



Full wwPDB X-ray Structure Validation Report ⓘ

May 25, 2020 – 03:56 am BST

PDB ID : 4HRF
Title : Atomic structure of DUSP26
Authors : Lokareddy, R.K.; Bhardwaj, A.; Cingolani, G.
Deposited on : 2012-10-27
Resolution : 1.68 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 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

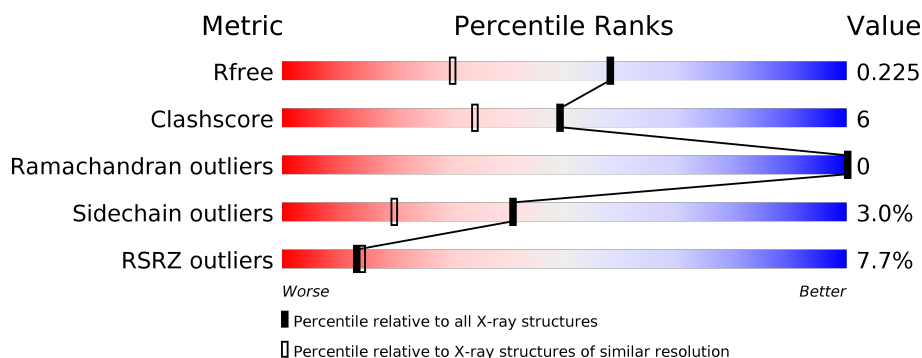
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.68 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	6780 (1.70-1.66)
Clashscore	141614	7310 (1.70-1.66)
Ramachandran outliers	138981	7173 (1.70-1.66)
Sidechain outliers	138945	7172 (1.70-1.66)
RSRZ outliers	127900	6661 (1.70-1.66)

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 ≥ 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	A	160	<div> <div>2%</div> <div> <div></div> <div>86%</div> <div>11%</div> <div></div> </div> <div></div> </div>
1	B	160	<div> <div>%</div> <div> <div></div> <div>89%</div> <div>7%</div> <div></div> </div> <div></div> </div>
1	C	160	<div> <div>10%</div> <div> <div></div> <div>84%</div> <div>7%</div> <div>5%</div> <div></div> </div> <div></div> </div>
1	D	160	<div> <div>16%</div> <div> <div></div> <div>84%</div> <div>9%</div> <div>6%</div> <div></div> </div> <div></div> </div>

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 5355 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 Dual specificity protein phosphatase 26.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	154	Total	C	N	O	S	0	0	0
			1226	775	234	213	4			
1	B	154	Total	C	N	O	S	0	0	0
			1226	775	234	213	4			
1	C	153	Total	C	N	O	S	0	0	0
			1217	770	233	210	4			
1	D	150	Total	C	N	O	S	0	0	0
			1195	756	230	205	4			

There are 40 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	60	MET	-	EXPRESSION TAG	UNP Q9BV47
A	152	SER	CYS	ENGINEERED MUTATION	UNP Q9BV47
A	212	LEU	-	EXPRESSION TAG	UNP Q9BV47
A	213	GLU	-	EXPRESSION TAG	UNP Q9BV47
A	214	HIS	-	EXPRESSION TAG	UNP Q9BV47
A	215	HIS	-	EXPRESSION TAG	UNP Q9BV47
A	216	HIS	-	EXPRESSION TAG	UNP Q9BV47
A	217	HIS	-	EXPRESSION TAG	UNP Q9BV47
A	218	HIS	-	EXPRESSION TAG	UNP Q9BV47
A	219	HIS	-	EXPRESSION TAG	UNP Q9BV47
B	60	MET	-	EXPRESSION TAG	UNP Q9BV47
B	152	SER	CYS	ENGINEERED MUTATION	UNP Q9BV47
B	212	LEU	-	EXPRESSION TAG	UNP Q9BV47
B	213	GLU	-	EXPRESSION TAG	UNP Q9BV47
B	214	HIS	-	EXPRESSION TAG	UNP Q9BV47
B	215	HIS	-	EXPRESSION TAG	UNP Q9BV47
B	216	HIS	-	EXPRESSION TAG	UNP Q9BV47
B	217	HIS	-	EXPRESSION TAG	UNP Q9BV47
B	218	HIS	-	EXPRESSION TAG	UNP Q9BV47
B	219	HIS	-	EXPRESSION TAG	UNP Q9BV47
C	60	MET	-	EXPRESSION TAG	UNP Q9BV47

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	152	SER	CYS	ENGINEERED MUTATION	UNP Q9BV47
C	212	LEU	-	EXPRESSION TAG	UNP Q9BV47
C	213	GLU	-	EXPRESSION TAG	UNP Q9BV47
C	214	HIS	-	EXPRESSION TAG	UNP Q9BV47
C	215	HIS	-	EXPRESSION TAG	UNP Q9BV47
C	216	HIS	-	EXPRESSION TAG	UNP Q9BV47
C	217	HIS	-	EXPRESSION TAG	UNP Q9BV47
C	218	HIS	-	EXPRESSION TAG	UNP Q9BV47
C	219	HIS	-	EXPRESSION TAG	UNP Q9BV47
D	60	MET	-	EXPRESSION TAG	UNP Q9BV47
D	152	SER	CYS	ENGINEERED MUTATION	UNP Q9BV47
D	212	LEU	-	EXPRESSION TAG	UNP Q9BV47
D	213	GLU	-	EXPRESSION TAG	UNP Q9BV47
D	214	HIS	-	EXPRESSION TAG	UNP Q9BV47
D	215	HIS	-	EXPRESSION TAG	UNP Q9BV47
D	216	HIS	-	EXPRESSION TAG	UNP Q9BV47
D	217	HIS	-	EXPRESSION TAG	UNP Q9BV47
D	218	HIS	-	EXPRESSION TAG	UNP Q9BV47
D	219	HIS	-	EXPRESSION TAG	UNP Q9BV47

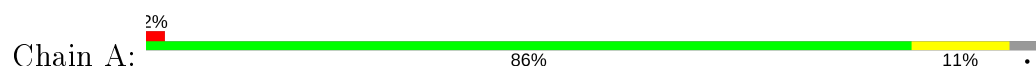
- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	140	Total O 140 140	0	0
2	B	167	Total O 167 167	0	0
2	C	89	Total O 89 89	0	0
2	D	95	Total O 95 95	0	0

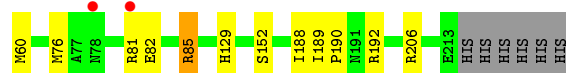
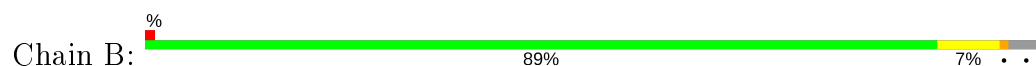
3 Residue-property plots [i](#)

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

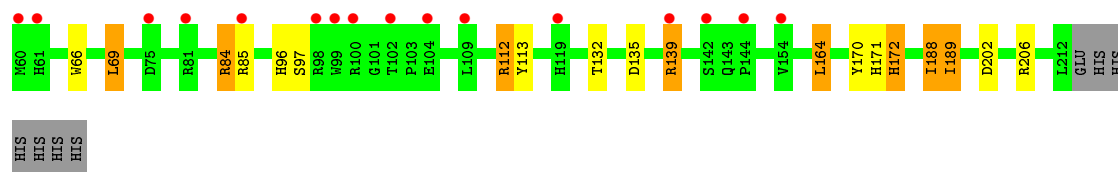
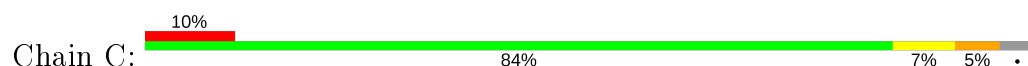
- Molecule 1: Dual specificity protein phosphatase 26



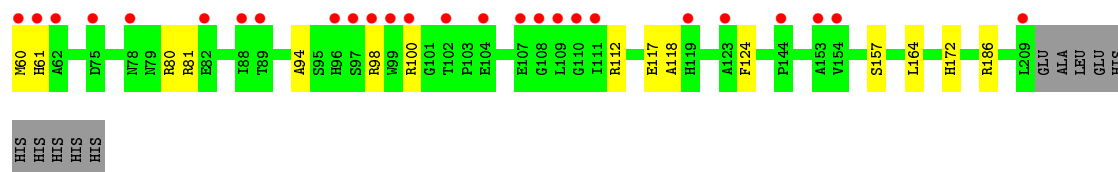
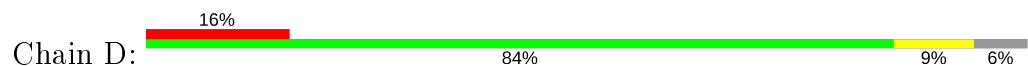
- Molecule 1: Dual specificity protein phosphatase 26



- Molecule 1: Dual specificity protein phosphatase 26



- Molecule 1: Dual specificity protein phosphatase 26



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	81.97Å 82.26Å 91.68Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	10.00 – 1.68 10.00 – 1.68	Depositor EDS
% Data completeness (in resolution range)	96.2 (10.00-1.68) 96.5 (10.00-1.68)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	4.18 (at 1.69Å)	Xtriage
Refinement program	REFMAC 5.7.0029, PHENIX (phenix.refine: 1.8.1_1161)	Depositor
R, R_{free}	0.185 , 0.216 0.197 , 0.225	Depositor DCC
R_{free} test set	3443 reflections (5.07%)	wwPDB-VP
Wilson B-factor (Å ²)	22.5	Xtriage
Anisotropy	0.331	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.46 , 60.8	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.000 for k,h,l	Xtriage
F_o, F_c correlation	0.97	EDS
Total number of atoms	5355	wwPDB-VP
Average B, all atoms (Å ²)	37.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 83.20 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 2.3013e-07. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

¹Intensities estimated from amplitudes.

²Theoretical 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	A	0.27	0/1253	0.46	0/1693
1	B	0.28	0/1253	0.46	0/1693
1	C	0.25	0/1244	0.45	0/1681
1	D	0.26	0/1222	0.44	0/1651
All	All	0.26	0/4972	0.45	0/6718

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

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	A	1226	0	1232	17	0
1	B	1226	0	1232	18	0
1	C	1217	0	1226	16	0
1	D	1195	0	1204	16	0
2	A	140	0	0	5	0
2	B	167	0	0	2	0
2	C	89	0	0	2	0
2	D	95	0	0	3	0
All	All	5355	0	4894	58	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 (58) 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:337:HOH:O	1:B:192:ARG:HD3	1.69	0.93
1:C:84:ARG:HB3	1:C:84:ARG:HH11	1.33	0.91
1:D:60:MET:HB3	1:D:61:HIS:HA	1.62	0.80
1:B:129:HIS:HD2	2:B:450:HOH:O	1.66	0.79
1:A:85:ARG:HE	1:B:85:ARG:HG3	1.47	0.78
1:C:84:ARG:CB	1:C:84:ARG:HH11	2.00	0.74
1:B:189:ILE:HD11	1:D:164:LEU:HD12	1.72	0.72
1:C:112:ARG:HD2	2:C:371:HOH:O	1.90	0.71
1:A:118:ALA:HB1	1:A:126:MET:HE1	1.74	0.70
1:D:60:MET:N	2:D:356:HOH:O	2.26	0.69
1:C:139:ARG:NH1	1:C:139:ARG:HG3	2.09	0.67
1:D:98:ARG:HD2	1:D:117:GLU:OE1	1.97	0.65
1:C:96:HIS:ND1	1:C:113:TYR:OH	2.27	0.65
1:D:172:HIS:ND1	2:D:333:HOH:O	2.31	0.62
1:B:85:ARG:HH21	1:B:85:ARG:CG	2.13	0.61
1:C:132:THR:HG23	2:C:358:HOH:O	1.99	0.61
1:C:139:ARG:HH11	1:C:139:ARG:HG3	1.66	0.60
1:A:85:ARG:NE	1:B:85:ARG:HG3	2.16	0.59
1:D:60:MET:CB	1:D:61:HIS:HA	2.30	0.59
1:C:139:ARG:HH11	1:C:139:ARG:CG	2.16	0.57
1:C:84:ARG:CG	1:C:84:ARG:HH11	2.19	0.56
1:D:100:ARG:HG3	1:D:100:ARG:O	2.06	0.55
1:B:189:ILE:HD11	1:D:164:LEU:CD1	2.37	0.53
1:A:80:ARG:NH1	2:A:416:HOH:O	2.42	0.51
1:A:82:GLU:OE2	1:A:82:GLU:HA	2.11	0.50
1:B:85:ARG:HG2	1:B:85:ARG:NH2	2.26	0.50
1:D:186:ARG:NH2	2:D:311:HOH:O	2.43	0.50
1:B:85:ARG:HH21	1:B:85:ARG:HG2	1.78	0.49
1:D:98:ARG:HD2	1:D:117:GLU:CD	2.32	0.49
1:B:190:PRO:HD2	1:D:157:SER:OG	2.13	0.48
1:B:85:ARG:CG	1:B:85:ARG:NH2	2.75	0.48
1:D:100:ARG:O	1:D:100:ARG:CG	2.61	0.48
1:A:154:VAL:HG22	1:A:155:GLY:N	2.28	0.47
1:D:118:ALA:HA	1:D:124:PHE:CZ	2.50	0.47
1:A:94:ALA:O	1:A:118:ALA:HB3	2.15	0.47
1:A:154:VAL:CG2	1:A:155:GLY:N	2.78	0.46
1:A:82:GLU:OE1	1:A:85:ARG:NH2	2.48	0.46
2:A:337:HOH:O	1:B:192:ARG:CD	2.44	0.46
1:C:66:TRP:HB3	1:C:69:LEU:HB3	1.98	0.45
1:B:188:ILE:O	1:B:189:ILE:HD13	2.16	0.45

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:87:GLY:HA2	1:B:192:ARG:NH1	2.32	0.44
1:D:81:ARG:HE	1:D:81:ARG:HB3	1.52	0.43
1:B:82:GLU:OE2	1:B:85:ARG:NE	2.51	0.43
1:B:60:MET:HE2	1:B:60:MET:HB2	1.89	0.43
1:B:206:ARG:HD3	2:B:464:HOH:O	2.19	0.43
1:A:156:VAL:HG13	1:C:188:ILE:O	2.19	0.43
1:A:143:GLN:NE2	2:A:363:HOH:O	2.40	0.42
1:C:135:ASP:OD1	1:C:170:TYR:OH	2.20	0.42
1:D:94:ALA:O	1:D:118:ALA:HB3	2.20	0.42
1:A:60:MET:HE1	1:A:79:ASN:HB2	2.01	0.41
1:A:198:LEU:HD13	1:C:164:LEU:HD23	2.02	0.41
1:A:85:ARG:HH22	1:A:86:LEU:HG	1.85	0.41
1:A:164:LEU:HD12	1:C:189:ILE:HD11	2.02	0.41
1:B:76:MET:SD	1:B:82:GLU:HG2	2.61	0.41
1:D:60:MET:N	1:D:61:HIS:HB2	2.36	0.41
1:A:119:HIS:HD2	2:A:428:HOH:O	2.03	0.40
1:C:171:HIS:C	1:C:172:HIS:CG	2.92	0.40
1:C:202:ASP:O	1:C:206:ARG:HG3	2.22	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	152/160 (95%)	150 (99%)	2 (1%)	0	100	100
1	B	152/160 (95%)	151 (99%)	1 (1%)	0	100	100
1	C	151/160 (94%)	150 (99%)	1 (1%)	0	100	100
1	D	148/160 (92%)	146 (99%)	2 (1%)	0	100	100
All	All	603/640 (94%)	597 (99%)	6 (1%)	0	100	100

There are no Ramachandran outliers to report.

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	126/132 (96%)	126 (100%)	0	100	100
1	B	126/132 (96%)	123 (98%)	3 (2%)	49	28
1	C	125/132 (95%)	115 (92%)	10 (8%)	12	2
1	D	123/132 (93%)	121 (98%)	2 (2%)	62	46
All	All	500/528 (95%)	485 (97%)	15 (3%)	41	20

All (15) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	81	ARG
1	B	85	ARG
1	B	152	SER
1	C	69	LEU
1	C	84	ARG
1	C	85	ARG
1	C	97	SER
1	C	112	ARG
1	C	139	ARG
1	C	164	LEU
1	C	172	HIS
1	C	188	ILE
1	C	189	ILE
1	D	80	ARG
1	D	112	ARG

Some 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 [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, 95th 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	154/160 (96%)	-0.02	3 (1%) 66 70	16, 28, 51, 82	0
1	B	154/160 (96%)	-0.15	2 (1%) 77 80	14, 24, 49, 60	0
1	C	153/160 (95%)	0.76	16 (10%) 6 6	19, 42, 67, 109	0
1	D	150/160 (93%)	0.80	26 (17%) 1 1	15, 38, 80, 107	0
All	All	611/640 (95%)	0.35	47 (7%) 13 14	14, 33, 68, 109	0

All (47) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	99	TRP	13.4
1	D	104	GLU	12.3
1	D	99	TRP	8.9
1	D	60	MET	5.3
1	D	107	GLU	5.2
1	D	108	GLY	5.1
1	C	109	LEU	5.0
1	D	100	ARG	4.8
1	C	154	VAL	4.6
1	C	142	SER	4.3
1	C	144	PRO	3.9
1	C	61	HIS	3.8
1	D	110	GLY	3.7
1	D	82	GLU	3.6
1	D	111	ILE	3.5
1	C	60	MET	3.5
1	B	81	ARG	3.5
1	D	98	ARG	3.5
1	C	75	ASP	3.3
1	D	78	ASN	3.3
1	D	123	ALA	3.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	D	154	VAL	3.1
1	C	85	ARG	3.0
1	C	102	THR	2.9
1	D	102	THR	2.9
1	D	97	SER	2.9
1	D	75	ASP	2.8
1	D	119	HIS	2.8
1	A	81	ARG	2.8
1	A	100	ARG	2.7
1	D	88	ILE	2.7
1	D	89	THR	2.6
1	C	98	ARG	2.5
1	D	61	HIS	2.5
1	D	109	LEU	2.5
1	D	153	ALA	2.5
1	A	99	TRP	2.5
1	C	81	ARG	2.4
1	D	62	ALA	2.4
1	D	209	LEU	2.3
1	C	119	HIS	2.3
1	C	139	ARG	2.3
1	C	104	GLU	2.2
1	B	78	ASN	2.2
1	D	96	HIS	2.1
1	C	100	ARG	2.0
1	D	144	PRO	2.0

6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

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

6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.

6.4 Ligands ⓘ

There are no ligands in this entry.

6.5 Other polymers

There are no such residues in this entry.