



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 06:10 PM GMT

PDB ID : 4KSD
Title : Structures of P-glycoprotein reveal its conformational flexibility and an epitope on the nucleotide-binding domain
Authors : Ward, A.; Szewczyk, P.; Grimard, V.; Lee, C.-W.; Martinez, L.; Doshi, R.; Caya, A.; Villaluz, M.; Pardon, E.; Cregger, C.; Swartz, D.J.; Falson, P.; Urbatsch, I.; Govaerts, C.; Steyaert, J.; Chang, G.
Deposited on : 2013-05-17
Resolution : 4.10 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

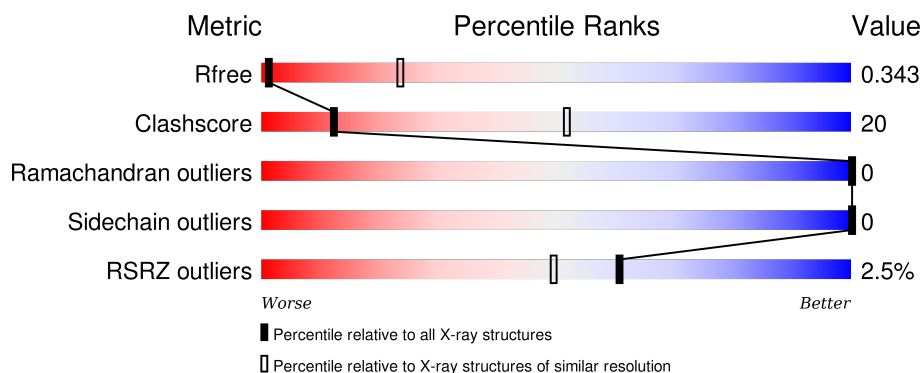
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 4.10 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	1018 (4.60-3.60)
Clashscore	102246	1117 (4.60-3.60)
Ramachandran outliers	100387	1063 (4.60-3.60)
Sidechain outliers	100360	1049 (4.60-3.60)
RSRZ outliers	91569	1022 (4.60-3.60)

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 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	1284	
2	B	124	

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 10084 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
			9171	5895	1552	1686	38			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
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

- Molecule 2 is a protein called R2 protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	118	Total	C	N	O	S	8	0	0
			913	573	159	177	4			

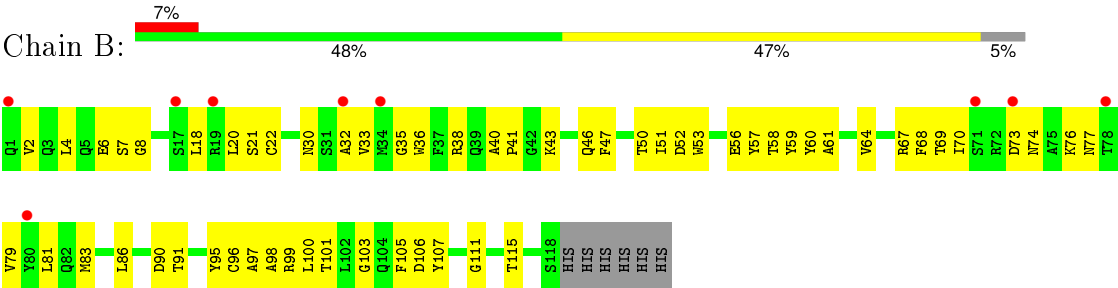
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according 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.

- Chain A: ■ 2% ■ 60% ■ 32% ■ 8%

Residue	Amino Acid	Frequency
1	Met	2%
2	Leu	2%
3	Leu	2%
4	Leu	2%
5	Leu	2%
6	Leu	2%
7	Leu	2%
8	Leu	2%
9	Leu	2%
10	Leu	2%
11	Leu	2%
12	Leu	2%
13	Leu	2%
14	Leu	2%
15	Leu	2%
16	Leu	2%
17	Leu	2%
18	Leu	2%
19	Leu	2%
20	Leu	2%
21	Leu	2%
22	Leu	2%
23	Leu	2%
24	Leu	2%
25	Leu	2%
26	Leu	2%
27	Leu	2%
28	Leu	2%
29	Leu	2%
30	Leu	2%
31	Leu	2%
32	Leu	2%
33	Leu	2%
34	Leu	2%
35	Leu	2%
36	Leu	2%
37	Leu	2%
38	Leu	2%
39	Leu	2%
40	Leu	2%
41	Leu	2%
42	Leu	2%
43	Leu	2%
44	Leu	2%
45	Leu	2%
46	Leu	2%
47	Leu	2%
48	Leu	2%
49	Leu	2%
50	Leu	2%
51	Leu	2%
52	Leu	2%
53	Leu	2%
54	Leu	2%
55	Leu	2%
56	Leu	2%
57	Leu	2%
58	Leu	2%
59	Leu	2%
60	Leu	2%
61	Leu	2%
62	Leu	2%
63	Leu	2%
64	Leu	2%
65	Leu	2%
66	Leu	2%
67	Leu	2%
68	Leu	2%
69	Leu	2%
70	Leu	2%
71	Leu	2%
72	Leu	2%
73	Leu	2%
74	Leu	2%
75	Leu	2%
76	Leu	2%
77	Leu	2%
78	Leu	2%
79	Leu	2%
80	Leu	2%
81	Leu	2%
82	Leu	2%
83	Leu	2%
84	Leu	2%
85	Leu	2%
86	Leu	2%
87	Leu	2%
88	Leu	2%
89	Leu	2%
90	Leu	2%
91	Leu	2%
92	Leu	2%
93	Leu	2%
94	Leu	2%
95	Leu	2%
96	Leu	2%
97	Leu	2%
98	Leu	2%
99	Leu	2%
100	Leu	2%
101	Leu	2%
102	Leu	2%
103	Leu	2%
104	Leu	2%
105	Leu	2%
106	Leu	2%
107	Leu	2%
108	Leu	2%
109	Leu	2%
110	Leu	2%
111	Leu	2%
112	Leu	2%
113	Leu	2%
114	Leu	2%
115	Leu	2%
116	Leu	2%
117	Leu	2%
118	Leu	2%
119	Leu	2%
120	Leu	2%
121	Leu	2%
122	Leu	2%
123	Leu	2%
124	Leu	2%
125	Leu	2%
126	Leu	2%
127	Leu	2%
128	Leu	2%
129	Leu	2%
130	Leu	2%
131	Leu	2%
132	Leu	2%
133	Leu	2%
134	Leu	2%
135	Leu	2%
136	Leu	2%
137	Leu	2%
138	Leu	2%
139	Leu	2%
140	Leu	2%
141	Leu	2%
142	Leu	2%
143	Leu	2%
144	Leu	2%
145	Leu	2%
146	Leu	2%
147	Leu	2%
148	Leu	2%
149	Leu	2%
150	Leu	2%
151	Leu	2%
152	Leu	2%
153	Leu	2%



• Molecule 2: R2 protein



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	87.10Å 102.47Å 312.10Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	61.07 – 4.10 64.91 – 4.10	Depositor EDS
% Data completeness (in resolution range)	95.7 (61.07-4.10) 95.7 (64.91-4.10)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	7.43 (at 4.14Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8.2_1309)	Depositor
R, R_{free}	0.323 , 0.344 0.319 , 0.343	Depositor DCC
R_{free} test set	1059 reflections (4.88%)	DCC
Wilson B-factor (Å ²)	95.8	Xtriage
Anisotropy	0.175	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.25 , 111.5	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.25$, $\langle L^2 \rangle = 0.09$	Xtriage
Outliers	0 of 21722 reflections	Xtriage
F_o, F_c correlation	0.66	EDS
Total number of atoms	10084	wwPDB-VP
Average B, all atoms (Å ²)	127.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.99% 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](#)

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	0.29	0/9339	0.53	2/12626 (0.0%)
2	B	0.34	0/932	0.68	1/1261 (0.1%)
All	All	0.30	0/10271	0.54	3/13887 (0.0%)

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	1042	THR	N-CA-C	-6.70	92.90	111.00
2	B	86	LEU	CA-CB-CG	5.24	127.35	115.30
1	A	993	ASP	CB-CG-OD2	5.17	122.95	118.30

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	9171	0	9344	342	0
2	B	913	0	877	71	0
All	All	10084	0	10221	400	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 20.

The worst 5 of 400 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:704:TRP:CZ2	1:A:707:PHE:HB2	1.43	1.51
1:A:310:PHE:CE1	1:A:332:PHE:CE2	2.20	1.29
1:A:310:PHE:CE1	1:A:332:PHE:CZ	2.25	1.24
1:A:704:TRP:CZ2	1:A:707:PHE:CB	2.29	1.15
1:A:1026:MET:C	1:A:1027:LEU:HD12	1.67	1.13

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%)	1071 (91%)	107 (9%)	0	100	100
2	B	116/124 (94%)	95 (82%)	21 (18%)	0	100	100
All	All	1294/1408 (92%)	1166 (90%)	128 (10%)	0	100	100

There are no Ramachandran outliers to report.

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	976/1065 (92%)	976 (100%)	0	100	100
2	B	94/100 (94%)	94 (100%)	0	100	100
All	All	1070/1165 (92%)	1070 (100%)	0	100	100

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

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	385	GLN
1	A	544	ASN
1	A	1211	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](#)

There are no ligands in this entry.

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	1182/1284 (92%)	0.16	23 (1%) 70 60	46, 123, 185, 232	0
2	B	117/124 (94%)	0.68	9 (7%) 16 12	123, 157, 179, 198	0
All	All	1299/1408 (92%)	0.20	32 (2%) 61 49	46, 130, 185, 232	0

The worst 5 of 32 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	319	SER	4.0
1	A	226	GLY	3.5
1	A	318	ILE	3.4
2	B	34	MET	3.2
1	A	230	LYS	3.0

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](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.