



Full wwPDB X-ray Structure Validation Report ⓘ

May 29, 2020 – 08:30 am BST

PDB ID : 5V3M
Title : mouseZFP568-ZnF1-11 in complex with DNA
Authors : Patel, A.; Cheng, X.
Deposited on : 2017-03-07
Resolution : 2.09 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at

<https://www.wwpdb.org/validation/2017/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
Xtrriage (Phenix) : 1.13
EDS : 2.11
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac : 5.8.0158
CCP4 : 7.0.044 (Gargrove)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.11

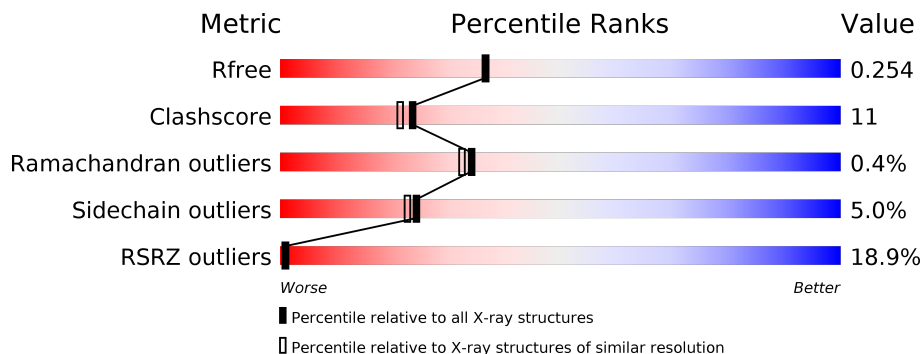
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.09 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}	130704	5197 (2.10-2.10)
Clashscore	141614	5710 (2.10-2.10)
Ramachandran outliers	138981	5647 (2.10-2.10)
Sidechain outliers	138945	5648 (2.10-2.10)
RSRZ outliers	127900	5083 (2.10-2.10)

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 ≥ 3 , 2, 1 and 0 types of geometric quality criteria respectively. 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	28	 4% 54% 46%
2	B	28	 4% 71% 29%
3	C	314	 19% 66% 20% 12%

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
4	ZN	C	1004	-	-	X	-

2 Entry composition i

There are 5 unique types of molecules in this entry. The entry contains 3360 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 DNA chain called DNA (28-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
1	A	28	585	276	120	162	27	0	0	0

- Molecule 2 is a DNA chain called DNA (28-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
2	B	28	557	268	92	170	27	0	0	0

- Molecule 3 is a protein called Zinc finger protein 568.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	C	277	2069	1266	414	364	25	0	1	0

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	358	GLY	-	expression tag	UNP E9PYI1
C	359	PRO	-	expression tag	UNP E9PYI1
C	360	GLY	-	expression tag	UNP E9PYI1
C	361	SER	-	expression tag	UNP E9PYI1
C	670	LEU	-	expression tag	UNP E9PYI1
C	671	GLU	-	expression tag	UNP E9PYI1

- Molecule 4 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	C	10	Total	Zn	0	0
			10	10		

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	34	Total 34	O 34	0	0
5	B	33	Total 33	O 33	0	0
5	C	72	Total 72	O 72	0	0

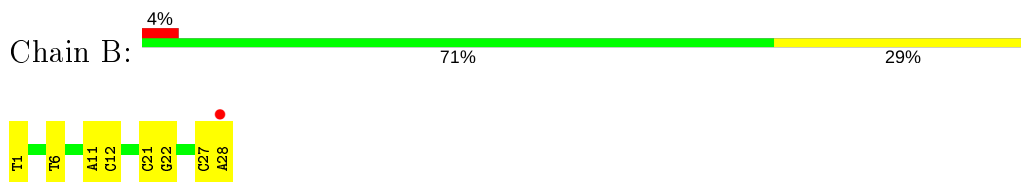
3 Residue-property plots [i](#)

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.

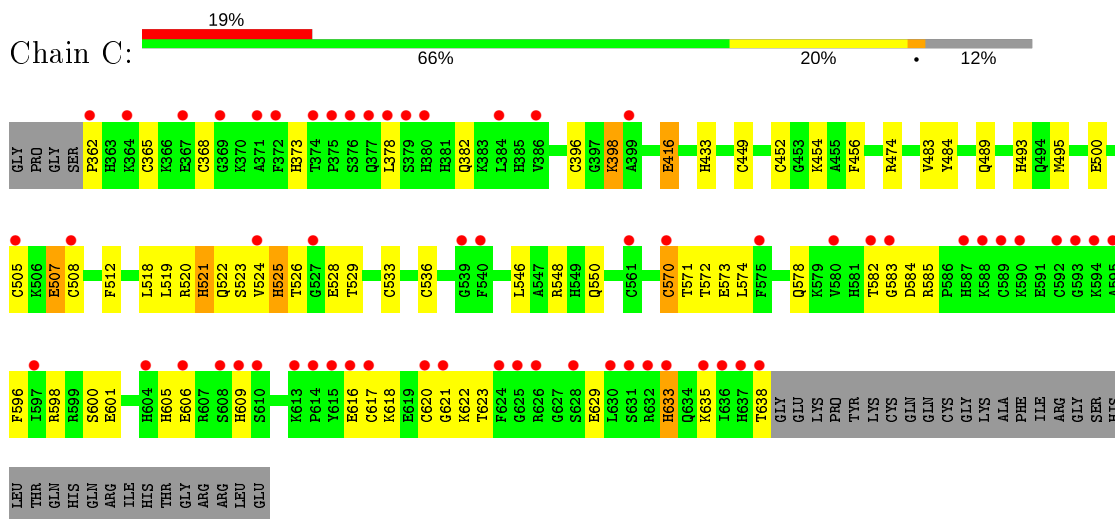
- Molecule 1: DNA (28-MER)



- Molecule 2: DNA (28-MER)



- Molecule 3: Zinc finger protein 568



4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α , β , γ	35.95Å 92.76Å 353.67Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	32.24 – 2.09 34.17 – 2.09	Depositor EDS
% Data completeness (in resolution range)	78.8 (32.24-2.09) 76.0 (34.17-2.09)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.17 (at 2.10Å)	Xtrriage
Refinement program	PHENIX (1.10.1_2155: ???)	Depositor
R, R_{free}	0.213 , 0.253 0.215 , 0.254	Depositor DCC
R_{free} test set	1432 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å ²)	36.5	Xtrriage
Anisotropy	0.951	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.27 , 55.2	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	3360	wwPDB-VP
Average B, all atoms (Å ²)	73.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.77% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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	1.07	1/660 (0.2%)	1.04	0/1020
2	B	0.95	0/620	1.10	3/952 (0.3%)
3	C	0.52	0/2125	0.68	2/2856 (0.1%)
All	All	0.75	1/3405 (0.0%)	0.86	5/4828 (0.1%)

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
3	C	0	1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	20	DA	C3'-O3'	-7.37	1.34	1.44

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed($^{\circ}$)	Ideal($^{\circ}$)
3	C	621	GLY	N-CA-C	-5.51	99.32	113.10
2	B	6	DT	C5-C4-O4	-5.17	121.28	124.90
2	B	1	DT	N3-C4-O4	5.09	122.95	119.90
2	B	6	DT	N3-C4-O4	5.09	122.95	119.90
3	C	508	CYS	CA-CB-SG	5.07	123.12	114.00

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
3	C	507	GLU	Mainchain

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	585	0	314	11	0
2	B	557	0	318	4	0
3	C	2069	0	1855	49	0
4	C	10	0	0	2	0
5	A	34	0	0	3	0
5	B	33	0	0	0	0
5	C	72	0	0	1	0
All	All	3360	0	2487	62	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 11.

All (62) 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)
3:C:507:GLU:HB2	3:C:525:HIS:HD2	1.31	0.94
3:C:507:GLU:HB2	3:C:525:HIS:CD2	2.10	0.85
3:C:452:CYS:SG	4:C:1004:ZN:ZN	1.70	0.81
3:C:616:GLU:CB	3:C:623:THR:HG22	2.15	0.77
1:A:5:DG:OP2	1:A:5:DG:H8	1.68	0.76
3:C:598:ARG:H	3:C:601:GLU:HG3	1.53	0.74
5:A:113:HOH:O	3:C:433:HIS:HD2	1.71	0.73
3:C:378:LEU:HG	3:C:382:GLN:HE21	1.55	0.72
1:A:21:DA:OP1	5:A:101:HOH:O	2.07	0.70
1:A:12:DC:OP2	5:A:102:HOH:O	2.12	0.66
3:C:524:VAL:HG12	3:C:525:HIS:HD1	1.61	0.64
2:B:21:DC:H2'	2:B:22:DG:C8	2.33	0.63
1:A:7:DC:H2''	1:A:8:DG:C8	2.34	0.63
3:C:635:LYS:O	3:C:638:THR:OG1	2.20	0.60
3:C:484:TYR:HB3	3:C:489:GLN:HE21	1.66	0.59
3:C:617:CYS:HB3	3:C:620:CYS:HB2	1.83	0.59

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:C:533:CYS:HB3	3:C:536:CYS:SG	2.42	0.58
3:C:362:PRO:HA	3:C:373:HIS:HA	1.87	0.57
1:A:17:DG:H2'	1:A:18:DT:C6	2.41	0.56
3:C:474[B]:ARG:HG2	3:C:483:VAL:HG21	1.88	0.55
3:C:507:GLU:CB	3:C:525:HIS:CD2	2.85	0.55
3:C:378:LEU:HG	3:C:382:GLN:NE2	2.21	0.54
3:C:574:LEU:O	3:C:578:GLN:HG3	2.09	0.53
3:C:493:HIS:CD2	3:C:493:HIS:C	2.84	0.52
2:B:28:DA:H8	2:B:28:DA:H5''	1.76	0.51
3:C:606:GLU:HA	3:C:609:HIS:HD2	1.74	0.51
3:C:620:CYS:C	3:C:622:LYS:H	2.14	0.51
1:A:5:DG:C8	1:A:5:DG:OP2	2.58	0.51
1:A:3:DT:H1'	1:A:4:DG:H5'	1.92	0.50
3:C:617:CYS:SG	3:C:618:LYS:N	2.85	0.50
3:C:570:CYS:SG	3:C:572:THR:HG22	2.52	0.49
3:C:452:CYS:HG	4:C:1004:ZN:ZN	1.26	0.49
3:C:454:LYS:HG2	3:C:456:PHE:CZ	2.48	0.48
3:C:582:THR:O	3:C:584:ASP:N	2.41	0.48
2:B:11:DA:C2	2:B:12:DC:C2	3.02	0.47
3:C:484:TYR:HH	3:C:493:HIS:HD1	1.62	0.47
3:C:571:THR:HG22	5:C:1134:HOH:O	2.14	0.47
3:C:526:THR:HG23	3:C:528:GLU:H	1.81	0.46
3:C:605:HIS:CE1	3:C:609:HIS:CE1	3.03	0.45
1:A:7:DC:H5	3:C:629:GLU:OE2	1.99	0.45
3:C:505:CYS:C	3:C:507:GLU:N	2.68	0.45
3:C:616:GLU:HA	3:C:623:THR:HA	1.99	0.45
1:A:12:DC:H2''	1:A:13:DA:C8	2.52	0.44
3:C:629:GLU:OE1	3:C:629:GLU:N	2.51	0.44
3:C:548:ARG:HD3	3:C:570:CYS:SG	2.58	0.44
3:C:449:CYS:SG	3:C:452:CYS:SG	3.02	0.44
3:C:585:ARG:HB3	3:C:596:PHE:O	2.19	0.43
2:B:27:DC:H4'	2:B:28:DA:OP2	2.18	0.43
3:C:365:CYS:HB3	3:C:368:CYS:SG	2.58	0.43
3:C:570:CYS:HB3	3:C:573:GLU:H	1.83	0.43
3:C:518:LEU:O	3:C:522:GLN:HG2	2.18	0.43
3:C:495:MET:HE3	3:C:500:GLU:HB3	2.01	0.42
1:A:5:DG:H2''	1:A:6:DG:C8	2.54	0.42
3:C:512:PHE:CZ	3:C:521:HIS:CG	3.07	0.42
3:C:546:LEU:O	3:C:550:GLN:HG3	2.19	0.41
3:C:523:SER:HA	3:C:526:THR:HG22	2.02	0.41
3:C:396:CYS:SG	3:C:398:LYS:HB2	2.60	0.41

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:C:416:GLU:H	3:C:416:GLU:HG2	1.41	0.41
3:C:484:TYR:HB3	3:C:489:GLN:NE2	2.35	0.41
1:A:25:DG:O6	3:C:433:HIS:HE1	2.04	0.41
3:C:522:GLN:O	3:C:526:THR:HG22	2.22	0.40
3:C:633:HIS:O	3:C:633:HIS:ND1	2.54	0.40

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
3	C	276/314 (88%)	259 (94%)	16 (6%)	1 (0%)	34 32

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	C	583	GLY

5.3.2 Protein sidechains [i](#)

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
3	C	202/273 (74%)	192 (95%)	10 (5%)	24 23

All (10) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	C	398	LYS
3	C	416	GLU
3	C	519	LEU
3	C	520	ARG
3	C	521	HIS
3	C	525	HIS
3	C	529	THR
3	C	570	CYS
3	C	600	SER
3	C	633	HIS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
3	C	382	GLN
3	C	433	HIS
3	C	489	GLN
3	C	522	GLN
3	C	550	GLN

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 10 ligands modelled in this entry, 10 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, 95th 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 ²)	Q<0.9
1	A	28/28 (100%)	0.41	1 (3%) 42 49	27, 60, 128, 145	0
2	B	28/28 (100%)	0.36	1 (3%) 42 49	27, 71, 157, 161	0
3	C	277/314 (88%)	1.11	61 (22%) 0 0	23, 68, 137, 172	0
All	All	333/370 (90%)	0.99	63 (18%) 1 1	23, 68, 138, 172	0

All (63) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	C	625	GLY	8.4
3	C	615	TYR	8.3
3	C	589	CYS	7.4
3	C	635	LYS	7.0
3	C	593	GLY	6.5
3	C	636	ILE	6.4
3	C	633	HIS	6.0
3	C	606	GLU	5.8
3	C	539	GLY	5.7
3	C	617	CYS	5.6
3	C	608	SER	4.8
3	C	592	CYS	4.8
3	C	371	ALA	4.5
3	C	620	CYS	4.4
3	C	590	LYS	4.3
3	C	628	SER	4.3
3	C	587	HIS	4.1
3	C	372	PHE	4.0
3	C	575	PHE	4.0
3	C	379	SER	3.9
3	C	376	SER	3.8
3	C	377	GLN	3.8
3	C	638	THR	3.8

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
3	C	378	LEU	3.6
3	C	624	PHE	3.5
3	C	384	LEU	3.4
3	C	508	CYS	3.3
3	C	380	HIS	3.3
3	C	621	GLY	3.3
3	C	588	LYS	3.2
3	C	610	SER	3.2
3	C	630	LEU	3.2
3	C	582	THR	3.1
3	C	631	SER	3.1
3	C	613	LYS	3.0
3	C	609	HIS	3.0
3	C	597	ILE	3.0
3	C	561	CYS	2.9
1	A	1	DT	2.9
3	C	374	THR	2.9
3	C	362	PRO	2.9
3	C	369	GLY	2.9
3	C	527	GLY	2.9
3	C	375	PRO	2.7
3	C	614	PRO	2.6
3	C	632	ARG	2.6
3	C	399	ALA	2.6
3	C	626	ARG	2.5
3	C	604	HIS	2.5
3	C	580	VAL	2.4
2	B	28	DA	2.3
3	C	583	GLY	2.3
3	C	540	PHE	2.3
3	C	616	GLU	2.3
3	C	505	CYS	2.2
3	C	595	ALA	2.2
3	C	570	CYS	2.1
3	C	364	LYS	2.1
3	C	367	GLU	2.1
3	C	386	VAL	2.1
3	C	594	LYS	2.1
3	C	637	HIS	2.1
3	C	524	VAL	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. The B-factors column lists the minimum, median, 95th 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	B-factors(\AA^2)	Q<0.9
4	ZN	C	1010	1/1	0.28	0.25	169,169,169,169	0
4	ZN	C	1006	1/1	0.53	0.12	72,72,72,72	0
4	ZN	C	1008	1/1	0.71	0.10	79,79,79,79	0
4	ZN	C	1009	1/1	0.78	0.19	160,160,160,160	0
4	ZN	C	1004	1/1	0.82	0.07	61,61,61,61	0
4	ZN	C	1005	1/1	0.94	0.07	56,56,56,56	0
4	ZN	C	1007	1/1	0.96	0.05	63,63,63,63	0
4	ZN	C	1001	1/1	0.98	0.05	76,76,76,76	0
4	ZN	C	1002	1/1	0.99	0.10	47,47,47,47	0
4	ZN	C	1003	1/1	1.00	0.13	33,33,33,33	0

6.5 Other polymers [i](#)

There are no such residues in this entry.