



# wwPDB X-ray Structure Validation Summary Report

Nov 6, 2023 – 09:33 PM EST

PDB ID : 4TZZ  
Title : Co-crystals of the Ternary Complex Containing a T-box Stem I RNA, its Cognate tRNAGly, and B. subtilis YbxF protein, treated by removing lithium sulfate and increasing PEG3350 concentration from 20% to 45% post crystallization  
Authors : Zhang, J.; Ferre-D'Amare, A.R.  
Deposited on : 2014-07-11  
Resolution : 3.64 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 : 2.36  
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

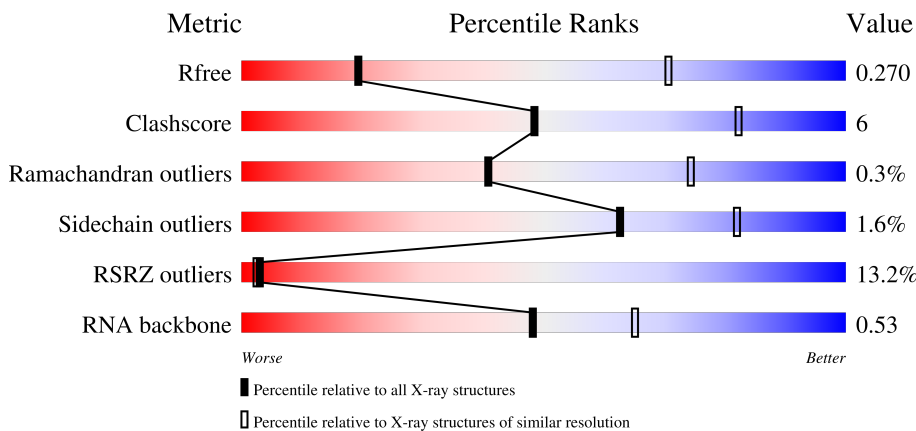
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.64 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	1341 (3.78-3.50)
Clashscore	141614	1439 (3.78-3.50)
Ramachandran outliers	138981	1391 (3.78-3.50)
Sidechain outliers	138945	1391 (3.78-3.50)
RSRZ outliers	127900	1242 (3.78-3.50)
RNA backbone	3102	1019 (4.26-3.00)

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	82	
1	D	82	
1	G	82	

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	J	82	
2	B	75	
2	E	75	
2	H	75	
2	K	75	
3	C	102	
3	F	102	
3	I	102	
3	L	102	

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
4	MG	B	102	-	-	-	X
4	MG	C	201	-	-	-	X
4	MG	E	103	-	-	-	X
4	MG	E	105	-	-	-	X
4	MG	E	107	-	-	-	X
4	MG	F	202	-	-	-	X
4	MG	F	204	-	-	-	X
4	MG	H	101	-	-	-	X
4	MG	H	103	-	-	-	X
4	MG	L	202	-	-	-	X

## 2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 17274 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 Ribosome-associated protein L7Ae-like.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	S	Se			
1	A	81	524	328	88	105	1	2	0	0	0
1	D	81	524	328	88	105	1	2	0	0	0
1	G	81	524	328	88	105	1	2	0	0	0
1	J	81	524	328	88	105	1	2	0	0	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	GLY	-	expression tag	UNP P46350
D	1	GLY	-	expression tag	UNP P46350
G	1	GLY	-	expression tag	UNP P46350
J	1	GLY	-	expression tag	UNP P46350

- Molecule 2 is a RNA chain called engineered tRNAGly.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
2	B	75	1592	708	289	520	75	0	0	0
2	E	75	1592	708	289	520	75	0	0	0
2	H	75	1592	708	289	520	75	0	0	0
2	K	75	1591	708	288	520	75	0	0	0

- Molecule 3 is a RNA chain called T-box Stem I RNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	102	Total	C	N	O	P	0	0	0
			2194	982	411	700	101			
3	F	102	Total	C	N	O	P	0	0	0
			2197	982	411	702	102			
3	I	102	Total	C	N	O	P	0	0	0
			2197	982	411	702	102			
3	L	102	Total	C	N	O	P	0	0	0
			2194	982	411	700	101			

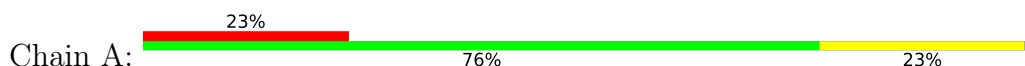
- Molecule 4 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	B	2	Total	Mg	0	0
			2	2		
4	C	4	Total	Mg	0	0
			4	4		
4	E	7	Total	Mg	0	0
			7	7		
4	F	4	Total	Mg	0	0
			4	4		
4	H	5	Total	Mg	0	0
			5	5		
4	I	5	Total	Mg	0	0
			5	5		
4	L	2	Total	Mg	0	0
			2	2		

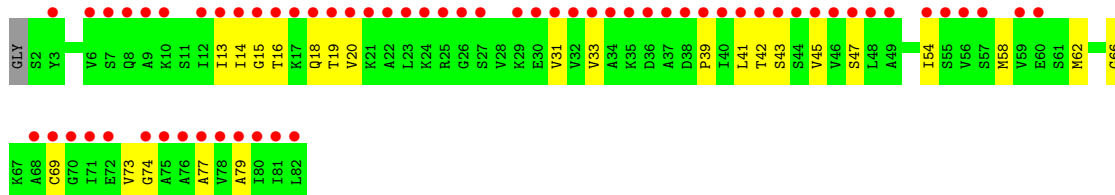
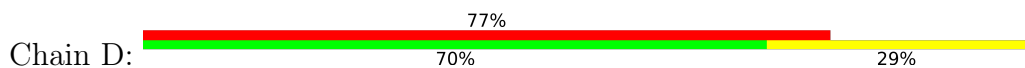
### 3 Residue-property plots [i](#)

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.

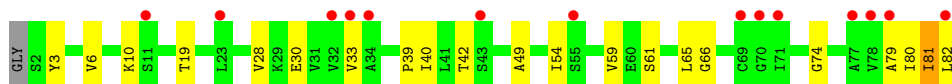
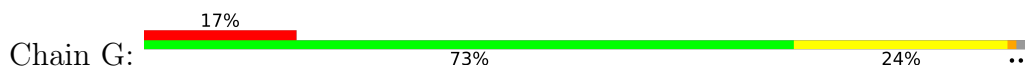
- Molecule 1: Ribosome-associated protein L7Ae-like



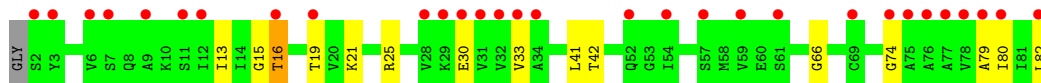
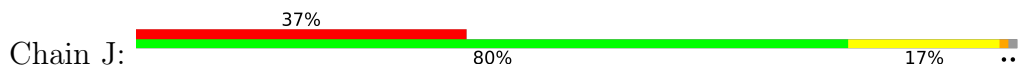
- Molecule 1: Ribosome-associated protein L7Ae-like



- Molecule 1: Ribosome-associated protein L7Ae-like

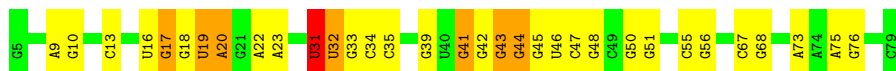


- Molecule 1: Ribosome-associated protein L7Ae-like



- Molecule 2: engineered tRNA<sup>Gly</sup>

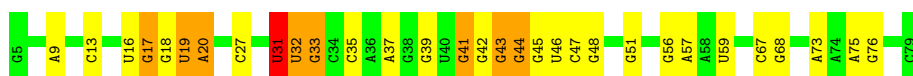




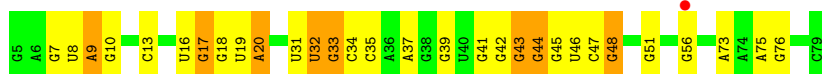
• Molecule 2: engineered tRNAGly



• Molecule 2: engineered tRNAGly



• Molecule 2: engineered tRNAGly



• Molecule 3: T-box Stem I RNA



• Molecule 3: T-box Stem I RNA

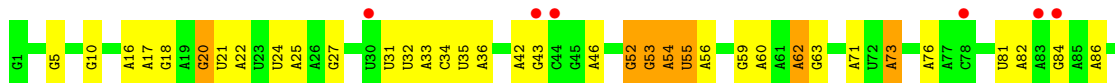


• Molecule 3: T-box Stem I RNA





- Molecule 3: T-box Stem I RNA





## 4 Data and refinement statistics i

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	70.64Å 260.73Å 70.70Å 90.00° 92.80° 90.00°	Depositor
Resolution (Å)	70.61 – 3.64 130.37 – 3.64	Depositor EDS
% Data completeness (in resolution range)	84.1 (70.61-3.64) 84.0 (130.37-3.64)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.54 (at 3.67Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8.2_1309)	Depositor
R, $R_{free}$	0.217 , 0.269 0.218 , 0.270	Depositor DCC
$R_{free}$ test set	1230 reflections (5.11%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	147.8	Xtriage
Anisotropy	0.242	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.29 , 84.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.43$ , $\langle L^2 \rangle = 0.26$	Xtriage
Estimated twinning fraction	0.076 for l,k,-h 0.096 for h,-k,-l 0.137 for l,-k,h	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	17274	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	117.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.60% 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: MG

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	0.19	0/523	0.43	0/709
1	D	0.23	0/523	0.47	0/709
1	G	0.21	0/523	0.46	0/709
1	J	0.19	0/523	0.41	0/709
2	B	0.18	0/1780	0.75	1/2775 (0.0%)
2	E	0.18	0/1780	0.76	0/2775
2	H	0.17	0/1780	0.74	1/2775 (0.0%)
2	K	0.17	0/1779	0.73	0/2773
3	C	0.18	0/2461	0.74	0/3838
3	F	0.18	0/2464	0.73	0/3842
3	I	0.18	0/2464	0.75	0/3842
3	L	0.18	0/2461	0.74	0/3838
All	All	0.18	0/19061	0.72	2/29294 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	H	31	U	P-O3'-C3'	6.88	127.96	119.70
2	B	31	U	P-O3'-C3'	5.48	126.28	119.70

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	524	0	505	10	0
1	D	524	0	505	14	0
1	G	524	0	505	13	0
1	J	524	0	505	8	0
2	B	1592	0	804	13	0
2	E	1592	0	804	12	0
2	H	1592	0	804	12	0
2	K	1591	0	802	12	0
3	C	2194	0	1100	14	0
3	F	2197	0	1100	17	0
3	I	2197	0	1098	17	0
3	L	2194	0	1101	20	0
4	B	2	0	0	0	0
4	C	4	0	0	0	0
4	E	7	0	0	0	0
4	F	4	0	0	0	0
4	H	5	0	0	0	0
4	I	5	0	0	0	0
4	L	2	0	0	0	0
All	All	17274	0	9633	154	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 6.

The worst 5 of 154 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)
2:K:10:G:N7	2:K:44:G:N2	2.23	0.87
2:K:32:U:O2'	2:K:34:C:OP2	1.95	0.83
3:L:35:U:H3	3:L:73:A:H61	1.33	0.76
2:H:31:U:H4'	2:H:32:U:O5'	1.84	0.76
2:E:32:U:H4'	2:E:33:G:OP1	1.84	0.76

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	79/82 (96%)	73 (92%)	6 (8%)	0	100	100
1	D	79/82 (96%)	72 (91%)	7 (9%)	0	100	100
1	G	79/82 (96%)	73 (92%)	5 (6%)	1 (1%)	12	48
1	J	79/82 (96%)	72 (91%)	7 (9%)	0	100	100
All	All	316/328 (96%)	290 (92%)	25 (8%)	1 (0%)	41	74

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	G	81	ILE

### 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	48/63 (76%)	48 (100%)	0	100	100
1	D	48/63 (76%)	47 (98%)	1 (2%)	53	77
1	G	48/63 (76%)	47 (98%)	1 (2%)	53	77
1	J	48/63 (76%)	47 (98%)	1 (2%)	53	77
All	All	192/252 (76%)	189 (98%)	3 (2%)	62	82

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	62	MSE
1	G	40	ILE
1	J	16	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 5.3.3 RNA [i](#)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
2	B	74/75 (98%)	20 (27%)	2 (2%)
2	E	74/75 (98%)	20 (27%)	3 (4%)
2	H	74/75 (98%)	20 (27%)	3 (4%)
2	K	74/75 (98%)	20 (27%)	2 (2%)
3	C	101/102 (99%)	21 (20%)	0
3	F	101/102 (99%)	20 (19%)	0
3	I	101/102 (99%)	20 (19%)	0
3	L	101/102 (99%)	21 (20%)	0
All	All	700/708 (98%)	162 (23%)	10 (1%)

5 of 162 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
2	B	9	A
2	B	13	C
2	B	16	U
2	B	17	G
2	B	18	G

5 of 10 RNA pucker outliers are listed below:

Mol	Chain	Res	Type
2	H	32	U
2	K	19	U
2	K	32	U
2	E	31	U
2	E	32	U

## 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 29 ligands modelled in this entry, 29 are monoatomic - leaving 0 for Mogul analysis.

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 [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	79/82 (96%)	1.08	19 (24%) 0 0	102, 170, 227, 242	0
1	D	79/82 (96%)	5.17	63 (79%) 0 0	146, 273, 306, 334	0
1	G	79/82 (96%)	0.78	14 (17%) 1 0	62, 105, 160, 170	0
1	J	79/82 (96%)	1.58	30 (37%) 0 0	93, 143, 193, 211	0
2	B	75/75 (100%)	0.18	0 100 100	73, 101, 125, 139	0
2	E	75/75 (100%)	0.22	0 100 100	61, 81, 146, 192	0
2	H	75/75 (100%)	0.05	0 100 100	57, 100, 139, 152	0
2	K	75/75 (100%)	0.28	1 (1%) 77 64	60, 87, 129, 165	0
3	C	102/102 (100%)	0.07	0 100 100	69, 99, 123, 144	0
3	F	102/102 (100%)	0.16	2 (1%) 65 50	63, 116, 170, 197	0
3	I	102/102 (100%)	0.10	0 100 100	63, 116, 163, 220	0
3	L	102/102 (100%)	0.40	6 (5%) 22 13	82, 144, 196, 217	0
All	All	1024/1036 (98%)	0.79	135 (13%) 3 2	57, 114, 242, 334	0

The worst 5 of 135 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	37	ALA	15.1
1	D	42	THR	14.8
1	D	45	VAL	11.7
1	D	41	LEU	11.6
1	D	57	SER	11.5

### 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
4	MG	H	103	1/1	0.26	0.68	69,69,69,69	0
4	MG	E	103	1/1	0.31	0.56	86,86,86,86	0
4	MG	F	204	1/1	0.38	0.76	86,86,86,86	0
4	MG	E	107	1/1	0.42	0.61	68,68,68,68	0
4	MG	F	202	1/1	0.46	0.40	61,61,61,61	0
4	MG	H	101	1/1	0.55	0.56	47,47,47,47	0
4	MG	B	102	1/1	0.61	0.67	60,60,60,60	0
4	MG	C	201	1/1	0.64	0.41	49,49,49,49	0
4	MG	I	202	1/1	0.68	0.15	48,48,48,48	0
4	MG	E	104	1/1	0.72	0.18	48,48,48,48	0
4	MG	I	203	1/1	0.72	0.22	72,72,72,72	0
4	MG	E	105	1/1	0.74	0.42	59,59,59,59	0
4	MG	I	201	1/1	0.74	0.16	66,66,66,66	0
4	MG	L	202	1/1	0.77	0.42	68,68,68,68	0
4	MG	B	101	1/1	0.82	0.79	77,77,77,77	0
4	MG	C	204	1/1	0.82	0.30	58,58,58,58	0
4	MG	I	204	1/1	0.83	0.45	48,48,48,48	0
4	MG	C	203	1/1	0.85	0.27	48,48,48,48	0
4	MG	I	205	1/1	0.86	0.27	56,56,56,56	0
4	MG	H	102	1/1	0.88	0.33	96,96,96,96	0
4	MG	E	106	1/1	0.89	0.49	92,92,92,92	0
4	MG	F	203	1/1	0.89	0.19	102,102,102,102	0
4	MG	H	104	1/1	0.90	0.59	60,60,60,60	0
4	MG	E	101	1/1	0.91	0.18	33,33,33,33	0
4	MG	F	201	1/1	0.92	0.17	69,69,69,69	0
4	MG	C	202	1/1	0.94	0.27	50,50,50,50	0
4	MG	L	201	1/1	0.95	0.16	63,63,63,63	0
4	MG	E	102	1/1	0.95	0.26	93,93,93,93	0
4	MG	H	105	1/1	0.98	0.22	91,91,91,91	0



## 6.5 Other polymers [i](#)

There are no such residues in this entry.