wwPDB X-ray Structure Validation Summary Report (i)

Sep 11, 2023 – 05:48 PM EDT

PDB ID	:	4LSG
Title	:	Structure of Mouse P-Glycoprotein
Authors	:	Chang, G.; Szewczyk, P.
Deposited on	:	2013-07-22
Resolution	:	3.80 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 (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references (i)) were used in the production of this report:

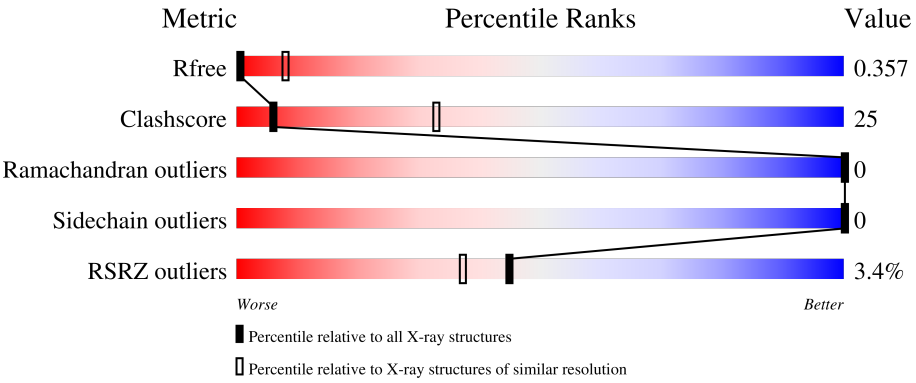
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35.1

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 3.80 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}	130704	1212 (4.00-3.60)
Clashscore	141614	1288 (4.00-3.60)
Ramachandran outliers	138981	1243 (4.00-3.60)
Sidechain outliers	138945	1237 (4.00-3.60)
RSRZ outliers	127900	1121 (4.00-3.60)

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 of 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 <=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	1284	4% 54% 38% 8%
1	B	1284	3% 55% 37% 8%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 18352 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 1A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	1182	Total	C	N	O	S	0	0	0
			9170	5895	1552	1686	37			
1	B	1182	Total	C	N	O	S	0	0	0
			9170	5895	1552	1686	37			

There are 18 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	952	ALA	CYS	engineered mutation	UNP P21447
A	1277	LEU	-	expression tag	UNP P21447
A	1278	GLU	-	expression tag	UNP P21447
A	1279	HIS	-	expression tag	UNP P21447
A	1280	HIS	-	expression tag	UNP P21447
A	1281	HIS	-	expression tag	UNP P21447
A	1282	HIS	-	expression tag	UNP P21447
A	1283	HIS	-	expression tag	UNP P21447
A	1284	HIS	-	expression tag	UNP P21447
B	952	ALA	CYS	engineered mutation	UNP P21447
B	1277	LEU	-	expression tag	UNP P21447
B	1278	GLU	-	expression tag	UNP P21447
B	1279	HIS	-	expression tag	UNP P21447
B	1280	HIS	-	expression tag	UNP P21447
B	1281	HIS	-	expression tag	UNP P21447
B	1282	HIS	-	expression tag	UNP P21447
B	1283	HIS	-	expression tag	UNP P21447
B	1284	HIS	-	expression tag	UNP P21447

- Molecule 2 is MERCURY (II) ION (three-letter code: HG) (formula: Hg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	6	Total	Hg	0	0
			6	6		

Continued on next page...

Continued from previous page...

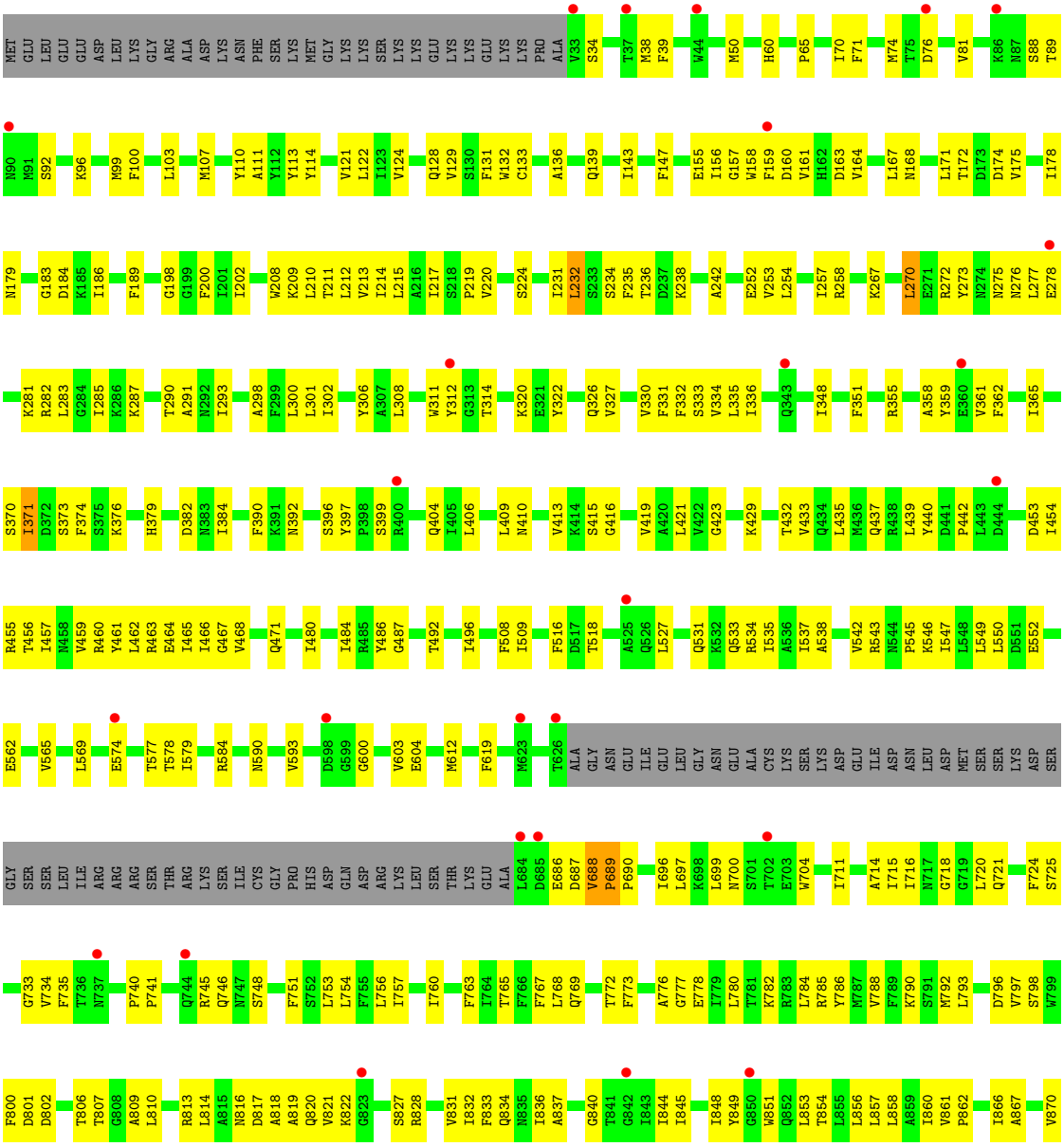
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Hg		
2	B	6	6	6	0	0

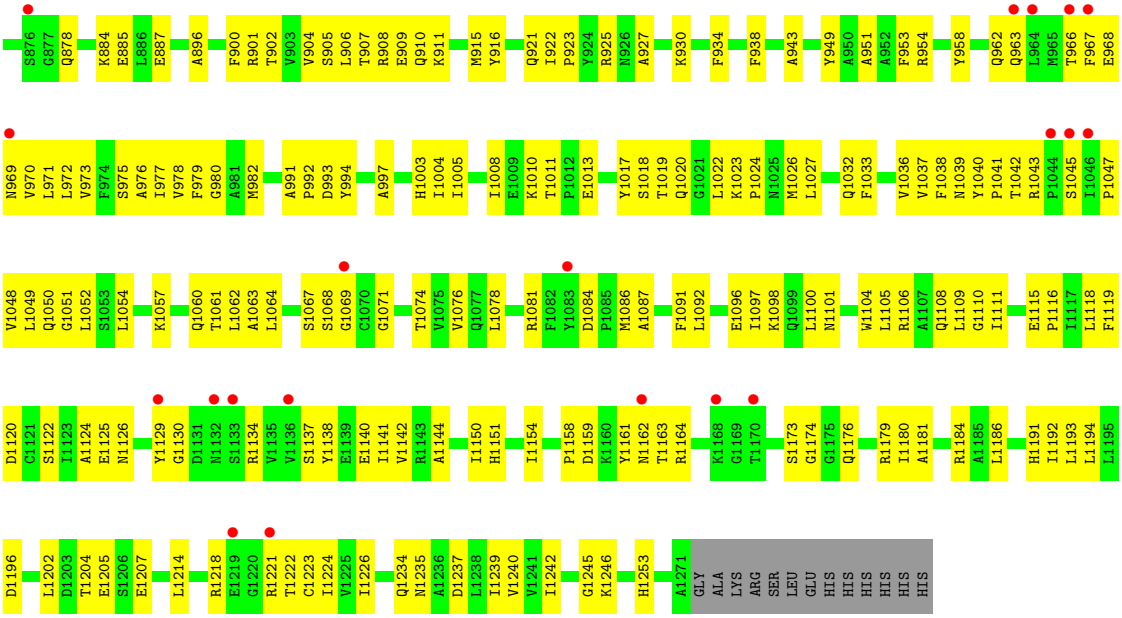
3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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: Multidrug resistance protein 1A

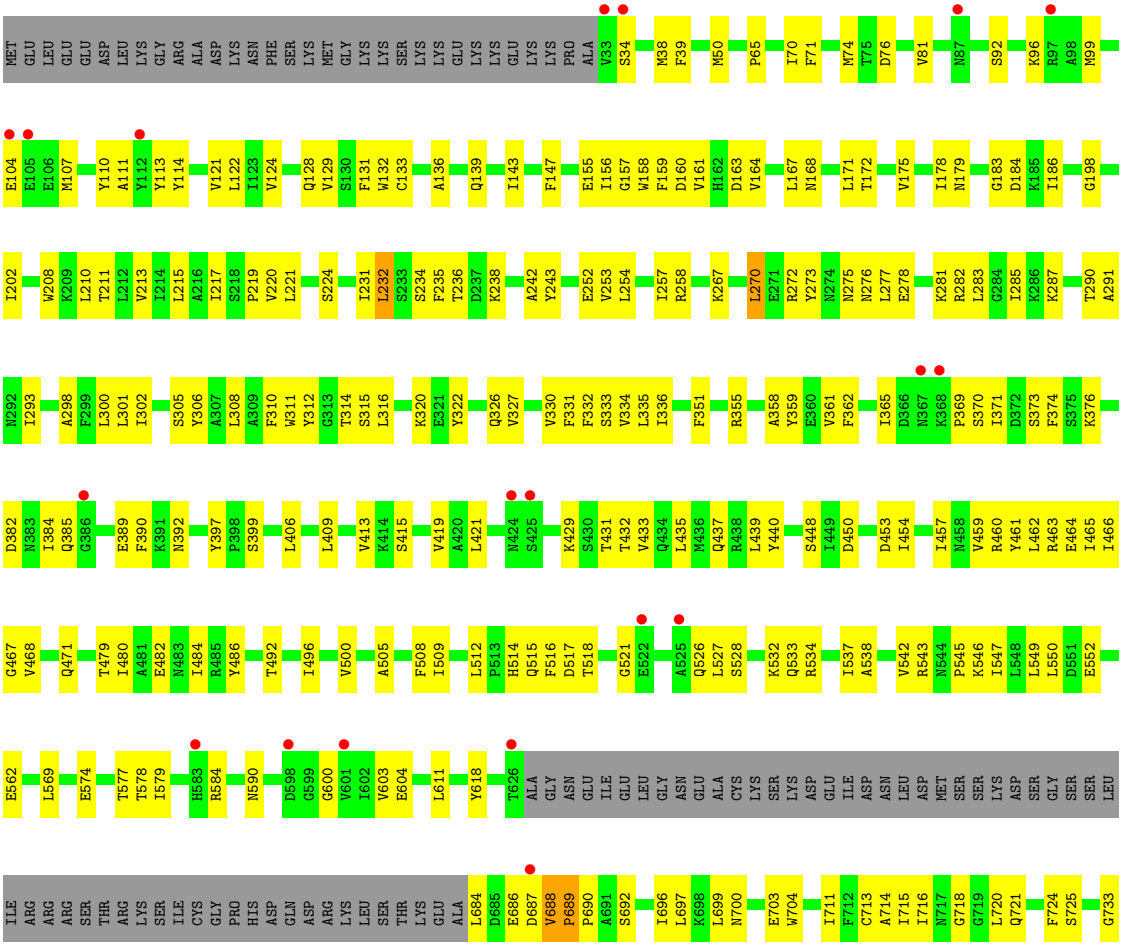
Chain A:

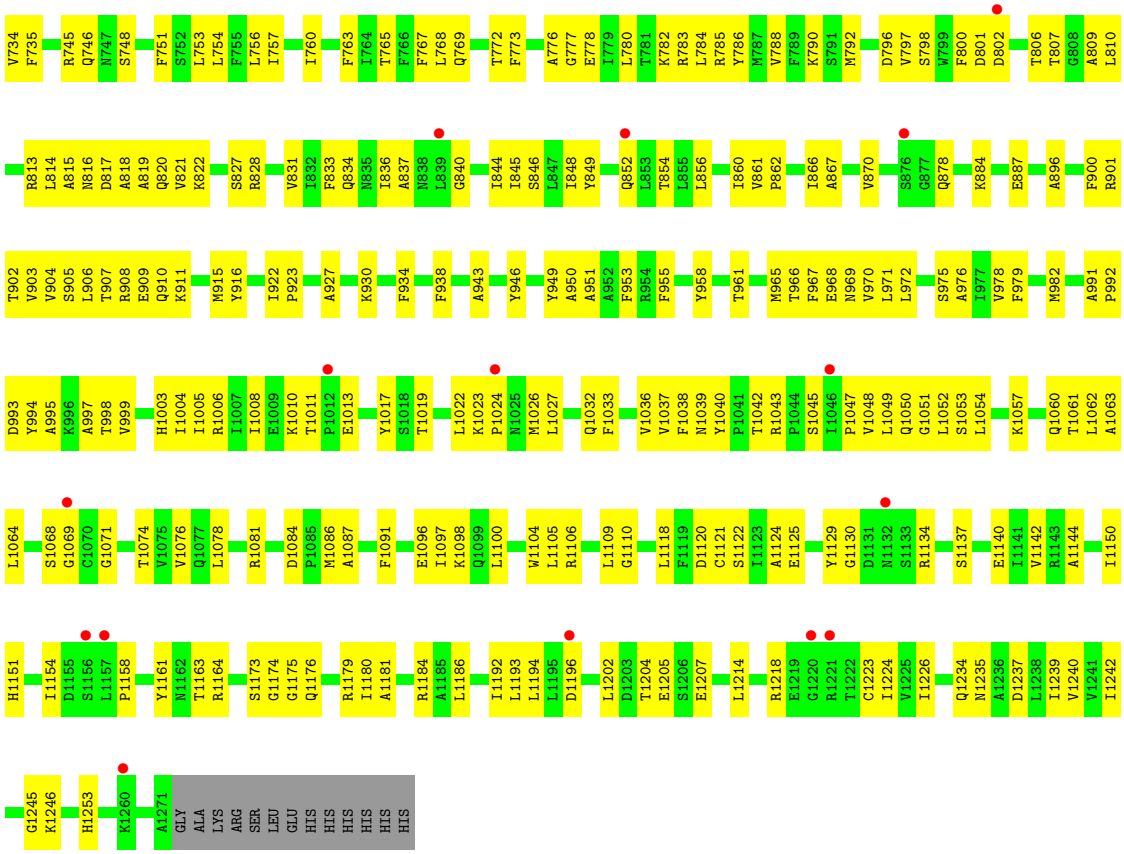




• Molecule 1: Multidrug resistance protein 1A

Chain B:




V734 F735 R745 Q746 N747 S748 F751 S752 L753 L754 F755 L756 I757 I760 F763 I764 T765 F766 F767 L768 Q769 T772 F773 A776 G777 E778 I779 L780 T781 K782 R783 L784 R785 Y786 M787 V788 F789 K790 S791 M792 D796 V797 S798 W799 F800 D801 D802 T806 T807 G808 A809 L810
R813 L814 A815 N816 D817 A818 A819 Q820 V821 K822 S827 R828 V831 I832 F833 Q834 N835 I836 A837 N838 L839 G840 I844 I845 S846 L847 I848 Y849 Q852 L853 T854 L855 L856 I860 V861 P862 I866 A867 V870 S876 G877 Q878 K884 E887 A896 F900 R901
T902 V903 V904 S905 L906 T907 R908 E909 Q910 K911 M915 Y916 I922 P923 A927 K930 F934 F938 A943 Y946 Y949 A950 A951 A952 F953 R954 F955 Y958 T961 M965 T966 F967 E968 N969 V970 L971 L972 S975 A976 I977 V978 F979 M982 A991 P992
D993 Y994 A995 K996 A997 T998 V999 H1003 I1004 I1005 R1006 I1007 I1008 E1009 K1010 T1011 P1012 E1013 Y1017 S1018 T1019 L1022 K1023 P1024 N1025 M1026 L1027 Q1032 F1033 V1036 V1037 F1038 N1039 Y1040 P1041 T1042 R1043 P1044 S1045 I1046 P1047 V1048 L1049 Q1050 G1051 L1052 S1053 L1054 K1057 Q1060 T1061 L1062 A1063
L1064 S1068 G1069 C1070 G1071 T1074 V1075 V1076 Q1077 L1078 R1081 D1084 P1085 M1086 A1087 F1091 E1096 I1097 K1098 Q1099 L1100 W1104 L1105 R1106 L1109 G1110 L1118 F1119 D1120 C1121 S1122 I1123 A1124 E1125 Y1129 G1130 D1131 N1132 S1133 R1134 S1137 E1140 I1141 V1142 R1143 A1144 I1150
H1151 I1154 D1155 S1156 L1157 P1158 Y1161 N1162 T1163 R1164 S1173 G1174 G1175 Q1176 R1179 I1180 A1181 R1184 A1185 L1186 I1192 L1193 L1194 L1195 D1196 L1202 D1203 T1204 E1205 S1206 E1207 L1214 R1218 E1219 G1220 R1221 T1222 C1223 I1224 V1225 I1226 Q1234 N1235 A1236 D1237 L1238 I1239 V1240 V1241 I1242
G1245 K1246 H1253 K1260 A1271 GLY ALA LYS ARG SER LEU GLU HIS HIS HIS HIS HIS

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	97.54Å 115.43Å 378.86Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	19.98 – 3.80 19.98 – 3.80	Depositor EDS
% Data completeness (in resolution range)	96.0 (19.98-3.80) 96.1 (19.98-3.80)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	11.69 (at 3.82Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8.2_1309)	Depositor
R, R_{free}	0.336 , 0.357 0.336 , 0.357	Depositor DCC
R_{free} test set	4203 reflections (10.22%)	wwPDB-VP
Wilson B-factor (Å^2)	133.9	Xtriage
Anisotropy	0.338	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.22 , 19.3	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.87	EDS
Total number of atoms	18352	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 2.39% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ 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

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

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	0.35	1/9338 (0.0%)	0.64	4/12625 (0.0%)
1	B	0.34	2/9338 (0.0%)	0.63	3/12625 (0.0%)
All	All	0.35	3/18676 (0.0%)	0.64	7/25250 (0.0%)

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	689	PRO	N-CD	5.25	1.55	1.47
1	A	689	PRO	N-CD	5.22	1.55	1.47
1	B	1121	CYS	CB-SG	5.10	1.91	1.82

The worst 5 of 7 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	688	VAL	C-N-CD	5.65	140.26	128.40
1	A	688	VAL	C-N-CD	5.55	140.06	128.40
1	B	270	LEU	CA-CB-CG	5.44	127.82	115.30
1	B	232	LEU	CA-CB-CG	-5.42	102.83	115.30
1	A	270	LEU	CA-CB-CG	5.41	127.75	115.30

There are no chirality outliers.

There are no planarity outliers.

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	9170	0	9338	471	0
1	B	9170	0	9338	451	0
2	A	6	0	0	0	0
2	B	6	0	0	0	0
All	All	18352	0	18676	916	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 25.

The worst 5 of 916 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:B:132:TRP:CD1	1:B:183:GLY:HA3	1.53	1.42
1:A:1019:THR:CG2	1:A:1101:ASN:HA	1.50	1.41
1:A:132:TRP:CD1	1:A:183:GLY:HA3	1.59	1.36
1:B:1017:TYR:CE2	1:B:1019:THR:HG23	1.65	1.30
1:B:133:CYS:SG	1:B:934:PHE:HD2	1.63	1.21

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	1178/1284 (92%)	1069 (91%)	109 (9%)	0	100	100
1	B	1178/1284 (92%)	1069 (91%)	109 (9%)	0	100	100
All	All	2356/2568 (92%)	2138 (91%)	218 (9%)	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	975/1064 (92%)	975 (100%)	0	100	100
1	B	975/1064 (92%)	975 (100%)	0	100	100
All	All	1950/2128 (92%)	1950 (100%)	0	100	100

There are no protein residues with a non-rotameric sidechain to report.

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 10 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	1060	GLN
1	B	1108	GLN
1	B	1253	HIS
1	A	1108	GLN
1	A	1253	HIS

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 monosaccharides in this entry.

5.6 Ligand geometry

Of 12 ligands modelled in this entry, 12 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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, 95^{th} 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	1182/1284 (92%)	-0.06	46 (3%)	39	32	48, 98, 147, 186	0
1	B	1182/1284 (92%)	-0.05	34 (2%)	51	42	50, 103, 148, 189	0
All	All	2364/2568 (92%)	-0.05	80 (3%)	45	37	48, 100, 147, 189	0

The worst 5 of 80 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	33	VAL	5.4
1	A	1044	PRO	5.0
1	A	963	GLN	4.5
1	B	598	ASP	4.4
1	A	964	LEU	4.4

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 monosaccharides 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. The B-factors column lists the minimum, median, 95^{th} 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(Å^2)	Q<0.9
2	HG	B	1302	1/1	0.62	0.31	109,109,109,109	1
2	HG	A	1302	1/1	0.87	0.39	166,166,166,166	1
2	HG	B	1304	1/1	0.87	0.29	146,146,146,146	0
2	HG	B	1301	1/1	0.89	0.08	146,146,146,146	0
2	HG	A	1306	1/1	0.90	0.14	146,146,146,146	0
2	HG	A	1301	1/1	0.92	0.13	146,146,146,146	0
2	HG	A	1304	1/1	0.93	0.22	146,146,146,146	0
2	HG	A	1305	1/1	0.93	0.06	146,146,146,146	0
2	HG	A	1303	1/1	0.94	0.10	146,146,146,146	0
2	HG	B	1306	1/1	0.95	0.11	146,146,146,146	0
2	HG	B	1303	1/1	0.96	0.07	146,146,146,146	0
2	HG	B	1305	1/1	0.97	0.08	146,146,146,146	0

6.5 Other polymers (i)

There are no such residues in this entry.