W O R L D W I D E

wwPDB

PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

Jan 6, 2021 – 07:17 pm GMT

PDB ID	:	6YRJ
Title	:	SFX structure of dye-type peroxidase DtpB in the ferric state
Authors	:	Lucic, M.; Axford, D.A.; Owen, R.L.; Worrall, J.A.R.; Hough, M.A.
Deposited on	:	2020-04-20
Resolution	:	1.85 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*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

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 1.85 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.
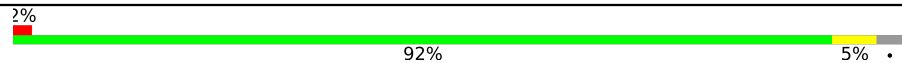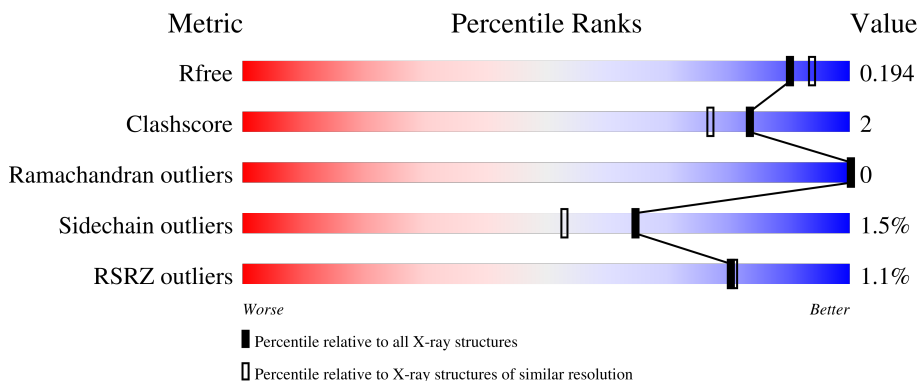

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	2469 (1.86-1.86)
Clashscore	141614	2625 (1.86-1.86)
Ramachandran outliers	138981	2592 (1.86-1.86)
Sidechain outliers	138945	2592 (1.86-1.86)
RSRZ outliers	127900	2436 (1.86-1.86)

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% 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	316	94%
1	B	316	91%
1	C	316	89%
1	D	316	90%
1	E	316	91%

Continued on next page...

Continued from previous page...

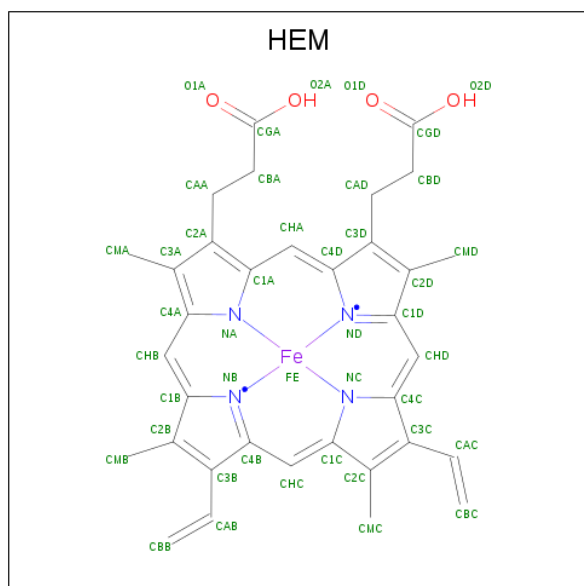| Mol	Chain	Length	Quality of chain
1	F	316	

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 15296 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 Putative iron-dependent peroxidase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	306	Total	C	N	O	S	0	2	0
			2351	1477	407	457	10			
1	B	306	Total	C	N	O	S	0	6	0
			2372	1494	408	461	9			
1	C	306	Total	C	N	O	S	0	9	0
			2399	1509	413	468	9			
1	D	304	Total	C	N	O	S	0	8	0
			2371	1492	408	462	9			
1	E	305	Total	C	N	O	S	0	7	0
			2370	1495	413	453	9			
1	F	306	Total	C	N	O	S	0	5	0
			2364	1485	409	460	10			

- Molecule 2 is PROTOPORPHYRIN IX CONTAINING FE (three-letter code: HEM) (formula: $C_{34}H_{32}FeN_4O_4$).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total 43	C 34	Fe 1	N 4	O 4	0	0
2	B	1	Total 43	C 34	Fe 1	N 4	O 4	0	0
2	C	1	Total 43	C 34	Fe 1	N 4	O 4	0	0
2	D	1	Total 43	C 34	Fe 1	N 4	O 4	0	0
2	E	1	Total 43	C 34	Fe 1	N 4	O 4	0	0
2	F	1	Total 43	C 34	Fe 1	N 4	O 4	0	0

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

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

- Molecule 4 is water.

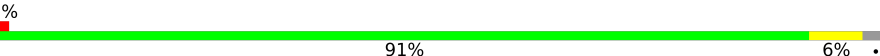| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	150	Total 150	O 150	0	0
4	B	124	Total 124	O 124	0	0
4	C	122	Total 122	O 122	0	0
4	D	147	Total 147	O 147	0	2
4	E	139	Total 139	O 139	0	0
4	F	128	Total 128	O 128	0	0

3 Residue-property plots ⓘ

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 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.
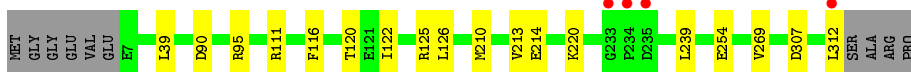
- Molecule 1: Putative iron-dependent peroxidase

Chain A:
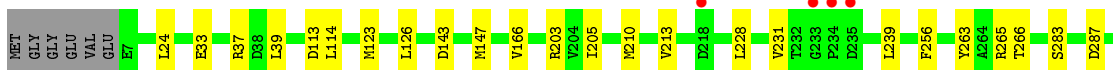

- Molecule 1: Putative iron-dependent peroxidase

Chain B:
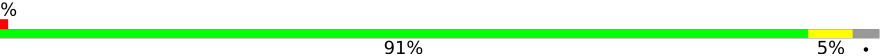

- Molecule 1: Putative iron-dependent peroxidase

Chain C:

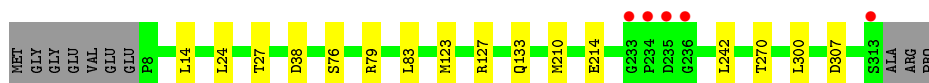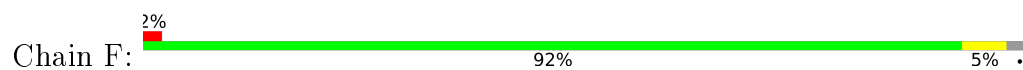- Molecule 1: Putative iron-dependent peroxidase

Chain D:

- Molecule 1: Putative iron-dependent peroxidase

Chain E:

• Molecule 1: Putative iron-dependent peroxidase

Chain F:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	86.00Å 120.40Å 197.50Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	13.00 – 1.85	Depositor
	102.80 – 1.85	EDS
% Data completeness	99.6 (13.00-1.85)	Depositor
(in resolution range)	99.9 (102.80-1.85)	EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$<I/\sigma(I)>$ [1]	1.69 (at 1.84Å)	Xtriage
Refinement program	REFMAC 5.8.0258	Depositor
R, R_{free}	0.159 , 0.192	Depositor
	0.163 , 0.194	DCC
R_{free} test set	1987 reflections (1.14%)	wwPDB-VP
Wilson B-factor (Å^2)	32.6	Xtriage
Anisotropy	0.042	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.37 , 47.1	EDS
L-test for twinning[2]	$<|L|> = 0.47$, $<L^2> = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	15296	wwPDB-VP
Average B, all atoms (Å^2)	36.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.76% 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: HEM, MG

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	0.76	0/2401	0.89	0/3260
1	B	0.70	0/2422	0.85	0/3293
1	C	0.73	0/2449	0.87	0/3329
1	D	0.72	0/2421	0.87	1/3289 (0.0%)
1	E	0.74	0/2420	0.88	1/3287 (0.0%)
1	F	0.72	1/2414 (0.0%)	0.87	0/3278
All	All	0.73	1/14527 (0.0%)	0.87	2/19736 (0.0%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	F	214	GLU	CD-OE2	-5.91	1.19	1.25

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	D	243	ARG	NE-CZ-NH2	-6.72	116.94	120.30
1	E	243	ARG	NE-CZ-NH2	-5.66	117.47	120.30

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	2351	0	2277	9	0
1	B	2372	0	2287	14	0
1	C	2399	0	2313	19	0
1	D	2371	0	2293	19	0
1	E	2370	0	2302	8	0
1	F	2364	0	2289	6	0
2	A	43	0	30	1	0
2	B	43	0	30	1	0
2	C	43	0	30	1	0
2	D	43	0	30	0	0
2	E	43	0	30	0	0
2	F	43	0	30	0	0
3	A	1	0	0	0	0
4	A	150	0	0	2	0
4	B	124	0	0	0	0
4	C	122	0	0	1	3
4	D	147	0	0	4	3
4	E	139	0	0	1	0
4	F	128	0	0	0	0
All	All	15296	0	13941	67	3

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 2.

The worst 5 of 67 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:39[B]:LEU:HD13	1:B:126:LEU:HG	1.60	0.83
1:D:231:VAL:HG23	1:D:239:LEU:HB2	1.63	0.79
1:D:113:ASP:OD2	4:D:501:HOH:O	2.03	0.75
1:B:214:GLU:OE2	1:B:220:LYS:NZ	2.19	0.71
1:C:37[B]:ARG:HH11	1:C:37[B]:ARG:HB3	1.56	0.71

All (3) 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	Interatomic distance (Å)	Clash overlap (Å)
4:C:615:HOH:O	4:D:644:HOH:O[4_545]	2.05	0.15
4:C:616:HOH:O	4:D:644:HOH:O[4_545]	2.11	0.09
4:C:517:HOH:O	4:D:644:HOH:O[4_545]	2.16	0.04

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	306/316 (97%)	301 (98%)	5 (2%)	0	100	100
1	B	310/316 (98%)	305 (98%)	5 (2%)	0	100	100
1	C	313/316 (99%)	308 (98%)	5 (2%)	0	100	100
1	D	310/316 (98%)	305 (98%)	5 (2%)	0	100	100
1	E	310/316 (98%)	306 (99%)	4 (1%)	0	100	100
1	F	309/316 (98%)	303 (98%)	6 (2%)	0	100	100
All	All	1858/1896 (98%)	1828 (98%)	30 (2%)	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	247/252 (98%)	244 (99%)	3 (1%)	71	62
1	B	247/252 (98%)	244 (99%)	3 (1%)	71	62
1	C	252/252 (100%)	248 (98%)	4 (2%)	62	49
1	D	250/252 (99%)	247 (99%)	3 (1%)	71	62
1	E	246/252 (98%)	243 (99%)	3 (1%)	71	62
1	F	249/252 (99%)	242 (97%)	7 (3%)	43	27
All	All	1491/1512 (99%)	1468 (98%)	23 (2%)	65	53

5 of 23 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	8	PRO
1	D	311	ASP
1	F	270[B]	THR
1	D	307	ASP
1	E	216	SER

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

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 7 ligands modelled in this entry, 1 is monoatomic - leaving 6 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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	306/316 (96%)	-0.15	0	100	100	23, 31, 57, 84	0
1	B	306/316 (96%)	-0.07	4 (1%)	77	78	25, 35, 62, 101	0
1	C	306/316 (96%)	-0.07	4 (1%)	77	78	24, 33, 62, 96	0
1	D	304/316 (96%)	-0.11	3 (0%)	82	82	23, 32, 59, 86	0
1	E	305/316 (96%)	-0.14	4 (1%)	77	78	24, 32, 57, 99	0
1	F	306/316 (96%)	-0.14	5 (1%)	72	72	24, 32, 59, 91	0
All	All	1833/1896 (96%)	-0.11	20 (1%)	80	81	23, 32, 60, 101	0

The worst 5 of 20 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	233	GLY	5.3
1	B	234	PRO	4.2
1	F	236	GLY	4.0
1	B	233	GLY	3.6
1	C	234	PRO	3.6

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 monosaccharides in this entry.

6.4 Ligands (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. 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	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	HEM	E	401	43/43	0.96	0.10	24,26,32,37	0
2	HEM	C	401	43/43	0.96	0.10	23,25,29,33	0
2	HEM	B	401	43/43	0.96	0.11	22,25,30,36	0
2	HEM	A	401	43/43	0.97	0.10	22,23,28,32	0
2	HEM	D	401	43/43	0.97	0.10	20,23,28,32	0
2	HEM	F	401	43/43	0.97	0.10	21,24,29,34	0
3	MG	A	402	1/1	0.99	0.05	25,25,25,25	0

6.5 Other polymers (i)

There are no such residues in this entry.