Full wwPDB X-ray Structure Validation Report (i)

Oct 30, 2023 – 09:07 am GMT

PDB ID	:	8BG5
Title	:	Crystal structure of the SARS-CoV-2 S RBD in complex with pT1631 scFV
Authors	:	Hansen, G.; Ssebyatika, G.L.; Krey, T.
Deposited on	:	2022-10-27
Resolution	:	2.05 Å(reported)

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

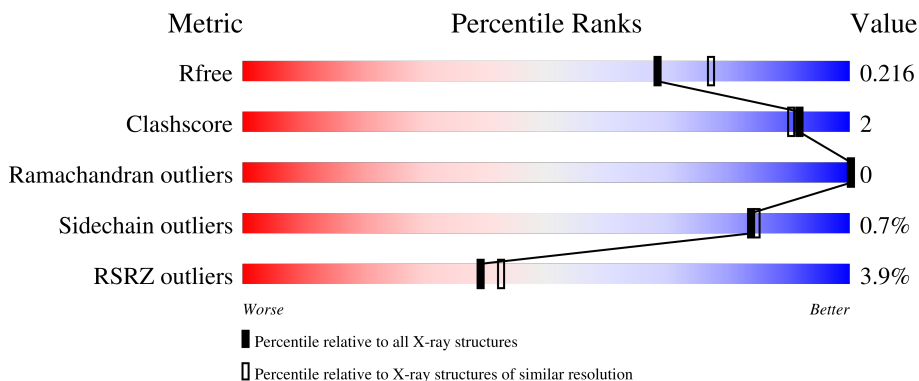|		
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	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 2.05 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		0.216
Clashscore		2
Ramachandran outliers		0
Sidechain outliers		0.7%
RSRZ outliers		3.9%

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	281	6% — 75% / 5% / 20%
1	C	281	3% — 77% / · / 20%
2	B	229	% — 82% / · / 15%
2	D	229	2% — 79% / 6% / 15%

2 Entry composition ⓘ

There are 4 unique types of molecules in this entry. The entry contains 6794 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 pT1631 single-chain Fv.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	225	Total	C	N	O	S	0	0	0
			1722	1094	288	334	6			
1	C	225	Total	C	N	O	S	0	0	0
			1722	1094	288	334	6			

- Molecule 2 is a protein called Spike protein S1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	194	Total	C	N	O	S	0	0	0
			1536	985	256	287	8			
2	D	194	Total	C	N	O	S	0	0	0
			1536	985	256	287	8			

There are 70 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	528	ASP	-	expression tag	UNP P0DTC2
B	529	ASP	-	expression tag	UNP P0DTC2
B	530	ASP	-	expression tag	UNP P0DTC2
B	531	ASP	-	expression tag	UNP P0DTC2
B	532	LYS	-	expression tag	UNP P0DTC2
B	533	ALA	-	expression tag	UNP P0DTC2
B	534	GLY	-	expression tag	UNP P0DTC2
B	535	TRP	-	expression tag	UNP P0DTC2
B	536	SER	-	expression tag	UNP P0DTC2
B	537	HIS	-	expression tag	UNP P0DTC2
B	538	PRO	-	expression tag	UNP P0DTC2
B	539	GLN	-	expression tag	UNP P0DTC2
B	540	PHE	-	expression tag	UNP P0DTC2
B	541	GLU	-	expression tag	UNP P0DTC2
B	542	LYS	-	expression tag	UNP P0DTC2
B	543	GLY	-	expression tag	UNP P0DTC2

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	544	GLY	-	expression tag	UNP P0DTC2
B	545	GLY	-	expression tag	UNP P0DTC2
B	546	SER	-	expression tag	UNP P0DTC2
B	547	GLY	-	expression tag	UNP P0DTC2
B	548	GLY	-	expression tag	UNP P0DTC2
B	549	GLY	-	expression tag	UNP P0DTC2
B	550	SER	-	expression tag	UNP P0DTC2
B	551	GLY	-	expression tag	UNP P0DTC2
B	552	GLY	-	expression tag	UNP P0DTC2
B	553	GLY	-	expression tag	UNP P0DTC2
B	554	SER	-	expression tag	UNP P0DTC2
B	555	TRP	-	expression tag	UNP P0DTC2
B	556	SER	-	expression tag	UNP P0DTC2
B	557	HIS	-	expression tag	UNP P0DTC2
B	558	PRO	-	expression tag	UNP P0DTC2
B	559	GLN	-	expression tag	UNP P0DTC2
B	560	PHE	-	expression tag	UNP P0DTC2
B	561	GLU	-	expression tag	UNP P0DTC2
B	562	LYS	-	expression tag	UNP P0DTC2
D	528	ASP	-	expression tag	UNP P0DTC2
D	529	ASP	-	expression tag	UNP P0DTC2
D	530	ASP	-	expression tag	UNP P0DTC2
D	531	ASP	-	expression tag	UNP P0DTC2
D	532	LYS	-	expression tag	UNP P0DTC2
D	533	ALA	-	expression tag	UNP P0DTC2
D	534	GLY	-	expression tag	UNP P0DTC2
D	535	TRP	-	expression tag	UNP P0DTC2
D	536	SER	-	expression tag	UNP P0DTC2
D	537	HIS	-	expression tag	UNP P0DTC2
D	538	PRO	-	expression tag	UNP P0DTC2
D	539	GLN	-	expression tag	UNP P0DTC2
D	540	PHE	-	expression tag	UNP P0DTC2
D	541	GLU	-	expression tag	UNP P0DTC2
D	542	LYS	-	expression tag	UNP P0DTC2
D	543	GLY	-	expression tag	UNP P0DTC2
D	544	GLY	-	expression tag	UNP P0DTC2
D	545	GLY	-	expression tag	UNP P0DTC2
D	546	SER	-	expression tag	UNP P0DTC2
D	547	GLY	-	expression tag	UNP P0DTC2
D	548	GLY	-	expression tag	UNP P0DTC2
D	549	GLY	-	expression tag	UNP P0DTC2
D	550	SER	-	expression tag	UNP P0DTC2

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	551	GLY	-	expression tag	UNP P0DTC2
D	552	GLY	-	expression tag	UNP P0DTC2
D	553	GLY	-	expression tag	UNP P0DTC2
D	554	SER	-	expression tag	UNP P0DTC2
D	555	TRP	-	expression tag	UNP P0DTC2
D	556	SER	-	expression tag	UNP P0DTC2
D	557	HIS	-	expression tag	UNP P0DTC2
D	558	PRO	-	expression tag	UNP P0DTC2
D	559	GLN	-	expression tag	UNP P0DTC2
D	560	PHE	-	expression tag	UNP P0DTC2
D	561	GLU	-	expression tag	UNP P0DTC2
D	562	LYS	-	expression tag	UNP P0DTC2

- Molecule 3 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).
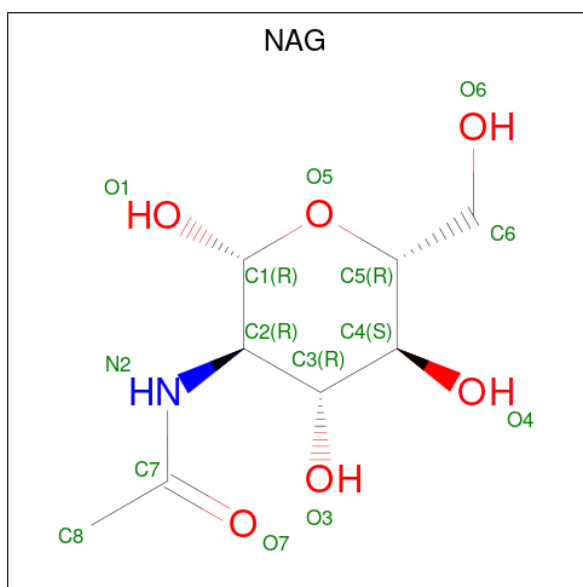
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	B	1	Total	C	N	O	0	0
			14	8	1	5		
3	D	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	56	Total	O	0	0
			56	56		

Continued from previous page...

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	B	62	Total 62	O 62	0	0
4	C	59	Total 59	O 59	0	0
4	D	73	Total 73	O 73	0	0

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.
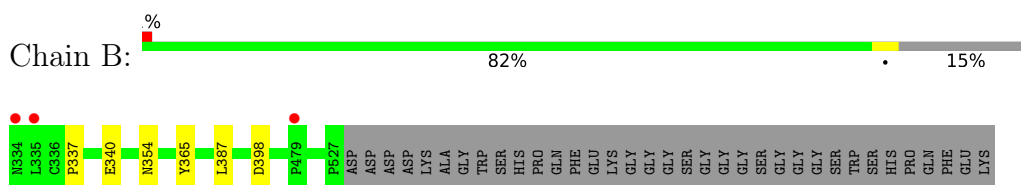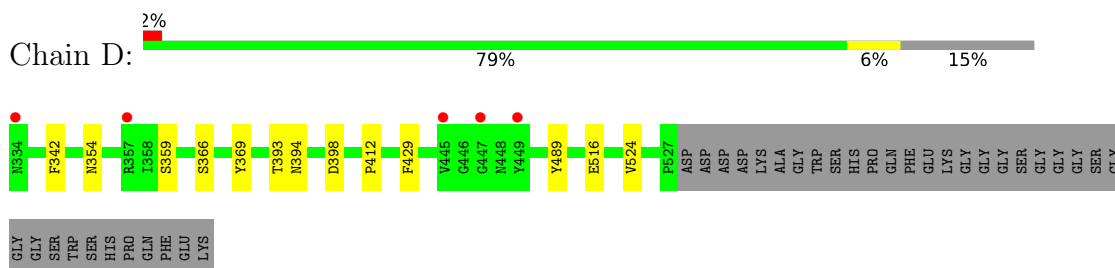
- Molecule 1: pT1631 single-chain Fv
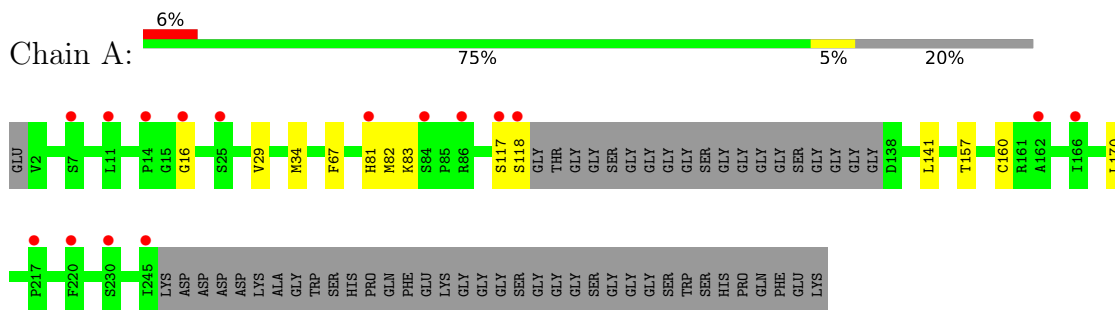
- Molecule 1: pT1631 single-chain Fv
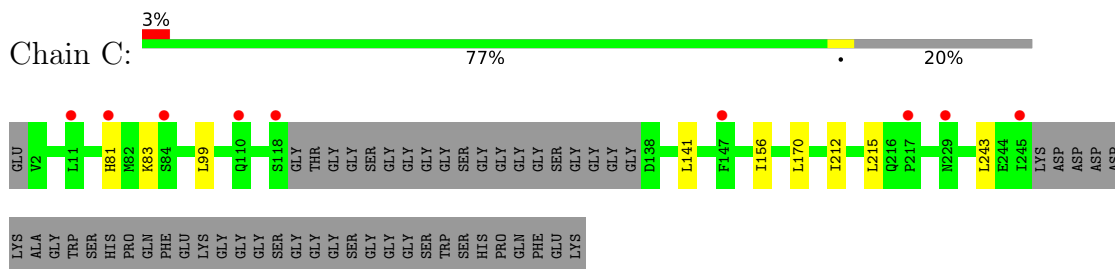
- Molecule 2: Spike protein S1

- Molecule 2: Spike protein S1

4   Data and refinement statistics   ⓘ

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	90.08Å 90.81Å 154.38Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.25 – 2.05 49.25 – 2.05	Depositor EDS
% Data completeness (in resolution range)	99.8 (49.25-2.05) 99.8 (49.25-2.05)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.13 (at 2.05Å)	Xtriage
Refinement program	PHENIX 1.20.1_4487	Depositor
R, R_{free}	0.196 , 0.217 0.195 , 0.216	Depositor DCC
R_{free} test set	2003 reflections (2.51%)	wwPDB-VP
Wilson B-factor (Å^2)	40.9	Xtriage
Anisotropy	0.435	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.38 , 45.1	EDS
L-test for twinning[2]	$< |L| > = 0.50, < L^2 > = 0.34$	Xtriage
Estimated twinning fraction	0.005 for k,h,-l	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	6794	wwPDB-VP
Average B, all atoms (Å^2)	42.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.20% 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

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	0.37	0/1765	0.63	0/2399
1	C	0.39	0/1765	0.61	0/2399
2	B	0.40	0/1580	0.58	0/2151
2	D	0.41	0/1580	0.59	0/2151
All	All	0.39	0/6690	0.61	0/9100

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	1722	0	1668	5	0
1	C	1722	0	1668	6	0
2	B	1536	0	1452	3	0
2	D	1536	0	1452	8	0
3	B	14	0	13	0	0
3	D	14	0	13	1	0
4	A	56	0	0	1	0
4	B	62	0	0	0	0
4	C	59	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	D	73	0	0	1	0
All	All	6794	0	6266	21	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 2.

All (21) 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:C:81:HIS:HB3	1:C:83:LYS:HE2	1.67	0.75
2:D:359:SER:HA	2:D:524:VAL:HG22	1.85	0.58
1:A:16:GLY:HA2	4:A:328:HOH:O	2.03	0.58
1:C:81:HIS:HB3	1:C:83:LYS:CE	2.33	0.57
2:D:393:THR:HG22	2:D:394:ASN:OD1	2.05	0.56
1:A:81:HIS:HB3	1:A:83:LYS:HE2	1.89	0.53
1:C:156:ILE:HD11	1:C:212:ILE:HD12	1.92	0.50
2:B:354:ASN:O	2:B:398:ASP:HA	2.13	0.49
1:A:67:PHE:CZ	1:A:82:MET:HG2	2.49	0.48
1:C:99:LEU:HD21	2:D:489:TYR:CE2	2.49	0.47
2:D:516:GLU:HG2	4:D:1106:HOH:O	2.15	0.46
2:D:342:PHE:HB2	3:D:1000:NAG:H82	1.96	0.46
2:D:366:SER:HA	2:D:369:TYR:CE2	2.51	0.45
1:C:215:LEU:HD11	1:C:243:LEU:HD21	1.99	0.45
1:A:29:VAL:HG13	1:A:34:MET:HG3	1.99	0.44
2:D:412:PRO:HG3	2:D:429:PHE:HB3	2.00	0.44
1:C:141:LEU:HA	1:C:141:LEU:HD12	1.86	0.44
1:A:141:LEU:HG	1:A:160:CYS:SG	2.58	0.43
2:B:337:PRO:HB2	2:B:340:GLU:HG3	2.00	0.43
2:D:354:ASN:O	2:D:398:ASP:HA	2.19	0.42
2:B:365:TYR:CD2	2:B:387:LEU:HB3	2.56	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	221/281 (79%)	211 (96%)	10 (4%)	0	100	100
1	C	221/281 (79%)	208 (94%)	13 (6%)	0	100	100
2	B	192/229 (84%)	187 (97%)	5 (3%)	0	100	100
2	D	192/229 (84%)	186 (97%)	6 (3%)	0	100	100
All	All	826/1020 (81%)	792 (96%)	34 (4%)	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	187/217 (86%)	183 (98%)	4 (2%)	53	48
1	C	187/217 (86%)	186 (100%)	1 (0%)	88	89
2	B	167/191 (87%)	167 (100%)	0	100	100
2	D	167/191 (87%)	167 (100%)	0	100	100
All	All	708/816 (87%)	703 (99%)	5 (1%)	84	84

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	117	SER
1	A	118	SER
1	A	157	THR
1	A	170	LEU
1	C	170	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
2	D	334	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 ⓘ

2 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	#$|Z| > 2$
3	NAG	B	1000	2	14,14,15	0.67	0	17,19,21	0.88	1 (5%)
3	NAG	D	1000	2	14,14,15	0.76	1 (7%)	17,19,21	0.59	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
3	NAG	B	1000	2	-	0/6/23/26	0/1/1/1
3	NAG	D	1000	2	-	2/6/23/26	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	D	1000	NAG	O5-C1	2.41	1.47	1.43

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	B	1000	NAG	C1-O5-C5	2.90	116.12	112.19

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	D	1000	NAG	O5-C5-C6-O6
3	D	1000	NAG	C4-C5-C6-O6

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	D	1000	NAG	1	0

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(Å^2)	Q<0.9
1	A	225/281 (80%)	0.70	16 (7%)	16	17	33, 42, 62, 68	0
1	C	225/281 (80%)	0.62	9 (4%)	38	41	30, 40, 56, 68	0
2	B	194/229 (84%)	0.42	3 (1%)	73	76	29, 40, 54, 66	0
2	D	194/229 (84%)	0.43	5 (2%)	56	60	30, 39, 52, 59	0
All	All	838/1020 (82%)	0.55	33 (3%)	39	42	29, 41, 57, 68	0

All (33) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	334	ASN	5.2
1	A	245	ILE	5.1
2	B	335	LEU	3.7
1	C	11	LEU	3.6
1	A	230	SER	3.6
1	A	217	PRO	3.6
1	A	11	LEU	3.4
2	D	334	ASN	3.1
1	A	14	PRO	3.0
2	D	449	TYR	3.0
1	C	118	SER	2.9
1	A	84	SER	2.9
2	D	445	VAL	2.8
1	A	117	SER	2.8
1	C	84	SER	2.7
1	A	81	HIS	2.7
1	A	118	SER	2.5
1	A	166	ILE	2.4
1	A	220	PHE	2.4
1	C	147	PHE	2.4
2	B	479	PRO	2.4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	16	GLY	2.3
1	C	217	PRO	2.3
1	C	245	ILE	2.3
1	C	110	GLN	2.3
1	A	7	SER	2.3
2	D	357	ARG	2.2
1	A	25	SER	2.2
1	A	162	ALA	2.1
1	C	81	HIS	2.1
2	D	447	GLY	2.0
1	C	229	ASN	2.0
1	A	86	ARG	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 ⓘ

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, 95th 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
3	NAG	B	1000	14/15	0.83	0.19	46,55,59,61	0
3	NAG	D	1000	14/15	0.90	0.22	50,62,66,69	0

6.5 Other polymers ⓘ

There are no such residues in this entry.