



wwPDB X-ray Structure Validation Summary Report ⓘ

Nov 18, 2022 – 03:28 am GMT

PDB ID : 7Q9H
Title : Peptide LLKAVAEKQ in complex with human cathepsin V C25A mutant
Authors : Loboda, J.; Sosnowski, P.; Tusar, L.; Vidmar, R.; Vizovisek, M.; Horvat, J.;
Kosec, G.; Impens, F.; Demol, H.; Turk, B.; Gevaert, K.; Turk, D.
Deposited on : 2021-11-12
Resolution : 1.40 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 : 1.8.4, CSD as541be (2020)
Xtriage (Phenix) : 1.13
EDS : 2.31.2
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac : 5.8.0267
CCP4 : 7.1.010 (Gargrove)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.31.2

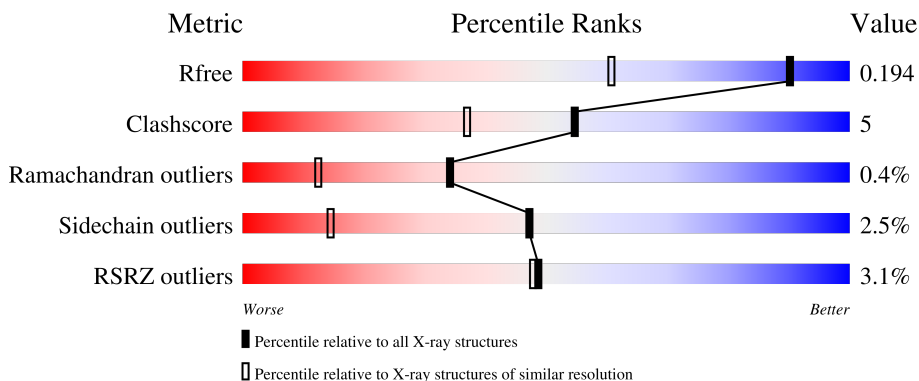
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.40 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	1714 (1.40-1.40)
Clashscore	141614	1812 (1.40-1.40)
Ramachandran outliers	138981	1763 (1.40-1.40)
Sidechain outliers	138945	1762 (1.40-1.40)
RSRZ outliers	127900	1674 (1.40-1.40)

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	AA	221	 91% 9%
1	BA	221	 2% 91% 9%
2	PAA	11	 27% 27% 9% 64%
2	PAC	11	 9% 27% 9% 55%
2	PB	11	 36% 64%

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	MPD	AA	302[B]	-	-	-	X

2 Entry composition [i](#)

There are 6 unique types of molecules in this entry. The entry contains 8708 atoms, of which 4516 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 Cathepsin L2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace	
			Total	C	H	N	O				S
1	AA	221	3476	1113	1713	302	338	10	1715	10	0
1	BA	221	3410	1095	1671	299	335	10	1673	6	0

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
AA	25	ALA	CYS	engineered mutation	UNP O60911
AA	108	GLN	ASN	engineered mutation	UNP O60911
AA	179	GLN	ASN	engineered mutation	UNP O60911
BA	246	ALA	CYS	engineered mutation	UNP O60911
BA	329	GLN	ASN	engineered mutation	UNP O60911
BA	400	GLN	ASN	engineered mutation	UNP O60911

- Molecule 2 is a protein called LLKAVAEKQ Peptide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	H	N	O			
2	PAA	4	66	20	38	4	4	38	4	0
2	PAC	5	73	22	38	6	7	38	5	0
2	PB	4	66	20	38	4	4	38	0	0

- Molecule 3 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	AA	3	Total Cl 3 3	0	1
3	BA	1	Total Cl 1 1	0	0

- Molecule 4 is (4S)-2-METHYL-2,4-PENTANEDIOL (three-letter code: MPD) (formula: $C_6H_{14}O_2$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	AA	1	Total C O 8 6 2	0	1
4	AA	1	Total C O 8 6 2	0	0
4	AA	1	Total C O 8 6 2	0	0
4	BA	1	Total C O 8 6 2	0	0
4	BA	1	Total C O 8 6 2	0	0
4	BA	1	Total C O 8 6 2	0	0
4	BA	1	Total C O 8 6 2	0	0

- Molecule 5 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	AA	1	Total C O 6 3 3	0	1
5	AA	1	Total C O 6 3 3	0	0
5	BA	1	Total C O 6 3 3	0	0
5	BA	1	Total C O 6 3 3	0	0
5	BA	1	Total C O 6 3 3	0	0

- Molecule 6 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	AA	262	Total H O 786 524 262	524	2
6	BA	238	Total H O 714 476 238	476	0
6	PAA	3	Total H O 9 6 3	6	0
6	PAC	2	Total H O 6 4 2	4	0
6	PB	4	Total H O 12 8 4	8	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: Cathepsin L2

Chain AA: 91% 9%



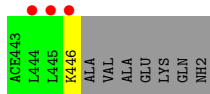
- Molecule 1: Cathepsin L2

Chain BA: 2% 91% 9%



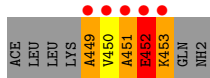
- Molecule 2: LLKAVAEKQ Peptide

Chain PAA: 27% 27% 9% 64%



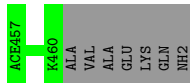
- Molecule 2: LLKAVAEKQ Peptide

Chain PAC: 9% 45% 27% 9% 55%



- Molecule 2: LLKAVAEKQ Peptide

Chain PB: 36% 64%



4 Data and refinement statistics

Property	Value	Source
Space group	P 43 21 2	Depositor
Cell constants a, b, c, α , β , γ	94.51Å 94.51Å 124.78Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	47.26 – 1.40 47.26 – 1.29	Depositor EDS
% Data completeness (in resolution range)	100.0 (47.26-1.40) 99.5 (47.26-1.29)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	0.95 (at 1.29Å)	Xtriage
Refinement program	MAIN	Depositor
R, R_{free}	0.182 , 0.209 0.188 , 0.194	Depositor DCC
R_{free} test set	2100 reflections (1.48%)	wwPDB-VP
Wilson B-factor (Å ²)	17.5	Xtriage
Anisotropy	0.102	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	(Not available) , (Not available)	EDS
L-test for twinning ²	$\langle L \rangle = 0.51$, $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.97	EDS
Total number of atoms	8708	wwPDB-VP
Average B, all atoms (Å ²)	11.0	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.*

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: MPD, GOL, ACE, CL

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	AA	1.02	2/1817 (0.1%)	1.00	0/2451
1	BA	1.00	2/1781 (0.1%)	0.98	0/2407
2	PAA	0.83	0/25	0.68	0/32
2	PAC	1.46	1/34 (2.9%)	1.49	0/44
2	PB	0.73	0/25	0.93	0/32
All	All	1.01	5/3682 (0.1%)	0.99	0/4966

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	BA	415	TRP	CB-CG	8.41	1.65	1.50
2	PAC	449[A]	ALA	N-CA	-6.79	1.32	1.46
1	AA	86	GLU	CG-CD	6.28	1.61	1.51
1	BA	405	TRP	CB-CG	-5.32	1.40	1.50
1	AA	193	GLU	CG-CD	5.27	1.59	1.51

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	AA	1763	1713	1699	14	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	BA	1739	1671	1661	10	0
2	PAA	28	38	38	2	0
2	PAC	35	38	37	7	0
2	PB	28	38	38	0	0
3	AA	3	0	0	0	0
3	BA	1	0	0	1	0
4	AA	24	0	41	6	0
4	BA	32	0	56	5	0
5	AA	12	0	11	2	0
5	BA	18	0	24	1	0
6	AA	262	524	0	6	1
6	BA	238	476	0	2	1
6	PAA	3	6	0	0	0
6	PAC	2	4	0	0	0
6	PB	4	8	0	0	0
All	All	4192	4516	3605	38	1

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 5.

The worst 5 of 38 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:PAA:446[A]:LYS:O	2:PAC:449[A]:ALA:N	1.71	1.23
4:AA:308:MPD:H12	6:AA:409:HOH:O	1.50	1.09
4:AA:308:MPD:C1	6:AA:409:HOH:O	2.07	1.00
4:BA:505:MPD:HM2	4:BA:505:MPD:H52	1.45	0.95
4:BA:505:MPD:HM2	4:BA:505:MPD:C5	2.15	0.75

All (1) 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)
6:AA:552:HOH:O	6:BA:612:HOH:O[3_554]	1.94	0.26

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	AA	229/221 (104%)	220 (96%)	9 (4%)	0	100	100
1	BA	225/221 (102%)	217 (96%)	8 (4%)	0	100	100
2	PAA	2/11 (18%)	2 (100%)	0	0	100	100
2	PAC	3/11 (27%)	1 (33%)	0	2 (67%)	0	0
2	PB	2/11 (18%)	2 (100%)	0	0	100	100
All	All	461/475 (97%)	442 (96%)	17 (4%)	2 (0%)	34	12

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	PAC	451[A]	ALA
2	PAC	452[A]	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	AA	188/178 (106%)	184 (98%)	4 (2%)	53	21
1	BA	184/178 (103%)	181 (98%)	3 (2%)	62	33
2	PAA	3/7 (43%)	3 (100%)	0	100	100
2	PAC	3/7 (43%)	1 (33%)	2 (67%)	0	0
2	PB	3/7 (43%)	3 (100%)	0	100	100
All	All	381/377 (101%)	372 (98%)	9 (2%)	47	16

5 of 9 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	PAC	452[A]	GLU
2	PAC	453[A]	LYS
1	AA	196	SER
1	BA	273	ASN
1	BA	310	TYR

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 16 ligands modelled in this entry, 4 are monoatomic - leaving 12 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	$\# Z > 2$
4	MPD	BA	503	-	7,7,7	1.04	1 (14%)	9,10,10	1.68	2 (22%)
5	GOL	BA	508	-	5,5,5	0.32	0	5,5,5	1.00	0
4	MPD	BA	506	-	7,7,7	0.24	0	9,10,10	0.32	0
5	GOL	AA	306	-	5,5,5	0.80	0	5,5,5	1.02	0
4	MPD	AA	307	-	7,7,7	0.44	0	9,10,10	1.61	1 (11%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
4	MPD	AA	302[B]	-	7,7,7	0.25	0	9,10,10	0.32	0
5	GOL	BA	501	-	5,5,5	0.46	0	5,5,5	0.76	0
4	MPD	BA	505	-	7,7,7	0.42	0	9,10,10	0.58	0
5	GOL	BA	507	-	5,5,5	0.61	0	5,5,5	0.82	0
5	GOL	AA	303[B]	-	5,5,5	0.14	0	5,5,5	0.32	0
4	MPD	BA	502	-	7,7,7	0.74	0	9,10,10	0.61	0
4	MPD	AA	308	-	7,7,7	0.51	0	9,10,10	0.57	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
4	MPD	BA	503	-	-	0/5/5/5	-
5	GOL	BA	508	-	-	0/4/4/4	-
4	MPD	BA	506	-	-	1/5/5/5	-
5	GOL	AA	306	-	-	1/4/4/4	-
4	MPD	AA	307	-	-	0/5/5/5	-
4	MPD	AA	302[B]	-	-	2/5/5/5	-
5	GOL	BA	501	-	-	0/4/4/4	-
4	MPD	BA	505	-	-	2/5/5/5	-
5	GOL	BA	507	-	-	2/4/4/4	-
5	GOL	AA	303[B]	-	-	0/4/4/4	-
4	MPD	BA	502	-	-	1/5/5/5	-
4	MPD	AA	308	-	-	1/5/5/5	-

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	BA	503	MPD	CM-C2	-2.18	1.45	1.52

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	AA	307	MPD	CM-C2-C1	-4.48	101.25	110.57
4	BA	503	MPD	CM-C2-C1	3.97	118.84	110.57
4	BA	503	MPD	O4-C4-C5	-2.39	99.01	109.38

There are no chirality outliers.

5 of 10 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	AA	302[B]	MPD	C2-C3-C4-O4
4	BA	505	MPD	C2-C3-C4-O4
4	AA	302[B]	MPD	C2-C3-C4-C5
4	BA	505	MPD	C2-C3-C4-C5
5	BA	507	GOL	O2-C2-C3-O3

There are no ring outliers.

8 monomers are involved in 14 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	BA	503	MPD	1	0
5	BA	508	GOL	1	0
5	AA	306	GOL	2	0
4	AA	307	MPD	2	0
4	AA	302[B]	MPD	1	0
4	BA	505	MPD	4	0
5	BA	507	GOL	1	0
4	AA	308	MPD	3	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 [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	AA	221/221 (100%)	-0.05	1 (0%) 91 89	9, 14, 26, 35	3 (1%)
1	BA	221/221 (100%)	0.20	5 (2%) 60 60	10, 17, 34, 68	2 (0%)
2	PAA	3/11 (27%)	3.79	3 (100%) 0 0	17, 17, 36, 40	3 (100%)
2	PAC	5/11 (45%)	4.83	5 (100%) 0 0	37, 50, 66, 67	5 (100%)
2	PB	3/11 (27%)	0.51	0 100 100	19, 19, 31, 32	0
All	All	453/475 (95%)	0.16	14 (3%) 49 48	9, 15, 33, 68	13 (2%)

The worst 5 of 14 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	PAC	451[A]	ALA	8.0
2	PAC	453[A]	LYS	6.1
2	PAC	452[A]	GLU	4.5
2	PAA	444[A]	LEU	4.3
2	PAA	445[A]	LEU	4.3

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, 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(Å ²)	Q<0.9
4	MPD	BA	506	8/8	0.39	0.40	76,79,86,88	0
4	MPD	BA	505	8/8	0.62	0.30	74,83,94,98	0
4	MPD	AA	302[B]	8/8	0.64	0.55	66,71,76,76	8
5	GOL	BA	507	6/6	0.68	0.17	64,72,83,87	0
4	MPD	BA	503	8/8	0.71	0.14	23,26,29,29	0
5	GOL	AA	303[B]	6/6	0.74	0.20	56,62,63,65	6
4	MPD	AA	308	8/8	0.75	0.28	64,75,76,77	0
3	CL	AA	301[B]	1/1	0.80	0.33	27,27,27,27	1
5	GOL	BA	508	6/6	0.81	0.17	35,72,81,84	0
5	GOL	AA	306	6/6	0.82	0.15	19,24,34,34	0
4	MPD	BA	502	8/8	0.83	0.11	27,31,37,39	0
5	GOL	BA	501	6/6	0.87	0.09	24,25,28,31	0
4	MPD	AA	307	8/8	0.89	0.13	38,46,77,88	0
3	CL	AA	304	1/1	0.94	0.07	20,20,20,20	0
3	CL	AA	305	1/1	0.99	0.04	17,17,17,17	0
3	CL	BA	504	1/1	0.99	0.04	19,19,19,19	0

6.5 Other polymers [i](#)

There are no such residues in this entry.