wwPDB X-ray Structure Validation Summary Report (i)

Feb 21, 2024 – 06:40 PM EST

PDB ID	:	4QU4
Title	:	Improved refinement of the Mtr4 apo crystal structure
Authors	:	Johnson, S.J.; Taylor, L.L.
Deposited on	:	2014-07-10
Resolution	:	3.39 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:

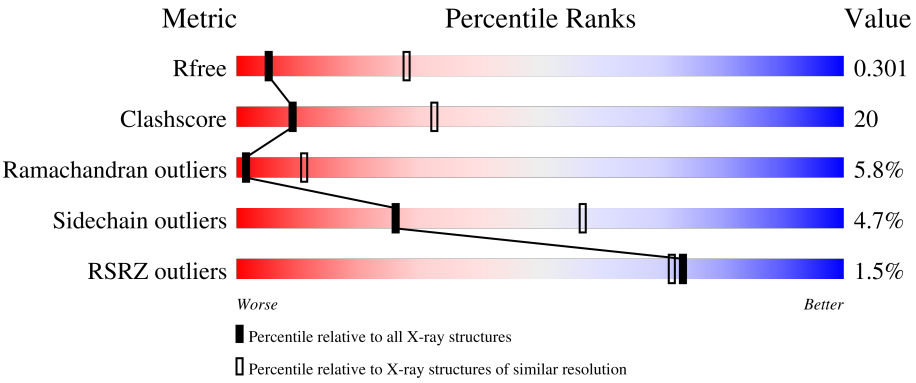
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, 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 3.39 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	1026 (3.48-3.32)
Clashscore	141614	1055 (3.48-3.32)
Ramachandran outliers	138981	1038 (3.48-3.32)
Sidechain outliers	138945	1038 (3.48-3.32)
RSRZ outliers	127900	2173 (3.50-3.30)

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	1108	% 52% 30% · 15%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 6680 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 ATP-dependent RNA helicase DOB1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	942	Total	C	N	O	S	0	0	0
			6665	4259	1104	1266	36			

There are 35 discrepancies between the modelled and reference sequences:

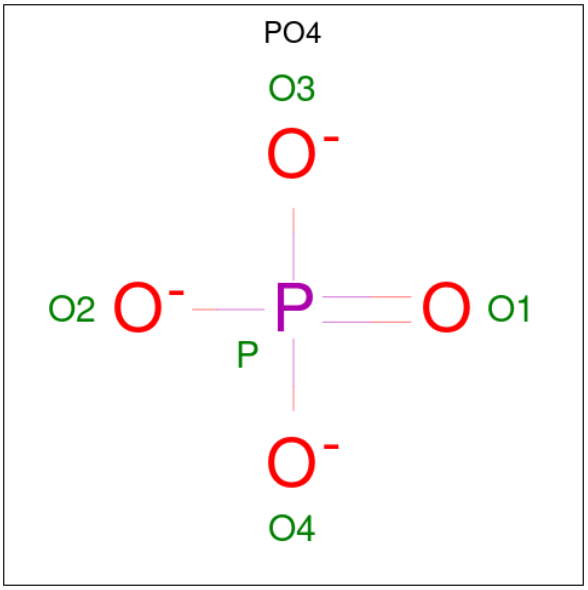
Chain	Residue	Modelled	Actual	Comment	Reference
A	-34	MET	-	expression tag	UNP P47047
A	-33	HIS	-	expression tag	UNP P47047
A	-32	HIS	-	expression tag	UNP P47047
A	-31	HIS	-	expression tag	UNP P47047
A	-30	HIS	-	expression tag	UNP P47047
A	-29	HIS	-	expression tag	UNP P47047
A	-28	HIS	-	expression tag	UNP P47047
A	-27	GLY	-	expression tag	UNP P47047
A	-26	LYS	-	expression tag	UNP P47047
A	-25	PRO	-	expression tag	UNP P47047
A	-24	ILE	-	expression tag	UNP P47047
A	-23	PRO	-	expression tag	UNP P47047
A	-22	ASN	-	expression tag	UNP P47047
A	-21	PRO	-	expression tag	UNP P47047
A	-20	LEU	-	expression tag	UNP P47047
A	-19	LEU	-	expression tag	UNP P47047
A	-18	GLY	-	expression tag	UNP P47047
A	-17	LEU	-	expression tag	UNP P47047
A	-16	ASP	-	expression tag	UNP P47047
A	-15	SER	-	expression tag	UNP P47047
A	-14	THR	-	expression tag	UNP P47047
A	-13	GLU	-	expression tag	UNP P47047
A	-12	ASN	-	expression tag	UNP P47047
A	-11	LEU	-	expression tag	UNP P47047
A	-10	TYR	-	expression tag	UNP P47047
A	-9	PHE	-	expression tag	UNP P47047
A	-8	GLN	-	expression tag	UNP P47047

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	-7	GLY	-	expression tag	UNP P47047
A	-6	ILE	-	expression tag	UNP P47047
A	-5	ASP	-	expression tag	UNP P47047
A	-4	PRO	-	expression tag	UNP P47047
A	-3	PHE	-	expression tag	UNP P47047
A	-2	THR	-	expression tag	UNP P47047
A	-1	GLU	-	expression tag	UNP P47047
A	0	PHE	-	expression tag	UNP P47047

- Molecule 2 is PHOSPHATE ION (three-letter code: PO4) (formula: O_4P).

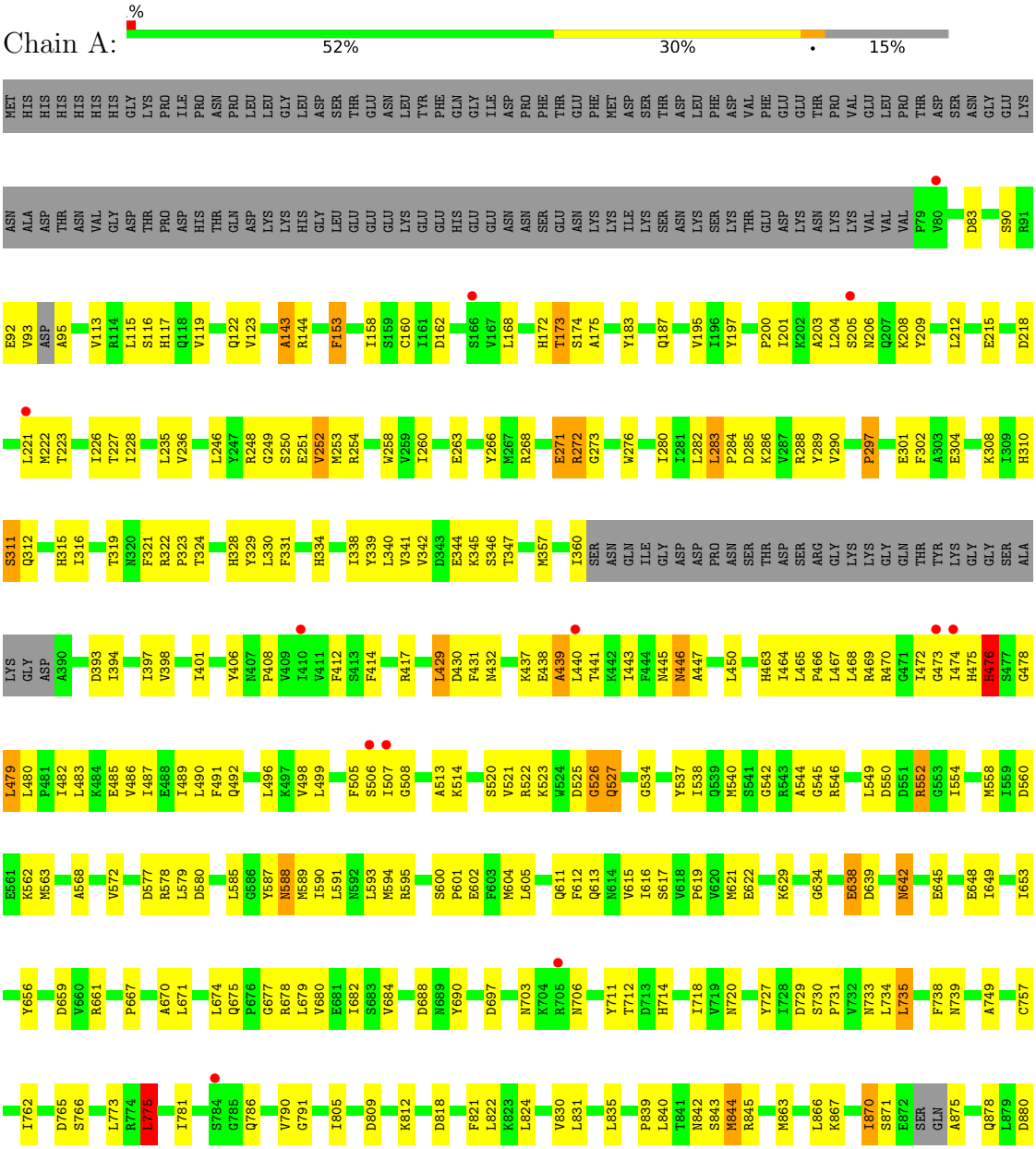


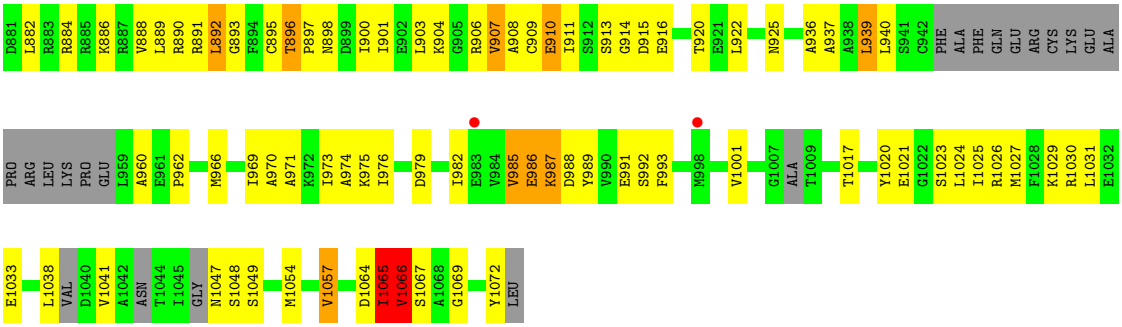
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total O P 5 4 1	0	0
2	A	1	Total O P 5 4 1	0	0
2	A	1	Total O P 5 4 1	0	0

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: ATP-dependent RNA helicase DOB1




D881 L882 R883 R884 R885 K886 R887 V888 L889 R890 R891 L892 G893 F894 C895 T896 P897 N898 D899 I900 I901 E902 L903 K904 G905 R906 V907 A908 C909 E910 I911 S912 S913 G914 D915 E916 T920 E921 L922 N925 A936 A937 A938 L939 L940 S941 C942 PHE ALA PHE GLN GLU ARG CYS LYS GLU ALA
PRO ARG LEU LYS PRO GLU L959 A960 E961 P962 M966 I969 A970 A971 K972 I973 A974 K975 I976 D979 I982 E983 V984 V985 E986 K987 D988 Y989 V990 E991 S992 F993 M998 V1001 G1007 ALA T1009 T1017 Y1020 E1021 G1022 S1023 L1024 I1025 R1026 M1027 F1028 K1029 R1030 L1031 E1032
E1033 L1038 VAL D1040 V1041 A1042 ASN T1044 I1045 GLY N1047 S1048 S1049 M1054 V1057 D1064 I1065 V1066 S1067 A1068 G1069 Y1072 LEU

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 31 2 1	Depositor
Cell constants a, b, c, α, β, γ	133.52Å 133.52Å 190.95Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	28.76 – 3.39 28.76 – 3.39	Depositor EDS
% Data completeness (in resolution range)	93.8 (28.76-3.39) 93.9 (28.76-3.39)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.81 (at 3.39Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.9_1692)	Depositor
R, R_{free}	0.247 , 0.299 0.254 , 0.301	Depositor DCC
R_{free} test set	1270 reflections (4.86%)	wwPDB-VP
Wilson B-factor (Å^2)	137.5	Xtriage
Anisotropy	0.363	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.28 , 153.3	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.022 for -h,-k,l	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	6680	wwPDB-VP
Average B, all atoms (Å^2)	166.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.54% 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

Bond lengths and bond angles in the following residue types are not validated in this section: PO4

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	0.49	0/6777	0.71	3/9247 (0.0%)

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	4

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	A	775	LEU	CA-CB-CG	7.66	132.91	115.30
1	A	735	LEU	CA-CB-CG	7.55	132.67	115.30
1	A	526	GLY	N-CA-C	5.93	127.93	113.10

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	334	HIS	Peptide
1	A	406	TYR	Peptide
1	A	476	HIS	Peptide
1	A	527	GLN	Peptide

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	6665	0	5969	252	0
2	A	15	0	0	0	0
All	All	6680	0	5969	252	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 20.

The worst 5 of 252 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:200:PRO:HG3	1:A:276:TRP:HZ2	1.43	0.83
1:A:1027:MET:HA	1:A:1030:ARG:HB3	1.59	0.83
1:A:974:ALA:HB1	1:A:985:VAL:HA	1.60	0.83
1:A:1054:MET:HA	1:A:1057:VAL:HB	1.61	0.81
1:A:986:GLU:O	1:A:988:ASP:N	2.15	0.79

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	924/1108 (83%)	771 (83%)	99 (11%)	54 (6%)	1 11

5 of 54 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	250	SER
1	A	253	MET
1	A	283	LEU
1	A	476	HIS
1	A	588	ASN

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	591/985 (60%)	563 (95%)	28 (5%)	26 57

5 of 28 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	675	GLN
1	A	1072	TYR
1	A	775	LEU
1	A	1064	ASP
1	A	735	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	187	GLN
1	A	475	HIS
1	A	838	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)

3 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	PO4	A	1103	-	4,4,4	0.74	0	6,6,6	0.49	0
2	PO4	A	1101	-	4,4,4	0.95	0	6,6,6	0.53	0
2	PO4	A	1102	-	4,4,4	0.76	0	6,6,6	0.70	0

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

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	942/1108 (85%)	-0.26	14 (1%)	73	72	106, 158, 259, 324	3 (0%)

The worst 5 of 14 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	705	ARG	4.6
1	A	440	LEU	3.9
1	A	506	SER	3.2
1	A	507	ILE	3.2
1	A	784	SER	3.0

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 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, 95^{th} 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	PO4	A	1103	5/5	0.79	0.15	190,197,203,204	0
2	PO4	A	1101	5/5	0.89	0.22	141,143,153,159	0
2	PO4	A	1102	5/5	0.93	0.09	188,192,197,205	0

6.5 Other polymers (i)

There are no such residues in this entry.