WORLDWIDE wwPDB

PROTEIN DATA BANK

wwPDB EM Validation Summary Report (i)

May 13, 2024 – 08:50 pm BST

PDB ID	:	6XZD
EMDB ID	:	EMD-10659
Title	:	Influenza C virus polymerase complex without chicken ANP32A - Subclass 2
Authors	:	Keown, J.R.; Carrique, L.; Fan, H.; Grimes, J.M.; Fodor, E.
Deposited on	:	2020-02-04
Resolution	:	3.40 Å(reported)

This is a wwPDB EM Validation Summary Report for a publicly released PDB entry.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/EMValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references (i)) were used in the production of this report:
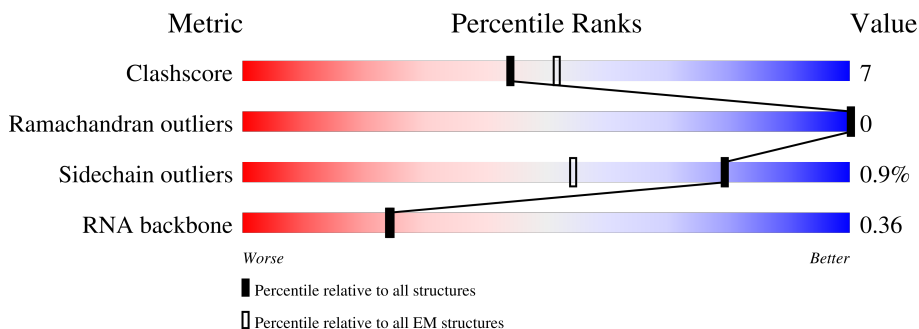
EMDB validation analysis	:	0.0.1.dev92
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ	:	FAILED
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36.2

1   Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:
ELECTRON MICROSCOPY

The reported resolution of this entry is 3.40 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.

Metric	Percentile Ranks	Value
Clashscore		7
Ramachandran outliers		0
Sidechain outliers		0.9%
RNA backbone		0.36

Worse   *Better*

▮ Percentile relative to all structures
▯ Percentile relative to all EM structures

Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	158937	4297
Ramachandran outliers	154571	4023
Sidechain outliers	154315	3826
RNA backbone	4643	859

The table below summarises the geometric issues observed across the polymeric chains and their fit to the map. The red, orange, yellow and green segments of the bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5%

Mol	Chain	Length	Quality of chain
1	IN1	47	17% 17% 6% • 57%
2	AP1	709	87% 11% •
2	DP1	709	62% 12% 26%
3	BP1	754	82% 12% 6%
3	EP1	754	63% 18% • 18%
4	CP1	774	84% 16%
4	FP1	774	60% 17% 23%

W O R L D W I D E
wwPDB
PROTEIN DATA BANK

2　Entry composition ⓘ

There are 4 unique types of molecules in this entry. The entry contains 63749 atoms, of which 31961 are hydrogens and 0 are deuteriums.

In the tables below, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a RNA chain called RNA (5'-R(*AP*GP*UP*AP*GP*AP*AP*AP*CP*AP* AP*GP*GP*GP*CP*CP*CP*UP*GP*C)-3').

Mol	Chain	Residues	Atoms						AltConf	Trace
1	IN1	20	Total	C	H	N	O	P	0	0
			650	193	221	84	133	19		

- Molecule 2 is a protein called Polymerase acidic protein.

Mol	Chain	Residues	Atoms						AltConf	Trace
2	AP1	696	Total	C	H	N	O	S	0	0
			11291	3599	5645	955	1049	43		
2	DP1	525	Total	C	H	N	O	S	0	0
			8558	2710	4306	729	779	34		

- Molecule 3 is a protein called RNA-directed RNA polymerase catalytic subunit.

Mol	Chain	Residues	Atoms						AltConf	Trace
3	BP1	712	Total	C	H	N	O	S	0	0
			11444	3602	5761	960	1068	53		
3	EP1	615	Total	C	H	N	O	S	0	0
			9855	3116	4961	822	911	45		

- Molecule 4 is a protein called Polymerase basic protein 2.

Mol	Chain	Residues	Atoms						AltConf	Trace
4	CP1	772	Total	C	H	N	O	S	0	0
			12409	3888	6261	1080	1142	38		
4	FP1	599	Total	C	H	N	O	S	0	0
			9542	3004	4806	820	884	28		

3 Residue-property plots ⓘ

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

• Molecule 1: RNA (5'-R(*AP*GP*UP*AP*GP*AP*AP*AP*CP*AP*AP*GP*GP*GP*CP*CP *CP*UP*GP*C)-3')

Chain IN1:

• Molecule 2: Polymerase acidic protein

Chain AP1:

• Molecule 2: Polymerase acidic protein

Chain DP1:

E595 W604 K608 L612 I613 L621 F629 C632 G641 E645 Q646 L653 K657 K660 S661 S662 C677 L678 I679 N680 L694 K707 THR ASP

• Molecule 3: RNA-directed RNA polymerase catalytic subunit

Chain BP1: 82% 12% 6%

M1 N12 T25 M30 SER HIS GLY SER S35 T36 Y47 S50 R51 R66 K67 F68 C69 N70 C71 D76 G83 E103 F106 T115 T131 T156 E137 G162 T163 T164 M165 L166 E167 I173 W178 K179 E180 I181 V186 LYS THR VAL

VAL ARG ARG GLU LYS ASP LYS ARG SER GLY LYS GLU ILE LYS THR LYS VAL PRO VAL MET GLY I211 D212 S213 R222 T225 I226 K237 R241 I249 V250 R251 L273 E279 N303 G306 D307 N308 Q316 I327 V341 V344 G354 R358 L359

E367 V368 I369 I370 M410 M415 M428 D429 E430 E431 Q443 S444 S445 D446 D447 F448 R466 N477 E481 S486 L487 P488 E489 L490 F497 F498 F502 V503 N519 D523 R561 I584 E585 M599 L612 V625 K637 THR ILE ASP ILE

PHE ARG ALA HIS GLY PRO ILE ARG VAL GLU E652 N653 T659 L671 M678 M691 F692 I710 K725 G729 A730 I731 D741 D745 R754

• Molecule 3: RNA-directed RNA polymerase catalytic subunit

Chain EP1: 63% 18% • 18%

M1 E2 I3 Y6 L7 M8 F9 L10 N11 N12 D13 S16 L17 T21 Y22 P23 M30 S31 H32 K37 E41 Y47 S50 R51 S58 K59 R65 F68 C69 N70 C71 L72 D76 E77 L78 G83 N84 V85 D86 I87 S88 H108 E112 D133

W134 D151 A152 T156 E157 T163 T164 M165 L166 N170 D177 E180 I181 K184 LYS VAL LYS THR VAL VAL ARG ARG GLU LYS ASP LYS ARG SER GLY LYS GLU ILE LYS THR LYS VAL PRO VAL MET GLY I211 D212 S213 T225 I226 A230 K231 D232 R235

Q239 R240 T245 I249 V250 F253 L268 K269 E270 N278 K281 Q299 V302 N303 N308 W311 Q316 D330 D337 L338 V341 N348 Q355 G356 I357 R358 K362 E367 V368 I369 I370 K374 M382 Y386 P405 L409 M410

V422 S423 T424 D429 L432 S444 S445 W462 R466 F467 N468 C471 K472 I476 N477 E481 S486 L487 P488 E489 L490 T494 S495 M496 F497 F498 D499 G500 V503 S504 N505 L506 E509 V518 N519 V522 N534 N538 S539 L540 M574

N578 E579 K582 E585 N586 S603 H606 E609 N623 R624 V625 F626 F632 THR ASN PHE ASP LYS THR ILE ASP ILE PHE ARG ALA HIS GLY PRO ILE ARG VAL GLU GLU ASN GLU A655 S658 T659 F662 R663 THR ARG ALA ASN ARG THR LEU LEU ASN THR ASP

MET ARG ALA MET ALA GLU LYS ARG GLN TYR MET VAL CYS ASP MET PHE LYS SER VAL PHE GLU SER ALA ASP ILE ASN PRO ILE GLY ALA HIS MET SER ILE GLY GLU VAL GLU ILE GLU GLU LYS LEU LEU GLU ARG ALA LYS MET LYS ARG ASP ILE GLY ALA ILE GLU ASP SER

GLU TYR GLU GLU ILE LYS ASP ILE ILE ARG ASP ALA LYS ALA ARG LEU GLU SER ARG

• Molecule 4: Polymerase basic protein 2

Chain CP1: 84% 16%

M1 S2 L3 L4 L14 Q22 K43 E44 K45 N46 R50 M51 R52 S56 E76 D85 T86 V95 V117 E139 L140 R141 F142 R151 Q155 Q159 D168 L169 D177 E178 A179 L205 N219 F224 F228 E234 G243 G244 E245

T246 W247 I248 S257 N258 N264 D265 K285 T302 V316 L327 T334 L335 L341 L342 K343 G349 V352 K353 D354 T355 I358 Q359 E367 S370 E371 Q372 S381 A384 R388 I398 D405 R406 L409 D421 K424 D425 A426 P427

I430 Q446 Y447 N451 W452 I453 L460 M481 G482 I483 I490 I498 Q499 P506 E507 L514 E515 L516 A522 F523 D524 V530 L536 F544 V547 D559 R574 L575 I582 P583 D590 Y591 L592 S593 E599 S616 R617 R620

Q626 L637 F646 R656 D662 S663 F664 R692 T704 S705 E710 R726 K751 R757 A758 D762 V763 N764 V767 K768 K769 M772 SER ASN

- Molecule 4: Polymerase basic protein 2

Chain FP1:

60%   17%   23%

MET SER LEU LEU LEU THR ILE ALA LYS GLU TYR LYS ARG LEU CYS GLN ASP ALA LYS ALA ALA GLN MET MET THR VAL GLY THR VAL SER ASN TYR THR THR PHE LYS LYS TRP THR THR SER ARG LYS GLU LYS ASN PRO SER LEU ARG MET ARG TRP ALA MET SER SER K58 I61

K65 R66 M67 L82 W83 GLU ASP THR GLU ASP VAL SER LYS ARG ASP HIS V95 L96 A97 S98 A99 S100 C101 F107 V112 N113 N114 I118 Y122 R132 R141 D146 ARG GLN ARG ARG ARG VAL ASP THR GLN PRO VAL GLU GLN ARG LEU ARG THR GLY GLU

ILE LYS ASP LEU GLN MET TRP THR LEU PHE GLU ASP ALA GLU PRO LEU LYS ALA SER LYS PHE ILE LEU ASP ASN TYR GLY LEU VAL LYS GLU MET ARG SER LYS PHE ALA ASN LYS PRO LEU ASN LYS GLU VAL VAL ALA HIS MET LEU GLU GLU LYS GLN PHE ASN PRO GLU ASN SER ARG PHE LEU

PRO VAL PHE GLY ALA ILE ARG P233 E234 R235 L238 E250 I256 V259 D260 Q261 L275 S280 I281 M282 I292 T295 R298 E301 K305 L306 E307 E308 L309 T313 V316 S317 T334 L335 L341 L342 K343 G347 S348 V361 R362 E367

E371 Q372 E373 R388 I402 R406 K407 R408 L409 R418 D419 G420 T429 I430 R441 P444 Y445 Q446 Y447 V448 W452 I453 Q454 K455 S456 E457 D458 N459 T472 K476 S486 K487 R488 L493 V496 T497 P506 E510 L516 S517 D518 R528

I529 Q545 F549 Q550 H551 P552 D553 L554 A555 V556 L557 K561 T565 R574 L575 I582 L585 D590 S593 L594 E599 Y615 R620 S633 E636 L637 I638 K639 V640 L641 F646 N671 N672 V673 C686 I691 R692 V693 P694 L695 E700

L706 E710 V714 C715 Q727 V732 Q736 L740 F741 V742 R752 T753 ALA SER GLN ARG ALA ALA SER SER ASP VAL ASN LYS ASN VAL LYS LYS ILE LYS MET SER ASN

4 Experimental information ⓘ

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	169000	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose $(e^-/\text{Å}^2)$	38.8	Depositor
Minimum defocus (nm)	2000	Depositor
Maximum defocus (nm)	3500	Depositor
Magnification	105000	Depositor
Image detector	GATAN K2 SUMMIT (4k x 4k)	Depositor

5   Model quality   ⓘ

5.1   Standard geometry   ⓘ

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	$\#|Z| > 5$	RMSZ	$\#|Z| > 5$
1	IN1	0.95	0/480	1.12	3/745 (0.4%)
2	AP1	0.68	0/5764	0.59	0/7746
2	DP1	0.65	0/4342	0.56	0/5835
3	BP1	0.69	0/5780	0.63	0/7762
3	EP1	0.43	0/4986	0.51	0/6707
4	CP1	0.65	0/6259	0.60	0/8425
4	FP1	0.31	0/4821	0.46	0/6493
All	All	0.60	0/32432	0.58	3/43713 (0.0%)

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	IN1	39	G	O4'-C1'-N9	6.47	113.38	108.20
1	IN1	5	G	O4'-C1'-N9	5.29	112.43	108.20
1	IN1	39	G	C4-N9-C1'	5.04	133.05	126.50

There are no chirality outliers.

There are no planarity outliers.

5.2   Too-close contacts   ⓘ

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	IN1	429	221	221	7	0
2	AP1	5646	5645	5645	55	0
2	DP1	4252	4306	4305	71	0

Continued on next page...

W O R L D W I D E
wwPDB
PROTEIN DATA BANK

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	BP1	5683	5761	5759	63	0
3	EP1	4894	4961	4959	106	0
4	CP1	6148	6261	6261	83	0
4	FP1	4736	4806	4806	94	0
All	All	31788	31961	31956	439	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 7.

The worst 5 of 439 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:BP1:303:ASN:ND2	3:BP1:488:PRO:O	1.89	1.05
4:FP1:575:LEU:HD12	4:FP1:582:ILE:HD13	1.45	0.97
2:DP1:449:CYS:SG	2:DP1:490:LYS:NZ	2.45	0.89
3:BP1:237:LYS:NZ	3:BP1:307:ASP:OD1	2.06	0.88
4:CP1:44:GLU:OE1	4:CP1:50:ARG:NH2	2.06	0.88

There are no symmetry-related clashes.

5.3 Torsion angles (i)

5.3.1 Protein backbone (i)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all PDB entries followed by that with respect to all EM entries.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	AP1	692/709 (98%)	661 (96%)	31 (4%)	0	100	100
2	DP1	523/709 (74%)	505 (97%)	18 (3%)	0	100	100
3	BP1	704/754 (93%)	661 (94%)	43 (6%)	0	100	100
3	EP1	609/754 (81%)	566 (93%)	43 (7%)	0	100	100
4	CP1	770/774 (100%)	729 (95%)	41 (5%)	0	100	100
4	FP1	593/774 (77%)	552 (93%)	41 (7%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
All	All	3891/4474 (87%)	3674 (94%)	217 (6%)	0	100	100

There are no Ramachandran outliers to report.

5.3.2　Protein sidechains　ⓘ

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all PDB entries followed by that with respect to all EM entries.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	AP1	620/631 (98%)	617 (100%)	3 (0%)	88	94
2	DP1	468/631 (74%)	463 (99%)	5 (1%)	73	86
3	BP1	632/669 (94%)	626 (99%)	6 (1%)	78	90
3	EP1	547/669 (82%)	539 (98%)	8 (2%)	65	82
4	CP1	677/679 (100%)	671 (99%)	6 (1%)	78	90
4	FP1	521/679 (77%)	517 (99%)	4 (1%)	81	91
All	All	3465/3958 (88%)	3433 (99%)	32 (1%)	79	90

5 of 32 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
4	FP1	118	ILE
4	FP1	261	GLN
4	CP1	371	GLU
4	CP1	52	ARG
4	FP1	388	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

5.3.3　RNA　ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	IN1	18/47 (38%)	8 (44%)	1 (5%)

5 of 8 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	IN1	6	A
1	IN1	8	A
1	IN1	11	A
1	IN1	14	G
1	IN1	36	C

All (1) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
1	IN1	38	U

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates ⓘ

There are no monosaccharides in this entry.

5.6 Ligand geometry ⓘ

There are no ligands in this entry.

5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

There are no chain breaks in this entry.

6   Map visualisation ⓘ

This section contains visualisations of the EMDB entry EMD-10659. These allow visual inspection of the internal detail of the map and identification of artifacts.

Images derived from a raw map, generated by summing the deposited half-maps, are presented below the corresponding image components of the primary map to allow further visual inspection and comparison with those of the primary map.

6.1   Orthogonal projections ⓘ

This section was not generated.

6.2   Central slices ⓘ

This section was not generated.

6.3   Largest variance slices ⓘ

This section was not generated.

6.4   Orthogonal standard-deviation projections (False-color) ⓘ

This section was not generated.

6.5   Orthogonal surface views ⓘ

This section was not generated.

6.6   Mask visualisation ⓘ

This section was not generated. No masks/segmentation were deposited.

7 Map analysis ⓘ

This section contains the results of statistical analysis of the map.

7.1 Map-value distribution ⓘ

This section was not generated.

7.2 Volume estimate versus contour level ⓘ

This section was not generated.

7.3 Rotationally averaged power spectrum ⓘ

This section was not generated. The rotationally averaged power spectrum had issues being displayed.

8 Fourier-Shell correlation (i)

This section was not generated. No FSC curve or half-maps provided.

9 Map-model fit (i)

This section was not generated.