Computational Studies on the Substrate Interactions of Influenza A Virus PB2 Subunit

Ya-Jun Wang[1], Jing-Fang Wang[1,2]*, Jie Ping[1], Yao Yu[2,4], Ying Wang[1], Peng Lian[1], Xuan Li[2,4], Yi-Xue Li[2,3], Pei Hao[2,5]*

1 Key Laboratory of Systems Biomedicine (Ministry of Education), Shanghai Jiao Tong University, Shanghai, China, 2 Shanghai Center for Bioinformation and Technology, Shanghai, China, 3 Bioinformatics Center, Key Laboratory of Systems Biology, Chinese Academy of Sciences, Shanghai, China, 4 Institute of Plant Physiology and Ecology, Chinese Academy of Sciences, Shanghai, China, 5 Institute of Pasteur, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai, China

Abstract

Influenza virus, which spreads around the world in seasonal epidemics and leads to large numbers of deaths every year, has several ribonucleoproteins in the central core of the viral particle. These viral ribonucleoproteins can specifically bind the conserved 3′ and 5′ caps of the viral RNAs with responsibility for replication and transcription of the viral RNA in the nucleus of infected cells. A fundamental question of most importance is that how the cap-binding proteins in the influenza virus discriminates between capped RNAs and non-capped ones. To get an answer, we performed molecular dynamics simulations and free energy calculations on the influenza A virus PB2 subunit, an important component of the RNP complexes, with a cap analog m7GTP. Our calculations showed that some key residues in the active site, such as Arg355, His357, Glu361 as well as Gln406, could offer significant hydrogen bonding and hydrophobic interactions with the guanine ring of the cap analog m7GTP to form an aromatic sandwich mechanism for the cap recognition and positioning in the active site. Subsequently, we applied this idea to a virtual screening procedure and identified 5 potential candidates that might be inhibitors against the PB2 subunit. Interestingly, 2 candidates Cpd1 and Cpd2 have been already reported to have inhibitory activities to the influenza virus cap-binding proteins. Further calculation also showed that they had comparatively higher binding affinities to the PB2 subunit than that of m7GTP. We believed that our findings could give an atomic insight into the deeper understanding of the cap recognition and binding mechanism, providing useful information for searching or designing novel drugs against influenza viruses.

Introduction

Influenza, commonly referred to the flu, is an acute viral-infection disease caused by a number of RNA viruses of the family Orthomyxoiridae (also known as influenza viruses) [1]. Typically, influenza viruses are transmitted through the air by coughs or sneezes, creating aerosols containing the viruses, or through direct contact with bird droppings or nasal secretions, or through contact with contaminated surfaces [2,3]. Nowadays, influenza virus spreads around the world in seasonal epidemics, leading to 25,000–500,000 deaths every year, which will be up to millions in the pandemic years [4,5]. Although having a number of subtypes, influenza viruses share a similar overall structure: the virus particle is roughly spherical with a diameter of about 80–120 nm [6]. The viral envelope contains a proton channel and two glycoproteins, wrapped around the central core, which contains the viral RNA genome and other viral proteins [7,8].

In the past few years, some powerful antiviral drugs have been developed to treat and prevent influenza infection targeted on the proteins in the viral envelope [9,10,11,12]. These antiviral drugs can be clustered into two major types: neuraminidase inhibitors (i.e., oseltamivir and zanamivir) and proton channel inhibitors (i.e., amantadine and rimantadine). Currently, neuraminidase inhibitors are preferred for influenza virus infections since they are less toxic and more effective [13]. However, increased resistance has been detected in patients with this kind of antiviral drugs [14,15]. Since then, a series of good attempts have been made by experimental and theoretical approaches to study the structural mechanism of drug inhibition and resistance for these antiviral drugs, with an aim of searching for an effective approach to prevent the known drug resistance [16–21]. However, to avoid the known resistance, an alter strategy is to develop novel antiviral drugs targeting on other proteins (or RNA) in the central core of influenza viruses, i.e., the polymerase complex of influenza viruses that is found to be essential for viral replication.

For influenza A viruses, the viral genome in the central core of the viral particle contains 8 single-stranded RNA segments of negative polarity with partially complementary ends, encoding totally 11 important viral proteins. Each single-stranded RNA segment can form several ribonucleoprotein (RNP) complexes via

the association with multiple monomers of the nucleoprotein (NP) and one single copy of the viral RNA-dependent RNA polymerase composed of three subunits: one polymerase acidic protein PA, and two polymerase basic proteins PB1 and PB2 [22,23]. The RNP complexes can bind the conserved 3′ and 5′ caps of each viral RNA segment, and are responsible for replication and transcription of the viral RNA in the nucleus of infected cells. Host-cell pre-mRNA is bound to the PB2 subunit by its 5′ caps, which is also considered as the initial step of viral mRNA transcription [24,25].

In 2008, Guilligay and his co-workers released an atomic-resolution structure of influenza A virus PB2 cap binding domain (residues 318–483) with bound m7GTP and provided functional analysis to show that the cap-binding site is essential for cap-dependent transcription by viral RNPs in vitro and in vivo [26]. They also suggested that PB2 cap binding domain is structurally distinct from other cap-binding proteins, and will be a good target for developing novel antiviral drugs. However, deeper under-standing of the structural flexibility and its interactions with 5′ cap RNAs is still needed. In comparison with the crystal and EM studies of the influenza virus PB2 subunit, computational approaches, especially molecular dynamics (MD) simulations and free energy calculations, have advantages to analyze the confor-mational fluctuations of the PB2 subunit and its interactions with RNAs [27,28,29]. In the current study, we employed MD simulations and free energy calculations to give an atomic description to the interactions of the influenza A virus PB2 subunit with cap analog m7GTP. Based on these structural analyses, we would provide a novel strategy for antiviral drug screening, which was further used to search for the potential leading compound in the related databases.

Materials and Methods

Initial structure and MD simulations

By now, only 1 crystal structure (2vqz.pdb) of the influenza A virus PB2 subunit is available in the Protein Data Bank [26], which was thus selected as the initial structure for the further calculations. Except for the polar hydrogen and heavy atoms of the protein and the ligand, all the other atoms including non-polar hydrogen in the crystal structure and the crystal water molecules were deleted. The pKa values for each residue in the influenza A virus PB2 subunit were computed using Delphi [30] as a Poisson-Boltzmann solver with a dielectric constant of 4. Hydrogen atoms were then added to the protein structure with t-Leap procedure in AMBER 11 package [31] based on the computational pKa values. Subse-quently the simulation systems were solvated in a simulation box with explicit TIP3P water models. To neutralize the solvated systems, chloride ions were added to random place equal number of water molecules in the simulation box. All the atoms of the influenza A virus PB2 subunit and its ligands were parameterized by Amber force field parameters [32].

After solvation, the simulation systems were subjected to steepest descent energy minimization for about 5000 steps followed by conjugate gradient for the next 5000 steps, and then equilibrated with the atoms in the protein and ligand fixed by a series short MD simulation (∼1 ns) to reduce the van der Waals conflicts in the simulation systems. Finally, 10-ns MD simulations were performed under the normal temperature (310 K) by AMBER 11 package [31] with periodic boundary conditions and NPT ensemble. The SHAKE algorithm with a tolerance of 10^{-6} was used to constrain all bonds in the simulation systems involving hydrogen atoms, and atomic velocities for start-up runs were obtained according to the Maxwell distribution at 310 K. 10

independent simulation trajectories with different starting vectors were for each simulation system were generated. The isothermal compressibility was set to be 4.5×10^{-5} per bar for solvent simulations. The electrostatic interactions were treated by PME algorithm with interpolation order of 4 and a grid spacing of 0.12 nm. The van der Waals interactions in our study were computed by using a cut-off value of 12 Å. All the MD simulations were run with a time step of 2 fs, and coordinates for the simulation systems were saved every 1 ps.

Free energy calculation

The molecular mechanics Poisson-Boltzmann surface area (MM-PB/SA) and Generalized Born surface area (MM-GB/SA) approaches [33] implemented in AMBER 11 package was applied to calculate the binding affinities for m7GTP and other ligands. Such approaches have been widely used to study protein folding [34,35], protein-DNA (or RNA) interactions [36,37], protein-ligand interaction [38,39,40], as well as structure-based drug design [41,42,43]. The principles of both MM-PB/SA and MM-GB/SA approaches can be summarized as following:

$$\Delta G_{bind} = G_{complex} - (G_{protein} + G_{ligand}) \qquad (1)$$

$$G \cong E_{gas} - TS_{config} + G_{sol} \qquad (2)$$

$$E_{gas} = E_{bond} + E_{angle} + E_{torsion} + E_{vdw} + E_{ele} \qquad (3)$$

$$G_{sol} = G_{ele} + G_{non-polar} \qquad (4)$$

In equation 1, the binding free energy change (ΔG) is computed as the difference between the free energies of the complex ($G_{complex}$), the protein ($G_{protein}$) as well as the ligand (G_{ligand}). These free energies are calculated through equation 2 by summing up its internal energy in the gas phase (E_{gas}), the solvation free energy (G_{sol}), and a vibrational entropy term ($T\Delta S$). The simulation systems involved in our study contained ∼150,000 atoms, while the variance of the atom number for different ligands was less than 100, indicating that the entropy differences for the same protein with different ligands was induced by less than 0.07% of the total atoms in the simulation system. Thus, we thought that the same protein with different ligands employed almost the same entropy, and the entropy contributions in the free energy calculations are neglected. E_{gas} is a standard force field energy calculated from equation 3 by the strain energies from covalent bonds (E_{bond} and E_{angle}) and torsion angles ($E_{torsion}$), non-covalent van der Waals (E_{vdw}), as well as electrostatic energies (E_{ele}). As described in equation 4, the solvation free energy (G_{sol}) is calculated by both an electrostatic term (G_{ele}) and a non-polar component ($G_{non-polar}$). The former can be obtained from either the Possion-Boltzmann (PB) approach or Generalized Born (GB) approach, while the latter is considered to be proportional to the molecular solvent accessible surface area. In our study, totally 200 snapshots retrieved from the last 1-ns segment on the MD trajectories with an interval of 5 ps were used for calculating the binding free energies.

Virtual screening procedure and multi-target selectivity

The cap analog m7GTP was selected as a structural template to search our in-house database which contained all the bioactive drug-like small molecules in ChEMBL and CNPD (Chinese

Natural Product Database, Neotrident Technology Ltd., China). MACCS keys [43] were employed to find the molecules that had highly similarity with the template m7GTP, and those having high similarity with the template m7GTP were further assessed by Lipinski's Rule Five [44]. The candidates in accordance with Lipinski's Rule Five were then subject to a flexible docking procedure using AutoDock 4.0 package, and the ones employed similar or higher docking scores were picked up for the molecular dynamics simulations and free energy calculations. During the docking procedure, a grid box with dimensions of $128 \times 128 \times 128$ cubic angstroms was used, and rotatable bonds were allowed for the candidates to rotate during docking operations. For each docking operation, 100 runs with 10 million energy evaluations were carried out using Lamarkian genetic algorithm.

The candidates found in the aforementioned procedures were then subjected to the multi-target selectivity studies. Totally 841 known drug targets were used in this multi-target selectivity procedure, which were collected from the Potential Drug Target Database (PDTD) [45]. The candidates were docked into all the targets using AutoDock 4.0 with Lamarkian genetic algorithm. The final results for each candidate were listed according to their docking scores.

Results and Discussion

Trajectory analysis and free energy calculation

The crystal structure of the influenza A virus PB2 subunit in the protein structure databases is the minimal cap binding domain (residues 318–483), in which a m7GTP molecule was also crystallized as a cap analog (Figure 1). During our MD simulations, RMS deviations for the Cα atoms of the PB2 cap-binding domain from the crystal structure were calculated (Figure S1), which was also believed as a crucial criterion for measuring the convergence of the simulation system concerned. For the current case, after about 800-ps MD simulations the PB2-m7GTP system reached to an equilibrated state with an average RMS
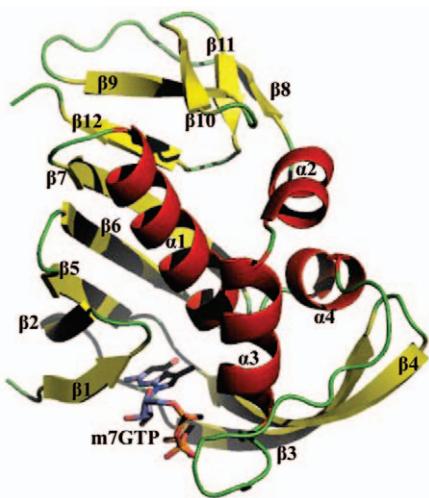
Figure 1. Illustratively showing the binding mode of the PB2 cap binding domain in the presence of the cap analog m7GTP. The protein structure was shown in a ribbon diagram, while m7GTP is in ball-and-stick representations. The secondary structure elements of the PB2 cap binding domain are labeled with α-helices in red and β-strands in yellow. The cap analog m7GTP is colored according to its atomic types.
doi:10.1371/journal.pone.0044079.g001

Table 1. Binding free energies (kcal/mol) of m7GTP and other ligands calculated from both MM-PB/SA and MM-GB/SA.

Energy term	PB2-m7GTP	PB2-CPD1	PB2-CPD2
ΔE_{ele}	-1048.00 ± 27.60	-39.20 ± 7.03	-25.42 ± 4.99
ΔE_{vdw}	-36.74 ± 4.09	-42.17 ± 3.41	-62.19 ± 3.45
ΔE_{gas}	-1084.75 ± 27.16	-81.37 ± 6.43	-87.62 ± 5.17
$\Delta G_{sur/PB}$	-5.86 ± 0.14	-5.68 ± 0.16	-7.08 ± 0.12
$\Delta G_{cal/PB}$	1048.68 ± 25.03	51.05 ± 4.98	55.58 ± 3.83
$\Delta G_{ele/PB}$	0.67 ± 7.86	11.85 ± 5.43	30.16 ± 4.41
$\Delta G_{sol/PB}$	1042.81 ± 25.03	45.37 ± 4.90	48.50 ± 3.80
$\Delta G_{bind/PB}$	-41.93 ± 6.52	-36.00 ± 4.70	-39.12 ± 4.20
$\Delta G_{sur/GB}$	-5.86 ± 0.14	-5.68 ± 0.16	-7.08 ± 0.12
ΔG_{GB}	1052.36 ± 25.19	42.05 ± 4.25	40.90 ± 3.45
$\Delta G_{ele/GB}$	4.36 ± 5.89	2.84 ± 4.26	15.48 ± 2.57
$\Delta G_{sol/GB}$	1046.50 ± 25.19	36.36 ± 4.21	33.82 ± 3.44
$\Delta G_{bind/GB}$	-38.25 ± 4.65	-45.01 ± 3.95	-53.80 ± 3.15

ΔE_{ele}: non-bonded electrostatic energy+1,4-electrostatic energy.
ΔE_{vdw}: non-bonded van der Waals energy+1,4-van der Waals energy
$\Delta E_{gas} = \Delta E_{ele} + \Delta E_{vdw}$
$\Delta G_{sur/PB}$: hydrophobic contributions to solvation free energy for the Possion Boltzmann calculations.
$\Delta G_{cal/PB}$: reaction field energy calculated by the Possion Boltzmann approach
$\Delta G_{sol/PB} = \Delta G_{sur/PB} + \Delta G_{cal/PB}$
$\Delta G_{ele/PB} = \Delta G_{cal/PB} + \Delta E_{ele}$
$\Delta G_{bind/PB} = \Delta E_{gas} + \Delta G_{sol/PB}$
$\Delta G_{sur/GB}$: hydrophobic contributions to solvatoion free energy for the Generalized Born calculations.
$\Delta G_{cal/PB}$: reaction field energy calculated by the Generalized Born approach
$\Delta G_{sol/GB} = \Delta G_{sur/GB} + \Delta G_{cal/GB}$
$\Delta G_{ele/GB} = \Delta G_{cal/GB} + \Delta E_{ele}$
$\Delta G_{bind/GB} = \Delta E_{gas} + \Delta G_{sol/GB}$
doi:10.1371/journal.pone.0044079.t001

deviation value of 1.55 ± 0.18 Å from the crystal structure (starting structure), which at the end of the simulations was ~1.53 Å. Additionally, the total energy of the PB2-m7GTP system in the equilibrated state was detected to be ~$6.742 \pm 0.017 \times 10^4$ kJ/mol. All these information gives an indication that our MD simulations are quite credible and the PB2-m7GTP system has been equilibrated. Thus, the free energy calculations discussed below were then performed based on the equilibrated simulation trajectories.

In the current study, we selected both MM-PB/SA and MM-GB/SA approach, the most popular methods for the protein-ligand free energy calculations, to compute the binding free energies of the PB2-m7GTP complex. The binding free energyies for the PB2-m7GTP complex were -41.93 ± 6.52 kcal/mol (MM-PB/SA) and -38.25 ± 4.65 (MM-GB/SA), as shown in Table 1. The variance of the binding free energies obtained from MM-PB/SA and MM-GB/SA was mainly caused by the hydrophobic contributions to the solvation free energies and reaction field energies, which were calculated by either Possion Boltzmann or Generalized Born methods. Additionally, we also calculated the contributions to the binding free energies for each residue in the PB2 cap binding domain. As shown in Figure 2, the residues Ser320, Phe323, Ser337, Lys339, Arg355, His357, Glu361, Lys376, Phe404, Gln406, Asn429, and His432 had significantly positive contributions to the binding free energies of the PB2-m7GTP complex. Among these residues, His357 had the most contributions to the binding free energies of the PB2-m7GTP complex. This residue employed both hydrogen bonding and π-π
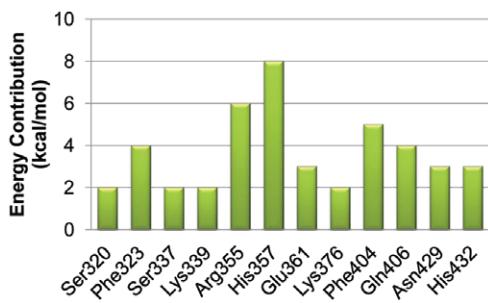
Figure 2. Energy contributions (kcal/mol) for each residue to the binding free energies of the PB2-m7GTP complex. Only the residues that have positive contributions to the binding free energies are listed in this figure. Among these residues, His357 has the most contributions owe to having both hydrogen bonding and π-π stacking interactions with the cap analog m7GTP.
doi:10.1371/journal.pone.0044079.g002

stacking interactions with the guanine ring of m7GTP (Figure 3), which was regarded as the key factor for recognizing and positioning substrates of the PB2 cap binding domain.

Structural analysis of the PB2-m7GTP complex

According to the crystal study, the cap analog m7GTP was non-covalently bound in the cap binding domain of influenza A virus PB2 subunit. Thus, our analyses of the interactions of the PB2 cap binding domain with m7GTP were focused on the non-covalent binding interactions, such as hydrogen bonding interactions and hydrophobic interactions. In the current study, we continued to use the definitions in our previous studies to define hydrogen bonds and hydrophobic interactions [46,47]. If the distance between the heavy atoms of a donor and an acceptor was less than 3.5 Å and the donor-H-acceptor angle was less than 30°, we thought that they could form a hydrogen bond. If the distance between the mass centers of the hydrophobic groups of the cap binding domain and m7GTP were less than 6.5 Å, we believed that they could have a hydrophobic interaction.

The detailed information of the hydrogen bonding interactions between the protein and m7GTP was described in Table 2. In Table 2, we only listed the information for the hydrogen bond having an occupancy more than 10%, which is believed to play crucial roles in the ligand binding. Hsi357, together with Ser320, Lys339, Arg355, Glu361, Lys376, and Gln406, formed an important hydrogen bonding interaction network, which was in agreement with the free energy calculations and believed to be essential for m7GTP binding and positioning. However, among
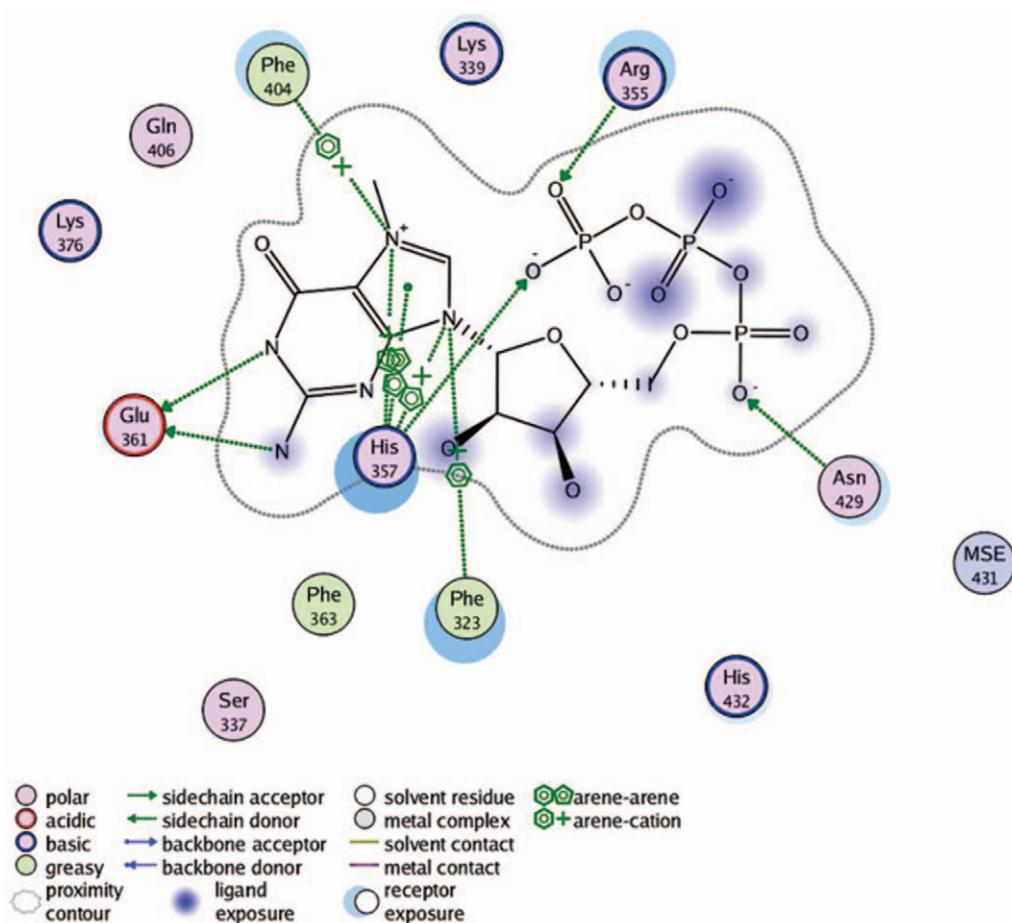
Figure 3. A 2D snapshot from MD simulations to show the binding interactions of the cap analog m7GTP with the key residues in the cap binding domain. The hydrogen bond is shown in green dashed line with the arrow pointed to the hydrogen acceptor. The π-π stacking and π-cation interactions are presented in green benzene ring symbol. The residues without any dashed line or benzene ring symbol have only van der Waals interactions with m7GTP.
doi:10.1371/journal.pone.0044079.g003

the aforementioned residues, His357 was considered to be the key component by reason that it was detected to contribute two important hydrogen bonds. One was formed by the side-chain nitrogen ($N\varepsilon_2$) of His357 and the oxygen (O6) on the β-phosphorus of m7GTP with an average distance of 3.95 ± 0.60 Å, while another was by the backbone nitrogen (N) of His357 and ketonic the oxygen (O) from the guanine ring of m7GTP with an average distance of 2.97 ± 0.30 Å. Notably, the hydrogen bond formed by His357 (N) and m7GTP (O6) had a comparatively highest occupancy (95.6%) among all the detected hydrogen bonds, functioning to position the guanine ring in a correct site. The hydrogen bonds formed by Ser320, Glu361, and Gln406 employed the similar function that they could keep the guanine ring of m7GTP in a correct direction and site in the active site. Otherwise, the hydrogen bonds formed by Lys339 and Arg355 were believed to play an important role in positioning the phosphorus tail.

Besides the hydrogen bonding interactions, hydrophobic interactions also play key roles in substrate binding. In the current case, the guanine ring was the major hydrophobic group of the cap analog m7GTP, which could provide hydrophobic interactions with the key residues in the active site. Addition to forming two significant hydrogen bonds, His357 also had strong hydrophobic interactions with the guanine ring of m7GTP with a probability of more than 95%. Additionally, Glu361, Lys376, and Gln406 were also detected to have strong hydrophobic interactions in over 90% frames from the MD simulations, which could also form significant hydrogen bonds according to the aforementioned discussions. Arg355, a key residue for hydrogen bonding interactions of m7GTP, also employed a significant hydrophobic interaction with the guanine ring of m7GTP in nearly 50% frames along the simulation trajectories.

Aromatic sandwich mechanism for the cap binding

How does the cap binding domain discriminate between capped RNAs and non-capped ones? This is of primary importance to the functions of the influenza A virus PB2 subunit. Besides the PB2 subunit, 3 cellular cap-binding proteins are well known by now: the eukaryotic translation initiation factor eIF4E [48], cap-binding complex (CBC [49]) consisting of 2 subunit cap-binding protein 20 (CBP20) in association with an ancillary protein CBP80, and the vaccinia viral protein 39 (VP39 [50]). Based on the discussions above and other cap binding proteins in the protein structure databases, we could get a direct im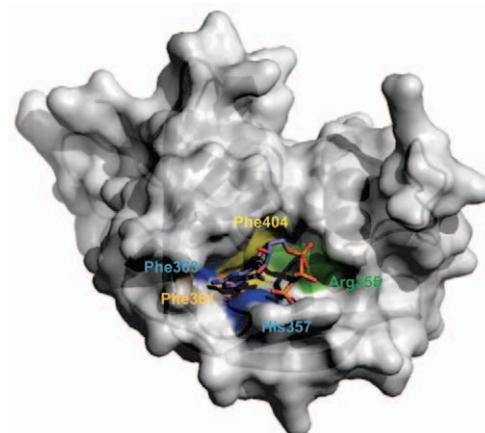pact on the active site of the PB2 cap binding domain, based on which 2 main features contributing to the cap-specific recognition were summarized.

The binding pocket that accommodated the m7GTP aromatic ring was believed to stack between two aromatic residues Phe363 and His357, which were Trp56 and Trp102 in eIF4E, Tyr20 and Tyr43 in CBC, Tyr22 and Phe180 in VP39. Thus, the cap recognition and binding mechanism was also known as an aromatic sandwich mechanism, which had been detected from other small-molecule models by crystallographic and theoretical studies [51–54]. In the aromatic sandwich mechanism, the stacked aromatic rings employed strong π-electron interactions due to the perfect alignment of the aromatic rings of Phe363, His357 as well as m7GTP, and extensive overlap regions in the two stacked rings (Figure 4). This was why the cap-binding protein had >100-fold low affinity for non-methylated cap analogues [55]. Additional experiments showed that the mutation of the aromatic residues involved in the aromatic sandwich mechanism could reduce the binding affinities and specificities for cap analogues.

Residue Phe404 was of a crucial importance for delocalization of the positive charge from N7-methylation of m7GTP purine ring. The delocalization of the positive charge was also good for salt bridges and hydrogen bonds between N1 and N2 of m7GTP and the carboxylate groups of Glu361 in the active sites. The residues having similar functions with Phe404 were also found in eIF4E (Glu103), CBC (Asp116), and VP39 (Glu223). When these critical acidic residues were substituted by other amino acids, the affinity and specificity of the cap analogues significantly decreased [56]. Thus, Phe404 was involved in these additional contacts in the case of the influenza A virus PB2 subunit, although this residue was also involved in the aromatic stacking with m7GTP.

Virtual screening for novel drugs

Based on the structural analysis of the binding mode for the PB2-m7GTP complex, we can obtain a deeper insight into the cap-binding pocket of the PB2 subunit. Based on this knowledge, we performed virtual screening on our in-house Finally, we found 5 candidates which might be potential leading compounds for the PB2 subunit, and their detailed information was listed in Table S1. Interestingly, Compound 1 (Cpd1) and Compound 2 (Cpd2) have

Table 2. Detailed information for the important hydrogen bonding interactions between the PB2 cap binding domain and cap analog m7GTP.

Rank	Residues	m7GTP	Distance/Å	Occupancy
1	His357 (N)	O6	2.97 ± 0.30	95.6%
2	Gln406 (N)	O1	3.01 ± 0.16	87.2%
3	Ser320 (Oγ)	O5	3.10 ± 0.59	72.0%
4	Arg355 (Nη$_1$)	O4	3.18 ± 1.08	64.0%
5	Lys376 (Nξ)	N2	3.36 ± 1.02	60.0%
6	Glu361 (Oε$_1$)	N3	3.51 ± 0.63	31.6%
7	Lys339 (Nξ)	O7	3.61 ± 0.65	22.7%
8	His357 (Nε$_2$)	O6	3.95 ± 0.60	12.9%

doi:10.1371/journal.pone.0044079.t002

Figure 4. The cap-binding pockets for the influenza A virus PB2 subunit. The residues around the cap-binding pocket are colored so that those aromatic amino acids forming the cap sandwich around the cap analogue m7GTP are in blue, those binding the functional groups of the guanine residue are in orange, those stabilizing the 7-methyl group are in yellow and those binding the triphosphate moiety are green.

doi:10.1371/journal.pone.0044079.g004

been reported to have inhibitory activities against the influenza A virus PB2 subunit [57,58]. In 2003, Hooker and his co-workers established a novel quantitative assay system to analysis the interactions between cap-binding proteins and cap analogues. Using this system, they have identified Cpd1 as a novel and selective inhibitor of the influenza cap binding protein [57]. Cpd2 is a phenethylphenylphthalimide analog derived from thalidomide, which is also proved to have strong inhibiting activity to the influenza virus cap-binding protein [58].

Their binding free energies calculated from MD trajectories by MM-PB/SA and MM-GB/SA were listed in Table 1, from which we found that Cpd1 and Cpd2 had a comparatively higher binding affinity (-45.01 ± 3.95 kcal/mol and -53.80 ± 3.15 kcal/mol respectively obtained from MM-GB/SA) than that of m7GTP (-38.25 ± 4.65 kcal/mol). The binding affinity calculated from MM-PB/SA for Cpd2 (-7.08 ± 0.12 kcal/mol) was also higher than that of m7GTP (-5.86 ± 0.14 kcal/mol). However, the one obtained from MM-PB/SA for Cpd1 (-5.68 ± 0.16 kcal/mol) was a little lower then that of m7GTP. This was mainly caused by the free energy calculation method. Due to having a number of negative charges in the phosphorus tail, the electrostatic contribution ($\Delta G_{ele/PB}$) to solvation free energy for m7GTP might be not correctly computed. In the current case, the electrostatic contribution to solvation free energy for m7GTP was 0.67 ± 7.86 kcal/mol using MM-PB/SA, while those for Cpd1 and Cpd2 were 11.85 ± 5.43 kcal/mol and 30.16 ± 4.41 kcal/mol, respectively.

Additionally, Cpd1 and Cpd2 had significantly higher van der Waals potential energies (-42.17 ± 3.41 kcal/mol and -62.19 ± 3.45 kcal/mol respectively) that that of m7GTP (-36.74 ± 4.09 kcal/mol). Detailed analyses showed that His357, Phe404 and Gln406 contributed strong hydrophobic interactions to Cpd1 and Cpd2 in over 90% frames along the simulation trajectories. According to the previous analysis, these residues also had significant hydrophobic interactions with m7GTP. In the PB2-Cpd1 complex, Met431 had also a strong hydrophobic interaction in more than 90% frames, which in the PB2-m7GTP and PB2-Cpd2 complexes were in 40% and 60% frames. Additions to the hydrophobic interactions, some residues, such as Arg355, Glu361 as well as Gln406 in the PB2-Cpd1 complex and Arg379 and Glu407 in the PB2-Cpd2 complex were also detected complexes to have important hydrogen bonding interactions.

Multi-target selectivity study

To further study the multi-target selectivity of the 5 candidates mentioned above, we also docked the candidates into the 841 known targets from the Potential Drug Target Database (PDTD) [45]. Based on their therapeutic areas, these targets can be categorized into 14 different types: Heoplastic disease, hormones and hormone antagonists, viral infections, blood and blood-forming organs, immunomodulation, bacterial infections, renal and cardiovascular functions, synaptic and neuroeffector junctional sites and central nervous system, parasitic infectious disease, inflammation, the vitamins, gastrointestinal functions, fungal infections and uterine motility. The docking results were ranked according to their docking scores, and the top 5 candidate-binding proteins were listed in Table S2. Interestingly, the top hit of the multi-target selectivity for all the 5 candidates were the influenza

virus cap-binding domain PB2 subunit (2vqz.pdb), indicating that influenza virus cap-binding domain PB2 subunit might be the preferred target for these candidates with respect to the other targets approved by the US FDA. Additionally, the top 5 binding proteins for all the candidates were involved in the viral infections according to the PDTD category.

In summary, influenza A virus PB2 subunit is responsible for replication and transcription of the viral RNA in the nucleus of infected cells. The initial step of viral mRNA transcription is the host-cell pre-mRNA bound to the PB2 subunit by its 5′ caps. How the cap-binding domain discriminates between capped RNAs and non-capped ones is of primary importance to the functions of the influenza A virus PB2 subunit. In the current study, we performed molecular dynamics simulations and free energy calculations on the influenza A virus PB2 subunit with a cap analog m7GTP. Based on the simulation trajectories and structural analyses, we identified some key residues, such as Arg355, His357, Glu361, and Gln406, to have significant hydrogen bonding and hydrophobic interactions with the guanine ring of m7GTP, recognizing and positioning the guanine ring with an aromatic sandwich mechanism, which would be the key strategy for searching and designing novel inhibitors against the influenza A virus PB2 subunit. Thus, we applied this idea to search our in-house database, and found 5 candidate structures that might employ significant inhibitory activities to the PB2 subunit. Interestingly, two of the candidates (Cpd1 and Cpd2) obtained from the virtual screening procedure had been reported to have inhibitory activities against the influenza virus cap-binding proteins. The free energy calculations for both PB2-Cpd1 and PB2-Cpd2 showed that they employed comparatively higher binding affinities with the PB2 subunit than that of m7GTP due to having much stronger non-bonded van der Waals energies. We believed that our findings could provide a deeper understanding of the cap recognition for the influenza A virus, offering useful information for searching or designing novel drugs against influenza A virus.

Supporting Information

Figure S1 The time-dependent RMS deviations for the Cα atoms of the PB2 cap-binding domain from the crystal structure. PB2-m7GTP system reached to its equilibrated state after about 800-ps MD simulations with an average RMS deviation value of 1.55 ± 0.18 Å from the crystal structure, which at the end of the simulations was ~1.53 Å.
(TIF)

Table S1 The candidates obtained from virtual screening using m7GTP as a structural template and aromatic sandwich mechanism as a designed strategy.
(DOC)

Table S2 The top 5 binding proteins of each candidates derived by the multi-target selectivity study.
(DOC)

Author Contributions

Conceived and designed the experiments: JFW PH XL YXL. Performed the experiments: YJW JP YW. Analyzed the data: YJW YY PL. Contributed reagents/materials/analysis tools: JP JFW PH. Wrote the paper: YJW JFW PH XL.

References

1. Eccles R (2005) Understanding the symptoms of the common cold and influenza. Lancet Infect Dis 5: 718–725.

2. Lee RV (2007) Transmission of influenza in human beings. Lancet Infect Dis 7: 759–760.

3. Tang JW, Li Y (2007) Transmission of influenza A in human beings. Lancet Infect Dis 7: 758.

4. Gambotto A, Barratt-Boyes SM, de Jong MD, Neumann G, Kawaoka Y (2008) Human infection with highly pathogenic H5N1 influenza virus. Lancet 371: 1464–1475.

5. Olsen B, Munster VJ, Wallensten A, Waldenström J, Osterhaus AD, et al. (2006) Global patterns of influenza A virus in wild birds. Science 312: 384–388.

6. Lamb RA, Choppin PW (1993) The gene structure and replication of influenza virus. Annu Rev Biochem 52: 467–506.

7. Bouiver NM, Palese P (2008) The biology of influenza viruses. Vaccine 26: D49–D53.

8. Ghedin E, Sengamalay NA, Shumway M, Zaborsky J, Feldblyum T, et al. (2005) Large-scale sequencing of human influenza reveals the dynamic nature of vrial genome evolution. Nature 437: 1162–1166.

9. Hayden FG, Treanor JJ, Fritz RS, Lobo M, Betts RF, et al. (1999) Use of the oral neuraminidase inhibitor oseltamivir in experimental human influenza: Randomized controlled trials for prevention and treatment. JAMA 282: 1240–1246.

10. Barroso L, Treanor J, Gubareva L, Hayden FG (2005) Efficacy and tolerability of the oral neuraminidase inhibitor peramivir in experimental human influenza: randomized, controlled trials for prophylaxis and treatment. Antivir Their 10: 901–910.

11. Monto AS, Robinson DP, Herlocher ML, Hinson JM, Elliott MJ, et al. (1999) Zanamivir in the prevention of influenza among healthy adults: a randomized controlled trial. JAMA 282: 31–35.

12. Hay AJ, Wolstenholme AJ, Skehel JJ, Smith MH (1985) The molecular basis of the specific anti-influenza action of amantadine. EMBO J 4: 3021–3024.

13. Beigel J, Bray M (2008) Current and future antiviral therapy of severe seasonal and avian influenza. Antivir Res 78: 91–102.

14. Bright RA, Shay DK, Shu B, Cox NJ, Klimov AI (2006) Adamantane resistance among influenza A viruses isolated early during the 2005–2006 influenza season in the United States. JAMA 295: 891–894.

15. Leung TWC, Tai ALS, Cheng PKC, Kong MSY, Lim W (2009) Detection of an oseltamivir-resistant pandemic influenza A/H1N1 virus in Hong Kong. J Clin Virol 46: 298–299.

16. Wang NX, Zheng JJ (2009) Computational studies of H5N1 influenza virus resistance to oseltamivir. Protein Sci 18: 707–715.

17. Pielak RM, Schnell JR, Chou JJ (2009) Mechanism of drug inhibition and drug resistance of influenza A M2 channel. Proc Natl Acad Sci USA 106: 7379–7384.

18. Wang JF, Wei DQ, Chou KC (2009) Insights from investigating the interactions of adamantane-based drugs with the M2 proton channel from the H1N1 swine virus. Biochem Biophys Res Commun 388: 413–417.

19. Pielak RM, Chou JJ (2010) Flu channel drug resistance: a tale of two sites. Protein Cell 1: 246–258.

20. Sharma M, Yi M, Dong H, Qin H, Peterson E, et al. (2010) Insight into the mechanism of the influenza A proton channel from a structure in a lipid bilayer. Science 330: 509–512.

21. Wang JF, Chou KC (2010) Insights from studying the mutation-induced allostery in the M2 proton channel by molecular dynamics. Protein Eng Des Sel 23: 663–666.

22. Coloma R, Valpuesta JM, Arranz R, Carrascosa JL, Ortin J, et al. (2009) The structure of a biologically active influenza virus ribonucleoprotein complex. PLoS Pathogens 5: e1000491.

23. Murti KG, Webster RG, Jones IM (1988) Localization of RNA polymerases of influenza viral ribonucleoproteins by immunogold labeling. Virology 164: 562–566.

24. Plotch SJ, Bouloy M, Ulmanen I, Krug RM (1981) A unique cap(m7GpppXm)-dependent influenza virion endonuclease cleaves capped RNAs to generate the primers that initiate viral RNA transcription. Cell 23: 847–858.

25. Li ML, Rao P, Krug RM (2001) The active sites of the influenza cap-dependent endonuclease are on different polymerase subunits. EMBO J 20: 2078–2086.

26. Guilligay D, Terendeau F, Resa IP, Coloma R, Crepin T, et al. (2008) The structural basis for cap binding by influenza virus polymerase subunit PB2. Nat Struct Mol Biol 15: 500–506.

27. Wang Y, Wei DQ, Wang JF (2010) Molecular dynamics studies on T1 lipase: insight into a double-flap mechanism. J Chem Inf Model 50: 875–878.

28. Wang JF, Chou KC (2011) Insights from modeling the 3D structure of New Delhi metallo-ß-lactamse and its binding interactions with antibiotic drugs. PLoS One 6: e18414.

29. Lian P, Wei DQ, Wang JF, Chou KC (2011) An allosteric mechanism inferred from molecular dynamics simulations on phospholamban pentamer in lipid membranes. PLoS One 6: e18587.

30. Georgescu RE, Alexov EG, Gunner MR (2002) Combining conformational flexibility and continuum electrostatics for calculating pK(a)s in proteins. Biophys J 83: 1731–1748.

31. Case DA, Cheatham TE III, Darden T, Gohlke H, Luo R, et al. (2005) The Amber biomolecular simulation program. J Comput Chem 26: 1668–1688.

32. Ponder JW, Case DA (2003) Force field for protein simulations. Adv Protein Chem 66: 27–85.

33. Kollman PA, Massova I, Reyes C, Kuhn B, Huo SH, et al. (2000) Cheatham, T.E. Calculating structures and free energies of complex molecules: Combining molecular mechanics and continuum models. Acc Chem Res 33: 889–897.

34. Chen HF, Luo R (2007) Binding induced folding in p53-MDM2 complex. J Am Chem Soc 129: 2930–2937.

35. Wang J, Tan C, Chen HF, Luo R (2008) All-atom computer simulations of amyloid fibrils disaggregation. Biophys J 95: 5037–5047.

36. Qin F, Chen Y, Li YX, Chen HF (2009) Induced fit for mRNA/TIS11d complex. J Chem Phys 131: 115103.

37. Qin F, Jiang Y, Chen Y, Wu M, Yan G, et al. (2011) Conformational selection or induced fit for Brinker and DNA recognition. Phys Chem Chem Phys 13: 1407–1412.

38. Wang JF, Chou KC (2012) Insights into the mutation-induced HHH syndrome from modling human mitochondrial ornithine transporter-1. PLoS One 7: e31048.

39. Wang JF, Hao P, Li YX, Dai JL, Li X (2012) Exploration of conformational transition in the aryl-binding site of human FXa using molecular dynamics simulations. J Mol Model 18: 2717–2725.

40. Chen HF (2009) Molecular dynamics simulation of phosphorylated KID post-translational modification. PLoS One 4: e6516.

41. Li J, Wei DQ, Wang JF, Li YX (2011) A negative cooerativity mechanism of human CYP2E1 inferred from molecular dynamics simulations and free energy calculations. J Chem Inf Model 51: 3217–3225.

42. Chen Y, Li Z, Chen HF (2010) Computational study of CCR5 antagonist with support vector machines and three dimensional quantitative structure activity relationship methods. Chem Biol Drug Des 75: 295–309.

43. Willett P (1998) Chemical similarity searching. J Chem Inf Comput Sci 38: 983–996.

44. Iwai Y, Takahashi H, Hatakeyama D, Motoshima K, Ishikawa M, et al. (2010) Anti-influenza activity of phenethylphenylphthalimide analogs derived from thalidomide. Bioorg Med Chem 18: 5379–5390.

45. Gao Z, Li H, Zhang H, Liu X, Kang L, et al. (2008) PDTD: a web-accessible protein database for drug target identification. BMC Bioinformatics 9: 104.

46. Gu H, Chen HF, Wei DQ, Wang JF (2010) Molecular dynamics simulations exploring drug resistance in HIV-1 protease. Chinese Sci Bull 55: 2677–2683.

47. Zeng QK, Du HL, Wang JF, Wei DQ, Wang XN, et al. (2009) Reversal of coenzyme specificity and improvement of catalytic efficiency of Pichia stipitis xylose reductase by rational site-directed mutagenesis. Biotechnol Lett 31: 1025–1029.

48. Macrotrigiano J, Gingras AC, Sonenberg N, Burley SK (1997) Cocrystal structure of the messenger RNA 5′ cap-binding protein (eIF4E) bound to 7-methyl-GDP. Cell 89: 951–961.

49. Calero G, Wilson KF, Ly T, Rios-Steiner JL, Clardy JC, et al. (2002) Structural basis of m7GpppG binding to the nuclear cap-binding protein complex. Nat Struct Biol 9: 912–917.

50. Hodel AE, Gershon PD, Quiocho FA (1998) Structural basis for sequence-nonspecific recognition of 5′-capped mRNA by a cap-modifying enzyme. Mol Cell 1: 443–447.

51. Ishida T, Doi M, Inoue M (1988) A selective recognition mode of a nucleic acid based by an aromatic amino acid: L-phenylalnine-7-methylguanosine 5′-monophosphate stacking interaction. Nucleic Acids Res 16: 6175–6190.

52. Darzynkiewicz E, Lonnberg H (1989) Base stacking of simple mRNA cap analogues. Association of 7,9-dimethylguanine, 7-methylguanosine and 7-methylguanosine 5′-monophosphate with indole and purine derivatives in aqueous solution. Biophys Chem 33: 289–293.

53. Ueda H, Iyo H, Doi M, Inoue M, Ishida T (1991) Cooperative stacking and hydrogen bond pairing interactions of fragment peptide in cap binding protein with mRNA cap structure. Biochim Biophys Acta 1075: 181–186.

54. Stolarski R, Sitek A, Stepinski J, Jankowska M, Oksman P, et al. (1996) 1H-NMR studies on association of mRNA cap-analogues with tryptophan-containing peptides. Biochim Biophys Acta 1293: 97–105.

55. Niedzwieck A, Marotrigiano J, Stepinski J, Jankowska-Anyszka M, Wyslouch-Cieszynska A, et al. (2002) Biophysical studies of eIF4E cap-binding protein: recognition of mRNA 5′ cap structure and synthetic fragments of eIF4G and 4E-BP1 proteins. J Mol Biol 319: 615–635.

56. Hu G, Gershon PD, Hodel AE, Quiocho FA (1999) mRNA cap recognition: dominant role of enhanced stacking interactions between methylated bases and protein aromatic side chains. Proc Natl Acad Sci USA 96: 7149–7154.

57. Hooker L, Sully R, Handa B, Ono N, Koyana H, et al. (2003) Quantitative analysis of influenza virus RNP interaction with RNA cap structures and comparison to human cap binding protein eIF4E. Biochemistry 42: 6234–6240.

58. Lipinski CA, Lombardo F, Dominy BW, Feeney PJ (2001) Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Adv Drug Del Rev 46: 2–36.