



# Full wwPDB X-ray Structure Validation Report ⓘ

Nov 22, 2023 – 09:22 PM JST

PDB ID : 7WN2  
Title : Crystal structure of SARS-CoV-2 spike receptor-binding domain (RBD) in complex with NCV2SG53 Fab  
Authors : Yamamoto, A.; Higashiura, A.  
Deposited on : 2022-01-17  
Resolution : 2.35 Å (reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto: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  
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

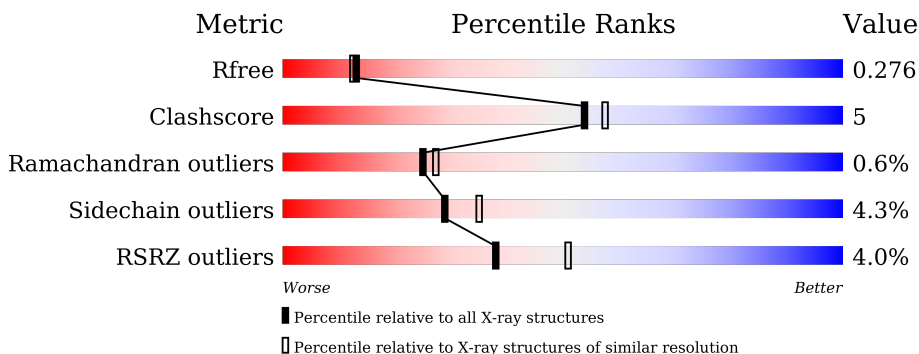
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.35 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	1164 (2.36-2.36)
Clashscore	141614	1232 (2.36-2.36)
Ramachandran outliers	138981	1211 (2.36-2.36)
Sidechain outliers	138945	1212 (2.36-2.36)
RSRZ outliers	127900	1150 (2.36-2.36)

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  $\geq 3$ , 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	224	 11% 73% 15% 12%
1	R	224	 5% 68% 18% 12%
2	B	227	 3% 81% 15% . .
2	H	227	 % 81% 13% . .
3	C	214	 % 88% 10% .
3	L	214	 2% 90% 8% .

## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 9987 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 Spike protein S1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	R	196	1552	995	259	290	8	13	0	0
1	A	197	1561	1001	261	291	8	11	0	0

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
R	537	THR	-	expression tag	UNP P0DTC2
R	538	GLY	-	expression tag	UNP P0DTC2
R	539	HIS	-	expression tag	UNP P0DTC2
R	540	HIS	-	expression tag	UNP P0DTC2
R	541	HIS	-	expression tag	UNP P0DTC2
R	542	HIS	-	expression tag	UNP P0DTC2
R	543	HIS	-	expression tag	UNP P0DTC2
R	544	HIS	-	expression tag	UNP P0DTC2
A	537	THR	-	expression tag	UNP P0DTC2
A	538	GLY	-	expression tag	UNP P0DTC2
A	539	HIS	-	expression tag	UNP P0DTC2
A	540	HIS	-	expression tag	UNP P0DTC2
A	541	HIS	-	expression tag	UNP P0DTC2
A	542	HIS	-	expression tag	UNP P0DTC2
A	543	HIS	-	expression tag	UNP P0DTC2
A	544	HIS	-	expression tag	UNP P0DTC2

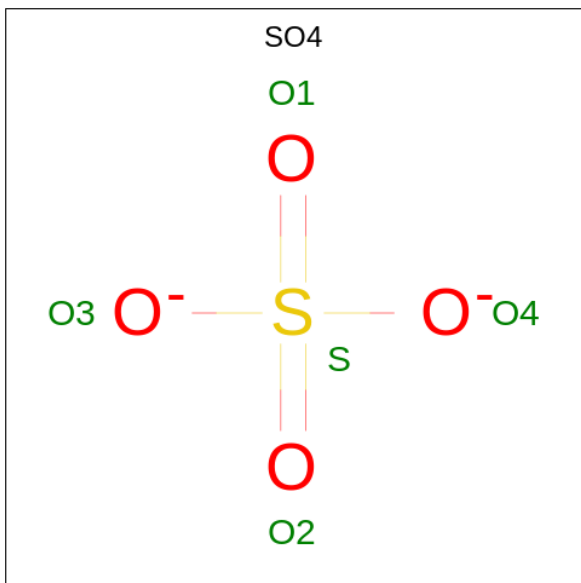
- Molecule 2 is a protein called Fab Heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	H	218	1644	1042	274	319	9	0	1	0
2	B	223	1671	1057	279	326	9	2	0	0

- Molecule 3 is a protein called Fab Light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	L	213	Total 1627	C 1021	N 273	O 329	S 4	0	1	0
3	C	213	Total 1625	C 1020	N 270	O 331	S 4	0	1	0

- Molecule 4 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	O	S		
4	R	1	Total 5	O 4	S 1	0	0
4	R	1	Total 5	O 4	S 1	0	0
4	H	1	Total 5	O 4	S 1	0	0
4	A	1	Total 5	O 4	S 1	0	0
4	B	1	Total 5	O 4	S 1	0	0
4	C	1	Total 5	O 4	S 1	0	0
4	C	1	Total 5	O 4	S 1	0	0

- Molecule 5 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: C<sub>8</sub>H<sub>15</sub>NO<sub>6</sub>).



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

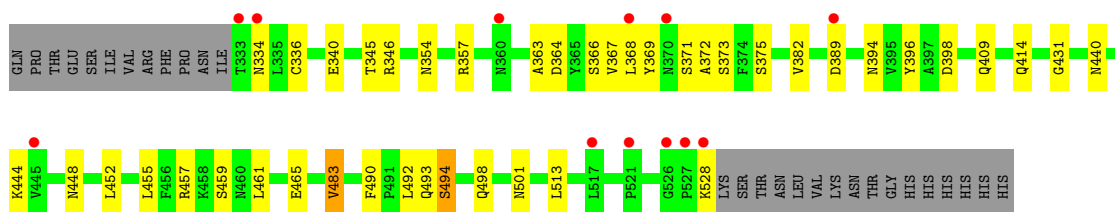
- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	R	33	Total	O	0	0
			33	33		
6	H	33	Total	O	0	0
			33	33		
6	L	57	Total	O	0	0
			57	57		
6	A	26	Total	O	0	0
			26	26		
6	B	34	Total	O	0	0
			34	34		
6	C	61	Total	O	0	0
			61	61		

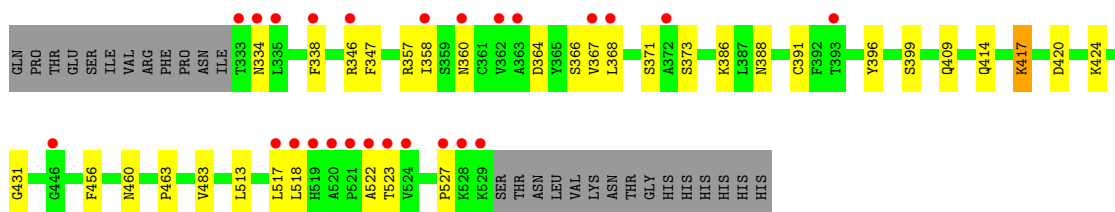
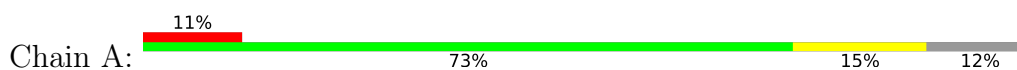
### 3 Residue-property plots [i](#)

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.

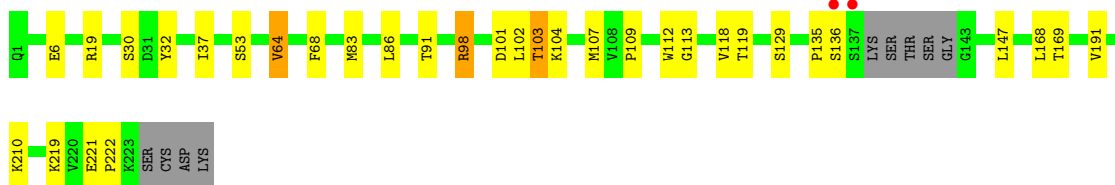
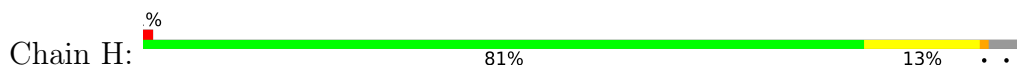
- Molecule 1: Spike protein S1



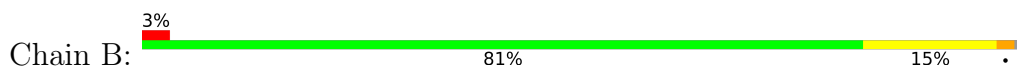
- Molecule 1: Spike protein S1

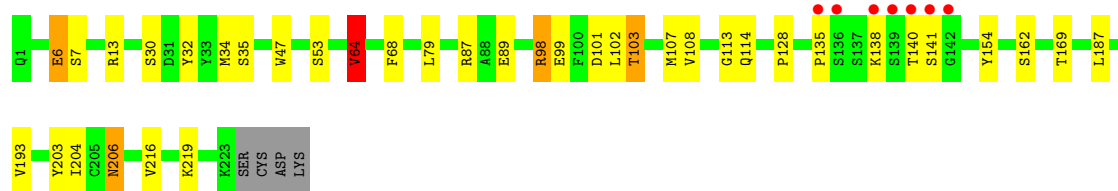


- Molecule 2: Fab Heavy chain

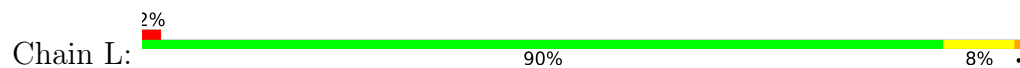


- Molecule 2: Fab Heavy chain

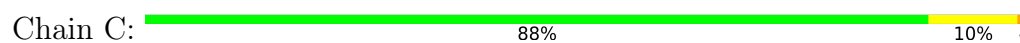




- Molecule 3: Fab Light chain



- Molecule 3: Fab Light chain



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	76.27Å 118.19Å 196.29Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	46.71 – 2.35 46.71 – 2.35	Depositor EDS
% Data completeness (in resolution range)	99.2 (46.71-2.35) 99.2 (46.71-2.35)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.39 (at 2.34Å)	Xtrriage
Refinement program	PHENIX (1.19.1_4122: ???)	Depositor
R, $R_{free}$	0.228 , 0.279 0.227 , 0.276	Depositor DCC
$R_{free}$ test set	3688 reflections (4.98%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	44.7	Xtrriage
Anisotropy	0.874	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 34.6	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.51$ , $\langle L^2 \rangle = 0.35$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	9987	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	54.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 43.26 % 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 1.8090e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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, SO4

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	0.46	0/1605	0.62	0/2183
1	R	0.45	0/1596	0.62	0/2172
2	B	0.43	0/1711	0.63	0/2326
2	H	0.45	0/1686	0.65	0/2292
3	C	0.45	0/1662	0.64	0/2258
3	L	0.45	0/1664	0.65	0/2260
All	All	0.45	0/9924	0.64	0/13491

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	1561	0	1485	16	0
1	R	1552	0	1472	25	0
2	B	1671	0	1651	20	0
2	H	1644	0	1623	17	0
3	C	1625	0	1594	16	0
3	L	1627	0	1601	15	0
4	A	5	0	0	0	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	B	5	0	0	0	0
4	C	10	0	0	0	0
4	H	5	0	0	0	0
4	R	10	0	0	0	0
5	A	14	0	13	0	0
5	R	14	0	13	0	0
6	A	26	0	0	1	0
6	B	34	0	0	0	0
6	C	61	0	0	1	0
6	H	33	0	0	0	0
6	L	57	0	0	1	0
6	R	33	0	0	1	0
All	All	9987	0	9452	101	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 5.

All (101) 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:409:GLN:HA	1:A:414:GLN:HG2	1.58	0.85
1:A:417:LYS:H	1:A:417:LYS:HE2	1.50	0.76
1:R:409:GLN:HA	1:R:414:GLN:HG2	1.72	0.71
1:A:391:CYS:HB3	1:A:522:ALA:HB1	1.72	0.71
2:H:107:MET:HG2	3:L:50:ALA:HB2	1.75	0.67
1:R:444:LYS:HG3	1:R:448:ASN:HB2	1.79	0.63
1:R:457:ARG:NH1	1:R:459:SER:O	2.28	0.61
1:R:440:ASN:ND2	1:R:440:ASN:H	2.00	0.60
3:L:18:ARG:HG2	3:L:18:ARG:HH11	1.67	0.59
2:B:64:VAL:HG13	2:B:68:PHE:HB2	1.85	0.58
2:H:135:PRO:HB3	2:H:147:LEU:HB3	1.86	0.57
1:R:364:ASP:O	1:R:367:VAL:HG12	2.06	0.56
1:R:440:ASN:H	1:R:440:ASN:HD22	1.52	0.56
2:B:204:ILE:HD13	2:B:219:LYS:HA	1.87	0.56
2:H:83:MET:HB3	2:H:86:LEU:HD21	1.89	0.55
1:R:389:ASP:HB3	1:R:528:LYS:HD3	1.88	0.55
1:A:424:LYS:HB3	1:A:463:PRO:HA	1.89	0.55
3:C:108:ARG:HG2	3:C:109:THR:N	2.21	0.55
1:R:357:ARG:HG3	1:R:396:TYR:HE1	1.73	0.54
3:L:61:ARG:NH2	3:L:82:ASP:OD1	2.40	0.54
2:H:91:THR:HG23	2:H:119:THR:HA	1.90	0.54

Continued on next page...

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:C:90:GLN:OE1	3:C:92:LYS:N	2.41	0.53
1:A:388:ASN:HB3	1:A:527:PRO:HD2	1.91	0.53
2:H:32:TYR:CD1	2:H:98:ARG:HD2	2.45	0.53
3:C:119:PRO:HB3	3:C:209:PHE:CZ	2.43	0.52
2:H:101:ASP:OD1	2:H:103:THR:HB	2.10	0.52
1:R:371:SER:O	1:R:373:SER:N	2.32	0.51
2:B:30:SER:O	2:B:53:SER:HB3	2.10	0.51
2:B:98:ARG:O	2:B:98:ARG:HG3	2.10	0.51
3:L:18:ARG:HG2	3:L:18:ARG:NH1	2.24	0.51
1:R:357:ARG:HH21	1:R:394:ASN:ND2	2.08	0.51
3:C:35:TRP:CZ3	3:C:88:CYS:HB3	2.46	0.50
3:L:90:GLN:OE1	3:L:92:LYS:N	2.44	0.50
2:H:6:GLU:OE2	2:H:113:GLY:HA3	2.12	0.50
2:H:135:PRO:HD2	2:H:222:PRO:HA	1.94	0.50
3:C:183:LYS:O	3:C:187:GLU:HG2	2.11	0.50
1:R:336:CYS:SG	1:R:363:ALA:HB2	2.52	0.50
1:R:490:PHE:CE2	1:R:492:LEU:HB2	2.46	0.50
1:R:357:ARG:HH21	1:R:394:ASN:HD21	1.59	0.49
2:B:87:ARG:HB3	2:B:89:GLU:OE1	2.12	0.49
1:A:456:PHE:HE2	2:B:102:LEU:HD11	1.78	0.49
2:B:107:MET:HE1	3:C:32:TYR:HA	1.95	0.48
2:H:219:LYS:HE3	2:H:221:GLU:OE1	2.14	0.48
2:B:128:PRO:HB3	2:B:154:TYR:HB3	1.96	0.47
1:A:364:ASP:OD2	1:A:367:VAL:HG23	2.15	0.47
3:C:163:VAL:HG22	3:C:175:LEU:HD12	1.96	0.47
1:A:460:ASN:N	6:A:706:HOH:O	2.47	0.47
1:R:357:ARG:HE	1:R:394:ASN:ND2	2.14	0.46
3:L:142[A]:ARG:CZ	3:L:163:VAL:HG21	2.44	0.46
1:A:431:GLY:HA3	1:A:513:LEU:O	2.15	0.46
1:A:368:LEU:HD12	1:A:368:LEU:HA	1.67	0.46
1:R:452:LEU:HD23	1:R:494:SER:HA	1.97	0.46
2:H:109:PRO:HD3	3:L:49:TYR:CE2	2.50	0.46
2:B:32:TYR:CD1	2:B:98:ARG:HD2	2.50	0.46
2:H:103:THR:HG22	2:H:104:LYS:HG3	1.98	0.46
2:B:6:GLU:OE2	2:B:113:GLY:HA3	2.17	0.45
2:B:107:MET:CE	3:C:32:TYR:HA	2.46	0.45
3:C:142:ARG:CZ	3:C:163:VAL:HG21	2.46	0.45
1:R:431:GLY:HA3	1:R:513:LEU:O	2.16	0.45
2:B:162:SER:OG	2:B:206:ASN:OD1	2.35	0.45
2:H:168:LEU:HD21	2:H:191:VAL:HG11	1.99	0.45
2:B:6:GLU:H	2:B:114:GLN:HE22	1.63	0.45

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:338:PHE:HE1	1:A:358:ILE:HG13	1.81	0.45
2:H:64:VAL:HG13	2:H:68:PHE:HB2	1.97	0.44
1:R:354:ASN:O	1:R:398:ASP:HA	2.18	0.44
1:R:357:ARG:HG3	1:R:396:TYR:CE1	2.52	0.44
1:A:357:ARG:HD2	1:A:396:TYR:HE1	1.83	0.44
2:B:89:GLU:H	2:B:89:GLU:CD	2.21	0.44
1:R:340:GLU:HB3	6:R:714:HOH:O	2.17	0.43
1:A:420:ASP:HA	1:A:424:LYS:NZ	2.33	0.43
1:R:498:GLN:HB2	1:R:501:ASN:OD1	2.18	0.43
3:L:30:SER:HB3	6:L:353:HOH:O	2.17	0.43
3:L:54:LEU:HD12	3:L:54:LEU:HA	1.69	0.43
1:A:360:ASN:HA	1:A:523:THR:HG22	1.99	0.43
1:A:456:PHE:CE2	2:B:102:LEU:HD11	2.53	0.43
2:B:35:SER:HB2	2:B:99:GLU:OE2	2.19	0.43
3:C:54:LEU:HA	3:C:54:LEU:HD12	1.82	0.43
3:L:35:TRP:CZ3	3:L:88:CYS:HB3	2.54	0.43
3:L:125:LEU:HA	3:L:125:LEU:HD23	1.87	0.43
3:L:163:VAL:HG22	3:L:175:LEU:HD12	2.01	0.43
2:H:30:SER:O	2:H:53:SER:HB3	2.18	0.43
3:L:116:PHE:CD1	3:L:135:LEU:HD23	2.54	0.43
2:B:34:MET:HE2	2:B:79:LEU:HD22	2.01	0.42
1:R:461:LEU:HD22	1:R:465:GLU:HB3	2.01	0.42
1:A:347:PHE:CE2	1:A:399:SER:HB2	2.54	0.42
2:B:101:ASP:OD1	2:B:103:THR:HB	2.19	0.42
3:L:120:PRO:HD3	3:L:132:VAL:HG22	2.02	0.42
3:C:119:PRO:HB3	3:C:209:PHE:CE1	2.55	0.42
3:C:147:GLN:HG2	3:C:154:LEU:HD13	2.01	0.41
3:C:207:LYS:HE2	6:C:435:HOH:O	2.19	0.41
3:C:154:LEU:HD12	3:C:155:GLN:N	2.35	0.41
2:H:68:PHE:CE2	2:H:83:MET:HG2	2.55	0.41
2:B:47:TRP:CZ3	3:C:95:PRO:HB3	2.55	0.41
1:R:368:LEU:HD23	1:R:368:LEU:HA	1.82	0.41
1:R:455:LEU:HB2	1:R:493:GLN:HG2	2.03	0.41
1:R:483:VAL:CG1	3:L:93:SER:HB2	2.51	0.41
2:B:193:VAL:HG21	2:B:203:TYR:CE1	2.55	0.41
1:R:366:SER:O	1:R:369:TYR:N	2.54	0.41
3:C:28:GLY:HA2	3:C:68:GLY:O	2.20	0.41
2:H:37:ILE:HD13	2:H:112:TRP:CZ2	2.56	0.40
2:H:68:PHE:CD1	2:H:68:PHE:N	2.89	0.40

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	195/224 (87%)	184 (94%)	11 (6%)	0	100	100
1	R	194/224 (87%)	178 (92%)	14 (7%)	2 (1%)	15	15
2	B	221/227 (97%)	206 (93%)	13 (6%)	2 (1%)	17	17
2	H	215/227 (95%)	204 (95%)	10 (5%)	1 (0%)	29	32
3	C	212/214 (99%)	202 (95%)	9 (4%)	1 (0%)	29	32
3	L	212/214 (99%)	203 (96%)	8 (4%)	1 (0%)	29	32
All	All	1249/1330 (94%)	1177 (94%)	65 (5%)	7 (1%)	25	27

All (7) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	R	372	ALA
2	H	136	SER
2	B	135	PRO
3	L	30	SER
1	R	382	VAL
2	B	64	VAL
3	C	29	ILE

### 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	170/196 (87%)	160 (94%)	10 (6%)	19	22

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	R	169/196 (86%)	163 (96%)	6 (4%)	35	43
2	B	189/193 (98%)	175 (93%)	14 (7%)	13	14
2	H	186/193 (96%)	177 (95%)	9 (5%)	25	30
3	C	185/185 (100%)	180 (97%)	5 (3%)	44	55
3	L	185/185 (100%)	182 (98%)	3 (2%)	62	75
All	All	1084/1148 (94%)	1037 (96%)	47 (4%)	29	35

All (47) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	R	334	ASN
1	R	345	THR
1	R	346	ARG
1	R	375	SER
1	R	483	VAL
1	R	494	SER
2	H	19	ARG
2	H	64	VAL
2	H	98	ARG
2	H	102	LEU
2	H	103	THR
2	H	118	VAL
2	H	129	SER
2	H	169	THR
2	H	210	LYS
3	L	90	GLN
3	L	93	SER
3	L	108	ARG
1	A	334	ASN
1	A	346	ARG
1	A	366	SER
1	A	371	SER
1	A	373	SER
1	A	386	LYS
1	A	417	LYS
1	A	483	VAL
1	A	517	LEU
1	A	518	LEU
2	B	6	GLU
2	B	7	SER

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
2	B	13	ARG
2	B	64	VAL
2	B	98	ARG
2	B	103	THR
2	B	108	VAL
2	B	138	LYS
2	B	140	THR
2	B	141	SER
2	B	169	THR
2	B	187	LEU
2	B	206	ASN
2	B	216	VAL
3	C	27	GLN
3	C	90	GLN
3	C	108	ARG
3	C	147	GLN
3	C	185	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	R	394	ASN
1	R	440	ASN
1	R	450	ASN
1	A	474	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

9 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
5	NAG	R	603	1	14,14,15	0.89	1 (7%)	17,19,21	0.71	0
4	SO4	R	602	-	4,4,4	0.17	0	6,6,6	0.18	0
4	SO4	A	601	-	4,4,4	0.17	0	6,6,6	0.18	0
4	SO4	C	301	-	4,4,4	0.13	0	6,6,6	0.18	0
4	SO4	H	301	-	4,4,4	0.10	0	6,6,6	0.25	0
4	SO4	R	601	-	4,4,4	0.10	0	6,6,6	0.23	0
4	SO4	B	301	-	4,4,4	0.13	0	6,6,6	0.17	0
5	NAG	A	602	1	14,14,15	0.54	0	17,19,21	0.89	1 (5%)
4	SO4	C	302	-	4,4,4	0.17	0	6,6,6	0.23	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. '2' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
5	NAG	R	603	1	-	0/6/23/26	0/1/1/1
5	NAG	A	602	1	-	0/6/23/26	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	R	603	NAG	O5-C1	2.97	1.48	1.43

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	A	602	NAG	C1-O5-C5	3.39	116.78	112.19

There are no chirality outliers.



There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 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<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	197/224 (87%)	0.67	25 (12%) 3 6	42, 57, 104, 140	3 (1%)
1	R	196/224 (87%)	0.45	12 (6%) 21 31	41, 54, 100, 149	3 (1%)
2	B	223/227 (98%)	0.18	7 (3%) 49 61	39, 50, 76, 146	1 (0%)
2	H	218/227 (96%)	0.23	2 (0%) 84 90	38, 50, 68, 97	0
3	C	213/214 (99%)	0.15	1 (0%) 91 95	37, 45, 61, 106	0
3	L	213/214 (99%)	0.26	4 (1%) 66 76	37, 45, 65, 91	0
All	All	1260/1330 (94%)	0.31	51 (4%) 38 51	37, 49, 88, 149	7 (0%)

All (51) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	362	VAL	4.8
1	A	528	LYS	4.5
1	A	358	ILE	4.3
1	A	522	ALA	4.2
1	R	517	LEU	4.2
1	R	528	LYS	4.1
2	H	137	SER	4.1
1	R	389	ASP	4.0
1	A	333	THR	3.8
1	R	521	PRO	3.8
1	A	520	ALA	3.7
1	A	517	LEU	3.6
1	A	518	LEU	3.4
1	R	527	PRO	3.4
2	B	136	SER	3.3
1	A	524	VAL	3.2
1	A	527	PRO	3.2
2	B	140	THR	3.1
2	B	138	LYS	3.1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	R	445	VAL	3.1
1	A	367	VAL	3.0
1	A	335	LEU	3.0
1	A	338	PHE	3.0
1	A	523	THR	3.0
1	R	526	GLY	2.9
1	A	519	HIS	2.9
2	B	142	GLY	2.9
1	A	334	ASN	2.8
1	R	333	THR	2.7
1	A	521	PRO	2.7
1	A	529	LYS	2.6
3	C	154	LEU	2.6
1	A	368	LEU	2.6
2	B	141	SER	2.5
2	B	135	PRO	2.5
3	L	125	LEU	2.4
1	R	334	ASN	2.3
1	R	368	LEU	2.3
3	L	123	GLU	2.3
2	B	139	SER	2.3
1	R	370	ASN	2.3
1	A	372	ALA	2.3
1	A	446	GLY	2.2
1	A	346	ARG	2.2
1	A	393	THR	2.2
1	A	360	ASN	2.2
3	L	150	VAL	2.2
2	H	136	SER	2.2
1	R	360	ASN	2.1
3	L	184	ALA	2.1
1	A	363	ALA	2.0

## 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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
5	NAG	R	603	14/15	0.74	0.16	59,77,83,83	0
4	SO4	C	302	5/5	0.83	0.17	96,98,117,125	0
5	NAG	A	602	14/15	0.83	0.20	66,77,83,83	0
4	SO4	R	601	5/5	0.88	0.19	95,106,116,122	0
4	SO4	C	301	5/5	0.92	0.20	86,92,98,107	0
4	SO4	R	602	5/5	0.92	0.17	90,105,114,128	0
4	SO4	H	301	5/5	0.93	0.10	78,89,105,106	0
4	SO4	A	601	5/5	0.93	0.23	101,105,112,126	0
4	SO4	B	301	5/5	0.96	0.10	80,87,99,99	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.