



# wwPDB X-ray Structure Validation Summary Report ⓘ

Sep 23, 2023 – 09:33 PM EDT

PDB ID : 5KPD  
Title : Mouse pgp 34 linker deleted double EQ mutant  
Authors : Xia, D.; Esser, L.; Zhou, F.  
Deposited on : 2016-07-03  
Resolution : 3.35 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

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ 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 ⓘ](#)) 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

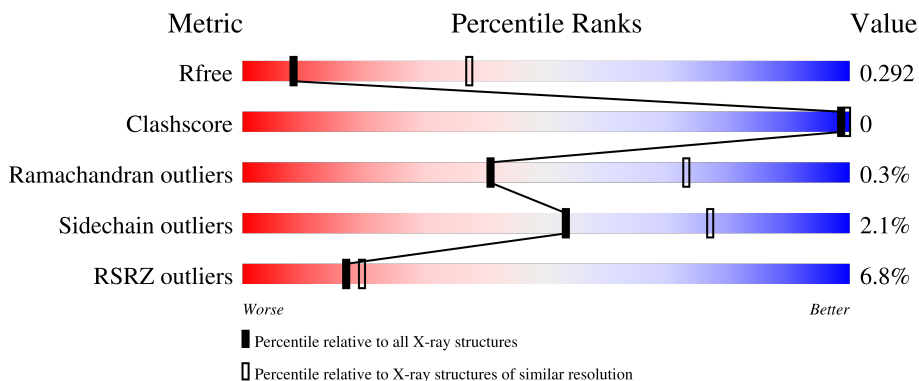
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

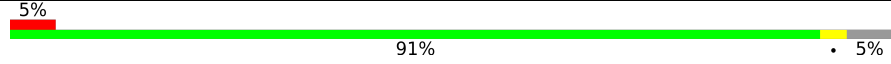
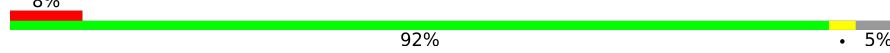
The reported resolution of this entry is 3.35 Å.

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	1558 (3.42-3.30)
Clashscore	141614	1627 (3.42-3.30)
Ramachandran outliers	138981	1599 (3.42-3.30)
Sidechain outliers	138945	1598 (3.42-3.30)
RSRZ outliers	127900	1507 (3.42-3.30)

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  $\geq 3$ , 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  $\leq 5\%$ . 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	1248	
1	B	1248	

## 2 Entry composition [i](#)

There is only 1 type of molecule in this entry. The entry contains 37084 atoms, of which 18710 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
			Total	C	H	N	O	S			
1	A	1185	18546	5907	9357	1559	1685	38	0	0	0
1	B	1184	18538	5905	9353	1558	1684	38	0	0	0

There are 84 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	552	GLN	GLU	engineered mutation	UNP P21447
A	?	-	MET	deletion	UNP P21447
A	?	-	SER	deletion	UNP P21447
A	?	-	SER	deletion	UNP P21447
A	?	-	LYS	deletion	UNP P21447
A	?	-	ASP	deletion	UNP P21447
A	?	-	SER	deletion	UNP P21447
A	?	-	GLY	deletion	UNP P21447
A	?	-	SER	deletion	UNP P21447
A	?	-	SER	deletion	UNP P21447
A	?	-	LEU	deletion	UNP P21447
A	?	-	ILE	deletion	UNP P21447
A	?	-	ARG	deletion	UNP P21447
A	?	-	ARG	deletion	UNP P21447
A	?	-	ARG	deletion	UNP P21447
A	?	-	SER	deletion	UNP P21447
A	?	-	THR	deletion	UNP P21447
A	?	-	ARG	deletion	UNP P21447
A	?	-	LYS	deletion	UNP P21447
A	?	-	SER	deletion	UNP P21447
A	?	-	ILE	deletion	UNP P21447
A	?	-	CYS	deletion	UNP P21447
A	?	-	GLY	deletion	UNP P21447
A	?	-	PRO	deletion	UNP P21447
A	?	-	HIS	deletion	UNP P21447

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	?	-	ASP	deletion	UNP P21447
A	?	-	GLN	deletion	UNP P21447
A	?	-	ASP	deletion	UNP P21447
A	?	-	ARG	deletion	UNP P21447
A	?	-	LYS	deletion	UNP P21447
A	?	-	LEU	deletion	UNP P21447
A	?	-	SER	deletion	UNP P21447
A	?	-	THR	deletion	UNP P21447
A	?	-	LYS	deletion	UNP P21447
A	?	-	GLU	deletion	UNP P21447
A	1197	GLN	GLU	engineered mutation	UNP P21447
A	1277	HIS	-	expression tag	UNP P21447
A	1278	HIS	-	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
B	552	GLN	GLU	engineered mutation	UNP P21447
B	?	-	MET	deletion	UNP P21447
B	?	-	SER	deletion	UNP P21447
B	?	-	SER	deletion	UNP P21447
B	?	-	LYS	deletion	UNP P21447
B	?	-	ASP	deletion	UNP P21447
B	?	-	SER	deletion	UNP P21447
B	?	-	GLY	deletion	UNP P21447
B	?	-	SER	deletion	UNP P21447
B	?	-	SER	deletion	UNP P21447
B	?	-	LEU	deletion	UNP P21447
B	?	-	ILE	deletion	UNP P21447
B	?	-	ARG	deletion	UNP P21447
B	?	-	ARG	deletion	UNP P21447
B	?	-	ARG	deletion	UNP P21447
B	?	-	SER	deletion	UNP P21447
B	?	-	THR	deletion	UNP P21447
B	?	-	ARG	deletion	UNP P21447
B	?	-	LYS	deletion	UNP P21447
B	?	-	SER	deletion	UNP P21447
B	?	-	ILE	deletion	UNP P21447
B	?	-	CYS	deletion	UNP P21447
B	?	-	GLY	deletion	UNP P21447
B	?	-	PRO	deletion	UNP P21447
B	?	-	HIS	deletion	UNP P21447

*Continued on next page...*

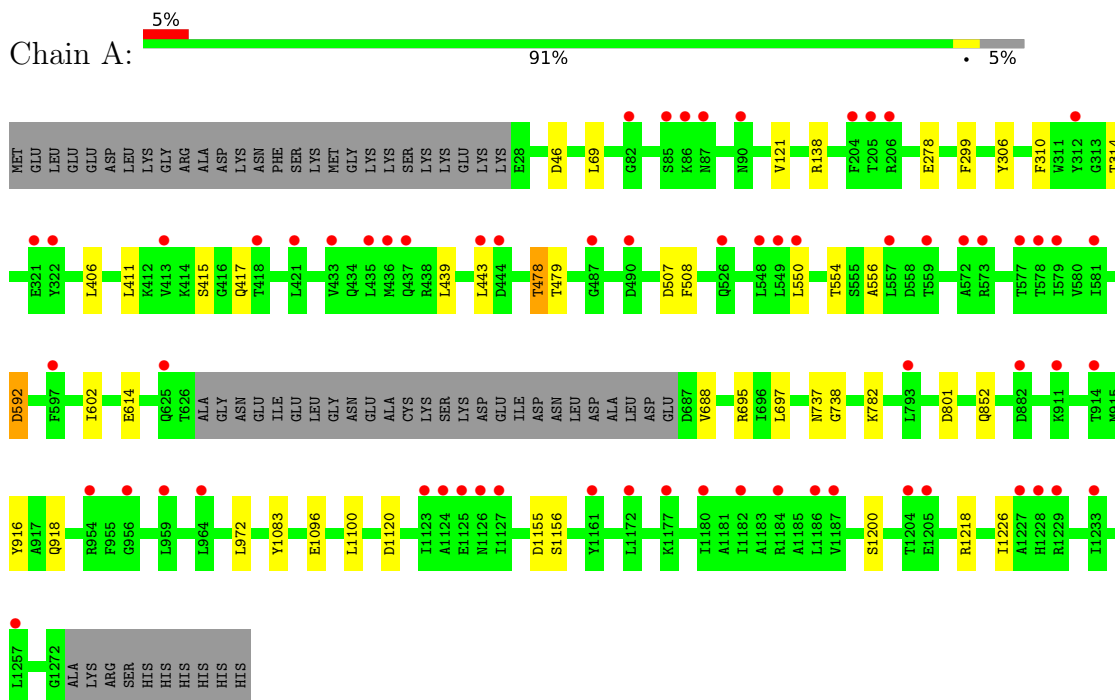
*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	?	-	ASP	deletion	UNP P21447
B	?	-	GLN	deletion	UNP P21447
B	?	-	ASP	deletion	UNP P21447
B	?	-	ARG	deletion	UNP P21447
B	?	-	LYS	deletion	UNP P21447
B	?	-	LEU	deletion	UNP P21447
B	?	-	SER	deletion	UNP P21447
B	?	-	THR	deletion	UNP P21447
B	?	-	LYS	deletion	UNP P21447
B	?	-	GLU	deletion	UNP P21447
B	1197	GLN	GLU	engineered mutation	UNP P21447
B	1277	HIS	-	expression tag	UNP P21447
B	1278	HIS	-	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

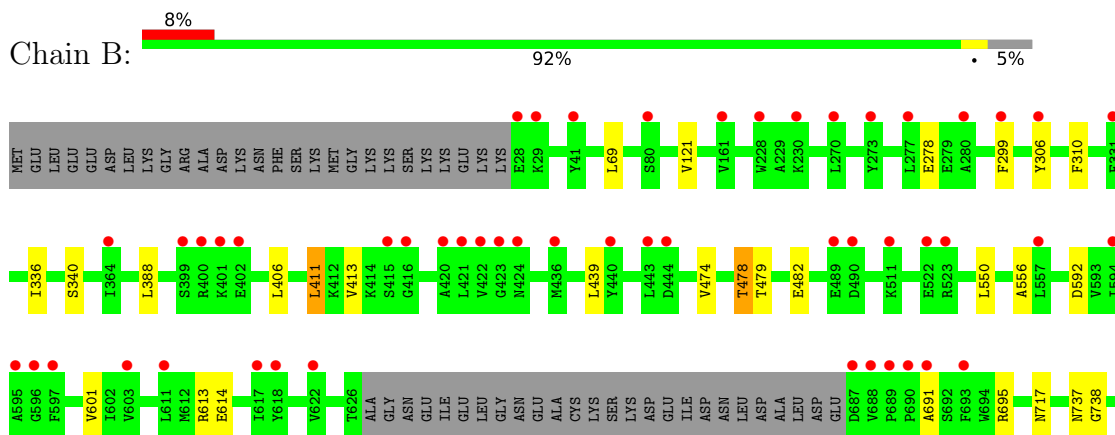
### 3 Residue-property plots [i](#)

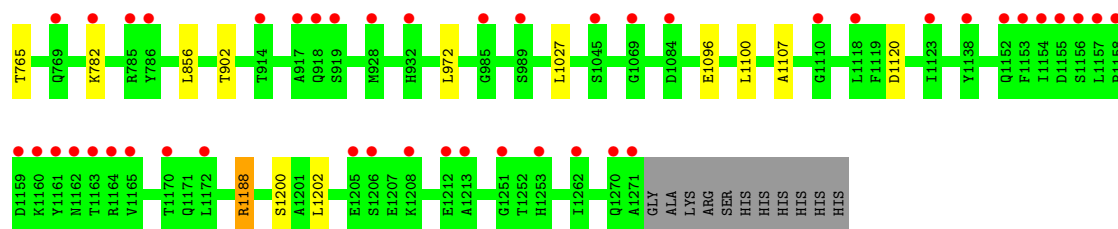
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. 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



- Molecule 1: Multidrug resistance protein 1A





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	97.95Å 116.59Å 375.56Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	22.42 – 3.35 43.90 – 3.31	Depositor EDS
% Data completeness (in resolution range)	92.9 (22.42-3.35) 91.5 (43.90-3.31)	Depositor EDS
$R_{merge}$	0.10	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.18 (at 3.32Å)	Xtrriage
Refinement program	PHENIX dev_2443	Depositor
R, $R_{free}$	0.240 , 0.283 0.247 , 0.292	Depositor DCC
$R_{free}$ test set	2408 reflections (4.04%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	120.8	Xtrriage
Anisotropy	0.171	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.26 , 46.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.36$ , $\langle L^2 \rangle = 0.19$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.89	EDS
Total number of atoms	37084	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	158.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.35% 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, 0.333 respectively for untwinned datasets, and 0.375, 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	0.37	0/9358	0.69	5/12650 (0.0%)
1	B	0.37	0/9354	0.71	3/12645 (0.0%)
All	All	0.37	0/18712	0.70	8/25295 (0.0%)

There are no bond length outliers.

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	1100	LEU	CA-CB-CG	6.93	131.24	115.30
1	A	1100	LEU	CA-CB-CG	6.45	130.12	115.30
1	A	443	LEU	CB-CG-CD1	5.84	120.93	111.00
1	A	406	LEU	CA-CB-CG	5.37	127.65	115.30
1	A	1218	ARG	NE-CZ-NH1	5.30	122.95	120.30

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	9189	9357	9374	9	1
1	B	9185	9353	9371	9	1
All	All	18374	18710	18745	17	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 0.

The worst 5 of 17 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:801:ASP:OD2	1:A:1083:TYR:OH	2.21	0.59
1:B:717:ASN:ND2	1:B:765:THR:OG1	2.37	0.54
1:A:278:GLU:OE1	1:A:782:LYS:NZ	2.41	0.52
1:A:478:THR:HG22	1:A:479:THR:H	1.75	0.52
1:B:1107:ALA:O	1:B:1188:ARG:NH2	2.46	0.47

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:688:VAL:O	1:B:613:ARG:NH1[1_455]	2.19	0.01

## 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	1181/1248 (95%)	1123 (95%)	55 (5%)	3 (0%)	41 73
1	B	1180/1248 (95%)	1120 (95%)	56 (5%)	4 (0%)	41 73
All	All	2361/2496 (95%)	2243 (95%)	111 (5%)	7 (0%)	41 73

5 of 7 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	1120	ASP
1	B	1120	ASP
1	A	411	LEU
1	A	556	ALA
1	B	411	LEU

### 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	977/1031 (95%)	955 (98%)	22 (2%)	50 75
1	B	977/1031 (95%)	958 (98%)	19 (2%)	57 79
All	All	1954/2062 (95%)	1913 (98%)	41 (2%)	53 77

5 of 41 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	439	LEU
1	B	972	LEU
1	B	550	LEU
1	B	614	GLU
1	B	1096	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 7 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	717	ASN
1	B	962	GLN
1	B	1077	GLN
1	B	1020	GLN
1	A	1077	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 monosaccharides 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, 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	1185/1248 (94%)	0.28	64 (5%) 25 28	71, 136, 214, 294	0
1	B	1184/1248 (94%)	0.48	96 (8%) 12 13	74, 141, 228, 343	0
All	All	2369/2496 (94%)	0.38	160 (6%) 17 20	71, 139, 220, 343	0

The worst 5 of 160 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	1161	TYR	8.5
1	A	87	ASN	7.7
1	B	690	PRO	7.1
1	B	1158	PRO	6.6
1	A	1161	TYR	6.4

### 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 monosaccharides 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.