



# wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 28, 2024 – 04:17 PM EDT

PDB ID : 9B8G  
Title : 2C9 Fab antibody fragment against the E protein of the Yellow Fever virus  
Authors : Balk, J.; Ting, Y.T.; Bibby, S.; Watterson, D.; Coulibaly, F.  
Deposited on : 2024-03-29  
Resolution : 2.66 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](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 : 4.02b-467  
Mogul : 2022.3.0, CSD as543be (2022)  
Xtriage (Phenix) : 1.20.1  
EDS : 3.0  
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)  
CCP4 : 9.0.003 (Gargrove)  
Density-Fitness : 1.0.11  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.39

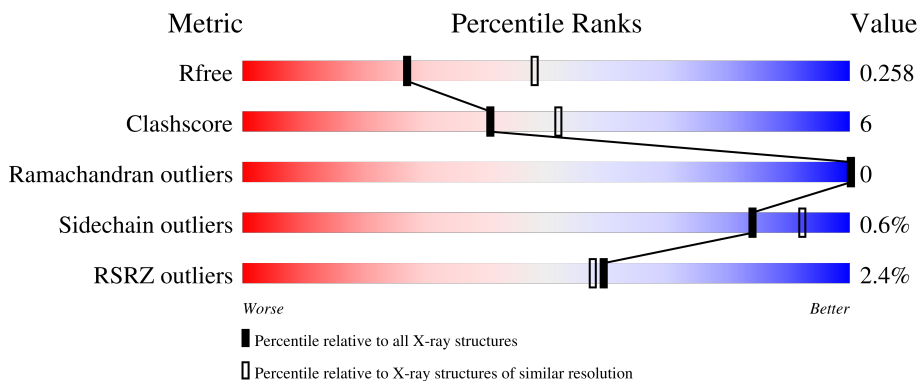
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.66 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.



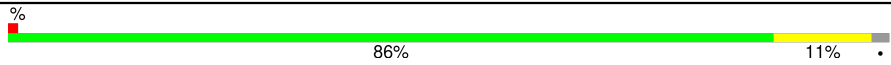
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	164625	1003 (2.66-2.66)
Clashscore	180529	1063 (2.66-2.66)
Ramachandran outliers	177936	1052 (2.66-2.66)
Sidechain outliers	177891	1052 (2.66-2.66)
RSRZ outliers	164620	1003 (2.66-2.66)

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	230	 3% 82% 13%
1	C	230	 0% 86% 11%
1	H	230	 4% 86% 9% 6%
2	B	218	 2% 83% 13%
2	D	218	 2% 82% 16%

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	L	218	 A horizontal bar chart representing the quality of chain. The bar is primarily green, indicating a high quality score of 86%. A small yellow segment at the end indicates a lower quality score of 11%. The bar is labeled with a '%' symbol at the start and ends with a small black dot.

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	SO4	L	301	-	-	X	-

## 2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 19485 atoms, of which 9285 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 2C9 Fab heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace	
			Total	C	H	N	O				S
1	A	220	3193	1048	1535	275	327	8	0	1	0
1	C	222	3219	1054	1551	278	328	8	0	0	0
1	H	217	3144	1032	1514	271	319	8	0	0	0

- Molecule 2 is a protein called 2C9 Fab light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace	
			Total	C	H	N	O				S
2	B	212	3220	1035	1571	275	332	7	0	0	0
2	D	212	3207	1041	1551	275	333	7	0	0	0
2	L	213	3224	1043	1563	275	336	7	0	0	0

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

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	1	Total	Cl	0	0
			1	1		

- Molecule 4 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	C	1	Total	O	S	0	0
			5	4	1		
4	C	1	Total	O	S	0	0
			5	4	1		
4	H	1	Total	O	S	0	0
			5	4	1		
4	H	1	Total	O	S	0	0
			5	4	1		
4	L	1	Total	O	S	0	0
			5	4	1		
4	L	1	Total	O	S	0	0
			5	4	1		

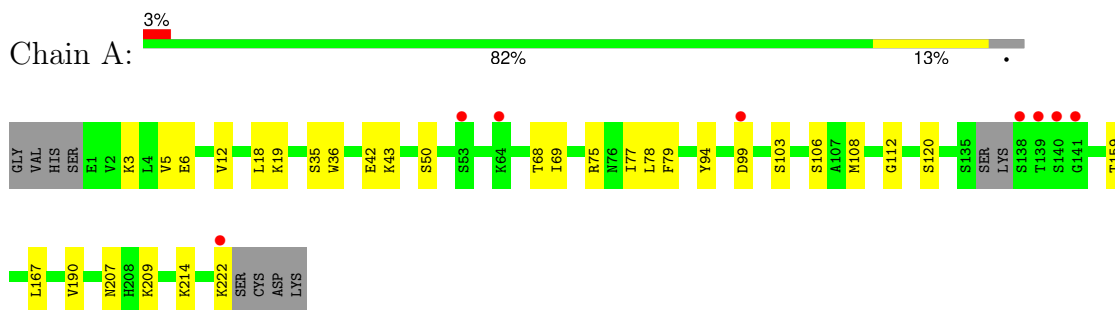
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	50	Total	O	0	0
			50	50		
5	B	37	Total	O	0	0
			37	37		
5	C	44	Total	O	0	0
			44	44		
5	D	28	Total	O	0	0
			28	28		
5	H	53	Total	O	0	0
			53	53		
5	L	35	Total	O	0	0
			35	35		

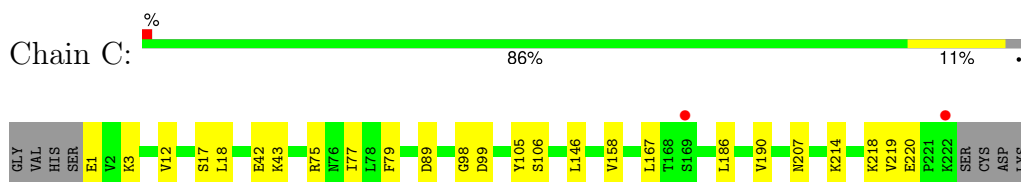
### 3 Residue-property plots [i](#)

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.

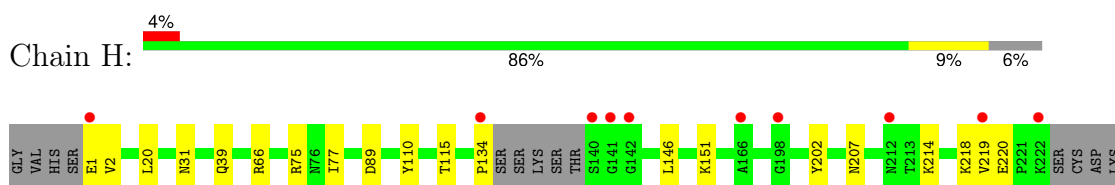
- Molecule 1: 2C9 Fab heavy chain



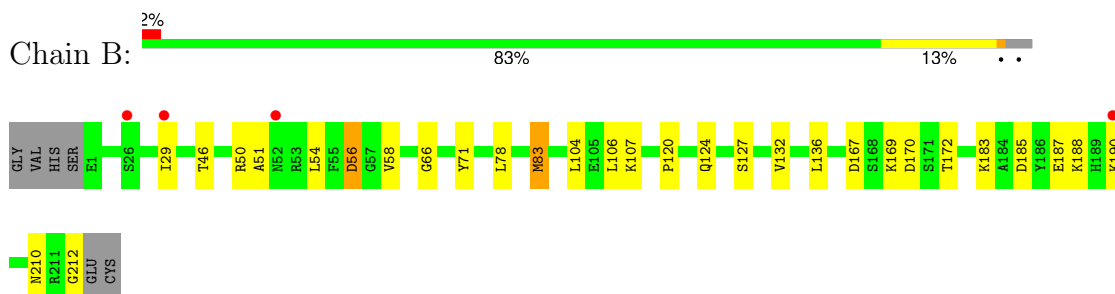
- Molecule 1: 2C9 Fab heavy chain



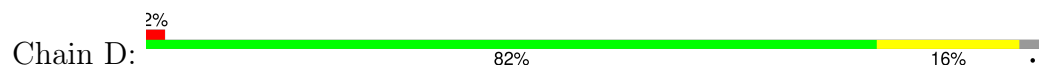
- Molecule 1: 2C9 Fab heavy chain



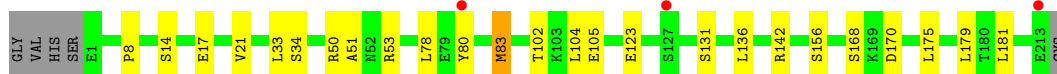
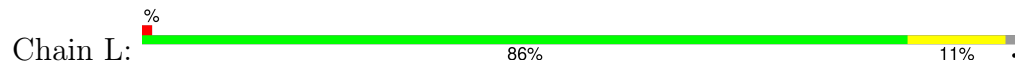
- Molecule 2: 2C9 Fab light chain



- Molecule 2: 2C9 Fab light chain



- Molecule 2: 2C9 Fab light chain



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	67.67Å 125.33Å 96.93Å 90.00° 93.95° 90.00°	Depositor
Resolution (Å)	49.32 – 2.66 49.32 – 2.66	Depositor EDS
% Data completeness (in resolution range)	99.9 (49.32-2.66) 100.0 (49.32-2.66)	Depositor EDS
$R_{merge}$	0.32	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.65 (at 2.65Å)	Xtrriage
Refinement program	PHENIX 1.21_5207	Depositor
R, $R_{free}$	0.201 , 0.258 0.200 , 0.258	Depositor DCC
$R_{free}$ test set	2294 reflections (4.95%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	39.7	Xtrriage
Anisotropy	0.532	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.39 , 38.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	19485	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	51.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.30% 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 [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: SO4, 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	0.38	0/1702	0.59	0/2315
1	C	0.39	0/1710	0.60	0/2326
1	H	0.37	0/1671	0.56	0/2274
2	B	0.39	0/1684	0.57	0/2278
2	D	0.38	0/1692	0.58	0/2289
2	L	0.38	0/1697	0.59	0/2297
All	All	0.38	0/10156	0.58	0/13779

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	1658	1535	1621	26	0
1	C	1668	1551	1636	21	0
1	H	1630	1514	1589	17	0
2	B	1649	1571	1602	21	0
2	D	1656	1551	1609	22	0
2	L	1661	1563	1604	18	0
3	A	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	C	10	0	0	0	0
4	H	10	0	0	1	0
4	L	10	0	0	3	0
5	A	50	0	0	6	0
5	B	37	0	0	0	0
5	C	44	0	0	2	0
5	D	28	0	0	2	0
5	H	53	0	0	3	0
5	L	35	0	0	5	0
All	All	10200	9285	9661	125	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 6.

The worst 5 of 125 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)
4:L:301:SO4:O4	5:L:401:HOH:O	1.87	0.93
1:A:12:VAL:O	5:A:401:HOH:O	1.94	0.85
2:B:190:LYS:HE3	2:B:210:ASN:HB3	1.59	0.84
4:L:301:SO4:S	5:L:401:HOH:O	2.36	0.83
2:B:83:MET:HE3	2:B:106:LEU:HB2	1.69	0.75

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	217/230 (94%)	212 (98%)	5 (2%)	0	100 100
1	C	220/230 (96%)	215 (98%)	5 (2%)	0	100 100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	H	213/230 (93%)	207 (97%)	6 (3%)	0	100	100
2	B	210/218 (96%)	207 (99%)	3 (1%)	0	100	100
2	D	210/218 (96%)	207 (99%)	3 (1%)	0	100	100
2	L	211/218 (97%)	207 (98%)	4 (2%)	0	100	100
All	All	1281/1344 (95%)	1255 (98%)	26 (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	A	187/195 (96%)	187 (100%)	0	100	100
1	C	188/195 (96%)	188 (100%)	0	100	100
1	H	182/195 (93%)	181 (100%)	1 (0%)	86	94
2	B	188/194 (97%)	186 (99%)	2 (1%)	70	84
2	D	189/194 (97%)	188 (100%)	1 (0%)	86	94
2	L	189/194 (97%)	186 (98%)	3 (2%)	58	76
All	All	1123/1167 (96%)	1116 (99%)	7 (1%)	84	92

5 of 7 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	H	31	ASN
2	L	83	MET
2	L	170	ASP
2	L	142	ARG
2	D	36	LEU

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

Mol	Chain	Res	Type
2	B	160	GLN
2	L	210	ASN

### 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 7 ligands modelled in this entry, 1 is 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	$\# Z  > 2$
4	SO4	L	301	-	4,4,4	0.77	0	6,6,6	0.28	0
4	SO4	C	301	-	4,4,4	0.81	0	6,6,6	0.17	0
4	SO4	L	302	-	4,4,4	0.85	0	6,6,6	0.26	0
4	SO4	H	302	-	4,4,4	0.73	0	6,6,6	0.13	0
4	SO4	C	302	-	4,4,4	0.80	0	6,6,6	0.14	0
4	SO4	H	301	-	4,4,4	0.83	0	6,6,6	0.14	0

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.

2 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	L	301	SO4	3	0
4	H	301	SO4	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 [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	220/230 (95%)	0.12	8 (3%) 46 44	30, 50, 73, 106	1 (0%)
1	C	222/230 (96%)	-0.07	2 (0%) 81 79	29, 46, 65, 105	0
1	H	217/230 (94%)	0.13	10 (4%) 38 35	32, 49, 86, 105	0
2	B	212/218 (97%)	-0.00	4 (1%) 66 64	34, 49, 69, 80	0
2	D	212/218 (97%)	0.02	4 (1%) 66 64	35, 48, 72, 94	0
2	L	213/218 (97%)	0.08	3 (1%) 73 71	34, 51, 77, 94	0
All	All	1296/1344 (96%)	0.05	31 (2%) 59 58	29, 49, 76, 106	1 (0%)

The worst 5 of 31 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	D	212	GLY	6.4
1	H	142	GLY	3.9
1	A	222	LYS	3.6
2	L	213	GLU	3.5
1	A	139	THR	3.4

### 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, 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
4	SO4	C	302	5/5	0.89	0.11	55,63,68,70	0
3	CL	A	301	1/1	0.91	0.16	56,56,56,56	0
4	SO4	C	301	5/5	0.92	0.12	48,51,67,70	0
4	SO4	H	302	5/5	0.92	0.14	58,59,77,94	0
4	SO4	H	301	5/5	0.93	0.10	48,52,65,68	0
4	SO4	L	301	5/5	0.96	0.08	42,46,54,58	0
4	SO4	L	302	5/5	0.96	0.08	34,39,46,51	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.