



wwPDB X-ray Structure Validation Summary Report ⓘ

Apr 4, 2016 – 02:02 PM EDT

PDB ID : 5F45
Title : Crystal structure of Fab H7.167 in complex with influenza virus hemagglutinin from A/Shanghai/02/2013 (H7N9)
Authors : Zhang, H.; Wilson, I.A.
Deposited on : 2015-12-03
Resolution : 4.66 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.
We welcome your comments at validation@mail.wwpdb.org
A user guide is available at
<http://wwpdb.org/validation/2016/XrayValidationReportHelp>
with specific help available everywhere you see the ⓘ symbol.

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

MolProbity : 4.02b-467
Mogul : unknown
Xtriage (Phenix) : 1.9-1692
EDS : rb-20027107
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac : 5.8.0135
CCP4 : 6.5.0
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : rb-20027107

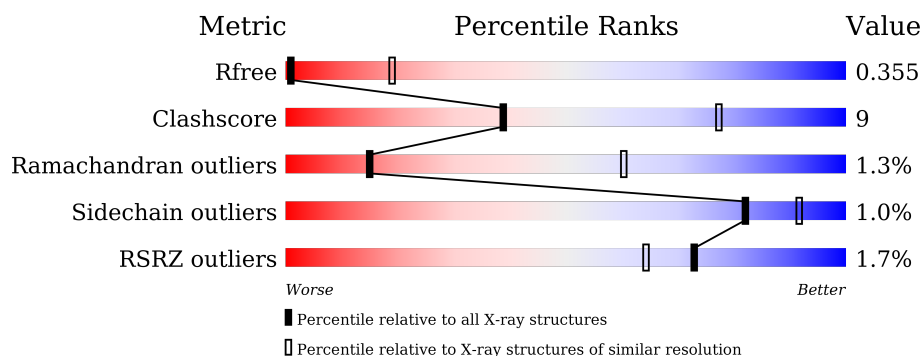
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 4.66 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}	91344	1093 (5.70-3.60)
Clashscore	102246	1007 (5.66-3.64)
Ramachandran outliers	100387	1135 (5.70-3.60)
Sidechain outliers	100360	1116 (5.70-3.60)
RSRZ outliers	91569	1096 (5.70-3.60)

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 on the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria. 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 $\leq 5\%$. 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	H	119	<div> <div>3%</div> <div> <div></div> <div>69%</div> <div>29%</div> <div>.</div> </div> </div>
2	A	318	<div> <div>%</div> <div> <div></div> <div>74%</div> <div>25%</div> <div>.</div> </div> </div>
3	B	183	<div> <div>3%</div> <div> <div></div> <div>77%</div> <div>16%</div> <div>8%</div> <div></div> </div> </div>
4	L	113	<div> <div></div> <div> <div></div> <div>78%</div> <div>21%</div> <div>.</div> </div> </div>

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 5599 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 Heavy chain of H7.167 Fab.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	H	119	Total	C	N	O	S	0	0	0
			931	594	156	177	4			

- Molecule 2 is a protein called hemagglutinin HA1 chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	A	318	Total	C	N	O	S	0	0	0
			2427	1509	438	465	15			

- Molecule 3 is a protein called hemagglutinin HA2 chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	B	169	Total	C	N	O	S	0	0	0
			1373	845	241	280	7			

There are 7 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	177	SER	-	expression tag	UNP A0A023HIQ5
B	178	GLY	-	expression tag	UNP A0A023HIQ5
B	179	ARG	-	expression tag	UNP A0A023HIQ5
B	180	LEU	-	expression tag	UNP A0A023HIQ5
B	181	VAL	-	expression tag	UNP A0A023HIQ5
B	182	PRO	-	expression tag	UNP A0A023HIQ5
B	183	ARG	-	expression tag	UNP A0A023HIQ5

- Molecule 4 is a protein called Light chain (kappa) of H7.167 Fab.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	L	113	Total	C	N	O	S	0	0	0
			868	546	144	175	3			

D1	R18	I21	M28	Y32	W35	Y36	Q37	Q42	P43	L47	I48	Y49	W50	S76	A80	Y83	A84	Y85	C88	Y91	T94	P95	G95A	L96	T97	F98	G101	T102	K103	V106
----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	------	-----	-----	-----	------	------	------	------

4 Data and refinement statistics

Property	Value	Source
Space group	I 21 3	Depositor
Cell constants a, b, c, α , β , γ	207.28 Å 207.28 Å 207.28 Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	46.35 – 4.66 48.86 – 4.66	Depositor EDS
% Data completeness (in resolution range)	100.0 (46.35-4.66) 100.0 (48.86-4.66)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	0.02	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.37 (at 4.64 Å)	Xtriage
Refinement program	PHENIX (1.10.1_2155: ???)	Depositor
R, R_{free}	0.326 , 0.354 0.324 , 0.355	Depositor DCC
R_{free} test set	806 reflections (9.97%)	DCC
Wilson B-factor (Å ²)	230.0	Xtriage
Anisotropy	0.000	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.21 , 211.1	EDS
Estimated twinning fraction	0.033 for -l,-k,-h	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Outliers	0 of 8096 reflections	Xtriage
F_o, F_c correlation	0.86	EDS
Total number of atoms	5599	wwPDB-VP
Average B, all atoms (Å ²)	301.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.06% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 0.2 for perfectly twinned datasets.

5 Model quality [i](#)

5.1 Standard geometry [i](#)

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	H	0.29	0/954	0.53	1/1292 (0.1%)
2	A	0.25	0/2474	0.49	0/3345
3	B	0.23	0/1396	0.44	0/1881
4	L	0.25	0/888	0.52	0/1208
All	All	0.25	0/5712	0.49	1/7726 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	H	0	1

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed($^{\circ}$)	Ideal($^{\circ}$)
1	H	94	LEU	CA-CB-CG	5.28	127.44	115.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	H	91	TYR	Peptide

5.2 Too-close contacts [i](#)

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	H	931	0	893	27	0
2	A	2427	0	2389	54	1
3	B	1373	0	1274	19	1
4	L	868	0	841	16	0
All	All	5599	0	5397	99	1

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 9.

The worst 5 of 99 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:H:87:THR:HG22	1:H:111:VAL:H	1.49	0.76
2:A:134:GLY:HA3	2:A:153:TRP:HB3	1.68	0.74
1:H:35:ASN:OD1	1:H:47:TRP:NE1	2.20	0.73
2:A:184:HIS:HB2	2:A:220:ARG:HH12	1.52	0.73
1:H:52:SER:O	1:H:71:ARG:NH1	2.27	0.68

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:A:28:THR:O	3:B:54:ARG:NH2[8_555]	2.06	0.14

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 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	H	117/119 (98%)	98 (84%)	16 (14%)	3 (3%)	7 47
2	A	316/318 (99%)	284 (90%)	29 (9%)	3 (1%)	21 67

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
3	B	167/183 (91%)	150 (90%)	14 (8%)	3 (2%)	11	54
4	L	111/113 (98%)	98 (88%)	13 (12%)	0	100	100
All	All	711/733 (97%)	630 (89%)	72 (10%)	9 (1%)	15	60

5 of 9 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	H	92	CYS
1	H	93	ARG
2	A	158(B)	ASN
2	A	143	SER
3	B	29	ALA

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	H	99/99 (100%)	98 (99%)	1 (1%)	82	91
2	A	267/267 (100%)	263 (98%)	4 (2%)	72	89
3	B	145/157 (92%)	145 (100%)	0	100	100
4	L	97/97 (100%)	96 (99%)	1 (1%)	82	91
All	All	608/620 (98%)	602 (99%)	6 (1%)	82	91

5 of 6 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	A	96	VAL
4	L	88	CYS
2	A	98	TYR
2	A	11	ASP
2	A	276(A)	ASN

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

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 carbohydrates in this entry.

5.6 Ligand geometry [i](#)

There are no ligands in this entry.

5.7 Other polymers [i](#)

There are no such residues in this entry.

5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

6 Fit of model and data [i](#)

6.1 Protein, DNA and RNA chains [i](#)

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, 95th 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	H	119/119 (100%)	-0.25	3 (2%) 61 51	247, 272, 307, 329	0
2	A	318/318 (100%)	-0.46	3 (0%) 85 80	246, 280, 359, 407	0
3	B	169/183 (92%)	-0.02	6 (3%) 46 38	244, 348, 419, 447	0
4	L	113/113 (100%)	-0.36	0 100 100	271, 298, 325, 334	0
All	All	719/733 (98%)	-0.31	12 (1%) 73 64	244, 288, 401, 447	0

The worst 5 of 12 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	H	100	LEU	4.1
3	B	4	GLY	3.2
2	A	219	ALA	3.2
3	B	172	GLN	3.2
3	B	171	ILE	2.9

6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

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

6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

6.4 Ligands [i](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.