



# Full wwPDB X-ray Structure Validation Report ⓘ

Mar 21, 2016 – 06:47 PM EDT

PDB ID : 5DSG  
Title : Structure of the M4 muscarinic acetylcholine receptor (M4-mT4L) bound to tiotropium  
Authors : Thal, D.M.; Kobilka, B.K.; Sexton, P.M.; Christopoulos, A.  
Deposited on : 2015-09-17  
Resolution : 2.60 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.  
We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)  
A user guide is available at  
<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 : 1.7.1 (RC1), CSD as537be (2016)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20027107  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Refmac : 5.8.0122  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : rb-20027107

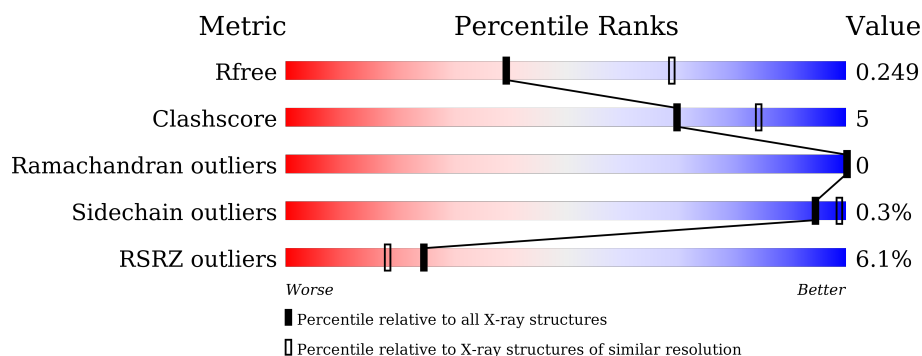
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.60 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))
$R_{free}$	91344	2328 (2.60-2.60)
Clashscore	102246	2679 (2.60-2.60)
Ramachandran outliers	100387	2635 (2.60-2.60)
Sidechain outliers	100360	2635 (2.60-2.60)
RSRZ outliers	91569	2334 (2.60-2.60)

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. 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	A	422	
1	B	422	

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
3	OLA	A	1203	-	-	-	X
3	OLA	A	1209	-	-	-	X
3	OLA	B	1203	-	-	-	X
3	OLA	B	1204	-	-	-	X
5	P6G	A	1205	-	-	-	X
5	P6G	B	1205	-	-	-	X
6	OLC	A	1206	-	-	-	X
6	OLC	A	1207	-	-	-	X
6	OLC	A	1208	-	-	-	X
6	OLC	B	1202	-	-	-	X
6	OLC	B	1206	-	-	-	X

## 2 Entry composition

There are 8 unique types of molecules in this entry. The entry contains 6516 atoms, of which 0 are hydrogens and 0 are deuteriums.

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

- Molecule 1 is a protein called Muscarinic acetylcholine receptor M4,Endolysin,Endolysin,Muscarinic acetylcholine receptor M4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	392	Total	C	N	O	S	0	0	0
			3058	2004	499	533	22			
1	B	399	Total	C	N	O	S	0	0	0
			3115	2041	512	541	21			

There are 38 discrepancies between the modelled and reference sequences:

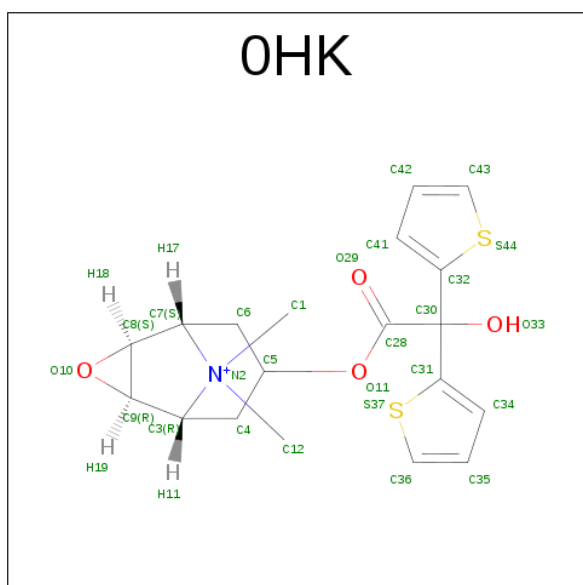
Chain	Residue	Modelled	Actual	Comment	Reference
A	20	GLY	-	expression tag	UNP P08173
A	21	PRO	-	expression tag	UNP P08173
A	1012	GLY	-	linker	UNP P00720
A	1013	GLY	-	linker	UNP P00720
A	1014	GLY	-	linker	UNP P00720
A	1015	SER	-	linker	UNP P00720
A	1016	GLY	-	linker	UNP P00720
A	1017	GLY	-	linker	UNP P00720
A	1054	ALA	CYS	engineered mutation	UNP P00720
A	1094	ARG	ILE	engineered mutation	UNP P00720
A	479	ARG	-	expression tag	UNP P08173
A	480	HIS	-	expression tag	UNP P08173
A	481	HIS	-	expression tag	UNP P08173
A	482	HIS	-	expression tag	UNP P08173
A	483	HIS	-	expression tag	UNP P08173
A	484	HIS	-	expression tag	UNP P08173
A	485	HIS	-	expression tag	UNP P08173
A	486	HIS	-	expression tag	UNP P08173
A	487	HIS	-	expression tag	UNP P08173
B	20	GLY	-	expression tag	UNP P08173
B	21	PRO	-	expression tag	UNP P08173
B	1012	GLY	-	linker	UNP P00720
B	1013	GLY	-	linker	UNP P00720
B	1014	GLY	-	linker	UNP P00720

*Continued on next page...*

Continued from previous page...

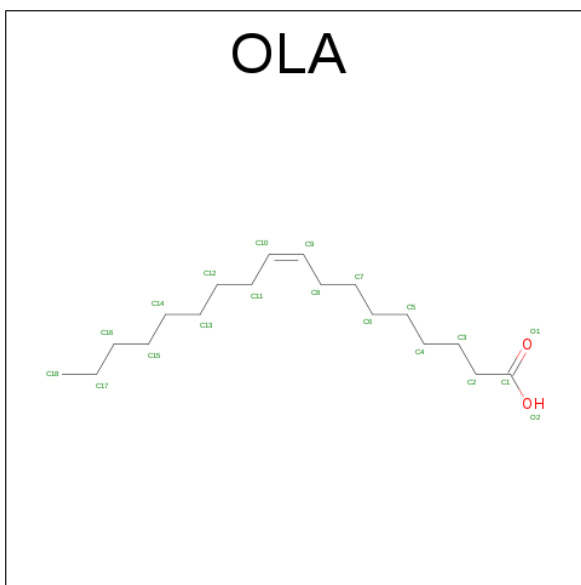
Chain	Residue	Modelled	Actual	Comment	Reference
B	1015	SER	-	linker	UNP P00720
B	1016	GLY	-	linker	UNP P00720
B	1017	GLY	-	linker	UNP P00720
B	1054	ALA	CYS	conflict	UNP P00720
B	1094	ARG	ILE	conflict	UNP P00720
B	479	ARG	-	expression tag	UNP P08173
B	480	HIS	-	expression tag	UNP P08173
B	481	HIS	-	expression tag	UNP P08173
B	482	HIS	-	expression tag	UNP P08173
B	483	HIS	-	expression tag	UNP P08173
B	484	HIS	-	expression tag	UNP P08173
B	485	HIS	-	expression tag	UNP P08173
B	486	HIS	-	expression tag	UNP P08173
B	487	HIS	-	expression tag	UNP P08173

- Molecule 2 is (1R,2R,4S,5S,7S)-7-{[hydroxy(dithiophen-2-yl)acetyl]oxy}-9,9-dimethyl-3-oxa-9-azoniatricyclo[3.3.1.0 2,4 ]nonane (three-letter code: OHK) (formula: C<sub>19</sub>H<sub>22</sub>NO<sub>4</sub>S<sub>2</sub>).



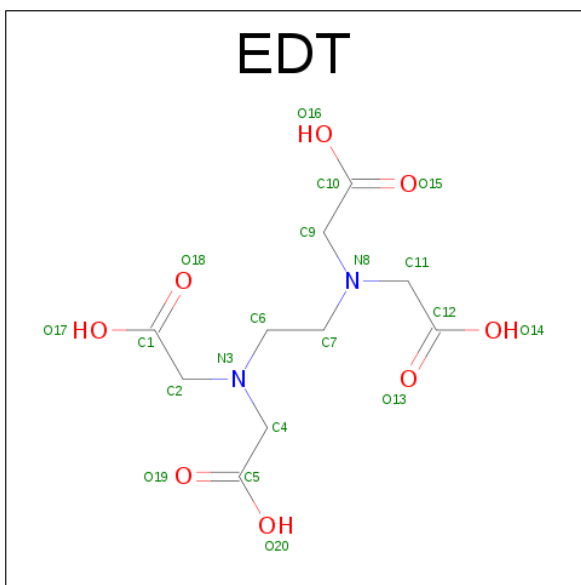
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	N	O	S	0	0
			26	19	1	4	2		
2	B	1	Total	C	N	O	S	0	0
			26	19	1	4	2		

- Molecule 3 is OLEIC ACID (three-letter code: OLA) (formula: C<sub>18</sub>H<sub>34</sub>O<sub>2</sub>).



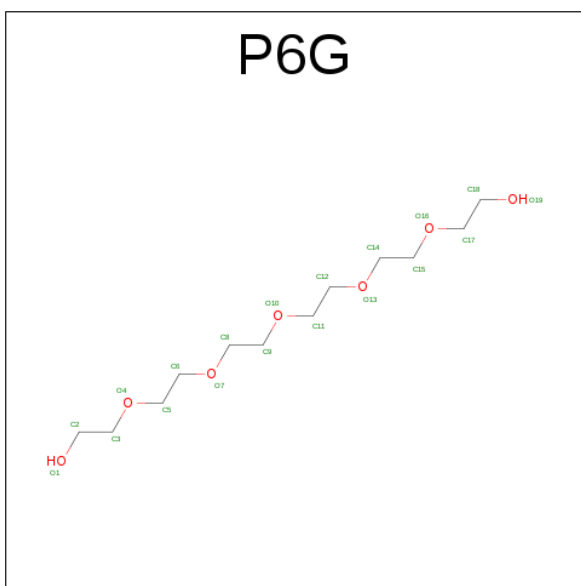
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	O	0	0
			12	10	2		
3	A	1	Total	C	O	0	0
			14	12	2		
3	A	1	Total	C	O	0	0
			20	18	2		
3	B	1	Total	C	O	0	0
			20	18	2		
3	B	1	Total	C	O	0	0
			20	18	2		

- Molecule 4 is {[-(BIS-CARBOXYMETHYL-AMINO)-ETHYL]-CARBOXYMETHYL-AMINO}-ACETIC ACID (three-letter code: EDT) (formula: C<sub>10</sub>H<sub>16</sub>N<sub>2</sub>O<sub>8</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	1	Total	C	N	O	0	0
			20	10	2	8		

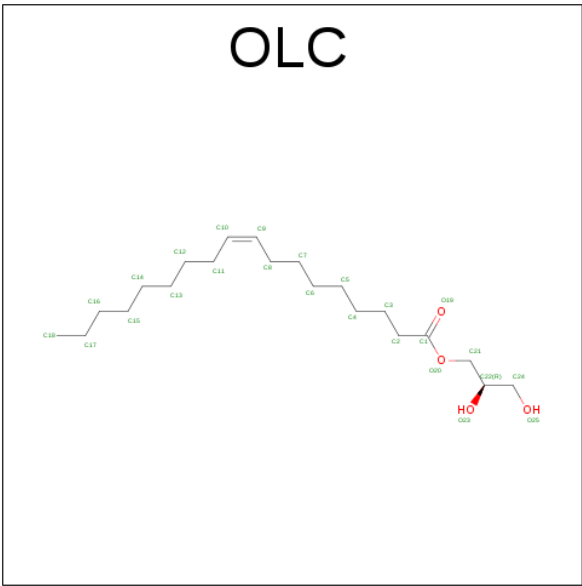
- Molecule 5 is HEXAETHYLENE GLYCOL (three-letter code: P6G) (formula:  $C_{12}H_{26}O_7$ ).



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

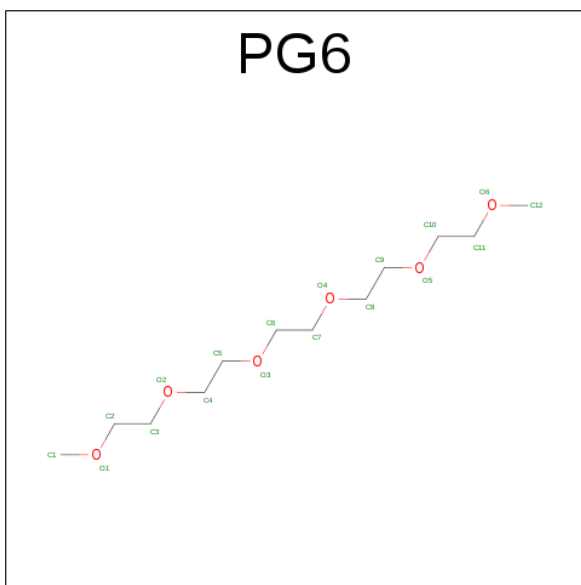
- Molecule 6 is (2R)-2,3-dihydroxypropyl (9Z)-octadec-9-enoate (three-letter code: OLC)

(formula: C<sub>21</sub>H<sub>40</sub>O<sub>4</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	A	1	Total	C	O	0	0
			25	21	4		
6	A	1	Total	C	O	0	0
			18	14	4		
6	A	1	Total	C	O	0	0
			25	21	4		
6	B	1	Total	C	O	0	0
			17	13	4		
6	B	1	Total	C	O	0	0
			25	21	4		

- Molecule 7 is 1-(2-METHOXY-ETHOXY)-2-{2-[2-(2-METHOXY-ETHOXY)-ETHOXY]-ETHANE (three-letter code: PG6) (formula: C<sub>12</sub>H<sub>26</sub>O<sub>6</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	B	1	Total	C	O	0	0
			10	6	4		

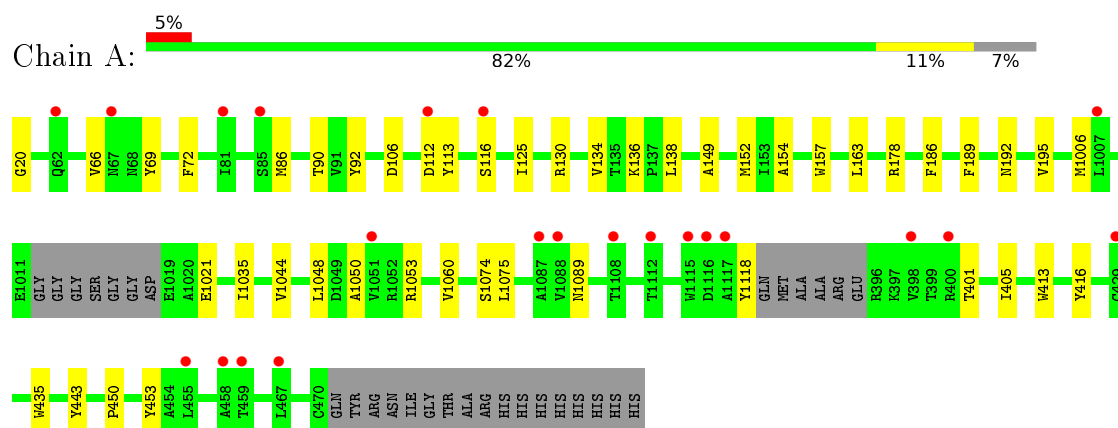
- Molecule 8 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	A	17	Total O 17 17	0	0
8	B	10	Total O 10 10	0	0

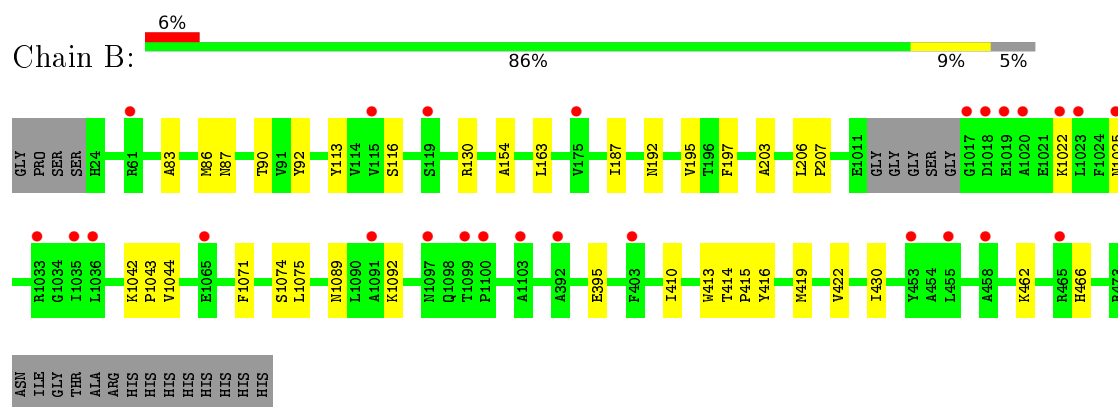
### 3 Residue-property plots

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: Muscarinic acetylcholine receptor M4,Endolysin,Endolysin,Muscarinic acetylcholine receptor M4



- Molecule 1: Muscarinic acetylcholine receptor M4,Endolysin,Endolysin,Muscarinic acetylcholine receptor M4



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	48.49 Å 172.04 Å 60.66 Å 90.00° 94.37° 90.00°	Depositor
Resolution (Å)	28.53 – 2.60 28.99 – 2.60	Depositor EDS
% Data completeness (in resolution range)	99.7 (28.53-2.60) 94.9 (28.99-2.60)	Depositor EDS
$R_{merge}$	0.21	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.93 (at 2.61 Å)	Xtriage
Refinement program	PHENIX	Depositor
R, $R_{free}$	0.227 , 0.240 0.236 , 0.249	Depositor DCC
$R_{free}$ test set	2261 reflections (8.47%)	DCC
Wilson B-factor (Å <sup>2</sup> )	50.2	Xtriage
Anisotropy	0.245	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 63.1	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	1 of 30337 reflections (0.003%)	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	6516	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	72.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.73% of the height of the origin peak. No significant pseudotranslation is detected.*

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

<sup>2</sup> Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 0.2 for perfectly twinned datasets.

## 5 Model quality

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: 0HK, OLA, OLC, EDT, PG6, 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	0.22	0/3132	0.38	0/4283
1	B	0.22	0/3190	0.38	0/4362
All	All	0.22	0/6322	0.38	0/8645

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

### 5.2 Too-close contacts

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	3058	0	3044	29	0
1	B	3115	0	3104	24	0
2	A	26	0	22	3	0
2	B	26	0	22	3	0
3	A	46	0	66	0	0
3	B	40	0	66	2	0
4	A	20	0	12	3	0
5	A	19	0	26	2	0
5	B	19	0	26	1	0
6	A	68	0	103	6	0
6	B	42	0	62	6	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
7	B	10	0	12	1	0
8	A	17	0	0	0	0
8	B	10	0	0	0	0
All	All	6516	0	6565	59	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 5.

All (59) 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)
1:A:136:LYS:HD2	6:A:1207:OLC:H24A	1.79	0.64
1:A:106:ASP:OD1	1:A:178:ARG:NH2	2.32	0.63
1:A:435:TRP:HB3	5:A:1205:P6G:H141	1.81	0.63
1:A:1044:VAL:HG21	1:A:1075:LEU:HB3	1.82	0.61
1:A:186:PHE:HA	5:A:1205:P6G:H61	1.84	0.60
1:B:462:LYS:O	1:B:466:HIS:ND1	2.35	0.60
1:A:416:TYR:HE2	2:A:1201:0HK:H1	1.69	0.58
1:A:20:GLY:N	4:A:1204:EDT:O15	2.38	0.57
3:B:1203:OLA:H71	3:B:1204:OLA:H31	1.84	0.57
1:B:416:TYR:HE2	2:B:1201:0HK:H1	1.70	0.56
1:A:450:PRO:HA	1:A:453:TYR:HD2	1.71	0.56
1:A:72:PHE:HD2	6:A:1206:OLC:H4A	1.73	0.52
1:B:203:ALA:HA	6:B:1206:OLC:H18A	1.92	0.52
1:B:422:VAL:HB	1:B:430:ILE:HD11	1.92	0.52
1:A:192:ASN:HD22	1:A:195:VAL:HG23	1.75	0.51
1:B:130:ARG:HH12	1:B:395:GLU:HB3	1.75	0.51
6:A:1206:OLC:H7	6:B:1202:OLC:H3	1.93	0.50
1:B:1071:PHE:HA	7:B:1207:PG6:H62	1.94	0.49
1:B:187:ILE:H	5:B:1205:P6G:H21	1.77	0.49
1:A:138:LEU:HD12	1:A:1021:GLU:HG2	1.95	0.48
1:B:192:ASN:HD22	1:B:195:VAL:HG23	1.77	0.48
1:A:69:TYR:CE1	6:A:1206:OLC:H21	2.49	0.47
1:B:1074:SER:OG	1:B:1089:ASN:ND2	2.37	0.47
1:B:1022:LYS:O	1:B:1025:ASN:HB2	2.14	0.47
1:A:113:TYR:CZ	2:A:1201:0HK:H19	2.50	0.47
1:A:154:ALA:HA	6:B:1202:OLC:H4A	1.97	0.46
4:A:1204:EDT:H111	4:A:1204:EDT:H061	1.74	0.46
1:A:189:PHE:HE1	6:A:1208:OLC:H5A	1.81	0.46
1:B:154:ALA:HA	3:B:1203:OLA:H61	1.96	0.46
1:A:1050:ALA:HA	1:A:1053:ARG:HG2	1.97	0.46

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:157:TRP:HB3	6:B:1202:OLC:H8	1.98	0.45
1:A:1074:SER:OG	1:A:1089:ASN:ND2	2.39	0.45
1:B:163:LEU:HD13	6:B:1206:OLC:H14	1.97	0.45
1:A:163:LEU:HD13	6:A:1208:OLC:H15A	1.99	0.45
1:A:130:ARG:O	1:A:134:VAL:HG23	2.17	0.45
1:B:1022:LYS:HA	1:B:1025:ASN:HD22	1.80	0.45
1:A:86:MET:O	1:A:90:THR:HG23	2.17	0.45
1:A:125:ILE:HD13	1:A:152:MET:HB3	2.00	0.44
1:B:86:MET:O	1:B:90:THR:HG23	2.17	0.44
1:A:116:SER:HB3	1:A:413:TRP:HH2	1.83	0.44
1:B:206:LEU:HB3	1:B:207:PRO:HD3	2.01	0.43
1:B:116:SER:HB3	1:B:413:TRP:HH2	1.82	0.43
1:B:83:ALA:O	1:B:87:ASN:ND2	2.51	0.43
1:B:415:PRO:O	1:B:419:MET:HG3	2.19	0.43
1:B:1089:ASN:HA	1:B:1092:LYS:HE2	2.00	0.43
4:A:1204:EDT:H072	4:A:1204:EDT:H022	1.58	0.43
1:A:66:VAL:HG13	1:A:149:ALA:HB2	2.01	0.43
1:A:1035:ILE:HG12	1:A:1060:VAL:HG21	2.01	0.42
1:B:1044:VAL:HG21	1:B:1075:LEU:HD22	2.02	0.42
1:B:410:ILE:O	1:B:414:THR:OG1	2.29	0.42
2:B:1201:0HK:H20	2:B:1201:0HK:O10	2.20	0.41
1:B:113:TYR:CZ	2:B:1201:0HK:H19	2.56	0.41
1:B:1042:LYS:HB3	1:B:1043:PRO:HD3	2.03	0.41
1:A:1006:MET:HG3	1:A:1118:TYR:CZ	2.56	0.41
2:A:1201:0HK:H20	2:A:1201:0HK:O10	2.21	0.41
1:A:1048:LEU:HD23	1:A:1053:ARG:HA	2.03	0.41
1:A:112:ASP:OD2	1:A:443:TYR:OH	2.35	0.40
1:A:401:THR:O	1:A:405:ILE:HG12	2.22	0.40
1:B:197:PHE:CE2	6:B:1202:OLC:H5A	2.57	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	386/422 (92%)	383 (99%)	3 (1%)	0	100	100
1	B	395/422 (94%)	391 (99%)	4 (1%)	0	100	100
All	All	781/844 (92%)	774 (99%)	7 (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	324/361 (90%)	323 (100%)	1 (0%)	94	99
1	B	327/361 (91%)	326 (100%)	1 (0%)	94	99
All	All	651/722 (90%)	649 (100%)	2 (0%)	94	99

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	92	TYR
1	B	92	TYR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	1089	ASN
1	B	1025	ASN
1	B	1080	GLN
1	B	1089	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 ⓘ

16 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$
2	0HK	A	1201	-	30,30,30	2.20	8 (26%)	32,48,48	5.50	17 (53%)
3	OLA	A	1202	-	8,11,19	0.99	1 (12%)	8,11,19	0.70	0
3	OLA	A	1203	-	10,13,19	0.89	1 (10%)	10,13,19	0.77	0
4	EDT	A	1204	-	7,19,19	0.19	0	12,24,24	1.51	3 (25%)
5	P6G	A	1205	-	18,18,18	0.47	0	17,17,17	0.33	0
6	OLC	A	1206	-	24,24,24	0.92	1 (4%)	25,25,25	0.86	1 (4%)
6	OLC	A	1207	-	17,17,24	1.07	1 (5%)	18,18,25	1.06	1 (5%)
6	OLC	A	1208	-	24,24,24	0.92	1 (4%)	25,25,25	0.84	1 (4%)
3	OLA	A	1209	-	16,19,19	0.70	1 (6%)	16,19,19	0.57	0
2	0HK	B	1201	-	30,30,30	2.21	8 (26%)	32,48,48	5.46	16 (50%)
6	OLC	B	1202	-	16,16,24	1.11	1 (6%)	17,17,25	1.06	1 (5%)
3	OLA	B	1203	-	16,19,19	0.70	1 (6%)	16,19,19	0.55	0
3	OLA	B	1204	-	16,19,19	0.70	1 (6%)	16,19,19	0.55	0
5	P6G	B	1205	-	18,18,18	0.47	0	17,17,17	0.33	0
6	OLC	B	1206	-	24,24,24	0.92	1 (4%)	25,25,25	0.79	1 (4%)
7	PG6	B	1207	-	9,9,17	0.46	0	8,8,16	0.36	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	0HK	A	1201	-	-	0/10/55/55	0/2/5/5
3	OLA	A	1202	-	-	0/7/9/17	0/0/0/0
3	OLA	A	1203	-	-	0/9/11/17	0/0/0/0
4	EDT	A	1204	-	-	0/13/21/21	0/0/0/0
5	P6G	A	1205	-	-	0/16/16/16	0/0/0/0
6	OLC	A	1206	-	-	0/24/24/24	0/0/0/0
6	OLC	A	1207	-	-	0/17/17/24	0/0/0/0
6	OLC	A	1208	-	-	0/24/24/24	0/0/0/0
3	OLA	A	1209	-	-	0/15/17/17	0/0/0/0
2	0HK	B	1201	-	-	0/10/55/55	0/2/5/5
6	OLC	B	1202	-	-	0/16/16/24	0/0/0/0
3	OLA	B	1203	-	-	0/15/17/17	0/0/0/0
3	OLA	B	1204	-	-	0/15/17/17	0/0/0/0
5	P6G	B	1205	-	-	0/16/16/16	0/0/0/0
6	OLC	B	1206	-	-	0/24/24/24	0/0/0/0
7	PG6	B	1207	-	-	0/7/7/15	0/0/0/0

All (26) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	1201	0HK	C3-N2	-6.09	1.40	1.53
2	A	1201	0HK	C3-N2	-6.03	1.40	1.53
2	B	1201	0HK	C7-N2	-5.51	1.41	1.53
2	A	1201	0HK	C7-N2	-5.50	1.42	1.53
2	B	1201	0HK	C1-N2	-4.15	1.43	1.51
2	A	1201	0HK	C1-N2	-4.14	1.43	1.51
2	A	1201	0HK	C12-N2	-2.61	1.46	1.51
2	B	1201	0HK	C12-N2	-2.58	1.46	1.51
2	B	1201	0HK	O33-C30	-2.58	1.37	1.43
2	A	1201	0HK	O33-C30	-2.47	1.38	1.43
2	A	1201	0HK	C31-S37	2.14	1.78	1.74
2	B	1201	0HK	C31-S37	2.16	1.78	1.74
2	B	1201	0HK	C32-S44	2.45	1.78	1.74
2	A	1201	0HK	C32-S44	2.49	1.78	1.74
3	B	1203	OLA	C10-C9	2.66	1.47	1.31
3	B	1204	OLA	C10-C9	2.67	1.47	1.31
3	A	1209	OLA	C10-C9	2.67	1.47	1.31
3	A	1203	OLA	C10-C9	2.68	1.47	1.31
3	A	1202	OLA	C10-C9	2.74	1.47	1.28
2	B	1201	0HK	O11-C28	3.80	1.41	1.34
2	A	1201	0HK	O11-C28	3.83	1.41	1.34

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	A	1207	OLC	O20-C1	4.19	1.45	1.33
6	B	1202	OLC	O20-C1	4.22	1.45	1.33
6	A	1208	OLC	O20-C1	4.25	1.46	1.33
6	A	1206	OLC	O20-C1	4.26	1.46	1.33
6	B	1206	OLC	O20-C1	4.28	1.46	1.33

All (41) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	1201	0HK	C42-C43-S44	-10.50	101.60	113.23
2	A	1201	0HK	C42-C43-S44	-10.47	101.63	113.23
2	A	1201	0HK	C35-C36-S37	-10.04	102.11	113.23
2	B	1201	0HK	C35-C36-S37	-10.04	102.12	113.23
2	A	1201	0HK	C7-C8-C9	-9.90	101.71	107.53
2	B	1201	0HK	C7-C8-C9	-9.69	101.83	107.53
2	B	1201	0HK	C3-C9-C8	-7.83	102.92	107.53
2	A	1201	0HK	C3-C9-C8	-7.61	103.06	107.53
2	A	1201	0HK	C6-C7-C8	-7.25	102.05	106.92
2	B	1201	0HK	C6-C7-C8	-6.80	102.36	106.92
2	A	1201	0HK	C4-C3-C9	-4.20	104.10	106.92
2	B	1201	0HK	C4-C3-C9	-4.17	104.12	106.92
2	B	1201	0HK	C7-C6-C5	-3.02	111.50	113.43
2	A	1201	0HK	C7-C6-C5	-2.80	111.63	113.43
4	A	1204	EDT	C12-C11-N8	-2.79	109.40	113.53
4	A	1204	EDT	C5-C4-N3	-2.55	109.75	113.53
4	A	1204	EDT	C10-C9-N8	-2.23	110.22	113.53
2	A	1201	0HK	C3-C4-C5	-2.04	112.12	113.43
2	B	1201	0HK	O33-C30-C31	2.09	111.03	108.41
6	B	1206	OLC	O20-C1-C2	2.45	119.39	111.85
2	A	1201	0HK	O33-C30-C31	2.56	111.61	108.41
2	A	1201	0HK	C5-O11-C28	2.61	122.80	117.79
2	B	1201	0HK	C5-O11-C28	2.62	122.82	117.79
6	A	1208	OLC	O20-C1-C2	2.64	119.98	111.85
6	A	1206	OLC	O20-C1-C2	2.77	120.38	111.85
2	B	1201	0HK	O33-C30-C32	2.87	112.00	108.41
6	A	1207	OLC	O20-C1-C2	2.88	120.72	111.85
2	A	1201	0HK	O33-C30-C32	2.90	112.04	108.41
6	B	1202	OLC	O20-C1-C2	2.95	120.93	111.85
2	A	1201	0HK	O10-C8-C7	6.19	122.12	117.04
2	B	1201	0HK	O10-C8-C7	6.25	122.17	117.04
2	B	1201	0HK	C9-C3-N2	6.67	107.95	103.56
2	A	1201	0HK	C9-C3-N2	6.73	107.98	103.56

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	1201	0HK	O10-C9-C3	7.15	122.91	117.04
2	B	1201	0HK	O10-C9-C3	7.23	122.97	117.04
2	B	1201	0HK	C8-C7-N2	7.25	108.33	103.56
2	A	1201	0HK	C8-C7-N2	7.42	108.44	103.56
2	B	1201	0HK	C4-C3-N2	10.56	114.83	108.79
2	A	1201	0HK	C4-C3-N2	10.63	114.87	108.79
2	B	1201	0HK	C6-C7-N2	13.15	116.31	108.79
2	A	1201	0HK	C6-C7-N2	13.33	116.42	108.79

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

13 monomers are involved in 26 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	1201	0HK	3	0
4	A	1204	EDT	3	0
5	A	1205	P6G	2	0
6	A	1206	OLC	3	0
6	A	1207	OLC	1	0
6	A	1208	OLC	2	0
2	B	1201	0HK	3	0
6	B	1202	OLC	4	0
3	B	1203	OLA	2	0
3	B	1204	OLA	1	0
5	B	1205	P6G	1	0
6	B	1206	OLC	2	0
7	B	1207	PG6	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 ⓘ

### 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	392/422 (92%)	0.34	22 (5%)	28 21	40, 64, 127, 167	0
1	B	399/422 (94%)	0.31	26 (6%)	22 16	42, 66, 114, 162	0
All	All	791/844 (93%)	0.33	48 (6%)	25 18	40, 65, 122, 167	0

All (48) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	1023	LEU	5.2
1	B	455	LEU	4.5
1	A	458	ALA	4.3
1	B	1019	GLU	4.0
1	B	1017	GLY	3.6
1	B	453	TYR	3.4
1	B	1097	ASN	3.4
1	A	1108	THR	3.3
1	B	458	ALA	3.2
1	A	1115	TRP	3.2
1	A	459	THR	2.8
1	A	1116	ASP	2.8
1	B	61	ARG	2.8
1	A	62	GLN	2.8
1	B	1025	ASN	2.8
1	B	1020	ALA	2.7
1	A	1088	VAL	2.7
1	B	1036	LEU	2.7
1	A	116	SER	2.6
1	B	1065	GLU	2.6
1	B	1018	ASP	2.6
1	B	1091	ALA	2.5
1	A	400	ARG	2.5
1	A	1117	ALA	2.5

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	455	LEU	2.5
1	B	1033	ARG	2.4
1	A	1087	ALA	2.4
1	A	1112	THR	2.4
1	B	1099	THR	2.3
1	B	1103	ALA	2.3
1	A	81	ILE	2.2
1	B	175	VAL	2.2
1	A	85	SER	2.2
1	B	403	PHE	2.2
1	B	465	ARG	2.2
1	A	112	ASP	2.1
1	A	398	VAL	2.1
1	B	392	ALA	2.1
1	A	429	CYS	2.1
1	B	1035	ILE	2.1
1	A	1051	VAL	2.1
1	B	1100	PRO	2.1
1	A	467	LEU	2.0
1	B	1022	LYS	2.0
1	A	1007	LEU	2.0
1	B	119	SER	2.0
1	A	67	ASN	2.0
1	B	115	VAL	2.0

## 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, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q < 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
6	OLC	A	1208	25/25	0.51	0.38	6.58	57,86,91,91	0
3	OLA	A	1203	14/20	0.82	0.33	5.39	83,84,89,91	0
3	OLA	B	1203	20/20	0.78	0.33	5.37	76,79,87,88	0
5	P6G	B	1205	19/19	0.76	0.39	4.62	75,81,85,86	0
6	OLC	B	1202	17/25	0.61	0.35	4.62	79,82,91,91	0
6	OLC	A	1207	18/25	0.75	0.26	3.40	73,79,94,95	0
3	OLA	A	1209	20/20	0.61	0.35	3.35	73,79,83,83	0
3	OLA	B	1204	20/20	0.83	0.33	3.14	59,70,73,73	0
6	OLC	A	1206	25/25	0.68	0.35	2.53	70,74,90,91	0
5	P6G	A	1205	19/19	0.85	0.27	2.36	65,69,86,86	0
6	OLC	B	1206	25/25	0.56	0.32	2.14	53,76,86,87	0
4	EDT	A	1204	20/20	0.90	0.21	0.95	70,74,80,80	0
2	0HK	A	1201	26/26	0.95	0.21	0.04	42,47,48,64	0
2	0HK	B	1201	26/26	0.96	0.21	0.04	40,47,54,67	0
7	PG6	B	1207	10/18	0.69	0.20	0.02	90,91,91,91	0
3	OLA	A	1202	12/20	0.72	0.34	-	79,82,91,92	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.