



wwPDB X-ray Structure Validation Summary Report ⓘ

Dec 13, 2021 – 06:13 pm GMT

PDB ID : 7NIT
Title : X-ray structure of a multidomain BbgIII from Bifidobacterium bifidum
Authors : Moroz, O.V.; Blagova, E.; Lebedev, A.A.; Sanchez Rodriguez, F.; Rigden, D.J.; Tams, J.W.; Wilting, R.; Vester, J.K.; Longhin, E.; Krogh, K.B.R.; Pache, R.A.; Davies, G.J.; Wilson, K.S.
Deposited on : 2021-02-14
Resolution : 2.89 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

<https://www.wwpdb.org/validation/2017/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	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	FAILED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.24

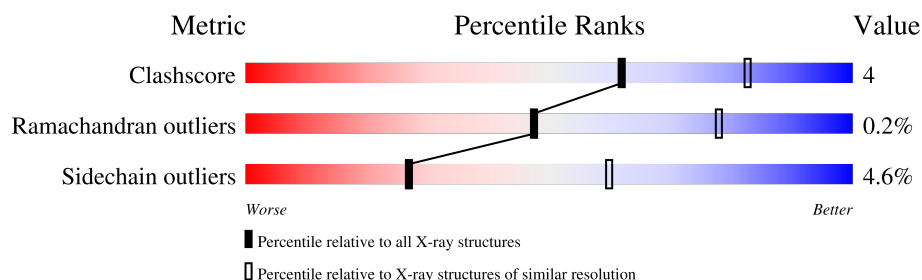
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.89 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))
Clashscore	141614	2172 (2.90-2.90)
Ramachandran outliers	138981	2115 (2.90-2.90)
Sidechain outliers	138945	2117 (2.90-2.90)

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 $\leq 5\%$.

Note EDS failed to run properly.

Mol	Chain	Length	Quality of chain
1	A	1304	84% 10% • 5%
1	B	1304	82% 12% • 5%
1	C	1304	83% 11% • 5%
1	D	1304	84% 11% • •
1	E	1304	83% 11% • 5%
1	F	1304	83% 11% • 5%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit crite-

ria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	GOL	B	2302	-	-	X	-
3	GOL	F	2302	-	-	X	-

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 55867 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 Beta-galactosidase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	1239	Total	C	N	O	S	0	0	0
			9306	5807	1585	1897	17			
1	B	1235	Total	C	N	O	S	0	0	0
			9246	5768	1574	1888	16			
1	C	1243	Total	C	N	O	S	0	0	0
			9326	5818	1585	1906	17			
1	D	1253	Total	C	N	O	S	0	0	0
			9390	5856	1601	1916	17			
1	E	1235	Total	C	N	O	S	0	0	0
			9195	5733	1567	1879	16			
1	F	1244	Total	C	N	O	S	0	0	0
			9324	5816	1586	1905	17			

There are 6 discrepancies between the modelled and reference sequences:

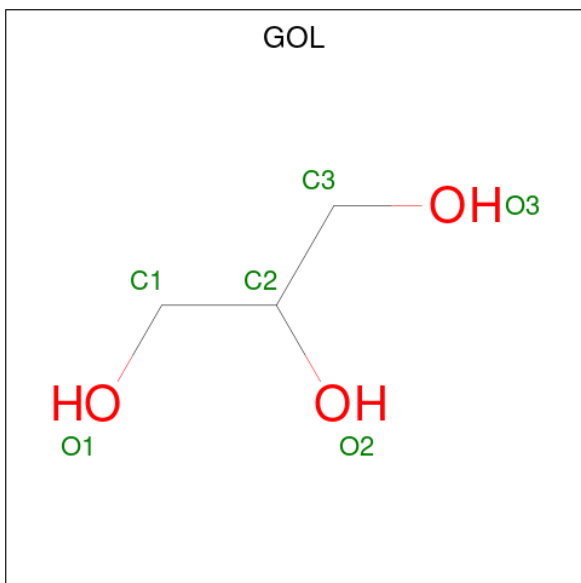
Chain	Residue	Modelled	Actual	Comment	Reference
A	1165	GLU	ASP	conflict	UNP A0A415C3Q2
B	1165	GLU	ASP	conflict	UNP A0A415C3Q2
C	1165	GLU	ASP	conflict	UNP A0A415C3Q2
D	1165	GLU	ASP	conflict	UNP A0A415C3Q2
E	1165	GLU	ASP	conflict	UNP A0A415C3Q2
F	1165	GLU	ASP	conflict	UNP A0A415C3Q2

- Molecule 2 is SULFATE ION (three-letter code: SO4) (formula: O₄S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	O	S	0	0
			5	4	1		
2	C	1	Total	O	S	0	0
			5	4	1		
2	D	1	Total	O	S	0	0
			5	4	1		
2	F	1	Total	O	S	0	0
			5	4	1		

- Molecule 3 is GLYCEROL (three-letter code: GOL) (formula: C₃H₈O₃).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C O 6 3 3	0	0
3	A	1	Total C O 6 3 3	0	0
3	B	1	Total C O 6 3 3	0	0
3	C	1	Total C O 6 3 3	0	0
3	D	1	Total C O 6 3 3	0	0
3	D	1	Total C O 6 3 3	0	0
3	E	1	Total C O 6 3 3	0	0
3	F	1	Total C O 6 3 3	0	0

- Molecule 4 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total Ca 1 1	0	0
4	B	1	Total Ca 1 1	0	0
4	C	1	Total Ca 1 1	0	0
4	D	1	Total Ca 1 1	0	0
4	E	1	Total Ca 1 1	0	0
4	F	1	Total Ca 1 1	0	0

- Molecule 5 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total O 1 1	0	0
5	B	1	Total O 1 1	0	0
5	C	1	Total O 1 1	0	0
5	D	1	Total O 1 1	0	0

Continued on next page...

Continued from previous page...

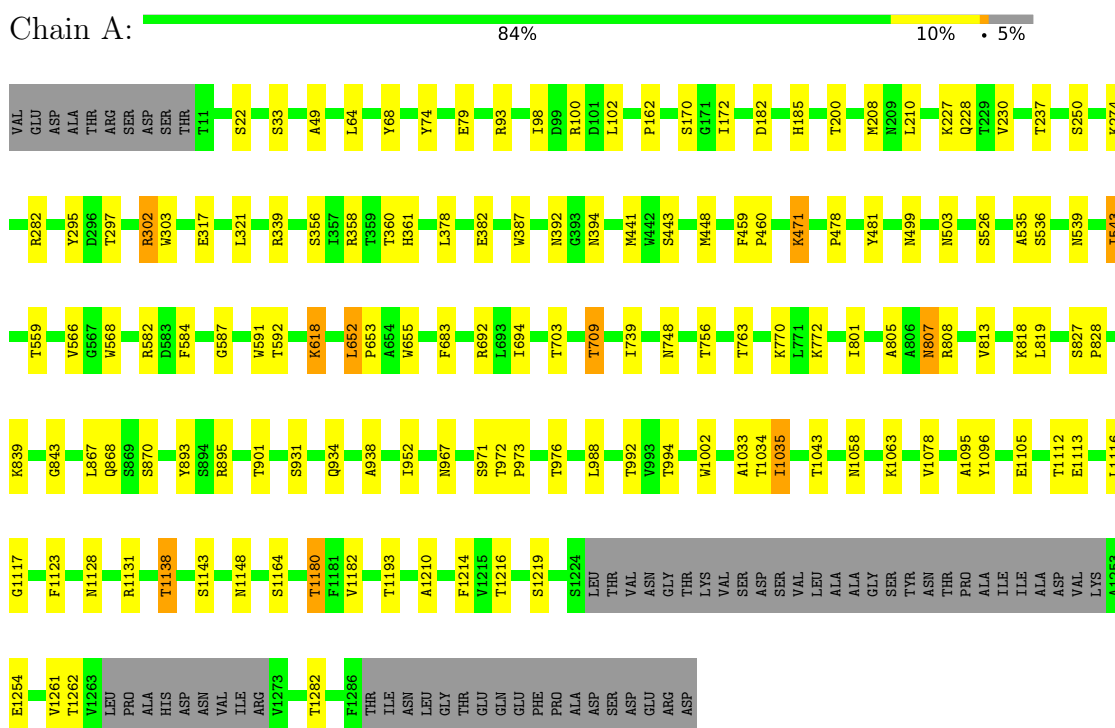
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	E	1	Total	O	0	0
			1	1		
5	F	1	Total	O	0	0
			1	1		

3 Residue-property plots

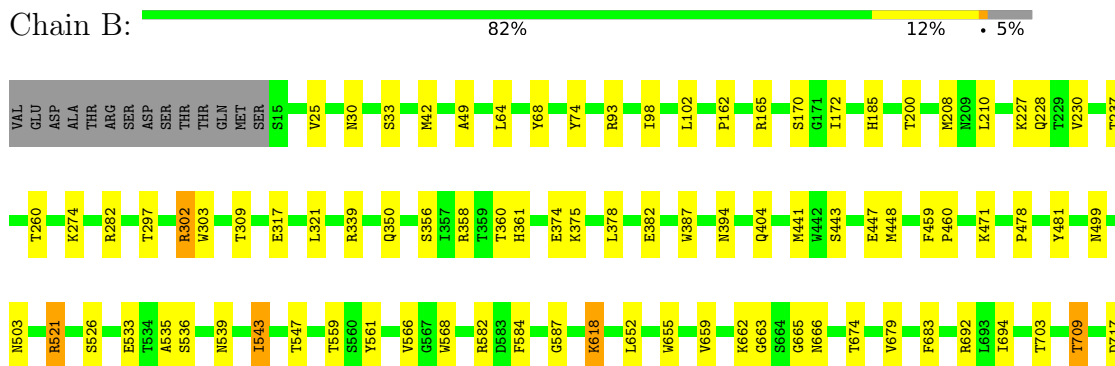
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. Residues are color-coded 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. 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.

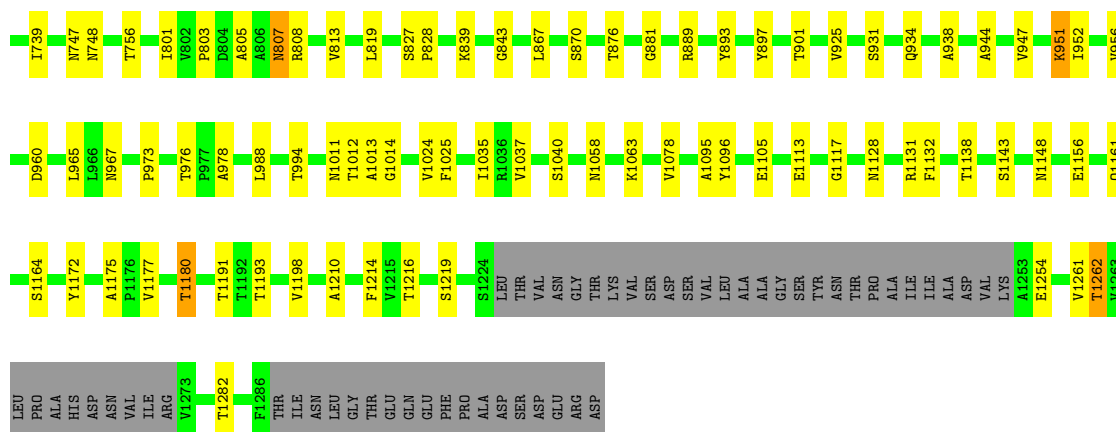
Note EDS failed to run properly.

- Molecule 1: Beta-galactosidase



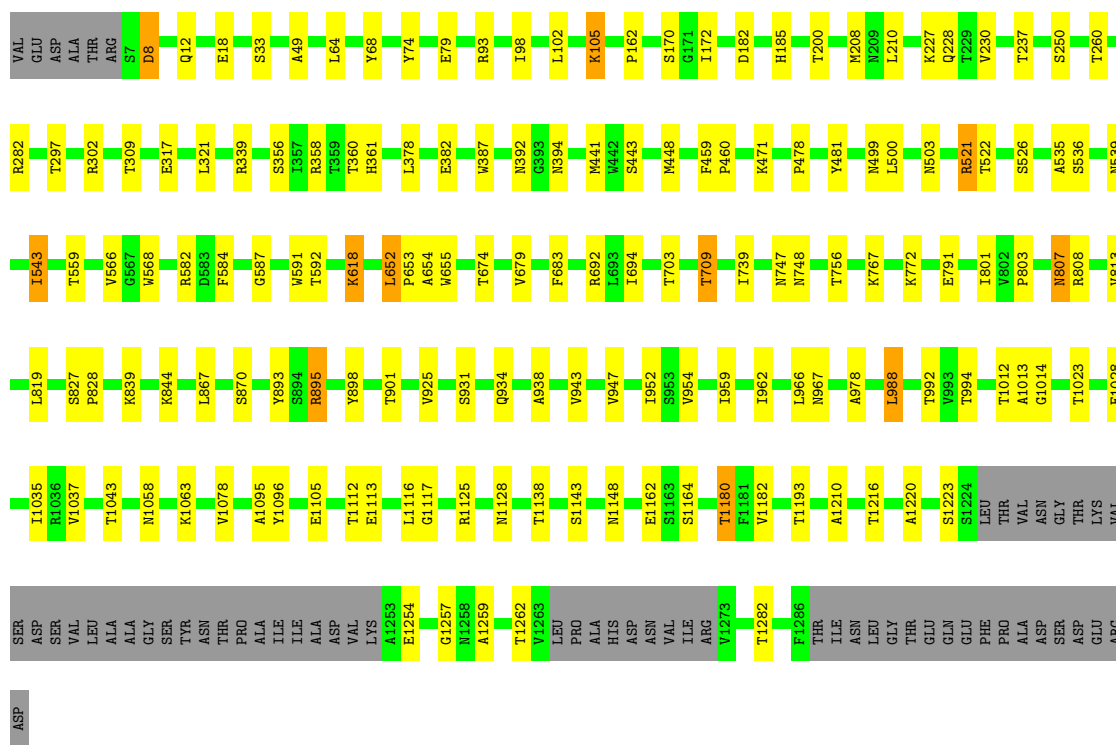
- Molecule 1: Beta-galactosidase





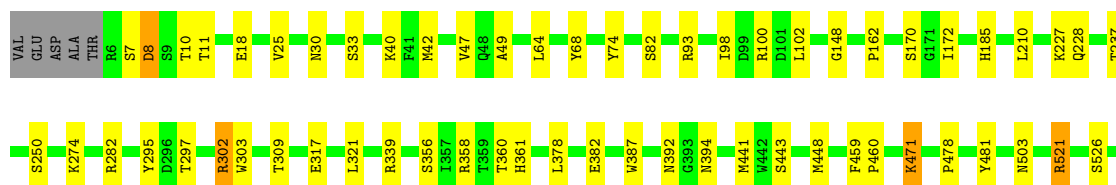
• Molecule 1: Beta-galactosidase

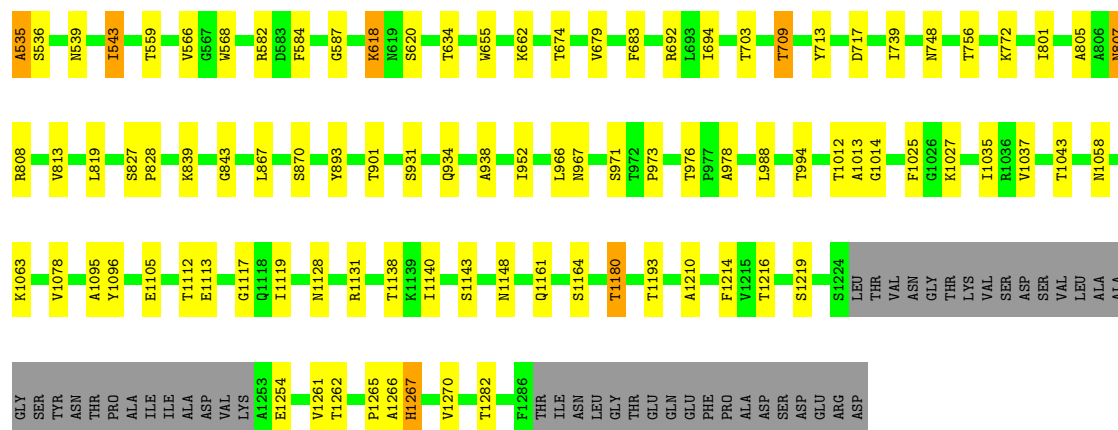
Chain C: 83% 11% 5%



• Molecule 1: Beta-galactosidase

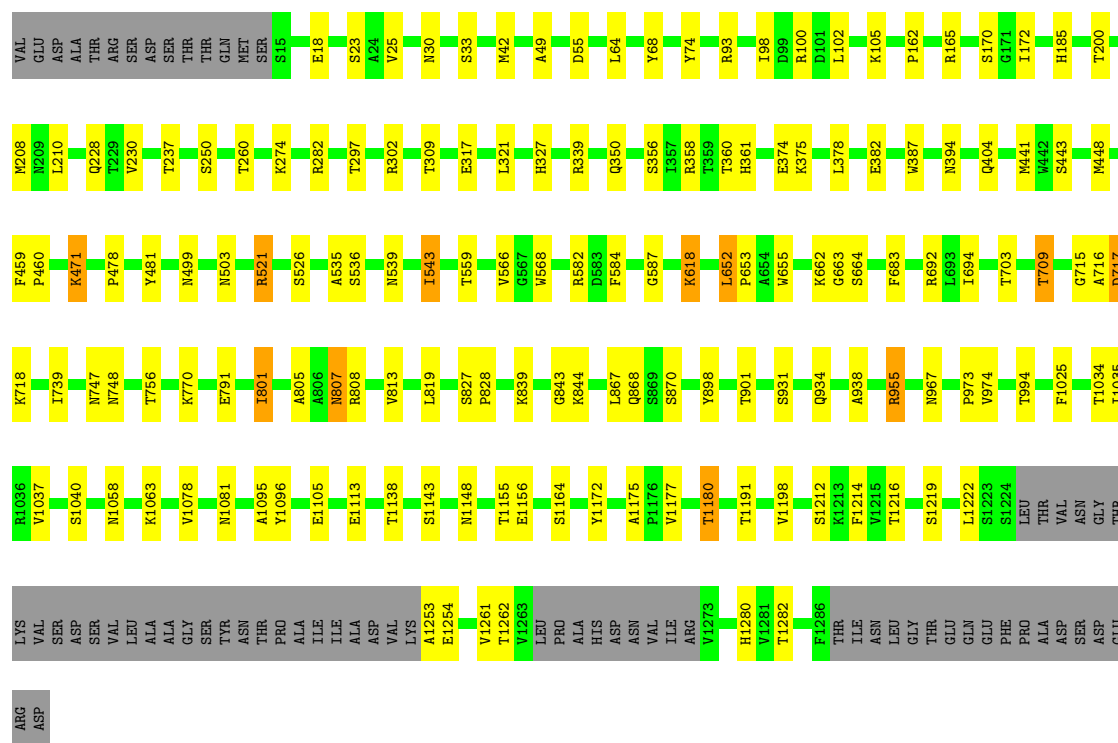
Chain D: 84% 11% 5%





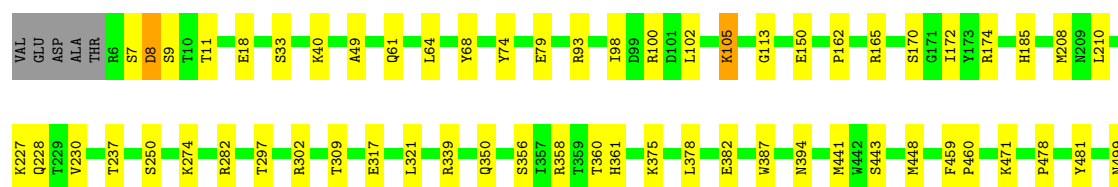
• Molecule 1: Beta-galactosidase

Chain E: 83% 11% • 5%



• Molecule 1: Beta-galactosidase

Chain F: 83% 11% • 5%





4 Data and refinement statistics

EDS failed to run properly - this section is therefore incomplete.

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α , β , γ	116.95Å 130.04Å 200.58Å 86.99° 84.83° 83.79°	Depositor
Resolution (Å)	199.58 – 2.89	Depositor
% Data completeness (in resolution range)	96.2 (199.58-2.89)	Depositor
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.13 (at 2.91Å)	Xtriage
Refinement program	REFMAC 5.8.0267	Depositor
R, R_{free}	0.214 , 0.247	Depositor
Wilson B-factor (Å ²)	88.8	Xtriage
Anisotropy	0.320	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.51$, $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	55867	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 4.81% 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

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: SO4, GOL, CA

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	0.36	0/9503	0.71	0/12962
1	B	0.39	0/9443	0.73	0/12889
1	C	0.37	0/9523	0.71	0/12990
1	D	0.36	0/9589	0.71	0/13082
1	E	0.36	0/9390	0.70	0/12825
1	F	0.33	0/9521	0.68	0/12990
All	All	0.36	0/56969	0.70	0/77738

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	9306	0	8880	70	0
1	B	9246	0	8784	92	0
1	C	9326	0	8895	93	0
1	D	9390	0	8946	82	0
1	E	9195	0	8696	73	0
1	F	9324	0	8877	74	0
2	A	5	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	C	5	0	0	0	0
2	D	5	0	0	0	0
2	F	5	0	0	0	0
3	A	12	0	16	0	0
3	B	6	0	8	5	0
3	C	6	0	8	0	0
3	D	12	0	16	0	0
3	E	6	0	8	1	0
3	F	6	0	8	9	0
4	A	1	0	0	0	0
4	B	1	0	0	0	0
4	C	1	0	0	0	0
4	D	1	0	0	0	0
4	E	1	0	0	0	0
4	F	1	0	0	0	0
5	A	1	0	0	0	0
5	B	1	0	0	0	0
5	C	1	0	0	0	0
5	D	1	0	0	0	0
5	E	1	0	0	0	0
5	F	1	0	0	0	0
All	All	55867	0	53142	474	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 4.

The worst 5 of 474 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:F:879:VAL:HB	1:F:882:THR:HG23	1.33	1.06
1:C:1023:THR:HG22	1:C:1028:GLU:HG3	1.52	0.91
1:C:959:ILE:HG21	1:C:962:ILE:HD11	1.59	0.85
1:B:1014:GLY:O	1:B:1037:VAL:HG22	1.79	0.82
1:E:1155:THR:HG22	1:F:704:THR:HA	1.65	0.79

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	1233/1304 (95%)	1180 (96%)	52 (4%)	1 (0%)	51	82
1	B	1229/1304 (94%)	1180 (96%)	47 (4%)	2 (0%)	47	78
1	C	1237/1304 (95%)	1186 (96%)	50 (4%)	1 (0%)	51	82
1	D	1249/1304 (96%)	1193 (96%)	53 (4%)	3 (0%)	47	78
1	E	1229/1304 (94%)	1179 (96%)	46 (4%)	4 (0%)	41	71
1	F	1238/1304 (95%)	1186 (96%)	49 (4%)	3 (0%)	47	78
All	All	7415/7824 (95%)	7104 (96%)	297 (4%)	14 (0%)	47	78

5 of 14 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	1267	HIS
1	F	9	SER
1	B	663	GLY
1	D	1270	VAL
1	E	663	GLY

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	982/1056 (93%)	938 (96%)	44 (4%)	27	61
1	B	972/1056 (92%)	927 (95%)	45 (5%)	27	60
1	C	986/1056 (93%)	943 (96%)	43 (4%)	28	61

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	D	989/1056 (94%)	944 (95%)	45 (5%)	27	60
1	E	960/1056 (91%)	911 (95%)	49 (5%)	24	56
1	F	983/1056 (93%)	936 (95%)	47 (5%)	25	58
All	All	5872/6336 (93%)	5599 (95%)	273 (5%)	27	60

5 of 273 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	F	100	ARG
1	F	297	THR
1	F	971	SER
1	C	237	THR
1	C	185	HIS

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

Mol	Chain	Res	Type
1	A	896	ASN
1	C	896	ASN
1	E	327	HIS
1	E	711	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 monosaccharides in this entry.

5.6 Ligand geometry ⓘ

Of 18 ligands modelled in this entry, 6 are monoatomic - leaving 12 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$
3	GOL	D	2302	-	5,5,5	0.15	0	5,5,5	0.48	0
3	GOL	E	2302	-	5,5,5	0.23	0	5,5,5	0.38	0
3	GOL	A	2302	-	5,5,5	0.18	0	5,5,5	0.59	0
3	GOL	F	2302	-	5,5,5	0.08	0	5,5,5	0.43	0
3	GOL	A	2303	-	5,5,5	0.20	0	5,5,5	0.61	0
3	GOL	B	2302	-	5,5,5	0.17	0	5,5,5	0.53	0
2	SO4	D	2301	-	4,4,4	0.33	0	6,6,6	0.23	0
3	GOL	C	2302	-	5,5,5	0.16	0	5,5,5	0.46	0
2	SO4	A	2301	-	4,4,4	0.17	0	6,6,6	0.31	0
3	GOL	D	2303	-	5,5,5	0.16	0	5,5,5	0.60	0
2	SO4	F	2301	-	4,4,4	0.32	0	6,6,6	0.07	0
2	SO4	C	2301	-	4,4,4	0.35	0	6,6,6	0.13	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	GOL	D	2302	-	-	0/4/4/4	-
3	GOL	E	2302	-	-	2/4/4/4	-
3	GOL	A	2302	-	-	4/4/4/4	-
3	GOL	F	2302	-	-	0/4/4/4	-
3	GOL	A	2303	-	-	2/4/4/4	-
3	GOL	B	2302	-	-	0/4/4/4	-
3	GOL	C	2302	-	-	2/4/4/4	-
3	GOL	D	2303	-	-	0/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 10 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	2302	GOL	C1-C2-C3-O3
3	A	2303	GOL	C1-C2-C3-O3
3	A	2303	GOL	O2-C2-C3-O3
3	A	2302	GOL	O1-C1-C2-C3
3	C	2302	GOL	O1-C1-C2-C3

There are no ring outliers.

3 monomers are involved in 15 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	E	2302	GOL	1	0
3	F	2302	GOL	9	0
3	B	2302	GOL	5	0

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

EDS failed to run properly - this section is therefore empty.

6.2 Non-standard residues in protein, DNA, RNA chains

EDS failed to run properly - this section is therefore empty.

6.3 Carbohydrates

EDS failed to run properly - this section is therefore empty.

6.4 Ligands

EDS failed to run properly - this section is therefore empty.

6.5 Other polymers

EDS failed to run properly - this section is therefore empty.