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 14, 2023 – 04:05 PM EDT

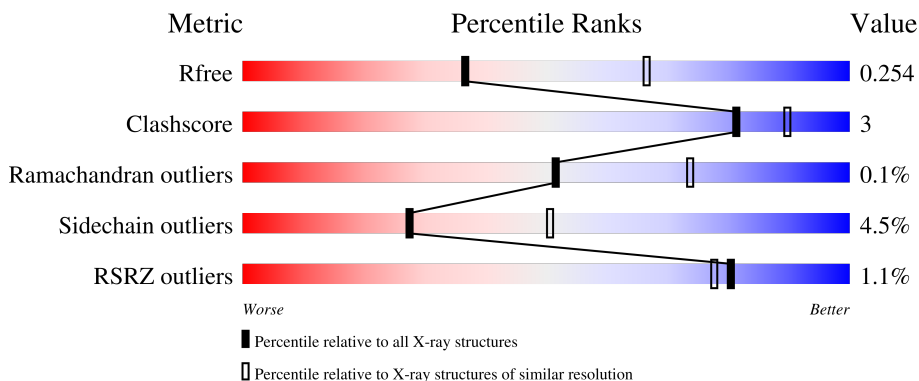|		
PDB ID	:	7KF4
Title	:	Crystal structure from SARS-CoV-2 NendoU NSP15
Authors	:	Godoy, A.S.; Nakamura, A.M.; Pereira, H.M.; Noske, G.D.; Gawriljuk, V.O.; Fernandes, R.S.; Oliveira, K.I.Z.; Oliva, G.
Deposited on	:	2020-10-13
Resolution	:	2.61 Å(reported)

1   Overall quality at a glance ⓘ

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

The reported resolution of this entry is 2.61 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.
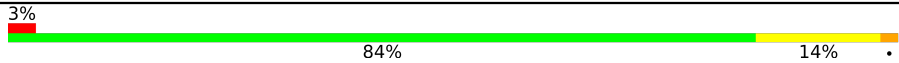
Metric	Percentile Ranks	Value
Rfree		0.254
Clashscore		3
Ramachandran outliers		0.1%
Sidechain outliers		4.5%
RSRZ outliers		1.1%

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$_{free}$	130704	3797 (2.64-2.60)
Clashscore	141614	4168 (2.64-2.60)
Ramachandran outliers	138981	4093 (2.64-2.60)
Sidechain outliers	138945	4093 (2.64-2.60)
RSRZ outliers	127900	3731 (2.64-2.60)

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	349	92% 7% •
1	B	349	2% 93% 6% •
1	C	349	% 92% 7%
1	D	349	% 90% 9% •
1	E	349	% 92% 7% •

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	349	3% 84% 14%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 16984 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 Uridylate-specific endoribonuclease.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	348	Total	C	N	O	S	0	1	0
			2751	1768	447	526	10			
1	B	348	Total	C	N	O	S	0	1	0
			2754	1770	447	526	11			
1	C	348	Total	C	N	O	S	0	1	0
			2742	1763	446	522	11			
1	D	348	Total	C	N	O	S	0	1	0
			2738	1759	444	524	11			
1	E	348	Total	C	N	O	S	0	1	0
			2750	1767	446	526	11			
1	F	348	Total	C	N	O	S	0	1	0
			2742	1763	446	522	11			

There are 18 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	GLY	-	expression tag	UNP P0DTD1
A	-1	ALA	-	expression tag	UNP P0DTD1
A	0	MET	-	expression tag	UNP P0DTD1
B	-2	GLY	-	expression tag	UNP P0DTD1
B	-1	ALA	-	expression tag	UNP P0DTD1
B	0	MET	-	expression tag	UNP P0DTD1
C	-2	GLY	-	expression tag	UNP P0DTD1
C	-1	ALA	-	expression tag	UNP P0DTD1
C	0	MET	-	expression tag	UNP P0DTD1
D	-2	GLY	-	expression tag	UNP P0DTD1
D	-1	ALA	-	expression tag	UNP P0DTD1
D	0	MET	-	expression tag	UNP P0DTD1
E	-2	GLY	-	expression tag	UNP P0DTD1
E	-1	ALA	-	expression tag	UNP P0DTD1
E	0	MET	-	expression tag	UNP P0DTD1
F	-2	GLY	-	expression tag	UNP P0DTD1
F	-1	ALA	-	expression tag	UNP P0DTD1

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
F	0	MET	-	expression tag	UNP P0DTD1

- Molecule 2 is CITRIC ACID (three-letter code: CIT) (formula: $C_6H_8O_7$).
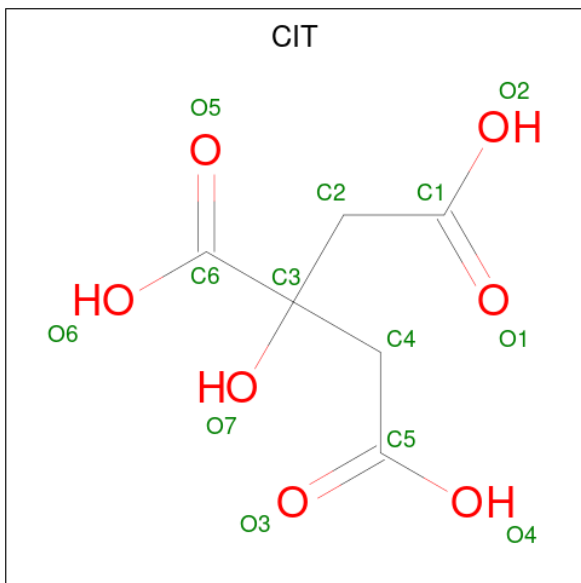
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total 13	C 6	O 7	0	0
2	B	1	Total 13	C 6	O 7	0	0
2	C	1	Total 13	C 6	O 7	0	0
2	D	1	Total 13	C 6	O 7	0	0
2	E	1	Total 13	C 6	O 7	0	0
2	F	1	Total 13	C 6	O 7	0	0

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	90	Total 90	O 90	0	0
3	B	77	Total 77	O 77	0	0
3	C	59	Total 59	O 59	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	D	71	Total 71	O 71	0	0
3	E	61	Total 61	O 61	0	0
3	F	71	Total 71	O 71	0	0

3 Residue-property plots ⓘ

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: Uridylate-specific endoribonuclease
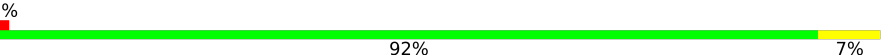
Chain A:

- Molecule 1: Uridylate-specific endoribonuclease

Chain B:

- Molecule 1: Uridylate-specific endoribonuclease

Chain C:

- Molecule 1: Uridylate-specific endoribonuclease

Chain D:

- Molecule 1: Uridylate-specific endoribonuclease

Chain E:

GLY A-1 M0 S1 V10 V22 I71 I79 V84 T98 V101 V121 F122 F123 V127 Q130 R135 S147 V148 L151 L200 I222 Y230 L245 E264 L265 E266 F279 D300 D301 F302 V314 S315 K316 D323 L331

V338 Q346

• Molecule 1: Uridylate-specific endoribonuclease

Chain F:

3%

84%　　　　　　　　14%

GLY A-1 M0 S1 V31 Y32 T33 K34 L42 F43 P50 V53 I71 L75 G76 V77 K89 M104 K110 A117 V127 D128 G129 Q130 R135 R138 V141 V148 L151 N163 K173 L200 Q201 E202 R206 E215 D219

K226 L227 L248 L254 L265 E266 S273 Y278 F279 I280 S293 L298 V303 K307 S308 Q309 D310 L311 S312 V313 V318 K319 D323 Y324 T325 E326 M330 L331 K334 D335 G336 H337 V338 E339 T340 F341 L345 Q346

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 2 21 21	Depositor
Cell constants a, b, c, α, β, γ	85.12Å 151.16Å 199.24Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.16 – 2.61 49.16 – 2.61	Depositor EDS
% Data completeness (in resolution range)	99.8 (49.16-2.61) 99.7 (49.16-2.61)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.22 (at 2.61Å)	Xtriage
Refinement program	BUSTER 2.10.3	Depositor
R, R_{free}	0.219 , 0.246 0.230 , 0.254	Depositor DCC
R_{free} test set	4004 reflections (5.09%)	wwPDB-VP
Wilson B-factor (Å^2)	59.3	Xtriage
Anisotropy	0.593	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 51.7	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	0.95	EDS
Total number of atoms	16984	wwPDB-VP
Average B, all atoms (Å^2)	74.0	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 45.42 % 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 1.3252e-04. 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 ⓘ

5.1   Standard geometry ⓘ

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

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	0.48	0/2811	0.60	0/3814
1	B	0.49	0/2814	0.61	0/3817
1	C	0.50	0/2802	0.68	0/3803
1	D	0.51	0/2798	0.68	0/3800
1	E	0.51	0/2810	0.68	0/3813
1	F	0.52	0/2802	0.72	0/3803
All	All	0.50	0/16837	0.66	0/22850

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	2751	0	2742	14	0
1	B	2754	0	2749	11	0
1	C	2742	0	2730	10	0
1	D	2738	0	2712	12	0
1	E	2750	0	2738	14	0
1	F	2742	0	2730	28	0
2	A	13	0	5	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	B	13	0	5	1	0
2	C	13	0	5	0	0
2	D	13	0	5	0	0
2	E	13	0	5	0	0
2	F	13	0	5	0	0
3	A	90	0	0	1	0
3	B	77	0	0	0	0
3	C	59	0	0	0	0
3	D	71	0	0	0	0
3	E	61	0	0	0	0
3	F	71	0	0	2	0
All	All	16984	0	16431	83	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 3.

The worst 5 of 83 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:322:ILE:HB	1:A:327:ILE:HD13	1.60	0.84
1:C:322:ILE:HB	1:C:327:ILE:HD13	1.60	0.83
1:D:322:ILE:HB	1:D:327:ILE:HD13	1.62	0.80
1:C:1:SER:HB2	1:D:1:SER:HB2	1.68	0.75
1:F:265:LEU:HD23	1:F:280:ILE:HG12	1.67	0.75

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	347/349 (99%)	342 (99%)	5 (1%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	B	347/349 (99%)	343 (99%)	4 (1%)	0	100	100
1	C	347/349 (99%)	334 (96%)	13 (4%)	0	100	100
1	D	347/349 (99%)	335 (96%)	11 (3%)	1 (0%)	41	62
1	E	347/349 (99%)	338 (97%)	9 (3%)	0	100	100
1	F	347/349 (99%)	338 (97%)	8 (2%)	1 (0%)	41	62
All	All	2082/2094 (99%)	2030 (98%)	50 (2%)	2 (0%)	51	74

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	F	226	LYS
1	D	334	LYS

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	309/309 (100%)	299 (97%)	10 (3%)	39	63
1	B	310/309 (100%)	297 (96%)	13 (4%)	30	53
1	C	307/309 (99%)	297 (97%)	10 (3%)	38	62
1	D	306/309 (99%)	291 (95%)	15 (5%)	25	46
1	E	309/309 (100%)	299 (97%)	10 (3%)	39	63
1	F	307/309 (99%)	283 (92%)	24 (8%)	12	24
All	All	1848/1854 (100%)	1766 (96%)	82 (4%)	27	52

5 of 82 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	338	VAL
1	F	227	LEU
1	F	33	THR
1	F	163	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	F	309	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	201	GLN
1	D	18	GLN
1	D	244	GLN
1	D	337	HIS
1	F	244	GLN

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 monosaccharides in this entry.

5.6 Ligand geometry ⓘ

6 ligands are modelled in this entry.

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	$\#|Z| > 2$
2	CIT	C	700	-	12,12,12	0.27	0	17,17,17	0.48	0
2	CIT	F	700	-	12,12,12	0.26	0	17,17,17	0.48	0
2	CIT	B	700	-	12,12,12	0.20	0	17,17,17	0.46	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	#|Z| > 2	Counts	RMSZ	#|Z| > 2
2	CIT	A	700	-	12,12,12	1.90	6 (50%)	17,17,17	2.50	6 (35%)
2	CIT	D	700	-	12,12,12	1.95	6 (50%)	17,17,17	2.44	6 (35%)
2	CIT	E	700	-	12,12,12	1.90	6 (50%)	17,17,17	2.44	6 (35%)

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
2	CIT	C	700	-	-	0/16/16/16	-
2	CIT	F	700	-	-	0/16/16/16	-
2	CIT	B	700	-	-	0/16/16/16	-
2	CIT	A	700	-	-	2/16/16/16	-
2	CIT	D	700	-	-	2/16/16/16	-
2	CIT	E	700	-	-	2/16/16/16	-

The worst 5 of 18 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	700	CIT	O5-C6	3.30	1.32	1.22
2	A	700	CIT	O5-C6	3.03	1.31	1.22
2	E	700	CIT	O4-C5	-3.03	1.20	1.30
2	A	700	CIT	O3-C5	2.91	1.31	1.22
2	E	700	CIT	O2-C1	-2.90	1.21	1.30

The worst 5 of 18 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	A	700	CIT	O5-C6-C3	-6.08	113.64	122.25
2	E	700	CIT	O5-C6-C3	-5.97	113.80	122.25
2	D	700	CIT	O5-C6-C3	-5.78	114.07	122.25
2	E	700	CIT	O6-C6-C3	5.44	122.50	113.05
2	A	700	CIT	O6-C6-C3	5.42	122.46	113.05

There are no chirality outliers.

5 of 6 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	D	700	CIT	O1-C1-C2-C3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
2	D	700	CIT	O2-C1-C2-C3
2	E	700	CIT	O2-C1-C2-C3
2	A	700	CIT	O1-C1-C2-C3
2	A	700	CIT	O2-C1-C2-C3

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	700	CIT	1	0
2	A	700	CIT	1	0

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 (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(Å²)	Q<0.9
1	A	348/349 (99%)	-0.09	0	100	100	43, 65, 93, 149	0
1	B	348/349 (99%)	-0.02	7 (2%)	65	60	43, 67, 105, 210	0
1	C	348/349 (99%)	0.04	2 (0%)	89	88	51, 74, 97, 160	0
1	D	348/349 (99%)	-0.02	4 (1%)	80	78	51, 69, 96, 170	0
1	E	348/349 (99%)	-0.10	2 (0%)	89	88	41, 68, 98, 238	0
1	F	348/349 (99%)	0.24	9 (2%)	56	50	50, 82, 133, 212	0
All	All	2088/2094 (99%)	0.01	24 (1%)	80	78	41, 71, 105, 238	0

The worst 5 of 24 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	F	346	GLN	3.7
1	F	313	VAL	3.5
1	F	330	MET	3.4
1	C	335	ASP	3.2
1	F	331	LEU	3.0

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 monosaccharides 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. 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	CIT	E	700	13/13	0.87	0.26	106,111,117,119	0
2	CIT	D	700	13/13	0.89	0.21	100,104,107,107	0
2	CIT	C	700	13/13	0.89	0.16	99,102,104,105	0
2	CIT	B	700	13/13	0.91	0.17	123,125,126,127	0
2	CIT	F	700	13/13	0.92	0.25	110,115,124,124	0
2	CIT	A	700	13/13	0.93	0.20	80,87,91,92	0

6.5 Other polymers ⓘ

There are no such residues in this entry.