Full wwPDB X-ray Structure Validation Report ⓘ

May 28, 2020 – 10:04 pm BST

PDB ID	:	2J9U
Title	:	2 Angstrom X-ray structure of the yeast ESCRT-I Vps28 C-terminus in complex with the NZF-N domain from ESCRT-II
Authors	:	Gill, D.J.; Teo, H.L.; Sun, J.; Perisic, O.; Veprintsev, D.B.; Emr, S.D.; Williams, R.L.
Deposited on	:	2006-11-16
Resolution	:	2.00 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 ①) were used in the production of this report:

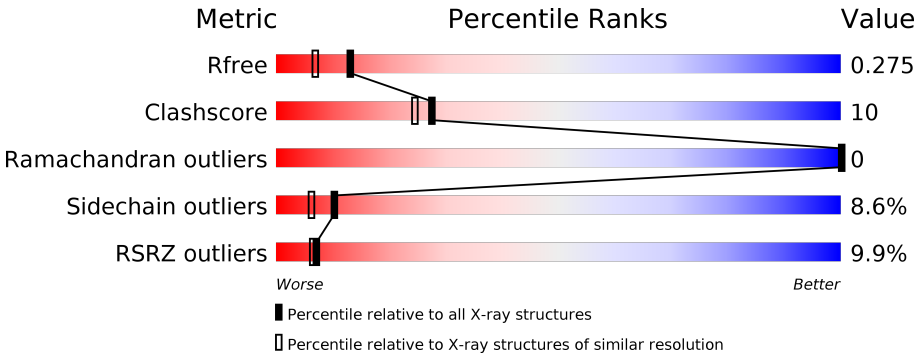
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.00 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	8085 (2.00-2.00)
Clashscore	141614	9178 (2.00-2.00)
Ramachandran outliers	138981	9054 (2.00-2.00)
Sidechain outliers	138945	9053 (2.00-2.00)
RSRZ outliers	127900	7900 (2.00-2.00)

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 >=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 <=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	96	7% 79% 16%
1	C	96	7% 82% 14%
2	B	76	11% 43% 12% 7% 38%
2	D	76	8% 43% 14% 38%

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 2295 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 VACUOLAR PROTEIN SORTING-ASSOCIATED PROTEIN 28.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	94	Total	C	N	O	S	0	0	0
			769	498	127	143	1			
1	C	94	Total	C	N	O	S	0	0	0
			769	498	127	143	1			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	147	MET	-	expression tag	UNP Q02767
C	147	MET	-	expression tag	UNP Q02767

- Molecule 2 is a protein called VACUOLAR PROTEIN SORTING-ASSOCIATED PROTEIN 36.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	47	Total	C	N	O	S	0	0	0
			352	220	54	72	6			
2	D	47	Total	C	N	O	S	0	0	0
			352	220	54	72	6			

There are 28 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	101	MET	-	expression tag	UNP Q06696
B	102	ALA	-	expression tag	UNP Q06696
B	103	HIS	-	expression tag	UNP Q06696
B	104	HIS	-	expression tag	UNP Q06696
B	105	HIS	-	expression tag	UNP Q06696
B	106	HIS	-	expression tag	UNP Q06696
B	107	HIS	-	expression tag	UNP Q06696
B	108	HIS	-	expression tag	UNP Q06696

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	109	MET	-	expression tag	UNP Q06696
B	172	ARG	-	expression tag	UNP Q06696
B	173	ASN	-	expression tag	UNP Q06696
B	174	GLN	-	expression tag	UNP Q06696
B	175	PHE	-	expression tag	UNP Q06696
B	176	GLY	-	expression tag	UNP Q06696
D	101	MET	-	expression tag	UNP Q06696
D	102	ALA	-	expression tag	UNP Q06696
D	103	HIS	-	expression tag	UNP Q06696
D	104	HIS	-	expression tag	UNP Q06696
D	105	HIS	-	expression tag	UNP Q06696
D	106	HIS	-	expression tag	UNP Q06696
D	107	HIS	-	expression tag	UNP Q06696
D	108	HIS	-	expression tag	UNP Q06696
D	109	MET	-	expression tag	UNP Q06696
D	172	ARG	-	expression tag	UNP Q06696
D	173	ASN	-	expression tag	UNP Q06696
D	174	GLN	-	expression tag	UNP Q06696
D	175	PHE	-	expression tag	UNP Q06696
D	176	GLY	-	expression tag	UNP Q06696

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

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	B	1	Total Zn 1 1	0	0
3	D	1	Total Zn 1 1	0	0

- Molecule 4 is water.

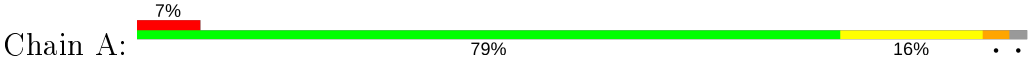
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	13	Total O 13 13	0	0
4	B	10	Total O 10 10	0	0
4	C	18	Total O 18 18	0	0
4	D	10	Total O 10 10	0	0

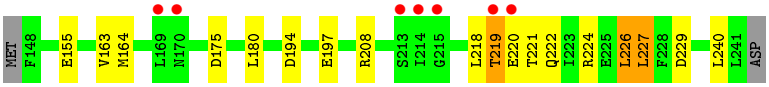
3 Residue-property plots (i)

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: VACUOLAR PROTEIN SORTING-ASSOCIATED PROTEIN 28

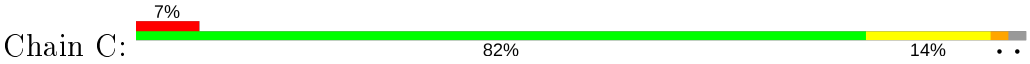
Chain A:

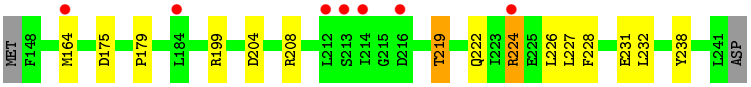




• Molecule 1: VACUOLAR PROTEIN SORTING-ASSOCIATED PROTEIN 28

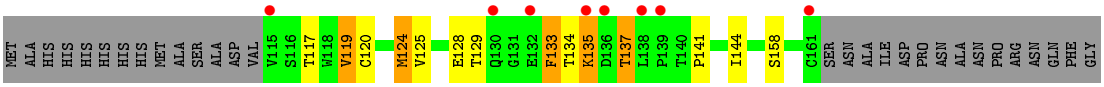
Chain C:





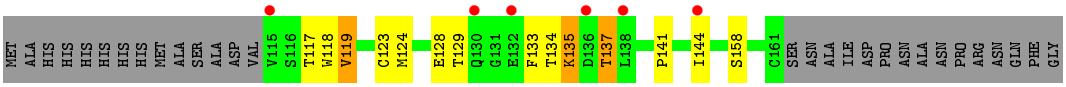
• Molecule 2: VACUOLAR PROTEIN SORTING-ASSOCIATED PROTEIN 36

Chain B:



• Molecule 2: VACUOLAR PROTEIN SORTING-ASSOCIATED PROTEIN 36

Chain D:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α, β, γ	66.64Å 99.96Å 115.62Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	32.03 – 2.00 32.02 – 2.00	Depositor EDS
% Data completeness (in resolution range)	99.8 (32.03-2.00) 99.2 (32.02-2.00)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.33 (at 2.00Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.223 , 0.260 0.251 , 0.275	Depositor DCC
R_{free} test set	1305 reflections (4.97%)	wwPDB-VP
Wilson B-factor (Å^2)	30.3	Xtriage
Anisotropy	0.220	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.37 , 49.9	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.91	EDS
Total number of atoms	2295	wwPDB-VP
Average B, all atoms (Å^2)	23.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.02% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ 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

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	1.02	0/781	0.97	7/1057 (0.7%)
1	C	1.02	0/781	0.86	3/1057 (0.3%)
2	B	0.92	1/359 (0.3%)	0.79	0/492
2	D	0.86	0/359	0.74	0/492
All	All	0.98	1/2280 (0.0%)	0.87	10/3098 (0.3%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	133	PHE	CE2-CZ	7.13	1.50	1.37

All (10) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	208	ARG	NE-CZ-NH2	-8.63	115.98	120.30
1	A	208	ARG	NE-CZ-NH1	8.03	124.31	120.30
1	A	226	LEU	CB-CG-CD2	6.76	122.49	111.00
1	C	208	ARG	NE-CZ-NH1	-5.72	117.44	120.30
1	C	208	ARG	NE-CZ-NH2	5.71	123.16	120.30
1	A	194	ASP	CB-CG-OD2	5.36	123.12	118.30
1	A	227	LEU	CB-CG-CD2	5.34	120.07	111.00
1	C	199	ARG	NE-CZ-NH2	-5.33	117.64	120.30
1	A	229	ASP	CB-CG-OD1	5.32	123.09	118.30
1	A	240	LEU	CA-CB-CG	5.29	127.48	115.30

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	769	0	786	7	1
1	C	769	0	786	7	1
2	B	352	0	338	16	0
2	D	352	0	338	16	0
3	B	1	0	0	0	0
3	D	1	0	0	0	0
4	A	13	0	0	3	0
4	B	10	0	0	0	0
4	C	18	0	0	1	0
4	D	10	0	0	0	0
All	All	2295	0	2248	45	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 10.

All (45) 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:219:THR:HG22	1:A:222:GLN:H	1.29	0.96
1:C:219:THR:HG22	1:C:222:GLN:H	1.36	0.91
2:D:134:THR:H	2:D:137:THR:CG2	1.86	0.88
2:B:134:THR:H	2:B:137:THR:CG2	1.87	0.87
1:A:164:MET:CE	4:A:2010:HOH:O	2.24	0.84
1:A:164:MET:HE3	4:A:2010:HOH:O	1.86	0.76
2:B:135:LYS:HD3	2:B:135:LYS:H	1.49	0.76
2:B:119:VAL:CG1	2:B:158:SER:HB2	2.22	0.69
1:C:204:ASP:OD1	4:C:2017:HOH:O	2.12	0.68
2:B:135:LYS:CD	2:B:135:LYS:H	2.08	0.67
2:D:135:LYS:H	2:D:135:LYS:HD3	1.61	0.65
2:D:134:THR:H	2:D:137:THR:HG22	1.61	0.64
2:D:119:VAL:CG1	2:D:158:SER:HB2	2.30	0.62
2:D:133:PHE:HA	2:D:137:THR:HG21	1.80	0.62
2:D:135:LYS:H	2:D:135:LYS:CD	2.12	0.62
2:D:129:THR:HG21	2:D:141:PRO:HB3	1.83	0.60
2:B:119:VAL:HG12	2:B:158:SER:HB2	1.85	0.59

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:D:129:THR:HG21	2:D:141:PRO:CB	2.33	0.58
2:D:118:TRP:HE1	2:D:129:THR:HG23	1.69	0.58
2:B:120:CYS:O	2:B:124:MET:HE3	2.03	0.57
2:B:133:PHE:HA	2:B:137:THR:HG21	1.86	0.57
2:B:134:THR:O	2:B:137:THR:HG23	2.05	0.57
1:A:220:GLU:O	4:A:2009:HOH:O	2.17	0.57
2:B:119:VAL:HG13	2:B:158:SER:CB	2.37	0.55
2:B:117:THR:HG22	2:B:128:GLU:HG2	1.88	0.54
2:D:119:VAL:HG13	2:D:158:SER:HB2	1.90	0.54
1:C:164:MET:HE1	1:C:231:GLU:HG2	1.88	0.54
2:D:119:VAL:HG13	2:D:158:SER:CB	2.38	0.52
2:B:135:LYS:HD3	2:B:135:LYS:N	2.23	0.51
1:C:219:THR:HG22	1:C:222:GLN:N	2.17	0.50
1:A:163:VAL:HG22	1:A:180:LEU:HB3	1.93	0.49
2:B:134:THR:H	2:B:137:THR:HG22	1.73	0.49
2:D:134:THR:O	2:D:137:THR:HG23	2.15	0.46
2:B:129:THR:HG21	2:B:141:PRO:CB	2.46	0.46
2:B:129:THR:HG21	2:B:141:PRO:HB3	1.99	0.45
1:A:197:GLU:HG3	1:C:179:PRO:HG2	1.99	0.45
1:C:224:ARG:NH2	1:C:228:PHE:HB2	2.32	0.45
2:B:119:VAL:HG13	2:B:158:SER:HB2	1.94	0.44
2:D:134:THR:N	2:D:137:THR:CG2	2.68	0.43
2:D:117:THR:HG22	2:D:128:GLU:HG2	2.01	0.43
1:C:164:MET:CE	1:C:231:GLU:HG2	2.49	0.42
1:A:219:THR:HG23	1:A:221:THR:H	1.85	0.42
2:B:134:THR:N	2:B:137:THR:CG2	2.69	0.42
2:D:119:VAL:HG12	2:D:158:SER:HB2	2.00	0.42
2:D:123:CYS:O	2:D:124:MET:HB2	2.21	0.41

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)
1:A:155:GLU:OE1	1:C:238:TYR:OH[4_555]	1.78	0.42

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	
1	A	92/96 (96%)	89 (97%)	3 (3%)	0	100	100
1	C	92/96 (96%)	90 (98%)	2 (2%)	0	100	100
2	B	45/76 (59%)	44 (98%)	1 (2%)	0	100	100
2	D	45/76 (59%)	44 (98%)	1 (2%)	0	100	100
All	All	274/344 (80%)	267 (97%)	7 (3%)	0	100	100

There are no Ramachandran outliers to report.

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	
1	A	84/86 (98%)	78 (93%)	6 (7%)	14	10
1	C	84/86 (98%)	78 (93%)	6 (7%)	14	10
2	B	44/67 (66%)	38 (86%)	6 (14%)	3	2
2	D	44/67 (66%)	40 (91%)	4 (9%)	9	5
All	All	256/306 (84%)	234 (91%)	22 (9%)	10	6

All (22) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	175	ASP
1	A	218	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	219	THR
1	A	224	ARG
1	A	226	LEU
1	A	227	LEU
2	B	119	VAL
2	B	124	MET
2	B	125	VAL
2	B	135	LYS
2	B	137	THR
2	B	144	ILE
1	C	175	ASP
1	C	219	THR
1	C	224	ARG
1	C	226	LEU
1	C	227	LEU
1	C	232	LEU
2	D	119	VAL
2	D	135	LYS
2	D	137	THR
2	D	144	ILE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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, 2 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

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^{th} 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	94/96 (97%)	0.59	7 (7%)	14	13	11, 18, 28, 34	0
1	C	94/96 (97%)	0.63	7 (7%)	14	13	11, 18, 28, 34	0
2	B	47/76 (61%)	1.09	8 (17%)	1	1	21, 29, 45, 47	0
2	D	47/76 (61%)	0.91	6 (12%)	3	3	22, 29, 45, 47	0
All	All	282/344 (81%)	0.74	28 (9%)	7	6	11, 22, 41, 47	0

All (28) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	161	CYS	5.7
1	A	214	ILE	5.7
1	A	220	GLU	4.7
1	C	214	ILE	4.3
2	B	130	GLN	3.9
2	B	132	GLU	3.4
2	D	138	LEU	3.4
2	D	130	GLN	3.3
1	C	213	SER	3.3
1	C	216	ASP	3.2
1	C	212	LEU	3.1
1	C	224	ARG	3.1
2	D	132	GLU	3.0
2	B	136	ASP	2.8
2	B	138	LEU	2.7
1	A	170	ASN	2.7
2	D	115	VAL	2.6
2	B	115	VAL	2.6
2	B	135	LYS	2.6
1	A	215	GLY	2.6
1	A	219	THR	2.5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	D	144	ILE	2.4
2	D	136	ASP	2.4
1	A	213	SER	2.3
2	B	139	PRO	2.3
1	C	184	LEU	2.2
1	C	164	MET	2.2
1	A	169	LEU	2.1

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. The B-factors column lists the minimum, median, 95^{th} 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(Å^2)	Q<0.9
3	ZN	D	1162	1/1	0.99	0.08	19,19,19,19	0
3	ZN	B	1162	1/1	0.99	0.11	19,19,19,19	0

6.5 Other polymers (i)

There are no such residues in this entry.