



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 15, 2017 – 01:20 am GMT

PDB ID : 5HBU  
Title : Structure of the E. coli nucleoid occlusion protein SlmA bound to DNA and the C-terminal tail of the cytoskeletal cell division protein FtsZ  
Authors : Schumacher, M.A.; Zeng, W.  
Deposited on : 2016-01-02  
Resolution : 2.60 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	:	trunk28620
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

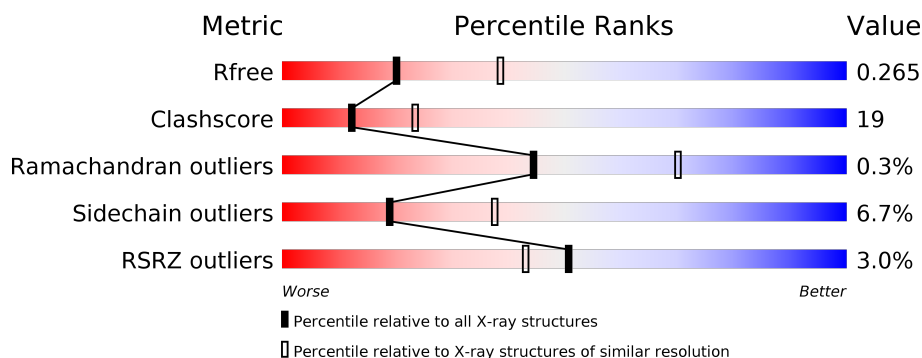
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.60 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	2542 (2.60-2.60)
Clashscore	112137	2895 (2.60-2.60)
Ramachandran outliers	110173	2848 (2.60-2.60)
Sidechain outliers	110143	2848 (2.60-2.60)
RSRZ outliers	101464	2550 (2.60-2.60)

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	196	<div> <div>6%</div> <div> <div></div> <div>54%</div> <div>41%</div> <div>...</div> </div> </div>
1	B	196	<div> <div>9%</div> <div> <div></div> <div>60%</div> <div>35%</div> <div>5%</div> </div> </div>
1	C	196	<div> <div>%</div> <div> <div></div> <div>62%</div> <div>33%</div> <div>..</div> </div> </div>
1	D	196	<div> <div></div> <div> <div></div> <div>63%</div> <div>30%</div> <div>..</div> </div> </div>
1	E	196	<div> <div>%</div> <div> <div></div> <div>59%</div> <div>34%</div> <div>..</div> </div> </div>
1	F	196	<div> <div>%</div> <div> <div></div> <div>68%</div> <div>24%</div> <div>5%</div> <div>.</div> </div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	G	196	<div><div></div><div>3%</div><div>65%</div><div>28%</div><div></div><div></div></div>
1	H	196	<div><div></div><div>2%</div><div>61%</div><div>33%</div><div></div><div></div></div>
2	R	12	<div><div></div><div>42%</div><div>58%</div><div></div><div></div></div>
2	T	12	<div><div></div><div>42%</div><div>58%</div><div></div><div></div></div>
2	W	12	<div><div></div><div>33%</div><div>67%</div><div></div><div></div></div>
2	Z	12	<div><div></div><div>33%</div><div>67%</div><div></div><div></div></div>
3	K	10	<div><div></div><div>60%</div><div>70%</div><div>10%</div><div>20%</div><div></div></div>

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 13710 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 Nucleoid occlusion factor SlmA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	190	Total	C	N	O	S	0	0	0
			1543	970	281	286	6			
1	B	196	Total	C	N	O	S	0	0	0
			1589	998	290	295	6			
1	C	190	Total	C	N	O	S	0	0	0
			1543	970	281	286	6			
1	D	189	Total	C	N	O	S	0	0	0
			1532	964	277	285	6			
1	E	190	Total	C	N	O	S	0	0	0
			1543	970	281	286	6			
1	F	189	Total	C	N	O	S	0	0	0
			1532	964	277	285	6			
1	G	190	Total	C	N	O	S	0	0	0
			1543	970	281	286	6			
1	H	190	Total	C	N	O	S	0	0	0
			1543	970	281	286	6			

There are 32 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	3	ALA	-	expression tag	UNP P0C093
A	4	GLU	-	expression tag	UNP P0C093
A	5	LYS	-	expression tag	UNP P0C093
A	6	GLN	-	expression tag	UNP P0C093
B	3	ALA	-	expression tag	UNP P0C093
B	4	GLU	-	expression tag	UNP P0C093
B	5	LYS	-	expression tag	UNP P0C093
B	6	GLN	-	expression tag	UNP P0C093
C	3	ALA	-	expression tag	UNP P0C093
C	4	GLU	-	expression tag	UNP P0C093
C	5	LYS	-	expression tag	UNP P0C093
C	6	GLN	-	expression tag	UNP P0C093
D	3	ALA	-	expression tag	UNP P0C093

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
D	4	GLU	-	expression tag	UNP P0C093
D	5	LYS	-	expression tag	UNP P0C093
D	6	GLN	-	expression tag	UNP P0C093
E	3	ALA	-	expression tag	UNP P0C093
E	4	GLU	-	expression tag	UNP P0C093
E	5	LYS	-	expression tag	UNP P0C093
E	6	GLN	-	expression tag	UNP P0C093
F	3	ALA	-	expression tag	UNP P0C093
F	4	GLU	-	expression tag	UNP P0C093
F	5	LYS	-	expression tag	UNP P0C093
F	6	GLN	-	expression tag	UNP P0C093
G	3	ALA	-	expression tag	UNP P0C093
G	4	GLU	-	expression tag	UNP P0C093
G	5	LYS	-	expression tag	UNP P0C093
G	6	GLN	-	expression tag	UNP P0C093
H	3	ALA	-	expression tag	UNP P0C093
H	4	GLU	-	expression tag	UNP P0C093
H	5	LYS	-	expression tag	UNP P0C093
H	6	GLN	-	expression tag	UNP P0C093

- Molecule 2 is a DNA chain called DNA (5'-D(\*GP\*TP\*GP\*AP\*GP\*TP\*AP\*CP\*TP\*CP\*AP\*C)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	W	12	Total	C	N	O	P	0	0	0
			243	117	45	70	11			
2	Z	12	Total	C	N	O	P	0	0	0
			243	117	45	70	11			
2	R	12	Total	C	N	O	P	0	0	0
			243	117	45	70	11			
2	T	12	Total	C	N	O	P	0	0	0
			243	117	45	70	11			

- Molecule 3 is a protein called FtsZ CTT peptide.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	K	8	Total	C	N	O	0	0	0
			67	45	11	11			

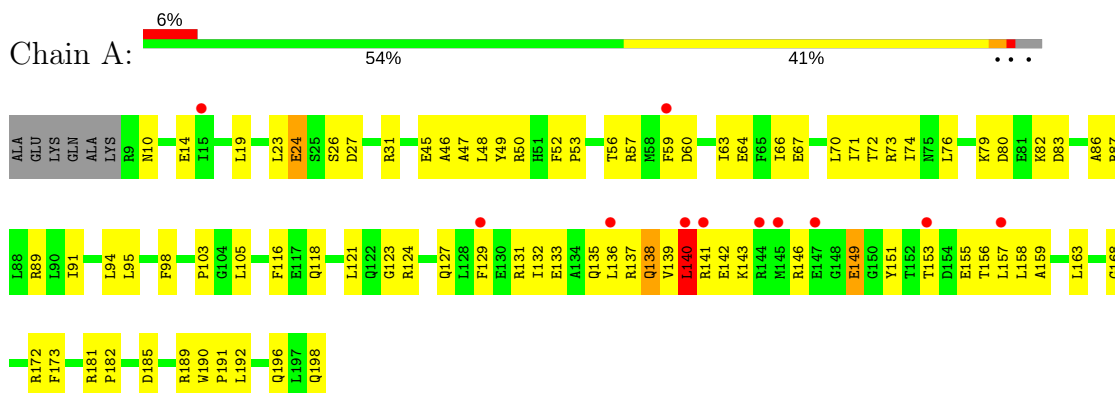
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	17	Total 17	O 17	0	0
4	B	10	Total 10	O 10	0	0
4	C	58	Total 58	O 58	0	0
4	D	48	Total 48	O 48	0	0
4	E	28	Total 28	O 28	0	0
4	F	29	Total 29	O 29	0	0
4	G	50	Total 50	O 50	0	0
4	H	36	Total 36	O 36	0	0
4	W	9	Total 9	O 9	0	0
4	Z	7	Total 7	O 7	0	0
4	R	7	Total 7	O 7	0	0
4	T	2	Total 2	O 2	0	0
4	K	2	Total 2	O 2	0	0

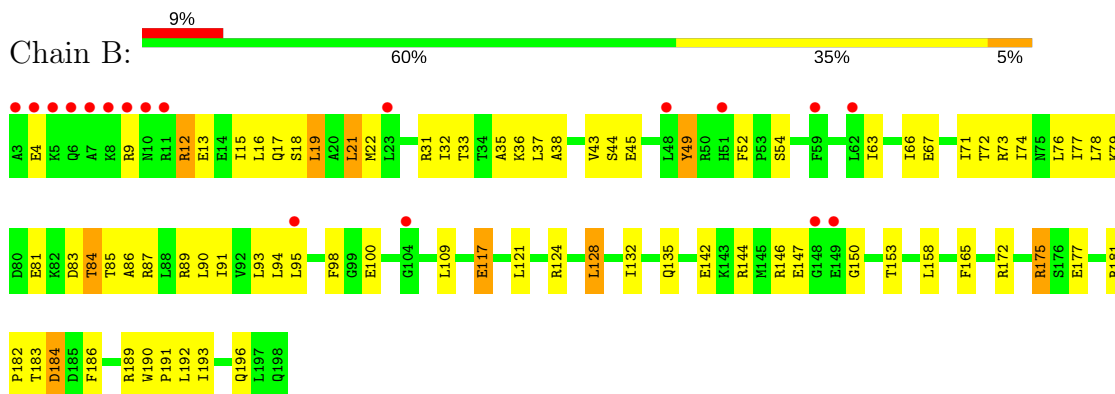
### 3 Residue-property plots [i](#)

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.

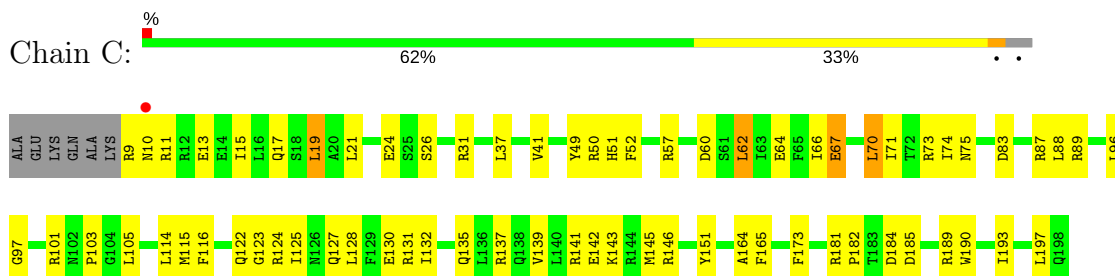
- Molecule 1: Nucleoid occlusion factor SlmA



- Molecule 1: Nucleoid occlusion factor SlmA

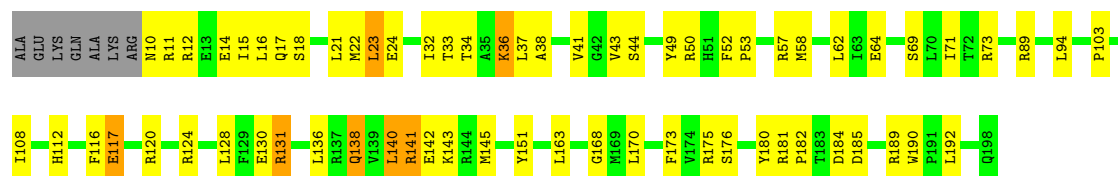


- Molecule 1: Nucleoid occlusion factor SlmA



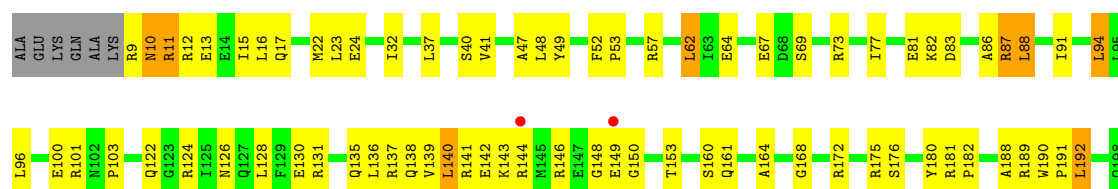
- Molecule 1: Nucleoid occlusion factor SlmA

Chain D:  63% 30%



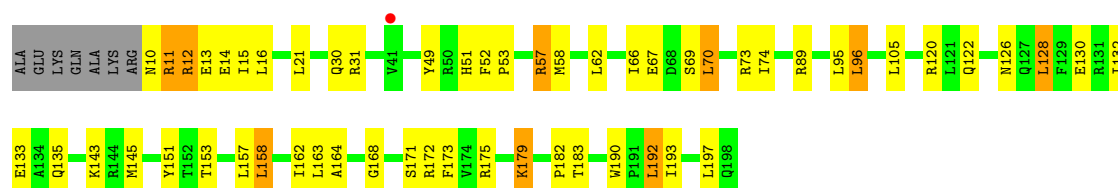
• Molecule 1: Nucleoid occlusion factor SlmA

Chain E:  59% 34%



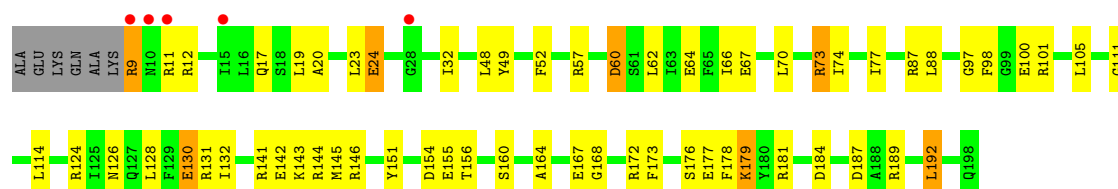
• Molecule 1: Nucleoid occlusion factor SlmA

Chain F:  68% 24% 5%



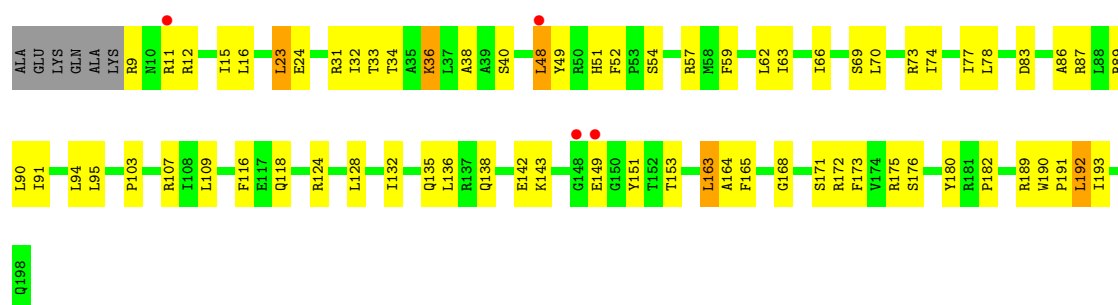
• Molecule 1: Nucleoid occlusion factor SlmA

Chain G:  65% 28% 3%



• Molecule 1: Nucleoid occlusion factor SlmA

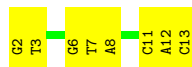
Chain H:  61% 33% 2%





- Molecule 2: DNA (5'-D(\*GP\*TP\*GP\*AP\*GP\*TP\*AP\*CP\*TP\*CP\*AP\*C)-3')

Chain W:  33% 67%



- Molecule 2: DNA (5'-D(\*GP\*TP\*GP\*AP\*GP\*TP\*AP\*CP\*TP\*CP\*AP\*C)-3')

Chain Z:  33% 67%



- Molecule 2: DNA (5'-D(\*GP\*TP\*GP\*AP\*GP\*TP\*AP\*CP\*TP\*CP\*AP\*C)-3')

Chain R:  42% 58%



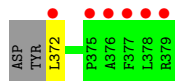
- Molecule 2: DNA (5'-D(\*GP\*TP\*GP\*AP\*GP\*TP\*AP\*CP\*TP\*CP\*AP\*C)-3')

Chain T:  42% 58%



- Molecule 3: FtsZ CTT peptide

Chain K:  60% 70% 10% 20%



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	70.11Å 158.89Å 200.20Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	100.10 – 2.60 100.10 – 2.60	Depositor EDS
% Data completeness (in resolution range)	99.5 (100.10-2.60) 99.7 (100.10-2.60)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.12	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.21 (at 2.62Å)	Xtriage
Refinement program	PHENIX	Depositor
R, $R_{free}$	0.226 , 0.270 0.220 , 0.265	Depositor DCC
$R_{free}$ test set	3805 reflections (2.88%)	DCC
Wilson B-factor (Å <sup>2</sup> )	37.3	Xtriage
Anisotropy	0.808	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 58.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	13710	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	55.0	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 36.13 % 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 5.2397e-04. 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	0.24	0/1563	0.42	1/2098 (0.0%)
1	B	0.25	0/1609	0.44	0/2158
1	C	0.21	0/1563	0.36	0/2098
1	D	0.26	0/1552	0.41	0/2084
1	E	0.22	0/1563	0.37	0/2098
1	F	0.21	0/1552	0.37	0/2084
1	G	0.23	0/1563	0.38	0/2098
1	H	0.22	0/1563	0.38	0/2098
2	R	0.44	0/272	0.97	0/418
2	T	0.42	0/272	1.00	0/418
2	W	0.42	0/272	0.99	0/418
2	Z	0.42	0/272	1.04	0/418
3	K	0.21	0/68	0.48	0/90
All	All	0.25	0/13684	0.48	1/18578 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	140	LEU	CB-CG-CD2	-5.03	102.46	111.00

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	1543	0	1579	84	0
1	B	1589	0	1629	87	0
1	C	1543	0	1579	58	0
1	D	1532	0	1566	67	0
1	E	1543	0	1579	68	0
1	F	1532	0	1566	46	0
1	G	1543	0	1579	47	0
1	H	1543	0	1579	59	0
2	R	243	0	137	9	0
2	T	243	0	137	8	0
2	W	243	0	137	6	0
2	Z	243	0	137	13	0
3	K	67	0	70	2	0
4	A	17	0	0	10	0
4	B	10	0	0	2	0
4	C	58	0	0	7	0
4	D	48	0	0	3	0
4	E	28	0	0	2	0
4	F	29	0	0	1	0
4	G	50	0	0	4	0
4	H	36	0	0	2	0
4	K	2	0	0	0	0
4	R	7	0	0	2	0
4	T	2	0	0	0	0
4	W	9	0	0	0	0
4	Z	7	0	0	0	0
All	All	13710	0	13274	505	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 19.

The worst 5 of 505 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:140:LEU:CD2	1:A:158:LEU:HD11	1.75	1.16
1:D:36:LYS:HB2	1:D:36:LYS:HZ3	1.08	1.11
1:A:140:LEU:HD21	1:A:158:LEU:CD1	1.86	1.05
1:D:64:GLU:HG3	1:D:124:ARG:HH22	1.25	1.01
1:C:73:ARG:HH12	1:C:101:ARG:HH12	1.07	1.00

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	
1	A	188/196 (96%)	176 (94%)	11 (6%)	1 (0%)	32	58
1	B	194/196 (99%)	171 (88%)	21 (11%)	2 (1%)	18	37
1	C	188/196 (96%)	181 (96%)	7 (4%)	0	100	100
1	D	187/196 (95%)	184 (98%)	3 (2%)	0	100	100
1	E	188/196 (96%)	180 (96%)	7 (4%)	1 (0%)	32	58
1	F	187/196 (95%)	185 (99%)	2 (1%)	0	100	100
1	G	188/196 (96%)	181 (96%)	7 (4%)	0	100	100
1	H	188/196 (96%)	181 (96%)	7 (4%)	0	100	100
3	K	6/10 (60%)	5 (83%)	1 (17%)	0	100	100
All	All	1514/1578 (96%)	1444 (95%)	66 (4%)	4 (0%)	44	70

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	12	ARG
1	E	10	ASN
1	A	53	PRO
1	B	84	THR

### 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	166/170 (98%)	162 (98%)	4 (2%)	54	80
1	B	170/170 (100%)	160 (94%)	10 (6%)	23	45
1	C	166/170 (98%)	156 (94%)	10 (6%)	22	44
1	D	165/170 (97%)	152 (92%)	13 (8%)	14	28
1	E	166/170 (98%)	152 (92%)	14 (8%)	13	25
1	F	165/170 (97%)	153 (93%)	12 (7%)	16	33
1	G	166/170 (98%)	150 (90%)	16 (10%)	10	18
1	H	166/170 (98%)	156 (94%)	10 (6%)	22	44
3	K	7/9 (78%)	7 (100%)	0	100	100
All	All	1337/1369 (98%)	1248 (93%)	89 (7%)	19	38

5 of 89 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	48	LEU
1	E	140	LEU
1	H	48	LEU
1	E	62	LEU
1	E	96	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	D	118	GLN
1	D	127	GLN
1	E	122	GLN
1	G	10	ASN

### 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 [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	190/196 (96%)	0.45	11 (5%) 24 18	27, 67, 132, 150	0
1	B	196/196 (100%)	0.74	18 (9%) 10 6	32, 73, 115, 192	0
1	C	190/196 (96%)	-0.05	1 (0%) 90 89	13, 36, 70, 138	0
1	D	189/196 (96%)	-0.00	0 100 100	12, 36, 86, 104	0
1	E	190/196 (96%)	0.01	2 (1%) 80 77	25, 46, 87, 127	0
1	F	189/196 (96%)	0.09	1 (0%) 90 89	27, 49, 95, 116	0
1	G	190/196 (96%)	-0.00	5 (2%) 56 49	20, 41, 79, 110	0
1	H	190/196 (96%)	0.07	4 (2%) 64 58	21, 43, 85, 118	0
2	R	12/12 (100%)	-0.27	0 100 100	43, 59, 84, 88	0
2	T	12/12 (100%)	-0.30	0 100 100	41, 54, 80, 82	0
2	W	12/12 (100%)	-0.07	0 100 100	35, 55, 89, 94	0
2	Z	12/12 (100%)	-0.20	0 100 100	48, 58, 83, 84	0
3	K	8/10 (80%)	2.33	6 (75%) 0 0	83, 117, 127, 132	0
All	All	1580/1626 (97%)	0.16	48 (3%) 51 43	12, 49, 102, 192	0

The worst 5 of 48 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	3	ALA	13.4
1	B	5	LYS	5.4
1	B	7	ALA	5.2
1	B	8	LYS	5.2
1	B	4	GLU	5.0

### 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.