



# Full wwPDB X-ray Structure Validation Report ⓘ

May 15, 2020 – 07:39 am BST

PDB ID : 3TOB  
Title : Human MOF E350Q crystal structure with active site lysine partially acetylated  
Authors : Yuan, H.; Ding, E.C.; Marmorstein, R.  
Deposited on : 2011-09-04  
Resolution : 2.70 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 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	:	2.11
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)	:	2.11

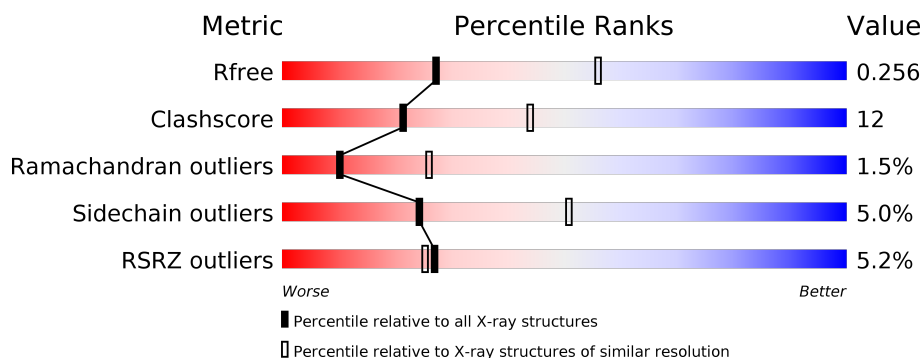
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.70 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	2808 (2.70-2.70)
Clashscore	141614	3122 (2.70-2.70)
Ramachandran outliers	138981	3069 (2.70-2.70)
Sidechain outliers	138945	3069 (2.70-2.70)
RSRZ outliers	127900	2737 (2.70-2.70)

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	304	<div> <div>5%</div> <div> <div></div> <div>67%</div> <div>18%</div> <div>• •</div> <div>11%</div> </div> </div>

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 2267 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 histone acetyltransferase MYST1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	269	Total	C	N	O	S	0	4	0
			2234	1456	362	405	11			

There are 23 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	156	HIS	-	EXPRESSION TAG	UNP Q9H7Z6
A	157	HIS	-	EXPRESSION TAG	UNP Q9H7Z6
A	158	HIS	-	EXPRESSION TAG	UNP Q9H7Z6
A	159	HIS	-	EXPRESSION TAG	UNP Q9H7Z6
A	160	HIS	-	EXPRESSION TAG	UNP Q9H7Z6
A	161	HIS	-	EXPRESSION TAG	UNP Q9H7Z6
A	162	ASP	-	EXPRESSION TAG	UNP Q9H7Z6
A	163	TYR	-	EXPRESSION TAG	UNP Q9H7Z6
A	164	ASP	-	EXPRESSION TAG	UNP Q9H7Z6
A	165	ILE	-	EXPRESSION TAG	UNP Q9H7Z6
A	166	PRO	-	EXPRESSION TAG	UNP Q9H7Z6
A	167	THR	-	EXPRESSION TAG	UNP Q9H7Z6
A	168	THR	-	EXPRESSION TAG	UNP Q9H7Z6
A	169	GLU	-	EXPRESSION TAG	UNP Q9H7Z6
A	170	ASN	-	EXPRESSION TAG	UNP Q9H7Z6
A	171	LEU	-	EXPRESSION TAG	UNP Q9H7Z6
A	172	TYR	-	EXPRESSION TAG	UNP Q9H7Z6
A	173	PHE	-	EXPRESSION TAG	UNP Q9H7Z6
A	174	GLN	-	EXPRESSION TAG	UNP Q9H7Z6
A	175	GLY	-	EXPRESSION TAG	UNP Q9H7Z6
A	176	SER	-	EXPRESSION TAG	UNP Q9H7Z6
A	274	ALY	LYS	MICROHETEROGENEITY	UNP Q9H7Z6
A	350	GLN	GLU	ENGINEERED MUTATION	UNP Q9H7Z6

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

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

- Molecule 3 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

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

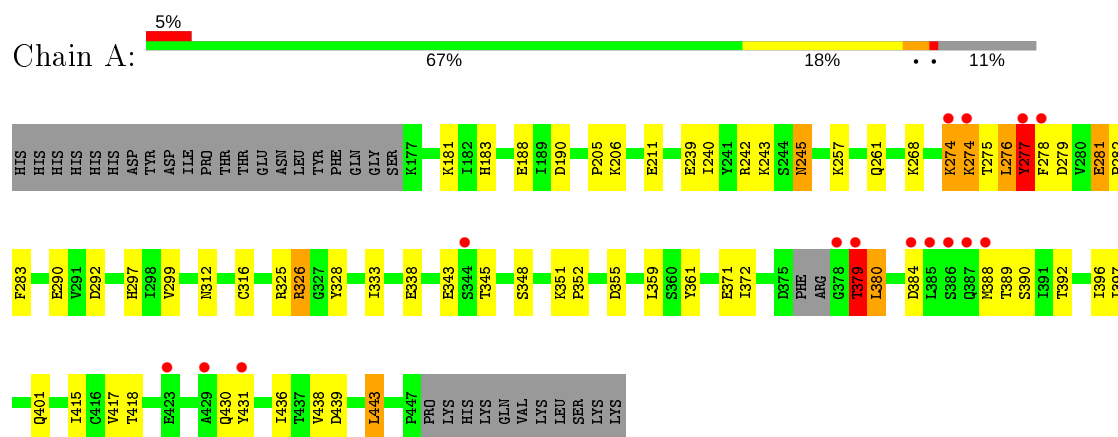
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	31	Total	O	0	0
			31	31		

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

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density ( $RSRZ > 2$ ). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

- Molecule 1: histone acetyltransferase MYST1



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	45.97Å 58.04Å 120.34Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	34.52 – 2.70 34.52 – 2.70	Depositor EDS
% Data completeness (in resolution range)	99.8 (34.52-2.70) 99.2 (34.52-2.70)	Depositor EDS
$R_{merge}$	0.10	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.71 (at 2.68Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.7_650)	Depositor
R, $R_{free}$	0.203 , 0.253 0.197 , 0.256	Depositor DCC
$R_{free}$ test set	443 reflections (4.76%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	38.5	Xtriage
Anisotropy	0.304	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 40.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.46$ , $\langle L^2 \rangle = 0.29$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	2267	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	36.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.64% 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: ZN, CL, 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	0.47	1/2286 (0.0%)	1.03	3/3101 (0.1%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	274[B]	LYS	C-N	-5.88	1.20	1.34

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	274[B]	LYS	C-N-CA	-35.03	34.12	121.70
1	A	274[B]	LYS	CA-C-N	-29.41	52.49	117.20
1	A	274[B]	LYS	O-C-N	15.72	147.85	122.70

There are no chirality outliers.

There are no planarity outliers.

### 5.2 Too-close contacts

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2234	0	2174	50	0
2	A	1	0	0	0	0
3	A	1	0	0	0	0
4	A	31	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	2267	0	2174	50	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 12.

All (50) 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:A:275[B]:THR:HG22	1:A:276[B]:LEU:H	1.11	1.09
1:A:275[B]:THR:HA	1:A:277:TYR:CZ	2.03	0.94
1:A:275[B]:THR:HG22	1:A:276[B]:LEU:N	1.92	0.83
1:A:245:ASN:HD22	1:A:245:ASN:H	1.29	0.80
1:A:275[B]:THR:CG2	1:A:276[B]:LEU:H	1.96	0.74
1:A:211:GLU:OE1	1:A:297:HIS:HD2	1.74	0.71
1:A:275[B]:THR:O	1:A:276[B]:LEU:CB	2.40	0.69
1:A:276[B]:LEU:HD21	1:A:279:ASP:O	1.94	0.67
1:A:275[B]:THR:O	1:A:276[B]:LEU:HB3	1.99	0.62
1:A:325:ARG:O	1:A:326:ARG:HD3	2.01	0.61
1:A:372:ILE:HD13	1:A:388:MET:HE2	1.82	0.61
1:A:292:ASP:OD1	1:A:297:HIS:HE1	1.82	0.61
1:A:380:LEU:HD12	1:A:384:ASP:HB3	1.83	0.60
1:A:379:THR:HG23	1:A:379:THR:O	2.02	0.60
1:A:276[B]:LEU:CD2	1:A:279:ASP:O	2.50	0.59
1:A:188:GLU:OE2	1:A:206:LYS:HE3	2.06	0.55
1:A:281:GLU:N	1:A:282:PRO:CD	2.70	0.54
1:A:257:LYS:O	1:A:261:GLN:HG3	2.08	0.53
1:A:181:LYS:HG2	1:A:206:LYS:HG3	1.92	0.51
1:A:338:GLU:HG3	1:A:438:VAL:HG11	1.92	0.51
1:A:275[B]:THR:HA	1:A:277:TYR:CE2	2.46	0.50
1:A:397:ILE:O	1:A:401:GLN:HG3	2.11	0.49
1:A:268:LYS:HD3	1:A:277:TYR:HE2	1.78	0.49
1:A:333:ILE:HG23	1:A:361:TYR:CE2	2.47	0.48
1:A:379:THR:O	1:A:379:THR:CG2	2.62	0.47
1:A:290:GLU:HG2	1:A:299:VAL:HG21	1.95	0.47
1:A:355:ASP:O	1:A:359:LEU:HD13	2.14	0.47
1:A:243:LYS:HE2	1:A:439:ASP:OD2	2.15	0.47
1:A:275[B]:THR:CG2	1:A:276[B]:LEU:N	2.64	0.47
1:A:312:ASN:HA	1:A:348:SER:O	2.16	0.46
1:A:276[B]:LEU:O	1:A:278:PHE:N	2.40	0.46
1:A:380:LEU:HD12	1:A:384:ASP:CB	2.46	0.46
1:A:275[B]:THR:O	1:A:276[B]:LEU:HB2	2.16	0.45

*Continued on next page...*



Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:430:GLN:HB2	1:A:431:TYR:CE2	2.52	0.45
1:A:343:GLU:HB2	1:A:345:THR:HG22	2.00	0.44
1:A:239:GLU:OE2	1:A:242:ARG:HD3	2.18	0.44
1:A:380:LEU:HD13	1:A:380:LEU:HA	1.87	0.43
1:A:392:THR:O	1:A:396:ILE:HG13	2.18	0.43
1:A:245:ASN:HD22	1:A:245:ASN:N	2.00	0.43
1:A:274[A]:ALY:HG2	1:A:316:CYS:SG	2.59	0.43
1:A:240:ILE:O	1:A:443:LEU:HD23	2.19	0.43
1:A:372:ILE:HD13	1:A:388:MET:CE	2.49	0.42
1:A:181:LYS:HE3	1:A:190:ASP:OD1	2.19	0.42
1:A:389:THR:O	1:A:390:SER:HB2	2.19	0.42
1:A:276[B]:LEU:HD21	1:A:283:PHE:CE2	2.55	0.42
1:A:183:HIS:HD2	1:A:188:GLU:OE2	2.03	0.42
1:A:351:LYS:HB3	1:A:352:PRO:HA	2.03	0.41
1:A:299:VAL:HG11	1:A:328:TYR:CD1	2.55	0.41
1:A:276[B]:LEU:HD21	1:A:283:PHE:HE2	1.86	0.41
1:A:181:LYS:HG2	1:A:206:LYS:CG	2.51	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	268/304 (88%)	246 (92%)	17 (6%)	5 (2%)	8 20

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	276[A]	LEU
1	A	276[B]	LEU
1	A	277	TYR
1	A	379	THR

Continued on next page...

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	205	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	244/281 (87%)	232 (95%)	12 (5%)	25 52

All (12) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	245	ASN
1	A	277	TYR
1	A	281	GLU
1	A	326	ARG
1	A	371	GLU
1	A	379	THR
1	A	380	LEU
1	A	415	ILE
1	A	417	VAL
1	A	418	THR
1	A	436	ILE
1	A	443	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (9) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	183	HIS
1	A	186	ASN
1	A	204	GLN
1	A	231	GLN
1	A	234	GLN
1	A	245	ASN
1	A	256	HIS
1	A	297	HIS

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	350	GLN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

1 non-standard protein/DNA/RNA residue is 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	$\# Z  > 2$
1	ALY	A	274[A]	1	10,11,12	0.46	0	7,12,14	0.77	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
1	ALY	A	274[A]	1	-	1/9/10/12	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	A	274[A]	ALY	CE-CD-CG-CB

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	A	274[A]	ALY	1	0

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 2 ligands modelled in this entry, 2 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

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 [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	268/304 (88%)	-0.02	13 (4%) 29 28	18, 33, 73, 89	0

All (13) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	378	GLY	5.0
1	A	277	TYR	4.3
1	A	278	PHE	4.3
1	A	431	TYR	2.8
1	A	385	LEU	2.7
1	A	387	GLN	2.6
1	A	379	THR	2.5
1	A	344	SER	2.5
1	A	388	MET	2.3
1	A	384	ASP	2.3
1	A	386	SER	2.2
1	A	423	GLU	2.1
1	A	429	ALA	2.1

### 6.2 Non-standard residues in protein, DNA, RNA chains [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
1	ALY	A	274[A]	12/13	0.93	0.26	23,27,33,35	12

### 6.3 Carbohydrates [i](#)

There are no carbohydrates 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( $\text{\AA}^2$ )	Q<0.9
2	ZN	A	501	1/1	0.98	0.13	34,34,34,34	0
3	CL	A	502	1/1	0.98	0.12	28,28,28,28	0

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

There are no such residues in this entry.