Full wwPDB X-ray Structure Validation Report (i)

Oct 14, 2024 – 07:10 pm BST

PDB ID	:	8S6M
Title	:	SARS-CoV-2 BQ.1.1 RBD bound to the S2V29 and the S2H97 Fab fragments
Authors	:	Errico, J.M.; Park, Y.J.; Rietz, T.; Czudnochowski, N.; Nix, J.C.; Cameroni, E.; Corti, D.; Snell, G.; Marco, A.D.; Pinto, D.; Seattle Structural Genomics Center for Infectious Disease (SSGCID); Veesler, D.
Deposited on	:	2024-02-28
Resolution	:	1.67 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*
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 (1)) were used in the production of this report:

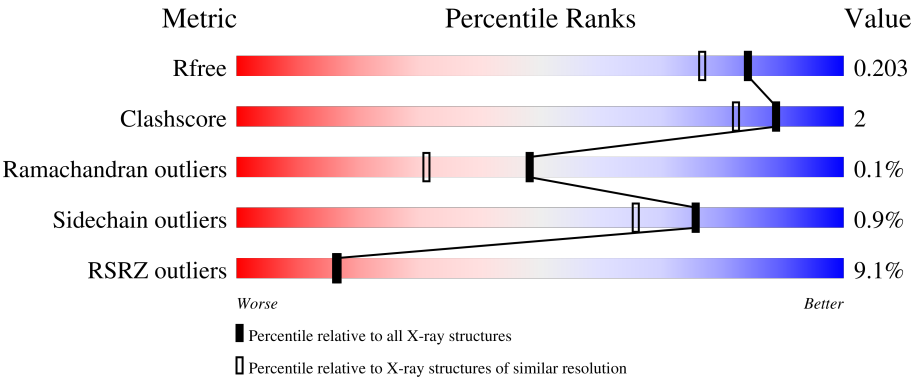
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.003 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.39

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 1.67 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}	164625	8422 (1.70-1.66)
Clashscore	180529	1005 (1.68-1.68)
Ramachandran outliers	177936	9065 (1.70-1.66)
Sidechain outliers	177891	9064 (1.70-1.66)
RSRZ outliers	164620	8421 (1.70-1.66)

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	H	226	2% 93% 5% ·
2	R	269	4% 71% · 27%
3	I	223	17% 87% 7% 5%
4	L	216	2% 93% 6% ·
5	M	218	17% 96% ··

2 Entry composition (i)

There are 10 unique types of molecules in this entry. The entry contains 8673 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 S2V29 Fab heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	H	221	Total	C	N	O	S	0	7	0
			1686	1077	272	328	9			

- Molecule 2 is a protein called Spike protein S1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	R	196	Total	C	N	O	S	0	1	0
			1539	988	258	285	8			

There are 85 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
R	309	MET	-	initiating methionine	UNP P0DTC2
R	310	GLU	-	expression tag	UNP P0DTC2
R	311	TRP	-	expression tag	UNP P0DTC2
R	312	SER	-	expression tag	UNP P0DTC2
R	313	TRP	-	expression tag	UNP P0DTC2
R	314	VAL	-	expression tag	UNP P0DTC2
R	315	PHE	-	expression tag	UNP P0DTC2
R	316	LEU	-	expression tag	UNP P0DTC2
R	317	PHE	-	expression tag	UNP P0DTC2
R	318	PHE	-	expression tag	UNP P0DTC2
R	319	LEU	-	expression tag	UNP P0DTC2
R	320	SER	-	expression tag	UNP P0DTC2
R	321	VAL	-	expression tag	UNP P0DTC2
R	322	THR	-	expression tag	UNP P0DTC2
R	323	THR	-	expression tag	UNP P0DTC2
R	324	GLY	-	expression tag	UNP P0DTC2
R	325	VAL	-	expression tag	UNP P0DTC2
R	326	HIS	-	expression tag	UNP P0DTC2
R	327	SER	-	expression tag	UNP P0DTC2
R	339	ASP	GLY	variant	UNP P0DTC2

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
R	346	THR	ARG	variant	UNP P0DTC2
R	371	PHE	SER	variant	UNP P0DTC2
R	373	PRO	SER	variant	UNP P0DTC2
R	375	PHE	SER	variant	UNP P0DTC2
R	376	ALA	THR	variant	UNP P0DTC2
R	405	ASN	ASP	variant	UNP P0DTC2
R	408	SER	ARG	variant	UNP P0DTC2
R	417	ASN	LYS	variant	UNP P0DTC2
R	440	LYS	ASN	variant	UNP P0DTC2
R	444	THR	LYS	variant	UNP P0DTC2
R	452	ARG	LEU	variant	UNP P0DTC2
R	460	LYS	ASN	variant	UNP P0DTC2
R	477	ASN	SER	variant	UNP P0DTC2
R	478	LYS	THR	variant	UNP P0DTC2
R	484	ALA	GLU	variant	UNP P0DTC2
R	486	VAL	PHE	variant	UNP P0DTC2
R	498	ARG	GLN	variant	UNP P0DTC2
R	501	TYR	ASN	variant	UNP P0DTC2
R	505	HIS	TYR	variant	UNP P0DTC2
R	532	GLY	-	expression tag	UNP P0DTC2
R	533	SER	-	expression tag	UNP P0DTC2
R	534	LEU	-	expression tag	UNP P0DTC2
R	535	VAL	-	expression tag	UNP P0DTC2
R	536	PRO	-	expression tag	UNP P0DTC2
R	537	ARG	-	expression tag	UNP P0DTC2
R	538	GLY	-	expression tag	UNP P0DTC2
R	539	SER	-	expression tag	UNP P0DTC2
R	540	SER	-	expression tag	UNP P0DTC2
R	541	ALA	-	expression tag	UNP P0DTC2
R	542	TRP	-	expression tag	UNP P0DTC2
R	543	SER	-	expression tag	UNP P0DTC2
R	544	HIS	-	expression tag	UNP P0DTC2
R	545	PRO	-	expression tag	UNP P0DTC2
R	546	GLN	-	expression tag	UNP P0DTC2
R	547	PHE	-	expression tag	UNP P0DTC2
R	548	GLU	-	expression tag	UNP P0DTC2
R	549	LYS	-	expression tag	UNP P0DTC2
R	550	GLY	-	expression tag	UNP P0DTC2
R	551	GLY	-	expression tag	UNP P0DTC2
R	552	GLY	-	expression tag	UNP P0DTC2
R	553	SER	-	expression tag	UNP P0DTC2
R	554	GLY	-	expression tag	UNP P0DTC2

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
R	555	GLY	-	expression tag	UNP P0DTC2
R	556	GLY	-	expression tag	UNP P0DTC2
R	557	SER	-	expression tag	UNP P0DTC2
R	558	GLY	-	expression tag	UNP P0DTC2
R	559	GLY	-	expression tag	UNP P0DTC2
R	560	SER	-	expression tag	UNP P0DTC2
R	561	ALA	-	expression tag	UNP P0DTC2
R	562	TRP	-	expression tag	UNP P0DTC2
R	563	SER	-	expression tag	UNP P0DTC2
R	564	HIS	-	expression tag	UNP P0DTC2
R	565	PRO	-	expression tag	UNP P0DTC2
R	566	GLN	-	expression tag	UNP P0DTC2
R	567	PHE	-	expression tag	UNP P0DTC2
R	568	GLU	-	expression tag	UNP P0DTC2
R	569	LYS	-	expression tag	UNP P0DTC2
R	570	HIS	-	expression tag	UNP P0DTC2
R	571	HIS	-	expression tag	UNP P0DTC2
R	572	HIS	-	expression tag	UNP P0DTC2
R	573	HIS	-	expression tag	UNP P0DTC2
R	574	HIS	-	expression tag	UNP P0DTC2
R	575	HIS	-	expression tag	UNP P0DTC2
R	576	HIS	-	expression tag	UNP P0DTC2
R	577	HIS	-	expression tag	UNP P0DTC2

- Molecule 3 is a protein called S2H97 Fab heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	I	211	Total	C	N	O	S	0	1	0
			1556	996	257	296	7			

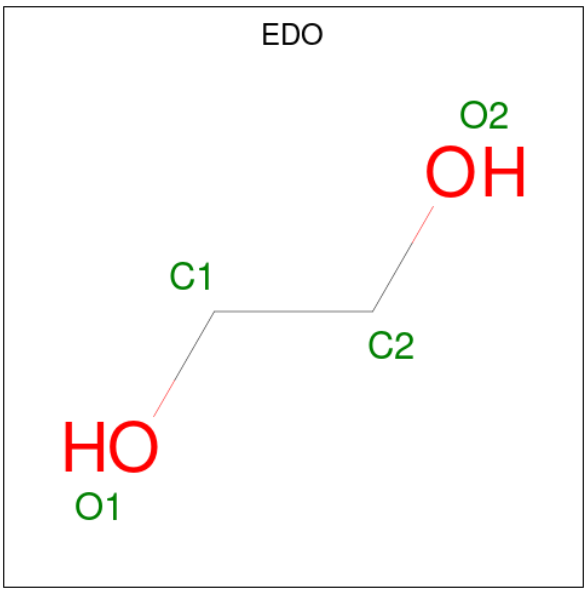
- Molecule 4 is a protein called S2V29 Fab light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	L	214	Total	C	N	O	S	0	0	0
			1573	993	255	320	5			

- Molecule 5 is a protein called S2H97 Fab light chain.

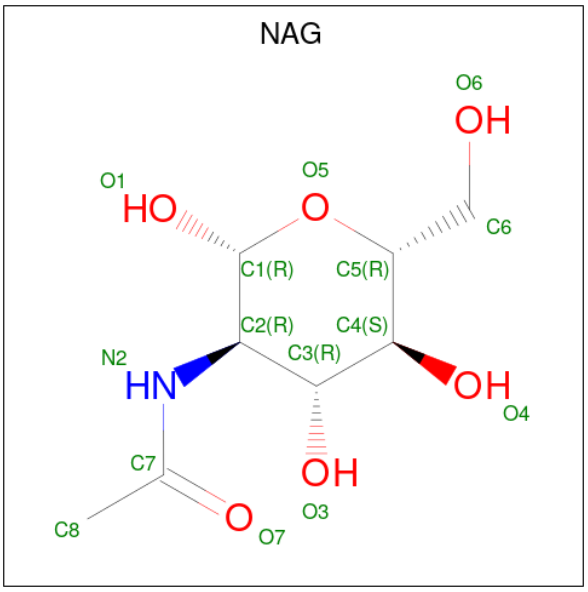
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	M	215	Total	C	N	O	S	0	1	0
			1490	930	249	306	5			

- Molecule 6 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: $C_2H_6O_2$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	H	1	Total	C	O	0	0
			4	2	2		
6	R	1	Total	C	O	0	0
			4	2	2		
6	R	1	Total	C	O	0	0
			4	2	2		
6	R	1	Total	C	O	0	0
			4	2	2		
6	R	1	Total	C	O	0	0
			4	2	2		
6	R	1	Total	C	O	0	0
			4	2	2		
6	I	1	Total	C	O	0	0
			4	2	2		
6	L	1	Total	C	O	0	0
			4	2	2		
6	M	1	Total	C	O	0	0
			4	2	2		

- Molecule 7 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).

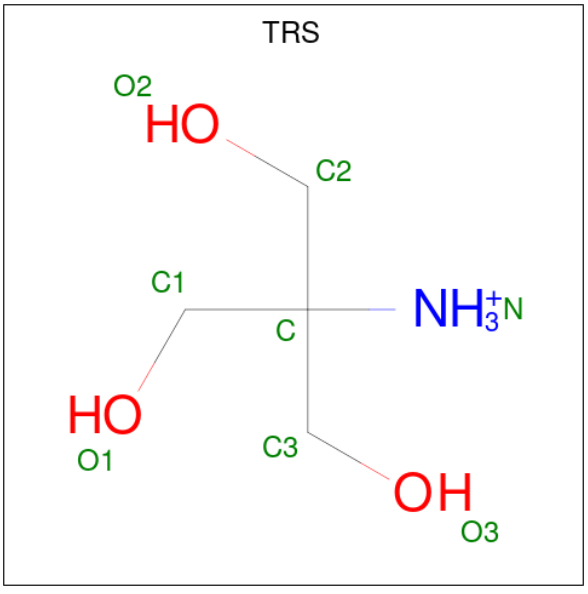


Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	R	1	Total C N O 14 8 1 5	0	0

- Molecule 8 is NICKEL (II) ION (three-letter code: NI) (formula: Ni).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	L	1	Total Ni 1 1	0	0

- Molecule 9 is 2-AMINO-2-HYDROXYMETHYL-PROPANE-1,3-DIOL (three-letter code: TRS) (formula: $C_4H_{12}NO_3$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
9	M	1	Total	C	N	O	0	0
			8	4	1	3		

- Molecule 10 is water.

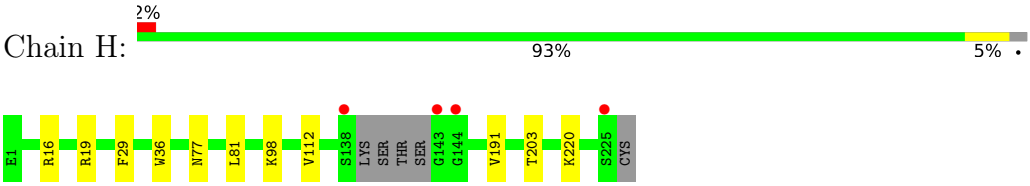
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
10	H	214	Total	O	0	0
			214	214		
10	R	167	Total	O	0	0
			167	167		
10	I	111	Total	O	0	0
			111	111		
10	L	147	Total	O	0	0
			147	147		
10	M	131	Total	O	0	0
			131	131		

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.

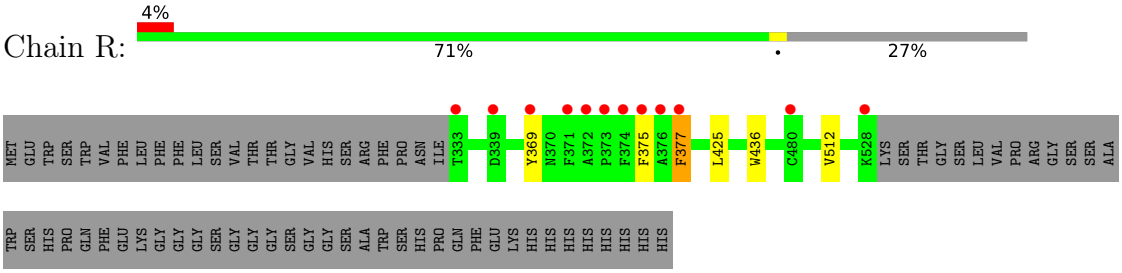
• Molecule 1: S2V29 Fab heavy chain

Chain H:



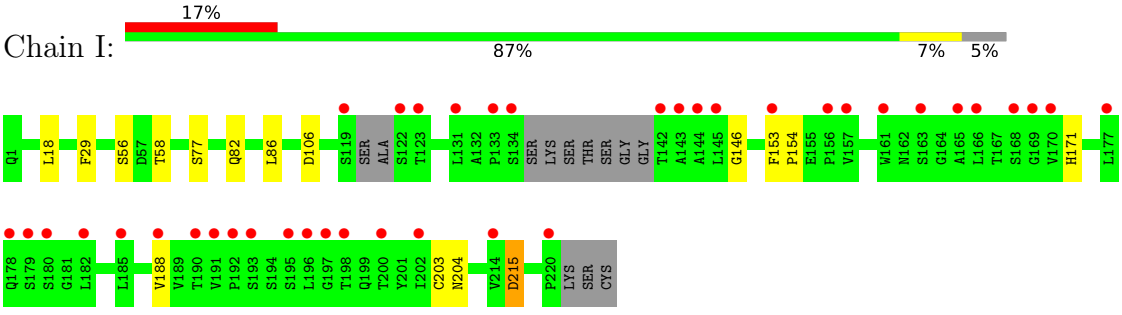
• Molecule 2: Spike protein S1

Chain R:



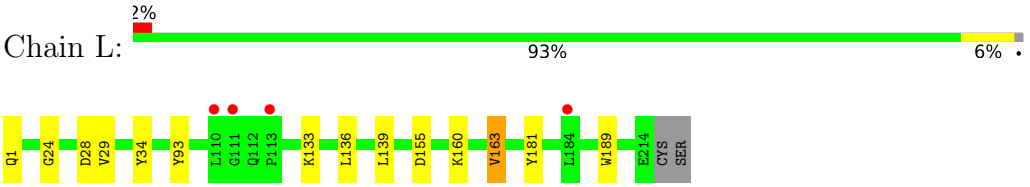
• Molecule 3: S2H97 Fab heavy chain

Chain I:



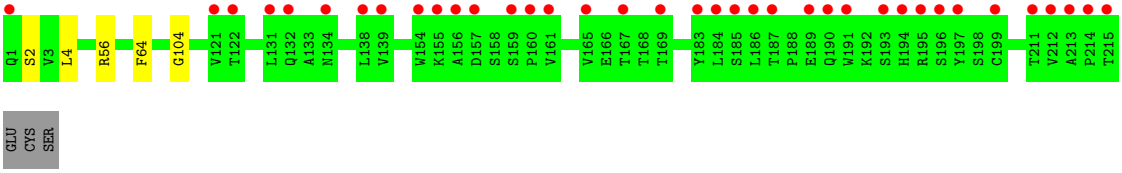
• Molecule 4: S2V29 Fab light chain

Chain L:



• Molecule 5: S2H97 Fab light chain

Chain M:


Q1
S2
V3
L4
R56
F64
G104
V121
T122
L131
Q132
A133
N134
L138
V139
W154
K155
A156
D157
S158
S159
P160
V161
V165
E166
T167
T168
T169
Y183
L184
S185
L186
T187
P188
E189
Q190
W191
K192
S193
H194
R195
S196
Y197
S198
C199
T211
V212
A213
P214
T215

GLU
CYS
SER

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	148.18Å 48.10Å 166.53Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	47.35 – 1.67 47.35 – 1.67	Depositor EDS
% Data completeness (in resolution range)	100.0 (47.35-1.67) 100.0 (47.35-1.67)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.23 (at 1.67Å)	Xtriage
Refinement program	PHENIX 1.18.2_3874	Depositor
R, R_{free}	0.184 , 0.203 0.184 , 0.203	Depositor DCC
R_{free} test set	6987 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å^2)	29.3	Xtriage
Anisotropy	0.034	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 39.8	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	8673	wwPDB-VP
Average B, all atoms (Å^2)	45.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.63% 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

Bond lengths and bond angles in the following residue types are not validated in this section: NAG, NI, EDO, PCA, TRS

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	H	0.43	0/1748	0.62	0/2383
2	R	0.36	0/1587	0.57	0/2165
3	I	0.31	0/1604	0.53	0/2196
4	L	0.37	0/1608	0.61	0/2203
5	M	0.34	0/1531	0.54	0/2106
All	All	0.36	0/8078	0.58	0/11053

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	H	1686	0	1653	7	0
2	R	1539	0	1443	4	0
3	I	1556	0	1404	7	0
4	L	1573	0	1480	7	0
5	M	1490	0	1320	3	0
6	H	4	0	6	0	0
6	I	4	0	6	0	0
6	L	4	0	4	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
6	M	4	0	6	0	0
6	R	20	0	30	0	0
7	R	14	0	13	0	0
8	L	1	0	0	0	0
9	M	8	0	12	0	0
10	H	214	0	0	2	1
10	I	111	0	0	1	0
10	L	147	0	0	0	0
10	M	131	0	0	1	0
10	R	167	0	0	0	0
All	All	8673	0	7377	27	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 2.

All (27) 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)
3:I:82:GLN:NE2	10:I:401:HOH:O	2.25	0.68
3:I:204:ASN:ND2	3:I:215:ASP:OD2	2.35	0.59
1:H:98:LYS:HG2	1:H:112:VAL:HB	1.86	0.58
2:R:369:TYR:HA	2:R:375:PHE:HE2	1.74	0.51
2:R:425:LEU:HD21	2:R:512:VAL:HG11	1.93	0.50
4:L:163:VAL:HA	4:L:181:TYR:O	2.12	0.49
2:R:375:PHE:HB3	2:R:436:TRP:HB3	1.94	0.49
2:R:375:PHE:HE1	2:R:377:PHE:CD2	2.32	0.48
4:L:24:GLY:HA3	4:L:29:VAL:HG13	1.95	0.48
4:L:136:LEU:HD21	4:L:189:TRP:CZ3	2.49	0.48
1:H:203:THR:HG23	1:H:220:LYS:HE3	1.96	0.47
5:M:2:SER:HB3	10:M:443:HOH:O	2.15	0.47
1:H:191:VAL:HB	4:L:139:LEU:HD13	1.97	0.46
10:H:573:HOH:O	4:L:133:LYS:HD2	2.16	0.46
3:I:146:GLY:HA3	3:I:188:VAL:HG12	1.97	0.46
5:M:4:LEU:HB2	5:M:104:GLY:HA2	1.97	0.46
3:I:56:SER:O	3:I:58[A]:THR:HG23	2.17	0.44
3:I:153:PHE:HA	3:I:154:PRO:HA	1.80	0.44
3:I:18:LEU:HB2	3:I:86:LEU:HD11	2.00	0.44
1:H:29:PHE:CD2	1:H:77:ASN:HA	2.53	0.43
4:L:34:TYR:HB2	4:L:93:TYR:HB2	2.01	0.43
1:H:16:ARG:HA	1:H:16:ARG:HD3	1.87	0.42
1:H:36:TRP:CE2	1:H:81:LEU:HB2	2.54	0.42

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
4:L:160:LYS:O	4:L:163:VAL:HG13	2.20	0.42
1:H:19:ARG:NH2	10:H:409:HOH:O	2.54	0.41
3:I:29:PHE:CD2	3:I:77:SER:HA	2.56	0.41
5:M:56:ARG:HD3	5:M:64:PHE:O	2.21	0.41

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)
10:H:571:HOH:O	10:H:582:HOH:O[1_565]	2.12	0.08

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	H	224/226 (99%)	219 (98%)	5 (2%)	0	100	100
2	R	195/269 (72%)	189 (97%)	6 (3%)	0	100	100
3	I	206/223 (92%)	204 (99%)	2 (1%)	0	100	100
4	L	212/216 (98%)	207 (98%)	4 (2%)	1 (0%)	25	11
5	M	214/218 (98%)	212 (99%)	2 (1%)	0	100	100
All	All	1051/1152 (91%)	1031 (98%)	19 (2%)	1 (0%)	48	31

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
4	L	155	ASP

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	H	189/191 (99%)	189 (100%)	0	100	100
2	R	163/228 (72%)	162 (99%)	1 (1%)	84	77
3	I	154/192 (80%)	150 (97%)	4 (3%)	41	22
4	L	169/181 (93%)	167 (99%)	2 (1%)	67	54
5	M	148/186 (80%)	148 (100%)	0	100	100
All	All	823/978 (84%)	816 (99%)	7 (1%)	75	66

All (7) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	R	377	PHE
3	I	106	ASP
3	I	171	HIS
3	I	203	CYS
3	I	215	ASP
4	L	28	ASP
4	L	163	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

5.3.3 RNA (i)

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains (i)

1 non-standard protein/DNA/RNA residue is 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	PCA	L	1	4	7,8,9	2.45	3 (42%)	9,10,12	1.02	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	PCA	L	1	4	-	0/0/11/13	0/1/1/1

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	L	1	PCA	CD-N	5.17	1.48	1.34
4	L	1	PCA	OE-CD	-2.22	1.18	1.23
4	L	1	PCA	CA-N	-2.16	1.43	1.46

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates (i)

There are no oligosaccharides in this entry.

5.6 Ligand geometry (i)

Of 12 ligands modelled in this entry, 1 is monoatomic - leaving 11 for Mogul analysis.

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	$\#\|Z\| > 2$
6	EDO	R	601	-	3,3,3	0.49	0	2,2,2	0.15	0
6	EDO	H	301	-	3,3,3	0.48	0	2,2,2	0.48	0
6	EDO	R	606	-	3,3,3	0.42	0	2,2,2	0.46	0
7	NAG	R	604	2	14,14,15	0.23	0	17,19,21	0.46	0
6	EDO	M	301	-	3,3,3	0.43	0	2,2,2	0.37	0
6	EDO	R	605	-	3,3,3	0.48	0	2,2,2	0.32	0
6	EDO	L	302	8	3,3,3	0.46	0	2,2,2	0.16	0
6	EDO	I	301	-	3,3,3	0.49	0	2,2,2	0.21	0
6	EDO	R	603	-	3,3,3	0.47	0	2,2,2	0.36	0
9	TRS	M	302	-	7,7,7	0.26	0	9,9,9	0.26	0
6	EDO	R	602	-	3,3,3	0.51	0	2,2,2	0.17	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
6	EDO	R	601	-	-	1/1/1/1	-
6	EDO	H	301	-	-	1/1/1/1	-
6	EDO	R	606	-	-	1/1/1/1	-
7	NAG	R	604	2	-	1/6/23/26	0/1/1/1
6	EDO	M	301	-	-	0/1/1/1	-
6	EDO	R	605	-	-	0/1/1/1	-
6	EDO	L	302	8	-	0/1/1/1	-
6	EDO	I	301	-	-	0/1/1/1	-
6	EDO	R	603	-	-	0/1/1/1	-
9	TRS	M	302	-	-	3/9/9/9	-
6	EDO	R	602	-	-	0/1/1/1	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (7) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
9	M	302	TRS	N-C-C2-O2
6	H	301	EDO	O1-C1-C2-O2
7	R	604	NAG	O5-C5-C6-O6
9	M	302	TRS	C3-C-C2-O2
9	M	302	TRS	C1-C-C2-O2
6	R	601	EDO	O1-C1-C2-O2
6	R	606	EDO	O1-C1-C2-O2

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^{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	H	221/226 (97%)	-0.23	4 (1%) 67 71	15, 30, 64, 106	7 (3%)
2	R	196/269 (72%)	0.11	12 (6%) 28 30	17, 33, 80, 129	1 (0%)
3	I	211/223 (94%)	0.68	39 (18%) 4 4	21, 50, 123, 140	1 (0%)
4	L	213/216 (98%)	0.07	4 (1%) 66 70	22, 38, 61, 81	0
5	M	215/218 (98%)	0.73	37 (17%) 5 4	26, 43, 131, 141	1 (0%)
All	All	1056/1152 (91%)	0.27	96 (9%) 16 16	15, 38, 111, 141	10 (0%)

All (96) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	R	371	PHE	6.1
2	R	377	PHE	5.6
2	R	374	PHE	5.6
2	R	373	PRO	5.2
3	I	177	LEU	5.0
1	H	225	SER	4.9
2	R	375	PHE	4.5
2	R	376	ALA	4.3
5	M	187	THR	4.2
3	I	166	LEU	4.2
1	H	138	SER	4.2
3	I	142	THR	4.2
5	M	161	VAL	4.2
5	M	186	LEU	4.2
1	H	143	GLY	4.1
5	M	167	THR	4.1
5	M	215	THR	4.0
3	I	134	SER	3.7
3	I	182	LEU	3.6
5	M	138	LEU	3.6

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
5	M	121	VAL	3.5
5	M	194	HIS	3.4
5	M	157	ASP	3.4
2	R	333	THR	3.4
3	I	119	SER	3.4
5	M	214	PRO	3.3
2	R	369	TYR	3.2
3	I	193	SER	3.2
3	I	185	LEU	3.2
3	I	122	SER	3.1
4	L	111	GLY	3.1
3	I	196	LEU	3.1
5	M	131	LEU	3.1
5	M	165	VAL	3.1
5	M	197	TYR	3.1
5	M	154	TRP	3.0
5	M	185	SER	3.0
3	I	220	PRO	3.0
3	I	163	SER	3.0
5	M	183	TYR	2.9
5	M	184	LEU	2.9
3	I	170	VAL	2.9
5	M	212	VAL	2.9
5	M	1	GLN	2.9
5	M	190	GLN	2.9
3	I	195	SER	2.9
3	I	156	PRO	2.9
5	M	191	TRP	2.8
5	M	189	GLU	2.8
3	I	143	ALA	2.7
3	I	197	GLY	2.7
3	I	192	PRO	2.7
3	I	144	ALA	2.7
5	M	134	ASN	2.7
3	I	168	SER	2.7
3	I	188	VAL	2.6
3	I	153	PHE	2.6
3	I	191	VAL	2.6
2	R	372	ALA	2.5
3	I	179	SER	2.5
5	M	132	GLN	2.5
2	R	528	LYS	2.5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
3	I	200	THR	2.5
3	I	165	ALA	2.5
3	I	131	LEU	2.4
5	M	199	CYS	2.4
5	M	196	SER	2.4
3	I	190	THR	2.4
3	I	161	TRP	2.4
3	I	178	GLN	2.4
5	M	159	SER	2.4
3	I	169	GLY	2.4
5	M	155	LYS	2.3
5	M	211	THR	2.3
5	M	213	ALA	2.3
5	M	169	THR	2.3
5	M	122	THR	2.3
3	I	145	LEU	2.3
2	R	339	ASP	2.2
4	L	110	LEU	2.2
4	L	184	LEU	2.2
5	M	139	VAL	2.2
5	M	156	ALA	2.1
3	I	202	ILE	2.1
3	I	123	THR	2.1
3	I	157	VAL	2.1
5	M	193	SER	2.1
3	I	133	PRO	2.1
3	I	180	SER	2.1
3	I	214	VAL	2.1
4	L	113	PRO	2.1
5	M	160	PRO	2.1
3	I	198	THR	2.0
5	M	195	ARG	2.0
2	R	480	CYS	2.0
1	H	144	GLY	2.0

6.2 Non-standard residues in protein, DNA, RNA chains (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^{th} 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	PCA	L	1	8/9	0.87	0.13	45,48,51,56	0

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^{th} 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
7	NAG	R	604	14/15	0.56	0.24	70,78,84,86	14
9	TRS	M	302	8/8	0.69	0.14	79,81,92,95	0
6	EDO	H	301	4/4	0.72	0.21	58,58,63,64	0
6	EDO	R	605	4/4	0.85	0.16	61,61,63,65	0
6	EDO	I	301	4/4	0.87	0.22	70,71,72,74	0
6	EDO	L	302	4/4	0.87	0.15	42,47,48,58	0
6	EDO	R	606	4/4	0.88	0.30	117,120,120,122	0
6	EDO	M	301	4/4	0.92	0.23	71,74,75,75	0
6	EDO	R	603	4/4	0.93	0.08	35,36,40,43	0
6	EDO	R	602	4/4	0.94	0.09	34,34,38,41	0
6	EDO	R	601	4/4	0.97	0.06	30,30,31,39	0
8	NI	L	301	1/1	0.98	0.04	44,44,44,44	0

6.5 Other polymers (i)

There are no such residues in this entry.