



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 08:55 PM GMT

PDB ID : 4U8Y
Title : Coupling of remote alternating-access transport mechanisms for protons and substrates in the multidrug efflux pump AcrB
Authors : Pos, K.M.
Deposited on : 2014-08-05
Resolution : 2.10 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
<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

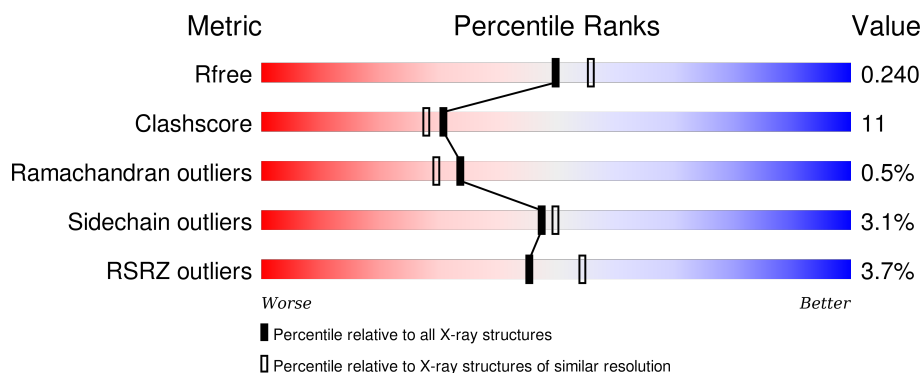
1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.10 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}	91344	3939 (2.10-2.10)
Clashscore	102246	4460 (2.10-2.10)
Ramachandran outliers	100387	4413 (2.10-2.10)
Sidechain outliers	100360	4414 (2.10-2.10)
RSRZ outliers	91569	3948 (2.10-2.10)

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 ≥ 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	1057	<div> <div>6%</div> <div>75% 22% ..</div> </div>
1	B	1057	<div> <div>3%</div> <div>78% 18% ..</div> </div>
1	C	1057	<div> <div>%</div> <div>80% 17% ..</div> </div>
2	D	169	<div> <div>%</div> <div>78% 14% • 8%</div> </div>
2	E	169	<div> <div>15%</div> <div>70% 19% • 10%</div> </div>

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 criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	LMT	A	1101	-	-	-	X
3	LMT	A	1102	-	-	-	X

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 28268 atoms, of which 27 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 Multidrug efflux pump subunit AcrB.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	1044	Total	C	N	O	S	0	0	0
			7943	5106	1316	1477	44			
1	B	1033	Total	C	N	O	S	0	0	0
			7849	5052	1296	1457	44			
1	C	1033	Total	C	N	O	S	0	0	0
			7849	5052	1296	1457	44			

There are 27 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	408	ASN	ASP	engineered mutation	UNP P31224
A	1050	LEU	-	expression tag	UNP P31224
A	1051	GLU	-	expression tag	UNP P31224
A	1052	HIS	-	expression tag	UNP P31224
A	1053	HIS	-	expression tag	UNP P31224
A	1054	HIS	-	expression tag	UNP P31224
A	1055	HIS	-	expression tag	UNP P31224
A	1056	HIS	-	expression tag	UNP P31224
A	1057	HIS	-	expression tag	UNP P31224
B	408	ASN	ASP	engineered mutation	UNP P31224
B	1050	LEU	-	expression tag	UNP P31224
B	1051	GLU	-	expression tag	UNP P31224
B	1052	HIS	-	expression tag	UNP P31224
B	1053	HIS	-	expression tag	UNP P31224
B	1054	HIS	-	expression tag	UNP P31224
B	1055	HIS	-	expression tag	UNP P31224
B	1056	HIS	-	expression tag	UNP P31224
B	1057	HIS	-	expression tag	UNP P31224
C	408	ASN	ASP	engineered mutation	UNP P31224
C	1050	LEU	-	expression tag	UNP P31224
C	1051	GLU	-	expression tag	UNP P31224
C	1052	HIS	-	expression tag	UNP P31224
C	1053	HIS	-	expression tag	UNP P31224

Continued on next page...

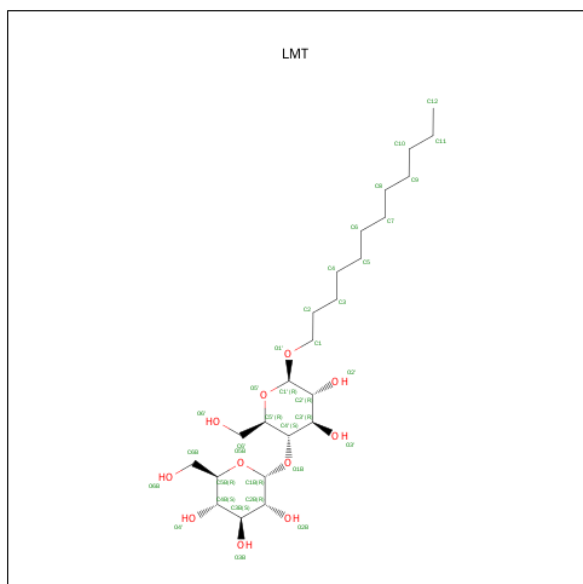
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	1054	HIS	-	expression tag	UNP P31224
C	1055	HIS	-	expression tag	UNP P31224
C	1056	HIS	-	expression tag	UNP P31224
C	1057	HIS	-	expression tag	UNP P31224

- Molecule 2 is a protein called DARPin.

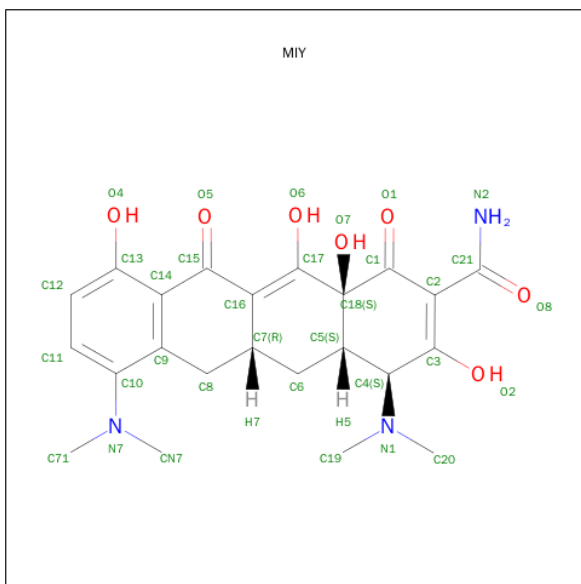
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	D	156	Total	C	N	O	S	0	0	0
			1177	741	206	229	1			
2	E	152	Total	C	N	O	S	0	0	0
			1151	726	202	222	1			

- Molecule 3 is DODECYL-BETA-D-MALTOSE (three-letter code: LMT) (formula: $C_{24}H_{46}O_{11}$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	O	0	0
			35	24	11		
3	A	1	Total	C	O	0	0
			32	21	11		
3	B	1	Total	C	O	0	0
			26	15	11		
3	C	1	Total	C	O	0	0
			35	24	11		

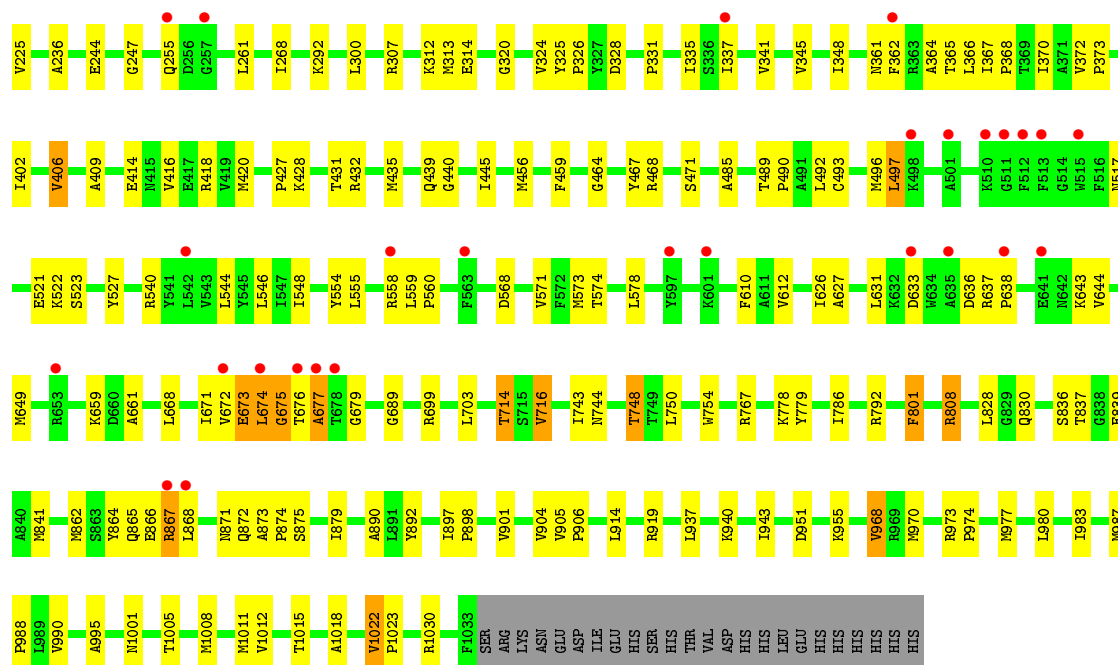
- Molecule 4 is (4S,4AS,5AR,12AS)-4,7-BIS(DIMETHYLAMINO)-3,10,12,12A-TETRAHYDROXY-1,11-DIOXO-1,4,4A,5,5A,6,11,12A-OCTAHYDROTETRACENE-2-CARBOXAMIDE (three-letter code: MIY) (formula: $C_{23}H_{27}N_3O_7$).



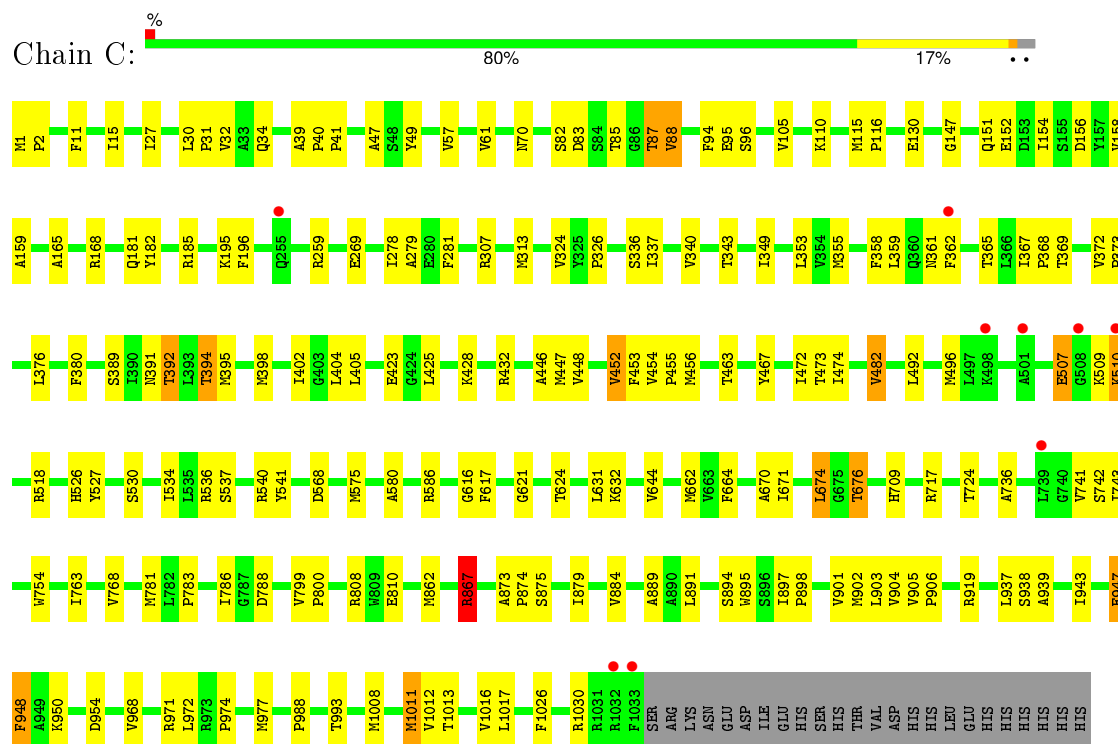
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	B	1	Total	C	H	N	O	
			60	23	27	3	7	

- Molecule 5 is water.

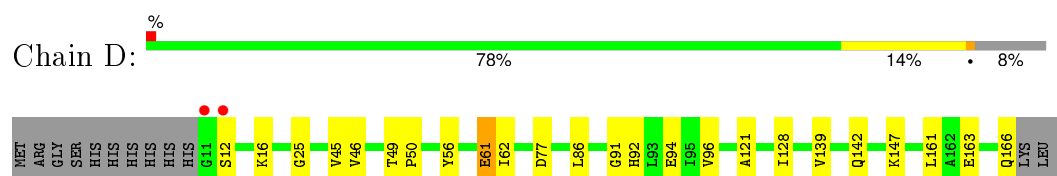
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	645	Total	O		
			645	645	0	0
5	B	566	Total	O		
			566	566	0	0
5	C	723	Total	O		
			723	723	0	0
5	D	103	Total	O		
			103	103	0	0
5	E	74	Total	O		
			74	74	0	0



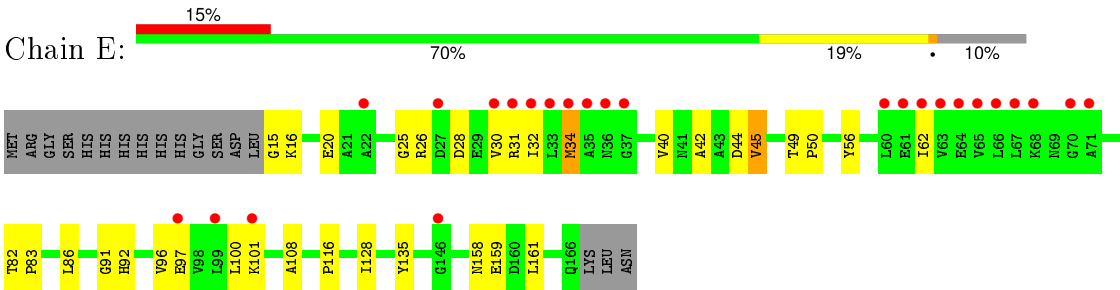
• Molecule 1: Multidrug efflux pump subunit AcrB



• Molecule 2: DARPin



● Molecule 2: DARPin



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	145.88Å 161.16Å 245.29Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.80 – 2.10 48.80 – 2.10	Depositor EDS
% Data completeness (in resolution range)	99.1 (48.80-2.10) 99.2 (48.80-2.10)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.48 (at 2.10Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.9_1692)	Depositor
R, R_{free}	0.193 , 0.236 0.200 , 0.240	Depositor DCC
R_{free} test set	16611 reflections (5.26%)	DCC
Wilson B-factor (Å ²)	28.7	Xtriage
Anisotropy	0.393	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.35 , 64.0	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.47$, $\langle L^2 \rangle = 0.29$	Xtriage
Outliers	1 of 332174 reflections (0.000%)	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	28268	wwPDB-VP
Average B, all atoms (Å ²)	41.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.91% 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, 0.375 respectively for untwinned datasets, and 0.333, 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: MIY, LMT

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	0.39	0/8095	0.52	0/10991
1	B	0.38	0/7999	0.52	0/10863
1	C	0.43	0/7999	0.55	2/10863 (0.0%)
2	D	0.34	0/1196	0.48	0/1626
2	E	0.31	0/1170	0.46	0/1591
All	All	0.39	0/26459	0.53	2/35934 (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	B	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	C	867	ARG	NE-CZ-NH1	5.60	123.10	120.30
1	C	867	ARG	NE-CZ-NH2	-5.53	117.54	120.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	689	GLY	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	7943	0	8086	208	0
1	B	7849	0	8003	168	0
1	C	7849	0	8003	166	0
2	D	1177	0	1159	14	0
2	E	1151	0	1136	29	0
3	A	67	0	83	5	0
3	B	26	0	25	5	0
3	C	35	0	46	2	0
4	B	33	27	25	0	0
5	A	645	0	0	15	0
5	B	566	0	0	18	0
5	C	723	0	0	21	0
5	D	103	0	0	3	0
5	E	74	0	0	3	0
All	All	28241	27	26566	575	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 11.

The worst 5 of 575 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:669:PRO:HD2	1:A:672:VAL:HG12	1.45	0.96
2:E:34:MET:HE3	2:E:40:VAL:HG12	1.48	0.95
1:B:445:ILE:HD13	1:B:940:LYS:HG3	1.51	0.93
1:C:527:TYR:CE2	1:C:968:VAL:HG13	2.07	0.89
3:B:1101:LMT:H6E	3:B:1101:LMT:H5B	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	1042/1057 (99%)	997 (96%)	33 (3%)	12 (1%)	16	10
1	B	1031/1057 (98%)	1000 (97%)	25 (2%)	6 (1%)	30	24
1	C	1031/1057 (98%)	1003 (97%)	28 (3%)	0	100	100
2	D	154/169 (91%)	152 (99%)	2 (1%)	0	100	100
2	E	150/169 (89%)	149 (99%)	1 (1%)	0	100	100
All	All	3408/3509 (97%)	3301 (97%)	89 (3%)	18 (0%)	34	30

5 of 18 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	672	VAL
1	A	867	ARG
1	A	1038	GLU
1	A	1039	ASP
1	A	1043	SER

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	850/863 (98%)	822 (97%)	28 (3%)	45	47
1	B	839/863 (97%)	812 (97%)	27 (3%)	46	48
1	C	839/863 (97%)	815 (97%)	24 (3%)	50	53
2	D	120/132 (91%)	117 (98%)	3 (2%)	55	59

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
2	E	117/132 (89%)	114 (97%)	3 (3%)	54 58
All	All	2765/2853 (97%)	2680 (97%)	85 (3%)	47 50

5 of 85 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	497	LEU
1	B	801	PHE
1	C	1011	MET
1	B	546	LEU
1	B	673	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	C	68	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](#)

5 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	$\# Z > 2$
3	LMT	A	1101	-	36,36,36	0.39	0	47,47,47	0.96	3 (6%)
3	LMT	A	1102	-	33,33,36	0.38	0	44,44,47	1.36	6 (13%)
3	LMT	B	1101	-	27,27,36	0.45	0	38,38,47	1.00	3 (7%)
4	MIY	B	1102	-	35,36,36	1.67	8 (22%)	40,58,58	2.78	19 (47%)
3	LMT	C	1101	-	36,36,36	0.42	0	47,47,47	0.91	1 (2%)

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
3	LMT	A	1101	-	-	0/21/61/61	0/2/2/2
3	LMT	A	1102	-	-	0/18/58/61	0/2/2/2
3	LMT	B	1101	-	-	0/12/52/61	0/2/2/2
4	MIY	B	1102	-	-	0/12/70/70	0/4/4/4
3	LMT	C	1101	-	-	0/21/61/61	0/2/2/2

The worst 5 of 8 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	B	1102	MIY	C7-C16	-3.14	1.48	1.51
4	B	1102	MIY	C14-C13	-2.49	1.37	1.41
4	B	1102	MIY	C2-C21	2.34	1.51	1.47
4	B	1102	MIY	O4-C13	2.38	1.41	1.36
4	B	1102	MIY	C18-C17	2.52	1.54	1.52

The worst 5 of 32 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	B	1102	MIY	C11-C12-C13	-6.94	113.38	120.49
4	B	1102	MIY	O6-C17-C16	-5.23	118.83	123.84
4	B	1102	MIY	O5-C15-C14	-4.08	114.35	122.01
4	B	1102	MIY	C11-C10-N7	-3.50	116.69	121.59
4	B	1102	MIY	O7-C18-C17	-2.98	104.50	109.85

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

4 monomers are involved in 12 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	1101	LMT	3	0
3	A	1102	LMT	2	0
3	B	1101	LMT	5	0
3	C	1101	LMT	2	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 ⓘ

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, 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	1044/1057 (98%)	-0.03	61 (5%) 26 34	15, 40, 88, 156	0
1	B	1033/1057 (97%)	-0.21	29 (2%) 56 64	16, 39, 67, 139	0
1	C	1033/1057 (97%)	-0.36	9 (0%) 85 88	16, 31, 58, 95	0
2	D	156/169 (92%)	-0.34	2 (1%) 79 84	27, 38, 62, 109	0
2	E	152/169 (89%)	0.61	25 (16%) 2 3	27, 46, 77, 98	0
All	All	3418/3509 (97%)	-0.17	126 (3%) 45 54	15, 37, 74, 156	0

The worst 5 of 126 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	D	11	GLY	10.7
1	A	870	GLY	7.6
1	B	674	LEU	6.7
2	E	35	ALA	5.8
2	E	33	LEU	5.7

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 ⓘ

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. LLDF column lists the quality of electron

density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(Å ²)	Q<0.9
3	LMT	A	1102	32/35	0.91	0.14	5.69	42,66,95,97	0
3	LMT	A	1101	35/35	0.91	0.15	4.63	45,60,82,83	0
3	LMT	C	1101	35/35	0.95	0.10	0.14	48,55,70,80	0
4	MIY	B	1102	33/33	0.89	0.14	0.12	48,65,83,92	0
3	LMT	B	1101	26/35	0.90	0.10	-0.04	54,60,77,89	0

6.5 Other polymers [i](#)

There are no such residues in this entry.