



wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 5, 2024 – 10:17 am BST

PDB ID : 8RD3
Title : Crystal structure of *Saccharomyces cerevisiae* Nmd4 protein bound to Upf1 helicase domain
Authors : Barbarin-Bocahu, I.; Graille, M.
Deposited on : 2023-12-07
Resolution : 2.40 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 : 4.02b-467
Mogul : 1.8.4, CSD as541be (2020)
Xtriage (Phenix) : 1.13
EDS : 2.37.1
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) : 2.37.1

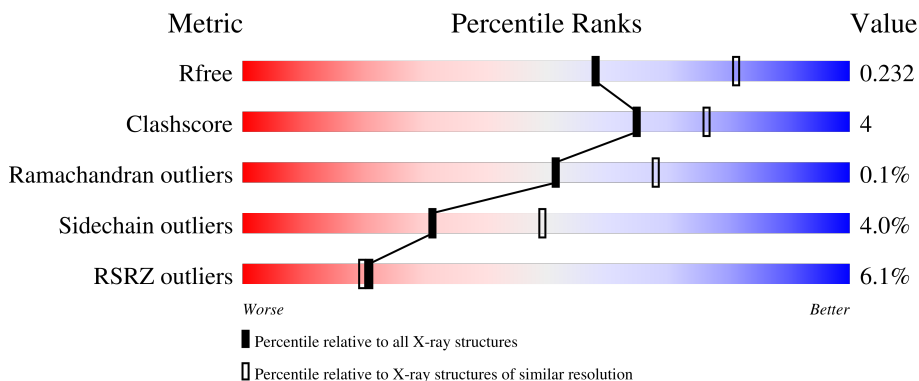
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.40 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	3907 (2.40-2.40)
Clashscore	141614	4398 (2.40-2.40)
Ramachandran outliers	138981	4318 (2.40-2.40)
Sidechain outliers	138945	4319 (2.40-2.40)
RSRZ outliers	127900	3811 (2.40-2.40)

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	640	 3% 85% 12%
2	B	221	 15% 89% 10%

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
3	EDO	A	903	-	-	X	-
5	FMT	A	907	-	-	X	-

2 Entry composition [i](#)

There are 7 unique types of molecules in this entry. The entry contains 6893 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 ATP-dependent helicase NAM7.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	628	4980	3174	871	916	19	0	0	0

There are 9 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	218	MET	-	initiating methionine	UNP P30771
A	219	ALA	-	expression tag	UNP P30771
A	220	SER	-	expression tag	UNP P30771
A	852	HIS	-	expression tag	UNP P30771
A	853	HIS	-	expression tag	UNP P30771
A	854	HIS	-	expression tag	UNP P30771
A	855	HIS	-	expression tag	UNP P30771
A	856	HIS	-	expression tag	UNP P30771
A	857	HIS	-	expression tag	UNP P30771

- Molecule 2 is a protein called Nonsense-mediated decay protein 4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	219	1779	1126	300	340	13	0	0	0

There are 4 discrepancies between the modelled and reference sequences:

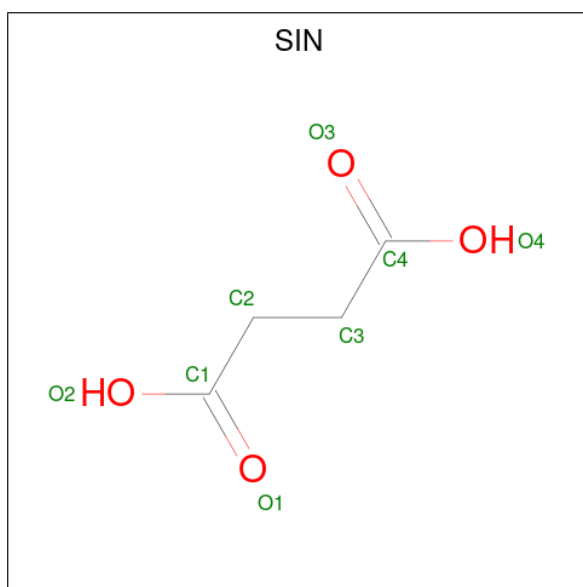
Chain	Residue	Modelled	Actual	Comment	Reference
B	-2	GLY	-	expression tag	UNP Q12129
B	-1	PRO	-	expression tag	UNP Q12129
B	0	GLY	-	expression tag	UNP Q12129
B	1	SER	-	expression tag	UNP Q12129

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



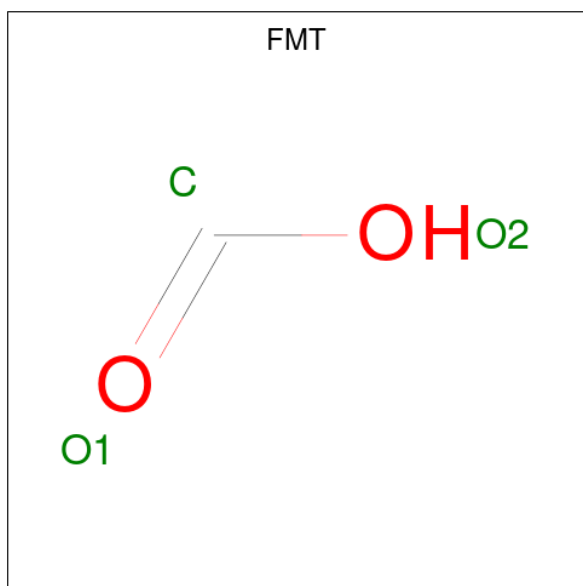
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C O 4 2 2	0	0
3	A	1	Total C O 4 2 2	0	0
3	A	1	Total C O 4 2 2	0	0
3	A	1	Total C O 4 2 2	0	0
3	A	1	Total C O 4 2 2	0	0
3	A	1	Total C O 4 2 2	0	0
3	A	1	Total C O 4 2 2	0	0
3	A	1	Total C O 4 2 2	0	0
3	B	1	Total C O 4 2 2	0	0

- Molecule 4 is SUCCINIC ACID (three-letter code: SIN) (formula: C₄H₆O₄).



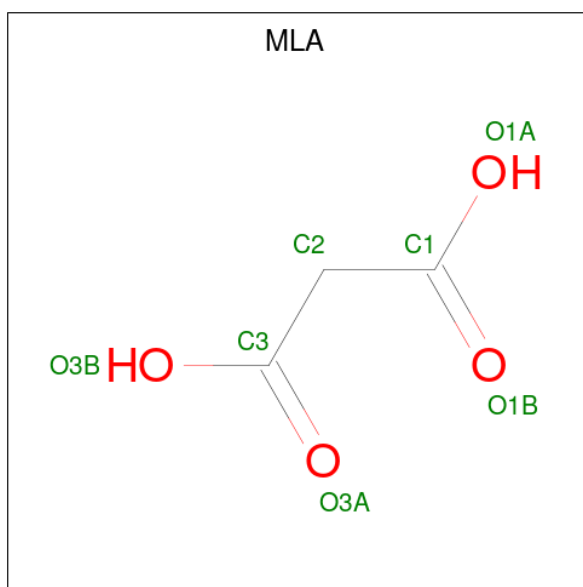
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			8	4	4		

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



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

- Molecule 6 is MALONIC ACID (three-letter code: MLA) (formula: $\text{C}_3\text{H}_4\text{O}_4$).



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

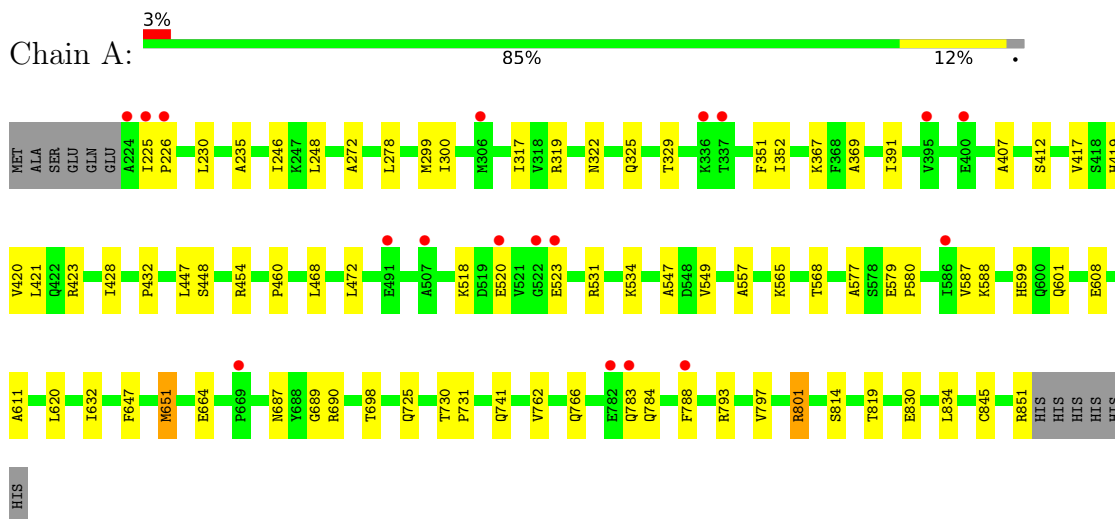
- Molecule 7 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	A	74	Total O 74 74	0	0
7	B	6	Total O 6 6	0	0

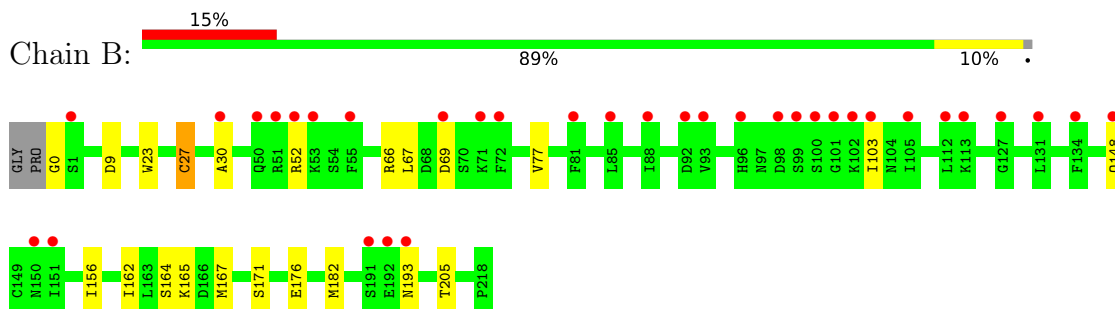
3 Residue-property plots [i](#)

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 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: ATP-dependent helicase NAM7



- Molecule 2: Nonsense-mediated decay protein 4



4 Data and refinement statistics

Property	Value	Source
Space group	I 2 2 2	Depositor
Cell constants a, b, c, α , β , γ	119.35Å 131.16Å 169.55Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.80 – 2.40 48.80 – 2.40	Depositor EDS
% Data completeness (in resolution range)	74.5 (48.80-2.40) 74.6 (48.80-2.40)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.27 (at 2.39Å)	Xtrriage
Refinement program	BUSTER 2.10.3	Depositor
R, R_{free}	0.202 , 0.222 0.208 , 0.232	Depositor DCC
R_{free} test set	1999 reflections (5.15%)	wwPDB-VP
Wilson B-factor (Å ²)	65.1	Xtrriage
Anisotropy	0.013	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.32 , 47.1	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	6893	wwPDB-VP
Average B, all atoms (Å ²)	76.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.53% 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: MLA, EDO, SIN, CME, FMT

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	0/5081	0.57	0/6878
2	B	0.38	0/1802	0.53	0/2422
All	All	0.41	0/6883	0.56	0/9300

There are no bond length outliers.

There are no bond angle outliers.

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	4980	0	5070	51	0
2	B	1779	0	1759	8	0
3	A	32	0	48	9	0
3	B	4	0	6	1	0
4	A	8	0	4	1	0
5	A	3	0	2	4	0
6	A	7	0	2	1	0
7	A	74	0	0	0	0
7	B	6	0	0	0	0
All	All	6893	0	6891	56	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 4.

The worst 5 of 56 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:647:PHE:HB3	3:A:901:EDO:H12	1.52	0.88
1:A:766:GLN:HG3	5:A:907:FMT:C	2.13	0.78
2:B:205:THR:HG23	3:B:301:EDO:H21	1.66	0.77
1:A:565:LYS:HE2	3:A:903:EDO:H11	1.70	0.73
2:B:103:ILE:H	2:B:148:GLN:HE22	1.39	0.70

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	626/640 (98%)	616 (98%)	9 (1%)	1 (0%)	47	62
2	B	216/221 (98%)	212 (98%)	4 (2%)	0	100	100
All	All	842/861 (98%)	828 (98%)	13 (2%)	1 (0%)	51	68

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	587	VAL

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	549/560 (98%)	530 (96%)	19 (4%)	36	55
2	B	199/200 (100%)	188 (94%)	11 (6%)	21	35
All	All	748/760 (98%)	718 (96%)	30 (4%)	31	49

5 of 30 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	788	PHE
2	B	171	SER
1	A	830	GLU
2	B	193	ASN
2	B	164	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	325	GLN
1	A	599	HIS
1	A	600	GLN
1	A	783	GLN
2	B	148	GLN

5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

1 non-standard protein/DNA/RNA residue is 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	CME	B	27	2	8,9,10	0.61	0	5,9,11	2.12	1 (20%)

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
2	CME	B	27	2	-	3/5/8/10	-

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	27	CME	CB-SG-SD	4.46	115.38	103.82

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	B	27	CME	CE-SD-SG-CB
2	B	27	CME	CA-CB-SG-SD
2	B	27	CME	N-CA-CB-SG

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	27	CME	1	0

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	B	301	-	3,3,3	0.49	0	2,2,2	0.34	0
4	SIN	A	902	-	7,7,7	0.55	0	8,8,8	0.59	0
3	EDO	A	903	-	3,3,3	0.85	0	2,2,2	0.09	0
3	EDO	A	911	-	3,3,3	0.61	0	2,2,2	0.27	0
3	EDO	A	909	-	3,3,3	0.98	0	2,2,2	0.09	0
6	MLA	A	910	-	6,6,6	0.45	0	7,7,7	0.74	0
3	EDO	A	906	-	3,3,3	0.54	0	2,2,2	0.31	0
3	EDO	A	905	-	3,3,3	0.78	0	2,2,2	0.25	0
3	EDO	A	901	-	3,3,3	1.12	0	2,2,2	0.95	0
3	EDO	A	904	-	3,3,3	0.44	0	2,2,2	0.41	0
3	EDO	A	908	-	3,3,3	0.69	0	2,2,2	0.08	0
5	FMT	A	907	-	2,2,2	1.75	1 (50%)	1,1,1	0.65	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	B	301	-	-	0/1/1/1	-
4	SIN	A	902	-	-	3/5/5/5	-
3	EDO	A	903	-	-	1/1/1/1	-
3	EDO	A	911	-	-	0/1/1/1	-
3	EDO	A	909	-	-	0/1/1/1	-
6	MLA	A	910	-	-	0/4/4/4	-
3	EDO	A	906	-	-	0/1/1/1	-
3	EDO	A	905	-	-	0/1/1/1	-
3	EDO	A	901	-	-	0/1/1/1	-
3	EDO	A	904	-	-	1/1/1/1	-
3	EDO	A	908	-	-	0/1/1/1	-

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	A	907	FMT	O2-C	2.39	1.40	1.28

There are no bond angle outliers.

There are no chirality outliers.

All (5) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	902	SIN	C1-C2-C3-C4
4	A	902	SIN	O1-C1-C2-C3
3	A	903	EDO	O1-C1-C2-O2
3	A	904	EDO	O1-C1-C2-O2
4	A	902	SIN	O2-C1-C2-C3

There are no ring outliers.

8 monomers are involved in 16 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	B	301	EDO	1	0
4	A	902	SIN	1	0
3	A	903	EDO	4	0
6	A	910	MLA	1	0
3	A	905	EDO	2	0
3	A	901	EDO	2	0
3	A	904	EDO	1	0
5	A	907	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 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	628/640 (98%)	0.12	18 (2%) 51 50	42, 63, 102, 131	0
2	B	218/221 (98%)	0.82	34 (15%) 2 1	55, 102, 149, 155	0
All	All	846/861 (98%)	0.30	52 (6%) 21 20	42, 69, 126, 155	0

The worst 5 of 52 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	193	ASN	5.4
1	A	782	GLU	4.9
2	B	192	GLU	4.6
2	B	55	PHE	4.5
2	B	191	SER	4.3

6.2 Non-standard residues in protein, DNA, RNA chains [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
2	CME	B	27	10/11	0.90	0.15	92,96,98,103	0

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	SIN	A	902	8/8	0.71	0.30	84,85,87,87	0
3	EDO	A	903	4/4	0.72	0.21	82,83,83,83	0
3	EDO	A	909	4/4	0.81	0.16	65,65,66,66	0
3	EDO	A	905	4/4	0.82	0.27	69,69,70,70	0
6	MLA	A	910	7/7	0.82	0.29	76,77,77,77	0
3	EDO	A	908	4/4	0.83	0.24	94,94,94,94	0
3	EDO	A	901	4/4	0.87	0.35	58,59,60,60	0
5	FMT	A	907	3/3	0.89	0.17	63,63,64,64	0
3	EDO	A	911	4/4	0.90	0.24	77,78,78,78	0
3	EDO	B	301	4/4	0.93	0.21	78,79,79,79	0
3	EDO	A	904	4/4	0.94	0.21	74,74,75,75	0
3	EDO	A	906	4/4	0.97	0.19	80,80,80,80	0

6.5 Other polymers [i](#)

There are no such residues in this entry.