



# Full wwPDB X-ray Structure Validation Report ⓘ

Jan 3, 2024 – 11:32 pm GMT

PDB ID : 4XIW  
Title : Carbonic anhydrase Cah3 from Chlamydomonas reinhardtii in complex with acetazolamide  
Authors : Hainzl, T.; Grundstrom, C.; Benloch, R.; Shevela, D.; Shutova, T.; Messinger, J.; Samuelsson, G.; Sauer-Eriksson, A.E.  
Deposited on : 2015-01-07  
Resolution : 2.60 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

<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 : 1.8.4, CSD as541be (2020)  
Xtrriage (Phenix) : 1.13  
EDS : 2.36  
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) : 2.36

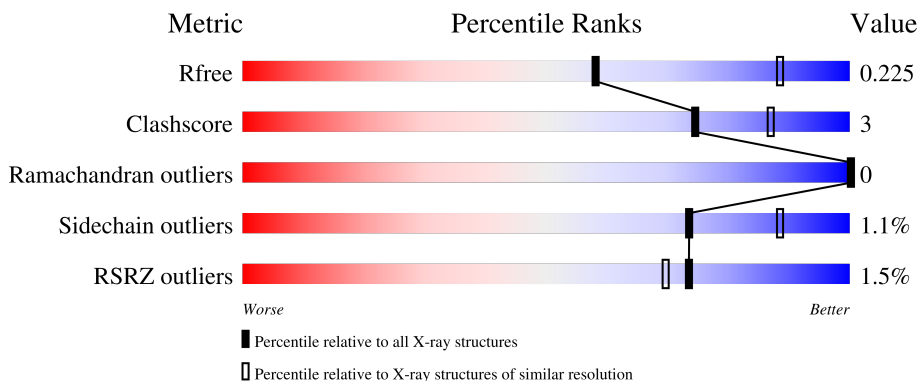
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.60 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	3163 (2.60-2.60)
Clashscore	141614	3518 (2.60-2.60)
Ramachandran outliers	138981	3455 (2.60-2.60)
Sidechain outliers	138945	3455 (2.60-2.60)
RSRZ outliers	127900	3104 (2.60-2.60)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions  $\leq 5\%$ . The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	239	
1	B	239	
1	C	239	
1	D	239	
1	E	239	

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	F	239	 3% 92% 8%
1	G	239	 % 93% 5%
1	H	239	 % 93% 7%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
4	2HP	E	404	-	X	-	-

## 2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 29766 atoms, of which 14607 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 Carbonic anhydrase, alpha type.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
1	A	237	3645	1172	1818	310	333	12	0	0	0
1	B	237	3644	1172	1817	310	333	12	0	0	0
1	C	239	3682	1186	1831	316	336	13	0	1	0
1	D	237	3639	1172	1812	310	333	12	0	0	0
1	E	238	3650	1175	1818	311	334	12	0	0	0
1	F	238	3655	1175	1823	311	334	12	0	0	0
1	G	238	3655	1175	1823	311	334	12	0	0	0
1	H	239	3657	1180	1817	312	335	13	0	0	0

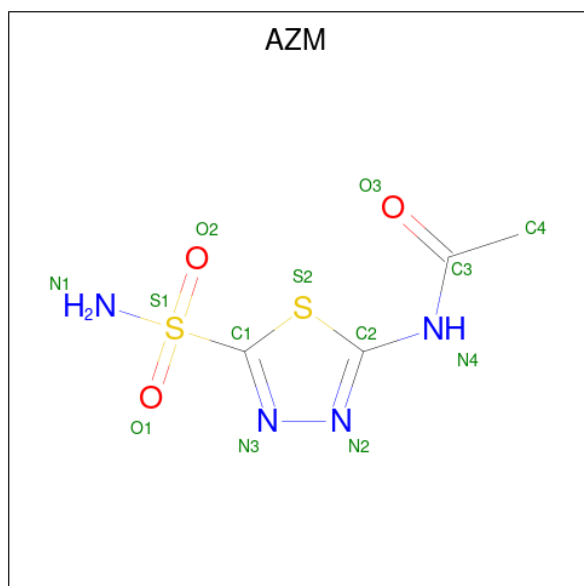
There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	72	MET	-	initiating methionine	UNP Q39588
B	72	MET	-	initiating methionine	UNP Q39588
C	72	MET	-	initiating methionine	UNP Q39588
D	72	MET	-	initiating methionine	UNP Q39588
E	72	MET	-	initiating methionine	UNP Q39588
F	72	MET	-	initiating methionine	UNP Q39588
G	72	MET	-	initiating methionine	UNP Q39588
H	72	MET	-	initiating methionine	UNP Q39588

- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn).

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

- Molecule 3 is 5-ACETAMIDO-1,3,4-THIADIAZOLE-2-SULFONAMIDE (three-letter code: AZM) (formula: C<sub>4</sub>H<sub>6</sub>N<sub>4</sub>O<sub>3</sub>S<sub>2</sub>).



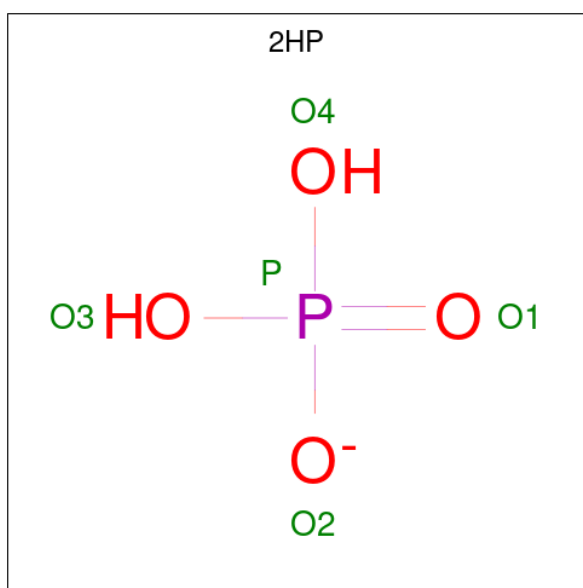
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C H N O S 19 4 6 4 3 2	0	0
3	B	1	Total C H N O S 19 4 6 4 3 2	0	0
3	C	1	Total C H N O S 19 4 6 4 3 2	0	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	
3	D	1	Total	C	H	N	O	S	0	0
			19	4	6	4	3	2		
3	E	1	Total	C	H	N	O	S	0	0
			19	4	6	4	3	2		
3	F	1	Total	C	H	N	O	S	0	0
			19	4	6	4	3	2		
3	G	1	Total	C	H	N	O	S	0	0
			19	4	6	4	3	2		
3	H	1	Total	C	H	N	O	S	0	0
			19	4	6	4	3	2		

- Molecule 4 is DIHYDROGENPHOSPHATE ION (three-letter code: 2HP) (formula:  $\text{H}_2\text{O}_4\text{P}$ ).



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

Continued on next page...

*Continued from previous page...*

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	C	1	Total	O	P	0	0
			5	4	1		
4	C	1	Total	O	P	0	0
			5	4	1		
4	D	1	Total	O	P	0	0
			5	4	1		
4	E	1	Total	O	P	0	0
			5	4	1		
4	E	1	Total	O	P	0	0
			5	4	1		
4	F	1	Total	O	P	0	0
			5	4	1		
4	G	1	Total	O	P	0	0
			5	4	1		
4	H	1	Total	O	P	0	0
			5	4	1		
4	H	1	Total	O	P	0	0
			5	4	1		
4	H	1	Total	O	P	0	0
			5	4	1		

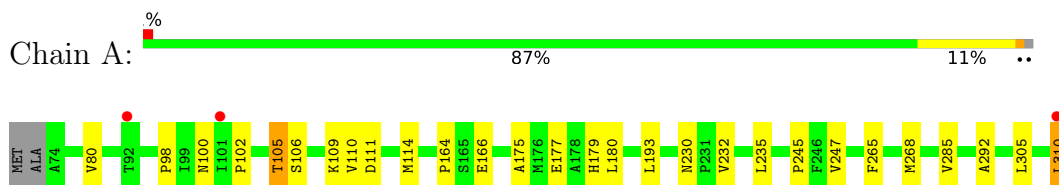
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	49	Total	O	0	0
			49	49		
5	B	37	Total	O	0	0
			37	37		
5	C	46	Total	O	0	0
			46	46		
5	D	29	Total	O	0	0
			29	29		
5	E	30	Total	O	0	0
			30	30		
5	F	28	Total	O	0	0
			28	28		
5	G	32	Total	O	0	0
			32	32		
5	H	48	Total	O	0	0
			48	48		

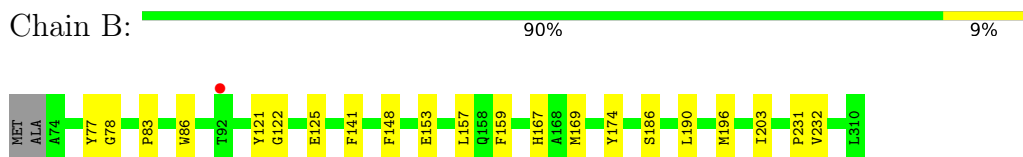
### 3 Residue-property plots [i](#)

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.

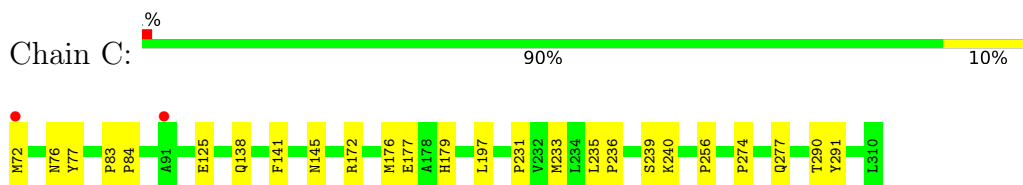
- Molecule 1: Carbonic anhydrase, alpha type



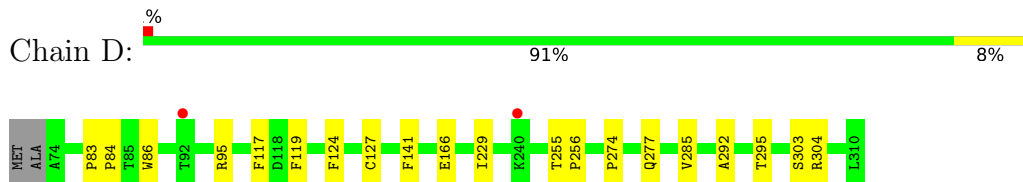
- Molecule 1: Carbonic anhydrase, alpha type



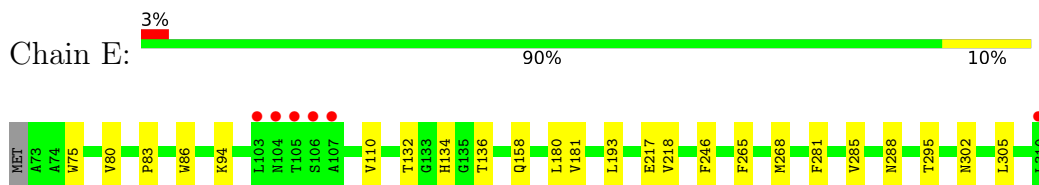
- Molecule 1: Carbonic anhydrase, alpha type



- Molecule 1: Carbonic anhydrase, alpha type

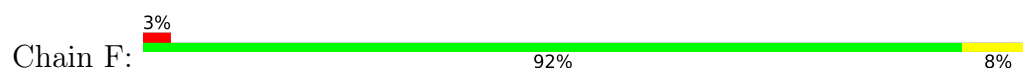


- Molecule 1: Carbonic anhydrase, alpha type

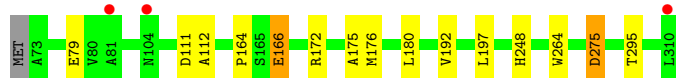
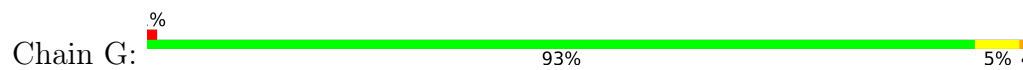


- Molecule 1: Carbonic anhydrase, alpha type

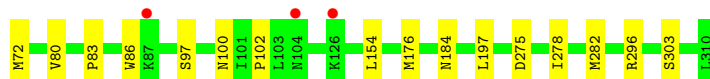
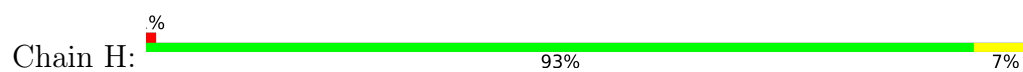




- Molecule 1: Carbonic anhydrase, alpha type



- Molecule 1: Carbonic anhydrase, alpha type



## 4 Data and refinement statistics i

Property	Value	Source
Space group	P 32 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	139.36Å 139.36Å 203.21Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	46.82 – 2.60 48.57 – 2.60	Depositor EDS
% Data completeness (in resolution range)	99.6 (46.82-2.60) 99.9 (48.57-2.60)	Depositor EDS
$R_{merge}$	0.16	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.56 (at 2.61Å)	Xtrriage
Refinement program	PHENIX (phenix.refine: 1.8.2_1309)	Depositor
R, $R_{free}$	0.161 , 0.226 0.165 , 0.225	Depositor DCC
$R_{free}$ test set	3562 reflections (5.04%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	52.2	Xtrriage
Anisotropy	0.753	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.39 , 41.2	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtrriage
Estimated twinning fraction	0.021 for -h,-k,l	Xtrriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	29766	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	35.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 23.06 % 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 5.0264e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

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

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

## 5 Model quality [i](#)

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

Bond lengths and bond angles in the following residue types are not validated in this section: 2HP, AZM, ZN

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	0.74	0/1879	0.71	0/2554
1	B	0.65	0/1879	0.71	0/2554
1	C	0.68	0/1903	0.74	0/2585
1	D	0.63	0/1879	0.69	0/2554
1	E	0.60	0/1884	0.66	0/2561
1	F	0.62	0/1884	0.68	0/2561
1	G	0.64	0/1884	0.69	0/2561
1	H	0.62	0/1892	0.70	0/2571
All	All	0.65	0/15084	0.70	0/20501

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

### 5.2 Too-close contacts [i](#)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1827	1818	1813	15	0
1	B	1827	1817	1813	16	0
1	C	1851	1831	1839	17	0
1	D	1827	1812	1813	12	0
1	E	1832	1818	1818	18	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	F	1832	1823	1818	11	0
1	G	1832	1823	1818	9	0
1	H	1840	1817	1827	11	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
2	C	1	0	0	0	0
2	D	1	0	0	0	0
2	E	1	0	0	0	0
2	F	1	0	0	0	0
2	G	1	0	0	0	0
2	H	1	0	0	0	0
3	A	13	6	5	0	0
3	B	13	6	5	0	0
3	C	13	6	5	0	0
3	D	13	6	5	0	0
3	E	13	6	5	0	0
3	F	13	6	5	0	0
3	G	13	6	5	0	0
3	H	13	6	5	0	0
4	A	15	0	0	0	0
4	B	10	0	0	0	0
4	C	15	0	0	0	0
4	D	5	0	0	0	0
4	E	10	0	0	0	0
4	F	5	0	0	0	0
4	G	5	0	0	0	0
4	H	15	0	0	0	0
5	A	49	0	0	0	0
5	B	37	0	0	0	0
5	C	46	0	0	0	0
5	D	29	0	0	0	0
5	E	30	0	0	0	0
5	F	28	0	0	0	0
5	G	32	0	0	0	0
5	H	48	0	0	0	0
All	All	15159	14607	14599	95	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 3.

All (95) 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:C:72:MET:HA	1:C:72:MET:HE2	1.69	0.72
1:A:230:ASN:OD1	1:A:232:VAL:HG12	1.93	0.69
1:H:97:SER:O	1:H:296:ARG:NH1	2.28	0.66
1:C:77:TYR:O	1:E:80:VAL:HG13	2.03	0.57
1:A:100:ASN:O	1:A:102:PRO:HD3	2.04	0.57
1:F:97:SER:O	1:F:296:ARG:NH1	2.37	0.56
1:C:176:MET:HB2	1:C:197:LEU:HB2	1.88	0.56
1:B:141:PHE:CE2	1:B:231:PRO:HG3	2.41	0.55
1:B:148:PHE:HE1	1:B:153:GLU:HG3	1.72	0.55
1:B:157:LEU:O	1:C:72:MET:HG3	2.06	0.55
1:A:285:VAL:HG12	1:A:292:ALA:HB3	1.88	0.55
1:A:247:VAL:HG13	1:A:268:MET:SD	2.46	0.55
1:H:100:ASN:O	1:H:102:PRO:HD3	2.07	0.54
1:E:132:THR:HG23	1:E:134:HIS:H	1.72	0.54
1:H:278:ILE:HG22	1:H:282:MET:CE	2.39	0.52
1:D:117:PHE:HB3	1:D:119:PHE:CE2	2.45	0.52
1:C:235:LEU:HB3	1:C:236:PRO:HD2	1.92	0.52
1:C:239:SER:OG	1:C:240:LYS:N	2.42	0.52
1:H:83:PRO:HA	1:H:86:TRP:CD2	2.45	0.51
1:D:127:CYS:SG	1:D:141:PHE:CE1	3.04	0.51
1:E:110:VAL:HG21	1:E:305:LEU:HD11	1.92	0.51
1:C:290:THR:OG1	1:C:291:TYR:N	2.43	0.50
1:H:176:MET:HB2	1:H:197:LEU:HB2	1.93	0.49
1:F:80:VAL:HG23	1:F:83:PRO:CD	2.42	0.49
1:G:111:ASP:OD1	1:G:112:ALA:N	2.46	0.49
1:D:95:ARG:CZ	1:D:304:ARG:HB3	2.42	0.49
1:A:230:ASN:CG	1:A:232:VAL:HG12	2.33	0.48
1:B:169:MET:HG3	1:B:174:TYR:HE1	1.77	0.48
1:F:80:VAL:HG23	1:F:83:PRO:HD2	1.94	0.48
1:E:193:LEU:HD22	1:E:265:PHE:HB2	1.96	0.48
1:E:217:GLU:OE2	1:H:72:MET:HE1	2.13	0.48
1:C:141:PHE:CE1	1:C:231:PRO:HG3	2.48	0.47
1:A:245:PRO:HB2	1:A:310:LEU:HD12	1.95	0.47
1:B:148:PHE:CE1	1:B:153:GLU:HG3	2.49	0.47
1:G:275:ASP:N	1:G:275:ASP:OD1	2.46	0.47
1:A:80:VAL:HA	1:G:79:GLU:O	2.14	0.47
1:A:110:VAL:HG21	1:A:305:LEU:HD21	1.97	0.47
1:G:176:MET:HB2	1:G:197:LEU:HB2	1.96	0.47
1:G:192:VAL:HG11	1:G:264:TRP:CH2	2.50	0.47
1:H:275:ASP:N	1:H:275:ASP:OD1	2.48	0.47
1:D:166:GLU:HG2	1:D:295:THR:HG23	1.97	0.46
1:C:77:TYR:HB3	1:E:80:VAL:HG11	1.97	0.46

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:98:PRO:HG3	1:A:166:GLU:HB3	1.98	0.46
1:B:78:GLY:O	1:H:80:VAL:HG13	2.15	0.46
1:B:141:PHE:CZ	1:B:231:PRO:HG3	2.51	0.46
1:B:77:TYR:HB3	1:H:80:VAL:HG11	1.98	0.46
1:F:176:MET:HB2	1:F:197:LEU:HB2	1.97	0.46
1:G:180:LEU:N	1:G:180:LEU:HD12	2.31	0.46
1:H:154:LEU:HD23	1:H:184:ASN:HA	1.98	0.46
1:C:172:ARG:HH22	1:H:72:MET:HG2	1.81	0.46
1:B:167:HIS:HB2	1:B:196:MET:HE1	1.98	0.45
1:B:141:PHE:O	1:C:72:MET:HB2	2.15	0.45
1:F:82:GLY:O	1:F:85:THR:HG22	2.16	0.45
1:F:177:GLU:OE2	1:F:179:HIS:NE2	2.41	0.45
1:B:121:TYR:HB2	1:B:232:VAL:HG23	1.99	0.44
1:E:158:GLN:NE2	1:E:181:VAL:HG21	2.33	0.44
1:E:281:PHE:O	1:E:285:VAL:HG23	2.18	0.44
1:E:217:GLU:OE2	1:E:288:ASN:ND2	2.50	0.44
1:B:159:PHE:C	1:B:159:PHE:CD1	2.91	0.44
1:D:127:CYS:SG	1:D:141:PHE:CD1	3.11	0.44
1:E:83:PRO:HA	1:E:86:TRP:CD2	2.53	0.43
1:A:193:LEU:HD22	1:A:265:PHE:HB2	2.00	0.43
1:D:83:PRO:N	1:D:84:PRO:HD2	2.34	0.43
1:D:285:VAL:HG12	1:D:292:ALA:HB3	1.99	0.43
1:D:256:PRO:HG2	1:F:75:TRP:CD1	2.53	0.43
1:G:164:PRO:HA	1:G:175:ALA:O	2.18	0.43
1:D:255:THR:HB	1:F:75:TRP:CH2	2.54	0.43
1:F:180:LEU:HD12	1:F:180:LEU:N	2.33	0.43
1:E:180:LEU:HD12	1:E:180:LEU:N	2.34	0.42
1:E:132:THR:HG22	1:E:136:THR:O	2.18	0.42
1:A:180:LEU:HD12	1:A:180:LEU:N	2.34	0.42
1:B:83:PRO:HA	1:B:86:TRP:CD2	2.54	0.42
1:B:122:GLY:HA3	1:B:125:GLU:OE2	2.19	0.42
1:D:274:PRO:HD2	1:D:277:GLN:OE1	2.20	0.42
1:E:132:THR:HG23	1:E:134:HIS:N	2.34	0.42
1:B:190:LEU:HD11	1:E:218:VAL:HG11	2.01	0.42
1:F:211:ALA:O	1:F:215:ALA:HB2	2.20	0.42
1:A:177:GLU:OE2	1:A:179:HIS:NE2	2.41	0.42
1:C:83:PRO:N	1:C:84:PRO:CD	2.83	0.42
1:A:105:THR:OG1	1:A:106:SER:N	2.53	0.41
1:D:83:PRO:HA	1:D:86:TRP:CD2	2.55	0.41
1:C:256:PRO:HG2	1:E:75:TRP:CG	2.55	0.41
1:E:94:LYS:HG2	1:E:302:ASN:OD1	2.20	0.41

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:177:GLU:OE2	1:C:179:HIS:NE2	2.44	0.41
1:C:274:PRO:HD2	1:C:277:GLN:OE1	2.20	0.41
1:A:164:PRO:HA	1:A:175:ALA:O	2.21	0.41
1:A:232:VAL:HG23	1:A:235:LEU:HD12	2.03	0.41
1:E:83:PRO:HA	1:E:86:TRP:CE2	2.56	0.41
1:D:124:PHE:HB2	1:D:229:ILE:O	2.21	0.41
1:E:246:PHE:HA	1:E:268:MET:HG2	2.03	0.40
1:G:166:GLU:HG3	1:G:295:THR:HG23	2.03	0.40
1:C:125:GLU:OE2	1:C:145:ASN:ND2	2.47	0.40
1:B:203:ILE:HD11	1:F:203:ILE:HD12	2.02	0.40
1:C:233:MET:CE	1:C:233:MET:HA	2.51	0.40
1:G:248:HIS:HA	1:G:264:TRP:O	2.22	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	235/239 (98%)	221 (94%)	14 (6%)	0	100	100
1	B	235/239 (98%)	219 (93%)	16 (7%)	0	100	100
1	C	238/239 (100%)	220 (92%)	18 (8%)	0	100	100
1	D	235/239 (98%)	222 (94%)	13 (6%)	0	100	100
1	E	236/239 (99%)	219 (93%)	17 (7%)	0	100	100
1	F	236/239 (99%)	221 (94%)	15 (6%)	0	100	100
1	G	236/239 (99%)	216 (92%)	20 (8%)	0	100	100
1	H	237/239 (99%)	226 (95%)	11 (5%)	0	100	100
All	All	1888/1912 (99%)	1764 (93%)	124 (7%)	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	198/199 (100%)	193 (98%)	5 (2%)	47	73
1	B	198/199 (100%)	197 (100%)	1 (0%)	88	96
1	C	200/199 (100%)	198 (99%)	2 (1%)	76	90
1	D	198/199 (100%)	197 (100%)	1 (0%)	88	96
1	E	198/199 (100%)	197 (100%)	1 (0%)	88	96
1	F	198/199 (100%)	194 (98%)	4 (2%)	55	78
1	G	198/199 (100%)	195 (98%)	3 (2%)	65	83
1	H	199/199 (100%)	198 (100%)	1 (0%)	88	96
All	All	1587/1592 (100%)	1569 (99%)	18 (1%)	73	88

All (18) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	105	THR
1	A	109	LYS
1	A	111	ASP
1	A	114	MET
1	A	310	LEU
1	B	186	SER
1	C	76	ASN
1	C	138	GLN
1	D	303	SER
1	E	295	THR
1	F	111	ASP
1	F	113	GLU
1	F	155	GLU
1	F	275	ASP
1	G	166	GLU
1	G	172	ARG
1	G	275	ASP
1	H	303	SER



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

Mol	Chain	Res	Type
1	H	248	HIS
1	H	269	GLN

### 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, 8 are monoatomic - leaving 24 for Mogul analysis.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
4	2HP	H	404	-	4,4,4	1.42	1 (25%)	6,6,6	0.59	0
4	2HP	A	403	-	4,4,4	1.41	1 (25%)	6,6,6	0.82	0
3	AZM	E	402	2	8,13,13	2.47	2 (25%)	9,19,19	2.36	4 (44%)
4	2HP	E	404	-	4,4,4	3.11	4 (100%)	6,6,6	0.78	0
4	2HP	H	403	-	4,4,4	1.12	0	6,6,6	0.68	0
3	AZM	B	402	2	8,13,13	2.44	2 (25%)	9,19,19	3.51	5 (55%)
4	2HP	A	405	-	4,4,4	1.58	1 (25%)	6,6,6	0.39	0
4	2HP	B	403	-	4,4,4	2.40	2 (50%)	6,6,6	0.45	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
4	2HP	C	405	-	4,4,4	1.86	1 (25%)	6,6,6	0.74	0
4	2HP	G	403	-	4,4,4	2.21	2 (50%)	6,6,6	0.49	0
4	2HP	A	404	-	4,4,4	1.50	1 (25%)	6,6,6	0.89	0
3	AZM	H	402	2	8,13,13	2.45	3 (37%)	9,19,19	2.30	2 (22%)
3	AZM	C	402	2	8,13,13	2.23	2 (25%)	9,19,19	2.67	5 (55%)
4	2HP	H	405	-	4,4,4	2.20	2 (50%)	6,6,6	0.65	0
3	AZM	F	402	2	8,13,13	2.40	2 (25%)	9,19,19	2.18	4 (44%)
3	AZM	G	402	2	8,13,13	2.50	3 (37%)	9,19,19	2.34	4 (44%)
4	2HP	E	403	-	4,4,4	2.17	1 (25%)	6,6,6	0.57	0
4	2HP	F	403	-	4,4,4	2.55	2 (50%)	6,6,6	0.80	0
4	2HP	C	403	-	4,4,4	1.57	1 (25%)	6,6,6	0.73	0
4	2HP	B	404	-	4,4,4	2.26	2 (50%)	6,6,6	0.73	0
3	AZM	D	402	2	8,13,13	2.25	2 (25%)	9,19,19	2.89	4 (44%)
4	2HP	C	404	-	4,4,4	1.76	1 (25%)	6,6,6	0.89	0
4	2HP	D	403	-	4,4,4	2.77	2 (50%)	6,6,6	0.58	0
3	AZM	A	402	2	8,13,13	2.14	2 (25%)	9,19,19	2.42	5 (55%)

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. '2' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	AZM	H	402	2	-	0/2/10/10	0/1/1/1
3	AZM	D	402	2	-	0/2/10/10	0/1/1/1
3	AZM	C	402	2	-	0/2/10/10	0/1/1/1
3	AZM	E	402	2	-	0/2/10/10	0/1/1/1
3	AZM	F	402	2	-	0/2/10/10	0/1/1/1
3	AZM	A	402	2	-	0/2/10/10	0/1/1/1
3	AZM	B	402	2	-	0/2/10/10	0/1/1/1
3	AZM	G	402	2	-	0/2/10/10	0/1/1/1

All (42) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	402	AZM	N3-N2	-5.75	1.26	1.37
3	D	402	AZM	N3-N2	-5.40	1.27	1.37
3	C	402	AZM	N3-N2	-5.40	1.27	1.37
3	H	402	AZM	N3-N2	-5.32	1.27	1.37

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	402	AZM	N3-N2	-5.17	1.27	1.37
3	E	402	AZM	N3-N2	-4.84	1.28	1.37
3	F	402	AZM	N3-N2	-4.75	1.28	1.37
3	E	402	AZM	S1-N1	4.60	1.69	1.60
3	G	402	AZM	N3-N2	-4.56	1.28	1.37
3	G	402	AZM	S1-N1	4.53	1.69	1.60
3	F	402	AZM	S1-N1	4.43	1.69	1.60
4	E	404	2HP	P-O3	3.94	1.66	1.54
4	F	403	2HP	P-O3	3.88	1.66	1.54
4	D	403	2HP	P-O2	3.83	1.66	1.54
4	E	403	2HP	P-O1	3.77	1.59	1.50
3	H	402	AZM	S1-N1	3.64	1.67	1.60
4	D	403	2HP	P-O1	3.57	1.59	1.50
4	C	404	2HP	P-O1	3.44	1.58	1.50
3	B	402	AZM	S1-N1	3.30	1.66	1.60
4	B	403	2HP	P-O1	3.26	1.58	1.50
4	B	404	2HP	P-O1	3.24	1.58	1.50
4	C	405	2HP	P-O1	3.19	1.58	1.50
4	H	405	2HP	P-O2	3.09	1.63	1.54
4	E	404	2HP	P-O2	2.95	1.63	1.54
4	H	405	2HP	P-O1	2.94	1.57	1.50
4	E	404	2HP	P-O1	2.91	1.57	1.50
4	C	403	2HP	P-O1	2.87	1.57	1.50
4	F	403	2HP	P-O1	2.87	1.57	1.50
4	B	404	2HP	P-O3	2.80	1.63	1.54
4	A	405	2HP	P-O1	2.77	1.57	1.50
4	G	403	2HP	P-O1	2.76	1.57	1.50
4	G	403	2HP	P-O4	2.69	1.62	1.54
4	B	403	2HP	P-O4	2.64	1.62	1.54
3	G	402	AZM	C2-N4	2.56	1.41	1.36
4	E	404	2HP	P-O4	2.47	1.62	1.54
3	A	402	AZM	S1-N1	2.45	1.65	1.60
3	D	402	AZM	S1-N1	2.44	1.65	1.60
4	A	404	2HP	P-O2	2.38	1.61	1.54
4	H	404	2HP	P-O1	2.28	1.56	1.50
3	C	402	AZM	C2-N4	2.03	1.40	1.36
3	H	402	AZM	C2-N4	2.01	1.40	1.36
4	A	403	2HP	P-O1	2.01	1.55	1.50

All (33) bond angle outliers are listed below:

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	402	AZM	O2-S1-C1	-7.52	99.76	108.59
3	D	402	AZM	O2-S1-C1	-5.69	101.91	108.59
3	H	402	AZM	O1-S1-C1	-5.01	102.71	108.59
3	E	402	AZM	O2-S1-O1	4.71	126.50	118.76
3	B	402	AZM	O2-S1-O1	4.66	126.41	118.76
3	A	402	AZM	O2-S1-C1	-4.44	103.38	108.59
3	D	402	AZM	O1-S1-N1	4.13	113.49	107.36
3	G	402	AZM	O2-S1-O1	4.08	125.47	118.76
3	C	402	AZM	C4-C3-N4	4.08	120.97	114.98
3	B	402	AZM	O1-S1-C1	-3.94	103.97	108.59
3	H	402	AZM	O2-S1-O1	3.91	125.19	118.76
3	C	402	AZM	O2-S1-O1	3.86	125.10	118.76
3	C	402	AZM	O2-S1-C1	-3.83	104.09	108.59
3	G	402	AZM	O1-S1-C1	-3.77	104.16	108.59
3	F	402	AZM	O2-S1-O1	3.60	124.68	118.76
3	D	402	AZM	O2-S1-O1	3.16	123.95	118.76
3	D	402	AZM	O2-S1-N1	-3.12	102.74	107.36
3	A	402	AZM	O2-S1-O1	3.09	123.84	118.76
3	G	402	AZM	C4-C3-N4	2.97	119.34	114.98
3	F	402	AZM	O1-S1-N1	-2.90	103.06	107.36
3	E	402	AZM	O2-S1-C1	-2.74	105.37	108.59
3	E	402	AZM	O1-S1-C1	-2.58	105.56	108.59
3	A	402	AZM	O1-S1-C1	-2.53	105.61	108.59
3	B	402	AZM	O1-S1-N1	2.47	111.02	107.36
3	F	402	AZM	O2-S1-C1	-2.42	105.75	108.59
3	C	402	AZM	C1-S1-N1	-2.40	103.82	108.25
3	A	402	AZM	C4-C3-N4	2.39	118.49	114.98
3	B	402	AZM	O2-S1-N1	-2.36	103.87	107.36
3	F	402	AZM	O1-S1-C1	-2.33	105.86	108.59
3	G	402	AZM	O3-C3-C4	-2.27	117.83	122.06
3	A	402	AZM	O1-S1-N1	2.23	110.67	107.36
3	E	402	AZM	O1-S1-N1	-2.14	104.19	107.36
3	C	402	AZM	O3-C3-C4	-2.11	118.14	122.06

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 [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

### 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	237/239 (99%)	-0.29	3 (1%) 77 73	15, 28, 45, 61	0
1	B	237/239 (99%)	-0.24	1 (0%) 92 91	16, 28, 46, 63	0
1	C	239/239 (100%)	-0.25	2 (0%) 86 84	17, 28, 44, 65	0
1	D	237/239 (99%)	-0.23	2 (0%) 86 84	18, 31, 53, 63	0
1	E	238/239 (99%)	-0.12	6 (2%) 57 51	16, 32, 56, 93	0
1	F	238/239 (99%)	-0.08	8 (3%) 45 38	16, 35, 63, 97	0
1	G	238/239 (99%)	-0.19	3 (1%) 77 73	14, 32, 56, 72	0
1	H	239/239 (100%)	-0.20	3 (1%) 77 73	19, 32, 54, 73	0
All	All	1903/1912 (99%)	-0.20	28 (1%) 73 70	14, 31, 53, 97	0

All (28) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	310	LEU	4.4
1	F	310	LEU	4.1
1	E	106	SER	3.9
1	C	72	MET	3.1
1	F	104	ASN	2.9
1	B	92	THR	2.7
1	H	104	ASN	2.7
1	E	103	LEU	2.6
1	F	105	THR	2.5
1	G	310	LEU	2.5
1	D	92	THR	2.5
1	A	92	THR	2.4
1	C	91	ALA	2.4
1	A	310	LEU	2.3
1	F	223	LYS	2.3
1	D	240	LYS	2.3

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	101	ILE	2.3
1	F	229	ILE	2.3
1	E	107	ALA	2.3
1	E	104	ASN	2.2
1	H	126	LYS	2.2
1	F	109	LYS	2.2
1	F	126	LYS	2.2
1	F	226	PRO	2.1
1	E	105	THR	2.1
1	G	104	ASN	2.1
1	G	81	ALA	2.0
1	H	87	LYS	2.0

## 6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

## 6.3 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
4	2HP	B	404	5/5	0.83	0.20	29,44,53,58	0
4	2HP	E	404	5/5	0.92	0.14	32,36,58,60	0
4	2HP	B	403	5/5	0.93	0.18	34,45,66,75	0
4	2HP	D	403	5/5	0.94	0.19	34,47,63,65	0
4	2HP	E	403	5/5	0.94	0.18	50,59,70,74	0
4	2HP	C	405	5/5	0.94	0.15	47,59,63,63	0
4	2HP	A	405	5/5	0.95	0.18	38,39,56,69	0
4	2HP	F	403	5/5	0.95	0.12	30,42,55,60	0
4	2HP	G	403	5/5	0.96	0.12	23,30,44,50	0
4	2HP	H	404	5/5	0.96	0.17	30,37,51,55	0
4	2HP	A	404	5/5	0.97	0.13	27,32,37,39	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
4	2HP	H	405	5/5	0.97	0.08	40,42,54,66	0
4	2HP	C	404	5/5	0.98	0.14	17,19,23,26	0
3	AZM	G	402	13/13	0.98	0.14	13,41,60,67	0
3	AZM	H	402	13/13	0.98	0.17	25,41,70,70	0
4	2HP	A	403	5/5	0.98	0.15	26,26,31,33	0
3	AZM	A	402	13/13	0.98	0.18	16,40,53,56	0
3	AZM	D	402	13/13	0.98	0.15	18,39,54,54	0
3	AZM	E	402	13/13	0.98	0.17	19,34,70,70	0
4	2HP	H	403	5/5	0.98	0.14	26,31,42,42	0
3	AZM	F	402	13/13	0.98	0.16	18,42,60,69	0
4	2HP	C	403	5/5	0.98	0.15	20,25,25,28	0
2	ZN	C	401	1/1	0.99	0.17	25,25,25,25	0
3	AZM	B	402	13/13	0.99	0.20	19,37,78,78	0
3	AZM	C	402	13/13	0.99	0.20	20,43,50,50	0
2	ZN	F	401	1/1	0.99	0.17	24,24,24,24	0
2	ZN	G	401	1/1	0.99	0.16	24,24,24,24	0
2	ZN	H	401	1/1	0.99	0.16	26,26,26,26	0
2	ZN	E	401	1/1	1.00	0.18	22,22,22,22	0
2	ZN	B	401	1/1	1.00	0.19	25,25,25,25	0
2	ZN	A	401	1/1	1.00	0.16	22,22,22,22	0
2	ZN	D	401	1/1	1.00	0.16	24,24,24,24	0

## 6.5 Other polymers [\(i\)](#)

There are no such residues in this entry.