



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 02:09 PM GMT

PDB ID : 3W99
Title : Crystal Structure of Human Nucleosome Core Particle lacking H4 N-terminal region
Authors : Iwasaki, W.; Miya, Y.; Horikoshi, N.; Osakabe, A.; Tachiwana, H.; Shibata, T.; Kagawa, W.; Kurumizaka, H.
Deposited on : 2013-04-01
Resolution : 3.00 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
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 (RC4), CSD as536be (2015)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20026688
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
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) : trunk26865

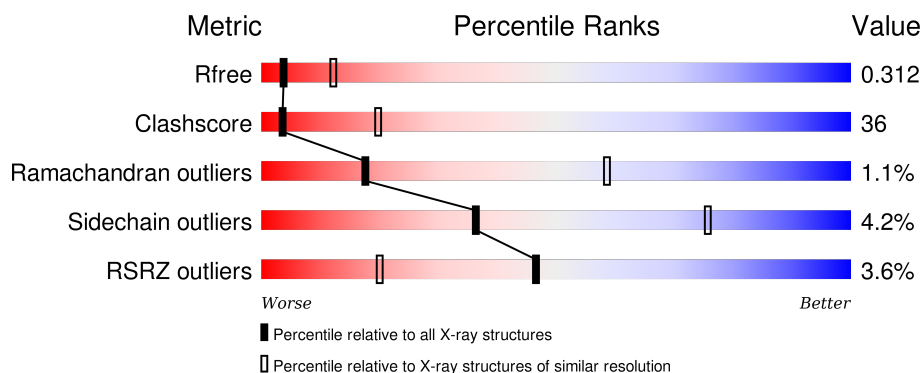
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.00 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}	91344	1578 (3.00-3.00)
Clashscore	102246	1912 (3.00-3.00)
Ramachandran outliers	100387	1853 (3.00-3.00)
Sidechain outliers	100360	1856 (3.00-3.00)
RSRZ outliers	91569	1592 (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 $\leq 5\%$. 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	139	 31% 34% 30%
1	E	139	 37% 35% 29%
2	B	91	 3% 52% 32% 14%
2	F	91	 2% 48% 40% 8%
3	C	133	 47% 29% 21%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	G	133	
4	D	129	
4	H	129	
5	I	146	
5	J	146	

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
6	MN	E	1001	-	-	-	X

2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 11912 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 H3.1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	97	Total	C	N	O	S	0	0	0
			801	505	155	137	4			
1	E	99	Total	C	N	O	S	0	0	0
			816	514	158	140	4			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-3	GLY	-	EXPRESSION TAG	UNP P68431
A	-2	SER	-	EXPRESSION TAG	UNP P68431
A	-1	HIS	-	EXPRESSION TAG	UNP P68431
E	-3	GLY	-	EXPRESSION TAG	UNP P68431
E	-2	SER	-	EXPRESSION TAG	UNP P68431
E	-1	HIS	-	EXPRESSION TAG	UNP P68431

- Molecule 2 is a protein called Histone H4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	78	Total	C	N	O	S	0	0	0
			619	391	120	107	1			
2	F	84	Total	C	N	O	S	0	0	0
			673	424	133	115	1			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	12	GLY	-	EXPRESSION TAG	UNP P62805
B	13	SER	-	EXPRESSION TAG	UNP P62805
B	14	HIS	-	EXPRESSION TAG	UNP P62805
B	15	MET	-	EXPRESSION TAG	UNP P62805
F	12	GLY	-	EXPRESSION TAG	UNP P62805
F	13	SER	-	EXPRESSION TAG	UNP P62805

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
F	14	HIS	-	EXPRESSION TAG	UNP P62805
F	15	MET	-	EXPRESSION TAG	UNP P62805

- Molecule 3 is a protein called Histone H2A type 1-B/E.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	C	105	Total	C	N	O	0	0	0
			810	511	158	141			
3	G	102	Total	C	N	O	0	0	0
			789	498	154	137			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	-3	GLY	-	EXPRESSION TAG	UNP P04908
C	-2	SER	-	EXPRESSION TAG	UNP P04908
C	-1	HIS	-	EXPRESSION TAG	UNP P04908
G	-3	GLY	-	EXPRESSION TAG	UNP P04908
G	-2	SER	-	EXPRESSION TAG	UNP P04908
G	-1	HIS	-	EXPRESSION TAG	UNP P04908

- Molecule 4 is a protein called Histone H2B type 1-J.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	95	Total	C	N	O	S	0	0	0
			745	468	136	139	2			
4	H	92	Total	C	N	O	S	0	0	0
			719	453	129	135	2			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	-3	GLY	-	EXPRESSION TAG	UNP P06899
D	-2	SER	-	EXPRESSION TAG	UNP P06899
D	-1	HIS	-	EXPRESSION TAG	UNP P06899
H	-3	GLY	-	EXPRESSION TAG	UNP P06899
H	-2	SER	-	EXPRESSION TAG	UNP P06899
H	-1	HIS	-	EXPRESSION TAG	UNP P06899

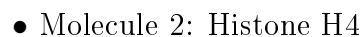
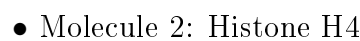
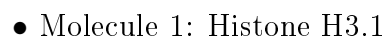
- Molecule 5 is a DNA chain called 146-mer DNA.

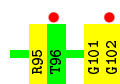
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	I	145	Total	C	N	O	P	0	0	0
			2970	1421	538	867	144			
5	J	145	Total	C	N	O	P	0	0	0
			2969	1421	535	869	144			

- Molecule 6 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	E	1	Total	Mn	0	0
			1	1		

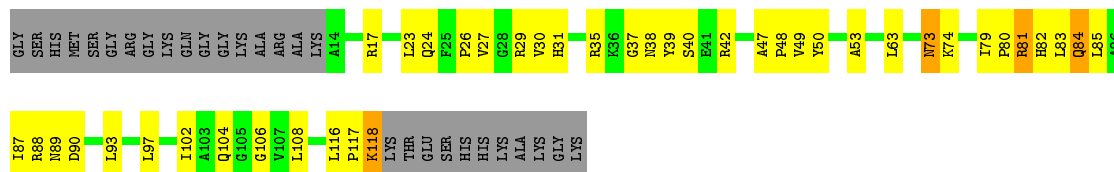
- Molecule 1: Histone H3.1





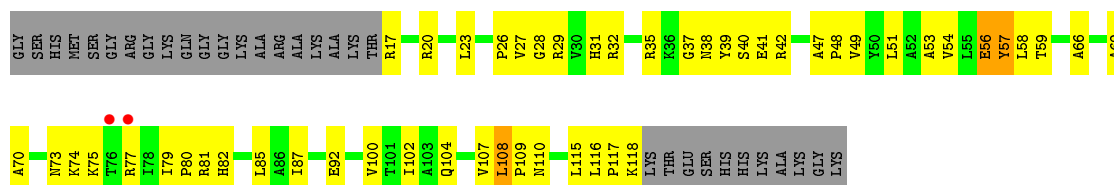
• Molecule 3: Histone H2A type 1-B/E

Chain C: 47% 29% 21%



• Molecule 3: Histone H2A type 1-B/E

Chain G: 38% 36% 23%



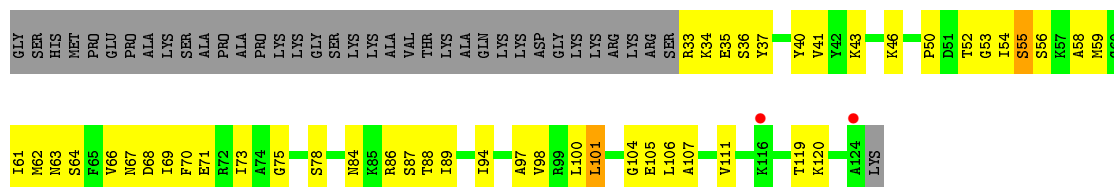
• Molecule 4: Histone H2B type 1-J

Chain D: 50% 22% 26%



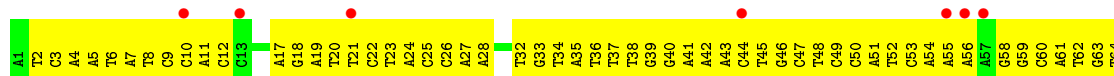
• Molecule 4: Histone H2B type 1-J

Chain H: 35% 35% 29%



• Molecule 5: 146-mer DNA

Chain I: 16% 83% 7%



T65	C66	A67	G68	C69	A73	T74	T75	T80	G81	A82	A83	C84	A85	T86	G87	C88	C89	T90	T91	T92	T93	G94	A95	T96	G97	G98	A99	G100	C101	A102	G103	T104	T105	T106	C107	C108	A109	A110	A111	T112	A113	C114	A115	C116	T117	T118	T119	T120	A126	A127	T128	C129	T130	G131	C132	A133
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------

G134	G135	T136	G137	G138	A139	T140	A141	T142	T143	G144	A145	DT
------	------	------	------	------	------	------	------	------	------	------	------	----

● Molecule 5: 146-mer DNA



DA	T143	C149	A150	A151	T152	A153	T154	C155	C156	A157	C158	C159	T160	G161	C162	G164	A165	T166	T167	C168	A170	C171	C172	A173	A174	A175	A176	G177	T178	T179	T180	A181	T182	T183	T184	G185	G186	A187	A188	A189	C190	T191	G192	C193	T194	C195	C196	A197	T198	C199	A200	A201	A202	A203	G204	G205	C206
----	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------

A207	T208	G209	T210	T211	C212	A213	G214	C215	T216	G217	A218	A219	T220	T221	C222	G223	G224	C225	T226	G227	A228	C229	C230	A231	T232	G233	C234	C235	T236	T237	T238	T239	G240	A241	T242	G243	G244	A245	G246	C247	A248	G249	T250	T251	T252	C253	C254	A255	A256	A257	T258	A259	C260	A261	C262	T263	T264	T265	T266
------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------

G267	G268	T269	A270	G271	A272	A273	T274	C275	T276	G277	C278	A279	G280	G281	T282	G283	G284	A285	T286	T288	T289	G290	A291	T292
------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------

4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	105.70Å 109.29Å 175.72Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.54 – 3.00 48.54 – 3.00	Depositor EDS
% Data completeness (in resolution range)	99.2 (48.54-3.00) 99.3 (48.54-3.00)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.09	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.43 (at 3.01Å)	Xtriage
Refinement program	CNS 1.2	Depositor
R, R_{free}	0.244 , 0.312 0.244 , 0.312	Depositor DCC
R_{free} test set	2064 reflections (5.03%)	DCC
Wilson B-factor (Å ²)	82.0	Xtriage
Anisotropy	0.816	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.28 , 74.2	EDS
Estimated twinning fraction	0.026 for k,h,l	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.46$, $\langle L^2 \rangle = 0.29$	Xtriage
Outliers	0 of 41066 reflections	Xtriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	11912	wwPDB-VP
Average B, all atoms (Å ²)	120.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.71% 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, 0.375 respectively for untwinned datasets, and 0.333, 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: MN

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	0.36	0/813	0.59	0/1090
1	E	0.38	0/828	0.63	0/1109
2	B	0.34	0/626	0.57	0/837
2	F	0.38	0/680	0.64	0/908
3	C	0.33	0/820	0.60	0/1107
3	G	0.30	0/799	0.57	0/1079
4	D	0.35	0/756	0.59	0/1015
4	H	0.33	0/730	0.58	0/982
5	I	0.36	0/3332	0.76	0/5141
5	J	0.36	0/3330	0.78	0/5138
All	All	0.35	0/12714	0.70	0/18406

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	801	0	839	72	0
1	E	816	0	856	63	0
2	B	619	0	659	50	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	F	673	0	722	60	0
3	C	810	0	866	59	0
3	G	789	0	841	68	0
4	D	745	0	771	34	0
4	H	719	0	740	44	0
5	I	2970	0	1640	201	0
5	J	2969	0	1641	247	0
6	E	1	0	0	0	0
All	All	11912	0	9575	773	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 36.

The worst 5 of 773 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)
2:B:78:ARG:HD3	5:J:248:DA:H5'	1.23	1.17
5:J:241:DA:H2''	5:J:242:DT:H5'	1.18	1.13
5:I:128:DT:H2''	5:I:129:DC:H5''	1.10	1.05
5:I:52:DT:H2''	5:I:53:DC:H5'	1.38	1.05
5:I:101:DC:H2''	5:I:102:DA:H5''	1.35	1.05

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	95/139 (68%)	84 (88%)	9 (10%)	2 (2%)	9	40
1	E	97/139 (70%)	88 (91%)	9 (9%)	0	100	100
2	B	76/91 (84%)	67 (88%)	8 (10%)	1 (1%)	15	53

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	F	82/91 (90%)	78 (95%)	4 (5%)	0	100	100
3	C	103/133 (77%)	95 (92%)	8 (8%)	0	100	100
3	G	100/133 (75%)	88 (88%)	10 (10%)	2 (2%)	9	41
4	D	93/129 (72%)	84 (90%)	8 (9%)	1 (1%)	17	58
4	H	90/129 (70%)	77 (86%)	11 (12%)	2 (2%)	8	38
All	All	736/984 (75%)	661 (90%)	67 (9%)	8 (1%)	17	58

5 of 8 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
4	H	55	SER
1	A	86	SER
4	H	104	GLY
3	G	56	GLU
3	G	57	TYR

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	85/113 (75%)	77 (91%)	8 (9%)	11	39
1	E	86/113 (76%)	86 (100%)	0	100	100
2	B	63/75 (84%)	61 (97%)	2 (3%)	46	82
2	F	69/75 (92%)	65 (94%)	4 (6%)	25	63
3	C	83/102 (81%)	78 (94%)	5 (6%)	24	62
3	G	81/102 (79%)	79 (98%)	2 (2%)	55	86
4	D	81/107 (76%)	79 (98%)	2 (2%)	55	86
4	H	78/107 (73%)	75 (96%)	3 (4%)	40	78
All	All	626/794 (79%)	600 (96%)	26 (4%)	36	76

5 of 26 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	C	81	ARG
3	C	118	LYS
4	H	101	LEU
3	C	84	GLN
3	C	90	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 30 such sidechains are listed below:

Mol	Chain	Res	Type
3	C	89	ASN
1	E	68	GLN
3	G	38	ASN
3	C	112	GLN
1	E	76	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 1 ligands modelled in this entry, 1 is 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	97/139 (69%)	0.21	2 (2%) 67 36	59, 92, 125, 160	0
1	E	99/139 (71%)	0.18	2 (2%) 68 39	54, 76, 117, 148	0
2	B	78/91 (85%)	0.23	3 (3%) 44 18	61, 91, 111, 135	0
2	F	84/91 (92%)	0.34	2 (2%) 62 32	48, 74, 95, 133	0
3	C	105/133 (78%)	0.14	0 100 100	57, 83, 114, 145	0
3	G	102/133 (76%)	0.07	2 (1%) 68 39	68, 92, 123, 136	0
4	D	95/129 (73%)	0.24	3 (3%) 51 23	58, 88, 117, 169	0
4	H	92/129 (71%)	0.34	2 (2%) 65 35	68, 92, 125, 139	0
5	I	145/146 (99%)	0.53	10 (6%) 20 7	75, 152, 191, 199	0
5	J	145/146 (99%)	0.62	12 (8%) 14 5	82, 153, 194, 200	0
All	All	1042/1276 (81%)	0.31	38 (3%) 46 20	48, 94, 179, 200	0

The worst 5 of 38 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	37	LYS	5.1
5	J	161	DG	4.7
4	D	31	ARG	4.5
4	D	33	ARG	4.0
5	I	138	DG	3.9

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, 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	LLDF	B-factors(\AA^2)	Q<0.9
6	MN	E	1001	1/1	0.92	0.33	2.53	75,75,75,75	0

6.5 Other polymers [i](#)

There are no such residues in this entry.