WORLDWIDE wwPDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

Dec 16, 2025 – 06:13 PM EST

PDB ID	:	9MK3 / pdb_00009mk3
Title	:	Engineered AsCas12a (M537R, F870L)
Authors	:	Dementiev, A.; White, A.; Olland, A.M.; Suto, R.K.
Deposited on	:	2024-12-16
Resolution	:	2.70 Å(reported)

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

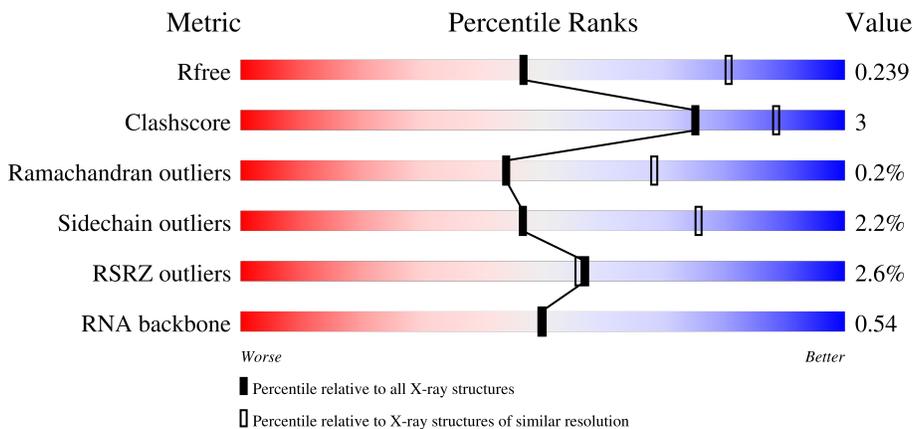|		
MolProbity	:	4-5-2 with Phenix2.0
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.47

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 2.70 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	Percentile Ranks	Value
Rfree		0.239
Clashscore		3
Ramachandran outliers		0.2%
Sidechain outliers		2.2%
RSRZ outliers		2.6%
RNA backbone		0.54

Worse — *Better*

■ Percentile relative to all X-ray structures
▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	164625	3333 (2.70-2.70)
Clashscore	180529	3684 (2.70-2.70)
Ramachandran outliers	177936	3633 (2.70-2.70)
Sidechain outliers	177891	3633 (2.70-2.70)
RSRZ outliers	164620	3333 (2.70-2.70)
RNA backbone	3690	1028 (2.94-2.46)

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 <=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	1307	2% / 89% / 9% •
1	E	1307	3% / 89% / 10% •
2	B	44	2% / 70% / 23% / 7%
2	F	44	73% / 18% • 7%

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	C	34	3% 71% 21% 9%
3	G	34	76% 12% 12%
4	D	10	90% 10%
4	H	10	100%

2 Entry composition ⓘ

There are 8 unique types of molecules in this entry. The entry contains 24407 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 CRISPR-associated endonuclease Cas12a.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	1292	Total 10430	C 6684	N 1764	O 1958	S 24	0	4	0
1	E	1295	Total 10436	C 6681	N 1769	O 1962	S 24	0	6	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	537	ARG	MET	engineered mutation	UNP U2UMQ6
A	870	LEU	PHE	engineered mutation	UNP U2UMQ6
E	537	ARG	MET	engineered mutation	UNP U2UMQ6
E	870	LEU	PHE	engineered mutation	UNP U2UMQ6

- Molecule 2 is a RNA chain called gRNA (41-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	41	Total 867	C 389	N 148	O 290	P 40	0	0	0
2	F	41	Total 849	C 379	N 143	O 287	P 40	0	0	0

- Molecule 3 is a DNA chain called TS DNA (31-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	31	Total 633	C 301	N 119	O 182	P 31	0	0	0
3	G	30	Total 613	C 291	N 117	O 175	P 30	0	0	0

- Molecule 4 is a DNA chain called NTS DNA (5'-D(*CP*AP*GP*TP*CP*CP*TP*TP*TP*T)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	10	Total 197	C 97	N 29	O 62	P 9	0	0	0
4	H	10	Total 197	C 97	N 29	O 62	P 9	0	0	0

- Molecule 5 is SODIUM ION (CCD ID: NA) (formula: Na).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	1	Total 1	Na 1	0	0
5	E	1	Total 1	Na 1	0	0

- Molecule 6 is CHLORIDE ION (CCD ID: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	1	Total 1	Cl 1	0	0
6	E	1	Total 1	Cl 1	0	0

- Molecule 7 is 1,2-ETHANEDIOL (CCD ID: EDO) (formula: $C_2H_6O_2$).
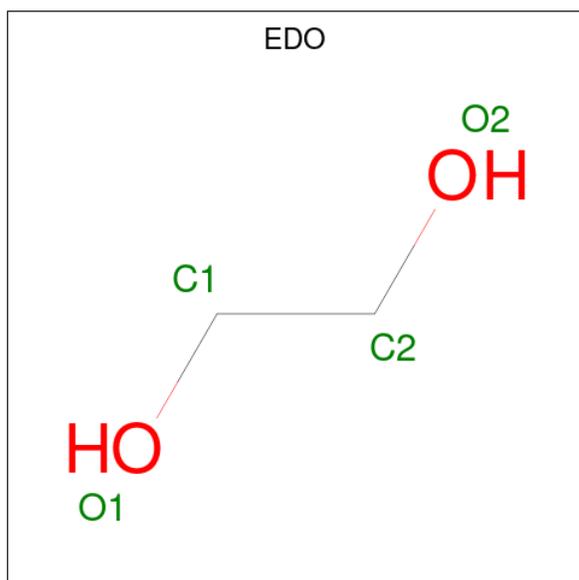
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	A	1	Total 4	C 2	O 2	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	A	1	Total C O 4 2 2			0	0
7	A	1	Total C O 4 2 2			0	0
7	A	1	Total C O 4 2 2			0	0
7	A	1	Total C O 4 2 2			0	0
7	A	1	Total C O 4 2 2			0	0
7	A	1	Total C O 4 2 2			0	0
7	E	1	Total C O 4 2 2			0	0
7	E	1	Total C O 4 2 2			0	0
7	E	1	Total C O 4 2 2			0	0
7	G	1	Total C O 4 2 2			0	0

- Molecule 8 is water.

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	A	34	Total O 34 34		0	0
8	B	17	Total O 17 17		0	0
8	C	5	Total O 5 5		0	0
8	E	51	Total O 51 51		0	0
8	F	21	Total O 21 21		0	0
8	G	9	Total O 9 9		0	0

3 Residue-property plots ⓘ

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: CRISPR-associated endonuclease Cas12a
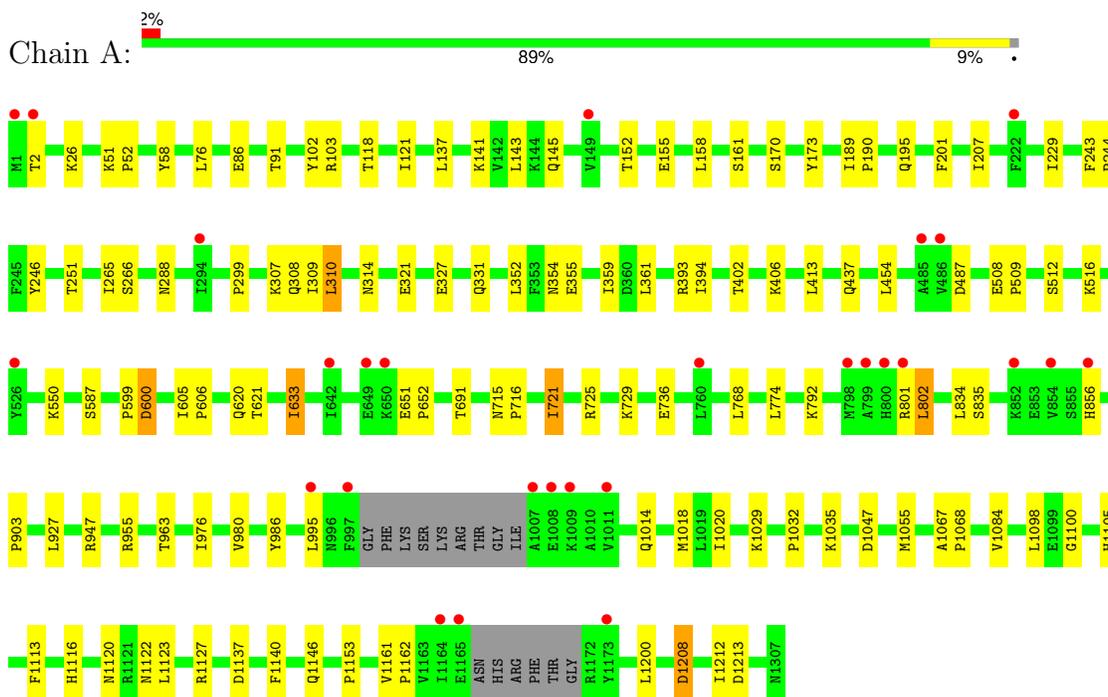
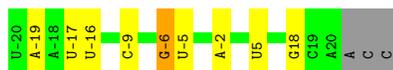• Molecule 1: CRISPR-associated endonuclease Cas12a
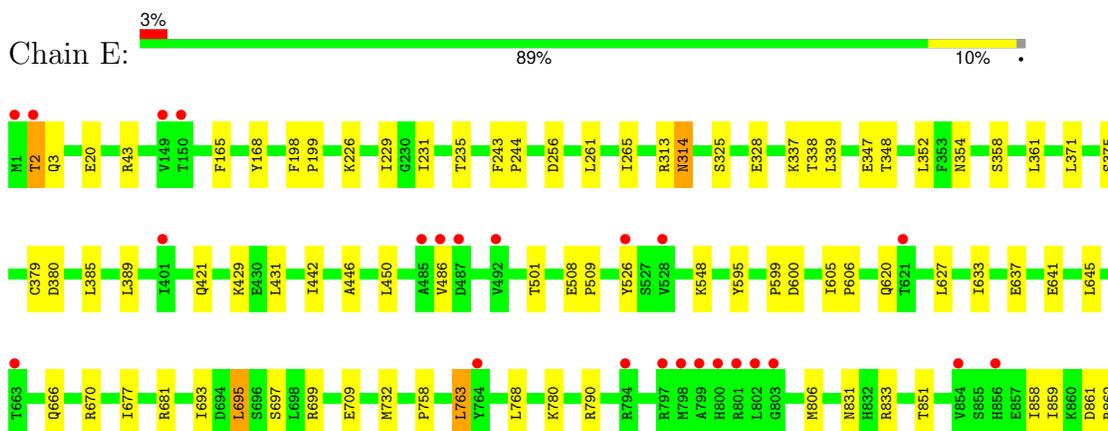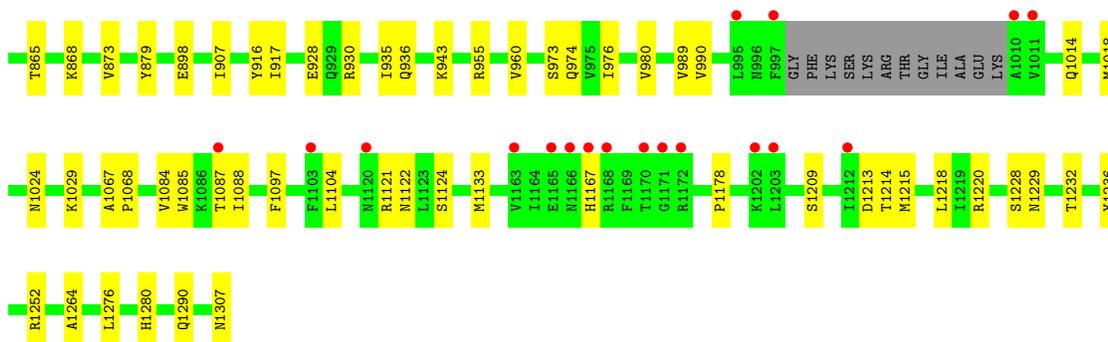
- Molecule 2: gRNA (41-MER)

Chain B:

- Molecule 2: gRNA (41-MER)

Chain F:

- Molecule 3: TS DNA (31-MER)

Chain C:

- Molecule 3: TS DNA (31-MER)

Chain G:

- Molecule 4: NTS DNA (5'-D(*CP*AP*GP*TP*CP*CP*TP*TP*TP*T)-3')

Chain D:

- Molecule 4: NTS DNA (5'-D(*CP*AP*GP*TP*CP*CP*TP*TP*TP*T)-3')

Chain H:

There are no outlier residues recorded for this chain.

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	81.92Å 198.59Å 140.36Å 90.00° 95.27° 90.00°	Depositor
Resolution (Å)	49.28 – 2.70 49.28 – 2.70	Depositor EDS
% Data completeness (in resolution range)	98.9 (49.28-2.70) 98.9 (49.28-2.70)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.36 (at 2.69Å)	Xtriage
Refinement program	REFMAC 5.8.0430 (refmacat 0.4.82)	Depositor
R, R_{free}	0.196 , 0.240 0.198 , 0.239	Depositor DCC
R_{free} test set	6068 reflections (4.95%)	wwPDB-VP
Wilson B-factor (Å^2)	61.7	Xtriage
Anisotropy	0.140	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , 41.6	EDS
L-test for twinning[2]	$< |L| > = 0.50, < L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	24407	wwPDB-VP
Average B, all atoms (Å^2)	68.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.94% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.
[2]Theoretical values of $< |L| >$, $< L^2 >$ 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: NA, EDO, CL

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	0.47	0/10676	0.93	3/14453 (0.0%)
1	E	0.47	0/10688	0.93	1/14475 (0.0%)
2	B	0.49	0/968	1.00	1/1506 (0.1%)
2	F	0.49	0/947	1.07	4/1474 (0.3%)
3	C	0.37	0/710	0.90	2/1092 (0.2%)
3	G	0.38	0/688	0.91	0/1058
4	D	0.36	0/218	1.04	1/334 (0.3%)
4	H	0.36	0/218	0.99	0/334
All	All	0.46	0/25113	0.94	12/34726 (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

There are no bond length outliers.

The worst 5 of 12 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	F	18	G	O3'-P-O5'	-6.01	94.99	104.00
2	F	5	U	O3'-P-O5'	-5.78	95.32	104.00
2	F	-17	U	O3'-P-O5'	-5.69	95.47	104.00
2	F	-19	A	O3'-P-O5'	-5.64	95.54	104.00
4	D	-6	DC	O3'-P-O5'	-5.49	95.77	104.00

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	947	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	10430	0	10241	57	0
1	E	10436	0	10210	72	0
2	B	867	0	438	3	0
2	F	849	0	426	2	0
3	C	633	0	348	3	0
3	G	613	0	336	4	0
4	D	197	0	117	0	0
4	H	197	0	117	0	0
5	A	1	0	0	0	0
5	E	1	0	0	0	0
6	A	1	0	0	0	0
6	E	1	0	0	0	0
7	A	28	0	42	0	0
7	E	12	0	18	0	0
7	G	4	0	6	0	0
8	A	34	0	0	1	0
8	B	17	0	0	1	0
8	C	5	0	0	0	0
8	E	51	0	0	0	0
8	F	21	0	0	0	0
8	G	9	0	0	0	0
All	All	24407	0	22299	136	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.

The worst 5 of 136 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:1100:GLY:HA3	1:A:1122:ASN:OD1	1.90	0.72
1:A:587:SER:O	1:A:725:ARG:HD3	1.93	0.68

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:E:1280:HIS:ND1	1:E:1290:GLN:OE1	2.23	0.67
1:E:1133:MET:HE3	1:E:1307:ASN:CB	2.27	0.65
1:A:1120:ASN:O	1:A:1123:LEU:HD21	1.97	0.65

There are no symmetry-related clashes.

5.3　Torsion angles ⓘ

5.3.1　Protein backbone ⓘ

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	1290/1307 (99%)	1245 (96%)	42 (3%)	3 (0%)	44	68
1	E	1297/1307 (99%)	1254 (97%)	42 (3%)	1 (0%)	48	73
All	All	2587/2614 (99%)	2499 (97%)	84 (3%)	4 (0%)	44	68

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	802	LEU
1	E	1167	HIS
1	A	600	ASP
1	A	310	LEU

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	1124/1170 (96%)	1103 (98%)	21 (2%)	52	79
1	E	1120/1170 (96%)	1091 (97%)	29 (3%)	41	70
All	All	2244/2340 (96%)	2194 (98%)	50 (2%)	47	76

5 of 50 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	358	SER
1	E	695	LEU
1	E	1228	SER
1	E	380	ASP
1	E	526	TYR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 24 such sidechains are listed below:

Mol	Chain	Res	Type
1	E	611	GLN
1	E	880	GLN
1	E	724	GLN
1	E	1014	GLN
1	A	1041	ASN

5.3.3   RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
2	B	40/44 (90%)	4 (10%)	0
2	F	39/44 (88%)	3 (7%)	0
All	All	79/88 (89%)	7 (8%)	0

5 of 7 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
2	B	-16	U
2	B	-9	C
2	B	-6	G
2	B	20	A
2	F	-16	U

There are no RNA pucker outliers to report.

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)

Of 15 ligands modelled in this entry, 4 are monoatomic - leaving 11 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	$\#|Z| > 2$
7	EDO	A	1404	-	3,3,3	0.11	0	2,2,2	0.20	0
7	EDO	A	1409	-	3,3,3	0.14	0	2,2,2	0.33	0
7	EDO	A	1408	-	3,3,3	0.09	0	2,2,2	0.17	0
7	EDO	A	1407	-	3,3,3	0.10	0	2,2,2	0.10	0
7	EDO	E	1403	-	3,3,3	0.10	0	2,2,2	0.23	0
7	EDO	E	1405	-	3,3,3	0.18	0	2,2,2	0.27	0
7	EDO	A	1406	-	3,3,3	0.12	0	2,2,2	0.24	0
7	EDO	G	101	-	3,3,3	0.25	0	2,2,2	0.37	0
7	EDO	A	1403	-	3,3,3	0.08	0	2,2,2	0.14	0
7	EDO	E	1404	-	3,3,3	0.08	0	2,2,2	0.13	0
7	EDO	A	1405	-	3,3,3	0.09	0	2,2,2	0.13	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
7	EDO	A	1404	-	-	1/1/1/1	-
7	EDO	A	1409	-	-	0/1/1/1	-
7	EDO	A	1408	-	-	0/1/1/1	-
7	EDO	A	1407	-	-	1/1/1/1	-

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
7	EDO	E	1403	-	-	1/1/1/1	-
7	EDO	E	1405	-	-	1/1/1/1	-
7	EDO	A	1406	-	-	1/1/1/1	-
7	EDO	G	101	-	-	1/1/1/1	-
7	EDO	A	1403	-	-	0/1/1/1	-
7	EDO	E	1404	-	-	1/1/1/1	-
7	EDO	A	1405	-	-	0/1/1/1	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 7 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
7	E	1405	EDO	O1-C1-C2-O2
7	E	1403	EDO	O1-C1-C2-O2
7	E	1404	EDO	O1-C1-C2-O2
7	A	1407	EDO	O1-C1-C2-O2
7	G	101	EDO	O1-C1-C2-O2

There are no ring outliers.

No monomer is involved in short contacts.

5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

There are no chain breaks in this entry.

6 Fit of model and data ⓘ

6.1 Protein, DNA and RNA chains ⓘ

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	1292/1307 (98%)	0.01	28 (2%)	62	61	32, 67, 105, 153	4 (0%)
1	E	1295/1307 (99%)	-0.00	42 (3%)	50	48	30, 64, 110, 168	6 (0%)
2	B	41/44 (93%)	-0.89	1 (2%)	59	58	43, 52, 79, 145	0
2	F	41/44 (93%)	-0.96	0	100	100	38, 49, 81, 102	0
3	C	31/34 (91%)	-0.61	1 (3%)	50	48	48, 56, 124, 145	0
3	G	30/34 (88%)	-0.71	0	100	100	42, 53, 96, 128	0
4	D	10/10 (100%)	-0.35	0	100	100	56, 67, 119, 125	0
4	H	10/10 (100%)	-0.66	0	100	100	49, 60, 113, 121	0
All	All	2750/2790 (98%)	-0.04	72 (2%)	57	56	30, 65, 108, 168	10 (0%)

The worst 5 of 72 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	1007	ALA	7.0
1	A	1008	GLU	5.4
1	A	1164	ILE	5.1
1	A	149	VAL	4.5
1	E	1010	ALA	4.3

6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

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

6.3 Carbohydrates ⓘ

There are no oligosaccharides in this entry.

6.4 Ligands ⓘ

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($\mathring{A}^2$)	Q<0.9
7	EDO	E	1404	4/4	0.73	0.25	80,80,82,82	0
7	EDO	A	1406	4/4	0.84	0.17	71,75,76,77	0
7	EDO	E	1405	4/4	0.85	0.23	64,68,68,70	0
7	EDO	A	1403	4/4	0.88	0.17	81,81,81,81	0
7	EDO	E	1403	4/4	0.89	0.14	69,70,70,71	0
7	EDO	A	1408	4/4	0.89	0.18	80,80,80,82	0
7	EDO	A	1409	4/4	0.89	0.16	64,66,67,68	0
7	EDO	A	1407	4/4	0.90	0.18	78,81,82,83	0
7	EDO	G	101	4/4	0.90	0.16	59,63,65,65	0
7	EDO	A	1404	4/4	0.92	0.15	72,72,74,74	0
5	NA	E	1401	1/1	0.95	0.10	54,54,54,54	0
7	EDO	A	1405	4/4	0.96	0.09	63,64,64,64	0
6	CL	E	1402	1/1	0.98	0.06	77,77,77,77	0
6	CL	A	1402	1/1	0.98	0.04	66,66,66,66	0
5	NA	A	1401	1/1	0.99	0.03	51,51,51,51	0

6.5 Other polymers ⓘ

There are no such residues in this entry.