



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 06:13 PM GMT

PDB ID : 4KXT  
Title : Structure of human ARGONAUTE1 in complex with guide RNA  
Authors : Nakanishi, K.; Patel, D.J.  
Deposited on : 2013-05-28  
Resolution : 2.29 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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) : 1.9-1692  
EDS : rb-20026688  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
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) : trunk26865

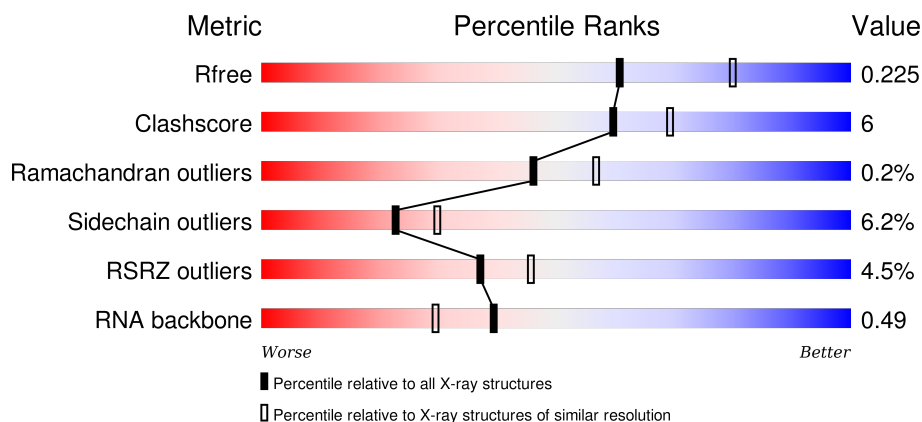
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.29 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))
$R_{free}$	91344	3852 (2.30-2.30)
Clashscore	102246	4452 (2.30-2.30)
Ramachandran outliers	100387	4410 (2.30-2.30)
Sidechain outliers	100360	4409 (2.30-2.30)
RSRZ outliers	91569	3857 (2.30-2.30)
RNA backbone	2183	1011 (2.84-1.76)

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>76%</div> <div>15%</div> <div>6%</div> </div>
2	B	12	<div> <div>83%</div> <div>8%</div> <div>8%</div> </div>

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 6983 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-1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	806	6473	4132	1153	1149	39	0	0	0

There are 5 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-4	GLY	-	EXPRESSION TAG	UNP Q9UL18
A	-3	ALA	-	EXPRESSION TAG	UNP Q9UL18
A	-2	MET	-	EXPRESSION TAG	UNP Q9UL18
A	-1	GLY	-	EXPRESSION TAG	UNP Q9UL18
A	0	SER	-	EXPRESSION TAG	UNP Q9UL18

- Molecule 2 is a RNA chain called RNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
2	B	12	222	98	44	69	11	0	0	1

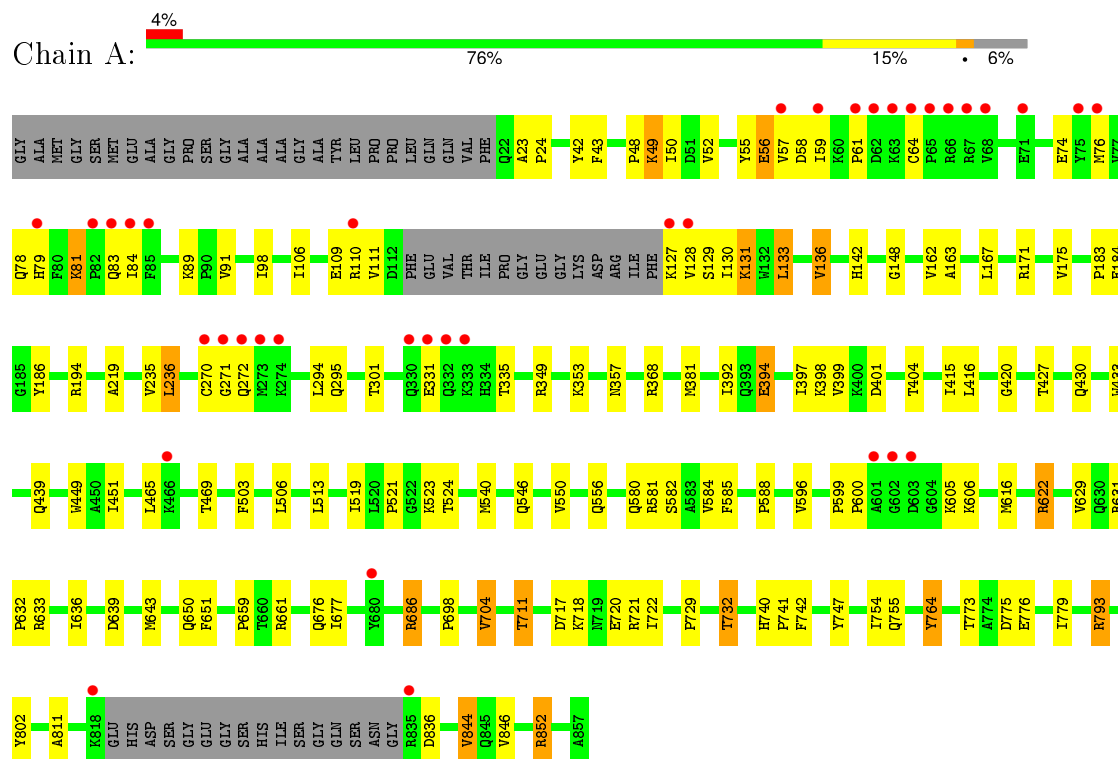
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	276	Total	O	0	0
			276	276		
3	B	12	Total	O	0	0
			12	12		

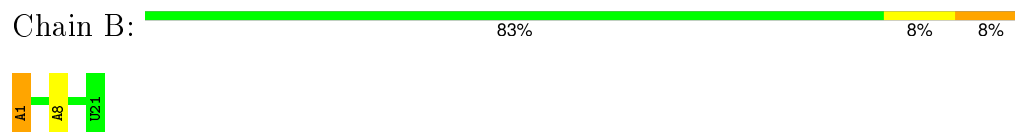
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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-1



#### • Molecule 2: RNA



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	69.88Å 100.25Å 136.70Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.86 – 2.29 48.86 – 2.29	Depositor EDS
% Data completeness (in resolution range)	99.5 (48.86-2.29) 96.5 (48.86-2.29)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.54 (at 2.29Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8.2_1309)	Depositor
R, $R_{free}$	0.175 , 0.228 0.170 , 0.225	Depositor DCC
$R_{free}$ test set	1933 reflections (4.58%)	DCC
Wilson B-factor (Å <sup>2</sup> )	40.0	Xtriage
Anisotropy	0.533	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 48.0	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	1 of 43579 reflections (0.002%)	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	6983	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	38.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.04% 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, 0.375 respectively for untwinned datasets, and 0.333, 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	0.40	0/6628	0.56	0/8980
2	B	0.88	1/248 (0.4%)	1.03	0/382
All	All	0.42	1/6876 (0.0%)	0.59	0/9362

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	1	A	OP3-P	-11.38	1.47	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	6473	0	6552	76	0
2	B	222	0	109	2	0
3	A	276	0	0	3	0
3	B	12	0	0	0	0
All	All	6983	0	6661	76	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.

All (76) 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:616:MET:SD	3:A:1085:HOH:O	2.33	0.87
1:A:639:ASP:HB2	1:A:643:MET:HE2	1.59	0.82
1:A:91:VAL:HG11	1:A:163:ALA:HA	1.64	0.80
1:A:83:GLN:HG3	1:A:84:ILE:HG23	1.66	0.78
1:A:89:LYS:HB2	1:A:394:GLU:HG3	1.70	0.73
1:A:616:MET:HE1	1:A:650:GLN:HG3	1.71	0.71
1:A:773:THR:HG22	1:A:776:GLU:HB2	1.72	0.70
1:A:127:LYS:N	3:A:1012:HOH:O	2.25	0.70
1:A:57:VAL:HG12	1:A:130:ILE:HG12	1.77	0.66
1:A:91:VAL:HG12	1:A:162:VAL:HG12	1.78	0.65
1:A:503:PHE:HB3	1:A:540:MET:HE1	1.78	0.65
1:A:76:MET:HE1	1:A:128:VAL:HG11	1.80	0.63
1:A:56:GLU:HB2	1:A:133:LEU:HD23	1.80	0.63
1:A:449:TRP:HZ3	1:A:451:ILE:HD11	1.64	0.63
1:A:56:GLU:OE1	1:A:110:ARG:NH2	2.31	0.63
1:A:793:ARG:NH1	3:A:924:HOH:O	2.32	0.62
1:A:42:TYR:O	1:A:404:THR:HG23	1.99	0.62
1:A:599:PRO:O	1:A:633:ARG:NH1	2.34	0.61
1:A:392:ILE:HG23	1:A:397:ILE:HB	1.83	0.60
1:A:773:THR:HG23	1:A:776:GLU:H	1.67	0.60
1:A:616:MET:HB2	1:A:622:ARG:HG2	1.84	0.59
1:A:729:PRO:O	1:A:732:THR:HG23	2.02	0.59
1:A:519:ILE:HD13	1:A:550:VAL:HG21	1.85	0.59
1:A:754:ILE:HG22	1:A:755:GLN:HG3	1.85	0.58
1:A:175:VAL:HG11	1:A:349:ARG:HD3	1.84	0.58
1:A:420:GLY:HA3	1:A:439:GLN:HG3	1.85	0.57
1:A:58:ASP:OD2	1:A:129:SER:OG	2.21	0.57
1:A:465:LEU:HD11	1:A:519:ILE:HD12	1.86	0.57
1:A:686:ARG:NH2	1:A:698:PRO:O	2.39	0.56
1:A:523:LYS:HA	2:B:1:A:C2	2.42	0.55
1:A:584:VAL:HG23	1:A:661:ARG:HD2	1.89	0.54
1:A:235:VAL:HG23	1:A:236:LEU:HD13	1.89	0.53
1:A:59:ILE:HG12	1:A:128:VAL:HG22	1.92	0.52
1:A:91:VAL:CG1	1:A:162:VAL:HG12	2.38	0.52
1:A:622:ARG:NH1	1:A:852:ARG:O	2.45	0.50
1:A:651:PHE:CD2	1:A:659:PRO:HD3	2.47	0.49
1:A:811:ALA:HB1	1:A:844:VAL:HG22	1.93	0.49
1:A:449:TRP:CZ3	1:A:451:ILE:HD11	2.47	0.49
1:A:106:ILE:HD11	1:A:111:VAL:HG11	1.94	0.49
1:A:331:GLU:N	1:A:331:GLU:OE1	2.46	0.49
1:A:718:LYS:O	1:A:721:ARG:HG3	2.13	0.48
1:A:353:LYS:HB2	1:A:353:LYS:HE3	1.70	0.48

Continued on next page...

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:55:TYR:HB2	1:A:98:ILE:HG23	1.95	0.48
1:A:616:MET:CE	1:A:650:GLN:HG3	2.42	0.48
1:A:295:GLN:HA	1:A:301:THR:HG22	1.96	0.47
1:A:588:PRO:HB2	1:A:659:PRO:HA	1.96	0.47
1:A:219:ALA:O	2:B:8:A:H5'	2.16	0.46
1:A:48:PRO:HB2	1:A:50:ILE:HG23	1.97	0.46
1:A:74:GLU:OE1	1:A:78:GLN:NE2	2.48	0.46
1:A:61:PRO:HG2	1:A:64:CYS:SG	2.55	0.46
1:A:52:VAL:HG22	1:A:136:VAL:HG22	1.98	0.46
1:A:704:VAL:HG13	1:A:764:TYR:HB3	1.98	0.45
1:A:148:GLY:HA2	1:A:742:PHE:CD2	2.51	0.45
1:A:740:HIS:CG	1:A:741:PRO:HD2	2.51	0.45
1:A:773:THR:CG2	1:A:776:GLU:H	2.30	0.45
1:A:717:ASP:HB2	1:A:720:GLU:HG3	1.98	0.45
1:A:465:LEU:HA	1:A:465:LEU:HD12	1.74	0.45
1:A:639:ASP:O	1:A:643:MET:HG3	2.18	0.44
1:A:636:ILE:HD13	1:A:677:ILE:HA	2.00	0.44
1:A:43:PHE:CE1	1:A:381:MET:HG2	2.53	0.44
1:A:711:THR:HG21	1:A:747:TYR:OH	2.17	0.44
1:A:581:ARG:HB3	1:A:585:PHE:CD2	2.53	0.44
1:A:23:ALA:HB1	1:A:24:PRO:HD2	2.00	0.43
1:A:91:VAL:HG11	1:A:163:ALA:CA	2.42	0.43
1:A:416:LEU:HD13	1:A:433:TRP:CZ2	2.54	0.43
1:A:49:LYS:HD2	1:A:142:HIS:CB	2.48	0.43
1:A:631:ARG:HA	1:A:632:PRO:HD3	1.91	0.42
1:A:270:CYS:O	1:A:272:GLN:N	2.52	0.42
1:A:110:ARG:HG2	1:A:131:LYS:HD3	2.02	0.42
1:A:430:GLN:OE1	1:A:722:ILE:HD13	2.20	0.41
1:A:775:ASP:O	1:A:779:ILE:HG12	2.21	0.41
1:A:81:LYS:HG3	1:A:81:LYS:H	1.58	0.41
1:A:183:PRO:HA	1:A:186:TYR:CZ	2.56	0.41
1:A:521:PRO:HD2	1:A:524:THR:HG21	2.03	0.40
1:A:415:ILE:HG12	1:A:580:GLN:OE1	2.21	0.40
1:A:606:LYS:HA	1:A:606:LYS:HE2	2.04	0.40

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	800/862 (93%)	768 (96%)	30 (4%)	2 (0%)	46	57

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	271	GLY
1	A	600	PRO

### 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	709/748 (95%)	665 (94%)	44 (6%)	23	30

All (44) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	49	LYS
1	A	56	GLU
1	A	79	HIS
1	A	81	LYS
1	A	109	GLU
1	A	131	LYS
1	A	133	LEU
1	A	136	VAL
1	A	167	LEU

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	171	ARG
1	A	184	GLU
1	A	194	ARG
1	A	236	LEU
1	A	294	LEU
1	A	335	THR
1	A	357	ASN
1	A	368	ARG
1	A	394	GLU
1	A	398	LYS
1	A	399	VAL
1	A	401	ASP
1	A	427	THR
1	A	469	THR
1	A	506	LEU
1	A	513	LEU
1	A	546	GLN
1	A	556	GLN
1	A	582	SER
1	A	596	VAL
1	A	605	LYS
1	A	622	ARG
1	A	629	VAL
1	A	676	GLN
1	A	686	ARG
1	A	704	VAL
1	A	711	THR
1	A	732	THR
1	A	764	TYR
1	A	793	ARG
1	A	802	TYR
1	A	836	ASP
1	A	844	VAL
1	A	846	VAL
1	A	852	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	546	GLN

### 5.3.3 RNA [i](#)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
2	B	9/12 (75%)	0	0

There are no RNA backbone outliers to report.

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 ⓘ

### 6.1 Protein, DNA and RNA chains ⓘ

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	806/862 (93%)	0.16	37 (4%) 36 45	16, 33, 76, 112	0
2	B	12/12 (100%)	0.03	0 100 100	32, 47, 86, 92	0
All	All	818/874 (93%)	0.16	37 (4%) 37 46	16, 33, 77, 112	0

All (37) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	271	GLY	10.1
1	A	272	GLN	6.2
1	A	835	ARG	5.8
1	A	110	ARG	5.6
1	A	75	TYR	5.4
1	A	84	ILE	4.9
1	A	83	GLN	4.6
1	A	68	VAL	4.5
1	A	273	MET	4.1
1	A	67	ARG	4.0
1	A	128	VAL	3.8
1	A	127	LYS	3.7
1	A	333	LYS	3.5
1	A	601	ALA	3.3
1	A	82	PRO	3.3
1	A	818	LYS	3.2
1	A	62	ASP	3.1
1	A	331	GLU	3.0
1	A	63	LYS	3.0
1	A	332	GLN	2.7
1	A	61	PRO	2.7
1	A	66	ARG	2.5
1	A	680	TYR	2.3
1	A	71	GLU	2.3

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	59	ILE	2.3
1	A	602	GLY	2.3
1	A	603	ASP	2.3
1	A	466	LYS	2.2
1	A	57	VAL	2.2
1	A	270	CYS	2.2
1	A	76	MET	2.2
1	A	65	PRO	2.1
1	A	274	LYS	2.1
1	A	85	PHE	2.1
1	A	330	GLN	2.1
1	A	79	HIS	2.0
1	A	64	CYS	2.0

## 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.