wwPDB X-ray Structure Validation Summary Report (i)

Mar 15, 2025 – 09:05 pm GMT

PDB ID	:	9EOS
Title	:	Human serum albumin with myristate
Authors	:	Freitag-Pohl, S.; Pohl, E.
Deposited on	:	2024-03-15
Resolution	:	2.10 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 (i)) were used in the production of this report:

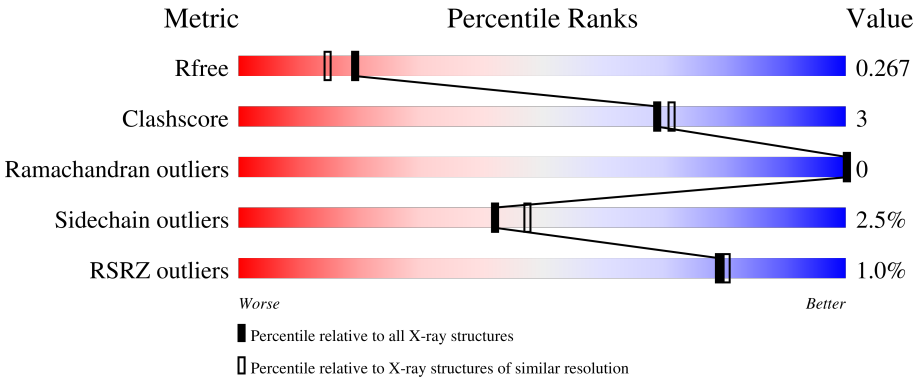
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.003 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.41

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 2.10 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}	164625	6234 (2.10-2.10)
Clashscore	180529	6893 (2.10-2.10)
Ramachandran outliers	177936	6839 (2.10-2.10)
Sidechain outliers	177891	6840 (2.10-2.10)
RSRZ outliers	164620	6234 (2.10-2.10)

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	610	% 88% 8% 5%
1	B	610	% 86% 10% .

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 17921 atoms, of which 8762 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.

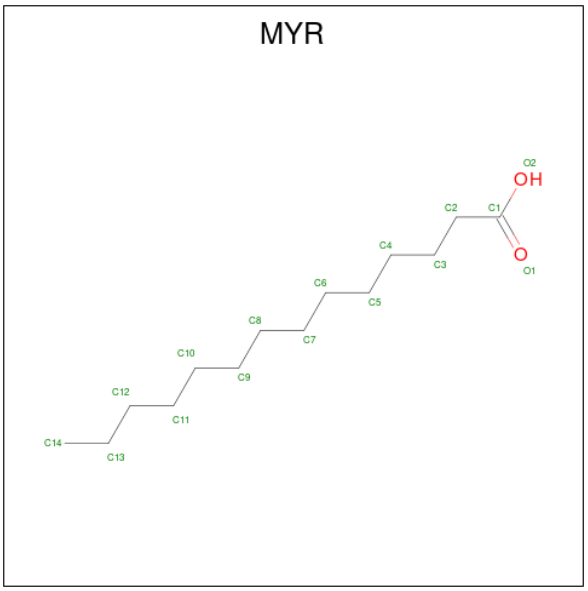
- Molecule 1 is a protein called Albumin.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	582	Total	C	H	N	O	S	153	2	0
			8736	2802	4314	753	826	41			
1	B	584	Total	C	H	N	O	S	158	0	0
			8658	2779	4254	746	838	41			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	586	LEU	-	expression tag	UNP P02768
B	586	LEU	-	expression tag	UNP P02768

- Molecule 2 is MYRISTIC ACID (three-letter code: MYR) (formula: $C_{14}H_{28}O_2$).



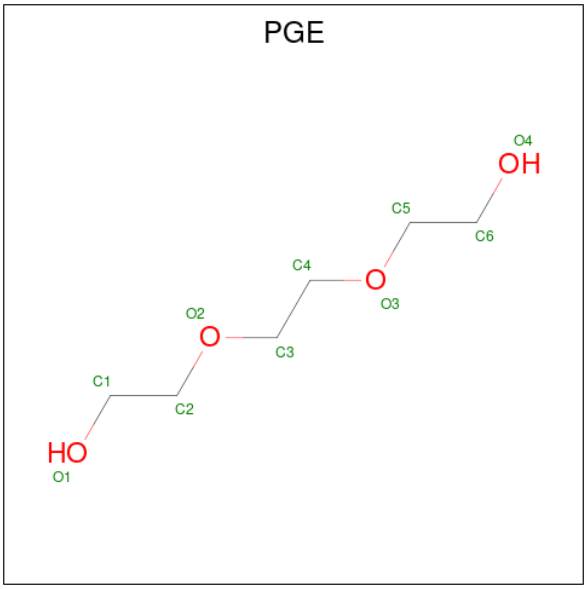
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	H	O	0	0
			43	14	27	2		

Continued on next page...

Continued from previous page...

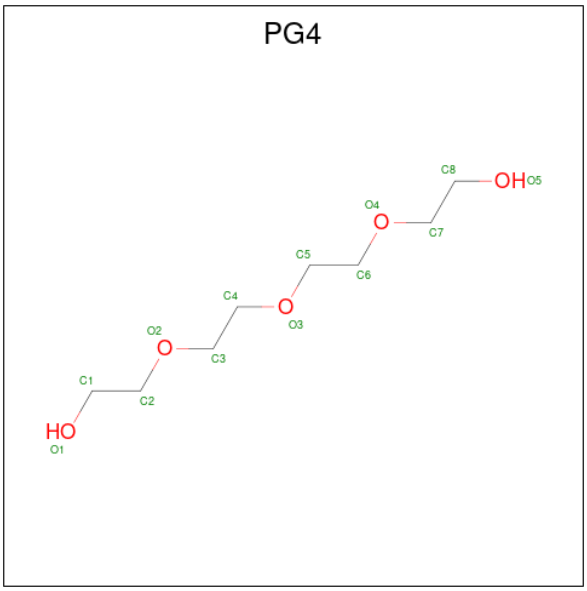
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total C H O 43 14 27 2	0	0
2	A	1	Total C H O 43 14 27 2	0	0
2	B	1	Total C H O 43 14 27 2	0	0
2	B	1	Total C H O 43 14 27 2	0	0
2	B	1	Total C H O 43 14 27 2	0	0

- Molecule 3 is TRIETHYLENE GLYCOL (three-letter code: PGE) (formula: $C_6H_{14}O_4$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C H O 24 6 14 4	1	0

- Molecule 4 is TETRAETHYLENE GLYCOL (three-letter code: PG4) (formula: $C_8H_{18}O_5$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	B	1	Total C H O 31 8 18 5	1	0

- Molecule 5 is water.

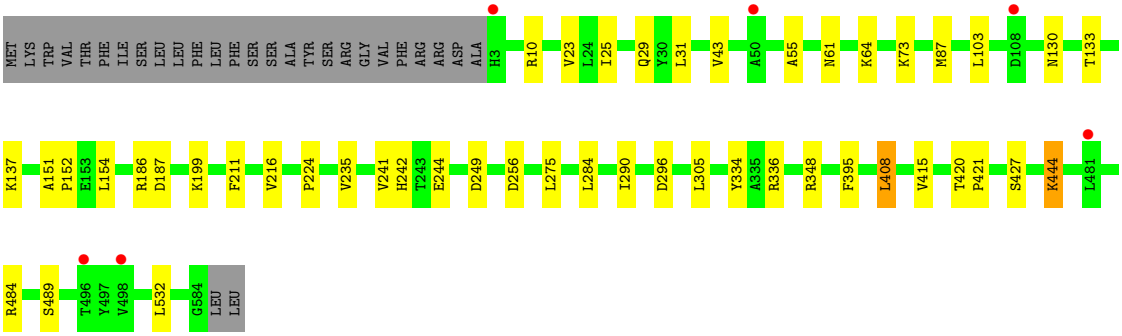
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	94	Total O 94 94	0	0
5	B	120	Total O 120 120	0	0

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.

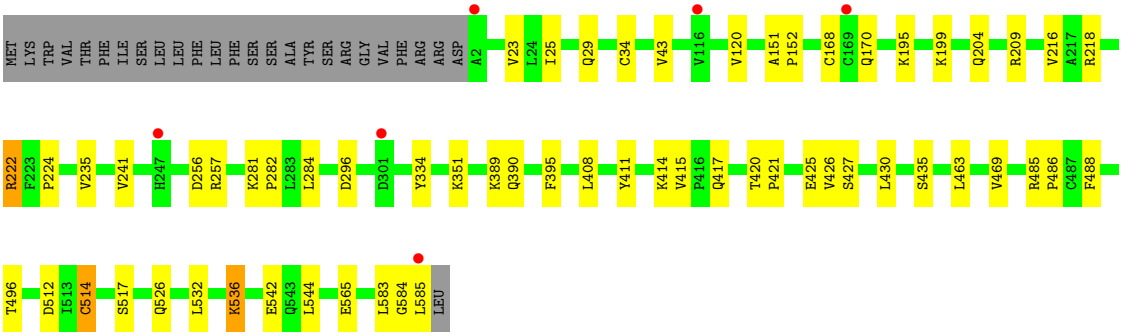
• Molecule 1: Albumin

Chain A:



• Molecule 1: Albumin

Chain B:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	38.32Å 87.11Å 97.47Å 74.44° 89.78° 78.57°	Depositor
Resolution (Å)	72.34 – 2.10 72.34 – 2.10	Depositor EDS
% Data completeness (in resolution range)	97.7 (72.34-2.10) 97.7 (72.34-2.10)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.55 (at 2.10Å)	Xtriage
Refinement program	REFMAC 5.8.0403	Depositor
R, R_{free}	0.211 , 0.269 0.211 , 0.267	Depositor DCC
R_{free} test set	3335 reflections (4.81%)	wwPDB-VP
Wilson B-factor (Å^2)	39.4	Xtriage
Anisotropy	0.109	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 32.3	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.108 for h,h-k,-l	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	17921	wwPDB-VP
Average B, all atoms (Å^2)	50.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.38% 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: PG4, MYR, PGE

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	0.40	0/4516	0.67	2/6127 (0.0%)
1	B	0.39	0/4490	0.67	0/6100
All	All	0.39	0/9006	0.67	2/12227 (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	2
1	B	0	4
All	All	0	6

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	336	ARG	NE-CZ-NH2	-5.74	117.43	120.30
1	A	484	ARG	NE-CZ-NH2	-5.23	117.69	120.30

There are no chirality outliers.

5 of 6 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	10	ARG	Sidechain
1	A	348	ARG	Sidechain
1	B	209	ARG	Sidechain
1	B	218	ARG	Sidechain

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Group
1	B	222	ARG	Sidechain

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	4422	4314	4190	27	0
1	B	4404	4254	4120	33	0
2	A	48	81	81	2	0
2	B	48	81	81	4	0
3	A	10	14	14	0	0
4	B	13	18	18	3	0
5	A	94	0	0	2	0
5	B	120	0	0	3	0
All	All	9159	8762	8504	60	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 60 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:444:LYS:HG3	5:A:713:HOH:O	1.71	0.89
1:A:199:LYS:HB3	1:A:242[B]:HIS:CE1	2.22	0.74
1:B:34:CYS:SG	5:B:746:HOH:O	2.48	0.70
1:A:186:ARG:NH1	1:A:187:ASP:OD1	2.25	0.69
1:B:120:VAL:HG23	5:B:737:HOH:O	1.92	0.69

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	582/610 (95%)	573 (98%)	9 (2%)	0	100	100
1	B	582/610 (95%)	570 (98%)	12 (2%)	0	100	100
All	All	1164/1220 (95%)	1143 (98%)	21 (2%)	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	454/534 (85%)	446 (98%)	8 (2%)	54	61
1	B	450/534 (84%)	435 (97%)	15 (3%)	33	36
All	All	904/1068 (85%)	881 (98%)	23 (2%)	42	47

5 of 23 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	425	GLU
1	B	514	CYS
1	B	496	THR
1	B	517	SER
1	A	408	LEU

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

Mol	Chain	Res	Type
1	A	130	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 oligosaccharides in this entry.

5.6 Ligand geometry (i)

8 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	MYR	B	603	-	15,15,15	0.57	0	15,15,15	0.68	0
3	PGE	A	604	-	9,9,9	0.39	0	8,8,8	0.19	0
4	PG4	B	604	-	12,12,12	0.23	0	11,11,11	0.21	0
2	MYR	A	602	-	15,15,15	0.64	0	15,15,15	0.48	0
2	MYR	B	602	-	15,15,15	0.55	0	15,15,15	0.64	0
2	MYR	B	601	-	15,15,15	0.57	0	15,15,15	0.64	0
2	MYR	A	603	-	15,15,15	0.47	0	15,15,15	0.67	0
2	MYR	A	601	-	15,15,15	0.62	0	15,15,15	0.52	0

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	MYR	B	603	-	-	10/13/13/13	-
3	PGE	A	604	-	-	3/7/7/7	-

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	PG4	B	604	-	-	5/10/10/10	-
2	MYR	A	602	-	-	9/13/13/13	-
2	MYR	B	602	-	-	4/13/13/13	-
2	MYR	B	601	-	-	8/13/13/13	-
2	MYR	A	603	-	-	3/13/13/13	-
2	MYR	A	601	-	-	5/13/13/13	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 47 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	604	PGE	O3-C5-C6-O4
4	B	604	PG4	O3-C5-C6-O4
2	A	603	MYR	C1-C2-C3-C4
4	B	604	PG4	O2-C3-C4-O3
2	A	602	MYR	C9-C10-C11-C12

There are no ring outliers.

5 monomers are involved in 9 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	603	MYR	3	0
4	B	604	PG4	3	0
2	A	602	MYR	1	0
2	B	602	MYR	1	0
2	A	603	MYR	1	0

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	582/610 (95%)	0.19	6 (1%) 79 80	21, 47, 65, 86	2 (0%)
1	B	584/610 (95%)	0.23	6 (1%) 79 80	31, 48, 67, 97	0
All	All	1166/1220 (95%)	0.21	12 (1%) 79 80	21, 48, 67, 97	2 (0%)

The worst 5 of 12 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	2	ALA	3.2
1	A	496	THR	3.0
1	A	3	HIS	2.7
1	B	301	ASP	2.7
1	B	116	VAL	2.7

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	MYR	A	603	16/16	0.87	0.14	40,53,71,74	0
2	MYR	A	602	16/16	0.88	0.14	47,54,65,74	0
2	MYR	B	603	16/16	0.88	0.17	50,66,79,81	0
4	PG4	B	604	13/13	0.90	0.14	50,64,68,71	1
2	MYR	A	601	16/16	0.92	0.13	38,57,69,70	0
2	MYR	B	602	16/16	0.92	0.12	43,49,66,74	0
3	PGE	A	604	10/10	0.93	0.12	51,56,61,65	1
2	MYR	B	601	16/16	0.93	0.13	38,59,72,74	0

6.5 Other polymers (i)

There are no such residues in this entry.