



wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 17, 2016 – 07:36 PM EDT

PDB ID : 5KJ7
Title : Structure of the Ca²⁺-bound synaptotagmin-1 SNARE complex (long unit cell form) - from XFEL diffraction
Authors : Lyubimov, A.Y.; Uervirojnangkoorn, M.; Zhou, Q.; Zhao, M.; Sauter, N.K.; Brewster, A.S.; Weis, W.I.; Brunger, A.T.
Deposited on : 2016-06-17
Resolution : 3.50 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

<http://wwpdb.org/validation/2016/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : unknown
Xtriage (Phenix) : 1.9-1692
EDS : rb-20027939
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac : 5.8.0135
CCP4 : 6.5.0
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : rb-20027939

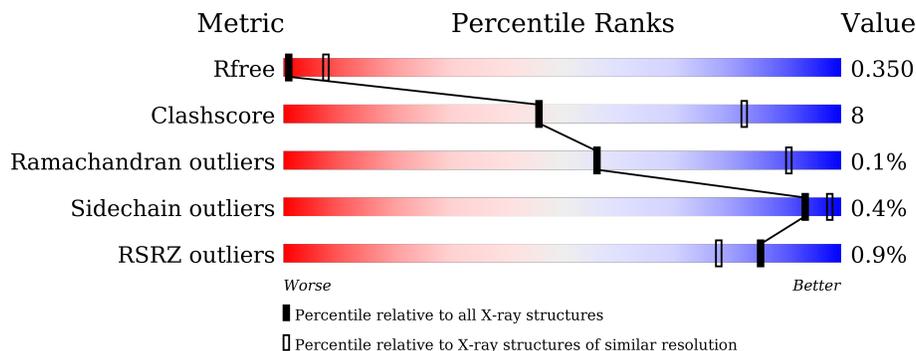
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.50 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}	91344	1051 (3.60-3.40)
Clashscore	102246	1157 (3.60-3.40)
Ramachandran outliers	100387	1120 (3.60-3.40)
Sidechain outliers	100360	1121 (3.60-3.40)
RSRZ outliers	91569	1058 (3.60-3.40)

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 on the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria. 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	63	81% 19%
1	G	63	81% 19%
2	B	66	76% 24%
2	H	66	74% 24% .
3	C	75	81% 17% .
3	I	75	79% 20% .

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
4	D	64	 91% 9%
4	J	64	 86% 13%
5	E	279	 87% 13%
5	F	279	 78% 20%
5	K	279	 76% 24%

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
6	CA	F	504	-	-	-	X
6	CA	K	503	-	-	-	X

2 Entry composition [i](#)

There are 7 unique types of molecules in this entry. The entry contains 10578 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 Vesicle-associated membrane protein 3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	63	502	305	95	101	1	0	0	0
1	G	63	492	299	91	101	1	0	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	37	ALA	ASN	conflict	UNP P63025
G	37	ALA	ASN	conflict	UNP P63025

- Molecule 2 is a protein called Syntaxin-1A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	66	515	315	87	108	5	0	0	0
2	H	65	510	314	84	107	5	0	0	0

- Molecule 3 is a protein called Synaptosomal-associated protein 25.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	C	74	579	343	102	129	5	0	0	0
3	I	74	549	325	98	121	5	0	0	0

- Molecule 4 is a protein called Synaptosomal-associated protein 25.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
4	D	64	498	291	96	106	5	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	J	63	Total	C	N	O	S	0	0	0
			496	292	94	105	5			

- Molecule 5 is a protein called Synaptotagmin-1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	E	279	Total	C	N	O	S	0	0	1
			2163	1391	363	402	7			
5	F	275	Total	C	N	O	S	0	0	1
			2076	1325	350	394	7			
5	K	279	Total	C	N	O	S	0	0	1
			2162	1392	360	403	7			

- Molecule 6 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	G	1	Total	Ca	0	0
			1	1		
6	D	1	Total	Ca	0	0
			1	1		
6	K	4	Total	Ca	0	0
			4	4		
6	E	4	Total	Ca	0	0
			4	4		
6	C	3	Total	Ca	0	0
			3	3		
6	A	2	Total	Ca	0	0
			2	2		
6	F	4	Total	Ca	0	0
			4	4		

- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	1	Total	O	0	0
			1	1		
7	B	1	Total	O	0	0
			1	1		
7	D	1	Total	O	0	0
			1	1		
7	E	4	Total	O	0	0
			4	4		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	F	4	Total O 4 4	0	0
7	I	1	Total O 1 1	0	0
7	J	2	Total O 2 2	0	0
7	K	3	Total O 3 3	0	0

3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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: Vesicle-associated membrane protein 3

Chain A:  81% 19%



- Molecule 1: Vesicle-associated membrane protein 3

Chain G:  81% 19%



- Molecule 2: Syntaxin-1A

Chain B:  76% 24%



- Molecule 2: Syntaxin-1A

Chain H:  74% 24%



- Molecule 3: Synaptosomal-associated protein 25

Chain C:  81% 17%

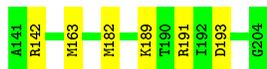


- Molecule 3: Synaptosomal-associated protein 25

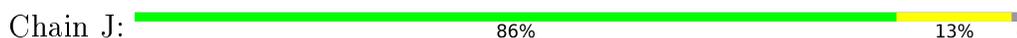
Chain I:  79% 20%



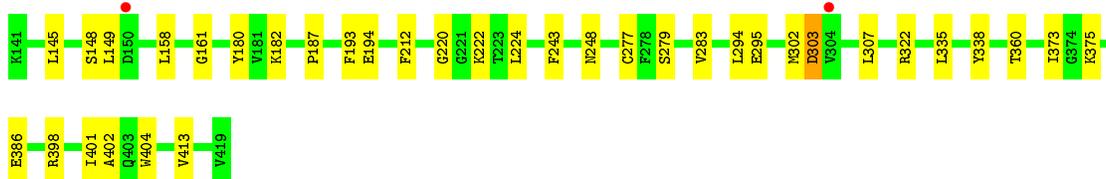
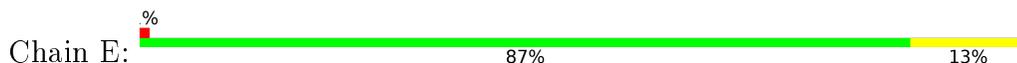
- Molecule 4: Synaptosomal-associated protein 25



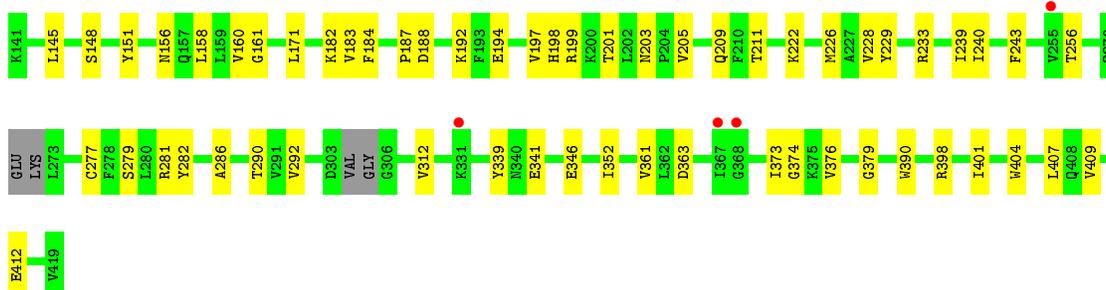
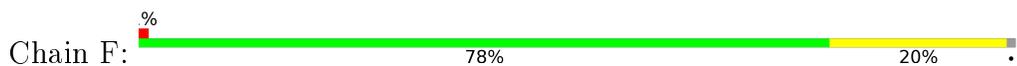
- Molecule 4: Synaptosomal-associated protein 25



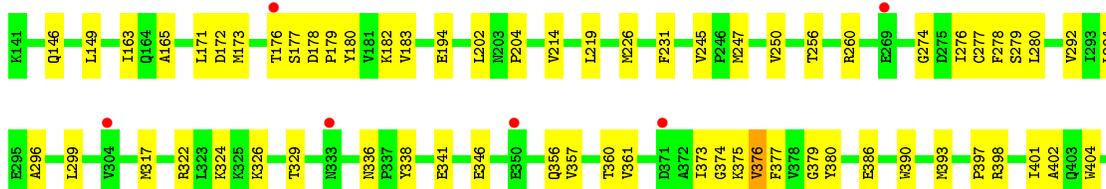
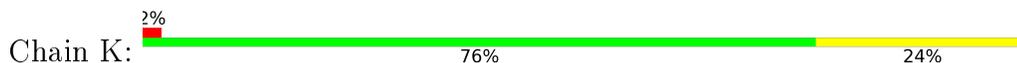
- Molecule 5: Synaptotagmin-1



- Molecule 5: Synaptotagmin-1



- Molecule 5: Synaptotagmin-1



L407
Q408
V409
E410
E411
E412
Y419

4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	69.39Å 170.68Å 291.01Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	19.96 – 3.50 19.96 – 3.49	Depositor EDS
% Data completeness (in resolution range)	91.3 (19.96-3.50) 82.3 (19.96-3.49)	Depositor EDS
R_{merge}	0.49	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.18 (at 3.52Å)	Xtrriage
Refinement program	PHENIX	Depositor
R, R_{free}	0.292 , 0.329 0.313 , 0.350	Depositor DCC
R_{free} test set	1851 reflections (5.06%)	DCC
Wilson B-factor (Å ²)	50.5	Xtrriage
Anisotropy	0.499	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.21 , 5.6	EDS
L-test for twinning ²	$\langle L \rangle = 0.31$, $\langle L^2 \rangle = 0.14$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.83	EDS
Total number of atoms	10578	wwPDB-VP
Average B, all atoms (Å ²)	105.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 9.22% 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: CA

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	0.23	0/505	0.33	0/675
1	G	0.24	0/495	0.36	0/667
2	B	0.23	0/519	0.36	0/697
2	H	0.24	0/515	0.37	0/693
3	C	0.23	0/579	0.34	0/774
3	I	0.23	0/549	0.35	0/730
4	D	0.23	0/498	0.35	0/662
4	J	0.23	0/496	0.35	0/661
5	E	0.25	0/2211	0.44	0/2997
5	F	0.24	0/2123	0.44	0/2878
5	K	0.24	0/2211	0.45	0/2997
All	All	0.24	0/10701	0.41	0/14431

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 [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	502	0	498	11	0
1	G	492	0	472	10	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	B	515	0	476	17	0
2	H	510	0	479	17	0
3	C	579	0	538	11	0
3	I	549	0	506	15	0
4	D	498	0	476	6	0
4	J	496	0	480	7	0
5	E	2163	0	2110	27	0
5	F	2076	0	1937	34	1
5	K	2162	0	2104	43	0
6	A	2	0	0	0	0
6	C	3	0	0	0	0
6	D	1	0	0	0	0
6	E	4	0	0	0	0
6	F	4	0	0	0	0
6	G	1	0	0	0	0
6	K	4	0	0	0	0
7	A	1	0	0	0	0
7	B	1	0	0	0	0
7	D	1	0	0	0	0
7	E	4	0	0	1	0
7	F	4	0	0	0	0
7	I	1	0	0	0	0
7	J	2	0	0	0	0
7	K	3	0	0	0	0
All	All	10578	0	10076	159	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 8.

The worst 5 of 159 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)
3:C:40:LYS:NZ	5:E:295:GLU:OE2	1.93	1.02
1:A:85:LYS:NZ	2:H:242:ASP:OD2	2.01	0.93
3:C:23:ASP:OD1	4:D:142:ARG:NH1	2.03	0.92
4:D:189:LYS:NZ	4:D:193:ASP:OD2	2.02	0.91
5:F:188:ASP:O	5:F:192:LYS:NZ	2.07	0.87

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)
5:F:233:ARG:N	5:F:346:GLU:OE2[4_445]	2.07	0.13

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	61/63 (97%)	61 (100%)	0	0	100	100
1	G	61/63 (97%)	61 (100%)	0	0	100	100
2	B	64/66 (97%)	63 (98%)	1 (2%)	0	100	100
2	H	63/66 (96%)	63 (100%)	0	0	100	100
3	C	72/75 (96%)	71 (99%)	1 (1%)	0	100	100
3	I	72/75 (96%)	72 (100%)	0	0	100	100
4	D	62/64 (97%)	61 (98%)	1 (2%)	0	100	100
4	J	61/64 (95%)	61 (100%)	0	0	100	100
5	E	277/279 (99%)	260 (94%)	17 (6%)	0	100	100
5	F	269/279 (96%)	255 (95%)	14 (5%)	0	100	100
5	K	277/279 (99%)	262 (95%)	14 (5%)	1 (0%)	39	81
All	All	1339/1373 (98%)	1290 (96%)	48 (4%)	1 (0%)	56	90

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
5	K	173	MET

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	53/54 (98%)	53 (100%)	0	100	100
1	G	51/54 (94%)	51 (100%)	0	100	100
2	B	55/60 (92%)	55 (100%)	0	100	100
2	H	56/60 (93%)	56 (100%)	0	100	100
3	C	62/67 (92%)	62 (100%)	0	100	100
3	I	57/67 (85%)	57 (100%)	0	100	100
4	D	53/55 (96%)	53 (100%)	0	100	100
4	J	54/55 (98%)	54 (100%)	0	100	100
5	E	229/249 (92%)	228 (100%)	1 (0%)	93	98
5	F	213/249 (86%)	211 (99%)	2 (1%)	84	94
5	K	229/249 (92%)	228 (100%)	1 (0%)	93	98
All	All	1112/1219 (91%)	1108 (100%)	4 (0%)	93	98

All (4) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
5	E	303	ASP
5	F	197	VAL
5	F	199	ARG
5	K	376	VAL

Some 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 carbohydrates in this entry.

5.6 Ligand geometry [i](#)

Of 19 ligands modelled in this entry, 19 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	63/63 (100%)	-0.39	0 100 100	48, 83, 116, 136	0
1	G	63/63 (100%)	-0.42	0 100 100	87, 120, 141, 149	0
2	B	66/66 (100%)	-0.32	0 100 100	46, 84, 129, 149	0
2	H	65/66 (98%)	-0.26	0 100 100	86, 114, 139, 158	0
3	C	74/75 (98%)	-0.26	0 100 100	55, 91, 145, 163	0
3	I	74/75 (98%)	-0.36	0 100 100	74, 112, 144, 153	0
4	D	64/64 (100%)	-0.17	0 100 100	54, 85, 126, 151	0
4	J	63/64 (98%)	-0.30	0 100 100	81, 116, 141, 154	0
5	E	279/279 (100%)	-0.09	2 (0%) 89 82	46, 106, 139, 171	0
5	F	275/279 (98%)	-0.13	4 (1%) 76 67	59, 105, 154, 181	0
5	K	279/279 (100%)	-0.10	6 (2%) 65 55	67, 109, 145, 188	0
All	All	1365/1373 (99%)	-0.19	12 (0%) 85 78	46, 106, 145, 188	0

The worst 5 of 12 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
5	E	304	VAL	3.4
5	F	368	GLY	3.1
5	K	371	ASP	3.0
5	K	304	VAL	2.8
5	K	269	GLU	2.8

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 carbohydrates 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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(Å ²)	Q<0.9
6	CA	F	504	1/1	0.92	0.51	4.34	66,66,66,66	0
6	CA	K	503	1/1	0.88	0.45	3.41	120,120,120,120	0
6	CA	E	503	1/1	0.89	0.31	-0.13	78,78,78,78	0
6	CA	E	502	1/1	0.60	0.20	-0.80	107,107,107,107	0
6	CA	F	501	1/1	0.83	0.21	-0.91	37,37,37,37	0
6	CA	E	504	1/1	0.91	0.10	-1.01	80,80,80,80	0
6	CA	K	504	1/1	0.72	0.17	-1.50	54,54,54,54	0
6	CA	K	502	1/1	0.96	0.14	-1.51	107,107,107,107	0
6	CA	F	503	1/1	0.80	0.13	-1.67	145,145,145,145	0
6	CA	K	501	1/1	0.65	0.15	-1.93	99,99,99,99	0
6	CA	F	502	1/1	0.95	0.12	-3.21	46,46,46,46	0
6	CA	C	101	1/1	0.89	0.10	-	66,66,66,66	0
6	CA	A	102	1/1	0.70	0.27	-	81,81,81,81	0
6	CA	G	101	1/1	0.74	0.37	-	107,107,107,107	0
6	CA	A	101	1/1	0.75	0.90	-	99,99,99,99	0
6	CA	E	501	1/1	0.69	0.14	-	83,83,83,83	0
6	CA	C	103	1/1	0.90	0.24	-	85,85,85,85	0
6	CA	C	102	1/1	0.77	0.23	-	95,95,95,95	0
6	CA	D	301	1/1	0.83	0.35	-	102,102,102,102	0

6.5 Other polymers [i](#)

There are no such residues in this entry.