



# Full wwPDB X-ray Structure Validation Report ⓘ

Jun 18, 2024 – 06:16 PM EDT

PDB ID : 4GTX  
Title : Crystal structure of mouse Enpp1 in complex with TMP  
Authors : Kato, K.; Nishimasu, H.; Ishitani, R.; Nureki, O.  
Deposited on : 2012-08-29  
Resolution : 3.20 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](mailto: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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

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

MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.20.1
EDS	:	2.37.1
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)	:	2.37.1

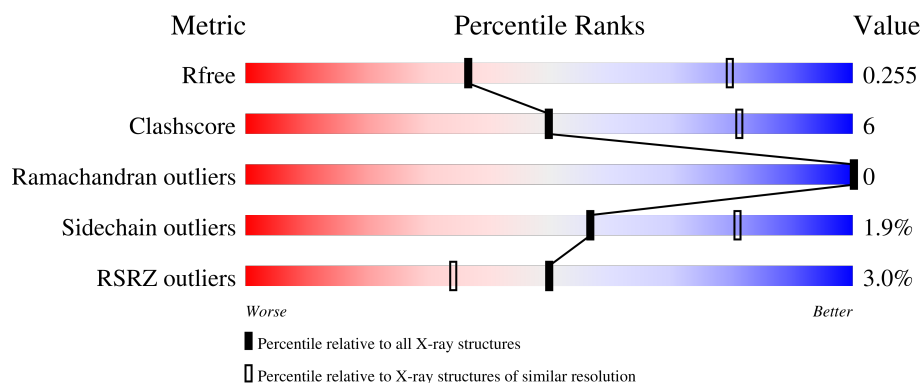
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.20 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	1133 (3.20-3.20)
Clashscore	141614	1253 (3.20-3.20)
Ramachandran outliers	138981	1234 (3.20-3.20)
Sidechain outliers	138945	1233 (3.20-3.20)
RSRZ outliers	127900	1095 (3.20-3.20)

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  $\geq 3$ , 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  $\leq 5\%$ . 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	823	<div> <div>4%</div> <div>73%</div> <div>13%</div> <div>14%</div> </div>
1	B	823	<div> <div>4%</div> <div>70%</div> <div>14%</div> <div>15%</div> </div>
2	C	6	<div> <div>33%</div> <div>17%</div> <div>50%</div> </div>
3	D	2	<div> <div>100%</div> </div>

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
4	NAG	B	1003	-	-	-	X

## 2 Entry composition [i](#)

There are 7 unique types of molecules in this entry. The entry contains 11211 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.

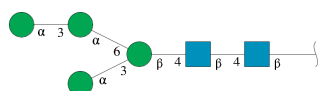
- Molecule 1 is a protein called Ectonucleotide pyrophosphatase/phosphodiesterase family member 2, Alkaline phosphodiesterase I.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	706	Total	C	N	O	S	0	0	0
			5534	3562	917	1026	29			
1	B	697	Total	C	N	O	S	0	0	0
			5473	3522	898	1024	29			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	59	ARG	LYS	engineered mutation	UNP Q9R1E6
B	59	ARG	LYS	engineered mutation	UNP Q9R1E6

- Molecule 2 is an oligosaccharide called alpha-D-mannopyranose-(1-3)-alpha-D-mannopyranose-(1-6)-[alpha-D-mannopyranose-(1-3)]beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose.



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	C	6	Total	C	N	O	0	0	0
			72	40	2	30			

- Molecule 3 is an oligosaccharide called 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose.



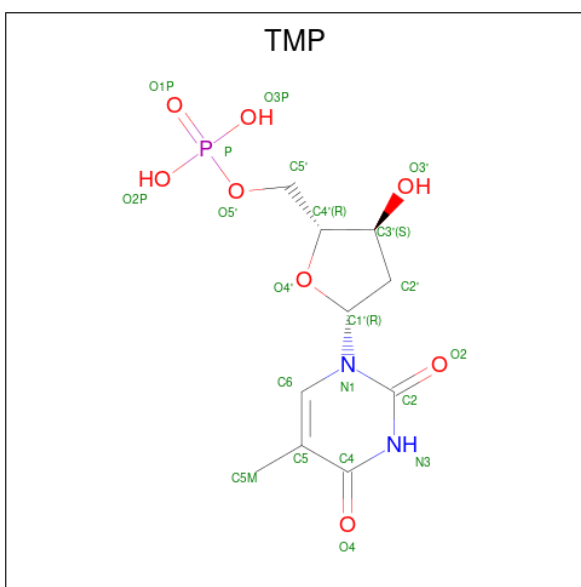
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	D	2	Total	C	N	O	0	0	0
			28	16	2	10			

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



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

- Molecule 5 is THYMIDINE-5'-PHOSPHATE (three-letter code: TMP) (formula:  $C_{10}H_{15}N_2O_8P$ ).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
5	A	1	Total	C	N	O	P	0	0
			21	10	2	8	1		
5	B	1	Total	C	N	O	P	0	0
			21	10	2	8	1		

- Molecule 6 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	2	Total	Zn	0	0
			2	2		
6	B	2	Total	Zn	0	0
			2	2		

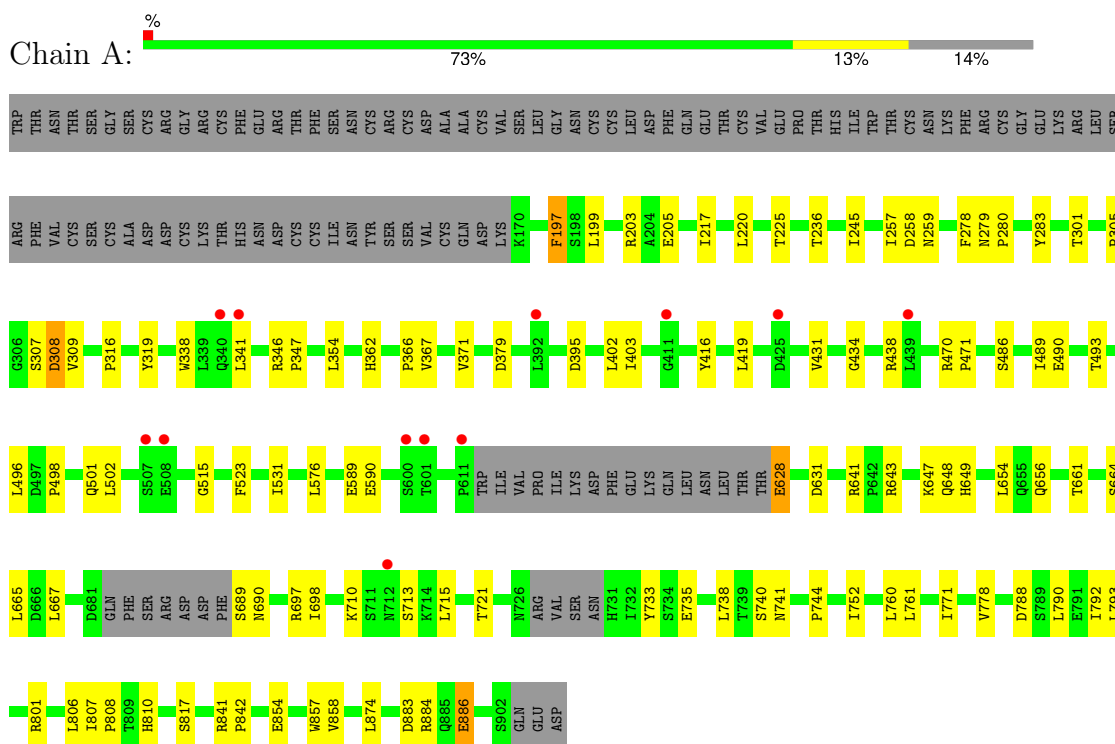
- Molecule 7 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	1	Total	Ca	0	0
			1	1		
7	B	1	Total	Ca	0	0
			1	1		

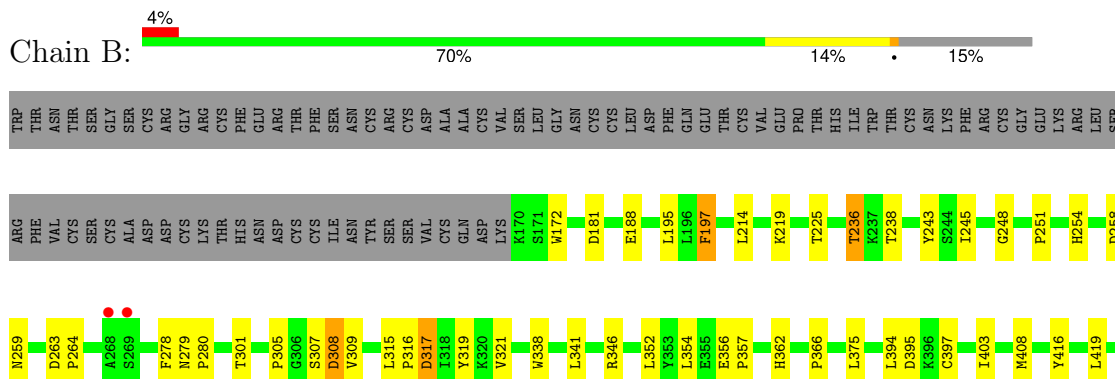
### 3 Residue-property plots [i](#)

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: Ectonucleotide pyrophosphatase/phosphodiesterase family member 2, Alkaline phosphodiesterase I

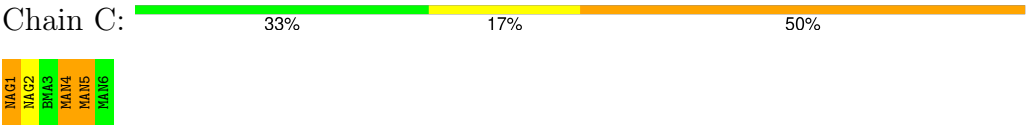


- Molecule 1: Ectonucleotide pyrophosphatase/phosphodiesterase family member 2, Alkaline phosphodiesterase I

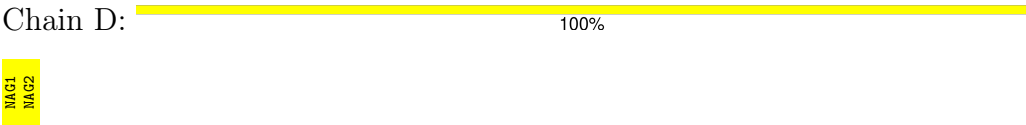




● Molecule 2:  $\alpha$ -D-mannopyranose-(1-3)- $\alpha$ -D-mannopyranose-(1-6)-[ $\alpha$ -D-mannopyranose-(1-3)] $\beta$ -D-mannopyranose-(1-4)-2-acetamido-2-deoxy- $\beta$ -D-glucopyranose-(1-4)-2-acetamido-2-deoxy- $\beta$ -D-glucopyranose



● Molecule 3: 2-acetamido-2-deoxy- $\beta$ -D-glucopyranose-(1-4)-2-acetamido-2-deoxy- $\beta$ -D-glucopyranose





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 31	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	105.72Å 105.72Å 174.44Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	49.09 – 3.20 49.09 – 3.20	Depositor EDS
% Data completeness (in resolution range)	97.1 (49.09-3.20) 97.1 (49.09-3.20)	Depositor EDS
$R_{merge}$	0.12	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.13 (at 3.19Å)	Xtriage
Refinement program	PHENIX 1.7.2_869	Depositor
R, $R_{free}$	0.218 , 0.265 0.212 , 0.255	Depositor DCC
$R_{free}$ test set	1765 reflections (5.06%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	59.5	Xtriage
Anisotropy	0.343	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 45.6	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.42$ , $\langle L^2 \rangle = 0.24$	Xtriage
Estimated twinning fraction	0.074 for -h,-k,l 0.086 for h,-h-k,-l 0.069 for -k,-h,-l	Xtriage
$F_o, F_c$ correlation	0.90	EDS
Total number of atoms	11211	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	81.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.44% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  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: BMA, NAG, TMP, ZN, CA, MAN

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	0.23	0/5698	0.43	0/7781
1	B	0.23	0/5633	0.43	0/7691
All	All	0.23	0/11331	0.43	0/15472

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	5534	0	5219	61	0
1	B	5473	0	5147	64	0
2	C	72	0	61	2	0
3	D	28	0	25	0	0
4	A	28	0	26	0	0
4	B	28	0	26	1	0
5	A	21	0	13	2	0
5	B	21	0	13	2	0
6	A	2	0	0	0	0
6	B	2	0	0	0	0
7	A	1	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
7	B	1	0	0	0	0
All	All	11211	0	10530	126	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 6.

All (126) 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:225:THR:HG21	1:A:576:LEU:HG	1.69	0.74
1:B:603:ASN:H	1:B:694:GLN:HE22	1.37	0.72
1:B:225:THR:HG21	1:B:576:LEU:HG	1.72	0.70
1:A:308:ASP:N	1:A:308:ASP:OD1	2.24	0.69
1:A:715:LEU:O	1:A:801:ARG:NH2	2.23	0.69
1:A:589:GLU:OE2	1:A:641:ARG:NH2	2.28	0.66
1:A:316:PRO:HG2	1:A:319:TYR:HB3	1.76	0.65
1:B:733:TYR:CZ	1:B:735:GLU:HB2	2.31	0.64
1:B:301:THR:HG21	1:B:307:SER:HB2	1.79	0.64
1:B:650:ARG:NH1	1:B:666:ASP:OD1	2.30	0.64
1:A:643:ARG:NH2	1:A:883:ASP:OD2	2.31	0.63
1:A:362:HIS:ND1	1:A:515:GLY:O	2.32	0.61
1:B:438:ARG:HB3	1:B:490:GLU:HB2	1.83	0.60
1:A:733:TYR:CZ	1:A:735:GLU:HB2	2.36	0.60
1:A:793:LEU:HD22	1:A:806:LEU:HD21	1.84	0.59
1:B:486:SER:HB3	1:B:489:ILE:HG13	1.84	0.59
1:B:667:LEU:HB3	1:B:669:MET:HG2	1.85	0.58
1:B:308:ASP:HB3	1:B:321:VAL:HA	1.85	0.57
1:B:315:LEU:HD12	1:B:316:PRO:HD2	1.85	0.57
1:B:496:LEU:HD21	1:B:502:LEU:HB2	1.86	0.57
1:A:656:GLN:OE1	1:A:697:ARG:NH2	2.37	0.56
1:A:778:VAL:HB	1:A:810:HIS:HB2	1.87	0.56
1:B:643:ARG:NH2	1:B:883:ASP:OD2	2.38	0.56
1:A:434:GLY:HA2	1:A:501:GLN:HE21	1.70	0.55
1:B:589:GLU:OE2	1:B:641:ARG:NH2	2.39	0.55
1:A:438:ARG:HB3	1:A:490:GLU:HB2	1.88	0.55
1:B:793:LEU:HD22	1:B:806:LEU:HD21	1.89	0.55
1:B:188:GLU:O	1:B:346:ARG:NH2	2.35	0.54
1:A:752:ILE:HG23	1:A:858:VAL:HG13	1.88	0.54
1:B:575:HIS:CE1	1:B:893:ARG:HD3	2.43	0.53
1:A:698:ILE:HD13	1:A:740:SER:HB3	1.90	0.53
1:B:607:CYS:HB3	1:B:691:CYS:HA	1.90	0.53

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:203:ARG:NH2	1:A:205:GLU:OE1	2.42	0.53
1:A:258:ASP:OD1	1:A:259:ASN:N	2.41	0.53
1:B:338:TRP:HA	1:B:341:LEU:HD13	1.91	0.53
1:A:710:LYS:H	1:A:801:ARG:NH1	2.07	0.52
1:A:854:GLU:HA	1:A:857:TRP:NE1	2.24	0.52
1:B:305:PRO:HG3	5:B:1005:TMP:H52	1.90	0.52
1:B:362:HIS:ND1	1:B:515:GLY:O	2.41	0.52
1:A:338:TRP:HA	1:A:341:LEU:HD13	1.92	0.52
1:A:649:HIS:NE2	1:A:664:SER:OG	2.33	0.52
1:A:628:GLU:N	1:A:631:ASP:OD1	2.44	0.51
1:B:554:LEU:HD11	1:B:577:LEU:HD21	1.92	0.51
1:A:735:GLU:HA	1:A:738:LEU:HD13	1.92	0.51
1:A:733:TYR:OH	1:A:735:GLU:OE1	2.26	0.50
1:B:278:PHE:HA	1:B:309:VAL:HG22	1.94	0.50
1:B:778:VAL:HB	1:B:810:HIS:HB2	1.93	0.50
1:A:301:THR:HG21	1:A:307:SER:HB3	1.93	0.50
1:A:438:ARG:HG2	1:A:493:THR:HG22	1.92	0.50
1:A:486:SER:HB3	1:A:489:ILE:HG13	1.93	0.50
1:A:689:SER:O	1:A:690:ASN:HB3	2.13	0.49
1:A:721:THR:HB	1:A:741:ASN:HD21	1.78	0.49
1:B:600:SER:OG	1:B:601:THR:N	2.46	0.48
1:B:316:PRO:HG2	1:B:319:TYR:HB3	1.94	0.48
1:A:419:LEU:HD12	1:A:431:VAL:HG21	1.94	0.48
1:B:197:PHE:HE1	1:B:354:LEU:HD13	1.78	0.48
1:A:788:ASP:HB3	1:A:792:ILE:HG23	1.96	0.48
1:A:496:LEU:HD21	1:A:502:LEU:HB2	1.96	0.48
1:B:394:LEU:HD23	1:B:397:CYS:HB3	1.96	0.48
1:A:643:ARG:NH1	1:B:181:ASP:OD1	2.47	0.47
1:B:317:ASP:OD1	1:B:317:ASP:N	2.46	0.47
1:B:445:PRO:HD2	4:B:1003:NAG:H82	1.96	0.47
1:A:884:ARG:HG3	1:A:886:GLU:HG2	1.96	0.47
1:B:480:ARG:HD3	1:B:866:ARG:O	2.15	0.47
1:A:245:ILE:HD13	1:A:403:ILE:HD13	1.97	0.47
1:B:214:LEU:HD21	1:B:375:LEU:HB3	1.96	0.47
1:B:744:PRO:HG2	1:B:808:PRO:HD3	1.97	0.47
1:B:537:PHE:O	1:B:578:LYS:NZ	2.41	0.46
1:A:305:PRO:HG3	5:A:1009:TMP:H52	1.98	0.46
1:A:220:LEU:HB3	1:A:531:ILE:HG21	1.98	0.46
1:A:713:SER:HB3	1:A:801:ARG:NH1	2.31	0.46
1:A:366:PRO:O	1:A:371:VAL:HG21	2.14	0.46
1:A:217:ILE:HG12	1:A:379:ASP:OD1	2.16	0.45

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:245:ILE:HD13	1:B:403:ILE:HD13	1.98	0.45
1:A:744:PRO:HG2	1:A:808:PRO:HD3	1.98	0.45
1:B:236:THR:HG21	1:B:526:MET:HB3	1.98	0.45
1:A:257:ILE:HG12	1:A:489:ILE:HG12	1.99	0.45
1:A:283:TYR:CE1	1:A:309:VAL:HG21	2.52	0.44
1:A:647:LYS:HA	1:A:648:GLN:HA	1.54	0.44
1:B:258:ASP:OD1	1:B:259:ASN:N	2.46	0.44
1:B:656:GLN:HE22	1:B:661:THR:HG22	1.82	0.44
1:A:259:ASN:ND2	5:A:1009:TMP:O3P	2.44	0.44
1:A:654:LEU:HB2	1:A:661:THR:HG23	1.99	0.44
1:B:853:ARG:O	1:B:856:SER:OG	2.27	0.44
1:A:279:ASN:HA	1:A:280:PRO:HD3	1.81	0.44
1:A:649:HIS:CE1	1:A:667:LEU:HD13	2.53	0.43
1:B:649:HIS:NE2	1:B:664:SER:OG	2.39	0.43
2:C:4:MAN:H2	2:C:5:MAN:H2	2.00	0.43
1:A:470:ARG:HA	1:A:471:PRO:HD3	1.84	0.43
1:A:854:GLU:HA	1:A:857:TRP:CD1	2.53	0.43
1:B:807:ILE:HA	1:B:808:PRO:HD3	1.84	0.43
1:A:649:HIS:HE1	1:A:667:LEU:HD13	1.83	0.43
1:B:841:ARG:HA	1:B:842:PRO:HD3	1.86	0.43
1:B:172:TRP:CZ3	1:B:219:LYS:HD3	2.54	0.43
1:B:419:LEU:HD12	1:B:431:VAL:HG21	2.01	0.43
1:A:199:LEU:HD11	1:A:402:LEU:HD11	2.01	0.42
1:A:197:PHE:HE1	1:A:354:LEU:HD13	1.84	0.42
1:A:807:ILE:HA	1:A:808:PRO:HD3	1.84	0.42
1:B:439:LEU:HB3	1:B:492:LEU:HD23	2.02	0.42
1:B:735:GLU:HA	1:B:738:LEU:HD13	2.01	0.42
1:A:278:PHE:HA	1:A:309:VAL:HG22	2.02	0.42
1:A:395:ASP:OD1	1:A:395:ASP:N	2.51	0.42
1:B:195:LEU:HD21	1:B:352:LEU:HD12	2.01	0.42
1:B:248:GLY:O	1:B:566:ASN:HB2	2.20	0.42
1:B:739:THR:O	1:B:742:ILE:HG13	2.20	0.42
1:B:238:THR:OG1	5:B:1005:TMP:O2P	2.38	0.42
1:B:279:ASN:HA	1:B:280:PRO:HD3	1.86	0.42
1:B:251:PRO:HA	1:B:254:HIS:CE1	2.54	0.42
1:A:761:LEU:HG	1:A:771:ILE:HD11	2.02	0.41
2:C:1:NAG:H61	2:C:2:NAG:N2	2.35	0.41
1:B:315:LEU:HD12	1:B:315:LEU:HA	1.83	0.41
1:A:841:ARG:HA	1:A:842:PRO:HD3	1.90	0.41
1:B:356:GLU:OE2	1:B:357:PRO:HA	2.20	0.41
1:B:854:GLU:HA	1:B:857:TRP:NE1	2.34	0.41

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:884:ARG:HG3	1:B:886:GLU:OE1	2.20	0.41
1:A:498:PRO:HD3	1:A:523:PHE:CZ	2.56	0.41
1:A:771:ILE:HG22	1:A:817:SER:OG	2.19	0.41
1:B:820:GLN:HE21	1:B:820:GLN:HB2	1.64	0.41
1:B:366:PRO:HG3	1:B:408:MET:HG3	2.02	0.41
1:B:490:GLU:HA	1:B:491:PRO:HD3	1.95	0.41
1:B:647:LYS:HA	1:B:648:GLN:HA	1.54	0.41
1:B:263:ASP:HA	1:B:264:PRO:HD2	1.95	0.41
1:A:346:ARG:HA	1:A:347:PRO:HD3	1.93	0.40
1:B:243:TYR:CE2	1:B:254:HIS:HB2	2.56	0.40
1:B:395:ASP:OD1	1:B:395:ASP:N	2.54	0.40
1:B:677:PHE:HE2	1:B:762:GLN:HA	1.86	0.40

There are no symmetry-related clashes.

## 5.3 Torsion angles ⓘ

### 5.3.1 Protein backbone ⓘ

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	698/823 (85%)	675 (97%)	23 (3%)	0	100	100
1	B	685/823 (83%)	661 (96%)	24 (4%)	0	100	100
All	All	1383/1646 (84%)	1336 (97%)	47 (3%)	0	100	100

There are no Ramachandran outliers to report.

### 5.3.2 Protein sidechains ⓘ

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	600/755 (80%)	588 (98%)	12 (2%)	55	80
1	B	595/755 (79%)	584 (98%)	11 (2%)	59	82
All	All	1195/1510 (79%)	1172 (98%)	23 (2%)	57	81

All (23) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	197	PHE
1	A	236	THR
1	A	308	ASP
1	A	367	VAL
1	A	416	TYR
1	A	590	GLU
1	A	628	GLU
1	A	665	LEU
1	A	760	LEU
1	A	790	LEU
1	A	874	LEU
1	A	886	GLU
1	B	197	PHE
1	B	236	THR
1	B	308	ASP
1	B	317	ASP
1	B	416	TYR
1	B	656	GLN
1	B	665	LEU
1	B	690	ASN
1	B	791	GLU
1	B	792	ILE
1	B	874	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (15) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	452	ASN
1	A	501	GLN
1	A	525	ASN
1	A	553	ASN
1	A	634	HIS

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	658	GLN
1	A	741	ASN
1	A	769	ASN
1	A	820	GLN
1	A	850	HIS
1	B	229	ASN
1	B	656	GLN
1	B	694	GLN
1	B	750	GLN
1	B	820	GLN

### 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 ⓘ

8 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	C	1	1,2	14,14,15	0.58	0	17,19,21	0.92	1 (5%)
2	NAG	C	2	2	14,14,15	0.50	0	17,19,21	0.65	0
2	BMA	C	3	2	11,11,12	0.54	0	15,15,17	0.78	0
2	MAN	C	4	2	11,11,12	0.51	0	15,15,17	2.40	4 (26%)
2	MAN	C	5	2	11,11,12	0.52	0	15,15,17	0.86	1 (6%)
2	MAN	C	6	2	11,11,12	0.68	0	15,15,17	0.62	0
3	NAG	D	1	1,3	14,14,15	0.56	0	17,19,21	1.18	1 (5%)
3	NAG	D	2	3	14,14,15	0.45	0	17,19,21	1.14	3 (17%)



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	C	1	1,2	-	2/6/23/26	0/1/1/1
2	NAG	C	2	2	-	4/6/23/26	0/1/1/1
2	BMA	C	3	2	-	0/2/19/22	0/1/1/1
2	MAN	C	4	2	-	2/2/19/22	0/1/1/1
2	MAN	C	5	2	-	1/2/19/22	1/1/1/1
2	MAN	C	6	2	-	0/2/19/22	0/1/1/1
3	NAG	D	1	1,3	-	0/6/23/26	0/1/1/1
3	NAG	D	2	3	-	2/6/23/26	0/1/1/1

There are no bond length outliers.

All (10) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	C	4	MAN	C1-O5-C5	7.35	122.04	112.19
2	C	4	MAN	C3-C4-C5	3.26	116.14	110.23
3	D	2	NAG	C1-O5-C5	2.99	116.19	112.19
2	C	4	MAN	O5-C5-C4	2.36	116.57	110.83
2	C	5	MAN	C1-O5-C5	2.28	115.24	112.19
3	D	1	NAG	C4-C3-C2	2.24	114.31	111.02
2	C	4	MAN	C1-C2-C3	2.16	112.79	109.64
3	D	2	NAG	C4-C3-C2	-2.16	107.86	111.02
2	C	1	NAG	C4-C3-C2	2.14	114.16	111.02
3	D	2	NAG	C2-N2-C7	-2.14	120.03	122.90

There are no chirality outliers.

All (11) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	C	2	NAG	O5-C5-C6-O6
2	C	4	MAN	O5-C5-C6-O6
2	C	2	NAG	C8-C7-N2-C2
2	C	2	NAG	C4-C5-C6-O6
2	C	4	MAN	C4-C5-C6-O6
2	C	2	NAG	O7-C7-N2-C2
2	C	5	MAN	O5-C5-C6-O6
2	C	1	NAG	C4-C5-C6-O6
3	D	2	NAG	C4-C5-C6-O6

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms
2	C	1	NAG	O5-C5-C6-O6
3	D	2	NAG	O5-C5-C6-O6

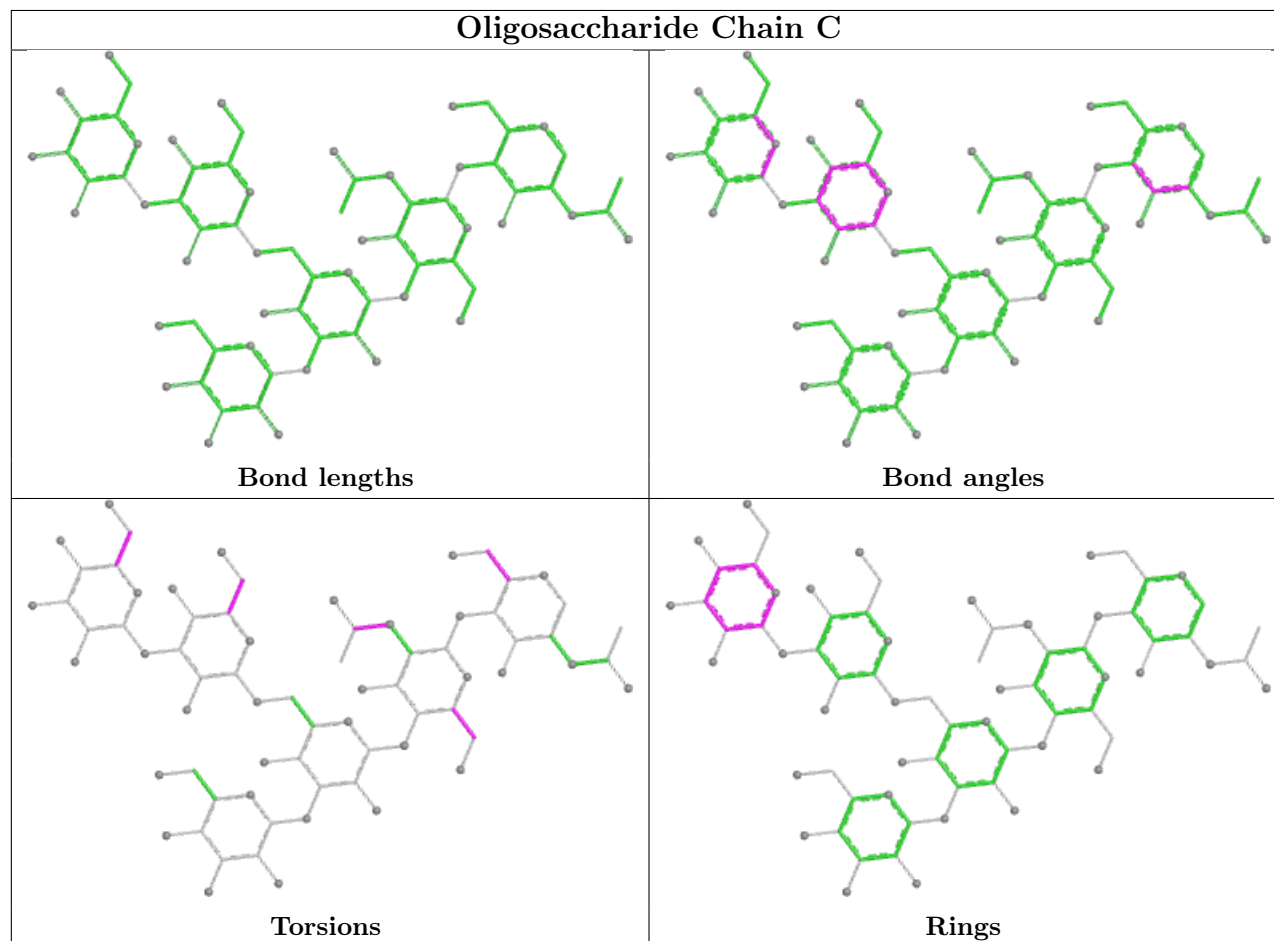
All (1) ring outliers are listed below:

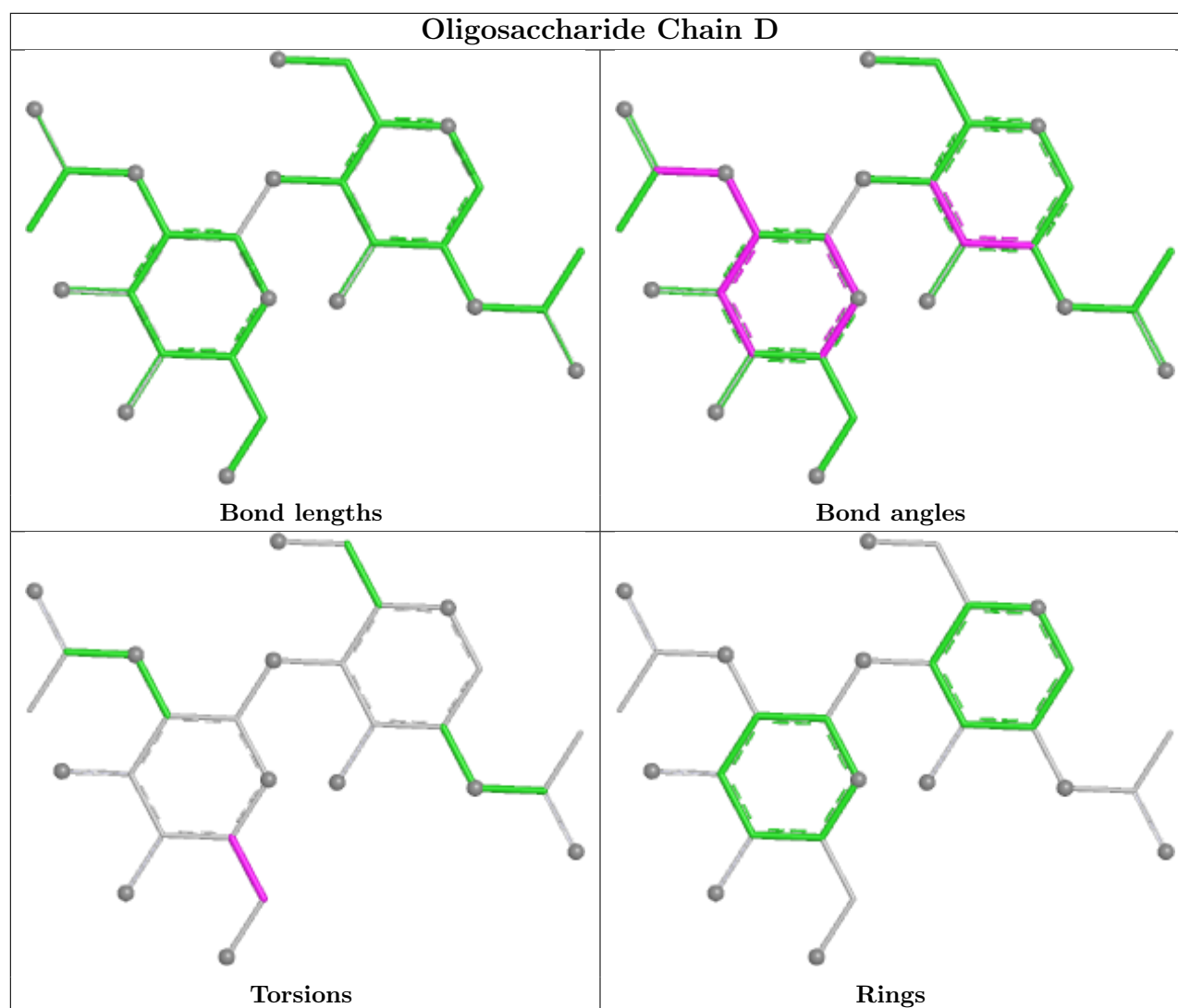
Mol	Chain	Res	Type	Atoms
2	C	5	MAN	C1-C2-C3-C4-C5-O5

4 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	C	5	MAN	1	0
2	C	1	NAG	1	0
2	C	2	NAG	1	0
2	C	4	MAN	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 oligosaccharide.





## 5.6 Ligand geometry [i](#)

Of 12 ligands modelled in this entry, 6 are monoatomic - leaving 6 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	$\# Z  > 2$
4	NAG	A	1008	1	14,14,15	0.50	0	17,19,21	0.99	1 (5%)
5	TMP	A	1009	6	22,22,22	2.61	6 (27%)	32,33,33	1.76	7 (21%)
5	TMP	B	1005	6	22,22,22	2.62	6 (27%)	32,33,33	1.82	9 (28%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
4	NAG	A	1007	1	14,14,15	0.50	0	17,19,21	0.87	1 (5%)
4	NAG	B	1004	1	14,14,15	0.50	0	17,19,21	0.97	1 (5%)
4	NAG	B	1003	1	14,14,15	0.59	0	17,19,21	1.03	1 (5%)

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
4	NAG	A	1008	1	-	2/6/23/26	0/1/1/1
5	TMP	A	1009	6	-	2/10/22/22	0/2/2/2
5	TMP	B	1005	6	-	2/10/22/22	0/2/2/2
4	NAG	A	1007	1	-	2/6/23/26	0/1/1/1
4	NAG	B	1004	1	-	0/6/23/26	0/1/1/1
4	NAG	B	1003	1	-	0/6/23/26	0/1/1/1

All (12) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	B	1005	TMP	C6-C5	8.21	1.48	1.34
5	A	1009	TMP	C6-C5	8.18	1.47	1.34
5	A	1009	TMP	C2-N1	4.64	1.45	1.38
5	B	1005	TMP	C2-N1	4.59	1.45	1.38
5	B	1005	TMP	C2-N3	3.87	1.44	1.38
5	A	1009	TMP	C2-N3	3.70	1.44	1.38
5	A	1009	TMP	C6-N1	3.32	1.43	1.38
5	B	1005	TMP	C6-N1	3.31	1.43	1.38
5	A	1009	TMP	C4-N3	2.61	1.43	1.38
5	A	1009	TMP	C3'-C4'	-2.57	1.46	1.53
5	B	1005	TMP	C4-N3	2.53	1.43	1.38
5	B	1005	TMP	C3'-C4'	-2.48	1.46	1.53

All (20) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	B	1005	TMP	C5-C6-N1	-4.25	118.70	123.31
5	A	1009	TMP	C5-C6-N1	-4.19	118.76	123.31
5	B	1005	TMP	O4-C4-C5	-3.77	120.60	124.92
5	A	1009	TMP	O4-C4-C5	-3.77	120.61	124.92

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	A	1009	TMP	C5-C4-N3	3.55	118.41	115.32
5	B	1005	TMP	C5-C4-N3	3.40	118.28	115.32
4	A	1008	NAG	C1-O5-C5	3.31	116.62	112.19
5	B	1005	TMP	O4'-C1'-N1	3.20	113.54	107.86
5	A	1009	TMP	O5'-C5'-C4'	2.82	118.61	108.99
5	B	1005	TMP	O5'-C5'-C4'	2.75	118.35	108.99
5	A	1009	TMP	O4'-C1'-N1	2.64	112.56	107.86
5	B	1005	TMP	C1'-N1-C6	-2.61	116.36	120.74
5	A	1009	TMP	C4-N3-C2	-2.56	123.99	127.34
5	A	1009	TMP	C1'-N1-C6	-2.49	116.56	120.74
4	B	1004	NAG	C1-O5-C5	2.44	115.46	112.19
4	A	1007	NAG	C1-O5-C5	2.42	115.44	112.19
5	B	1005	TMP	C4-N3-C2	-2.38	124.22	127.34
5	B	1005	TMP	C6-N1-C2	2.13	123.43	121.30
4	B	1003	NAG	O5-C1-C2	2.07	114.50	111.29
5	B	1005	TMP	C5M-C5-C6	-2.04	120.10	122.85

There are no chirality outliers.

All (8) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	A	1009	TMP	O4'-C4'-C5'-O5'
5	B	1005	TMP	C3'-C4'-C5'-O5'
4	A	1007	NAG	O5-C5-C6-O6
5	A	1009	TMP	C3'-C4'-C5'-O5'
4	A	1008	NAG	C8-C7-N2-C2
4	A	1007	NAG	C4-C5-C6-O6
4	A	1008	NAG	O7-C7-N2-C2
5	B	1005	TMP	O4'-C4'-C5'-O5'

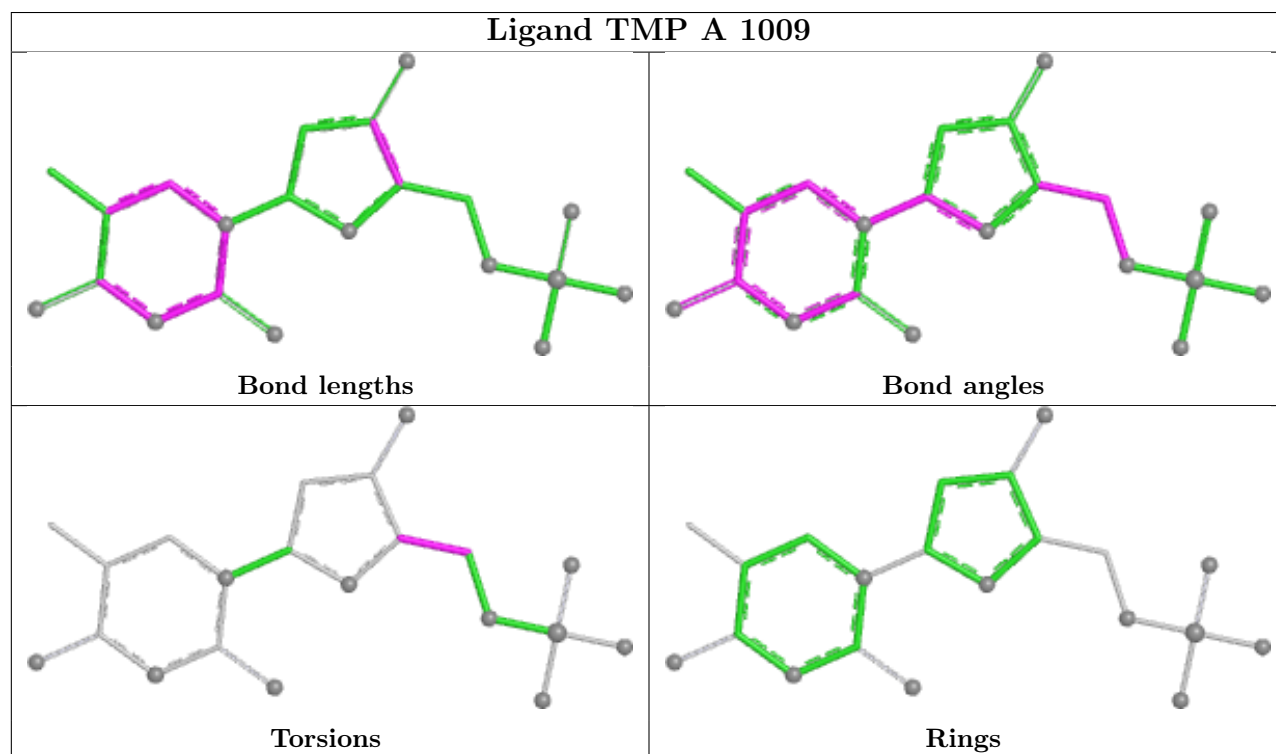
There are no ring outliers.

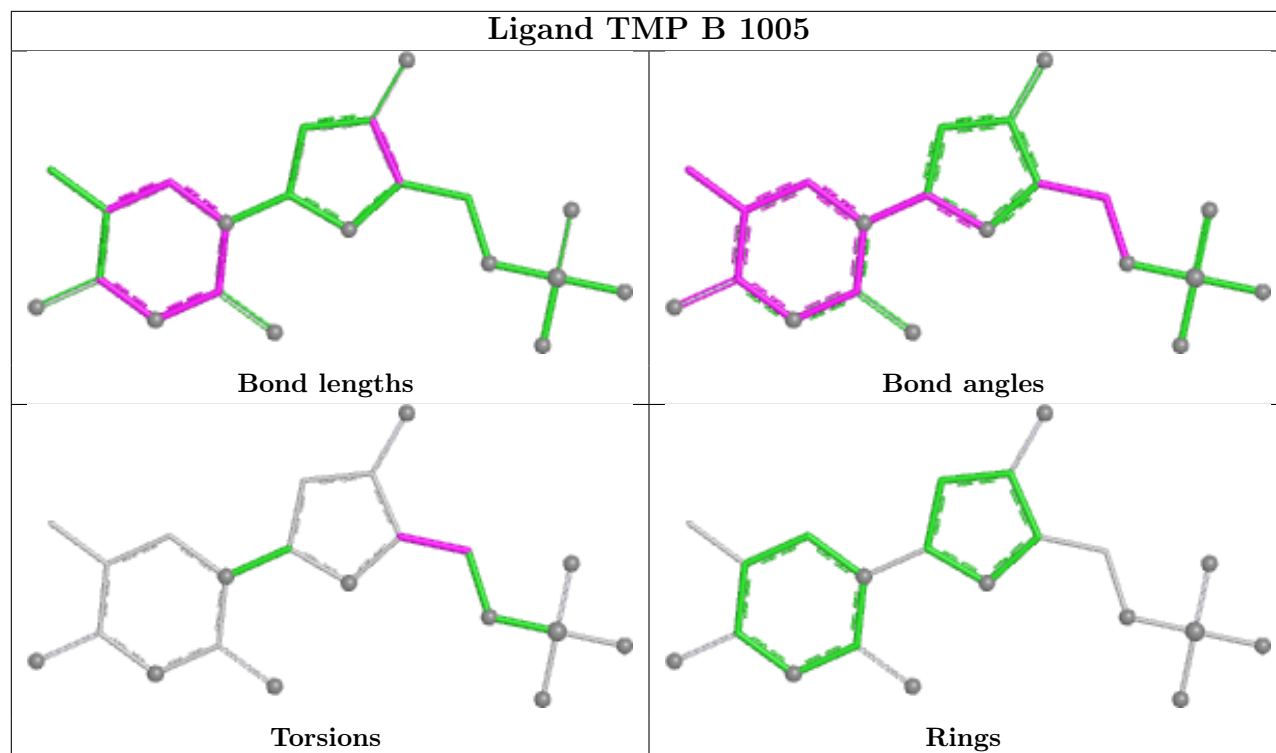
3 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	A	1009	TMP	2	0
5	B	1005	TMP	2	0
4	B	1003	NAG	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 than 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.





## 5.7 Other polymers [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

## 6 Fit of model and data ⓘ

### 6.1 Protein, DNA and RNA chains ⓘ

In the following table, the column labelled ‘#RSRZ > 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	706/823 (85%)	-0.07	12 (1%) 70 57	42, 77, 118, 161	0
1	B	697/823 (84%)	0.02	30 (4%) 35 22	35, 81, 123, 178	0
All	All	1403/1646 (85%)	-0.03	42 (2%) 50 34	35, 79, 121, 178	0

All (42) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	600	SER	4.8
1	B	441	PRO	4.2
1	B	698	ILE	4.1
1	A	712	ASN	3.4
1	A	411	GLY	3.3
1	B	608	THR	3.3
1	A	601	THR	3.1
1	A	508	GLU	3.0
1	B	442	THR	3.0
1	B	448	TYR	2.9
1	B	605	LEU	2.9
1	B	692	LEU	2.7
1	B	700	LEU	2.6
1	A	600	SER	2.6
1	A	507	SER	2.5
1	B	609	CYS	2.5
1	B	805	ILE	2.5
1	A	340	GLN	2.5
1	B	428	ASN	2.5
1	B	507	SER	2.5
1	B	726	ASN	2.5
1	B	601	THR	2.5
1	B	704	HIS	2.3
1	A	425	ASP	2.3

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	715	LEU	2.3
1	B	699	PRO	2.3
1	B	696	LEU	2.2
1	B	604	ASP	2.2
1	A	439	LEU	2.2
1	A	392	LEU	2.2
1	B	268	ALA	2.2
1	B	707	SER	2.2
1	A	611	PRO	2.2
1	A	341	LEU	2.1
1	B	439	LEU	2.1
1	B	603	ASN	2.1
1	B	269	SER	2.1
1	B	691	CYS	2.1
1	B	602	SER	2.1
1	B	737	LEU	2.1
1	B	611	PRO	2.0
1	B	694	GLN	2.0

## 6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

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<sup>th</sup> 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.

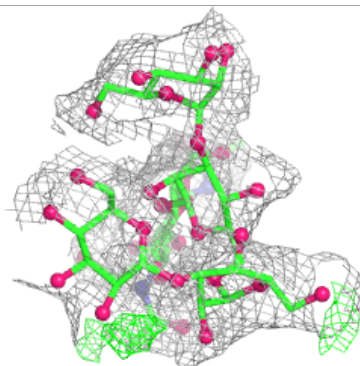
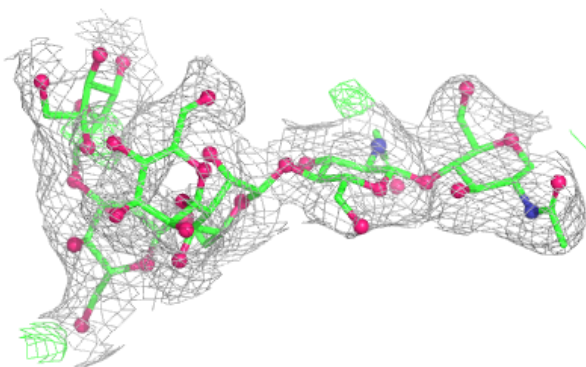
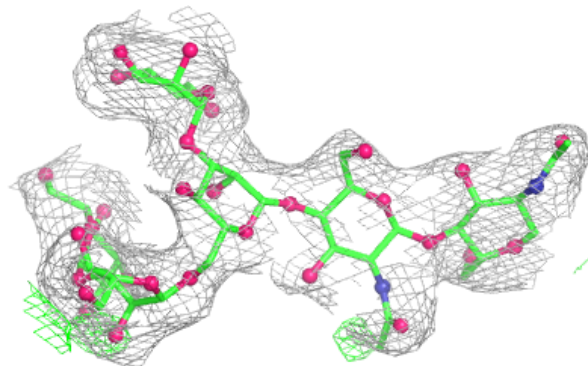
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
2	MAN	C	5	11/12	0.75	0.18	102,112,121,123	0
2	MAN	C	4	11/12	0.82	0.19	89,99,108,110	0
3	NAG	D	2	14/15	0.84	0.25	97,104,110,110	0
2	BMA	C	3	11/12	0.86	0.14	88,101,116,117	0
2	MAN	C	6	11/12	0.88	0.14	70,84,99,99	0
2	NAG	C	2	14/15	0.94	0.20	62,82,90,92	0
3	NAG	D	1	14/15	0.95	0.21	58,67,77,94	0
2	NAG	C	1	14/15	0.95	0.19	49,59,66,74	0

The following is a graphical depiction of the model fit to experimental electron density for oligosac-

charide. Each fit is shown from different orientation to approximate a three-dimensional view.

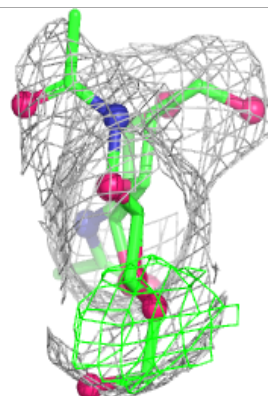
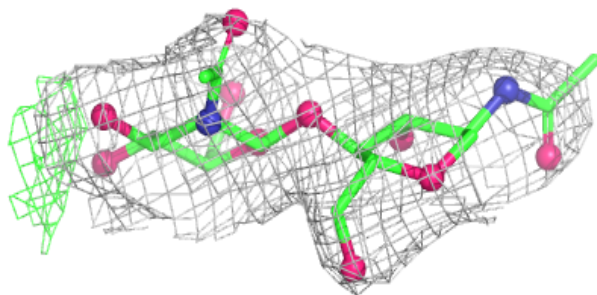
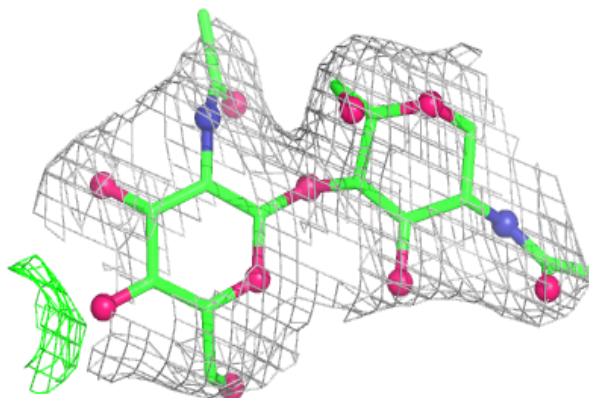
**Electron density around Chain C:**

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



**Electron density around Chain D:**

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



## 6.4 Ligands ⓘ

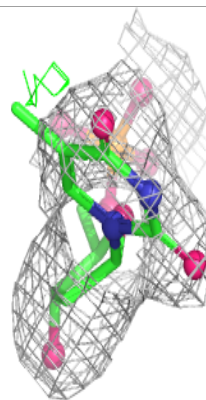
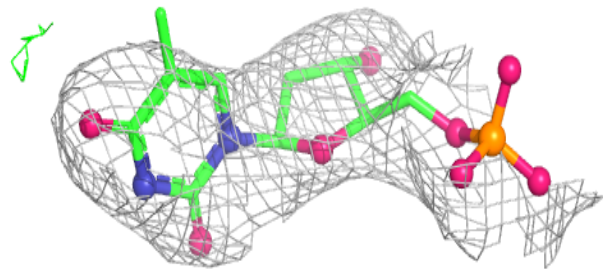
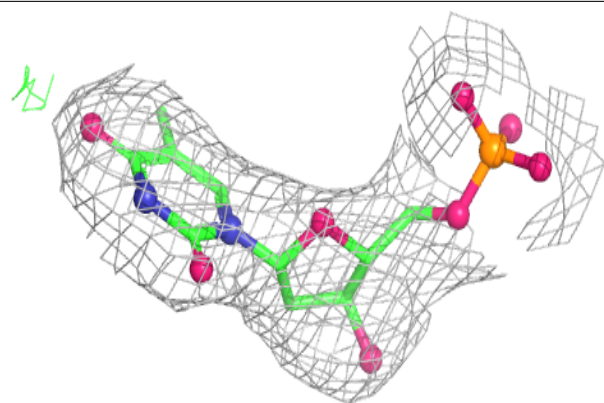
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
4	NAG	B	1003	14/15	0.78	0.57	99,123,131,131	0
4	NAG	A	1007	14/15	0.79	0.34	101,111,120,121	0
4	NAG	B	1004	14/15	0.86	0.26	95,108,114,117	0
4	NAG	A	1008	14/15	0.92	0.31	93,104,117,118	0
5	TMP	B	1005	21/21	0.96	0.17	45,76,85,94	0
5	TMP	A	1009	21/21	0.97	0.15	56,73,82,99	0
7	CA	A	1012	1/1	0.97	0.12	68,68,68,68	0
6	ZN	B	1006	1/1	0.98	0.17	73,73,73,73	0
7	CA	B	1008	1/1	0.98	0.15	92,92,92,92	0
6	ZN	B	1007	1/1	0.99	0.14	68,68,68,68	0
6	ZN	A	1011	1/1	0.99	0.14	70,70,70,70	0
6	ZN	A	1010	1/1	0.99	0.15	72,72,72,72	0

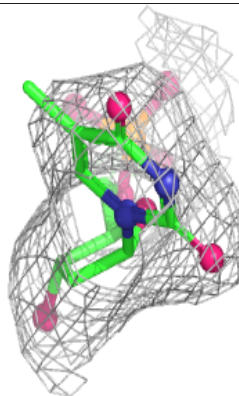
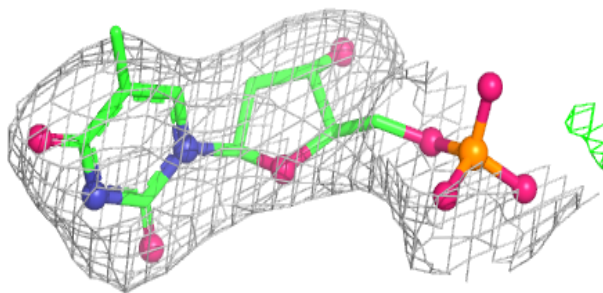
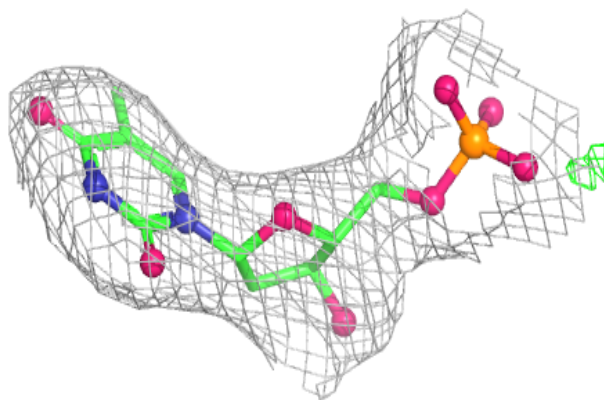
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

**Electron density around TMP B 1005:**

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

**Electron density around TMP A 1009:**

$2mF_o-DF_c$  (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.