



# wwPDB X-ray Structure Validation Summary Report ⓘ

Sep 12, 2016 – 02:53 PM EDT

PDB ID : 5J3X  
Title : Structure of c-CBL Y371F  
Authors : Huang, D.T.; Buetow, L.; Dou, H.  
Deposited on : 2016-03-31  
Resolution : 2.82 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

<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	:	unknown
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20027939
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
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-20027939

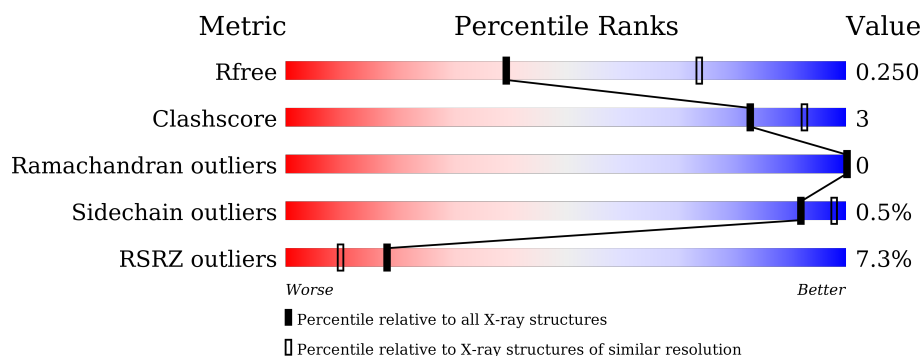
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.82 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))
$R_{free}$	91344	2676 (2.84-2.80)
Clashscore	102246	3124 (2.84-2.80)
Ramachandran outliers	100387	3072 (2.84-2.80)
Sidechain outliers	100360	3074 (2.84-2.80)
RSRZ outliers	91569	2690 (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 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	391	<div> <div>3%</div> <div>92%</div> <div>7%</div> </div>
1	B	391	<div> <div>6%</div> <div>91%</div> <div>8%</div> </div>
1	C	391	<div> <div>6%</div> <div>90%</div> <div>9%</div> </div>
1	D	391	<div> <div>9%</div> <div>90%</div> <div>10%</div> </div>
1	E	391	<div> <div>7%</div> <div>88%</div> <div>9%</div> </div>
1	F	391	<div> <div>12%</div> <div>90%</div> <div>8%</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
2	ZN	E	501	-	-	-	X

## 2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 18284 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 E3 ubiquitin-protein ligase CBL.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	386	Total	C	N	O	S	0	0	0
			3053	1962	513	553	25			
1	B	389	Total	C	N	O	S	0	0	0
			3055	1967	514	549	25			
1	C	389	Total	C	N	O	S	0	0	0
			3067	1975	516	551	25			
1	D	389	Total	C	N	O	S	0	0	0
			3031	1954	511	541	25			
1	E	381	Total	C	N	O	S	0	0	0
			2981	1921	500	535	25			
1	F	384	Total	C	N	O	S	0	0	0
			2985	1921	501	538	25			

There are 18 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	45	GLY	-	expression tag	UNP P22681
A	46	SER	-	expression tag	UNP P22681
A	371	PHE	TYR	engineered mutation	UNP P22681
B	45	GLY	-	expression tag	UNP P22681
B	46	SER	-	expression tag	UNP P22681
B	371	PHE	TYR	engineered mutation	UNP P22681
C	45	GLY	-	expression tag	UNP P22681
C	46	SER	-	expression tag	UNP P22681
C	371	PHE	TYR	engineered mutation	UNP P22681
D	45	GLY	-	expression tag	UNP P22681
D	46	SER	-	expression tag	UNP P22681
D	371	PHE	TYR	engineered mutation	UNP P22681
E	45	GLY	-	expression tag	UNP P22681
E	46	SER	-	expression tag	UNP P22681
E	371	PHE	TYR	engineered mutation	UNP P22681
F	45	GLY	-	expression tag	UNP P22681
F	46	SER	-	expression tag	UNP P22681

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
F	371	PHE	TYR	engineered mutation	UNP P22681

- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	D	2	Total 2	Zn 2	0	0
2	E	2	Total 2	Zn 2	0	0
2	B	2	Total 2	Zn 2	0	0
2	C	2	Total 2	Zn 2	0	0
2	A	2	Total 2	Zn 2	0	0
2	F	2	Total 2	Zn 2	0	0

- Molecule 3 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	D	1	Total 1	Ca 1	0	0
3	E	1	Total 1	Ca 1	0	0
3	B	1	Total 1	Ca 1	0	0
3	C	1	Total 1	Ca 1	0	0
3	A	1	Total 1	Ca 1	0	0
3	F	1	Total 1	Ca 1	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	30	Total 30	O 30	0	0
4	B	20	Total 20	O 20	0	0

*Continued on next page...*

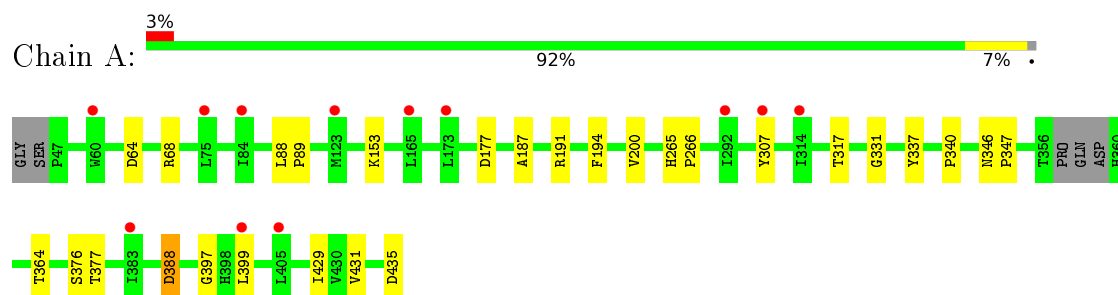
*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	C	7	Total 7	O 7	0	0
4	D	7	Total 7	O 7	0	0
4	E	14	Total 14	O 14	0	0
4	F	16	Total 16	O 16	0	0

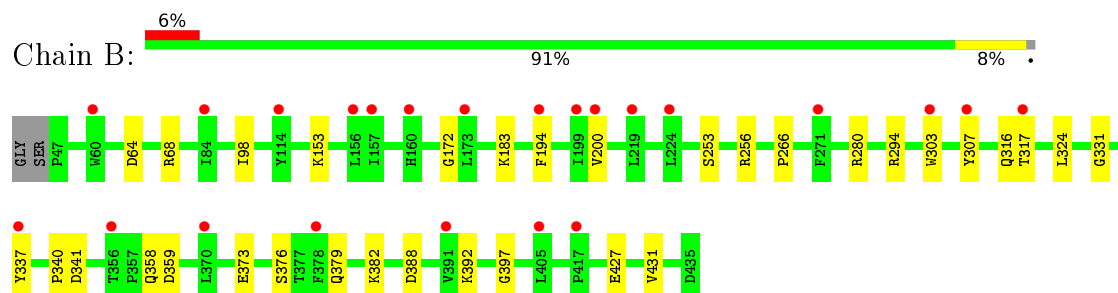
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

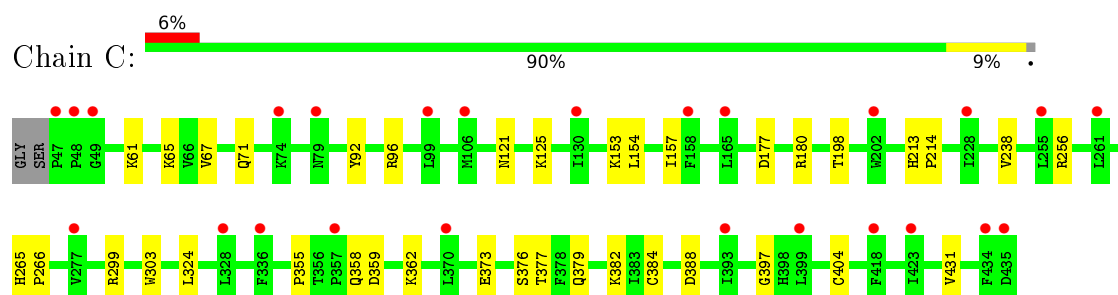
#### • Molecule 1: E3 ubiquitin-protein ligase CBL



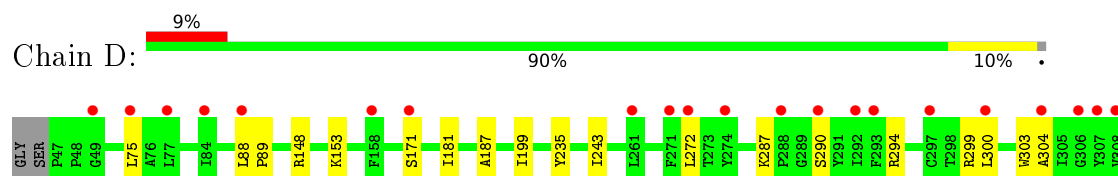
#### • Molecule 1: E3 ubiquitin-protein ligase CBL

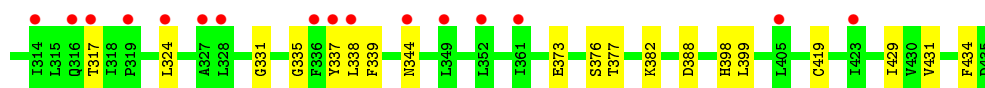


#### • Molecule 1: E3 ubiquitin-protein ligase CBL

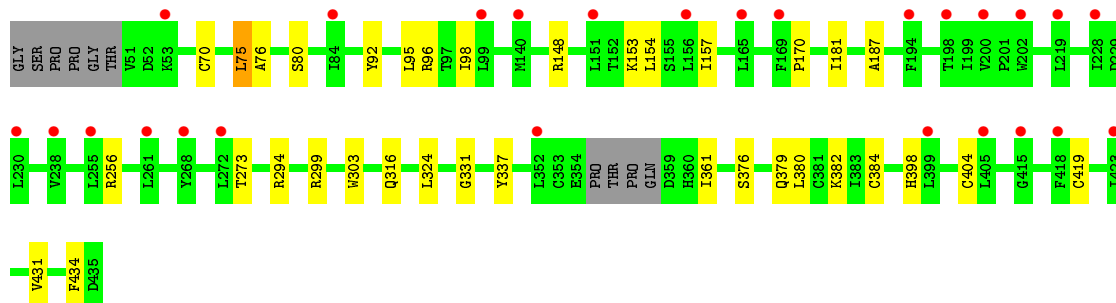
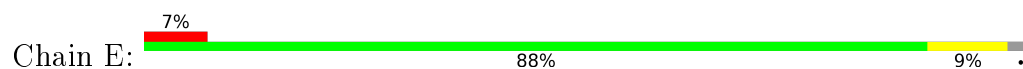


#### • Molecule 1: E3 ubiquitin-protein ligase CBL

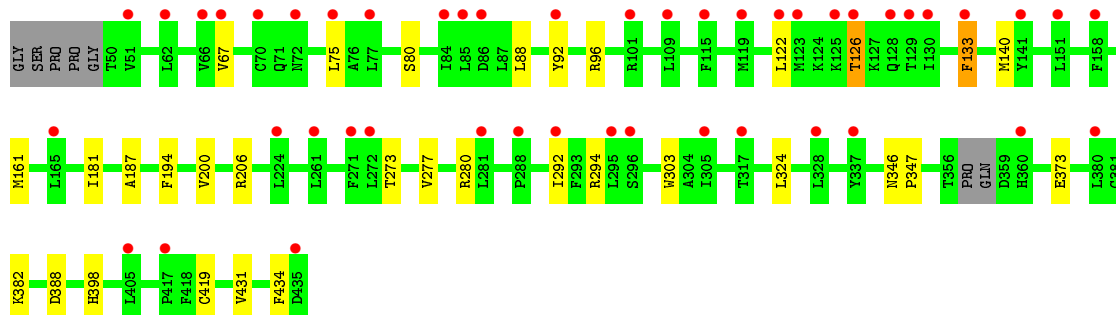
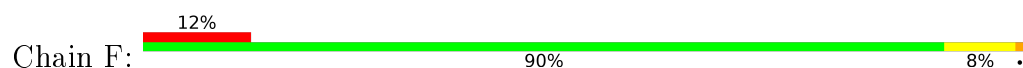




- Molecule 1: E3 ubiquitin-protein ligase CBL



- Molecule 1: E3 ubiquitin-protein ligase CBL





## 4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	147.84Å 149.04Å 344.12Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.37 – 2.82 100.39 – 2.82	Depositor EDS
% Data completeness (in resolution range)	99.6 (50.37-2.82) 91.5 (100.39-2.82)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.66 (at 2.82Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.9_1692)	Depositor
R, $R_{free}$	0.212 , 0.254 0.211 , 0.250	Depositor DCC
$R_{free}$ test set	4164 reflections (4.99%)	DCC
Wilson B-factor (Å <sup>2</sup> )	64.1	Xtriage
Anisotropy	0.779	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 71.7	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	0.043 for -k,-h,-l	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	18284	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	105.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.04% 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: ZN, CA

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	0.21	0/3133	0.35	0/4248
1	B	0.22	0/3136	0.36	0/4254
1	C	0.21	0/3148	0.35	0/4267
1	D	0.21	0/3112	0.36	0/4226
1	E	0.21	0/3057	0.35	0/4146
1	F	0.22	0/3063	0.35	0/4160
All	All	0.21	0/18649	0.35	0/25301

There are no bond length outliers.

There are no bond angle outliers.

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	3053	0	2944	17	0
1	B	3055	0	2948	19	0
1	C	3067	0	2974	23	0
1	D	3031	0	2917	22	0
1	E	2981	0	2859	20	0
1	F	2985	0	2827	17	0
2	A	2	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	B	2	0	0	0	0
2	C	2	0	0	0	0
2	D	2	0	0	0	0
2	E	2	0	0	0	0
2	F	2	0	0	0	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
3	C	1	0	0	0	0
3	D	1	0	0	0	0
3	E	1	0	0	0	0
3	F	1	0	0	0	0
4	A	30	0	0	1	0
4	B	20	0	0	0	0
4	C	7	0	0	0	0
4	D	7	0	0	0	0
4	E	14	0	0	0	0
4	F	16	0	0	1	0
All	All	18284	0	17469	112	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 3.

The worst 5 of 112 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:E:256:ARG:HH12	1:E:361:ILE:H	1.41	0.69
1:A:388:ASP:HB2	1:A:431:VAL:HG11	1.81	0.62
1:D:75:LEU:O	1:D:148:ARG:NH2	2.33	0.60
1:E:76:ALA:O	1:E:148:ARG:NH2	2.35	0.59
1:E:95:LEU:HD23	1:E:98:ILE:HD12	1.85	0.58

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	382/391 (98%)	372 (97%)	10 (3%)	0	100	100
1	B	387/391 (99%)	373 (96%)	14 (4%)	0	100	100
1	C	387/391 (99%)	369 (95%)	18 (5%)	0	100	100
1	D	387/391 (99%)	367 (95%)	20 (5%)	0	100	100
1	E	377/391 (96%)	364 (97%)	13 (3%)	0	100	100
1	F	380/391 (97%)	371 (98%)	9 (2%)	0	100	100
All	All	2300/2346 (98%)	2216 (96%)	84 (4%)	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	328/351 (93%)	326 (99%)	2 (1%)	90	97
1	B	325/351 (93%)	325 (100%)	0	100	100
1	C	328/351 (93%)	328 (100%)	0	100	100
1	D	320/351 (91%)	319 (100%)	1 (0%)	94	98
1	E	314/351 (90%)	312 (99%)	2 (1%)	90	97
1	F	311/351 (89%)	307 (99%)	4 (1%)	76	94
All	All	1926/2106 (92%)	1917 (100%)	9 (0%)	92	98

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

Mol	Chain	Res	Type
1	E	434	PHE
1	F	434	PHE
1	F	133	PHE
1	D	434	PHE
1	F	126	THR

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

### 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 ⓘ

Of 18 ligands modelled in this entry, 18 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 ⓘ

There are no such residues in this entry.

## 5.8 Polymer linkage issues ⓘ

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	386/391 (98%)	0.58	12 (3%)	52	40	55, 82, 127, 184	0
1	B	389/391 (99%)	0.66	23 (5%)	26	16	67, 96, 140, 237	0
1	C	389/391 (99%)	0.67	25 (6%)	23	13	74, 104, 158, 255	0
1	D	389/391 (99%)	0.80	37 (9%)	10	5	66, 104, 156, 214	0
1	E	381/391 (97%)	0.66	26 (6%)	20	12	71, 111, 160, 201	0
1	F	384/391 (98%)	0.84	46 (11%)	6	3	64, 112, 181, 214	0
All	All	2318/2346 (98%)	0.70	169 (7%)	18	10	55, 101, 159, 255	0

The worst 5 of 169 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	F	130	ILE	6.0
1	F	85	LEU	5.9
1	D	292	ILE	5.8
1	D	328	LEU	5.6
1	D	324	LEU	5.5

### 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(Å <sup>2</sup> )	Q<0.9
2	ZN	E	501	1/1	0.43	0.34	2.22	221,221,221,221	0
2	ZN	C	501	1/1	0.85	0.34	1.86	199,199,199,199	0
2	ZN	A	502	1/1	0.98	0.29	1.03	105,105,105,105	0
2	ZN	D	501	1/1	0.95	0.23	0.48	84,84,84,84	0
2	ZN	E	502	1/1	0.92	0.25	0.23	151,151,151,151	0
2	ZN	C	502	1/1	0.95	0.26	0.15	111,111,111,111	0
2	ZN	D	502	1/1	0.97	0.25	0.14	78,78,78,78	0
2	ZN	A	501	1/1	0.97	0.24	0.01	82,82,82,82	0
2	ZN	F	501	1/1	0.90	0.17	-0.99	93,93,93,93	0
3	CA	D	503	1/1	0.63	0.15	-1.22	155,155,155,155	0
3	CA	B	503	1/1	0.89	0.14	-1.25	164,164,164,164	0
3	CA	A	503	1/1	0.87	0.14	-1.49	160,160,160,160	0
3	CA	F	503	1/1	0.49	0.13	-1.78	212,212,212,212	0
3	CA	E	503	1/1	0.84	0.09	-2.00	177,177,177,177	0
3	CA	C	503	1/1	0.86	0.07	-2.69	197,197,197,197	0
2	ZN	B	502	1/1	0.94	0.24	-	93,93,93,93	0
2	ZN	F	502	1/1	0.69	0.46	-	224,224,224,224	0
2	ZN	B	501	1/1	0.98	0.22	-	93,93,93,93	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.