



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 31, 2016 – 09:32 PM GMT

PDB ID : 1PD5  
Title : Crystal structure of E.coli chloramphenicol acetyltransferase type I at 2.5 Angstrom resolution  
Authors : Roidis, A.; Kokkinidis, M.  
Deposited on : 2003-05-19  
Resolution : 2.50 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  
Mogul : 1.7 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026688  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
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) : trunk26865

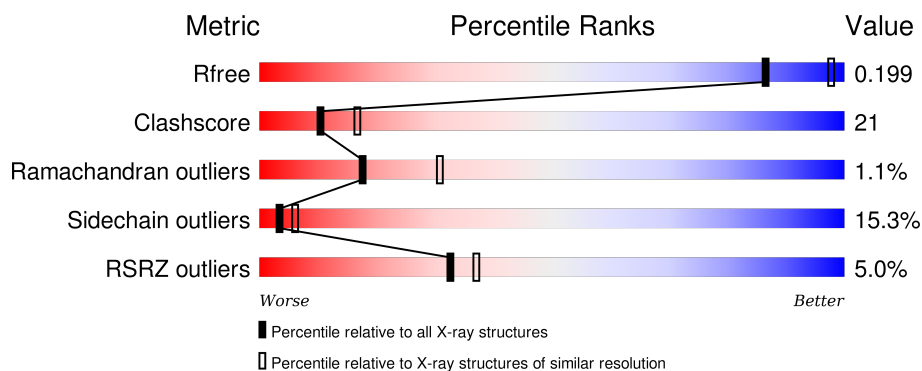
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.50 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))
$R_{free}$	91344	3553 (2.50-2.50)
Clashscore	102246	4242 (2.50-2.50)
Ramachandran outliers	100387	4156 (2.50-2.50)
Sidechain outliers	100360	4158 (2.50-2.50)
RSRZ outliers	91569	3562 (2.50-2.50)

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	219	<div> <div>45%</div> <div>37%</div> <div>12%</div> <div>• •</div> </div>
1	B	219	<div> <div>57%</div> <div>29%</div> <div>9%</div> <div>• •</div> </div>
1	C	219	<div> <div>47%</div> <div>38%</div> <div>9%</div> <div>• •</div> </div>
1	D	219	<div> <div>46%</div> <div>40%</div> <div>8%</div> <div>• •</div> </div>
1	E	219	<div> <div>54%</div> <div>35%</div> <div>8%</div> <div>• •</div> </div>

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	219	<div><div></div><div>2%</div><div>48%</div><div>38%</div><div>10%</div><div></div><div></div></div>
1	G	219	<div><div></div><div>3%</div><div>51%</div><div>35%</div><div>8%</div><div></div><div></div></div>
1	H	219	<div><div></div><div>5%</div><div>49%</div><div>40%</div><div>8%</div><div></div><div></div></div>
1	I	219	<div><div></div><div>2%</div><div>44%</div><div>41%</div><div>10%</div><div></div><div></div></div>
1	J	219	<div><div></div><div>14%</div><div>43%</div><div>36%</div><div>13%</div><div></div><div></div></div>
1	K	219	<div><div></div><div>11%</div><div>47%</div><div>39%</div><div>10%</div><div></div><div></div></div>
1	L	219	<div><div></div><div>15%</div><div>41%</div><div>39%</div><div>16%</div><div></div><div></div></div>

## 2 Entry composition [i](#)

There are 2 unique types of molecules in this entry. The entry contains 21418 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 Chloramphenicol acetyltransferase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	213	Total	C	N	O	S	0	0	0
			1763	1144	290	316	13			
1	B	216	Total	C	N	O	S	0	0	0
			1789	1162	295	319	13			
1	C	213	Total	C	N	O	S	0	0	0
			1763	1144	290	316	13			
1	D	213	Total	C	N	O	S	0	0	0
			1763	1144	290	316	13			
1	E	215	Total	C	N	O	S	0	0	0
			1785	1160	294	318	13			
1	F	212	Total	C	N	O	S	0	0	0
			1759	1142	289	315	13			
1	G	213	Total	C	N	O	S	0	0	0
			1763	1144	290	316	13			
1	H	216	Total	C	N	O	S	0	0	0
			1785	1159	294	319	13			
1	I	214	Total	C	N	O	S	0	0	0
			1776	1154	292	317	13			
1	J	211	Total	C	N	O	S	0	0	0
			1755	1140	288	314	13			
1	K	212	Total	C	N	O	S	0	0	0
			1763	1147	289	314	13			
1	L	210	Total	C	N	O	S	0	0	0
			1746	1135	286	312	13			

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	19	Total	O	0	0
			19	19		
2	B	22	Total	O	0	0
			22	22		

*Continued on next page...*

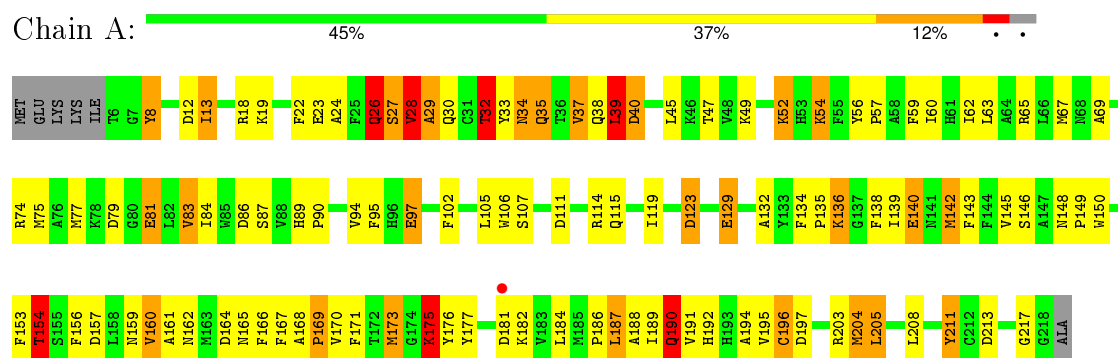
*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	C	19	Total 19	O 19	0	0
2	D	21	Total 21	O 21	0	0
2	E	15	Total 15	O 15	0	0
2	F	16	Total 16	O 16	0	0
2	G	20	Total 20	O 20	0	0
2	H	36	Total 36	O 36	0	0
2	I	17	Total 17	O 17	0	0
2	J	11	Total 11	O 11	0	0
2	K	4	Total 4	O 4	0	0
2	L	8	Total 8	O 8	0	0

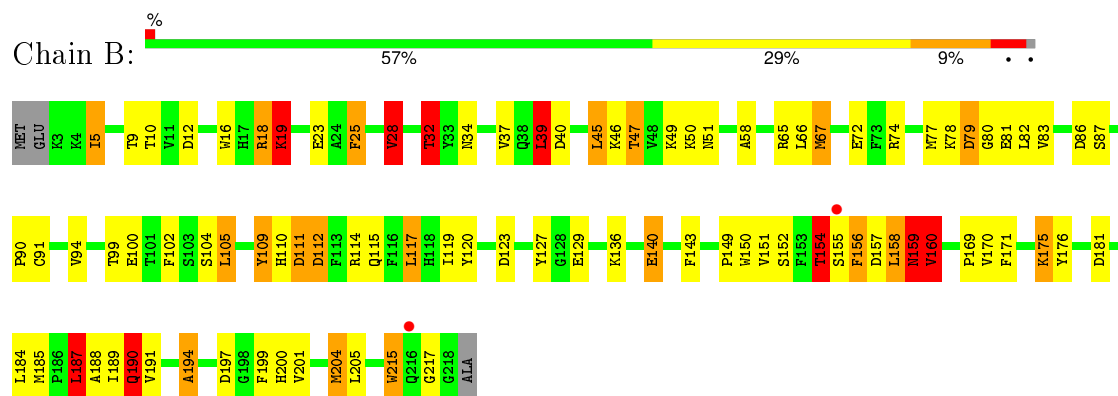
### 3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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 ( $\text{RSRZ} > 2$ ). 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.

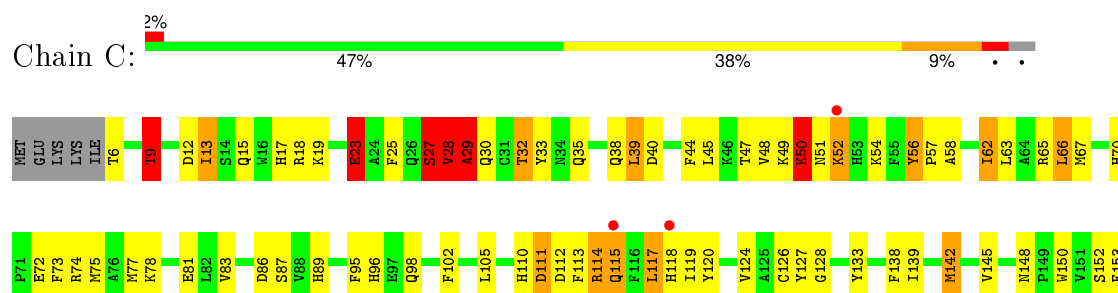
#### • Molecule 1: Chloramphenicol acetyltransferase

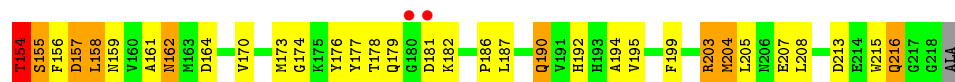


#### • Molecule 1: Chloramphenicol acetyltransferase

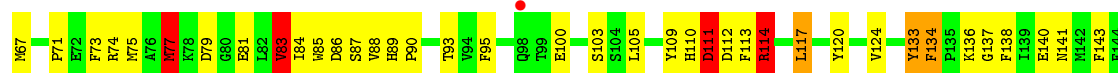
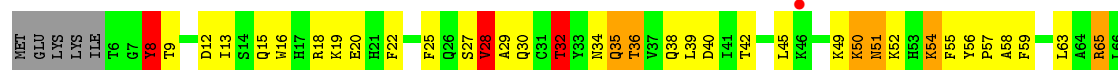


#### • Molecule 1: Chloramphenicol acetyltransferase

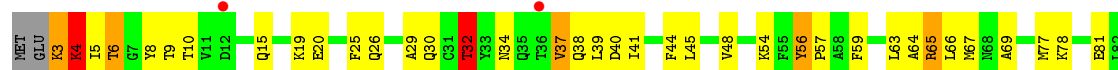




• Molecule 1: Chloramphenicol acetyltransferase



• Molecule 1: Chloramphenicol acetyltransferase

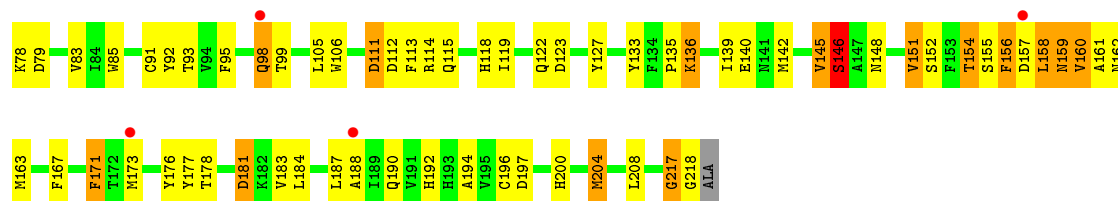


• Molecule 1: Chloramphenicol acetyltransferase

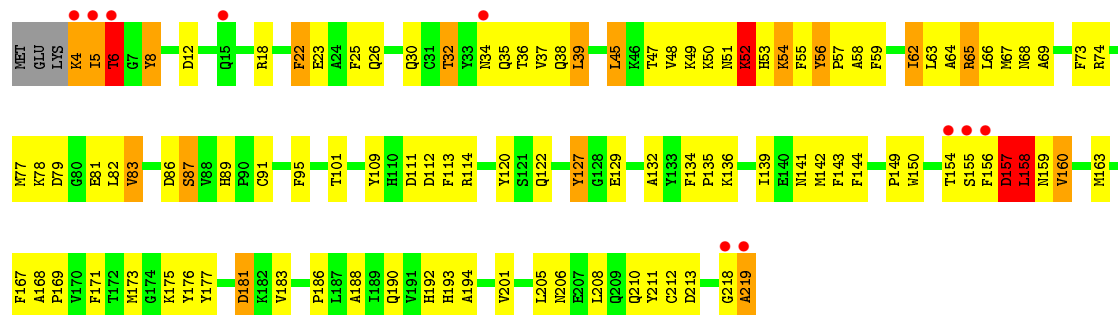


• Molecule 1: Chloramphenicol acetyltransferase

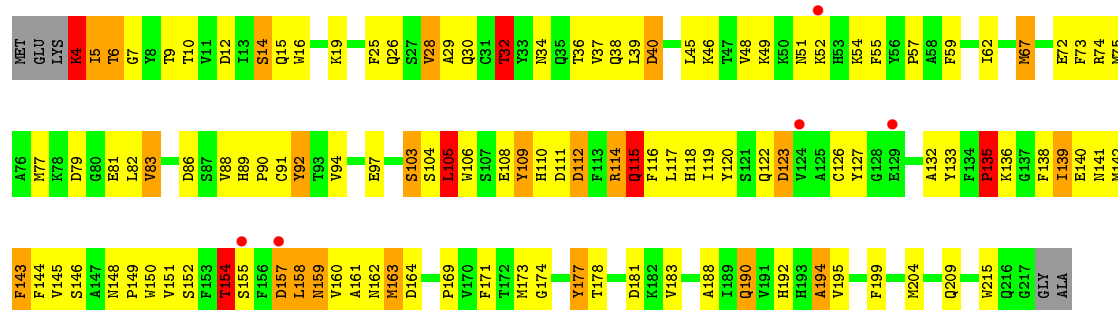




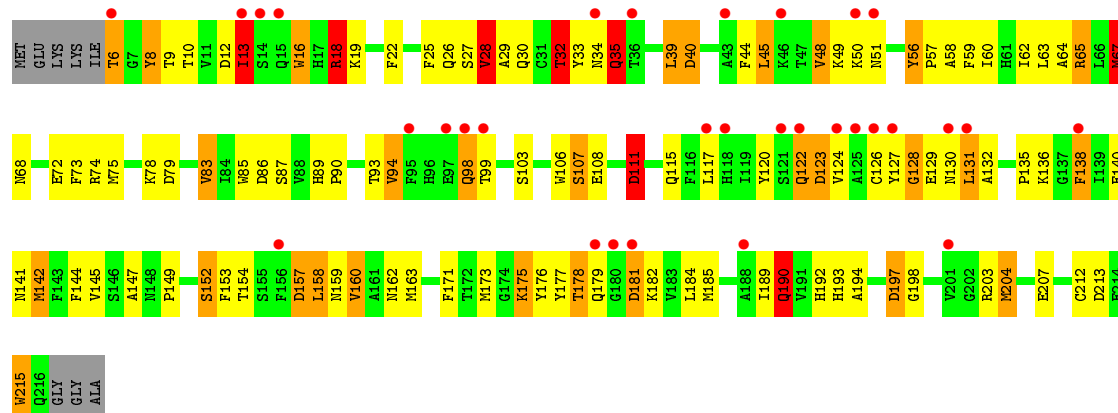
• Molecule 1: Chloramphenicol acetyltransferase



• Molecule 1: Chloramphenicol acetyltransferase



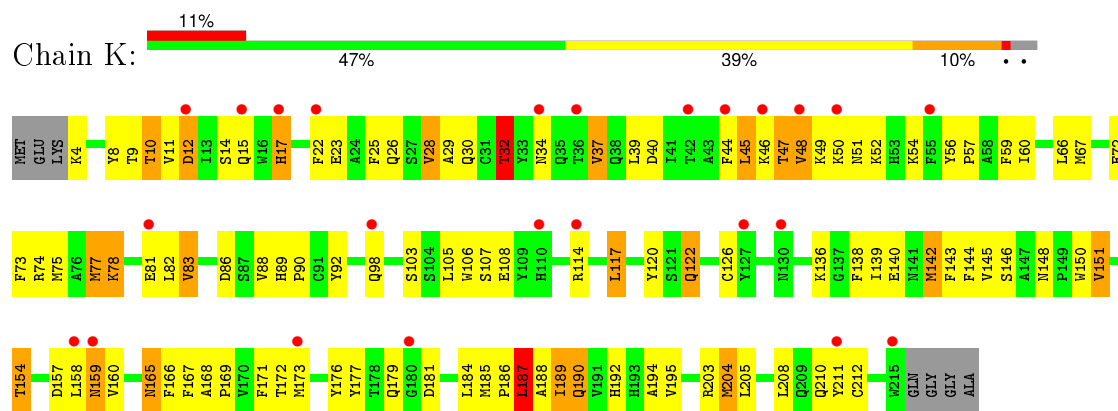
• Molecule 1: Chloramphenicol acetyltransferase





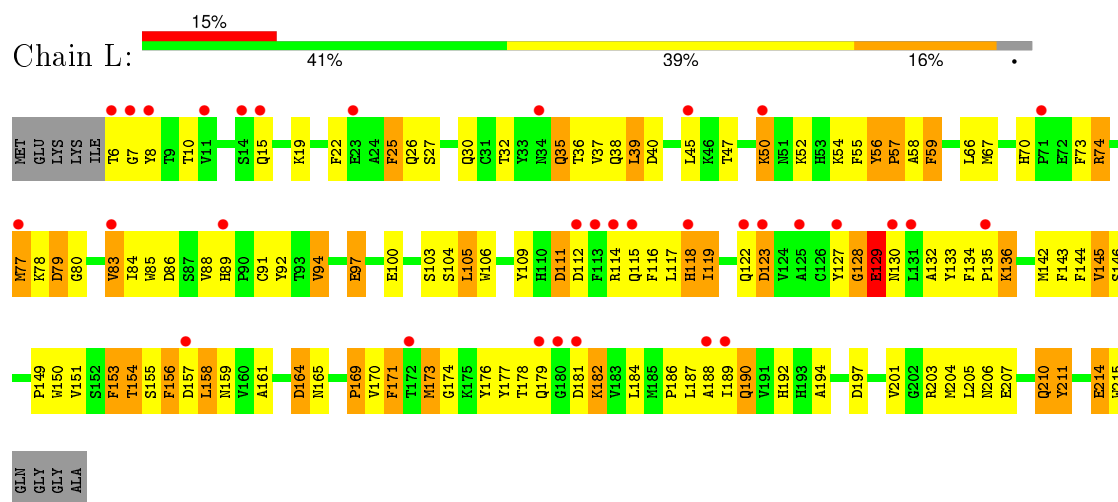
## ● Molecule 1: Chloramphenicol acetyltransferase

Chain K:



## ● Molecule 1: Chloramphenicol acetyltransferase

Chain L:



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	115.74Å 129.70Å 117.98Å 90.00° 108.38° 90.00°	Depositor
Resolution (Å)	10.00 – 2.50 39.33 – 2.50	Depositor EDS
% Data completeness (in resolution range)	98.6 (10.00-2.50) 98.7 (39.33-2.50)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	0.06	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.61 (at 2.51Å)	Xtriage
Refinement program	REFMAC 5.1.24	Depositor
R, $R_{free}$	0.195 , 0.281 0.207 , 0.199	Depositor DCC
$R_{free}$ test set	5598 reflections (5.32%)	DCC
Wilson B-factor (Å <sup>2</sup> )	37.9	Xtriage
Anisotropy	0.643	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 51.5	EDS
Estimated twinning fraction	0.012 for l,-k,h	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.34$	Xtriage
Outliers	0 of 112711 reflections	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	21418	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	46.0	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 21.31 % 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 7.2096e-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, 0.375 respectively for untwinned datasets, and 0.333, 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	2.04	58/1822 (3.2%)	1.62	38/2473 (1.5%)
1	B	1.76	22/1848 (1.2%)	1.59	34/2506 (1.4%)
1	C	1.84	35/1822 (1.9%)	1.58	36/2473 (1.5%)
1	D	1.83	38/1822 (2.1%)	1.52	25/2473 (1.0%)
1	E	1.67	20/1844 (1.1%)	1.39	14/2501 (0.6%)
1	F	1.76	26/1818 (1.4%)	1.53	31/2468 (1.3%)
1	G	1.71	18/1822 (1.0%)	1.53	30/2473 (1.2%)
1	H	1.85	33/1844 (1.8%)	1.55	30/2502 (1.2%)
1	I	1.85	38/1835 (2.1%)	1.50	23/2490 (0.9%)
1	J	1.57	15/1814 (0.8%)	1.47	26/2463 (1.1%)
1	K	1.55	14/1822 (0.8%)	1.35	9/2473 (0.4%)
1	L	1.67	21/1805 (1.2%)	1.41	19/2451 (0.8%)
All	All	1.76	338/21918 (1.5%)	1.51	315/29746 (1.1%)

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	C	0	2
1	G	0	1
1	I	0	1
1	J	0	3
1	K	0	1
1	L	0	1
All	All	0	10

The worst 5 of 338 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	150	TRP	CE3-CZ3	10.79	1.56	1.38

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	95	PHE	CE2-CZ	10.43	1.57	1.37
1	A	87	SER	CB-OG	-10.09	1.29	1.42
1	I	127	TYR	CD1-CE1	10.07	1.54	1.39
1	I	140	GLU	CD-OE1	9.99	1.36	1.25

The worst 5 of 315 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	F	111	ASP	CB-CG-OD2	13.02	130.02	118.30
1	A	65	ARG	NE-CZ-NH1	11.59	126.09	120.30
1	C	157	ASP	CB-CG-OD2	11.45	128.60	118.30
1	C	18	ARG	NE-CZ-NH1	11.39	126.00	120.30
1	I	181	ASP	CB-CG-OD2	11.01	128.21	118.30

There are no chirality outliers.

5 of 10 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	168	ALA	Mainchain
1	C	28	VAL	Peptide
1	C	52	LYS	Peptide
1	G	217	GLY	Peptide
1	I	4	LYS	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	1763	0	1647	72	0
1	B	1789	0	1684	58	0
1	C	1763	0	1647	83	0
1	D	1763	0	1647	79	0
1	E	1785	0	1681	61	0
1	F	1759	0	1644	56	0
1	G	1763	0	1647	83	0
1	H	1785	0	1675	72	0
1	I	1776	0	1668	104	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	J	1755	0	1641	115	0
1	K	1763	0	1657	92	0
1	L	1746	0	1633	94	0
2	A	19	0	0	1	0
2	B	22	0	0	1	0
2	C	19	0	0	1	0
2	D	21	0	0	2	0
2	E	15	0	0	1	0
2	F	16	0	0	1	0
2	G	20	0	0	3	0
2	H	36	0	0	3	0
2	I	17	0	0	0	0
2	J	11	0	0	5	0
2	K	4	0	0	1	0
2	L	8	0	0	0	0
All	All	21418	0	19871	881	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 21.

The worst 5 of 881 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:C:13:ILE:CG1	1:C:13:ILE:CD1	1.83	1.56
1:I:6:THR:CG2	1:I:6:THR:CB	1.74	1.55
1:I:139:ILE:N	1:I:142:MET:HE3	1.19	1.44
1:I:139:ILE:H	1:I:142:MET:CE	1.36	1.37
1:I:171:PHE:CZ	1:I:204:MET:HE3	1.61	1.35

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	211/219 (96%)	203 (96%)	7 (3%)	1 (0%)	34	55
1	B	214/219 (98%)	204 (95%)	8 (4%)	2 (1%)	21	37
1	C	211/219 (96%)	200 (95%)	8 (4%)	3 (1%)	14	24
1	D	211/219 (96%)	196 (93%)	12 (6%)	3 (1%)	14	24
1	E	213/219 (97%)	199 (93%)	12 (6%)	2 (1%)	21	37
1	F	210/219 (96%)	195 (93%)	14 (7%)	1 (0%)	34	55
1	G	211/219 (96%)	196 (93%)	14 (7%)	1 (0%)	34	55
1	H	214/219 (98%)	204 (95%)	8 (4%)	2 (1%)	21	37
1	I	212/219 (97%)	199 (94%)	11 (5%)	2 (1%)	21	37
1	J	209/219 (95%)	191 (91%)	16 (8%)	2 (1%)	19	34
1	K	210/219 (96%)	195 (93%)	10 (5%)	5 (2%)	7	11
1	L	208/219 (95%)	188 (90%)	16 (8%)	4 (2%)	10	16
All	All	2534/2628 (96%)	2370 (94%)	136 (5%)	28 (1%)	17	31

5 of 28 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	28	VAL
1	B	79	ASP
1	D	164	ASP
1	K	48	VAL
1	L	119	ILE

### 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	189/194 (97%)	160 (85%)	29 (15%)	3	6
1	B	192/194 (99%)	163 (85%)	29 (15%)	3	6
1	C	189/194 (97%)	163 (86%)	26 (14%)	4	8
1	D	189/194 (97%)	162 (86%)	27 (14%)	4	7

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	E	192/194 (99%)	161 (84%)	31 (16%)	3	5
1	F	189/194 (97%)	161 (85%)	28 (15%)	4	7
1	G	189/194 (97%)	164 (87%)	25 (13%)	5	9
1	H	191/194 (98%)	171 (90%)	20 (10%)	8	16
1	I	191/194 (98%)	163 (85%)	28 (15%)	4	7
1	J	189/194 (97%)	158 (84%)	31 (16%)	3	5
1	K	190/194 (98%)	150 (79%)	40 (21%)	1	2
1	L	188/194 (97%)	154 (82%)	34 (18%)	2	3
All	All	2278/2328 (98%)	1930 (85%)	348 (15%)	3	6

5 of 348 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	F	115	GLN
1	G	204	MET
1	L	52	LYS
1	F	148	ASN
1	G	45	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 75 such sidechains are listed below:

Mol	Chain	Res	Type
1	F	110	HIS
1	G	53	HIS
1	K	122	GLN
1	F	148	ASN
1	G	26	GLN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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

## 5.5 Carbohydrates [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	213/219 (97%)	-0.02	1 (0%) 91 92	21, 34, 50, 57	0
1	B	216/219 (98%)	-0.05	2 (0%) 85 88	23, 39, 59, 71	0
1	C	213/219 (97%)	0.09	5 (2%) 64 67	24, 38, 62, 76	0
1	D	213/219 (97%)	0.12	3 (1%) 78 80	27, 42, 60, 80	0
1	E	215/219 (98%)	0.13	4 (1%) 70 73	30, 45, 64, 84	0
1	F	212/219 (96%)	0.17	5 (2%) 62 66	29, 45, 62, 74	0
1	G	213/219 (97%)	0.19	6 (2%) 56 61	31, 41, 59, 69	0
1	H	216/219 (98%)	0.14	10 (4%) 36 41	19, 37, 52, 72	0
1	I	214/219 (97%)	0.25	5 (2%) 64 67	32, 44, 62, 72	0
1	J	211/219 (96%)	0.72	31 (14%) 3 3	38, 55, 77, 81	0
1	K	212/219 (96%)	0.70	24 (11%) 7 7	43, 59, 76, 88	0
1	L	210/219 (95%)	0.82	33 (15%) 3 2	40, 55, 74, 84	0
All	All	2558/2628 (97%)	0.27	129 (5%) 32 37	19, 44, 69, 88	0

The worst 5 of 129 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	J	124	VAL	6.3
1	J	125	ALA	5.8
1	H	5	ILE	5.0
1	J	6	THR	4.8
1	L	6	THR	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 [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.