Full wwPDB X-ray Structure Validation Report (i)

Oct 10, 2023 – 02:55 AM EDT

PDB ID	:	7T68
Title	:	Co-crystal structure of Chaetomium glucosidase with compound UV-5
Authors	:	Karade, S.S.; Mariuzza, R.A.
Deposited on	:	2021-12-13
Resolution	:	2.32 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:

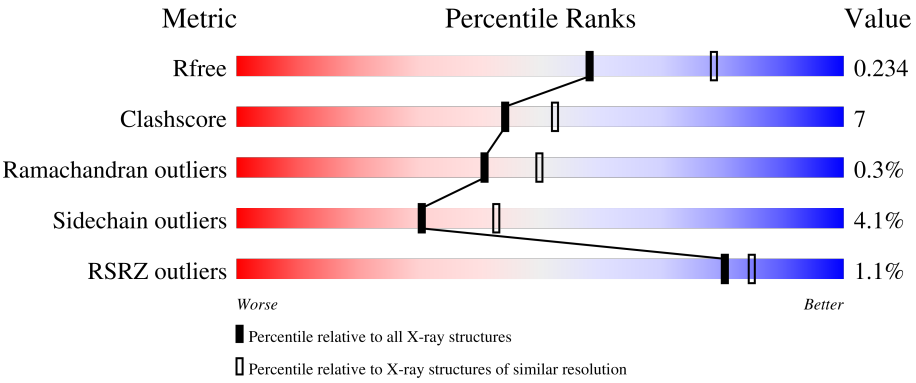
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35.1

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 2.32 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	5974 (2.34-2.30)
Clashscore	141614	6604 (2.34-2.30)
Ramachandran outliers	138981	6523 (2.34-2.30)
Sidechain outliers	138945	6523 (2.34-2.30)
RSRZ outliers	127900	5855 (2.34-2.30)

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% 11% • 7%
1	B	819	79% 13% • 7%

2 Entry composition (i)

There are 7 unique types of molecules in this entry. The entry contains 12634 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	4	0
			6096	3918	1029	1135	14			
1	B	764	Total	C	N	O	S	0	3	0
			6078	3897	1020	1147	14			

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...

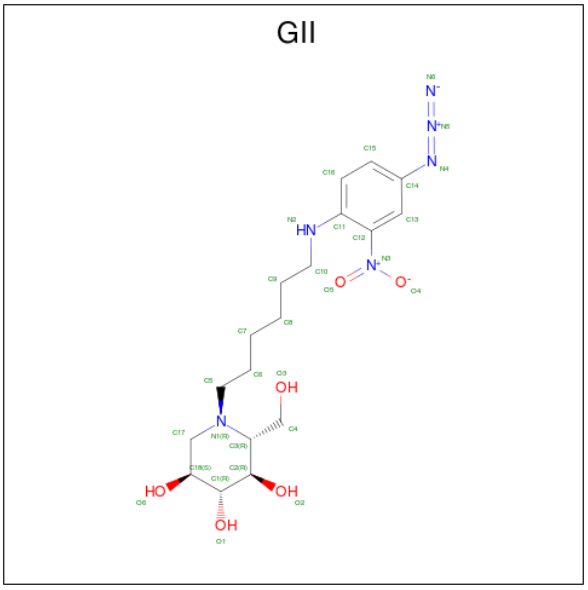
Chain	Residue	Modelled	Actual	Comment	Reference
A	24	CYS	-	expression tag	UNP G0SFD1
A	25	VAL	-	expression tag	UNP G0SFD1
A	26	ALA	-	expression tag	UNP G0SFD1
A	27	GLU	-	expression tag	UNP G0SFD1
A	28	THR	-	expression tag	UNP G0SFD1
A	29	GLY	-	expression tag	UNP G0SFD1
A	810	SER	-	expression tag	UNP 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...

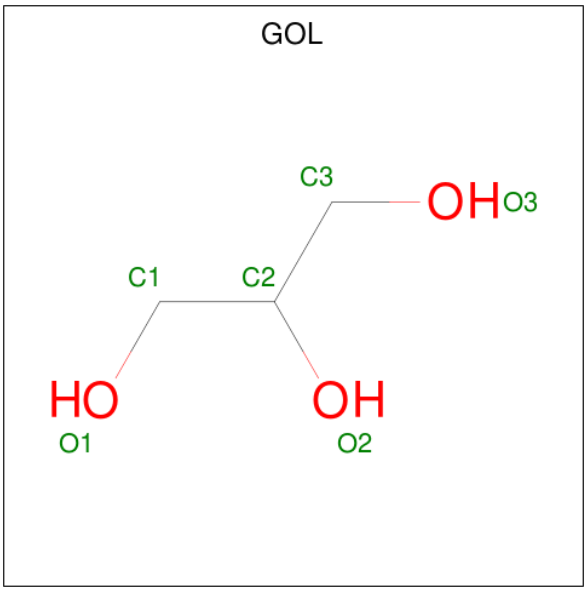
Chain	Residue	Modelled	Actual	Comment	Reference
B	27	GLU	-	expression tag	UNP G0SFD1
B	28	THR	-	expression tag	UNP G0SFD1
B	29	GLY	-	expression tag	UNP G0SFD1
B	810	SER	-	expression tag	UNP G0SFD1
B	811	GLY	-	expression tag	UNP G0SFD1
B	812	HIS	-	expression tag	UNP G0SFD1
B	813	HIS	-	expression tag	UNP 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)-1-[6-(4-azido-2-nitroanilino)hexyl]-2-(hydroxymethyl)piperidine-3,4,5-triol (three-letter code: GII) (formula: $C_{18}H_{28}N_6O_6$) (labeled as "Ligand of Interest" by depositor).



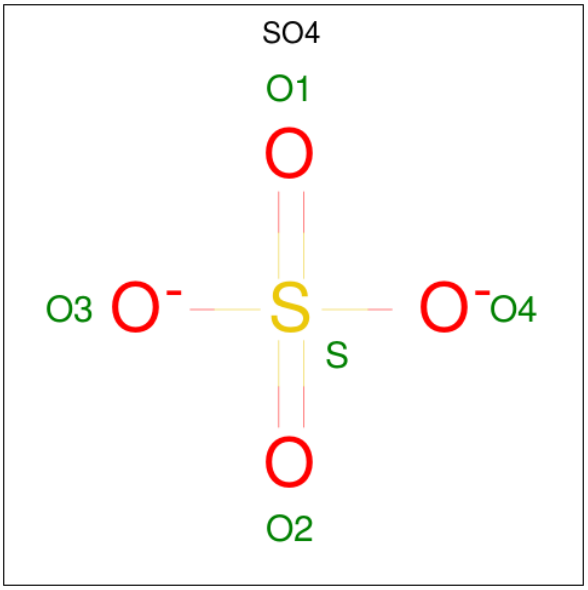
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	N	O	0	0
			30	18	6	6		
2	B	1	Total	C	N	O	0	0
			30	18	6	6		

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



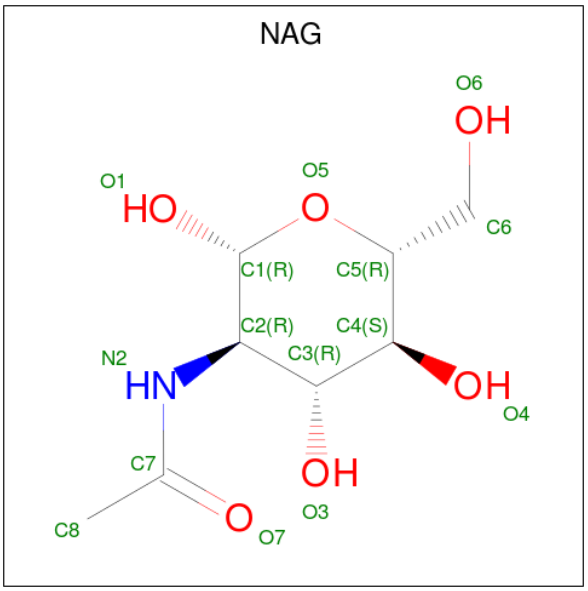
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total O S 5 4 1	0	0
4	A	1	Total O S 5 4 1	0	0

Continued on next page...

Continued from previous page...

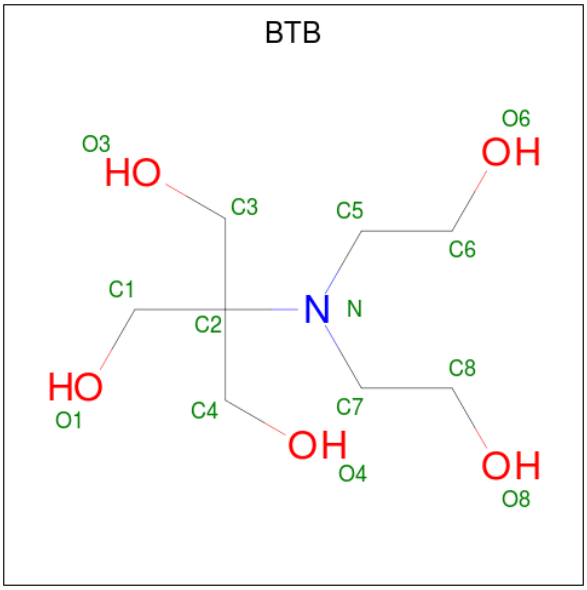
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total O S 5 4 1	0	0
4	A	1	Total O S 5 4 1	0	0
4	A	1	Total O S 5 4 1	0	0
4	A	1	Total O S 5 4 1	0	0
4	A	1	Total O S 5 4 1	0	0
4	A	1	Total O S 5 4 1	0	0
4	B	1	Total O S 5 4 1	0	0
4	B	1	Total O S 5 4 1	0	0
4	B	1	Total O S 5 4 1	0	0
4	B	1	Total O S 5 4 1	0	0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

- 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	0	0
			14	8	1	5		

- Molecule 7 is water.

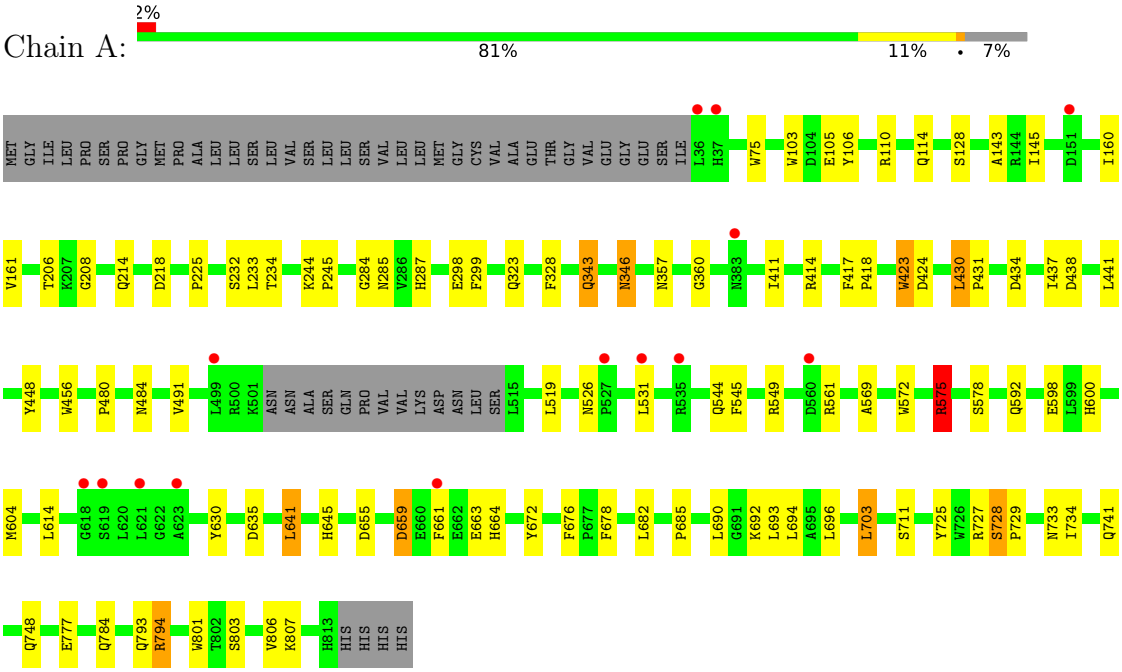
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	141	Total	O	0	0
			141	141		
7	B	154	Total	O	0	0
			154	154		

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.

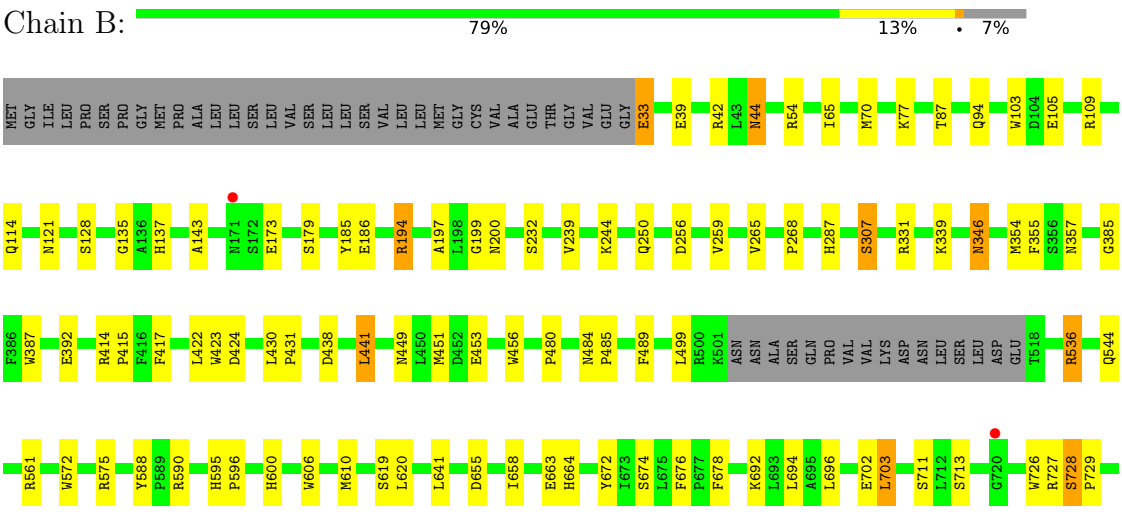
• Molecule 1: Chaetomium alpha glucosidase

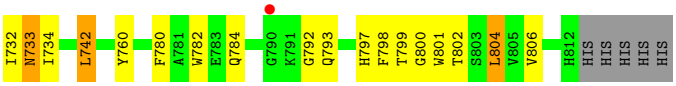
Chain A:



• Molecule 1: Chaetomium alpha glucosidase

Chain B:




I732
N733
I734
L742
Y760
F780
A781
W782
E783
Q784
G790
K791
G792
Q793
H797
F798
T799
G800
W801
T802
S803
L804
V805
V806
H812
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	135.75Å 178.38Å 179.57Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	46.72 – 2.32 46.67 – 2.32	Depositor EDS
% Data completeness (in resolution range)	99.7 (46.72-2.32) 99.7 (46.67-2.32)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.94 (at 2.32Å)	Xtriage
Refinement program	REFMAC 5.8.0258	Depositor
R, R_{free}	0.190 , 0.234 0.191 , 0.234	Depositor DCC
R_{free} test set	4710 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å^2)	40.2	Xtriage
Anisotropy	0.409	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 27.1	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	0.95	EDS
Total number of atoms	12634	wwPDB-VP
Average B, all atoms (Å^2)	41.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.43% 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: BTB, GOL, SO4, NAG, GII

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	0.40	0/6281	0.67	2/8549 (0.0%)
1	B	0.41	0/6254	0.67	1/8517 (0.0%)
All	All	0.41	0/12535	0.67	3/17066 (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	2

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	575	ARG	CB-CA-C	6.77	123.94	110.40
1	A	777	GLU	CB-CA-C	5.68	121.76	110.40
1	B	536	ARG	CG-CD-NE	-5.14	101.00	111.80

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	732	ILE	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	6096	0	5784	79	0
1	B	6078	0	5705	81	0
2	A	30	0	0	0	0
2	B	30	0	0	0	0
3	A	6	0	8	3	0
3	B	6	0	8	0	0
4	A	40	0	0	1	0
4	B	25	0	0	0	0
5	B	14	0	13	0	0
6	B	14	0	19	1	0
7	A	141	0	0	0	0
7	B	154	0	0	2	0
All	All	12634	0	11537	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 7.

All (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:B:561:ARG:HE	1:B:664:HIS:HD2	1.05	0.99
1:A:659:ASP:HB3	1:A:661:PHE:H	1.33	0.93
1:B:572:TRP:H	1:B:600:HIS:HD2	1.06	0.92
1:B:346:ASN:HD22	1:B:346:ASN:H	1.19	0.90
1:B:561:ARG:HE	1:B:664:HIS:CD2	1.91	0.89
1:B:572:TRP:H	1:B:600:HIS:CD2	1.90	0.87
1:A:784:GLN:HE21	1:A:793:GLN:HE21	1.26	0.83
1:B:135:GLY:HA3	1:B:307:SER:HB3	1.57	0.83
1:B:33:GLU:HB3	1:B:536:ARG:NH2	1.96	0.81
1:B:114:GLN:NE2	1:B:414:ARG:HH12	1.81	0.77
1:A:561:ARG:HE	1:A:664:HIS:HD2	1.32	0.77
1:A:208:GLY:HA2	3:A:902:GOL:H31	1.65	0.76
1:A:572:TRP:H	1:A:600:HIS:HD2	1.33	0.75
1:B:561:ARG:NE	1:B:664:HIS:HD2	1.84	0.74
1:A:690:LEU:C	1:A:690:LEU:HD23	2.07	0.74

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:784:GLN:NE2	1:A:793:GLN:HE21	1.85	0.74
1:A:234[B]:THR:HG22	7:B:1101:HOH:O	1.89	0.73
1:B:87:THR:HB	1:B:121:ASN:HD21	1.54	0.71
1:A:659:ASP:HB2	1:A:663:GLU:H	1.56	0.70
1:B:801:TRP:O	1:B:804:LEU:HB2	1.91	0.70
1:B:65:ILE:HD13	1:B:197:ALA:HB1	1.75	0.69
1:A:346:ASN:H	1:A:346:ASN:HD22	1.40	0.69
1:A:343:GLN:HE21	1:A:343:GLN:H	1.41	0.67
1:B:346:ASN:HD22	1:B:346:ASN:N	1.92	0.66
1:B:250:GLN:OE1	6:B:1003:BTB:H61	1.95	0.66
1:A:545:PHE:CZ	1:A:549:ARG:HD2	2.32	0.64
1:B:728:SER:N	1:B:729:PRO:HD3	2.13	0.63
1:A:206:THR:OG1	3:A:902:GOL:H32	2.00	0.62
1:A:114:GLN:NE2	1:A:414:ARG:HH12	1.96	0.62
1:B:572:TRP:N	1:B:600:HIS:HD2	1.88	0.61
1:B:728:SER:H	1:B:729:PRO:CD	2.14	0.61
1:B:728:SER:N	1:B:729:PRO:CD	2.64	0.61
1:A:423:TRP:H	1:A:484:ASN:ND2	1.99	0.60
1:B:784:GLN:HE21	1:B:793:GLN:HE21	1.48	0.60
1:A:232:SER:OG	1:A:287:HIS:HD2	1.84	0.60
1:B:137:HIS:HD2	1:B:307:SER:OG	1.85	0.60
1:A:346:ASN:HD22	1:A:346:ASN:N	2.01	0.58
1:B:33:GLU:HB3	1:B:536:ARG:HH21	1.68	0.58
1:B:77:LYS:H	1:B:121:ASN:ND2	2.02	0.58
1:B:696:LEU:HG	1:B:703:LEU:HD22	1.85	0.58
1:A:598:GLU:OE2	1:A:600:HIS:HE1	1.87	0.58
1:A:110:ARG:NH2	1:A:323:GLN:HG3	2.19	0.57
1:A:561:ARG:HE	1:A:664:HIS:CD2	2.19	0.57
1:B:672:TYR:CG	1:B:734:ILE:HG21	2.39	0.57
1:B:733:ASN:ND2	1:B:801:TRP:CZ2	2.73	0.57
1:A:423:TRP:H	1:A:484:ASN:HD22	1.51	0.57
1:A:456:TRP:CE2	1:A:480:PRO:HA	2.40	0.57
1:A:727:ARG:O	1:A:728:SER:HB3	2.05	0.56
1:B:727:ARG:O	1:B:728:SER:HB3	2.06	0.56
1:B:137:HIS:CD2	1:B:307:SER:OG	2.59	0.56
1:A:690:LEU:HD23	1:A:690:LEU:O	2.05	0.56
1:A:784:GLN:HE21	1:A:793:GLN:NE2	1.99	0.55
1:B:33:GLU:N	1:B:536:ARG:NH2	2.54	0.55
1:A:696:LEU:HG	1:A:703:LEU:HD22	1.89	0.55
1:A:244:LYS:HB3	1:A:245:PRO:HD3	1.89	0.55
1:A:685:PRO:HG3	1:A:748:GLN:HE22	1.72	0.54

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:94:GLN:OE1	1:B:797:HIS:HE1	1.90	0.54
1:A:575:ARG:HH22	1:A:592:GLN:HE22	1.54	0.53
1:B:346:ASN:H	1:B:346:ASN:ND2	1.97	0.53
1:B:784:GLN:NE2	1:B:793:GLN:HE21	2.07	0.53
1:A:569:ALA:HA	1:A:604:MET:CE	2.37	0.53
1:A:733:ASN:HB3	1:A:801:TRP:CG	2.43	0.53
1:A:234[B]:THR:HG21	1:B:268:PRO:HG3	1.90	0.52
1:A:233:LEU:HD22	1:B:265:VAL:HG23	1.92	0.52
1:A:343:GLN:H	1:A:343:GLN:NE2	2.06	0.52
1:B:256:ASP:O	1:B:259:VAL:HG22	2.10	0.52
1:B:44:ASN:ND2	1:B:109:ARG:HH12	2.07	0.52
1:B:430:LEU:HB2	1:B:431:PRO:HD3	1.92	0.52
1:B:782:TRP:CG	1:B:792:GLY:HA3	2.45	0.52
1:A:232:SER:OG	1:A:287:HIS:CD2	2.64	0.51
1:B:114:GLN:HE21	1:B:414:ARG:HH12	1.56	0.51
1:A:128:SER:O	1:A:143:ALA:HA	2.10	0.51
1:A:161:VAL:HG11	1:A:299:PHE:HE1	1.75	0.51
1:A:114:GLN:HE22	1:A:414:ARG:HH22	1.58	0.50
1:B:33:GLU:HB3	1:B:536:ARG:HH22	1.74	0.50
1:A:578:SER:HB2	4:A:908:SO4:O1	2.11	0.50
1:A:690:LEU:C	1:A:690:LEU:CD2	2.77	0.50
1:B:678:PHE:HZ	1:B:742:LEU:HD13	1.76	0.50
1:B:456:TRP:CE2	1:B:480:PRO:HA	2.46	0.50
1:A:114:GLN:HE21	1:A:414:ARG:HH12	1.60	0.49
1:B:128[A]:SER:O	1:B:143:ALA:HA	2.13	0.49
1:B:451:MET:CE	1:B:544:GLN:HB2	2.42	0.49
1:B:456:TRP:CD2	1:B:480:PRO:HA	2.48	0.49
1:B:451:MET:HE2	1:B:544:GLN:HB2	1.94	0.49
1:A:234[B]:THR:CG2	7:B:1101:HOH:O	2.54	0.48
1:B:54:ARG:CZ	1:B:70:MET:HE2	2.42	0.48
1:B:194:ARG:HA	1:B:199:GLY:O	2.13	0.48
1:A:672:TYR:CE1	1:A:711:SER:HA	2.49	0.48
1:B:179:SER:HB3	1:B:185:TYR:CD2	2.48	0.48
1:B:600:HIS:HA	1:B:655:ASP:OD1	2.13	0.48
1:B:595:HIS:ND1	1:B:596:PRO:HD2	2.29	0.47
1:B:385:GLY:HA2	1:B:387:TRP:CZ3	2.50	0.47
1:A:105:GLU:HG2	1:A:328:PHE:CD1	2.49	0.47
1:B:742:LEU:HB3	1:B:760:TYR:HB2	1.97	0.47
1:A:434:ASP:OD2	1:A:807:LYS:NZ	2.43	0.47
1:A:728:SER:N	1:A:729:PRO:CD	2.77	0.47
1:A:208:GLY:HA2	3:A:902:GOL:C3	2.40	0.47

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:672:TYR:CZ	1:A:711:SER:HA	2.50	0.47
1:B:726:TRP:HH2	1:B:733:ASN:HD22	1.62	0.46
1:B:799:THR:N	1:B:800:GLY:HA3	2.29	0.46
1:B:39:GLU:OE1	1:B:42:ARG:NH1	2.48	0.46
1:A:672:TYR:CG	1:A:734:ILE:HG21	2.51	0.46
1:A:234[A]:THR:HG23	1:A:284:GLY:HA2	1.97	0.46
1:A:430:LEU:HB2	1:A:431:PRO:HD3	1.96	0.46
1:A:161:VAL:HG11	1:A:299:PHE:CE1	2.51	0.46
1:B:798:PHE:C	1:B:800:GLY:HA3	2.37	0.45
1:A:614:LEU:HB3	1:A:630:TYR:CE2	2.52	0.45
1:B:103:TRP:H	1:B:357:ASN:ND2	2.15	0.44
1:B:422:LEU:HB3	1:B:484:ASN:HD22	1.82	0.44
1:A:411:ILE:HD12	1:A:418:PRO:HA	1.99	0.44
1:A:234[B]:THR:HA	1:A:285:ASN:OD1	2.18	0.44
1:A:803:SER:O	1:A:806[A]:VAL:HG23	2.17	0.44
1:B:179:SER:HB3	1:B:185:TYR:CE2	2.52	0.44
1:A:569:ALA:HA	1:A:604:MET:HE1	1.99	0.44
1:B:499:LEU:HD12	1:B:499:LEU:HA	1.72	0.44
1:B:54:ARG:NE	1:B:70:MET:HE2	2.32	0.43
1:B:489:PHE:CZ	1:B:610:MET:HE2	2.53	0.43
1:A:106:TYR:CD2	1:A:360:GLY:HA2	2.53	0.43
1:A:456:TRP:CD2	1:A:480:PRO:HA	2.53	0.43
1:B:346:ASN:N	1:B:346:ASN:ND2	2.63	0.43
1:B:485:PRO:HB3	1:B:606:TRP:CE2	2.53	0.43
1:B:727:ARG:O	1:B:728:SER:CB	2.64	0.43
1:B:354:MET:HE1	1:B:802:THR:HG22	2.01	0.43
1:B:658:ILE:HA	1:B:663:GLU:O	2.18	0.43
1:A:145:ILE:O	1:A:298:GLU:HA	2.18	0.43
1:A:103:TRP:H	1:A:357:ASN:ND2	2.17	0.43
1:A:600:HIS:HA	1:A:655:ASP:OD1	2.18	0.43
1:A:794:ARG:HD2	1:A:794:ARG:HA	1.78	0.42
1:A:214:GLN:NE2	1:A:225:PRO:HB2	2.34	0.42
1:A:430:LEU:HD13	1:A:491:VAL:HG22	2.01	0.42
1:B:355:PHE:HB2	1:B:806:VAL:HG21	2.01	0.42
1:B:441:LEU:HD12	1:B:441:LEU:HA	1.90	0.42
1:B:672:TYR:CE1	1:B:711:SER:HA	2.55	0.42
1:A:641:LEU:O	1:A:645:HIS:HB2	2.20	0.42
1:B:385:GLY:HA2	1:B:387:TRP:CH2	2.55	0.42
1:A:448:TYR:OH	1:A:544:GLN:NE2	2.50	0.42
1:A:690:LEU:HD23	1:A:694:LEU:HG	2.01	0.41
1:B:414:ARG:N	1:B:415:PRO:HD2	2.35	0.41

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:437:ILE:O	1:A:441:LEU:HG	2.20	0.41
1:A:598:GLU:OE2	1:A:600:HIS:CE1	2.70	0.41
1:B:33:GLU:CB	1:B:536:ARG:NH2	2.78	0.41
1:A:75:TRP:HA	1:A:160:ILE:O	2.21	0.41
1:A:430:LEU:CB	1:A:431:PRO:HD3	2.50	0.41
1:B:194:ARG:HB2	1:B:200:ASN:ND2	2.36	0.41
1:B:588:TYR:O	1:B:590:ARG:HG2	2.20	0.41
1:A:110:ARG:HH22	1:A:323:GLN:HG3	1.84	0.41
1:B:44:ASN:HD22	1:B:109:ARG:HH12	1.69	0.41
1:A:641:LEU:HD12	1:A:641:LEU:HA	1.94	0.41
1:B:114:GLN:HE22	1:B:414:ARG:HH22	1.68	0.41
1:B:702:GLU:OE1	1:B:713:SER:OG	2.30	0.41
1:A:234[A]:THR:CG2	1:A:284:GLY:HA2	2.51	0.41
1:A:678:PHE:O	1:A:741:GLN:HG3	2.21	0.41
1:B:733:ASN:HB3	1:B:734:ILE:H	1.40	0.41
1:B:232:SER:OG	1:B:287:HIS:CD2	2.74	0.40
1:A:575:ARG:HH22	1:A:592:GLN:NE2	2.18	0.40
1:A:110:ARG:HA	1:A:110:ARG:HD2	1.95	0.40
1:A:234[A]:THR:HA	1:A:285:ASN:OD1	2.21	0.40
1:B:620:LEU:HD12	1:B:620:LEU:HA	1.94	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	765/819 (93%)	743 (97%)	20 (3%)	2 (0%)	41	50
1	B	763/819 (93%)	742 (97%)	19 (2%)	2 (0%)	41	50
All	All	1528/1638 (93%)	1485 (97%)	39 (3%)	4 (0%)	41	50

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	728	SER
1	B	728	SER
1	B	733	ASN
1	A	659	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	633/707 (90%)	612 (97%)	21 (3%)	38	52
1	B	630/707 (89%)	600 (95%)	30 (5%)	25	35
All	All	1263/1414 (89%)	1212 (96%)	51 (4%)	30	44

All (51) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	218	ASP
1	A	343	GLN
1	A	346	ASN
1	A	417	PHE
1	A	423	TRP
1	A	424	ASP
1	A	430	LEU
1	A	438	ASP
1	A	519	LEU
1	A	526	ASN
1	A	531	LEU
1	A	575	ARG
1	A	635	ASP
1	A	641	LEU
1	A	676	PHE
1	A	682	LEU
1	A	692	LYS
1	A	693	LEU
1	A	703	LEU
1	A	725	TYR

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	794	ARG
1	B	33	GLU
1	B	44	ASN
1	B	105	GLU
1	B	173	GLU
1	B	186	GLU
1	B	194	ARG
1	B	239	VAL
1	B	244	LYS
1	B	307	SER
1	B	331	ARG
1	B	339	LYS
1	B	346	ASN
1	B	392	GLU
1	B	417	PHE
1	B	423	TRP
1	B	424	ASP
1	B	438	ASP
1	B	441	LEU
1	B	449	ASN
1	B	453	GLU
1	B	575	ARG
1	B	619	SER
1	B	641	LEU
1	B	674	SER
1	B	676	PHE
1	B	692	LYS
1	B	694	LEU
1	B	703	LEU
1	B	742	LEU
1	B	804	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (38) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	85	GLN
1	A	114	GLN
1	A	137	HIS
1	A	157	GLN
1	A	167	GLN
1	A	214	GLN
1	A	242	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	287	HIS
1	A	343	GLN
1	A	346	ASN
1	A	357	ASN
1	A	383	ASN
1	A	484	ASN
1	A	494	ASN
1	A	544	GLN
1	A	592	GLN
1	A	600	HIS
1	A	664	HIS
1	A	748	GLN
1	A	784	GLN
1	A	797	HIS
1	B	44	ASN
1	B	114	GLN
1	B	121	ASN
1	B	137	HIS
1	B	157	GLN
1	B	167	GLN
1	B	200	ASN
1	B	287	HIS
1	B	346	ASN
1	B	357	ASN
1	B	476	GLN
1	B	484	ASN
1	B	600	HIS
1	B	664	HIS
1	B	733	ASN
1	B	784	GLN
1	B	797	HIS

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)

19 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	$\#\|Z\| > 2$
4	SO4	A	906	-	4,4,4	0.38	0	6,6,6	0.10	0
3	GOL	B	1004	-	5,5,5	0.30	0	5,5,5	0.36	0
6	BTB	B	1003	-	13,13,13	1.06	1 (7%)	7,16,16	0.55	0
4	SO4	A	904	-	4,4,4	0.31	0	6,6,6	0.10	0
4	SO4	B	1009	-	4,4,4	0.22	0	6,6,6	0.12	0
4	SO4	A	908	-	4,4,4	0.43	0	6,6,6	0.12	0
4	SO4	A	903	-	4,4,4	0.36	0	6,6,6	0.04	0
4	SO4	A	907	-	4,4,4	0.37	0	6,6,6	0.07	0
4	SO4	B	1008	-	4,4,4	0.30	0	6,6,6	0.08	0
5	NAG	B	1002	1	14,14,15	0.92	0	17,19,21	2.50	7 (41%)
4	SO4	B	1007	-	4,4,4	0.55	0	6,6,6	0.13	0
4	SO4	A	905	-	4,4,4	0.39	0	6,6,6	0.20	0
4	SO4	A	909	-	4,4,4	0.28	0	6,6,6	0.06	0
4	SO4	B	1006	-	4,4,4	0.32	0	6,6,6	0.05	0
2	GII	A	901	-	29,31,31	4.18	13 (44%)	34,41,41	1.36	4 (11%)
4	SO4	A	910	-	4,4,4	0.37	0	6,6,6	0.09	0
2	GII	B	1001	-	29,31,31	4.27	12 (41%)	34,41,41	1.25	4 (11%)
3	GOL	A	902	-	5,5,5	0.16	0	5,5,5	0.47	0
4	SO4	B	1005	-	4,4,4	0.33	0	6,6,6	0.12	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	GII	B	1001	-	-	10/17/39/39	0/2/2/2
3	GOL	B	1004	-	-	3/4/4/4	-
6	BTB	B	1003	-	-	5/21/21/21	-
3	GOL	A	902	-	-	2/4/4/4	-
2	GII	A	901	-	-	1/17/39/39	0/2/2/2
5	NAG	B	1002	1	-	2/6/23/26	0/1/1/1

All (26) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	1001	GII	C5-N1	-15.71	1.19	1.47
2	A	901	GII	C5-N1	-14.51	1.21	1.47
2	A	901	GII	O5-N3	9.65	1.39	1.22
2	B	1001	GII	O5-N3	9.64	1.39	1.22
2	B	1001	GII	C11-N2	6.75	1.55	1.37
2	A	901	GII	C11-N2	6.40	1.54	1.37
2	A	901	GII	N5-N4	6.37	1.42	1.24
2	B	1001	GII	N5-N4	6.03	1.41	1.24
2	A	901	GII	C18-C1	-5.75	1.44	1.52
2	B	1001	GII	C4-C3	5.39	1.61	1.52
2	B	1001	GII	C18-C1	-4.27	1.46	1.52
2	A	901	GII	C4-C3	3.90	1.58	1.52
2	A	901	GII	O1-C1	3.01	1.50	1.43
2	A	901	GII	C6-C5	2.95	1.63	1.51
2	A	901	GII	C13-C12	2.92	1.44	1.39
2	B	1001	GII	O6-C18	2.76	1.49	1.43
2	A	901	GII	C12-N3	2.75	1.50	1.45
2	B	1001	GII	C16-C11	2.74	1.44	1.39
2	B	1001	GII	C2-C1	-2.69	1.45	1.52
2	A	901	GII	C16-C11	2.47	1.43	1.39
6	B	1003	BTB	C2-N	2.37	1.53	1.48
2	A	901	GII	C2-C1	-2.27	1.46	1.52
2	A	901	GII	C14-N4	2.19	1.48	1.43
2	B	1001	GII	C6-C5	2.08	1.60	1.51
2	B	1001	GII	C13-C12	2.07	1.43	1.39
2	B	1001	GII	C16-C15	2.01	1.42	1.38

All (15) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	B	1002	NAG	C1-O5-C5	7.90	122.89	112.19
2	A	901	GII	C4-C3-C2	-3.42	107.66	112.90

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	1001	GII	C4-C3-C2	-3.31	107.83	112.90
2	A	901	GII	C14-N4-N5	3.11	122.19	116.02
5	B	1002	NAG	C6-C5-C4	-2.90	106.21	113.00
2	B	1001	GII	C14-N4-N5	2.72	121.42	116.02
5	B	1002	NAG	O3-C3-C2	-2.67	103.93	109.47
2	B	1001	GII	O1-C1-C2	-2.63	104.27	110.35
2	A	901	GII	C17-C18-C1	-2.55	107.33	110.24
2	A	901	GII	C13-C12-C11	-2.55	119.22	121.53
5	B	1002	NAG	C4-C3-C2	2.33	114.43	111.02
5	B	1002	NAG	O5-C5-C4	2.19	116.16	110.83
2	B	1001	GII	O6-C18-C1	2.18	114.51	110.14
5	B	1002	NAG	C2-N2-C7	2.14	125.95	122.90
5	B	1002	NAG	O5-C1-C2	2.06	114.54	111.29

There are no chirality outliers.

All (23) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	B	1001	GII	C9-C10-N2-C11
3	A	902	GOL	C1-C2-C3-O3
3	B	1004	GOL	C1-C2-C3-O3
6	B	1003	BTB	O1-C1-C2-C3
6	B	1003	BTB	O1-C1-C2-C4
2	B	1001	GII	C16-C11-N2-C10
2	B	1001	GII	C12-C11-N2-C10
5	B	1002	NAG	C8-C7-N2-C2
5	B	1002	NAG	O7-C7-N2-C2
3	A	902	GOL	O2-C2-C3-O3
3	B	1004	GOL	O2-C2-C3-O3
2	B	1001	GII	C5-C6-C7-C8
2	B	1001	GII	C7-C8-C9-C10
6	B	1003	BTB	N-C5-C6-O6
2	B	1001	GII	C14-N4-N5-N6
6	B	1003	BTB	O1-C1-C2-N
2	B	1001	GII	C6-C7-C8-C9
2	A	901	GII	C7-C8-C9-C10
2	B	1001	GII	C13-C14-N4-N5
2	B	1001	GII	C15-C14-N4-N5
2	B	1001	GII	N2-C10-C9-C8
3	B	1004	GOL	O1-C1-C2-C3
6	B	1003	BTB	C3-C2-C4-O4

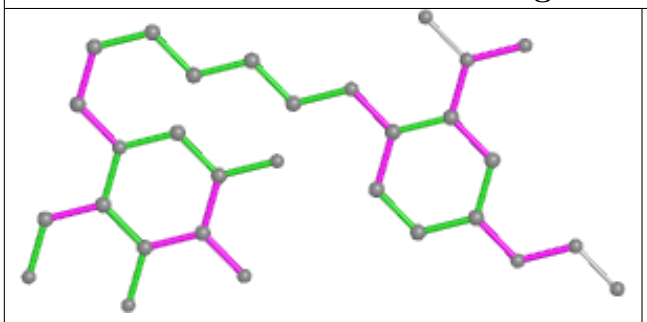
There are no ring outliers.

3 monomers are involved in 5 short contacts:

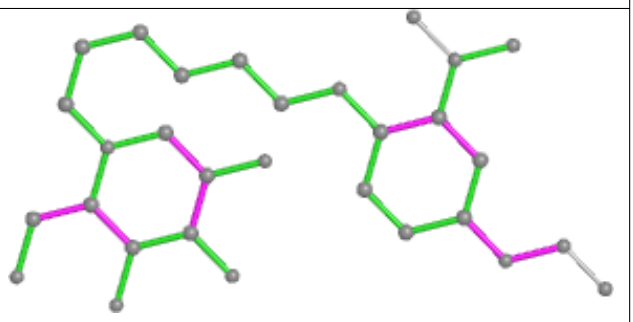
Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	B	1003	BTB	1	0
4	A	908	SO4	1	0
3	A	902	GOL	3	0

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.

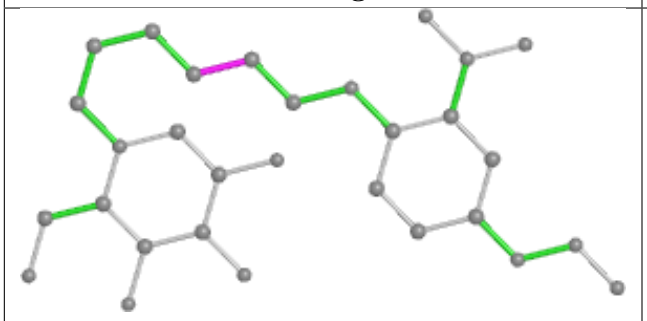
Ligand GII A 901



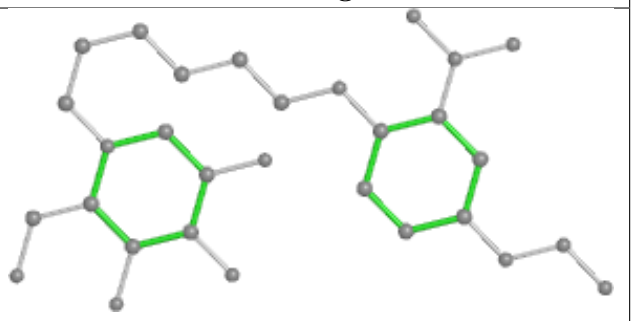
Bond lengths



Bond angles

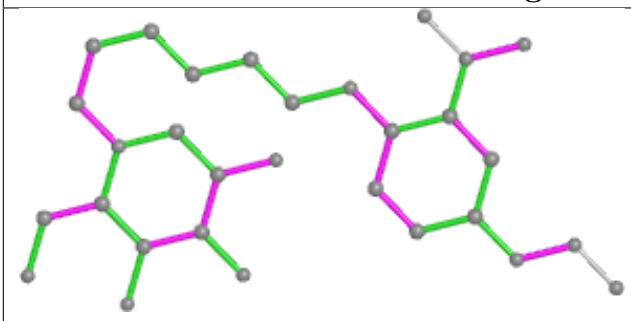


Torsions

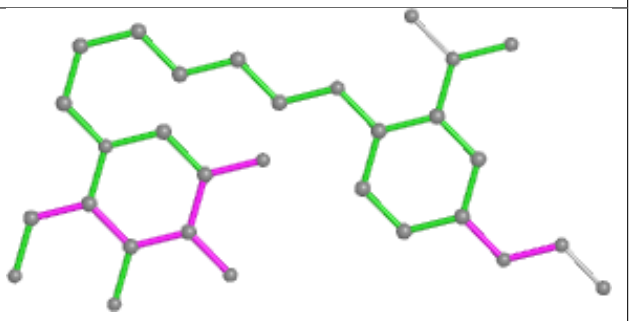


Rings

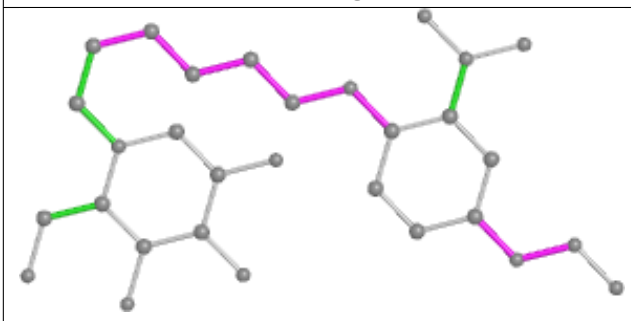
Ligand GII B 1001



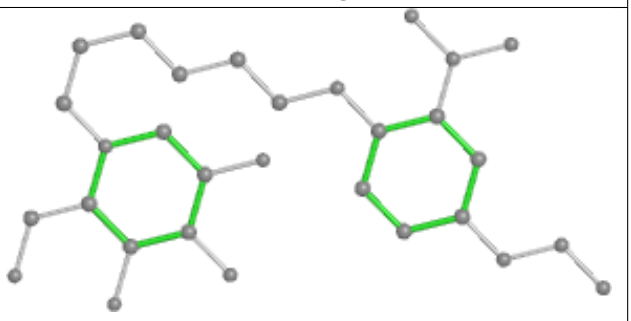
Bond lengths



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, 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%)	-0.16	14 (1%) 68 75	26, 39, 64, 86	0
1	B	764/819 (93%)	-0.28	3 (0%) 92 95	25, 39, 59, 87	0
All	All	1529/1638 (93%)	-0.22	17 (1%) 80 85	25, 39, 62, 87	0

All (17) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	621	LEU	4.0
1	A	36	LEU	3.8
1	A	531	LEU	3.6
1	A	37	HIS	3.1
1	A	560	ASP	2.9
1	A	499	LEU	2.7
1	A	535	ARG	2.4
1	B	171	ASN	2.4
1	A	383	ASN	2.4
1	A	661	PHE	2.4
1	A	619	SER	2.2
1	A	151	ASP	2.2
1	A	623	ALA	2.1
1	A	618	GLY	2.0
1	B	790	GLY	2.0
1	A	527	PRO	2.0
1	B	720	GLY	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)

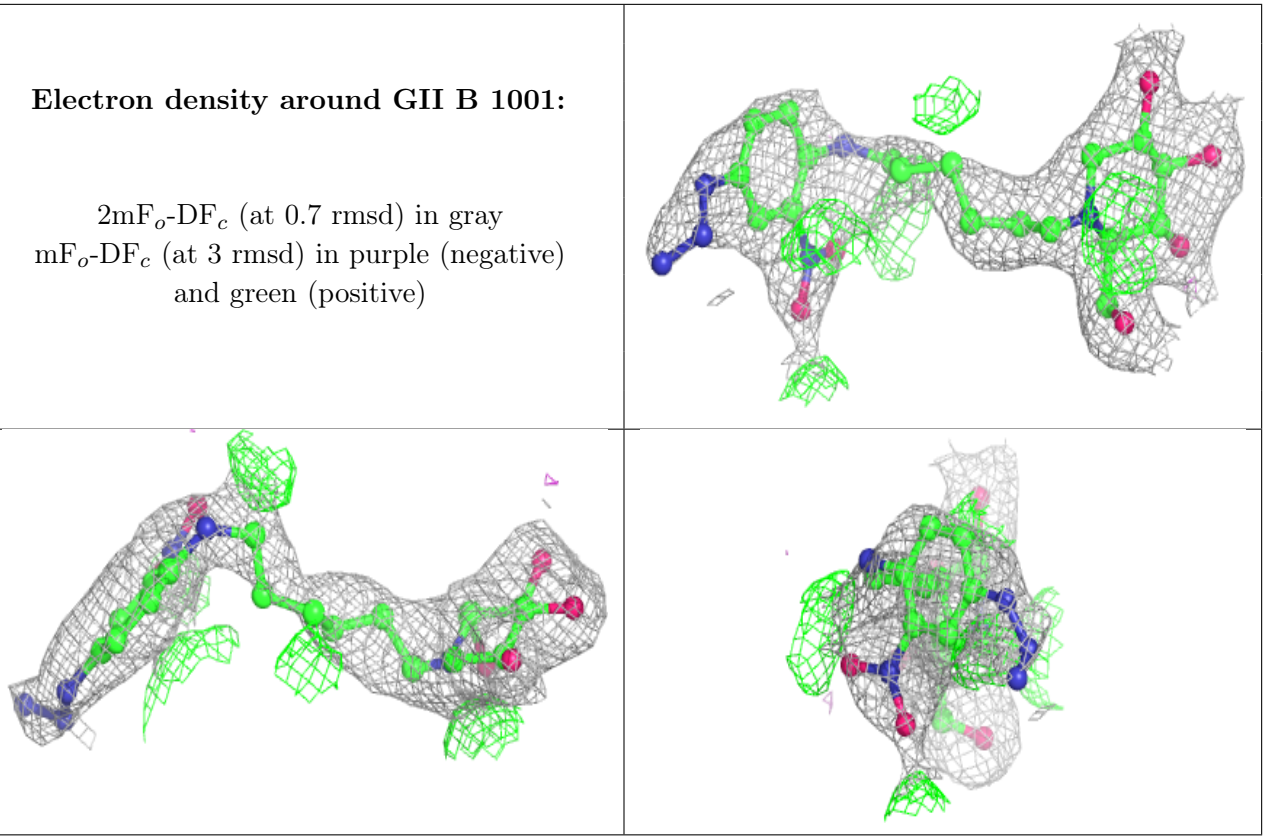
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' 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
5	NAG	B	1002	14/15	0.84	0.25	48,69,78,81	0
4	SO4	A	909	5/5	0.86	0.24	85,89,95,100	0
6	BTB	B	1003	14/14	0.86	0.17	53,68,72,74	0
3	GOL	A	902	6/6	0.87	0.20	52,60,66,69	0
4	SO4	B	1006	5/5	0.87	0.29	81,91,110,112	0
4	SO4	A	903	5/5	0.93	0.35	94,99,102,110	0
4	SO4	A	907	5/5	0.93	0.19	101,112,120,122	0
3	GOL	B	1004	6/6	0.93	0.17	34,43,46,49	0
2	GII	B	1001	30/30	0.94	0.18	26,75,98,99	0
4	SO4	A	904	5/5	0.95	0.12	71,75,83,98	0
4	SO4	A	908	5/5	0.95	0.15	84,84,87,90	0
4	SO4	A	910	5/5	0.96	0.25	93,97,100,104	0
2	GII	A	901	30/30	0.97	0.20	25,50,88,108	0
4	SO4	B	1009	5/5	0.97	0.18	46,63,66,67	0
4	SO4	B	1005	5/5	0.98	0.15	53,67,69,70	0
4	SO4	A	906	5/5	0.98	0.13	54,55,57,73	0
4	SO4	B	1008	5/5	0.98	0.12	62,62,70,77	0
4	SO4	A	905	5/5	0.99	0.16	36,40,41,42	0
4	SO4	B	1007	5/5	0.99	0.12	38,38,42,50	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.

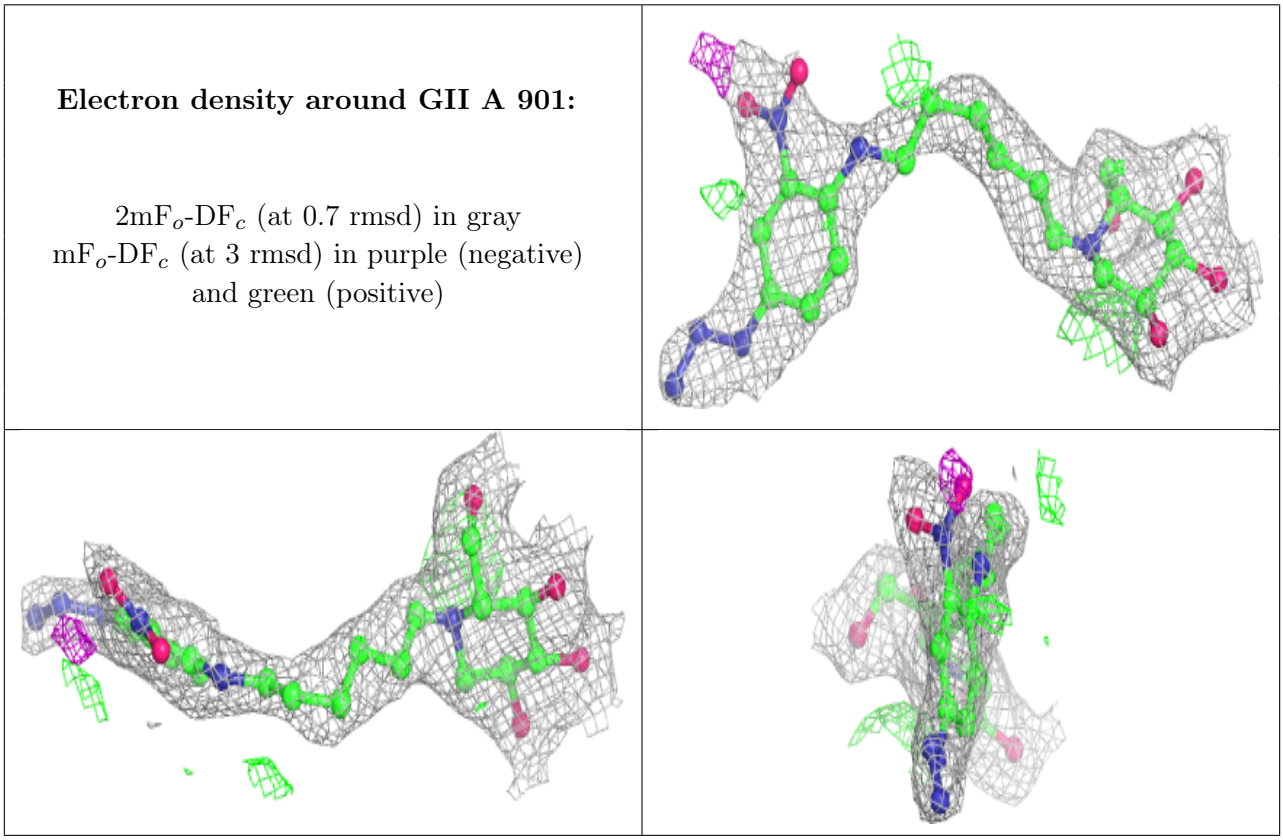
Electron density around GII B 1001:

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



Electron density around GII 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.