



wwPDB X-ray Structure Validation Summary Report ⓘ

May 14, 2020 – 09:50 pm BST

PDB ID : 4IOX
Title : The structure of the herpes simplex virus DNA-packaging motor pUL15 C-terminal nuclease domain provides insights into cleavage of concatemeric viral genome precursors
Authors : Selvarajan Sigamani, S.; Zhao, H.; Kamau, Y.; Tang, L.
Deposited on : 2013-01-08
Resolution : 2.46 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

<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

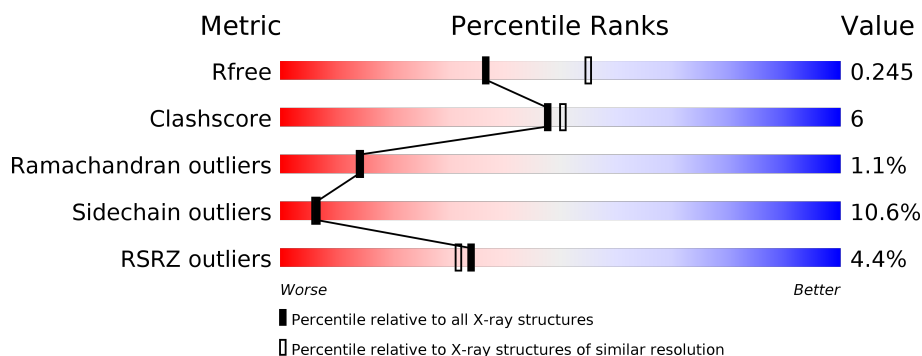
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.46 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	1544 (2.48-2.44)
Clashscore	141614	1613 (2.48-2.44)
Ramachandran outliers	138981	1598 (2.48-2.44)
Sidechain outliers	138945	1598 (2.48-2.44)
RSRZ outliers	127900	1523 (2.48-2.44)

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 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	286	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 0%, red 1%, orange 1%, orange 15%, yellow 15%, yellow 62%, green 62%, green 100%);"></div> <div style="display: flex; justify-content: space-between; margin-top: 5px;"> % 62% 15% • 21% </div> </div>
1	B	286	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 0%, red 3%, orange 3%, orange 11%, yellow 11%, yellow 60%, green 60%, green 100%);"></div> <div style="display: flex; justify-content: space-between; margin-top: 5px;"> 3% 60% 11% • 24% </div> </div>
1	C	286	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 0%, red 5%, orange 5%, orange 12%, yellow 12%, yellow 62%, green 62%, green 100%);"></div> <div style="display: flex; justify-content: space-between; margin-top: 5px;"> 5% 62% 12% • 23% </div> </div>
2	D	6	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, green 0%, green 100%);"></div> <div style="display: flex; justify-content: space-between; margin-top: 5px;"> 100% </div> </div>

2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 5172 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 Tripartite terminase subunit UL15.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	225	Total	C	N	O	S	0	0	0
			1727	1109	295	316	7			
1	B	216	Total	C	N	O	S	0	0	0
			1653	1064	279	303	7			
1	C	220	Total	C	N	O	S	0	0	0
			1682	1085	283	307	7			

There are 63 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	450	MET	-	EXPRESSION TAG	UNP P04295
A	451	GLY	-	EXPRESSION TAG	UNP P04295
A	452	SER	-	EXPRESSION TAG	UNP P04295
A	453	SER	-	EXPRESSION TAG	UNP P04295
A	454	HIS	-	EXPRESSION TAG	UNP P04295
A	455	HIS	-	EXPRESSION TAG	UNP P04295
A	456	HIS	-	EXPRESSION TAG	UNP P04295
A	457	HIS	-	EXPRESSION TAG	UNP P04295
A	458	HIS	-	EXPRESSION TAG	UNP P04295
A	459	HIS	-	EXPRESSION TAG	UNP P04295
A	460	SER	-	EXPRESSION TAG	UNP P04295
A	461	SER	-	EXPRESSION TAG	UNP P04295
A	462	GLY	-	EXPRESSION TAG	UNP P04295
A	463	LEU	-	EXPRESSION TAG	UNP P04295
A	464	VAL	-	EXPRESSION TAG	UNP P04295
A	465	PRO	-	EXPRESSION TAG	UNP P04295
A	466	ARG	-	EXPRESSION TAG	UNP P04295
A	467	GLY	-	EXPRESSION TAG	UNP P04295
A	468	SER	-	EXPRESSION TAG	UNP P04295
A	469	HIS	-	EXPRESSION TAG	UNP P04295
A	470	MET	-	EXPRESSION TAG	UNP P04295
B	450	MET	-	EXPRESSION TAG	UNP P04295
B	451	GLY	-	EXPRESSION TAG	UNP P04295

Continued on next page...

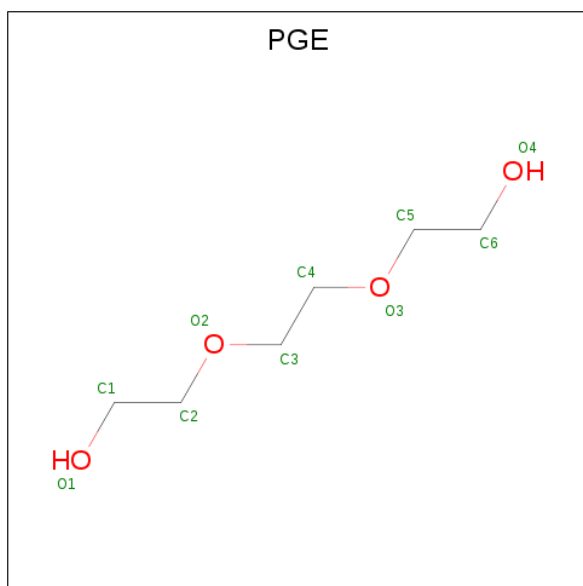
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	452	SER	-	EXPRESSION TAG	UNP P04295
B	453	SER	-	EXPRESSION TAG	UNP P04295
B	454	HIS	-	EXPRESSION TAG	UNP P04295
B	455	HIS	-	EXPRESSION TAG	UNP P04295
B	456	HIS	-	EXPRESSION TAG	UNP P04295
B	457	HIS	-	EXPRESSION TAG	UNP P04295
B	458	HIS	-	EXPRESSION TAG	UNP P04295
B	459	HIS	-	EXPRESSION TAG	UNP P04295
B	460	SER	-	EXPRESSION TAG	UNP P04295
B	461	SER	-	EXPRESSION TAG	UNP P04295
B	462	GLY	-	EXPRESSION TAG	UNP P04295
B	463	LEU	-	EXPRESSION TAG	UNP P04295
B	464	VAL	-	EXPRESSION TAG	UNP P04295
B	465	PRO	-	EXPRESSION TAG	UNP P04295
B	466	ARG	-	EXPRESSION TAG	UNP P04295
B	467	GLY	-	EXPRESSION TAG	UNP P04295
B	468	SER	-	EXPRESSION TAG	UNP P04295
B	469	HIS	-	EXPRESSION TAG	UNP P04295
B	470	MET	-	EXPRESSION TAG	UNP P04295
C	450	MET	-	EXPRESSION TAG	UNP P04295
C	451	GLY	-	EXPRESSION TAG	UNP P04295
C	452	SER	-	EXPRESSION TAG	UNP P04295
C	453	SER	-	EXPRESSION TAG	UNP P04295
C	454	HIS	-	EXPRESSION TAG	UNP P04295
C	455	HIS	-	EXPRESSION TAG	UNP P04295
C	456	HIS	-	EXPRESSION TAG	UNP P04295
C	457	HIS	-	EXPRESSION TAG	UNP P04295
C	458	HIS	-	EXPRESSION TAG	UNP P04295
C	459	HIS	-	EXPRESSION TAG	UNP P04295
C	460	SER	-	EXPRESSION TAG	UNP P04295
C	461	SER	-	EXPRESSION TAG	UNP P04295
C	462	GLY	-	EXPRESSION TAG	UNP P04295
C	463	LEU	-	EXPRESSION TAG	UNP P04295
C	464	VAL	-	EXPRESSION TAG	UNP P04295
C	465	PRO	-	EXPRESSION TAG	UNP P04295
C	466	ARG	-	EXPRESSION TAG	UNP P04295
C	467	GLY	-	EXPRESSION TAG	UNP P04295
C	468	SER	-	EXPRESSION TAG	UNP P04295
C	469	HIS	-	EXPRESSION TAG	UNP P04295
C	470	MET	-	EXPRESSION TAG	UNP P04295

- Molecule 2 is a protein called peptide.

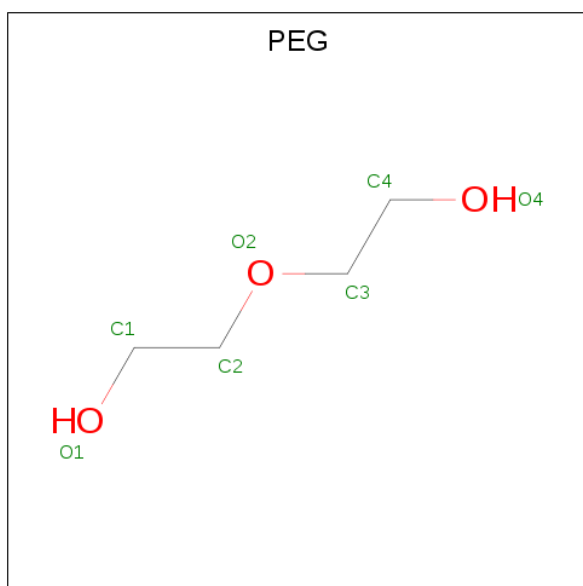
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	D	6	Total	C	N	O	0	0	0
			30	18	6	6			

- Molecule 3 is TRIETHYLENE GLYCOL (three-letter code: PGE) (formula: $C_6H_{14}O_4$).



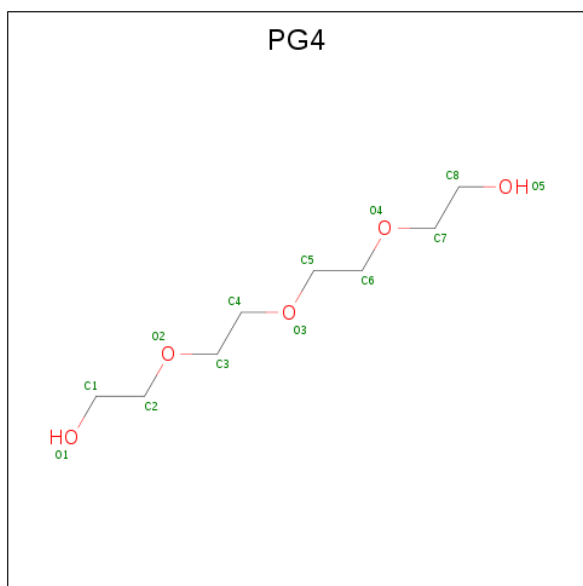
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	O	0	0
			10	6	4		

- Molecule 4 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula: $C_4H_{10}O_3$).



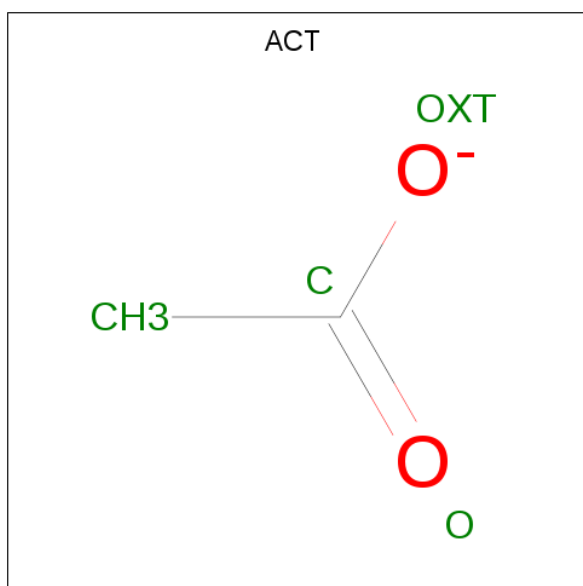
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			7	4	3		

- Molecule 5 is TETRAETHYLENE GLYCOL (three-letter code: PG4) (formula: $C_8H_{18}O_5$).



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

- Molecule 6 is ACETATE ION (three-letter code: ACT) (formula: $C_2H_3O_2$).

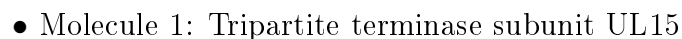
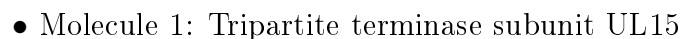


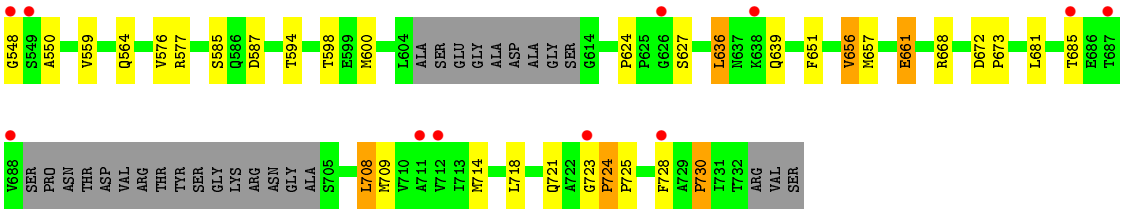
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	A	1	Total	C	O	0	0
			4	2	2		

- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	16	Total	O	0	0
			16	16		
7	B	21	Total	O	0	0
			21	21		
7	C	9	Total	O	0	0
			9	9		

- Molecule 1: Tripartite terminase subunit UL15





● Molecule 2: peptide

Chain D:  100%

There are no outlier residues recorded for this chain.

4 Data and refinement statistics

Property	Value	Source
Space group	P 43 21 2	Depositor
Cell constants a, b, c, α , β , γ	96.94Å 96.94Å 194.03Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	19.80 – 2.46 33.75 – 2.46	Depositor EDS
% Data completeness (in resolution range)	99.5 (19.80-2.46) 95.7 (33.75-2.46)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	0.09	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.57 (at 2.45Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8.1_1168)	Depositor
R, R_{free}	0.203 , 0.247 0.206 , 0.245	Depositor DCC
R_{free} test set	1722 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å ²)	53.9	Xtriage
Anisotropy	0.332	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.34 , 47.2	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	5172	wwPDB-VP
Average B, all atoms (Å ²)	56.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.69% 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

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: PG4, PEG, PGE, ACT

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	0.48	0/1770	0.61	0/2412
1	B	0.49	0/1692	0.67	1/2303 (0.0%)
1	C	0.45	0/1722	0.61	0/2345
All	All	0.47	0/5184	0.63	1/7060 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	1
1	B	0	2
All	All	0	3

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	723	GLY	N-CA-C	9.96	138.00	113.10

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	721	GLN	Peptide
1	B	721	GLN	Peptide
1	B	723	GLY	Peptide

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	1727	0	1710	27	0
1	B	1653	0	1627	21	0
1	C	1682	0	1661	19	0
2	D	30	0	9	0	0
3	A	10	0	14	1	0
4	A	7	0	10	0	0
5	A	13	0	18	3	0
6	A	4	0	3	0	0
7	A	16	0	0	1	0
7	B	21	0	0	0	0
7	C	9	0	0	0	0
All	All	5172	0	5052	62	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 6.

The worst 5 of 62 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:681:LEU:HD23	1:A:709:MET:HG3	1.53	0.91
1:A:525:VAL:HG13	1:A:537:ALA:HB3	1.59	0.84
1:A:621:HIS:HE1	5:A:803:PG4:H31	1.45	0.80
1:A:619:PHE:H	5:A:803:PG4:H32	1.48	0.79
1:C:681:LEU:HD23	1:C:709:MET:HG3	1.68	0.76

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	219/286 (77%)	210 (96%)	8 (4%)	1 (0%)	29	34
1	B	206/286 (72%)	195 (95%)	8 (4%)	3 (2%)	10	9
1	C	210/286 (73%)	197 (94%)	10 (5%)	3 (1%)	11	9
All	All	635/858 (74%)	602 (95%)	26 (4%)	7 (1%)	14	14

5 of 7 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	730	PRO
1	C	730	PRO
1	B	724	PRO
1	B	725	PRO
1	C	724	PRO

5.3.2 Protein sidechains ⓘ

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

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	184/233 (79%)	165 (90%)	19 (10%)	7	7
1	B	176/233 (76%)	155 (88%)	21 (12%)	5	4
1	C	179/233 (77%)	162 (90%)	17 (10%)	8	9
All	All	539/699 (77%)	482 (89%)	57 (11%)	6	6

5 of 57 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	562	LEU
1	B	636	LEU
1	C	657	MET
1	B	576	VAL
1	B	618	LEU

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

Mol	Chain	Res	Type
1	A	621	HIS

5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

5.6 Ligand geometry [i](#)

4 ligands are modelled in this entry.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z > 2$	Counts	RMSZ	$\# Z > 2$
3	PGE	A	801	-	9,9,9	0.62	0	8,8,8	1.49	0
5	PG4	A	803	-	12,12,12	0.69	0	11,11,11	1.44	0
4	PEG	A	802	-	6,6,6	0.58	0	5,5,5	1.51	1 (20%)
6	ACT	A	804	-	1,3,3	1.47	0	0,3,3	0.00	-

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	PGE	A	801	-	-	5/7/7/7	-
5	PG4	A	803	-	-	3/10/10/10	-
4	PEG	A	802	-	-	3/4/4/4	-

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	A	802	PEG	O2-C2-C1	2.03	118.99	110.07

There are no chirality outliers.

5 of 11 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	801	PGE	C1-C2-O2-C3
3	A	801	PGE	O1-C1-C2-O2
3	A	801	PGE	C6-C5-O3-C4
4	A	802	PEG	O2-C3-C4-O4
4	A	802	PEG	O1-C1-C2-O2

There are no ring outliers.

2 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	801	PGE	1	0
5	A	803	PG4	3	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	A	225/286 (78%)	-0.21	4 (1%) 68 65	35, 49, 82, 114	0
1	B	216/286 (75%)	-0.09	10 (4%) 32 30	35, 51, 87, 106	0
1	C	220/286 (76%)	0.21	15 (6%) 17 13	34, 54, 92, 121	0
2	D	0/6	-	-	-	-
All	All	661/864 (76%)	-0.03	29 (4%) 34 32	34, 51, 87, 121	0

The worst 5 of 29 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	543	LEU	9.2
1	C	687	THR	4.1
1	B	543	LEU	3.7
1	B	728	PHE	3.6
1	C	548	GLY	3.3

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. 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	B-factors(\AA^2)	Q<0.9
4	PEG	A	802	7/7	0.74	0.13	72,76,83,85	0
6	ACT	A	804	4/4	0.77	0.19	78,85,87,90	0
3	PGE	A	801	10/10	0.83	0.19	50,61,69,73	0
5	PG4	A	803	13/13	0.85	0.20	66,80,86,88	0

6.5 Other polymers [i](#)

There are no such residues in this entry.