



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 15, 2020 – 03:36 am BST

PDB ID : 5U1A  
Title : Ferritin with Gc MtrE loop 1 inserted at His34  
Authors : Wang, S.  
Deposited on : 2016-11-28  
Resolution : 2.00 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

<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	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.11
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)	:	2.11

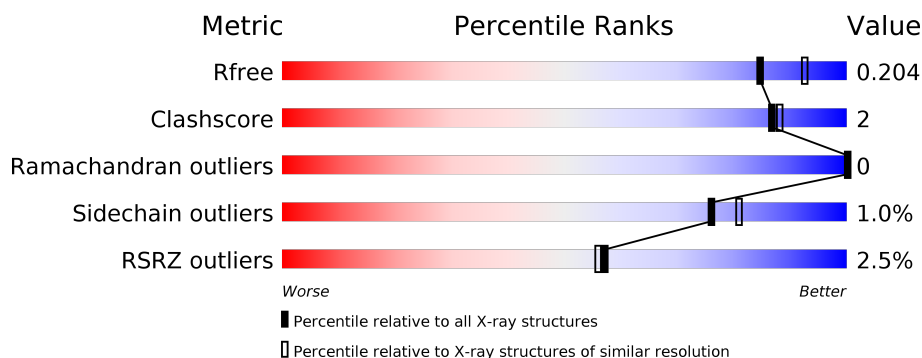
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.00 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	8085 (2.00-2.00)
Clashscore	141614	9178 (2.00-2.00)
Ramachandran outliers	138981	9054 (2.00-2.00)
Sidechain outliers	138945	9053 (2.00-2.00)
RSRZ outliers	127900	7900 (2.00-2.00)

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 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	182	<div> <div>3%</div> <div> <div></div> <div>86%</div> <div>7%</div> <div>7%</div> </div> </div>
1	B	182	<div> <div>3%</div> <div> <div></div> <div>82%</div> <div>8%</div> <div>8%</div> </div> </div>
1	C	182	<div> <div>3%</div> <div> <div></div> <div>83%</div> <div>9%</div> <div>8%</div> </div> </div>
1	D	182	<div> <div>4%</div> <div> <div></div> <div>81%</div> <div>12%</div> <div>7%</div> </div> </div>
1	E	182	<div> <div>3%</div> <div> <div></div> <div>81%</div> <div>12%</div> <div>7%</div> </div> </div>
1	F	182	<div> <div>2%</div> <div> <div></div> <div>82%</div> <div>10%</div> <div>8%</div> </div> </div>

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	G	182	<div><div><div></div><div></div><div></div></div><div><div></div><div></div><div></div></div><div>2%80%13%7%</div></div>
1	H	182	<div><div><div></div><div></div><div></div></div><div><div></div><div></div><div></div></div><div>2%77%14%8%</div></div>
1	I	182	<div><div><div></div><div></div><div></div></div><div><div></div><div></div><div></div></div><div>2%86%7%7%</div></div>
1	J	182	<div><div><div></div><div></div><div></div></div><div><div></div><div></div><div></div></div><div>3%84%8%8%</div></div>
1	K	182	<div><div><div></div><div></div><div></div></div><div><div></div><div></div><div></div></div><div>%86%6%8%</div></div>
1	L	182	<div><div><div></div><div></div><div></div></div><div><div></div><div></div><div></div></div><div>%80%10%9%</div></div>

## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 17961 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,MtrE protein chimera.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	169	Total	C	N	O	S	0	3	0
			1392	888	227	271	6			
1	B	167	Total	C	N	O	S	0	0	0
			1360	866	224	265	5			
1	C	168	Total	C	N	O	S	0	0	0
			1369	871	225	267	6			
1	D	169	Total	C	N	O	S	0	0	0
			1378	875	229	269	5			
1	E	169	Total	C	N	O	S	0	1	0
			1379	877	228	269	5			
1	F	168	Total	C	N	O	S	0	0	0
			1368	871	225	266	6			
1	G	170	Total	C	N	O	S	0	0	0
			1380	877	227	270	6			
1	H	167	Total	C	N	O	S	0	0	0
			1360	866	224	265	5			
1	I	169	Total	C	N	O	S	0	0	0
			1375	874	226	269	6			
1	J	168	Total	C	N	O	S	0	0	0
			1366	869	225	267	5			
1	K	168	Total	C	N	O	S	0	0	0
			1366	869	225	267	5			
1	L	166	Total	C	N	O	S	0	0	0
			1354	863	223	263	5			

There are 48 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	34	GLY	HIS	conflict	UNP A0A0B2EHC8
A	54	CYS	SER	conflict	UNP A0A0B2EHC8
A	56	THR	LEU	conflict	UNP A0A0B2EHC8
A	140	SER	ALA	conflict	UNP A0A0B2EHC8
B	34	GLY	HIS	conflict	UNP A0A0B2EHC8

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	54	CYS	SER	conflict	UNP A0A0B2EHC8
B	56	THR	LEU	conflict	UNP A0A0B2EHC8
B	140	SER	ALA	conflict	UNP A0A0B2EHC8
C	34	GLY	HIS	conflict	UNP A0A0B2EHC8
C	54	CYS	SER	conflict	UNP A0A0B2EHC8
C	56	THR	LEU	conflict	UNP A0A0B2EHC8
C	140	SER	ALA	conflict	UNP A0A0B2EHC8
D	34	GLY	HIS	conflict	UNP A0A0B2EHC8
D	54	CYS	SER	conflict	UNP A0A0B2EHC8
D	56	THR	LEU	conflict	UNP A0A0B2EHC8
D	140	SER	ALA	conflict	UNP A0A0B2EHC8
E	34	GLY	HIS	conflict	UNP A0A0B2EHC8
E	54	CYS	SER	conflict	UNP A0A0B2EHC8
E	56	THR	LEU	conflict	UNP A0A0B2EHC8
E	140	SER	ALA	conflict	UNP A0A0B2EHC8
F	34	GLY	HIS	conflict	UNP A0A0B2EHC8
F	54	CYS	SER	conflict	UNP A0A0B2EHC8
F	56	THR	LEU	conflict	UNP A0A0B2EHC8
F	140	SER	ALA	conflict	UNP A0A0B2EHC8
G	34	GLY	HIS	conflict	UNP A0A0B2EHC8
G	54	CYS	SER	conflict	UNP A0A0B2EHC8
G	56	THR	LEU	conflict	UNP A0A0B2EHC8
G	140	SER	ALA	conflict	UNP A0A0B2EHC8
H	34	GLY	HIS	conflict	UNP A0A0B2EHC8
H	54	CYS	SER	conflict	UNP A0A0B2EHC8
H	56	THR	LEU	conflict	UNP A0A0B2EHC8
H	140	SER	ALA	conflict	UNP A0A0B2EHC8
I	34	GLY	HIS	conflict	UNP A0A0B2EHC8
I	54	CYS	SER	conflict	UNP A0A0B2EHC8
I	56	THR	LEU	conflict	UNP A0A0B2EHC8
I	140	SER	ALA	conflict	UNP A0A0B2EHC8
J	34	GLY	HIS	conflict	UNP A0A0B2EHC8
J	54	CYS	SER	conflict	UNP A0A0B2EHC8
J	56	THR	LEU	conflict	UNP A0A0B2EHC8
J	140	SER	ALA	conflict	UNP A0A0B2EHC8
K	34	GLY	HIS	conflict	UNP A0A0B2EHC8
K	54	CYS	SER	conflict	UNP A0A0B2EHC8
K	56	THR	LEU	conflict	UNP A0A0B2EHC8
K	140	SER	ALA	conflict	UNP A0A0B2EHC8
L	34	GLY	HIS	conflict	UNP A0A0B2EHC8
L	54	CYS	SER	conflict	UNP A0A0B2EHC8
L	56	THR	LEU	conflict	UNP A0A0B2EHC8

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
L	140	SER	ALA	conflict	UNP A0A0B2EHC8

- Molecule 2 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	G	3	Total Na 3 3	0	0
2	J	1	Total Na 1 1	0	0
2	D	1	Total Na 1 1	0	0
2	K	2	Total Na 2 2	0	0
2	E	3	Total Na 3 3	0	0
2	H	4	Total Na 4 4	0	0
2	B	2	Total Na 2 2	0	0
2	I	2	Total Na 2 2	0	0
2	C	1	Total Na 1 1	0	0
2	A	2	Total Na 2 2	0	0
2	L	5	Total Na 5 5	0	0
2	F	4	Total Na 4 4	0	0

- Molecule 3 is FE (III) ION (three-letter code: FE) (formula: Fe).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	G	1	Total Fe 1 1	0	0
3	K	1	Total Fe 1 1	0	0
3	E	1	Total Fe 1 1	0	0
3	A	1	Total Fe 1 1	0	0
3	L	1	Total Fe 1 1	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	F	1	Total	Fe	0	0
			1	1		

- Molecule 4 is GLYCEROL (three-letter code: GOL) (formula:  $C_3H_8O_3$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	B	1	Total	C	O	0	0
			6	3	3		
4	C	1	Total	C	O	0	0
			6	3	3		
4	C	1	Total	C	O	0	0
			6	3	3		
4	D	1	Total	C	O	0	0
			6	3	3		
4	D	1	Total	C	O	0	0
			6	3	3		
4	F	1	Total	C	O	0	0
			6	3	3		
4	F	1	Total	C	O	0	0
			6	3	3		
4	F	1	Total	C	O	0	0
			6	3	3		
4	H	1	Total	C	O	0	0
			6	3	3		
4	I	1	Total	C	O	0	0
			6	3	3		

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	I	1	Total	C	O	0	0
			6	3	3		
4	L	1	Total	C	O	0	0
			6	3	3		

- Molecule 5 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	G	1	Total	Cl	0	0
			1	1		
5	D	1	Total	Cl	0	0
			1	1		
5	K	1	Total	Cl	0	0
			1	1		
5	C	1	Total	Cl	0	0
			1	1		
5	L	3	Total	Cl	0	0
			3	3		
5	F	1	Total	Cl	0	0
			1	1		

- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	96	Total	O	0	0
			96	96		
6	B	106	Total	O	0	0
			106	106		
6	C	106	Total	O	0	0
			106	106		
6	D	95	Total	O	0	0
			95	95		
6	E	110	Total	O	0	0
			110	110		
6	F	146	Total	O	0	0
			146	146		
6	G	133	Total	O	0	0
			133	133		
6	H	102	Total	O	0	0
			102	102		
6	I	119	Total	O	0	0
			119	119		

*Continued on next page...*



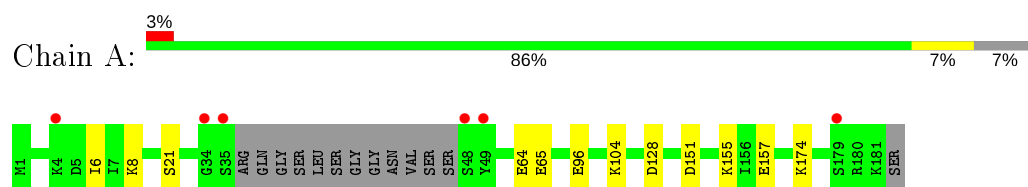
*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	J	102	Total 102	O 102	0	0
6	K	137	Total 137	O 137	0	0
6	L	146	Total 146	O 146	0	0

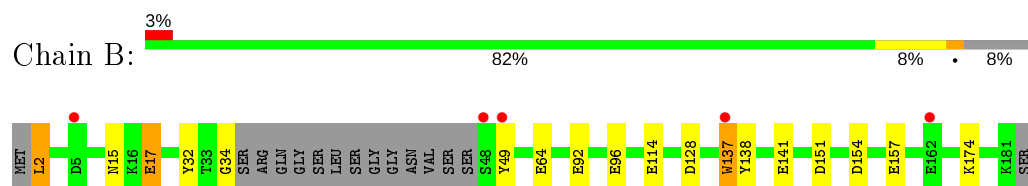
### 3 Residue-property plots [i](#)

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.

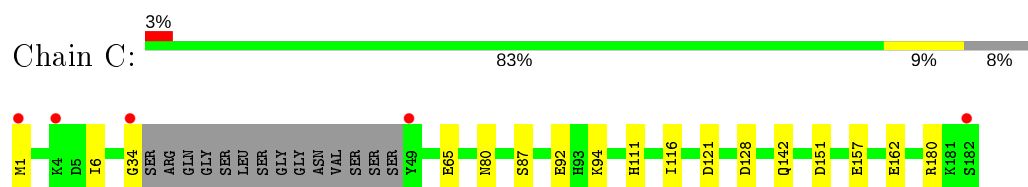
- Molecule 1: Ferritin,MtrE protein chimera



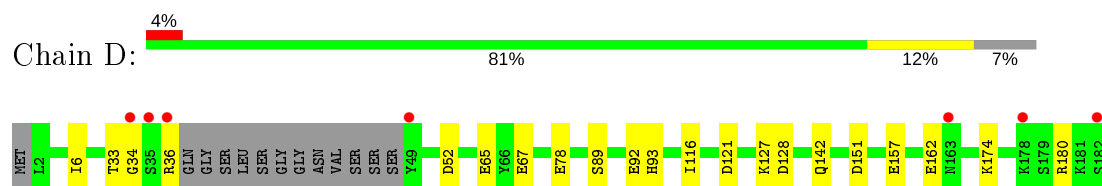
- Molecule 1: Ferritin,MtrE protein chimera



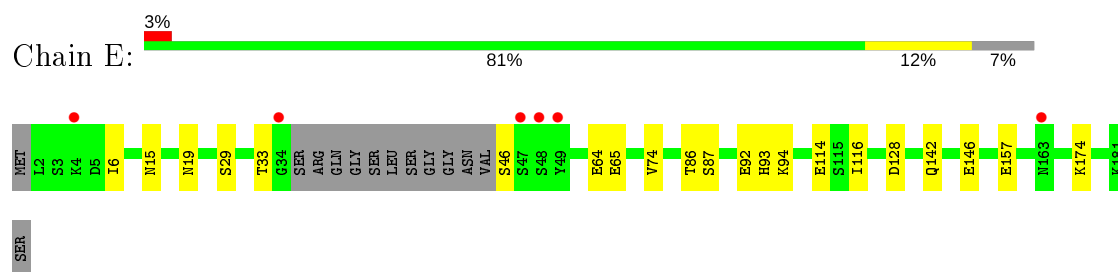
- Molecule 1: Ferritin,MtrE protein chimera



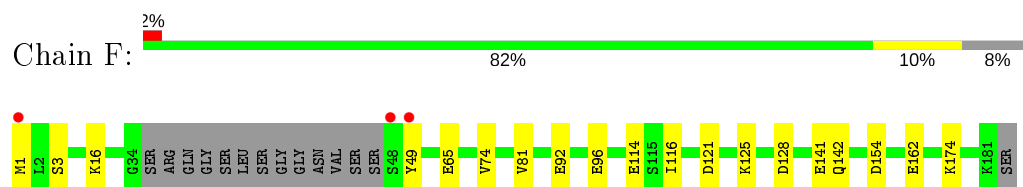
- Molecule 1: Ferritin,MtrE protein chimera



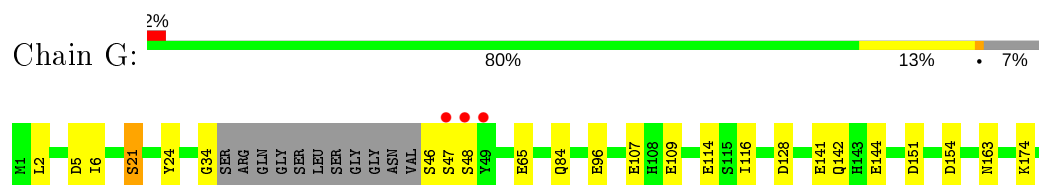
- Molecule 1: Ferritin,MtrE protein chimera



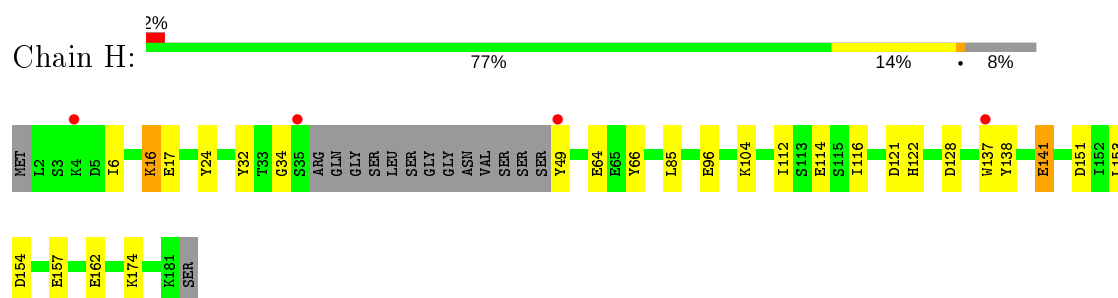
- Molecule 1: Ferritin,MtrE protein chimera



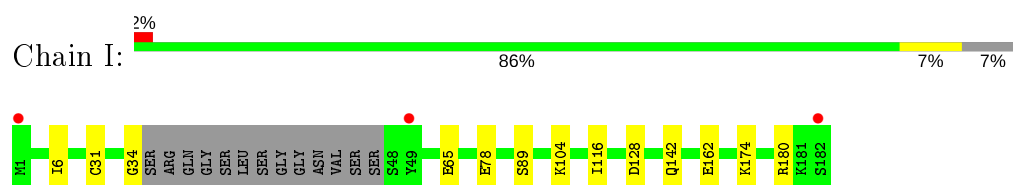
- Molecule 1: Ferritin,MtrE protein chimera



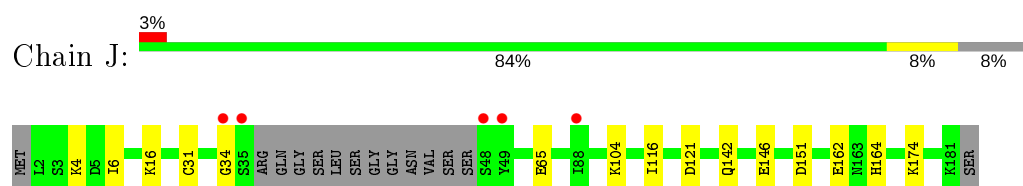
- Molecule 1: Ferritin,MtrE protein chimera



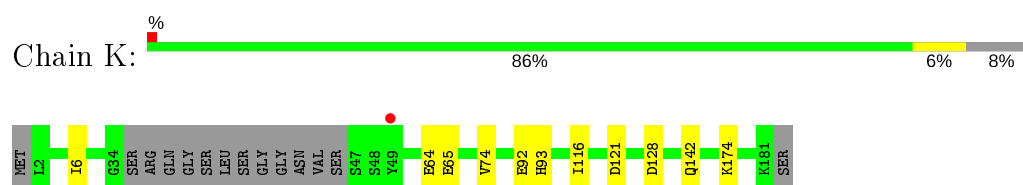
- Molecule 1: Ferritin,MtrE protein chimera



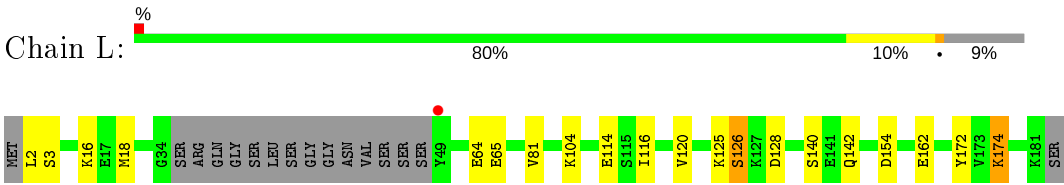
- Molecule 1: Ferritin,MtrE protein chimera



- Molecule 1: Ferritin,MtrE protein chimera



- Molecule 1: Ferritin,MtrE protein chimera



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 4	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	128.41Å 128.41Å 165.04Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 2.00 19.95 – 2.00	Depositor EDS
% Data completeness (in resolution range)	99.7 (20.00-2.00) 99.8 (19.95-2.00)	Depositor EDS
$R_{merge}$	0.14	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.17 (at 2.01Å)	Xtriage
Refinement program	REFMAC 5.8.0155	Depositor
R, $R_{free}$	0.168 , 0.196 0.177 , 0.204	Depositor DCC
$R_{free}$ test set	9002 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	22.9	Xtriage
Anisotropy	0.044	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.43 , 57.6	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	0.022 for h,-k,-l	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	17961	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	27.0	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 28.59 % 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 1.7938e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<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: GOL, CL, FE, NA

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	1.24	8/1434 (0.6%)	1.07	4/1933 (0.2%)
1	B	1.30	10/1389 (0.7%)	1.15	4/1872 (0.2%)
1	C	1.22	4/1398 (0.3%)	1.08	6/1882 (0.3%)
1	D	1.24	10/1407 (0.7%)	1.10	10/1894 (0.5%)
1	E	1.23	5/1412 (0.4%)	1.01	4/1903 (0.2%)
1	F	1.32	8/1397 (0.6%)	1.08	7/1882 (0.4%)
1	G	1.30	7/1409 (0.5%)	1.12	10/1898 (0.5%)
1	H	1.34	8/1389 (0.6%)	1.15	8/1872 (0.4%)
1	I	1.21	4/1404 (0.3%)	1.02	5/1890 (0.3%)
1	J	1.21	4/1395 (0.3%)	1.02	5/1880 (0.3%)
1	K	1.23	3/1395 (0.2%)	1.04	5/1880 (0.3%)
1	L	1.32	8/1383 (0.6%)	1.01	4/1864 (0.2%)
All	All	1.26	79/16812 (0.5%)	1.07	72/22650 (0.3%)

The worst 5 of 79 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	J	65	GLU	CD-OE1	-9.86	1.14	1.25
1	C	65	GLU	CD-OE1	-9.68	1.15	1.25
1	D	65	GLU	CD-OE1	8.57	1.35	1.25
1	E	87	SER	CB-OG	-8.15	1.31	1.42
1	F	3	SER	CB-OG	8.11	1.52	1.42

The worst 5 of 72 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	128	ASP	CB-CG-OD1	12.32	129.39	118.30
1	D	128	ASP	CB-CG-OD2	-12.19	107.33	118.30
1	K	128	ASP	CB-CG-OD1	10.61	127.85	118.30
1	H	154	ASP	CB-CG-OD2	-10.45	108.89	118.30
1	E	128	ASP	CB-CG-OD1	10.41	127.67	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	1392	0	1342	1	0
1	B	1360	0	1308	6	0
1	C	1369	0	1320	4	0
1	D	1378	0	1326	4	0
1	E	1379	0	1325	8	0
1	F	1368	0	1320	7	0
1	G	1380	0	1330	6	0
1	H	1360	0	1307	14	0
1	I	1375	0	1325	4	0
1	J	1366	0	1313	4	0
1	K	1366	0	1313	5	0
1	L	1354	0	1301	8	0
2	A	2	0	0	0	0
2	B	2	0	0	0	0
2	C	1	0	0	0	0
2	D	1	0	0	0	0
2	E	3	0	0	0	0
2	F	4	0	0	0	0
2	G	3	0	0	0	0
2	H	4	0	0	0	0
2	I	2	0	0	0	0
2	J	1	0	0	0	0
2	K	2	0	0	0	0
2	L	5	0	0	0	0
3	A	1	0	0	0	0
3	E	1	0	0	0	0
3	F	1	0	0	0	0
3	G	1	0	0	0	0
3	K	1	0	0	0	0
3	L	1	0	0	0	0
4	B	6	0	8	2	0
4	C	12	0	16	0	0
4	D	12	0	16	1	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	F	18	0	23	2	0
4	H	6	0	8	2	0
4	I	12	0	16	1	0
4	L	6	0	8	0	0
5	C	1	0	0	0	0
5	D	1	0	0	1	0
5	F	1	0	0	0	0
5	G	1	0	0	0	0
5	K	1	0	0	0	0
5	L	3	0	0	2	0
6	A	96	0	0	0	1
6	B	106	0	0	2	0
6	C	106	0	0	1	0
6	D	95	0	0	0	0
6	E	110	0	0	2	0
6	F	146	0	0	2	1
6	G	133	0	0	3	0
6	H	102	0	0	1	0
6	I	119	0	0	2	0
6	J	102	0	0	1	0
6	K	137	0	0	0	0
6	L	146	0	0	2	0
All	All	17961	0	15925	68	1

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 2.

The worst 5 of 68 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:L:126:SER:OG	1:L:128:ASP:OD2	1.58	1.19
1:E:15:ASN:HB2	6:E:386:HOH:O	1.82	0.79
1:J:34:GLY:O	6:J:301:HOH:O	2.02	0.78
1:H:17:GLU:OE2	1:H:138:TYR:OH	2.02	0.77
1:G:34:GLY:HA3	1:G:47:SER:O	1.90	0.71

All (1) 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)
6:A:377:HOH:O	6:F:384:HOH:O[3_765]	2.13	0.07

## 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	168/182 (92%)	165 (98%)	3 (2%)	0	100	100
1	B	163/182 (90%)	163 (100%)	0	0	100	100
1	C	164/182 (90%)	163 (99%)	1 (1%)	0	100	100
1	D	165/182 (91%)	164 (99%)	1 (1%)	0	100	100
1	E	166/182 (91%)	164 (99%)	2 (1%)	0	100	100
1	F	164/182 (90%)	163 (99%)	1 (1%)	0	100	100
1	G	166/182 (91%)	164 (99%)	2 (1%)	0	100	100
1	H	163/182 (90%)	162 (99%)	1 (1%)	0	100	100
1	I	165/182 (91%)	164 (99%)	1 (1%)	0	100	100
1	J	164/182 (90%)	162 (99%)	2 (1%)	0	100	100
1	K	164/182 (90%)	161 (98%)	3 (2%)	0	100	100
1	L	162/182 (89%)	161 (99%)	1 (1%)	0	100	100
All	All	1974/2184 (90%)	1956 (99%)	18 (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	156/163 (96%)	155 (99%)	1 (1%)	86	90
1	B	151/163 (93%)	150 (99%)	1 (1%)	84	88
1	C	152/163 (93%)	149 (98%)	3 (2%)	55	58
1	D	153/163 (94%)	151 (99%)	2 (1%)	69	74
1	E	154/163 (94%)	151 (98%)	3 (2%)	57	61
1	F	152/163 (93%)	151 (99%)	1 (1%)	84	88
1	G	154/163 (94%)	151 (98%)	3 (2%)	57	61
1	H	151/163 (93%)	151 (100%)	0	100	100
1	I	153/163 (94%)	152 (99%)	1 (1%)	84	88
1	J	152/163 (93%)	149 (98%)	3 (2%)	55	58
1	K	152/163 (93%)	151 (99%)	1 (1%)	84	88
1	L	150/163 (92%)	150 (100%)	0	100	100
All	All	1830/1956 (94%)	1811 (99%)	19 (1%)	76	81

5 of 19 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	19	ASN
1	F	1	MET
1	J	4	LYS
1	E	6	ILE
1	J	6	ILE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 7 such sidechains are listed below:

Mol	Chain	Res	Type
1	E	15	ASN
1	K	79	ASN
1	G	19	ASN
1	D	22	ASN
1	H	122	HIS

### 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](#)

Of 56 ligands modelled in this entry, 44 are monoatomic - leaving 12 for Mogul analysis.

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	$\# Z  > 2$
4	GOL	F	207	2	5,5,5	0.44	0	5,5,5	0.58	0
4	GOL	B	203	-	5,5,5	0.70	0	5,5,5	1.03	0
4	GOL	C	203	-	5,5,5	0.88	0	5,5,5	0.82	0
4	GOL	D	203	-	5,5,5	0.51	0	5,5,5	0.86	0
4	GOL	H	204	-	5,5,5	0.49	0	5,5,5	0.86	0
4	GOL	L	207	-	5,5,5	0.90	0	5,5,5	1.19	0
4	GOL	I	203	-	5,5,5	0.44	0	5,5,5	0.99	0
4	GOL	D	202	-	5,5,5	0.49	0	5,5,5	0.57	0
4	GOL	I	204	-	5,5,5	0.47	0	5,5,5	0.79	0
4	GOL	F	201	2	5,5,5	0.42	0	5,5,5	0.99	0
4	GOL	C	202	-	5,5,5	0.34	0	5,5,5	0.82	0
4	GOL	F	208	-	5,5,5	0.40	0	5,5,5	0.72	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
4	GOL	F	207	2	-	2/4/4/4	-
4	GOL	B	203	-	-	2/4/4/4	-

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	GOL	C	203	-	-	0/4/4/4	-
4	GOL	D	203	-	-	0/4/4/4	-
4	GOL	H	204	-	-	2/4/4/4	-
4	GOL	L	207	-	-	2/4/4/4	-
4	GOL	I	203	-	-	3/4/4/4	-
4	GOL	D	202	-	-	0/4/4/4	-
4	GOL	I	204	-	-	1/4/4/4	-
4	GOL	F	201	2	-	2/4/4/4	-
4	GOL	C	202	-	-	2/4/4/4	-
4	GOL	F	208	-	-	0/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 16 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	C	202	GOL	C1-C2-C3-O3
4	B	203	GOL	C1-C2-C3-O3
4	F	201	GOL	O1-C1-C2-O2
4	F	201	GOL	O1-C1-C2-C3
4	H	204	GOL	O1-C1-C2-C3

There are no ring outliers.

5 monomers are involved in 8 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	B	203	GOL	2	0
4	D	203	GOL	1	0
4	H	204	GOL	2	0
4	I	204	GOL	1	0
4	F	201	GOL	2	0

## 5.7 Other polymers [i](#)

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 ⓘ

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	169/182 (92%)	-0.29	6 (3%) 42 42	15, 24, 47, 66	0
1	B	167/182 (91%)	-0.29	5 (2%) 50 49	14, 26, 47, 61	0
1	C	168/182 (92%)	-0.31	5 (2%) 50 49	16, 24, 46, 81	0
1	D	169/182 (92%)	-0.16	7 (4%) 37 36	16, 27, 48, 79	0
1	E	169/182 (92%)	-0.28	6 (3%) 42 42	15, 24, 47, 55	0
1	F	168/182 (92%)	-0.39	3 (1%) 68 66	15, 22, 43, 76	0
1	G	170/182 (93%)	-0.32	4 (2%) 59 57	13, 23, 45, 65	0
1	H	167/182 (91%)	-0.32	4 (2%) 59 57	13, 24, 48, 62	0
1	I	169/182 (92%)	-0.31	3 (1%) 68 66	14, 23, 44, 77	0
1	J	168/182 (92%)	-0.24	5 (2%) 50 49	14, 25, 47, 64	0
1	K	168/182 (92%)	-0.40	1 (0%) 89 88	14, 23, 47, 63	0
1	L	166/182 (91%)	-0.46	1 (0%) 89 88	14, 21, 41, 59	0
All	All	2018/2184 (92%)	-0.31	50 (2%) 57 56	13, 24, 47, 81	0

The worst 5 of 50 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	182	SER	6.1
1	D	182	SER	5.8
1	I	182	SER	5.4
1	A	35	SER	4.9
1	J	49	TYR	4.9

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

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

## 6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 6.4 Ligands ⓘ

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. 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	B-factors(Å <sup>2</sup> )	Q<0.9
3	FE	L	206	1/1	0.65	0.08	42,42,42,42	1
3	FE	E	204	1/1	0.66	0.14	37,37,37,37	1
4	GOL	D	203	6/6	0.75	0.20	43,46,47,49	0
4	GOL	I	203	6/6	0.82	0.16	36,38,40,43	0
2	NA	H	205	1/1	0.84	0.19	49,49,49,49	0
2	NA	F	205	1/1	0.88	0.32	53,53,53,53	0
4	GOL	I	204	6/6	0.89	0.11	36,48,51,53	0
2	NA	I	202	1/1	0.89	0.15	51,51,51,51	0
3	FE	F	206	1/1	0.89	0.04	37,37,37,37	1
4	GOL	C	203	6/6	0.89	0.12	32,35,36,38	0
4	GOL	D	202	6/6	0.89	0.14	39,46,55,58	0
2	NA	G	205	1/1	0.89	0.23	44,44,44,44	0
4	GOL	F	201	6/6	0.90	0.15	38,44,49,50	0
2	NA	L	204	1/1	0.90	0.16	52,52,52,52	0
2	NA	H	202	1/1	0.91	0.15	49,49,49,49	0
2	NA	L	205	1/1	0.91	0.30	43,43,43,43	0
4	GOL	C	202	6/6	0.91	0.20	50,52,59,59	0
2	NA	H	203	1/1	0.91	0.34	38,38,38,38	0
4	GOL	L	207	6/6	0.92	0.13	33,44,46,50	0
4	GOL	B	203	6/6	0.92	0.21	32,45,46,54	0
4	GOL	H	204	6/6	0.92	0.14	30,39,42,45	0
4	GOL	F	208	6/6	0.92	0.14	40,46,53,54	0
2	NA	E	203	1/1	0.93	0.15	46,46,46,46	0
4	GOL	F	207	6/6	0.93	0.15	34,43,46,50	0
3	FE	K	203	1/1	0.93	0.12	38,38,38,38	1
2	NA	B	202	1/1	0.94	0.09	46,46,46,46	0
2	NA	K	202	1/1	0.94	0.09	46,46,46,46	0
2	NA	F	203	1/1	0.95	0.16	40,40,40,40	0
2	NA	L	203	1/1	0.96	0.08	29,29,29,29	0
2	NA	H	201	1/1	0.96	0.07	25,25,25,25	0
2	NA	F	204	1/1	0.97	0.06	23,23,23,23	0
5	CL	L	209	1/1	0.97	0.04	35,35,35,35	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
5	CL	D	204	1/1	0.97	0.12	34,34,34,34	1
2	NA	A	202	1/1	0.97	0.09	41,41,41,41	0
2	NA	E	201	1/1	0.97	0.05	22,22,22,22	0
2	NA	G	202	1/1	0.97	0.14	36,36,36,36	0
2	NA	E	202	1/1	0.97	0.09	29,29,29,29	0
2	NA	C	201	1/1	0.97	0.05	23,23,23,23	0
5	CL	G	204	1/1	0.98	0.07	34,34,34,34	0
2	NA	J	201	1/1	0.98	0.04	24,24,24,24	0
2	NA	D	201	1/1	0.98	0.07	26,26,26,26	0
2	NA	A	201	1/1	0.98	0.05	22,22,22,22	0
2	NA	B	201	1/1	0.98	0.05	23,23,23,23	0
2	NA	L	202	1/1	0.98	0.05	26,26,26,26	0
2	NA	G	201	1/1	0.98	0.05	22,22,22,22	0
2	NA	I	201	1/1	0.98	0.06	23,23,23,23	0
2	NA	L	201	1/1	0.99	0.03	19,19,19,19	0
2	NA	K	201	1/1	0.99	0.06	24,24,24,24	0
3	FE	A	203	1/1	0.99	0.02	41,41,41,41	0
3	FE	G	203	1/1	0.99	0.02	39,39,39,39	0
5	CL	L	210	1/1	0.99	0.04	35,35,35,35	0
5	CL	C	204	1/1	0.99	0.02	24,24,24,24	1
5	CL	K	204	1/1	0.99	0.09	32,32,32,32	0
2	NA	F	202	1/1	0.99	0.04	22,22,22,22	0
5	CL	L	208	1/1	0.99	0.03	21,21,21,21	0
5	CL	F	209	1/1	0.99	0.04	22,22,22,22	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.