wwPDB X-ray Structure Validation Summary Report (i)

Nov 19, 2023 – 06:05 PM JST

PDB ID	:	6LZ9
Title	:	t8E4 antibody Fab complexed with the active form of HGF
Authors	:	Kitago, Y.; Umitsu, M.; Takagi, J.
Deposited on	:	2020-02-18
Resolution	:	2.80 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 (1)) were used in the production of this report:

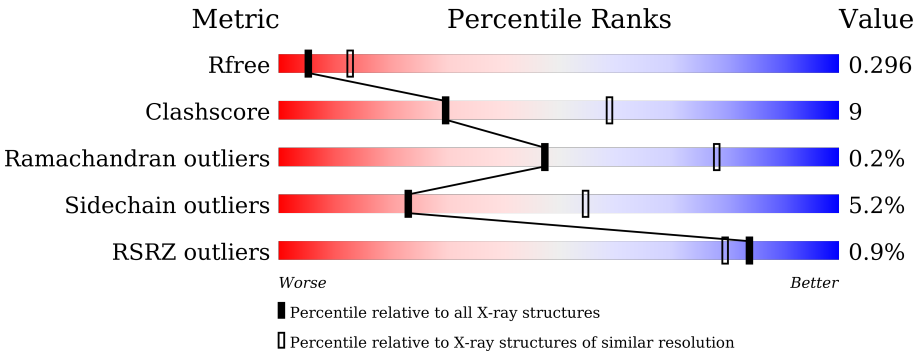
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 2.80 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	3140 (2.80-2.80)
Clashscore	141614	3569 (2.80-2.80)
Ramachandran outliers	138981	3498 (2.80-2.80)
Sidechain outliers	138945	3500 (2.80-2.80)
RSRZ outliers	127900	3078 (2.80-2.80)

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	108	6% 94%
2	B	242	67% 22% 8%
3	H	223	77% 20%
4	L	214	76% 23%

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 5028 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 Hepatocyte growth factor.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	7	48	31	8	8	1	0	0	0

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	387	SER	-	expression tag	UNP P14210
A	402	GLN	ASN	engineered mutation	UNP P14210
A	476	GLY	THR	engineered mutation	UNP P14210
A	491	ILE	LYS	engineered mutation	UNP P14210
A	492	GLU	GLN	engineered mutation	UNP P14210
A	493	GLY	LEU	engineered mutation	UNP P14210

- Molecule 2 is a protein called Hepatocyte growth factor.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	223	1654	1060	275	305	14	0	0	0

There are 11 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	561	SER	CYS	engineered mutation	UNP P14210
B	566	GLN	ASN	engineered mutation	UNP P14210
B	653	GLN	ASN	engineered mutation	UNP P14210
B	729	ARG	-	expression tag	UNP P14210
B	730	LEU	-	expression tag	UNP P14210
B	731	GLU	-	expression tag	UNP P14210
B	732	ASN	-	expression tag	UNP P14210
B	733	LEU	-	expression tag	UNP P14210
B	734	TYR	-	expression tag	UNP P14210
B	735	PHE	-	expression tag	UNP P14210

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	736	GLN	-	expression tag	UNP P14210

- Molecule 3 is a protein called Heavy chain of t8E4 Fab fragment.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	H	218	Total	C	N	O	S	0	0	0
			1656	1053	269	326	8			

- Molecule 4 is a protein called Light chain of t8E4 Fab fragment.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	L	213	Total	C	N	O	S	0	0	0
			1665	1034	281	344	6			

- Molecule 5 is water.

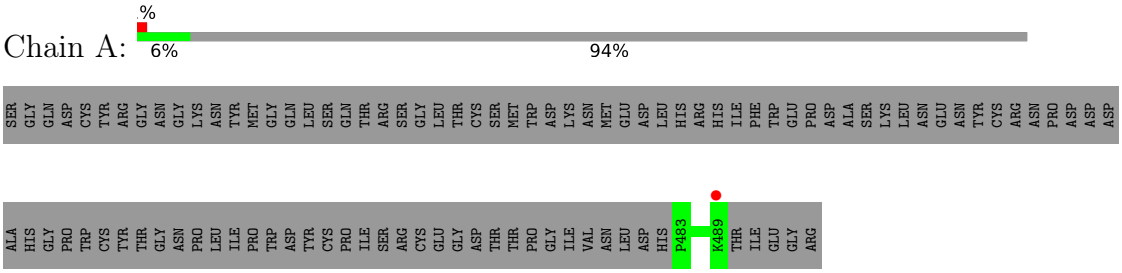
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	B	1	Total	O	0	0
			1	1		
5	H	2	Total	O	0	0
			2	2		
5	L	2	Total	O	0	0
			2	2		

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.

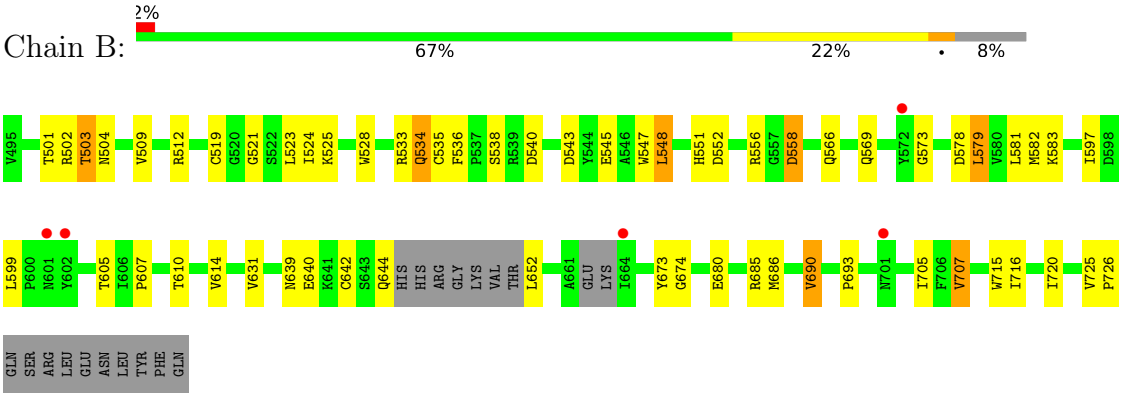
• Molecule 1: Hepatocyte growth factor

Chain A:



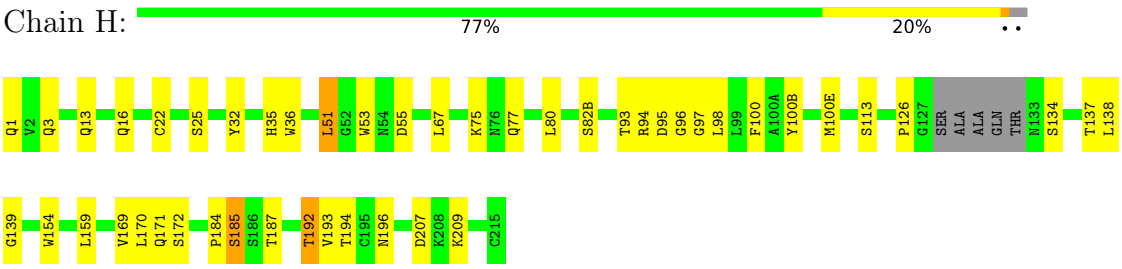
• Molecule 2: Hepatocyte growth factor

Chain B:



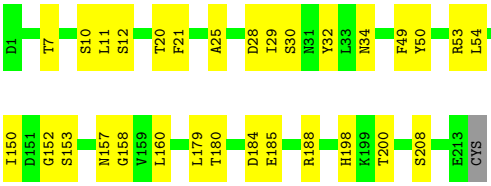
• Molecule 3: Heavy chain of t8E4 Fab fragment

Chain H:



• Molecule 4: Light chain of t8E4 Fab fragment

Chain L:


D1
T7
S10
L11
S12
T20
F21
A25
D28
I29
S30
N31
Y32
L33
N34
F49
Y50
R53
L54
P59
S60
R61
F62
S65
G68
L73
T74
T85
D91
S92
F96
S100
K103
L104
E105
R108
A109
S122
T126
K149
I150
D151
G152
S153
N157
G158
V159
L160
L179
T180
D184
E185
R188
H198
K199
T200
S208
E213
CYS

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 63 2 2	Depositor
Cell constants a, b, c, α, β, γ	176.18Å 176.18Å 122.96Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	47.87 – 2.80 47.87 – 2.80	Depositor EDS
% Data completeness (in resolution range)	99.7 (47.87-2.80) 99.9 (47.87-2.80)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.52 (at 2.81Å)	Xtriage
Refinement program	PHENIX 1.17.1_3660	Depositor
R, R_{free}	0.248 , 0.297 0.247 , 0.296	Depositor DCC
R_{free} test set	1366 reflections (4.85%)	wwPDB-VP
Wilson B-factor (Å^2)	71.2	Xtriage
Anisotropy	0.579	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , 40.7	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	0.92	EDS
Total number of atoms	5028	wwPDB-VP
Average B, all atoms (Å^2)	78.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.10% 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

Bond lengths and bond angles in the following residue types are not validated in this section: PCA

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	0.38	0/48	0.45	0/63
2	B	0.37	0/1694	0.53	0/2314
3	H	0.41	0/1693	0.62	0/2318
4	L	0.39	0/1702	0.59	0/2309
All	All	0.39	0/5137	0.58	0/7004

There are no bond length outliers.

There are no bond angle outliers.

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	48	0	54	0	0
2	B	1654	0	1564	42	0
3	H	1656	0	1615	28	0
4	L	1665	0	1572	29	0
5	B	1	0	0	0	0
5	H	2	0	0	0	0
5	L	2	0	0	0	0
All	All	5028	0	4805	92	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 9.

The worst 5 of 92 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:502:ARG:HH12	4:L:91:ASP:HB3	1.31	0.94
2:B:690:VAL:HG13	2:B:707:VAL:HG13	1.53	0.91
4:L:25:ALA:HB2	4:L:29:ILE:HD11	1.65	0.79
2:B:579:LEU:HD11	2:B:690:VAL:HG11	1.70	0.74
2:B:685:ARG:NH2	3:H:98:LEU:HD12	2.04	0.73

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	5/108 (5%)	5 (100%)	0	0	100 100
2	B	217/242 (90%)	206 (95%)	10 (5%)	1 (0%)	29 61
3	H	214/223 (96%)	204 (95%)	10 (5%)	0	100 100
4	L	211/214 (99%)	198 (94%)	13 (6%)	0	100 100
All	All	647/787 (82%)	613 (95%)	33 (5%)	1 (0%)	47 78

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	534	GLN

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	6/95 (6%)	6 (100%)	0	100	100
2	B	171/206 (83%)	161 (94%)	10 (6%)	20	50
3	H	190/194 (98%)	183 (96%)	7 (4%)	34	68
4	L	191/193 (99%)	179 (94%)	12 (6%)	18	46
All	All	558/688 (81%)	529 (95%)	29 (5%)	23	55

5 of 29 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	H	169	VAL
4	L	185	GLU
4	L	7	THR
4	L	100	SER
3	H	192	THR

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

Mol	Chain	Res	Type
4	L	124	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)

1 non-standard protein/DNA/RNA residue is modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\|>2$	Counts	RMSZ	$\#\|Z\|>2$
3	PCA	H	1	3	7,8,9	1.82	1 (14%)	9,10,12	2.26	5 (55%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	PCA	H	1	3	-	0/0/11/13	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	H	1	PCA	CD-N	4.64	1.46	1.34

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	H	1	PCA	CA-N-CD	-3.40	101.94	113.58
3	H	1	PCA	CB-CA-N	3.01	111.94	103.30
3	H	1	PCA	OE-CD-CG	-2.71	122.04	126.76
3	H	1	PCA	CG-CD-N	2.53	114.95	108.39
3	H	1	PCA	CB-CA-C	-2.51	109.25	112.70

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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	7/108 (6%)	0.64	1 (14%) 2 1	94, 105, 123, 133	0
2	B	223/242 (92%)	0.05	5 (2%) 62 52	57, 88, 117, 140	0
3	H	217/223 (97%)	-0.27	0 100 100	48, 70, 104, 149	0
4	L	213/214 (99%)	-0.21	0 100 100	48, 71, 96, 124	0
All	All	660/787 (83%)	-0.13	6 (0%) 84 80	48, 76, 111, 149	0

The worst 5 of 6 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	601	ASN	4.7
2	B	602	TYR	4.3
2	B	701	ASN	2.9
2	B	572	TYR	2.3
1	A	489	LYS	2.2

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

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
3	PCA	H	1	8/9	0.96	0.13	98,100,104,106	0

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.