



# wwPDB X-ray Structure Validation Summary Report ⓘ

Nov 29, 2017 – 09:40 PM EST

PDB ID : 5VM9  
Title : Human Argonaute3 bound to guide RNA  
Authors : Park, M.S.; Nakanishi, K.  
Deposited on : unknown  
Resolution : 3.28 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
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20030345
Percentile statistics	:	20161228.v01 (using entries in the PDB archive December 28th 2016)
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-20030345



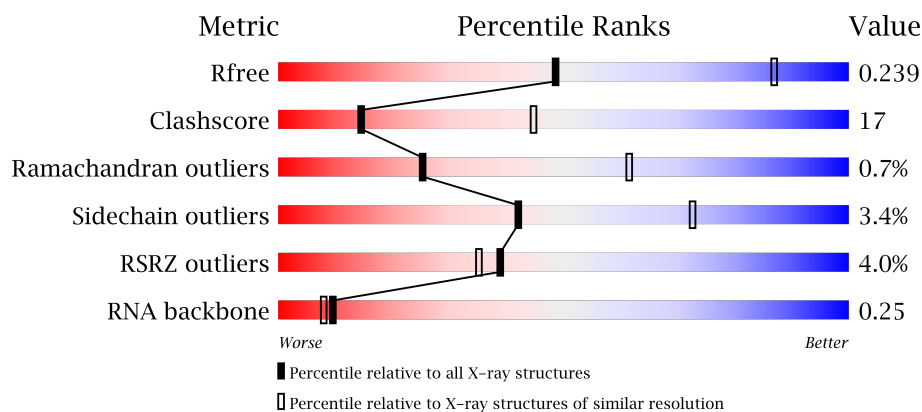
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 3.28 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))
$R_{free}$	100719	1006 (3.34-3.22)
Clashscore	112137	1070 (3.34-3.22)
Ramachandran outliers	110173	1050 (3.34-3.22)
Sidechain outliers	110143	1049 (3.34-3.22)
RSRZ outliers	101464	1011 (3.34-3.22)
RNA backbone	2435	1089 (3.76-2.80)

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.

Mol	Chain	Length	Quality of chain
1	A	862	<div> <div>4%</div> <div> <div></div> <div>55%</div> <div>34%</div> <div>• 10%</div> </div> </div>
1	C	862	<div> <div>3%</div> <div> <div></div> <div>54%</div> <div>36%</div> <div>• 9%</div> </div> </div>
2	B	13	<div> <div>8%</div> <div> <div></div> <div>23%</div> <div>54%</div> <div>23%</div> </div> </div>
3	D	15	<div> <div></div> <div> <div></div> <div>13%</div> <div>53%</div> <div>33%</div> </div> </div>



## 2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 13121 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-3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	779	Total	C	N	O	S	0	0	0
			6262	3992	1128	1102	40			
1	C	785	Total	C	N	O	S	0	0	0
			6295	4014	1132	1109	40			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	GLY	-	expression tag	UNP Q9H9G7
A	0	SER	-	expression tag	UNP Q9H9G7
C	-1	GLY	-	expression tag	UNP Q9H9G7
C	0	SER	-	expression tag	UNP Q9H9G7

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	13	Total	C	N	O	P	0	0	0
			265	118	54	80	13			

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

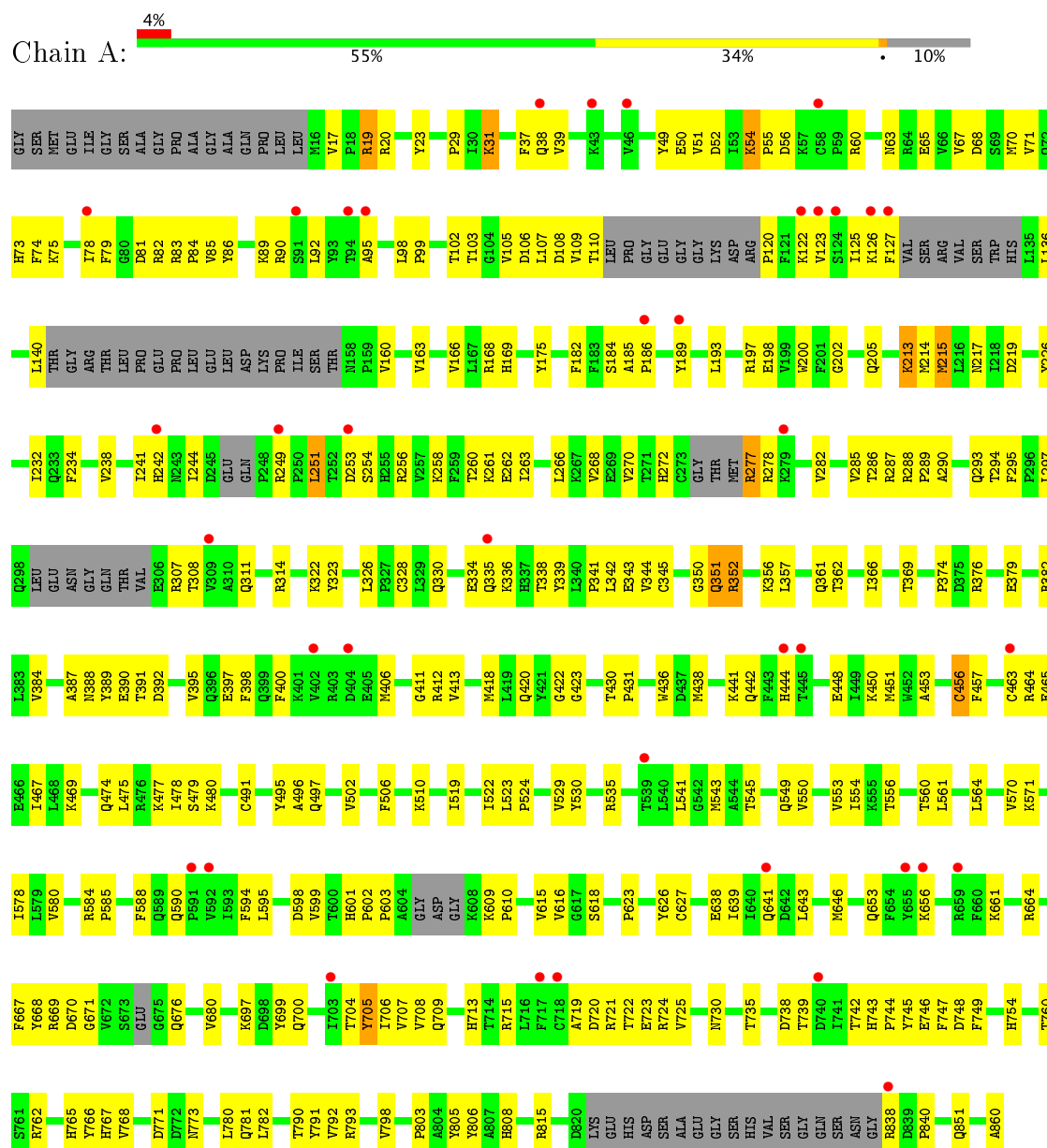
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	D	15	Total	C	N	O	P	0	0	0
			299	133	59	92	15			



### 3 Residue-property plots

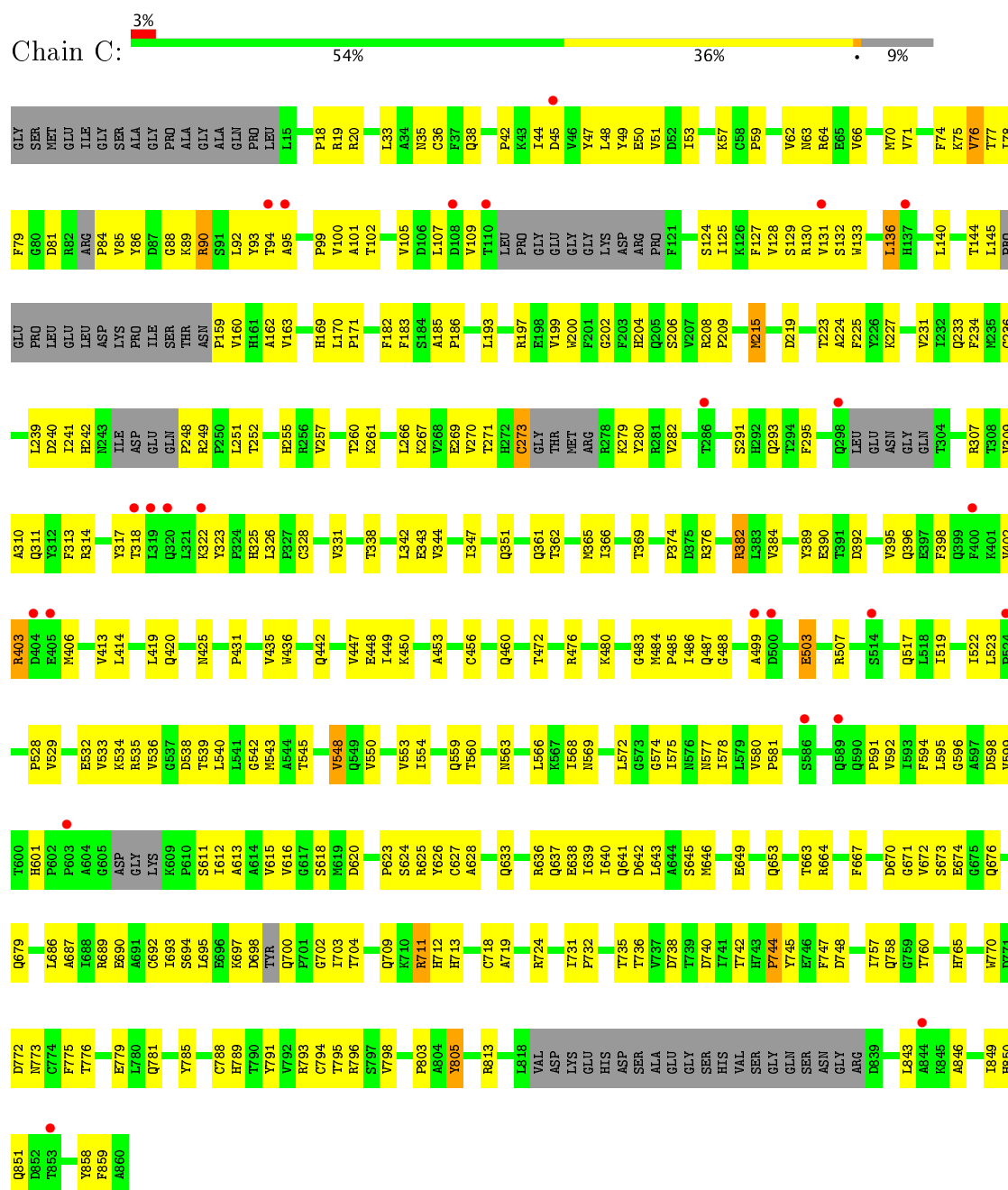
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are 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 ( $RSRZ > 2$ ). 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: Protein argonaute-3



#### • Molecule 1: Protein argonaute-3







A1	A2	A3	A4	A5	A6	A7	A8	A14	A18	A19	U20	U21
----	----	----	----	----	----	----	----	-----	-----	-----	-----	-----



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 43	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	68.41Å 68.41Å 408.80Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	67.48 – 3.28 64.88 – 3.28	Depositor EDS
% Data completeness (in resolution range)	100.0 (67.48-3.28) 100.0 (64.88-3.28)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.72 (at 3.26Å)	Xtriage
Refinement program	PHENIX (1.10.1_2155: ???)	Depositor
R, $R_{free}$	0.187 , 0.239 0.187 , 0.239	Depositor DCC
$R_{free}$ test set	1995 reflections (6.99%)	DCC
Wilson B-factor (Å <sup>2</sup> )	89.9	Xtriage
Anisotropy	0.094	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 82.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.34$ , $\langle L^2 \rangle = 0.17$	Xtriage
Estimated twinning fraction	0.339 for h,-k,-l	Xtriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	13121	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	97.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.99% 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 [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	0.30	0/6410	0.52	0/8677
1	C	0.32	0/6444	0.52	0/8725
2	B	0.75	1/297 (0.3%)	1.03	0/457
3	D	0.68	1/334 (0.3%)	0.89	0/512
All	All	0.34	2/13485 (0.0%)	0.55	0/18371

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	1	A	OP3-P	-10.65	1.48	1.61
3	D	1	A	OP3-P	-10.60	1.48	1.61

There are no bond angle outliers.

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	6262	0	6314	213	1
1	C	6295	0	6348	232	1
2	B	265	0	131	17	0
3	D	299	0	149	18	0
All	All	13121	0	12942	447	1

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 17.

The worst 5 of 447 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:38:GLN:HB3	1:A:213:LYS:HD2	1.44	0.98
1:C:128:VAL:HG12	1:C:129:SER:H	1.39	0.88
1:A:251:LEU:HD22	1:A:256:ARG:HB2	1.57	0.86
1:C:100:VAL:HG21	1:C:105:VAL:HG22	1.60	0.84
1:C:331:VAL:HB	1:C:338:THR:HB	1.59	0.84

All (1) 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)
1:A:697:LYS:NZ	1:C:636:ARG:O[4_545]	2.18	0.02

## 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	759/862 (88%)	693 (91%)	62 (8%)	4 (0%)	32	67
1	C	765/862 (89%)	705 (92%)	54 (7%)	6 (1%)	22	58
All	All	1524/1724 (88%)	1398 (92%)	116 (8%)	10 (1%)	25	61

5 of 10 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	251	LEU
1	C	271	THR
1	C	131	VAL
1	A	744	PRO
1	C	76	VAL



### 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	689/753 (92%)	661 (96%)	28 (4%)	35	69
1	C	693/753 (92%)	674 (97%)	19 (3%)	50	77
All	All	1382/1506 (92%)	1335 (97%)	47 (3%)	42	72

5 of 47 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	535	ARG
1	A	747	PHE
1	C	543	MET
1	A	641	GLN
1	A	805	TYR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 7 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	676	GLN
1	C	217	ASN
1	A	730	ASN
1	A	361	GLN
1	A	765	HIS

### 5.3.3 RNA [i](#)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
2	B	10/13 (76%)	3 (30%)	0
3	D	11/15 (73%)	5 (45%)	0
All	All	21/28 (75%)	8 (38%)	0

5 of 8 RNA backbone outliers are listed below:



Mol	Chain	Res	Type
2	B	6	A
2	B	8	A
2	B	20	U
3	D	8	A
3	D	18	A

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](#)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
3	D	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	D	8:A	O3'	14:A	P	15.64



## 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	A	779/862 (90%)	0.23	38 (4%) 30 27	52, 92, 148, 221	0
1	C	785/862 (91%)	0.15	25 (3%) 48 45	44, 92, 142, 202	0
2	B	13/13 (100%)	1.04	1 (7%) 14 13	69, 88, 219, 254	0
3	D	15/15 (100%)	-0.03	0 100 100	73, 92, 159, 166	0
All	All	1592/1752 (90%)	0.19	64 (4%) 39 35	44, 92, 147, 254	0

The worst 5 of 64 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	108	ASP	6.4
2	B	10	A	5.5
1	C	94	THR	5.2
1	A	655	TYR	4.9
1	C	319	LEU	4.8

### 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.