Full wwPDB X-ray Structure Validation Report (i)

Oct 15, 2023 – 07:09 PM EDT

PDB ID	:	8E3J
Title	:	Co-crystal structure of Chaetomium glucosidase with compound 4
Authors	:	Karade, S.S.; Mariuzza, R.A.
Deposited on	:	2022-08-17
Resolution	:	2.71 Å(reported)

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

|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (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 ⓘ

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

The reported resolution of this entry is 2.71 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.

Metric	Percentile Ranks	Value
Rfree		0.245
Clashscore		6
Ramachandran outliers		0.3%
Sidechain outliers		5.3%
RSRZ outliers		1.2%

Worse *Better*

▮ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution
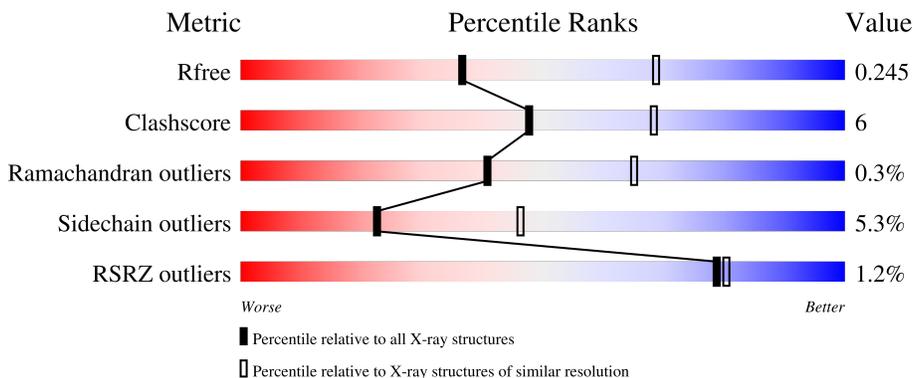
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R$_{free}$	130704	3359 (2.74-2.70)
Clashscore	141614	3686 (2.74-2.70)
Ramachandran outliers	138981	3622 (2.74-2.70)
Sidechain outliers	138945	3623 (2.74-2.70)
RSRZ outliers	127900	3276 (2.74-2.70)

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% · 77% · 15% · 7%
1	B	819	77% · 15% · 6%

2 Entry composition ⓘ

There are 6 unique types of molecules in this entry. The entry contains 12295 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 6032	C 3874	N 1019	O 1126	S 13	0	2	0
1	B	767	Total 6059	C 3887	N 1020	O 1139	S 13	0	0	0

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	817	HIS	-	expression tag	UNP G0SFD1
B	818	HIS	-	expression tag	UNP G0SFD1
B	819	HIS	-	expression tag	UNP G0SFD1
B	820	HIS	-	expression tag	UNP G0SFD1
B	821	HIS	-	expression tag	UNP G0SFD1

- Molecule 2 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).
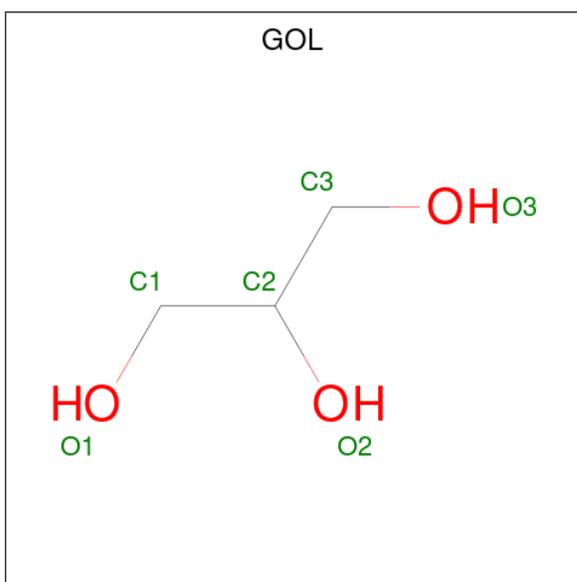
| Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total 6	C 3	O 3	0	0
2	B	1	Total 6	C 3	O 3	0	0
2	B	1	Total 6	C 3	O 3	0	0
2	B	1	Total 6	C 3	O 3	0	0
2	B	1	Total 6	C 3	O 3	0	0

- Molecule 3 is (2R,3R,4R,5S)-2-(hydroxymethyl)-1-{6-[2-nitro-4-(pyrimidin-2-yl)anilino]hexy

l}piperidine-3,4,5-triol (three-letter code: UGO) (formula: $C_{22}H_{32}N_5O_6$) (labeled as "Ligand of Interest" by depositor).
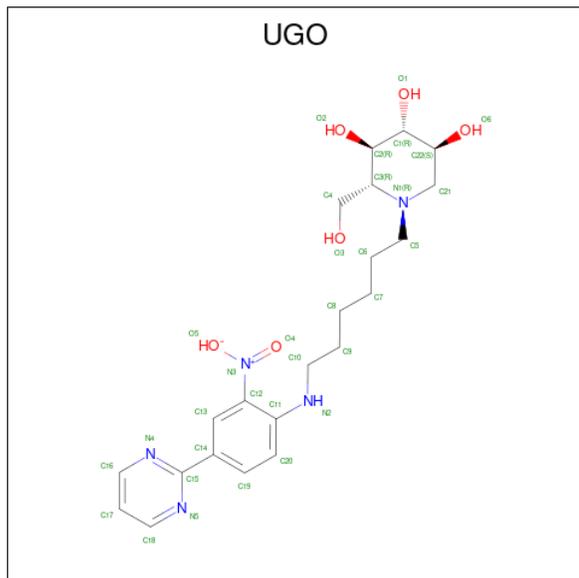
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	N	O	0	0
			33	22	5	6		
3	B	1	Total	C	N	O	0	0
			33	22	5	6		

- Molecule 4 is SULFATE ION (three-letter code: SO4) (formula: O_4S).
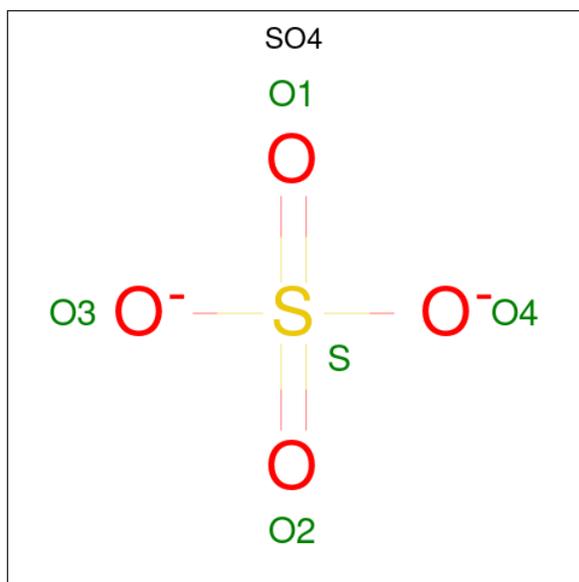
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total 5	O 4	S 1	0	0
4	A	1	Total 5	O 4	S 1	0	0
4	A	1	Total 5	O 4	S 1	0	0
4	A	1	Total 5	O 4	S 1	0	0
4	A	1	Total 5	O 4	S 1	0	0
4	A	1	Total 5	O 4	S 1	0	0
4	A	1	Total 5	O 4	S 1	0	0
4	A	1	Total 5	O 4	S 1	0	0
4	B	1	Total 5	O 4	S 1	0	0
4	B	1	Total 5	O 4	S 1	0	0
4	B	1	Total 5	O 4	S 1	0	0
4	B	1	Total 5	O 4	S 1	0	0
4	B	1	Total 5	O 4	S 1	0	0

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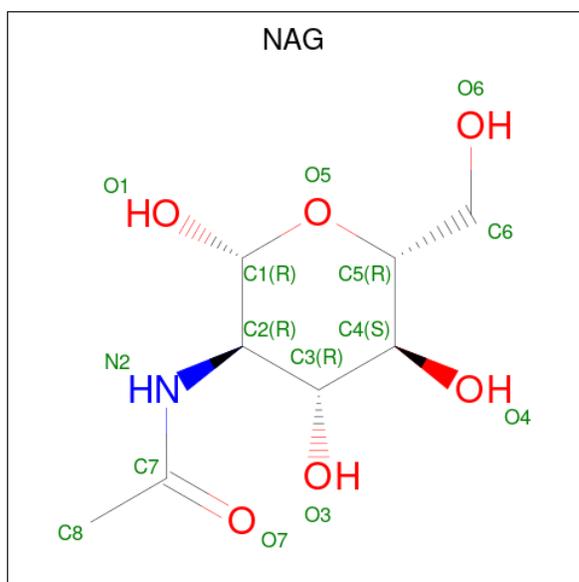
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	B	1	Total 14	C 8	N 1	O 5	0	0

- Molecule 6 is water.

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	10	Total 10	O 10	0	0
6	B	19	Total 19	O 19	0	0

3 Residue-property plots ⓘ

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

- Molecule 1: Chaetomium alpha glucosidase
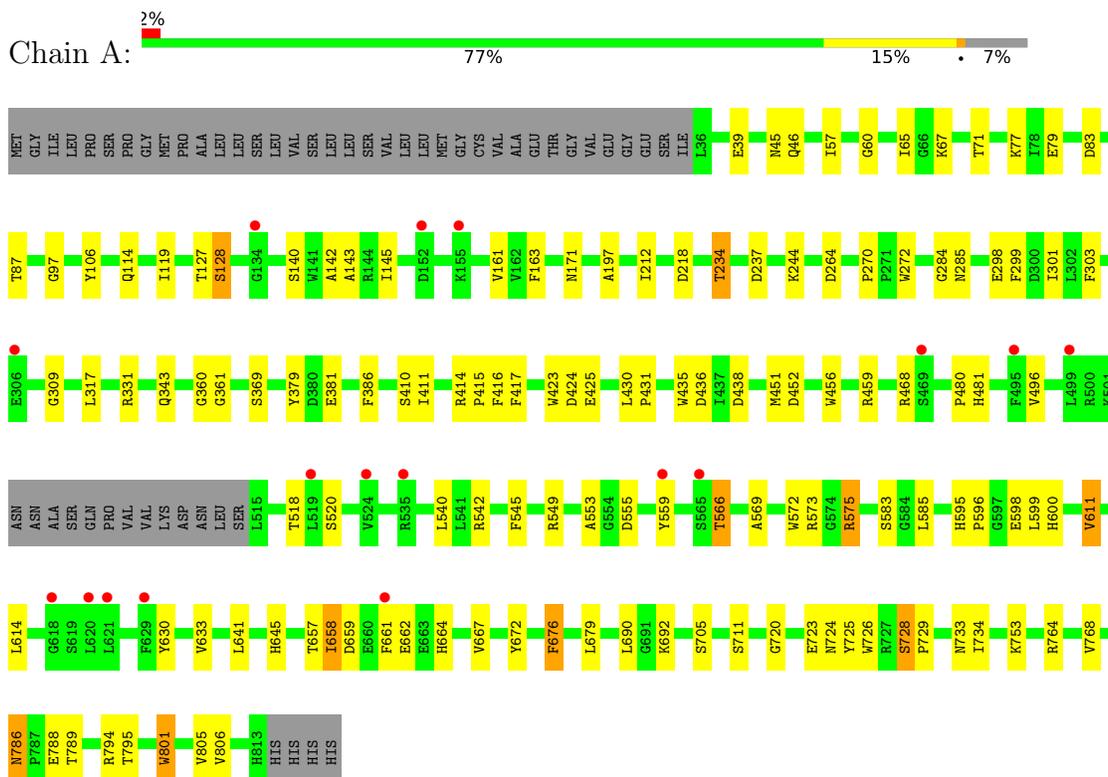
- Molecule 1: Chaetomium alpha glucosidase
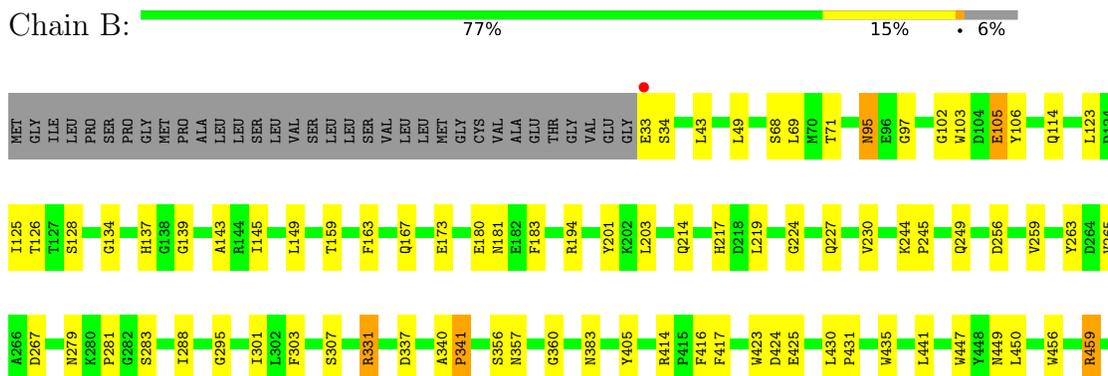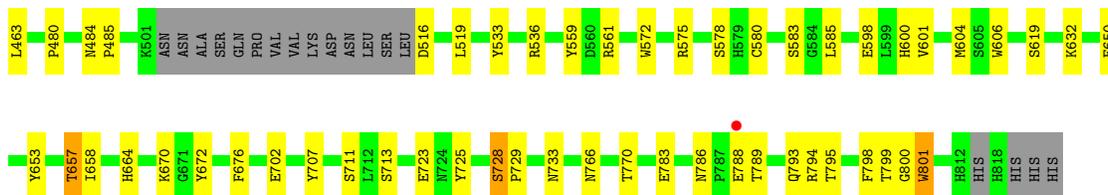
4 Data and refinement statistics (i)

Property	Value	Source
Space group	I 2 2 2	Depositor
Cell constants a, b, c, α, β, γ	135.21Å　179.21Å　180.59Å 90.00°　　90.00°　　90.00°	Depositor
Resolution (Å)	46.91　–　2.71 46.87　–　2.71	Depositor EDS
% Data completeness (in resolution range)	99.5 (46.91-2.71) 99.5 (46.87-2.71)	Depositor EDS
R_{merge}	0.23	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.07 (at 2.69Å)	Xtriage
Refinement program	REFMAC 5.8.0258	Depositor
R, R_{free}	0.194　,　0.244 0.196　,　0.245	Depositor DCC
R_{free} test set	3039 reflections (5.10%)	wwPDB-VP
Wilson B-factor (Å^2)	55.4	Xtriage
Anisotropy	0.551	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 24.9	EDS
L-test for twinning[2]	$< |L| > = 0.48, < L^2 > = 0.31$	Xtriage
Estimated twinning fraction	0.013 for -h,-l,-k	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	12295	wwPDB-VP
Average B, all atoms (Å^2)	60.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.09% 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, SO4, UGO, 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	0.69	0/6211	0.83	1/8468 (0.0%)
1	B	0.68	0/6235	0.82	0/8495
All	All	0.69	0/12446	0.82	1/16963 (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	B	0	1

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	575	ARG	CB-CA-C	6.21	122.81	110.40

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	657	THR	Peptide

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	6032	0	5648	69	0
1	B	6059	0	5667	74	0
2	A	6	0	8	1	0
2	B	24	0	32	2	0
3	A	33	0	0	0	0
3	B	33	0	0	0	0
4	A	40	0	0	1	0
4	B	25	0	0	1	0
5	B	14	0	13	0	0
6	A	10	0	0	0	0
6	B	19	0	0	0	0
All	All	12295	0	11368	143	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 (143) 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:B:572:TRP:H	1:B:600:HIS:HD2	1.18	0.88
1:A:598:GLU:OE2	1:A:600:HIS:HE1	1.56	0.87
1:A:553:ALA:O	1:A:573:ARG:NH2	2.11	0.83
1:A:598:GLU:OE2	1:A:600:HIS:CE1	2.31	0.83
1:B:128:SER:O	1:B:143:ALA:HA	1.79	0.83
1:A:658:ILE:HD12	1:A:662:GLU:HA	1.61	0.83
1:B:572:TRP:H	1:B:600:HIS:CD2	2.01	0.78
1:B:702:GLU:OE1	1:B:713:SER:OG	2.03	0.77
1:B:786:ASN:HD22	1:B:789:THR:H	1.33	0.75
1:B:561:ARG:HE	1:B:664:HIS:HD2	1.34	0.74
1:B:786:ASN:HB2	1:B:793:GLN:NE2	2.07	0.69
1:B:766:ASN:O	1:B:770:THR:HG23	1.93	0.68
1:B:244:LYS:HB3	1:B:245:PRO:HD3	1.76	0.68
1:B:430:LEU:HB2	1:B:431:PRO:HD3	1.75	0.68
1:B:450:LEU:HB3	4:B:909:SO4:O4	1.96	0.65
1:B:707:TYR:HA	1:B:770:THR:HG21	1.80	0.64
1:B:575:ARG:NH1	1:B:598:GLU:OE2	2.31	0.64
1:B:561:ARG:HE	1:B:664:HIS:CD2	2.14	0.63
1:B:95:ASN:HD22	1:B:97:GLY:H	1.45	0.63
1:B:786:ASN:ND2	1:B:788:GLU:H	1.97	0.62
1:A:234[A]:THR:HG23	1:A:284:GLY:HA2	1.82	0.61

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:217:HIS:HD2	1:B:219:LEU:H	1.49	0.60
1:B:134:GLY:O	1:B:139:GLY:HA2	2.02	0.60
1:A:234[B]:THR:HA	1:A:285:ASN:OD1	2.04	0.58
1:B:783:GLU:HA	1:B:798:PHE:CD2	2.39	0.58
1:B:485:PRO:HB3	1:B:606:TRP:CE2	2.42	0.55
1:A:614:LEU:HB3	1:A:630:TYR:CE2	2.43	0.54
1:B:137:HIS:HD2	1:B:307:SER:OG	1.90	0.54
1:B:786:ASN:HD21	1:B:788:GLU:HB2	1.73	0.53
1:A:456:TRP:CE2	1:A:480:PRO:HA	2.43	0.53
1:A:676:PHE:HA	1:A:679:LEU:HD12	1.90	0.53
1:A:171[B]:ASN:ND2	1:A:171[B]:ASN:H	2.05	0.53
1:B:102:GLY:CA	1:B:357:ASN:HD21	2.20	0.53
1:B:230:VAL:HA	1:B:288:ILE:O	2.08	0.53
1:A:720:GLY:HA2	1:A:724:ASN:OD1	2.08	0.53
1:B:559:TYR:CE1	1:B:658:ILE:HG13	2.44	0.53
1:B:103:TRP:H	1:B:357:ASN:ND2	2.06	0.52
1:B:114:GLN:NE2	1:B:414:ARG:HH12	2.07	0.52
1:B:723:GLU:O	1:B:794:ARG:NH2	2.43	0.52
1:A:106:TYR:CD2	1:A:360:GLY:HA2	2.46	0.51
1:B:145:ILE:CD1	1:B:301:ILE:CD1	2.88	0.51
1:A:77:LYS:HE3	1:A:79:GLU:OE1	2.11	0.51
1:B:601:VAL:HG21	1:B:653:TYR:HB3	1.93	0.51
1:B:585:LEU:HD12	1:B:585:LEU:N	2.24	0.51
1:B:580:CYS:SG	1:B:583:SER:HB3	2.51	0.51
1:A:161:VAL:HG11	1:A:299:PHE:HE2	1.76	0.50
1:A:672:TYR:CG	1:A:734:ILE:HG21	2.47	0.50
1:B:733:ASN:HB3	1:B:801:TRP:CG	2.47	0.50
1:A:71:THR:HB	1:A:163:PHE:CZ	2.46	0.50
1:A:331:ARG:NH1	1:A:435:TRP:O	2.45	0.49
1:A:234[A]:THR:HA	1:A:285:ASN:OD1	2.12	0.49
1:B:263:TYR:CE1	1:B:265:VAL:HG12	2.47	0.49
1:A:106:TYR:CE2	1:A:360:GLY:HA2	2.48	0.49
1:A:128:SER:O	1:A:143:ALA:HA	2.12	0.49
1:A:451:MET:HE2	1:A:540:LEU:HB3	1.93	0.49
1:A:410:SER:OG	1:A:425:GLU:OE2	2.30	0.48
1:A:39:GLU:HA	1:A:39:GLU:OE2	2.14	0.48
1:B:43:LEU:HB3	1:B:519:LEU:HD11	1.96	0.48
1:B:331:ARG:NH1	1:B:435:TRP:O	2.46	0.47
1:B:106:TYR:CE2	1:B:360:GLY:HA2	2.49	0.47
1:B:180:GLU:C	1:B:181:ASN:HD22	2.18	0.47
1:A:733:ASN:HB3	1:A:801:TRP:CG	2.48	0.47

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	1.97	0.47
1:B:103:TRP:H	1:B:357:ASN:HD21	1.63	0.47
1:B:214:GLN:NE2	1:B:214:GLN:HA	2.29	0.47
1:B:167:GLN:NE2	1:B:201:TYR:OH	2.47	0.47
1:A:234[A]:THR:HG23	1:A:284:GLY:CA	2.43	0.47
1:A:786:ASN:HD22	1:A:789:THR:H	1.62	0.47
1:A:57:ILE:HD11	1:A:60:GLY:HA3	1.97	0.47
1:A:145:ILE:O	1:A:298:GLU:HA	2.15	0.46
1:A:414:ARG:N	1:A:415:PRO:CD	2.78	0.46
1:A:659:ASP:HB3	1:A:661:PHE:H	1.80	0.46
1:A:545:PHE:CZ	1:A:549:ARG:HD2	2.50	0.46
1:A:430:LEU:HB2	1:A:431:PRO:HD3	1.97	0.46
1:A:555:ASP:HB3	1:A:559:TYR:HE2	1.80	0.46
1:A:786:ASN:ND2	1:A:788:GLU:H	2.13	0.46
1:A:728:SER:N	1:A:729:PRO:CD	2.79	0.46
1:A:452:ASP:OD1	1:A:452:ASP:C	2.53	0.46
1:B:71:THR:HB	1:B:163:PHE:CZ	2.50	0.46
1:B:728:SER:OG	1:B:729:PRO:HD3	2.16	0.46
1:A:171[B]:ASN:H	1:A:171[B]:ASN:HD22	1.63	0.46
1:A:317:LEU:HD12	1:A:317:LEU:O	2.16	0.46
1:B:68:SER:OG	1:B:69:LEU:N	2.47	0.46
1:B:456:TRP:CE2	1:B:480:PRO:HA	2.51	0.46
1:A:566:THR:OG1	2:A:1101:GOL:O3	2.31	0.45
1:B:707:TYR:CA	1:B:770:THR:HG21	2.45	0.45
1:A:161:VAL:HG11	1:A:299:PHE:CE2	2.51	0.45
1:B:430:LEU:CB	1:B:431:PRO:HD3	2.46	0.45
1:B:799:THR:N	1:B:800:GLY:HA3	2.32	0.45
1:A:764:ARG:O	1:A:768:VAL:HG23	2.16	0.45
1:B:114:GLN:HE22	1:B:414:ARG:HH22	1.65	0.45
1:B:123:LEU:HD11	1:B:159:THR:HG21	1.98	0.45
1:B:279:ASN:ND2	1:B:281:PRO:HD3	2.31	0.45
1:B:183:PHE:CD2	1:B:224:GLY:HA2	2.52	0.44
1:A:468:ARG:NE	4:A:1107:SO4:O1	2.43	0.44
1:A:572:TRP:CZ2	1:A:583:SER:HA	2.52	0.44
1:B:106:TYR:CD2	1:B:360:GLY:HA2	2.52	0.44
1:B:650:GLU:OE1	1:B:670:LYS:NZ	2.46	0.44
1:B:672:TYR:CE1	1:B:711:SER:HA	2.52	0.44
1:B:149:LEU:HD11	1:B:295:GLY:HA2	1.99	0.44
1:B:201:TYR:CD1	1:B:303:PHE:HE1	2.35	0.44
1:A:569:ALA:HB1	1:A:599:LEU:HD22	1.99	0.44
1:B:561:ARG:NE	1:B:664:HIS:HD2	2.09	0.44

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:456:TRP:CD2	1:A:480:PRO:HA	2.53	0.44
1:A:414:ARG:N	1:A:415:PRO:HD2	2.33	0.44
1:B:33:GLU:HB2	1:B:536:ARG:HH12	1.83	0.43
1:B:405:TYR:HB2	1:B:463:LEU:HD22	1.99	0.43
1:B:459:ARG:HB3	1:B:484:ASN:HB2	2.00	0.43
1:A:142:ALA:HA	1:A:301:ILE:O	2.18	0.43
1:A:83:ASP:O	1:A:87:THR:HG23	2.18	0.43
1:A:140:SER:HB3	1:A:303:PHE:O	2.19	0.43
1:A:381:GLU:HA	1:A:386:PHE:HB2	2.00	0.43
1:A:415:PRO:HG2	1:A:416:PHE:CD2	2.54	0.42
1:A:451:MET:CE	1:A:540:LEU:HD22	2.49	0.42
1:A:572:TRP:H	1:A:600:HIS:HD2	1.66	0.42
1:A:65:ILE:HD13	1:A:197:ALA:HB1	2.01	0.42
1:A:658:ILE:HD12	1:A:662:GLU:CA	2.42	0.42
1:B:105:GLU:HA	1:B:356:SER:OG	2.18	0.42
1:A:585:LEU:HD23	1:A:726:TRP:CE3	2.54	0.42
1:B:203:LEU:HD11	1:B:301:ILE:HG23	2.00	0.42
1:B:575:ARG:HE	1:B:575:ARG:HB2	1.54	0.42
1:A:611:VAL:HG11	1:A:633:VAL:HG12	2.01	0.42
1:A:672:TYR:CE2	1:A:711:SER:HA	2.55	0.42
1:B:416:PHE:CD2	1:B:795:THR:HB	2.53	0.42
1:B:441:LEU:HD22	1:B:533:TYR:CE1	2.55	0.42
1:A:595:HIS:ND1	1:A:596:PRO:HD2	2.35	0.42
1:B:145:ILE:CD1	1:B:301:ILE:HD11	2.50	0.42
1:A:97:GLY:O	1:A:119:ILE:HB	2.20	0.42
1:A:641:LEU:O	1:A:645:HIS:HB2	2.20	0.42
1:B:244:LYS:NZ	2:B:902:GOL:H12	2.35	0.42
1:A:436:ASP:OD1	1:A:520:SER:HA	2.21	0.41
1:A:142:ALA:HB1	1:A:317:LEU:HD22	2.03	0.41
1:A:270:PRO:HB2	1:A:272:TRP:CD1	2.56	0.41
1:A:361:GLY:O	1:A:411:ILE:HG12	2.20	0.41
1:A:379:TYR:HB3	1:A:481:HIS:CE1	2.55	0.41
1:A:114:GLN:HB3	1:A:127:THR:OG1	2.20	0.41
1:B:430:LEU:HD23	1:B:430:LEU:HA	1.93	0.41
1:A:728:SER:OG	1:A:729:PRO:HD3	2.21	0.40
1:B:256:ASP:O	1:B:259:VAL:HG22	2.20	0.40
1:B:340:ALA:HA	1:B:341:PRO:HA	1.90	0.40
1:A:723:GLU:O	1:A:794:ARG:NH1	2.54	0.40
1:B:244:LYS:HZ1	2:B:902:GOL:H12	1.86	0.40
1:B:425:GLU:HG2	1:B:447:TRP:NE1	2.36	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	763/819 (93%)	726 (95%)	35 (5%)	2 (0%)	41	65
1	B	762/819 (93%)	728 (96%)	32 (4%)	2 (0%)	41	65
All	All	1525/1638 (93%)	1454 (95%)	67 (4%)	4 (0%)	41	65

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	728	SER
1	B	728	SER
1	A	309	GLY
1	B	337	ASP

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.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	617/707 (87%)	578 (94%)	39 (6%)	18	38
1	B	623/707 (88%)	595 (96%)	28 (4%)	27	53
All	All	1240/1414 (88%)	1173 (95%)	67 (5%)	22	45

All (67) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	45	ASN
1	A	46	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	67	LYS
1	A	128	SER
1	A	212	ILE
1	A	218	ASP
1	A	234[A]	THR
1	A	234[B]	THR
1	A	237	ASP
1	A	244	LYS
1	A	264	ASP
1	A	343	GLN
1	A	369	SER
1	A	417	PHE
1	A	423	TRP
1	A	424	ASP
1	A	438	ASP
1	A	459	ARG
1	A	496	VAL
1	A	518	THR
1	A	542	ARG
1	A	566	THR
1	A	575	ARG
1	A	611	VAL
1	A	657	THR
1	A	658	ILE
1	A	664	HIS
1	A	667	VAL
1	A	676	PHE
1	A	690	LEU
1	A	692	LYS
1	A	705	SER
1	A	725	TYR
1	A	753	LYS
1	A	786	ASN
1	A	795	THR
1	A	801	TRP
1	A	805	VAL
1	A	806	VAL
1	B	34	SER
1	B	49	LEU
1	B	95	ASN
1	B	105	GLU
1	B	125	ILE

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	126	THR
1	B	173	GLU
1	B	194	ARG
1	B	227	GLN
1	B	249	GLN
1	B	283	SER
1	B	331	ARG
1	B	341	PRO
1	B	383	ASN
1	B	417	PHE
1	B	423	TRP
1	B	424	ASP
1	B	449	ASN
1	B	459	ARG
1	B	516	ASP
1	B	578	SER
1	B	604	MET
1	B	619	SER
1	B	632	LYS
1	B	657	THR
1	B	676	PHE
1	B	725	TYR
1	B	801	TRP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (34) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	85	GLN
1	A	120	GLN
1	A	167	GLN
1	A	231	GLN
1	A	250	GLN
1	A	323	GLN
1	A	397	HIS
1	A	461	GLN
1	A	484	ASN
1	A	600	HIS
1	A	639	HIS
1	A	748	GLN
1	A	786	ASN
1	A	793	GLN
1	B	95	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	114	GLN
1	B	120	GLN
1	B	137	HIS
1	B	167	GLN
1	B	181	ASN
1	B	217	HIS
1	B	227	GLN
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	664	HIS
1	B	786	ASN
1	B	793	GLN
1	B	812	HIS

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 ⓘ

21 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	$\#|Z| > 2$
2	GOL	A	1101	-	5,5,5	0.12	0	5,5,5	0.33	0
2	GOL	B	905	-	5,5,5	0.16	0	5,5,5	0.40	0
4	SO4	A	1103	-	4,4,4	0.46	0	6,6,6	0.27	0
4	SO4	B	910	-	4,4,4	0.37	0	6,6,6	0.10	0
4	SO4	B	911	-	4,4,4	0.33	0	6,6,6	0.21	0
2	GOL	B	903	-	5,5,5	0.14	0	5,5,5	0.28	0
4	SO4	B	908	-	4,4,4	0.32	0	6,6,6	0.12	0
3	UGO	B	906	-	34,35,35	3.92	13 (38%)	41,47,47	1.69	7 (17%)
4	SO4	A	1104	-	4,4,4	0.33	0	6,6,6	0.09	0
4	SO4	A	1109	-	4,4,4	0.36	0	6,6,6	0.07	0
4	SO4	A	1108	-	4,4,4	0.40	0	6,6,6	0.08	0
4	SO4	A	1106	-	4,4,4	0.33	0	6,6,6	0.19	0
5	NAG	B	901	1	14,14,15	1.09	1 (7%)	17,19,21	2.04	6 (35%)
4	SO4	A	1110	-	4,4,4	0.32	0	6,6,6	0.09	0
2	GOL	B	902	-	5,5,5	0.20	0	5,5,5	0.50	0
4	SO4	A	1107	-	4,4,4	0.32	0	6,6,6	0.11	0
3	UGO	A	1102	-	34,35,35	3.96	15 (44%)	41,47,47	1.82	9 (21%)
4	SO4	A	1105	-	4,4,4	0.33	0	6,6,6	0.18	0
4	SO4	B	907	-	4,4,4	0.42	0	6,6,6	0.24	0
4	SO4	B	909	-	4,4,4	0.38	0	6,6,6	0.24	0
2	GOL	B	904	-	5,5,5	0.12	0	5,5,5	0.31	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	GOL	B	902	-	-	4/4/4/4	-
2	GOL	B	903	-	-	4/4/4/4	-
2	GOL	A	1101	-	-	2/4/4/4	-
3	UGO	B	906	-	-	7/18/40/40	0/3/3/3
3	UGO	A	1102	-	-	2/18/40/40	0/3/3/3
5	NAG	B	901	1	-	0/6/23/26	0/1/1/1
2	GOL	B	905	-	-	2/4/4/4	-
2	GOL	B	904	-	-	0/4/4/4	-

All (29) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	906	UGO	C5-N1	-13.73	1.23	1.47
3	A	1102	UGO	C5-N1	-13.18	1.23	1.47
3	A	1102	UGO	O4-N3	12.06	1.43	1.22
3	B	906	UGO	O4-N3	11.00	1.41	1.22
3	B	906	UGO	C22-C1	-6.33	1.43	1.52
3	A	1102	UGO	C22-C1	-6.07	1.43	1.52
3	B	906	UGO	C11-N2	6.07	1.53	1.37
3	A	1102	UGO	C11-N2	5.89	1.53	1.37
3	B	906	UGO	C4-C3	-5.11	1.44	1.52
3	A	1102	UGO	C13-C12	4.60	1.48	1.39
3	B	906	UGO	C13-C12	4.05	1.47	1.39
3	A	1102	UGO	C19-C14	4.01	1.47	1.39
3	A	1102	UGO	C20-C11	3.99	1.46	1.39
3	A	1102	UGO	C4-C3	-3.81	1.46	1.52
3	A	1102	UGO	O1-C1	3.76	1.51	1.43
3	B	906	UGO	C19-C14	3.72	1.47	1.39
3	B	906	UGO	O1-C1	3.71	1.51	1.43
3	B	906	UGO	C20-C11	3.58	1.45	1.39
3	A	1102	UGO	C2-C1	-3.09	1.44	1.52
3	B	906	UGO	C2-C1	-3.04	1.44	1.52
3	A	1102	UGO	C2-C3	2.94	1.59	1.53
3	A	1102	UGO	O6-C22	2.90	1.49	1.43
3	B	906	UGO	O6-C22	2.90	1.49	1.43
3	A	1102	UGO	C19-C20	2.54	1.43	1.38
5	B	901	NAG	C1-C2	2.41	1.55	1.52
3	A	1102	UGO	C17-C18	2.38	1.44	1.37
3	B	906	UGO	C17-C18	2.31	1.44	1.37
3	B	906	UGO	C21-N1	-2.12	1.44	1.47
3	A	1102	UGO	C3-N1	2.12	1.53	1.48

All (22) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	A	1102	UGO	C18-N5-C15	4.77	121.06	116.08
3	A	1102	UGO	C16-N4-C15	4.72	121.02	116.08
3	B	906	UGO	C16-N4-C15	4.66	120.95	116.08
5	B	901	NAG	C2-N2-C7	4.23	128.92	122.90
3	A	1102	UGO	C12-C11-N2	-3.90	116.48	123.33
3	B	906	UGO	C18-N5-C15	3.86	120.12	116.08
3	B	906	UGO	C4-C3-C2	-3.84	107.01	112.90
5	B	901	NAG	C4-C3-C2	3.76	116.53	111.02
3	A	1102	UGO	C17-C16-N4	-3.17	118.26	123.43
3	A	1102	UGO	C17-C18-N5	-3.12	118.33	123.43

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	1102	UGO	C21-C22-C1	3.05	113.73	110.24
5	B	901	NAG	C1-O5-C5	2.98	116.22	112.19
3	B	906	UGO	C17-C16-N4	-2.96	118.58	123.43
3	B	906	UGO	C17-C18-N5	-2.83	118.80	123.43
5	B	901	NAG	O3-C3-C2	-2.56	104.17	109.47
3	A	1102	UGO	C4-C3-C2	-2.35	109.29	112.90
3	B	906	UGO	C20-C11-N2	-2.29	117.90	121.80
5	B	901	NAG	O5-C5-C6	2.22	110.68	107.20
5	B	901	NAG	C1-C2-N2	2.18	114.21	110.49
3	A	1102	UGO	C20-C11-N2	2.10	125.37	121.80
3	A	1102	UGO	C18-C17-C16	2.08	119.64	116.67
3	B	906	UGO	O3-C4-C3	-2.00	107.13	111.42

There are no chirality outliers.

All (21) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	1101	GOL	O1-C1-C2-O2
2	A	1101	GOL	O1-C1-C2-C3
2	B	902	GOL	C1-C2-C3-O3
2	B	903	GOL	O1-C1-C2-C3
2	B	903	GOL	C1-C2-C3-O3
2	B	905	GOL	C1-C2-C3-O3
3	A	1102	UGO	C11-C12-N3-O4
3	B	906	UGO	C12-C11-N2-C10
3	B	906	UGO	C20-C11-N2-C10
2	B	903	GOL	O2-C2-C3-O3
3	B	906	UGO	C9-C10-N2-C11
3	A	1102	UGO	C13-C12-N3-O4
3	B	906	UGO	C6-C7-C8-C9
2	B	902	GOL	O2-C2-C3-O3
2	B	905	GOL	O2-C2-C3-O3
2	B	903	GOL	O1-C1-C2-O2
2	B	902	GOL	O1-C1-C2-O2
3	B	906	UGO	N2-C10-C9-C8
3	B	906	UGO	C13-C12-N3-O4
3	B	906	UGO	C5-C6-C7-C8
2	B	902	GOL	O1-C1-C2-C3

There are no ring outliers.

4 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	1101	GOL	1	0
2	B	902	GOL	2	0
4	A	1107	SO4	1	0
4	B	909	SO4	1	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 UGO B 906

Bond lengths

Bond angles

Torsions

Rings

Ligand UGO A 1102

Bond lengths

Bond angles

Torsions

Rings
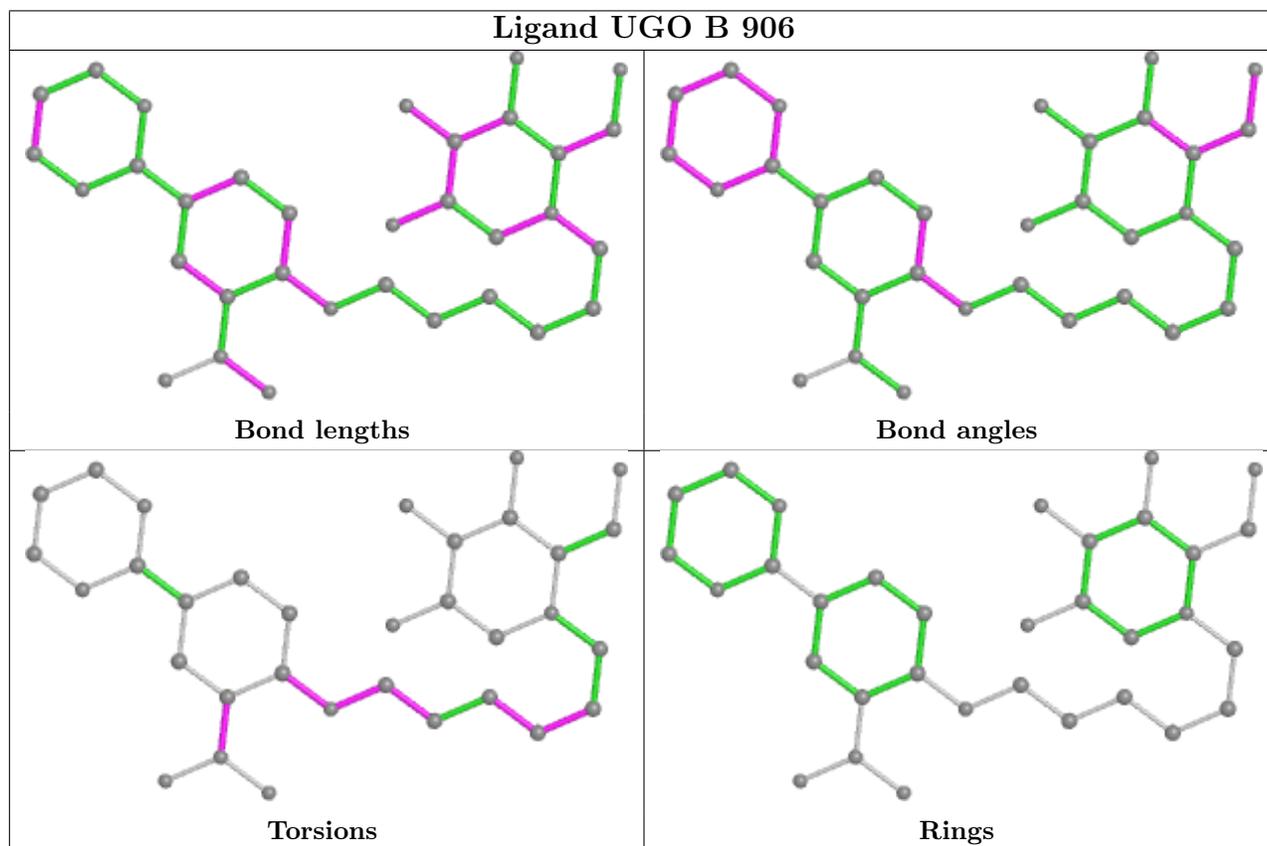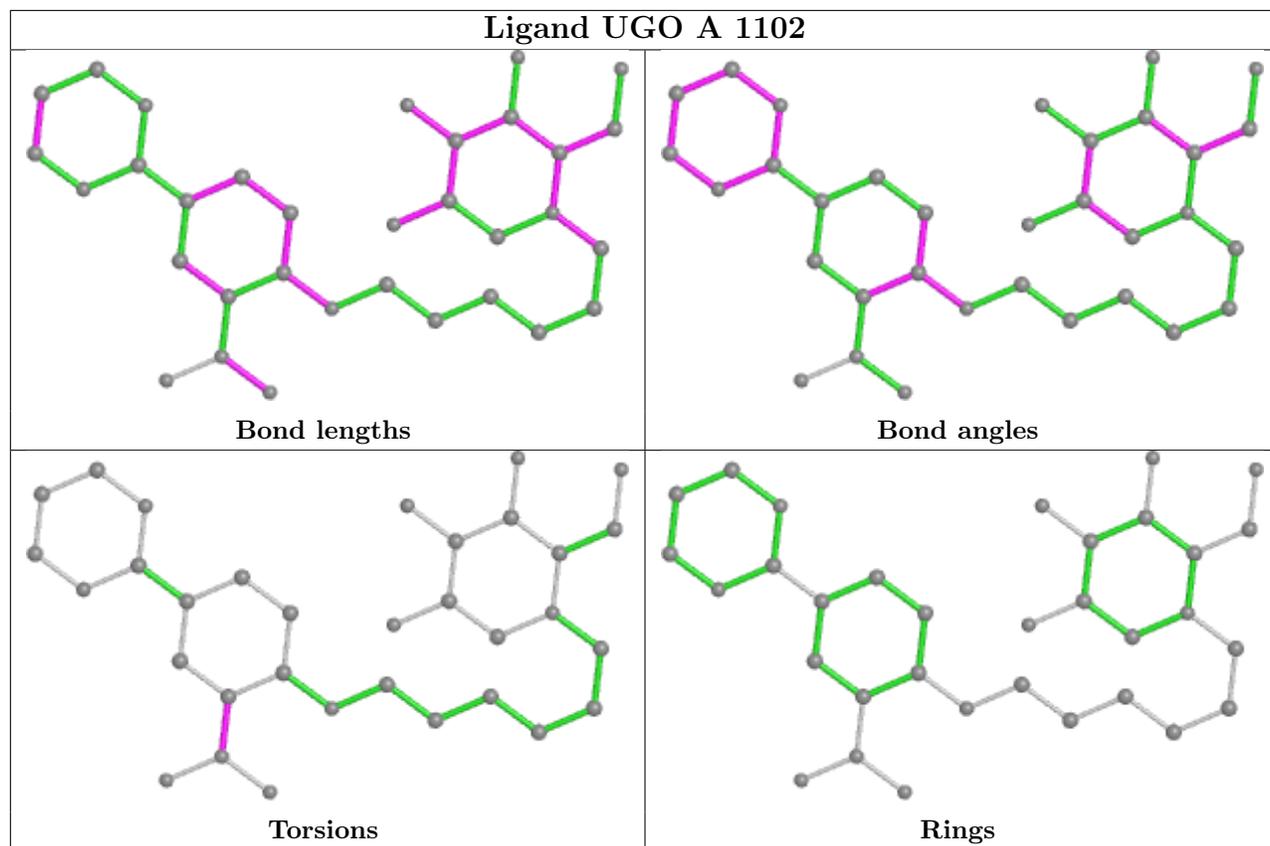
5.7 Other polymers ⓘ

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	765/819 (93%)	-0.14	17 (2%)	62	63	37, 60, 85, 105	0
1	B	767/819 (93%)	-0.28	2 (0%)	94	95	40, 56, 78, 114	0
All	All	1532/1638 (93%)	-0.21	19 (1%)	79	80	37, 58, 82, 114	0

All (19) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	621	LEU	4.6
1	A	155	LYS	3.1
1	A	661	PHE	3.0
1	A	535	ARG	3.0
1	A	499	LEU	3.0
1	A	629	PHE	2.8
1	A	620	LEU	2.8
1	B	788	GLU	2.7
1	A	519	LEU	2.7
1	B	33	GLU	2.6
1	A	134	GLY	2.5
1	A	524	VAL	2.4
1	A	495	PHE	2.3
1	A	618	GLY	2.2
1	A	565	SER	2.2
1	A	306	GLU	2.2
1	A	559	TYR	2.1
1	A	469	SER	2.1
1	A	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 ⓘ

There are no monosaccharides in this entry.

6.4 Ligands ⓘ

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

| Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
4	SO4	A	1109	5/5	0.80	0.42	125,129,139,142	0
2	GOL	B	902	6/6	0.83	0.29	73,85,95,96	0
4	SO4	A	1105	5/5	0.84	0.29	107,114,126,127	0
2	GOL	B	903	6/6	0.85	0.23	67,71,73,73	0
2	GOL	A	1101	6/6	0.89	0.24	87,92,101,103	0
2	GOL	B	904	6/6	0.89	0.15	76,90,97,100	0
4	SO4	A	1110	5/5	0.89	0.31	98,108,119,122	0
2	GOL	B	905	6/6	0.90	0.26	91,92,93,94	0
5	NAG	B	901	14/15	0.91	0.34	65,82,88,92	0
3	UGO	A	1102	33/33	0.93	0.22	32,66,100,104	0
4	SO4	B	910	5/5	0.94	0.10	101,101,105,120	0
4	SO4	A	1108	5/5	0.94	0.12	100,110,117,121	0
4	SO4	B	909	5/5	0.95	0.10	65,68,93,99	0
4	SO4	A	1106	5/5	0.95	0.22	85,88,104,122	0
3	UGO	B	906	33/33	0.95	0.20	37,89,128,132	0
4	SO4	A	1104	5/5	0.96	0.07	86,97,104,105	0
4	SO4	A	1107	5/5	0.96	0.10	85,92,95,99	0
4	SO4	B	908	5/5	0.97	0.11	75,80,85,86	0
4	SO4	B	911	5/5	0.98	0.12	62,77,79,83	0
4	SO4	B	907	5/5	0.99	0.08	41,55,62,76	0
4	SO4	A	1103	5/5	0.99	0.16	50,51,56,57	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 UGO A 1102:

$2mF_o\text{-}DF_c$ (at 0.7 rmsd) in gray
$mF_o\text{-}DF_c$ (at 3 rmsd) in purple (negative)
and green (positive)
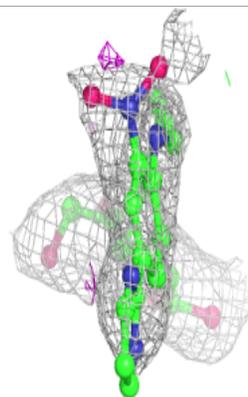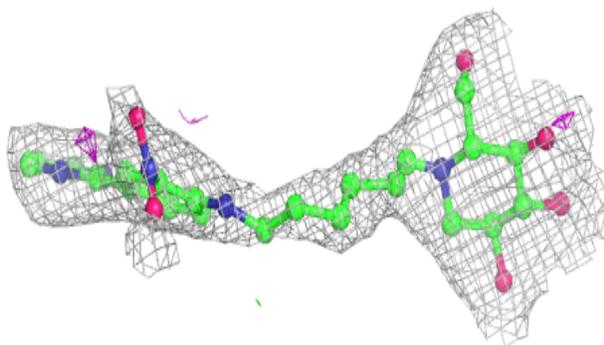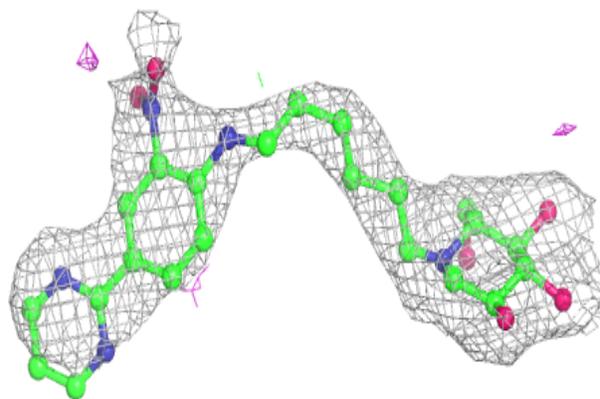
Electron density around UGO B 906:

$2mF_o\text{-}DF_c$ (at 0.7 rmsd) in gray
$mF_o\text{-}DF_c$ (at 3 rmsd) in purple (negative)
and green (positive)
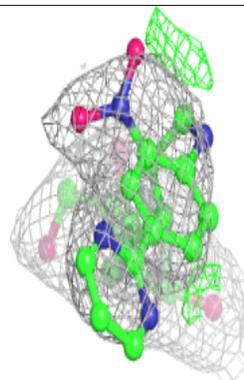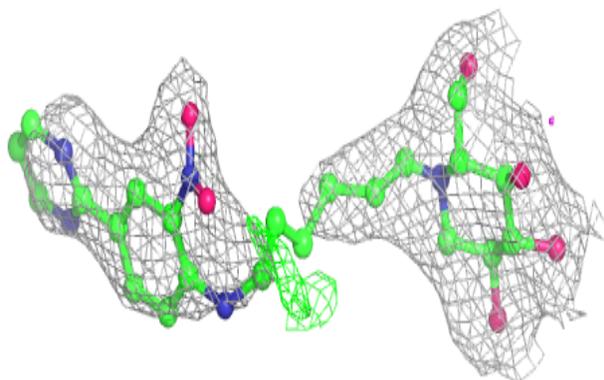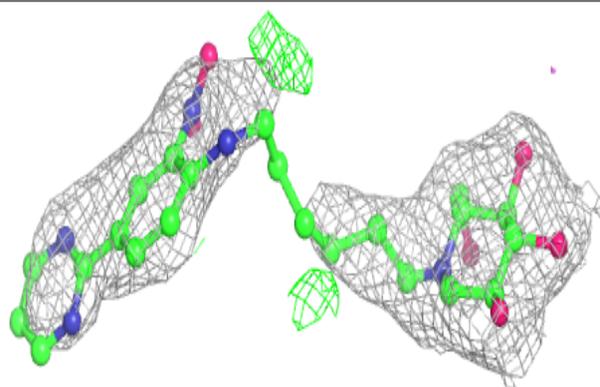
6.5 Other polymers (i)

There are no such residues in this entry.