



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 22, 2020 – 09:54 am BST

PDB ID : 2F8T  
Title : Crystal structure of Aa-Ago with externally-bound siRNA  
Authors : Yuan, Y.R.; Chen, H.Y.; Patel, D.J.  
Deposited on : 2005-12-03  
Resolution : 3.10 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](mailto: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

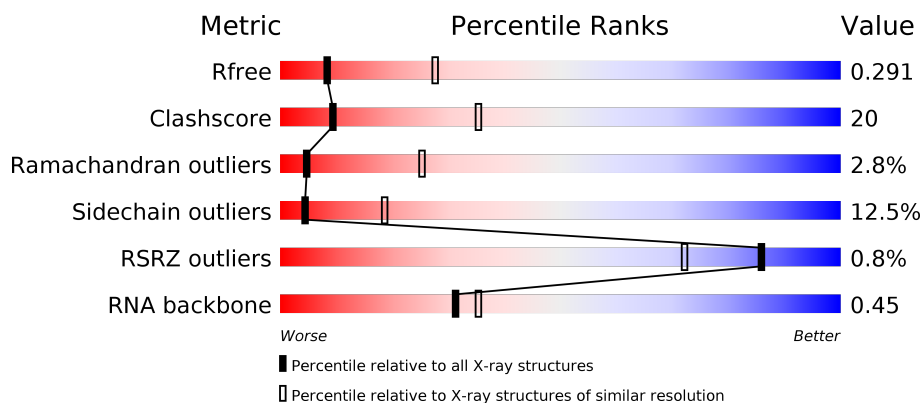
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.10 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	1094 (3.10-3.10)
Clashscore	141614	1184 (3.10-3.10)
Ramachandran outliers	138981	1141 (3.10-3.10)
Sidechain outliers	138945	1141 (3.10-3.10)
RSRZ outliers	127900	1067 (3.10-3.10)
RNA backbone	3102	1116 (3.40-2.80)

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 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	C	26	<div> <div>15%</div> <div>46%</div> <div>19%</div> <div>19%</div> </div>
1	D	26	<div> <div>15%</div> <div>38%</div> <div>35%</div> <div>8%</div> </div>
2	A	706	<div> <div>58%</div> <div>36%</div> <div>6%</div> </div>
2	B	706	<div> <div>52%</div> <div>39%</div> <div>8%</div> </div>

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 12841 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 RNA chain called 26-MER.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	C	26	Total	C	N	O	P	0	0	0
			551	246	94	185	26			
1	D	24	Total	C	N	O	P	0	0	0
			511	228	90	169	24			

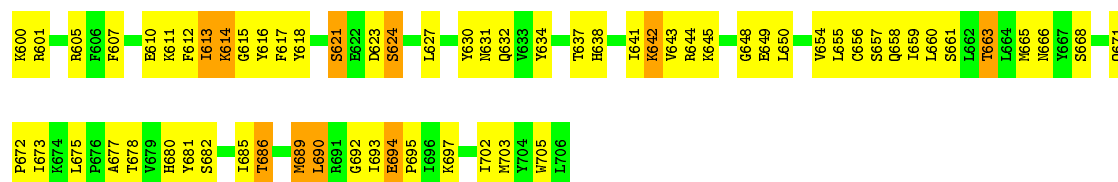
- Molecule 2 is a protein called Argonaute protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	A	704	Total	C	N	O	S	0	0	0
			5864	3818	993	1042	11			
2	B	704	Total	C	N	O	S	0	0	0
			5864	3818	993	1042	11			

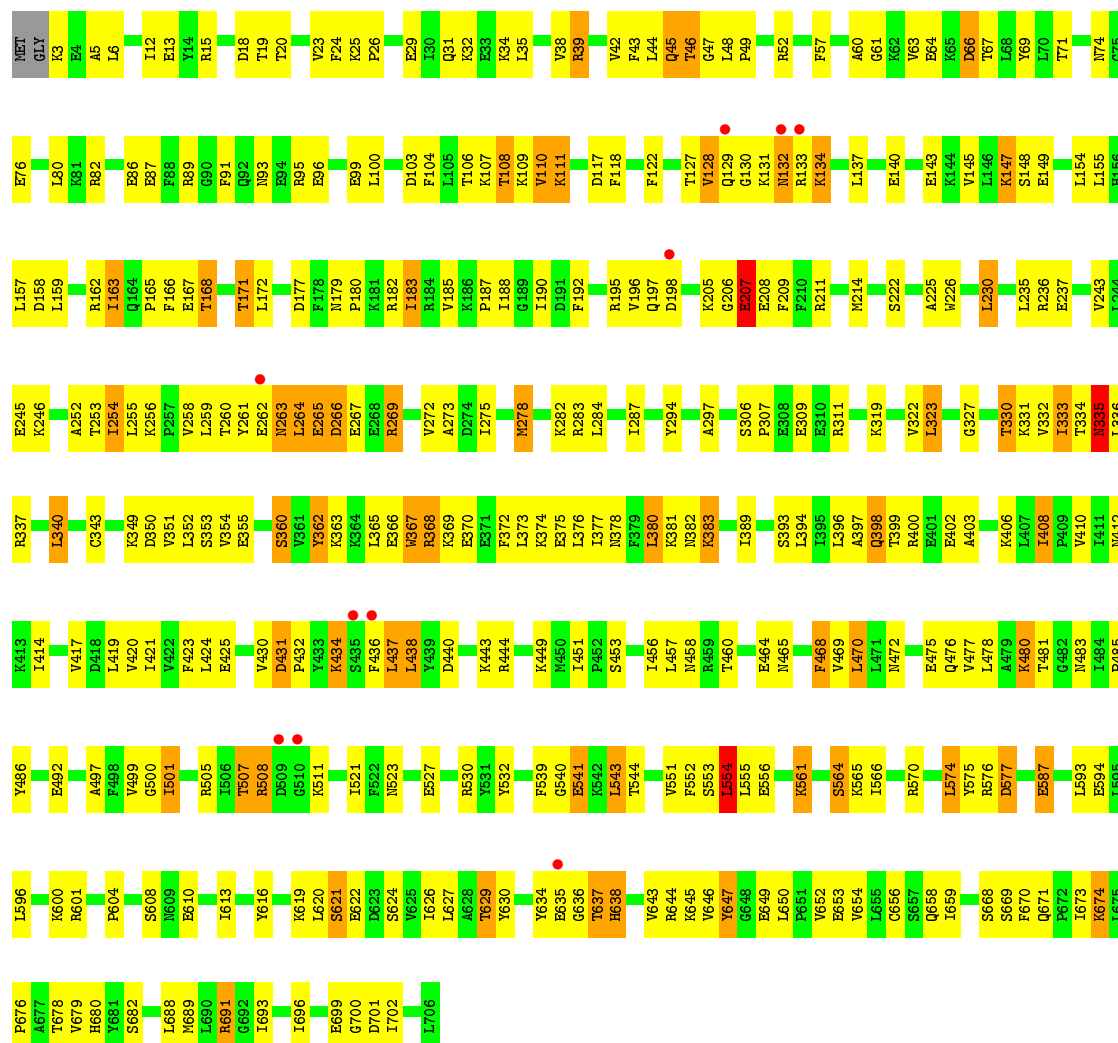
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	C	2	Total	O	0	0
			2	2		
3	D	2	Total	O	0	0
			2	2		
3	A	24	Total	O	0	0
			24	24		
3	B	23	Total	O	0	0
			23	23		





• Molecule 2: Argonaute protein



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	81.20 Å 118.08 Å 98.50 Å 90.00° 99.57° 90.00°	Depositor
Resolution (Å)	40.00 – 3.10 32.44 – 3.10	Depositor EDS
% Data completeness (in resolution range)	99.9 (40.00-3.10) 100.0 (32.44-3.10)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	4.71 (at 3.12 Å)	Xtriage
Refinement program	REFMAC 5.2.0005	Depositor
R, $R_{free}$	0.204 , 0.297 0.201 , 0.291	Depositor DCC
$R_{free}$ test set	1672 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	58.5	Xtriage
Anisotropy	0.033	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.28 , 61.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	12841	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	56.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.23% 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

### 5.1 Standard geometry

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	C	2.45	24/614 (3.9%)	2.20	31/952 (3.3%)
1	D	1.33	5/570 (0.9%)	1.75	14/884 (1.6%)
2	A	0.76	5/5979 (0.1%)	0.82	2/8027 (0.0%)
2	B	0.72	0/5979	0.81	2/8027 (0.0%)
All	All	0.93	34/13142 (0.3%)	1.00	49/17890 (0.3%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	C	0	1
2	A	0	1
2	B	0	4
All	All	0	6

The worst 5 of 34 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	2	G	N9-C8	21.71	1.53	1.37
1	C	2	G	C5-C4	16.53	1.50	1.38
1	C	1	A	N7-C5	16.30	1.49	1.39
1	C	1	A	N9-C4	14.03	1.46	1.37
1	C	2	G	N3-C4	13.16	1.44	1.35

The worst 5 of 49 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	2	G	C8-N9-C4	-17.28	99.49	106.40
1	C	2	G	C2-N3-C4	-13.63	105.08	111.90
1	C	2	G	N3-C2-N2	-11.04	112.17	119.90
1	C	2	G	N3-C4-N9	-10.15	119.91	126.00

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	2	G	N7-C8-N9	10.09	118.15	113.10

There are no chirality outliers.

5 of 6 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	A	335	ASN	Peptide
2	B	507	THR	Peptide
2	B	576	ARG	Peptide
2	B	621	SER	Peptide
1	C	1	A	Sidechain

## 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	C	551	0	279	25	0
1	D	511	0	259	31	0
2	A	5864	0	6128	210	0
2	B	5864	0	6128	251	0
3	A	24	0	0	3	0
3	B	23	0	0	1	0
3	C	2	0	0	0	0
3	D	2	0	0	1	0
All	All	12841	0	12794	504	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 20.

The worst 5 of 504 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:A:642:LYS:CD	2:A:642:LYS:CG	1.76	1.61
2:B:532:TYR:CZ	2:B:554:LEU:HD21	1.70	1.27
2:B:532:TYR:CE1	2:B:554:LEU:HD21	1.78	1.18

*Continued on next page...*



*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:A:66:ASP:HB3	3:A:708:HOH:O	1.51	1.11
2:B:122:PHE:CZ	2:B:275:ILE:HG21	1.87	1.09

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	
2	A	702/706 (99%)	605 (86%)	77 (11%)	20 (3%)	5	25
2	B	702/706 (99%)	605 (86%)	78 (11%)	19 (3%)	5	25
All	All	1404/1412 (99%)	1210 (86%)	155 (11%)	39 (3%)	5	25

5 of 39 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	A	434	LYS
2	A	435	SER
2	B	130	GLY
2	B	132	ASN
2	B	265	GLU

### 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	
2	A	646/647 (100%)	580 (90%)	66 (10%)	7	27
2	B	646/647 (100%)	550 (85%)	96 (15%)	3	13
All	All	1292/1294 (100%)	1130 (88%)	162 (12%)	4	18

5 of 162 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	64	GLU
2	B	166	PHE
2	B	629	THR
2	B	67	THR
2	B	117	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 16 such sidechains are listed below:

Mol	Chain	Res	Type
2	A	569	HIS
2	A	602	ASN
2	B	263	ASN
2	A	458	ASN
2	B	335	ASN

### 5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	C	25/26 (96%)	6 (24%)	0
1	D	23/26 (88%)	6 (26%)	1 (4%)
All	All	48/52 (92%)	12 (25%)	1 (2%)

5 of 12 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	C	3	A
1	C	6	G
1	C	9	U
1	C	10	G
1	C	11	C

All (1) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
1	D	23	C

#### 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 [i](#)

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, 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	C	26/26 (100%)	-0.10	0 100 100	31, 58, 79, 82	0
1	D	24/26 (92%)	0.23	0 100 100	26, 59, 89, 91	0
2	A	704/706 (99%)	-0.48	2 (0%) 94 88	27, 54, 82, 96	0
2	B	704/706 (99%)	-0.31	10 (1%) 75 56	30, 55, 78, 93	0
All	All	1458/1464 (99%)	-0.38	12 (0%) 86 72	26, 54, 80, 96	0

The worst 5 of 12 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	132	ASN	3.0
2	A	179	ASN	2.9
2	B	509	ASP	2.8
2	B	435	SER	2.5
2	B	129	GLN	2.5

### 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.