Full wwPDB X-ray Structure Validation Report ⓘ

Oct 31, 2023 – 07:53 PM JST

PDB ID	:	5H2V
Title	:	Crystal structure of the karyopherin Kap121p bound to the SUMO protease Ulp1p
Authors	:	Kobayashi, J.; Matsuura, Y.
Deposited on	:	2016-10-18
Resolution	:	2.80 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:

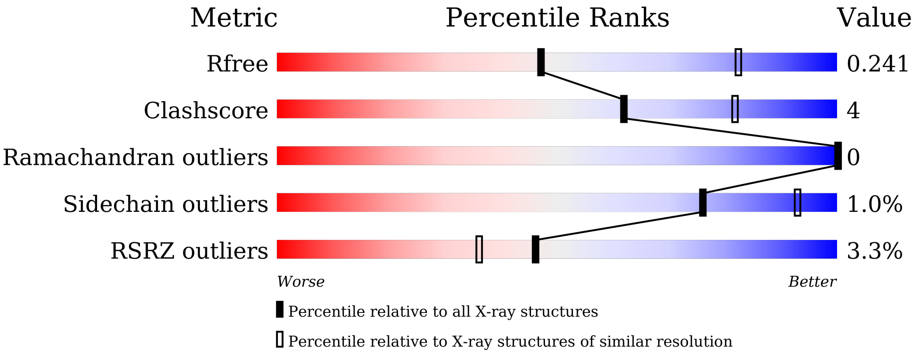
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

1 Overall quality at a glance

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

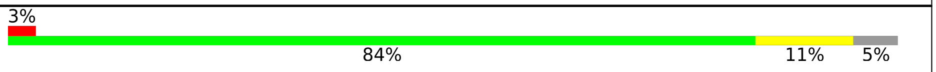
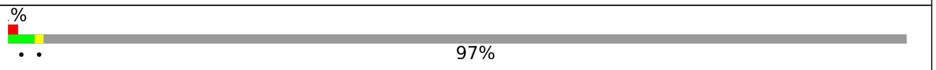
The reported resolution of this entry is 2.80 Å.

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	3140 (2.80-2.80)
Clashscore	141614	3569 (2.80-2.80)
Ramachandran outliers	138981	3498 (2.80-2.80)
Sidechain outliers	138945	3500 (2.80-2.80)
RSRZ outliers	127900	3078 (2.80-2.80)

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 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	1078	
2	B	150	

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 7988 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 Importin subunit beta-3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	1027	7956	5102	1288	1529	37	0	0	0

There are 11 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	?	-	PRO	deletion	UNP P32337
A	?	-	SER	deletion	UNP P32337
A	?	-	SER	deletion	UNP P32337
A	?	-	LYS	deletion	UNP P32337
A	?	-	LEU	deletion	UNP P32337
A	?	-	MET	deletion	UNP P32337
A	?	-	ILE	deletion	UNP P32337
A	?	-	MET	deletion	UNP P32337
A	?	-	SER	deletion	UNP P32337
A	?	-	LYS	deletion	UNP P32337
A	?	-	ASN	deletion	UNP P32337

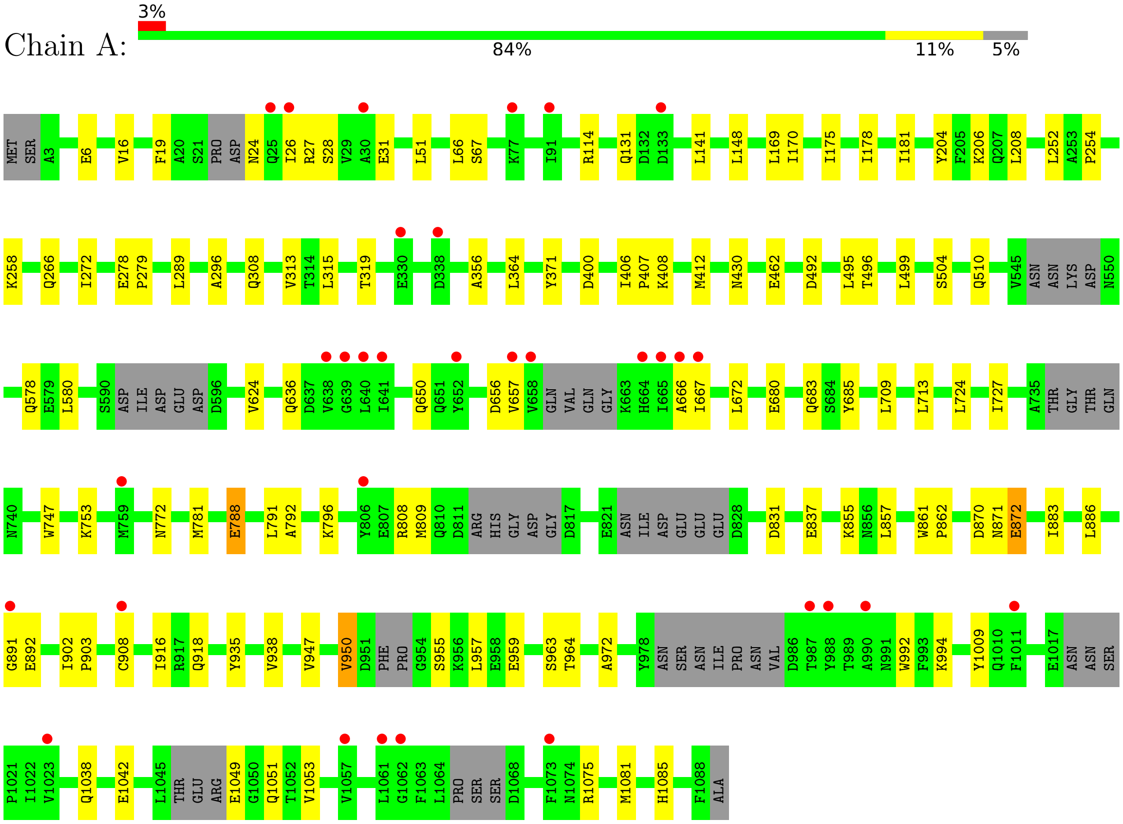
- Molecule 2 is a protein called Ubiquitin-like-specific protease 1.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
2	B	5	32	21	6	5	0	0	0

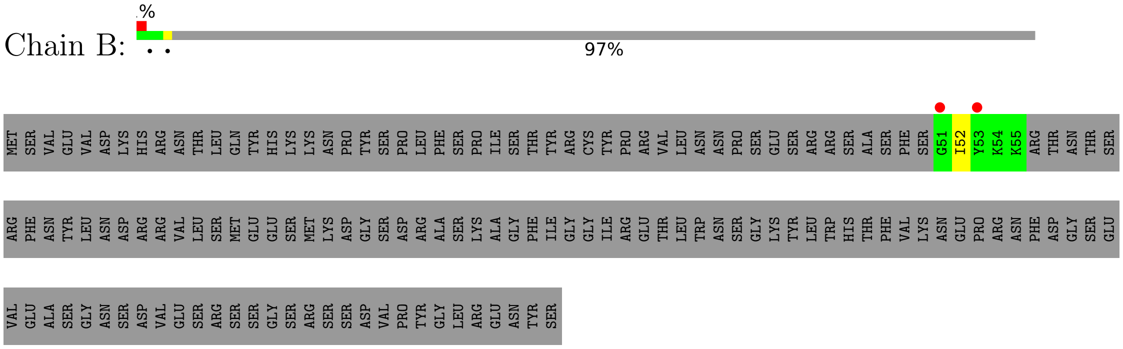
3 Residue-property plots (i)

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: Importin subunit beta-3



- Molecule 2: Ubiquitin-like-specific protease 1



4 Data and refinement statistics ⓘ

Property	**Value**	**Source**
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	76.83Å 124.71Å 130.00Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	36.71 – 2.80 36.71 – 2.80	Depositor EDS
% Data completeness (in resolution range)	100.0 (36.71-2.80) 100.0 (36.71-2.80)	Depositor EDS
R_{merge}	0.17	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.71 (at 2.81Å)	Xtriage
Refinement program	PHENIX 1.9_1692	Depositor
R, R_{free}	0.198 , 0.241 0.200 , 0.241	Depositor DCC
R_{free} test set	1595 reflections (5.07%)	wwPDB-VP
Wilson B-factor (Å^2)	46.1	Xtriage
Anisotropy	0.642	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 43.7	EDS
L-test for twinning[2]	$< \|L\| > = 0.46$, $< L^2 > = 0.29$	Xtriage
Estimated twinning fraction	0.020 for -h,l,k	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	7988	wwPDB-VP
Average B, all atoms (Å^2)	59.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.58% 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	0.28	0/8092	0.45	0/11001
2	B	0.77	0/31	0.66	0/39
All	All	0.29	0/8123	0.45	0/11040

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	7956	0	7955	70	0
2	B	32	0	32	1	0
All	All	7988	0	7987	70	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 4.

All (70) 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:683:GLN:HG2	1:A:724:LEU:HD13	1.59	0.85
1:A:504:SER:O	1:A:510:GLN:NE2	2.10	0.84

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:28:SER:HA	1:A:31:GLU:HG2	1.68	0.76
1:A:788:GLU:HG3	1:A:791:LEU:HD12	1.71	0.73
1:A:1049:GLU:N	1:A:1049:GLU:OE1	2.27	0.68
1:A:657:VAL:HG12	1:A:666:ALA:HA	1.76	0.66
1:A:857:LEU:HD22	1:A:886:LEU:HD21	1.77	0.65
1:A:957:LEU:HD23	1:A:959:GLU:H	1.62	0.64
1:A:1009:TYR:HB3	1:A:1053:VAL:HG21	1.82	0.62
1:A:636:GLN:HE22	1:A:713:LEU:H	1.46	0.61
1:A:16:VAL:HG23	1:A:66:LEU:HD12	1.81	0.60
1:A:170:ILE:HG23	1:A:178:ILE:HG12	1.84	0.59
1:A:406:ILE:HG23	1:A:407:PRO:HD3	1.86	0.57
1:A:955:SER:HB2	1:A:964:THR:HG21	1.87	0.56
1:A:683:GLN:OE1	1:A:727:ILE:HD12	2.07	0.55
1:A:918:GLN:HB2	1:A:963:SER:HA	1.88	0.54
1:A:278:GLU:HG3	1:A:279:PRO:HD2	1.89	0.54
1:A:148:LEU:HD13	1:A:181:ILE:HG21	1.89	0.54
1:A:809:MET:HG2	1:A:872:GLU:OE1	2.09	0.52
1:A:406:ILE:CG2	1:A:407:PRO:HD3	2.38	0.52
1:A:430:ASN:HA	2:B:52:ILE:HD13	1.91	0.52
1:A:356:ALA:HA	1:A:364:LEU:HD12	1.91	0.51
1:A:492:ASP:O	1:A:496:THR:HG23	2.11	0.50
1:A:6:GLU:H	1:A:6:GLU:CD	2.14	0.50
1:A:206:LYS:HG3	1:A:252:LEU:HD11	1.95	0.49
1:A:902:ILE:HD13	1:A:938:VAL:HG11	1.94	0.48
1:A:902:ILE:HD11	1:A:935:TYR:CD1	2.49	0.48
1:A:680:GLU:O	1:A:683:GLN:HB2	2.13	0.48
1:A:772:ASN:HB2	1:A:837:GLU:HG2	1.96	0.48
1:A:650:GLN:NE2	1:A:657:VAL:H	2.11	0.47
1:A:972:ALA:HB2	1:A:992:TRP:NE1	2.30	0.47
1:A:994:LYS:HA	1:A:994:LYS:HD2	1.63	0.47
1:A:636:GLN:NE2	1:A:713:LEU:H	2.12	0.46
1:A:51:LEU:HD22	1:A:67:SER:HA	1.98	0.46
1:A:747:TRP:CG	1:A:781:MET:HG3	2.51	0.46
1:A:855:LYS:HD2	1:A:855:LYS:C	2.36	0.46
1:A:308:GLN:NE2	1:A:371:TYR:HE2	2.14	0.45
1:A:495:LEU:O	1:A:499:LEU:HG	2.16	0.45
1:A:462:GLU:HA	1:A:462:GLU:OE1	2.17	0.45
1:A:808:ARG:NH2	1:A:831:ASP:OD1	2.45	0.45
1:A:709:LEU:HD12	1:A:753:LYS:HG2	1.99	0.45
1:A:272:ILE:HG21	1:A:313:VAL:HB	1.98	0.45
1:A:624:VAL:HG11	1:A:685:TYR:CZ	2.52	0.45

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:947:VAL:O	1:A:950:VAL:HG12	2.16	0.44
1:A:883:ILE:HA	1:A:886:LEU:HD12	1.99	0.44
1:A:175:ILE:HA	1:A:178:ILE:HG22	1.99	0.43
1:A:891:GLY:HA2	1:A:892:GLU:HA	1.66	0.43
1:A:902:ILE:N	1:A:903:PRO:HD2	2.32	0.43
1:A:808:ARG:HH12	1:A:831:ASP:CG	2.21	0.43
1:A:861:TRP:N	1:A:862:PRO:HD2	2.34	0.43
1:A:315:LEU:O	1:A:319:THR:HG23	2.19	0.42
1:A:141:LEU:HD21	1:A:169:LEU:HD22	1.99	0.42
1:A:258:LYS:HB3	1:A:258:LYS:HE3	1.66	0.42
1:A:636:GLN:HA	1:A:672:LEU:HD21	2.01	0.42
1:A:870:ASP:OD2	1:A:871:ASN:N	2.52	0.42
1:A:1038:GLN:O	1:A:1042:GLU:HB2	2.20	0.42
1:A:400:ASP:OD2	1:A:400:ASP:N	2.53	0.42
1:A:1081:MET:O	1:A:1085:HIS:HD2	2.04	0.41
1:A:254:PRO:HB2	1:A:296:ALA:HB2	2.02	0.41
1:A:792:ALA:O	1:A:796:LYS:HG3	2.21	0.41
1:A:580:LEU:HD12	1:A:580:LEU:HA	1.96	0.41
1:A:667:ILE:O	1:A:667:ILE:HG13	2.20	0.41
1:A:908:CYS:HB3	1:A:916:ILE:HG22	2.03	0.41
1:A:24:ASN:HA	1:A:27:ARG:HB3	2.02	0.41
1:A:114:ARG:HA	1:A:114:ARG:HD2	1.93	0.41
1:A:204:TYR:CE1	1:A:208:LEU:HD22	2.56	0.40
1:A:408:LYS:O	1:A:412:MET:HG3	2.20	0.40
1:A:289:LEU:HD23	1:A:289:LEU:HA	1.94	0.40
1:A:1051:GLN:CD	1:A:1051:GLN:H	2.25	0.40
1:A:26:ILE:HD13	1:A:26:ILE:HA	1.90	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	1001/1078 (93%)	978 (98%)	23 (2%)	0	100	100
2	B	3/150 (2%)	3 (100%)	0	0	100	100
All	All	1004/1228 (82%)	981 (98%)	23 (2%)	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	871/937 (93%)	862 (99%)	9 (1%)	76	93
2	B	2/136 (2%)	2 (100%)	0	100	100
All	All	873/1073 (81%)	864 (99%)	9 (1%)	76	93

All (9) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	19	PHE
1	A	131	GLN
1	A	266	GLN
1	A	578	GLN
1	A	656	ASP
1	A	788	GLU
1	A	872	GLU
1	A	950	VAL
1	A	1075	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	308	GLN
1	A	636	GLN
1	A	650	GLN
1	A	651	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	960	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 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, 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	1027/1078 (95%)	0.01	32 (3%)	49	39	20, 55, 103, 124	0
2	B	5/150 (3%)	1.21	2 (40%)	0	0	45, 57, 63, 78	0
All	All	1032/1228 (84%)	0.02	34 (3%)	46	36	20, 55, 103, 124	0

All (34) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	638	VAL	5.6
1	A	658	VAL	5.2
1	A	665	ILE	5.1
1	A	652	TYR	4.1
1	A	639	GLY	4.1
1	A	987	THR	4.0
1	A	1061	LEU	3.5
1	A	338	ASP	3.2
1	A	1057	VAL	3.1
1	A	640	LEU	3.1
1	A	891	GLY	2.9
1	A	641	ILE	2.8
1	A	26	ILE	2.7
1	A	657	VAL	2.7
1	A	664	HIS	2.6
1	A	330	GLU	2.6
1	A	667	ILE	2.5
1	A	25	GLN	2.5
1	A	1062	GLY	2.5
2	B	51	GLY	2.5
1	A	988	TYR	2.4
1	A	666	ALA	2.4
1	A	1023	VAL	2.3
1	A	30	ALA	2.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	77	LYS	2.3
1	A	759	MET	2.3
1	A	806	TYR	2.2
1	A	990	ALA	2.2
1	A	908	CYS	2.1
1	A	1073	PHE	2.1
1	A	133	ASP	2.0
2	B	53	TYR	2.0
1	A	1011	PHE	2.0
1	A	91	ILE	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 monosaccharides 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.