



wwPDB X-ray Structure Validation Summary Report ⓘ

Jul 14, 2026 – 01:17 pm BST

PDB ID : 9HZA / pdb_00009hza
Title : SMAD4 MH2, residues 272-552
Authors : Torner, C.; Condeminas, M.; Pluta, R.; Pous, J.; Macias, M.J.
Deposited on : 2025-01-13
Resolution : 2.03 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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity	:	4-5-2 with Phenix2.0
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.015 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.50

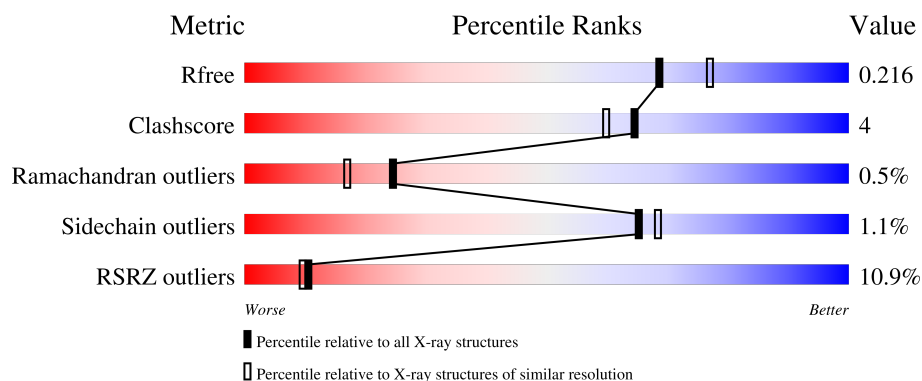
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.03 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}	180053	13299 (2.04-2.00)
Clashscore	190562	1022 (2.02-2.02)
Ramachandran outliers	187476	1014 (2.02-2.02)
Sidechain outliers	187428	1014 (2.02-2.02)
RSRZ outliers	180081	13314 (2.04-2.00)

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	282	5% 68% 8% 24%
1	B	282	13% 76% 7% 17%
1	C	282	8% 72% 9% 18%

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
2	SO4	C	604	-	-	X	-

2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 10863 atoms, of which 5227 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.

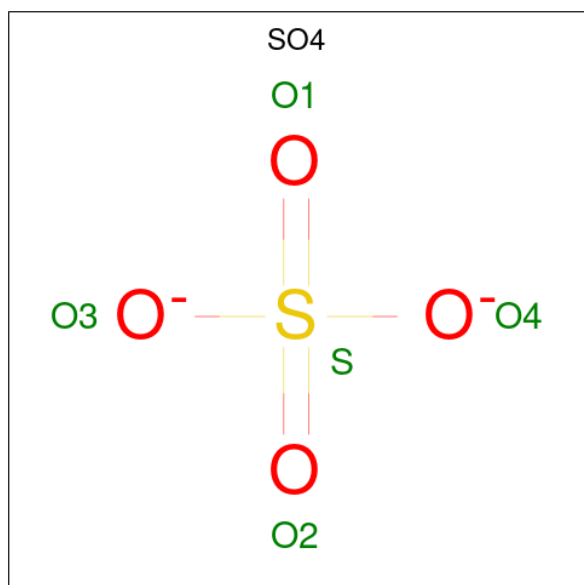
- Molecule 1 is a protein called Mothers against decapentaplegic homolog 4.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	215	Total	C	H	N	O	S	0	1	0
			3361	1078	1659	307	303	14			
1	B	235	Total	C	H	N	O	S	0	1	0
			3581	1147	1763	329	329	13			
1	C	230	Total	C	H	N	O	S	0	1	0
			3542	1132	1749	324	324	13			

There are 3 discrepancies between the modelled and reference sequences:

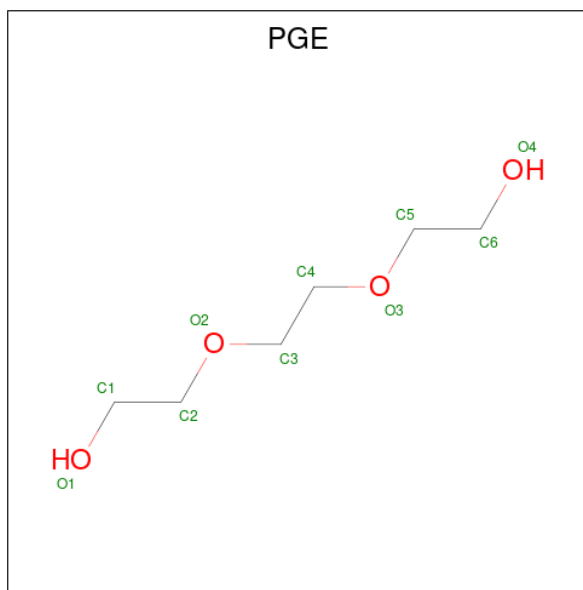
Chain	Residue	Modelled	Actual	Comment	Reference
A	271	GLY	-	expression tag	UNP Q13485
B	271	GLY	-	expression tag	UNP Q13485
C	271	GLY	-	expression tag	UNP Q13485

- Molecule 2 is SULFATE ION (CCD ID: SO4) (formula: O₄S).



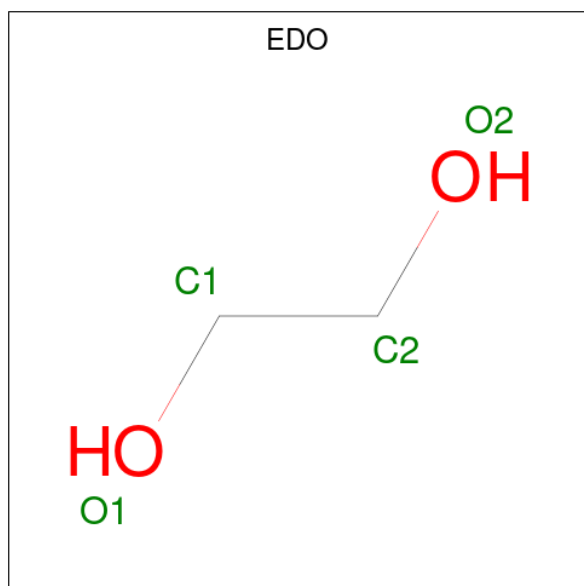
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total O S 5 4 1	0	0
2	A	1	Total O S 5 4 1	0	0
2	A	1	Total O S 5 4 1	0	0
2	A	1	Total O S 5 4 1	0	0
2	A	1	Total O S 5 4 1	0	0
2	B	1	Total O S 5 4 1	0	0
2	B	1	Total O S 5 4 1	0	0
2	B	1	Total O S 5 4 1	0	0
2	B	1	Total O S 5 4 1	0	0
2	B	1	Total O S 5 4 1	0	0
2	C	1	Total O S 5 4 1	0	0
2	C	1	Total O S 5 4 1	0	0

- Molecule 3 is TRIETHYLENE GLYCOL (CCD ID: PGE) (formula: C₆H₁₄O₄).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	H	O	0	0
			24	6	14	4		

- Molecule 4 is 1,2-ETHANEDIOL (CCD ID: EDO) (formula: $C_2H_6O_2$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	1	Total	C	H	O	0	0
			10	2	6	2		
4	A	1	Total	C	H	O	0	0
			10	2	6	2		
4	A	1	Total	C	H	O	0	0
			10	2	6	2		
4	B	1	Total	C	H	O	0	0
			10	2	6	2		
4	B	1	Total	C	H	O	0	0
			10	2	6	2		
4	C	1	Total	C	H	O	0	0
			10	2	6	2		
4	C	1	Total	C	H	O	0	0
			10	2	6	2		

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	73	Total	O	0	0
			73	73		
5	B	75	Total	O	0	0
			75	75		

Continued on next page...

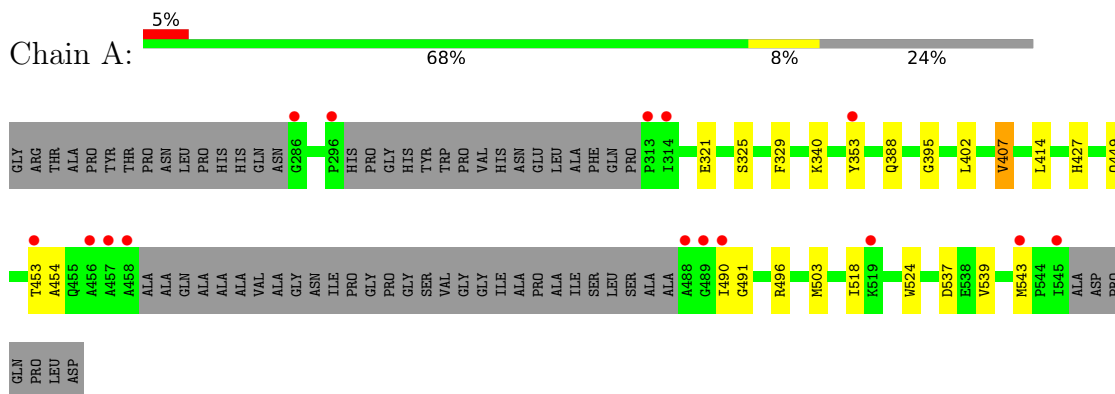
Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	C	77	Total	O	0	0
			77	77		

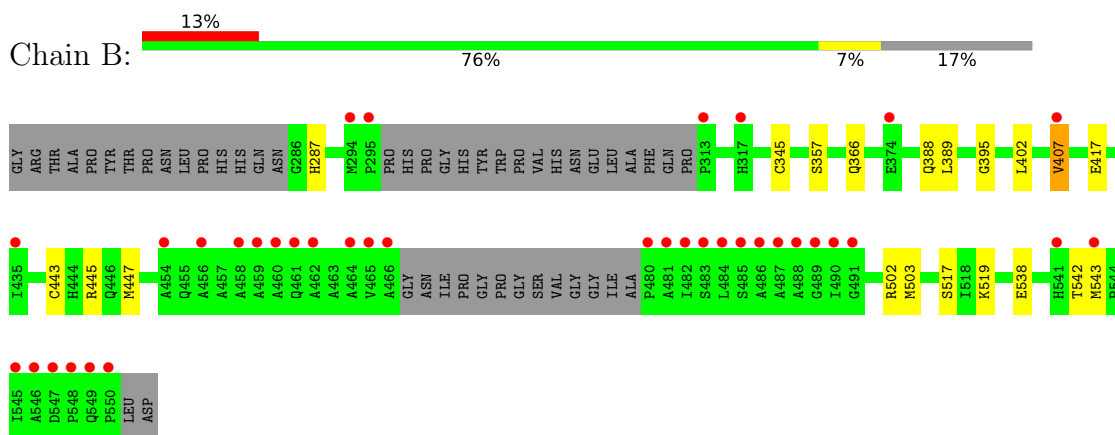
3 Residue-property plots [i](#)

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.

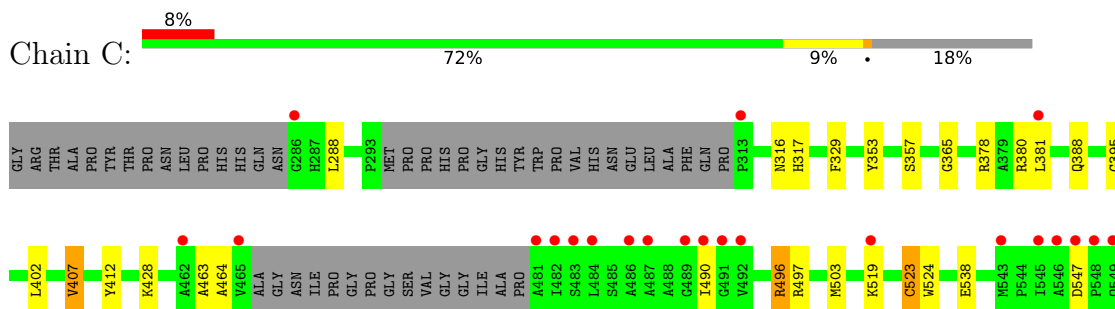
• Molecule 1: Mothers against decapentaplegic homolog 4



• Molecule 1: Mothers against decapentaplegic homolog 4



• Molecule 1: Mothers against decapentaplegic homolog 4



PRO
LEU
ASP

4 Data and refinement statistics

Property	Value	Source
Space group	I 41 2 2	Depositor
Cell constants a, b, c, α , β , γ	140.86Å 140.86Å 191.95Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	99.60 – 2.03 99.60 – 2.03	Depositor EDS
% Data completeness (in resolution range)	84.1 (99.60-2.03) 84.1 (99.60-2.03)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.90 (at 2.03Å)	Xtriage
Refinement program	PHENIX 1.21rc1_5107	Depositor
R, R_{free}	0.190 , 0.213 0.194 , 0.216	Depositor DCC
R_{free} test set	2613 reflections (4.22%)	wwPDB-VP
Wilson B-factor (Å ²)	36.6	Xtriage
Anisotropy	0.006	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.44 , 57.6	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	0.001 for -1/2*h+1/2*k-1/2*l,1/2*h-1/2*k-1/2*l,-h-k 0.007 for -1/2*h-1/2*k+1/2*l,-1/2*h-1/2*k-1/2*l,h-k	Xtriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	10863	wwPDB-VP
Average B, all atoms (Å ²)	51.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.88% 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: SO4, EDO, PGE

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	0.43	1/1752 (0.1%)	0.50	0/2378
1	B	0.27	0/1875	0.47	0/2549
1	C	0.35	0/1842	0.51	0/2502
All	All	0.36	1/5469 (0.0%)	0.49	0/7429

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	C	0	2

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	353	TYR	C-O	-11.26	1.12	1.23

There are no bond angle outliers.

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	C	496	ARG	Sidechain
1	C	497	ARG	Sidechain

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	1702	1659	1659	15	0
1	B	1818	1763	1758	13	1
1	C	1793	1749	1749	16	0
2	A	25	0	0	1	0
2	B	25	0	0	0	0
2	C	10	0	0	2	0
3	A	10	14	14	0	0
4	A	12	18	18	2	0
4	B	8	12	12	0	0
4	C	8	12	12	1	0
5	A	73	0	0	0	0
5	B	75	0	0	3	0
5	C	77	0	0	2	0
All	All	5636	5227	5222	41	1

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 41 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:B:345:CYS:SG	5:B:771:HOH:O	2.34	0.84
1:C:490:ILE:HD12	1:C:490:ILE:O	1.84	0.77
1:A:449:GLN:O	1:A:453:THR:HG23	1.90	0.70
2:C:604:SO4:O2	5:C:701:HOH:O	2.07	0.69
1:B:517:SER:OG	1:B:519:LYS:HG3	1.98	0.64

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:417:GLU:OE2	1:B:445:ARG:HH11[5_454]	1.59	0.01

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	210/282 (74%)	203 (97%)	6 (3%)	1 (0%)	24	17
1	B	230/282 (82%)	221 (96%)	8 (4%)	1 (0%)	30	22
1	C	225/282 (80%)	217 (96%)	7 (3%)	1 (0%)	30	22
All	All	665/846 (79%)	641 (96%)	21 (3%)	3 (0%)	24	17

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	395	GLY
1	C	395	GLY
1	A	395	GLY

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	183/229 (80%)	182 (100%)	1 (0%)	81	81
1	B	191/229 (83%)	190 (100%)	1 (0%)	81	81
1	C	189/229 (82%)	184 (97%)	5 (3%)	40	38
All	All	563/687 (82%)	556 (99%)	7 (1%)	65	65

5 of 7 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	519	LYS
1	C	523[A]	CYS
1	C	547	ASP
1	C	523[B]	CYS
1	C	407	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 10 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	410	GLN
1	C	455	GLN
1	C	516	GLN
1	B	316	ASN
1	B	405	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 oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

20 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	SO4	C	604	-	4,4,4	0.75	0	6,6,6	0.39	0
2	SO4	A	608	-	4,4,4	0.46	0	6,6,6	0.05	0
4	EDO	C	602	-	3,3,3	0.29	0	2,2,2	0.32	0
3	PGE	A	603	-	9,9,9	0.11	0	8,8,8	0.14	0
2	SO4	B	607	-	4,4,4	0.81	0	6,6,6	0.50	0
4	EDO	A	604	-	3,3,3	0.31	0	2,2,2	0.37	0
2	SO4	C	603	-	4,4,4	0.46	0	6,6,6	0.10	0
4	EDO	B	603	-	3,3,3	0.26	0	2,2,2	0.26	0
2	SO4	B	602	-	4,4,4	0.59	0	6,6,6	0.07	0
2	SO4	B	606	-	4,4,4	0.54	0	6,6,6	0.11	0
2	SO4	A	602	-	4,4,4	0.65	0	6,6,6	0.08	0
2	SO4	A	601	-	4,4,4	0.53	0	6,6,6	0.28	0
2	SO4	B	604	-	4,4,4	0.58	0	6,6,6	0.32	0
2	SO4	A	609	-	4,4,4	0.47	0	6,6,6	0.07	0
4	EDO	A	605	-	3,3,3	0.31	0	2,2,2	0.34	0
2	SO4	B	605	-	4,4,4	0.47	0	6,6,6	0.04	0
2	SO4	A	606	-	4,4,4	0.47	0	6,6,6	0.09	0
4	EDO	C	601	-	3,3,3	0.28	0	2,2,2	0.29	0
4	EDO	A	607	-	3,3,3	0.24	0	2,2,2	0.27	0
4	EDO	B	601	-	3,3,3	0.09	0	2,2,2	0.14	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	EDO	C	602	-	-	0/1/1/1	-
3	PGE	A	603	-	-	5/7/7/7	-
4	EDO	B	603	-	-	0/1/1/1	-
4	EDO	A	604	-	-	0/1/1/1	-
4	EDO	C	601	-	-	0/1/1/1	-
4	EDO	A	607	-	-	1/1/1/1	-
4	EDO	B	601	-	-	0/1/1/1	-
4	EDO	A	605	-	-	0/1/1/1	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 6 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	603	PGE	O2-C3-C4-O3
3	A	603	PGE	O1-C1-C2-O2
3	A	603	PGE	O3-C5-C6-O4
4	A	607	EDO	O1-C1-C2-O2
3	A	603	PGE	C3-C4-O3-C5

There are no ring outliers.

4 monomers are involved in 6 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	C	604	SO4	2	0
2	A	609	SO4	1	0
4	C	601	EDO	1	0
4	A	607	EDO	2	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	215/282 (76%)	0.24	15 (6%) 22 21	23, 45, 84, 111	1 (0%)
1	B	235/282 (83%)	0.66	37 (15%) 5 4	29, 49, 100, 121	0
1	C	230/282 (81%)	0.35	22 (9%) 13 12	21, 46, 89, 104	1 (0%)
All	All	680/846 (80%)	0.42	74 (10%) 10 9	21, 47, 96, 121	2 (0%)

The worst 5 of 74 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	482	ILE	7.0
1	B	484	LEU	6.6
1	B	546	ALA	6.4
1	A	458	ALA	6.3
1	B	550	PRO	4.7

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 oligosaccharides 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($\text{\AA}^2$)	Q<0.9
2	SO4	A	606	5/5	0.66	0.17	85,86,101,116	0
2	SO4	B	605	5/5	0.67	0.14	84,86,91,118	0
2	SO4	C	603	5/5	0.67	0.14	72,76,87,94	0
2	SO4	A	609	5/5	0.74	0.17	88,92,102,121	0
2	SO4	B	604	5/5	0.79	0.22	44,59,60,61	5
2	SO4	C	604	5/5	0.79	0.27	47,60,75,78	5
4	EDO	A	604	4/4	0.81	0.21	53,64,67,71	0
2	SO4	B	606	5/5	0.83	0.10	63,67,80,85	5
2	SO4	A	608	5/5	0.86	0.12	98,107,115,125	5
3	PGE	A	603	10/10	0.87	0.15	51,61,76,77	0
2	SO4	B	607	5/5	0.87	0.25	47,55,71,91	5
4	EDO	A	607	4/4	0.88	0.16	40,52,62,66	0
4	EDO	C	601	4/4	0.88	0.16	46,56,65,70	0
4	EDO	B	603	4/4	0.90	0.13	40,50,67,67	0
4	EDO	A	605	4/4	0.90	0.14	55,66,76,76	0
2	SO4	A	602	5/5	0.91	0.12	34,49,54,61	5
4	EDO	B	601	4/4	0.92	0.12	58,70,83,83	0
4	EDO	C	602	4/4	0.93	0.10	58,70,74,85	0
2	SO4	A	601	5/5	0.96	0.11	41,44,53,53	0
2	SO4	B	602	5/5	0.99	0.05	48,49,51,52	5

6.5 Other polymers [i](#)

There are no such residues in this entry.