



# Full wwPDB X-ray Structure Validation Report ⓘ

Nov 15, 2017 – 10:57 PM EST

PDB ID : 4PJT  
Title : Structure of PARP1 catalytic domain bound to inhibitor BMN 673  
Authors : Aoyagi-Scharber, M.; Gardberg, A.S.; Arakaki, T.L.  
Deposited on : unknown  
Resolution : 2.35 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

<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 : 1.7.2 (RC1), CSD as538be (2017)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20030345  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
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-20030345

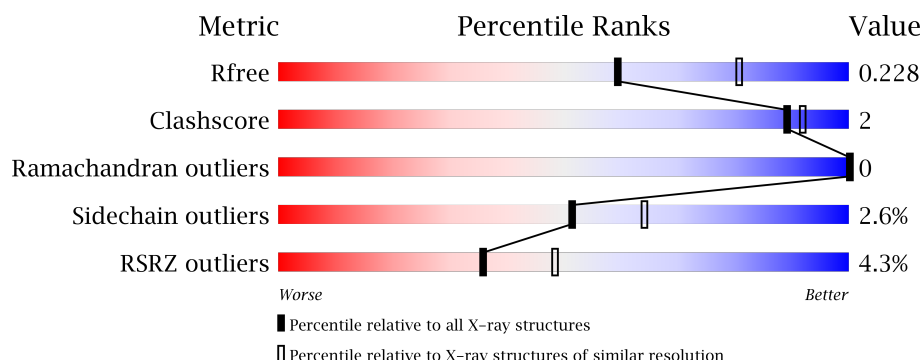
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.35 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))
$R_{free}$	100719	1522 (2.38-2.34)
Clashscore	112137	1626 (2.38-2.34)
Ramachandran outliers	110173	1605 (2.38-2.34)
Sidechain outliers	110143	1606 (2.38-2.34)
RSRZ outliers	101464	1528 (2.38-2.34)

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  $\geq 3$ , 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	370	<div> <div>0.1%</div> <div>85%</div> <div>6%</div> <div>9%</div> </div>
1	B	370	<div> <div>4%</div> <div>84%</div> <div>5%</div> <div>10%</div> </div>
1	C	370	<div> <div>3%</div> <div>87%</div> <div>5%</div> <div>6%</div> </div>
1	D	370	<div> <div>9%</div> <div>81%</div> <div>6%</div> <div>12%</div> </div>

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 crite-

ria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	SO4	A	1101	-	-	-	X
2	SO4	B	1104	-	-	-	X
2	SO4	B	1106	-	-	-	X
2	SO4	D	1102	-	-	-	X
4	GOL	A	1107	-	-	-	X

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 10711 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 Poly [ADP-ribose] polymerase 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	338	Total	C	N	O	S	0	1	0
			2595	1657	428	500	10			
1	B	332	Total	C	N	O	S	0	0	0
			2504	1603	412	479	10			
1	C	346	Total	C	N	O	S	0	2	0
			2647	1690	440	506	11			
1	D	326	Total	C	N	O	S	0	0	0
			2444	1562	408	464	10			

There are 80 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	642	MET	-	expression tag	UNP P09874
A	643	GLY	-	expression tag	UNP P09874
A	644	SER	-	expression tag	UNP P09874
A	645	SER	-	expression tag	UNP P09874
A	646	HIS	-	expression tag	UNP P09874
A	647	HIS	-	expression tag	UNP P09874
A	648	HIS	-	expression tag	UNP P09874
A	649	HIS	-	expression tag	UNP P09874
A	650	HIS	-	expression tag	UNP P09874
A	651	HIS	-	expression tag	UNP P09874
A	652	SER	-	expression tag	UNP P09874
A	653	GLY	-	expression tag	UNP P09874
A	654	LEU	-	expression tag	UNP P09874
A	655	VAL	-	expression tag	UNP P09874
A	656	PRO	-	expression tag	UNP P09874
A	657	ARG	-	expression tag	UNP P09874
A	658	GLY	-	expression tag	UNP P09874
A	659	SER	-	expression tag	UNP P09874
A	660	HIS	-	expression tag	UNP P09874
A	661	MET	-	expression tag	UNP P09874
B	642	MET	-	expression tag	UNP P09874

*Continued on next page...*

*Continued from previous page...*

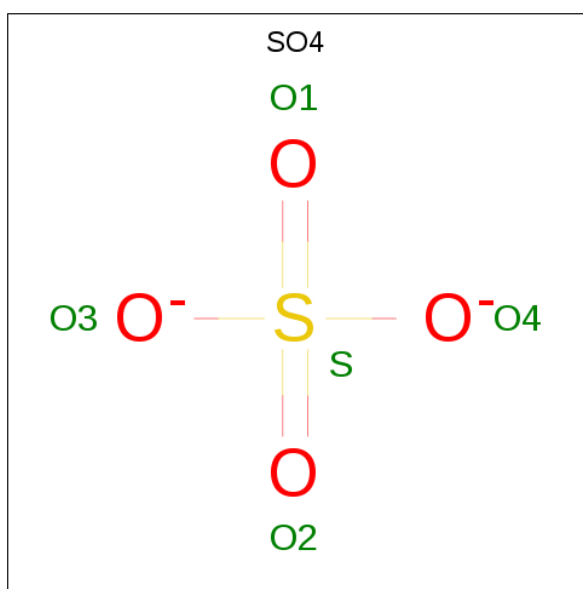
Chain	Residue	Modelled	Actual	Comment	Reference
B	643	GLY	-	expression tag	UNP P09874
B	644	SER	-	expression tag	UNP P09874
B	645	SER	-	expression tag	UNP P09874
B	646	HIS	-	expression tag	UNP P09874
B	647	HIS	-	expression tag	UNP P09874
B	648	HIS	-	expression tag	UNP P09874
B	649	HIS	-	expression tag	UNP P09874
B	650	HIS	-	expression tag	UNP P09874
B	651	HIS	-	expression tag	UNP P09874
B	652	SER	-	expression tag	UNP P09874
B	653	GLY	-	expression tag	UNP P09874
B	654	LEU	-	expression tag	UNP P09874
B	655	VAL	-	expression tag	UNP P09874
B	656	PRO	-	expression tag	UNP P09874
B	657	ARG	-	expression tag	UNP P09874
B	658	GLY	-	expression tag	UNP P09874
B	659	SER	-	expression tag	UNP P09874
B	660	HIS	-	expression tag	UNP P09874
B	661	MET	-	expression tag	UNP P09874
C	642	MET	-	expression tag	UNP P09874
C	643	GLY	-	expression tag	UNP P09874
C	644	SER	-	expression tag	UNP P09874
C	645	SER	-	expression tag	UNP P09874
C	646	HIS	-	expression tag	UNP P09874
C	647	HIS	-	expression tag	UNP P09874
C	648	HIS	-	expression tag	UNP P09874
C	649	HIS	-	expression tag	UNP P09874
C	650	HIS	-	expression tag	UNP P09874
C	651	HIS	-	expression tag	UNP P09874
C	652	SER	-	expression tag	UNP P09874
C	653	GLY	-	expression tag	UNP P09874
C	654	LEU	-	expression tag	UNP P09874
C	655	VAL	-	expression tag	UNP P09874
C	656	PRO	-	expression tag	UNP P09874
C	657	ARG	-	expression tag	UNP P09874
C	658	GLY	-	expression tag	UNP P09874
C	659	SER	-	expression tag	UNP P09874
C	660	HIS	-	expression tag	UNP P09874
C	661	MET	-	expression tag	UNP P09874
D	642	MET	-	expression tag	UNP P09874
D	643	GLY	-	expression tag	UNP P09874
D	644	SER	-	expression tag	UNP P09874

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
D	645	SER	-	expression tag	UNP P09874
D	646	HIS	-	expression tag	UNP P09874
D	647	HIS	-	expression tag	UNP P09874
D	648	HIS	-	expression tag	UNP P09874
D	649	HIS	-	expression tag	UNP P09874
D	650	HIS	-	expression tag	UNP P09874
D	651	HIS	-	expression tag	UNP P09874
D	652	SER	-	expression tag	UNP P09874
D	653	GLY	-	expression tag	UNP P09874
D	654	LEU	-	expression tag	UNP P09874
D	655	VAL	-	expression tag	UNP P09874
D	656	PRO	-	expression tag	UNP P09874
D	657	ARG	-	expression tag	UNP P09874
D	658	GLY	-	expression tag	UNP P09874
D	659	SER	-	expression tag	UNP P09874
D	660	HIS	-	expression tag	UNP P09874
D	661	MET	-	expression tag	UNP P09874

- Molecule 2 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



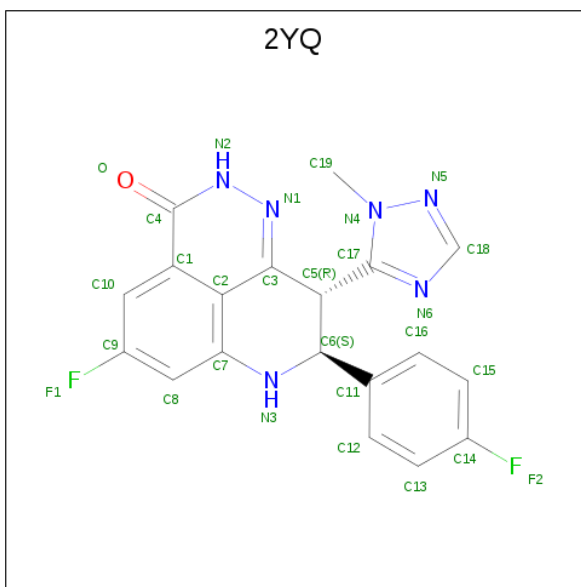
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	O	S	0	0
			5	4	1		
2	A	1	Total	O	S	0	0
			5	4	1		
2	A	1	Total	O	S	0	0
			5	4	1		

*Continued on next page...*

*Continued from previous page...*

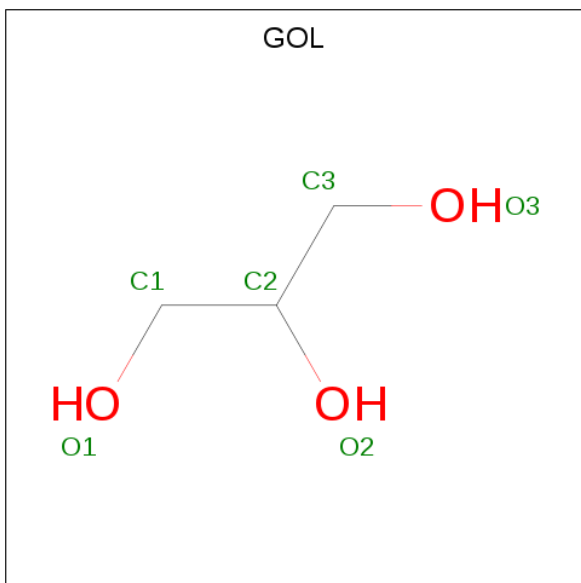
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	O	S	0	0
			5	4	1		
2	B	1	Total	O	S	0	0
			5	4	1		
2	B	1	Total	O	S	0	0
			5	4	1		
2	B	1	Total	O	S	0	0
			5	4	1		
2	B	1	Total	O	S	0	0
			5	4	1		
2	B	1	Total	O	S	0	0
			5	4	1		
2	C	1	Total	O	S	0	0
			5	4	1		
2	C	1	Total	O	S	0	0
			5	4	1		
2	C	1	Total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

- Molecule 3 is (8S,9R)-5-fluoro-8-(4-fluorophenyl)-9-(1-methyl-1H-1,2,4-triazol-5-yl)-2,7,8,9-tetrahydro-3H-pyrido[4,3,2-de]phthalazin-3-one (three-letter code: 2YQ) (formula: C<sub>19</sub>H<sub>14</sub>F<sub>2</sub>N<sub>6</sub>O).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total	C	F	N	O	0	0
			28	19	2	6	1		
3	B	1	Total	C	F	N	O	0	0
			28	19	2	6	1		
3	C	1	Total	C	F	N	O	0	0
			28	19	2	6	1		
3	D	1	Total	C	F	N	O	0	0
			28	19	2	6	1		

- Molecule 4 is GLYCEROL (three-letter code: GOL) (formula:  $C_3H_8O_3$ ).





Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			6	3	3		
4	A	1	Total	C	O	0	0
			6	3	3		
4	C	1	Total	C	O	0	0
			6	3	3		

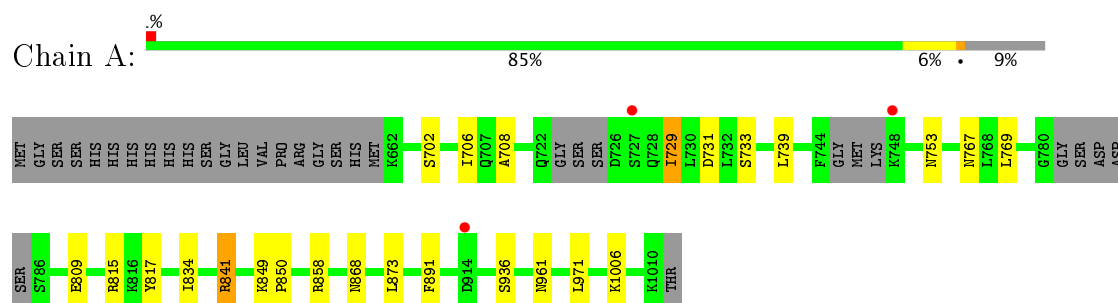
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	108	Total	O	0	0
			108	108		
5	B	55	Total	O	0	0
			55	55		
5	C	100	Total	O	0	0
			100	100		
5	D	53	Total	O	0	0
			53	53		

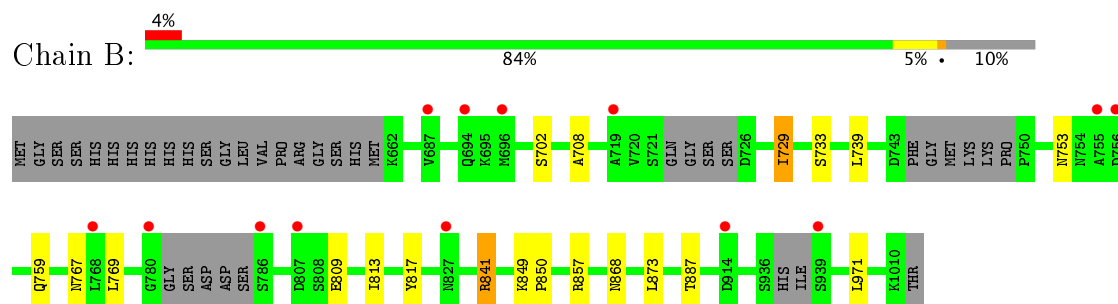
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of 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.

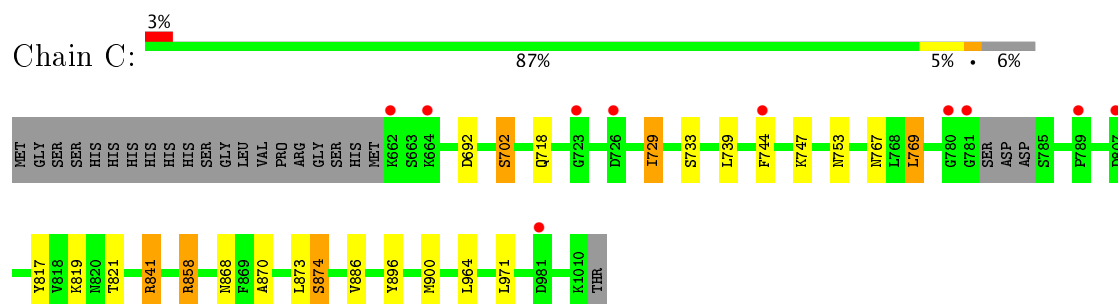
- Molecule 1: Poly [ADP-ribose] polymerase 1



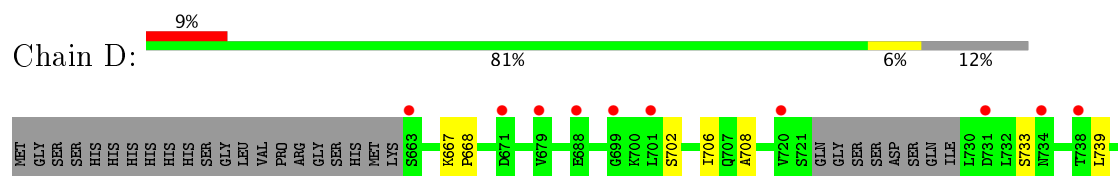
- Molecule 1: Poly [ADP-ribose] polymerase 1

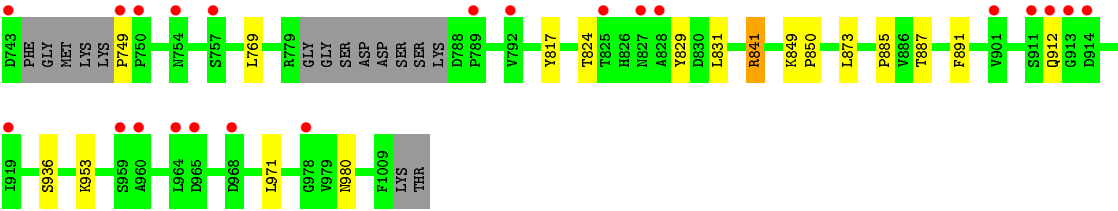


- Molecule 1: Poly [ADP-ribose] polymerase 1



- Molecule 1: Poly [ADP-ribose] polymerase 1





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	103.69Å 108.15Å 142.00Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	19.94 – 2.35 19.94 – 2.35	Depositor EDS
% Data completeness (in resolution range)	99.6 (19.94-2.35) 99.8 (19.94-2.35)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.52 (at 2.35Å)	Xtriage
Refinement program	REFMAC 5.8.0071	Depositor
R, $R_{free}$	0.188 , 0.228 0.192 , 0.228	Depositor DCC
$R_{free}$ test set	3387 reflections (5.33%)	DCC
Wilson B-factor (Å <sup>2</sup> )	35.9	Xtriage
Anisotropy	0.010	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 47.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	0.020 for k,h,-l	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	10711	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	41.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.60% 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: GOL, SO4, 2YQ

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	0.65	0/2647	0.78	2/3590 (0.1%)
1	B	0.58	0/2549	0.74	3/3462 (0.1%)
1	C	0.63	0/2706	0.77	5/3675 (0.1%)
1	D	0.53	0/2490	0.71	3/3388 (0.1%)
All	All	0.60	0/10392	0.75	13/14115 (0.1%)

There are no bond length outliers.

All (13) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	841	ARG	NE-CZ-NH2	-10.50	115.05	120.30
1	B	841	ARG	NE-CZ-NH2	-9.83	115.38	120.30
1	C	841	ARG	NE-CZ-NH1	9.47	125.04	120.30
1	C	841	ARG	NE-CZ-NH2	-9.37	115.62	120.30
1	D	841	ARG	NE-CZ-NH2	-8.98	115.81	120.30
1	A	841	ARG	NE-CZ-NH1	8.36	124.48	120.30
1	B	841	ARG	NE-CZ-NH1	8.26	124.43	120.30
1	D	841	ARG	NE-CZ-NH1	7.37	123.99	120.30
1	D	749	PRO	N-CA-CB	6.00	110.50	103.30
1	C	692	ASP	CB-CG-OD1	5.90	123.61	118.30
1	B	857	ARG	NE-CZ-NH2	-5.59	117.51	120.30
1	C	858	ARG	CG-CD-NE	-5.42	100.42	111.80
1	C	964	LEU	CA-CB-CG	5.12	127.08	115.30

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	2595	0	2532	15	0
1	B	2504	0	2412	9	0
1	C	2647	0	2575	13	0
1	D	2444	0	2340	10	0
2	A	20	0	0	1	0
2	B	25	0	0	0	0
2	C	15	0	0	3	0
2	D	15	0	0	0	0
3	A	28	0	14	0	0
3	B	28	0	14	0	0
3	C	28	0	14	0	0
3	D	28	0	14	0	0
4	A	12	0	16	0	0
4	C	6	0	8	1	0
5	A	108	0	0	1	0
5	B	55	0	0	0	0
5	C	100	0	0	2	0
5	D	53	0	0	0	0
All	All	10711	0	9939	44	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 (44) 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:759:GLN:HG3	1:B:887:THR:HG22	1.48	0.92
1:B:841:ARG:HD2	1:B:873:LEU:O	1.83	0.78
1:A:809:GLU:CD	1:C:886:VAL:HG11	2.06	0.76
1:A:841:ARG:HD2	1:A:873:LEU:O	1.86	0.74
1:A:809:GLU:OE1	1:C:886:VAL:HG11	1.87	0.73
1:C:841:ARG:HD2	1:C:873:LEU:O	1.87	0.73
1:D:841:ARG:HD2	1:D:873:LEU:O	1.88	0.73
1:C:769:LEU:HD22	5:C:1283:HOH:O	1.94	0.67
1:B:729:ILE:HD11	1:B:753:ASN:CB	2.27	0.64

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:729:ILE:HD11	1:B:753:ASN:HB3	1.78	0.64
1:C:702:SER:HB2	2:C:1103:SO4:O1	2.01	0.61
1:B:759:GLN:CG	1:B:887:THR:HG22	2.27	0.59
1:D:829:TYR:HE1	1:D:831:LEU:HD21	1.69	0.56
1:D:885:PRO:HB2	1:D:887:THR:HG22	1.87	0.55
1:A:961:ASN:ND2	5:A:1243:HOH:O	2.39	0.55
1:C:744:PHE:HB2	1:C:747:LYS:O	2.09	0.52
1:A:767:ASN:HA	1:A:868:ASN:HD21	1.75	0.52
2:C:1101:SO4:O2	5:C:1266:HOH:O	2.18	0.50
1:A:729:ILE:HD11	1:A:753:ASN:CB	2.43	0.49
2:C:1102:SO4:O3	1:D:849:LYS:NZ	2.41	0.49
1:A:708:ALA:HB3	1:A:739:LEU:HD21	1.95	0.48
1:B:767:ASN:HA	1:B:868:ASN:HD21	1.78	0.48
1:C:767:ASN:HA	1:C:868:ASN:HD21	1.79	0.48
1:A:858:ARG:NH2	2:A:1102:SO4:O3	2.46	0.45
1:B:849:LYS:HB3	1:B:850:PRO:HD3	1.98	0.45
1:C:896:TYR:CE2	4:C:1105:GOL:H32	2.51	0.45
1:B:809:GLU:O	1:B:813:ILE:HG13	2.17	0.45
1:B:708:ALA:HB3	1:B:739:LEU:HD21	1.98	0.44
1:C:870:ALA:O	1:C:874:SER:HB3	2.18	0.44
1:D:849:LYS:HB3	1:D:850:PRO:HD3	1.99	0.44
1:A:849:LYS:HB3	1:A:850:PRO:HD3	1.99	0.43
1:A:729:ILE:HD11	1:A:753:ASN:HB3	1.99	0.42
1:C:821:THR:HB	1:C:900:MET:HA	2.02	0.42
1:A:815:ARG:NH1	1:C:718:GLN:HG2	2.35	0.42
1:D:708:ALA:HB3	1:D:739:LEU:HD21	2.00	0.42
1:A:891:PHE:HA	1:A:936:SER:O	2.20	0.42
1:A:809:GLU:OE1	1:C:886:VAL:CG1	2.62	0.41
1:D:891:PHE:HA	1:D:936:SER:O	2.21	0.41
1:A:834:ILE:HD11	1:A:1006:LYS:HB2	2.01	0.41
1:D:667:LYS:N	1:D:668:PRO:CD	2.84	0.41
1:D:953:LYS:NZ	1:D:980:ASN:O	2.51	0.41
1:D:706:ILE:HG21	1:D:769:LEU:HG	2.03	0.41
1:A:706:ILE:HG21	1:A:769:LEU:HG	2.03	0.40
1:C:729:ILE:HD11	1:C:753:ASN:CB	2.52	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	331/370 (90%)	329 (99%)	2 (1%)	0	100	100
1	B	322/370 (87%)	321 (100%)	1 (0%)	0	100	100
1	C	344/370 (93%)	342 (99%)	2 (1%)	0	100	100
1	D	318/370 (86%)	316 (99%)	2 (1%)	0	100	100
All	All	1315/1480 (89%)	1308 (100%)	7 (0%)	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	279/325 (86%)	273 (98%)	6 (2%)	57	70
1	B	260/325 (80%)	254 (98%)	6 (2%)	56	68
1	C	282/325 (87%)	272 (96%)	10 (4%)	41	52
1	D	252/325 (78%)	246 (98%)	6 (2%)	54	67
All	All	1073/1300 (82%)	1045 (97%)	28 (3%)	51	64

All (28) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	702	SER
1	A	729	ILE

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type
1	A	731	ASP
1	A	733	SER
1	A	817	TYR
1	A	971	LEU
1	B	702	SER
1	B	729	ILE
1	B	733	SER
1	B	769	LEU
1	B	817	TYR
1	B	971	LEU
1	C	702	SER
1	C	729	ILE
1	C	733	SER
1	C	739	LEU
1	C	769	LEU
1	C	817	TYR
1	C	819	LYS
1	C	858	ARG
1	C	874	SER
1	C	971	LEU
1	D	702	SER
1	D	733	SER
1	D	817	TYR
1	D	824	THR
1	D	912	GLN
1	D	971	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (8) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	767	ASN
1	A	868	ASN
1	A	961	ASN
1	B	767	ASN
1	B	868	ASN
1	C	767	ASN
1	C	868	ASN
1	D	998	ASN

### 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 carbohydrates in this entry.

## 5.6 Ligand geometry ⓘ

22 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	1101	-	4,4,4	1.16	1 (25%)	6,6,6	0.44	0
2	SO4	A	1102	-	4,4,4	0.35	0	6,6,6	0.77	0
2	SO4	A	1103	-	4,4,4	0.72	0	6,6,6	1.18	1 (16%)
2	SO4	A	1104	-	4,4,4	0.39	0	6,6,6	0.58	0
3	2YQ	A	1105	-	28,32,32	2.04	6 (21%)	33,48,48	2.25	9 (27%)
4	GOL	A	1106	-	5,5,5	0.45	0	5,5,5	0.82	0
4	GOL	A	1107	-	5,5,5	0.48	0	5,5,5	0.42	0
2	SO4	B	1101	-	4,4,4	0.51	0	6,6,6	0.88	0
2	SO4	B	1102	-	4,4,4	0.41	0	6,6,6	0.54	0
2	SO4	B	1103	-	4,4,4	0.23	0	6,6,6	0.42	0
2	SO4	B	1104	-	4,4,4	0.35	0	6,6,6	0.70	0
3	2YQ	B	1105	-	28,32,32	1.68	5 (17%)	33,48,48	1.71	5 (15%)
2	SO4	B	1106	-	4,4,4	0.46	0	6,6,6	0.39	0
2	SO4	C	1101	-	4,4,4	1.04	0	6,6,6	0.79	0
2	SO4	C	1102	-	4,4,4	0.40	0	6,6,6	0.36	0
2	SO4	C	1103	-	4,4,4	0.28	0	6,6,6	0.74	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
3	2YQ	C	1104	-	28,32,32	1.90	6 (21%)	33,48,48	1.97	10 (30%)
4	GOL	C	1105	-	5,5,5	0.86	0	5,5,5	1.41	2 (40%)
2	SO4	D	1101	-	4,4,4	0.46	0	6,6,6	0.60	0
2	SO4	D	1102	-	4,4,4	0.84	0	6,6,6	0.59	0
2	SO4	D	1103	-	4,4,4	0.42	0	6,6,6	0.26	0
3	2YQ	D	1104	-	28,32,32	1.87	9 (32%)	33,48,48	2.09	8 (24%)

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	SO4	A	1101	-	-	0/0/0/0	0/0/0/0
2	SO4	A	1102	-	-	0/0/0/0	0/0/0/0
2	SO4	A	1103	-	-	0/0/0/0	0/0/0/0
2	SO4	A	1104	-	-	0/0/0/0	0/0/0/0
3	2YQ	A	1105	-	-	0/4/20/20	0/4/5/5
4	GOL	A	1106	-	-	0/4/4/4	0/0/0/0
4	GOL	A	1107	-	-	0/4/4/4	0/0/0/0
2	SO4	B	1101	-	-	0/0/0/0	0/0/0/0
2	SO4	B	1102	-	-	0/0/0/0	0/0/0/0
2	SO4	B	1103	-	-	0/0/0/0	0/0/0/0
2	SO4	B	1104	-	-	0/0/0/0	0/0/0/0
3	2YQ	B	1105	-	-	0/4/20/20	0/4/5/5
2	SO4	B	1106	-	-	0/0/0/0	0/0/0/0
2	SO4	C	1101	-	-	0/0/0/0	0/0/0/0
2	SO4	C	1102	-	-	0/0/0/0	0/0/0/0
2	SO4	C	1103	-	-	0/0/0/0	0/0/0/0
3	2YQ	C	1104	-	-	0/4/20/20	0/4/5/5
4	GOL	C	1105	-	-	0/4/4/4	0/0/0/0
2	SO4	D	1101	-	-	0/0/0/0	0/0/0/0
2	SO4	D	1102	-	-	0/0/0/0	0/0/0/0
2	SO4	D	1103	-	-	0/0/0/0	0/0/0/0
3	2YQ	D	1104	-	-	0/4/20/20	0/4/5/5

All (27) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	1105	2YQ	C3-C2	-6.53	1.38	1.43
3	C	1104	2YQ	C3-C2	-6.34	1.38	1.43
3	A	1105	2YQ	C7-N3	-5.08	1.35	1.40

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	1105	2YQ	C7-N3	-3.81	1.36	1.40
3	D	1104	2YQ	C11-C6	-3.70	1.47	1.51
3	C	1104	2YQ	C7-N3	-3.45	1.36	1.40
3	A	1105	2YQ	C17-C5	-3.31	1.45	1.50
3	B	1105	2YQ	C3-C2	-3.04	1.40	1.43
3	D	1104	2YQ	C7-N3	-2.93	1.37	1.40
3	A	1105	2YQ	C7-C2	-2.77	1.38	1.42
3	D	1104	2YQ	C3-C2	-2.76	1.41	1.43
3	C	1104	2YQ	C17-C5	-2.33	1.47	1.50
3	D	1104	2YQ	C18-N5	2.03	1.36	1.32
2	A	1101	SO4	O2-S	2.10	1.57	1.45
3	C	1104	2YQ	C3-N1	2.23	1.35	1.33
3	A	1105	2YQ	C10-C9	2.41	1.40	1.36
3	D	1104	2YQ	C17-N6	2.44	1.38	1.34
3	A	1105	2YQ	N2-N1	2.44	1.40	1.34
3	B	1105	2YQ	C10-C9	2.58	1.40	1.36
3	C	1104	2YQ	N2-N1	2.60	1.41	1.34
3	B	1105	2YQ	N2-N1	2.71	1.41	1.34
3	D	1104	2YQ	C19-N4	2.78	1.50	1.47
3	D	1104	2YQ	C10-C9	3.22	1.41	1.36
3	D	1104	2YQ	C3-N1	3.26	1.37	1.33
3	C	1104	2YQ	C19-N4	3.50	1.51	1.47
3	D	1104	2YQ	N2-N1	3.64	1.43	1.34
3	B	1105	2YQ	C19-N4	3.72	1.51	1.47

All (35) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	D	1104	2YQ	C10-C9-C8	-5.55	120.19	124.11
3	D	1104	2YQ	C11-C6-C5	-5.16	103.91	112.11
3	A	1105	2YQ	C11-C6-C5	-4.96	104.23	112.11
3	C	1104	2YQ	C10-C9-C8	-3.99	121.30	124.11
3	A	1105	2YQ	C10-C9-C8	-3.36	121.74	124.11
3	B	1105	2YQ	C11-C6-C5	-3.31	106.86	112.11
3	C	1104	2YQ	C11-C6-C5	-3.26	106.93	112.11
3	B	1105	2YQ	C2-C3-N1	-3.20	119.31	121.50
3	A	1105	2YQ	C2-C3-N1	-3.06	119.40	121.50
3	C	1104	2YQ	C2-C3-N1	-2.61	119.71	121.50
4	C	1105	GOL	O2-C2-C3	-2.14	98.74	108.84
3	A	1105	2YQ	C15-C14-C13	-2.14	119.89	122.86
2	A	1103	SO4	O3-S-O2	2.00	120.30	109.26
3	C	1104	2YQ	C11-C6-N3	2.01	114.22	111.00

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	C	1104	2YQ	C5-C3-N1	2.02	120.31	118.18
3	A	1105	2YQ	C3-C5-C6	2.08	113.43	109.86
3	D	1104	2YQ	C9-C10-C1	2.08	120.71	119.40
3	D	1104	2YQ	C5-C3-N1	2.09	120.38	118.18
4	C	1105	GOL	C3-C2-C1	2.11	119.91	111.52
3	A	1105	2YQ	C5-C17-N6	2.13	129.05	124.96
3	B	1105	2YQ	C5-C3-N1	2.23	120.53	118.18
3	D	1104	2YQ	C3-C5-C6	2.39	113.97	109.86
3	D	1104	2YQ	F1-C9-C10	2.39	122.78	119.25
3	C	1104	2YQ	C9-C10-C1	2.41	120.92	119.40
3	C	1104	2YQ	C3-C5-C6	2.57	114.28	109.86
3	C	1104	2YQ	C5-C17-N4	2.71	128.25	123.88
3	A	1105	2YQ	C16-C15-C14	2.72	121.21	118.35
3	A	1105	2YQ	C3-N1-N2	2.72	121.83	119.73
3	D	1104	2YQ	C5-C17-N6	2.75	130.25	124.96
3	B	1105	2YQ	C3-N1-N2	3.49	122.43	119.73
3	C	1104	2YQ	C3-N1-N2	4.54	123.24	119.73
3	B	1105	2YQ	C2-C7-N3	5.11	120.61	117.41
3	C	1104	2YQ	C2-C7-N3	5.51	120.86	117.41
3	D	1104	2YQ	C2-C7-N3	6.03	121.19	117.41
3	A	1105	2YQ	C2-C7-N3	7.95	122.39	117.41

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

5 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	1102	SO4	1	0
2	C	1101	SO4	1	0
2	C	1102	SO4	1	0
2	C	1103	SO4	1	0
4	C	1105	GOL	1	0

## 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, 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	338/370 (91%)	-0.09	3 (0%) 84 91	17, 34, 61, 91	0
1	B	332/370 (89%)	0.20	13 (3%) 40 52	22, 46, 72, 86	0
1	C	346/370 (93%)	-0.01	10 (2%) 52 63	17, 34, 60, 81	0
1	D	326/370 (88%)	0.40	32 (9%) 8 12	26, 51, 80, 91	0
All	All	1342/1480 (90%)	0.12	58 (4%) 36 48	17, 40, 72, 91	0

All (58) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	939	SER	6.0
1	C	807	ASP	4.7
1	C	744	PHE	4.2
1	D	663	SER	4.0
1	B	786	SER	3.9
1	C	780	GLY	3.8
1	C	781	GLY	3.7
1	D	827	ASN	3.6
1	B	780	GLY	3.5
1	D	792	VAL	3.5
1	D	738	THR	3.4
1	D	965	ASP	3.4
1	B	687	VAL	3.3
1	D	749	PRO	3.3
1	B	807	ASP	3.2
1	B	768	LEU	3.2
1	A	727	SER	3.2
1	D	913	GLY	3.2
1	D	828	ALA	3.0
1	D	750	PRO	3.0
1	D	720	VAL	3.0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	D	960	ALA	2.9
1	D	959	SER	2.8
1	D	743	ASP	2.8
1	D	914	ASP	2.8
1	B	755	ALA	2.8
1	B	756	ASP	2.8
1	B	827	ASN	2.7
1	D	964	LEU	2.6
1	C	726	ASP	2.6
1	C	664	LYS	2.6
1	D	911	SER	2.5
1	D	731	ASP	2.5
1	C	981	ASP	2.5
1	D	789	PRO	2.5
1	A	748	LYS	2.4
1	D	919	ILE	2.4
1	D	912	GLN	2.4
1	B	694	GLN	2.4
1	D	825	THR	2.3
1	D	701	LEU	2.3
1	B	696	MET	2.3
1	B	914	ASP	2.3
1	D	757	SER	2.3
1	D	901	VAL	2.2
1	D	699	GLY	2.2
1	D	671	ASP	2.2
1	D	734	ASN	2.2
1	C	662	LYS	2.1
1	D	679	VAL	2.1
1	B	719	ALA	2.1
1	D	754	ASN	2.1
1	C	723	GLY	2.1
1	D	978	GLY	2.1
1	C	789	PRO	2.0
1	D	688	GLU	2.0
1	A	914	ASP	2.0
1	D	968	ASP	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 carbohydrates 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. 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, 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	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
2	SO4	B	1106	5/5	0.86	0.25	6.74	46,49,54,54	5
4	GOL	A	1107	6/6	0.81	0.17	2.54	54,60,62,65	0
2	SO4	A	1101	5/5	0.91	0.16	2.34	29,37,37,41	0
2	SO4	D	1102	5/5	0.87	0.17	2.30	50,51,61,66	0
2	SO4	B	1104	5/5	0.95	0.20	2.19	25,26,28,29	5
4	GOL	A	1106	6/6	0.91	0.21	1.68	42,53,56,62	0
2	SO4	D	1101	5/5	0.98	0.17	1.57	43,44,50,51	0
2	SO4	A	1102	5/5	0.94	0.17	1.13	30,31,33,34	5
2	SO4	B	1103	5/5	0.99	0.16	0.63	31,31,34,34	5
2	SO4	C	1101	5/5	0.92	0.13	0.39	36,40,53,54	0
2	SO4	A	1103	5/5	0.97	0.14	0.26	37,39,44,50	0
2	SO4	B	1101	5/5	0.88	0.16	0.20	47,54,56,60	0
2	SO4	D	1103	5/5	0.96	0.13	-0.05	37,37,40,42	5
2	SO4	A	1104	5/5	0.95	0.14	-0.23	28,28,29,32	5
2	SO4	C	1102	5/5	0.98	0.15	-0.32	55,59,64,67	0
4	GOL	C	1105	6/6	0.94	0.11	-0.63	28,31,35,36	0
3	2YQ	D	1104	28/28	0.95	0.11	-0.72	23,27,32,34	0
2	SO4	B	1102	5/5	0.98	0.11	-0.78	52,52,57,57	0
3	2YQ	C	1104	28/28	0.96	0.10	-0.97	17,19,23,28	0
3	2YQ	B	1105	28/28	0.97	0.09	-1.26	22,27,37,45	0
3	2YQ	A	1105	28/28	0.97	0.10	-1.39	18,19,23,24	0
2	SO4	C	1103	5/5	0.95	0.12	-	30,31,33,33	5

### 6.5 Other polymers ⓘ

There are no such residues in this entry.