WORLDWIDE wwPDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

Feb 14, 2023 – 10:46 AM EST

|		
PDB ID	:	6BTJ
Title	:	Crystal structure of pan-H7, anti-hemagglutinin monoclonal antibody H7.5 (Fab fragment)
Authors	:	Shanshan, L.; Ian, A.W.
Deposited on	:	2017-12-06
Resolution	:	2.00 Å(reported)

---

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 2.00 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.
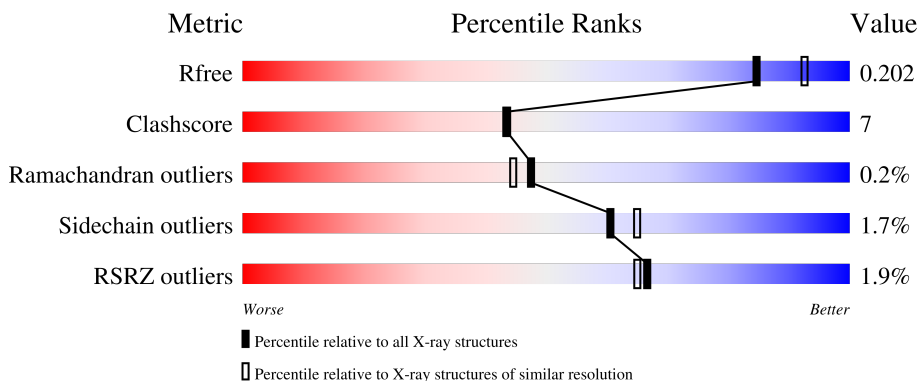
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	8085 (2.00-2.00)
Clashscore	141614	9178 (2.00-2.00)
Ramachandran outliers	138981	9054 (2.00-2.00)
Sidechain outliers	138945	9053 (2.00-2.00)
RSRZ outliers	127900	7900 (2.00-2.00)

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	233	
1	H	233	
2	B	214	
2	L	214	

2 Entry composition ⓘ

There are 3 unique types of molecules in this entry. The entry contains 7300 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 H7.5 Fab.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	216	Total 1630	C 1035	N 270	O 316	S 9	0	0	0
1	H	218	Total 1643	C 1043	N 273	O 318	S 9	0	0	0

- Molecule 2 is a protein called Light chain of H7.5 Fab.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	213	Total 1641	C 1028	N 272	O 336	S 5	0	0	0
2	L	213	Total 1641	C 1028	N 272	O 336	S 5	0	0	0

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	199	Total 199	O 199	0	0
3	B	189	Total 189	O 189	0	0
3	H	195	Total 195	O 195	0	0
3	L	162	Total 162	O 162	0	0

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: Heavy chain of H7.5 Fab
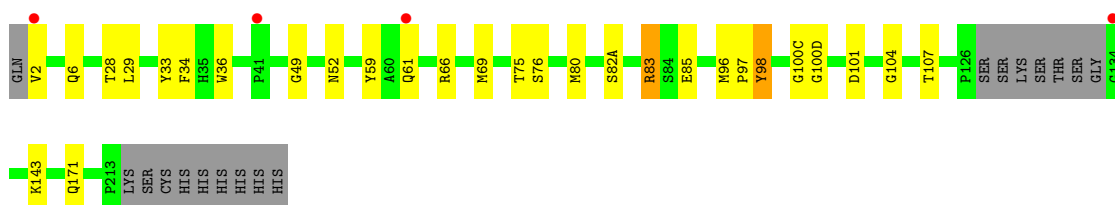
- Molecule 1: Heavy chain of H7.5 Fab
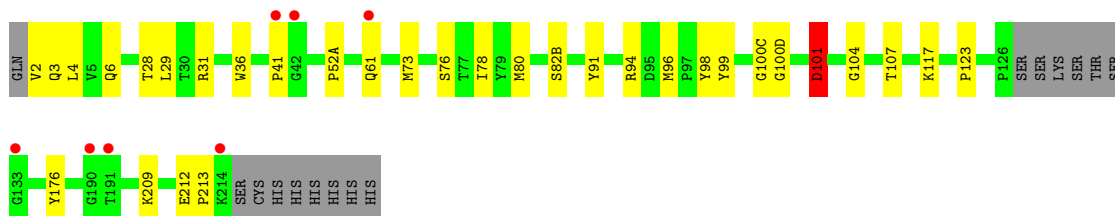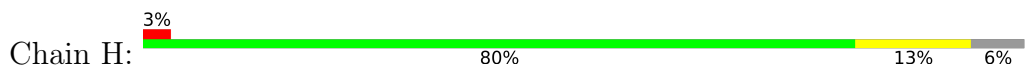
- Molecule 2: Light chain of H7.5 Fab

- Molecule 2: Light chain of H7.5 Fab

4 Data and refinement statistics ⓘ
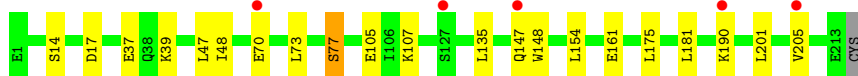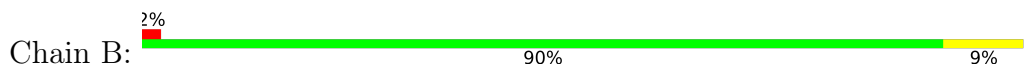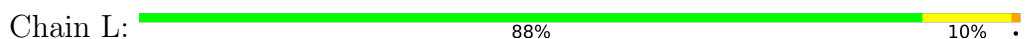
Property	Value	Source
Space group	P 2 21 21	Depositor
Cell constants a, b, c, α, β, γ	65.51Å 99.04Å 208.56Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.52 – 2.00 49.52 – 2.00	Depositor EDS
% Data completeness (in resolution range)	98.5 (49.52-2.00) 98.5 (49.52-2.00)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.52 (at 2.00Å)	Xtriage
Refinement program	PHENIX (1.12_2829: ???)	Depositor
R, R_{free}	0.200 , 0.201 0.199 , 0.202	Depositor DCC
R_{free} test set	4560 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å²)	30.6	Xtriage
Anisotropy	0.567	Xtriage
Bulk solvent k_{sol}(e/Å³), B_{sol}(Å²)	0.35 , 46.1	EDS
L-test for twinning[2]	$< |L| > = 0.49, < L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	7300	wwPDB-VP
Average B, all atoms (Å²)	35.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 21.93 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 6.3244e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

---

[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	0.77	4/1673 (0.2%)	0.67	0/2280
1	H	0.71	3/1686 (0.2%)	0.70	1/2296 (0.0%)
2	B	0.45	0/1677	0.69	2/2277 (0.1%)
2	L	0.44	0/1677	0.67	2/2277 (0.1%)
All	All	0.61	7/6713 (0.1%)	0.68	5/9130 (0.1%)

The worst 5 of 7 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	101	ASP	CB-CG	-6.83	1.37	1.51
1	A	100(D)	GLY	C-O	-5.87	1.14	1.23
1	H	100(C)	GLY	C-O	-5.83	1.14	1.23
1	A	100(C)	GLY	C-O	-5.72	1.14	1.23
1	H	99	TYR	CE1-CZ	-5.51	1.31	1.38

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	181	LEU	CA-CB-CG	6.55	130.37	115.30
1	H	101	ASP	CB-CG-OD1	5.93	123.64	118.30
2	L	135	LEU	CA-CB-CG	5.47	127.88	115.30
2	B	154	LEU	CA-CB-CG	5.12	127.09	115.30
2	L	154	LEU	CA-CB-CG	5.05	126.91	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	1630	0	1584	27	0
1	H	1643	0	1600	22	0
2	B	1641	0	1593	15	0
2	L	1641	0	1593	23	0
3	A	199	0	0	7	1
3	B	189	0	0	9	1
3	H	195	0	0	6	1
3	L	162	0	0	12	1
All	All	7300	0	6370	86	2

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 86 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)
2:B:39:LYS:CE	3:B:302:HOH:O	1.99	1.08
1:H:96:MET:SD	3:H:394:HOH:O	2.14	1.05
2:B:190:LYS:NZ	3:B:301:HOH:O	1.88	1.02
2:B:39:LYS:NZ	3:B:302:HOH:O	1.97	0.93
1:A:61:GLN:OE1	3:A:301:HOH:O	1.92	0.86

All (2) 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)
3:A:465:HOH:O	3:L:376:HOH:O[4_455]	1.83	0.37
3:B:376:HOH:O	3:H:462:HOH:O[4_555]	2.18	0.02

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	212/233 (91%)	208 (98%)	3 (1%)	1 (0%)	29	23
1	H	214/233 (92%)	210 (98%)	3 (1%)	1 (0%)	29	23
2	B	211/214 (99%)	207 (98%)	4 (2%)	0	100	100
2	L	211/214 (99%)	207 (98%)	4 (2%)	0	100	100
All	All	848/894 (95%)	832 (98%)	14 (2%)	2 (0%)	47	44

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	98	TYR
1	H	98	TYR

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	182/198 (92%)	180 (99%)	2 (1%)	73	78
1	H	183/198 (92%)	181 (99%)	2 (1%)	73	78
2	B	190/191 (100%)	186 (98%)	4 (2%)	53	57
2	L	190/191 (100%)	185 (97%)	5 (3%)	46	48
All	All	745/778 (96%)	732 (98%)	13 (2%)	60	65

5 of 13 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	H	107	THR
2	L	14	SER
2	L	208	SER
2	L	135	LEU
2	L	142	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such

sidechains are listed below:

Mol	Chain	Res	Type
1	A	6	GLN
1	H	6	GLN

5.3.3   RNA ⓘ

There are no RNA molecules in this entry.

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 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, 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(Å²)	Q<0.9
1	A	216/233 (92%)	0.17	4 (1%)	66	65	24, 32, 52, 66	4 (1%)
1	H	218/233 (93%)	0.31	7 (3%)	47	46	23, 32, 52, 77	2 (0%)
2	B	213/214 (99%)	0.17	5 (2%)	60	59	23, 33, 48, 67	0
2	L	213/214 (99%)	0.17	0	100	100	24, 36, 46, 69	0
All	All	860/894 (96%)	0.20	16 (1%)	66	65	23, 33, 49, 77	6 (0%)

The worst 5 of 16 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	H	133	GLY	7.4
1	H	41	PRO	3.5
1	A	134	GLY	3.3
1	H	214	LYS	3.0
2	B	127	SER	2.8

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.