wwPDB X-ray Structure Validation Summary Report (i)

Oct 10, 2021 – 06:12 PM EDT

PDB ID	:	3G5U
Title	:	Structure of P-glycoprotein Reveals a Molecular Basis for Poly-Specific Drug Binding
Authors	:	Aller, S.G.; Yu, J.; Ward, A.; Weng, Y.; Chittaboina, S.; Zhuo, R.; Harrell, P.M.; Trinh, Y.T.; Zhang, Q.; Urbatsch, I.L.; Chang, G.
Deposited on	:	2009-02-05
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 following versions of software and data (see references (1)) were used in the production of this report:

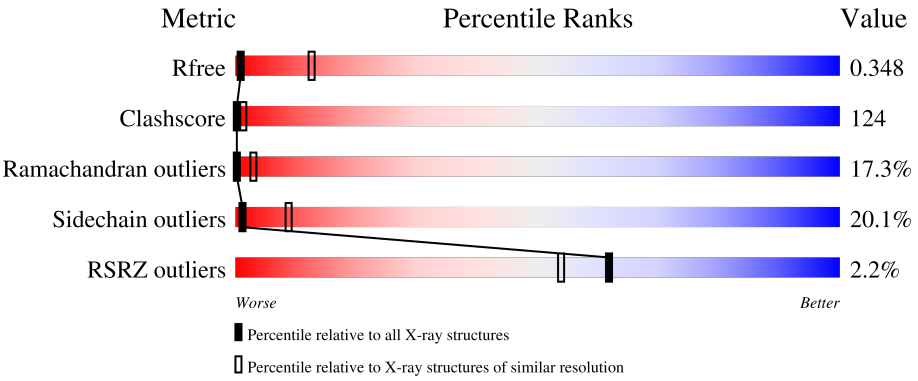
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.23.2
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.23.2

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	2% 11% 53% 24% · 8%
1	B	1284	2% 11% 55% 23% · 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 Q5I1Y5
A	1277	TYR	-	expression tag	UNP Q5I1Y5
A	1278	VAL	-	expression tag	UNP Q5I1Y5
A	1279	HIS	-	expression tag	UNP Q5I1Y5
A	1280	HIS	-	expression tag	UNP Q5I1Y5
A	1281	HIS	-	expression tag	UNP Q5I1Y5
A	1282	HIS	-	expression tag	UNP Q5I1Y5
A	1283	HIS	-	expression tag	UNP Q5I1Y5
A	1284	HIS	-	expression tag	UNP Q5I1Y5
B	952	ALA	CYS	engineered mutation	UNP Q5I1Y5
B	1277	TYR	-	expression tag	UNP Q5I1Y5
B	1278	VAL	-	expression tag	UNP Q5I1Y5
B	1279	HIS	-	expression tag	UNP Q5I1Y5
B	1280	HIS	-	expression tag	UNP Q5I1Y5
B	1281	HIS	-	expression tag	UNP Q5I1Y5
B	1282	HIS	-	expression tag	UNP Q5I1Y5
B	1283	HIS	-	expression tag	UNP Q5I1Y5
B	1284	HIS	-	expression tag	UNP Q5I1Y5

- 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...

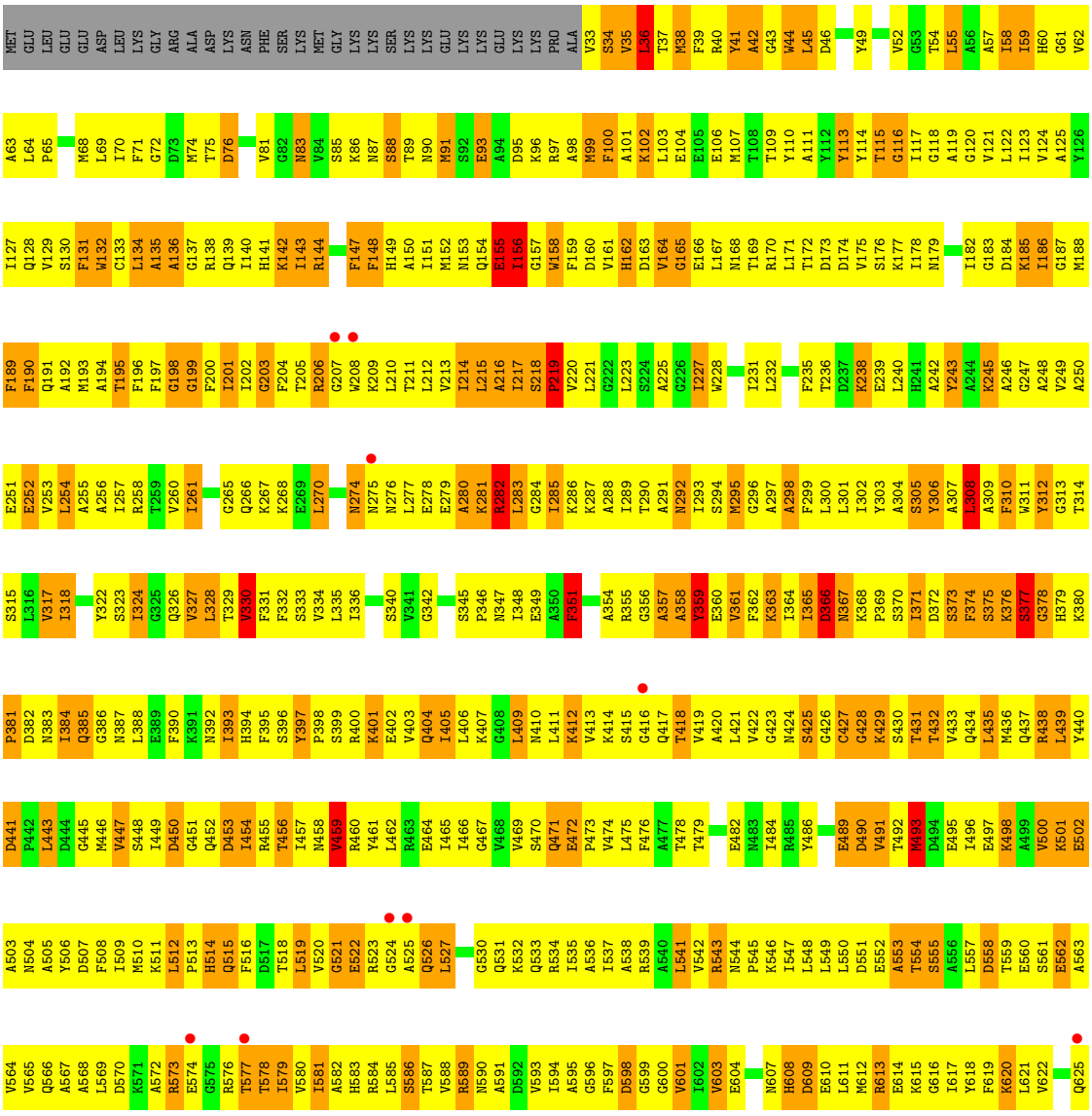
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Hg		
2	B	6	6	6	0	0

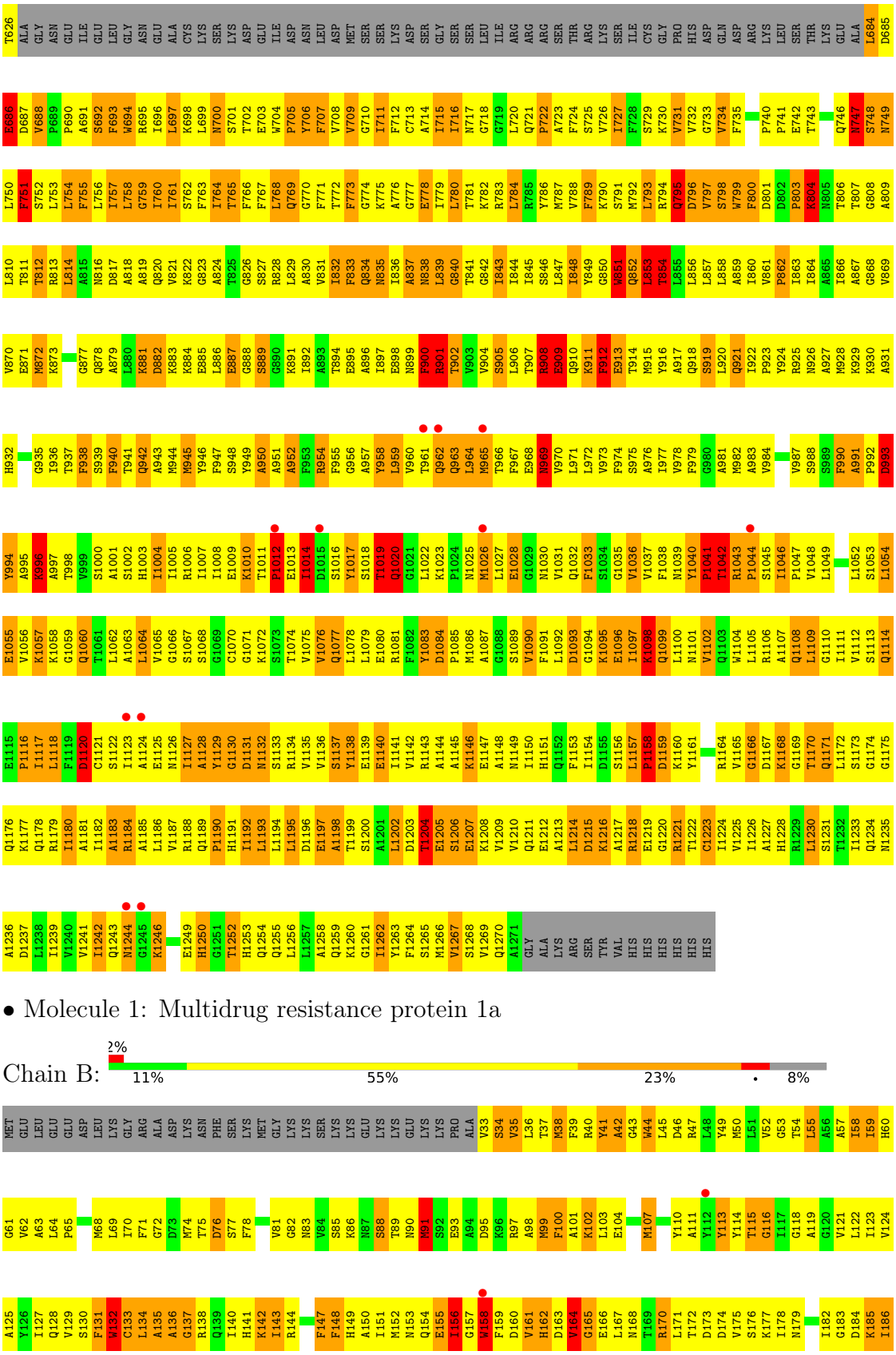
3 Residue-property plots (i)

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

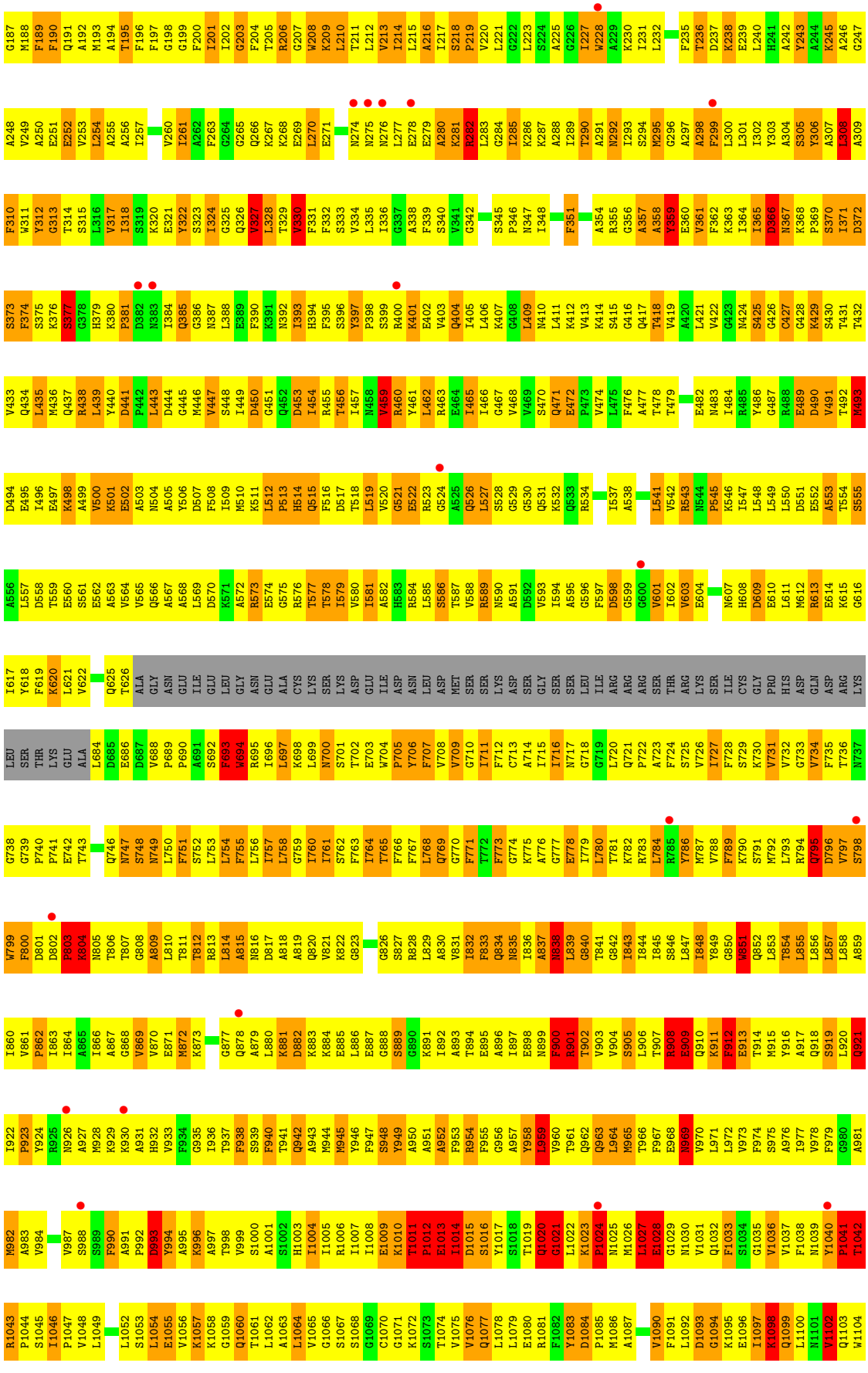
Chain A:

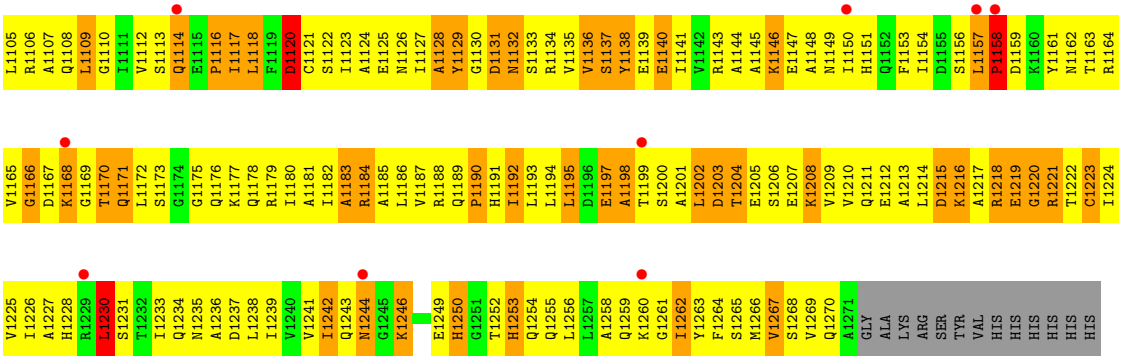




• Molecule 1: Multidrug resistance protein 1a

Chain B:




L1105 R1106 A1107 Q1108 L1109 G1110 I1111 V1112 S1113 Q1114 E1115 P1116 I1117 L1118 F1119 D1120 C1121 S1122 I1123 A1124 E1125 N1126 I1127 A1128 Y1129 G1130 D1131 N1132 S1133 R1134 V1135 V1136 S1137 Y1138 E1139 E1140 I1141 V1142 R1143 A1144 A1145 K1146 E1147 A1148 N1149 I1150 H1151 Q1152 F1153 I1154 D1155 S1156 L1157 P1158 D1159 K1160 Y1161 N1162 T1163 R1164
V1165 G1166 D1167 K1168 G1169 T1170 Q1171 L1172 S1173 G1174 G1175 Q1176 K1177 Q1178 R1179 I1180 A1181 I1182 A1183 R1184 A1185 L1186 V1187 R1188 Q1189 P1190 H1191 I1192 L1193 L1194 L1195 D1196 E1197 A1198 T1199 S1200 A1201 L1202 D1203 T1204 E1205 S1206 E1207 K1208 V1209 V1210 Q1211 E1212 A1213 L1214 D1215 K1216 A1217 R1218 E1219 G1220 R1221 T1222 C1223 I1224
V1225 I1226 A1227 H1228 R1229 L1230 S1231 T1232 I1233 Q1234 N1235 A1236 D1237 L1238 I1239 V1240 V1241 I1242 Q1243 N1244 G1245 K1246 E1249 H1250 G1251 T1252 H1253 Q1254 Q1255 L1256 L1257 A1258 Q1259 K1260 G1261 I1262 Y1263 F1264 S1265 M1266 V1267 S1268 V1269 Q1270 A1271 GLY ALA LYS ARG SER TYR VAL HIS 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 42.33 – 3.79	Depositor EDS
% Data completeness (in resolution range)	96.1 (19.98-3.80) 95.5 (42.33-3.79)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.08	Depositor
$< I/\sigma(I) >$ [1]	4.19 (at 3.76Å)	Xtriage
Refinement program	CNS 1.2	Depositor
R, R_{free}	0.306 , 0.347 0.309 , 0.348	Depositor DCC
R_{free} test set	4245 reflections (10.21%)	wwPDB-VP
Wilson B-factor (Å^2)	132.2	Xtriage
Anisotropy	0.330	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.20 , 113.6	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	0.86	EDS
Total number of atoms	18352	wwPDB-VP
Average B, all atoms (Å^2)	136.0	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	0.54	1/9338 (0.0%)	0.88	22/12625 (0.2%)
1	B	0.51	0/9338	0.86	25/12625 (0.2%)
All	All	0.53	1/18676 (0.0%)	0.87	47/25250 (0.2%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	600	GLY	CA-C	6.60	1.62	1.51

The worst 5 of 47 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	165	GLY	N-CA-C	-9.98	88.15	113.10
1	A	1097	ILE	N-CA-C	-8.94	86.87	111.00
1	A	994	TYR	N-CA-C	-8.30	88.59	111.00
1	A	1098	LYS	N-CA-C	-8.12	89.09	111.00
1	B	693	PHE	N-CA-C	7.98	132.54	111.00

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	2363	1
1	B	9170	0	9337	2245	1
2	A	6	0	0	0	0
2	B	6	0	0	0	0
All	All	18352	0	18675	4595	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 124.

The worst 5 of 4595 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:830:ALA:HB2	1:B:990:PHE:CE2	1.25	1.61
1:B:830:ALA:CB	1:B:990:PHE:CE2	2.06	1.36
1:B:830:ALA:CB	1:B:990:PHE:HE2	1.38	1.36
1:B:263:PHE:HE2	1:B:266:GLN:NE2	1.32	1.27
1:A:856:LEU:HD13	1:A:955:PHE:CD1	1.70	1.25

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)
1:A:1099:GLN:CG	1:B:450:ASP:OD1[1_455]	2.03	0.17

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%)	678 (58%)	299 (25%)	201 (17%)	0 3
1	B	1178/1284 (92%)	676 (57%)	295 (25%)	207 (18%)	0 2
All	All	2356/2568 (92%)	1354 (58%)	594 (25%)	408 (17%)	0 3

5 of 408 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	34	SER
1	A	115	THR
1	A	135	ALA
1	A	156	ILE
1	A	164	VAL

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	975/1064 (92%)	775 (80%)	200 (20%)	1 8
1	B	975/1064 (92%)	783 (80%)	192 (20%)	1 9
All	All	1950/2128 (92%)	1558 (80%)	392 (20%)	1 9

5 of 392 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	206	ARG
1	B	577	THR
1	B	245	LYS
1	B	401	LYS
1	B	751	PHE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 82 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	437	GLN
1	B	963	GLN
1	B	625	GLN
1	B	820	GLN
1	B	1099	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 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

There are no such residues in this entry.

5.8 Polymer linkage issues

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%)	-0.31	20 (1%)	70	62	36, 127, 195, 207	0
1	B	1182/1284 (92%)	-0.23	31 (2%)	56	47	47, 141, 200, 207	0
All	All	2364/2568 (92%)	-0.27	51 (2%)	62	54	36, 135, 198, 207	0

The worst 5 of 51 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	275	ASN	4.5
1	A	1244	ASN	4.3
1	B	382	ASP	4.3
1	A	961	THR	4.2
1	B	383	ASN	4.1

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	A	1286	1/1	0.58	0.40	166,166,166,166	1
2	HG	B	1288	1/1	0.90	0.18	147,147,147,147	0
2	HG	A	1288	1/1	0.94	0.11	147,147,147,147	0
2	HG	B	1286	1/1	0.95	0.28	109,109,109,109	1
2	HG	B	1287	1/1	0.96	0.05	147,147,147,147	0
2	HG	A	1290	1/1	0.96	0.06	147,147,147,147	0
2	HG	B	1290	1/1	0.96	0.04	147,147,147,147	0
2	HG	A	1289	1/1	0.97	0.05	147,147,147,147	0
2	HG	B	1289	1/1	0.97	0.06	147,147,147,147	0
2	HG	B	1285	1/1	0.97	0.05	147,147,147,147	0
2	HG	A	1287	1/1	0.98	0.06	147,147,147,147	0
2	HG	A	1285	1/1	0.98	0.05	147,147,147,147	0

6.5 Other polymers (i)

There are no such residues in this entry.