Full wwPDB X-ray Structure Validation Report ⓘ

May 14, 2020 – 09:05 am BST

PDB ID	:	4P1T
Title	:	Crystal structure of the DBL3X-DBL4epsilon double domain from the extracellular part of VAR2CSA PfEMP1 from Plasmodium falciparum
Authors	:	Gangnard, S.; Dechavanne, S.; Srivastava, A.; Amirat, F.; Gamain, B.; Lewit-Bentley, A.; Bentley, G.A.
Deposited on	:	2014-02-27
Resolution	:	2.90 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 following versions of software and data (see references ①) were used in the production of this report:

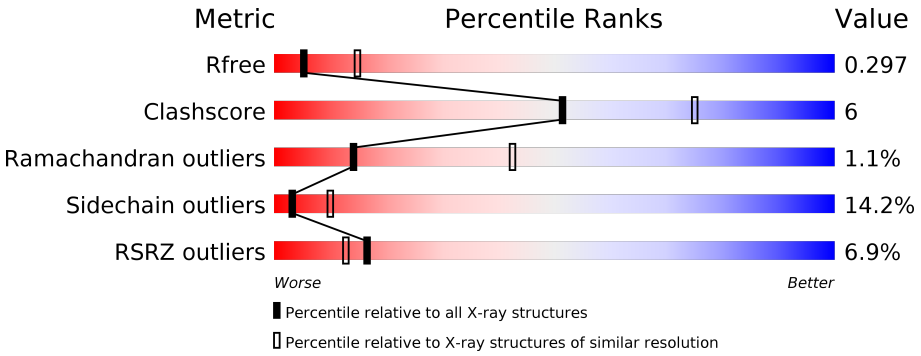
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 2.90 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	1957 (2.90-2.90)
Clashscore	141614	2172 (2.90-2.90)
Ramachandran outliers	138981	2115 (2.90-2.90)
Sidechain outliers	138945	2117 (2.90-2.90)
RSRZ outliers	127900	1906 (2.90-2.90)

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 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 <=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	746	6% 67% 17% • 13%

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 5090 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 Erythrocyte membrane protein 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	649	Total	C	N	O	S	0	1	0
			5076	3203	852	984	37			

There are 18 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1211	MET	-	expression tag	UNP Q6UDW7
A	1212	ALA	-	expression tag	UNP Q6UDW7
A	1213	SER	-	expression tag	UNP Q6UDW7
A	1214	MET	-	expression tag	UNP Q6UDW7
A	1222	GLN	ASN	engineered mutation	UNP Q6UDW7
A	1290	GLN	ASN	engineered mutation	UNP Q6UDW7
A	1430	GLY	SER	engineered mutation	UNP Q6UDW7
A	1594	LEU	SER	engineered mutation	UNP Q6UDW7
A	1746	ALA	THR	engineered mutation	UNP Q6UDW7
A	1751	PRO	THR	engineered mutation	UNP Q6UDW7
A	1846	ALA	THR	engineered mutation	UNP Q6UDW7
A	1916	GLN	ASN	engineered mutation	UNP Q6UDW7
A	1951	HIS	-	expression tag	UNP Q6UDW7
A	1952	HIS	-	expression tag	UNP Q6UDW7
A	1953	HIS	-	expression tag	UNP Q6UDW7
A	1954	HIS	-	expression tag	UNP Q6UDW7
A	1955	HIS	-	expression tag	UNP Q6UDW7
A	1956	HIS	-	expression tag	UNP Q6UDW7

- Molecule 2 is water.

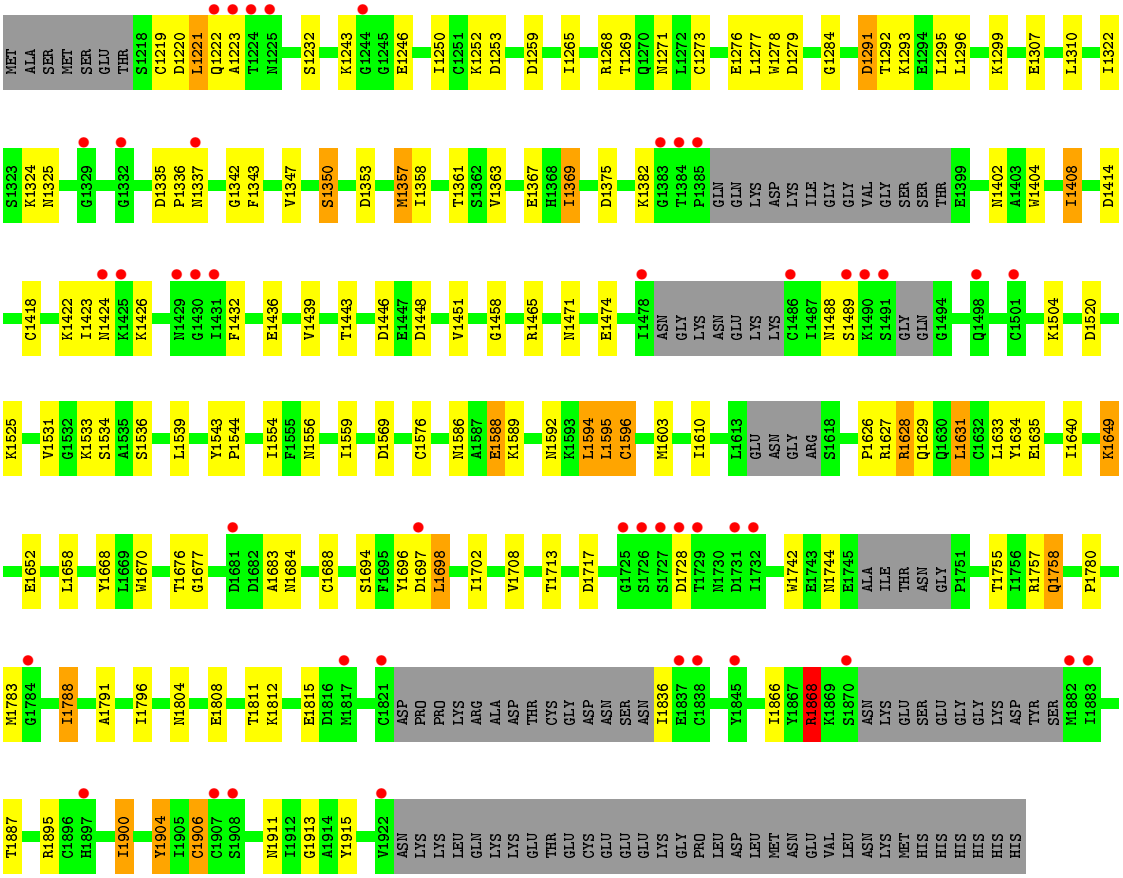
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	14	Total	O	0	0
			14	14		

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 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: Erythrocyte membrane protein 1

Chain A:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α, β, γ	99.79Å 129.34Å 64.06Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	28.16 – 2.90 27.73 – 2.90	Depositor EDS
% Data completeness (in resolution range)	92.2 (28.16-2.90) 92.2 (27.73-2.90)	Depositor EDS
R_{merge}	0.26	Depositor
R_{sym}	0.24	Depositor
$< I/\sigma(I) >$ [1]	1.33 (at 2.90Å)	Xtriage
Refinement program	BUSTER 2.11.4	Depositor
R, R_{free}	0.203 , 0.276 0.222 , 0.297	Depositor DCC
R_{free} test set	903 reflections (5.16%)	wwPDB-VP
Wilson B-factor (Å^2)	52.3	Xtriage
Anisotropy	0.631	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , 81.7	EDS
L-test for twinning[2]	$< \|L\| > = 0.47$, $< L^2 > = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.90	EDS
Total number of atoms	5090	wwPDB-VP
Average B, all atoms (Å^2)	60.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.10% 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	0.52	0/5190	0.75	0/7014

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	5076	0	4650	54	0
2	A	14	0	0	0	0
All	All	5090	0	4650	54	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 6.

All (54) 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:1335:ASP:HB2	1:A:1336:PRO:HD2	1.64	0.80
1:A:1424:ASN:HB2	1:A:1432:PHE:HB2	1.72	0.71
1:A:1278:TRP:HH2	1:A:1367:GLU:HG2	1.59	0.68
1:A:1357:MET:HE1	1:A:1369:ILE:HG12	1.77	0.67

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:1232:SER:HA	1:A:1310:LEU:HD22	1.78	0.65
1:A:1592:ASN:C	1:A:1594:LEU:H	2.03	0.61
1:A:1322:ILE:HG22	1:A:1342:GLY:HA3	1.84	0.59
1:A:1279:ASP:O	1:A:1284:GLY:HA2	2.07	0.55
1:A:1658:LEU:HD21	1:A:1702:ILE:HD11	1.89	0.55
1:A:1900:ILE:HG12	1:A:1906:CYS:HA	1.87	0.54
1:A:1677:GLY:HA3	1:A:1683:ALA:HA	1.90	0.54
1:A:1361:THR:HG22	1:A:1363:VAL:H	1.73	0.53
1:A:1423:ILE:HA	1:A:1426:LYS:HD3	1.90	0.53
1:A:1291:ASP:OD1	1:A:1295:LEU:HB3	2.09	0.52
1:A:1250:ILE:HG23	1:A:1265:ILE:HD12	1.93	0.51
1:A:1696:TYR:CD1	1:A:1791:ALA:HB2	2.45	0.51
1:A:1628:ARG:HH22	1:A:1713:THR:HG21	1.74	0.50
1:A:1596:CYS:HB3	1:A:1634:TYR:HE2	1.76	0.50
1:A:1868:ARG:HE	1:A:1887:THR:HG22	1.77	0.49
1:A:1788:ILE:HD12	1:A:1788:ILE:H	1.76	0.49
1:A:1268:ARG:HH21	1:A:1350:SER:HA	1.78	0.48
1:A:1458:GLY:HA3	1:A:1539:LEU:HD22	1.96	0.47
1:A:1811:THR:O	1:A:1815:GLU:HB3	2.14	0.47
1:A:1629:GLN:O	1:A:1631:LEU:HD22	2.14	0.47
1:A:1276:GLU:HG3	1:A:1299:LYS:HG2	1.96	0.46
1:A:1631:LEU:HG	1:A:1633:LEU:HG	1.98	0.46
1:A:1268:ARG:HB3	1:A:1268:ARG:HH11	1.81	0.45
1:A:1471:ASN:HB3	1:A:1504:LYS:HD2	1.98	0.45
1:A:1596:CYS:HB3	1:A:1634:TYR:CE2	2.52	0.45
1:A:1223:ALA:HB1	1:A:1422:LYS:HE3	1.99	0.44
1:A:1559:ILE:HD12	1:A:1559:ILE:H	1.82	0.44
1:A:1595:LEU:HA	1:A:1668:TYR:CE2	2.52	0.44
1:A:1543:TYR:HA	1:A:1544:PRO:HD2	1.85	0.44
1:A:1292:THR:HG23	1:A:1295:LEU:HD22	2.00	0.43
1:A:1465:ARG:HD2	1:A:1569:ASP:OD2	2.18	0.43
1:A:1273:CYS:HB3	1:A:1307:GLU:HB2	2.01	0.43
1:A:1694:SER:O	1:A:1698:LEU:HB2	2.19	0.43
1:A:1755:THR:OG1	1:A:1758:GLN:HB2	2.19	0.43
1:A:1343:PHE:O	1:A:1347:VAL:HG23	2.19	0.43
1:A:1603:MET:HB3	1:A:1626:PRO:HB3	2.01	0.43
1:A:1649:LYS:HA	1:A:1652:GLU:HB2	2.00	0.43
1:A:1448:ASP:HB3	1:A:1451:VAL:HG22	2.01	0.43
1:A:1269:THR:HG23	1:A:1353:ASP:OD2	2.19	0.42
1:A:1688:CYS:SG	1:A:1783:MET:HG2	2.59	0.42
1:A:1322:ILE:CG2	1:A:1342:GLY:HA3	2.49	0.42

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:1780:PRO:HG2	1:A:1783:MET:HB2	2.02	0.42
1:A:1219[A]:CYS:HB2	1:A:1221:LEU:HG	2.02	0.42
1:A:1592:ASN:C	1:A:1594:LEU:N	2.73	0.42
1:A:1757:ARG:HE	1:A:1788:ILE:HD13	1.84	0.42
1:A:1588:GLU:HB2	1:A:1589:LYS:H	1.68	0.41
1:A:1404:TRP:NE1	1:A:1408:ILE:HD13	2.35	0.41
1:A:1554:ILE:HG12	1:A:1569:ASP:HB2	2.03	0.41
1:A:1742:TRP:HE1	1:A:1757:ARG:HG3	1.85	0.41
1:A:1335:ASP:CB	1:A:1336:PRO:HD2	2.44	0.41

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	634/746 (85%)	566 (89%)	61 (10%)	7 (1%)	14 42

All (7) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	1443	THR
1	A	1595	LEU
1	A	1868	ARG
1	A	1904	TYR
1	A	1222	GLN
1	A	1594	LEU
1	A	1913	GLY

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	515/661 (78%)	442 (86%)	73 (14%)	3 10

All (73) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1220	ASP
1	A	1221	LEU
1	A	1243	LYS
1	A	1246	GLU
1	A	1252	LYS
1	A	1253	ASP
1	A	1259	ASP
1	A	1271	ASN
1	A	1277	LEU
1	A	1291	ASP
1	A	1293	LYS
1	A	1296	LEU
1	A	1324	LYS
1	A	1325	ASN
1	A	1337	ASN
1	A	1350	SER
1	A	1357	MET
1	A	1358	ILE
1	A	1369	ILE
1	A	1375	ASP
1	A	1382	LYS
1	A	1402	ASN
1	A	1408	ILE
1	A	1414	ASP
1	A	1418	CYS
1	A	1436	GLU
1	A	1439	VAL
1	A	1446	ASP
1	A	1474	GLU
1	A	1488	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	1489	SER
1	A	1520	ASP
1	A	1525	LYS
1	A	1531	VAL
1	A	1533	LYS
1	A	1534	SER
1	A	1536	SER
1	A	1556	ASN
1	A	1576	CYS
1	A	1586	ASN
1	A	1588	GLU
1	A	1596	CYS
1	A	1610	ILE
1	A	1627	ARG
1	A	1628	ARG
1	A	1631	LEU
1	A	1635	GLU
1	A	1640	ILE
1	A	1649	LYS
1	A	1670	TRP
1	A	1676	THR
1	A	1684	ASN
1	A	1697	ASP
1	A	1698	LEU
1	A	1708	VAL
1	A	1717	ASP
1	A	1728	ASP
1	A	1744	ASN
1	A	1758	GLN
1	A	1788	ILE
1	A	1796	ILE
1	A	1804	ASN
1	A	1808	GLU
1	A	1812	LYS
1	A	1836	ILE
1	A	1866	ILE
1	A	1868	ARG
1	A	1895	ARG
1	A	1900	ILE
1	A	1904	TYR
1	A	1906	CYS
1	A	1911	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	1915	TYR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	1368	HIS
1	A	1433	ASN
1	A	1556	ASN
1	A	1585	ASN

5.3.3 RNA

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates

There are no carbohydrates in this entry.

5.6 Ligand geometry

There are no ligands in this entry.

5.7 Other polymers

There are no such residues in this entry.

5.8 Polymer linkage issues

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^{th} 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	649/746 (86%)	0.22	45 (6%)	16	13	29, 58, 99, 126	0

All (45) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	1224	THR	7.1
1	A	1727	SER	5.8
1	A	1431	ILE	5.1
1	A	1486	CYS	4.5
1	A	1385	PRO	4.3
1	A	1728	ASP	4.2
1	A	1332	GLY	4.1
1	A	1908	SER	3.9
1	A	1329	GLY	3.8
1	A	1478	ILE	3.7
1	A	1489	SER	3.5
1	A	1424	ASN	3.5
1	A	1225	ASN	3.3
1	A	1429	ASN	3.3
1	A	1732	ILE	3.3
1	A	1725	GLY	3.2
1	A	1731	ASP	3.2
1	A	1817	MET	3.0
1	A	1501	CYS	2.9
1	A	1223	ALA	2.9
1	A	1907	CYS	2.9
1	A	1430	GLY	2.8
1	A	1729	THR	2.8
1	A	1883	ILE	2.8
1	A	1897	HIS	2.7
1	A	1384	THR	2.7
1	A	1784	GLY	2.7

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	1222	GLN	2.6
1	A	1726	SER	2.4
1	A	1697	ASP	2.3
1	A	1821	CYS	2.3
1	A	1383	GLY	2.3
1	A	1882	MET	2.3
1	A	1490	LYS	2.2
1	A	1244	GLY	2.2
1	A	1425	LYS	2.2
1	A	1838	CYS	2.2
1	A	1837	GLU	2.2
1	A	1870	SER	2.1
1	A	1498	GLN	2.1
1	A	1491	SER	2.1
1	A	1922	VAL	2.1
1	A	1845	TYR	2.1
1	A	1337	ASN	2.0
1	A	1681	ASP	2.0

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.