wwPDB X-ray Structure Validation Summary Report ⓘ

Jun 24, 2024 – 04:04 PM EDT

PDB ID	:	8TWP
Title	:	Influenza A virus (A/Aichi/2/1968(H3N2) nucleoprotein mutant - 2-7 deleted, R416A
Authors	:	Yoon, J.; Zhang, Y.M.; Grant, R.A.; Shoulders, M.D.
Deposited on	:	2023-08-21
Resolution	:	2.90 Å(reported)

The following versions of software and data (see references ⓘ) were used in the production of this report:

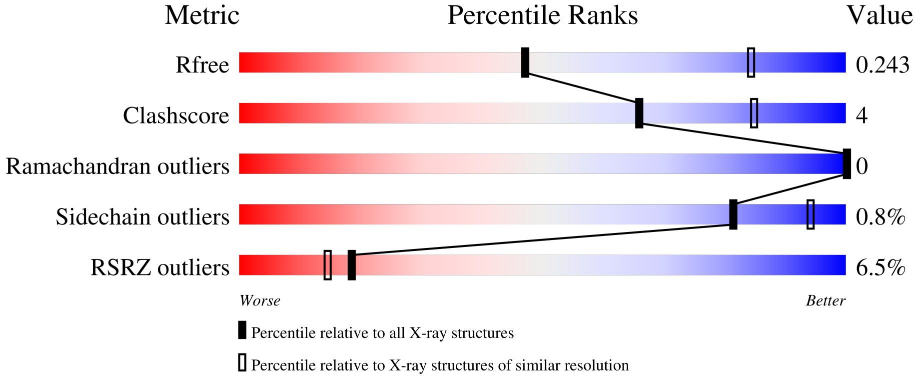
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.37.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

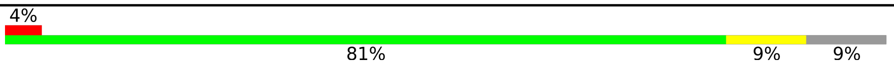
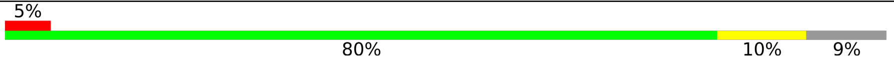
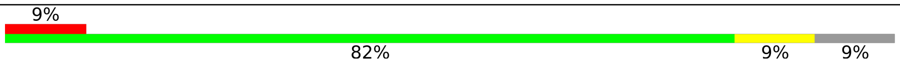
The reported resolution of this entry is 2.90 Å.

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	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1957 (2.90-2.90)
Clashscore	141614	2172 (2.90-2.90)
Ramachandran outliers	138981	2115 (2.90-2.90)
Sidechain outliers	138945	2117 (2.90-2.90)
RSRZ outliers	127900	1906 (2.90-2.90)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower 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% The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	500	
1	B	500	
1	C	500	

2 Entry composition ⓘ

There is only 1 type of molecule in this entry. The entry contains 10734 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, 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 protein called Nucleoprotein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	453	3578	2221	659	669	29	0	0	0
1	B	453	3578	2221	659	669	29	0	0	0
1	C	453	3578	2221	659	669	29	0	0	0

There are 30 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	7	MET	-	initiating methionine	UNP I6TAH8
A	416	ALA	ARG	engineered mutation	UNP I6TAH8
A	499	LEU	-	expression tag	UNP I6TAH8
A	500	GLU	-	expression tag	UNP I6TAH8
A	501	HIS	-	expression tag	UNP I6TAH8
A	502	HIS	-	expression tag	UNP I6TAH8
A	503	HIS	-	expression tag	UNP I6TAH8
A	504	HIS	-	expression tag	UNP I6TAH8
A	505	HIS	-	expression tag	UNP I6TAH8
A	506	HIS	-	expression tag	UNP I6TAH8
B	7	MET	-	initiating methionine	UNP I6TAH8
B	416	ALA	ARG	engineered mutation	UNP I6TAH8
B	499	LEU	-	expression tag	UNP I6TAH8
B	500	GLU	-	expression tag	UNP I6TAH8
B	501	HIS	-	expression tag	UNP I6TAH8
B	502	HIS	-	expression tag	UNP I6TAH8
B	503	HIS	-	expression tag	UNP I6TAH8
B	504	HIS	-	expression tag	UNP I6TAH8
B	505	HIS	-	expression tag	UNP I6TAH8
B	506	HIS	-	expression tag	UNP I6TAH8
C	7	MET	-	initiating methionine	UNP I6TAH8
C	416	ALA	ARG	engineered mutation	UNP I6TAH8
C	499	LEU	-	expression tag	UNP I6TAH8

Continued on next page...

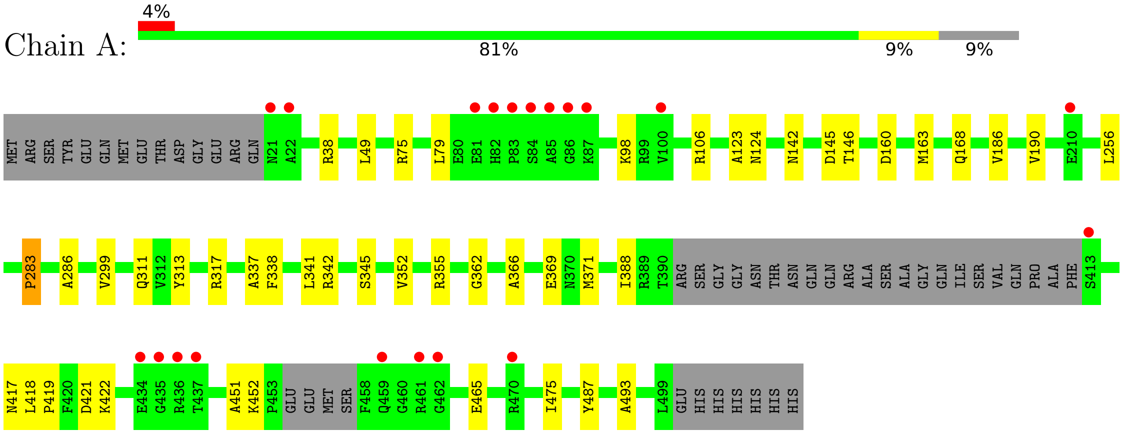
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	500	GLU	-	expression tag	UNP I6TAH8
C	501	HIS	-	expression tag	UNP I6TAH8
C	502	HIS	-	expression tag	UNP I6TAH8
C	503	HIS	-	expression tag	UNP I6TAH8
C	504	HIS	-	expression tag	UNP I6TAH8
C	505	HIS	-	expression tag	UNP I6TAH8
C	506	HIS	-	expression tag	UNP I6TAH8

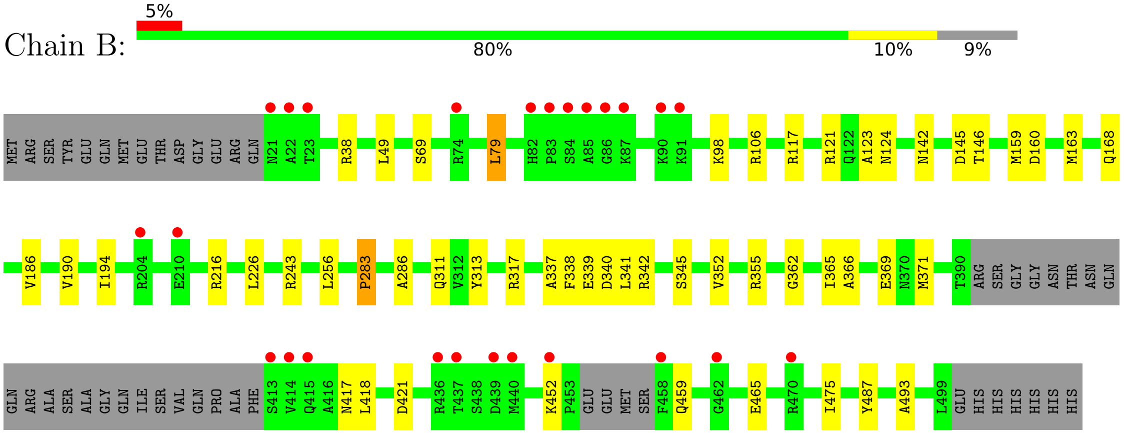
3 Residue-property plots ⓘ

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 and electron density. 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. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). 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.

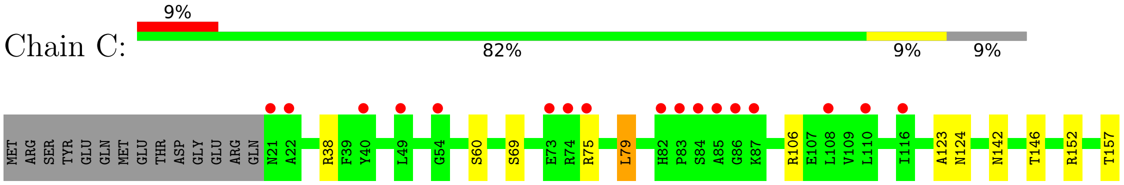
- Molecule 1: Nucleoprotein

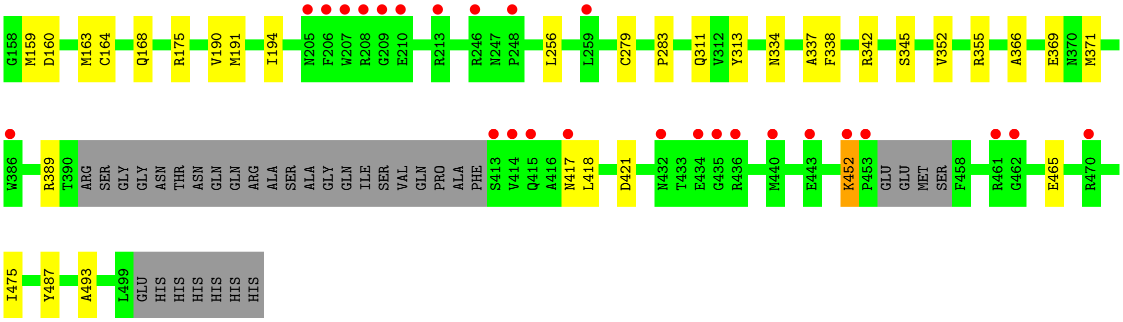


- Molecule 1: Nucleoprotein



- Molecule 1: Nucleoprotein




G158 M159 D160 M163 C164 Q168 R175 V190 M191 I194 N205 F206 W207 R208 G209 E210 R213 R246 N247 P248 L256 L259 C279 P283 Q311 V312 Y313 N334 A337 F338 R342 S345 V352 R355 A366 E369 N370 M371
W386 R389 T390 ARG SER GLY GLY ASN THR ASN GLN GLN ARG ALA SER ALA GLY GLN ILE SER VAL GLN PRO ALA PHE S413 V414 Q415 A416 N417 L418 D421 N432 T433 E434 G435 R436 M440 E443 K452 P453 GLU GLU MET SER F458 R461 G462 E465 R470
I475 Y487 A493 L499 GLU HIS HIS HIS HIS HIS HIS

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α, β, γ	163.84Å 282.92Å 116.14Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.68 – 2.90 48.68 – 2.90	Depositor EDS
% Data completeness (in resolution range)	98.1 (48.68-2.90) 98.3 (48.68-2.90)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.89 (at 2.91Å)	Xtriage
Refinement program	PHENIX 1.17.1	Depositor
R, R_{free}	0.220 , 0.244 0.220 , 0.243	Depositor DCC
R_{free} test set	2000 reflections (3.39%)	wwPDB-VP
Wilson B-factor ($Å^2$)	82.5	Xtriage
Anisotropy	0.673	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 69.3	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.000 for 1/2*h-1/2*k,-3/2*h-1/2*k,-l 0.000 for 1/2*h+1/2*k,3/2*h-1/2*k,-l	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	10734	wwPDB-VP
Average B, all atoms ($Å^2$)	102.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.03% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

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	A	0.24	0/3639	0.42	1/4890 (0.0%)
1	B	0.24	0/3639	0.42	1/4890 (0.0%)
1	C	0.24	0/3639	0.41	1/4890 (0.0%)
All	All	0.24	0/10917	0.42	3/14670 (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	A	417	ASN	C-N-CA	8.23	142.27	121.70
1	B	417	ASN	C-N-CA	8.00	141.69	121.70
1	C	417	ASN	C-N-CA	7.43	140.27	121.70

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	A	3578	0	3555	30	0
1	B	3578	0	3555	32	0
1	C	3578	0	3555	27	0
All	All	10734	0	10665	85	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 4.

The worst 5 of 85 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)
1:A:146:THR:HG22	1:A:337:ALA:HB2	1.52	0.87
1:C:146:THR:HG22	1:C:337:ALA:HB2	1.55	0.87
1:B:146:THR:HG22	1:B:337:ALA:HB2	1.56	0.86
1:A:106:ARG:HE	1:A:371:MET:HE3	1.45	0.82
1:C:106:ARG:HE	1:C:371:MET:HE3	1.49	0.78

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

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	
1	A	447/500 (89%)	440 (98%)	7 (2%)	0	100	100
1	B	447/500 (89%)	440 (98%)	7 (2%)	0	100	100
1	C	447/500 (89%)	440 (98%)	7 (2%)	0	100	100
All	All	1341/1500 (89%)	1320 (98%)	21 (2%)	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 X-ray entries followed by that with respect to entries of similar resolution.

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	
1	A	380/420 (90%)	377 (99%)	3 (1%)	81	94

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	B	380/420 (90%)	377 (99%)	3 (1%)	81	94
1	C	380/420 (90%)	377 (99%)	3 (1%)	81	94
All	All	1140/1260 (90%)	1131 (99%)	9 (1%)	81	94

5 of 9 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	283	PRO
1	C	452	LYS
1	B	79	LEU
1	B	283	PRO
1	B	452	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	311	GLN
1	B	311	GLN
1	C	311	GLN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

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 Fit of model and data

6.1 Protein, DNA and RNA chains

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95^{th} percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å^2)	Q<0.9
1	A	453/500 (90%)	0.37	20 (4%)	34	30	64, 89, 150, 200	0
1	B	453/500 (90%)	0.37	25 (5%)	25	21	63, 88, 143, 186	0
1	C	453/500 (90%)	0.65	43 (9%)	8	6	80, 113, 163, 201	0
All	All	1359/1500 (90%)	0.47	88 (6%)	18	14	63, 96, 153, 201	0

The worst 5 of 88 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	435	GLY	9.9
1	C	414	VAL	8.3
1	A	82	HIS	8.0
1	B	85	ALA	7.7
1	C	415	GLN	7.2

6.2 Non-standard residues in protein, DNA, RNA chains

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

6.3 Carbohydrates

There are no monosaccharides in this entry.

6.4 Ligands

There are no ligands in this entry.

6.5 Other polymers ⓘ

There are no such residues in this entry.