



# Full wwPDB X-ray Structure Validation Report ⓘ

May 28, 2020 – 09:32 pm BST

PDB ID : 2ANU  
Title : Crystal structure of Predicted metal-dependent phosphoesterase (PHP family)  
(tm0559) from THERMOTOGA MARITIMA at 2.40 Å resolution  
Authors : Joint Center for Structural Genomics (JCSG)  
Deposited on : 2005-08-11  
Resolution : 2.40 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

<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

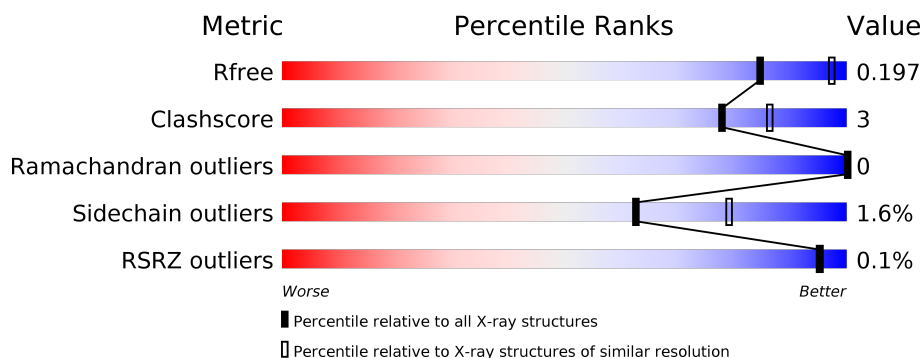
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.40 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))
$R_{free}$	130704	3907 (2.40-2.40)
Clashscore	141614	4398 (2.40-2.40)
Ramachandran outliers	138981	4318 (2.40-2.40)
Sidechain outliers	138945	4319 (2.40-2.40)
RSRZ outliers	127900	3811 (2.40-2.40)

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 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	255	<div> <div>82%</div> <div>6%</div> <div>12%</div> </div>
1	B	255	<div> <div>76%</div> <div>10%</div> <div>13%</div> </div>
1	C	255	<div> <div>80%</div> <div>7%</div> <div>13%</div> </div>
1	D	255	<div> <div>78%</div> <div>8%</div> <div>14%</div> </div>
1	E	255	<div> <div>80%</div> <div>7%</div> <div>13%</div> </div>
1	F	255	<div> <div>83%</div> <div>5%</div> <div>12%</div> </div>

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 11171 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 hypothetical protein TM0559.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	224	Total	C	N	O	S	Se	0	0	0
			1835	1183	314	333	1	4			
1	B	221	Total	C	N	O	S	Se	0	0	0
			1816	1172	308	331	1	4			
1	C	222	Total	C	N	O	S	Se	0	0	0
			1810	1170	305	330	1	4			
1	D	220	Total	C	N	O	S	Se	0	0	0
			1813	1169	309	330	1	4			
1	E	223	Total	C	N	O	S	Se	0	0	0
			1819	1177	309	328	1	4			
1	F	224	Total	C	N	O	S	Se	0	0	0
			1827	1182	310	330	1	4			

There are 102 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-11	MSE	-	MODIFIED RESIDUE	UNP Q9WZ29
A	-10	GLY	-	LEADER SEQUENCE	UNP Q9WZ29
A	-9	SER	-	LEADER SEQUENCE	UNP Q9WZ29
A	-8	ASP	-	LEADER SEQUENCE	UNP Q9WZ29
A	-7	LYS	-	LEADER SEQUENCE	UNP Q9WZ29
A	-6	ILE	-	LEADER SEQUENCE	UNP Q9WZ29
A	-5	HIS	-	LEADER SEQUENCE	UNP Q9WZ29
A	-4	HIS	-	LEADER SEQUENCE	UNP Q9WZ29
A	-3	HIS	-	LEADER SEQUENCE	UNP Q9WZ29
A	-2	HIS	-	LEADER SEQUENCE	UNP Q9WZ29
A	-1	HIS	-	LEADER SEQUENCE	UNP Q9WZ29
A	0	HIS	-	LEADER SEQUENCE	UNP Q9WZ29
A	1	MSE	MET	MODIFIED RESIDUE	UNP Q9WZ29
A	18	MSE	MET	MODIFIED RESIDUE	UNP Q9WZ29
A	90	MSE	MET	MODIFIED RESIDUE	UNP Q9WZ29
A	159	MSE	MET	MODIFIED RESIDUE	UNP Q9WZ29
A	231	MSE	MET	MODIFIED RESIDUE	UNP Q9WZ29

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	-11	MSE	-	MODIFIED RESIDUE	UNP Q9WZ29
B	-10	GLY	-	LEADER SEQUENCE	UNP Q9WZ29
B	-9	SER	-	LEADER SEQUENCE	UNP Q9WZ29
B	-8	ASP	-	LEADER SEQUENCE	UNP Q9WZ29
B	-7	LYS	-	LEADER SEQUENCE	UNP Q9WZ29
B	-6	ILE	-	LEADER SEQUENCE	UNP Q9WZ29
B	-5	HIS	-	LEADER SEQUENCE	UNP Q9WZ29
B	-4	HIS	-	LEADER SEQUENCE	UNP Q9WZ29
B	-3	HIS	-	LEADER SEQUENCE	UNP Q9WZ29
B	-2	HIS	-	LEADER SEQUENCE	UNP Q9WZ29
B	-1	HIS	-	LEADER SEQUENCE	UNP Q9WZ29
B	0	HIS	-	LEADER SEQUENCE	UNP Q9WZ29
B	1	MSE	MET	MODIFIED RESIDUE	UNP Q9WZ29
B	18	MSE	MET	MODIFIED RESIDUE	UNP Q9WZ29
B	90	MSE	MET	MODIFIED RESIDUE	UNP Q9WZ29
B	159	MSE	MET	MODIFIED RESIDUE	UNP Q9WZ29
B	231	MSE	MET	MODIFIED RESIDUE	UNP Q9WZ29
C	-11	MSE	-	MODIFIED RESIDUE	UNP Q9WZ29
C	-10	GLY	-	LEADER SEQUENCE	UNP Q9WZ29
C	-9	SER	-	LEADER SEQUENCE	UNP Q9WZ29
C	-8	ASP	-	LEADER SEQUENCE	UNP Q9WZ29
C	-7	LYS	-	LEADER SEQUENCE	UNP Q9WZ29
C	-6	ILE	-	LEADER SEQUENCE	UNP Q9WZ29
C	-5	HIS	-	LEADER SEQUENCE	UNP Q9WZ29
C	-4	HIS	-	LEADER SEQUENCE	UNP Q9WZ29
C	-3	HIS	-	LEADER SEQUENCE	UNP Q9WZ29
C	-2	HIS	-	LEADER SEQUENCE	UNP Q9WZ29
C	-1	HIS	-	LEADER SEQUENCE	UNP Q9WZ29
C	0	HIS	-	LEADER SEQUENCE	UNP Q9WZ29
C	1	MSE	MET	MODIFIED RESIDUE	UNP Q9WZ29
C	18	MSE	MET	MODIFIED RESIDUE	UNP Q9WZ29
C	90	MSE	MET	MODIFIED RESIDUE	UNP Q9WZ29
C	159	MSE	MET	MODIFIED RESIDUE	UNP Q9WZ29
C	231	MSE	MET	MODIFIED RESIDUE	UNP Q9WZ29
D	-11	MSE	-	MODIFIED RESIDUE	UNP Q9WZ29
D	-10	GLY	-	LEADER SEQUENCE	UNP Q9WZ29
D	-9	SER	-	LEADER SEQUENCE	UNP Q9WZ29
D	-8	ASP	-	LEADER SEQUENCE	UNP Q9WZ29
D	-7	LYS	-	LEADER SEQUENCE	UNP Q9WZ29
D	-6	ILE	-	LEADER SEQUENCE	UNP Q9WZ29
D	-5	HIS	-	LEADER SEQUENCE	UNP Q9WZ29
D	-4	HIS	-	LEADER SEQUENCE	UNP Q9WZ29

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
D	-3	HIS	-	LEADER SEQUENCE	UNP Q9WZ29
D	-2	HIS	-	LEADER SEQUENCE	UNP Q9WZ29
D	-1	HIS	-	LEADER SEQUENCE	UNP Q9WZ29
D	0	HIS	-	LEADER SEQUENCE	UNP Q9WZ29
D	1	MSE	MET	MODIFIED RESIDUE	UNP Q9WZ29
D	18	MSE	MET	MODIFIED RESIDUE	UNP Q9WZ29
D	90	MSE	MET	MODIFIED RESIDUE	UNP Q9WZ29
D	159	MSE	MET	MODIFIED RESIDUE	UNP Q9WZ29
D	231	MSE	MET	MODIFIED RESIDUE	UNP Q9WZ29
E	-11	MSE	-	MODIFIED RESIDUE	UNP Q9WZ29
E	-10	GLY	-	LEADER SEQUENCE	UNP Q9WZ29
E	-9	SER	-	LEADER SEQUENCE	UNP Q9WZ29
E	-8	ASP	-	LEADER SEQUENCE	UNP Q9WZ29
E	-7	LYS	-	LEADER SEQUENCE	UNP Q9WZ29
E	-6	ILE	-	LEADER SEQUENCE	UNP Q9WZ29
E	-5	HIS	-	LEADER SEQUENCE	UNP Q9WZ29
E	-4	HIS	-	LEADER SEQUENCE	UNP Q9WZ29
E	-3	HIS	-	LEADER SEQUENCE	UNP Q9WZ29
E	-2	HIS	-	LEADER SEQUENCE	UNP Q9WZ29
E	-1	HIS	-	LEADER SEQUENCE	UNP Q9WZ29
E	0	HIS	-	LEADER SEQUENCE	UNP Q9WZ29
E	1	MSE	MET	MODIFIED RESIDUE	UNP Q9WZ29
E	18	MSE	MET	MODIFIED RESIDUE	UNP Q9WZ29
E	90	MSE	MET	MODIFIED RESIDUE	UNP Q9WZ29
E	159	MSE	MET	MODIFIED RESIDUE	UNP Q9WZ29
E	231	MSE	MET	MODIFIED RESIDUE	UNP Q9WZ29
F	-11	MSE	-	MODIFIED RESIDUE	UNP Q9WZ29
F	-10	GLY	-	LEADER SEQUENCE	UNP Q9WZ29
F	-9	SER	-	LEADER SEQUENCE	UNP Q9WZ29
F	-8	ASP	-	LEADER SEQUENCE	UNP Q9WZ29
F	-7	LYS	-	LEADER SEQUENCE	UNP Q9WZ29
F	-6	ILE	-	LEADER SEQUENCE	UNP Q9WZ29
F	-5	HIS	-	LEADER SEQUENCE	UNP Q9WZ29
F	-4	HIS	-	LEADER SEQUENCE	UNP Q9WZ29
F	-3	HIS	-	LEADER SEQUENCE	UNP Q9WZ29
F	-2	HIS	-	LEADER SEQUENCE	UNP Q9WZ29
F	-1	HIS	-	LEADER SEQUENCE	UNP Q9WZ29
F	0	HIS	-	LEADER SEQUENCE	UNP Q9WZ29
F	1	MSE	MET	MODIFIED RESIDUE	UNP Q9WZ29
F	18	MSE	MET	MODIFIED RESIDUE	UNP Q9WZ29
F	90	MSE	MET	MODIFIED RESIDUE	UNP Q9WZ29
F	159	MSE	MET	MODIFIED RESIDUE	UNP Q9WZ29

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
F	231	MSE	MET	MODIFIED RESIDUE	UNP Q9WZ29

- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	D	4	Total Zn 4 4	0	0
2	E	4	Total Zn 4 4	0	0
2	B	4	Total Zn 4 4	0	0
2	C	4	Total Zn 4 4	0	0
2	A	4	Total Zn 4 4	0	0
2	F	4	Total Zn 4 4	0	0

- Molecule 3 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	D	2	Total Cl 2 2	0	0
3	E	3	Total Cl 3 3	0	0
3	B	2	Total Cl 2 2	0	0
3	C	3	Total Cl 3 3	0	0
3	A	3	Total Cl 3 3	0	0
3	F	2	Total Cl 2 2	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	44	Total O 44 44	0	0
4	B	31	Total O 31 31	0	0

*Continued on next page...*

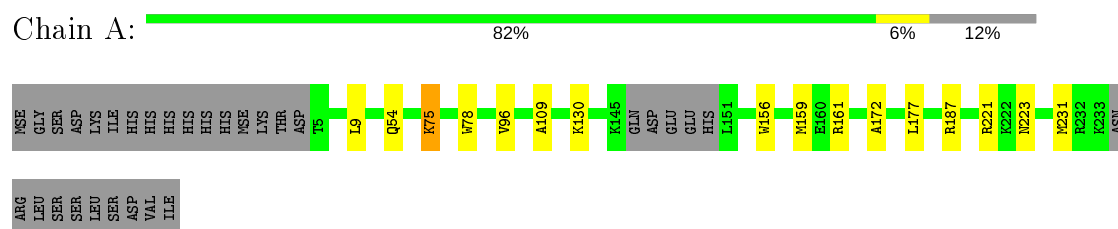
*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	C	22	Total 22	O 22	0	0
4	D	30	Total 30	O 30	0	0
4	E	46	Total 46	O 46	0	0
4	F	39	Total 39	O 39	0	0

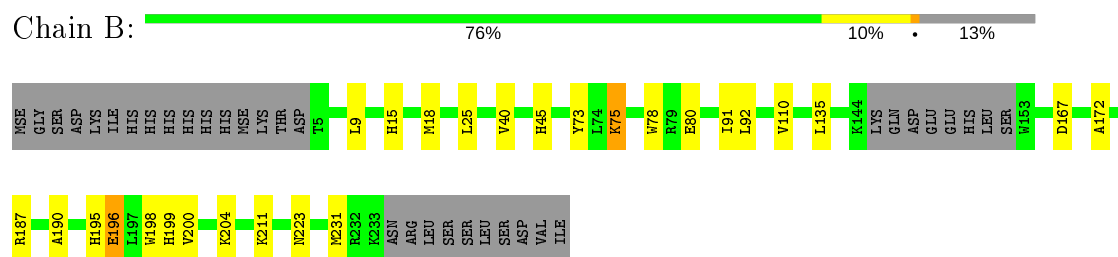
### 3 Residue-property plots [i](#)

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.

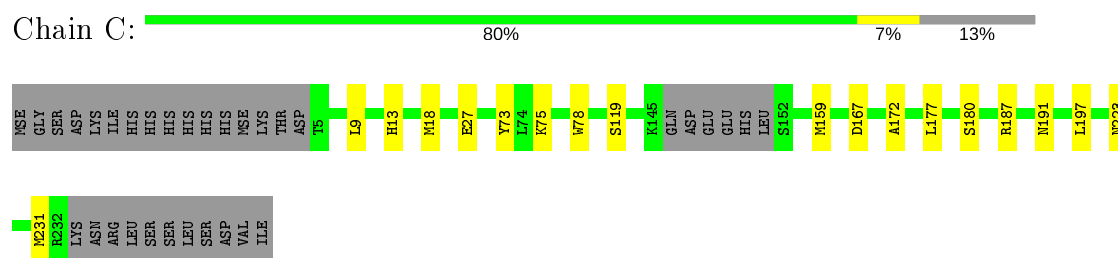
- Molecule 1: hypothetical protein TM0559



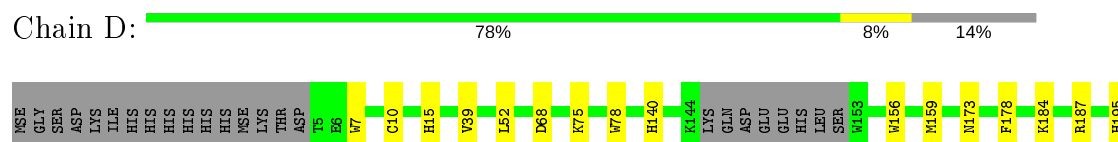
- Molecule 1: hypothetical protein TM0559



- Molecule 1: hypothetical protein TM0559



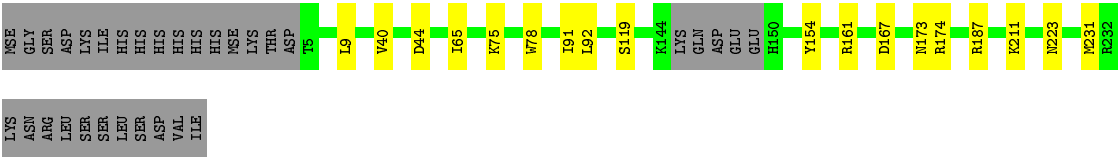
- Molecule 1: hypothetical protein TM0559



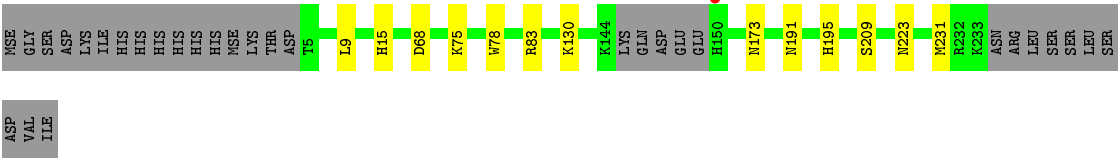
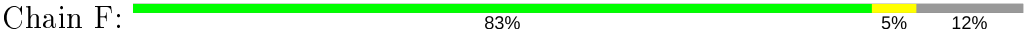




- Molecule 1: hypothetical protein TM0559



- Molecule 1: hypothetical protein TM0559



## 4 Data and refinement statistics

Property	Value	Source
Space group	H 3	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	111.31Å 111.31Å 383.22Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	29.00 – 2.40 29.00 – 2.25	Depositor EDS
% Data completeness (in resolution range)	100.0 (29.00-2.40) 99.3 (29.00-2.25)	Depositor EDS
$R_{merge}$	0.13	Depositor
$R_{sym}$	0.13	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.80 (at 2.24Å)	Xtriage
Refinement program	REFMAC 5.2.0005	Depositor
R, $R_{free}$	0.169 , 0.220 0.181 , 0.197	Depositor DCC
$R_{free}$ test set	3506 reflections (4.20%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	29.7	Xtriage
Anisotropy	0.154	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 30.2	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	0.148 for -h-k,k,-l	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	11171	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	32.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.71% 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](#)

Bond lengths and bond angles in the following residue types are not validated in this section: ZN, 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	0.66	0/1876	0.72	0/2536
1	B	0.67	0/1856	0.68	0/2511
1	C	0.62	0/1851	0.66	0/2503
1	D	0.63	0/1854	0.70	0/2506
1	E	0.66	0/1860	0.72	0/2515
1	F	0.61	0/1869	0.66	0/2526
All	All	0.64	0/11166	0.69	0/15097

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	B	0	1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	45	HIS	Peptide

### 5.2 Too-close contacts [i](#)

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	1835	0	1786	8	0
1	B	1816	0	1760	18	0
1	C	1810	0	1750	10	0
1	D	1813	0	1764	10	0
1	E	1819	0	1769	12	0
1	F	1827	0	1764	5	0
2	A	4	0	0	0	0
2	B	4	0	0	0	0
2	C	4	0	0	0	0
2	D	4	0	0	0	0
2	E	4	0	0	0	0
2	F	4	0	0	0	0
3	A	3	0	0	1	0
3	B	2	0	0	0	0
3	C	3	0	0	0	0
3	D	2	0	0	0	0
3	E	3	0	0	0	0
3	F	2	0	0	0	0
4	A	44	0	0	1	0
4	B	31	0	0	0	0
4	C	22	0	0	0	0
4	D	30	0	0	1	0
4	E	46	0	0	0	0
4	F	39	0	0	0	0
All	All	11171	0	10593	63	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 3.

All (63) 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)
3:A:1001:CL:CL	4:A:1055:HOH:O	2.24	0.92
1:C:9:LEU:HD21	1:C:231:MSE:CE	2.06	0.86
1:E:9:LEU:HD21	1:E:231:MSE:CE	2.14	0.77
1:B:9:LEU:HD21	1:B:231:MSE:HE3	1.65	0.77
1:C:9:LEU:HD21	1:C:231:MSE:HE3	1.67	0.76
1:A:9:LEU:HD21	1:A:231:MSE:HE3	1.69	0.74
1:E:40:VAL:HG13	1:E:92:LEU:HD12	1.71	0.70
1:B:196:GLU:OE2	1:B:198:TRP:CZ2	2.49	0.66

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:161:ARG:HA	1:C:27:GLU:HG3	1.79	0.63
1:A:9:LEU:HD21	1:A:231:MSE:CE	2.28	0.61
1:E:75:LYS:HA	1:E:78:TRP:CE3	2.35	0.61
1:B:40:VAL:HG13	1:B:92:LEU:HD12	1.83	0.60
1:E:40:VAL:HG13	1:E:92:LEU:CD1	2.33	0.59
1:B:9:LEU:HD21	1:B:231:MSE:CE	2.33	0.58
1:E:9:LEU:HD21	1:E:231:MSE:HE3	1.86	0.58
1:E:167:ASP:O	1:E:187:ARG:HD2	2.05	0.56
1:F:9:LEU:HD21	1:F:231:MSE:HE3	1.87	0.55
1:F:75:LYS:HA	1:F:78:TRP:CE3	2.42	0.55
1:D:75:LYS:HA	1:D:78:TRP:CE3	2.43	0.54
1:D:187:ARG:NH2	4:D:1033:HOH:O	2.41	0.54
1:F:15:HIS:CD2	1:F:195:HIS:CE1	2.96	0.53
1:B:196:GLU:OE2	1:B:198:TRP:HZ2	1.91	0.53
1:A:75:LYS:HA	1:A:78:TRP:CE3	2.44	0.52
1:C:18:MSE:HE1	1:C:73:TYR:HA	1.91	0.52
1:E:91:ILE:HD11	1:E:211:LYS:HG2	1.91	0.52
1:D:202:SER:OG	1:D:204:LYS:HE2	2.10	0.51
1:D:159:MSE:HB2	1:D:184:LYS:HZ2	1.76	0.50
1:B:167:ASP:O	1:B:187:ARG:HD2	2.12	0.50
1:F:173:ASN:HA	1:F:191:ASN:O	2.12	0.49
1:E:9:LEU:HD21	1:E:231:MSE:HE1	1.92	0.49
1:A:187:ARG:HD3	1:A:221:ARG:O	2.13	0.48
1:B:75:LYS:HG2	1:B:78:TRP:CZ3	2.49	0.48
1:A:172:ALA:HB2	1:A:177:LEU:HD23	1.96	0.48
1:A:156:TRP:HE3	1:A:159:MSE:HE2	1.78	0.47
1:B:18:MSE:HE1	1:B:73:TYR:HA	1.96	0.47
1:D:173:ASN:HB2	1:D:178:PHE:HE2	1.81	0.46
1:E:154:TYR:OH	1:E:161:ARG:NH1	2.49	0.46
1:B:25:LEU:HG	1:B:80:GLU:HG3	1.97	0.45
1:C:197:LEU:HD23	1:C:197:LEU:O	2.16	0.45
1:B:75:LYS:HA	1:B:78:TRP:CE3	2.51	0.45
1:B:91:ILE:HD11	1:B:211:LYS:HG2	2.00	0.44
1:B:196:GLU:HG2	1:B:199:HIS:CE1	2.53	0.44
1:E:44:ASP:HB3	1:E:65:ILE:HD11	2.00	0.43
1:A:96:VAL:HG12	1:A:109:ALA:HB3	2.01	0.43
1:B:15:HIS:CD2	1:B:195:HIS:CE1	3.07	0.43
1:D:15:HIS:CD2	1:D:195:HIS:CE1	3.07	0.42
1:D:10:CYS:HA	1:D:39:VAL:O	2.18	0.42
1:B:9:LEU:CD2	1:B:231:MSE:CE	2.97	0.42
1:C:172:ALA:HB2	1:C:177:LEU:HD23	2.02	0.42

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:75:LYS:HA	1:C:78:TRP:CE3	2.54	0.42
1:E:44:ASP:CB	1:E:65:ILE:HD11	2.50	0.42
1:B:110:VAL:HB	1:B:135:LEU:HB3	2.01	0.42
1:F:9:LEU:HD21	1:F:231:MSE:CE	2.50	0.42
1:B:25:LEU:CD2	1:B:80:GLU:HG3	2.49	0.42
1:C:159:MSE:HE1	1:C:180:SER:HB2	2.03	0.41
1:D:140:HIS:HB2	1:D:156:TRP:CZ2	2.56	0.41
1:D:52:LEU:HA	1:D:52:LEU:HD23	1.96	0.41
1:C:167:ASP:O	1:C:187:ARG:HD2	2.21	0.41
1:C:13:HIS:HB3	1:C:191:ASN:HB3	2.03	0.41
1:B:172:ALA:HB3	1:B:190:ALA:HB1	2.01	0.41
1:E:173:ASN:O	1:E:174:ARG:C	2.59	0.41
1:D:7:TRP:HB3	1:D:206:LEU:HG	2.03	0.40
1:B:200:VAL:O	1:B:204:LYS:NZ	2.52	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	220/255 (86%)	210 (96%)	10 (4%)	0	100	100
1	B	217/255 (85%)	209 (96%)	8 (4%)	0	100	100
1	C	218/255 (86%)	207 (95%)	11 (5%)	0	100	100
1	D	216/255 (85%)	207 (96%)	9 (4%)	0	100	100
1	E	219/255 (86%)	211 (96%)	8 (4%)	0	100	100
1	F	220/255 (86%)	210 (96%)	10 (4%)	0	100	100
All	All	1310/1530 (86%)	1254 (96%)	56 (4%)	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	192/226 (85%)	188 (98%)	4 (2%)	53	72
1	B	189/226 (84%)	186 (98%)	3 (2%)	62	79
1	C	190/226 (84%)	188 (99%)	2 (1%)	73	87
1	D	190/226 (84%)	188 (99%)	2 (1%)	73	87
1	E	188/226 (83%)	186 (99%)	2 (1%)	73	87
1	F	188/226 (83%)	183 (97%)	5 (3%)	44	65
All	All	1137/1356 (84%)	1119 (98%)	18 (2%)	62	79

All (18) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	54	GLN
1	A	75	LYS
1	A	130	LYS
1	A	223	ASN
1	B	75	LYS
1	B	196	GLU
1	B	223	ASN
1	C	119	SER
1	C	223	ASN
1	D	68	ASP
1	D	223	ASN
1	E	119	SER
1	E	223	ASN
1	F	68	ASP
1	F	83	ARG
1	F	130	LYS
1	F	209	SER
1	F	223	ASN

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

Mol	Chain	Res	Type
1	E	54	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 39 ligands modelled in this entry, 39 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 [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, 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(Å²)	Q<0.9	
1	A	220/255 (86%)	-0.59	0	100	100	25, 31, 42, 54	0
1	B	217/255 (85%)	-0.51	0	100	100	24, 31, 42, 54	0
1	C	218/255 (85%)	-0.43	0	100	100	24, 31, 40, 63	0
1	D	216/255 (84%)	-0.48	0	100	100	25, 31, 41, 54	0
1	E	219/255 (85%)	-0.58	0	100	100	24, 31, 42, 55	0
1	F	220/255 (86%)	-0.40	1 (0%)	91	89	23, 30, 40, 59	0
All	All	1310/1530 (85%)	-0.50	1 (0%)	95	95	23, 31, 41, 63	0

All (1) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	F	150	HIS	2.1

### 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. 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( $\text{\AA}^2$ )	Q<0.9
2	ZN	C	705	1/1	0.95	0.03	73,73,73,73	0
3	CL	C	1006	1/1	0.95	0.10	44,44,44,44	0
3	CL	E	1010	1/1	0.96	0.08	32,32,32,32	0
2	ZN	B	605	1/1	0.97	0.05	62,62,62,62	0
2	ZN	A	705	1/1	0.97	0.07	67,67,67,67	0
2	ZN	B	705	1/1	0.97	0.05	79,79,79,79	0
3	CL	B	1004	1/1	0.97	0.04	41,41,41,41	0
2	ZN	F	705	1/1	0.97	0.04	61,61,61,61	0
3	CL	D	1008	1/1	0.97	0.12	40,40,40,40	0
3	CL	B	1003	1/1	0.97	0.14	27,27,27,27	0
3	CL	C	1005	1/1	0.97	0.16	31,31,31,31	0
3	CL	A	1002	1/1	0.98	0.10	33,33,33,33	0
2	ZN	E	705	1/1	0.98	0.07	65,65,65,65	0
2	ZN	E	405	1/1	0.98	0.11	37,37,37,37	0
3	CL	F	1011	1/1	0.98	0.15	24,24,24,24	0
3	CL	E	1009	1/1	0.98	0.14	25,25,25,25	0
2	ZN	D	405	1/1	0.98	0.09	40,40,40,40	0
3	CL	A	1001	1/1	0.98	0.18	25,25,25,25	0
3	CL	F	1012	1/1	0.98	0.08	33,33,33,33	0
2	ZN	A	605	1/1	0.99	0.07	46,46,46,46	0
2	ZN	D	605	1/1	0.99	0.06	53,53,53,53	0
2	ZN	E	605	1/1	0.99	0.04	47,47,47,47	0
2	ZN	F	505	1/1	0.99	0.07	31,31,31,31	0
3	CL	D	1007	1/1	0.99	0.14	31,31,31,31	0
2	ZN	E	505	1/1	0.99	0.05	30,30,30,30	0
2	ZN	A	405	1/1	0.99	0.05	32,32,32,32	0
2	ZN	A	505	1/1	0.99	0.03	29,29,29,29	0
2	ZN	D	705	1/1	0.99	0.05	73,73,73,73	0
2	ZN	C	405	1/1	0.99	0.06	42,42,42,42	0
2	ZN	F	405	1/1	0.99	0.08	34,34,34,34	0
2	ZN	B	405	1/1	0.99	0.04	39,39,39,39	0
2	ZN	C	505	1/1	0.99	0.09	33,33,33,33	0
2	ZN	F	605	1/1	0.99	0.05	41,41,41,41	0
2	ZN	C	605	1/1	1.00	0.03	54,54,54,54	0
3	CL	A	1013	1/1	1.00	0.03	24,24,24,24	0
2	ZN	B	505	1/1	1.00	0.06	30,30,30,30	0
3	CL	C	1014	1/1	1.00	0.04	24,24,24,24	0
3	CL	E	1015	1/1	1.00	0.02	19,19,19,19	0
2	ZN	D	505	1/1	1.00	0.06	35,35,35,35	0

## 6.5 Other polymers ⓘ

There are no such residues in this entry.