



Full wwPDB X-ray Structure Validation Report ⓘ

Jan 31, 2016 – 08:34 PM GMT

PDB ID : 1KXM
Title : Crystal structure of Cytochrome c Peroxidase with a Proposed Electron Transfer Pathway Excised to Form a Ligand Binding Channel.
Authors : Rosenfeld, R.J.; Hayes, A.M.A.; Musah, R.A.; Goodin, D.B.
Deposited on : 2002-02-01
Resolution : 1.74 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
<http://wwpdb.org/validation/2016/XrayValidationReportHelp>
with specific help available everywhere you see the ⓘ symbol.

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

MolProbity : 4.02b-467
Mogul : 1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix) : **NOT EXECUTED**
EDS : **NOT EXECUTED**
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : trunk26865

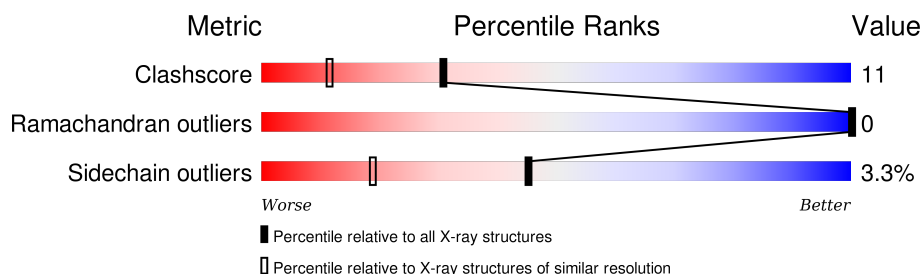
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.74 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))
Clashscore	102246	2570 (1.76-1.72)
Ramachandran outliers	100387	2544 (1.76-1.72)
Sidechain outliers	100360	2544 (1.76-1.72)

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 $\leq 5\%$. 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.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	290	

2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 2716 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 Cytochrome c Peroxidase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	290	Total	C	N	O	S	0	0	0
			2317	1477	385	449	6			

There are 4 discrepancies between the modelled and reference sequences:

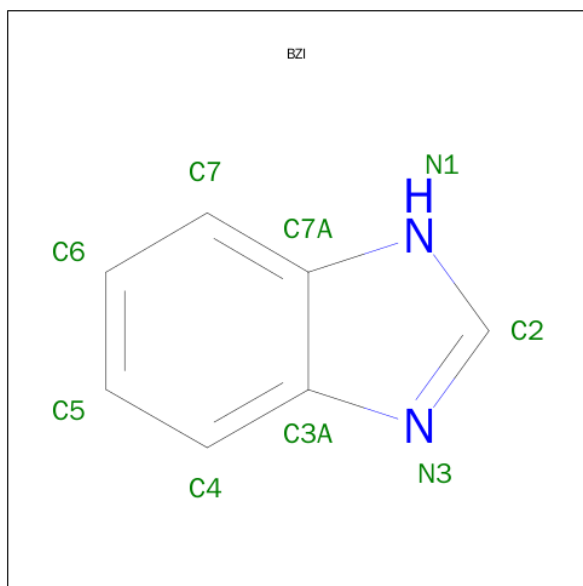
Chain	Residue	Modelled	Actual	Comment	Reference
A	3	THR	-	CLONING ARTIFACT	UNP P00431
A	190	GLY	PRO	ENGINEERED	UNP P00431
A	?	-	TRP	DELETION	UNP P00431
A	?	-	ALA	DELETION	UNP P00431

- 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	C	Fe	N	O	0	0
			43	34	1	4	4		

- Molecule 3 is BENZIMIDAZOLE (three-letter code: BZI) (formula: $C_7H_6N_2$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	N	0	0
			9	7	2		

- Molecule 4 is water.

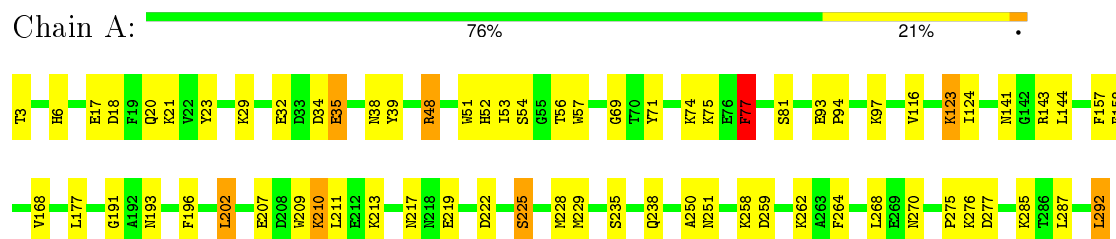
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	347	Total	O	0	0
			347	347		

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 ($\text{RSRZ} > 2$). 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.

Note EDS was not executed.

• Molecule 1: Cytochrome c Peroxidase



4 Data and refinement statistics

Xtriage (Phenix) and EDS were not executed - this section will therefore be incomplete.

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	106.50 Å 75.00 Å 50.90 Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	10.00 – 1.74	Depositor
% Data completeness (in resolution range)	77.8 (10.00-1.74)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	0.05	Depositor
Refinement program	SHELXL-96	Depositor
R, R_{free}	0.194 , 0.214	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	2716	wwPDB-VP
Average B, all atoms (Å ²)	16.0	wwPDB-VP

5 Model quality ⓘ

5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: HEM, BZI

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	1.18	2/2379 (0.1%)	1.07	10/3222 (0.3%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	292	LEU	C-OXT	41.53	2.02	1.23
1	A	123	LYS	CE-NZ	34.47	2.35	1.49

All (10) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	276	LYS	C-N-CA	13.77	156.13	121.70
1	A	48	ARG	NE-CZ-NH2	10.34	125.47	120.30
1	A	48	ARG	NE-CZ-NH1	-7.49	116.56	120.30
1	A	23	TYR	CB-CG-CD2	7.24	125.35	121.00
1	A	23	TYR	CG-CD2-CE2	7.05	126.94	121.30
1	A	157	PHE	CB-CG-CD1	-6.10	116.53	120.80
1	A	48	ARG	CD-NE-CZ	6.09	132.12	123.60
1	A	191	GLY	C-N-CA	5.75	136.08	121.70
1	A	123	LYS	CD-CE-NZ	-5.29	99.54	111.70
1	A	77	PHE	CA-CB-CG	5.16	126.28	113.90

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-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 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	2317	0	2164	49	1
2	A	43	0	30	1	0
3	A	9	0	6	1	0
4	A	347	0	0	7	2
All	All	2716	0	2200	50	2

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 11.

All (50) 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:141:ASN:OD1	4:A:712:HOH:O	1.77	1.01
1:A:292:LEU:OXT	1:A:292:LEU:O	1.77	1.00
1:A:292:LEU:OXT	1:A:292:LEU:C	2.02	0.98
1:A:20:GLN:HE22	1:A:285:LYS:H	1.14	0.94
1:A:123:LYS:NZ	1:A:123:LYS:CE	2.35	0.90
1:A:210:LYS:HE2	1:A:222:ASP:O	1.80	0.82
1:A:287:LEU:HB3	1:A:292:LEU:OXT	1.88	0.74
1:A:116:VAL:HG11	1:A:124:ILE:HD11	1.78	0.66
1:A:69:GLY:HA2	4:A:499:HOH:O	1.97	0.63
1:A:20:GLN:HE22	1:A:285:LYS:N	1.92	0.62
1:A:53:ILE:HG22	1:A:71:TYR:HB2	1.86	0.57
1:A:258:LYS:HG2	1:A:262:LYS:HE2	1.86	0.56
1:A:17:GLU:O	1:A:21:LYS:HG3	2.07	0.55
1:A:6:HIS:HD2	4:A:493:HOH:O	1.89	0.54
1:A:81:SER:HA	4:A:785:HOH:O	2.07	0.54
1:A:235:SER:HA	1:A:238:GLN:HE21	1.73	0.52
1:A:202:LEU:HD12	1:A:250:ALA:HB1	1.92	0.52
1:A:93:GLU:O	1:A:97:LYS:HG2	2.09	0.51
1:A:18:ASP:HA	1:A:21:LYS:HE3	1.93	0.51
1:A:93:GLU:HG3	1:A:97:LYS:HZ3	1.77	0.49
1:A:207:GLU:HB2	1:A:209:TRP:CE2	2.48	0.49
1:A:34:ASP:O	1:A:39:TYR:HA	2.12	0.49
1:A:29:LYS:HE2	1:A:94:PRO:CG	2.43	0.48
2:A:1:HEM:CHA	3:A:2:BZI:H6	2.44	0.48
1:A:259:ASP:HA	1:A:262:LYS:HE3	1.96	0.47
1:A:207:GLU:OE1	1:A:225:SER:OG	2.30	0.47
1:A:54:SER:OG	1:A:268:LEU:HD23	2.14	0.47

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:48:ARG:NH1	4:A:434:HOH:O	2.49	0.46
1:A:228:MET:HG2	1:A:229:MET:N	2.33	0.44
1:A:38:ASN:O	1:A:39:TYR:HB2	2.18	0.44
1:A:52:HIS:ND1	1:A:144:LEU:HD13	2.33	0.44
1:A:29:LYS:HE2	1:A:94:PRO:HG3	2.00	0.43
1:A:211:LEU:HD11	1:A:219:GLU:HB3	2.00	0.43
1:A:285:LYS:HD3	4:A:761:HOH:O	2.18	0.43
1:A:56:THR:O	1:A:143:ARG:HD3	2.19	0.43
1:A:93:GLU:HG3	1:A:97:LYS:NZ	2.34	0.43
1:A:158:PHE:CG	1:A:168:VAL:HG22	2.54	0.43
1:A:32:GLU:OE2	1:A:32:GLU:HA	2.17	0.43
1:A:213:LYS:NZ	1:A:217:ASN:HD21	2.16	0.43
1:A:124:ILE:HD13	1:A:264:PHE:HD1	1.83	0.43
1:A:57:TRP:CD2	1:A:270:ASN:HB3	2.54	0.42
1:A:75:LYS:O	4:A:712:HOH:O	2.21	0.42
1:A:20:GLN:NE2	1:A:285:LYS:H	1.97	0.42
1:A:210:LYS:HB2	1:A:210:LYS:HE3	1.24	0.42
1:A:177:LEU:HD11	1:A:196:PHE:HD2	1.86	0.41
1:A:124:ILE:HD13	1:A:264:PHE:CD1	2.56	0.41
1:A:202:LEU:HD11	1:A:250:ALA:O	2.21	0.41
1:A:74:LYS:HA	1:A:77:PHE:CE2	2.56	0.41
1:A:275:PRO:HB2	1:A:277:ASP:HB3	2.03	0.40
1:A:35:GLU:H	1:A:35:GLU:CD	2.24	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	Interatomic distance (Å)	Clash overlap (Å)
4:A:706:HOH:O	4:A:821:HOH:O[2_574]	1.97	0.23
1:A:251:ASN:OD1	4:A:723:HOH:O[2_574]	2.19	0.01

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	288/290 (99%)	285 (99%)	3 (1%)	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	241/248 (97%)	233 (97%)	8 (3%)	45	19

All (8) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	3	THR
1	A	35	GLU
1	A	51	TRP
1	A	77	PHE
1	A	193	ASN
1	A	202	LEU
1	A	210	LYS
1	A	225	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (13) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	6	HIS
1	A	20	GLN
1	A	24	ASN
1	A	78	ASN
1	A	87	ASN
1	A	120	GLN
1	A	159	GLN
1	A	193	ASN
1	A	206	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	217	ASN
1	A	218	ASN
1	A	238	GLN
1	A	290	GLN

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](#)

2 ligands are modelled in this entry.

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	HEM	A	1	1,4	30,50,50	3.74	13 (43%)	24,82,82	2.48	10 (41%)
3	BZI	A	2	-	8,10,10	0.67	0	6,13,13	1.00	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	HEM	A	1	1,4	-	0/10/54/54	0/0/8/8
3	BZI	A	2	-	-	0/0/0/0	0/2/2/2

All (13) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	1	HEM	C3B-C4B	-11.92	1.41	1.51
2	A	1	HEM	C3D-C4D	-8.87	1.40	1.51
2	A	1	HEM	C2C-C1C	-6.58	1.40	1.52
2	A	1	HEM	C3B-CAB	-5.75	1.40	1.51
2	A	1	HEM	C3C-CAC	-5.33	1.41	1.51
2	A	1	HEM	C2D-C3D	-4.96	1.39	1.54
2	A	1	HEM	C2D-C1D	-3.91	1.39	1.51
2	A	1	HEM	C2B-C1B	-3.35	1.40	1.51
2	A	1	HEM	FE-ND	2.04	2.08	1.97
2	A	1	HEM	FE-NB	2.45	2.10	1.97
2	A	1	HEM	CBB-CAB	2.48	1.43	1.29
2	A	1	HEM	CHC-C1C	2.63	1.42	1.36
2	A	1	HEM	CBC-CAC	2.65	1.44	1.29

All (10) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	1	HEM	CAA-C2A-C1A	-4.01	122.65	127.01
2	A	1	HEM	C3B-CAB-CBB	2.50	128.30	124.46
2	A	1	HEM	CMD-C2D-C3D	2.58	125.76	114.35
2	A	1	HEM	C3C-CAC-CBC	2.61	128.45	124.46
2	A	1	HEM	C2D-C3D-C4D	2.72	106.12	101.50
2	A	1	HEM	C1D-CHD-C4C	2.77	130.45	125.82
2	A	1	HEM	CAD-C3D-C2D	3.72	123.92	113.22
2	A	1	HEM	CMC-C2C-C3C	4.61	128.04	116.53
2	A	1	HEM	CAD-C3D-C4D	4.95	129.91	112.47
2	A	1	HEM	CMB-C2B-C3B	5.29	129.74	116.53

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	1	HEM	1	0
3	A	2	BZI	1	0

5.7 Other polymers

There are no such residues in this entry.

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 ⓘ

EDS was not executed - this section will therefore be empty.

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

EDS was not executed - this section will therefore be empty.

6.3 Carbohydrates ⓘ

EDS was not executed - this section will therefore be empty.

6.4 Ligands ⓘ

EDS was not executed - this section will therefore be empty.

6.5 Other polymers ⓘ

EDS was not executed - this section will therefore be empty.