



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 13, 2017 – 05:03 pm GMT

PDB ID : 4EQA  
Title : Crystal structure of PA1844 in complex with PA1845 from *Pseudomonas aeruginosa* PAO1  
Authors : Shang, G.; Li, N.; Zhang, J.; Lu, D.; Yu, Q.; Zhao, Y.; Liu, X.; Xu, S.; Gu, L.  
Deposited on : 2012-04-18  
Resolution : 1.60 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  
Xtriage (Phenix) : 1.9-1692  
EDS : trunk28620  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
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) : recalc28949

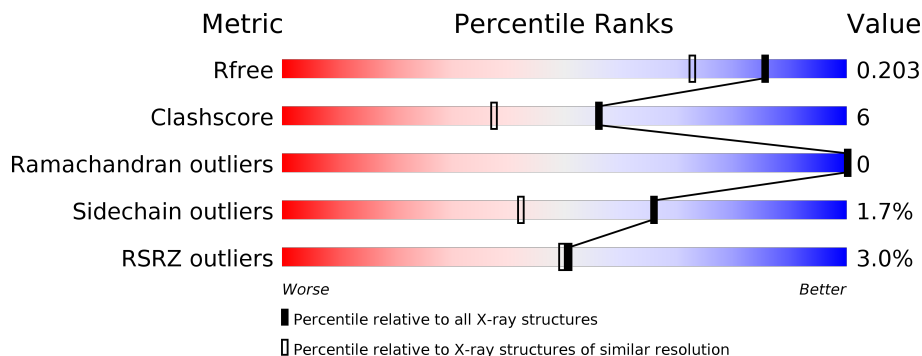
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 1.60 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))
$R_{free}$	100719	2696 (1.60-1.60)
Clashscore	112137	2967 (1.60-1.60)
Ramachandran outliers	110173	2887 (1.60-1.60)
Sidechain outliers	110143	2886 (1.60-1.60)
RSRZ outliers	101464	2714 (1.60-1.60)

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  $\geq 3$ , 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	146	<div> <div>97%</div> <div> <div></div> <div></div> <div></div> <div></div> </div> </div>
1	B	146	<div> <div>%</div> <div> <div></div> <div></div> <div></div> <div></div> </div> <div> <div>92%</div> <div>7%</div> </div> </div>
2	C	153	<div> <div>4%</div> <div> <div></div> <div></div> <div></div> <div></div> </div> <div> <div>83%</div> <div>13%</div> </div> </div>
2	D	153	<div> <div>7%</div> <div> <div></div> <div></div> <div></div> <div></div> </div> <div> <div>86%</div> <div>12%</div> </div> </div>

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 5329 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 Putative uncharacterized protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	146	Total	C	N	O	S	0	1	0
			1097	687	197	203	10			
1	B	146	Total	C	N	O	S	0	1	0
			1099	688	198	203	10			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	3	SER	-	EXPRESSION TAG	UNP Q9I2Q1
A	4	HIS	-	EXPRESSION TAG	UNP Q9I2Q1
A	5	MET	-	EXPRESSION TAG	UNP Q9I2Q1
B	3	SER	-	EXPRESSION TAG	UNP Q9I2Q1
B	4	HIS	-	EXPRESSION TAG	UNP Q9I2Q1
B	5	MET	-	EXPRESSION TAG	UNP Q9I2Q1

- Molecule 2 is a protein called Putative uncharacterized protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	150	Total	C	N	O	S	0	0	0
			1155	708	204	234	9			
2	D	150	Total	C	N	O	S	0	8	0
			1199	743	208	239	9			

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	211	Total	O	0	0
			211	211		
3	B	167	Total	O	0	0
			167	167		
3	C	226	Total	O	0	0
			226	226		

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	D	175	Total	O	0	0
			175	175		

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

These plots are drawn for all protein, RNA and DNA 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.

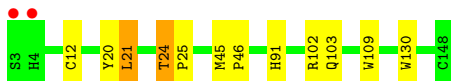
- Molecule 1: Putative uncharacterized protein

Chain A:  97% .




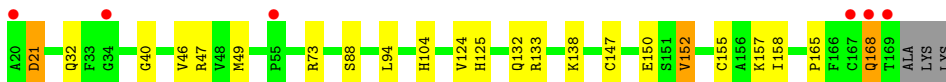
- Molecule 1: Putative uncharacterized protein

Chain B:  92% 7% .




- Molecule 2: Putative uncharacterized protein

Chain C:  83% 13% 4% . .



- Molecule 2: Putative uncharacterized protein

Chain D:  86% 12% 7% .



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	48.51Å 99.20Å 56.33Å 90.00° 98.82° 90.00°	Depositor
Resolution (Å)	27.22 – 1.60 27.22 – 1.60	Depositor EDS
% Data completeness (in resolution range)	97.8 (27.22-1.60) 97.6 (27.22-1.60)	Depositor EDS
$R_{merge}$	0.11	Depositor
$R_{sym}$	0.11	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	5.62 (at 1.60Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.6.4_486)	Depositor
R, $R_{free}$	0.174 , 0.206 0.171 , 0.203	Depositor DCC
$R_{free}$ test set	1966 reflections (2.90%)	DCC
Wilson B-factor (Å <sup>2</sup> )	17.1	Xtriage
Anisotropy	0.243	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 48.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	5329	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	26.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.02% 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	0.35	0/1123	0.50	0/1521
1	B	0.32	0/1125	0.47	0/1522
2	C	0.35	0/1179	0.53	0/1593
2	D	0.30	0/1252	0.49	0/1689
All	All	0.33	0/4679	0.50	0/6325

There are no bond length outliers.

There are no bond angle outliers.

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	1097	0	1079	3	0
1	B	1099	0	1083	9	0
2	C	1155	0	1070	25	0
2	D	1199	0	1144	15	0
3	A	211	0	0	2	0
3	B	167	0	0	2	0
3	C	226	0	0	4	0
3	D	175	0	0	2	0
All	All	5329	0	4376	52	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 6.

All (52) 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)
2:C:155:CYS:HB2	2:C:158:ILE:HD11	1.35	1.06
2:D:155:CYS:HB2	2:D:158:ILE:HD11	1.58	0.84
2:C:155:CYS:HB2	2:C:158:ILE:CD1	2.11	0.79
2:C:152:VAL:HG22	2:C:158:ILE:HD12	1.67	0.76
2:D:152:VAL:HG22	2:D:158:ILE:HD12	1.70	0.73
2:C:124:VAL:HG13	2:C:168:GLN:HG3	1.71	0.72
2:C:40:GLY:HA3	2:C:49:MET:HE1	1.72	0.72
2:D:46:VAL:HG11	2:D:94:LEU:HD22	1.71	0.71
2:C:124:VAL:CG1	2:C:168:GLN:HG3	2.23	0.68
2:C:88:SER:OG	2:C:104:HIS:HD2	1.77	0.66
2:D:155:CYS:HB2	2:D:158:ILE:CD1	2.28	0.64
2:C:46:VAL:HG11	2:C:94:LEU:HD22	1.82	0.61
2:C:47:ARG:NE	2:C:49:MET:HE3	2.16	0.60
2:D:28:GLU:HG2	3:D:355:HOH:O	2.00	0.60
2:D:20:ALA:HB2	2:D:162:SER:OG	2.03	0.59
2:D:88[A]:SER:OG	2:D:104:HIS:HD2	1.87	0.57
2:D:53:GLU:O	2:D:55:PRO:HD3	2.05	0.56
1:B:102:ARG:O	1:B:103:GLN:HG2	2.06	0.55
2:C:88:SER:OG	2:C:104:HIS:CD2	2.60	0.54
2:C:21:ASP:HB3	3:C:325:HOH:O	2.08	0.53
2:C:40:GLY:HA3	2:C:49:MET:CE	2.39	0.52
2:C:32:GLN:HE21	2:C:47:ARG:NH2	2.08	0.51
2:C:125:HIS:CD2	2:C:165:PRO:HB2	2.45	0.51
2:D:45:HIS:HD2	2:D:74:ASP:OD2	1.93	0.51
2:C:32:GLN:HG3	2:C:49:MET:SD	2.53	0.48
2:C:46:VAL:HG13	2:C:94:LEU:HD13	1.94	0.48
1:B:24:THR:HG21	3:B:313:HOH:O	2.14	0.48
2:D:46:VAL:CG1	2:D:94:LEU:HD22	2.41	0.48
2:C:47:ARG:HD3	3:C:407:HOH:O	2.14	0.48
2:C:73:ARG:HG2	3:C:388:HOH:O	2.14	0.48
1:B:20:TYR:CE2	1:B:21:LEU:HD22	2.49	0.47
1:B:24:THR:HG23	1:B:25:PRO:HD2	1.96	0.47
2:D:98:GLY:HA2	3:D:358:HOH:O	2.15	0.47
2:D:20:ALA:N	2:D:164:ALA:N	2.65	0.45
2:D:88[A]:SER:OG	2:D:104:HIS:CD2	2.68	0.45
1:B:20:TYR:O	1:B:21:LEU:HB2	2.17	0.44
2:C:147:CYS:HA	2:C:155:CYS:HA	1.99	0.44
2:C:133:ARG:CZ	2:C:150:GLU:HA	2.49	0.43
2:D:45:HIS:CD2	2:D:74:ASP:OD2	2.72	0.43
2:C:47:ARG:CZ	2:C:49:MET:HE3	2.49	0.42

*Continued on next page...*



Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:A:205:HOH:O	2:C:104:HIS:HE1	2.01	0.42
2:C:138:LYS:HG2	3:C:340:HOH:O	2.20	0.42
2:C:157:LYS:O	2:C:158:ILE:HD13	2.20	0.42
1:A:107:MET:HA	1:A:125:SER:HA	2.02	0.42
3:B:245:HOH:O	2:D:104:HIS:HE1	2.02	0.42
1:B:12:CYS:HB3	1:B:109:TRP:CD2	2.55	0.41
1:B:102:ARG:C	1:B:103:GLN:HG2	2.41	0.41
1:A:47:ARG:HD3	3:A:264:HOH:O	2.21	0.41
1:A:55:ASP:O	1:A:59:GLN:HG3	2.21	0.41
2:C:152:VAL:O	2:C:158:ILE:HD11	2.20	0.41
1:B:91:HIS:HB3	1:B:130:TRP:CZ2	2.56	0.40
1:B:45:MET:SD	1:B:46:PRO:HD2	2.61	0.40

There are no symmetry-related clashes.

## 5.3 Torsion angles

### 5.3.1 Protein backbone

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	145/146 (99%)	141 (97%)	4 (3%)	0	100	100
1	B	145/146 (99%)	141 (97%)	4 (3%)	0	100	100
2	C	148/153 (97%)	142 (96%)	6 (4%)	0	100	100
2	D	157/153 (103%)	151 (96%)	6 (4%)	0	100	100
All	All	595/598 (100%)	575 (97%)	20 (3%)	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	114/113 (101%)	114 (100%)	0	100	100
1	B	114/113 (101%)	112 (98%)	2 (2%)	64	40
2	C	128/130 (98%)	124 (97%)	4 (3%)	45	18
2	D	137/130 (105%)	135 (98%)	2 (2%)	70	48
All	All	493/486 (101%)	485 (98%)	8 (2%)	66	45

All (8) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	21	LEU
1	B	24	THR
2	C	21	ASP
2	C	132	GLN
2	C	152	VAL
2	C	168	GLN
2	D	21	ASP
2	D	22	CYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (6) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	139	ASN
1	B	139	ASN
2	C	32	GLN
2	C	104	HIS
2	D	45	HIS
2	D	104	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 [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

### 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	146/146 (100%)	-0.36	0 <span>100</span> <span>100</span>	12, 17, 28, 45	0
1	B	146/146 (100%)	-0.25	2 (1%) <span>75</span> <span>76</span>	13, 22, 36, 71	0
2	C	150/153 (98%)	-0.07	6 (4%) <span>39</span> <span>37</span>	12, 21, 48, 76	0
2	D	150/153 (98%)	0.38	10 (6%) <span>19</span> <span>18</span>	14, 28, 57, 82	0
All	All	592/598 (98%)	-0.07	18 (3%) <span>51</span> <span>49</span>	12, 21, 46, 82	0

All (18) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	D	20	ALA	9.0
2	C	20	ALA	8.3
2	C	169	THR	5.0
2	C	167	CYS	4.4
2	D	169	THR	4.3
2	D	55	PRO	4.1
2	D	56	ASN	3.9
2	C	55	PRO	3.6
1	B	3	SER	3.5
2	D	21	ASP	3.3
2	D	152	VAL	3.1
2	D	58	ASP	2.8
2	D	34	GLY	2.7
2	C	168	GLN	2.7
1	B	4	HIS	2.6
2	D	22	CYS	2.3
2	C	34	GLY	2.2
2	D	145	GLN	2.1

## 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.