Full wwPDB X-ray Structure Validation Report ⓘ

Mar 26, 2024 – 12:05 pm GMT

PDB ID	:	8P6I
Title	:	Crystal structure of the 139H2 Fab fragment bound to Muc1 peptide epitope
Authors	:	Beugelink, J.W.; Peng, W.; Siborova, M.; Pronker, M.F.; Snijder, J.; Janssen, B.J.C.
Deposited on	:	2023-05-26
Resolution	:	2.50 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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 ⓘ 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 ⓘ) were used in the production of this report:

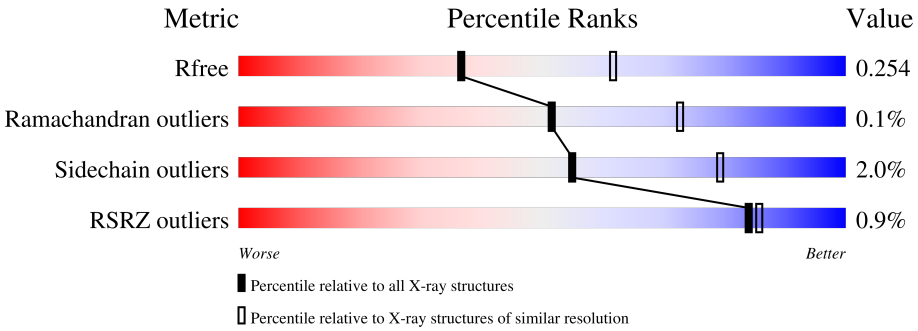
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 2.50 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	4661 (2.50-2.50)
Ramachandran outliers	138981	5231 (2.50-2.50)
Sidechain outliers	138945	5233 (2.50-2.50)
RSRZ outliers	127900	4559 (2.50-2.50)

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	9	100%
1	a	9	100%
2	H	236	% 88% · 11%
2	h	236	3% 86% · 12%
3	L	220	98% ·
3	l	220	96% ·

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 6948 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 Mucin-1.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
1	A	9	Total	C	N	O	0	0	0
			61	37	12	12			
1	a	9	Total	C	N	O	0	0	0
			61	37	12	12			

- Molecule 2 is a protein called 139H2 HC.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	H	210	Total	C	N	O	S	0	0	0
			1595	1010	260	317	8			
2	h	208	Total	C	N	O	S	0	0	0
			1582	1002	258	314	8			

- Molecule 3 is a protein called 139H2 LC.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	L	219	Total	C	N	O	S	0	0	0
			1710	1071	285	347	7			
3	l	219	Total	C	N	O	S	0	0	0
			1710	1071	285	347	7			

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	3	Total	O	0	0
			3	3		
4	H	61	Total	O	0	0
			61	61		
4	L	65	Total	O	0	0
			65	65		
4	a	3	Total	O	0	0
			3	3		

Continued on next page...

Continued from previous page...

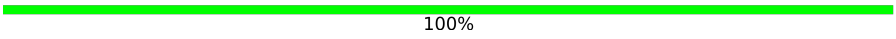
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	h	34	Total O 34 34	0	0
4	l	63	Total O 63 63	0	0

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: Mucin-1

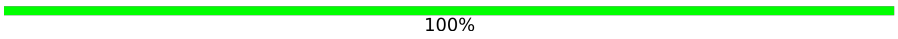
Chain A:



There are no outlier residues recorded for this chain.

- Molecule 1: Mucin-1

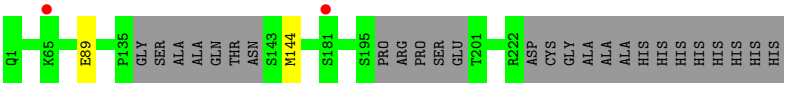
Chain a:



There are no outlier residues recorded for this chain.

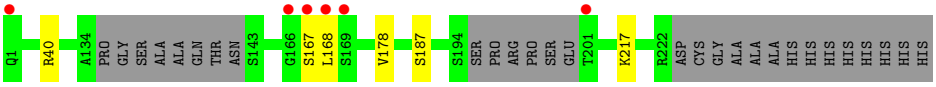
- Molecule 2: 139H2 HC

Chain H:



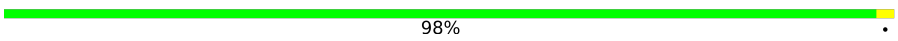
- Molecule 2: 139H2 HC

Chain h:



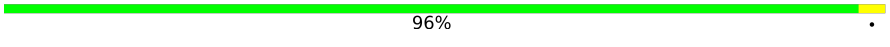
- Molecule 3: 139H2 LC

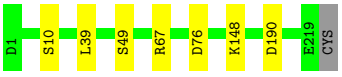
Chain L:



- Molecule 3: 139H2 LC

Chain l:





4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	211.83Å 42.76Å 129.06Å 90.00° 122.31° 90.00°	Depositor
Resolution (Å)	56.12 – 2.50 56.05 – 2.50	Depositor EDS
% Data completeness (in resolution range)	99.9 (56.12-2.50) 100.0 (56.05-2.50)	Depositor EDS
R_{merge}	0.39	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.47 (at 2.51Å)	Xtriage
Refinement program	REFMAC 5.8.0405	Depositor
R, R_{free}	0.203 , 0.254 0.207 , 0.254	Depositor DCC
R_{free} test set	1758 reflections (5.09%)	wwPDB-VP
Wilson B-factor (Å^2)	30.8	Xtriage
Anisotropy	0.259	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 37.3	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	6948	wwPDB-VP
Average B, all atoms (Å^2)	34.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.06% 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	0.39	0/63	0.80	0/87
1	a	0.43	0/63	0.97	0/87
2	H	0.41	0/1634	0.70	0/2230
2	h	0.40	0/1620	0.69	0/2210
3	L	0.43	0/1750	0.70	0/2376
3	l	0.44	0/1750	0.74	0/2376
All	All	0.42	0/6880	0.71	0/9366

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
2	h	0	2
3	l	0	1
All	All	0	3

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	h	168	LEU	Peptide
2	h	40	ARG	Sidechain
3	l	67	ARG	Sidechain

5.2 Too-close contacts

Due to software issues we are unable to calculate clashes - this section is therefore empty.

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	7/9 (78%)	7 (100%)	0	0	100 100
1	a	7/9 (78%)	7 (100%)	0	0	100 100
2	H	204/236 (86%)	198 (97%)	6 (3%)	0	100 100
2	h	202/236 (86%)	196 (97%)	5 (2%)	1 (0%)	29 48
3	L	217/220 (99%)	211 (97%)	6 (3%)	0	100 100
3	l	217/220 (99%)	211 (97%)	6 (3%)	0	100 100
All	All	854/930 (92%)	830 (97%)	23 (3%)	1 (0%)	51 73

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	h	167	SER

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	6/6 (100%)	6 (100%)	0	100 100
1	a	6/6 (100%)	6 (100%)	0	100 100
2	H	183/202 (91%)	181 (99%)	2 (1%)	73 89
2	h	181/202 (90%)	178 (98%)	3 (2%)	60 82
3	L	196/197 (100%)	192 (98%)	4 (2%)	55 79
3	l	196/197 (100%)	190 (97%)	6 (3%)	40 67

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	768/810 (95%)	753 (98%)	15 (2%)	55 79

All (15) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	H	89	GLU
2	H	144	MET
3	L	39	LEU
3	L	49	SER
3	L	186	THR
3	L	219	GLU
2	h	178	VAL
2	h	187	SER
2	h	217	LYS
3	l	10	SER
3	l	39	LEU
3	l	49	SER
3	l	76	ASP
3	l	148	LYS
3	l	190	ASP

Sometimes 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 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	9/9 (100%)	-0.17	0 100 100	18, 21, 34, 41	0
1	a	9/9 (100%)	-0.01	0 100 100	23, 29, 35, 44	0
2	H	210/236 (88%)	-0.13	2 (0%) 82 84	18, 32, 56, 81	0
2	h	208/236 (88%)	0.08	6 (2%) 51 55	20, 37, 68, 123	0
3	L	219/220 (99%)	-0.11	0 100 100	15, 29, 57, 64	0
3	l	219/220 (99%)	-0.25	0 100 100	17, 30, 52, 67	0
All	All	874/930 (93%)	-0.10	8 (0%) 84 86	15, 32, 58, 123	0

All (8) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	h	167	SER	6.1
2	h	166	GLY	5.1
2	h	169	SER	4.4
2	h	168	LEU	3.5
2	h	1	GLN	2.6
2	H	181	SER	2.4
2	H	65	LYS	2.1
2	h	201	THR	2.0

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.