Full wwPDB X-ray Structure Validation Report (i)

Oct 31, 2023 – 07:01 AM EDT

PDB ID	:	3IWT
Title	:	Structure of hypothetical molybdenum cofactor biosynthesis protein B from Sulfolobus tokodaii
Authors	:	Antonyuk, S.V.; Ellis, M.J.; Strange, R.W.; Hasnain, S.S.; Bessho, Y.; Kuramitsu, S.; Yokoyama, S.; RIKEN Structural Genomics/Proteomics Initiative (RSGI)
Deposited on	:	2009-09-03
Resolution	:	1.90 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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 types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

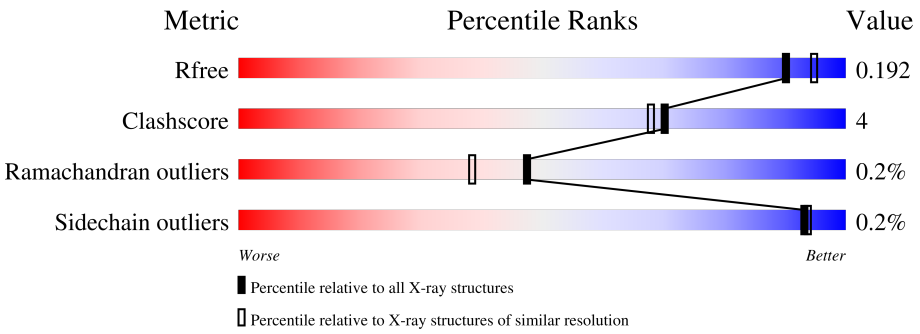
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 1.90 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	6207 (1.90-1.90)
Clashscore	141614	6847 (1.90-1.90)
Ramachandran outliers	138981	6760 (1.90-1.90)
Sidechain outliers	138945	6760 (1.90-1.90)

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 <=5%

Mol	Chain	Length	Quality of chain
1	A	178	87% 7% 7%
1	B	178	83% 10% 7%
1	C	178	85% 8% 7%

2 Entry composition (i)

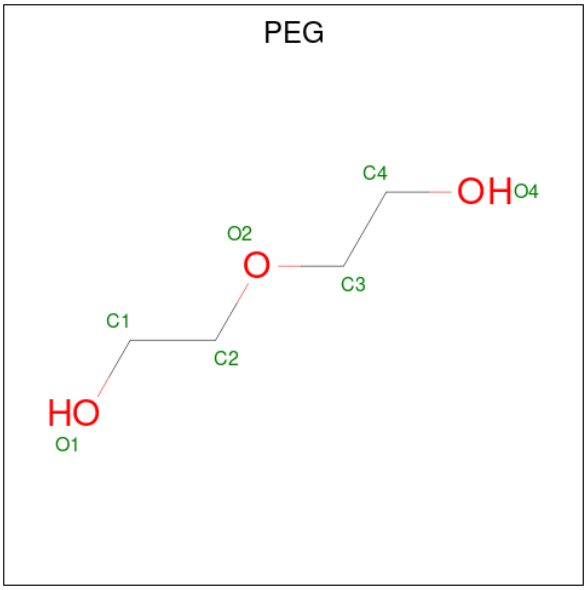
There are 5 unique types of molecules in this entry. The entry contains 4698 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 178aa long hypothetical molybdenum cofactor biosynthesis protein B.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
1	A	166	1352	882	209	261	0	12	0
1	B	166	1355	892	206	257	0	13	0
1	C	166	1332	871	210	251	0	7	0

- Molecule 2 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula: $C_4H_{10}O_3$).



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

Continued on next page...

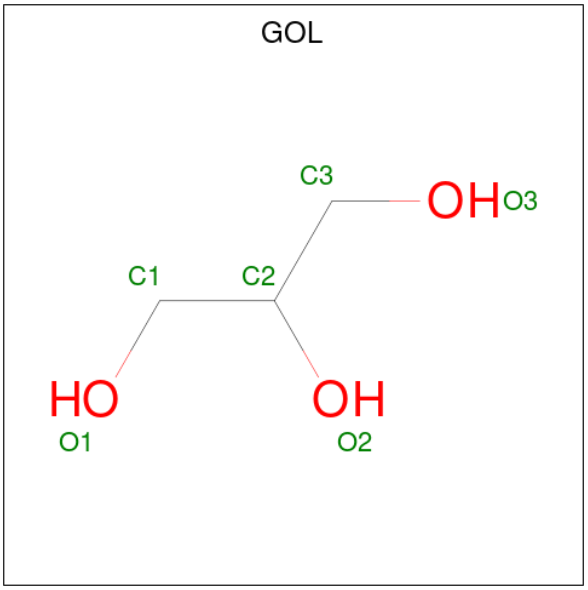
Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	B	1	Total C O 14 8 6	0	1
2	B	1	Total C O 7 4 3	0	0
2	B	1	Total C O 7 4 3	0	0
2	B	1	Total C O 7 4 3	0	0
2	C	1	Total C O 7 4 3	0	0

- Molecule 3 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total Mg 1 1	0	0
3	B	1	Total Mg 1 1	0	1

- Molecule 4 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



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

- Molecule 5 is water.

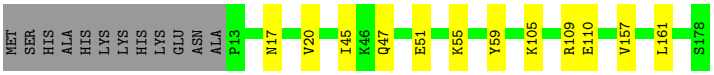
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
5	A	202	202	202	0	4
			Total	O		
5	B	217	217	217	0	5
			Total	O		
5	C	162	163	163	0	2

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. 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.

• Molecule 1: 178aa long hypothetical molybdenum cofactor biosynthesis protein B

Chain A:



• Molecule 1: 178aa long hypothetical molybdenum cofactor biosynthesis protein B

Chain B:

• Molecule 1: 178aa long hypothetical molybdenum cofactor biosynthesis protein B

Chain C:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 64 2 2	Depositor
Cell constants a, b, c, α, β, γ	136.68Å 136.68Å 210.53Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	29.74 – 1.90 29.74 – 1.90	Depositor EDS
% Data completeness (in resolution range)	98.2 (29.74-1.90) 98.2 (29.74-1.90)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.98 (at 1.91Å)	Xtriage
Refinement program	REFMAC 5.3.0037	Depositor
R, R_{free}	0.168 , 0.187 0.175 , 0.192	Depositor DCC
R_{free} test set	4506 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å^2)	32.5	Xtriage
Anisotropy	0.048	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 38.1	EDS
L-test for twinning[2]	$< \|L\| > = 0.51$, $< L^2 > = 0.35$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	4698	wwPDB-VP
Average B, all atoms (Å^2)	47.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.95% 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 (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: MG, GOL, PEG

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	0.73	2/1409 (0.1%)	0.72	0/1904
1	B	0.70	0/1413	0.75	0/1912
1	C	0.57	0/1374	0.71	0/1858
All	All	0.67	2/4196 (0.0%)	0.73	0/5674

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	51[A]	GLU	CD-OE2	5.42	1.31	1.25
1	A	51[B]	GLU	CD-OE2	5.42	1.31	1.25

There are no bond angle outliers.

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	1352	0	1429	10	0
1	B	1355	0	1450	15	0
1	C	1332	0	1415	10	0
2	A	7	0	10	0	0
2	B	49	0	70	3	0
2	C	7	0	10	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	A	1	0	0	0	0
3	B	1	0	0	0	0
4	A	6	0	8	0	0
4	B	6	0	8	0	0
5	A	202	0	0	3	0
5	B	217	0	0	2	0
5	C	163	0	0	0	0
All	All	4698	0	4400	31	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.

All (31) 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:C:101:GLU:OE2	1:C:136:LYS:NZ	2.24	0.71
1:B:45:ILE:HA	1:B:161[A]:LEU:HD11	1.78	0.66
1:B:158:LYS:HE3	2:B:209:PEG:H42	1.83	0.59
1:A:109[B]:ARG:HH21	1:C:136:LYS:NZ	2.03	0.56
1:C:45:ILE:HA	1:C:161:LEU:HD11	1.91	0.53
1:A:105:LYS:NZ	5:A:213:HOH:O	2.41	0.53
1:B:162:LYS:HB2	2:B:209:PEG:H12	1.91	0.52
1:A:109[B]:ARG:HH21	1:C:136:LYS:HZ1	1.57	0.51
1:B:134:LEU:HD22	1:C:172:LEU:HD13	1.93	0.50
1:B:100[B]:VAL:HG12	5:B:679:HOH:O	2.11	0.50
1:A:109[B]:ARG:NH2	1:C:136:LYS:HZ1	2.09	0.50
1:B:158:LYS:CE	2:B:209:PEG:H42	2.43	0.49
1:C:104:ARG:HD2	1:C:110:GLU:OE2	2.13	0.49
1:A:110[B]:GLU:CG	5:A:806:HOH:O	2.61	0.48
1:B:45:ILE:HA	1:B:161[A]:LEU:CD1	2.45	0.47
1:A:45:ILE:HA	1:A:161:LEU:HD11	1.98	0.45
1:A:45:ILE:HG13	1:A:157:VAL:HG11	1.98	0.45
1:C:120:LEU:HD23	1:C:120:LEU:C	2.37	0.45
1:C:30:LYS:HB2	1:C:37[A]:ILE:HG22	2.00	0.44
1:A:47[B]:GLN:HB3	5:A:216:HOH:O	2.18	0.44
1:B:101:GLU:OE2	1:B:136[A]:LYS:NZ	2.46	0.44
1:B:40[B]:GLU:HG2	5:B:709:HOH:O	2.17	0.43
1:C:175:LEU:O	1:C:178:SER:HB2	2.18	0.43
1:B:50[A]:ILE:HD11	1:B:56[A]:ILE:CD1	2.48	0.43
1:B:50[A]:ILE:HG13	1:B:56[A]:ILE:CD1	2.50	0.42
1:B:45:ILE:HG13	1:B:157:VAL:HG11	2.02	0.42

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:144[B]:LYS:HB2	1:B:177:ARG:HG2	2.02	0.41
1:A:20:VAL:O	1:A:59:TYR:HA	2.21	0.41
1:B:21:ILE:O	1:B:87:SER:HA	2.21	0.41
1:B:20:VAL:O	1:B:59:TYR:HA	2.21	0.40
1:A:17:ASN:HD22	1:A:55:LYS:HB2	1.86	0.40

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	175/178 (98%)	173 (99%)	2 (1%)	0	100	100
1	B	176/178 (99%)	173 (98%)	3 (2%)	0	100	100
1	C	171/178 (96%)	169 (99%)	1 (1%)	1 (1%)	25	15
All	All	522/534 (98%)	515 (99%)	6 (1%)	1 (0%)	47	38

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	38	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	157/156 (101%)	157 (100%)	0	100	100
1	B	158/156 (101%)	158 (100%)	0	100	100
1	C	153/156 (98%)	152 (99%)	1 (1%)	84	84
All	All	468/468 (100%)	467 (100%)	1 (0%)	93	94

All (1) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	14	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	17	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 monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 13 ligands modelled in this entry, 2 are monoatomic - leaving 11 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	PEG	B	181[B]	-	6,6,6	0.55	0	5,5,5	0.20	0
2	PEG	B	183	-	6,6,6	0.54	0	5,5,5	0.24	0
2	PEG	B	209	-	6,6,6	0.72	0	5,5,5	0.96	0
2	PEG	A	180	-	6,6,6	0.52	0	5,5,5	0.19	0
2	PEG	B	179	-	6,6,6	0.43	0	5,5,5	0.45	0
4	GOL	A	179	-	5,5,5	0.75	0	5,5,5	1.13	0
4	GOL	B	180	-	5,5,5	0.47	0	5,5,5	0.52	0
2	PEG	B	181[A]	-	6,6,6	0.37	0	5,5,5	0.33	0
2	PEG	B	208	-	6,6,6	0.61	0	5,5,5	0.49	0
2	PEG	B	182	-	6,6,6	0.49	0	5,5,5	0.32	0
2	PEG	C	180	-	6,6,6	0.38	0	5,5,5	0.27	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	PEG	B	181[B]	-	-	3/4/4/4	-
2	PEG	B	183	-	-	2/4/4/4	-
2	PEG	B	209	-	-	2/4/4/4	-
2	PEG	A	180	-	-	2/4/4/4	-
2	PEG	B	179	-	-	3/4/4/4	-
4	GOL	A	179	-	-	2/4/4/4	-
4	GOL	B	180	-	-	3/4/4/4	-
2	PEG	B	181[A]	-	-	2/4/4/4	-
2	PEG	B	208	-	-	2/4/4/4	-
2	PEG	B	182	-	-	3/4/4/4	-
2	PEG	C	180	-	-	2/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (26) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	179	GOL	C1-C2-C3-O3
4	A	179	GOL	O2-C2-C3-O3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
2	A	180	PEG	O1-C1-C2-O2
2	B	209	PEG	O1-C1-C2-O2
2	B	181[B]	PEG	O2-C3-C4-O4
4	B	180	GOL	C1-C2-C3-O3
2	B	182	PEG	O2-C3-C4-O4
4	B	180	GOL	O2-C2-C3-O3
2	B	209	PEG	O2-C3-C4-O4
2	B	181[B]	PEG	O1-C1-C2-O2
2	B	182	PEG	O1-C1-C2-O2
2	B	179	PEG	O2-C3-C4-O4
4	B	180	GOL	O1-C1-C2-O2
2	B	183	PEG	O2-C3-C4-O4
2	B	208	PEG	O2-C3-C4-O4
2	C	180	PEG	C1-C2-O2-C3
2	B	182	PEG	C1-C2-O2-C3
2	B	181[A]	PEG	C1-C2-O2-C3
2	B	208	PEG	C4-C3-O2-C2
2	B	179	PEG	C1-C2-O2-C3
2	B	183	PEG	C4-C3-O2-C2
2	B	181[B]	PEG	C1-C2-O2-C3
2	C	180	PEG	O1-C1-C2-O2
2	A	180	PEG	C4-C3-O2-C2
2	B	179	PEG	C4-C3-O2-C2
2	B	181[A]	PEG	O2-C3-C4-O4

There are no ring outliers.

1 monomer is involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	209	PEG	3	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)

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates (i)

Unable to reproduce the depositors R factor - this section is therefore empty.

6.4 Ligands (i)

Unable to reproduce the depositors R factor - this section is therefore empty.

6.5 Other polymers (i)

Unable to reproduce the depositors R factor - this section is therefore empty.