



wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 9, 2023 – 01:57 PM EDT

PDB ID : 7U8M
Title : Crystal structure of chimeric hemagglutinin cH15/3 in complex with broad protective antibodies 31.a.83 and FluA-20
Authors : Zhu, X.; Wilson, I.A.
Deposited on : 2022-03-08
Resolution : 5.39 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 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 : 2.35.1
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) : 2.35.1

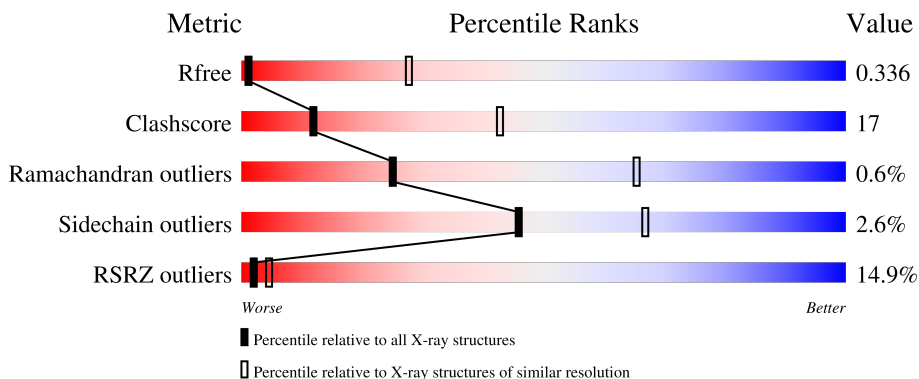
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 5.39 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	1207 (7.00-3.80)
Clashscore	141614	1016 (6.92-3.86)
Ramachandran outliers	138981	1210 (7.00-3.80)
Sidechain outliers	138945	1181 (7.00-3.80)
RSRZ outliers	127900	1021 (7.04-3.76)

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	333	
1	C	333	
1	G	333	
2	B	176	
2	D	176	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	I	176	
3	E	214	
3	J	214	
3	L	214	
4	F	234	
4	H	234	
4	K	234	
5	N	229	
5	Q	229	
5	T	229	
6	O	214	
6	R	214	
6	U	214	

2 Entry composition [i](#)

There are 6 unique types of molecules in this entry. The entry contains 31560 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 HA1 subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	322	Total 2482	C 1545	N 444	O 480	S 13	0	0	0
1	C	321	Total 2474	C 1539	N 443	O 479	S 13	0	0	0
1	G	322	Total 2482	C 1545	N 444	O 480	S 13	0	0	0

- Molecule 2 is a protein called hemagglutinin HA2 subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	173	Total 1396	C 869	N 247	O 274	S 6	0	0	0
2	D	172	Total 1390	C 866	N 246	O 272	S 6	0	0	0
2	I	173	Total 1396	C 869	N 247	O 274	S 6	0	0	0

- Molecule 3 is a protein called Antibody 31.a.83 Fab light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	L	213	Total 1643	C 1025	N 281	O 331	S 6	0	0	0
3	E	213	Total 1643	C 1025	N 281	O 331	S 6	0	0	0
3	J	213	Total 1643	C 1025	N 281	O 331	S 6	0	0	0

- Molecule 4 is a protein called antibody 31.a.83 Fab heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
4	H	226	Total 1699	C 1074	N 280	O 338	S 7	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	F	226	Total	C	N	O	S	0	0	0
			1699	1074	280	338	7			
4	K	226	Total	C	N	O	S	0	0	0
			1699	1074	280	338	7			

- Molecule 5 is a protein called antibody FluA-20 Fab heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	N	222	Total	C	N	O	S	0	0	0
			1662	1050	273	332	7			
5	Q	221	Total	C	N	O	S	0	0	0
			1658	1048	272	331	7			
5	T	221	Total	C	N	O	S	0	0	0
			1656	1047	272	330	7			

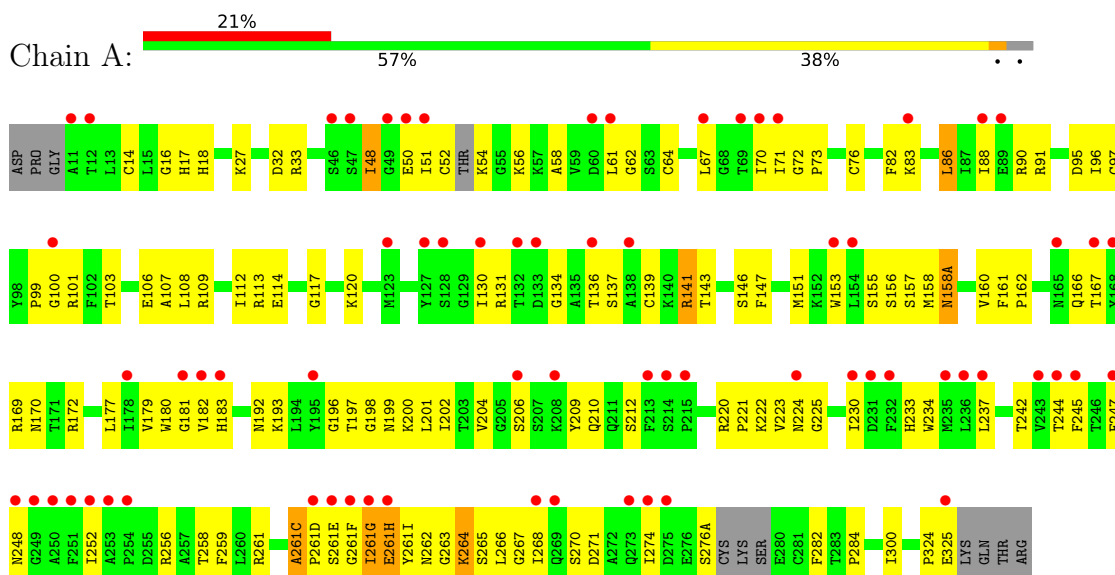
- Molecule 6 is a protein called antibody FluA-20 Fab light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	O	213	Total	C	N	O	S	0	0	0
			1646	1031	281	329	5			
6	R	213	Total	C	N	O	S	0	0	0
			1646	1031	281	329	5			
6	U	213	Total	C	N	O	S	0	0	0
			1646	1031	281	329	5			

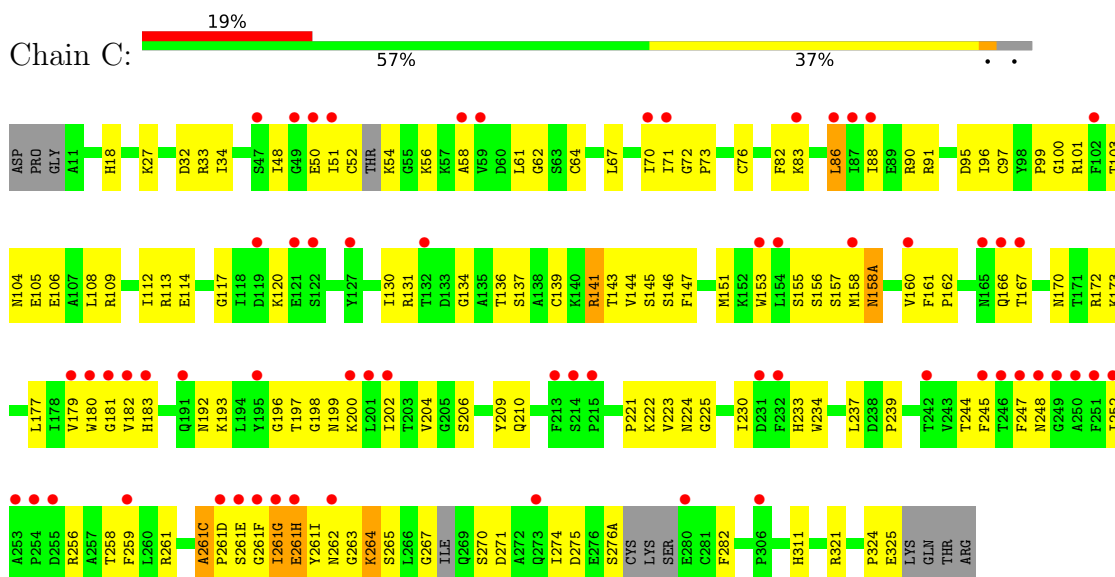
3 Residue-property plots [i](#)

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.

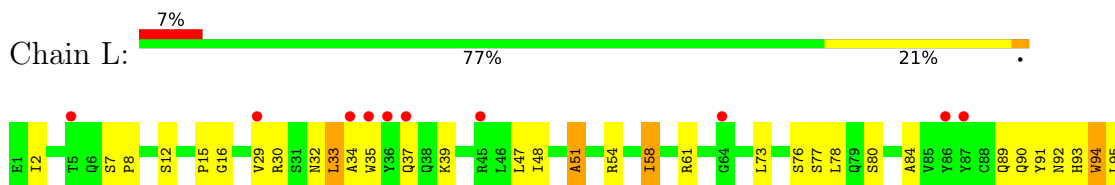
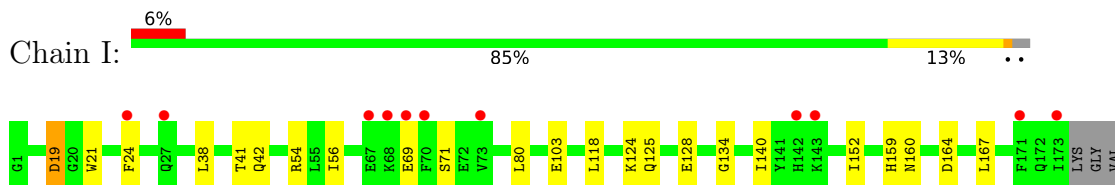
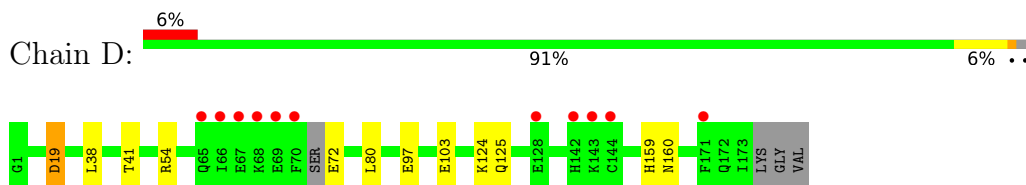
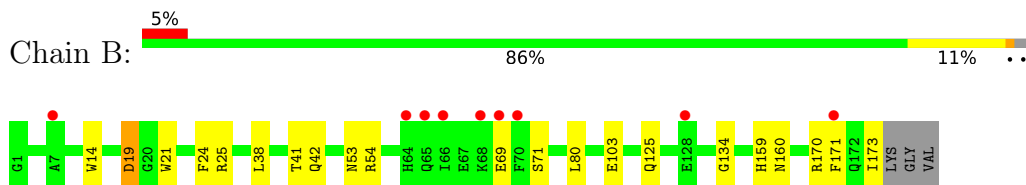
- Molecule 1: hemagglutinin HA1 subunit



- Molecule 1: hemagglutinin HA1 subunit

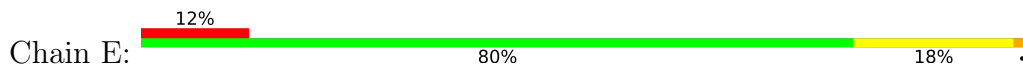


- Molecule 1: hemagglutinin HA1 subunit

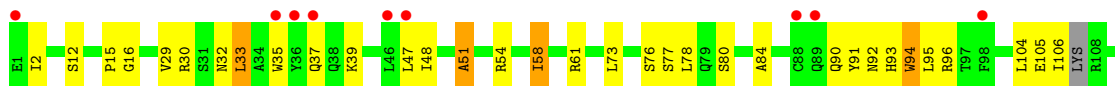
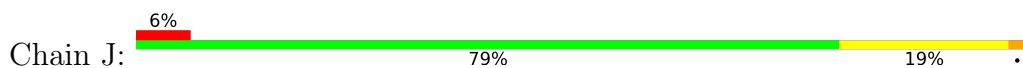




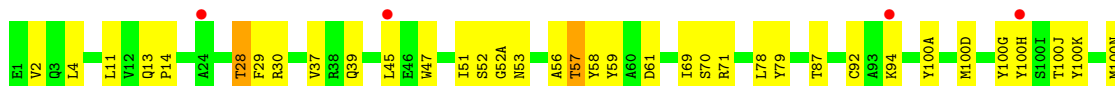
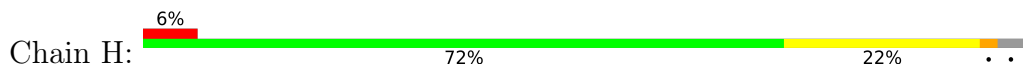
- Molecule 3: Antibody 31.a.83 Fab light chain



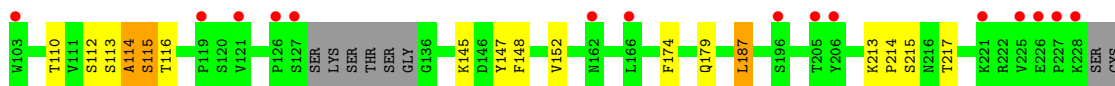
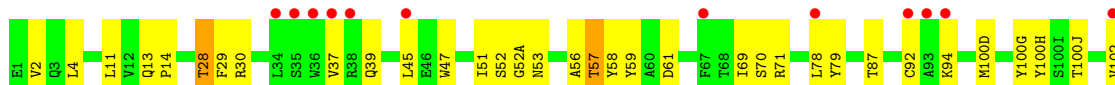
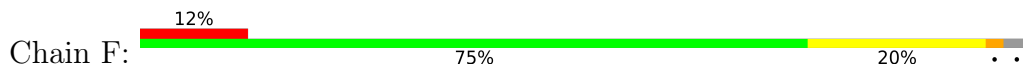
- Molecule 3: Antibody 31.a.83 Fab light chain



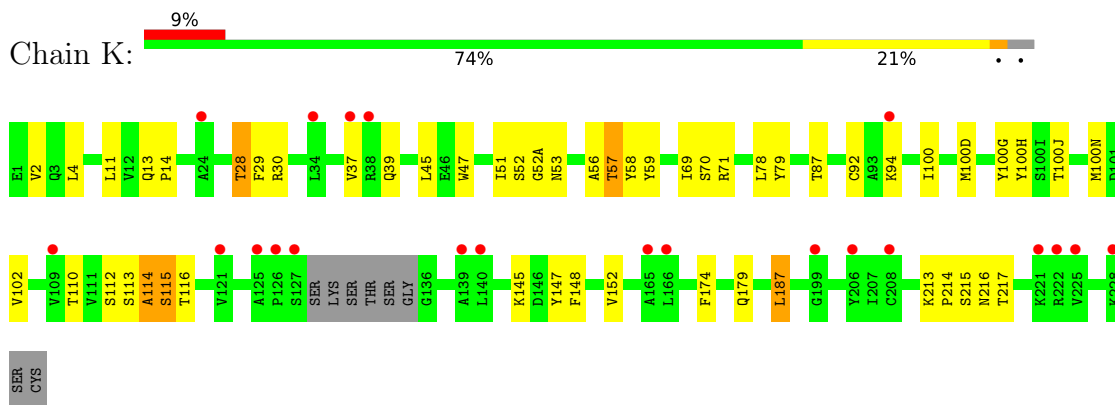
- Molecule 4: antibody 31.a.83 Fab heavy chain



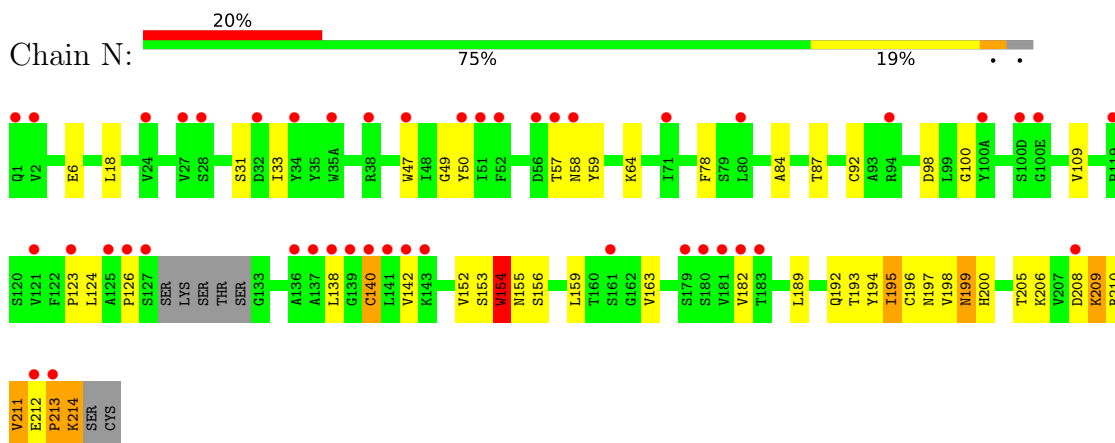
- Molecule 4: antibody 31.a.83 Fab heavy chain



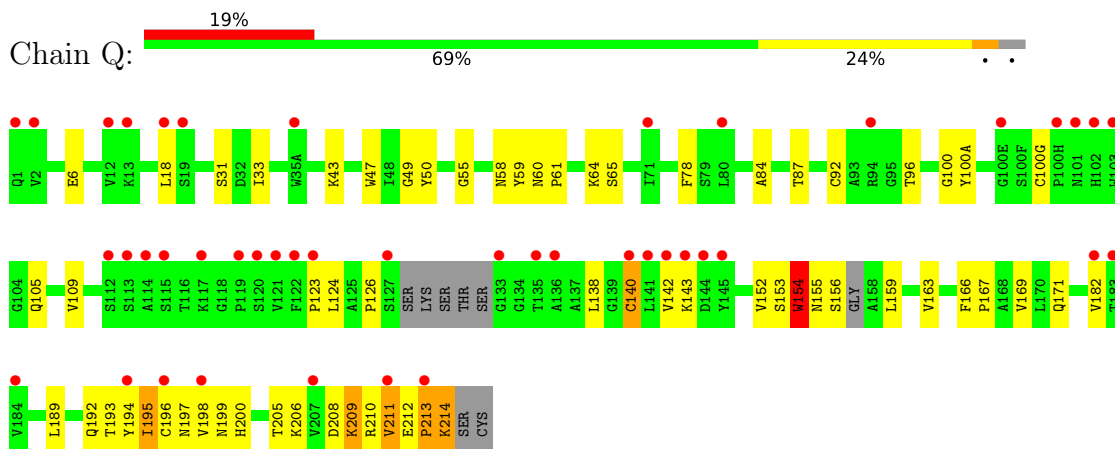
- Molecule 4: antibody 31.a.83 Fab heavy chain



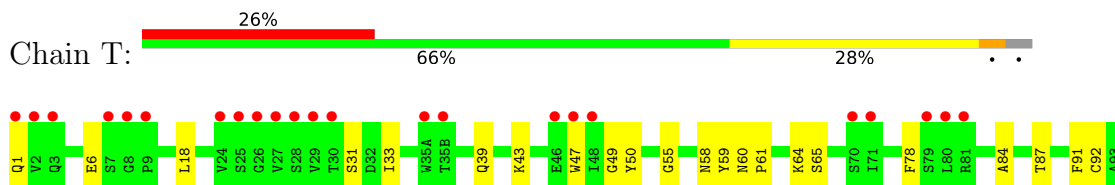
- Molecule 5: antibody FluA-20 Fab heavy chain

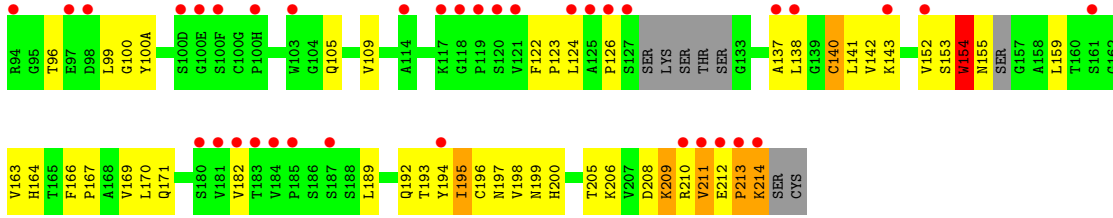


- Molecule 5: antibody FluA-20 Fab heavy chain

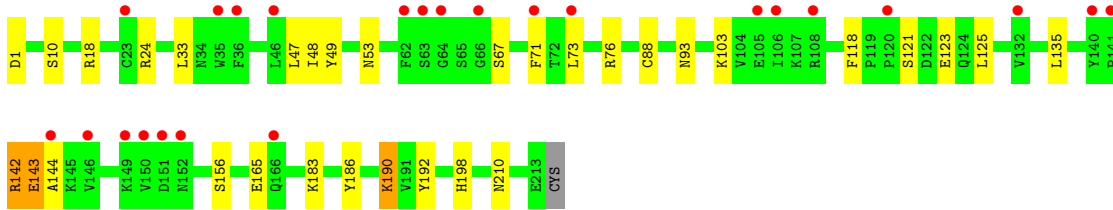
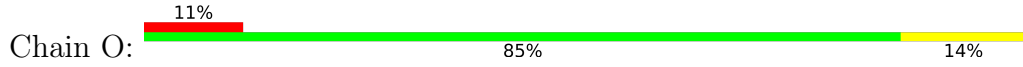


- Molecule 5: antibody FluA-20 Fab heavy chain

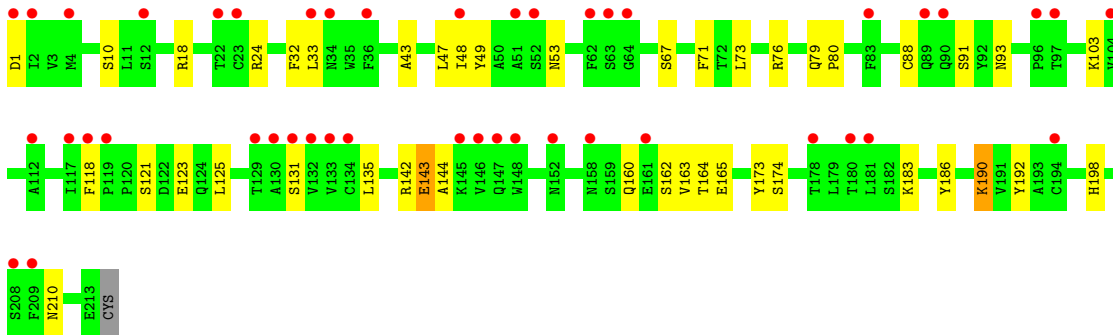
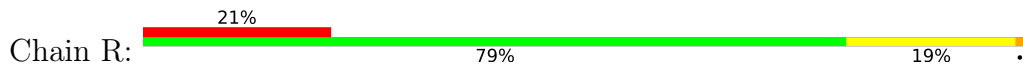




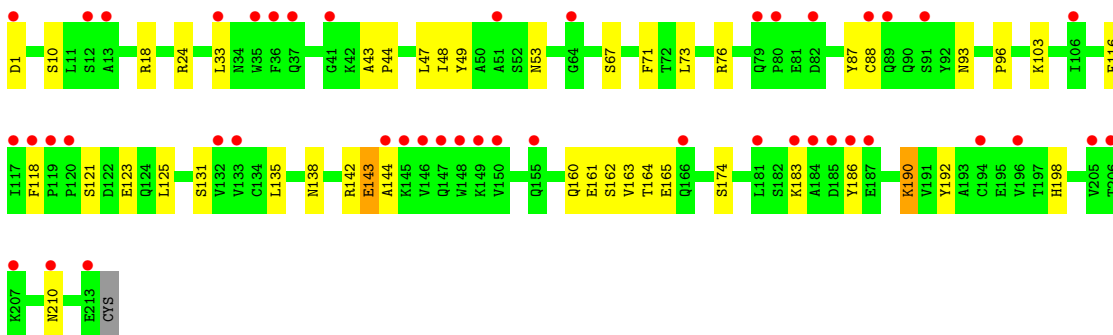
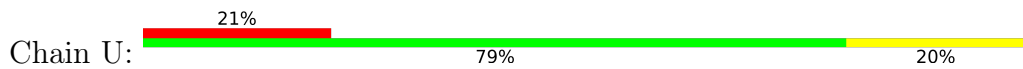
• Molecule 6: antibody FluA-20 Fab light chain



• Molecule 6: antibody FluA-20 Fab light chain



• Molecule 6: antibody FluA-20 Fab light chain



4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	272.92Å 155.73Å 156.89Å 90.00° 94.63° 90.00°	Depositor
Resolution (Å)	49.94 – 5.39 49.94 – 5.39	Depositor EDS
% Data completeness (in resolution range)	95.9 (49.94-5.39) 96.0 (49.94-5.39)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.10	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.46 (at 5.39Å)	Xtrriage
Refinement program	PHENIX 1.19.2_4158	Depositor
R, R_{free}	0.304 , 0.342 0.304 , 0.336	Depositor DCC
R_{free} test set	1073 reflections (4.92%)	wwPDB-VP
Wilson B-factor (Å ²)	248.6	Xtrriage
Anisotropy	0.330	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.32 , 294.7	EDS
L-test for twinning ²	$\langle L \rangle = 0.40$, $\langle L^2 \rangle = 0.23$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.81	EDS
Total number of atoms	31560	wwPDB-VP
Average B, all atoms (Å ²)	312.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.21% of the height of the origin peak. No significant pseudotranslation is detected.*

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](#)

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	0.49	0/2531	0.75	5/3422 (0.1%)
1	C	0.49	0/2522	0.75	5/3408 (0.1%)
1	G	0.49	0/2531	0.75	5/3422 (0.1%)
2	B	0.31	0/1420	0.46	0/1906
2	D	0.31	0/1413	0.46	0/1895
2	I	0.31	0/1420	0.46	0/1906
3	E	0.43	0/1677	0.55	0/2277
3	J	0.42	0/1677	0.55	0/2277
3	L	0.43	0/1677	0.55	0/2277
4	F	0.36	0/1741	0.56	0/2376
4	H	0.36	0/1741	0.56	0/2376
4	K	0.36	0/1741	0.56	0/2376
5	N	0.70	1/1704 (0.1%)	0.91	4/2330 (0.2%)
5	Q	0.70	1/1699 (0.1%)	0.91	3/2322 (0.1%)
5	T	0.70	1/1697 (0.1%)	0.91	3/2319 (0.1%)
6	O	0.66	1/1683 (0.1%)	0.95	6/2285 (0.3%)
6	R	0.66	1/1683 (0.1%)	0.95	6/2285 (0.3%)
6	U	0.66	1/1683 (0.1%)	0.94	6/2285 (0.3%)
All	All	0.51	6/32240 (0.0%)	0.73	43/43744 (0.1%)

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
5	N	0	1
5	Q	0	1
5	T	0	1
6	O	0	1
6	R	0	1
6	U	0	1
All	All	0	6

The worst 5 of 6 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	Q	140	CYS	CB-SG	-6.67	1.71	1.82
5	T	140	CYS	CB-SG	-6.67	1.71	1.82
5	N	140	CYS	CB-SG	-6.63	1.71	1.82
6	U	93	ASN	C-N	-5.47	1.21	1.34
6	O	93	ASN	C-N	-5.47	1.21	1.34

The worst 5 of 43 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	261(C)	ALA	C-N-CD	-9.17	100.43	120.60
1	G	261(C)	ALA	C-N-CD	-9.15	100.47	120.60
1	A	261(C)	ALA	C-N-CD	-9.15	100.48	120.60
6	R	24	ARG	NE-CZ-NH1	9.14	124.87	120.30
6	O	24	ARG	NE-CZ-NH1	9.10	124.85	120.30

There are no chirality outliers.

5 of 6 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
5	N	154	TRP	Mainchain
6	O	143	GLU	Mainchain
5	Q	154	TRP	Mainchain
6	R	143	GLU	Mainchain
5	T	154	TRP	Mainchain

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	2482	0	2431	162	0
1	C	2474	0	2418	148	0
1	G	2482	0	2430	183	0
2	B	1396	0	1331	49	0
2	D	1390	0	1324	26	0
2	I	1396	0	1330	40	0
3	E	1643	0	1592	41	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	J	1643	0	1592	45	0
3	L	1643	0	1592	46	0
4	F	1699	0	1650	67	0
4	H	1699	0	1650	85	0
4	K	1699	0	1650	56	0
5	N	1662	0	1618	59	0
5	Q	1658	0	1614	90	0
5	T	1656	0	1612	121	0
6	O	1646	0	1608	38	0
6	R	1646	0	1608	64	0
6	U	1646	0	1608	74	0
All	All	31560	0	30658	1078	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 17.

The worst 5 of 1078 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:G:51:ILE:HD11	1:G:268:ILE:CG2	1.37	1.51
1:A:106:GLU:CB	2:B:71:SER:HB2	1.47	1.43
1:C:101:ARG:NE	5:N:100:GLY:HA2	1.40	1.32
1:G:106:GLU:CB	2:I:71:SER:HB2	1.60	1.31
1:G:106:GLU:HB3	2:I:71:SER:CB	1.60	1.31

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	316/333 (95%)	307 (97%)	7 (2%)	2 (1%)	25 65

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	C	313/333 (94%)	306 (98%)	5 (2%)	2 (1%)	25	65
1	G	316/333 (95%)	308 (98%)	6 (2%)	2 (1%)	25	65
2	B	171/176 (97%)	164 (96%)	7 (4%)	0	100	100
2	D	168/176 (96%)	161 (96%)	7 (4%)	0	100	100
2	I	171/176 (97%)	164 (96%)	7 (4%)	0	100	100
3	E	209/214 (98%)	193 (92%)	13 (6%)	3 (1%)	11	46
3	J	209/214 (98%)	193 (92%)	13 (6%)	3 (1%)	11	46
3	L	209/214 (98%)	193 (92%)	13 (6%)	3 (1%)	11	46
4	F	222/234 (95%)	209 (94%)	11 (5%)	2 (1%)	17	56
4	H	222/234 (95%)	209 (94%)	11 (5%)	2 (1%)	17	56
4	K	222/234 (95%)	209 (94%)	11 (5%)	2 (1%)	17	56
5	N	218/229 (95%)	209 (96%)	8 (4%)	1 (0%)	29	69
5	Q	215/229 (94%)	208 (97%)	6 (3%)	1 (0%)	29	69
5	T	215/229 (94%)	208 (97%)	6 (3%)	1 (0%)	29	69
6	O	211/214 (99%)	207 (98%)	4 (2%)	0	100	100
6	R	211/214 (99%)	207 (98%)	4 (2%)	0	100	100
6	U	211/214 (99%)	207 (98%)	4 (2%)	0	100	100
All	All	4029/4200 (96%)	3862 (96%)	143 (4%)	24 (1%)	25	65

5 of 24 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	L	94	TRP
3	E	94	TRP
3	J	94	TRP
1	A	141	ARG
3	L	76	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	275/285 (96%)	269 (98%)	6 (2%)	52	71
1	C	274/285 (96%)	268 (98%)	6 (2%)	52	71
1	G	275/285 (96%)	269 (98%)	6 (2%)	52	71
2	B	146/148 (99%)	144 (99%)	2 (1%)	67	81
2	D	145/148 (98%)	143 (99%)	2 (1%)	67	81
2	I	146/148 (99%)	144 (99%)	2 (1%)	67	81
3	E	186/187 (100%)	182 (98%)	4 (2%)	52	71
3	J	186/187 (100%)	182 (98%)	4 (2%)	52	71
3	L	186/187 (100%)	182 (98%)	4 (2%)	52	71
4	F	189/196 (96%)	182 (96%)	7 (4%)	34	58
4	H	189/196 (96%)	182 (96%)	7 (4%)	34	58
4	K	189/196 (96%)	181 (96%)	8 (4%)	30	54
5	N	192/199 (96%)	182 (95%)	10 (5%)	23	49
5	Q	192/199 (96%)	182 (95%)	10 (5%)	23	49
5	T	191/199 (96%)	182 (95%)	9 (5%)	26	52
6	O	189/190 (100%)	187 (99%)	2 (1%)	73	85
6	R	189/190 (100%)	187 (99%)	2 (1%)	73	85
6	U	189/190 (100%)	187 (99%)	2 (1%)	73	85
All	All	3528/3615 (98%)	3435 (97%)	93 (3%)	46	67

5 of 93 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
4	K	216	ASN
5	Q	33	ILE
5	N	33	ILE
5	N	209	LYS
5	Q	193	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 21 such sidechains are listed below:

Mol	Chain	Res	Type
4	K	172	HIS
6	R	138	ASN
6	U	199	GLN
6	R	199	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
6	O	199	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 [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	322/333 (96%)	1.22	69 (21%) 0 2	219, 365, 393, 401	0
1	C	321/333 (96%)	1.13	62 (19%) 1 2	219, 348, 383, 391	0
1	G	322/333 (96%)	1.39	72 (22%) 0 2	219, 375, 402, 405	0
2	B	173/176 (98%)	0.66	9 (5%) 27 25	213, 245, 324, 353	0
2	D	172/176 (97%)	0.71	11 (6%) 19 17	214, 248, 328, 355	0
2	I	173/176 (98%)	0.68	11 (6%) 19 17	215, 248, 325, 359	0
3	E	213/214 (99%)	0.66	26 (12%) 4 7	246, 298, 321, 344	0
3	J	213/214 (99%)	0.64	13 (6%) 21 19	244, 289, 328, 366	0
3	L	213/214 (99%)	0.49	14 (6%) 18 16	245, 279, 309, 331	0
4	F	226/234 (96%)	0.77	27 (11%) 4 7	230, 276, 345, 359	0
4	H	226/234 (96%)	0.61	15 (6%) 18 16	239, 268, 306, 315	0
4	K	226/234 (96%)	0.68	21 (9%) 8 10	228, 267, 327, 340	0
5	N	222/229 (96%)	0.98	45 (20%) 1 2	303, 333, 349, 359	0
5	Q	221/229 (96%)	1.02	44 (19%) 1 2	328, 347, 373, 381	0
5	T	221/229 (96%)	1.34	59 (26%) 0 1	337, 370, 384, 388	0
6	O	213/214 (99%)	0.81	24 (11%) 5 7	287, 309, 343, 348	0
6	R	213/214 (99%)	1.13	44 (20%) 1 2	330, 347, 366, 372	0
6	U	213/214 (99%)	1.09	45 (21%) 1 2	336, 357, 385, 388	0
All	All	4103/4200 (97%)	0.92	611 (14%) 2 4	213, 317, 386, 405	0

The worst 5 of 611 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	G	51	ILE	13.2
1	A	181	GLY	11.1
1	A	182	VAL	10.4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	C	51	ILE	9.7
1	G	59	VAL	8.7

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](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.