

wwPDB X-ray Structure Validation Summary Report (i)

May 14, 2020 – 01:10 pm BST

PDB ID : 3A5W

Title: Peroxiredoxin (wild type) from Aeropyrum pernix K1 (reduced form)

Authors: Nakamura, T.; Kado, Y.; Yamaguchi, T.; Matsumura, H.; Ishikawa, K.; Inoue,

Τ.

Deposited on : 2009-08-12

Resolution : 2.20 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.

The following versions of software and data (see references (1)) were used in the production of this report:

MolProbity : 4.02b-467 Xtriage (Phenix) : 1.13

EDS : 2.11

Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)

Refmac: 5.8.0158

CCP4 : 7.0.044 (Gargrove) roteins) : Engh & Huber (2001)

Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)

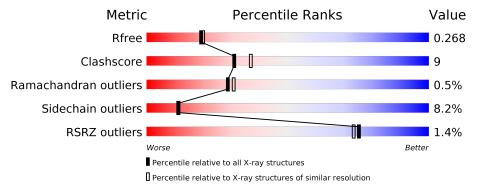
Validation Pipeline (wwPDB-VP) : 2.11

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 2.20 Å.

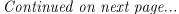
Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	$\begin{array}{c} \textbf{Whole archive} \\ (\#\text{Entries}) \end{array}$	$\begin{array}{c} {\rm Similar \; resolution} \\ (\#{\rm Entries, \; resolution \; range(\AA)}) \end{array}$
R_{free}	130704	4898 (2.20-2.20)
Clashscore	141614	5594 (2.20-2.20)
Ramachandran outliers	138981	5503 (2.20-2.20)
Sidechain outliers	138945	5504 (2.20-2.20)
RSRZ outliers	127900	4800 (2.20-2.20)

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% 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	249	71%	20%	5% •	-	
1	В	249	72%	20%	6%	-	
1	С	249	72%	20%	6%	-	
1	D	249	68%	25%	5%	-	
1	Е	249	63%	31%	•	-	
1	F	249	70%	22%		-	





Continued from previous page...

Mol	Chain	Length	Quality of chain		
1	G	249	67%	24%	5% • •
1	Н	249	72%	20%	5% •
1	I	249	67%	26%	5% • •
1	J	249	69%	21%	7% •



2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 19832 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 peroxiredoxin.

Mol	Chain	Residues		Ato	oms			ZeroOcc	AltConf	Trace
1	A	241	Total	С	N	О	S	0	0	0
1	Λ	241	1953	1257	344	345	7	0	U	
1	В	244	Total	$^{\mathrm{C}}$	N	Ο	S	0	0	
1	Ъ	244	1973	1268	347	351	7	U	U	U
1	C	244	Total	\mathbf{C}	N	Ο	S	0	0	
1		244	1973	1268	347	351	7	U	U	U
1	D	244	Total	$^{\mathrm{C}}$	N	Ο	S	0	0	0
	D	244	1973	1268	347	351	7	U	Ü	U
1	E	242	Total	\mathbf{C}	N	О	S	0	0	
1	L	242	1958	1260	345	346	7	· ·	Ü	0
1	F	244	Total	$^{\mathrm{C}}$	N	О	S	0	0	
	1	244	1973	1268	347	351	7	U	U	U
1	G	242	Total	$^{\mathrm{C}}$	N	О	S	0	0	
	G G	242	1958	1260	345	346	7	U	Ü	U
1	Н	243	Total	С	N	О	\mathbf{S}	0	0	
	11	240	1968	1265	346	350	7	U	Ü	U
1	I	244	Total	С	N	О	S	0	0	0
	1	244	1973	1268	347	351	7	U	U	U
1	J	241	Total	С	N	О	S	0	0	0
1	J	241	1953	1257	344	345	7		U	

• Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	16	Total O 16 16	0	0
2	В	17	Total O 17 17	0	0
2	С	23	Total O 23 23	0	0
2	D	20	Total O 20 20	0	0

Continued on next page...



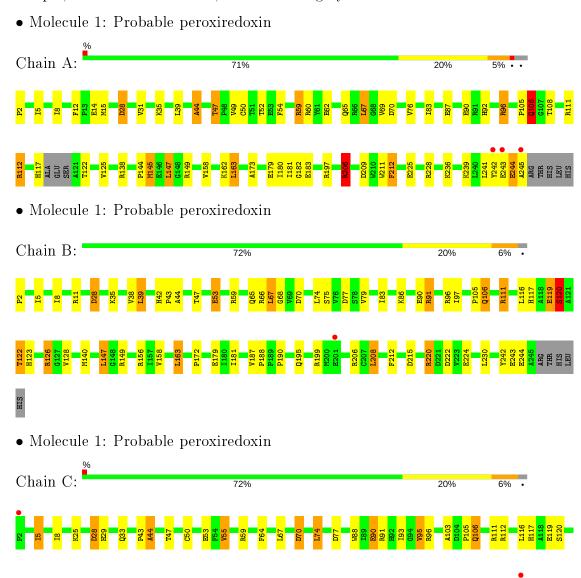
 $Continued\ from\ previous\ page...$

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	Е	21	Total O 21 21	0	0
2	F	18	Total O 18 18	0	0
2	G	11	Total O 11 11	0	0
2	Н	15	Total O 15 15	0	0
2	I	16	Total O 16 16	0	0
2	J	20	Total O 20 20	0	0



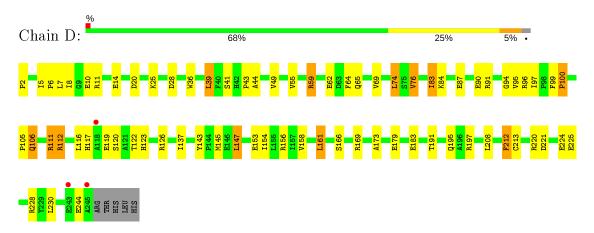
3 Residue-property plots (i)

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 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.

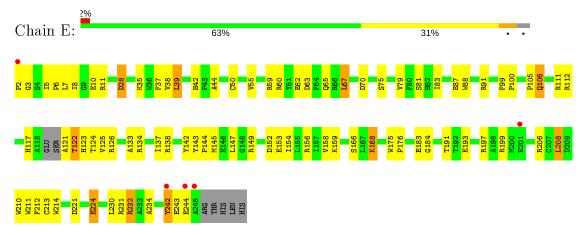


• Molecule 1: Probable peroxiredoxin

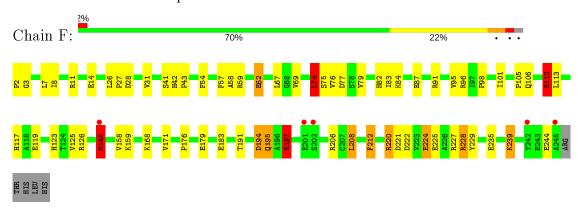




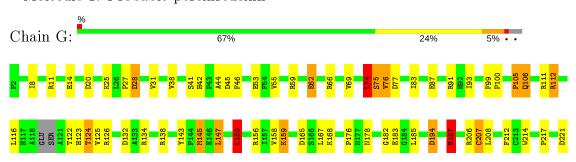
• Molecule 1: Probable peroxiredoxin



• Molecule 1: Probable peroxiredoxin



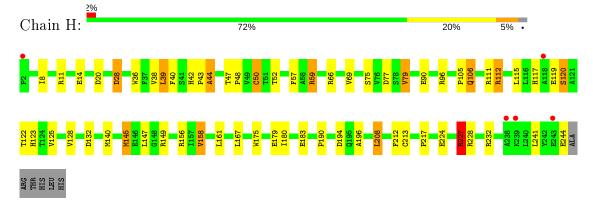
• Molecule 1: Probable peroxiredoxin



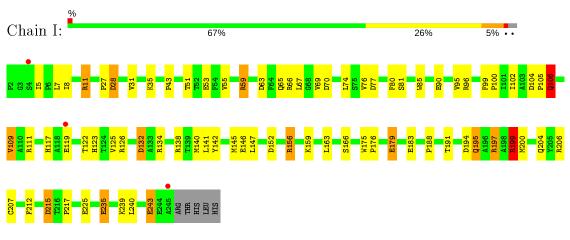




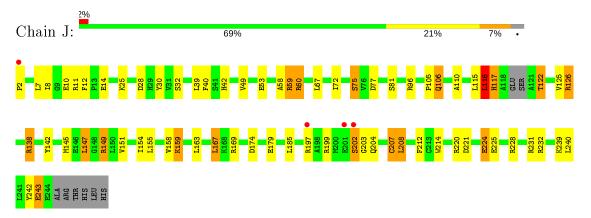
• Molecule 1: Probable peroxiredoxin



• Molecule 1: Probable peroxiredoxin



• Molecule 1: Probable peroxiredoxin





4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1	Depositor
Cell constants	77.61Å 104.78Å 106.17Å	Depositor
a, b, c, α , β , γ	105.91° 104.92° 92.88°	Depositor
Resolution (Å)	46.83 - 2.20	Depositor
resolution (A)	46.81 - 2.20	EDS
% Data completeness	94.3 (46.83-2.20)	Depositor
(in resolution range)	94.3 (46.81-2.20)	EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) > 1$	1.93 (at 2.20Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
P. P.	0.209 , 0.269	Depositor
R, R_{free}	0.209 , 0.268	DCC
R_{free} test set	7414 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å ²)	30.5	Xtriage
Anisotropy	0.015	Xtriage
Bulk solvent $k_{sol}(e/Å^3)$, $B_{sol}(Å^2)$	0.37, 51.5	EDS
L-test for twinning ²	$ < L >=0.49, < L^2>=0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	19832	wwPDB-VP
Average B, all atoms (Å ²)	31.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: The largest off-origin peak in the Patterson function is 9.27% of the height of the origin peak. No significant pseudotranslation is detected.

²Theoretical 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.



¹Intensities estimated from amplitudes.

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	Mol Chain		ond lengths	Е	ond angles
WIOI	Chain	RMSZ	# Z > 5	RMSZ	# Z >5
1	Α	1.58	20/2007~(1.0%)	1.19	$12/2727 \ (0.4\%)$
1	В	1.47	$15/2028 \; (0.7\%)$	1.25	$14/2757 \; (0.5\%)$
1	С	1.54	$17/2028 \; (0.8\%)$	1.25	$15/2757 \ (0.5\%)$
1	D	1.68	$26/2028 \; (1.3\%)$	1.24	$16/2757 \ (0.6\%)$
1	E	1.52	$23/2012 \ (1.1\%)$	1.24	$13/2734 \ (0.5\%)$
1	F	1.54	$18/2028 \; (0.9\%)$	1.21	$14/2757 \; (0.5\%)$
1	G	1.39	5/2012~(0.2%)	1.22	18/2734 (0.7%)
1	Н	1.42	$13/2023 \; (0.6\%)$	1.17	$11/2750 \ (0.4\%)$
1	I	1.59	$18/2028 \; (0.9\%)$	1.22	$13/2757 \ (0.5\%)$
1	J	1.57	18/2007~(0.9%)	1.30	18/2727 (0.7%)
All	All	1.53	$173/20201 \ (0.9\%)$	1.23	$144/27457 \ (0.5\%)$

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	Е	0	1

The worst 5 of 173 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	\mathbf{Z}	${f Observed(\AA)}$	$\operatorname{Ideal}(ext{\AA})$
1	С	50	CYS	CB-SG	10.41	2.00	1.82
1	E	88	TRP	CE3-CZ3	9.00	1.53	1.38
1	I	55	VAL	CB-CG2	8.72	1.71	1.52
1	D	119	GLU	CD-OE2	8.45	1.34	1.25
1	J	224	GLU	CB-CG	8.37	1.68	1.52

The worst 5 of 144 bond angle outliers are listed below:

		Type	Atoms		$ $ Observed $(^o)$	14341()
1 .	J 138	ARG	NE-CZ-NH2	-15.63	112.48	120.30

Continued on next page...



Continued from previous page...

Mol	Chain	${f Res}$	Type	Atoms	\mathbf{Z}	$Observed(^o)$	$\operatorname{Ideal}({}^{o})$
1	J	138	ARG	NE-CZ-NH1	11.86	126.23	120.30
1	G	208	LEU	CA-CB-CG	10.66	139.81	115.30
1	E	134	ARG	NE-CZ-NH1	10.49	125.55	120.30
1	E	149	ARG	NE-CZ-NH1	8.71	124.66	120.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	Е	121	ALA	Peptide

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)	$\mathbf{H}(\mathbf{added})$	Clashes	Symm-Clashes
1	A	1953	0	1942	48	0
1	В	1973	0	1957	39	0
1	С	1973	0	1959	30	0
1	D	1973	0	1959	44	0
1	Ε	1958	0	1945	36	0
1	F	1973	0	1957	38	0
1	G	1958	0	1945	42	0
1	Н	1968	0	1952	41	0
1	I	1973	0	1957	54	0
1	J	1953	0	1940	36	0
2	A	16	0	0	1	0
2	В	17	0	0	2	0
2	С	23	0	0	5	0
2	D	20	0	0	0	0
2	Ε	21	0	0	3	0
2	F	18	0	0	0	0
2	G	11	0	0	0	0
2	Н	15	0	0	2	0
2	I	16	0	0	4	0
2	J	20	0	0	3	0
All	All	19832	0	19513	356	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 9.

The worst 5 of 356 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	$egin{array}{ll} ext{Interatomic} \ ext{distance} & (ext{Å}) \end{array}$	$egin{array}{c} ext{Clash} \ ext{overlap } (ext{Å}) \end{array}$
1:E:137:ILE:CD1	1:E:137:ILE:CG1	1.77	1.62
1:D:97:ILE:CD1	1:D:97:ILE:CG1	1.77	1.56
1:D:100:PRO:CB	1:D:100:PRO:CG	1.75	1.40
2:C:268:HOH:O	1:E:83:ILE:HD11	1.42	1.14
1:I:11:ARG:HG2	1:I:11:ARG:HH11	1.11	1.09

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	Perce	ntiles
1	A	$237/249 \ (95\%)$	222 (94%)	13 (6%)	2 (1%)	19	19
1	В	242/249 (97%)	226 (93%)	15 (6%)	1 (0%)	34	37
1	С	242/249 (97%)	233 (96%)	8 (3%)	1 (0%)	34	37
1	D	242/249 (97%)	229 (95%)	12 (5%)	1 (0%)	34	37
1	Е	238/249 (96%)	226 (95%)	11 (5%)	1 (0%)	34	37
1	F	242/249 (97%)	228 (94%)	14 (6%)	0	100	100
1	G	238/249 (96%)	225 (94%)	11 (5%)	2 (1%)	19	19
1	Н	241/249 (97%)	230 (95%)	9 (4%)	2 (1%)	19	19
1	I	242/249 (97%)	234 (97%)	7 (3%)	1 (0%)	34	37
1	J	237/249 (95%)	225 (95%)	10 (4%)	2 (1%)	19	19
All	All	2401/2490 (96%)	2278 (95%)	110 (5%)	13 (0%)	29	31

5 of 13 Ramachandran outliers are listed below:



Mol	Chain	Res	Type
1	В	120	SER
1	D	120	SER
1	С	201	GLU
1	G	106	GLN
1	Н	106	GLN

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	$208/215 \; (97\%)$	190 (91%)	18 (9%)	10 10
1	В	210/215~(98%)	197 (94%)	13 (6%)	18 21
1	С	210/215~(98%)	192 (91%)	18 (9%)	10 10
1	D	210/215~(98%)	196 (93%)	14 (7%)	16 18
1	E	$208/215 \ (97\%)$	190 (91%)	18 (9%)	10 10
1	F	210/215~(98%)	190 (90%)	20 (10%)	8 8
1	G	$208/215 \; (97\%)$	188 (90%)	20 (10%)	8 8
1	Н	210/215~(98%)	198 (94%)	12 (6%)	20 24
1	I	210/215~(98%)	191 (91%)	19 (9%)	9 9
1	J	$208/215 \ (97\%)$	189 (91%)	19 (9%)	9 9
All	All	$2092/2150 \ (97\%)$	1921 (92%)	171 (8%)	11 11

5 of 171 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	208	LEU
1	F	208	LEU
1	J	116	LEU
1	E	224	GLU
1	F	79	VAL

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 27 such sidechains are listed below:



Mol	Chain	Res	Type
1	Е	123	HIS
1	F	117	HIS
1	J	92	HIS
1	F	92	HIS
1	С	117	HIS

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, 95^{th} 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>	#RSR	Z>2		$OWAB(m \AA^2)$	Q < 0.9
1	A	$241/249 \ (96\%)$	-0.24	3 (1%) 7	79 7	7	18, 28, 50, 61	0
1	В	244/249 (97%)	-0.33	1 (0%) 9	9	1	18, 30, 52, 65	0
1	С	244/249 (97%)	-0.42	2 (0%) 8	86 8	55	15, 27, 51, 66	0
1	D	244/249 (97%)	-0.30	3 (1%) 7	79 7	7	15, 26, 48, 63	0
1	E	242/249 (97%)	-0.42	5 (2%) 6	63	1	16, 28, 47, 64	0
1	F	244/249 (97%)	-0.21	5 (2%) 6	65	3	20, 31, 53, 70	0
1	G	$242/249 \ (97\%)$	-0.31	3 (1%) 7	79 7	7	21, 33, 53, 73	0
1	Н	243/249 (97%)	-0.35	5 (2%) 6	63	1	19, 32, 56, 68	0
1	I	244/249 (97%)	-0.29	3 (1%) 7	79 7	7	16, 29, 53, 58	0
1	J	241/249 (96%)	-0.40	4 (1%) 7	70 6	8	18, 28, 49, 64	0
All	All	2429/2490 (97%)	-0.33	34 (1%)	75	73	15, 29, 52, 73	0

The worst 5 of 34 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	F	245	ALA	4.0
1	Н	118	ALA	3.8
1	A	245	ALA	3.4
1	Н	238	ALA	3.4
1	I	245	ALA	3.3

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.

