WORLDWIDE PDB

PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

May 25, 2020 – 02:42 am BST

PDB ID	:	2YXD
Title	:	Crystal Structure of Cobalamin biosynthesis precorrin 8W decarboxylase (cbiT)
Authors	:	Padmanabhan, B.; Bessho, Y.; Yokoyama, S.; RIKEN Structural Genomics/Proteomics Initiative (RSGI)
Deposited on	:	2007-04-26
Resolution	:	2.30 Å(reported)

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

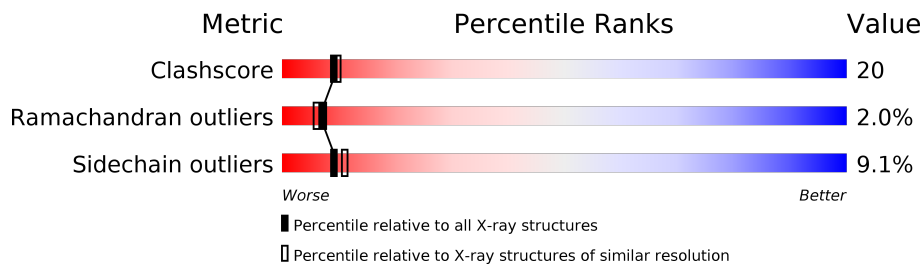|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

1 Overall quality at a glance ⓘ

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

The reported resolution of this entry is 2.30 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
Clashscore		20
Ramachandran outliers		2.0%
Sidechain outliers		9.1%

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))
Clashscore	141614	5643 (2.30-2.30)
Ramachandran outliers	138981	5575 (2.30-2.30)
Sidechain outliers	138945	5575 (2.30-2.30)

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%

Note EDS was not executed.

| Mol	Chain	Length	Quality of chain
1	A	183	60% 31% 7% ··
1	B	183	61% 34% ···

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 2893 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 Probable cobalt-precorrin-6Y C(15)-methyltransferase [decarboxylating].
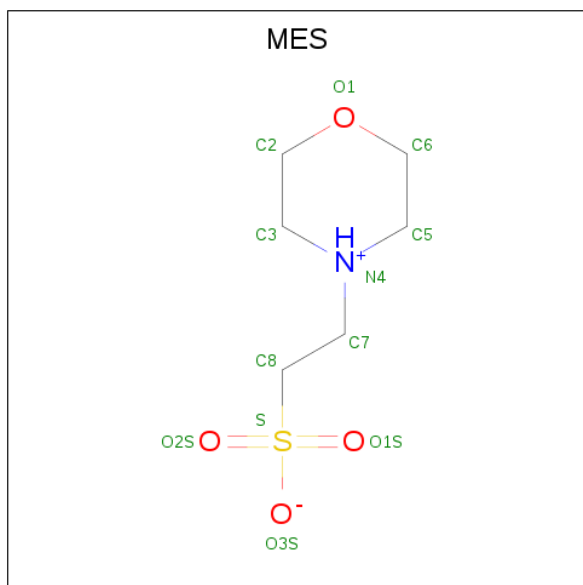
Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	180	Total	C	N	O	S	Se	0	0	0
			1378	881	235	257	3	2			
1	B	180	Total	C	N	O	S	Se	0	0	0
			1385	886	235	259	3	2			

- Molecule 2 is SULFATE ION (three-letter code: SO4) (formula: O_4S).
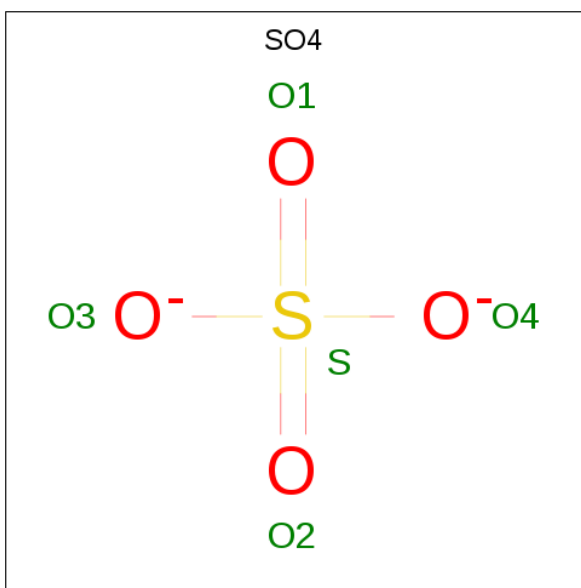
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	O	S	0	0
			5	4	1		
2	B	1	Total	O	S	0	0
			5	4	1		

- Molecule 3 is 2-(N-MORPHOLINO)-ETHANESULFONIC ACID (three-letter code: MES) (formula: $C_6H_{13}NO_4S$).

MES

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	B	1	Total	C	N	O	S	0	0
			12	6	1	4	1		

• Molecule 4 is water.

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	61	Total	O	0	0
			61	61		
4	B	47	Total	O	0	0
			47	47		

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. 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. 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.
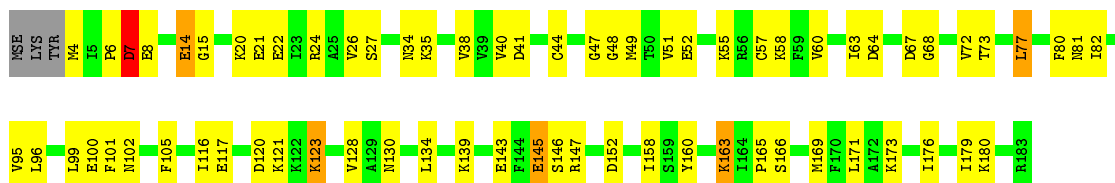
Note EDS was not executed.

• Molecule 1: Probable cobalt-precorrin-6Y C(15)-methyltransferase [decarboxylating]

Chain A:
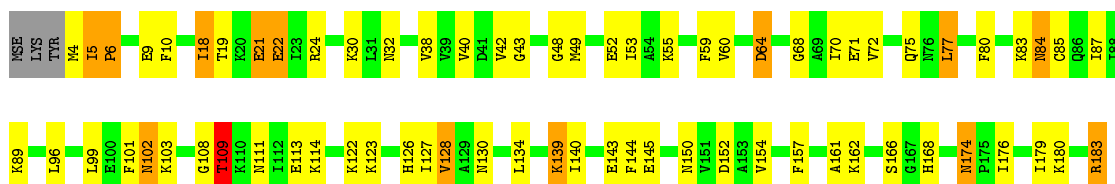
• Molecule 1: Probable cobalt-precorrin-6Y C(15)-methyltransferase [decarboxylating]

Chain B:

4 Data and refinement statistics (i)

Xtriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	P 41 21 2	Depositor
Cell constants a, b, c, α, β, γ	93.44Å 93.44Å 81.04Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 2.30	Depositor
% Data completeness (in resolution range)	94.2 (20.00-2.30)	Depositor
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.198 , 0.273	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	2893	wwPDB-VP
Average B, all atoms (Å^2)	49.0	wwPDB-VP

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: SO4, MES

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	1.34	5/1392 (0.4%)	1.13	5/1870 (0.3%)
1	B	1.47	7/1399 (0.5%)	1.18	7/1879 (0.4%)
All	All	1.41	12/2791 (0.4%)	1.16	12/3749 (0.3%)

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	145	GLU	CB-CG	16.43	1.83	1.52
1	B	145	GLU	CD-OE2	15.24	1.42	1.25
1	A	128	VAL	CB-CG1	7.08	1.67	1.52
1	A	21	GLU	CG-CD	6.84	1.62	1.51
1	A	60	VAL	CB-CG2	-6.68	1.38	1.52

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

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	145	GLU	CG-CD-OE1	-10.25	97.80	118.30
1	B	145	GLU	CG-CD-OE2	8.94	136.19	118.30
1	A	64	ASP	CB-CG-OD1	7.82	125.33	118.30
1	B	67	ASP	CB-CG-OD2	-7.01	111.99	118.30
1	A	109	THR	N-CA-C	6.99	129.86	111.00

There are no chirality outliers.

There are no planarity outliers.

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	1378	0	1421	58	0
1	B	1385	0	1434	60	0
2	A	5	0	0	1	0
2	B	5	0	0	0	0
3	B	12	0	12	0	0
4	A	61	0	0	3	0
4	B	47	0	0	3	0
All	All	2893	0	2867	113	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 20.

The worst 5 of 113 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:B:145:GLU:CG	1:B:145:GLU:CB	1.83	1.57
1:B:169:MSE:HE2	1:B:171:LEU:CD1	1.75	1.16
1:B:169:MSE:HE2	1:B:171:LEU:HD12	1.05	1.04
1:B:169:MSE:CE	1:B:171:LEU:HD12	1.96	0.91
1:A:38:VAL:H	1:A:102:ASN:HD21	1.14	0.87

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	178/183 (97%)	171 (96%)	2 (1%)	5 (3%)	5	3
1	B	178/183 (97%)	166 (93%)	10 (6%)	2 (1%)	14	15
All	All	356/366 (97%)	337 (95%)	12 (3%)	7 (2%)	7	6

5 of 7 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	5	ILE
1	A	6	PRO
1	A	108	GLY
1	A	109	THR
1	A	174	ASN

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	147/153 (96%)	132 (90%)	15 (10%)	7	8
1	B	149/153 (97%)	137 (92%)	12 (8%)	11	15
All	All	296/306 (97%)	269 (91%)	27 (9%)	9	11

5 of 27 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	157	PHE
1	B	7	ASP
1	B	163	LYS
1	A	166	SER
1	A	77	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 11 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	102	ASN
1	A	126	HIS
1	B	86	GLN
1	A	86	GLN
1	A	174	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)

3 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	MES	B	2001	-	12,12,12	1.48	2 (16%)	14,16,16	6.33	8 (57%)
2	SO4	B	1002	-	4,4,4	0.39	0	6,6,6	1.20	0
2	SO4	A	1001	-	4,4,4	0.48	0	6,6,6	0.84	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	MES	B	2001	-	-	3/6/14/14	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	2001	MES	O2S-S	2.32	1.51	1.45

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	2001	MES	C7-C8	2.02	1.58	1.52

The worst 5 of 8 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	2001	MES	O3S-S-C8	-15.92	80.02	105.77
3	B	2001	MES	O2S-S-C8	-11.01	93.66	106.92
3	B	2001	MES	O1S-S-C8	-7.23	98.21	106.92
3	B	2001	MES	O3S-S-O2S	6.91	128.15	111.27
3	B	2001	MES	C5-N4-C3	6.37	123.17	108.83

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	B	2001	MES	C8-C7-N4-C5
3	B	2001	MES	N4-C7-C8-S
3	B	2001	MES	C8-C7-N4-C3

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	1001	SO4	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)

EDS was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.

6.3 Carbohydrates (i)

EDS was not executed - this section is therefore empty.

6.4 Ligands (i)

EDS was not executed - this section is therefore empty.

6.5 Other polymers (i)

EDS was not executed - this section is therefore empty.