



# wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 25, 2020 – 03:32 PM BST

PDB ID : 6QGA  
Title : Crystal structure of Ideonella sakaiensis MHETase bound to the non-hydrolyzable ligand MHETA  
Authors : Palm, G.J.; Reisky, L.; Boettcher, D.; Mueller, H.; Michels, E.A.P.; Walczak, C.; Berndt, L.; Weiss, M.S.; Bornscheuer, U.T.; Weber, G.  
Deposited on : 2019-01-10  
Resolution : 2.10 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 following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.13
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.13

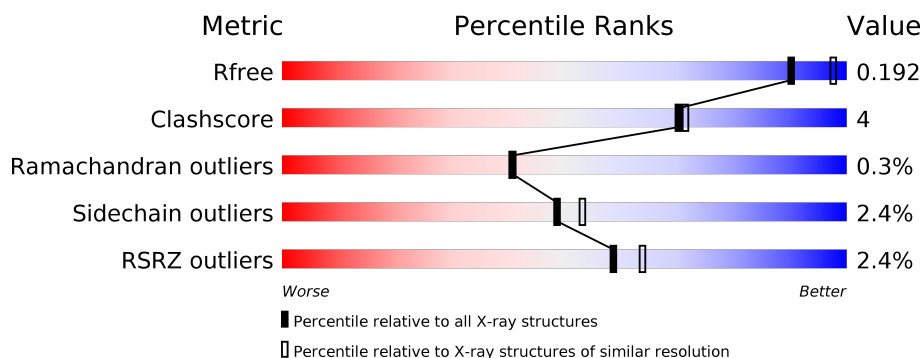
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.10 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))
$R_{free}$	130704	5197 (2.10-2.10)
Clashscore	141614	5710 (2.10-2.10)
Ramachandran outliers	138981	5647 (2.10-2.10)
Sidechain outliers	138945	5648 (2.10-2.10)
RSRZ outliers	127900	5083 (2.10-2.10)

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 on 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	596	<div> <div>5%</div> <div> <div></div> <div>85%</div> <div>8%</div> <div>6%</div> </div> </div>
1	B	596	<div> <div>2%</div> <div> <div></div> <div>86%</div> <div>7%</div> <div>6%</div> </div> </div>
1	C	596	<div> <div>%</div> <div> <div></div> <div>84%</div> <div>9%</div> <div>7%</div> </div> </div>
1	D	596	<div> <div>3%</div> <div> <div></div> <div>86%</div> <div>8%</div> <div>6%</div> </div> </div>
1	E	596	<div> <div>2%</div> <div> <div></div> <div>87%</div> <div>6%</div> <div>6%</div> </div> </div>
1	F	596	<div> <div>%</div> <div> <div></div> <div>86%</div> <div>7%</div> <div>6%</div> </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 criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
6	CL	B	707	-	-	X	-
6	CL	C	704	-	-	X	-
6	CL	E	706	-	-	X	-

## 2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 27633 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 Mono(2-hydroxyethyl) terephthalate hydrolase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	558	Total	C	N	O	S	0	0	0
			4121	2581	724	788	28			
1	B	558	Total	C	N	O	S	0	1	0
			4136	2590	728	790	28			
1	C	557	Total	C	N	O	S	5	1	0
			4122	2584	723	786	29			
1	D	559	Total	C	N	O	S	0	0	0
			4129	2586	725	790	28			
1	E	558	Total	C	N	O	S	0	0	0
			4125	2584	724	789	28			
1	F	559	Total	C	N	O	S	0	1	0
			4134	2590	725	790	29			

There are 72 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	8	MET	-	initiating methionine	UNP A0A0K8P8E7
A	9	ASN	-	expression tag	UNP A0A0K8P8E7
A	10	HIS	-	expression tag	UNP A0A0K8P8E7
A	11	LYS	-	expression tag	UNP A0A0K8P8E7
A	12	VAL	-	expression tag	UNP A0A0K8P8E7
A	13	HIS	-	expression tag	UNP A0A0K8P8E7
A	14	HIS	-	expression tag	UNP A0A0K8P8E7
A	15	HIS	-	expression tag	UNP A0A0K8P8E7
A	16	HIS	-	expression tag	UNP A0A0K8P8E7
A	17	HIS	-	expression tag	UNP A0A0K8P8E7
A	18	HIS	-	expression tag	UNP A0A0K8P8E7
A	19	MET	-	expression tag	UNP A0A0K8P8E7
B	8	MET	-	initiating methionine	UNP A0A0K8P8E7
B	9	ASN	-	expression tag	UNP A0A0K8P8E7
B	10	HIS	-	expression tag	UNP A0A0K8P8E7
B	11	LYS	-	expression tag	UNP A0A0K8P8E7
B	12	VAL	-	expression tag	UNP A0A0K8P8E7

*Continued on next page...*

*Continued from previous page...*

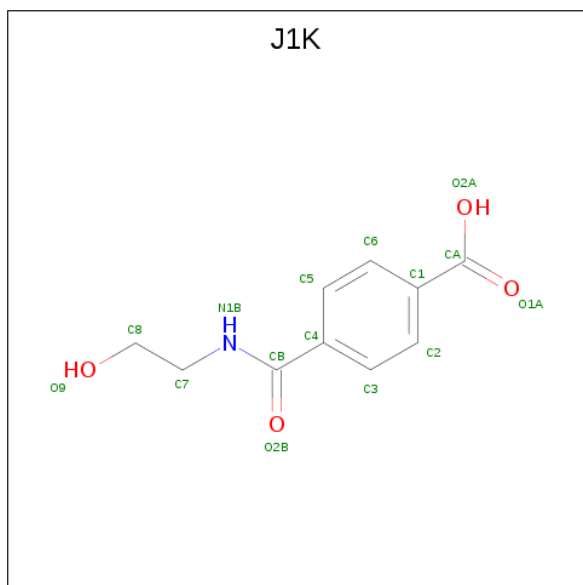
Chain	Residue	Modelled	Actual	Comment	Reference
B	13	HIS	-	expression tag	UNP A0A0K8P8E7
B	14	HIS	-	expression tag	UNP A0A0K8P8E7
B	15	HIS	-	expression tag	UNP A0A0K8P8E7
B	16	HIS	-	expression tag	UNP A0A0K8P8E7
B	17	HIS	-	expression tag	UNP A0A0K8P8E7
B	18	HIS	-	expression tag	UNP A0A0K8P8E7
B	19	MET	-	expression tag	UNP A0A0K8P8E7
C	8	MET	-	initiating methionine	UNP A0A0K8P8E7
C	9	ASN	-	expression tag	UNP A0A0K8P8E7
C	10	HIS	-	expression tag	UNP A0A0K8P8E7
C	11	LYS	-	expression tag	UNP A0A0K8P8E7
C	12	VAL	-	expression tag	UNP A0A0K8P8E7
C	13	HIS	-	expression tag	UNP A0A0K8P8E7
C	14	HIS	-	expression tag	UNP A0A0K8P8E7
C	15	HIS	-	expression tag	UNP A0A0K8P8E7
C	16	HIS	-	expression tag	UNP A0A0K8P8E7
C	17	HIS	-	expression tag	UNP A0A0K8P8E7
C	18	HIS	-	expression tag	UNP A0A0K8P8E7
C	19	MET	-	expression tag	UNP A0A0K8P8E7
D	8	MET	-	initiating methionine	UNP A0A0K8P8E7
D	9	ASN	-	expression tag	UNP A0A0K8P8E7
D	10	HIS	-	expression tag	UNP A0A0K8P8E7
D	11	LYS	-	expression tag	UNP A0A0K8P8E7
D	12	VAL	-	expression tag	UNP A0A0K8P8E7
D	13	HIS	-	expression tag	UNP A0A0K8P8E7
D	14	HIS	-	expression tag	UNP A0A0K8P8E7
D	15	HIS	-	expression tag	UNP A0A0K8P8E7
D	16	HIS	-	expression tag	UNP A0A0K8P8E7
D	17	HIS	-	expression tag	UNP A0A0K8P8E7
D	18	HIS	-	expression tag	UNP A0A0K8P8E7
D	19	MET	-	expression tag	UNP A0A0K8P8E7
E	8	MET	-	initiating methionine	UNP A0A0K8P8E7
E	9	ASN	-	expression tag	UNP A0A0K8P8E7
E	10	HIS	-	expression tag	UNP A0A0K8P8E7
E	11	LYS	-	expression tag	UNP A0A0K8P8E7
E	12	VAL	-	expression tag	UNP A0A0K8P8E7
E	13	HIS	-	expression tag	UNP A0A0K8P8E7
E	14	HIS	-	expression tag	UNP A0A0K8P8E7
E	15	HIS	-	expression tag	UNP A0A0K8P8E7
E	16	HIS	-	expression tag	UNP A0A0K8P8E7
E	17	HIS	-	expression tag	UNP A0A0K8P8E7
E	18	HIS	-	expression tag	UNP A0A0K8P8E7

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
E	19	MET	-	expression tag	UNP A0A0K8P8E7
F	8	MET	-	initiating methionine	UNP A0A0K8P8E7
F	9	ASN	-	expression tag	UNP A0A0K8P8E7
F	10	HIS	-	expression tag	UNP A0A0K8P8E7
F	11	LYS	-	expression tag	UNP A0A0K8P8E7
F	12	VAL	-	expression tag	UNP A0A0K8P8E7
F	13	HIS	-	expression tag	UNP A0A0K8P8E7
F	14	HIS	-	expression tag	UNP A0A0K8P8E7
F	15	HIS	-	expression tag	UNP A0A0K8P8E7
F	16	HIS	-	expression tag	UNP A0A0K8P8E7
F	17	HIS	-	expression tag	UNP A0A0K8P8E7
F	18	HIS	-	expression tag	UNP A0A0K8P8E7
F	19	MET	-	expression tag	UNP A0A0K8P8E7

- Molecule 2 is 4-(2-hydroxyethylcarbamoyl)benzoic acid (three-letter code: J1K) (formula:  $C_{10}H_{11}NO_4$ ).



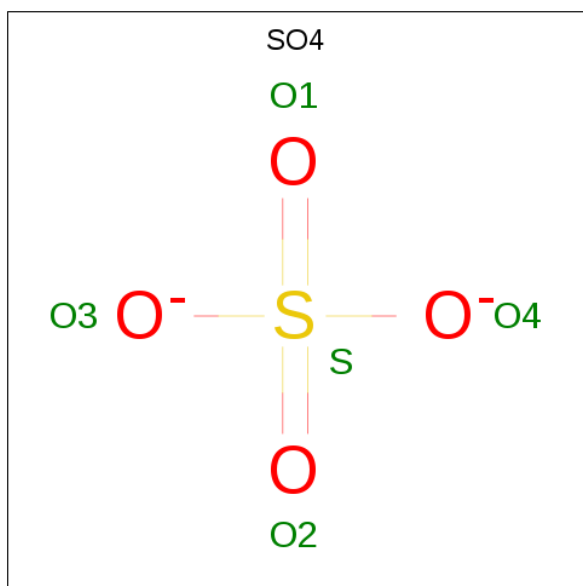
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	N	O	0	0
			15	10	1	4		
2	B	1	Total	C	N	O	0	0
			15	10	1	4		
2	C	1	Total	C	N	O	0	0
			15	10	1	4		
2	C	1	Total	C	N	O	0	0
			15	10	1	4		

Continued on next page...

*Continued from previous page...*

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	D	1	Total	C	N	O	0	0
			15	10	1	4		
2	E	1	Total	C	N	O	0	0
			15	10	1	4		
2	E	1	Total	C	N	O	0	0
			15	10	1	4		
2	E	1	Total	C	N	O	0	0
			15	10	1	4		
2	F	1	Total	C	N	O	0	0
			15	10	1	4		
2	F	1	Total	C	N	O	0	0
			15	10	1	4		

- Molecule 3 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



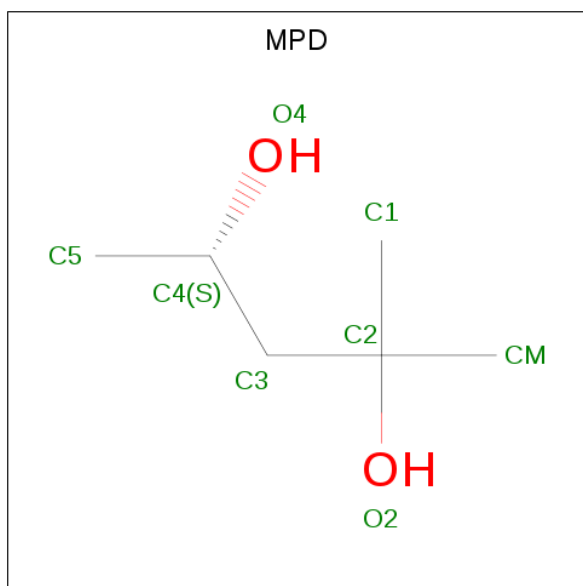
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	O	S	0	0
			5	4	1		
3	A	1	Total	O	S	0	0
			5	4	1		
3	A	1	Total	O	S	0	0
			5	4	1		
3	B	1	Total	O	S	0	0
			5	4	1		
3	B	1	Total	O	S	0	0
			5	4	1		

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	B	1	Total	O	S	0	0
			5	4	1		
3	E	1	Total	O	S	0	0
			5	4	1		
3	F	1	Total	O	S	0	0
			5	4	1		
3	F	1	Total	O	S	0	0
			5	4	1		
3	F	1	Total	O	S	0	0
			5	4	1		

- Molecule 4 is (4S)-2-METHYL-2,4-PENTANEDIOL (three-letter code: MPD) (formula:  $C_6H_{14}O_2$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			8	6	2		
4	B	1	Total	C	O	0	0
			8	6	2		

- Molecule 5 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	D	1	Total	Ca	0	0
			1	1		
5	E	1	Total	Ca	0	0
			1	1		

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	B	1	Total 1	Ca 1	0	0
5	C	1	Total 1	Ca 1	0	0
5	A	1	Total 1	Ca 1	0	0
5	F	1	Total 1	Ca 1	0	0

- Molecule 6 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	B	1	Total 1	Cl 1	0	0
6	C	2	Total 2	Cl 2	0	0
6	E	1	Total 1	Cl 1	0	0

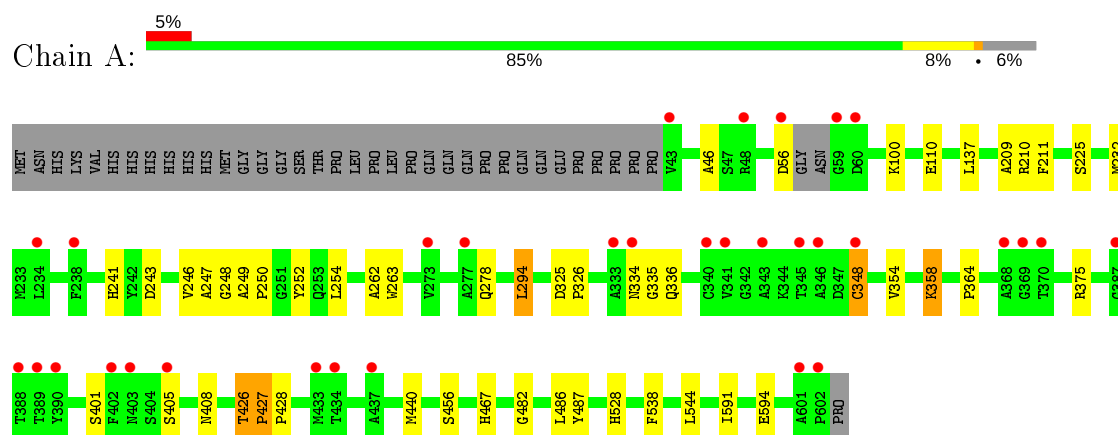
- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	390	Total 390	O 390	0	0
7	B	488	Total 488	O 488	0	0
7	C	424	Total 424	O 424	0	0
7	D	412	Total 412	O 412	0	0
7	E	490	Total 490	O 490	0	0
7	F	436	Total 436	O 436	0	0

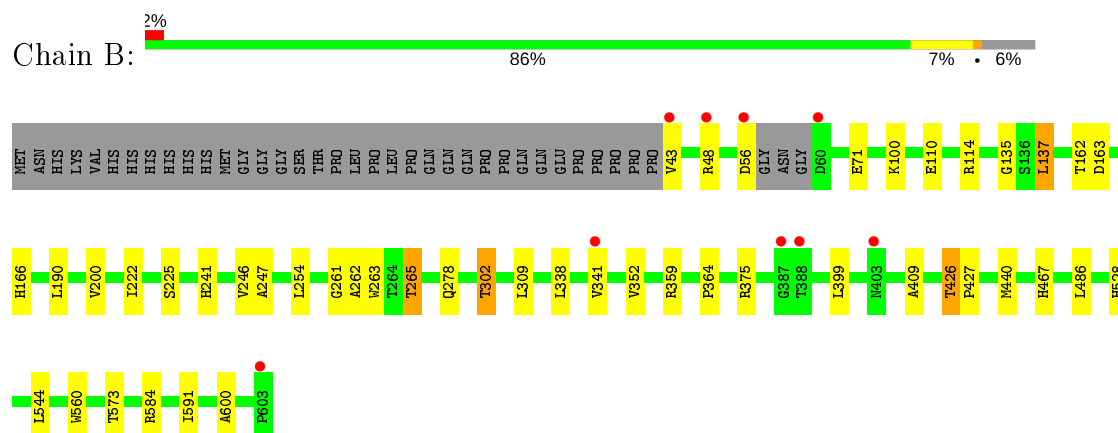
### 3 Residue-property plots [i](#)

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.

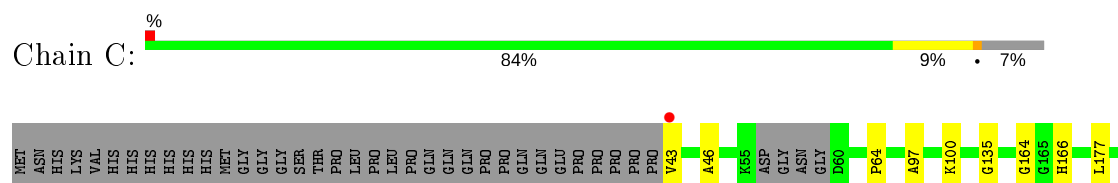
- Molecule 1: Mono(2-hydroxyethyl) terephthalate hydrolase

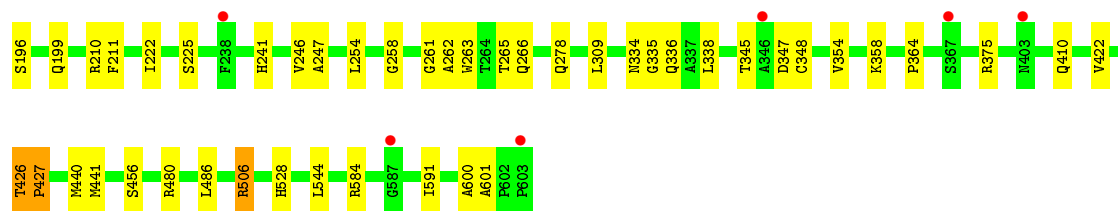


- Molecule 1: Mono(2-hydroxyethyl) terephthalate hydrolase

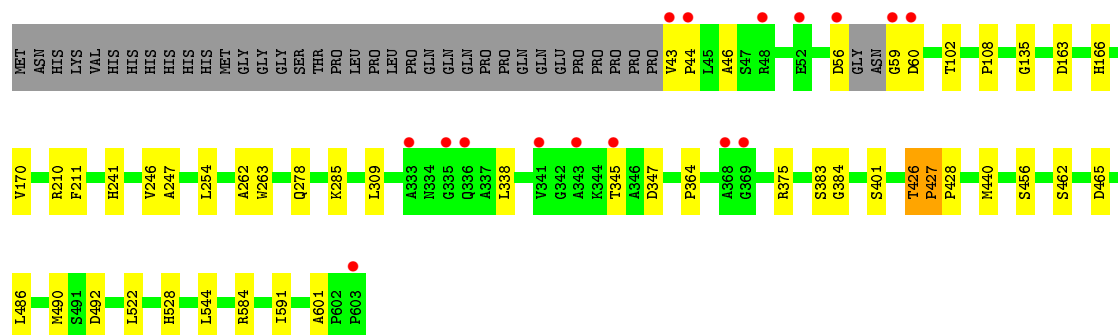
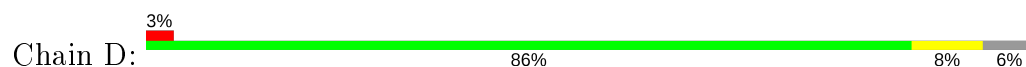


- Molecule 1: Mono(2-hydroxyethyl) terephthalate hydrolase

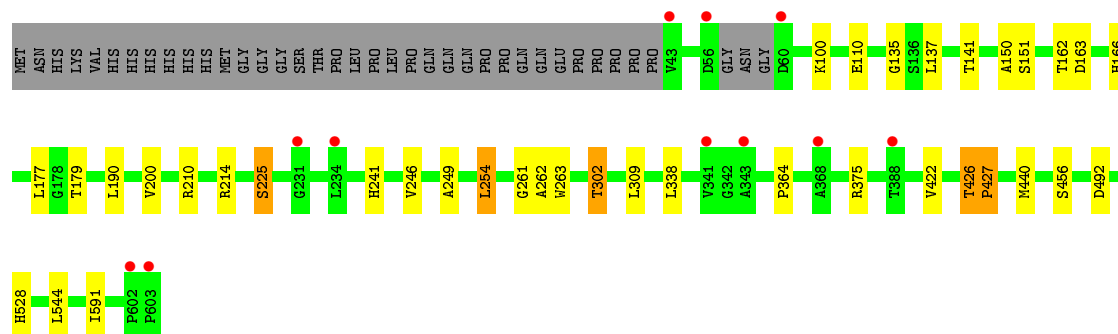
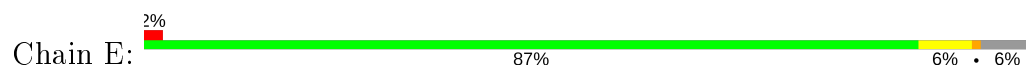




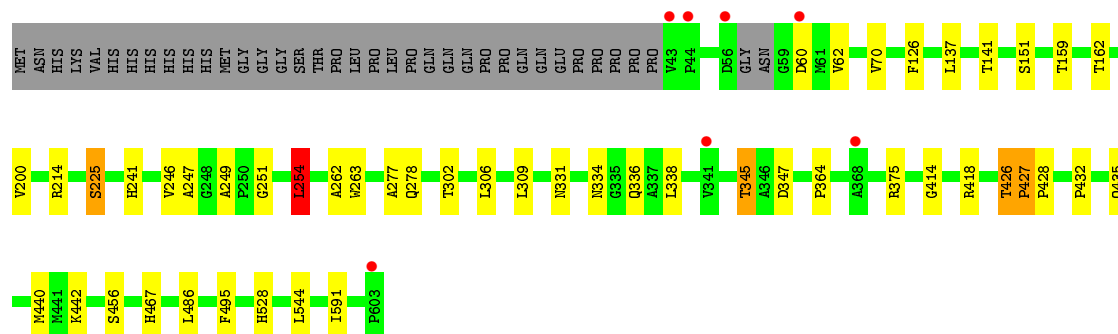
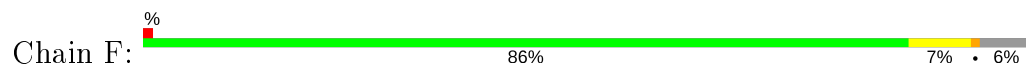
• Molecule 1: Mono(2-hydroxyethyl) terephthalate hydrolase



• Molecule 1: Mono(2-hydroxyethyl) terephthalate hydrolase



• Molecule 1: Mono(2-hydroxyethyl) terephthalate hydrolase



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	112.15Å 183.65Å 246.75Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.11 – 2.10 49.11 – 2.10	Depositor EDS
% Data completeness (in resolution range)	98.9 (49.11-2.10) 98.9 (49.11-2.10)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.20	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.31 (at 2.10Å)	Xtriage
Refinement program	PHENIX (1.10.1 _2155: ???)	Depositor
R, $R_{free}$	0.173 , 0.198 0.173 , 0.192	Depositor DCC
$R_{free}$ test set	3081 reflections (1.05%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	28.0	Xtriage
Anisotropy	0.605	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 44.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.51$ , $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	27633	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	29.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 8.98% 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: SO4, MPD, CA, J1K, 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	0.46	1/4227 (0.0%)	0.60	1/5754 (0.0%)
1	B	0.50	0/4243	0.66	3/5775 (0.1%)
1	C	0.47	0/4232	0.64	3/5760 (0.1%)
1	D	0.47	0/4236	0.61	2/5766 (0.0%)
1	E	0.48	0/4232	0.64	3/5761 (0.1%)
1	F	0.46	0/4244	0.62	2/5776 (0.0%)
All	All	0.48	1/25414 (0.0%)	0.63	14/34592 (0.0%)

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	1
1	B	0	1
1	C	0	1
1	D	0	1
1	E	0	1
1	F	0	1
All	All	0	6

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	348	CYS	CB-SG	-6.06	1.72	1.82

The worst 5 of 14 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	426	THR	C-N-CD	-12.82	92.41	120.60
1	B	426	THR	C-N-CD	-12.28	93.59	120.60
1	F	426	THR	C-N-CD	-11.75	94.75	120.60
1	E	426	THR	C-N-CD	-11.30	95.74	120.60
1	D	426	THR	C-N-CD	-10.62	97.23	120.60

There are no chirality outliers.

5 of 6 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	426	THR	Peptide
1	B	426	THR	Peptide
1	C	426	THR	Peptide
1	D	426	THR	Peptide
1	E	426	THR	Peptide

## 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	4121	0	3912	27	0
1	B	4136	0	3928	27	0
1	C	4122	0	3921	39	0
1	D	4129	0	3919	26	0
1	E	4125	0	3916	22	0
1	F	4134	0	3928	29	0
2	A	15	0	0	0	0
2	B	15	0	0	0	0
2	C	30	0	0	1	0
2	D	15	0	0	1	0
2	E	45	0	0	3	0
2	F	30	0	0	1	0
3	A	15	0	0	1	0
3	B	15	0	0	1	0
3	E	5	0	0	1	0
3	F	15	0	0	0	0
4	A	8	0	14	1	0
4	B	8	0	14	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	A	1	0	0	0	0
5	B	1	0	0	0	0
5	C	1	0	0	0	0
5	D	1	0	0	0	0
5	E	1	0	0	0	0
5	F	1	0	0	0	0
6	B	1	0	0	3	0
6	C	2	0	0	5	0
6	E	1	0	0	3	0
7	A	390	0	0	4	2
7	B	488	0	0	6	2
7	C	424	0	0	4	2
7	D	412	0	0	1	1
7	E	490	0	0	3	4
7	F	436	0	0	3	2
All	All	27633	0	23552	172	7

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 4.

The worst 5 of 172 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:D:345:THR:HG22	1:D:347:ASP:H	1.23	0.99
1:A:232:MET:HE2	1:A:250:PRO:HB2	1.45	0.98
1:A:232:MET:HE3	1:A:252:TYR:HB3	1.44	0.98
1:D:490:MET:HE3	1:D:522:LEU:HB3	1.45	0.97
1:C:345:THR:HG22	1:C:347:ASP:H	1.33	0.92

The worst 5 of 7 symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
7:A:1171:HOH:O	7:E:1063:HOH:O[4_545]	2.00	0.20
7:D:1202:HOH:O	7:E:1252:HOH:O[2_554]	2.09	0.11
7:B:1273:HOH:O	7:E:1196:HOH:O[4_445]	2.12	0.08
7:C:1220:HOH:O	7:F:1105:HOH:O[2_455]	2.12	0.08
7:C:1191:HOH:O	7:F:1187:HOH:O[2_455]	2.12	0.08

## 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	554/596 (93%)	536 (97%)	16 (3%)	2 (0%)	34	32
1	B	555/596 (93%)	538 (97%)	16 (3%)	1 (0%)	47	49
1	C	554/596 (93%)	534 (96%)	18 (3%)	2 (0%)	34	32
1	D	555/596 (93%)	541 (98%)	13 (2%)	1 (0%)	47	49
1	E	554/596 (93%)	537 (97%)	15 (3%)	2 (0%)	34	32
1	F	556/596 (93%)	536 (96%)	18 (3%)	2 (0%)	34	32
All	All	3328/3576 (93%)	3222 (97%)	96 (3%)	10 (0%)	41	41

5 of 10 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	427	PRO
1	B	427	PRO
1	C	427	PRO
1	D	427	PRO
1	E	427	PRO

### 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	405/439 (92%)	393 (97%)	12 (3%)	41	44
1	B	407/439 (93%)	394 (97%)	13 (3%)	39	41
1	C	406/439 (92%)	400 (98%)	6 (2%)	65	71

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	D	406/439 (92%)	397 (98%)	9 (2%)	52	57
1	E	406/439 (92%)	398 (98%)	8 (2%)	55	60
1	F	407/439 (93%)	396 (97%)	11 (3%)	44	48
All	All	2437/2634 (92%)	2378 (98%)	59 (2%)	49	53

5 of 59 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	241	HIS
1	D	241	HIS
1	F	302	THR
1	C	254	LEU
1	C	338	LEU

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

Mol	Chain	Res	Type
1	B	528	HIS
1	C	528	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 monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 32 ligands modelled in this entry, 10 are monoatomic - leaving 22 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
3	SO4	E	704	-	4,4,4	0.18	0	6,6,6	0.14	0
2	J1K	A	701	-	13,15,15	0.83	1 (7%)	16,19,19	1.15	1 (6%)
2	J1K	E	703	-	13,15,15	1.09	1 (7%)	16,19,19	1.39	2 (12%)
3	SO4	F	703	-	4,4,4	0.16	0	6,6,6	0.18	0
3	SO4	A	704	-	4,4,4	0.17	0	6,6,6	0.15	0
2	J1K	D	701	-	13,15,15	0.98	1 (7%)	16,19,19	1.21	1 (6%)
4	MPD	B	705	-	7,7,7	0.37	0	9,10,10	0.68	0
2	J1K	B	701	-	13,15,15	0.59	0	16,19,19	1.53	2 (12%)
2	J1K	E	701	-	13,15,15	1.40	1 (7%)	16,19,19	1.10	0
3	SO4	B	703	-	4,4,4	0.19	0	6,6,6	0.12	0
2	J1K	C	701	-	13,15,15	0.37	0	16,19,19	1.29	2 (12%)
2	J1K	E	702	-	13,15,15	0.87	1 (7%)	16,19,19	0.72	0
3	SO4	F	705	-	4,4,4	0.15	0	6,6,6	0.12	0
3	SO4	A	702	-	4,4,4	0.20	0	6,6,6	0.15	0
3	SO4	A	703	-	4,4,4	0.17	0	6,6,6	0.17	0
2	J1K	F	701	-	13,15,15	0.95	1 (7%)	16,19,19	0.86	0
2	J1K	C	702	-	13,15,15	0.82	1 (7%)	16,19,19	1.39	2 (12%)
4	MPD	A	705	-	7,7,7	0.26	0	9,10,10	0.59	0
3	SO4	B	702	-	4,4,4	0.17	0	6,6,6	0.11	0
2	J1K	F	702	-	13,15,15	0.54	0	16,19,19	1.31	1 (6%)
3	SO4	F	704	-	4,4,4	0.18	0	6,6,6	0.13	0
3	SO4	B	704	-	4,4,4	0.15	0	6,6,6	0.14	0

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	J1K	C	702	-	-	3/8/12/12	0/1/1/1
2	J1K	E	702	-	-	0/8/12/12	0/1/1/1
2	J1K	A	701	-	-	1/8/12/12	0/1/1/1
4	MPD	A	705	-	-	1/5/5/5	-
2	J1K	E	703	-	-	1/8/12/12	0/1/1/1

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	J1K	F	702	-	-	2/8/12/12	0/1/1/1
4	MPD	B	705	-	-	4/5/5/5	-
2	J1K	B	701	-	-	3/8/12/12	0/1/1/1
2	J1K	E	701	-	-	0/8/12/12	0/1/1/1
2	J1K	C	701	-	-	3/8/12/12	0/1/1/1
2	J1K	F	701	-	-	3/8/12/12	0/1/1/1
2	J1K	D	701	-	-	3/8/12/12	0/1/1/1

The worst 5 of 7 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	E	701	J1K	C1-CA	4.67	1.52	1.47
2	E	703	J1K	C1-CA	3.40	1.50	1.47
2	D	701	J1K	C1-CA	2.97	1.50	1.47
2	E	702	J1K	C1-CA	2.90	1.50	1.47
2	F	701	J1K	C1-CA	2.62	1.50	1.47

The worst 5 of 11 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	F	702	J1K	C2-C1-CA	-3.02	116.31	120.37
2	C	702	J1K	O2B-CB-C4	2.97	126.23	120.94
2	E	703	J1K	O2B-CB-C4	2.88	126.07	120.94
2	B	701	J1K	C7-N1B-CB	-2.65	116.05	122.08
2	E	703	J1K	C7-N1B-CB	-2.56	116.25	122.08

There are no chirality outliers.

5 of 24 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	701	J1K	N1B-C7-C8-O9
2	E	703	J1K	N1B-C7-C8-O9
2	D	701	J1K	C4-CB-N1B-C7
4	B	705	MPD	C1-C2-C3-C4
2	B	701	J1K	N1B-C7-C8-O9

There are no ring outliers.

9 monomers are involved in 10 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	E	704	SO4	1	0
2	E	703	J1K	1	0
3	A	704	SO4	1	0
2	D	701	J1K	1	0
2	E	701	J1K	2	0
3	B	703	SO4	1	0
2	C	701	J1K	1	0
2	F	701	J1K	1	0
4	A	705	MPD	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	558/596 (93%)	0.06	32 (5%) 23 29	20, 29, 51, 77	0
1	B	558/596 (93%)	-0.27	9 (1%) 72 75	18, 24, 36, 84	0
1	C	557/596 (93%)	-0.17	7 (1%) 77 80	19, 27, 44, 62	0
1	D	559/596 (93%)	-0.15	16 (2%) 51 57	19, 27, 44, 76	0
1	E	558/596 (93%)	-0.11	11 (1%) 65 69	18, 24, 38, 93	0
1	F	559/596 (93%)	-0.21	7 (1%) 77 80	20, 27, 42, 78	0
All	All	3349/3576 (93%)	-0.14	82 (2%) 59 64	18, 26, 44, 93	0

The worst 5 of 82 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	56	ASP	5.6
1	D	341	VAL	5.2
1	E	56	ASP	4.7
1	F	43	VAL	4.6
1	A	341	VAL	4.5

### 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 ⓘ

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
2	J1K	E	701	15/15	0.74	0.21	34,48,59,61	0
4	MPD	A	705	8/8	0.78	0.28	54,62,64,69	0
2	J1K	F	702	15/15	0.80	0.21	54,55,61,62	0
2	J1K	C	701	15/15	0.81	0.19	40,48,63,64	0
3	SO4	F	705	5/5	0.85	0.20	110,111,113,114	0
2	J1K	E	702	15/15	0.86	0.14	44,46,55,60	0
3	SO4	F	704	5/5	0.86	0.31	102,103,106,108	0
3	SO4	B	704	5/5	0.90	0.12	78,82,85,86	0
3	SO4	A	703	5/5	0.91	0.20	92,93,97,97	0
3	SO4	A	704	5/5	0.91	0.23	100,100,105,106	0
3	SO4	B	702	5/5	0.91	0.31	102,102,107,109	0
3	SO4	E	704	5/5	0.92	0.23	97,98,99,103	0
3	SO4	A	702	5/5	0.92	0.14	75,77,78,79	0
3	SO4	F	703	5/5	0.93	0.13	85,86,86,88	0
3	SO4	B	703	5/5	0.93	0.18	90,92,96,98	0
2	J1K	C	702	15/15	0.95	0.11	16,21,51,56	0
4	MPD	B	705	8/8	0.95	0.14	26,31,36,37	0
2	J1K	D	701	15/15	0.95	0.11	19,24,38,44	0
2	J1K	B	701	15/15	0.96	0.10	12,17,42,51	0
2	J1K	A	701	15/15	0.96	0.11	23,25,51,59	0
6	CL	B	707	1/1	0.96	0.17	61,61,61,61	0
2	J1K	F	701	15/15	0.97	0.10	17,21,39,45	0
6	CL	E	706	1/1	0.97	0.09	51,51,51,51	0
6	CL	C	704	1/1	0.97	0.16	58,58,58,58	0
6	CL	C	705	1/1	0.97	0.20	66,66,66,66	0
2	J1K	E	703	15/15	0.97	0.14	15,20,46,50	0
5	CA	C	703	1/1	0.98	0.03	30,30,30,30	0
5	CA	A	706	1/1	0.99	0.04	25,25,25,25	0
5	CA	F	706	1/1	0.99	0.04	24,24,24,24	0
5	CA	E	705	1/1	0.99	0.05	26,26,26,26	0
5	CA	D	702	1/1	0.99	0.03	21,21,21,21	0
5	CA	B	706	1/1	1.00	0.03	23,23,23,23	0

## 6.5 Other polymers ⓘ

There are no such residues in this entry.