



# wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 20, 2017 – 05:07 PM EDT

PDB ID : 3M2U  
Title : Structural Insight into Methyl-Coenzyme M Reductase Chemistry using Coenzyme B Analogues  
Authors : Cedervall, P.E.; Dey, M.; Ragsdale, S.W.; Wilmot, C.M.  
Deposited on : 2010-03-08  
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](mailto:validation@mail.wwpdb.org)

A user guide is available at

<http://wwpdb.org/validation/2016/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 : 1.7.2 (RC1), CSD as538be (2017)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20029077  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : rb-20029077

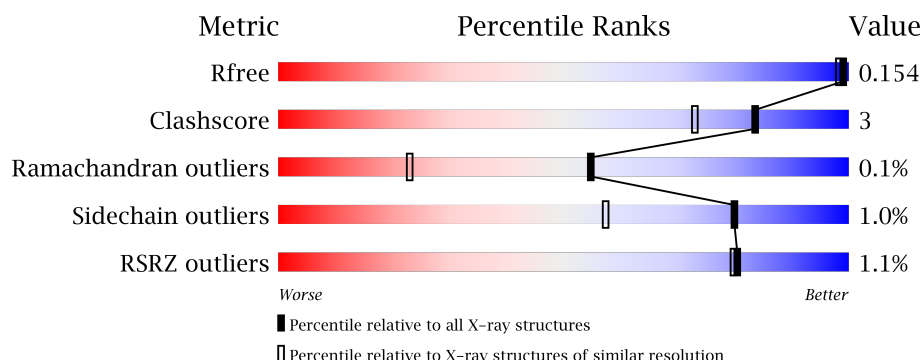
# 1 Overall quality at a glance

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))
$R_{free}$	100719	1307 (1.40-1.40)
Clashscore	112137	1411 (1.40-1.40)
Ramachandran outliers	110173	1373 (1.40-1.40)
Sidechain outliers	110143	1372 (1.40-1.40)
RSRZ outliers	101464	1315 (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  $\geq 3$ , 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	549	 85% 13% .
1	D	549	 85% 13% .
2	B	442	 86% 14% .
2	E	442	 83% 16% .
3	C	248	 81% 17% ..

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
3	F	248	

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
4	MG	A	552[B]	-	-	-	X
4	MG	C	250	-	-	-	X
4	MG	F	250	-	-	-	X
7	COM	A	554[A]	-	-	-	X
7	COM	D	555[A]	-	-	-	X
9	ACT	C	1	-	-	-	X

## 2 Entry composition [i](#)

There are 13 unique types of molecules in this entry. The entry contains 22594 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 Methyl-coenzyme M reductase I subunit alpha.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	548	Total	C	N	O	S	0	31	0
			4446	2817	734	874	21			
1	D	548	Total	C	N	O	S	0	28	0
			4410	2799	724	867	20			

- Molecule 2 is a protein called Methyl-coenzyme M reductase I subunit beta.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	442	Total	C	N	O	S	0	24	0
			3464	2212	563	666	23			
2	E	442	Total	C	N	O	S	0	34	0
			3527	2250	571	683	23			

- Molecule 3 is a protein called Methyl-coenzyme M reductase I subunit gamma.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	246	Total	C	N	O	S	0	14	0
			2072	1287	359	412	14			
3	F	246	Total	C	N	O	S	0	10	0
			2041	1268	356	404	13			

- Molecule 4 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

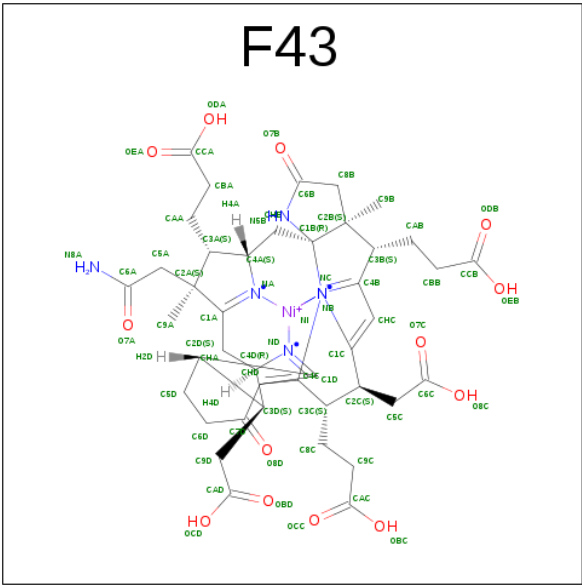
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	D	3	Total	Mg	0	0
			3	3		
4	E	2	Total	Mg	0	1
			2	2		
4	B	2	Total	Mg	0	1
			2	2		
4	C	1	Total	Mg	0	0
			1	1		

*Continued on next page...*

Continued from previous page...

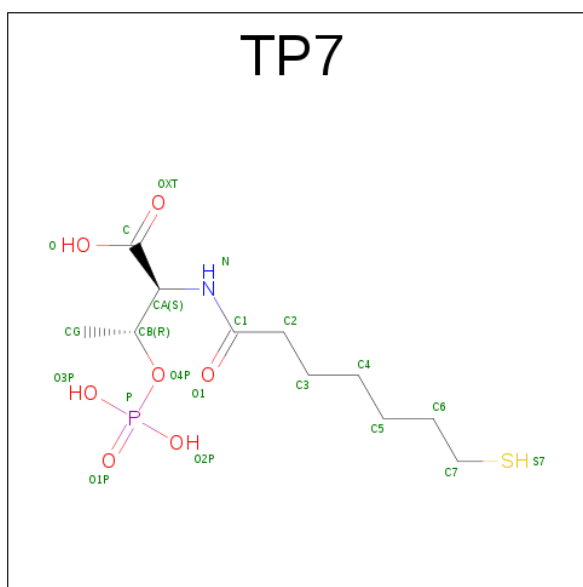
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	2	Total	Mg	0	2
			2	2		
4	F	1	Total	Mg	0	0
			1	1		

- Molecule 5 is FACTOR 430 (three-letter code: F43) (formula: C<sub>42</sub>H<sub>51</sub>N<sub>6</sub>NiO<sub>13</sub>).



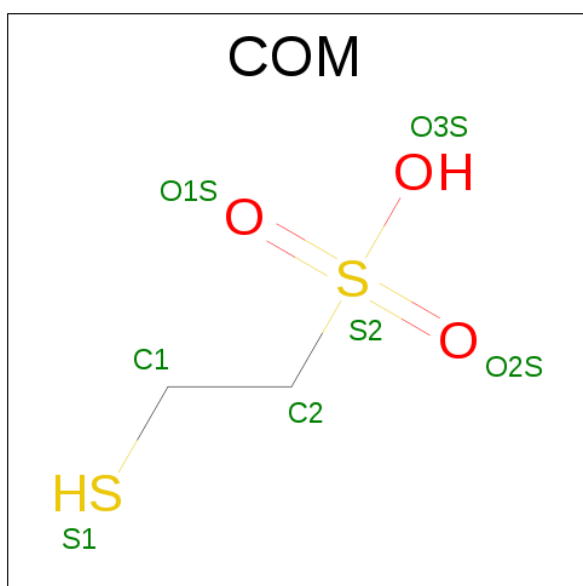
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
5	A	1	Total	C	N	Ni	O	0	0
			62	42	6	1	13		
5	D	1	Total	C	N	Ni	O	0	0
			62	42	6	1	13		

- Molecule 6 is COENZYME B (three-letter code: TP7) (formula: C<sub>11</sub>H<sub>22</sub>NO<sub>7</sub>PS).



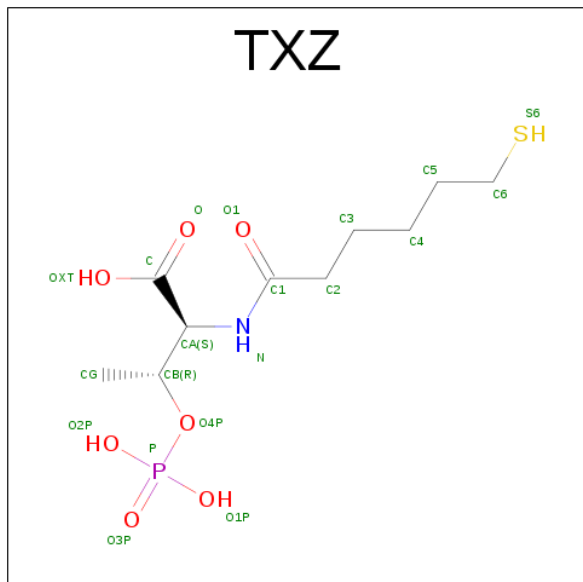
Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
6	A	1	Total	C	N	O	P	S	0	1
			21	11	1	7	1	1		
6	D	1	Total	C	N	O	P	S	0	1
			21	11	1	7	1	1		

- Molecule 7 is 1-THIOETHANESULFONIC ACID (three-letter code: COM) (formula:  $C_2H_6O_3S_2$ ).



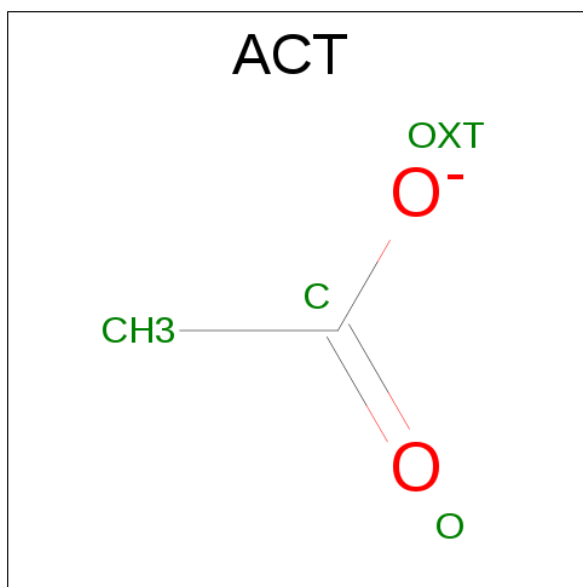
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
7	A	1	Total	C	O	S	0	1
			7	2	3	2		
7	D	1	Total	C	O	S	0	1
			7	2	3	2		

- Molecule 8 is O-PHOSPHONO-N-(6-SULFANYLHEXANOYL)-L-THREONINE (three-letter code: TXZ) (formula:  $C_{10}H_{20}NO_7PS$ ).



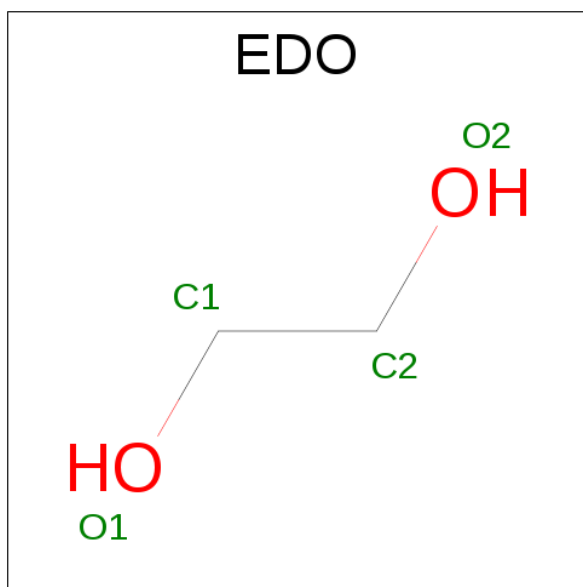
Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
8	A	1	Total	C	N	O	P	S	0	1
			20	10	1	7	1	1		
8	D	1	Total	C	N	O	P	S	0	1
			20	10	1	7	1	1		

- Molecule 9 is ACETATE ION (three-letter code: ACT) (formula:  $C_2H_3O_2$ ).



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

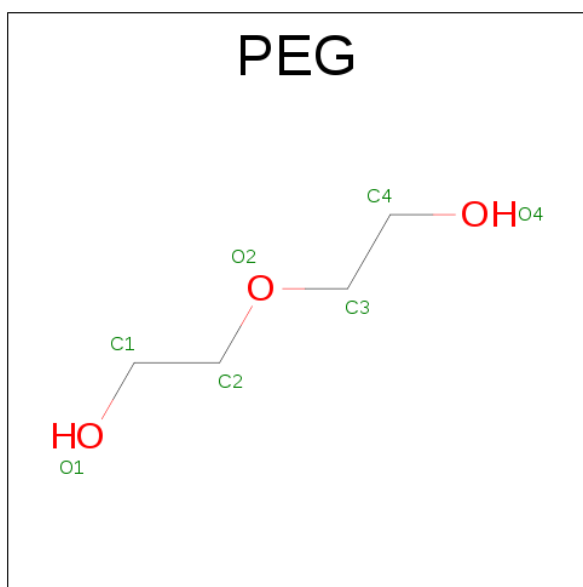
- Molecule 10 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C<sub>2</sub>H<sub>6</sub>O<sub>2</sub>).



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

- Molecule 11 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula: C<sub>4</sub>H<sub>10</sub>O<sub>3</sub>).





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

- Molecule 12 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
12	A	1	Total	Zn	0	0
			1	1		

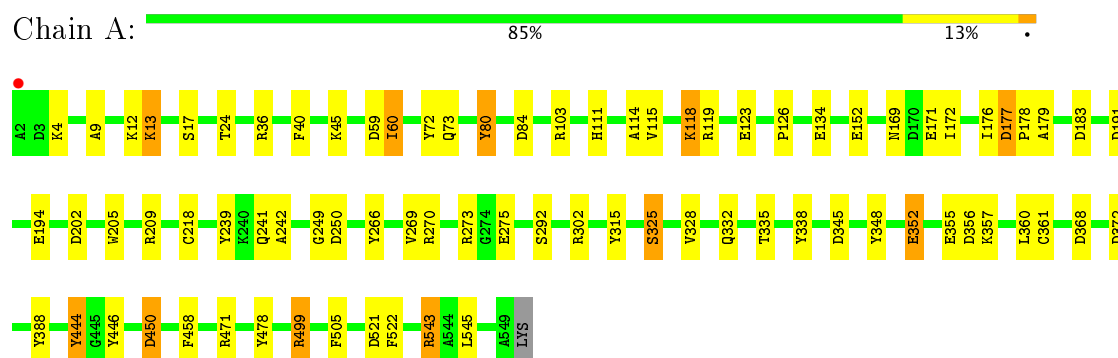
- Molecule 13 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
13	A	482	Total	O	0	24
			491	491		
13	B	436	Total	O	0	26
			454	454		
13	C	245	Total	O	0	15
			255	255		
13	D	488	Total	O	0	23
			502	502		
13	E	393	Total	O	0	11
			400	400		
13	F	246	Total	O	0	10
			250	250		

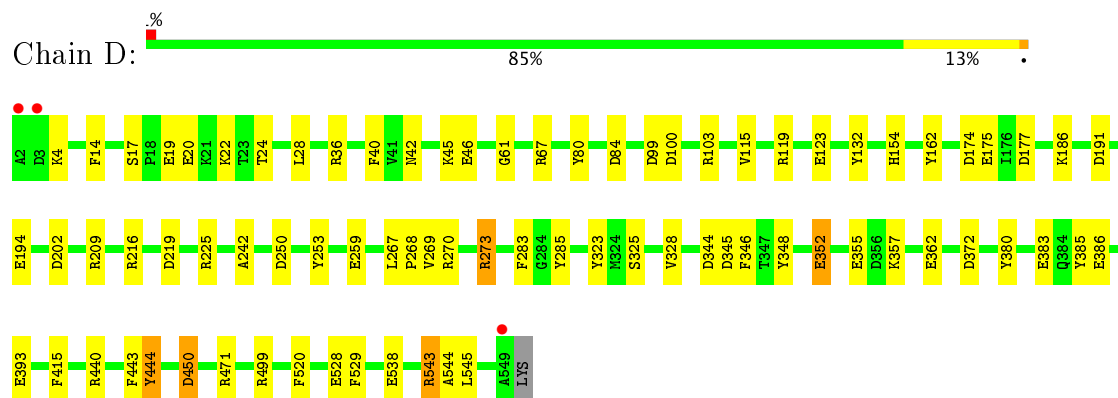
### 3 Residue-property plots [i](#)

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.

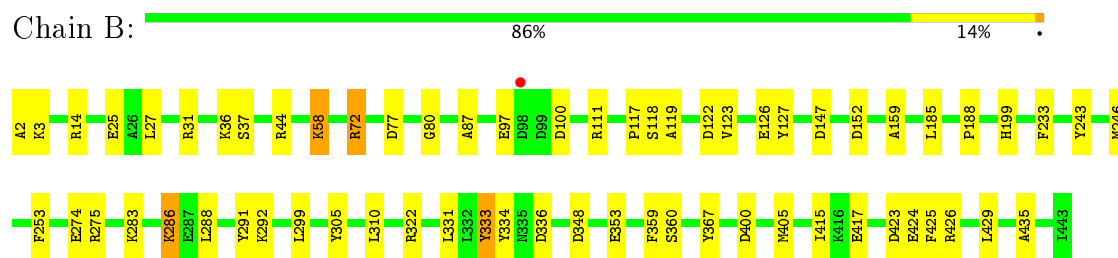
- Molecule 1: Methyl-coenzyme M reductase I subunit alpha



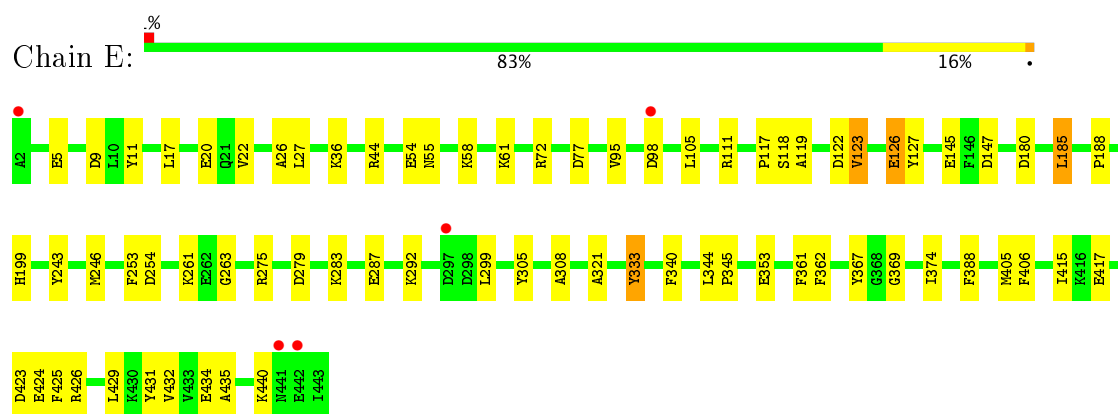
- Molecule 1: Methyl-coenzyme M reductase I subunit alpha



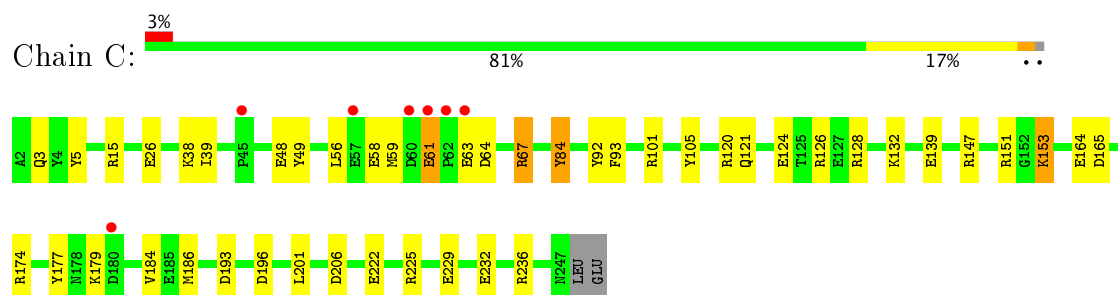
- Molecule 2: Methyl-coenzyme M reductase I subunit beta



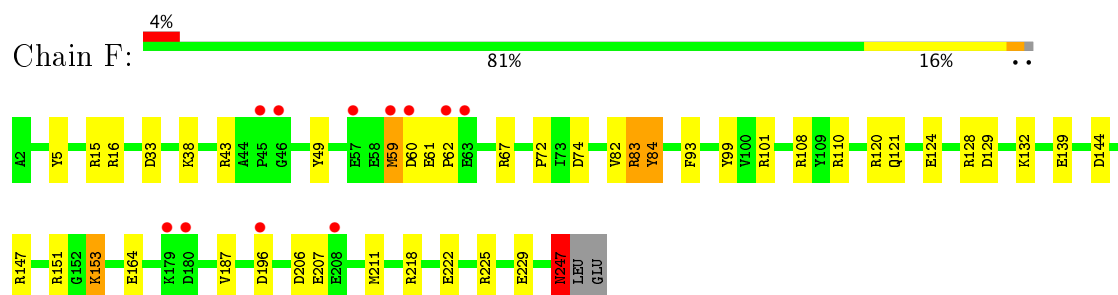
- Molecule 2: Methyl-coenzyme M reductase I subunit beta



• Molecule 3: Methyl-coenzyme M reductase I subunit gamma



• Molecule 3: Methyl-coenzyme M reductase I subunit gamma



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	82.02Å 118.26Å 122.39Å 90.00° 91.84° 90.00°	Depositor
Resolution (Å)	19.89 – 1.40 19.89 – 1.40	Depositor EDS
% Data completeness (in resolution range)	98.5 (19.89-1.40) 98.6 (19.89-1.40)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	5.77 (at 1.40Å)	Xtriage
Refinement program	REFMAC 5.5.0072	Depositor
R, $R_{free}$	0.130 , 0.155 0.129 , 0.154	Depositor DCC
$R_{free}$ test set	22348 reflections (5.27%)	DCC
Wilson B-factor (Å <sup>2</sup> )	11.7	Xtriage
Anisotropy	0.067	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.38 , 51.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.000 for -h,l,k 0.007 for -h,-l,-k 0.013 for h,-k,-l	Xtriage
$F_o, F_c$ correlation	0.98	EDS
Total number of atoms	22594	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	14.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.36% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup> Intensities estimated from amplitudes.

<sup>2</sup> Theoretical 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: ZN, PEG, SMC, ACT, MG, F43, MGN, TP7, AGM, TXZ, EDO, GL3, COM, MHS

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	1.60	28/4555 (0.6%)	1.38	44/6181 (0.7%)
1	D	1.61	27/4531 (0.6%)	1.43	49/6148 (0.8%)
2	B	1.58	28/3580 (0.8%)	1.38	24/4841 (0.5%)
2	E	1.61	31/3656 (0.8%)	1.40	26/4943 (0.5%)
3	C	1.64	20/2143 (0.9%)	1.52	28/2885 (1.0%)
3	F	1.67	14/2105 (0.7%)	1.53	32/2833 (1.1%)
All	All	1.61	148/20570 (0.7%)	1.43	203/27831 (0.7%)

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	1	1
1	D	1	0
2	B	0	2
2	E	0	1
All	All	2	4

The worst 5 of 148 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	134	GLU	CD-OE1	9.96	1.36	1.25
2	B	3	LYS	CD-CE	-9.19	1.28	1.51
3	C	153	LYS	CE-NZ	8.90	1.71	1.49
2	B	58	LYS	CB-CG	-8.75	1.28	1.52
1	A	171	GLU	CB-CG	7.69	1.66	1.52

The worst 5 of 203 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	C	147	ARG	NE-CZ-NH1	18.90	129.75	120.30
2	B	44	ARG	NE-CZ-NH1	14.68	127.64	120.30
3	F	147	ARG	NE-CZ-NH1	13.93	127.26	120.30
2	E	353	GLU	OE1-CD-OE2	13.62	139.64	123.30
3	C	147	ARG	NE-CZ-NH2	-11.91	114.34	120.30

All (2) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	A	450	ASP	CA
1	D	450	ASP	CA

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	499	ARG	Sidechain
2	B	333	TYR	Sidechain
2	B	72	ARG	Sidechain
2	E	333	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	4446	0	4259	24	0
1	D	4410	0	4244	22	0
2	B	3464	0	3524	17	0
2	E	3527	0	3584	28	0
3	C	2072	0	2014	14	0
3	F	2041	0	1991	14	0
4	A	2	0	0	0	0
4	B	2	0	0	0	0
4	C	1	0	0	0	0
4	D	3	0	0	0	0
4	E	2	0	0	0	0
4	F	1	0	0	0	0
5	A	62	0	43	1	0
5	D	62	0	43	1	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
6	A	21	0	19	0	0
6	D	21	0	19	0	0
7	A	7	0	5	1	0
7	D	7	0	5	0	0
8	A	20	0	17	0	0
8	D	20	0	17	0	0
9	A	4	0	3	0	0
9	C	4	0	3	0	0
10	A	8	0	12	1	0
10	C	4	0	6	0	0
10	D	8	0	12	2	0
10	F	8	0	12	0	0
11	A	7	0	10	0	0
11	C	7	0	10	0	0
12	A	1	0	0	0	0
13	A	491	0	0	7	0
13	B	454	0	0	4	0
13	C	255	0	0	5	0
13	D	502	0	0	7	0
13	E	400	0	0	11	0
13	F	250	0	0	5	0
All	All	22594	0	19852	112	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 112 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)
3:C:153:LYS:CE	3:C:153:LYS:NZ	1.71	1.54
1:A:352[B]:GLU:HG2	13:A:4123:HOH:O	1.56	1.04
3:C:186[B]:MET:CE	3:C:201:LEU:HD11	2.02	0.87
1:A:24:THR:HG23	13:C:3803:HOH:O	1.73	0.87
3:C:186[B]:MET:HE2	3:C:201:LEU:HD11	1.58	0.85

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	572/549 (104%)	553 (97%)	18 (3%)	1 (0%)	51	21
1	D	569/549 (104%)	555 (98%)	13 (2%)	1 (0%)	51	21
2	B	465/442 (105%)	454 (98%)	11 (2%)	0	100	100
2	E	475/442 (108%)	463 (98%)	12 (2%)	0	100	100
3	C	258/248 (104%)	251 (97%)	7 (3%)	0	100	100
3	F	254/248 (102%)	247 (97%)	7 (3%)	0	100	100
All	All	2593/2478 (105%)	2523 (97%)	68 (3%)	2 (0%)	55	23

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	325	SER
1	D	325	SER

### 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	463/434 (107%)	455 (98%)	8 (2%)	66	33
1	D	460/434 (106%)	455 (99%)	5 (1%)	78	54
2	B	366/341 (107%)	364 (100%)	2 (0%)	91	77
2	E	375/341 (110%)	372 (99%)	3 (1%)	85	65
3	C	228/216 (106%)	226 (99%)	2 (1%)	82	61

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
3	F	224/216 (104%)	222 (99%)	2 (1%)	82	61
All	All	2116/1982 (107%)	2094 (99%)	22 (1%)	80	57

5 of 22 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	283	LYS
1	D	352[A]	GLU
3	F	121	GLN
3	C	61	GLU
3	C	121	GLN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	42	ASN
3	C	121	GLN
3	F	121	GLN
3	F	247	ASN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

10 non-standard protein/DNA/RNA residues 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	$\# Z  > 2$
1	MHS	A	257	1	9,11,12	1.82	3 (33%)	9,14,16	2.37	5 (55%)
1	AGM	A	271	1	11,11,12	1.24	1 (9%)	8,13,15	1.92	3 (37%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
1	MGN	A	400	1	7,9,10	0.99	1 (14%)	5,12,14	0.99	0
1	GL3	A	445	1	3,3,4	2.22	1 (33%)	1,2,4	0.29	0
1	SMC	A	452	1	6,6,7	1.48	1 (16%)	3,6,8	2.36	2 (66%)
1	MHS	D	257	1	9,11,12	1.46	1 (11%)	9,14,16	2.54	4 (44%)
1	AGM	D	271	1	11,11,12	1.51	2 (18%)	8,13,15	0.71	0
1	MGN	D	400	1	7,9,10	1.67	2 (28%)	5,12,14	0.65	0
1	GL3	D	445	1	3,3,4	2.00	1 (33%)	1,2,4	0.29	0
1	SMC	D	452	1	6,6,7	2.15	1 (16%)	3,6,8	1.99	1 (33%)

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
1	MHS	A	257	1	-	0/4/6/8	0/1/1/1
1	AGM	A	271	1	-	0/9/11/13	0/0/0/0
1	MGN	A	400	1	-	0/7/9/12	0/0/0/0
1	GL3	A	445	1	-	0/1/1/2	0/0/0/0
1	SMC	A	452	1	-	0/3/5/7	0/0/0/0
1	MHS	D	257	1	-	0/4/6/8	0/1/1/1
1	AGM	D	271	1	-	0/9/11/13	0/0/0/0
1	MGN	D	400	1	-	0/7/9/12	0/0/0/0
1	GL3	D	445	1	-	0/1/1/2	0/0/0/0
1	SMC	D	452	1	-	0/3/5/7	0/0/0/0

The worst 5 of 14 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	445	GL3	C-S	-3.85	1.68	1.80
1	D	445	GL3	C-S	-3.43	1.69	1.80
1	D	257	MHS	O-C	2.01	1.28	1.19
1	A	271	AGM	CZ-NE1	2.05	1.36	1.33
1	A	400	MGN	O-C	2.10	1.27	1.19

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	257	MHS	O-C-CA	-3.91	114.22	125.02
1	D	452	SMC	CA-CB-SG	-3.34	109.63	114.22

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	257	MHS	O-C-CA	-3.33	115.84	125.02
1	A	452	SMC	CA-CB-SG	-3.24	109.76	114.22
1	A	257	MHS	CG-CB-CA	-2.57	109.18	114.12

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 31 ligands modelled in this entry, 12 are monoatomic - leaving 19 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	F43	A	1	1,13,7	47,71,71	2.85	12 (25%)	48,118,118	1.61	9 (18%)
6	TP7	A	553[A]	-	16,20,20	1.12	1 (6%)	18,26,26	1.25	1 (5%)
7	COM	A	554[A]	5	6,6,6	1.91	2 (33%)	8,8,8	2.43	3 (37%)
8	TXZ	A	555[B]	-	15,19,19	1.17	2 (13%)	17,25,25	1.12	1 (5%)
9	ACT	A	556[A]	4	1,3,3	3.13	1 (100%)	0,3,3	0.00	-
10	EDO	A	557	-	3,3,3	0.37	0	2,2,2	0.12	0
10	EDO	A	558	-	3,3,3	0.50	0	2,2,2	0.82	0
11	PEG	A	559	-	6,6,6	0.96	0	5,5,5	2.11	2 (40%)
9	ACT	C	1	-	1,3,3	3.27	1 (100%)	0,3,3	0.00	-
10	EDO	C	251	-	3,3,3	0.34	0	2,2,2	0.73	0
11	PEG	C	252	-	6,6,6	0.60	0	5,5,5	0.78	0
5	F43	D	553	1,13,7	47,71,71	2.93	11 (23%)	48,118,118	1.80	13 (27%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
6	TP7	D	554[A]	-	16,20,20	1.00	1 (6%)	18,26,26	1.68	4 (22%)
7	COM	D	555[A]	5	6,6,6	2.62	3 (50%)	8,8,8	1.31	1 (12%)
8	TXZ	D	556[B]	-	15,19,19	1.12	1 (6%)	17,25,25	1.01	0
10	EDO	D	557	-	3,3,3	0.61	0	2,2,2	0.11	0
10	EDO	D	558	-	3,3,3	0.64	0	2,2,2	0.69	0
10	EDO	F	251	-	3,3,3	0.54	0	2,2,2	0.50	0
10	EDO	F	252	-	3,3,3	0.54	0	2,2,2	0.41	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
5	F43	A	1	1,13,7	-	0/18/185/185	0/0/10/10
6	TP7	A	553[A]	-	-	0/20/24/24	0/0/0/0
7	COM	A	554[A]	5	-	0/4/4/4	0/0/0/0
8	TXZ	A	555[B]	-	-	0/19/23/23	0/0/0/0
9	ACT	A	556[A]	4	-	0/0/0/0	0/0/0/0
10	EDO	A	557	-	-	0/1/1/1	0/0/0/0
10	EDO	A	558	-	-	0/1/1/1	0/0/0/0
11	PEG	A	559	-	-	0/4/4/4	0/0/0/0
9	ACT	C	1	-	-	0/0/0/0	0/0/0/0
10	EDO	C	251	-	-	0/1/1/1	0/0/0/0
11	PEG	C	252	-	-	0/4/4/4	0/0/0/0
5	F43	D	553	1,13,7	-	0/18/185/185	0/0/10/10
6	TP7	D	554[A]	-	-	0/20/24/24	0/0/0/0
7	COM	D	555[A]	5	-	0/4/4/4	0/0/0/0
8	TXZ	D	556[B]	-	-	0/19/23/23	0/0/0/0
10	EDO	D	557	-	-	0/1/1/1	0/0/0/0
10	EDO	D	558	-	-	0/1/1/1	0/0/0/0
10	EDO	F	251	-	-	0/1/1/1	0/0/0/0
10	EDO	F	252	-	-	0/1/1/1	0/0/0/0

The worst 5 of 35 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
7	D	555[A]	COM	O1S-S2	-4.31	1.32	1.45
7	A	554[A]	COM	C2-S2	-2.91	1.73	1.77
7	A	554[A]	COM	C1-C2	-2.71	1.44	1.51
5	A	1	F43	C2A-C1A	-2.28	1.48	1.51
6	D	554[A]	TP7	C7-S7	-2.24	1.73	1.80

The worst 5 of 34 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	D	553	F43	O8D-C7D-C6D	-5.68	111.03	120.82
6	D	554[A]	TP7	C5-C6-C7	-4.89	104.38	113.09
5	A	1	F43	O8D-C7D-C6D	-4.62	112.86	120.82
5	A	1	F43	C4A-NA-C1A	-3.89	103.44	108.52
5	D	553	F43	CAB-C3B-C2B	-3.60	112.29	119.03

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

5 monomers are involved in 6 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	A	1	F43	1	0
7	A	554[A]	COM	1	0
10	A	558	EDO	1	0
5	D	553	F43	1	0
10	D	558	EDO	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 [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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	543/549 (98%)	-0.69	1 (0%) 94 93	7, 10, 20, 32	0
1	D	543/549 (98%)	-0.70	3 (0%) 89 88	7, 10, 20, 43	0
2	B	442/442 (100%)	-0.69	1 (0%) 94 93	7, 12, 21, 38	0
2	E	442/442 (100%)	-0.60	5 (1%) 80 80	8, 13, 25, 44	0
3	C	246/248 (99%)	-0.50	7 (2%) 53 53	9, 13, 29, 52	0
3	F	246/248 (99%)	-0.41	11 (4%) 34 34	9, 14, 32, 57	0
All	All	2462/2478 (99%)	-0.63	28 (1%) 80 80	7, 12, 23, 57	0

The worst 5 of 28 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	F	60	ASP	7.1
3	C	60	ASP	6.5
1	D	549	ALA	5.0
3	C	62	PRO	4.9
2	B	98	ASP	3.7

### 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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95<sup>th</sup> 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	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
1	MHS	A	257	11/12	0.97	0.05	-	9,11,14,15	0
1	AGM	A	271	12/13	0.98	0.05	-	5,7,7,8	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
1	SMC	A	452	7/8	0.99	0.05	-	7,7,9,11	0
1	MGN	D	400	10/11	0.98	0.04	-	6,8,9,9	0
1	GL3	A	445	4/5	1.00	0.06	-	6,7,8,8	0
1	GL3	D	445	4/5	1.00	0.04	-	7,7,7,8	0
1	MGN	A	400	10/11	0.98	0.05	-	6,8,9,9	0
1	MHS	D	257	11/12	0.97	0.05	-	8,10,13,14	0
1	SMC	D	452	7/8	0.99	0.05	-	8,9,10,12	0
1	AGM	D	271	12/13	0.98	0.06	-	6,7,8,9	0

## 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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95<sup>th</sup> 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	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
7	COM	D	555[A]	7/7	0.96	0.11	6.97	6,10,11,14	7
7	COM	A	554[A]	7/7	0.97	0.09	6.31	9,10,13,14	7
4	MG	C	250	1/1	0.99	0.07	4.53	16,16,16,16	0
4	MG	F	250	1/1	0.98	0.09	3.95	15,15,15,15	0
9	ACT	C	1	4/4	0.90	0.16	3.44	35,36,36,38	0
4	MG	A	552[B]	1/1	0.96	0.18	2.68	20,20,20,20	1
11	PEG	A	559	7/7	0.89	0.10	1.80	21,36,45,45	0
10	EDO	D	558	4/4	0.91	0.09	-0.15	24,33,38,39	0
10	EDO	F	251	4/4	0.86	0.09	-0.15	36,42,45,46	0
6	TP7	D	554[A]	21/21	0.99	0.05	-0.22	7,8,11,12	21
5	F43	D	553	62/62	0.99	0.05	-0.23	6,8,10,12	0
6	TP7	A	553[A]	21/21	0.99	0.05	-0.24	8,9,12,15	21
8	TXZ	D	556[B]	20/20	0.99	0.05	-0.41	7,8,10,11	20
5	F43	A	1	62/62	0.99	0.05	-0.45	7,8,11,13	0
8	TXZ	A	555[B]	20/20	0.99	0.05	-0.48	5,8,10,12	20
9	ACT	A	556[A]	4/4	0.98	0.07	-0.96	14,15,16,16	4
12	ZN	A	560	1/1	1.00	0.03	-2.70	11,11,11,11	1
4	MG	A	551[A]	1/1	0.99	0.21	-	17,17,17,17	1

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
10	EDO	A	557	4/4	0.82	0.09	-	40,46,46,48	0
11	PEG	C	252	7/7	0.79	0.14	-	35,38,45,47	0
4	MG	B	444	1/1	0.93	0.20	-	29,29,29,29	0
4	MG	D	552	1/1	0.96	0.22	-	20,20,20,20	1
10	EDO	A	558	4/4	0.74	0.20	-	48,50,50,51	0
4	MG	D	1	1/1	0.94	0.12	-	23,23,23,23	1
4	MG	E	444	1/1	0.99	0.22	-	21,21,21,21	0
10	EDO	D	557	4/4	0.83	0.11	-	43,43,45,48	0
4	MG	E	445[B]	1/1	0.97	0.14	-	20,20,20,20	1
10	EDO	C	251	4/4	0.93	0.08	-	40,42,44,44	0
4	MG	B	445[A]	1/1	0.97	0.25	-	20,20,20,20	1
4	MG	D	551	1/1	0.98	0.13	-	18,18,18,18	1
10	EDO	F	252	4/4	0.87	0.09	-	35,36,44,44	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.