



# wwPDB X-ray Structure Validation Summary Report ⓘ

Sep 23, 2023 – 08:46 PM EDT

PDB ID : 5TDS  
Title : Toluene bound in the resting active site of toluene 4-monoxygenase (T4moH)  
Authors : Acheson, J.F.; Fox, B.G.  
Deposited on : 2016-09-19  
Resolution : 1.72 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](mailto: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 ⓘ 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 ⓘ](#)) 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 : 2.35.1  
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

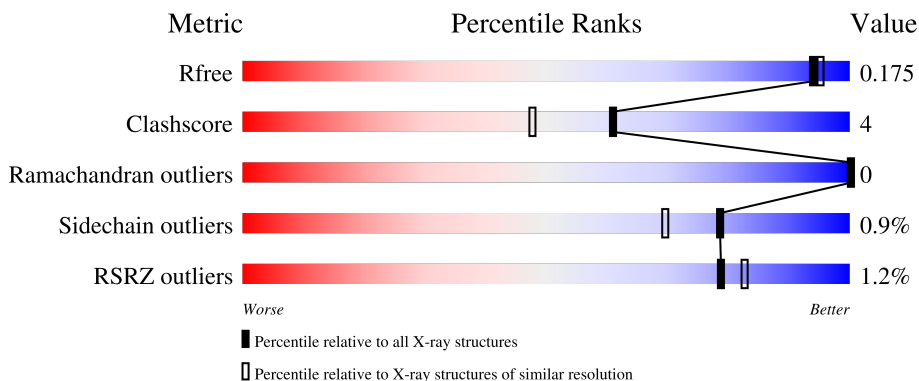
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.72 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.



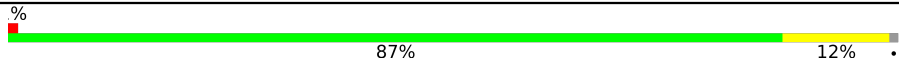
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	5722 (1.74-1.70)
Clashscore	141614	6152 (1.74-1.70)
Ramachandran outliers	138981	6051 (1.74-1.70)
Sidechain outliers	138945	6051 (1.74-1.70)
RSRZ outliers	127900	5629 (1.74-1.70)

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  $\geq 3$ , 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  $\leq 5\%$ . 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	493	87% 12% .
1	D	493	91% 8% .
2	B	327	87% 5% 8%
2	E	327	87% 6% 7%
3	C	84	85% 11% . .

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
3	F	84	 A horizontal bar chart representing the quality of chain. The bar is primarily green, indicating a high quality score of 87%. A small yellow segment at the end indicates a lower quality score of 12%. The bar is labeled with a '%' symbol at the start and ends with a dot.

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
6	MBN	A	506	-	-	X	-
6	MBN	D	507	-	-	X	-

## 2 Entry composition [i](#)

There are 8 unique types of molecules in this entry. The entry contains 16655 atoms, of which 72 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 Toluene-4-monooxygenase system protein A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	490	4064	2607	686	748	23	0	5	0
1	D	490	4052	2599	684	745	24	0	3	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	336	TRP	LEU	conflict	UNP Q00456
A	337	TYR	ASP	conflict	UNP Q00456
D	336	TRP	LEU	conflict	UNP Q00456
D	337	TYR	ASP	conflict	UNP Q00456

- Molecule 2 is a protein called Toluene-4-monooxygenase system protein E.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	301	2519	1595	434	475	15	0	4	0
2	E	304	2542	1609	442	476	15	0	3	0

- Molecule 3 is a protein called Toluene-4-monooxygenase system protein B.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	C	82	662	417	118	122	5	0	1	0
3	F	83	674	423	122	125	4	0	1	0

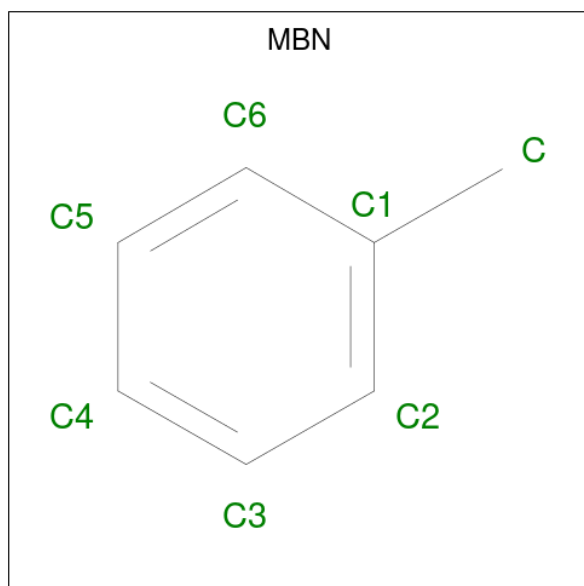
- Molecule 4 is FE (III) ION (three-letter code: FE) (formula: Fe).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	2	Total Fe 2 2	0	0
4	D	2	Total Fe 2 2	0	0

- Molecule 5 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total Mg 1 1	0	0
5	B	4	Total Mg 4 4	0	0
5	C	1	Total Mg 1 1	0	0
5	D	1	Total Mg 1 1	0	0
5	E	2	Total Mg 2 2	0	0

- Molecule 6 is TOLUENE (three-letter code: MBN) (formula: C<sub>7</sub>H<sub>8</sub>).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	1	Total C H 15 7 8	0	0
6	A	1	Total C H 15 7 8	0	0
6	A	1	Total C H 15 7 8	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	1	Total C H 15 7 8	0	0
6	B	1	Total C H 15 7 8	0	0
6	D	1	Total C H 15 7 8	0	0
6	D	1	Total C H 15 7 8	0	0
6	D	1	Total C H 15 7 8	0	0
6	D	1	Total C H 15 7 8	0	0

- Molecule 7 is CHLORIDE ION (three-letter code: CL) (formula: Cl).


Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	D	1	Total Cl 1 1	0	0

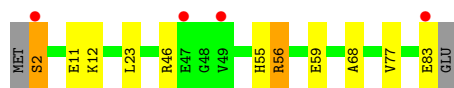
- Molecule 8 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	A	480	Total O 480 480	0	0
8	B	386	Total O 386 386	0	0
8	C	98	Total O 98 98	0	0
8	D	508	Total O 508 508	0	0
8	E	385	Total O 385 385	0	0
8	F	136	Total O 136 136	0	0




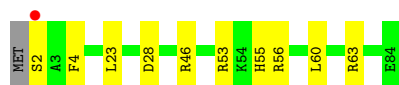
- Molecule 3: Toluene-4-monooxygenase system protein B

Chain C: 



- Molecule 3: Toluene-4-monooxygenase system protein B

Chain F: 





## 4 Data and refinement statistics i

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	169.39Å 176.91Å 55.94Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	40.78 – 1.72 40.78 – 1.72	Depositor EDS
% Data completeness (in resolution range)	98.8 (40.78-1.72) 95.4 (40.78-1.72)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.69 (at 1.72Å)	Xtrriage
Refinement program	PHENIX dev_1839	Depositor
R, $R_{free}$	0.130 , 0.173 0.132 , 0.175	Depositor DCC
$R_{free}$ test set	2009 reflections (1.13%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	13.1	Xtrriage
Anisotropy	0.265	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 47.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtrriage
Estimated twinning fraction	0.013 for k,h,-l	Xtrriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	16655	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	16.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 9.41% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  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: MBN, MG, CL, FE

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	0.33	0/4199	0.47	0/5700
1	D	0.34	0/4184	0.48	0/5679
2	B	0.34	0/2600	0.50	0/3535
2	E	0.34	0/2620	0.50	0/3560
3	C	0.28	0/674	0.49	0/912
3	F	0.35	0/686	0.52	0/928
All	All	0.33	0/14963	0.49	0/20314

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 [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	4064	0	3829	51	0
1	D	4052	0	3815	32	0
2	B	2519	0	2384	14	0
2	E	2542	0	2424	14	0
3	C	662	0	657	8	0
3	F	674	0	667	7	0
4	A	2	0	0	0	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	D	2	0	0	0	0
5	A	1	0	0	0	0
5	B	4	0	0	0	0
5	C	1	0	0	0	0
5	D	1	0	0	0	0
5	E	2	0	0	0	0
6	A	28	32	32	8	0
6	B	7	8	8	0	0
6	D	28	32	32	7	0
7	D	1	0	0	0	0
8	A	480	0	0	16	2
8	B	386	0	0	6	0
8	C	98	0	0	2	0
8	D	508	0	0	11	0
8	E	385	0	0	7	3
8	F	136	0	0	0	0
All	All	16583	72	13848	128	3

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 128 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:147:GLU:OE1	8:A:601:HOH:O	1.89	0.90
1:A:204[A]:GLN:HG3	1:A:269:PHE:HE1	1.44	0.82
1:D:481:LYS:HE2	8:D:616:HOH:O	1.77	0.82
1:D:6:ARG:NH2	8:D:602:HOH:O	2.13	0.80
1:D:231:GLU:OE1	8:D:601:HOH:O	1.99	0.79

All (3) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
8:A:1019:HOH:O	8:E:566:HOH:O[4_555]	1.96	0.24
8:A:757:HOH:O	8:E:699:HOH:O[1_556]	2.07	0.13
8:E:694:HOH:O	8:E:732:HOH:O[2_555]	2.19	0.01

## 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	493/493 (100%)	482 (98%)	11 (2%)	0	100	100
1	D	491/493 (100%)	483 (98%)	8 (2%)	0	100	100
2	B	301/327 (92%)	297 (99%)	4 (1%)	0	100	100
2	E	305/327 (93%)	301 (99%)	4 (1%)	0	100	100
3	C	81/84 (96%)	78 (96%)	3 (4%)	0	100	100
3	F	82/84 (98%)	80 (98%)	2 (2%)	0	100	100
All	All	1753/1808 (97%)	1721 (98%)	32 (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	420/418 (100%)	416 (99%)	4 (1%)	76	65
1	D	418/418 (100%)	414 (99%)	4 (1%)	76	65
2	B	276/296 (93%)	273 (99%)	3 (1%)	73	62
2	E	278/296 (94%)	277 (100%)	1 (0%)	91	86
3	C	74/75 (99%)	71 (96%)	3 (4%)	30	11
3	F	75/75 (100%)	75 (100%)	0	100	100
All	All	1541/1578 (98%)	1526 (99%)	15 (1%)	78	65

5 of 15 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	C	2	SER
1	D	490	ASP
3	C	56	ARG
2	E	17	GLU
1	D	123	ASN

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

Mol	Chain	Res	Type
2	E	100	GLN

### 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 23 ligands modelled in this entry, 14 are monoatomic - leaving 9 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	$\# Z  > 2$
6	MBN	B	405	-	7,7,7	0.35	0	8,8,8	0.48	0
6	MBN	D	505	-	7,7,7	0.24	0	8,8,8	0.44	0
6	MBN	A	507	-	7,7,7	0.34	0	8,8,8	0.35	0
6	MBN	A	505	-	7,7,7	0.25	0	8,8,8	0.41	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
6	MBN	D	504	-	7,7,7	0.24	0	8,8,8	0.35	0
6	MBN	A	504	-	7,7,7	0.27	0	8,8,8	0.29	0
6	MBN	D	508	-	7,7,7	0.30	0	8,8,8	0.32	0
6	MBN	A	506	-	7,7,7	0.29	0	8,8,8	0.60	0
6	MBN	D	507	-	7,7,7	0.32	0	8,8,8	0.45	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
6	MBN	B	405	-	-	-	0/1/1/1
6	MBN	D	505	-	-	-	0/1/1/1
6	MBN	A	507	-	-	-	0/1/1/1
6	MBN	A	505	-	-	-	0/1/1/1
6	MBN	D	504	-	-	-	0/1/1/1
6	MBN	A	504	-	-	-	0/1/1/1
6	MBN	D	508	-	-	-	0/1/1/1
6	MBN	A	506	-	-	-	0/1/1/1
6	MBN	D	507	-	-	-	0/1/1/1

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.

5 monomers are involved in 15 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	D	505	MBN	1	0
6	A	505	MBN	1	0
6	D	508	MBN	1	0
6	A	506	MBN	7	0
6	D	507	MBN	5	0

## 5.7 Other polymers [i](#)

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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	490/493 (99%)	-0.50	4 (0%) 86 89	6, 14, 29, 75	0
1	D	490/493 (99%)	-0.50	6 (1%) 79 83	6, 12, 28, 68	0
2	B	301/327 (92%)	-0.63	3 (0%) 82 85	6, 10, 23, 60	1 (0%)
2	E	304/327 (92%)	-0.60	3 (0%) 82 85	6, 11, 27, 69	0
3	C	82/84 (97%)	-0.20	4 (4%) 29 33	11, 19, 39, 55	0
3	F	83/84 (98%)	-0.58	1 (1%) 79 83	7, 12, 27, 37	0
All	All	1750/1808 (96%)	-0.53	21 (1%) 79 83	6, 12, 29, 75	1 (0%)

The worst 5 of 21 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	491	LYS	7.0
1	A	491	LYS	6.3
2	E	18	MET	5.9
1	D	295	TYR	4.6
3	C	2	SER	4.6

### 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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
6	MBN	A	507	7/7	0.69	0.29	28,31,37,37	15
6	MBN	D	508	7/7	0.77	0.23	24,27,33,33	15
6	MBN	B	405	7/7	0.78	0.20	19,22,26,27	15
6	MBN	D	507	7/7	0.80	0.24	33,36,43,44	15
6	MBN	A	506	7/7	0.87	0.17	28,33,38,39	0
5	MG	B	404	1/1	0.92	0.16	42,42,42,42	0
5	MG	B	403	1/1	0.94	0.21	41,41,41,41	0
6	MBN	A	505	7/7	0.95	0.09	17,21,25,26	0
6	MBN	D	505	7/7	0.95	0.09	18,21,24,25	0
5	MG	C	101	1/1	0.97	0.20	34,34,34,34	0
5	MG	E	401	1/1	0.98	0.10	28,28,28,28	0
6	MBN	A	504	7/7	0.98	0.05	9,11,13,13	0
5	MG	B	402	1/1	0.98	0.20	29,29,29,29	0
5	MG	E	402	1/1	0.99	0.24	34,34,34,34	0
5	MG	A	503	1/1	0.99	0.11	25,25,25,25	0
6	MBN	D	504	7/7	0.99	0.08	5,8,10,10	0
5	MG	D	503	1/1	1.00	0.04	17,17,17,17	0
5	MG	B	401	1/1	1.00	0.01	10,10,10,10	0
4	FE	A	502	1/1	1.00	0.05	14,14,14,14	0
4	FE	D	501	1/1	1.00	0.05	9,9,9,9	0
4	FE	D	502	1/1	1.00	0.05	13,13,13,13	0
4	FE	A	501	1/1	1.00	0.06	11,11,11,11	0
7	CL	D	506	1/1	1.00	0.02	16,16,16,16	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.