



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 14, 2017 – 11:09 pm GMT

PDB ID : 5JHF  
Title : Crystal structure of Atg13(17BR)-Atg13(17LR)-Atg17-Atg29-Atg31 complex  
Authors : Fujioka, Y.; Noda, N.N.  
Deposited on : 2016-04-21  
Resolution : 3.21 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)

A user guide is available at

<http://wwpdb.org/validation/2016/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

---

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

MolProbity : 4.02b-467  
Xtriage (Phenix) : 1.9-1692  
EDS : trunk28683  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : recalc28949

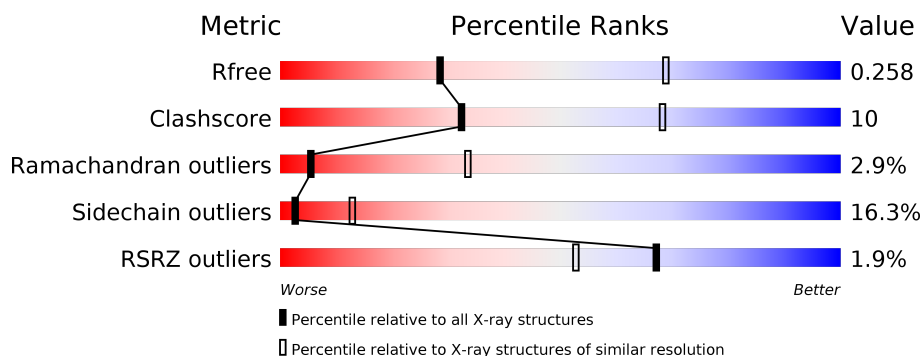
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.21 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	100719	1036 (3.24-3.20)
Clashscore	112137	1161 (3.24-3.20)
Ramachandran outliers	110173	1140 (3.24-3.20)
Sidechain outliers	110143	1139 (3.24-3.20)
RSRZ outliers	101464	1040 (3.24-3.20)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions  $\leq 5\%$ . The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	87	<div> <div>3%</div> <div> <div></div> <div>55%</div> <div>29%</div> <div>•</div> <div>13%</div> </div> </div>
1	D	87	<div> <div>3%</div> <div> <div></div> <div>71%</div> <div>22%</div> <div>•</div> <div>6%</div> </div> </div>
2	B	151	<div> <div></div> <div> <div></div> <div>56%</div> <div>25%</div> <div>7%</div> <div>•</div> <div>11%</div> </div> </div>
2	E	151	<div> <div>%</div> <div> <div></div> <div>54%</div> <div>23%</div> <div>7%</div> <div>•</div> <div>15%</div> </div> </div>
3	C	413	<div> <div>%</div> <div> <div></div> <div>67%</div> <div>25%</div> <div>5%</div> <div>•</div> </div> </div>
3	F	413	<div> <div>3%</div> <div> <div></div> <div>74%</div> <div>20%</div> <div>•</div> </div> </div>

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
4	G	13	<div><div></div><div></div><div></div><div>31%</div><div>31%</div><div>38%</div></div>
4	H	13	<div><div></div><div></div><div></div><div>38%</div><div>8%</div><div>54%</div></div>
5	I	13	<div><div></div><div></div><div></div><div>15%</div><div>62%</div><div>8%</div><div>31%</div></div>
5	J	13	<div><div></div><div></div><div></div><div>54%</div><div>8%</div><div>8%</div><div>31%</div></div>

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 9843 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 KLTH0D11660p.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	76	Total	C	N	O	S	0	0	0
			583	377	104	101	1			
1	D	82	Total	C	N	O	S	0	0	0
			596	378	111	106	1			

- Molecule 2 is a protein called KLTH0C07942p.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	135	Total	C	N	O	S	0	0	0
			1044	659	176	207	2			
2	E	128	Total	C	N	O	S	0	0	0
			1005	634	165	204	2			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	146	HIS	-	expression tag	UNP C5DEB9
B	147	HIS	-	expression tag	UNP C5DEB9
B	148	HIS	-	expression tag	UNP C5DEB9
B	149	HIS	-	expression tag	UNP C5DEB9
B	150	HIS	-	expression tag	UNP C5DEB9
B	151	HIS	-	expression tag	UNP C5DEB9
E	146	HIS	-	expression tag	UNP C5DEB9
E	147	HIS	-	expression tag	UNP C5DEB9
E	148	HIS	-	expression tag	UNP C5DEB9
E	149	HIS	-	expression tag	UNP C5DEB9
E	150	HIS	-	expression tag	UNP C5DEB9
E	151	HIS	-	expression tag	UNP C5DEB9

- Molecule 3 is a protein called KLTH0D15642p.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	404	Total	C	N	O	S	0	0	0
			3206	2012	556	625	13			
3	F	405	Total	C	N	O	S	0	0	0
			3200	2004	557	627	12			

- Molecule 4 is a protein called Atg13 17BR.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
4	G	8	Total	C	N	O	0	0	0
			57	35	11	11			
4	H	6	Total	C	N	O	0	0	0
			40	26	6	8			

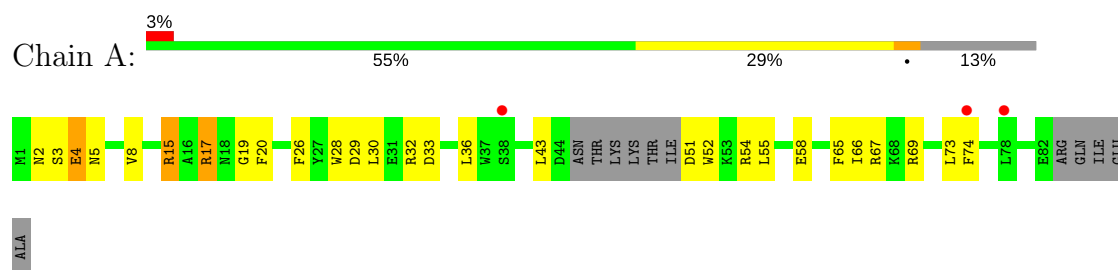
- Molecule 5 is a protein called Atg13 17LR.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
5	I	9	Total	C	N	O	0	0	0
			54	35	9	10			
5	J	9	Total	C	N	O	0	0	0
			58	38	10	10			

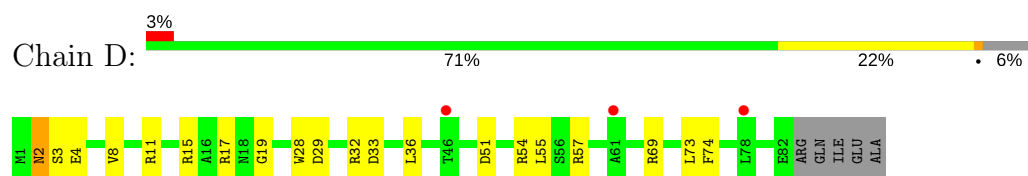
### 3 Residue-property plots [i](#)

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.

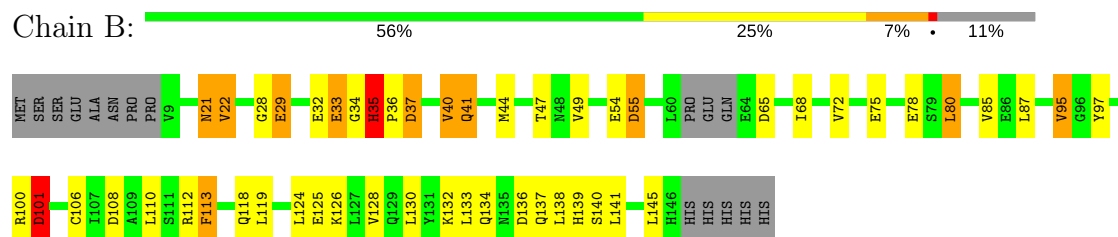
#### • Molecule 1: KLTH0D11660p



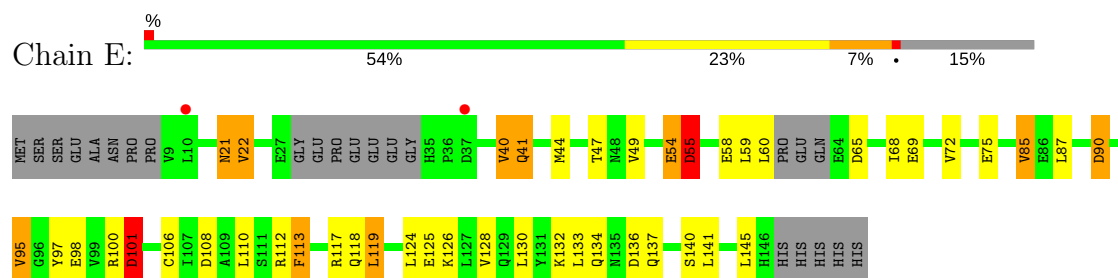
#### • Molecule 1: KLTH0D11660p



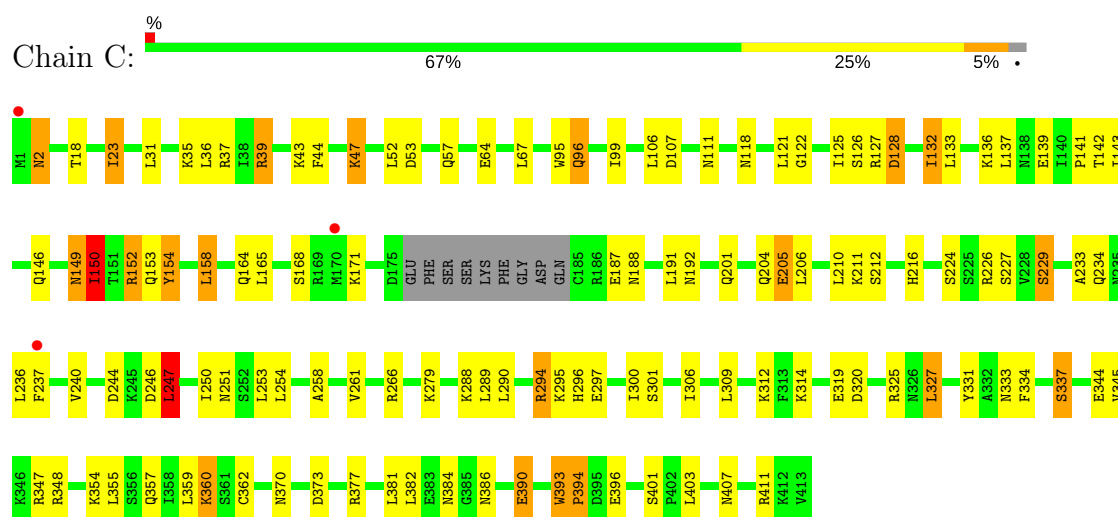
#### • Molecule 2: KLTH0C07942p



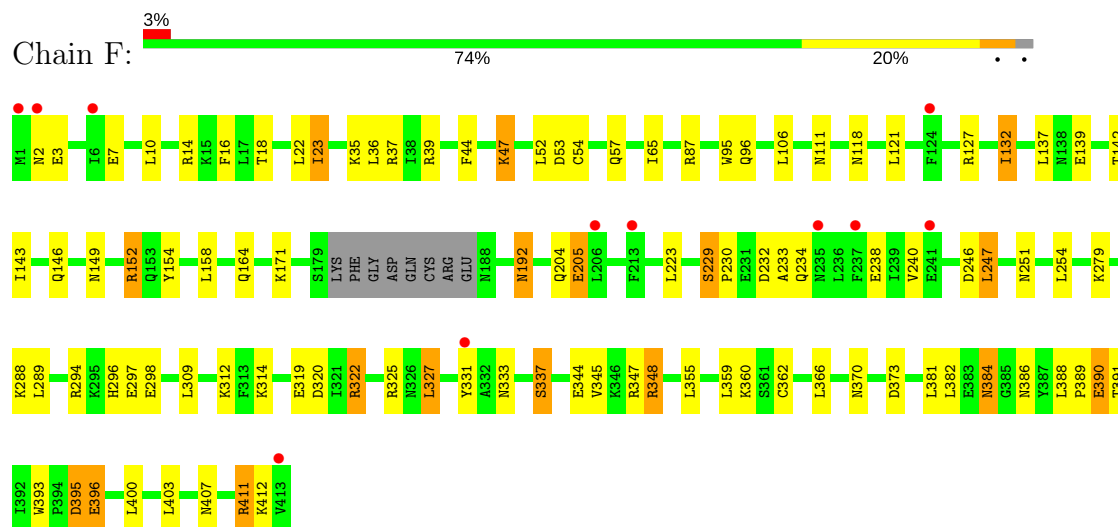
#### • Molecule 2: KLTH0C07942p



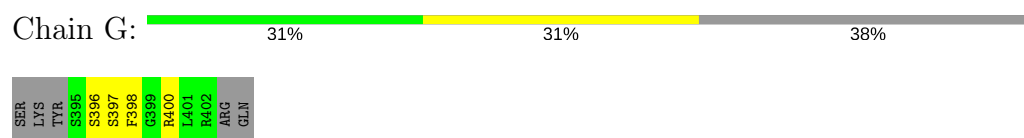
#### • Molecule 3: KLTH0D15642p



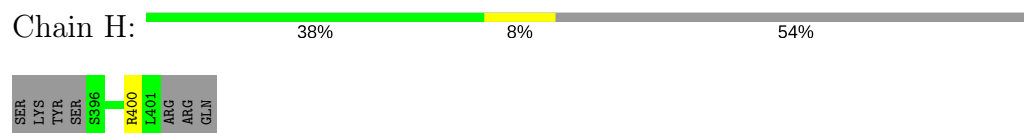
• Molecule 3: KLTH0D15642p



• Molecule 4: Atg13 17BR

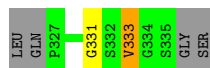


• Molecule 4: Atg13 17BR



• Molecule 5: Atg13 17LR







## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	147.96Å 64.04Å 184.39Å 90.00° 109.91° 90.00°	Depositor
Resolution (Å)	49.22 – 3.21 49.17 – 3.21	Depositor EDS
% Data completeness (in resolution range)	98.3 (49.22-3.21) 98.3 (49.17-3.21)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.62 (at 3.19Å)	Xtriage
Refinement program	REFMAC 5.7.0032	Depositor
R, $R_{free}$	0.224 , 0.265 0.217 , 0.258	Depositor DCC
$R_{free}$ test set	2700 reflections (5.36%)	DCC
Wilson B-factor (Å <sup>2</sup> )	91.2	Xtriage
Anisotropy	0.333	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 97.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	9843	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	142.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.60% 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	0.58	0/599	0.78	0/815
1	D	0.57	0/611	0.80	0/834
2	B	0.77	0/1061	0.97	1/1445 (0.1%)
2	E	0.65	0/1018	0.90	0/1383
3	C	0.74	1/3248 (0.0%)	0.89	1/4384 (0.0%)
3	F	0.63	0/3242	0.84	0/4378
4	G	0.74	0/57	1.02	0/74
4	H	0.55	0/40	0.62	0/52
5	I	0.58	0/55	0.69	0/73
5	J	0.62	0/59	0.99	0/77
All	All	0.68	1/9990 (0.0%)	0.87	2/13515 (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	D	0	1
2	E	0	1
3	C	0	2
3	F	0	1
All	All	0	6

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	C	154	TYR	CG-CD2	-5.13	1.32	1.39

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	C	150	ILE	CB-CA-C	6.19	123.99	111.60
2	B	80	LEU	CA-CB-CG	-5.01	103.78	115.30

There are no chirality outliers.

5 of 6 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	28	TRP	Peptide
3	C	227	SER	Peptide
3	C	393	TRP	Peptide
1	D	28	TRP	Peptide
2	E	58	GLU	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	583	0	504	17	0
1	D	596	0	470	8	0
2	B	1044	0	978	36	0
2	E	1005	0	947	33	0
3	C	3206	0	3110	70	0
3	F	3200	0	3070	49	0
4	G	57	0	52	2	0
4	H	40	0	34	0	0
5	I	54	0	45	0	0
5	J	58	0	56	2	0
All	All	9843	0	9266	191	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 10.

The worst 5 of 191 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:34:GLY:O	2:B:35:HIS:HB2	1.52	1.03

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:F:344:GLU:OE2	3:F:347:ARG:NH2	2.15	0.78
2:E:112:ARG:O	2:E:113:PHE:HB2	1.82	0.78
1:A:54:ARG:O	1:A:58:GLU:HG3	1.86	0.76
2:B:112:ARG:O	2:B:113:PHE:HB2	1.85	0.74

There are no symmetry-related clashes.

## 5.3 Torsion angles [i](#)

### 5.3.1 Protein backbone [i](#)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	72/87 (83%)	65 (90%)	6 (8%)	1 (1%)	13	52
1	D	80/87 (92%)	70 (88%)	9 (11%)	1 (1%)	14	54
2	B	131/151 (87%)	110 (84%)	11 (8%)	10 (8%)	1	8
2	E	122/151 (81%)	108 (88%)	7 (6%)	7 (6%)	2	16
3	C	400/413 (97%)	359 (90%)	34 (8%)	7 (2%)	10	47
3	F	401/413 (97%)	355 (88%)	39 (10%)	7 (2%)	11	48
4	G	6/13 (46%)	2 (33%)	2 (33%)	2 (33%)	0	0
4	H	4/13 (31%)	3 (75%)	0	1 (25%)	0	0
5	I	7/13 (54%)	4 (57%)	3 (43%)	0	100	100
5	J	7/13 (54%)	5 (71%)	2 (29%)	0	100	100
All	All	1230/1354 (91%)	1081 (88%)	113 (9%)	36 (3%)	5	33

5 of 36 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	29	GLU
2	B	33	GLU
2	B	35	HIS

Continued on next page...

*Continued from previous page...*

Mol	Chain	Res	Type
2	B	37	ASP
2	B	40	VAL

### 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	50/79 (63%)	43 (86%)	7 (14%)	4	18
1	D	43/79 (54%)	35 (81%)	8 (19%)	2	9
2	B	112/140 (80%)	96 (86%)	16 (14%)	4	17
2	E	110/140 (79%)	90 (82%)	20 (18%)	2	10
3	C	343/383 (90%)	283 (82%)	60 (18%)	2	11
3	F	338/383 (88%)	287 (85%)	51 (15%)	3	16
4	G	6/12 (50%)	5 (83%)	1 (17%)	2	12
4	H	4/12 (33%)	4 (100%)	0	100	100
5	I	4/9 (44%)	3 (75%)	1 (25%)	1	2
5	J	5/9 (56%)	4 (80%)	1 (20%)	1	7
All	All	1015/1246 (82%)	850 (84%)	165 (16%)	3	13

5 of 165 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	C	354	LYS
2	E	22	VAL
3	F	355	LEU
3	C	357	GLN
1	D	2	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 23 such sidechains are listed below:

Mol	Chain	Res	Type
3	C	326	ASN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
3	C	386	ASN
3	F	384	ASN
3	C	384	ASN
2	E	21	ASN

### 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	76/87 (87%)	0.19	3 (3%)	40	27	113, 164, 221, 234	0
1	D	82/87 (94%)	0.11	3 (3%)	42	28	140, 179, 243, 258	0
2	B	135/151 (89%)	-0.05	0	100	100	81, 116, 190, 243	0
2	E	128/151 (84%)	-0.12	2 (1%)	72	60	98, 136, 203, 222	0
3	C	404/413 (97%)	-0.09	3 (0%)	87	80	77, 119, 193, 249	0
3	F	405/413 (98%)	-0.01	11 (2%)	55	41	91, 145, 227, 258	0
4	G	8/13 (61%)	-0.45	0	100	100	128, 137, 155, 173	0
4	H	6/13 (46%)	0.14	0	100	100	188, 198, 208, 212	0
5	I	9/13 (69%)	0.64	2 (22%)	1	1	150, 178, 204, 205	0
5	J	9/13 (69%)	0.88	0	100	100	134, 161, 180, 182	0
All	All	1262/1354 (93%)	-0.02	24 (1%)	67	53	77, 137, 219, 258	0

The worst 5 of 24 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
5	I	328	PHE	3.6
3	C	1	MET	3.5
3	F	241	GLU	3.4
3	F	237	PHE	3.3
3	F	1	MET	3.2

### 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 [i](#)

There are no carbohydrates in this entry.

### 6.4 Ligands [i](#)

There are no ligands in this entry.

### 6.5 Other polymers [i](#)

There are no such residues in this entry.