



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 12:42 PM GMT

PDB ID : 3RUJ  
Title : Crystal Structure of N-terminal region of yeast Atg7  
Authors : Hong, S.B.; Kim, B.W.; Song, H.K.  
Deposited on : 2011-05-05  
Resolution : 2.10 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

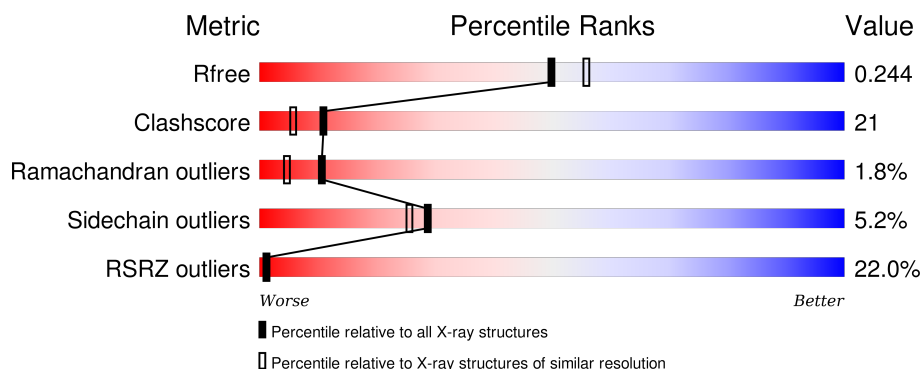
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.10 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	3939 (2.10-2.10)
Clashscore	102246	4460 (2.10-2.10)
Ramachandran outliers	100387	4413 (2.10-2.10)
Sidechain outliers	100360	4414 (2.10-2.10)
RSRZ outliers	91569	3948 (2.10-2.10)

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	296	

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 2490 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 Ubiquitin-like modifier-activating enzyme ATG7.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	288	Total	C	N	O	S	Se	0	0	0
			2335	1500	392	434	7	2			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	GLY	-	EXPRESSION TAG	UNP P38862
A	0	SER	-	EXPRESSION TAG	UNP P38862

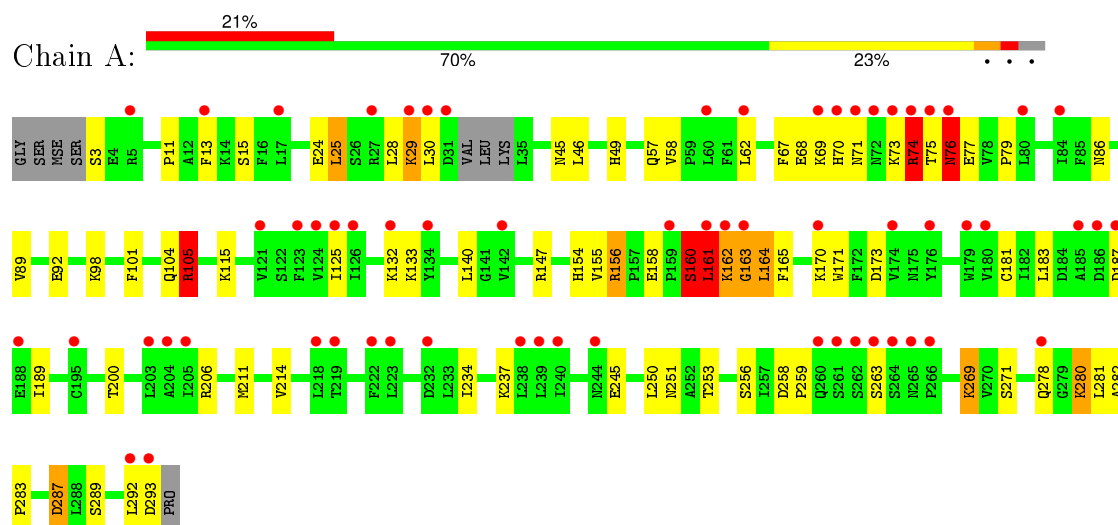
- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	155	Total	O	0	0
			155	155		

### 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: Ubiquitin-like modifier-activating enzyme ATG7



## 4 Data and refinement statistics

Property	Value	Source
Space group	I 41	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	112.85Å 112.85Å 102.25Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	29.26 – 2.10 29.26 – 2.10	Depositor EDS
% Data completeness (in resolution range)	85.0 (29.26-2.10) 84.3 (29.26-2.10)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.95 (at 2.10Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.7_650)	Depositor
R, $R_{free}$	0.228 , 0.254 0.221 , 0.244	Depositor DCC
$R_{free}$ test set	1701 reflections (5.36%)	DCC
Wilson B-factor (Å <sup>2</sup> )	40.1	Xtriage
Anisotropy	0.186	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.39 , 53.1	EDS
Estimated twinning fraction	0.033 for -k,-h,-l	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.33$	Xtriage
Outliers	0 of 37028 reflections	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	2490	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	50.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.38% 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

### 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	0.43	0/2386	0.66	4/3229 (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	4

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
1	A	105	ARG	NE-CZ-NH2	-8.72	115.94	120.30
1	A	163	GLY	N-CA-C	-5.67	98.91	113.10
1	A	25	LEU	CB-CG-CD1	-5.48	101.69	111.00
1	A	105	ARG	NE-CZ-NH1	5.46	123.03	120.30

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	29	LYS	Peptide
1	A	69	LYS	Peptide
1	A	74	ARG	Peptide
1	A	76	ASN	Peptide

## 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	2335	0	2330	98	0
2	A	155	0	0	9	1
All	All	2490	0	2330	98	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 21.

All (98) 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:104:GLN:O	1:A:105:ARG:HB2	1.46	1.08
1:A:25:LEU:HG	1:A:29:LYS:HD2	1.07	1.04
1:A:158:GLU:HG3	1:A:161:LEU:HD11	1.43	0.98
1:A:162:LYS:HE3	1:A:164:LEU:HD12	1.43	0.96
1:A:76:ASN:HD21	1:A:206:ARG:HH12	1.13	0.93
1:A:25:LEU:HG	1:A:29:LYS:CD	1.99	0.91
1:A:29:LYS:HE2	1:A:86:ASN:ND2	1.86	0.89
1:A:280:LYS:HE3	1:A:281:LEU:H	1.36	0.88
1:A:3:SER:HB2	1:A:154:HIS:HB3	1.57	0.84
1:A:25:LEU:O	1:A:29:LYS:HG3	1.79	0.82
1:A:29:LYS:HE2	1:A:86:ASN:CG	2.00	0.81
1:A:158:GLU:HG2	1:A:251:ASN:ND2	1.99	0.78
1:A:29:LYS:HZ1	1:A:86:ASN:HB2	1.46	0.78
1:A:289:SER:O	1:A:292:LEU:HG	1.85	0.77
1:A:29:LYS:NZ	1:A:86:ASN:HB2	2.00	0.76
1:A:68:GLU:OE1	2:A:377:HOH:O	2.04	0.75
1:A:292:LEU:HD12	1:A:293:ASP:N	2.02	0.75
1:A:161:LEU:HD22	1:A:237:LYS:HD3	1.68	0.75
1:A:245:GLU:OE2	2:A:409:HOH:O	2.04	0.74
1:A:25:LEU:HD21	1:A:29:LYS:HZ2	1.53	0.72
1:A:76:ASN:ND2	1:A:206:ARG:HH12	1.87	0.72
1:A:115:LYS:HE3	2:A:358:HOH:O	1.90	0.71
1:A:73:LYS:O	1:A:74:ARG:O	2.08	0.70
1:A:158:GLU:OE1	1:A:161:LEU:HG	1.93	0.69
1:A:45:ASN:ND2	1:A:70:HIS:NE2	2.39	0.69

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:292:LEU:HD12	1:A:293:ASP:H	1.57	0.69
1:A:25:LEU:HD21	1:A:29:LYS:NZ	2.10	0.67
1:A:104:GLN:O	1:A:105:ARG:CB	2.30	0.66
1:A:29:LYS:HE2	1:A:86:ASN:CB	2.26	0.65
1:A:25:LEU:HD11	1:A:29:LYS:HZ2	1.61	0.65
1:A:11:PRO:O	1:A:57:GLN:HG2	1.97	0.65
1:A:30:LEU:HG	2:A:436:HOH:O	1.97	0.64
1:A:25:LEU:CG	1:A:29:LYS:HZ2	2.10	0.64
1:A:29:LYS:NZ	1:A:86:ASN:CB	2.61	0.63
1:A:29:LYS:CE	1:A:86:ASN:CG	2.67	0.63
1:A:25:LEU:CD2	1:A:29:LYS:HZ2	2.12	0.63
1:A:3:SER:HA	1:A:156:ARG:HG2	1.80	0.62
1:A:160:SER:O	1:A:162:LYS:HG2	2.00	0.62
1:A:162:LYS:HD2	1:A:165:PHE:H	1.65	0.62
1:A:29:LYS:CE	1:A:86:ASN:CB	2.77	0.61
1:A:62:LEU:HD13	1:A:125:ILE:HD11	1.85	0.59
1:A:162:LYS:CE	1:A:164:LEU:HD12	2.25	0.58
1:A:280:LYS:HE3	1:A:281:LEU:N	2.15	0.58
1:A:29:LYS:HZ3	1:A:86:ASN:CG	2.08	0.58
1:A:45:ASN:OD1	1:A:70:HIS:NE2	2.36	0.58
1:A:162:LYS:HB2	1:A:164:LEU:N	2.19	0.57
1:A:45:ASN:OD1	1:A:70:HIS:CE1	2.58	0.57
1:A:29:LYS:NZ	1:A:86:ASN:CG	2.60	0.55
1:A:25:LEU:CD1	1:A:29:LYS:HZ2	2.19	0.54
1:A:160:SER:O	1:A:161:LEU:C	2.46	0.53
1:A:101:PHE:O	1:A:104:GLN:O	2.26	0.53
1:A:161:LEU:CD2	1:A:237:LYS:HD3	2.38	0.53
1:A:160:SER:OG	1:A:162:LYS:HE2	2.09	0.53
1:A:70:HIS:HE1	1:A:79:PRO:HB3	1.74	0.52
1:A:161:LEU:HD22	1:A:237:LYS:CD	2.38	0.52
1:A:162:LYS:HB2	1:A:164:LEU:H	1.75	0.51
1:A:161:LEU:HD13	1:A:237:LYS:HD3	1.92	0.51
1:A:183:LEU:HD23	1:A:189:ILE:HD13	1.92	0.51
1:A:183:LEU:CD2	1:A:189:ILE:HD13	2.42	0.50
1:A:258:ASP:N	1:A:259:PRO:HD3	2.26	0.50
1:A:62:LEU:CD1	1:A:125:ILE:HD11	2.41	0.50
1:A:155:VAL:HG11	1:A:250:LEU:HD13	1.94	0.50
1:A:46:LEU:O	1:A:77:GLU:HA	2.12	0.49
1:A:76:ASN:O	2:A:426:HOH:O	2.20	0.49
1:A:278:GLN:CD	1:A:278:GLN:N	2.65	0.49
1:A:163:GLY:HA3	1:A:189:ILE:HD11	1.94	0.49

*Continued on next page...*



*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:24:GLU:HG3	1:A:28:LEU:HD13	1.95	0.49
1:A:133:LYS:HB3	2:A:378:HOH:O	2.13	0.48
1:A:25:LEU:HD11	1:A:29:LYS:NZ	2.29	0.48
1:A:25:LEU:HD13	1:A:67:PHE:HZ	1.78	0.47
1:A:269:LYS:HE2	2:A:439:HOH:O	2.14	0.46
1:A:160:SER:C	1:A:162:LYS:HG2	2.36	0.46
1:A:45:ASN:HD21	1:A:70:HIS:CD2	2.31	0.46
1:A:76:ASN:HB2	1:A:173:ASP:OD1	2.16	0.46
1:A:161:LEU:HD22	1:A:237:LYS:CE	2.46	0.45
1:A:45:ASN:CG	1:A:70:HIS:NE2	2.70	0.45
1:A:89:VAL:CG1	1:A:92:GLU:HG3	2.45	0.45
1:A:49:HIS:CD2	1:A:76:ASN:HB3	2.51	0.45
1:A:13:PHE:HD1	1:A:58:VAL:HG13	1.83	0.45
1:A:105:ARG:HA	1:A:105:ARG:HD3	1.65	0.44
1:A:132:LYS:HD2	1:A:132:LYS:HA	1.86	0.44
1:A:147:ARG:NH1	2:A:437:HOH:O	2.50	0.44
1:A:15:SER:HB3	1:A:140:LEU:HD22	1.98	0.44
1:A:160:SER:O	1:A:162:LYS:N	2.51	0.43
1:A:70:HIS:CE1	1:A:79:PRO:HB3	2.52	0.43
1:A:75:THR:O	1:A:76:ASN:C	2.57	0.43
1:A:171:TRP:CZ3	1:A:181:CYS:HB3	2.53	0.43
1:A:271:SER:OG	2:A:351:HOH:O	2.20	0.43
1:A:49:HIS:HD2	1:A:76:ASN:HB3	1.84	0.43
1:A:256:SER:OG	1:A:259:PRO:HG3	2.18	0.42
1:A:234:ILE:O	1:A:253:THR:HA	2.20	0.42
1:A:161:LEU:CD1	1:A:161:LEU:O	2.69	0.41
1:A:71:ASN:OD1	1:A:71:ASN:O	2.38	0.41
1:A:287:ASP:OD2	1:A:287:ASP:C	2.59	0.41
1:A:282:ALA:HB1	1:A:283:PRO:HD2	2.03	0.41
1:A:25:LEU:HA	1:A:25:LEU:HD12	1.80	0.41
1:A:3:SER:CB	1:A:154:HIS:HB3	2.40	0.41
1:A:211:MSE:HE2	1:A:214:VAL:HG11	2.02	0.40

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)
2:A:422:HOH:O	2:A:448:HOH:O 4_464	2.01	0.19

## 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	284/296 (96%)	269 (95%)	10 (4%)	5 (2%)	11 5

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	74	ARG
1	A	105	ARG
1	A	161	LEU
1	A	164	LEU
1	A	160	SER

### 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	269/273 (98%)	255 (95%)	14 (5%)	29 25

All (14) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	76	ASN
1	A	98	LYS
1	A	105	ARG
1	A	156	ARG
1	A	160	SER
1	A	161	LEU

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	162	LYS
1	A	170	LYS
1	A	187	ASP
1	A	200	THR
1	A	263	SER
1	A	269	LYS
1	A	280	LYS
1	A	287	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	76	ASN
1	A	251	ASN

### 5.3.3 RNA [i](#)

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, 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	286/296 (96%)	1.13	63 (22%) ⓘ ⓘ	27, 48, 77, 102	0

All (63) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	263	SER	8.5
1	A	262	SER	8.2
1	A	264	SER	7.4
1	A	161	LEU	7.0
1	A	74	ARG	5.7
1	A	75	THR	5.5
1	A	205	ILE	5.5
1	A	203	LEU	5.4
1	A	185	ALA	5.4
1	A	265	ASN	5.4
1	A	69	LYS	5.1
1	A	261	SER	5.1
1	A	76	ASN	5.1
1	A	29	LYS	5.0
1	A	222	PHE	4.3
1	A	134	TYR	4.3
1	A	238	LEU	4.3
1	A	223	LEU	4.2
1	A	132	LYS	4.2
1	A	292	LEU	4.2
1	A	73	LYS	4.1
1	A	163	GLY	4.1
1	A	240	ILE	4.1
1	A	293	ASP	4.0
1	A	72	ASN	3.9
1	A	186	ASP	3.9
1	A	188	GLU	3.8

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	239	LEU	3.8
1	A	218	LEU	3.7
1	A	71	ASN	3.7
1	A	195	CYS	3.4
1	A	27	ARG	3.4
1	A	187	ASP	3.3
1	A	70	HIS	3.3
1	A	162	LYS	3.1
1	A	266	PRO	3.1
1	A	219	THR	3.0
1	A	126	ILE	2.9
1	A	17	LEU	2.8
1	A	62	LEU	2.7
1	A	204	ALA	2.7
1	A	232	ASP	2.7
1	A	125	ILE	2.7
1	A	60	LEU	2.6
1	A	176	TYR	2.5
1	A	179	TRP	2.5
1	A	278	GLN	2.5
1	A	80	LEU	2.5
1	A	13	PHE	2.4
1	A	260	GLN	2.4
1	A	5	ARG	2.4
1	A	159	PRO	2.4
1	A	244	ASN	2.3
1	A	124	VAL	2.3
1	A	174	VAL	2.3
1	A	180	VAL	2.3
1	A	142	VAL	2.2
1	A	170	LYS	2.2
1	A	84	ILE	2.1
1	A	30	LEU	2.1
1	A	31	ASP	2.0
1	A	123	PHE	2.0
1	A	121	VAL	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 [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.