



# Full wwPDB X-ray Structure Validation Report ⓘ

May 16, 2016 – 07:51 PM EDT

PDB ID : 4Q9I  
Title : P-glycoprotein cocrystallised with QZ-Ala  
Authors : McGrath, A.P.; Szewczyk, P.; Chang, G.  
Deposited on : 2014-05-01  
Resolution : 3.78 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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 : 1.7.1 (RC1), CSD as537be (2016)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20027457  
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) : rb-20027457

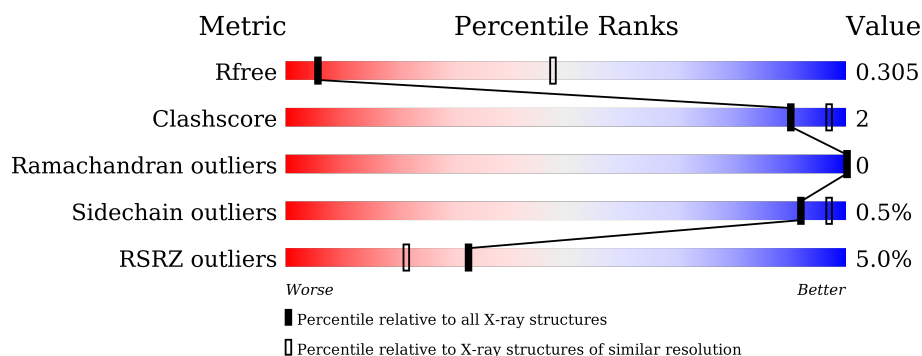
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.78 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	1273 (4.06-3.50)
Clashscore	102246	1412 (4.06-3.50)
Ramachandran outliers	100387	1351 (4.06-3.50)
Sidechain outliers	100360	1347 (4.06-3.50)
RSRZ outliers	91569	1281 (4.06-3.50)

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	1284	<div> <div>4%</div> <div>86%</div> <div>6%</div> <div>8%</div> </div>
2	B	6	<div> <div>33%</div> <div>83%</div> <div>17%</div> </div>
2	C	6	<div> <div>50%</div> <div>67%</div> <div>33%</div> </div>

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 9248 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	1184	Total	C	N	O	S	0	0	0
			9188	5906	1559	1685	38			

There are 11 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	83	GLN	ASN	CONFLICT	UNP P21447
A	87	GLN	ASN	CONFLICT	UNP P21447
A	90	GLN	ASN	CONFLICT	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

- Molecule 2 is a protein called (30F)A(30F)A(30F)A Peptide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	6	Total	C	N	O	Se	0	0	0
			30	18	6	3	3			
2	C	6	Total	C	N	O	Se	0	0	0
			30	18	6	3	3			



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	85.64Å 138.40Å 183.71Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	91.86 – 3.78 91.86 – 3.78	Depositor EDS
% Data completeness (in resolution range)	98.8 (91.86-3.78) 98.8 (91.86-3.78)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.10	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.64 (at 3.78Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8.2_1309)	Depositor
R, $R_{free}$	0.263 , 0.295 0.270 , 0.305	Depositor DCC
$R_{free}$ test set	1120 reflections (5.04%)	DCC
Wilson B-factor (Å <sup>2</sup> )	157.1	Xtriage
Anisotropy	0.340	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.24 , 78.2	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.85	EDS
Total number of atoms	9248	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	156.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.56% 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 [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: 30F

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	0.25	0/9355	0.42	0/12641
2	B	0.73	0/9	1.08	0/9
2	C	0.63	0/9	1.48	0/9
All	All	0.25	0/9373	0.42	0/12659

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

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	237	ASP	Peptide

### 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	9188	0	9378	42	0
2	B	30	0	15	1	0
2	C	30	0	15	3	0
All	All	9248	0	9408	42	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 2.

All (42) 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:160:ASP:OD2	1:A:901:ARG:NH2	2.26	0.69
1:A:1173:SER:OG	1:A:1176:GLN:OE1	2.21	0.58
1:A:729:SER:O	1:A:733:GLY:N	2.36	0.58
1:A:604:GLU:OE1	1:A:617:ILE:N	2.36	0.57
1:A:972:LEU:O	1:A:975:SER:OG	2.23	0.56
1:A:728:PHE:CE2	2:C:2:ALA:HB1	2.42	0.54
1:A:387:ASN:O	1:A:450:ASP:N	2.40	0.54
1:A:922:ILE:HB	1:A:923:PRO:HD3	1.90	0.54
1:A:801:ASP:OD2	1:A:1083:TYR:OH	2.25	0.54
1:A:495:GLU:N	1:A:495:GLU:OE1	2.42	0.53
1:A:1207:GLU:OE1	1:A:1229:ARG:NH2	2.42	0.52
1:A:843:ILE:HD13	1:A:858:LEU:HD11	1.93	0.51
1:A:1110:GLY:N	1:A:1192:ILE:O	2.43	0.50
1:A:729:SER:OG	1:A:968:GLU:O	2.28	0.50
1:A:728:PHE:CZ	2:C:2:ALA:HB1	2.46	0.50
1:A:734:VAL:O	1:A:738:GLY:N	2.45	0.50
1:A:906:LEU:O	1:A:907:THR:OG1	2.26	0.50
1:A:1204:THR:OG1	1:A:1205:GLU:N	2.46	0.49
1:A:544:ASN:ND2	1:A:544:ASN:O	2.46	0.49
1:A:93:GLU:N	1:A:93:GLU:OE1	2.46	0.48
1:A:861:VAL:HG11	1:A:981:ALA:HB2	1.97	0.47
1:A:727:ILE:HD12	1:A:758:LEU:HD23	1.97	0.47
1:A:762:SER:HA	1:A:765:THR:HG22	1.98	0.46
1:A:1165:VAL:HG12	1:A:1166:GLY:N	2.32	0.44
1:A:979:PHE:CE1	2:C:1:30F:C09	3.00	0.44
1:A:1032:GLN:NE2	1:A:1055:GLU:OE2	2.51	0.44
1:A:704:TRP:N	1:A:705:PRO:HD2	2.34	0.42
1:A:721:GLN:HB3	1:A:722:PRO:HD3	2.01	0.42
1:A:217:ILE:HG23	1:A:305:SER:HB2	2.00	0.42
1:A:339:PHE:CZ	2:B:6:ALA:HB1	2.54	0.42
1:A:78:PHE:O	1:A:82:GLY:N	2.50	0.42

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:1170:THR:HG22	1:A:1171:GLN:N	2.36	0.41
1:A:861:VAL:HB	1:A:862:PRO:HD3	2.02	0.41
1:A:396:SER:OG	1:A:403:VAL:O	2.30	0.41
1:A:408:GLY:O	1:A:409:LEU:HB2	2.20	0.41
1:A:372:ASP:O	1:A:373:SER:CB	2.68	0.41
1:A:388:LEU:HB2	1:A:413:VAL:CG1	2.51	0.40
1:A:573:ARG:O	1:A:574:GLU:HB3	2.21	0.40
1:A:1189:GLN:N	1:A:1190:PRO:HD3	2.36	0.40
1:A:214:ILE:O	1:A:218:SER:N	2.54	0.40
1:A:313:GLY:O	1:A:317:VAL:HG23	2.22	0.40
1:A:1150:ILE:HB	1:A:1179:ARG:HB3	2.04	0.40

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	1180/1284 (92%)	1129 (96%)	51 (4%)	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%)	971 (100%)	5 (0%)	92 97

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	310	PHE
1	A	372	ASP
1	A	700	ASN
1	A	750	LEU
1	A	872	MET

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

6 non-standard protein/DNA/RNA residues 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	30F	B	1	2	3,5,6	3.32	2 (66%)	1,5,7	0.20	0
2	30F	B	3	2	3,5,6	3.07	2 (66%)	1,5,7	1.01	0
2	30F	B	5	2	3,5,6	3.17	2 (66%)	1,5,7	0.15	0
2	30F	C	1	2	3,5,6	3.24	2 (66%)	1,5,7	1.19	0
2	30F	C	3	2	3,5,6	3.13	2 (66%)	1,5,7	0.94	0
2	30F	C	5	2	3,5,6	3.19	2 (66%)	1,5,7	0.22	0

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	30F	B	1	2	-	0/0/4/6	0/0/0/0
2	30F	B	3	2	-	0/0/4/6	0/0/0/0
2	30F	B	5	2	-	0/0/4/6	0/0/0/0
2	30F	C	1	2	-	0/0/4/6	0/0/0/0
2	30F	C	3	2	-	0/0/4/6	0/0/0/0
2	30F	C	5	2	-	0/0/4/6	0/0/0/0

All (12) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	1	30F	C10-C08	-4.85	1.37	1.45
2	C	1	30F	C10-C08	-4.75	1.37	1.45
2	C	3	30F	C10-C08	-4.73	1.37	1.45
2	B	5	30F	C10-C08	-4.66	1.37	1.45
2	C	5	30F	C10-C08	-4.62	1.37	1.45
2	B	3	30F	C10-C08	-4.52	1.37	1.45
2	B	1	30F	C08-N04	-3.09	1.27	1.35
2	C	5	30F	C08-N04	-3.00	1.27	1.35
2	C	1	30F	C08-N04	-3.00	1.27	1.35
2	B	5	30F	C08-N04	-2.86	1.27	1.35
2	B	3	30F	C08-N04	-2.80	1.28	1.35
2	C	3	30F	C08-N04	-2.64	1.28	1.35

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	C	1	30F	1	0

## 5.5 Carbohydrates ⓘ

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

### 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<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	1184/1284 (92%)	0.32	55 (4%) 36 24	69, 145, 253, 461	0
2	B	3/6 (50%)	2.87	2 (66%) 0 1	141, 141, 141, 141	0
2	C	3/6 (50%)	3.65	3 (100%) 0 0	141, 141, 141, 141	0
All	All	1190/1296 (91%)	0.34	60 (5%) 32 21	69, 145, 253, 461	0

All (60) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	95	ASP	6.8
1	A	96	LYS	6.6
1	A	234	SER	6.2
1	A	231	ILE	5.2
1	A	554	THR	5.1
1	A	574	GLU	4.6
1	A	405	ILE	4.4
1	A	233	SER	4.2
2	B	4	ALA	4.2
2	C	2	ALA	4.1
1	A	986	GLN	4.1
1	A	324	ILE	4.1
2	C	6	ALA	3.8
1	A	230	LYS	3.7
1	A	92	SER	3.5
1	A	982	MET	3.5
1	A	306	TYR	3.4
1	A	1219	GLU	3.4
1	A	240	LEU	3.3
1	A	1225	VAL	3.3
1	A	1032	GLN	3.3
1	A	60	HIS	3.3
1	A	1079	LEU	3.1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	516	PHE	3.1
2	C	4	ALA	3.1
1	A	742	GLU	3.0
1	A	1055	GLU	3.0
1	A	987	VAL	3.0
1	A	599	GLY	3.0
1	A	755	PHE	2.8
1	A	468	VAL	2.8
1	A	1195	LEU	2.8
1	A	1033	PHE	2.7
2	B	6	ALA	2.7
1	A	1193	LEU	2.6
1	A	99	MET	2.6
1	A	1110	GLY	2.5
1	A	331	PHE	2.5
1	A	71	PHE	2.5
1	A	1125	GLU	2.5
1	A	269	GLU	2.4
1	A	983	ALA	2.4
1	A	310	PHE	2.4
1	A	904	VAL	2.4
1	A	732	VAL	2.4
1	A	277	LEU	2.3
1	A	971	LEU	2.3
1	A	438	ARG	2.3
1	A	897	ILE	2.2
1	A	572	ALA	2.2
1	A	476	PHE	2.2
1	A	86	LYS	2.2
1	A	1117	ILE	2.2
1	A	89	THR	2.2
1	A	1080	GLU	2.1
1	A	112	TYR	2.1
1	A	1004	ILE	2.0
1	A	624	THR	2.0
1	A	379	HIS	2.0
1	A	900	PHE	2.0

## 6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

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( $\text{\AA}^2$ )	Q<0.9
2	30F	C	5	6/7	0.71	0.32	-	140,140,140,140	0
2	30F	B	1	6/7	0.78	0.49	-	140,140,140,140	0
2	30F	C	1	6/7	0.77	0.62	-	140,140,140,140	0
2	30F	B	3	6/7	0.49	0.66	-	140,140,140,140	0
2	30F	C	3	6/7	0.80	0.83	-	140,140,140,140	0
2	30F	B	5	6/7	0.77	0.54	-	140,140,140,140	0

### 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.