



# wwPDB X-ray Structure Validation Summary Report ⓘ

Nov 16, 2023 – 02:54 AM JST

PDB ID : 6KKM  
Title : Crystal structure of RbcL-Raf1 complex from Anabaena sp. PCC 7120  
Authors : Xia, L.Y.; Jiang, Y.L.; Kong, W.W.; Chen, Y.; Zhou, C.Z.  
Deposited on : 2019-07-26  
Resolution : 3.00 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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

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 : 2.36  
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) : 2.36

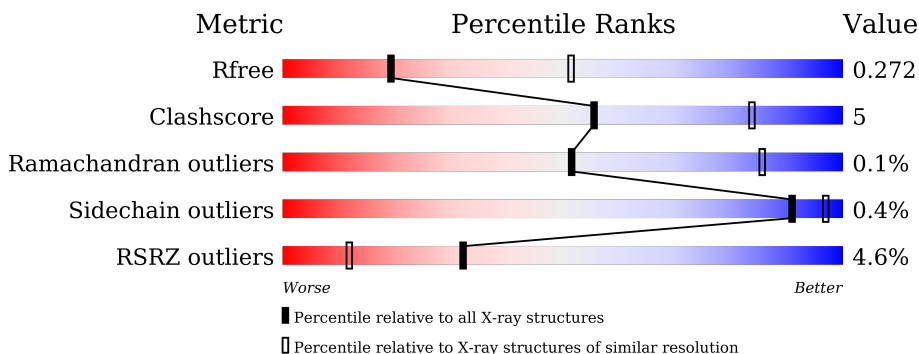
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.00 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.



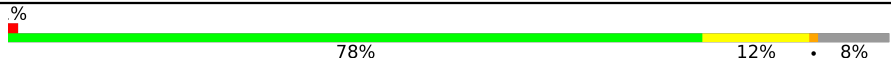

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	2092 (3.00-3.00)
Clashscore	141614	2416 (3.00-3.00)
Ramachandran outliers	138981	2333 (3.00-3.00)
Sidechain outliers	138945	2336 (3.00-3.00)
RSRZ outliers	127900	1990 (3.00-3.00)

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 of 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	E	361	
1	F	361	
1	G	361	
1	H	361	
2	A	491	
2	B	491	

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	C	491	 <p>% 78% 12% 8%</p>
2	D	491	 <p>% 81% 9% 9%</p>

## 2 Entry composition i

There are 2 unique types of molecules in this entry. The entry contains 24810 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 All5250 protein.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
1	E	334	2705	1731	473	501	0	0	0
1	F	333	2693	1723	472	498	0	0	0
1	G	329	2656	1700	461	495	0	0	0
1	H	332	2683	1716	468	499	0	0	0

- Molecule 2 is a protein called Ribulose biphosphate carboxylase large chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	A	451	3541	2255	620	652	14	0	0	0
2	B	445	3490	2219	616	642	13	0	0	0
2	C	452	3552	2261	624	653	14	0	0	0
2	D	445	3490	2219	616	642	13	0	0	0

There are 60 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-14	MET	-	expression tag	UNP P00879
A	-13	GLY	-	expression tag	UNP P00879
A	-12	HIS	-	expression tag	UNP P00879
A	-11	HIS	-	expression tag	UNP P00879
A	-10	HIS	-	expression tag	UNP P00879
A	-9	HIS	-	expression tag	UNP P00879
A	-8	HIS	-	expression tag	UNP P00879
A	-7	HIS	-	expression tag	UNP P00879

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	-6	HIS	-	expression tag	UNP P00879
A	-5	HIS	-	expression tag	UNP P00879
A	-4	HIS	-	expression tag	UNP P00879
A	-3	HIS	-	expression tag	UNP P00879
A	-2	SER	-	expression tag	UNP P00879
A	-1	SER	-	expression tag	UNP P00879
A	0	GLY	-	expression tag	UNP P00879
B	-14	MET	-	expression tag	UNP P00879
B	-13	GLY	-	expression tag	UNP P00879
B	-12	HIS	-	expression tag	UNP P00879
B	-11	HIS	-	expression tag	UNP P00879
B	-10	HIS	-	expression tag	UNP P00879
B	-9	HIS	-	expression tag	UNP P00879
B	-8	HIS	-	expression tag	UNP P00879
B	-7	HIS	-	expression tag	UNP P00879
B	-6	HIS	-	expression tag	UNP P00879
B	-5	HIS	-	expression tag	UNP P00879
B	-4	HIS	-	expression tag	UNP P00879
B	-3	HIS	-	expression tag	UNP P00879
B	-2	SER	-	expression tag	UNP P00879
B	-1	SER	-	expression tag	UNP P00879
B	0	GLY	-	expression tag	UNP P00879
C	-14	MET	-	expression tag	UNP P00879
C	-13	GLY	-	expression tag	UNP P00879
C	-12	HIS	-	expression tag	UNP P00879
C	-11	HIS	-	expression tag	UNP P00879
C	-10	HIS	-	expression tag	UNP P00879
C	-9	HIS	-	expression tag	UNP P00879
C	-8	HIS	-	expression tag	UNP P00879
C	-7	HIS	-	expression tag	UNP P00879
C	-6	HIS	-	expression tag	UNP P00879
C	-5	HIS	-	expression tag	UNP P00879
C	-4	HIS	-	expression tag	UNP P00879
C	-3	HIS	-	expression tag	UNP P00879
C	-2	SER	-	expression tag	UNP P00879
C	-1	SER	-	expression tag	UNP P00879
C	0	GLY	-	expression tag	UNP P00879
D	-14	MET	-	expression tag	UNP P00879
D	-13	GLY	-	expression tag	UNP P00879
D	-12	HIS	-	expression tag	UNP P00879
D	-11	HIS	-	expression tag	UNP P00879
D	-10	HIS	-	expression tag	UNP P00879

*Continued on next page...*

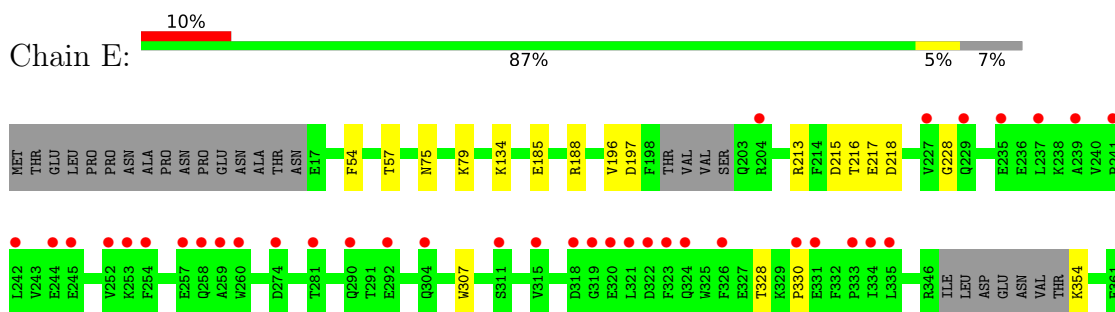
*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
D	-9	HIS	-	expression tag	UNP P00879
D	-8	HIS	-	expression tag	UNP P00879
D	-7	HIS	-	expression tag	UNP P00879
D	-6	HIS	-	expression tag	UNP P00879
D	-5	HIS	-	expression tag	UNP P00879
D	-4	HIS	-	expression tag	UNP P00879
D	-3	HIS	-	expression tag	UNP P00879
D	-2	SER	-	expression tag	UNP P00879
D	-1	SER	-	expression tag	UNP P00879
D	0	GLY	-	expression tag	UNP P00879

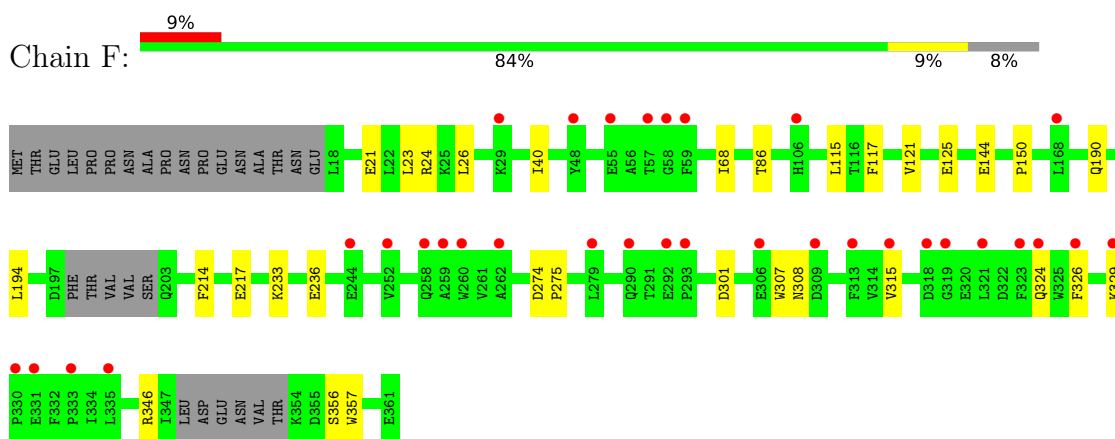
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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.

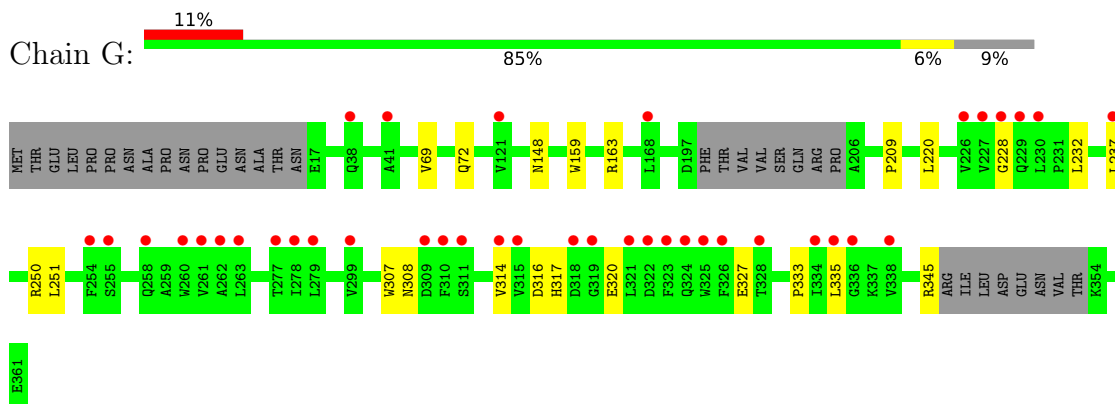
- Molecule 1: All5250 protein



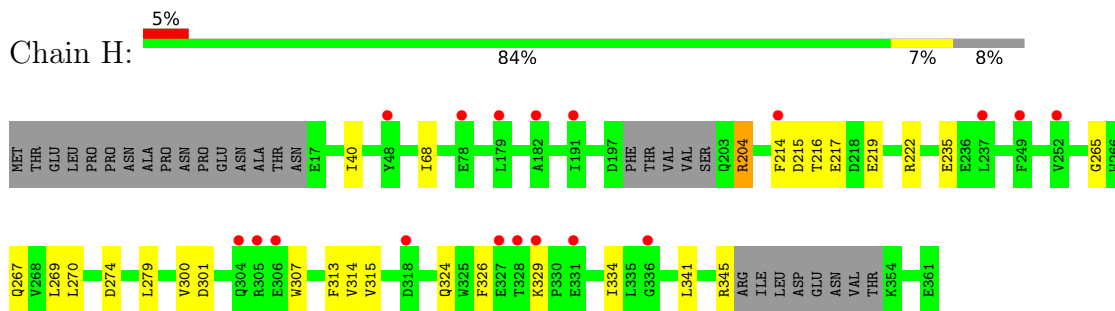
- Molecule 1: All5250 protein



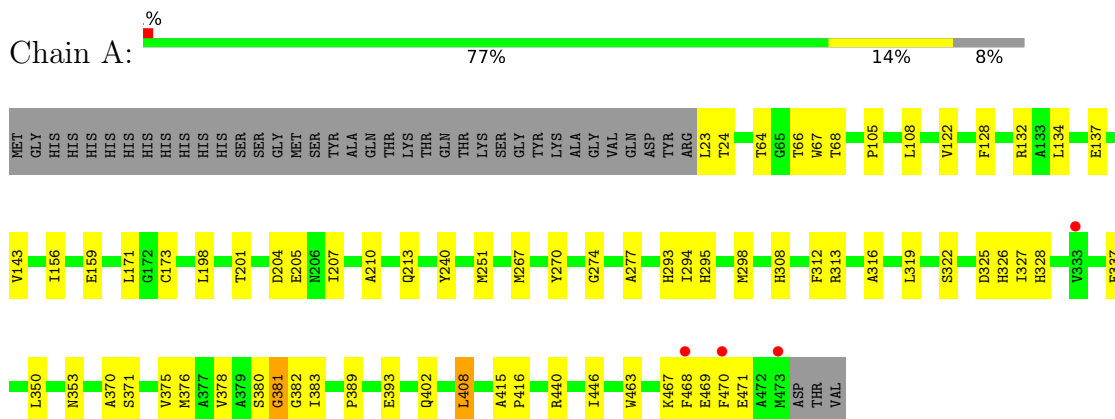
- Molecule 1: All5250 protein



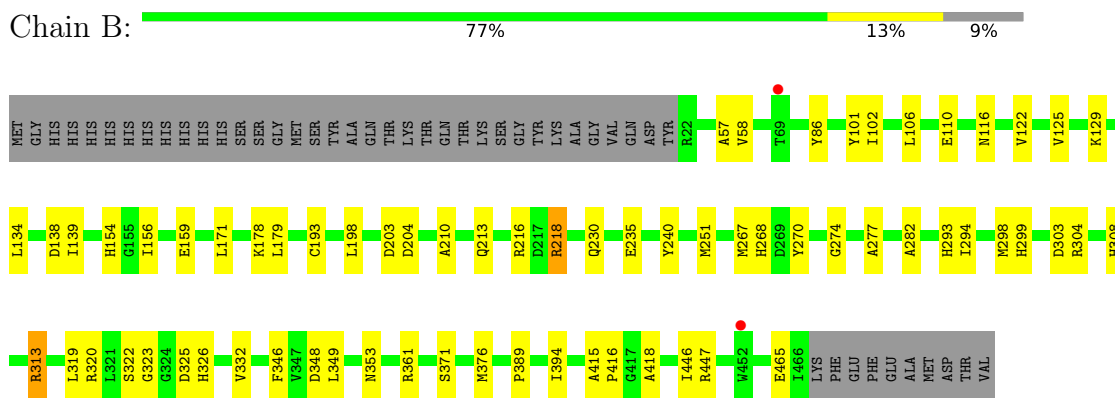
• Molecule 1: All5250 protein



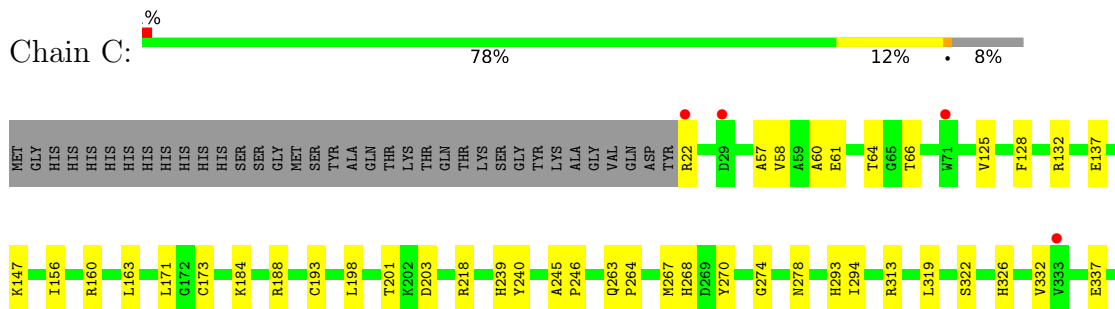
• Molecule 2: Ribulose biphosphate carboxylase large chain



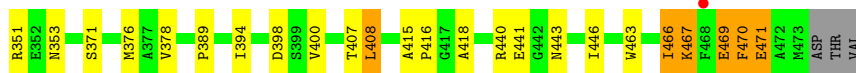
• Molecule 2: Ribulose biphosphate carboxylase large chain



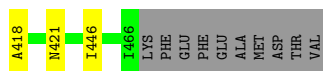
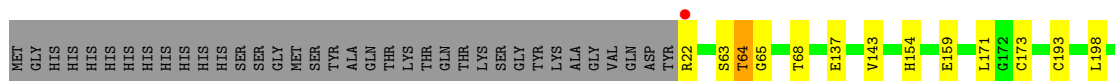
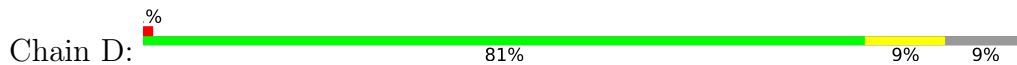
• Molecule 2: Ribulose biphosphate carboxylase large chain







● Molecule 2: Ribulose biphosphate carboxylase large chain



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	222.12Å 228.54Å 162.34Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.00 – 3.00 49.94 – 3.00	Depositor EDS
% Data completeness (in resolution range)	88.5 (49.00-3.00) 88.6 (49.94-3.00)	Depositor EDS
$R_{merge}$	0.20	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.30 (at 3.01Å)	Xtrriage
Refinement program	REFMAC 5.8.0158	Depositor
R, $R_{free}$	0.229 , 0.278 0.230 , 0.272	Depositor DCC
$R_{free}$ test set	3725 reflections (5.08%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	56.1	Xtrriage
Anisotropy	0.203	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 43.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	0.003 for $-1/2^*h-1/2^*k+1,-1/2^*h-1/2^*k-1,1/2^*h-1/2^*k$ 0.002 for $-1/2^*h-1/2^*k-1,-1/2^*h-1/2^*k+1,-1/2^*h+1/2^*k$ 0.000 for $-1/2^*h+1/2^*k-1,1/2^*h-1/2^*k-1,-1/2^*h-1/2^*k$ 0.000 for $-1/2^*h+1/2^*k+1,1/2^*h-1/2^*k+1,1/2^*h+1/2^*k$ 0.015 for $-k,-h,-l$	Xtrriage
$F_o, F_c$ correlation	0.90	EDS
Total number of atoms	24810	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	73.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.11% 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 [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	E	0.44	0/2771	0.69	0/3761
1	F	0.42	0/2758	0.67	0/3744
1	G	0.46	0/2720	0.67	0/3693
1	H	0.43	0/2748	0.67	0/3731
2	A	0.45	0/3628	0.76	4/4927 (0.1%)
2	B	0.43	0/3575	0.72	2/4857 (0.0%)
2	C	0.43	0/3639	0.75	4/4941 (0.1%)
2	D	0.45	1/3575 (0.0%)	0.74	3/4857 (0.1%)
All	All	0.44	1/25414 (0.0%)	0.71	13/34511 (0.0%)

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
2	B	0	1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	65	GLY	N-CA	5.82	1.54	1.46

The worst 5 of 13 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	D	64	THR	N-CA-C	-14.23	72.56	111.00
2	C	467	LYS	N-CA-C	8.54	134.07	111.00
2	C	470	PHE	N-CA-CB	7.46	124.03	110.60
2	D	63	SER	C-N-CA	6.32	137.49	121.70
2	A	469	GLU	N-CA-C	6.16	127.64	111.00

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	B	216	ARG	Sidechain

## 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	E	2705	0	2662	10	0
1	F	2693	0	2658	26	0
1	G	2656	0	2612	28	0
1	H	2683	0	2640	19	1
2	A	3541	0	3469	50	1
2	B	3490	0	3425	54	0
2	C	3552	0	3480	59	0
2	D	3490	0	3424	45	0
All	All	24810	0	24370	251	1

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 251 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:C:407:THR:CG2	2:D:22:ARG:HH22	1.13	1.59
2:C:407:THR:CG2	2:D:22:ARG:NH2	1.97	1.26
2:C:407:THR:HG21	2:D:22:ARG:NH2	1.52	1.24
1:G:307:TRP:CE2	1:G:327:GLU:O	1.96	1.18
2:C:407:THR:OG1	2:D:22:ARG:NH2	1.79	1.16

All (1) 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:H:235:GLU:OE2	2:A:440:ARG:NH2[8_575]	1.73	0.47

## 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	E	328/361 (91%)	321 (98%)	6 (2%)	1 (0%)	41	76
1	F	327/361 (91%)	320 (98%)	6 (2%)	1 (0%)	41	76
1	G	323/361 (90%)	317 (98%)	5 (2%)	1 (0%)	41	76
1	H	326/361 (90%)	316 (97%)	10 (3%)	0	100	100
2	A	449/491 (91%)	435 (97%)	14 (3%)	0	100	100
2	B	443/491 (90%)	432 (98%)	11 (2%)	0	100	100
2	C	450/491 (92%)	436 (97%)	13 (3%)	1 (0%)	47	82
2	D	443/491 (90%)	433 (98%)	10 (2%)	0	100	100
All	All	3089/3408 (91%)	3010 (97%)	75 (2%)	4 (0%)	51	85

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	C	470	PHE
1	F	275	PRO
1	E	228	GLY
1	G	228	GLY

### 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	E	286/311 (92%)	284 (99%)	2 (1%)	84	94

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	F	285/311 (92%)	284 (100%)	1 (0%)	91	97
1	G	281/311 (90%)	281 (100%)	0	100	100
1	H	284/311 (91%)	283 (100%)	1 (0%)	91	97
2	A	364/399 (91%)	364 (100%)	0	100	100
2	B	359/399 (90%)	359 (100%)	0	100	100
2	C	365/399 (92%)	361 (99%)	4 (1%)	73	90
2	D	359/399 (90%)	357 (99%)	2 (1%)	86	95
All	All	2583/2840 (91%)	2573 (100%)	10 (0%)	91	97

5 of 10 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	C	471	GLU
2	D	64	THR
2	D	193	CYS
1	H	204	ARG
2	C	193	CYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 54 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	278	ASN
2	C	116	ASN
2	D	278	ASN
2	B	299	HIS
2	B	402	GLN

### 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 monosaccharides 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	E	334/361 (92%)	0.40	37 (11%) 5 1	30, 68, 160, 191	0
1	F	333/361 (92%)	0.65	33 (9%) 7 2	55, 106, 140, 170	0
1	G	329/361 (91%)	0.75	39 (11%) 4 1	55, 100, 157, 187	0
1	H	332/361 (91%)	0.41	18 (5%) 25 9	59, 93, 141, 162	0
2	A	451/491 (91%)	-0.38	4 (0%) 84 63	19, 39, 79, 138	0
2	B	445/491 (90%)	-0.31	2 (0%) 92 79	22, 46, 95, 146	0
2	C	452/491 (92%)	-0.22	5 (1%) 80 56	29, 53, 108, 163	0
2	D	445/491 (90%)	-0.16	6 (1%) 77 51	31, 59, 103, 136	0
All	All	3121/3408 (91%)	0.08	144 (4%) 32 12	19, 65, 136, 191	0

The worst 5 of 144 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	331	GLU	6.7
1	E	318	ASP	6.3
1	F	258	GLN	5.6
1	G	258	GLN	5.4
1	F	331	GLU	4.9

### 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 monosaccharides in this entry.



## 6.4 Ligands

There are no ligands in this entry.

## 6.5 Other polymers

There are no such residues in this entry.