



# wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 14, 2018 – 12:11 am GMT

PDB ID : 1KB0  
Title : Crystal Structure of Quinohemoprotein Alcohol Dehydrogenase from Comamonas testosteroni  
Authors : Rozeboom, H.J.; Oubrie, A.  
Deposited on : 2001-11-05  
Resolution : 1.44 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 following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.7.3 (157068), CSD as539be (2018)  
Xtriage (Phenix) : 1.13  
EDS : trunk31020  
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)  
Refmac : 5.8.0158  
CCP4 : 7.0 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk31020

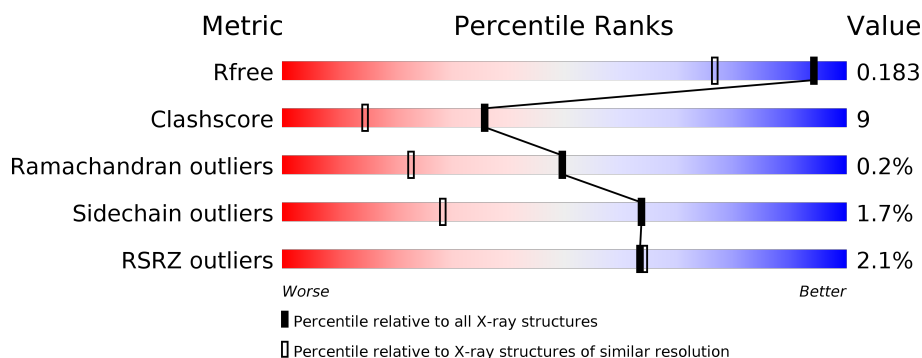
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.44 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))
$R_{free}$	111664	1592 (1.46-1.42)
Clashscore	122126	1643 (1.46-1.42)
Ramachandran outliers	120053	1618 (1.46-1.42)
Sidechain outliers	120020	1618 (1.46-1.42)
RSRZ outliers	108989	1566 (1.46-1.42)

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	677	<div> <div>2%</div> <div> <div></div> <div>90%</div> <div>8% ..</div> </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
6	GOL	A	1823	-	-	X	-

## 2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 6286 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 quinohemoprotein alcohol dehydrogenase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	670	Total	C	N	O	S	0	6	0
			5172	3290	903	960	19			

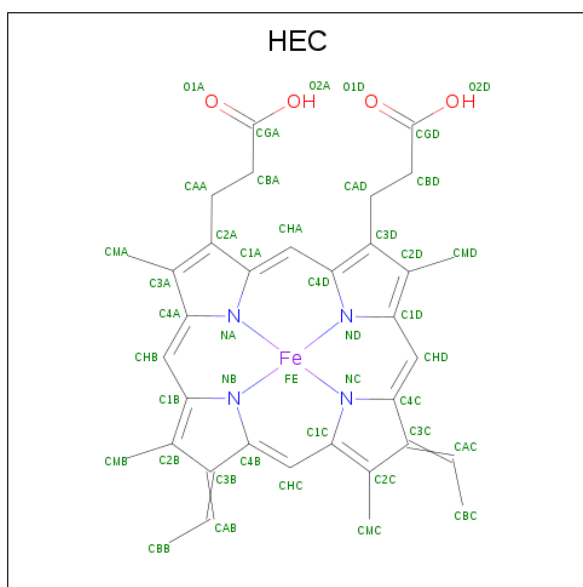
There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	181	ASN	LYS	see remark 999	UNP Q46444
A	182	GLY	ARG	see remark 999	UNP Q46444
A	566	GLY	ALA	see remark 999	UNP Q46444
A	512	TRO	TRP	MODIFIED RESIDUE	UNP Q46444

- Molecule 2 is CALCIUM ION (three-letter code: CA) (formula: Ca).

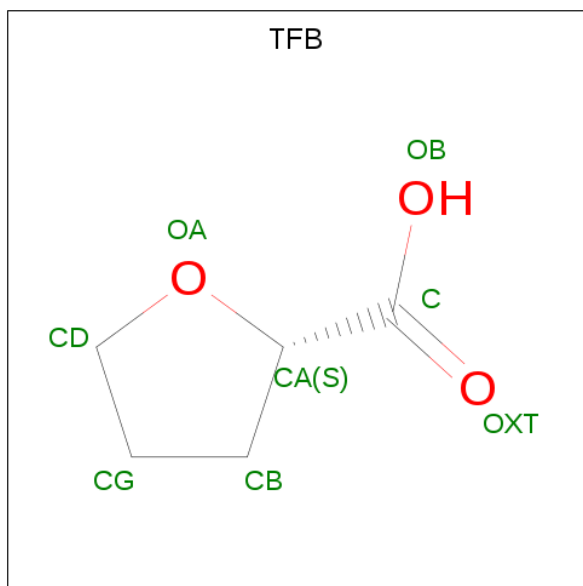
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	1	Total	Ca	0	0
			1	1		

- Molecule 3 is HEME C (three-letter code: HEC) (formula: C<sub>34</sub>H<sub>34</sub>FeN<sub>4</sub>O<sub>4</sub>).



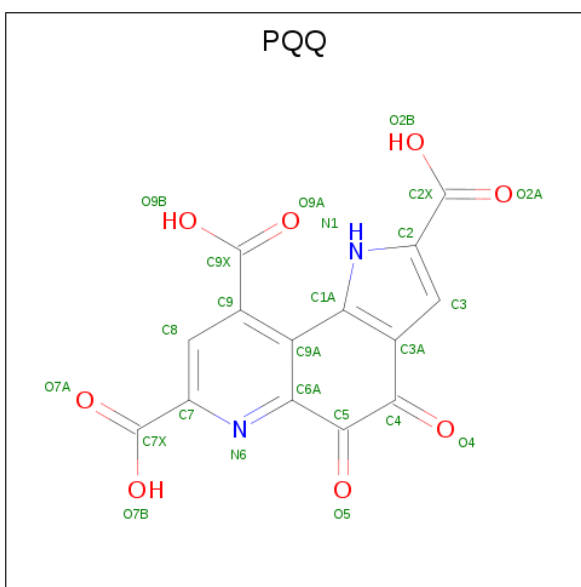
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		

- Molecule 4 is TETRAHYDROFURAN-2-CARBOXYLIC ACID (three-letter code: TFB) (formula:  $C_5H_8O_3$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			8	5	3		

- Molecule 5 is PYRROLOQUINOLINE QUINONE (three-letter code: PQQ) (formula:  $C_{14}H_6N_2O_8$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	A	1	Total	C	N	O	0	0
			24	14	2	8		

- Molecule 6 is GLYCEROL (three-letter code: GOL) (formula:  $C_3H_8O_3$ ).



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

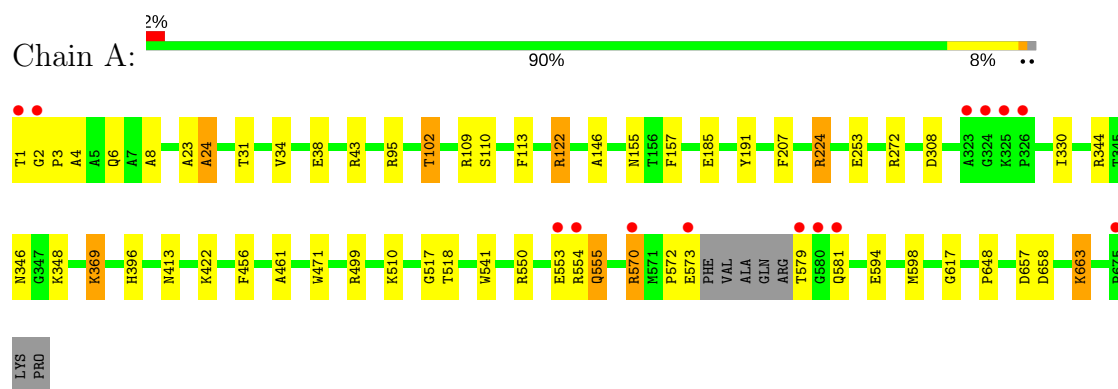
- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	1016	Total 1020	O 1020	0	5

### 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: quinoheprotein alcohol dehydrogenase



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	97.92Å 74.33Å 92.33Å 90.00° 105.67° 90.00°	Depositor
Resolution (Å)	29.63 – 1.44 29.63 – 1.44	Depositor EDS
% Data completeness (in resolution range)	96.3 (29.63-1.44) 96.5 (29.63-1.44)	Depositor EDS
$R_{merge}$	0.06	Depositor
$R_{sym}$	0.06	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.60 (at 1.44Å)	Xtriage
Refinement program	CNS 1.0	Depositor
R, $R_{free}$	0.160 , 0.188 0.155 , 0.183	Depositor DCC
$R_{free}$ test set	5671 reflections (5.10%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	16.4	Xtriage
Anisotropy	0.025	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 56.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	6286	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 4.57% 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: GOL, TRO, CA, TFB, PQQ, HEC

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	0.82	0/5307	0.96	4/7227 (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( $^{\circ}$ )	Ideal( $^{\circ}$ )
1	A	272	ARG	NE-CZ-NH1	-6.35	117.12	120.30
1	A	272	ARG	NE-CZ-NH2	5.95	123.27	120.30
1	A	157	PHE	CB-CG-CD1	-5.41	117.02	120.80
1	A	122	ARG	NE-CZ-NH2	-5.16	117.72	120.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	122	ARG	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	5172	0	4979	86	0
2	A	1	0	0	0	0
3	A	43	0	30	0	0
4	A	8	0	7	3	0
5	A	24	0	3	0	0
6	A	18	0	23	14	0
7	A	1020	0	0	40	0
All	All	6286	0	5042	90	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 9.

The worst 5 of 90 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:594:GLU:HG3	7:A:2679:HOH:O	1.70	0.92
1:A:253:GLU:HG3	7:A:2235:HOH:O	1.72	0.89
1:A:663:LYS:NZ	1:A:663:LYS:HB3	1.93	0.83
1:A:550:ARG:HH11	1:A:579:THR:HG23	1.45	0.80
1:A:541:TRP:HD1	6:A:1823:GOL:H2	1.46	0.80

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	671/677 (99%)	644 (96%)	25 (4%)	2 (0%)	43 19

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	24[A]	ALA
1	A	24[B]	ALA

### 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	521/523 (100%)	512 (98%)	9 (2%)	63	28

5 of 9 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	369	LYS
1	A	663	LYS
1	A	555	GLN
1	A	185	GLU
1	A	518	THR

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	6	GLN
1	A	396	HIS

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

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	TRO	A	512	1	14,16,17	1.32	2 (14%)	13,22,24	1.12	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	TRO	A	512	1	-	0/2/6/8	0/2/2/2

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	512	TRO	CA-C	2.04	1.52	1.50
1	A	512	TRO	CB-CG	3.12	1.55	1.51

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.

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 7 ligands modelled in this entry, 1 is monoatomic - leaving 6 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	PQQ	A	1800	2	17,26,26	3.05	9 (52%)	16,40,40	1.89	5 (31%)
4	TFB	A	1810	2	5,8,8	1.34	2 (40%)	5,10,10	0.84	0
6	GOL	A	1820	-	5,5,5	1.49	1 (20%)	5,5,5	1.12	1 (20%)
6	GOL	A	1823	-	5,5,5	0.94	0	5,5,5	0.56	0
6	GOL	A	1824	-	5,5,5	0.72	0	5,5,5	0.44	0
3	HEC	A	802	1	26,50,50	1.09	2 (7%)	18,82,82	1.63	3 (16%)

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	PQQ	A	1800	2	-	0/0/28/28	0/3/3/3
4	TFB	A	1810	2	-	0/0/11/11	0/1/1/1
6	GOL	A	1820	-	-	0/4/4/4	0/0/0/0
6	GOL	A	1823	-	-	0/4/4/4	0/0/0/0
6	GOL	A	1824	-	-	0/4/4/4	0/0/0/0
3	HEC	A	802	1	-	0/6/54/54	0/0/8/8

The worst 5 of 14 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	A	1800	PQQ	C3A-C4	-5.18	1.38	1.48
5	A	1800	PQQ	C5-C4	-4.08	1.41	1.53
5	A	1800	PQQ	C9A-C1A	-3.41	1.42	1.48
3	A	802	HEC	C3C-C4C	-2.50	1.38	1.43
6	A	1820	GOL	O2-C2	-2.45	1.36	1.43

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	802	HEC	CBA-CAA-C2A	-3.66	105.48	112.47
5	A	1800	PQQ	O5-C5-C6A	-3.47	118.12	121.83
5	A	1800	PQQ	C9A-C1A-N1	2.20	129.18	124.45
5	A	1800	PQQ	C9-C8-C7	2.25	122.84	120.36
6	A	1820	GOL	O2-C2-C1	2.39	119.93	109.00

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

3 monomers are involved in 17 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	1810	TFB	3	0
6	A	1820	GOL	2	0
6	A	1823	GOL	12	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	669/677 (98%)	-0.18	14 (2%) 63 64	11, 17, 29, 50	0

The worst 5 of 14 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	579	THR	6.4
1	A	1	THR	6.1
1	A	324	GLY	3.4
1	A	580	GLY	3.0
1	A	675	PRO	3.0

### 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	TRO	A	512	15/16	0.95	0.07	14,17,20,22	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. 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
6	GOL	A	1824	6/6	0.67	0.14	57,58,58,59	0
4	TFB	A	1810	8/8	0.81	0.16	27,31,32,33	0
6	GOL	A	1823	6/6	0.84	0.26	36,38,39,39	0
6	GOL	A	1820	6/6	0.89	0.15	17,24,29,32	0
3	HEC	A	802	43/43	0.98	0.08	11,13,24,29	0
5	PQQ	A	1800	24/24	0.98	0.07	10,11,13,14	0
2	CA	A	801	1/1	1.00	0.04	21,21,21,21	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.