Full wwPDB X-ray Structure Validation Report ⓘ

Apr 9, 2014 – 07:29 PM EDT

PDB ID : 4MLM
Title : Crystal Structure of PhnZ from uncultured bacterium HF130_AEPn_1
Authors : van Staalduinen, L.M.; McSorley, F.R.; Zechel, D.L.; Jia, Z.; Montreal-Kingston Bacterial Structural Genomics Initiative (BSGI)
Deposited on : 2013-09-06
Resolution : 1.70 Å(reported)

This is a full wwPDB validation report for a publicly released PDB entry.
We welcome your comments at validation@mail.wwpdb.org
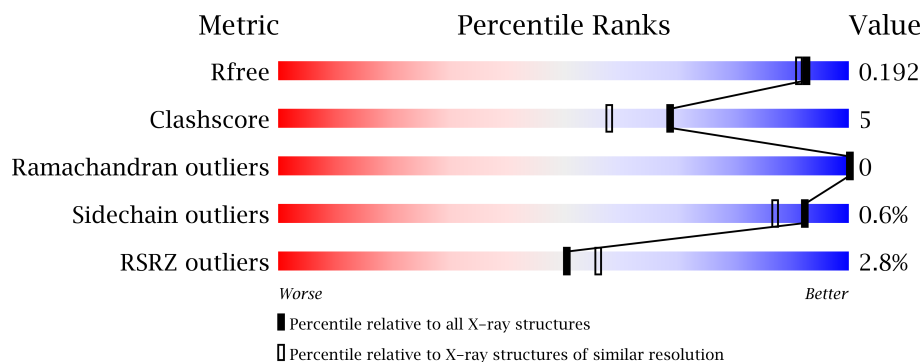A user guide is available at http://wwpdb.org/ValidationPDFNotes.html

---

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

MolProbity	:	4.02b-467
Mogul	:	FAILED
Xtriage (Phenix)	:	dev-1439
EDS	:	stable22978
Percentile statistics	:	21963
Refmac	:	5.8.0049
CCP4	:	6.3.0 (Settle)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	stable22978

1 Overall quality at a glance ⓘ

The reported resolution of this entry is 1.70 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
Rfree		0.192
Clashscore		5
Ramachandran outliers		0
Sidechain outliers		0.6%
RSRZ outliers		2.8%

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))
R_{free}	66092	2456 (1.70-1.70)
Clashscore	79885	2929 (1.70-1.70)
Ramachandran outliers	78287	2878 (1.70-1.70)
Sidechain outliers	78261	2878 (1.70-1.70)
RSRZ outliers	66119	2456 (1.70-1.70)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	196	
1	B	196	

The following table lists non-polymeric compounds that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Geometry	Electron density
4	SO4	A	204	-	X
4	SO4	A	205	-	X
5	GOL	A	206	-	X
7	NA	B	205	-	X

2   Entry composition ⓘ

There are 8 unique types of molecules in this entry. The entry contains 3469 atoms, of which 8 are hydrogens and 0 are deuterium.

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 Predicted HD phosphohydrolase PhnZ.

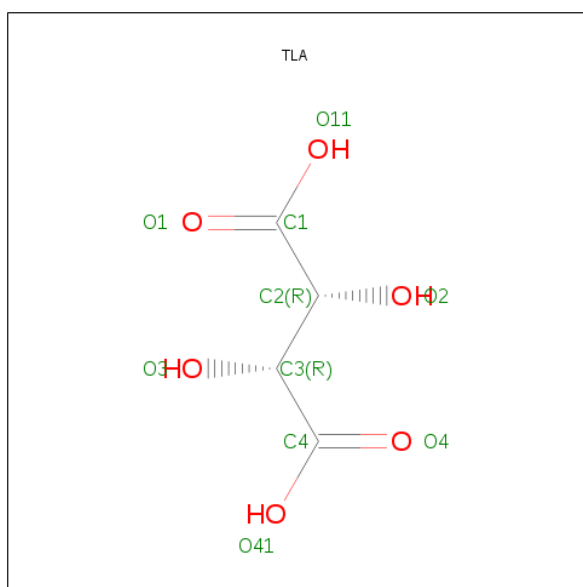| Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	188	Total	C	N	O	S	Se	11	4	0
			1513	942	263	298	4	6			
1	B	187	Total	C	N	O	S	Se	11	3	0
			1511	940	265	297	4	5			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	191	HIS	-	EXPRESSION TAG	UNP D0E8I5
A	192	HIS	-	EXPRESSION TAG	UNP D0E8I5
A	193	HIS	-	EXPRESSION TAG	UNP D0E8I5
A	194	HIS	-	EXPRESSION TAG	UNP D0E8I5
A	195	HIS	-	EXPRESSION TAG	UNP D0E8I5
A	196	HIS	-	EXPRESSION TAG	UNP D0E8I5
B	191	HIS	-	EXPRESSION TAG	UNP D0E8I5
B	192	HIS	-	EXPRESSION TAG	UNP D0E8I5
B	193	HIS	-	EXPRESSION TAG	UNP D0E8I5
B	194	HIS	-	EXPRESSION TAG	UNP D0E8I5
B	195	HIS	-	EXPRESSION TAG	UNP D0E8I5
B	196	HIS	-	EXPRESSION TAG	UNP D0E8I5

- Molecule 2 is FE (III) ION (three-letter code: FE) (formula: Fe).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	B	2	Total	Fe	0	0
			2	2		
2	A	2	Total	Fe	0	0
			2	2		

- Molecule 3 is L(+)-TARTARIC ACID (three-letter code: TLA) (formula: $C_4H_6O_6$).

![Worldwide Protein Data Bank logo]

TLA

O11
OH

O1 O   C1

C2(R) ·······||||OH O2

O3 HO||||||···· C3(R)

C4   O O4

HO
O41

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total 14	C 4	H 4	O 6	0	0
3	B	1	Total 14	C 4	H 4	O 6	0	0

- Molecule 4 is SULFATE ION (three-letter code: SO4) (formula: O_4S).

SO4

O1

O

O3 O⁻ —— S —— O⁻ O4

S

O

O2
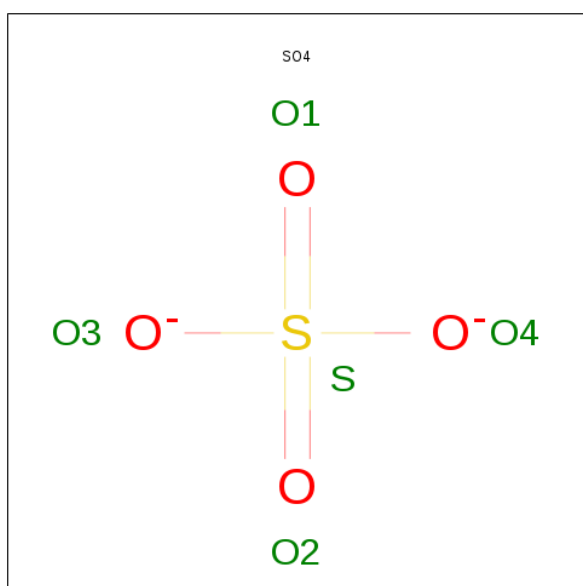
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total 5	O 4	S 1	0	0
4	A	1	Total 5	O 4	S 1	0	0

- Molecule 5 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).
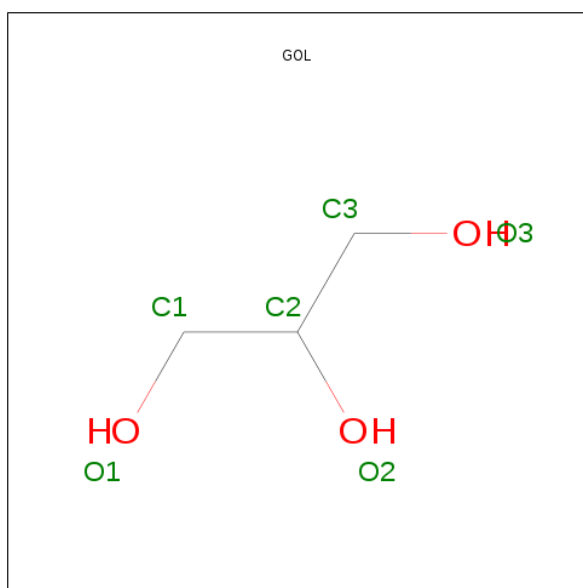
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total 6	C 3	O 3	0	0

- Molecule 6 is SUGAR (B-OCTYLGLUCOSIDE) (three-letter code: BOG) (formula: $C_{14}H_{28}O_6$).
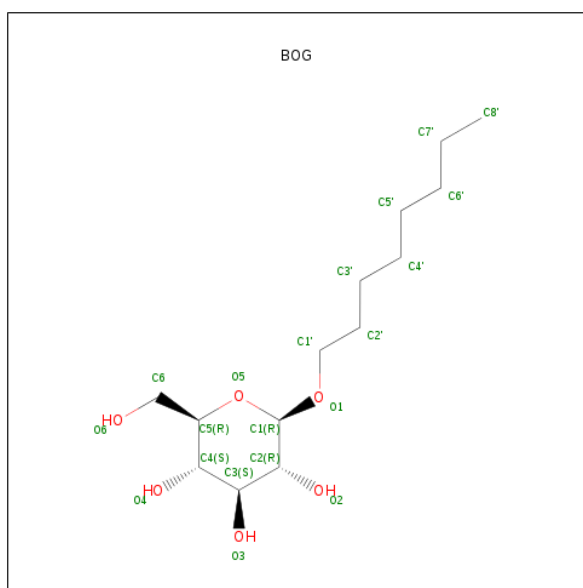
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	B	1	Total 20	C 14	O 6	0	0

- Molecule 7 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	B	1	Total 1	Na 1	0	0

- Molecule 8 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	A	188	Total 188	O 188	0	0
8	B	188	Total 188	O 188	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: Predicted HD phosphohydrolase PhnZ

Chain A:
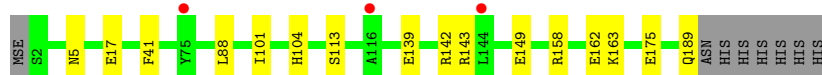
• Molecule 1: Predicted HD phosphohydrolase PhnZ

Chain B:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α, β, γ	104.67Å 76.04Å 60.70Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	19.80 – 1.70 19.82 – 1.70	Depositor EDS
% Data completeness (in resolution range)	100.0 (19.80-1.70) 100.0 (19.82-1.70)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.06	Depositor
$< I/\sigma(I) >$ [1]	4.38 (at 1.70Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.7.3_928)	Depositor
R, R_{free}	0.172 , 0.198 0.164 , 0.192	Depositor DCC
R_{free} test set	2742 reflections (5.08%)	DCC
Wilson B-factor (Å^2)	19.0	Xtriage
Anisotropy	0.014	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.40 , 42.4	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$< |L| > = 0.47, < L^2 > = 0.30$	Xtriage
Outliers	0 of 53982 reflections	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	3469	wwPDB-VP
Average B, all atoms (Å^2)	22.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.92% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.

5 Model quality ⓘ

5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: GOL, NA, SO4, TLA, FE, BOG

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	0/1547	0.45	0/2067
1	B	0.29	0/1541	0.48	0/2057
All	All	0.29	0/3088	0.46	0/4124

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 Close contacts ⓘ

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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1513	0	1466	16	0
1	B	1511	0	1459	13	1
2	A	2	0	0	0	0
2	B	2	0	0	0	0
3	A	10	4	0	0	0
3	B	10	4	0	0	0
4	A	10	0	0	0	0
5	A	6	0	8	1	0
6	B	20	0	28	1	0
7	B	1	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
8	A	188	0	0	10	2
8	B	188	0	0	3	1
All	All	3461	8	2961	29	2

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 5.

All (29) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:17:GLU:HG3	8:A:483:HOH:O	1.68	0.92
1:A:143:ARG:NH2	8:A:475:HOH:O	1.96	0.92
1:B:43:GLN:NE2	8:B:341:HOH:O	2.09	0.80
1:A:17:GLU:CG	8:A:483:HOH:O	2.31	0.75
1:A:113:SER:OG	1:A:142:ARG:NH1	2.18	0.75
1:B:127:ARG:HD3	8:B:313:HOH:O	1.87	0.73
1:A:17:GLU:OE2	8:A:483:HOH:O	2.06	0.72
1:A:143:ARG:NH1	8:A:443:HOH:O	2.24	0.70
1:B:40:TYR:O	1:B:44[A]:ARG:HG3	1.92	0.70
1:B:40:TYR:CZ	1:B:44[A]:ARG:HD2	2.28	0.68
1:A:149:GLU:OE2	8:A:413:HOH:O	2.13	0.66
1:A:17:GLU:CD	8:A:483:HOH:O	2.38	0.60
1:A:175:GLU:OE2	8:A:365:HOH:O	2.16	0.60
1:A:41:PHE:CD1	5:A:206:GOL:H12	2.45	0.51
1:B:183:GLU:O	1:B:187:GLU:HG3	2.13	0.48
1:A:163:LYS:NZ	8:A:431:HOH:O	2.02	0.47
1:B:113:SER:OG	1:B:142:ARG:HD3	2.15	0.47
1:B:115:LYS:HG2	1:B:117:SER:OG	2.15	0.46
1:A:88:LEU:HD12	1:A:101:ILE:HD12	1.98	0.46
1:B:147:GLU:HG3	6:B:201:BOG:H7'1	1.99	0.45
1:B:104:HIS:O	1:B:158:ARG:HD3	2.17	0.44
1:A:163:LYS:CD	8:A:431:HOH:O	2.64	0.44
1:B:178:ARG:NH1	8:B:324:HOH:O	2.14	0.43
1:A:5:ASN:ND2	1:A:189:GLN:OE1	2.45	0.43
1:B:118:TYR:OH	1:B:162[A]:GLU:HG3	2.19	0.42
1:A:139:GLU:O	1:A:143:ARG:HG2	2.20	0.41
1:B:179:LYS:HE2	1:B:179:LYS:HB3	1.57	0.40
1:A:104:HIS:O	1:A:158:ARG:HD3	2.22	0.40
1:B:40:TYR:CE2	1:B:44[A]:ARG:HD2	2.56	0.40

All (2) 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	Distance(Å)	Clash(Å)
1:B:44[B]:ARG:NH2	8:A:448:HOH:O[2_755]	1.50	0.70
8:A:455:HOH:O	8:B:454:HOH:O[3_657]	1.95	0.25

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	190/196 (97%)	185 (97%)	5 (3%)	0	100	100
1	B	186/196 (95%)	182 (98%)	4 (2%)	0	100	100
All	All	376/392 (96%)	367 (98%)	9 (2%)	0	100	100

There are no Ramachandran outliers to report.

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	164/162 (101%)	163 (99%)	1 (1%)	92	87
1	B	163/162 (101%)	162 (99%)	1 (1%)	92	87
All	All	327/324 (101%)	325 (99%)	2 (1%)	92	87

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	162	GLU
1	B	179	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

5.3.3 RNA ⓘ

There are no RNA chains in this entry.

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

Mogul failed to run properly - this section will therefore be empty.

5.5 Carbohydrates ⓘ

Mogul failed to run properly - this section will therefore be empty.

5.6 Ligand geometry ⓘ

Mogul failed to run properly - this section will therefore be empty.

5.7 Other polymers ⓘ

Mogul failed to run properly - this section will therefore be empty.

5.8 Polymer linkage issues

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, 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	188/196 (95%)	-0.13	3 (1%)	68	74	11, 19, 35, 43	11 (5%)
1	B	187/196 (95%)	-0.10	7 (3%)	39	44	11, 19, 34, 55	10 (5%)
All	All	375/392 (95%)	-0.12	10 (2%)	50	57	11, 19, 35, 55	21 (5%)

All (10) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	189	GLN	3.6
1	B	168	ASP	3.4
1	A	116	ALA	3.1
1	B	27	GLU	3.1
1	B	28	ALA	3.0
1	B	117	SER	2.8
1	A	75	TYR	2.4
1	B	121	ASN	2.3
1	A	144	LEU	2.1
1	B	2	SER	2.1

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. 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, 95th 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	RSR	LLDF	B-factors(Å²)	Q<0.9
5	GOL	A	206	6/6	0.45	29.24	21,23,29,37	6
4	SO4	A	205	5/5	0.17	5.87	27,28,32,36	5
4	SO4	A	204	5/5	0.12	5.17	18,19,24,26	5
7	NA	B	205	1/1	0.29	3.27	40,40,40,40	0
6	BOG	B	201	20/20	0.08	-0.05	16,22,31,32	0
3	TLA	A	203	10/10	0.06	-0.96	13,15,18,19	0
3	TLA	B	204	10/10	0.06	-1.24	11,14,17,19	0
2	FE	B	203	1/1	0.05	-1.40	11,11,11,11	0
2	FE	B	202	1/1	0.06	-1.50	14,14,14,14	0
2	FE	A	202	1/1	0.06	-1.95	11,11,11,11	0
2	FE	A	201	1/1	0.04	-3.02	12,12,12,12	0

6.5 Other polymers ⓘ

There are no such residues in this entry.