



wwPDB X-ray Structure Validation Summary Report ⓘ

May 21, 2020 – 06:26 am BST

PDB ID : 6S9I
Title : Human Brr2 Helicase Region D534C/N1866C in complex with C-tail deleted Jab1
Authors : Vester, K.; Santos, K.F.; Wahl, M.C.
Deposited on : 2019-07-13
Resolution : 2.60 Å(reported)

This is a wwPDB X-ray Structure Validation Summary 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
Xtrriage (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

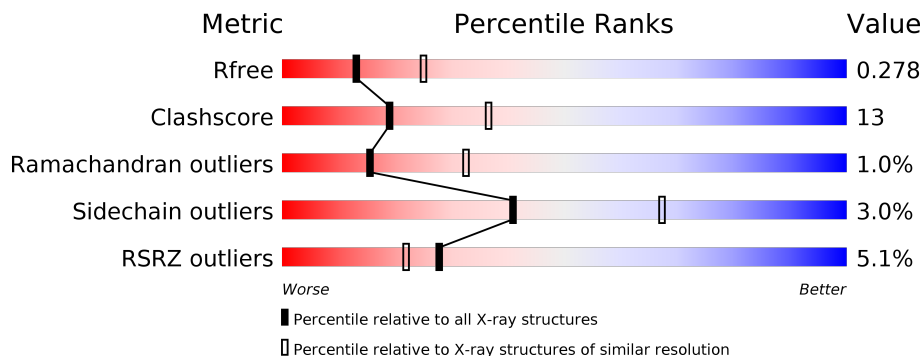
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.60 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	3163 (2.60-2.60)
Clashscore	141614	3518 (2.60-2.60)
Ramachandran outliers	138981	3455 (2.60-2.60)
Sidechain outliers	138945	3455 (2.60-2.60)
RSRZ outliers	127900	3104 (2.60-2.60)

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	B	1747	
2	J	263	

2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 15928 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 U5 small nuclear ribonucleoprotein 200 kDa helicase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	B	1712	13780	8808	2359	2539	74	0	1	0

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	390	GLY	-	expression tag	UNP O75643
B	391	ALA	-	expression tag	UNP O75643
B	392	GLU	-	expression tag	UNP O75643
B	393	PHE	-	expression tag	UNP O75643
B	534	CYS	ASP	engineered mutation	UNP O75643
B	1866	CYS	ASN	engineered mutation	UNP O75643

- Molecule 2 is a protein called Pre-mRNA-processing-splicing factor 8.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	J	252	2046	1312	353	369	12	0	0	0

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
J	2058	GLY	-	expression tag	UNP Q6P2Q9
J	2059	PRO	-	expression tag	UNP Q6P2Q9
J	2060	LEU	-	expression tag	UNP Q6P2Q9
J	2061	GLY	-	expression tag	UNP Q6P2Q9
J	2062	SER	-	expression tag	UNP Q6P2Q9
J	2063	MET	-	expression tag	UNP Q6P2Q9

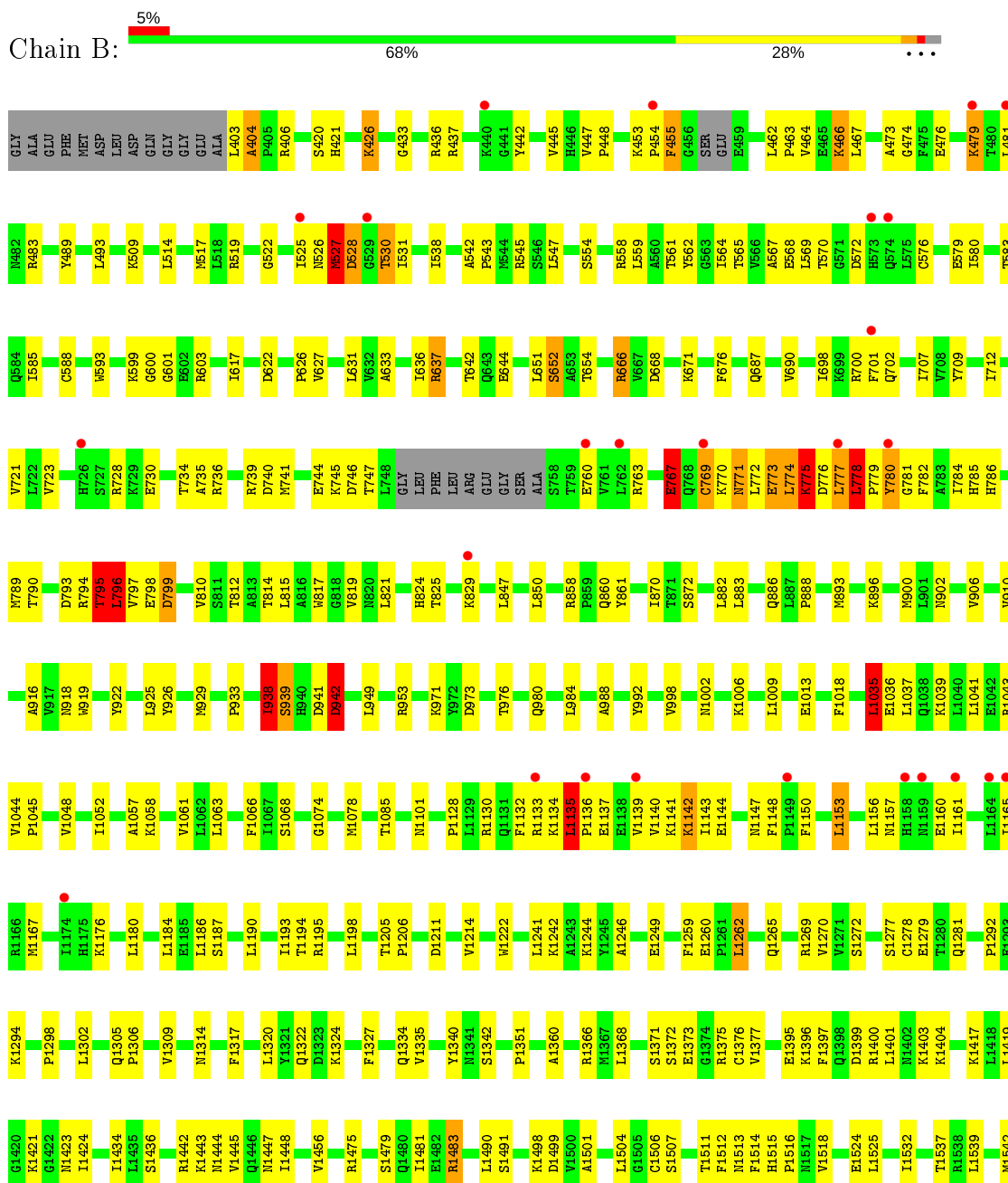
- Molecule 3 is water.

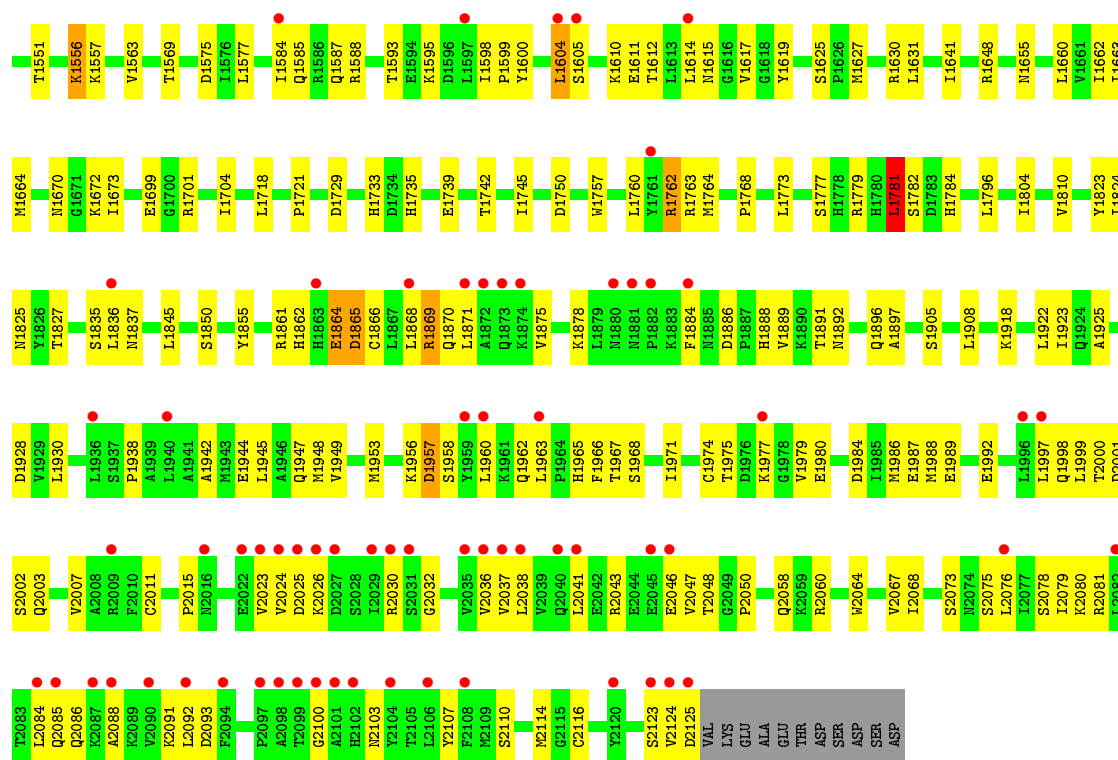
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	86	Total 86	O 86	0	0
3	J	16	Total 16	O 16	0	0

3 Residue-property plots

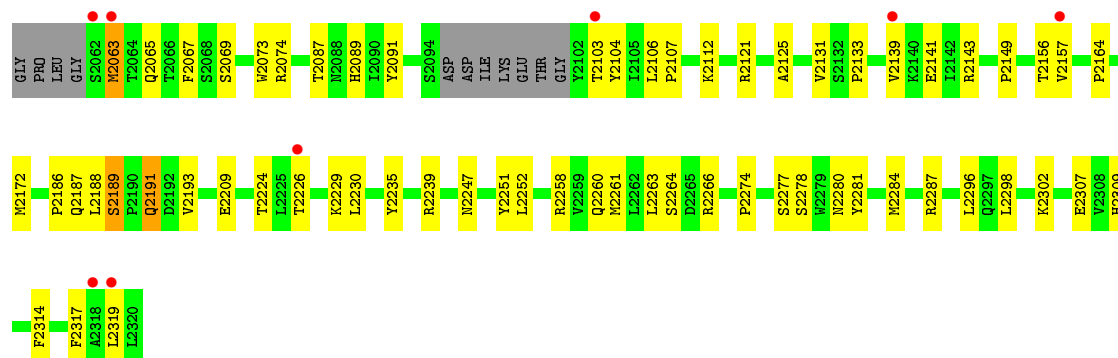
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are 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: U5 small nuclear ribonucleoprotein 200 kDa helicase





• Molecule 2: Pre-mRNA-processing-splicing factor 8



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	99.24Å 119.08Å 187.91Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	47.98 – 2.60 49.27 – 2.60	Depositor EDS
% Data completeness (in resolution range)	98.9 (47.98-2.60) 98.9 (49.27-2.60)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.35 (at 2.61Å)	Xtrriage
Refinement program	PHENIX 1.15.2_3472	Depositor
R, R_{free}	0.197 , 0.278 0.197 , 0.278	Depositor DCC
R_{free} test set	2100 reflections (3.06%)	wwPDB-VP
Wilson B-factor (Å ²)	54.0	Xtrriage
Anisotropy	0.130	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.29 , 50.4	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	15928	wwPDB-VP
Average B, all atoms (Å ²)	69.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.27% of the height of the origin peak. No significant pseudotranslation is detected.*

¹ Intensities estimated from amplitudes.

² Theoretical 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 [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	B	0.34	2/14071 (0.0%)	0.59	17/19063 (0.1%)
2	J	0.30	0/2111	0.48	0/2874
All	All	0.34	2/16182 (0.0%)	0.57	17/21937 (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	B	0	9

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	476	GLU	CD-OE2	8.52	1.35	1.25
1	B	426	LYS	CE-NZ	5.30	1.62	1.49

The worst 5 of 17 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed($^{\circ}$)	Ideal($^{\circ}$)
1	B	796	LEU	CA-CB-CG	18.17	157.09	115.30
1	B	778	LEU	CA-CB-CG	15.48	150.91	115.30
1	B	426	LYS	CD-CE-NZ	-13.53	80.58	111.70
1	B	796	LEU	CB-CG-CD2	9.60	127.32	111.00
1	B	799	ASP	CB-CG-OD2	8.83	126.25	118.30

There are no chirality outliers.

5 of 9 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	479	LYS	Peptide

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Group
1	B	527	MET	Peptide
1	B	603	ARG	Peptide
1	B	767	GLU	Peptide
1	B	773	GLU	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	B	13780	0	13925	383	2
2	J	2046	0	1990	42	1
3	B	86	0	0	5	0
3	J	16	0	0	1	0
All	All	15928	0	15915	413	2

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 13.

The worst 5 of 413 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:B:1130:ARG:NH1	1:B:1144:GLU:OE2	1.57	1.32
1:B:796:LEU:HB2	1:B:799:ASP:HB2	1.33	1.06
1:B:1130:ARG:NH1	1:B:1144:GLU:CD	2.09	1.05
1:B:1928:ASP:OD1	1:B:2081:ARG:NH1	1.94	0.98
1:B:896:LYS:HB2	1:B:900:MET:HE2	1.49	0.93

All (2) 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)
1:B:1782:SER:OG	2:J:2065:GLN:OE1[4_545]	2.01	0.19
1:B:860:GLN:NE2	1:B:941:ASP:OD1[4_555]	2.14	0.06

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	B	1707/1747 (98%)	1584 (93%)	103 (6%)	20 (1%)	13	27
2	J	248/263 (94%)	234 (94%)	14 (6%)	0	100	100
All	All	1955/2010 (97%)	1818 (93%)	117 (6%)	20 (1%)	15	32

5 of 20 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	528	ASP
1	B	780	TYR
1	B	795	THR
1	B	938	ILE
1	B	1133	ARG

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	B	1535/1560 (98%)	1491 (97%)	44 (3%)	42	68
2	J	228/236 (97%)	220 (96%)	8 (4%)	36	62
All	All	1763/1796 (98%)	1711 (97%)	52 (3%)	41	68

5 of 52 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	1187	SER

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	1605	SER
2	J	2247	ASN
1	B	1342	SER
1	B	1498	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	573	HIS
1	B	2085	GLN
1	B	2103	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 [i](#)

6.1 Protein, DNA and RNA chains [i](#)

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	B	1712/1747 (97%)	0.18	92 (5%) 25 20	29, 63, 123, 183	0
2	J	252/263 (95%)	0.10	8 (3%) 47 40	37, 60, 100, 172	0
All	All	1964/2010 (97%)	0.17	100 (5%) 28 22	29, 62, 120, 183	0

The worst 5 of 100 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	J	2318	ALA	8.0
1	B	1880	ASN	6.5
2	J	2319	LEU	6.4
1	B	2099	THR	6.1
1	B	1165	ILE	5.8

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.