

# wwPDB X-ray Structure Validation Summary Report (i)

Dec 17, 2023 – 11:51 AM EST

PDB ID : 4TUG

Title : Crystal structure of MjMre11-DNA2 complex

Authors : Sung, S.; Cho, Y.

Deposited on : 2014-06-24

Resolution : 3.55 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:

 $\begin{array}{ccc} \text{MolProbity} & : & 4.02\text{b-}467 \\ \text{Xtriage (Phenix)} & : & 1.13 \end{array}$ 

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)

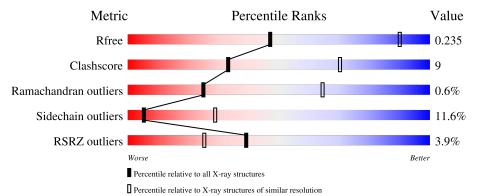
Validation Pipeline (wwPDB-VP) : 2.36

## 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 3.55 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	Similar resolution	
Metric	$(\# \mathrm{Entries})$	$(\#  ext{Entries},  ext{ resolution range}( ext{Å}))$	
$R_{free}$	130704	1020 (3.62-3.50)	
Clashscore	141614	1100 (3.62-3.50)	
Ramachandran outliers	138981	1065 (3.62-3.50)	
Sidechain outliers	138945	1066 (3.62-3.50)	
RSRZ outliers	127900	1009 (3.64-3.48)	

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	337	65%	25%	• 6%			
1	В	337	66%	21%	• 9%			
1	С	337	66%	23%	• 7%			
1	D	337	64%	24%	• 9%			
1	Е	337	65%	25%	• 7%			

Continued on next page...



Continued from previous page...

Mol	Chain	Length	Quality of chain					
1	F	337	10%	61%		26%	• 9%	
2	G	14	29%		50%		21%	
3	Н	15		67%		27%	7%	



# 2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 16104 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 DNA double-strand break repair protein Mre11.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	С	313	Total	С	N	О	S	0	0	0
1		313	2601	1680	434	476	11	U	U	
1	A	317	Total	С	N	О	S	0	0	0
1	Λ	911	2635	1698	444	482	11	U	0	
1	В	306	Total	С	N	О	S	0	0	0
1	Ъ	300	2546	1644	427	465	10			
1	D	307	Total	С	N	О	S	0	0	0
1	D	307	2556	1650	430	466	10	U	0	
1	Е	314	Total	С	N	О	S	0	0	0
1	12	314	2605	1682	437	476	10	U	0	
1	F	305	Total	С	N	О	S	0	0	0
1	I'	300	2537	1640	428	459	10	U	U	

There are 24 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
С	-3	ARG	-	expression tag	UNP Q58719
С	-2	GLY	-	expression tag	UNP Q58719
С	-1	SER	-	expression tag	UNP Q58719
С	0	HIS	-	expression tag	UNP Q58719
A	-3	ARG	-	expression tag	UNP Q58719
A	-2	GLY	-	expression tag	UNP Q58719
A	-1	SER	-	expression tag	UNP Q58719
A	0	HIS	-	expression tag	UNP Q58719
В	-3	ARG	-	expression tag	UNP Q58719
В	-2	GLY	-	expression tag	UNP Q58719
В	-1	SER	-	expression tag	UNP Q58719
В	0	HIS	-	expression tag	UNP Q58719
D	-3	ARG	-	expression tag	UNP Q58719
D	-2	GLY	-	expression tag	UNP Q58719
D	-1	SER	-	expression tag	UNP Q58719
D	0	HIS	-	expression tag	UNP Q58719
Е	-3	ARG	-	expression tag	UNP Q58719

Continued on next page...



Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
E	-2	GLY	-	expression tag	UNP Q58719
E	-1	SER	-	expression tag	UNP Q58719
E	0	HIS	-	expression tag	UNP Q58719
F	-3	ARG	-	expression tag	UNP Q58719
F	-2	GLY	-	expression tag	UNP Q58719
F	-1	SER	-	expression tag	UNP Q58719
F	0	HIS	-	expression tag	UNP Q58719

• Molecule 2 is a DNA chain called DNA (5'-D(P\*GP\*CP\*AP\*CP\*GP\*TP\*AP\*GP\*GP\*AP\*CP\*AP\*GP\*C)-3').

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace		
2	С	1.4	Total	С	N	О	Р	0	0	0
	G	14	290	136	59	81	14	0	U	

• Molecule 3 is a DNA chain called DNA (5'-D(P\*CP\*TP\*GP\*TP\*CP\*CP\*TP\*AP\*CP\*GP \*TP\*GP\*CP\*CP\*A)-3').

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace		
3	Н	15	Total 302	C 144	N 51	O 92	P 15	0	0	0

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

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	С	2	Total Mg 2 2	0	0
4	A	2	Total Mg 2 2	0	0
4	В	2	Total Mg 2 2	0	0
4	D	2	Total Mg 2 2	0	0
4	E	2	Total Mg 2 2	0	0
4	F	2	Total Mg 2 2	0	0

• Molecule 5 is water.



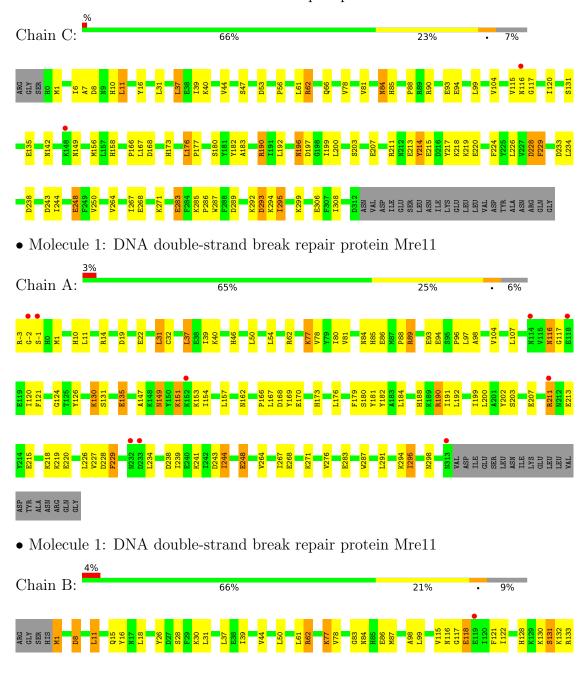
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	С	2	Total O 2 2	0	0
5	A	7	Total O 7 7	0	0
5	В	5	Total O 5 5	0	0
5	D	2	Total O 2 2	0	0
5	E	3	Total O 3 3	0	0
5	F	1	Total O 1 1	0	0



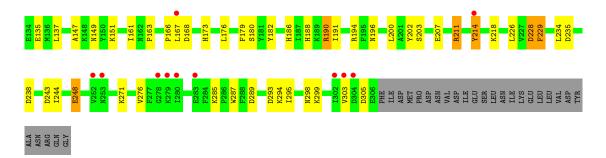
## 3 Residue-property plots (i)

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.

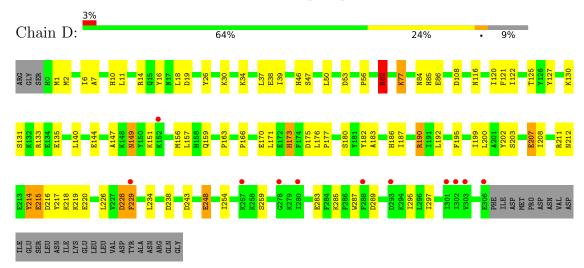
• Molecule 1: DNA double-strand break repair protein Mre11



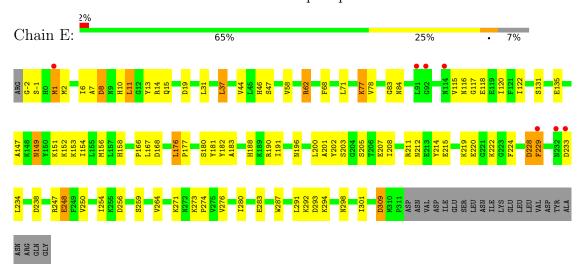




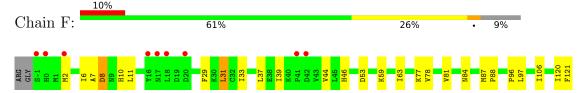
• Molecule 1: DNA double-strand break repair protein Mre11



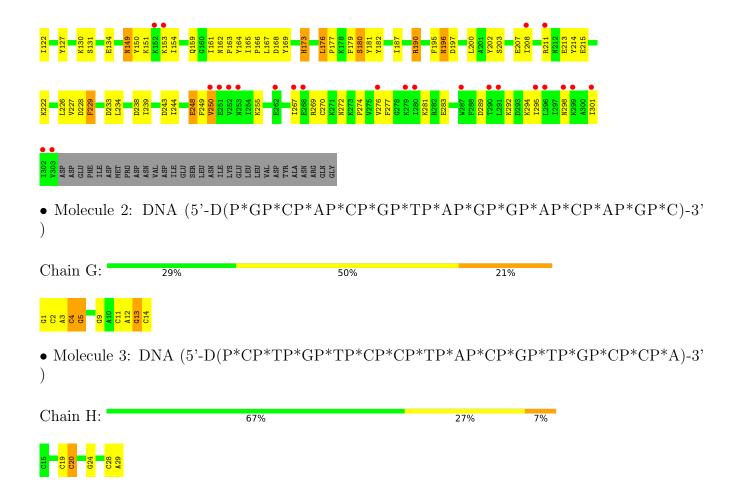
• Molecule 1: DNA double-strand break repair protein Mre11



 $\bullet$  Molecule 1: DNA double-strand break repair protein Mre11









# 4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants	91.53Å 185.62Å 106.21Å	Donositor
a, b, c, $\alpha$ , $\beta$ , $\gamma$	90.00° 99.86° 90.00°	Depositor
Resolution (Å)	42.42 - 3.55	Depositor
rtesolution (A)	43.05 - 3.55	EDS
% Data completeness	94.1 (42.42-3.55)	Depositor
(in resolution range)	94.2 (43.05-3.55)	EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$< I/\sigma(I) > 1$	3.31 (at 3.57Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8_1069)	Depositor
D D.	0.184 , 0.236	Depositor
$R, R_{free}$	0.187 , $0.235$	DCC
$R_{free}$ test set	1996 reflections $(5.04\%)$	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	79.8	Xtriage
Anisotropy	0.097	Xtriage
Bulk solvent $k_{sol}(e/Å^3)$ , $B_{sol}(Å^2)$	0.32 , 101.3	EDS
L-test for twinning <sup>2</sup>	$< L >=0.48, < L^2>=0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	16104	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	144.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: The largest off-origin peak in the Patterson function is 3.00% of the height of the origin peak. No significant pseudotranslation is detected.

<sup>&</sup>lt;sup>2</sup>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.



<sup>&</sup>lt;sup>1</sup>Intensities estimated from amplitudes.

## 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		
MIOI		RMSZ	# Z  > 5	RMSZ	# Z  > 5	
1	A	0.35	0/2692	0.56	0/3613	
1	В	0.33	0/2600	0.55	0/3489	
1	С	0.35	0/2657	0.56	0/3567	
1	D	0.33	0/2611	0.57	1/3504~(0.0%)	
1	Е	0.33	0/2662	0.53	0/3574	
1	F	0.29	0/2592	0.52	0/3478	
2	G	0.58	0/326	1.53	6/501 (1.2%)	
3	Н	0.75	0/336	1.63	4/515 (0.8%)	
All	All	0.35	0/16476	0.63	11/22241 (0.0%)	

There are no bond length outliers.

The worst 5 of 11 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	$\mathbf{Z}$	$\mathbf{Observed}(^{o})$	$\operatorname{Ideal}({}^{o})$
3	Н	28	DC	O4'-C1'-N1	10.39	115.27	108.00
2	G	11	DC	O4'-C1'-N1	9.35	114.55	108.00
2	G	4	DC	O4'-C1'-N1	9.24	114.47	108.00
2	G	14	DC	O4'-C1'-N1	7.22	113.05	108.00
3	Н	29	DA	O4'-C1'-N9	6.95	112.87	108.00

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	2635	0	2650	56	0
1	В	2546	0	2572	39	0
1	С	2601	0	2618	40	0
1	D	2556	0	2579	42	0
1	Ε	2605	0	2620	46	0
1	F	2537	0	2570	59	0
2	G	290	0	156	7	0
3	Н	302	0	170	1	0
4	A	2	0	0	0	0
4	В	2	0	0	0	0
4	С	2	0	0	0	0
4	D	2	0	0	0	0
4	Ε	2	0	0	0	0
4	F	2	0	0	0	0
5	A	7	0	0	1	0
5	В	5	0	0	0	0
5	С	2	0	0	0	0
5	D	2	0	0	0	0
5	Ε	3	0	0	0	0
5	F	1	0	0	0	0
All	All	16104	0	15935	275	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 9.

The worst 5 of 275 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	$\begin{array}{c} {\rm Interatomic} \\ {\rm distance} \ ({\rm \AA}) \end{array}$	$\begin{array}{c} \text{Clash} \\ \text{overlap } (\text{\AA}) \end{array}$
1:B:1:MET:HG3	1:B:115:VAL:HG21	1.57	0.84
1:F:179:PHE:HB2	1:F:182:TYR:HE1	1.53	0.74
2:G:2:DC:H2'	2:G:3:DA:H8	1.54	0.73
1:F:127:TYR:HA	1:F:159:GLN:HE22	1.55	0.71
1:A:14:ARG:NH2	1:A:19:ASP:OD2	2.26	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	Perce	entiles
1	A	315/337 (94%)	299 (95%)	14 (4%)	2 (1%)	25	64
1	В	304/337 (90%)	290 (95%)	12 (4%)	2 (1%)	22	62
1	С	311/337 (92%)	293 (94%)	16 (5%)	2 (1%)	25	64
1	D	305/337 (90%)	290 (95%)	13 (4%)	2 (1%)	22	62
1	E	312/337 (93%)	295 (95%)	15 (5%)	2 (1%)	25	64
1	F	303/337 (90%)	289 (95%)	13 (4%)	1 (0%)	41	74
All	All	1850/2022 (92%)	1756 (95%)	83 (4%)	11 (1%)	25	64

#### 5 of 11 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	186	HIS
1	Е	309	ASP
1	С	166	PRO
1	В	186	HIS
1	С	213	GLU

#### 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	290/308~(94%)	258 (89%)	32 (11%)	6	31
1	В	280/308 (91%)	246 (88%)	34 (12%)	5	26
1	С	286/308 (93%)	249 (87%)	37 (13%)	4	24

Continued on next page...



Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles		
1	D	281/308 (91%)	248 (88%)	33 (12%)		5	28
1	E	286/308 (93%)	255 (89%)	31 (11%)		6	32
1	F	279/308 (91%)	249 (89%)	30 (11%)		6	32
All	All	1702/1848 (92%)	1505 (88%)	197 (12%)		5	29

5 of 197 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	135	GLU
1	Е	84	ASN
1	D	176	LEU
1	D	234	LEU
1	Е	180	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 11 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	159	GLN
1	Е	159	GLN
1	F	159	GLN
1	F	128	HIS
1	В	253	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)

Of 12 ligands modelled in this entry, 12 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,  $95^{th}$  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>	$\#\mathrm{RSRZ}{>}2$	$\mathrm{OWAB}(\mathrm{\AA}^2)$	Q < 0.9
1	A	317/337~(94%)	-0.17	9 (2%) 53 37	70, 114, 179, 243	0
1	В	306/337~(90%)	-0.04	12 (3%) 39 25	66, 130, 208, 285	0
1	С	313/337 (92%)	-0.20	2 (0%) 89 80	78, 114, 178, 221	0
1	D	307/337 (91%)	0.08	11 (3%) 42 29	78, 129, 232, 306	0
1	E	314/337~(93%)	-0.20	7 (2%) 62 45	79, 129, 208, 267	0
1	F	305/337~(90%)	0.36	33 (10%) 5 4	103, 191, 279, 339	0
2	G	14/14 (100%)	0.15	0 100 100	131, 184, 233, 233	0
3	Н	15/15 (100%)	0.29	0 100 100	134, 174, 222, 393	0
All	All	$1891/2051 \ (92\%)$	-0.03	74 (3%) 39 25	66, 132, 240, 393	0

The worst 5 of 74 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	F	211	ARG	6.2
1	F	253	ASN	6.1
1	D	278	GLY	5.7
1	F	279	LYS	5.3
1	F	296	LEU	5.3

#### 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,  $95^{th}$  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	$\mathbf{B} ext{-}\mathbf{factors}(\mathring{\mathbf{A}}^2)$	Q<0.9
4	MG	F	401	1/1	0.92	0.20	77,77,77,77	0
4	MG	В	401	1/1	0.93	0.36	57,57,57,57	0
4	MG	В	402	1/1	0.95	0.33	77,77,77,77	0
4	MG	F	402	1/1	0.95	0.21	121,121,121,121	0
4	MG	С	401	1/1	0.98	0.24	50,50,50,50	0
4	MG	С	402	1/1	0.98	0.18	77,77,77,77	0
4	MG	Е	402	1/1	0.98	0.23	102,102,102,102	0
4	MG	A	401	1/1	0.98	0.28	54,54,54,54	0
4	MG	A	402	1/1	0.98	0.28	55,55,55,55	0
4	MG	D	401	1/1	0.99	0.24	45,45,45,45	0
4	MG	D	402	1/1	0.99	0.24	61,61,61,61	0
4	MG	Е	401	1/1	0.99	0.20	70,70,70,70	0

## 6.5 Other polymers (i)

There are no such residues in this entry.

