Full wwPDB X-ray Structure Validation Report (i)

Jul 18, 2022 – 10:07 am BST

PDB ID	:	7Z28
Title	:	High-resolution crystal structure of ERAP1 with bound bestatin analogue inhibitor
Authors	:	Giastas, P.; Papakyriakou, A.; Stratikos, E.; Vourloumis, D.
Deposited on	:	2022-02-26
Resolution	:	1.55 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:

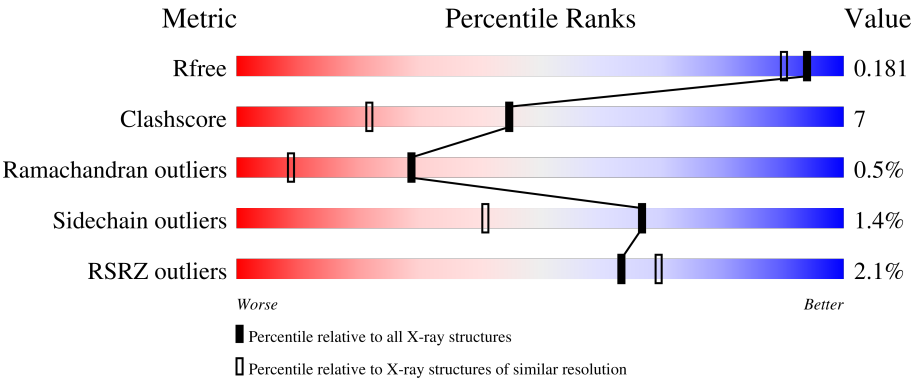
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.29
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0267
CCP4	:	7.1.010 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.29

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 1.55 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	1483 (1.56-1.56)
Clashscore	141614	1529 (1.56-1.56)
Ramachandran outliers	138981	1498 (1.56-1.56)
Sidechain outliers	138945	1495 (1.56-1.56)
RSRZ outliers	127900	1465 (1.56-1.56)

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	859	2%, 88%, 11%
2	B	2	100%

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	NAG	A	1001	-	-	-	X

2 Entry composition (i)

There are 9 unique types of molecules in this entry. The entry contains 8168 atoms, of which 126 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 Endoplasmic reticulum aminopeptidase 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	859	Total	C	N	O	S	0	21	0
			6993	4511	1149	1295	38			

There are 35 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	?	-	CYS	deletion	UNP Q9NZ08
A	?	-	PRO	deletion	UNP Q9NZ08
A	?	-	THR	deletion	UNP Q9NZ08
A	?	-	ASP	deletion	UNP Q9NZ08
A	?	-	GLY	deletion	UNP Q9NZ08
A	?	-	VAL	deletion	UNP Q9NZ08
A	?	-	LYS	deletion	UNP Q9NZ08
A	?	-	GLY	deletion	UNP Q9NZ08
A	?	-	MET	deletion	UNP Q9NZ08
A	?	-	ASP	deletion	UNP Q9NZ08
A	?	-	GLY	deletion	UNP Q9NZ08
A	?	-	PHE	deletion	UNP Q9NZ08
A	?	-	CYS	deletion	UNP Q9NZ08
A	?	-	SER	deletion	UNP Q9NZ08
A	?	-	ARG	deletion	UNP Q9NZ08
A	?	-	SER	deletion	UNP Q9NZ08
A	?	-	GLN	deletion	UNP Q9NZ08
A	?	-	HIS	deletion	UNP Q9NZ08
A	?	-	SER	deletion	UNP Q9NZ08
A	?	-	SER	deletion	UNP Q9NZ08
A	?	-	SER	deletion	UNP Q9NZ08
A	?	-	SER	deletion	UNP Q9NZ08
A	?	-	SER	deletion	UNP Q9NZ08
A	?	-	HIS	deletion	UNP Q9NZ08
A	?	-	TRP	deletion	UNP Q9NZ08
A	?	-	HIS	deletion	UNP Q9NZ08
A	?	-	GLN	deletion	UNP Q9NZ08

Continued on next page...

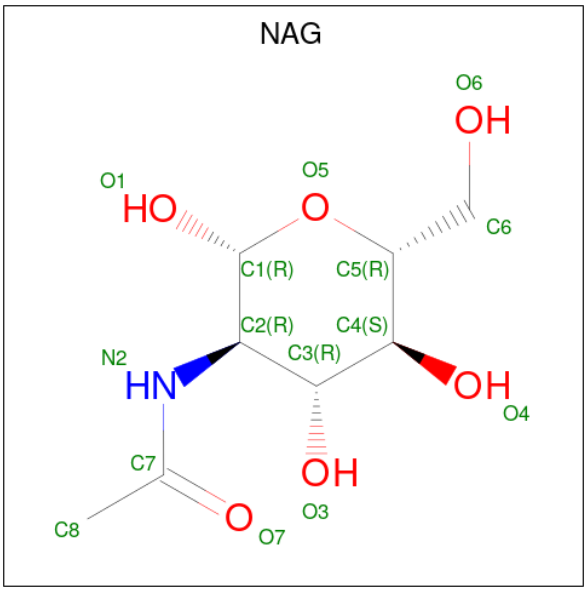
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	?	-	GLU	deletion	UNP Q9NZ08
A	487	SER	GLY	conflict	UNP Q9NZ08
A	515	GLY	VAL	conflict	UNP Q9NZ08
A	?	-	SER	deletion	UNP Q9NZ08
A	?	-	ASP	deletion	UNP Q9NZ08
A	?	-	GLY	deletion	UNP Q9NZ08
A	?	-	ALA	deletion	UNP Q9NZ08
A	?	-	PRO	deletion	UNP Q9NZ08

- Molecule 2 is an oligosaccharide called 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf	Trace
2	B	2	Total C N O 28 16 2 10	0	0	0

- Molecule 3 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).

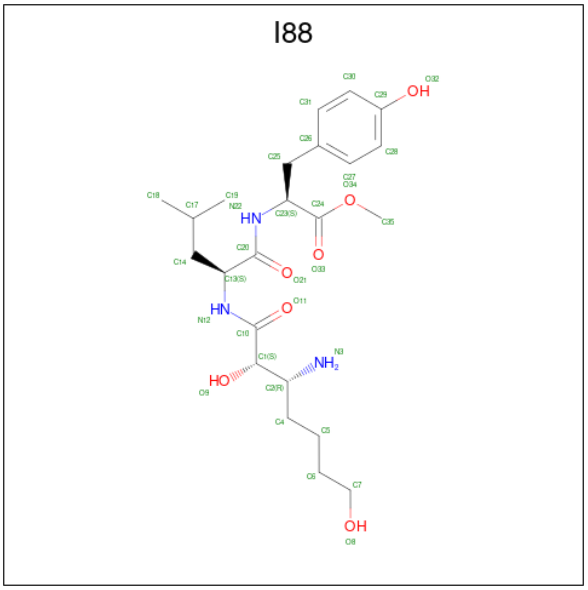


Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C N O 14 8 1 5	0	0

- Molecule 4 is ZINC ION (three-letter code: ZN) (formula: Zn).

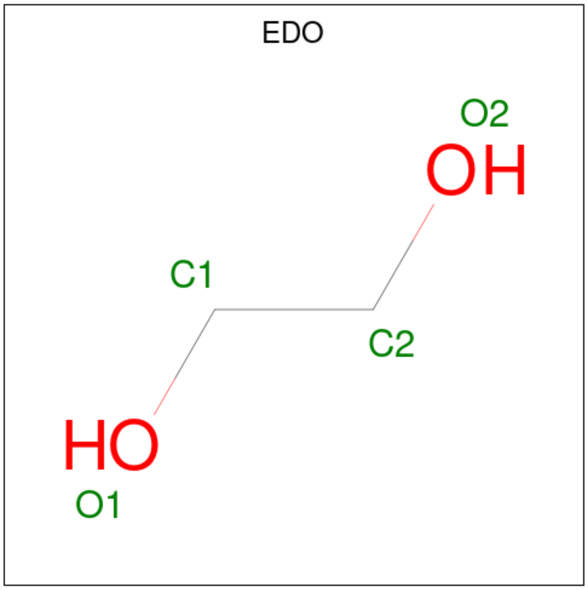
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total Zn 1 1	0	0

- Molecule 5 is methyl (2 {S})-2-[[(2 {S})-2-[[(2 {S},3 {R})-3-azanyl-2,7-bis(oxidanyl)heptanoyl]amino]-4-methyl-pentanoyl]amino]-3-(4-hydroxyphenyl)propanoate (three-letter code: I88) (formula: $C_{23}H_{37}N_3O_7$) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total C N O 33 23 3 7	0	0

- Molecule 6 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: $C_2H_6O_2$).



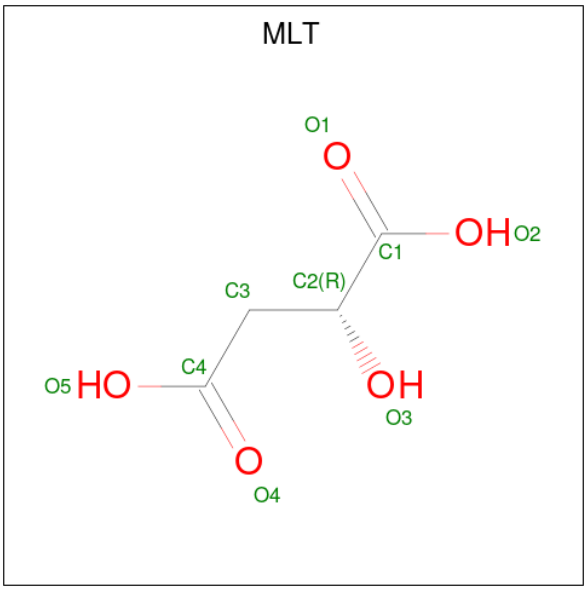
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
6	A	1	Total 10	C 2	H 6	O 2	0	0
6	A	1	Total 10	C 2	H 6	O 2	0	0
6	A	1	Total 10	C 2	H 6	O 2	0	0
6	A	1	Total 10	C 2	H 6	O 2	0	0
6	A	1	Total 10	C 2	H 6	O 2	0	0
6	A	1	Total 10	C 2	H 6	O 2	0	0
6	A	1	Total 10	C 2	H 6	O 2	0	0
6	A	1	Total 10	C 2	H 6	O 2	0	0
6	A	1	Total 10	C 2	H 6	O 2	0	0
6	A	1	Total 10	C 2	H 6	O 2	0	0
6	A	1	Total 10	C 2	H 6	O 2	0	0
6	A	1	Total 10	C 2	H 6	O 2	0	0
6	A	1	Total 10	C 2	H 6	O 2	0	0
6	A	1	Total 10	C 2	H 6	O 2	0	0

Continued on next page...

Continued from previous page...

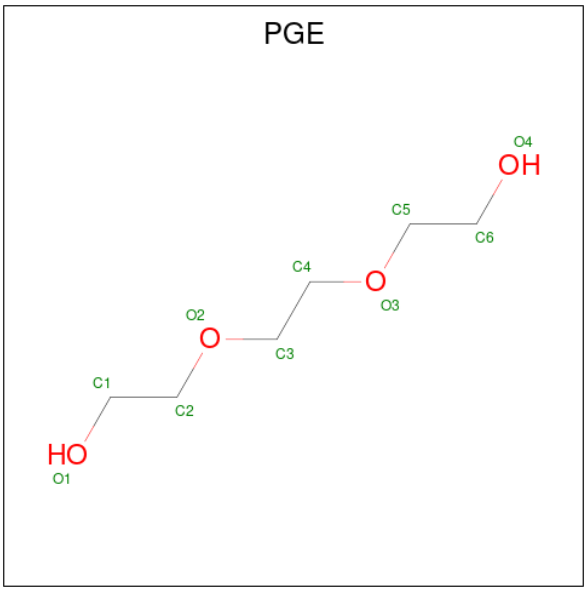
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	1	Total C H O 10 2 6 2	0	0
6	A	1	Total C H O 10 2 6 2	0	0
6	A	1	Total C H O 10 2 6 2	0	0
6	A	1	Total C H O 10 2 6 2	0	0

- Molecule 7 is D-MALATE (three-letter code: MLT) (formula: $C_4H_6O_5$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	A	1	Total C H O 13 4 4 5	0	0

- Molecule 8 is TRIETHYLENE GLYCOL (three-letter code: PGE) (formula: $C_6H_{14}O_4$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
8	A	1	Total	C	H	O	0	0
			24	6	14	4		

- Molecule 9 is water.

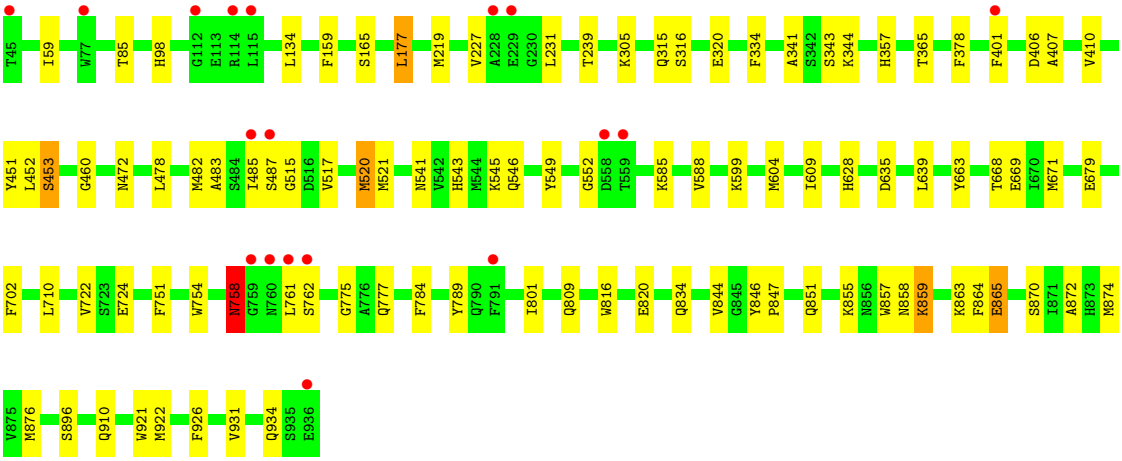
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	A	882	Total	O	0	0
			882	882		

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.

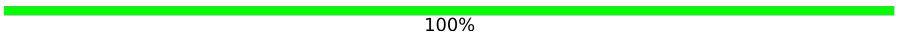
• Molecule 1: Endoplasmic reticulum aminopeptidase 1

Chain A:



• Molecule 2: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain B:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 2 21 21	Depositor
Cell constants a, b, c, α, β, γ	57.65Å 113.17Å 140.16Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	88.05 – 1.55 88.05 – 1.55	Depositor EDS
% Data completeness (in resolution range)	99.6 (88.05-1.55) 99.6 (88.05-1.55)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.56 (at 1.55Å)	Xtriage
Refinement program	PHENIX 1.19.2_4158	Depositor
R, R_{free}	0.165 , 0.182 0.163 , 0.181	Depositor DCC
R_{free} test set	6621 reflections (4.98%)	wwPDB-VP
Wilson B-factor (Å^2)	18.5	Xtriage
Anisotropy	0.248	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	(Not available) , (Not available)	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	8168	wwPDB-VP
Average B, all atoms (Å^2)	24.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.00% 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: MLT, PGE, I88, EDO, NAG, ZN

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	0.39	0/7228	0.60	0/9795

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	6993	0	6951	90	0
2	B	28	0	25	0	0
3	A	14	0	13	0	0
4	A	1	0	0	0	0
5	A	33	0	0	2	0
6	A	72	108	108	12	0
7	A	9	4	4	0	0
8	A	10	14	14	1	0
9	A	882	0	0	25	0
All	All	8042	126	7115	93	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 (93) 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:859:LYS:HE3	9:A:1128:HOH:O	1.68	0.92
1:A:460:GLY:HA2	1:A:485:ILE:HD11	1.52	0.90
1:A:859:LYS:NZ	9:A:1102:HOH:O	1.92	0.89
1:A:406:ASP:OD1	9:A:1101:HOH:O	1.91	0.89
1:A:541[A]:ASN:OD1	9:A:1103:HOH:O	1.94	0.86
1:A:487:SER:C	1:A:515:GLY:HA3	1.96	0.85
1:A:334:PHE:CZ	1:A:344:LYS:HE3	2.12	0.83
1:A:858:ASN:N	9:A:1102:HOH:O	1.96	0.79
1:A:517:VAL:HG13	1:A:521:MET:HG2	1.71	0.73
1:A:85[B]:THR:HG22	9:A:1113:HOH:O	1.88	0.72
1:A:460:GLY:CA	1:A:485:ILE:HD11	2.19	0.72
1:A:401:PHE:CE1	1:A:671:MET:HE2	2.25	0.71
1:A:59:ILE:HD12	9:A:1564:HOH:O	1.92	0.69
1:A:227:VAL:HG23	1:A:231:LEU:HD23	1.79	0.65
1:A:98:HIS:HD2	9:A:1270:HOH:O	1.80	0.64
1:A:671:MET:HE1	9:A:1812:HOH:O	1.97	0.64
1:A:478:LEU:O	1:A:482:MET:HG2	1.96	0.63
1:A:859:LYS:H	1:A:859:LYS:CE	2.12	0.63
1:A:546:GLN:O	1:A:585:LYS:HD2	1.99	0.62
1:A:668:THR:HG22	6:A:1020:EDO:H22	1.82	0.61
1:A:754:TRP:HB2	1:A:761:LEU:HD23	1.81	0.61
1:A:316[A]:SER:HB3	5:A:1003:I88:O21	2.00	0.61
1:A:401:PHE:CZ	1:A:671:MET:HE2	2.37	0.59
1:A:549:TYR:CD2	1:A:609:ILE:HD11	2.38	0.59
1:A:922[B]:MET:HE3	1:A:922[B]:MET:HA	1.85	0.57
1:A:777:GLN:HG3	9:A:1789:HOH:O	2.03	0.57
1:A:316[B]:SER:HB2	5:A:1003:I88:O21	2.05	0.57
1:A:588:VAL:O	6:A:1012:EDO:H21	2.05	0.56
1:A:870:SER:O	1:A:874:MET:HG3	2.05	0.56
6:A:1012:EDO:H22	9:A:1332:HOH:O	2.04	0.56
1:A:485:ILE:N	1:A:485:ILE:HD12	2.21	0.56
1:A:451:TYR:CE2	1:A:520:MET:HE1	2.41	0.56
1:A:482:MET:O	1:A:485:ILE:HD13	2.06	0.55
1:A:85[B]:THR:HG23	9:A:1564:HOH:O	2.07	0.55
1:A:628:HIS:HD2	9:A:1623:HOH:O	1.89	0.55
1:A:669:GLU:HA	6:A:1009:EDO:H11	1.89	0.54
1:A:851:GLN:HG2	1:A:855:LYS:HD2	1.89	0.54
1:A:401:PHE:CZ	1:A:671:MET:CE	2.91	0.54
1:A:517:VAL:HG13	1:A:521:MET:CG	2.38	0.54
1:A:896:SER:HB3	6:A:1014:EDO:H12	1.88	0.54

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:754:TRP:CA	1:A:761:LEU:HD23	2.39	0.52
1:A:846:TYR:CG	1:A:847:PRO:HD3	2.45	0.52
1:A:710:LEU:HD13	6:A:1020:EDO:H21	1.93	0.51
1:A:922[B]:MET:CE	1:A:926:PHE:HB2	2.40	0.51
1:A:483:ALA:HB2	9:A:1513:HOH:O	2.09	0.51
1:A:872:ALA:O	1:A:876[B]:MET:HG2	2.11	0.51
1:A:543[A]:HIS:ND1	9:A:1104:HOH:O	2.04	0.50
1:A:859:LYS:H	1:A:859:LYS:HE2	1.77	0.50
1:A:931:VAL:O	1:A:934:GLN:HG2	2.12	0.49
1:A:320:GLU:HG3	1:A:357:HIS:HB3	1.94	0.49
6:A:1010:EDO:O2	9:A:1105:HOH:O	2.17	0.49
1:A:365:THR:O	1:A:472:ASN:HA	2.13	0.49
1:A:834:GLN:HG2	9:A:1112:HOH:O	2.13	0.49
1:A:863:LYS:HZ3	6:A:1015:EDO:C2	2.26	0.49
1:A:227:VAL:CG2	1:A:231:LEU:HD23	2.42	0.48
1:A:816:TRP:CE2	1:A:820:GLU:HG3	2.48	0.48
1:A:517:VAL:CG1	1:A:517:VAL:O	2.62	0.47
1:A:754:TRP:CB	1:A:761:LEU:HD23	2.44	0.47
1:A:857:TRP:N	9:A:1102:HOH:O	2.47	0.47
1:A:859:LYS:N	1:A:859:LYS:HD3	2.30	0.47
1:A:758:ASN:O	1:A:784:PHE:HZ	1.96	0.47
6:A:1019:EDO:H11	9:A:1129:HOH:O	2.15	0.47
1:A:344:LYS:NZ	9:A:1122:HOH:O	2.47	0.46
1:A:599[B]:LYS:HE2	1:A:635:ASP:OD1	2.16	0.46
1:A:165:SER:HB3	1:A:177:LEU:HG	1.97	0.46
1:A:341:ALA:HB1	1:A:722:VAL:HG13	1.97	0.45
1:A:859:LYS:N	1:A:859:LYS:CD	2.79	0.45
1:A:922[B]:MET:HE3	1:A:922[B]:MET:CA	2.44	0.45
1:A:552:GLY:HA3	9:A:1321:HOH:O	2.17	0.45
1:A:517:VAL:CG1	1:A:521:MET:CG	2.94	0.45
1:A:679:GLU:OE1	8:A:1023:PGE:H32	2.18	0.44
1:A:341:ALA:HB3	6:A:1007:EDO:H12	1.98	0.44
1:A:754:TRP:HA	1:A:761:LEU:HB3	1.99	0.44
1:A:921:TRP:CD2	6:A:1006:EDO:H22	2.53	0.43
1:A:761:LEU:O	1:A:761:LEU:HD12	2.17	0.43
1:A:599[B]:LYS:NZ	1:A:639:LEU:HG	2.34	0.43
1:A:159:PHE:CB	1:A:315:GLN:HB2	2.49	0.43
1:A:517:VAL:CG1	1:A:521:MET:HG2	2.45	0.43
1:A:219:MET:HG2	1:A:239:THR:HG22	2.02	0.42
1:A:858:ASN:HB2	1:A:859:LYS:NZ	2.33	0.42
1:A:545:LYS:HD2	9:A:1915:HOH:O	2.19	0.42

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:628:HIS:HE1	9:A:1340:HOH:O	2.02	0.42
1:A:751:PHE:CD1	1:A:775:GLY:HA3	2.54	0.42
1:A:809:GLN:HA	1:A:844:VAL:HG21	2.01	0.42
1:A:485:ILE:HD12	1:A:485:ILE:H	1.84	0.42
1:A:789:TYR:HB2	1:A:801:ILE:HG21	2.02	0.42
1:A:846:TYR:CD2	1:A:847:PRO:HD3	2.55	0.42
1:A:452:LEU:O	1:A:453:SER:CB	2.68	0.41
1:A:910:GLN:HG2	9:A:1457:HOH:O	2.20	0.41
1:A:724:GLU:OE2	9:A:1106:HOH:O	2.21	0.41
1:A:724:GLU:HG3	6:A:1009:EDO:O2	2.21	0.41
1:A:864:PHE:O	1:A:865:GLU:HB2	2.21	0.40
1:A:407:ALA:O	1:A:410:VAL:HG12	2.21	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	873/859 (102%)	845 (97%)	24 (3%)	4 (0%)	29 9

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	758	ASN
1	A	453	SER
1	A	762	SER
1	A	865	GLU

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	780/766 (102%)	769 (99%)	11 (1%)	67 41

All (11) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	134	LEU
1	A	177	LEU
1	A	305	LYS
1	A	343	SER
1	A	378	PHE
1	A	520	MET
1	A	604	MET
1	A	663	TYR
1	A	702	PHE
1	A	758	ASN
1	A	859	LYS

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

Mol	Chain	Res	Type
1	A	98	HIS
1	A	446	ASN
1	A	628	HIS
1	A	645	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)

2 monosaccharides 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	NAG	B	1	2,1	14,14,15	0.34	0	17,19,21	0.37	0
2	NAG	B	2	2	14,14,15	0.36	0	17,19,21	0.48	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	NAG	B	1	2,1	-	0/6/23/26	0/1/1/1
2	NAG	B	2	2	-	2/6/23/26	0/1/1/1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (2) torsion outliers are listed below:

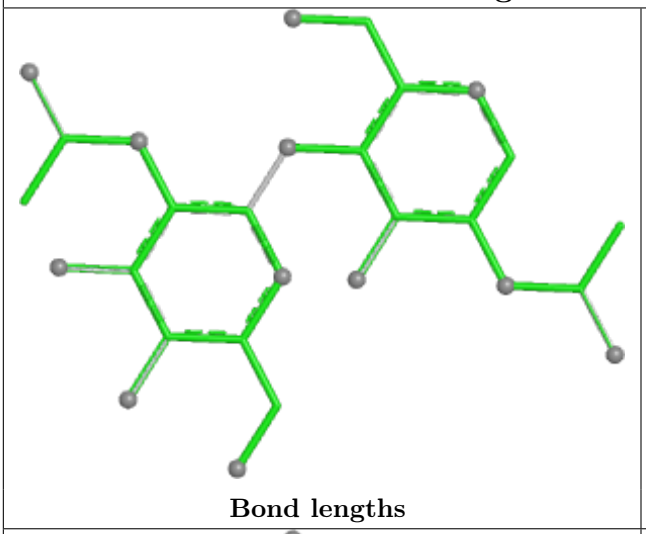
Mol	Chain	Res	Type	Atoms
2	B	2	NAG	O5-C5-C6-O6
2	B	2	NAG	C4-C5-C6-O6

There are no ring outliers.

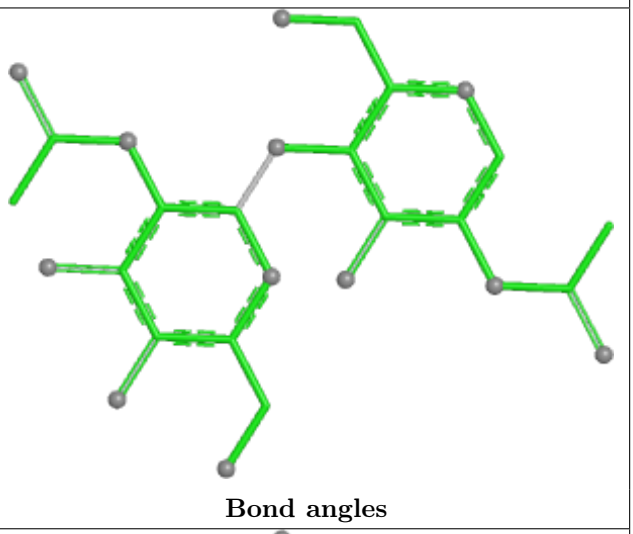
No monomer is involved in short contacts.

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for oligosaccharide.

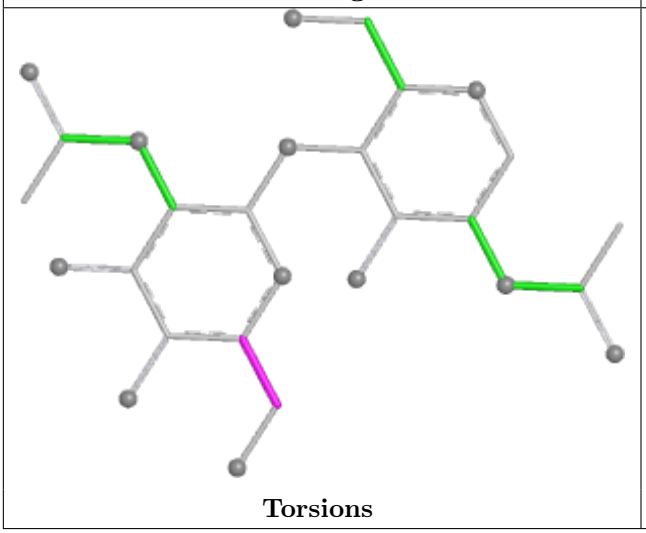
Oligosaccharide Chain B



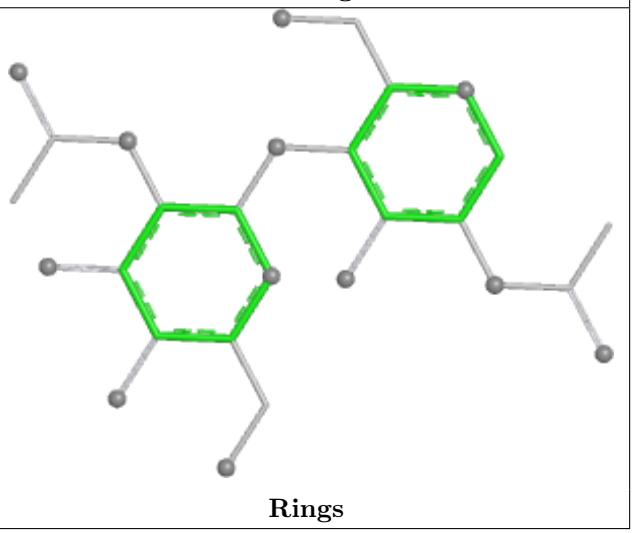
Bond lengths



Bond angles



Torsions



Rings

5.6 Ligand geometry

Of 23 ligands modelled in this entry, 1 is monoatomic - leaving 22 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
6	EDO	A	1012	-	3,3,3	0.44	0	2,2,2	0.21	0
6	EDO	A	1015	-	3,3,3	0.51	0	2,2,2	0.47	0
6	EDO	A	1008	-	3,3,3	0.48	0	2,2,2	0.08	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\|>2$	Counts	RMSZ	$\#\|Z\|>2$
3	NAG	A	1001	1	14,14,15	0.64	0	17,19,21	0.58	0
6	EDO	A	1004	-	3,3,3	0.48	0	2,2,2	0.56	0
6	EDO	A	1013	-	3,3,3	0.56	0	2,2,2	0.25	0
6	EDO	A	1009	-	3,3,3	0.47	0	2,2,2	0.39	0
6	EDO	A	1010	-	3,3,3	0.51	0	2,2,2	0.27	0
6	EDO	A	1006	-	3,3,3	0.49	0	2,2,2	0.34	0
6	EDO	A	1020	-	3,3,3	0.49	0	2,2,2	0.34	0
5	I88	A	1003	4	33,33,33	1.03	1 (3%)	41,43,43	1.06	3 (7%)
6	EDO	A	1017	-	3,3,3	0.47	0	2,2,2	0.36	0
6	EDO	A	1019	-	3,3,3	0.42	0	2,2,2	0.20	0
6	EDO	A	1021	-	3,3,3	0.47	0	2,2,2	0.38	0
6	EDO	A	1014	-	3,3,3	0.51	0	2,2,2	0.30	0
7	MLT	A	1022	-	8,8,8	1.06	0	10,10,10	1.87	4 (40%)
8	PGE	A	1023	-	9,9,9	0.29	0	8,8,8	0.63	0
6	EDO	A	1005	-	3,3,3	0.54	0	2,2,2	0.45	0
6	EDO	A	1007	-	3,3,3	0.42	0	2,2,2	0.42	0
6	EDO	A	1011	-	3,3,3	0.49	0	2,2,2	0.13	0
6	EDO	A	1016	-	3,3,3	0.46	0	2,2,2	0.26	0
6	EDO	A	1018	-	3,3,3	0.53	0	2,2,2	0.33	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
6	EDO	A	1012	-	-	1/1/1/1	-
6	EDO	A	1015	-	-	0/1/1/1	-
6	EDO	A	1008	-	-	0/1/1/1	-
3	NAG	A	1001	1	-	4/6/23/26	0/1/1/1
6	EDO	A	1004	-	-	0/1/1/1	-
6	EDO	A	1013	-	-	0/1/1/1	-
6	EDO	A	1009	-	-	0/1/1/1	-
6	EDO	A	1010	-	-	1/1/1/1	-
6	EDO	A	1006	-	-	0/1/1/1	-
6	EDO	A	1020	-	-	1/1/1/1	-
5	I88	A	1003	4	-	6/39/39/39	0/1/1/1
6	EDO	A	1017	-	-	1/1/1/1	-
6	EDO	A	1019	-	-	0/1/1/1	-
6	EDO	A	1021	-	-	1/1/1/1	-
6	EDO	A	1014	-	-	1/1/1/1	-
7	MLT	A	1022	-	-	4/8/8/8	-

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
8	PGE	A	1023	-	-	3/7/7/7	-
6	EDO	A	1005	-	-	0/1/1/1	-
6	EDO	A	1007	-	-	0/1/1/1	-
6	EDO	A	1011	-	-	0/1/1/1	-
6	EDO	A	1016	-	-	1/1/1/1	-
6	EDO	A	1018	-	-	1/1/1/1	-

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	A	1003	I88	O34-C24	5.02	1.45	1.33

All (7) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
7	A	1022	MLT	O2-C1-C2	3.66	120.76	112.72
5	A	1003	I88	O34-C24-C23	3.39	120.19	111.52
5	A	1003	I88	O9-C1-C10	-3.18	103.78	110.63
5	A	1003	I88	O34-C24-O33	-2.51	118.93	123.84
7	A	1022	MLT	C2-C3-C4	-2.35	106.31	112.13
7	A	1022	MLT	O1-C1-C2	-2.31	118.02	122.54
7	A	1022	MLT	O5-C4-C3	2.07	120.70	114.07

There are no chirality outliers.

All (25) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	A	1003	I88	C10-C1-C2-C4
5	A	1003	I88	C10-C1-C2-N3
5	A	1003	I88	O9-C1-C2-N3
7	A	1022	MLT	O1-C1-C2-O3
7	A	1022	MLT	O1-C1-C2-C3
7	A	1022	MLT	O2-C1-C2-O3
7	A	1022	MLT	O2-C1-C2-C3
3	A	1001	NAG	O5-C5-C6-O6
3	A	1001	NAG	C4-C5-C6-O6
3	A	1001	NAG	C8-C7-N2-C2
3	A	1001	NAG	O7-C7-N2-C2
6	A	1010	EDO	O1-C1-C2-O2
6	A	1012	EDO	O1-C1-C2-O2
5	A	1003	I88	O9-C1-C10-N12

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
8	A	1023	PGE	C6-C5-O3-C4
8	A	1023	PGE	O3-C5-C6-O4
5	A	1003	I88	O9-C1-C10-O11
5	A	1003	I88	O9-C1-C2-C4
6	A	1016	EDO	O1-C1-C2-O2
6	A	1017	EDO	O1-C1-C2-O2
6	A	1018	EDO	O1-C1-C2-O2
6	A	1020	EDO	O1-C1-C2-O2
6	A	1021	EDO	O1-C1-C2-O2
8	A	1023	PGE	O2-C3-C4-O3
6	A	1014	EDO	O1-C1-C2-O2

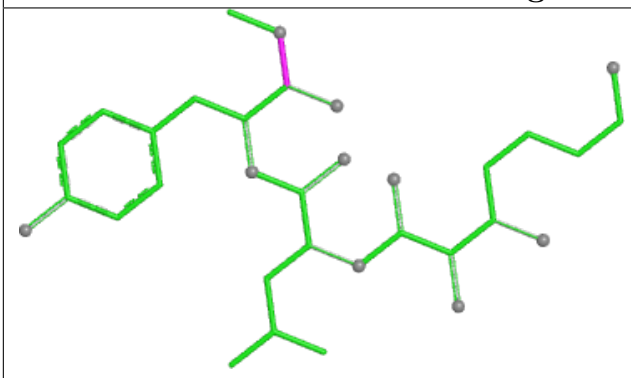
There are no ring outliers.

11 monomers are involved in 15 short contacts:

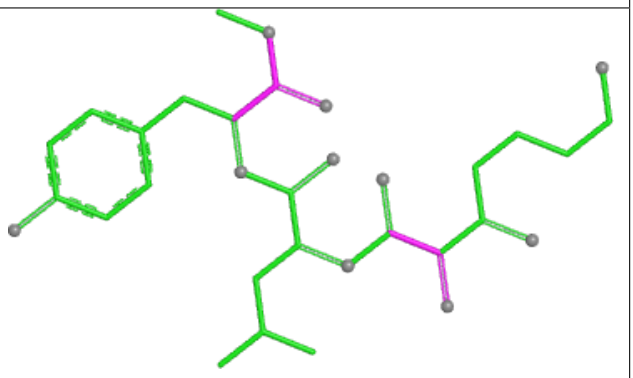
Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	A	1012	EDO	2	0
6	A	1015	EDO	1	0
6	A	1009	EDO	2	0
6	A	1010	EDO	1	0
6	A	1006	EDO	1	0
6	A	1020	EDO	2	0
5	A	1003	I88	2	0
6	A	1019	EDO	1	0
6	A	1014	EDO	1	0
8	A	1023	PGE	1	0
6	A	1007	EDO	1	0

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.

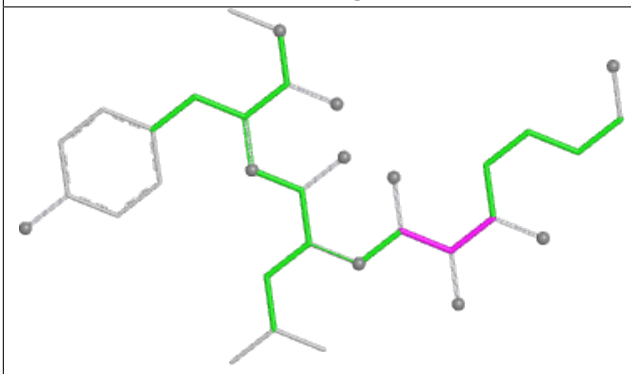
Ligand I88 A 1003



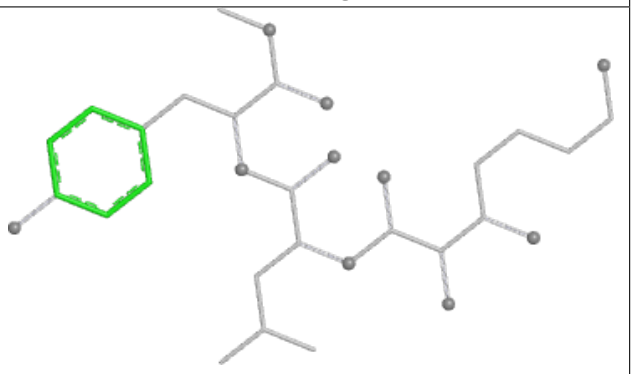
Bond lengths



Bond angles



Torsions



Rings

5.7 Other polymers

There are no such residues in this entry.

5.8 Polymer linkage issues

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	A	3

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	A	552:GLY	C	558:ASP	N	9.46
1	A	487:SER	C	515:GLY	N	3.80
1	A	485:ILE	C	487:SER	N	2.99

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	859/859 (100%)	-0.10	18 (2%) 63 69	13, 20, 35, 85	0

All (18) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	761	LEU	6.3
1	A	485	ILE	4.8
1	A	558	ASP	4.6
1	A	487	SER	3.7
1	A	228	ALA	3.7
1	A	760	ASN	3.4
1	A	229	GLU	3.2
1	A	791	PHE	3.0
1	A	45	THR	2.7
1	A	559	THR	2.6
1	A	762	SER	2.4
1	A	936	GLU	2.4
1	A	759	GLY	2.2
1	A	401	PHE	2.1
1	A	114	ARG	2.1
1	A	112	GLY	2.1
1	A	115	LEU	2.1
1	A	77	TRP	2.0

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

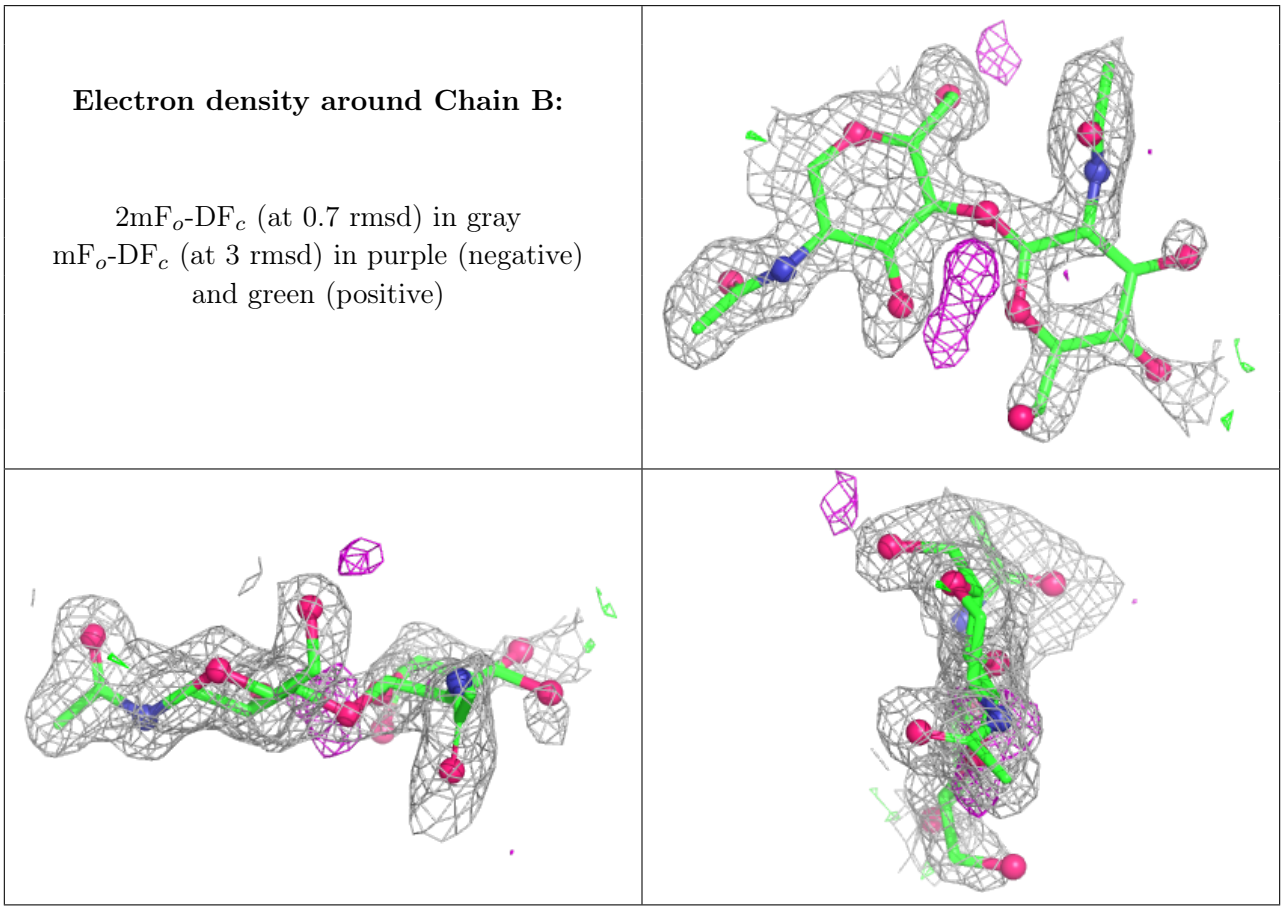
There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates (i)

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
2	NAG	B	2	14/15	0.62	0.29	55,59,67,69	0
2	NAG	B	1	14/15	0.86	0.15	26,33,41,48	0

The following is a graphical depiction of the model fit to experimental electron density for oligosaccharide. Each fit is shown from different orientation to approximate a three-dimensional view.



6.4 Ligands (i)

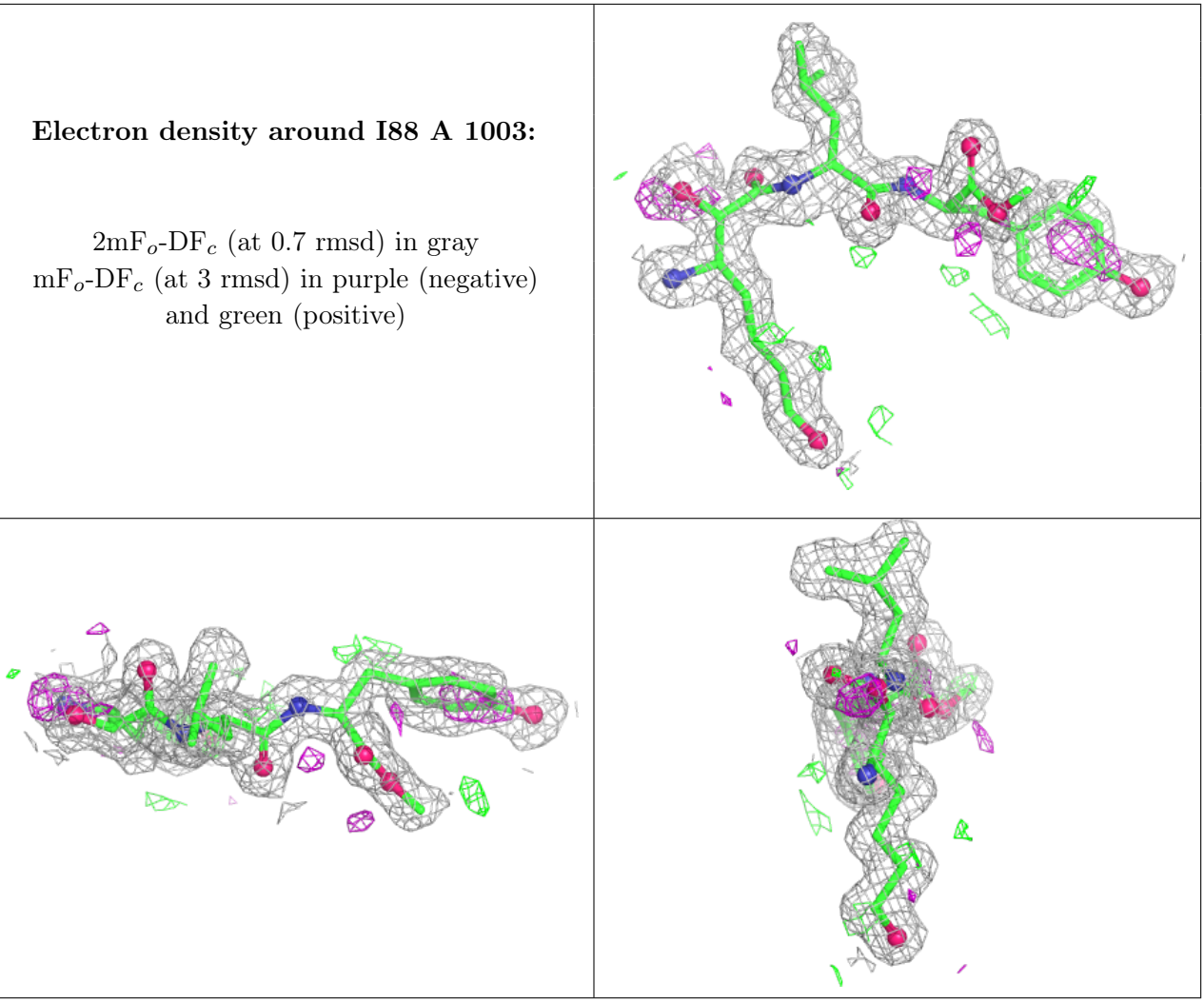
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' 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
3	NAG	A	1001	14/15	0.41	0.43	47,61,78,78	0
6	EDO	A	1010	4/4	0.67	0.17	41,49,54,56	0
8	PGE	A	1023	10/10	0.74	0.33	32,44,51,57	0
6	EDO	A	1014	4/4	0.76	0.21	28,37,47,56	0
6	EDO	A	1018	4/4	0.80	0.12	36,43,48,51	0
6	EDO	A	1016	4/4	0.81	0.17	37,44,50,53	0
7	MLT	A	1022	9/9	0.83	0.26	30,45,54,55	0
6	EDO	A	1012	4/4	0.86	0.23	30,40,45,49	0
6	EDO	A	1015	4/4	0.86	0.23	37,44,51,52	0
6	EDO	A	1008	4/4	0.89	0.15	28,39,40,46	0
6	EDO	A	1020	4/4	0.89	0.19	36,44,54,54	0
6	EDO	A	1005	4/4	0.90	0.13	21,26,29,31	0
5	I88	A	1003	33/33	0.90	0.14	14,25,32,33	0
6	EDO	A	1013	4/4	0.90	0.15	22,29,42,50	0
6	EDO	A	1017	4/4	0.90	0.16	33,40,52,62	0
6	EDO	A	1019	4/4	0.93	0.17	26,34,39,41	0
6	EDO	A	1007	4/4	0.94	0.12	21,29,38,38	0
6	EDO	A	1021	4/4	0.94	0.29	30,39,50,61	0
6	EDO	A	1009	4/4	0.95	0.16	24,36,47,55	0
6	EDO	A	1006	4/4	0.95	0.07	24,29,34,35	0
6	EDO	A	1011	4/4	0.95	0.16	25,35,46,55	0
6	EDO	A	1004	4/4	0.98	0.06	16,21,24,28	0
4	ZN	A	1002	1/1	1.00	0.07	15,15,15,15	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 I88 A 1003:

$2mF_o$-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.