



Full wwPDB X-ray Structure Validation Report ⓘ

Feb 21, 2017 – 10:30 PM EST

PDB ID : 5G0Y
Title : Pseudomonas aeruginosa HDAH unliganded.
Authors : Kraemer, A.; Meyer-Almes, F.J.; Yildiz, O.
Deposited on : 2016-03-23
Resolution : 2.29 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

A user guide is available at

<http://wwpdb.org/validation/2016/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

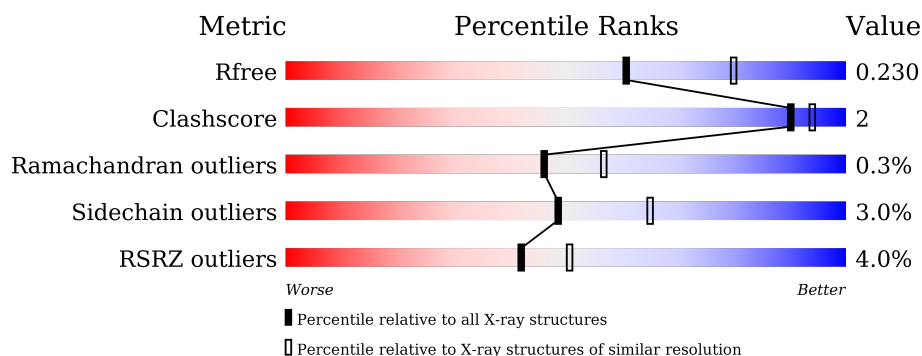
The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	unknown
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20028442
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-20028442

i

X-RAY DIFFRACTION

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	3852 (2.30-2.30)
Clashscore	102246	4452 (2.30-2.30)
Ramachandran outliers	100387	4410 (2.30-2.30)
Sidechain outliers	100360	4409 (2.30-2.30)
RSRZ outliers	91569	3857 (2.30-2.30)

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 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	379	<div> <div>5%</div> <div>89%</div> <div>8%</div> </div>
1	B	379	<div> <div>3%</div> <div>91%</div> <div>8%</div> </div>

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 5929 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 HDAH.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	372	Total	C	N	O	S	0	0	0
			2833	1781	522	518	12			
1	B	377	Total	C	N	O	S	0	0	0
			2867	1803	528	524	12			

- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	B	1	Total	Zn	0	0
			1	1		
2	A	1	Total	Zn	0	0
			1	1		

- Molecule 3 is POTASSIUM ION (three-letter code: K) (formula: K).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	2	Total	K	0	0
			2	2		
3	A	2	Total	K	0	0
			2	2		

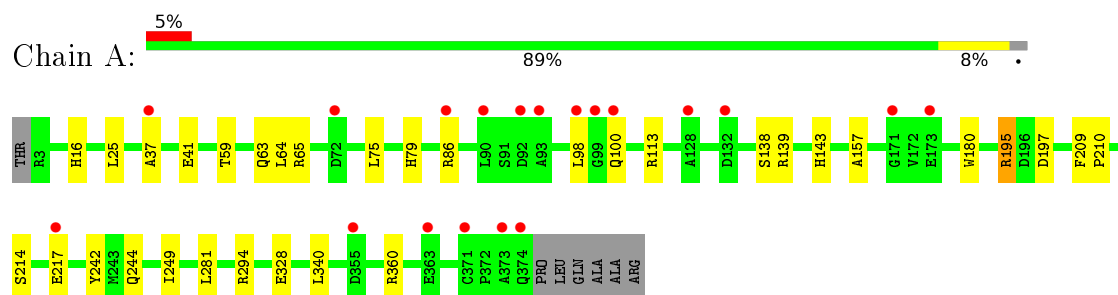
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	115	Total	O	0	0
			115	115		
4	B	108	Total	O	0	0
			108	108		

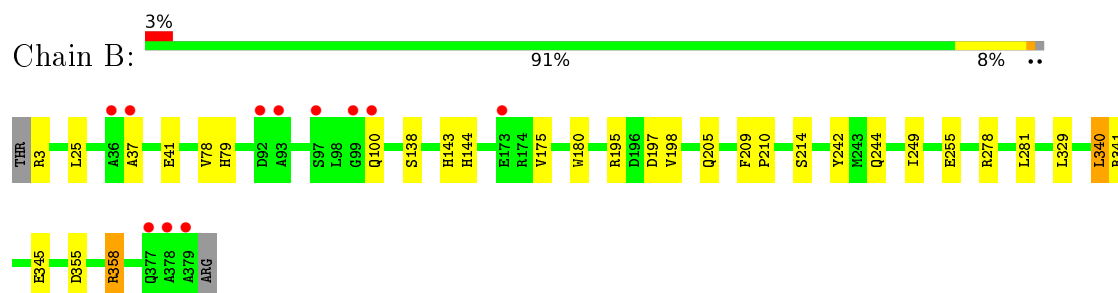
3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of 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 ($\text{RSRZ} > 2$). 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: HDAH



• Molecule 1: HDAH



4 Data and refinement statistics

Property	Value	Source
Space group	P 41 21 2	Depositor
Cell constants a, b, c, α , β , γ	81.53Å 81.53Å 203.22Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	75.67 – 2.29 40.77 – 2.29	Depositor EDS
% Data completeness (in resolution range)	99.8 (75.67-2.29) 99.9 (40.77-2.29)	Depositor EDS
R_{merge}	0.23	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.39 (at 2.29Å)	Xtriage
Refinement program	REFMAC	Depositor
R, R_{free}	0.187 , 0.227 0.193 , 0.230	Depositor DCC
R_{free} test set	1565 reflections (5.20%)	DCC
Wilson B-factor (Å ²)	20.8	Xtriage
Anisotropy	0.803	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.41 , 36.1	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	5929	wwPDB-VP
Average B, all atoms (Å ²)	25.0	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 36.61 % 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 4.8657e-04. 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](#)

Bond lengths and bond angles in the following residue types are not validated in this section: ZN, K

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	0.46	0/2903	0.73	1/3943 (0.0%)
1	B	0.46	0/2938	0.72	1/3992 (0.0%)
All	All	0.46	0/5841	0.73	2/7935 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	195	ARG	NE-CZ-NH2	-5.04	117.78	120.30
1	B	358	ARG	NE-CZ-NH2	-5.01	117.79	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	2833	0	2752	15	0
1	B	2867	0	2788	13	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
3	A	2	0	0	0	0
3	B	2	0	0	0	0
4	A	115	0	0	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	B	108	0	0	1	0
All	All	5929	0	5540	26	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 (26) 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:16:HIS:O	1:A:139:ARG:NH2	2.10	0.85
1:A:294:ARG:NH1	1:A:328:GLU:OE2	2.26	0.69
1:A:340:LEU:HD13	1:B:25:LEU:HD21	1.81	0.63
1:B:144:HIS:NE2	4:B:2047:HOH:O	2.31	0.63
1:A:86:ARG:NH2	1:A:98:LEU:O	2.34	0.61
1:B:205:GLN:HE22	1:B:278:ARG:H	1.49	0.60
1:B:78:VAL:HG22	1:B:79:HIS:CD2	2.42	0.54
1:A:25:LEU:HD21	1:B:340:LEU:HD13	1.89	0.54
1:A:360:ARG:NH1	4:A:2081:HOH:O	2.46	0.48
1:B:3:ARG:NH1	1:B:329:LEU:O	2.45	0.48
1:B:355:ASP:OD1	1:B:358:ARG:NH2	2.47	0.48
1:A:195:ARG:HD2	1:A:197:ASP:OD1	2.14	0.48
1:B:195:ARG:HD2	1:B:197:ASP:OD1	2.14	0.48
1:B:341:ARG:HH12	1:B:345:GLU:HB2	1.79	0.47
1:A:209:PHE:HA	1:A:210:PRO:C	2.34	0.47
1:B:242:TYR:CD1	1:B:281:LEU:HD21	2.51	0.46
1:B:205:GLN:NE2	1:B:278:ARG:H	2.13	0.46
1:A:59:THR:HG23	1:A:64:LEU:HD21	1.96	0.46
1:B:209:PHE:HA	1:B:210:PRO:C	2.37	0.45
1:A:100:GLN:HE21	1:A:100:GLN:HA	1.82	0.45
1:A:63:GLN:HE21	1:A:65:ARG:HD2	1.82	0.45
1:A:75:LEU:HD11	1:A:113:ARG:HG2	2.00	0.44
1:A:242:TYR:CD1	1:A:281:LEU:HD21	2.55	0.42
1:A:79:HIS:CE1	1:A:157:ALA:H	2.39	0.41
1:B:175:VAL:O	1:B:198:VAL:HA	2.21	0.41
1:A:217:GLU:H	1:A:217:GLU:CD	2.24	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	370/379 (98%)	356 (96%)	13 (4%)	1 (0%)	46	57
1	B	375/379 (99%)	360 (96%)	14 (4%)	1 (0%)	46	57
All	All	745/758 (98%)	716 (96%)	27 (4%)	2 (0%)	46	57

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	37	ALA
1	B	37	ALA

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	283/288 (98%)	276 (98%)	7 (2%)	55	73
1	B	286/288 (99%)	276 (96%)	10 (4%)	43	58
All	All	569/576 (99%)	552 (97%)	17 (3%)	48	65

All (17) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	41	GLU
1	A	138	SER
1	A	143	HIS
1	A	180	TRP

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	214	SER
1	A	244	GLN
1	A	249	ILE
1	B	41	GLU
1	B	100	GLN
1	B	138	SER
1	B	143	HIS
1	B	180	TRP
1	B	214	SER
1	B	244	GLN
1	B	249	ILE
1	B	255	GLU
1	B	340	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (13) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	79	HIS
1	A	100	GLN
1	A	228	ASN
1	A	230	ASN
1	A	239	HIS
1	A	309	HIS
1	B	39	HIS
1	B	63	GLN
1	B	100	GLN
1	B	205	GLN
1	B	228	ASN
1	B	247	GLN
1	B	282	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](#)

Of 6 ligands modelled in this entry, 6 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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, 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	372/379 (98%)	0.08	19 (5%)	32 41	13, 24, 43, 97	0
1	B	377/379 (99%)	-0.03	11 (2%)	55 64	13, 22, 40, 70	0
All	All	749/758 (98%)	0.02	30 (4%)	42 51	13, 23, 42, 97	0

All (30) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	378	ALA	5.9
1	A	373	ALA	4.7
1	A	92	ASP	4.5
1	A	37	ALA	4.3
1	B	100	GLN	4.2
1	A	93	ALA	4.0
1	A	355	ASP	3.9
1	A	171	GLY	3.4
1	A	86	ARG	3.3
1	A	100	GLN	3.2
1	B	93	ALA	3.2
1	B	97	SER	3.2
1	A	98	LEU	3.1
1	B	36	ALA	3.1
1	B	92	ASP	3.0
1	A	90	LEU	2.9
1	B	37	ALA	2.9
1	A	99	GLY	2.8
1	B	99	GLY	2.7
1	B	377	GLN	2.7
1	B	379	ALA	2.3
1	A	173	GLU	2.3
1	A	72	ASP	2.3
1	A	217	GLU	2.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	363	GLU	2.2
1	A	371	CYS	2.2
1	A	374	GLN	2.2
1	A	128	ALA	2.1
1	B	173	GLU	2.1
1	A	132	ASP	2.0

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](#)

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, 95th 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(Å ²)	Q<0.9
3	K	A	1376	1/1	0.98	0.03	-3.87	16,16,16,16	0
3	K	A	1377	1/1	1.00	0.06	-3.88	16,16,16,16	0
3	K	B	1381	1/1	0.99	0.07	-4.77	16,16,16,16	0
3	K	B	1382	1/1	0.99	0.04	-4.98	15,15,15,15	0
2	ZN	B	1380	1/1	0.97	0.04	-5.36	22,22,22,22	1
2	ZN	A	1375	1/1	0.96	0.05	-5.65	20,20,20,20	1

6.5 Other polymers [i](#)

There are no such residues in this entry.