



# Full wwPDB X-ray Structure Validation Report

Feb 28, 2014 – 09:39 PM GMT

PDB ID : 2PGZ  
Title : Crystal structure of Cocaine bound to an ACh-Binding Protein  
Authors : Hansen, S.B.; Taylor, P.  
Deposited on : 2007-04-10  
Resolution : 1.76 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>

---

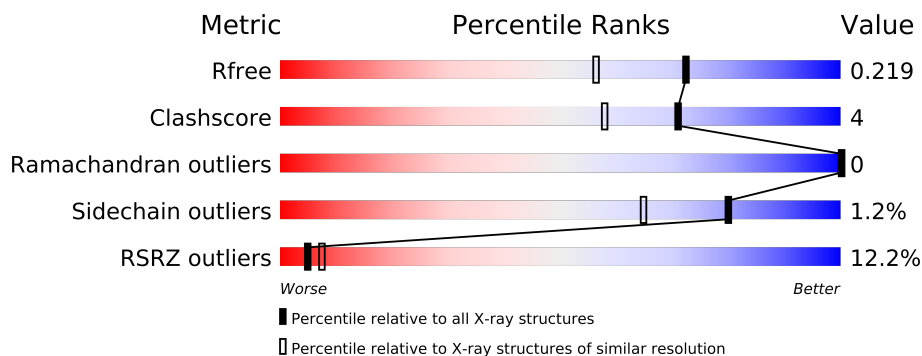
The following versions of software and data (see [references](#)) were used in the production of this report:

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 1.76 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	1134 (1.76-1.76)
Clashscore	79885	1304 (1.76-1.76)
Ramachandran outliers	78287	1288 (1.76-1.76)
Sidechain outliers	78261	1288 (1.76-1.76)
RSRZ outliers	66119	1135 (1.76-1.76)

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	230	
1	B	230	
1	C	230	
1	D	230	
1	E	230	

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

Mol	Type	Chain	Res	Geometry	Electron density
4	COC	D	401	-	X
5	PG4	B	403	-	X
5	PG4	C	407	-	X

## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 9900 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 Soluble acetylcholine receptor.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	211	Total	C	N	O	S	0	6	0
			1721	1085	285	343	8			
1	B	211	Total	C	N	O	S	0	2	0
			1695	1072	276	339	8			
1	C	211	Total	C	N	O	S	0	7	0
			1728	1091	286	341	10			
1	D	214	Total	C	N	O	S	0	5	0
			1742	1100	285	348	9			
1	E	209	Total	C	N	O	S	0	9	0
			1718	1086	282	341	9			

There are 55 discrepancies between the modelled and reference sequences:

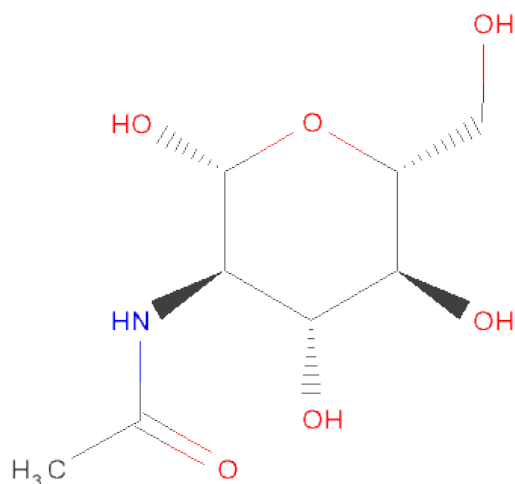
Chain	Residue	Modelled	Actual	Comment	Reference
A	-8	ASP	-	CLONING ARTIFACT	UNP Q8WSF8
A	-7	TYR	-	CLONING ARTIFACT	UNP Q8WSF8
A	-6	LYS	-	CLONING ARTIFACT	UNP Q8WSF8
A	-5	ASP	-	CLONING ARTIFACT	UNP Q8WSF8
A	-4	ASP	-	CLONING ARTIFACT	UNP Q8WSF8
A	-3	ASP	-	CLONING ARTIFACT	UNP Q8WSF8
A	-2	ASP	-	CLONING ARTIFACT	UNP Q8WSF8
A	-1	LYS	-	CLONING ARTIFACT	UNP Q8WSF8
A	0	LEU	-	CLONING ARTIFACT	UNP Q8WSF8
A	220	SER	-	CLONING ARTIFACT	UNP Q8WSF8
A	221	ARG	-	CLONING ARTIFACT	UNP Q8WSF8
B	-8	ASP	-	CLONING ARTIFACT	UNP Q8WSF8
B	-7	TYR	-	CLONING ARTIFACT	UNP Q8WSF8
B	-6	LYS	-	CLONING ARTIFACT	UNP Q8WSF8
B	-5	ASP	-	CLONING ARTIFACT	UNP Q8WSF8
B	-4	ASP	-	CLONING ARTIFACT	UNP Q8WSF8
B	-3	ASP	-	CLONING ARTIFACT	UNP Q8WSF8
B	-2	ASP	-	CLONING ARTIFACT	UNP Q8WSF8
B	-1	LYS	-	CLONING ARTIFACT	UNP Q8WSF8

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	0	LEU	-	CLONING ARTIFACT	UNP Q8WSF8
B	220	SER	-	CLONING ARTIFACT	UNP Q8WSF8
B	221	ARG	-	CLONING ARTIFACT	UNP Q8WSF8
C	-8	ASP	-	CLONING ARTIFACT	UNP Q8WSF8
C	-7	TYR	-	CLONING ARTIFACT	UNP Q8WSF8
C	-6	LYS	-	CLONING ARTIFACT	UNP Q8WSF8
C	-5	ASP	-	CLONING ARTIFACT	UNP Q8WSF8
C	-4	ASP	-	CLONING ARTIFACT	UNP Q8WSF8
C	-3	ASP	-	CLONING ARTIFACT	UNP Q8WSF8
C	-2	ASP	-	CLONING ARTIFACT	UNP Q8WSF8
C	-1	LYS	-	CLONING ARTIFACT	UNP Q8WSF8
C	0	LEU	-	CLONING ARTIFACT	UNP Q8WSF8
C	220	SER	-	CLONING ARTIFACT	UNP Q8WSF8
C	221	ARG	-	CLONING ARTIFACT	UNP Q8WSF8
D	-8	ASP	-	CLONING ARTIFACT	UNP Q8WSF8
D	-7	TYR	-	CLONING ARTIFACT	UNP Q8WSF8
D	-6	LYS	-	CLONING ARTIFACT	UNP Q8WSF8
D	-5	ASP	-	CLONING ARTIFACT	UNP Q8WSF8
D	-4	ASP	-	CLONING ARTIFACT	UNP Q8WSF8
D	-3	ASP	-	CLONING ARTIFACT	UNP Q8WSF8
D	-2	ASP	-	CLONING ARTIFACT	UNP Q8WSF8
D	-1	LYS	-	CLONING ARTIFACT	UNP Q8WSF8
D	0	LEU	-	CLONING ARTIFACT	UNP Q8WSF8
D	220	SER	-	CLONING ARTIFACT	UNP Q8WSF8
D	221	ARG	-	CLONING ARTIFACT	UNP Q8WSF8
E	-8	ASP	-	CLONING ARTIFACT	UNP Q8WSF8
E	-7	TYR	-	CLONING ARTIFACT	UNP Q8WSF8
E	-6	LYS	-	CLONING ARTIFACT	UNP Q8WSF8
E	-5	ASP	-	CLONING ARTIFACT	UNP Q8WSF8
E	-4	ASP	-	CLONING ARTIFACT	UNP Q8WSF8
E	-3	ASP	-	CLONING ARTIFACT	UNP Q8WSF8
E	-2	ASP	-	CLONING ARTIFACT	UNP Q8WSF8
E	-1	LYS	-	CLONING ARTIFACT	UNP Q8WSF8
E	0	LEU	-	CLONING ARTIFACT	UNP Q8WSF8
E	220	SER	-	CLONING ARTIFACT	UNP Q8WSF8
E	221	ARG	-	CLONING ARTIFACT	UNP Q8WSF8

- Molecule 2 is SUGAR (N-ACETYL-D-GLUCOSAMINE) (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



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

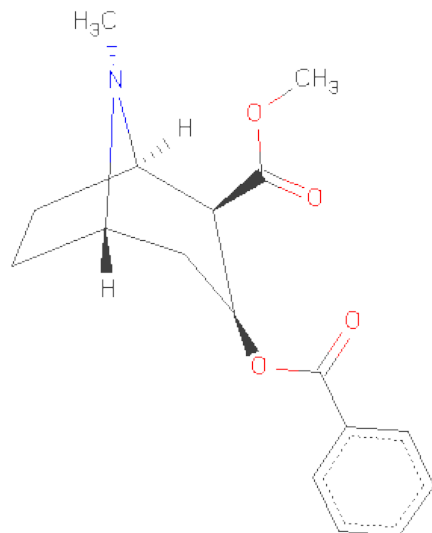
- Molecule 3 is a polymer of unknown type called SUGAR (5-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	C	5	Total	C	N	O	0	0
			61	34	2	25		

There are 11 discrepancies between the modelled and reference sequences:

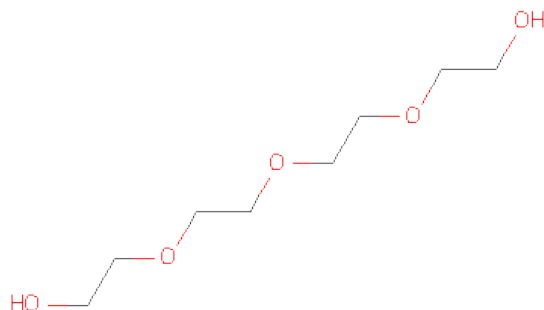
Chain	Residue	Modelled	Actual	Comment	Reference
C	-8	ASP	-	CLONING ARTIFACT	UNP Q8WSF8
C	-7	TYR	-	CLONING ARTIFACT	UNP Q8WSF8
C	-6	LYS	-	CLONING ARTIFACT	UNP Q8WSF8
C	-5	ASP	-	CLONING ARTIFACT	UNP Q8WSF8
C	-4	ASP	-	CLONING ARTIFACT	UNP Q8WSF8
C	-3	ASP	-	CLONING ARTIFACT	UNP Q8WSF8
C	-2	ASP	-	CLONING ARTIFACT	UNP Q8WSF8
C	-1	LYS	-	CLONING ARTIFACT	UNP Q8WSF8
C	0	LEU	-	CLONING ARTIFACT	UNP Q8WSF8
C	220	SER	-	CLONING ARTIFACT	UNP Q8WSF8
C	221	ARG	-	CLONING ARTIFACT	UNP Q8WSF8

- Molecule 4 is COCAINE (three-letter code: COC) (formula: C<sub>17</sub>H<sub>21</sub>NO<sub>4</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	1	Total	C	N	O	0	0
			22	17	1	4		
4	D	1	Total	C	N	O	0	0
			22	17	1	4		

- Molecule 5 is TETRAETHYLENE GLYCOL (three-letter code: PG4) (formula: C<sub>8</sub>H<sub>18</sub>O<sub>5</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	B	1	Total	C	O	0	0
			13	8	5		
5	C	1	Total	C	O	0	0
			10	6	4		

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	E	1	Total	C	O	0	0
			13	8	5		

- Molecule 6 is water.

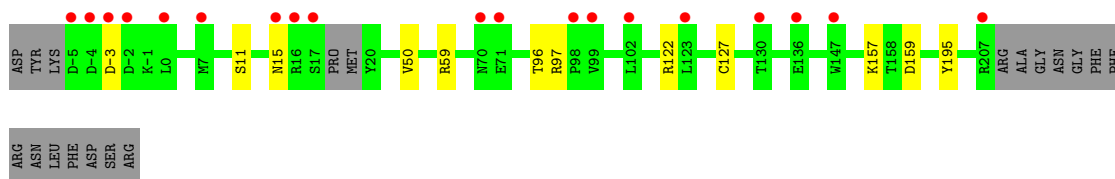
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	251	Total	O	0	6
			251	251		
6	B	229	Total	O	0	2
			229	229		
6	C	230	Total	O	0	8
			230	230		
6	D	223	Total	O	0	2
			223	223		
6	E	208	Total	O	0	6
			208	208		

### 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.

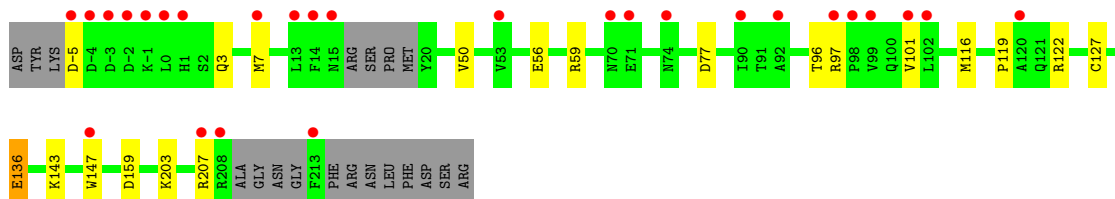
- Molecule 1: Soluble acetylcholine receptor

Chain A: 



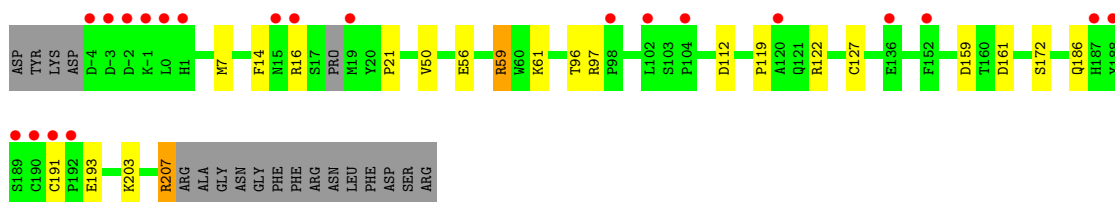
- Molecule 1: Soluble acetylcholine receptor

Chain B: 



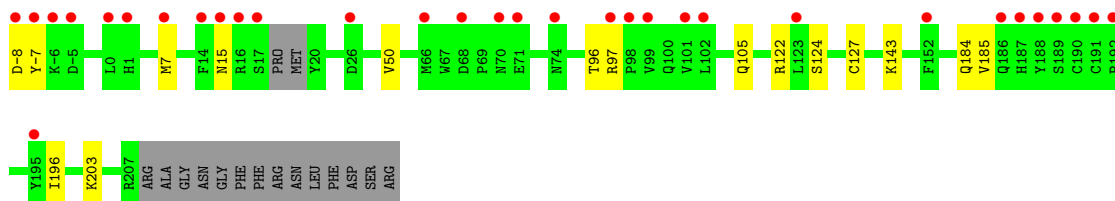
- Molecule 1: Soluble acetylcholine receptor

Chain C: 



- Molecule 1: Soluble acetylcholine receptor

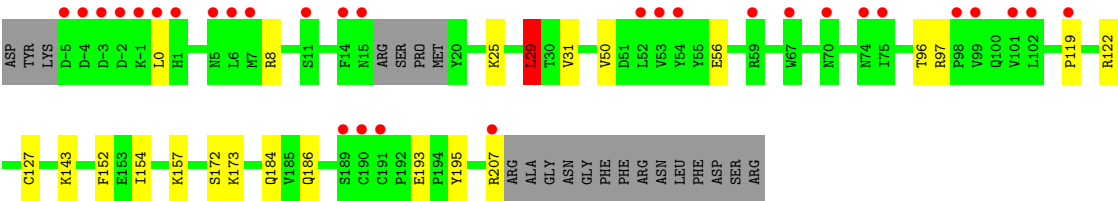
Chain D: 





● Molecule 1: Soluble acetylcholine receptor

Chain E: 



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	86.82Å 115.59Å 130.88Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.00 – 1.76 47.62 – 1.76	Depositor EDS
% Data completeness (in resolution range)	99.5 (50.00-1.76) 99.5 (47.62-1.76)	Depositor EDS
$R_{merge}$	0.06	Depositor
$R_{sym}$	0.05	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.37 (at 1.76Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, $R_{free}$	0.180 , 0.210 0.191 , 0.219	Depositor DCC
$R_{free}$ test set	1314 reflections (1.02%)	DCC
Wilson B-factor (Å <sup>2</sup> )	25.8	Xtriage
Anisotropy	0.057	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 31.9	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.51$ , $\langle L^2 \rangle = 0.34$	Xtriage
Outliers	0 of 129999 reflections	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	9900	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	29.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.27% 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: PG4, COC, BMA, NAG, 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	0.53	0/1777	0.61	0/2418
1	B	0.54	0/1740	0.63	0/2369
1	C	0.49	0/1787	0.64	0/2428
1	D	0.61	1/1796 (0.1%)	0.66	0/2443
1	E	0.65	3/1787 (0.2%)	0.76	3/2430 (0.1%)
All	All	0.57	4/8887 (0.0%)	0.66	3/12088 (0.0%)

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	E	8	ARG	NE-CZ	11.91	1.48	1.33
1	E	8	ARG	CG-CD	8.34	1.72	1.51
1	D	-8	ASP	N-CA	6.38	1.59	1.46
1	E	8	ARG	CZ-NH2	5.43	1.40	1.33

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	E	8	ARG	NE-CZ-NH1	-16.74	111.93	120.30
1	E	8	ARG	NE-CZ-NH2	10.15	125.37	120.30
1	E	29	LEU	CA-CB-CG	5.31	127.52	115.30

There are no chirality outliers.

There are no planarity outliers.

### 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	1721	0	1659	9	0
1	B	1695	0	1615	19	0
1	C	1728	0	1675	21	0
1	D	1742	0	1675	15	1
1	E	1718	0	1661	16	1
2	B	14	0	13	0	0
3	C	61	0	52	0	0
4	A	22	0	21	1	0
4	D	22	0	21	0	0
5	B	13	0	18	3	0
5	C	10	0	13	0	0
5	E	13	0	18	2	0
6	A	251	0	0	3	0
6	B	229	0	0	8	0
6	C	230	0	0	7	0
6	D	223	0	0	2	0
6	E	208	0	0	5	0
All	All	9900	0	8441	77	1

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 4.

All (77) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:C:97[B]:ARG:HG3	6:C:633[B]:HOH:O	1.45	1.12
1:A:97[B]:ARG:NE	6:A:650[B]:HOH:O	1.65	1.00
5:B:403:PG4:H62	6:B:611:HOH:O	1.78	0.83
1:A:97[B]:ARG:CZ	6:A:650[B]:HOH:O	2.16	0.79
1:C:207[A]:ARG:HD3	6:C:635[A]:HOH:O	1.82	0.78
1:C:207[A]:ARG:HH11	1:C:207[A]:ARG:HG2	1.49	0.78
1:C:207[A]:ARG:HH11	1:C:207[A]:ARG:CG	1.97	0.78
1:A:97[B]:ARG:NH2	6:A:650[B]:HOH:O	2.17	0.75
1:C:59[B]:ARG:NH1	1:C:159:ASP:OD2	2.17	0.75
1:D:105:GLN:OE1	6:D:580:HOH:O	2.04	0.74
1:B:101:VAL:HG13	6:B:632:HOH:O	1.87	0.73
1:B:59:ARG:NH1	1:B:159:ASP:OD2	2.21	0.73
1:A:59[A]:ARG:NH1	1:A:159:ASP:OD2	2.23	0.71
1:C:191:CYS:HB3	1:C:193:GLU:OE2	1.93	0.68

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:D:105:GLN:OE1	6:D:594:HOH:O	2.12	0.67
5:E:403:PG4:H32	6:E:587:HOH:O	1.95	0.67
1:B:147:TRP:O	5:B:403:PG4:H82	1.95	0.66
1:D:185:VAL:HG13	1:D:196:ILE:CD1	2.27	0.65
1:A:11:SER:HA	1:A:15:ASN:HD22	1.62	0.63
1:E:29:LEU:HD21	1:E:31:VAL:HG23	1.80	0.63
1:D:97[A]:ARG:HH21	1:D:124:SER:CB	2.13	0.62
1:C:203[B]:LYS:HG3	6:C:565:HOH:O	1.99	0.61
1:D:143:LYS:HZ2	1:D:184:GLN:HE22	1.50	0.60
1:C:21:PRO:HB3	1:D:7:MET:CE	2.31	0.60
5:E:403:PG4:H21	6:E:587:HOH:O	2.01	0.59
1:D:143:LYS:NZ	1:D:184:GLN:HE22	2.00	0.59
1:E:29:LEU:HD21	1:E:31:VAL:CG2	2.34	0.57
1:B:203[B]:LYS:HG3	6:B:623:HOH:O	2.05	0.57
1:D:7:MET:CE	1:D:7:MET:HA	2.34	0.57
1:C:97[B]:ARG:CG	6:C:633[B]:HOH:O	2.24	0.56
1:E:143[A]:LYS:HE2	1:E:184:GLN:HE22	1.70	0.56
1:C:161[A]:ASP:HB3	6:C:624:HOH:O	2.05	0.56
1:B:59:ARG:HG2	1:B:116:MET:HG3	1.88	0.56
1:B:96:THR:O	1:C:122:ARG:HD2	2.06	0.56
1:C:7:MET:HG2	6:C:505:HOH:O	2.06	0.55
1:C:207[A]:ARG:CG	1:C:207[A]:ARG:NH1	2.66	0.55
1:C:97[A]:ARG:NH1	6:C:442:HOH:O	2.39	0.54
1:B:136:GLU:CD	1:B:136:GLU:H	2.12	0.53
1:D:7:MET:HA	1:D:7:MET:HE2	1.90	0.52
1:B:59:ARG:CG	1:B:116:MET:HG3	2.40	0.52
1:C:96:THR:O	1:D:122:ARG:HD2	2.10	0.52
1:E:29:LEU:CD2	1:E:31:VAL:HG23	2.39	0.52
1:D:50:VAL:HG21	1:D:127:CYS:SG	2.50	0.51
1:A:96:THR:O	1:B:122:ARG:HD2	2.11	0.51
1:B:59:ARG:HD3	1:B:116:MET:HE2	1.93	0.50
1:A:122:ARG:HD2	1:E:96:THR:O	2.12	0.50
1:E:157:LYS:NZ	6:E:562:HOH:O	2.43	0.50
1:B:3:GLN:O	1:B:7:MET:HG3	2.11	0.50
1:C:50:VAL:HG21	1:C:127:CYS:SG	2.52	0.49
1:D:96:THR:O	1:E:122:ARG:HD2	2.13	0.49
1:B:59:ARG:HD3	1:B:116:MET:CE	2.43	0.48
1:B:97[B]:ARG:HD3	6:B:476:HOH:O	2.13	0.48
1:B:143:LYS:NZ	6:B:532:HOH:O	2.45	0.48
1:E:25:LYS:HE2	1:E:152:PHE:CD2	2.48	0.48
1:B:50:VAL:HG21	1:B:127:CYS:SG	2.54	0.47
1:A:195:TYR:CG	4:A:401:COC:H15	2.50	0.47

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:E:25:LYS:HG3	1:E:152:PHE:HB3	1.96	0.46
1:C:61:LYS:HE3	1:C:112:ASP:O	2.15	0.46
1:E:50:VAL:HG21	1:E:127:CYS:SG	2.56	0.46
1:E:29:LEU:HD22	1:E:154:ILE:HG12	1.97	0.45
1:C:21:PRO:HB3	1:D:7:MET:HE1	1.97	0.45
1:E:97[A]:ARG:HD2	6:E:609[A]:HOH:O	2.17	0.45
1:B:56:GLU:O	1:B:119:PRO:HD2	2.16	0.44
5:B:403:PG4:H12	6:B:611:HOH:O	2.18	0.44
1:E:97[A]:ARG:HD3	6:E:476:HOH:O	2.17	0.44
1:E:172:SER:O	1:E:207[A]:ARG:NH2	2.50	0.44
1:B:3:GLN:HB3	1:B:7:MET:HE3	1.99	0.44
1:A:50:VAL:HG21	1:A:127:CYS:SG	2.59	0.43
1:D:185:VAL:HG13	1:D:196:ILE:HD13	2.01	0.42
1:C:14:PHE:C	1:C:16:ARG:H	2.23	0.42
1:B:7:MET:HG2	6:B:597:HOH:O	2.19	0.41
1:E:193:GLU:HG2	1:E:195:TYR:CE2	2.55	0.41
1:E:56:GLU:O	1:E:119:PRO:HD2	2.21	0.41
1:C:172:SER:N	1:C:207[B]:ARG:HH11	2.18	0.40
1:B:207:ARG:HD3	6:B:523:HOH:O	2.20	0.40
1:C:56:GLU:O	1:C:119:PRO:HD2	2.22	0.40

All (1) 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)
1:D:-7:TYR:OH	1:E:173:LYS:O[2_554]	2.12	0.08

## 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	212/230 (92%)	210 (99%)	2 (1%)	0	100	100
1	B	208/230 (90%)	206 (99%)	2 (1%)	0	100	100
1	C	213/230 (93%)	208 (98%)	5 (2%)	0	100	100

Continued on next page...

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	D	214/230 (93%)	212 (99%)	2 (1%)	0	100	100
1	E	213/230 (93%)	210 (99%)	3 (1%)	0	100	100
All	All	1060/1150 (92%)	1046 (99%)	14 (1%)	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	198/208 (95%)	196 (99%)	2 (1%)	85	72
1	B	193/208 (93%)	190 (98%)	3 (2%)	75	55
1	C	199/208 (96%)	194 (98%)	5 (2%)	60	33
1	D	200/208 (96%)	199 (100%)	1 (0%)	94	88
1	E	199/208 (96%)	196 (98%)	3 (2%)	76	59
All	All	989/1040 (95%)	975 (99%)	14 (1%)	80	62

All (14) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	-3	ASP
1	A	157	LYS
1	B	-5	ASP
1	B	77	ASP
1	B	136	GLU
1	C	59[A]	ARG
1	C	59[B]	ARG
1	C	186	GLN
1	C	207[A]	ARG
1	C	207[B]	ARG
1	D	15	ASN
1	E	0	LEU
1	E	29	LEU
1	E	186	GLN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (14) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	15	ASN
1	A	63	ASN
1	B	3	GLN
1	B	184	GLN
1	C	184	GLN
1	C	186	GLN
1	D	3	GLN
1	D	15	ASN
1	D	74	ASN
1	D	184	GLN
1	E	3	GLN
1	E	15	ASN
1	E	184	GLN
1	E	186	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 ⓘ

5 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	$\# Z  > 2$
3	NAG	C	402	1,3	12,14,15	0.52	0	15,19,21	1.28	2 (13%)
3	NAG	C	403	3	12,14,15	0.50	0	15,19,21	1.43	2 (13%)
3	BMA	C	404	3	10,11,12	0.45	0	11,15,17	0.94	0



Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
3	MAN	C	405	3	10,11,12	0.44	0	11,15,17	0.67	0
3	MAN	C	406	3	10,11,12	0.46	0	11,15,17	0.76	1 (9%)

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	C	402	1,3	-	0/6/23/26	0/1/1/1
3	NAG	C	403	3	-	0/6/23/26	0/1/1/1
3	BMA	C	404	3	-	0/2/19/22	0/1/1/1
3	MAN	C	405	3	-	0/2/19/22	0/1/1/1
3	MAN	C	406	3	-	0/2/19/22	0/1/1/1

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	C	403	NAG	C3-C2-N2	-4.37	105.11	111.76
3	C	402	NAG	C2-N2-C7	3.27	128.58	123.09
3	C	403	NAG	O5-C5-C6	2.41	109.51	106.98
3	C	402	NAG	O5-C5-C6	2.19	109.28	106.98
3	C	406	MAN	O5-C5-C6	2.09	109.17	106.98

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

## 5.6 Ligand geometry

6 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
4	COC	A	401	-	24,24,24	1.55	4 (16%)	34,34,34	3.72	9 (26%)
2	NAG	B	402	1	12,14,15	3.37	4 (33%)	15,19,21	2.87	4 (26%)
5	PG4	B	403	-	12,12,12	1.55	3 (25%)	11,11,11	0.65	0
5	PG4	C	407	-	9,9,12	1.52	2 (22%)	8,8,11	1.00	0
4	COC	D	401	-	24,24,24	1.67	3 (12%)	34,34,34	3.46	15 (44%)
5	PG4	E	403	-	12,12,12	0.53	0	11,11,11	0.27	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
4	COC	A	401	-	-	0/14/39/39	0/1/3/3
2	NAG	B	402	1	-	0/6/23/26	0/1/1/1
5	PG4	B	403	-	-	0/10/10/10	0/0/0/0
5	PG4	C	407	-	-	0/7/7/10	0/0/0/0
4	COC	D	401	-	-	0/14/39/39	0/1/3/3
5	PG4	E	403	-	-	0/10/10/10	0/0/0/0

All (16) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	402	NAG	C7-N2	7.54	1.65	1.34
2	B	402	NAG	O7-C7	7.30	1.39	1.23
4	D	401	COC	O1-C2	5.00	1.45	1.34
4	D	401	COC	O5-C6	4.99	1.46	1.33
4	A	401	COC	O1-C2	4.33	1.44	1.34
2	B	402	NAG	O6-C6	4.19	1.60	1.42
4	A	401	COC	O5-C6	4.05	1.43	1.33
5	B	403	PG4	O3-C5	3.36	1.56	1.42
5	C	407	PG4	O2-C3	3.10	1.50	1.38
4	A	401	COC	O1-C29	-3.01	1.40	1.46
5	C	407	PG4	O3-C4	2.26	1.51	1.42
4	D	401	COC	O1-C29	-2.25	1.42	1.46
2	B	402	NAG	C8-C7	2.13	1.55	1.50
5	B	403	PG4	C4-C3	2.11	1.59	1.48
5	B	403	PG4	O2-C3	2.09	1.51	1.42
4	A	401	COC	C24-N12	-2.08	1.45	1.48

All (28) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	A	401	COC	C24-N12-C10	14.28	114.71	100.95
4	D	401	COC	C24-N12-C10	13.83	114.28	100.95
4	A	401	COC	C18-C10-N12	-11.41	92.26	105.10
2	B	402	NAG	C2-N2-C7	-8.11	109.47	123.09
4	D	401	COC	C18-C10-N12	-7.40	96.77	105.10
4	A	401	COC	C21-C24-N12	-6.30	93.77	104.59
4	D	401	COC	O5-C6-C8	5.73	119.38	111.00
2	B	402	NAG	C8-C7-N2	-5.56	105.23	116.11
4	D	401	COC	C14-N12-C24	5.05	123.06	113.19
4	A	401	COC	O5-C6-C8	4.75	117.95	111.00
4	A	401	COC	C14-N12-C24	4.16	121.33	113.19
2	B	402	NAG	C3-C2-N2	-4.11	105.50	111.76
4	D	401	COC	C21-C24-N12	-4.03	97.67	104.59
4	D	401	COC	C14-N12-C10	3.98	122.60	113.52
4	A	401	COC	C14-N12-C10	3.70	121.97	113.52
4	D	401	COC	O1-C2-C4	3.30	117.72	111.91
4	A	401	COC	O5-C6-O7	-2.79	118.11	123.79
4	D	401	COC	C1-O5-C6	2.79	122.66	116.02
4	D	401	COC	C26-C29-C8	2.44	116.17	112.14
4	D	401	COC	O5-C6-O7	-2.40	118.91	123.79
4	D	401	COC	C21-C18-C10	2.40	107.83	104.04
4	D	401	COC	C29-O1-C2	-2.33	113.14	117.43
4	A	401	COC	O1-C2-C4	2.32	116.00	111.91
2	B	402	NAG	O6-C6-C5	-2.27	103.53	111.36
4	A	401	COC	C21-C18-C10	2.24	107.59	104.04
4	D	401	COC	O1-C2-O3	-2.09	119.89	123.52
4	D	401	COC	O7-C6-C8	-2.09	121.70	125.13
4	D	401	COC	C26-C24-N12	-2.08	104.69	107.64

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	211/230 (91%)	0.64	19 (9%) 10 12	21, 26, 34, 46	0
1	B	211/230 (91%)	0.97	27 (12%) 4 6	21, 27, 43, 51	0
1	C	211/230 (91%)	0.77	21 (9%) 8 10	22, 28, 47, 57	0
1	D	214/230 (93%)	0.94	32 (14%) 3 4	20, 27, 43, 57	0
1	E	209/230 (90%)	0.99	30 (14%) 3 5	22, 26, 54, 60	0
All	All	1056/1150 (91%)	0.86	129 (12%) 5 7	20, 27, 46, 60	0

All (129) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	-5	ASP	7.3
1	D	189	SER	7.3
1	B	-4	ASP	7.2
1	E	1	HIS	6.9
1	B	-3	ASP	6.9
1	C	-3	ASP	6.8
1	C	190	CYS	6.6
1	E	-5	ASP	6.6
1	B	0	LEU	6.4
1	C	0	LEU	6.4
1	E	-4	ASP	6.3
1	D	16	ARG	6.3
1	E	189	SER	6.1
1	D	15	ASN	6.1
1	E	190	CYS	6.1
1	E	-2	ASP	5.8
1	C	189	SER	5.6
1	D	187	HIS	5.6
1	E	-1	LYS	5.4
1	E	0	LEU	5.4

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	D	192	PRO	5.3
1	C	-4	ASP	5.2
1	E	-3	ASP	5.1
1	C	188	TYR	4.9
1	A	16	ARG	4.9
1	B	-2	ASP	4.9
1	C	-1	LYS	4.8
1	E	11	SER	4.8
1	A	-4	ASP	4.7
1	C	191	CYS	4.7
1	A	0	LEU	4.4
1	B	1	HIS	4.3
1	B	14	PHE	4.3
1	E	14	PHE	4.3
1	C	187	HIS	4.2
1	E	191	CYS	4.0
1	A	-5	ASP	4.0
1	C	19	MET	3.9
1	B	-1	LYS	3.9
1	C	-2	ASP	3.8
1	C	1	HIS	3.8
1	D	188	TYR	3.6
1	E	74	ASN	3.6
1	D	74	ASN	3.5
1	B	208	ARG	3.5
1	D	7	MET	3.4
1	D	14	PHE	3.4
1	C	15	ASN	3.4
1	D	190	CYS	3.4
1	B	70	ASN	3.3
1	A	-3	ASP	3.3
1	D	99[A]	VAL	3.2
1	D	152	PHE	3.2
1	D	-6	LYS	3.2
1	D	0	LEU	3.2
1	D	70	ASN	3.1
1	D	26	ASP	3.1
1	C	102	LEU	3.0
1	E	53	VAL	3.0
1	B	7	MET	3.0
1	B	15	ASN	2.9
1	B	213	PHE	2.9

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	D	-5	ASP	2.9
1	E	75	ILE	2.9
1	D	17	SER	2.9
1	B	207	ARG	2.8
1	C	192	PRO	2.8
1	E	6	LEU	2.8
1	D	186	GLN	2.7
1	B	101	VAL	2.7
1	D	-8	ASP	2.7
1	A	-2	ASP	2.7
1	A	136	GLU	2.7
1	D	191	CYS	2.7
1	E	15	ASN	2.6
1	A	99[A]	VAL	2.6
1	B	98	PRO	2.6
1	A	102	LEU	2.6
1	D	1	HIS	2.5
1	B	74	ASN	2.5
1	D	98	PRO	2.5
1	A	7	MET	2.5
1	A	207[A]	ARG	2.5
1	E	101	VAL	2.5
1	D	195	TYR	2.5
1	E	102	LEU	2.5
1	A	15	ASN	2.5
1	D	101	VAL	2.4
1	B	102	LEU	2.4
1	E	99	VAL	2.4
1	E	70	ASN	2.4
1	E	7	MET	2.4
1	E	5	ASN	2.4
1	C	16	ARG	2.4
1	B	92	ALA	2.4
1	D	66	MET	2.4
1	D	71	GLU	2.4
1	C	136	GLU	2.4
1	B	90	ILE	2.3
1	E	52	LEU	2.3
1	A	147	TRP	2.3
1	C	104	PRO	2.3
1	E	54	TYR	2.3
1	C	152	PHE	2.3

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	17	SER	2.2
1	D	-7	TYR	2.2
1	C	98	PRO	2.2
1	E	67	TRP	2.2
1	A	123	LEU	2.2
1	D	123	LEU	2.2
1	D	68	ASP	2.2
1	B	71	GLU	2.2
1	A	98	PRO	2.2
1	A	71	GLU	2.2
1	B	99	VAL	2.2
1	B	97[A]	ARG	2.1
1	D	102	LEU	2.1
1	C	120	ALA	2.1
1	B	147	TRP	2.1
1	B	120	ALA	2.1
1	B	53	VAL	2.1
1	E	98	PRO	2.1
1	E	59[A]	ARG	2.1
1	B	13	LEU	2.1
1	A	130	THR	2.0
1	D	97[A]	ARG	2.0
1	E	119	PRO	2.0
1	A	70	ASN	2.0
1	E	207[A]	ARG	2.0

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

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

## 6.3 Carbohydrates ⓘ

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. 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	NAG	C	402	14/15	0.24	2.84	44,47,49,49	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
3	MAN	C	405	11/12	0.30	1.09	39,40,42,42	0
3	BMA	C	404	11/12	0.22	0.47	39,40,41,42	0
3	MAN	C	406	11/12	0.43	-	43,44,44,44	0
3	NAG	C	403	14/15	0.21	-	44,47,50,50	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( $\text{\AA}^2$ )	Q<0.9
5	PG4	B	403	13/13	0.40	9.56	78,85,88,88	0
5	PG4	C	407	10/13	0.28	2.33	44,45,47,47	0
4	COC	D	401	22/22	0.22	2.26	32,34,37,38	0
2	NAG	B	402	14/15	0.35	1.85	42,43,44,44	0
5	PG4	E	403	13/13	0.23	1.15	57,58,60,60	0
4	COC	A	401	22/22	0.14	-0.08	20,22,26,27	0

## 6.5 Other polymers

There are no such residues in this entry.