



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 13, 2020 – 03:06 am BST

PDB ID : 5JTV  
Title : USP7CD-UBL45 in complex with Ubiquitin  
Authors : Murray, J.M.; Rouge, L.  
Deposited on : 2016-05-09  
Resolution : 3.31 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
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

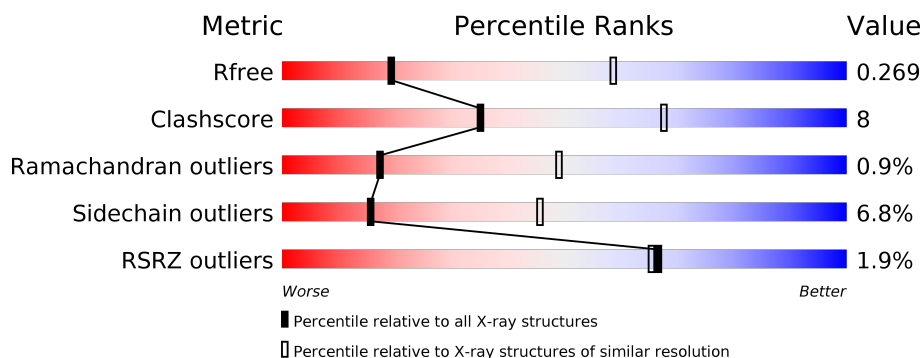
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 3.31 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	1089 (3.36-3.28)
Clashscore	141614	1137 (3.36-3.28)
Ramachandran outliers	138981	1115 (3.36-3.28)
Sidechain outliers	138945	1114 (3.36-3.28)
RSRZ outliers	127900	1059 (3.36-3.28)

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	594	<div> <div>2%</div> <div> <div></div> <div>70%</div> <div>23%</div> <div>• 5%</div> </div> </div>
1	C	594	<div> <div>%</div> <div> <div></div> <div>74%</div> <div>19%</div> <div>• 5%</div> </div> </div>
1	E	594	<div> <div>2%</div> <div> <div></div> <div>72%</div> <div>20%</div> <div>• 5%</div> </div> </div>
1	G	594	<div> <div>2%</div> <div> <div></div> <div>72%</div> <div>20%</div> <div>• 5%</div> </div> </div>
2	B	76	<div> <div></div> <div> <div>84%</div> <div>16%</div> </div> </div>
2	D	76	<div> <div></div> <div> <div>84%</div> <div>16%</div> </div> </div>

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	F	76	<div><div></div><div>4%</div><div>89%</div><div>11%</div></div>
2	H	76	<div><div></div><div>68%</div><div>30%</div><div></div></div>

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 20520 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 Ubiquitin carboxyl-terminal hydrolase 7.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	564	Total	C	N	O	S	2	0	0
			4562	2899	784	854	25			
1	C	564	Total	C	N	O	S	26	0	0
			4533	2878	784	846	25			
1	E	565	Total	C	N	O	S	11	0	0
			4530	2881	778	846	25			
1	G	563	Total	C	N	O	S	0	0	0
			4507	2861	776	845	25			

There are 100 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	192	MET	-	initiating methionine	UNP Q93009
A	193	HIS	-	expression tag	UNP Q93009
A	194	HIS	-	expression tag	UNP Q93009
A	195	HIS	-	expression tag	UNP Q93009
A	196	HIS	-	expression tag	UNP Q93009
A	197	HIS	-	expression tag	UNP Q93009
A	198	HIS	-	expression tag	UNP Q93009
A	199	GLY	-	expression tag	UNP Q93009
A	200	GLU	-	expression tag	UNP Q93009
A	201	ASN	-	expression tag	UNP Q93009
A	202	LEU	-	expression tag	UNP Q93009
A	203	TYR	-	expression tag	UNP Q93009
A	204	PHE	-	expression tag	UNP Q93009
A	205	GLN	-	expression tag	UNP Q93009
A	206	GLY	-	expression tag	UNP Q93009
A	872	GLY	-	linker	UNP Q93009
A	873	GLY	-	linker	UNP Q93009
A	874	SER	-	linker	UNP Q93009
A	875	GLY	-	linker	UNP Q93009
A	876	GLY	-	linker	UNP Q93009
A	877	SER	-	linker	UNP Q93009

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	878	GLY	-	linker	UNP Q93009
A	879	GLY	-	linker	UNP Q93009
A	880	SER	-	linker	UNP Q93009
A	881	GLY	-	linker	UNP Q93009
C	192	MET	-	initiating methionine	UNP Q93009
C	193	HIS	-	expression tag	UNP Q93009
C	194	HIS	-	expression tag	UNP Q93009
C	195	HIS	-	expression tag	UNP Q93009
C	196	HIS	-	expression tag	UNP Q93009
C	197	HIS	-	expression tag	UNP Q93009
C	198	HIS	-	expression tag	UNP Q93009
C	199	GLY	-	expression tag	UNP Q93009
C	200	GLU	-	expression tag	UNP Q93009
C	201	ASN	-	expression tag	UNP Q93009
C	202	LEU	-	expression tag	UNP Q93009
C	203	TYR	-	expression tag	UNP Q93009
C	204	PHE	-	expression tag	UNP Q93009
C	205	GLN	-	expression tag	UNP Q93009
C	206	GLY	-	expression tag	UNP Q93009
C	872	GLY	-	linker	UNP Q93009
C	873	GLY	-	linker	UNP Q93009
C	874	SER	-	linker	UNP Q93009
C	875	GLY	-	linker	UNP Q93009
C	876	GLY	-	linker	UNP Q93009
C	877	SER	-	linker	UNP Q93009
C	878	GLY	-	linker	UNP Q93009
C	879	GLY	-	linker	UNP Q93009
C	880	SER	-	linker	UNP Q93009
C	881	GLY	-	linker	UNP Q93009
E	192	MET	-	initiating methionine	UNP Q93009
E	193	HIS	-	expression tag	UNP Q93009
E	194	HIS	-	expression tag	UNP Q93009
E	195	HIS	-	expression tag	UNP Q93009
E	196	HIS	-	expression tag	UNP Q93009
E	197	HIS	-	expression tag	UNP Q93009
E	198	HIS	-	expression tag	UNP Q93009
E	199	GLY	-	expression tag	UNP Q93009
E	200	GLU	-	expression tag	UNP Q93009
E	201	ASN	-	expression tag	UNP Q93009
E	202	LEU	-	expression tag	UNP Q93009
E	203	TYR	-	expression tag	UNP Q93009
E	204	PHE	-	expression tag	UNP Q93009

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
E	205	GLN	-	expression tag	UNP Q93009
E	206	GLY	-	expression tag	UNP Q93009
E	872	GLY	-	linker	UNP Q93009
E	873	GLY	-	linker	UNP Q93009
E	874	SER	-	linker	UNP Q93009
E	875	GLY	-	linker	UNP Q93009
E	876	GLY	-	linker	UNP Q93009
E	877	SER	-	linker	UNP Q93009
E	878	GLY	-	linker	UNP Q93009
E	879	GLY	-	linker	UNP Q93009
E	880	SER	-	linker	UNP Q93009
E	881	GLY	-	linker	UNP Q93009
G	192	MET	-	initiating methionine	UNP Q93009
G	193	HIS	-	expression tag	UNP Q93009
G	194	HIS	-	expression tag	UNP Q93009
G	195	HIS	-	expression tag	UNP Q93009
G	196	HIS	-	expression tag	UNP Q93009
G	197	HIS	-	expression tag	UNP Q93009
G	198	HIS	-	expression tag	UNP Q93009
G	199	GLY	-	expression tag	UNP Q93009
G	200	GLU	-	expression tag	UNP Q93009
G	201	ASN	-	expression tag	UNP Q93009
G	202	LEU	-	expression tag	UNP Q93009
G	203	TYR	-	expression tag	UNP Q93009
G	204	PHE	-	expression tag	UNP Q93009
G	205	GLN	-	expression tag	UNP Q93009
G	206	GLY	-	expression tag	UNP Q93009
G	872	GLY	-	linker	UNP Q93009
G	873	GLY	-	linker	UNP Q93009
G	874	SER	-	linker	UNP Q93009
G	875	GLY	-	linker	UNP Q93009
G	876	GLY	-	linker	UNP Q93009
G	877	SER	-	linker	UNP Q93009
G	878	GLY	-	linker	UNP Q93009
G	879	GLY	-	linker	UNP Q93009
G	880	SER	-	linker	UNP Q93009
G	881	GLY	-	linker	UNP Q93009

- Molecule 2 is a protein called Polyubiquitin-B.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	76	Total	C	N	O	S	0	0	0
			597	375	104	117	1			

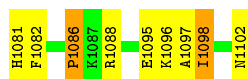
*Continued on next page...*

*Continued from previous page...*

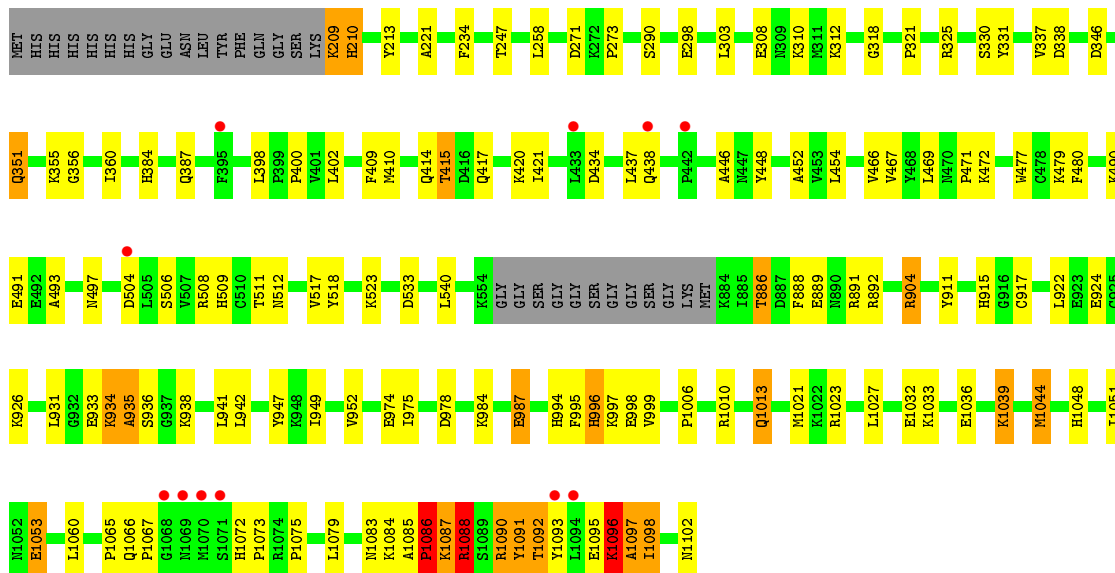
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	D	76	Total	C	N	O	S	0	0	0
			597	375	104	117	1			
2	F	76	Total	C	N	O	S	0	0	0
			597	375	104	117	1			
2	H	76	Total	C	N	O	S	0	0	0
			597	375	104	117	1			



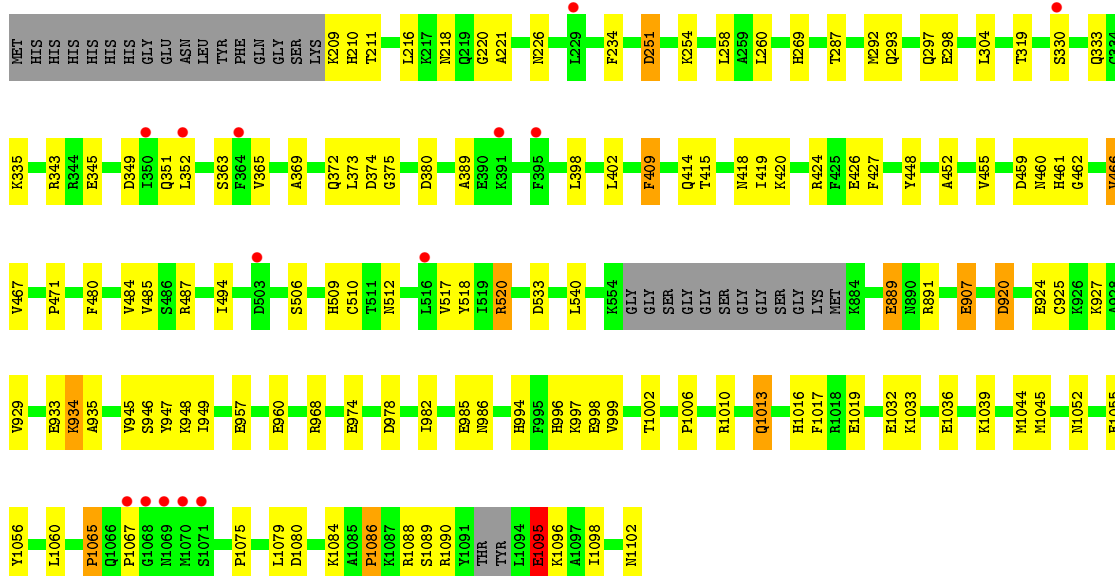
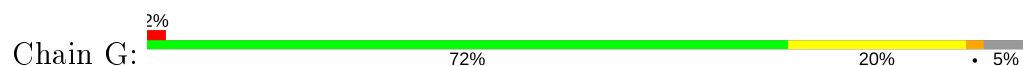




• Molecule 1: Ubiquitin carboxyl-terminal hydrolase 7



• Molecule 1: Ubiquitin carboxyl-terminal hydrolase 7

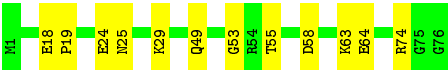
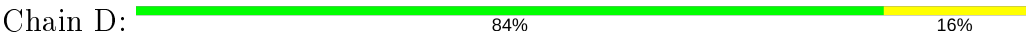


• Molecule 2: Polyubiquitin-B

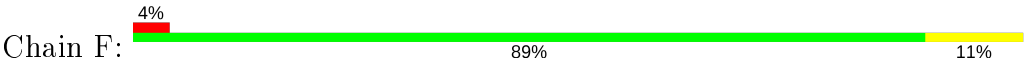




● Molecule 2: Polyubiquitin-B



● Molecule 2: Polyubiquitin-B



● Molecule 2: Polyubiquitin-B



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	113.61Å 115.52Å 257.82Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	46.95 – 3.31 46.95 – 3.31	Depositor EDS
% Data completeness (in resolution range)	99.4 (46.95-3.31) 92.6 (46.95-3.31)	Depositor EDS
$R_{merge}$	0.13	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.01 (at 3.32Å)	Xtriage
Refinement program	PHENIX dev_2210	Depositor
R, $R_{free}$	0.239 , 0.269 0.239 , 0.269	Depositor DCC
$R_{free}$ test set	2596 reflections (5.10%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	94.0	Xtriage
Anisotropy	0.346	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.26 , 28.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.29$	Xtriage
Estimated twinning fraction	0.036 for k,h,-l	Xtriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	20520	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	114.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.77% of the height of the origin peak. No significant pseudotranslation is detected.*

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

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

## 5 Model quality

### 5.1 Standard geometry

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	0.34	3/4663 (0.1%)	0.53	1/6290 (0.0%)
1	C	0.26	0/4634	0.49	1/6254 (0.0%)
1	E	0.28	0/4630	0.54	4/6249 (0.1%)
1	G	0.29	0/4605	0.54	4/6215 (0.1%)
2	B	0.32	0/603	0.68	1/812 (0.1%)
2	D	0.27	0/603	0.52	0/812
2	F	0.28	0/603	0.60	0/812
2	H	0.30	0/603	0.67	1/812 (0.1%)
All	All	0.29	3/20944 (0.0%)	0.54	12/28256 (0.0%)

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	E	0	2
1	G	0	2
All	All	0	5

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	498	TYR	CD2-CE2	-7.40	1.28	1.39
1	A	498	TYR	CD1-CE1	-6.54	1.29	1.39
1	A	498	TYR	CE2-CZ	-5.01	1.32	1.38

The worst 5 of 12 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	G	506	SER	CB-CA-C	13.36	135.49	110.10
1	E	1088	ARG	N-CA-C	8.57	134.13	111.00

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	E	1096	LYS	CD-CE-NZ	-8.38	92.44	111.70
2	B	71	LEU	CB-CA-C	-7.57	95.81	110.20
1	G	996	HIS	N-CA-C	-7.00	92.11	111.00

There are no chirality outliers.

All (5) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	504	ASP	Peptide
1	E	1092	THR	Peptide
1	E	504	ASP	Peptide
1	G	1090	ARG	Peptide
1	G	1095	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	4562	0	4423	91	1
1	C	4533	0	4370	71	0
1	E	4530	0	4366	86	0
1	G	4507	0	4335	88	1
2	B	597	0	618	12	0
2	D	597	0	618	8	0
2	F	597	0	618	12	0
2	H	597	0	618	21	0
All	All	20520	0	19966	330	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 8.

The worst 5 of 330 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)
2:B:74:ARG:NH1	1:E:1096:LYS:HB2	1.64	1.12

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:1096:LYS:HB3	2:F:74:ARG:NH1	1.69	1.08
1:C:292:MET:HE1	1:G:1095:GLU:O	1.59	1.03
1:E:934:LYS:HZ3	1:E:934:LYS:H	1.03	1.01
1:C:292:MET:CE	1:G:1095:GLU:O	2.09	1.00

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)
1:A:383:GLU:OE2	1:G:287:THR:OG1[4_545]	2.08	0.12

## 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	560/594 (94%)	523 (93%)	33 (6%)	4 (1%)	22	55
1	C	560/594 (94%)	522 (93%)	31 (6%)	7 (1%)	12	40
1	E	561/594 (94%)	516 (92%)	37 (7%)	8 (1%)	11	39
1	G	557/594 (94%)	517 (93%)	37 (7%)	3 (0%)	29	61
2	B	74/76 (97%)	74 (100%)	0	0	100	100
2	D	74/76 (97%)	74 (100%)	0	0	100	100
2	F	74/76 (97%)	74 (100%)	0	0	100	100
2	H	74/76 (97%)	74 (100%)	0	0	100	100
All	All	2534/2680 (95%)	2374 (94%)	138 (5%)	22 (1%)	17	49

5 of 22 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	1093	TYR

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	1097	ALA
1	E	1087	LYS
1	C	475	GLY
1	C	1096	LYS

### 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	494/535 (92%)	453 (92%)	41 (8%)	11	37
1	C	487/535 (91%)	447 (92%)	40 (8%)	11	37
1	E	484/535 (90%)	448 (93%)	36 (7%)	13	41
1	G	483/535 (90%)	453 (94%)	30 (6%)	18	49
2	B	67/68 (98%)	65 (97%)	2 (3%)	41	69
2	D	67/68 (98%)	67 (100%)	0	100	100
2	F	67/68 (98%)	67 (100%)	0	100	100
2	H	67/68 (98%)	66 (98%)	1 (2%)	65	81
All	All	2216/2412 (92%)	2066 (93%)	150 (7%)	16	46

5 of 150 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	454	LEU
1	E	209	LYS
1	G	934	LYS
1	C	484	VAL
1	C	945	VAL

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	E	387	GLN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	E	915	HIS
1	G	915	HIS
1	E	230	GLN
1	G	1012	HIS

### 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](#)

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	564/594 (94%)	-0.05	14 (2%) 57 55	91, 114, 145, 166	1 (0%)
1	C	562/594 (94%)	-0.12	7 (1%) 79 79	88, 109, 137, 161	1 (0%)
1	E	565/594 (95%)	-0.13	11 (1%) 66 65	89, 110, 144, 168	2 (0%)
1	G	563/594 (94%)	-0.10	14 (2%) 57 55	88, 114, 153, 181	0
2	B	76/76 (100%)	-0.07	0 100 100	94, 105, 115, 123	0
2	D	76/76 (100%)	-0.08	0 100 100	89, 104, 113, 118	0
2	F	76/76 (100%)	-0.01	3 (3%) 39 38	103, 124, 139, 155	0
2	H	76/76 (100%)	-0.12	0 100 100	105, 135, 152, 162	0
All	All	2558/2680 (95%)	-0.10	49 (1%) 66 65	88, 112, 145, 181	4 (0%)

The worst 5 of 49 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	1094	LEU	11.2
1	E	504	ASP	6.6
1	E	1094	LEU	5.2
1	G	1070	MET	5.0
1	A	504	ASP	4.8

### 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 [i](#)

There are no ligands in this entry.

## 6.5 Other polymers [i](#)

There are no such residues in this entry.