wwPDB X-ray Structure Validation Summary Report (i)

Mar 15, 2023 – 06:24 pm GMT

PDB ID	:	8BBN
Title	:	SARS-CoV-2 Delta-RBD complexed with BA.2-10 and EY6A Fabs
Authors	:	Zhou, D.; Ren, J.; Stuart, D.I.
Deposited on	:	2022-10-14
Resolution	:	3.58 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 (i)) were used in the production of this report:

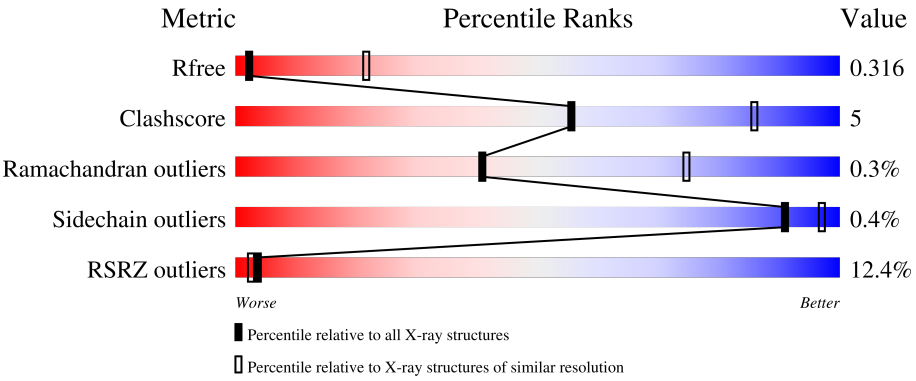
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.32.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.32.1

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 3.58 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.



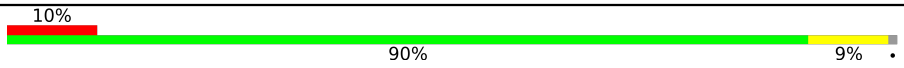
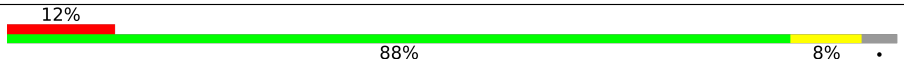
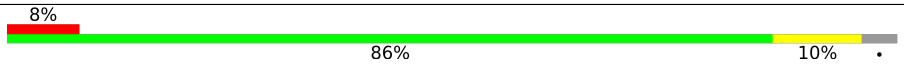
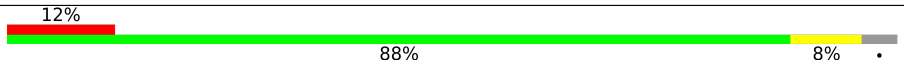
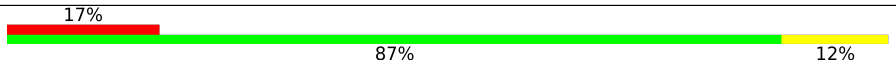
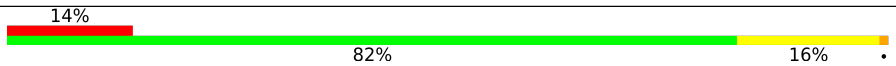
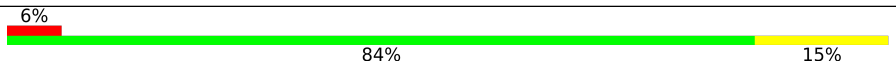
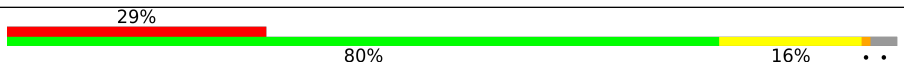
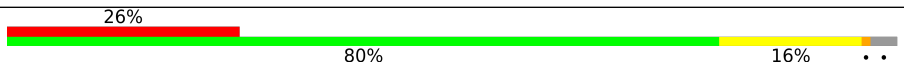
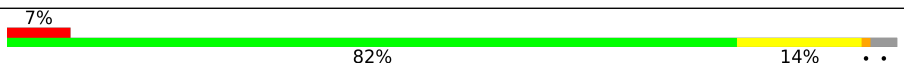
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1094 (3.66-3.50)
Clashscore	141614	1181 (3.66-3.50)
Ramachandran outliers	138981	1143 (3.66-3.50)
Sidechain outliers	138945	1143 (3.66-3.50)
RSRZ outliers	127900	1012 (3.66-3.50)

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	E	226	8% 82% 13% 5%
1	J	226	8% 81% 14% 5%
1	M	226	3% 81% 15% 5%
2	F	215	7% 88% 11% .
2	K	215	13% 86% 13% .

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	N	215	
3	R	202	
3	X	202	
3	Y	202	
4	B	214	
4	D	214	
4	L	214	
5	A	230	
5	C	230	
5	H	230	

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 24354 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 EY6A Heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	M	215	Total	C	N	O	S	0	0	0
			1637	1041	273	317	6			
1	J	215	Total	C	N	O	S	0	0	0
			1637	1041	273	317	6			
1	E	215	Total	C	N	O	S	0	0	0
			1637	1041	273	317	6			

- Molecule 2 is a protein called EY6A light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	N	212	Total	C	N	O	S	0	0	0
			1618	1012	270	331	5			
2	K	212	Total	C	N	O	S	0	0	0
			1618	1012	270	331	5			
2	F	212	Total	C	N	O	S	0	0	0
			1618	1012	270	331	5			

- Molecule 3 is a protein called Spike protein S1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	R	194	Total	C	N	O	S	0	1	0
			1547	990	263	286	8			
3	Y	194	Total	C	N	O	S	0	1	0
			1547	990	263	286	8			
3	X	194	Total	C	N	O	S	0	1	0
			1547	990	263	286	8			

There are 27 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
R	327	HIS	-	expression tag	UNP P0DTC2
R	328	HIS	-	expression tag	UNP P0DTC2

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
R	329	HIS	-	expression tag	UNP P0DTC2
R	330	HIS	-	expression tag	UNP P0DTC2
R	331	HIS	-	expression tag	UNP P0DTC2
R	332	HIS	-	expression tag	UNP P0DTC2
R	452	ARG	LEU	variant	UNP P0DTC2
R	478	LYS	THR	variant	UNP P0DTC2
R	527	LYS	PRO	variant	UNP P0DTC2
Y	327	HIS	-	expression tag	UNP P0DTC2
Y	328	HIS	-	expression tag	UNP P0DTC2
Y	329	HIS	-	expression tag	UNP P0DTC2
Y	330	HIS	-	expression tag	UNP P0DTC2
Y	331	HIS	-	expression tag	UNP P0DTC2
Y	332	HIS	-	expression tag	UNP P0DTC2
Y	452	ARG	LEU	variant	UNP P0DTC2
Y	478	LYS	THR	variant	UNP P0DTC2
Y	527	LYS	PRO	variant	UNP P0DTC2
X	327	HIS	-	expression tag	UNP P0DTC2
X	328	HIS	-	expression tag	UNP P0DTC2
X	329	HIS	-	expression tag	UNP P0DTC2
X	330	HIS	-	expression tag	UNP P0DTC2
X	331	HIS	-	expression tag	UNP P0DTC2
X	332	HIS	-	expression tag	UNP P0DTC2
X	452	ARG	LEU	variant	UNP P0DTC2
X	478	LYS	THR	variant	UNP P0DTC2
X	527	LYS	PRO	variant	UNP P0DTC2

- Molecule 4 is a protein called BA.2-10 light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	L	213	Total	C	N	O	S	0	0	0
			1642	1028	274	336	4			
4	D	213	Total	C	N	O	S	0	0	0
			1642	1028	274	336	4			
4	B	213	Total	C	N	O	S	0	0	0
			1642	1028	274	336	4			

- Molecule 5 is a protein called BA.2-10 heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	H	223	Total	C	N	O	S	0	0	0
			1674	1057	286	323	8			

Continued on next page...

Continued from previous page...

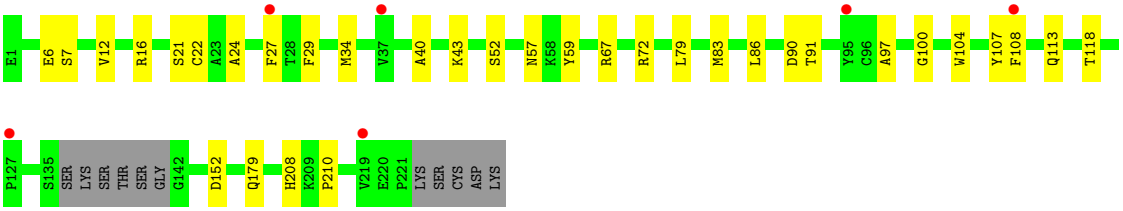
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	C	223	Total	C	N	O	S	0	0	0
			1674	1057	286	323	8			
5	A	223	Total	C	N	O	S	0	0	0
			1674	1057	286	323	8			

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.

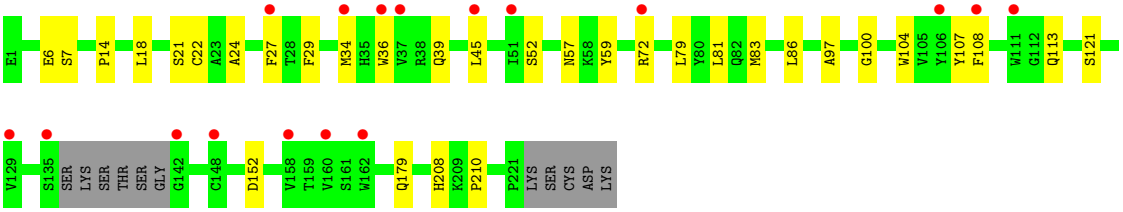
• Molecule 1: EY6A Heavy chain

Chain M:



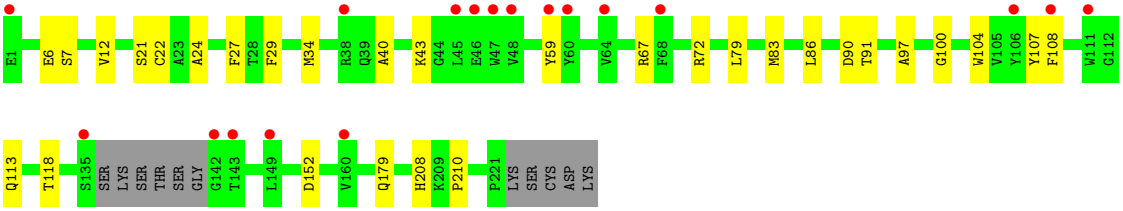
• Molecule 1: EY6A Heavy chain

Chain J:



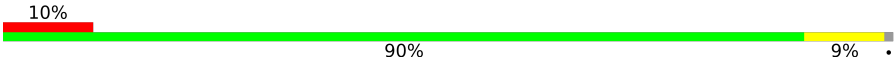
• Molecule 1: EY6A Heavy chain

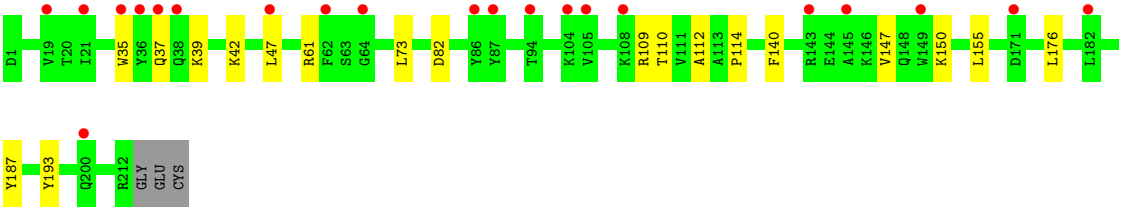
Chain E:



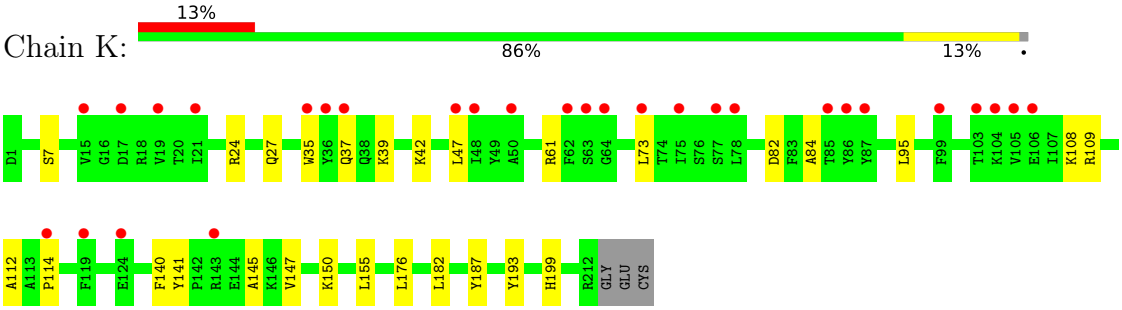
• Molecule 2: EY6A light chain

Chain N:

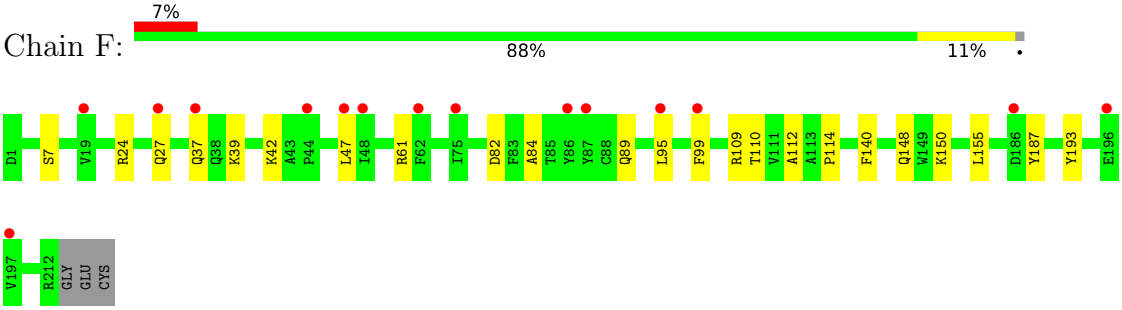




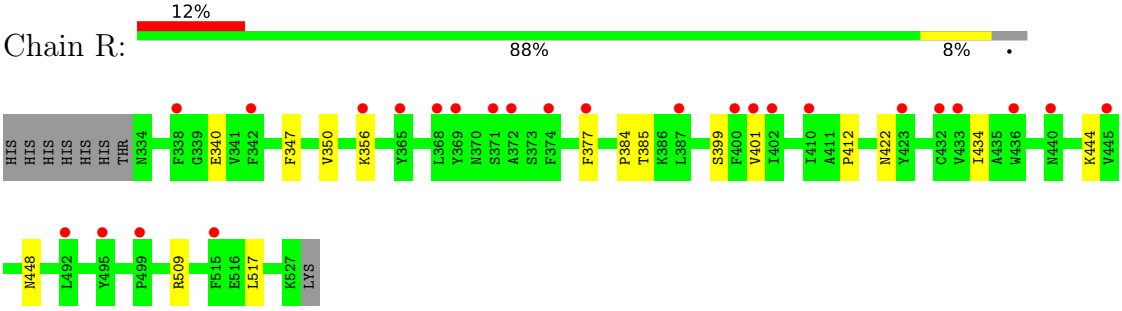
• Molecule 2: EY6A light chain



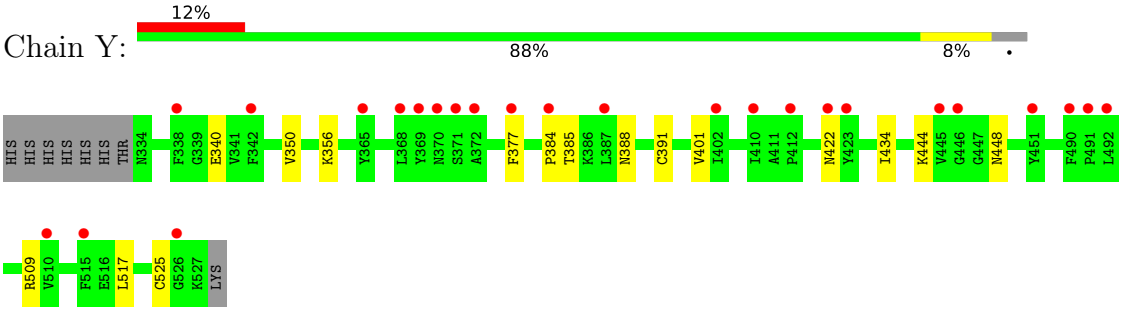
• Molecule 2: EY6A light chain



• Molecule 3: Spike protein S1

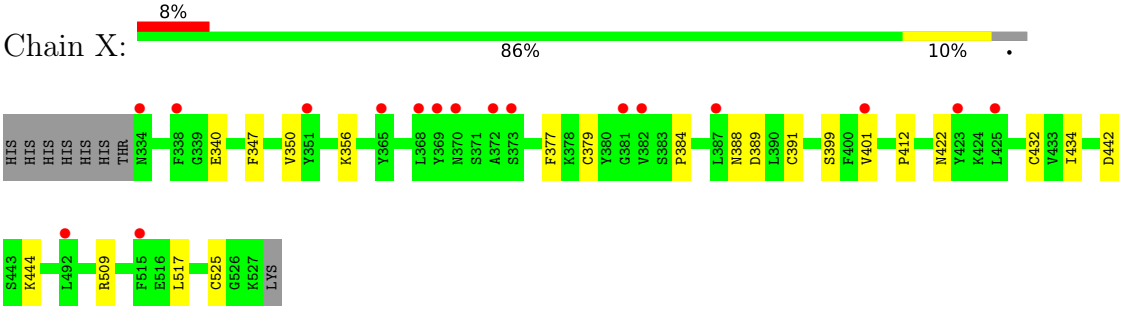


• Molecule 3: Spike protein S1



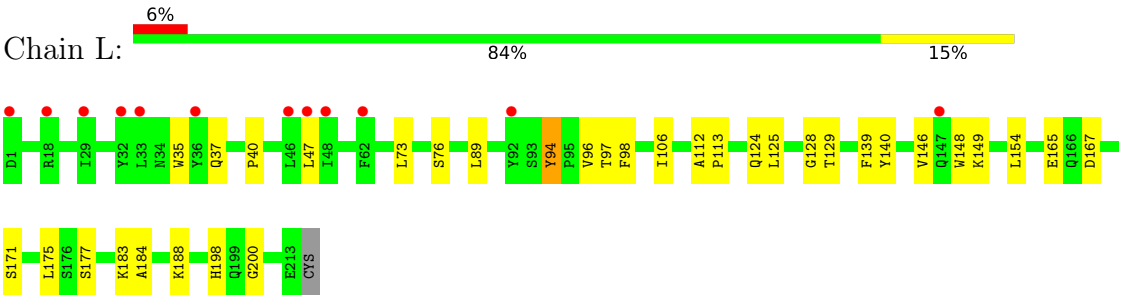
• Molecule 3: Spike protein S1

Chain X:



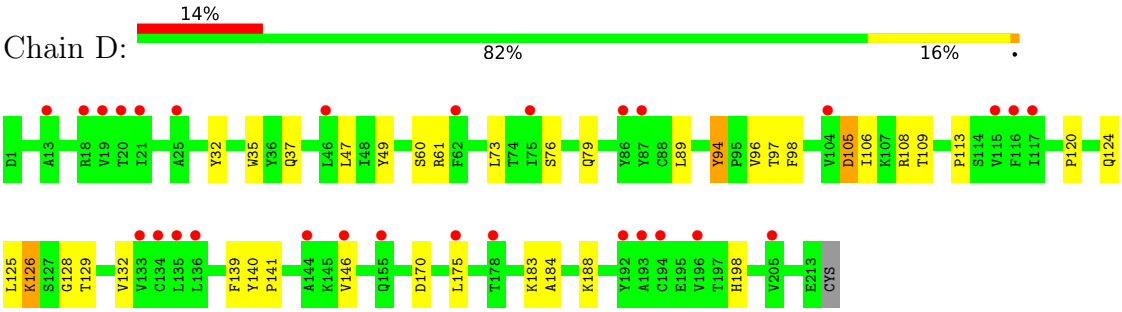
• Molecule 4: BA.2-10 light chain

Chain L:



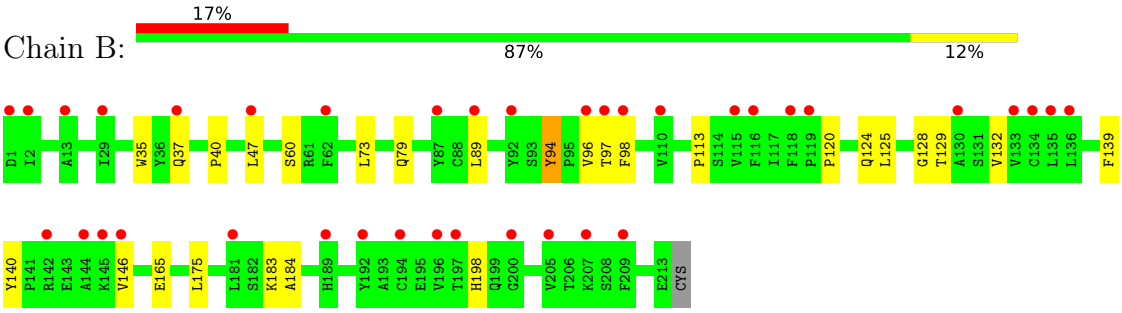
• Molecule 4: BA.2-10 light chain

Chain D:



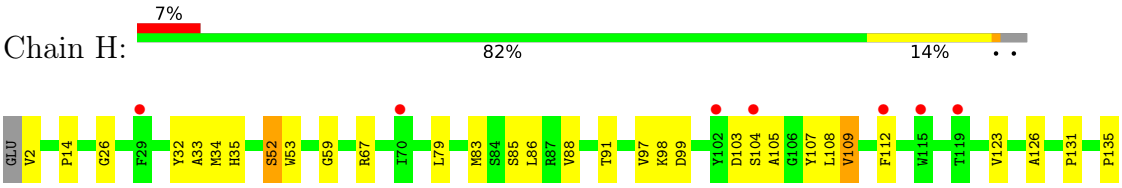
• Molecule 4: BA.2-10 light chain

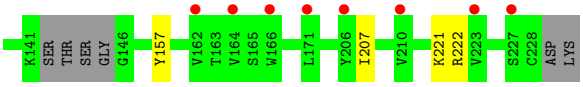
Chain B:



• Molecule 5: BA.2-10 heavy chain

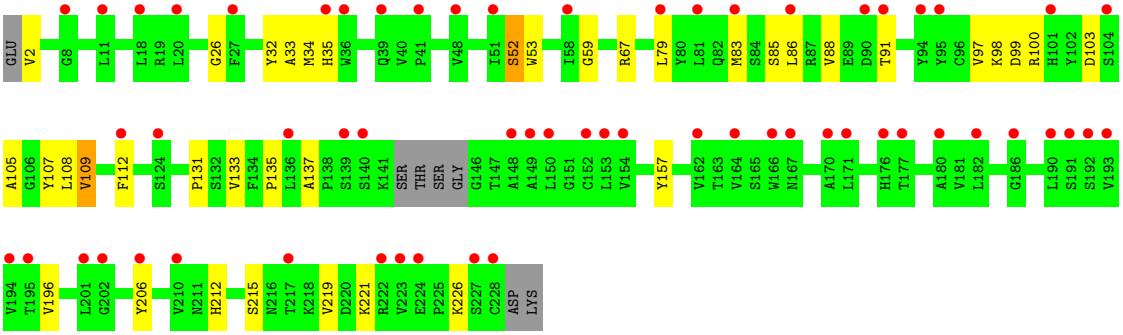
Chain H:





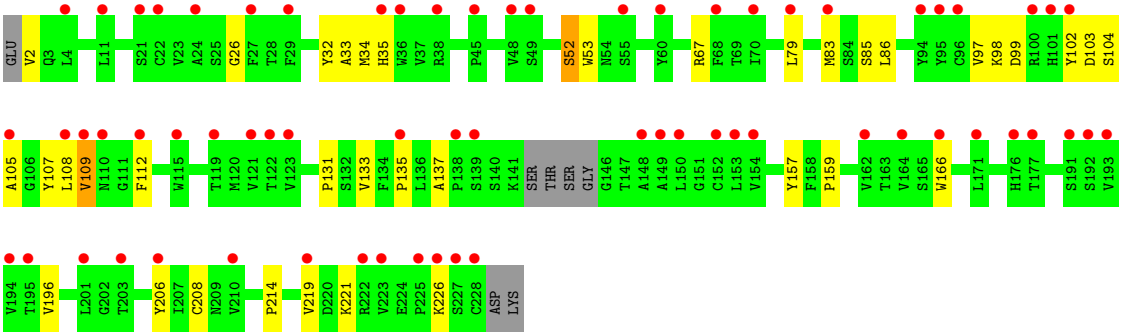
• Molecule 5: BA.2-10 heavy chain

Chain C:



• Molecule 5: BA.2-10 heavy chain

Chain A:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	170.66Å 171.69Å 176.99Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	55.79 – 3.58 56.89 – 3.58	Depositor EDS
% Data completeness (in resolution range)	89.6 (55.79-3.58) 90.0 (56.89-3.58)	Depositor EDS
R_{merge}	0.47	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.06 (at 3.57Å)	Xtriage
Refinement program	PHENIX 1.20.1_4487	Depositor
R, R_{free}	0.265 , 0.314 0.273 , 0.316	Depositor DCC
R_{free} test set	2853 reflections (5.13%)	wwPDB-VP
Wilson B-factor (Å^2)	126.7	Xtriage
Anisotropy	0.520	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 107.4	EDS
L-test for twinning[2]	$< \|L\| > = 0.30$, $< L^2 > = 0.15$	Xtriage
Estimated twinning fraction	0.218 for -h,l,k 0.109 for -l,-k,-h 0.105 for k,h,-l 0.098 for k,l,h 0.098 for l,h,k	Xtriage
F_o,F_c correlation	0.90	EDS
Total number of atoms	24354	wwPDB-VP
Average B, all atoms (Å^2)	173.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.64% 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	E	0.24	0/1679	0.48	0/2288
1	J	0.25	0/1679	0.47	0/2288
1	M	0.24	0/1679	0.48	0/2288
2	F	0.26	0/1651	0.50	0/2241
2	K	0.26	0/1651	0.50	0/2241
2	N	0.26	0/1651	0.50	0/2241
3	R	0.25	0/1593	0.46	0/2164
3	X	0.25	0/1593	0.47	0/2164
3	Y	0.25	0/1593	0.46	0/2164
4	B	0.26	0/1677	0.51	0/2281
4	D	0.27	0/1677	0.53	1/2281 (0.0%)
4	L	0.27	0/1677	0.52	0/2281
5	A	0.26	0/1716	0.50	0/2337
5	C	0.26	0/1716	0.51	0/2337
5	H	0.27	0/1716	0.50	0/2337
All	All	0.26	0/24948	0.49	1/33933 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	D	126	LYS	CD-CE-NZ	5.10	123.43	111.70

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	E	1637	0	1591	19	0
1	J	1637	0	1591	21	0
1	M	1637	0	1591	22	0
2	F	1618	0	1582	15	0
2	K	1618	0	1582	15	0
2	N	1618	0	1582	11	0
3	R	1547	0	1469	13	0
3	X	1547	0	1469	16	0
3	Y	1547	0	1469	14	0
4	B	1642	0	1596	21	0
4	D	1642	0	1596	27	1
4	L	1642	0	1596	23	1
5	A	1674	0	1621	24	0
5	C	1674	0	1621	21	0
5	H	1674	0	1621	19	0
All	All	24354	0	23577	246	1

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 5.

The worst 5 of 246 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)
4:D:79:GLN:NE2	3:X:391:CYS:SG	2.35	1.00
4:D:79:GLN:NE2	3:X:525:CYS:SG	2.49	0.85
3:Y:525:CYS:SG	4:B:79:GLN:NE2	2.49	0.85
3:Y:391:CYS:SG	4:B:79:GLN:NE2	2.59	0.76
5:A:131:PRO:HB3	5:A:157:TYR:HB3	1.72	0.72

All (1) 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)
4:L:106:ILE:O	4:D:126:LYS:NZ[3_454]	1.99	0.21

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	E	211/226 (93%)	201 (95%)	10 (5%)	0	100	100
1	J	211/226 (93%)	201 (95%)	10 (5%)	0	100	100
1	M	211/226 (93%)	200 (95%)	11 (5%)	0	100	100
2	F	210/215 (98%)	204 (97%)	6 (3%)	0	100	100
2	K	210/215 (98%)	203 (97%)	7 (3%)	0	100	100
2	N	210/215 (98%)	204 (97%)	6 (3%)	0	100	100
3	R	193/202 (96%)	186 (96%)	7 (4%)	0	100	100
3	X	193/202 (96%)	186 (96%)	7 (4%)	0	100	100
3	Y	193/202 (96%)	186 (96%)	7 (4%)	0	100	100
4	B	211/214 (99%)	202 (96%)	9 (4%)	0	100	100
4	D	211/214 (99%)	202 (96%)	9 (4%)	0	100	100
4	L	211/214 (99%)	202 (96%)	9 (4%)	0	100	100
5	A	219/230 (95%)	200 (91%)	16 (7%)	3 (1%)	11	48
5	C	219/230 (95%)	201 (92%)	15 (7%)	3 (1%)	11	48
5	H	219/230 (95%)	201 (92%)	14 (6%)	4 (2%)	8	43
All	All	3132/3261 (96%)	2979 (95%)	143 (5%)	10 (0%)	41	74

5 of 10 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
5	H	52	SER
5	H	105	ALA
5	C	52	SER
5	C	53	TRP
5	C	105	ALA

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	E	181/191 (95%)	181 (100%)	0	100 100
1	J	181/191 (95%)	181 (100%)	0	100 100
1	M	181/191 (95%)	181 (100%)	0	100 100
2	F	186/188 (99%)	186 (100%)	0	100 100
2	K	186/188 (99%)	185 (100%)	1 (0%)	88 95
2	N	186/188 (99%)	186 (100%)	0	100 100
3	R	167/175 (95%)	167 (100%)	0	100 100
3	X	167/175 (95%)	167 (100%)	0	100 100
3	Y	167/175 (95%)	167 (100%)	0	100 100
4	B	189/190 (100%)	188 (100%)	1 (0%)	88 95
4	D	189/190 (100%)	186 (98%)	3 (2%)	62 83
4	L	189/190 (100%)	188 (100%)	1 (0%)	88 95
5	A	187/195 (96%)	186 (100%)	1 (0%)	88 95
5	C	187/195 (96%)	185 (99%)	2 (1%)	73 88
5	H	187/195 (96%)	185 (99%)	2 (1%)	73 88
All	All	2730/2817 (97%)	2719 (100%)	11 (0%)	91 97

5 of 11 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
5	C	107	TYR
5	C	109	VAL
5	A	109	VAL
4	B	94	TYR
4	D	32	TYR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
2	N	148	GLN
3	R	501	ASN
4	L	38	GLN
5	H	39	GLN
3	Y	501	ASN

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, 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	E	215/226 (95%)	0.37	18 (8%) 11 6	115, 164, 214, 296	0
1	J	215/226 (95%)	0.44	17 (7%) 12 6	125, 168, 216, 285	0
1	M	215/226 (95%)	0.25	6 (2%) 53 35	127, 158, 212, 270	0
2	F	212/215 (98%)	0.53	15 (7%) 16 8	106, 153, 197, 252	0
2	K	212/215 (98%)	0.86	29 (13%) 3 2	124, 172, 229, 312	0
2	N	212/215 (98%)	0.58	21 (9%) 7 4	112, 150, 192, 249	0
3	R	194/202 (96%)	0.63	25 (12%) 3 2	119, 163, 244, 339	0
3	X	194/202 (96%)	0.63	17 (8%) 10 5	123, 166, 228, 366	0
3	Y	194/202 (96%)	0.82	25 (12%) 3 2	113, 159, 230, 310	0
4	B	213/214 (99%)	0.81	37 (17%) 1 0	116, 174, 246, 335	0
4	D	213/214 (99%)	0.64	29 (13%) 3 2	141, 194, 247, 292	0
4	L	213/214 (99%)	0.53	12 (5%) 24 13	122, 170, 207, 265	0
5	A	223/230 (96%)	1.57	66 (29%) 0 0	137, 187, 254, 324	0
5	C	223/230 (96%)	1.19	60 (26%) 0 0	139, 206, 270, 337	0
5	H	223/230 (96%)	0.43	15 (6%) 17 9	105, 160, 217, 284	0
All	All	3171/3261 (97%)	0.69	392 (12%) 4 2	105, 168, 236, 366	0

The worst 5 of 392 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
5	A	149	ALA	9.3
5	C	150	LEU	7.8
5	A	222	ARG	7.7
5	A	194	VAL	7.6
5	C	149	ALA	7.4

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.