



# Full wwPDB X-ray Structure Validation Report ⓘ

Oct 30, 2017 – 03:14 PM EDT

PDB ID : 3PLN  
Title : Crystal structure of *Klebsiella pneumoniae* UDP-glucose 6-dehydrogenase complexed with UDP-glucose  
Authors : Chen, Y.-Y.; Ko, T.-P.; Lin, C.-H.; Chen, W.-H.; Wang, A.H.-J.  
Deposited on : unknown  
Resolution : 1.50 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

<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 : 1.7.2 (RC1), CSD as538be (2017)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20030345  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
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) : rb-20030345

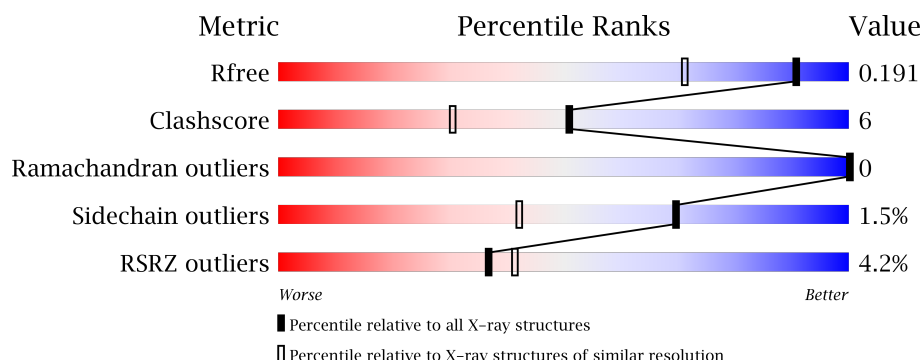
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.50 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))
$R_{free}$	100719	2279 (1.50-1.50)
Clashscore	112137	2503 (1.50-1.50)
Ramachandran outliers	110173	2445 (1.50-1.50)
Sidechain outliers	110143	2443 (1.50-1.50)
RSRZ outliers	101464	2305 (1.50-1.50)

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 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	424	<div> <div>4%</div> <div>80%</div> <div>8%</div> <div>11%</div> </div>

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 3679 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 UDP-glucose 6-dehydrogenase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	379	Total	C	N	O	S	0	0	0
			2994	1901	511	574	8			

There are 36 discrepancies between the modelled and reference sequences:

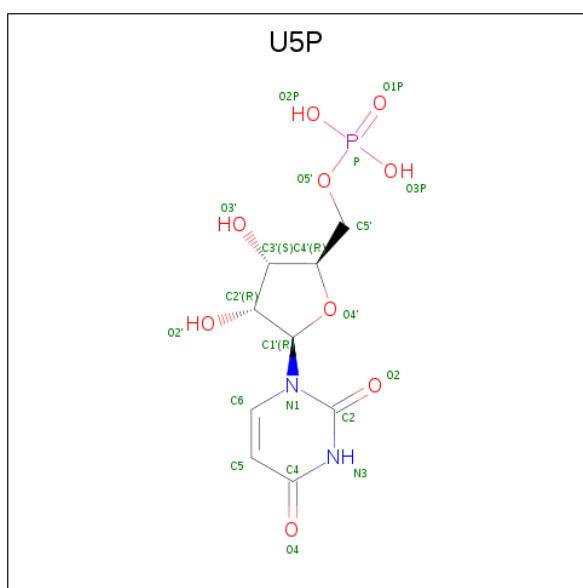
Chain	Residue	Modelled	Actual	Comment	Reference
A	964	MET	-	EXPRESSION TAG	UNP C4XAX5
A	965	GLY	-	EXPRESSION TAG	UNP C4XAX5
A	966	SER	-	EXPRESSION TAG	UNP C4XAX5
A	967	SER	-	EXPRESSION TAG	UNP C4XAX5
A	968	HIS	-	EXPRESSION TAG	UNP C4XAX5
A	969	HIS	-	EXPRESSION TAG	UNP C4XAX5
A	970	HIS	-	EXPRESSION TAG	UNP C4XAX5
A	971	HIS	-	EXPRESSION TAG	UNP C4XAX5
A	972	HIS	-	EXPRESSION TAG	UNP C4XAX5
A	973	HIS	-	EXPRESSION TAG	UNP C4XAX5
A	974	SER	-	EXPRESSION TAG	UNP C4XAX5
A	975	SER	-	EXPRESSION TAG	UNP C4XAX5
A	976	GLY	-	EXPRESSION TAG	UNP C4XAX5
A	977	LEU	-	EXPRESSION TAG	UNP C4XAX5
A	978	VAL	-	EXPRESSION TAG	UNP C4XAX5
A	979	PRO	-	EXPRESSION TAG	UNP C4XAX5
A	980	ARG	-	EXPRESSION TAG	UNP C4XAX5
A	981	GLY	-	EXPRESSION TAG	UNP C4XAX5
A	982	SER	-	EXPRESSION TAG	UNP C4XAX5
A	983	HIS	-	EXPRESSION TAG	UNP C4XAX5
A	984	MET	-	EXPRESSION TAG	UNP C4XAX5
A	985	ALA	-	EXPRESSION TAG	UNP C4XAX5
A	986	SER	-	EXPRESSION TAG	UNP C4XAX5
A	987	MET	-	EXPRESSION TAG	UNP C4XAX5
A	988	THR	-	EXPRESSION TAG	UNP C4XAX5
A	989	GLY	-	EXPRESSION TAG	UNP C4XAX5
A	990	GLY	-	EXPRESSION TAG	UNP C4XAX5

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	991	GLN	-	EXPRESSION TAG	UNP C4XAX5
A	992	GLN	-	EXPRESSION TAG	UNP C4XAX5
A	993	MET	-	EXPRESSION TAG	UNP C4XAX5
A	994	GLY	-	EXPRESSION TAG	UNP C4XAX5
A	995	ARG	-	EXPRESSION TAG	UNP C4XAX5
A	996	GLY	-	EXPRESSION TAG	UNP C4XAX5
A	997	SER	-	EXPRESSION TAG	UNP C4XAX5
A	998	GLU	-	EXPRESSION TAG	UNP C4XAX5
A	999	PHE	-	EXPRESSION TAG	UNP C4XAX5

- Molecule 2 is URIDINE-5'-MONOPHOSPHATE (three-letter code: U5P) (formula:  $C_9H_{13}N_2O_9P$ ).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	N	O	P	0	0
			21	9	2	9	1		

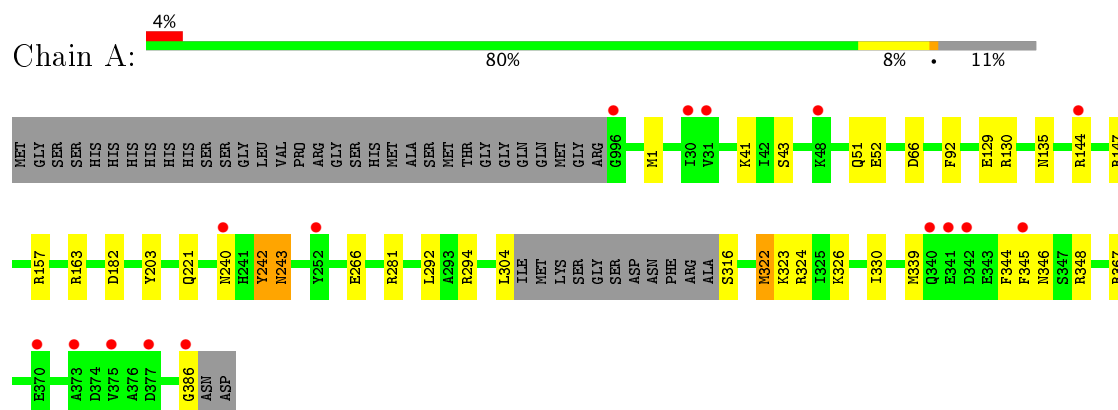
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	664	Total	O	0	0
			664	664		

### 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: UDP-glucose 6-dehydrogenase



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	106.94Å 63.22Å 79.15Å 90.00° 114.09° 90.00°	Depositor
Resolution (Å)	30.00 – 1.50 28.96 – 1.50	Depositor EDS
% Data completeness (in resolution range)	95.9 (30.00-1.50) 95.4 (28.96-1.50)	Depositor EDS
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	4.33 (at 1.50Å)	Xtriage
Refinement program	CNS	Depositor
R, $R_{free}$	0.164 , 0.190 0.165 , 0.191	Depositor DCC
$R_{free}$ test set	3776 reflections (5.10%)	DCC
Wilson B-factor (Å <sup>2</sup> )	14.4	Xtriage
Anisotropy	0.409	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.37 , 49.6	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.015 for $1/2^*h+3/2^*k, 1/2^*h-1/2^*k, -1/2^*h-1/2^*k-l$ 0.010 for $1/2^*h-3/2^*k, -1/2^*h-1/2^*k, -1/2^*h+1/2^*k-l$	Xtriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	3679	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	21.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.04% 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: U5P

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	0.89	3/3045 (0.1%)	0.95	7/4124 (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	0	1

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	129	GLU	CB-CG	-6.66	1.39	1.52
1	A	1	MET	SD-CE	-5.47	1.47	1.77
1	A	242	TYR	CD1-CE1	-5.40	1.31	1.39

All (7) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	66	ASP	CB-CG-OD2	9.08	126.47	118.30
1	A	242	TYR	CB-CG-CD1	-7.32	116.61	121.00
1	A	324	ARG	NE-CZ-NH1	6.41	123.50	120.30
1	A	348	ARG	NE-CZ-NH2	-6.13	117.23	120.30
1	A	367	ARG	NE-CZ-NH2	-5.98	117.31	120.30
1	A	130	ARG	NE-CZ-NH1	5.88	123.24	120.30
1	A	242	TYR	CG-CD2-CE2	-5.23	117.12	121.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	203	TYR	Sidechain

## 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	2994	0	3008	39	0
2	A	21	0	11	1	0
3	A	664	0	0	24	4
All	All	3679	0	3019	39	4

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 6.

All (39) 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:242:TYR:HE2	3:A:518:HOH:O	1.15	1.26
1:A:41:LYS:HE2	3:A:632:HOH:O	1.52	1.07
1:A:135:ASN:HB2	3:A:454:HOH:O	1.59	1.03
1:A:135:ASN:HB2	3:A:453:HOH:O	1.62	0.99
1:A:182:ASP:HB2	3:A:659:HOH:O	1.63	0.96
1:A:221:GLN:NE2	3:A:519:HOH:O	2.03	0.90
1:A:316:SER:HA	3:A:661:HOH:O	1.71	0.89
1:A:240:ASN:ND2	3:A:660:HOH:O	2.19	0.74
1:A:135:ASN:CB	3:A:454:HOH:O	2.27	0.71
1:A:240:ASN:ND2	1:A:240:ASN:O	2.24	0.70
1:A:41:LYS:CE	3:A:632:HOH:O	2.23	0.70
1:A:316:SER:HB3	3:A:1337:HOH:O	1.95	0.67
1:A:144:ARG:HB2	1:A:147:ARG:HG3	1.77	0.66
1:A:316:SER:CA	3:A:661:HOH:O	2.35	0.65
1:A:292:LEU:CD2	1:A:330:ILE:HD12	2.27	0.65
1:A:43:SER:H	1:A:51:GLN:HE22	1.43	0.65
1:A:281:ARG:HG3	3:A:564:HOH:O	1.98	0.63
1:A:163:ARG:CZ	3:A:648:HOH:O	2.47	0.62
1:A:163:ARG:NH2	3:A:648:HOH:O	2.32	0.61
1:A:43:SER:H	1:A:51:GLN:NE2	1.98	0.61

*Continued on next page...*



*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:243:ASN:HD22	1:A:243:ASN:H	1.48	0.60
1:A:339:MET:HE1	1:A:344:PHE:HB2	1.83	0.60
1:A:304:LEU:C	3:A:1331:HOH:O	2.39	0.60
1:A:52:GLU:HG3	3:A:1075:HOH:O	2.02	0.59
1:A:281:ARG:CG	3:A:564:HOH:O	2.52	0.58
1:A:326:LYS:NZ	1:A:345:PHE:O	2.38	0.56
1:A:144:ARG:HH22	2:A:901:U5P:P	2.30	0.55
1:A:304:LEU:HD21	1:A:322:MET:HE1	1.91	0.53
1:A:243:ASN:HD22	1:A:243:ASN:N	2.06	0.52
1:A:386:GLY:HA3	3:A:535:HOH:O	2.11	0.51
1:A:242:TYR:CE2	3:A:518:HOH:O	2.07	0.51
1:A:144:ARG:HB2	1:A:147:ARG:CG	2.41	0.51
1:A:316:SER:N	3:A:661:HOH:O	2.44	0.50
1:A:316:SER:N	3:A:1258:HOH:O	2.47	0.46
1:A:294:ARG:NH1	3:A:523:HOH:O	2.46	0.45
1:A:339:MET:HE3	1:A:344:PHE:HD1	1.84	0.42
1:A:345:PHE:O	1:A:346:ASN:HB2	2.19	0.42
1:A:240:ASN:CG	1:A:240:ASN:O	2.58	0.41
1:A:323:LYS:CE	3:A:468:HOH:O	2.69	0.40

All (4) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:A:1051:HOH:O	3:A:1080:HOH:O[4_647]	1.94	0.26
3:A:1271:HOH:O	3:A:1273:HOH:O[2_656]	2.07	0.13
3:A:1161:HOH:O	3:A:1236:HOH:O[2_656]	2.17	0.03
3:A:475:HOH:O	3:A:649:HOH:O[3_545]	2.19	0.01

## 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	375/424 (88%)	368 (98%)	7 (2%)	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	326/362 (90%)	321 (98%)	5 (2%)	70	42

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	92	PHE
1	A	157	ARG
1	A	243	ASN
1	A	266	GLU
1	A	322	MET

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	51	GLN
1	A	240	ASN
1	A	243	ASN
1	A	264	ASN
1	A	319	GLN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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

## 5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 5.6 Ligand geometry ⓘ

1 ligand 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$
2	U5P	A	901	-	18,22,22	1.64	4 (22%)	22,33,33	1.71	2 (9%)

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	U5P	A	901	-	-	0/6/26/26	0/2/2/2

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	901	U5P	O4'-C1'	2.11	1.44	1.41
2	A	901	U5P	P-O1P	3.15	1.61	1.50
2	A	901	U5P	C6-N1	3.31	1.40	1.35
2	A	901	U5P	C4-N3	3.96	1.40	1.33

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
2	A	901	U5P	O3P-P-O2P	2.85	119.11	107.61
2	A	901	U5P	C4-N3-C2	7.04	120.17	114.13

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	901	U5P	1	0

## 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	379/424 (89%)	0.19	16 (4%) 37 41	9, 15, 32, 42	3 (0%)

All (16) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	345	PHE	6.1
1	A	252	TYR	5.8
1	A	30	ILE	4.3
1	A	370	GLU	3.8
1	A	377	ASP	3.7
1	A	996	GLY	3.4
1	A	373	ALA	3.4
1	A	386	GLY	3.1
1	A	375	VAL	2.7
1	A	340	GLN	2.5
1	A	342	ASP	2.2
1	A	31	VAL	2.2
1	A	240	ASN	2.2
1	A	144	ARG	2.1
1	A	341	GLU	2.1
1	A	48	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 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( $\text{\AA}^2$ )	Q<0.9
2	U5P	A	901	21/21	0.97	0.08	-0.38	13,16,21,24	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.