



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 8, 2018 – 11:05 pm GMT

PDB ID : 5Y9H
Title : Crystal structure of SafDAA-dsc complex
Authors : Zeng, L.H.; Zhang, L.; Wang, P.R.; Meng, G.
Deposited on : 2017-08-25
Resolution : 2.80 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

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 : trunk30967
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) : trunk30967

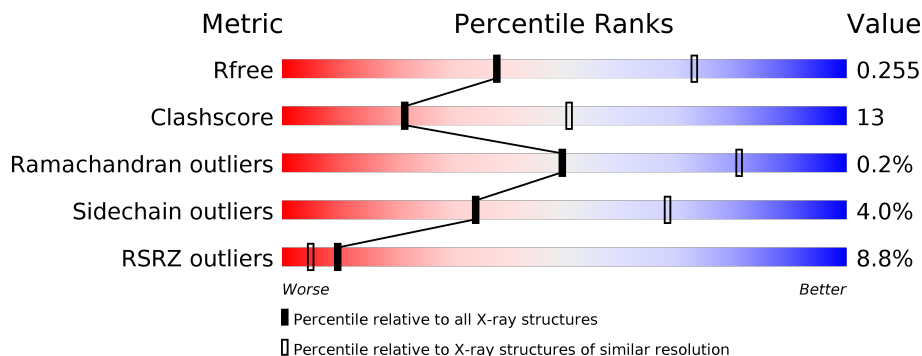
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.80 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	2792 (2.80-2.80)
Clashscore	122126	3209 (2.80-2.80)
Ramachandran outliers	120053	3158 (2.80-2.80)
Sidechain outliers	120020	3160 (2.80-2.80)
RSRZ outliers	108989	2726 (2.80-2.80)

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 ≥ 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	466	<div> <div>2%</div> <div> <div></div> <div>72%</div> <div>19%</div> <div>•</div> <div>8%</div> </div> </div>
1	B	466	<div> <div>17%</div> <div> <div></div> <div>61%</div> <div>28%</div> <div>•</div> <div>9%</div> </div> </div>
1	C	466	<div> <div>5%</div> <div> <div></div> <div>66%</div> <div>22%</div> <div>•</div> <div>10%</div> </div> </div>

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 18834 atoms, of which 9244 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 SafD,Putative outer membrane protein,Putative outer membrane protein,Putative outer membrane protein.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	430	Total	C	H	N	O	S	0	0	0
			6313	1999	3099	563	647	5			
1	B	424	Total	C	H	N	O	S	0	0	0
			6235	1969	3070	555	636	5			
1	C	421	Total	C	H	N	O	S	0	0	0
			6236	1965	3075	555	637	4			

There are 99 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MET	-	initiating methionine	UNP H9L4M7
A	2	GLY	-	expression tag	UNP H9L4M7
A	3	SER	-	expression tag	UNP H9L4M7
A	4	SER	-	expression tag	UNP H9L4M7
A	5	HIS	-	expression tag	UNP H9L4M7
A	6	HIS	-	expression tag	UNP H9L4M7
A	7	HIS	-	expression tag	UNP H9L4M7
A	8	HIS	-	expression tag	UNP H9L4M7
A	9	HIS	-	expression tag	UNP H9L4M7
A	10	HIS	-	expression tag	UNP H9L4M7
A	11	SER	-	expression tag	UNP H9L4M7
A	12	SER	-	expression tag	UNP H9L4M7
A	13	GLY	-	expression tag	UNP H9L4M7
A	14	LEU	-	expression tag	UNP H9L4M7
A	15	VAL	-	expression tag	UNP H9L4M7
A	16	PRO	-	expression tag	UNP H9L4M7
A	17	ARG	-	expression tag	UNP H9L4M7
A	18	GLY	-	expression tag	UNP H9L4M7
A	19	SER	-	expression tag	UNP H9L4M7
A	20	HIS	-	expression tag	UNP H9L4M7
A	21	MET	-	expression tag	UNP H9L4M7
A	148	ASP	-	linker	UNP H9L4M7

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	149	ASN	-	linker	UNP H9L4M7
A	150	LYS	-	linker	UNP H9L4M7
A	151	GLN	-	linker	UNP H9L4M7
A	296	ASP	-	linker	UNP Q8ZRK4
A	297	ASN	-	linker	UNP Q8ZRK4
A	298	LYS	-	linker	UNP Q8ZRK4
A	299	GLN	-	linker	UNP Q8ZRK4
A	444	ASP	-	linker	UNP Q8ZRK4
A	445	ASN	-	linker	UNP Q8ZRK4
A	446	LYS	-	linker	UNP Q8ZRK4
A	447	GLN	-	linker	UNP Q8ZRK4
B	1	MET	-	initiating methionine	UNP H9L4M7
B	2	GLY	-	expression tag	UNP H9L4M7
B	3	SER	-	expression tag	UNP H9L4M7
B	4	SER	-	expression tag	UNP H9L4M7
B	5	HIS	-	expression tag	UNP H9L4M7
B	6	HIS	-	expression tag	UNP H9L4M7
B	7	HIS	-	expression tag	UNP H9L4M7
B	8	HIS	-	expression tag	UNP H9L4M7
B	9	HIS	-	expression tag	UNP H9L4M7
B	10	HIS	-	expression tag	UNP H9L4M7
B	11	SER	-	expression tag	UNP H9L4M7
B	12	SER	-	expression tag	UNP H9L4M7
B	13	GLY	-	expression tag	UNP H9L4M7
B	14	LEU	-	expression tag	UNP H9L4M7
B	15	VAL	-	expression tag	UNP H9L4M7
B	16	PRO	-	expression tag	UNP H9L4M7
B	17	ARG	-	expression tag	UNP H9L4M7
B	18	GLY	-	expression tag	UNP H9L4M7
B	19	SER	-	expression tag	UNP H9L4M7
B	20	HIS	-	expression tag	UNP H9L4M7
B	21	MET	-	expression tag	UNP H9L4M7
B	148	ASP	-	linker	UNP H9L4M7
B	149	ASN	-	linker	UNP H9L4M7
B	150	LYS	-	linker	UNP H9L4M7
B	151	GLN	-	linker	UNP H9L4M7
B	296	ASP	-	linker	UNP Q8ZRK4
B	297	ASN	-	linker	UNP Q8ZRK4
B	298	LYS	-	linker	UNP Q8ZRK4
B	299	GLN	-	linker	UNP Q8ZRK4
B	444	ASP	-	linker	UNP Q8ZRK4
B	445	ASN	-	linker	UNP Q8ZRK4

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	446	LYS	-	linker	UNP Q8ZRK4
B	447	GLN	-	linker	UNP Q8ZRK4
C	1	MET	-	initiating methionine	UNP H9L4M7
C	2	GLY	-	expression tag	UNP H9L4M7
C	3	SER	-	expression tag	UNP H9L4M7
C	4	SER	-	expression tag	UNP H9L4M7
C	5	HIS	-	expression tag	UNP H9L4M7
C	6	HIS	-	expression tag	UNP H9L4M7
C	7	HIS	-	expression tag	UNP H9L4M7
C	8	HIS	-	expression tag	UNP H9L4M7
C	9	HIS	-	expression tag	UNP H9L4M7
C	10	HIS	-	expression tag	UNP H9L4M7
C	11	SER	-	expression tag	UNP H9L4M7
C	12	SER	-	expression tag	UNP H9L4M7
C	13	GLY	-	expression tag	UNP H9L4M7
C	14	LEU	-	expression tag	UNP H9L4M7
C	15	VAL	-	expression tag	UNP H9L4M7
C	16	PRO	-	expression tag	UNP H9L4M7
C	17	ARG	-	expression tag	UNP H9L4M7
C	18	GLY	-	expression tag	UNP H9L4M7
C	19	SER	-	expression tag	UNP H9L4M7
C	20	HIS	-	expression tag	UNP H9L4M7
C	21	MET	-	expression tag	UNP H9L4M7
C	148	ASP	-	linker	UNP H9L4M7
C	149	ASN	-	linker	UNP H9L4M7
C	150	LYS	-	linker	UNP H9L4M7
C	151	GLN	-	linker	UNP H9L4M7
C	296	ASP	-	linker	UNP Q8ZRK4
C	297	ASN	-	linker	UNP Q8ZRK4
C	298	LYS	-	linker	UNP Q8ZRK4
C	299	GLN	-	linker	UNP Q8ZRK4
C	444	ASP	-	linker	UNP Q8ZRK4
C	445	ASN	-	linker	UNP Q8ZRK4
C	446	LYS	-	linker	UNP Q8ZRK4
C	447	GLN	-	linker	UNP Q8ZRK4

- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	20	Total O 20 20	0	0
2	B	11	Total O 11 11	0	0

Continued on next page...

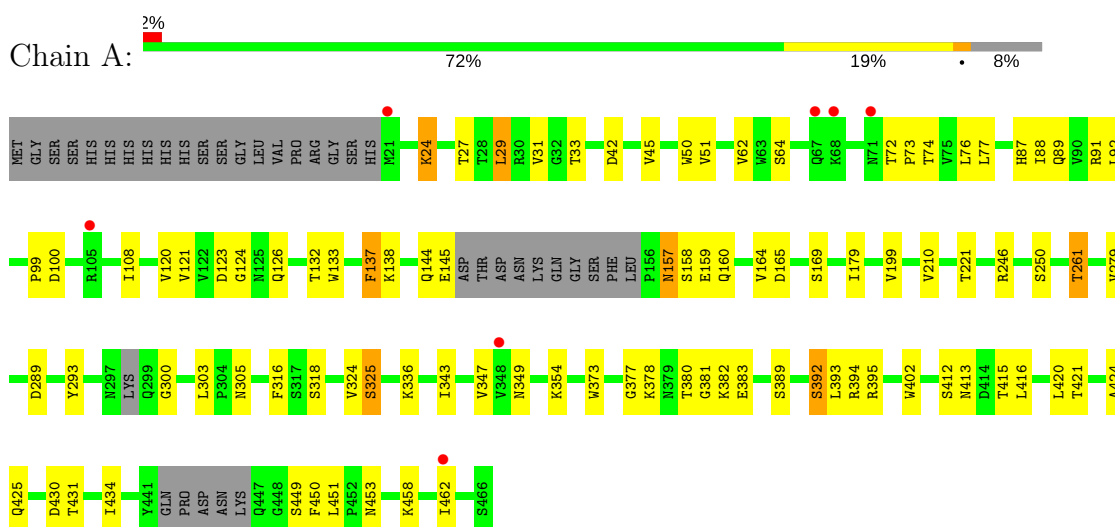
Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	C	19	Total	O	0	0
			19	19		

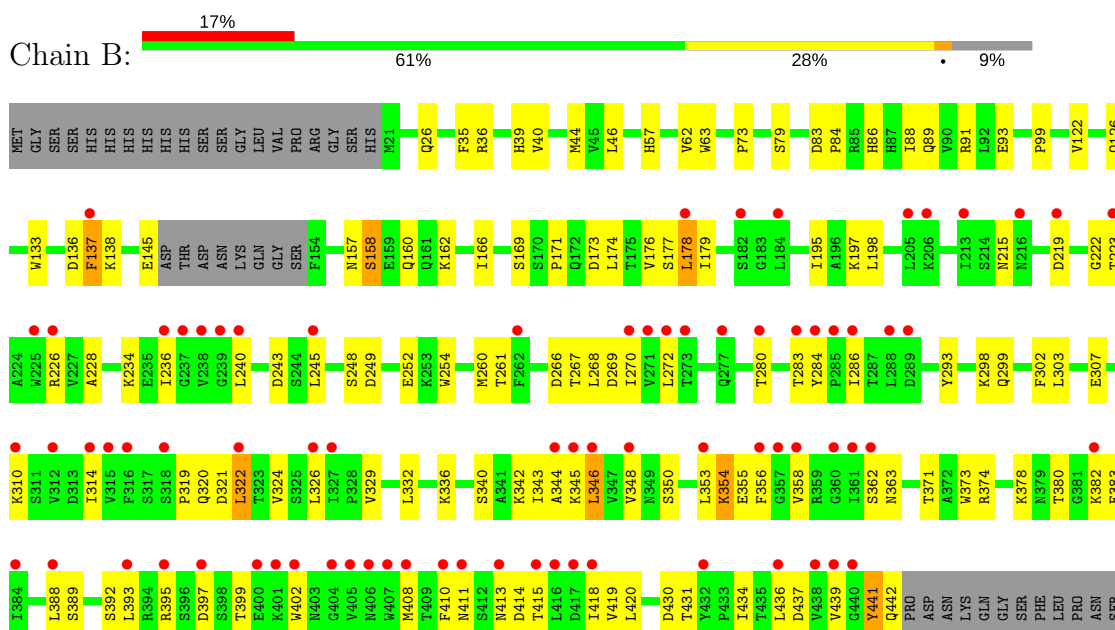
3 Residue-property plots [i](#)

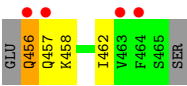
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are 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: SafD, Putative outer membrane protein, Putative outer membrane protein, Putative outer membrane protein

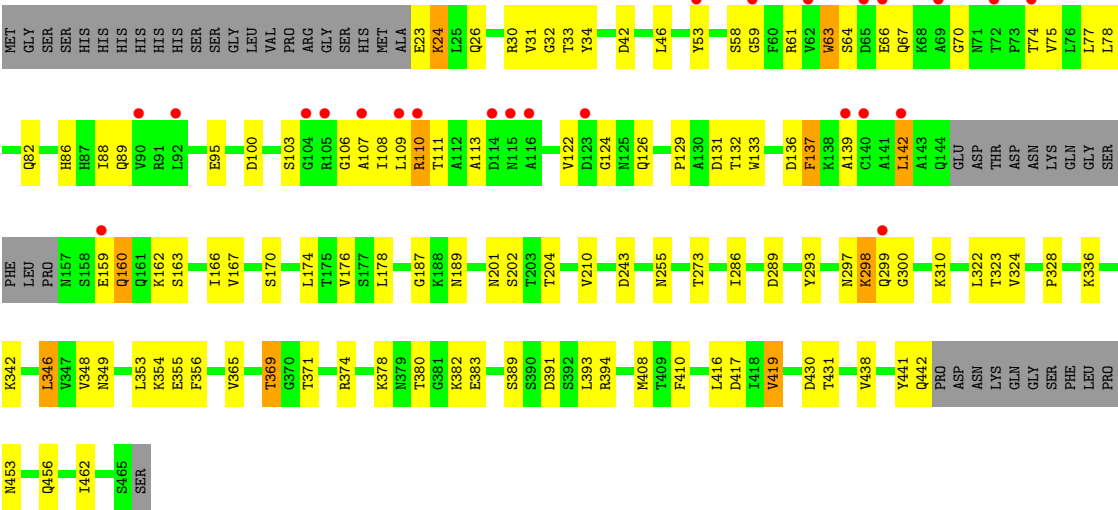


- Molecule 1: SafD, Putative outer membrane protein, Putative outer membrane protein, Putative outer membrane protein





● Molecule 1: SafD,Putative outer membrane protein,Putative outer membrane protein,Putative outer membrane protein



4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	133.27Å 66.07Å 187.68Å 90.00° 96.22° 90.00°	Depositor
Resolution (Å)	93.29 – 2.80 93.29 – 2.80	Depositor EDS
% Data completeness (in resolution range)	95.6 (93.29-2.80) 95.7 (93.29-2.80)	Depositor EDS
R_{merge}	0.14	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.50 (at 2.82Å)	Xtriage
Refinement program	PHENIX (1.11.1 _2575: ???)	Depositor
R, R_{free}	0.219 , 0.255 0.219 , 0.255	Depositor DCC
R_{free} test set	1894 reflections (4.89%)	wwPDB-VP
Wilson B-factor (Å ²)	71.7	Xtriage
Anisotropy	0.063	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.36 , 70.2	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	18834	wwPDB-VP
Average B, all atoms (Å ²)	102.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.20% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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

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	0.42	0/3256	0.62	0/4433
1	B	0.41	0/3209	0.65	2/4373 (0.0%)
1	C	0.33	0/3204	0.62	0/4364
All	All	0.39	0/9669	0.63	2/13170 (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
1	A	0	1

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	303	LEU	CB-CG-CD2	-6.39	100.14	111.00
1	B	346	LEU	CA-CB-CG	5.25	127.37	115.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	453	ASN	Peptide

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	3214	3099	3101	62	0
1	B	3165	3070	3062	111	0
1	C	3161	3075	3070	72	0
2	A	20	0	0	3	0
2	B	11	0	0	5	0
2	C	19	0	0	2	0
All	All	9590	9244	9233	241	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 13.

The worst 5 of 241 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:C:442:GLN:NE2	2:C:501:HOH:O	1.82	1.11
1:B:248:SER:OG	1:B:249:ASP:OD1	1.83	0.94
1:B:442:GLN:NE2	2:B:501:HOH:O	1.93	0.93
1:B:26:GLN:NE2	2:B:502:HOH:O	2.06	0.88
1:B:458:LYS:CD	1:B:458:LYS:CB	2.56	0.83

There are no symmetry-related clashes.

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	422/466 (91%)	404 (96%)	18 (4%)	0	100	100
1	B	418/466 (90%)	402 (96%)	14 (3%)	2 (0%)	31	65
1	C	415/466 (89%)	389 (94%)	25 (6%)	1 (0%)	49	81
All	All	1255/1398 (90%)	1195 (95%)	57 (4%)	3 (0%)	49	81

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	158	SER
1	C	34	TYR
1	B	320	GLN

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	A	347/395 (88%)	334 (96%)	13 (4%)	37	71
1	B	342/395 (87%)	332 (97%)	10 (3%)	45	79
1	C	344/395 (87%)	326 (95%)	18 (5%)	25	58
All	All	1033/1185 (87%)	992 (96%)	41 (4%)	34	68

5 of 41 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	322	LEU
1	B	456	GLN
1	C	346	LEU
1	B	336	LYS
1	B	354	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	86	HIS
1	B	258	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, 95th 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 ²)	Q<0.9
1	A	430/466 (92%)	0.38	7 (1%) 72 65	35, 74, 115, 159	1 (0%)
1	B	424/466 (90%)	0.98	81 (19%) 1 0	53, 111, 151, 179	0
1	C	421/466 (90%)	0.46	24 (5%) 24 15	39, 82, 160, 201	2 (0%)
All	All	1275/1398 (91%)	0.60	112 (8%) 10 5	35, 89, 151, 201	3 (0%)

The worst 5 of 112 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	284	TYR	5.7
1	B	439	VAL	5.2
1	B	356	PHE	5.0
1	B	357	GLY	4.6
1	C	116	ALA	4.4

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.