wwPDB X-ray Structure Validation Summary Report (i)

May 15, 2020 – 06:20 am BST

PDB ID	:	1F07
Title	:	STRUCTURE OF COENZYME F420 DEPENDENT TETRAHYDROME THANOPTERIN REDUCTASE FROM METHANOBACTERIUM THERMOAUTOTROPHICUM
Authors	:	Shima, S.; Warkentin, E.; Grabarse, W.; Thauer, R.K.; Ermler, U.
Deposited on	:	2000-05-15
Resolution	:	2.00 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 (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

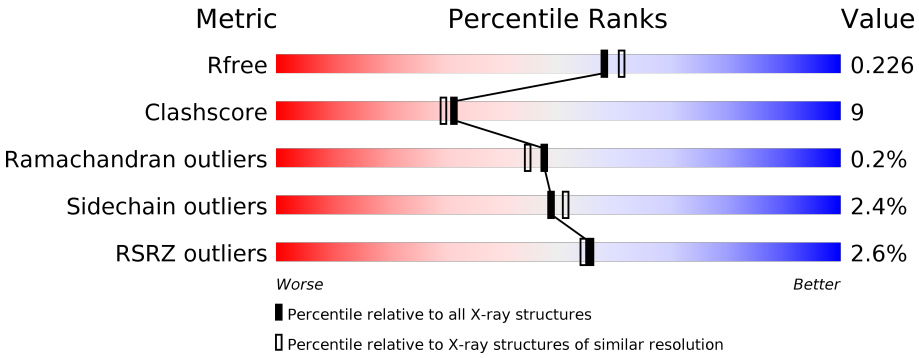
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.00 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	8085 (2.00-2.00)
Clashscore	141614	9178 (2.00-2.00)
Ramachandran outliers	138981	9054 (2.00-2.00)
Sidechain outliers	138945	9053 (2.00-2.00)
RSRZ outliers	127900	7900 (2.00-2.00)

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 on 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	321	2% 84% 15%
1	B	321	5% 80% 18%
1	C	321	% 84% 15%
1	D	321	3% 78% 21%

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
2	MPO	A	331	-	-	X	-
3	MPD	A	335	X	-	-	-
3	MPD	B	336	X	-	-	-
3	MPD	C	337	X	-	-	-
3	MPD	D	338	X	-	-	-

2 Entry composition

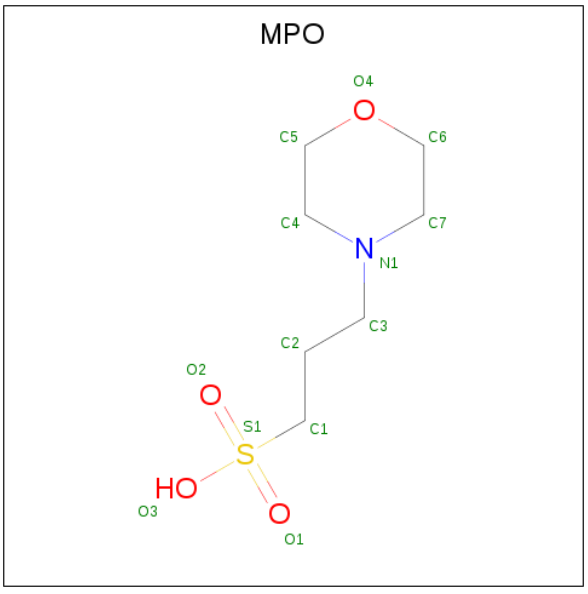
There are 5 unique types of molecules in this entry. The entry contains 11407 atoms, of which 1404 are hydrogens and 0 are deuteriums.

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 COENZYME F420-DEPENDENT N5,N10-METHYLENETETRAHYDROMETHANOPTERIN REDUCTASE.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	321	Total	C	N	O	S		0	0	0
			2353	1503	378	461	11				
1	B	321	Total	C	H	N	O	S	0	0	0
			2821	1503	468	378	461	11			
1	C	321	Total	C	H	N	O	S	0	0	0
			2821	1503	468	378	461	11			
1	D	321	Total	C	H	N	O	S	0	0	0
			2821	1503	468	378	461	11			

- Molecule 2 is 3[N-MORPHOLINO]PROPANE SULFONIC ACID (three-letter code: MPO) (formula: $C_7H_{15}NO_4S$).



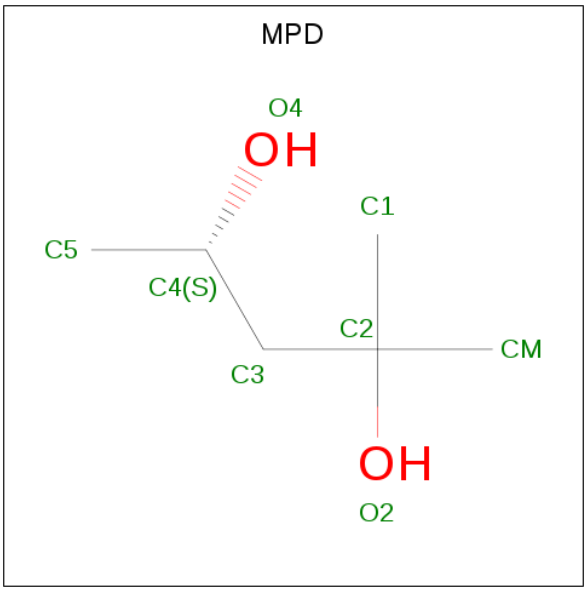
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	N	O	S	0	0
			13	7	1	4	1		

Continued on next page...

Continued from previous page...

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

- Molecule 3 is (4S)-2-METHYL-2,4-PENTANEDIOL (three-letter code: MPD) (formula: $C_6H_{14}O_2$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C O 8 6 2	0	0
3	B	1	Total C O 8 6 2	0	0
3	C	1	Total C O 8 6 2	0	0
3	D	1	Total C O 8 6 2	0	0

- Molecule 4 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	B	1	Total Cl 1 1	0	0
4	C	1	Total Cl 1 1	0	0

- Molecule 5 is water.

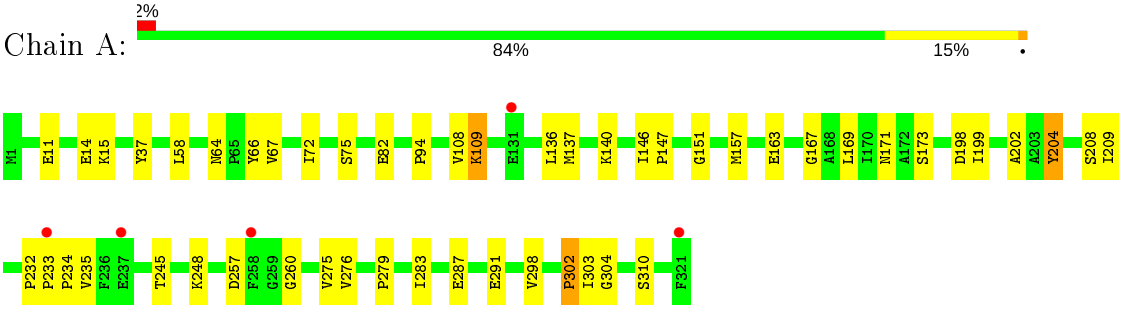
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	143	Total O 143 143	0	0
5	B	97	Total O 97 97	0	0
5	C	146	Total O 146 146	0	0
5	D	119	Total O 119 119	0	0

3 Residue-property plots

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.

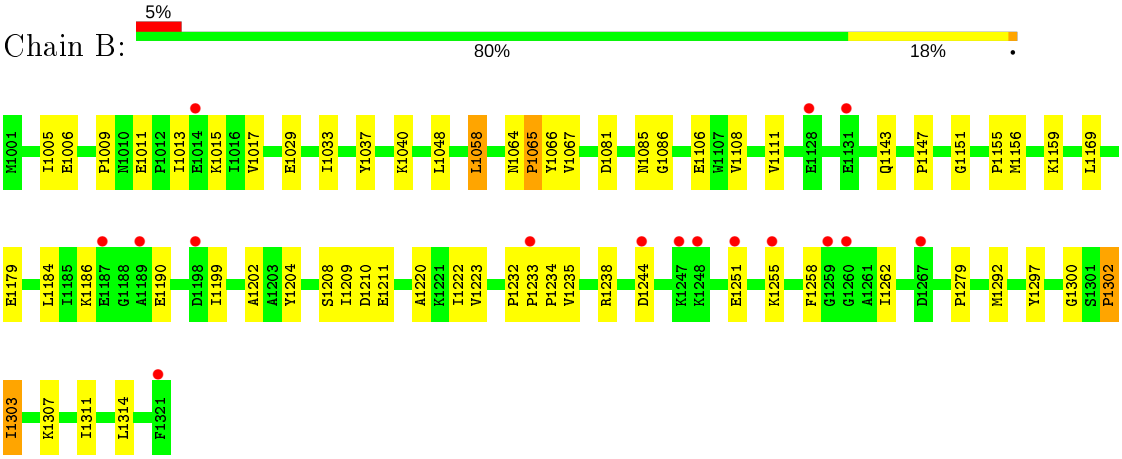
• Molecule 1: COENZYME F420-DEPENDENT N5,N10-METHYLENETETRAHYDROMETH ANOPTERIN REDUCTASE

Chain A:



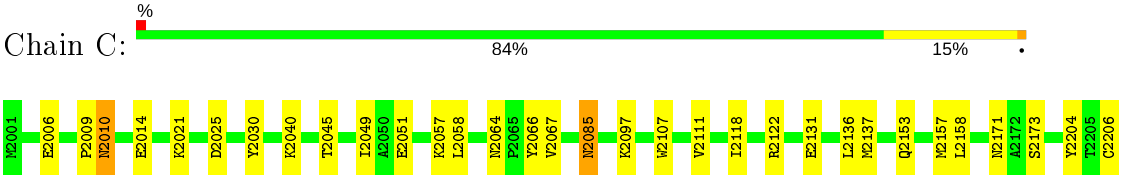
• Molecule 1: COENZYME F420-DEPENDENT N5,N10-METHYLENETETRAHYDROMETH ANOPTERIN REDUCTASE

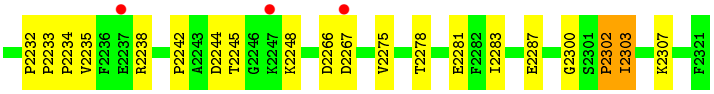
Chain B:



• Molecule 1: COENZYME F420-DEPENDENT N5,N10-METHYLENETETRAHYDROMETH ANOPTERIN REDUCTASE

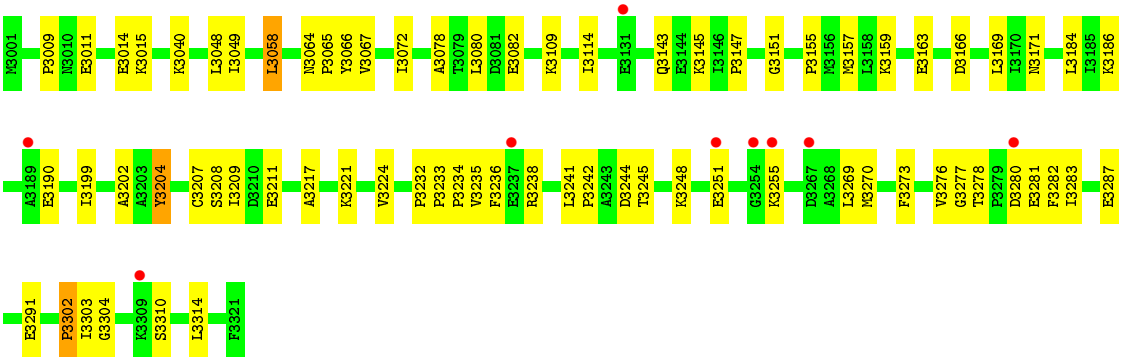
Chain C:





• Molecule 1: COENZYME F420-DEPENDENT N5,N10-METHYLENETETRAHYDROMETHANOPTERIN REDUCTASE

Chain D:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	162.80Å 128.20Å 109.20Å 90.00° 131.40° 90.00°	Depositor
Resolution (Å)	30.00 – 2.00 9.99 – 2.00	Depositor EDS
% Data completeness (in resolution range)	95.0 (30.00-2.00) 95.7 (9.99-2.00)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.05 (at 2.01Å)	Xtriage
Refinement program	CNS 0.3	Depositor
R, R_{free}	0.208 , 0.237 0.198 , 0.226	Depositor DCC
R_{free} test set	5402 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å^2)	30.6	Xtriage
Anisotropy	0.075	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.46 , 62.5	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	0.002 for k+l,h+l,-l 0.000 for -k+l,-h-l,-l 0.013 for -h-2*l,-k,l	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	11407	wwPDB-VP
Average B, all atoms (Å^2)	35.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.86% 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: MPO, MPD, 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	0.39	0/2396	0.67	1/3249 (0.0%)
1	B	0.35	0/2396	0.66	1/3249 (0.0%)
1	C	0.39	0/2396	0.66	1/3249 (0.0%)
1	D	0.37	0/2396	0.67	1/3249 (0.0%)
All	All	0.38	0/9584	0.66	4/12996 (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	B	0	1

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	1151	GLY	N-CA-C	-5.62	99.05	113.10
1	A	151	GLY	N-CA-C	-5.51	99.33	113.10
1	C	2206	CYS	N-CA-C	-5.23	96.88	111.00
1	D	3151	GLY	N-CA-C	-5.13	100.27	113.10

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	1297	TYR	Sidechain

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	A	2353	0	2378	43	1
1	B	2353	468	2375	44	0
1	C	2353	468	2375	39	1
1	D	2353	468	2375	51	0
2	A	13	0	14	6	0
2	B	13	0	14	0	0
2	C	13	0	14	0	0
2	D	13	0	14	0	0
3	A	8	0	14	0	0
3	B	8	0	14	1	0
3	C	8	0	14	2	0
3	D	8	0	14	1	0
4	B	1	0	0	0	0
4	C	1	0	0	0	0
5	A	143	0	0	1	0
5	B	97	0	0	1	0
5	C	146	0	0	4	0
5	D	119	0	0	2	0
All	All	10003	1404	9615	172	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 9.

The worst 5 of 172 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:64:ASN:HD22	1:A:66:TYR:H	1.06	0.99
1:B:1106:GLU:HG2	1:B:1108:VAL:HG13	1.47	0.95
1:B:1064:ASN:HD22	1:B:1066:TYR:H	1.19	0.89
1:D:3064:ASN:HD22	1:D:3066:TYR:H	1.25	0.85
1:C:2244:ASP:OD2	1:C:2248:LYS:HE2	1.76	0.84

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:A:163:GLU:OE1	1:C:2085:ASN:HD21[4_546]	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	319/321 (99%)	313 (98%)	6 (2%)	0	100	100
1	B	319/321 (99%)	313 (98%)	5 (2%)	1 (0%)	41	37
1	C	319/321 (99%)	312 (98%)	6 (2%)	1 (0%)	41	37
1	D	319/321 (99%)	310 (97%)	8 (2%)	1 (0%)	41	37
All	All	1276/1284 (99%)	1248 (98%)	25 (2%)	3 (0%)	47	44

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	1303	ILE
1	D	3109	LYS
1	C	2303	ILE

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	241/241 (100%)	236 (98%)	5 (2%)	53	57
1	B	241/241 (100%)	235 (98%)	6 (2%)	47	49

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	C	241/241 (100%)	235 (98%)	6 (2%)	47 49
1	D	241/241 (100%)	235 (98%)	6 (2%)	47 49
All	All	964/964 (100%)	941 (98%)	23 (2%)	49 51

5 of 23 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	1302	PRO
1	C	2014	GLU
1	D	3280	ASP
1	C	2010	ASN
1	C	2085	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 20 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	2010	ASN
1	C	2036	HIS
1	D	3064	ASN
1	B	1143	GLN
1	B	1218	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 10 ligands modelled in this entry, 2 are monoatomic - leaving 8 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	MPO	A	331	-	13,13,13	0.85	1 (7%)	17,17,17	1.06	1 (5%)
3	MPD	A	335	-	7,7,7	1.20	1 (14%)	9,10,10	3.07	4 (44%)
2	MPO	D	334	-	13,13,13	0.96	1 (7%)	17,17,17	0.84	1 (5%)
3	MPD	B	336	-	7,7,7	1.22	1 (14%)	9,10,10	3.17	4 (44%)
3	MPD	D	338	-	7,7,7	1.08	1 (14%)	9,10,10	3.16	4 (44%)
2	MPO	C	333	-	13,13,13	0.96	0	17,17,17	0.87	0
3	MPD	C	337	-	7,7,7	1.25	1 (14%)	9,10,10	3.04	4 (44%)
2	MPO	B	332	-	13,13,13	0.90	1 (7%)	17,17,17	1.01	2 (11%)

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	MPO	A	331	-	-	3/7/15/15	0/1/1/1
3	MPD	A	335	-	1/1/2/2	2/5/5/5	-
2	MPO	D	334	-	-	3/7/15/15	0/1/1/1
3	MPD	B	336	-	1/1/2/2	2/5/5/5	-
3	MPD	C	337	-	1/1/2/2	2/5/5/5	-
2	MPO	C	333	-	-	1/7/15/15	0/1/1/1
3	MPD	D	338	-	1/1/2/2	1/5/5/5	-
2	MPO	B	332	-	-	1/7/15/15	0/1/1/1

The worst 5 of 7 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	C	337	MPD	O2-C2	2.68	1.51	1.44
3	B	336	MPD	O2-C2	2.67	1.51	1.44
3	A	335	MPD	O2-C2	2.46	1.50	1.44
3	D	338	MPD	O2-C2	2.14	1.50	1.44
2	D	334	MPO	C7-N1	2.13	1.52	1.46

The worst 5 of 20 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	B	336	MPD	CM-C2-C1	8.23	127.72	110.57
3	D	338	MPD	CM-C2-C1	8.22	127.69	110.57
3	A	335	MPD	CM-C2-C1	8.04	127.32	110.57
3	C	337	MPD	CM-C2-C1	8.01	127.26	110.57
3	D	338	MPD	O2-C2-CM	-3.06	98.28	108.08

All (4) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
3	A	335	MPD	C4
3	B	336	MPD	C4
3	D	338	MPD	C4
3	C	337	MPD	C4

5 of 15 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	331	MPO	C2-C1-S1-O1
2	A	331	MPO	C2-C1-S1-O2
2	D	334	MPO	C2-C3-N1-C7
2	D	334	MPO	C2-C3-N1-C4
2	B	332	MPO	C2-C3-N1-C7

There are no ring outliers.

4 monomers are involved in 10 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	331	MPO	6	0
3	B	336	MPD	1	0
3	D	338	MPD	1	0
3	C	337	MPD	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

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, 95^{th} 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	321/321 (100%)	-0.23	5 (1%) 72 70	17, 31, 51, 61	0
1	B	321/321 (100%)	0.13	16 (4%) 28 28	18, 38, 64, 73	0
1	C	321/321 (100%)	-0.33	3 (0%) 84 83	18, 30, 46, 53	0
1	D	321/321 (100%)	-0.11	9 (2%) 53 51	20, 34, 54, 66	0
All	All	1284/1284 (100%)	-0.13	33 (2%) 56 54	17, 33, 55, 73	0

The worst 5 of 33 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	1244	ASP	5.1
1	B	1247	LYS	5.0
1	D	3251	GLU	4.1
1	A	233	PRO	3.5
1	B	1255	LYS	3.0

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 carbohydrates 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, 95^{th} 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
3	MPD	D	338	8/8	0.81	0.14	68,70,71,72	0
2	MPO	A	331	13/13	0.83	0.19	38,40,43,44	13
3	MPD	B	336	8/8	0.90	0.12	48,50,50,51	0
2	MPO	B	332	13/13	0.92	0.12	33,34,35,36	13
2	MPO	D	334	13/13	0.93	0.11	35,38,39,39	13
2	MPO	C	333	13/13	0.94	0.10	32,34,38,39	13
3	MPD	C	337	8/8	0.95	0.07	37,39,39,39	0
3	MPD	A	335	8/8	0.95	0.08	40,41,41,41	0
4	CL	B	324	1/1	0.97	0.05	30,30,30,30	0
4	CL	C	323	1/1	0.99	0.06	30,30,30,30	0

6.5 Other polymers (i)

There are no such residues in this entry.