wwPDB X-ray Structure Validation Summary Report (i)

Apr 27, 2023 – 10:24 pm BST

PDB ID	:	8BI3
Title	:	Structure of E. coli Class 2 L-asparaginase EcAIII, mutant M200W (crystal M200W#1)
Authors	:	Sciuk, A.; Ruszkowski, M.; Jaskolski, M.; Loch, J.I.
Deposited on	:	2022-11-01
Resolution	:	1.45 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:

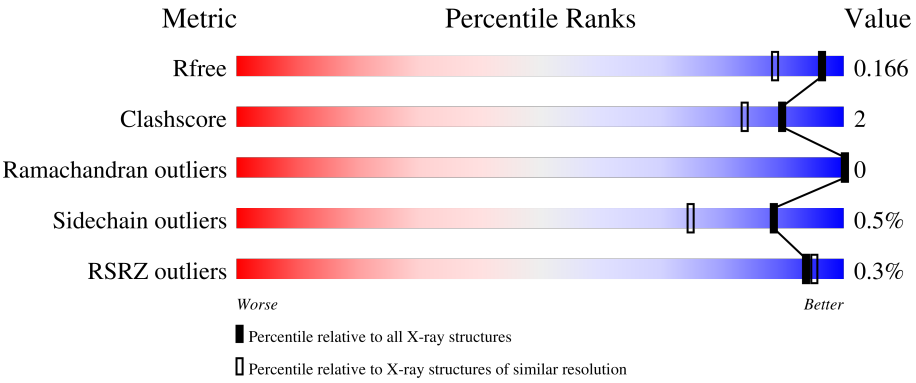
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.32.2
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.32.2

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 1.45 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	1156 (1.46-1.46)
Clashscore	141614	1202 (1.46-1.46)
Ramachandran outliers	138981	1178 (1.46-1.46)
Sidechain outliers	138945	1178 (1.46-1.46)
RSRZ outliers	127900	1139 (1.46-1.46)

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	AAA	178	87% · 11%
1	CCC	178	87% · 12%
2	BBB	143	90% 5% 6%
2	DDD	143	90% · 6%

2 Entry composition (i)

There are 9 unique types of molecules in this entry. The entry contains 5029 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 Isoaspartyl peptidase subunit alpha.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	AAA	158	Total	C	N	O	S	0	1	0
			1187	740	210	227	10			
1	CCC	157	Total	C	N	O	S	0	2	0
			1191	743	212	226	10			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
AAA	1	MET	-	initiating methionine	UNP P37595
CCC	1	MET	-	initiating methionine	UNP P37595

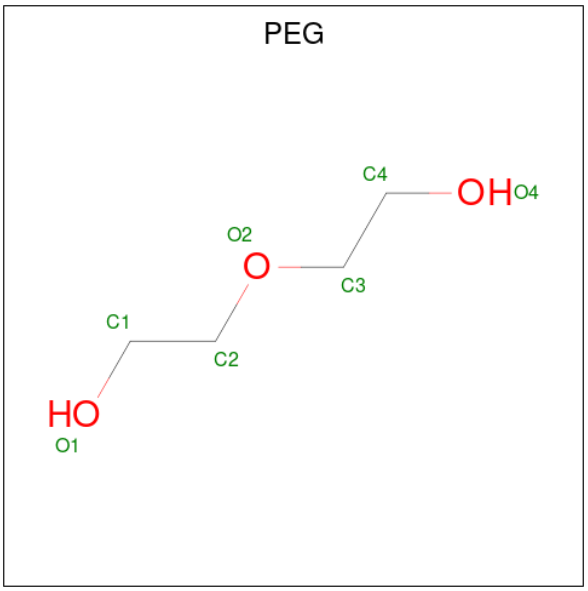
- Molecule 2 is a protein called Isoaspartyl peptidase subunit beta.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	BBB	135	Total	C	N	O	S	0	3	0
			981	618	166	191	6			
2	DDD	134	Total	C	N	O	S	0	2	0
			957	601	161	189	6			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
BBB	200	TRP	MET	engineered mutation	UNP P37595
DDD	200	TRP	MET	engineered mutation	UNP P37595

- Molecule 3 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula: $C_4H_{10}O_3$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	AAA	1	Total C O 7 4 3	0	0
3	BBB	1	Total C O 7 4 3	0	0
3	CCC	1	Total C O 7 4 3	0	0
3	DDD	1	Total C O 7 4 3	0	0

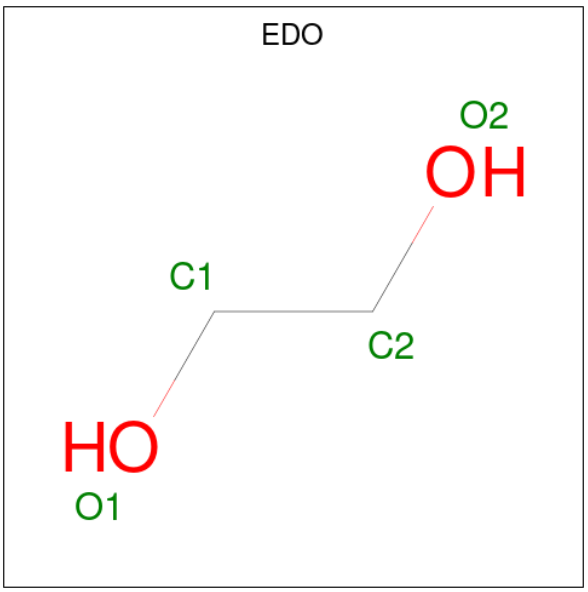
- Molecule 4 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	AAA	1	Total Na 1 1	0	0
4	CCC	1	Total Na 1 1	0	0

- Molecule 5 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

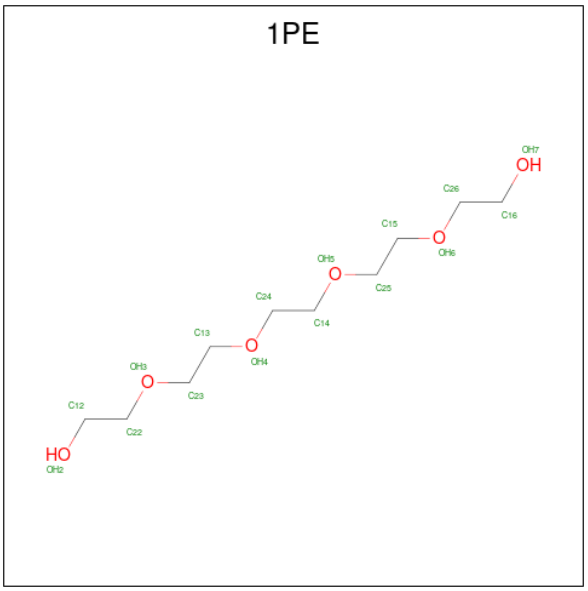
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	AAA	2	Total Cl 2 2	0	0
5	CCC	1	Total Cl 1 1	0	0
5	DDD	1	Total Cl 1 1	0	0

- Molecule 6 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: $C_2H_6O_2$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	BBB	1	Total C O 4 2 2	0	0

- Molecule 7 is PENTAETHYLENE GLYCOL (three-letter code: 1PE) (formula: $C_{10}H_{22}O_6$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	CCC	1	Total C O 16 10 6	0	0

- Molecule 8 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	CCC	1	Total Ca 1 1	0	0
8	DDD	1	Total Ca 1 1	0	0

- Molecule 9 is water.

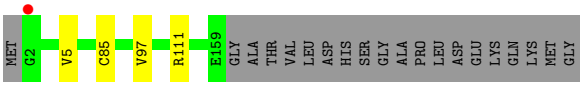
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
9	AAA	189	Total O 189 189	0	0
9	BBB	111	Total O 111 111	0	0
9	CCC	218	Total O 218 218	0	0
9	DDD	138	Total O 139 139	0	1

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.

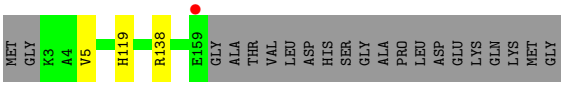
• Molecule 1: Isoaspartyl peptidase subunit alpha

Chain AAA:



• Molecule 1: Isoaspartyl peptidase subunit alpha

Chain CCC:

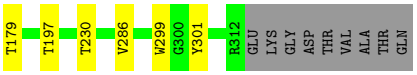


• Molecule 2: Isoaspartyl peptidase subunit beta

Chain BBB:

• Molecule 2: Isoaspartyl peptidase subunit beta

Chain DDD:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	62.69Å 70.84Å 149.58Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	74.79 – 1.45 74.79 – 1.45	Depositor EDS
% Data completeness (in resolution range)	99.6 (74.79-1.45) 99.6 (74.79-1.45)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.63 (at 1.45Å)	Xtriage
Refinement program	REFMAC 5.8.0267	Depositor
R, R_{free}	0.127 , 0.162 0.131 , 0.166	Depositor DCC
R_{free} test set	1068 reflections (0.91%)	wwPDB-VP
Wilson B-factor (Å^2)	17.0	Xtriage
Anisotropy	0.692	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 46.2	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.98	EDS
Total number of atoms	5029	wwPDB-VP
Average B, all atoms (Å^2)	27.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.10% 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: EDO, 1PE, NA, CL, CA, PEG

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	AAA	0.71	0/1204	0.77	0/1622
1	CCC	0.70	0/1211	0.79	0/1631
2	BBB	0.67	0/1009	0.79	0/1376
2	DDD	0.66	0/980	0.79	0/1337
All	All	0.69	0/4404	0.78	0/5966

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

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	AAA	1187	0	1197	3	0
1	CCC	1191	0	1207	4	0
2	BBB	981	0	960	8	0
2	DDD	957	0	937	6	0
3	AAA	7	0	10	0	0
3	BBB	7	0	10	0	0
3	CCC	7	0	10	0	0
3	DDD	7	0	10	1	0
4	AAA	1	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	CCC	1	0	0	0	0
5	AAA	2	0	0	0	0
5	CCC	1	0	0	0	0
5	DDD	1	0	0	0	0
6	BBB	4	0	6	1	0
7	CCC	16	0	22	0	0
8	CCC	1	0	0	0	0
8	DDD	1	0	0	0	0
9	AAA	189	0	0	0	0
9	BBB	111	0	0	3	0
9	CCC	218	0	0	1	0
9	DDD	139	0	0	2	0
All	All	5029	0	4369	16	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 2.

The worst 5 of 16 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:BBB:200[A]:TRP:CZ2	9:BBB:588:HOH:O	1.78	1.29
2:DDD:179[B]:THR:OG1	9:DDD:501:HOH:O	1.97	0.82
1:CCC:5:VAL:HG21	2:DDD:286[B]:VAL:HG21	1.77	0.66
1:CCC:138:ARG:NH1	9:CCC:301:HOH:O	2.28	0.65
2:BBB:200[A]:TRP:CH2	9:BBB:588:HOH:O	2.22	0.61

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	AAA	157/178 (88%)	155 (99%)	2 (1%)	0	100	100
1	CCC	157/178 (88%)	155 (99%)	2 (1%)	0	100	100
2	BBB	136/143 (95%)	131 (96%)	5 (4%)	0	100	100
2	DDD	133/143 (93%)	128 (96%)	5 (4%)	0	100	100
All	All	583/642 (91%)	569 (98%)	14 (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	AAA	122/136 (90%)	121 (99%)	1 (1%)	81	62
1	CCC	123/136 (90%)	123 (100%)	0	100	100
2	BBB	96/99 (97%)	96 (100%)	0	100	100
2	DDD	94/99 (95%)	93 (99%)	1 (1%)	73	48
All	All	435/470 (93%)	433 (100%)	2 (0%)	88	75

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	AAA	111	ARG
2	DDD	230	THR

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

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 14 ligands modelled in this entry, 8 are monoatomic - leaving 6 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	EDO	BBB	401	-	3,3,3	0.09	0	2,2,2	0.31	0
3	PEG	BBB	402	-	6,6,6	0.17	0	5,5,5	0.11	0
7	1PE	CCC	202	-	15,15,15	0.14	0	14,14,14	0.18	0
3	PEG	DDD	401	-	6,6,6	0.13	0	5,5,5	0.12	0
3	PEG	CCC	201	-	6,6,6	0.17	0	5,5,5	0.13	0
3	PEG	AAA	201	-	6,6,6	0.19	0	5,5,5	0.10	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	EDO	BBB	401	-	-	0/1/1/1	-
3	PEG	BBB	402	-	-	2/4/4/4	-
7	1PE	CCC	202	-	-	3/13/13/13	-
3	PEG	DDD	401	-	-	2/4/4/4	-
3	PEG	CCC	201	-	-	2/4/4/4	-
3	PEG	AAA	201	-	-	2/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 11 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	CCC	201	PEG	O1-C1-C2-O2
3	BBB	402	PEG	O1-C1-C2-O2
3	AAA	201	PEG	O1-C1-C2-O2
3	DDD	401	PEG	O2-C3-C4-O4
7	CCC	202	1PE	C12-C22-OH3-C23

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	BBB	401	EDO	1	0
3	DDD	401	PEG	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	AAA	158/178 (88%)	-0.74	1 (0%) 89 91	16, 22, 46, 74	0
1	CCC	157/178 (88%)	-0.75	1 (0%) 89 91	15, 19, 38, 68	0
2	BBB	135/143 (94%)	-0.76	0 100 100	16, 23, 39, 75	0
2	DDD	134/143 (93%)	-0.75	0 100 100	15, 22, 33, 49	0
All	All	584/642 (90%)	-0.75	2 (0%) 94 95	15, 22, 39, 75	0

All (2) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	AAA	2	GLY	5.1
1	CCC	159	GLU	2.4

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
6	EDO	BBB	401	4/4	0.84	0.14	69,72,79,79	0
3	PEG	AAA	201	7/7	0.90	0.09	51,58,66,74	0
3	PEG	DDD	401	7/7	0.92	0.10	37,46,50,58	0
3	PEG	BBB	402	7/7	0.94	0.12	47,50,59,60	0
3	PEG	CCC	201	7/7	0.94	0.07	58,64,68,71	0
7	1PE	CCC	202	16/16	0.94	0.11	44,51,67,68	0
5	CL	AAA	204	1/1	0.95	0.07	63,63,63,63	0
4	NA	AAA	202	1/1	1.00	0.03	19,19,19,19	0
5	CL	CCC	204	1/1	1.00	0.05	24,24,24,24	0
5	CL	DDD	402	1/1	1.00	0.03	29,29,29,29	0
4	NA	CCC	203	1/1	1.00	0.04	16,16,16,16	0
5	CL	AAA	203	1/1	1.00	0.04	37,37,37,37	0
8	CA	CCC	205	1/1	1.00	0.04	22,22,22,22	0
8	CA	DDD	403	1/1	1.00	0.07	30,30,30,30	0

6.5 Other polymers (i)

There are no such residues in this entry.