ELSEVIER

Structural analysis of inhibition mechanisms of Aurintricarboxylic Acid on SARS-CoV polymerase and other proteins

YeeLeng Yap [a], XueWu Zhang [b], Anton Andonov [c], RunTao He [c,*]

[a] *Institute Pasteur, Unité de Génétique des Génomes Bactériens, CNRS URA 2171, 28 rue du Docteur Roux, 75724 Paris Cedex 15, France*
[b] *Department of Biology, College of Life and Environment Sciences, Shanghai Normal University, 100 GuiLin Road, Shanghai, 200234, China*
[c] *National Microbiology Laboratory, Public Health Agency of Canada, 1015 Arlington St., Winnipeg, MB, R3E 3R2, Canada*

Abstract

We recently published experimental results that indicated Aurintricarboxylic Acid (ATA) could selectively inhibit SARS-CoV replication inside host cells by greater than 1000 times. This inhibition suggested that ATA could be developed as potent anti-viral drug. Here, to extend our experimental observation, we have incorporated protein structural studies (with positive/negative controls) to investigate the potential binding modes/sites of ATA onto RNA-dependent RNA polymerase (RdRp) from SARS-CoV and other pathogenic positive-strand RNA-viruses, as well as other proteins in SARS-CoV based on the fact that ATA binds to Ca^{2+}-activated neutral protease (*m*-calpain), the protein tyrosine phosphatase (PTP) and HIV integrase which have existing crystal structures. Eight regions with homologous 3D-conformation were derived for 10 proteins of interest. One of the region, $R_{binding}$ (754–766 in SARS-CoV's RdRp), located in the palm sub-domain mainly constituted of anti-parallel β-strand-turn-β-strand hairpin structures that covers two of the three RdRp catalytic sites (Asp 760, Asp761), was also predicted by molecular docking method (based on free energy of binding ΔG) to be important binding motif recognized by ATA. The existence of this strictly conserved region that incorporated catalytic residues, coupled with the homologous ATA binding pockets and their consistent ΔG values, suggested strongly ATA may be involved in an analogous inhibition mechanism of SARS-COV's RdRp in concomitant to the case in *m*-calpain, PTP and HIV integrase.
© 2005 Elsevier Ltd. All rights reserved.

Keywords: SARS-CoV; RNA-dependent RNA polymerase; Inhibitor; HIV

1. Introduction

Severe Acute Respiratory Syndrome (SARS) is an infectious disease that appeared in November 2002 (Drosten et al., 2003; Marra et al., 2003; Rota et al., 2003), and has imperilled the health of human population in more than 30 nations during its outbreak. It has claimed over 770 lives and infected more than 8050 people (∼9.6% death rate) around the globe. On May 15, 2003, the primary etiological agent for SARS that belongs to the coronavirus family was found to fulfil Koch's postulate through experimental infection of cynomol-

gus macaques (*Macaca fascicularis*). By May 18, 2004, the last human case of infection was successfully contained in China. The detailed chronicles for the discovery of SARS CoronaVirus (SARS-CoV) can be found in websites (James, 2003).

Despite almost two years of intensive SARS-CoV research around the world, the key question regarding the original reservoir of SARS-CoV remains widely speculative and unanswered. This raises the concerns that another outbreak could be imminent. This concern, coupled with the fact that preventive vaccine projects underway now (Marshall and Enserink, 2004) are expected to take years to be developed with possible adverse side effects (Glansbeek et al., 2002; Marshall and Enserink, 2004), the anti-viral agents such as an effective protein inhibitor must be studied exhaustively to define alternative options. In the past, some studies have

---

* Corresponding author. Canadian Science Centre for Human & Animal Health 1015 Arlington Street, Winnipeg, MB R3E 3R2, Canada.
Tel.: +1 204 789 5018; fax: +1 204 789 2082.
E-mail address: runtao_he@phac-aspc.gc.ca (R. He).

reported potent inhibitors that block key proteins from coronaviruses, including those from SARS-CoV (e.g. 3C-Like protease, S proteins) which are directly involved in the viral entry machinery (Anand et al., 2003; Bermejo Martin et al., 2003; Chou et al., 2003; Koren et al., 2003; Pace et al., 2004; Takeda-Shitaka et al., 2004; Wu et al., 2004; Zhang and Yap, 2004). The successful inhibitory binding of these anti-viral agents to the targeted proteins will disrupt the normal viral cell cycle and therefore render the viral infection to the human host cell unsuccessful (Anand et al., 2003; Bermejo Martin et al., 2003; Chou et al., 2003; Koren et al., 2003; Pace et al., 2004; Takeda-Shitaka et al., 2004; Wu et al., 2004; Zhang and Yap, 2004). In contrast to the widely studied proteins mentioned above, there are limited documentations reporting the inhibition of the RNA-dependent RNA polymerase (RdRp) from any coronaviruses by any potent inhibitor despite the fact that this attractive protein target that is crucial for the overall viral RNA replication inside the host cells (Dhanak et al., 2002; Sarisky, 2004). Recently, we have shown by experiment, that Aurintricarboxylic Acid (ATA) could selec-

tively inhibit SARS-CoV replication inside host cells with their viral mRNA transcripts production 1000-fold less than that in the untreated control (He et al., 2004). This inhibition of viral replication is possibly related to SARS-CoV's RdRp. However, more extensive experiments will be needed to prove that ATA does bind to RdRp and inhibit the viral replication. We searched the appearance of term "ATA" in MedLine and found out that ATA is tightly correlated to polymerase activity (Akiyama et al., 1977; Givens and Manly, 1976; Swennen et al., 1978; Swennen et al., 1979). Additionally, when ATA was compared with interferons (IFNs) alpha (α) and beta (β), which is the current standard treatment for SARS patients, ATA is reported to be approximately 10 times more potent than IFN alpha and 100 times more than interferon beta at their highest concentrations designed in our previous experiment. All these observations suggested that further exploration on ATA might be scientifically a rewarding task, and this chemical compound (ATA) can possibly be developed into an effective anti-SARS drug with high selectivity. Here, to extend our experimental observations,

```
1C2PA  (  98 )  -kSk-f-----gY-gAkdVrn-ls----skAvn--hIh-svwkdLledt-
1DF0A  (   2 )          agiamklakdreaaeglGsherAikylnqdyetlrnecle
1HHSA  ( 151 )  -TcIpyfsn--dmgtKieiAerALekAeeAGnl--MlqgkfddAyqlhq-
1KHVA  ( 126 )  -CG-PfVp----g-kkidhvkdg--vmdqvLak--hLy-kcwsvAnsGk-
1O5SA  ( 502 )  -aGfpfnk---Wg-kArly---y----ds-msy--Edq-dalfaytkrn-
1QZ0A  ( 187 )              spygp---earae--Lss-rLttlrntLa-
1RDR0  (  20 )  -psktkle---pS-aFhyvFegvitevdeyMke--AVd-hYAgqLmsld-
1S4FA  ( 222 )  -aGf-lEk----k-niGeVLd-------sekh--lVe-qlvrdLkaGr-
1SH0A  ( 119 )  -Sg-hphh----m-rKndcwnge--sFtgkLad--qAs-kAnlmFeegk-
3HVTA  ( 132 )  IpsinnetpgirY-gynVLPqgWkGSpaIfqss--Mtk-ILepFkkqnp-
                                  aa  aaa aaaaaaaa

1C2PA  ( 219 )  ydTr-------cFdSTVt---enDIrvEesiYqcCdLa-peArqAIkSLT
1DF0A  ( 154 )  VDDrLPTkdgelLFVhSaegseFWSALLEKAYAKIngcYEALsgGatteG
1HHSA  ( 323 )  TdVs-------dHdtfWP---GwLRdlICdELlnMgYa-pwWVkLFeTSL
1KHVA  ( 249 )  LdYs-------kWdStMS---pcVvrlAIdILadCCeqteLTkSVvlTLk
1O5SA  ( 617 )  WdYp-------kcdraMP---nmLRimAslVLArKHntccnlshRfyrLA
1QZ0A  ( 257 )  IACQ-----------YPl---qsQLesHFrMLaen---------r---TP
1RDR0  ( 232 )  FdYt-------gYdasLs---pAWfeALkmVLekI-----gFgdrvdyId
1S4FA  ( 344 )  FdTk-------aWdtgVT---skDLqLIgeIQkyYYk--keWhkFIdtit
1SH0A  ( 241 )  AdYs-------rWdStQQ---raVLaaALeIMvkfSsephlAqvVAedLl
3HVTA  ( 243 )  pivlpe-----k-dswtv---ndIqklvgklnwasqIYpgIkvrqLckll
                                 aaaaaaaaaaaaaa          aa

1C2PA  ( 258 )  eRLYigGpLtNsk---gqnCGyR------rCRAsGVLTtsCGnTLTcylK
1DF0A  ( 204 )  FEDFTGGIAEwyeLr--------------kppp-----------nLfk
1HHSA  ( 362 )  k--LPVYVgaPApeqGhtLLGdpsnPdLevGLSsgQgAtdLMGtLLMSIT
1KHVA  ( 289 )  s----hpmTILda----mIVqTk------rGLPsgMpftsVINSICHWLL
1O5SA  ( 657 )  nECaqvLsEmvmc---gGslYvkp------gGTssgdattayanSVFNICQ
1QZ0A  ( 281 )  -----VLAVL--------------------As--ssEIanqr
1RDR0  ( 267 )  -----ylnsG------------------------tsIFnSMINNLI
1S4FA  ( 382 )  dh-mteVpViTad---gevyiRn------GQRGsgQpdtsAGnSMLNVLT
1SH0A  ( 281 )  s----pSvVdVgd----ftIsIn------eGLPsgVpctsqWNSIAHWLL
3HVTA  ( 284 )  r-------gtka--lteviplt---------------eeAelelae
                        b                          aaaaaaaa
```
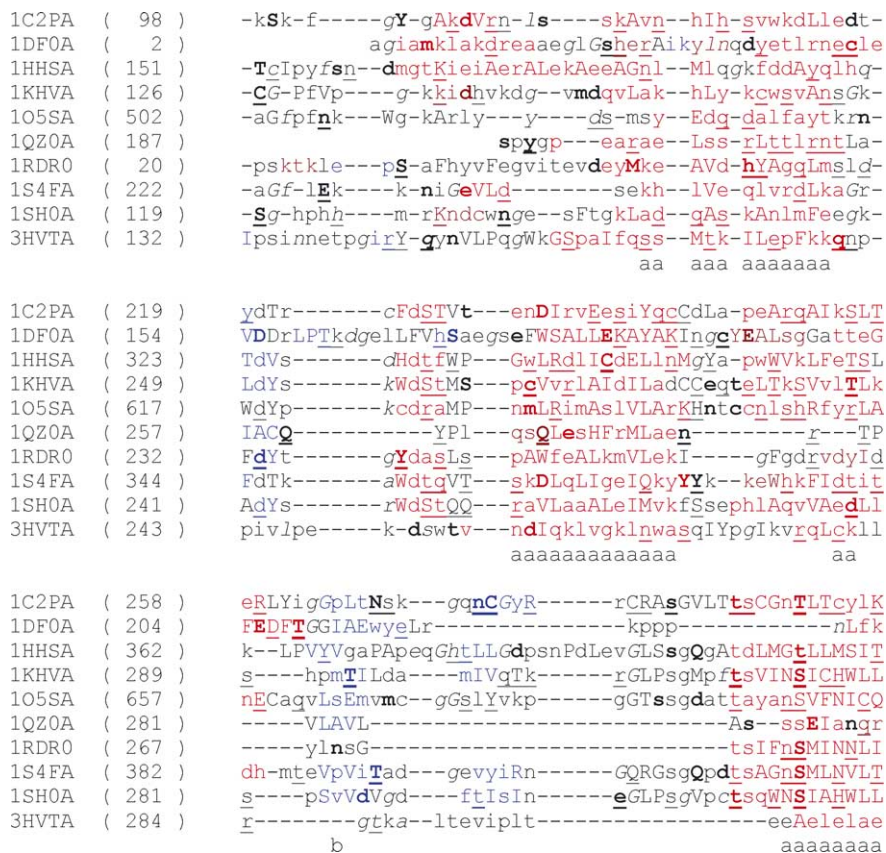
Fig. 1. Structural alignment between proteins from SARS-CoV and other RNA viruses. The structural alignment between proteins from SARS-CoV and other RNA viruses based on their 3D atomic coordinate files were performed using Dali and LGA with manual alignment. Solvent inaccessible residues were represented in upper case, while the solvent accessible residues in lower case. The alpha helix residues in red, specifically the 3–10 helix residues in maroon, and beta strand residues in blue. The residues that have hydrogen bond to main chain amide in bold, and residues that have hydrogen bond to main chain carbonyl underlined. Finally, residues that are joined by disulphide bonds are represented in cedilla (ç) (1C2PA, Hepatitis C virus RNA-dependent RNA polymerase; 1DF0A, m-calpain; 1hhsa, bacteriophage phi6 RNA-dependent RNA polymerase; 1KHVA, Rabbit hemorrhagic disease virus RNA-dependent RNA polymerase; 1O5SA, SARS-CoV RNA-dependent RNA polymerase; 1QZ0A, Yersinia pestis phosphatase yoph; 1RDR0, Poliovirus RNA-dependent RNA polymerase; 1S4FA, Bovine viral diarrhea virus RNA-dependent RNA polymerase; 1SH0A, Norwalk virus RNA-dependent RNA polymerase; 3HVTA, HIV Type 1 RNA-dependent RNA polymerase).

```
1C2PA  ( 299 )  asAAcraA-------------------k---------------------
1DF0A  ( 227 )  IIqkAleKgSlLgCsidighaysVtgAeevqklIrIrnPwgqgefwmsfs
1HHSA  ( 410 )  YLVMQLdhTA-----phlnsr--Ik-----------dmpsAcr-----FL
1KHVA  ( 325 )  WSAAVYkSCa-----ei------g---------lh---csnLye-------
1O5SA  ( 699 )  AvTAnvnall-----st------d--------------------------
1QZ0A  ( 296 )  fgMpdYFrq-----------------------------------------
1RDR0  ( 305 )  IRTLLLktyk----------------------gid----Ld---------
1S4FA  ( 422 )  MMYAFCeSTg----------------------vpyksFnr----------
1SH0A  ( 317 )  TLCALSevtn-----------------------ls---PdiIqa-------
3HVTA  ( 306 )  nreilke-----pvhGv------yYdpskdliaeiqkqgqg---------
                aaaaaaa

1C2PA  ( 308 )  -------LqqctLVNgd---------dLVVIces---ag----vqeda-a
1DF0A  ( 330 )  dFlrhYsrLeICNLtpdTLtcd-sykkWkltkmdgnWr-rGsTAGGcr-n
1HHSA  ( 437 )  dsYWqGheeIrQISksd---------dAMLGWtkg------rAlvGGh-r
1KHVA  ( 346 )  ------d--ApFYTygd---------dGVYAMt------pmMvsllp-a
1O5SA  ( 752 )  ------h--fsMMIlsd---------dAVVCYns---nyaaqGlvaSi-k
1QZ0A  ( 326 )  ---------sgtygsIt--VeSkmtqqvgLgdgimAdMytLtIreagqk
1RDR0  ( 320 )  ------h--LkMIAygd---------dVIASyph---------eVdas-l
1S4FA  ( 440 )  --------vArIHVcgd---------dGFLITek------glGlkFAn-k
1SH0A  ( 335 )  ------n--SlFSFygd---------dEIVSTdi---k----L--dpe-k
3HVTA  ( 336 )  ---------qwtYqIyQep--fknLKTgkyarmrgahtndVkqlTeAVqk
                           bbbb              bbbbbb           aa a

1C2PA  ( 335 )  sLraFte-AtrYs--appgd------------------------------
1DF0A  ( 377 )  ypntf---wMNPQYlIkLeeeDeddedgrGCtFLVGLIQkhrrrqrkmge
1HHSA  ( 471 )  LfemLkeg---------------k--------------------------
1KHVA  ( 371 )  IienLrd-YgL-s--PtAadkte---------------------------
1O5SA  ( 781 )  NfkAVly-YQNnv--FmSeakcw---------------------------
1QZ0A  ( 364 )  tisVpVVHVgn---Wpdg--tAvsS-eVTkaLAs-lVd-q-----tae--
1RDR0  ( 343 )  LaqsGkd-YgL-t--Mt-padks---------------------------
1S4FA  ( 466 )  GmqiLhe-AGK-p--Qkiteg-----------------------------
1SH0A  ( 358 )  LtakLke-YgL-k--Ptrpdkte---------------------------
3HVTA  ( 375 )  ItteSiviWgktP-kFkLpIqketWetwwteywqatwipewefvntpplv

1C2PA  ( 361 )  elI--tscsSnVSvAh-da-sgkrvyYlTr--d-----------PttPLA
1DF0A  ( 491 )  phk--ng--dFCIRVFSek---kAdyqtvddeieAnieeianeedIgdgf
1HHSA  ( 492 )  hGg--aFlgdILlydsrrePgs---AiFVG--n--------------Ins
1KHVA  ( 397 )  nkI--sFlkrtFelt----d---igwvSkl--d--------------kss
1O5SA  ( 808 )  gpH--eFcsqhTmlv---kqgddyvylpYp--d--------------psr
1QZ0A  ( 380 )  rnmye--sk----gs--sAvaddsklrpVI--H--------------Cra
1RDR0  ( 369 )  enVtF--LkRfFrad--e--kypflihPvM--p--------------mke
1S4FA  ( 492 )  edI--eFcshtPVpVrWs--dntsshMAGr--d--------------TAv
1SH0A  ( 386 )  ngl--tFlrRtVtrd----p---agwfGkl--e--------------qss
3HVTA  ( 449 )  e-------tklGkAGYvTnk-grgkvvpltntt-----------nqkTeL
                           bbbb              bbb              aaa

1C2PA  ( 394 )  rAAwET--Arhtp----v-ns-WLgniIyApTl-WArILTHFFsilLaqe
1DF0A  ( 535 )  rrlFaqlA-----gedaeIsafeLqtiLrrvlakksdgFsieTCkiMVdM
1HHSA  ( 521 )  MLnnqFSpeygVqs---------------gvrdrSkrkrPFpGLawaSm
1KHVA  ( 422 )  IlrqLE--WSktt----srhm-vieetydLake-eRgvqLeeLQVaAA--
1O5SA  ( 837 )  IlGagC--fvddi----vk-t-d--gt-------lmierFvsLaidAy--
1QZ0A  ( 406 )  GVg----------------------------RTAQLIGAMCMnds--
1RDR0  ( 397 )  IhesIr--wTkdp------------------r-ntqdhVrsLCllAW--
1S4FA  ( 522 )  ILSkMA--trldss--gergtt-aYEKAVAFSFLLmYSWNplVrrIClLVL
1SH0A  ( 411 )  IlrqMy--Wtrg----pnhe-dPsetmi-phs-qrpiqLmsLLGeAA--
3HVTA  ( 480 )  qAIylALqdsgleV-nIvTDs-qyAlgiIqagPdkseselVnqIIeqLik
                aaaaa                                    aaaaaaaa

1C2PA  ( 438 )  qlekaldCqiygAcYsi--ePldLPQIIerLhglsAFsLh----syspgE
1DF0A  ( 584 )  ldedg-------sgkLglkefyiLwtkIqkYqkIYreIdvdrsgtMn-Sy
1HHSA  ( 555 )  k---------dtygaCp--IysdVleAIerCWwnafg---------esYr
1KHVA  ( 462 )  Ahg-------------qefFnfVcrELerQq-ayT--qF----svysyd
1O5SA  ( 868 )  plt-------------khpnqeyadvfhlyl-qyi--rk
1QZ0A  ( 423 )  rNs-------------qlSVedMVsqMRvqR--NGiM------Vqk--d
1RDR0  ( 423 )  Hng-------------eeeYnkFLakIrsvp-iGral------lLpeys
1S4FA  ( 568 )  sqqpetdPskhAtyyyk----gdpigaYkdvIgrnlseLk----rtgfek
1SH0A  ( 449 )  Lhg-------------pafYskIsklViael-keggmdf----yvprqe
3HVTA  ( 528 )  kekVyLawvpahkgig---gNeqVdklVsagi
                                aaaaaaaaaaa
```

Fig. 1. (*Continued*).

we have incorporated multiple bioinformatic tools to investigate the structural relationships between ATA and RdRps from SARS-CoV and other pathogenic positive-strand RNA viruses (with positive controls). Such a study attempts to reveal the potential binding modes/sites of ATA onto RdRps from these viruses. The positive controls used constituted of three known ATA's protein targets: (1) the Ca^{2+}-activated neutral protease (*m*-calpain) (Posner et al., 1995), (2) the protein tyrosine phosphatase (PTP) of Yersinia pestis (Liang et al., 2003; Sun et al., 2003) and (3) HIV integrase (Cushman and Sherman, 1992), which have existing 3D crystal structures. Eventually, current study attempts to elucidate the plausible conformations for the inhibition mode of ATA, and hence provide clues to rational anti-SARS drug design based on ATA.

2. Materials and methods

In current study, the protein interactions with small molecule, ATA, were made up of theoretically-determined (six structures) and experimentally-determined (13 structures) structures as stated the following. The three-dimensional theoretical models for RNA-dependent RNA polymerase (RdRp, ID = 1O5S) (Xu et al., 2003), spike protein subunit 1 (S_1, ID = 1Q4Z) (Spiga et al., 2003; Zhang and Yap, 2004), spike protein subunit 2 (S_2, ID = 1Q4Y) (Spiga et al., 2003; Zhang and Yap, 2004) of SARS-CoV, the 3-dimensional crystal structures for nucleocapsid protein (N protein, ID = 1SSK) (Huang et al., 2004), non-structural protein 9 (Nsp9, ID = 1QZ8) (Egloff et al., 2004), 3C-like main protease (3Clpro, ID = 1Q2W) (Anand et al., 2003) of SARS-CoV, the crystal structure for YopH from Yersinia pestis (ID = 1QZ0) (Sun et al., 2003), *m*-calpain from Rattus norvigecus (ID = 1DF0) (Strobl et al., 2000), RNA-dependent RNA polymerase from Dsrna bacteriophage φ6 (ID = 1HHS), Rabbit hemorrhagic disease virus (ID = 1KHW), Poliovirus (ID = 1RDR), Bovine viral diarrhea virus (IVDV, ID = 1S4F), Norwalk virus (ID = 1SH0), Hepatitis C virus (ID = 1C2PA) and HIV reverse transcriptase (ID = 3HVT) were retrieved from Protein Data Bank. The three-dimensional geometry of all the molecules were validated for restraint violations using PROCHECK (Laskowski et al., 1996). The three-dimensional coordinate for the small molecule Aurintricarboxylic Acid (ATA—$C_{22}H_{14}O_9$) was derived using PRODRG2 (Schuttelkopf and Aalten, 2004) and JME editor (http://www.cem.msu.edu/~reusch/VirtualText/Questions/MOLEDITOR/jme_window.html). VAST and DALI programs were used to locate similar structural patterns between crystal structure of all molecules. Sequential structural alignment was done by CE (Shindyalov and Bourne, 1998) and COMPARER (Sali and Blundell, 1990). Finally, three-dimensional structural comparative analysis was performed by LGA (Zemla, 2003). Preparation of Macromolecule and ligand prior molecular docking was done using WhatIF software (Vriend, 1990). Molecular docking to determine the best conformation in term of lowest Gibbs free energy and shape complementary was performed using Autodock 3.0 (Morris et al., 1998). The visualisation of the 3-dimensional structural data was generated by Rasmol (Bernstein, 2000).

3. Results

3.1. Binding motifs

Protein 3D structural alignments were performed on 10 amino acid (mainly RdRps) sequences of interest, namely: (1) Hepatitis C virus RNA-dependent RNA polymerase; (2) *m*-calpain; (3) bacteriophage phi6 RNA-dependent RNA polymerase; (4) Rabbit hemorrhagic disease virus RNA-dependent RNA polymerase; (5) SARS-CoV RNA-dependent RNA polymerase; (6) Yersinia pestis phosphatase yoph; (7) Poliovirus RNA-dependent RNA polymerase; (8) Bovine viral diarrhea virus RNA-dependent RNA polymerase; (9) Norwalk virus RNA-dependent RNA polymerase; (10) HIV Type 1 RNA-dependent RNA polymerase. Regions with homologous 3D-conformation were identified together with their conserved secondary structures (underneath their amino acid alignment in Fig. 1). In total, there are eight structurally conserved motif blocks (CMBs) with each block extending at least eight amino acid residues. The secondary structures for all CMBs include six α-helices and two β-strands regions. The exact positions of all CMBs in each protein are provided in Fig. 1.

The molecular docking method based on free energy of ligand binding, ΔG, between ATA and all proteins (Figs. 2 and 3) revealed that ATA binds favorably to one structurally conserved region ($R_{binding}$) among all proteins (Fig. 1 in red box). As shown in Fig. 4, the corresponding $R_{binding}$ region in SARS-CoV's polymerase (Ser754–Tyr766) overlapped with one CMB. $R_{binding}$ is located in the palm sub-domain and constituted mainly of anti-parallel β-strand-turn-β-strand hairpin structures. This conserved region is similar to the majority of the remaining nine proteins in term of their secondary structures. Surprising, this $R_{binding}$ region also contains a highly conserved 'XSDD' amino acid motif that is especially prominent among viral RdRps, of which two of the highly conserved aspartic acid (D) residues form the catalytic center important for polymerase activity (Xu et al., 2003).

3.2. Binding energy

The free energy of ligand binding (final intermolecular energy + torsional free energy), ΔG, between ATA and all proteins were compared (Table 1). For the proteins that were documented to be inhibited by ATA molecule were assigned as positive controls (HIV integrase, yoph and *m*-calpain), their estimated free energies of binding were −11.88 kcal/mol, −7.79 kcal/mol and −7.67 kcal/mol,
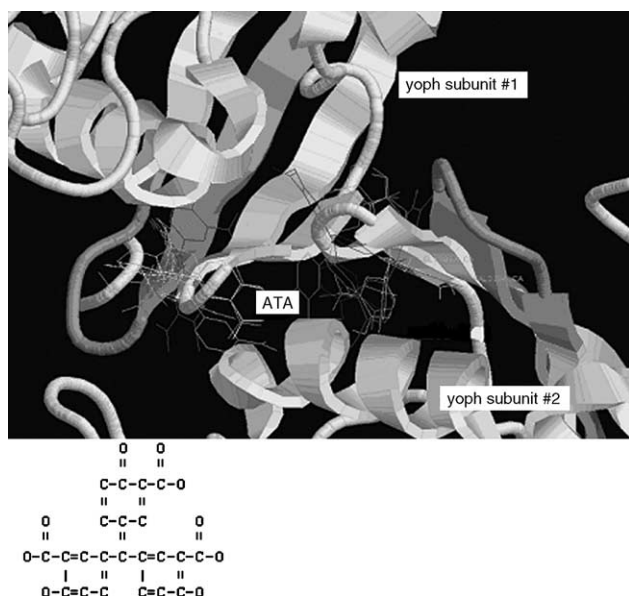
Fig. 2. Calculated structure (using Autodock) for the interaction of ATA with ypoH. YpoH is a protein tyrosine phosphatase which is essential for virulence in the Yersinia pestis. It is known that the functionality of this protein is inhibited strongly by ATA molecule (Liang et al., 2003). This figure shows the 10 most possible confirmations of ypoH-ATA complexes. The border residues that have contact with ATA are highlighted in red. This region is structurally conserved between ypoH, *m*-calpain and SARS-CoV RdRps, and constituted mainly of anti-parallel β-strand-turn-β-strand hairpin structures. Below, the two dimensional chemical structure for ATA ($C_{22}H_{14}O_9$) was shown.
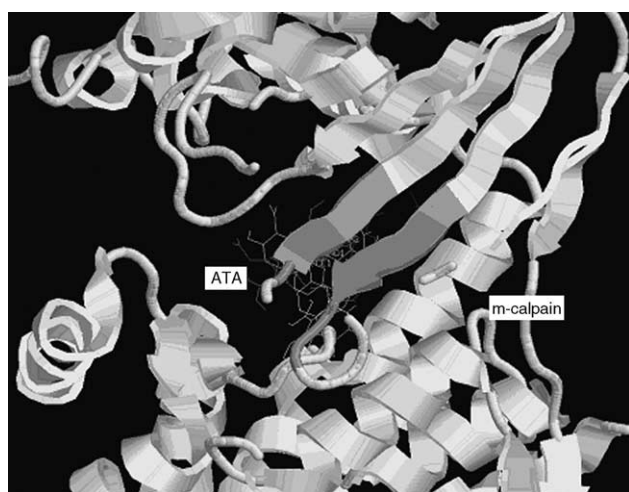
Fig. 3. Calculated structure (using Autodock) for the interaction of ATA with *m*-calpain. The neutral protease (calpain) is a class of cytosolic enzyme that is activated during apoptosis. It is known to be inhibited strongly by ATA molecule (Posner et al., 1995). This figure shows the 10 most possible confirmations of *m*-calpain-ATA complexes. The border residues that have contact with ATA are highlighted in red. This region is structurally conserved between ypoH, *m*-calpain and SARS-CoV RdRps, and constituted mainly of anti-parallel β-strand-turn-β-strand hairpin structures.
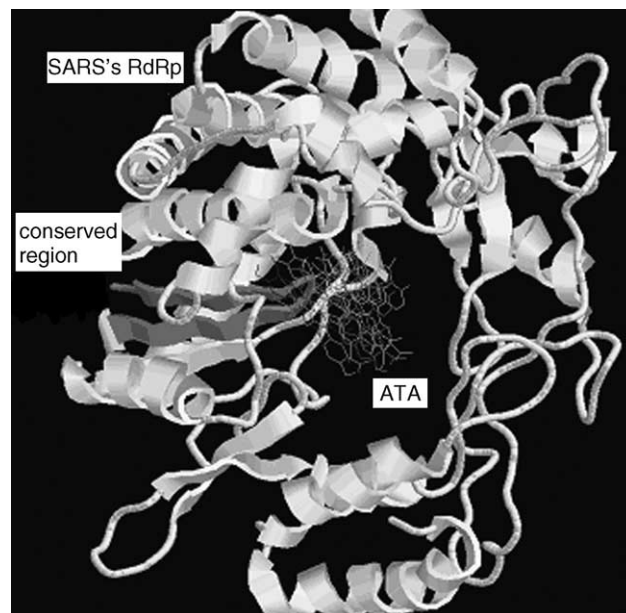

Fig. 4. Calculated structure (using Autodock) for the interaction of ATA with RdRp of SARS-CoV. This figure shows the 10 most possible confirmations of RdRp-ATA complexes. The structurally conserved residues between *ypoH*, *m*-calpain and SARS-CoV RdRp are colored red, and constituted mainly of anti-parallel β-strand-turn-β-strand hairpin structures.

respectively. Any estimated free energy around/lower than −7.67 kcal/mol will potentially suggest similar inhibitory binding mechanism by ATA molecule, if the corresponded binding motif incorporates catalytic sides of that specific protein. When we studied the binding of ATA onto various RNA dependent RNA polymerases (RdRps) from other genomes, the free energy of binding for most RdRps were significantly higher than −7.67 kcal/mol, and therefore suggested a lower strength of inhibition upon binding by ATA molecule (Bovine viral diarrhea virus, Dsrna bacteriophage, Feline calicivirus, Hepatitis C virus, HIV, Poliovirus and Rabbit hemorrhagic disease virus). Only the RdRps from SARS-CoV ($\Delta G = -7.68$) and Norwalk virus ($\Delta G = -14.92$) were estimated to have a lower ΔG (≥ 7.67 kcal/mol). This potentiated the possibility that ATA molecule could inhibit strongly the polymerase activity of these two RdRps as in HIV integrase, yoph and *m*-calpain.

4. Discussion

In summary, we first identified the homologous amino acid regions shared by 10 proteins (CMBs) using 3D structural alignments. Of these 10 proteins studied, three proteins were documented to be strongly inhibited by ATA molecules. In parallel, we employed the molecular docking method (based on free energy of ligand binding) to predict the most energy favorable binding conformation of ATA onto these 10 proteins and eventually determined one binding pocket ($R_{binding}$)

Table 1
Estimated free energy (final intermolecular energy + torsional free energy) of different targeted proteins in complex with ATA using Autodock

Targeted protein	Type of structure used	Estimated free energy of binding in complex with ATA (kcal/mol)
RNA dependent RNA polymerase		
RdRp (Bovine viral diarrhea virus)	Crystal structure	−5.82
RdRp (Dsrna bacteriophage)	Crystal structure	−5.87
RdRp (Feline calicivirus)	Theoretical model	−5.83
RdRp (Hepatitis C virus)	Crystal structure	−5.8
RdRp (HIV)	Crystal structure	−5.82
RdRp (Norwalk virus)	Crystal structure	−14.92
RdRp (Poliovirus)	Crystal structure	−5.83
RdRp (Rabbit hemorrhagic disease virus)	Crystal structure	−5.94
RdRp (SARS-CoV)	Theoretical model	−7.68
Protein known to be inhibited by ATA[a]		
Yoph[a]	Crystal structure	−7.79
HIV integrase[a]	Crystal structure	−11.88
m-Calpain[a]	Crystal structure	−7.67
Other SARS' proteins		
Main protease (3CL)	Crystal structure	−7.78
N protein	Crystal structure	−8.32
Nsp 9	Crystal structure	−7.64
S_1 (Spiga et al., 2003)	Theoretical model	−18.59
S_2 (Spiga et al., 2003)	Theoretical model	−7.66
S_1 (Zhang and Yap, 2004)	Theoretical model	−14.79
S_2 (Zhang and Yap, 2004)	Theoretical model	−15.22

[a] Cushman and Sherman (1992), Liang et al. (2003), Posner et al. (1995).

recognized by ATA, that also overlapped with CMBs. This binding pocket ($R_{binding}$), denoted to be 754–766 in SARS-CoV's RdRp, not only shared a highly conserved secondary structure with other RdRps, but also incorporates two of the three predicted RdRp catalytic residues (Asp 760, Asp761) identified by other researchers (Xu et al., 2003). These residues are expected to be important for metal ion chelation (Beese and Steitz, 1991; Bressanelli et al., 2002; Huang et al., 1998). Structurally, this binding pocket is located in the palm sub-domain constituted mainly of anti-parallel β-strand-turn-β-strand hairpin structures.

As reported in Section 3, there are altogether eight structurally conserved motif blocks (CMBs) and their secondary structures include six α-helices and two β-strands regions. Among these structurally conserved regions, we subsequently identified that there was one common region recognized by ATA. The binding of ATA to this region also fulfilled the lowest free energy of ligand binding. By using this free energy of ligand binding, we were able to quantify the binding strength between a macromolecule and a ligand. Therefore, if a ligand binds strongly (with lower free energy of ligand binding ΔG) to the active domains of one specific protein, it will presumably inhibit the activity for that specific protein. On the other hand, if ATA does not bind to the active domains of the protein, ATA will not be able to inhibit the function of the protein regardless of the strength of binding (e.g. 3CL main protease and ATA). Based on this rationale, it is likely that ATA could have inhibited the function of RdRp of SARS-CoV because the binding pocket for ATA and RdRp includes two strictly conserved aspartate residues (D)

involved in binding divalent metal ions required for RdRp catalysis (Beese and Steitz, 1991; Bressanelli et al., 2002; Huang et al., 1998). When ATA binds to this β-strand-turn-β-strand hairpin structures ($R_{binding}$), it covers the two aspartate residues and therefore inhibits the metal ion chelation process crucial for the functionality of this protein during viral replication. In contrast, although we found out that ATA binds to 3CL main protease at a much stronger binding strength (Table 1), we do not anticipate a similar inhibitory binding because ATA does not bind to the active domains of 3CL main protease (Yang et al., 2003).

Furthermore, we also predicted the ATA binding regions for Ca^{2+}-activated neutral protease of Rattus norvegicus (*m*-calpain), the protein tyrosine phosphatase (PTP) of Yersinia pestis and HIV integrase of Human immunodeficiency virus 1, of which all were experimentally proven to be inhibited by ATA and contain existing crystal structures. However, to the best of our knowledge, there are no literatures attempting to explain explicitly how ATA inhibits the proteins mentioned above. Surprisingly, we performed molecular docking analysis and predicted that ATA binds to the region of these three proteins (Fig. 1) that have correspondence to $R_{binding}$ in SARS-CoV and other viral RdRps. This region (Fig. 1 in box) was identified to be highly conserved among viral RdRps, in term of their 3D structures and secondary structures. In essence, ATA selectively binds to the β-strand-turn-β-strand hairpin structures. Such a strong conservation signifies that this region might be evolutionarily crucial and important for the function of the proteins. Therefore, it remains plausible that ATA inhibits the function of SARS-CoV's RdRp by bind-

ing to the homologous region as in *m*-calpain, PTP and HIV integrase.

Finally, we had also analyzed the binding of ATA onto viral RdRps from the standpoint of free energy of ligand binding (ΔG). The results of ΔG between ATA and all viral RdRps, as well as the three proteins known to be readily inhibited by ATA molecules (Table 1), suggested that ATA binds to RdRp from SARS-CoV at the equivalent binding strength similar to yoph and *m*-calpain. On the other hand, ATA was predicted to bind to the remaining viral RdRps (except for Norwalk virus) at a much lower binding strength whereas ATA binds to RdRp from Norwalk virus at a significantly higher binding strength (approx. two times). If our hypothesis that ATA inhibits the functionality of RdRps by binding to its catalytic sites is indeed true, these numerical representations of ΔGs would likely represent how well ATA will inhibit the viral replication inside virus-infected cell culture. To prove our hypothesis, we had also carried out some experimental study on how well ATA interrupts the viral replication in SARS-infected cell culture and demonstrated successfully that the viral load (e.g. HCV, West Nile virus and Feline calicivirus, etc.), with the addition of ATA at the highest concentration, corresponds to the ranking of the estimated free energy of ligand binding (experimental data to be published).

5. Conclusions

The following conclusions can be drawn from this structural study on the binding of ATA to proteins from SARS-CoV and other positive-strand RNA viruses:

1. ATA could bind to SARS-CoV's RdRp and other SARS-CoV's proteins. The strongest binding was predicted to be for S_1 protein. By binding to a protein, ATA could inhibit their activity if the ATA binding motifs incorporate catalytic domains of that specific protein, for example: ATA binds to the catalytic sites of SARS-CoV's RdRp and inhibit their function.
2. Derived from the inhibitory binding of ATA on SARS-CoV's RdRp, ATA was also predicted to inhibit the polymerase activity of RdRps from other RNA viruses in a weaker manner (Bovine viral diarrhea virus, Dsrna bacteriophage, Feline calicivirus, Hepatitis C virus, HIV, Poliovirus, Rabbit hemorrhagic disease virus). However, stronger inhibition (approx. two times) was predicted for the RdRp from Norwalk virus (experimental data to be published).
3. This inhibitory binding mechanism, especially for RdRps, could serve to explain why mRNA transcripts were decreased by >1000 times when SARS-infected cell culture was treated with ATA in our previous study. It could also explicate why the inhibition on the replication on other RNA viruses are milder when virus-infected cells was treated with ATA (experimental data to be published).
4. ATA could serve as a template for rational anti-SARS drug design, targeting the replication machinery possibly over large groups of positive-strand RNA viruses.

References

Schuttelkopf, A.W., van Aalten, D.M., 2004. PRODRG: a tool for high-throughput crystallography of protein-ligand complexes. Acta Crystallogr. D. Biol. Crystallogr. 60(Pt 8), 1355–1363.

Akiyama, S., Kuwano, M., Yamamoto, M., Endo, H., 1977. The effect of aurintricarboxylic acid on RNA polymerase from rat liver. J. Biochem. (Tokyo) 81, 135–141.

Anand, K., Ziebuhr, J., Wadhwani, P., Mesters, J.R., Hilgenfeld, R., 2003. Coronavirus main proteinase (3CLpro) structure: basis for design of anti-SARS drugs. Science 300, 1763–1767.

Beese, L.S., Steitz, T.A., 1991. Structural basis for the 3′–5′ exonuclease activity of *Escherichia coli* DNA polymerase I: a two metal ion mechanism. EMBO. J. 10, 25–33.

Bermejo Martin, J.F., Jimenez, J.L., Munoz-Fernandez, A., 2003. Pentoxifylline and severe acute respiratory syndrome (SARS): a drug to be considered. Med. Sci. Monit. 9, SR29–SR34.

Bernstein, H.J., 2000. Recent changes to RasMol, recombining the variants. Trends Biochem. Sci. 25, 453–455.

Bressanelli, S., Tomei, L., Rey, F.A., De Francesco, R., 2002. Structural analysis of the hepatitis C virus RNA polymerase in complex with ribonucleotides. J. Virol. 76, 3482–3492.

Chou, K.C., Wei, D.Q., Zhong, W.Z., 2003. Binding mechanism of coronavirus main proteinase with ligands and its implication to drug design against SARS. Biochem. Biophys. Res. Commun. 308, 148–151.

Cushman, M., Sherman, P., 1992. Inhibition of HIV-1 integration protein by aurintricarboxylic acid monomers, monomer analogs, and polymer fractions. Biochem. Biophys. Res. Commun. 185, 85–90.

Dhanak, D., Duffy, K.J., Johnston, V.K., Lin-Goerke, J., Darcy, M., Shaw, A.N., Gu, B., Silverman, C., Gates, A.T., Nonnemacher, M.R., Earnshaw, D.L., Casper, D.J., Kaura, A., Baker, A., Greenwood, C., Gutshall, L.L., Maley, D., DelVecchio, A., Macarron, R., Hofmann, G.A., Alnoah, Z., Cheng, H.Y., Chan, G., Khandekar, S., Keenan, R.M., Sarisky, R.T., 2002. Identification and biological characterization of heterocyclic inhibitors of the hepatitis C virus RNA-dependent RNA polymerase. J. Biol. Chem. 277, 38322–38327.

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., Escriou, 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., 2003. Identification of a novel coronavirus in patients with severe acute respiratory syndrome. N. Engl. J. Med. 348, 1967–1976.

Egloff, M.P., Ferron, F., Campanacci, V., Longhi, S., Rancurel, C., Dutartre, H., Snijder, E.J., Gorbalenya, A.E., Cambillau, C., Canard,

B., 2004. The severe acute respiratory syndrome-coronavirus replicative protein nsp9 is a single-stranded RNA-binding subunit unique in the RNA virus world. Proc. Natl. Acad. Sci. USA 101, 3792–3796.

Givens, J.F., Manly, K.F., 1976. Inhibition of RNA-directed DNA polymerase by aurintricarboxylic acid. Nucleic Acids Res. 3, 405–418.

Glansbeek, H.L., Haagmans, B.L., te Lintelo, E.G., Egberink, H.F., Duquesne, V., Aubert, A., Horzinek, M.C., Rottier, P.J., 2002. Adverse effects of feline IL-12 during DNA vaccination against feline infectious peritonitis virus. J. Gen. Virol. 83, 1–10.

He, R., Adonov, A., Traykova-Adonova, M., Cao, J., Cutts, T., Grudesky, E., Deschambaul, Y., Berry, J., Drebot, M., Li, X., 2004. Potent and selective inhibition of SARS coronavirus replication by aurintricarboxylic acid. Biochem. Biophys. Res. Commun. 320, 1199–1203.

Huang, H., Chopra, R., Verdine, G.L., Harrison, S.C., 1998. Structure of a covalently trapped catalytic complex of HIV-1 reverse transcriptase: implications for drug resistance. Science 282, 1669–1675.

Huang, Q., Yu, L., Petros, A.M., Gunasekera, A., Liu, Z., Xu, N., Hajduk, P., Mack, J., Fesik, S.W., Olejniczak, E.T., 2004. Structure of the N-terminal RNA-binding domain of the SARS CoV nucleocapsid protein. Biochemistry 43, 6059–6063.

James, J.S., 2003. SARS Web information. AIDS Treat. News, 6.

Koren, G., King, S., Knowles, S., Phillips, E., 2003. Ribavirin in the treatment of SARS: a new trick for an old drug. CMAJ 168, 1289–1292.

Laskowski, R.A., Rullmannn, J.A., MacArthur, M.W., Kaptein, R., Thornton, J.M., 1996. AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. J. Biomol. NMR 8, 477–486.

Liang, F., Huang, Z., Lee, S.Y., Liang, J., Ivanov, M.I., Alonso, A., Bliska, J.B., Lawrence, D.S., Mustelin, T., Zhang, Z.Y., 2003. Aurintricarboxylic acid blocks in vitro and in vivo activity of YopH, an essential virulent factor of Yersinia pestis, the agent of plague. J. Biol. Chem. 278, 41734–41741.

Marra, M.A., Jones, S.J., Astell, C.R., Holt, R.A., Brooks-Wilson, A., Butterfield, Y.S., Khattra, J., Asano, J.K., Barber, S.A., Chan, S.Y., Cloutier, A., Coughlin, S.M., Freeman, D., Girn, N., Griffith, O.L., Leach, S.R., Mayo, M., McDonald, H., Montgomery, S.B., Pandoh, P.K., Petrescu, A.S., Robertson, A.G., Schein, J.E., Siddiqui, A., Smailus, D.E., Stott, J.M., Yang, G.S., Plummer, F., Andonov, A., Artsob, H., Bastien, N., Bernard, K., Booth, T.F., Bowness, D., Czub, M., Drebot, M., Fernando, L., Flick, R., Garbutt, M., Gray, M., Grolla, A., Jones, S., Feldmann, H., Meyers, A., Kabani, A., Li, Y., Normand, S., Stroher, U., Tipples, G.A., Tyler, S., Vogrig, R., Ward, D., Watson, B., Brunham, R.C., Krajden, M., Petric, M., Skowronski, D.M., Upton, C., Roper, R.L., 2003. The Genome sequence of the SARS-associated coronavirus. Science 300, 1399–1404.

Marshall, E., Enserink, M., 2004. Medicine. Caution urged on SARS vaccines. Science 303, 944–946.

Morris, G.M., Goodsell, D.S., Halliday, R.S., HUEY, R., Hart, W.E., Belew, R.K., Olson, A.J., 1998. Automated docking using a lamarckian genetic algorithm and empirical binding free energy function. J. Comput. Chem. 19, 1639–1662.

Pace, P., Nizi, E., Pacini, B., Pesci, S., Matassa, V., De Francesco, R., Altamura, S., Summa, V., 2004. The monoethyl ester of meconic acid is an active site inhibitor of HCV NS5B RNA-dependent RNA polymerase. Bioorg. Med. Chem. Lett. 14, 3257–3261.

Posner, A., Raser, K.J., Hajimohammadreza, I., Yuen, P.W., Wang, K.K., 1995. Aurintricarboxylic acid is an inhibitor of mu- and *m*-calpain. Biochem. Mol. Biol. Int. 36, 291–299.

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., 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., Drosten, C., Pallansch, M.A., Anderson, L.J., Bellini, W.J., 2003. Characterization of a novel coronavirus associated with severe acute respiratory syndrome. Science 300, 1394–1399.

Sali, A., Blundell, T.L., 1990. Definition of general topological equivalence in protein structures. A procedure involving comparison of properties and relationships through simulated annealing and dynamic programming. J. Mol. Biol. 212, 403–428.

Sarisky, R.T., 2004. Non-nucleoside inhibitors of the HCV polymerase. J. Antimicrob. Chemother. 54, 14–16.

Shindyalov, I.N., Bourne, P.E., 1998. Protein structure alignment by incremental combinatorial extension (CE) of the optimal path. Protein Eng. 11, 739–747.

Spiga, O., Bernini, A., Ciutti, A., Chiellini, S., Menciassi, N., Finetti, F., Causarono, V., Anselmi, F., Prischi, F., Niccolai, N., 2003. Molecular modelling of S1 and S2 subunits of SARS coronavirus spike glycoprotein. Biochem. Biophys Res. Commun. 310, 78–83.

Strobl, S., Fernandez-Catalan, C., Braun, M., Huber, R., Masumoto, H., Nakagawa, K., Irie, A., Sorimachi, H., Bourenkow, G., Bartunik, H., Suzuki, K., Bode, W., 2000. The crystal structure of calcium-free human *m*-calpain suggests an electrostatic switch mechanism for activation by calcium. Proc. Natl. Acad. Sci. USA 97, 588–592.

Sun, J.P., Wu, L., Fedorov, A.A., Almo, S.C., Zhang, Z.Y., 2003. Crystal structure of the Yersinia protein-tyrosine phosphatase YopH complexed with a specific small molecule inhibitor. J. Biol. Chem. 278, 33392–33399.

Swennen, L., Moens, L., Kondo, M., 1978. Aurintricarboxylic acid stimulates specifically DNA-dependent RNA polymerase II-directed RNA synthesis in isolated larval nuclei of Artemia saline. Arch. Int. Physiol. Biochim. 86, 925–926.

Swennen, L., Moens, L., Kondo, M., 1979. Mechanism of preferential stimulation of DNA-dependent RNA polymerase II by aurintricarboxylic acid in isolated Artemia larval nuclei. Arch. Int. Physiol. Biochim. 87, 842–843.

Takeda-Shitaka, M., Nojima, H., Takaya, D., Kanou, K., Iwadate, M., Umeyama, H., 2004. Evaluation of homology modeling of the severe acute respiratory syndrome (SARS) coronavirus main protease for structure based drug design. Chem. Pharm. Bull (Tokyo) 52, 643–645.

Vriend, G., 1990. WHAT IF: a molecular modeling and drug design program. J. Mol. Graph 8, 52–56.

Wu, C.Y., Jan, J.T., Ma, S.H., Kuo, C.J., Juan, H.F., Cheng, 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., 2004. Small molecules targeting severe acute respiratory syndrome human coronavirus. Proc. Natl. Acad. Sci. USA 101, 10012–10017.

Xu, X., Liu, Y., Weiss, S., Arnold, E., Sarafianos, S.G., Ding, J., 2003. Molecular model of SARS coronavirus polymerase: implications for biochemical functions and drug design. Nucleic Acids Res. 31, 7117–7130.

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., 2003. The crystal structures of severe acute respiratory syndrome virus main protease and its complex with an inhibitor. Proc. Natl. Acad. Sci. USA 100, 13190–13195.

Zemla, A., 2003. LGA: a method for finding 3D similarities in protein structures. Nucleic Acids Res. 31, 3370–3374.

Zhang, X.W., Yap, Y.L., 2004a. The 3D structure analysis of SARS-CoV S1 protein reveals a link to influenza virus neuraminidase and implications for drug and antibody discovery. J. Mol. Struct.: THEOCHEM 681, 137–141.

Zhang, X.W., Yap, Y.L., 2004b. Exploring the binding mechanism of the main proteinase in SARS-associated coronavirus and its implication to anti-SARS drug design. Bioorg. Med. Chem. 12, 2219–2223.

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

Zhang, X.W., Yap, Y.L., 2004d. Structural similarity between HIV-1 gp41 and SARS-CoV S2 proteins suggests an analogous membrane fusion mechanism. J. Mol. Struct.: THEOCHEM 677, 73–76.