wwPDB X-ray Structure Validation Summary Report (i)

Jun 18, 2024 – 04:52 PM EDT

PDB ID	:	4ADY
Title	:	Crystal structure of 26S proteasome subunit Rpn2
Authors	:	Kulkarni, K.; He, J.; Da Fonseca, P.C.A.; Krutauz, D.; Glickman, M.H.; Barford, D.; Morris, E.P.
Deposited on	:	2012-01-04
Resolution	:	2.70 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 (i) 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 (i)) were used in the production of this report:

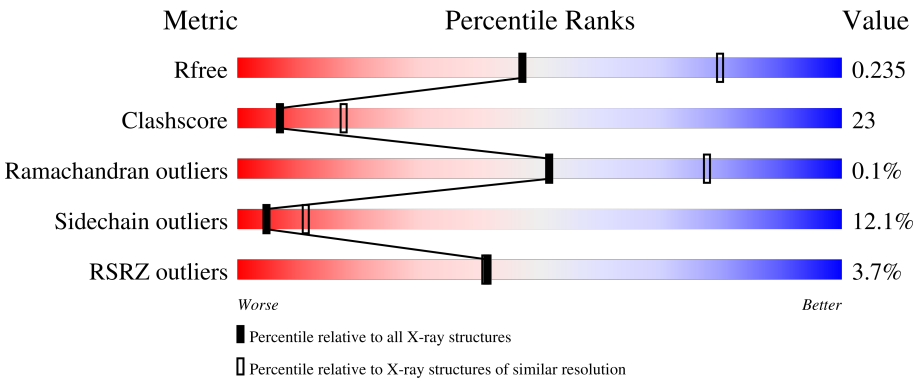
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.20.1
EDS	:	2.37.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

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 2.70 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	2808 (2.70-2.70)
Clashscore	141614	3122 (2.70-2.70)
Ramachandran outliers	138981	3069 (2.70-2.70)
Sidechain outliers	138945	3069 (2.70-2.70)
RSRZ outliers	127900	2737 (2.70-2.70)

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	963	3% 56% 27% 5% 12%
1	B	963	4% 50% 28% 6% 17%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 12581 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 26S PROTEASOME REGULATORY SUBUNIT RPN2.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	S	Se			
1	A	849	6498	4138	1076	1256	8	20	0	0	0
1	B	802	6024	3851	993	1152	8	20	0	0	0

There are 38 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-10	MSE	-	expression tag	UNP P32565
A	-9	ARG	-	expression tag	UNP P32565
A	-8	GLY	-	expression tag	UNP P32565
A	-7	SER	-	expression tag	UNP P32565
A	-6	HIS	-	expression tag	UNP P32565
A	-5	HIS	-	expression tag	UNP P32565
A	-4	HIS	-	expression tag	UNP P32565
A	-3	HIS	-	expression tag	UNP P32565
A	-2	HIS	-	expression tag	UNP P32565
A	-1	HIS	-	expression tag	UNP P32565
A	0	GLY	-	expression tag	UNP P32565
A	1	SER	-	expression tag	UNP P32565
A	946	LEU	-	expression tag	UNP P32565
A	947	GLN	-	expression tag	UNP P32565
A	948	PRO	-	expression tag	UNP P32565
A	949	SER	-	expression tag	UNP P32565
A	950	LEU	-	expression tag	UNP P32565
A	951	ILE	-	expression tag	UNP P32565
A	952	SER	-	expression tag	UNP P32565
B	-10	MSE	-	expression tag	UNP P32565
B	-9	ARG	-	expression tag	UNP P32565
B	-8	GLY	-	expression tag	UNP P32565
B	-7	SER	-	expression tag	UNP P32565
B	-6	HIS	-	expression tag	UNP P32565
B	-5	HIS	-	expression tag	UNP P32565

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	-4	HIS	-	expression tag	UNP P32565
B	-3	HIS	-	expression tag	UNP P32565
B	-2	HIS	-	expression tag	UNP P32565
B	-1	HIS	-	expression tag	UNP P32565
B	0	GLY	-	expression tag	UNP P32565
B	1	SER	-	expression tag	UNP P32565
B	946	LEU	-	expression tag	UNP P32565
B	947	GLN	-	expression tag	UNP P32565
B	948	PRO	-	expression tag	UNP P32565
B	949	SER	-	expression tag	UNP P32565
B	950	LEU	-	expression tag	UNP P32565
B	951	ILE	-	expression tag	UNP P32565
B	952	SER	-	expression tag	UNP P32565

- Molecule 2 is water.

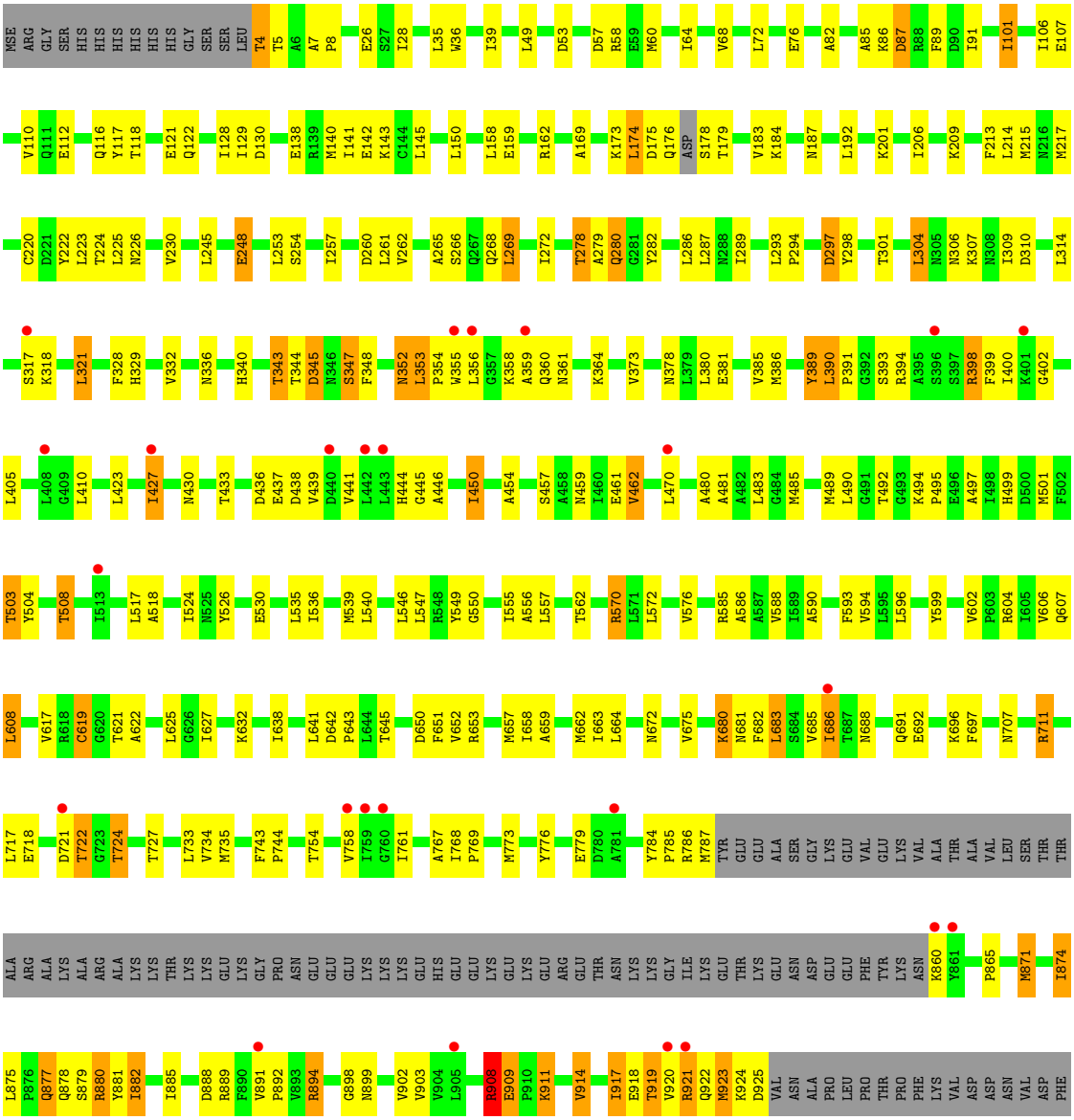
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
2	A	29	29	29	0	0
			Total	O		
2	B	30	30	30	0	0

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: 26S PROTEASOME REGULATORY SUBUNIT RPN2

Chain A:



PRO SER ALA LEU GLN PRO SER LEU ILE SER

- Molecule 1: 26S PROTEASOME REGULATORY SUBUNIT RPN2

Chain B: 4% 50% 28% 6% 17%

MSE ARG GLY SER HIS HIS HIS HIS HIS HIS GLY SER SER LEU THR THR ALA ALA PRO LEU LEU ALA LEU LEU ARG GLU ASN GLN ASP SER VAL LYS THR TYR ALA LEU GLU SER ILE ASN ASN VAL VAL ASP GLN LEU TRP SER GLU ILE SER ASN GLU LEU PRO ASP ILE GLU ALA LEU

TYR ASP ASP ASP THR PHE SER ASP R58 E59 M60 A61 A62 L63 I64 A65 S66 K67 V68 Y69 L72 G73 E74 Y75 V79 K80 Y81 A82 L83 F89 D90 I91 D92 E93 K94 S95 Q96 F97 V98 E99 T100 I101 V102 S103 K104 S105 I106 E107 M108 Y109 V110 S114 Y117 D120

E121 Q122 F123 Y124 T125 D130 L133 T134 S135 I136 F137 M140 I141 E142 K143 C144 A147 S148 E149 L150 K151 L152 A153 L154 L158 E159 G160 Y161 R162 L163 D164 I165 I166 L170 D175 Q176 D177 S178 T179 S180 E181 N182 V183 K184 I185 I186 N187 Y188 L189 L190 T191 L192 A193

I194 V197 F202 R203 S204 S205 I206 L207 R208 K209 S210 F211 D212 F213 L214 M217 P218 N219 C220 D221 Y222 L223 T224 L225 N226 K227 V230 N231 D234 A235 G236 L237 A238 L239 Q240 L241 F242 K243 K244 L245 E248 G252 L253 S254 I257 A258 F259 V262 Q268 L269

L270 E271 I272 L273 V274 L277 T278 A279 Q280 G281 Y282 D283 P284 A285 L286 L287 N288 I289 L290 S291 G292 L293 P294 T295 Y299 L303 L304 K307 N308 I309 D310 I311 L314 D322 G323 K324 H329 H340 D345 M346 S347 F348 I349 K350 A351 N352 L353 P354 M355 L356

A359 Q360 N361 K364 L371 G372 V373 I374 H375 L380 E381 V385 M386 A387 P388 Y389 L390 S393 R394 A395 S396 S397 R398 F399 I400 K401 G407 L410 I411 G414 F415 D418 T419 K424 I427 V428 E429 N430 S431 G432 T433 E437 D438 V439 D440 V441 L442

M455 N459 V462 Y463 E464 L470 Y471 N472 T476 S477 G478 A481 A482 L483 G484 M485 M489 K494 A497 I498 H499 D500 M501 F502 T503 Q506 E507 T508 I513 I514 R515 G516 L517 L521 I524 M539 D543 L546 L547 R548 Y549 G550 G551 A556

G561 N565 R570 D578 D581 A590 F593 L596 R597 T600 Q607 L608 L609 S610 H616 V617 T621 A622 F623 I627 I638 D642 P643 I644 T645 K646 D647 P648 V649 D650 F651 V652 R653 Q654 M657 I658 A659 M662 I663 L664 K670

V675 N681 F682 L683 S684 V685 I686 T687 N688 E692 G693 L694 C700 R711 Q716 L717 E718 N719 T722 G723 T724 L725 K728 L733 V734 M735 F743 P744 T754 P755 V758 I761 R762 D765 Q766 A767 I768 P769 K770 F771 Q772 M773 D780 A781 F782

P785 R786 M787 Y788 E789 GLU ALA SER GLY LYS GLU VAL GLU LYS VAL ALA THR ALA VAL LEU SER THR THR ALA ARG ALA LYS ALA ARG ALA LYS LYS THR LYS LYS GLU LYS GLY PRO ASN GLU GLU GLU LYS LYS LYS GLU HIS GLU GLU LYS GLU LYS GLU ARG GLU THR ASN LYS LYS

GLY ILE LYS GLU THR LYS GLU GLU ASN ASP GLU GLU F856 K864 P865 Y866 K867 V868 I874 Q878 S879 I882 S883 F884 I885 K886 D887 D888 R889 F890 V893 R894 K895 F896 V902 L905 E909 P910 K911 E912 P913 V914 A915 L916 I917 E918 T919 V920 R921 Q922 M923 K924 D925

VAL ASN ALA PRO LEU PRO THR PRO PHE LYS VAL ASP ASP ASN VAL ASP PHE PRO SER ALA LEU GLN PRO SER LEU ILE SER

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	73.96Å 148.66Å 114.03Å 90.00° 107.39° 90.00°	Depositor
Resolution (Å)	70.58 – 2.70 74.33 – 2.70	Depositor EDS
% Data completeness (in resolution range)	99.7 (70.58-2.70) 99.9 (74.33-2.70)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.26 (at 2.69Å)	Xtriage
Refinement program	PHENIX 1.7.2_869	Depositor
R, R_{free}	0.183 , 0.240 0.183 , 0.235	Depositor DCC
R_{free} test set	1323 reflections (2.05%)	wwPDB-VP
Wilson B-factor (Å^2)	56.7	Xtriage
Anisotropy	0.639	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 67.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	0.028 for h,-k,-h-l	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	12581	wwPDB-VP
Average B, all atoms (Å^2)	79.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.44% 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 (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	A	0.43	0/6586	0.60	0/8893
1	B	0.41	0/6106	0.59	0/8252
All	All	0.42	0/12692	0.60	0/17145

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

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	248	GLU	Peptide
1	A	908	ARG	Sidechain

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	6498	0	6471	249	0
1	B	6024	0	5933	323	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	A	29	0	0	2	0
2	B	30	0	0	1	0
All	All	12581	0	12404	570	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 23.

The worst 5 of 570 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:918:GLU:CD	1:A:923:MSE:SE	2.38	1.12
1:A:398:ARG:HD3	1:A:438:ASP:HA	1.25	1.11
1:A:919:THR:HG22	1:A:922:GLN:H	1.08	1.09
1:B:371:LEU:HB3	1:B:386:MSE:HE1	1.33	1.08
1:A:570:ARG:HG3	1:A:570:ARG:HH11	1.21	1.05

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	843/963 (88%)	829 (98%)	14 (2%)	0	100	100
1	B	798/963 (83%)	783 (98%)	14 (2%)	1 (0%)	51	78
All	All	1641/1926 (85%)	1612 (98%)	28 (2%)	1 (0%)	51	78

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	910	PRO

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	694/792 (88%)	612 (88%)	82 (12%)	5 12
1	B	623/792 (79%)	546 (88%)	77 (12%)	4 11
All	All	1317/1584 (83%)	1158 (88%)	159 (12%)	5 11

5 of 159 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	257	ILE
1	B	600	THR
1	B	289	ILE
1	B	398	ARG
1	B	728	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 15 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	300	ASN
1	B	681	ASN
1	B	329	HIS
1	B	899	ASN
1	B	510	HIS

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 (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	A	829/963 (86%)	0.42	25 (3%)	50	51	40, 73, 114, 144	0
1	B	782/963 (81%)	0.43	35 (4%)	33	31	35, 80, 136, 173	0
All	All	1611/1926 (83%)	0.42	60 (3%)	41	41	35, 75, 129, 173	0

The worst 5 of 60 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	884	PHE	6.5
1	B	83	LEU	5.3
1	A	920	VAL	4.2
1	B	396	SER	3.8
1	A	860	LYS	3.6

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.