WORLDWIDE wwPDB PROTEIN DATA BANK

Full wwPDB X-ray Structure Validation Report ⓘ

Oct 11, 2023 – 03:37 PM EDT

PDB ID	:	7RK1
Title	:	Crystal structure of the human astrovirus serotype 8 capsid spike in complex with scFv 3E8, an astrovirus-neutralizing antibody, at 2.05-A resolution
Authors	:	Meyer, L.; DuBois, R.M.
Deposited on	:	2021-07-21
Resolution	:	2.05 Å(reported)

The following versions of software and data (see references ⓘ) were used in the production of this report:

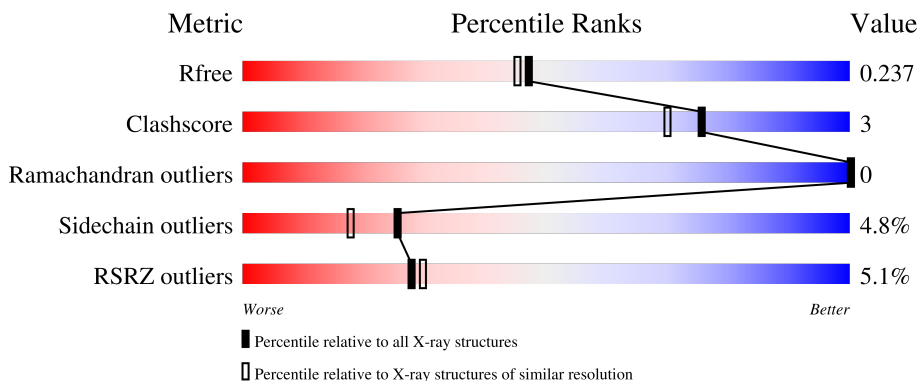|		
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation 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 2.05 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	Percentile Ranks	Value
Rfree		0.237
Clashscore		3
Ramachandran outliers		0
Sidechain outliers		4.8%
RSRZ outliers		5.1%

Worse   *Better*

▌ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1692 (2.04-2.04)
Clashscore	141614	1773 (2.04-2.04)
Ramachandran outliers	138981	1752 (2.04-2.04)
Sidechain outliers	138945	1752 (2.04-2.04)
RSRZ outliers	127900	1672 (2.04-2.04)

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	230	3% 86% 8% 5%
1	B	230	6% 85% 10% 5%
2	C	248	% 76% 15% 9%
2	D	248	9% 80% 8% • 11%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 7378 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 Capsid polyprotein VP70.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	218	Total 1756	C 1122	N 297	O 331	S 6	0	1	0
1	B	218	Total 1761	C 1125	N 299	O 331	S 6	0	2	0

There are 22 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	427	MET	-	initiating methionine	UNP Q9IFX1
A	428	GLY	-	expression tag	UNP Q9IFX1
A	648	ALA	-	expression tag	UNP Q9IFX1
A	649	ALA	-	expression tag	UNP Q9IFX1
A	650	GLU	-	expression tag	UNP Q9IFX1
A	651	LEU	-	expression tag	UNP Q9IFX1
A	652	ALA	-	expression tag	UNP Q9IFX1
A	653	LEU	-	expression tag	UNP Q9IFX1
A	654	VAL	-	expression tag	UNP Q9IFX1
A	655	PRO	-	expression tag	UNP Q9IFX1
A	656	ARG	-	expression tag	UNP Q9IFX1
B	427	MET	-	initiating methionine	UNP Q9IFX1
B	428	GLY	-	expression tag	UNP Q9IFX1
B	648	ALA	-	expression tag	UNP Q9IFX1
B	649	ALA	-	expression tag	UNP Q9IFX1
B	650	GLU	-	expression tag	UNP Q9IFX1
B	651	LEU	-	expression tag	UNP Q9IFX1
B	652	ALA	-	expression tag	UNP Q9IFX1
B	653	LEU	-	expression tag	UNP Q9IFX1
B	654	VAL	-	expression tag	UNP Q9IFX1
B	655	PRO	-	expression tag	UNP Q9IFX1
B	656	ARG	-	expression tag	UNP Q9IFX1

- Molecule 2 is a protein called scFv 3E8.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	225	Total	C	N	O	S	0	1	0
			1783	1128	295	353	7			
2	D	221	Total	C	N	O	S	0	0	0
			1743	1105	283	348	7			

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	101	Total	O	0	0
			101	101		
3	C	98	Total	O	0	0
			98	98		
3	B	77	Total	O	0	0
			77	77		
3	D	59	Total	O	0	0
			59	59		

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.

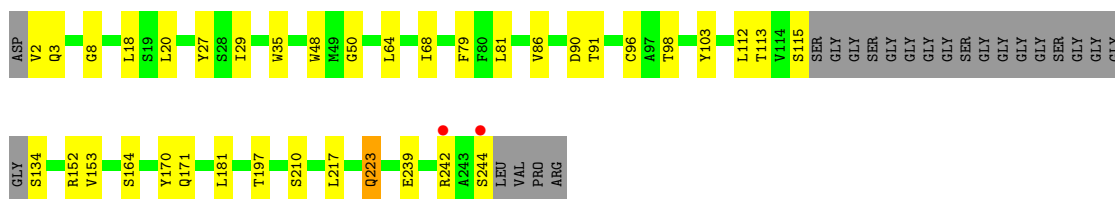• Molecule 1: Capsid polyprotein VP70
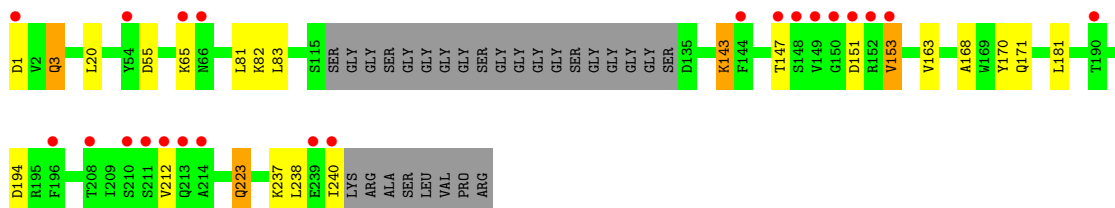
Chain A:
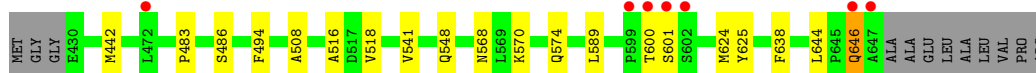
• Molecule 1: Capsid polyprotein VP70

Chain B:

• Molecule 2: scFv 3E8

Chain C:

• Molecule 2: scFv 3E8

Chain D:

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	48.83Å　　80.43Å　　117.21Å 90.00°　　90.67°　　　90.00°	Depositor
Resolution (Å)	66.32　–　2.05 66.32　–　2.05	Depositor EDS
% Data completeness (in resolution range)	97.6 (66.32-2.05) 97.6 (66.32-2.05)	Depositor EDS
R_{merge}	0.18	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.21 (at 2.05Å)	Xtriage
Refinement program	PHENIX 1.19_4092	Depositor
R, R_{free}	0.223　,　0.238 0.222　,　0.237	Depositor DCC
R_{free} test set	1992 reflections (3.58%)	wwPDB-VP
Wilson B-factor (Å^2)	25.3	Xtriage
Anisotropy	1.076	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 44.9	EDS
L-test for twinning[2]	$< |L| > = 0.50, < L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.053 for h,-k,-l	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	7378	wwPDB-VP
Average B, all atoms (Å^2)	32.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.34% 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	A	0.54	0/1809	0.74	0/2478
1	B	0.56	0/1814	0.77	0/2485
2	C	0.57	0/1829	0.82	0/2487
2	D	0.56	0/1789	0.78	0/2436
All	All	0.56	0/7241	0.78	0/9886

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	1756	0	1697	10	0
1	B	1761	0	1703	8	0
2	C	1783	0	1697	15	0
2	D	1743	0	1651	11	0
3	A	101	0	0	2	0
3	B	77	0	0	0	0
3	C	98	0	0	0	0
3	D	59	0	0	0	0
All	All	7378	0	6748	43	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 3.

All (43) 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:C:170:TYR:HE1	2:C:223:GLN:HG2	1.58	0.69
1:B:442:MET:HG3	1:B:483:PRO:HG3	1.76	0.67
2:D:170:TYR:HE1	2:D:223:GLN:HG2	1.61	0.66
2:D:147:THR:HG21	2:D:153:VAL:CG2	2.27	0.65
1:A:442:MET:HG3	1:A:483:PRO:HG3	1.80	0.64
1:A:624:MET:HB3	1:B:624:MET:HB3	1.83	0.61
2:C:152[A]:ARG:HG3	2:C:210:SER:HA	1.83	0.60
2:C:20:LEU:HD12	2:C:81:LEU:HD23	1.87	0.56
2:C:18:LEU:HD22	2:C:112:LEU:HD22	1.87	0.56
2:D:147:THR:HG21	2:D:153:VAL:HG23	1.87	0.55
1:B:444:PRO:HG3	1:B:578:THR:HG21	1.89	0.53
2:C:2:VAL:HG22	2:C:27:TYR:HB2	1.91	0.53
2:D:20:LEU:HD12	2:D:81:LEU:HD23	1.91	0.52
1:A:574:GLN:HG2	3:A:788:HOH:O	2.11	0.50
1:B:525:LYS:HG2	1:B:535:ILE:HG23	1.97	0.47
2:D:171:GLN:HB2	2:D:181:LEU:HD11	1.96	0.47
1:A:508:ALA:HB3	1:A:516:ALA:HB3	1.97	0.46
2:C:171:GLN:HB2	2:C:181:LEU:HD11	1.97	0.46
1:B:516:ALA:HB2	1:B:543:VAL:HA	1.97	0.46
2:C:91:THR:HG23	2:C:113:THR:HA	1.96	0.46
2:C:98:THR:O	2:C:103:TYR:HA	2.16	0.46
2:D:153:VAL:HG21	2:D:238:LEU:HD21	1.97	0.46
1:A:518:VAL:HG22	1:A:541:VAL:HG12	1.98	0.46
2:C:8:GLY:HA3	2:C:20:LEU:CD2	2.46	0.45
2:D:147:THR:HG21	2:D:153:VAL:HG22	1.98	0.45
2:C:29:ILE:HA	2:C:35:TRP:CZ2	2.51	0.45
1:A:625:TYR:CE1	1:A:638:PHE:HZ	2.35	0.45
2:C:64:LEU:O	2:C:68:ILE:HG22	2.17	0.45
2:C:8:GLY:HA3	2:C:20:LEU:HD23	1.99	0.44
1:A:548:GLN:NE2	3:A:701:HOH:O	2.49	0.44
2:D:3:GLN:HE21	2:D:3:GLN:HB2	1.54	0.43
1:B:570:LYS:HE3	1:B:614:ASN:HA	1.99	0.43
2:C:86:VAL:HG13	2:C:90:ASP:HB2	2.01	0.43
2:D:168:ALA:HB3	2:D:223:GLN:HG2	2.00	0.43
2:C:79:PHE:CZ	2:C:96:CYS:HB2	2.53	0.42
1:A:589:LEU:HD23	1:A:589:LEU:HA	1.89	0.42
1:B:528:HIS:NE2	1:B:618[B]:ARG:HD3	2.36	0.41
1:A:568:ASN:HB2	1:A:570:LYS:HE2	2.02	0.41
1:B:440:PRO:O	1:B:483:PRO:HG2	2.20	0.41
2:D:151:ASP:O	2:D:212:VAL:HG23	2.20	0.41

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:C:48:TRP:CZ2	2:C:50:GLY:HA2	2.56	0.40
2:D:143:LYS:HD2	2:D:143:LYS:H	1.86	0.40
1:A:646:GLN:H	1:A:646:GLN:HG2	1.43	0.40

There are no symmetry-related clashes.

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	217/230 (94%)	213 (98%)	4 (2%)	0	100	100
1	B	218/230 (95%)	215 (99%)	3 (1%)	0	100	100
2	C	222/248 (90%)	216 (97%)	6 (3%)	0	100	100
2	D	217/248 (88%)	208 (96%)	9 (4%)	0	100	100
All	All	874/956 (91%)	852 (98%)	22 (2%)	0	100	100

There are no Ramachandran outliers to report.

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	198/204 (97%)	192 (97%)	6 (3%)	41	34
1	B	198/204 (97%)	190 (96%)	8 (4%)	31	24

Continued on next page...

W O R L D W I D E
wwPDB
PROTEIN DATA BANK

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	C	199/207 (96%)	188 (94%)	11 (6%)	21	13
2	D	195/207 (94%)	182 (93%)	13 (7%)	16	9
All	All	790/822 (96%)	752 (95%)	38 (5%)	25	18

All (38) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	486	SER
1	A	494	PHE
1	A	600	THR
1	A	601	SER
1	A	644	LEU
1	A	646	GLN
2	C	3	GLN
2	C	115	SER
2	C	134	SER
2	C	153	VAL
2	C	164	SER
2	C	197	THR
2	C	217	LEU
2	C	223	GLN
2	C	239	GLU
2	C	242	ARG
2	C	244	SER
1	B	472	LEU
1	B	486	SER
1	B	498	ASP
1	B	514	LYS
1	B	530	ASN
1	B	600	THR
1	B	601	SER
1	B	646	GLN
2	D	1	ASP
2	D	3	GLN
2	D	55	ASP
2	D	65	LYS
2	D	82	LYS
2	D	83	LEU
2	D	143	LYS
2	D	153	VAL
2	D	163	VAL

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	D	194	ASP
2	D	223	GLN
2	D	237	LYS
2	D	240	ILE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (8) such sidechains are listed below:

Mol	Chain	Res	Type
2	C	3	GLN
2	C	223	GLN
1	B	548	GLN
1	B	584	GLN
2	D	3	GLN
2	D	40	GLN
2	D	59	ASN
2	D	172	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	218/230 (94%)	0.34	7 (3%)	47	52	18, 26, 47, 67	0
1	B	218/230 (94%)	0.58	14 (6%)	19	21	19, 31, 52, 68	0
2	C	225/248 (90%)	0.36	2 (0%)	84	86	20, 29, 41, 51	0
2	D	221/248 (89%)	0.82	22 (9%)	7	7	22, 35, 60, 67	0
All	All	882/956 (92%)	0.52	45 (5%)	28	30	18, 30, 54, 68	0

All (45) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	600	THR	10.0
1	B	647	ALA	6.6
2	D	210	SER	5.3
2	D	1	ASP	5.2
1	A	601	SER	4.8
1	B	646	GLN	4.7
1	B	601	SER	4.6
2	D	214	ALA	4.4
2	D	240	ILE	4.4
2	D	150	GLY	4.1
2	D	144	PHE	4.1
2	D	65	LYS	4.1
2	D	151	ASP	3.8
1	B	531	ASN	3.7
1	B	600	THR	3.6
2	D	152	ARG	3.6
2	D	211	SER	3.5
2	D	153	VAL	3.4
1	B	599	PRO	3.4
2	C	242	ARG	3.4
2	D	212	VAL	3.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	D	147	THR	3.2
1	A	602	SER	3.2
1	B	648	ALA	3.2
2	D	196	PHE	3.2
2	D	213	GLN	3.1
2	D	148	SER	3.0
1	A	646	GLN	3.0
2	D	239	GLU	2.9
1	B	504	PRO	2.9
1	B	498	ASP	2.7
1	A	647	ALA	2.6
1	B	598	GLY	2.5
1	A	472	LEU	2.5
2	D	190	THR	2.5
1	A	599	PRO	2.4
2	D	149	VAL	2.4
2	D	66	ASN	2.4
2	D	54	TYR	2.4
2	C	244	SER	2.3
1	B	495	GLN	2.3
1	B	501	SER	2.3
2	D	208	THR	2.2
1	B	530	ASN	2.2
1	B	597	ASP	2.0

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.