wwPDB X-ray Structure Validation Summary Report (i)

Nov 5, 2023 – 04:32 AM EST

PDB ID	:	4YS3
Title	:	Nucleosome disassembly by RSC and SWI/SNF is enhanced by H3 acetylation near the nucleosome dyad axis
Authors	:	Dechassa, M.L.; Luger, K.; Chatterjee, N.; North, J.A.; Manohar, M.; Prasad, R.; Ottessen, J.J.; Poirier, M.G.; Bartholomew, B.
Deposited on	:	2015-03-16
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
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

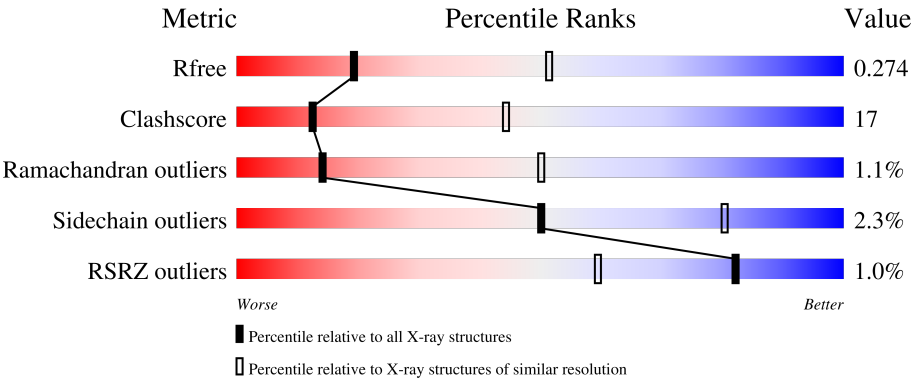
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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}	130704	2092 (3.00-3.00)
Clashscore	141614	2416 (3.00-3.00)
Ramachandran outliers	138981	2333 (3.00-3.00)
Sidechain outliers	138945	2336 (3.00-3.00)
RSRZ outliers	127900	1990 (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 of 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	98	69% 28% .
1	E	98	76% 20% .
2	B	79	71% 28% .
2	F	79	% 72% 28%
3	C	107	3% 68% 29% .

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	G	107	69% 27%
4	D	93	62% 35%
4	H	93	68% 32%
5	I	147	44% 55%
6	J	147	46% 52%

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
1	ALY	A	522	-	-	X	-

2 Entry composition (i)

There are 7 unique types of molecules in this entry. The entry contains 12058 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 H3.2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	98	Total	C	N	O	S	0	0	0
			814	513	156	142	3			
1	E	98	Total	C	N	O	S	0	0	0
			814	513	156	142	3			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	502	ALA	GLY	engineered mutation	UNP P84233
E	702	ALA	GLY	engineered mutation	UNP P84233

- Molecule 2 is a protein called Histone H4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	79	Total	C	N	O	S	0	0	0
			627	395	121	110	1			
2	F	79	Total	C	N	O	S	0	0	0
			627	395	121	110	1			

- Molecule 3 is a protein called Histone H2A.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	C	107	Total	C	N	O	0	0	0
			826	520	161	145			
3	G	106	Total	C	N	O	0	0	0
			818	516	160	142			

- Molecule 4 is a protein called Histone H2B 1.1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	93	Total	C	N	O	S	0	0	0
			729	459	131	137	2			

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	H	93	Total	C	N	O	S	0	0	0
			729	459	131	137	2			

- Molecule 5 is a DNA chain called DNA (147-MER).

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

- Molecule 6 is a DNA chain called DNA (147-MER).

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

- Molecule 7 is water.

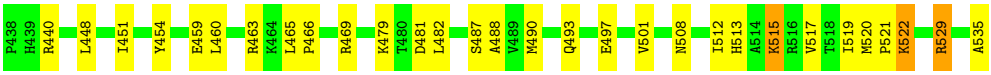
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	2	Total	O	0	0
			2	2		
7	B	2	Total	O	0	0
			2	2		
7	C	5	Total	O	0	0
			5	5		
7	D	2	Total	O	0	0
			2	2		
7	E	7	Total	O	0	0
			7	7		
7	F	3	Total	O	0	0
			3	3		
7	G	3	Total	O	0	0
			3	3		
7	I	17	Total	O	0	0
			17	17		
7	J	12	Total	O	0	0
			12	12		

3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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 H3.2

Chain A:



• Molecule 1: Histone H3.2

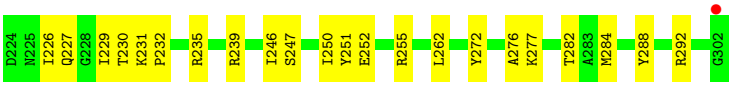
Chain E:

• Molecule 2: Histone H4

Chain B:

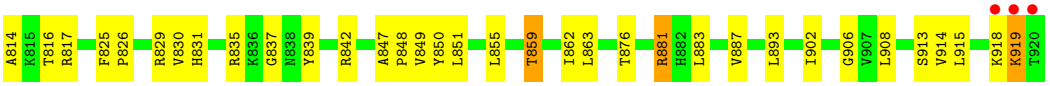
• Molecule 2: Histone H4

Chain F:



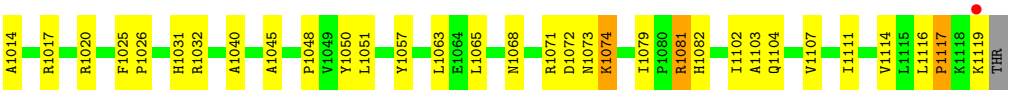
• Molecule 3: Histone H2A

Chain C:



• Molecule 3: Histone H2A

Chain G:



• Molecule 4: Histone H2B 1.1

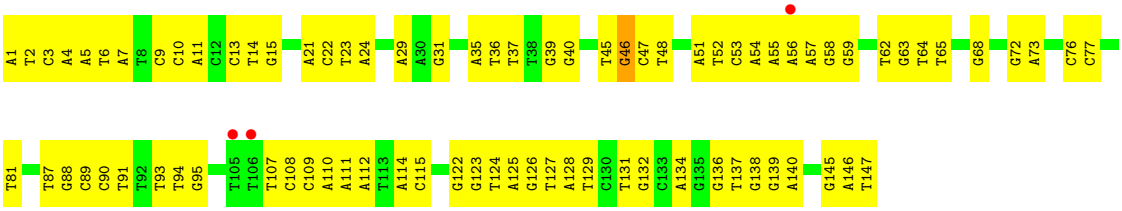
Chain D:

• Molecule 4: Histone H2B 1.1

Chain H:

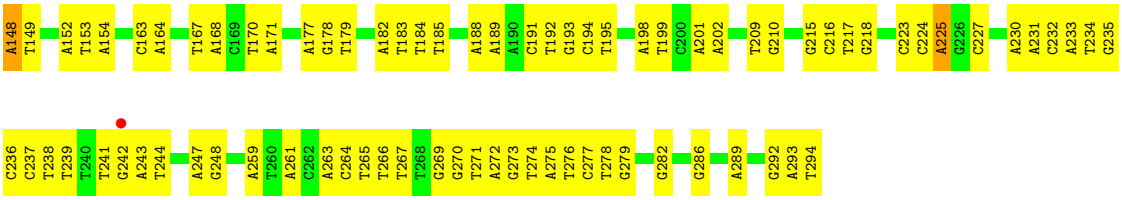
• Molecule 5: DNA (147-MER)

Chain I:



• Molecule 6: DNA (147-MER)

Chain J:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	105.51Å 109.62Å 181.23Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	47.45 – 3.00 47.54 – 3.00	Depositor EDS
% Data completeness (in resolution range)	97.8 (47.45-3.00) 99.3 (47.54-3.00)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.02 (at 3.01Å)	Xtriage
Refinement program	CNS	Depositor
R, R_{free}	0.226 , 0.279 0.223 , 0.274	Depositor DCC
R_{free} test set	2423 reflections (2.99%)	wwPDB-VP
Wilson B-factor (Å^2)	66.4	Xtriage
Anisotropy	0.129	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.28 , 61.1	EDS
L-test for twinning[2]	$< \|L\| > = 0.51$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	0.000 for k,h,-l	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	12058	wwPDB-VP
Average B, all atoms (Å^2)	80.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.91% 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 (i)

5.1 Standard geometry (i)

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

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	0.43	0/800	0.60	0/1071
1	E	0.48	0/800	0.69	0/1071
2	B	0.43	0/634	0.65	0/848
2	F	0.51	0/634	0.71	0/848
3	C	0.41	0/836	0.63	0/1127
3	G	0.36	0/828	0.63	0/1117
4	D	0.45	0/740	0.63	0/994
4	H	0.42	0/740	0.61	0/994
5	I	0.42	0/3378	0.72	0/5212
6	J	0.41	0/3376	0.73	0/5209
All	All	0.43	0/12766	0.69	0/18491

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
5	I	0	4
6	J	0	6
All	All	0	10

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 10 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
5	I	134	DA	Sidechain

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Group
5	I	145	DG	Sidechain
5	I	46	DG	Sidechain
5	I	68	DG	Sidechain
6	J	148	DA	Sidechain

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	814	0	847	48	0
1	E	814	0	847	28	0
2	B	627	0	663	25	0
2	F	627	0	663	21	0
3	C	826	0	884	38	0
3	G	818	0	877	30	0
4	D	729	0	753	41	0
4	H	729	0	753	24	0
5	I	3011	0	1662	83	0
6	J	3010	0	1663	87	0
7	A	2	0	0	1	0
7	B	2	0	0	0	0
7	C	5	0	0	0	0
7	D	2	0	0	1	0
7	E	7	0	0	1	0
7	F	3	0	0	0	0
7	G	3	0	0	0	0
7	I	17	0	0	2	0
7	J	12	0	0	3	0
All	All	12058	0	9612	360	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 17.

The worst 5 of 360 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)
6:J:273:DG:H2"	6:J:274:DT:H5"	1.23	1.15
5:I:137:DT:H2"	5:I:138:DG:H5"	1.35	1.04
1:A:515:ALY:HH32	1:E:714:ALA:HB1	1.39	1.04
5:I:125:DA:H2"	5:I:126:DG:H5"	1.39	1.02
4:D:1295:VAL:HG13	4:D:1299:LEU:HD12	1.40	1.02

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	94/98 (96%)	90 (96%)	3 (3%)	1 (1%)	14 50
1	E	94/98 (96%)	88 (94%)	6 (6%)	0	100 100
2	B	77/79 (98%)	74 (96%)	3 (4%)	0	100 100
2	F	77/79 (98%)	75 (97%)	2 (3%)	0	100 100
3	C	105/107 (98%)	97 (92%)	6 (6%)	2 (2%)	8 36
3	G	104/107 (97%)	92 (88%)	10 (10%)	2 (2%)	8 36
4	D	91/93 (98%)	82 (90%)	8 (9%)	1 (1%)	14 50
4	H	91/93 (98%)	83 (91%)	6 (7%)	2 (2%)	6 31
All	All	733/754 (97%)	681 (93%)	44 (6%)	8 (1%)	14 50

5 of 8 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	481	ASP
3	C	918	LYS
3	C	914	VAL
4	D	1301	GLY
4	H	1501	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
1	A	83/83 (100%)	81 (98%)	2 (2%)	49 79
1	E	83/83 (100%)	81 (98%)	2 (2%)	49 79
2	B	64/64 (100%)	62 (97%)	2 (3%)	40 75
2	F	64/64 (100%)	64 (100%)	0	100 100
3	C	85/85 (100%)	82 (96%)	3 (4%)	36 71
3	G	84/85 (99%)	82 (98%)	2 (2%)	49 79
4	D	79/79 (100%)	77 (98%)	2 (2%)	47 79
4	H	79/79 (100%)	78 (99%)	1 (1%)	69 89
All	All	621/622 (100%)	607 (98%)	14 (2%)	50 80

5 of 14 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
4	D	1269	ARG
4	D	1309	SER
4	H	1516	THR
3	G	1074	LYS
3	G	1081	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 9 such sidechains are listed below:

Mol	Chain	Res	Type
3	G	1038	ASN
3	G	1073	ASN
4	D	1279	HIS
4	D	1292	GLN
2	F	227	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)

4 non-standard protein/DNA/RNA residues are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
1	ALY	E	715	1	10,11,12	2.53	4 (40%)	7,12,14	1.85	1 (14%)
1	ALY	A	515	1	10,11,12	2.45	3 (30%)	7,12,14	1.75	1 (14%)
1	ALY	A	522	1	10,11,12	2.82	5 (50%)	7,12,14	1.77	1 (14%)
1	ALY	E	722	1	10,11,12	1.35	1 (10%)	7,12,14	2.21	1 (14%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
1	ALY	E	715	1	-	6/9/10/12	-
1	ALY	A	515	1	-	2/9/10/12	-
1	ALY	A	522	1	-	4/9/10/12	-
1	ALY	E	722	1	-	2/9/10/12	-

The worst 5 of 13 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	522	ALY	CH-NZ	6.15	1.52	1.34
1	E	715	ALY	CH-NZ	5.06	1.48	1.34
1	A	515	ALY	CH-NZ	4.93	1.48	1.34
1	E	715	ALY	CE-NZ	4.15	1.55	1.46
1	A	515	ALY	CE-NZ	4.10	1.55	1.46

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	E	722	ALY	CE-NZ-CH	-5.53	114.05	122.56

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	E	715	ALY	CE-NZ-CH	-4.81	115.16	122.56
1	A	522	ALY	CE-NZ-CH	-4.54	115.57	122.56
1	A	515	ALY	CE-NZ-CH	-3.84	116.65	122.56

There are no chirality outliers.

5 of 14 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	E	715	ALY	N-CA-CB-CG
1	E	715	ALY	C-CA-CB-CG
1	E	715	ALY	CE-CD-CG-CB
1	A	522	ALY	CD-CE-NZ-CH
1	E	715	ALY	CG-CD-CE-NZ

There are no ring outliers.

4 monomers are involved in 22 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	E	715	ALY	1	0
1	A	515	ALY	5	0
1	A	522	ALY	12	0
1	E	722	ALY	5	0

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

There are no ligands in this entry.

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, 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	96/98 (97%)	-0.24	0 100 100	26, 62, 96, 140	0
1	E	96/98 (97%)	-0.28	0 100 100	26, 54, 81, 154	0
2	B	79/79 (100%)	-0.37	0 100 100	38, 60, 91, 152	0
2	F	79/79 (100%)	-0.12	1 (1%) 77 51	29, 52, 78, 132	0
3	C	107/107 (100%)	-0.18	3 (2%) 53 25	35, 59, 94, 184	0
3	G	106/107 (99%)	-0.18	1 (0%) 84 63	44, 64, 112, 179	0
4	D	93/93 (100%)	-0.17	1 (1%) 80 56	37, 57, 94, 150	0
4	H	93/93 (100%)	-0.19	0 100 100	37, 65, 103, 153	0
5	I	147/147 (100%)	-0.25	3 (2%) 65 36	53, 95, 131, 171	0
6	J	147/147 (100%)	-0.35	1 (0%) 87 69	56, 99, 131, 161	0
All	All	1043/1048 (99%)	-0.24	10 (0%) 82 59	26, 67, 122, 184	0

The worst 5 of 10 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	C	920	THR	8.7
4	D	1322	LYS	4.4
5	I	56	DA	3.9
3	C	919	LYS	3.8
5	I	105	DT	3.6

6.2 Non-standard residues in protein, DNA, RNA chains

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
1	ALY	A	522	12/13	0.80	0.28	21,25,26,30	0
1	ALY	E	722	12/13	0.81	0.40	21,25,26,30	12
1	ALY	A	515	12/13	0.84	0.23	21,24,26,30	0
1	ALY	E	715	12/13	0.88	0.40	21,24,26,30	0

6.3 Carbohydrates (i)

There are no monosaccharides in this entry.

6.4 Ligands (i)

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.