



wwPDB X-ray Structure Validation Summary Report ⓘ

May 14, 2020 – 10:56 am BST

PDB ID : 4Y68
Title : Structure of a lipoprotein from *Streptococcus agalactiae*
Authors : Khosa, S.; Hoepfner, A.; Smits, S.H.
Deposited on : 2015-02-12
Resolution : 2.21 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
Mogul : 1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix) : 1.13
EDS : 2.11
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac : 5.8.0158
CCP4 : 7.0.044 (Gargrove)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.11

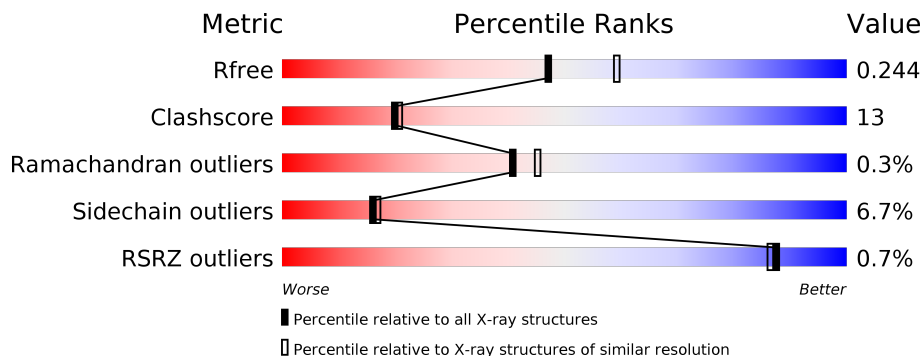
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.21 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}	130704	5912 (2.24-2.20)
Clashscore	141614	6646 (2.24-2.20)
Ramachandran outliers	138981	6543 (2.24-2.20)
Sidechain outliers	138945	6544 (2.24-2.20)
RSRZ outliers	127900	5797 (2.24-2.20)

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 respectively. 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	312	
1	B	312	
1	C	312	
1	D	312	

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit crite-

ria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	MES	C	401	-	-	X	-
2	MES	D	401	-	-	X	-

2 Entry composition i

There are 3 unique types of molecules in this entry. The entry contains 9601 atoms, of which 13 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 Putative nisin-resistance protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	C	287	2295	1450	408	431	6	0	0	0
1	A	287	2251	1424	391	430	6	0	0	0
1	B	286	2186	1378	379	423	6	0	0	0
1	D	287	2285	1445	405	429	6	0	0	0

There are 88 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	9	MET	-	expression tag	UNP Q3DLX7
C	10	GLY	-	expression tag	UNP Q3DLX7
C	11	SER	-	expression tag	UNP Q3DLX7
C	12	SER	-	expression tag	UNP Q3DLX7
C	13	HIS	-	expression tag	UNP Q3DLX7
C	14	HIS	-	expression tag	UNP Q3DLX7
C	15	HIS	-	expression tag	UNP Q3DLX7
C	16	HIS	-	expression tag	UNP Q3DLX7
C	17	HIS	-	expression tag	UNP Q3DLX7
C	18	HIS	-	expression tag	UNP Q3DLX7
C	19	HIS	-	expression tag	UNP Q3DLX7
C	20	HIS	-	expression tag	UNP Q3DLX7
C	21	SER	-	expression tag	UNP Q3DLX7
C	22	SER	-	expression tag	UNP Q3DLX7
C	23	GLY	-	expression tag	UNP Q3DLX7
C	24	LEU	-	expression tag	UNP Q3DLX7
C	25	VAL	-	expression tag	UNP Q3DLX7
C	26	PRO	-	expression tag	UNP Q3DLX7
C	27	ARG	-	expression tag	UNP Q3DLX7
C	28	GLY	-	expression tag	UNP Q3DLX7
C	29	SER	-	expression tag	UNP Q3DLX7

Continued on next page...

Continued from previous page...

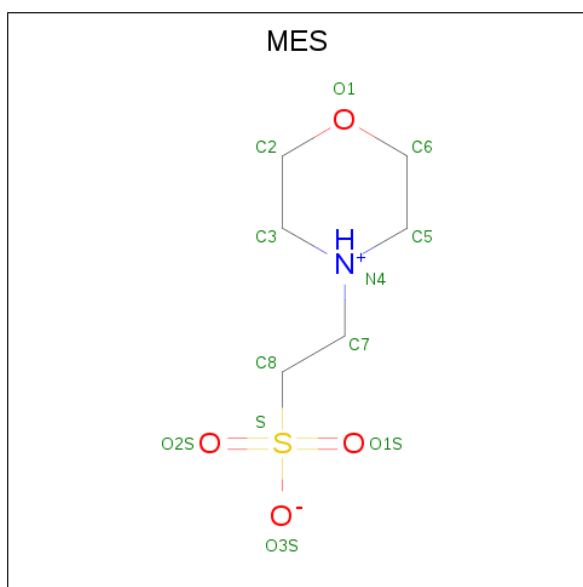
Chain	Residue	Modelled	Actual	Comment	Reference
C	30	HIS	-	expression tag	UNP Q3DLX7
A	9	MET	-	expression tag	UNP Q3DLX7
A	10	GLY	-	expression tag	UNP Q3DLX7
A	11	SER	-	expression tag	UNP Q3DLX7
A	12	SER	-	expression tag	UNP Q3DLX7
A	13	HIS	-	expression tag	UNP Q3DLX7
A	14	HIS	-	expression tag	UNP Q3DLX7
A	15	HIS	-	expression tag	UNP Q3DLX7
A	16	HIS	-	expression tag	UNP Q3DLX7
A	17	HIS	-	expression tag	UNP Q3DLX7
A	18	HIS	-	expression tag	UNP Q3DLX7
A	19	HIS	-	expression tag	UNP Q3DLX7
A	20	HIS	-	expression tag	UNP Q3DLX7
A	21	SER	-	expression tag	UNP Q3DLX7
A	22	SER	-	expression tag	UNP Q3DLX7
A	23	GLY	-	expression tag	UNP Q3DLX7
A	24	LEU	-	expression tag	UNP Q3DLX7
A	25	VAL	-	expression tag	UNP Q3DLX7
A	26	PRO	-	expression tag	UNP Q3DLX7
A	27	ARG	-	expression tag	UNP Q3DLX7
A	28	GLY	-	expression tag	UNP Q3DLX7
A	29	SER	-	expression tag	UNP Q3DLX7
A	30	HIS	-	expression tag	UNP Q3DLX7
B	9	MET	-	expression tag	UNP Q3DLX7
B	10	GLY	-	expression tag	UNP Q3DLX7
B	11	SER	-	expression tag	UNP Q3DLX7
B	12	SER	-	expression tag	UNP Q3DLX7
B	13	HIS	-	expression tag	UNP Q3DLX7
B	14	HIS	-	expression tag	UNP Q3DLX7
B	15	HIS	-	expression tag	UNP Q3DLX7
B	16	HIS	-	expression tag	UNP Q3DLX7
B	17	HIS	-	expression tag	UNP Q3DLX7
B	18	HIS	-	expression tag	UNP Q3DLX7
B	19	HIS	-	expression tag	UNP Q3DLX7
B	20	HIS	-	expression tag	UNP Q3DLX7
B	21	SER	-	expression tag	UNP Q3DLX7
B	22	SER	-	expression tag	UNP Q3DLX7
B	23	GLY	-	expression tag	UNP Q3DLX7
B	24	LEU	-	expression tag	UNP Q3DLX7
B	25	VAL	-	expression tag	UNP Q3DLX7
B	26	PRO	-	expression tag	UNP Q3DLX7
B	27	ARG	-	expression tag	UNP Q3DLX7

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	28	GLY	-	expression tag	UNP Q3DLX7
B	29	SER	-	expression tag	UNP Q3DLX7
B	30	HIS	-	expression tag	UNP Q3DLX7
D	9	MET	-	expression tag	UNP Q3DLX7
D	10	GLY	-	expression tag	UNP Q3DLX7
D	11	SER	-	expression tag	UNP Q3DLX7
D	12	SER	-	expression tag	UNP Q3DLX7
D	13	HIS	-	expression tag	UNP Q3DLX7
D	14	HIS	-	expression tag	UNP Q3DLX7
D	15	HIS	-	expression tag	UNP Q3DLX7
D	16	HIS	-	expression tag	UNP Q3DLX7
D	17	HIS	-	expression tag	UNP Q3DLX7
D	18	HIS	-	expression tag	UNP Q3DLX7
D	19	HIS	-	expression tag	UNP Q3DLX7
D	20	HIS	-	expression tag	UNP Q3DLX7
D	21	SER	-	expression tag	UNP Q3DLX7
D	22	SER	-	expression tag	UNP Q3DLX7
D	23	GLY	-	expression tag	UNP Q3DLX7
D	24	LEU	-	expression tag	UNP Q3DLX7
D	25	VAL	-	expression tag	UNP Q3DLX7
D	26	PRO	-	expression tag	UNP Q3DLX7
D	27	ARG	-	expression tag	UNP Q3DLX7
D	28	GLY	-	expression tag	UNP Q3DLX7
D	29	SER	-	expression tag	UNP Q3DLX7
D	30	HIS	-	expression tag	UNP Q3DLX7

- Molecule 2 is 2-(N-MORPHOLINO)-ETHANESULFONIC ACID (three-letter code: MES) (formula: C₆H₁₃NO₄S).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	S		
2	C	1	12	6	1	4	1	0	0
2	C	1	25	6	13	1	4	1	0
2	A	1	12	6	1	4	1	0	0
2	D	1	12	6	1	4	1	0	0

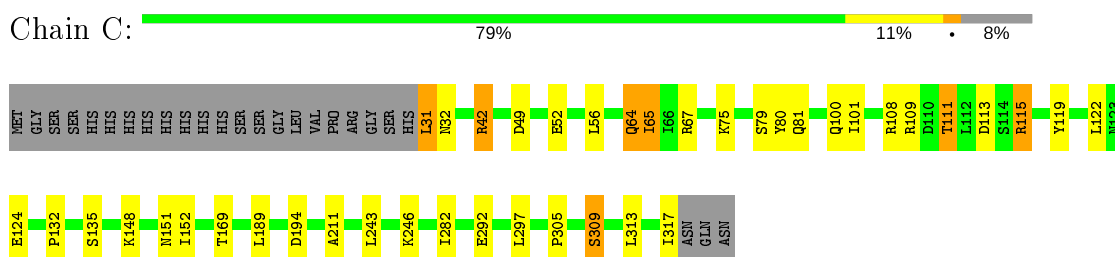
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	C	193	Total	O	0	0
			193	193		
3	A	110	Total	O	0	0
			110	110		
3	B	73	Total	O	0	0
			73	73		
3	D	147	Total	O	0	0
			147	147		

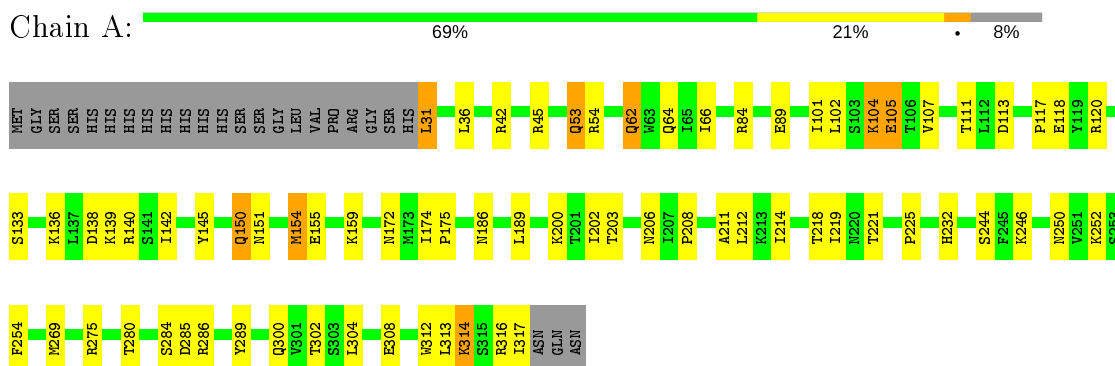
3 Residue-property plots [i](#)

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.

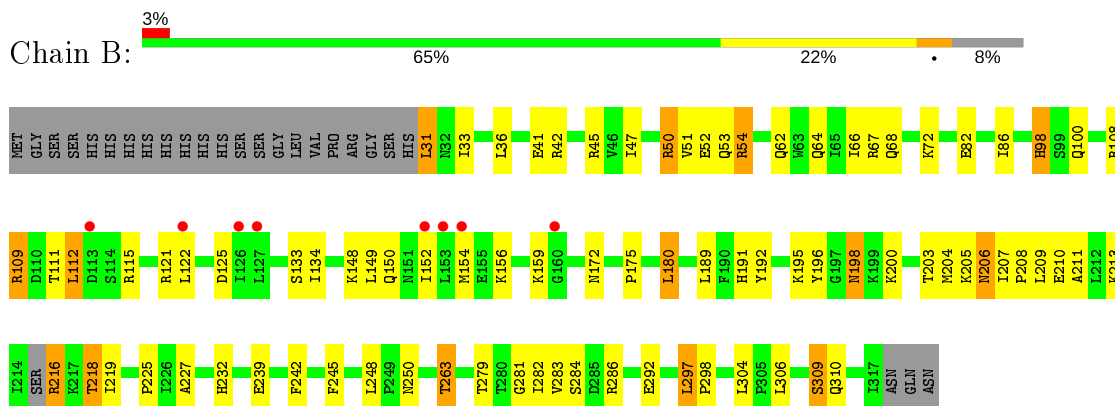
- Molecule 1: Putative nisin-resistance protein



- Molecule 1: Putative nisin-resistance protein

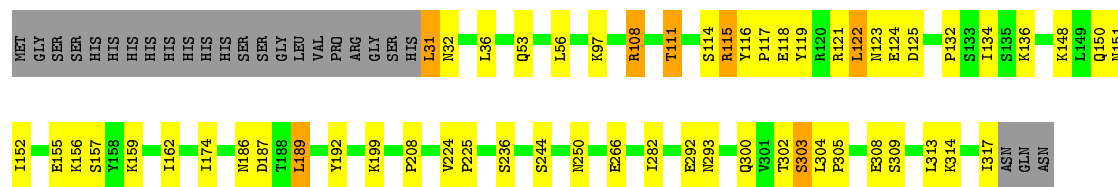


- Molecule 1: Putative nisin-resistance protein



- Molecule 1: Putative nisin-resistance protein

Chain D:  74% 16% 8%



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	58.81Å 137.24Å 164.00Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	55.36 – 2.21 55.36 – 2.21	Depositor EDS
% Data completeness (in resolution range)	99.6 (55.36-2.21) 99.6 (55.36-2.21)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.55 (at 2.20Å)	Xtrriage
Refinement program	PHENIX (phenix.refine: 1.9_1692)	Depositor
R, R_{free}	0.194 , 0.243 0.198 , 0.244	Depositor DCC
R_{free} test set	2000 reflections (2.98%)	wwPDB-VP
Wilson B-factor (Å ²)	34.8	Xtrriage
Anisotropy	0.040	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.32 , 48.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	9601	wwPDB-VP
Average B, all atoms (Å ²)	40.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.81% 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 [i](#)

5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: MES

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	0.38	0/2291	0.54	0/3100
1	B	0.34	0/2224	0.52	0/3018
1	C	0.47	0/2335	0.59	0/3149
1	D	0.44	0/2325	0.59	0/3137
All	All	0.41	0/9175	0.56	0/12404

There are no bond length outliers.

There are no bond angle outliers.

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	2251	0	2272	64	0
1	B	2186	0	2138	77	0
1	C	2295	0	2366	40	0
1	D	2285	0	2351	57	0
2	A	12	0	12	0	0
2	C	24	13	24	8	0
2	D	12	0	12	9	0
3	A	110	0	0	9	0
3	B	73	0	0	8	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	C	193	0	0	13	1
3	D	147	0	0	7	1
All	All	9588	13	9175	229	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 13.

The worst 5 of 229 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:D:121:ARG:HE	2:D:401:MES:H52	1.09	1.16
1:C:42:ARG:NH2	3:C:607:HOH:O	1.89	1.03
1:A:53:GLN:HG2	1:A:54:ARG:HG2	1.47	0.96
1:C:49:ASP:OD2	3:C:607:HOH:O	1.83	0.95
1:C:80:TYR:H	2:C:401:MES:H51	1.27	0.95

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)
3:C:526:HOH:O	3:D:524:HOH:O[4_475]	2.04	0.16

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	285/312 (91%)	272 (95%)	13 (5%)	0	100 100
1	B	282/312 (90%)	264 (94%)	17 (6%)	1 (0%)	34 37
1	C	285/312 (91%)	278 (98%)	6 (2%)	1 (0%)	34 37
1	D	285/312 (91%)	277 (97%)	7 (2%)	1 (0%)	34 37

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
All	All	1137/1248 (91%)	1091 (96%)	43 (4%)	3 (0%)	41	45

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	65	ILE
1	D	236	SER
1	B	208	PRO

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	249/280 (89%)	227 (91%)	22 (9%)	10	9
1	B	234/280 (84%)	214 (92%)	20 (8%)	10	9
1	C	258/280 (92%)	244 (95%)	14 (5%)	22	25
1	D	256/280 (91%)	245 (96%)	11 (4%)	29	35
All	All	997/1120 (89%)	930 (93%)	67 (7%)	16	17

5 of 67 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	214	ILE
1	B	42	ARG
1	D	118	GLU
1	A	246	LYS
1	A	304	LEU

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	206	ASN
1	B	100	GLN

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](#)

4 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
2	MES	A	401	-	12,12,12	2.17	1 (8%)	14,16,16	2.56	7 (50%)
2	MES	D	401	-	12,12,12	2.30	1 (8%)	14,16,16	2.65	6 (42%)
2	MES	C	402	-	12,12,12	2.36	1 (8%)	14,16,16	2.36	8 (57%)
2	MES	C	401	-	12,12,12	2.16	1 (8%)	14,16,16	2.57	7 (50%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	MES	A	401	-	-	5/6/14/14	0/1/1/1
2	MES	D	401	-	-	5/6/14/14	0/1/1/1
2	MES	C	402	-	-	6/6/14/14	0/1/1/1
2	MES	C	401	-	-	6/6/14/14	0/1/1/1

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	402	MES	C8-S	-7.81	1.66	1.77
2	D	401	MES	C8-S	-7.57	1.66	1.77
2	A	401	MES	C8-S	-7.20	1.67	1.77
2	C	401	MES	C8-S	-7.10	1.67	1.77

The worst 5 of 28 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	D	401	MES	C2-C3-N4	-5.40	101.91	110.10
2	A	401	MES	C5-N4-C3	4.80	119.62	108.83
2	C	401	MES	C6-C5-N4	-4.56	103.19	110.10
2	C	402	MES	O3S-S-C8	4.40	112.89	105.77
2	C	401	MES	C5-N4-C3	4.34	118.60	108.83

There are no chirality outliers.

5 of 22 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	401	MES	C8-C7-N4-C3
2	D	401	MES	C7-C8-S-O1S
2	D	401	MES	C7-C8-S-O3S
2	C	402	MES	C8-C7-N4-C3
2	C	402	MES	C8-C7-N4-C5

There are no ring outliers.

3 monomers are involved in 17 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	D	401	MES	9	0
2	C	402	MES	2	0
2	C	401	MES	6	0

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	287/312 (91%)	-0.34	0 100 100	28, 41, 58, 66	0
1	B	286/312 (91%)	0.03	8 (2%) 53 51	35, 52, 71, 81	0
1	C	287/312 (91%)	-0.29	0 100 100	19, 29, 44, 57	0
1	D	287/312 (91%)	-0.32	0 100 100	20, 34, 54, 61	0
All	All	1147/1248 (91%)	-0.23	8 (0%) 87 86	19, 39, 61, 81	0

The worst 5 of 8 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	160	GLY	4.2
1	B	122	LEU	3.6
1	B	127	LEU	3.0
1	B	153	LEU	2.7
1	B	152	ILE	2.6

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](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(\AA^2)	Q<0.9
2	MES	A	401	12/12	0.75	0.30	105,113,119,120	0
2	MES	C	402	12/12	0.81	0.27	25,43,54,61	0
2	MES	D	401	12/12	0.82	0.27	58,68,81,88	0
2	MES	C	401	12/12	0.89	0.22	34,44,60,64	0

6.5 Other polymers [i](#)

There are no such residues in this entry.