Full wwPDB X-ray Structure Validation Report (i)

Jan 31, 2016 – 11:14 PM GMT

PDB ID	:	1WU2
Title	:	Crystal Structure of molybdopterin biosynthesis moeA protein from Pyrococcus horikoshii OT3
Authors	:	Lokanath, N.K.; Kunishima, N.; RIKEN Structural Genomics/Proteomics Initiative (RSGI)
Deposited on	:	2004-11-30
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	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
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)	:	trunk26865

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 2.30 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		3
Ramachandran outliers		0
Sidechain outliers		8.0%
RSRZ outliers		5.5%

Worse   *Better*

▮ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution
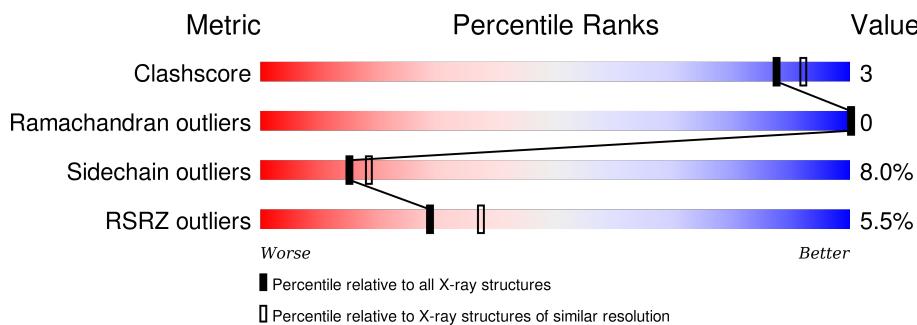
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
Clashscore	102246	4452 (2.30-2.30)
Ramachandran outliers	100387	4410 (2.30-2.30)
Sidechain outliers	100360	4409 (2.30-2.30)
RSRZ outliers	91569	3857 (2.30-2.30)

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. 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	396	5% / 83% / 12% / · ·
1	B	396	5% / 84% / 10% / · ·

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 6401 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 molybdopterin biosynthesis moeA protein.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	383	Total 3009	C 1943	N 494	O 565	S 1	Se 6	0	0	0
1	B	379	Total 2969	C 1917	N 487	O 558	S 1	Se 6	0	0	0

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MSE	MET	MODIFIED RESIDUE	GB 14591417
A	2	MSE	MET	MODIFIED RESIDUE	GB 14591417
A	118	MSE	MET	MODIFIED RESIDUE	GB 14591417
A	121	MSE	MET	MODIFIED RESIDUE	GB 14591417
A	168	MSE	MET	MODIFIED RESIDUE	GB 14591417
A	215	MSE	MET	MODIFIED RESIDUE	GB 14591417
A	296	MSE	MET	MODIFIED RESIDUE	GB 14591417
A	318	MSE	MET	MODIFIED RESIDUE	GB 14591417
B	1	MSE	MET	MODIFIED RESIDUE	GB 14591417
B	2	MSE	MET	MODIFIED RESIDUE	GB 14591417
B	118	MSE	MET	MODIFIED RESIDUE	GB 14591417
B	121	MSE	MET	MODIFIED RESIDUE	GB 14591417
B	168	MSE	MET	MODIFIED RESIDUE	GB 14591417
B	215	MSE	MET	MODIFIED RESIDUE	GB 14591417
B	296	MSE	MET	MODIFIED RESIDUE	GB 14591417
B	318	MSE	MET	MODIFIED RESIDUE	GB 14591417

- Molecule 2 is water.

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	226	Total 226	O 226	0	0
2	B	197	Total 197	O 197	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 errors 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: molybdopterin biosynthesis moeA protein
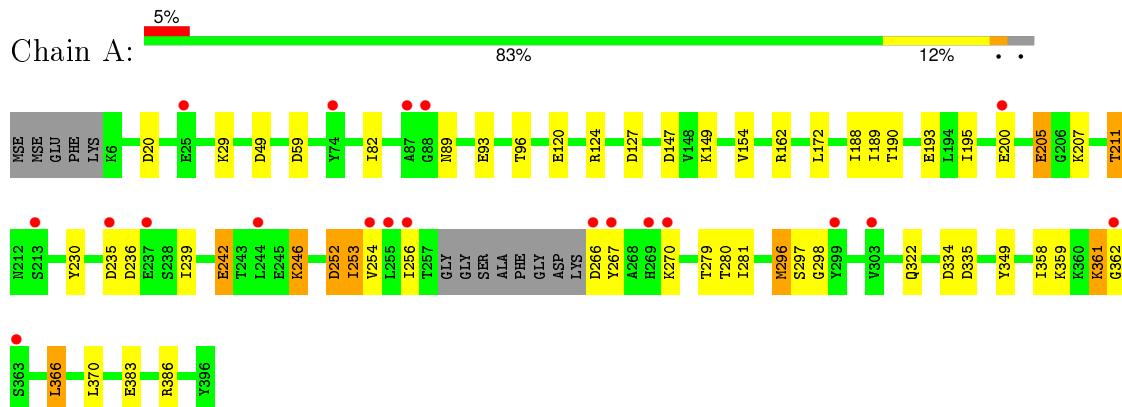
- Molecule 1: molybdopterin biosynthesis moeA protein
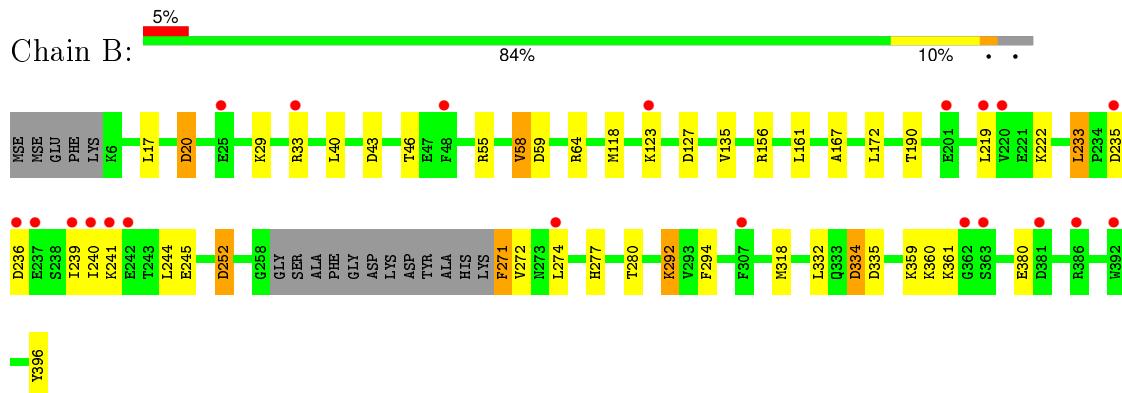
4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	173.15Å 65.91Å 76.68Å 90.00° 102.17° 90.00°	Depositor
Resolution (Å)	20.00 – 2.30 28.45 – 2.28	Depositor EDS
% Data completeness (in resolution range)	98.1 (20.00-2.30) 97.0 (28.45-2.28)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$<I/\sigma(I)>$ [1]	2.50 (at 2.29Å)	Xtriage
Refinement program	REFMAC 5.1.27	Depositor
R, R_{free}	0.233 , 0.282 0.236 , (Not available)	Depositor DCC
R_{free} test set	No test flags present.	DCC
Wilson B-factor (Å^2)	33.7	Xtriage
Anisotropy	0.507	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 35.1	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning[2]	$<|L|> = 0.46$, $<L^2> = 0.29$	Xtriage
Outliers	0 of 37435 reflections	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	6401	wwPDB-VP
Average B, all atoms (Å^2)	38.0	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.*

---

[1]Intensities estimated from amplitudes.

[2]Theoretical values of $<|L|>$, $<L^2>$ for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 0.2 for perfectly twinned datasets.

5 Model quality (i)

5.1 Standard geometry (i)

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	0.30	0/3056	0.66	11/4117 (0.3%)
1	B	0.29	0/3014	0.62	7/4060 (0.2%)
All	All	0.29	0/6070	0.64	18/8177 (0.2%)

There are no bond length outliers.

All (18) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	127	ASP	CB-CG-OD2	6.16	123.84	118.30
1	B	127	ASP	CB-CG-OD2	6.09	123.78	118.30
1	B	334	ASP	CB-CG-OD2	5.79	123.51	118.30
1	A	266	ASP	CB-CG-OD2	5.77	123.50	118.30
1	A	59	ASP	CB-CG-OD2	5.76	123.48	118.30
1	A	335	ASP	CB-CG-OD2	5.76	123.48	118.30
1	A	252	ASP	CB-CG-OD2	5.50	123.25	118.30
1	B	59	ASP	CB-CG-OD2	5.50	123.25	118.30
1	B	20	ASP	CB-CG-OD2	5.49	123.24	118.30
1	A	147	ASP	CB-CG-OD2	5.46	123.21	118.30
1	A	20	ASP	CB-CG-OD2	5.41	123.17	118.30
1	B	252	ASP	CB-CG-OD2	5.23	123.01	118.30
1	A	49	ASP	CB-CG-OD2	5.21	122.99	118.30
1	A	235	ASP	CB-CG-OD2	5.19	122.97	118.30
1	B	335	ASP	CB-CG-OD2	5.13	122.92	118.30
1	A	334	ASP	CB-CG-OD2	5.06	122.85	118.30
1	B	43	ASP	CB-CG-OD2	5.02	122.82	118.30
1	A	236	ASP	CB-CG-OD2	5.01	122.81	118.30

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	3009	0	3105	19	0
1	B	2969	0	3070	14	0
2	A	226	0	0	0	0
2	B	197	0	0	1	0
All	All	6401	0	6175	32	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.

All (32) 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:361:LYS:CB	1:A:362:GLY:HA3	2.18	0.74
1:A:361:LYS:HB3	1:A:362:GLY:HA3	1.73	0.70
1:A:239:ILE:O	1:A:242:GLU:HG2	1.97	0.64
1:A:256:ILE:HG23	1:A:298:GLY:HA2	1.79	0.64
1:A:230:TYR:CE2	1:A:246:LYS:HG3	2.38	0.58
1:B:190:THR:HG22	1:B:233:LEU:HD21	1.90	0.52
1:B:46:THR:HG23	1:B:172:LEU:HG	1.92	0.51
1:A:361:LYS:CG	1:A:362:GLY:HA3	2.41	0.50
1:A:211:THR:HG23	1:B:167:ALA:HB1	1.94	0.49
1:A:242:GLU:C	1:A:246:LYS:HZ2	2.18	0.47
1:B:222:LYS:NZ	1:B:396:TYR:OXT	2.48	0.47
1:B:274:LEU:HD21	1:B:277:HIS:HB2	1.95	0.47
1:A:242:GLU:O	1:A:246:LYS:HD3	2.16	0.46
1:A:349:TYR:HB2	1:A:358:ILE:HD11	1.98	0.46
1:A:361:LYS:HB3	1:A:362:GLY:CA	2.43	0.45
1:A:361:LYS:HG2	1:A:362:GLY:HA3	1.97	0.45
1:B:20:ASP:O	1:B:292:LYS:NZ	2.50	0.45
1:A:205:GLU:HB3	1:A:207:LYS:HZ2	1.83	0.44
1:A:188:ILE:HD12	1:A:253:ILE:HD11	2.01	0.43
1:A:366:LEU:HD13	1:A:370:LEU:HD13	2.00	0.43
1:B:64:ARG:NH1	2:B:438:HOH:O	2.52	0.43
1:B:235:ASP:HA	1:B:239:ILE:HD11	2.01	0.43
1:B:58:VAL:HG21	1:B:118:MSE:HG2	2.02	0.42

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:254:VAL:HG13	1:A:296:MSE:SE	2.69	0.42
1:B:280:THR:HG21	1:B:380:GLU:OE2	2.18	0.42
1:B:40:LEU:HD13	1:B:161:LEU:HD21	2.02	0.42
1:A:154:VAL:HG21	1:A:172:LEU:HD11	2.02	0.42
1:B:17:LEU:HD11	1:B:294:PHE:CE2	2.54	0.42
1:B:252:ASP:CB	1:B:318:MSE:HE1	2.51	0.41
1:A:279:THR:HG23	1:A:281:ILE:H	1.86	0.41
1:A:242:GLU:OE2	1:A:246:LYS:NZ	2.40	0.41
1:B:240:ILE:HG23	1:B:271:PHE:N	2.36	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	379/396 (96%)	364 (96%)	15 (4%)	0	100	100
1	B	375/396 (95%)	362 (96%)	13 (4%)	0	100	100
All	All	754/792 (95%)	726 (96%)	28 (4%)	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	325/326 (100%)	294 (90%)	31 (10%)	11	12
1	B	321/326 (98%)	300 (94%)	21 (6%)	21	27
All	All	646/652 (99%)	594 (92%)	52 (8%)	15	18

All (52) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	29	LYS
1	A	82	ILE
1	A	89	ASN
1	A	93	GLU
1	A	96	THR
1	A	120	GLU
1	A	124	ARG
1	A	149	LYS
1	A	162	ARG
1	A	189	ILE
1	A	190	THR
1	A	193	GLU
1	A	195	ILE
1	A	200	GLU
1	A	205	GLU
1	A	211	THR
1	A	242	GLU
1	A	246	LYS
1	A	252	ASP
1	A	253	ILE
1	A	267	TYR
1	A	270	LYS
1	A	280	THR
1	A	296	MSE
1	A	297	SER
1	A	322	GLN
1	A	359	LYS
1	A	361	LYS
1	A	366	LEU
1	A	383	GLU
1	A	386	ARG
1	B	29	LYS
1	B	33	ARG
1	B	55	ARG
1	B	58	VAL

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	123	LYS
1	B	135	VAL
1	B	156	ARG
1	B	219	LEU
1	B	233	LEU
1	B	236	ASP
1	B	241	LYS
1	B	244	LEU
1	B	245	GLU
1	B	271	PHE
1	B	272	VAL
1	B	292	LYS
1	B	332	LEU
1	B	334	ASP
1	B	359	LYS
1	B	360	LYS
1	B	361	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	217	GLN
1	A	322	GLN
1	A	352	ASN
1	B	273	ASN
1	B	352	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)

There are no ligands in this entry.

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	377/396 (95%)	0.42	20 (5%)	30	39	23, 35, 57, 63	0
1	B	373/396 (94%)	0.41	21 (5%)	28	36	24, 37, 55, 77	0
All	All	750/792 (94%)	0.41	41 (5%)	29	37	23, 36, 57, 77	0

All (41) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	240	ILE	7.5
1	B	241	LYS	6.3
1	B	363	SER	5.6
1	A	266	ASP	4.9
1	B	235	ASP	4.6
1	A	363	SER	4.2
1	B	239	ILE	4.0
1	A	269	HIS	3.7
1	A	362	GLY	3.7
1	A	235	ASP	3.4
1	A	299	TYR	3.3
1	A	267	TYR	3.2
1	A	254	VAL	3.1
1	B	242	GLU	3.0
1	B	274	LEU	3.0
1	B	25	GLU	2.9
1	B	201	GLU	2.9
1	A	74	TYR	2.8
1	B	362	GLY	2.7
1	A	256	ILE	2.7
1	B	236	ASP	2.7
1	A	200	GLU	2.7
1	A	25	GLU	2.7
1	B	392	TRP	2.6

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	381	ASP	2.6
1	A	270	LYS	2.5
1	B	237	GLU	2.4
1	A	88	GLY	2.3
1	A	255	LEU	2.3
1	B	48	PHE	2.2
1	A	303	VAL	2.2
1	B	219	LEU	2.2
1	A	244	LEU	2.2
1	B	33	ARG	2.2
1	B	386	ARG	2.1
1	B	123	LYS	2.1
1	A	87	ALA	2.0
1	A	237	GLU	2.0
1	B	220	VAL	2.0
1	B	307	PHE	2.0
1	A	213	SER	2.0

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)

There are no ligands in this entry.

6.5   Other polymers (i)

There are no such residues in this entry.