



# Full wwPDB X-ray Structure Validation Report

Feb 28, 2014 – 04:41 AM GMT

PDB ID : 1J06  
Title : Crystal structure of mouse acetylcholinesterase in the apo form  
Authors : Bourne, Y.; Taylor, P.; Radic, Z.; Marchot, P.  
Deposited on : 2002-11-07  
Resolution : 2.35 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>

---

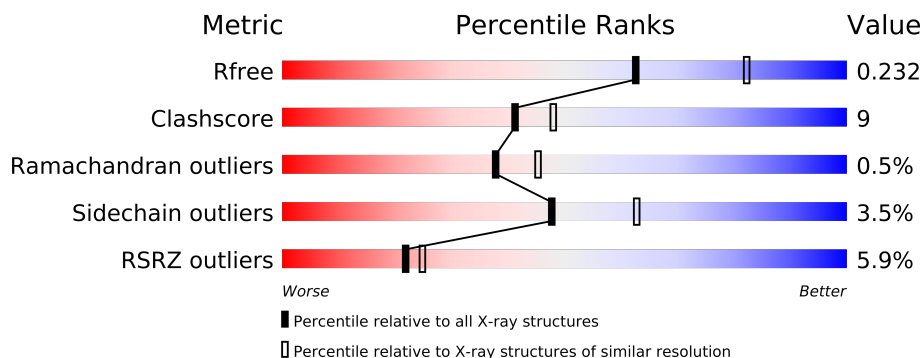
The following versions of software and data (see [references](#)) were used in the production of this report:

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 2.35 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	3327 (2.40-2.32)
Clashscore	79885	1064 (2.38-2.34)
Ramachandran outliers	78287	1048 (2.38-2.34)
Sidechain outliers	78261	1049 (2.38-2.34)
RSRZ outliers	66119	3330 (2.40-2.32)

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	543	
1	B	543	

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

Mol	Type	Chain	Res	Geometry	Electron density
3	NAG	A	701	-	X
3	NAG	B	601	-	X
4	CO3	B	952	-	X
5	P6G	A	901	-	X
6	AE3	A	902	-	X
6	AE3	B	903	-	X
7	PG4	B	904	-	X

## 2 Entry composition i

There are 8 unique types of molecules in this entry. The entry contains 8804 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.

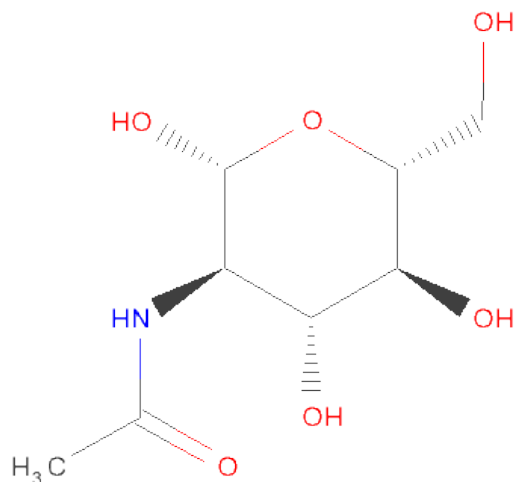
- Molecule 1 is a protein called acetylcholinesterase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	535	Total	C	N	O	S	0	0	0
			4177	2679	725	759	14			
1	B	533	Total	C	N	O	S	0	0	0
			4159	2670	718	757	14			

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

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

- Molecule 3 is SUGAR (N-ACETYL-D-GLUCOSAMINE) (three-letter code: NAG) (formula: C<sub>8</sub>H<sub>15</sub>NO<sub>6</sub>).



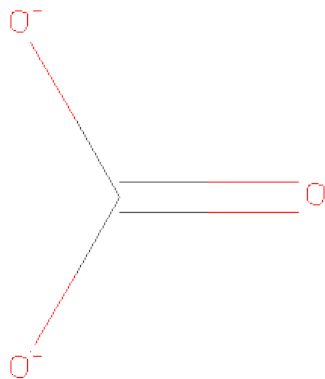
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	B	1	Total	C	N	O	0	0
			14	8	1	5		

*Continued on next page...*

*Continued from previous page...*

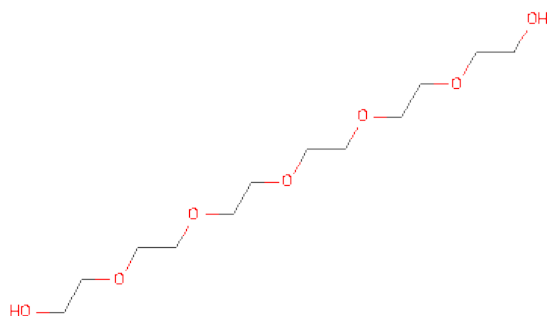
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 4 is CARBONATE ION (three-letter code: CO3) (formula: CO<sub>3</sub>).



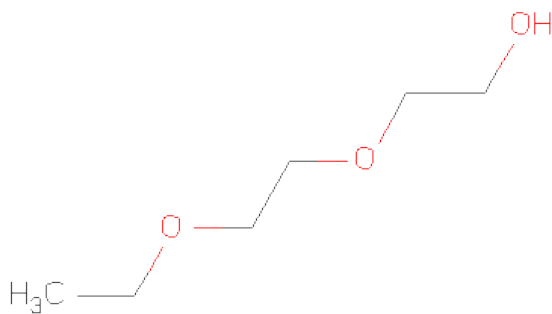
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			4	1	3		
4	B	1	Total	C	O	0	0
			4	1	3		

- Molecule 5 is HEXAETHYLENE GLYCOL (three-letter code: P6G) (formula: C<sub>12</sub>H<sub>26</sub>O<sub>7</sub>).



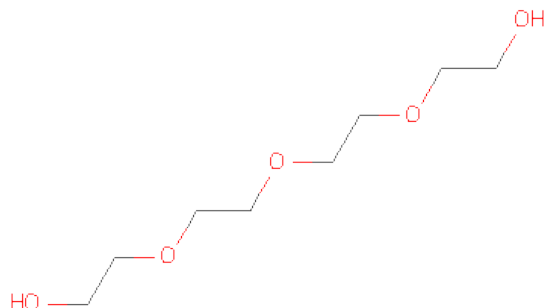
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	C	O	0	0
			19	12	7		

- Molecule 6 is 2-(2-ethoxyethoxy)ETHANOL (three-letter code: AE3) (formula:  $C_6H_{14}O_3$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	A	1	Total	C	O	0	0
			9	6	3		
6	B	1	Total	C	O	0	0
			9	6	3		

- Molecule 7 is TETRAETHYLENE GLYCOL (three-letter code: PG4) (formula:  $C_8H_{18}O_5$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	B	1	Total	C	O	0	0
			11	8	3		

- Molecule 8 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	A	210	Total	O	0	0
			210	210		
8	B	150	Total	O	0	0
			150	150		



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	79.20Å 111.97Å 226.83Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 2.35 39.84 – 2.35	Depositor EDS
% Data completeness (in resolution range)	(Not available) (20.00-2.35) 99.0 (39.84-2.35)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.06	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	4.25 (at 2.34Å)	Xtriage
Refinement program	CNS 1.0	Depositor
R, $R_{free}$	0.209 , 0.232 0.207 , 0.232	Depositor DCC
$R_{free}$ test set	1692 reflections (2.06%)	DCC
Wilson B-factor (Å <sup>2</sup> )	41.9	Xtriage
Anisotropy	0.847	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 41.8	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.51$ , $\langle L^2 \rangle = 0.35$	Xtriage
Outliers	0 of 83956 reflections	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	8804	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	49.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.88% 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: AE3, NAG, CO3, PG4, FUC, P6G

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	0.69	0/4300	0.80	0/5875
1	B	0.61	0/4282	0.78	1/5851 (0.0%)
All	All	0.65	0/8582	0.79	1/11726 (0.0%)

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
1	B	0	1
All	All	0	2

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
1	B	161	LEU	N-CA-C	-5.82	95.28	111.00

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	449	TYR	Sidechain
1	B	70	TYR	Sidechain

## 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	4177	0	4063	72	0
1	B	4159	0	4046	85	0
2	A	24	0	22	0	0
3	A	14	0	13	0	0
3	B	14	0	13	0	0
4	A	4	0	0	1	0
4	B	4	0	0	1	0
5	A	19	0	26	1	0
6	A	9	0	14	0	0
6	B	9	0	14	2	0
7	B	11	0	12	0	0
8	A	210	0	0	6	0
8	B	150	0	0	2	0
All	All	8804	0	8223	154	0

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 9.

All (154) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:48:MET:HE3	1:A:166:GLU:HA	1.29	1.09
1:B:48:MET:HE1	1:B:166:GLU:HA	1.32	1.09
1:A:197:VAL:H	1:A:223:HIS:HD2	1.14	0.94
1:B:197:VAL:H	1:B:223:HIS:HD2	1.03	0.93
1:A:485:ARG:HG3	8:A:1107:HOH:O	1.70	0.92
1:B:497:SER:HB2	1:B:498:PRO:O	1.72	0.90
1:A:4:GLU:OE2	1:A:18:ARG:HD3	1.73	0.88
1:B:245:ARG:O	1:B:249:THR:HG23	1.80	0.82
1:B:197:VAL:H	1:B:223:HIS:CD2	1.94	0.80
1:B:497:SER:HB2	1:B:498:PRO:C	2.02	0.79
1:A:197:VAL:H	1:A:223:HIS:CD2	1.99	0.78
1:A:245:ARG:O	1:A:249:THR:HG23	1.85	0.77
1:A:113:PRO:HG2	1:A:485:ARG:HG2	1.73	0.70
1:B:424:ARG:HG3	1:B:424:ARG:NH1	2.06	0.70

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:B:424:ARG:HH11	1:B:424:ARG:HG3	1.55	0.70
1:A:203:SER:OG	4:A:951:CO3:C	2.39	0.70
1:B:112:THR:HG21	1:B:143:GLY:O	1.92	0.69
1:A:48:MET:CE	1:A:166:GLU:HA	2.18	0.66
1:B:203:SER:OG	4:B:952:CO3:C	2.43	0.66
1:B:243:GLU:OE1	1:B:246:ARG:NH1	2.29	0.65
1:B:161:LEU:HD11	1:B:269:LEU:HD22	1.79	0.64
1:A:243:GLU:OE1	1:A:246:ARG:NH1	2.31	0.64
1:B:48:MET:CE	1:B:166:GLU:HA	2.18	0.62
1:A:376:GLU:HG2	1:B:538:LYS:NZ	2.15	0.61
1:A:376:GLU:HG2	1:B:538:LYS:CE	2.29	0.61
1:B:460:ASP:HB3	1:B:463:LEU:HD12	1.83	0.61
1:A:161:LEU:HD12	1:A:270:ILE:HD11	1.83	0.60
1:A:115:LEU:HD21	1:A:484:ALA:HB2	1.83	0.60
1:B:458:PRO:HA	1:B:465:TYR:CD2	2.37	0.60
1:A:88:PRO:HD3	8:A:1062:HOH:O	2.01	0.59
1:A:433:ARG:HD3	8:A:1080:HOH:O	2.02	0.59
1:B:353:LEU:HB3	1:B:391:PRO:HB2	1.84	0.58
1:A:161:LEU:HD12	1:A:270:ILE:CG1	2.32	0.58
1:A:112:THR:HG21	1:A:143:GLY:O	2.03	0.58
1:B:374:ALA:HB2	1:B:540:LEU:HD21	1.86	0.57
1:B:11:ARG:NH1	1:B:16:GLN:HG2	2.19	0.57
1:B:495:SER:O	1:B:496:LYS:O	2.23	0.57
1:A:13:ARG:NH2	8:A:1029:HOH:O	2.38	0.57
1:A:537:PRO:O	1:A:540:LEU:HB3	2.04	0.57
1:A:414:LEU:HG	1:A:418:LEU:HD22	1.86	0.56
1:A:161:LEU:HD11	1:A:269:LEU:HD22	1.88	0.56
1:B:11:ARG:HH12	1:B:16:GLN:CG	2.20	0.54
1:B:161:LEU:HD12	1:B:270:ILE:CG1	2.38	0.54
1:A:245:ARG:O	1:A:249:THR:CG2	2.55	0.54
1:B:437:LEU:HD11	1:B:449:TYR:CD2	2.43	0.53
1:B:80:PHE:O	1:B:84:GLU:HG2	2.09	0.53
1:B:408:VAL:HG11	1:B:525:ARG:HG3	1.90	0.53
1:A:541:SER:O	1:A:542:ALA:CB	2.57	0.53
1:A:458:PRO:HA	1:A:465:TYR:CD2	2.44	0.52
1:B:424:ARG:CG	1:B:424:ARG:HH11	2.20	0.52
1:B:304:ASP:OD2	1:B:306:ASP:HB3	2.09	0.52
1:A:541:SER:O	1:A:542:ALA:HB2	2.10	0.52
1:B:340:VAL:HG11	1:B:443:MET:CE	2.40	0.51
1:A:420:ALA:HB2	1:A:505:THR:HG21	1.92	0.51
1:B:122:GLY:O	1:B:123:PHE:HB2	2.09	0.51
1:B:210:GLY:HA3	1:B:232:PRO:HD3	1.93	0.51

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:75:THR:O	1:A:78:PRO:HD3	2.11	0.51
1:A:224:ARG:HH11	1:A:224:ARG:HG3	1.75	0.51
1:A:326:VAL:HG12	1:A:328:VAL:HG13	1.93	0.51
1:A:527:GLN:HE21	5:A:901:P6G:H21	1.77	0.50
1:B:224:ARG:HG3	1:B:224:ARG:NH1	2.26	0.50
1:A:329:GLY:HA3	1:A:428:TYR:CZ	2.47	0.50
1:B:296:ARG:NH2	1:B:406:ASN:OD1	2.46	0.49
1:B:328:VAL:O	1:B:427:ALA:HA	2.12	0.49
1:B:67:ASN:OD1	1:B:91:GLU:HB2	2.11	0.49
1:B:243:GLU:O	1:B:247:ARG:HG3	2.13	0.49
1:B:11:ARG:NH1	1:B:16:GLN:CG	2.74	0.49
1:B:104:PRO:HG2	1:B:108:PRO:HD3	1.94	0.49
1:A:216:LEU:HB3	1:A:217:PRO:HD3	1.94	0.49
1:A:161:LEU:HD12	1:A:270:ILE:CD1	2.42	0.49
1:A:224:ARG:HG2	1:A:325:GLN:HB2	1.95	0.48
1:B:541:SER:O	1:B:542:ALA:HB2	2.13	0.48
1:B:369:GLN:HB2	8:B:988:HOH:O	2.12	0.48
1:A:380:LEU:HD22	1:A:385:TRP:CZ2	2.49	0.48
1:A:498:PRO:HG2	1:A:518:LEU:HB2	1.96	0.48
1:A:5:ASP:HB3	1:A:8:LEU:HD12	1.96	0.48
1:A:312:PRO:O	1:A:316:ILE:HG23	2.14	0.48
1:B:329:GLY:HA3	1:B:428:TYR:CZ	2.49	0.47
1:B:66:GLN:HG3	1:B:98:TYR:CD1	2.49	0.47
1:B:7:GLN:OE1	1:B:107:ARG:N	2.43	0.47
1:A:24:ALA:HB3	1:A:140:GLN:HG3	1.96	0.47
1:B:346:PHE:HE2	1:B:395:ARG:HG2	1.80	0.47
1:B:211:MET:HG2	1:B:308:LEU:HD21	1.97	0.47
1:A:393:HIS:HB2	8:A:1033:HOH:O	2.14	0.47
1:B:10:VAL:HG23	1:B:32:PHE:CE2	2.50	0.47
1:A:122:GLY:O	1:A:123:PHE:HB2	2.14	0.47
1:A:369:GLN:HB2	8:A:1027:HOH:O	2.14	0.46
1:A:159:LEU:C	1:A:159:LEU:HD23	2.36	0.46
1:A:66:GLN:HG3	1:A:98:TYR:CD1	2.51	0.46
1:A:88:PRO:HG2	1:A:92:LEU:HD21	1.97	0.46
1:A:224:ARG:NH1	1:A:224:ARG:HG3	2.29	0.46
1:B:206:ALA:HB3	1:B:230:GLY:HA3	1.97	0.46
1:A:380:LEU:HD22	1:A:385:TRP:HZ2	1.81	0.46
1:A:300:VAL:HB	1:A:301:PRO:HD2	1.98	0.46
1:A:327:LEU:HD11	1:A:500:TRP:CZ2	2.51	0.45
1:A:166:GLU:HB2	1:A:270:ILE:HD13	1.98	0.45
1:B:532:TRP:CE3	1:B:536:LEU:HD12	2.51	0.45
1:B:75:THR:O	1:B:78:PRO:HD3	2.17	0.45

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:104:PRO:HG3	1:A:143:GLY:HA2	1.98	0.45
1:A:45:ARG:CZ	1:A:51:GLU:OE1	2.64	0.45
1:B:367:VAL:HG12	1:B:370:ALA:HB2	1.98	0.45
1:B:339:LEU:HD11	1:B:399:SER:HA	1.99	0.45
1:A:433:ARG:NH2	1:A:440:PRO:O	2.50	0.45
1:B:296:ARG:HG2	8:B:1066:HOH:O	2.16	0.44
1:B:286:TRP:CD2	6:B:903:AE3:H3C1	2.52	0.44
1:B:200:PHE:HB2	1:B:226:VAL:HB	1.99	0.44
1:A:293:SER:HA	1:A:365:ILE:HG23	2.00	0.44
1:B:161:LEU:HD12	1:B:270:ILE:HD11	2.00	0.44
1:B:496:LYS:NZ	1:B:496:LYS:HB2	2.33	0.44
1:B:66:GLN:HG3	1:B:98:TYR:CG	2.52	0.44
1:B:528:THR:O	1:B:531:PHE:HB3	2.17	0.44
1:B:536:LEU:HA	1:B:536:LEU:HD23	1.83	0.43
1:A:386:LEU:HA	1:A:386:LEU:HD23	1.78	0.43
1:A:7:GLN:O	1:A:107:ARG:NH1	2.52	0.43
1:B:197:VAL:N	1:B:223:HIS:HD2	1.88	0.43
1:B:202:GLU:HA	1:B:228:GLN:O	2.18	0.43
1:B:224:ARG:HG3	1:B:224:ARG:HH11	1.84	0.43
1:B:113:PRO:HG3	1:B:485:ARG:HG2	2.01	0.43
1:A:200:PHE:HB2	1:A:226:VAL:HB	2.01	0.43
1:B:326:VAL:HG12	1:B:328:VAL:HG13	2.01	0.43
1:B:497:SER:CB	1:B:498:PRO:C	2.82	0.42
1:A:360:LEU:CD2	1:A:379:VAL:HG11	2.49	0.42
1:A:103:THR:HG22	1:A:145:VAL:HG22	2.00	0.42
1:B:161:LEU:CD1	1:B:269:LEU:HD22	2.48	0.42
1:A:376:GLU:HG2	1:B:538:LYS:HE3	2.00	0.42
1:B:76:LEU:HD22	1:B:341:TYR:CE2	2.54	0.42
1:A:243:GLU:CD	1:A:246:ARG:NH1	2.73	0.42
1:B:491:ASP:HA	1:B:492:PRO:HD3	1.66	0.42
1:B:356:ARG:HA	1:B:394:LEU:HD13	2.02	0.42
1:B:107:ARG:HA	1:B:108:PRO:HD3	1.94	0.42
1:A:328:VAL:O	1:A:427:ALA:HA	2.19	0.41
1:A:200:PHE:CB	1:A:226:VAL:HB	2.50	0.41
1:B:491:ASP:HB3	1:B:494:ASP:HB3	2.02	0.41
1:B:496:LYS:HB2	1:B:496:LYS:HZ1	1.85	0.41
1:B:48:MET:HE1	1:B:165:ARG:O	2.21	0.41
1:B:88:PRO:HG2	1:B:92:LEU:HD21	2.02	0.41
1:A:290:PRO:HG2	1:A:291:GLN:NE2	2.36	0.41
1:B:161:LEU:HD23	1:B:161:LEU:HA	1.88	0.41
1:B:495:SER:C	1:B:496:LYS:O	2.58	0.41
1:B:104:PRO:CG	1:B:108:PRO:HG3	2.50	0.41

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:5:ASP:HA	1:A:6:PRO:HD2	1.96	0.41
1:B:286:TRP:CE2	6:B:903:AE3:H3C1	2.55	0.41
1:B:38:ALA:HB2	1:B:178:LEU:HD23	2.02	0.41
1:A:160:ALA:HB2	1:A:169:GLY:HA2	2.02	0.41
1:B:115:LEU:HD23	1:B:198:THR:HB	2.03	0.41
1:A:206:ALA:HB3	1:A:230:GLY:HA3	2.03	0.41
1:B:273:LEU:HA	1:B:273:LEU:HD23	1.93	0.40
1:A:524:LEU:O	1:A:525:ARG:C	2.60	0.40
1:B:454:ILE:HD12	1:B:480:TRP:CE2	2.57	0.40
1:B:380:LEU:HD22	1:B:385:TRP:CZ2	2.57	0.40
1:A:437:LEU:HD11	1:A:449:TYR:CD2	2.57	0.40
1:B:134:ASP:OD1	1:B:136:ARG:HD2	2.20	0.40
1:A:183:VAL:HG13	1:A:187:ILE:HB	2.04	0.40
1:A:390:ASP:HA	1:A:391:PRO:HD3	1.95	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	531/543 (98%)	506 (95%)	25 (5%)	0	100	100
1	B	529/543 (97%)	507 (96%)	17 (3%)	5 (1%)	25	27
All	All	1060/1086 (98%)	1013 (96%)	42 (4%)	5 (0%)	38	45

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	542	ALA
1	B	496	LYS
1	B	494	ASP
1	B	497	SER
1	B	108	PRO

### 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	439/443 (99%)	423 (96%)	16 (4%)	47	61
1	B	438/443 (99%)	423 (97%)	15 (3%)	49	64
All	All	877/886 (99%)	846 (96%)	31 (4%)	48	63

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

Mol	Chain	Res	Type
1	A	3	ARG
1	A	13	ARG
1	A	21	ARG
1	A	23	LYS
1	A	45	ARG
1	A	70	TYR
1	A	249	THR
1	A	291	GLN
1	A	295	PHE
1	A	322	GLN
1	A	356	ARG
1	A	410	PRO
1	A	413	GLN
1	A	418	LEU
1	A	424	ARG
1	A	525	ARG
1	B	11	ARG
1	B	136	ARG
1	B	203	SER
1	B	246	ARG
1	B	249	THR
1	B	281	LEU
1	B	291	GLN
1	B	295	PHE
1	B	296	ARG
1	B	410	PRO
1	B	424	ARG
1	B	467	THR
1	B	496	LYS

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	B	497	SER
1	B	543	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (6) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	223	HIS
1	A	291	GLN
1	A	527	GLN
1	B	223	HIS
1	B	291	GLN
1	B	499	GLN

### 5.3.3 RNA ⓘ

There are no RNA chains 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 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	$\# Z  > 2$
2	NAG	A	544	1,2	12,14,15	0.65	0	15,19,21	0.92	1 (6%)
2	FUC	A	545	2	9,10,11	0.60	0	10,14,16	0.76	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical



component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	NAG	A	544	1,2	-	0/6/23/26	0/1/1/1
2	FUC	A	545	2	-	0/0/17/20	0/1/1/1

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	544	NAG	C2-N2-C7	-2.08	119.60	123.09

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

## 5.6 Ligand geometry

8 ligands 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
3	NAG	A	701	1	12,14,15	0.60	0	15,19,21	0.76	1 (6%)
5	P6G	A	901	-	18,18,18	2.07	6 (33%)	17,17,17	1.08	1 (5%)
6	AE3	A	902	-	8,8,8	1.58	2 (25%)	7,7,7	1.25	1 (14%)
4	CO3	A	951	-	0,3,3	0.00	-	0,3,3	0.00	-
3	NAG	B	601	1	12,14,15	0.77	0	15,19,21	0.70	0
6	AE3	B	903	-	8,8,8	1.71	2 (25%)	7,7,7	1.38	1 (14%)
7	PG4	B	904	-	10,10,12	11.10	5 (50%)	7,9,11	1.02	0
4	CO3	B	952	-	0,3,3	0.00	-	0,3,3	0.00	-

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
3	NAG	A	701	1	-	0/6/23/26	0/1/1/1
5	P6G	A	901	-	-	0/16/16/16	0/0/0/0
6	AE3	A	902	-	-	0/6/6/6	0/0/0/0
4	CO3	A	951	-	-	0/0/0/0	0/0/0/0
3	NAG	B	601	1	-	0/6/23/26	0/1/1/1
6	AE3	B	903	-	-	0/6/6/6	0/0/0/0
7	PG4	B	904	-	-	0/8/8/10	0/0/0/0
4	CO3	B	952	-	-	0/0/0/0	0/0/0/0

All (15) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
7	B	904	PG4	C1-C2	-25.83	1.44	1.55
7	B	904	PG4	C8-C7	-22.68	1.46	1.55
7	B	904	PG4	O2-C3	4.78	1.62	1.42
7	B	904	PG4	O4-C7	4.09	1.57	1.42
5	A	901	P6G	O10-C9	4.05	1.59	1.42
5	A	901	P6G	O13-C12	3.68	1.57	1.42
5	A	901	P6G	O16-C15	3.56	1.57	1.42
6	B	903	AE3	O3-C5	3.54	1.57	1.42
5	A	901	P6G	O4-C3	3.34	1.56	1.42
6	A	902	AE3	O3-C5	3.30	1.56	1.42
5	A	901	P6G	O19-C18	3.12	1.58	1.42
6	B	903	AE3	O2-C3	2.97	1.54	1.42
5	A	901	P6G	O7-C6	2.87	1.54	1.42
7	B	904	PG4	O3-C5	2.80	1.54	1.42
6	A	902	AE3	O2-C3	2.61	1.53	1.42

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	A	901	P6G	O1-C2-C3	3.04	131.48	111.80
6	B	903	AE3	O4-C6-C5	2.96	130.98	111.80
6	A	902	AE3	O4-C6-C5	2.77	129.76	111.80
3	A	701	NAG	C2-N2-C7	-2.14	119.50	123.09

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	535/543 (98%)	0.05	25 (4%)	30 34	30, 44, 66, 98	0
1	B	533/543 (98%)	0.23	38 (7%)	16 18	35, 49, 71, 99	0
All	All	1068/1086 (98%)	0.14	63 (5%)	22 25	30, 46, 69, 99	0

All (63) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	543	THR	6.5
1	B	497	SER	6.0
1	B	118	ILE	4.7
1	A	495	SER	4.5
1	B	540	LEU	4.5
1	A	493	ARG	4.3
1	B	541	SER	4.1
1	B	117	TRP	3.9
1	A	118	ILE	3.8
1	A	492	PRO	3.6
1	B	207	ALA	3.6
1	B	542	ALA	3.5
1	B	204	ALA	3.4
1	A	496	LYS	3.4
1	B	200	PHE	3.2
1	A	540	LEU	3.2
1	A	494	ASP	3.2
1	A	497	SER	3.2
1	B	109	ALA	3.2
1	A	542	ALA	3.1
1	B	4	GLU	3.1
1	B	227	LEU	3.1
1	B	119	TYR	3.0
1	A	230	GLY	3.0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	1	GLU	3.0
1	A	207	ALA	2.9
1	A	236	TRP	2.8
1	B	201	GLY	2.8
1	B	78	PRO	2.8
1	A	117	TRP	2.7
1	B	265	ASN	2.7
1	B	349	ASP	2.7
1	A	541	SER	2.7
1	B	203	SER	2.7
1	B	255	VAL	2.6
1	B	230	GLY	2.5
1	B	228	GLN	2.5
1	A	231	THR	2.5
1	B	471	ILE	2.4
1	B	133	TYR	2.4
1	B	147	VAL	2.4
1	B	467	THR	2.4
1	A	268	GLU	2.4
1	A	498	PRO	2.4
1	B	236	TRP	2.3
1	A	201	GLY	2.3
1	B	268	GLU	2.3
1	B	229	SER	2.3
1	A	229	SER	2.2
1	A	237	ALA	2.2
1	B	206	ALA	2.2
1	A	300	VAL	2.2
1	B	322	GLN	2.2
1	B	208	SER	2.2
1	B	205	GLY	2.2
1	A	206	ALA	2.2
1	B	9	LEU	2.2
1	B	455	PHE	2.2
1	A	204	ALA	2.2
1	A	119	TYR	2.2
1	B	132	VAL	2.1
1	B	495	SER	2.1
1	B	323	ASP	2.1

## 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. 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
2	FUC	A	545	10/11	0.47	22.46	91,93,95,95	0
2	NAG	A	544	14/15	0.35	7.20	76,79,85,89	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
6	AE3	A	902	9/9	0.41	7.00	81,82,83,83	0
6	AE3	B	903	9/9	0.35	5.50	83,87,89,89	0
7	PG4	B	904	11/13	0.17	5.09	87,88,90,90	0
3	NAG	A	701	14/15	0.31	4.28	81,85,88,89	0
3	NAG	B	601	14/15	0.36	3.43	87,89,91,92	0
5	P6G	A	901	19/19	0.16	2.20	61,71,75,76	0
4	CO3	B	952	4/4	0.31	2.01	63,64,64,65	0
4	CO3	A	951	4/4	0.28	1.67	50,50,51,52	0

## 6.5 Other polymers ⓘ

There are no such residues in this entry.