



Full wwPDB X-ray Structure Validation Report ⓘ

Mar 9, 2018 – 08:21 am GMT

PDB ID : 1AC4
Title : VARIATION IN THE STRENGTH OF A CH TO O HYDROGEN BOND IN AN ARTIFICIAL PROTEIN CAVITY (2,3,4-TRIMETHYL-1,3-THIAZOLE)
Authors : Musah, R.A.; Jensen, G.M.; Bunte, S.W.; Rosenfeld, R.; Mcree, D.E.; Goodin, D.B.
Deposited on : 1997-02-12
Resolution : 2.10 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

<https://www.wwpdb.org/validation/2017/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 : 1.7.3 (157068), CSD as539be (2018)
Xtriage (Phenix) : 1.13
EDS : trunk30967
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)
Refmac : 5.8.0158
CCP4 : 7.0 (Gargrove)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : trunk30967

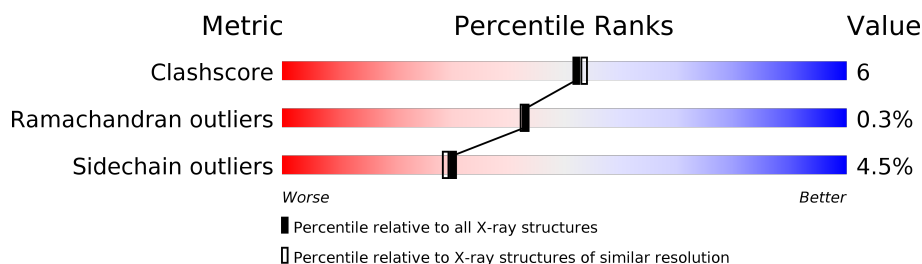
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.10 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	122126	5109 (2.10-2.10)
Ramachandran outliers	120053	5059 (2.10-2.10)
Sidechain outliers	120020	5060 (2.10-2.10)

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 $\leq 5\%$.

Mol	Chain	Length	Quality of chain
1	A	294	

2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 2941 atoms, of which 503 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.

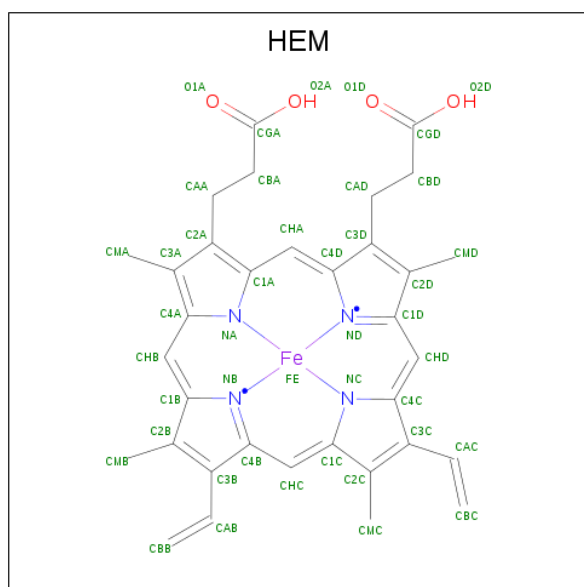
- Molecule 1 is a protein called CYTOCHROME C PEROXIDASE.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	291	Total	C	H	N	O	S	0	0	0
			2841	1492	503	388	452	6			

There are 4 discrepancies between the modelled and reference sequences:

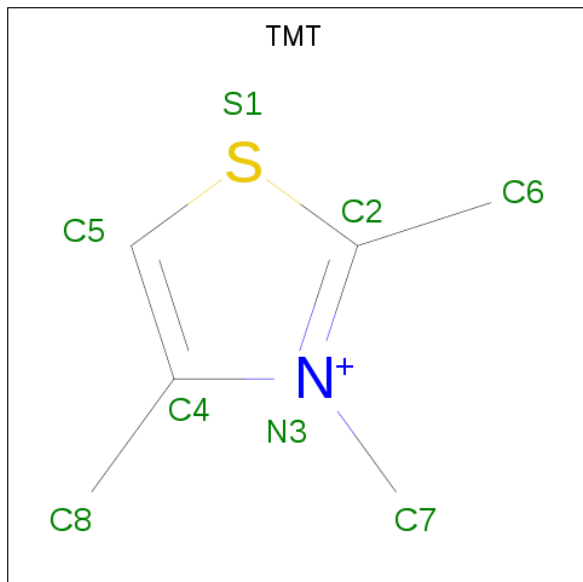
Chain	Residue	Modelled	Actual	Comment	Reference
A	53	ILE	THR	SEE REMARK 999	UNP P00431
A	152	GLY	ASP	SEE REMARK 999	UNP P00431
A	191	GLY	TRP	ENGINEERED	UNP P00431
A	272	ASP	ASN	SEE REMARK 999	UNP P00431

- 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 2,3,4-TRIMETHYL-1,3-THIAZOLE (three-letter code: TMT) (formula: $C_6H_{10}NS$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	N	S	0	0
			8	6	1	1		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	49	Total	O	0	0
			49	49		

i

- Molecule 1: CYTOCHROME C PEROXIDASE

[illegible]

4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	105.20Å 74.30Å 45.40Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	7.00 – 2.10 21.47 – 1.95	Depositor EDS
% Data completeness (in resolution range)	79.0 (7.00-2.10) 70.7 (21.47-1.95)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	0.05	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.68 (at 1.95Å)	Xtriage
Refinement program	XTALVIEW	Depositor
R, R_{free}	(Not available) , (Not available) 0.201 , (Not available)	Depositor DCC
R_{free} test set	No test flags present.	wwPDB-VP
Wilson B-factor (Å ²)	18.8	Xtriage
Anisotropy	0.676	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.36 , 58.4	EDS
L-test for twinning ²	$\langle L \rangle = 0.47$, $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	2941	wwPDB-VP
Average B, all atoms (Å ²)	14.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.12% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

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

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	0.90	0/2402	1.62	37/3250 (1.1%)

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	A	0	1

There are no bond length outliers.

All (37) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	31	ARG	NE-CZ-NH2	-13.13	113.73	120.30
1	A	48	ARG	NE-CZ-NH1	10.81	125.70	120.30
1	A	163	MET	CG-SD-CE	-10.26	83.78	100.20
1	A	127	ARG	NE-CZ-NH2	-9.80	115.40	120.30
1	A	143	ARG	NE-CZ-NH2	-9.49	115.56	120.30
1	A	31	ARG	NE-CZ-NH1	9.44	125.02	120.30
1	A	229	TYR	CB-CG-CD2	-8.86	115.69	121.00
1	A	223	TRP	CD1-CG-CD2	8.62	113.19	106.30
1	A	127	ARG	NE-CZ-NH1	8.49	124.55	120.30
1	A	51	TRP	CD1-CG-CD2	8.22	112.88	106.30
1	A	223	TRP	CE2-CD2-CG	-7.95	100.94	107.30
1	A	101	TRP	CD1-CG-CD2	7.65	112.42	106.30
1	A	126	TRP	CE2-CD2-CG	-7.57	101.24	107.30
1	A	57	TRP	CE2-CD2-CG	-7.53	101.28	107.30
1	A	162	ASN	CA-C-N	-7.11	101.56	117.20
1	A	51	TRP	CE2-CD2-CG	-7.07	101.64	107.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	101	TRP	CE2-CD2-CG	-7.01	101.69	107.30
1	A	126	TRP	CD1-CG-CD2	6.90	111.82	106.30
1	A	278	LYS	CA-CB-CG	-6.80	98.45	113.40
1	A	211	TRP	CD1-CG-CD2	6.45	111.46	106.30
1	A	57	TRP	CD1-CG-CD2	6.45	111.46	106.30
1	A	211	TRP	CE2-CD2-CG	-6.21	102.33	107.30
1	A	277	PRO	CA-C-N	5.91	130.21	117.20
1	A	162	ASN	N-CA-C	5.70	126.38	111.00
1	A	143	ARG	NE-CZ-NH1	5.55	123.07	120.30
1	A	51	TRP	CG-CD1-NE1	-5.49	104.61	110.10
1	A	101	TRP	CB-CG-CD1	-5.35	120.04	127.00
1	A	51	TRP	CB-CG-CD1	-5.24	120.19	127.00
1	A	285	ILE	O-C-N	-5.20	114.38	122.70
1	A	101	TRP	CG-CD2-CE3	5.19	138.57	133.90
1	A	57	TRP	CG-CD2-CE3	5.18	138.56	133.90
1	A	71	TYR	CB-CG-CD2	-5.17	117.90	121.00
1	A	101	TRP	CG-CD1-NE1	-5.17	104.93	110.10
1	A	16	TYR	CB-CG-CD2	-5.16	117.91	121.00
1	A	130	ARG	CG-CD-NE	-5.11	101.07	111.80
1	A	162	ASN	CA-C-O	5.07	130.75	120.10
1	A	223	TRP	CG-CD1-NE1	-5.05	105.05	110.10

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	72	ARG	Sidechain

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	2338	503	2209	27	0
2	A	43	0	30	1	0
3	A	8	0	10	3	0
4	A	49	0	0	3	0
All	All	2438	503	2249	28	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 6.

All (28) 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:195:ASN:HD22	1:A:195:ASN:H	1.41	0.69
1:A:74:LYS:H	1:A:74:LYS:HD2	1.61	0.66
1:A:20:GLN:HE22	1:A:287:LYS:H	1.47	0.62
2:A:295:HEM:HAD2	3:A:296:TMT:H63	1.82	0.62
1:A:235:ASP:OD2	3:A:296:TMT:H5	2.01	0.59
1:A:200:ASN:H	1:A:255:GLN:HE21	1.51	0.57
1:A:75:LYS:HE2	1:A:138:THR:HG22	1.86	0.56
1:A:166:ARG:HH21	1:A:257:LYS:HZ2	1.55	0.54
1:A:20:GLN:HE22	1:A:287:LYS:N	2.05	0.54
1:A:164:ASN:O	1:A:168:VAL:HG23	2.12	0.49
1:A:206:LEU:HD13	1:A:231:MET:SD	2.52	0.49
1:A:125:PRO:HG3	1:A:285:ILE:CD1	2.43	0.49
1:A:166:ARG:HH21	1:A:257:LYS:NZ	2.09	0.49
1:A:74:LYS:H	1:A:74:LYS:CD	2.26	0.47
1:A:147:ALA:O	1:A:234:THR:HG23	2.15	0.47
1:A:236:TYR:CE1	1:A:239:ILE:HD11	2.50	0.46
1:A:195:ASN:ND2	1:A:195:ASN:H	2.12	0.44
1:A:218:ALA:O	1:A:219:ASN:HB2	2.18	0.43
1:A:48:ARG:NH2	4:A:300:HOH:O	2.51	0.43
1:A:130:ARG:NE	4:A:326:HOH:O	2.53	0.42
1:A:72:ARG:NH1	1:A:73:PHE:CZ	2.88	0.42
1:A:130:ARG:NH2	4:A:326:HOH:O	2.53	0.42
1:A:195:ASN:HD22	1:A:195:ASN:N	2.14	0.41
1:A:4:LEU:HB2	1:A:62:ASN:HB3	2.02	0.41
1:A:179:LYS:C	3:A:296:TMT:H62	2.41	0.41
1:A:23:TYR:CD1	1:A:23:TYR:C	2.94	0.40
1:A:199:THR:HB	1:A:255:GLN:NE2	2.37	0.40
1:A:20:GLN:NE2	1:A:287:LYS:H	2.16	0.40

There are no symmetry-related clashes.

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	289/294 (98%)	278 (96%)	10 (4%)	1 (0%)	43 43

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	162	ASN

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	246/251 (98%)	235 (96%)	11 (4%)	30 29

All (11) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	23	TYR
1	A	28	LEU
1	A	35	GLU
1	A	51	TRP
1	A	74	LYS
1	A	80	PRO
1	A	159	GLN
1	A	166	ARG
1	A	171	LEU
1	A	195	ASN
1	A	279	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (8) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	20	GLN
1	A	24	ASN
1	A	87	ASN
1	A	195	ASN
1	A	208	ASN
1	A	240	GLN
1	A	255	GLN
1	A	292	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	295	1,4	27,50,50	1.41	4 (14%)	17,82,82	1.80	3 (17%)
3	TMT	A	296	-	5,8,8	0.81	0	4,11,11	1.95	2 (50%)

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	295	1,4	-	0/6/54/54	0/0/8/8
3	TMT	A	296	-	-	0/0/0/0	0/1/1/1

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	295	HEM	C3B-CAB	-3.93	1.40	1.47
2	A	295	HEM	C3C-CAC	-3.44	1.40	1.47
2	A	295	HEM	CBB-CAB	2.07	1.43	1.29
2	A	295	HEM	CBC-CAC	2.10	1.43	1.29

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	296	TMT	C7-N3-C4	-2.66	121.49	124.09
3	A	296	TMT	C6-C2-S1	2.15	123.26	120.28
2	A	295	HEM	CMC-C2C-C3C	3.02	130.38	124.88
2	A	295	HEM	CMB-C2B-C3B	3.86	131.91	124.88
2	A	295	HEM	CAD-CBD-CGD	4.01	119.51	112.66

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	295	HEM	1	0
3	A	296	TMT	3	0

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](#)

Unable to reproduce the depositors R factor - this section is therefore empty.

6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates [i](#)

Unable to reproduce the depositors R factor - this section is therefore empty.

6.4 Ligands [i](#)

Unable to reproduce the depositors R factor - this section is therefore empty.

6.5 Other polymers [i](#)

Unable to reproduce the depositors R factor - this section is therefore empty.