Full wwPDB X-ray Structure Validation Report (i)

Oct 15, 2023 – 05:38 PM EDT

PDB ID	:	8EPJ
Title	:	Co-crystal structure of Chaetomium glucosidase with compound 17
Authors	:	Karade, S.S.; Mariuzza, R.A.
Deposited on	:	2022-10-05
Resolution	:	2.15 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 (i)) were used in the production of this report:

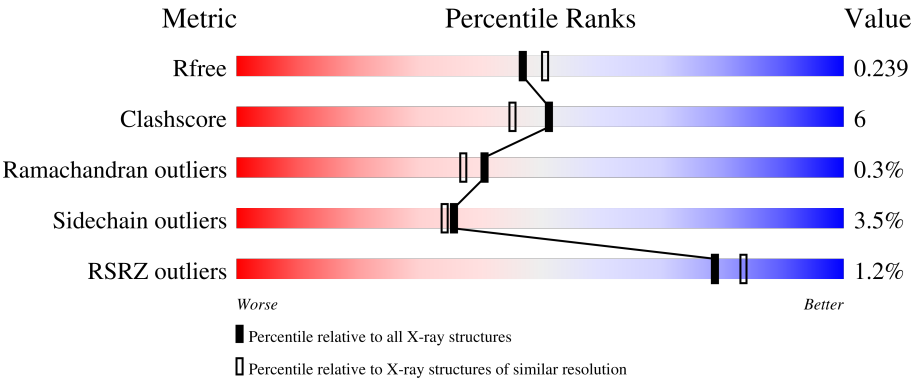
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.15 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}	130704	1479 (2.16-2.16)
Clashscore	141614	1585 (2.16-2.16)
Ramachandran outliers	138981	1560 (2.16-2.16)
Sidechain outliers	138945	1559 (2.16-2.16)
RSRZ outliers	127900	1456 (2.16-2.16)

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	819	2% 81% 12% 7%
1	B	819	81% 11% 7%

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 criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	GOL	A	902	-	-	X	-

2 Entry composition (i)

There are 7 unique types of molecules in this entry. The entry contains 12794 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 Chaetomium alpha glucosidase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	765	Total	C	N	O	S	0	3	0
			6079	3910	1024	1132	13			
1	B	764	Total	C	N	O	S	0	0	0
			6083	3903	1021	1146	13			

There are 78 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	MET	-	initiating methionine	UNP G0SFD1
A	0	GLY	-	expression tag	UNP G0SFD1
A	1	ILE	-	expression tag	UNP G0SFD1
A	2	LEU	-	expression tag	UNP G0SFD1
A	3	PRO	-	expression tag	UNP G0SFD1
A	4	SER	-	expression tag	UNP G0SFD1
A	5	PRO	-	expression tag	UNP G0SFD1
A	6	GLY	-	expression tag	UNP G0SFD1
A	7	MET	-	expression tag	UNP G0SFD1
A	8	PRO	-	expression tag	UNP G0SFD1
A	9	ALA	-	expression tag	UNP G0SFD1
A	10	LEU	-	expression tag	UNP G0SFD1
A	11	LEU	-	expression tag	UNP G0SFD1
A	12	SER	-	expression tag	UNP G0SFD1
A	13	LEU	-	expression tag	UNP G0SFD1
A	14	VAL	-	expression tag	UNP G0SFD1
A	15	SER	-	expression tag	UNP G0SFD1
A	16	LEU	-	expression tag	UNP G0SFD1
A	17	LEU	-	expression tag	UNP G0SFD1
A	18	SER	-	expression tag	UNP G0SFD1
A	19	VAL	-	expression tag	UNP G0SFD1
A	20	LEU	-	expression tag	UNP G0SFD1
A	21	LEU	-	expression tag	UNP G0SFD1
A	22	MET	-	expression tag	UNP G0SFD1
A	23	GLY	-	expression tag	UNP G0SFD1

Continued on next page...

Continued from previous page...

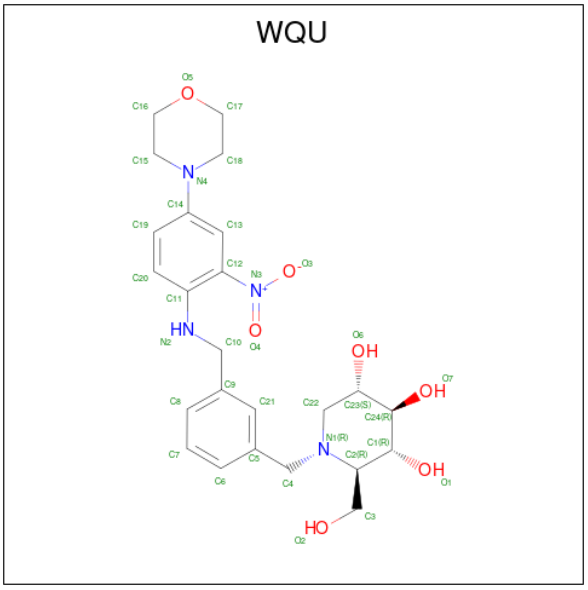
Chain	Residue	Modelled	Actual	Comment	Reference
A	24	CYS	-	expression tag	UNP G0SFD1
A	25	VAL	-	expression tag	UNP G0SFD1
A	26	ALA	-	expression tag	UNP G0SFD1
A	27	GLU	-	expression tag	UNP G0SFD1
A	28	THR	-	expression tag	UNP G0SFD1
A	29	GLY	-	expression tag	UNP G0SFD1
A	810	SER	-	expression tag	UNP G0SFD1
A	811	GLY	-	expression tag	UNP G0SFD1
A	812	HIS	-	expression tag	UNP G0SFD1
A	813	HIS	-	expression tag	UNP G0SFD1
A	814	HIS	-	expression tag	UNP G0SFD1
A	815	HIS	-	expression tag	UNP G0SFD1
A	816	HIS	-	expression tag	UNP G0SFD1
A	817	HIS	-	expression tag	UNP G0SFD1
B	-1	MET	-	initiating methionine	UNP G0SFD1
B	0	GLY	-	expression tag	UNP G0SFD1
B	1	ILE	-	expression tag	UNP G0SFD1
B	2	LEU	-	expression tag	UNP G0SFD1
B	3	PRO	-	expression tag	UNP G0SFD1
B	4	SER	-	expression tag	UNP G0SFD1
B	5	PRO	-	expression tag	UNP G0SFD1
B	6	GLY	-	expression tag	UNP G0SFD1
B	7	MET	-	expression tag	UNP G0SFD1
B	8	PRO	-	expression tag	UNP G0SFD1
B	9	ALA	-	expression tag	UNP G0SFD1
B	10	LEU	-	expression tag	UNP G0SFD1
B	11	LEU	-	expression tag	UNP G0SFD1
B	12	SER	-	expression tag	UNP G0SFD1
B	13	LEU	-	expression tag	UNP G0SFD1
B	14	VAL	-	expression tag	UNP G0SFD1
B	15	SER	-	expression tag	UNP G0SFD1
B	16	LEU	-	expression tag	UNP G0SFD1
B	17	LEU	-	expression tag	UNP G0SFD1
B	18	SER	-	expression tag	UNP G0SFD1
B	19	VAL	-	expression tag	UNP G0SFD1
B	20	LEU	-	expression tag	UNP G0SFD1
B	21	LEU	-	expression tag	UNP G0SFD1
B	22	MET	-	expression tag	UNP G0SFD1
B	23	GLY	-	expression tag	UNP G0SFD1
B	24	CYS	-	expression tag	UNP G0SFD1
B	25	VAL	-	expression tag	UNP G0SFD1
B	26	ALA	-	expression tag	UNP G0SFD1

Continued on next page...

Continued from previous page...

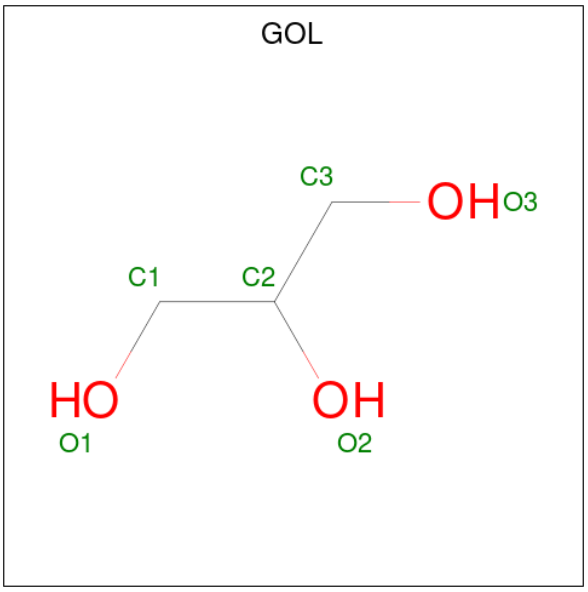
Chain	Residue	Modelled	Actual	Comment	Reference
B	27	GLU	-	expression tag	UNP G0SFD1
B	28	THR	-	expression tag	UNP G0SFD1
B	29	GLY	-	expression tag	UNP G0SFD1
B	810	SER	-	expression tag	UNP G0SFD1
B	811	GLY	-	expression tag	UNP G0SFD1
B	812	HIS	-	expression tag	UNP G0SFD1
B	813	HIS	-	expression tag	UNP G0SFD1
B	814	HIS	-	expression tag	UNP G0SFD1
B	815	HIS	-	expression tag	UNP G0SFD1
B	816	HIS	-	expression tag	UNP G0SFD1
B	817	HIS	-	expression tag	UNP G0SFD1

- Molecule 2 is (2R,3R,4R,5S)-2-(hydroxymethyl)-1-[(3-{[4-(morpholin-4-yl)-2-nitroanilino]methyl}phenyl)methyl]piperidine-3,4,5-triol (three-letter code: WQU) (formula: $C_{24}H_{32}N_4O_7$) (labeled as "Ligand of Interest" by depositor).



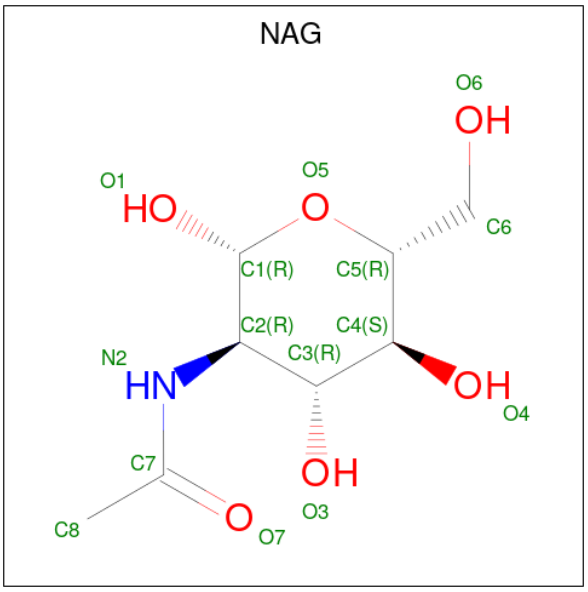
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	N	O	0	0
			35	24	4	7		
2	B	1	Total	C	N	O	0	0
			35	24	4	7		

- Molecule 3 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



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

- Molecule 4 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).



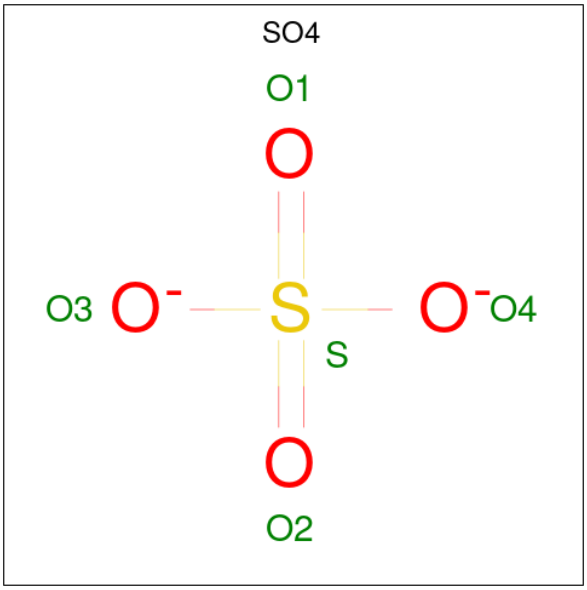
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total C N O 14 8 1 5	0	0

Continued on next page...

Continued from previous page...

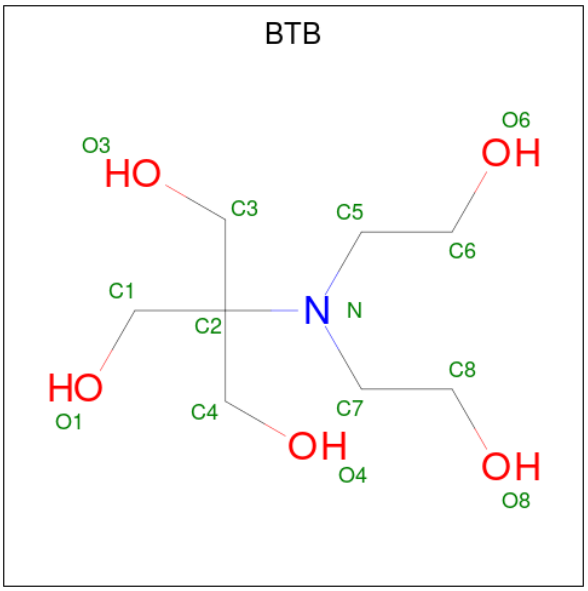
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	B	1	Total C N O 14 8 1 5	0	0

- Molecule 5 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



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

- Molecule 6 is 2-[BIS-(2-HYDROXY-ETHYL)-AMINO]-2-HYDROXYMETHYL-PROPANE-1,3-DIOL (three-letter code: BTB) (formula: $C_8H_{19}NO_5$).



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

- Molecule 7 is water.

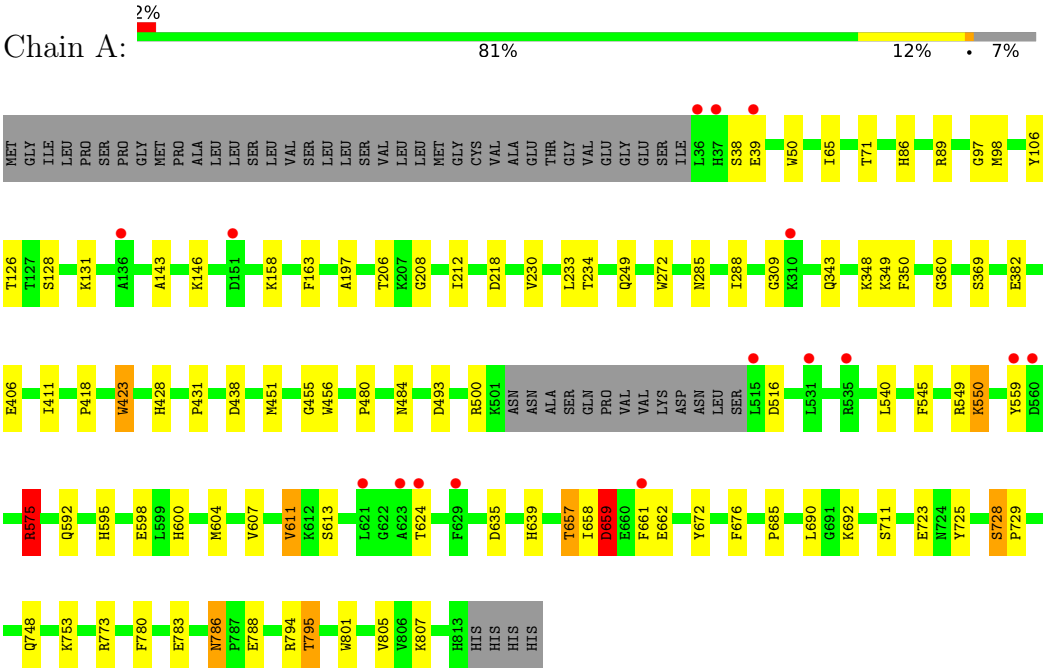
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	A	218	Total O 218 218	0	0
7	B	245	Total O 245 245	0	0

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.

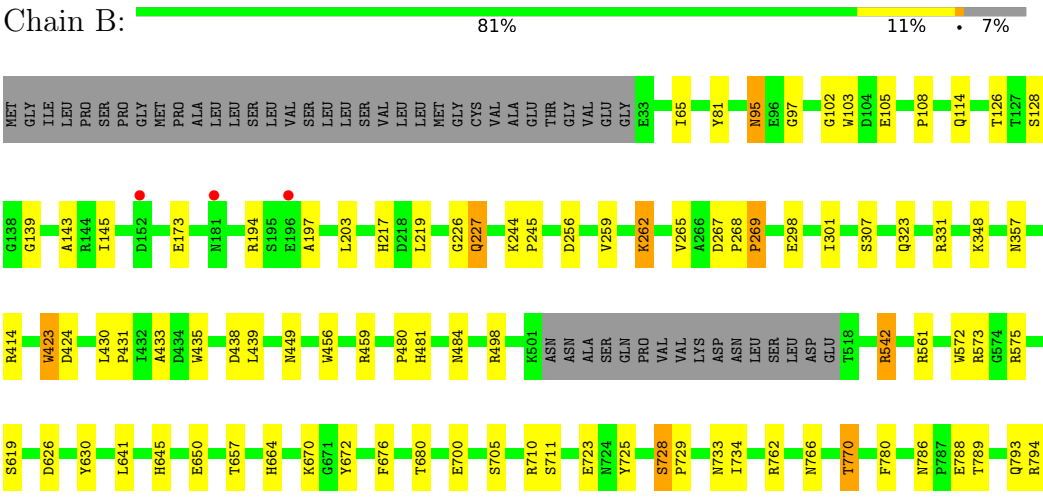
• Molecule 1: Chaetomium alpha glucosidase

Chain A:



• Molecule 1: Chaetomium alpha glucosidase

Chain B:




HIS
HIS
HIS
HIS
HIS

4 Data and refinement statistics (i)

Property	Value	Source
Space group	I 2 2 2	Depositor
Cell constants a, b, c, α, β, γ	135.26Å 180.01Å 180.43Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	46.95 – 2.15 46.90 – 2.15	Depositor EDS
% Data completeness (in resolution range)	99.9 (46.95-2.15) 99.9 (46.90-2.15)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.35 (at 2.16Å)	Xtriage
Refinement program	REFMAC 5.8.0258	Depositor
R, R_{free}	0.192 , 0.236 0.197 , 0.239	Depositor DCC
R_{free} test set	5830 reflections (4.92%)	wwPDB-VP
Wilson B-factor (Å^2)	42.2	Xtriage
Anisotropy	0.250	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 39.1	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.010 for -h,-l,-k	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	12794	wwPDB-VP
Average B, all atoms (Å^2)	47.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.56% 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, WQU, BTB, SO4, GOL

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	0.66	0/6264	0.81	2/8529 (0.0%)
1	B	0.65	0/6259	0.82	4/8518 (0.0%)
All	All	0.66	0/12523	0.81	6/17047 (0.0%)

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	A	0	1
1	B	0	1
All	All	0	2

There are no bond length outliers.

All (6) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	575	ARG	CB-CA-C	7.45	125.30	110.40
1	B	459	ARG	NE-CZ-NH2	-6.58	117.01	120.30
1	B	573	ARG	NE-CZ-NH2	-6.43	117.08	120.30
1	B	459	ARG	NE-CZ-NH1	5.83	123.21	120.30
1	A	659	ASP	CB-CA-C	-5.43	99.54	110.40
1	B	542	ARG	CG-CD-NE	5.01	122.33	111.80

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	780	PHE	Peptide
1	B	780	PHE	Peptide

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	A	6079	0	5761	71	0
1	B	6083	0	5743	65	0
2	A	35	0	0	0	0
2	B	35	0	0	0	0
3	A	6	0	8	6	0
3	B	6	0	8	0	0
4	A	14	0	13	3	0
4	B	14	0	13	0	0
5	A	25	0	0	0	0
5	B	20	0	0	1	0
6	B	14	0	19	0	0
7	A	218	0	0	5	0
7	B	245	0	0	3	0
All	All	12794	0	11565	133	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 6.

All (133) 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:A:406:GLU:OE2	4:A:903:NAG:H81	1.37	1.22
1:A:657:THR:HG21	7:A:1025:HOH:O	1.59	1.00
1:B:561:ARG:HE	1:B:664:HIS:HD2	1.17	0.93
1:A:451:MET:HE1	1:A:540:LEU:HB3	1.57	0.86
1:B:786:ASN:HD22	1:B:789:THR:H	1.26	0.83
1:A:685:PRO:HG3	1:A:748:GLN:HE22	1.42	0.83
1:A:206:THR:HB	3:A:902:GOL:H32	1.62	0.82
1:B:331:ARG:NH1	1:B:435:TRP:O	2.20	0.74
1:A:382:GLU:HB2	7:A:1187:HOH:O	1.88	0.73

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:217:HIS:HE1	1:B:267:ASP:O	1.72	0.73
1:A:658:ILE:HD12	1:A:662:GLU:HA	1.69	0.72
1:A:493:ASP:OD1	1:A:613:SER:OG	2.06	0.72
1:A:559:TYR:CE1	1:A:658:ILE:HD13	2.25	0.72
1:A:451:MET:CE	1:A:540:LEU:HB3	2.23	0.69
1:A:451:MET:HE1	1:A:540:LEU:CB	2.25	0.66
1:B:561:ARG:HE	1:B:664:HIS:CD2	2.08	0.66
4:A:903:NAG:O7	4:A:903:NAG:O3	2.10	0.66
1:B:766:ASN:O	1:B:770:THR:HG23	1.98	0.64
1:A:559:TYR:CD1	1:A:658:ILE:HD13	2.33	0.64
4:A:903:NAG:HO3	4:A:903:NAG:C7	2.10	0.63
1:A:598:GLU:OE2	1:A:600:HIS:HE1	1.83	0.62
1:B:217:HIS:HD2	1:B:219:LEU:H	1.47	0.62
1:B:572:TRP:H	1:B:600:HIS:HD2	1.47	0.60
1:A:128:SER:O	1:A:143:ALA:HA	2.02	0.59
1:B:786:ASN:HD21	1:B:788:GLU:HB2	1.67	0.59
1:B:430:LEU:HB2	1:B:431:PRO:HD3	1.85	0.59
1:B:762:ARG:NH1	5:B:906:SO4:O1	2.36	0.58
1:A:249:GLN:HE21	1:B:262:LYS:HZ3	1.51	0.58
1:A:456:TRP:CE2	1:A:480:PRO:HA	2.40	0.56
1:B:95:ASN:HD22	1:B:97:GLY:H	1.54	0.56
1:A:575:ARG:HH22	1:A:592:GLN:HE22	1.52	0.56
1:B:561:ARG:NE	1:B:664:HIS:HD2	1.96	0.55
1:B:766:ASN:O	1:B:770:THR:CG2	2.54	0.55
1:A:206:THR:HB	3:A:902:GOL:C3	2.35	0.55
1:A:451:MET:HE3	1:A:540:LEU:HD22	1.87	0.55
1:A:234[B]:THR:HA	1:A:285:ASN:OD1	2.07	0.54
1:A:126:THR:HG22	7:A:1118:HOH:O	2.07	0.54
1:B:65:ILE:HD13	1:B:197:ALA:HB1	1.90	0.54
1:B:456:TRP:CE2	1:B:480:PRO:HA	2.43	0.54
1:A:350:PHE:CZ	1:A:805[B]:VAL:HG11	2.43	0.54
1:B:650:GLU:OE1	1:B:670:LYS:NZ	2.41	0.53
1:B:723:GLU:O	1:B:794:ARG:NH2	2.42	0.53
1:B:226:GLY:H	1:B:227:GLN:HE22	1.56	0.53
1:A:249:GLN:HE21	1:B:262:LYS:NZ	2.07	0.52
1:B:203:LEU:HD11	1:B:301:ILE:CG2	2.40	0.52
1:A:690:LEU:HD21	1:A:745:ILE:HD13	1.91	0.52
1:B:542:ARG:HD3	1:B:630:TYR:OH	2.09	0.52
1:B:226:GLY:H	1:B:227:GLN:NE2	2.08	0.52
1:A:206:THR:CB	3:A:902:GOL:H32	2.35	0.52
1:A:659:ASP:HB3	1:A:661:PHE:H	1.75	0.52

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:685:PRO:HG3	1:A:748:GLN:NE2	2.20	0.51
1:A:208:GLY:H	3:A:902:GOL:C3	2.23	0.51
1:A:550:LYS:HE2	1:A:550:LYS:HA	1.94	0.50
1:A:86:HIS:HD2	7:A:1210:HOH:O	1.94	0.50
1:A:233:LEU:HD22	1:B:265:VAL:HG23	1.93	0.50
1:A:234[B]:THR:HG21	1:B:268:PRO:HG3	1.94	0.50
1:B:102:GLY:HA3	1:B:357:ASN:HD21	1.77	0.50
1:A:234[A]:THR:HA	1:A:285:ASN:OD1	2.12	0.49
1:B:256:ASP:O	1:B:259:VAL:HG22	2.12	0.49
1:B:128:SER:O	1:B:143:ALA:HA	2.13	0.49
1:B:244:LYS:HB3	1:B:245:PRO:HD3	1.95	0.49
1:A:569:ALA:HA	1:A:604:MET:CE	2.43	0.49
1:B:700:GLU:O	1:B:710:ARG:NH1	2.45	0.48
1:B:786:ASN:ND2	1:B:788:GLU:H	2.11	0.48
1:B:145:ILE:HD11	1:B:301:ILE:HD12	1.95	0.48
1:A:451:MET:HE1	1:A:540:LEU:CA	2.44	0.48
1:A:71:THR:HB	1:A:163:PHE:CZ	2.49	0.48
1:A:106:TYR:CD2	1:A:360:GLY:HA2	2.49	0.48
1:A:208:GLY:N	3:A:902:GOL:O3	2.43	0.47
1:B:134:GLY:O	1:B:139:GLY:HA2	2.13	0.47
1:A:723:GLU:O	1:A:794:ARG:NH1	2.48	0.47
1:A:595:HIS:CE1	1:A:658:ILE:HG12	2.50	0.47
1:B:498:ARG:HD3	7:B:1221:HOH:O	2.14	0.47
1:B:626:ASP:HB3	1:B:630:TYR:CE2	2.49	0.47
1:A:158:LYS:HD3	1:A:272:TRP:CD1	2.50	0.47
1:A:249:GLN:NE2	1:B:262:LYS:NZ	2.63	0.47
1:B:114:GLN:NE2	1:B:414:ARG:HH12	2.12	0.47
1:B:103:TRP:H	1:B:357:ASN:ND2	2.12	0.46
1:A:728:SER:N	1:A:729:PRO:CD	2.79	0.46
1:A:572:TRP:H	1:A:600:HIS:HD2	1.64	0.46
1:A:575:ARG:HH22	1:A:592:GLN:NE2	2.13	0.45
1:B:498:ARG:CD	7:B:1221:HOH:O	2.64	0.45
1:A:411:ILE:HD12	1:A:418:PRO:HA	1.99	0.45
1:A:545:PHE:CE2	1:A:549:ARG:HD2	2.52	0.45
1:B:81:TYR:OH	1:B:269:PRO:HD2	2.17	0.45
1:A:598:GLU:OE2	1:A:600:HIS:CE1	2.67	0.45
1:B:433:ALA:O	1:B:498:ARG:NH2	2.50	0.45
1:B:145:ILE:O	1:B:298:GLU:HA	2.16	0.45
1:B:641:LEU:O	1:B:645:HIS:HB2	2.17	0.45
1:B:672:TYR:CG	1:B:734:ILE:HG21	2.52	0.45
1:A:423:TRP:H	1:A:484:ASN:ND2	2.15	0.44

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:672:TYR:CE1	1:B:711:SER:HA	2.53	0.44
1:B:733:ASN:HB3	1:B:801:TRP:CG	2.52	0.44
1:A:39:GLU:HA	1:A:39:GLU:OE2	2.17	0.44
1:A:786:ASN:ND2	1:A:788:GLU:H	2.15	0.44
1:B:609:VAL:CG2	1:B:680:THR:HB	2.48	0.44
1:B:385:GLY:HA2	1:B:387:TRP:CH2	2.53	0.44
1:A:672:TYR:CZ	1:A:711:SER:HA	2.52	0.44
1:A:397:HIS:HD2	1:A:399:GLU:OE2	2.00	0.43
1:A:545:PHE:CZ	1:A:549:ARG:HD2	2.53	0.43
1:B:481:HIS:HD2	7:B:1230:HOH:O	2.02	0.43
1:B:145:ILE:HD11	1:B:301:ILE:CD1	2.49	0.43
1:B:728:SER:N	1:B:729:PRO:CD	2.82	0.43
1:A:672:TYR:CG	1:A:734:ILE:HG21	2.53	0.43
1:A:451:MET:SD	1:A:455:GLY:HA2	2.58	0.43
1:B:137:HIS:HD2	1:B:307:SER:OG	2.02	0.43
1:B:572:TRP:H	1:B:600:HIS:CD2	2.34	0.43
1:A:208:GLY:HA2	3:A:902:GOL:C1	2.49	0.42
1:A:635:ASP:O	1:A:639:HIS:HD2	2.02	0.42
1:A:97:GLY:HA3	7:A:1031:HOH:O	2.19	0.42
1:B:786:ASN:ND2	1:B:788:GLU:HB2	2.34	0.42
1:A:733:ASN:HB3	1:A:801:TRP:CG	2.53	0.42
1:B:786:ASN:HB2	1:B:793:GLN:NE2	2.35	0.42
1:A:230:VAL:HA	1:A:288:ILE:O	2.20	0.42
1:B:95:ASN:HD22	1:B:95:ASN:C	2.23	0.42
1:B:108:PRO:HG2	1:B:439:LEU:HD22	2.01	0.42
1:A:607:VAL:O	1:A:611:VAL:HG22	2.19	0.42
1:B:102:GLY:CA	1:B:357:ASN:HD21	2.33	0.42
1:A:456:TRP:CD2	1:A:480:PRO:HA	2.55	0.41
1:A:50:TRP:O	1:A:131:LYS:NZ	2.42	0.41
1:A:89:ARG:NH2	1:A:98:MET:HE2	2.35	0.41
1:A:125:ILE:HA	1:A:146:LYS:O	2.18	0.41
1:B:114:GLN:HE22	1:B:414:ARG:HH22	1.69	0.41
1:B:423:TRP:H	1:B:484:ASN:ND2	2.18	0.41
1:A:110:ARG:HD3	1:A:516:ASP:OD2	2.20	0.41
1:B:203:LEU:HD11	1:B:301:ILE:HG23	2.01	0.41
1:B:385:GLY:HA2	1:B:387:TRP:CZ3	2.55	0.41
1:A:672:TYR:CE2	1:A:711:SER:HA	2.56	0.40
1:A:783:GLU:OE1	1:A:795:THR:HB	2.21	0.40
1:B:95:ASN:HD22	1:B:97:GLY:N	2.18	0.40
1:A:65:ILE:HD13	1:A:197:ALA:HB1	2.04	0.40
1:A:428:HIS:O	1:A:431:PRO:HD2	2.21	0.40

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:456:TRP:CD2	1:B:480:PRO:HA	2.56	0.40

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	A	764/819 (93%)	739 (97%)	21 (3%)	4 (0%)	29	22
1	B	760/819 (93%)	736 (97%)	23 (3%)	1 (0%)	51	53
All	All	1524/1638 (93%)	1475 (97%)	44 (3%)	5 (0%)	41	37

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	728	SER
1	B	728	SER
1	A	659	ASP
1	A	500	ARG
1	A	309	GLY

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.

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	630/707 (89%)	608 (96%)	22 (4%)	36	34
1	B	633/707 (90%)	611 (96%)	22 (4%)	36	34
All	All	1263/1414 (89%)	1219 (96%)	44 (4%)	36	34

All (44) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	38	SER
1	A	212	ILE
1	A	218	ASP
1	A	343	GLN
1	A	348	LYS
1	A	349	LYS
1	A	369	SER
1	A	423	TRP
1	A	438	ASP
1	A	550	LYS
1	A	575	ARG
1	A	611	VAL
1	A	624	THR
1	A	657	THR
1	A	676	PHE
1	A	692	LYS
1	A	725	TYR
1	A	753	LYS
1	A	773	ARG
1	A	786	ASN
1	A	795	THR
1	A	807	LYS
1	B	95	ASN
1	B	105	GLU
1	B	126	THR
1	B	173	GLU
1	B	194	ARG
1	B	227	GLN
1	B	262	LYS
1	B	269	PRO
1	B	323	GLN
1	B	348	LYS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	423	TRP
1	B	424	ASP
1	B	438	ASP
1	B	449	ASN
1	B	575	ARG
1	B	619	SER
1	B	657	THR
1	B	676	PHE
1	B	705	SER
1	B	725	TYR
1	B	770	THR
1	B	801	TRP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (39) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	120	GLN
1	A	157	GLN
1	A	167	GLN
1	A	231	GLN
1	A	242	GLN
1	A	249	GLN
1	A	250	GLN
1	A	323	GLN
1	A	397	HIS
1	A	398	GLN
1	A	461	GLN
1	A	484	ASN
1	A	592	GLN
1	A	600	HIS
1	A	639	HIS
1	A	748	GLN
1	A	786	ASN
1	A	793	GLN
1	B	85	GLN
1	B	95	ASN
1	B	114	GLN
1	B	120	GLN
1	B	137	HIS
1	B	167	GLN
1	B	217	HIS
1	B	227	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	249	GLN
1	B	323	GLN
1	B	357	ASN
1	B	368	HIS
1	B	461	GLN
1	B	481	HIS
1	B	484	ASN
1	B	600	HIS
1	B	639	HIS
1	B	664	HIS
1	B	748	GLN
1	B	786	ASN
1	B	793	GLN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

16 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	$\#\|Z\| > 2$
5	SO4	B	906	-	4,4,4	0.38	0	6,6,6	0.16	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\|>2$	Counts	RMSZ	$\#\|Z\|>2$
5	SO4	B	905	-	4,4,4	0.59	0	6,6,6	0.22	0
5	SO4	A	907	-	4,4,4	0.36	0	6,6,6	0.15	0
2	WQU	B	901	-	37,38,38	3.33	13 (35%)	48,53,53	2.07	13 (27%)
2	WQU	A	901	-	37,38,38	3.54	14 (37%)	48,53,53	2.09	12 (25%)
5	SO4	A	904	-	4,4,4	0.43	0	6,6,6	0.16	0
5	SO4	B	908	-	4,4,4	0.29	0	6,6,6	0.11	0
5	SO4	A	905	-	4,4,4	0.38	0	6,6,6	0.15	0
4	NAG	A	903	1	14,14,15	0.54	0	17,19,21	2.16	3 (17%)
5	SO4	A	908	-	4,4,4	0.37	0	6,6,6	0.09	0
6	BTB	B	903	-	13,13,13	1.28	2 (15%)	7,16,16	0.43	0
3	GOL	A	902	-	5,5,5	0.13	0	5,5,5	0.37	0
3	GOL	B	902	-	5,5,5	0.23	0	5,5,5	0.40	0
5	SO4	B	907	-	4,4,4	0.33	0	6,6,6	0.09	0
4	NAG	B	904	1	14,14,15	0.76	0	17,19,21	1.58	3 (17%)
5	SO4	A	906	-	4,4,4	0.31	0	6,6,6	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
2	WQU	B	901	-	-	5/17/47/47	0/4/4/4
2	WQU	A	901	-	-	4/17/47/47	0/4/4/4
4	NAG	A	903	1	-	5/6/23/26	0/1/1/1
6	BTB	B	903	-	-	2/21/21/21	-
3	GOL	A	902	-	-	0/4/4/4	-
3	GOL	B	902	-	-	3/4/4/4	-
4	NAG	B	904	1	-	0/6/23/26	0/1/1/1

All (29) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	901	WQU	C4-N1	-11.53	1.27	1.47
2	A	901	WQU	C4-N1	-11.48	1.28	1.47
2	A	901	WQU	O4-N3	10.51	1.40	1.22
2	B	901	WQU	O4-N3	10.45	1.40	1.22
2	A	901	WQU	C23-C24	-8.16	1.40	1.52
2	B	901	WQU	C23-C24	-7.58	1.41	1.52
2	A	901	WQU	C11-N2	5.03	1.51	1.37

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	901	WQU	C4-C5	4.69	1.59	1.51
2	A	901	WQU	O7-C24	4.27	1.53	1.43
2	A	901	WQU	C11-C12	4.27	1.47	1.40
2	A	901	WQU	C3-C2	-4.07	1.46	1.52
2	B	901	WQU	O7-C24	3.90	1.52	1.43
2	A	901	WQU	C12-N3	3.67	1.52	1.45
2	B	901	WQU	C11-N2	3.66	1.47	1.37
2	A	901	WQU	C4-C5	3.61	1.57	1.51
2	A	901	WQU	C1-C2	3.59	1.61	1.53
6	B	903	BTB	C2-N	3.36	1.55	1.48
2	B	901	WQU	C1-C2	3.18	1.60	1.53
2	A	901	WQU	C14-N4	2.84	1.46	1.38
2	A	901	WQU	C1-C24	-2.84	1.45	1.52
2	B	901	WQU	O6-C23	2.80	1.49	1.43
2	B	901	WQU	C12-N3	2.54	1.50	1.45
2	B	901	WQU	C14-N4	2.42	1.45	1.38
2	A	901	WQU	O6-C23	2.36	1.48	1.43
2	A	901	WQU	C21-C9	2.27	1.43	1.39
2	B	901	WQU	C1-C24	-2.20	1.46	1.52
6	B	903	BTB	C7-N	2.10	1.51	1.48
2	B	901	WQU	C7-C6	2.05	1.43	1.38
2	B	901	WQU	C21-C5	-2.01	1.35	1.39

All (31) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	A	901	WQU	C20-C11-N2	-6.92	110.01	121.80
2	A	901	WQU	C13-C12-C11	-6.70	115.46	121.53
4	A	903	NAG	C1-O5-C5	6.50	121.00	112.19
2	B	901	WQU	C13-C12-C11	-5.90	116.19	121.53
4	A	903	NAG	O5-C1-C2	-5.08	103.27	111.29
4	B	904	NAG	C1-O5-C5	4.94	118.89	112.19
2	B	901	WQU	C3-C2-C1	-4.27	106.35	112.90
2	B	901	WQU	C13-C14-N4	-4.18	116.83	121.33
2	B	901	WQU	C10-C9-C21	-3.98	112.03	120.64
2	A	901	WQU	C4-N1-C2	3.64	120.33	112.97
2	B	901	WQU	C7-C8-C9	-3.53	115.22	120.63
2	A	901	WQU	C12-C11-N2	3.51	129.50	123.33
2	B	901	WQU	C20-C11-N2	-3.24	116.27	121.80
2	A	901	WQU	C3-C2-C1	-3.24	107.94	112.90
2	B	901	WQU	C20-C11-C12	3.21	123.33	118.58
2	B	901	WQU	C8-C9-C21	3.01	122.76	118.54

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	901	WQU	C4-C5-C6	2.99	126.39	120.77
2	A	901	WQU	O4-N3-C12	2.94	124.07	119.03
4	B	904	NAG	C3-C4-C5	-2.65	105.51	110.24
2	B	901	WQU	O7-C24-C1	-2.60	104.33	110.35
2	A	901	WQU	O2-C3-C2	-2.46	106.14	111.42
2	A	901	WQU	C10-N2-C11	2.38	128.04	121.97
2	B	901	WQU	C4-N1-C22	2.21	114.55	110.31
4	A	903	NAG	O3-C3-C4	-2.16	105.36	110.35
2	A	901	WQU	C23-C24-C1	-2.11	107.25	110.89
2	A	901	WQU	C19-C14-N4	-2.10	118.48	121.38
2	B	901	WQU	C10-C9-C8	2.07	125.21	120.91
2	A	901	WQU	C16-O5-C17	2.05	116.75	109.89
4	B	904	NAG	O4-C4-C5	2.04	114.37	109.30
2	A	901	WQU	C4-N1-C22	2.03	114.20	110.31
2	B	901	WQU	C20-C19-C14	-2.01	117.68	120.32

There are no chirality outliers.

All (19) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	901	WQU	C13-C12-N3-O4
2	A	901	WQU	C11-C12-N3-O4
2	A	901	WQU	C12-C11-N2-C10
2	B	901	WQU	C20-C11-N2-C10
2	B	901	WQU	C12-C11-N2-C10
3	B	902	GOL	O1-C1-C2-C3
4	A	903	NAG	C3-C2-N2-C7
2	A	901	WQU	C20-C11-N2-C10
4	A	903	NAG	C8-C7-N2-C2
4	A	903	NAG	O5-C5-C6-O6
2	B	901	WQU	C13-C14-N4-C18
2	B	901	WQU	C19-C14-N4-C18
6	B	903	BTB	N-C7-C8-O8
4	A	903	NAG	O7-C7-N2-C2
4	A	903	NAG	C4-C5-C6-O6
3	B	902	GOL	O1-C1-C2-O2
6	B	903	BTB	N-C5-C6-O6
3	B	902	GOL	C1-C2-C3-O3
2	B	901	WQU	C1-C2-C3-O2

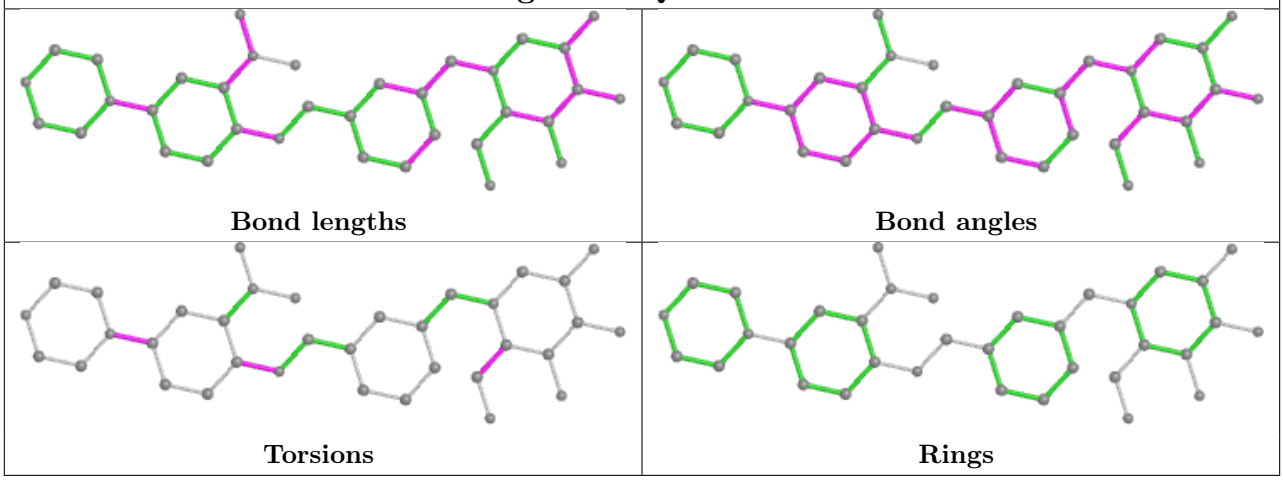
There are no ring outliers.

3 monomers are involved in 10 short contacts:

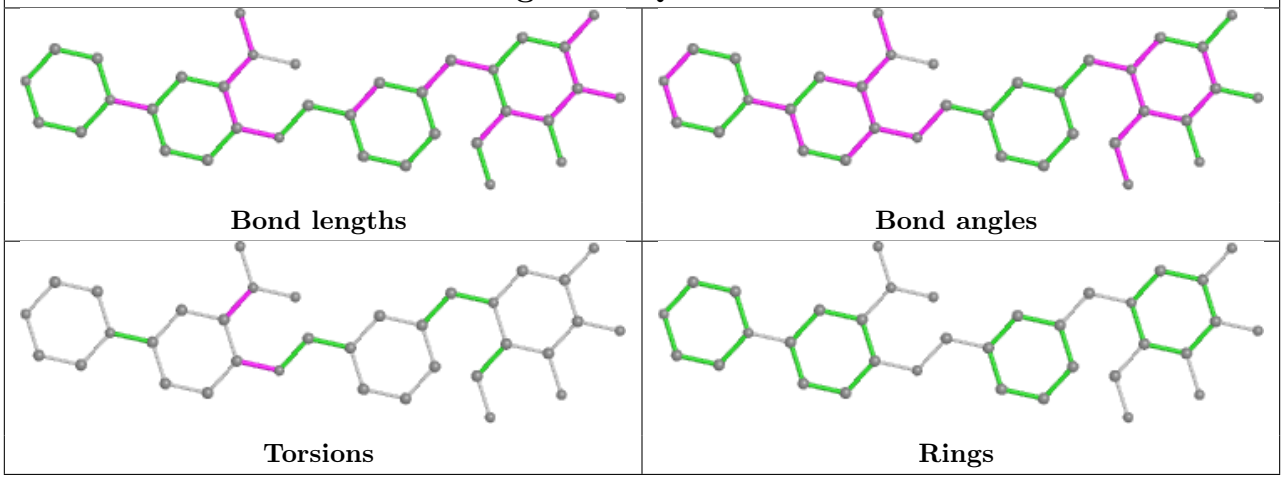
Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	B	906	SO4	1	0
4	A	903	NAG	3	0
3	A	902	GOL	6	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less then 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.

Ligand WQU B 901



Ligand WQU A 901



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	A	765/819 (93%)	-0.07	16 (2%) 63 71	30, 46, 67, 85	0
1	B	764/819 (93%)	-0.24	3 (0%) 92 94	31, 44, 62, 84	0
All	All	1529/1638 (93%)	-0.15	19 (1%) 79 83	30, 45, 65, 85	0

All (19) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	531	LEU	3.9
1	A	37	HIS	3.5
1	A	136	ALA	2.9
1	A	621	LEU	2.8
1	A	623	ALA	2.7
1	A	624	THR	2.7
1	A	515	LEU	2.7
1	A	39	GLU	2.5
1	A	151	ASP	2.4
1	A	560	ASP	2.4
1	A	535	ARG	2.3
1	A	661	PHE	2.3
1	A	310	LYS	2.3
1	B	181	ASN	2.2
1	A	629	PHE	2.2
1	A	559	TYR	2.1
1	A	36	LEU	2.1
1	B	196	GLU	2.1
1	B	152	ASP	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 (i)

There are no monosaccharides in this entry.

6.4 Ligands (i)

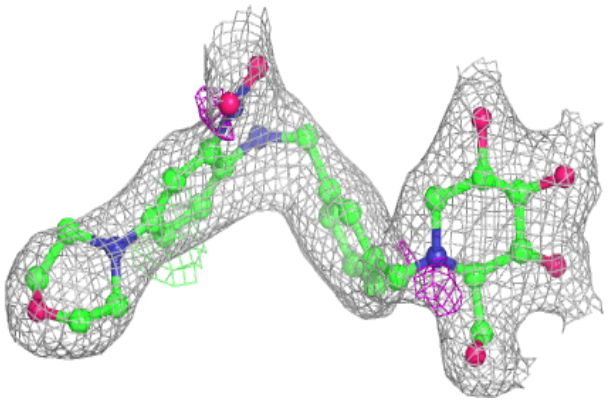
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

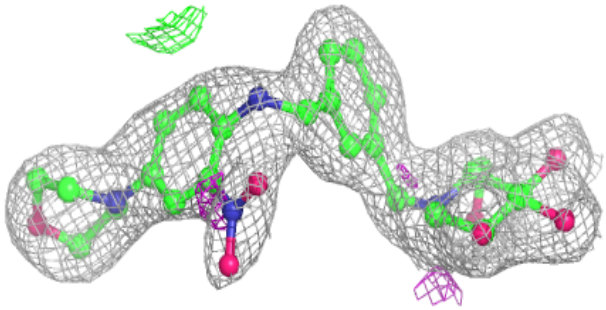
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
4	NAG	A	903	14/15	0.46	0.35	84,97,99,107	0
4	NAG	B	904	14/15	0.81	0.19	60,65,70,71	0
6	BTB	B	903	14/14	0.84	0.21	67,84,87,94	0
5	SO4	A	907	5/5	0.89	0.22	85,104,110,119	0
3	GOL	A	902	6/6	0.90	0.22	71,75,76,78	0
2	WQU	B	901	35/35	0.94	0.13	32,49,77,84	0
3	GOL	B	902	6/6	0.94	0.15	38,46,54,60	0
2	WQU	A	901	35/35	0.95	0.15	31,47,85,99	0
5	SO4	B	908	5/5	0.95	0.18	61,64,81,85	0
5	SO4	A	906	5/5	0.95	0.14	73,77,85,86	0
5	SO4	A	908	5/5	0.96	0.18	92,92,98,99	0
5	SO4	B	907	5/5	0.96	0.12	79,79,85,85	0
5	SO4	A	905	5/5	0.97	0.14	58,67,75,80	0
5	SO4	B	906	5/5	0.98	0.13	59,63,68,68	0
5	SO4	A	904	5/5	0.99	0.14	43,43,45,48	0
5	SO4	B	905	5/5	0.99	0.09	40,41,47,63	0

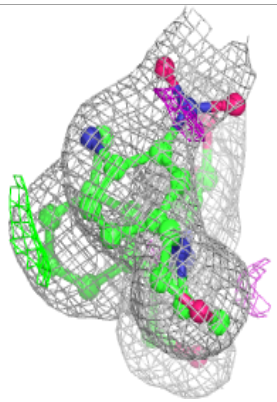
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

Electron density around WQU B 901:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

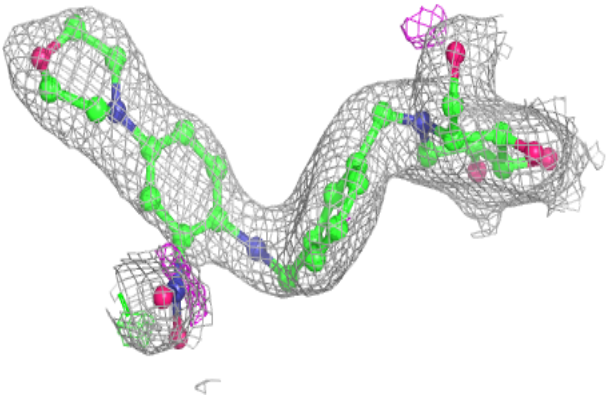


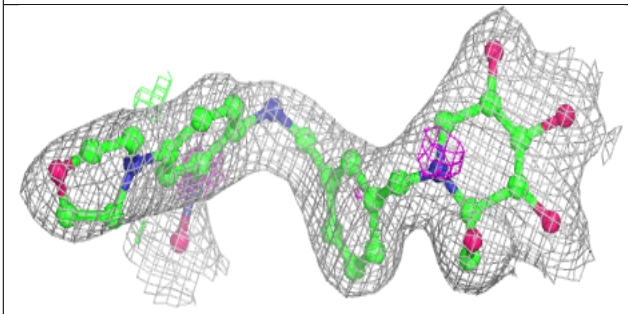


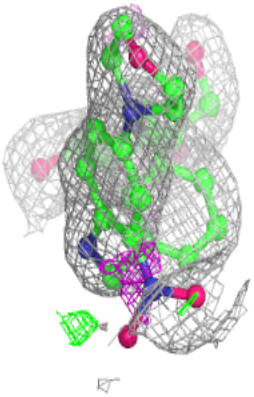


Electron density around WQU A 901:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)







6.5 Other polymers (i)

There are no such residues in this entry.