



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jul 19, 2017 – 12:00 PM EDT

PDB ID : 3GTV  
Title : Human-mouse SOD1 chimera  
Authors : Seetharaman, S.V.; Taylor, A.B.; Hart, P.J.  
Deposited on : unknown  
Resolution : 2.20 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	:	rb-20029824
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)	:	rb-20029824

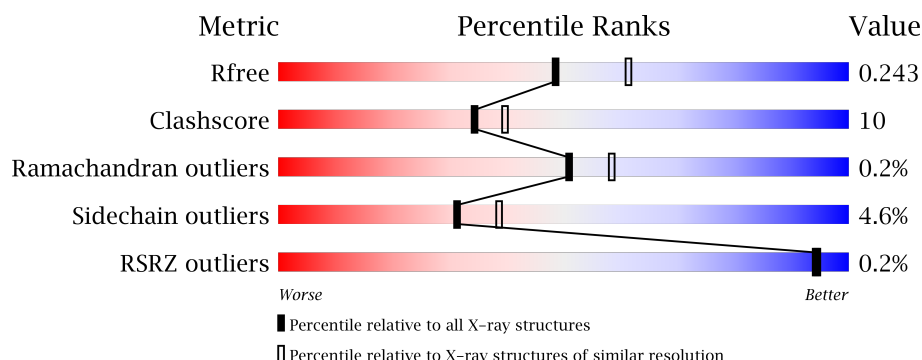
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.20 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	4002 (2.20-2.20)
Clashscore	112137	4730 (2.20-2.20)
Ramachandran outliers	110173	4656 (2.20-2.20)
Sidechain outliers	110143	4657 (2.20-2.20)
RSRZ outliers	101464	4033 (2.20-2.20)

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	153	
1	B	153	
1	C	153	
1	D	153	
1	E	153	

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	F	153	
1	G	153	
1	H	153	
1	I	153	
1	J	153	
1	K	153	
1	L	153	

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	B	155	-	-	-	X
2	ZN	F	155	-	-	-	X
2	ZN	H	155	-	-	-	X
2	ZN	L	155	-	-	-	X

## 2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 15217 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 Superoxide dismutase [Cu-Zn].

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	153	Total	C	N	O	S	0	0	0
			1116	682	208	222	4			
1	B	153	Total	C	N	O	S	0	0	0
			1116	682	208	222	4			
1	C	153	Total	C	N	O	S	0	0	0
			1116	682	208	222	4			
1	D	153	Total	C	N	O	S	0	0	0
			1116	682	208	222	4			
1	E	153	Total	C	N	O	S	0	0	0
			1116	682	208	222	4			
1	F	153	Total	C	N	O	S	0	0	0
			1116	682	208	222	4			
1	G	153	Total	C	N	O	S	0	1	0
			1121	685	208	224	4			
1	H	153	Total	C	N	O	S	0	0	0
			1116	682	208	222	4			
1	I	153	Total	C	N	O	S	0	0	0
			1116	682	208	222	4			
1	J	153	Total	C	N	O	S	0	0	0
			1116	682	208	222	4			
1	K	153	Total	C	N	O	S	0	0	0
			1116	682	208	222	4			
1	L	153	Total	C	N	O	S	0	0	0
			1116	682	208	222	4			

- 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	J	1	Total	Zn	0	0
			1	1		

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	D	1	Total 1	Zn 1	0	0
2	K	1	Total 1	Zn 1	0	0
2	E	1	Total 1	Zn 1	0	0
2	H	1	Total 1	Zn 1	0	0
2	B	1	Total 1	Zn 1	0	0
2	I	1	Total 1	Zn 1	0	0
2	C	1	Total 1	Zn 1	0	0
2	A	1	Total 1	Zn 1	0	0
2	L	1	Total 1	Zn 1	0	0
2	F	1	Total 1	Zn 1	0	0

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	173	Total 173	O 173	0	0
3	B	189	Total 189	O 189	0	0
3	C	184	Total 184	O 184	0	0
3	D	177	Total 177	O 177	0	0
3	E	91	Total 91	O 91	0	0
3	F	130	Total 130	O 130	0	0
3	G	166	Total 166	O 166	0	0
3	H	185	Total 185	O 185	0	0
3	I	156	Total 156	O 156	0	0

*Continued on next page...*

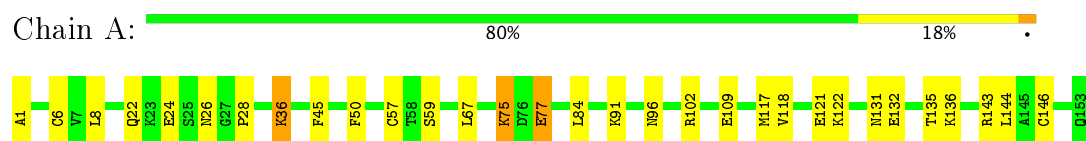
*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	J	137	Total 137	O 137	0	0
3	K	112	Total 112	O 112	0	0
3	L	108	Total 108	O 108	0	0

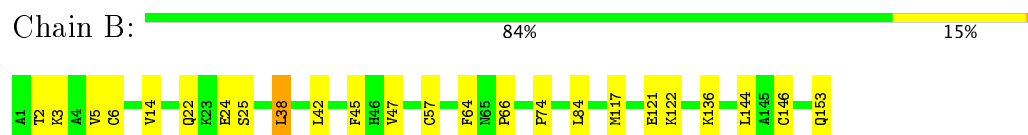
### 3 Residue-property plots [i](#)

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.

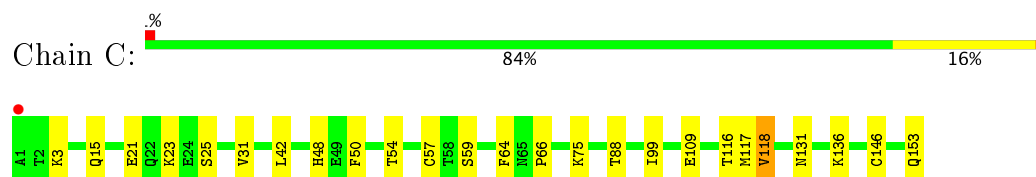
- Molecule 1: Superoxide dismutase [Cu-Zn]



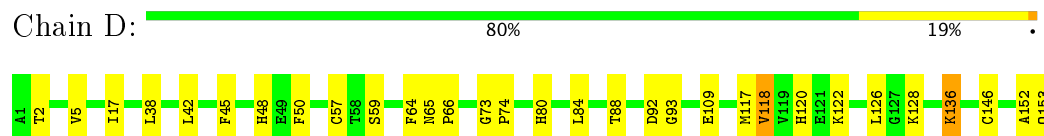
- Molecule 1: Superoxide dismutase [Cu-Zn]



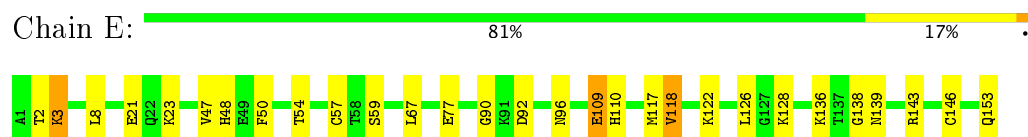
- Molecule 1: Superoxide dismutase [Cu-Zn]



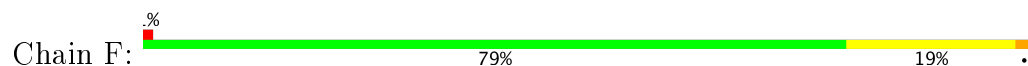
- Molecule 1: Superoxide dismutase [Cu-Zn]

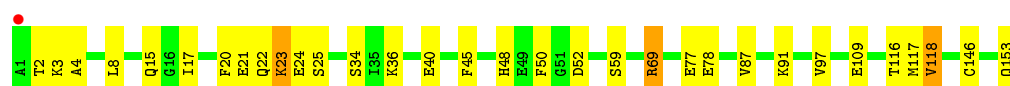


- Molecule 1: Superoxide dismutase [Cu-Zn]



- Molecule 1: Superoxide dismutase [Cu-Zn]





- Molecule 1: Superoxide dismutase [Cu-Zn]

Chain G: 80% 18%



- Molecule 1: Superoxide dismutase [Cu-Zn]

Chain H: 82% 17%



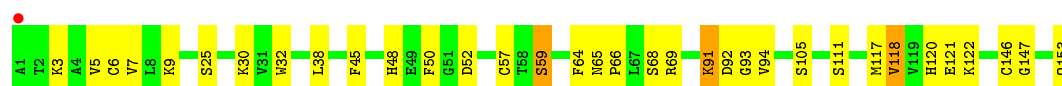
- Molecule 1: Superoxide dismutase [Cu-Zn]

Chain I: 80% 20%



- Molecule 1: Superoxide dismutase [Cu-Zn]

Chain J: 78% 20%



- Molecule 1: Superoxide dismutase [Cu-Zn]

Chain K: 75% 24%



- Molecule 1: Superoxide dismutase [Cu-Zn]

Chain L: 78% 20%





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	112.62Å 144.01Å 112.61Å 90.00° 119.82° 90.00°	Depositor
Resolution (Å)	40.35 – 2.20 40.35 – 2.20	Depositor EDS
% Data completeness (in resolution range)	93.8 (40.35-2.20) 93.8 (40.35-2.20)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.11	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.49 (at 2.20Å)	Xtriage
Refinement program	PHENIX	Depositor
R, $R_{free}$	0.184 , 0.231 0.196 , 0.243	Depositor DCC
$R_{free}$ test set	7463 reflections (5.04%)	DCC
Wilson B-factor (Å <sup>2</sup> )	25.2	Xtriage
Anisotropy	0.630	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 25.2	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.44$ , $\langle L^2 \rangle = 0.26$	Xtriage
Estimated twinning fraction	0.066 for l,k,-h-l 0.066 for -h-l,k,h 0.068 for -h-l,-k,l 0.087 for h,-k,-h-l 0.126 for l,-k,h	Xtriage
Reported twinning fraction	0.124 for l,-k,h	Depositor
Outliers	1 of 148404 reflections (0.001%)	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	15217	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	28.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 15.07% 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

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	0.40	0/1134	0.57	0/1525
1	B	0.42	0/1134	0.59	0/1525
1	C	0.41	0/1134	0.58	0/1525
1	D	0.43	0/1134	0.59	0/1525
1	E	0.38	0/1134	0.56	0/1525
1	F	0.44	0/1134	0.59	0/1525
1	G	0.38	0/1142	0.58	0/1536
1	H	0.40	0/1134	0.58	0/1525
1	I	0.40	0/1134	0.58	0/1525
1	J	0.39	0/1134	0.59	0/1525
1	K	0.39	0/1134	0.56	0/1525
1	L	0.39	0/1134	0.55	0/1525
All	All	0.40	0/13616	0.58	0/18311

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	1116	0	1091	25	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	B	1116	0	1091	17	0
1	C	1116	0	1091	23	0
1	D	1116	0	1091	24	0
1	E	1116	0	1091	22	0
1	F	1116	0	1091	21	0
1	G	1121	0	1095	22	0
1	H	1116	0	1091	18	0
1	I	1116	0	1091	24	0
1	J	1116	0	1091	25	0
1	K	1116	0	1091	26	0
1	L	1116	0	1091	22	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
2	C	1	0	0	0	0
2	D	1	0	0	0	0
2	E	1	0	0	0	0
2	F	1	0	0	0	0
2	G	1	0	0	0	0
2	H	1	0	0	0	0
2	I	1	0	0	0	0
2	J	1	0	0	0	0
2	K	1	0	0	0	0
2	L	1	0	0	0	0
3	A	173	0	0	5	0
3	B	189	0	0	2	0
3	C	184	0	0	2	0
3	D	177	0	0	3	0
3	E	91	0	0	1	0
3	F	130	0	0	2	0
3	G	166	0	0	2	0
3	H	185	0	0	3	0
3	I	156	0	0	2	0
3	J	137	0	0	4	0
3	K	112	0	0	3	0
3	L	108	0	0	0	0
All	All	15217	0	13096	243	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 10.

The worst 5 of 243 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:J:48:HIS:CD2	1:J:118:VAL:HG13	1.97	0.99
1:F:48:HIS:CD2	1:F:118:VAL:HG13	2.01	0.95
1:B:57:CYS:HG	1:B:146:CYS:HG	0.95	0.90
1:C:109:GLU:CD	1:C:109:GLU:H	1.77	0.87
1:G:144:LEU:HD23	3:G:1337:HOH:O	1.76	0.84

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	151/153 (99%)	150 (99%)	1 (1%)	0	100	100
1	B	151/153 (99%)	150 (99%)	1 (1%)	0	100	100
1	C	151/153 (99%)	147 (97%)	4 (3%)	0	100	100
1	D	151/153 (99%)	148 (98%)	3 (2%)	0	100	100
1	E	151/153 (99%)	143 (95%)	8 (5%)	0	100	100
1	F	151/153 (99%)	148 (98%)	3 (2%)	0	100	100
1	G	152/153 (99%)	147 (97%)	4 (3%)	1 (1%)	25	24
1	H	151/153 (99%)	150 (99%)	1 (1%)	0	100	100
1	I	151/153 (99%)	146 (97%)	5 (3%)	0	100	100
1	J	151/153 (99%)	143 (95%)	6 (4%)	2 (1%)	14	11
1	K	151/153 (99%)	146 (97%)	5 (3%)	0	100	100
1	L	151/153 (99%)	144 (95%)	7 (5%)	0	100	100
All	All	1813/1836 (99%)	1762 (97%)	48 (3%)	3 (0%)	51	58

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	J	25	SER
1	G	25	SER
1	J	68	SER

### 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	118/118 (100%)	112 (95%)	6 (5%)	28	33
1	B	118/118 (100%)	115 (98%)	3 (2%)	53	65
1	C	118/118 (100%)	114 (97%)	4 (3%)	42	53
1	D	118/118 (100%)	112 (95%)	6 (5%)	28	33
1	E	118/118 (100%)	111 (94%)	7 (6%)	23	26
1	F	118/118 (100%)	110 (93%)	8 (7%)	18	20
1	G	119/118 (101%)	114 (96%)	5 (4%)	34	43
1	H	118/118 (100%)	113 (96%)	5 (4%)	34	43
1	I	118/118 (100%)	114 (97%)	4 (3%)	42	53
1	J	118/118 (100%)	115 (98%)	3 (2%)	53	65
1	K	118/118 (100%)	111 (94%)	7 (6%)	23	26
1	L	118/118 (100%)	111 (94%)	7 (6%)	23	26
All	All	1417/1416 (100%)	1352 (95%)	65 (5%)	31	39

5 of 65 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	F	34	SER
1	G	59	SER
1	L	38	LEU
1	F	69	ARG
1	F	118	VAL

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 12 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	110	HIS
1	E	110	HIS
1	G	15	GLN
1	C	153	GLN
1	F	53	ASN

### 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 [i](#)

Of 12 ligands modelled in this entry, 12 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 [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	153/153 (100%)	-0.66	0 100 100	15, 22, 35, 45	0
1	B	153/153 (100%)	-0.66	0 100 100	15, 21, 33, 43	0
1	C	153/153 (100%)	-0.67	1 (0%) 87 86	16, 21, 38, 51	0
1	D	153/153 (100%)	-0.69	0 100 100	17, 23, 36, 43	0
1	E	153/153 (100%)	-0.59	0 100 100	21, 32, 49, 57	0
1	F	153/153 (100%)	-0.61	1 (0%) 87 86	19, 25, 42, 51	0
1	G	153/153 (100%)	-0.65	0 100 100	17, 23, 34, 42	0
1	H	153/153 (100%)	-0.69	0 100 100	16, 22, 34, 43	0
1	I	153/153 (100%)	-0.63	0 100 100	18, 24, 38, 56	0
1	J	153/153 (100%)	-0.61	1 (0%) 87 86	18, 26, 45, 54	0
1	K	153/153 (100%)	-0.61	0 100 100	23, 32, 45, 52	0
1	L	153/153 (100%)	-0.55	0 100 100	22, 34, 49, 64	0
All	All	1836/1836 (100%)	-0.64	3 (0%) 94 94	15, 25, 43, 64	0

All (3) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	F	1	ALA	2.4
1	J	1	ALA	2.3
1	C	1	ALA	2.1

### 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 [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	H	155	1/1	1.00	0.14	4.56	21,21,21,21	0
2	ZN	L	155	1/1	1.00	0.15	4.54	28,28,28,28	0
2	ZN	F	155	1/1	1.00	0.14	3.17	24,24,24,24	0
2	ZN	B	155	1/1	1.00	0.13	2.19	23,23,23,23	0
2	ZN	G	155	1/1	1.00	0.14	1.96	22,22,22,22	0
2	ZN	A	155	1/1	0.99	0.13	1.52	23,23,23,23	0
2	ZN	C	155	1/1	1.00	0.13	1.47	22,22,22,22	0
2	ZN	D	155	1/1	1.00	0.12	1.23	23,23,23,23	0
2	ZN	J	155	1/1	0.99	0.12	0.94	27,27,27,27	0
2	ZN	I	155	1/1	1.00	0.13	0.90	24,24,24,24	0
2	ZN	K	155	1/1	0.99	0.13	0.88	29,29,29,29	0
2	ZN	E	155	1/1	1.00	0.10	0.15	31,31,31,31	0

### 6.5 Other polymers [i](#)

There are no such residues in this entry.