



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 17, 2020 – 06:33 am BST

PDB ID : 5X11  
Title : Crystal structure of Bacillus subtilis PadR in complex with operator DNA  
Authors : Park, S.C.; Kwak, Y.M.; Song, W.S.; Hong, M.; Yoon, S.I.  
Deposited on : 2017-01-24  
Resolution : 2.65 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

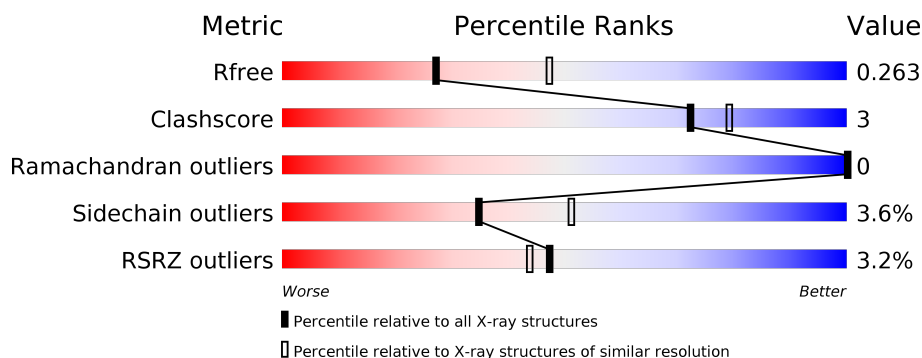
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.65 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	I	28	<div> <div></div> <div> <div>39%</div> <div>43%</div> <div>18%</div> </div> </div>
1	M	28	<div> <div>4%</div> <div> <div>46%</div> <div>43%</div> <div>11%</div> </div> </div>
2	J	28	<div> <div>25%</div> <div>71%</div> <div>•</div> </div>
2	N	28	<div> <div>54%</div> <div>39%</div> <div>7%</div> </div>
3	A	188	<div> <div>6%</div> <div> <div>86%</div> <div>9%</div> <div>5%</div> </div> </div>
3	B	188	<div> <div>2%</div> <div> <div>82%</div> <div>11%</div> <div>• 5%</div> </div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	C	188	<div><div></div><div>2%</div><div>84%</div><div>10%</div><div>6%</div></div>
3	D	188	<div><div></div><div>3%</div><div>84%</div><div>10%</div><div>6%</div></div>
3	E	188	<div><div></div><div>3%</div><div>88%</div><div>9%</div><div>•</div></div>
3	F	188	<div><div></div><div>4%</div><div>90%</div><div>•</div><div>5%</div></div>
3	G	188	<div><div></div><div>2%</div><div>84%</div><div>11%</div><div>6%</div></div>
3	H	188	<div><div></div><div>5%</div><div>85%</div><div>7%</div><div>•</div><div>8%</div></div>

## 2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 13422 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 DNA chain called DNA (28-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	I	28	Total	C	N	O	P	0	0	0
			574	276	108	163	27			
1	M	28	Total	C	N	O	P	0	0	0
			574	276	108	163	27			

- Molecule 2 is a DNA chain called DNA (28-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	J	28	Total	C	N	O	P	0	0	0
			568	275	97	169	27			
2	N	28	Total	C	N	O	P	0	0	0
			568	275	97	169	27			

- Molecule 3 is a protein called Transcriptional regulator.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	A	178	Total	C	N	O	S	0	0	0
			1411	920	229	256	6			
3	B	179	Total	C	N	O	S	0	0	0
			1427	930	231	260	6			
3	C	176	Total	C	N	O	S	0	0	0
			1391	908	223	256	4			
3	D	177	Total	C	N	O	S	0	0	0
			1413	922	225	261	5			
3	E	181	Total	C	N	O	S	0	0	0
			1395	914	223	252	6			
3	F	178	Total	C	N	O	S	0	0	0
			1383	900	227	251	5			
3	G	177	Total	C	N	O	S	0	0	0
			1383	902	225	252	4			
3	H	173	Total	C	N	O	S	0	0	0
			1318	861	215	238	4			

There are 48 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-5	GLY	-	expression tag	UNP E0TW95
A	-4	SER	-	expression tag	UNP E0TW95
A	-3	ALA	-	expression tag	UNP E0TW95
A	-2	LYS	-	expression tag	UNP E0TW95
A	-1	ASP	-	expression tag	UNP E0TW95
A	0	PRO	-	expression tag	UNP E0TW95
B	-5	GLY	-	expression tag	UNP E0TW95
B	-4	SER	-	expression tag	UNP E0TW95
B	-3	ALA	-	expression tag	UNP E0TW95
B	-2	LYS	-	expression tag	UNP E0TW95
B	-1	ASP	-	expression tag	UNP E0TW95
B	0	PRO	-	expression tag	UNP E0TW95
C	-5	GLY	-	expression tag	UNP E0TW95
C	-4	SER	-	expression tag	UNP E0TW95
C	-3	ALA	-	expression tag	UNP E0TW95
C	-2	LYS	-	expression tag	UNP E0TW95
C	-1	ASP	-	expression tag	UNP E0TW95
C	0	PRO	-	expression tag	UNP E0TW95
D	-5	GLY	-	expression tag	UNP E0TW95
D	-4	SER	-	expression tag	UNP E0TW95
D	-3	ALA	-	expression tag	UNP E0TW95
D	-2	LYS	-	expression tag	UNP E0TW95
D	-1	ASP	-	expression tag	UNP E0TW95
D	0	PRO	-	expression tag	UNP E0TW95
E	-5	GLY	-	expression tag	UNP E0TW95
E	-4	SER	-	expression tag	UNP E0TW95
E	-3	ALA	-	expression tag	UNP E0TW95
E	-2	LYS	-	expression tag	UNP E0TW95
E	-1	ASP	-	expression tag	UNP E0TW95
E	0	PRO	-	expression tag	UNP E0TW95
F	-5	GLY	-	expression tag	UNP E0TW95
F	-4	SER	-	expression tag	UNP E0TW95
F	-3	ALA	-	expression tag	UNP E0TW95
F	-2	LYS	-	expression tag	UNP E0TW95
F	-1	ASP	-	expression tag	UNP E0TW95
F	0	PRO	-	expression tag	UNP E0TW95
G	-5	GLY	-	expression tag	UNP E0TW95
G	-4	SER	-	expression tag	UNP E0TW95
G	-3	ALA	-	expression tag	UNP E0TW95
G	-2	LYS	-	expression tag	UNP E0TW95
G	-1	ASP	-	expression tag	UNP E0TW95
G	0	PRO	-	expression tag	UNP E0TW95

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
H	-5	GLY	-	expression tag	UNP E0TW95
H	-4	SER	-	expression tag	UNP E0TW95
H	-3	ALA	-	expression tag	UNP E0TW95
H	-2	LYS	-	expression tag	UNP E0TW95
H	-1	ASP	-	expression tag	UNP E0TW95
H	0	PRO	-	expression tag	UNP E0TW95

- Molecule 4 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	I	4	Total O 4 4	0	0
4	J	4	Total O 4 4	0	0
4	N	1	Total O 1 1	0	0
4	A	2	Total O 2 2	0	0
4	B	1	Total O 1 1	0	0
4	C	1	Total O 1 1	0	0
4	D	2	Total O 2 2	0	0
4	G	2	Total O 2 2	0	0

### 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: DNA (28-MER)

Chain I: 



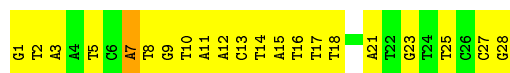
- Molecule 1: DNA (28-MER)

Chain M: 



- Molecule 2: DNA (28-MER)

Chain J: 




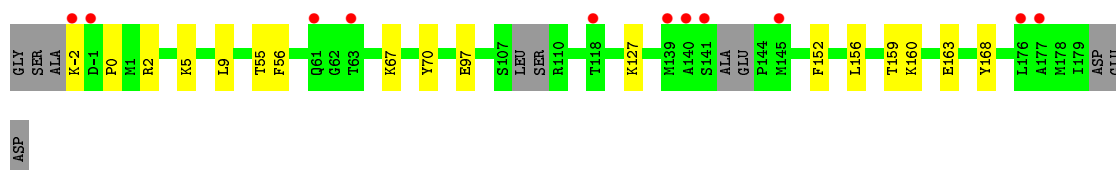
- Molecule 2: DNA (28-MER)

Chain N: 

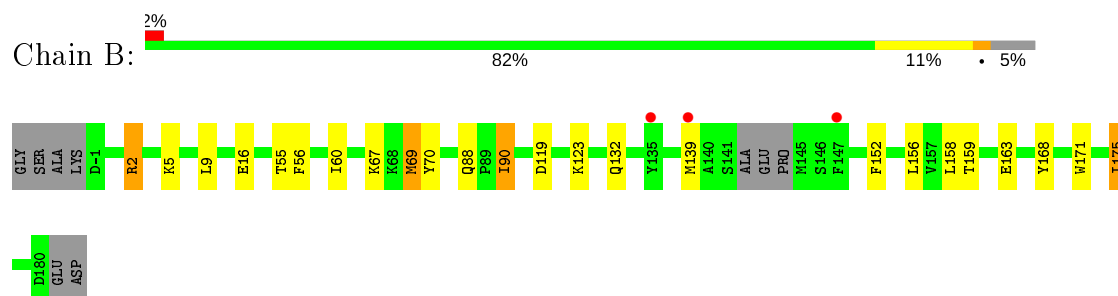


- Molecule 3: Transcriptional regulator

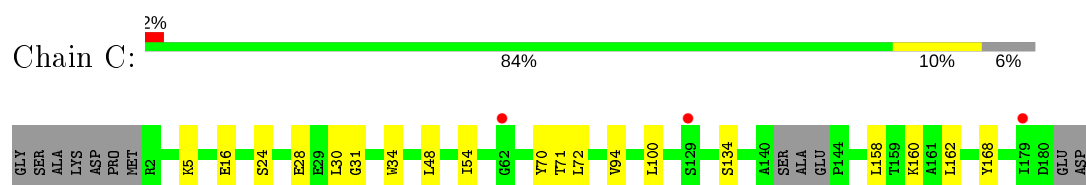
Chain A: 



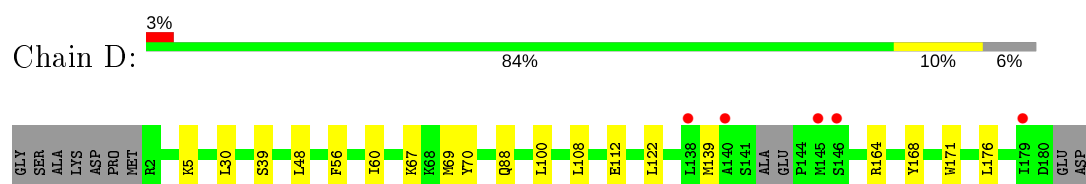
- Molecule 3: Transcriptional regulator



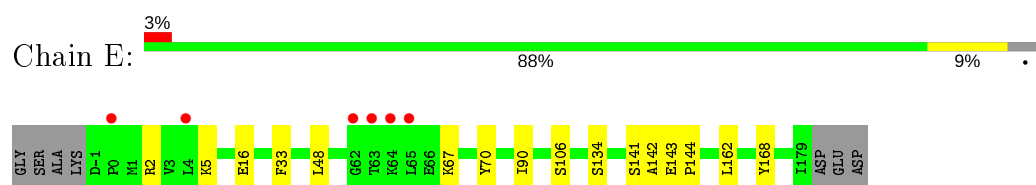
- Molecule 3: Transcriptional regulator



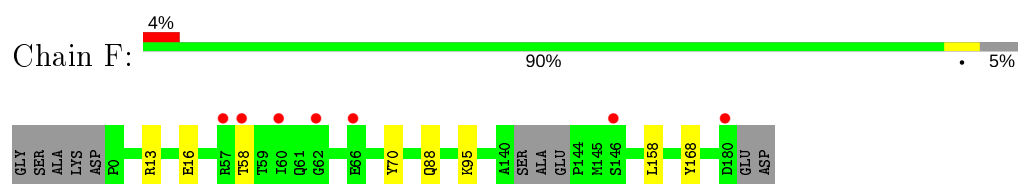
- Molecule 3: Transcriptional regulator



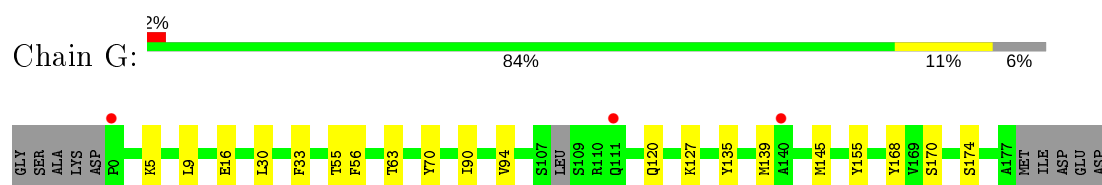
- Molecule 3: Transcriptional regulator



- Molecule 3: Transcriptional regulator

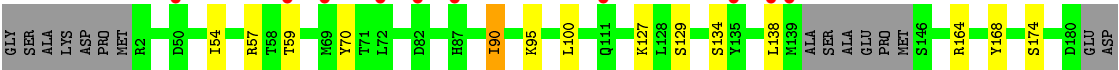
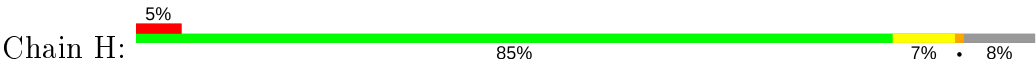


- Molecule 3: Transcriptional regulator



- Molecule 3: Transcriptional regulator





## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	341.81Å 61.57Å 120.55Å 90.00° 107.94° 90.00°	Depositor
Resolution (Å)	30.00 – 2.65 29.73 – 2.64	Depositor EDS
% Data completeness (in resolution range)	99.3 (30.00-2.65) 98.6 (29.73-2.64)	Depositor EDS
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.53 (at 2.64Å)	Xtriage
Refinement program	REFMAC 5.5.0109	Depositor
R, $R_{free}$	0.228 , 0.268 0.225 , 0.263	Depositor DCC
$R_{free}$ test set	3346 reflections (4.79%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	57.9	Xtriage
Anisotropy	0.688	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 48.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.51$ , $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	0.008 for -h-2*k,l	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	13422	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	75.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.47% 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	I	1.08	0/645	1.78	22/994 (2.2%)
1	M	0.75	0/645	1.46	11/994 (1.1%)
2	J	1.05	0/635	1.92	32/978 (3.3%)
2	N	0.78	0/635	1.61	16/978 (1.6%)
3	A	0.45	0/1442	0.56	0/1940
3	B	0.43	0/1458	0.54	0/1963
3	C	0.42	0/1422	0.54	0/1919
3	D	0.47	0/1444	0.57	1/1945 (0.1%)
3	E	0.35	0/1428	0.50	0/1934
3	F	0.36	0/1415	0.48	0/1910
3	G	0.44	0/1415	0.56	0/1909
3	H	0.37	0/1347	0.51	0/1823
All	All	0.55	0/13931	0.90	82/19287 (0.4%)

There are no bond length outliers.

The worst 5 of 82 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
1	I	21	DA	O4'-C1'-N9	10.54	115.38	108.00
2	J	9	DG	C1'-O4'-C4'	-10.29	99.81	110.10
2	N	2	DT	O4'-C1'-N1	10.09	115.06	108.00
1	I	7	DA	O4'-C4'-C3'	-9.73	100.16	106.00
2	J	9	DG	O4'-C1'-N9	9.02	114.31	108.00

There are no chirality outliers.

There are no planarity outliers.

### 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	I	574	0	318	12	0
1	M	574	0	318	10	0
2	J	568	0	321	2	0
2	N	568	0	321	4	0
3	A	1411	0	1372	15	0
3	B	1427	0	1414	15	0
3	C	1391	0	1352	10	0
3	D	1413	0	1388	10	0
3	E	1395	0	1330	10	0
3	F	1383	0	1318	3	0
3	G	1383	0	1336	11	0
3	H	1318	0	1237	9	0
4	A	2	0	0	0	0
4	B	1	0	0	0	0
4	C	1	0	0	0	0
4	D	2	0	0	0	0
4	G	2	0	0	0	0
4	I	4	0	0	0	0
4	J	4	0	0	0	0
4	N	1	0	0	0	0
All	All	13422	0	12025	83	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 3.

The worst 5 of 83 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)
3:A:127:LYS:HD2	3:B:90:ILE:HD11	1.28	1.16
1:I:20:DC:H2"	1:I:21:DA:H5"	1.35	1.07
3:C:5:LYS:HG3	3:C:48:LEU:HD21	1.61	0.80
3:B:159:THR:O	3:B:163:GLU:HG2	1.89	0.73
3:D:5:LYS:HG3	3:D:48:LEU:HD21	1.71	0.72

There are no symmetry-related clashes.

## 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	
3	A	172/188 (92%)	170 (99%)	2 (1%)	0	100	100
3	B	175/188 (93%)	172 (98%)	3 (2%)	0	100	100
3	C	172/188 (92%)	168 (98%)	4 (2%)	0	100	100
3	D	173/188 (92%)	171 (99%)	2 (1%)	0	100	100
3	E	179/188 (95%)	176 (98%)	3 (2%)	0	100	100
3	F	174/188 (93%)	171 (98%)	3 (2%)	0	100	100
3	G	173/188 (92%)	169 (98%)	4 (2%)	0	100	100
3	H	169/188 (90%)	168 (99%)	1 (1%)	0	100	100
All	All	1387/1504 (92%)	1365 (98%)	22 (2%)	0	100	100

There are no Ramachandran outliers to report.

### 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	
3	A	141/168 (84%)	139 (99%)	2 (1%)	67	81
3	B	148/168 (88%)	138 (93%)	10 (7%)	16	24
3	C	141/168 (84%)	137 (97%)	4 (3%)	43	61
3	D	146/168 (87%)	141 (97%)	5 (3%)	37	53
3	E	134/168 (80%)	131 (98%)	3 (2%)	52	70
3	F	135/168 (80%)	131 (97%)	4 (3%)	41	59

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
3	G	137/168 (82%)	131 (96%)	6 (4%)	28	43
3	H	124/168 (74%)	118 (95%)	6 (5%)	25	39
All	All	1106/1344 (82%)	1066 (96%)	40 (4%)	35	51

5 of 40 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	D	139	MET
3	E	162	LEU
3	H	134	SER
3	D	176	LEU
3	E	168	TYR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 15 such sidechains are listed below:

Mol	Chain	Res	Type
3	E	154	HIS
3	E	167	ASN
3	G	154	HIS
3	E	120	GLN
3	F	167	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	I	28/28 (100%)	-0.30	0 100 100	35, 44, 74, 90	0
1	M	28/28 (100%)	-0.04	1 (3%) 42 39	67, 100, 120, 125	0
2	J	28/28 (100%)	-0.10	0 100 100	33, 43, 85, 93	0
2	N	28/28 (100%)	-0.06	0 100 100	74, 92, 131, 134	0
3	A	178/188 (94%)	0.27	11 (6%) 20 17	35, 62, 104, 126	0
3	B	179/188 (95%)	0.06	3 (1%) 70 67	38, 66, 105, 126	0
3	C	176/188 (93%)	0.14	3 (1%) 70 67	49, 75, 109, 133	0
3	D	177/188 (94%)	0.14	5 (2%) 53 49	36, 59, 91, 111	0
3	E	181/188 (96%)	0.31	6 (3%) 46 43	61, 89, 115, 155	0
3	F	178/188 (94%)	0.31	7 (3%) 39 35	59, 82, 131, 152	0
3	G	177/188 (94%)	0.06	3 (1%) 70 67	37, 60, 123, 148	0
3	H	173/188 (92%)	0.34	10 (5%) 23 19	45, 101, 154, 166	0
All	All	1531/1616 (94%)	0.18	49 (3%) 47 44	33, 74, 128, 166	0

The worst 5 of 49 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	A	140	ALA	5.4
3	A	118	THR	4.4
3	A	-1	ASP	4.2
3	A	177	ALA	4.2
3	H	139	MET	3.8

### 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.