



# wwPDB X-ray Structure Validation Summary Report ⓘ

Sep 19, 2017 – 11:37 AM EDT

PDB ID : 5XAZ  
Title : Crystal structure of full length native tylp, a tetr regulator from streptomyces fradiae  
Authors : Ray, S.; Panjikar, S.; Anand, R.  
Deposited on : unknown  
Resolution : 2.30 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-20029824
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-20029824

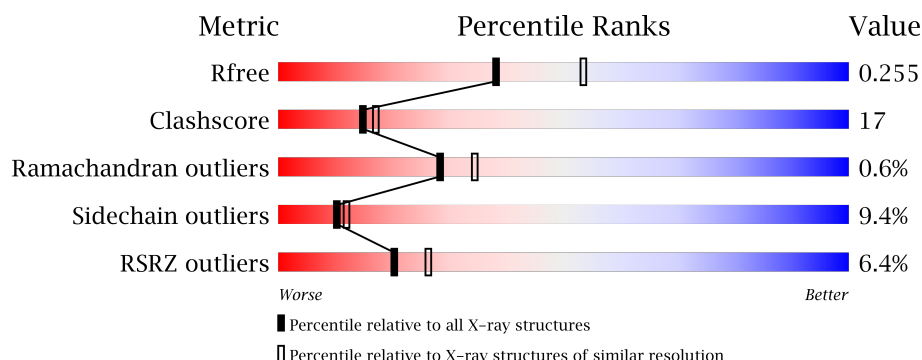
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.30 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	4130 (2.30-2.30)
Clashscore	112137	4751 (2.30-2.30)
Ramachandran outliers	110173	4705 (2.30-2.30)
Sidechain outliers	110143	4704 (2.30-2.30)
RSRZ outliers	101464	4156 (2.30-2.30)

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	228	<div> <div>7%</div> <div> <div></div> <div>71%</div> <div>20%</div> <div>6%</div> <div>..</div> </div> </div>
1	B	228	<div> <div>6%</div> <div> <div></div> <div>71%</div> <div>20%</div> <div>6%</div> <div>..</div> </div> </div>
1	C	228	<div> <div>7%</div> <div> <div></div> <div>67%</div> <div>22%</div> <div>7%</div> <div>.</div> </div> </div>
1	D	228	<div> <div>5%</div> <div> <div></div> <div>73%</div> <div>20%</div> <div>..</div> <div>..</div> </div> </div>
1	E	228	<div> <div>7%</div> <div> <div></div> <div>61%</div> <div>25%</div> <div>7%</div> <div>7%</div> </div> </div>

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	228	<div><div></div><div>5%</div><div>76%</div><div>16%</div><div>5%</div><div></div></div>
1	G	228	<div><div></div><div>7%</div><div>66%</div><div>19%</div><div>7%</div><div></div><div></div><div>7%</div></div>
1	H	228	<div><div></div><div>5%</div><div>69%</div><div>19%</div><div></div><div></div><div></div><div>7%</div></div>

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 13820 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 Gamma-butyrolactone receptor protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	223	Total	C	N	O	S	0	0	0
			1709	1067	318	319	5			
1	H	211	Total	C	N	O	S	0	0	0
			1620	1015	296	304	5			
1	B	223	Total	C	N	O	S	0	0	0
			1709	1067	318	319	5			
1	C	213	Total	C	N	O	S	0	0	0
			1634	1024	299	306	5			
1	D	220	Total	C	N	O	S	0	1	0
			1702	1062	318	316	6			
1	E	213	Total	C	N	O	S	0	1	0
			1643	1030	304	304	5			
1	F	222	Total	C	N	O	S	0	0	0
			1701	1063	317	316	5			
1	G	213	Total	C	N	O	S	0	1	0
			1643	1030	304	304	5			

There are 32 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	113	ALA	ARG	engineered mutation	UNP Q9XCC7
A	227	LEU	-	expression tag	UNP Q9XCC7
A	228	GLU	-	expression tag	UNP Q9XCC7
A	229	HIS	-	expression tag	UNP Q9XCC7
H	113	ALA	ARG	engineered mutation	UNP Q9XCC7
H	227	LEU	-	expression tag	UNP Q9XCC7
H	228	GLU	-	expression tag	UNP Q9XCC7
H	229	HIS	-	expression tag	UNP Q9XCC7
B	113	ALA	ARG	engineered mutation	UNP Q9XCC7
B	227	LEU	-	expression tag	UNP Q9XCC7
B	228	GLU	-	expression tag	UNP Q9XCC7
B	229	HIS	-	expression tag	UNP Q9XCC7
C	113	ALA	ARG	engineered mutation	UNP Q9XCC7

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
C	227	LEU	-	expression tag	UNP Q9XCC7
C	228	GLU	-	expression tag	UNP Q9XCC7
C	229	HIS	-	expression tag	UNP Q9XCC7
D	113	ALA	ARG	engineered mutation	UNP Q9XCC7
D	227	LEU	-	expression tag	UNP Q9XCC7
D	228	GLU	-	expression tag	UNP Q9XCC7
D	229	HIS	-	expression tag	UNP Q9XCC7
E	113	ALA	ARG	engineered mutation	UNP Q9XCC7
E	227	LEU	-	expression tag	UNP Q9XCC7
E	228	GLU	-	expression tag	UNP Q9XCC7
E	229	HIS	-	expression tag	UNP Q9XCC7
F	113	ALA	ARG	engineered mutation	UNP Q9XCC7
F	227	LEU	-	expression tag	UNP Q9XCC7
F	228	GLU	-	expression tag	UNP Q9XCC7
F	229	HIS	-	expression tag	UNP Q9XCC7
G	113	ALA	ARG	engineered mutation	UNP Q9XCC7
G	227	LEU	-	expression tag	UNP Q9XCC7
G	228	GLU	-	expression tag	UNP Q9XCC7
G	229	HIS	-	expression tag	UNP Q9XCC7

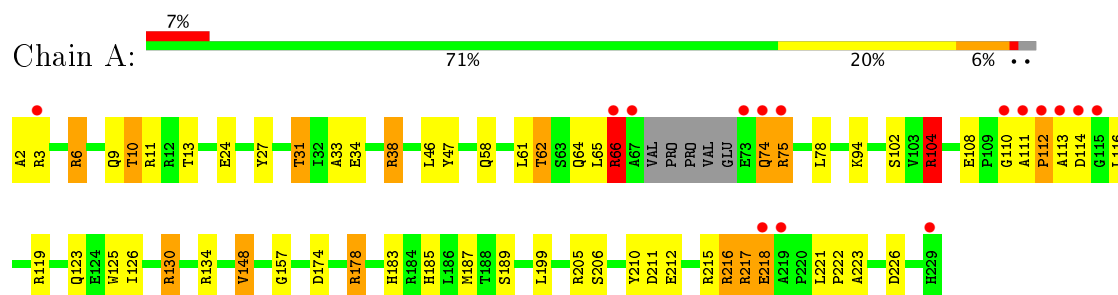
- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	58	Total O 58 58	0	0
2	H	58	Total O 58 58	0	0
2	B	68	Total O 68 68	0	0
2	C	57	Total O 57 57	0	0
2	D	62	Total O 62 62	0	0
2	E	45	Total O 45 45	0	0
2	F	62	Total O 62 62	0	0
2	G	49	Total O 49 49	0	0

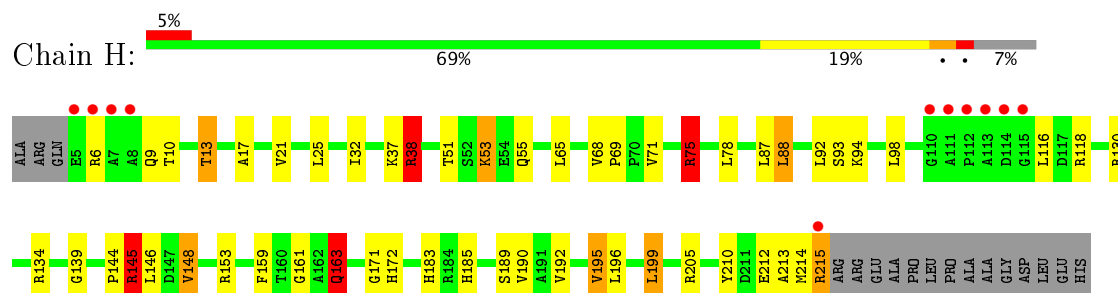
### 3 Residue-property plots [i](#)

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.

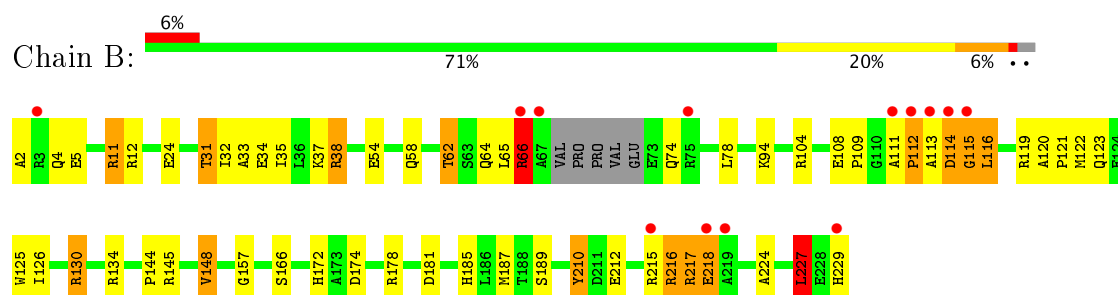
- Molecule 1: Gamma-butyrolactone receptor protein



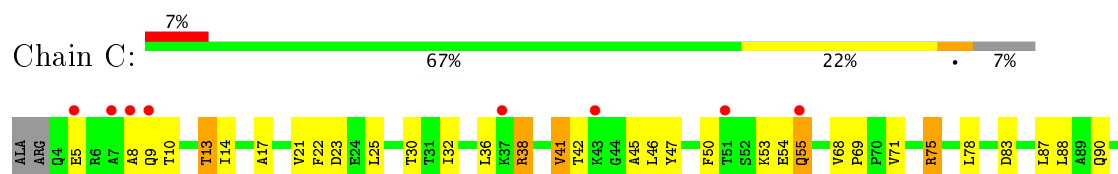
- Molecule 1: Gamma-butyrolactone receptor protein



- Molecule 1: Gamma-butyrolactone receptor protein

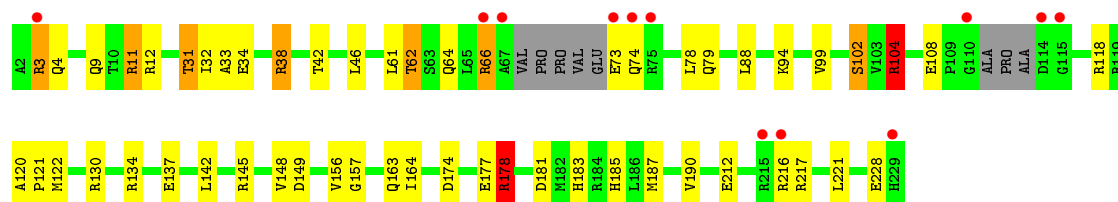


- Molecule 1: Gamma-butyrolactone receptor protein

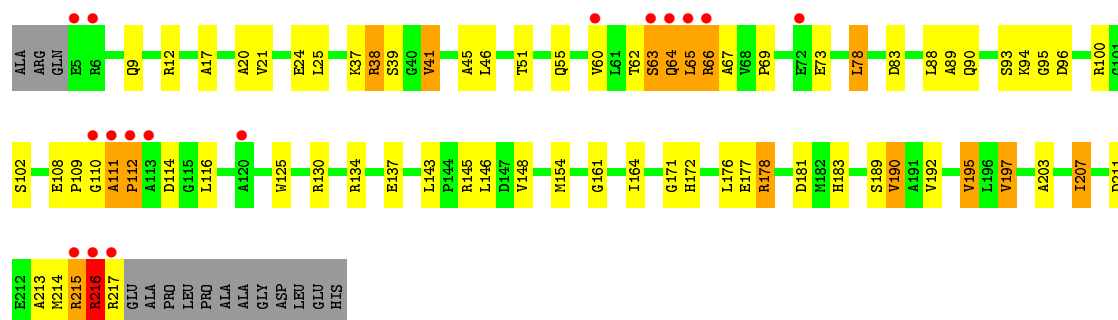




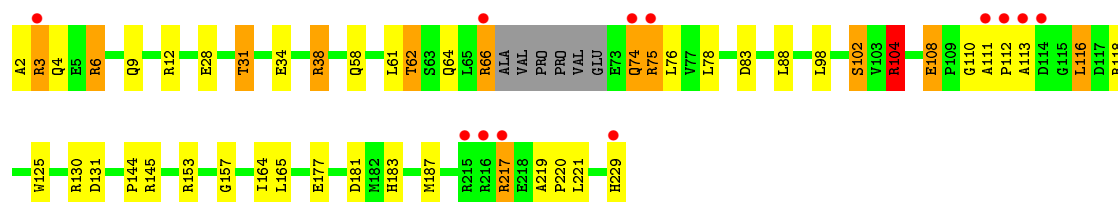
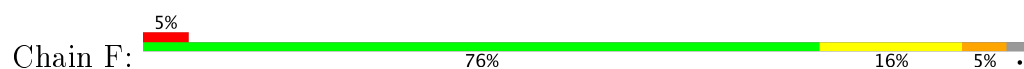
- Molecule 1: Gamma-butyrolactone receptor protein



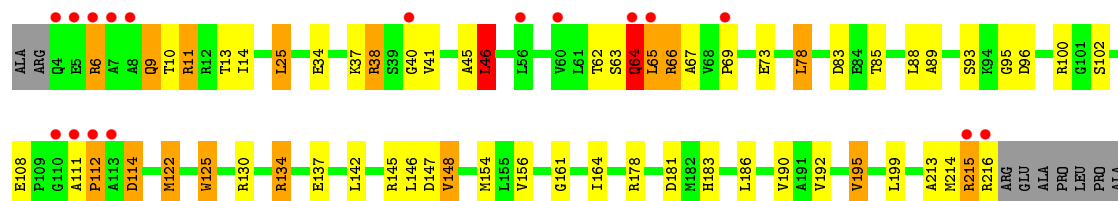
- Molecule 1: Gamma-butyrolactone receptor protein



- Molecule 1: Gamma-butyrolactone receptor protein



- Molecule 1: Gamma-butyrolactone receptor protein



ALA
GLY
ASP
LEU
GLU
HIS



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	101.48Å 71.57Å 158.58Å 90.00° 102.92° 90.00°	Depositor
Resolution (Å)	19.81 – 2.30 19.81 – 2.30	Depositor EDS
% Data completeness (in resolution range)	99.2 (19.81-2.30) 99.4 (19.81-2.30)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.08 (at 2.30Å)	Xtriage
Refinement program	REFMAC 5.6.0117	Depositor
R, $R_{free}$	0.204 , 0.253 0.208 , 0.255	Depositor DCC
$R_{free}$ test set	992 reflections (1.02%)	DCC
Wilson B-factor (Å <sup>2</sup> )	24.3	Xtriage
Anisotropy	0.097	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.38 , 42.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.46$ , $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	13820	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	32.0	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 65.10 % 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 7.4642e-06. 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	1.04	1/1731 (0.1%)	1.15	9/2343 (0.4%)
1	B	1.04	2/1731 (0.1%)	1.10	11/2343 (0.5%)
1	C	0.96	0/1656	1.04	7/2245 (0.3%)
1	D	0.96	1/1722 (0.1%)	1.11	8/2326 (0.3%)
1	E	0.99	1/1669 (0.1%)	1.09	10/2262 (0.4%)
1	F	0.98	3/1723 (0.2%)	1.08	10/2332 (0.4%)
1	G	1.00	1/1669 (0.1%)	1.11	12/2262 (0.5%)
1	H	0.89	0/1642	1.05	7/2227 (0.3%)
All	All	0.98	9/13543 (0.1%)	1.09	74/18340 (0.4%)

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	G	0	1

The worst 5 of 9 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	F	125	TRP	CD2-CE2	8.28	1.51	1.41
1	A	125	TRP	CD2-CE2	6.64	1.49	1.41
1	E	125	TRP	CD2-CE2	6.32	1.49	1.41
1	F	102	SER	CB-OG	-6.21	1.34	1.42
1	G	125	TRP	CD2-CE2	6.05	1.48	1.41

The worst 5 of 74 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	178	ARG	NE-CZ-NH2	-10.62	114.99	120.30
1	H	145	ARG	NE-CZ-NH1	-10.19	115.20	120.30
1	D	178	ARG	NE-CZ-NH1	9.82	125.21	120.30

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	F	153	ARG	NE-CZ-NH2	-9.44	115.58	120.30
1	F	6	ARG	NE-CZ-NH1	-9.10	115.75	120.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	G	40	GLY	Peptide

## 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	1709	0	1741	102	0
1	B	1709	0	1741	70	0
1	C	1634	0	1669	39	0
1	D	1702	0	1736	49	0
1	E	1643	0	1684	62	1
1	F	1701	0	1734	46	0
1	G	1643	0	1683	60	0
1	H	1620	0	1655	52	0
2	A	58	0	0	7	0
2	B	68	0	0	19	0
2	C	57	0	0	3	0
2	D	62	0	0	11	0
2	E	45	0	0	4	0
2	F	62	0	0	12	0
2	G	49	0	0	5	0
2	H	58	0	0	8	1
All	All	13820	0	13643	459	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 17.

The worst 5 of 459 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:H:51:THR:HG22	1:H:55:GLN:NE2	1.07	1.38
1:A:111:ALA:HB3	1:A:114:ASP:CB	1.51	1.38
1:B:35:ILE:HA	2:B:304:HOH:O	1.20	1.34
1:H:51:THR:CG2	1:H:55:GLN:NE2	1.91	1.34
1:H:51:THR:CG2	1:H:55:GLN:HE22	1.44	1.24

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:E:207:ILE:CG2	2:H:352:HOH:O[2_645]	1.93	0.27

## 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	219/228 (96%)	211 (96%)	6 (3%)	2 (1%)	20	23
1	B	219/228 (96%)	212 (97%)	4 (2%)	3 (1%)	13	13
1	C	211/228 (92%)	199 (94%)	11 (5%)	1 (0%)	32	39
1	D	215/228 (94%)	209 (97%)	6 (3%)	0	100	100
1	E	212/228 (93%)	199 (94%)	11 (5%)	2 (1%)	20	23
1	F	218/228 (96%)	206 (94%)	11 (5%)	1 (0%)	32	39
1	G	212/228 (93%)	202 (95%)	8 (4%)	2 (1%)	20	23
1	H	209/228 (92%)	202 (97%)	7 (3%)	0	100	100
All	All	1715/1824 (94%)	1640 (96%)	64 (4%)	11 (1%)	28	34

5 of 11 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	216	ARG

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	B	112	PRO
1	B	216	ARG
1	C	112	PRO
1	E	112	PRO

### 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	175/182 (96%)	160 (91%)	15 (9%)	12	14
1	B	175/182 (96%)	161 (92%)	14 (8%)	14	17
1	C	170/182 (93%)	153 (90%)	17 (10%)	9	10
1	D	175/182 (96%)	161 (92%)	14 (8%)	14	17
1	E	171/182 (94%)	151 (88%)	20 (12%)	6	7
1	F	174/182 (96%)	161 (92%)	13 (8%)	16	19
1	G	171/182 (94%)	153 (90%)	18 (10%)	8	9
1	H	169/182 (93%)	150 (89%)	19 (11%)	7	7
All	All	1380/1456 (95%)	1250 (91%)	130 (9%)	10	12

5 of 130 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	94	LYS
1	D	104	ARG
1	G	65	LEU
1	C	148	VAL
1	D	4	GLN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 21 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	185	HIS

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	C	183	HIS
1	F	185	HIS
1	B	64	GLN
1	F	229	HIS

### 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	A	223/228 (97%)	-0.10	15 (6%)	19	25	15, 24, 62, 138	0
1	B	223/228 (97%)	-0.14	13 (5%)	24	30	13, 24, 62, 129	0
1	C	213/228 (93%)	-0.02	15 (7%)	17	23	16, 31, 78, 130	0
1	D	220/228 (96%)	-0.23	12 (5%)	26	32	16, 24, 56, 96	0
1	E	213/228 (93%)	-0.00	16 (7%)	15	20	16, 29, 70, 100	0
1	F	222/228 (97%)	-0.14	12 (5%)	26	33	14, 24, 62, 134	0
1	G	213/228 (93%)	0.04	17 (7%)	13	17	16, 30, 69, 104	0
1	H	211/228 (92%)	-0.04	11 (5%)	28	35	16, 32, 71, 98	0
All	All	1738/1824 (95%)	-0.08	111 (6%)	20	26	13, 26, 69, 138	0

The worst 5 of 111 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	113	ALA	11.2
1	G	112	PRO	10.3
1	F	112	PRO	10.1
1	A	113	ALA	10.0
1	G	111	ALA	9.5

### 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 ⓘ

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.