



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 14, 2017 – 11:13 am GMT

PDB ID : 4XP9
Title : X-ray structure of Drosophila dopamine transporter bound to psychostimulant
D-amphetamine
Authors : Aravind, P.; Wang, K.; Gouaux, E.
Deposited on : 2015-01-16
Resolution : 2.80 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at

<http://wwpdb.org/validation/2016/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 : 1.7.2 (RC1), CSD as538be (2017)
Xtriage (Phenix) : 1.9-1692
EDS : trunk28620
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)
Refmac : 5.8.0135
CCP4 : 6.5.0
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : recalc28949

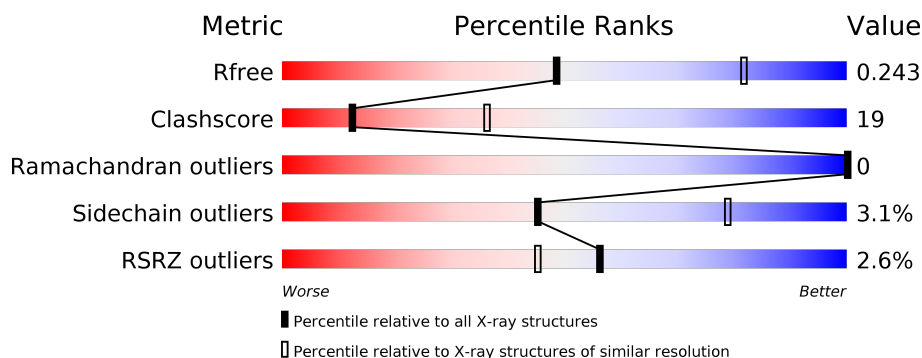
1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.80 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}	100719	2583 (2.80-2.80)
Clashscore	112137	3033 (2.80-2.80)
Ramachandran outliers	110173	2983 (2.80-2.80)
Sidechain outliers	110143	2985 (2.80-2.80)
RSRZ outliers	101464	2610 (2.80-2.80)

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 ≥ 3 , 2, 1 and 0 types of geometric quality criteria. 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	C	541	<div> <div>2%</div> <div> <div></div> <div>69%</div> <div>29%</div> <div>..</div> </div> </div>
2	L	213	<div> <div>%</div> <div> <div></div> <div>69%</div> <div>30%</div> <div>.</div> </div> </div>
3	H	219	<div> <div>5%</div> <div> <div></div> <div>71%</div> <div>28%</div> </div> </div>

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
11	NA	C	710	-	-	-	X

2 Entry composition

There are 13 unique types of molecules in this entry. The entry contains 7701 atoms, of which 36 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called Transporter.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	C	537	Total	C	N	O	S	0	0	0
			4229	2835	655	721	18			

There are 48 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	74	ALA	VAL	engineered mutation	UNP A0A0B4KEX2
C	?	-	SER	deletion	UNP A0A0B4KEX2
C	?	-	GLN	deletion	UNP A0A0B4KEX2
C	?	-	ASN	deletion	UNP A0A0B4KEX2
C	?	-	ALA	deletion	UNP A0A0B4KEX2
C	?	-	SER	deletion	UNP A0A0B4KEX2
C	?	-	ARG	deletion	UNP A0A0B4KEX2
C	?	-	VAL	deletion	UNP A0A0B4KEX2
C	?	-	PRO	deletion	UNP A0A0B4KEX2
C	?	-	VAL	deletion	UNP A0A0B4KEX2
C	?	-	ILE	deletion	UNP A0A0B4KEX2
C	?	-	GLY	deletion	UNP A0A0B4KEX2
C	?	-	ASN	deletion	UNP A0A0B4KEX2
C	?	-	TYR	deletion	UNP A0A0B4KEX2
C	?	-	SER	deletion	UNP A0A0B4KEX2
C	?	-	ASP	deletion	UNP A0A0B4KEX2
C	?	-	LEU	deletion	UNP A0A0B4KEX2
C	?	-	TYR	deletion	UNP A0A0B4KEX2
C	?	-	ALA	deletion	UNP A0A0B4KEX2
C	?	-	MET	deletion	UNP A0A0B4KEX2
C	?	-	GLY	deletion	UNP A0A0B4KEX2
C	?	-	ASN	deletion	UNP A0A0B4KEX2
C	?	-	GLN	deletion	UNP A0A0B4KEX2
C	?	-	SER	deletion	UNP A0A0B4KEX2
C	?	-	LEU	deletion	UNP A0A0B4KEX2
C	?	-	LEU	deletion	UNP A0A0B4KEX2
C	?	-	TYR	deletion	UNP A0A0B4KEX2

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	?	-	ASN	deletion	UNP A0A0B4KEX2
C	?	-	GLU	deletion	UNP A0A0B4KEX2
C	?	-	THR	deletion	UNP A0A0B4KEX2
C	?	-	TYR	deletion	UNP A0A0B4KEX2
C	?	-	MET	deletion	UNP A0A0B4KEX2
C	?	-	ASN	deletion	UNP A0A0B4KEX2
C	?	-	GLY	deletion	UNP A0A0B4KEX2
C	?	-	SER	deletion	UNP A0A0B4KEX2
C	?	-	SER	deletion	UNP A0A0B4KEX2
C	?	-	LEU	deletion	UNP A0A0B4KEX2
C	?	-	ASP	deletion	UNP A0A0B4KEX2
C	?	-	THR	deletion	UNP A0A0B4KEX2
C	?	-	SER	deletion	UNP A0A0B4KEX2
C	?	-	ALA	deletion	UNP A0A0B4KEX2
C	?	-	VAL	deletion	UNP A0A0B4KEX2
C	415	ALA	LEU	engineered mutation	UNP A0A0B4KEX2
C	602	LEU	-	expression tag	UNP A0A0B4KEX2
C	603	LEU	-	expression tag	UNP A0A0B4KEX2
C	604	VAL	-	expression tag	UNP A0A0B4KEX2
C	605	PRO	-	expression tag	UNP A0A0B4KEX2
C	606	ARG	-	expression tag	UNP A0A0B4KEX2

- Molecule 2 is a protein called ANTIBODY FRAGMENT HEAVY CHAIN-PROTEIN, 9D5-HEAVY CHAIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	L	213	Total	C	N	O	S	0	0	0
			1615	1005	267	335	8			

- Molecule 3 is a protein called ANTIBODY FRAGMENT LIGHT CHAIN-PROTEIN, 9D5-LIGHT CHAIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	H	219	Total	C	N	O	S	0	0	0
			1640	1031	277	324	8			

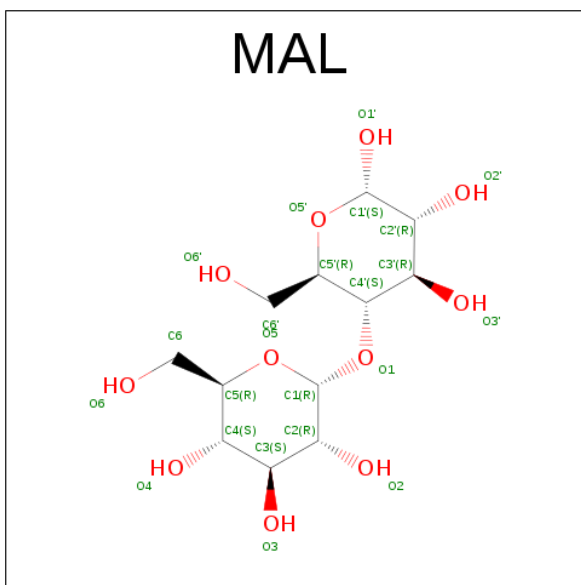
- Molecule 4 is CHOLESTEROL HEMISUCCINATE (three-letter code: Y01) (formula: C₃₁H₅₀O₄).



- Molecule 5 is N-ACETYL-D-GLUCOSAMINE (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).

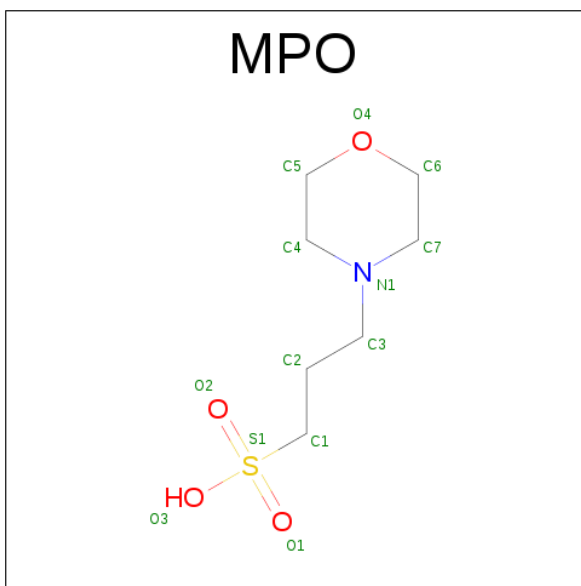


- Molecule 6 is MALTOSE (three-letter code: MAL) (formula: $C_{12}H_{22}O_{11}$).



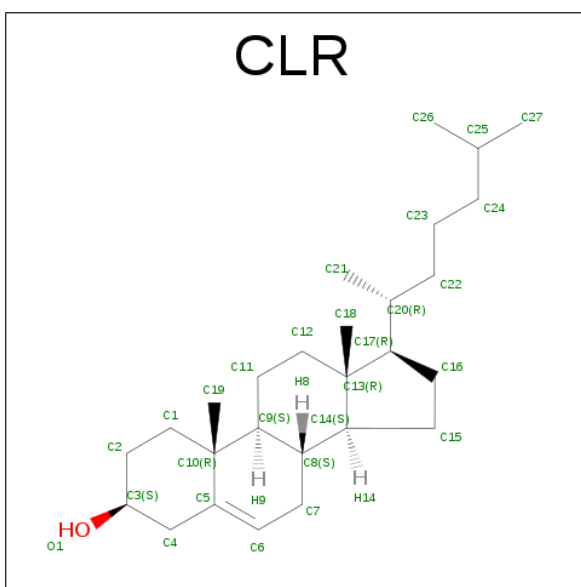
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	C	1	Total	C	O	0	0
			23	12	11		

- Molecule 7 is 3[N-MORPHOLINO]PROPANE SULFONIC ACID (three-letter code: MPO) (formula: $C_7H_{15}NO_4S$).



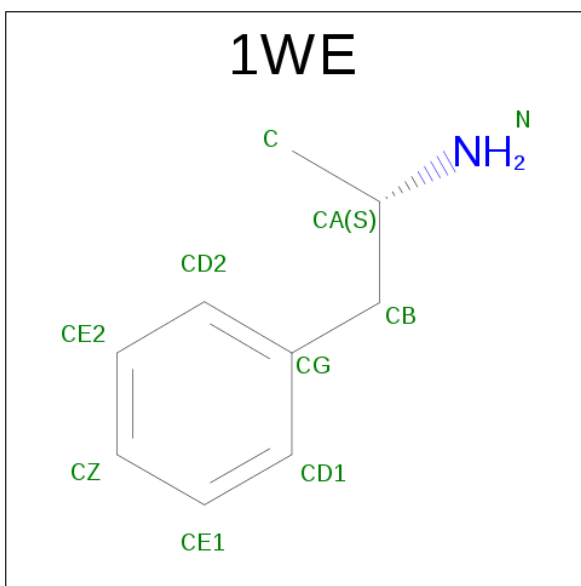
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
7	C	1	Total	C	N	O	S	0	0
			13	7	1	4	1		

- Molecule 8 is CHOLESTEROL (three-letter code: CLR) (formula: $C_{27}H_{46}O$).



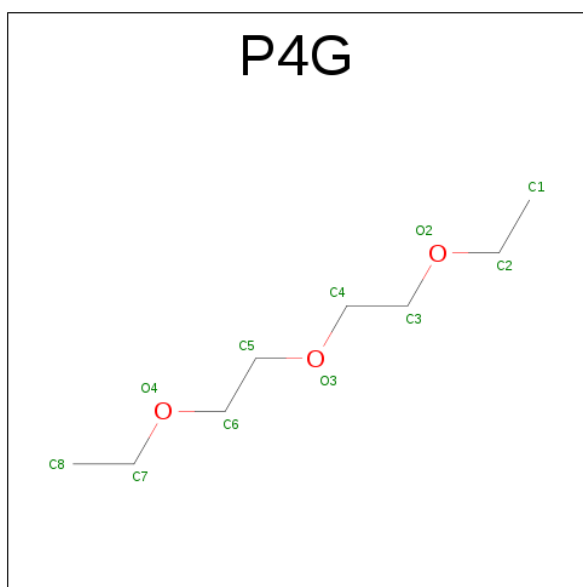
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
8	C	1	Total	C	O	0	0
			28	27	1		

- Molecule 9 is (2S)-1-phenylpropan-2-amine (three-letter code: 1WE) (formula: $C_9H_{13}N$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
9	C	1	Total	C	N	0	0
			10	9	1		

- Molecule 10 is 1-ETHOXY-2-(2-ETHOXYETHOXY)ETHANE (three-letter code: P4G) (formula: $C_8H_{18}O_3$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
10	C	1	Total	C	H	O	0	0
			29	8	18	3		
10	C	1	Total	C	H	O	0	0
			29	8	18	3		

- Molecule 11 is SODIUM ION (three-letter code: NA) (formula: Na).

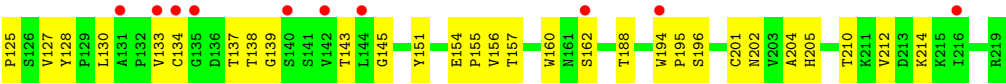
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
11	C	2	Total	Na	0	0
			2	2		

- Molecule 12 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
12	C	1	Total	Cl	0	0
			1	1		

- Molecule 13 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
13	C	16	Total	O	0	0
			16	16		
13	L	9	Total	O	0	0
			9	9		
13	H	8	Total	O	0	0
			8	8		



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	97.35Å 140.75Å 166.76Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	46.72 – 2.80 46.72 – 2.80	Depositor EDS
% Data completeness (in resolution range)	96.8 (46.72-2.80) 96.6 (46.72-2.80)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.17 (at 2.81Å)	Xtriage
Refinement program	PHENIX	Depositor
R, R_{free}	0.211 , 0.242 0.209 , 0.243	Depositor DCC
R_{free} test set	2693 reflections (4.86%)	DCC
Wilson B-factor (Å ²)	86.2	Xtriage
Anisotropy	0.049	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.32 , 62.6	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	7701	wwPDB-VP
Average B, all atoms (Å ²)	85.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.29% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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 [i](#)

5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: Y01, NAG, CL, P4G, NA, 1WE, MPO, MAL, CLR

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	C	0.40	0/4371	0.53	0/5977
2	L	0.39	0/1653	0.58	0/2248
3	H	0.40	0/1679	0.60	0/2289
All	All	0.40	0/7703	0.56	0/10514

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 [i](#)

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	C	4229	0	4130	157	0
2	L	1615	0	1523	61	0
3	H	1640	0	1585	70	0
4	C	35	0	49	2	0
5	C	14	0	13	1	0
6	C	23	0	22	4	0
7	C	13	0	14	1	0
8	C	28	0	46	0	0
9	C	10	0	13	2	0
10	C	22	36	36	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
11	C	2	0	0	0	0
12	C	1	0	0	0	0
13	C	16	0	0	2	0
13	H	8	0	0	0	0
13	L	9	0	0	0	0
All	All	7665	36	7431	283	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 19.

The worst 5 of 283 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)
2:L:121:PRO:HB2	2:L:126:LEU:HD11	1.37	1.05
2:L:92:PHE:HE2	3:H:105:GLY:HA2	1.23	1.02
3:H:137:THR:HA	3:H:138:THR:HG23	1.48	0.96
3:H:138:THR:HB	3:H:139:GLY:HA3	1.48	0.94
1:C:133:ARG:HH22	6:C:703:MAL:H6'2	1.33	0.93

There are no symmetry-related clashes.

5.3 Torsion angles [i](#)

5.3.1 Protein backbone [i](#)

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	C	535/541 (99%)	523 (98%)	12 (2%)	0	100	100
2	L	211/213 (99%)	203 (96%)	8 (4%)	0	100	100
3	H	217/219 (99%)	210 (97%)	7 (3%)	0	100	100
All	All	963/973 (99%)	936 (97%)	27 (3%)	0	100	100

There are no Ramachandran outliers to report.

5.3.2 Protein sidechains ⓘ

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	C	432/446 (97%)	416 (96%)	16 (4%)	39	73
2	L	182/186 (98%)	175 (96%)	7 (4%)	38	72
3	H	181/187 (97%)	179 (99%)	2 (1%)	78	94
All	All	795/819 (97%)	770 (97%)	25 (3%)	45	79

5 of 25 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	482	SER
1	C	566	SER
3	H	96	CYS
1	C	529	ILE
1	C	582	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
2	L	54	ASN

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 ⓘ

There are no carbohydrates in this entry.

5.6 Ligand geometry

Of 11 ligands modelled in this entry, 3 are monoatomic - leaving 8 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
4	Y01	C	701	-	35,38,38	4.32	13 (37%)	54,57,57	1.97	16 (29%)
5	NAG	C	702	1	14,14,15	0.89	1 (7%)	15,19,21	0.79	1 (6%)
6	MAL	C	703	-	24,24,24	0.55	0	35,35,35	0.89	0
7	MPO	C	704	-	13,13,13	2.70	2 (15%)	16,17,17	1.79	4 (25%)
8	CLR	C	705	-	31,31,31	0.66	0	48,48,48	1.36	7 (14%)
9	1WE	C	706	-	10,10,10	0.46	0	10,12,12	0.54	0
10	P4G	C	707	-	10,10,10	0.82	0	9,9,9	0.38	0
10	P4G	C	708	-	10,10,10	0.88	0	9,9,9	0.46	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	Y01	C	701	-	-	0/17/77/77	0/4/4/4
5	NAG	C	702	1	-	0/6/23/26	0/1/1/1
6	MAL	C	703	-	-	0/8/48/48	0/2/2/2
7	MPO	C	704	-	-	0/7/15/15	0/1/1/1
8	CLR	C	705	-	-	0/10/68/68	0/4/4/4
9	1WE	C	706	-	-	0/4/4/4	0/1/1/1
10	P4G	C	707	-	-	0/8/8/8	0/0/0/0
10	P4G	C	708	-	-	0/8/8/8	0/0/0/0

The worst 5 of 16 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	C	701	Y01	CBB-CBE	-9.73	1.36	1.54
4	C	701	Y01	CAU-CBI	-6.79	1.41	1.54
7	C	704	MPO	C1-S1	-5.68	1.69	1.77

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	C	701	Y01	CBH-CAZ	-4.24	1.44	1.52
4	C	701	Y01	CAE-CBI	-2.03	1.50	1.54

The worst 5 of 28 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	C	701	Y01	CBI-CBE-CBB	-3.95	113.12	119.47
4	C	701	Y01	CAE-CBI-CBE	-3.83	104.47	111.73
4	C	701	Y01	CAK-CAI-CAZ	-3.75	117.77	125.06
4	C	701	Y01	CAE-CBI-CBG	-3.39	105.32	111.73
4	C	701	Y01	CAD-CBH-CBF	-3.38	107.50	111.68

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

6 monomers are involved in 11 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	C	701	Y01	2	0
5	C	702	NAG	1	0
6	C	703	MAL	4	0
7	C	704	MPO	1	0
9	C	706	1WE	2	0
10	C	707	P4G	1	0

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 [i](#)

6.1 Protein, DNA and RNA chains [i](#)

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, 95th 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 ²)	Q<0.9
1	C	537/541 (99%)	0.03	12 (2%) 62 52	65, 83, 109, 145	0
2	L	213/213 (100%)	-0.08	3 (1%) 75 69	63, 80, 109, 129	0
3	H	219/219 (100%)	0.11	10 (4%) 33 23	66, 81, 121, 161	0
All	All	969/973 (99%)	0.03	25 (2%) 56 45	63, 82, 111, 161	0

The worst 5 of 25 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	586	LEU	6.0
3	H	142	VAL	4.4
3	H	134	CYS	4.1
3	H	216	ILE	3.8
3	H	144	LEU	3.5

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

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

6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

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. 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, 95th 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	LLDF	B-factors(\AA^2)	Q<0.9
11	NA	C	710	1/1	0.94	0.29	2.20	70,70,70,70	0
8	CLR	C	705	28/28	0.94	0.27	1.99	75,89,94,98	0
11	NA	C	709	1/1	0.94	0.28	1.72	72,72,72,72	0
4	Y01	C	701	35/35	0.93	0.20	0.47	75,95,111,117	0
6	MAL	C	703	23/23	0.92	0.17	-0.11	89,98,103,104	0
9	1WE	C	706	10/10	0.97	0.22	-0.39	75,79,84,84	0
7	MPO	C	704	13/13	0.84	0.19	-0.63	94,108,136,150	0
10	P4G	C	707	11/11	0.89	0.15	-0.86	111,133,145,149	0
12	CL	C	711	1/1	0.94	0.18	-1.31	77,77,77,77	0
5	NAG	C	702	14/15	0.79	0.34	-	114,118,122,122	0
10	P4G	C	708	11/11	0.59	0.25	-	109,133,139,140	0

6.5 Other polymers [i](#)

There are no such residues in this entry.