



wwPDB X-ray Structure Validation Summary Report ⓘ

May 25, 2020 – 11:20 pm BST

PDB ID : 6Q1H
Title : Structure of *P. aeruginosa* ATCC27853 NucC, cAAA-bound form
Authors : Ye, Q.; Lau, R.K.; Corbett, K.D.
Deposited on : 2019-08-04
Resolution : 1.45 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 ⓘ symbol.

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.11
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)	:	2.11

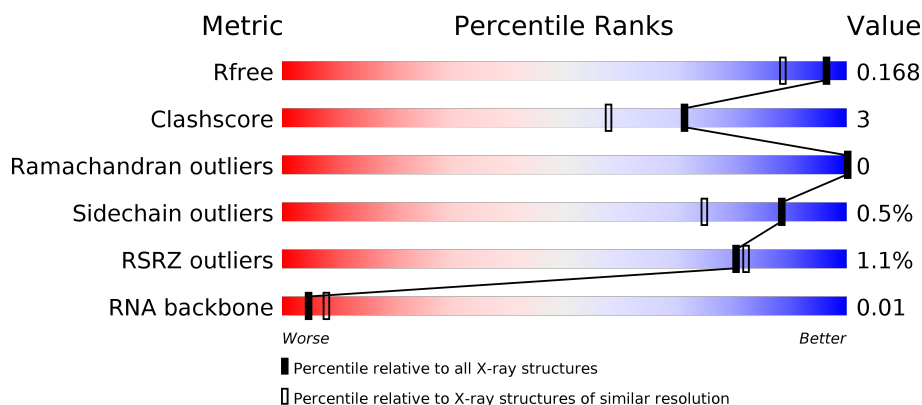
1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.45 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	1156 (1.46-1.46)
Clashscore	141614	1202 (1.46-1.46)
Ramachandran outliers	138981	1178 (1.46-1.46)
Sidechain outliers	138945	1178 (1.46-1.46)
RSRZ outliers	127900	1139 (1.46-1.46)
RNA backbone	3102	1000 (2.34-0.62)

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 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	241	<div> <div>2%</div> <div> <div></div> <div>93%</div> <div>6%</div> </div> </div>
1	B	241	<div> <div>2%</div> <div> <div></div> <div>95%</div> <div>5%</div> </div> </div>
1	C	241	<div> <div>%</div> <div> <div></div> <div>94%</div> <div>6%</div> </div> </div>
1	E	241	<div> <div></div> <div> <div></div> <div>94%</div> <div>5%</div> </div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	241	<div><div></div><div>2%</div><div>95%</div><div>5%</div></div>
1	G	241	<div><div></div><div>94%</div><div>5%</div></div>
2	D	3	<div><div></div><div>100%</div></div>
2	H	3	<div><div></div><div>100%</div></div>

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 24734 atoms, of which 11172 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 Bacterial protein ORF C62.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	240	Total	C	H	N	O	S	0	4	0
			3723	1209	1837	316	357	4			
1	B	240	Total	C	H	N	O	S	0	5	0
			3759	1218	1857	319	361	4			
1	C	240	Total	C	H	N	O	S	0	7	0
			3757	1220	1856	318	359	4			
1	E	240	Total	C	H	N	O	S	0	2	0
			3744	1215	1852	317	356	4			
1	F	240	Total	C	H	N	O	S	0	8	0
			3782	1226	1871	320	361	4			
1	G	239	Total	C	H	N	O	S	0	4	0
			3713	1206	1833	315	355	4			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	73	ASN	ASP	engineered mutation	UNP Q8GQ48
B	73	ASN	ASP	engineered mutation	UNP Q8GQ48
C	73	ASN	ASP	engineered mutation	UNP Q8GQ48
E	73	ASN	ASP	engineered mutation	UNP Q8GQ48
F	73	ASN	ASP	engineered mutation	UNP Q8GQ48
G	73	ASN	ASP	engineered mutation	UNP Q8GQ48

- Molecule 2 is a RNA chain called RNA (5'-R(P*AP*AP*A)-3').

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	D	3	Total	C	H	N	O	P	0	0	0
			99	30	33	15	18	3			
2	H	3	Total	C	H	N	O	P	0	0	0
			99	30	33	15	18	3			

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	347	Total 347	O 347	0	0
3	B	332	Total 332	O 332	0	0
3	C	330	Total 330	O 330	0	0
3	D	9	Total 9	O 9	0	0
3	E	356	Total 356	O 356	0	0
3	F	329	Total 329	O 329	0	0
3	G	348	Total 348	O 348	0	0
3	H	7	Total 7	O 7	0	0

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: Bacterial protein ORF C62



- Molecule 1: Bacterial protein ORF C62



- Molecule 1: Bacterial protein ORF C62



- Molecule 1: Bacterial protein ORF C62



- Molecule 1: Bacterial protein ORF C62



- Molecule 1: Bacterial protein ORF C62

Chain G:

94%

5% •



- Molecule 2: RNA (5'-R(P*AP*AP*A)-3')

Chain D:

100%



- Molecule 2: RNA (5'-R(P*AP*AP*A)-3')

Chain H:

100%



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	80.50Å 80.45Å 262.08Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	76.95 – 1.45 76.95 – 1.45	Depositor EDS
% Data completeness (in resolution range)	97.4 (76.95-1.45) 97.2 (76.95-1.45)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.02 (at 1.45Å)	Xtriage
Refinement program	PHENIX 1.15.2_3472	Depositor
R, R_{free}	0.147 , 0.167 0.147 , 0.168	Depositor DCC
R_{free} test set	14695 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å ²)	16.8	Xtriage
Anisotropy	0.279	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.38 , 43.8	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	0.015 for k,h,-l	Xtriage
F_o, F_c correlation	0.98	EDS
Total number of atoms	24734	wwPDB-VP
Average B, all atoms (Å ²)	21.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 27.79 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 2.0659e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

¹ Intensities estimated from amplitudes.

² 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	0.45	0/1951	0.63	0/2656
1	B	0.43	0/1965	0.62	0/2674
1	C	0.44	0/1975	0.63	0/2690
1	E	0.45	0/1942	0.62	0/2643
1	F	0.45	0/1984	0.63	0/2701
1	G	0.47	0/1945	0.65	0/2648
2	D	1.55	0/74	2.34	4/113 (3.5%)
2	H	1.72	2/74 (2.7%)	2.43	8/113 (7.1%)
All	All	0.48	2/11910 (0.0%)	0.69	12/16238 (0.1%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	H	3	A	N9-C4	-5.89	1.34	1.37
2	H	1	A	N9-C4	-5.10	1.34	1.37

The worst 5 of 12 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	H	3	A	C2-N3-C4	-10.67	105.27	110.60
2	D	3	A	C2-N3-C4	-9.92	105.64	110.60
2	D	1	A	C2-N3-C4	-8.92	106.14	110.60
2	H	1	A	C2-N3-C4	-8.77	106.22	110.60
2	H	2	A	C2-N3-C4	-7.89	106.66	110.60

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	1886	1837	1822	15	0
1	B	1902	1857	1841	11	0
1	C	1901	1856	1829	16	0
1	E	1892	1852	1850	15	0
1	F	1911	1871	1846	10	0
1	G	1880	1833	1817	13	0
2	D	66	33	33	0	0
2	H	66	33	33	0	0
3	A	347	0	0	10	0
3	B	332	0	0	12	1
3	C	330	0	0	11	0
3	D	9	0	0	0	0
3	E	356	0	0	11	0
3	F	329	0	0	5	0
3	G	348	0	0	7	1
3	H	7	0	0	0	0
All	All	13562	11172	11071	79	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 3.

The worst 5 of 79 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:C:170:LYS:NZ	3:C:301:HOH:O	2.11	0.82
1:B:241:LYS:O	3:B:301:HOH:O	1.99	0.78
1:C:19:GLN:NE2	3:C:303:HOH:O	2.12	0.77
1:B:16:ASP:OD1	3:B:302:HOH:O	2.02	0.77
1:G:15:GLU:OE1	3:G:301:HOH:O	2.06	0.73

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)
3:B:319:HOH:O	3:G:554:HOH:O[1_545]	2.00	0.20

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	242/241 (100%)	236 (98%)	6 (2%)	0	100	100
1	B	244/241 (101%)	241 (99%)	3 (1%)	0	100	100
1	C	245/241 (102%)	241 (98%)	4 (2%)	0	100	100
1	E	240/241 (100%)	235 (98%)	5 (2%)	0	100	100
1	F	246/241 (102%)	240 (98%)	6 (2%)	0	100	100
1	G	241/241 (100%)	235 (98%)	6 (2%)	0	100	100
All	All	1458/1446 (101%)	1428 (98%)	30 (2%)	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	202/200 (101%)	199 (98%)	3 (2%)	65	35
1	B	205/200 (102%)	204 (100%)	1 (0%)	88	75
1	C	205/200 (102%)	205 (100%)	0	100	100
1	E	201/200 (100%)	200 (100%)	1 (0%)	88	75
1	F	207/200 (104%)	206 (100%)	1 (0%)	88	75
1	G	201/200 (100%)	201 (100%)	0	100	100
All	All	1221/1200 (102%)	1215 (100%)	6 (0%)	88	75

5 of 6 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	168	MET
1	F	2	SER
1	B	161	SER
1	A	19	GLN
1	E	168	MET

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	19	GLN

5.3.3 RNA [i](#)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
2	D	2/3 (66%)	0	0
2	H	2/3 (66%)	0	0
All	All	4/6 (66%)	0	0

There are no RNA backbone outliers to report.

There are no RNA pucker outliers to report.

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 ⓘ

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, 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	240/241 (99%)	-0.42	4 (1%) 70 70	11, 16, 33, 60	0
1	B	240/241 (99%)	-0.34	5 (2%) 63 65	11, 17, 37, 67	0
1	C	240/241 (99%)	-0.37	2 (0%) 86 87	11, 16, 33, 57	0
1	E	240/241 (99%)	-0.46	1 (0%) 92 94	10, 15, 32, 55	0
1	F	240/241 (99%)	-0.38	4 (1%) 70 70	11, 17, 32, 59	0
1	G	239/241 (99%)	-0.43	0 100 100	10, 15, 30, 40	0
2	D	3/3 (100%)	-0.65	0 100 100	13, 13, 13, 13	0
2	H	3/3 (100%)	-0.63	0 100 100	12, 12, 12, 13	0
All	All	1445/1452 (99%)	-0.40	16 (1%) 80 82	10, 16, 33, 67	0

The worst 5 of 16 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	160	TRP	3.7
1	F	2	SER	3.6
1	B	2	SER	3.4
1	B	241	LYS	3.1
1	F	240	THR	3.1

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 carbohydrates in this entry.

6.4 Ligands

There are no ligands in this entry.

6.5 Other polymers

There are no such residues in this entry.