



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 8, 2026 – 03:30 PM UTC

PDB ID : 6NIT / pdb_00006nit
Title : Human Argonaute2-miR-122 bound to a target RNA with four central mismatches (bu4)
Authors : Sheu-Gruttadauria, J.; MacRae, I.J.
Deposited on : 2018-12-31
Resolution : 3.80 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

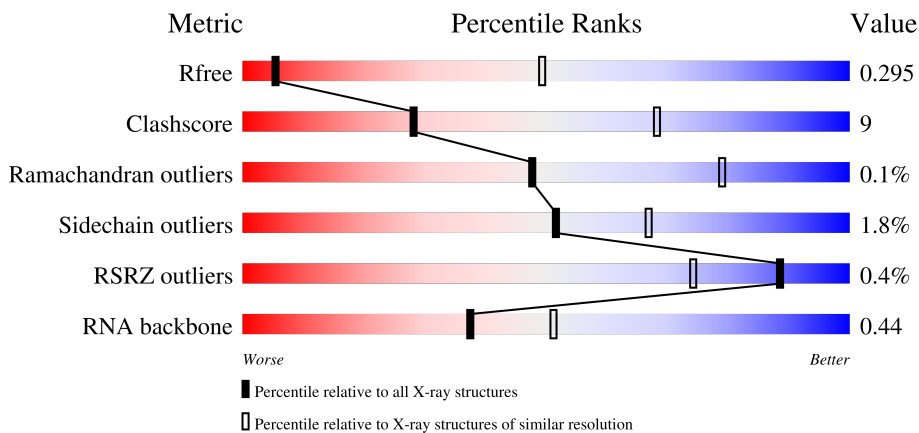
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.80 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.



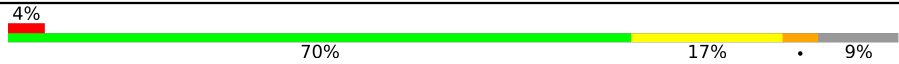

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	180053	1065 (3.96-3.64)
Clashscore	190562	1012 (3.94-3.66)
Ramachandran outliers	187476	1048 (3.96-3.64)
Sidechain outliers	187428	1043 (3.96-3.64)
RSRZ outliers	180081	1064 (3.96-3.64)
RNA backbone	3983	1007 (4.50-3.10)

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	859	 71% 21% • 7%
1	B	859	 73% 19% 8%
2	C	21	 33% 38% 14% 14%
2	D	21	 10% 33% 14% 5% 38%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	E	23	 4% 70% 17% 9%
3	F	23	 4% 52% 17% 13%

2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 14206 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 Protein argonaute-2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	798	Total 6398	C 4073	N 1151	O 1133	S 41	0	0	0
1	B	793	Total 6360	C 4051	N 1143	O 1126	S 40	0	0	0

There are 12 discrepancies between the modelled and reference sequences:

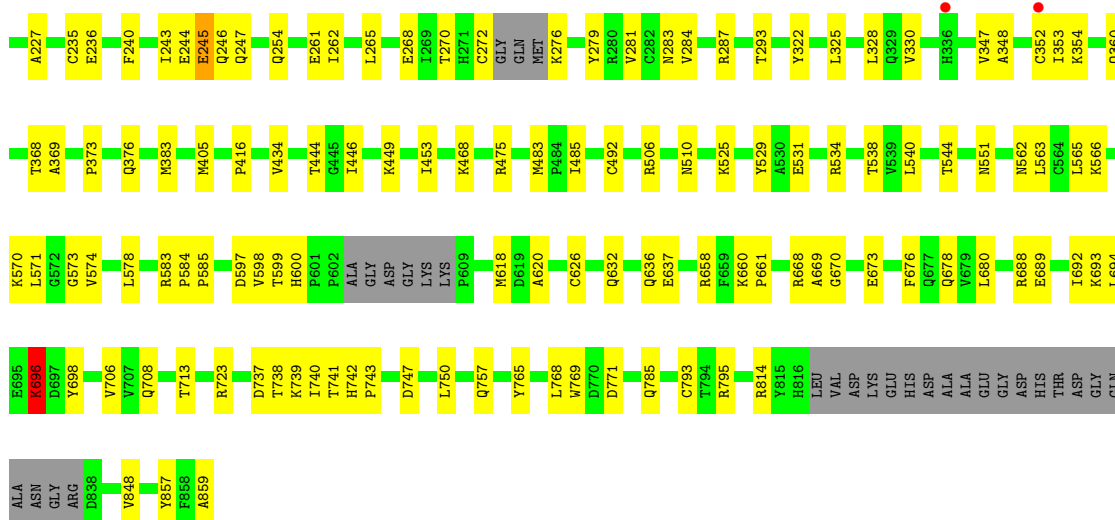
Chain	Residue	Modelled	Actual	Comment	Reference
A	387	ASP	SER	engineered mutation	UNP Q9UKV8
A	669	ALA	ASP	engineered mutation	UNP Q9UKV8
A	824	ALA	SER	engineered mutation	UNP Q9UKV8
A	828	ASP	SER	engineered mutation	UNP Q9UKV8
A	831	ASP	SER	engineered mutation	UNP Q9UKV8
A	834	ALA	SER	engineered mutation	UNP Q9UKV8
B	387	ASP	SER	engineered mutation	UNP Q9UKV8
B	669	ALA	ASP	engineered mutation	UNP Q9UKV8
B	824	ALA	SER	engineered mutation	UNP Q9UKV8
B	828	ASP	SER	engineered mutation	UNP Q9UKV8
B	831	ASP	SER	engineered mutation	UNP Q9UKV8
B	834	ALA	SER	engineered mutation	UNP Q9UKV8

- Molecule 2 is a RNA chain called RNA (5'-R(P*UP*GP*GP*AP*GP*UP*GP*UP*GP*AP*CP*AP*AP*UP*GP*GP*UP*GP*UP*UP*U)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
2	C	18	Total 372	C 163	N 64	O 127	P 18	0	0	0
2	D	13	Total 267	C 116	N 48	O 90	P 13	0	0	0

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
3	E	21	406	180	73	132	21	0	0	0
3	F	20	403	180	74	129	20	0	0	0



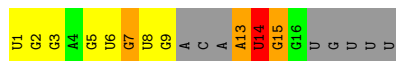
- Molecule 2: RNA (5'-R(P*UP*GP*GP*AP*GP*UP*GP*UP*GP*AP*CP*AP*AP*UP*GP*GP*UP*GP*UP*UP*U)-3')

Chain C: 33% 38% 14% 14%



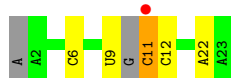
- Molecule 2: RNA (5'-R(P*UP*GP*GP*AP*GP*UP*GP*UP*GP*AP*CP*AP*AP*UP*GP*GP*UP*GP*UP*UP*U)-3')

Chain D: 10% 33% 14% 5% 38%



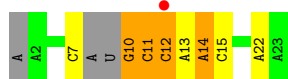
- Molecule 3: RNA (5'-R(P*AP*AP*CP*AP*CP*CP*AP*UP*CP*CP*AP*AP*CP*AP*CP*U*P*CP*CP*AP*AP*A)-3')

Chain E: 4% 70% 17% 9%



- Molecule 3: RNA (5'-R(P*AP*AP*CP*AP*CP*CP*AP*UP*CP*CP*AP*AP*CP*AP*CP*U*P*CP*CP*AP*AP*A)-3')

Chain F: 4% 52% 17% 17% 13%



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	106.71Å 137.09Å 154.06Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	39.30 – 3.80 39.30 – 3.80	Depositor EDS
% Data completeness (in resolution range)	98.9 (39.30-3.80) 99.1 (39.30-3.80)	Depositor EDS
R_{merge}	0.28	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.54 (at 3.76Å)	Xtrriage
Refinement program	PHENIX 1.9_1692	Depositor
R, R_{free}	0.243 , 0.294 0.253 , 0.295	Depositor DCC
R_{free} test set	1163 reflections (3.02%)	wwPDB-VP
Wilson B-factor (Å ²)	101.9	Xtrriage
Anisotropy	0.276	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.30 , 71.7	EDS
L-test for twinning ²	$\langle L \rangle = 0.46$, $\langle L^2 \rangle = 0.28$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.90	EDS
Total number of atoms	14206	wwPDB-VP
Average B, all atoms (Å ²)	117.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 28.95 % 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 1.6859e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

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](#)

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	0.31	0/6548	0.82	6/8862 (0.1%)
1	B	0.30	0/6509	0.80	6/8809 (0.1%)
2	C	0.10	0/414	0.26	0/641
2	D	0.19	0/297	0.58	1/459 (0.2%)
3	E	0.13	0/449	0.41	0/689
3	F	0.12	0/447	0.33	0/688
All	All	0.29	0/14664	0.77	13/20148 (0.1%)

There are no bond length outliers.

The worst 5 of 13 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	245	GLU	N-CA-C	-10.43	95.11	110.23
1	A	539	VAL	N-CA-C	6.52	116.55	110.42
1	B	24	LYS	CA-C-N	6.18	124.11	119.66
1	B	24	LYS	C-N-CA	6.18	124.11	119.66
1	A	97	ARG	N-CA-C	5.96	117.16	107.32

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	6398	0	6461	119	0
1	B	6360	0	6413	110	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	C	372	0	181	7	0
2	D	267	0	129	16	0
3	E	406	0	209	4	0
3	F	403	0	210	12	0
All	All	14206	0	13603	245	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 9.

The worst 5 of 245 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:566:LYS:NZ	1:B:859:ALA:OXT	1.77	1.18
1:B:551:ASN:ND2	2:D:2:G:OP2	2.02	0.91
1:B:531:GLU:OE2	1:B:534:ARG:NH1	2.06	0.88
1:A:384:ARG:NH1	1:B:658:ARG:HD3	1.91	0.84
1:A:551:ASN:ND2	2:C:2:G:OP2	2.11	0.83

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	784/859 (91%)	747 (95%)	36 (5%)	1 (0%)	48 79
1	B	777/859 (90%)	738 (95%)	38 (5%)	1 (0%)	48 79
All	All	1561/1718 (91%)	1485 (95%)	74 (5%)	2 (0%)	48 79

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	696	LYS
1	A	98	LYS

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	708/749 (94%)	695 (98%)	13 (2%)	51 67
1	B	703/749 (94%)	691 (98%)	12 (2%)	53 67
All	All	1411/1498 (94%)	1386 (98%)	25 (2%)	51 67

5 of 25 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	77	HIS
1	B	210	LEU
1	B	708	GLN
1	B	137	CYS
1	B	245	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 23 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	729	ASN
1	B	334	GLN
1	B	247	GLN
1	B	432	GLN
1	A	324	HIS

5.3.3 RNA [i](#)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
2	C	15/21 (71%)	6 (40%)	2 (13%)
2	D	11/21 (52%)	3 (27%)	2 (18%)

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
3	E	16/23 (69%)	1 (6%)	1 (6%)
3	F	17/23 (73%)	3 (17%)	2 (11%)
All	All	59/88 (67%)	13 (22%)	7 (11%)

5 of 13 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
2	C	13	A
2	C	14	U
2	C	15	G
2	C	17	U
2	C	18	G

5 of 7 RNA pucker outliers are listed below:

Mol	Chain	Res	Type
2	D	14	U
3	E	11	C
3	F	11	C
3	F	10	G
2	D	13	A

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 oligosaccharides 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 [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	798/859 (92%)	-0.01	3 (0%) 88 74	61, 111, 190, 247	0
1	B	793/859 (92%)	-0.07	2 (0%) 90 78	53, 97, 175, 249	0
2	C	18/21 (85%)	0.17	0 100 100	82, 142, 270, 275	0
2	D	13/21 (61%)	0.22	0 100 100	76, 90, 188, 203	0
3	E	21/23 (91%)	0.36	1 (4%) 35 26	115, 184, 224, 245	0
3	F	20/23 (86%)	0.57	1 (5%) 34 25	83, 157, 271, 299	0
All	All	1663/1806 (92%)	-0.02	7 (0%) 88 74	53, 104, 190, 299	0

The worst 5 of 7 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	454	ALA	5.3
1	A	520	VAL	3.3
1	B	336	HIS	2.4
3	E	11	C	2.4
1	B	352	CYS	2.3

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](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.