



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 12, 2017 – 08:32 pm GMT

PDB ID : 3KA6  
Title : Frog M-ferritin, EED mutant, with cobalt  
Authors : Tosha, T.; Ng, H.L.; Theil, E.; Alber, T.; Bhattasali, O.  
Deposited on : 2009-10-18  
Resolution : 1.40 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](mailto: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  
Xtriage (Phenix) : 1.9-1692  
EDS : trunk28620  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : recalc28949

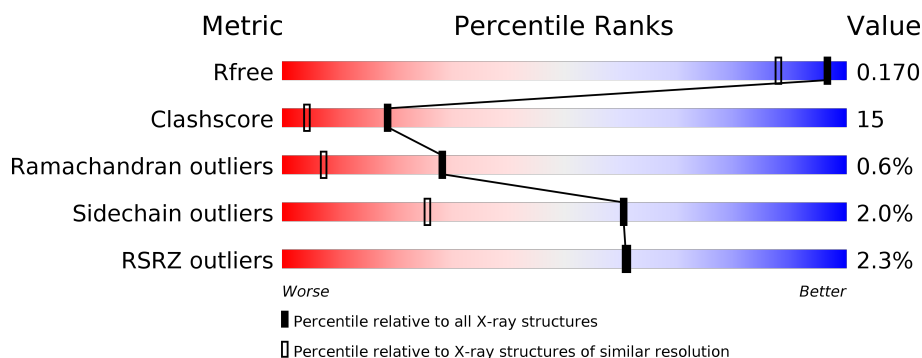
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.40 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))
$R_{free}$	100719	1307 (1.40-1.40)
Clashscore	112137	1411 (1.40-1.40)
Ramachandran outliers	110173	1373 (1.40-1.40)
Sidechain outliers	110143	1372 (1.40-1.40)
RSRZ outliers	101464	1315 (1.40-1.40)

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  $\geq 3$ , 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.

Mol	Chain	Length	Quality of chain
1	A	176	<div> <div>2%</div> <div> <div></div> <div>75%</div> <div>19%</div> <div>...</div> </div> </div>

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 1838 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 Ferritin, middle subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	172	Total	C	N	O	S	0	14	0
			1497	948	258	284	7			

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	137	GLU	GLN	ENGINEERED	UNP P07798

- Molecule 2 is COBALT (II) ION (three-letter code: CO) (formula: Co).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	10	Total	Co	0	0
			10	10		

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

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	5	Total	Mg	0	0
			5	5		

- Molecule 4 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	4	Total	Cl	0	0
			4	4		

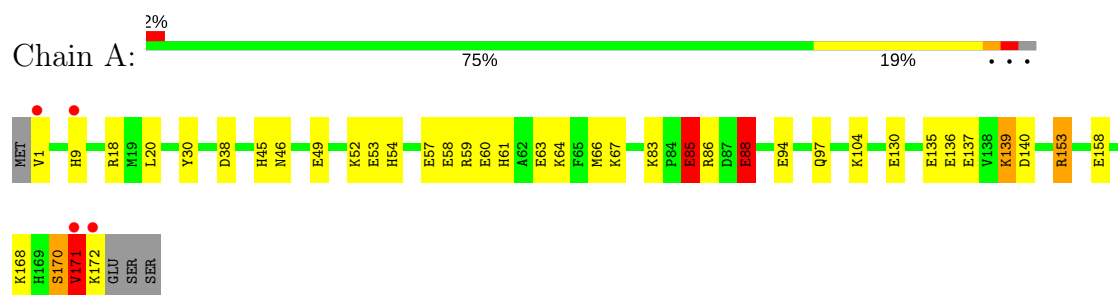
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	322	Total	O	0	0
			322	322		

### 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: Ferritin, middle subunit



## 4 Data and refinement statistics

Property	Value	Source
Space group	F 4 3 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	183.71Å 183.71Å 183.71Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.00 – 1.40 64.95 – 1.40	Depositor EDS
% Data completeness (in resolution range)	100.0 (50.00-1.40) 100.0 (64.95-1.40)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	0.07	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.66 (at 1.40Å)	Xtriage
Refinement program	REFMAC 5.4.0077	Depositor
R, $R_{free}$	0.139 , 0.166 0.142 , 0.170	Depositor DCC
$R_{free}$ test set	2678 reflections (5.37%)	DCC
Wilson B-factor (Å <sup>2</sup> )	14.6	Xtriage
Anisotropy	0.000	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 46.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	1838	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	16.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.05% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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: MG, CO, CL

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	1.31	20/1571 (1.3%)	1.13	12/2112 (0.6%)

All (20) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	88[A]	GLU	CD-OE1	11.26	1.38	1.25
1	A	88[B]	GLU	CD-OE1	11.26	1.38	1.25
1	A	88[A]	GLU	CG-CD	9.87	1.66	1.51
1	A	88[B]	GLU	CG-CD	9.87	1.66	1.51
1	A	85	GLU	CG-CD	9.05	1.65	1.51
1	A	18	ARG	CZ-NH2	-7.34	1.23	1.33
1	A	137	GLU	CB-CG	-6.71	1.39	1.52
1	A	63	GLU	CD-OE2	6.38	1.32	1.25
1	A	83	LYS	CE-NZ	-6.23	1.33	1.49
1	A	58	GLU	CB-CG	-5.87	1.41	1.52
1	A	64	LYS	CG-CD	5.81	1.72	1.52
1	A	60	GLU	CD-OE2	5.81	1.32	1.25
1	A	139	LYS	CD-CE	5.67	1.65	1.51
1	A	64	LYS	CB-CG	-5.57	1.37	1.52
1	A	94	GLU	CB-CG	-5.54	1.41	1.52
1	A	83	LYS	CD-CE	-5.33	1.38	1.51
1	A	137	GLU	CD-OE2	-5.21	1.20	1.25
1	A	85	GLU	CD-OE1	5.21	1.31	1.25
1	A	139	LYS	CE-NZ	5.10	1.61	1.49
1	A	130	GLU	CD-OE2	-5.00	1.20	1.25

All (12) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	104	LYS	CD-CE-NZ	-11.05	86.28	111.70

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	153	ARG	NE-CZ-NH2	-9.49	115.56	120.30
1	A	139	LYS	CD-CE-NZ	7.87	129.80	111.70
1	A	153	ARG	NE-CZ-NH1	7.29	123.95	120.30
1	A	170	SER	O-C-N	-6.80	111.82	122.70
1	A	88[A]	GLU	CG-CD-OE1	6.25	130.81	118.30
1	A	88[B]	GLU	CG-CD-OE1	6.25	130.81	118.30
1	A	171	VAL	CA-C-N	-5.67	104.73	117.20
1	A	171	VAL	CA-CB-CG2	5.30	118.86	110.90
1	A	64	LYS	CB-CG-CD	-5.16	98.20	111.60
1	A	88[A]	GLU	CG-CD-OE2	-5.01	108.28	118.30
1	A	88[B]	GLU	CG-CD-OE2	-5.01	108.28	118.30

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	1497	0	1466	44	4
2	A	10	0	0	0	0
3	A	5	0	0	0	0
4	A	4	0	0	0	0
5	A	322	0	0	35	11
All	All	1838	0	1466	44	11

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 15.

All (44) 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:9[B]:HIS:HE1	5:A:328:HOH:O	1.20	1.22
1:A:54[B]:HIS:HE1	1:A:140[B]:ASP:OD2	1.20	1.21
1:A:97:GLN:HG2	5:A:514:HOH:O	1.39	1.20
1:A:153:ARG:HD2	5:A:206:HOH:O	1.46	1.13

Continued on next page...

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:54[B]:HIS:CE1	1:A:140[B]:ASP:OD2	2.12	1.03
1:A:136[B]:GLU:OE1	5:A:419:HOH:O	1.78	1.02
1:A:45[B]:HIS:ND1	5:A:204:HOH:O	1.91	1.02
1:A:61[A]:HIS:HE1	5:A:347:HOH:O	1.42	1.00
1:A:136[B]:GLU:HG2	5:A:345:HOH:O	1.63	0.98
1:A:61[B]:HIS:CE1	5:A:501:HOH:O	2.20	0.94
1:A:136[B]:GLU:OE2	5:A:322:HOH:O	1.95	0.85
1:A:52:LYS:HE3	5:A:232:HOH:O	1.77	0.84
1:A:45[B]:HIS:CE1	5:A:204:HOH:O	2.31	0.83
1:A:45[A]:HIS:ND1	5:A:204:HOH:O	2.12	0.82
1:A:61[A]:HIS:CE1	5:A:347:HOH:O	2.21	0.81
1:A:20[B]:LEU:HD12	5:A:360:HOH:O	1.82	0.79
1:A:170:SER:O	1:A:171:VAL:C	2.15	0.79
1:A:88[B]:GLU:OE2	5:A:201:HOH:O	2.01	0.79
1:A:67[A]:LYS:HE3	5:A:321:HOH:O	1.83	0.78
1:A:49[B]:GLU:OE1	5:A:478:HOH:O	2.01	0.77
1:A:158:GLU:O	5:A:320:HOH:O	2.06	0.74
1:A:136[B]:GLU:OE2	5:A:338:HOH:O	2.06	0.72
1:A:85:GLU:HG2	1:A:86:ARG:HG3	1.73	0.69
1:A:53:GLU:HG3	5:A:257:HOH:O	1.93	0.69
1:A:57:GLU:OE1	5:A:349:HOH:O	2.13	0.67
1:A:136[B]:GLU:HG2	5:A:344:HOH:O	1.98	0.62
1:A:170:SER:O	1:A:171:VAL:O	2.17	0.61
1:A:136[B]:GLU:HG3	5:A:507:HOH:O	2.01	0.61
1:A:9[B]:HIS:HD2	5:A:367:HOH:O	1.85	0.58
1:A:46:ASN:OD1	5:A:478:HOH:O	2.17	0.58
1:A:20[A]:LEU:HD11	1:A:66:MET:HG2	1.85	0.58
1:A:1:VAL:HB	5:A:496:HOH:O	2.04	0.56
1:A:136[B]:GLU:CG	5:A:345:HOH:O	2.35	0.55
1:A:9[B]:HIS:CD2	5:A:367:HOH:O	2.59	0.54
1:A:135:GLU:HG3	1:A:139:LYS:HE2	1.92	0.52
1:A:86:ARG:NH2	5:A:201:HOH:O	2.41	0.52
1:A:136[B]:GLU:CD	5:A:322:HOH:O	2.43	0.51
1:A:97:GLN:NE2	5:A:308:HOH:O	1.99	0.51
1:A:168:LYS:HG2	5:A:195:HOH:O	2.13	0.49
1:A:61[B]:HIS:CD2	5:A:197:HOH:O	2.67	0.47
1:A:168:LYS:HE2	5:A:195:HOH:O	2.17	0.45
1:A:54[A]:HIS:CD2	5:A:198:HOH:O	2.72	0.43
1:A:30:TYR:CE2	1:A:54[B]:HIS:CD2	3.09	0.41

All (11) 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)
5:A:487:HOH:O	5:A:487:HOH:O[19_566]	1.69	0.51
5:A:225:HOH:O	5:A:348:HOH:O[88_555]	1.86	0.34
1:A:38:ASP:OD2	5:A:350:HOH:O[19_566]	1.87	0.33
5:A:346:HOH:O	5:A:421:HOH:O[19_566]	2.00	0.20
1:A:45[A]:HIS:CD2	5:A:335:HOH:O[87_455]	2.03	0.17
5:A:321:HOH:O	5:A:470:HOH:O[60_565]	2.03	0.17
5:A:204:HOH:O	5:A:333:HOH:O[87_455]	2.11	0.09
1:A:172:LYS:NZ	5:A:343:HOH:O[88_555]	2.13	0.07
1:A:45[A]:HIS:CE1	5:A:333:HOH:O[87_455]	2.18	0.02
5:A:336:HOH:O	5:A:337:HOH:O[88_555]	2.18	0.02
5:A:244:HOH:O	5:A:320:HOH:O[88_555]	2.18	0.02

## 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	184/176 (104%)	177 (96%)	6 (3%)	1 (0%)	32 10

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	171	VAL

### 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.

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
-----	-------	----------	-----------	----------	-------------

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	A	166/157 (106%)	162 (98%)	4 (2%)	54 18

All (4) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	59	ARG
1	A	85	GLU
1	A	88[A]	GLU
1	A	88[B]	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	97	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](#)

Of 19 ligands modelled in this entry, 19 are monoatomic - leaving 0 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

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 [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<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	172/176 (97%)	0.03	4 (2%) 61 61	9, 12, 20, 40	0

All (4) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	172	LYS	5.1
1	A	1	VAL	4.3
1	A	171	VAL	2.6
1	A	9[A]	HIS	2.0

### 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](#)

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, 95<sup>th</sup> 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	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
2	CO	A	184	1/1	0.99	0.14	0.26	19,19,19,19	1

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
2	CO	A	178	1/1	1.00	0.08	-1.19	16,16,16,16	1
4	CL	A	190	1/1	1.00	0.07	-1.21	22,22,22,22	0
2	CO	A	177	1/1	1.00	0.06	-1.63	14,14,14,14	1
4	CL	A	194	1/1	1.00	0.04	-1.85	29,29,29,29	0
2	CO	A	183	1/1	0.90	0.07	-4.96	28,28,28,28	1
3	MG	A	187	1/1	0.99	0.17	-	22,22,22,22	0
4	CL	A	193	1/1	1.00	0.05	-	20,20,20,20	0
3	MG	A	192	1/1	0.92	0.30	-	35,35,35,35	0
2	CO	A	180	1/1	1.00	0.06	-	14,14,14,14	1
2	CO	A	181	1/1	0.99	0.14	-	22,22,22,22	1
3	MG	A	186	1/1	1.00	0.09	-	13,13,13,13	1
3	MG	A	188	1/1	0.99	0.40	-	28,28,28,28	0
2	CO	A	185	1/1	1.00	0.10	-	19,19,19,19	1
3	MG	A	191	1/1	0.99	0.08	-	18,18,18,18	0
2	CO	A	176	1/1	1.00	0.07	-	13,13,13,13	1
2	CO	A	179	1/1	0.99	0.10	-	20,20,20,20	1
4	CL	A	189	1/1	1.00	0.06	-	13,13,13,13	1
2	CO	A	182	1/1	0.98	0.10	-	28,28,28,28	1

## 6.5 Other polymers [i](#)

There are no such residues in this entry.