wwPDB X-ray Structure Validation Summary Report (i)

Mar 27, 2024 – 10:12 AM JST

PDB ID	:	8YTI
Title	:	Crystal Structure of Nucleosome-H1x Linker Histone Assembly (sticky-169a DNA fragment)
Authors	:	Adhireksan, Z.; Qiuye, B.; Padavattan, S.; Davey, C.A.
Deposited on	:	2024-03-26
Resolution	:	2.70 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 (1)) were used in the production of this report:

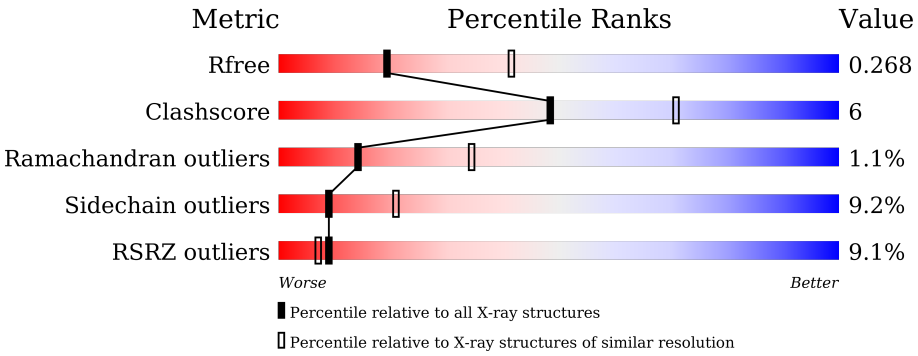
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation 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 2.70 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.



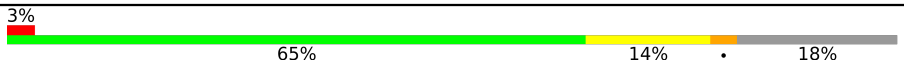
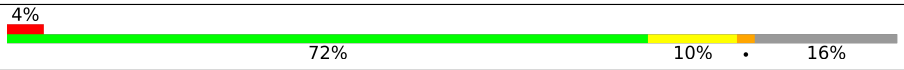
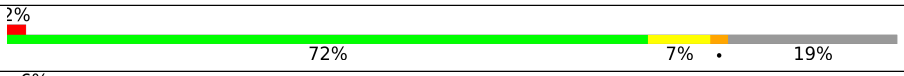
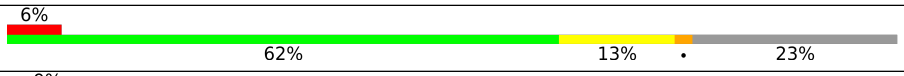
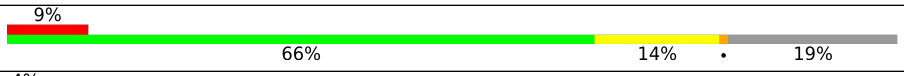
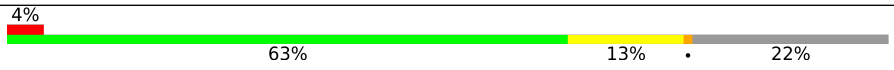
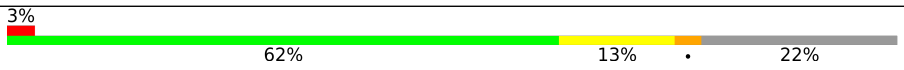
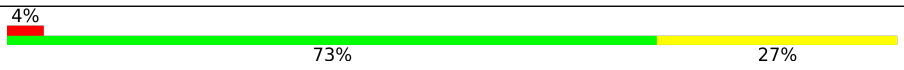
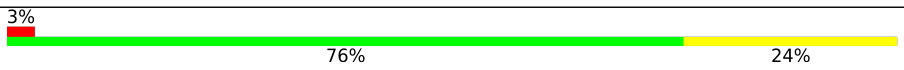
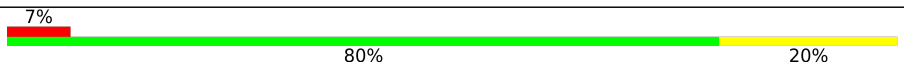
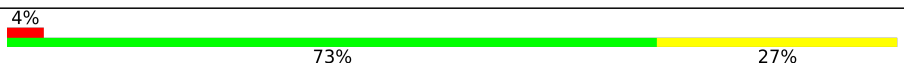
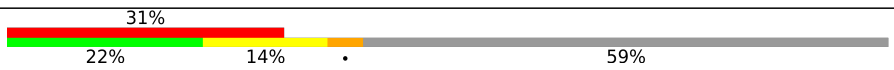
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	2808 (2.70-2.70)
Clashscore	141614	3122 (2.70-2.70)
Ramachandran outliers	138981	3069 (2.70-2.70)
Sidechain outliers	138945	3069 (2.70-2.70)
RSRZ outliers	127900	2737 (2.70-2.70)

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	136	4% 66% 5% 26%
1	E	136	% 66% 7% 24%
1	K	136	5% 70% 6% 22%
1	O	136	65% 8% 26%
2	B	103	4% 68% 11% 20%
2	F	103	4% 69% 11% 19%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	L	103	
2	P	103	
3	C	130	
3	G	130	
3	M	130	
3	Q	130	
4	D	126	
4	H	126	
4	N	126	
4	R	126	
5	I	169	
5	S	169	
6	J	169	
6	T	169	
7	U	213	
7	V	213	

2 Entry composition (i)

There are 11 unique types of molecules in this entry. The entry contains 28083 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	100	Total	C	N	O	S	0	0	0
			825	520	160	141	4			
1	E	103	Total	C	N	O	S	0	0	0
			840	529	163	144	4			
1	K	106	Total	C	N	O	S	0	0	0
			859	541	166	148	4			
1	O	100	Total	C	N	O	S	0	0	0
			825	520	160	141	4			

- Molecule 2 is a protein called Histone H4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	82	Total	C	N	O	S	0	0	0
			653	412	127	113	1			
2	F	83	Total	C	N	O	S	0	0	0
			662	418	129	114	1			
2	L	84	Total	C	N	O	S	0	0	0
			673	424	133	115	1			
2	P	102	Total	C	N	O	S	0	0	0
			792	494	163	134	1			

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

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	C	109	Total	C	N	O	0	0	0
			844	532	167	145			
3	G	104	Total	C	N	O	0	0	0
			805	508	157	140			
3	M	109	Total	C	N	O	0	0	0
			844	532	167	145			
3	Q	105	Total	C	N	O	0	0	0
			814	514	159	141			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	97	Total	C	N	O	S	0	0	0
			766	480	142	142	2			
4	H	102	Total	C	N	O	S	0	0	0
			805	504	150	149	2			
4	N	98	Total	C	N	O	S	0	0	0
			775	486	144	143	2			
4	R	98	Total	C	N	O	S	0	0	0
			775	486	144	143	2			

- Molecule 5 is a DNA chain called DNA (169-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	I	169	Total	C	N	O	P	0	0	0
			3462	1646	637	1011	168			
5	S	169	Total	C	N	O	P	0	0	0
			3462	1646	637	1011	168			

- Molecule 6 is a DNA chain called DNA (169-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	J	169	Total	C	N	O	P	0	0	0
			3461	1646	634	1013	168			
6	T	169	Total	C	N	O	P	0	0	0
			3461	1646	634	1013	168			

- Molecule 7 is a protein called Histone H1x.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
7	U	87	Total	C	N	O	0	0	0
			703	443	135	125			
7	V	94	Total	C	N	O	0	0	0
			752	471	145	136			

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

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	A	1	Total	Cl	0	0
			1	1		
8	C	1	Total	Cl	0	0
			1	1		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	E	1	Total Cl 1 1	0	0
8	G	1	Total Cl 1 1	0	0
8	K	1	Total Cl 1 1	0	0
8	M	1	Total Cl 1 1	0	0
8	O	1	Total Cl 1 1	0	0
8	R	1	Total Cl 1 1	0	0

- Molecule 9 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
9	C	3	Total Ca 3 3	0	0
9	G	3	Total Ca 3 3	0	0
9	I	8	Total Ca 8 8	0	0
9	J	9	Total Ca 9 9	0	0
9	N	1	Total Ca 1 1	0	0
9	Q	3	Total Ca 3 3	0	0
9	R	1	Total Ca 1 1	0	0
9	S	18	Total Ca 18 18	0	0
9	T	6	Total Ca 6 6	0	0

- Molecule 10 is POTASSIUM ION (three-letter code: K) (formula: K).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
10	C	1	Total K 1 1	0	0
10	I	2	Total K 2 2	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
10	J	1	Total K 1 1	0	0
10	M	1	Total K 1 1	0	0
10	S	2	Total K 2 2	0	0
10	T	1	Total K 1 1	0	0

- Molecule 11 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
11	A	1	Total O 1 1	0	0
11	B	5	Total O 5 5	0	0
11	C	4	Total O 4 4	0	0
11	D	3	Total O 3 3	0	0
11	E	4	Total O 4 4	0	0
11	F	3	Total O 3 3	0	0
11	G	4	Total O 4 4	0	0
11	H	3	Total O 3 3	0	0
11	I	3	Total O 3 3	0	0
11	J	7	Total O 7 7	0	0
11	K	11	Total O 11 11	0	0
11	L	12	Total O 12 12	0	0
11	M	11	Total O 11 11	0	0
11	N	10	Total O 10 10	0	0
11	O	14	Total O 14 14	0	0

Continued on next page...

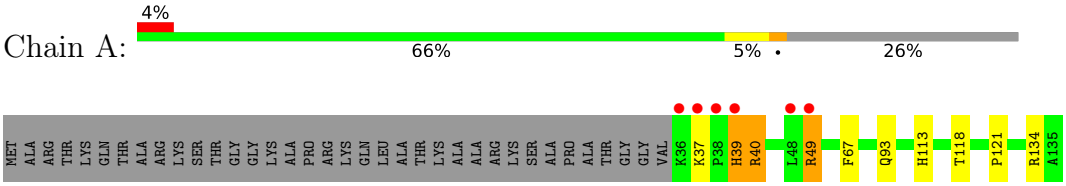
Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
11	P	12	Total O 12 12	0	0
11	Q	9	Total O 9 9	0	0
11	R	8	Total O 8 8	0	0
11	S	13	Total O 13 13	0	0
11	T	20	Total O 20 20	0	0

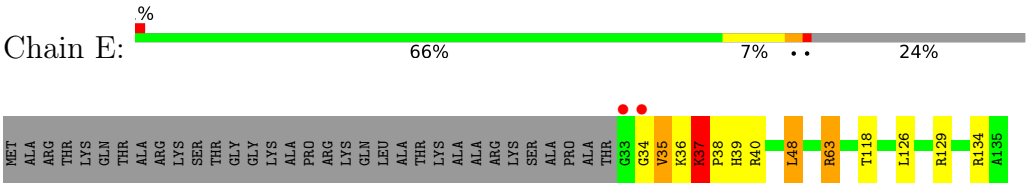
3 Residue-property plots (i)

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.

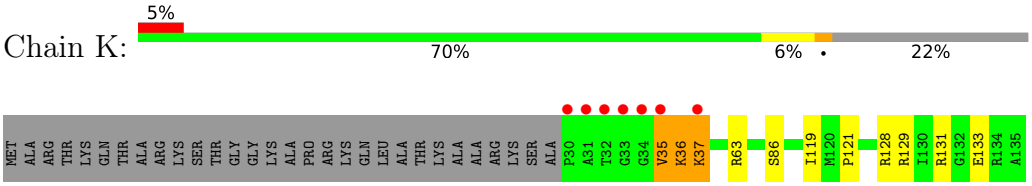
• Molecule 1: Histone H3.1



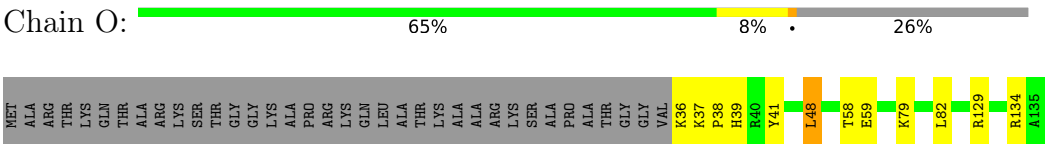
• Molecule 1: Histone H3.1



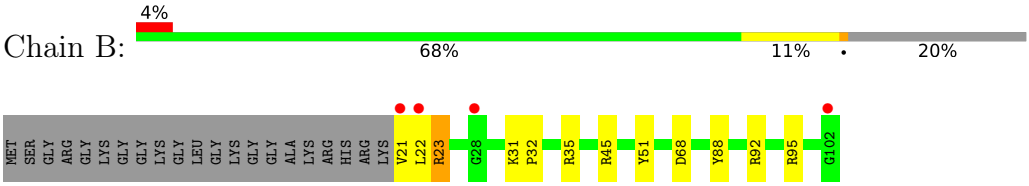
• Molecule 1: Histone H3.1



• Molecule 1: Histone H3.1

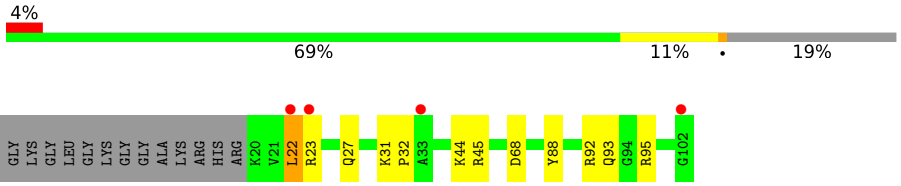


• Molecule 2: Histone H4



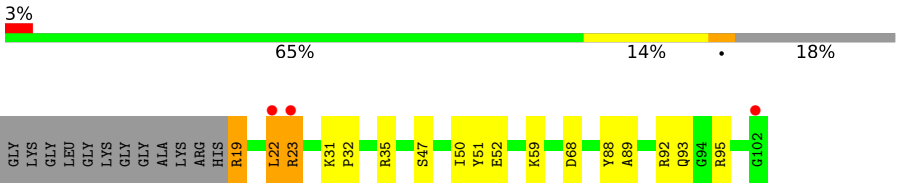
• Molecule 2: Histone H4

Chain F:



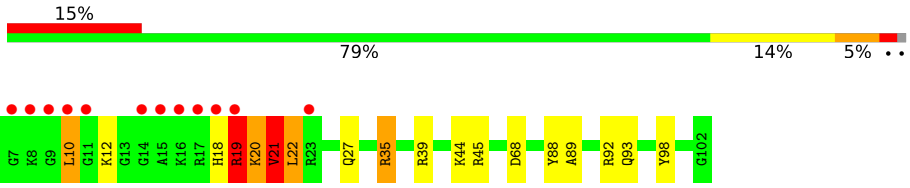
• Molecule 2: Histone H4

Chain L:



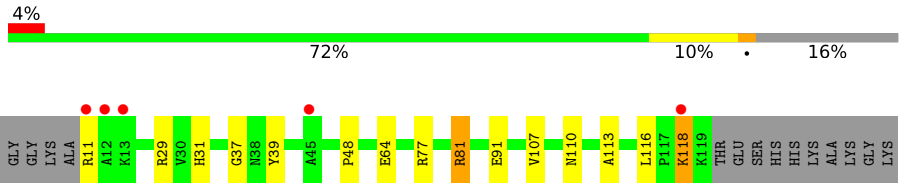
• Molecule 2: Histone H4

Chain P:



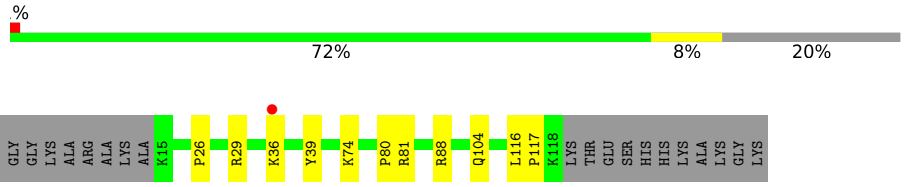
• Molecule 3: Histone H2A type 1-B/E

Chain C:



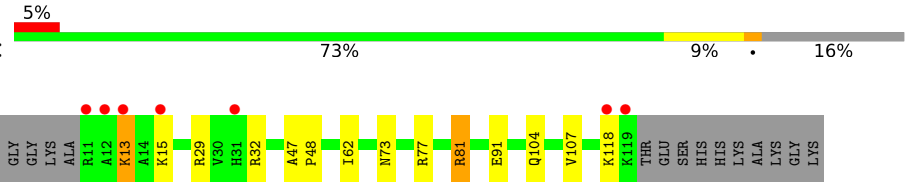
• Molecule 3: Histone H2A type 1-B/E

Chain G:



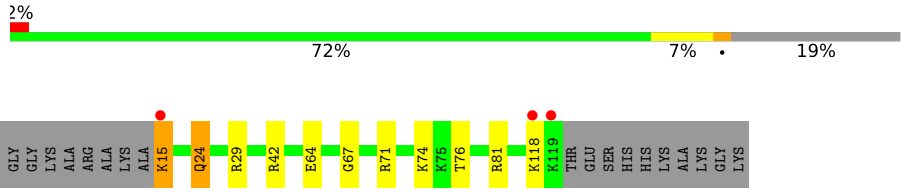
• Molecule 3: Histone H2A type 1-B/E

Chain M:



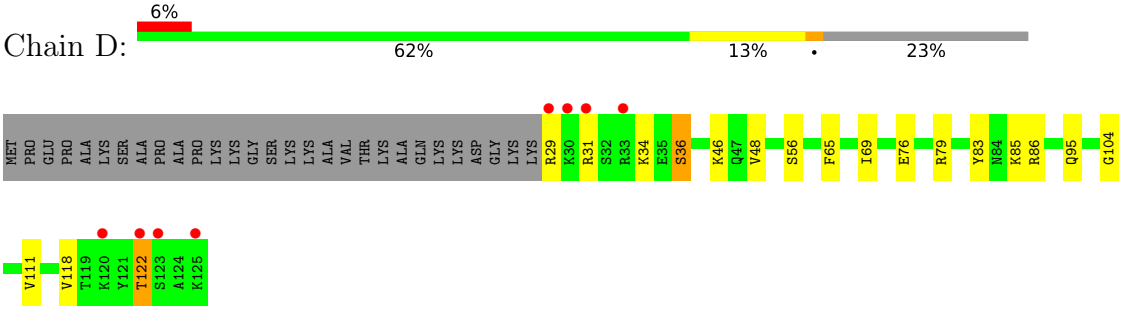
• Molecule 3: Histone H2A type 1-B/E

Chain Q:



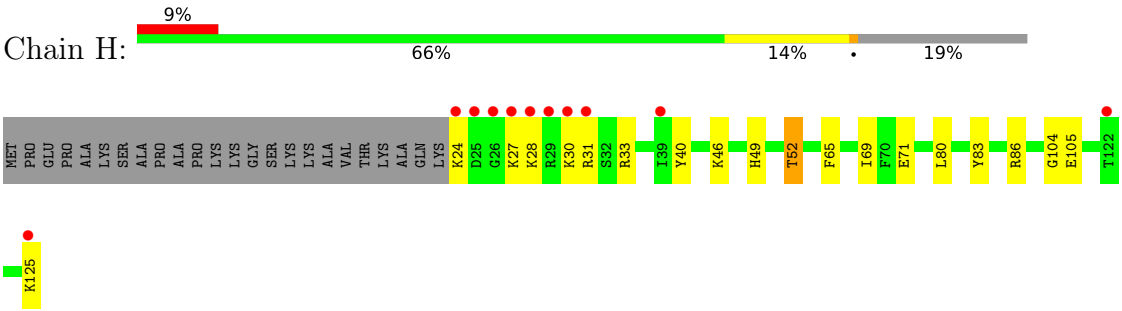
• Molecule 4: Histone H2B type 1-J

Chain D:



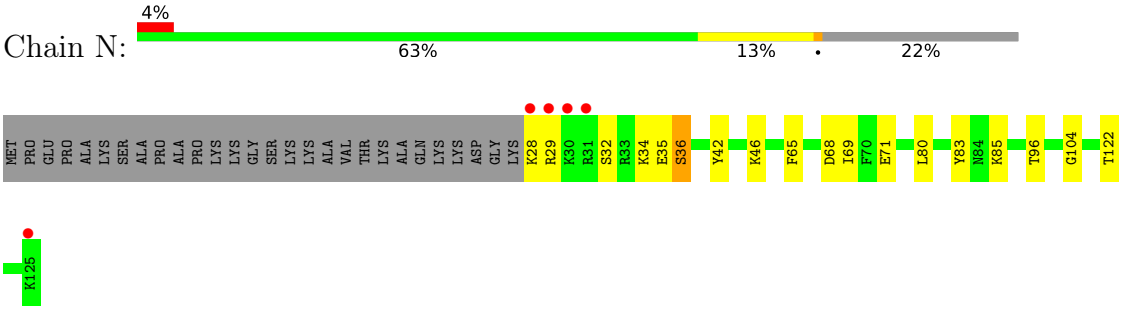
• Molecule 4: Histone H2B type 1-J

Chain H:



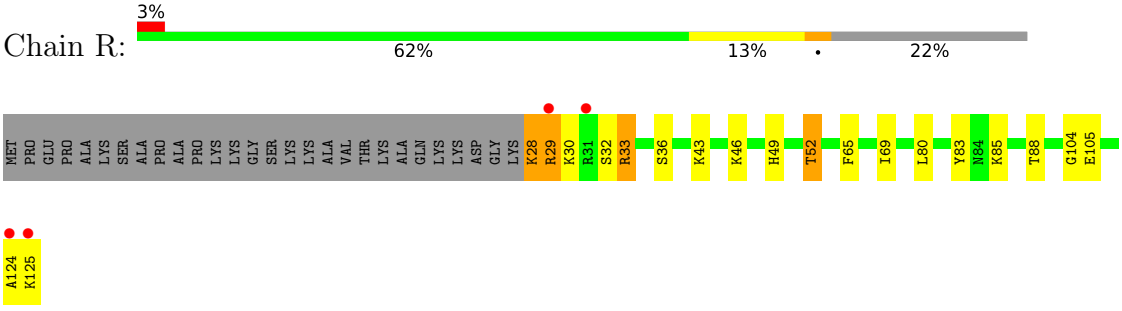
• Molecule 4: Histone H2B type 1-J

Chain N:



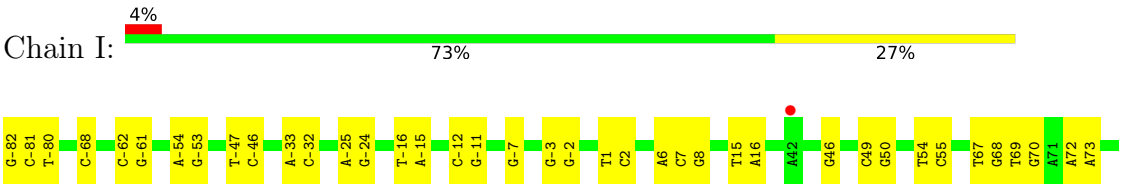
• Molecule 4: Histone H2B type 1-J

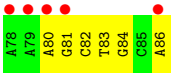
Chain R:



• Molecule 5: DNA (169-MER)

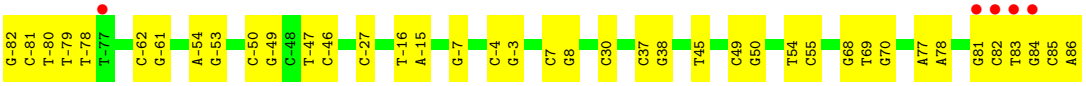
Chain I:





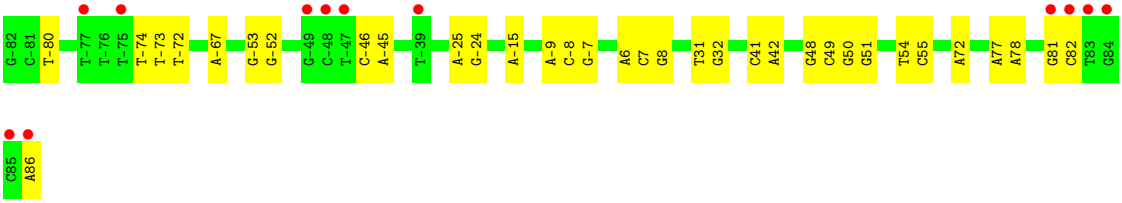
• Molecule 5: DNA (169-MER)

Chain S:



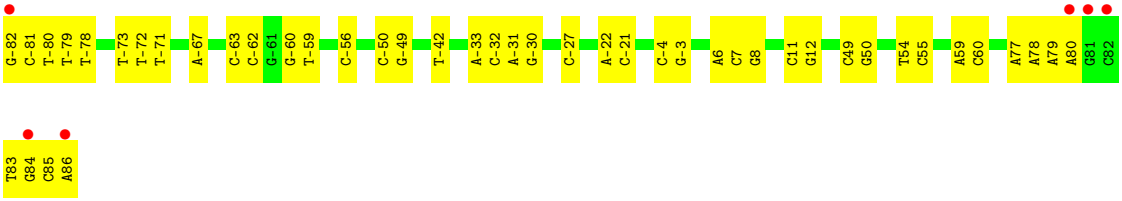
• Molecule 6: DNA (169-MER)

Chain J:



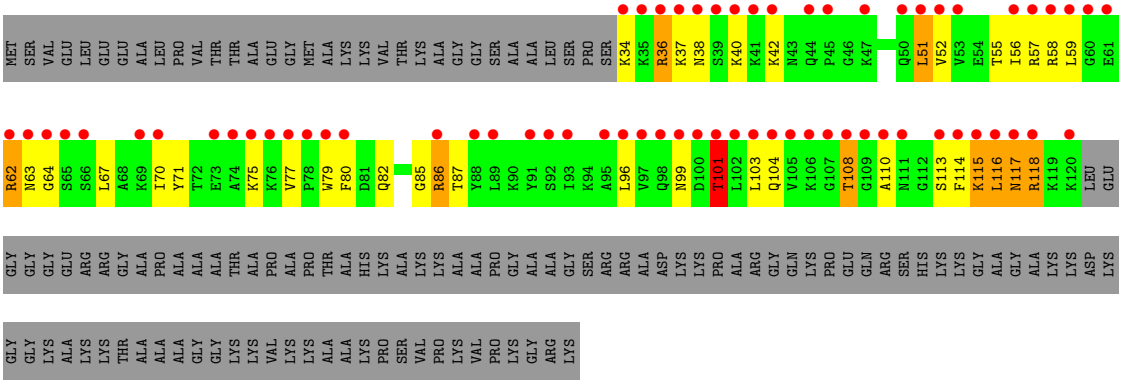
• Molecule 6: DNA (169-MER)

Chain T:



• Molecule 7: Histone H1x

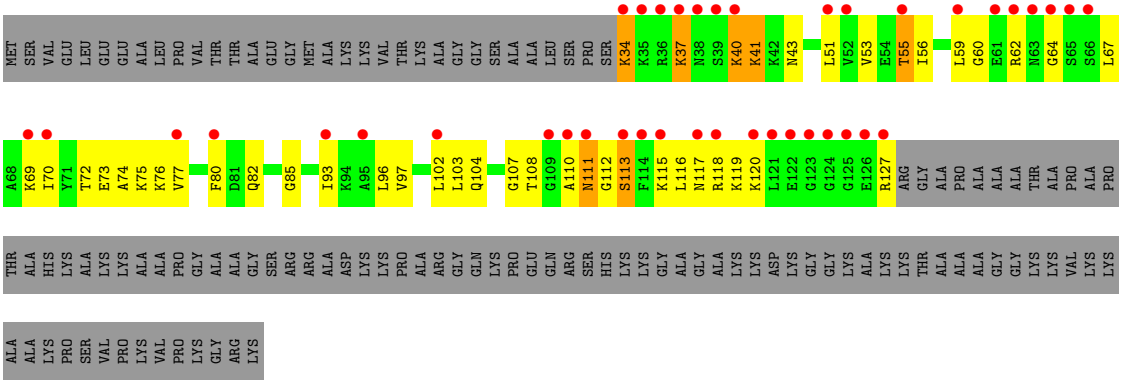
Chain U:



• Molecule 7: Histone H1x

Chain V:


19%
23%
17%
56%


MET SER VAL GLU LEU GLU GLU ALA LEU PRO VAL THR THR ALA GLU GLY MET ALA LYS LYS VAL THR LYS ALA GLY GLY SER ALA ALA LEU SER PRO SER
K34 K35 R36 K37 N38 S39 K40 K41 K42 N43 L51 V52 V53 E54 T55 I56 L59 G60 E61 R62 N63 G64 S65 S66 L67
A68 K69 I70 Y71 T72 E73 A74 K75 K76 V77 F80 D81 Q82 G85 I93 K94 A95 L96 V97 L102 L103 Q104 G107 T108 G109 A110 N111 G112 S113 F114 K115 L116 N117 R118 K119 K120 L121 E122 G123 G124 G125 E126 R127
ARG GLY ALA PRO ALA ALA ALA THR ALA PRO ALA PRO
THR ALA HIS LYS ALA LYS LYS ALA ALA PRO GLY ALA ALA GLY SER ARG ARG ALA ASP LYS LYS PRO ALA ARG GLY GLN LYS PRO GLU GLN ARG SER HIS LYS LYS GLY ALA GLY ALA LYS LYS ASP LYS GLY GLY LYS ALA ALA LYS LYS THR ALA ALA ALA GLY GLY LYS LYS VAL LYS LYS
ALA ALA LYS PRO SER VAL PRO LYS VAL PRO LYS GLY ARG LYS

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	104.85Å 102.11Å 216.18Å 90.00° 96.34° 90.00°	Depositor
Resolution (Å)	92.22 – 2.70 92.23 – 2.70	Depositor EDS
% Data completeness (in resolution range)	99.6 (92.22-2.70) 99.6 (92.23-2.70)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.65 (at 2.69Å)	Xtriage
Refinement program	REFMAC 5.8.0258	Depositor
R, R_{free}	0.210 , 0.267 0.212 , 0.268	Depositor DCC
R_{free} test set	2402 reflections (1.93%)	wwPDB-VP
Wilson B-factor (Å^2)	67.9	Xtriage
Anisotropy	0.301	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 61.3	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	28083	wwPDB-VP
Average B, all atoms (Å^2)	95.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 34.70 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 6.5434e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

[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: K, CL, CA

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	0.62	0/837	0.80	0/1120
1	E	0.63	0/852	0.80	0/1140
1	K	0.64	0/872	0.84	0/1168
1	O	0.64	0/837	0.82	0/1120
2	B	0.65	0/660	0.80	0/883
2	F	0.66	0/669	0.83	0/894
2	L	0.62	0/680	0.82	0/908
2	P	0.70	0/800	0.92	1/1061 (0.1%)
3	C	0.65	0/854	0.79	0/1150
3	G	0.63	0/815	0.77	0/1100
3	M	0.66	0/854	0.86	0/1150
3	Q	0.64	0/824	0.79	0/1111
4	D	0.65	0/777	0.74	0/1040
4	H	0.66	0/816	0.78	0/1089
4	N	0.66	0/786	0.79	0/1051
4	R	0.65	0/786	0.80	0/1051
5	I	0.36	0/3884	0.78	0/5993
5	S	0.43	0/3884	0.79	0/5993
6	J	0.38	0/3882	0.77	0/5990
6	T	0.44	0/3882	0.81	0/5990
7	U	0.69	0/712	0.89	0/947
7	V	0.73	0/761	0.87	0/1011
All	All	0.54	0/29724	0.80	1/42960 (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	K	0	1

Continued on next page...

Continued from previous page...

Mol	Chain	#Chirality outliers	#Planarity outliers
2	P	0	2
7	U	0	1
All	All	0	4

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	P	35	ARG	CB-CA-C	5.27	120.93	110.40

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	K	36	LYS	Peptide
2	P	18	HIS	Peptide
2	P	21	VAL	Peptide
7	U	101	THR	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	825	0	869	6	0
1	E	840	0	884	13	0
1	K	859	0	904	9	0
1	O	825	0	869	6	0
2	B	653	0	696	9	0
2	F	662	0	709	9	0
2	L	673	0	722	14	0
2	P	792	0	855	16	0
3	C	844	0	910	10	0
3	G	805	0	861	5	0
3	M	844	0	910	12	0
3	Q	814	0	874	6	0
4	D	766	0	797	12	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	H	805	0	843	11	0
4	N	775	0	810	16	0
4	R	775	0	810	18	0
5	I	3462	0	1901	38	0
5	S	3462	0	1901	33	0
6	J	3461	0	1902	31	0
6	T	3461	0	1902	35	0
7	U	703	0	751	27	0
7	V	752	0	796	26	0
8	A	1	0	0	1	0
8	C	1	0	0	0	0
8	E	1	0	0	0	0
8	G	1	0	0	0	0
8	K	1	0	0	1	0
8	M	1	0	0	0	0
8	O	1	0	0	0	0
8	R	1	0	0	0	0
9	C	3	0	0	0	0
9	G	3	0	0	0	0
9	I	8	0	0	0	0
9	J	9	0	0	0	0
9	N	1	0	0	0	0
9	Q	3	0	0	0	0
9	R	1	0	0	0	0
9	S	18	0	0	0	0
9	T	6	0	0	0	0
10	C	1	0	0	0	0
10	I	2	0	0	0	0
10	J	1	0	0	0	0
10	M	1	0	0	0	0
10	S	2	0	0	0	0
10	T	1	0	0	0	0
11	A	1	0	0	0	0
11	B	5	0	0	1	0
11	C	4	0	0	0	0
11	D	3	0	0	0	0
11	E	4	0	0	0	0
11	F	3	0	0	0	0
11	G	4	0	0	0	0
11	H	3	0	0	0	0
11	I	3	0	0	0	0
11	J	7	0	0	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
11	K	11	0	0	0	0
11	L	12	0	0	1	0
11	M	11	0	0	0	0
11	N	10	0	0	0	0
11	O	14	0	0	0	0
11	P	12	0	0	0	0
11	Q	9	0	0	0	0
11	R	8	0	0	0	0
11	S	13	0	0	1	0
11	T	20	0	0	2	0
All	All	28083	0	22476	301	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 6.

The worst 5 of 301 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)
4:H:49:HIS:HB3	4:H:52:THR:HG23	1.42	0.99
6:J:-74:DT:H2"	6:J:-73:DT:O5'	1.61	0.99
7:V:37:LYS:HE2	7:V:127:ARG:HD3	1.45	0.96
4:R:49:HIS:HB3	4:R:52:THR:HG23	1.47	0.95
4:H:49:HIS:HB3	4:H:52:THR:CG2	1.99	0.93

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	98/136 (72%)	95 (97%)	2 (2%)	1 (1%)	15 37
1	E	101/136 (74%)	98 (97%)	1 (1%)	2 (2%)	7 19

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	K	104/136 (76%)	101 (97%)	1 (1%)	2 (2%)	8 20
1	O	98/136 (72%)	97 (99%)	1 (1%)	0	100 100
2	B	80/103 (78%)	75 (94%)	5 (6%)	0	100 100
2	F	81/103 (79%)	76 (94%)	5 (6%)	0	100 100
2	L	82/103 (80%)	80 (98%)	2 (2%)	0	100 100
2	P	100/103 (97%)	90 (90%)	7 (7%)	3 (3%)	4 10
3	C	107/130 (82%)	104 (97%)	3 (3%)	0	100 100
3	G	102/130 (78%)	97 (95%)	5 (5%)	0	100 100
3	M	107/130 (82%)	103 (96%)	4 (4%)	0	100 100
3	Q	103/130 (79%)	101 (98%)	2 (2%)	0	100 100
4	D	95/126 (75%)	93 (98%)	1 (1%)	1 (1%)	14 34
4	H	100/126 (79%)	97 (97%)	2 (2%)	1 (1%)	15 37
4	N	96/126 (76%)	93 (97%)	2 (2%)	1 (1%)	15 37
4	R	96/126 (76%)	90 (94%)	3 (3%)	3 (3%)	4 9
7	U	85/213 (40%)	72 (85%)	10 (12%)	3 (4%)	3 8
7	V	92/213 (43%)	79 (86%)	11 (12%)	2 (2%)	6 17
All	All	1727/2406 (72%)	1641 (95%)	67 (4%)	19 (1%)	14 34

5 of 19 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	39	HIS
1	E	37	LYS
1	K	35	VAL
1	K	37	LYS
2	P	21	VAL

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	87/111 (78%)	82 (94%)	5 (6%)	20 44
1	E	88/111 (79%)	82 (93%)	6 (7%)	16 36
1	K	90/111 (81%)	85 (94%)	5 (6%)	21 45
1	O	87/111 (78%)	83 (95%)	4 (5%)	27 54
2	B	67/79 (85%)	65 (97%)	2 (3%)	41 70
2	F	68/79 (86%)	65 (96%)	3 (4%)	28 56
2	L	69/79 (87%)	62 (90%)	7 (10%)	7 17
2	P	78/79 (99%)	72 (92%)	6 (8%)	13 30
3	C	86/100 (86%)	81 (94%)	5 (6%)	20 43
3	G	83/100 (83%)	79 (95%)	4 (5%)	25 53
3	M	86/100 (86%)	81 (94%)	5 (6%)	20 43
3	Q	84/100 (84%)	76 (90%)	8 (10%)	8 20
4	D	83/105 (79%)	75 (90%)	8 (10%)	8 19
4	H	87/105 (83%)	76 (87%)	11 (13%)	4 10
4	N	84/105 (80%)	77 (92%)	7 (8%)	11 25
4	R	84/105 (80%)	76 (90%)	8 (10%)	8 20
7	U	76/157 (48%)	56 (74%)	20 (26%)	0 1
7	V	80/157 (51%)	59 (74%)	21 (26%)	0 1
All	All	1467/1894 (78%)	1332 (91%)	135 (9%)	9 21

5 of 135 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
7	V	37	LYS
7	V	51	LEU
7	V	113	SER
1	K	129	ARG
1	K	86	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 18 such sidechains are listed below:

Mol	Chain	Res	Type
7	V	82	GLN
7	V	111	ASN
7	V	104	GLN
3	M	31	HIS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
7	U	117	ASN

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 monosaccharides in this entry.

5.6 Ligand geometry

Of 68 ligands modelled in this entry, 68 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	100/136 (73%)	0.36	6 (6%) 21 20	55, 76, 115, 146	0
1	E	103/136 (75%)	0.39	2 (1%) 66 69	54, 69, 132, 159	0
1	K	106/136 (77%)	0.49	7 (6%) 18 16	38, 52, 137, 162	0
1	O	100/136 (73%)	0.36	0 100 100	36, 51, 91, 138	0
2	B	82/103 (79%)	0.38	4 (4%) 29 28	55, 71, 100, 148	0
2	F	83/103 (80%)	0.45	4 (4%) 30 28	52, 68, 91, 152	0
2	L	84/103 (81%)	0.45	3 (3%) 42 42	38, 50, 109, 154	0
2	P	102/103 (99%)	1.06	15 (14%) 2 1	39, 53, 162, 184	0
3	C	109/130 (83%)	0.59	5 (4%) 32 31	57, 75, 110, 135	0
3	G	104/130 (80%)	0.31	1 (0%) 82 83	53, 74, 104, 110	0
3	M	109/130 (83%)	0.68	7 (6%) 19 18	43, 55, 94, 128	0
3	Q	105/130 (80%)	0.42	3 (2%) 51 52	41, 55, 85, 123	0
4	D	97/126 (76%)	0.86	8 (8%) 11 9	55, 77, 112, 149	0
4	H	102/126 (80%)	0.82	11 (10%) 5 4	57, 74, 138, 161	0
4	N	98/126 (77%)	0.58	5 (5%) 28 26	40, 57, 103, 138	0
4	R	98/126 (77%)	0.55	4 (4%) 37 36	42, 55, 117, 145	0
5	I	169/169 (100%)	-0.05	6 (3%) 42 42	76, 121, 190, 219	0
5	S	169/169 (100%)	-0.19	5 (2%) 50 51	58, 93, 181, 205	0
6	J	169/169 (100%)	0.02	12 (7%) 16 14	80, 121, 190, 226	0
6	T	169/169 (100%)	-0.22	6 (3%) 42 42	59, 93, 182, 216	0
7	U	87/213 (40%)	3.75	67 (77%) 0 0	115, 145, 170, 202	0
7	V	94/213 (44%)	2.09	40 (42%) 0 0	88, 120, 158, 164	0
All	All	2439/3082 (79%)	0.54	221 (9%) 9 7	36, 77, 158, 226	0

The worst 5 of 221 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
4	H	25	ASP	10.9
7	U	111	ASN	10.3
4	D	29	ARG	8.8
4	H	31	ARG	8.5
7	U	110	ALA	8.1

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 monosaccharides 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
9	CA	I	107	1/1	0.15	0.35	123,123,123,123	0
9	CA	S	104	1/1	0.39	0.13	120,120,120,120	0
9	CA	G	203	1/1	0.52	0.18	132,132,132,132	0
9	CA	S	116	1/1	0.62	0.16	106,106,106,106	0
9	CA	S	106	1/1	0.65	0.10	110,110,110,110	0
9	CA	I	104	1/1	0.65	0.18	123,123,123,123	0
9	CA	R	201	1/1	0.66	0.14	117,117,117,117	0
9	CA	I	105	1/1	0.67	0.11	133,133,133,133	0
9	CA	I	103	1/1	0.67	0.15	121,121,121,121	0
9	CA	T	103	1/1	0.67	0.23	122,122,122,122	0
9	CA	S	112	1/1	0.71	0.27	118,118,118,118	0
10	K	I	110	1/1	0.71	0.15	113,113,113,113	0
9	CA	S	118	1/1	0.72	0.11	121,121,121,121	0
9	CA	Q	203	1/1	0.73	0.25	103,103,103,103	0
9	CA	Q	201	1/1	0.75	0.16	114,114,114,114	0
9	CA	C	201	1/1	0.78	0.32	103,103,103,103	0
9	CA	S	115	1/1	0.78	0.10	114,114,114,114	0
8	CL	A	201	1/1	0.81	0.16	98,98,98,98	0
9	CA	C	202	1/1	0.82	0.18	114,114,114,114	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
9	CA	G	201	1/1	0.82	0.12	121,121,121,121	0
9	CA	T	104	1/1	0.83	0.20	110,110,110,110	0
9	CA	T	102	1/1	0.83	0.22	113,113,113,113	0
9	CA	J	108	1/1	0.84	0.19	111,111,111,111	0
9	CA	I	101	1/1	0.85	0.22	89,89,89,89	0
10	K	S	119	1/1	0.85	0.14	100,100,100,100	0
9	CA	S	107	1/1	0.86	0.16	98,98,98,98	0
9	CA	C	203	1/1	0.87	0.18	115,115,115,115	0
8	CL	C	204	1/1	0.87	0.18	91,91,91,91	0
9	CA	J	101	1/1	0.87	0.09	106,106,106,106	0
9	CA	J	104	1/1	0.87	0.21	114,114,114,114	0
9	CA	S	109	1/1	0.89	0.22	114,114,114,114	0
9	CA	Q	202	1/1	0.89	0.27	101,101,101,101	0
10	K	C	205	1/1	0.90	0.12	105,105,105,105	0
9	CA	I	108	1/1	0.90	0.11	102,102,102,102	0
10	K	J	110	1/1	0.90	0.10	99,99,99,99	0
10	K	M	202	1/1	0.90	0.12	92,92,92,92	0
9	CA	S	105	1/1	0.90	0.20	110,110,110,110	0
10	K	S	120	1/1	0.90	0.06	106,106,106,106	0
9	CA	S	110	1/1	0.91	0.22	114,114,114,114	0
8	CL	E	201	1/1	0.92	0.29	82,82,82,82	0
9	CA	J	106	1/1	0.92	0.07	122,122,122,122	0
9	CA	S	114	1/1	0.92	0.11	91,91,91,91	0
9	CA	J	109	1/1	0.93	0.07	111,111,111,111	0
9	CA	T	106	1/1	0.93	0.12	107,107,107,107	0
9	CA	S	111	1/1	0.93	0.18	102,102,102,102	0
8	CL	R	202	1/1	0.93	0.14	65,65,65,65	0
9	CA	J	107	1/1	0.94	0.12	123,123,123,123	0
9	CA	S	102	1/1	0.94	0.25	75,75,75,75	0
9	CA	I	106	1/1	0.94	0.17	93,93,93,93	0
8	CL	O	201	1/1	0.94	0.20	71,71,71,71	0
9	CA	S	113	1/1	0.95	0.15	113,113,113,113	0
9	CA	S	117	1/1	0.95	0.13	114,114,114,114	0
9	CA	T	105	1/1	0.95	0.16	86,86,86,86	0
8	CL	K	201	1/1	0.95	0.14	76,76,76,76	0
9	CA	J	103	1/1	0.95	0.13	102,102,102,102	0
10	K	T	107	1/1	0.95	0.14	95,95,95,95	0
9	CA	J	105	1/1	0.96	0.50	125,125,125,125	0
9	CA	G	202	1/1	0.96	0.15	107,107,107,107	0
9	CA	S	108	1/1	0.96	0.18	85,85,85,85	0
8	CL	G	204	1/1	0.96	0.06	75,75,75,75	0
8	CL	M	201	1/1	0.96	0.18	63,63,63,63	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
9	CA	S	103	1/1	0.97	0.05	99,99,99,99	0
9	CA	S	101	1/1	0.97	0.08	81,81,81,81	0
9	CA	T	101	1/1	0.97	0.21	100,100,100,100	0
9	CA	I	102	1/1	0.98	0.06	94,94,94,94	0
10	K	I	109	1/1	0.98	0.04	91,91,91,91	0
9	CA	N	201	1/1	0.98	0.23	94,94,94,94	0
9	CA	J	102	1/1	0.98	0.22	118,118,118,118	0

6.5 Other polymers (i)

There are no such residues in this entry.