



# wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 14, 2018 – 12:12 pm GMT

PDB ID : 3X0X  
Title : Crystal structure of apo-DszC from Rhodococcus erythropolis D-1  
Authors : Guan, L.J.; Lee, W.C.; Wang, S.P.; Ohtsuka, J.; Tanokura, M.  
Deposited on : 2014-10-23  
Resolution : 2.11 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	:	trunk31020
Percentile statistics	:	20171227.v01 (using entries in the PDB archive December 27th 2017)
Refmac	:	5.8.0158
CCP4	:	7.0 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk31020

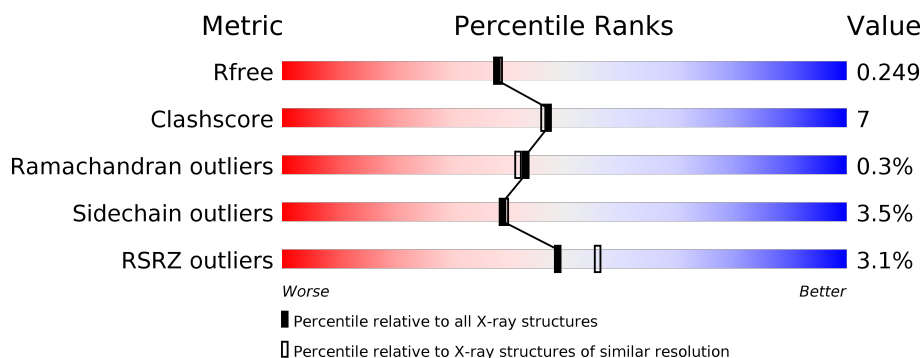
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.11 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))
$R_{free}$	111664	5449 (2.14-2.10)
Clashscore	122126	5972 (2.14-2.10)
Ramachandran outliers	120053	5914 (2.14-2.10)
Sidechain outliers	120020	5915 (2.14-2.10)
RSRZ outliers	108989	5321 (2.14-2.10)

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	417	<div> <div>84%</div> <div>12%</div> <div>•</div> </div>
1	B	417	<div> <div>82%</div> <div>13%</div> <div>• •</div> </div>
1	C	417	<div> <div>81%</div> <div>13%</div> <div>• •</div> </div>
1	D	417	<div> <div>82%</div> <div>13%</div> <div>• •</div> </div>
1	E	417	<div> <div>80%</div> <div>15%</div> <div>•</div> </div>
1	F	417	<div> <div>2%</div> <div>80%</div> <div>14%</div> <div>• •</div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	G	417	<div><div>%</div><div><div></div><div>80%</div><div>15%</div><div></div><div></div></div></div>
1	H	417	<div><div>17%</div><div><div></div><div>72%</div><div>20%</div><div></div><div></div></div></div>

## 2 Entry composition [i](#)

There are 2 unique types of molecules in this entry. The entry contains 25802 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 DszC.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	400	Total	C	N	O	S	0	0	0
			3054	1918	545	587	4			
1	B	400	Total	C	N	O	S	0	0	0
			3054	1918	545	587	4			
1	D	400	Total	C	N	O	S	0	0	0
			3054	1918	545	587	4			
1	C	400	Total	C	N	O	S	0	0	0
			3054	1918	545	587	4			
1	E	400	Total	C	N	O	S	0	0	0
			3054	1918	545	587	4			
1	F	400	Total	C	N	O	S	0	0	0
			3054	1918	545	587	4			
1	G	400	Total	C	N	O	S	0	0	0
			3054	1918	545	587	4			
1	H	400	Total	C	N	O	S	0	0	0
			3054	1918	545	587	4			

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	189	Total	O	0	0
			189	189		
2	B	223	Total	O	0	0
			223	223		
2	D	196	Total	O	0	0
			196	196		
2	C	196	Total	O	0	0
			196	196		
2	E	191	Total	O	0	0
			191	191		
2	F	153	Total	O	0	0
			153	153		

*Continued on next page...*

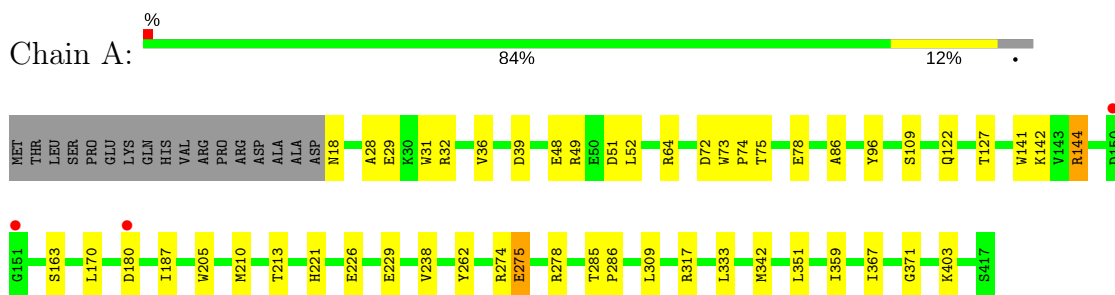
*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	G	146	Total 146	O 146	0	0
2	H	76	Total 76	O 76	0	0

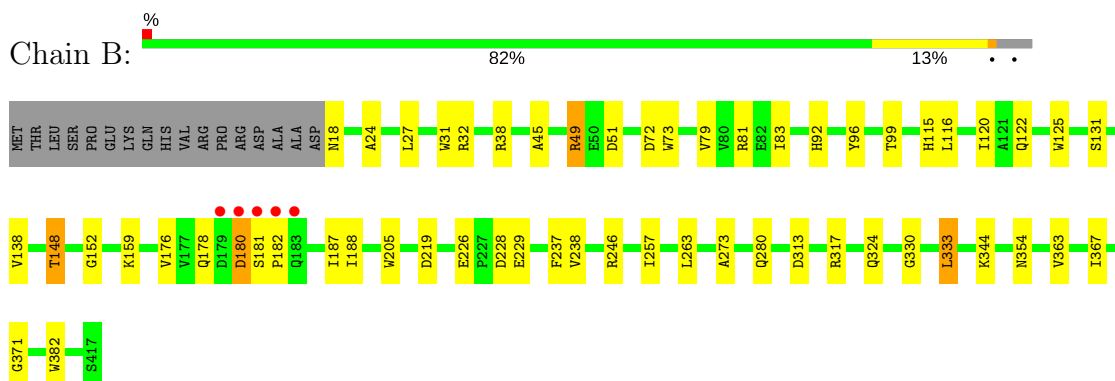
### 3 Residue-property plots [i](#)

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.

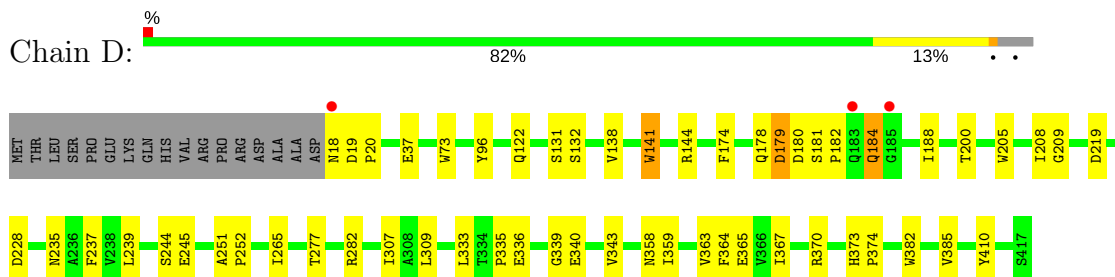
#### • Molecule 1: DszC



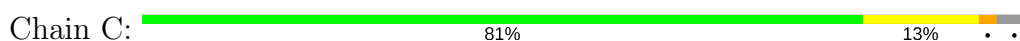
#### • Molecule 1: DszC

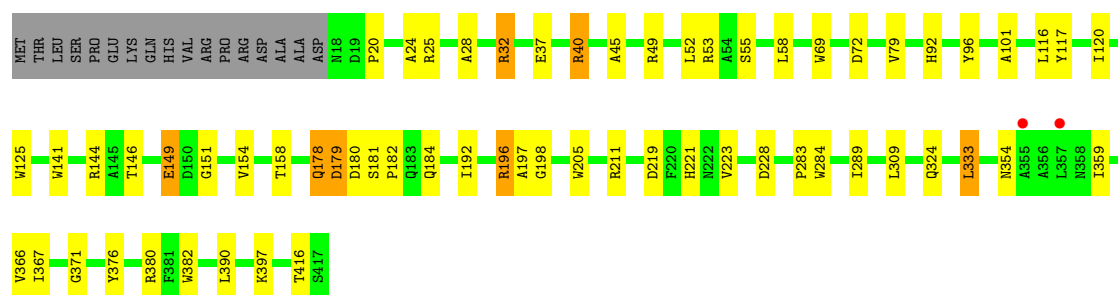


#### • Molecule 1: DszC



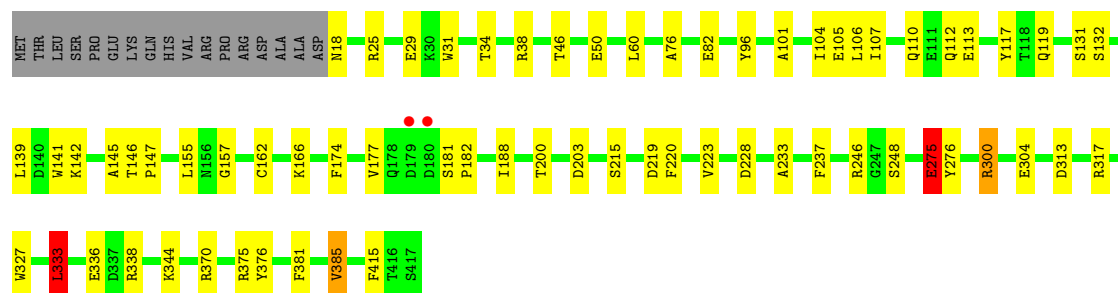
#### • Molecule 1: DszC





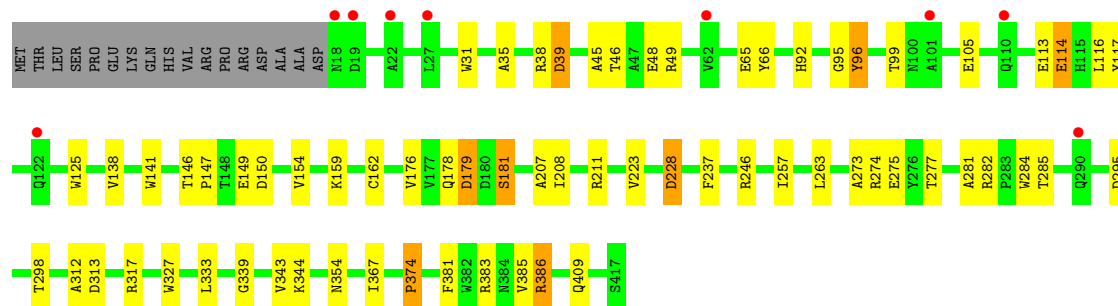
• Molecule 1: DszC

Chain E: 80% 15% .



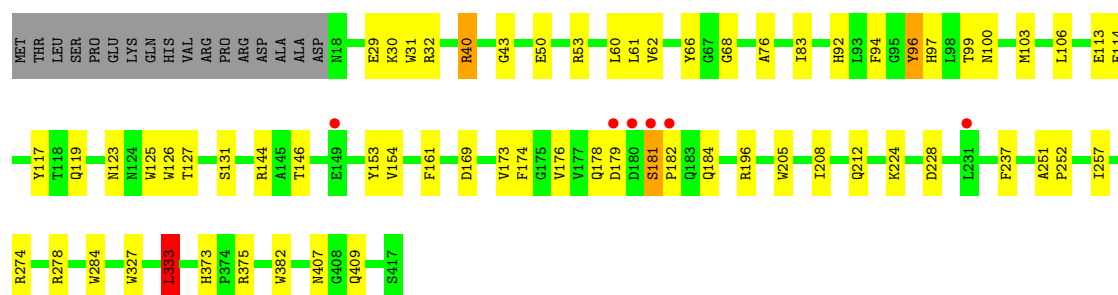
• Molecule 1: DszC

Chain F: 80% 14% . .

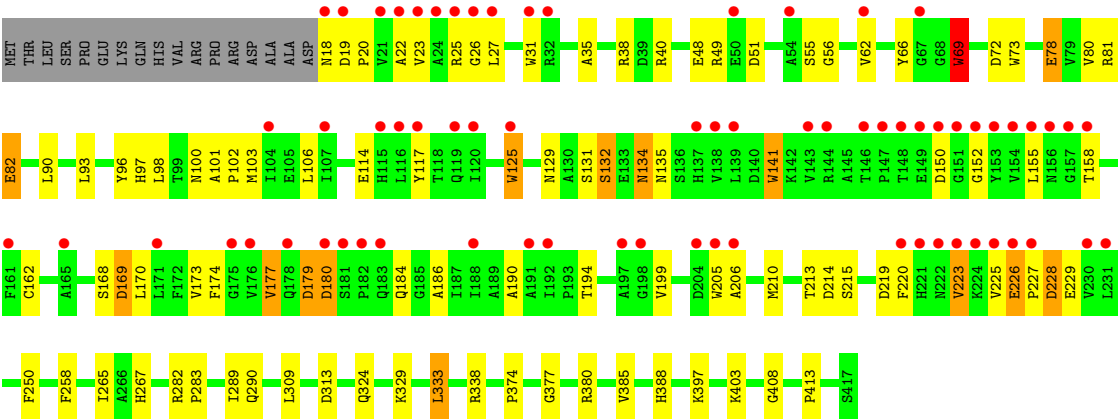


• Molecule 1: DszC

Chain G: 80% 15% . .



• Molecule 1: DszC





## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	146.13Å 123.25Å 184.27Å 90.00° 101.21° 90.00°	Depositor
Resolution (Å)	43.91 – 2.11 43.91 – 2.11	Depositor EDS
% Data completeness (in resolution range)	98.2 (43.91-2.11) 98.2 (43.91-2.11)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.37 (at 2.12Å)	Xtriage
Refinement program	REFMAC 5.6.0117	Depositor
R, $R_{free}$	0.186 , 0.253 0.184 , 0.249	Depositor DCC
$R_{free}$ test set	8996 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	30.1	Xtriage
Anisotropy	0.092	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 42.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.44$ , $\langle L^2 \rangle = 0.26$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	25802	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	40.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.99% 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	A	0.89	4/3129 (0.1%)	0.86	3/4265 (0.1%)
1	B	0.94	4/3129 (0.1%)	0.90	5/4265 (0.1%)
1	C	0.96	5/3129 (0.2%)	0.90	4/4265 (0.1%)
1	D	0.92	5/3129 (0.2%)	0.87	0/4265
1	E	0.91	4/3129 (0.1%)	0.88	3/4265 (0.1%)
1	F	0.87	4/3129 (0.1%)	0.85	1/4265 (0.0%)
1	G	0.85	6/3129 (0.2%)	0.84	1/4265 (0.0%)
1	H	0.78	4/3129 (0.1%)	0.77	2/4265 (0.0%)
All	All	0.89	36/25032 (0.1%)	0.86	19/34120 (0.1%)

The worst 5 of 36 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	E	275	GLU	CD-OE2	7.89	1.34	1.25
1	F	141	TRP	CD2-CE2	6.69	1.49	1.41
1	B	125	TRP	CD2-CE2	6.33	1.49	1.41
1	C	382	TRP	CD2-CE2	6.17	1.48	1.41
1	B	205	TRP	CD2-CE2	6.14	1.48	1.41

The worst 5 of 19 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	49	ARG	NE-CZ-NH2	-9.19	115.71	120.30
1	F	386	ARG	NE-CZ-NH1	6.80	123.70	120.30
1	H	333	LEU	CA-CB-CG	6.42	130.05	115.30
1	E	300	ARG	NE-CZ-NH1	6.22	123.41	120.30
1	E	385	VAL	CB-CA-C	-6.20	99.63	111.40

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	A	3054	0	2932	26	0
1	B	3054	0	2932	36	0
1	C	3054	0	2932	31	0
1	D	3054	0	2932	36	0
1	E	3054	0	2932	40	0
1	F	3054	0	2932	48	0
1	G	3054	0	2932	36	0
1	H	3054	0	2932	76	0
2	A	189	0	0	10	0
2	B	223	0	0	8	0
2	C	196	0	0	3	0
2	D	196	0	0	7	0
2	E	191	0	0	7	0
2	F	153	0	0	3	0
2	G	146	0	0	1	0
2	H	76	0	0	9	0
All	All	25802	0	23456	319	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 7.

The worst 5 of 319 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:31:TRP:O	1:H:38:ARG:NH2	1.87	1.08
1:D:282:ARG:HD2	1:D:370:ARG:NH1	1.73	1.02
1:G:97:HIS:ND1	1:G:127:THR:HG22	1.75	1.01
1:E:370:ARG:HD3	1:G:161:PHE:CZ	2.06	0.91
1:B:181:SER:HB2	1:B:182:PRO:HD2	1.52	0.89

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	398/417 (95%)	381 (96%)	17 (4%)	0	100	100
1	B	398/417 (95%)	388 (98%)	10 (2%)	0	100	100
1	C	398/417 (95%)	384 (96%)	13 (3%)	1 (0%)	43	42
1	D	398/417 (95%)	388 (98%)	9 (2%)	1 (0%)	43	42
1	E	398/417 (95%)	385 (97%)	12 (3%)	1 (0%)	43	42
1	F	398/417 (95%)	383 (96%)	12 (3%)	3 (1%)	21	15
1	G	398/417 (95%)	381 (96%)	16 (4%)	1 (0%)	43	42
1	H	398/417 (95%)	365 (92%)	29 (7%)	4 (1%)	17	11
All	All	3184/3336 (95%)	3055 (96%)	118 (4%)	11 (0%)	43	42

5 of 11 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	179	ASP
1	G	181	SER
1	H	131	SER
1	D	179	ASP
1	H	179	ASP

### 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	308/323 (95%)	299 (97%)	9 (3%)	45	47

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	B	308/323 (95%)	301 (98%)	7 (2%)	53	57
1	C	308/323 (95%)	298 (97%)	10 (3%)	42	43
1	D	308/323 (95%)	301 (98%)	7 (2%)	53	57
1	E	308/323 (95%)	298 (97%)	10 (3%)	42	43
1	F	308/323 (95%)	296 (96%)	12 (4%)	35	35
1	G	308/323 (95%)	294 (96%)	14 (4%)	30	28
1	H	308/323 (95%)	290 (94%)	18 (6%)	22	18
All	All	2464/2584 (95%)	2377 (96%)	87 (4%)	39	39

5 of 87 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	228	ASP
1	F	150	ASP
1	H	219	ASP
1	E	275	GLU
1	F	46	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 9 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	178	GLN
1	H	184	GLN
1	G	137	HIS
1	D	184	GLN
1	E	280	GLN

### 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 ⓘ

### 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	400/417 (95%)	-0.20	3 (0%) 86 88	17, 35, 56, 84	0
1	B	400/417 (95%)	-0.20	5 (1%) 77 80	16, 30, 54, 108	0
1	C	400/417 (95%)	-0.01	2 (0%) 90 92	17, 34, 57, 104	0
1	D	400/417 (95%)	-0.10	3 (0%) 86 88	16, 33, 56, 95	0
1	E	400/417 (95%)	-0.06	2 (0%) 90 92	17, 34, 60, 105	0
1	F	400/417 (95%)	0.08	9 (2%) 60 65	19, 39, 67, 109	0
1	G	400/417 (95%)	-0.04	6 (1%) 73 77	20, 40, 63, 101	0
1	H	400/417 (95%)	0.74	69 (17%) 1 1	20, 57, 94, 125	0
All	All	3200/3336 (95%)	0.03	99 (3%) 49 56	16, 36, 72, 125	0

The worst 5 of 99 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	H	151	GLY	6.2
1	H	148	THR	5.4
1	H	147	PRO	5.3
1	H	150	ASP	5.2
1	H	175	GLY	5.2

### 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.