



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 15, 2017 – 01:12 am GMT

PDB ID : 4CQ1  
Title : Crystal structure of the neuronal isoform of PTB  
Authors : Joshi, A.; Buckroyd, A.N.; Curry, S.  
Deposited on : 2014-02-10  
Resolution : 1.69 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
Xtriage (Phenix)	:	1.9-1692
EDS	:	trunk28620
Percentile statistics	:	20161228.v01 (using entries in the PDB archive December 28th 2016)
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)	:	recalc28949

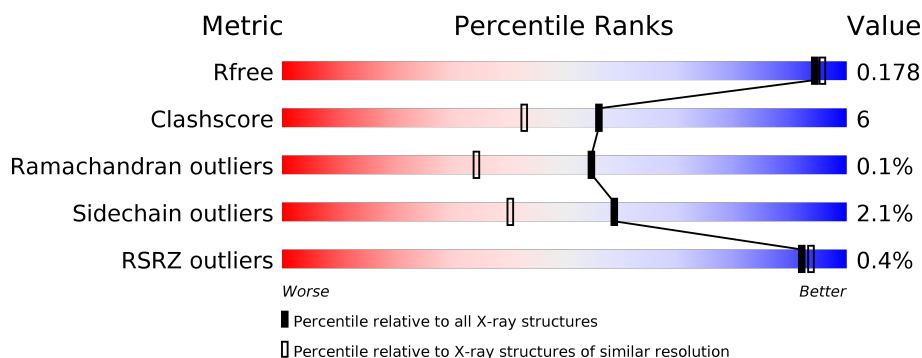
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.69 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))
$R_{free}$	100719	3453 (1.70-1.70)
Clashscore	112137	3876 (1.70-1.70)
Ramachandran outliers	110173	3815 (1.70-1.70)
Sidechain outliers	110143	3815 (1.70-1.70)
RSRZ outliers	101464	3491 (1.70-1.70)

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	196	<div> <div>%</div> <div> <div></div> <div>82%</div> <div>15%</div> <div>• •</div> </div> </div>
1	B	196	<div> <div>%</div> <div> <div></div> <div>90%</div> <div>8%</div> <div>•</div> </div> </div>
1	C	196	<div> <div></div> <div> <div></div> <div>79%</div> <div>16%</div> <div>• •</div> </div> </div>
1	D	196	<div> <div></div> <div> <div></div> <div>85%</div> <div>11%</div> <div>• •</div> </div> </div>
1	E	196	<div> <div>%</div> <div> <div></div> <div>80%</div> <div>17%</div> <div>•</div> </div> </div>
1	F	196	<div> <div>%</div> <div> <div></div> <div>88%</div> <div>9%</div> <div>• •</div> </div> </div>

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	G	196	<div><div>%</div><div><div></div><div></div><div></div></div><div>90%7% . .</div></div>
1	H	196	<div><div></div><div></div><div></div></div> <div>84%14% .</div>

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 12397 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 POLYPYRIMIDINE TRACT-BINDING PROTEIN 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	191	Total	C	N	O	S	0	0	0
			1453	924	252	271	6			
1	B	192	Total	C	N	O	S	0	0	0
			1462	932	255	269	6			
1	C	188	Total	C	N	O	S	0	0	0
			1434	916	249	263	6			
1	D	189	Total	C	N	O	S	0	0	0
			1447	923	252	266	6			
1	E	189	Total	C	N	O	S	0	0	0
			1438	917	248	267	6			
1	F	191	Total	C	N	O	S	0	0	0
			1452	926	253	267	6			
1	G	191	Total	C	N	O	S	0	0	0
			1447	921	251	269	6			
1	H	191	Total	C	N	O	S	0	0	0
			1462	929	256	271	6			

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

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	G	1	Total	Zn	0	0
			1	1		
2	E	2	Total	Zn	0	0
			2	2		
2	B	3	Total	Zn	0	0
			3	3		
2	C	1	Total	Zn	0	0
			1	1		
2	A	2	Total	Zn	0	0
			2	2		
2	F	3	Total	Zn	0	0
			3	3		

- Molecule 3 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	G	1	Total Cl 1 1	0	0
3	D	1	Total Cl 1 1	0	0
3	E	1	Total Cl 1 1	0	0
3	B	1	Total Cl 1 1	0	0
3	C	1	Total Cl 1 1	0	0
3	A	1	Total Cl 1 1	0	0
3	F	2	Total Cl 2 2	0	0

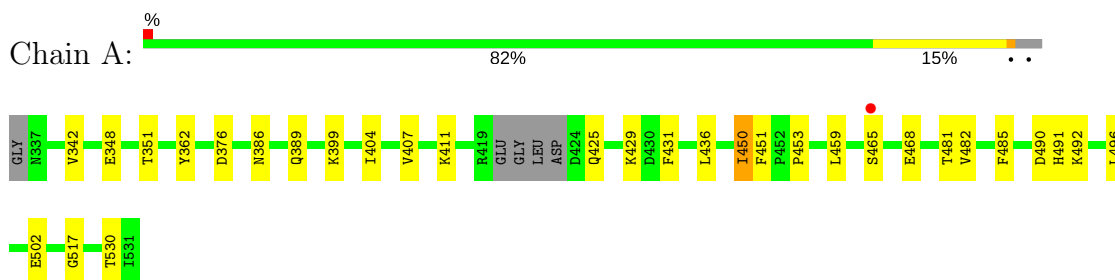
- Molecule 4 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	89	Total O 89 89	0	0
4	B	118	Total O 118 118	0	0
4	C	87	Total O 87 87	0	0
4	D	108	Total O 108 108	0	0
4	E	88	Total O 88 88	0	0
4	F	99	Total O 99 99	0	0
4	G	85	Total O 85 85	0	0
4	H	108	Total O 108 108	0	0

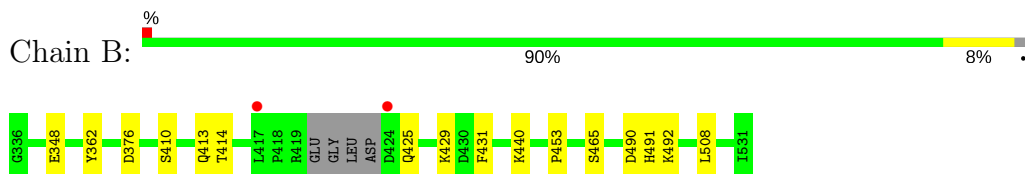
### 3 Residue-property plots [i](#)

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

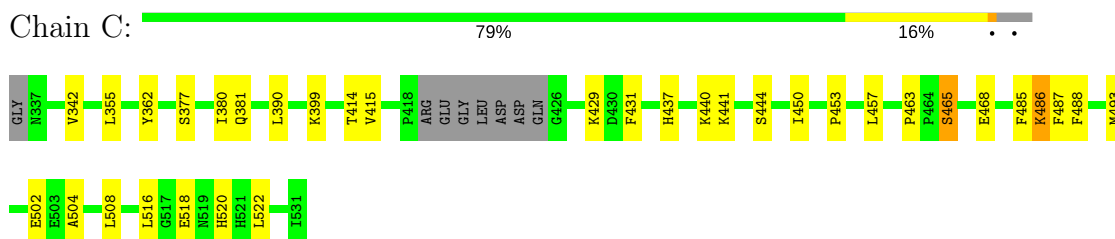
#### • Molecule 1: POLYPYRIMIDINE TRACT-BINDING PROTEIN 2



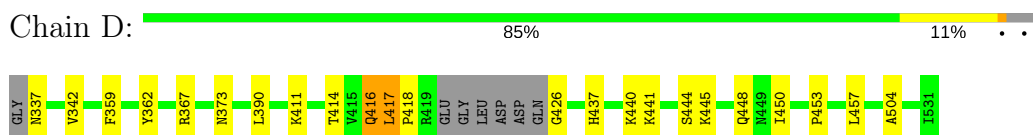
#### • Molecule 1: POLYPYRIMIDINE TRACT-BINDING PROTEIN 2



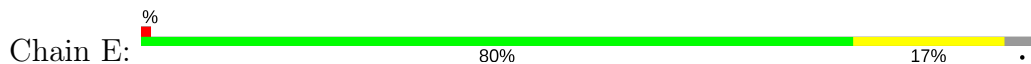
#### • Molecule 1: POLYPYRIMIDINE TRACT-BINDING PROTEIN 2

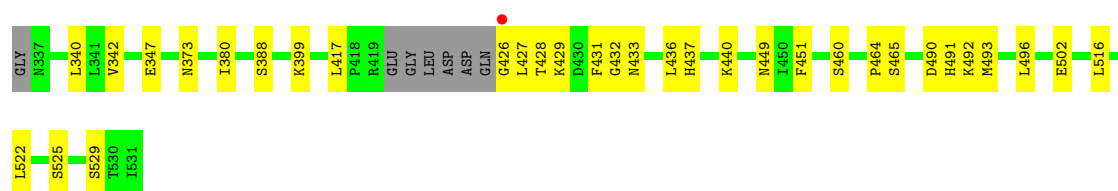


#### • Molecule 1: POLYPYRIMIDINE TRACT-BINDING PROTEIN 2

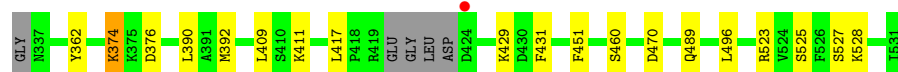
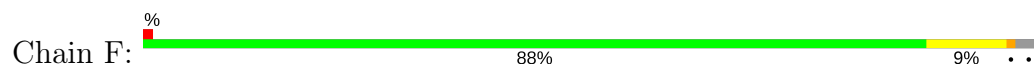


#### • Molecule 1: POLYPYRIMIDINE TRACT-BINDING PROTEIN 2

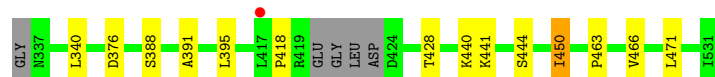
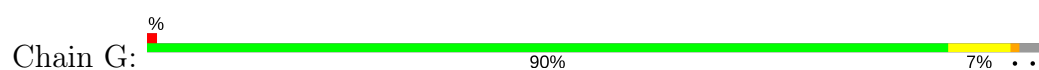




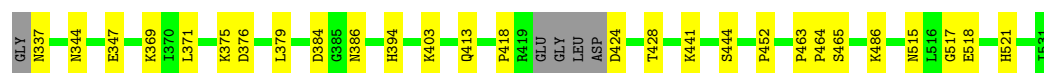
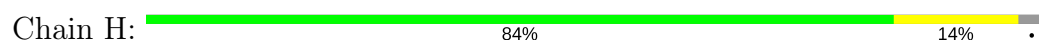
• Molecule 1: POLYPYRIMIDINE TRACT-BINDING PROTEIN 2



• Molecule 1: POLYPYRIMIDINE TRACT-BINDING PROTEIN 2



• Molecule 1: POLYPYRIMIDINE TRACT-BINDING PROTEIN 2



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	61.00Å 65.81Å 99.58Å 89.99° 90.00° 89.99°	Depositor
Resolution (Å)	40.81 – 1.69 40.81 – 1.69	Depositor EDS
% Data completeness (in resolution range)	88.9 (40.81-1.69) 88.9 (40.81-1.69)	Depositor EDS
$R_{merge}$	0.10	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.60 (at 1.69Å)	Xtriage
Refinement program	REFMAC 5.8.0049	Depositor
R, $R_{free}$	0.170 , 0.208 0.178 , 0.178	Depositor DCC
$R_{free}$ test set	7740 reflections (5.30%)	DCC
Wilson B-factor (Å <sup>2</sup> )	13.8	Xtriage
Anisotropy	0.644	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 22.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.45$ , $\langle L^2 \rangle = 0.27$	Xtriage
Estimated twinning fraction	0.457 for h,-k,-l 0.358 for -h,k,-l 0.366 for -h,-k,l	Xtriage
Reported twinning fraction	0.153 for H, K, L 0.245 for -H, -K, L 0.176 for H, -K, -L 0.425 for -H, K, -L	Depositor
Outliers	0 of 153727 reflections	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	12397	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	18.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 44.01 % of the origin peak, indicating pseudo translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo translational symmetry is equal to 1.6234e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<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: ZN, CL

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	0.34	0/1481	0.49	0/2006
1	B	0.37	0/1490	0.51	0/2015
1	C	0.33	0/1462	0.49	0/1980
1	D	0.36	0/1475	0.50	0/1997
1	E	0.36	0/1466	0.51	0/1988
1	F	0.33	0/1480	0.49	0/2004
1	G	0.34	0/1475	0.50	0/2002
1	H	0.34	0/1490	0.50	0/2017
All	All	0.35	0/11819	0.50	0/16009

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	2
1	C	0	1
All	All	0	3

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	465	SER	Peptide
1	A	517	GLY	Peptide
1	C	465	SER	Peptide

## 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	1453	0	1411	26	0
1	B	1462	0	1442	15	0
1	C	1434	0	1409	23	0
1	D	1447	0	1422	20	0
1	E	1438	0	1397	24	0
1	F	1452	0	1419	15	0
1	G	1447	0	1389	8	0
1	H	1462	0	1430	14	0
2	A	2	0	0	0	0
2	B	3	0	0	0	0
2	C	1	0	0	0	0
2	E	2	0	0	0	0
2	F	3	0	0	0	0
2	G	1	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	2	0	0	0	0
3	G	1	0	0	0	0
4	A	89	0	0	2	1
4	B	118	0	0	4	0
4	C	87	0	0	0	1
4	D	108	0	0	3	0
4	E	88	0	0	1	0
4	F	99	0	0	3	0
4	G	85	0	0	0	0
4	H	108	0	0	1	0
All	All	12397	0	11319	134	1

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 6.

The worst 5 of 134 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:F:411:LYS:CB	4:F:2046:HOH:O	1.87	1.20
1:D:390:LEU:HD21	1:D:450:ILE:HG21	1.41	1.02
4:B:2085:HOH:O	1:D:414:THR:HG22	1.59	1.02
1:B:490:ASP:HB2	4:B:2102:HOH:O	1.61	0.99
1:A:411:LYS:HE3	1:C:465:SER:HA	1.45	0.97

All (1) 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)
4:A:2078:HOH:O	4:C:2045:HOH:O[1_554]	1.48	0.72

## 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	187/196 (95%)	183 (98%)	3 (2%)	1 (0%)	32	15
1	B	188/196 (96%)	183 (97%)	5 (3%)	0	100	100
1	C	184/196 (94%)	182 (99%)	2 (1%)	0	100	100
1	D	185/196 (94%)	181 (98%)	3 (2%)	1 (0%)	32	15
1	E	185/196 (94%)	179 (97%)	6 (3%)	0	100	100
1	F	187/196 (95%)	183 (98%)	4 (2%)	0	100	100
1	G	187/196 (95%)	185 (99%)	2 (1%)	0	100	100
1	H	187/196 (95%)	184 (98%)	3 (2%)	0	100	100
All	All	1490/1568 (95%)	1460 (98%)	28 (2%)	2 (0%)	55	34

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	450	ILE

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	D	418	PRO

### 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	154/172 (90%)	150 (97%)	4 (3%)	51	31
1	B	156/172 (91%)	154 (99%)	2 (1%)	73	60
1	C	152/172 (88%)	150 (99%)	2 (1%)	73	60
1	D	154/172 (90%)	152 (99%)	2 (1%)	73	60
1	E	152/172 (88%)	148 (97%)	4 (3%)	51	31
1	F	153/172 (89%)	149 (97%)	4 (3%)	51	31
1	G	150/172 (87%)	147 (98%)	3 (2%)	60	42
1	H	156/172 (91%)	151 (97%)	5 (3%)	44	22
All	All	1227/1376 (89%)	1201 (98%)	26 (2%)	59	40

5 of 26 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	436	LEU
1	F	374	LYS
1	H	413	GLN
1	E	525	SER
1	E	529	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 14 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	497	GLN
1	E	449	ASN
1	H	413	GLN
1	D	416	GLN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	H	394	HIS

### 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 20 ligands modelled in this entry, 20 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	191/196 (97%)	-0.32	1 (0%) 90 93	8, 17, 33, 47	0
1	B	192/196 (97%)	-0.37	2 (1%) 82 86	7, 17, 32, 63	0
1	C	188/196 (95%)	-0.40	0 100 100	8, 17, 32, 37	0
1	D	189/196 (96%)	-0.45	0 100 100	7, 15, 31, 42	0
1	E	189/196 (96%)	-0.40	1 (0%) 90 93	9, 15, 33, 46	0
1	F	191/196 (97%)	-0.34	1 (0%) 90 93	8, 16, 33, 63	0
1	G	191/196 (97%)	-0.35	1 (0%) 90 93	8, 17, 34, 43	0
1	H	191/196 (97%)	-0.40	0 100 100	8, 16, 31, 46	0
All	All	1522/1568 (97%)	-0.38	6 (0%) 92 93	7, 16, 33, 63	0

The worst 5 of 6 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	424	ASP	4.3
1	G	417	LEU	3.9
1	F	424	ASP	3.2
1	A	465	SER	3.1
1	E	426	GLY	3.1

### 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
3	CL	D	1532	1/1	0.99	0.07	0.08	26,26,26,26	0
3	CL	C	1533	1/1	0.99	0.05	-1.67	24,24,24,24	0
2	ZN	E	1534	1/1	0.99	0.03	-2.06	24,24,24,24	0
2	ZN	A	1532	1/1	0.99	0.03	-2.80	22,22,22,22	0
2	ZN	A	1534	1/1	0.99	0.03	-2.85	23,23,23,23	0
2	ZN	B	1535	1/1	0.99	0.02	-3.41	28,28,28,28	0
3	CL	F	1533	1/1	0.99	0.06	-3.52	19,19,19,19	0
3	CL	A	1533	1/1	1.00	0.03	-5.20	20,20,20,20	0
2	ZN	E	1532	1/1	0.99	0.04	-	20,20,20,20	0
2	ZN	F	1534	1/1	1.00	0.02	-	19,19,19,19	0
3	CL	G	1533	1/1	0.98	0.04	-	25,25,25,25	0
3	CL	F	1535	1/1	1.00	0.05	-	28,28,28,28	0
2	ZN	F	1532	1/1	1.00	0.03	-	16,16,16,16	0
2	ZN	B	1534	1/1	1.00	0.02	-	18,18,18,18	0
2	ZN	C	1532	1/1	0.99	0.03	-	22,22,22,22	0
2	ZN	G	1532	1/1	1.00	0.03	-	21,21,21,21	0
2	ZN	F	1536	1/1	0.99	0.04	-	24,24,24,24	0
3	CL	E	1533	1/1	0.99	0.03	-	23,23,23,23	0
3	CL	B	1533	1/1	0.99	0.06	-	22,22,22,22	0
2	ZN	B	1532	1/1	1.00	0.03	-	19,19,19,19	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.