



wwPDB X-ray Structure Validation Summary Report ⓘ

Apr 12, 2022 – 02:51 PM EDT

PDB ID : 6UIG
Title : Crystal structure of human monoclonal antibody H7.200 in complex with H7N9 hemagglutinin HA1
Authors : Dong, J.; Crowe, J.E.
Deposited on : 2019-09-30
Resolution : 3.20 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 ⓘ symbol.

The following versions of software and data (see [references ⓘ](#)) 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	:	2.27
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)	:	2.27

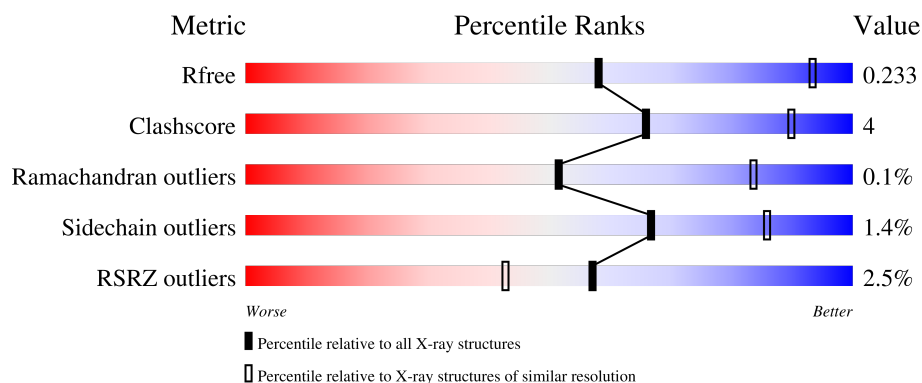
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.20 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	1133 (3.20-3.20)
Clashscore	141614	1253 (3.20-3.20)
Ramachandran outliers	138981	1234 (3.20-3.20)
Sidechain outliers	138945	1233 (3.20-3.20)
RSRZ outliers	127900	1095 (3.20-3.20)

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 $\leq 5\%$. 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	A	308	<div> <div>4%</div> <div>78%</div> <div>9%</div> <div>13%</div> </div>
1	G	308	<div> <div>6%</div> <div>76%</div> <div>12%</div> <div>13%</div> </div>
2	B	227	<div> <div>%</div> <div>85%</div> <div>10%</div> <div>.</div> </div>
2	H	227	<div> <div>%</div> <div>88%</div> <div>7%</div> <div>.</div> </div>
3	C	214	<div> <div>87%</div> <div>10%</div> <div>.</div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	L	214	 90%9% .

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 10231 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 Hemagglutinin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	G	269	Total	C	N	O	S	0	0	0
			1958	1223	342	381	12			
1	A	269	Total	C	N	O	S	0	0	0
			1906	1192	342	361	11			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
G	303	HIS	-	expression tag	UNP A0A4Y5QYN9
G	304	HIS	-	expression tag	UNP A0A4Y5QYN9
G	305	HIS	-	expression tag	UNP A0A4Y5QYN9
G	306	HIS	-	expression tag	UNP A0A4Y5QYN9
G	307	HIS	-	expression tag	UNP A0A4Y5QYN9
G	308	HIS	-	expression tag	UNP A0A4Y5QYN9
A	303	HIS	-	expression tag	UNP A0A4Y5QYN9
A	304	HIS	-	expression tag	UNP A0A4Y5QYN9
A	305	HIS	-	expression tag	UNP A0A4Y5QYN9
A	306	HIS	-	expression tag	UNP A0A4Y5QYN9
A	307	HIS	-	expression tag	UNP A0A4Y5QYN9
A	308	HIS	-	expression tag	UNP A0A4Y5QYN9

- Molecule 2 is a protein called H7.200 Fab heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	H	217	Total	C	N	O	S	0	0	0
			1603	1023	266	309	5			
2	B	218	Total	C	N	O	S	0	0	0
			1594	1020	269	300	5			

- Molecule 3 is a protein called H7.200 Fab light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	L	212	Total	C	N	O	S	0	0	0
			1590	998	272	315	5			
3	C	210	Total	C	N	O	S	0	0	0
			1552	979	263	305	5			

- Molecule 4 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).

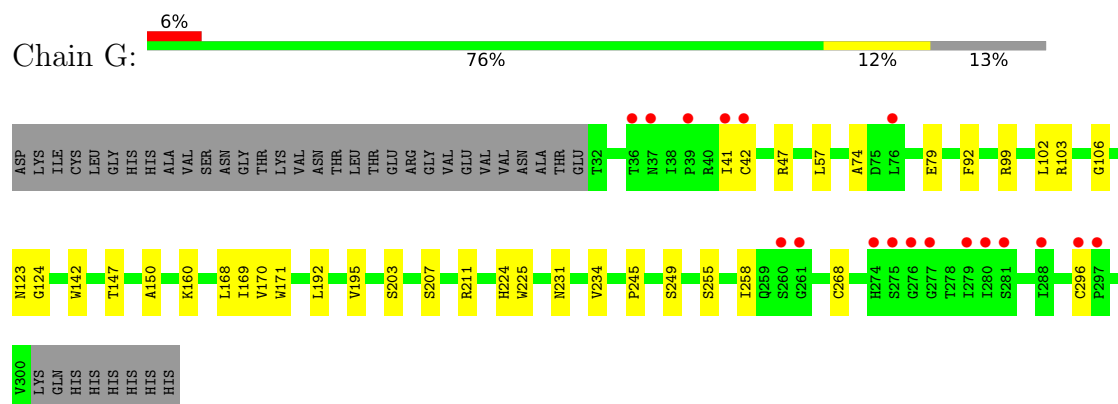


Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	G	1	Total	C	N	O	0	0
			14	8	1	5		
4	A	1	Total	C	N	O	0	0
			14	8	1	5		

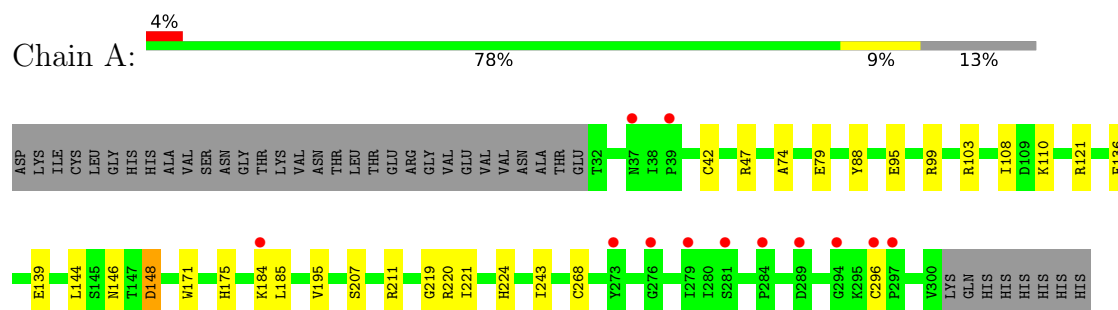
3 Residue-property plots [i](#)

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.

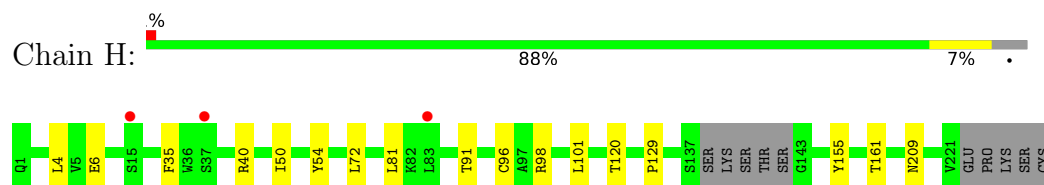
- Molecule 1: Hemagglutinin



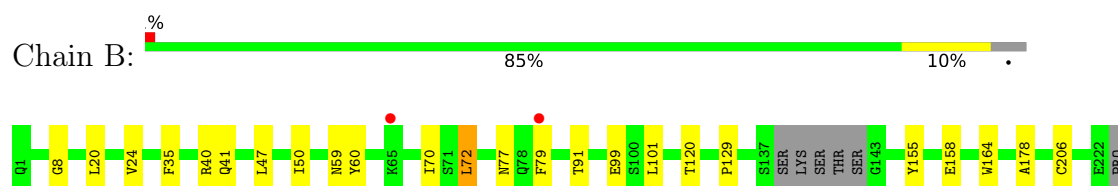
- Molecule 1: Hemagglutinin



- Molecule 2: H7.200 Fab heavy chain

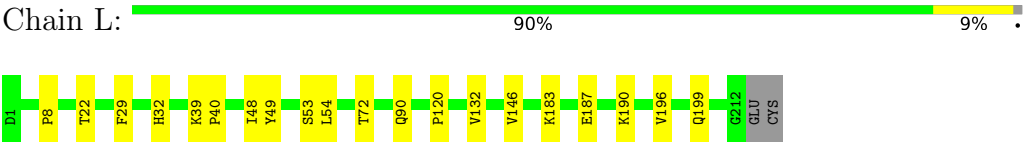


- Molecule 2: H7.200 Fab heavy chain

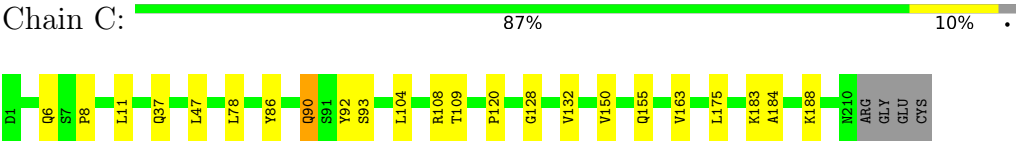


LYS
SER
CYS

- Molecule 3: H7.200 Fab light chain



- Molecule 3: H7.200 Fab light chain



4 Data and refinement statistics

Property	Value	Source
Space group	P 43 2 2	Depositor
Cell constants a, b, c, α , β , γ	156.49Å 156.49Å 229.08Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.82 – 3.20 49.82 – 3.20	Depositor EDS
% Data completeness (in resolution range)	99.8 (49.82-3.20) 99.9 (49.82-3.20)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.48 (at 3.19Å)	Xtriage
Refinement program	PHENIX 1.16_3549	Depositor
R, R_{free}	0.207 , 0.234 0.207 , 0.233	Depositor DCC
R_{free} test set	2380 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å ²)	93.0	Xtriage
Anisotropy	0.107	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.31 , 72.7	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	10231	wwPDB-VP
Average B, all atoms (Å ²)	88.0	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 42.45 % 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 2.0352e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

¹Intensities estimated from amplitudes.

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

5 Model quality [i](#)

5.1 Standard geometry [i](#)

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

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	0.25	0/1948	0.45	0/2650
1	G	0.25	0/2001	0.45	0/2717
2	B	0.25	0/1637	0.47	0/2239
2	H	0.26	0/1646	0.49	0/2255
3	C	0.25	0/1588	0.45	0/2167
3	L	0.26	0/1626	0.46	0/2216
All	All	0.25	0/10446	0.46	0/14244

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 [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	A	1906	0	1648	16	0
1	G	1958	0	1785	18	0
2	B	1594	0	1479	12	0
2	H	1603	0	1518	9	0
3	C	1552	0	1403	14	0
3	L	1590	0	1463	9	0
4	A	14	0	13	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	G	14	0	13	0	0
All	All	10231	0	9322	78	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 4.

The worst 5 of 78 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:79:GLU:OE2	1:A:103:ARG:NH2	2.13	0.81
2:B:40:ARG:HB3	2:B:50:ILE:HD11	1.77	0.67
2:H:4:LEU:HD23	2:H:96:CYS:HB3	1.78	0.66
3:L:120:PRO:HD3	3:L:132:VAL:HG22	1.78	0.64
1:A:110:LYS:NZ	1:A:139:GLU:OE2	2.18	0.64

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	267/308 (87%)	251 (94%)	16 (6%)	0	100	100
1	G	267/308 (87%)	256 (96%)	11 (4%)	0	100	100
2	B	214/227 (94%)	209 (98%)	5 (2%)	0	100	100
2	H	213/227 (94%)	209 (98%)	4 (2%)	0	100	100
3	C	208/214 (97%)	199 (96%)	8 (4%)	1 (0%)	29	67
3	L	210/214 (98%)	201 (96%)	8 (4%)	1 (0%)	29	67
All	All	1379/1498 (92%)	1325 (96%)	52 (4%)	2 (0%)	51	83

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	L	8	PRO
3	C	8	PRO

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	167/259 (64%)	165 (99%)	2 (1%)	71	88
1	G	195/259 (75%)	192 (98%)	3 (2%)	65	85
2	B	161/198 (81%)	159 (99%)	2 (1%)	71	88
2	H	172/198 (87%)	169 (98%)	3 (2%)	60	83
3	C	156/189 (82%)	155 (99%)	1 (1%)	86	94
3	L	165/189 (87%)	162 (98%)	3 (2%)	59	82
All	All	1016/1292 (79%)	1002 (99%)	14 (1%)	67	86

5 of 14 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	L	190	LYS
3	L	199	GLN
3	C	90	GLN
2	B	59	ASN
2	B	72	LEU

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

Mol	Chain	Res	Type
2	B	59	ASN

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](#)

2 ligands are 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	$\# Z > 2$
4	NAG	A	401	1	14,14,15	0.24	0	17,19,21	0.39	0
4	NAG	G	401	1	14,14,15	0.22	0	17,19,21	0.39	0

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
4	NAG	A	401	1	-	2/6/23/26	0/1/1/1
4	NAG	G	401	1	-	0/6/23/26	0/1/1/1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	401	NAG	O5-C5-C6-O6
4	A	401	NAG	C4-C5-C6-O6

There are no ring outliers.

No monomer is involved in short contacts.

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 [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	A	269/308 (87%)	0.05	12 (4%) 33 21	71, 99, 143, 162	1 (0%)
1	G	269/308 (87%)	0.29	18 (6%) 17 10	71, 102, 148, 170	1 (0%)
2	B	218/227 (96%)	-0.05	2 (0%) 84 75	57, 78, 106, 141	0
2	H	217/227 (95%)	-0.00	3 (1%) 75 63	57, 79, 109, 134	0
3	C	210/214 (98%)	-0.15	0 100 100	52, 76, 103, 122	0
3	L	212/214 (99%)	-0.09	0 100 100	53, 73, 95, 114	0
All	All	1395/1498 (93%)	0.02	35 (2%) 57 43	52, 85, 133, 170	2 (0%)

The worst 5 of 35 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	G	279	ILE	5.2
1	G	37	ASN	5.2
1	G	288	ILE	5.2
1	G	280	ILE	4.2
1	A	279	ILE	3.7

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 monosaccharides in this entry.

6.4 Ligands [i](#)

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, 95th 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
4	NAG	A	401	14/15	0.84	0.40	136,150,154,156	0
4	NAG	G	401	14/15	0.85	0.34	155,161,164,164	0

6.5 Other polymers [i](#)

There are no such residues in this entry.