



# wwPDB X-ray Structure Validation Summary Report ⓘ

Nov 6, 2023 – 03:33 am GMT

PDB ID : 6GYT  
Title : Transcription factor dimerization activates the p300 acetyltransferase  
Authors : Panne, D.; Ortega, E.  
Deposited on : 2018-07-01  
Resolution : 2.50 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

<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)  
Xtriage (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

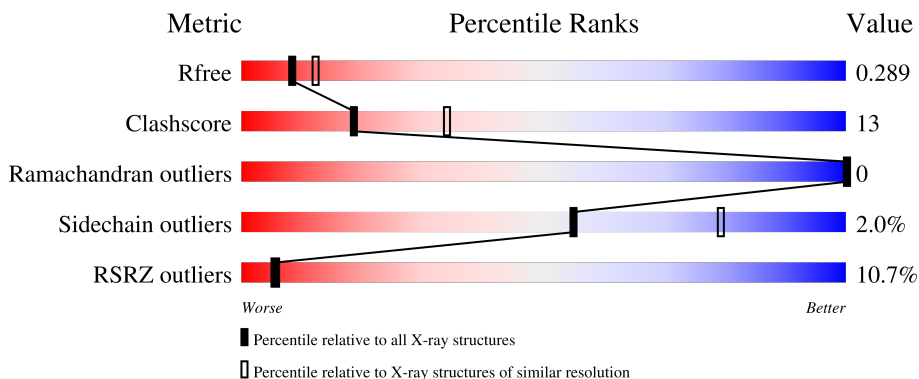
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.50 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	4661 (2.50-2.50)
Clashscore	141614	5346 (2.50-2.50)
Ramachandran outliers	138981	5231 (2.50-2.50)
Sidechain outliers	138945	5233 (2.50-2.50)
RSRZ outliers	127900	4559 (2.50-2.50)

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	161	 11% 88% 11%
2	B	168	 10% 90% 9%
3	C	9	 11% 78% 22%

## 2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 5512 atoms, of which 2715 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 p300.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace	
			Total	C	H	N	O				S
1	A	161	2634	856	1306	220	239	13	0	0	0

There are 39 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1169	GLY	-	expression tag	UNP Q09472
A	1242	GLU	-	expression tag	UNP Q09472
A	1243	LEU	-	expression tag	UNP Q09472
A	1244	PHE	-	expression tag	UNP Q09472
A	1245	VAL	-	expression tag	UNP Q09472
A	1246	GLU	-	expression tag	UNP Q09472
A	1247	CYS	-	expression tag	UNP Q09472
A	1248	THR	-	expression tag	UNP Q09472
A	1249	GLU	-	expression tag	UNP Q09472
A	1250	CYS	-	expression tag	UNP Q09472
A	1251	GLY	-	expression tag	UNP Q09472
A	1252	ARG	-	expression tag	UNP Q09472
A	1253	LYS	-	expression tag	UNP Q09472
A	1254	MET	-	expression tag	UNP Q09472
A	1255	HIS	-	expression tag	UNP Q09472
A	1256	GLN	-	expression tag	UNP Q09472
A	1257	ILE	-	expression tag	UNP Q09472
A	1258	CYS	-	expression tag	UNP Q09472
A	1259	VAL	-	expression tag	UNP Q09472
A	1260	LEU	-	expression tag	UNP Q09472
A	1261	HIS	-	expression tag	UNP Q09472
A	1262	HIS	-	expression tag	UNP Q09472
A	1263	GLU	-	expression tag	UNP Q09472
A	1264	ILE	-	expression tag	UNP Q09472
A	1265	ILE	-	expression tag	UNP Q09472
A	1266	TRP	-	expression tag	UNP Q09472
A	1267	PRO	-	expression tag	UNP Q09472

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	1268	ALA	-	expression tag	UNP Q09472
A	1269	GLY	-	expression tag	UNP Q09472
A	1270	PHE	-	expression tag	UNP Q09472
A	1271	VAL	-	expression tag	UNP Q09472
A	1272	CYS	-	expression tag	UNP Q09472
A	1273	ASP	-	expression tag	UNP Q09472
A	1274	GLY	-	expression tag	UNP Q09472
A	1275	CYS	-	expression tag	UNP Q09472
A	1276	LEU	-	expression tag	UNP Q09472
A	1277	LYS	-	expression tag	UNP Q09472
A	1278	LYS	-	expression tag	UNP Q09472
A	1279	SER	-	expression tag	UNP Q09472

- Molecule 2 is a protein called Histone acetyltransferase p300.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace	
			Total	C	H	N	O				S
2	B	168	2722	880	1351	231	247	13	0	0	0

There are 48 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	1045	ALA	-	expression tag	UNP Q09472
B	1046	GLY	-	expression tag	UNP Q09472
B	1048	ALA	ILE	conflict	UNP Q09472
B	1137	ALA	ARG	conflict	UNP Q09472
B	1241	GLY	-	expression tag	UNP Q09472
B	1242	GLU	-	expression tag	UNP Q09472
B	1243	LEU	-	expression tag	UNP Q09472
B	1244	PHE	-	expression tag	UNP Q09472
B	1245	VAL	-	expression tag	UNP Q09472
B	1246	GLU	-	expression tag	UNP Q09472
B	1247	CYS	-	expression tag	UNP Q09472
B	1248	THR	-	expression tag	UNP Q09472
B	1249	GLU	-	expression tag	UNP Q09472
B	1250	CYS	-	expression tag	UNP Q09472
B	1251	GLY	-	expression tag	UNP Q09472
B	1252	ARG	-	expression tag	UNP Q09472
B	1253	LYS	-	expression tag	UNP Q09472
B	1254	MET	-	expression tag	UNP Q09472
B	1255	HIS	-	expression tag	UNP Q09472
B	1256	GLN	-	expression tag	UNP Q09472

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	1257	ILE	-	expression tag	UNP Q09472
B	1258	CYS	-	expression tag	UNP Q09472
B	1259	VAL	-	expression tag	UNP Q09472
B	1260	LEU	-	expression tag	UNP Q09472
B	1261	HIS	-	expression tag	UNP Q09472
B	1262	HIS	-	expression tag	UNP Q09472
B	1263	GLU	-	expression tag	UNP Q09472
B	1264	ILE	-	expression tag	UNP Q09472
B	1265	ILE	-	expression tag	UNP Q09472
B	1266	TRP	-	expression tag	UNP Q09472
B	1267	PRO	-	expression tag	UNP Q09472
B	1268	ALA	-	expression tag	UNP Q09472
B	1269	GLY	-	expression tag	UNP Q09472
B	1270	PHE	-	expression tag	UNP Q09472
B	1271	VAL	-	expression tag	UNP Q09472
B	1272	CYS	-	expression tag	UNP Q09472
B	1273	ASP	-	expression tag	UNP Q09472
B	1274	GLY	-	expression tag	UNP Q09472
B	1275	CYS	-	expression tag	UNP Q09472
B	1276	LEU	-	expression tag	UNP Q09472
B	1277	LYS	-	expression tag	UNP Q09472
B	1278	LYS	-	expression tag	UNP Q09472
B	1279	SER	-	expression tag	UNP Q09472
B	1280	ALA	-	expression tag	UNP Q09472
B	1281	ARG	-	expression tag	UNP Q09472
B	1282	THR	-	expression tag	UNP Q09472
B	1283	ARG	-	expression tag	UNP Q09472
B	1284	LYS	-	expression tag	UNP Q09472

- Molecule 3 is a protein called Histone H4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	9	Total	C	H	N	O	0	0	0
			116	36	58	11	11			

- Molecule 4 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	2	Total	Zn	0	0
			2	2		
4	B	2	Total	Zn	0	0
			2	2		

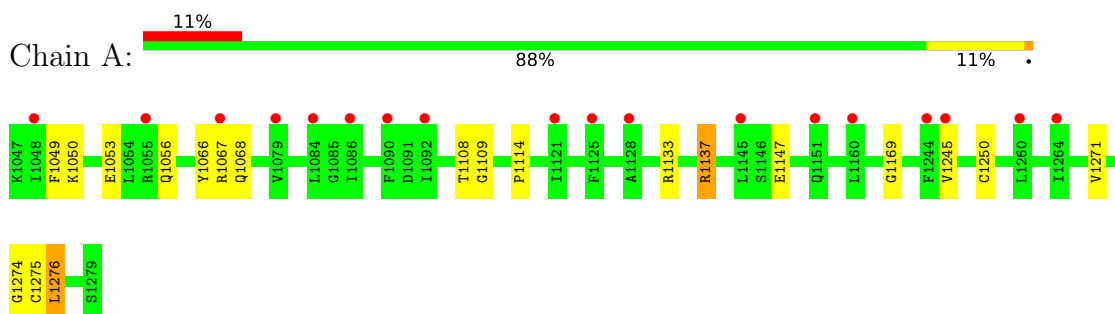
- Molecule 5 is water.

<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>		<b>ZeroOcc</b>	<b>AltConf</b>
5	A	19	Total 19	O 19	0	0
5	B	17	Total 17	O 17	0	0

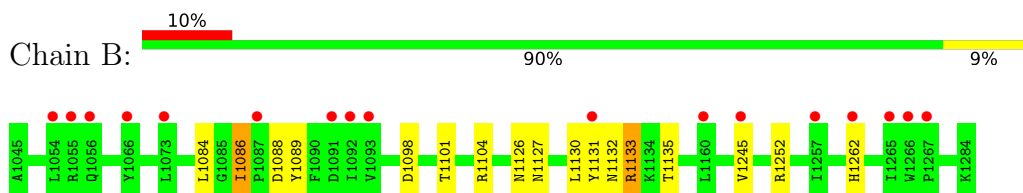
### 3 Residue-property plots [i](#)

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.

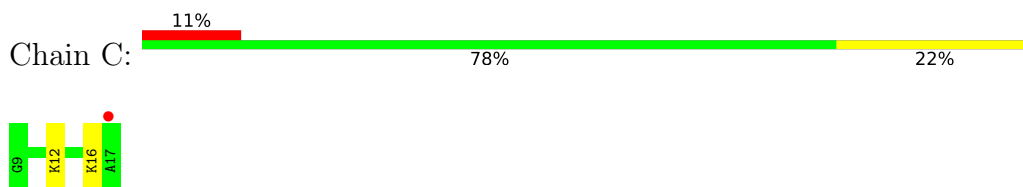
- Molecule 1: Histone acetyltransferase p300



- Molecule 2: Histone acetyltransferase p300



- Molecule 3: Histone H4



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	49.64Å 83.71Å 165.62Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	46.09 – 2.50 46.09 – 2.50	Depositor EDS
% Data completeness (in resolution range)	97.3 (46.09-2.50) 97.3 (46.09-2.50)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.82 (at 2.51Å)	Xtrriage
Refinement program	PHENIX (1.12_2829: ???)	Depositor
R, $R_{free}$	0.261 , 0.288 0.261 , 0.289	Depositor DCC
$R_{free}$ test set	1200 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	64.2	Xtrriage
Anisotropy	0.560	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.38 , 23.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	5512	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	60.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.65% 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 [i](#)

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

Bond lengths and bond angles in the following residue types are not validated in this section: ALY, 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	0.58	0/1364	0.80	2/1845 (0.1%)
2	B	0.59	0/1407	0.81	2/1902 (0.1%)
3	C	0.98	0/31	1.26	0/36
All	All	0.59	0/2802	0.81	4/3783 (0.1%)

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	0	1

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	1250	CYS	N-CA-C	-5.71	95.57	111.00
2	B	1104	ARG	NE-CZ-NH1	5.63	123.11	120.30
1	A	1137	ARG	NE-CZ-NH1	5.22	122.91	120.30
2	B	1252	ARG	NE-CZ-NH1	5.07	122.84	120.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	1169	GLY	Peptide

## 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	1328	1306	1306	46	0
2	B	1371	1351	1351	26	0
3	C	58	58	58	0	0
4	A	2	0	0	0	0
4	B	2	0	0	0	0
5	A	19	0	0	46	0
5	B	17	0	0	24	0
All	All	2797	2715	2715	72	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 13.

The worst 5 of 72 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)
2:B:1101:THR:HB	5:B:1402:HOH:O	1.29	1.27
1:A:1049:PHE:HB3	5:A:1407:HOH:O	1.25	1.25
2:B:1262:HIS:HB3	5:B:1407:HOH:O	1.34	1.23
1:A:1114:PRO:CB	5:A:1419:HOH:O	1.90	1.15
1:A:1114:PRO:CG	5:A:1419:HOH:O	1.93	1.14

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	159/161 (99%)	154 (97%)	5 (3%)	0	100	100
2	B	166/168 (99%)	162 (98%)	4 (2%)	0	100	100
3	C	5/9 (56%)	2 (40%)	3 (60%)	0	100	100
All	All	330/338 (98%)	318 (96%)	12 (4%)	0	100	100

There are no Ramachandran outliers to report.

### 5.3.2 Protein sidechains [i](#)

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	149/149 (100%)	146 (98%)	3 (2%)	55	79
2	B	151/151 (100%)	148 (98%)	3 (2%)	55	79
3	C	1/1 (100%)	1 (100%)	0	100	100
All	All	301/301 (100%)	295 (98%)	6 (2%)	55	79

5 of 6 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	1086	ILE
2	B	1133	ARG
2	B	1245	VAL
1	A	1245	VAL
1	A	1133	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

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
3	ALY	C	16	3	10,11,12	0.67	0	7,12,14	1.84	2 (28%)
3	ALY	C	12	3	10,11,12	0.59	0	7,12,14	1.75	1 (14%)

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
3	ALY	C	16	3	-	5/9/10/12	-
3	ALY	C	12	3	-	1/9/10/12	-

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
3	C	16	ALY	CE-NZ-CH	3.73	128.29	122.56
3	C	12	ALY	CE-NZ-CH	3.39	127.77	122.56
3	C	16	ALY	CH3-CH-NZ	2.59	120.67	116.09

There are no chirality outliers.

5 of 6 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	C	12	ALY	CD-CE-NZ-CH
3	C	16	ALY	OH-CH-NZ-CE
3	C	16	ALY	CH3-CH-NZ-CE
3	C	16	ALY	N-CA-CB-CG
3	C	16	ALY	CE-CD-CG-CB

There are no ring outliers.

No monomer is involved in short contacts.

## 5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 4 ligands modelled in this entry, 4 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	161/161 (100%)	0.92	18 (11%) 5 4	47, 57, 74, 93	0
2	B	168/168 (100%)	0.96	17 (10%) 7 6	47, 60, 75, 87	0
3	C	7/9 (77%)	0.79	1 (14%) 2 2	51, 53, 60, 62	0
All	All	336/338 (99%)	0.93	36 (10%) 6 5	47, 59, 75, 93	0

The worst 5 of 36 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	1091	ASP	3.5
2	B	1245	VAL	2.9
2	B	1055	ARG	2.8
2	B	1262	HIS	2.8
1	A	1260	LEU	2.7

### 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
3	ALY	C	16	12/13	0.85	0.31	53,55,58,62	0
3	ALY	C	12	12/13	0.93	0.19	44,46,49,50	0

### 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( $\text{\AA}^2$ )	Q<0.9
4	ZN	A	1302	1/1	0.93	0.29	82,82,82,82	0
4	ZN	B	1302	1/1	0.97	0.35	72,72,72,72	0
4	ZN	B	1301	1/1	0.99	0.33	76,76,76,76	0
4	ZN	A	1301	1/1	0.99	0.32	70,70,70,70	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.