



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 14, 2018 – 03:14 am GMT

PDB ID : 4LXT
Title : Crystal structure WlaRD, a sugar 3N-formyl transferase in the presence of dTdp-Qui3N and 5-N-Formyl-THF
Authors : Thoden, J.B.; Goneau, M.-F.; Gilbert, M.; Holden, H.M.
Deposited on : 2013-07-30
Resolution : 1.40 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 following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : 1.7.3 (157068), CSD as539be (2018)
Xtriage (Phenix) : 1.13
EDS : trunk31020
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)
Refmac : 5.8.0158
CCP4 : 7.0 (Gargrove)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : trunk31020

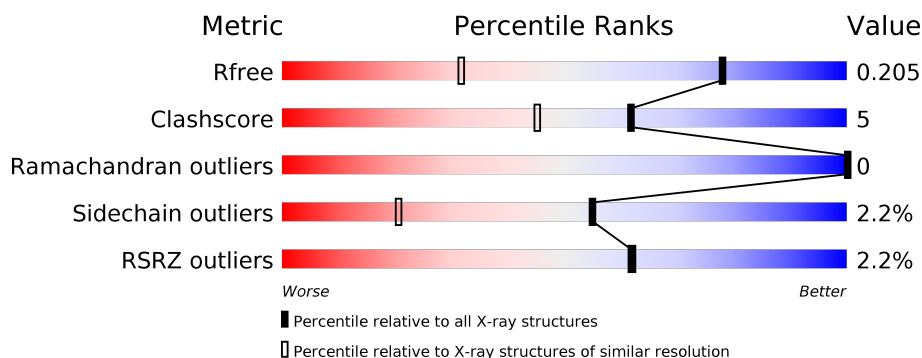
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.40 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}	111664	1450 (1.40-1.40)
Clashscore	122126	1541 (1.40-1.40)
Ramachandran outliers	120053	1500 (1.40-1.40)
Sidechain outliers	120020	1499 (1.40-1.40)
RSRZ outliers	108989	1412 (1.40-1.40)

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 on the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria. 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	274	<div> <div>%</div> <div> <div></div> <div>89%</div> <div>9%</div> </div> </div>
1	B	274	<div> <div>3%</div> <div> <div></div> <div>87%</div> <div>12%</div> </div> </div>

2 Entry composition [i](#)

There are 7 unique types of molecules in this entry. The entry contains 5801 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.

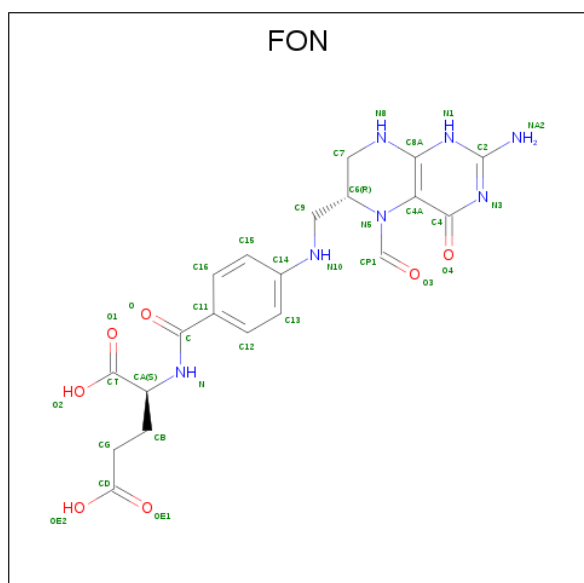
- Molecule 1 is a protein called WlaRD, a sugar 3N-formyl transferase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	273	Total	C	N	O	S	0	14	0
			2320	1521	375	421	3			
1	B	274	Total	C	N	O	S	0	10	0
			2313	1514	375	421	3			

There are 4 discrepancies between the modelled and reference sequences:

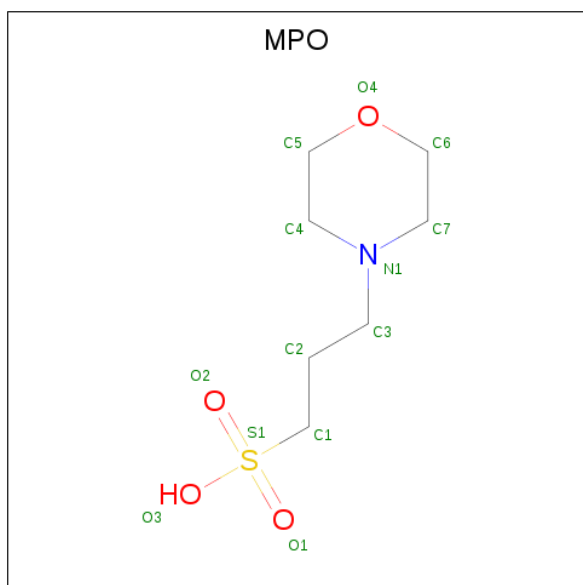
Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	GLY	-	EXPRESSION TAG	UNP A8FMJ3
A	0	HIS	-	EXPRESSION TAG	UNP A8FMJ3
B	-1	GLY	-	EXPRESSION TAG	UNP A8FMJ3
B	0	HIS	-	EXPRESSION TAG	UNP A8FMJ3

- Molecule 2 is N-{[4-({[(6R)-2-amino-5-formyl-4-oxo-1,4,5,6,7,8-hexahydropteridin-6-yl]methyl}amino)phenyl]carbonyl}-L-glutamic acid (three-letter code: FON) (formula: C₂₀H₂₃N₇O₇).



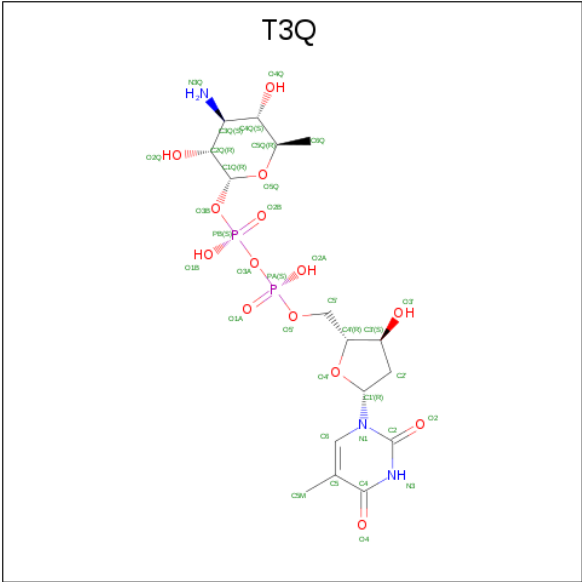
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	N	O	0	0
			34	20	7	7		
2	B	1	Total	C	N	O	0	0
			34	20	7	7		

- Molecule 3 is 3[N-MORPHOLINO]PROPANE SULFONIC ACID (three-letter code: MPO) (formula: C₇H₁₅NO₄S).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total	C	N	O	S	0	0
			13	7	1	4	1		
3	A	1	Total	C	N	O	S	0	0
			13	7	1	4	1		
3	A	1	Total	C	N	O	S	0	0
			13	7	1	4	1		
3	B	1	Total	C	N	O	S	0	0
			13	7	1	4	1		

- Molecule 4 is [(3R,4S,5S,6R)-4-amino-3,5-dihydroxy-6-methyloxan-2-yl][hydroxy-[(2R,3S,5R)-3-hydroxy-5-(5-methyl-2,4-dioxypyrimidin-1-yl)oxolan-2-yl]methoxy]phosphoryl] hydrogen phosphate (three-letter code: T3Q) (formula: C₁₆H₂₇N₃O₁₄P₂).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
4	A	1	Total 35	C 16	N 3	O 14	P 2	0	0
4	B	1	Total 35	C 16	N 3	O 14	P 2	0	0

- Molecule 5 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C₂H₆O₂).



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

Continued on next page...

Continued from previous page...

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

- Molecule 6 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

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

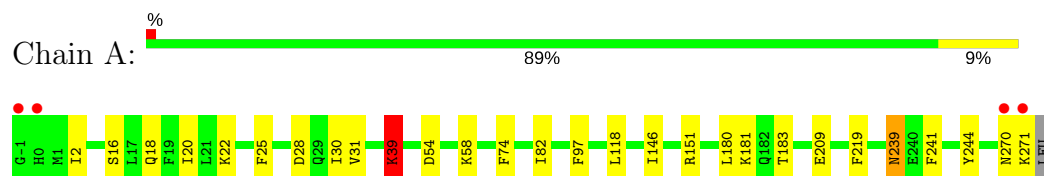
- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	526	Total	O	0	4
			530	530		
7	B	420	Total	O	0	2
			422	422		

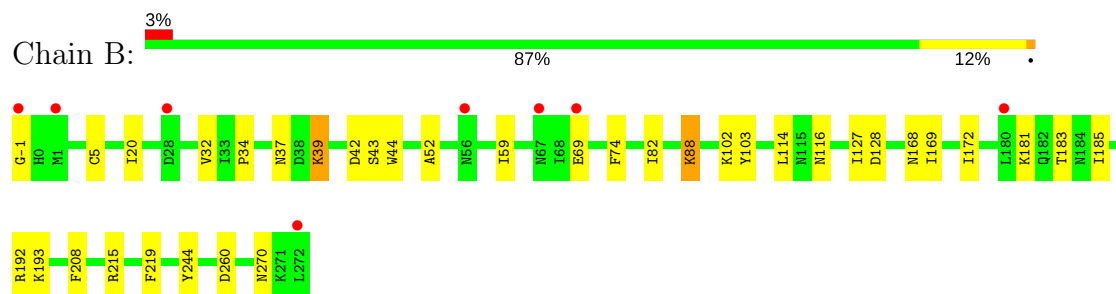
3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA 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: WlaRD, a sugar 3N-formyl transferase



- Molecule 1: WlaRD, a sugar 3N-formyl transferase



4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	96.78Å 65.15Å 134.72Å 90.00° 110.51° 90.00°	Depositor
Resolution (Å)	30.00 – 1.40 28.69 – 1.40	Depositor EDS
% Data completeness (in resolution range)	92.9 (30.00-1.40) 92.9 (28.69-1.40)	Depositor EDS
R_{merge}	0.04	Depositor
R_{sym}	0.04	Depositor
$\langle I/\sigma(I) \rangle$ ¹	10.26 (at 1.40Å)	Xtriage
Refinement program	REFMAC 5.5.0109	Depositor
R, R_{free}	0.177 , 0.206 0.177 , 0.205	Depositor DCC
R_{free} test set	7171 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å ²)	13.5	Xtriage
Anisotropy	0.020	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.37 , 52.9	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.013 for h,-k,-h-l	Xtriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	5801	wwPDB-VP
Average B, all atoms (Å ²)	18.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.66% 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: MPO, FON, T3Q, 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	0.72	0/2408	1.00	6/3246 (0.2%)
1	B	0.67	1/2392 (0.0%)	1.01	5/3226 (0.2%)
All	All	0.70	1/4800 (0.0%)	1.01	11/6472 (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	B	0	1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	5	CYS	CB-SG	-5.41	1.73	1.81

The worst 5 of 11 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	215	ARG	NE-CZ-NH1	7.43	124.02	120.30
1	B	260	ASP	CB-CG-OD2	-7.12	111.89	118.30
1	A	209	GLU	OE1-CD-OE2	-7.07	114.82	123.30
1	A	39	LYS	CB-CG-CD	-6.43	94.88	111.60
1	A	151	ARG	NE-CZ-NH2	5.59	123.10	120.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	-1	GLY	Peptide

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	2320	0	2395	21	0
1	B	2313	0	2375	21	1
2	A	34	0	21	2	0
2	B	34	0	21	2	0
3	A	39	0	45	8	0
3	B	13	0	15	0	0
4	A	35	0	25	0	0
4	B	35	0	25	0	0
5	A	12	0	18	1	0
5	B	12	0	18	1	0
6	A	1	0	0	0	0
6	B	1	0	0	0	0
7	A	530	0	0	9	1
7	B	422	0	0	4	2
All	All	5801	0	4958	47	3

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 5.

The worst 5 of 47 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)
7:A:691:HOH:O	1:B:193:LYS:HE3	1.68	0.91
1:B:32[A]:VAL:HG12	1:B:34:PRO:HD3	1.60	0.81
1:A:82:ILE:H	2:A:301:FON:HN1	1.31	0.79
1:B:82:ILE:H	2:B:301:FON:HN1	1.30	0.76
3:A:304:MPO:H71	7:B:499:HOH:O	1.87	0.74

All (3) 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)
7:B:703:HOH:O	7:B:749:HOH:O[2_556]	2.04	0.16
7:A:741:HOH:O	7:A:824:HOH:O[4_455]	2.07	0.13
1:B:43[B]:SER:OG	7:B:608:HOH:O[2_556]	2.17	0.03

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	285/274 (104%)	276 (97%)	9 (3%)	0	100	100
1	B	283/274 (103%)	273 (96%)	10 (4%)	0	100	100
All	All	568/548 (104%)	549 (97%)	19 (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	267/254 (105%)	261 (98%)	6 (2%)	55	20
1	B	265/254 (104%)	259 (98%)	6 (2%)	53	19
All	All	532/508 (105%)	520 (98%)	12 (2%)	55	19

5 of 12 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	239[B]	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	39	LYS
1	B	181	LYS
1	A	239[A]	ASN
1	B	88	LYS

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

Mol	Chain	Res	Type
1	A	80	GLN
1	A	94	ASN
1	B	94	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 carbohydrates in this entry.

5.6 Ligand geometry [i](#)

Of 16 ligands modelled in this entry, 2 are monoatomic - leaving 14 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
2	FON	A	301	-	25,36,36	1.06	1 (4%)	30,50,50	1.78	7 (23%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
3	MPO	A	302	-	13,13,13	1.82	1 (7%)	16,17,17	2.28	4 (25%)
3	MPO	A	303	-	13,13,13	2.05	1 (7%)	16,17,17	2.62	8 (50%)
3	MPO	A	304	-	13,13,13	2.05	1 (7%)	16,17,17	2.48	8 (50%)
4	T3Q	A	305	-	29,37,37	1.17	2 (6%)	41,57,57	2.01	7 (17%)
5	EDO	A	306	-	3,3,3	0.45	0	2,2,2	0.51	0
5	EDO	A	307	-	3,3,3	0.57	0	2,2,2	0.60	0
5	EDO	A	308	-	3,3,3	0.49	0	2,2,2	0.80	0
2	FON	B	301	-	25,36,36	1.08	1 (4%)	30,50,50	1.64	6 (20%)
3	MPO	B	302	-	13,13,13	2.09	1 (7%)	16,17,17	2.56	6 (37%)
4	T3Q	B	303	-	29,37,37	1.08	2 (6%)	41,57,57	1.59	8 (19%)
5	EDO	B	304	-	3,3,3	0.41	0	2,2,2	0.49	0
5	EDO	B	305	-	3,3,3	0.52	0	2,2,2	0.51	0
5	EDO	B	307	-	3,3,3	0.50	0	2,2,2	0.72	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
2	FON	A	301	-	-	0/18/37/37	0/2/3/3
3	MPO	A	302	-	-	0/7/15/15	0/1/1/1
3	MPO	A	303	-	-	0/7/15/15	0/1/1/1
3	MPO	A	304	-	-	0/7/15/15	0/1/1/1
4	T3Q	A	305	-	-	0/17/53/53	0/3/3/3
5	EDO	A	306	-	-	0/1/1/1	0/0/0/0
5	EDO	A	307	-	-	0/1/1/1	0/0/0/0
5	EDO	A	308	-	-	0/1/1/1	0/0/0/0
2	FON	B	301	-	-	0/18/37/37	0/2/3/3
3	MPO	B	302	-	-	0/7/15/15	0/1/1/1
4	T3Q	B	303	-	-	0/17/53/53	0/3/3/3
5	EDO	B	304	-	-	0/1/1/1	0/0/0/0
5	EDO	B	305	-	-	0/1/1/1	0/0/0/0
5	EDO	B	307	-	-	0/1/1/1	0/0/0/0

The worst 5 of 10 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	302	MPO	C1-S1	-7.13	1.67	1.77
3	A	303	MPO	C1-S1	-6.81	1.67	1.77
3	A	304	MPO	C1-S1	-6.78	1.67	1.77

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	302	MPO	C1-S1	-6.13	1.68	1.77
4	A	305	T3Q	C6-C5	-3.09	1.31	1.40

The worst 5 of 54 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	A	305	T3Q	O5Q-C5Q-C4Q	-4.22	102.76	109.62
4	B	303	T3Q	O4Q-C4Q-C3Q	-3.94	103.14	110.22
3	A	303	MPO	O3-S1-O1	-3.86	101.83	111.27
4	A	305	T3Q	O4Q-C4Q-C3Q	-3.69	103.59	110.22
3	A	304	MPO	O3-S1-O1	-3.32	103.16	111.27

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

6 monomers are involved in 14 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	301	FON	2	0
3	A	303	MPO	4	0
3	A	304	MPO	4	0
5	A	307	EDO	1	0
2	B	301	FON	2	0
5	B	305	EDO	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	273/274 (99%)	-0.40	4 (1%) 73 74	9, 13, 24, 45	0
1	B	274/274 (100%)	-0.09	8 (2%) 51 51	9, 16, 32, 51	0
All	All	547/548 (99%)	-0.24	12 (2%) 62 62	9, 14, 29, 51	0

The worst 5 of 12 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	-1	GLY	6.0
1	B	-1	GLY	5.9
1	A	271	LYS	3.1
1	A	270	ASN	3.0
1	B	56[A]	ASN	2.7

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 carbohydrates 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(\AA^2)	Q<0.9
3	MPO	A	303	13/13	0.86	0.20	20,40,56,57	0
3	MPO	A	304	13/13	0.87	0.24	22,43,64,65	0
5	EDO	B	305	4/4	0.93	0.14	26,31,35,36	0
2	FON	B	301	34/34	0.94	0.08	14,16,26,32	0
2	FON	A	301	34/34	0.95	0.08	9,13,20,29	0
5	EDO	A	308	4/4	0.96	0.07	19,21,30,32	0
3	MPO	A	302	13/13	0.96	0.12	19,22,26,27	0
3	MPO	B	302	13/13	0.97	0.10	19,23,29,31	0
5	EDO	A	306	4/4	0.97	0.07	20,20,21,26	0
5	EDO	B	304	4/4	0.98	0.05	16,21,21,25	0
5	EDO	A	307	4/4	0.98	0.08	16,19,21,28	0
4	T3Q	A	305	35/35	0.98	0.06	8,11,14,17	0
5	EDO	B	307	4/4	0.98	0.06	16,22,31,31	0
4	T3Q	B	303	35/35	0.98	0.05	10,12,17,20	0
6	CL	A	309	1/1	0.99	0.03	16,16,16,16	0
6	CL	B	306	1/1	0.99	0.04	21,21,21,21	0

6.5 Other polymers [i](#)

There are no such residues in this entry.