



# Full wwPDB X-ray Structure Validation Report ⓘ

Sep 22, 2014 – 07:34 PM EDT

PDB ID : 4QA7  
Title : Crystal structure of H334R/Y306F HDAC8 in complex with a tetrapeptide substrate  
Authors : Decroos, C.; Bowman, C.B.; Moser, J.-A.S.; Christianson, K.E.; Deardorff, M.A.; Christianson, D.W.  
Deposited on : 2014-05-02  
Resolution : 2.31 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>

---

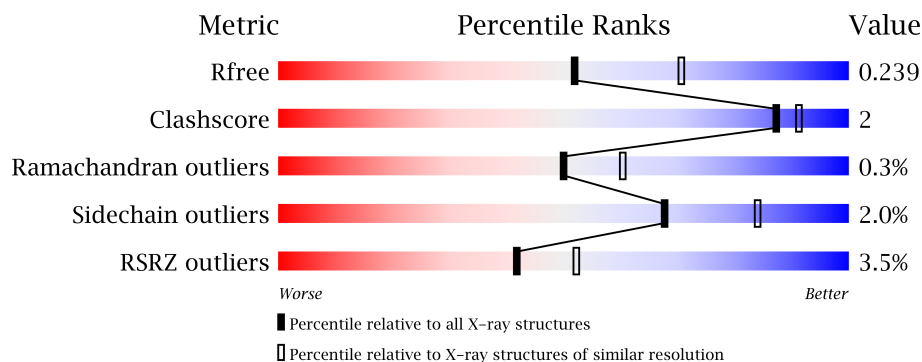
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.16 November 2013  
Xtriage (Phenix) : dev-1439  
EDS : stable23489  
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) : stable23489

# 1 Overall quality at a glance

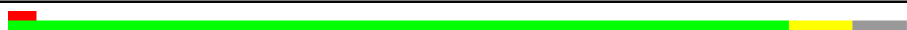

The reported resolution of this entry is 2.31 Å.

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	3293 (2.34-2.30)
Clashscore	79885	4097 (2.34-2.30)
Ramachandran outliers	78287	4055 (2.34-2.30)
Sidechain outliers	78261	4054 (2.34-2.30)
RSRZ outliers	66119	3294 (2.34-2.30)

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	389	
2	B	5	

## 2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 2942 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 Histone deacetylase 8.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	365	Total	C	N	O	S	0	2	0
			2823	1809	469	526	19			

There are 14 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	306	PHE	TYR	ENGINEERED MUTATION	UNP Q9BY41
A	334	ARG	HIS	ENGINEERED MUTATION	UNP Q9BY41
A	378	ILE	-	EXPRESSION TAG	UNP Q9BY41
A	379	GLU	-	EXPRESSION TAG	UNP Q9BY41
A	380	GLY	-	EXPRESSION TAG	UNP Q9BY41
A	381	ARG	-	EXPRESSION TAG	UNP Q9BY41
A	382	GLY	-	EXPRESSION TAG	UNP Q9BY41
A	383	SER	-	EXPRESSION TAG	UNP Q9BY41
A	384	HIS	-	EXPRESSION TAG	UNP Q9BY41
A	385	HIS	-	EXPRESSION TAG	UNP Q9BY41
A	386	HIS	-	EXPRESSION TAG	UNP Q9BY41
A	387	HIS	-	EXPRESSION TAG	UNP Q9BY41
A	388	HIS	-	EXPRESSION TAG	UNP Q9BY41
A	389	HIS	-	EXPRESSION TAG	UNP Q9BY41

- Molecule 2 is a protein called tetrapeptide substrate.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	B	5	Total	C	N	O	0	0	0
			48	30	11	7			

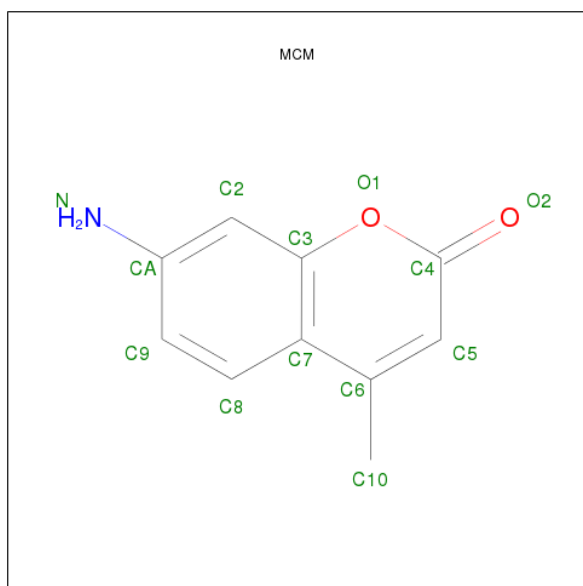
- Molecule 3 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	1	Total	Zn	0	0
			1	1		

- Molecule 4 is POTASSIUM ION (three-letter code: K) (formula: K).

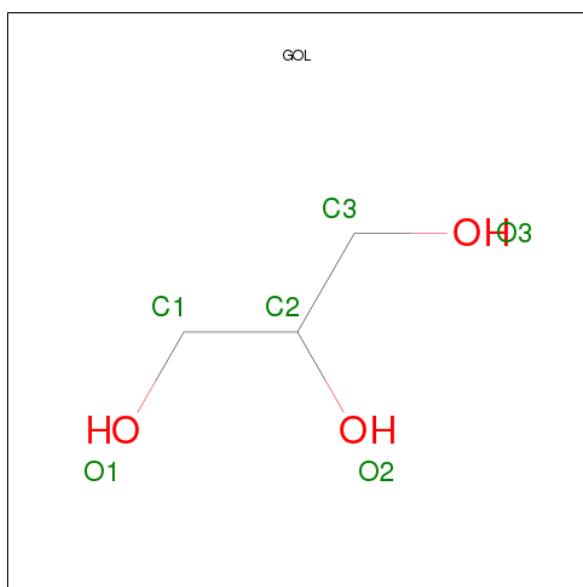
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	2	Total K 2 2	0	0

- Molecule 5 is 7-AMINO-4-METHYL-CHROMEN-2-ONE (three-letter code: MCM) (formula: C<sub>10</sub>H<sub>9</sub>NO<sub>2</sub>).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	B	1	Total C N O 13 10 1 2	0	0

- Molecule 6 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
6	B	1	Total	C	O	0	0
			6	3	3		

- Molecule 7 is water.

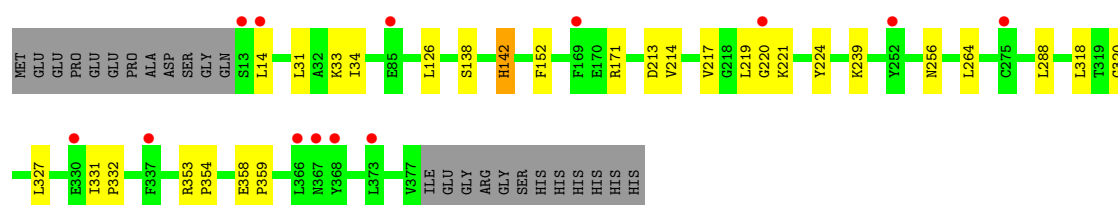
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	46	Total	O	0	0
			46	46		
7	B	3	Total	O	0	0
			3	3		

### 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: Histone deacetylase 8

Chain A: 



- Molecule 2: tetrapeptide substrate

Chain B: 

There are no outlier residues recorded for this chain.

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 31 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	80.41Å 80.41Å 106.06Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	42.19 – 2.31 42.19 – 2.31	Depositor EDS
% Data completeness (in resolution range)	99.9 (42.19-2.31) 100.0 (42.19-2.31)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.74 (at 2.32Å)	Xtriage
Refinement program	PHENIX (phenix.refine: dev_1370)	Depositor
R, $R_{free}$	0.204 , 0.241 0.203 , 0.239	Depositor DCC
$R_{free}$ test set	1813 reflections (10.14%)	DCC
Wilson B-factor (Å <sup>2</sup> )	54.9	Xtriage
Anisotropy	0.244	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 41.4	EDS
Estimated twinning fraction	0.033 for -h,-k,l	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Outliers	0 of 17893 reflections	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	2942	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	57.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.07% of the height of the origin peak. No significant pseudotranslation is detected.*

<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, ZN, ACE, K, MCM, ALY

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	0.21	0/2899	0.38	0/3939
2	B	0.14	0/22	0.31	0/28
All	All	0.20	0/2921	0.38	0/3967

There are no bond length outliers.

There are no bond angle outliers.

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	2823	0	2750	14	0
2	B	48	0	49	0	0
3	A	1	0	0	0	0
4	A	2	0	0	0	0
5	B	13	0	7	0	0
6	B	6	0	8	0	0
7	A	46	0	0	1	0
7	B	3	0	0	0	0
All	All	2942	0	2814	14	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 2.

All (14) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:171:ARG:NH1	1:A:256:ASN:O	2.27	0.68
1:A:320:GLY:HA3	1:A:327:LEU:HD21	1.83	0.60
1:A:142:HIS:NE2	7:A:522:HOH:O	2.31	0.59
1:A:264:LEU:HD13	1:A:318:LEU:HD13	1.92	0.52
1:A:353:ARG:HG3	1:A:354:PRO:HD2	1.94	0.49
1:A:31:LEU:HD23	1:A:34:ILE:HG12	1.95	0.48
1:A:171:ARG:HH12	1:A:256:ASN:HB3	1.78	0.48
1:A:33:LYS:HG3	1:A:152:PHE:CD2	2.49	0.46
1:A:219:LEU:O	1:A:221:LYS:N	2.50	0.44
1:A:358:GLU:HA	1:A:359:PRO:HD3	1.87	0.43
1:A:217:VAL:O	1:A:224:TYR:N	2.51	0.42
1:A:142:HIS:CD2	1:A:142:HIS:H	2.38	0.42
1:A:331:ILE:HA	1:A:332:PRO:HD3	1.76	0.42
1:A:213:ASP:OD2	1:A:214:VAL:N	2.53	0.41

There are no symmetry-related clashes.

## 5.3 Torsion angles

### 5.3.1 Protein backbone ⓘ

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	365/389 (94%)	354 (97%)	10 (3%)	1 (0%)	50	60
2	B	2/5 (40%)	2 (100%)	0	0	100	100
All	All	367/394 (93%)	356 (97%)	10 (3%)	1 (0%)	50	60

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	220	GLY

### 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	300/326 (92%)	293 (98%)	7 (2%)	63	80
2	B	2/2 (100%)	2 (100%)	0	100	100
All	All	302/328 (92%)	295 (98%)	7 (2%)	68	80

All (7) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	14	LEU
1	A	126	LEU
1	A	138	SER
1	A	142	HIS
1	A	239[A]	LYS
1	A	239[B]	LYS
1	A	288	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	A	236	GLN
1	A	363	GLN

### 5.3.3 RNA ⓘ

There are no RNA chains in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

2 non-standard protein/DNA/RNA residues 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	ALY	B	504	2	11,11,12	5.48	2 (18%)	10,12,14	2.36	2 (20%)
2	ALY	B	505	3,2,5	11,11,12	5.60	1 (9%)	10,12,14	1.04	1 (10%)

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	ALY	B	504	2	-	0/8/10/12	0/0/0/0
2	ALY	B	505	3,2,5	-	0/8/10/12	0/0/0/0

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	505	ALY	O-C	18.34	1.24	1.11
2	B	504	ALY	O-C	17.90	1.23	1.11
2	B	504	ALY	CH-NZ	2.04	1.38	1.33

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	504	ALY	C-CA-N	-6.72	107.11	113.83
2	B	504	ALY	CB-CA-N	2.59	117.72	110.35
2	B	505	ALY	C-CA-N	-2.23	111.60	113.83

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

## 5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 5.6 Ligand geometry ⓘ

Of 5 ligands modelled in this entry, 3 are monoatomic - leaving 2 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	$\# Z  > 2$
5	MCM	B	601	2	14,14,14	3.19	3 (21%)	18,20,20	1.76	4 (22%)
6	GOL	B	602	-	5,5,5	0.35	0	5,5,5	0.24	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
5	MCM	B	601	2	-	0/0/0/0	0/2/2/2
6	GOL	B	602	-	-	0/4/4/4	0/0/0/0

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	B	601	MCM	O1-C4	9.76	1.38	1.33
5	B	601	MCM	C7-C3	5.46	1.48	1.41
5	B	601	MCM	C6-C7	3.01	1.48	1.42

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	B	601	MCM	O1-C3-C2	4.55	121.53	115.97
5	B	601	MCM	C2-C3-C7	-3.59	119.03	123.11
5	B	601	MCM	C8-C7-C3	2.47	119.41	116.39
5	B	601	MCM	O1-C4-C5	2.14	122.12	119.53

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	365/389 (93%)	0.30	13 (3%) 41 51	44, 56, 75, 88	0
2	B	5/5 (100%)	0.71	0 100 100	48, 52, 67, 73	0
All	All	370/394 (93%)	0.30	13 (3%) 42 52	44, 56, 75, 88	0

All (13) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	13	SER	5.4
1	A	14	LEU	4.4
1	A	220	GLY	3.5
1	A	366	LEU	2.4
1	A	85	GLU	2.4
1	A	252	TYR	2.3
1	A	368	TYR	2.3
1	A	367	ASN	2.2
1	A	169	PHE	2.2
1	A	373	LEU	2.2
1	A	330	GLU	2.1
1	A	275	CYS	2.0
1	A	337	PHE	2.0

### 6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

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( $\text{\AA}^2$ )	Q<0.9
2	ALY	B	505	12/13	0.21	1.37	44,48,51,51	0
2	ALY	B	504	12/13	0.22	1.17	49,51,54,61	0

### 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( $\text{\AA}^2$ )	Q<0.9
4	K	A	402	1/1	0.16	0.52	48,48,48,48	0
3	ZN	A	401	1/1	0.16	-0.12	47,47,47,47	0
6	GOL	B	602	6/6	0.16	-0.57	62,64,70,73	0
5	MCM	B	601	13/13	0.18	-0.59	45,49,51,52	0
4	K	A	403	1/1	0.09	-1.63	59,59,59,59	0

### 6.5 Other polymers

There are no such residues in this entry.