



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 15, 2017 – 12:57 am GMT

PDB ID : 4TZ4  
Title : Crystal Structure of Human Cereblon in Complex with DDB1 and Lenalido-  
mide  
Authors : Chamberlain, P.P.; Pagarigan, B.; Delker, S.; Leon, B.; Riley, M.  
Deposited on : 2014-07-09  
Resolution : 3.01 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 : 1.7.2 (RC1), CSD as538be (2017)  
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

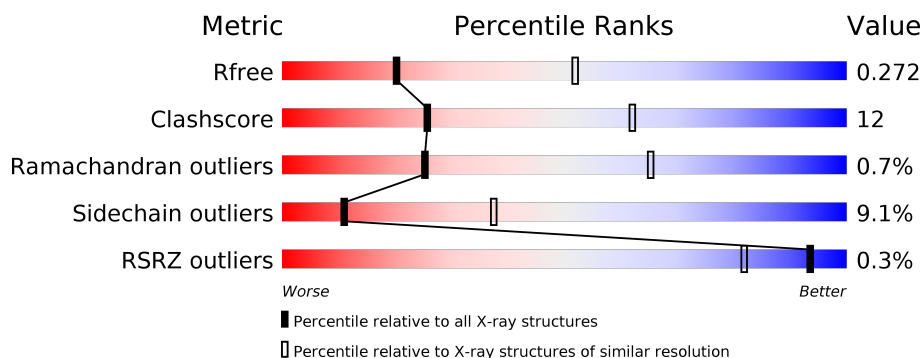
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 3.01 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	1924 (3.04-3.00)
Clashscore	112137	2279 (3.04-3.00)
Ramachandran outliers	110173	2207 (3.04-3.00)
Sidechain outliers	110143	2210 (3.04-3.00)
RSRZ outliers	101464	1948 (3.04-3.00)

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	1146	 70% 20% 7%
2	C	381	 67% 21% 6% 5%

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 11081 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 DNA damage-binding protein 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	1070	Total	C	N	O	S	0	0	0
			8191	5227	1367	1552	45			

There are 7 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1141	TRP	-	expression tag	UNP Q16531
A	1142	SER	-	expression tag	UNP Q16531
A	1143	HIS	-	expression tag	UNP Q16531
A	1144	PRO	-	expression tag	UNP Q16531
A	1145	GLN	-	expression tag	UNP Q16531
A	1146	PHE	-	expression tag	UNP Q16531
A	1147	GLU	-	expression tag	UNP Q16531

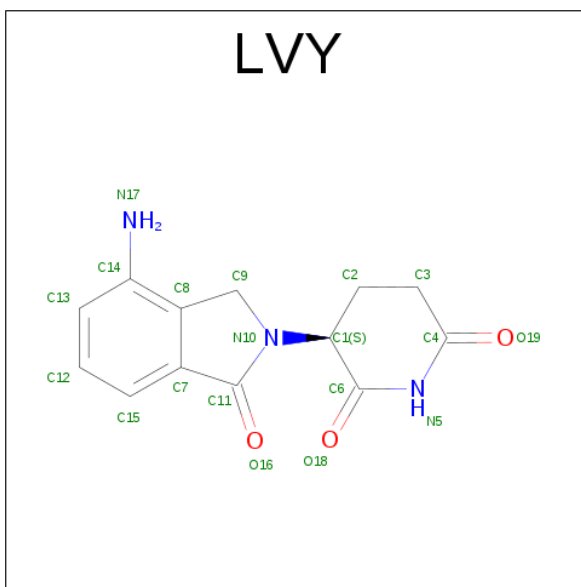
- Molecule 2 is a protein called Protein cereblon.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	361	Total	C	N	O	S	0	0	0
			2852	1832	485	513	22			

- Molecule 3 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	C	1	Total	Zn	0	0
			1	1		

- Molecule 4 is S-Lenalidomide (three-letter code: LVY) (formula: C<sub>13</sub>H<sub>13</sub>N<sub>3</sub>O<sub>3</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	C	1	Total	C	N	O	0	0
			19	13	3	3		

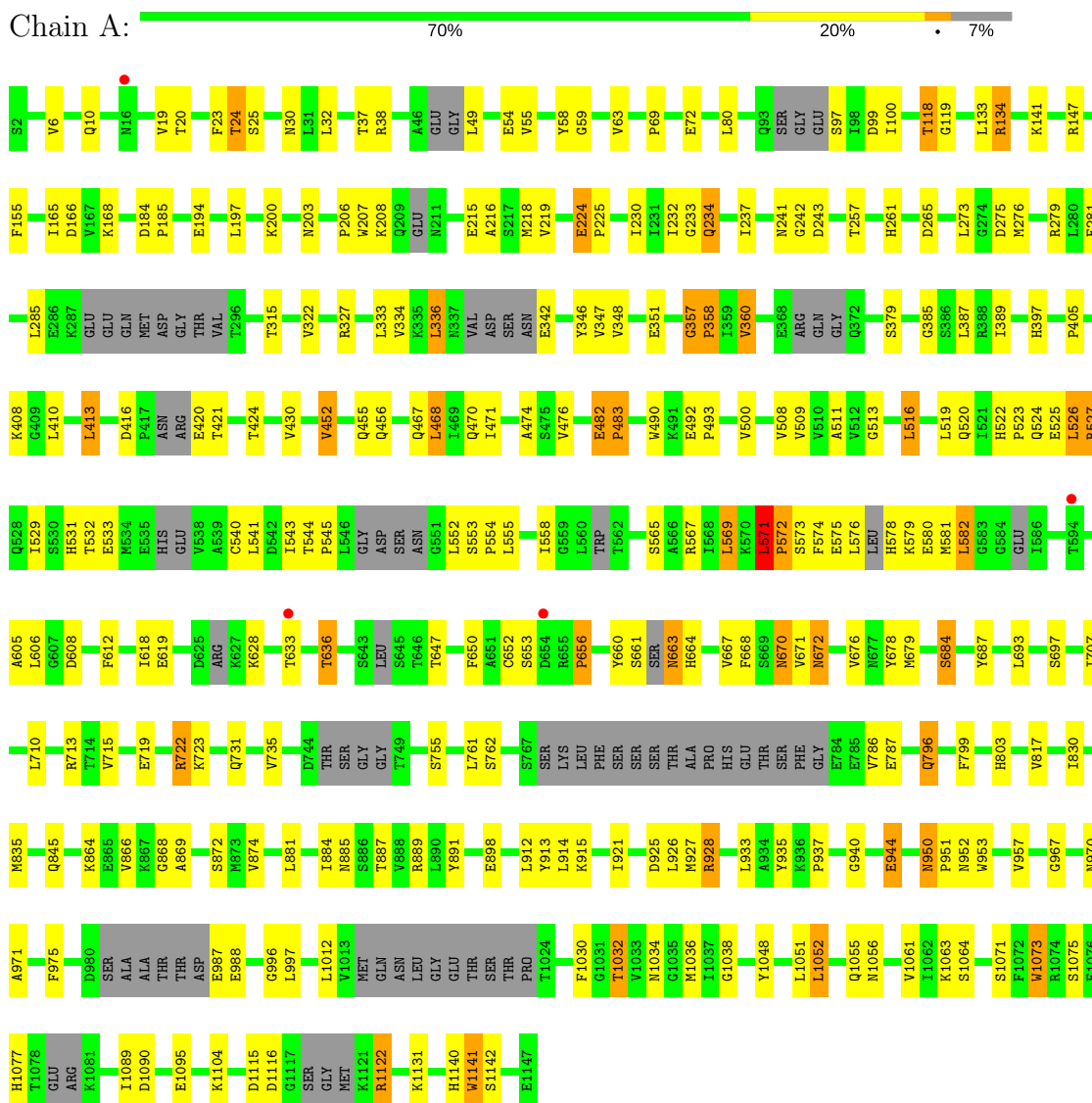
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	13	Total	O	0	0
			13	13		
5	C	5	Total	O	0	0
			5	5		

### 3 Residue-property plots

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: DNA damage-binding protein 1



- Molecule 2: Protein cereblon



S396	H397	T405	K406	K407	D408	M409	S410	P411	Q412	K413	P414	W415	L423	I426	P427	D428	E263	W264	GLU	ASN	LEU	LYS	ASP	ASP	S272	L273	R283	V284	I290	D291	D292	V293	L294	R295	L298	L299	K300	I301	C310	K317	F338	S339	L340	P345	M346	A347	V350	N351	P352	E358	T359	L360	N367	L368	N369	W380	V388	K392	I393	C394	A395
K48	F49	H57	L60	G61	A62	D63	G69	R70	T71	D75	D76	S77	M88	M89	E187	L91	I92	Q95	T96	L97	P98	L101	F102	V110	R111	N112	L113	T114	Q115	R118	T119	F120	S126	ASN	VAL	GLN	GLU	ARG	E132	Y144	R145	Q148	D149	T154	G161																
F165	L168	E169	L170	R171	T172	GLN	GLN	ASP	ASP	G176	I177	Q178	K181	V182	Q183	P186	E187	C188	T193	V197	Q198	L199	Q206	V213	SER	ARG	GLU	ASP	GLN	CYS	S220	Y221	K222	W223	F232	L237	T238	S239	K240	P241	R242	W243	L244	Y245	D249	A250	E251	Q260													

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	71.86Å 129.12Å 198.68Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.00 – 3.01 48.70 – 3.01	Depositor EDS
% Data completeness (in resolution range)	99.4 (50.00-3.01) 99.5 (48.70-3.01)	Depositor EDS
$R_{merge}$	0.18	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.21 (at 3.01Å)	Xtriage
Refinement program	REFMAC 5.6.0117	Depositor
R, $R_{free}$	0.202 , 0.271 0.202 , 0.272	Depositor DCC
$R_{free}$ test set	1857 reflections (5.25%)	DCC
Wilson B-factor (Å <sup>2</sup> )	51.7	Xtriage
Anisotropy	0.063	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 37.6	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	11081	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	51.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.57% 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: ZN, LVY

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	0.56	3/8331 (0.0%)	0.62	2/11295 (0.0%)
2	C	0.62	5/2921 (0.2%)	0.63	0/3970
All	All	0.58	8/11252 (0.1%)	0.63	2/15265 (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	5
2	C	0	1
All	All	0	6

The worst 5 of 8 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	240	TRP	CD2-CE2	6.97	1.49	1.41
2	C	243	TRP	CD2-CE2	5.61	1.48	1.41
2	C	415	TRP	CD2-CE2	5.58	1.48	1.41
2	C	264	TRP	CD2-CE2	5.33	1.47	1.41
1	A	1141	TRP	CD2-CE2	5.23	1.47	1.41

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	357	GLY	C-N-CD	-5.30	108.94	120.60
1	A	49	LEU	CA-CB-CG	5.12	127.08	115.30

There are no chirality outliers.



5 of 6 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	224	GLU	Peptide
1	A	357	GLY	Peptide
1	A	482	GLU	Peptide
1	A	571	LEU	Peptide
1	A	940	GLY	Peptide

## 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	8191	0	7989	197	0
2	C	2852	0	2808	75	0
3	C	1	0	0	0	0
4	C	19	0	13	0	0
5	A	13	0	0	0	0
5	C	5	0	0	0	0
All	All	11081	0	10810	261	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.

The worst 5 of 261 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:927:MET:O	1:A:928:ARG:HB3	1.44	1.15
1:A:565:SER:OG	1:A:580:GLU:O	1.66	1.13
2:C:295:ARG:HH11	2:C:295:ARG:HG2	1.20	1.07
1:A:660:TYR:O	1:A:667:VAL:N	1.92	1.02
1:A:20:THR:HG23	1:A:315:THR:HG21	1.45	0.96

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	1026/1146 (90%)	956 (93%)	62 (6%)	8 (1%)	22	62
2	C	351/381 (92%)	324 (92%)	26 (7%)	1 (0%)	44	80
All	All	1377/1527 (90%)	1280 (93%)	88 (6%)	9 (1%)	25	65

5 of 9 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	224	GLU
1	A	358	PRO
1	A	483	PRO
1	A	572	PRO
2	C	411	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	882/1005 (88%)	817 (93%)	65 (7%)	16	48
2	C	310/345 (90%)	266 (86%)	44 (14%)	4	17
All	All	1192/1350 (88%)	1083 (91%)	109 (9%)	11	37

5 of 109 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	874	VAL

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	1116	ASP
2	C	340	LEU
1	A	898	GLU
1	A	1052	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 26 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	670	ASN
1	A	950	ASN
2	C	367	ASN
1	A	672	ASN
1	A	677	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 2 ligands modelled in this entry, 1 is 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
4	LVY	C	502	-	20,21,21	1.71	4 (20%)	23,31,31	2.81	8 (34%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	LVY	C	502	-	-	0/4/29/29	0/3/3/3

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	C	502	LVY	C7-C11	-5.63	1.39	1.48
4	C	502	LVY	C11-N10	-2.50	1.33	1.36
4	C	502	LVY	C6-N5	2.07	1.40	1.37
4	C	502	LVY	C4-N5	2.36	1.41	1.37

The worst 5 of 8 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	C	502	LVY	C9-N10-C11	-6.42	110.79	113.04
4	C	502	LVY	C4-N5-C6	-3.54	121.18	126.54
4	C	502	LVY	O16-C11-C7	-2.32	124.12	128.68
4	C	502	LVY	C2-C3-C4	-2.22	110.50	114.31
4	C	502	LVY	C15-C7-C11	2.13	132.96	129.65

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 [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	1070/1146 (93%)	-0.39	4 (0%) 92 77	21, 44, 108, 136	5 (0%)
2	C	361/381 (94%)	-0.33	0 100 100	30, 50, 78, 100	0
All	All	1431/1527 (93%)	-0.37	4 (0%) 93 82	21, 46, 102, 136	5 (0%)

All (4) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	594	THR	2.4
1	A	633	THR	2.3
1	A	654	ASP	2.3
1	A	16	ASN	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 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( $\text{\AA}^2$ )	Q<0.9
4	LVY	C	502	19/19	0.97	0.19	0.34	35,37,39,39	0
3	ZN	C	501	1/1	0.99	0.09	-3.26	45,45,45,45	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.