



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 02:18 AM GMT

PDB ID : 2GJE  
Title : Structure of a guideRNA-binding protein complex bound to a gRNA  
Authors : Schumacher, M.A.; Karamooz, E.; Zikova, A.; Trantirek, L.; Lukes, J.  
Deposited on : 2006-03-30  
Resolution : 3.37 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  
<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 : 1.7 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : **NOT EXECUTED**  
EDS : **NOT EXECUTED**  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk26865

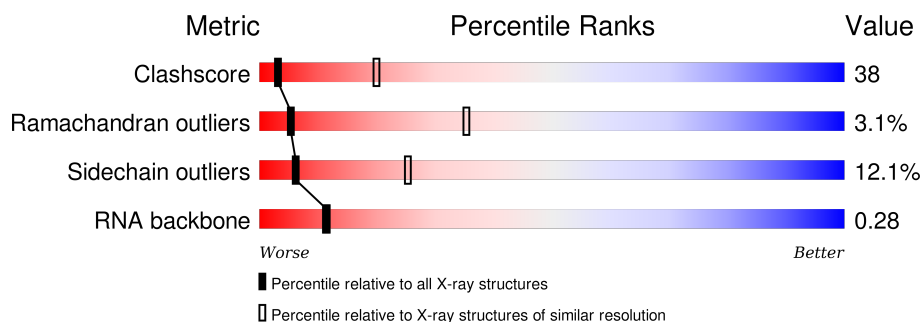
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.37 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))
Clashscore	102246	1158 (3.46-3.30)
Ramachandran outliers	100387	1139 (3.46-3.30)
Sidechain outliers	100360	1138 (3.46-3.30)
RNA backbone	2183	1024 (3.96-2.80)

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. 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.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	R	39	
2	S	5	
3	A	195	
4	D	187	

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 2769 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 guide RNA 40-mer.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	R	13	Total	C	N	O	P	0	0	0
			272	124	49	87	12			

- Molecule 2 is a RNA chain called RNA tetramer.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	S	5	Total	C	N	O	P	0	0	0
			105	48	19	34	4			

- Molecule 3 is a protein called mitochondrial RNA-binding protein 2.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
3	A	145	Total	C	N	O	S	Se	0	0	0
			1187	752	216	215	1	3			

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	110	MSE	MET	MODIFIED RESIDUE	UNP Q952G2
A	125	MSE	MET	MODIFIED RESIDUE	UNP Q952G2
A	204	MSE	MET	MODIFIED RESIDUE	UNP Q952G2

- Molecule 4 is a protein called mitochondrial RNA-binding protein 1.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
4	D	147	Total	C	N	O	S	Se	0	0	0
			1204	759	222	219	1	3			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	43	GLU	LEU	CONFLICT	UNP P90629

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
D	46	MSE	MET	MODIFIED RESIDUE	UNP P90629
D	109	MSE	MET	MODIFIED RESIDUE	UNP P90629
D	139	MSE	MET	MODIFIED RESIDUE	UNP P90629

- Molecule 5 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	D	1	Total O 1 1	0	0



4173	THR
	ARG
	ALA
	ALA
	GLU
	GLY
	GLY
	GLU
	LYS
	ILE
	ALA
	ALA
	THR
	SER
	SER
	ALA
	GLU
	GLY
	GLY
	ALA
	ARG
	ARG
	LYS
	ARG
	SER
	VAL
	SER
	ASP
	THR
	SER
	ARG
	TYR
	HIS

## 4 Data and refinement statistics

Xtriage (Phenix) and EDS were not executed - this section will therefore be incomplete.

Property	Value	Source
Space group	P 61 2 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	157.60 Å   157.60 Å   81.30 Å 90.00°   90.00°   120.00°	Depositor
Resolution (Å)	69.85 – 3.37	Depositor
% Data completeness (in resolution range)	91.8 (69.85-3.37)	Depositor
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
Refinement program	CNS 1.1	Depositor
R, $R_{free}$	0.282 , 0.300	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	2769	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	60.0	wwPDB-VP

## 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	R	0.57	0/304	1.86	5/471 (1.1%)
2	S	0.75	0/117	1.20	1/181 (0.6%)
3	A	0.59	0/1213	0.88	2/1643 (0.1%)
4	D	0.54	0/1231	0.78	1/1657 (0.1%)
All	All	0.57	0/2865	1.03	9/3952 (0.2%)

There are no bond length outliers.

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	R	19	C	O5'-P-OP2	-20.36	86.27	110.70
1	R	18	C	OP2-P-O3'	-18.18	65.21	105.20
1	R	18	C	OP1-P-O3'	-16.19	69.58	105.20
1	R	19	C	O5'-P-OP1	-15.56	91.69	105.70
2	S	40	U	O4'-C1'-N1	7.93	114.54	108.20

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	R	272	0	141	21	0
2	S	105	0	55	7	0
3	A	1187	0	1158	112	0
4	D	1204	0	1172	77	2

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	D	1	0	0	0	0
All	All	2769	0	2526	201	2

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 38.

The worst 5 of 201 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)
3:A:117:LYS:HD2	3:A:117:LYS:H	1.11	1.10
3:A:77:VAL:HG21	3:A:197:PHE:HE1	1.27	0.98
3:A:174:ARG:HH21	3:A:175:PRO:HG2	1.28	0.94
3:A:77:VAL:HG22	3:A:201:GLU:HG2	1.53	0.91
1:R:18:C:OP2	3:A:117:LYS:NZ	2.07	0.87

All (2) 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)
4:D:67:ARG:NH2	4:D:67:ARG:NH2[8_556]	1.89	0.31
4:D:67:ARG:NH1	4:D:67:ARG:NH1[8_556]	1.91	0.29

## 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	
3	A	141/195 (72%)	121 (86%)	16 (11%)	4 (3%)	6	37
4	D	145/187 (78%)	124 (86%)	16 (11%)	5 (3%)	5	33
All	All	286/382 (75%)	245 (86%)	32 (11%)	9 (3%)	5	35

5 of 9 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	A	161	ALA
4	D	28	SER
3	A	85	ILE
3	A	199	VAL
4	D	148	ASN

### 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	
3	A	127/159 (80%)	112 (88%)	15 (12%)	6	26
4	D	130/155 (84%)	114 (88%)	16 (12%)	6	24
All	All	257/314 (82%)	226 (88%)	31 (12%)	6	25

5 of 31 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	A	204	MSE
4	D	47	THR
4	D	137	GLN
4	D	27	GLN
4	D	69	VAL

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 8 such sidechains are listed below:

Mol	Chain	Res	Type
3	A	216	ASN
4	D	171	HIS
4	D	128	HIS
3	A	153	ASN
4	D	60	GLN

### 5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	R	12/39 (30%)	6 (50%)	0
2	S	4/5 (80%)	0	0
All	All	16/44 (36%)	6 (37%)	0

5 of 6 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	R	12	A
1	R	15	U
1	R	17	A
1	R	18	C
1	R	19	C

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 [i](#)

There are no chain breaks in this entry.

## 6 Fit of model and data [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

EDS was not executed - this section will therefore be empty.

### 6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

EDS was not executed - this section will therefore be empty.

### 6.3 Carbohydrates [i](#)

EDS was not executed - this section will therefore be empty.

### 6.4 Ligands [i](#)

EDS was not executed - this section will therefore be empty.

### 6.5 Other polymers [i](#)

EDS was not executed - this section will therefore be empty.