



wwPDB X-ray Structure Validation Summary Report ⓘ

Jun 23, 2024 – 01:36 AM EDT

PDB ID : 6JOB
Title : Ferritin variant with "GMG" motif
Authors : Zheng, B.W.; Zhou, K.; Zhang, T.; Lv, C.; Wang, H.; Zhao, G.
Deposited on : 2019-03-20
Resolution : 2.93 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 ⓘ symbol.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

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

MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.20.1
EDS	:	2.37.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

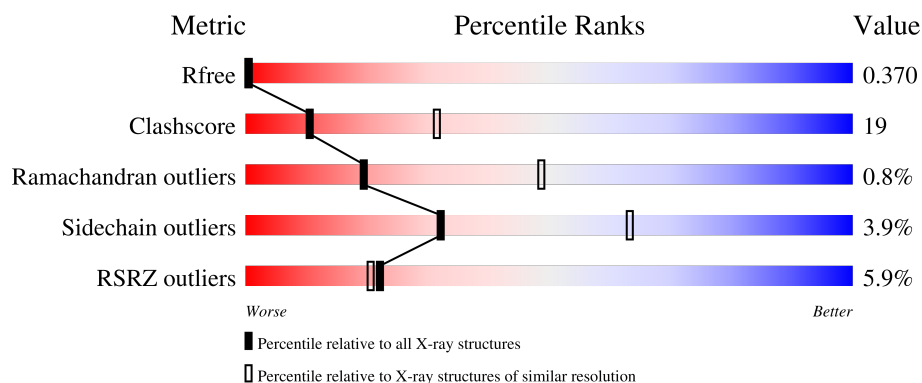
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION, SOLUTION SCATTERING

The reported resolution of this entry is 2.93 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}	130704	2969 (2.98-2.90)
Clashscore	141614	3218 (2.98-2.90)
Ramachandran outliers	138981	3122 (2.98-2.90)
Sidechain outliers	138945	3124 (2.98-2.90)
RSRZ outliers	127900	2902 (2.98-2.90)

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 of 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 $\leq 5\%$. 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	183	 4% 60% 30% •• 6%
1	B	183	 7% 55% 34% •• 6%
1	C	183	 7% 56% 33% 5% 6%
1	D	183	 8% 57% 34% • 6%
1	E	183	 3% 63% 27% •• 6%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	183	<div><div></div><div>4%</div><div>56%</div><div>32%</div><div>6%</div><div>6%</div></div>
1	G	183	<div><div></div><div>4%</div><div>59%</div><div>34%</div><div>•</div><div>6%</div></div>
1	H	183	<div><div></div><div>7%</div><div>50%</div><div>43%</div><div>•</div><div>6%</div></div>
1	I	183	<div><div></div><div>8%</div><div>57%</div><div>34%</div><div>•</div><div>6%</div></div>
1	J	183	<div><div></div><div>5%</div><div>58%</div><div>34%</div><div>•</div><div>6%</div></div>
1	K	183	<div><div></div><div>4%</div><div>56%</div><div>37%</div><div>•</div><div>6%</div></div>
1	L	183	<div><div></div><div>5%</div><div>57%</div><div>34%</div><div>•</div><div>6%</div></div>

2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 16956 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	172	Total	C	N	O	S	0	0	0
			1413	887	248	270	8			
1	B	172	Total	C	N	O	S	0	0	0
			1413	887	248	270	8			
1	C	172	Total	C	N	O	S	0	0	0
			1413	887	248	270	8			
1	D	172	Total	C	N	O	S	0	0	0
			1413	887	248	270	8			
1	E	172	Total	C	N	O	S	0	0	0
			1413	887	248	270	8			
1	F	172	Total	C	N	O	S	0	0	0
			1413	887	248	270	8			
1	G	172	Total	C	N	O	S	0	0	0
			1413	887	248	270	8			
1	H	172	Total	C	N	O	S	0	0	0
			1413	887	248	270	8			
1	I	172	Total	C	N	O	S	0	0	0
			1413	887	248	270	8			
1	J	172	Total	C	N	O	S	0	0	0
			1413	887	248	270	8			
1	K	172	Total	C	N	O	S	0	0	0
			1413	887	248	270	8			
1	L	172	Total	C	N	O	S	0	0	0
			1413	887	248	270	8			

There are 60 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	86	GLN	LYS	conflict	UNP P02794
A	160	LEU	ALA	conflict	UNP P02794
A	161	MET	PRO	conflict	UNP P02794
A	162	VAL	GLU	conflict	UNP P02794
A	163	GLY	SER	conflict	UNP P02794

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	86	GLN	LYS	conflict	UNP P02794
B	160	LEU	ALA	conflict	UNP P02794
B	161	MET	PRO	conflict	UNP P02794
B	162	VAL	GLU	conflict	UNP P02794
B	163	GLY	SER	conflict	UNP P02794
C	86	GLN	LYS	conflict	UNP P02794
C	160	LEU	ALA	conflict	UNP P02794
C	161	MET	PRO	conflict	UNP P02794
C	162	VAL	GLU	conflict	UNP P02794
C	163	GLY	SER	conflict	UNP P02794
D	86	GLN	LYS	conflict	UNP P02794
D	160	LEU	ALA	conflict	UNP P02794
D	161	MET	PRO	conflict	UNP P02794
D	162	VAL	GLU	conflict	UNP P02794
D	163	GLY	SER	conflict	UNP P02794
E	86	GLN	LYS	conflict	UNP P02794
E	160	LEU	ALA	conflict	UNP P02794
E	161	MET	PRO	conflict	UNP P02794
E	162	VAL	GLU	conflict	UNP P02794
E	163	GLY	SER	conflict	UNP P02794
F	86	GLN	LYS	conflict	UNP P02794
F	160	LEU	ALA	conflict	UNP P02794
F	161	MET	PRO	conflict	UNP P02794
F	162	VAL	GLU	conflict	UNP P02794
F	163	GLY	SER	conflict	UNP P02794
G	86	GLN	LYS	conflict	UNP P02794
G	160	LEU	ALA	conflict	UNP P02794
G	161	MET	PRO	conflict	UNP P02794
G	162	VAL	GLU	conflict	UNP P02794
G	163	GLY	SER	conflict	UNP P02794
H	86	GLN	LYS	conflict	UNP P02794
H	160	LEU	ALA	conflict	UNP P02794
H	161	MET	PRO	conflict	UNP P02794
H	162	VAL	GLU	conflict	UNP P02794
H	163	GLY	SER	conflict	UNP P02794
I	86	GLN	LYS	conflict	UNP P02794
I	160	LEU	ALA	conflict	UNP P02794
I	161	MET	PRO	conflict	UNP P02794
I	162	VAL	GLU	conflict	UNP P02794
I	163	GLY	SER	conflict	UNP P02794
J	86	GLN	LYS	conflict	UNP P02794
J	160	LEU	ALA	conflict	UNP P02794

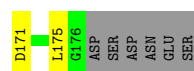
Continued on next page...

Continued from previous page...

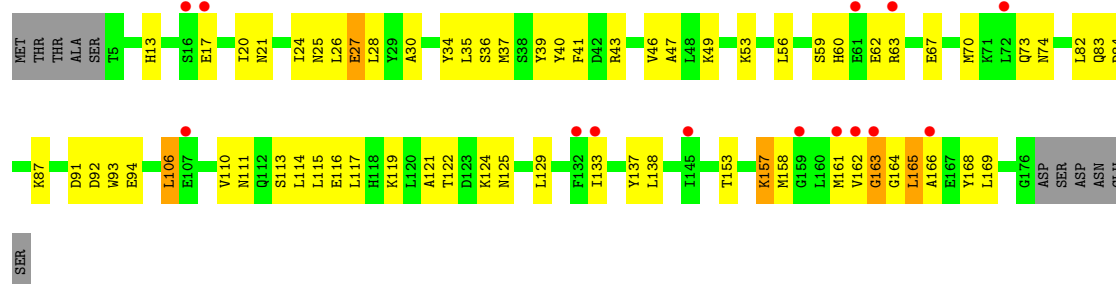
Chain	Residue	Modelled	Actual	Comment	Reference
J	161	MET	PRO	conflict	UNP P02794
J	162	VAL	GLU	conflict	UNP P02794
J	163	GLY	SER	conflict	UNP P02794
K	86	GLN	LYS	conflict	UNP P02794
K	160	LEU	ALA	conflict	UNP P02794
K	161	MET	PRO	conflict	UNP P02794
K	162	VAL	GLU	conflict	UNP P02794
K	163	GLY	SER	conflict	UNP P02794
L	86	GLN	LYS	conflict	UNP P02794
L	160	LEU	ALA	conflict	UNP P02794
L	161	MET	PRO	conflict	UNP P02794
L	162	VAL	GLU	conflict	UNP P02794
L	163	GLY	SER	conflict	UNP P02794

- Molecule 1: Ferritin heavy chain

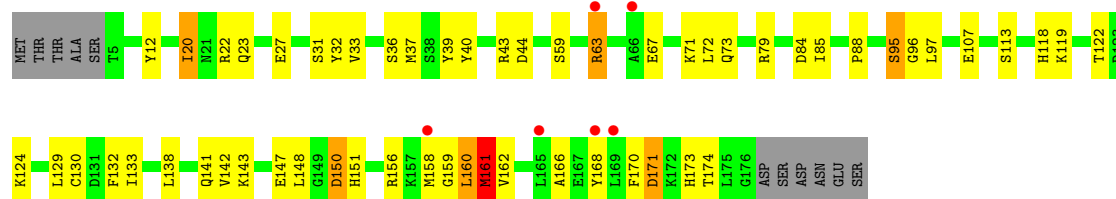




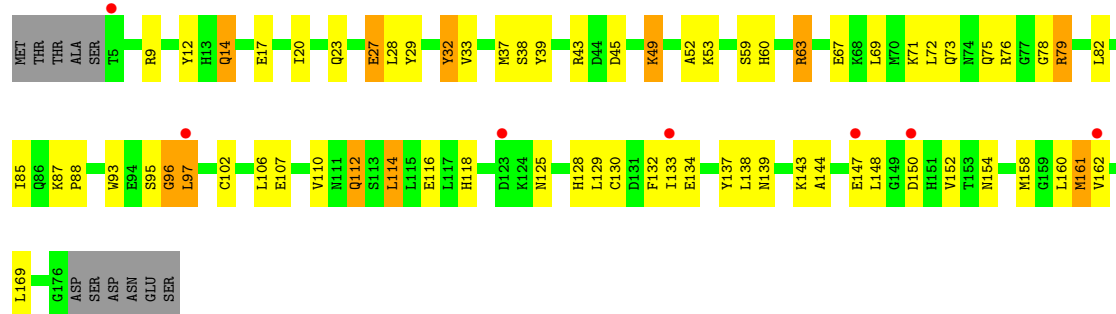
• Molecule 1: Ferritin heavy chain



• Molecule 1: Ferritin heavy chain

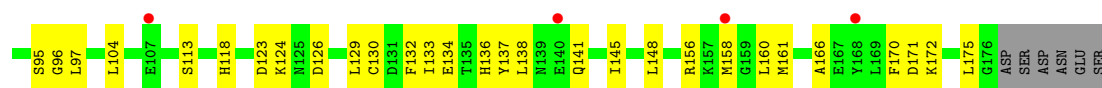


• Molecule 1: Ferritin heavy chain

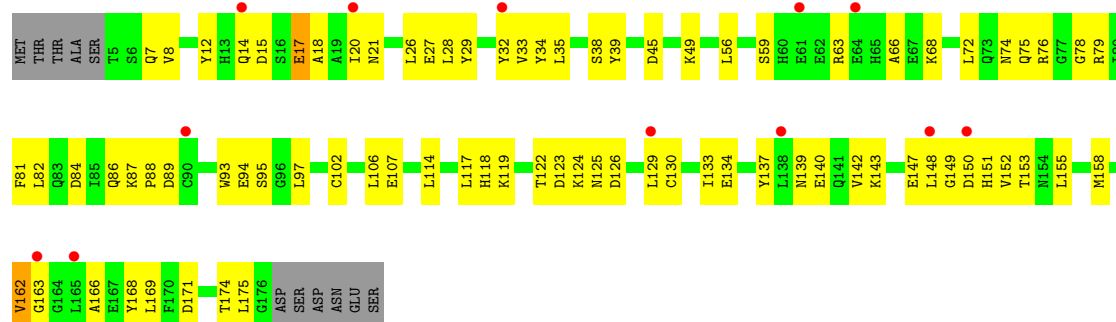


• Molecule 1: Ferritin heavy chain

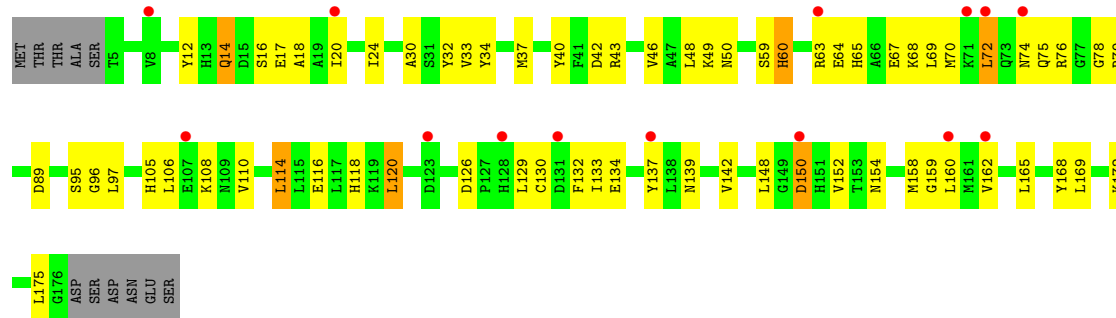




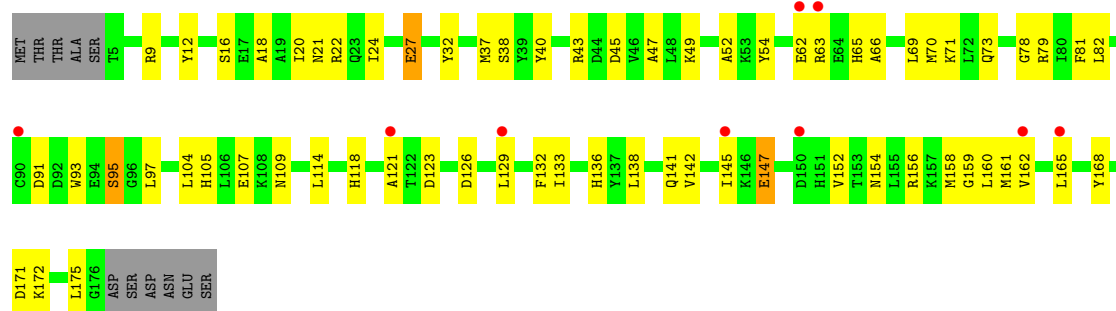
• Molecule 1: Ferritin heavy chain



• Molecule 1: Ferritin heavy chain

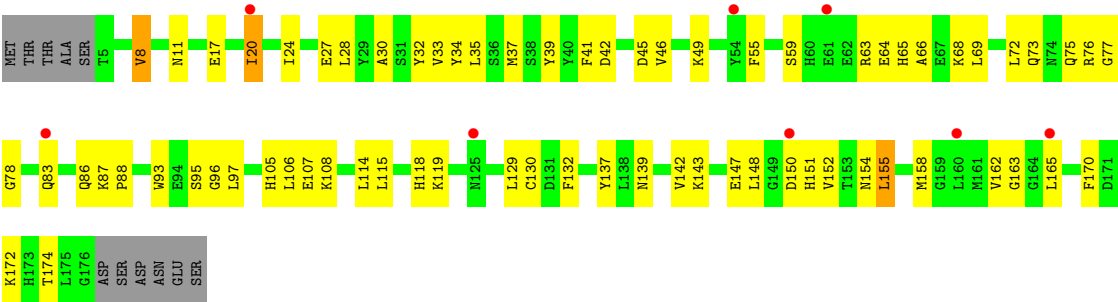


• Molecule 1: Ferritin heavy chain

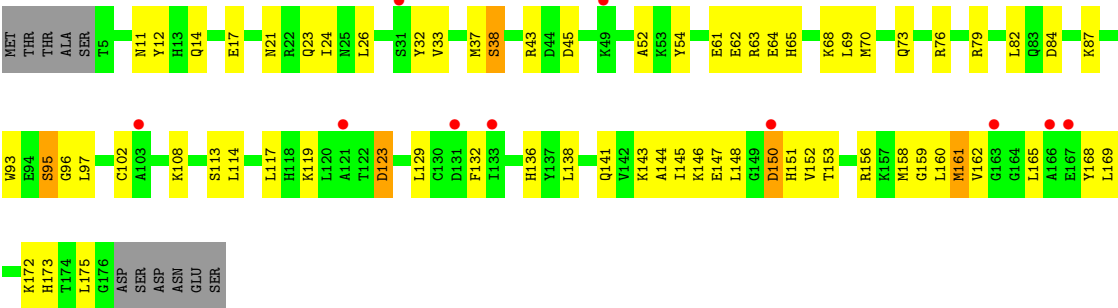


• Molecule 1: Ferritin heavy chain





● Molecule 1: Ferritin heavy chain



4 Data and refinement statistics

Property	Value	Source
Space group	P 4 ₂ 2 ₁ 2	Depositor
Cell constants a, b, c, α , β , γ	219.17Å 219.17Å 147.95Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	107.01 – 2.93 107.01 – 2.93	Depositor EDS
% Data completeness (in resolution range)	97.5 (107.01-2.93) 98.0 (107.01-2.93)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.55 (at 2.91Å)	Xtriage
Refinement program	PHENIX (1.14_3260: ???)	Depositor
R, R_{free}	0.284 , 0.368 0.291 , 0.370	Depositor DCC
R_{free} test set	3859 reflections (5.08%)	wwPDB-VP
Wilson B-factor (Å ²)	59.9	Xtriage
Anisotropy	0.089	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 60.8	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.90	EDS
Total number of atoms	16956	wwPDB-VP
Average B, all atoms (Å ²)	69.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 47.52 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 9.8143e-05. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

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

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	0.52	0/1441	0.83	5/1940 (0.3%)
1	B	0.58	0/1441	0.87	5/1940 (0.3%)
1	C	0.61	0/1441	0.94	8/1940 (0.4%)
1	D	0.58	1/1441 (0.1%)	0.80	2/1940 (0.1%)
1	E	0.57	0/1441	0.79	6/1940 (0.3%)
1	F	0.58	1/1441 (0.1%)	0.89	7/1940 (0.4%)
1	G	0.54	0/1441	0.73	2/1940 (0.1%)
1	H	0.61	1/1441 (0.1%)	0.80	3/1940 (0.2%)
1	I	0.60	0/1441	0.95	13/1940 (0.7%)
1	J	0.64	1/1441 (0.1%)	0.73	1/1940 (0.1%)
1	K	0.55	0/1441	0.83	6/1940 (0.3%)
1	L	0.57	0/1441	0.79	2/1940 (0.1%)
All	All	0.58	4/17292 (0.0%)	0.83	60/23280 (0.3%)

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
1	B	0	1
1	F	0	1
1	G	0	3
1	J	0	1
All	All	0	7

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	J	27	GLU	CB-CG	-10.30	1.32	1.52
1	H	17	GLU	CG-CD	-9.02	1.38	1.51
1	F	27	GLU	CG-CD	-7.64	1.40	1.51

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	62	GLU	CB-CG	-5.33	1.42	1.52

The worst 5 of 60 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	35	LEU	CA-CB-CG	10.98	140.55	115.30
1	C	115	LEU	CA-CB-CG	10.12	138.57	115.30
1	D	35	LEU	CA-CB-CG	9.40	136.91	115.30
1	C	35	LEU	CB-CG-CD1	-9.16	95.43	111.00
1	I	72	LEU	CB-CG-CD2	-9.11	95.51	111.00

There are no chirality outliers.

5 of 7 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	17	GLU	Sidechain
1	B	138	LEU	Peptide
1	F	161	MET	Peptide
1	G	160	LEU	Peptide
1	G	61	GLU	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	1413	0	1361	42	1
1	B	1413	0	1361	58	0
1	C	1413	0	1361	68	0
1	D	1413	0	1361	51	1
1	E	1413	0	1361	54	0
1	F	1413	0	1361	55	0
1	G	1413	0	1361	59	1
1	H	1413	0	1361	73	1
1	I	1413	0	1361	57	1
1	J	1413	0	1361	52	1
1	K	1413	0	1361	59	0
1	L	1413	0	1361	55	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	16956	0	16332	616	3

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 19.

The worst 5 of 616 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:I:63:ARG:O	1:I:67:GLU:OE2	1.53	1.25
1:K:64:GLU:O	1:K:68:LYS:HD3	1.49	1.13
1:K:20:ILE:HD12	1:K:73:GLN:HG2	1.24	1.12
1:C:114:LEU:HD22	1:C:138:LEU:CD1	1.81	1.11
1:K:20:ILE:CD1	1:K:73:GLN:HG2	1.80	1.09

All (3) 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)
1:G:124:LYS:NZ	1:J:123:ASP:O[3_555]	1.88	0.32
1:A:174:THR:OG1	1:I:168:TYR:OH[6_554]	2.05	0.15
1:D:168:TYR:OH	1:H:174:THR:OG1[3_454]	2.15	0.05

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	170/183 (93%)	161 (95%)	6 (4%)	3 (2%)	8	27
1	B	170/183 (93%)	155 (91%)	12 (7%)	3 (2%)	8	27
1	C	170/183 (93%)	159 (94%)	10 (6%)	1 (1%)	25	56
1	D	170/183 (93%)	159 (94%)	10 (6%)	1 (1%)	25	56

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	E	170/183 (93%)	163 (96%)	5 (3%)	2 (1%)	13	38
1	F	170/183 (93%)	161 (95%)	8 (5%)	1 (1%)	25	56
1	G	170/183 (93%)	161 (95%)	9 (5%)	0	100	100
1	H	170/183 (93%)	161 (95%)	7 (4%)	2 (1%)	13	38
1	I	170/183 (93%)	163 (96%)	7 (4%)	0	100	100
1	J	170/183 (93%)	164 (96%)	5 (3%)	1 (1%)	25	56
1	K	170/183 (93%)	165 (97%)	4 (2%)	1 (1%)	25	56
1	L	170/183 (93%)	160 (94%)	8 (5%)	2 (1%)	13	38
All	All	2040/2196 (93%)	1932 (95%)	91 (4%)	17 (1%)	19	49

5 of 17 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	139	ASN
1	C	162	VAL
1	D	163	GLY
1	E	160	LEU
1	E	161	MET

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	153/163 (94%)	144 (94%)	9 (6%)	19	47
1	B	153/163 (94%)	146 (95%)	7 (5%)	27	57
1	C	153/163 (94%)	148 (97%)	5 (3%)	38	69
1	D	153/163 (94%)	144 (94%)	9 (6%)	19	47
1	E	153/163 (94%)	148 (97%)	5 (3%)	38	69
1	F	153/163 (94%)	142 (93%)	11 (7%)	14	37
1	G	153/163 (94%)	152 (99%)	1 (1%)	84	94
1	H	153/163 (94%)	152 (99%)	1 (1%)	84	94

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	I	153/163 (94%)	147 (96%)	6 (4%)	32	63
1	J	153/163 (94%)	147 (96%)	6 (4%)	32	63
1	K	153/163 (94%)	149 (97%)	4 (3%)	46	75
1	L	153/163 (94%)	145 (95%)	8 (5%)	23	53
All	All	1836/1956 (94%)	1764 (96%)	72 (4%)	32	63

5 of 72 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	J	91	ASP
1	L	165	LEU
1	J	126	ASP
1	L	32	TYR
1	D	87	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 17 such sidechains are listed below:

Mol	Chain	Res	Type
1	L	7	GLN
1	L	151	HIS
1	H	7	GLN
1	H	73	GLN
1	H	125	ASN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates ⓘ

There are no monosaccharides 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, 95th 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 ²)	Q<0.9
1	A	172/183 (93%)	0.74	8 (4%) 31 31	46, 65, 105, 157	0
1	B	172/183 (93%)	0.79	13 (7%) 13 12	41, 66, 105, 161	0
1	C	172/183 (93%)	0.81	12 (6%) 16 14	44, 65, 101, 149	0
1	D	172/183 (93%)	0.88	14 (8%) 12 10	45, 65, 100, 147	0
1	E	172/183 (93%)	0.62	6 (3%) 44 42	41, 63, 103, 136	0
1	F	172/183 (93%)	0.72	7 (4%) 37 36	40, 66, 108, 137	0
1	G	172/183 (93%)	0.68	8 (4%) 31 31	41, 64, 93, 142	0
1	H	172/183 (93%)	0.69	12 (6%) 16 14	39, 66, 103, 143	0
1	I	172/183 (93%)	0.70	14 (8%) 12 10	38, 66, 103, 140	0
1	J	172/183 (93%)	0.65	9 (5%) 27 26	40, 66, 98, 140	0
1	K	172/183 (93%)	0.64	8 (4%) 31 31	46, 65, 98, 155	0
1	L	172/183 (93%)	0.73	10 (5%) 23 21	39, 67, 97, 148	0
All	All	2064/2196 (93%)	0.72	121 (5%) 22 20	38, 65, 103, 161	0

The worst 5 of 121 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	161	MET	5.7
1	H	150	ASP	5.1
1	F	5	THR	4.7
1	F	150	ASP	4.6
1	L	150	ASP	4.2

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 monosaccharides 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.