



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 02:10 PM GMT

PDB ID : 3W9J  
Title : Structural basis for the inhibition of bacterial multidrug exporters  
Authors : Sakurai, K.; Nakashima, R.; Hayashi, K.; Yamaguchi, A.  
Deposited on : 2013-04-04  
Resolution : 3.15 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

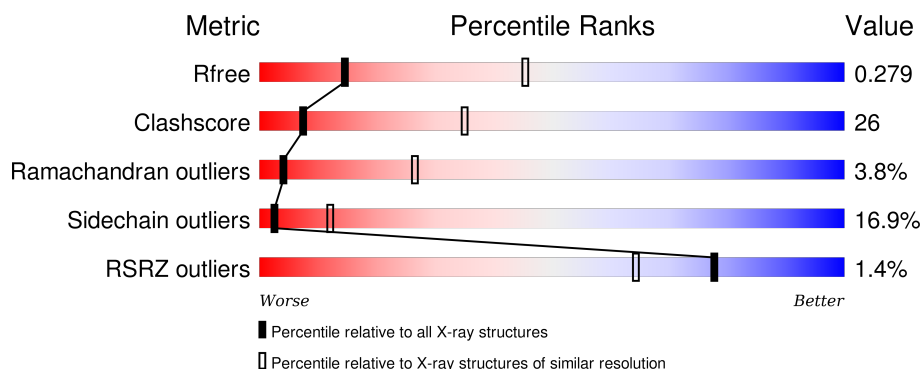
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 3.15 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	1112 (3.20-3.12)
Clashscore	102246	1249 (3.20-3.12)
Ramachandran outliers	100387	1222 (3.20-3.12)
Sidechain outliers	100360	1221 (3.20-3.12)
RSRZ outliers	91569	1117 (3.20-3.12)

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	1052	<div> <div> <div></div> <div>52%</div> <div>37%</div> <div>8%</div> <div>.</div> </div> </div>
1	B	1052	<div> <div> <div></div> <div>52%</div> <div>38%</div> <div>8%</div> <div>.</div> </div> </div>
1	C	1052	<div> <div> <div>3%</div> <div>46%</div> <div>40%</div> <div>12%</div> <div>.</div> </div> </div>
1	D	1052	<div> <div> <div></div> <div>48%</div> <div>40%</div> <div>8%</div> <div>.</div> </div> </div>
1	E	1052	<div> <div> <div></div> <div>50%</div> <div>38%</div> <div>10%</div> <div>.</div> </div> </div>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	F	1052	

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
2	LMT	A	2001	-	-	-	X
2	LMT	A	2002	-	-	-	X
2	LMT	B	2002	-	-	-	X
2	LMT	B	2003	-	-	-	X
2	LMT	B	2004	-	-	-	X
2	LMT	C	2001	-	-	-	X
2	LMT	D	2001	-	-	-	X
2	LMT	D	2002	-	-	-	X
2	LMT	E	2003	-	-	-	X
2	LMT	E	2004	-	-	-	X
2	LMT	F	2002	-	-	-	X
3	P9D	E	2001	-	-	-	X

## 2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 47358 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 Multidrug resistance protein MexB.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	1018	Total	C	N	O	S	0	0	0
			7724	4975	1280	1429	40			
1	B	1030	Total	C	N	O	S	0	0	0
			7812	5027	1298	1447	40			
1	C	1030	Total	C	N	O	S	0	0	0
			7812	5027	1298	1447	40			
1	D	1019	Total	C	N	O	S	0	0	0
			7735	4984	1281	1430	40			
1	E	1030	Total	C	N	O	S	0	0	0
			7812	5027	1298	1447	40			
1	F	1033	Total	C	N	O	S	0	0	0
			7840	5046	1302	1452	40			

There are 36 discrepancies between the modelled and reference sequences:

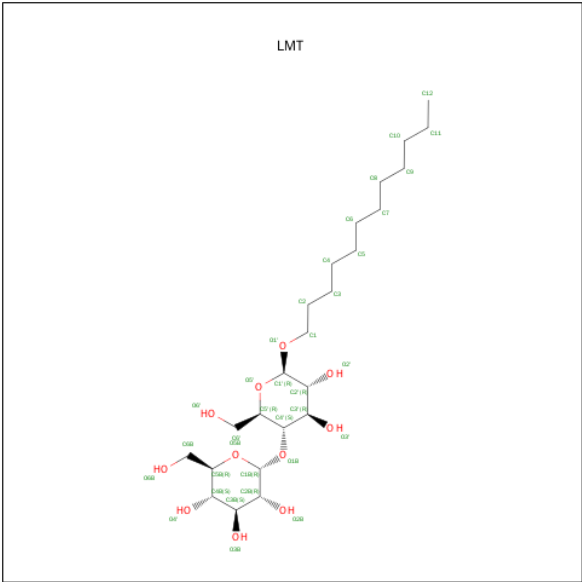
Chain	Residue	Modelled	Actual	Comment	Reference
A	1047	HIS	-	EXPRESSION TAG	UNP P52002
A	1048	HIS	-	EXPRESSION TAG	UNP P52002
A	1049	HIS	-	EXPRESSION TAG	UNP P52002
A	1050	HIS	-	EXPRESSION TAG	UNP P52002
A	1051	HIS	-	EXPRESSION TAG	UNP P52002
A	1052	HIS	-	EXPRESSION TAG	UNP P52002
B	1047	HIS	-	EXPRESSION TAG	UNP P52002
B	1048	HIS	-	EXPRESSION TAG	UNP P52002
B	1049	HIS	-	EXPRESSION TAG	UNP P52002
B	1050	HIS	-	EXPRESSION TAG	UNP P52002
B	1051	HIS	-	EXPRESSION TAG	UNP P52002
B	1052	HIS	-	EXPRESSION TAG	UNP P52002
C	1047	HIS	-	EXPRESSION TAG	UNP P52002
C	1048	HIS	-	EXPRESSION TAG	UNP P52002
C	1049	HIS	-	EXPRESSION TAG	UNP P52002
C	1050	HIS	-	EXPRESSION TAG	UNP P52002
C	1051	HIS	-	EXPRESSION TAG	UNP P52002

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	1052	HIS	-	EXPRESSION TAG	UNP P52002
D	1047	HIS	-	EXPRESSION TAG	UNP P52002
D	1048	HIS	-	EXPRESSION TAG	UNP P52002
D	1049	HIS	-	EXPRESSION TAG	UNP P52002
D	1050	HIS	-	EXPRESSION TAG	UNP P52002
D	1051	HIS	-	EXPRESSION TAG	UNP P52002
D	1052	HIS	-	EXPRESSION TAG	UNP P52002
E	1047	HIS	-	EXPRESSION TAG	UNP P52002
E	1048	HIS	-	EXPRESSION TAG	UNP P52002
E	1049	HIS	-	EXPRESSION TAG	UNP P52002
E	1050	HIS	-	EXPRESSION TAG	UNP P52002
E	1051	HIS	-	EXPRESSION TAG	UNP P52002
E	1052	HIS	-	EXPRESSION TAG	UNP P52002
F	1047	HIS	-	EXPRESSION TAG	UNP P52002
F	1048	HIS	-	EXPRESSION TAG	UNP P52002
F	1049	HIS	-	EXPRESSION TAG	UNP P52002
F	1050	HIS	-	EXPRESSION TAG	UNP P52002
F	1051	HIS	-	EXPRESSION TAG	UNP P52002
F	1052	HIS	-	EXPRESSION TAG	UNP P52002

- Molecule 2 is DODECYL-BETA-D-MALTOSE (three-letter code: LMT) (formula: C<sub>24</sub>H<sub>46</sub>O<sub>11</sub>).



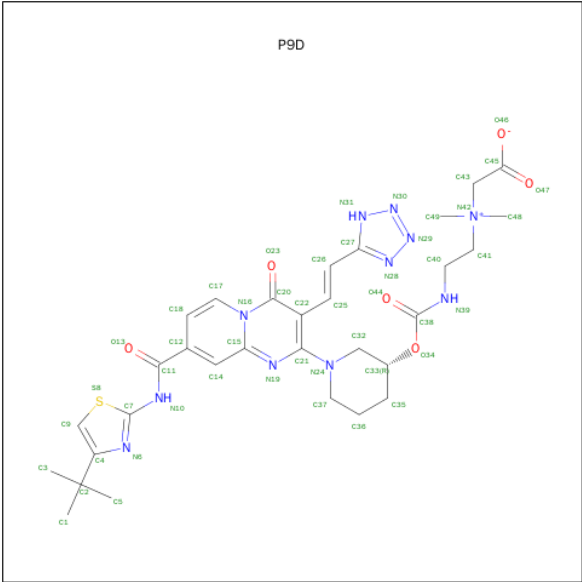
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	C	O	0	0
			35	24	11		

Continued on next page...

*Continued from previous page...*

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	C	O	0	0
			35	24	11		
2	A	1	Total	C	O	0	0
			35	24	11		
2	B	1	Total	C	O	0	0
			35	24	11		
2	B	1	Total	C	O	0	0
			35	24	11		
2	B	1	Total	C	O	0	0
			35	24	11		
2	C	1	Total	C	O	0	0
			35	24	11		
2	D	1	Total	C	O	0	0
			35	24	11		
2	D	1	Total	C	O	0	0
			35	24	11		
2	D	1	Total	C	O	0	0
			35	24	11		
2	E	1	Total	C	O	0	0
			35	24	11		
2	E	1	Total	C	O	0	0
			35	24	11		
2	E	1	Total	C	O	0	0
			35	24	11		
2	F	1	Total	C	O	0	0
			35	24	11		
2	F	1	Total	C	O	0	0
			35	24	11		

- Molecule 3 is [{2-[(3R)-1-{8-[(4-TERT-BUTYL-1,3-THIAZOL-2-YL)CARBAMOYL]-4-OXO-3-[(E)-2-(1H-TETRAZOL-5-YL)ETHENYL]-4H-PYRIDO[1,2-A]PYRIMIDIN-2-YL} PIPERIDIN-3-YL]OXY} CARBONYL)AMINO]ETHYL}(DIMETHYL)AMMONIO]ACETATE (three-letter code: P9D) (formula: C<sub>31</sub>H<sub>39</sub>N<sub>11</sub>O<sub>6</sub>S).

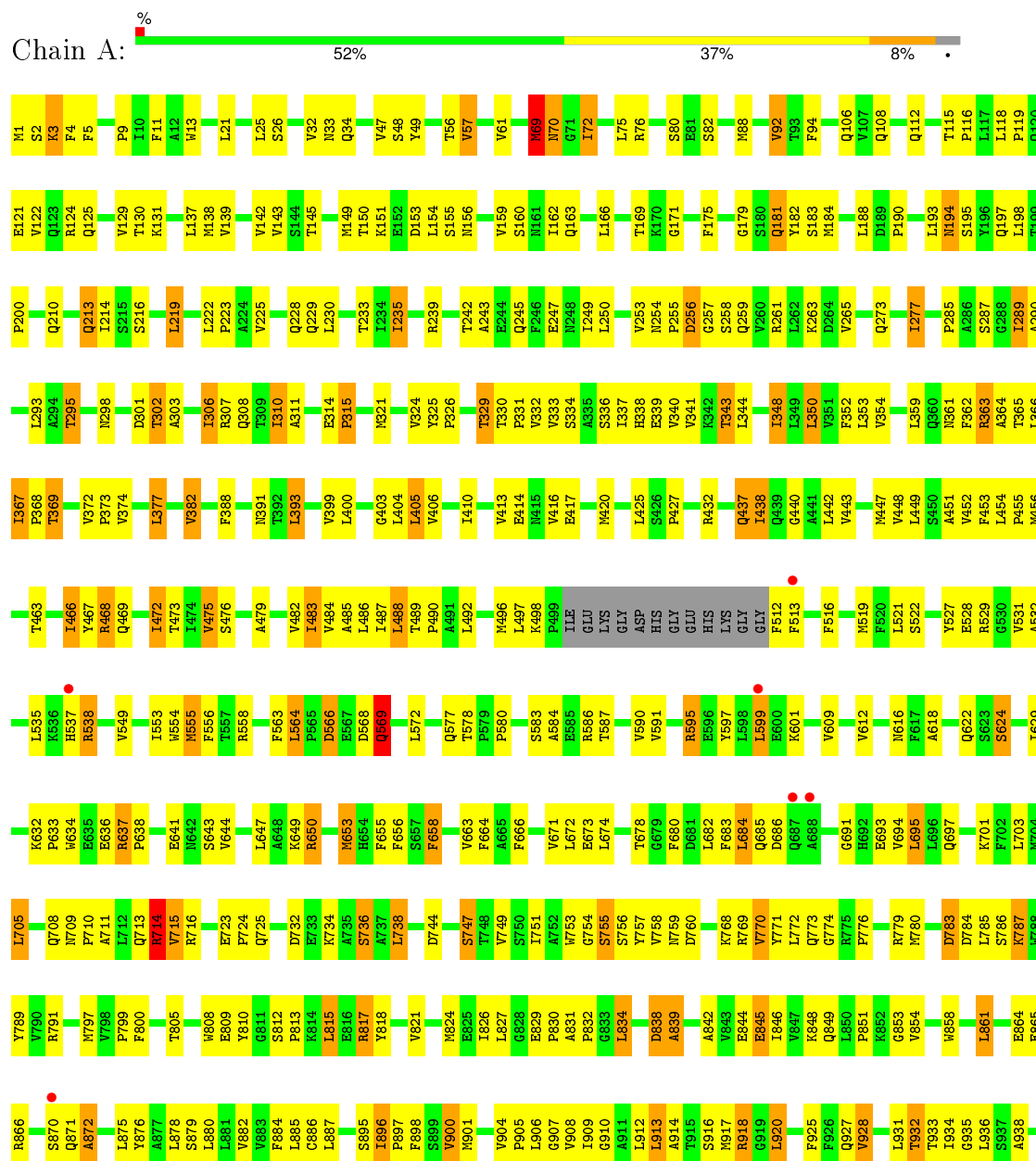


Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	B	1	Total	C	N	O	S	0	0
			49	31	11	6	1		
3	E	1	Total	C	N	O	S	0	0
			49	31	11	6	1		

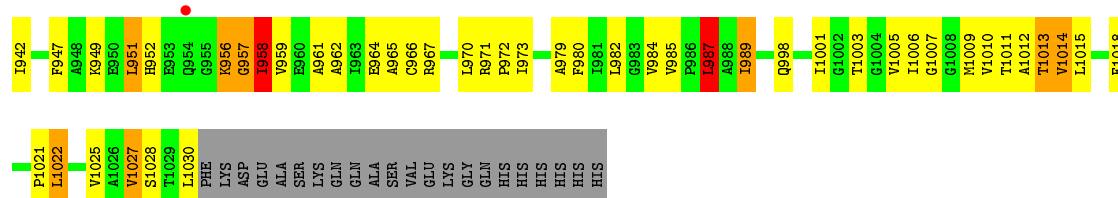
### 3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

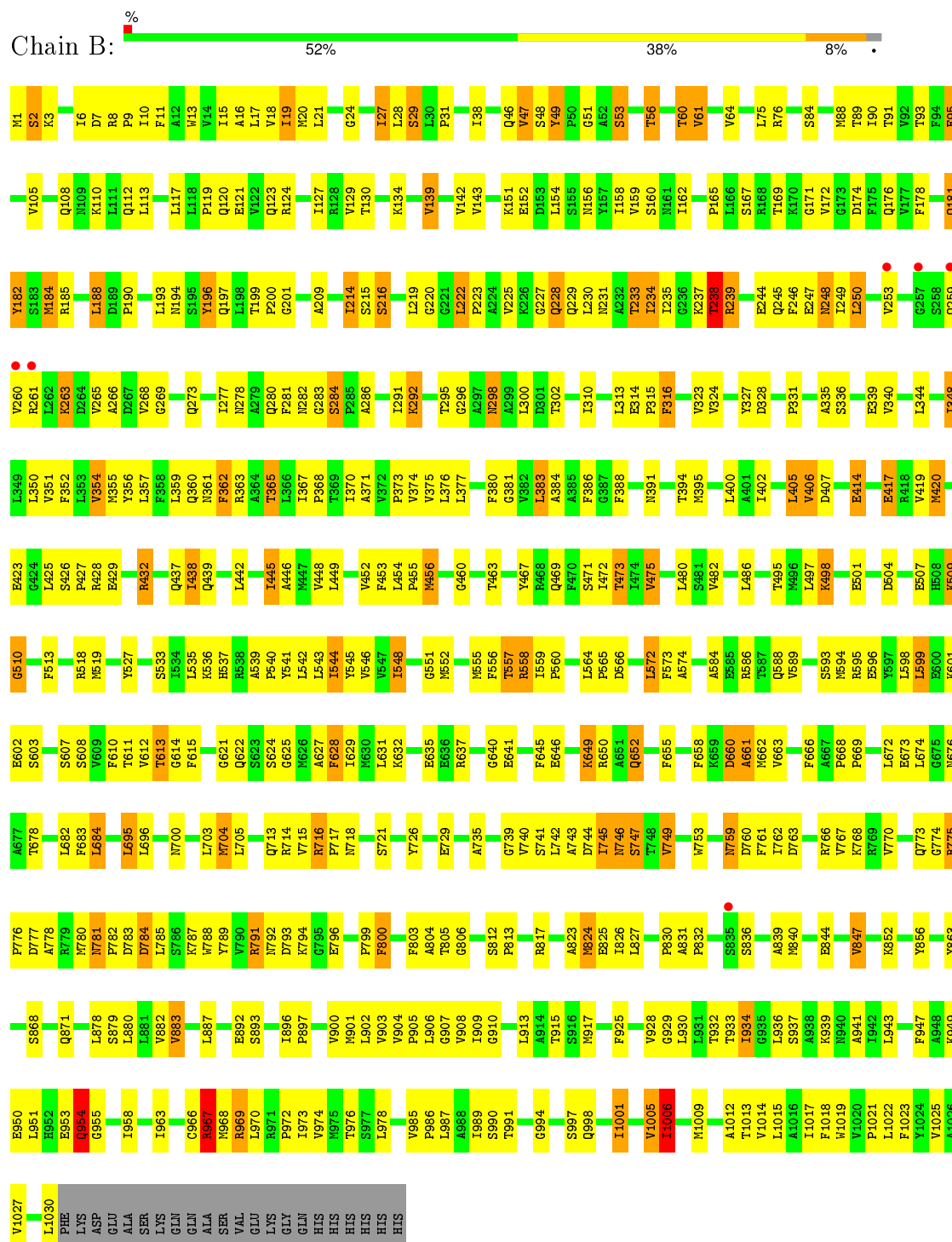
- Molecule 1: Multidrug resistance protein MexB



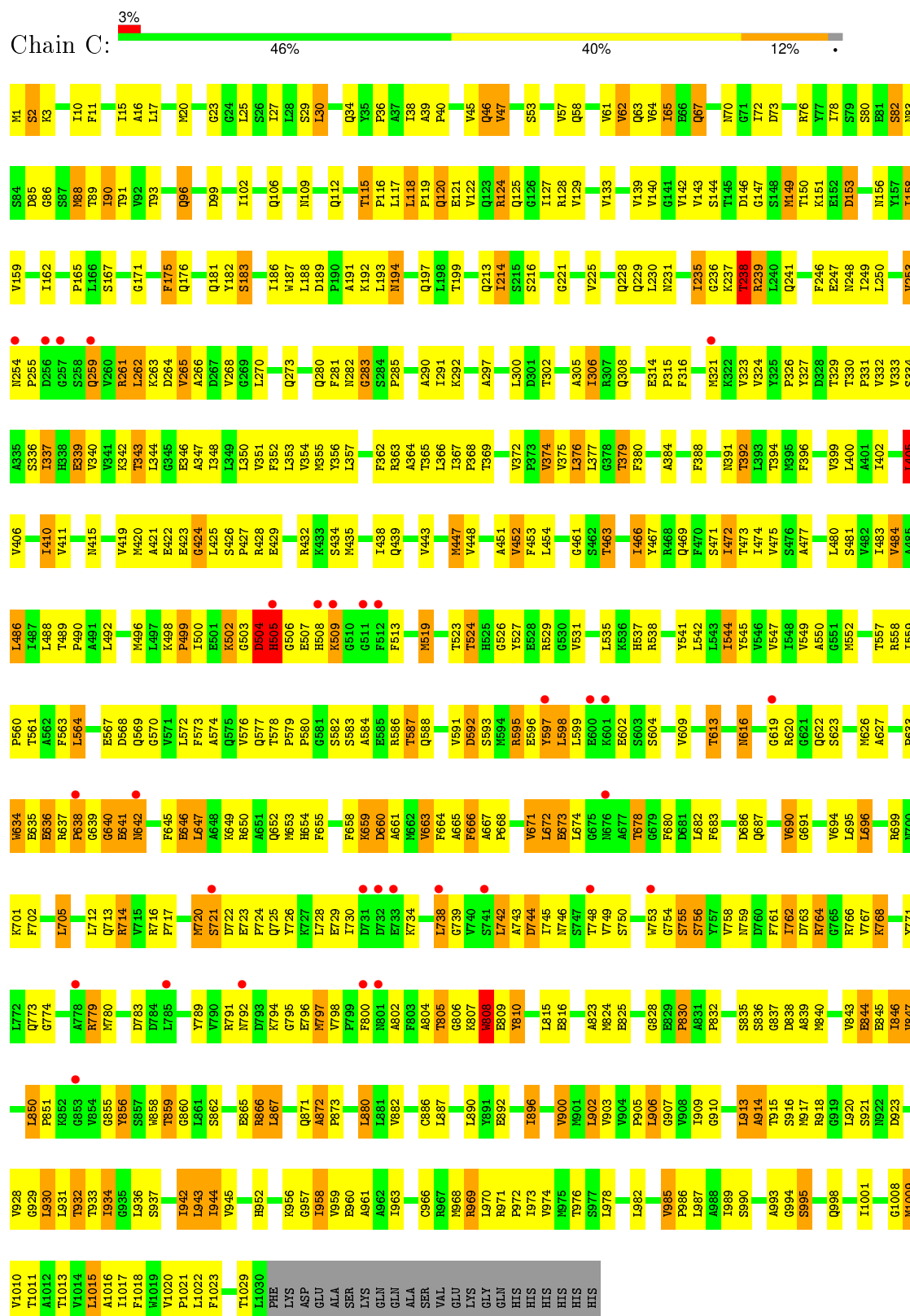




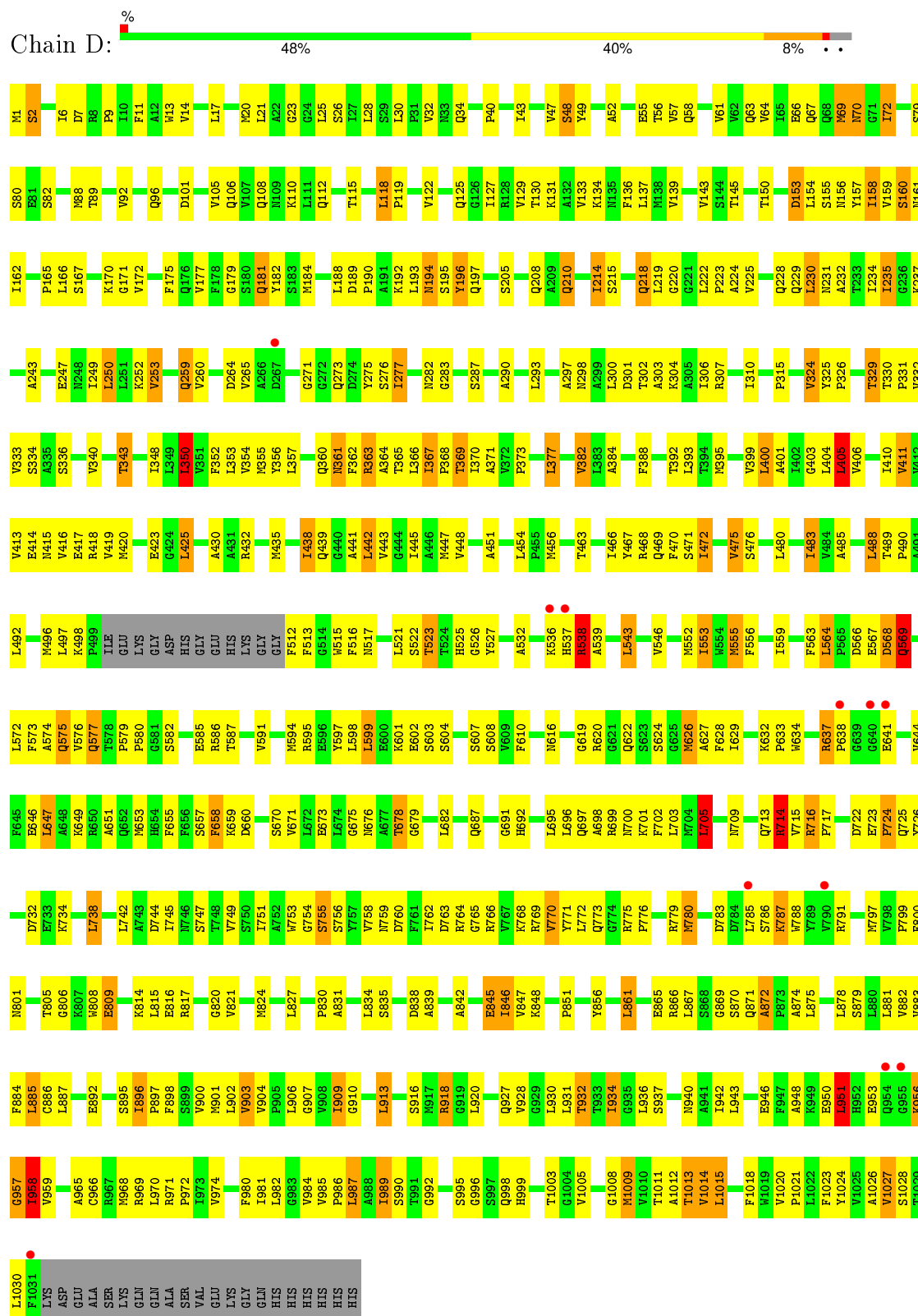
### • Molecule 1: Multidrug resistance protein MexB

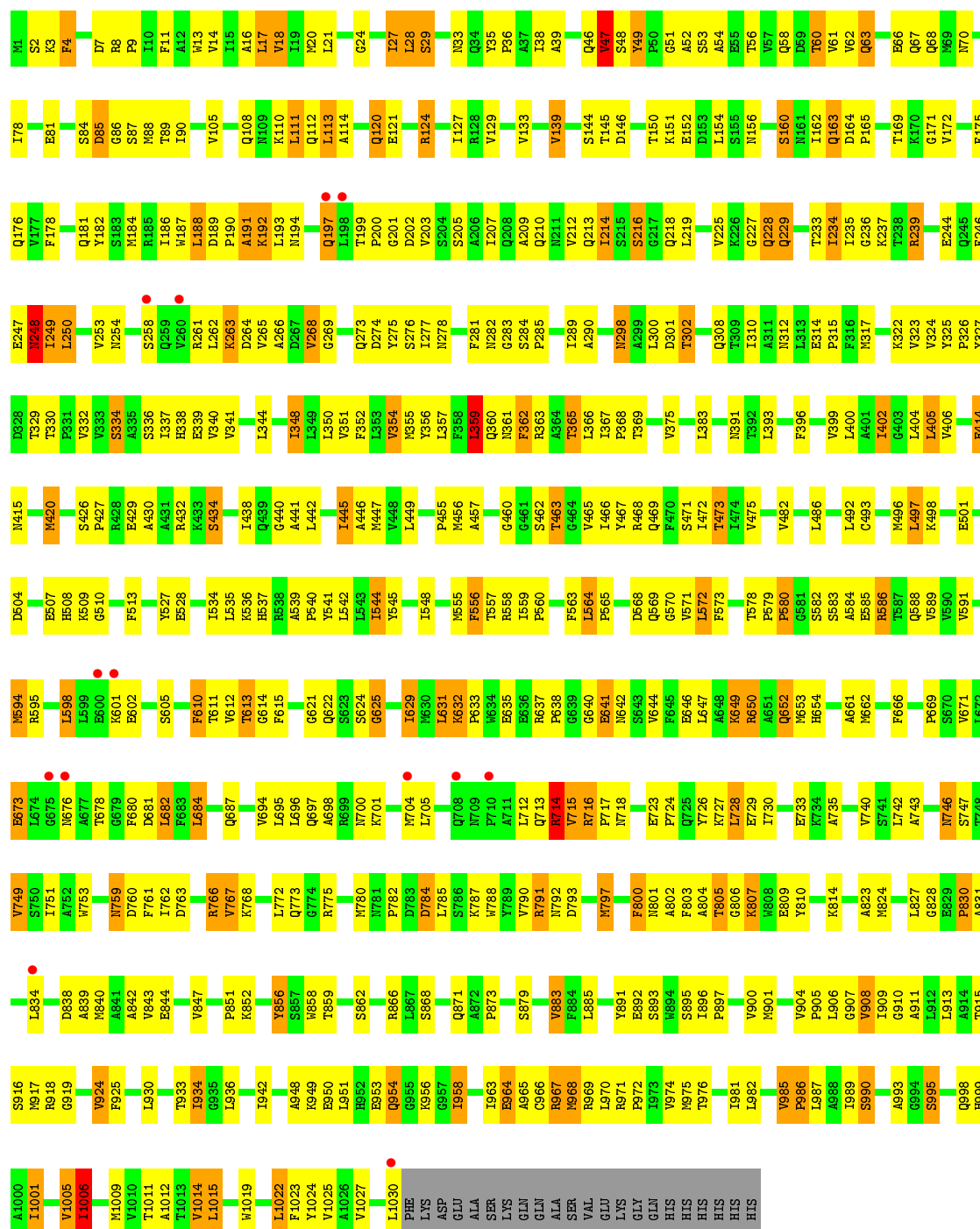


• Molecule 1: Multidrug resistance protein MexB

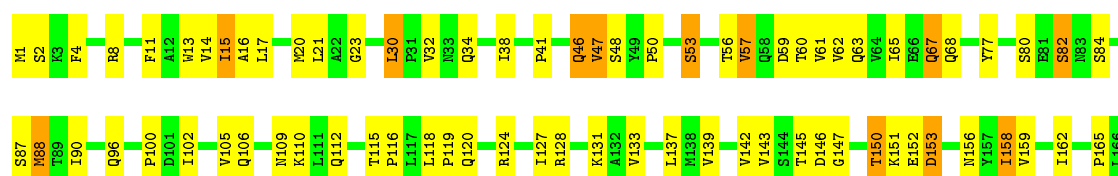


• Molecule 1: Multidrug resistance protein MexB





- Molecule 1: Multidrug resistance protein MexB



GLU	V959	V882	M797	R714	M634	I544	A479	G403	V333	P256	S167
LYS	I963	V883	V798		E635	I552	L400	L404	S334	G257	Q181
GLY		F884	F799	M720	E636	M552	V481	L405	A335	S258	Y182
HIS	C966	C886	F800	S721	R637		V482	D407	S336	Q259	
HIS			N801	D722	P638	M555	I483	D408	I337	V260	R185
HIS	P969	H894	A802	E723	G639		V484	A409	H338	R261	I186
HIS	L970	S895	F803	P724	G640	I559	A485	A409	E339	L282	W187
HIS	P971	I896	A804	Q725	E641	P560	L486	L410	V340	K263	L188
HIS	P972	P897	T805	Q726	N642	T561	I487	V416	V341	D264	D189
HIS	P973	F898	G806	K727	S643	A562	L488	E417	K342	V265	P190
		F898	K807	L728	V644	F563	T489	E417	T343	A266	
		S899	N808	L728	F645	L564	P490	R418	L344		A191
	L978	T900	E809	E729	E646		A491	V419		V267	K192
		P901	E809		L647	E567	L492	M420	A347	V268	L193
	I982	I902	Y810	D732	A648		C493		I348	G269	N194
		V903	P813	E733	R650	Q569	A494	E423		I277	S195
	V985	V904		L738			T495	G424			Y196
P986		P905	E816	Q739	L572	F573	M496	G425	V354	Q280	
L987		G907	V821	V740	E654	A574	L497	S426	M355	P281	T199
A988		V908	P822	S741	F655	Q575	K498	P427	T356	N282	
I989		A823	V822	L742	F656	A874	P499	R428			V203
S890		I909	A823	A743	S657	V576	I500	E429	F358	P285	S204
T991		G910	M824	D744	F658	Q577	E501	A430	L359		
G992		A911	E825	I745	E659	T578	K502		Q360	I289	
A993		I826	I826	M746	D660		G503	S434	M361		N211
		L912			E661	A584	D504	M435	F362	K292	V212
		L913			V663	E585	H505	G436	R363	L293	Q213
		A914				E586	G506	Q437	A294	A294	I214
G996						T587	E507	I438	T365	T295	S216
S997							H508	Q439	L366		
Q998		R918	S836	N759	F666	S593	K509	L442	I367	N298	L219
		G919	G837	D760	P669	M594			P368	I299	G220
I1006		L920	D838	F761	S670	R595	F513	M447	L300	G221	
G1007		S921	A839	I762	V671	E596	V516	V448	I370	D301	
G1008		D823	P840	R766		E597	M517	L449	A371	L222	
M1009		V924	V843	V767	L674	L598	R518		V372	T302	
V1010		F925	E844	K768		L599	M519	S450	P373	G227	
T1011		T926	E845		T678	E800		S451	P373	Q228	
		Q927	I846			K601		V452	V374	Q229	
	L1015	V928	V847			E602		F453	V375	L230	
		G929				S603			T309	N231	
F1018		I930	P851	P776	L682	S604	T523		I306	Q228	
W1019		L931	K852	G774	L683	S604	T524		R307		
V1020		T932			L684	S605	H525		Q308		
P1021		T932			L685	S606	T525		T309		
L1022		T933	G855	M780	L686	V606	G526		I310	A232	
F1023		I934	T856	M781	L687	S607	Y527		A311	T233	
Y1024		G935	S857	P782			E528		G378	T234	
V1025		L936							T379	I235	
					V690	F610	R529		A384	G236	
					G691		V531	T463		K237	
	L1030	I939	G860	L765	H692	F615	A532		F388	P316	
		N940	L861	L765	E693	N616	S533		R239	R317	
		A941		K787	V694	A618	I534		N391	Q319	
D1033		I942	R866		L695				G320	G320	
GLU		L943	L867		L696		R468		M321		F246
ALA		I944			L697		Q469		K322		E247
SER		I944			Q697		F470		T394		M248
LYS		V945	Q871		L698		S471		M395		L249
GLN			A872		A698		R538		V323		L250
GLN			N792		S623		T473		V324		L251
ALA			D793		S624		A539		Y325		K252
SER			K794		L631		P540		P326		V253
					L712		Y541		Y327		M254
					Q713		L542		V332		P255
							L543				

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	126.20 Å   137.03 Å   152.28 Å 85.75°   68.93°   87.39°	Depositor
Resolution (Å)	42.19 – 3.15 42.19 – 3.15	Depositor EDS
% Data completeness (in resolution range)	97.3 (42.19-3.15) 95.8 (42.19-3.15)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.47 (at 3.12 Å)	Xtriage
Refinement program	REFMAC 5.7.0029	Depositor
R, $R_{free}$	0.202   ,   0.279 0.202   ,   0.279	Depositor DCC
$R_{free}$ test set	7980 reflections (5.27%)	DCC
Wilson B-factor (Å <sup>2</sup> )	80.1	Xtriage
Anisotropy	0.052	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.28 , 58.7	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.52$ , $\langle L^2 \rangle = 0.36$	Xtriage
Outliers	0 of 159527 reflections	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	47358	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	83.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.64% 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

Bond lengths and bond angles in the following residue types are not validated in this section: P9D, 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	0.55	0/7880	0.81	4/10712 (0.0%)
1	B	0.54	0/7971	0.79	0/10833
1	C	0.51	0/7971	0.78	3/10833 (0.0%)
1	D	0.52	0/7892	0.79	6/10728 (0.1%)
1	E	0.53	0/7971	0.79	3/10833 (0.0%)
1	F	0.53	0/8000	0.80	6/10871 (0.1%)
All	All	0.53	0/47685	0.79	22/64810 (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
1	D	0	2
All	All	0	3

There are no bond length outliers.

The worst 5 of 22 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	405	LEU	CA-CB-CG	8.26	134.30	115.30
1	D	705	LEU	CA-CB-CG	7.54	132.64	115.30
1	E	359	LEU	CA-CB-CG	7.48	132.51	115.30
1	D	405	LEU	CA-CB-CG	7.38	132.28	115.30
1	A	393	LEU	CA-CB-CG	-7.24	98.65	115.30

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	951	LEU	Peptide
1	D	691	GLY	Peptide
1	D	951	LEU	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	7724	0	7864	375	0
1	B	7812	0	7944	395	0
1	C	7812	0	7944	484	0
1	D	7735	0	7873	417	0
1	E	7812	0	7944	421	0
1	F	7840	0	7970	439	0
2	A	105	0	138	6	0
2	B	105	0	138	5	0
2	C	35	0	46	6	0
2	D	105	0	138	5	0
2	E	105	0	138	10	0
2	F	70	0	92	7	0
3	B	49	0	39	8	0
3	E	49	0	39	8	0
All	All	47358	0	48307	2452	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 26.

The worst 5 of 2452 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:780:MET:CE	1:F:220:GLY:HA2	1.75	1.17
1:E:958:ILE:H	1:E:958:ILE:HD12	1.11	1.10
1:F:239:ARG:HG2	1:F:239:ARG:HH11	1.11	1.10
1:B:469:GLN:O	1:B:473:THR:HG22	1.52	1.10
1:C:239:ARG:HG2	1:C:239:ARG:HH11	0.96	1.09

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	1014/1052 (96%)	861 (85%)	122 (12%)	31 (3%)	5	32
1	B	1028/1052 (98%)	874 (85%)	127 (12%)	27 (3%)	7	38
1	C	1028/1052 (98%)	844 (82%)	129 (12%)	55 (5%)	2	17
1	D	1015/1052 (96%)	846 (83%)	130 (13%)	39 (4%)	4	26
1	E	1028/1052 (98%)	865 (84%)	127 (12%)	36 (4%)	4	29
1	F	1031/1052 (98%)	862 (84%)	125 (12%)	44 (4%)	3	23
All	All	6144/6312 (97%)	5152 (84%)	760 (12%)	232 (4%)	4	26

5 of 232 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	69	MET
1	A	70	ASN
1	A	498	LYS
1	A	673	GLU
1	A	714	ARG

### 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	833/860 (97%)	696 (84%)	137 (16%)	3	13
1	B	841/860 (98%)	703 (84%)	138 (16%)	3	13
1	C	841/860 (98%)	699 (83%)	142 (17%)	2	12

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	D	834/860 (97%)	696 (84%)	138 (16%)	3	13
1	E	841/860 (98%)	697 (83%)	144 (17%)	2	12
1	F	844/860 (98%)	693 (82%)	151 (18%)	2	11
All	All	5034/5160 (98%)	4184 (83%)	850 (17%)	2	12

5 of 850 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	783	ASP
1	D	411	VAL
1	F	604	SER
1	C	859	THR
1	D	88	MET

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 167 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	687	GLN
1	D	210	GLN
1	F	569	GLN
1	C	718	ASN
1	D	96	GLN

### 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 ⓘ

17 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	LMT	A	2001	-	36,36,36	0.72	1 (2%)	47,47,47	1.29	5 (10%)
2	LMT	A	2002	-	36,36,36	0.82	1 (2%)	47,47,47	1.75	16 (34%)
2	LMT	A	2003	-	36,36,36	0.66	1 (2%)	47,47,47	1.69	13 (27%)
3	P9D	B	2001	-	42,53,53	2.20	12 (28%)	50,77,77	2.37	17 (34%)
2	LMT	B	2002	-	36,36,36	0.75	1 (2%)	47,47,47	1.79	11 (23%)
2	LMT	B	2003	-	36,36,36	0.80	1 (2%)	47,47,47	1.62	10 (21%)
2	LMT	B	2004	-	36,36,36	0.47	0	47,47,47	1.06	3 (6%)
2	LMT	C	2001	-	36,36,36	0.69	1 (2%)	47,47,47	1.34	6 (12%)
2	LMT	D	2001	-	36,36,36	0.92	2 (5%)	47,47,47	1.47	7 (14%)
2	LMT	D	2002	-	36,36,36	0.65	0	47,47,47	1.35	5 (10%)
2	LMT	D	2003	-	36,36,36	0.78	1 (2%)	47,47,47	1.46	5 (10%)
3	P9D	E	2001	-	42,53,53	2.26	9 (21%)	50,77,77	1.71	7 (14%)
2	LMT	E	2002	-	36,36,36	0.94	1 (2%)	47,47,47	1.48	9 (19%)
2	LMT	E	2003	-	36,36,36	0.73	1 (2%)	47,47,47	1.52	7 (14%)
2	LMT	E	2004	-	36,36,36	0.66	1 (2%)	47,47,47	1.15	3 (6%)
2	LMT	F	2001	-	36,36,36	0.80	1 (2%)	47,47,47	2.02	14 (29%)
2	LMT	F	2002	-	36,36,36	0.82	1 (2%)	47,47,47	1.40	6 (12%)

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	LMT	A	2001	-	-	0/21/61/61	0/2/2/2
2	LMT	A	2002	-	-	0/21/61/61	0/2/2/2
2	LMT	A	2003	-	-	0/21/61/61	0/2/2/2
3	P9D	B	2001	-	-	0/33/49/49	0/5/5/5

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	LMT	B	2002	-	-	0/21/61/61	0/2/2/2
2	LMT	B	2003	-	-	0/21/61/61	0/2/2/2
2	LMT	B	2004	-	-	0/21/61/61	0/2/2/2
2	LMT	C	2001	-	-	0/21/61/61	0/2/2/2
2	LMT	D	2001	-	-	0/21/61/61	0/2/2/2
2	LMT	D	2002	-	-	0/21/61/61	0/2/2/2
2	LMT	D	2003	-	-	0/21/61/61	0/2/2/2
3	P9D	E	2001	-	-	0/33/49/49	1/5/5/5
2	LMT	E	2002	-	-	0/21/61/61	0/2/2/2
2	LMT	E	2003	-	-	0/21/61/61	0/2/2/2
2	LMT	E	2004	-	-	0/21/61/61	0/2/2/2
2	LMT	F	2001	-	-	0/21/61/61	0/2/2/2
2	LMT	F	2002	-	-	0/21/61/61	0/2/2/2

The worst 5 of 35 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	2001	P9D	C9-S8	-3.71	1.64	1.70
3	E	2001	P9D	C9-S8	-3.40	1.65	1.70
3	B	2001	P9D	C43-N42	-2.09	1.47	1.53
3	B	2001	P9D	C41-N42	-2.04	1.47	1.52
2	B	2002	LMT	O1'-C1'	2.05	1.43	1.40

The worst 5 of 144 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	2001	P9D	C9-C4-C2	-6.90	120.64	129.15
3	E	2001	P9D	C9-C4-C2	-5.85	121.94	129.15
2	F	2001	LMT	C3'-C4'-C5'	-4.99	99.55	110.84
3	B	2001	P9D	N31-N30-N29	-4.84	106.80	109.59
3	E	2001	P9D	C32-N24-C21	-4.36	105.62	120.30

There are no chirality outliers.

There are no torsion outliers.

All (1) ring outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	E	2001	P9D	C32-C33-C35-C36-C37-N24

16 monomers are involved in 55 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	2001	LMT	2	0
2	A	2002	LMT	2	0
2	A	2003	LMT	2	0
3	B	2001	P9D	8	0
2	B	2002	LMT	2	0
2	B	2003	LMT	2	0
2	B	2004	LMT	1	0
2	C	2001	LMT	6	0
2	D	2001	LMT	2	0
2	D	2002	LMT	1	0
2	D	2003	LMT	2	0
3	E	2001	P9D	8	0
2	E	2002	LMT	9	0
2	E	2003	LMT	1	0
2	F	2001	LMT	4	0
2	F	2002	LMT	3	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, 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	1018/1052 (96%)	-0.34	7 (0%)	89 82	39, 75, 112, 164	0
1	B	1030/1052 (97%)	-0.38	6 (0%)	90 84	40, 73, 113, 152	0
1	C	1030/1052 (97%)	-0.21	31 (3%)	54 37	43, 87, 143, 201	0
1	D	1019/1052 (96%)	-0.36	11 (1%)	82 71	47, 78, 117, 160	0
1	E	1030/1052 (97%)	-0.26	13 (1%)	79 67	46, 78, 124, 162	0
1	F	1033/1052 (98%)	-0.24	16 (1%)	76 62	45, 79, 129, 264	0
All	All	6160/6312 (97%)	-0.30	84 (1%)	78 64	39, 78, 124, 264	0

The worst 5 of 84 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	F	253	VAL	5.8
1	F	259	GLN	4.9
1	F	742	LEU	4.4
1	C	600	GLU	4.1
1	C	256	ASP	3.9

### 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. 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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
2	LMT	D	2001	35/35	0.79	0.45	5.18	47,65,79,89	35
2	LMT	F	2002	35/35	0.85	0.34	4.69	76,118,147,148	0
2	LMT	E	2003	35/35	0.90	0.38	4.17	67,94,108,124	0
2	LMT	B	2003	35/35	0.93	0.28	3.34	71,82,104,111	0
2	LMT	E	2004	35/35	0.95	0.32	3.15	67,74,100,103	0
2	LMT	B	2004	35/35	0.94	0.23	2.90	62,69,78,84	0
2	LMT	A	2001	35/35	0.86	0.34	2.79	43,66,94,103	35
3	P9D	E	2001	49/49	0.92	0.27	2.77	60,78,108,123	0
2	LMT	C	2001	35/35	0.88	0.34	2.62	80,105,146,151	0
2	LMT	A	2002	35/35	0.91	0.26	2.61	60,86,132,142	0
2	LMT	D	2002	35/35	0.93	0.31	2.57	69,111,137,144	0
2	LMT	B	2002	35/35	0.90	0.23	2.28	63,106,135,137	0
2	LMT	F	2001	35/35	0.84	0.28	1.95	82,104,126,139	0
3	P9D	B	2001	49/49	0.94	0.24	1.86	53,71,85,108	0
2	LMT	E	2002	35/35	0.90	0.25	1.63	67,106,124,132	0
2	LMT	D	2003	35/35	0.89	0.25	0.91	70,98,134,144	0
2	LMT	A	2003	35/35	0.91	0.23	0.90	64,90,115,120	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.