



# Full wwPDB X-ray Structure Validation Report ⓘ

Mar 9, 2026 – 06:03 AM UTC

PDB ID : 9P55 / pdb\_00009p55  
Title : Structure of DNA-free long form XPD from *Thermoplasma acidophilum*  
Authors : Bravo, M.; Fan, L.  
Deposited on : 2025-06-17  
Resolution : 2.13 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](mailto: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  
Mogul : 2022.3.0, CSD as543be (2022)  
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

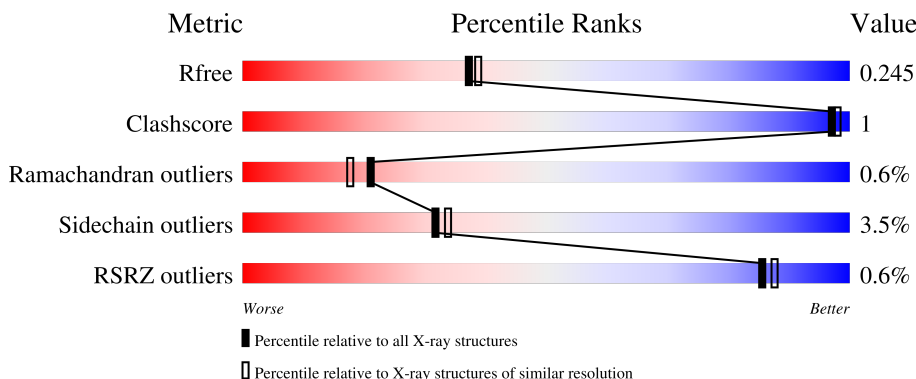
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*


The reported resolution of this entry is 2.13 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	8290 (2.14-2.10)
Clashscore	190562	8817 (2.14-2.10)
Ramachandran outliers	187476	8738 (2.14-2.10)
Sidechain outliers	187428	8739 (2.14-2.10)
RSRZ outliers	180081	8294 (2.14-2.10)

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  $\geq 3$ , 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	642	

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 10029 atoms, of which 4880 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 ATP-dependent DNA helicase XPD.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
1	A	622	9920	3202	4880	853	955	30	0	0	0

There are 44 discrepancies between the modelled and reference sequences:

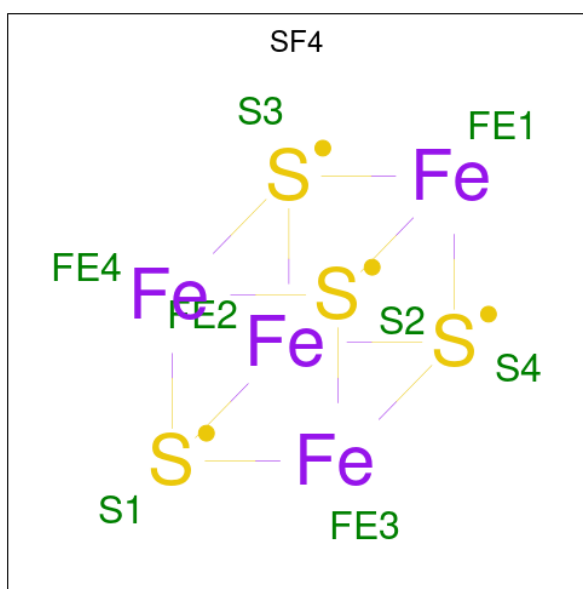
Chain	Residue	Modelled	Actual	Comment	Reference
A	-24	HIS	-	expression tag	UNP Q9HM14
A	-23	HIS	-	expression tag	UNP Q9HM14
A	-22	HIS	-	expression tag	UNP Q9HM14
A	-21	HIS	-	expression tag	UNP Q9HM14
A	-20	HIS	-	expression tag	UNP Q9HM14
A	-19	HIS	-	expression tag	UNP Q9HM14
A	-18	SER	-	expression tag	UNP Q9HM14
A	-17	SER	-	expression tag	UNP Q9HM14
A	-16	GLY	-	expression tag	UNP Q9HM14
A	-15	LEU	-	expression tag	UNP Q9HM14
A	-14	GLU	-	expression tag	UNP Q9HM14
A	-13	VAL	-	expression tag	UNP Q9HM14
A	-12	LEU	-	expression tag	UNP Q9HM14
A	-11	PHE	-	expression tag	UNP Q9HM14
A	-10	GLN	-	expression tag	UNP Q9HM14
A	-9	GLY	-	expression tag	UNP Q9HM14
A	-8	PRO	-	expression tag	UNP Q9HM14
A	-7	HIS	-	expression tag	UNP Q9HM14
A	-6	ARG	-	expression tag	UNP Q9HM14
A	-5	GLY	-	expression tag	UNP Q9HM14
A	-4	SER	-	expression tag	UNP Q9HM14
A	-3	GLU	-	expression tag	UNP Q9HM14
A	-2	PHE	-	expression tag	UNP Q9HM14
A	-1	GLU	-	expression tag	UNP Q9HM14
A	0	LEU	-	expression tag	UNP Q9HM14
A	1	MET	-	expression tag	UNP Q9HM14
A	2	TYR	-	expression tag	UNP Q9HM14

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	3	GLU	-	expression tag	UNP Q9HM14
A	4	ASN	-	expression tag	UNP Q9HM14
A	5	ARG	-	expression tag	UNP Q9HM14
A	6	GLN	-	expression tag	UNP Q9HM14
A	7	TYR	-	expression tag	UNP Q9HM14
A	8	GLN	-	expression tag	UNP Q9HM14
A	9	VAL	-	expression tag	UNP Q9HM14
A	10	GLU	-	expression tag	UNP Q9HM14
A	11	ALA	-	expression tag	UNP Q9HM14
A	12	ILE	-	expression tag	UNP Q9HM14
A	13	ASP	-	expression tag	UNP Q9HM14
A	14	PHE	-	expression tag	UNP Q9HM14
A	15	LEU	-	expression tag	UNP Q9HM14
A	16	ARG	-	expression tag	UNP Q9HM14
A	17	SER	-	expression tag	UNP Q9HM14
A	18	SER	-	expression tag	UNP Q9HM14
A	19	LEU	-	expression tag	UNP Q9HM14

- Molecule 2 is IRON/SULFUR CLUSTER (CCD ID: SF4) (formula:  $\text{Fe}_4\text{S}_4$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	Fe	S	0	0
			8	4	4		

- Molecule 3 is CHLORIDE ION (CCD ID: CL) (formula:  $\text{Cl}$ ).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	13	Total	Cl	0	0
			13	13		

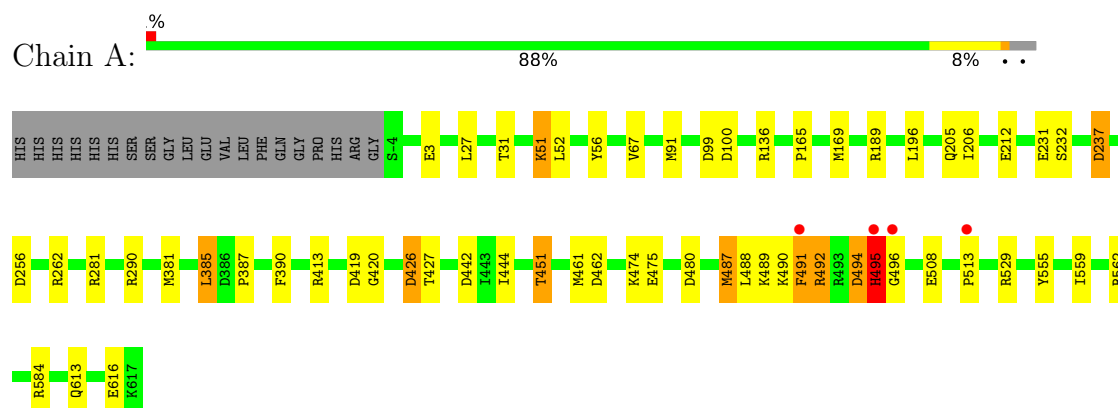
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	88	Total	O	0	0
			88	88		

### 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: ATP-dependent DNA helicase XPD



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	59.53Å 96.00Å 159.09Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	12.82 – 2.13 12.82 – 2.13	Depositor EDS
% Data completeness (in resolution range)	99.4 (12.82-2.13) 99.4 (12.82-2.13)	Depositor EDS
$R_{merge}$	0.16	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.10 (at 2.12Å)	Xtriage
Refinement program	REFMAC 5.8.0431 (refmacat 0.4.105)	Depositor
R, $R_{free}$	0.202 , 0.245 0.203 , 0.245	Depositor DCC
$R_{free}$ test set	2544 reflections (4.89%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	50.2	Xtriage
Anisotropy	0.502	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.42 , 39.6	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	10029	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	70.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.19% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: SF4, 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	0.69	0/5141	1.44	22/6913 (0.3%)

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	13

There are no bond length outliers.

All (22) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	495	HIS	CA-CB-CG	12.72	126.52	113.80
1	A	495	HIS	CB-CG-CD2	-9.69	118.60	131.20
1	A	495	HIS	CB-CG-ND1	8.99	136.19	122.70
1	A	387	PRO	CB-CA-C	7.28	123.57	111.56
1	A	442	ASP	CA-CB-CG	7.04	119.64	112.60
1	A	100	ASP	CA-CB-CG	6.75	119.35	112.60
1	A	480	ASP	CA-CB-CG	6.50	119.09	112.60
1	A	390	PHE	CA-CB-CG	6.29	120.09	113.80
1	A	426	ASP	CA-CB-CG	6.27	118.87	112.60
1	A	419	ASP	CA-CB-CG	6.19	118.79	112.60
1	A	237	ASP	CA-CB-CG	5.83	118.43	112.60
1	A	256	ASP	CA-CB-CG	5.70	118.30	112.60
1	A	212	GLU	CB-CA-C	5.60	119.36	111.86
1	A	492	ARG	CA-CB-CG	5.56	125.23	114.10
1	A	462	ASP	CA-CB-CG	5.19	117.79	112.60
1	A	385	LEU	CA-C-N	5.14	128.57	120.97

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	385	LEU	C-N-CA	5.14	128.57	120.97
1	A	451	THR	CA-CB-OG1	-5.12	101.92	109.60
1	A	495	HIS	CB-CA-C	5.12	119.63	111.24
1	A	487	MET	CG-SD-CE	5.03	111.98	100.90
1	A	51	LYS	CA-CB-CG	5.01	124.13	114.10
1	A	3	GLU	CB-CG-CD	5.00	121.10	112.60

There are no chirality outliers.

All (13) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	136	ARG	Sidechain
1	A	189	ARG	Sidechain
1	A	262	ARG	Sidechain
1	A	281	ARG	Sidechain
1	A	290	ARG	Sidechain
1	A	385	LEU	Peptide
1	A	413	ARG	Sidechain
1	A	420	GLY	Peptide
1	A	494	ASP	Peptide
1	A	495	HIS	Peptide
1	A	529	ARG	Sidechain
1	A	562	ARG	Sidechain
1	A	584	ARG	Sidechain

## 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	A	5040	4880	4998	9	0
2	A	8	0	0	0	0
3	A	13	0	0	0	0
4	A	88	0	0	0	0
All	All	5149	4880	4998	9	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 1.

All (9) 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:196:LEU:HD22	1:A:206:ILE:HD11	1.92	0.51
1:A:91:MET:HE3	1:A:165:PRO:HB2	1.94	0.49
1:A:56:TYR:CE1	1:A:67:VAL:HG21	2.49	0.48
1:A:461:MET:HE2	1:A:474:LYS:HD2	1.96	0.47
1:A:91:MET:HE2	1:A:169:MET:CE	2.48	0.42
1:A:27:LEU:HD23	1:A:381:MET:HE1	2.01	0.42
1:A:52:LEU:HD22	1:A:205:GLN:HA	2.01	0.42
1:A:444:ILE:HD11	1:A:451:THR:HG21	2.02	0.40
1:A:555:TYR:CE1	1:A:559:ILE:HD11	2.56	0.40

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	620/642 (97%)	600 (97%)	16 (3%)	4 (1%)	21	18

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	496	GLY
1	A	491	PHE
1	A	494	ASP
1	A	513	PRO

### 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	549/566 (97%)	530 (96%)	19 (4%)	32	34

All (19) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	31	THR
1	A	51	LYS
1	A	99	ASP
1	A	231	GLU
1	A	232	SER
1	A	237	ASP
1	A	426	ASP
1	A	427	THR
1	A	475	GLU
1	A	487	MET
1	A	488	LEU
1	A	489	LYS
1	A	490	LYS
1	A	491	PHE
1	A	492	ARG
1	A	495	HIS
1	A	508	GLU
1	A	613	GLN
1	A	616	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	412	ASN
1	A	481	GLN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

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 14 ligands modelled in this entry, 13 are monoatomic - leaving 1 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	$\# Z  > 2$
2	SF4	A	701	1	0,12,12	-	-	-		

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
2	SF4	A	701	1	-	-	0/6/5/5

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 [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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	622/642 (96%)	-0.58	4 (0%) 85 87	45, 65, 103, 189	0

All (4) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	496	GLY	6.7
1	A	491	PHE	6.1
1	A	495	HIS	4.6
1	A	513	PRO	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 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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
3	CL	A	704	1/1	0.66	0.25	102,102,102,102	0
3	CL	A	706	1/1	0.84	0.16	91,91,91,91	0
3	CL	A	705	1/1	0.88	0.08	94,94,94,94	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
3	CL	A	714	1/1	0.88	0.25	100,100,100,100	0
3	CL	A	708	1/1	0.90	0.17	132,132,132,132	0
3	CL	A	703	1/1	0.92	0.09	91,91,91,91	0
3	CL	A	709	1/1	0.92	0.26	109,109,109,109	0
3	CL	A	713	1/1	0.92	0.07	105,105,105,105	0
3	CL	A	707	1/1	0.92	0.06	114,114,114,114	0
3	CL	A	712	1/1	0.94	0.12	98,98,98,98	0
3	CL	A	711	1/1	0.95	0.14	102,102,102,102	0
3	CL	A	710	1/1	0.96	0.05	98,98,98,98	0
2	SF4	A	701	8/8	0.99	0.04	52,57,60,62	0
3	CL	A	702	1/1	0.99	0.06	58,58,58,58	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.