wwPDB X-ray Structure Validation Summary Report (i)

May 20, 2024 – 04:08 PM JST

PDB ID	:	8JSB
Title	:	Antibody scFv against the Matrix protein 2 of influenza virus
Authors	:	Kumar, U.; Gaur, V.; Salunke, D.M.
Deposited on	:	2023-06-19
Resolution	:	3.03 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
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
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:

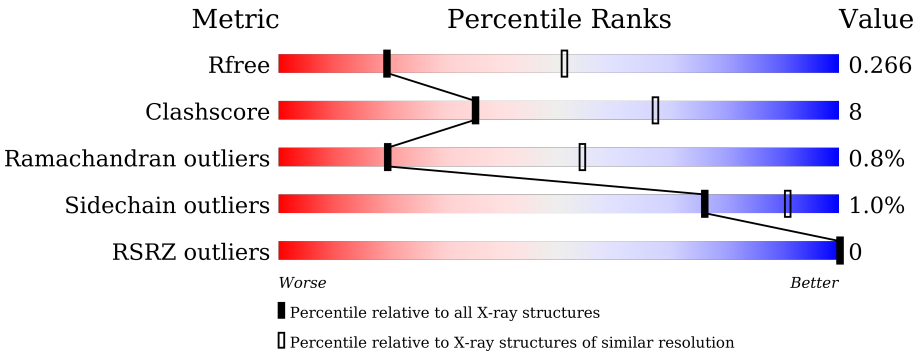
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36.2
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36.2

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 3.03 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	2752 (3.08-3.00)
Clashscore	141614	3096 (3.08-3.00)
Ramachandran outliers	138981	2986 (3.08-3.00)
Sidechain outliers	138945	2988 (3.08-3.00)
RSRZ outliers	127900	2636 (3.08-3.00)

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	253	72% 15% 13%
1	B	253	65% 23% 11%
1	C	253	74% 14% 12%
1	D	253	69% 18% 13%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 6657 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 single chain variable fragment.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	221	Total	C	N	O	S	0	0	0
			1637	1031	277	323	6			
1	B	224	Total	C	N	O	S	0	0	0
			1645	1033	276	329	7			
1	C	223	Total	C	N	O	S	0	0	0
			1667	1048	284	328	7			
1	D	220	Total	C	N	O	S	0	0	0
			1634	1028	278	321	7			

- Molecule 2 is water.

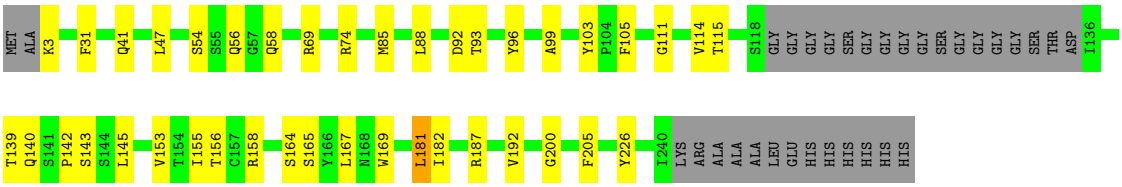
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	15	Total	O	0	0
			15	15		
2	B	18	Total	O	0	0
			18	18		
2	C	20	Total	O	0	0
			20	20		
2	D	21	Total	O	0	0
			21	21		

3 Residue-property plots (i)

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.

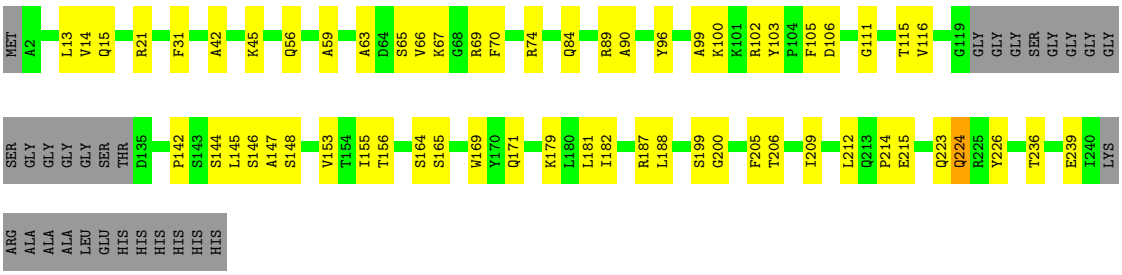
• Molecule 1: single chain variable fragment

Chain A:



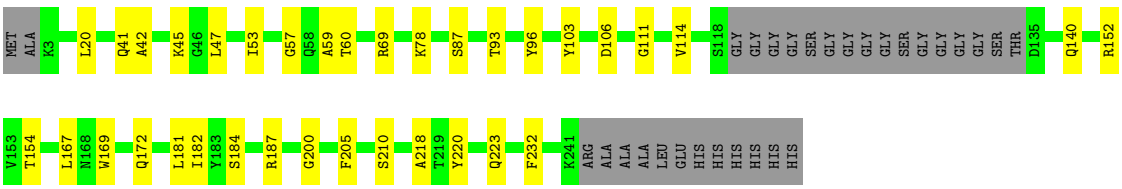
• Molecule 1: single chain variable fragment

Chain B:



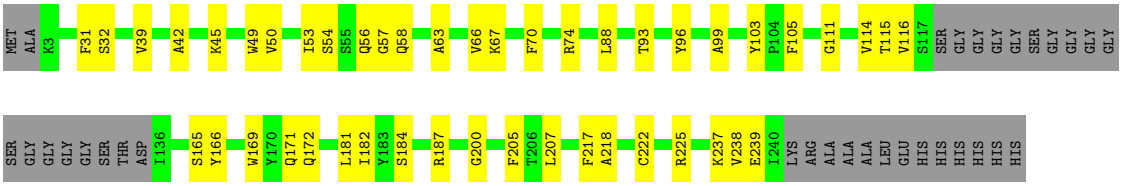
• Molecule 1: single chain variable fragment

Chain C:



• Molecule 1: single chain variable fragment

Chain D:


MET ALA K3 F31 S32 V39 A42 K45 W49 V50 I53 S54 S55 Q56 G57 Q58 A63 V66 K67 F70 R74 L88 T93 Y96 A99 Y103 P104 F105 G111 V114 T115 V116 S117 SER GLY GLY GLY GLY SER GLY GLY GLY GLY
SER GLY GLY GLY GLY SER THR ASP I136 S165 Y166 W169 Y170 Q171 Q172 L181 I182 Y183 S184 R187 G200 F205 T206 L207 F217 A218 C222 R225 K237 V238 E239 I240 LYS ARG ALA ALA ALA LEU GLU HIS HIS HIS HIS HIS HIS

4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	101.10Å 98.39Å 94.56Å 90.00° 94.33° 90.00°	Depositor
Resolution (Å)	50.41 – 3.03 94.29 – 3.03	Depositor EDS
% Data completeness (in resolution range)	99.7 (50.41-3.03) 99.8 (94.29-3.03)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.55 (at 3.01Å)	Xtriage
Refinement program	PHENIX 1.20.1_4487	Depositor
R, R_{free}	0.223 , 0.267 0.221 , 0.266	Depositor DCC
R_{free} test set	857 reflections (4.73%)	wwPDB-VP
Wilson B-factor (Å^2)	50.8	Xtriage
Anisotropy	0.233	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 45.7	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.89	EDS
Total number of atoms	6657	wwPDB-VP
Average B, all atoms (Å^2)	46.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.34% 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	0.28	0/1672	0.56	1/2270 (0.0%)
1	B	0.27	0/1679	0.54	0/2280
1	C	0.28	0/1702	0.53	0/2308
1	D	0.27	0/1668	0.54	0/2263
All	All	0.28	0/6721	0.54	1/9121 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	181	LEU	CA-CB-CG	5.91	128.90	115.30

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	1637	0	1538	26	0
1	B	1645	0	1549	34	0
1	C	1667	0	1590	24	0
1	D	1634	0	1551	25	0
2	A	15	0	0	1	0
2	B	18	0	0	0	0
2	C	20	0	0	1	0
2	D	21	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	6657	0	6228	105	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 8.

The worst 5 of 105 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:169:TRP:HB2	1:A:182:ILE:HB	1.76	0.67
1:C:42:ALA:HB3	1:C:45:LYS:HB2	1.79	0.64
1:B:148:SER:O	1:B:212:LEU:HD12	1.98	0.63
1:D:169:TRP:HB2	1:D:182:ILE:HB	1.79	0.63
1:A:167:LEU:HD12	1:A:205:PHE:CG	2.34	0.62

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 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	217/253 (86%)	202 (93%)	14 (6%)	1 (0%)	29	65
1	B	220/253 (87%)	200 (91%)	17 (8%)	3 (1%)	11	40
1	C	219/253 (87%)	207 (94%)	10 (5%)	2 (1%)	17	52
1	D	216/253 (85%)	203 (94%)	12 (6%)	1 (0%)	29	65
All	All	872/1012 (86%)	812 (93%)	53 (6%)	7 (1%)	19	54

5 of 7 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	210	SER
1	B	215	GLU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	103	TYR
1	C	103	TYR
1	B	103	TYR

5.3.2 Protein sidechains (i)

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	169/202 (84%)	168 (99%)	1 (1%)	86 94
1	B	171/202 (85%)	166 (97%)	5 (3%)	42 74
1	C	176/202 (87%)	176 (100%)	0	100 100
1	D	170/202 (84%)	169 (99%)	1 (1%)	86 94
All	All	686/808 (85%)	679 (99%)	7 (1%)	76 91

5 of 7 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	188	LEU
1	B	224	GLN
1	D	32	SER
1	B	226	TYR
1	B	102	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	171	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 (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

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	221/253 (87%)	-0.09	0	100	100	26, 44, 76, 87	0
1	B	224/253 (88%)	-0.02	0	100	100	29, 48, 76, 82	0
1	C	223/253 (88%)	-0.29	0	100	100	28, 43, 63, 70	0
1	D	220/253 (86%)	-0.31	0	100	100	28, 43, 61, 69	0
All	All	888/1012 (87%)	-0.18	0	100	100	26, 44, 71, 87	0

There are no RSRZ outliers to report.

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.