wwPDB X-ray Structure Validation Summary Report (i)

Aug 21, 2023 – 03:52 PM EDT

PDB ID	:	2PVX
Title	:	NMR and X-ray Analysis of Structural Additivity in Metal Binding Site-Swapped Hybrids of Rubredoxin
Authors	:	Wang, L.; LeMaster, D.M.; Hernandez, G.; Li, H.
Deposited on	:	2007-05-10
Resolution	:	1.04 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:

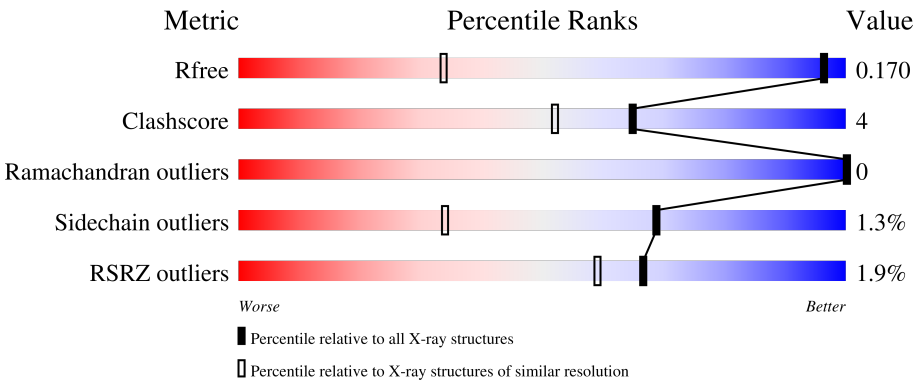
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.35
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

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 1.04 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	1596 (1.10-0.98)
Clashscore	141614	1677 (1.10-0.98)
Ramachandran outliers	138981	1591 (1.10-0.98)
Sidechain outliers	138945	1589 (1.10-0.98)
RSRZ outliers	127900	1557 (1.10-0.98)

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	54	85% 11% • •
1	B	54	2% 91% 7% •
1	C	54	4% 89% 9% •
1	D	54	4% 89% 11%
1	E	54	6% 80% 19% •

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	54	89% 9%
1	G	54	93% 7%
1	H	54	93% 7%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 7286 atoms, of which 3160 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 Rubredoxin.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
1	A	53	810	267	391	62	85	5	0	5	0
1	B	54	819	270	391	64	89	5	0	3	0
1	C	54	824	271	394	64	90	5	0	6	0
1	D	54	808	267	385	63	88	5	0	5	0
1	E	54	853	280	414	64	90	5	0	8	0
1	F	54	835	275	400	64	91	5	0	3	0
1	G	54	814	269	387	63	90	5	0	1	0
1	H	54	829	273	398	64	89	5	0	2	0

There are 64 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	2	LYS	ALA	engineered mutation	UNP P24297
A	7	THR	LYS	engineered mutation	UNP P24297
A	8	VAL	ILE	engineered mutation	UNP P24297
A	41	LEU	ILE	engineered mutation	UNP P24297
A	44	VAL	ALA	engineered mutation	UNP P24297
A	45	GLY	PRO	engineered mutation	UNP P24297
A	47	ASP	SER	engineered mutation	UNP P24297
A	48	GLN	GLU	engineered mutation	UNP P24297
B	102	LYS	ALA	engineered mutation	UNP P24297
B	107	THR	LYS	engineered mutation	UNP P24297
B	108	VAL	ILE	engineered mutation	UNP P24297
B	141	LEU	ILE	engineered mutation	UNP P24297
B	144	VAL	ALA	engineered mutation	UNP P24297

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	145	GLY	PRO	engineered mutation	UNP P24297
B	147	ASP	SER	engineered mutation	UNP P24297
B	148	GLN	GLU	engineered mutation	UNP P24297
C	202	LYS	ALA	engineered mutation	UNP P24297
C	207	THR	LYS	engineered mutation	UNP P24297
C	208	VAL	ILE	engineered mutation	UNP P24297
C	241	LEU	ILE	engineered mutation	UNP P24297
C	244	VAL	ALA	engineered mutation	UNP P24297
C	245	GLY	PRO	engineered mutation	UNP P24297
C	247	ASP	SER	engineered mutation	UNP P24297
C	248	GLN	GLU	engineered mutation	UNP P24297
D	302	LYS	ALA	engineered mutation	UNP P24297
D	307	THR	LYS	engineered mutation	UNP P24297
D	308	VAL	ILE	engineered mutation	UNP P24297
D	341	LEU	ILE	engineered mutation	UNP P24297
D	344	VAL	ALA	engineered mutation	UNP P24297
D	345	GLY	PRO	engineered mutation	UNP P24297
D	347	ASP	SER	engineered mutation	UNP P24297
D	348	GLN	GLU	engineered mutation	UNP P24297
E	402	LYS	ALA	engineered mutation	UNP P24297
E	407	THR	LYS	engineered mutation	UNP P24297
E	408	VAL	ILE	engineered mutation	UNP P24297
E	441	LEU	ILE	engineered mutation	UNP P24297
E	444	VAL	ALA	engineered mutation	UNP P24297
E	445	GLY	PRO	engineered mutation	UNP P24297
E	447	ASP	SER	engineered mutation	UNP P24297
E	448	GLN	GLU	engineered mutation	UNP P24297
F	502	LYS	ALA	engineered mutation	UNP P24297
F	507	THR	LYS	engineered mutation	UNP P24297
F	508	VAL	ILE	engineered mutation	UNP P24297
F	541	LEU	ILE	engineered mutation	UNP P24297
F	544	VAL	ALA	engineered mutation	UNP P24297
F	545	GLY	PRO	engineered mutation	UNP P24297
F	547	ASP	SER	engineered mutation	UNP P24297
F	548	GLN	GLU	engineered mutation	UNP P24297
G	602	LYS	ALA	engineered mutation	UNP P24297
G	607	THR	LYS	engineered mutation	UNP P24297
G	608	VAL	ILE	engineered mutation	UNP P24297
G	641	LEU	ILE	engineered mutation	UNP P24297
G	644	VAL	ALA	engineered mutation	UNP P24297
G	645	GLY	PRO	engineered mutation	UNP P24297
G	647	ASP	SER	engineered mutation	UNP P24297

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
G	648	GLN	GLU	engineered mutation	UNP P24297
H	702	LYS	ALA	engineered mutation	UNP P24297
H	707	THR	LYS	engineered mutation	UNP P24297
H	708	VAL	ILE	engineered mutation	UNP P24297
H	741	LEU	ILE	engineered mutation	UNP P24297
H	744	VAL	ALA	engineered mutation	UNP P24297
H	745	GLY	PRO	engineered mutation	UNP P24297
H	747	ASP	SER	engineered mutation	UNP P24297
H	748	GLN	GLU	engineered mutation	UNP P24297

- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total Zn 1 1	0	0
2	B	1	Total Zn 1 1	0	0
2	C	1	Total Zn 1 1	0	0
2	D	1	Total Zn 1 1	0	0
2	E	1	Total Zn 1 1	0	0
2	F	1	Total Zn 1 1	0	0
2	G	1	Total Zn 1 1	0	0
2	H	1	Total Zn 1 1	0	0

- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	80	Total O 82 82	0	2
3	B	77	Total O 79 79	0	2
3	C	87	Total O 89 89	0	2
3	D	74	Total O 74 74	0	0
3	E	76	Total O 79 79	0	3

Continued on next page...

Continued from previous page...

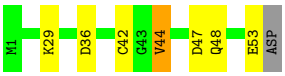
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
3	F	90	93	93	0	4
			Total	O		
3	G	76	80	80	0	6
			Total	O		
3	H	103	110	110	0	7

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.

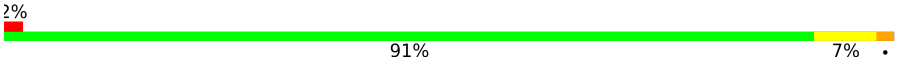
• Molecule 1: Rubredoxin

Chain A:



• Molecule 1: Rubredoxin

Chain B:

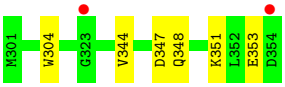


• Molecule 1: Rubredoxin

Chain C:

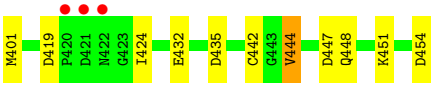
• Molecule 1: Rubredoxin

Chain D:



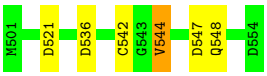
• Molecule 1: Rubredoxin

Chain E:

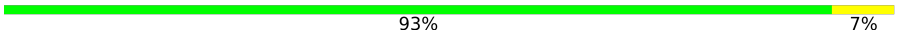


• Molecule 1: Rubredoxin

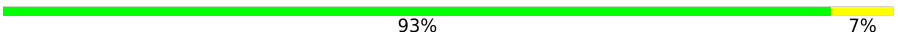
Chain F:

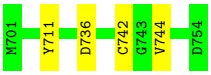


- Molecule 1: Rubredoxin

Chain G: 

- Molecule 1: Rubredoxin

Chain H: 



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	45.26Å 45.84Å 95.08Å 90.00° 98.43° 90.00°	Depositor
Resolution (Å)	10.00 – 1.04 27.87 – 1.00	Depositor EDS
% Data completeness (in resolution range)	95.4 (10.00-1.04) 85.8 (27.87-1.00)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	0.06	Depositor
$< I/\sigma(I) >$ [1]	0.63 (at 1.00Å)	Xtriage
Refinement program	SHELXL-97	Depositor
R, R_{free}	0.139 , 0.180 0.136 , 0.170	Depositor DCC
R_{free} test set	9834 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	9.9	Xtriage
Anisotropy	0.506	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.40 , 56.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.98	EDS
Total number of atoms	7286	wwPDB-VP
Average B, all atoms (Å^2)	19.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 20.22 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 9.1191e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

[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: ZN

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	0.76	0/438	1.12	2/595 (0.3%)
1	B	0.72	0/443	1.20	2/599 (0.3%)
1	C	0.72	0/449	1.26	2/608 (0.3%)
1	D	0.74	0/434	1.38	2/587 (0.3%)
1	E	0.75	0/466	1.35	4/630 (0.6%)
1	F	0.76	0/458	1.32	4/620 (0.6%)
1	G	0.82	0/442	1.24	2/598 (0.3%)
1	H	0.82	0/450	1.06	2/609 (0.3%)
All	All	0.76	0/3580	1.24	20/4846 (0.4%)

There are no bond length outliers.

The worst 5 of 20 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	D	353[A]	GLU	CA-CB-CG	10.43	136.34	113.40
1	F	536	ASP	CB-CG-OD2	-8.84	110.34	118.30
1	H	711	TYR	CG-CD2-CE2	7.69	127.45	121.30
1	E	447[A]	ASP	CB-CG-OD2	7.59	125.13	118.30
1	E	447[B]	ASP	CB-CG-OD2	7.59	125.13	118.30

There are no chirality outliers.

There are no planarity outliers.

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	419	391	373	5	0
1	B	428	391	382	3	0
1	C	430	394	371	4	0
1	D	423	385	356	2	0
1	E	439	414	394	7	0
1	F	435	400	391	3	0
1	G	427	387	382	1	0
1	H	431	398	391	1	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
2	C	1	0	0	0	0
2	D	1	0	0	0	0
2	E	1	0	0	0	0
2	F	1	0	0	0	0
2	G	1	0	0	0	0
2	H	1	0	0	0	0
3	A	82	0	0	1	0
3	B	79	0	0	2	0
3	C	89	0	0	3	0
3	D	74	0	0	0	0
3	E	79	0	0	5	0
3	F	93	0	0	1	0
3	G	80	0	0	0	0
3	H	110	0	0	0	0
All	All	4126	3160	3040	26	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.

The worst 5 of 26 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:E:419:ASP:OD1	3:E:966:HOH:O	2.07	0.71
1:E:424[A]:ILE:HD13	3:E:977:HOH:O	1.95	0.66
1:A:53:GLU:HG3	3:F:992:HOH:O	1.95	0.65
1:B:121:ASP:HB3	3:B:959:HOH:O	2.01	0.60
1:C:231[A]:GLU:HG3	3:C:974:HOH:O	2.03	0.58

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	53/54 (98%)	53 (100%)	0	0	100	100
1	B	53/54 (98%)	53 (100%)	0	0	100	100
1	C	54/54 (100%)	54 (100%)	0	0	100	100
1	D	52/54 (96%)	51 (98%)	1 (2%)	0	100	100
1	E	56/54 (104%)	55 (98%)	1 (2%)	0	100	100
1	F	55/54 (102%)	55 (100%)	0	0	100	100
1	G	53/54 (98%)	53 (100%)	0	0	100	100
1	H	54/54 (100%)	54 (100%)	0	0	100	100
All	All	430/432 (100%)	428 (100%)	2 (0%)	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	48/47 (102%)	46 (96%)	2 (4%)	30	4
1	B	48/47 (102%)	47 (98%)	1 (2%)	53	17
1	C	49/47 (104%)	48 (98%)	1 (2%)	55	19
1	D	47/47 (100%)	47 (100%)	0	100	100
1	E	51/47 (108%)	49 (96%)	2 (4%)	32	5
1	F	50/47 (106%)	48 (96%)	2 (4%)	31	5

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	G	48/47 (102%)	48 (100%)	0	100	100
1	H	49/47 (104%)	49 (100%)	0	100	100
All	All	390/376 (104%)	382 (98%)	8 (2%)	69	17

5 of 8 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	F	544[B]	VAL
1	F	544[A]	VAL
1	E	444[A]	VAL
1	C	221[A]	ASP
1	E	444[B]	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	H	722	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)

Of 8 ligands modelled in this entry, 8 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	53/54 (98%)	-0.41	0	100	100	8, 12, 29, 39	0
1	B	54/54 (100%)	-0.36	1 (1%)	66	59	8, 14, 35, 54	1 (1%)
1	C	54/54 (100%)	-0.34	2 (3%)	41	35	8, 14, 30, 41	0
1	D	54/54 (100%)	-0.11	2 (3%)	41	35	9, 17, 42, 50	2 (3%)
1	E	54/54 (100%)	-0.20	3 (5%)	24	23	10, 16, 38, 48	2 (3%)
1	F	54/54 (100%)	-0.61	0	100	100	8, 12, 20, 27	0
1	G	54/54 (100%)	-0.63	0	100	100	8, 13, 26, 33	0
1	H	54/54 (100%)	-0.75	0	100	100	8, 11, 15, 23	0
All	All	431/432 (99%)	-0.43	8 (1%)	66	59	8, 13, 33, 54	5 (1%)

The worst 5 of 8 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	221[A]	ASP	4.0
1	D	354[A]	ASP	4.0
1	D	323	GLY	3.7
1	E	421[A]	ASP	3.3
1	E	420[A]	PRO	3.3

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	ZN	A	901	1/1	1.00	0.05	9,9,9,9	0
2	ZN	B	902	1/1	1.00	0.04	9,9,9,9	0
2	ZN	C	903	1/1	1.00	0.04	10,10,10,10	0
2	ZN	D	904	1/1	1.00	0.04	10,10,10,10	0
2	ZN	E	905	1/1	1.00	0.03	12,12,12,12	0
2	ZN	F	906	1/1	1.00	0.03	11,11,11,11	0
2	ZN	G	907	1/1	1.00	0.03	10,10,10,10	0
2	ZN	H	908	1/1	1.00	0.03	9,9,9,9	0

6.5 Other polymers (i)

There are no such residues in this entry.