, S.; Manuguerra, J. C.; Muller, S.; Rickerts, V.; Sturmer, M.; Vieth, S.; Klenk, H. D.; Osterhaus, A. D.; Schmitz, H.; Doerr, H. W. *N. Engl. J. Med.* **2003**, *348*, 1967.
- Baker, S. C. *Coronaviruses: Molecular Biology*, 3rd ed. In *Encyclopedia of Virology*; Academic Press: New York, 2008; Vol. 1, pp 554–562.
- (a) Wu, C.-Y.; Jan, J.-T.; Ma, S.-H.; Kuo, C.-J.; Juan, H.-F.; Cheng, E. Y.-S.; Hsu, H.-H.; Huang, H.-C.; Wu, D.; Brik, A.; Liang, F.-S.; Liu, R.-S.; Fang, J.-M.; Chen, S.-T.; Liang, P.-H.; Wong, C.-H. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 10012; (b) Shao, Y.-M.; Yang, W.-B.; Peng, H.-P.; Hsu, M.-F.; Tsai, K.-C.; Kuo, T.-H.; Wang, A. H.-J.; Liang, P.-H.; Lin, C.-H.; Yang, A.-S.; Wong, C.-H. *ChemBioChem* **2007**, *8*, 1654.
- Kao, R. Y.; Tsui, W. H. W.; Lee, T. S. W.; Tanner, J. A.; Watt, R. M.; Huang, J. D.; Hu, L. H.; Chen, G. H.; Chen, Z. W.; Zhang, L. Q.; He, T.; Chan, K. H.; Tse, H.; To, A. P. C.; Ng, L. W. Y.; Wong, B. C. W.; Tsoi, H. W.; Yang, D.; Ho, D. D.; Yuen, K. Y. *Chem. Biol.* **2004**, *11*, 1293.
- Blanchard, J. E.; Elowe, N. H.; Huitema, C.; Fortin, P. D.; Cechetto, J. D.; Eltis, L. D.; Brown, E. D. *Chem. Biol.* **2004**, *11*, 1445–1453.
- (a) Chen, L. L.; Gui, C. S.; Luo, X. M.; Yang, Q. G.; Gunther, S.; Scandella, E.; Drosten, C.; Bai, D.; He, X. C.; Ludewig, B.; Chen, J.; Luo, H. B.; Yang, Y. M.; Yang, Y. F.; Zou, J. P.; Thiel, V.; Chen, K.; Shen, J. H.; Xu, S.; Jiang, H. L. *J. Virol.* **2005**, *79*, 7095; (b) Yang, Q.; Chen, L.; He, X.; Gao, Z.; Shen, X.; Bai, D. *Chem. Pharm. Bull.* **2008**, *56*, 1400.
- Liu, Z.; Huang, C.; Fan, K.; Wei, P.; Chen, H.; Liu, S.; Pei, J.; Shi, L.; Li, B.; Yang, K.; Liu, Y.; Lai, L. *J. Chem. Inf. Model.* **2005**, *45*, 10–17.
- Jain, R. P.; Petterson, H. I.; Zhang, J.; Aull, K. D.; Fortin, P. D.; Huitema, C.; Eltis, L. D.; Parrish, J. C.; James, M. N. G.; Wishart, D. S.; Vederas, J. C. *J. Med. Chem.* **2004**, *47*, 6113.
- Shie, J.-J.; Fang, J.-M.; Kuo, T.-H.; Kuo, C.-J.; Liang, P.-H.; Huang, H.-J.; Yang, W.-B.; Lin, C.-H.; Chen, J.-L.; Wu, Y.-T.; Wong, C.-H. *J. Med. Chem.* **2005**, *48*, 4469.
- Bacha, U.; Barrila, J.; Velasquez-Campoy, A.; Leavitt, S. A.; Freire, E. *Biochemistry* **2004**, *43*, 4906.
- Chen, L.-R.; Wang, Y.-C.; Lin, Y.-W.; Chou, S.-Y.; Chen, S.-F.; Liu, L.-T.; Wu, Y.-T.; Kuo, C.-J.; Chen, T. S.-S.; Juang, S.-H. *Bioorg. Med. Chem. Lett.* **2005**, *15*, 3058.
- Wu, C.-Y.; King, K.-Y.; Kuo, C.-J.; Fang, J.-M.; Wu, Y.-T.; Ho, M.-Y.; Liao, C.-L.; Shie, J.-J.; Liang, P.-H.; Wong, C.-H. *Chem. Biol.* **2006**, *13*, 4469.
- (a) Bacha, U.; Barrila, J.; Gabelli, B.; Kiso, Y.; Amzel, L. M.; Freire, E. *Chem. Biol. Drug Des.* **2008**, *72*, 34; (b) Regnier, T.; Sarma, D.; Hidaka, K.; Bacha, U.; Freire, E.; Hayashi, Y.; Kiso, Y. *Bioorg. Med. Chem. Lett.* **2009**, *19*, 2722.
- (a) Ghosh, A. K.; Xi, K.; Ratia, K.; Santarsiero, B. D.; Fu, W.; Harcourt, B. H.; Rota, P. A.; Baker, S. C.; Johnson, M. E.; Mesecar, A. D. *J. Med. Chem.* **2005**, *48*, 6767; (b) Shie, J.-J.; Fang, J.-M.; Kuo, T.-H.; Kuo, C.-J.; Liang, P.-H.; Huang, H.-J.; Wu, Y.-T.; Jan, J.-T.; Cheng, E. Y.-S.; Wong, C.-H. *Bioorg. Med. Chem.* **2005**, *13*, 5240; (c) Ghosh, A. K.; Xi, K.; Grum-Tokars, V.; Xu, X.; Ratia, K.; Fu, W.; Houser, K. V.; Baker, S. C.; Johnson, M. E.; Mesecar, A. D. *Bioorg. Med. Chem. Lett.* **2007**, *17*, 5876.
- Kaeppler, U.; Stiefl, N.; Schiller, M.; Vicik, B.; Breuning, A.; Schmitz, W.; Rupprecht, D.; Schmuck, C.; Baumann, K.; Ziebuhr, J.; Schirmeister, T. *A. J. Med. Chem.* **2005**, *48*, 6832.
- Hsu, J. T. A.; Kuo, C. J.; Hsieh, H. P.; Wang, Y. C.; Huang, K. K.; Lin, C. P. C.; Huang, P. F.; Chen, X.; Liang, P. H. *FEBS Lett.* **2004**, *574*, 116.
- Kuo, C. J.; Liu, H. G.; Lo, Y. K.; Seong, C. M.; Lee, K. I.; Jung, Y. S.; Liang, P. H. *FEBS Lett.* **2009**, *583*, 549.
- Ahn, T. Y.; Kuo, C. J.; Liu, H. G.; Ha, D. C.; Liang, P. H.; Jung, Y. S. *Bull. Korean Chem. Soc.* **2010**, *31*, 87.
- Ram, V. J.; Vanden Berghe, D. A.; Vlietinck, A. J. *Heterocycl. Chem.* **1984**, *21*, 307.
- Ramajayam, R.; Mahera, N. B.; Neamati, N.; Yadav, M. R.; Giridhar, R.; *Arch. Pharm.* **2009**, *342*, 710. *General procedure for the synthesis of compound 6a–n*: To a mixture of 6-aryl-5-cyano-2-thiouracil (1 mmol) and K<sub>2</sub>CO<sub>3</sub> (1.5 mmol) in DMF (10 mL), alkyl iodide (1.2 mmol) was added dropwise with stirring while maintaining the temperature of the reaction mixture at 0–5 °C. Stirring was continued for 3 h at this temperature and continued for additional 2 h at room temperature. Water was added to the mixture and filtered. The aqueous filtrate was neutralized with acetic acid and the precipitate was filtered and purified. The selected data of representative compounds **6k** and **6m** were as follows: *Compound 6k*: Yield: 51%; mp 236–237 °C; IR (KBr): 3078, 2221, 1666, 1521, 1473, 1346, 1249, 1112, 1010, 916, 891, 785 cm<sup>-1</sup>; <sup>1</sup>H NMR (400 MHz, DMSO-d<sub>6</sub>) δ: 4.59 (s, 2H, CH<sub>2</sub>), 7.63–7.65 (d, J = 8.72 Hz, 2H, ArH), 7.79 (t, J = 8.04 Hz, 1H, ArH), 8.17–8.20 (d, J = 8.72 Hz, 2H, ArH), 8.31–8.33 (d, J = 8.0 Hz, 1H, ArH), 8.41–8.44 (m, 1H, ArH), 8.69–8.70 (m, 1H, ArH), 12.9 (br, 1H, NH); MS (CI) m/z: 410 [M+H]<sup>+</sup>. Anal. Calcd for C<sub>18</sub>H<sub>11</sub>N<sub>5</sub>O<sub>5</sub>S: C, 52.81; H, 2.71; N, 17.11. Found: C, 52.68; H, 2.85; N, 17.04.

**Compound 6m**: Yield: 66%; mp 254–257 °C; IR (KBr): 3000, 2218, 1651, 1517, 1467, 1346, 1244, 1009, 1004, 997, 856, 779  $\text{cm}^{-1}$ ;  $^1\text{H}$  NMR (400 MHz, DMSO- $d_6$ ):  $\delta$ : 4.60 (s, 2H,  $\text{CH}_2$ ), 7.51–7.53 (d,  $J = 8.64$  Hz, 2H, ArH), 7.60–7.62 (d,  $J = 8.76$  Hz, 2H, ArH), 7.89–7.91 (d,  $J = 8.64$  Hz, 2H, ArH), 8.13–8.15 (d,  $J = 8.72$  Hz, 2H, ArH), 12.88 (br, 1H, NH); MS (CI)  $m/z$ : 399  $[\text{M}+\text{H}]^+$ . Anal. Calcd for  $\text{C}_{18}\text{H}_{11}\text{ClN}_4\text{O}_3\text{S}$ : C, 54.21; H, 2.78; N, 14.05. Found: C, 54.18; H, 2.85; N, 14.01.

25. Kuo, C. J.; Chi, Y. H.; Hsu, J. T. A.; Liang, P. H. *Biochem. Biophys. Res. Commun.* **2004**, 318, 862.
26. Yang, H.; Yang, M.; Ding, Y.; Liu, Y.; Lou, Z.; Zhou, Z.; Sun, L.; Mo, L.; Ye, S.; Pang, H.; Gao, G. F.; Anard, K.; Bartlam, M.; Hilgenfeld, R.; Rao, Z. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, 100, 13190.





## Synthesis and evaluation of pyrazolone compounds as SARS-coronavirus 3C-like protease inhibitors

R. Ramajayam<sup>a</sup>, Kian-Pin Tan<sup>b</sup>, Hun-Ge Liu<sup>b</sup>, Po-Huang Liang<sup>a,b,\*</sup>

<sup>a</sup>Institute of Biological Chemistry, Academia Sinica, 128 Academia Road, Taipei 11529, Taiwan

<sup>b</sup>Institute of Biochemical Sciences, National Taiwan University, Taipei 106, Taiwan

### ARTICLE INFO

#### Article history:

Received 10 August 2010

Revised 20 September 2010

Accepted 21 September 2010

Available online 25 September 2010

#### Keywords:

Pyrazolone

3CL protease

SARS-CoV

Coxsackievirus

Computer modeling

### ABSTRACT

A series of pyrazolone compounds as possible SARS-CoV 3CL protease inhibitors were designed, synthesized, and evaluated by in vitro protease assay using fluorogenic substrate peptide in which several showed potent inhibition against the 3CL protease. Interestingly, one of the inhibitors was also active against 3C protease from coxsackievirus B3. These inhibitors could be potentially developed into anti-coronaviral and anti-picornaviral agents.

© 2010 Elsevier Ltd. All rights reserved.

### 1. Introduction

Severe acute respiratory syndrome (SARS) is a newly emerging infectious disease caused by a novel coronavirus, SARS-coronavirus (SARS-CoV). SARS has been recognized as a global threat since the initial outbreak of SARS first identified in Guangdong Province, China in November 2002. This outbreak spread to several countries and has had significant health and economic impact. SARS is a life-threatening form of atypical pneumonia characterized by high fever, malaise rigor, headache, chills, cough, and progressive radiographic changes of the chest and lymphopenia.<sup>1–3</sup> The mortality rate is nearly 10%.<sup>4</sup> The SARS-CoV is a positive-strand RNA virus that uses a complex set of enzymes to replicate the largest RNA genomes currently known for RNA viruses and synthesize an extensive set of 5' leader-containing subgenomic mRNAs that encode the viral structural proteins and several species-specific proteins with unknown functions. These processes are mediated primarily by the 3C-like protease (3CL<sup>PRO</sup>) with chymotrypsin fold. The active site of SARS-CoV 3CL<sup>PRO</sup> contains Cys145 and His41 to constitute a catalytic dyad, in which cysteine functions as the common nucleophile in the proteolytic process.<sup>5,6</sup> Because of the essential role in viral processing, the 3CL<sup>PRO</sup> is considered as an attractive target for anti-SARS and other coronavirus infections. 3CL<sup>PRO</sup> is named after the 3C proteases (3C<sup>PRO</sup>) from picornaviruses such as enterovirus (EV), coxsackievirus (CV), and rhinovirus (RV)

which cause life-threatening infectious diseases. The 3C<sup>PRO</sup> essential for viral replication had served as a drug target.<sup>7</sup> Both 3C<sup>PRO</sup> and 3CL<sup>PRO</sup> have similar 3-D structures, but unlike the dimeric 3CL<sup>PRO</sup>, 3C<sup>PRO</sup> is monomeric and utilizes Glu-His-Cys triad for catalysis.<sup>8</sup>

To date various SARS-CoV protease inhibitors have been reported from both screened compound libraries and designed compounds based on the substrate structure or active site properties. Their scaffolds are diverse, including C<sub>2</sub>-symmetric diols,<sup>9</sup> 3-quinolinecarboxylic acid derivatives,<sup>10</sup> thiophene-2-carboxylate derivatives,<sup>11</sup> cinanserin,<sup>12</sup> calmodulin,<sup>13</sup> keto-glutamine analogs,<sup>14</sup> anilide,<sup>15</sup> bifunctional boronic acid compounds,<sup>16</sup> isatin derivatives,<sup>17</sup> pyrimidinone,<sup>18</sup> benzotriazole<sup>19</sup> as well as glutamic acid and glutamine peptides possessing a trifluoromethyl ketone group,<sup>20</sup> α,β-unsaturated esters,<sup>21</sup> and etacrynic acid derivatives.<sup>22</sup> With metal-coordinated structures, some molecules make a covalent bond with Cys-145 at the active site of SARS-CoV 3CL<sup>PRO</sup>.<sup>23</sup> However, no effective therapy has been developed so far and recent isolation of strains of SARS-CoV emphasizes the possibility of a reemergence. Therefore, it is still a great challenge to explore new chemical classes of SARS-CoV 3CL<sup>PRO</sup> inhibitors that can be used in anti-SARS therapy in case the disease re-emerges.

Compounds containing a pyrazole and its related analogs have received significant attention in chemical, medicinal, and pharmaceutical research as this structural scaffold is found in a variety of drugs. As shown in Figure 1, a new pyrazolone compound, edaravone (**A**), also known as MCI-186, has been developed as a medical drug for brain ischemia<sup>24</sup> and has also been reported to be effective for myocardial ischemia.<sup>25</sup> Compound (**B**) is claimed

\* Corresponding author. Tel.: +886 2 2362 0261x3091; fax: +886 2 2363 5038.

E-mail address: [phliang@gate.sinica.edu.tw](mailto:phliang@gate.sinica.edu.tw) (P.-H. Liang).

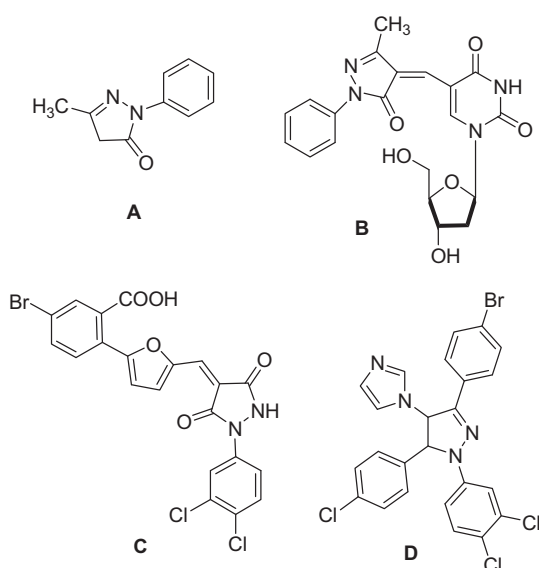


Figure 1. Pyrazole compounds as drugs or enzyme inhibitors.

to have potent anti-orthopoxvirus activity.<sup>26</sup> Compound (**C**) 3,5-dioxypyrazolidine has been reported as SARS-CoV 3CL<sup>pro</sup> inhibitor.<sup>27</sup> Recently, we identified from high throughput screening certain pyrazolines, particularly those displaying a 1,3,5-triaryl substitution pattern (**D**), were active against SARS-CoV 3CL<sup>pro</sup>, CoV-229E 3CL<sup>pro</sup>, CVB3 3C<sup>pro</sup>, EV71 3C<sup>pro</sup>, and RV14 3C<sup>pro</sup>.<sup>28</sup> In the present study, we synthesized the pyrazolone compounds as SARS 3CL protease inhibitors and explored their structure–activity relationship (SAR) in inhibiting 3CL<sup>pro</sup> and 3C<sup>pro</sup>.

In view of the facts mentioned above, 21 compounds containing the pyrazolone template were synthesized and screened for their 3C and 3CL protease inhibitory activities. Some of the synthesized compounds displayed potent inhibition against CVB3 3C<sup>pro</sup> and SARS-CoV 3CL<sup>pro</sup>. The synthesis of the target compounds **2a–u** was envisioned following the synthetic route illustrated in Scheme 1.

## 2. Results and discussion

Compounds **1a–k** were synthesized by refluxing the corresponding  $\beta$ -ketoester and the substituted phenylhydrazine hydrochloride in acetic acid.<sup>29</sup> The pyrazolones were treated with the appropriate aromatic aldehyde in presence of piperidine in ethanol to obtain target compounds **2a–u** in 70–87% yields.<sup>30</sup>

From the preliminary investigation, as summarized in Table 1, we noted that compounds with substituent R<sup>2</sup>, carboxyl group at 4th position in benzylidene aryl ring shows significant inhibition against SARS-CoV 3CL<sup>pro</sup>. Compounds having R<sup>1</sup> substitution like halogens, cyano, and nitro group increase the inhibitory action

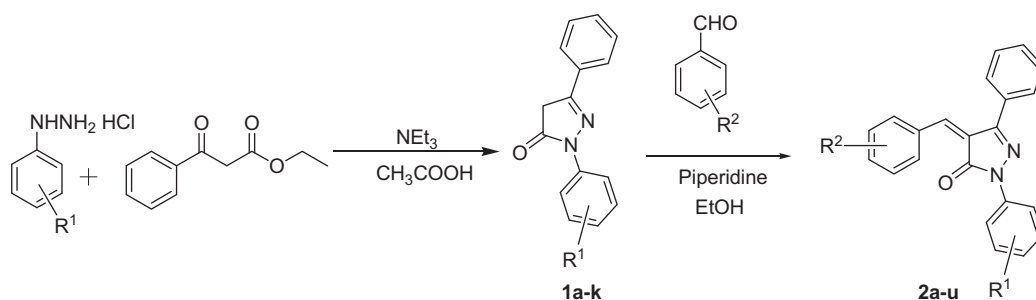
Table 1  
Structure and IC<sub>50</sub> ( $\mu$ M) of compounds **2a–u**

2a–u				
Compd	R <sup>1</sup>	R <sup>2</sup>	SARS ( $\mu$ M)	CVB3 ( $\mu$ M)
<b>2a</b>	H	H	N.I.	N.I.
<b>2b</b>	H	3-OCH <sub>3</sub>	N.I.	N.I.
<b>2c</b>	H	4-NHCOCH <sub>3</sub>	N.I.	N.I.
<b>2d</b>	H	4-COOH	18.0	51.1
<b>2e</b>	H	4-N(CH <sub>3</sub> ) <sub>2</sub>	N.I.	N.I.
<b>2f</b>	H	3-NO <sub>2</sub>	N.I.	N.I.
<b>2g</b>	4-Cl	H	N.I.	N.I.
<b>2h</b>	4-Cl	4-Cl	N.I.	N.I.
<b>2i</b>	4-Cl	4-COOH	13.9	25.0
<b>2j</b>	4-Cl	4-NHCOCH <sub>3</sub>	N.I.	N.I.
<b>2k</b>	4-Cl	4-OCH <sub>3</sub>	N.I.	N.I.
<b>2l</b>	4-Cl	4-OH	N.I.	N.I.
<b>2m</b>	4-OCH <sub>3</sub>	4-COOH	12.0	16.6
<b>2n</b>	4-CH(CH <sub>3</sub> ) <sub>2</sub>	4-COOH	N.I.	N.I.
<b>2o</b>	4-C(CH <sub>3</sub> ) <sub>3</sub>	4-COOH	N.I.	N.I.
<b>2p</b>	4-CN	4-COOH	5.5	20.8
<b>2q</b>	4-OCF <sub>3</sub>	4-COOH	42.0	98.8
<b>2r</b>	3-Cl	4-COOH	10.8	17.3
<b>2s</b>	3,4-Cl <sub>2</sub>	4-COOH	24.3	125.5
<b>2t</b>	4-F	4-COOH	6.8	22.4
<b>2u</b>	3-NO <sub>2</sub>	4-COOH	8.4	9.6

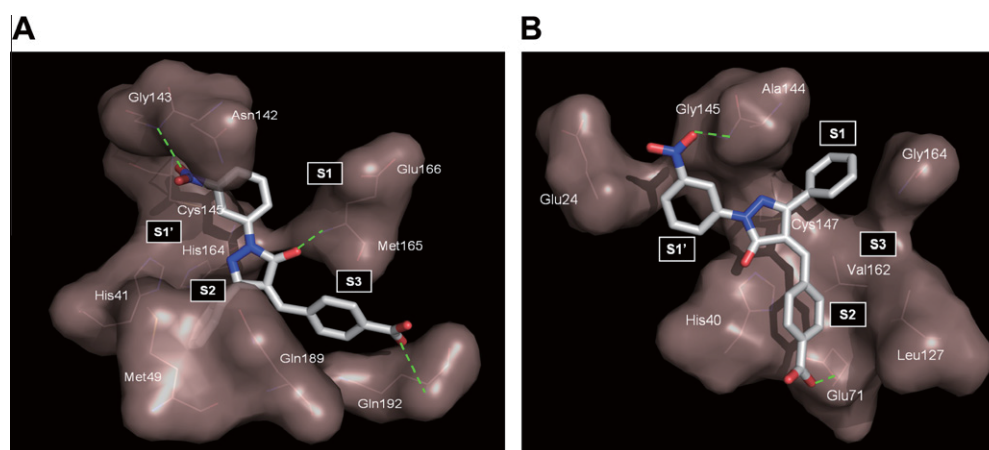
N.I.: no inhibition at 50  $\mu$ M.

(see Table 1). Compound **2p** is the most potent inhibitor showing an IC<sub>50</sub> of 5.5  $\mu$ M and **2t** is the second with IC<sub>50</sub> of 6.8  $\mu$ M against SARS-CoV 3CL<sup>pro</sup>. Interestingly, **2u** showed inhibitory activity significantly against both SARS-CoV 3CL<sup>pro</sup> (IC<sub>50</sub> = 8.4  $\mu$ M) and CVB3 3C<sup>pro</sup> (IC<sub>50</sub> = 9.6  $\mu$ M). The cytotoxicity of the test compounds was tested by performing the MTT assay and found that all compounds are devoid of cytotoxicity at 200  $\mu$ M.

In search of a computer model of the associated complex between the compound **2u** and the proteases to rationalize its inhibitory activities, the orientation of the ligand has the N1-phenyl group situated in the S1' pocket of the 3CL<sup>pro</sup>. One of the oxygen of the nitro group is in close proximity 2.7 Å and forms H-bond to the Gly-143 (Fig. 2A). The C=O in the central pyrazolone ring is close to Glu-166 with the distance of 3.0 Å to form a H-bond. C-3 phenyl ring fits into the S2 pocket, having hydrophobic interactions with Met-49, Arg-188, and Gln-189 (hiding behind these residues in Fig. 2A). The carboxyl benzylidene group is situated in the S3 pocket of the 3CL<sup>pro</sup>. The oxygen of the carboxyl group forms a hydrogen bond with the side chain of Gln-192 at a distance of



Scheme 1. General synthesis of compounds **2a–u**.



**Figure 2.** Docking studies of **2u** binding in the active site of SARS 3CL<sup>pro</sup> (A) and CVB3 3C<sup>pro</sup> (B).

3.2 Å. It is important for inhibition activity since the compounds lacking carboxy functionality in the benzylidene lost the activity. Electron withdrawing R<sup>1</sup> substituents like cyano (**2p**), fluoro (**2t**), and nitro (**2u**), accompanied with R<sup>2</sup> carboxyl group favors the inhibitory activity.

In further evaluating the inhibitors against CVB3 3C<sup>pro</sup>, we found **2p** and **2t** were moderate inhibitors against CVB3 3C<sup>pro</sup> (IC<sub>50</sub> = 20.8 and 22.4 μM, respectively), but **2u** was more active against CVB3 3C<sup>pro</sup> (IC<sub>50</sub> = 9.6 μM). According to the modeling shown in Figure 2B, the R<sup>1</sup> nitro group of **2u** forms H-bond with Gly-145 (**2p** and **2t** without nitro group fail to form such a H-bond) and benzylidene carboxylate of **2u** is H-bonded to Glu-71 in the active site of 3C<sup>pro</sup>. It was predicted that the C-3 phenyl ring of **2u** is pointed to S1 site and the carboxyl benzylidene group is relocated to S2 in order to form the H-bond in 3C<sup>pro</sup> due to the subtle differences between the structures of 3CL<sup>pro</sup> and 3C<sup>pro</sup>.<sup>7</sup> However, it should be noted that computer modeling is speculation based on energy minimization to fit the SAR data. In conclusion, **2p** and **2t** are selective against 3CL<sup>pro</sup>, but **2u** is a common inhibitor of 3CL<sup>pro</sup> and 3C<sup>pro</sup>, which may be potentially developed into anti-coronaviral and anti-picornaviral drugs.

### 3. Conclusion

As reported here, pyrazolone compounds (**2p**, **2t**, and **2u**) with a 4-carboxylbenzylidene aryl ring attaching to C4 of pyrazolone showed potent 3CL<sup>pro</sup> inhibition, while 3-nitro-phenyl group attached to N1 atom (**2u**) gave the simultaneously inhibitory activity against 3C<sup>pro</sup> from CVB3.

## 4. Experimental

### 4.1. Chemistry

#### 4.1.1. General

All chemicals (reagent grade) used were purchased from Sigma-Aldrich (USA) and Acros organics Co., Ltd (USA). ESITOF-MS spectra were recorded on a Bruker BioTOF II mass spectrometer and <sup>1</sup>H NMR spectra were recorded on a AV-400 or AV-500 spectrometer at 25 °C with TMS as an internal standard. Chemical shifts (δ) are reported in ppm and were adjusted relative to the residual solvent peak.

#### 4.1.2. General procedure for the preparation of compounds 1a–k

An equimolar solution of ethyl benzoylacetate and substituted phenylhydrazine hydrochloride was treated with triethylamine.

The mixture was stirred at reflux temperature for 20 h. The solvent was removed by evaporation, and the residue was extracted with AcOEt. The organic phase was dried over anhydrous Na<sub>2</sub>SO<sub>4</sub> and evaporated under reduced pressure to get crude solid. The crude product was recrystallised from methanol to yield pure pyrazolone 1a–k. The spectroscopic data of the 1a–k are described below in details.

**4.1.2.1. 1,3-Diphenyl-5-pyrazol-5(4H)-one (1a).** 89% yield, mp: 139–141 °C. <sup>1</sup>H NMR (CDCl<sub>3</sub>, 500 MHz) δ 3.86 (s, 2H), 7.23–7.26 (m, 1H), 7.44–7.49 (m, 5H), 7.80 (d, *J* = 8.0 Hz, 2H), 8.01 (d, *J* = 7.9 Hz, 2H); ESI-TOF-MS: 237.10 (C<sub>15</sub>H<sub>13</sub>N<sub>2</sub>O, [M+H]<sup>+</sup>).

**4.1.2.2. 3-Phenyl-1-(4-chlorophenyl)pyrazol-5(4H)-one (1b).** 82% yield, mp: 162–164 °C. <sup>1</sup>H NMR (CDCl<sub>3</sub>, 500 MHz) δ 3.88 (s, 2H), 7.41 (d, *J* = 8.8 Hz, 2H), 7.49 (m, 3H), 7.78–7.80 (m, 2H), 7.99 (d, *J* = 8.8 Hz, 2H); ESI-TOF-MS: 271.06 (C<sub>15</sub>H<sub>12</sub>ClN<sub>2</sub>O, [M+H]<sup>+</sup>).

**4.1.2.3. 3-Phenyl-1-(4-methoxyphenyl)pyrazol-5(4H)-one (1c).** 77% yield, mp: 127–129 °C. <sup>1</sup>H NMR (DMSO-*d*<sub>6</sub>, 500 MHz) δ 3.75 (s, 3H), 3.86 (s, 2H), 6.88 (d, *J* = 8.15 Hz, 2H), 7.19–7.22 (m, 1H), 7.29–7.31 (m, 2H), 7.64 (d, *J* = 8.3 Hz, 2H), 7.70–7.74 (m, 2H); ESI-TOF-MS: 267.11 (C<sub>16</sub>H<sub>15</sub>N<sub>2</sub>O<sub>2</sub>, [M+H]<sup>+</sup>).

**4.1.2.4. 3-Phenyl-1-(4-isopropylphenyl)pyrazol-5(4H)-one (1d).** 81% yield, mp: 132–133 °C. <sup>1</sup>H NMR (CDCl<sub>3</sub>, 500 MHz) δ 1.29 (s, 6H), 2.93–2.99 (m, 1H), 3.85 (s, 2H), 7.31 (d, *J* = 8.4 Hz, 2H), 7.47–7.48 (m, 3H), 7.79 (d, 2H), 7.88 (d, *J* = 8.4 Hz, 2H); ESI-TOF-MS: 279.15 (C<sub>18</sub>H<sub>19</sub>N<sub>2</sub>O, [M+H]<sup>+</sup>).

**4.1.2.5. 3-Phenyl-1-(4-*tert*-butylphenyl)pyrazol-5(4H)-one (1e).** 78% yield, mp: 120–122 °C. <sup>1</sup>H NMR (CDCl<sub>3</sub>, 500 MHz) δ 1.35 (s, 9H), 3.86 (s, 2H), 7.46–7.48 (m, 5H), 7.80 (d, *J* = 8.6 Hz, 2H), 7.88 (d, *J* = 8.6 Hz, 2H); ESI-TOF-MS: 293.16 (C<sub>19</sub>H<sub>21</sub>N<sub>2</sub>O, [M+H]<sup>+</sup>).

**4.1.2.6. 3-Phenyl-1-(4-cyanophenyl)pyrazol-5(4H)-one (1f).** 73% yield, mp: 211–212 °C. <sup>1</sup>H NMR (CDCl<sub>3</sub>, 500 MHz) δ 3.93 (s, 2H), 7.51–7.52 (m, 3H), 7.74 (d, *J* = 8.75 Hz, 2H), 7.80–7.82 (m, 2H), 8.22 (d, *J* = 8.75 Hz, 2H); ESI-TOF-MS: 262.10 (C<sub>16</sub>H<sub>12</sub>N<sub>3</sub>O, [M+H]<sup>+</sup>).

**4.1.2.7. 3-Phenyl-1-(4-trifluoromethoxyphenyl)pyrazol-5(4H)-one (1g).** 65% yield, mp: 114–116 °C. <sup>1</sup>H NMR (CDCl<sub>3</sub>, 500 MHz) δ 3.90 (s, 2H), 7.30 (d, *J* = 8.5 Hz, 2H), 7.49–7.50 (m, 3H), 7.79–7.80 (m, 2H), 8.07 (d, *J* = 8.95 Hz, 2H); ESI-TOF-MS: 321.08 (C<sub>16</sub>H<sub>12</sub>F<sub>3</sub>N<sub>2</sub>O<sub>2</sub>, [M+H]<sup>+</sup>).

**4.1.2.8. 3-Phenyl-1-(3-chlorophenyl)pyrazol-5(4H)-one (1h).** 61% yield, mp: 101–103 °C. <sup>1</sup>H NMR (CDCl<sub>3</sub>, 500 MHz) δ 3.89 (s, 2H), 7.41–7.43 (m, 3H), 7.47–7.50 (m, 1H), 7.56–7.58 (m, 2H), 7.61 (d, *J* = 7.8 Hz, 1H), 7.63 (d, *J* = 8.1 Hz, 1H), 7.91 (s, 1H); ESI-TOF-MS: 271.06 (C<sub>15</sub>H<sub>12</sub>ClN<sub>2</sub>O, [M+H]<sup>+</sup>).

**4.1.2.9. 3-Phenyl-1-(3,4-dichlorophenyl)pyrazol-5(4H)-one (1i).** 71% yield, mp: 139–141 °C. <sup>1</sup>H NMR (CDCl<sub>3</sub>, 500 MHz) δ 3.89 (s, 2H), 7.49–7.50 (m, 4H), 7.79–7.81 (m, 2H), 7.95–7.98 (m, 1H), 8.20 (d, 1H); ESI-TOF-MS: 305.02 (C<sub>15</sub>H<sub>11</sub>Cl<sub>2</sub>N<sub>2</sub>O, [M+H]<sup>+</sup>).

**4.1.2.10. 3-Phenyl-1-(4-fluorophenyl)pyrazol-5(4H)-one (1j).** 65% yield, mp: 158–160 °C. <sup>1</sup>H NMR (CDCl<sub>3</sub>, 500 MHz) δ 3.88 (s, 2H), 7.13–7.16 (m, 2H), 7.49–7.51 (m, 3H), 7.78–7.80 (m, 2H), 7.96–7.99 (m, 2H); ESI-TOF-MS: 255.09 (C<sub>15</sub>H<sub>12</sub>FN<sub>2</sub>O, [M+H]<sup>+</sup>).

**4.1.2.11. 3-Phenyl-1-(3-nitrophenyl)pyrazol-5(4H)-one (1k).** 80% yield, mp: 172–174 °C. <sup>1</sup>H NMR (CDCl<sub>3</sub>, 500 MHz) δ 3.94 (s, 2H), 7.50–7.52 (m, 3H), 7.62 (t, *J* = 8.2 Hz, 1H), 7.82–7.84 (m, 2H), 8.08 (d, *J* = 7.85 Hz, 1H), 8.48 (d, *J* = 8.05 Hz, 1H), 8.89 (s, 1H); ESI-TOF-MS: 282.08 (C<sub>15</sub>H<sub>12</sub>N<sub>3</sub>O<sub>3</sub>, [M+H]<sup>+</sup>).

#### 4.1.3. General procedure for the preparation of compounds 2a–u

An equimolar of pyrazolone **1a–k**, substituted benzaldehyde and piperidine in ethanol (50 ml) were refluxed for 3–5 h. The excess of ethanol was evaporated and the residue was poured into water. The solid product was filtered, dried, and recrystallized from methanol. The spectroscopic data of the synthesized compounds are described below in details.

**4.1.3.1. 1,3-Diphenyl-4-benzylidenepyrazol-5(4H)-one (2a).** 74% yield, mp: 232–233 °C. <sup>1</sup>H NMR (DMSO-*d*<sub>6</sub>, 500 MHz) δ 7.08–7.14 (m, 4H), 7.16–7.27 (m, 6H), 7.34–7.37 (m, 4H), 8.02–8.03 (m, 2H); ESI-TOF-MS: 325.13 (C<sub>22</sub>H<sub>17</sub>N<sub>2</sub>O, [M+H]<sup>+</sup>).

**4.1.3.2. 1,3-Diphenyl-4-(4-methoxybenzylidene)pyrazol-5(4H)-one (2b).** 81% yield, mp: 222–224 °C. <sup>1</sup>H NMR (DMSO-*d*<sub>6</sub>, 500 MHz) δ 3.64 (s, 3H), 7.10–7.15 (m, 5H), 7.22–7.27 (m, 5H), 7.34–7.37 (m, 3H), 8.01–8.03 (m, 2H); ESI-TOF-MS: 355.14 (C<sub>23</sub>H<sub>19</sub>N<sub>2</sub>O<sub>2</sub>, [M+H]<sup>+</sup>).

**4.1.3.3. 1,3-Diphenyl-4-(4-acetamidobenzylidene)pyrazol-5(4H)-one (2c).** 75% yield, mp: 184–186 °C. <sup>1</sup>H NMR (DMSO-*d*<sub>6</sub>, 500 MHz) δ 2.0 (s, 3H), 7.09–7.15 (m, 6H), 7.23–7.26 (m, 5H), 7.36 (d, *J* = 8.2 Hz, 2H), 8.02 (d, *J* = 8.15 Hz, 2H), 9.76 (s, 1H); ESI-TOF-MS: 382.15 (C<sub>24</sub>H<sub>20</sub>N<sub>3</sub>O<sub>2</sub>, [M+H]<sup>+</sup>).

**4.1.3.4. 1,3-Diphenyl-4-(4-carboxybenzylidene)pyrazol-5(4H)-one (2d).** 83% yield, mp: 194–196 °C. <sup>1</sup>H NMR (DMSO-*d*<sub>6</sub>, 500 MHz) δ 7.09–7.16 (m, 4H), 7.22–7.28 (m, 4H), 7.31–7.38 (m, 3H), 7.80 (d, *J* = 8.95 Hz, 2H), 8.01 (d, *J* = 8.85 Hz, 2H), 12.65 (br, 1H); ESI-TOF-MS: 369.12 (C<sub>23</sub>H<sub>17</sub>N<sub>2</sub>O<sub>3</sub>, [M+H]<sup>+</sup>).

**4.1.3.5. 1,3-Diphenyl-4-(4-dimethylaminobenzylidene)pyrazol-5(4H)-one (2e).** 76% yield, mp: 194–196 °C. <sup>1</sup>H NMR (DMSO-*d*<sub>6</sub>, 500 MHz) δ 2.80 (s, 6H), 6.60 (d, *J* = 8.85 Hz, 2H), 7.03 (d, *J* = 8.65 Hz, 2H), 7.09–7.15 (m, 4H), 7.24–7.26 (m, 4H), 7.34–7.37 (m, 3H); ESI-TOF-MS: 368.17 (C<sub>24</sub>H<sub>22</sub>N<sub>3</sub>O, [M+H]<sup>+</sup>).

**4.1.3.6. 1,3-Diphenyl-4-(3-nitrobenzylidene)pyrazol-5(4H)-one (2f).** 79% yield, mp: 230–232 °C. <sup>1</sup>H NMR (DMSO-*d*<sub>6</sub>, 500 MHz) δ 7.11–7.16 (m, 5H), 7.23–7.29 (m, 5H), 7.37 (t, 2H), 7.52–7.62 (m, 1H), 8.01 (d, 1H), 8.09 (s, 1H); ESI-TOF-MS: 370.12 (C<sub>22</sub>H<sub>16</sub>N<sub>3</sub>O<sub>3</sub>, [M+H]<sup>+</sup>).

**4.1.3.7. 3-Phenyl-1-(4-chlorophenyl)-4-benzylidenepyrazol-5(4H)-one (2g).** 73% yield, mp: 245–247 °C. <sup>1</sup>H NMR (DMSO-*d*<sub>6</sub>, 500 MHz) δ 7.19–7.21 (m, 4H), 7.23–7.32 (m, 5H), 7.40–7.41 (m, 2H), 7.55 (d, *J* = 8.85 Hz, 2H), 7.88 (d, *J* = 8.9 Hz, 2H); ESI-TOF-MS: 359.09 (C<sub>22</sub>H<sub>16</sub>ClN<sub>2</sub>O, [M+H]<sup>+</sup>).

**4.1.3.8. 3-Phenyl-1-(4-chlorophenyl)-4-(4-chlorobenzylidene)pyrazol-5(4H)-one (2h).** 70% yield, mp: 174–176 °C. <sup>1</sup>H NMR (DMSO-*d*<sub>6</sub>, 500 MHz) δ 7.20 (d, *J* = 8.35 Hz, 2H), 7.25–7.28 (m, 4H), 7.36 (d, *J* = 8.5 Hz, 2H), 7.40 (m, 2H), 7.55 (d, *J* = 8.8 Hz, 2H), 7.86 (d, *J* = 8.85 Hz, 2H); ESI-TOF-MS: 393.05 (C<sub>22</sub>H<sub>15</sub>Cl<sub>2</sub>N<sub>2</sub>O, [M+H]<sup>+</sup>).

**4.1.3.9. 3-Phenyl-1-(4-chlorophenyl)-4-(4-carboxybenzylidene)pyrazol-5(4H)-one (2i).** 85% yield, mp: 249–250 °C. <sup>1</sup>H NMR (DMSO-*d*<sub>6</sub>, 500 MHz) δ 7.23–7.33 (m, 5H), 7.40–7.44 (m, 4H), 7.54 (d, *J* = 8.85 Hz, 2H), 7.88 (d, *J* = 8.9 Hz, 2H), 7.89 (s, 1H), 12.52 (br, 1H); ESI-TOF-MS: 403.08 (C<sub>23</sub>H<sub>16</sub>ClN<sub>2</sub>O<sub>3</sub>, [M+H]<sup>+</sup>).

**4.1.3.10. 3-Phenyl-1-(4-chlorophenyl)-4-(4-acetamidobenzylidene)pyrazol-5(4H)-one (2j).** 69% yield, mp: 182–184 °C. <sup>1</sup>H NMR (DMSO-*d*<sub>6</sub>, 500 MHz) δ 1.99 (s, 3H), 7.10 (d, *J* = 8.55 Hz, 2H), 7.18–7.21 (m, 3H), 7.26 (d, *J* = 6.95 Hz, 2H), 7.32–7.35 (m, 2H), 7.42–7.43 (m, 1H), 7.48 (d, *J* = 9.0 Hz, 2H), 7.98 (d, *J* = 7.1 Hz, 2H), 9.83 (s, 1H); ESI-TOF-MS: 416.11 (C<sub>24</sub>H<sub>19</sub>ClN<sub>3</sub>O<sub>2</sub>, [M+H]<sup>+</sup>).

**4.1.3.11. 3-Phenyl-1-(4-chlorophenyl)-4-(4-methoxybenzylidene)pyrazol-5(4H)-one (2k).** 80% yield, mp: 190–191 °C. <sup>1</sup>H NMR (DMSO-*d*<sub>6</sub>, 500 MHz) δ 3.65 (s, 3H), 6.74 (d, *J* = 7.25 Hz, 2H), 7.14–7.18 (m, 4H), 7.24 (d, *J* = 7.05 Hz, 2H), 7.31–7.34 (m, 2H), 7.46 (d, *J* = 9.0 Hz, 2H), 8.0 (d, *J* = 9.0 Hz, 2H); ESI-TOF-MS: 389.10 (C<sub>23</sub>H<sub>18</sub>ClN<sub>2</sub>O<sub>2</sub>, [M+H]<sup>+</sup>).

**4.1.3.12. 3-Phenyl-1-(4-chlorophenyl)-4-(4-hydroxybenzylidene)pyrazol-5(4H)-one (2l).** 83% yield, mp: 218–220 °C. <sup>1</sup>H NMR (DMSO-*d*<sub>6</sub>, 500 MHz) δ 6.97 (d, *J* = 8.5 Hz, 2H), 7.14–7.17 (m, 2H), 7.24 (d, *J* = 6.95 Hz, 2H), 7.28–7.31 (m, 1H), 7.45 (d, *J* = 7.15 Hz, 2H), 7.52–7.54 (m, 1H), 7.70–7.72 (m, 1H), 8.03 (d, *J* = 6.9 Hz, 2H), 8.58 (d, *J* = 8.85 Hz, 1H), 9.03 (s, 1H); ESI-TOF-MS: 375.09 (C<sub>22</sub>H<sub>16</sub>ClN<sub>2</sub>O<sub>2</sub>, [M+H]<sup>+</sup>).

**4.1.3.13. 3-Phenyl-1-(4-methoxyphenyl)-4-(4-carboxybenzylidene)pyrazol-5(4H)-one (2m).** 72% yield, mp: 230–232 °C. <sup>1</sup>H NMR (DMSO-*d*<sub>6</sub>, 500 MHz) δ 3.74 (s, 3H), 7.0 (d, *J* = 8.15 Hz, 2H), 7.23 (s, 1H), 7.27 (d, *J* = 8.25 Hz, 2H), 7.32–7.35 (m, 5H), 7.65 (d, *J* = 8.2 Hz, 2H), 7.84 (d, *J* = 8.35 Hz, 2H), 14.28 (br, 1H); ESI-TOF-MS: 399.13 (C<sub>24</sub>H<sub>19</sub>N<sub>2</sub>O<sub>4</sub>, [M+H]<sup>+</sup>).

**4.1.3.14. 3-Phenyl-1-(4-isopropylphenyl)-4-(4-carboxybenzylidene)pyrazol-5(4H)-one (2n).** 69% yield, mp: 206–208 °C. <sup>1</sup>H NMR (DMSO-*d*<sub>6</sub>, 500 MHz) δ 1.21 (s, 6H), 2.86–2.89 (m, 1H), 7.16 (d, *J* = 7.55 Hz, 2H), 7.22–7.24 (m, 2H), 7.25–7.28 (m, 4H), 7.31 (d, *J* = 8.25 Hz, 2H), 7.80 (d, *J* = 8.35 Hz, 2H), 7.86 (d, *J* = 8.5 Hz, 2H), 12.59 (br, 1H); ESI-TOF-MS: 411.17 (C<sub>26</sub>H<sub>23</sub>N<sub>2</sub>O<sub>3</sub>, [M+H]<sup>+</sup>).

**4.1.3.15. 3-Phenyl-1-(4-*tert*-butylphenyl)-4-(4-carboxybenzylidene)pyrazol-5(4H)-one (2o).** 71% yield, mp: 198–200 °C. <sup>1</sup>H NMR (DMSO-*d*<sub>6</sub>, 500 MHz) δ 1.29 (s, 9H), 7.21–7.24 (m, 4H), 7.27–7.28 (m, 2H), 7.31 (d, *J* = 8.2 Hz, 2H), 7.48 (d, *J* = 8.75 Hz, 2H), 7.74 (d, *J* = 8.75 Hz, 2H), 7.87 (d, *J* = 8.35 Hz, 2H), 12.47 (br, 1H); ESI-TOF-MS: 425.18 (C<sub>27</sub>H<sub>25</sub>N<sub>2</sub>O<sub>3</sub>, [M+H]<sup>+</sup>).

**4.1.3.16. 3-Phenyl-1-(4-cyanophenyl)-4-(4-carboxybenzylidene)pyrazol-5(4H)-one (2p).** 87% yield, mp: 179–181 °C. <sup>1</sup>H NMR (DMSO-*d*<sub>6</sub>, 500 MHz) δ 7.21–7.25 (m, 2H), 7.28 (s, 1H), 7.29–7.32 (m, 1H), 7.38–7.41 (m, 1H), 7.87 (d, *J* = 8.4 Hz, 2H), 7.92 (d,



$J = 8.75$  Hz, 2H), 8.01–8.03 (m, 1H), 8.13 (d,  $J = 8.2$  Hz, 2H), 8.15 (d,  $J = 8.85$  Hz, 2H), 12.58 (br, 1H); ESI-TOF-MS: 394.11 ( $C_{24}H_{16}N_3O_3$ ,  $[M+H]^+$ ).

**4.1.3.17. 3-Phenyl-1-(4-trifluoromethoxyphenyl)-4-(4-carboxybenzylidene)pyrazol-5(4H)-one (2q).** 81% yield, mp: 164–166 °C.  $^1H$  NMR (DMSO- $d_6$ , 500 MHz)  $\delta$  7.13–7.15 (m, 3H), 7.19–7.32 (m, 5H), 7.36 (d,  $J = 8.75$  Hz, 2H), 7.75 (d,  $J = 8.2$  Hz, 2H), 8.14 (d,  $J = 9.05$  Hz, 2H), 12.45 (br, 1H); ESI-TOF-MS: 453.11 ( $C_{24}H_{16}F_3N_2O_4$ ,  $[M+H]^+$ ).

**4.1.3.18. 3-Phenyl-1-(3-chlorophenyl)-4-(4-carboxybenzylidene)pyrazol-5(4H)-one (2r).** 85% yield, mp: 161–163 °C.  $^1H$  NMR (DMSO- $d_6$ , 500 MHz)  $\delta$  7.14–7.18 (m, 3H), 7.23–7.30 (m, 5H), 7.38–7.42 (m, 1H), 7.81 (d,  $J = 8.35$  Hz, 2H), 8.03 (d,  $J = 9.45$  Hz, 2H), 8.13 (t,  $J = 9.0$  Hz, 1H), 12.74 (br, 1H); ESI-TOF-MS: 403.08 ( $C_{23}H_{16}ClN_2O_3$ ,  $[M+H]^+$ ).

**4.1.3.19. 3-Phenyl-1-(3,4-dichlorophenyl)-4-(4-carboxybenzylidene)pyrazol-5(4H)-one (2s).** 83% yield, mp: 185–187 °C.  $^1H$  NMR (DMSO- $d_6$ , 500 MHz)  $\delta$  7.14–7.30 (m, 5H), 7.62 (d,  $J = 9.0$  Hz, 2H), 7.81 (d,  $J = 8.35$  Hz, 2H), 8.08 (d,  $J = 8.9$  Hz, 2H), 8.31 (m, 2H), 12.62 (br, 1H); ESI-TOF-MS: 437.04 ( $C_{23}H_{15}Cl_2N_2O_3$ ,  $[M+H]^+$ ).

**4.1.3.20. 3-Phenyl-1-(4-fluorophenyl)-4-(4-carboxybenzylidene)pyrazol-5(4H)-one (2t).** 78% yield, mp: 190–192 °C.  $^1H$  NMR (DMSO- $d_6$ , 400 MHz)  $\delta$  7.12–7.28 (m, 5H), 7.31 (d,  $J = 8.24$  Hz, 2H), 7.80 (d,  $J = 8.28$  Hz, 2H), 8.0 (d,  $J = 8.52$  Hz, 2H), 8.03 (d,  $J = 8.4$  Hz, 2H), 8.1 (s, 1H), 12.62 (br, 1H); ESI-TOF-MS: 387.11 ( $C_{23}H_{16}FN_2O_3$ ,  $[M+H]^+$ ).

**4.1.3.21. 3-Phenyl-1-(3-nitrophenyl)-4-(4-carboxybenzylidene)pyrazol-5(4H)-one (2u).** 72% yield, mp: 175–177 °C.  $^1H$  NMR (DMSO- $d_6$ , 500 MHz)  $\delta$  7.13–7.36 (m, 7H), 7.69–7.30 (m, 1H), 7.85 (d,  $J = 8.1$  Hz, 2H), 8.03 (d,  $J = 7.3$  Hz, 1H), 8.48 (d,  $J = 8.1$  Hz, 2H), 8.88 (s, 1H), 12.71 (br, 1H); ESI-TOF-MS: 414.10 ( $C_{23}H_{16}N_3O_5$ ,  $[M+H]^+$ ).

## 4.2. 3CL<sup>PRO</sup> and 3C<sup>PRO</sup> activity assays

A fluorogenic peptide substrate (Dabcyl-KTSAVL QSGFRKME-Edans) was used for assays of 3CL<sup>PRO</sup> and 3C<sup>PRO</sup> activities. SARS-CoV 3CL<sup>PRO</sup> and CVB3 3C<sup>PRO</sup> were prepared as previously reported.<sup>8,31</sup> The proteases were stored in the buffer containing 12 mM Tris-HCl (pH 7.5), 120 mM NaCl, 0.1 mM EDTA, 7.5 mM  $\beta$ -ME, and 1 mM DTT at –70 °C before use. The anti-SARS-3CL<sup>PRO</sup> activity of the test compounds were performed in the solution containing 0.05  $\mu$ M SARS 3CL<sup>PRO</sup>, 6  $\mu$ M fluorogenic substrate, and 50  $\mu$ M of test compounds at 25 °C and the anti-CVB3 3C<sup>PRO</sup> activity was assayed using 0.05  $\mu$ M CVB3 3C<sup>PRO</sup>. Enhanced fluorescence of the reactions in the buffer of 20 mM Bis-Tris at pH 7.0 was monitored at 538 nm with excitation at 355 nm using a fluorescence plate reader (Fluoroskan Ascent; ThermoLabsystems, Helsinki, Finland). The compounds which inhibited more than 50% of the protease activity at 50  $\mu$ M were selected for the next assay run.

## 4.3. Cytotoxicity assay

Cell viability was determined by MTT 3-(4,5-dimethyl thiazol-2-yl)-2,5-diphenyl tetrazolium bromide,<sup>32</sup> using Vybrant<sup>®</sup> MTT cell proliferation assay kit purchased from Molecular Probes, USA. Human embryonic kidney (HEK) 293 cells ( $2 \times 10^5$ /ml) were seeded into a 96-well culture plate containing 0.1 ml of Minimum Essential Medium (MEM) (Gibico, Invitrogen, CA, USA) supplemented with 10% fetal bovine serum (FBS) (Gibico) and cultured

in 5% CO<sub>2</sub> at 37 °C. Cells with 70% confluence at density were treated with each compound at designated concentrations for 24 h. After the incubation, 10  $\mu$ L of MTT (3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide) stock solution was added into each well. The conversion of MTT to formazan by viable cells was performed at 37 °C for another 4 h. After the reaction, 100  $\mu$ L of DMSO solution were added into each well following the removal of culture media in order to solubilize the formazan precipitates. The levels of formazan were determined by optical density at 540 nm using an ELISA reader and represented as cell viability.

## 4.4. Docking studies

To gain further molecular insight into the mode of inhibition of active compound, we conducted docking studies in the 3CL<sup>PRO</sup> active site. For modeling analysis, the crystal structure of SARS 3CL<sup>PRO</sup> in complex with a peptide inhibitor (PDB code 1UK4) was used.<sup>33</sup> Docking process was performed using an automated ligand-docking subprogram of the Discovery Studio Modeling 1.2 SBD (Accelrys Inc., San Diego, CA), with a set of parameters chosen to control the precise operation of the genetic algorithm. Docking runs were carried out using standard default settings 'grid resolution' of 5 Å, 'site opening' of 12 Å, and 'binding site' selected for defining the active site cavity.

## References and notes

- Ksiazek, T. G.; Erdman, D.; Goldsmith, C. S.; Zaki, S. R.; Peret, T.; Emery, S.; Tong, S.; Urbani, C.; Comer, J. A.; Lim, W.; Rollin, P. E.; Dowell, S. F.; Ling, A.-E.; Humphrey, C. D.; Shieh, W.-J.; Guarner, J.; Paddock, C. D.; Rota, P.; Fields, B.; DeRisi, J.; Yang, J.-Y.; Cox, N.; Hughes, J. M.; LeDuc, J. W.; Bellini, W. J.; Anderson, L. J. *N. Engl. J. Med.* **2003**, *348*, 1953.
- Peiris, J. S. M.; Lai, S.-T.; Poon, L. L.-M.; Guan, Y.; Yam, L. Y.-C.; Lim, W.; Nicholls, J.; Yee, W. K.-S.; Yan, W. W.; Cheung, M.-T.; Cheng, V. C.-C.; Chan, K.-H.; Tsang, D. N.-C.; Yung, R. W.-H.; Ng, T. K.; Yuen, K.-Y. *Lancet* **2003**, *361*, 1319.
- Rota, P. A.; Oberste, M. S.; Monroe, S. S.; Nix, W. A.; Campagnoli, R.; Icenogle, J. P.; Penaranda, S.; Bankamp, B.; Maher, K.; Chen, M.-H.; Tong, S.; Tamin, A.; Lowe, L.; Frace, M.; DeRisi, J. L.; Chen, Q.; Wang, X.-N.; Chen, S.-J.; Chen, Z.; Hao, P.; Tang, H.; Ksiazek, T. G.; Rollin, P. E.; Sanchez, A.; Liffick, S.; Holloway, B.; Limor, J.; McCaustland, K.; Olsen-Rasmussen, M.; Fouchier, R.; Gunther, S.; Osterhaus, A. D. M. E.; Drosten, C.; Pallansch, M. A.; Anderson, L. J.; Bellini, W. J. *Science* **2003**, *300*, 1394.
- He, J.-F.; Peng, G.-W.; Min, J.; Yu, D.-W.; Liang, W.-L.; Zhang, S.-Y.; Xu, R.-H.; Zheng, H.-Y.; Wu, X.-W.; Xu, J.; Wang, Z.-H.; Fang, L.; Zhang, X.; Li, H.; Yan, X.-G.; Lu, J.-H.; Hu, Z.-H.; Huang, J.-C.; Wan, Z.-Y.; Hou, J.-L.; Lin, J.-Y.; Song, H.-D.; Wang, S.-Y.; Zhou, X.-J.; Zhang, G.-W.; Gu, B.-W.; Zheng, H.-J.; Zhang, X.-L.; He, M.; Zheng, K.; Wang, B.-F.; Fu, G.; Wang, X.-N.; Chen, S.-J.; Chen, Z.; Hao, P.; Tang, H.; Ren, S.-X.; Zhong, Y.; Guo, Z.-M.; Liu, Q.; Miao, Y.-G.; Kong, X.-Y.; He, W.-Z.; Li, Y.-X.; Wu, C.-L.; Zhao, G.-P.; Chiu, R. W. K.; Chim, S. S. C.; Tong, Y.-K.; Chan, P. K. S.; Tam, J. S.; Lo, Y. M. D. *Science* **2004**, *303*, 1666.
- Anand, K.; Ziebuhr, J.; Wadhwani, P.; Mesters, J. R.; Hilgenfeld, R. *Science* **2003**, *300*, 1763.
- Du, Q. S.; Wang, S. Q.; Zhu, Y.; Wei, D. Q.; Guo, H.; Sirois, S.; Chou, K. C. *Peptides* **2004**, *25*, 1857.
- (a) Binford, S. L.; Maldonado, F.; Brothers, M. A.; Weady, P. T.; Zalman, L. S.; Meador, J. W., 3rd; Matthews, D. A.; Patick, A. K. *Antimicrob. Agents Chemother.* **2005**, *49*, 619; (b) Lee, E. S.; Lee, W. G.; Yun, S. H.; Rho, S. H.; Im, I.; Yang, S. T.; Sellamuthu, S.; Lee, Y. J.; Kwon, S. J.; Park, O. K.; Jeon, E. S.; Park, W. J.; Kim, Y. C. *Biochem. Biophys. Res. Commun.* **2007**, *358*, 7; (c) Kuo, C. J.; Shie, J. J.; Fang, J. M.; Yen, G. R.; Hsu, J. T.; Liu, H. G.; Tseng, S. N.; Chang, S. C.; Lee, C. Y.; Shih, S. R.; Liang, P. H. *Bioorg. Med. Chem.* **2008**, *16*, 7388.
- Lee, C. C.; Kuo, C. J.; Ko, T. P.; Hsu, M. F.; Tsui, Y. C.; Chang, S. C.; Yang, S.; Chen, S. J.; Chen, H. C.; Hsu, M. C.; Shih, S. R.; Liang, P. H.; Wang, A. H. *J. Biol. Chem.* **2008**, *284*, 7646.
- (a) Wu, C.-Y.; Jan, J.-T.; Ma, S.-H.; Kuo, C.-J.; Juan, H.-F.; Cheng, E. Y.-S.; Hsu, H.-H.; Huang, H.-C.; Wu, D.; Briik, A.; Liang, F.-S.; Liu, R.-S.; Fang, J.-M.; Chen, S.-T.; Liang, P.-H.; Wong, C.-H. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 10012; (b) Shao, Y.-M.; Yang, W.-B.; Peng, H.-P.; Hsu, M.-F.; Tsai, K.-C.; Kuo, T.-H.; Wang, A. H.-J.; Liang, P.-H.; Lin, C.-H.; Yang, A.-S.; Wong, C.-H. *ChemBioChem* **2007**, *8*, 1654.
- Kao, R. Y.; Tsui, W. H. W.; Lee, T. S. W.; Tanner, J. A.; Watt, R. M.; Huang, J. D.; Hu, L. H.; Chen, G. H.; Chen, Z. W.; Zhang, L. Q.; He, T.; Chan, K. H.; Tse, H.; To, A. P. C.; Ng, L. W. Y.; Wong, B. C. W.; Tsoi, H. W.; Yang, D.; Ho, D. D.; Yuen, K. Y. *Chem. Biol.* **2004**, *11*, 1293.
- Blanchard, J. E.; Elowe, N. H.; Huitema, C.; Fortin, P. D.; Cechetto, J. D.; Eltis, L. D.; Brown, E. D. *Chem. Biol.* **2004**, *11*, 1445.
- (a) Chen, L. D.; Gui, C. S.; Luo, X. M.; Yang, Q. G.; Gunther, S.; Scandella, E.; Drosten, C.; Bai, D.; He, X. C.; Ludewig, B.; Chen, J.; Luo, H. B.; Yang, Y. M.; Yang,

- Y. F.; Zou, J. P.; Thiel, V.; Chen, K.; Shen, J. H.; Xu, S.; Jiang, H. L. *J. Virol.* **2005**, *79*, 7095; (b) Yang, Q.; Chen, L.; He, X.; Gao, Z.; Shen, X.; Bai, D. *Chem. Pharm. Bull.* **2008**, *56*, 1400.
13. Liu, Z.; Huang, C.; Fan, K.; Wei, P.; Chen, H.; Liu, S.; Pei, J.; Shi, L.; Li, B.; Yang, K.; Liu, Y.; Lai, L. *J. Chem. Inf. Model.* **2005**, *45*, 10.
14. Jain, R. P.; Petterson, H. I.; Zhang, J.; Aull, K. D.; Fortin, P. D.; Huitema, C.; Eltis, L. D.; Parrish, J. C.; James, M. N. G.; Wishart, D. S.; Vederas, J. C. *J. Med. Chem.* **2004**, *47*, 6113.
15. Shie, J.-J.; Fang, J.-M.; Kuo, T.-H.; Kuo, C.-J.; Liang, P.-H.; Huang, H.-J.; Yang, W.-B.; Lin, C.-H.; Chen, J.-L.; Wu, Y.-T.; Wong, C.-H. *J. Med. Chem.* **2005**, *48*, 4469.
16. Bacha, U.; Barrila, J.; Velasquez-Campoy, A.; Leavitt, S. A.; Freire, E. *Biochemistry* **2004**, *43*, 4906.
17. Chen, L.-R.; Wang, Y.-C.; Lin, Y.-W.; Chou, S.-Y.; Chen, S.-F.; Liu, L.-T.; Wu, Y.-T.; Kuo, C.-J.; Chen, T. S.-S.; Juang, S.-H. *Bioorg. Med. Chem. Lett.* **2005**, *15*, 3058.
18. Ramajayam, R.; Tan, K.-P.; Liu, H.-G.; Liang, P.-H. *Bioorg. Med. Chem. Lett.* **2010**, *20*, 3569.
19. Wu, C.-Y.; King, K.-Y.; Kuo, C.-J.; Fang, J.-M.; Wu, Y.-T.; Ho, M.-Y.; Liao, C.-L.; Shie, J.-J.; Liang, P.-H.; Wong, C.-H. *Chem. Biol.* **2006**, *13*, 4469.
20. (a) Bacha, U.; Barrila, J.; Gabelli, B.; Kiso, Y.; Amzel, L. M.; Freire, E. *Chem. Biol. Drug Des.* **2008**, *72*, 34; (b) Regnier, T.; Sarma, D.; Hidaka, K.; Bacha, U.; Freire, E.; Hayashi, Y.; Kiso, Y. *Bioorg. Med. Chem. Lett.* **2009**, *19*, 2722.
21. (a) Ghosh, A. K.; Xi, K.; Ratia, K.; Santarsiero, B. D.; Fu, W.; Harcourt, B. H.; Rota, P. A.; Baker, S. C.; Johnson, M. E.; Mesecar, A. D. *J. Med. Chem.* **2005**, *48*, 6767; (b) Shie, J.-J.; Fang, J.-M.; Kuo, T.-H.; Kuo, C.-J.; Liang, P.-H.; Huang, H.-J.; Wu, Y.-T.; Jan, J.-T.; Cheng, E. Y.-S.; Wong, C.-H. *Bioorg. Med. Chem.* **2005**, *13*, 5240; (c) Ghosh, A. K.; Xi, K.; Grum-Tokars, V.; Xu, X.; Ratia, K.; Fu, W.; Houser, K. V.; Baker, S. C.; Johnson, M. E.; Mesecar, A. D. *Bioorg. Med. Chem. Lett.* **2007**, *17*, 5876.
22. Kaeppler, U.; Stiefl, N.; Schiller, M.; Vicik, R.; Breuning, A.; Schmitz, W.; Rupprecht, D.; Schmuck, C.; Baumann, K.; Ziebuhr, J.; Schirmeister, T. *A. J. Med. Chem.* **2005**, *48*, 6832.
23. Hsu, J. T. A.; Kuo, C. J.; Hsieh, H. P.; Wang, Y. C.; Huang, K. K.; Lin, C. P. C.; Huang, P. F.; Chen, X.; Liang, P. H. *FEBS Lett.* **2004**, *574*, 116.
24. (a) Watanabe, T.; Yuki, S.; Egawa, M.; Nishi, H. *J. Pharmacol. Exp. Ther.* **1994**, *268*, 1597; (b) Kawai, H.; Nakai, H.; Suga, M.; Yuki, S.; Watanabe, T.; Saito, K. I. *J. Pharmacol. Exp. Ther.* **1997**, *281*, 921.
25. Wu, T. W.; Zeng, L. H.; Wu, J.; Fung, K. P. *Life Sci.* **2002**, *71*, 2249.
26. Fan, X.; Zhang, X.; Zhou, L.; Keith, K. A.; Kernb, E. R.; Torrence, P. F. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 3224.
27. Chen, L.; Chen, S.; Gui, C.; Shen, J.; Shen, X.; Jiang, H. *J. Biomol. Screen.* **2006**, *11*, 915.
28. Kuo, C. J.; Liu, H. G.; Lo, Y. K.; Seong, C. M.; Lee, K. I.; Jung, Y. S.; Liang, P. H. *FEBS Lett.* **2009**, *583*, 549.
29. Wang, X.-J.; Tan, J.; Grozinger, K. *Tetrahedron Lett.* **2000**, *41*, 4713.
30. Dhal, P. N.; Achary, T. E.; Nayak, A. J. *Indian Chem. Soc.* **1975**, *52*, 1196.
31. Kuo, C. J.; Chi, Y. H.; Hsu, J. T. A.; Liang, P. H. *Biochem. Biophys. Res. Commun.* **2004**, *318*, 862.
32. Terachi, T.; Stanescu, G.; Pontes, J. E.; Medof, M. E.; Caulfield, M. J. *Cancer Res.* **1991**, *51*, 2515.
33. 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.; Rao, Z. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 13190.