W O R L D W I D E

wwPDB

P R O T E I N   D A T A   B A N K

wwPDB X-ray Structure Validation Summary Report (i)

Oct 30, 2014 – 10:43 PM EDT

PDB ID	:	4CCR
Title	:	Crystal structure of the thioredoxin reductase apoenzyme from Entamoeba histolytica in the absence of the NADP cofactor
Authors	:	Parsonage, D.; Kells, P.M.; Hirata, K.; Debnath, A.; Poole, L.B.; McKerrow, J.H.; Reed, S.L.; Podust, L.M.
Deposited on	:	2013-10-25
Resolution	:	2.28 Å(reported)

The following versions of software and data (see references) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.16 November 2013
Xtriage (Phenix)	:	dev-1439
EDS	:	stable24103
Percentile statistics	:	21963
Refmac	:	5.8.0049
CCP4	:	6.1.3
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	stable24103

1 Overall quality at a glance ⓘ

The reported resolution of this entry is 2.28 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	Percentile Ranks	Value
Rfree		0.268
Clashscore		4
Ramachandran outliers		0.2%
Sidechain outliers		4.0%
RSRZ outliers		1.0%

Worse   *Better*

▮ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution
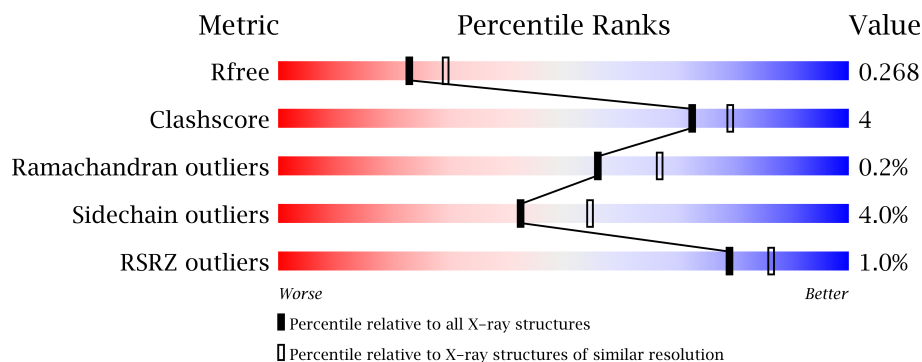
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	66092	3861 (2.30-2.26)
Clashscore	79885	4801 (2.30-2.26)
Ramachandran outliers	78287	4729 (2.30-2.26)
Sidechain outliers	78261	4728 (2.30-2.26)
RSRZ outliers	66119	3864 (2.30-2.26)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	314	
1	B	314	
1	C	314	
1	D	314	

2 Entry composition ⓘ

There are 4 unique types of molecules in this entry. The entry contains 8845 atoms, of which 0 are hydrogen and 0 are deuterium.

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 THIOREDOXIN REDUCTASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	311	Total 2331	C 1476	N 398	O 441	S 16	0	0	0
1	B	312	Total 2328	C 1475	N 397	O 440	S 16	0	0	0
1	C	313	Total 2338	C 1480	N 400	O 442	S 16	0	0	0
1	D	191	Total 1395	C 877	N 234	O 274	S 10	0	0	0

- Molecule 2 is GOLD ION (three-letter code: AU) (formula: Au).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	B	1	Total 1	Au 1	0	0
2	A	1	Total 1	Au 1	0	0
2	D	1	Total 1	Au 1	0	0
2	C	1	Total 1	Au 1	0	0

- Molecule 3 is FLAVIN-ADENINE DINUCLEOTIDE (three-letter code: FAD) (formula: $C_{27}H_{33}N_9O_{15}P_2$).

FAD
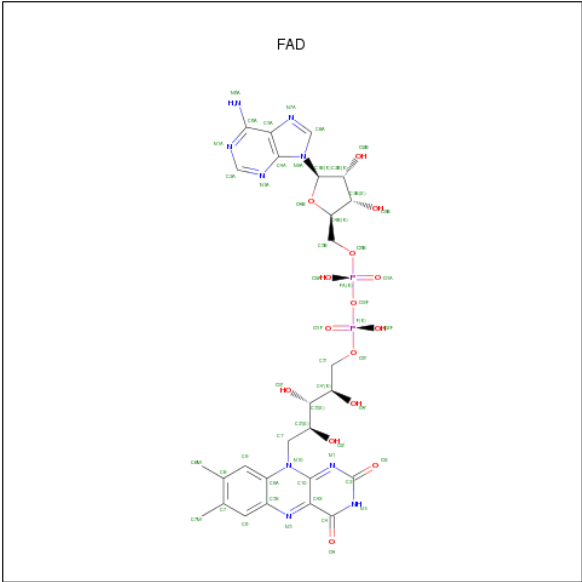
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total	C	N	O	P	0	0
			53	27	9	15	2		
3	B	1	Total	C	N	O	P	0	0
			53	27	9	15	2		
3	C	1	Total	C	N	O	P	0	0
			53	27	9	15	2		
3	D	1	Total	C	N	O	P	0	0
			53	27	9	15	2		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	89	Total	O	0	0
			89	89		
4	B	80	Total	O	0	0
			80	80		
4	C	46	Total	O	0	0
			46	46		
4	D	22	Total	O	0	0
			22	22		

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 errors 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.
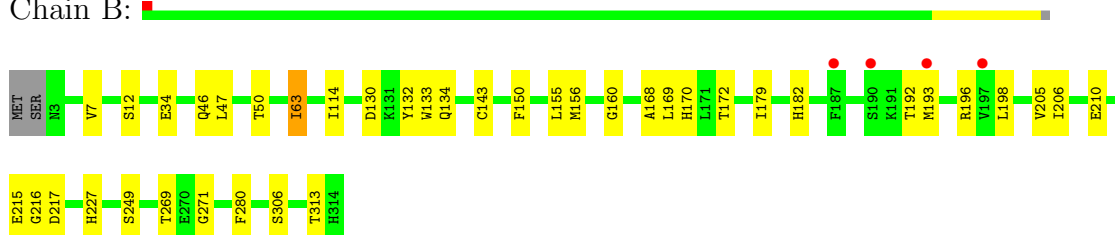
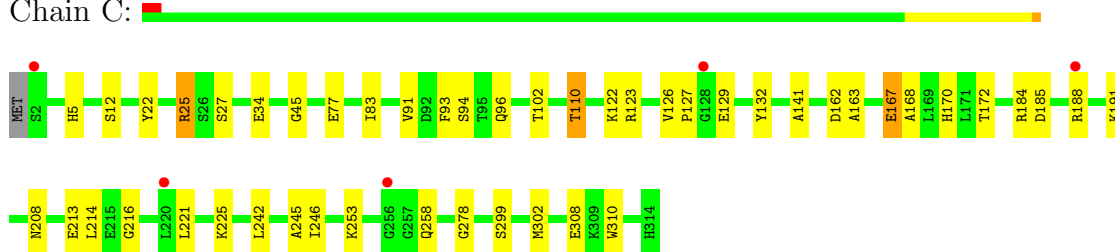- Molecule 1: THIOREDOXIN REDUCTASE

Chain A:
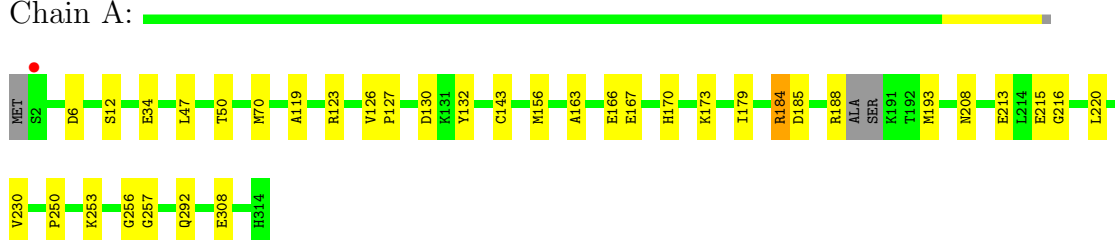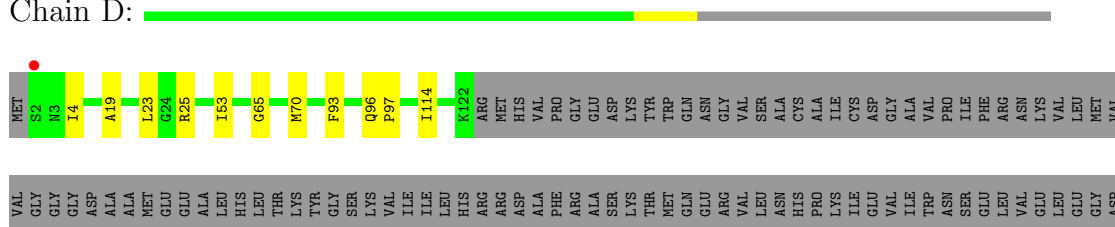

- Molecule 1: THIOREDOXIN REDUCTASE

Chain B:

- Molecule 1: THIOREDOXIN REDUCTASE

Chain C:

- Molecule 1: THIOREDOXIN REDUCTASE

Chain D:

GLY ASP LEU LEU ASN GLY ALA LYS ILE HIS ASN LEU VAL SER GLY GLU TYR LYS VAL VAL PRO VAL ALA GLY LEU PHE TYR A245 K253 Q258 A262 D263 D264 T269 E270 G271 C286 D287 R288 A304 E308 T313 H314

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	83.78Å 91.03Å 90.15Å 90.00° 105.80° 90.00°	Depositor
Resolution (Å)	91.03 – 2.28 91.03 – 2.28	Depositor EDS
% Data completeness (in resolution range)	99.6 (91.03-2.28) 99.6 (91.03-2.28)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.83 (at 2.27Å)	Xtriage
Refinement program	REFMAC 5.7.0032	Depositor
R, R_{free}	0.203 , 0.265 0.205 , 0.268	Depositor DCC
R_{free} test set	2992 reflections (5.31%)	DCC
Wilson B-factor (Å^2)	40.8	Xtriage
Anisotropy	0.364	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 23.4	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$< |L| > = 0.50, < L^2 > = 0.33$	Xtriage
Outliers	0 of 59326 reflections	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	8845	wwPDB-VP
Average B, all atoms (Å^2)	44.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.11% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.

wwPDB X-ray Structure Validation Summary Report

5 Model quality ⓘ

5.1 Standard geometry ⓘ

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

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	0.69	0/2376	0.81	1/3217 (0.0%)
1	B	0.67	0/2374	0.76	0/3218
1	C	0.65	0/2384	0.76	2/3231 (0.1%)
1	D	0.63	0/1420	0.72	0/1925
All	All	0.67	0/8554	0.77	3/11591 (0.0%)

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	C	25	ARG	NE-CZ-NH1	6.74	123.67	120.30
1	C	25	ARG	NE-CZ-NH2	-5.50	117.55	120.30
1	A	6	ASP	CB-CG-OD1	5.05	122.84	118.30

There are no chirality outliers.

There are no planarity outliers.

5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2331	0	2299	20	0
1	B	2328	0	2293	19	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	C	2338	0	2303	28	0
1	D	1395	0	1354	14	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
3	A	53	0	31	0	0
3	B	53	0	31	1	0
3	C	53	0	31	1	0
3	D	53	0	31	0	0
4	A	89	0	0	4	0
4	B	80	0	0	0	0
4	C	46	0	0	1	0
4	D	22	0	0	1	0
All	All	8845	0	8373	73	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 4.

The worst 5 of 73 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:256:GLY:HA3	4:A:2072:HOH:O	1.72	0.89
1:C:185:ASP:HB3	1:C:208:ASN:HD21	1.44	0.83
1:A:170:HIS:HD2	1:C:308:GLU:OE2	1.63	0.82
1:C:245:ALA:O	4:C:2023:HOH:O	2.01	0.78
1:A:143:CYS:O	1:C:25:ARG:NH2	2.15	0.78

There are no symmetry-related clashes.

5.3 Torsion angles

5.3.1 Protein backbone ⓘ

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	307/314 (98%)	295 (96%)	12 (4%)	0	100	100
1	B	310/314 (99%)	296 (96%)	14 (4%)	0	100	100
1	C	311/314 (99%)	293 (94%)	17 (6%)	1 (0%)	50	59
1	D	187/314 (60%)	178 (95%)	8 (4%)	1 (0%)	38	43
All	All	1115/1256 (89%)	1062 (95%)	51 (5%)	2 (0%)	56	67

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	122	LYS
1	D	262	ALA

5.3.2　Protein sidechains　ⓘ

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	244/251 (97%)	235 (96%)	9 (4%)	45	58
1	B	242/251 (96%)	231 (96%)	11 (4%)	38	48
1	C	243/251 (97%)	231 (95%)	12 (5%)	35	44
1	D	145/251 (58%)	142 (98%)	3 (2%)	66	81
All	All	874/1004 (87%)	839 (96%)	35 (4%)	42	54

5 of 35 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	215	GLU
1	B	313	THR
1	D	70	MET
1	B	217	ASP
1	B	249	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 27 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	251	ASN
1	C	61	ASN
1	D	251	ASN
1	B	258	GLN
1	A	251	ASN

5.3.3 RNA ⓘ

There are no RNA chains 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 ⓘ

Of 8 ligands modelled in this entry, 4 are monoatomic - leaving 4 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	#$|Z| > 2$	Counts	RMSZ	#$|Z| > 2$
3	FAD	A	1316	-	58,58,58	1.35	8 (13%)	85,89,89	1.75	13 (15%)
3	FAD	B	1316	-	58,58,58	1.40	9 (15%)	85,89,89	1.96	15 (17%)
3	FAD	C	1316	-	58,58,58	1.31	8 (13%)	85,89,89	2.00	14 (16%)
3	FAD	D	1316	-	58,58,58	1.32	8 (13%)	85,89,89	2.17	14 (16%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	FAD	A	1316	-	-	0/34/50/50	0/6/6/6
3	FAD	B	1316	-	-	0/34/50/50	0/6/6/6
3	FAD	C	1316	-	-	0/34/50/50	0/6/6/6
3	FAD	D	1316	-	-	0/34/50/50	0/6/6/6

The worst 5 of 33 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	1316	FAD	C9A-C5X	4.15	1.51	1.42
3	B	1316	FAD	C4A-N9A	-3.84	1.32	1.37
3	A	1316	FAD	C4A-N9A	-3.79	1.32	1.37
3	D	1316	FAD	C9A-C5X	3.78	1.50	1.42
3	A	1316	FAD	C4X-C10	3.65	1.48	1.41

The worst 5 of 56 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	D	1316	FAD	N3A-C2A-N1A	-8.61	121.32	128.89
3	D	1316	FAD	C5A-C4A-N3A	-8.29	117.89	125.98
3	B	1316	FAD	N3A-C2A-N1A	-7.84	121.99	128.89
3	C	1316	FAD	N3A-C2A-N1A	-7.49	122.31	128.89
3	C	1316	FAD	C5A-C4A-N3A	-7.39	118.78	125.98

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

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, 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	311/314 (99%)	-0.21	1 (0%)	91	96	24, 37, 60, 93	0
1	B	312/314 (99%)	-0.15	4 (1%)	74	81	25, 40, 73, 85	0
1	C	313/314 (99%)	-0.13	5 (1%)	68	77	25, 44, 74, 92	0
1	D	191/314 (60%)	-0.22	1 (0%)	88	94	28, 44, 62, 83	0
All	All	1127/1256 (89%)	-0.17	11 (0%)	79	86	24, 42, 71, 93	0

The worst 5 of 11 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	2	SER	3.7
1	B	197	VAL	3.5
1	B	193	MET	2.9
1	B	190	SER	2.7
1	C	2	SER	2.6

6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

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

6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.

6.4 Ligands ⓘ

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	RSR	LLDF	B-factors(Å^2)	Q<0.9
3	FAD	A	1316	53/53	0.12	0.35	27,31,41,49	0
3	FAD	B	1316	53/53	0.11	0.06	26,32,49,56	0
3	FAD	C	1316	53/53	0.12	0.05	31,39,44,45	0
3	FAD	D	1316	53/53	0.11	-0.13	33,40,46,49	0
2	AU	C	1315	1/1	0.06	-2.28	71,71,71,71	1
2	AU	D	1315	1/1	0.05	-3.35	106,106,106,106	1
2	AU	A	1315	1/1	0.04	-3.96	97,97,97,97	1
2	AU	B	1315	1/1	0.04	-4.30	80,80,80,80	1

6.5 Other polymers ⓘ

There are no such residues in this entry.