



# Full wwPDB X-ray Structure Validation Report

Oct 13, 2014 – 04:53 PM EDT

PDB ID : 4K6T  
Title : Crystal structure of Ad37 fiber knob in complex with trivalent sialic acid inhibitor ME0385  
Authors : Stehle, T.; Bauer, J.  
Deposited on : 2013-04-16  
Resolution : 2.00 Å(reported)

This is a full wwPDB 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/ValidationPDFNotes.html>

---

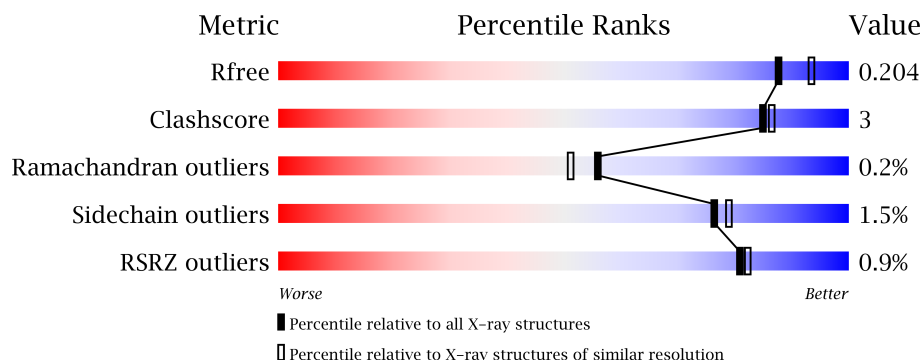
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.16 November 2013  
Xtriage (Phenix) : dev-1439  
EDS : stable23828  
Percentile statistics : 21963  
Refmac : 5.8.0049  
CCP4 : 6.3.0 (Settle)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)  
Validation Pipeline (wwPDB-VP) : stable23828

# 1 Overall quality at a glance

The reported resolution of this entry is 2.00 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))
$R_{free}$	66092	4888 (2.00-2.00)
Clashscore	79885	6188 (2.00-2.00)
Ramachandran outliers	78287	6102 (2.00-2.00)
Sidechain outliers	78261	6100 (2.00-2.00)
RSRZ outliers	66119	4890 (2.00-2.00)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	194	
1	B	194	
1	C	194	
1	E	194	
1	F	194	
1	G	194	

The following table lists non-polymeric compounds that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Geometry	Electron density
2	SIA	E	403	-	X
2	SIA	F	403	-	X

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Geometry	Electron density
2	SIA	G	404	-	X
4	ACT	A	406	-	X
4	ACT	E	404	-	X
4	ACT	E	405	-	X
4	ACT	E	406	-	X
4	ACT	G	405	-	X
5	EDO	E	408	-	X
5	EDO	E	409	-	X
5	EDO	E	410	-	X
5	EDO	F	405	-	X
5	EDO	F	407	-	X
8	MG	C	404	-	X
9	GOL	G	406	-	X

## 2 Entry composition

There are 11 unique types of molecules in this entry. The entry contains 9858 atoms, of which 0 are hydrogen and 0 are deuterium.

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 Fiber protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	186	Total	C	N	O	S	0	5	0
			1466	943	231	286	6			
1	B	186	Total	C	N	O	S	0	4	0
			1460	941	231	283	5			
1	C	182	Total	C	N	O	S	0	8	0
			1459	940	230	284	5			
1	E	185	Total	C	N	O	S	0	4	0
			1458	939	231	282	6			
1	F	186	Total	C	N	O	S	0	3	0
			1448	927	232	284	5			
1	G	185	Total	C	N	O	S	0	6	0
			1469	946	231	287	5			

There are 30 discrepancies between the modelled and reference sequences:

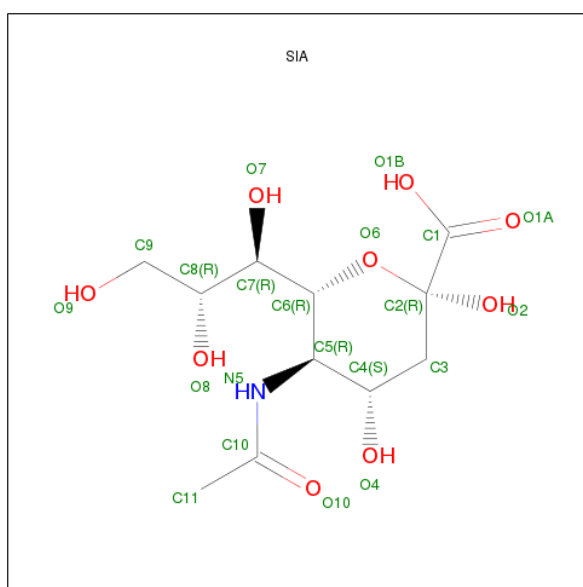
Chain	Residue	Modelled	Actual	Comment	Reference
A	172	GLY	-	EXPRESSION TAG	UNP Q64823
A	173	ALA	-	EXPRESSION TAG	UNP Q64823
A	174	MET	-	EXPRESSION TAG	UNP Q64823
A	175	GLY	-	EXPRESSION TAG	UNP Q64823
A	176	SER	-	EXPRESSION TAG	UNP Q64823
B	172	GLY	-	EXPRESSION TAG	UNP Q64823
B	173	ALA	-	EXPRESSION TAG	UNP Q64823
B	174	MET	-	EXPRESSION TAG	UNP Q64823
B	175	GLY	-	EXPRESSION TAG	UNP Q64823
B	176	SER	-	EXPRESSION TAG	UNP Q64823
C	172	GLY	-	EXPRESSION TAG	UNP Q64823
C	173	ALA	-	EXPRESSION TAG	UNP Q64823
C	174	MET	-	EXPRESSION TAG	UNP Q64823
C	175	GLY	-	EXPRESSION TAG	UNP Q64823
C	176	SER	-	EXPRESSION TAG	UNP Q64823
E	172	GLY	-	EXPRESSION TAG	UNP Q64823
E	173	ALA	-	EXPRESSION TAG	UNP Q64823

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
E	174	MET	-	EXPRESSION TAG	UNP Q64823
E	175	GLY	-	EXPRESSION TAG	UNP Q64823
E	176	SER	-	EXPRESSION TAG	UNP Q64823
F	172	GLY	-	EXPRESSION TAG	UNP Q64823
F	173	ALA	-	EXPRESSION TAG	UNP Q64823
F	174	MET	-	EXPRESSION TAG	UNP Q64823
F	175	GLY	-	EXPRESSION TAG	UNP Q64823
F	176	SER	-	EXPRESSION TAG	UNP Q64823
G	172	GLY	-	EXPRESSION TAG	UNP Q64823
G	173	ALA	-	EXPRESSION TAG	UNP Q64823
G	174	MET	-	EXPRESSION TAG	UNP Q64823
G	175	GLY	-	EXPRESSION TAG	UNP Q64823
G	176	SER	-	EXPRESSION TAG	UNP Q64823

- Molecule 2 is SUGAR (O-SIALIC ACID) (three-letter code: SIA) (formula:  $C_{11}H_{19}NO_9$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	N	O	0	0
			20	11	1	8		
2	B	1	Total	C	N	O	0	0
			20	11	1	8		
2	C	1	Total	C	N	O	0	0
			20	11	1	8		
2	E	1	Total	C	N	O	0	0
			20	11	1	8		
2	F	1	Total	C	N	O	0	0
			20	11	1	8		

Continued on next page...

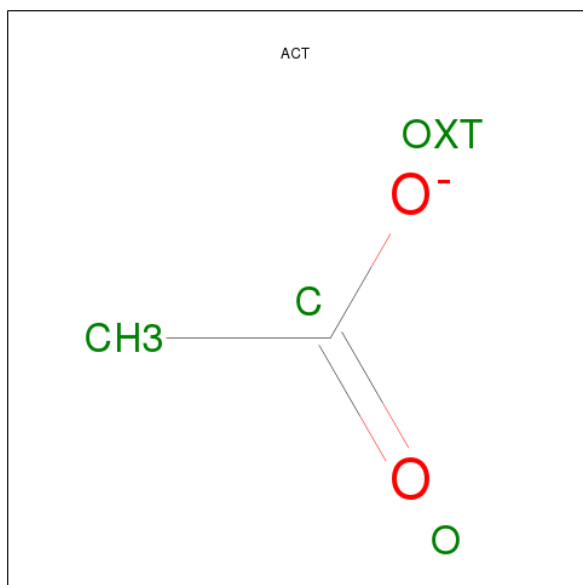
*Continued from previous page...*

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	G	1	Total	C	N	O	0	0
			20	11	1	8		

- Molecule 3 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	G	3	Total	Zn	0	0
			3	3		
3	E	2	Total	Zn	0	0
			2	2		
3	B	3	Total	Zn	0	0
			3	3		
3	C	1	Total	Zn	0	0
			1	1		
3	A	3	Total	Zn	0	0
			3	3		
3	F	2	Total	Zn	0	0
			2	2		

- Molecule 4 is ACETATE ION (three-letter code: ACT) (formula: C<sub>2</sub>H<sub>3</sub>O<sub>2</sub>).



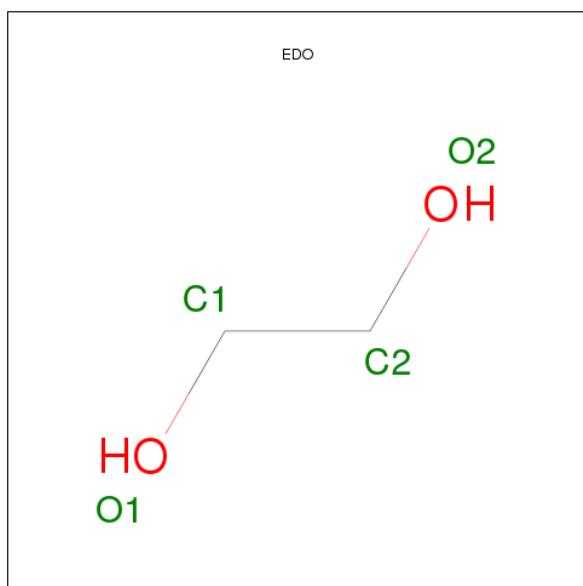
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			4	2	2		
4	A	1	Total	C	O	0	0
			4	2	2		

*Continued on next page...*

Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	B	1	Total	C	O	0	0
			4	2	2		
4	B	1	Total	C	O	0	0
			4	2	2		
4	E	1	Total	C	O	0	0
			4	2	2		
4	E	1	Total	C	O	0	0
			4	2	2		
4	E	1	Total	C	O	0	0
			4	2	2		
4	F	1	Total	C	O	0	0
			4	2	2		
4	G	1	Total	C	O	0	0
			4	2	2		

- Molecule 5 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C<sub>2</sub>H<sub>6</sub>O<sub>2</sub>).



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

Continued on next page...

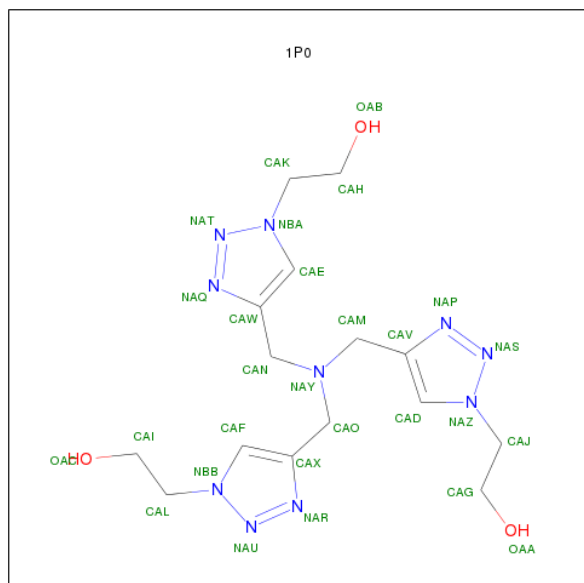
Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	B	1	Total C O 4 2 2	0	0
5	C	1	Total C O 4 2 2	0	0
5	E	1	Total C O 4 2 2	0	0
5	E	1	Total C O 4 2 2	0	0
5	E	1	Total C O 4 2 2	0	0
5	E	1	Total C O 4 2 2	0	0
5	F	1	Total C O 4 2 2	0	0
5	F	1	Total C O 4 2 2	0	0
5	F	1	Total C O 4 2 2	0	0

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

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

- Molecule 7 is 2,2',2''-[NITRILOTRIS(METHANEDIYL-1H-1,2,3-TRIAZOLE-4,1-DIYL)]TRIETHANOL (three-letter code: 1P0) (formula: C<sub>15</sub>H<sub>24</sub>N<sub>10</sub>O<sub>3</sub>).



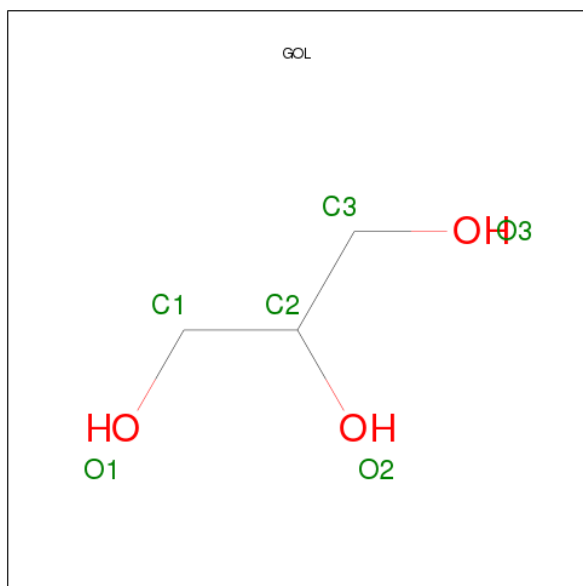


Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
7	B	1	Total	C	N	O	0	0
			28	15	10	3		

- Molecule 8 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	B	1	Total	Mg	0	0
			1	1		
8	C	2	Total	Mg	0	0
			2	2		
8	F	2	Total	Mg	0	0
			2	2		
8	E	1	Total	Mg	0	0
			1	1		

- Molecule 9 is GLYCEROL (three-letter code: GOL) (formula: C<sub>3</sub>H<sub>8</sub>O<sub>3</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
9	E	1	Total	C	O	0	0
			6	3	3		
9	G	1	Total	C	O	0	0
			6	3	3		
9	G	1	Total	C	O	0	0
			6	3	3		

- Molecule 10 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
10	G	1	Total 1	Ca 1	0	0

- Molecule 11 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
11	A	143	Total 143	O 143	0	0
11	B	146	Total 146	O 146	0	0
11	C	131	Total 131	O 131	0	0
11	E	129	Total 129	O 129	0	0
11	F	127	Total 127	O 127	0	0
11	G	146	Total 146	O 146	0	0

### 3 Residue-property plots

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

Chain A: 



- Molecule 1: Fiber protein

Chain B: 



- Molecule 1: Fiber protein

Chain C: 



- Molecule 1: Fiber protein

Chain E: 



- Molecule 1: Fiber protein

Chain F: 



- Molecule 1: Fiber protein

Chain G: 



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	74.94Å 67.18Å 121.43Å 90.00° 96.81° 90.00°	Depositor
Resolution (Å)	37.26 – 2.00 37.26 – 2.00	Depositor EDS
% Data completeness (in resolution range)	99.3 (37.26-2.00) 99.4 (37.26-2.00)	Depositor EDS
$R_{merge}$	0.12	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.36 (at 2.00Å)	Xtriage
Refinement program	REFMAC 5.8.0025	Depositor
R, $R_{free}$	0.160 , 0.200 0.170 , 0.204	Depositor DCC
$R_{free}$ test set	1615 reflections (2.04%)	DCC
Wilson B-factor (Å <sup>2</sup> )	23.6	Xtriage
Anisotropy	0.306	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.39 , 45.9	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.34$	Xtriage
Outliers	0 of 80719 reflections	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	9858	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	27.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 39.79 % of the origin peak, indicating pseudo translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo translational symmetry is equal to 3.0119e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

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

## 5 Model quality

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: GOL, MG, CL, CA, ZN, EDO, SIA, ACT, 1P0

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	0.58	0/1515	0.69	0/2064
1	B	0.61	0/1506	0.69	0/2052
1	C	0.56	0/1514	0.67	0/2062
1	E	0.57	0/1504	0.70	0/2046
1	F	0.54	1/1490 (0.1%)	0.66	1/2030 (0.0%)
1	G	0.59	0/1518	0.70	1/2069 (0.0%)
All	All	0.57	1/9047 (0.0%)	0.69	2/12323 (0.0%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	F	343	TRP	CD2-CE2	5.37	1.47	1.41

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	F	304	ARG	NE-CZ-NH2	-5.49	117.56	120.30
1	G	318	ASP	CB-CG-OD1	5.00	122.80	118.30

There are no chirality outliers.

There are no planarity outliers.

### 5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1466	0	1431	9	0
1	B	1460	0	1423	13	0
1	C	1459	0	1446	7	0
1	E	1458	0	1431	13	0
1	F	1448	0	1407	6	0
1	G	1469	0	1440	5	0
2	A	20	0	17	0	0
2	B	20	0	17	0	0
2	C	20	0	17	0	0
2	E	20	0	17	0	0
2	F	20	0	17	0	0
2	G	20	0	17	0	0
3	A	3	0	0	0	0
3	B	3	0	0	0	0
3	C	1	0	0	0	0
3	E	2	0	0	0	0
3	F	2	0	0	0	0
3	G	3	0	0	0	0
4	A	8	0	6	0	0
4	B	8	0	6	2	0
4	E	12	0	9	0	0
4	F	4	0	3	0	0
4	G	4	0	3	0	0
5	A	8	0	12	0	0
5	B	12	0	18	3	0
5	C	4	0	6	0	0
5	E	16	0	24	3	0
5	F	12	0	18	2	0
6	A	1	0	0	0	0
7	B	28	0	0	1	0
8	B	1	0	0	0	0
8	C	2	0	0	0	0
8	E	1	0	0	0	0
8	F	2	0	0	0	0
9	E	6	0	8	0	0
9	G	12	0	16	0	0
10	G	1	0	0	0	0
11	A	143	0	0	0	0
11	B	146	0	0	0	0
11	C	131	0	0	0	0
11	E	129	0	0	1	0
11	F	127	0	0	0	0
11	G	146	0	0	1	0

Continued on next page...

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	9858	0	8809	49	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 3.

All (49) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
7:B:402:1P0:CAN	7:B:402:1P0:NAR	2.36	0.89
1:E:285:GLY:H	5:E:407:EDO:H22	1.38	0.88
1:B:231:HIS:HE1	4:B:407:ACT:O	1.67	0.77
1:B:296:PRO:HA	1:B:297:SER:CB	2.22	0.69
1:A:213[B]:CYS:HG	1:C:213[B]:CYS:HB2	1.59	0.67
1:B:285:GLY:H	5:B:409:EDO:H21	1.62	0.64
1:E:213[B]:CYS:HG	1:G:213[B]:CYS:HB2	1.60	0.64
1:A:330:GLU:OE2	1:A:338[A]:THR:HG21	1.99	0.63
1:B:231:HIS:CE1	4:B:407:ACT:O	2.50	0.63
1:A:213[B]:CYS:HB2	1:B:213[B]:CYS:HG	1.66	0.61
1:B:322:VAL:HG23	5:B:410:EDO:H22	1.83	0.60
1:B:213[B]:CYS:HB2	1:C:213[B]:CYS:HG	1.68	0.59
1:E:330:GLU:OE2	1:E:338[A]:THR:HG21	2.03	0.58
1:E:301:LYS:NZ	5:E:408:EDO:O1	2.26	0.57
1:G:330:GLU:OE2	1:G:338[A]:THR:HG21	2.04	0.57
1:B:296:PRO:CA	1:B:297:SER:CB	2.83	0.56
1:E:234:ASN:OD1	1:E:236:LYS:HB2	2.07	0.54
1:G:329:GLN:NE2	11:G:585:HOH:O	2.28	0.53
1:B:285:GLY:H	5:B:409:EDO:C2	2.21	0.53
1:G:189[A]:THR:HG23	1:G:273:ASN:O	2.10	0.51
1:E:270:ARG:HH11	1:F:215:SER:HB2	1.76	0.51
1:B:330:GLU:OE2	1:B:338[B]:THR:HG21	2.10	0.51
1:A:300:LYS:HE3	1:A:302:TYR:CE1	2.46	0.50
1:C:330:GLU:OE2	1:C:338[A]:THR:HG21	2.12	0.49
1:E:241:ILE:HG22	1:E:241:ILE:O	2.12	0.48
1:F:266:TYR:CB	5:F:407:EDO:H11	2.43	0.48
1:A:215:SER:HB2	1:C:270:ARG:HH11	1.81	0.45
1:A:295:LYS:HD3	1:A:334:GLU:O	2.17	0.45
1:E:238:ASN:CB	1:E:241:ILE:HD12	2.47	0.45
1:F:330:GLU:OE2	1:F:338:THR:HG21	2.17	0.45
1:B:251:ASN:HB3	1:B:257:LEU:HD21	2.00	0.44
1:E:251:ASN:HB3	1:E:257:LEU:HD21	2.00	0.44
1:E:238:ASN:HB3	1:E:241:ILE:HD12	2.00	0.43
1:A:189:THR:HB	1:A:190:PRO:HD2	2.01	0.42
5:E:408:EDO:C1	11:E:613:HOH:O	2.68	0.41

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:F:228:GLY:H	5:F:405:EDO:C2	2.34	0.41
1:B:234:ASN:OD1	1:B:236:LYS:HB2	2.21	0.41
1:F:311:ILE:C	1:F:311:ILE:HD12	2.40	0.41
1:C:311:ILE:HD12	1:C:311:ILE:C	2.41	0.41
1:C:251:ASN:HB3	1:C:257:LEU:HD21	2.03	0.41
1:E:243:SER:HA	1:E:341:PHE:O	2.21	0.41
1:E:298:ASN:O	1:E:299:SER:CB	2.69	0.41
1:A:251:ASN:HB3	1:A:257:LEU:HD21	2.04	0.40

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	189/194 (97%)	182 (96%)	7 (4%)	0	100	100
1	B	188/194 (97%)	179 (95%)	8 (4%)	1 (0%)	38	29
1	C	188/194 (97%)	180 (96%)	8 (4%)	0	100	100
1	E	187/194 (96%)	179 (96%)	7 (4%)	1 (0%)	38	29
1	F	187/194 (96%)	179 (96%)	8 (4%)	0	100	100
1	G	189/194 (97%)	182 (96%)	7 (4%)	0	100	100
All	All	1128/1164 (97%)	1081 (96%)	45 (4%)	2 (0%)	56	51

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	297	SER
1	E	299	SER



### 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	163/171 (95%)	162 (99%)	1 (1%)	92	94
1	B	160/171 (94%)	158 (99%)	2 (1%)	80	82
1	C	166/171 (97%)	162 (98%)	4 (2%)	61	61
1	E	162/171 (95%)	159 (98%)	3 (2%)	69	71
1	F	161/171 (94%)	158 (98%)	3 (2%)	69	71
1	G	164/171 (96%)	161 (98%)	3 (2%)	71	73
All	All	976/1026 (95%)	960 (98%)	16 (2%)	76	77

All (16) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	187	TRP
1	B	187	TRP
1	B	301	LYS
1	C	187	TRP
1	C	189[A]	THR
1	C	189[B]	THR
1	C	264	LYS
1	E	187	TRP
1	E	278	THR
1	E	301	LYS
1	F	187	TRP
1	F	236	LYS
1	F	353	GLU
1	G	187	TRP
1	G	248[A]	LEU
1	G	248[B]	LEU

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

Mol	Chain	Res	Type
1	B	298	ASN
1	C	253	ASN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	G	319	GLN

### 5.3.3 RNA ⓘ

There are no RNA chains 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 ⓘ

Of 54 ligands modelled in this entry, 22 are monoatomic - leaving 32 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z  > 2$	Counts	RMSZ	$\# Z  > 2$
2	SIA	A	401	7	20,20,21	4.04	1 (5%)	23,28,31	2.51	3 (13%)
4	ACT	A	405	3	1,3,3	2.13	1 (100%)	0,3,3	0.00	-
4	ACT	A	406	3	1,3,3	1.69	0	0,3,3	0.00	-
5	EDO	A	407	-	3,3,3	0.61	0	2,2,2	0.26	0
5	EDO	A	408	-	3,3,3	0.60	0	2,2,2	0.27	0
2	SIA	B	401	7	20,20,21	3.99	1 (5%)	23,28,31	2.88	2 (8%)
7	1P0	B	402	2	30,30,30	3.09	10 (33%)	39,39,39	2.97	18 (46%)
4	ACT	B	406	3	1,3,3	1.33	0	0,3,3	0.00	-
4	ACT	B	407	3	1,3,3	1.61	0	0,3,3	0.00	-
5	EDO	B	408	-	3,3,3	0.61	0	2,2,2	0.21	0
5	EDO	B	409	-	3,3,3	0.48	0	2,2,2	0.31	0
5	EDO	B	410	-	3,3,3	0.47	0	2,2,2	0.48	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
2	SIA	C	401	7	20,20,21	4.34	1 (5%)	23,28,31	2.30	2 (8%)
5	EDO	C	403	-	3,3,3	0.50	0	2,2,2	0.30	0
2	SIA	E	403	-	20,20,21	4.27	1 (5%)	23,28,31	2.67	4 (17%)
4	ACT	E	404	-	1,3,3	1.60	0	0,3,3	0.00	-
4	ACT	E	405	-	1,3,3	1.14	0	0,3,3	0.00	-
4	ACT	E	406	-	1,3,3	1.77	0	0,3,3	0.00	-
5	EDO	E	407	-	3,3,3	0.50	0	2,2,2	0.62	0
5	EDO	E	408	-	3,3,3	0.28	0	2,2,2	0.50	0
5	EDO	E	409	-	3,3,3	0.56	0	2,2,2	0.25	0
5	EDO	E	410	-	3,3,3	0.49	0	2,2,2	0.44	0
9	GOL	E	411	-	5,5,5	0.42	0	5,5,5	0.45	0
2	SIA	F	403	-	20,20,21	4.03	1 (5%)	23,28,31	2.85	4 (17%)
4	ACT	F	404	3	1,3,3	2.13	1 (100%)	0,3,3	0.00	-
5	EDO	F	405	-	3,3,3	0.62	0	2,2,2	0.48	0
5	EDO	F	406	-	3,3,3	0.45	0	2,2,2	0.43	0
5	EDO	F	407	-	3,3,3	0.64	0	2,2,2	0.31	0
2	SIA	G	404	-	20,20,21	3.97	1 (5%)	23,28,31	2.61	3 (13%)
4	ACT	G	405	10	1,3,3	1.14	0	0,3,3	0.00	-
9	GOL	G	406	-	5,5,5	0.28	0	5,5,5	0.38	0
9	GOL	G	407	-	5,5,5	0.40	0	5,5,5	0.85	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	SIA	A	401	7	-	0/15/34/38	0/1/1/1
4	ACT	A	405	3	-	0/0/0/0	0/0/0/0
4	ACT	A	406	3	-	0/0/0/0	0/0/0/0
5	EDO	A	407	-	-	0/1/1/1	0/0/0/0
5	EDO	A	408	-	-	0/1/1/1	0/0/0/0
2	SIA	B	401	7	-	0/15/34/38	0/1/1/1
7	1P0	B	402	2	-	0/21/21/21	0/3/3/3
4	ACT	B	406	3	-	0/0/0/0	0/0/0/0
4	ACT	B	407	3	-	0/0/0/0	0/0/0/0
5	EDO	B	408	-	-	0/1/1/1	0/0/0/0
5	EDO	B	409	-	-	0/1/1/1	0/0/0/0
5	EDO	B	410	-	-	0/1/1/1	0/0/0/0
2	SIA	C	401	7	-	0/15/34/38	0/1/1/1
5	EDO	C	403	-	-	0/1/1/1	0/0/0/0

Continued on next page...

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	SIA	E	403	-	-	0/15/34/38	0/1/1/1
4	ACT	E	404	-	-	0/0/0/0	0/0/0/0
4	ACT	E	405	-	-	0/0/0/0	0/0/0/0
4	ACT	E	406	-	-	0/0/0/0	0/0/0/0
5	EDO	E	407	-	-	0/1/1/1	0/0/0/0
5	EDO	E	408	-	-	0/1/1/1	0/0/0/0
5	EDO	E	409	-	-	0/1/1/1	0/0/0/0
5	EDO	E	410	-	-	0/1/1/1	0/0/0/0
9	GOL	E	411	-	-	0/4/4/4	0/0/0/0
2	SIA	F	403	-	-	0/15/34/38	0/1/1/1
4	ACT	F	404	3	-	0/0/0/0	0/0/0/0
5	EDO	F	405	-	-	0/1/1/1	0/0/0/0
5	EDO	F	406	-	-	0/1/1/1	0/0/0/0
5	EDO	F	407	-	-	0/1/1/1	0/0/0/0
2	SIA	G	404	-	-	0/15/34/38	0/1/1/1
4	ACT	G	405	10	-	0/0/0/0	0/0/0/0
9	GOL	G	406	-	-	0/4/4/4	0/0/0/0
9	GOL	G	407	-	-	0/4/4/4	0/0/0/0

All (18) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	401	SIA	O6-C2	19.09	1.45	1.28
2	E	403	SIA	O6-C2	18.78	1.45	1.28
2	A	401	SIA	O6-C2	17.63	1.44	1.28
2	F	403	SIA	O6-C2	17.61	1.44	1.28
2	B	401	SIA	O6-C2	17.55	1.44	1.28
2	G	404	SIA	O6-C2	17.38	1.44	1.28
7	B	402	1P0	NAR-NAU	-9.82	1.20	1.34
7	B	402	1P0	NAQ-NAT	-8.27	1.22	1.34
7	B	402	1P0	NAU-NBB	-5.12	1.24	1.34
7	B	402	1P0	NAS-NAZ	-4.22	1.26	1.34
7	B	402	1P0	CAO-CAX	3.39	1.56	1.51
7	B	402	1P0	CAJ-NAZ	3.21	1.54	1.47
7	B	402	1P0	CAE-CAW	-3.06	1.31	1.36
7	B	402	1P0	CAE-NBA	-3.05	1.32	1.35
7	B	402	1P0	NAT-NBA	-2.52	1.29	1.34
7	B	402	1P0	CAN-NAY	2.16	1.51	1.47
4	A	405	ACT	CH3-C	2.13	1.51	1.48
4	F	404	ACT	CH3-C	2.13	1.51	1.48

All (36) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	401	SIA	O6-C2-C3	-13.13	107.75	124.91
2	F	403	SIA	O6-C2-C3	-11.68	109.65	124.91
2	G	404	SIA	O6-C2-C3	-11.32	110.12	124.91
2	A	401	SIA	O6-C2-C3	-11.16	110.33	124.91
2	E	403	SIA	O6-C2-C3	-10.75	110.87	124.91
2	C	401	SIA	O6-C2-C3	-10.05	111.77	124.91
7	B	402	1P0	CAW-CAN-NAY	-9.78	94.62	112.87
7	B	402	1P0	CAX-CAO-NAY	-7.13	99.57	112.87
7	B	402	1P0	NAR-NAU-NBB	5.49	111.45	107.31
7	B	402	1P0	NAQ-NAT-NBA	5.24	111.27	107.31
2	E	403	SIA	O6-C6-C7	4.79	114.14	105.67
2	F	403	SIA	O6-C6-C7	4.38	113.43	105.67
7	B	402	1P0	CAF-CAX-NAR	-3.80	105.34	111.35
2	G	404	SIA	O6-C6-C7	3.70	112.21	105.67
7	B	402	1P0	CAK-NBA-NAT	3.57	130.14	121.63
7	B	402	1P0	NAP-NAS-NAZ	3.36	109.85	107.31
7	B	402	1P0	CAK-NBA-CAE	-3.31	121.72	129.93
7	B	402	1P0	CAJ-NAZ-NAS	2.99	128.76	121.63
7	B	402	1P0	CAL-NBB-NAU	-2.96	114.58	121.63
7	B	402	1P0	CAJ-NAZ-CAD	-2.81	122.96	129.93
7	B	402	1P0	CAL-NBB-CAF	2.64	136.49	129.93
7	B	402	1P0	CAX-CAF-NBB	-2.64	103.99	107.51
7	B	402	1P0	CAH-CAK-NBA	-2.53	104.98	110.91
7	B	402	1P0	CAW-CAE-NBA	-2.52	104.14	107.51
7	B	402	1P0	CAD-CAV-NAP	-2.48	107.43	111.35
2	C	401	SIA	O6-C6-C7	2.42	109.95	105.67
2	E	403	SIA	C7-C6-C5	-2.34	110.86	114.25
2	F	403	SIA	C6-C5-N5	-2.30	106.99	110.99
2	E	403	SIA	C3-C2-C1	-2.23	108.61	121.08
2	B	401	SIA	O6-C6-C7	2.17	109.51	105.67
2	G	404	SIA	C6-C5-N5	-2.15	107.24	110.99
2	A	401	SIA	C6-C5-N5	-2.12	107.29	110.99
7	B	402	1P0	CAN-CAW-CAE	-2.08	125.88	129.28
2	F	403	SIA	O7-C7-C8	2.06	113.92	108.73
7	B	402	1P0	CAG-CAJ-NAZ	2.06	115.73	110.91
2	A	401	SIA	O6-C6-C7	2.05	109.29	105.67

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

## 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	186/194 (95%)	-0.54	0	100	100	15, 23, 39, 68	0
1	B	186/194 (95%)	-0.35	0	100	100	17, 24, 39, 72	0
1	C	182/194 (93%)	-0.16	5 (2%)	52	52	17, 28, 50, 75	0
1	E	185/194 (95%)	-0.36	1 (0%)	88	89	16, 24, 43, 80	0
1	F	186/194 (95%)	-0.23	4 (2%)	59	59	17, 28, 46, 75	0
1	G	185/194 (95%)	-0.53	0	100	100	14, 22, 36, 46	0
All	All	1110/1164 (95%)	-0.36	10 (0%)	81	82	14, 25, 43, 80	0

All (10) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	F	181	TYR	4.3
1	F	180	LYS	4.2
1	C	277	SER	3.3
1	C	266	TYR	3.2
1	F	277	SER	3.2
1	C	278	THR	2.9
1	C	276	VAL	2.8
1	C	271	SER	2.4
1	E	332	GLY	2.2
1	F	278	THR	2.2

### 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	RSR	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
4	ACT	E	404	4/4	0.18	34.75	27,33,33,36	0
9	GOL	G	406	6/6	0.23	13.49	32,44,53,57	0
4	ACT	A	406	4/4	0.11	9.95	18,21,22,27	0
2	SIA	G	404	20/21	0.25	9.22	37,57,75,77	0
2	SIA	F	403	20/21	0.23	7.55	33,56,69,76	0
4	ACT	G	405	4/4	0.26	5.16	48,52,53,54	0
5	EDO	E	408	4/4	0.14	3.74	31,41,41,41	0
4	ACT	E	405	4/4	0.19	3.39	27,30,38,46	0
5	EDO	E	410	4/4	0.17	3.18	54,56,58,59	0
5	EDO	F	407	4/4	0.20	3.14	28,29,40,50	0
5	EDO	F	405	4/4	0.14	2.91	31,36,47,50	0
5	EDO	E	409	4/4	0.19	2.36	48,49,54,57	0
4	ACT	E	406	4/4	0.16	2.35	24,36,44,53	0
2	SIA	E	403	20/21	0.23	2.29	39,58,72,74	0
8	MG	C	404	1/1	0.08	2.20	40,40,40,40	0
5	EDO	F	406	4/4	0.13	1.64	33,42,44,45	0
5	EDO	A	408	4/4	0.15	1.48	31,32,42,51	0
7	1P0	B	402	28/28	0.18	1.43	19,44,56,68	0
9	GOL	E	411	6/6	0.19	1.27	41,54,57,60	0
3	ZN	B	405	1/1	0.10	1.21	81,81,81,81	0
2	SIA	A	401	20/21	0.12	1.13	23,30,42,44	0
5	EDO	A	407	4/4	0.13	1.05	23,28,32,37	0
5	EDO	E	407	4/4	0.12	0.95	34,36,39,47	0
2	SIA	B	401	20/21	0.13	0.85	24,34,45,55	0
6	CL	A	409	1/1	0.09	0.81	48,48,48,48	0
9	GOL	G	407	6/6	0.13	0.62	35,43,53,57	0
5	EDO	B	409	4/4	0.12	0.27	26,34,36,40	0
3	ZN	F	402	1/1	0.08	0.26	73,73,73,73	0
5	EDO	B	408	4/4	0.12	0.26	46,55,59,62	0
3	ZN	B	404	1/1	0.13	0.24	71,71,71,71	0

*Continued on next page...*



*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
4	ACT	F	404	4/4	0.08	0.24	25,25,32,34	0
4	ACT	A	405	4/4	0.09	0.21	15,22,31,40	0
5	EDO	B	410	4/4	0.11	0.13	31,32,36,39	0
2	SIA	C	401	20/21	0.10	0.01	26,33,43,44	0
4	ACT	B	406	4/4	0.08	-0.21	23,24,29,35	0
5	EDO	C	403	4/4	0.09	-0.30	34,36,50,53	0
4	ACT	B	407	4/4	0.09	-0.77	24,31,34,40	0
8	MG	E	412	1/1	0.06	-0.83	39,39,39,39	0
3	ZN	E	401	1/1	0.06	-0.86	21,21,21,21	1
8	MG	C	405	1/1	0.06	-1.25	37,37,37,37	0
8	MG	F	409	1/1	0.06	-1.26	36,36,36,36	0
10	CA	G	408	1/1	0.06	-1.43	45,45,45,45	0
3	ZN	G	402	1/1	0.03	-1.71	35,35,35,35	0
3	ZN	G	403	1/1	0.07	-1.83	80,80,80,80	0
8	MG	F	408	1/1	0.07	-2.03	41,41,41,41	0
8	MG	B	411	1/1	0.03	-2.14	28,28,28,28	0
3	ZN	A	403	1/1	0.04	-2.45	23,23,23,23	0
3	ZN	F	401	1/1	0.04	-2.52	33,33,33,33	0
3	ZN	B	403	1/1	0.02	-2.59	40,40,40,40	0
3	ZN	A	402	1/1	0.02	-2.87	24,24,24,24	0
3	ZN	C	402	1/1	0.04	-2.98	37,37,37,37	0
3	ZN	G	401	1/1	0.03	-3.02	28,28,28,28	0
3	ZN	E	402	1/1	0.06	-17.26	87,87,87,87	0
3	ZN	A	404	1/1	0.06	-	68,68,68,68	0

## 6.5 Other polymers ⓘ

There are no such residues in this entry.