



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 22, 2020 – 02:43 pm BST

PDB ID : 6BHL  
Title : Phosphotriesterase variant S5deltaL7  
Authors : Miton, C.M.; Campbell, E.C.; Jackson, C.J.; Tokuriki, N.  
Deposited on : 2017-10-30  
Resolution : 1.40 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](mailto:validation@mail.wwpdb.org)

A user guide is available at

<https://www.wwpdb.org/validation/2017/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 : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.11  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0158  
CCP4 : 7.0.044 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.11

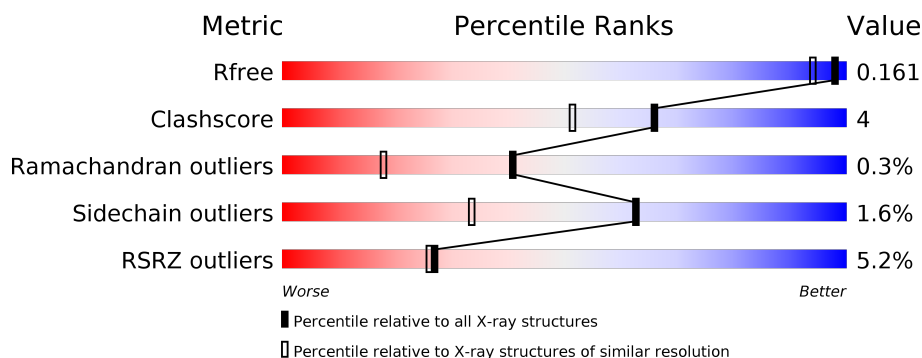
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 1.40 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))
$R_{free}$	130704	1714 (1.40-1.40)
Clashscore	141614	1812 (1.40-1.40)
Ramachandran outliers	138981	1763 (1.40-1.40)
Sidechain outliers	138945	1762 (1.40-1.40)
RSRZ outliers	127900	1674 (1.40-1.40)

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 respectively. 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	324	<div> <div>6%</div> <div> <div></div> <div>90%</div> <div>8% ..</div> </div> </div>
1	G	324	<div> <div>4%</div> <div> <div></div> <div>92%</div> <div>6% .</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
3	MPD	A	2404	-	-	X	-

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 13146 atoms, of which 6434 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 Phosphotriesterase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	320	Total	C	H	N	O	S	326	108	0
			6484	2016	3303	573	586	6			
1	G	319	Total	C	H	N	O	S	93	86	0
			5934	1846	3019	519	543	7			

There are 30 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	216	LEU	PHE	conflict	UNP A0A060GSX0
A	233	ALA	ASP	conflict	UNP A0A060GSX0
A	254	SER	ARG	conflict	UNP A0A060GSX0
A	?	-	ASN	deletion	UNP A0A060GSX0
A	?	-	ALA	deletion	UNP A0A060GSX0
A	?	-	SER	deletion	UNP A0A060GSX0
A	?	-	ALA	deletion	UNP A0A060GSX0
A	?	-	SER	deletion	UNP A0A060GSX0
A	?	-	ALA	deletion	UNP A0A060GSX0
A	?	-	PHE	deletion	UNP A0A060GSX0
A	?	-	MET	deletion	UNP A0A060GSX0
A	?	-	GLY	deletion	UNP A0A060GSX0
A	284	THR	MET	conflict	UNP A0A060GSX0
A	304	ILE	PHE	conflict	UNP A0A060GSX0
A	311	GLY	VAL	conflict	UNP A0A060GSX0
G	216	LEU	PHE	conflict	UNP A0A060GSX0
G	233	ALA	ASP	conflict	UNP A0A060GSX0
G	254	SER	ARG	conflict	UNP A0A060GSX0
G	?	-	ASN	deletion	UNP A0A060GSX0
G	?	-	ALA	deletion	UNP A0A060GSX0
G	?	-	SER	deletion	UNP A0A060GSX0
G	?	-	ALA	deletion	UNP A0A060GSX0
G	?	-	SER	deletion	UNP A0A060GSX0
G	?	-	ALA	deletion	UNP A0A060GSX0
G	?	-	PHE	deletion	UNP A0A060GSX0

*Continued on next page...*

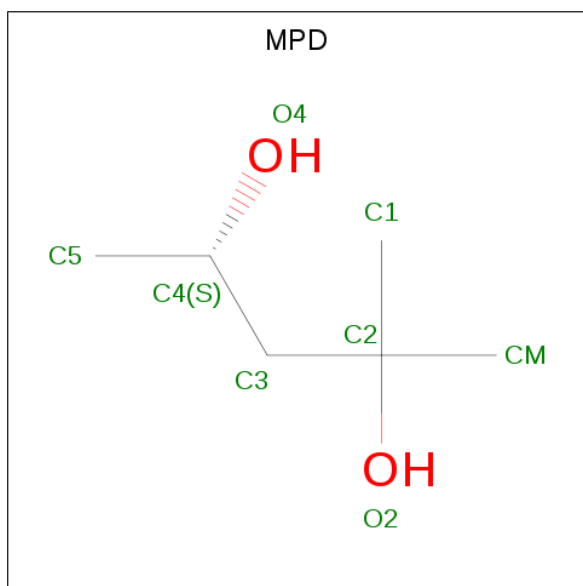
*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
G	?	-	MET	deletion	UNP A0A060GSX0
G	?	-	GLY	deletion	UNP A0A060GSX0
G	284	THR	MET	conflict	UNP A0A060GSX0
G	304	ILE	PHE	conflict	UNP A0A060GSX0
G	311	GLY	VAL	conflict	UNP A0A060GSX0

- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	G	2	Total	Zn	0	0
			2	2		
2	A	2	Total	Zn	0	0
			2	2		

- Molecule 3 is (4S)-2-METHYL-2,4-PENTANEDIOL (three-letter code: MPD) (formula: C<sub>6</sub>H<sub>14</sub>O<sub>2</sub>).



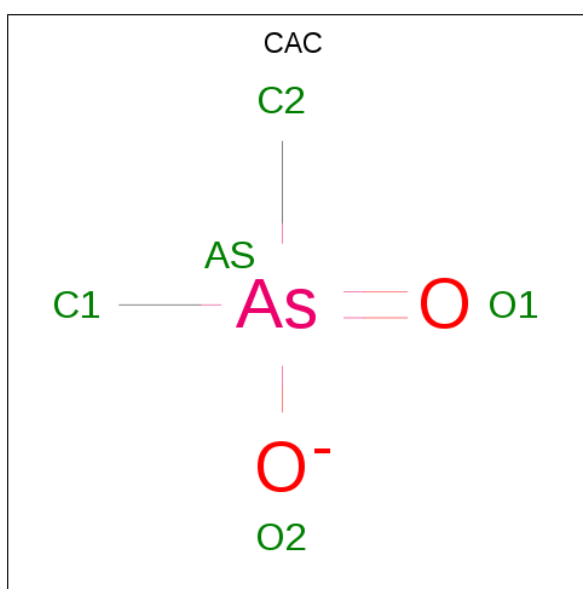
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	H	O	0	0
			22	6	14	2		
3	A	1	Total	C	H	O	0	0
			22	6	14	2		
3	A	1	Total	C	H	O	0	0
			22	6	14	2		
3	A	1	Total	C	H	O	0	0
			22	6	14	2		

*Continued on next page...*

Continued from previous page...

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	G	1	Total	C	H	O	0	0
			22	6	14	2		
3	G	1	Total	C	H	O	0	0
			22	6	14	2		
3	G	1	Total	C	H	O	0	0
			22	6	14	2		
3	G	1	Total	C	H	O	0	0
			22	6	14	2		

- Molecule 4 is CACODYLATE ION (three-letter code: CAC) (formula:  $C_2H_6AsO_2$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	1	Total	As	C	O	0	0
			5	1	2	2		
4	G	1	Total	As	C	O	0	0
			5	1	2	2		

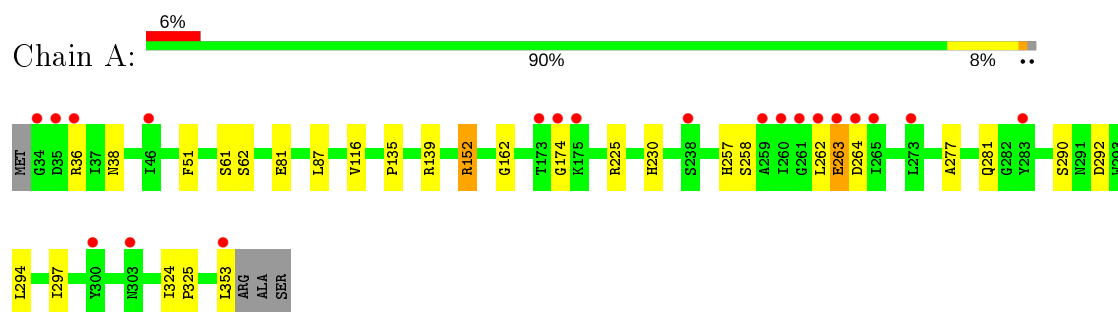
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	243	Total	O	0	0
			243	243		
5	G	295	Total	O	0	0
			295	295		

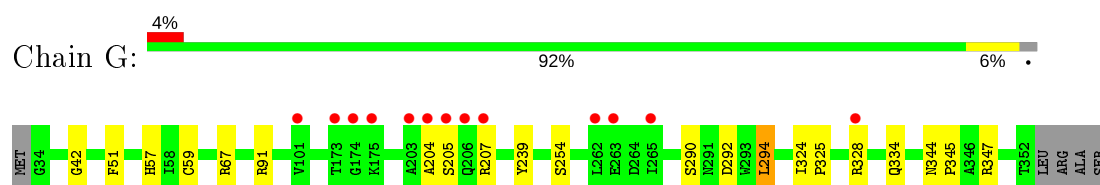
### 3 Residue-property plots [i](#)

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



#### • Molecule 1: Phosphotriesterase



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	85.15Å 85.48Å 88.53Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	30.75 – 1.40 30.75 – 1.40	Depositor EDS
% Data completeness (in resolution range)	99.0 (30.75-1.40) 99.0 (30.75-1.40)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.29 (at 1.40Å)	Xtriage
Refinement program	PHENIX 1.9_1692	Depositor
R, $R_{free}$	0.134 , 0.161 0.134 , 0.161	Depositor DCC
$R_{free}$ test set	6195 reflections (4.91%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	12.6	Xtriage
Anisotropy	0.034	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.42 , 52.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.015 for -h,l,k 0.015 for -l,-k,-h 0.015 for k,h,-l 0.005 for k,l,h 0.005 for l,h,k	Xtriage
$F_o, F_c$ correlation	0.98	EDS
Total number of atoms	13146	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	18.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.06% 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, 0.333 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: CAC, ZN, MPD, KCX

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	0.42	0/3487	0.63	3/4739 (0.1%)
1	G	0.41	0/3199	0.61	0/4342
All	All	0.42	0/6686	0.62	3/9081 (0.0%)

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	139[A]	ARG	NE-CZ-NH1	6.03	123.31	120.30
1	A	139[B]	ARG	NE-CZ-NH1	6.03	123.31	120.30
1	A	225	ARG	NE-CZ-NH2	-5.14	117.73	120.30

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	A	3181	3303	3097	20	0
1	G	2915	3019	2809	18	0
2	A	2	0	0	0	0
2	G	2	0	0	0	0
3	A	32	56	56	9	0
3	G	32	56	56	2	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	A	5	0	0	3	0
4	G	5	0	0	3	0
5	A	243	0	0	11	0
5	G	295	0	0	11	0
All	All	6712	6434	6018	48	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 4.

The worst 5 of 48 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)
4:A:2407:CAC:AS	5:A:2501:HOH:O	2.21	1.15
1:G:292:ASP:OD2	5:G:2501:HOH:O	1.81	0.97
3:A:2404:MPD:H13	5:A:2501:HOH:O	1.64	0.97
4:A:2407:CAC:O2	5:A:2501:HOH:O	1.82	0.95
4:G:2407:CAC:AS	5:G:2501:HOH:O	2.48	0.90

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	A	445/324 (137%)	428 (96%)	15 (3%)	2 (0%)	34	12
1	G	406/324 (125%)	386 (95%)	20 (5%)	0	100	100
All	All	851/648 (131%)	814 (96%)	35 (4%)	2 (0%)	41	21

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	174	GLY
1	A	263	GLU

### 5.3.2 Protein sidechains [i](#)

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	359/259 (139%)	351 (98%)	8 (2%)	52	19
1	G	337/259 (130%)	333 (99%)	4 (1%)	71	47
All	All	696/518 (134%)	684 (98%)	12 (2%)	62	31

5 of 12 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	290[A]	SER
1	A	290[B]	SER
1	G	290	SER
1	A	258	SER
1	G	51	PHE

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

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

2 non-standard protein/DNA/RNA residues 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	$\# Z  > 2$
1	KCX	G	169	1,2	7,11,12	0.90	0	4,12,14	0.47	0
1	KCX	A	169	1,2	7,11,12	0.87	0	4,12,14	0.64	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
1	KCX	G	169	1,2	-	0/7/10/12	-
1	KCX	A	169	1,2	-	0/7/10/12	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 14 ligands modelled in this entry, 4 are monoatomic - leaving 10 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
3	MPD	G	2405	-	7,7,7	0.37	0	9,10,10	0.74	0
3	MPD	A	2406	-	7,7,7	0.33	0	9,10,10	0.38	0
4	CAC	A	2407	2	0,4,4	0.00	-	0,6,6	0.00	-
4	CAC	G	2407	2	0,4,4	0.00	-	0,6,6	0.00	-
3	MPD	G	2403	-	7,7,7	0.34	0	9,10,10	0.66	0
3	MPD	G	2406	-	7,7,7	0.34	0	9,10,10	0.30	0
3	MPD	A	2403	-	7,7,7	0.42	0	9,10,10	0.27	0
3	MPD	G	2404	-	7,7,7	0.37	0	9,10,10	0.16	0
3	MPD	A	2404	-	7,7,7	0.51	0	9,10,10	1.16	1 (11%)
3	MPD	A	2405	-	7,7,7	0.31	0	9,10,10	0.45	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
3	MPD	G	2405	-	-	3/5/5/5	-
3	MPD	A	2406	-	-	0/5/5/5	-
3	MPD	G	2403	-	-	2/5/5/5	-
3	MPD	G	2406	-	-	0/5/5/5	-
3	MPD	A	2403	-	-	3/5/5/5	-
3	MPD	G	2404	-	-	0/5/5/5	-
3	MPD	A	2404	-	-	3/5/5/5	-
3	MPD	A	2405	-	-	2/5/5/5	-

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	2404	MPD	O4-C4-C3	2.39	121.00	111.36

There are no chirality outliers.

5 of 13 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	G	2405	MPD	C2-C3-C4-C5
3	A	2404	MPD	C1-C2-C3-C4
3	A	2404	MPD	O2-C2-C3-C4
3	A	2405	MPD	C2-C3-C4-O4

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms
3	G	2405	MPD	O2-C2-C3-C4

There are no ring outliers.

6 monomers are involved in 17 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	G	2405	MPD	1	0
4	A	2407	CAC	3	0
4	G	2407	CAC	3	0
3	G	2406	MPD	1	0
3	A	2404	MPD	8	0
3	A	2405	MPD	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, 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	319/324 (98%)	0.30	20 (6%) 20 18	8, 15, 33, 83	0
1	G	318/324 (98%)	-0.01	13 (4%) 37 37	8, 14, 29, 53	0
All	All	637/648 (98%)	0.14	33 (5%) 27 26	8, 14, 32, 83	0

The worst 5 of 33 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	262	LEU	12.0
1	A	265	ILE	9.4
1	A	263	GLU	8.1
1	A	261	GLY	7.7
1	A	353	LEU	7.3

### 6.2 Non-standard residues in protein, DNA, RNA chains [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. 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	B-factors(Å <sup>2</sup> )	Q<0.9
1	KCX	G	169	12/13	0.97	0.11	8,11,12,13	0
1	KCX	A	169	12/13	0.98	0.10	8,12,14,15	0

### 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. 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	B-factors(Å <sup>2</sup> )	Q<0.9
3	MPD	A	2405	8/8	0.48	0.32	64,77,79,79	0
3	MPD	G	2405	8/8	0.51	0.38	40,48,50,52	22
3	MPD	G	2406	8/8	0.53	0.31	52,63,65,66	0
3	MPD	A	2406	8/8	0.59	0.31	63,75,76,76	0
3	MPD	G	2403	8/8	0.62	0.30	38,46,50,50	0
3	MPD	G	2404	8/8	0.70	0.25	40,49,53,53	0
3	MPD	A	2403	8/8	0.73	0.26	64,77,79,79	0
3	MPD	A	2404	8/8	0.74	0.25	37,47,50,50	0
4	CAC	A	2407	5/5	0.99	0.11	10,11,13,14	5
4	CAC	G	2407	5/5	0.99	0.11	10,11,13,13	5
2	ZN	G	2402	1/1	1.00	0.07	12,12,12,12	1
2	ZN	A	2401	1/1	1.00	0.07	9,9,9,9	0
2	ZN	G	2401	1/1	1.00	0.07	10,10,10,10	0
2	ZN	A	2402	1/1	1.00	0.07	12,12,12,12	1

## 6.5 Other polymers [i](#)

There are no such residues in this entry.