Full wwPDB X-ray Structure Validation Report ⓘ

Apr 17, 2025 – 12:09 PM EDT

PDB ID	:	9O8J / pdb_00009o8j
Title	:	Crystal structure of Phosphoglycerate mutase from Trichomonas vaginalis (sulfate bound)
Authors	:	Seattle Structural Genomics Center for Infectious Disease; Seattle Structural Genomics Center for Infectious Disease (SSGCID)
Deposited on	:	2025-04-16
Resolution	:	2.05 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:

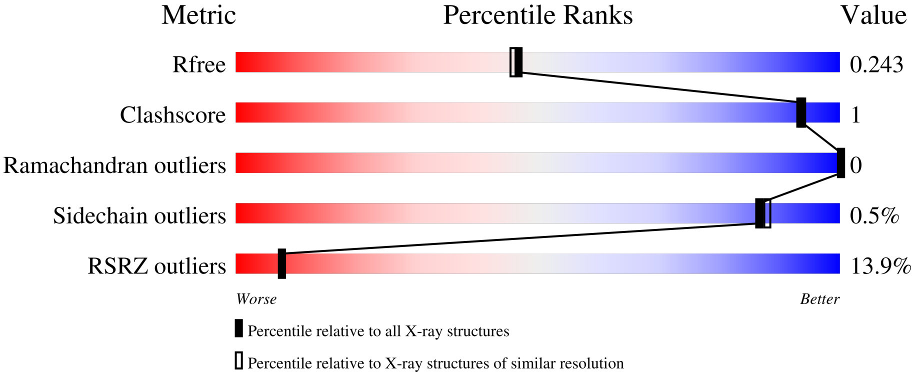
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	2.0rc1
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.006 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.42

1 Overall quality at a glance

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

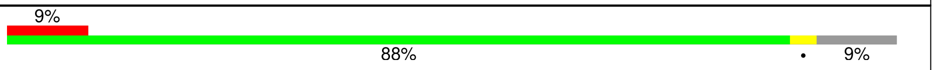
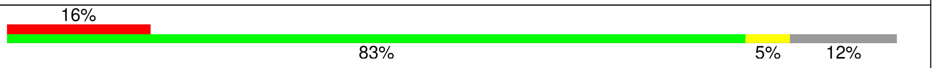
The reported resolution of this entry is 2.05 Å.

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	2096 (2.04-2.04)
Clashscore	180529	2229 (2.04-2.04)
Ramachandran outliers	177936	2217 (2.04-2.04)
Sidechain outliers	177891	2217 (2.04-2.04)
RSRZ outliers	164620	2096 (2.04-2.04)

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 <=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	258	
1	B	258	

2 Entry composition ⓘ

There are 3 unique types of molecules in this entry. The entry contains 3936 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.

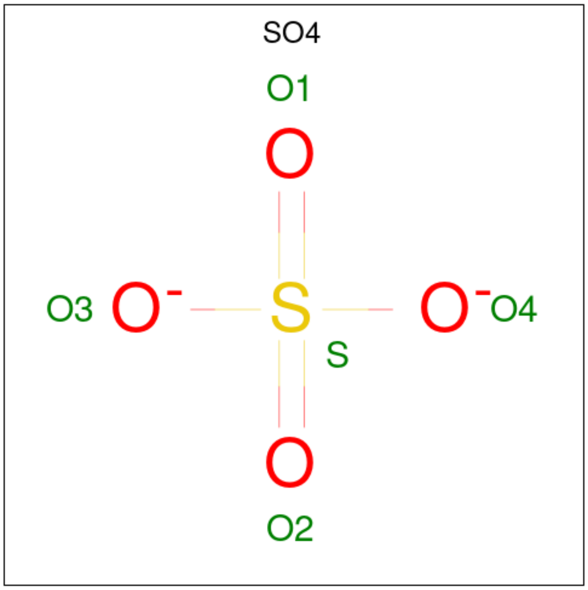
- Molecule 1 is a protein called Phosphoglycerate mutase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	235	Total 1913	C 1218	N 324	O 356	S 15	0	1	0
1	B	226	Total 1824	C 1162	N 310	O 339	S 13	0	1	0

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-7	MET	-	initiating methionine	UNP A2DUN8
A	-6	ALA	-	expression tag	UNP A2DUN8
A	-5	HIS	-	expression tag	UNP A2DUN8
A	-4	HIS	-	expression tag	UNP A2DUN8
A	-3	HIS	-	expression tag	UNP A2DUN8
A	-2	HIS	-	expression tag	UNP A2DUN8
A	-1	HIS	-	expression tag	UNP A2DUN8
A	0	HIS	-	expression tag	UNP A2DUN8
B	-7	MET	-	initiating methionine	UNP A2DUN8
B	-6	ALA	-	expression tag	UNP A2DUN8
B	-5	HIS	-	expression tag	UNP A2DUN8
B	-4	HIS	-	expression tag	UNP A2DUN8
B	-3	HIS	-	expression tag	UNP A2DUN8
B	-2	HIS	-	expression tag	UNP A2DUN8
B	-1	HIS	-	expression tag	UNP A2DUN8
B	0	HIS	-	expression tag	UNP A2DUN8

- Molecule 2 is SULFATE ION (CCD ID: SO4) (formula: O_4S).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total O S 5 4 1	0	0
2	A	1	Total O S 5 4 1	0	0
2	A	1	Total O S 5 4 1	0	0
2	A	1	Total O S 5 4 1	0	0
2	A	1	Total O S 5 4 1	0	0
2	A	1	Total O S 5 4 1	0	0
2	A	1	Total O S 10 8 2	0	1
2	A	1	Total O S 5 4 1	0	0
2	A	1	Total O S 5 4 1	0	0
2	A	1	Total O S 5 4 1	0	0
2	A	1	Total O S 5 4 1	0	0
2	A	1	Total O S 5 4 1	0	0
2	B	1	Total O S 5 4 1	0	0
2	B	1	Total O S 5 4 1	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	B	1	Total O S 5 4 1	0	0
2	B	1	Total O S 10 8 2	0	1
2	B	1	Total O S 5 4 1	0	0
2	B	1	Total O S 5 4 1	0	0

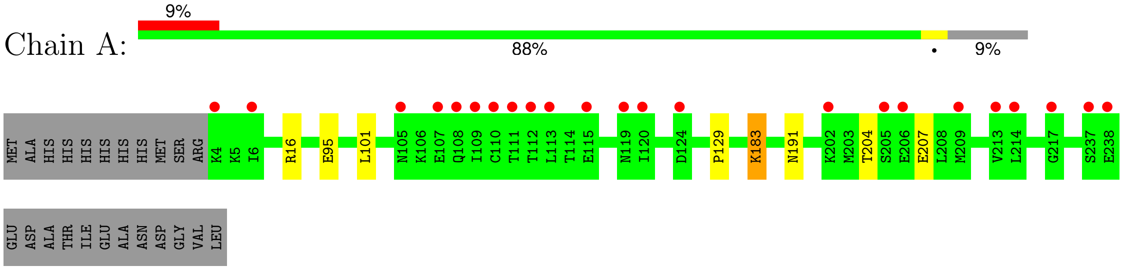
- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	76	Total O 76 76	0	0
3	B	23	Total O 23 23	0	0

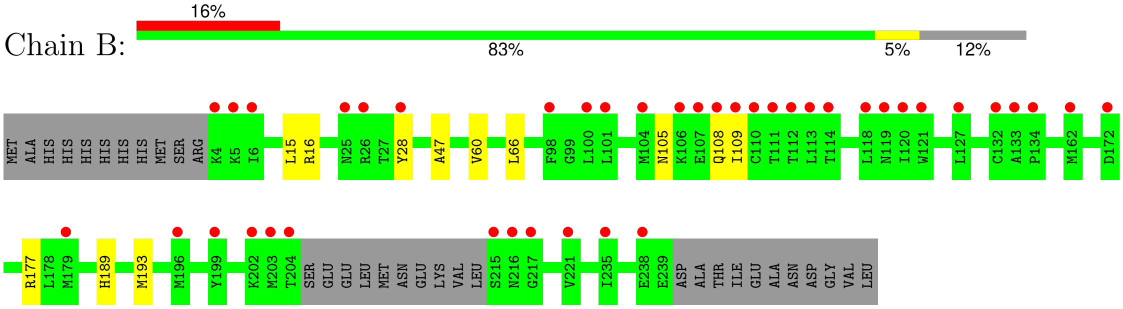
3 Residue-property plots ⓘ

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: Phosphoglycerate mutase



- Molecule 1: Phosphoglycerate mutase



4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α, β, γ	92.36Å 95.80Å 73.52Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	47.90 – 2.05 47.90 – 2.05	Depositor EDS
% Data completeness (in resolution range)	99.9 (47.90-2.05) 99.9 (47.90-2.05)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.84 (at 2.05Å)	Xtriage
Refinement program	PHENIX dev_5660	Depositor
R, R_{free}	0.205 , 0.239 0.211 , 0.243	Depositor DCC
R_{free} test set	2100 reflections (5.05%)	wwPDB-VP
Wilson B-factor (Å^2)	39.6	Xtriage
Anisotropy	0.157	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 56.1	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.199 for k,h,-l	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	3936	wwPDB-VP
Average B, all atoms (Å^2)	64.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.06% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ 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: SO4

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	0.25	0/1958	0.48	0/2643
1	B	0.24	0/1868	0.48	0/2525
All	All	0.24	0/3826	0.48	0/5168

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 ⓘ

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	1913	0	1913	4	0
1	B	1824	0	1796	6	0
2	A	65	0	0	0	0
2	B	35	0	0	0	0
3	A	76	0	0	0	0
3	B	23	0	0	0	0
All	All	3936	0	3709	10	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 (10) 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:B:28:TYR:HE2	1:B:109:ILE:HG21	1.72	0.55
1:A:204:THR:HG23	1:A:207:GLU:OE2	2.07	0.53
1:A:183:LYS:N	1:A:183:LYS:HD2	2.25	0.52
1:B:66:LEU:HD12	1:B:189:HIS:CE1	2.46	0.50
1:B:15:LEU:HD11	1:B:47:ALA:HB1	1.97	0.46
1:B:105:ASN:H	1:B:108:GLN:NE2	2.15	0.45
1:B:60:VAL:HG11	1:B:177:ARG:HD3	1.99	0.43
1:B:16:ARG:N	1:B:193:MET:HE1	2.34	0.42
1:A:101:LEU:HD22	1:A:129:PRO:HB3	2.02	0.41
1:A:95:GLU:OE2	1:A:191:ASN:HB2	2.21	0.41

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	234/258 (91%)	233 (100%)	1 (0%)	0	100	100
1	B	223/258 (86%)	220 (99%)	3 (1%)	0	100	100
All	All	457/516 (89%)	453 (99%)	4 (1%)	0	100	100

There are no Ramachandran outliers to report.

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	216/234 (92%)	214 (99%)	2 (1%)	75	77
1	B	202/234 (86%)	202 (100%)	0	100	100
All	All	418/468 (89%)	416 (100%)	2 (0%)	86	88

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	16	ARG
1	A	183	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	105	ASN
1	B	108	GLN
1	B	128	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 ⓘ

There are no oligosaccharides in this entry.

5.6 Ligand geometry ⓘ

20 ligands are modelled in this entry.

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	SO4	A	303	-	4,4,4	0.65	0	6,6,6	0.10	0
2	SO4	A	301	-	4,4,4	0.67	0	6,6,6	0.06	0
2	SO4	A	302	-	4,4,4	0.68	0	6,6,6	0.07	0
2	SO4	A	306	-	4,4,4	0.67	0	6,6,6	0.08	0
2	SO4	A	304	-	4,4,4	0.67	0	6,6,6	0.08	0
2	SO4	A	305	-	4,4,4	0.68	0	6,6,6	0.09	0
2	SO4	A	308	-	4,4,4	0.67	0	6,6,6	0.09	0
2	SO4	A	309	-	4,4,4	0.68	0	6,6,6	0.08	0
2	SO4	B	303	-	4,4,4	0.67	0	6,6,6	0.09	0
2	SO4	B	301	-	4,4,4	0.67	0	6,6,6	0.09	0
2	SO4	B	306	-	4,4,4	0.67	0	6,6,6	0.10	0
2	SO4	A	310	-	4,4,4	0.67	0	6,6,6	0.07	0
2	SO4	A	307[A]	-	4,4,4	0.66	0	6,6,6	0.09	0
2	SO4	B	302	-	4,4,4	0.67	0	6,6,6	0.08	0
2	SO4	A	307[B]	-	4,4,4	0.66	0	6,6,6	0.09	0
2	SO4	B	304[A]	-	4,4,4	0.67	0	6,6,6	0.08	0
2	SO4	B	304[B]	-	4,4,4	0.67	0	6,6,6	0.08	0
2	SO4	A	311	-	4,4,4	0.66	0	6,6,6	0.09	0
2	SO4	A	312	-	4,4,4	0.67	0	6,6,6	0.09	0
2	SO4	B	305	-	4,4,4	0.67	0	6,6,6	0.07	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.

No monomer is involved in short contacts.

5.7 Other polymers ⓘ

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	235/258 (91%)	0.53	23 (9%)	14	15	20, 43, 94, 127	1 (0%)
1	B	226/258 (87%)	1.22	41 (18%)	4	3	34, 75, 127, 150	1 (0%)
All	All	461/516 (89%)	0.87	64 (13%)	7	7	20, 57, 118, 150	2 (0%)

All (64) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	204	THR	5.6
1	B	109	ILE	5.5
1	A	111	THR	4.2
1	A	112	THR	4.0
1	A	217	GLY	3.9
1	B	113	LEU	3.9
1	B	111	THR	3.9
1	B	25	ASN	3.9
1	A	109	ILE	3.7
1	B	101	LEU	3.6
1	A	115	GLU	3.5
1	B	4	LYS	3.5
1	A	214	LEU	3.3
1	B	118	LEU	3.2
1	A	4	LYS	3.2
1	A	238	GLU	3.2
1	B	235	ILE	3.1
1	B	121	TRP	3.1
1	A	110	CYS	3.1
1	B	199	TYR	3.1
1	B	202	LYS	3.1
1	A	209	MET	3.0
1	B	5	LYS	3.0
1	B	120	ILE	2.9

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	112	THR	2.9
1	A	119	ASN	2.9
1	B	98	PHE	2.8
1	B	108	GLN	2.8
1	A	206	GLU	2.8
1	B	28	TYR	2.8
1	B	6	ILE	2.8
1	B	114	THR	2.8
1	A	213	VAL	2.7
1	A	108	GLN	2.7
1	B	238	GLU	2.7
1	B	217	GLY	2.6
1	B	215	SER	2.6
1	B	107	GLU	2.5
1	B	110	CYS	2.5
1	B	132	CYS	2.5
1	A	113	LEU	2.5
1	B	119	ASN	2.5
1	A	202	LYS	2.4
1	B	162	MET	2.4
1	A	107	GLU	2.4
1	B	127	LEU	2.3
1	B	104	MET	2.3
1	B	203	MET	2.3
1	A	105	ASN	2.3
1	B	216	ASN	2.3
1	B	106	LYS	2.3
1	A	6	ILE	2.3
1	A	237	SER	2.3
1	B	172	ASP	2.2
1	B	196	MET	2.2
1	A	120	ILE	2.2
1	A	124	ASP	2.2
1	B	133	ALA	2.1
1	B	221	VAL	2.1
1	B	134	PRO	2.1
1	A	205	SER	2.1
1	B	100	LEU	2.1
1	B	179	MET	2.1
1	B	26	ARG	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 ⓘ

There are no monosaccharides in this entry.

6.4 Ligands ⓘ

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
2	SO4	A	311	5/5	0.62	0.10	138,148,161,161	0
2	SO4	A	312	5/5	0.63	0.11	94,100,113,118	0
2	SO4	B	304[A]	5/5	0.71	0.15	64,64,71,75	5
2	SO4	B	304[B]	5/5	0.71	0.15	68,69,75,78	5
2	SO4	B	306	5/5	0.71	0.10	108,111,116,119	0
2	SO4	A	309	5/5	0.73	0.10	86,102,116,133	0
2	SO4	B	305	5/5	0.74	0.12	84,93,106,110	0
2	SO4	B	302	5/5	0.74	0.12	91,93,102,103	0
2	SO4	A	310	5/5	0.75	0.10	60,73,92,109	0
2	SO4	B	301	5/5	0.78	0.12	59,65,82,92	0
2	SO4	A	304	5/5	0.78	0.09	89,89,96,99	0
2	SO4	A	305	5/5	0.79	0.14	57,66,76,82	0
2	SO4	A	306	5/5	0.80	0.08	75,86,103,114	0
2	SO4	A	302	5/5	0.82	0.12	53,61,77,93	0
2	SO4	B	303	5/5	0.82	0.12	64,72,76,82	0
2	SO4	A	308	5/5	0.84	0.09	78,78,96,98	0
2	SO4	A	307[A]	5/5	0.84	0.13	40,42,47,51	5
2	SO4	A	307[B]	5/5	0.84	0.13	51,51,55,62	5
2	SO4	A	301	5/5	0.87	0.11	51,51,66,79	0
2	SO4	A	303	5/5	0.94	0.15	49,49,51,52	0

6.5 Other polymers ⓘ

There are no such residues in this entry.