



# Full wwPDB X-ray Structure Validation Report ⓘ

Aug 9, 2020 – 04:18 PM BST

PDB ID : 1XC9  
Title : Structure of a high-fidelity polymerase bound to a benzo[a]pyrene adduct that blocks replication  
Authors : Hsu, G.W.; Huang, X.; Luneva, N.P.; Geacintov, N.E.; Beese, L.S.  
Deposited on : 2004-09-01  
Resolution : 1.90 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 following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.13.1  
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) : 2.13.1

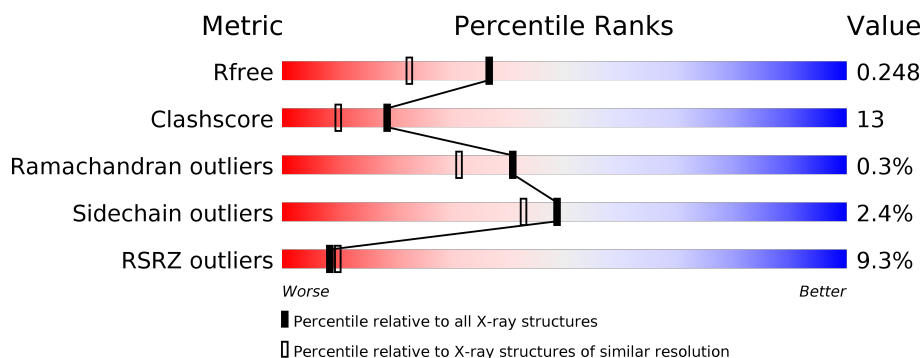
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.90 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	6207 (1.90-1.90)
Clashscore	141614	6847 (1.90-1.90)
Ramachandran outliers	138981	6760 (1.90-1.90)
Sidechain outliers	138945	6760 (1.90-1.90)
RSRZ outliers	127900	6082 (1.90-1.90)

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 on 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	B	10	<div> <div>40%</div> <div> <div>60%</div> <div>40%</div> </div> </div>
2	C	15	<div> <div>20%</div> <div> <div>7%</div> <div>73%</div> <div>20%</div> </div> </div>
3	A	580	<div> <div>8%</div> <div> <div>75%</div> <div>23%</div> </div> </div>
4	D	2	<div> <div>50%</div> <div>50%</div> </div>

## 2 Entry composition

There are 8 unique types of molecules in this entry. The entry contains 5533 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 DNA chain called DNA primer strand.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	B	10	Total	C	N	O	P	0	0	0
			209	99	42	59	9			

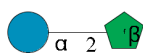
- Molecule 2 is a DNA chain called DNA template strand with benzo[a]pyrene adduct.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	12	Total	C	N	O	P	0	0	0
			234	113	40	70	11			

- Molecule 3 is a protein called DNA polymerase I.

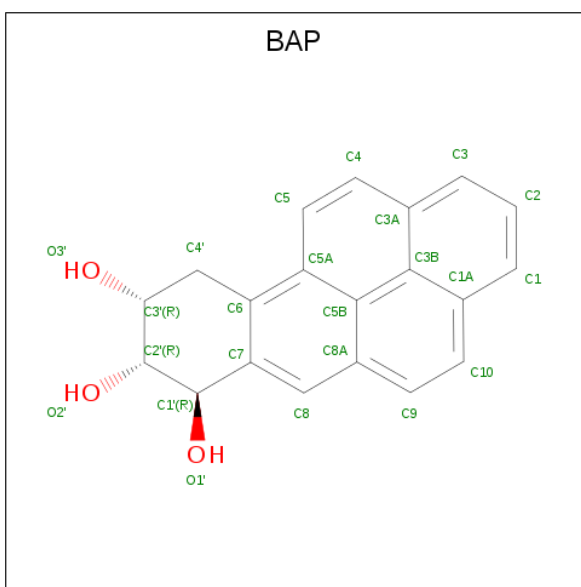
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	A	580	Total	C	N	O	S	0	0	0
			4650	2956	807	870	17			

- Molecule 4 is an oligosaccharide called beta-D-fructofuranose-(2-1)-alpha-D-glucopyranose.



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace
4	D	2	Total	C	O	0	0	0
			23	12	11			

- Molecule 5 is 1,2,3-TRIHYDROXY-1,2,3,4-TETRAHYDROBENZO[A]PYRENE (three-letter code: BAP) (formula: C<sub>20</sub>H<sub>16</sub>O<sub>3</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	C	1	Total	C	O	0	0
			23	20	3		

- Molecule 6 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



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

*Continued on next page...*

*Continued from previous page...*

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

- Molecule 7 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	1	Total	Mg	0	0
			1	1		

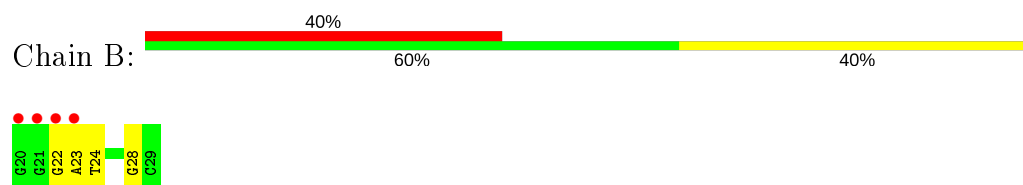
- Molecule 8 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	B	4	Total	O	0	0
			4	4		
8	C	8	Total	O	0	0
			8	8		
8	A	361	Total	O	0	0
			361	361		

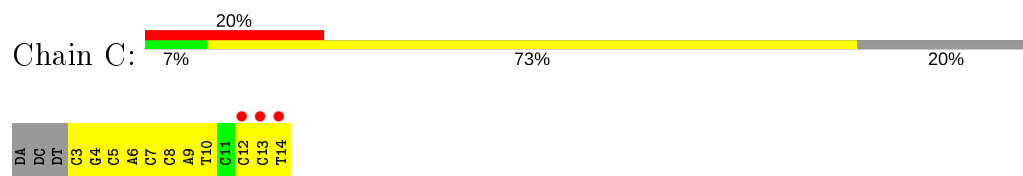
### 3 Residue-property plots

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

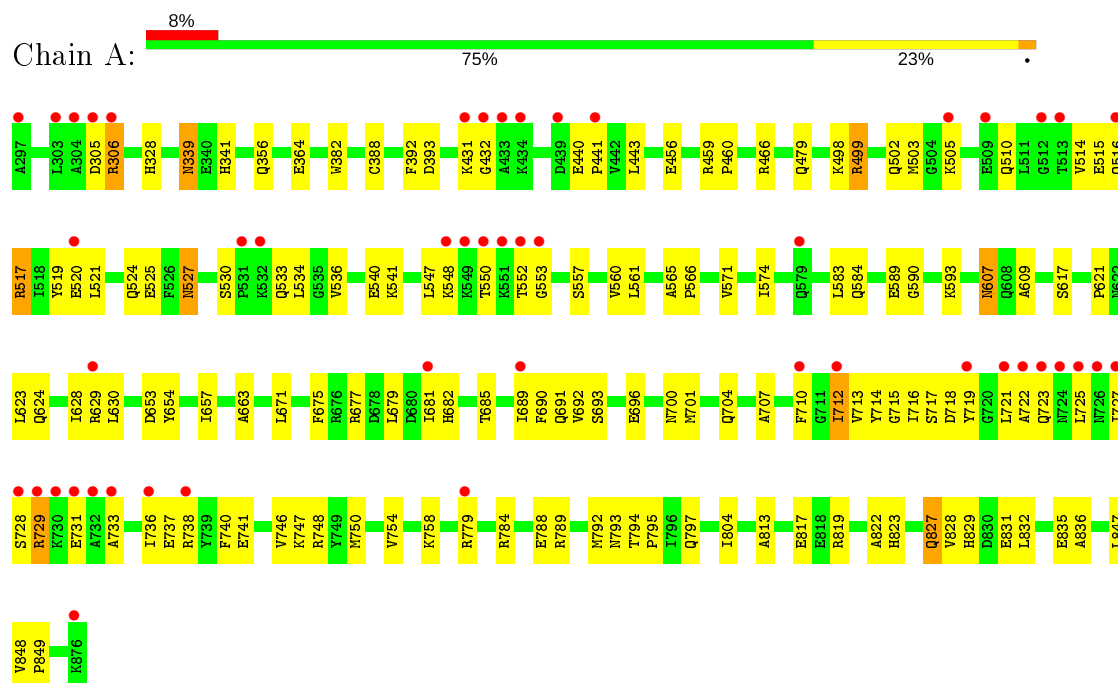
- Molecule 1: DNA primer strand



- Molecule 2: DNA template strand with benzo[a]pyrene adduct



- Molecule 3: DNA polymerase I



- Molecule 4: beta-D-fructofuranose-(2-1)-alpha-D-glucopyranose

Chain D:



GLA1  
FRU2

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	88.75Å 93.73Å 104.51Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	42.76 – 1.90 42.76 – 1.85	Depositor EDS
% Data completeness (in resolution range)	96.8 (42.76-1.90) 95.8 (42.76-1.85)	Depositor EDS
$R_{merge}$	0.04	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.54 (at 1.86Å)	Xtriage
Refinement program	CNS	Depositor
R, $R_{free}$	0.221 , 0.255 0.216 , 0.248	Depositor DCC
$R_{free}$ test set	3414 reflections (4.75%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	25.0	Xtriage
Anisotropy	0.099	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.38 , 49.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	5533	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	35.0	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.*

<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: MG, GLC, FRU, SO4, BAP

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	B	0.29	0/235	0.68	0/363
2	C	0.33	0/260	0.77	0/397
3	A	0.32	0/4734	0.56	0/6398
All	All	0.32	0/5229	0.58	0/7158

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	B	209	0	114	3	0
2	C	234	0	135	12	0
3	A	4650	0	4698	123	0
4	D	23	0	21	1	0
5	C	23	0	15	0	0
6	A	20	0	0	0	0
7	A	1	0	0	0	0
8	A	361	0	0	10	0
8	B	4	0	0	0	0
8	C	8	0	0	0	0
All	All	5533	0	4983	134	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 13.

All (134) 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)
3:A:712:ILE:H	3:A:712:ILE:HD13	1.36	0.86
3:A:364:GLU:HG2	3:A:388:CYS:HB3	1.58	0.85
3:A:590:GLY:HA2	3:A:593:LYS:HE2	1.59	0.84
3:A:527:ASN:H	3:A:533:GLN:NE2	1.78	0.81
3:A:729:ARG:H	3:A:729:ARG:NE	1.79	0.80
3:A:534:LEU:HD11	3:A:574:ILE:HD13	1.65	0.78
3:A:328:HIS:HD2	3:A:382:TRP:HE1	1.34	0.74
3:A:456:GLU:HG2	8:A:2809:HOH:O	1.89	0.71
2:C:8:DC:H2'	2:C:9:DA:C8	2.26	0.71
3:A:691:GLN:HG3	3:A:738:ARG:HH22	1.57	0.69
3:A:729:ARG:HE	3:A:729:ARG:H	1.39	0.69
3:A:431:LYS:HD2	3:A:431:LYS:N	2.08	0.68
3:A:675:PHE:CZ	3:A:681:ILE:HG21	2.29	0.68
3:A:804:ILE:HD11	8:A:2772:HOH:O	1.94	0.66
2:C:13:DC:H2''	2:C:14:DT:H72	1.78	0.66
3:A:339:ASN:HD22	3:A:339:ASN:C	1.97	0.65
3:A:779:ARG:NH1	3:A:779:ARG:HB2	2.11	0.65
3:A:718:ASP:HB2	3:A:733:ALA:HB2	1.79	0.65
3:A:737:GLU:O	3:A:741:GLU:HG3	1.97	0.64
1:B:22:DG:H2''	1:B:23:DA:H5'	1.77	0.64
3:A:712:ILE:HB	3:A:750:MET:HE1	1.78	0.64
3:A:624:GLN:HG2	3:A:828:VAL:HG12	1.80	0.63
3:A:677:ARG:HB2	3:A:679:LEU:HD13	1.80	0.62
3:A:728:SER:OG	3:A:731:GLU:HG3	1.99	0.62
3:A:589:GLU:O	3:A:593:LYS:HG3	1.99	0.61
3:A:707:ALA:HB3	3:A:725:LEU:HD21	1.82	0.61
3:A:681:ILE:HD12	3:A:682:HIS:N	2.16	0.61
3:A:754:VAL:HG12	3:A:758:LYS:HE2	1.83	0.60
3:A:657:ILE:HG21	8:A:2772:HOH:O	2.02	0.59
3:A:459:ARG:HB3	3:A:460:PRO:HD3	1.84	0.59
3:A:690:PHE:CD2	3:A:701:MET:HE2	2.37	0.59
3:A:505:LYS:NZ	3:A:505:LYS:HB3	2.18	0.59
3:A:813:ALA:O	3:A:817:GLU:HG3	2.02	0.59
2:C:6:DA:H2''	2:C:7:DC:H5'	1.84	0.58
3:A:547:LEU:C	3:A:548:LYS:HD2	2.24	0.58
3:A:339:ASN:ND2	3:A:341:HIS:H	2.02	0.57
3:A:729:ARG:H	3:A:729:ARG:CD	2.18	0.56

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:A:692:VAL:HB	3:A:696:GLU:HB2	1.87	0.56
3:A:675:PHE:CE1	3:A:681:ILE:HG21	2.41	0.55
3:A:733:ALA:O	3:A:736:ILE:HG12	2.06	0.55
3:A:550:THR:HG23	3:A:553:GLY:H	1.71	0.55
1:B:24:DT:OP1	3:A:552:THR:HG22	2.07	0.54
2:C:3:DC:H1'	3:A:716:ILE:HG21	1.90	0.54
3:A:498:LYS:O	3:A:502:GLN:HG3	2.08	0.54
2:C:9:DA:H2''	2:C:10:DT:H5'	1.90	0.53
3:A:519:TYR:CD1	3:A:525:GLU:HA	2.43	0.53
3:A:621:PRO:HG2	3:A:623:LEU:CD2	2.39	0.53
3:A:547:LEU:HB2	3:A:548:LYS:HD2	1.90	0.52
3:A:607:ASN:ND2	4:D:2:FRU:H4	2.24	0.52
3:A:725:LEU:O	3:A:727:ILE:HG23	2.10	0.52
3:A:517:ARG:O	3:A:520:GLU:HG2	2.09	0.52
3:A:685:THR:O	3:A:689:ILE:HG12	2.10	0.52
3:A:550:THR:HG22	3:A:553:GLY:O	2.10	0.51
2:C:6:DA:H2''	2:C:7:DC:C5'	2.40	0.51
3:A:779:ARG:HH11	3:A:779:ARG:HB2	1.74	0.51
3:A:677:ARG:CB	3:A:679:LEU:HD13	2.39	0.51
3:A:784:ARG:O	3:A:788:GLU:HG3	2.10	0.51
3:A:754:VAL:CG1	3:A:758:LYS:HE2	2.41	0.50
3:A:505:LYS:HZ3	3:A:505:LYS:HB3	1.75	0.50
3:A:630:LEU:HD12	3:A:630:LEU:N	2.27	0.50
1:B:28:DG:H2''	3:A:829:HIS:CE1	2.46	0.50
3:A:713:VAL:HG13	3:A:797:GLN:HE21	1.77	0.50
2:C:4:DG:H2''	2:C:5:DC:O5'	2.12	0.50
3:A:392:PHE:HA	3:A:479:GLN:HE22	1.76	0.49
3:A:527:ASN:H	3:A:533:GLN:HE22	1.54	0.49
3:A:510:GLN:O	3:A:514:VAL:HG23	2.11	0.49
2:C:13:DC:C2'	2:C:14:DT:H72	2.42	0.49
2:C:13:DC:H2''	2:C:14:DT:C7	2.43	0.49
3:A:681:ILE:HG13	8:A:2123:HOH:O	2.12	0.49
3:A:393:ASP:H	3:A:479:GLN:NE2	2.10	0.49
3:A:621:PRO:HG2	3:A:623:LEU:HD21	1.93	0.49
3:A:617:SER:HA	3:A:624:GLN:HE22	1.77	0.49
3:A:328:HIS:CD2	3:A:382:TRP:HE1	2.23	0.49
3:A:550:THR:CG2	3:A:553:GLY:H	2.25	0.49
3:A:607:ASN:HD22	3:A:609:ALA:H	1.61	0.48
2:C:7:DC:H2''	2:C:8:DC:O5'	2.13	0.48
3:A:583:LEU:HD13	8:A:2802:HOH:O	2.13	0.48
3:A:729:ARG:N	3:A:729:ARG:CD	2.76	0.48

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:A:629:ARG:HH11	3:A:629:ARG:HG2	1.79	0.48
3:A:736:ILE:HG13	3:A:737:GLU:N	2.29	0.48
3:A:550:THR:HG22	3:A:553:GLY:C	2.33	0.48
3:A:823:HIS:HD2	8:A:2159:HOH:O	1.97	0.47
3:A:748:ARG:HD2	8:A:2960:HOH:O	2.13	0.47
3:A:822:ALA:CB	3:A:836:ALA:HB2	2.44	0.47
3:A:456:GLU:HB3	8:A:3025:HOH:O	2.14	0.47
3:A:653:ASP:HB3	3:A:831:GLU:HB3	1.97	0.47
3:A:847:LEU:C	3:A:847:LEU:HD23	2.34	0.47
3:A:466:ARG:HG3	3:A:466:ARG:HH11	1.80	0.47
3:A:530:SER:OG	3:A:533:GLN:HG3	2.15	0.47
2:C:12:DC:H2"	2:C:13:DC:C5	2.50	0.47
3:A:663:ALA:HB2	3:A:671:LEU:HG	1.97	0.47
3:A:548:LYS:HD2	3:A:548:LYS:N	2.30	0.46
3:A:607:ASN:HD22	3:A:607:ASN:C	2.18	0.46
3:A:548:LYS:HD3	3:A:560:VAL:HG22	1.97	0.46
3:A:515:GLU:HG3	3:A:516:GLN:N	2.30	0.46
3:A:654:TYR:HB3	3:A:657:ILE:HB	1.98	0.45
3:A:716:ILE:HD12	3:A:721:LEU:N	2.32	0.45
3:A:536:VAL:O	3:A:540:GLU:HB2	2.18	0.44
3:A:693:SER:OG	3:A:696:GLU:HG3	2.17	0.44
3:A:700:ASN:O	3:A:704:GLN:HG2	2.17	0.44
3:A:431:LYS:N	3:A:431:LYS:CD	2.79	0.44
3:A:717:SER:N	3:A:789:ARG:HH11	2.16	0.44
3:A:524:GLN:HE21	3:A:541:LYS:NZ	2.15	0.43
3:A:565:ALA:HB3	3:A:566:PRO:HD3	2.00	0.43
3:A:719:TYR:O	3:A:722:ALA:HB3	2.18	0.43
3:A:819:ARG:HG2	8:A:2827:HOH:O	2.17	0.43
3:A:823:HIS:HE1	3:A:835:GLU:OE2	2.01	0.43
3:A:750:MET:SD	3:A:792:MET:HB3	2.59	0.43
3:A:305:ASP:HB2	3:A:306:ARG:HD2	2.00	0.43
3:A:725:LEU:HB2	3:A:727:ILE:HG12	2.00	0.43
3:A:557:SER:HB3	3:A:560:VAL:HG23	2.01	0.43
3:A:499:ARG:HG3	3:A:503:MET:HE3	2.01	0.43
3:A:584:GLN:O	3:A:589:GLU:HG3	2.19	0.42
3:A:607:ASN:ND2	3:A:609:ALA:H	2.17	0.42
3:A:848:VAL:HB	3:A:849:PRO:HD3	2.00	0.42
3:A:547:LEU:CB	3:A:548:LYS:HD2	2.49	0.42
3:A:710:PHE:HB3	3:A:714:TYR:CD2	2.54	0.42
3:A:715:GLY:O	3:A:789:ARG:HD2	2.19	0.42
3:A:827:GLN:HG2	3:A:832:LEU:HD23	2.01	0.42

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:A:624:GLN:CG	3:A:828:VAL:HG12	2.49	0.42
3:A:561:LEU:O	3:A:571:VAL:HG11	2.20	0.42
3:A:715:GLY:HA2	8:A:2798:HOH:O	2.20	0.41
3:A:356:GLN:H	3:A:356:GLN:NE2	2.19	0.41
3:A:517:ARG:HG2	3:A:517:ARG:HH11	1.85	0.41
3:A:746:VAL:O	3:A:750:MET:HG2	2.21	0.41
3:A:794:THR:HB	3:A:795:PRO:HD3	2.03	0.41
3:A:653:ASP:CB	3:A:831:GLU:HB3	2.51	0.41
2:C:3:DC:O3'	3:A:716:ILE:HG22	2.20	0.41
3:A:499:ARG:HG3	3:A:503:MET:CE	2.51	0.40
3:A:740:PHE:CD1	3:A:747:LYS:HB2	2.56	0.40
3:A:339:ASN:ND2	3:A:339:ASN:C	2.70	0.40
3:A:440:GLU:HB3	3:A:441:PRO:HD3	2.04	0.40
3:A:517:ARG:NE	3:A:521:LEU:HD11	2.36	0.40
3:A:681:ILE:CD1	3:A:682:HIS:CE1	3.05	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
3	A	578/580 (100%)	562 (97%)	14 (2%)	2 (0%)	41 31

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	A	432	GLY
3	A	628	ILE

### 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
3	A	495/496 (100%)	483 (98%)	12 (2%)	49 43

All (12) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	A	306	ARG
3	A	339	ASN
3	A	443	LEU
3	A	499	ARG
3	A	517	ARG
3	A	527	ASN
3	A	607	ASN
3	A	712	ILE
3	A	723	GLN
3	A	729	ARG
3	A	793	ASN
3	A	827	GLN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (20) such sidechains are listed below:

Mol	Chain	Res	Type
3	A	328	HIS
3	A	339	ASN
3	A	356	GLN
3	A	479	GLN
3	A	502	GLN
3	A	510	GLN
3	A	524	GLN
3	A	527	ASN
3	A	533	GLN
3	A	543	GLN
3	A	579	GLN
3	A	607	ASN
3	A	608	GLN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
3	A	624	GLN
3	A	656	GLN
3	A	723	GLN
3	A	724	ASN
3	A	797	GLN
3	A	823	HIS
3	A	827	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 ⓘ

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
4	GLC	D	1	4	11,11,12	3.41	4 (36%)	15,15,17	1.60	2 (13%)
4	FRU	D	2	4	11,12,12	1.71	2 (18%)	10,18,18	0.96	1 (10%)

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	GLC	D	1	4	-	0/2/19/22	0/1/1/1
4	FRU	D	2	4	-	4/5/24/24	0/1/1/1

All (6) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	D	1	GLC	C2-C3	9.90	1.67	1.52
4	D	2	FRU	O2-C2	4.28	1.48	1.40
4	D	2	FRU	C1-C2	3.19	1.57	1.52
4	D	1	GLC	O5-C1	3.03	1.48	1.43
4	D	1	GLC	O5-C5	2.73	1.49	1.43
4	D	1	GLC	C4-C5	2.33	1.57	1.53

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	D	1	GLC	C1-O5-C5	4.12	117.77	112.19
4	D	1	GLC	C1-C2-C3	-3.16	105.78	109.67
4	D	2	FRU	O2-C2-O5	-2.07	105.50	109.50

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	D	2	FRU	O5-C5-C6-O6
4	D	2	FRU	C4-C5-C6-O6
4	D	2	FRU	O1-C1-C2-C3
4	D	2	FRU	O1-C1-C2-O5

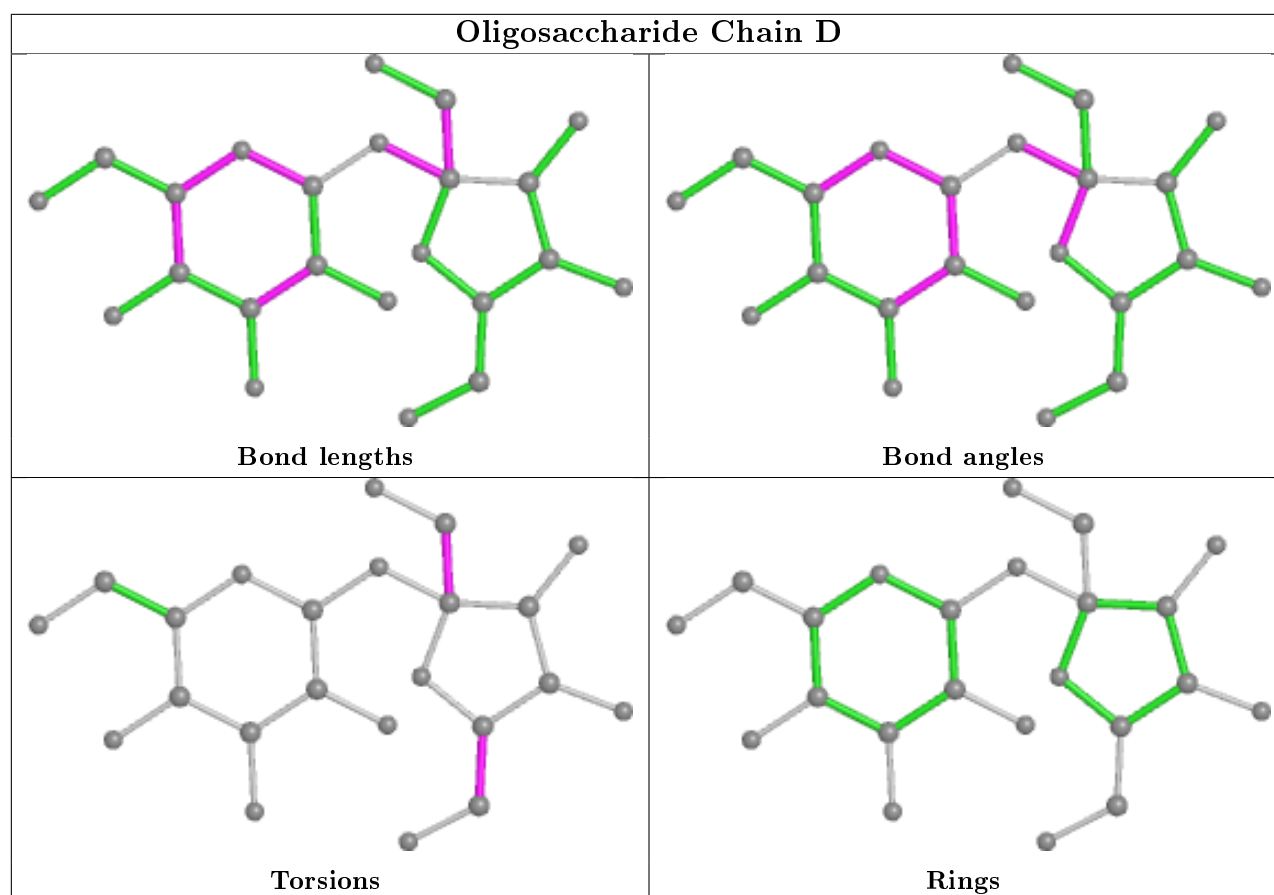
There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	D	2	FRU	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 6 ligands modelled in this entry, 1 is monoatomic - leaving 5 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	BAP	C	444	2	27,27,27	3.30	14 (51%)	39,42,42	1.53	5 (12%)
6	SO4	A	913	-	4,4,4	0.27	0	6,6,6	0.07	0
6	SO4	A	911	-	4,4,4	0.26	0	6,6,6	0.10	0
6	SO4	A	912	-	4,4,4	0.24	0	6,6,6	0.07	0
6	SO4	A	910	-	4,4,4	0.28	0	6,6,6	0.06	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
5	BAP	C	444	2	-	-	0/5/5/5

All (14) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	C	444	BAP	C6-C7	9.66	1.53	1.37
5	C	444	BAP	C8-C7	7.18	1.48	1.36
5	C	444	BAP	C4'-C6	5.73	1.58	1.51
5	C	444	BAP	C10-C9	4.03	1.46	1.35
5	C	444	BAP	C2-C3	3.90	1.45	1.36
5	C	444	BAP	C2-C1	3.82	1.45	1.36
5	C	444	BAP	C6-C5A	3.33	1.48	1.43
5	C	444	BAP	C5-C5A	2.61	1.47	1.42
5	C	444	BAP	C5-C4	2.49	1.42	1.35
5	C	444	BAP	C3B-C3A	2.43	1.47	1.42
5	C	444	BAP	C3B-C1A	2.15	1.47	1.42
5	C	444	BAP	C10-C1A	2.12	1.47	1.41
5	C	444	BAP	C5A-C5B	2.04	1.46	1.42
5	C	444	BAP	C9-C8A	2.01	1.46	1.41

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	C	444	BAP	O1'-C1'-C7	-6.66	97.36	110.49
5	C	444	BAP	C6-C7-C1'	-3.77	114.44	120.20
5	C	444	BAP	O1'-C1'-C2'	2.84	114.52	108.55
5	C	444	BAP	C8-C7-C6	-2.27	115.92	119.59
5	C	444	BAP	C6-C4'-C3'	2.25	116.54	111.19

There are no chirality outliers.

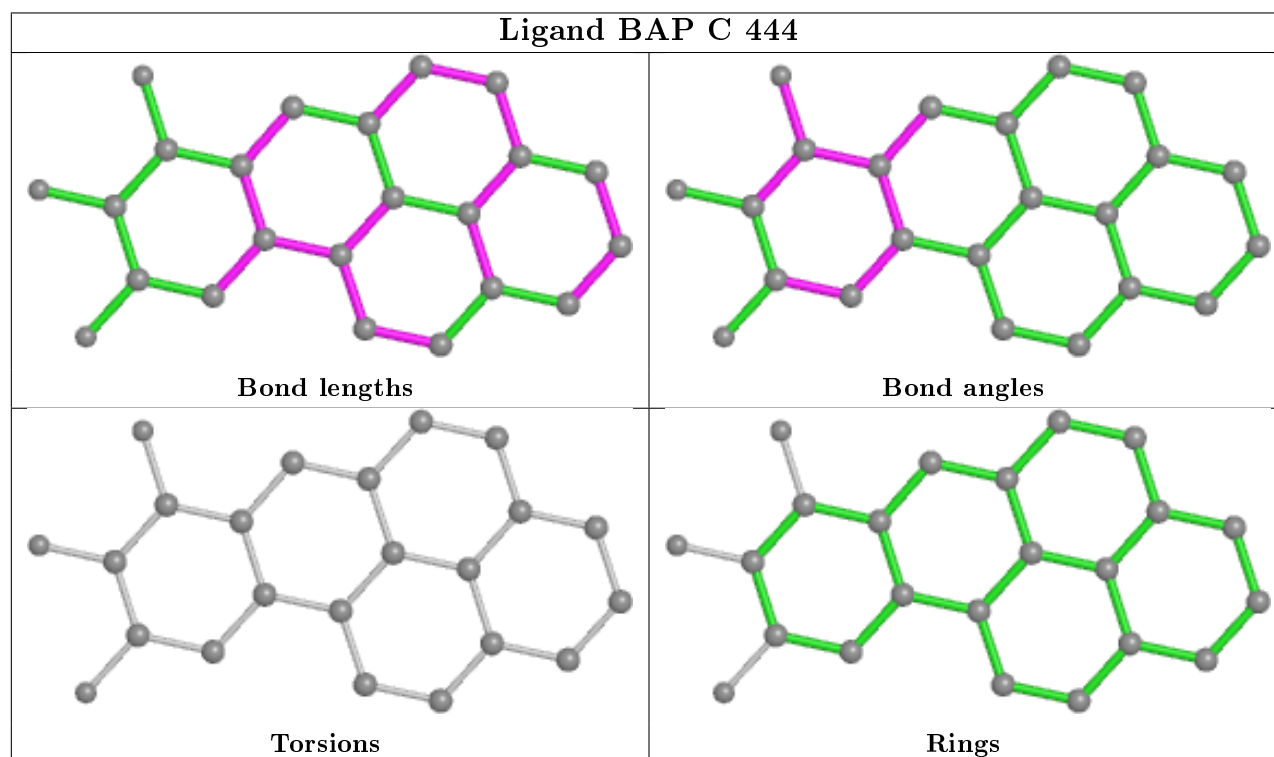
There are no torsion outliers.

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 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	B	10/10 (100%)	1.24	4 (40%) 0 0	38, 47, 82, 85	0
2	C	12/15 (80%)	1.34	3 (25%) 0 0	42, 59, 85, 92	0
3	A	580/580 (100%)	0.48	49 (8%) 11 12	16, 28, 57, 72	0
All	All	602/605 (99%)	0.51	56 (9%) 8 10	16, 29, 61, 92	0

All (56) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	A	719	TYR	7.9
3	A	297	ALA	7.4
3	A	729	ARG	6.1
3	A	552	THR	5.9
3	A	431	LYS	5.5
3	A	551	LYS	5.3
3	A	433	ALA	5.0
3	A	550	THR	5.0
3	A	434	LYS	4.9
2	C	14	DT	4.9
3	A	732	ALA	4.7
3	A	723	GLN	4.3
3	A	725	LEU	4.3
2	C	13	DC	4.0
3	A	548	LYS	3.8
3	A	553	GLY	3.7
3	A	629	ARG	3.4
3	A	531	PRO	3.4
3	A	532	LYS	3.4
3	A	736	ILE	3.3
1	B	22	DG	3.3
3	A	520	GLU	3.1
3	A	712	ILE	3.1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
3	A	513	THR	3.1
3	A	306	ARG	3.1
3	A	779	ARG	3.1
3	A	516	GLN	3.0
3	A	876	LYS	3.0
3	A	304	ALA	3.0
3	A	509	GLU	2.9
3	A	303	LEU	2.9
3	A	549	LYS	2.9
3	A	579	GLN	2.7
3	A	441	PRO	2.7
3	A	726	ASN	2.7
3	A	305	ASP	2.7
3	A	738	ARG	2.7
3	A	505	LYS	2.6
3	A	727	ILE	2.6
3	A	722	ALA	2.6
1	B	20	DG	2.6
3	A	733	ALA	2.5
1	B	21	DG	2.5
2	C	12	DC	2.4
3	A	512	GLY	2.4
3	A	710	PHE	2.4
3	A	724	ASN	2.4
1	B	23	DA	2.4
3	A	730	LYS	2.3
3	A	432	GLY	2.3
3	A	681	ILE	2.2
3	A	728	SER	2.1
3	A	439	ASP	2.1
3	A	731	GLU	2.0
3	A	689	ILE	2.0
3	A	721	LEU	2.0

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

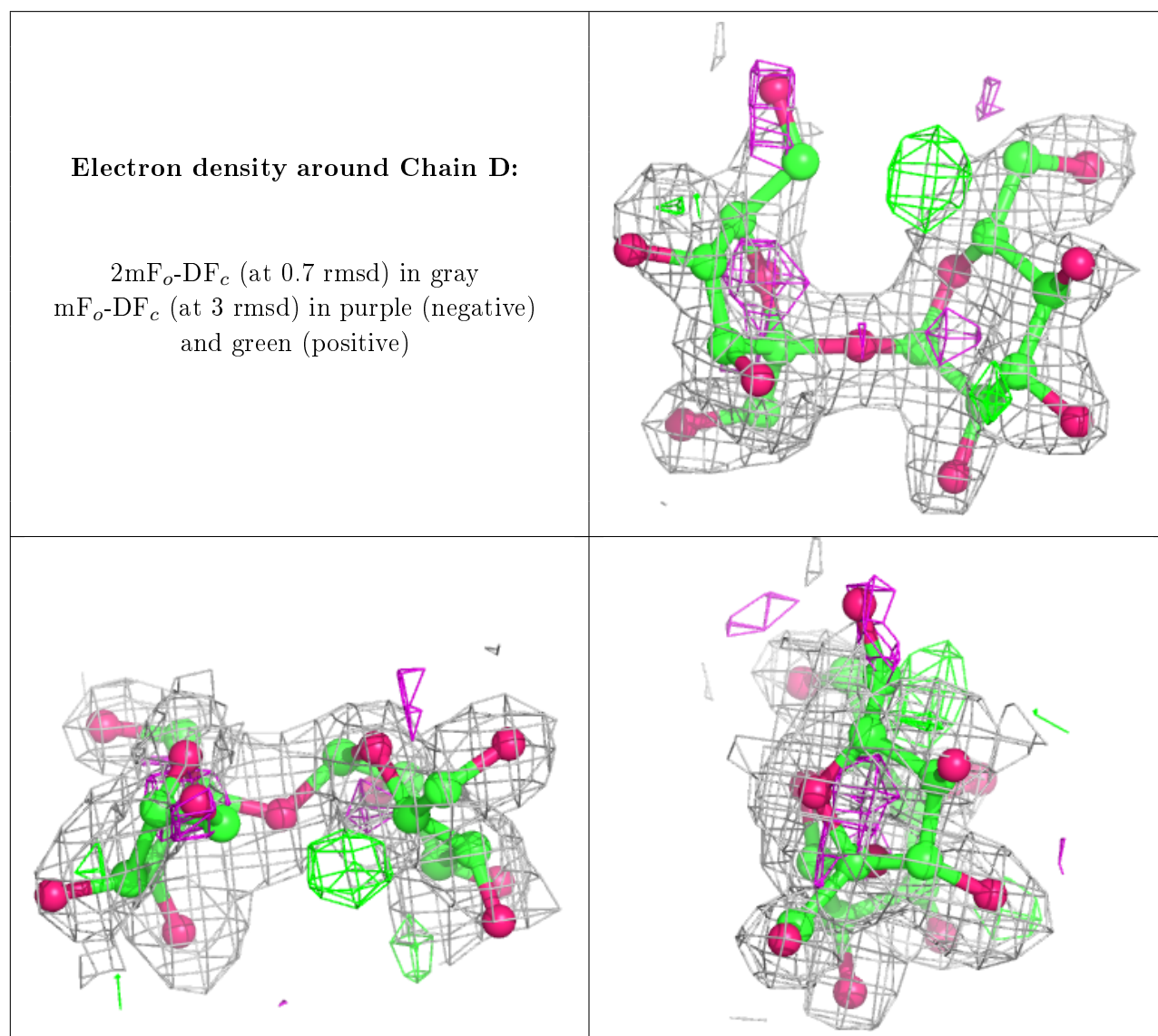
There are no non-standard protein/DNA/RNA residues in this entry.

### 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<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	FRU	D	2	12/12	0.71	0.28	50,51,53,54	0
4	GLC	D	1	11/12	0.78	0.21	48,50,50,51	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.

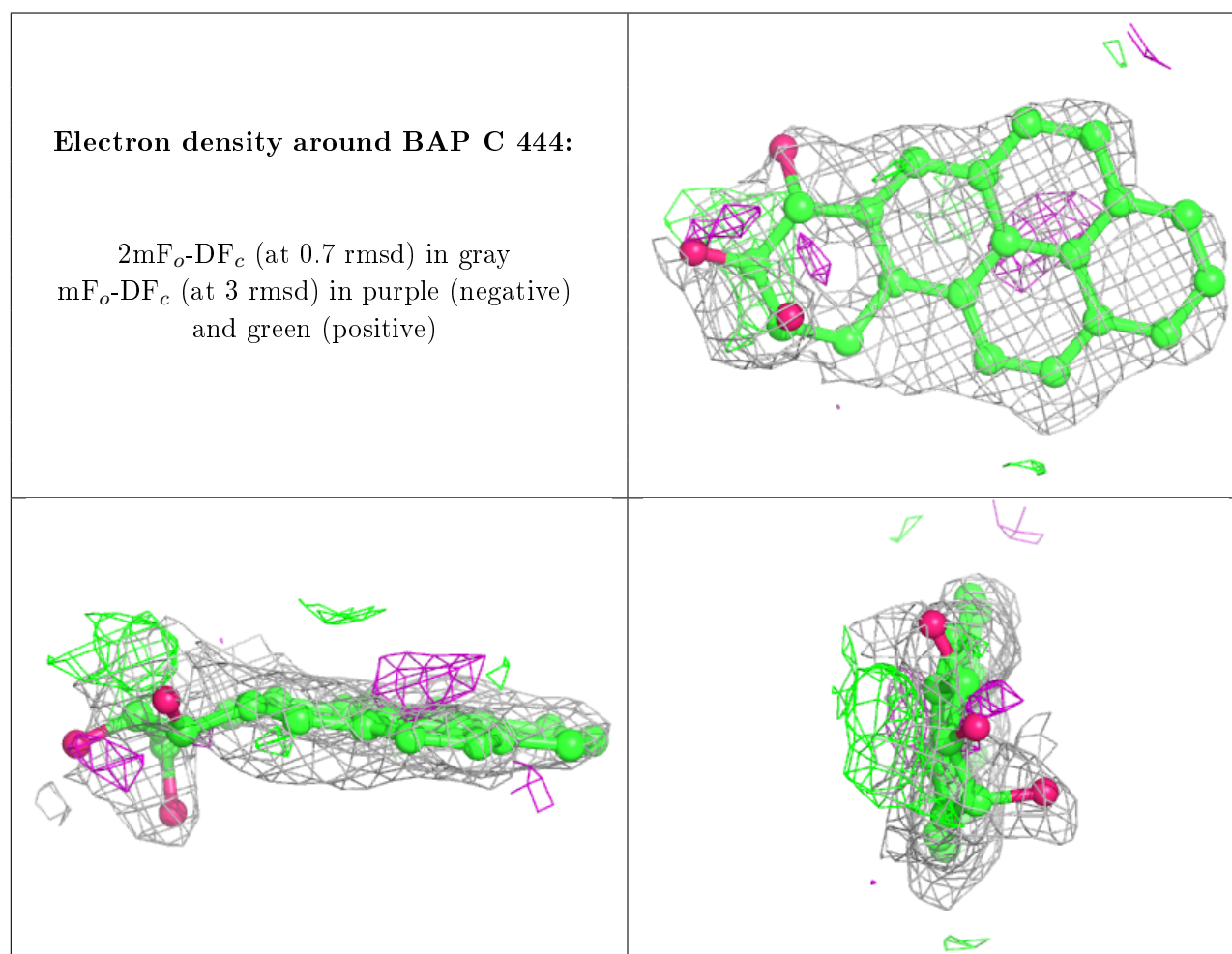


## 6.4 Ligands [i](#)

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<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
5	BAP	C	444	23/23	0.56	0.28	57,57,60,61	0
6	SO4	A	913	5/5	0.87	0.17	75,75,76,76	0
6	SO4	A	911	5/5	0.89	0.17	75,75,76,76	0
6	SO4	A	912	5/5	0.95	0.13	51,51,52,52	0
7	MG	A	920	1/1	0.96	0.06	49,49,49,49	0
6	SO4	A	910	5/5	0.96	0.09	62,62,62,62	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.



## 6.5 Other polymers [i](#)

There are no such residues in this entry.