



# wwPDB X-ray Structure Validation Summary Report i

Feb 26, 2014 – 04:59 PM GMT

PDB ID : 2CE3  
Title : CRYSTAL STRUCTURE OF THE ATP-DEPENDENT CLP PROTEASE  
PROTEOLYTIC SUBUNIT 1 (CLPP1) FROM MYCOBACTERIUM TU-  
BERCULOSIS  
Authors : Segelke, B.; Kim, C.Y.; Ortiz-Lombardia, M.; Alzari, P.M.; Legin, T.  
Deposited on : 2006-02-03  
Resolution : 2.60 Å(reported)

This is a wwPDB 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/ValidationPDFNotes.html>

---

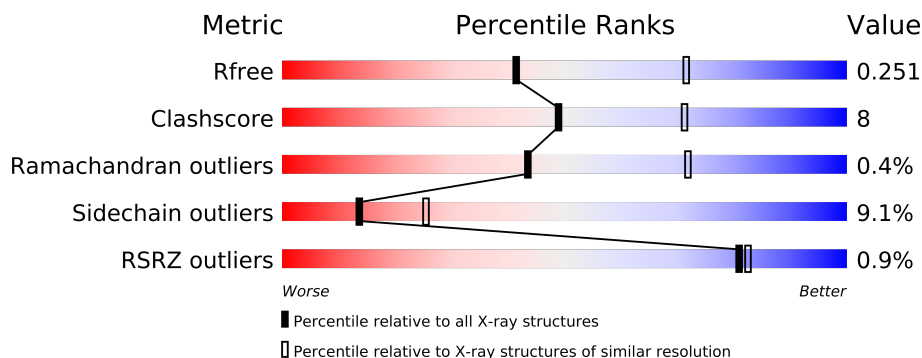
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.15 2013  
Xtriage (Phenix) : dev-1323  
EDS : stable22639  
Percentile statistics : 21963  
Refmac : 5.8.0049  
CCP4 : 6.3.0 (Settle)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)  
Validation Pipeline (wwPDB-VP) : stable22683

# 1 Overall quality at a glance

The reported resolution of this entry is 2.60 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))
$R_{free}$	66092	1718 (2.60-2.60)
Clashscore	79885	2154 (2.60-2.60)
Ramachandran outliers	78287	2113 (2.60-2.60)
Sidechain outliers	78261	2113 (2.60-2.60)
RSRZ outliers	66119	1718 (2.60-2.60)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	200	
1	B	200	
1	C	200	
1	D	200	
1	E	200	
1	F	200	
1	G	200	
1	H	200	
1	I	200	
1	J	200	
1	K	200	
1	L	200	
1	M	200	
1	N	200	

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 18231 atoms, of which 0 are hydrogen and 0 are deuterium.

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 ATP-DEPENDENT CLP PROTEASE PROTEOLYTIC SUB-UNIT 1.

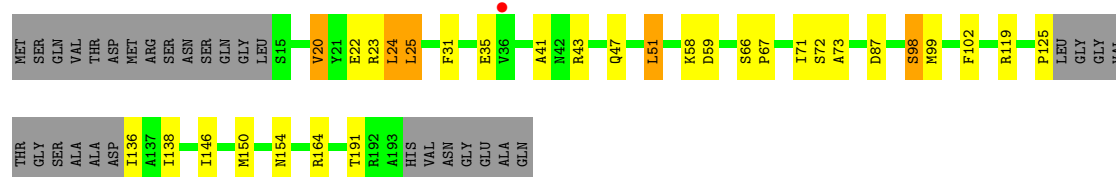
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	167	Total	C	N	O	S	0	1	0
			1295	823	217	246	9			
1	B	167	Total	C	N	O	S	0	2	0
			1303	828	220	246	9			
1	C	167	Total	C	N	O	S	0	1	0
			1295	823	217	246	9			
1	D	167	Total	C	N	O	S	0	0	0
			1288	818	215	246	9			
1	E	169	Total	C	N	O	S	0	0	0
			1304	827	220	248	9			
1	F	167	Total	C	N	O	S	0	1	0
			1295	823	217	246	9			
1	G	167	Total	C	N	O	S	0	1	0
			1295	823	217	246	9			
1	H	167	Total	C	N	O	S	0	0	0
			1288	818	215	246	9			
1	I	167	Total	C	N	O	S	0	0	0
			1288	818	215	246	9			
1	J	167	Total	C	N	O	S	0	0	0
			1288	818	215	246	9			
1	K	167	Total	C	N	O	S	0	0	0
			1288	818	215	246	9			
1	L	169	Total	C	N	O	S	0	0	0
			1304	827	220	248	9			
1	M	167	Total	C	N	O	S	0	1	0
			1295	823	217	246	9			
1	N	167	Total	C	N	O	S	0	0	0
			1288	818	215	246	9			

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	5	Total 5	O 5	0	0
2	B	21	Total 21	O 21	0	0
2	C	11	Total 11	O 11	0	0
2	D	3	Total 3	O 3	0	0
2	E	4	Total 4	O 4	0	0
2	F	12	Total 12	O 12	0	0
2	G	17	Total 17	O 17	0	0
2	H	3	Total 3	O 3	0	0
2	I	8	Total 8	O 8	0	0
2	J	7	Total 7	O 7	0	0
2	K	7	Total 7	O 7	0	0
2	L	10	Total 10	O 10	0	0
2	M	6	Total 6	O 6	0	0
2	N	3	Total 3	O 3	0	0

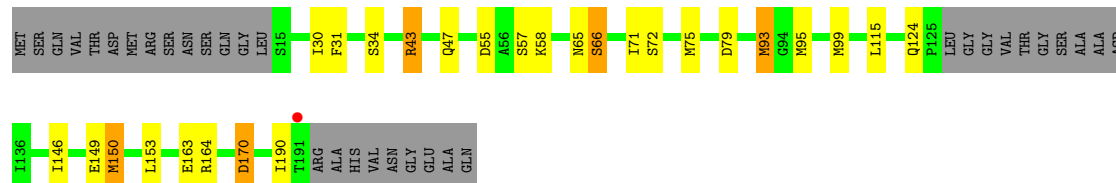


Chain E:



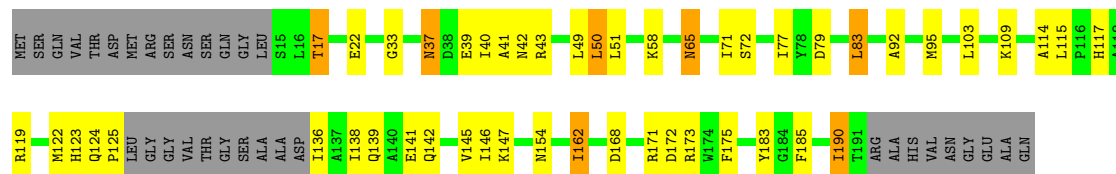
- Molecule 1: ATP-DEPENDENT CLP PROTEASE PROTEOLYTIC SUBUNIT 1

Chain F:



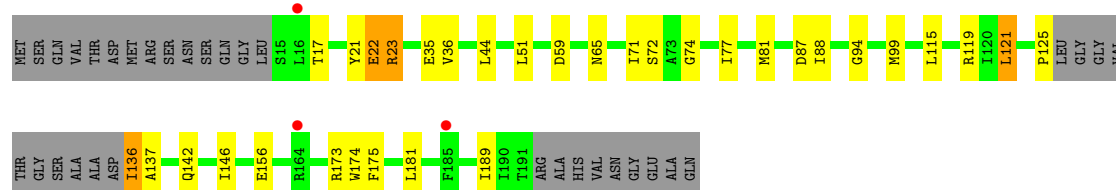
- Molecule 1: ATP-DEPENDENT CLP PROTEASE PROTEOLYTIC SUBUNIT 1

Chain G:



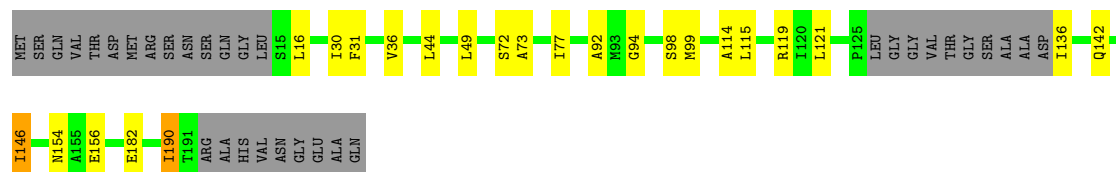
- Molecule 1: ATP-DEPENDENT CLP PROTEASE PROTEOLYTIC SUBUNIT 1

Chain H:



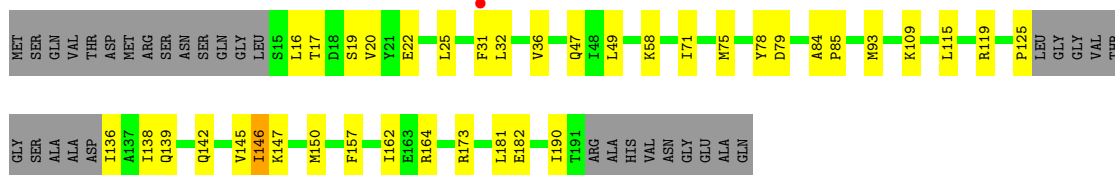
- Molecule 1: ATP-DEPENDENT CLP PROTEASE PROTEOLYTIC SUBUNIT 1

Chain I:



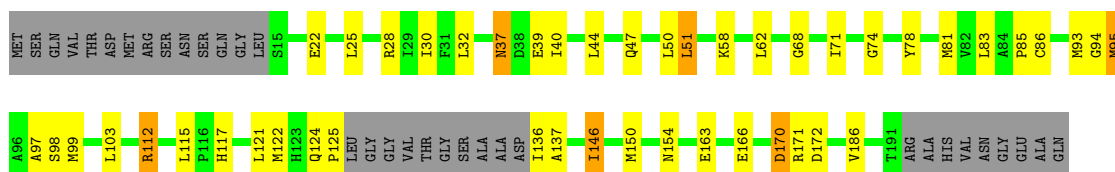
- Molecule 1: ATP-DEPENDENT CLP PROTEASE PROTEOLYTIC SUBUNIT 1

Chain J:



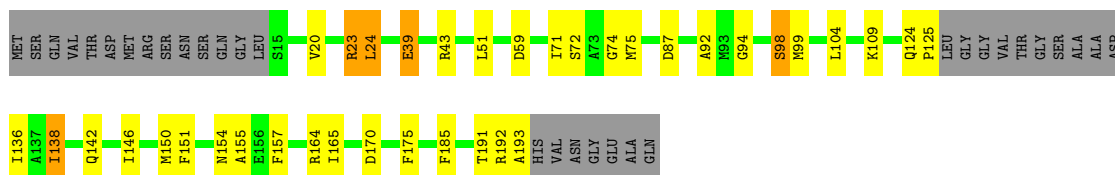
• Molecule 1: ATP-DEPENDENT CLP PROTEASE PROTEOLYTIC SUBUNIT 1

Chain K:



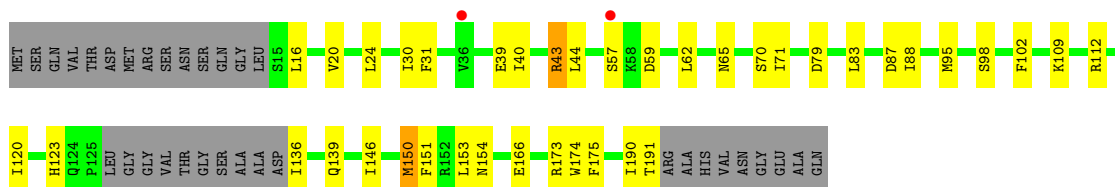
• Molecule 1: ATP-DEPENDENT CLP PROTEASE PROTEOLYTIC SUBUNIT 1

Chain L:



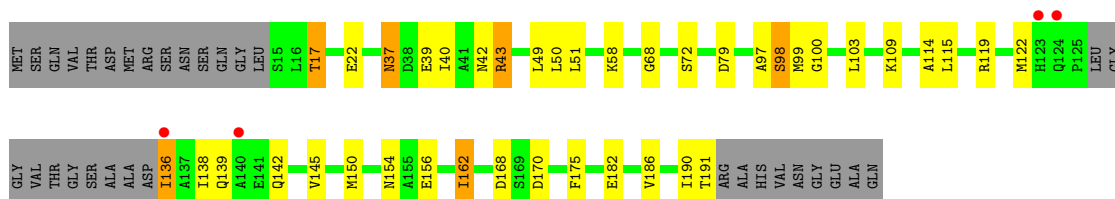
• Molecule 1: ATP-DEPENDENT CLP PROTEASE PROTEOLYTIC SUBUNIT 1

Chain M:



• Molecule 1: ATP-DEPENDENT CLP PROTEASE PROTEOLYTIC SUBUNIT 1

Chain N:



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	97.72Å 168.95Å 104.36Å 90.00° 114.83° 90.00°	Depositor
Resolution (Å)	169.03 – 2.60 49.50 – 2.60	Depositor EDS
% Data completeness (in resolution range)	95.7 (169.03-2.60) 95.7 (49.50-2.60)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.53 (at 2.61Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, $R_{free}$	0.202 , 0.261 0.196 , 0.251	Depositor DCC
$R_{free}$ test set	4541 reflections (5.28%)	DCC
Wilson B-factor (Å <sup>2</sup> )	51.8	Xtriage
Anisotropy	0.773	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 2.1	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Outliers	0 of 90592 reflections	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	18231	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	29.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.44% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.



## 5 Model quality

### 5.1 Standard geometry

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	0.54	0/1320	0.67	0/1783
1	B	0.69	0/1331	0.78	0/1797
1	C	0.54	0/1320	0.70	0/1783
1	D	0.51	0/1309	0.64	0/1768
1	E	0.57	0/1325	0.68	0/1789
1	F	0.62	0/1320	0.73	0/1783
1	G	0.63	0/1320	0.77	1/1783 (0.1%)
1	H	0.52	0/1309	0.64	0/1768
1	I	0.66	0/1309	0.73	0/1768
1	J	0.55	0/1309	0.67	0/1768
1	K	0.59	0/1309	0.71	0/1768
1	L	0.68	0/1325	0.77	0/1789
1	M	0.59	0/1320	0.68	0/1783
1	N	0.54	0/1309	0.64	0/1768
All	All	0.59	0/18435	0.70	1/24898 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	G	83	LEU	CA-CB-CG	6.41	130.05	115.30

There are no chirality outliers.

There are no planarity outliers.

### 5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the

chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1295	0	1287	10	0
1	B	1303	0	1300	31	0
1	C	1295	0	1287	12	0
1	D	1288	0	1280	22	0
1	E	1304	0	1298	16	0
1	F	1295	0	1287	17	0
1	G	1295	0	1287	34	0
1	H	1288	0	1280	20	0
1	I	1288	0	1280	18	0
1	J	1288	0	1280	20	0
1	K	1288	0	1280	34	0
1	L	1304	0	1298	29	0
1	M	1295	0	1287	21	0
1	N	1288	0	1280	26	0
2	A	5	0	0	0	0
2	B	21	0	0	0	0
2	C	11	0	0	1	0
2	D	3	0	0	0	0
2	E	4	0	0	0	0
2	F	12	0	0	0	0
2	G	17	0	0	0	0
2	H	3	0	0	0	0
2	I	8	0	0	0	0
2	J	7	0	0	0	0
2	K	7	0	0	0	0
2	L	10	0	0	0	0
2	M	6	0	0	0	0
2	N	3	0	0	0	0
All	All	18231	0	18011	276	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 8.

The worst 5 of 276 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:I:115:LEU:HD23	1:I:190:ILE:HD12	1.25	1.13
1:G:122:MET:CE	1:G:168:ASP:HB3	1.83	1.08
1:G:122:MET:HE3	1:G:168:ASP:HB3	1.34	1.05
1:A:65:ASN:HD22	1:G:42:ASN:HD21	1.19	0.90
1:I:115:LEU:CD2	1:I:190:ILE:HD12	2.04	0.87

There are no symmetry-related clashes.

## 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	164/200 (82%)	159 (97%)	4 (2%)	1 (1%)	33	63
1	B	165/200 (82%)	159 (96%)	5 (3%)	1 (1%)	33	63
1	C	164/200 (82%)	158 (96%)	5 (3%)	1 (1%)	33	63
1	D	163/200 (82%)	157 (96%)	6 (4%)	0	100	100
1	E	165/200 (82%)	156 (94%)	9 (6%)	0	100	100
1	F	164/200 (82%)	159 (97%)	5 (3%)	0	100	100
1	G	164/200 (82%)	157 (96%)	7 (4%)	0	100	100
1	H	163/200 (82%)	152 (93%)	9 (6%)	2 (1%)	19	39
1	I	163/200 (82%)	156 (96%)	6 (4%)	1 (1%)	33	63
1	J	163/200 (82%)	157 (96%)	6 (4%)	0	100	100
1	K	163/200 (82%)	156 (96%)	6 (4%)	1 (1%)	33	63
1	L	165/200 (82%)	159 (96%)	5 (3%)	1 (1%)	33	63
1	M	164/200 (82%)	154 (94%)	10 (6%)	0	100	100
1	N	163/200 (82%)	155 (95%)	7 (4%)	1 (1%)	33	63
All	All	2293/2800 (82%)	2194 (96%)	90 (4%)	9 (0%)	43	72

5 of 9 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	K	94	GLY
1	C	94	GLY
1	H	94	GLY
1	I	94	GLY
1	L	94	GLY

### 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	134/157 (85%)	127 (95%)	7 (5%)	32	59
1	B	135/157 (86%)	124 (92%)	11 (8%)	17	32
1	C	134/157 (85%)	127 (95%)	7 (5%)	32	59
1	D	133/157 (85%)	122 (92%)	11 (8%)	16	30
1	E	134/157 (85%)	118 (88%)	16 (12%)	8	14
1	F	134/157 (85%)	123 (92%)	11 (8%)	17	31
1	G	134/157 (85%)	114 (85%)	20 (15%)	4	7
1	H	133/157 (85%)	125 (94%)	8 (6%)	27	51
1	I	133/157 (85%)	125 (94%)	8 (6%)	27	51
1	J	133/157 (85%)	120 (90%)	13 (10%)	12	21
1	K	133/157 (85%)	120 (90%)	13 (10%)	12	21
1	L	134/157 (85%)	121 (90%)	13 (10%)	12	22
1	M	134/157 (85%)	121 (90%)	13 (10%)	12	22
1	N	133/157 (85%)	114 (86%)	19 (14%)	5	8
All	All	1871/2198 (85%)	1701 (91%)	170 (9%)	14	25

5 of 170 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	G	136	ILE
1	I	154	ASN
1	N	51	LEU
1	G	145	VAL
1	H	72	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 30 such sidechains are listed below:

Mol	Chain	Res	Type
1	I	154	ASN
1	K	37	ASN
1	N	124	GLN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	J	124	GLN
1	K	47	GLN

### 5.3.3 RNA ⓘ

There are no RNA chains 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 ⓘ

There are no ligands in this entry.

### 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 ⓘ

### 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	167/200 (83%)	0.06	3 (1%) 65 64	13, 29, 37, 42	2 (1%)
1	B	167/200 (83%)	0.21	1 (0%) 86 89	18, 28, 37, 47	0
1	C	167/200 (83%)	0.29	3 (1%) 65 64	22, 29, 37, 39	0
1	D	167/200 (83%)	0.05	2 (1%) 75 77	20, 29, 37, 45	0
1	E	169/200 (84%)	0.04	1 (0%) 86 89	14, 28, 39, 46	0
1	F	167/200 (83%)	0.02	1 (0%) 86 89	20, 29, 38, 42	0
1	G	167/200 (83%)	0.13	0 100 100	19, 29, 37, 42	1 (0%)
1	H	167/200 (83%)	0.20	3 (1%) 65 64	18, 29, 38, 46	0
1	I	167/200 (83%)	0.20	0 100 100	17, 29, 38, 46	1 (0%)
1	J	167/200 (83%)	0.10	1 (0%) 86 89	19, 29, 37, 41	0
1	K	167/200 (83%)	-0.02	0 100 100	18, 29, 37, 45	1 (0%)
1	L	169/200 (84%)	0.16	0 100 100	18, 29, 36, 45	1 (0%)
1	M	167/200 (83%)	0.15	2 (1%) 75 77	20, 29, 38, 45	1 (0%)
1	N	167/200 (83%)	0.15	4 (2%) 56 53	22, 29, 35, 41	0
All	All	2342/2800 (83%)	0.12	21 (0%) 81 82	13, 29, 37, 47	7 (0%)

The worst 5 of 21 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	16	LEU	5.4
1	E	36	VAL	3.7
1	N	140	ALA	3.7
1	D	124	GLN	3.5
1	M	36	VAL	3.2

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

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

## 6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 6.4 Ligands ⓘ

There are no ligands in this entry.

## 6.5 Other polymers ⓘ

There are no such residues in this entry.