



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 21, 2017 – 04:52 pm GMT

PDB ID : 5GOU  
Title : Structure of a 16-mer protein nanocage fabricated from its 24-mer analogue  
by subunit interface redesign  
Authors : Zhang, S.; Zang, J.; Wang, W.; Wang, H.; Zhao, G.  
Deposited on : 2016-07-29  
Resolution : 2.91 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](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	:	recalc29047
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)	:	recalc29047

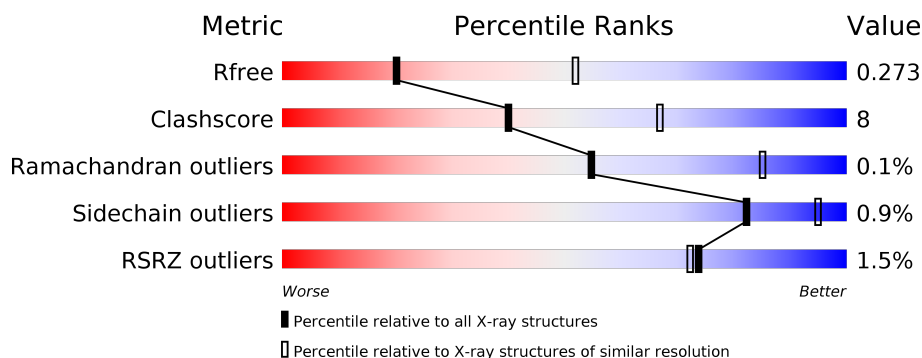
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.91 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	1813 (2.94-2.90)
Clashscore	112137	2045 (2.94-2.90)
Ramachandran outliers	110173	1997 (2.94-2.90)
Sidechain outliers	110143	1999 (2.94-2.90)
RSRZ outliers	101464	1825 (2.94-2.90)

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	190	<div> <div>2%</div> <div>57% 17% 26%</div> </div>
1	B	190	<div> <div>2%</div> <div>56% 13% 31%</div> </div>
1	C	190	<div> <div>66% 12% 22%</div> </div>
1	D	190	<div> <div>2%</div> <div>51% 17% 31%</div> </div>
1	E	190	<div> <div>2%</div> <div>48% 21% 31%</div> </div>
1	F	190	<div> <div>71% 9% 20%</div> </div>

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	G	190	
1	H	190	
1	I	190	
1	J	190	
1	K	190	
1	L	190	
1	M	190	
1	N	190	
1	O	190	
1	P	190	

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 18625 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 heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	141	Total	C	N	O	S	0	0	0
			1153	725	201	221	6			
1	B	131	Total	C	N	O	S	0	0	0
			1088	680	191	211	6			
1	C	148	Total	C	N	O	S	0	0	0
			1207	758	211	232	6			
1	D	131	Total	C	N	O	S	0	0	0
			1088	680	191	211	6			
1	F	152	Total	C	N	O	S	0	0	0
			1246	783	219	238	6			
1	E	132	Total	C	N	O	S	0	0	0
			1103	691	194	212	6			
1	H	145	Total	C	N	O	S	0	0	0
			1183	744	206	227	6			
1	G	134	Total	C	N	O	S	0	0	0
			1109	692	196	215	6			
1	J	146	Total	C	N	O	S	0	0	0
			1188	747	207	228	6			
1	I	133	Total	C	N	O	S	0	0	0
			1104	689	195	214	6			
1	L	146	Total	C	N	O	S	0	0	0
			1192	749	208	229	6			
1	K	133	Total	C	N	O	S	0	0	0
			1104	689	195	214	6			
1	N	144	Total	C	N	O	S	0	0	0
			1176	741	205	224	6			
1	M	133	Total	C	N	O	S	0	0	0
			1110	695	195	214	6			
1	P	143	Total	C	N	O	S	0	0	0
			1168	735	204	223	6			
1	O	133	Total	C	N	O	S	0	0	0
			1110	695	195	214	6			

There are 112 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	139	LEU	-	insertion	UNP P02794
A	140	ASN	-	insertion	UNP P02794
A	141	GLU	-	insertion	UNP P02794
A	142	GLN	-	insertion	UNP P02794
A	143	VAL	-	insertion	UNP P02794
A	144	LYS	-	insertion	UNP P02794
A	145	ALA	-	insertion	UNP P02794
B	139	LEU	-	insertion	UNP P02794
B	140	ASN	-	insertion	UNP P02794
B	141	GLU	-	insertion	UNP P02794
B	142	GLN	-	insertion	UNP P02794
B	143	VAL	-	insertion	UNP P02794
B	144	LYS	-	insertion	UNP P02794
B	145	ALA	-	insertion	UNP P02794
C	139	LEU	-	insertion	UNP P02794
C	140	ASN	-	insertion	UNP P02794
C	141	GLU	-	insertion	UNP P02794
C	142	GLN	-	insertion	UNP P02794
C	143	VAL	-	insertion	UNP P02794
C	144	LYS	-	insertion	UNP P02794
C	145	ALA	-	insertion	UNP P02794
D	139	LEU	-	insertion	UNP P02794
D	140	ASN	-	insertion	UNP P02794
D	141	GLU	-	insertion	UNP P02794
D	142	GLN	-	insertion	UNP P02794
D	143	VAL	-	insertion	UNP P02794
D	144	LYS	-	insertion	UNP P02794
D	145	ALA	-	insertion	UNP P02794
F	139	LEU	-	insertion	UNP P02794
F	140	ASN	-	insertion	UNP P02794
F	141	GLU	-	insertion	UNP P02794
F	142	GLN	-	insertion	UNP P02794
F	143	VAL	-	insertion	UNP P02794
F	144	LYS	-	insertion	UNP P02794
F	145	ALA	-	insertion	UNP P02794
E	139	LEU	-	insertion	UNP P02794
E	140	ASN	-	insertion	UNP P02794
E	141	GLU	-	insertion	UNP P02794
E	142	GLN	-	insertion	UNP P02794
E	143	VAL	-	insertion	UNP P02794
E	144	LYS	-	insertion	UNP P02794
E	145	ALA	-	insertion	UNP P02794

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
H	139	LEU	-	insertion	UNP P02794
H	140	ASN	-	insertion	UNP P02794
H	141	GLU	-	insertion	UNP P02794
H	142	GLN	-	insertion	UNP P02794
H	143	VAL	-	insertion	UNP P02794
H	144	LYS	-	insertion	UNP P02794
H	145	ALA	-	insertion	UNP P02794
G	139	LEU	-	insertion	UNP P02794
G	140	ASN	-	insertion	UNP P02794
G	141	GLU	-	insertion	UNP P02794
G	142	GLN	-	insertion	UNP P02794
G	143	VAL	-	insertion	UNP P02794
G	144	LYS	-	insertion	UNP P02794
G	145	ALA	-	insertion	UNP P02794
J	139	LEU	-	insertion	UNP P02794
J	140	ASN	-	insertion	UNP P02794
J	141	GLU	-	insertion	UNP P02794
J	142	GLN	-	insertion	UNP P02794
J	143	VAL	-	insertion	UNP P02794
J	144	LYS	-	insertion	UNP P02794
J	145	ALA	-	insertion	UNP P02794
I	139	LEU	-	insertion	UNP P02794
I	140	ASN	-	insertion	UNP P02794
I	141	GLU	-	insertion	UNP P02794
I	142	GLN	-	insertion	UNP P02794
I	143	VAL	-	insertion	UNP P02794
I	144	LYS	-	insertion	UNP P02794
I	145	ALA	-	insertion	UNP P02794
L	139	LEU	-	insertion	UNP P02794
L	140	ASN	-	insertion	UNP P02794
L	141	GLU	-	insertion	UNP P02794
L	142	GLN	-	insertion	UNP P02794
L	143	VAL	-	insertion	UNP P02794
L	144	LYS	-	insertion	UNP P02794
L	145	ALA	-	insertion	UNP P02794
K	139	LEU	-	insertion	UNP P02794
K	140	ASN	-	insertion	UNP P02794
K	141	GLU	-	insertion	UNP P02794
K	142	GLN	-	insertion	UNP P02794
K	143	VAL	-	insertion	UNP P02794
K	144	LYS	-	insertion	UNP P02794
K	145	ALA	-	insertion	UNP P02794

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
N	139	LEU	-	insertion	UNP P02794
N	140	ASN	-	insertion	UNP P02794
N	141	GLU	-	insertion	UNP P02794
N	142	GLN	-	insertion	UNP P02794
N	143	VAL	-	insertion	UNP P02794
N	144	LYS	-	insertion	UNP P02794
N	145	ALA	-	insertion	UNP P02794
M	139	LEU	-	insertion	UNP P02794
M	140	ASN	-	insertion	UNP P02794
M	141	GLU	-	insertion	UNP P02794
M	142	GLN	-	insertion	UNP P02794
M	143	VAL	-	insertion	UNP P02794
M	144	LYS	-	insertion	UNP P02794
M	145	ALA	-	insertion	UNP P02794
P	139	LEU	-	insertion	UNP P02794
P	140	ASN	-	insertion	UNP P02794
P	141	GLU	-	insertion	UNP P02794
P	142	GLN	-	insertion	UNP P02794
P	143	VAL	-	insertion	UNP P02794
P	144	LYS	-	insertion	UNP P02794
P	145	ALA	-	insertion	UNP P02794
O	139	LEU	-	insertion	UNP P02794
O	140	ASN	-	insertion	UNP P02794
O	141	GLU	-	insertion	UNP P02794
O	142	GLN	-	insertion	UNP P02794
O	143	VAL	-	insertion	UNP P02794
O	144	LYS	-	insertion	UNP P02794
O	145	ALA	-	insertion	UNP P02794

- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	22	Total O 22 22	0	0
2	B	13	Total O 13 13	0	0
2	C	29	Total O 29 29	0	0
2	D	13	Total O 13 13	0	0
2	F	24	Total O 24 24	0	0

*Continued on next page...*

*Continued from previous page...*

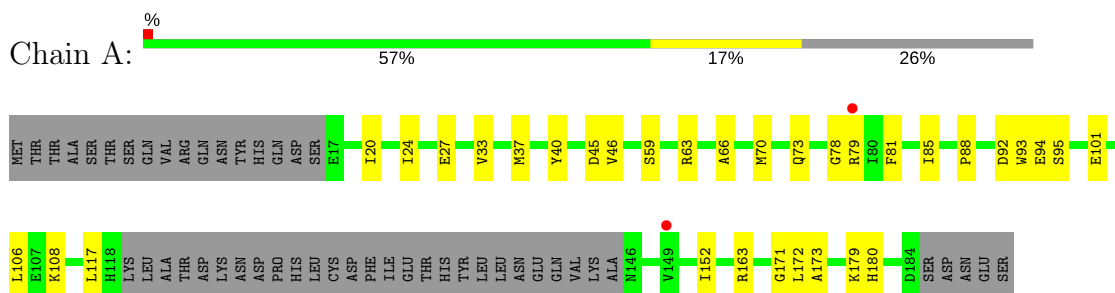
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	E	14	Total 14	O 14	0	0
2	H	20	Total 20	O 20	0	0
2	G	12	Total 12	O 12	0	0
2	J	22	Total 22	O 22	0	0
2	I	11	Total 11	O 11	0	0
2	L	15	Total 15	O 15	0	0
2	K	8	Total 8	O 8	0	0
2	N	26	Total 26	O 26	0	0
2	M	20	Total 20	O 20	0	0
2	P	30	Total 30	O 30	0	0
2	O	17	Total 17	O 17	0	0



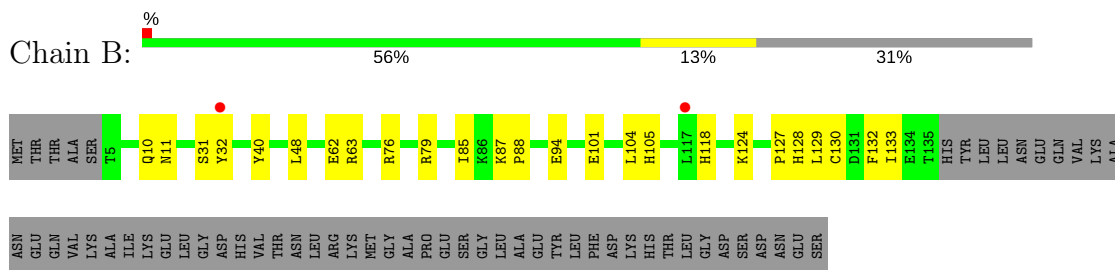
### 3 Residue-property plots [i](#)

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.

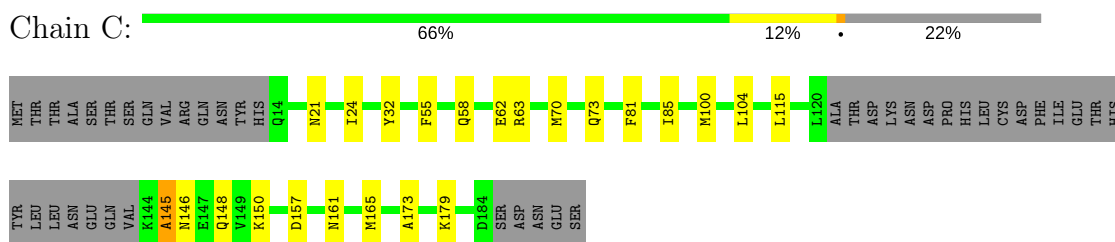
- Molecule 1: Ferritin heavy chain



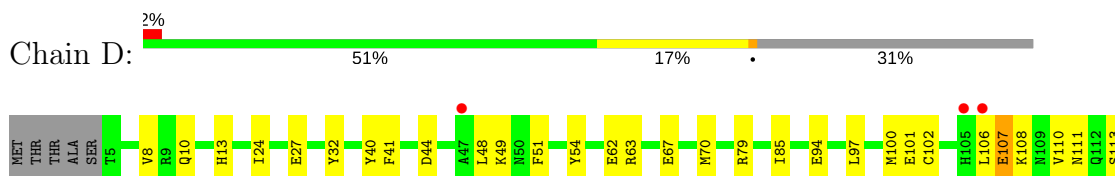
- Molecule 1: Ferritin heavy chain

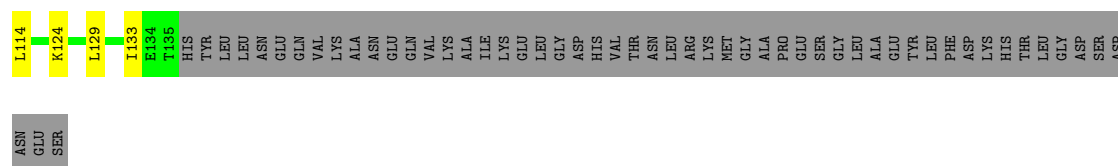


- Molecule 1: Ferritin heavy chain



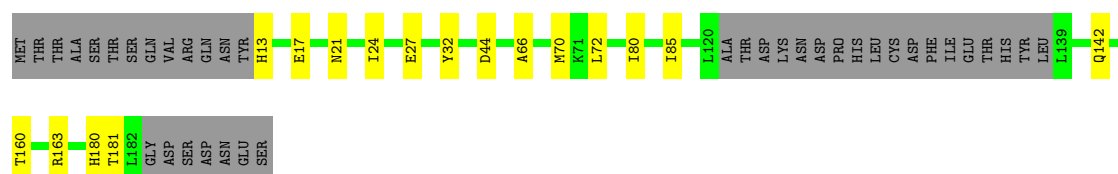
- Molecule 1: Ferritin heavy chain





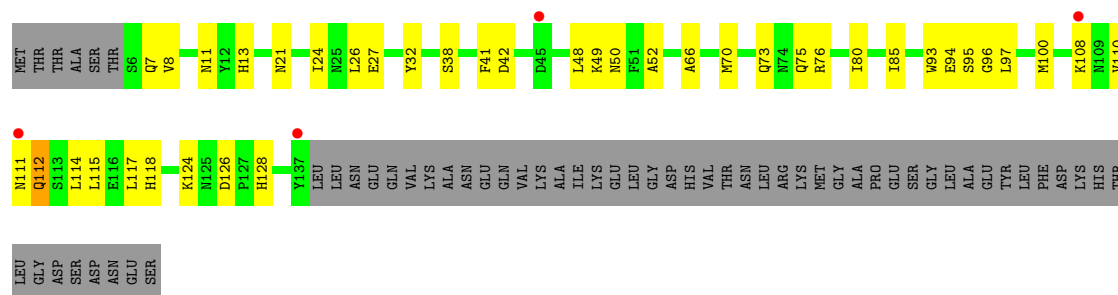
• Molecule 1: Ferritin heavy chain

Chain F: 71% 9% 20%



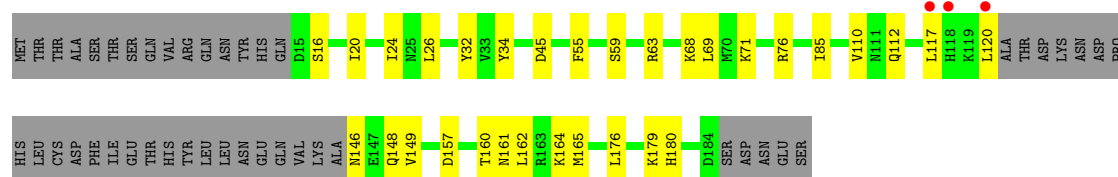
• Molecule 1: Ferritin heavy chain

Chain E: 2% 48% 21% 31%



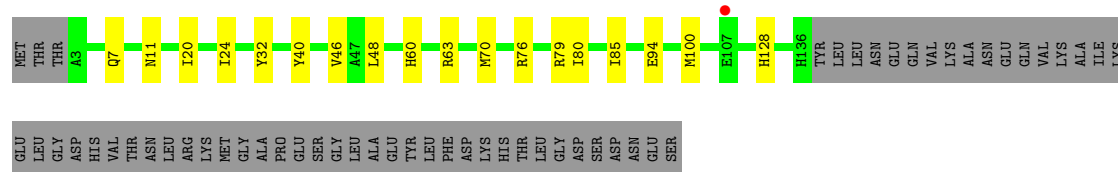
• Molecule 1: Ferritin heavy chain

Chain H: 2% 60% 16% 24%

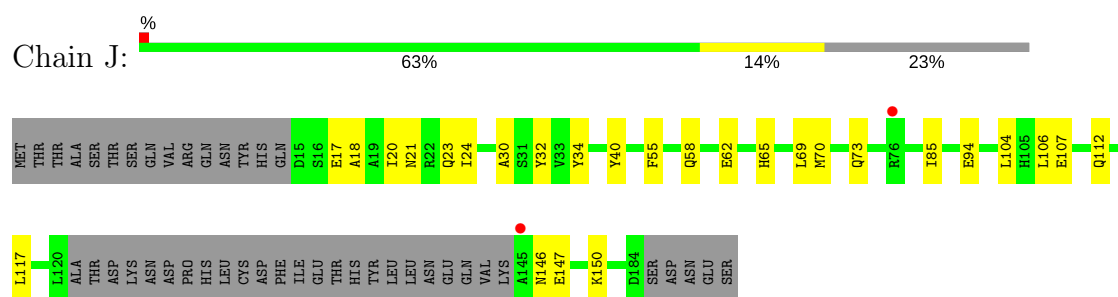


• Molecule 1: Ferritin heavy chain

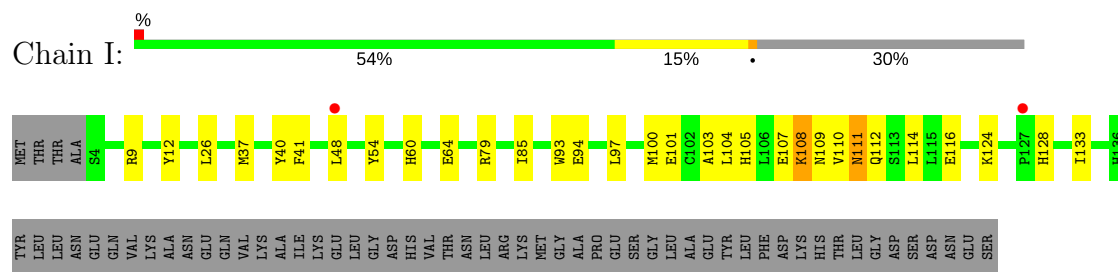
Chain G: 61% 9% 29%



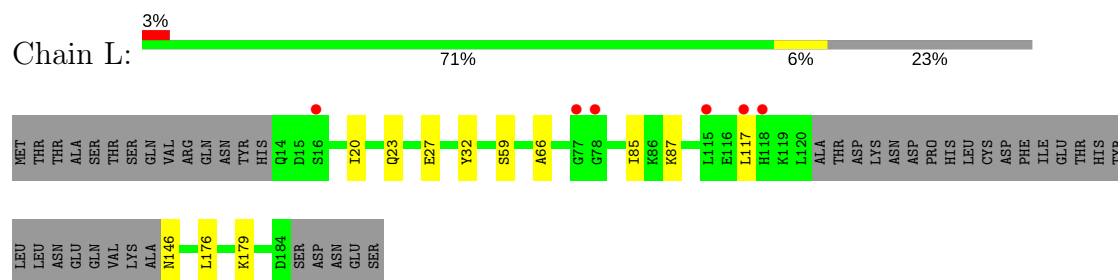
• Molecule 1: Ferritin heavy chain



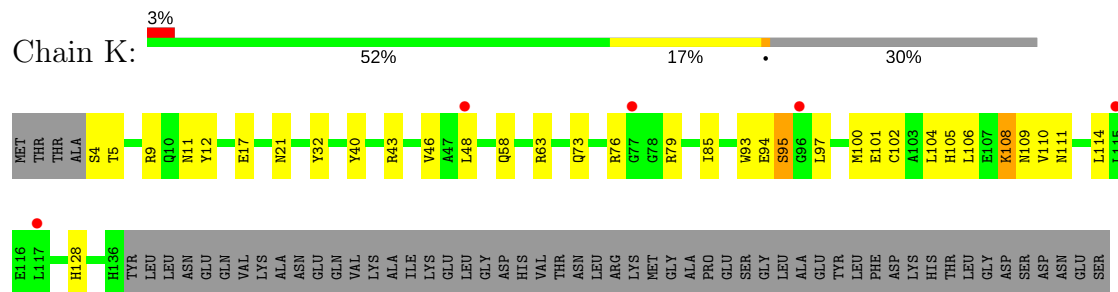
- Molecule 1: Ferritin heavy chain



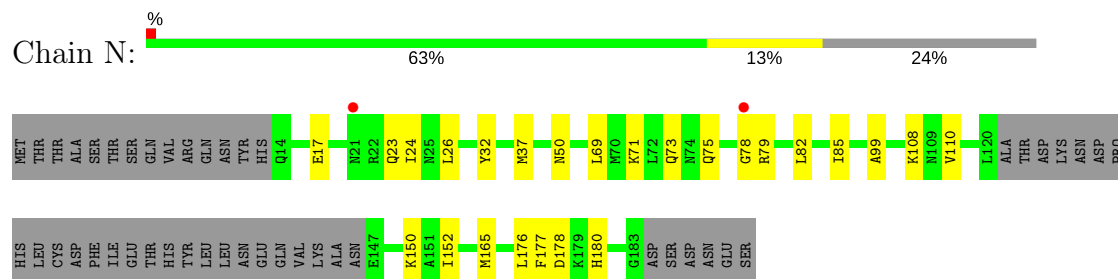
- Molecule 1: Ferritin heavy chain



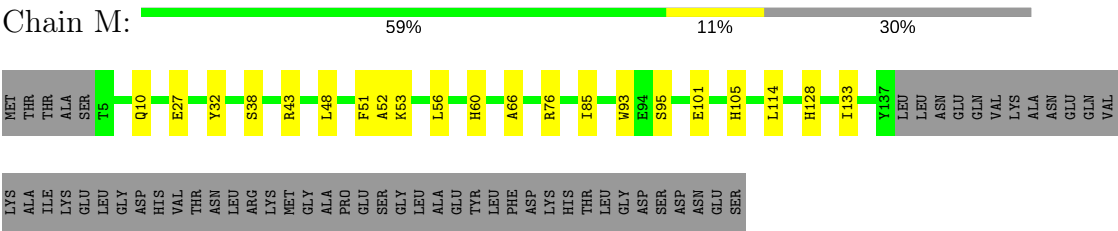
- Molecule 1: Ferritin heavy chain



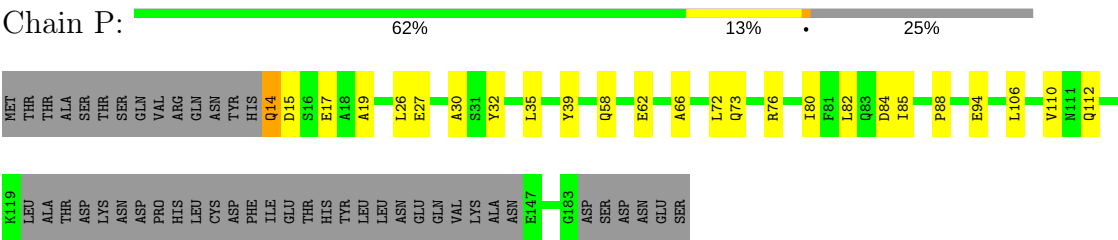
- Molecule 1: Ferritin heavy chain



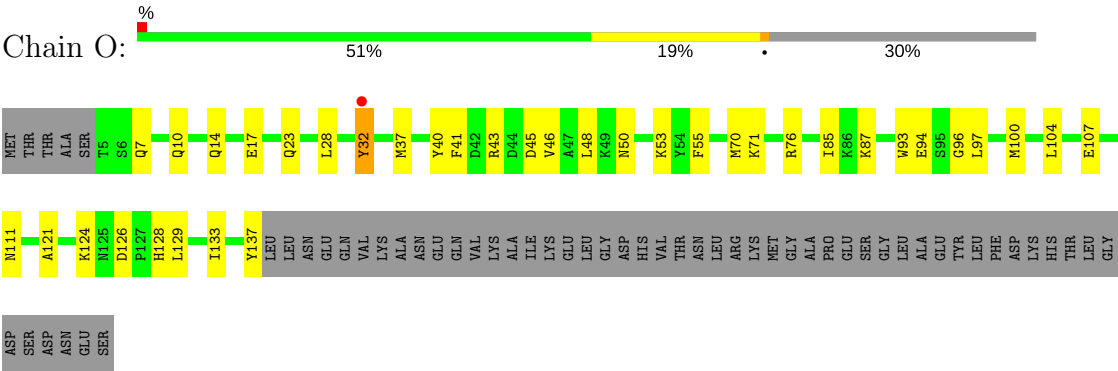
- Molecule 1: Ferritin heavy chain



• Molecule 1: Ferritin heavy chain



• Molecule 1: Ferritin heavy chain



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	99.59Å 95.46Å 111.97Å 67.08° 68.29° 70.06°	Depositor
Resolution (Å)	49.55 – 2.91 49.93 – 2.91	Depositor EDS
% Data completeness (in resolution range)	85.1 (49.55-2.91) 73.3 (49.93-2.91)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	4.02 (at 2.91Å)	Xtriage
Refinement program	PHENIX 1.10.1 _2155	Depositor
R, $R_{free}$	0.228 , 0.275 0.227 , 0.273	Depositor DCC
$R_{free}$ test set	3162 reflections (4.93%)	DCC
Wilson B-factor (Å <sup>2</sup> )	47.5	Xtriage
Anisotropy	0.391	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 49.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.89	EDS
Total number of atoms	18625	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	55.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.90% 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 [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	Chain	Bond lengths		Bond angles	
		RMSZ	# Z  >5	RMSZ	# Z  >5
1	A	0.24	0/1174	0.38	0/1576
1	B	0.28	0/1110	0.39	0/1495
1	C	0.23	0/1228	0.35	0/1647
1	D	0.27	0/1110	0.41	0/1495
1	E	0.24	0/1127	0.39	0/1518
1	F	0.23	0/1268	0.37	0/1702
1	G	0.25	0/1132	0.39	0/1525
1	H	0.24	0/1204	0.36	0/1615
1	I	0.24	0/1127	0.37	0/1518
1	J	0.23	0/1209	0.35	0/1622
1	K	0.24	0/1127	0.42	0/1518
1	L	0.23	0/1213	0.35	0/1627
1	M	0.23	0/1134	0.36	0/1528
1	N	0.23	0/1197	0.36	0/1606
1	O	0.24	0/1134	0.37	0/1528
1	P	0.24	0/1189	0.35	0/1595
All	All	0.24	0/18683	0.37	0/25115

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	C	0	1
1	D	0	1
1	E	0	1
1	I	0	2
1	P	0	1
All	All	0	6

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 6 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	C	145	ALA	Peptide
1	D	107	GLU	Peptide
1	E	112	GLN	Peptide
1	I	108	LYS	Peptide
1	I	111	ASN	Peptide

## 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	1153	0	1114	35	0
1	B	1088	0	1035	39	0
1	C	1207	0	1173	17	0
1	D	1088	0	1035	33	0
1	E	1103	0	1044	32	0
1	F	1246	0	1213	14	0
1	G	1109	0	1052	18	0
1	H	1183	0	1147	26	0
1	I	1104	0	1047	30	0
1	J	1188	0	1152	16	0
1	K	1104	0	1047	25	0
1	L	1192	0	1155	8	0
1	M	1110	0	1051	13	0
1	N	1176	0	1145	14	0
1	O	1110	0	1051	25	0
1	P	1168	0	1134	19	0
2	A	22	0	0	0	0
2	B	13	0	0	0	0
2	C	29	0	0	1	0
2	D	13	0	0	0	0
2	E	14	0	0	0	0
2	F	24	0	0	2	0
2	G	12	0	0	0	0
2	H	20	0	0	0	0
2	I	11	0	0	1	0
2	J	22	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	K	8	0	0	0	0
2	L	15	0	0	0	0
2	M	20	0	0	0	0
2	N	26	0	0	0	0
2	O	17	0	0	1	0
2	P	30	0	0	0	0
All	All	18625	0	17595	299	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 8.

The worst 5 of 299 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:118:HIS:CD2	1:B:133:ILE:CD1	1.97	1.46
1:B:118:HIS:CD2	1:B:133:ILE:HD11	1.50	1.45
1:B:118:HIS:HD2	1:B:133:ILE:CD1	1.39	1.22
1:A:73:GLN:HE21	1:A:79:ARG:N	1.41	1.18
1:B:118:HIS:CD2	1:B:133:ILE:HD12	1.76	1.10

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	Percentiles	
1	A	137/190 (72%)	128 (93%)	9 (7%)	0	100	100
1	B	129/190 (68%)	120 (93%)	9 (7%)	0	100	100
1	C	144/190 (76%)	140 (97%)	4 (3%)	0	100	100
1	D	129/190 (68%)	124 (96%)	5 (4%)	0	100	100
1	E	130/190 (68%)	123 (95%)	6 (5%)	1 (1%)	22	55

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	F	148/190 (78%)	145 (98%)	3 (2%)	0	100	100
1	G	132/190 (70%)	127 (96%)	5 (4%)	0	100	100
1	H	141/190 (74%)	139 (99%)	2 (1%)	0	100	100
1	I	131/190 (69%)	128 (98%)	3 (2%)	0	100	100
1	J	142/190 (75%)	135 (95%)	7 (5%)	0	100	100
1	K	131/190 (69%)	125 (95%)	5 (4%)	1 (1%)	22	55
1	L	142/190 (75%)	138 (97%)	4 (3%)	0	100	100
1	M	131/190 (69%)	128 (98%)	3 (2%)	0	100	100
1	N	140/190 (74%)	136 (97%)	4 (3%)	0	100	100
1	O	131/190 (69%)	125 (95%)	6 (5%)	0	100	100
1	P	139/190 (73%)	136 (98%)	2 (1%)	1 (1%)	25	59
All	All	2177/3040 (72%)	2097 (96%)	77 (4%)	3 (0%)	55	85

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	E	94	GLU
1	K	95	SER
1	P	94	GLU

### 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	123/169 (73%)	123 (100%)	0	100	100
1	B	119/169 (70%)	119 (100%)	0	100	100
1	C	129/169 (76%)	128 (99%)	1 (1%)	85	95
1	D	119/169 (70%)	117 (98%)	2 (2%)	66	89
1	E	120/169 (71%)	119 (99%)	1 (1%)	85	95
1	F	134/169 (79%)	133 (99%)	1 (1%)	87	96

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	G	121/169 (72%)	120 (99%)	1 (1%)	85	95
1	H	127/169 (75%)	126 (99%)	1 (1%)	85	95
1	I	121/169 (72%)	121 (100%)	0	100	100
1	J	127/169 (75%)	126 (99%)	1 (1%)	85	95
1	K	121/169 (72%)	119 (98%)	2 (2%)	66	89
1	L	128/169 (76%)	126 (98%)	2 (2%)	68	89
1	M	121/169 (72%)	119 (98%)	2 (2%)	66	89
1	N	126/169 (75%)	125 (99%)	1 (1%)	85	95
1	O	121/169 (72%)	118 (98%)	3 (2%)	53	83
1	P	125/169 (74%)	125 (100%)	0	100	100
All	All	1982/2704 (73%)	1964 (99%)	18 (1%)	82	95

5 of 18 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	L	32	TYR
1	L	87	LYS
1	M	43	ARG
1	G	32	TYR
1	J	32	TYR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 16 such sidechains are listed below:

Mol	Chain	Res	Type
1	F	73	GLN
1	E	7	GLN
1	G	60	HIS
1	F	21	ASN
1	J	73	GLN

### 5.3.3 RNA ⓘ

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 ⓘ

### 6.1 Protein, DNA and RNA chains ⓘ

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	141/190 (74%)	0.09	2 (1%) 75 75	36, 50, 78, 97	0
1	B	131/190 (68%)	-0.02	2 (1%) 74 72	36, 58, 79, 87	0
1	C	148/190 (77%)	-0.11	0 100 100	35, 48, 72, 111	0
1	D	131/190 (68%)	0.07	3 (2%) 61 58	45, 66, 83, 98	0
1	E	132/190 (69%)	0.16	4 (3%) 51 45	37, 60, 90, 107	0
1	F	152/190 (80%)	-0.04	0 100 100	30, 47, 82, 102	0
1	G	134/190 (70%)	-0.09	1 (0%) 87 86	34, 55, 75, 87	0
1	H	145/190 (76%)	-0.01	3 (2%) 64 61	31, 45, 64, 90	0
1	I	133/190 (70%)	0.12	2 (1%) 74 72	43, 62, 81, 90	0
1	J	146/190 (76%)	-0.09	2 (1%) 75 75	32, 46, 69, 81	0
1	K	133/190 (70%)	0.16	5 (3%) 41 36	45, 69, 92, 99	0
1	L	146/190 (76%)	0.15	6 (4%) 38 33	38, 55, 87, 101	0
1	M	133/190 (70%)	-0.14	0 100 100	34, 52, 73, 83	0
1	N	144/190 (75%)	0.04	2 (1%) 75 75	33, 50, 83, 99	0
1	O	133/190 (70%)	0.02	1 (0%) 86 85	31, 52, 74, 84	0
1	P	143/190 (75%)	-0.09	0 100 100	28, 40, 66, 97	0
All	All	2225/3040 (73%)	0.01	33 (1%) 74 72	28, 53, 82, 111	0

The worst 5 of 33 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	137	TYR	3.8
1	D	105	HIS	3.6
1	L	118	HIS	3.3
1	A	79	ARG	3.3
1	B	117	LEU	3.2

## 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.