



# Full wwPDB X-ray Structure Validation Report ⓘ

Jun 5, 2023 – 10:08 pm BST

PDB ID : 8OFD  
Title : Crystal structure of beta-conglutin from Lupinus albus refined to 2.81 Å  
Authors : Dolot, R.M.; O'Sullivan, C.K.; Jauset-Rubio, M.  
Deposited on : 2023-03-15  
Resolution : 2.81 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 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 : 2.33  
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) : 2.33

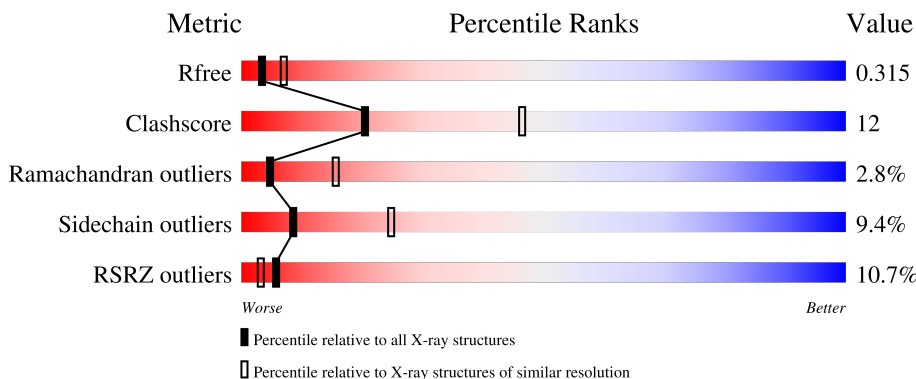
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.81 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	3617 (2.84-2.80)
Clashscore	141614	4060 (2.84-2.80)
Ramachandran outliers	138981	3978 (2.84-2.80)
Sidechain outliers	138945	3980 (2.84-2.80)
RSRZ outliers	127900	3552 (2.84-2.80)

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	423	

## 2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 3387 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 Conglutin beta 1.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
1	A	401	3299	2059	595	645	0	2	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	176	ILE	LEU	variant	UNP Q53HY0

- Molecule 2 is ACETATE ION (three-letter code: ACT) (formula: C<sub>2</sub>H<sub>3</sub>O<sub>2</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
2	A	1	4	2	2	0	0

- Molecule 3 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total Na 1 1	0	0

- Molecule 4 is POTASSIUM ION (three-letter code: K) (formula: K).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total K 1 1	0	0

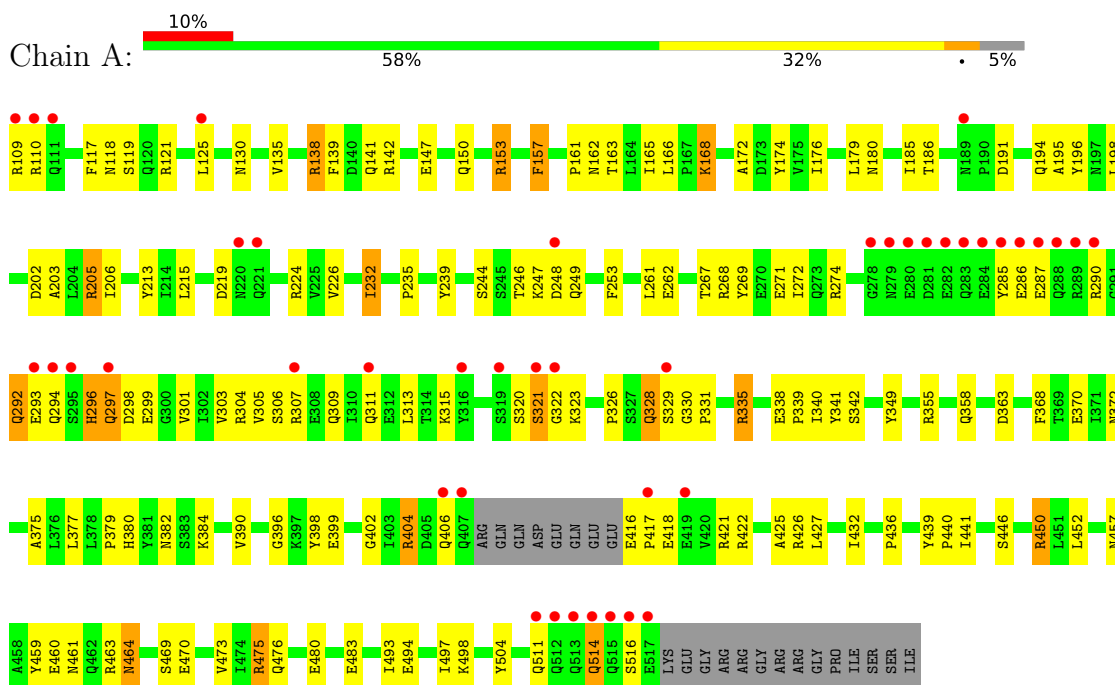
- Molecule 5 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	82	Total O 82 82	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: Conglutin beta 1



## 4 Data and refinement statistics i

Property	Value	Source
Space group	P 3 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	166.63Å 166.63Å 40.10Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	21.77 – 2.81 21.77 – 2.81	Depositor EDS
% Data completeness (in resolution range)	99.8 (21.77-2.81) 99.4 (21.77-2.81)	Depositor EDS
$R_{merge}$	0.17	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.39 (at 2.79Å)	Xtrriage
Refinement program	REFMAC 5.8.0411	Depositor
R, $R_{free}$	0.204 , 0.277 0.231 , 0.315	Depositor DCC
$R_{free}$ test set	775 reflections (4.93%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	58.6	Xtrriage
Anisotropy	0.011	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.40 , 62.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	0.000 for -h,-k,l	Xtrriage
$F_o, F_c$ correlation	0.90	EDS
Total number of atoms	3387	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	64.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 9.02% 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: ACT, K, NA

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	0.35	0/3362	0.73	1/4544 (0.0%)

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	8

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
1	A	239	TYR	CB-CA-C	-5.14	100.12	110.40

There are no chirality outliers.

All (8) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	121	ARG	Sidechain
1	A	153	ARG	Sidechain
1	A	205	ARG	Sidechain
1	A	304	ARG	Sidechain
1	A	355	ARG	Sidechain
1	A	426	ARG	Sidechain
1	A	450	ARG	Sidechain
1	A	475	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	3299	0	3200	81	0
2	A	4	0	3	0	0
3	A	1	0	0	0	0
4	A	1	0	0	0	0
5	A	82	0	0	2	0
All	All	3387	0	3203	81	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 12.

All (81) 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:157:PHE:CE1	1:A:165:ILE:HD11	2.13	0.83
1:A:297:GLN:HB2	1:A:303:VAL:HB	1.66	0.78
1:A:380:HIS:O	1:A:440:PRO:HA	1.92	0.70
1:A:176:ILE:O	1:A:203:ALA:HA	1.98	0.64
1:A:384:LYS:NZ	1:A:457:ASN:O	2.31	0.63
1:A:186:THR:HB	1:A:213:TYR:CE1	2.35	0.62
1:A:379:PRO:HA	1:A:441:ILE:O	2.00	0.62
1:A:125:LEU:HD21	1:A:135:VAL:HG12	1.83	0.61
1:A:186:THR:HG23	1:A:194:GLN:O	2.01	0.60
1:A:372:ASN:O	1:A:375:ALA:HB3	2.02	0.59
1:A:157:PHE:CZ	1:A:165:ILE:HD11	2.38	0.58
1:A:328:GLN:HA	1:A:358:GLN:HG3	1.87	0.56
1:A:186:THR:HG21	1:A:213:TYR:OH	2.05	0.56
1:A:247:LYS:HG3	1:A:248:ASP:OD1	2.05	0.56
1:A:269:TYR:HA	1:A:272:ILE:HD12	1.88	0.55
1:A:339:PRO:HB2	1:A:342:SER:HB2	1.89	0.54
1:A:399:GLU:HG2	1:A:422:ARG:HD3	1.90	0.54
1:A:335:ARG:HA	1:A:349:TYR:CD2	2.43	0.54
1:A:469:SER:O	1:A:470:GLU:HG2	2.08	0.53
1:A:292:GLN:HG3	1:A:293:GLU:HG2	1.91	0.53
1:A:464:ASN:HD22	1:A:473:VAL:HG11	1.73	0.53
1:A:130:ASN:HD21	1:A:298:ASP:HB3	1.72	0.52

*Continued on next page...*



*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:138:ARG:HG3	1:A:153:ARG:HE	1.74	0.52
1:A:404:ARG:NH2	5:A:703:HOH:O	2.42	0.52
1:A:436:PRO:HB2	1:A:439:TYR:CD1	2.45	0.52
1:A:377:LEU:HG	1:A:380:HIS:HE1	1.76	0.51
1:A:396:GLY:O	1:A:427:LEU:N	2.44	0.51
1:A:185:ILE:O	1:A:195:ALA:HA	2.12	0.50
1:A:262:GLU:HG2	1:A:267:THR:O	2.11	0.50
1:A:335:ARG:HD2	1:A:368:PHE:CD2	2.45	0.50
1:A:402:GLY:O	1:A:421:ARG:N	2.41	0.49
1:A:464:ASN:N	1:A:464:ASN:OD1	2.45	0.49
1:A:475:ARG:HH12	1:A:498:LYS:HE2	1.78	0.49
1:A:516:SER:HB2	5:A:751:HOH:O	2.11	0.49
1:A:296:HIS:N	1:A:296:HIS:ND1	2.60	0.49
1:A:125:LEU:HD21	1:A:135:VAL:CG1	2.43	0.48
1:A:297:GLN:O	1:A:298:ASP:C	2.51	0.48
1:A:380:HIS:N	1:A:380:HIS:ND1	2.61	0.48
1:A:294:GLN:HB3	1:A:299:GLU:HB2	1.94	0.48
1:A:180:ASN:HB3	1:A:224:ARG:HB2	1.95	0.48
1:A:215:LEU:HD11	1:A:305:VAL:HG11	1.96	0.47
1:A:246:THR:HG23	1:A:249:GLN:NE2	2.30	0.47
1:A:297:GLN:HB2	1:A:303:VAL:CB	2.39	0.47
1:A:162:ASN:C	1:A:305:VAL:HG12	2.34	0.47
1:A:166:LEU:HD13	1:A:244:SER:CB	2.44	0.47
1:A:179:LEU:HD12	1:A:226:VAL:CG2	2.44	0.46
1:A:307:ARG:O	1:A:311:GLN:HG2	2.14	0.46
1:A:174:TYR:HB2	1:A:206:ILE:HB	1.98	0.46
1:A:253:PHE:O	1:A:261:LEU:HD21	2.16	0.46
1:A:330:GLY:O	1:A:358:GLN:NE2	2.41	0.46
1:A:198:LEU:HD23	1:A:202:ASP:HB3	1.98	0.45
1:A:416:GLU:HB3	1:A:417:PRO:CD	2.46	0.45
1:A:226:VAL:HG11	1:A:390:VAL:HG11	1.98	0.45
1:A:157:PHE:CD1	1:A:165:ILE:HD11	2.52	0.45
1:A:320:SER:O	1:A:322:GLY:N	2.50	0.44
1:A:130:ASN:ND2	1:A:298:ASP:HB3	2.32	0.44
1:A:117:PHE:HB2	1:A:432:ILE:HB	2.00	0.44
1:A:118:ASN:OD1	1:A:118:ASN:N	2.48	0.44
1:A:163:THR:O	1:A:215:LEU:HD12	2.18	0.44
1:A:382:ASN:OD1	1:A:461:ASN:ND2	2.50	0.43
1:A:338:GLU:HG3	1:A:339:PRO:HD2	2.01	0.42
1:A:459:TYR:CD2	1:A:460:GLU:HG2	2.54	0.42
1:A:341:TYR:HB3	1:A:504:TYR:CD2	2.54	0.42

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:168[A]:LYS:HE2	1:A:168[A]:LYS:HB2	1.87	0.42
1:A:271:GLU:OE1	1:A:274:ARG:NH1	2.53	0.42
1:A:309:GLN:O	1:A:313:LEU:HG	2.21	0.41
1:A:196:TYR:CD2	1:A:331:PRO:HD3	2.54	0.41
1:A:377:LEU:HG	1:A:380:HIS:CE1	2.53	0.41
1:A:404:ARG:HG2	1:A:421:ARG:HB2	2.03	0.41
1:A:390:VAL:HB	1:A:452:LEU:HB3	2.03	0.41
1:A:339:PRO:HG3	1:A:349:TYR:CZ	2.55	0.41
1:A:363:ASP:OD1	1:A:459:TYR:N	2.50	0.41
1:A:493:ILE:HG22	1:A:497:ILE:HD12	2.01	0.41
1:A:138:ARG:HH21	1:A:150:GLN:HB2	1.85	0.41
1:A:172:ALA:HA	1:A:232:ILE:HD13	2.03	0.41
1:A:398:TYR:CE1	1:A:425:ALA:HB3	2.56	0.41
1:A:268:ARG:O	1:A:269:TYR:C	2.59	0.41
1:A:370:GLU:HG2	1:A:450:ARG:HG2	2.02	0.41
1:A:480:GLU:O	1:A:483:GLU:HB3	2.21	0.41
1:A:416:GLU:CB	1:A:417:PRO:CD	2.99	0.40
1:A:130:ASN:HD21	1:A:298:ASP:CB	2.34	0.40

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	399/423 (94%)	337 (84%)	51 (13%)	11 (3%)	<b>5</b>   <b>15</b>

All (11) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	321	SER
1	A	514	GLN
1	A	141	GLN

Continued on next page...

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	290	ARG
1	A	292	GLN
1	A	326	PRO
1	A	139	PHE
1	A	418	GLU
1	A	119	SER
1	A	286	GLU
1	A	235	PRO

### 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	362/379 (96%)	327 (90%)	35 (10%)	<b>8</b> <b>23</b>

All (35) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	109	ARG
1	A	110	ARG
1	A	138	ARG
1	A	142	ARG
1	A	147	GLU
1	A	157	PHE
1	A	161	PRO
1	A	168[A]	LYS
1	A	168[B]	LYS
1	A	191	ASP
1	A	205	ARG
1	A	219	ASP
1	A	232	ILE
1	A	285	TYR
1	A	287	GLU
1	A	296	HIS
1	A	297	GLN
1	A	301	VAL

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	306	SER
1	A	315	LYS
1	A	321	SER
1	A	323	LYS
1	A	328	GLN
1	A	329	SER
1	A	335	ARG
1	A	340	ILE
1	A	404	ARG
1	A	406	GLN
1	A	446	SER
1	A	463	ARG
1	A	464	ASN
1	A	476	GLN
1	A	494	GLU
1	A	511	GLN
1	A	514	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (6) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	194	GLN
1	A	234	ASN
1	A	249	GLN
1	A	273	GLN
1	A	294	GLN
1	A	297	GLN

### 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 monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 3 ligands modelled in this entry, 2 are monoatomic - leaving 1 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
2	ACT	A	601	-	3,3,3	1.10	0	3,3,3	0.63	0

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

### 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	401/423 (94%)	0.46	43 (10%) <b>6</b>   <b>3</b>	35, 58, 112, 146	0

All (43) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	288	GLN	9.5
1	A	285	TYR	6.8
1	A	516	SER	6.3
1	A	294	GLN	5.9
1	A	290	ARG	5.7
1	A	220	ASN	5.7
1	A	289	ARG	5.6
1	A	293	GLU	5.4
1	A	279	ASN	5.4
1	A	515	GLN	5.2
1	A	514	GLN	5.2
1	A	513	GLN	4.9
1	A	517	GLU	4.4
1	A	512	GLN	4.3
1	A	406	GLN	4.3
1	A	283	GLN	3.9
1	A	280	GLU	3.8
1	A	282	GLU	3.7
1	A	321	SER	3.6
1	A	417	PRO	3.5
1	A	329	SER	3.5
1	A	316	TYR	3.5
1	A	419	GLU	3.5
1	A	287	GLU	3.4
1	A	319	SER	3.2
1	A	286	GLU	3.0
1	A	284	GLU	3.0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	511	GLN	2.8
1	A	322	GLY	2.8
1	A	297	GLN	2.8
1	A	189	ASN	2.7
1	A	221	GLN	2.7
1	A	407	GLN	2.6
1	A	110	ARG	2.6
1	A	111	GLN	2.6
1	A	278	GLY	2.5
1	A	295	SER	2.5
1	A	109	ARG	2.5
1	A	248	ASP	2.2
1	A	281	ASP	2.2
1	A	307	ARG	2.2
1	A	311	GLN	2.1
1	A	125	LEU	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 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(Å <sup>2</sup> )	Q<0.9
3	NA	A	602	1/1	0.92	0.28	50,50,50,50	1
2	ACT	A	601	4/4	0.94	0.14	60,60,62,66	0
4	K	A	603	1/1	0.99	0.57	66,66,66,66	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.