



Full wwPDB X-ray Structure Validation Report ⓘ

Feb 21, 2022 – 12:08 PM JST

PDB ID : 7E6B
Title : Crystal structure of PMP-bound form of cysteine desulfurase SufS C361A from *Bacillus subtilis*
Authors : Nakamura, R.; Takahashi, Y.; Fujishiro, T.
Deposited on : 2021-02-22
Resolution : 1.84 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

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 : 2.26
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) : 2.26

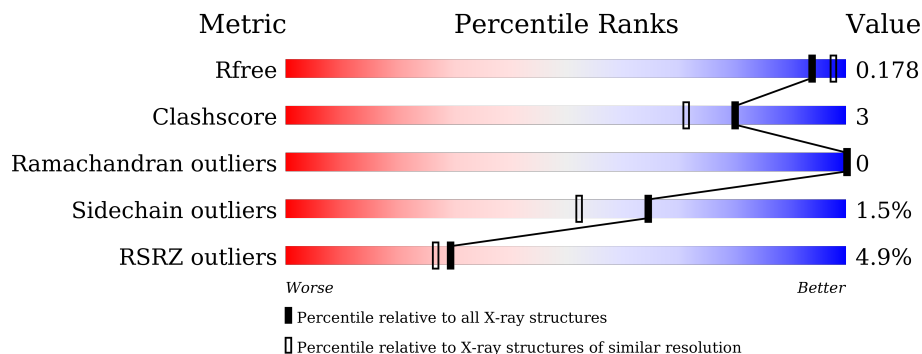
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.84 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	4003 (1.86-1.82)
Clashscore	141614	4233 (1.86-1.82)
Ramachandran outliers	138981	4185 (1.86-1.82)
Sidechain outliers	138945	4186 (1.86-1.82)
RSRZ outliers	127900	3957 (1.86-1.82)

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 of 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	419	

2 Entry composition [i](#)

There are 6 unique types of molecules in this entry. The entry contains 3540 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.

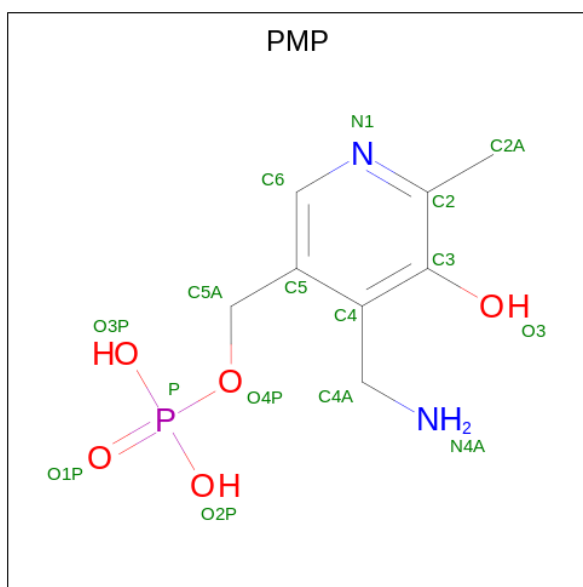
- Molecule 1 is a protein called Cysteine desulfurase SufS.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	410	3236	2056	549	619	12	0	7	0

There are 14 discrepancies between the modelled and reference sequences:

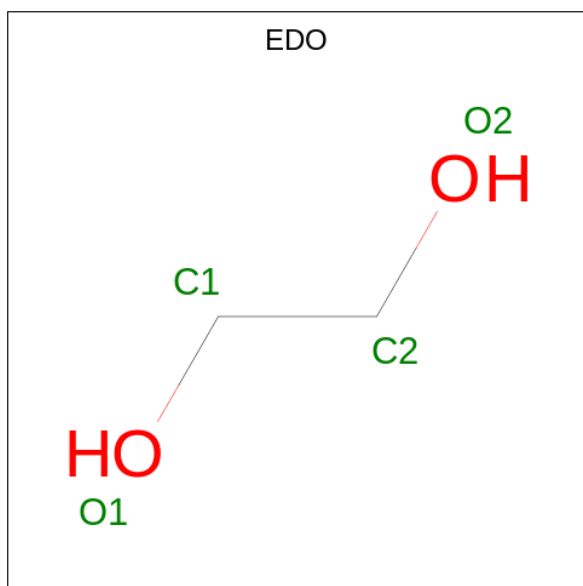
Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	MET	-	initiating methionine	UNP O32164
A	-1	GLY	-	expression tag	UNP O32164
A	0	HIS	-	expression tag	UNP O32164
A	361	ALA	CYS	engineered mutation	UNP O32164
A	407	VAL	-	expression tag	UNP O32164
A	408	ASP	-	expression tag	UNP O32164
A	409	LEU	-	expression tag	UNP O32164
A	410	GLU	-	expression tag	UNP O32164
A	411	HIS	-	expression tag	UNP O32164
A	412	HIS	-	expression tag	UNP O32164
A	413	HIS	-	expression tag	UNP O32164
A	414	HIS	-	expression tag	UNP O32164
A	415	HIS	-	expression tag	UNP O32164
A	416	HIS	-	expression tag	UNP O32164

- Molecule 2 is 4'-DEOXY-4'-AMINOPYRIDOXAL-5'-PHOSPHATE (three-letter code: PMP) (formula: C₈H₁₃N₂O₅P).



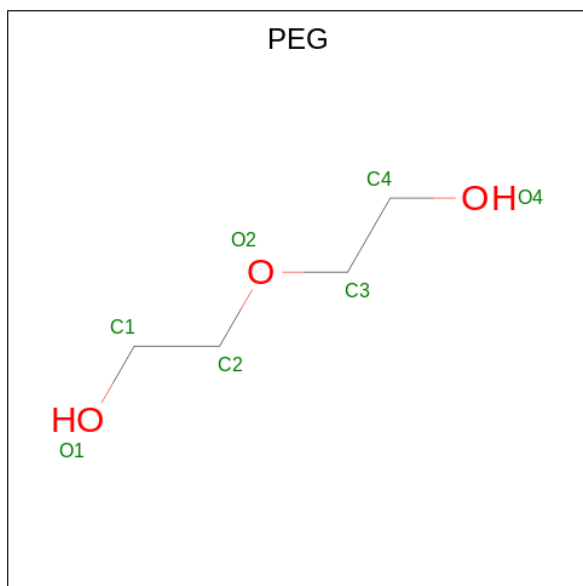
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	P		
2	A	1	16	8	2	5	1	0	0

- Molecule 3 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C₂H₆O₂).



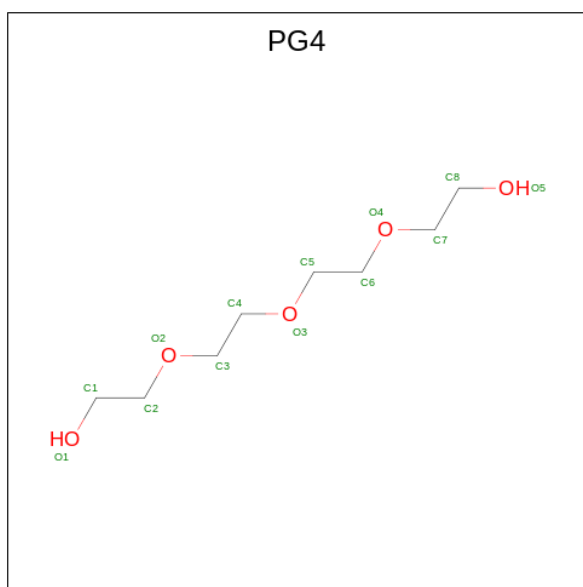
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
3	A	1	4	2	2	0	0
3	A	1	4	2	2	0	0
3	A	1	4	2	2	0	0

- Molecule 4 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula: C₄H₁₀O₃).



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

- Molecule 5 is TETRAETHYLENE GLYCOL (three-letter code: PG4) (formula: C₈H₁₈O₅).



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

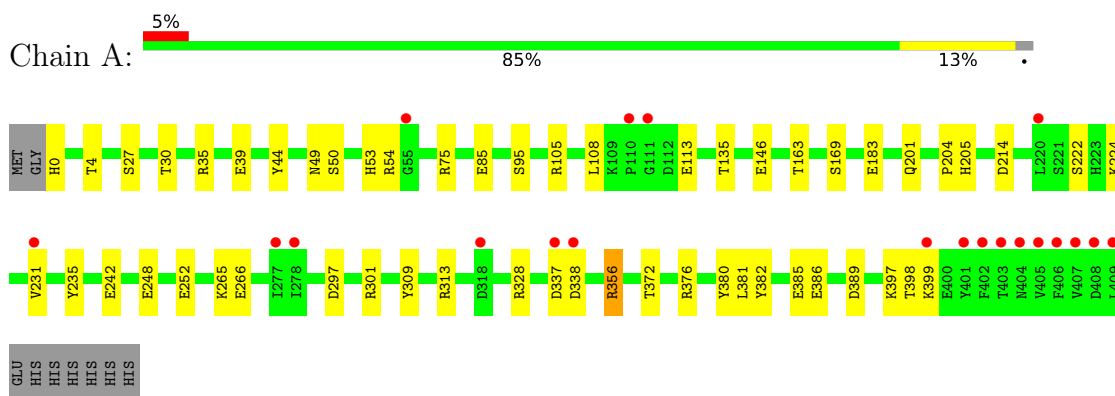
- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	214	Total	O	0	0
			214	214		

3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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: Cysteine desulfurase SufS



4 Data and refinement statistics i

Property	Value	Source
Space group	P 31 2 1	Depositor
Cell constants a, b, c, α , β , γ	92.80Å 92.80Å 129.00Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	46.44 – 1.84 46.40 – 1.84	Depositor EDS
% Data completeness (in resolution range)	100.0 (46.44-1.84) 100.0 (46.40-1.84)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.76 (at 1.84Å)	Xtrriage
Refinement program	REFMAC 5.8.0267	Depositor
R, R_{free}	0.157 , 0.178 0.157 , 0.178	Depositor DCC
R_{free} test set	2816 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å ²)	35.3	Xtrriage
Anisotropy	0.014	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.38 , 55.3	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtrriage
Estimated twinning fraction	0.028 for -h,-k,l	Xtrriage
F_o, F_c correlation	0.97	EDS
Total number of atoms	3540	wwPDB-VP
Average B, all atoms (Å ²)	42.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.52% 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: PMP, EDO, PEG, PG4

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	1.06	11/3321 (0.3%)	1.22	19/4510 (0.4%)

All (11) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	252	GLU	CD-OE2	14.95	1.42	1.25
1	A	183	GLU	CD-OE1	9.07	1.35	1.25
1	A	385	GLU	CD-OE1	7.99	1.34	1.25
1	A	248	GLU	CD-OE2	-7.30	1.17	1.25
1	A	386	GLU	CD-OE1	7.09	1.33	1.25
1	A	95	SER	CB-OG	5.57	1.49	1.42
1	A	266	GLU	CD-OE2	5.55	1.31	1.25
1	A	50	SER	CA-CB	5.50	1.61	1.52
1	A	242	GLU	CD-OE1	-5.30	1.19	1.25
1	A	53	HIS	CE1-NE2	5.16	1.44	1.32
1	A	39	GLU	CD-OE1	-5.02	1.20	1.25

All (19) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	356[A]	ARG	CG-CD-NE	8.82	130.33	111.80
1	A	356[B]	ARG	CG-CD-NE	8.82	130.33	111.80
1	A	313	ARG	NE-CZ-NH1	7.33	123.96	120.30
1	A	35	ARG	CG-CD-NE	-7.04	97.02	111.80
1	A	248	GLU	CB-CA-C	-6.56	97.29	110.40
1	A	75	ARG	NE-CZ-NH2	-6.55	117.03	120.30
1	A	309	TYR	CB-CG-CD1	-6.55	117.07	121.00
1	A	328	ARG	NE-CZ-NH2	-6.26	117.17	120.30
1	A	372	THR	CA-CB-OG1	-6.14	96.10	109.00
1	A	169	SER	N-CA-CB	-5.76	101.86	110.50

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	265	LYS	CB-CA-C	-5.64	99.12	110.40
1	A	297	ASP	CB-CG-OD1	5.58	123.32	118.30
1	A	389	ASP	CB-CG-OD1	5.33	123.10	118.30
1	A	309	TYR	CB-CG-CD2	5.22	124.13	121.00
1	A	301	ARG	NE-CZ-NH2	5.19	122.90	120.30
1	A	214	ASP	CB-CG-OD1	5.15	122.93	118.30
1	A	398	THR	CA-CB-OG1	-5.04	98.42	109.00
1	A	356[A]	ARG	NE-CZ-NH1	5.02	122.81	120.30
1	A	356[B]	ARG	NE-CZ-NH1	5.02	122.81	120.30

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	3236	0	3209	18	0
2	A	16	0	10	0	0
3	A	12	0	18	0	0
4	A	49	0	70	2	0
5	A	13	0	18	0	0
6	A	214	0	0	2	0
All	All	3540	0	3325	18	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 3.

All (18) 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:201:GLN:HE21	1:A:224:LYS:HZ3	1.25	0.84
1:A:201:GLN:NE2	1:A:224:LYS:HZ3	1.90	0.68
1:A:201:GLN:NE2	1:A:224:LYS:NZ	2.46	0.64
1:A:146:GLU:HG2	6:A:713:HOH:O	1.97	0.63
1:A:337:ASP:O	1:A:338:ASP:HB2	2.06	0.56

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:222:SER:CB	1:A:231[A]:VAL:HG13	2.36	0.56
1:A:4:THR:OG1	4:A:506:PEG:H41	2.08	0.54
1:A:381:LEU:HD23	1:A:381:LEU:H	1.72	0.54
1:A:356[A]:ARG:HH11	1:A:376:ARG:HH22	1.62	0.47
1:A:108:LEU:HB2	1:A:135:THR:HG21	1.98	0.46
1:A:204:PRO:HB2	1:A:205:HIS:CD2	2.50	0.46
1:A:44:TYR:CZ	1:A:49:ASN:HA	2.51	0.45
1:A:201:GLN:HE21	1:A:224:LYS:NZ	2.01	0.45
1:A:113:GLU:HB2	1:A:163:THR:HA	1.98	0.44
1:A:85:GLU:O	1:A:235:TYR:HA	2.20	0.42
1:A:105:ARG:HD3	6:A:665:HOH:O	2.20	0.41
1:A:397:LYS:HB3	4:A:509:PEG:H41	2.03	0.41
1:A:27:SER:HA	1:A:30:THR:O	2.21	0.40

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	415/419 (99%)	403 (97%)	12 (3%)	0	100 100

There are no Ramachandran outliers to report.

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	347/348 (100%)	342 (99%)	5 (1%)	67 55

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	0	HIS
1	A	54	ARG
1	A	380	TYR
1	A	382	TYR
1	A	399	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	18	ASN
1	A	201	GLN
1	A	359	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 monosaccharides in this entry.

5.6 Ligand geometry [i](#)

12 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	EDO	A	503	-	3,3,3	0.13	0	2,2,2	0.45	0
2	PMP	A	501	-	16,16,16	0.92	0	21,23,23	1.18	2 (9%)
5	PG4	A	512	-	12,12,12	0.21	0	11,11,11	0.26	0
4	PEG	A	507	-	6,6,6	0.20	0	5,5,5	0.27	0
4	PEG	A	510	-	6,6,6	0.17	0	5,5,5	0.11	0
4	PEG	A	506	-	6,6,6	0.41	0	5,5,5	0.13	0
4	PEG	A	511	-	6,6,6	0.16	0	5,5,5	0.27	0
3	EDO	A	504	-	3,3,3	0.20	0	2,2,2	0.21	0
3	EDO	A	502	-	3,3,3	0.42	0	2,2,2	0.29	0
4	PEG	A	509	-	6,6,6	0.17	0	5,5,5	0.12	0
4	PEG	A	508	-	6,6,6	0.15	0	5,5,5	0.07	0
4	PEG	A	505	-	6,6,6	0.17	0	5,5,5	0.10	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	EDO	A	503	-	-	0/1/1/1	-
2	PMP	A	501	-	-	2/8/8/8	0/1/1/1
5	PG4	A	512	-	-	1/10/10/10	-
4	PEG	A	507	-	-	2/4/4/4	-
4	PEG	A	510	-	-	3/4/4/4	-
4	PEG	A	506	-	-	1/4/4/4	-
4	PEG	A	511	-	-	0/4/4/4	-
3	EDO	A	504	-	-	1/1/1/1	-
3	EDO	A	502	-	-	0/1/1/1	-
4	PEG	A	509	-	-	2/4/4/4	-
4	PEG	A	508	-	-	2/4/4/4	-
4	PEG	A	505	-	-	1/4/4/4	-

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	501	PMP	O3P-P-O4P	-2.79	99.32	106.73
2	A	501	PMP	O4P-C5A-C5	-2.60	104.40	109.35

There are no chirality outliers.

All (15) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	501	PMP	C3-C4-C4A-N4A
2	A	501	PMP	C5-C4-C4A-N4A
4	A	510	PEG	C1-C2-O2-C3
4	A	508	PEG	O1-C1-C2-O2
4	A	507	PEG	O1-C1-C2-O2
4	A	509	PEG	O2-C3-C4-O4
4	A	505	PEG	O2-C3-C4-O4
4	A	510	PEG	O2-C3-C4-O4
3	A	504	EDO	O1-C1-C2-O2
5	A	512	PG4	C5-C6-O4-C7
4	A	508	PEG	C4-C3-O2-C2
4	A	509	PEG	C4-C3-O2-C2
4	A	507	PEG	C1-C2-O2-C3
4	A	510	PEG	O1-C1-C2-O2
4	A	506	PEG	O1-C1-C2-O2

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	506	PEG	1	0
4	A	509	PEG	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	A	410/419 (97%)	0.07	20 (4%) 29 27	25, 38, 63, 116	0

All (20) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	406	PHE	10.0
1	A	407	VAL	7.9
1	A	409	LEU	7.0
1	A	405	VAL	7.0
1	A	408	ASP	5.1
1	A	401	TYR	4.4
1	A	404	ASN	4.0
1	A	402	PHE	3.5
1	A	338	ASP	3.5
1	A	231[A]	VAL	3.4
1	A	277	ILE	2.9
1	A	111	GLY	2.8
1	A	403	THR	2.8
1	A	220	LEU	2.5
1	A	55	GLY	2.5
1	A	318	ASP	2.4
1	A	337	ASP	2.2
1	A	278	ILE	2.1
1	A	110	PRO	2.1
1	A	399	LYS	2.1

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 monosaccharides 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(Å ²)	Q<0.9
4	PEG	A	509	7/7	0.81	0.39	77,80,90,93	0
3	EDO	A	503	4/4	0.87	0.10	54,59,62,67	0
4	PEG	A	511	7/7	0.88	0.19	67,72,83,84	0
4	PEG	A	510	7/7	0.89	0.14	65,70,73,84	0
4	PEG	A	508	7/7	0.90	0.14	63,70,82,84	0
4	PEG	A	507	7/7	0.91	0.14	57,61,72,72	0
5	PG4	A	512	13/13	0.91	0.13	55,67,74,76	0
3	EDO	A	504	4/4	0.92	0.09	53,61,61,66	0
4	PEG	A	505	7/7	0.92	0.11	51,67,72,77	0
3	EDO	A	502	4/4	0.93	0.07	50,55,59,60	0
4	PEG	A	506	7/7	0.94	0.13	44,51,58,68	0
2	PMP	A	501	16/16	0.98	0.12	30,33,44,53	0

6.5 Other polymers [i](#)

There are no such residues in this entry.