



# Full wwPDB X-ray Structure Validation Report ⓘ

May 16, 2020 – 04:33 pm BST

PDB ID : 6UGL  
Title : VqmA bound to DPO  
Authors : Paczkowski, J.E.; Huang, X.  
Deposited on : 2019-09-26  
Resolution : 2.03 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

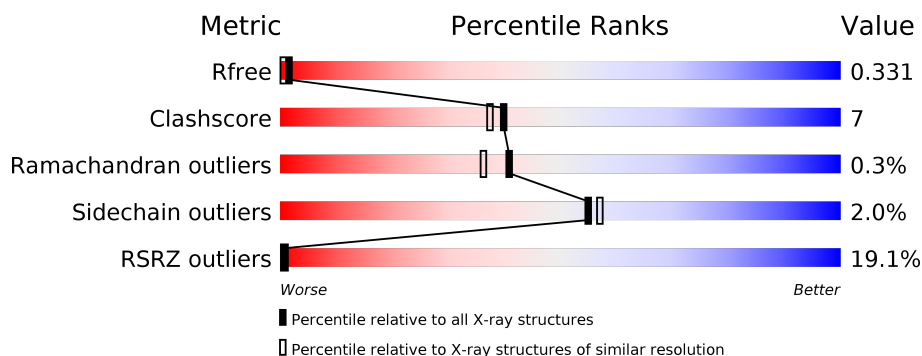
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.03 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	10434 (2.04-2.00)
Clashscore	141614	11643 (2.04-2.00)
Ramachandran outliers	138981	11493 (2.04-2.00)
Sidechain outliers	138945	11492 (2.04-2.00)
RSRZ outliers	127900	10220 (2.04-2.00)

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	246	
1	B	246	

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
2	QOS	A	301	-	X	-	-
2	QOS	B	301	-	X	-	-

## 2 Entry composition [i](#)

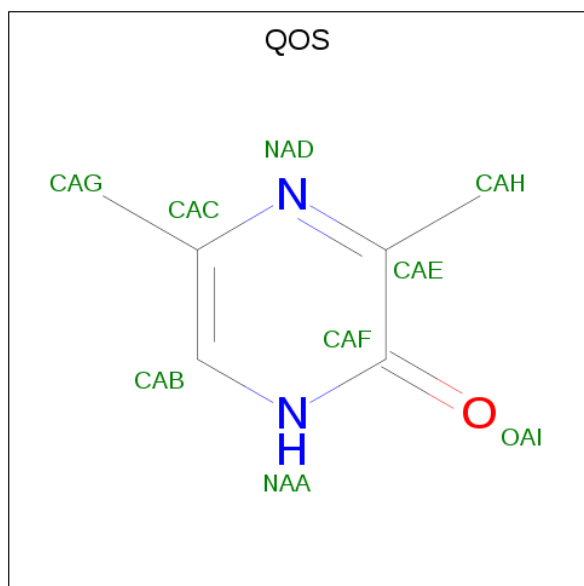
There are 2 unique types of molecules in this entry. The entry contains 3225 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 Helix-turn-helix transcriptional regulator.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	200	Total	C	N	O	S	0	0	0
			1590	1024	273	288	5			
1	B	203	Total	C	N	O	S	0	0	0
			1617	1043	276	293	5			

- Molecule 2 is 3,5-dimethylpyrazin-2(1H)-one (three-letter code: QOS) (formula: C<sub>6</sub>H<sub>8</sub>N<sub>2</sub>O).

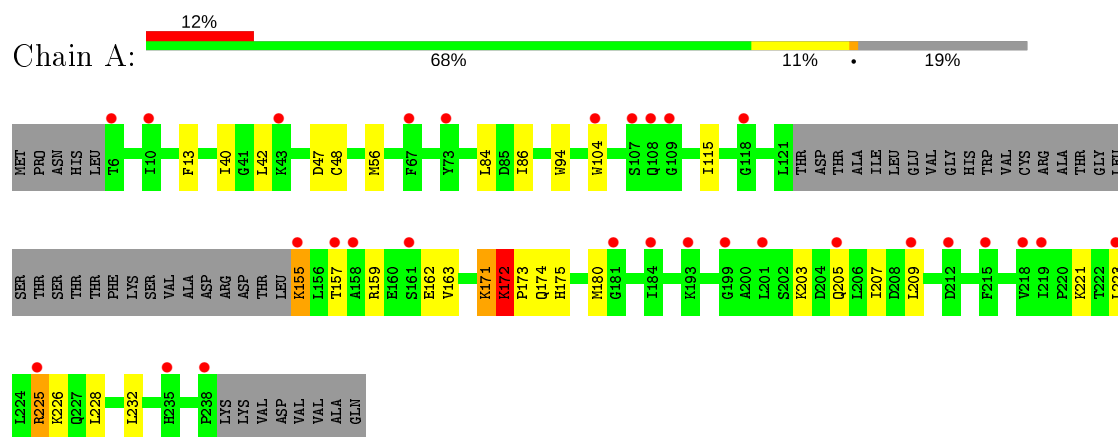


Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	N	O	9	0
			9	6	2	1		
2	B	1	Total	C	N	O	9	0
			9	6	2	1		

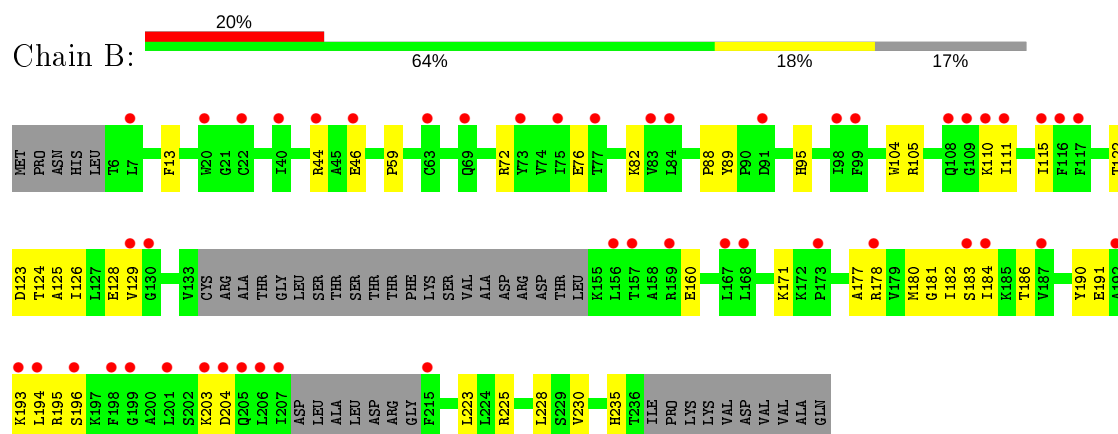
### 3 Residue-property plots [i](#)

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: Helix-turn-helix transcriptional regulator



- Molecule 1: Helix-turn-helix transcriptional regulator



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	50.46Å 84.68Å 116.04Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	28.86 – 2.03 28.86 – 2.03	Depositor EDS
% Data completeness (in resolution range)	98.7 (28.86-2.03) 98.7 (28.86-2.03)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.02 (at 2.03Å)	Xtriage
Refinement program	PHENIX 1.16_3549, PHENIX 1.16_3549	Depositor
R, $R_{free}$	0.277 , 0.331 0.277 , 0.331	Depositor DCC
$R_{free}$ test set	1990 reflections (6.12%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	34.5	Xtriage
Anisotropy	0.981	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.38 , 47.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.51$ , $\langle L^2 \rangle = 0.35$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	3225	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	44.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.51% 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: QOS

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	0.41	0/1627	0.69	2/2198 (0.1%)
1	B	0.39	0/1655	0.57	0/2237
All	All	0.40	0/3282	0.63	2/4435 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	225	ARG	NE-CZ-NH2	15.23	127.91	120.30
1	A	225	ARG	NE-CZ-NH1	-7.81	116.39	120.30

There are no chirality outliers.

There are no planarity outliers.

### 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	1590	0	1599	21	3
1	B	1617	0	1618	27	3
2	A	9	0	0	0	0
2	B	9	0	0	0	0
All	All	3225	0	3217	47	3

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 7.

All (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:171:LYS:O	1:A:175:HIS:ND1	2.17	0.77
1:B:184:ILE:H	1:B:184:ILE:HD12	1.62	0.65
1:A:94:TRP:HE1	1:A:225:ARG:NH2	1.98	0.62
1:A:94:TRP:NE1	1:A:225:ARG:NH2	2.48	0.61
1:B:223:LEU:HG	1:B:228:LEU:HD21	1.83	0.60
1:B:82:LYS:HE2	1:B:129:VAL:HG12	1.85	0.58
1:A:42:LEU:HD13	1:A:47:ASP:HB2	1.86	0.57
1:B:183:SER:HB2	1:B:186:THR:OG1	2.04	0.57
1:A:171:LYS:HE2	1:A:171:LYS:HA	1.89	0.55
1:A:84:LEU:HB2	1:A:232:LEU:HD11	1.88	0.55
1:A:172:LYS:HE3	1:A:173:PRO:HD2	1.89	0.54
1:A:86:ILE:HD12	1:A:228:LEU:HD11	1.89	0.53
1:B:124:THR:O	1:B:128:GLU:HG3	2.08	0.53
1:B:180:MET:O	1:B:182:ILE:HG12	2.08	0.53
1:A:205:GLN:O	1:A:209:LEU:HD13	2.10	0.52
1:B:171:LYS:HE3	1:B:230:VAL:HG22	1.90	0.52
1:B:72:ARG:O	1:B:76:GLU:HG3	2.10	0.51
1:A:172:LYS:HG3	1:A:173:PRO:HD2	1.93	0.50
1:B:191:GLU:OE1	1:B:203:LYS:HD3	2.11	0.50
1:B:160:GLU:HB2	1:B:194:LEU:HD11	1.95	0.49
1:B:89:TYR:CE2	1:B:95:HIS:HB2	2.49	0.48
1:B:181:GLY:O	1:B:182:ILE:HD13	2.14	0.48
1:B:177:ALA:HB1	1:B:182:ILE:O	2.14	0.47
1:A:157:THR:HG22	1:A:159:ARG:N	2.29	0.47
1:A:163:VAL:HG22	1:A:180:MET:HE1	1.96	0.47
1:A:155:LYS:HB2	1:A:155:LYS:HE2	1.81	0.46
1:B:190:TYR:HA	1:B:193:LYS:HZ3	1.80	0.46
1:B:110:LYS:HA	1:B:110:LYS:HD3	1.70	0.45
1:B:195:ARG:HE	1:B:195:ARG:HB2	1.51	0.45
1:A:104:TRP:HB2	1:A:115:ILE:HG13	1.99	0.45
1:A:162:GLU:HG2	1:A:223:LEU:HD13	1.99	0.45
1:A:203:LYS:O	1:A:207:ILE:HG13	2.16	0.45
1:A:42:LEU:HD12	1:A:48:CYS:SG	2.58	0.44
1:B:105:ARG:NH2	1:B:111:ILE:HD11	2.33	0.44
1:A:40:ILE:HG21	1:A:56:MET:HE2	1.99	0.43
1:A:225:ARG:HE	1:A:225:ARG:HB3	1.22	0.43
1:B:123:ASP:O	1:B:126:ILE:HG13	2.18	0.43
1:A:13:PHE:HD1	1:B:13:PHE:CD1	2.37	0.43

*Continued on next page...*



Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:59:PRO:HB2	1:B:88:PRO:HG2	2.00	0.42
1:B:104:TRP:HB2	1:B:115:ILE:HG13	2.01	0.42
1:B:44:ARG:HG3	1:B:46:GLU:OE2	2.20	0.41
1:B:46:GLU:H	1:B:46:GLU:CD	2.22	0.41
1:B:44:ARG:HG2	1:B:46:GLU:HG2	2.03	0.41
1:B:122:THR:H	1:B:125:ALA:HB3	1.86	0.41
1:B:178:ARG:HD3	1:B:178:ARG:HH11	1.60	0.41
1:B:171:LYS:CE	1:B:230:VAL:HG22	2.52	0.40
1:A:174:GLN:HG2	1:A:175:HIS:H	1.86	0.40

All (3) 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)
1:A:221:LYS:NZ	1:B:204:ASP:OD2[4_466]	1.85	0.35
1:A:225:ARG:CD	1:B:178:ARG:NH2[1_455]	2.16	0.04
1:A:225:ARG:NE	1:B:178:ARG:NH1[1_455]	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	196/246 (80%)	194 (99%)	1 (0%)	1 (0%)	29	22
1	B	197/246 (80%)	192 (98%)	5 (2%)	0	100	100
All	All	393/492 (80%)	386 (98%)	6 (2%)	1 (0%)	41	36

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	172	LYS

### 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	171/211 (81%)	167 (98%)	4 (2%)	50	51
1	B	174/211 (82%)	171 (98%)	3 (2%)	60	63
All	All	345/422 (82%)	338 (98%)	7 (2%)	55	57

All (7) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	155	LYS
1	A	171	LYS
1	A	172	LYS
1	A	226	LYS
1	B	196	SER
1	B	225	ARG
1	B	235	HIS

Some 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](#)

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

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry

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	QOS	B	301	-	8,9,9	7.00	7 (87%)	8,12,12	1.63	2 (25%)
2	QOS	A	301	-	8,9,9	6.86	7 (87%)	8,12,12	1.50	2 (25%)

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	QOS	B	301	-	-	-	0/1/1/1
2	QOS	A	301	-	-	-	0/1/1/1

All (14) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	301	QOS	CAE-NAD	15.24	1.45	1.32
2	A	301	QOS	CAE-NAD	14.90	1.45	1.32
2	B	301	QOS	CAF-NAA	7.22	1.45	1.33
2	A	301	QOS	CAF-NAA	7.20	1.45	1.33
2	B	301	QOS	CAB-CAC	5.60	1.53	1.39
2	A	301	QOS	CAB-CAC	5.59	1.53	1.39
2	B	301	QOS	CAC-NAD	5.56	1.45	1.34
2	A	301	QOS	CAC-NAD	5.44	1.45	1.34
2	B	301	QOS	CAB-NAA	5.31	1.45	1.34
2	A	301	QOS	CAB-NAA	5.30	1.45	1.34
2	B	301	QOS	CAH-CAE	3.23	1.52	1.50
2	A	301	QOS	CAH-CAE	2.69	1.52	1.50
2	A	301	QOS	OAI-CAF	-2.31	1.18	1.24
2	B	301	QOS	OAI-CAF	-2.29	1.18	1.24

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
2	B	301	QOS	CAH-CAE-NAD	2.97	120.16	116.01
2	A	301	QOS	CAH-CAE-NAD	2.60	119.65	116.01
2	A	301	QOS	CAC-CAB-NAA	-2.38	119.95	123.21
2	B	301	QOS	CAC-CAB-NAA	-2.34	120.01	123.21

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

### 6.1 Protein, DNA and RNA chains

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	200/246 (81%)	1.18	29 (14%)	2 2	31, 42, 57, 68	0
1	B	203/246 (82%)	1.49	48 (23%)	0 0	30, 43, 66, 75	0
All	All	403/492 (81%)	1.33	77 (19%)	1 0	30, 42, 63, 75	0

All (77) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	194	LEU	6.0
1	B	215	PHE	5.1
1	B	156	LEU	4.8
1	B	207	ILE	4.6
1	B	205	GLN	4.2
1	A	157	THR	4.2
1	A	238	PRO	4.2
1	B	157	THR	4.1
1	B	192	ALA	4.1
1	A	43	LYS	3.9
1	B	73	TYR	3.8
1	B	109	GLY	3.7
1	B	108	GLN	3.6
1	B	183	SER	3.6
1	B	130	GLY	3.5
1	A	155	LYS	3.5
1	B	98	ILE	3.5
1	B	184	ILE	3.5
1	B	69	GLN	3.5
1	A	215	PHE	3.4
1	B	199	GLY	3.4
1	A	109	GLY	3.3
1	A	184	ILE	3.3
1	B	159	ARG	3.2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	204	ASP	3.1
1	A	223	LEU	3.1
1	B	168	LEU	3.1
1	B	196	SER	3.1
1	B	198	PHE	3.1
1	A	225	ARG	2.9
1	B	77	THR	2.9
1	B	178	ARG	2.9
1	B	7	LEU	2.9
1	A	107	SER	2.8
1	B	201	LEU	2.8
1	B	111	ILE	2.8
1	A	209	LEU	2.8
1	B	203	LYS	2.7
1	A	158	ALA	2.7
1	B	129	VAL	2.6
1	A	118	GLY	2.6
1	A	199	GLY	2.6
1	B	116	PHE	2.5
1	B	63	CYS	2.4
1	A	10	ILE	2.4
1	A	73	TYR	2.4
1	A	193	LYS	2.3
1	B	187	VAL	2.3
1	B	99	PHE	2.3
1	B	173	PRO	2.3
1	B	75	ILE	2.3
1	A	108	GLN	2.3
1	A	201	LEU	2.3
1	B	206	LEU	2.3
1	B	44	ARG	2.3
1	B	167	LEU	2.3
1	B	83	VAL	2.3
1	A	104	TRP	2.3
1	A	212	ASP	2.3
1	B	117	PHE	2.2
1	A	218	VAL	2.2
1	A	161	SER	2.2
1	B	84	LEU	2.2
1	A	67	PHE	2.2
1	A	235	HIS	2.2
1	B	91	ASP	2.2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	110	LYS	2.2
1	B	40	ILE	2.2
1	B	193	LYS	2.1
1	B	20	TRP	2.1
1	A	6	THR	2.1
1	B	46	GLU	2.1
1	B	115	ILE	2.1
1	A	205	GLN	2.0
1	B	22	CYS	2.0
1	A	181	GLY	2.0
1	A	219	ILE	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. 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	QOS	B	301	9/9	-	-	33,34,35,35	9
2	QOS	A	301	9/9	-	-	32,33,33,33	9

## 6.5 Other polymers [i](#)

There are no such residues in this entry.