wwPDB X-ray Structure Validation Summary Report (i)

Oct 30, 2023 – 08:44 PM JST

PDB ID	:	4ZG5
Title	:	Structural and functional insights into Survival endonuclease, an important virulence factor of Brucella abortus
Authors	:	Tarique, K.F.; Abdul Rehman, S.A.; Devi, S.; Gourinath, S.
Deposited on	:	2015-04-22
Resolution	:	1.90 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 (1)) were used in the production of this report:

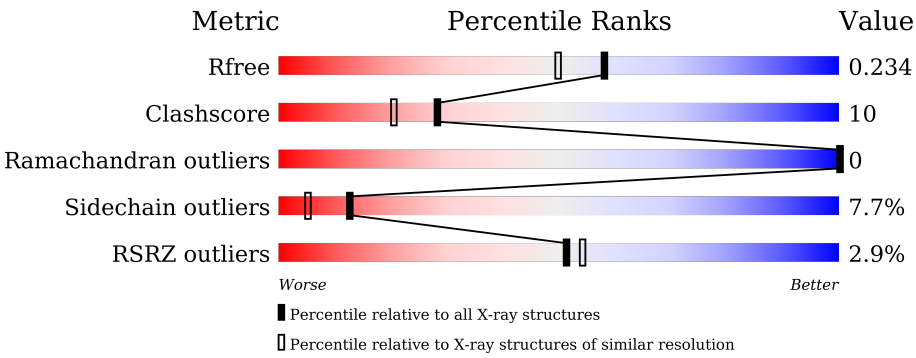
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation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

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 1.90 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	6207 (1.90-1.90)
Clashscore	141614	6847 (1.90-1.90)
Ramachandran outliers	138981	6760 (1.90-1.90)
Sidechain outliers	138945	6760 (1.90-1.90)
RSRZ outliers	127900	6082 (1.90-1.90)

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	263	3% 75% 15% • 6%
1	C	263	3% 73% 18% • • 5%
1	D	263	2% 78% 15% • •
1	G	263	3% 73% 17% • 6%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 7696 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 5'-nucleotidase SurE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	D	252	1894	1186	342	357	9	0	0	0
1	A	247	1861	1167	337	348	9	0	0	0
1	C	249	1869	1172	334	354	9	0	0	0
1	G	247	1853	1160	337	347	9	0	0	0

There are 32 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	256	LEU	-	expression tag	UNP B2S5B9
D	257	GLU	-	expression tag	UNP B2S5B9
D	258	HIS	-	expression tag	UNP B2S5B9
D	259	HIS	-	expression tag	UNP B2S5B9
D	260	HIS	-	expression tag	UNP B2S5B9
D	261	HIS	-	expression tag	UNP B2S5B9
D	262	HIS	-	expression tag	UNP B2S5B9
D	263	HIS	-	expression tag	UNP B2S5B9
A	256	LEU	-	expression tag	UNP B2S5B9
A	257	GLU	-	expression tag	UNP B2S5B9
A	258	HIS	-	expression tag	UNP B2S5B9
A	259	HIS	-	expression tag	UNP B2S5B9
A	260	HIS	-	expression tag	UNP B2S5B9
A	261	HIS	-	expression tag	UNP B2S5B9
A	262	HIS	-	expression tag	UNP B2S5B9
A	263	HIS	-	expression tag	UNP B2S5B9
C	256	LEU	-	expression tag	UNP B2S5B9
C	257	GLU	-	expression tag	UNP B2S5B9
C	258	HIS	-	expression tag	UNP B2S5B9
C	259	HIS	-	expression tag	UNP B2S5B9
C	260	HIS	-	expression tag	UNP B2S5B9

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	261	HIS	-	expression tag	UNP B2S5B9
C	262	HIS	-	expression tag	UNP B2S5B9
C	263	HIS	-	expression tag	UNP B2S5B9
G	256	LEU	-	expression tag	UNP B2S5B9
G	257	GLU	-	expression tag	UNP B2S5B9
G	258	HIS	-	expression tag	UNP B2S5B9
G	259	HIS	-	expression tag	UNP B2S5B9
G	260	HIS	-	expression tag	UNP B2S5B9
G	261	HIS	-	expression tag	UNP B2S5B9
G	262	HIS	-	expression tag	UNP B2S5B9
G	263	HIS	-	expression tag	UNP B2S5B9

- Molecule 2 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	D	1	Total Mg 1 1	0	0
2	A	1	Total Mg 1 1	0	0
2	C	1	Total Mg 1 1	0	0
2	G	1	Total Mg 1 1	0	0

- Molecule 3 is water.

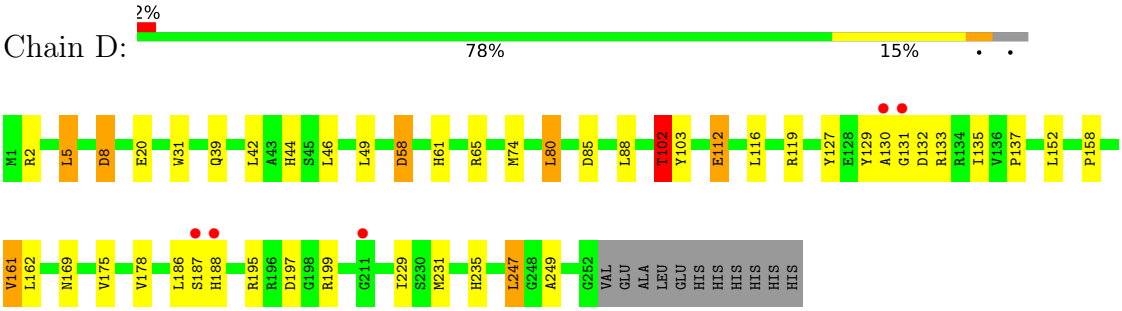
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	D	64	Total O 64 64	0	0
3	A	51	Total O 51 51	0	0
3	C	56	Total O 56 56	0	0
3	G	44	Total O 44 44	0	0

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.

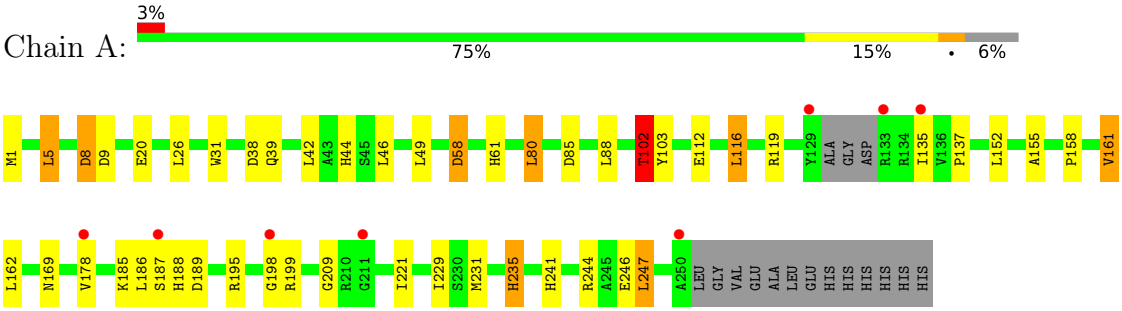
- Molecule 1: 5'-nucleotidase SurE

Chain D:



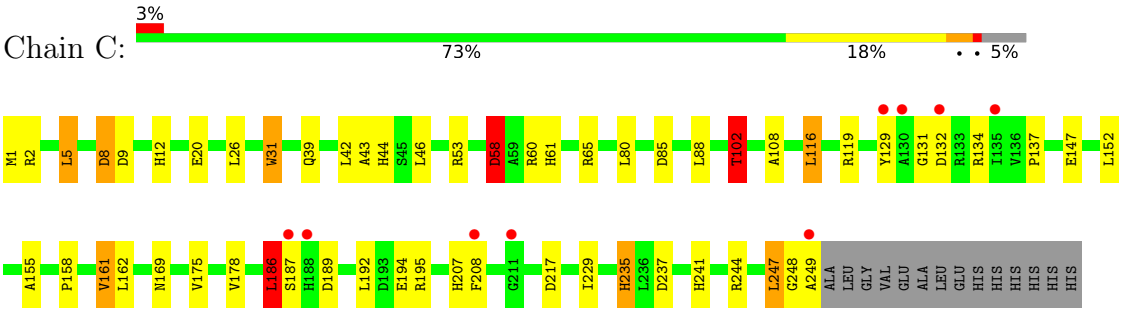
- Molecule 1: 5'-nucleotidase SurE

Chain A:



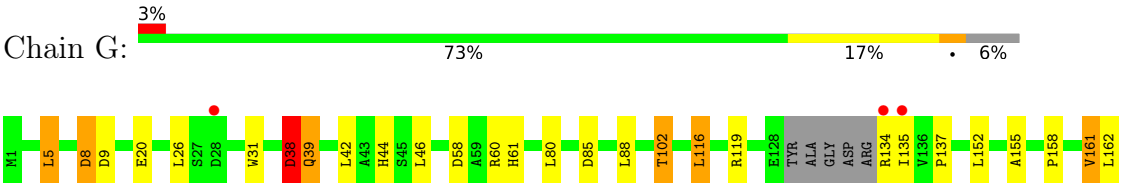
- Molecule 1: 5'-nucleotidase SurE

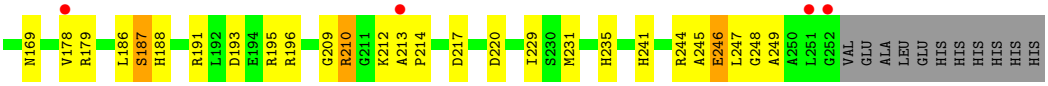
Chain C:



- Molecule 1: 5'-nucleotidase SurE

Chain G:




N169
V178
R179
L186
S187
H188
R191
L192
D193
E194
R195
R196
G209
R210
G211
K212
A213
P214
D217
D220
I229
S230
M231
H235
H241
R244
A245
E246
L247
G248
A249
A250
L251
G252
VAL
GLU
ALA
LEU
GLU
HIS
HIS
HIS
HIS
HIS
HIS

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	50.18Å 121.07Å 82.43Å 90.00° 93.58° 90.00°	Depositor
Resolution (Å)	50.00 – 1.90 41.37 – 1.90	Depositor EDS
% Data completeness (in resolution range)	97.7 (50.00-1.90) 97.7 (41.37-1.90)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.66 (at 1.89Å)	Xtriage
Refinement program	REFMAC 5.8.0049	Depositor
R, R_{free}	0.198 , 0.227 0.207 , 0.234	Depositor DCC
R_{free} test set	3807 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å^2)	30.4	Xtriage
Anisotropy	0.707	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 46.9	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	7696	wwPDB-VP
Average B, all atoms (Å^2)	40.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.29% 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)

Bond lengths and bond angles in the following residue types are not validated in this section: MG

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	0.96	1/1898 (0.1%)	1.14	8/2582 (0.3%)
1	C	1.02	4/1906 (0.2%)	1.17	15/2595 (0.6%)
1	D	0.99	3/1932 (0.2%)	1.10	11/2629 (0.4%)
1	G	0.91	0/1889	1.07	13/2569 (0.5%)
All	All	0.97	8/7625 (0.1%)	1.12	47/10375 (0.5%)

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	3
1	C	0	1
1	D	0	1
All	All	0	5

The worst 5 of 8 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	102	THR	CB-CG2	-6.97	1.29	1.52
1	C	58	ASP	CB-CG	-6.65	1.37	1.51
1	C	31	TRP	CB-CG	-6.57	1.38	1.50
1	D	102	THR	CB-CG2	-6.38	1.31	1.52
1	C	102	THR	CB-CG2	-6.07	1.32	1.52

The worst 5 of 47 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	8	ASP	CB-CG-OD1	15.76	132.48	118.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	8	ASP	OD1-CG-OD2	-8.36	107.43	123.30
1	C	195	ARG	NE-CZ-NH1	7.58	124.09	120.30
1	D	195	ARG	NE-CZ-NH1	7.45	124.02	120.30
1	A	9	ASP	CB-CG-OD2	-7.28	111.75	118.30

There are no chirality outliers.

All (5) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	198	GLY	Peptide
1	A	246	GLU	Peptide
1	A	247	LEU	Peptide
1	C	58	ASP	Mainchain
1	D	130	ALA	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	A	1861	0	1833	57	0
1	C	1869	0	1838	33	0
1	D	1894	0	1864	46	0
1	G	1853	0	1827	47	0
2	A	1	0	0	0	0
2	C	1	0	0	0	0
2	D	1	0	0	0	0
2	G	1	0	0	0	0
3	A	51	0	0	2	0
3	C	56	0	0	2	0
3	D	64	0	0	1	0
3	G	44	0	0	3	0
All	All	7696	0	7362	146	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 10.

The worst 5 of 146 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:D:39:GLN:NE2	1:D:42:LEU:CD2	2.23	1.01
1:A:49:LEU:HD13	1:G:39:GLN:OE1	1.69	0.90
1:A:186:LEU:HD11	1:A:221:ILE:CD1	2.04	0.88
1:D:39:GLN:HE21	1:D:42:LEU:CD2	1.90	0.83
1:G:246:GLU:CD	1:G:246:GLU:H	1.81	0.82

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	243/263 (92%)	236 (97%)	7 (3%)	0	100	100
1	C	247/263 (94%)	242 (98%)	5 (2%)	0	100	100
1	D	250/263 (95%)	241 (96%)	9 (4%)	0	100	100
1	G	243/263 (92%)	236 (97%)	7 (3%)	0	100	100
All	All	983/1052 (93%)	955 (97%)	28 (3%)	0	100	100

There are no Ramachandran outliers to report.

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	190/207 (92%)	179 (94%)	11 (6%)	20	10
1	C	191/207 (92%)	173 (91%)	18 (9%)	8	3

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	D	193/207 (93%)	180 (93%)	13 (7%)	16 7
1	G	189/207 (91%)	172 (91%)	17 (9%)	9 3
All	All	763/828 (92%)	704 (92%)	59 (8%)	13 5

5 of 59 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	58	ASP
1	G	210	ARG
1	C	175	VAL
1	G	187	SER
1	G	134	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 23 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	96	ASN
1	C	241	HIS
1	C	169	ASN
1	G	61	HIS
1	D	188	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)

Of 4 ligands modelled in this entry, 4 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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	247/263 (93%)	0.15	8 (3%)	47	50	24, 38, 67, 94	0
1	C	249/263 (94%)	0.08	9 (3%)	42	45	24, 36, 68, 83	0
1	D	252/263 (95%)	-0.04	5 (1%)	65	68	24, 35, 66, 84	0
1	G	247/263 (93%)	0.13	7 (2%)	53	56	26, 41, 71, 79	0
All	All	995/1052 (94%)	0.08	29 (2%)	51	54	24, 37, 68, 94	0

The worst 5 of 29 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	198	GLY	4.8
1	C	129	TYR	4.7
1	G	135	ILE	4.1
1	A	135	ILE	4.0
1	G	28	ASP	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)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	MG	A	301	1/1	0.85	0.06	31,31,31,31	0
2	MG	D	301	1/1	0.98	0.04	28,28,28,28	0
2	MG	C	301	1/1	0.98	0.05	26,26,26,26	0
2	MG	G	301	1/1	0.99	0.07	31,31,31,31	0

6.5 Other polymers (i)

There are no such residues in this entry.