



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 07:04 AM GMT

PDB ID : 2ZFN  
Title : Self-acetylation mediated histone H3 lysine 56 acetylation by rtt109  
Authors : Yuan, Y.A.  
Deposited on : 2008-01-08  
Resolution : 1.90 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.  
We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)  
A user guide is available at  
<http://wwpdb.org/validation/2016/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 : 1.7 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026688  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk26865

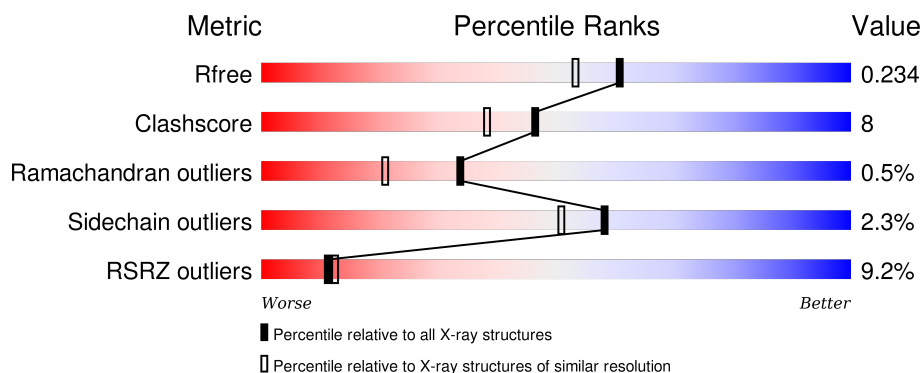
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 1.90 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))
$R_{free}$	91344	4755 (1.90-1.90)
Clashscore	102246	5398 (1.90-1.90)
Ramachandran outliers	100387	5338 (1.90-1.90)
Sidechain outliers	100360	5339 (1.90-1.90)
RSRZ outliers	91569	4766 (1.90-1.90)

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. 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	460	<div> <div>7%</div> <div>66%</div> <div>13%</div> <div>•</div> <div>18%</div> </div>

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

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
1	ALY	A	290	X	-	-	-

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	GOL	A	601	-	-	-	X

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 3324 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.

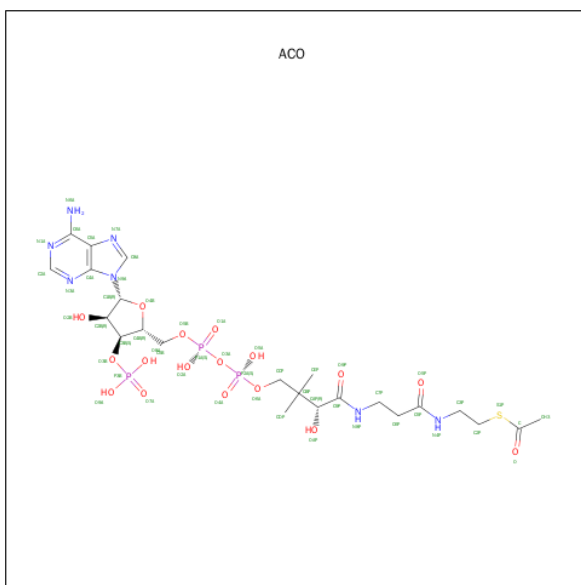
- Molecule 1 is a protein called Regulator of Ty1 transposition protein 109.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	S	Se			
1	A	376	3052	1973	506	563	4	6	1	0	0

There are 24 discrepancies between the modelled and reference sequences:

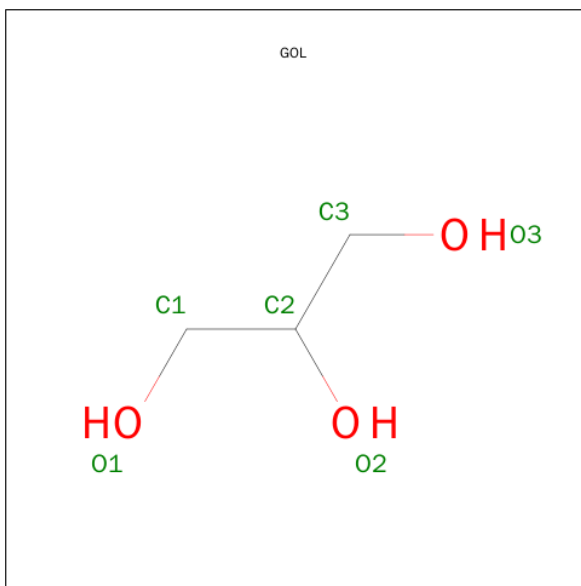
Chain	Residue	Modelled	Actual	Comment	Reference
A	-19	MET	-	EXPRESSION TAG	UNP Q07794
A	-18	GLY	-	EXPRESSION TAG	UNP Q07794
A	-17	SER	-	EXPRESSION TAG	UNP Q07794
A	-16	SER	-	EXPRESSION TAG	UNP Q07794
A	-15	HIS	-	EXPRESSION TAG	UNP Q07794
A	-14	HIS	-	EXPRESSION TAG	UNP Q07794
A	-13	HIS	-	EXPRESSION TAG	UNP Q07794
A	-12	HIS	-	EXPRESSION TAG	UNP Q07794
A	-11	HIS	-	EXPRESSION TAG	UNP Q07794
A	-10	HIS	-	EXPRESSION TAG	UNP Q07794
A	-9	SER	-	EXPRESSION TAG	UNP Q07794
A	-8	SER	-	EXPRESSION TAG	UNP Q07794
A	-7	GLY	-	EXPRESSION TAG	UNP Q07794
A	-6	LEU	-	EXPRESSION TAG	UNP Q07794
A	-5	VAL	-	EXPRESSION TAG	UNP Q07794
A	-4	PRO	-	EXPRESSION TAG	UNP Q07794
A	-3	ARG	-	EXPRESSION TAG	UNP Q07794
A	-2	GLY	-	EXPRESSION TAG	UNP Q07794
A	-1	SER	-	EXPRESSION TAG	UNP Q07794
A	0	HIS	-	EXPRESSION TAG	UNP Q07794
A	1D	MET	-	EXPRESSION TAG	UNP Q07794
A	1C	ALA	-	EXPRESSION TAG	UNP Q07794
A	1B	ALA	-	EXPRESSION TAG	UNP Q07794
A	1A	MSE	-	EXPRESSION TAG	UNP Q07794

- Molecule 2 is ACETYL COENZYME \*A (three-letter code: ACO) (formula: C<sub>23</sub>H<sub>38</sub>N<sub>7</sub>O<sub>17</sub>P<sub>3</sub>S).



Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
2	A	1	Total	C	N	O	P	S	0	0
			51	23	7	17	3	1		

- Molecule 3 is GLYCEROL (three-letter code: GOL) (formula:  $C_3H_8O_3$ ).



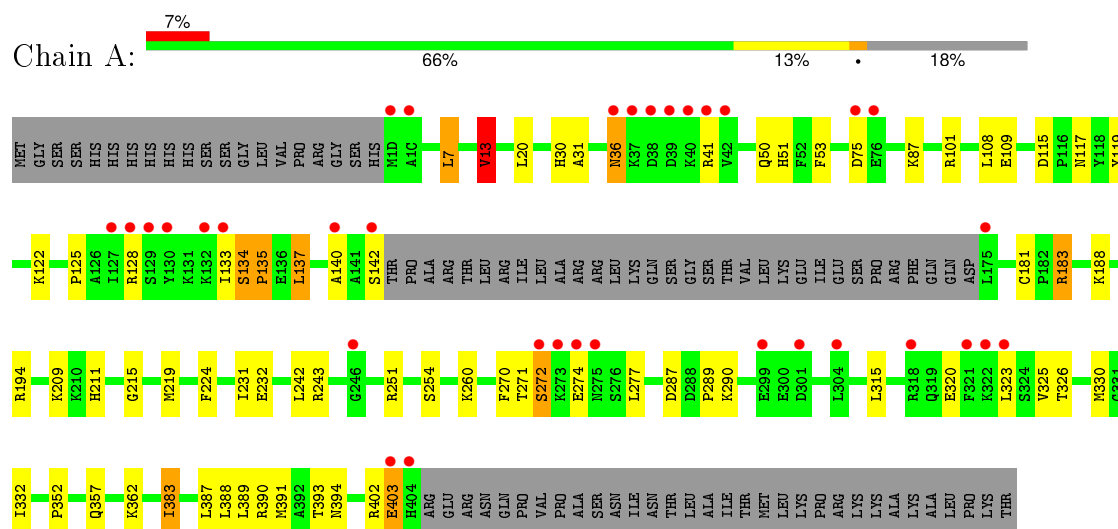
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	O	0	0
			6	3	3		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	215	Total 215	O 215	0	0



● Molecule 1: Regulator of Ty1 transposition protein 109



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	146.59Å 68.61Å 55.47Å 90.00° 95.02° 90.00°	Depositor
Resolution (Å)	42.32 – 1.90 42.32 – 1.90	Depositor EDS
% Data completeness (in resolution range)	99.9 (42.32-1.90) 99.9 (42.32-1.90)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	5.71 (at 1.89Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, $R_{free}$	0.185 , 0.229 0.191 , 0.234	Depositor DCC
$R_{free}$ test set	2178 reflections (5.30%)	DCC
Wilson B-factor (Å <sup>2</sup> )	26.1	Xtriage
Anisotropy	0.082	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.39 , 52.6	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	1 of 43234 reflections (0.002%)	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	3324	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	31.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.88% 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, 0.375 respectively for untwinned datasets, and 0.333, 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: GOL, ACO, 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	1.09	10/3104 (0.3%)	0.91	10/4190 (0.2%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	1	2

The worst 5 of 10 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	134	SER	C-O	23.36	1.67	1.23
1	A	137	LEU	C-O	15.24	1.52	1.23
1	A	134	SER	CA-C	13.12	1.87	1.52
1	A	135	PRO	N-CD	11.30	1.63	1.47
1	A	140	ALA	C-N	10.12	1.57	1.34

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	134	SER	CB-CA-C	-15.78	80.11	110.10
1	A	243	ARG	NE-CZ-NH1	8.11	124.35	120.30
1	A	243	ARG	NE-CZ-NH2	-7.53	116.53	120.30
1	A	134	SER	O-C-N	-6.83	108.13	121.10
1	A	134	SER	N-CA-C	-6.57	93.27	111.00

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	A	290	ALY	CA

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	134	SER	Mainchain
1	A	274	GLU	Peptide

## 5.2 Too-close contacts

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	3052	0	3078	47	0
2	A	51	0	34	1	0
3	A	6	0	8	0	0
4	A	215	0	0	1	0
All	All	3324	0	3120	47	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 8.

The worst 5 of 47 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:194:ARG:HH22	1:A:320:GLU:CD	1.45	1.18
1:A:194:ARG:HD3	1:A:323:LEU:HD22	1.46	0.95
1:A:388:LEU:HD21	1:A:394:ASN:OD1	1.70	0.92
1:A:194:ARG:NH2	1:A:320:GLU:CD	2.30	0.83
1:A:211:HIS:HD2	2:A:500:ACO:O4A	1.66	0.79

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	371/460 (81%)	360 (97%)	9 (2%)	2 (0%)	34	21

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	272	SER
1	A	135	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	341/410 (83%)	333 (98%)	8 (2%)	58	51

5 of 8 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	277	LEU
1	A	403	GLU
1	A	383	ILE
1	A	122	LYS
1	A	325	VAL

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 10 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	57	HIS
1	A	117	ASN
1	A	239	GLN
1	A	51	HIS
1	A	211	HIS

### 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	290	1	10,11,12	7.37	4 (40%)	10,12,14	3.06	3 (30%)

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	290	1	1/1/2/4	2/8/10/12	0/0/0/0

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	290	ALY	CB-CA	-20.55	1.33	1.53
1	A	290	ALY	CA-N	-4.18	1.34	1.47
1	A	290	ALY	CH-NZ	3.55	1.43	1.33
1	A	290	ALY	OH-CH	9.31	1.44	1.23

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	290	ALY	OH-CH-CH3	-7.40	108.48	122.06
1	A	290	ALY	OH-CH-NZ	-4.63	113.56	121.79
1	A	290	ALY	O-C-CA	2.43	131.82	125.49

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	A	290	ALY	CA

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	A	290	ALY	CH3-CH-NZ-CE
1	A	290	ALY	OH-CH-NZ-CE

There are no ring outliers.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	A	290	ALY	2	0

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

2 ligands 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	ACO	A	500	-	43,53,53	0.98	2 (4%)	55,79,79	1.49	9 (16%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
3	GOL	A	601	-	5,5,5	0.52	0	5,5,5	1.62	1 (20%)

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	ACO	A	500	-	-	0/47/67/67	0/3/3/3
3	GOL	A	601	-	-	0/4/4/4	0/0/0/0

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	500	ACO	C2A-N3A	2.07	1.35	1.32
2	A	500	ACO	C5A-C4A	3.35	1.48	1.40

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	500	ACO	N3A-C2A-N1A	-6.42	123.98	128.89
2	A	500	ACO	C1B-N9A-C4A	-2.82	122.69	126.94
2	A	500	ACO	C4A-C5A-N7A	-2.28	107.38	109.48
2	A	500	ACO	O3A-P2A-O6A	-2.27	96.92	102.94
2	A	500	ACO	O5P-C5P-C6P	-2.08	118.40	121.98

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	500	ACO	1	0

## 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 [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	369/460 (80%)	0.37	34 (9%) <b>11</b> <b>12</b>	18, 27, 52, 65	2 (0%)

The worst 5 of 34 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	1(C)	ALA	8.5
1	A	323	LEU	6.9
1	A	130	TYR	6.9
1	A	274	GLU	5.2
1	A	321	PHE	5.2

### 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. 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	RSCC	RSR	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
1	ALY	A	290	12/13	0.83	0.17	-	20,23,24,29	0

### 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. 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	RSCC	RSR	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
3	GOL	A	601	6/6	0.89	0.18	2.34	32,42,44,47	0
2	ACO	A	500	51/51	0.97	0.08	-0.92	15,22,50,52	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.