



# wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 30, 2017 – 01:35 PM EDT

PDB ID : 3LVG  
Title : Crystal structure of a clathrin heavy chain and clathrin light chain complex  
Authors : Wilbur, J.D.; Hwang, P.K.; Ybe, J.A.; Lane, M.; Sellers, B.D.; Jacobson, M.P.; Fletterick, R.J.; Brodsky, F.M.  
Deposited on : unknown  
Resolution : 7.94 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	:	rb-20030345
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)	:	rb-20030345

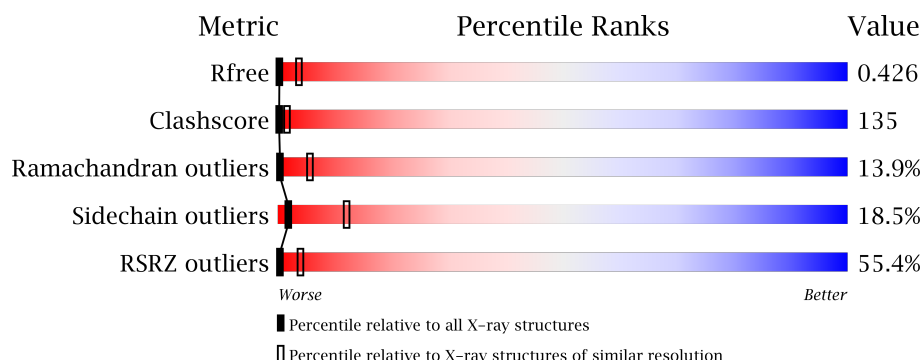
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 7.94 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	1100 (10.00-3.70)
Clashscore	112137	1036 (11.50-3.80)
Ramachandran outliers	110173	1004 (11.50-3.76)
Sidechain outliers	110143	1099 (11.50-3.70)
RSRZ outliers	101464	1004 (11.50-3.72)

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	624	<div> <div>54%</div> <div> <div>14%</div> <div>49%</div> <div>21%</div> <div>6%</div> <div>11%</div> </div> </div>
1	B	624	<div> <div>57%</div> <div> <div>12%</div> <div>52%</div> <div>20%</div> <div>5%</div> <div>11%</div> </div> </div>
1	C	624	<div> <div>37%</div> <div> <div>13%</div> <div>50%</div> <div>20%</div> <div>6%</div> <div>11%</div> </div> </div>
2	D	190	<div> <div>32%</div> <div> <div>35%</div> <div>46%</div> <div>12%</div> <div>• 5%</div> </div> </div>
2	E	190	<div> <div>13%</div> <div> <div>16%</div> <div>32%</div> <div>9%</div> <div>•</div> <div>39%</div> </div> </div>

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	F	190	<p>17% 18% 36% 13% 31%</p>

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 16504 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 Clathrin heavy chain 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	553	Total	C	N	O	S	0	0	0
			4543	2896	767	855	25			
1	B	553	Total	C	N	O	S	0	0	0
			4543	2896	767	855	25			
1	C	553	Total	C	N	O	S	0	0	0
			4543	2896	767	855	25			

There are 66 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1052	MET	-	EXPRESSION TAG	UNP P49951
A	1053	GLY	-	EXPRESSION TAG	UNP P49951
A	1054	SER	-	EXPRESSION TAG	UNP P49951
A	1055	SER	-	EXPRESSION TAG	UNP P49951
A	1056	HIS	-	EXPRESSION TAG	UNP P49951
A	1057	HIS	-	EXPRESSION TAG	UNP P49951
A	1058	HIS	-	EXPRESSION TAG	UNP P49951
A	1059	HIS	-	EXPRESSION TAG	UNP P49951
A	1060	HIS	-	EXPRESSION TAG	UNP P49951
A	1061	HIS	-	EXPRESSION TAG	UNP P49951
A	1062	SER	-	EXPRESSION TAG	UNP P49951
A	1063	SER	-	EXPRESSION TAG	UNP P49951
A	1064	GLY	-	EXPRESSION TAG	UNP P49951
A	1065	LEU	-	EXPRESSION TAG	UNP P49951
A	1066	VAL	-	EXPRESSION TAG	UNP P49951
A	1067	PRO	-	EXPRESSION TAG	UNP P49951
A	1068	ARG	-	EXPRESSION TAG	UNP P49951
A	1069	GLY	-	EXPRESSION TAG	UNP P49951
A	1070	SER	-	EXPRESSION TAG	UNP P49951
A	1071	HIS	-	EXPRESSION TAG	UNP P49951
A	1072	MET	-	EXPRESSION TAG	UNP P49951
A	1073	LEU	-	EXPRESSION TAG	UNP P49951
B	1052	MET	-	EXPRESSION TAG	UNP P49951

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	1053	GLY	-	EXPRESSION TAG	UNP P49951
B	1054	SER	-	EXPRESSION TAG	UNP P49951
B	1055	SER	-	EXPRESSION TAG	UNP P49951
B	1056	HIS	-	EXPRESSION TAG	UNP P49951
B	1057	HIS	-	EXPRESSION TAG	UNP P49951
B	1058	HIS	-	EXPRESSION TAG	UNP P49951
B	1059	HIS	-	EXPRESSION TAG	UNP P49951
B	1060	HIS	-	EXPRESSION TAG	UNP P49951
B	1061	HIS	-	EXPRESSION TAG	UNP P49951
B	1062	SER	-	EXPRESSION TAG	UNP P49951
B	1063	SER	-	EXPRESSION TAG	UNP P49951
B	1064	GLY	-	EXPRESSION TAG	UNP P49951
B	1065	LEU	-	EXPRESSION TAG	UNP P49951
B	1066	VAL	-	EXPRESSION TAG	UNP P49951
B	1067	PRO	-	EXPRESSION TAG	UNP P49951
B	1068	ARG	-	EXPRESSION TAG	UNP P49951
B	1069	GLY	-	EXPRESSION TAG	UNP P49951
B	1070	SER	-	EXPRESSION TAG	UNP P49951
B	1071	HIS	-	EXPRESSION TAG	UNP P49951
B	1072	MET	-	EXPRESSION TAG	UNP P49951
B	1073	LEU	-	EXPRESSION TAG	UNP P49951
C	1052	MET	-	EXPRESSION TAG	UNP P49951
C	1053	GLY	-	EXPRESSION TAG	UNP P49951
C	1054	SER	-	EXPRESSION TAG	UNP P49951
C	1055	SER	-	EXPRESSION TAG	UNP P49951
C	1056	HIS	-	EXPRESSION TAG	UNP P49951
C	1057	HIS	-	EXPRESSION TAG	UNP P49951
C	1058	HIS	-	EXPRESSION TAG	UNP P49951
C	1059	HIS	-	EXPRESSION TAG	UNP P49951
C	1060	HIS	-	EXPRESSION TAG	UNP P49951
C	1061	HIS	-	EXPRESSION TAG	UNP P49951
C	1062	SER	-	EXPRESSION TAG	UNP P49951
C	1063	SER	-	EXPRESSION TAG	UNP P49951
C	1064	GLY	-	EXPRESSION TAG	UNP P49951
C	1065	LEU	-	EXPRESSION TAG	UNP P49951
C	1066	VAL	-	EXPRESSION TAG	UNP P49951
C	1067	PRO	-	EXPRESSION TAG	UNP P49951
C	1068	ARG	-	EXPRESSION TAG	UNP P49951
C	1069	GLY	-	EXPRESSION TAG	UNP P49951
C	1070	SER	-	EXPRESSION TAG	UNP P49951
C	1071	HIS	-	EXPRESSION TAG	UNP P49951
C	1072	MET	-	EXPRESSION TAG	UNP P49951

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
C	1073	LEU	-	EXPRESSION TAG	UNP P49951

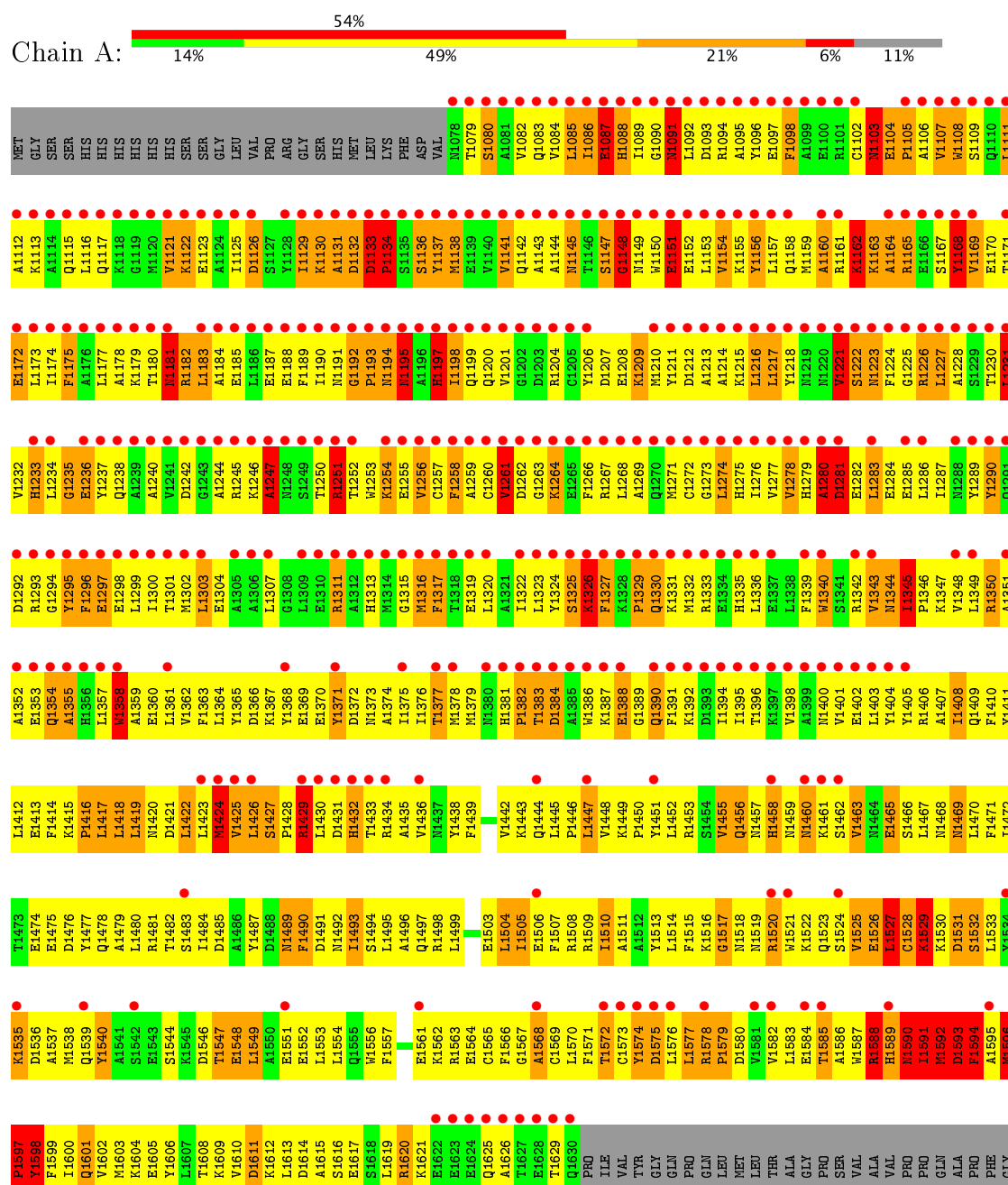
- Molecule 2 is a protein called Clathrin light chain B.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	D	180	Total	C	N	O	S	0	0	0
			1146	690	227	228	1			
2	E	116	Total	C	N	O	S	0	0	0
			823	497	163	162	1			
2	F	132	Total	C	N	O	S	0	0	0
			906	546	179	180	1			

### 3 Residue-property plots

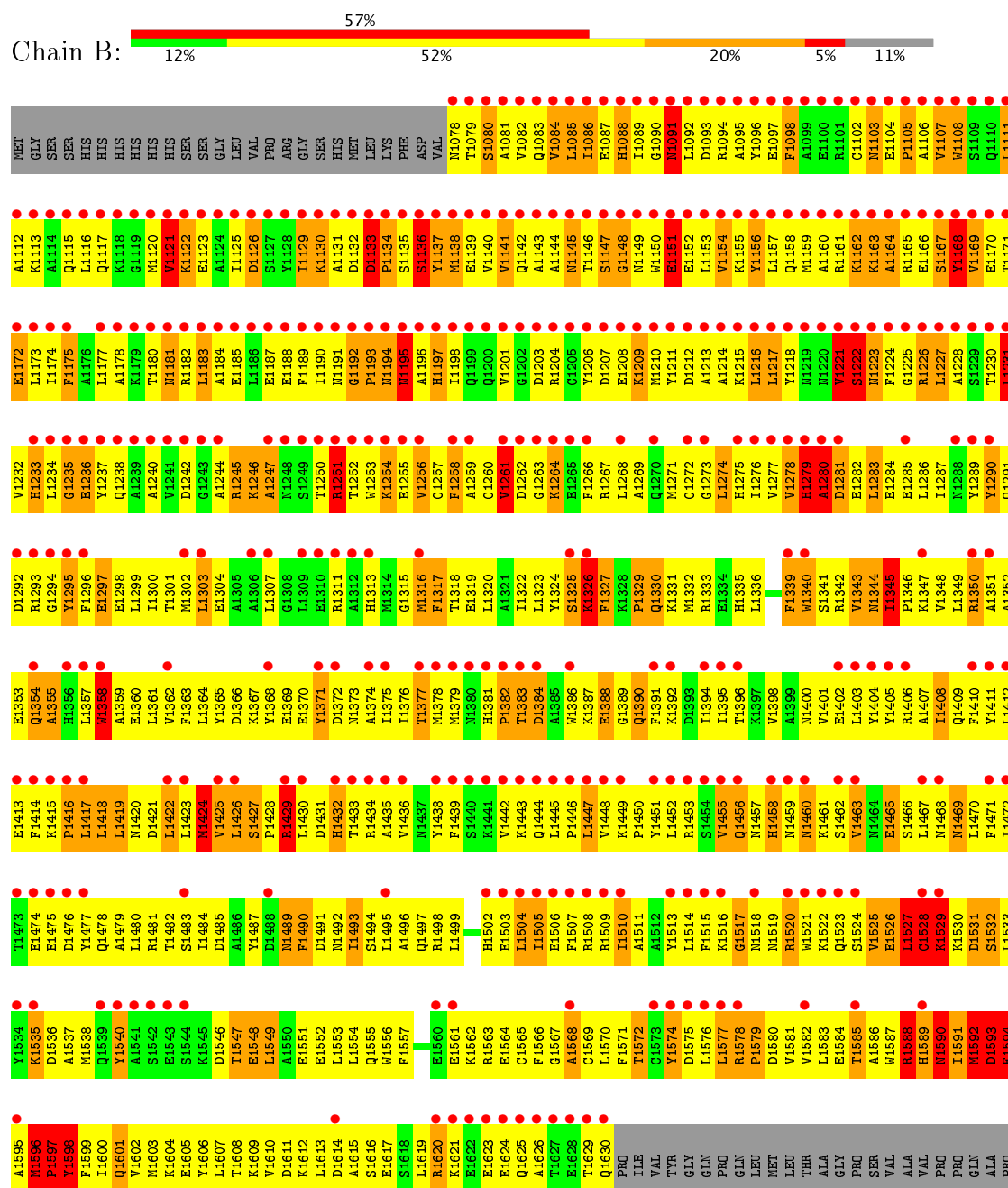
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are 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: Clathrin heavy chain 1

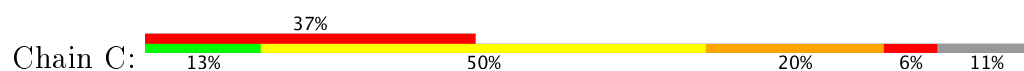


TYR  
GLY  
THR  
ALA  
PRO  
ALA  
TYR  
GLY  
GLN  
PRO  
GLN  
GLY  
PHE  
GLY  
TYR  
SER  
MET

• Molecule 1: Clathrin heavy chain 1



• Molecule 1: Clathrin heavy chain 1

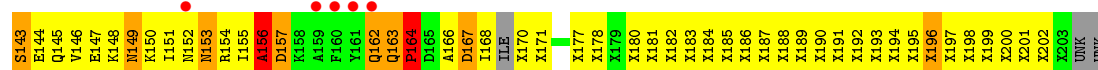




- Molecule 2: Clathrin light chain B



T96	X1
Q97	X5
E98	X10
P99	UNK
S100	UNK
S101	X12
I102	X13
K103	X14
K104	X15
M105	X16
R106	X19
E107	X23
E108	UNK
Q109	X25
R110	X26
K111	X27
L112	X28
L113	X29
Q114	X36
E115	UNK
E116	X38
D117	X45
A118	UNK
A119	X47
S120	X48
K121	X49
M122	X50
M123	X51
E124	X52
Q125	X53
E126	X54
M127	X55
R128	X58
R129	UNK
K130	X60
K131	X61
K132	X62
K133	X63
D134	X64
L135	X65
E136	X66
E137	X69
M138	X70
M139	X71
Q140	X72
R141	UNK
Q142	X76
S143	X77
E144	X78
Q145	X79
V146	X81
E147	X82
K148	X83
M149	X84
K150	X85
I151	X86
M152	X87
M153	X88
R154	X89
R155	X90



## 4 Data and refinement statistics

Property	Value	Source
Space group	I 41 2 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	228.56 Å   228.56 Å   710.32 Å 90.00°   90.00°   90.00°	Depositor
Resolution (Å)	100.00 – 7.94 82.22 – 4.99	Depositor EDS
% Data completeness (in resolution range)	99.6 (100.00-7.94) 63.0 (82.22-4.99)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.07	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	0.97 (at 5.12 Å)	Xtriage
Refinement program	CNS	Depositor
R, $R_{free}$	0.419   ,   0.425 0.423   ,   0.426	Depositor DCC
$R_{free}$ test set	No test flags present.	DCC
Wilson B-factor (Å <sup>2</sup> )	310.7	Xtriage
Anisotropy	0.082	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.41 , 471.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.30$ , $\langle L^2 \rangle = 0.15$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.83	EDS
Total number of atoms	16504	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	295.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.39% 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	0.83	20/4638 (0.4%)	1.27	83/6266 (1.3%)
1	B	0.80	17/4638 (0.4%)	1.26	83/6266 (1.3%)
1	C	0.77	13/4638 (0.3%)	1.20	73/6266 (1.2%)
2	D	0.75	0/647	1.15	8/866 (0.9%)
2	E	0.69	0/589	1.24	5/785 (0.6%)
2	F	0.77	0/642	1.20	8/859 (0.9%)
All	All	0.79	50/15792 (0.3%)	1.24	260/21308 (1.2%)

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	1	8
1	B	0	3
1	C	2	3
2	F	0	1
All	All	3	15

The worst 5 of 50 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	1162	LYS	C-O	-19.68	0.85	1.23
1	A	1222	SER	C-O	18.12	1.57	1.23
1	C	1182	ARG	C-O	17.19	1.56	1.23
1	C	1136	SER	C-O	-17.10	0.90	1.23
1	C	1248	ASN	N-CA	17.04	1.80	1.46

The worst 5 of 260 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	1162	LYS	CA-C-O	22.12	166.55	120.10

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	1279	HIS	O-C-N	-22.04	87.43	122.70
1	C	1223	ASN	N-CA-C	18.52	161.00	111.00
1	A	1103	ASN	C-N-CA	-17.51	77.93	121.70
1	A	1162	LYS	O-C-N	-16.77	95.87	122.70

All (3) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	A	1104	GLU	CA
1	C	1137	TYR	CA
1	C	1223	ASN	CA

5 of 15 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	1103	ASN	Mainchain
1	A	1104	GLU	Mainchain
1	A	1133	ASP	Mainchain
1	A	1147	SER	Mainchain
1	A	1162	LYS	Mainchain

## 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	4543	0	4454	1296	1
1	B	4543	0	4456	1248	3
1	C	4543	0	4455	1291	4
2	D	1146	0	735	252	0
2	E	823	0	633	246	0
2	F	906	0	669	239	0
All	All	16504	0	15402	4313	5

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 135.

The worst 5 of 4313 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:A:1258:PHE:HB2	1:A:1289:TYR:CD2	1.19	1.69
1:A:1253:TRP:CZ3	1:A:1276:ILE:HG22	1.25	1.64
1:C:1253:TRP:CZ3	1:C:1276:ILE:HG22	1.25	1.64
1:B:1253:TRP:CZ3	1:B:1276:ILE:HG22	1.25	1.63
1:A:1258:PHE:CB	1:A:1289:TYR:CE2	1.75	1.63

All (5) 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:B:1340:TRP:CZ2	1:C:1222:SER:OG[12_655]	1.80	0.40
1:C:1304:GLU:OE2	1:C:1334:GLU:OE2[15_645]	1.98	0.22
1:A:1199:GLN:NE2	1:A:1431:ASP:OD2[10_555]	2.13	0.07
1:B:1340:TRP:CZ2	1:C:1222:SER:CB[12_655]	2.15	0.05
1:B:1341:SER:OG	1:C:1203:ASP:OD2[12_655]	2.18	0.02

## 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	551/624 (88%)	322 (58%)	152 (28%)	77 (14%)	0	5
1	B	551/624 (88%)	319 (58%)	161 (29%)	71 (13%)	0	6
1	C	551/624 (88%)	309 (56%)	164 (30%)	78 (14%)	0	5
2	D	77/190 (40%)	50 (65%)	16 (21%)	11 (14%)	0	5
2	E	66/190 (35%)	35 (53%)	21 (32%)	10 (15%)	0	5
2	F	76/190 (40%)	42 (55%)	20 (26%)	14 (18%)	0	3
All	All	1872/2442 (77%)	1077 (58%)	534 (28%)	261 (14%)	0	5

5 of 261 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	1087	GLU
1	A	1091	ASN
1	A	1105	PRO
1	A	1122	LYS
1	A	1130	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	485/541 (90%)	409 (84%)	76 (16%)	3	18
1	B	485/541 (90%)	404 (83%)	81 (17%)	2	16
1	C	485/541 (90%)	395 (81%)	90 (19%)	2	11
2	D	62/73 (85%)	45 (73%)	17 (27%)	0	4
2	E	61/73 (84%)	42 (69%)	19 (31%)	0	2
2	F	62/73 (85%)	42 (68%)	20 (32%)	0	2
All	All	1640/1842 (89%)	1337 (82%)	303 (18%)	2	12

5 of 303 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	1527	LEU
1	C	1108	TRP
2	E	140	GLN
1	B	1549	LEU
1	B	1601	GLN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 63 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	1390	GLN
1	B	1523	GLN
2	D	153	ASN
1	B	1456	GLN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	B	1468	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 [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(Å <sup>2</sup> )	Q<0.9
1	A	553/624 (88%)	3.88	336 (60%) 0 3	306, 348, 348, 348	0
1	B	553/624 (88%)	4.65	357 (64%) 0 3	257, 257, 318, 319	0
1	C	553/624 (88%)	2.64	232 (41%) 0 5	232, 232, 308, 309	0
2	D	79/190 (41%)	4.01	61 (77%) 0 3	298, 298, 298, 298	0
2	E	68/190 (35%)	1.85	25 (36%) 0 5	314, 314, 314, 314	0
2	F	78/190 (41%)	3.00	33 (42%) 0 5	339, 339, 339, 339	0
All	All	1884/2442 (77%)	3.64	1044 (55%) 0 4	232, 307, 348, 348	0

The worst 5 of 1044 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	1091	ASN	37.4
1	C	1078	ASN	35.2
1	C	1090	GLY	33.1
1	C	1079	THR	31.5
1	B	1105	PRO	27.7

### 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](#)

There are no ligands in this entry.

## 6.5 Other polymers [i](#)

There are no such residues in this entry.