

Article

# Characterization of the PB2 Cap Binding Domain Accelerates Inhibitor Design

Amanda E. Constantinides <sup>1</sup> , Chelsea C. Severin <sup>1,2</sup>, Ryan H. Gumpper <sup>1</sup>, Xiaofeng Zheng <sup>3,4</sup>   
and Ming Luo <sup>1,2,\*</sup>

<sup>1</sup> Department of Chemistry, Georgia State University, Atlanta, GA 30302, USA; Aconstantinides1@student.gsu.edu (A.E.C.); Cseverin1@student.gsu.edu (C.C.S.); Rgumpper1@student.gsu.edu (R.H.G.)

<sup>2</sup> Center for Diagnostics and Therapeutics, Georgia State University, Atlanta, GA 30302, USA

<sup>3</sup> State Key Lab of Protein and Plant Gene Research, School of Life Sciences, Peking University, Beijing 100871, China; xiaofengz@pku.edu.cn

<sup>4</sup> Department of Biochemistry and Molecular Biology, School of Life Sciences, Peking University, Beijing 100871, China

\* Correspondence: mluo@gsu.edu; Tel.: +1-404-413-6608

Received: 7 December 2017; Accepted: 23 January 2018; Published: 28 January 2018

**Abstract:** X-ray crystallographic structural determinations of the PB2 cap binding domain (PB2cap) have improved the conformational characterization of the RNA-dependent RNA polymerase machinery (PA, PB2, and PB1) of the influenza virus. Geometrically, the catalytic PB1 subunit resembles the palm of a human hand. PA lies near the thumb region, and PB2 lies near the finger region. PB2 binds the cap moiety in the pre-mRNA of the host cell, while the endonuclease of PA cleaves the pre-mRNA 10–13 nucleotides downstream. The truncated RNA piece performs as a primer for PB1 to synthesize the viral mRNA. Precisely targeting PB2cap with a small molecule inhibitor will halt viral proliferation via interference of the cap-snatching activity. Wild-type and mutant PB2cap from A/California/07/2009 H1N1 were expressed in *Escherichia coli*, purified by nickel affinity and size exclusion chromatography, crystallized, and subjected to X-ray diffraction experiments. The crystal of mutant PB2cap liganded with m<sup>7</sup>GTP was prepared by co-crystallization. Structures were solved by the molecular replacement method, refined, and deposited in the Protein Data Bank (PDB). Structural determination and comparative analyses of these structures revealed the functions of Glu361, Lys376, His357, Phe404, Phe323, Lys339, His432, Asn429, Gln406, and Met401 in PB2cap, and the dissociation of the influenza A PB2cap C-terminal subdomain (residues 446–479) upon ligand binding. Understanding the role of these residues will aid in the ultimate development of a small-molecule inhibitor that binds both Influenza A and B virus PB2cap.

**Keywords:** cap-snatching; polymerase basic protein 2; RNA polymerase; domain movement

## 1. Introduction

The influenza virus continues to have a devastating impact on the global population as it evolves in defiance of current countermeasures. Adamantane derivatives, which were once effective against the influenza A virus, have lost their clinical practicality and potency against current strains of the virus. A new generation of antiviral drugs, neuraminidase inhibitors, has recently been developed as effective treatment against infection by influenza A and B viruses. The rate of developing resistance against this class of inhibitors is of critical importance as single-molecule inhibition strategies may fail. The inspiration for a novel small-molecule inhibitor originates from the difference between cap-binding patterns of human and influenza proteins. Unlike the influenza virus, 7-methyl-G(5')ppp(5')N at the 5' end of human capped pre-mRNA must first be recognized by the eukaryotic initiation factor

4E (eIF4E) [1] to form the nuclear cap binding complex (CBC) for translation initiation [2]. On the contrary, polymerase basic protein 2 (PB2), part of the viral heterotrimeric RNA-dependent RNA polymerase machinery, interacts with polymerase basic protein 1 (PB1) [3–5] and binds the cap of the host pre-mRNA [6] to allow the N-terminal endonuclease of polymerase acidic (PA) to cleave the pre-mRNA 10–13 nucleotides downstream. The truncated capped RNA is then used as the new primer by PB1 to synthesize the viral mRNA [7–9]. To halt viral transcription, a novel competitive inhibitor must bind at the PB2 cap-binding domain, named PB2cap (residues 318–483 in PB2 from A/California/07/2009 H1N1), with a higher binding affinity than the host pre-mRNA.

X-ray crystallographic structural determinations of mutant influenza A virus PB2cap bound to  $m^7$ GTP, a 7-methyl-G(5')ppp(5')N cap analogue, have revealed key residues involved in cap binding (PDB codes: 5EG7, 4CB4, and 4EQK) [10–12] by comparison to unliganded PB2cap (PDB codes: 5WOP, 5EG8, and 4ENF) [10,11,13]. Wild-type and mutant influenza A PB2cap from A/California/07/2009 (H1N1) were expressed, purified, crystallized, and the crystals were subjected to X-ray diffraction studies. The crystal of mutant PB2cap liganded with  $m^7$ GTP was prepared by co-crystallization. High resolution structures of liganded and unliganded PB2cap were solved by molecular replacement and deposited in the Protein Data Bank (PDB) (PDB codes: 5EG7, 5WOP and 5EG8). These structures were also compared to similar ligand-bound structures (PDB codes: 4CB5, 4P1U) [12,14]. As seen in the binding site of 4P1U, an azaindole inhibitor, VX-787, has been identified to exclusively bind and inhibit influenza A virus PB2cap. It was also discovered that changes in the conformation of the PB2cap C-terminal subdomain were induced upon ligand binding. Additional structural comparisons were composed to highlight the differences between the influenza A and B virus PB2cap (PDB code: 4Q46) [15] and between PB2cap and  $m^7$ GTP-bound human cap-binding proteins (PDB codes: 4N49, 4TQB) [16,17]. These conclusions serve as a solid platform for the future design of a potent small-molecule inhibitor that targets both the influenza A and B virus PB2cap.

## 2. Materials and Methods

### 2.1. Expression and Purification

For the structure determination of the wild-type (PDB code: 5EG8),  $m^7$ GTP-mutant (PDB code: 5EG7), and mutant CA09-PB2cap (PDB code: 5WOP) (Table 1), the A/California/07/2009 H1N1 DNA coding sequence was first inserted between NdeI and XhoI restrictions sites into the pET28a expression vector to generate a His<sub>6</sub>-tag chimeric protein [10]. A QuickChange kit from Agilent Technologies, Santa Clara, CA, USA, was used to generate the expression vector for the mutant PB2cap. Between 0.6–0.8 OD<sub>600</sub>, protein expression was induced with 0.5 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) at 293.15 K for 18 h in *Escherichia coli* BL21 (DE3) cells. A nickel-affinity column (Ni<sup>2+</sup>-charged HiTrap chelating HP column from GE Healthcare, Little Chalfont, UK) was used to purify the recombinant protein at 281.15 K. Thrombin protease was used to cleave the His<sub>6</sub> tag from the protein while dialyzing in 50 mM Tris pH 8.0, 200 mM NaCl, 15 mM imidazole, and 5 mM β-mercaptoethanol. A second nickel-affinity purification was conducted to remove the protein still containing the His<sub>6</sub>-tag. For the purification of the mutant CA09-PB2cap, a Superdex 75 column from GE Healthcare was equilibrated with 50 mM Tris pH 8.0, 200 mM NaCl, and 2 mM DTT. For the purification of the wild-type and  $m^7$ GTP-mutant CA09-PB2cap, the Superdex 75 column was equilibrated with 10 mM Tris pH 8.0 and 200 mM NaCl. The proteins were purified by size-exclusion chromatography.

**Table 1.** Data collection, processing, structure solution, and refinement (values for the outer resolution shell are given in parentheses) [10,13].

Protein	CA09-PB2cap	m <sup>7</sup> GTP-Mutant CA09-PB2cap	Mutant CA09-PB2cap
PDB code	5EG8	5EG7	5WOP
Data collection			
Space-group	P3 <sub>1</sub> 21	P4 <sub>3</sub> 2 <sub>1</sub> 2	P12 <sub>1</sub> 1
Unit-cell parameters			
a (Å)	54.61	45.26	36.86
b (Å)	54.61	45.26	72.43
c (Å)	196.49	156.4	61.32
α (°)	90	90	90
β (°)	90	90	103.18
γ (°)	120	90	90
Molecules in asymmetric unit	2	1	2
Wavelength (Å)	0.9795	1	1
Resolution (Å)	15–1.54 (1.59–1.54)	25–1.40 (1.42–1.40)	36.21–1.52 (1.574–1.52)
R <sub>merge</sub> <sup>†</sup> (%)	7.0 (21.4)	7.1 (61.1)	4.6 (47.3)
⟨I/σ(I)⟩	57.9 (14.2)	33.2 (1.8)	11.05 (2.03)
CC1/2	0.997 (0.992)	0.912 (0.581)	0.998 (0.806)
Completeness (%)	97.2 (96.9)	99.8 (97.0)	99.17 (99.94)
Multiplicity	15.1 (16.7)	10.4 (4.0)	2.0 (2.0)
Refinement			
Resolution (Å)	14.63–1.54	25–1.40	36.21–1.52
No. of reflections	47,547	31,406	93,971
R <sub>work</sub> /R <sub>free</sub> <sup>‡</sup> (%)	12.2/17.7	12.4/16.9	0.1498/0.2044
Average B factor (Å <sup>2</sup> )	24.9	16.1	22.01
R.m.s. deviations			
Bond lengths (°)	0.019	0.02	0.026
Bond angles (°)	1.93	2.35	2.22
Ramachandran plot (%)			
Favored region	97.5	98.1	98.09
Allowed region	2.2	1.9	1.91
Outlier region <sup>§</sup>	0.3	—	0

<sup>†</sup>  $R_{\text{merge}} = \frac{\sum_{hkl} \sum_i |I_i(hkl) - I(hkl)|}{\sum_{hkl} \sum_i I_i(hkl)}$ . <sup>‡</sup>  $R = \frac{\sum_{hkl} ||F_{\text{obs}}| - |F_{\text{calc}}||}{\sum_{hkl} |F_{\text{obs}}|}$ . <sup>§</sup> Outliers are in loop regions lacking clear density.

## 2.2. Crystallization

For the wild-type and m<sup>7</sup>GTP-mutant CA09-PB2cap, the purified protein solution was concentrated to 4–10 mg·mL<sup>−1</sup> for crystallization. The proteins were subjected to crystal screens (Index, Natrix, PEG/Ion, Crystal Screen, and Crystal Screen 2) from Hampton Research, Aliso Viejo, CA, USA. Conditions were optimized and then crystals were grown via hanging-drop vapor diffusion with 10 mg·mL<sup>−1</sup> wild-type CA09-PB2cap protein mixed with 0.2 M magnesium nitrate, 0.1 M HEPES pH 8.0, and 20% (*w/v*) PEG 3350 at 293.15 K. 0.2 mM m<sup>7</sup>GTP was added to the mutant CA09-PB2cap solution, and then the resultant solution was concentrated to 4 mg·mL<sup>−1</sup>. M<sup>7</sup>GTP-mutant CA09-PB2cap cocrystals were grown using reservoir solution composed of 0.1 M Bis-Tris pH 6.5 and 15% (*w/v*) PEG 3350.

11 mg·mL<sup>−1</sup> mutant CA09-PB2cap was incubated with 5% 10 mM candidate small-molecule inhibitor in DMSO for two hours and then mixed at a 1:1 ratio (2 μL total) with the crystallization solution discovered in well H6 of Hampton Research's Joint Center for Structural Genomics (JCSG) suite at 277.15 K (0.1 M ammonium acetate, 0.1 M Bis Tris pH 5.86, 20% PEG 8000 (*v/v*)) by hanging drop vapor diffusion. Two weeks later, crystals in the hanging drops were observed.

## 2.3. Data Collection

For the wild-type and m<sup>7</sup>GTP-mutant CA-09 PB2cap, crystals were transferred into mother liquor with 20% glycerol (*v/v*). For the mutant CA-09 PB2cap, 10-μL of a cryoprotectant solution containing 20% glycerol (*v/v*) in 0.1 M ammonium acetate, 0.1 M Bis Tris pH 5.86, 20% PEG 8000 (*v/v*) was pipetted into the hanging drop. The drop was allowed two minutes of soaking and was then aspirated and sequentially pipetted into 500-μL of cryoprotectant.

To fish crystals for flash freezing in liquid nitrogen, the smallest fit cryo-pin was used to capture the crystal in a layer of cryoprotectant. Once it was microscopically confirmed that each crystal was in the center of its loop, they were quickly flash-frozen into a puck for shipment. The wild-type CA-09 PB2cap was shipped to the Shanghai Synchrotron Radiation Facility at the Shanghai National Synchrotron Radiation Centre in the People's Republic of China. The m<sup>7</sup>GTP-mutant CA-09 PB2cap and the mutant CA-09 PB2cap were sent to the Southeast Regional Collective Access Team (SER-CAT) at Argonne National Lab's Advanced Photon Source in Chicago, IL, USA.

#### 2.4. Data Processing and Refinement

To determine the structure of wild-type CA-09 PB2cap and m<sup>7</sup>GTP mutant CA-09 PB2cap, HKL-2000 [18] was used to process the data. 4ENF was used as the search model for molecular replacement using CCP4 Molrep [19]. Refinement (Table 1) was performed using REFMAC5 [19]. The coordinates with PDB codes: 5EG8 and 5EG7 were deposited in RCSB PDB.

To determine the structure of mutant CA-09 PB2cap, XDS [20] was used to process the data and CCP4 AIMLESS [19] was used for scaling. Xtriage in the PHENIX graphic user interface [21] revealed that the crystal may have a pseudo-translational symmetry. Intensity (I) was converted to structure amplitude (F) values using CCP4 Truncate [19]. 5EG7 was used as the search model for the molecular replacement structural determination of the mutant CA-09 PB2cap. CCP4 Molrep [19] was employed to conduct the molecular replacement calculations. CCP4 Refmac [19] was used to perform isotropic and anisotropic refinement and Coot [22] was used for model-building. Structural validation throughout the refinement process was generated by Molprobity and Polygon in PHENIX [21]. Translation, Libration, and Screw Rotation (TLS) input files were generated by CCP4 Tlsanl [19] for TLS refinement in Refmac (Table 1) [19]. No ligand appeared in the binding site and there are two chains of PB2cap per asymmetric unit. The structure with PDB code: 5WOP was deposited in RCSB PDB.

PYMOL [23] and Chimera 1.11.2 [24] were used to perform structural superpositions (Figures 1–4) and respectively generate root-mean-square deviation (RMSD) values (Table 2). PyMol [23] was also used for the generation of hydrogen bond representations (Figures 2 and 3) and for superposition to generate angular differences upon ligand binding (Table 2 and Figure 4). CCP4 Baverage [19] was used to generate lists of main-chain residual B-factors used in graphical representations of B-factor differences (Figure 5). The min-max rescaling normalization method was applied to normalize B-factor values prior to calculating the differences. Normalized B-factor values from liganded PB2cap were subtracted from native PB2cap normalized B-factor values. Normalized B-factor differences and RMSD values were plotted in Microsoft Excel. PyMol [23] was used for B-putty drawings in Figure 6, for the ribbon drawing in Figure 7, for the superpositions in Figures 8 and 9, and for the generation of hydrogen bonds in Figures 8–10. BIOVIA Discovery Studio 2017 was used to visualize  $\pi$ - $\pi$  force interactions [25].

**Table 2.** Alternative superposition of PB2cap structures and their subdomains [13].

Structures in Comparison	N-Subdomain RMSD (Å)	C-Subdomain RMSD (Å)	C <sup>†</sup> -Subdomain RMSD (Å)	C <sup>†</sup> -Subdomain Angular Distance (°)
5WOP Chain A vs. 5EG7	0.39	0.40	1.10	5.1
5WOP Chain B vs. 5EG7	0.36	0.40	1.06	4.7
4ENF vs. 5EG7	0.55	0.49	0.63	4.8
4CB7 Chain A vs. 4CB5	0.50	0.33	0.53	3.5

<sup>†</sup> indicates that RMSD was calculated between the C-subdomains after superposition using only the coordinates from the N-subdomains [23,24].

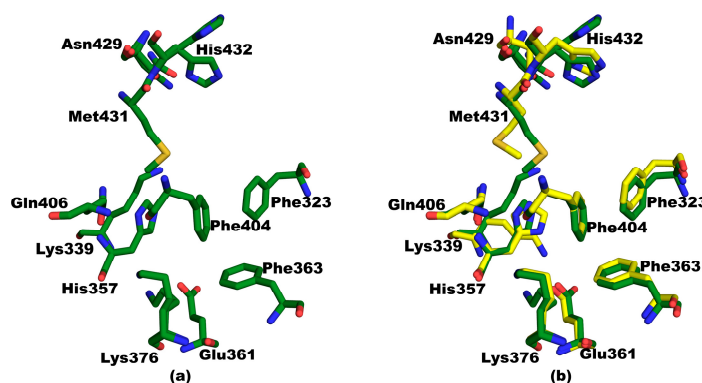
### 3. Results and Discussion

#### 3.1. Structure of PB2cap from A/California/07/2009 (H1N1) Compared to PB2cap from A/PR/8/34 (H1N1)

The crystal structure of unliganded wild-type PB2cap from A/California/07/2009 (H1N1) (PDB code: 5EG8) [10] at 1.54 Å revealed the key residues available for interaction with the ligand

(Figure 1). Glu361 has one stable conformation and Lys376 has two, with only one suitable for hydrogen bonding to the guanine moiety of a cap analogue (Figure 1a) [10]. The sidechain of His357 is oriented out of plane but could rotate in plane to sandwich the guanine moiety together with the sidechain of Phe404 to achieve optimal  $\pi$ -force interactions. Phe323 is available for hydrophobic interactions with the pyranose ring. When the cap-binding site is empty, Lys339 is rotated outside hydrogen-bonding distance from a potential interaction with the pyranose ring. Upon ligand binding, it could rotate into the cap-binding site and form a hydrogen bond with the 2'OH of the pyranose ring. The flexible sidechain of His432 and the dual-conforming sidechain of Asn429 could both participate in charge interactions with the  $\alpha$ -phosphate in the cap. A water molecule is bound to the carbonyl of Phe404 that could easily be displaced upon ligand binding [10].

Alternatively, the structure of PB2cap from A/PR/8/34 (H1N1) (PDB code: 4ENF) [11] does not feature multiple conformations of Lys376, Asn429, or His432 (Figure 1b). This comparison suggests that the sidechain orientation of binding site residues may be slightly different in unliganded PB2scap from different strains of H1N1 influenza A virus, despite highly conserved amino acid sequences [10].



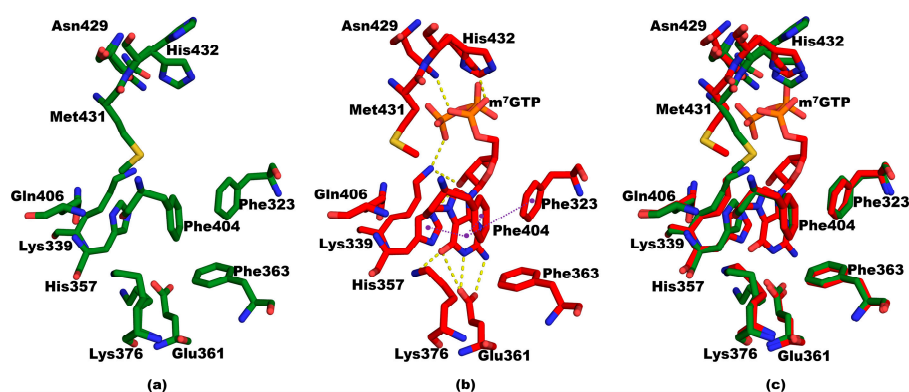
**Figure 1.** (a) The cap-binding site in PB2cap from A/California/07/2009 (H1N1) (PDB code: 5EG8) [10]; (b) Comparison between the cap-binding site in 5EG8 (forest) and that in PB2cap from A/PR/8/34 (H1N1) (PDB code: 4ENF) (yellow) [10,11,23,25].

### 3.2. Structure of Mutant PB2cap from A/California/07/2009 (H1N1) Bound to $m^7$ GTP

Co-crystallization with wild-type PB2cap from A/California/07/2009 (H1N1) and  $m^7$ GTP were unsuccessful because the loop containing Arg423 occupies the cap-binding site of neighboring molecules within the asymmetric unit [10]. To shorten the loop, a GSG linker was engineered between Val421 and Arg427 to generate the mutant PB2cap. The mutant PB2cap has the same structure as the wild-type PB2cap (R.M.S.D = 0.379 Å for 905 superposed atoms), except His357 in the active site of the mutant is slightly rotated. However, soaking crystals of mutant PB2cap in the mother liquor containing  $m^7$ GTP resulted in complete dissolution, indicating a major conformational change upon ligand binding.

The successful co-crystallization with  $m^7$ GTP led to the determination of the liganded mutant PB2cap (PDB code: 5EG7) structure at 1.40 Å resolution [10]. Glu361 is conveniently rotated to form a hydrogen bond with the guanine moiety of the cap analogue,  $m^7$ GTP (Figure 2). The comparison between 5EG7 and 5EG8 revealed that Lys376 has only one conformation when bound to the 6th positioned carbonyl of the guanine moiety instead of two when unbound (Figures 1a and 2). The guanine moiety is firmly snugged between sidechains of His357 and Phe404 due to the adaptive orientation of His357 (Figure 2). To participate in optimal  $\pi$ - $\pi$  stacking on one side of the guanine moiety, the sidechain of His357 is rotated from its unliganded conformation in plane with the guanine moiety and the sidechain of Phe404 (Figures 1 and 2). The water bound to the carbonyl of Phe404 in the unliganded structure is displaced upon binding to facilitate  $\pi$ - $\pi$  stacking. Phe323 is also available to participate in  $\pi$ - $\pi$  stacking on the same side of the guanine moiety as Phe404 by forming T-shaped interactions (Figure 2). His432 and

Asn429 form salt-bridges and hydrogen bonds with the  $\alpha$ -phosphate group in the cap. When PB2cap binding interactions are placed in the context of the full-length viral polymerase assembly, the interactions with the  $\beta$ -phosphate will be visible with more clarity. Like 5EG8, the sidechain of His432 assumes two conformations (Figures 1a and 2), whereas the sidechain of Asn429 assumes only one conformation when bound to  $m^7$ GTP, like the single conformation observed in 4ENF (Figures 1b and 2) [10,11]. Lastly, it was observed that sidechains of Phe404, Met431, and Gln406 form a hydrophobic pocket to accommodate the methyl group at position 7 of the cap (Figure 2). Comparisons with 5EG7 and unliganded mutant PB2cap from A/California/07/2009 H1N1 (mutant CA09-PB2cap) (PDB code: 5WOP) revealed similar interactions. 5WOP has the same structure as 5EG8 (R.M.S.D = 0.464 Å for 135 superposed atoms) except His432 and Asn429 adopts only one conformation instead of two, Met431 is rotated 180°, and His357 is slightly rotated [10].



**Figure 2.** (a) The cap-binding site in PB2cap from A/California/07/2009 (H1N1) (PDB code: 5EG8) (forest); (b) The cap-binding site in mutant PB2cap from A/California/07/2009 (H1N1) bound to  $m^7$ GTP (PDB code: 5EG7) (red); (c) 5EG8 (forest) aligned with 5EG7 (red). Hydrogen bonds are shown with dotted yellow lines, and  $\pi$ - $\pi$  force interactions are shown with dotted purple lines [10,23,25].

### 3.3. Structural Comparisons between 5EG7 and Other Structures

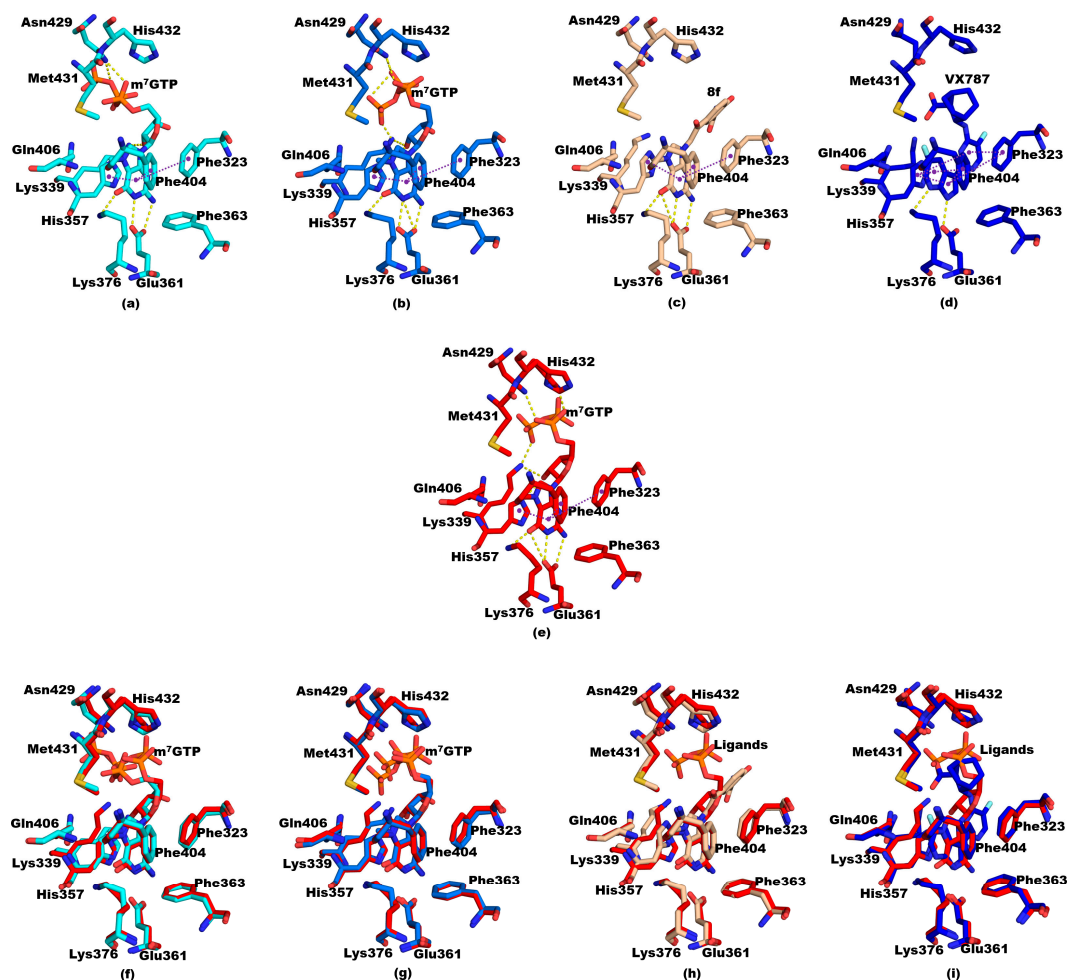
Comparison between 5EG7 [10] and PB2cap from A/Duck/Shantou/4610/2003 (H5N1) bound to  $m^7$ GTP (PDB code: 4CB4) [12] revealed similar interactions except the moved sidechain of His357 in 4CB4 no longer forms a hydrogen bond with the 2'-OH of the pyranose in the cap (Figure 3a,f) [10]. The side chain of Lys339 replaces this interaction in 4CB4. In addition, interactions with  $\alpha$ ,  $\beta$ , and  $\gamma$ -phosphates in the cap are different in 4CB4 versus 5EG7. Comparisons between 5EG7 and PB2cap from A/Hong Kong/1/68 (H3N2) bound to  $m^7$ GTP (PDB code: 4EQK) [11] revealed similar interactions with the cap, except the side chain of Phe323 is shifted in 4EQK, reducing T-shaped interactions with the guanine moiety (Figure 3b,g) [10].

5EG7 and PB2cap from A/Duck/Shantou/4610/2003 (H5N1) bound to compound **8f**, an inhibitor, (PDB code: 4CB5) [12] both have similar interactions due to the structural similarity between  $m^7$ GTP and its derivative, compound **8f** (Figure 3c,h) [10]. However, the hydroxyl group on the phenyl ring of compound **8f** forms a hydrogen bond with the mainchain amide of Phe303 via a water molecule. The interactions between compound **8f** and the PB2cap binding domain are relatively weak compared to inhibitor VX-787 [14].

VX-787 is a potent inhibitor and the structure of PB2cap from A/Victoria/3/1975 (H3N2) bound to VX-787 (PDB code: 4P1U) [14] showed conserved interactions in the cap-binding site (Figure 3d,i). VX-787 retains hydrogen bond interactions with Lys376 and Glu361 as seen in the  $m^7$ GTP-PB2cap complex. Similarly, the sidechain of His357 is in the position to sandwich the guanine moiety on one side of the binding pocket, and the sidechains of Phe404 and Phe323 on the other [10]. In addition to conserved interactions, the bicyclooctane group of VX-787 contributes to the high affinity via hydrophobic interactions with the sidechains of Met431 and Phe325. The carboxy group of the

bicyclooctane moiety forms a salt bridge with the side chain of Arg355 and hydrogen bonds to the sidechains of His357 and Gln406 by way of water molecules. It is also possible that Lys339 forms a salt bridge with the carboxy group of the bicyclooctane moiety [10].

These comparisons confirm that the interactions between the cap and PB2cap are mostly conserved among different virus strains and ligands although the unliganded structures of PB2cap may be slightly different.

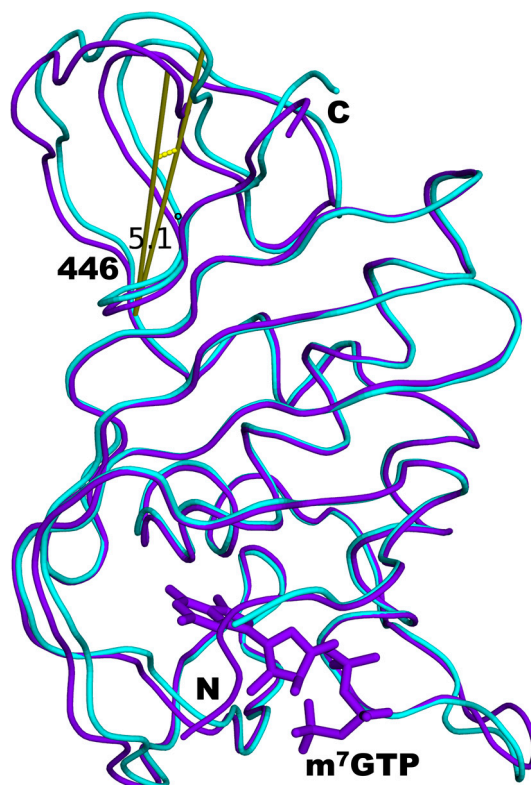


**Figure 3.** (a)  $m^7$ GTP-bound PB2cap from A/Duck/Shantou/4610/2003 (H5N1) (PDB code: 4CB4) (cyan) [12]; (b)  $m^7$ GTP-bound PB2cap from A/Hong Kong/1/68 (H3N2) (PDB code: 4EQK) (marine) [11]; (c) PB2cap from A/Duck/Shantou/4610/2003 (H5N1) bound to compound 8f (PDB code: 4CB5) (wheat) [12]; (d) PB2cap from A/Victoria/3/1975 (H3N2) bound to VX-787 (PDB code: 4P1U) (blue) [14]; (e) The cap-binding site in  $m^7$ GTP-bound mutant PB2cap from A/California/07/2009 (H1N1) (PDB code: 5EG7) (red) [10]; (f) 4CB4 (cyan) aligned with 5EG7 (red); (g) 4EQK (marine) aligned with 5EG7 (red); (h) 4CB5 (wheat) aligned with 5EG7 (red); (i) 4P1U (blue) aligned with 5EG7 (red). Hydrogen bonds are shown with yellow dotted lines, and  $\pi$ - $\pi$  interactions are shown with purple dotted lines [10,14,23,25].

### 3.4. Structural Comparisons between 5EG7 and Unliganded Mutant PB2cap

A clear crystal prism that diffracted X-rays to 1.52 Å resolution was used to solve the structure of unliganded mutant PB2cap from A/California/07/2009 H1N1 (mutant CA09-PB2cap) (PDB code: 5WOP) [13] by standard crystallography methods [19,22]. Superposition of only the N-terminal subdomains (residues 319–445) of 5WOP Chain A and  $m^7$ GTP bound 5EG7 revealed a 5.1° shift of the C-terminal subdomain (residues 446–479) upon ligand binding (Table 2 and Figure 4) [24,26].

B-putty drawings (Figure 4) along with RMSD and residual B-factor comparisons between 5WOP and 5EG7 (Figure 5a,b) demonstrated an increased flexibility of the PB2cap C-terminal subdomain (residues 446–479) and an increased rigidity of the N-terminal subdomain upon ligand occupation of the cap-binding domain [13].



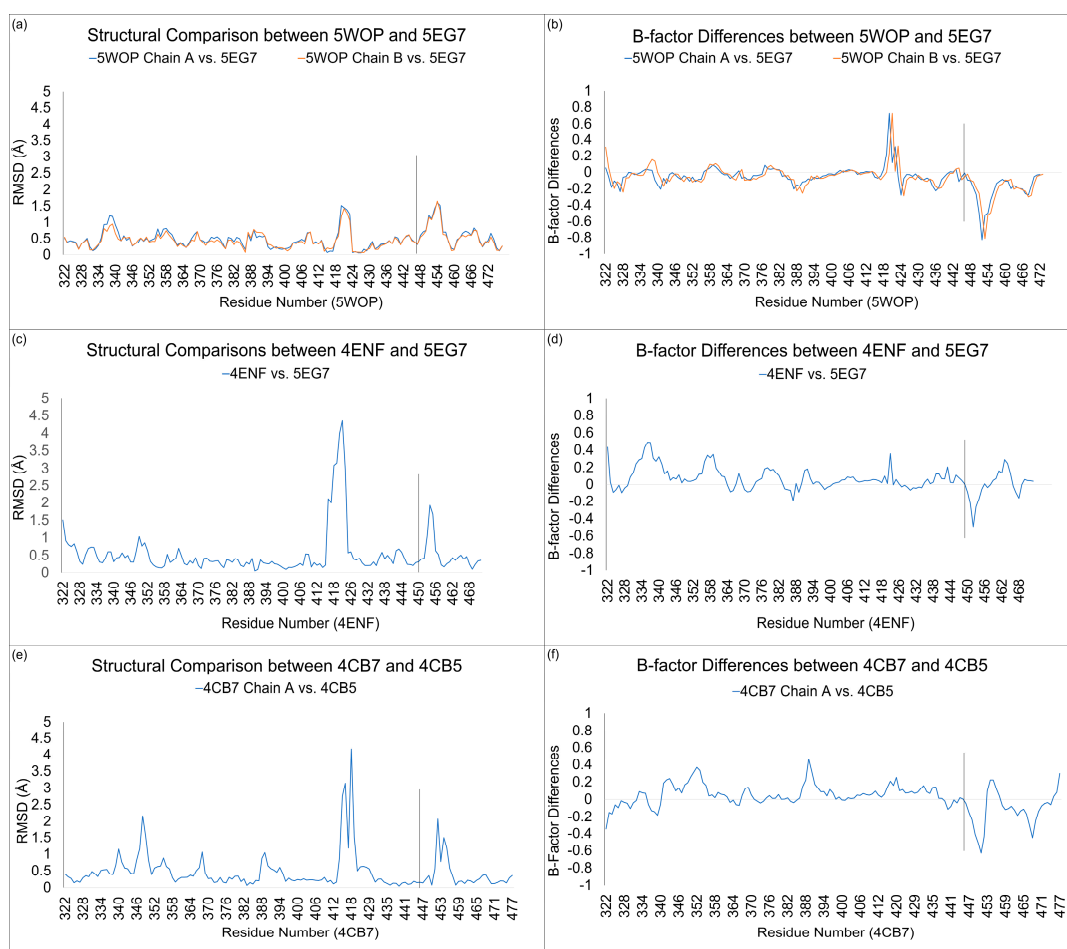
**Figure 4.** Using only the coordinates from the N-subdomain, 5WOP Chain A (cyan) was superimposed with 5EG7 (purple). The yellow line represents a  $5.1^\circ$  rotation of the C-terminal subdomain [10,26].

### 3.5. Structural Comparisons between Other PB2cap Structures

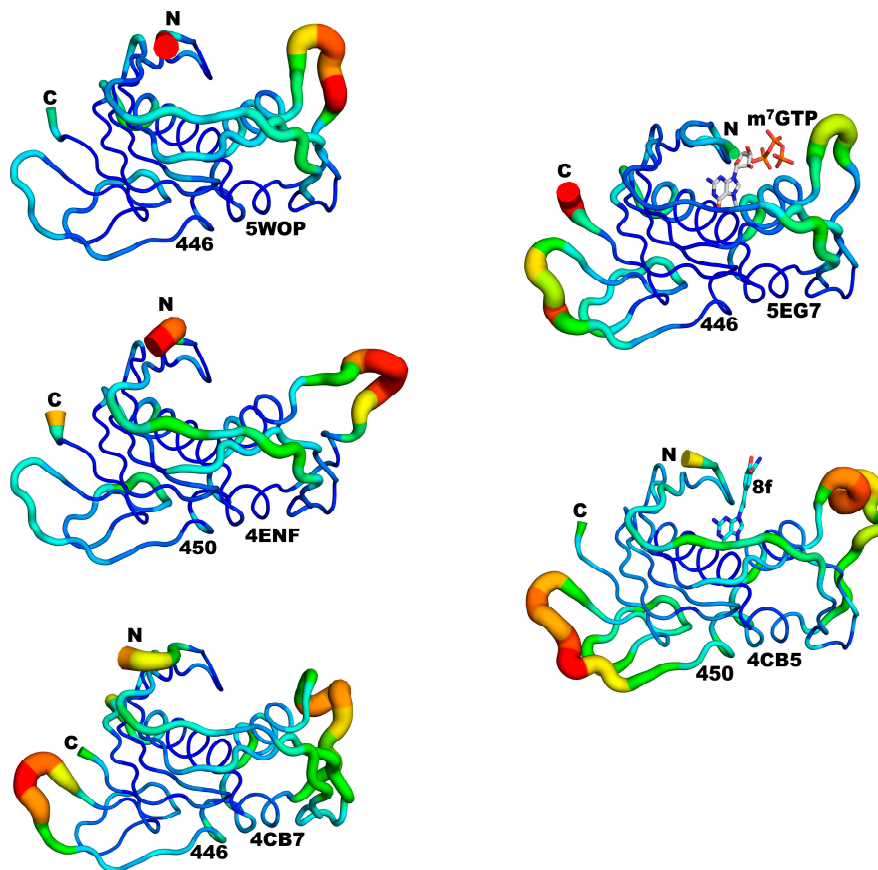
In addition to the comparison mentioned in Section 2.4, additional B-factor and RMSD comparisons were drawn between 5WOP [13] and 5EG7 [10], 4ENF [11] and 5EG7, and between PB2cap from A/Duck/Shantou/4610/2003 (H5N1) (PDB code: 4CB7) Chain A and PB2cap from A/Duck/Shantou/4610/2003 (H5N1) in complex with compound 8f (PDB code: 4CB5) [12] (Table 2 and Figure 5). These comparisons indicate that these structural changes are universal across different strains. The changes in the structural flexibility are more visible in the B-putty drawings (Figure 6). Influenza A PB2cap can be divided into two subdomains: the ligand binding N-subdomain (residues 319–445) and the C-terminal subdomain (residues 446–479). These residues were numbered respective to those of PB2cap from A/California/07/2009 H1N1. Rotations up to  $5.1^\circ$  of the C<sup>+</sup>-terminal subdomain accompanied by increased RMSD values were observed when only the N-subdomains of ligand-bound and unbound PB2cap were superposed (Table 2). The higher RMSD values of the C<sup>+</sup>-subdomains compared to the RMSD of the individually superposed C-subdomains confirm the separation and rotation of the C-subdomain from the N-subdomain upon ligand binding (Table 2). Additionally, the B-putty drawings of each structure in Figure 6 reveal the structural changes induced by ligand binding. Additional RMSD and B-factor comparisons between 4ENF and 5EG7 indicate that these morphologies are present in other H1N1 PB2cap structures (Figure 5c,d, Figure 6, and Table 2). It was determined that these changes are also present among 4CB7 Chain A and 4CB5 by RMSD and B-factor comparisons (Figure 5e,f, Figure 6, and Table 2). These observations indicate that the ligand-induced structural metamorphosis is universal in different strains of influenza A PB2cap [13].



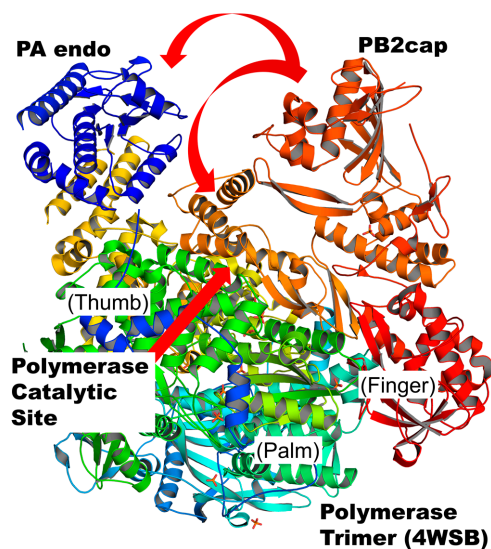
If compared to the anatomy of a right-sided human hand, the polymerase complex of the influenza virus consists of three regions, the fingers, thumb, and palm region (Figure 7). PB2cap is located near the finger region, the PA endonuclease is located near the thumb region, and PB1 is the palm region. We suggest that when the cap binds in the N-subdomain, the induced conformational change results in the separation of the ligand binding domain from the C-subdomain, facilitating the movement of the snatched pre-mRNA toward the PA endonuclease for cleavage, and then towards PB1, the catalytic active site, to serve as the primer for transcription [13].



**Figure 5.** (a) RMSD and (b) B-factor comparisons between chain A and chain B of 5WOP [13] and 5EG7 [10]. Regions of residues 334–340, 418–424 and 448–454 have notable differences. The shrinkage of B-factors between residues 418–424 in 5EG7 confirm that the ligand binding domain is stabilized by m<sup>7</sup>GTP binding, whereas the increase of B-factors between residues 446–479 in 5EG7 suggests that the C-terminal domain is no longer stabilized by the ligand binding domain. (c) RMSD and (d) B-factor comparisons between 4ENF [11] and 5EG7. (e) RMSD and (f) B-factor comparisons between chain A of 4CB7 and 4CB5 [12]. In this figure, CCP4 Baverage [19] was used to calculate B-factor differences and Chimera 1.11.2 [24] was used to calculate RMSD values. The vertical line represents the division between the N- and C-subdomains [13].



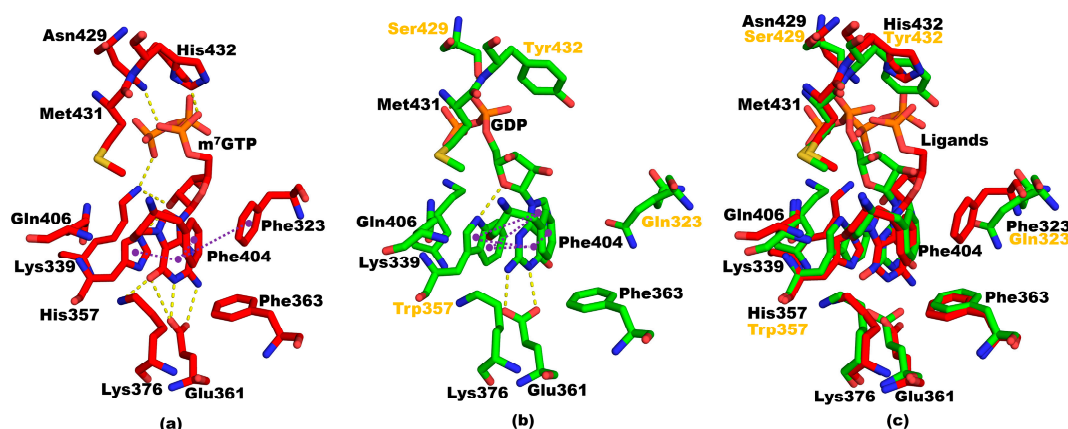
**Figure 6.** B-putty drawings [26] of 5WOP Chain A [13], 5EG7 [10], 4ENF [11], 4CB5, and 4CB7 Chain A [12]. The diameter of the tubes is proportional to the B-factor. The darkest shade of red corresponds to a higher B-factor, whereas the darkest shade of blue corresponds to a lower B-factor [13,23].



**Figure 7.** The ribbon drawing of influenza A virus polymerase complex using coordinates from PDB 4WSB [27]. PB2cap is colored in orange, and the PA endonuclease domain is colored in blue. Arrows indicate possible movements of PB2cap induced by cap binding. Regions respective to a right-sided human hand are shown in parenthesis [13,23].

### 3.6. Structural Comparisons between 5EG7 and PB2cap from B/Jiangxi/BV/2006 PB2cap Bound to GDP

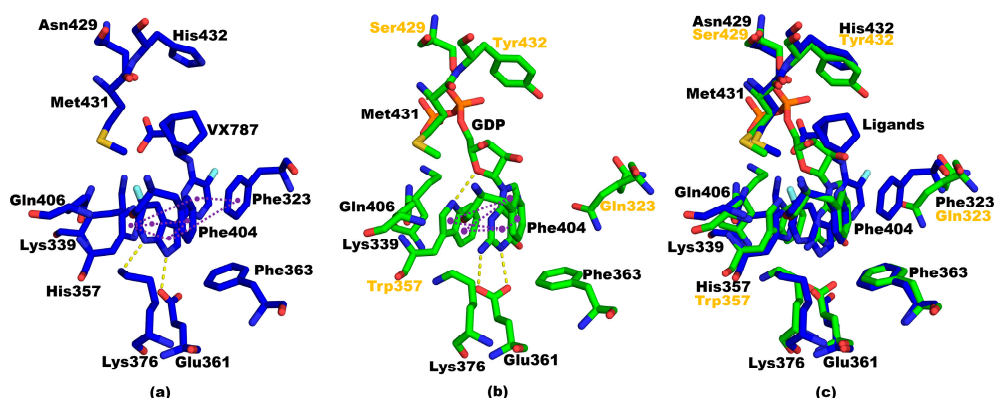
The guanine in GDP bound to influenza B virus PB2cap from B/Jiangxi/BV/2006 (PDB code: 4Q46) [15] is rotated 180° relative to the orientation of the guanine moiety in m<sup>7</sup>GTP bound to influenza A virus PB2cap (Figure 8) [10]. The carbonyl at position 6 of GDP forms hydrogen bonds with the sidechain of Arg322 in 4Q46 instead of Lys376 like in 5EG7. Gln323, unable to participate in  $\pi$ - $\pi$  stacking with the guanine moiety, replaces Phe323 in the cap-binding site in 4Q46. This leaves Phe404 as the only participant in  $\pi$ - $\pi$  stacking interactions on this side of the binding site. On the other side of the guanine moiety, Trp357 replaces His357 and participates in stronger  $\pi$ - $\pi$  stacking interactions [10].



**Figure 8.** (a) Mutant PB2cap from A/California/07/2009 (H1N1) bound to m<sup>7</sup>GTP (PDB code: 5EG7) (red) [10]; (b) PB2cap from B/Jiangxi/BV/2006 PB2cap bound to GDP (PDB code: 4Q46) (green) [15]; (c) 5EG7 (red) superposed with 4Q46 (green). The gold residue labels highlight the differences in 4Q46 relative to 5EG7. Hydrogen bonds are depicted with yellow dotted lines, and  $\pi$ - $\pi$  force interactions are depicted with purple dotted lines [10,23,25].

### 3.7. Structural Comparisons between 4Q46 and 4P1U 6

Comparisons between 4Q46 and 4P1U identified an important residue that may account for the inability of VX-787 to bind to the influenza B virus PB2cap (Figure 9) [10]. In 4Q46, His432 is replaced by Tyr432 which has a bulkier side-chain, causing steric hindrance to the accommodation of VX-787 in the cap-binding site of influenza B virus PB2cap [10].

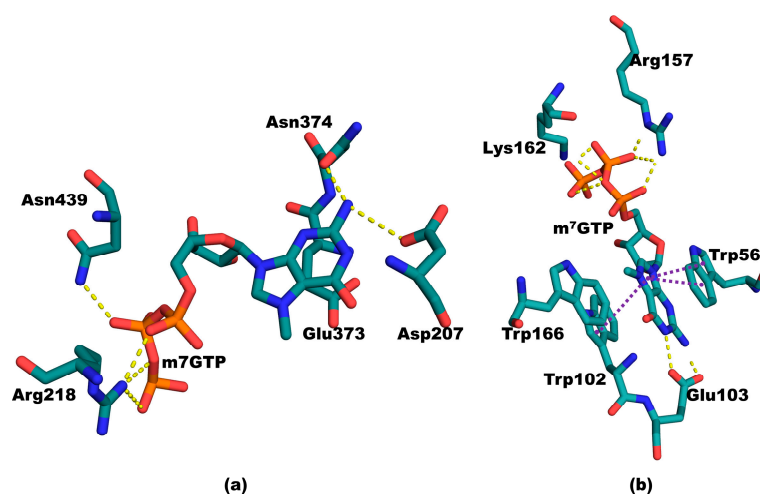


**Figure 9.** (a) PB2cap from A/Victoria/3/1975 (H3N2) bound to VX-787 (PDB code: 4P1U) (blue) [14]; (b) PB2cap from B/Jiangxi/BV/2006 PB2cap bound to GDP (PDB code: 4Q46) (green) [15]; (c) 4P1U superposed with 4Q46. Gold labelled residues represent 4Q46 differences relative to 4P1U. Hydrogen bonds are labelled with yellow dotted lines, and  $\pi$ - $\pi$  force interactions are labelled with purple dotted lines [10,23,25].

### 3.8. The Difference between PB2cap and Human Cap-Binding Proteins: Cap-Specific mRNA (Nucleoside-2'-O-)-methyltransferase 1 and Eukaryotic Translation Initiation Factor 4E

A desirable cap-binding inhibitor must bind the influenza A and B virus PB2cap without binding the host cap-binding proteins, such as the cap-specific mRNA (nucleoside-2'-O-)-methyltransferase 1 and the eukaryotic translation initiation factor 4E (eIF4E).  $m^7$ GTP-bound structures of the methyltransferase (PDB code: 4N49) and of eIF4E (PDB code: 4TQB) highlight the interactions of the cap with human proteins (Figure 10a,b) [10].

In the binding site of 4N49, the position 1 nitrogen atom forms a hydrogen bond with the side chain of Asp207, and the position 2 amino group forms a hydrogen bond with the side chain of Asn374 (Figure 10a) [10]. Glu373 exhibits anion-aromatic stacking with the guanine moiety in lieu of any  $\pi$ - $\pi$  force interactions as seen in other structures. There are clear hydrogen bonding or salt bridge interactions between Asn439 and Arg218 with the  $\alpha$ -,  $\beta$ -, and  $\gamma$ -phosphate groups. In the binding site of 4TQB, the main-chain amide nitrogen atom of Trp102 forms a hydrogen bond with the carbonyl oxygen atom in the guanine moiety (Figure 10b). The guanine moiety is sandwiched between Trp56 and Trp102. The side chain of Trp166 hangs near the position 7 methyl group. Unlike what is seen in PB2cap, the position 7 methyl group is not enclosed in a hydrophobic pocket. The  $\alpha$ - and  $\beta$ -phosphate groups interact with the side chain of Arg157, and the  $\beta$ - and  $\gamma$ -phosphate groups interact with the side chain of Lys162. The nitrogen atom at position 1 and the amino group linked to position 2 form two hydrogen bonds to the side chain of Glu103 (Figure 10b) [10]. The binding sites of 4N49 and 4TQB do not interact with the pyranose ring, which serves merely as a linker between the guanine moiety and the triphosphate. Primarily, the binding site interacts with the guanine moiety and the triphosphate.



**Figure 10.** (a)  $M^7$ GTP bound to the cap-binding site in the human cap-specific mRNA (nucleoside-2'-O-)-methyltransferase 1 (pdb code: 4N49) [16] and (b)  $m^7$ GTP bound to the Eukaryotic translation initiation factor 4E (eIF4E) cap-binding site (PDB code: 4TQB). Hydrogen bonds are shown with yellow dotted lines, and  $\pi$ - $\pi$  interactions are shown with purple dotted lines [10,17,23,25].

## 4. Conclusions

Wild-type,  $m^7$ GTP-mutant, and mutant PB2cap from A/California/07/2009 H1N1 were expressed, purified, crystallized, and subjected to X-ray diffraction experiments. Structures are reported at 1.54 Å, 1.40 Å, and 1.52 Å, respectively. Structural comparisons of known liganded and unliganded structures reveal the conformation of key residues in the influenza A and B PB2cap binding domains and the dissociation of the C-terminal subdomain upon ligand binding. We theorize that this movement assists the transport of the host cell's 5'-cap snatched pre-mRNA towards the PA endonuclease for cleavage and then to the catalytic subunit, PB1, to function as a primer for transcription. A novel inhibitor that binds Influenza A and B PB2cap will overcome the steric hindrance imposed by Tyr432 in

the FluB PB2cap binding domain and will retain hydrogen bonding interactions with Arg322, Lys339, Glu361 and Lys376,  $\pi$ - $\pi$  force interactions with Phe404 and His357 (Trp357 in the FluB PB2cap binding domain), salt bridge interactions with Arg355, Lys339, His432, Asn429, and hydrophobic interactions with Phe325, Phe404, Gln406, and Met431.

For inhibitor design, the ligand should maintain the hydrophobic interactions of the guanine moiety with PB2cap. The hydrogen bonds with Glu361 and Lys376 must be satisfied. Most importantly, an appropriate group linked to the scaffold must accommodate hydrophobic interactions with either a His or Tyr sidechain in the pocket at residue 432.

**Acknowledgments:** Data were collected at Southeast Regional Collaborative Access Team (SER-CAT) 22-ID (or 22-BM) beamline at the Advanced Photon Source, Argonne National Laboratory. Supporting institutions may be found at [www.ser-cat.org/members.html](http://www.ser-cat.org/members.html). Use of the Advanced Photon Source was supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, under Contract No. W-31-109-Eng-38. This work was supported in part by a grant to Xiaofeng Zheng from Beijing Natural Science Foundation (5152012).

**Author Contributions:** Ming Luo, Xiaofeng Zheng, Amanda E. Constantinides, and Chelsea C. Severin conceived and designed the experiments; Amanda E. Constantinides and Chelsea C. Severin performed the experiments; Amanda E. Constantinides, Chelsea C. Severin, and Ryan H. Gumpfer analyzed the data; Ming Luo and Xiaofeng Zheng contributed reagents/materials/analysis tools; Amanda E. Constantinides and Ming Luo wrote the paper.

**Conflicts of Interest:** The authors declare no conflict of interest. The founding sponsors had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, and in the decision to publish the results.

## References

1. Marcotrigiano, J.; Gingras, A.C.; Sonenberg, N.; Burley, S.K. Cap-dependent translation initiation in eukaryotes is regulated by a molecular mimic of eIF4G. *Mol. Cell* **1999**, *3*, 707–716. [[CrossRef](#)]
2. Mazza, C.; Segref, A.; Mattaj, I.W.; Cusack, S. Large-scale induced fit recognition of an m(7)gpppg cap analogue by the human nuclear cap-binding complex. *EMBO J.* **2002**, *21*, 5548–5557. [[CrossRef](#)] [[PubMed](#)]
3. Li, M.L.; Rao, P.; Krug, R.M. The active sites of the influenza cap-dependent endonuclease are on different polymerase subunits. *EMBO J.* **2001**, *20*, 2078–2086. [[CrossRef](#)] [[PubMed](#)]
4. Guilligay, D.; Tarendeau, F.; Resa-Infante, P.; Coloma, R.; Crepin, T.; Sehr, P.; Lewis, J.; Ruigrok, R.W.; Ortin, J.; Hart, D.J.; et al. The structural basis for cap binding by influenza virus polymerase subunit PB2. *Nat. Struct. Mol. Biol.* **2008**, *15*, 500–506. [[CrossRef](#)] [[PubMed](#)]
5. Sugiyama, K.; Obayashi, E.; Kawaguchi, A.; Suzuki, Y.; Tame, J.R.; Nagata, K.; Park, S.Y. Structural insight into the essential pb1-pb2 subunit contact of the influenza virus RNA polymerase. *EMBO J.* **2009**, *28*, 1803–1811. [[CrossRef](#)] [[PubMed](#)]
6. Plotch, S.J.; Bouloy, M.; Ulmanen, I.; Krug, R.M. A unique cap(m7gpppxm)-dependent influenza virion endonuclease cleaves capped RNAs to generate the primers that initiate viral RNA transcription. *Cell* **1981**, *23*, 847–858. [[CrossRef](#)]
7. Jung, T.E.; Brownlee, G.G. A new promoter-binding site in the PB1 subunit of the influenza A virus polymerase. *J. Gen. Virol.* **2006**, *87*, 679–688. [[CrossRef](#)] [[PubMed](#)]
8. Yuan, P.; Bartlam, M.; Lou, Z.; Chen, S.; Zhou, J.; He, X.; Lv, Z.; Ge, R.; Li, X.; Deng, T.; et al. Crystal structure of an avian influenza polymerase PA(n) reveals an endonuclease active site. *Nature* **2009**, *458*, 909–913. [[CrossRef](#)] [[PubMed](#)]
9. Dias, A.; Bouvier, D.; Crepin, T.; McCarthy, A.A.; Hart, D.J.; Baudin, F.; Cusack, S.; Ruigrok, R.W. The cap-snatching endonuclease of influenza virus polymerase resides in the PA subunit. *Nature* **2009**, *458*, 914–918. [[CrossRef](#)] [[PubMed](#)]
10. Severin, C.; Rocha de Moura, T.; Liu, Y.; Li, K.; Zheng, X.; Luo, M. The cap-binding site of influenza virus protein pb2 as a drug target. *Acta Crystallogr. D Struct. Biol.* **2016**, *72*, 245–253. [[CrossRef](#)] [[PubMed](#)]
11. Liu, Y.; Qin, K.; Meng, G.; Zhang, J.; Zhou, J.; Zhao, G.; Luo, M.; Zheng, X. Structural and functional characterization of k339t substitution identified in the PB2 subunit cap-binding pocket of influenza A virus. *J. Biol. Chem.* **2013**, *288*, 11013–11023. [[CrossRef](#)] [[PubMed](#)]

12. Pautus, S.; Sehr, P.; Lewis, J.; Fortune, A.; Wolkerstorfer, A.; Szolar, O.; Guilligay, D.; Lunardi, T.; Decout, J.L.; Cusack, S. New 7-methylguanine derivatives targeting the influenza polymerase PB2 cap-binding domain. *J. Med. Chem.* **2013**, *56*, 8915–8930. [[CrossRef](#)] [[PubMed](#)]
13. Constantinides, A.; Gumpfer, R.H.; Severin, C.; Luo, M. High-resolution structure of influenza a PB2cap binding domain illuminates changes induced by ligand binding. *Acta Crystallogr. F* **2018**. Accepted.
14. Clark, M.P.; Ledebor, M.W.; Davies, I.; Byrn, R.A.; Jones, S.M.; Perola, E.; Tsai, A.; Jacobs, M.; Nti-Addae, K.; Bandarage, U.K.; et al. Discovery of a novel, first-in-class, orally bioavailable azaindole inhibitor (vx-787) of influenza PB2. *J. Med. Chem.* **2014**, *57*, 6668–6678. [[CrossRef](#)] [[PubMed](#)]
15. Liu, Y.; Yang, Y.; Fan, J.; He, R.; Luo, M.; Zheng, X. The crystal structure of the PB2 cap-binding domain of influenza b virus reveals a novel cap recognition mechanism. *J. Biol. Chem.* **2015**, *290*, 9141–9149. [[CrossRef](#)] [[PubMed](#)]
16. Smietanski, M.; Werner, M.; Purta, E.; Kaminska, K.H.; Stepinski, J.; Darzynkiewicz, E.; Nowotny, M.; Bujnicki, J.M. Structural analysis of human 2'-o-ribose methyltransferases involved in mRNA cap structure formation. *Nat. Commun.* **2014**, *5*, 3004. [[CrossRef](#)] [[PubMed](#)]
17. Papadopoulos, E.; Jenni, S.; Kabha, E.; Takroui, K.J.; Yi, T.; Salvi, N.; Luna, R.E.; Gavathiotis, E.; Mahalingam, P.; Arthanari, H.; et al. Structure of the eukaryotic translation initiation factor eIF4E in complex with 4EIG-1 reveals an allosteric mechanism for dissociating eIF4G. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, E3187–E3195. [[CrossRef](#)] [[PubMed](#)]
18. Otwinowski, Z.; Minor, W. [20] Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **1997**, *276*, 307–326. [[PubMed](#)]
19. Winn, M.D.; Ballard, C.C.; Cowtan, K.D.; Dodson, E.J.; Emsley, P.; Evans, P.R.; Keegan, R.M.; Krissinel, E.B.; Leslie, A.G.; McCoy, A.; et al. Overview of the CCP4 suite and current developments. *Acta Crystallogr. D Biol. Crystallogr.* **2011**, *67*, 235–242. [[CrossRef](#)] [[PubMed](#)]
20. Kabsch, W. Xds. *Acta Crystallogr. D Biol. Crystallogr.* **2010**, *66*, 125–132. [[CrossRef](#)] [[PubMed](#)]
21. Adams, P.D.; Afonine, P.V.; Bunkoczi, G.; Chen, V.B.; Davis, I.W.; Echols, N.; Headd, J.J.; Hung, L.W.; Kapral, G.J.; Grosse-Kunstleve, R.W.; et al. Phenix: A comprehensive python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **2010**, *66*, 213–221. [[CrossRef](#)] [[PubMed](#)]
22. Emsley, P.; Cowtan, K. Coot: Model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **2004**, *60*, 2126–2132. [[CrossRef](#)] [[PubMed](#)]
23. *The Pymol Molecular Graphics System*, version 1.3; Schrödinger, LLC: New York, NY, USA, 2010.
24. Pettersen, E.F.; Goddard, T.D.; Huang, C.C.; Couch, G.S.; Greenblatt, D.M.; Meng, E.C.; Ferrin, T.E. Ucsf chimera—A visualization system for exploratory research and analysis. *J. Comput. Chem.* **2004**, *25*, 1605–1612. [[CrossRef](#)] [[PubMed](#)]
25. Biovia, D.S. Biovia Discovery Studio 2017 r2: A Comprehensive Predictive Science Application for the Life Sciences. Available online: <http://accelrys.com/products/collaborative-science/biovia-discovery-studio> (accessed on 29 December 2017).
26. *The Pymol Molecular Graphics System*, version 1.8; Schrödinger, LLC: New York, NY, USA, 2015.
27. Reich, S.; Guilligay, D.; Pflug, A.; Malet, H.; Berger, I.; Crepin, T.; Hart, D.; Lunardi, T.; Nanao, M.; Ruigrok, R.W.; et al. Structural insight into cap-snatching and rna synthesis by influenza polymerase. *Nature* **2014**, *516*, 361–366. [[CrossRef](#)] [[PubMed](#)]

