



wwPDB X-ray Structure Validation Summary Report ⓘ

Apr 30, 2024 – 07:13 PM JST

PDB ID : 8WTB
Title : Crystal structure of McsA/McsB complex truncated by chymotrypsin
Authors : Arifuzzaman, M.; Kwon, E.; Kim, D.Y.
Deposited on : 2023-10-18
Resolution : 2.90 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 ⓘ 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 : 2.36.2
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) : 2.36.2

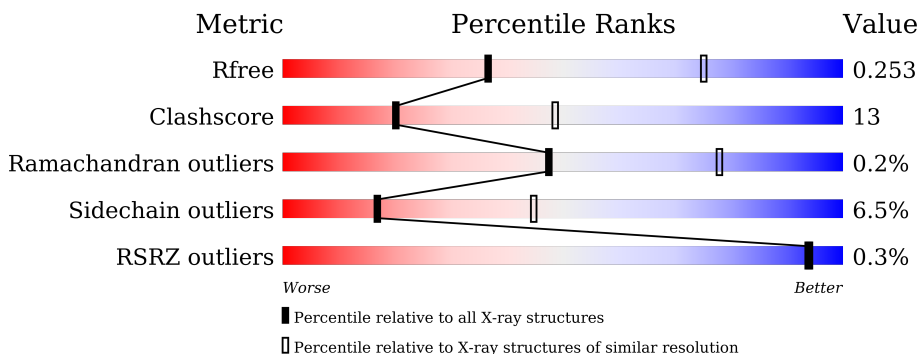
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.90 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	1957 (2.90-2.90)
Clashscore	141614	2172 (2.90-2.90)
Ramachandran outliers	138981	2115 (2.90-2.90)
Sidechain outliers	138945	2117 (2.90-2.90)
RSRZ outliers	127900	1906 (2.90-2.90)

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 $\leq 5\%$. 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	365	
1	C	365	
2	B	187	
2	D	187	

2 Entry composition i

There are 3 unique types of molecules in this entry. The entry contains 5288 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 Protein-arginine kinase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	S	Se			
1	A	233	1845	1158	324	355	3	5	0	0	0
1	C	233	1845	1158	324	355	3	5	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 P37570
A	0	SER	-	expression tag	UNP P37570
C	-1	GLY	-	expression tag	UNP P37570
C	0	SER	-	expression tag	UNP P37570

- Molecule 2 is a protein called Protein-arginine kinase activator protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	99	798	500	154	138	6	0	0	0
2	D	99	798	500	154	138	6	0	0	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	-1	GLY	-	expression tag	UNP P37569
B	0	SER	-	expression tag	UNP P37569
D	-1	GLY	-	expression tag	UNP P37569
D	0	SER	-	expression tag	UNP P37569

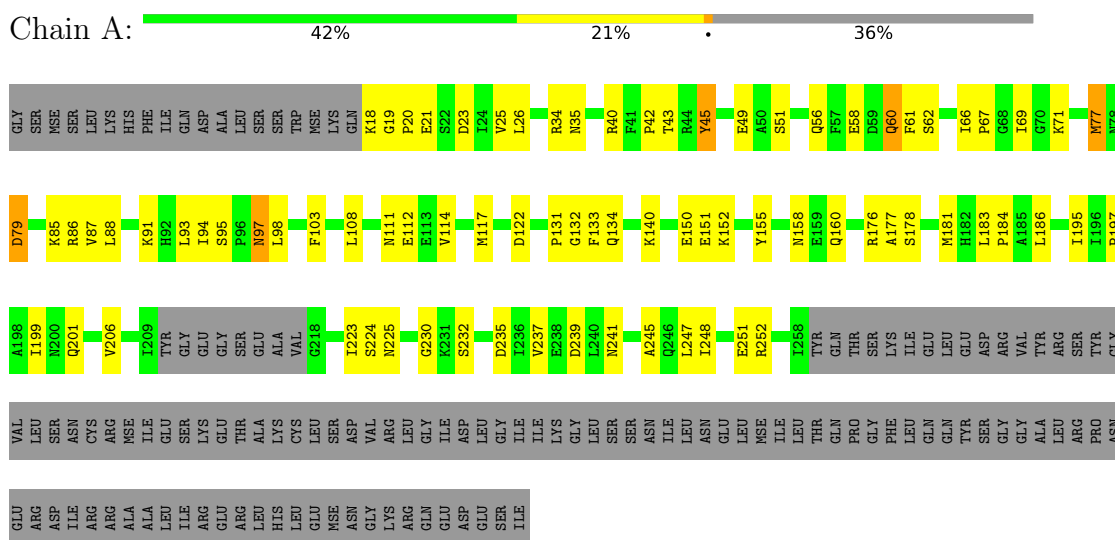
- Molecule 3 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	1	Total 1	Zn 1	0	0
3	D	1	Total 1	Zn 1	0	0

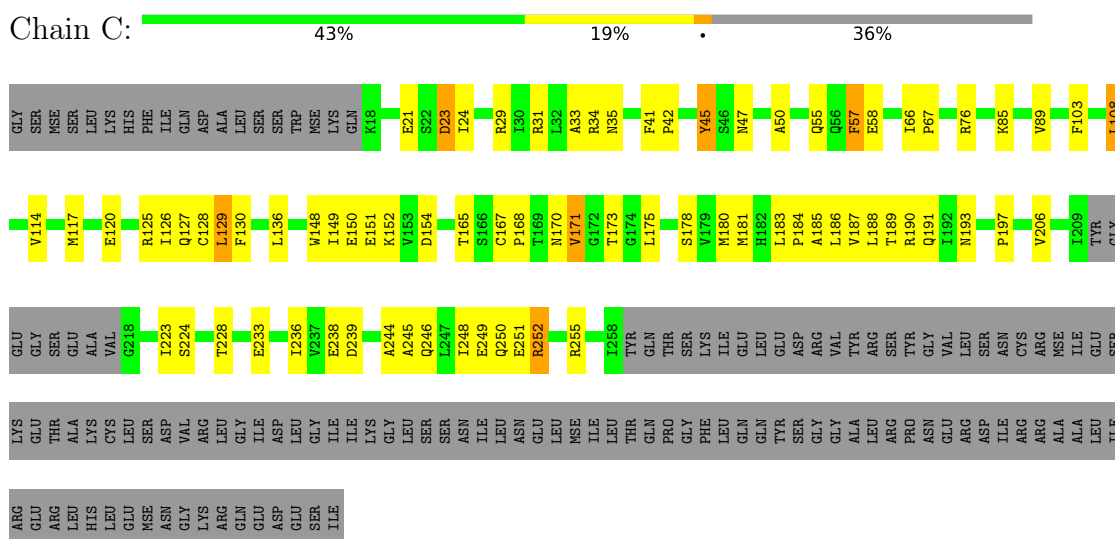
3 Residue-property plots [i](#)

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.

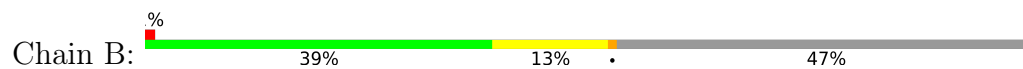
- Molecule 1: Protein-arginine kinase



- Molecule 1: Protein-arginine kinase



- Molecule 2: Protein-arginine kinase activator protein

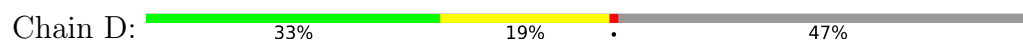


GLY	SER	GLN	CYS	GLU	GLU	HIS	HIS	ARG	PRO	ALA	THR	PHE	HIS	HIS	HIS	THR	THR	LYS	VAL	VAL	ASN	GLY	GLY	LYS	ILE	GLU	VAL	HIS	ILE	CYS	GLU	GLM	CYS	ALA	ALA	GLU	ASN	SER	SER	TYR	GLY	ILE	SER	ALA	ASN	GLN	GLY	PHE	SER	ILE	HIS	ASN	LEU	LEU	SER	GLY	LEU	LEU
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

ASN	MET	ASP	SER	PHE	GLN	ASN	ALA	THR	GLN	MET	PHE	SER	HIS	HIS	SER	GLU	GLN	LYS	I76	S79	P82	Q89	R92	T106	F106	H107	S108	M109	I110	T111	P112	I113	M121	T122	V123	R131	I132	G133	G134	N135	N136	H137	V138	R139	M144	L145	K146	V164	I168	R169
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------

Q173	S174	L175	K176	THR	ASP	ASP	GLU	GLU	GLN	GLU	GLU
------	------	------	------	-----	-----	-----	-----	-----	-----	-----	-----

- Molecule 2: Protein-arginine kinase activator protein



GLY	SER	MET	ILE	CYS	GLN	GLU	CYS	HIS	HIS	ARG	PRO	ALA	THR	PHE	HIS	HIS	PHE	THR	THR	LYS	VAL	VAL	ASN	GLY	GLU	LYS	ILE	GLU	VAL	ILE	ILE	CYS	GLU	GLN	CYS	ALA	ALA	LYS	GLU	ASN	SER	ASP	TYR	GLY	ILE	SER	ALA	ASN	GLN	GLY	PHE	SER	ILE	HIS	ASN	LEU	LEU	SER	GLY	LEU	LEU
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

ASN	MET	ASP	SER	PHE	GLN	ASN	ALA	THR	GLN	MET	PHE	SER	HIS	HIS	SER	GLU	GLN	I76	C81	P82	K83	C84	G85	M86	T87	F88	Q89	Q90	F91	R92	G98	C99	S100	E101	C102	T105	F106	H107	S108	H109	I110	L114	R115	G120	R131	G134	V138	L149
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------

L152	I153	E159	N160	A161	R165	D166	Q167	I168	R169	L170	L171	E172	Q173	S174	L175	K176	SER	THR	ASP	SER	GLU	GLU	GLN	GLU
------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	-----	-----	-----	-----	-----	-----	-----	-----

4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	179.87Å 75.59Å 79.19Å 90.00° 99.78° 90.00°	Depositor
Resolution (Å)	88.63 – 2.90 88.63 – 2.90	Depositor EDS
% Data completeness (in resolution range)	95.5 (88.63-2.90) 95.5 (88.63-2.90)	Depositor EDS
R_{merge}	0.16	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.54 (at 2.91Å)	Xtrriage
Refinement program	PHENIX (1.19rc5_4047: ???)	Depositor
R, R_{free}	0.205 , 0.259 0.200 , 0.253	Depositor DCC
R_{free} test set	1103 reflections (4.92%)	wwPDB-VP
Wilson B-factor (Å ²)	59.8	Xtrriage
Anisotropy	0.632	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 47.6	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	5288	wwPDB-VP
Average B, all atoms (Å ²)	57.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.75% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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:
ZN

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	0.54	0/1867	0.71	0/2514
1	C	0.51	0/1867	0.71	1/2514 (0.0%)
2	B	0.48	0/812	0.69	0/1084
2	D	0.55	2/812 (0.2%)	0.67	1/1084 (0.1%)
All	All	0.52	2/5358 (0.0%)	0.70	2/7196 (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	1
1	C	0	1
All	All	0	2

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	84	CYS	CB-SG	-7.94	1.68	1.82
2	D	102	CYS	CB-SG	-5.40	1.73	1.81

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	D	98	GLY	N-CA-C	-5.46	99.46	113.10
1	C	171	VAL	C-N-CA	-5.26	111.25	122.30

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	86	ARG	Sidechain
1	C	170	ASN	Peptide

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	1845	0	1842	54	0
1	C	1845	0	1842	50	0
2	B	798	0	815	18	0
2	D	798	0	815	21	0
3	B	1	0	0	0	0
3	D	1	0	0	0	0
All	All	5288	0	5314	133	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 13.

The worst 5 of 133 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:230:GLY:HA3	2:B:121:ASN:HD22	1.43	0.84
1:A:181:MSE:HE3	1:A:223:ILE:HD13	1.60	0.83
1:A:66:ILE:HD12	1:A:67:PRO:HD2	1.63	0.81
1:A:95:SER:H	1:A:117:MSE:HE1	1.48	0.77
1:C:42:PRO:HA	1:C:45:TYR:CE1	2.22	0.75

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	229/365 (63%)	215 (94%)	14 (6%)	0	100	100
1	C	229/365 (63%)	219 (96%)	9 (4%)	1 (0%)	34	66
2	B	97/187 (52%)	94 (97%)	3 (3%)	0	100	100
2	D	97/187 (52%)	90 (93%)	7 (7%)	0	100	100
All	All	652/1104 (59%)	618 (95%)	33 (5%)	1 (0%)	47	78

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	236	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	205/311 (66%)	191 (93%)	14 (7%)	16	42
1	C	205/311 (66%)	190 (93%)	15 (7%)	14	38
2	B	88/166 (53%)	83 (94%)	5 (6%)	20	51
2	D	88/166 (53%)	84 (96%)	4 (4%)	27	61
All	All	586/954 (61%)	548 (94%)	38 (6%)	17	45

5 of 38 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	180	MSE
2	D	100	SER
1	C	190	ARG
1	C	239	ASP
2	D	174	SER

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	250	GLN
2	B	163	HIS
1	C	97	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 2 ligands modelled in this entry, 2 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, 95th 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	228/365 (62%)	0.11	0 100 100	32, 51, 72, 89	0
1	C	228/365 (62%)	0.09	0 100 100	34, 53, 74, 92	0
2	B	99/187 (52%)	0.12	2 (2%) 65 63	40, 59, 84, 90	0
2	D	99/187 (52%)	0.06	0 100 100	44, 61, 86, 91	0
All	All	654/1104 (59%)	0.10	2 (0%) 94 94	32, 55, 81, 92	0

All (2) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	144	MET	2.1
2	B	132	ILE	2.0

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, 95th 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(\AA^2)	Q<0.9
3	ZN	D	201	1/1	0.98	0.20	51,51,51,51	0
3	ZN	B	201	1/1	0.99	0.20	49,49,49,49	0

6.5 Other polymers [i](#)

There are no such residues in this entry.