



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 14, 2020 – 07:25 pm BST

PDB ID : 5IG9  
Title : Crystal structure of macrocyclase MdnC bound with precursor peptide MdnA from *Microcystis aeruginosa* MRC  
Authors : Li, K.; Condurso, H.L.; Bruner, S.D.  
Deposited on : 2016-02-27  
Resolution : 2.67 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

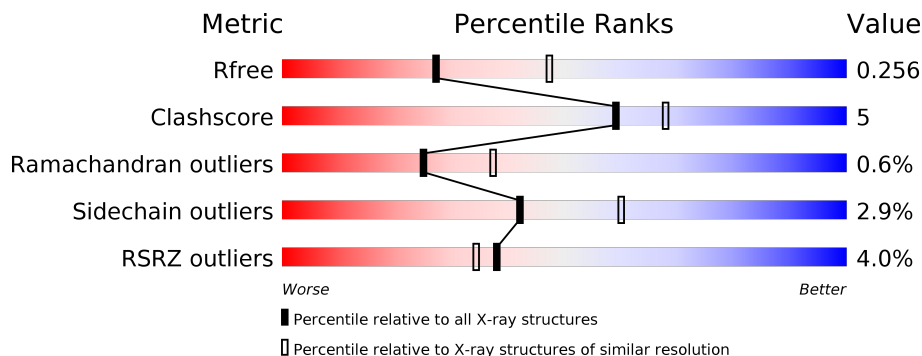
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.67 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	1332 (2.68-2.64)
Clashscore	141614	1374 (2.68-2.64)
Ramachandran outliers	138981	1349 (2.68-2.64)
Sidechain outliers	138945	1349 (2.68-2.64)
RSRZ outliers	127900	1318 (2.68-2.64)

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	333	
1	B	333	
1	C	333	
1	D	333	
1	E	333	
1	F	333	

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	G	333	<p>3% 85% 8% 6%</p>
1	H	333	<p>4% 77% 14% 7%</p>
2	I	49	<p>10% 24% 76%</p>
2	J	49	<p>6% 24% 76%</p>
2	K	49	<p>2% 18% 6% 76%</p>
2	L	49	<p>8% 22% 76%</p>
2	M	49	<p>6% 18% 78%</p>
2	N	49	<p>2% 27% 73%</p>
2	O	49	<p>6% 22% 76%</p>
2	P	49	<p>2% 18% 76%</p>

## 2 Entry composition i

There are 2 unique types of molecules in this entry. The entry contains 20643 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 ATP grasp ligase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	319	Total 2559	C 1651	N 421	O 476	S 11	0	0	0
1	B	305	Total 2446	C 1582	N 398	O 455	S 11	0	1	0
1	C	312	Total 2503	C 1619	N 408	O 464	S 12	0	0	0
1	D	306	Total 2447	C 1580	N 399	O 456	S 12	0	0	0
1	E	312	Total 2502	C 1616	N 411	O 463	S 12	0	0	0
1	F	302	Total 2412	C 1560	N 392	O 448	S 12	0	0	0
1	G	313	Total 2514	C 1625	N 412	O 465	S 12	0	0	0
1	H	310	Total 2489	C 1608	N 406	O 463	S 12	0	1	0

There are 72 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	325	ALA	-	expression tag	UNP B2G3D0
A	326	ALA	-	expression tag	UNP B2G3D0
A	327	ALA	-	expression tag	UNP B2G3D0
A	328	HIS	-	expression tag	UNP B2G3D0
A	329	HIS	-	expression tag	UNP B2G3D0
A	330	HIS	-	expression tag	UNP B2G3D0
A	331	HIS	-	expression tag	UNP B2G3D0
A	332	HIS	-	expression tag	UNP B2G3D0
A	333	HIS	-	expression tag	UNP B2G3D0
B	325	ALA	-	expression tag	UNP B2G3D0
B	326	ALA	-	expression tag	UNP B2G3D0
B	327	ALA	-	expression tag	UNP B2G3D0
B	328	HIS	-	expression tag	UNP B2G3D0

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	329	HIS	-	expression tag	UNP B2G3D0
B	330	HIS	-	expression tag	UNP B2G3D0
B	331	HIS	-	expression tag	UNP B2G3D0
B	332	HIS	-	expression tag	UNP B2G3D0
B	333	HIS	-	expression tag	UNP B2G3D0
C	325	ALA	-	expression tag	UNP B2G3D0
C	326	ALA	-	expression tag	UNP B2G3D0
C	327	ALA	-	expression tag	UNP B2G3D0
C	328	HIS	-	expression tag	UNP B2G3D0
C	329	HIS	-	expression tag	UNP B2G3D0
C	330	HIS	-	expression tag	UNP B2G3D0
C	331	HIS	-	expression tag	UNP B2G3D0
C	332	HIS	-	expression tag	UNP B2G3D0
C	333	HIS	-	expression tag	UNP B2G3D0
D	325	ALA	-	expression tag	UNP B2G3D0
D	326	ALA	-	expression tag	UNP B2G3D0
D	327	ALA	-	expression tag	UNP B2G3D0
D	328	HIS	-	expression tag	UNP B2G3D0
D	329	HIS	-	expression tag	UNP B2G3D0
D	330	HIS	-	expression tag	UNP B2G3D0
D	331	HIS	-	expression tag	UNP B2G3D0
D	332	HIS	-	expression tag	UNP B2G3D0
D	333	HIS	-	expression tag	UNP B2G3D0
E	325	ALA	-	expression tag	UNP B2G3D0
E	326	ALA	-	expression tag	UNP B2G3D0
E	327	ALA	-	expression tag	UNP B2G3D0
E	328	HIS	-	expression tag	UNP B2G3D0
E	329	HIS	-	expression tag	UNP B2G3D0
E	330	HIS	-	expression tag	UNP B2G3D0
E	331	HIS	-	expression tag	UNP B2G3D0
E	332	HIS	-	expression tag	UNP B2G3D0
E	333	HIS	-	expression tag	UNP B2G3D0
F	325	ALA	-	expression tag	UNP B2G3D0
F	326	ALA	-	expression tag	UNP B2G3D0
F	327	ALA	-	expression tag	UNP B2G3D0
F	328	HIS	-	expression tag	UNP B2G3D0
F	329	HIS	-	expression tag	UNP B2G3D0
F	330	HIS	-	expression tag	UNP B2G3D0
F	331	HIS	-	expression tag	UNP B2G3D0
F	332	HIS	-	expression tag	UNP B2G3D0
F	333	HIS	-	expression tag	UNP B2G3D0
G	325	ALA	-	expression tag	UNP B2G3D0

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
G	326	ALA	-	expression tag	UNP B2G3D0
G	327	ALA	-	expression tag	UNP B2G3D0
G	328	HIS	-	expression tag	UNP B2G3D0
G	329	HIS	-	expression tag	UNP B2G3D0
G	330	HIS	-	expression tag	UNP B2G3D0
G	331	HIS	-	expression tag	UNP B2G3D0
G	332	HIS	-	expression tag	UNP B2G3D0
G	333	HIS	-	expression tag	UNP B2G3D0
H	325	ALA	-	expression tag	UNP B2G3D0
H	326	ALA	-	expression tag	UNP B2G3D0
H	327	ALA	-	expression tag	UNP B2G3D0
H	328	HIS	-	expression tag	UNP B2G3D0
H	329	HIS	-	expression tag	UNP B2G3D0
H	330	HIS	-	expression tag	UNP B2G3D0
H	331	HIS	-	expression tag	UNP B2G3D0
H	332	HIS	-	expression tag	UNP B2G3D0
H	333	HIS	-	expression tag	UNP B2G3D0

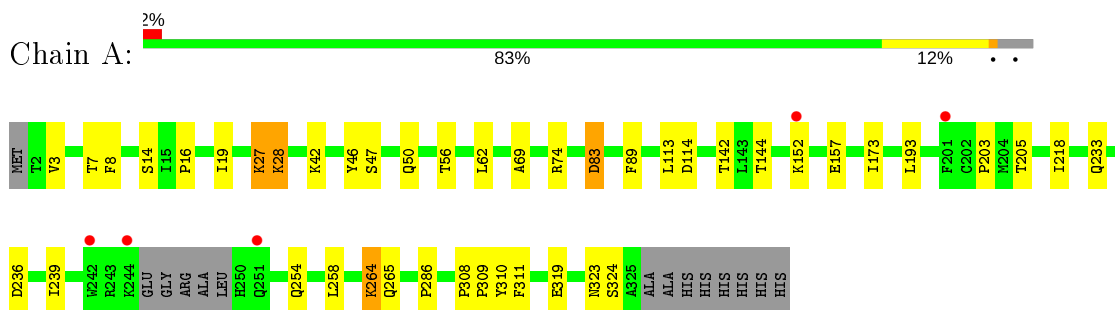
- Molecule 2 is a protein called Microviridin.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	P	12	Total	C	N	O	0	0	0
			96	67	15	14			
2	I	12	Total	C	N	O	0	0	0
			96	67	15	14			
2	J	12	Total	C	N	O	0	0	0
			96	67	15	14			
2	K	12	Total	C	N	O	0	0	0
			96	67	15	14			
2	L	12	Total	C	N	O	0	0	0
			96	67	15	14			
2	M	11	Total	C	N	O	0	0	0
			90	64	14	12			
2	N	13	Total	C	N	O	0	0	0
			105	73	17	15			
2	O	12	Total	C	N	O	0	0	0
			96	67	15	14			

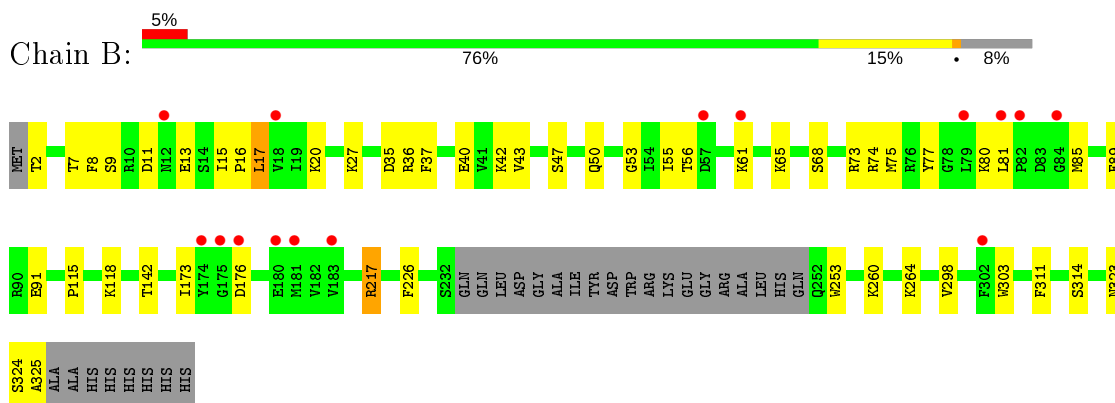
### 3 Residue-property plots [i](#)

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.

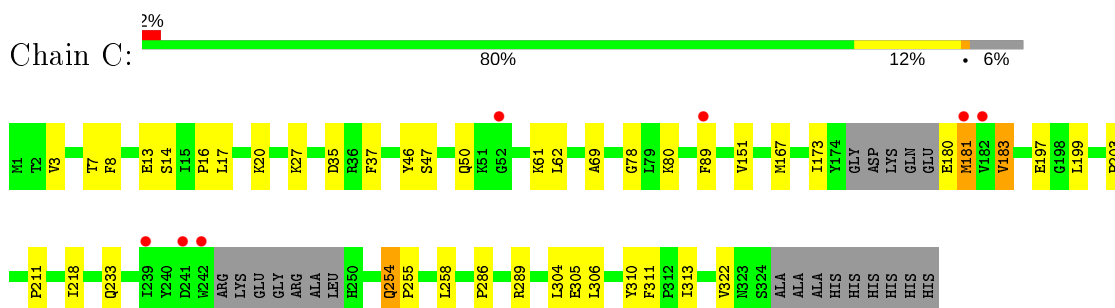
- Molecule 1: ATP grasp ligase



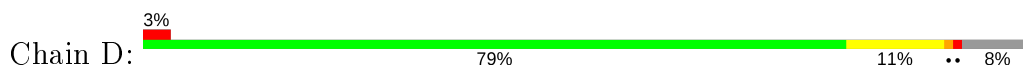
- Molecule 1: ATP grasp ligase



- Molecule 1: ATP grasp ligase

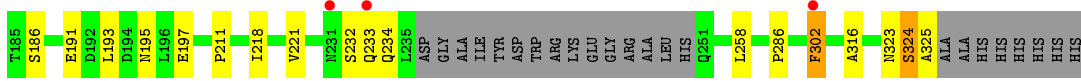


- Molecule 1: ATP grasp ligase

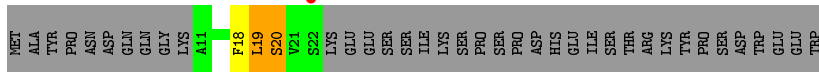




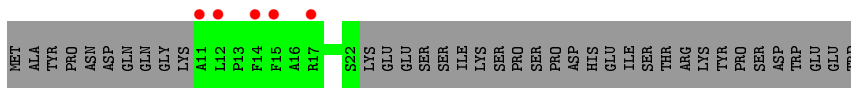




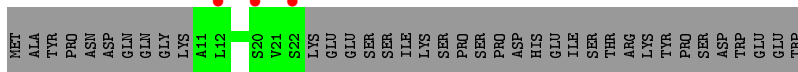
• Molecule 2: Microviridin



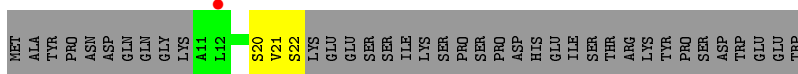
• Molecule 2: Microviridin



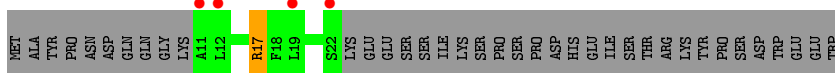
• Molecule 2: Microviridin



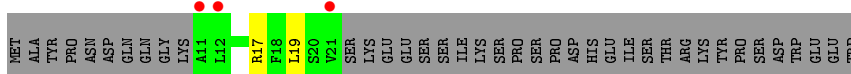
• Molecule 2: Microviridin



• Molecule 2: Microviridin

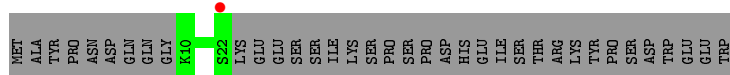


• Molecule 2: Microviridin

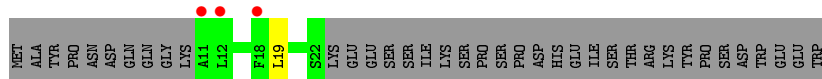


• Molecule 2: Microviridin





- Molecule 2: Microviridin



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 41	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	132.56Å 132.56Å 198.16Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	39.68 – 2.67 39.68 – 2.66	Depositor EDS
% Data completeness (in resolution range)	99.6 (39.68-2.67) 99.6 (39.68-2.66)	Depositor EDS
$R_{merge}$	0.11	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.42 (at 2.65Å)	Xtrriage
Refinement program	PHENIX 1.10.1_2155	Depositor
R, $R_{free}$	0.219 , 0.261 0.218 , 0.256	Depositor DCC
$R_{free}$ test set	4861 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	42.0	Xtrriage
Anisotropy	1.247	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 28.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	0.269 for h,-k,-l	Xtrriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	20643	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	55.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 21.62 % 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 6.7498e-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, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

## 5 Model quality [i](#)

### 5.1 Standard geometry [i](#)

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	0.27	0/2614	0.43	0/3536
1	B	0.26	0/2498	0.42	0/3379
1	C	0.25	0/2557	0.44	0/3460
1	D	0.27	0/2498	0.45	0/3378
1	E	0.27	0/2555	0.44	0/3456
1	F	0.27	0/2462	0.44	0/3329
1	G	0.28	0/2568	0.43	0/3474
1	H	0.26	0/2541	0.44	0/3436
2	I	0.28	0/99	0.44	0/133
2	J	0.31	0/99	0.40	0/133
2	K	0.27	0/99	0.36	0/133
2	L	0.27	0/99	0.40	0/133
2	M	0.26	0/93	0.40	0/125
2	N	0.27	0/108	0.39	0/144
2	O	0.27	0/99	0.36	0/133
2	P	0.33	0/99	0.80	0/133
All	All	0.27	0/21088	0.44	0/28515

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

### 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	2559	0	2566	25	0
1	B	2446	0	2460	33	0
1	C	2503	0	2512	29	0
1	D	2447	0	2467	24	0
1	E	2502	0	2516	22	0
1	F	2412	0	2433	31	0
1	G	2514	0	2525	19	0
1	H	2489	0	2507	32	0
2	I	96	0	97	0	0
2	J	96	0	97	0	0
2	K	96	0	97	4	0
2	L	96	0	97	3	0
2	M	90	0	92	2	0
2	N	105	0	110	0	0
2	O	96	0	97	1	0
2	P	96	0	97	8	0
All	All	20643	0	20770	195	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 5.

The worst 5 of 195 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:L:17:ARG:HH11	2:L:17:ARG:HG2	1.45	0.80
1:B:11:ASP:HB3	1:B:15:ILE:HD11	1.67	0.77
1:H:76:ARG:HB3	1:H:79:LEU:HD23	1.67	0.76
1:D:76:ARG:HB3	1:D:79:LEU:HD23	1.69	0.74
1:C:181:MET:H	2:K:21:VAL:HA	1.54	0.73

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	315/333 (95%)	296 (94%)	18 (6%)	1 (0%)	41	56
1	B	302/333 (91%)	285 (94%)	15 (5%)	2 (1%)	22	33
1	C	306/333 (92%)	292 (95%)	13 (4%)	1 (0%)	41	56
1	D	302/333 (91%)	285 (94%)	14 (5%)	3 (1%)	15	23
1	E	306/333 (92%)	287 (94%)	17 (6%)	2 (1%)	22	33
1	F	296/333 (89%)	281 (95%)	13 (4%)	2 (1%)	22	33
1	G	307/333 (92%)	292 (95%)	14 (5%)	1 (0%)	41	56
1	H	307/333 (92%)	289 (94%)	15 (5%)	3 (1%)	15	23
2	I	10/49 (20%)	9 (90%)	1 (10%)	0	100	100
2	J	10/49 (20%)	9 (90%)	1 (10%)	0	100	100
2	K	10/49 (20%)	9 (90%)	1 (10%)	0	100	100
2	L	10/49 (20%)	10 (100%)	0	0	100	100
2	M	9/49 (18%)	8 (89%)	1 (11%)	0	100	100
2	N	11/49 (22%)	10 (91%)	1 (9%)	0	100	100
2	O	10/49 (20%)	9 (90%)	1 (10%)	0	100	100
2	P	10/49 (20%)	8 (80%)	1 (10%)	1 (10%)	0	0
All	All	2521/3056 (82%)	2379 (94%)	126 (5%)	16 (1%)	25	37

5 of 16 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	176	ASP
1	D	178	GLN
2	P	20	SER
1	B	324	SER
1	G	301	PHE

### 5.3.2 Protein sidechains [i](#)

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	278/288 (96%)	268 (96%)	10 (4%)	35	51
1	B	267/288 (93%)	261 (98%)	6 (2%)	52	70
1	C	273/288 (95%)	266 (97%)	7 (3%)	46	64
1	D	268/288 (93%)	258 (96%)	10 (4%)	34	50
1	E	273/288 (95%)	267 (98%)	6 (2%)	52	70
1	F	264/288 (92%)	254 (96%)	10 (4%)	33	49
1	G	274/288 (95%)	272 (99%)	2 (1%)	84	91
1	H	272/288 (94%)	260 (96%)	12 (4%)	28	43
2	I	10/45 (22%)	10 (100%)	0	100	100
2	J	10/45 (22%)	10 (100%)	0	100	100
2	K	10/45 (22%)	10 (100%)	0	100	100
2	L	10/45 (22%)	9 (90%)	1 (10%)	7	10
2	M	9/45 (20%)	8 (89%)	1 (11%)	6	8
2	N	11/45 (24%)	11 (100%)	0	100	100
2	O	10/45 (22%)	10 (100%)	0	100	100
2	P	10/45 (22%)	9 (90%)	1 (10%)	7	10
All	All	2249/2664 (84%)	2183 (97%)	66 (3%)	42	60

5 of 66 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	176	ASP
1	E	75	MET
1	H	234	GLN
1	D	178	GLN
1	E	10	ARG

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

Mol	Chain	Res	Type
1	C	254	GLN

### 5.3.3 RNA ⓘ

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	319/333 (95%)	0.33	5 (1%) 72 69	38, 50, 85, 116	0
1	B	305/333 (91%)	0.45	15 (4%) 29 26	40, 53, 89, 107	0
1	C	312/333 (93%)	0.34	7 (2%) 62 57	35, 51, 79, 96	0
1	D	306/333 (91%)	0.38	10 (3%) 46 43	38, 52, 84, 108	0
1	E	312/333 (93%)	0.34	12 (3%) 40 36	36, 52, 79, 105	0
1	F	302/333 (90%)	0.35	11 (3%) 42 39	37, 50, 77, 98	0
1	G	313/333 (93%)	0.32	9 (2%) 51 48	37, 48, 78, 101	0
1	H	310/333 (93%)	0.36	12 (3%) 39 35	38, 50, 80, 106	0
2	I	12/49 (24%)	1.71	5 (41%) 0 0	76, 87, 97, 104	0
2	J	12/49 (24%)	1.56	3 (25%) 0 0	69, 77, 94, 95	0
2	K	12/49 (24%)	0.84	1 (8%) 11 9	69, 77, 91, 94	0
2	L	12/49 (24%)	1.94	4 (33%) 0 0	66, 75, 95, 97	0
2	M	11/49 (22%)	1.42	3 (27%) 0 0	69, 79, 89, 94	0
2	N	13/49 (26%)	0.75	1 (7%) 13 10	54, 60, 85, 92	0
2	O	12/49 (24%)	1.09	3 (25%) 0 0	68, 77, 95, 98	0
2	P	12/49 (24%)	1.19	1 (8%) 11 9	60, 68, 84, 87	0
All	All	2575/3056 (84%)	0.39	102 (3%) 38 34	35, 52, 84, 116	0

The worst 5 of 102 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	L	12	LEU	6.2
2	I	12	LEU	4.7
1	D	177	LYS	4.4
1	C	239	ILE	4.2
1	B	84	GLY	3.9

## 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.