Full wwPDB X-ray Structure Validation Report ⓘ

Jun 30, 2022 – 10:06 AM EDT

PDB ID	:	7R8H
Title	:	Crystal structure of Pseudooceanicola lipolyticus Argonaute bound to 5' p guide DNA
Authors	:	Shin, Y.; Murakami, K.S.
Deposited on	:	2021-06-26
Resolution	:	2.54 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*
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:

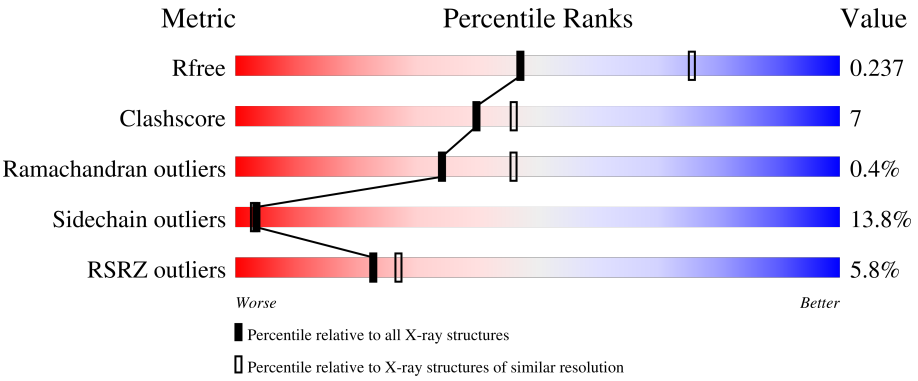
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.29
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.29

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.54 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	1284 (2.56-2.52)
Clashscore	141614	1332 (2.56-2.52)
Ramachandran outliers	138981	1315 (2.56-2.52)
Sidechain outliers	138945	1315 (2.56-2.52)
RSRZ outliers	127900	1272 (2.56-2.52)

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 <=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	789	5% 72% 22%
2	T	18	17% 50% 50%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 6408 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 Argonaute.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	769	Total	C	N	O	S	0	0	0
			6036	3832	1079	1096	29			

There are 5 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	13	THR	LEU	conflict	UNP A0A2M8J4C7
A	14	GLY	GLU	conflict	UNP A0A2M8J4C7
A	15	ALA	GLY	conflict	UNP A0A2M8J4C7
A	16	CYS	LEU	conflict	UNP A0A2M8J4C7
A	17	GLY	THR	conflict	UNP A0A2M8J4C7

- Molecule 2 is a DNA chain called DNA (5'-D(*TP*TP*AP*CP*TP*GP*CP*AP*CP*AP*GP*GP*TP*GP*AP*CP*GP*A)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	T	18	Total	C	N	O	P	0	0	0
			372	176	70	108	18			

3 Residue-property plots (i)

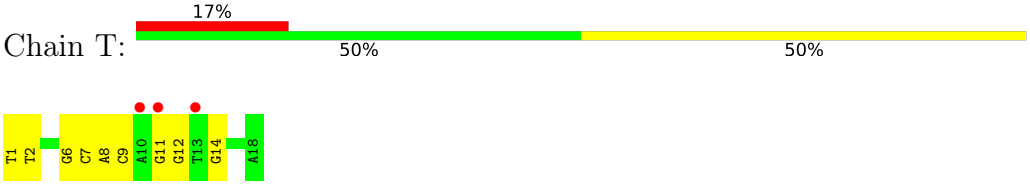
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. 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: Argonaute

Chain A:

• Molecule 2: DNA (5'-D(*TP*TP*AP*CP*TP*GP*CP*AP*CP*AP*GP*GP*TP*GP*AP*CP *GP*A)-3')

Chain T:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	I 2 3	Depositor
Cell constants a, b, c, α, β, γ	199.82Å 199.82Å 199.82Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	44.68 – 2.54 44.68 – 2.54	Depositor EDS
% Data completeness (in resolution range)	99.4 (44.68-2.54) 99.4 (44.68-2.54)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.32 (at 2.54Å)	Xtriage
Refinement program	PHENIX 1.20.1_4487	Depositor
R, R_{free}	0.200 , 0.239 0.200 , 0.237	Depositor DCC
R_{free} test set	1993 reflections (4.61%)	wwPDB-VP
Wilson B-factor (Å^2)	78.8	Xtriage
Anisotropy	0.000	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.28 , 54.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.025 for -l,-k,-h	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	6408	wwPDB-VP
Average B, all atoms (Å^2)	98.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.95% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ 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	A	0.47	0/6171	0.70	3/8345 (0.0%)
2	T	1.30	3/417 (0.7%)	1.05	3/640 (0.5%)
All	All	0.56	3/6588 (0.0%)	0.73	6/8985 (0.1%)

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	A	0	2

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	T	1	DT	OP3-P	-11.47	1.47	1.61
2	T	1	DT	C3'-O3'	-5.24	1.37	1.44
2	T	2	DT	C3'-O3'	-5.22	1.37	1.44

All (6) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	690	LEU	C-N-CA	-6.76	108.11	122.30
2	T	11	DG	O4'-C1'-N9	6.21	112.35	108.00
1	A	576	GLU	C-N-CA	-6.12	109.44	122.30
1	A	490	PHE	C-N-CA	-5.59	110.56	122.30
2	T	1	DT	N3-C4-O4	5.42	123.16	119.90
2	T	1	DT	C5-C4-O4	-5.37	121.14	124.90

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	577	GLY	Peptide
1	A	707	GLY	Peptide

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	6036	0	6072	89	0
2	T	372	0	203	5	0
All	All	6408	0	6275	92	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 7.

All (92) 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:746:PRO:HD2	1:A:752:MET:HE1	1.40	1.04
1:A:703:ILE:HG13	1:A:734:LEU:HD22	1.65	0.79
1:A:104:ASP:OD2	1:A:110:ARG:NH2	2.19	0.76
1:A:591:ARG:HG3	1:A:592:GLU:HG3	1.70	0.73
1:A:80:ARG:HA	1:A:85:MET:HG3	1.71	0.72
1:A:56:VAL:HG22	1:A:112:ARG:HD2	1.71	0.72
1:A:507:LYS:HG2	1:A:520:GLN:HB3	1.72	0.72
1:A:600:THR:O	1:A:604:THR:OG1	2.09	0.71
1:A:554:LEU:HD22	1:A:736:ASP:HB2	1.72	0.71
1:A:590:GLN:HG3	1:A:593:GLN:HB3	1.73	0.70
1:A:400:LYS:HE3	1:A:430:GLU:HB3	1.74	0.69
1:A:35:GLU:O	1:A:39:ASN:ND2	2.28	0.66
1:A:23:TYR:O	1:A:65:PHE:HA	1.96	0.65
1:A:293:GLN:O	1:A:296:THR:HG22	1.96	0.65
1:A:55:PHE:HE2	1:A:68:VAL:HG22	1.65	0.62
1:A:311:ILE:HD13	1:A:315:LEU:HD12	1.83	0.61
1:A:34:THR:HG21	1:A:595:LEU:HD22	1.83	0.60
1:A:573:MET:O	1:A:581:VAL:HA	2.02	0.59
1:A:754:LEU:HD13	1:A:758:ILE:HD11	1.83	0.59
2:T:7:DC:H2"	2:T:8:DA:C8	2.39	0.57

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:15:ALA:H	1:A:98:ARG:HH21	1.53	0.56
1:A:593:GLN:HG2	1:A:594:LEU:H	1.72	0.55
1:A:32:ASP:HB2	1:A:35:GLU:HB2	1.90	0.54
1:A:25:VAL:HG12	1:A:86:VAL:HB	1.88	0.53
1:A:3:LEU:HD11	1:A:692:SER:HB3	1.89	0.53
1:A:703:ILE:H	1:A:703:ILE:HD12	1.73	0.53
2:T:12:DG:H3'	2:T:14:DG:H22	1.73	0.53
1:A:320:ILE:O	1:A:323:ALA:HB3	2.09	0.52
1:A:629:ALA:N	1:A:660:GLU:OE1	2.43	0.51
1:A:55:PHE:CE2	1:A:68:VAL:HG22	2.45	0.49
1:A:40:VAL:O	1:A:44:VAL:HG23	2.11	0.49
1:A:23:TYR:CE2	1:A:90:PRO:HG3	2.48	0.49
1:A:169:ILE:HG22	1:A:171:CYS:SG	2.53	0.48
1:A:11:GLU:HB2	1:A:326:SER:HB3	1.96	0.48
1:A:233:ILE:HG13	1:A:237:GLN:HG2	1.94	0.48
1:A:46:ARG:HH22	1:A:81:VAL:HB	1.79	0.48
1:A:663:LYS:O	1:A:753:ARG:NH2	2.47	0.48
1:A:251:ILE:HG23	1:A:252:ILE:HG23	1.96	0.47
1:A:593:GLN:HG2	1:A:594:LEU:N	2.30	0.47
1:A:157:VAL:O	1:A:158:ARG:HD2	2.15	0.47
1:A:179:THR:HG23	1:A:280:LEU:O	2.15	0.47
1:A:567:ASN:H	1:A:591:ARG:HA	1.80	0.47
1:A:690:LEU:HD13	1:A:690:LEU:HA	1.80	0.46
1:A:412:VAL:O	1:A:416:GLN:HG3	2.15	0.46
1:A:334:PHE:CD2	1:A:336:THR:HB	2.50	0.46
1:A:156:SER:O	1:A:171:CYS:HA	2.16	0.46
1:A:6:THR:O	1:A:170:VAL:HA	2.16	0.46
1:A:488:LYS:HB3	1:A:488:LYS:HE3	1.59	0.45
1:A:675:GLU:HB3	1:A:686:ILE:HD13	1.98	0.45
1:A:395:LEU:O	1:A:430:GLU:HA	2.16	0.45
1:A:659:PHE:CE1	1:A:723:LYS:HB2	2.52	0.45
1:A:711:LEU:HD23	1:A:711:LEU:HA	1.80	0.45
1:A:77:VAL:HG13	1:A:88:LEU:O	2.17	0.44
1:A:564:VAL:HG21	1:A:594:LEU:HG	1.99	0.44
1:A:452:GLU:OE2	1:A:490:PHE:HZ	2.00	0.44
1:A:18:ALA:HB3	1:A:97:LEU:HB2	2.00	0.44
1:A:596:GLU:HB2	1:A:632:ALA:HB1	2.00	0.43
1:A:589:ARG:HB2	1:A:590:GLN:NE2	2.33	0.43
1:A:32:ASP:C	1:A:34:THR:H	2.20	0.43
1:A:81:VAL:HG12	1:A:82:VAL:H	1.83	0.43
1:A:703:ILE:HG21	1:A:738:PHE:HB2	2.01	0.43

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:118:LEU:HD23	1:A:118:LEU:HA	1.83	0.43
1:A:23:TYR:HA	1:A:90:PRO:HA	2.00	0.43
1:A:47:LEU:HD21	1:A:79:HIS:NE2	2.33	0.43
1:A:171:CYS:SG	1:A:311:ILE:HD11	2.58	0.43
1:A:194:LYS:HB3	1:A:194:LYS:HE2	1.85	0.43
1:A:121:PRO:HG2	1:A:318:LEU:HG	2.01	0.42
1:A:288:ARG:CZ	1:A:472:PRO:HG3	2.49	0.42
1:A:616:LYS:HG2	1:A:617:GLN:O	2.19	0.42
1:A:500:ASP:O	1:A:504:ARG:HG3	2.20	0.42
1:A:7:LEU:HD23	1:A:169:ILE:O	2.19	0.42
1:A:567:ASN:N	1:A:591:ARG:HA	2.35	0.42
2:T:8:DA:H2"	2:T:9:DC:C6	2.55	0.42
1:A:88:LEU:HA	1:A:88:LEU:HD23	1.90	0.42
1:A:15:ALA:H	1:A:98:ARG:NH2	2.17	0.42
1:A:42:LEU:HD23	1:A:82:VAL:HG13	2.02	0.42
1:A:564:VAL:O	1:A:565:LYS:HG2	2.19	0.42
1:A:754:LEU:HD22	1:A:758:ILE:HD11	2.02	0.42
1:A:642:ALA:O	1:A:646:GLU:HG3	2.20	0.41
1:A:76:PRO:O	1:A:78:ASP:N	2.48	0.41
1:A:486:PHE:HB3	1:A:492:LYS:HB2	2.02	0.41
1:A:10:LEU:HD12	1:A:10:LEU:HA	1.57	0.41
1:A:119:GLN:HB2	1:A:137:PHE:HZ	1.85	0.41
1:A:303:ARG:NH1	2:T:6:DG:OP1	2.54	0.41
1:A:15:ALA:N	1:A:98:ARG:HH21	2.18	0.41
1:A:341:PRO:HG3	1:A:364:MET:CE	2.51	0.41
1:A:509:ILE:HA	1:A:509:ILE:HD13	1.57	0.41
1:A:581:VAL:O	1:A:582:ARG:HD3	2.21	0.41
1:A:303:ARG:HD2	1:A:715:THR:HG21	2.03	0.40
1:A:561:GLY:HA2	1:A:624:HIS:O	2.21	0.40
1:A:296:THR:HG21	2:T:6:DG:H2"	2.04	0.40
1:A:680:GLU:O	1:A:682:PRO:HD3	2.22	0.40

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	767/789 (97%)	719 (94%)	45 (6%)	3 (0%)	34 46

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	81	VAL
1	A	581	VAL
1	A	64	GLN

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	643/657 (98%)	554 (86%)	89 (14%)	3 3

All (89) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	7	LEU
1	A	10	LEU
1	A	13	THR
1	A	21	GLN
1	A	25	VAL
1	A	26	LYS
1	A	29	SER
1	A	33	GLU
1	A	35	GLU
1	A	51	MET
1	A	59	SER
1	A	70	ASN
1	A	73	THR
1	A	83	ARG
1	A	85	MET
1	A	87	LYS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	89	VAL
1	A	112	ARG
1	A	133	SER
1	A	146	LEU
1	A	147	ASP
1	A	179	THR
1	A	189	ARG
1	A	192	LEU
1	A	195	LEU
1	A	226	LEU
1	A	230	ASN
1	A	233	ILE
1	A	270	ARG
1	A	284	THR
1	A	294	ARG
1	A	303	ARG
1	A	319	ASN
1	A	320	ILE
1	A	324	LYS
1	A	325	LEU
1	A	330	ARG
1	A	335	PHE
1	A	356	THR
1	A	360	ARG
1	A	364	MET
1	A	365	GLN
1	A	366	GLU
1	A	378	LYS
1	A	380	VAL
1	A	384	HIS
1	A	401	LYS
1	A	425	ARG
1	A	449	GLU
1	A	461	VAL
1	A	463	VAL
1	A	473	ARG
1	A	487	GLU
1	A	494	VAL
1	A	509	ILE
1	A	510	PHE
1	A	514	LYS
1	A	521	ARG

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	523	ARG
1	A	525	VAL
1	A	544	SER
1	A	550	LEU
1	A	559	THR
1	A	590	GLN
1	A	591	ARG
1	A	595	LEU
1	A	597	SER
1	A	604	THR
1	A	607	LEU
1	A	616	LYS
1	A	619	GLN
1	A	628	ARG
1	A	650	ILE
1	A	660	GLU
1	A	676	GLU
1	A	678	THR
1	A	679	GLN
1	A	680	GLU
1	A	683	LYS
1	A	690	LEU
1	A	703	ILE
1	A	712	LEU
1	A	713	GLN
1	A	715	THR
1	A	721	VAL
1	A	758	ILE
1	A	760	LEU
1	A	761	CYS
1	A	767	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

5.3.3 RNA

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates (i)

There are no monosaccharides 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^{th} 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	769/789 (97%)	0.18	43 (5%) 24 29	52, 88, 159, 204	0
2	T	18/18 (100%)	0.42	3 (16%) 1 1	61, 99, 254, 268	0
All	All	787/807 (97%)	0.19	46 (5%) 23 27	52, 88, 161, 268	0

All (46) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	84	ALA	5.1
1	A	88	LEU	4.3
1	A	81	VAL	4.2
1	A	58	LEU	4.1
1	A	80	ARG	3.9
1	A	335	PHE	3.8
1	A	15	ALA	3.7
1	A	473	ARG	3.7
1	A	79	HIS	3.7
2	T	13	DT	3.7
1	A	65	PHE	3.6
1	A	765	LEU	3.5
1	A	82	VAL	3.3
1	A	762	ASP	3.3
1	A	62	GLY	3.1
1	A	708	ALA	3.1
1	A	57	ALA	3.0
1	A	59	SER	3.0
1	A	83	ARG	3.0
1	A	34	THR	2.9
1	A	43	LEU	2.9
1	A	66	ILE	2.8
1	A	523	ARG	2.8
1	A	25	VAL	2.8

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	565	LYS	2.8
1	A	63	GLU	2.7
1	A	76	PRO	2.8
1	A	60	ARG	2.7
1	A	287	ARG	2.7
1	A	470	GLY	2.7
1	A	766	PHE	2.7
1	A	599	VAL	2.7
1	A	85	MET	2.6
1	A	20	TYR	2.6
1	A	95	LEU	2.5
1	A	192	LEU	2.5
1	A	28	LEU	2.5
1	A	27	GLY	2.4
1	A	77	VAL	2.3
1	A	78	ASP	2.3
1	A	106	TYR	2.2
2	T	11	DG	2.2
1	A	110	ARG	2.2
2	T	10	DA	2.1
1	A	47	LEU	2.1
1	A	61	ASP	2.1

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 monosaccharides 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.