wwPDB EM Validation Summary Report

Apr 3, 2023 – 04:10 PM EDT

PDB ID	:	7UJJ
EMDB ID	:	EMD-26563
Title	:	Stx2a and DARPin complex
Authors	:	Jiang, M.; Zhang, J.
Deposited on	:	2022-03-30
Resolution	:	6.50 Å(reported)

This is a wwPDB EM 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/EMValidationReportHelp
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:

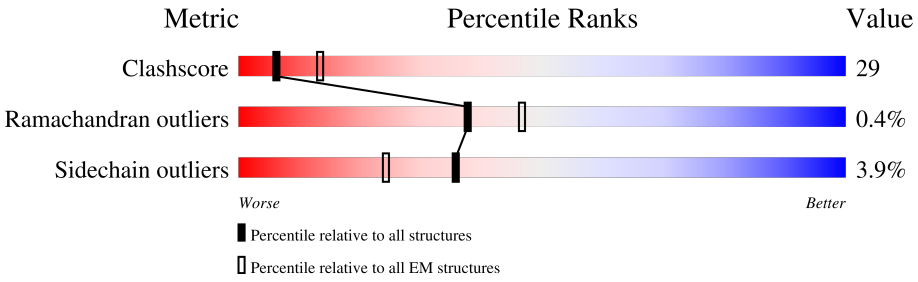
EMDB validation analysis	:	0.0.1.dev50
Mogul	:	1.8.5 (274361), CSD as541be (2020)
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ	:	FAILED
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.32.2

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
ELECTRON MICROSCOPY

The reported resolution of this entry is 6.50 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)	EM structures (#Entries)
Clashscore	158937	4297
Ramachandran outliers	154571	4023
Sidechain outliers	154315	3826

The table below summarises the geometric issues observed across the polymeric chains and their fit to the map. The red, orange, yellow and green segments of the 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 <=5%

Mol	Chain	Length	Quality of chain
1	G	185	70% 25%
2	A	297	82% 12% 5%
3	B	70	64% 36%
3	C	70	79% 19%
3	D	70	90% 10%
3	E	70	86% 14%
3	F	70	86% 14%

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 crite-

ria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
4	1PS	G	201	-	-	X	-
5	FMT	G	202	-	-	X	-
7	EDO	D	102	-	X	-	-
7	EDO	F	104	-	X	-	-

2 Entry composition (i)

There are 8 unique types of molecules in this entry. The entry contains 8403 atoms, of which 1388 are hydrogens and 0 are deuteriums.

In the tables below, 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 DARPin.

Mol	Chain	Residues	Atoms						AltConf	Trace
1	G	185	Total	C	H	N	O	S	0	0
			2786	875	1388	253	268	2		

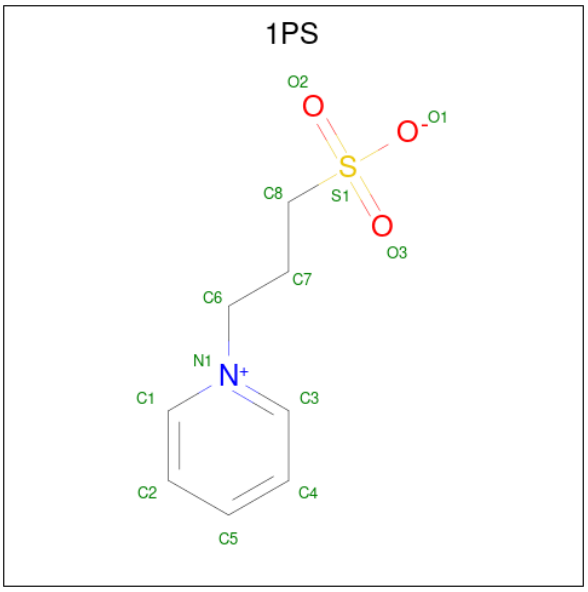
- Molecule 2 is a protein called Shiga-like toxin 2 subunit A.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	A	281	Total	C	N	O	S	5	0
			2225	1393	392	433	7		

- Molecule 3 is a protein called Shiga-like toxin 2 subunit B.

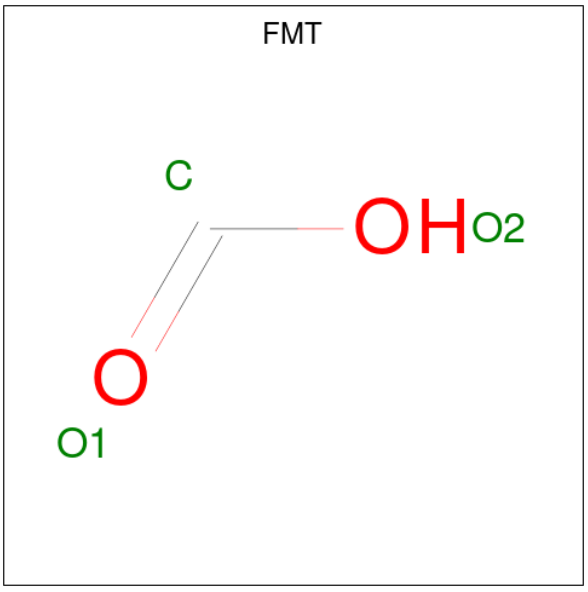
Mol	Chain	Residues	Atoms					AltConf	Trace
3	B	70	Total	C	N	O	S	0	0
			549	342	89	115	3		
3	C	70	Total	C	N	O	S	4	0
			559	348	89	119	3		
3	D	70	Total	C	N	O	S	0	0
			549	342	89	115	3		
3	E	70	Total	C	N	O	S	0	0
			549	342	89	115	3		
3	F	70	Total	C	N	O	S	1	0
			550	342	89	116	3		

- Molecule 4 is 3-PYRIDINIUM-1-YLPROPANE-1-SULFONATE (three-letter code: 1PS) (formula: $C_8H_{11}NO_3S$).



Mol	Chain	Residues	Atoms					AltConf
4	G	1	Total	C	N	O	S	0
			13	8	1	3	1	
4	C	1	Total	C	N	O	S	0
			13	8	1	3	1	
4	D	1	Total	C	N	O	S	0
			13	8	1	3	1	
4	F	1	Total	C	N	O	S	0
			13	8	1	3	1	

- Molecule 5 is FORMIC ACID (three-letter code: FMT) (formula: CH_2O_2).

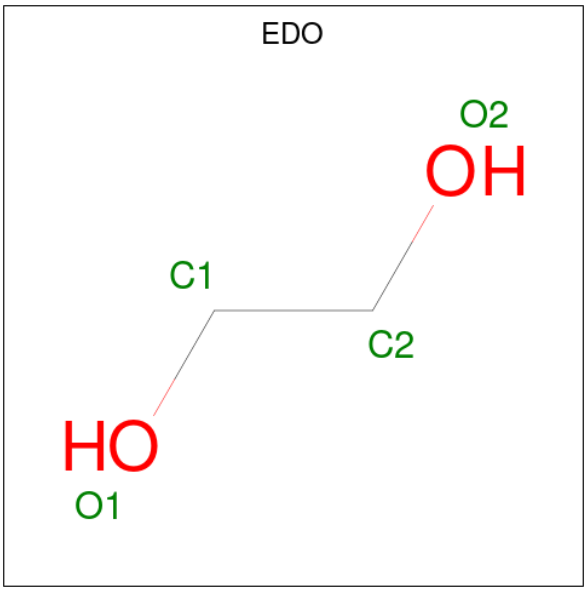


Mol	Chain	Residues	Atoms	AltConf
5	G	1	Total C O 3 1 2	0
5	A	1	Total C O 3 1 2	0
5	A	1	Total C O 3 1 2	0
5	A	1	Total C O 3 1 2	0
5	A	1	Total C O 3 1 2	0
5	A	1	Total C O 3 1 2	0
5	A	1	Total C O 6 2 4	1
5	B	1	Total C O 3 1 2	0
5	C	1	Total C O 3 1 2	0
5	C	1	Total C O 3 1 2	0
5	D	1	Total C O 3 1 2	0
5	E	1	Total C O 3 1 2	0
5	E	1	Total C O 3 1 2	0
5	E	1	Total C O 3 1 2	0
5	F	1	Total C O 3 1 2	0

- Molecule 6 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms	AltConf
6	A	3	Total Na 3 3	0
6	B	1	Total Na 1 1	0
6	F	1	Total Na 1 1	0

- Molecule 7 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: $C_2H_6O_2$).



Mol	Chain	Residues	Atoms	AltConf
7	A	1	Total C O 4 2 2	0
7	A	1	Total C O 4 2 2	0
7	D	1	Total C O 4 2 2	0
7	F	1	Total C O 4 2 2	0
7	F	1	Total C O 4 2 2	0

- Molecule 8 is water.

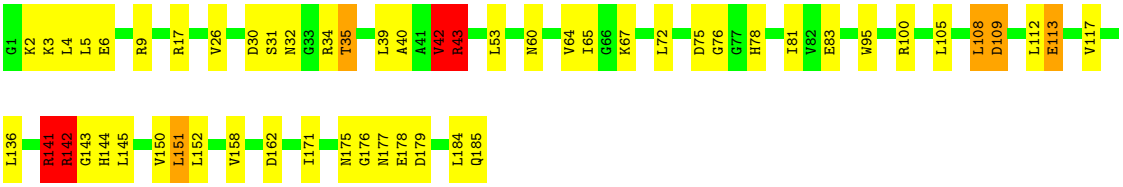
Mol	Chain	Residues	Atoms	AltConf
8	G	13	Total O 13 13	0
8	A	234	Total O 234 234	0
8	B	42	Total O 42 42	0
8	C	71	Total O 71 71	0
8	D	60	Total O 60 60	0
8	E	39	Total O 39 39	0
8	F	52	Total O 52 52	0

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. 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. 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.

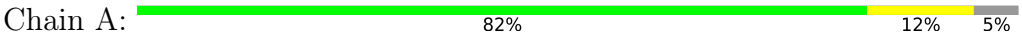
• Molecule 1: DARPin

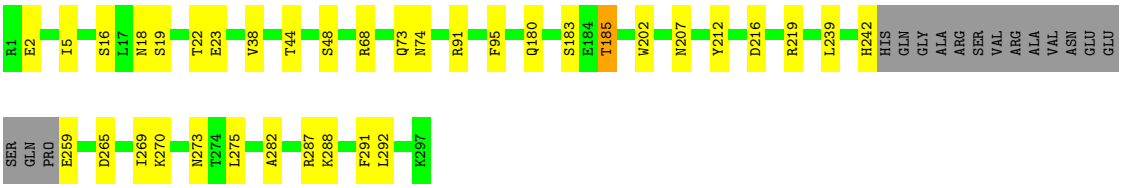
Chain G:



• Molecule 2: Shiga-like toxin 2 subunit A

Chain A:



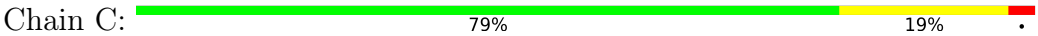


• Molecule 3: Shiga-like toxin 2 subunit B

Chain B:

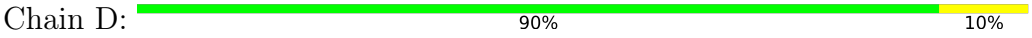
• Molecule 3: Shiga-like toxin 2 subunit B

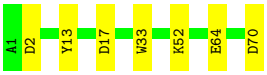
Chain C:



• Molecule 3: Shiga-like toxin 2 subunit B

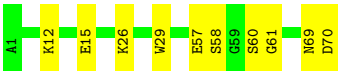
Chain D:





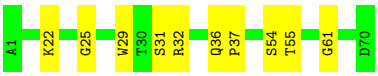
• Molecule 3: Shiga-like toxin 2 subunit B

Chain E:



• Molecule 3: Shiga-like toxin 2 subunit B

Chain F:



4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	108744	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	FEI TALOS ARCTICA	Depositor
Voltage (kV)	200	Depositor
Electron dose ($e^-/Å^2$)	50	Depositor
Minimum defocus (nm)	1000	Depositor
Maximum defocus (nm)	3000	Depositor
Magnification	Not provided	
Image detector	GATAN K2 SUMMIT (4k x 4k)	Depositor

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: FMT, 1PS, EDO, NA

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	G	0.61	0/1420	1.15	7/1930 (0.4%)
2	A	0.77	0/2291	0.98	5/3114 (0.2%)
3	B	0.78	0/559	0.90	0/753
3	C	0.84	0/589	1.04	3/794 (0.4%)
3	D	0.90	0/559	0.96	0/753
3	E	0.84	0/559	0.94	0/753
3	F	0.77	0/565	0.87	0/761
All	All	0.76	0/6542	1.01	15/8858 (0.2%)

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	G	0	4

There are no bond length outliers.

The worst 5 of 15 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	A	219	ARG	NE-CZ-NH2	-8.37	116.11	120.30
1	G	142	ARG	NE-CZ-NH1	8.22	124.41	120.30
3	C	44[A]	LEU	CA-CB-CG	8.18	134.11	115.30
3	C	44[B]	LEU	CA-CB-CG	8.18	134.11	115.30
1	G	141	ARG	NE-CZ-NH1	7.75	124.17	120.30

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	G	100	ARG	Sidechain
1	G	141	ARG	Sidechain
1	G	150	VAL	Peptide
1	G	151	LEU	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	G	1398	1388	1384	294	0
2	A	2225	0	2192	32	0
3	B	549	0	511	184	0
3	C	559	0	531	42	0
3	D	549	0	518	6	0
3	E	549	0	518	14	0
3	F	550	0	514	34	0
4	C	13	0	11	0	0
4	D	13	0	11	0	0
4	F	13	0	11	0	0
4	G	13	0	9	29	0
5	A	21	0	7	1	0
5	B	3	0	1	0	0
5	C	6	0	2	0	0
5	D	3	0	1	0	0
5	E	9	0	3	1	0
5	F	3	0	1	0	0
5	G	3	0	1	4	0
6	A	3	0	0	0	0
6	B	1	0	0	0	0
6	F	1	0	0	0	0
7	A	8	0	12	3	0
7	D	4	0	6	2	0
7	F	8	0	12	3	0
8	A	234	0	0	2	0
8	B	42	0	0	11	0
8	C	71	0	0	0	0
8	D	60	0	0	0	0
8	E	39	0	0	1	0
8	F	52	0	0	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
8	G	13	0	0	27	0
All	All	7015	1388	6256	369	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 29.

The worst 5 of 369 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:G:142:ARG:CZ	3:B:14:ASN:CB	1.82	1.56
1:G:9:ARG:HH21	3:B:1:ALA:CB	1.14	1.53
1:G:75:ASP:CG	3:B:29:TRP:HZ2	1.06	1.50
1:G:5:LEU:CD1	3:B:1:ALA:HB2	1.42	1.49
1:G:39:LEU:HD13	3:B:55:THR:CA	1.37	1.49

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 PDB entries followed by that with respect to all EM entries.

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	G	183/185 (99%)	168 (92%)	12 (7%)	3 (2%)	9 44
2	A	282/297 (95%)	277 (98%)	5 (2%)	0	100 100
3	B	68/70 (97%)	67 (98%)	1 (2%)	0	100 100
3	C	72/70 (103%)	72 (100%)	0	0	100 100
3	D	68/70 (97%)	67 (98%)	1 (2%)	0	100 100
3	E	68/70 (97%)	66 (97%)	2 (3%)	0	100 100
3	F	69/70 (99%)	68 (99%)	1 (1%)	0	100 100
All	All	810/832 (97%)	785 (97%)	22 (3%)	3 (0%)	38 72

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	G	142	ARG
1	G	152	LEU
1	G	151	LEU

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 PDB entries followed by that with respect to all EM entries.

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	G	143/143 (100%)	123 (86%)	20 (14%)	3 17
2	A	253/261 (97%)	251 (99%)	2 (1%)	81 89
3	B	61/61 (100%)	60 (98%)	1 (2%)	62 79
3	C	65/61 (107%)	61 (94%)	4 (6%)	18 43
3	D	61/61 (100%)	60 (98%)	1 (2%)	62 79
3	E	61/61 (100%)	61 (100%)	0	100 100
3	F	62/61 (102%)	62 (100%)	0	100 100
All	All	706/709 (100%)	678 (96%)	28 (4%)	36 55

5 of 28 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	G	113	GLU
3	D	70	ASP
1	G	142	ARG
3	C	44[A]	LEU
1	G	136	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 6 such sidechains are listed below:

Mol	Chain	Res	Type
2	A	180	GLN
2	A	207	ASN
2	A	226	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	A	42	ASN
1	G	177	ASN

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)

Of 30 ligands modelled in this entry, 5 are monoatomic - leaving 25 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	$\#\|Z\| > 2$
5	FMT	A	4006	-	2,2,2	0.57	0	1,1,1	0.05	0
5	FMT	A	4009	-	2,2,2	0.42	0	1,1,1	0.11	0
5	FMT	A	4008	-	2,2,2	0.53	0	1,1,1	0.17	0
5	FMT	A	4007	6	2,2,2	0.64	0	1,1,1	0.10	0
5	FMT	E	101	-	2,2,2	0.58	0	1,1,1	0.06	0
4	1PS	F	102	-	13,13,13	0.95	1 (7%)	17,17,17	0.55	0
5	FMT	E	103	-	2,2,2	0.76	0	1,1,1	0.23	0
7	EDO	A	4005	-	3,3,3	2.17	2 (66%)	2,2,2	0.68	0
5	FMT	E	102	-	2,2,2	0.61	0	1,1,1	0.06	0
5	FMT	A	4010	-	2,2,2	0.54	0	1,1,1	0.15	0
5	FMT	F	105	-	2,2,2	0.37	0	1,1,1	0.09	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\|>2$	Counts	RMSZ	$\#\|Z\|>2$
7	EDO	A	4004	-	3,3,3	1.45	0	2,2,2	0.90	0
7	EDO	F	104	-	3,3,3	2.43	2 (66%)	2,2,2	0.46	0
5	FMT	D	103	-	2,2,2	0.59	0	1,1,1	0.03	0
5	FMT	C	103	-	2,2,2	0.57	0	1,1,1	0.11	0
5	FMT	B	102	-	2,2,2	0.38	0	1,1,1	0.12	0
5	FMT	A	4011[A]	-	2,2,2	0.43	0	1,1,1	0.32	0
4	1PS	G	201	-	13,13,13	0.91	0	17,17,17	0.41	0
7	EDO	D	102	-	3,3,3	2.19	2 (66%)	2,2,2	0.57	0
5	FMT	C	102	-	2,2,2	0.64	0	1,1,1	0.06	0
7	EDO	F	103	-	3,3,3	1.75	1 (33%)	2,2,2	0.30	0
5	FMT	G	202	-	2,2,2	0.61	0	1,1,1	0.16	0
5	FMT	A	4011[B]	-	2,2,2	0.71	0	1,1,1	0.17	0
4	1PS	D	101	-	13,13,13	1.00	0	17,17,17	0.50	0
4	1PS	C	101	-	13,13,13	0.57	0	17,17,17	0.85	1 (5%)

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
7	EDO	A	4004	-	-	1/1/1/1	-
7	EDO	F	104	-	-	1/1/1/1	-
7	EDO	F	103	-	-	0/1/1/1	-
4	1PS	G	201	-	-	0/7/7/7	0/1/1/1
7	EDO	A	4005	-	-	0/1/1/1	-
4	1PS	D	101	-	-	0/7/7/7	0/1/1/1
4	1PS	C	101	-	-	0/7/7/7	0/1/1/1
4	1PS	F	102	-	-	0/7/7/7	0/1/1/1
7	EDO	D	102	-	-	1/1/1/1	-

The worst 5 of 8 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
7	F	104	EDO	O2-C2	3.40	1.59	1.42
7	A	4005	EDO	O2-C2	3.00	1.57	1.42
7	D	102	EDO	O2-C2	2.84	1.56	1.42
7	F	103	EDO	O2-C2	2.65	1.55	1.42
7	D	102	EDO	O1-C1	2.39	1.54	1.42

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	C	101	1PS	O3-S1-C8	-2.04	104.46	106.92

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
7	A	4004	EDO	O1-C1-C2-O2
7	D	102	EDO	O1-C1-C2-O2
7	F	104	EDO	O1-C1-C2-O2

There are no ring outliers.

9 monomers are involved in 43 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	A	4007	FMT	1	0
5	E	103	FMT	1	0
7	A	4005	EDO	1	0
7	A	4004	EDO	2	0
7	F	104	EDO	2	0
4	G	201	1PS	29	0
7	D	102	EDO	2	0
7	F	103	EDO	2	0
5	G	202	FMT	4	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 Map visualisation

This section contains visualisations of the EMDB entry EMD-26563. These allow visual inspection of the internal detail of the map and identification of artifacts.

Images derived from a raw map, generated by summing the deposited half-maps, are presented below the corresponding image components of the primary map to allow further visual inspection and comparison with those of the primary map.

6.1 Orthogonal projections

This section was not generated.

6.2 Central slices

This section was not generated.

6.3 Largest variance slices

This section was not generated.

6.4 Orthogonal standard-deviation projections (False-color)

This section was not generated.

6.5 Orthogonal surface views

This section was not generated.

6.6 Mask visualisation

This section was not generated. No masks/segmentation were deposited.

7 Map analysis

This section contains the results of statistical analysis of the map.

7.1 Map-value distribution

This section was not generated.

7.2 Volume estimate versus contour level

This section was not generated.

7.3 Rotationally averaged power spectrum

This section was not generated. The rotationally averaged power spectrum had issues being displayed.

8 Fourier-Shell correlation

This section was not generated. No FSC curve or half-maps provided.

9 Map-model fit

This section was not generated.