



wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 8, 2017 – 07:04 AM EDT

PDB ID : 2B4C
Title : Crystal structure of HIV-1 JR-FL gp120 core protein containing the third variable region (V3) complexed with CD4 and the X5 antibody
Authors : Huang, C.; Tang, M.; Zhang, M.Y.; Majeed, S.; Montabana, E.; Stanfield, R.L.; Dimitrov, D.S.; Korber, B.; Sodroski, J.; Wilson, I.A.; Wyatt, R.; Kwong, P.D.
Deposited on : unknown
Resolution : 3.30 Å(reported)

This is a wwPDB X-ray Structure 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/validation/2016/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

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

MolProbity	:	4.02b-467
Mogul	:	1.7.2 (RC1), CSD as538be (2017)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20029824
Percentile statistics	:	20161228.v01 (using entries in the PDB archive December 28th 2016)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	rb-20029824

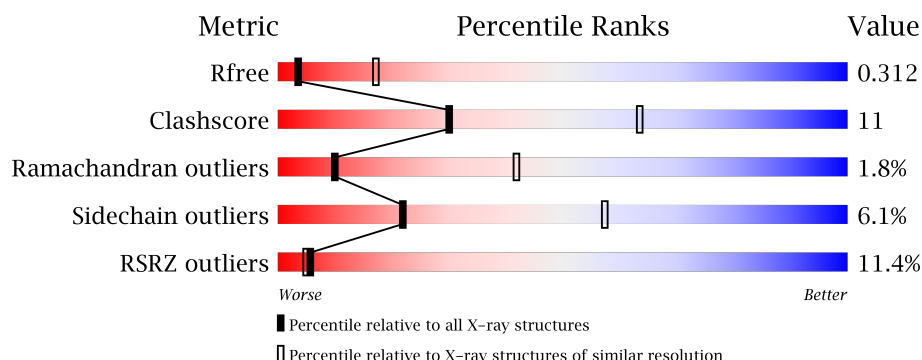
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.30 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}	100719	1034 (3.36-3.24)
Clashscore	112137	1100 (3.36-3.24)
Ramachandran outliers	110173	1081 (3.36-3.24)
Sidechain outliers	110143	1080 (3.36-3.24)
RSRZ outliers	101464	1039 (3.36-3.24)

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. 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	G	344	
2	C	181	
3	L	215	
4	H	235	

2 Entry composition

There are 9 unique types of molecules in this entry. The entry contains 7543 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 envelope glycoprotein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	G	339	Total	C	N	O	S	0	4	0
			2681	1674	477	509	21			

- Molecule 2 is a protein called T-cell surface glycoprotein CD4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	175	Total	C	N	O	S	0	0	0
			1363	851	239	269	4			

- Molecule 3 is a protein called anti-HIV-1 gp120 immunoglobulin X5 light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	L	215	Total	C	N	O	S	0	0	0
			1636	1018	280	333	5			

- Molecule 4 is a protein called anti-HIV-1 gp120 immunoglobulin X5 heavy chain.

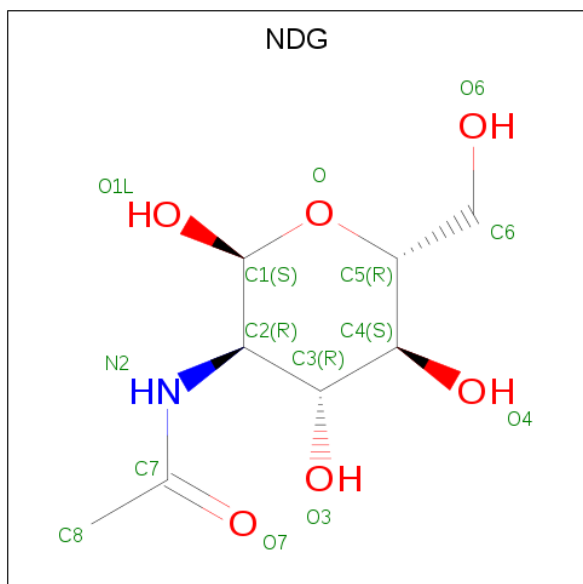
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	H	233	Total	C	N	O	S	0	0	0
			1742	1098	290	346	8			

- Molecule 5 is N-ACETYL-D-GLUCOSAMINE (three-letter code: NAG) (formula: C₈H₁₅NO₆).



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

- Molecule 6 is 2-(ACETYLAMINO)-2-DEOXY-A-D-GLUCOPYRANOSE (three-letter code: NDG) (formula: C₈H₁₅NO₆).



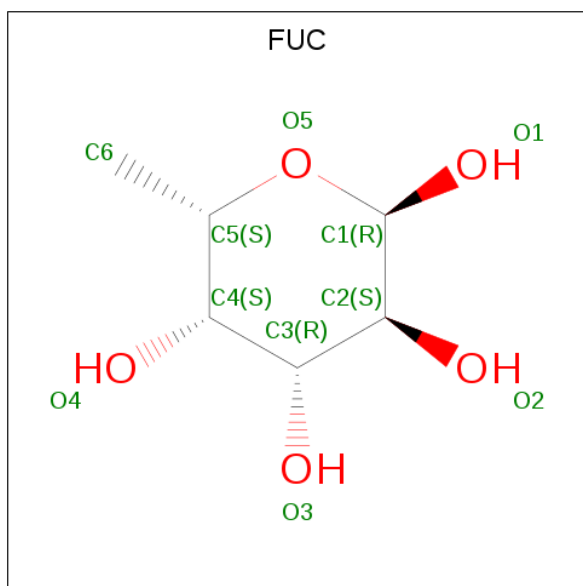
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
6	G	1	Total	C	N	O	0	0
			14	8	1	5		

Continued on next page...

Continued from previous page...

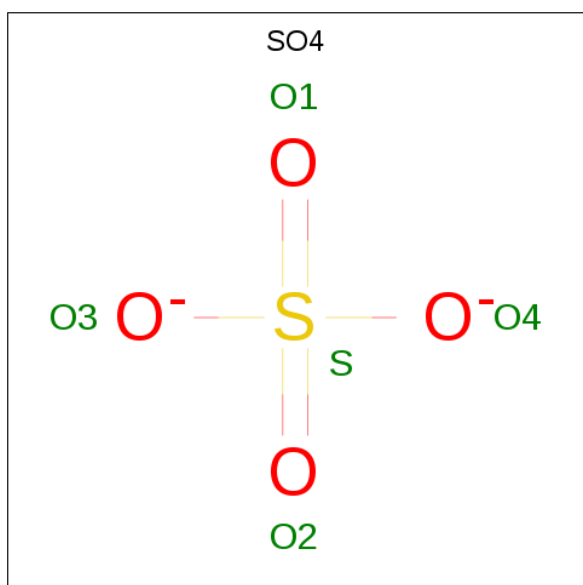
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
6	G	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 7 is ALPHA-L-FUCOSE (three-letter code: FUC) (formula: C₆H₁₂O₅).



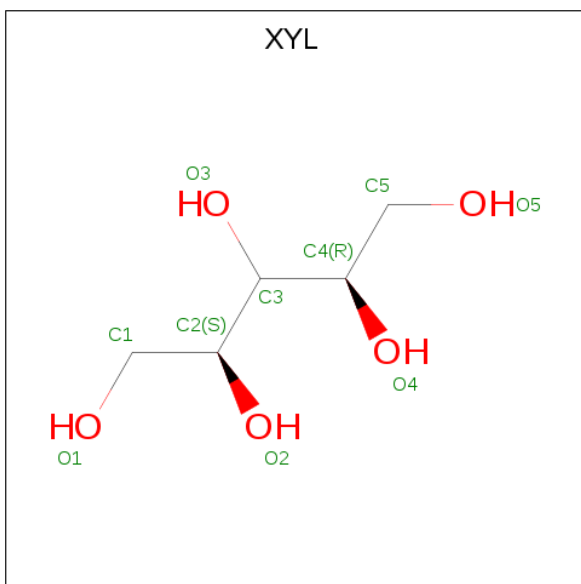
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	G	1	Total	C	O	0	0
			10	6	4		

- Molecule 8 is SULFATE ION (three-letter code: SO₄) (formula: O₄S).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	G	1	Total O S 5 4 1	0	0
8	L	1	Total O S 5 4 1	0	0
8	H	1	Total O S 5 4 1	0	0
8	H	1	Total O S 5 4 1	0	0
8	H	1	Total O S 5 4 1	0	0

- Molecule 9 is D-XYLITOL (three-letter code: XYL) (formula: $C_5H_{12}O_5$).

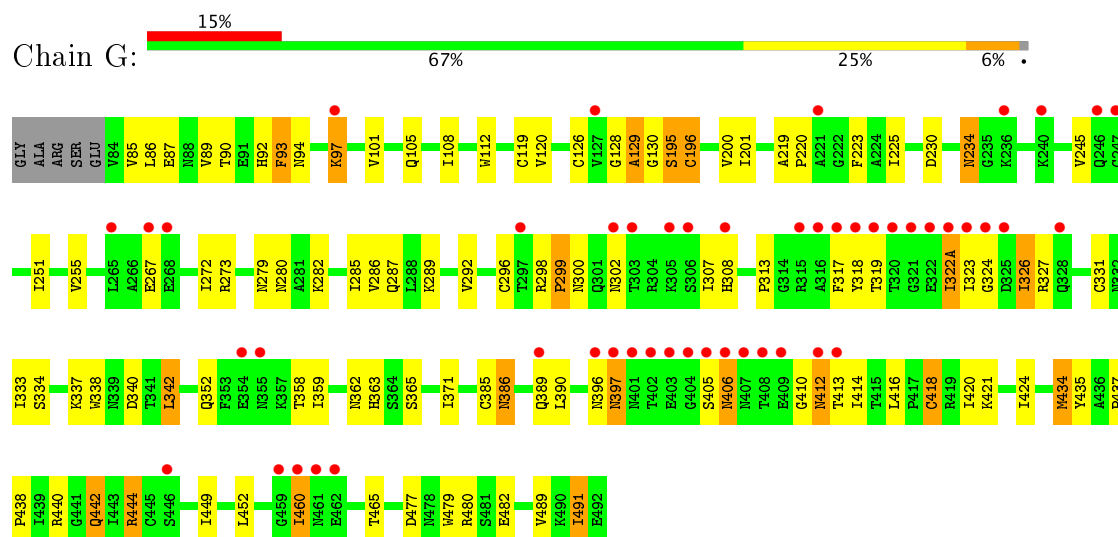


Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
9	H	1	Total C O 10 5 5	0	0
9	H	1	Total C O 10 5 5	0	0
9	H	1	Total C O 10 5 5	0	0

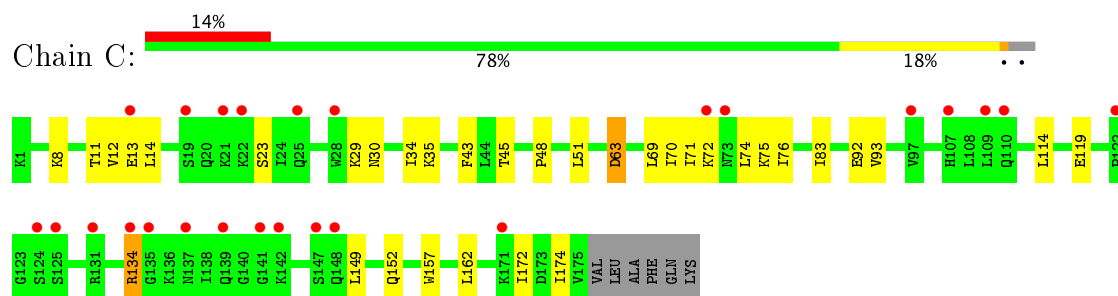
3 Residue-property plots

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

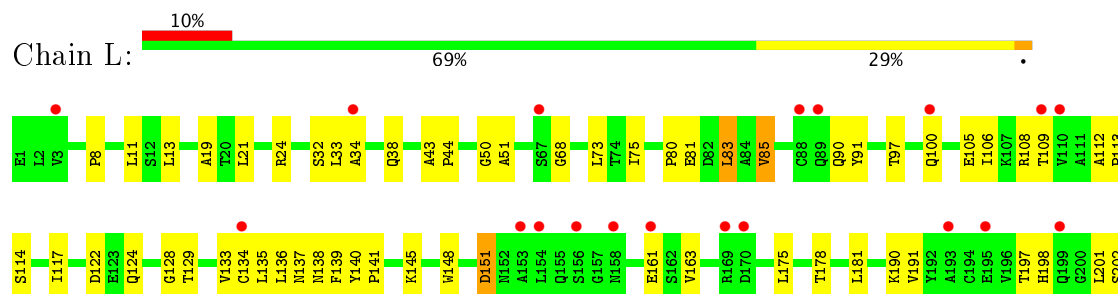
- Molecule 1: envelope glycoprotein

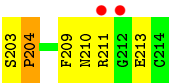


- Molecule 2: T-cell surface glycoprotein CD4

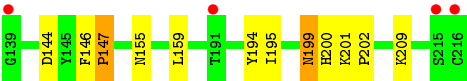
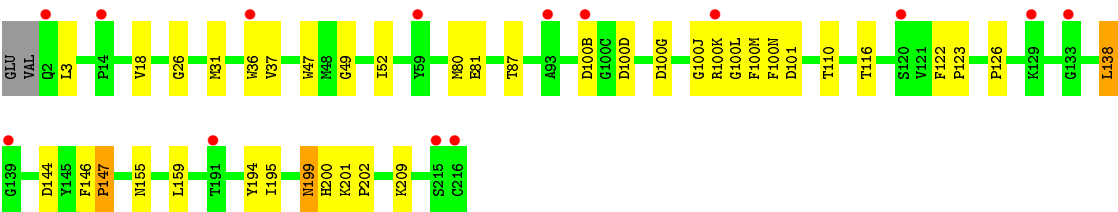
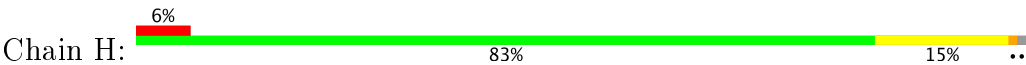


- Molecule 3: anti-HIV-1 gp120 immunoglobulin X5 light chain





● Molecule 4: anti-HIV-1 gp120 immunoglobulin X5 heavy chain



4 Data and refinement statistics

Property	Value	Source
Space group	P 6 2 2	Depositor
Cell constants a, b, c, α , β , γ	225.99Å 225.99Å 97.97Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	20.00 – 3.30 19.88 – 3.31	Depositor EDS
% Data completeness (in resolution range)	86.2 (20.00-3.30) 86.3 (19.88-3.31)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	0.08	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.43 (at 3.29Å)	Xtriage
Refinement program	CNS	Depositor
R, R_{free}	0.319 , 0.349 0.293 , 0.312	Depositor DCC
R_{free} test set	1928 reflections (11.06%)	DCC
Wilson B-factor (Å ²)	134.8	Xtriage
Anisotropy	0.213	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.31 , 56.3	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.90	EDS
Total number of atoms	7543	wwPDB-VP
Average B, all atoms (Å ²)	114.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.02% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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: FUC, XYZ, SO4, NDG, 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	G	0.33	1/2734 (0.0%)	0.52	0/3701
2	C	0.26	0/1382	0.53	0/1863
3	L	0.28	0/1669	0.53	0/2264
4	H	0.28	0/1788	0.53	0/2432
All	All	0.30	1/7573 (0.0%)	0.53	0/10260

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	G	87	GLU	C-N	-7.20	1.17	1.34

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	G	2681	0	2628	79	0
2	C	1363	0	1389	28	0
3	L	1636	0	1591	38	0
4	H	1742	0	1662	22	0
5	G	28	0	26	0	0
6	G	28	0	25	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
7	G	10	0	10	0	0
8	G	5	0	0	0	0
8	H	15	0	0	0	0
8	L	5	0	0	0	0
9	H	30	0	36	0	0
All	All	7543	0	7367	158	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 11.

The worst 5 of 158 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:G:326:ILE:HD13	1:G:326:ILE:H	1.35	0.92
4:H:126:PRO:HG3	4:H:138:LEU:HB3	1.52	0.91
1:G:371:ILE:HD13	2:C:45:THR:HG22	1.53	0.90
2:C:76:ILE:H	2:C:76:ILE:HD12	1.35	0.90
3:L:198:HIS:HB3	3:L:201:LEU:HB3	1.66	0.78

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	G	341/344 (99%)	290 (85%)	40 (12%)	11 (3%)	5	29
2	C	173/181 (96%)	158 (91%)	15 (9%)	0	100	100
3	L	213/215 (99%)	194 (91%)	14 (7%)	5 (2%)	7	36
4	H	231/235 (98%)	208 (90%)	21 (9%)	2 (1%)	20	55
All	All	958/975 (98%)	850 (89%)	90 (9%)	18 (2%)	10	41

5 of 18 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	G	129	ALA
1	G	299	PRO
1	G	323	ILE
1	G	412	ASN
3	L	202	SER

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	G	302/303 (100%)	275 (91%)	27 (9%)	11	39
2	C	159/164 (97%)	154 (97%)	5 (3%)	45	75
3	L	182/182 (100%)	173 (95%)	9 (5%)	29	65
4	H	192/194 (99%)	182 (95%)	10 (5%)	27	63
All	All	835/843 (99%)	784 (94%)	51 (6%)	22	57

5 of 51 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	G	444	ARG
2	C	119	GLU
4	H	138	LEU
1	G	460	ILE
2	C	134	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 21 such sidechains are listed below:

Mol	Chain	Res	Type
1	G	389	GLN
1	G	397	ASN
3	L	90	GLN
1	G	386	ASN
3	L	124	GLN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

5.6 Ligand geometry ⓘ

13 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$
8	SO4	G	5	-	4,4,4	0.14	0	6,6,6	0.05	0
5	NAG	G	762	1	14,14,15	0.56	0	15,19,21	0.73	0
6	NDG	G	776	1,7	14,14,15	0.74	0	15,19,21	0.63	0
7	FUC	G	777	6	9,10,11	0.61	0	13,14,16	0.76	0
6	NDG	G	832	1	14,14,15	0.56	0	15,19,21	0.82	1 (6%)
5	NAG	G	862	1	14,14,15	0.57	0	15,19,21	0.73	0
9	XYL	H	217	-	9,9,9	1.22	1 (11%)	11,11,11	0.84	0
9	XYL	H	218	-	9,9,9	1.22	1 (11%)	11,11,11	0.82	0
9	XYL	H	219	-	9,9,9	1.22	0	11,11,11	0.85	0
8	SO4	H	220	-	4,4,4	0.37	0	6,6,6	0.05	0
8	SO4	H	221	-	4,4,4	0.37	0	6,6,6	0.05	0
8	SO4	H	222	-	4,4,4	0.13	0	6,6,6	0.06	0
8	SO4	L	215	-	4,4,4	0.37	0	6,6,6	0.06	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
8	SO4	G	5	-	-	0/0/0/0	0/0/0/0
5	NAG	G	762	1	-	0/6/23/26	0/1/1/1
6	NDG	G	776	1,7	-	0/6/23/26	0/1/1/1
7	FUC	G	777	6	-	0/0/17/20	0/1/1/1
6	NDG	G	832	1	-	0/6/23/26	0/1/1/1
5	NAG	G	862	1	-	0/6/23/26	0/1/1/1
9	XYL	H	217	-	-	0/12/12/12	0/0/0/0
9	XYL	H	218	-	-	0/12/12/12	0/0/0/0
9	XYL	H	219	-	-	0/12/12/12	0/0/0/0
8	SO4	H	220	-	-	0/0/0/0	0/0/0/0
8	SO4	H	221	-	-	0/0/0/0	0/0/0/0
8	SO4	H	222	-	-	0/0/0/0	0/0/0/0
8	SO4	L	215	-	-	0/0/0/0	0/0/0/0

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
9	H	217	XYL	C4-C3	2.05	1.57	1.53
9	H	218	XYL	C4-C3	2.10	1.57	1.53

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	G	832	NDG	C2-N2-C7	-2.24	119.67	122.94

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](#)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	G	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	G	87:GLU	C	88:ASN	N	1.17

6 Fit of model and data [i](#)

6.1 Protein, DNA and RNA chains [i](#)

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	G	339/344 (98%)	0.60	50 (14%) 3 2	58, 122, 148, 160	0
2	C	175/181 (96%)	0.67	25 (14%) 3 3	100, 139, 161, 166	0
3	L	215/215 (100%)	0.54	21 (9%) 8 7	65, 111, 141, 154	0
4	H	233/235 (99%)	0.28	14 (6%) 23 21	48, 83, 121, 150	0
All	All	962/975 (98%)	0.52	110 (11%) 6 5	48, 116, 154, 166	0

The worst 5 of 110 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	G	404	GLY	7.6
1	G	460	ILE	6.7
1	G	406	ASN	5.8
1	G	127	VAL	5.7
2	C	122	PRO	4.8

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 carbohydrates 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. 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	RSCC	RSR	LLDF	B-factors(\AA^2)	Q<0.9
5	NAG	G	762	14/15	0.91	0.36	0.52	105,108,110,110	0
9	XYL	H	217	10/10	0.52	0.32	0.08	145,150,150,150	0
8	SO4	H	221	5/5	0.72	0.23	-0.17	189,189,190,190	0
9	XYL	H	218	10/10	0.79	0.20	-0.47	134,136,137,137	0
6	NDG	G	832	14/15	0.82	0.28	-0.88	141,144,147,147	0
6	NDG	G	776	14/15	0.73	0.41	-	138,139,140,141	0
5	NAG	G	862	14/15	0.77	0.35	-	135,138,140,140	0
8	SO4	L	215	5/5	0.84	0.13	-	186,187,187,187	0
8	SO4	H	222	5/5	0.96	0.16	-	143,143,143,143	5
7	FUC	G	777	10/11	0.81	0.35	-	141,142,143,143	0
8	SO4	G	5	5/5	0.92	0.14	-	163,163,164,164	5
8	SO4	H	220	5/5	0.71	0.17	-	161,161,161,161	0
9	XYL	H	219	10/10	0.65	0.27	-	153,155,156,157	0

6.5 Other polymers [i](#)

There are no such residues in this entry.