



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 6, 2026 – 01:12 PM UTC

PDB ID : 5AWH / pdb_00005awh
Title : Rhodobacter sphaeroides Argonaute in complex with guide RNA/target DNA heteroduplex
Authors : Miyoshi, T.; Ito, K.; Murakami, R.; Uchiumi, T.
Deposited on : 2015-07-03
Resolution : 2.00 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 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	:	4-5-2 with Phenix2.0
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

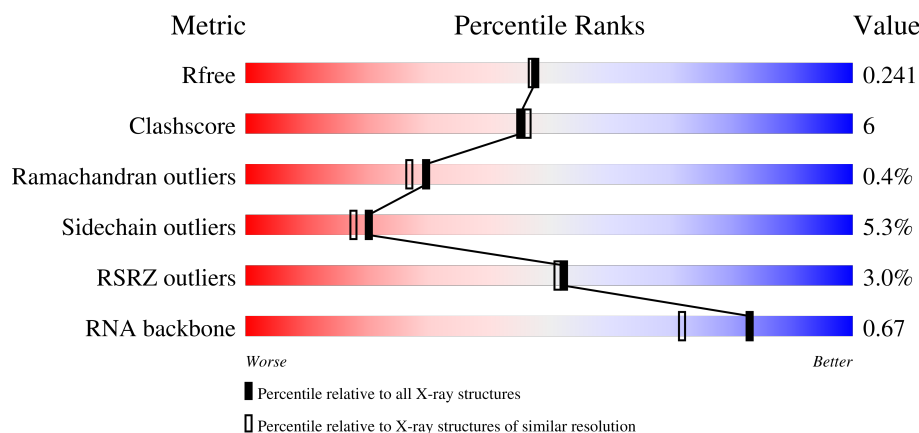
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.00 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}	180053	10052 (2.00-2.00)
Clashscore	190562	11152 (2.00-2.00)
Ramachandran outliers	187476	11031 (2.00-2.00)
Sidechain outliers	187428	11029 (2.00-2.00)
RSRZ outliers	180081	10067 (2.00-2.00)
RNA backbone	3983	1000 (2.36-1.64)

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 ≥ 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	778	3% 80% 15% . .
1	B	778	3% 79% 11% . 8%
2	C	18	56% 33% 11%
2	E	18	67% 22% 6% 6%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	D	18	 83%17%
3	F	18	 72%28%

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 13805 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 Uncharacterized protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	759	Total	C	N	O	S	0	0	0
			6005	3808	1084	1096	17			
1	B	715	Total	C	N	O	S	0	0	0
			5664	3600	1027	1022	15			

There are 42 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	MET	-	expression tag	UNP A4WYU7
A	1	GLY	-	expression tag	UNP A4WYU7
A	2	SER	-	expression tag	UNP A4WYU7
A	3	SER	-	expression tag	UNP A4WYU7
A	4	HIS	-	expression tag	UNP A4WYU7
A	5	HIS	-	expression tag	UNP A4WYU7
A	6	HIS	-	expression tag	UNP A4WYU7
A	7	HIS	-	expression tag	UNP A4WYU7
A	8	HIS	-	expression tag	UNP A4WYU7
A	9	HIS	-	expression tag	UNP A4WYU7
A	10	SER	-	expression tag	UNP A4WYU7
A	11	SER	-	expression tag	UNP A4WYU7
A	12	GLY	-	expression tag	UNP A4WYU7
A	13	LEU	-	expression tag	UNP A4WYU7
A	14	VAL	-	expression tag	UNP A4WYU7
A	15	PRO	-	expression tag	UNP A4WYU7
A	16	ALA	-	expression tag	UNP A4WYU7
A	17	GLY	-	expression tag	UNP A4WYU7
A	18	SER	-	expression tag	UNP A4WYU7
A	19	HIS	-	expression tag	UNP A4WYU7
A	20	MET	-	expression tag	UNP A4WYU7
B	0	MET	-	expression tag	UNP A4WYU7
B	1	GLY	-	expression tag	UNP A4WYU7
B	2	SER	-	expression tag	UNP A4WYU7
B	3	SER	-	expression tag	UNP A4WYU7

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	4	HIS	-	expression tag	UNP A4WYU7
B	5	HIS	-	expression tag	UNP A4WYU7
B	6	HIS	-	expression tag	UNP A4WYU7
B	7	HIS	-	expression tag	UNP A4WYU7
B	8	HIS	-	expression tag	UNP A4WYU7
B	9	HIS	-	expression tag	UNP A4WYU7
B	10	SER	-	expression tag	UNP A4WYU7
B	11	SER	-	expression tag	UNP A4WYU7
B	12	GLY	-	expression tag	UNP A4WYU7
B	13	LEU	-	expression tag	UNP A4WYU7
B	14	VAL	-	expression tag	UNP A4WYU7
B	15	PRO	-	expression tag	UNP A4WYU7
B	16	ALA	-	expression tag	UNP A4WYU7
B	17	GLY	-	expression tag	UNP A4WYU7
B	18	SER	-	expression tag	UNP A4WYU7
B	19	HIS	-	expression tag	UNP A4WYU7
B	20	MET	-	expression tag	UNP A4WYU7

- Molecule 2 is a RNA chain called RNA (5'-D(P*UP*UP*AP*CP*AP*AP*CP*CP*UP*AP*CP*UP*AP*CP*CP*UP*CP*G)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	18	Total	C	N	O	P	0	0	0
			374	168	61	127	18			
2	E	18	Total	C	N	O	P	0	0	0
			374	168	61	127	18			

- Molecule 3 is a DNA chain called DNA (5'-D(*CP*GP*AP*GP*GP*TP*AP*GP*TP*AP*GP*GP*TP*TP*GP*TP*AP*A)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	D	18	Total	C	N	O	P	0	0	0
			375	179	73	106	17			
3	F	18	Total	C	N	O	P	0	0	0
			375	179	73	106	17			

- Molecule 4 is MAGNESIUM ION (CCD ID: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	C	1	Total	Mg	0	0
			1	1		

Continued on next page...

Continued from previous page...

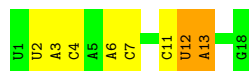
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	E	1	Total 1	Mg 1	0	0

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	292	Total 292	O 292	0	0
5	B	182	Total 182	O 182	0	0
5	C	50	Total 50	O 50	0	0
5	D	57	Total 57	O 57	0	0
5	E	30	Total 30	O 30	0	0
5	F	25	Total 25	O 25	0	0



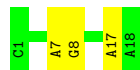
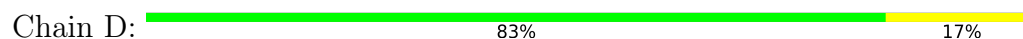
- Molecule 2: RNA (5'-D(P*UP*UP*AP*CP*AP*AP*CP*CP*UP*AP*CP*UP*AP*CP*CP*UP*CP*G)-3')



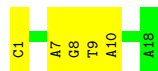
- Molecule 2: RNA (5'-D(P*UP*UP*AP*CP*AP*AP*CP*CP*UP*AP*CP*UP*AP*CP*CP*UP*CP*G)-3')



- Molecule 3: DNA (5'-D(*CP*GP*AP*GP*GP*TP*AP*GP*TP*AP*GP*GP*TP*TP*GP*TP*AP*A)-3')



- Molecule 3: DNA (5'-D(*CP*GP*AP*GP*GP*TP*AP*GP*TP*AP*GP*GP*TP*TP*GP*TP*AP*A)-3')



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	68.19Å 118.30Å 118.48Å 90.00° 95.83° 90.00°	Depositor
Resolution (Å)	35.63 – 2.00 35.63 – 2.00	Depositor EDS
% Data completeness (in resolution range)	97.1 (35.63-2.00) 97.6 (35.63-2.00)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.56 (at 2.00Å)	Xtriage
Refinement program	REFMAC 5.7.0032	Depositor
R, R_{free}	0.184 , 0.233 0.194 , 0.241	Depositor DCC
R_{free} test set	6232 reflections (4.89%)	wwPDB-VP
Wilson B-factor (Å ²)	39.9	Xtriage
Anisotropy	0.074	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.36 , 51.2	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	13805	wwPDB-VP
Average B, all atoms (Å ²)	58.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.22% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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: MG

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	1.18	9/6136 (0.1%)	1.11	8/8305 (0.1%)
1	B	1.04	2/5787 (0.0%)	1.07	12/7826 (0.2%)
2	C	0.82	0/415	1.16	2/640 (0.3%)
2	E	0.63	0/415	1.09	2/640 (0.3%)
3	D	0.56	0/422	1.01	2/652 (0.3%)
3	F	0.46	0/422	0.95	1/652 (0.2%)
All	All	1.07	11/13597 (0.1%)	1.08	27/18715 (0.1%)

The worst 5 of 11 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	117	ARG	C-O	-9.50	1.19	1.23
1	A	468	SER	C-O	-7.90	1.15	1.24
1	A	549	THR	CA-C	7.34	1.61	1.52
1	A	314	ASN	N-CA	5.71	1.52	1.46
1	A	394	PHE	CA-CB	5.35	1.57	1.53

The worst 5 of 27 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	C	12	U	C4'-C3'-O3'	-8.32	100.52	113.00
1	B	196	GLY	N-CA-C	-7.85	105.35	114.69
1	B	641	HIS	N-CA-C	-7.18	101.68	108.07
1	B	31	LEU	CA-C-N	-6.12	113.47	119.78
1	B	31	LEU	C-N-CA	-6.12	113.47	119.78

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts ⓘ

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	6005	0	6039	87	0
1	B	5664	0	5713	57	0
2	C	374	0	194	5	0
2	E	374	0	194	3	0
3	D	375	0	205	1	0
3	F	375	0	205	3	0
4	C	1	0	0	0	0
4	E	1	0	0	0	0
5	A	292	0	0	9	0
5	B	182	0	0	4	0
5	C	50	0	0	0	0
5	D	57	0	0	0	0
5	E	30	0	0	0	0
5	F	25	0	0	0	0
All	All	13805	0	12550	152	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.

The worst 5 of 152 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:B:148:VAL:HG12	1:B:176:MET:CE	1.99	0.93
1:B:606:ARG:HG2	5:B:966:HOH:O	1.78	0.84
1:B:363:ARG:H	1:B:443:HIS:HD2	1.24	0.83
1:B:363:ARG:H	1:B:443:HIS:CD2	1.99	0.80
1:A:610:ARG:CZ	5:A:802:HOH:O	2.33	0.77

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	757/778 (97%)	728 (96%)	25 (3%)	4 (0%)	24	21
1	B	705/778 (91%)	686 (97%)	17 (2%)	2 (0%)	36	35
All	All	1462/1556 (94%)	1414 (97%)	42 (3%)	6 (0%)	30	27

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	518	LYS
1	A	76	GLY
1	A	518	LYS
1	A	522	ASP
1	B	768	ILE

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	643/658 (98%)	608 (95%)	35 (5%)	20	17
1	B	605/658 (92%)	574 (95%)	31 (5%)	21	19
All	All	1248/1316 (95%)	1182 (95%)	66 (5%)	20	18

5 of 66 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	359	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	479	GLN
1	B	769	LYS
1	A	421	ARG
1	A	395	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 32 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	561	ASN
1	B	590	ASN
1	A	479	GLN
1	A	465	HIS
1	B	686	ASN

5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
2	C	17/18 (94%)	0	0
2	E	17/18 (94%)	2 (11%)	0
All	All	34/36 (94%)	2 (5%)	0

All (2) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
2	E	12	U
2	E	13	A

There are no RNA pucker outliers to report.

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 ⓘ

There are no oligosaccharides in this entry.

5.6 Ligand geometry

Of 2 ligands modelled in this entry, 2 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

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 [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, 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	759/778 (97%)	0.27	20 (2%) 57 56	29, 51, 85, 110	0
1	B	715/778 (91%)	0.49	27 (3%) 44 43	34, 62, 99, 130	0
2	C	18/18 (100%)	-0.76	0 100 100	34, 41, 62, 80	0
2	E	18/18 (100%)	-0.43	0 100 100	43, 53, 70, 86	0
3	D	18/18 (100%)	-0.33	0 100 100	39, 47, 74, 99	0
3	F	18/18 (100%)	-0.02	0 100 100	41, 64, 74, 106	0
All	All	1546/1628 (94%)	0.34	47 (3%) 52 51	29, 56, 92, 130	0

The worst 5 of 47 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	239	ALA	3.7
1	B	438	GLY	3.6
1	B	519	ALA	3.5
1	A	519	ALA	3.4
1	B	694	ALA	3.4

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 oligosaccharides 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. 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	B-factors(Å ²)	Q<0.9
4	MG	E	101	1/1	0.98	0.04	46,46,46,46	0
4	MG	C	101	1/1	0.99	0.03	33,33,33,33	0

6.5 Other polymers [i](#)

There are no such residues in this entry.