Full wwPDB X-ray Structure Validation Report ⓘ

Apr 5, 2022 – 10:07 am BST

PDB ID	:	7O48
Title	:	Crystal structure of carbonic anhydrase from schistosoma mansoni with 4-(2-(3-(4-iodophenyl)thioureido)ethyl)benzenesulfonamide
Authors	:	Angeli, A.; Ferraroni, M.
Deposited on	:	2021-04-05
Resolution	:	2.00 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 ⓘ symbol.

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

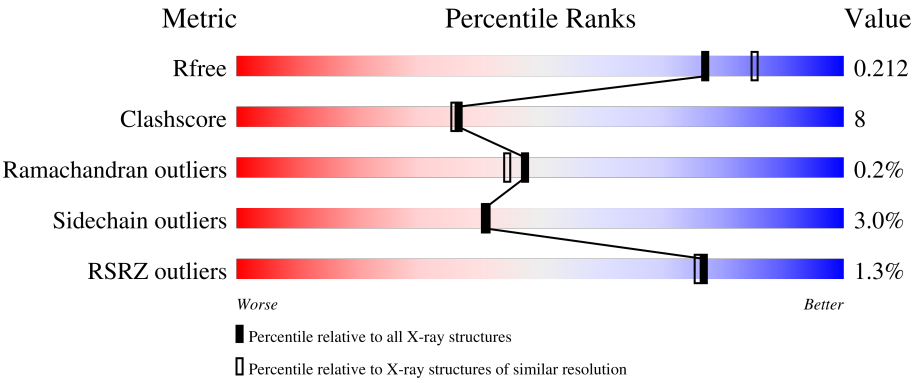
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.27
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0267
CCP4	:	7.1.010 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.27

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 2.00 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	8085 (2.00-2.00)
Clashscore	141614	9178 (2.00-2.00)
Ramachandran outliers	138981	9054 (2.00-2.00)
Sidechain outliers	138945	9053 (2.00-2.00)
RSRZ outliers	127900	7900 (2.00-2.00)

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	312	% 75% 13% • 11%
1	B	312	% 78% 11% • 11%
2	G	2	100%

2 Entry composition (i)

There are 7 unique types of molecules in this entry. The entry contains 5042 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 Carbonic anhydrase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	278	Total	C	N	O	S	0	3	0
			2270	1450	389	424	7			
1	B	279	Total	C	N	O	S	0	3	0
			2272	1450	389	426	7			

There are 68 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	12	ASP	-	expression tag	UNP A0A3Q0KSG2
A	13	ALA	-	expression tag	UNP A0A3Q0KSG2
A	14	ALA	-	expression tag	UNP A0A3Q0KSG2
A	15	GLN	-	expression tag	UNP A0A3Q0KSG2
A	16	PRO	-	expression tag	UNP A0A3Q0KSG2
A	17	ALA	-	expression tag	UNP A0A3Q0KSG2
A	18	ARG	-	expression tag	UNP A0A3Q0KSG2
A	19	ARG	-	expression tag	UNP A0A3Q0KSG2
A	20	ALA	-	expression tag	UNP A0A3Q0KSG2
A	299	ARG	-	expression tag	UNP A0A3Q0KSG2
A	300	GLY	-	expression tag	UNP A0A3Q0KSG2
A	301	GLY	-	expression tag	UNP A0A3Q0KSG2
A	302	PRO	-	expression tag	UNP A0A3Q0KSG2
A	303	GLU	-	expression tag	UNP A0A3Q0KSG2
A	304	GLN	-	expression tag	UNP A0A3Q0KSG2
A	305	LYS	-	expression tag	UNP A0A3Q0KSG2
A	306	LEU	-	expression tag	UNP A0A3Q0KSG2
A	307	ILE	-	expression tag	UNP A0A3Q0KSG2
A	308	SER	-	expression tag	UNP A0A3Q0KSG2
A	309	GLU	-	expression tag	UNP A0A3Q0KSG2
A	310	GLU	-	expression tag	UNP A0A3Q0KSG2
A	311	ASP	-	expression tag	UNP A0A3Q0KSG2
A	312	LEU	-	expression tag	UNP A0A3Q0KSG2
A	313	ASN	-	expression tag	UNP A0A3Q0KSG2
A	314	SER	-	expression tag	UNP A0A3Q0KSG2

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	315	ALA	-	expression tag	UNP A0A3Q0KSG2
A	316	VAL	-	expression tag	UNP A0A3Q0KSG2
A	317	ASP	-	expression tag	UNP A0A3Q0KSG2
A	318	HIS	-	expression tag	UNP A0A3Q0KSG2
A	319	HIS	-	expression tag	UNP A0A3Q0KSG2
A	320	HIS	-	expression tag	UNP A0A3Q0KSG2
A	321	HIS	-	expression tag	UNP A0A3Q0KSG2
A	322	HIS	-	expression tag	UNP A0A3Q0KSG2
A	323	HIS	-	expression tag	UNP A0A3Q0KSG2
B	12	ASP	-	expression tag	UNP A0A3Q0KSG2
B	13	ALA	-	expression tag	UNP A0A3Q0KSG2
B	14	ALA	-	expression tag	UNP A0A3Q0KSG2
B	15	GLN	-	expression tag	UNP A0A3Q0KSG2
B	16	PRO	-	expression tag	UNP A0A3Q0KSG2
B	17	ALA	-	expression tag	UNP A0A3Q0KSG2
B	18	ARG	-	expression tag	UNP A0A3Q0KSG2
B	19	ARG	-	expression tag	UNP A0A3Q0KSG2
B	20	ALA	-	expression tag	UNP A0A3Q0KSG2
B	299	ARG	-	expression tag	UNP A0A3Q0KSG2
B	300	GLY	-	expression tag	UNP A0A3Q0KSG2
B	301	GLY	-	expression tag	UNP A0A3Q0KSG2
B	302	PRO	-	expression tag	UNP A0A3Q0KSG2
B	303	GLU	-	expression tag	UNP A0A3Q0KSG2
B	304	GLN	-	expression tag	UNP A0A3Q0KSG2
B	305	LYS	-	expression tag	UNP A0A3Q0KSG2
B	306	LEU	-	expression tag	UNP A0A3Q0KSG2
B	307	ILE	-	expression tag	UNP A0A3Q0KSG2
B	308	SER	-	expression tag	UNP A0A3Q0KSG2
B	309	GLU	-	expression tag	UNP A0A3Q0KSG2
B	310	GLU	-	expression tag	UNP A0A3Q0KSG2
B	311	ASP	-	expression tag	UNP A0A3Q0KSG2
B	312	LEU	-	expression tag	UNP A0A3Q0KSG2
B	313	ASN	-	expression tag	UNP A0A3Q0KSG2
B	314	SER	-	expression tag	UNP A0A3Q0KSG2
B	315	ALA	-	expression tag	UNP A0A3Q0KSG2
B	316	VAL	-	expression tag	UNP A0A3Q0KSG2
B	317	ASP	-	expression tag	UNP A0A3Q0KSG2
B	318	HIS	-	expression tag	UNP A0A3Q0KSG2
B	319	HIS	-	expression tag	UNP A0A3Q0KSG2
B	320	HIS	-	expression tag	UNP A0A3Q0KSG2
B	321	HIS	-	expression tag	UNP A0A3Q0KSG2
B	322	HIS	-	expression tag	UNP A0A3Q0KSG2

Continued on next page...

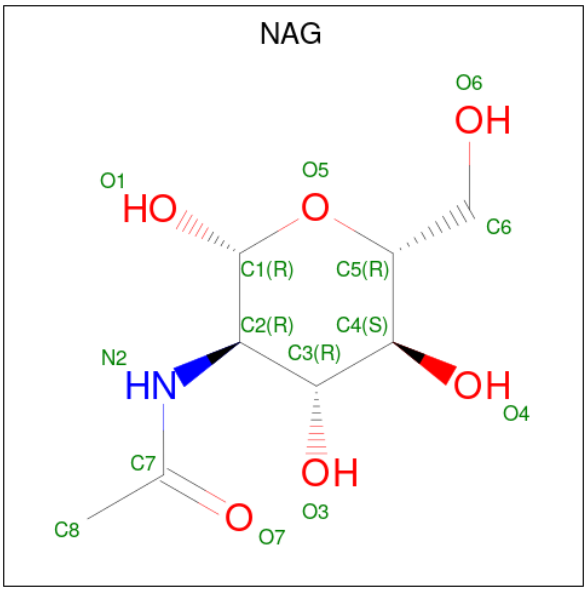
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	323	HIS	-	expression tag	UNP A0A3Q0KSG2

- 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	G	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: $C_8H_{15}NO_6$).

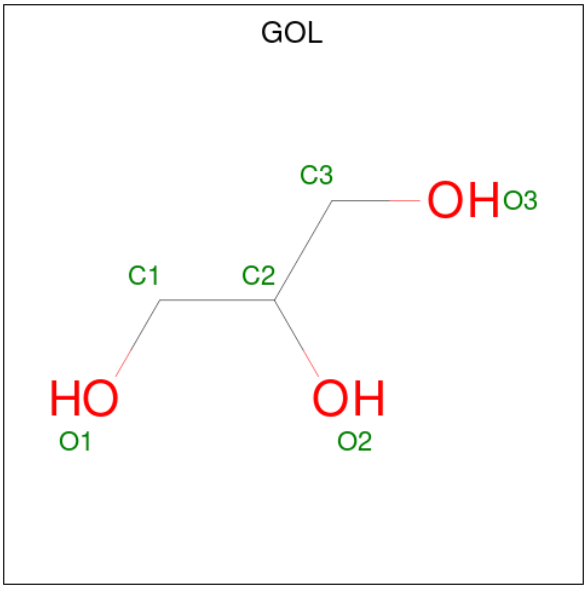


Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C N O 14 8 1 5	0	0
3	B	1	Total C N O 14 8 1 5	0	0
3	B	1	Total C N O 14 8 1 5	0	0
3	G	1	Total C N O 14 8 1 5	0	0

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

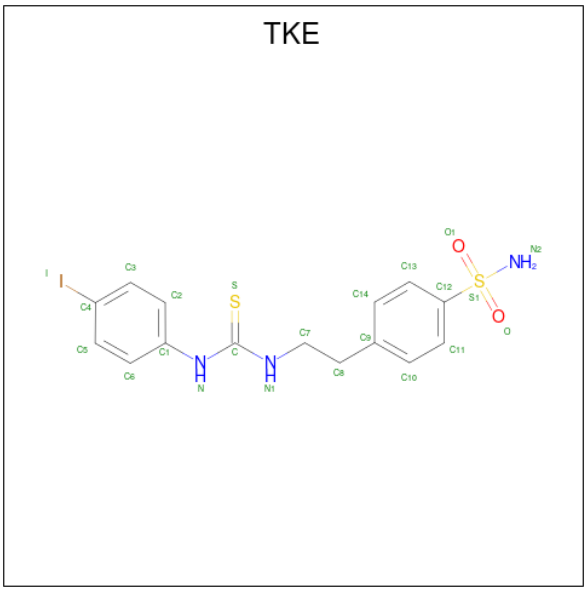
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total Zn 1 1	0	0
4	B	1	Total Zn 1 1	0	0

- Molecule 5 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



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

- Molecule 6 is 1-(4-iodophenyl)-3-[2-(4-sulfamoylphenyl)ethyl]thiourea (three-letter code: TKE) (formula: $C_{15}H_{16}IN_3O_2S_2$) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	1	Total C I N O S 23 15 1 3 2 2	7	0
6	B	1	Total C I N O S 23 15 1 3 2 2	7	0

- Molecule 7 is water.

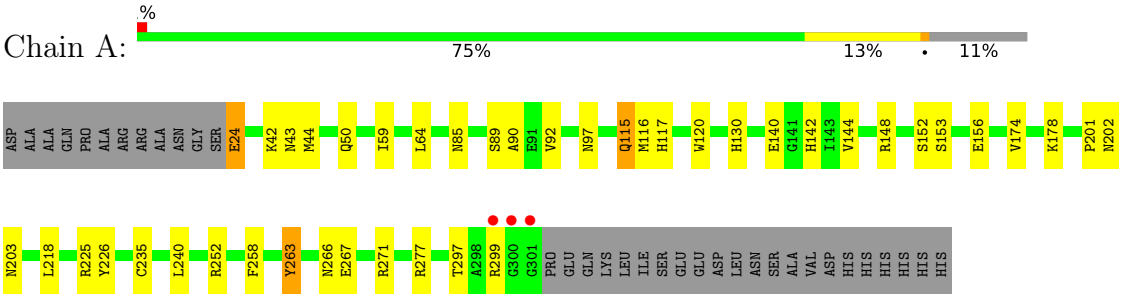
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	A	182	Total O 182 182	0	0
7	B	180	Total O 180 180	0	0

3 Residue-property plots (i)

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

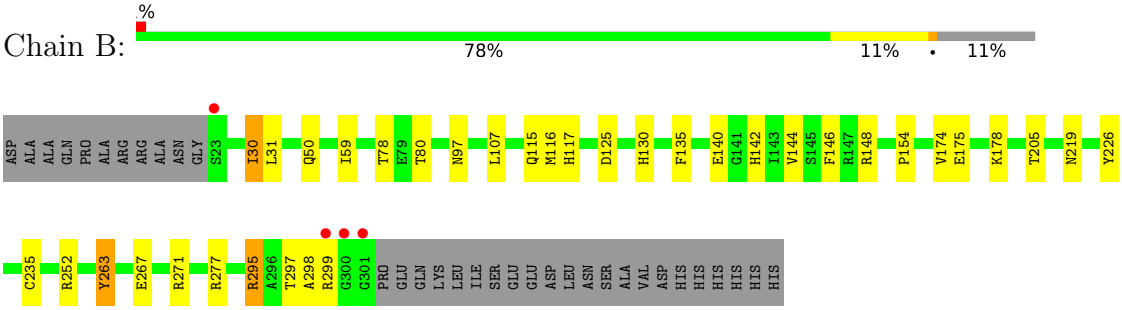
• Molecule 1: Carbonic anhydrase

Chain A:



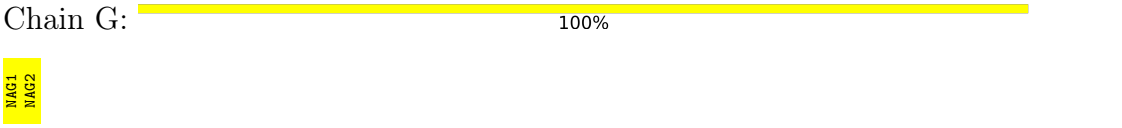
• Molecule 1: Carbonic anhydrase

Chain B:



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

Chain G:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 32 2 1	Depositor
Cell constants a, b, c, α, β, γ	103.53Å 103.53Å 133.11Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	30.00 – 2.00 89.66 – 2.00	Depositor EDS
% Data completeness (in resolution range)	99.6 (30.00-2.00) 99.7 (89.66-2.00)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	7.31 (at 2.00Å)	Xtriage
Refinement program	REFMAC 5.8.0258	Depositor
R, R_{free}	0.168 , 0.212 0.169 , 0.212	Depositor DCC
R_{free} test set	2910 reflections (5.19%)	wwPDB-VP
Wilson B-factor (Å^2)	28.7	Xtriage
Anisotropy	0.045	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	0.019 for -h,-k,l	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	5042	wwPDB-VP
Average B, all atoms (Å^2)	31.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.14% 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: GOL, TKE, 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	0.55	0/2341	0.85	0/3185
1	B	0.55	0/2343	0.87	1/3189 (0.0%)
All	All	0.55	0/4684	0.86	1/6374 (0.0%)

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	B	295	ARG	CB-CA-C	5.37	121.14	110.40

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	2270	0	2205	38	0
1	B	2272	0	2202	37	0
2	G	28	0	25	0	0
3	A	14	0	13	4	0
3	B	28	0	26	5	0
3	G	14	0	13	0	0
4	A	1	0	0	0	0
4	B	1	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	A	6	0	8	0	0
6	A	23	0	0	2	0
6	B	23	0	0	3	0
7	A	182	0	0	5	0
7	B	180	0	0	5	0
All	All	5042	0	4492	72	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 8.

All (72) 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:115[B]:GLN:OE1	1:B:117[B]:HIS:CD2	1.78	1.34
1:B:115[B]:GLN:NE2	1:B:117[B]:HIS:NE2	1.95	1.14
1:A:115[C]:GLN:NE2	6:A:404:TKE:S	2.23	1.12
1:A:97:ASN:HB3	3:A:401:NAG:H83	1.37	1.04
1:B:115[B]:GLN:OE1	1:B:117[B]:HIS:HD2	1.38	0.98
1:B:97:ASN:HB3	3:B:401:NAG:H83	1.47	0.94
1:B:115[A]:GLN:NE2	6:B:404:TKE:S	2.47	0.87
1:B:148:ARG:HH12	3:B:401:NAG:H81	1.40	0.85
1:A:117[B]:HIS:ND1	7:A:503:HOH:O	2.09	0.84
1:B:97:ASN:HB3	3:B:401:NAG:C8	2.08	0.84
1:B:115[B]:GLN:CD	1:B:117[B]:HIS:NE2	2.31	0.83
1:A:115[A]:GLN:HE22	1:A:117[A]:HIS:HD1	1.21	0.83
1:B:115[B]:GLN:OE1	1:B:117[B]:HIS:NE2	2.13	0.81
1:A:97:ASN:HB3	3:A:401:NAG:C8	2.10	0.80
1:B:115[B]:GLN:CD	1:B:117[B]:HIS:CD2	2.55	0.80
1:A:130:HIS:HE1	1:A:226:TYR:OH	1.67	0.77
1:A:271:ARG:CZ	7:A:654:HOH:O	2.33	0.76
1:B:117[B]:HIS:ND1	7:B:501:HOH:O	2.19	0.74
1:B:130:HIS:HE1	1:B:226:TYR:OH	1.69	0.74
1:A:152:SER:HB2	1:A:156:GLU:OE2	1.90	0.71
1:A:148:ARG:HH12	3:A:401:NAG:H81	1.57	0.68
1:A:117[B]:HIS:CE1	7:A:503:HOH:O	2.45	0.67
1:B:298:ALA:CB	7:B:664:HOH:O	2.42	0.67
1:B:50:GLN:HE22	1:B:235:CYS:HB3	1.60	0.67
1:A:115[A]:GLN:NE2	1:A:117[A]:HIS:HD1	1.93	0.66
1:B:115[A]:GLN:CD	6:B:404:TKE:S	2.74	0.65
1:A:50:GLN:HE22	1:A:235:CYS:HB3	1.62	0.64
1:A:115[C]:GLN:CD	6:A:404:TKE:S	2.77	0.64

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:24:GLU:HG3	1:A:266:ASN:HD22	1.65	0.62
1:B:50:GLN:HE21	1:B:277:ARG:HH22	1.51	0.59
1:A:174:VAL:HG12	1:B:174:VAL:HG12	1.84	0.59
1:B:115[A]:GLN:OE1	6:B:404:TKE:S	2.60	0.58
1:B:130:HIS:HD2	1:B:140:GLU:OE2	1.87	0.57
1:A:263:TYR:CD1	1:A:267:GLU:HG3	2.40	0.56
1:A:50:GLN:NE2	1:A:277:ARG:HH12	2.03	0.56
1:A:130:HIS:HD2	1:A:140:GLU:OE2	1.88	0.56
1:B:148:ARG:NH1	3:B:401:NAG:H81	2.17	0.56
1:A:24:GLU:HA	1:A:24:GLU:OE1	2.05	0.56
1:B:50:GLN:NE2	1:B:277:ARG:HH12	2.02	0.56
1:B:298:ALA:HB2	7:B:664:HOH:O	2.06	0.54
1:B:97:ASN:CB	3:B:401:NAG:H83	2.32	0.53
1:B:175:GLU:OE2	1:B:252:ARG:NH1	2.42	0.52
1:A:178:LYS:HE3	7:A:521:HOH:O	2.11	0.51
1:B:30:ILE:HD12	1:B:31:LEU:N	2.27	0.50
1:A:225:ARG:HD2	1:A:240:LEU:HD11	1.93	0.50
1:A:50:GLN:HE21	1:A:277:ARG:HH22	1.59	0.50
1:A:24:GLU:HG3	1:A:266:ASN:ND2	2.28	0.48
1:A:178:LYS:HD2	1:B:135:PHE:HB2	1.97	0.47
1:A:271:ARG:HD3	1:A:271:ARG:HA	1.62	0.47
1:B:117[B]:HIS:CE1	7:B:501:HOH:O	2.62	0.47
1:A:89:SER:OG	1:A:90:ALA:N	2.50	0.45
1:A:252:ARG:HB2	1:A:252:ARG:NH1	2.33	0.44
1:A:115[C]:GLN:HG2	1:A:144:VAL:HB	2.00	0.44
1:B:115[B]:GLN:NE2	1:B:117[B]:HIS:CD2	2.76	0.44
1:A:218:LEU:O	1:B:219:ASN:HB2	2.18	0.43
1:A:85:ASN:O	1:A:201:PRO:HA	2.18	0.43
1:B:125:ASP:OD1	1:B:125:ASP:N	2.50	0.43
1:A:116:MET:HA	1:A:142:HIS:O	2.19	0.42
1:B:263:TYR:CD1	1:B:267:GLU:HG3	2.54	0.42
1:B:271:ARG:HG2	7:B:648:HOH:O	2.19	0.42
1:A:153:SER:OG	1:A:156:GLU:HG3	2.20	0.42
1:A:148:ARG:NH1	3:A:401:NAG:H81	2.30	0.42
1:B:59:ILE:O	1:B:59:ILE:HG22	2.20	0.42
1:B:116:MET:HA	1:B:142:HIS:O	2.20	0.42
1:A:59:ILE:HG22	1:A:59:ILE:O	2.20	0.41
1:A:92:VAL:O	1:A:115[C]:GLN:HB2	2.19	0.41
1:B:146:PHE:CE1	1:B:154:PRO:HG3	2.56	0.41
1:B:80:THR:OG1	1:B:205:THR:HG23	2.20	0.40
1:A:120:TRP:CE2	1:A:258:PHE:HB3	2.55	0.40

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:115[A]:GLN:HG2	1:B:144:VAL:HB	2.02	0.40
1:A:43:ASN:HA	1:A:44:MET:HA	1.85	0.40
1:A:202:ASN:HA	7:A:608:HOH:O	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	279/312 (89%)	269 (96%)	10 (4%)	0	100	100
1	B	280/312 (90%)	272 (97%)	7 (2%)	1 (0%)	34	30
All	All	559/624 (90%)	541 (97%)	17 (3%)	1 (0%)	47	44

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	299	ARG

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	254/279 (91%)	245 (96%)	9 (4%)	36	35
1	B	254/279 (91%)	247 (97%)	7 (3%)	43	44

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	508/558 (91%)	492 (97%)	16 (3%)	41 40

All (16) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	24	GLU
1	A	42	LYS
1	A	64	LEU
1	A	115[A]	GLN
1	A	115[C]	GLN
1	A	203	ASN
1	A	263	TYR
1	A	297	THR
1	A	299	ARG
1	B	30	ILE
1	B	78	THR
1	B	107	LEU
1	B	178	LYS
1	B	263	TYR
1	B	295	ARG
1	B	297	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (13) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	38	HIS
1	A	50	GLN
1	A	130	HIS
1	A	202	ASN
1	A	203	ASN
1	A	256	ASN
1	B	38	HIS
1	B	50	GLN
1	B	130	HIS
1	B	149	GLN
1	B	188	ASN
1	B	256	ASN
1	B	260	ASN

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

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	G	1	2,1	14,14,15	0.68	0	17,19,21	1.46	3 (17%)
2	NAG	G	2	2	14,14,15	0.66	0	17,19,21	2.06	8 (47%)

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	G	1	2,1	-	0/6/23/26	0/1/1/1
2	NAG	G	2	2	-	2/6/23/26	0/1/1/1

There are no bond length outliers.

All (11) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	G	1	NAG	C1-O5-C5	3.72	117.23	112.19
2	G	2	NAG	O5-C1-C2	-3.69	105.47	111.29
2	G	2	NAG	C1-C2-N2	3.60	116.64	110.49
2	G	2	NAG	C2-N2-C7	-3.47	117.96	122.90

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	G	2	NAG	O5-C5-C6	3.03	111.95	107.20
2	G	1	NAG	O3-C3-C2	-2.99	103.29	109.47
2	G	2	NAG	C1-O5-C5	2.21	115.19	112.19
2	G	1	NAG	O5-C1-C2	2.18	114.74	111.29
2	G	2	NAG	C3-C4-C5	2.08	113.94	110.24
2	G	2	NAG	C6-C5-C4	-2.05	108.20	113.00
2	G	2	NAG	C4-C3-C2	2.04	114.00	111.02

There are no chirality outliers.

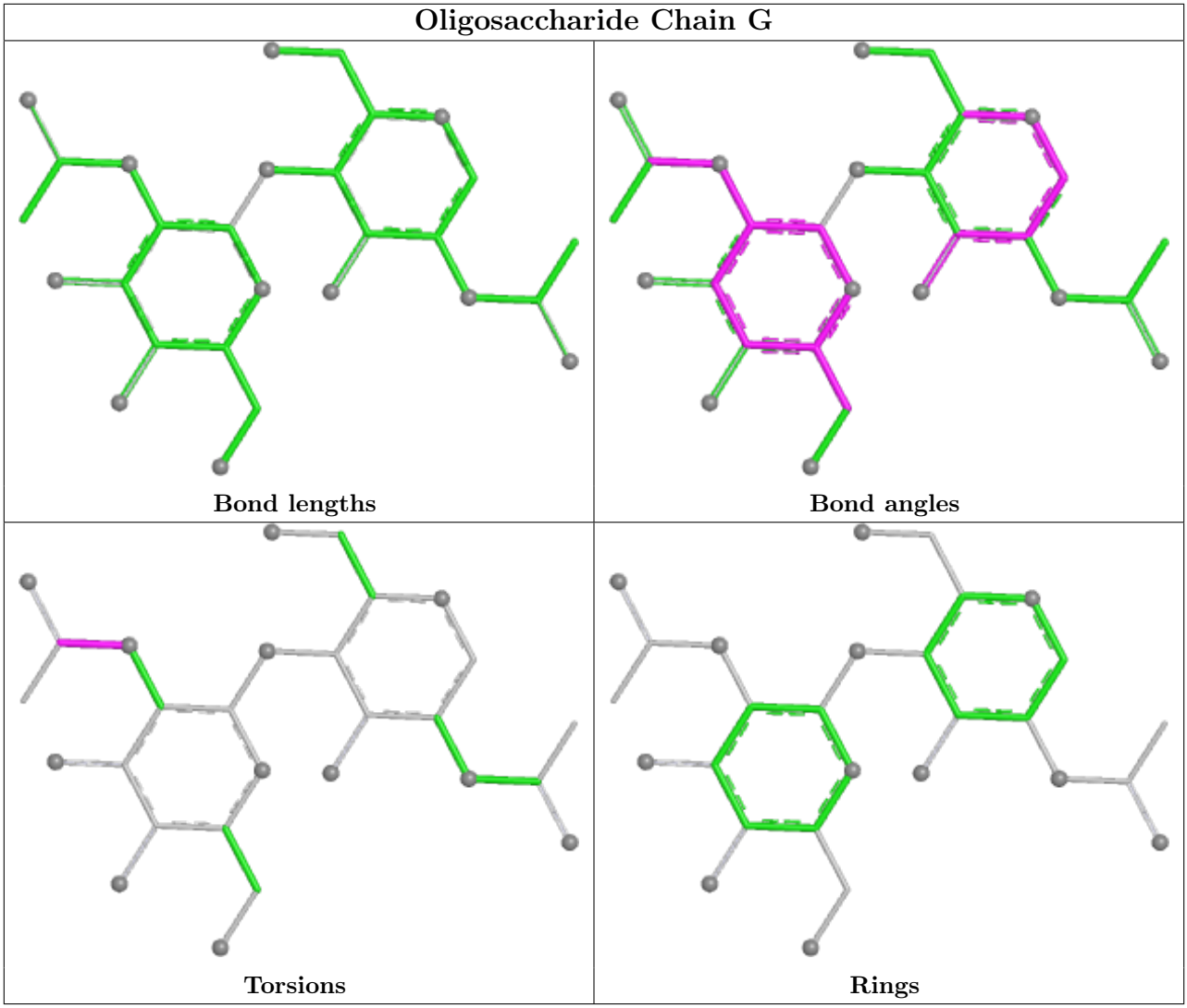
All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	G	2	NAG	C8-C7-N2-C2
2	G	2	NAG	O7-C7-N2-C2

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.



5.6 Ligand geometry (i)

Of 9 ligands modelled in this entry, 2 are monoatomic - leaving 7 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	$\#\|Z\| > 2$
5	GOL	A	403	-	5,5,5	0.44	0	5,5,5	0.96	0
6	TKE	B	404	4	24,24,24	0.88	1 (4%)	33,33,33	0.84	1 (3%)
6	TKE	A	404	4	24,24,24	1.19	1 (4%)	33,33,33	0.65	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\|>2$	Counts	RMSZ	$\#\|Z\|>2$
3	NAG	G	3	1	14,14,15	0.96	0	17,19,21	3.04	8 (47%)
3	NAG	A	401	1	14,14,15	0.94	0	17,19,21	1.66	4 (23%)
3	NAG	B	402	1	14,14,15	1.07	0	17,19,21	2.26	4 (23%)
3	NAG	B	401	1	14,14,15	1.08	1 (7%)	17,19,21	1.66	4 (23%)

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
5	GOL	A	403	-	-	0/4/4/4	-
6	TKE	B	404	4	-	8/16/16/16	0/2/2/2
6	TKE	A	404	4	-	10/16/16/16	0/2/2/2
3	NAG	G	3	1	-	2/6/23/26	0/1/1/1
3	NAG	A	401	1	-	4/6/23/26	0/1/1/1
3	NAG	B	402	1	-	2/6/23/26	0/1/1/1
3	NAG	B	401	1	-	3/6/23/26	0/1/1/1

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	A	404	TKE	C1-N	5.05	1.51	1.41
6	B	404	TKE	C1-N	-3.44	1.34	1.41
3	B	401	NAG	O5-C5	2.55	1.48	1.43

All (21) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	G	3	NAG	C1-O5-C5	8.94	124.30	112.19
3	B	402	NAG	C1-O5-C5	6.57	121.09	112.19
3	G	3	NAG	C6-C5-C4	-4.22	103.12	113.00
3	G	3	NAG	O4-C4-C5	3.83	118.82	109.30
3	G	3	NAG	C3-C4-C5	3.75	116.92	110.24
3	B	401	NAG	C2-N2-C7	3.56	127.98	122.90
6	B	404	TKE	C1-N-C	3.55	138.49	128.57
3	B	402	NAG	O4-C4-C5	3.22	117.28	109.30
3	A	401	NAG	O5-C5-C6	3.16	112.16	107.20
3	B	401	NAG	O4-C4-C5	3.15	117.11	109.30
3	A	401	NAG	O4-C4-C5	2.80	116.26	109.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	G	3	NAG	O3-C3-C2	-2.78	103.71	109.47
3	G	3	NAG	C1-C2-N2	2.68	115.07	110.49
3	B	402	NAG	C3-C4-C5	-2.56	105.67	110.24
3	B	401	NAG	O5-C5-C6	2.53	111.17	107.20
3	B	402	NAG	O5-C1-C2	2.48	115.21	111.29
3	A	401	NAG	O7-C7-C8	-2.48	117.46	122.06
3	G	3	NAG	O5-C1-C2	2.46	115.17	111.29
3	A	401	NAG	C3-C4-C5	-2.32	106.10	110.24
3	B	401	NAG	O3-C3-C2	-2.28	104.74	109.47
3	G	3	NAG	O6-C6-C5	2.13	118.60	111.29

There are no chirality outliers.

All (29) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	A	404	TKE	S-C-N1-C7
6	A	404	TKE	N-C-N1-C7
6	B	404	TKE	N1-C7-C8-C9
6	B	404	TKE	N-C-N1-C7
3	G	3	NAG	O5-C5-C6-O6
3	B	402	NAG	C8-C7-N2-C2
3	B	402	NAG	O7-C7-N2-C2
3	G	3	NAG	C4-C5-C6-O6
6	A	404	TKE	N1-C7-C8-C9
3	A	401	NAG	C8-C7-N2-C2
3	A	401	NAG	O7-C7-N2-C2
3	B	401	NAG	C8-C7-N2-C2
3	B	401	NAG	O7-C7-N2-C2
3	A	401	NAG	O5-C5-C6-O6
3	A	401	NAG	C4-C5-C6-O6
6	B	404	TKE	S-C-N1-C7
6	B	404	TKE	C7-C8-C9-C14
6	B	404	TKE	C7-C8-C9-C10
6	A	404	TKE	C13-C12-S1-O
3	B	401	NAG	C4-C5-C6-O6
6	A	404	TKE	C11-C12-S1-O
6	A	404	TKE	C13-C12-S1-N2
6	A	404	TKE	C2-C1-N-C
6	A	404	TKE	C8-C7-N1-C
6	B	404	TKE	C8-C7-N1-C
6	A	404	TKE	C6-C1-N-C
6	B	404	TKE	C13-C12-S1-O1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
6	A	404	TKE	C11-C12-S1-N2
6	B	404	TKE	C11-C12-S1-O1

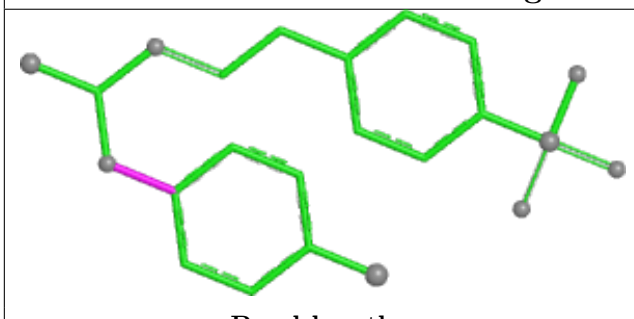
There are no ring outliers.

4 monomers are involved in 14 short contacts:

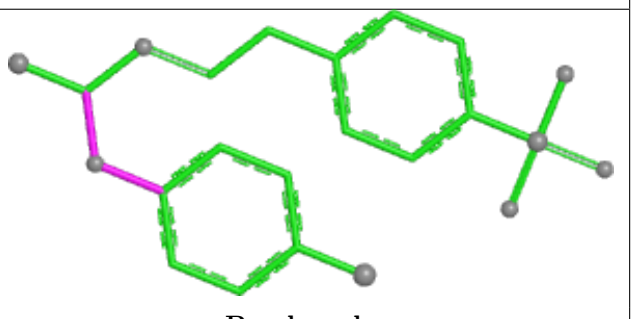
Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	B	404	TKE	3	0
6	A	404	TKE	2	0
3	A	401	NAG	4	0
3	B	401	NAG	5	0

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.

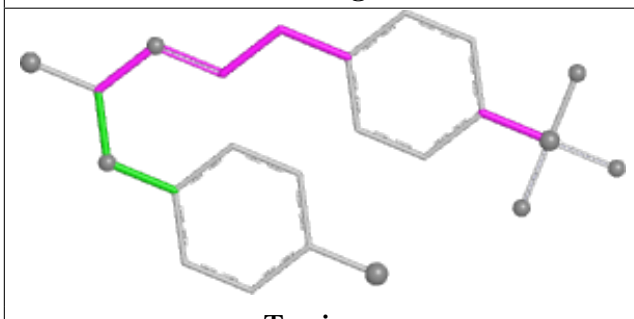
Ligand TKE B 404



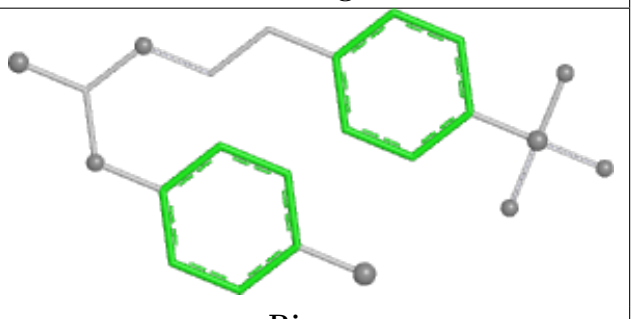
Bond lengths



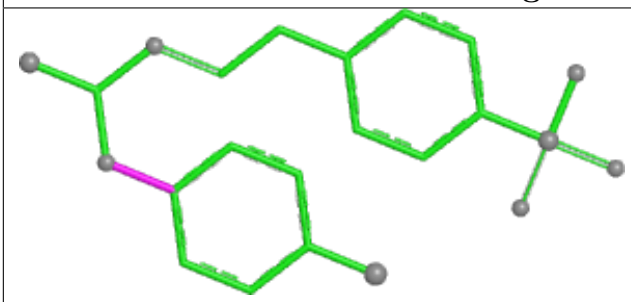
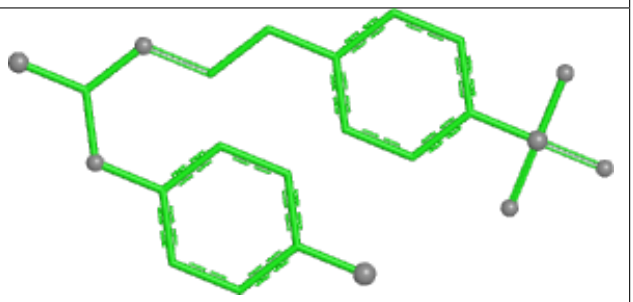
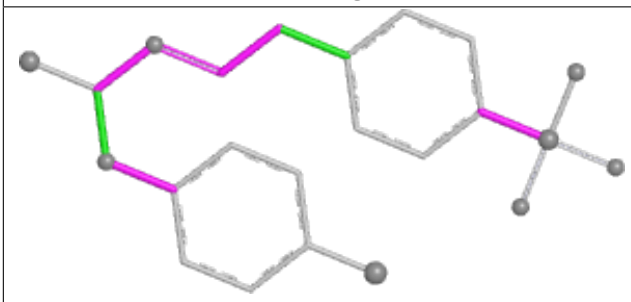
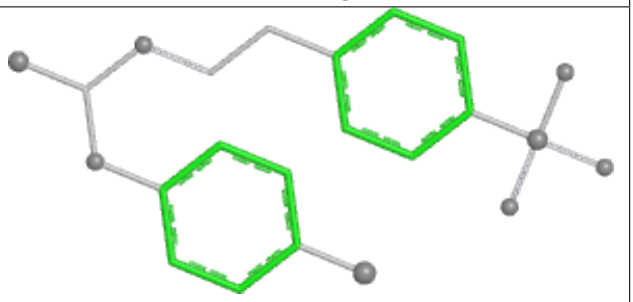
Bond angles



Torsions



Rings

Ligand TKE A 404	
 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	278/312 (89%)	-0.40	3 (1%)	80	79	18, 26, 48, 93	0
1	B	279/312 (89%)	-0.43	4 (1%)	75	74	18, 28, 51, 86	0
All	All	557/624 (89%)	-0.41	7 (1%)	77	76	18, 27, 49, 93	0

All (7) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	300	GLY	4.6
1	B	300	GLY	3.0
1	B	23	SER	2.7
1	B	301	GLY	2.6
1	A	301	GLY	2.5
1	B	299	ARG	2.5
1	A	299	ARG	2.0

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

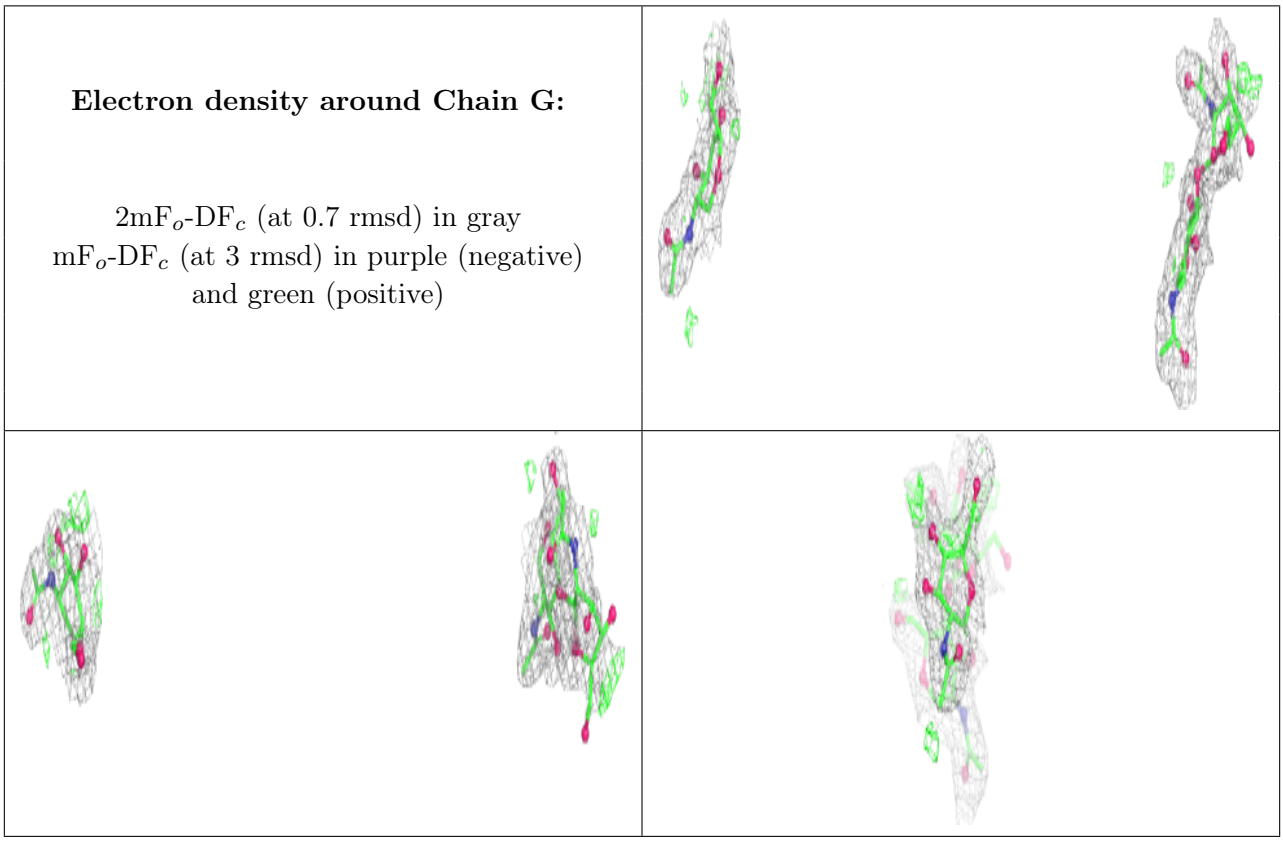
There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates

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	G	2	14/15	0.81	0.21	59,79,99,108	0
2	NAG	G	1	14/15	0.91	0.12	42,52,63,70	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

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.

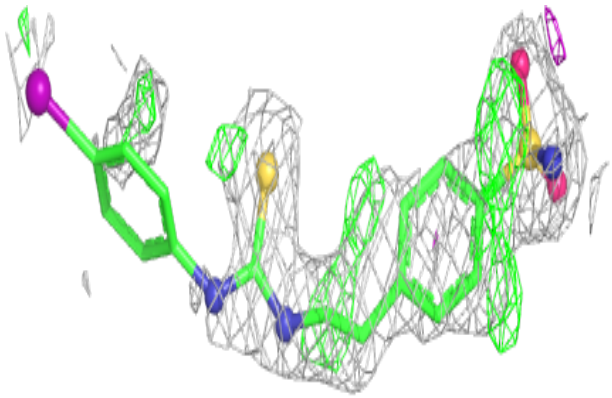
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
3	NAG	B	402	14/15	0.61	0.18	74,83,93,94	0
3	NAG	G	3	14/15	0.88	0.13	43,54,62,67	0
3	NAG	A	401	14/15	0.90	0.10	33,47,52,53	0
3	NAG	B	401	14/15	0.92	0.10	37,47,53,58	0
6	TKE	A	404	23/23	0.92	0.21	26,38,137,203	23
6	TKE	B	404	23/23	0.93	0.25	27,44,141,256	23
5	GOL	A	403	6/6	0.94	0.11	23,27,32,41	0
4	ZN	A	402	1/1	0.98	0.08	27,27,27,27	1
4	ZN	B	403	1/1	0.99	0.05	32,32,32,32	1

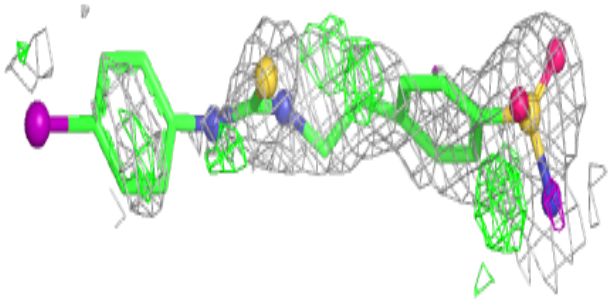
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers

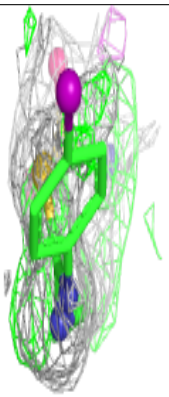
as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

Electron density around TKE A 404:

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

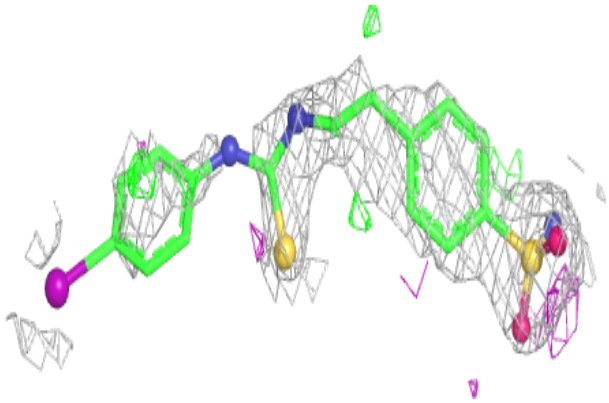


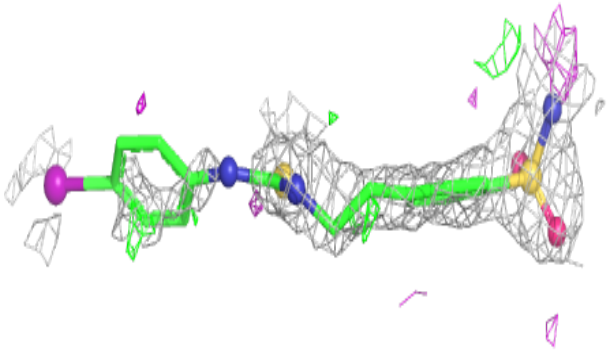


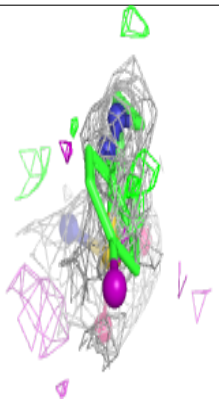


Electron density around TKE B 404:

$2mF_o$-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.