



# Full wwPDB X-ray Structure Validation Report

Feb 27, 2014 – 02:04 AM GMT

PDB ID : 2Y1K  
Title : STRUCTURE OF HUMAN BUTYRYLCHOLINESTERASEINHIBITED  
BY CBDP (12H SOAK): PHOSPHOSERINE ADDUCT  
Authors : Carletti, E.; Colletier, J.P.; Nachon, F.; Weik, M.  
Deposited on : 2010-12-08  
Resolution : 2.50 Å(reported)

This is a full wwPDB 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 <http://wwpdb.org/ValidationPDFNotes.html>

---

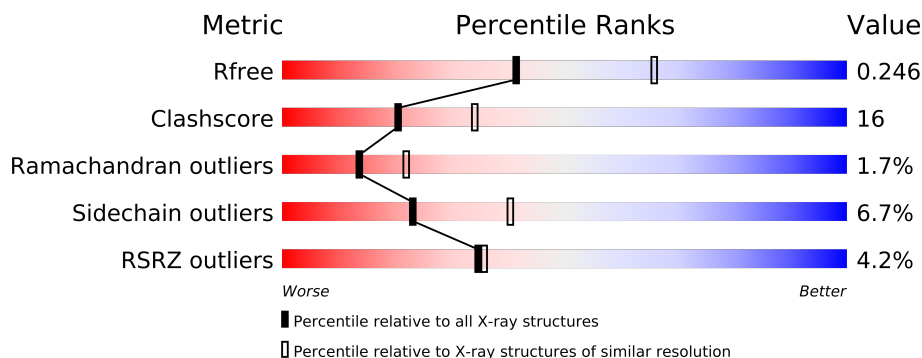
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.15 2013  
Xtriage (Phenix) : dev-1323  
EDS : stable22639  
Percentile statistics : 21963  
Refmac : 5.8.0049  
CCP4 : 6.3.0 (Settle)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)  
Validation Pipeline (wwPDB-VP) : stable22683

# 1 Overall quality at a glance

The reported resolution of this entry is 2.50 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))
$R_{free}$	66092	2784 (2.50-2.50)
Clashscore	79885	3562 (2.50-2.50)
Ramachandran outliers	78287	3480 (2.50-2.50)
Sidechain outliers	78261	3482 (2.50-2.50)
RSRZ outliers	66119	2785 (2.50-2.50)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	529	

The following table lists non-polymeric compounds that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Geometry	Electron density
2	SO4	A	1535	-	X
3	NA	A	1532	-	X
7	NAG	A	1547	-	X
8	FU4	A	1543	-	X
8	FU4	A	1549	-	X

## 2 Entry composition i

There are 9 unique types of molecules in this entry. The entry contains 4597 atoms, of which 0 are hydrogen and 0 are deuterium.

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 CHOLINESTERASE.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	P	S			
1	A	525	4193	2704	705	768	1	15	0	2	0

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	17	GLN	ASN	ENGINEERED MUTATION	UNP P06276
A	455	GLN	ASN	ENGINEERED MUTATION	UNP P06276
A	481	GLN	ASN	ENGINEERED MUTATION	UNP P06276

- Molecule 2 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



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

*Continued on next page...*

Continued from previous page...

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

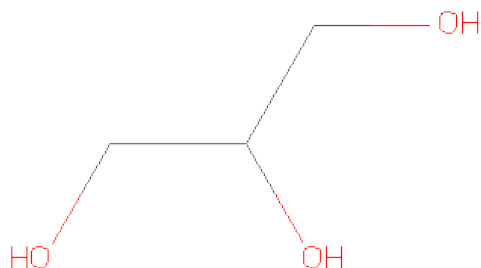
- Molecule 3 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	1	Total	Na	0	0
			1	1		

- Molecule 4 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

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

- Molecule 5 is GLYCEROL (three-letter code: GOL) (formula: C<sub>3</sub>H<sub>8</sub>O<sub>3</sub>).



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

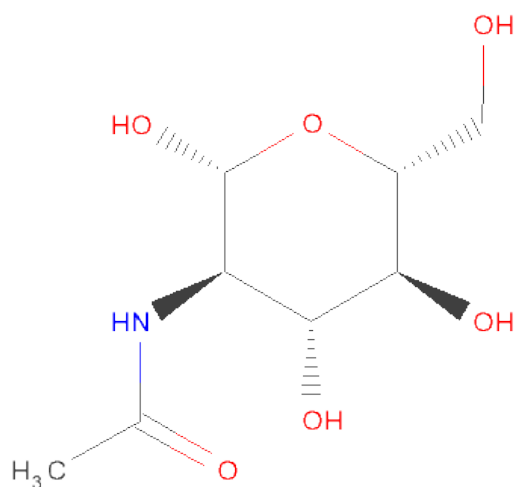
- Molecule 6 is a polymer of unknown type called SUGAR (2-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
6	A	2	Total	C	N	O	0	0
			24	14	1	9		

There are 3 discrepancies between the modelled and reference sequences:

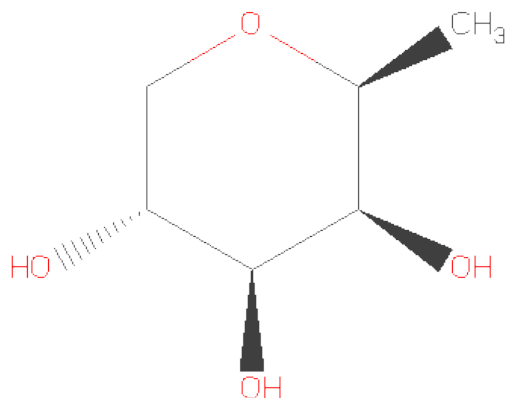
Chain	Residue	Modelled	Actual	Comment	Reference
A	17	GLN	ASN	ENGINEERED MUTATION	UNP P06276
A	455	GLN	ASN	ENGINEERED MUTATION	UNP P06276
A	481	GLN	ASN	ENGINEERED MUTATION	UNP P06276

- Molecule 7 is SUGAR (N-ACETYL-D-GLUCOSAMINE) (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



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

- Molecule 8 is SUGAR (2,6-ANHYDRO-1-DEOXY-D-GALACTITOL) (three-letter code: FU4) (formula:  $C_6H_{12}O_4$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
8	A	1	Total	C	O	0	0
			10	6	4		
8	A	1	Total	C	O	0	0
			10	6	4		

- Molecule 9 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	A	229	Total	O	0	0
			229	229		

### 3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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: CHOLINESTERASE



## 4 Data and refinement statistics

Property	Value	Source
Space group	I 4 2 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	153.77Å 153.77Å 127.60Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	41.38 – 2.50 41.38 – 2.50	Depositor EDS
% Data completeness (in resolution range)	94.8 (41.38-2.50) 94.8 (41.38-2.50)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	4.76 (at 2.51Å)	Xtriage
Refinement program	REFMAC 5.6.0093	Depositor
R, $R_{free}$	0.183 , 0.247 0.183 , 0.246	Depositor DCC
$R_{free}$ test set	760 reflections (3.09%)	DCC
Wilson B-factor (Å <sup>2</sup> )	52.5	Xtriage
Anisotropy	0.142	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 49.4	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Outliers	0 of 25337 reflections	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	4597	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	61.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.00% of the height of the origin peak. No significant pseudotranslation is detected.*

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



## 5 Model quality

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: GOL, FU4, NAG, SEP, NA, CL, FUC, SO4

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	0.90	1/4306 (0.0%)	0.86	4/5841 (0.1%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	66	CYS	CB-SG	-5.11	1.73	1.81

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	265	ARG	NE-CZ-NH2	-5.41	117.59	120.30
1	A	361	VAL	N-CA-C	5.39	125.55	111.00
1	A	265	ARG	NE-CZ-NH1	5.08	122.84	120.30
1	A	448	LEU	CA-CB-CG	5.07	126.96	115.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	156	LEU	Peptide

## 5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	4193	0	0	60	0
2	A	25	0	0	2	0
3	A	1	0	0	0	0
4	A	1	0	0	2	0
5	A	6	0	8	0	0
6	A	24	0	22	4	0
7	A	98	0	91	18	0
8	A	20	0	24	5	0
9	A	229	0	0	35	3
All	All	4597	0	145	71	3

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 16.

All (71) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:391:ASP:O	9:A:2144:HOH:O	1.52	1.26
7:A:1542:NAG:O6	8:A:1543:FU4:H1C2	1.46	1.15
1:A:395:ASP:N	9:A:2144:HOH:O	1.82	1.11
1:A:522:TRP:O	9:A:2214:HOH:O	1.66	1.10
6:A:1536:NAG:O4	7:A:1538:NAG:C1	2.00	1.09
1:A:475:ALA:CB	9:A:2179:HOH:O	1.98	1.09
1:A:347:ARG:NH1	9:A:2129:HOH:O	1.93	1.00
1:A:300:THR:CG2	9:A:2111:HOH:O	2.10	0.98
4:A:1533:CL:CL	9:A:2186:HOH:O	2.18	0.98
1:A:496:THR:N	9:A:2192:HOH:O	2.01	0.94
6:A:1536:NAG:HO4	7:A:1538:NAG:C1	1.77	0.93
1:A:522:TRP:C	9:A:2214:HOH:O	2.06	0.92
1:A:381:ARG:N	9:A:2138:HOH:O	2.01	0.92
1:A:514:LEU:O	9:A:2210:HOH:O	1.89	0.91
6:A:1536:NAG:O4	7:A:1538:NAG:O5	1.86	0.91
1:A:485:ASN:ND2	7:A:1545:NAG:H2	1.87	0.90
1:A:500:TYR:O	9:A:2199:HOH:O	1.89	0.90
1:A:495:SER:CA	9:A:2192:HOH:O	2.24	0.86

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:510:ILE:O	9:A:2204:HOH:O	1.96	0.81
1:A:256:ASN:ND2	7:A:1546:NAG:C1	2.45	0.80
1:A:527:PRO:CD	9:A:2214:HOH:O	2.29	0.79
7:A:1544:NAG:H61	9:A:2224:HOH:O	1.85	0.77
1:A:172:GLN:NE2	9:A:2081:HOH:O	2.17	0.76
7:A:1542:NAG:O6	8:A:1543:FU4:C1	2.33	0.74
7:A:1542:NAG:HO6	8:A:1543:FU4:H1C2	1.52	0.73
1:A:241:ASN:ND2	7:A:1547:NAG:H4	2.06	0.71
1:A:372:HIS:ND1	9:A:2134:HOH:O	2.26	0.68
1:A:254:ARG:NE	9:A:2101:HOH:O	2.27	0.66
1:A:106:ASN:CG	7:A:1542:NAG:H83	2.16	0.66
1:A:485:ASN:ND2	7:A:1545:NAG:H62	2.12	0.65
1:A:360:GLY:N	9:A:2132:HOH:O	2.28	0.65
1:A:157:PRO:O	9:A:2078:HOH:O	2.14	0.65
1:A:311:GLN:NE2	9:A:2116:HOH:O	2.32	0.63
1:A:138:ARG:NH2	9:A:2071:HOH:O	2.32	0.63
1:A:361:VAL:O	1:A:366:LYS:NZ	2.33	0.61
1:A:106:ASN:ND2	7:A:1542:NAG:H83	2.16	0.60
7:A:1548:NAG:H5	8:A:1549:FU4:O2	2.03	0.59
1:A:452:ARG:NE	9:A:2170:HOH:O	2.37	0.58
1:A:485:ASN:ND2	7:A:1545:NAG:C2	2.65	0.57
1:A:240:ARG:CG	9:A:2076:HOH:O	2.53	0.57
6:A:1536:NAG:O3	7:A:1538:NAG:H61	2.04	0.56
1:A:265:ARG:NE	9:A:2102:HOH:O	2.37	0.56
1:A:501:LEU:CA	9:A:2199:HOH:O	2.53	0.56
1:A:188:ASN:ND2	9:A:2088:HOH:O	2.40	0.55
7:A:1542:NAG:C6	8:A:1543:FU4:H1C2	2.36	0.53
1:A:509:ARG:CG	1:A:509:ARG:NH1	2.70	0.53
1:A:106:ASN:ND2	7:A:1542:NAG:C1	2.72	0.52
1:A:347:ARG:NH2	2:A:1531:SO4:O4	2.42	0.52
1:A:471:TRP:O	9:A:2179:HOH:O	2.20	0.50
1:A:508:THR:CG2	4:A:1533:CL:CL	2.97	0.49
1:A:424:ARG:NH1	1:A:428:LEU:CD1	2.75	0.49
1:A:482:GLU:OE1	1:A:487:SER:OG	2.31	0.49
1:A:425:SER:N	9:A:2161:HOH:O	2.45	0.48
1:A:455:GLN:N	9:A:2173:HOH:O	2.46	0.48
1:A:247:ALA:O	1:A:252:CYS:N	2.47	0.47
1:A:420:TYR:OH	2:A:1540:SO4:O3	2.33	0.47
1:A:227:PHE:CD1	1:A:227:PHE:C	2.88	0.46
1:A:377:VAL:N	9:A:2137:HOH:O	2.48	0.45
1:A:395:ASP:CB	9:A:2144:HOH:O	2.65	0.44
1:A:448:LEU:N	1:A:449:PRO:CD	2.80	0.44

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:240:ARG:CA	9:A:2076:HOH:O	2.66	0.44
1:A:474:PHE:O	1:A:478:GLY:N	2.51	0.44
1:A:50:THR:O	1:A:51:LYS:CB	2.65	0.44
1:A:495:SER:O	1:A:496:THR:OG1	2.35	0.43
1:A:198:SEP:C	1:A:200:GLY:N	2.79	0.43
1:A:519:CYS:O	1:A:523:THR:N	2.52	0.42
1:A:427:LYS:O	1:A:428:LEU:C	2.56	0.42
1:A:268:ASP:O	1:A:269:PRO:C	2.59	0.41
1:A:518:GLN:NE2	1:A:518:GLN:N	2.69	0.41
1:A:482:GLU:N	9:A:2183:HOH:O	2.54	0.40
1:A:320:GLY:CA	1:A:419:TYR:CE2	3.04	0.40

All (3) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
9:A:2089:HOH:O	9:A:2091:HOH:O[7_556]	1.50	0.70
9:A:2204:HOH:O	9:A:2208:HOH:O[7_556]	1.56	0.64
9:A:2198:HOH:O	9:A:2207:HOH:O[7_556]	2.11	0.09

## 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	522/529 (99%)	480 (92%)	33 (6%)	9 (2%)	14 22

All (9) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	51	LYS
1	A	281	PRO
1	A	361	VAL
1	A	496	THR
1	A	4	ILE
1	A	50	THR

Continued on next page...

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	52	TRP
1	A	67	GLN
1	A	360	GLY

### 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	449/453 (99%)	418 (93%)	31 (7%)	22	39

All (31) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	4	ILE
1	A	9	LYS
1	A	16	MET
1	A	50	THR
1	A	51	LYS
1	A	72	SER
1	A	110	LEU
1	A	144	MET
1	A	176	GLN
1	A	195	PHE
1	A	236	LEU
1	A	255	GLU
1	A	262	LYS
1	A	268	ASP
1	A	278	PHE
1	A	286	LEU
1	A	299	LEU
1	A	305	ILE
1	A	348	LYS
1	A	363[A]	GLU
1	A	363[B]	GLU
1	A	375	ASP
1	A	376	TRP
1	A	383	GLU
1	A	407	LYS

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	428	LEU
1	A	453	ARG
1	A	497	GLU
1	A	506	GLU
1	A	508	THR
1	A	518	GLN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 5.3.3 RNA ⓘ

There are no RNA chains in this entry.

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

1 non-standard protein/DNA/RNA residue is 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
1	SEP	A	198	1	9,9,10	6.47	4 (44%)	10,12,14	2.06	5 (50%)

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
1	SEP	A	198	1	-	0/6/8/10	0/0/0/0

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	198	SEP	O-C	17.06	1.23	1.11
1	A	198	SEP	P-OG	7.56	1.86	1.60
1	A	198	SEP	P-O2P	3.53	1.67	1.54
1	A	198	SEP	CA-C	3.18	1.54	1.48

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	198	SEP	P-OG-CB	-3.69	107.51	118.19
1	A	198	SEP	C-CA-N	-3.05	110.78	113.83
1	A	198	SEP	O3P-P-O1P	-2.14	103.43	110.44
1	A	198	SEP	O3P-P-O2P	2.11	115.84	107.61
1	A	198	SEP	OG-P-O1P	-2.02	100.77	106.71

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

## 5.5 Carbohydrates i

2 carbohydrates 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	$\# Z  > 2$
6	NAG	A	1536	1,6	12,14,15	1.14	2 (16%)	15,19,21	2.44	6 (40%)
6	FUC	A	1539	6	9,10,11	0.92	0	10,14,16	1.81	2 (20%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
6	NAG	A	1536	1,6	-	0/6/23/26	0/1/1/1

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
6	FUC	A	1539	6	-	0/0/17/20	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	A	1536	NAG	C2-N2	-2.25	1.43	1.46
6	A	1536	NAG	O5-C5	-2.19	1.41	1.45

All (8) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	A	1536	NAG	O5-C5-C6	5.01	112.24	106.98
6	A	1539	FUC	C4-C3-C2	4.65	116.75	110.50
6	A	1536	NAG	O5-C5-C4	-4.31	105.19	110.65
6	A	1536	NAG	C8-C7-N2	3.07	122.10	116.11
6	A	1536	NAG	C2-N2-C7	2.93	128.00	123.09
6	A	1536	NAG	C6-C5-C4	2.72	119.56	113.00
6	A	1539	FUC	O5-C5-C6	2.25	111.65	108.03
6	A	1536	NAG	C3-C4-C5	-2.09	106.48	110.20

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

## 5.6 Ligand geometry ⓘ

Of 17 ligands modelled in this entry, 2 are monoatomic - leaving 15 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	$\# Z  > 2$
2	SO4	A	1531	-	4,4,4	0.69	0	6,6,6	0.43	0
5	GOL	A	1534	-	5,5,5	0.74	0	5,5,5	1.09	0
2	SO4	A	1535	-	4,4,4	0.92	0	6,6,6	3.21	4 (66%)



Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
2	SO4	A	1537	-	4,4,4	0.72	0	6,6,6	0.35	0
7	NAG	A	1538	-	12,14,15	0.75	0	15,19,21	1.62	1 (6%)
2	SO4	A	1540	-	4,4,4	0.85	0	6,6,6	0.39	0
2	SO4	A	1541	-	4,4,4	0.94	0	6,6,6	0.52	0
7	NAG	A	1542	-	12,14,15	0.62	0	15,19,21	1.09	2 (13%)
8	FU4	A	1543	-	10,10,10	0.70	0	14,14,14	1.10	2 (14%)
7	NAG	A	1544	-	12,14,15	0.63	0	15,19,21	1.37	3 (20%)
7	NAG	A	1545	-	12,14,15	0.77	0	15,19,21	4.11	8 (53%)
7	NAG	A	1546	-	12,14,15	0.87	0	15,19,21	1.53	4 (26%)
7	NAG	A	1547	-	12,14,15	0.72	0	15,19,21	1.58	5 (33%)
7	NAG	A	1548	-	12,14,15	0.65	0	15,19,21	1.71	4 (26%)
8	FU4	A	1549	-	10,10,10	0.78	0	14,14,14	1.38	1 (7%)

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	SO4	A	1531	-	-	0/0/0/0	0/0/0/0
5	GOL	A	1534	-	-	0/4/4/4	0/0/0/0
2	SO4	A	1535	-	-	0/0/0/0	0/0/0/0
2	SO4	A	1537	-	-	0/0/0/0	0/0/0/0
7	NAG	A	1538	-	-	0/6/23/26	0/1/1/1
2	SO4	A	1540	-	-	0/0/0/0	0/0/0/0
2	SO4	A	1541	-	-	0/0/0/0	0/0/0/0
7	NAG	A	1542	-	-	0/6/23/26	0/1/1/1
8	FU4	A	1543	-	-	0/0/17/17	0/1/1/1
7	NAG	A	1544	-	-	0/6/23/26	0/1/1/1
7	NAG	A	1545	-	-	0/6/23/26	0/1/1/1
7	NAG	A	1546	-	-	0/6/23/26	0/1/1/1
7	NAG	A	1547	-	-	0/6/23/26	0/1/1/1
7	NAG	A	1548	-	-	0/6/23/26	0/1/1/1
8	FU4	A	1549	-	-	0/0/17/17	0/1/1/1

There are no bond length outliers.

All (34) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
7	A	1545	NAG	C2-N2-C7	9.75	139.45	123.09

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
7	A	1545	NAG	C3-C2-N2	8.12	124.13	111.76
2	A	1535	SO4	O2-S-O1	-6.63	87.60	109.53
7	A	1545	NAG	O5-C5-C6	5.38	112.63	106.98
7	A	1545	NAG	O7-C7-N2	5.32	133.00	121.90
7	A	1538	NAG	O5-C5-C6	5.12	112.35	106.98
7	A	1548	NAG	C3-C2-N2	-3.87	105.86	111.76
7	A	1548	NAG	O5-C5-C6	3.71	110.87	106.98
7	A	1545	NAG	C8-C7-N2	-3.24	109.78	116.11
8	A	1549	FU4	O5-C5-C6	3.08	111.55	106.20
7	A	1546	NAG	C3-C4-C5	2.74	115.10	110.20
8	A	1543	FU4	O5-C5-C6	2.68	110.87	106.20
2	A	1535	SO4	O4-S-O1	-2.63	86.23	110.12
7	A	1546	NAG	C4-C3-C2	2.57	117.61	111.32
7	A	1547	NAG	O3-C3-C4	-2.51	104.73	110.35
7	A	1545	NAG	O7-C7-C8	-2.49	117.19	122.04
7	A	1544	NAG	O7-C7-C8	-2.48	117.21	122.04
7	A	1547	NAG	C2-N2-C7	2.44	127.19	123.09
7	A	1542	NAG	O7-C7-C8	-2.40	117.36	122.04
7	A	1546	NAG	O5-C5-C6	2.29	109.38	106.98
7	A	1544	NAG	C8-C7-N2	2.26	120.52	116.11
7	A	1547	NAG	C3-C4-C5	2.23	114.19	110.20
7	A	1544	NAG	C3-C2-N2	2.23	115.16	111.76
7	A	1546	NAG	C2-N2-C7	2.18	126.75	123.09
7	A	1545	NAG	C3-C4-C5	2.16	114.07	110.20
7	A	1548	NAG	C3-C4-C5	2.16	114.06	110.20
7	A	1548	NAG	C4-C3-C2	2.15	116.59	111.32
7	A	1542	NAG	C8-C7-N2	2.11	120.23	116.11
7	A	1547	NAG	O5-C5-C4	2.08	113.29	110.65
8	A	1543	FU4	C3-C4-C5	2.06	113.27	109.84
7	A	1545	NAG	O3-C3-C2	2.05	113.38	109.09
2	A	1535	SO4	O4-S-O3	2.05	117.72	109.08
7	A	1547	NAG	C3-C2-N2	2.02	114.84	111.76
2	A	1535	SO4	O3-S-O1	-2.00	91.92	110.12

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

## 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<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	525/529 (99%)	0.18	20 (3%) 38 40	33, 56, 93, 147	10 (1%)

All (20) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	3	ASP	5.2
1	A	376	TRP	4.6
1	A	237	TYR	4.0
1	A	40	ARG	3.8
1	A	259	GLU	3.7
1	A	4	ILE	3.6
1	A	377	VAL	3.6
1	A	261	ILE	3.5
1	A	255	GLU	3.5
1	A	258	THR	3.3
1	A	55	ILE	3.3
1	A	256	ASN	3.3
1	A	17	GLN	3.0
1	A	381	ARG	3.0
1	A	14	ARG	2.5
1	A	240	ARG	2.5
1	A	22	GLY	2.3
1	A	385	TYR	2.2
1	A	260	ILE	2.2
1	A	278	PHE	2.0

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

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains.

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	RSR	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
1	SEP	A	198	10/11	0.20	0.11	36,40,42,52	0

### 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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	RSR	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
6	FUC	A	1539	10/11	0.27	5.07	74,83,92,99	0
6	NAG	A	1536	14/15	0.13	-0.64	47,57,68,68	0

### 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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	RSR	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
3	NA	A	1532	1/1	0.68	21.88	70,70,70,70	1
8	FU4	A	1543	10/10	0.40	8.87	36,46,58,76	10
2	SO4	A	1535	5/5	0.20	7.26	57,60,66,76	5
8	FU4	A	1549	10/10	0.51	4.98	64,99,120,122	0
7	NAG	A	1547	14/15	0.53	2.99	81,133,148,149	0
2	SO4	A	1540	5/5	0.17	1.96	61,68,72,95	5
2	SO4	A	1541	5/5	0.17	1.82	56,58,71,82	5
5	GOL	A	1534	6/6	0.20	0.86	54,60,66,79	0
7	NAG	A	1545	14/15	0.17	0.65	51,89,107,108	0
7	NAG	A	1544	14/15	0.34	0.47	64,88,99,121	0
2	SO4	A	1531	5/5	0.21	0.20	80,83,111,112	0
7	NAG	A	1542	14/15	0.16	0.01	50,87,99,102	0
7	NAG	A	1546	14/15	0.37	-0.60	75,97,112,118	0
2	SO4	A	1537	5/5	0.12	-1.74	78,82,91,97	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
4	CL	A	1533	1/1	0.10	-2.60	73,73,73,73	0
7	NAG	A	1538	14/15	0.19	-	76,94,113,114	0
7	NAG	A	1548	14/15	0.54	-	89,106,135,135	0

## 6.5 Other polymers ⓘ

There are no such residues in this entry.