



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 14, 2017 – 01:34 am GMT

PDB ID : 5IIP  
Title : Staphylococcus aureus OpuCA  
Authors : Tosi, T.; Campeotto, I.; Freemont, P.S.; Grundling, A.  
Deposited on : 2016-03-01  
Resolution : 2.50 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

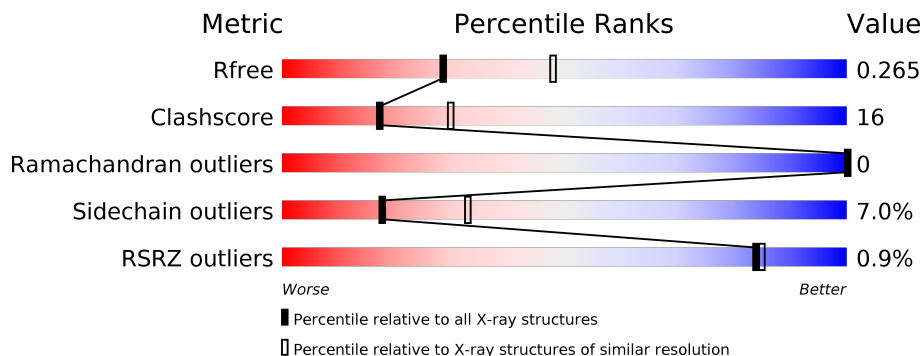
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.50 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	3846 (2.50-2.50)
Clashscore	112137	4554 (2.50-2.50)
Ramachandran outliers	110173	4463 (2.50-2.50)
Sidechain outliers	110143	4465 (2.50-2.50)
RSRZ outliers	101464	3876 (2.50-2.50)

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	194	<div> <div>%</div> <div> <div></div> <div>49%</div> <div>8%</div> <div>•</div> <div>40%</div> </div> </div>
1	B	194	<div> <div></div> <div>40%</div> <div>18%</div> <div>•</div> <div>41%</div> </div>
1	C	194	<div> <div>%</div> <div> <div></div> <div>37%</div> <div>22%</div> <div>•</div> <div>40%</div> </div> </div>
1	D	194	<div> <div>%</div> <div> <div></div> <div>45%</div> <div>14%</div> <div>•</div> <div>40%</div> </div> </div>

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 3732 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 Glycine betaine/carnitine/choline ABC transporter%2C ATP-binding protein%2C putative.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	116	Total	C	N	O	S	97	0	0
			929	580	172	174	3			
1	B	115	Total	C	N	O	S	70	0	0
			917	570	168	176	3			
1	C	116	Total	C	N	O	S	78	0	0
			929	580	171	175	3			
1	D	117	Total	C	N	O	S	68	0	0
			942	590	174	175	3			

There are 88 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	215	GLY	-	expression tag	UNP A0A0D6GYR3
A	216	SER	-	expression tag	UNP A0A0D6GYR3
A	217	SER	-	expression tag	UNP A0A0D6GYR3
A	218	HIS	-	expression tag	UNP A0A0D6GYR3
A	219	HIS	-	expression tag	UNP A0A0D6GYR3
A	220	HIS	-	expression tag	UNP A0A0D6GYR3
A	221	HIS	-	expression tag	UNP A0A0D6GYR3
A	222	HIS	-	expression tag	UNP A0A0D6GYR3
A	223	HIS	-	expression tag	UNP A0A0D6GYR3
A	224	SER	-	expression tag	UNP A0A0D6GYR3
A	225	SER	-	expression tag	UNP A0A0D6GYR3
A	226	GLY	-	expression tag	UNP A0A0D6GYR3
A	227	LEU	-	expression tag	UNP A0A0D6GYR3
A	228	VAL	-	expression tag	UNP A0A0D6GYR3
A	229	PRO	-	expression tag	UNP A0A0D6GYR3
A	230	ARG	-	expression tag	UNP A0A0D6GYR3
A	231	GLY	-	expression tag	UNP A0A0D6GYR3
A	232	SER	-	expression tag	UNP A0A0D6GYR3
A	233	HIS	-	expression tag	UNP A0A0D6GYR3
A	234	MET	-	expression tag	UNP A0A0D6GYR3

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	235	ALA	-	expression tag	UNP A0A0D6GYR3
A	236	SER	-	expression tag	UNP A0A0D6GYR3
B	215	GLY	-	expression tag	UNP A0A0D6GYR3
B	216	SER	-	expression tag	UNP A0A0D6GYR3
B	217	SER	-	expression tag	UNP A0A0D6GYR3
B	218	HIS	-	expression tag	UNP A0A0D6GYR3
B	219	HIS	-	expression tag	UNP A0A0D6GYR3
B	220	HIS	-	expression tag	UNP A0A0D6GYR3
B	221	HIS	-	expression tag	UNP A0A0D6GYR3
B	222	HIS	-	expression tag	UNP A0A0D6GYR3
B	223	HIS	-	expression tag	UNP A0A0D6GYR3
B	224	SER	-	expression tag	UNP A0A0D6GYR3
B	225	SER	-	expression tag	UNP A0A0D6GYR3
B	226	GLY	-	expression tag	UNP A0A0D6GYR3
B	227	LEU	-	expression tag	UNP A0A0D6GYR3
B	228	VAL	-	expression tag	UNP A0A0D6GYR3
B	229	PRO	-	expression tag	UNP A0A0D6GYR3
B	230	ARG	-	expression tag	UNP A0A0D6GYR3
B	231	GLY	-	expression tag	UNP A0A0D6GYR3
B	232	SER	-	expression tag	UNP A0A0D6GYR3
B	233	HIS	-	expression tag	UNP A0A0D6GYR3
B	234	MET	-	expression tag	UNP A0A0D6GYR3
B	235	ALA	-	expression tag	UNP A0A0D6GYR3
B	236	SER	-	expression tag	UNP A0A0D6GYR3
C	215	GLY	-	expression tag	UNP A0A0D6GYR3
C	216	SER	-	expression tag	UNP A0A0D6GYR3
C	217	SER	-	expression tag	UNP A0A0D6GYR3
C	218	HIS	-	expression tag	UNP A0A0D6GYR3
C	219	HIS	-	expression tag	UNP A0A0D6GYR3
C	220	HIS	-	expression tag	UNP A0A0D6GYR3
C	221	HIS	-	expression tag	UNP A0A0D6GYR3
C	222	HIS	-	expression tag	UNP A0A0D6GYR3
C	223	HIS	-	expression tag	UNP A0A0D6GYR3
C	224	SER	-	expression tag	UNP A0A0D6GYR3
C	225	SER	-	expression tag	UNP A0A0D6GYR3
C	226	GLY	-	expression tag	UNP A0A0D6GYR3
C	227	LEU	-	expression tag	UNP A0A0D6GYR3
C	228	VAL	-	expression tag	UNP A0A0D6GYR3
C	229	PRO	-	expression tag	UNP A0A0D6GYR3
C	230	ARG	-	expression tag	UNP A0A0D6GYR3
C	231	GLY	-	expression tag	UNP A0A0D6GYR3
C	232	SER	-	expression tag	UNP A0A0D6GYR3

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
C	233	HIS	-	expression tag	UNP A0A0D6GYR3
C	234	MET	-	expression tag	UNP A0A0D6GYR3
C	235	ALA	-	expression tag	UNP A0A0D6GYR3
C	236	SER	-	expression tag	UNP A0A0D6GYR3
D	215	GLY	-	expression tag	UNP A0A0D6GYR3
D	216	SER	-	expression tag	UNP A0A0D6GYR3
D	217	SER	-	expression tag	UNP A0A0D6GYR3
D	218	HIS	-	expression tag	UNP A0A0D6GYR3
D	219	HIS	-	expression tag	UNP A0A0D6GYR3
D	220	HIS	-	expression tag	UNP A0A0D6GYR3
D	221	HIS	-	expression tag	UNP A0A0D6GYR3
D	222	HIS	-	expression tag	UNP A0A0D6GYR3
D	223	HIS	-	expression tag	UNP A0A0D6GYR3
D	224	SER	-	expression tag	UNP A0A0D6GYR3
D	225	SER	-	expression tag	UNP A0A0D6GYR3
D	226	GLY	-	expression tag	UNP A0A0D6GYR3
D	227	LEU	-	expression tag	UNP A0A0D6GYR3
D	228	VAL	-	expression tag	UNP A0A0D6GYR3
D	229	PRO	-	expression tag	UNP A0A0D6GYR3
D	230	ARG	-	expression tag	UNP A0A0D6GYR3
D	231	GLY	-	expression tag	UNP A0A0D6GYR3
D	232	SER	-	expression tag	UNP A0A0D6GYR3
D	233	HIS	-	expression tag	UNP A0A0D6GYR3
D	234	MET	-	expression tag	UNP A0A0D6GYR3
D	235	ALA	-	expression tag	UNP A0A0D6GYR3
D	236	SER	-	expression tag	UNP A0A0D6GYR3

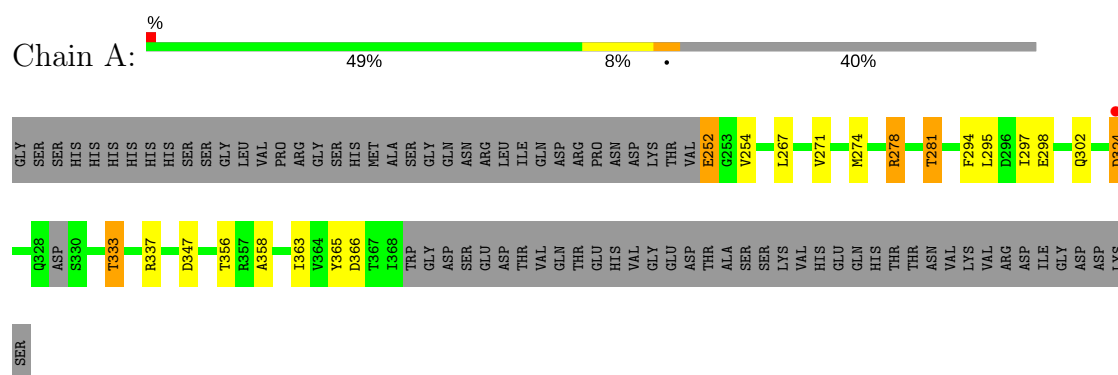
- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	2	Total O 2 2	0	0
2	B	3	Total O 3 3	0	0
2	C	3	Total O 3 3	0	0
2	D	7	Total O 7 7	0	0

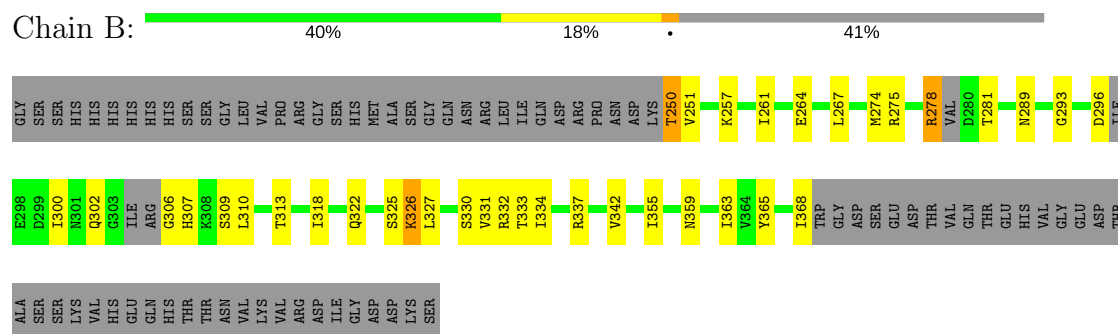
### 3 Residue-property plots

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.

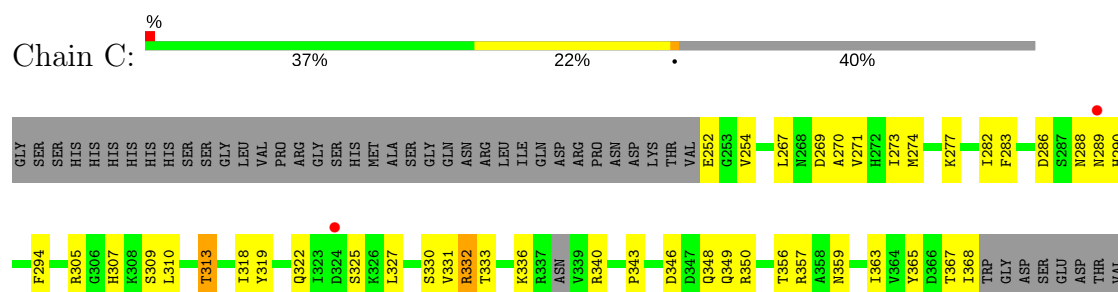
- Molecule 1: Glycine betaine/carnitine/choline ABC transporter%2C ATP-binding protein%2C putative



- Molecule 1: Glycine betaine/carnitine/choline ABC transporter%2C ATP-binding protein%2C putative



- Molecule 1: Glycine betaine/carnitine/choline ABC transporter%2C ATP-binding protein%2C putative



GLN	THR	GLU	HIS	VAL	GLY	GLU	ASP	THR	ALA	SER	SER	LYS	VAL	HIS	GLU	GLN	HIS	THR	THR	ASN	VAL	LYS	VAL	ARG	ASP	ILE	GLY	ASP	ASP	LYS	SER
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

● Molecule 1: Glycine betaine/carnitine/choline ABC transporter%2C ATP-binding protein%2C putative



GLY	SER	HIS	HIS	HIS	HIS	HIS	SER	SER	GLY	LEU	VAL	PRO	ARG	GLY	SER	HIS	MET	ALA	SER	GLY	GLN	ASN	ARG	LEU	ILE	GLN	ASP	ARG	PRO	ASN	ASP	LYS	THR	VAL	GLU	G253		L259		L267	N268	D269		L273		M274	R275	Q276	K277	R278	T281		N289	D286		D299
Q302	R305		D324	S325	K326	L327	Q328	V339	V342		D346		Q349	R350		I355		I363	V364	Y365	D366	T367	I368	W369	GLY	ASP	SER	GLU	ASP	THR	VAL	GLN	THR	GLU	GLU	HIS	VAL	GLY	GLU	ASP	THR	ALA	SER	SER	LYS	VAL	HIS	GLU	GLN	HIS	THR	THR	ASN	VAL	LYS	VAL

ARG	ASP	ILE	GLY	ASP	LYS	SER
-----	-----	-----	-----	-----	-----	-----

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	52.15Å 88.26Å 61.79Å 90.00° 114.79° 90.00°	Depositor
Resolution (Å)	47.34 – 2.50 47.34 – 2.50	Depositor EDS
% Data completeness (in resolution range)	99.8 (47.34-2.50) 99.9 (47.34-2.50)	Depositor EDS
$R_{merge}$	0.14	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.77 (at 2.51Å)	Xtriage
Refinement program	PHENIX	Depositor
R, $R_{free}$	0.244 , 0.280 0.233 , 0.265	Depositor DCC
$R_{free}$ test set	880 reflections (4.99%)	DCC
Wilson B-factor (Å <sup>2</sup> )	50.5	Xtriage
Anisotropy	0.530	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 63.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	0.038 for h,-k,-h-l	Xtriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	3732	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	50.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.98% 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	1.24	0/938	0.89	1/1270 (0.1%)
1	B	1.21	0/923	0.82	0/1246
1	C	1.22	0/937	0.88	1/1267 (0.1%)
1	D	1.23	0/953	0.88	0/1292
All	All	1.23	0/3751	0.87	2/5075 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	332	ARG	NE-CZ-NH1	6.97	123.78	120.30
1	A	324	ASP	CB-CG-OD1	5.70	123.43	118.30

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	929	0	958	21	11
1	B	917	0	934	35	8
1	C	929	0	957	33	18
1	D	942	0	966	19	1
2	A	2	0	0	0	0
2	B	3	0	0	0	0
2	C	3	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	D	7	0	0	0	0
All	All	3732	0	3815	105	19

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 16.

The worst 5 of 105 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:281:THR:HG21	1:A:294:PHE:CE1	1.36	1.55
1:A:281:THR:CG2	1:A:294:PHE:HE1	1.27	1.48
1:C:254:VAL:HG13	1:C:363:ILE:CD1	1.48	1.42
1:A:281:THR:CG2	1:A:294:PHE:CE1	2.05	1.30
1:C:254:VAL:HG13	1:C:363:ILE:HD12	1.32	1.11

The worst 5 of 19 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:A:365:TYR:CE1	1:C:332:ARG:NH2[1_655]	0.49	1.71
1:A:365:TYR:CG	1:C:332:ARG:NH1[1_655]	0.69	1.51
1:B:275:ARG:NH2	1:C:368:ILE:O[2_456]	0.71	1.49
1:B:275:ARG:NH2	1:C:368:ILE:C[2_456]	0.72	1.48
1:B:275:ARG:CZ	1:C:368:ILE:O[2_456]	0.85	1.35

## 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	A	114/194 (59%)	110 (96%)	4 (4%)	0	100	100
1	B	107/194 (55%)	104 (97%)	3 (3%)	0	100	100

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	C	112/194 (58%)	109 (97%)	3 (3%)	0	100	100
1	D	115/194 (59%)	115 (100%)	0	0	100	100
All	All	448/776 (58%)	438 (98%)	10 (2%)	0	100	100

There are no Ramachandran outliers to report.

### 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	107/176 (61%)	99 (92%)	8 (8%)	16	29
1	B	106/176 (60%)	99 (93%)	7 (7%)	19	36
1	C	107/176 (61%)	101 (94%)	6 (6%)	25	45
1	D	108/176 (61%)	99 (92%)	9 (8%)	13	25
All	All	428/704 (61%)	398 (93%)	30 (7%)	18	33

5 of 30 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	307	HIS
1	C	282	ILE
1	D	328	GLN
1	C	252	GLU
1	C	294	PHE

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

Mol	Chain	Res	Type
1	A	290	HIS
1	A	315	GLN
1	C	322	GLN
1	D	301	ASN

### 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 [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	115/194 (59%)	-0.00	1 (0%) 84 85	27, 45, 66, 73	23 (20%)
1	B	115/194 (59%)	0.02	0 100 100	33, 48, 68, 74	17 (14%)
1	C	116/194 (59%)	-0.04	2 (1%) 70 72	34, 52, 67, 78	20 (17%)
1	D	117/194 (60%)	0.14	1 (0%) 84 85	35, 49, 67, 80	16 (13%)
All	All	463/776 (59%)	0.03	4 (0%) 84 85	27, 49, 67, 80	76 (16%)

All (4) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	299	ASP	2.6
1	C	289	ASN	2.3
1	C	324	ASP	2.3
1	A	324	ASP	2.1

### 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.