



wwPDB X-ray Structure Validation Summary Report

Feb 27, 2014 – 01:26 AM GMT

PDB ID : 2B8H
Title : A/NWS/whale/Maine/1/84(H1N9) reassortant influenza virus neuraminidase
Authors : Smith, B.J.; Platis, D.; Cox, M.M.J.; Huyton, T.; Joosten, R.P.; McKimm-Breschkin, J.L.; Zhang, J.-G.; Luo, C.S.; Lou, M.-Z.; Garrett, T.P.J.; Labrou, N.E.
Deposited on : 2005-10-07
Resolution : 2.20 Å(reported)

This is a wwPDB validation summary report for a publicly released PDB entry.
We welcome your comments at validation@mail.wwpdb.org
A user guide is available at <http://wwpdb.org/ValidationPDFNotes.html>

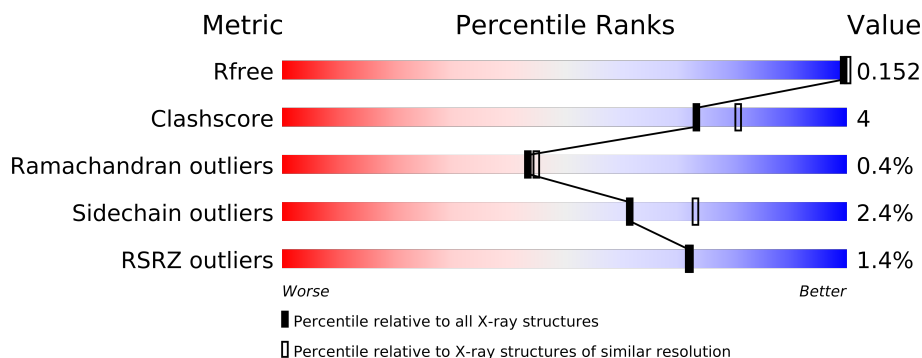
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : 1.15 2013
Xtriage (Phenix) : dev-1323
EDS : stable22639
Percentile statistics : 21963
Refmac : 5.8.0049
CCP4 : 6.3.0 (Settle)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP) : stable22683

1 Overall quality at a glance

The reported resolution of this entry is 2.20 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}	66092	2938 (2.20-2.20)
Clashscore	79885	3751 (2.20-2.20)
Ramachandran outliers	78287	3681 (2.20-2.20)
Sidechain outliers	78261	3682 (2.20-2.20)
RSRZ outliers	66119	2939 (2.20-2.20)

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 on the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	388	
1	B	388	
1	C	388	
1	D	388	

The following table lists non-polymeric compounds that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Geometry	Electron density
2	NAG	A	500	-	X
2	NAG	C	500	-	X
2	NAG	D	500	-	X
6	SO4	A	1515	-	X

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Geometry	Electron density
6	SO4	A	1517	-	X
6	SO4	A	1518	-	X
6	SO4	A	1519	-	X
6	SO4	A	1520	-	X
6	SO4	A	2516	-	X
6	SO4	B	2513	-	X
6	SO4	B	2515	-	X
6	SO4	B	2519	-	X
6	SO4	B	3516	-	X
6	SO4	C	3515	-	X
6	SO4	C	3517	-	X
6	SO4	C	3519	-	X
6	SO4	C	3520	-	X
6	SO4	C	4516	-	X
6	SO4	D	1516	-	X
6	SO4	D	4515	-	X
6	SO4	D	4519	-	X
6	SO4	D	4520	-	X
8	GOL	A	5001	-	X
8	GOL	A	5002	-	X

2 Entry composition

There are 9 unique types of molecules in this entry. The entry contains 14424 atoms, of which 0 are hydrogen and 0 are deuterium.

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.

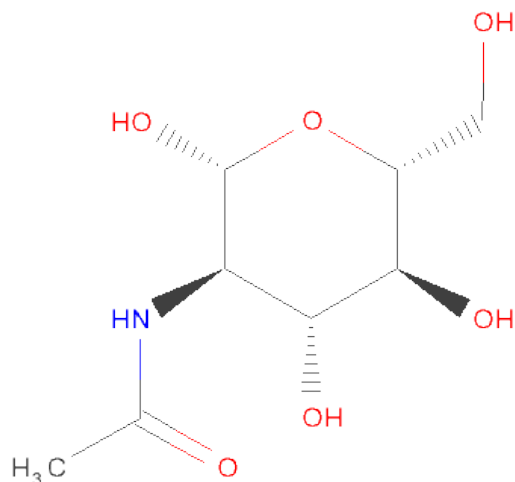
- Molecule 1 is a protein called Neuraminidase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	388	Total	C	N	O	S	0	4	0
			3069	1910	539	595	25			
1	B	388	Total	C	N	O	S	0	4	0
			3067	1908	540	594	25			
1	C	388	Total	C	N	O	S	0	4	0
			3067	1909	540	593	25			
1	D	388	Total	C	N	O	S	0	6	0
			3082	1919	544	594	25			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	368	VAL	ILE	CONFLICT	UNP P05803
B	368	VAL	ILE	CONFLICT	UNP P05803
C	368	VAL	ILE	CONFLICT	UNP P05803
D	368	VAL	ILE	CONFLICT	UNP P05803

- Molecule 2 is SUGAR (N-ACETYL-D-GLUCOSAMINE) (three-letter code: NAG) (formula: C₈H₁₅NO₆).



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

- Molecule 3 is a polymer of unknown type called SUGAR (2-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	2	Total	C	N	O	0	0
			28	16	2	10		
3	B	2	Total	C	N	O	0	0
			28	16	2	10		
3	C	2	Total	C	N	O	0	0
			28	16	2	10		
3	D	2	Total	C	N	O	0	0
			28	16	2	10		

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	368	VAL	ILE	CONFLICT	UNP P05803
B	368	VAL	ILE	CONFLICT	UNP P05803
C	368	VAL	ILE	CONFLICT	UNP P05803

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	368	VAL	ILE	CONFLICT	UNP P05803

- Molecule 4 is a polymer of unknown type called SUGAR (9-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	9	Total	C	N	O	0	0
			105	58	2	45		
4	C	9	Total	C	N	O	0	0
			105	58	2	45		
4	D	9	Total	C	N	O	0	0
			105	58	2	45		

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	368	VAL	ILE	CONFLICT	UNP P05803
C	368	VAL	ILE	CONFLICT	UNP P05803
D	368	VAL	ILE	CONFLICT	UNP P05803

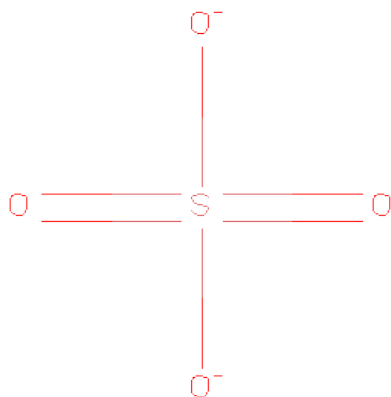
- Molecule 5 is a polymer of unknown type called SUGAR (10-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	B	10	Total	C	N	O	0	0
			116	64	2	50		

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	368	VAL	ILE	CONFLICT	UNP P05803

- Molecule 6 is SULFATE ION (three-letter code: SO4) (formula: O₄S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	A	1	Total	O	S	0	0
			5	4	1		
6	A	1	Total	O	S	0	0
			5	4	1		
6	A	1	Total	O	S	0	0
			5	4	1		
6	D	1	Total	O	S	0	0
			5	4	1		
6	A	1	Total	O	S	0	0
			5	4	1		
6	A	1	Total	O	S	0	0
			5	4	1		
6	A	1	Total	O	S	0	0
			5	4	1		
6	A	1	Total	O	S	0	0
			5	4	1		
6	B	1	Total	O	S	0	0
			5	4	1		
6	B	1	Total	O	S	0	0
			5	4	1		
6	B	1	Total	O	S	0	0
			5	4	1		
6	A	1	Total	O	S	0	0
			5	4	1		
6	B	1	Total	O	S	0	0
			5	4	1		
6	B	1	Total	O	S	0	0
			5	4	1		

Continued on next page...

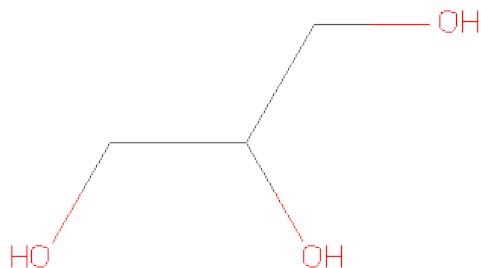
Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	C	1	Total	O	S	0	0
			5	4	1		
6	C	1	Total	O	S	0	0
			5	4	1		
6	C	1	Total	O	S	0	0
			5	4	1		
6	B	1	Total	O	S	0	0
			5	4	1		
6	C	1	Total	O	S	0	0
			5	4	1		
6	C	1	Total	O	S	0	0
			5	4	1		
6	C	1	Total	O	S	0	0
			5	4	1		
6	C	1	Total	O	S	0	0
			5	4	1		
6	D	1	Total	O	S	0	0
			5	4	1		
6	D	1	Total	O	S	0	0
			5	4	1		
6	D	1	Total	O	S	0	0
			5	4	1		
6	C	1	Total	O	S	0	0
			5	4	1		
6	D	1	Total	O	S	0	0
			5	4	1		
6	D	1	Total	O	S	0	0
			5	4	1		

- Molecule 7 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	B	1	Total	Cl	0	0
			1	1		
7	A	1	Total	Cl	0	0
			1	1		
7	D	1	Total	Cl	0	0
			1	1		
7	C	1	Total	Cl	0	0
			1	1		

- Molecule 8 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
8	A	1	Total	C	O	0	0
			6	3	3		
8	A	1	Total	C	O	0	0
			6	3	3		
8	A	1	Total	C	O	0	0
			6	3	3		

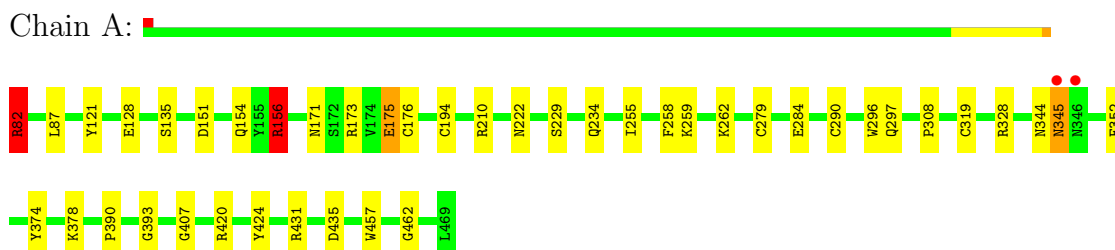
- Molecule 9 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	A	369	Total	O	0	0
			369	369		
9	B	315	Total	O	0	0
			315	315		
9	C	324	Total	O	0	0
			324	324		
9	D	365	Total	O	0	0
			365	365		

3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

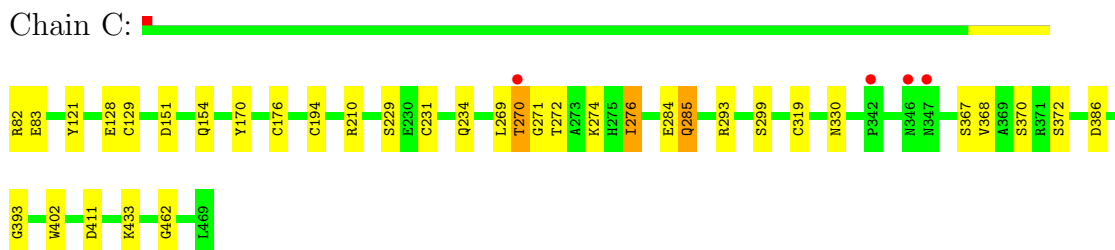
- Molecule 1: Neuraminidase



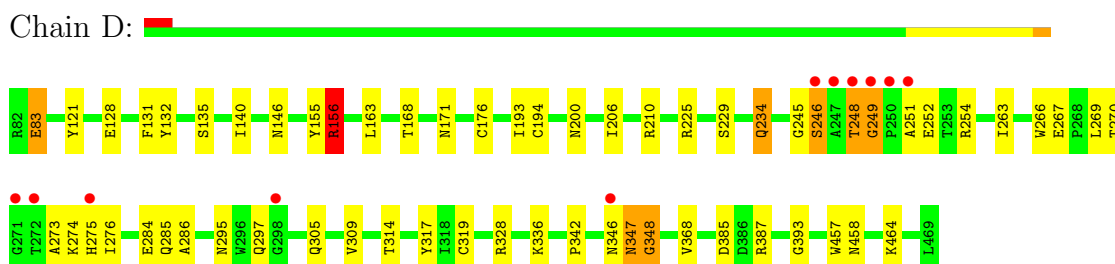
- Molecule 1: Neuraminidase



- Molecule 1: Neuraminidase



- Molecule 1: Neuraminidase



4 Data and refinement statistics

Property	Value	Source
Space group	P 32 2 1	Depositor
Cell constants a, b, c, α , β , γ	107.52Å 107.52Å 338.49Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	20.00 – 2.20 19.95 – 2.20	Depositor EDS
% Data completeness (in resolution range)	99.7 (20.00-2.20) 99.7 (19.95-2.20)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	4.21 (at 2.19Å)	Xtriage
Refinement program	REFMAC 5.2.0003	Depositor
R, R_{free}	0.139 , 0.189 0.151 , 0.152	Depositor DCC
R_{free} test set	5837 reflections (5.31%)	DCC
Wilson B-factor (Å ²)	21.1	Xtriage
Anisotropy	0.003	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.39 , 34.8	EDS
Estimated twinning fraction	0.025 for -h,-k,l	Xtriage
L-test for twinning	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Outliers	0 of 115856 reflections	Xtriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	14424	wwPDB-VP
Average B, all atoms (Å ²)	22.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.65% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

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: GOL, BMA, NAG, CL, SO4, MAN

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	1.11	7/3169 (0.2%)	0.99	9/4316 (0.2%)
1	B	1.10	2/3167 (0.1%)	0.98	7/4314 (0.2%)
1	C	1.13	5/3169 (0.2%)	0.98	8/4315 (0.2%)
1	D	1.15	8/3195 (0.3%)	1.02	7/4352 (0.2%)
All	All	1.12	22/12700 (0.2%)	0.99	31/17297 (0.2%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	C	0	1
1	D	0	1
All	All	0	2

The worst 5 of 22 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	128	GLU	CD-OE2	8.05	1.34	1.25
1	D	319	CYS	CB-SG	-7.90	1.68	1.82
1	A	128	GLU	CG-CD	7.25	1.62	1.51
1	C	128	GLU	CD-OE2	6.97	1.33	1.25
1	C	319	CYS	CB-SG	-6.88	1.70	1.82

The worst 5 of 31 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	156	ARG	NE-CZ-NH2	-14.08	113.26	120.30
1	D	156	ARG	NE-CZ-NH1	12.58	126.59	120.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	156	ARG	NE-CZ-NH2	-11.53	114.54	120.30
1	A	156	ARG	NE-CZ-NH1	11.17	125.88	120.30
1	C	293	ARG	NE-CZ-NH1	9.74	125.17	120.30

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	C	270	THR	Peptide
1	D	246	SER	Peptide

5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	3069	0	2880	23	0
1	B	3067	0	2878	21	0
1	C	3067	0	2878	20	0
1	D	3082	0	2891	37	0
2	A	14	0	13	1	0
2	B	14	0	13	1	0
2	C	14	0	13	1	0
2	D	14	0	13	0	0
3	A	28	0	25	0	0
3	B	28	0	25	1	0
3	C	28	0	25	0	0
3	D	28	0	25	1	0
4	A	105	0	87	1	0
4	C	105	0	88	1	0
4	D	105	0	88	2	0
5	B	116	0	97	1	0
6	A	40	0	0	1	0
6	B	30	0	0	1	0
6	C	45	0	0	3	0
6	D	30	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
7	A	1	0	0	0	0
7	B	1	0	0	1	0
7	C	1	0	0	0	0
7	D	1	0	0	0	0
8	A	18	0	24	1	0
9	A	369	0	0	3	1
9	B	315	0	0	5	0
9	C	324	0	0	4	0
9	D	365	0	0	10	1
All	All	14424	0	12063	105	1

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 4.

The worst 5 of 105 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:B:275:HIS:HB3	1:B:295[A]:ASN:HD21	1.04	1.16
1:B:275:HIS:HB3	1:B:295[A]:ASN:ND2	1.78	0.98
6:A:2516:SO4:O1	9:A:5368:HOH:O	1.84	0.93
1:B:271:GLY:O	1:B:274:LYS:NZ	2.02	0.93
6:B:3516:SO4:O4	9:B:3772:HOH:O	1.86	0.93

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	Distance(Å)	Clash(Å)
9:A:5115:HOH:O	9:D:4752:HOH:O[4_565]	2.19	0.01

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	390/388 (100%)	374 (96%)	16 (4%)	0	100	100
1	B	390/388 (100%)	369 (95%)	19 (5%)	2 (0%)	38	38

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	C	390/388 (100%)	366 (94%)	24 (6%)	0	100	100
1	D	392/388 (101%)	365 (93%)	23 (6%)	4 (1%)	22	18
All	All	1562/1552 (101%)	1474 (94%)	82 (5%)	6 (0%)	43	45

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	249	GLY
1	B	348	GLY
1	D	348	GLY
1	D	246	SER
1	B	347	ASN

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	343/340 (101%)	338 (98%)	5 (2%)	76	86
1	B	343/340 (101%)	333 (97%)	10 (3%)	55	66
1	C	343/340 (101%)	332 (97%)	11 (3%)	51	62
1	D	345/340 (102%)	336 (97%)	9 (3%)	59	70
All	All	1374/1360 (101%)	1339 (98%)	35 (2%)	61	71

5 of 35 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	82	ARG
1	C	272	THR
1	D	284	GLU
1	C	83	GLU
1	C	234	GLN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 10 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	95	ASN
1	C	234	GLN
1	D	200	ASN
1	B	235	ASN
1	C	326	ASN

5.3.3 RNA ⓘ

There are no RNA chains 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 ⓘ

45 carbohydrates 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	A	501	1,3	12,14,15	1.05	1 (8%)	15,19,21	1.44	4 (26%)
3	NAG	A	502	3	12,14,15	0.69	0	15,19,21	1.36	2 (13%)
4	NAG	A	503	1,4	12,14,15	0.88	1 (8%)	15,19,21	1.81	3 (20%)
4	NAG	A	504	4	12,14,15	1.18	2 (16%)	15,19,21	1.85	6 (40%)
4	BMA	A	505	4	10,11,12	0.65	0	11,15,17	1.18	2 (18%)
4	MAN	A	506	4	10,11,12	0.70	0	11,15,17	1.33	2 (18%)
4	MAN	A	507	4	10,11,12	1.32	1 (10%)	11,15,17	1.61	2 (18%)
4	MAN	A	508	4	10,11,12	0.60	0	11,15,17	1.11	1 (9%)
4	MAN	A	509	4	10,11,12	0.93	0	11,15,17	1.63	2 (18%)
4	MAN	A	510	4	10,11,12	0.84	0	11,15,17	2.85	5 (45%)
4	MAN	A	511	4	10,11,12	0.78	0	11,15,17	1.54	2 (18%)
3	NAG	B	501	1,3	12,14,15	1.02	1 (8%)	15,19,21	1.50	3 (20%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
3	NAG	B	502	3	12,14,15	0.55	0	15,19,21	1.71	2 (13%)
5	NAG	B	503	1,5	12,14,15	1.04	1 (8%)	15,19,21	1.76	5 (33%)
5	NAG	B	504	5	12,14,15	1.07	0	15,19,21	1.97	4 (26%)
5	BMA	B	505	5	10,11,12	1.05	1 (10%)	11,15,17	1.77	4 (36%)
5	MAN	B	506	5	10,11,12	1.04	0	11,15,17	1.57	2 (18%)
5	MAN	B	507	5	10,11,12	0.86	0	11,15,17	1.56	1 (9%)
5	MAN	B	508	5	10,11,12	0.74	0	11,15,17	0.86	0
5	MAN	B	509	5	10,11,12	1.01	1 (10%)	11,15,17	1.45	2 (18%)
5	MAN	B	510	5	10,11,12	0.82	0	11,15,17	1.67	3 (27%)
5	MAN	B	511	5	10,11,12	0.97	0	11,15,17	1.10	1 (9%)
5	MAN	B	512	5	10,11,12	0.88	0	11,15,17	1.65	2 (18%)
3	NAG	C	501	1,3	12,14,15	1.14	1 (8%)	15,19,21	2.32	4 (26%)
3	NAG	C	502	3	12,14,15	0.64	0	15,19,21	2.61	3 (20%)
4	NAG	C	503	1,4	12,14,15	0.92	0	15,19,21	1.87	5 (33%)
4	NAG	C	504	4	12,14,15	0.72	1 (8%)	15,19,21	1.40	2 (13%)
4	BMA	C	505	4	10,11,12	1.00	0	11,15,17	2.16	2 (18%)
4	MAN	C	506	4	10,11,12	0.96	0	11,15,17	1.63	4 (36%)
4	MAN	C	507	4	10,11,12	1.03	1 (10%)	11,15,17	1.29	1 (9%)
4	MAN	C	508	4	10,11,12	0.91	0	11,15,17	1.49	2 (18%)
4	MAN	C	509	4	10,11,12	1.18	2 (20%)	11,15,17	0.71	0
4	MAN	C	510	4	10,11,12	0.79	0	11,15,17	2.05	4 (36%)
4	MAN	C	511	4	10,11,12	0.81	0	11,15,17	1.76	3 (27%)
3	NAG	D	501	1,3	12,14,15	1.01	1 (8%)	15,19,21	2.08	5 (33%)
3	NAG	D	502	3	12,14,15	0.99	0	15,19,21	1.92	4 (26%)
4	NAG	D	503	1,4	12,14,15	0.87	1 (8%)	15,19,21	1.72	1 (6%)
4	NAG	D	504	4	12,14,15	0.61	0	15,19,21	2.10	5 (33%)
4	BMA	D	505	4	10,11,12	0.64	0	11,15,17	1.30	2 (18%)
4	MAN	D	506	4	10,11,12	0.90	0	11,15,17	2.02	3 (27%)
4	MAN	D	507	4	10,11,12	0.71	0	11,15,17	1.43	3 (27%)
4	MAN	D	508	4	10,11,12	0.97	0	11,15,17	1.09	0
4	MAN	D	509	4	10,11,12	0.92	0	11,15,17	0.80	0
4	MAN	D	510	4	10,11,12	0.78	0	11,15,17	1.22	2 (18%)
4	MAN	D	511	4	10,11,12	0.69	0	11,15,17	1.93	1 (9%)

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	A	501	1,3	-	0/6/23/26	0/1/1/1
3	NAG	A	502	3	-	0/6/23/26	0/1/1/1
4	NAG	A	503	1,4	-	0/6/23/26	0/1/1/1
4	NAG	A	504	4	-	0/6/23/26	0/1/1/1
4	BMA	A	505	4	-	0/2/19/22	0/1/1/1
4	MAN	A	506	4	-	0/2/19/22	0/1/1/1
4	MAN	A	507	4	-	0/2/19/22	0/1/1/1
4	MAN	A	508	4	-	0/2/19/22	0/1/1/1
4	MAN	A	509	4	-	0/2/19/22	0/1/1/1
4	MAN	A	510	4	-	0/2/19/22	0/1/1/1
4	MAN	A	511	4	-	0/2/19/22	0/1/1/1
3	NAG	B	501	1,3	-	0/6/23/26	0/1/1/1
3	NAG	B	502	3	-	0/6/23/26	0/1/1/1
5	NAG	B	503	1,5	-	0/6/23/26	0/1/1/1
5	NAG	B	504	5	-	0/6/23/26	0/1/1/1
5	BMA	B	505	5	-	0/2/19/22	0/1/1/1
5	MAN	B	506	5	-	0/2/19/22	0/1/1/1
5	MAN	B	507	5	-	0/2/19/22	0/1/1/1
5	MAN	B	508	5	-	0/2/19/22	0/1/1/1
5	MAN	B	509	5	-	0/2/19/22	0/1/1/1
5	MAN	B	510	5	-	0/2/19/22	0/1/1/1
5	MAN	B	511	5	-	0/2/19/22	0/1/1/1
5	MAN	B	512	5	-	0/2/19/22	0/1/1/1
3	NAG	C	501	1,3	-	0/6/23/26	0/1/1/1
3	NAG	C	502	3	-	0/6/23/26	0/1/1/1
4	NAG	C	503	1,4	-	0/6/23/26	0/1/1/1
4	NAG	C	504	4	-	0/6/23/26	0/1/1/1
4	BMA	C	505	4	-	0/2/19/22	0/1/1/1
4	MAN	C	506	4	-	0/2/19/22	0/1/1/1
4	MAN	C	507	4	-	0/2/19/22	0/1/1/1
4	MAN	C	508	4	-	0/2/19/22	0/1/1/1
4	MAN	C	509	4	-	0/2/19/22	0/1/1/1
4	MAN	C	510	4	-	0/2/19/22	0/1/1/1
4	MAN	C	511	4	-	0/2/19/22	0/1/1/1
3	NAG	D	501	1,3	-	0/6/23/26	0/1/1/1
3	NAG	D	502	3	-	0/6/23/26	0/1/1/1
4	NAG	D	503	1,4	-	0/6/23/26	0/1/1/1
4	NAG	D	504	4	-	0/6/23/26	0/1/1/1
4	BMA	D	505	4	-	0/2/19/22	0/1/1/1
4	MAN	D	506	4	-	0/2/19/22	0/1/1/1
4	MAN	D	507	4	-	0/2/19/22	0/1/1/1

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	MAN	D	508	4	-	0/2/19/22	0/1/1/1
4	MAN	D	509	4	-	0/2/19/22	0/1/1/1
4	MAN	D	510	4	-	0/2/19/22	0/1/1/1
4	MAN	D	511	4	-	0/2/19/22	0/1/1/1

The worst 5 of 16 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	507	MAN	O3-C3	-3.39	1.34	1.43
3	D	501	NAG	O7-C7	2.80	1.29	1.23
5	B	509	MAN	O2-C2	-2.77	1.37	1.43
3	B	501	NAG	O7-C7	2.74	1.29	1.23
4	A	504	NAG	O5-C5	-2.62	1.40	1.45

The worst 5 of 116 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	C	502	NAG	C3-C2-N2	-6.49	101.88	111.76
4	C	505	BMA	O5-C5-C6	6.07	113.35	106.98
3	C	502	NAG	O5-C5-C4	6.04	118.32	110.65
4	D	511	MAN	O5-C5-C6	5.98	113.25	106.98
4	A	510	MAN	O3-C3-C2	5.56	120.11	109.94

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

5.6 Ligand geometry

Of 40 ligands modelled in this entry, 4 are monoatomic - leaving 36 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	SO4	A	1513	-	4,4,4	0.66	0	6,6,6	0.62	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
6	SO4	A	1514	-	4,4,4	0.70	0	6,6,6	0.60	0
6	SO4	A	1515	-	4,4,4	0.42	0	6,6,6	0.42	0
6	SO4	A	1517	-	4,4,4	0.41	0	6,6,6	0.31	0
6	SO4	A	1518	-	4,4,4	0.37	0	6,6,6	0.95	1 (16%)
6	SO4	A	1519	-	4,4,4	0.47	0	6,6,6	0.31	0
6	SO4	A	1520	-	4,4,4	0.25	0	6,6,6	1.03	0
6	SO4	A	2516	-	4,4,4	0.40	0	6,6,6	0.96	1 (16%)
2	NAG	A	500	1	12,14,15	0.94	0	15,19,21	2.87	9 (60%)
8	GOL	A	5001	-	5,5,5	0.72	0	5,5,5	0.84	0
8	GOL	A	5002	-	5,5,5	0.76	0	5,5,5	1.20	0
8	GOL	A	5003	-	5,5,5	0.58	0	5,5,5	0.34	0
6	SO4	B	2513	-	4,4,4	0.54	0	6,6,6	0.32	0
6	SO4	B	2514	-	4,4,4	0.20	0	6,6,6	0.44	0
6	SO4	B	2515	-	4,4,4	0.62	0	6,6,6	0.40	0
6	SO4	B	2519	-	4,4,4	0.79	0	6,6,6	0.43	0
6	SO4	B	2520	-	4,4,4	0.36	0	6,6,6	0.69	0
6	SO4	B	3516	-	4,4,4	0.15	0	6,6,6	0.89	0
2	NAG	B	500	1	12,14,15	0.85	0	15,19,21	2.22	1 (6%)
6	SO4	C	3513	-	4,4,4	0.68	0	6,6,6	0.37	0
6	SO4	C	3514	-	4,4,4	0.42	0	6,6,6	0.68	0
6	SO4	C	3515	-	4,4,4	0.36	0	6,6,6	1.01	0
6	SO4	C	3517	-	4,4,4	0.31	0	6,6,6	0.83	0
6	SO4	C	3518	-	4,4,4	0.46	0	6,6,6	1.05	1 (16%)
6	SO4	C	3519	-	4,4,4	0.37	0	6,6,6	0.46	0
6	SO4	C	3520	-	4,4,4	0.19	0	6,6,6	0.36	0
6	SO4	C	3521	-	4,4,4	0.93	0	6,6,6	0.65	0
6	SO4	C	4516	-	4,4,4	0.23	0	6,6,6	0.61	0
2	NAG	C	500	1	12,14,15	0.60	0	15,19,21	1.83	2 (13%)
6	SO4	D	1516	-	4,4,4	0.28	0	6,6,6	0.62	0
6	SO4	D	4513	-	4,4,4	0.79	0	6,6,6	0.37	0
6	SO4	D	4514	-	4,4,4	0.73	0	6,6,6	0.73	0
6	SO4	D	4515	-	4,4,4	0.52	0	6,6,6	0.82	0
6	SO4	D	4519	-	4,4,4	0.96	0	6,6,6	0.81	0
6	SO4	D	4520	-	4,4,4	0.48	0	6,6,6	0.54	0
2	NAG	D	500	1	12,14,15	1.10	1 (8%)	15,19,21	1.71	5 (33%)

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	SO4	A	1513	-	-	0/0/0/0	0/0/0/0
6	SO4	A	1514	-	-	0/0/0/0	0/0/0/0
6	SO4	A	1515	-	-	0/0/0/0	0/0/0/0
6	SO4	A	1517	-	-	0/0/0/0	0/0/0/0
6	SO4	A	1518	-	-	0/0/0/0	0/0/0/0
6	SO4	A	1519	-	-	0/0/0/0	0/0/0/0
6	SO4	A	1520	-	-	0/0/0/0	0/0/0/0
6	SO4	A	2516	-	-	0/0/0/0	0/0/0/0
2	NAG	A	500	1	-	0/6/23/26	0/1/1/1
8	GOL	A	5001	-	-	0/4/4/4	0/0/0/0
8	GOL	A	5002	-	-	0/4/4/4	0/0/0/0
8	GOL	A	5003	-	-	0/4/4/4	0/0/0/0
6	SO4	B	2513	-	-	0/0/0/0	0/0/0/0
6	SO4	B	2514	-	-	0/0/0/0	0/0/0/0
6	SO4	B	2515	-	-	0/0/0/0	0/0/0/0
6	SO4	B	2519	-	-	0/0/0/0	0/0/0/0
6	SO4	B	2520	-	-	0/0/0/0	0/0/0/0
6	SO4	B	3516	-	-	0/0/0/0	0/0/0/0
2	NAG	B	500	1	-	0/6/23/26	0/1/1/1
6	SO4	C	3513	-	-	0/0/0/0	0/0/0/0
6	SO4	C	3514	-	-	0/0/0/0	0/0/0/0
6	SO4	C	3515	-	-	0/0/0/0	0/0/0/0
6	SO4	C	3517	-	-	0/0/0/0	0/0/0/0
6	SO4	C	3518	-	-	0/0/0/0	0/0/0/0
6	SO4	C	3519	-	-	0/0/0/0	0/0/0/0
6	SO4	C	3520	-	-	0/0/0/0	0/0/0/0
6	SO4	C	3521	-	-	0/0/0/0	0/0/0/0
6	SO4	C	4516	-	-	0/0/0/0	0/0/0/0
2	NAG	C	500	1	-	0/6/23/26	0/1/1/1
6	SO4	D	1516	-	-	0/0/0/0	0/0/0/0
6	SO4	D	4513	-	-	0/0/0/0	0/0/0/0
6	SO4	D	4514	-	-	0/0/0/0	0/0/0/0
6	SO4	D	4515	-	-	0/0/0/0	0/0/0/0
6	SO4	D	4519	-	-	0/0/0/0	0/0/0/0
6	SO4	D	4520	-	-	0/0/0/0	0/0/0/0
2	NAG	D	500	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)
2	D	500	NAG	C4-C5	2.32	1.58	1.53

The worst 5 of 20 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed($^{\circ}$)	Ideal($^{\circ}$)
2	B	500	NAG	O5-C5-C6	7.57	114.92	106.98
2	A	500	NAG	O5-C5-C6	7.16	114.50	106.98
2	C	500	NAG	O5-C5-C6	5.40	112.65	106.98
2	A	500	NAG	O3-C3-C4	-4.01	101.36	110.35
2	A	500	NAG	O5-C5-C4	-3.22	106.56	110.65

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

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	388/388 (100%)	-0.86	2 (0%) 88 90	11, 18, 31, 44	5 (1%)
1	B	388/388 (100%)	-0.70	6 (1%) 70 71	12, 20, 37, 53	5 (1%)
1	C	388/388 (100%)	-0.73	4 (1%) 79 80	12, 20, 43, 58	0
1	D	388/388 (100%)	-0.74	11 (2%) 50 51	11, 18, 38, 58	3 (0%)
All	All	1552/1552 (100%)	-0.76	23 (1%) 72 71	11, 19, 38, 58	13 (0%)

The worst 5 of 23 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	248	THR	9.2
1	D	250	PRO	5.4
1	D	247	ALA	5.3
1	B	346	ASN	4.5
1	D	249	GLY	4.2

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 ⓘ

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	RSR	LLDF	B-factors(\AA^2)	Q<0.9
5	MAN	B	512	11/12	0.30	24.37	52,56,57,57	0
4	MAN	A	509	11/12	0.13	21.00	21,26,32,37	0
5	MAN	B	511	11/12	0.30	10.05	44,50,54,56	0
3	NAG	A	501	14/15	0.14	8.99	30,37,41,45	0
3	NAG	B	501	14/15	0.17	8.69	33,39,47,48	0
3	NAG	D	501	14/15	0.17	6.02	25,33,47,47	0
4	MAN	C	509	11/12	0.13	3.34	29,35,38,42	0
5	MAN	B	509	11/12	0.13	3.23	35,39,44,44	0
3	NAG	C	501	14/15	0.18	2.88	33,44,48,58	0
3	NAG	B	502	14/15	0.29	2.83	44,52,57,59	0
5	MAN	B	510	11/12	0.22	2.02	49,52,53,55	0
4	MAN	D	510	11/12	0.13	1.69	35,38,41,43	0
4	MAN	C	506	11/12	0.11	1.20	18,22,24,31	0
5	MAN	B	506	11/12	0.10	1.16	16,20,26,33	0
4	MAN	C	507	11/12	0.12	0.96	18,21,26,27	0
5	MAN	B	507	11/12	0.10	0.94	22,28,32,33	0
4	NAG	A	503	14/15	0.09	0.85	15,18,27,29	0
5	NAG	B	504	14/15	0.10	0.84	17,22,26,38	0
4	MAN	C	508	11/12	0.09	0.71	16,19,22,22	0
4	NAG	D	503	14/15	0.09	0.40	17,22,28,30	0
4	NAG	A	504	14/15	0.07	0.40	15,18,24,24	0
4	MAN	D	509	11/12	0.11	0.34	24,27,29,31	0
4	MAN	A	506	11/12	0.08	0.21	16,19,25,26	0
4	MAN	A	508	11/12	0.07	0.21	15,17,18,21	0
5	MAN	B	508	11/12	0.08	0.20	18,25,26,26	0
5	NAG	B	503	14/15	0.09	0.06	18,22,25,26	0
4	BMA	C	505	11/12	0.08	0.05	18,23,25,29	0
4	MAN	D	507	11/12	0.08	-0.07	18,20,24,27	0
4	NAG	D	504	14/15	0.07	-0.17	17,21,25,31	0
4	NAG	C	503	14/15	0.08	-0.35	20,21,25,26	0
4	MAN	D	506	11/12	0.08	-0.57	16,21,28,34	0
4	MAN	D	511	11/12	0.17	-0.68	32,35,38,39	0
4	MAN	D	508	11/12	0.07	-0.74	19,21,23,23	0
4	BMA	A	505	11/12	0.05	-0.92	14,19,22,23	0
5	BMA	B	505	11/12	0.06	-0.95	20,22,28,29	0
4	NAG	C	504	14/15	0.06	-1.09	17,21,27,32	0
4	MAN	A	507	11/12	0.06	-1.10	18,20,22,24	0
4	BMA	D	505	11/12	0.06	-1.86	20,24,26,26	0
3	NAG	C	502	14/15	0.44	-	63,67,69,69	0
4	MAN	C	511	11/12	0.22	-	32,39,44,46	0
4	MAN	C	510	11/12	0.26	-	46,51,56,57	0
4	MAN	A	511	11/12	0.16	-	32,36,38,43	0
3	NAG	D	502	14/15	0.31	-	46,51,55,56	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSR	LLDF	B-factors(\AA^2)	Q<0.9
3	NAG	A	502	14/15	0.32	-	43,51,56,58	0
4	MAN	A	510	11/12	0.28	-	43,48,50,53	0

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	RSR	LLDF	B-factors(\AA^2)	Q<0.9
6	SO4	A	1519	5/5	0.23	335.00	40,43,45,45	0
6	SO4	B	2519	5/5	0.25	78.00	46,46,48,52	0
6	SO4	D	1516	5/5	0.20	29.97	52,55,57,58	0
6	SO4	A	1517	5/5	0.28	25.82	53,56,58,59	0
6	SO4	C	3517	5/5	0.28	19.76	59,61,64,66	0
6	SO4	B	3516	5/5	0.19	13.99	72,73,75,76	0
6	SO4	C	3519	5/5	0.27	13.85	57,61,61,63	0
6	SO4	A	2516	5/5	0.21	12.11	60,61,62,64	0
6	SO4	D	4519	5/5	0.22	10.31	40,42,44,45	0
8	GOL	A	5002	6/6	0.13	7.98	34,40,43,44	0
6	SO4	A	1518	5/5	0.28	7.60	76,77,81,81	0
6	SO4	C	4516	5/5	0.11	7.17	56,57,58,60	0
6	SO4	C	3520	5/5	0.28	5.35	63,64,65,67	0
6	SO4	C	3515	5/5	0.26	5.03	55,56,58,61	0
2	NAG	A	500	14/15	0.22	4.19	40,48,52,52	0
6	SO4	D	4515	5/5	0.22	3.87	50,54,58,59	0
6	SO4	A	1515	5/5	0.23	3.47	50,54,57,60	0
6	SO4	B	2515	5/5	0.25	3.33	53,55,57,59	0
2	NAG	C	500	14/15	0.26	2.83	47,55,58,59	0
2	NAG	D	500	14/15	0.24	2.72	45,50,57,58	0
6	SO4	A	1520	5/5	0.19	2.58	60,63,64,65	0
6	SO4	D	4520	5/5	0.22	2.42	63,64,64,64	0
6	SO4	B	2513	5/5	0.22	2.27	70,71,73,75	0
8	GOL	A	5001	6/6	0.21	2.24	55,57,59,59	0
2	NAG	B	500	14/15	0.24	1.96	45,50,59,61	0
6	SO4	B	2520	5/5	0.20	1.55	47,50,52,52	0
6	SO4	C	3513	5/5	0.23	1.42	63,65,66,67	0
8	GOL	A	5003	6/6	0.18	1.36	56,56,58,59	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSR	LLDF	B-factors(\AA^2)	Q<0.9
6	SO4	D	4513	5/5	0.18	1.12	55,59,62,63	0
6	SO4	D	4514	5/5	0.14	0.90	35,48,50,52	0
6	SO4	A	1514	5/5	0.07	0.21	29,29,32,32	0
6	SO4	B	2514	5/5	0.14	0.08	40,41,45,45	0
6	SO4	C	3514	5/5	0.13	-0.01	32,38,41,43	0
6	SO4	A	1513	5/5	0.12	-0.46	40,41,47,47	0
6	SO4	C	3518	5/5	0.06	-0.77	21,22,24,25	0
7	CL	A	1522	1/1	0.06	-0.95	46,46,46,46	0
7	CL	C	3522	1/1	0.05	-1.40	50,50,50,50	0
7	CL	B	2522	1/1	0.04	-1.45	45,45,45,45	0
7	CL	D	4522	1/1	0.06	-1.45	46,46,46,46	0
6	SO4	C	3521	5/5	0.07	-1.75	17,20,21,23	0

6.5 Other polymers ⓘ

There are no such residues in this entry.