Full wwPDB X-ray Structure Validation Report (i)

Feb 14, 2017 – 12:06 am GMT

PDB ID	:	3PHD
Title	:	Crystal structure of human HDAC6 in complex with ubiquitin
Authors	:	Dong, A.; Qui, W.; Ravichandran, M.; Schuetz, A.; Loppnau, P.; Li, F.; Mackenzie, F.; Kozieradzki, I.; Ouyang, H.; Structural Genomics Consortium (SGC)
Deposited on	:	2010-11-03
Resolution	:	3.00 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
http://wwpdb.org/validation/2016/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

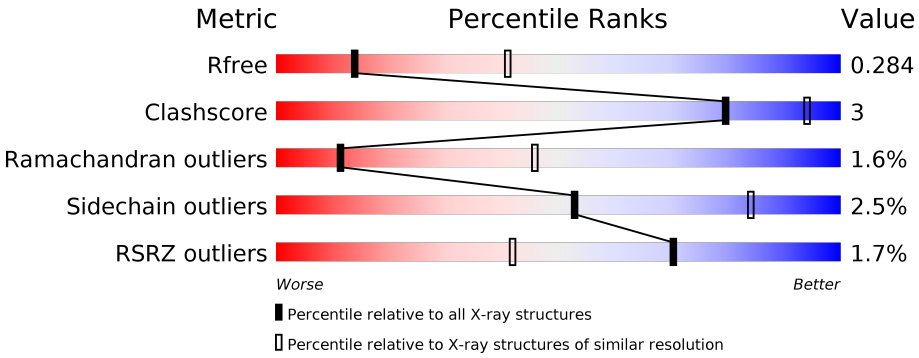
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.9-1692
EDS	:	trunk28620
Percentile statistics	:	20161228.v01 (using entries in the PDB archive December 28th 2016)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	recalc28949

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 3.00 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.



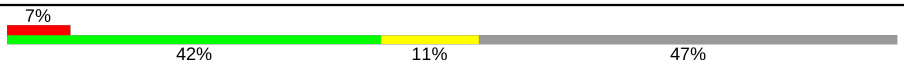
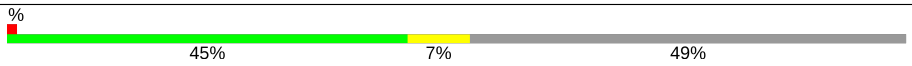
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	100719	1692 (3.00-3.00)
Clashscore	112137	2037 (3.00-3.00)
Ramachandran outliers	110173	1973 (3.00-3.00)
Sidechain outliers	110143	1976 (3.00-3.00)
RSRZ outliers	101464	1716 (3.00-3.00)

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 >=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 <=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	107	91% 5% 5%
1	B	107	2% 91% 7%
1	C	107	69% 10% 21%
1	D	107	70% 8% 21%
2	E	76	18% 82%
2	F	76	7% 93%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	G	76	
2	H	76	

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 3321 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 Histone deacetylase 6.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	102	Total	C	N	O	S	0	0	0
			761	482	135	134	10			
1	B	100	Total	C	N	O	S	0	0	0
			758	481	133	134	10			
1	C	85	Total	C	N	O	S	0	0	0
			629	396	111	112	10			
1	D	84	Total	C	N	O	S	0	0	0
			623	392	110	111	10			

- Molecule 2 is a protein called Polyubiquitin.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	E	14	Total	C	N	O	0	0	0
			74	43	16	15			
2	F	5	Total	C	N	O	0	0	0
			30	16	8	6			
2	G	40	Total	C	N	O	0	0	0
			223	134	46	43			
2	H	39	Total	C	N	O	0	0	0
			209	126	42	41			

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

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

- Molecule 4 is water.

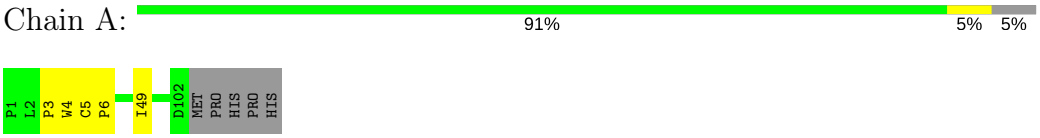
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	B	1	Total O 1 1	0	0
4	G	1	Total O 1 1	0	0

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.

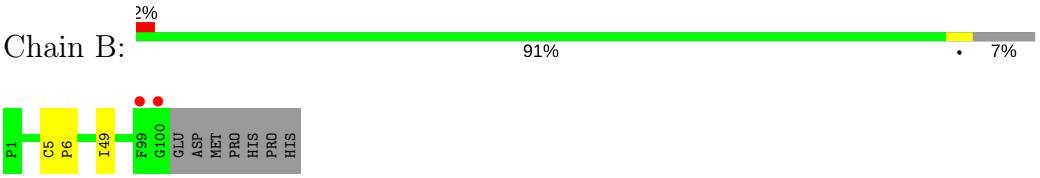
• Molecule 1: Histone deacetylase 6

Chain A:



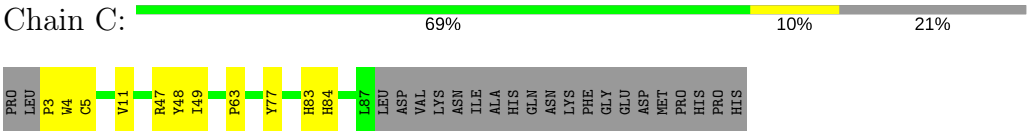
• Molecule 1: Histone deacetylase 6

Chain B:



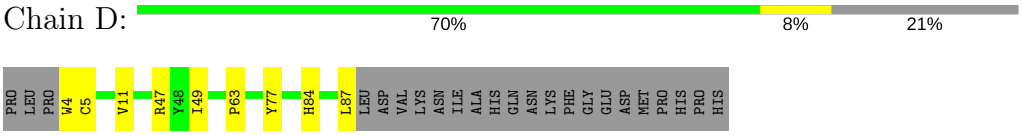
• Molecule 1: Histone deacetylase 6

Chain C:



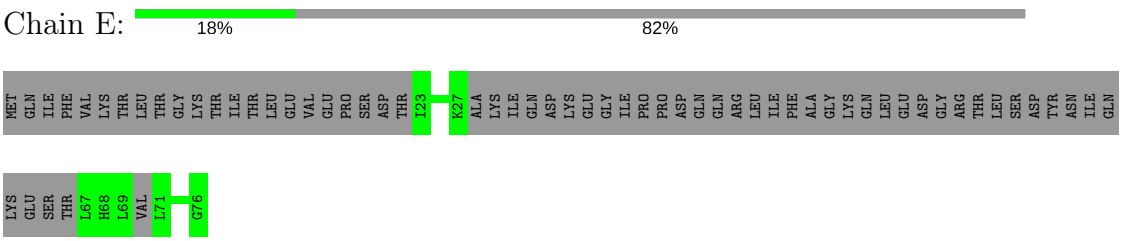
• Molecule 1: Histone deacetylase 6

Chain D:



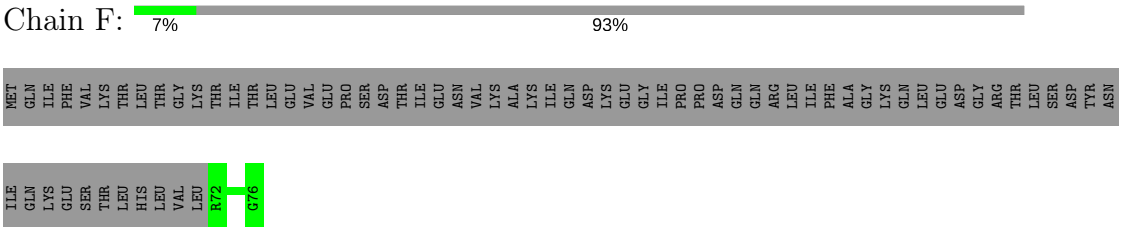
• Molecule 2: Polyubiquitin

Chain E:



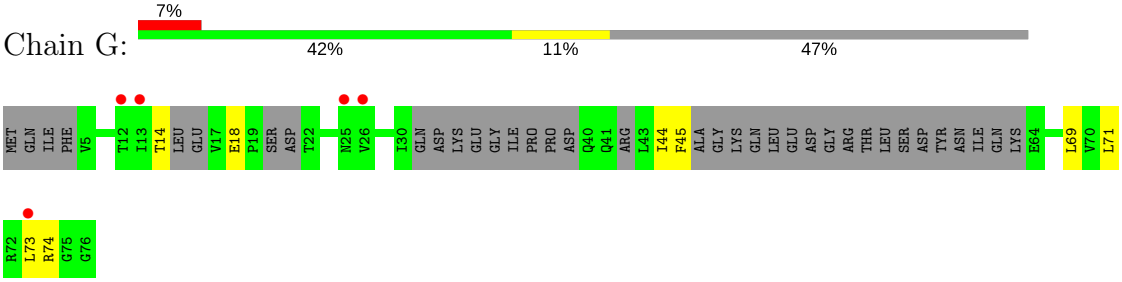
• Molecule 2: Polyubiquitin

Chain F:



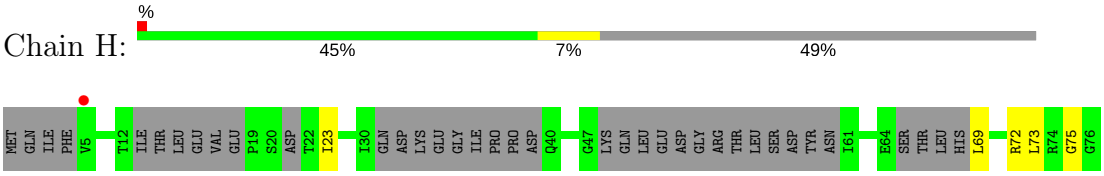
• Molecule 2: Polyubiquitin

Chain G:



• Molecule 2: Polyubiquitin

Chain H:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 43 21 2	Depositor
Cell constants a, b, c, α, β, γ	133.75Å 133.75Å 118.77Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	32.72 – 3.00 44.40 – 3.00	Depositor EDS
% Data completeness (in resolution range)	(Not available) (32.72-3.00) 100.0 (44.40-3.00)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	0.09	Depositor
$< I/\sigma(I) >$ [1]	4.62 (at 3.01Å)	Xtriage
Refinement program	BUSTER 2.8.0	Depositor
R, R_{free}	0.235 , 0.265 0.250 , 0.284	Depositor DCC
R_{free} test set	463 reflections (2.14%)	DCC
Wilson B-factor (Å^2)	78.0	Xtriage
Anisotropy	0.119	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , 69.9	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.90	EDS
Total number of atoms	3321	wwPDB-VP
Average B, all atoms (Å^2)	72.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.82% 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	0.49	0/788	0.65	0/1082
1	B	0.49	0/785	0.64	0/1077
1	C	0.48	0/654	0.67	0/900
1	D	0.47	0/644	0.68	0/884
2	E	0.74	0/72	0.79	0/93
2	F	0.61	0/29	0.66	0/35
2	G	0.62	0/218	0.78	0/289
2	H	0.61	0/204	0.87	1/270 (0.4%)
All	All	0.51	0/3394	0.69	1/4630 (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
2	G	0	1

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	H	72	ARG	C-N-CA	6.43	137.77	121.70

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	G	18	GLU	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	761	0	649	5	0
1	B	758	0	668	4	0
1	C	629	0	516	10	0
1	D	623	0	535	9	0
2	E	74	0	32	0	0
2	F	30	0	22	0	0
2	G	223	0	127	2	0
2	H	209	0	114	2	0
3	A	3	0	0	0	0
3	B	3	0	0	0	0
3	C	3	0	0	0	0
3	D	3	0	0	0	0
4	B	1	0	0	0	0
4	G	1	0	0	0	0
All	All	3321	0	2663	19	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 3.

All (19) 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:B:6:PRO:HB2	1:C:11:VAL:HG11	1.38	1.05
1:A:6:PRO:HB2	1:D:11:VAL:HG11	1.38	1.01
1:A:6:PRO:HB2	1:D:11:VAL:CG1	2.01	0.91
1:B:6:PRO:HB2	1:C:11:VAL:CG1	2.01	0.91
1:D:84:HIS:CD2	2:H:69:LEU:HD13	2.34	0.62
1:A:3:PRO:HG2	1:D:87:LEU:HA	1.86	0.57
1:A:4:TRP:HB3	1:D:87:LEU:HD23	1.86	0.57
1:C:84:HIS:CD2	2:G:69:LEU:HD13	2.44	0.53
2:G:44:ILE:CB	2:G:71:LEU:H	2.22	0.53
1:A:3:PRO:HD2	1:D:87:LEU:HD22	1.93	0.50
1:D:47:ARG:NH2	2:H:75:GLY:O	2.45	0.49
1:C:4:TRP:HH2	1:D:4:TRP:CB	2.28	0.47
1:C:47:ARG:HG3	1:C:48:TYR:HD2	1.81	0.46

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:47:ARG:HG3	1:C:48:TYR:CD2	2.51	0.46
1:C:3:PRO:HA	1:C:83:HIS:HB3	2.01	0.43
1:B:6:PRO:CB	1:C:11:VAL:CG1	2.87	0.42
1:D:63:PRO:HB3	1:D:77:TYR:CD2	2.56	0.40
1:B:6:PRO:CB	1:C:11:VAL:HG11	2.29	0.40
1:C:63:PRO:HB3	1:C:77:TYR:CD2	2.56	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
1	A	100/107 (94%)	93 (93%)	6 (6%)	1 (1%)	18 59
1	B	98/107 (92%)	91 (93%)	6 (6%)	1 (1%)	18 59
1	C	83/107 (78%)	79 (95%)	3 (4%)	1 (1%)	15 53
1	D	82/107 (77%)	78 (95%)	3 (4%)	1 (1%)	15 53
2	E	8/76 (10%)	6 (75%)	2 (25%)	0	100 100
2	F	3/76 (4%)	3 (100%)	0	0	100 100
2	G	28/76 (37%)	24 (86%)	3 (11%)	1 (4%)	4 22
2	H	27/76 (36%)	23 (85%)	2 (7%)	2 (7%)	1 6
All	All	429/732 (59%)	397 (92%)	25 (6%)	7 (2%)	11 46

All (7) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	G	73	LEU
2	H	73	LEU
1	A	49	ILE
1	B	49	ILE

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	C	49	ILE
1	D	49	ILE
2	H	23	ILE

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	72/92 (78%)	71 (99%)	1 (1%)	71 91
1	B	75/92 (82%)	74 (99%)	1 (1%)	73 92
1	C	59/92 (64%)	58 (98%)	1 (2%)	66 89
1	D	61/92 (66%)	60 (98%)	1 (2%)	68 90
2	E	1/68 (2%)	1 (100%)	0	100 100
2	F	1/68 (2%)	1 (100%)	0	100 100
2	G	6/68 (9%)	3 (50%)	3 (50%)	0 0
2	H	4/68 (6%)	4 (100%)	0	100 100
All	All	279/640 (44%)	272 (98%)	7 (2%)	53 84

All (7) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	5	CYS
1	B	5	CYS
1	C	5	CYS
1	D	5	CYS
2	G	14	THR
2	G	45	PHE
2	G	74	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	C	50	ASN
1	C	84	HIS
1	D	84	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 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

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	102/107 (95%)	-0.40	0 100 100	49, 62, 85, 123	0
1	B	100/107 (93%)	-0.29	2 (2%) 65 36	48, 63, 89, 113	0
1	C	85/107 (79%)	-0.40	0 100 100	51, 60, 84, 98	0
1	D	84/107 (78%)	-0.46	0 100 100	51, 66, 82, 105	0
2	E	14/76 (18%)	0.12	0 100 100	60, 115, 145, 146	0
2	F	5/76 (6%)	-0.83	0 100 100	60, 63, 79, 80	0
2	G	40/76 (52%)	0.67	5 (12%) 4 2	61, 123, 155, 156	0
2	H	39/76 (51%)	-0.20	1 (2%) 56 27	68, 119, 145, 148	0
All	All	469/732 (64%)	-0.27	8 (1%) 70 42	48, 66, 132, 156	0

All (8) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	G	25	ASN	4.0
2	G	26	VAL	3.0
2	H	5	VAL	2.7
2	G	12	THR	2.6
2	G	13	ILE	2.2
2	G	73	LEU	2.1
1	B	100	GLY	2.0
1	B	99	PHE	2.0

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^{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	LLDF	B-factors(Å^2)	Q<0.9
3	ZN	B	2001	1/1	0.99	0.13	-0.27	62,62,62,62	0
3	ZN	D	2000	1/1	0.99	0.13	-0.39	66,66,66,66	0
3	ZN	A	1002	1/1	0.99	0.16	-0.90	57,57,57,57	0
3	ZN	B	2000	1/1	1.00	0.15	-0.90	55,55,55,55	0
3	ZN	C	2002	1/1	0.99	0.15	-0.91	67,67,67,67	0
3	ZN	A	1001	1/1	0.99	0.11	-1.11	66,66,66,66	0
3	ZN	C	2000	1/1	1.00	0.14	-1.25	56,56,56,56	0
3	ZN	D	2002	1/1	0.99	0.13	-1.32	66,66,66,66	0
3	ZN	A	1000	1/1	1.00	0.13	-1.41	57,57,57,57	0
3	ZN	B	2002	1/1	1.00	0.12	-1.45	62,62,62,62	0
3	ZN	D	2001	1/1	0.98	0.10	-1.59	71,71,71,71	0
3	ZN	C	2001	1/1	1.00	0.11	-1.68	57,57,57,57	0

6.5 Other polymers (i)

There are no such residues in this entry.