



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 31, 2016 – 10:16 PM GMT

PDB ID : 1SS8  
Title : GroEL  
Authors : Chaudhry, C.; Horwich, A.L.; Brunger, A.T.; Adams, P.D.  
Deposited on : 2004-03-23  
Resolution : 2.70 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  
Mogul : 1.7 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026688  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
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) : trunk26865

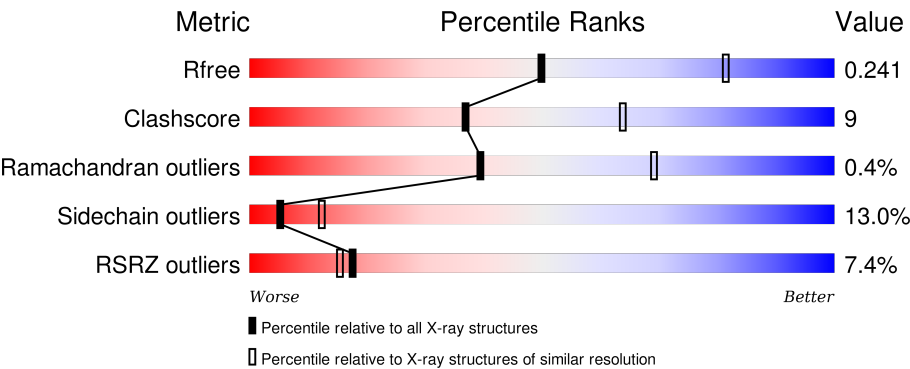
# 1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.70 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 <sub>free</sub>	91344	2103 (2.70-2.70)
Clashscore	102246	2422 (2.70-2.70)
Ramachandran outliers	100387	2382 (2.70-2.70)
Sidechain outliers	100360	2382 (2.70-2.70)
RSRZ outliers	91569	2107 (2.70-2.70)

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	524	<div><div>6%</div><div>71%23%5%</div></div>
1	B	524	<div><div>9%</div><div>71%22%5%</div></div>
1	C	524	<div><div>6%</div><div>73%21%</div></div>
1	D	524	<div><div>6%</div><div>71%23%5%</div></div>
1	E	524	<div><div>4%</div><div>72%22%6%</div></div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	524	<div><div></div><div>14%</div><div></div><div>74%</div><div></div><div>22%</div><div></div><div>• •</div></div>
1	G	524	<div><div></div><div>6%</div><div></div><div>72%</div><div></div><div>23%</div><div></div><div>• •</div></div>

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 27064 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 groEL protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	524	Total	C	N	O	S	0	0	0
			3851	2395	662	774	20			
1	B	524	Total	C	N	O	S	0	0	0
			3851	2395	662	774	20			
1	C	524	Total	C	N	O	S	0	0	0
			3851	2395	662	774	20			
1	D	524	Total	C	N	O	S	0	0	0
			3851	2395	662	774	20			
1	E	524	Total	C	N	O	S	0	0	0
			3851	2395	662	774	20			
1	F	524	Total	C	N	O	S	0	0	0
			3851	2395	662	774	20			
1	G	524	Total	C	N	O	S	0	0	0
			3851	2395	662	774	20			

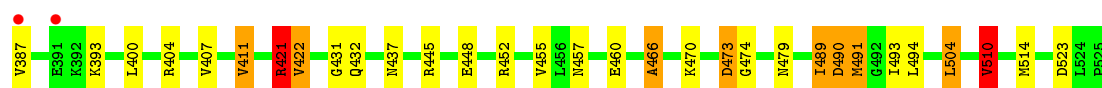
There are 14 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	13	GLY	ARG	SEE REMARK 999	UNP P0A6F5
A	126	VAL	ALA	SEE REMARK 999	UNP P0A6F5
B	13	GLY	ARG	SEE REMARK 999	UNP P0A6F5
B	126	VAL	ALA	SEE REMARK 999	UNP P0A6F5
C	13	GLY	ARG	SEE REMARK 999	UNP P0A6F5
C	126	VAL	ALA	SEE REMARK 999	UNP P0A6F5
D	13	GLY	ARG	SEE REMARK 999	UNP P0A6F5
D	126	VAL	ALA	SEE REMARK 999	UNP P0A6F5
E	13	GLY	ARG	SEE REMARK 999	UNP P0A6F5
E	126	VAL	ALA	SEE REMARK 999	UNP P0A6F5
F	13	GLY	ARG	SEE REMARK 999	UNP P0A6F5
F	126	VAL	ALA	SEE REMARK 999	UNP P0A6F5
G	13	GLY	ARG	SEE REMARK 999	UNP P0A6F5
G	126	VAL	ALA	SEE REMARK 999	UNP P0A6F5

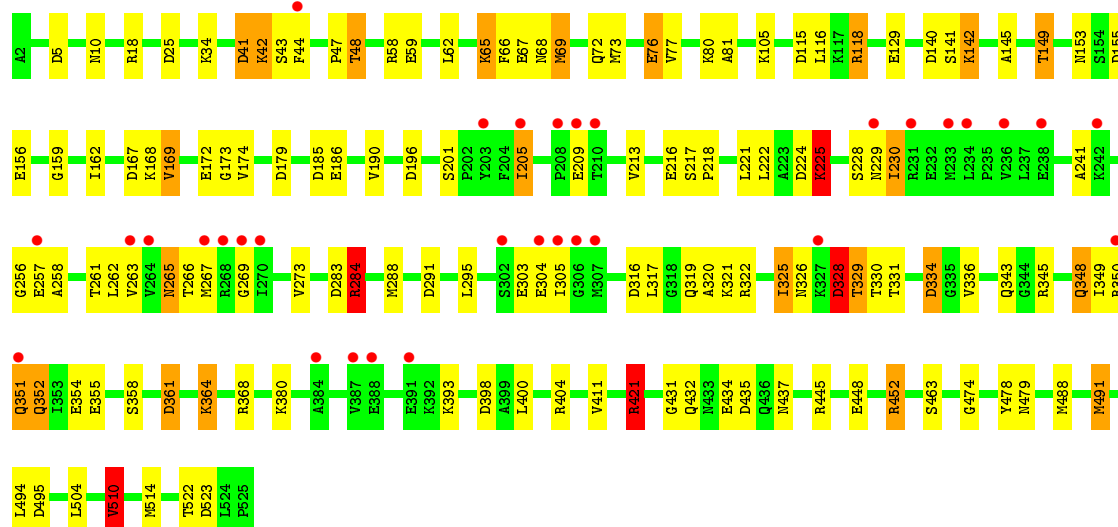
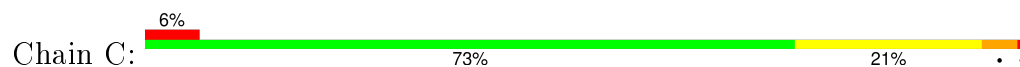
- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	18	Total 18	O 18	0	0
2	B	21	Total 21	O 21	0	0
2	C	15	Total 15	O 15	0	0
2	D	16	Total 16	O 16	0	0
2	E	17	Total 17	O 17	0	0
2	F	7	Total 7	O 7	0	0
2	G	13	Total 13	O 13	0	0

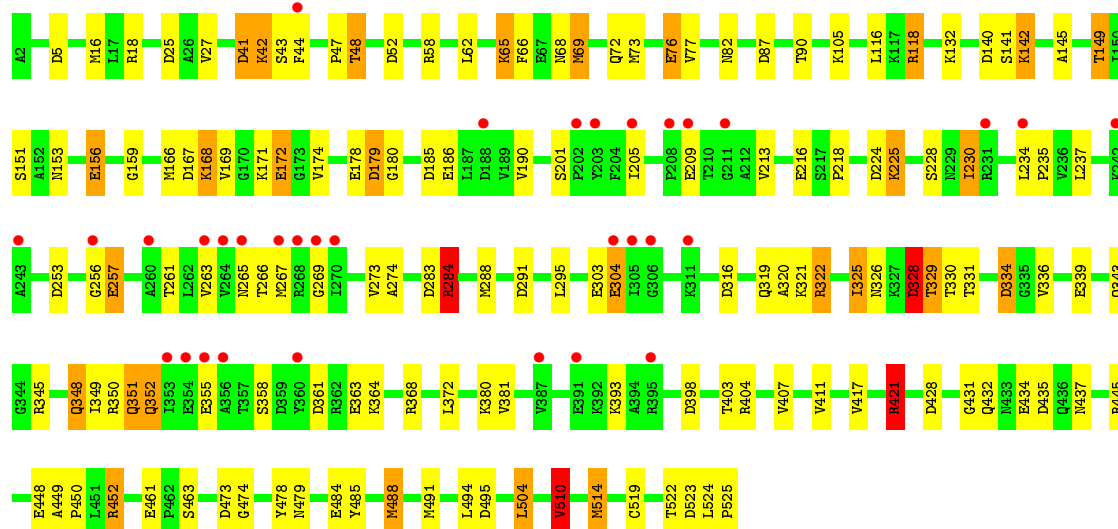




• Molecule 1: groEL protein

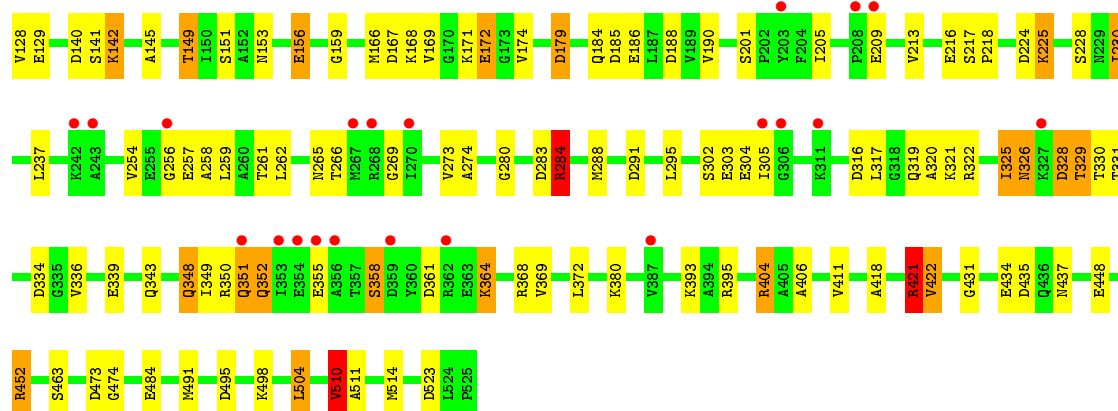


• Molecule 1: groEL protein

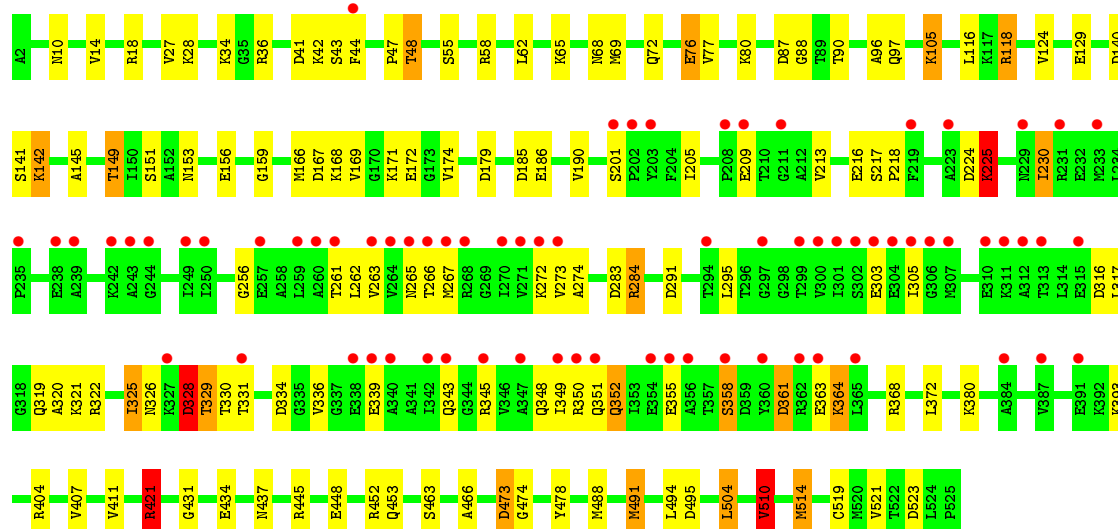
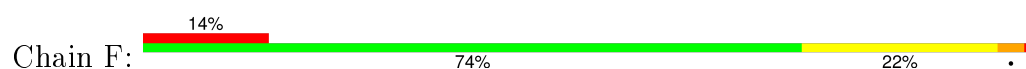


• Molecule 1: groEL protein

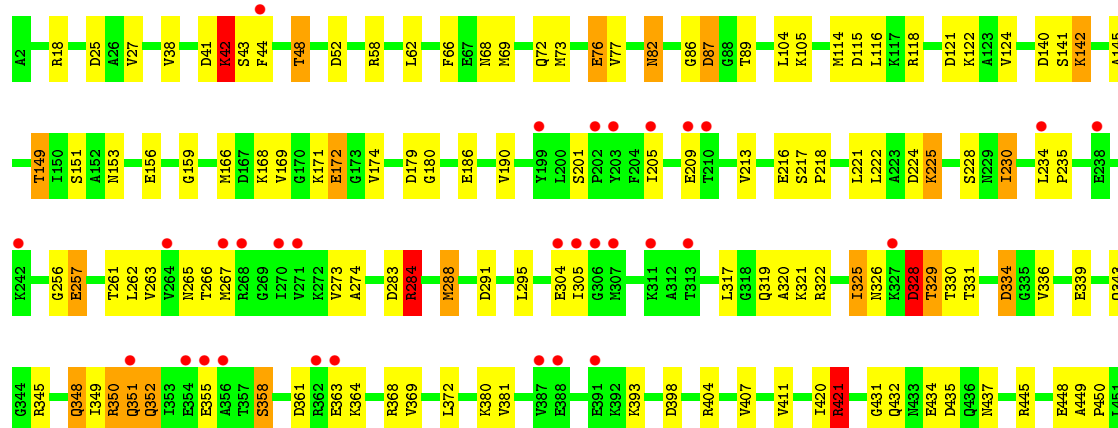
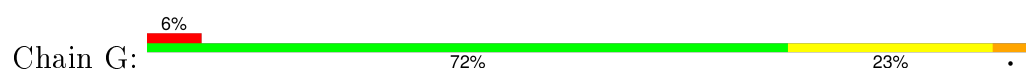




• Molecule 1: groEL protein



• Molecule 1: groEL protein



R452	S463	D473 G474	M479	E484	M491	L494 D495	S502 A503 L504	V510	M514	E518	T522 D523 L524 P525
------	------	--------------	------	------	------	--------------	----------------------	------	------	------	------------------------------

## 4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	178.38Å 204.98Å 280.98Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	30.00 – 2.70 29.96 – 2.70	Depositor EDS
% Data completeness (in resolution range)	(Not available) (30.00-2.70) 83.9 (29.96-2.70)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	0.59 (at 2.68Å)	Xtriage
Refinement program	REFMAC 5	Depositor
R, $R_{free}$	0.215 , 0.249 0.209 , 0.241	Depositor DCC
$R_{free}$ test set	2966 reflections (2.58%)	DCC
Wilson B-factor (Å <sup>2</sup> )	38.5	Xtriage
Anisotropy	0.354	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 45.3	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtriage
Outliers	0 of 123506 reflections	Xtriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	27064	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	2.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.19% 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, 0.375 respectively for untwinned datasets, and 0.333, 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	A	1.24	28/3879 (0.7%)	1.11	22/5237 (0.4%)
1	B	1.28	33/3879 (0.9%)	1.15	27/5237 (0.5%)
1	C	1.19	30/3879 (0.8%)	1.13	26/5237 (0.5%)
1	D	1.19	27/3879 (0.7%)	1.14	31/5237 (0.6%)
1	E	1.34	35/3879 (0.9%)	1.21	30/5237 (0.6%)
1	F	1.09	22/3879 (0.6%)	1.08	18/5237 (0.3%)
1	G	1.21	21/3879 (0.5%)	1.11	19/5237 (0.4%)
All	All	1.22	196/27153 (0.7%)	1.13	173/36659 (0.5%)

The worst 5 of 196 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	E	129	GLU	CD-OE2	15.67	1.42	1.25
1	E	129	GLU	CD-OE1	15.50	1.42	1.25
1	E	156	GLU	CD-OE2	11.11	1.37	1.25
1	E	76	GLU	CD-OE1	11.03	1.37	1.25
1	B	473	ASP	CB-CG	11.01	1.74	1.51

The worst 5 of 173 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	421	ARG	NE-CZ-NH1	12.62	126.61	120.30
1	E	421	ARG	NE-CZ-NH2	-12.43	114.08	120.30
1	E	452	ARG	NE-CZ-NH1	12.14	126.37	120.30
1	B	491	MET	CG-SD-CE	-11.48	81.84	100.20
1	E	421	ARG	NE-CZ-NH1	11.13	125.86	120.30

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	3851	0	3970	77	0
1	B	3851	0	3970	83	0
1	C	3851	0	3970	77	1
1	D	3851	0	3970	77	0
1	E	3851	0	3970	77	0
1	F	3851	0	3970	69	0
1	G	3851	0	3970	76	1
2	A	18	0	0	1	0
2	B	21	0	0	4	0
2	C	15	0	0	0	0
2	D	16	0	0	0	0
2	E	17	0	0	0	0
2	F	7	0	0	1	0
2	G	13	0	0	1	0
All	All	27064	0	27790	517	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 9.

The worst 5 of 517 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:B:489:ILE:CD1	1:B:489:ILE:CG1	1.85	1.52
1:B:473:ASP:CG	1:B:473:ASP:CB	1.74	1.51
1:A:288:MET:CE	1:A:288:MET:SD	2.03	1.47
1:G:114:MET:SD	1:G:114:MET:CE	2.01	1.46
1:C:288:MET:CE	1:C:288:MET:SD	2.02	1.46

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:C:354:GLU:OE1	1:G:350:ARG:NH1[5_455]	2.05	0.15

## 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	522/524 (100%)	507 (97%)	12 (2%)	3 (1%)	30	59
1	B	522/524 (100%)	505 (97%)	15 (3%)	2 (0%)	39	69
1	C	522/524 (100%)	506 (97%)	13 (2%)	3 (1%)	30	59
1	D	522/524 (100%)	507 (97%)	13 (2%)	2 (0%)	39	69
1	E	522/524 (100%)	504 (97%)	17 (3%)	1 (0%)	52	80
1	F	522/524 (100%)	506 (97%)	14 (3%)	2 (0%)	39	69
1	G	522/524 (100%)	504 (97%)	15 (3%)	3 (1%)	30	59
All	All	3654/3668 (100%)	3539 (97%)	99 (3%)	16 (0%)	39	69

5 of 16 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	256	GLY
1	B	256	GLY
1	C	256	GLY
1	D	256	GLY
1	E	256	GLY

### 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	404/404 (100%)	351 (87%)	53 (13%)	5	12
1	B	404/404 (100%)	347 (86%)	57 (14%)	4	10

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	C	404/404 (100%)	356 (88%)	48 (12%)	6	15
1	D	404/404 (100%)	350 (87%)	54 (13%)	5	11
1	E	404/404 (100%)	354 (88%)	50 (12%)	6	13
1	F	404/404 (100%)	350 (87%)	54 (13%)	5	11
1	G	404/404 (100%)	351 (87%)	53 (13%)	5	12
All	All	2828/2828 (100%)	2459 (87%)	369 (13%)	5	12

5 of 369 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	132	LYS
1	D	380	LYS
1	G	265	ASN
1	D	153	ASN
1	D	303	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 34 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	68	ASN
1	E	68	ASN
1	G	351	GLN
1	D	351	GLN
1	B	265	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 ⓘ

There are no carbohydrates in this entry.

## 5.6 Ligand geometry

There are no ligands in this entry.

## 5.7 Other polymers

There are no such residues in this entry.

## 5.8 Polymer linkage issues

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	524/524 (100%)	0.10	34 (6%) 22 20	2, 2, 4, 6	0
1	B	524/524 (100%)	0.20	47 (8%) 12 9	2, 2, 4, 6	0
1	C	524/524 (100%)	0.12	32 (6%) 25 23	2, 2, 4, 6	0
1	D	524/524 (100%)	0.11	33 (6%) 23 22	2, 2, 4, 6	0
1	E	524/524 (100%)	-0.11	22 (4%) 40 39	2, 2, 4, 6	0
1	F	524/524 (100%)	0.45	73 (13%) 4 3	2, 2, 4, 6	0
1	G	524/524 (100%)	0.04	31 (5%) 26 24	2, 2, 4, 6	0
All	All	3668/3668 (100%)	0.13	272 (7%) 17 15	2, 2, 4, 6	0

The worst 5 of 272 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	F	270	ILE	8.1
1	C	203	TYR	8.1
1	A	305	ILE	7.4
1	G	268	ARG	7.0
1	G	44	PHE	6.7

### 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

There are no ligands in this entry.

## 6.5 Other polymers

There are no such residues in this entry.