



# wwPDB X-ray Structure Validation Summary Report

Feb 28, 2014 – 05:44 AM GMT

PDB ID : 2PYO  
Title : Drosophila nucleosome core  
Authors : Clapier, C.R.; Petosa, C.; Mueller, C.W.  
Deposited on : 2007-05-16  
Resolution : 2.43 Å(reported)

This is a wwPDB 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/ValidationPDFNotes.html>

---

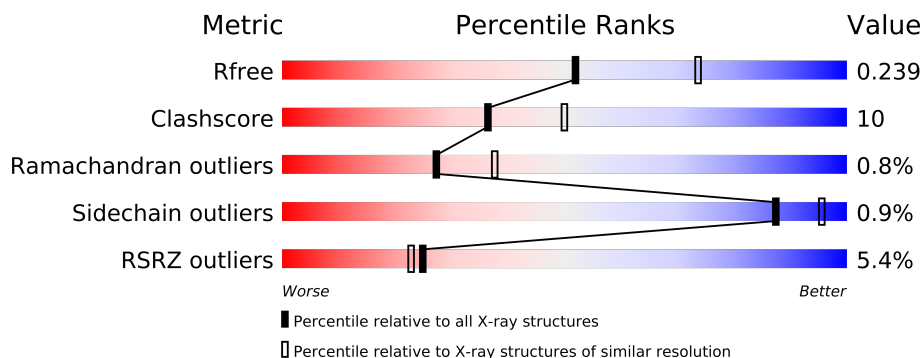
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.15 2013
Xtriage (Phenix)	:	dev-1323
EDS	:	stable22639
Percentile statistics	:	21963
Refmac	:	5.8.0049
CCP4	:	6.3.0 (Settle)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	stable22683

# 1 Overall quality at a glance

The reported resolution of this entry is 2.43 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))
$R_{free}$	66092	2989 (2.48-2.40)
Clashscore	79885	3698 (2.48-2.40)
Ramachandran outliers	78287	3639 (2.48-2.40)
Sidechain outliers	78261	3640 (2.48-2.40)
RSRZ outliers	66119	2993 (2.48-2.40)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	I	147	
2	J	147	
3	A	135	
3	E	135	
4	B	102	
4	F	102	
5	C	120	
5	G	120	
6	D	122	
6	H	122	

The following table lists non-polymeric compounds that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Geometry	Electron density
7	MN	E	1002	-	X
7	MN	I	1008	-	X
7	MN	J	1004	-	X
7	MN	J	1005	-	X
8	CL	D	1016	-	X

## 2 Entry composition

There are 9 unique types of molecules in this entry. The entry contains 12239 atoms, of which 0 are hydrogen and 0 are deuterium.

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 (147-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	I	147	Total	C	N	O	P	0	0	0
			3011	1440	546	879	146			

- Molecule 2 is a DNA chain called DNA (147-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	J	147	Total	C	N	O	P	0	0	0
			3010	1440	543	881	146			

- Molecule 3 is a protein called Histone H3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	A	99	Total	C	N	O	S	0	0	0
			816	514	158	141	3			
3	E	99	Total	C	N	O	S	0	0	0
			816	514	158	141	3			

- Molecule 4 is a protein called Histone H4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	B	80	Total	C	N	O	S	0	0	0
			631	397	122	111	1			
4	F	88	Total	C	N	O	S	0	0	0
			708	445	143	119	1			

- Molecule 5 is a protein called Histone H2A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	C	108	Total	C	N	O	S	0	0	0
			811	508	158	144	1			
5	G	109	Total	C	N	O	S	0	0	0
			824	518	161	144	1			

- Molecule 6 is a protein called Histone H2B.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	D	95	Total	C	N	O	S	0	0	0
			753	474	137	140	2			
6	H	95	Total	C	N	O	S	0	0	0
			753	474	137	140	2			

- Molecule 7 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	J	6	Total	Mn	0	0
			6	6		
7	I	7	Total	Mn	0	0
			7	7		
7	E	1	Total	Mn	0	0
			1	1		

- Molecule 8 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	H	1	Total	Cl	0	0
			1	1		
8	A	1	Total	Cl	0	0
			1	1		
8	D	1	Total	Cl	0	0
			1	1		
8	E	1	Total	Cl	0	0
			1	1		

- Molecule 9 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	I	10	Total	O	0	0
			10	10		
9	B	7	Total	O	0	0
			7	7		
9	J	2	Total	O	0	0
			2	2		
9	A	8	Total	O	0	0
			8	8		
9	C	10	Total	O	0	0
			10	10		
9	D	9	Total	O	0	0
			9	9		

*Continued on next page...*

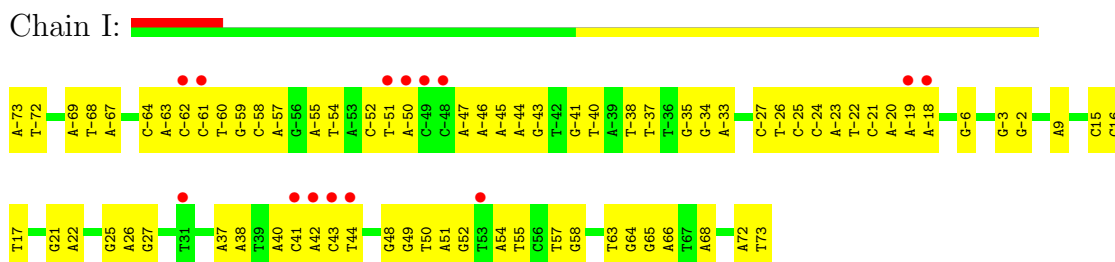
*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	E	15	Total 15	O 15	0	0
9	F	16	Total 16	O 16	0	0
9	G	7	Total 7	O 7	0	0
9	H	4	Total 4	O 4	0	0

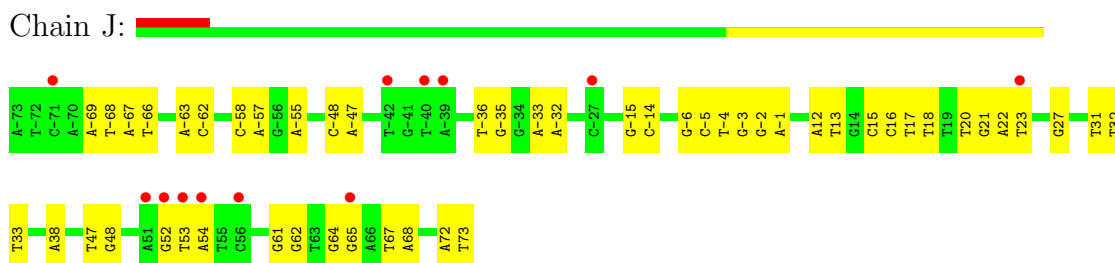
### 3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

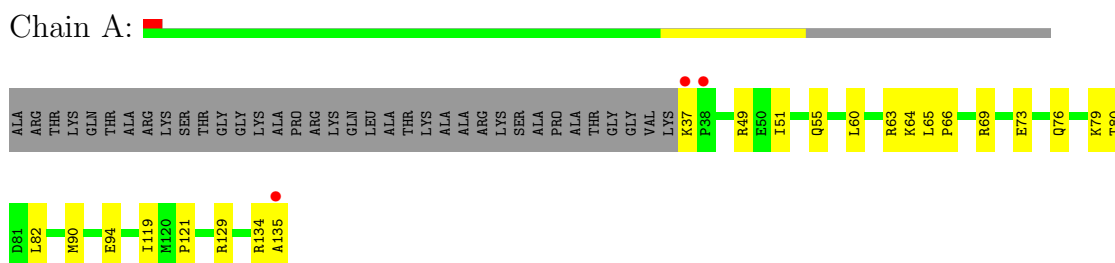
- Molecule 1: DNA (147-MER)



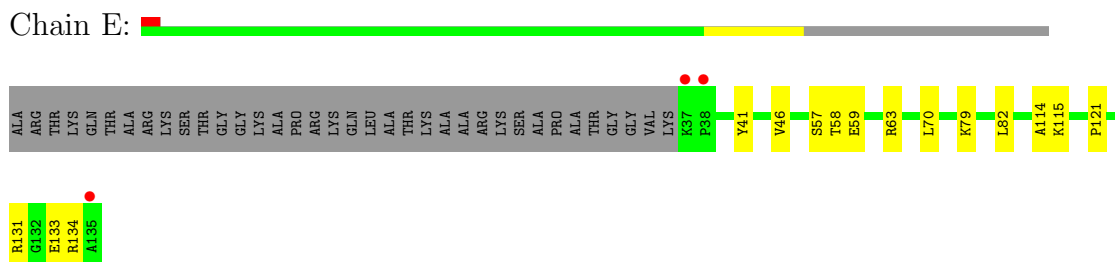
- Molecule 2: DNA (147-MER)



- Molecule 3: Histone H3

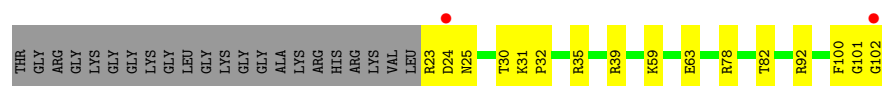


- Molecule 3: Histone H3



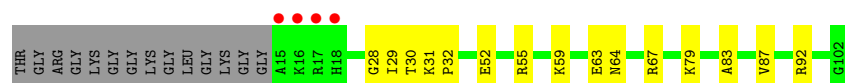
- Molecule 4: Histone H4

Chain B:



- Molecule 4: Histone H4

Chain F:



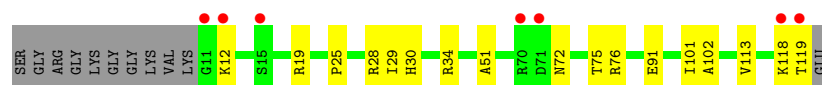
- Molecule 5: Histone H2A

Chain C:



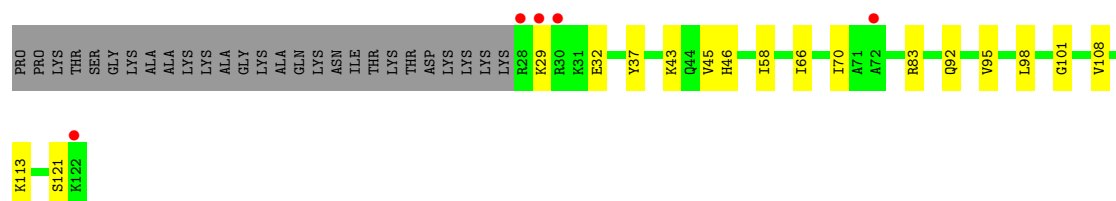
- Molecule 5: Histone H2A

Chain G:



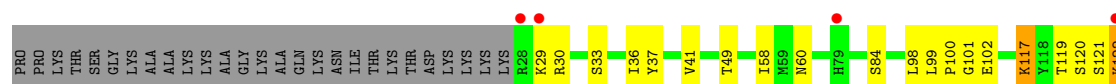
- Molecule 6: Histone H2B

Chain D:



- Molecule 6: Histone H2B

Chain H:





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	106.03Å 182.04Å 109.42Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	500.00 – 2.43 48.62 – 2.43	Depositor EDS
% Data completeness (in resolution range)	93.6 (500.00-2.43) 93.6 (48.62-2.43)	Depositor EDS
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.36 (at 2.42Å)	Xtriage
Refinement program	CNS	Depositor
R, $R_{free}$	0.229 , 0.262 0.234 , 0.239	Depositor DCC
$R_{free}$ test set	3789 reflections (5.30%)	DCC
Wilson B-factor (Å <sup>2</sup> )	55.2	Xtriage
Anisotropy	0.351	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 47.4	EDS
Estimated twinning fraction	0.025 for l,-k,h	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtriage
Outliers	0 of 75234 reflections	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	12239	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	76.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.48% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

## 5 Model quality

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: MN, CL

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	I	0.41	0/3378	0.72	0/5212
2	J	0.42	0/3376	0.75	1/5209 (0.0%)
3	A	0.53	0/828	0.68	0/1109
3	E	0.60	0/828	0.74	0/1109
4	B	0.57	0/638	0.79	0/853
4	F	0.64	0/716	0.84	0/955
5	C	0.55	0/821	0.65	0/1106
5	G	0.46	0/834	0.70	0/1123
6	D	0.57	0/764	0.68	0/1024
6	H	0.57	0/764	0.69	0/1024
All	All	0.49	0/12947	0.73	1/18724 (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	I	0	1
2	J	0	1
All	All	0	2

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	J	27	DG	N9-C4-C5	5.09	107.44	105.40

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	I	-6	DG	Sidechain
2	J	-6	DG	Sidechain

## 5.2 Close contacts ⓘ

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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	I	3011	0	1662	78	0
2	J	3010	0	1663	51	0
3	A	816	0	856	21	0
3	E	816	0	856	12	0
4	B	631	0	664	16	0
4	F	708	0	760	16	0
5	C	811	0	844	24	0
5	G	824	0	871	15	0
6	D	753	0	785	16	0
6	H	753	0	785	21	0
7	E	1	0	0	0	0
7	I	7	0	0	0	0
7	J	6	0	0	0	0
8	A	1	0	0	1	0
8	D	1	0	0	0	0
8	E	1	0	0	1	0
8	H	1	0	0	0	0
9	A	8	0	0	1	0
9	B	7	0	0	0	0
9	C	10	0	0	1	0
9	D	9	0	0	0	0
9	E	15	0	0	1	0
9	F	16	0	0	1	0
9	G	7	0	0	0	0
9	H	4	0	0	0	0
9	I	10	0	0	4	0
9	J	2	0	0	3	0
All	All	12239	0	9746	226	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including

hydrogens) of the entry. The overall clashscore for this entry is 10.

The worst 5 of 226 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:I:-51:DT:H2"	1:I:-50:DA:H5'	1.39	1.03
1:I:72:DA:H2"	1:I:73:DT:H5"	1.41	0.98
1:I:50:DT:H5"	6:H:29:LYS:O	1.77	0.83
4:F:92:ARG:HH21	6:H:98:LEU:HD23	1.41	0.82
4:F:32:PRO:HG3	9:F:520:HOH:O	1.80	0.81

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	
3	A	97/135 (72%)	94 (97%)	3 (3%)	0	100	100
3	E	97/135 (72%)	94 (97%)	3 (3%)	0	100	100
4	B	78/102 (76%)	75 (96%)	3 (4%)	0	100	100
4	F	86/102 (84%)	82 (95%)	4 (5%)	0	100	100
5	C	106/120 (88%)	100 (94%)	5 (5%)	1 (1%)	25	33
5	G	107/120 (89%)	100 (94%)	6 (6%)	1 (1%)	25	33
6	D	93/122 (76%)	91 (98%)	0	2 (2%)	10	10
6	H	93/122 (76%)	90 (97%)	1 (1%)	2 (2%)	10	10
All	All	757/958 (79%)	726 (96%)	25 (3%)	6 (1%)	27	37

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
5	C	117	LYS
6	D	101	GLY
6	H	101	GLY
5	G	118	LYS
6	D	29	LYS

### 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	
3	A	86/110 (78%)	86 (100%)	0	100	100
3	E	86/110 (78%)	85 (99%)	1 (1%)	82	93
4	B	64/78 (82%)	64 (100%)	0	100	100
4	F	72/78 (92%)	72 (100%)	0	100	100
5	C	80/91 (88%)	80 (100%)	0	100	100
5	G	82/91 (90%)	80 (98%)	2 (2%)	61	80
6	D	82/103 (80%)	81 (99%)	1 (1%)	82	93
6	H	82/103 (80%)	80 (98%)	2 (2%)	61	80
All	All	634/764 (83%)	628 (99%)	6 (1%)	87	95

5 of 6 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
5	G	12	LYS
6	H	122	LYS
5	G	72	ASN
3	E	63	ARG
6	H	117	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 12 such sidechains are listed below:

Mol	Chain	Res	Type
3	E	125	GLN
5	G	18	ASN
6	H	44	GLN
3	E	39	HIS
5	G	109	ASN

### 5.3.3 RNA ⓘ

There are no RNA chains 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 18 ligands modelled in this entry, 18 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.

## 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<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	I	147/147 (100%)	0.36	14 (9%) 8 7	48, 99, 140, 161	0
2	J	147/147 (100%)	0.36	12 (8%) 12 10	55, 101, 143, 162	0
3	A	99/135 (73%)	0.31	3 (3%) 48 46	35, 51, 78, 101	0
3	E	99/135 (73%)	0.36	3 (3%) 48 46	31, 42, 64, 98	0
4	B	80/102 (78%)	0.32	2 (2%) 54 52	38, 48, 70, 115	0
4	F	88/102 (86%)	0.65	4 (4%) 32 30	29, 41, 89, 136	0
5	C	108/120 (90%)	0.54	5 (4%) 31 29	32, 45, 75, 115	0
5	G	109/120 (90%)	0.33	7 (6%) 19 17	38, 53, 93, 120	0
6	D	95/122 (77%)	0.45	5 (5%) 25 23	30, 46, 78, 137	0
6	H	95/122 (77%)	0.55	4 (4%) 35 33	39, 54, 94, 124	0
All	All	1067/1252 (85%)	0.42	59 (5%) 25 22	29, 54, 129, 162	0

The worst 5 of 59 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
5	C	119	THR	10.8
3	E	37	LYS	8.0
3	E	135	ALA	7.4
6	D	28	ARG	7.2
5	C	120	GLU	6.0

### 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 ⓘ

There are no carbohydrates in this entry.

### 6.4 Ligands ⓘ

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	RSR	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
7	MN	J	1004	1/1	0.17	17.68	89,89,89,89	0
7	MN	E	1002	1/1	0.25	4.68	45,45,45,45	0
7	MN	J	1005	1/1	0.20	3.00	86,86,86,86	0
7	MN	I	1008	1/1	0.21	2.71	90,90,90,90	0
8	CL	D	1016	1/1	0.21	2.28	56,56,56,56	0
7	MN	J	1001	1/1	0.17	1.63	87,87,87,87	0
7	MN	I	1007	1/1	0.17	1.30	100,100,100,100	0
7	MN	J	1003	1/1	0.17	0.57	76,76,76,76	0
8	CL	E	1018	1/1	0.14	-0.99	73,73,73,73	0
8	CL	H	1015	1/1	0.12	-2.13	59,59,59,59	0
7	MN	I	1013	1/1	0.06	-2.37	94,94,94,94	0
7	MN	I	1006	1/1	0.10	-2.76	86,86,86,86	0
7	MN	I	1011	1/1	0.10	-3.39	98,98,98,98	0
7	MN	I	1010	1/1	0.05	-3.48	115,115,115,115	0
7	MN	I	1012	1/1	0.10	-3.54	94,94,94,94	0
7	MN	J	1014	1/1	0.10	-4.36	121,121,121,121	0
7	MN	J	1009	1/1	0.10	-6.04	116,116,116,116	0
8	CL	A	1017	1/1	0.08	-6.99	73,73,73,73	0

### 6.5 Other polymers ⓘ

There are no such residues in this entry.