High-resolution structure of the *Influenza A virus* PB2cap binding domain illuminates the changes induced by ligand binding

Amanda Constantinides, Ryan Gumpfer, Chelsea Severin and Ming Luo*

Department of Chemistry, Georgia State University, 50 Decatur Street, Atlanta, GA 30303, USA. *Correspondence e-mail: mluo@gsu.edu

Received 26 September 2017

Accepted 14 January 2018

Edited by N. Sträter, University of Leipzig, Germany

Keywords: high resolution structures; transcription; polymerase basic protein 2; RNA polymerase; cap-snatching; domain movement; *Influenza A virus*.

PDB reference: native influenza A PB2cap, 5wop

Supporting information: this article has supporting information at journals.iucr.org/f

In the face of increasing drug resistance and the rapidly increasing necessity for practicality in clinical settings, drugs targeting different viral proteins are needed in order to control influenza A and B. A small molecule that tenaciously adheres to the PB2cap binding domain, part of the heterotrimeric RNA polymerase machinery of *influenza A virus* and *influenza B virus*, is a promising drug candidate. Understanding the anatomic behavior of PB2cap upon ligand binding will aid in the development of a more robust inhibitor. In this report, the anatomic behavior of the *influenza A virus* PB2cap domain is established by solving the crystal structure of native *influenza A virus* PB2cap at 1.52 Å resolution. By comparing it with the ligand-bound structure, the dissociation and rotation of the ligand-binding domain in PB2cap from the C-terminal domain is identified. This domain movement is present in many PB2cap structures, suggesting its functional relevance for polymerase activity.

1. Introduction

Resistance to available drug treatments for influenza emerges in sync with the development of new treatment strategies. To gain the upper hand, scientists are compiling a greater understanding of the structure and function of influenza viruses. *Influenza A virus* (FluA) and *influenza B virus* (FluB) are negative-sense single-stranded viruses, and their genomes contain eight RNA segments that code for 11 proteins: PB2, PB1, PB1-F2, PA, HA, NP, NA, M1, M2, NS1 and NEP. Polymerase basic protein 2 (PB2), which is part of the influenza virus RNA polymerase machinery (PB2, PB1 and PA), ‘snatches’ the cap of the host cell’s pre-mRNA; the N-terminal endonuclease domain of PA then cleaves it 10–13 nucleotides downstream (Tarendeau *et al.*, 2007). The PB2 domain that has a specific cap-binding site is named PB2cap. The 5'-capped RNA fragment serves as the primer for the synthesis of viral mRNA by the PB1 subunit (Doan *et al.*, 1999). After cap binding to PB2, how the host cell’s pre-mRNA is cleaved by PA is not fully understood. According to the crystal structure of the *influenza A virus* polymerase complex, PB2cap is located away from PA (Pflug *et al.*, 2014). We compared our new structure of ligand-free PB2cap with that of ligand-bound PB2cap, and found a universal movement of PB2cap induced by ligand binding. This structural movement may contribute to the delivery of the snatched pre-mRNA to PA for cleavage and then to the active site of the polymerase for the synthesis of viral mRNA.

Table 1

Recombinant protein.

Source organism	<i>Influenza A virus</i> strain A/California/08/2009 (H1N1)
DNA source	<i>Influenza A virus</i> strain A/California/08/2009 (H1N1)
Expression vector	pET-28a
Expression host	<i>Escherichia coli</i>
Complete amino-acid sequence of the construct produced	SSFSFGGFTFKRTSGSSVKKEEVLTGNLQ TLKIRVHEGYEFTMVGRRTAILRKAT RRLIQLIVSGRDEQSIAEAIIVAMVFSQ EDCMIKAVRGDLNFGSGLNPMHQLLRHF QKDAKVLFGNWGIESIDNVGMIGILPD MTPSTEMSLRGIRVSKM

Table 2

Crystallization.

Method	Manual hanging-drop vapor diffusion
Plate type	24-well plate
Temperature (K)	277
Protein concentration (mg ml ⁻¹)	11
Buffer composition of protein solution	50 mM Tris pH 8.0, 200 mM NaCl, 2 mM DTT, 5%(v/v) 10 mM inhibitor in DMSO
Composition of reservoir solution	0.1 M ammonium acetate, 0.1 M bis-tris pH 5.86, 20%(v/v) PEG 8000
Volume and ratio of drop	2 µl (1:1 ratio)
Volume of reservoir (µl)	500

2. Materials and methods

2.1. Macromolecule production

The expression and purification of PB2cap from *influenza A virus* strain A/California/07/2009H1N1 (CA09-PB2cap) has previously been established (Table 1; Severin *et al.*, 2016).

2.2. Crystallization

CA09-PB2cap was incubated with a small-molecule inhibitor for 2 h and subjected to crystallization conditions

Table 3

Data collection and processing.

Values in parentheses are for the outer shell.

Diffraction source	22-BM, APS
Wavelength (Å)	1
Detector	MAR300
Crystal-to-detector distance (mm)	132
Rotation range per image (°)	1
Total rotation range (°)	180
Exposure time per image (s)	1
Space group	<i>P</i> 12 ₁ 1
<i>a</i> , <i>b</i> , <i>c</i> (Å)	36.86, 72.43, 61.32
α , β , γ (°)	90.00, 103.18, 90.00
Mosaicity (°)	0.249
Resolution range (Å)	36.21–1.52 (1.574–1.520)
Total No. of reflections	93971 (9544)
No. of unique reflections	47925 (4803)
Completeness (%)	99.17 (99.94)
Multiplicity	2.0 (2.0)
$\langle I/\sigma(I) \rangle$	11.05 (2.03)
$R_{\text{p.i.m.}}$	0.04558 (0.4732)
Overall <i>B</i> factor from Wilson plot (Å ²)	13.4

discovered using the JCSG-*plus* suite from Molecular Dimensions (well H6) at 277 K as described in Table 2. Crystals in hanging drops were observed in the manually set up tray approximately two weeks later (Fig. 1*a*). Diffraction data (Fig. 1*b*) were collected from a clear crystal prism buried within a cluster of rods.

2.3. Data collection and processing

To fish out the crystal to be flash-cooled in liquid nitrogen, 10 µl cryoprotectant solution (JCSG-*plus* condition H6 with 20% glycerol) was added to the hanging drop. After 2 min, the drop was aspirated and pipetted into 500 µl cryoprotectant. Using the smallest fitting cryo-pin, the crystal was fished out of the cryoprotectant solution and flash-cooled in liquid nitrogen for shipping to the Southeast Regional Collaborative Access Team (SER-CAT). A complete data set was collected on

Figure 1

(*a*) A cluster of PB2cap crystal rods. (*b*) A diffraction image from data collected at the Advanced Photon Source at Argonne National Laboratory.

Figure 2

(a) R.m.s.d. and (b) *B*-factor comparisons between each chain of PDB entries 5wop and 5eg7 (Severin *et al.*, 2016). The main structural changes are at residues 334–340, 418–424 and 448–454. The decrease in *B* factors between residues 418 and 424 in PDB entry 5eg7 (Severin *et al.*, 2016) confirmed that the ligand-binding domain was stabilized by m⁷GTP binding, whereas the increase in *B* factors between residues 446 and 479 in PDB entry 5eg7 (Severin *et al.*, 2016) suggested that the C-subdomain was no longer stabilized by the N-subdomain after its separation induced by ligand binding. A vertical line distinguishes the boundary between the two subdomains in all panels. (c) R.m.s.d. and (d) *B*-factor comparisons between PDB entries 4enf (Liu *et al.*, 2013) and 5eg7 (Severin *et al.*, 2016). (e) R.m.s.d. and (f) *B*-factor comparisons between each chain of PDB entry 4cb7 chain A and PDB entry 4cb5 (Pautus *et al.*, 2013). Structural changes in all comparisons are similar. (g) and (h) are comparisons between chain A and chain B of PDB entry 5wop, which do not show any significant differences except for the *B* factors of the undefined loop. This confirms that the differences observed from other structures were independent of crystal contacts. In this figure, *B*-factor differences were calculated using *BAVERAGE* in *CCP4* (Winn *et al.*, 2011) and r.m.s.d. calculations were performed using *UCSF Chimera* v.1.11.2 (Pettersen *et al.*, 2004).

Table 4

Structure solution and refinement.

Values in parentheses are for the outer shell.

Resolution range (Å)	36.21–1.52 (1.559–1.520)
Completeness (%)	99.17
σ Cutoff	$F > 0.000\sigma(F)$
No. of reflections, working set	45487 (3369)
No. of reflections, test set	2438 (194)
Final R_{cryst}	0.1498 (0.2830)
Final R_{free}	0.2044 (0.3490)
No. of non-H atoms	
Protein	2489
Ligand (glycerol)	24
Solvent	336
Total	2849
R.m.s. deviations	
Bonds (Å)	0.026
Angles (°)	2.22
Average B factors (Å ²)	
Overall	22.01
Protein	20.10
Ligand	87.37
Solvent	31.54
Ramachandran plot	
Favored (%)	98.09
Allowed (%)	1.91
Outliers (%)	0.00

beamline 22-BM at the Advanced Photon Source, Argonne National Laboratory as described in Table 3. Data were processed with *XDS* and scaled with *AIMLESS* in *CCP4* (Winn *et al.*, 2011). This crystal was determined to have pseudo-translational symmetry using *phenix.xtriage* in the *PHENIX* graphical user interface (Adams *et al.*, 2010).

2.4. Structure solution and refinement

TRUNCATE in *CCP4* was used to convert intensity values (I) to structure amplitudes (F) (Winn *et al.*, 2011). Molecular replacement using PDB entry 5eg7 (Severin *et al.*, 2016) as the probe structure was performed using *MOLREP* in *CCP4* (Winn *et al.*, 2011). Isotropic and anisotropic refinement was conducted with *REFMAC* and *Coot* in *CCP4* (Winn *et al.*, 2011; Emsley & Cowtan, 2004). *MolProbity* and *POLYGON* in *PHENIX* were used for structural validation throughout refinement (Adams *et al.*, 2010). *TLSANL* in *CCP4* was used to generate translation, libration and screw rotation (TLS) input files to conduct TLS refinement in *REFMAC* (Table 4; Winn *et al.*, 2011). There are two chains of PB2cap per asymmetric unit and no ligand was found in the binding site. The structure was deposited in the Protein Data Bank (PDB) as entry 5wop.

2.5. Structural comparisons

Structural superpositions were performed using *UCSF Chimera* v.1.11.2 (Pettersen *et al.*, 2004) to generate r.m.s.d. values, and *BAVERAGE* in *CCP4* (Winn *et al.*, 2011) was used to generate lists of main-chain residual B factors (Fig. 2). B -factor values were normalized using the min–max rescaled normalization method before calculating the differences. Normalized B -factor values from PB2cap with ligand bound were subtracted from PB2cap without ligand bound. R.m.s.d.

Table 5

Structural comparisons of subdomains.

C' denotes that the r.m.s.d. was calculated between the C-subdomains when only the coordinates of the N-subdomains were used to superimpose the two structures using *UCSF Chimera* (Pettersen *et al.*, 2004).

	R.m.s.d. (Å)			Angular distance (°)	
	N-subdomain	C-subdomain	C'-subdomain	C'-subdomain	C-subdomain rigid body
5wop chain A versus 5eg7	0.3922	0.3978	1.1049	5.1	5.7
5wop chain B versus 5eg7	0.3558	0.4022	1.0638	4.7	5.3
4enf versus 5eg7	0.5547	0.4939	0.6266	4.8	4.1
4cb7 chain A versus 4cb5	0.5032	0.3176	0.5342	3.5	3.1

values and B -factor differences were plotted in *Microsoft Excel*. Superposition was also carried out with *PyMOL* (v.1.3; Schrödinger) and r.m.s.d. values were calculated using *UCSF Chimera* v.1.11.2 (Table 5). To confirm the rigid-body movements of the C-terminal domains relative to the N-terminal domains by an independent method, *domainSelect* in *DynDom* v.2.0 (Hayward & Berendsen, 1998) was used to perform additional superpositions independent of any previous assumptions (Table 5 and Fig. 3). The results from these two methods are consistent with each other.

3. Results and discussion

A histidine tag-fused PB2cap protein containing a thrombin cleavage site was expressed, purified, crystallized and subjected to X-ray diffraction experiments.

Figure 3

Using *domainSelect* in *DynDom* v.2.0 (Hayward & Berendsen, 1998), chain A of PDB entry 5wop (blue) was superimposed with PDB entry 5eg7 (yellow; Severin *et al.*, 2016) independent of any previous assumptions. Relative to the fixed N-terminal domain (residues 446–479), a rigid-body rotation of 5.7° of the C-terminal domain (residues 446–479) is observed along the axis of rotation (red). This drawing was prepared with *RASMOL* (Sayle & Milner-White, 1995). The angle of rotation, the N-terminus, C-terminus, residue 446 and m⁷GTP (yellow) are labelled. *POV-ray* v.3.6 was used to ray-trace this image (<http://www.povray.org/>).

A crystallization screen was performed to assess potential crystallization conditions for CA09-PB2cap incubated with a candidate small-molecule inhibitor compound dissolved in 5% DMSO. Based on the visible crystal quality, condition H6 from the JCSG-*plus* suite [0.1 M ammonium acetate, 0.1 M bis-tris pH 5.86, 20%(v/v) PEG 8000] was selected for optimization *via* manual hanging-drop vapor diffusion. Diffraction-quality crystals of CA09-PB2cap were obtained and X-ray diffraction data were collected at 100 K using 20% glycerol as a cryoprotectant (§§2.3 and 2.4).

Figure 4
B-putty drawings of PDB entries 5wop, 5eg7, 4enf, 4cb7 chain A and 4cb5 (Severin *et al.*, 2016; Liu *et al.*, 2013; Pautus *et al.*, 2013). The diameter of the tubes is proportional to the *B* factor. The blue color corresponds to a lower *B* factor, whereas the red color corresponds to a higher *B* factor. The N-terminus, C-terminus, m⁷GTP (PDB entry 5eg7; Severin *et al.*, 2016), compound 8f (PDB entry 4cb5; Pautus *et al.*, 2013) and the residue that separates the two subdomains are labelled.

A crystal obtained from this experiment diffracted X-rays to 1.52 Å resolution. X-ray diffraction experiments were performed on SER-CAT beamline 22-BM at the Advanced Photon Source, Argonne National Laboratory, Chicago, Illinois, USA. The structure of native CA09-PB2cap in space group *P*12₁1 was solved with an *R*_{work} of 0.1498 and an *R*_{free} of 0.2044, and was deposited in the PDB as entry 5wop.

Previously, the structure of CA09-PB2cap in complex with m⁷GTP (a cap analogue) was solved at 1.40 Å resolution (Severin *et al.*, 2016). PB2cap may be further divided into two subdomains: the N-subdomain (residues 319–445; the ligand-binding domain) and the C-subdomain (residues 446–479). The residue numbers refer to those of PB2cap from *influenza virus A* strain A/California/07/2009H1N1. R.m.s.d. and residual *B*-factor comparisons of the new CA09-PB2cap structure (PDB entry 5wop) with those of m⁷GTP-bound CA09-PB2cap (PDB entry 5eg7; Severin *et al.*, 2016) revealed stabilization of the N-subdomain and increased flexibility of the C-subdomain upon ligand binding (Figs. 2a and 2b). When only the N-subdomain of the two structures was used in superposition, the C-subdomain showed a rotation of up to 5.1° accompanied by an increased r.m.s.d. (Table 5). The higher r.m.s.d. between the C'-subdomains using the N-subdomains for superposition compared with the r.m.s.d. between individually superimposed C-subdomains confirmed the separation and rotation of the

Figure 5
Ribbon drawing of the *influenza A virus* polymerase complex using coordinates from PDB entry 4wsb (Reich *et al.*, 2014). The PA subunit is colored green, PB1 cyan and PB2 magenta. The structure of the RNA promoter is also shown. PB2cap is colored yellow and the PA endonuclease domain is colored blue. Arrows indicate the possible movements of PB2cap induced by cap binding.

C-subdomain from the N-subdomain upon ligand binding (Table 5). These claims were substantiated by additional nonbiased, rigid-body superpositions. On selecting the N-subdomain as the fixed rigid body and the C-subdomain as the moving rigid body, rotations of the C-subdomain up to 5.7° were observed upon ligand binding (Fig. 3). Other structural changes are also visible in the *B*-putty drawings of each structure (Fig. 4).

Additional r.m.s.d. and *B*-factor comparisons between PB2cap from *influenza A virus* strain A/Puerto Rico/8/34 (H1N1) (PDB entry 4enf; Liu *et al.*, 2013) and CA09-PB2cap in complex with m⁷GTP (PDB entry 5eg7; Severin *et al.*, 2016) indicated that these structural changes are present in other H1N1 PB2cap structures, suggesting that structural changes induced by ligand binding are common (Figs. 2c, 2d and 4; Table 5).

R.m.s.d. and *B*-factor comparisons were also performed on PB2cap from *influenza A virus* strain Duck/Shantou/4610/2003 (H5N1) (PDB entry 4cb7, chain A) and its complex with compound 8f (PDB entry 4cb5) (Pautus *et al.*, 2013). Similar structural changes induced by ligand binding were found (Figs. 2e, 2f and 4; Table 5). This observation indicated that the ligand-induced conformational change is universal in different strains of influenza virus.

In the polymerase complex of *influenza A virus*, PB2cap is located near the fingers region, whereas the PA endonuclease is located near the thumb region (Fig. 5; Reich *et al.*, 2014, Pflug *et al.*, 2014). We postulate that when the cap binds in the ligand-binding site of PB2cap, the ligand-induced conformational change results in separation of the ligand-binding domain from the C-subdomain. This separation facilitates movement of the snatched pre-mRNA towards the PA endonuclease for cleavage, and then towards the palm region, which is the catalytic active site, to serve as the primer for transcription.

Acknowledgements

We thank Dr Weike Li, Dr Xin Li, Ross Terrell and Ziba Enadi for collaboration. We thank SER-CAT for facilitating data collection at Argonne National Laboratory. We thank Professor Herbert J. Bernstein for ray-tracing (Fig. 3). We also thank members of the *PHENIX* users mailing list (PHENIXbb) and the *UCSF Chimera* users mailing list for contributions.

References

- Adams, P. D. *et al.* (2010). *Acta Cryst.* **D66**, 213–221.
- Doan, L., Handa, B., Roberts, N. A. & Klumpp, K. (1999). *Biochemistry*, **38**, 5612–5619.
- Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* **D60**, 2126–2132.
- Hayward, S. & Berendsen, H. J. C. (1998). *Proteins*, **30**, 144–154.
- Liu, Y., Qin, K., Meng, G., Zhang, J., Zhou, J., Zhao, G., Luo, M. & Zheng, X. (2013). *J. Biol. Chem.* **288**, 11013–11023.
- Pautus, S., Sehr, P., Lewis, J., Fortuné, A., Wolkerstorfer, A., Szolar, O., Guilligay, D., Lunardi, T., Décout, J.-L. & Cusack, S. (2013). *J. Med. Chem.* **56**, 8915–8930.
- Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C. & Ferrin, T. E. (2004). *J. Comput. Chem.* **25**, 1605–1612.
- Pflug, A., Guilligay, D., Reich, S. & Cusack, S. (2014). *Nature (London)*, **516**, 355–360.
- Reich, S., Guilligay, D., Pflug, A., Malet, H., Berger, I., Crépin, T., Hart, D., Lunardi, T., Nanao, M., Ruigrok, R. W. & Cusack, S. (2014). *Nature (London)*, **516**, 361–366.
- Sayle, R. A. & Milner-White, E. J. (1995). *Trends Biochem. Sci.* **20**, 374.
- Severin, C., Rocha de Moura, T., Liu, Y., Li, K., Zheng, X. & Luo, M. (2016). *Acta Cryst.* **D72**, 245–253.
- Tarendeau, F., Boudet, J., Guilligay, D., Mas, P. J., Bougault, C. M., Boulo, S., Baudin, F., Ruigrok, R. W. H., Daigle, N., Ellenberg, J., Cusack, S., Simorre, J.-P. & Hart, D. J. (2007). *Nature Struct. Mol. Biol.* **14**, 229–233.
- Winn, M. D. *et al.* (2011). *Acta Cryst.* **D67**, 235–242.