



Full wwPDB X-ray Structure Validation Report ⓘ

Mar 20, 2026 – 04:43 PM UTC

PDB ID : 5VBA / pdb_00005vba
Title : Structure of EspG1 chaperone from the type VII (ESX-1) secretion system determined with the assistance of N-terminal T4 lysozyme fusion
Authors : Korotkov, K.V.
Deposited on : 2017-03-29
Resolution : 2.27 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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	:	4-5-2 with Phenix2.0
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

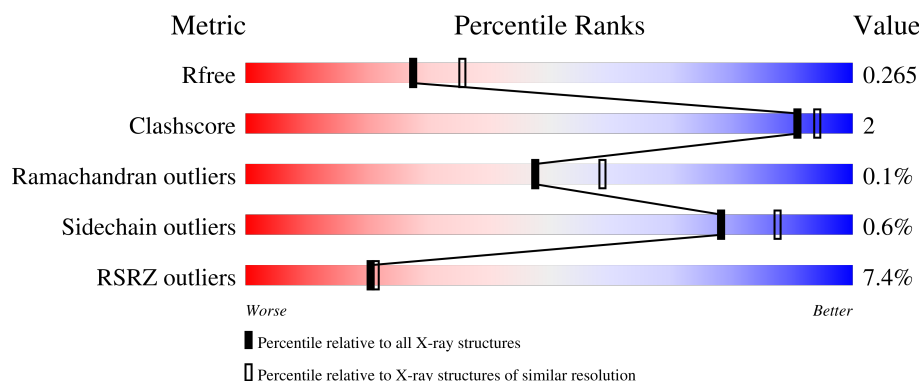
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.27 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}	180053	9078 (2.30-2.26)
Clashscore	190562	9802 (2.30-2.26)
Ramachandran outliers	187476	9690 (2.30-2.26)
Sidechain outliers	187428	9691 (2.30-2.26)
RSRZ outliers	180081	9085 (2.30-2.26)

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	437	8% 88% 5% 7%
1	B	437	6% 90% • 7%

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
2	CL	B	1205	-	-	X	-

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 12710 atoms, of which 6356 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 Lysozyme, ESX-1 secretion-associated protein EspG1 chimera.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	408	Total	C	H	N	O	S	0	0	0
			6352	1996	3193	568	586	9			
1	B	406	Total	C	H	N	O	S	0	0	0
			6292	1976	3163	563	581	9			

There are 52 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	982	MET	-	expression tag	UNP D9IEF7
A	983	GLY	-	expression tag	UNP D9IEF7
A	984	SER	-	expression tag	UNP D9IEF7
A	985	SER	-	expression tag	UNP D9IEF7
A	986	HIS	-	expression tag	UNP D9IEF7
A	987	HIS	-	expression tag	UNP D9IEF7
A	988	HIS	-	expression tag	UNP D9IEF7
A	989	HIS	-	expression tag	UNP D9IEF7
A	990	HIS	-	expression tag	UNP D9IEF7
A	991	HIS	-	expression tag	UNP D9IEF7
A	992	SER	-	expression tag	UNP D9IEF7
A	993	SER	-	expression tag	UNP D9IEF7
A	994	GLY	-	expression tag	UNP D9IEF7
A	995	GLU	-	expression tag	UNP D9IEF7
A	996	ASN	-	expression tag	UNP D9IEF7
A	997	LEU	-	expression tag	UNP D9IEF7
A	998	TYR	-	expression tag	UNP D9IEF7
A	999	PHE	-	expression tag	UNP D9IEF7
A	1000	GLN	-	expression tag	UNP D9IEF7
A	1001	GLY	-	expression tag	UNP D9IEF7
A	1054	THR	CYS	engineered mutation	UNP D9IEF7
A	1097	ALA	CYS	engineered mutation	UNP D9IEF7
A	1162	ALA	LYS	engineered mutation	UNP D9IEF7
A	16	MET	VAL	engineered mutation	UNP X7YCN8
A	114	ALA	CYS	engineered mutation	UNP X7YCN8

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	170	ALA	CYS	engineered mutation	UNP X7YCN8
B	982	MET	-	expression tag	UNP D9IEF7
B	983	GLY	-	expression tag	UNP D9IEF7
B	984	SER	-	expression tag	UNP D9IEF7
B	985	SER	-	expression tag	UNP D9IEF7
B	986	HIS	-	expression tag	UNP D9IEF7
B	987	HIS	-	expression tag	UNP D9IEF7
B	988	HIS	-	expression tag	UNP D9IEF7
B	989	HIS	-	expression tag	UNP D9IEF7
B	990	HIS	-	expression tag	UNP D9IEF7
B	991	HIS	-	expression tag	UNP D9IEF7
B	992	SER	-	expression tag	UNP D9IEF7
B	993	SER	-	expression tag	UNP D9IEF7
B	994	GLY	-	expression tag	UNP D9IEF7
B	995	GLU	-	expression tag	UNP D9IEF7
B	996	ASN	-	expression tag	UNP D9IEF7
B	997	LEU	-	expression tag	UNP D9IEF7
B	998	TYR	-	expression tag	UNP D9IEF7
B	999	PHE	-	expression tag	UNP D9IEF7
B	1000	GLN	-	expression tag	UNP D9IEF7
B	1001	GLY	-	expression tag	UNP D9IEF7
B	1054	THR	CYS	engineered mutation	UNP D9IEF7
B	1097	ALA	CYS	engineered mutation	UNP D9IEF7
B	1162	ALA	LYS	engineered mutation	UNP D9IEF7
B	16	MET	VAL	engineered mutation	UNP X7YCN8
B	114	ALA	CYS	engineered mutation	UNP X7YCN8
B	170	ALA	CYS	engineered mutation	UNP X7YCN8

- Molecule 2 is CHLORIDE ION (CCD ID: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	5	Total Cl 5 5	0	0
2	B	5	Total Cl 5 5	0	0

- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	27	Total O 27 27	0	0

Continued on next page...

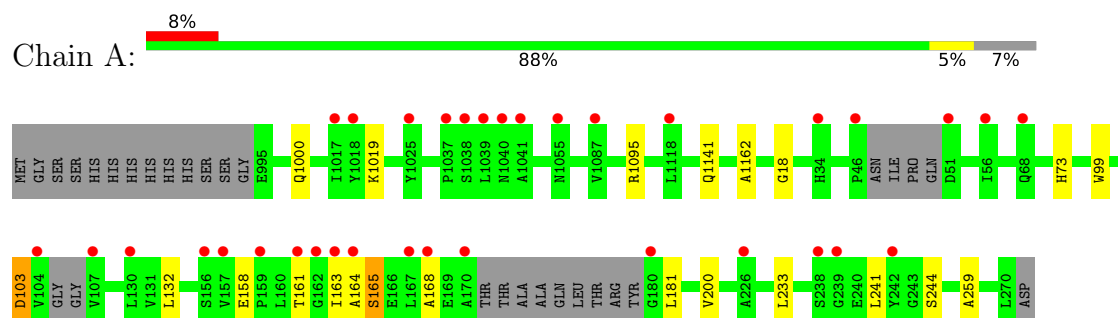
Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	29	Total	O	0	0
			29	29		

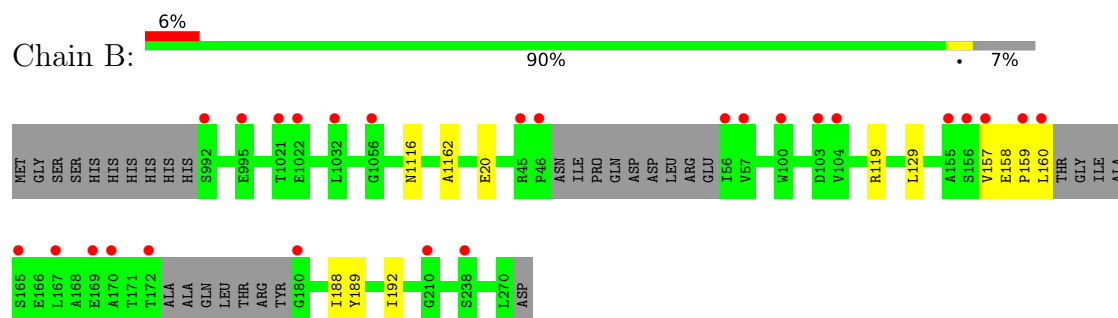
3 Residue-property plots

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: Lysozyme, ESX-1 secretion-associated protein EspG1 chimera



- Molecule 1: Lysozyme, ESX-1 secretion-associated protein EspG1 chimera



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	64.14Å 81.69Å 160.02Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	59.54 – 2.27 59.54 – 2.27	Depositor EDS
% Data completeness (in resolution range)	99.2 (59.54-2.27) 99.8 (59.54-2.27)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.84 (at 2.27Å)	Xtriage
Refinement program	PHENIX DEV_2722	Depositor
R, R_{free}	0.214 , 0.251 0.229 , 0.265	Depositor DCC
R_{free} test set	1882 reflections (4.74%)	wwPDB-VP
Wilson B-factor (Å ²)	48.4	Xtriage
Anisotropy	0.170	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.39 , 34.1	EDS
L-test for twinning ²	$\langle L \rangle = 0.45$, $\langle L^2 \rangle = 0.28$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	12710	wwPDB-VP
Average B, all atoms (Å ²)	67.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.91% 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: 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	A	0.22	1/3216 (0.0%)	0.40	1/4373 (0.0%)
1	B	0.24	1/3186 (0.0%)	0.39	0/4332
All	All	0.23	2/6402 (0.0%)	0.40	1/8705 (0.0%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	1162	ALA	C-N	7.72	1.43	1.33
1	A	1162	ALA	C-N	7.56	1.43	1.33

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	1162	ALA	O-C-N	6.12	128.70	122.09

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	3159	3193	3192	13	0
1	B	3129	3163	3161	8	0
2	A	5	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	B	5	0	0	3	0
3	A	27	0	0	1	0
3	B	29	0	0	1	0
All	All	6354	6356	6353	22	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.

All (22) 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:B:1205:CL:CL	3:B:1321:HOH:O	2.49	0.67
1:A:1019:LYS:O	1:A:73:HIS:NE2	2.33	0.57
1:B:157:VAL:O	1:B:159:PRO:HD3	2.06	0.55
1:A:163:ILE:HG21	1:A:168:ALA:HB2	1.91	0.53
1:B:159:PRO:HB2	1:B:189:TYR:CE2	2.46	0.51
1:B:159:PRO:HB2	1:B:189:TYR:HE2	1.76	0.50
1:B:1116:ASN:ND2	2:B:1205:CL:CL	2.77	0.50
1:B:158:GLU:OE1	1:B:158:GLU:N	2.45	0.50
1:A:1000:GLN:NE2	2:B:1201:CL:CL	2.84	0.47
1:A:103:ASP:N	1:A:103:ASP:OD1	2.49	0.45
1:A:158:GLU:OE1	1:A:158:GLU:N	2.50	0.44
1:A:1095:ARG:NH1	3:A:1302:HOH:O	2.43	0.43
1:A:233:LEU:HB3	1:A:244:SER:OG	2.19	0.43
1:B:159:PRO:C	1:B:160:LEU:HD12	2.43	0.43
1:A:164:ALA:O	1:A:165:SER:C	2.61	0.42
1:A:165:SER:OG	1:A:241:LEU:O	2.16	0.41
1:A:161:THR:HA	1:A:244:SER:HA	2.02	0.41
1:A:18:GLY:HA2	1:A:132:LEU:O	2.20	0.41
1:B:20:GLU:OE2	1:B:129:LEU:HD21	2.21	0.41
1:B:188:ILE:O	1:B:192:ILE:HG12	2.21	0.41
1:A:233:LEU:HG	1:A:259:ALA:HB3	2.03	0.41
1:A:99:TRP:CD1	1:A:200:VAL:HG22	2.56	0.40

There are no symmetry-related clashes.

5.3 Torsion angles

5.3.1 Protein backbone

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	400/437 (92%)	391 (98%)	8 (2%)	1 (0%)	36	44
1	B	398/437 (91%)	389 (98%)	9 (2%)	0	100	100
All	All	798/874 (91%)	780 (98%)	17 (2%)	1 (0%)	48	59

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	165	SER

5.3.2 Protein sidechains

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	334/357 (94%)	331 (99%)	3 (1%)	70	82
1	B	331/357 (93%)	330 (100%)	1 (0%)	86	92
All	All	665/714 (93%)	661 (99%)	4 (1%)	78	87

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

Mol	Chain	Res	Type
1	A	1141	GLN
1	A	103	ASP
1	A	181	LEU
1	B	119	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (6) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	1132	ASN
1	A	68	GLN
1	A	195	ASN
1	A	215	HIS
1	B	996	ASN
1	B	215	HIS

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 oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 10 ligands modelled in this entry, 10 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 ⓘ

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, 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	408/437 (93%)	0.73	34 (8%) 17 18	37, 69, 113, 158	0
1	B	406/437 (92%)	0.52	26 (6%) 25 26	35, 59, 104, 224	0
All	All	814/874 (93%)	0.63	60 (7%) 20 21	35, 64, 108, 224	0

All (60) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	995	GLU	5.4
1	A	163	ILE	5.1
1	A	46	PRO	4.7
1	B	157	VAL	4.7
1	A	162	GLY	4.5
1	B	159	PRO	4.5
1	B	46	PRO	4.4
1	A	107	VAL	4.1
1	B	172	THR	4.1
1	B	160	LEU	4.0
1	B	56	ILE	3.9
1	A	1037	PRO	3.7
1	A	1039	LEU	3.5
1	A	164	ALA	3.5
1	A	167	LEU	3.5
1	A	170	ALA	3.4
1	A	168	ALA	3.1
1	A	161	THR	3.1
1	B	180	GLY	3.1
1	B	238	SER	3.0
1	A	1087	VAL	3.0
1	A	104	VAL	3.0
1	A	1055	ASN	3.0
1	A	34	HIS	2.9

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	1038	SER	2.9
1	B	103	ASP	2.9
1	B	57	VAL	2.9
1	A	159	PRO	2.7
1	A	242	TYR	2.6
1	A	238	SER	2.6
1	B	992	SER	2.6
1	B	165	SER	2.6
1	A	1041	ALA	2.6
1	A	157	VAL	2.6
1	A	239	GLY	2.6
1	B	100	TRP	2.5
1	B	169	GLU	2.5
1	A	1017	ILE	2.5
1	A	156	SER	2.4
1	A	56	ILE	2.4
1	A	130	LEU	2.3
1	B	167	LEU	2.2
1	A	1040	ASN	2.2
1	B	1032	LEU	2.2
1	A	51	ASP	2.2
1	B	45	ARG	2.2
1	A	226	ALA	2.2
1	B	104	VAL	2.1
1	A	1018	TYR	2.1
1	A	1025	TYR	2.1
1	A	1118	LEU	2.1
1	B	1056	GLY	2.1
1	B	210	GLY	2.1
1	B	155	ALA	2.1
1	B	170	ALA	2.1
1	A	180	GLY	2.1
1	B	1021	THR	2.0
1	B	1022	GLU	2.0
1	A	68	GLN	2.0
1	B	156	SER	2.0

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 [i](#)

There are no oligosaccharides 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($\text{\AA}^2$)	Q<0.9
2	CL	A	1201	1/1	0.87	0.12	82,82,82,82	0
2	CL	B	1205	1/1	0.91	0.08	83,83,83,83	0
2	CL	B	1204	1/1	0.93	0.09	64,64,64,64	0
2	CL	A	1204	1/1	0.93	0.10	76,76,76,76	0
2	CL	B	1202	1/1	0.94	0.10	55,55,55,55	0
2	CL	A	1202	1/1	0.95	0.08	64,64,64,64	0
2	CL	A	1205	1/1	0.96	0.07	71,71,71,71	0
2	CL	A	1203	1/1	0.97	0.07	64,64,64,64	0
2	CL	B	1201	1/1	0.98	0.19	52,52,52,52	0
2	CL	B	1203	1/1	0.98	0.04	64,64,64,64	0

6.5 Other polymers [i](#)

There are no such residues in this entry.