Full wwPDB X-ray Structure Validation Report ⓘ

Mar 26, 2024 – 10:14 AM JST

PDB ID	:	7XVL
Title	:	Crystal Structure of Nucleosome-H1.0 Linker Histone Assembly (sticky-169an DNA fragment)
Authors	:	Adhireksan, Z.; Qiuye, B.; Lee, P.L.; Sharma, D.; Padavattan, S.; Davey, C.A.
Deposited on	:	2022-05-24
Resolution	:	3.51 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the ⓘ 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 ⓘ) were used in the production of this report:

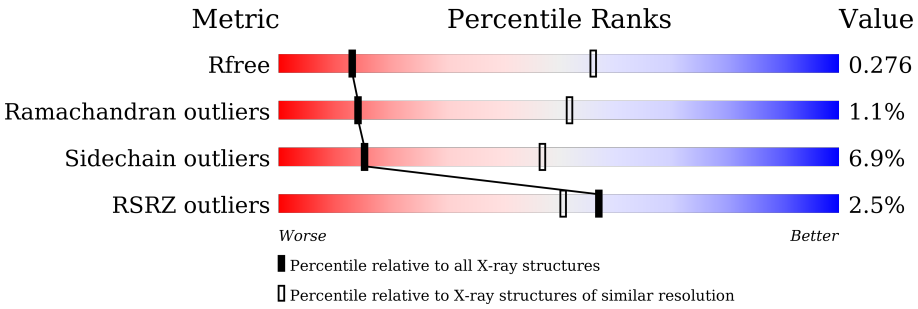
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation 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 3.51 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	1659 (3.60-3.40)
Ramachandran outliers	138981	1005 (3.58-3.42)
Sidechain outliers	138945	1006 (3.58-3.42)
RSRZ outliers	127900	1559 (3.60-3.40)

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 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	138	69% · 29%
1	E	138	% 67% · 29%
1	K	138	69% · 29%
1	O	138	67% · 29%
1	U	138	% 68% · 29%
1	Y	138	70% · 29%
1	e	138	66% 5% 29%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	i	138	67% 29%
2	B	105	74% 25%
2	F	105	81% 17%
2	L	105	71% 25%
2	P	105	74% 21%
2	V	105	4% 70% 6% 25%
2	Z	105	75% 6% 17%
2	f	105	71% 25%
2	j	105	73% 24%
3	C	132	80% 18%
3	G	132	76% 20%
3	M	132	77% 5% 18%
3	Q	132	73% 6% 20%
3	W	132	77% 5% 18%
3	a	132	11% 75% 21%
3	g	132	77% 20%
3	k	132	71% 7% 21%
4	D	128	67% 8% 25%
4	H	128	65% 9% 25%
4	N	128	71% 25%
4	R	128	65% 9% 25%
4	X	128	66% 9% 25%
4	b	128	14% 66% 9% 25%
4	h	128	68% 7% 25%
4	l	128	69% 6% 25%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
5	I	169	96%
5	S	169	97%
5	c	169	98%
5	m	169	99%
6	J	169	95% 5%
6	T	169	96%
6	d	169	98%
6	n	169	96%
7	o	195	28% 32% 8% 60%

2 Entry composition (i)

There are 7 unique types of molecules in this entry. The entry contains 52574 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	98	Total	C	N	O	S	0	0	0
			807	508	156	139	4			
1	E	98	Total	C	N	O	S	0	0	0
			807	508	156	139	4			
1	K	98	Total	C	N	O	S	0	0	0
			807	508	156	139	4			
1	O	98	Total	C	N	O	S	0	0	0
			807	508	156	139	4			
1	U	98	Total	C	N	O	S	0	0	0
			807	508	156	139	4			
1	Y	98	Total	C	N	O	S	0	0	0
			807	508	156	139	4			
1	e	98	Total	C	N	O	S	0	0	0
			807	508	156	139	4			
1	i	98	Total	C	N	O	S	0	0	0
			807	508	156	139	4			

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	GLY	-	expression tag	UNP P68431
A	-1	SER	-	expression tag	UNP P68431
E	-2	GLY	-	expression tag	UNP P68431
E	-1	SER	-	expression tag	UNP P68431
K	-2	GLY	-	expression tag	UNP P68431
K	-1	SER	-	expression tag	UNP P68431
O	-2	GLY	-	expression tag	UNP P68431
O	-1	SER	-	expression tag	UNP P68431
U	-2	GLY	-	expression tag	UNP P68431
U	-1	SER	-	expression tag	UNP P68431
Y	-2	GLY	-	expression tag	UNP P68431
Y	-1	SER	-	expression tag	UNP P68431
e	-2	GLY	-	expression tag	UNP P68431

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
e	-1	SER	-	expression tag	UNP P68431
i	-2	GLY	-	expression tag	UNP P68431
i	-1	SER	-	expression tag	UNP P68431

- 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	87	Total	C	N	O	S	0	0	0
			703	442	142	118	1			
2	L	79	Total	C	N	O	S	0	0	0
			627	395	121	110	1			
2	P	83	Total	C	N	O	S	0	0	0
			662	418	129	114	1			
2	V	79	Total	C	N	O	S	0	0	0
			627	395	121	110	1			
2	Z	87	Total	C	N	O	S	0	0	0
			703	442	142	118	1			
2	f	79	Total	C	N	O	S	0	0	0
			627	395	121	110	1			
2	j	80	Total	C	N	O	S	0	0	0
			638	401	125	111	1			

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	-2	GLY	-	expression tag	UNP P62805
B	-1	SER	-	expression tag	UNP P62805
F	-2	GLY	-	expression tag	UNP P62805
F	-1	SER	-	expression tag	UNP P62805
L	-2	GLY	-	expression tag	UNP P62805
L	-1	SER	-	expression tag	UNP P62805
P	-2	GLY	-	expression tag	UNP P62805
P	-1	SER	-	expression tag	UNP P62805
V	-2	GLY	-	expression tag	UNP P62805
V	-1	SER	-	expression tag	UNP P62805
Z	-2	GLY	-	expression tag	UNP P62805
Z	-1	SER	-	expression tag	UNP P62805
f	-2	GLY	-	expression tag	UNP P62805
f	-1	SER	-	expression tag	UNP P62805
j	-2	GLY	-	expression tag	UNP P62805

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
j	-1	SER	-	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	108	Total 835	C 526	N 165	O 144	0	0	0
3	G	105	Total 810	C 511	N 158	O 141	0	0	0
3	M	108	Total 835	C 526	N 165	O 144	0	0	0
3	Q	105	Total 810	C 511	N 158	O 141	0	0	0
3	W	108	Total 835	C 526	N 165	O 144	0	0	0
3	a	104	Total 805	C 508	N 157	O 140	0	0	0
3	g	105	Total 810	C 511	N 158	O 141	0	0	0
3	k	104	Total 805	C 508	N 157	O 140	0	0	0

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	-2	GLY	-	expression tag	UNP P04908
C	-1	SER	-	expression tag	UNP P04908
G	-2	GLY	-	expression tag	UNP P04908
G	-1	SER	-	expression tag	UNP P04908
M	-2	GLY	-	expression tag	UNP P04908
M	-1	SER	-	expression tag	UNP P04908
Q	-2	GLY	-	expression tag	UNP P04908
Q	-1	SER	-	expression tag	UNP P04908
W	-2	GLY	-	expression tag	UNP P04908
W	-1	SER	-	expression tag	UNP P04908
a	-2	GLY	-	expression tag	UNP P04908
a	-1	SER	-	expression tag	UNP P04908
g	-2	GLY	-	expression tag	UNP P04908
g	-1	SER	-	expression tag	UNP P04908
k	-2	GLY	-	expression tag	UNP P04908
k	-1	SER	-	expression tag	UNP P04908

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	96	Total	C	N	O	S	0	0	0
			755	474	138	141	2			
4	H	96	Total	C	N	O	S	0	0	0
			755	474	138	141	2			
4	N	96	Total	C	N	O	S	0	0	0
			755	474	138	141	2			
4	R	96	Total	C	N	O	S	0	0	0
			755	474	138	141	2			
4	X	96	Total	C	N	O	S	0	0	0
			755	474	138	141	2			
4	b	96	Total	C	N	O	S	0	0	0
			755	474	138	141	2			
4	h	96	Total	C	N	O	S	0	0	0
			755	474	138	141	2			
4	l	96	Total	C	N	O	S	0	0	0
			755	474	138	141	2			

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	-2	GLY	-	expression tag	UNP P06899
D	-1	SER	-	expression tag	UNP P06899
H	-2	GLY	-	expression tag	UNP P06899
H	-1	SER	-	expression tag	UNP P06899
N	-2	GLY	-	expression tag	UNP P06899
N	-1	SER	-	expression tag	UNP P06899
R	-2	GLY	-	expression tag	UNP P06899
R	-1	SER	-	expression tag	UNP P06899
X	-2	GLY	-	expression tag	UNP P06899
X	-1	SER	-	expression tag	UNP P06899
b	-2	GLY	-	expression tag	UNP P06899
b	-1	SER	-	expression tag	UNP P06899
h	-2	GLY	-	expression tag	UNP P06899
h	-1	SER	-	expression tag	UNP P06899
l	-2	GLY	-	expression tag	UNP P06899
l	-1	SER	-	expression tag	UNP P06899

- 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
			3457	1639	656	993	169			

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	S	169	Total	C	N	O	P	0	0	0
			3457	1639	656	993	169			
5	c	169	Total	C	N	O	P	0	0	0
			3457	1639	656	993	169			
5	m	169	Total	C	N	O	P	0	0	0
			3457	1639	656	993	169			

- 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
			3474	1652	616	1037	169			
6	T	169	Total	C	N	O	P	0	0	0
			3474	1652	616	1037	169			
6	d	169	Total	C	N	O	P	0	0	0
			3474	1652	616	1037	169			
6	n	169	Total	C	N	O	P	0	0	0
			3474	1652	616	1037	169			

- Molecule 7 is a protein called Histone H1.0.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	o	78	Total	C	N	O	S	0	0	0
			595	368	111	115	1			

There are 2 discrepancies between the modelled and reference sequences:

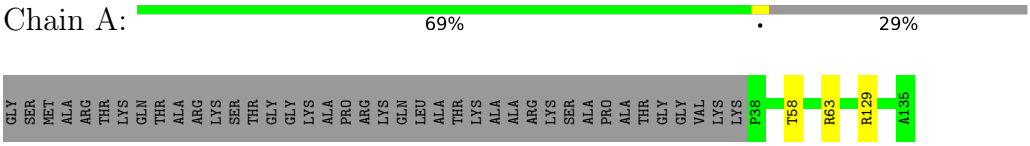
Chain	Residue	Modelled	Actual	Comment	Reference
o	0	GLY	-	expression tag	UNP P07305
o	1	PRO	-	expression tag	UNP P07305

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.

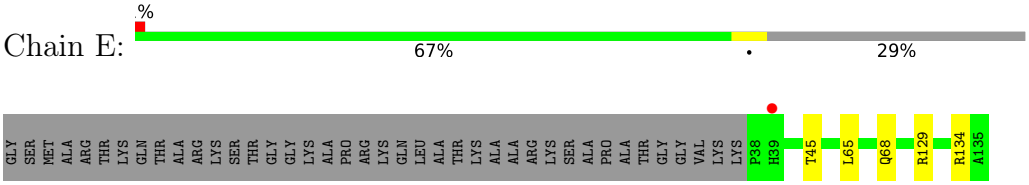
• Molecule 1: Histone H3.1

Chain A:



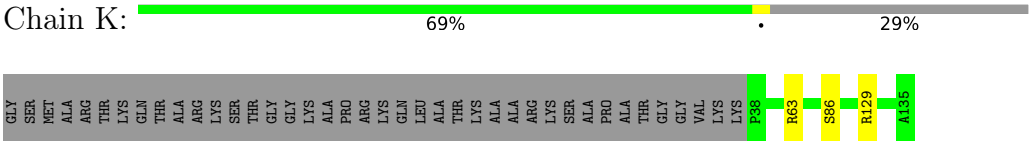
• Molecule 1: Histone H3.1

Chain E:



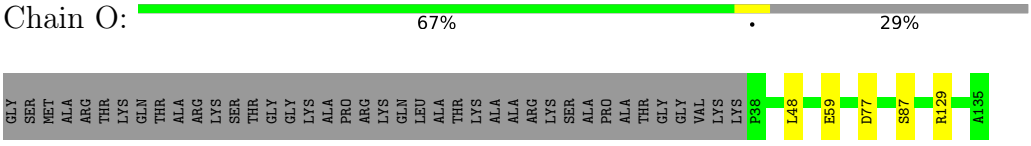
• Molecule 1: Histone H3.1

Chain K:



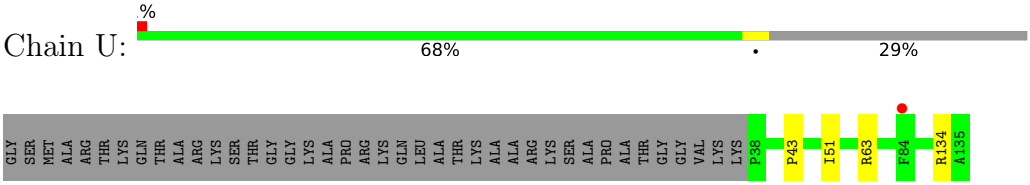
• Molecule 1: Histone H3.1

Chain O:



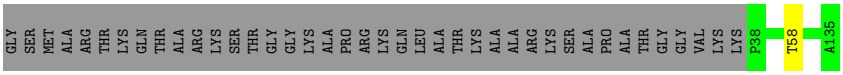
• Molecule 1: Histone H3.1

Chain U:



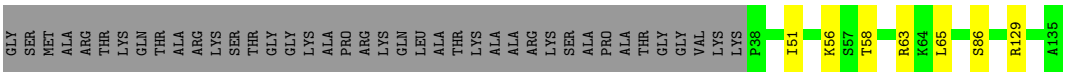
• Molecule 1: Histone H3.1

Chain Y:



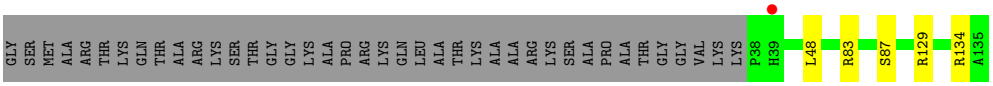
• Molecule 1: Histone H3.1

Chain e:



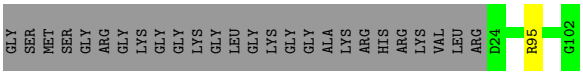
• Molecule 1: Histone H3.1

Chain i:



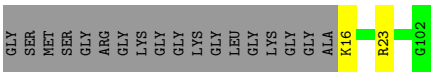
• Molecule 2: Histone H4

Chain B:



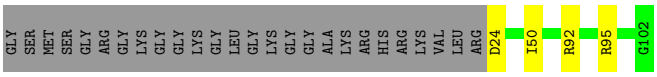
• Molecule 2: Histone H4

Chain F:



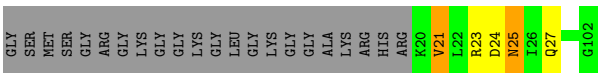
• Molecule 2: Histone H4

Chain L:



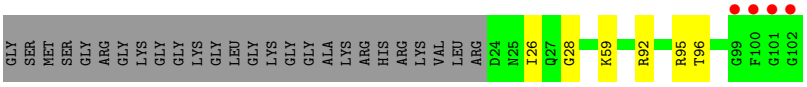
• Molecule 2: Histone H4

Chain P:

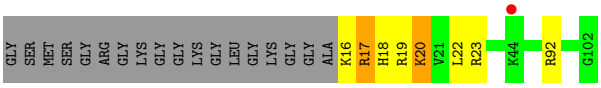


• Molecule 2: Histone H4

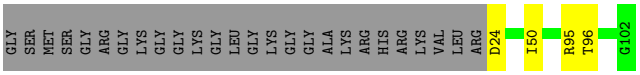
Chain V:



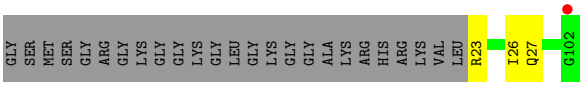
• Molecule 2: Histone H4



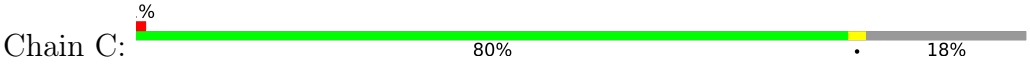
• Molecule 2: Histone H4

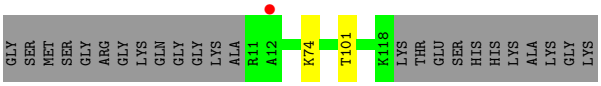


• Molecule 2: Histone H4

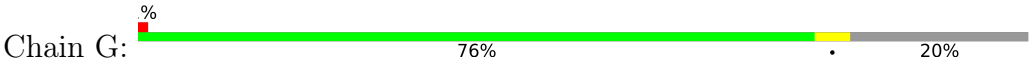


• Molecule 3: Histone H2A type 1-B/E

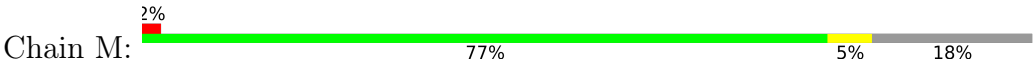


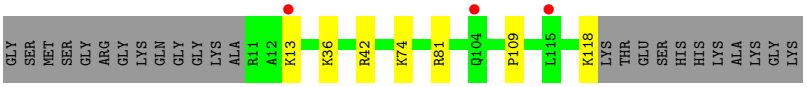


• Molecule 3: Histone H2A type 1-B/E



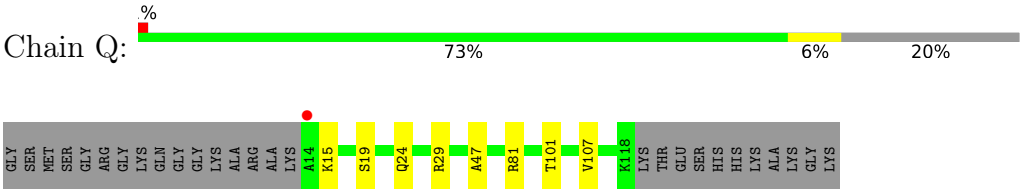
• Molecule 3: Histone H2A type 1-B/E





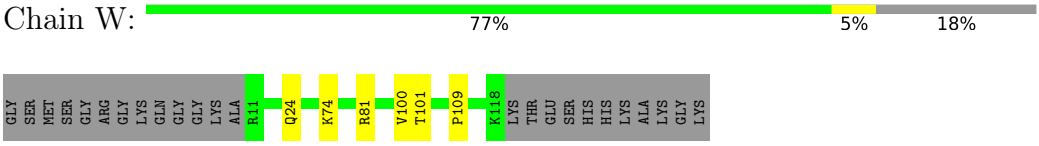
• Molecule 3: Histone H2A type 1-B/E

Chain Q:



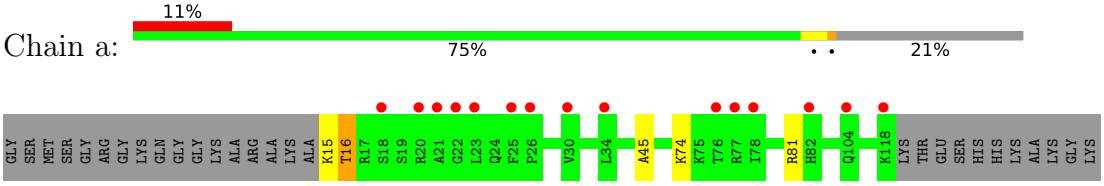
• Molecule 3: Histone H2A type 1-B/E

Chain W:



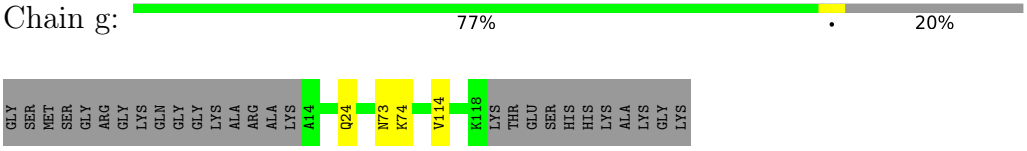
• Molecule 3: Histone H2A type 1-B/E

Chain a:



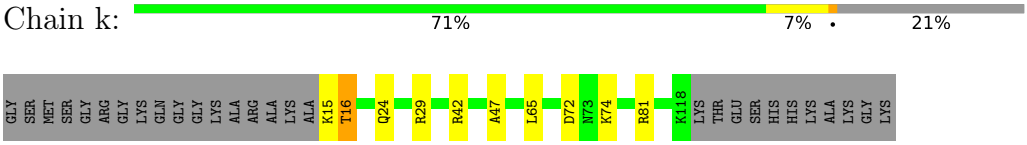
• Molecule 3: Histone H2A type 1-B/E

Chain g:



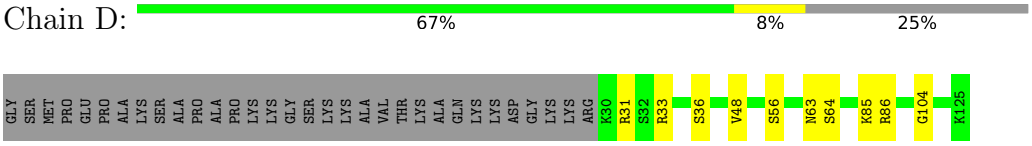
• Molecule 3: Histone H2A type 1-B/E

Chain k:



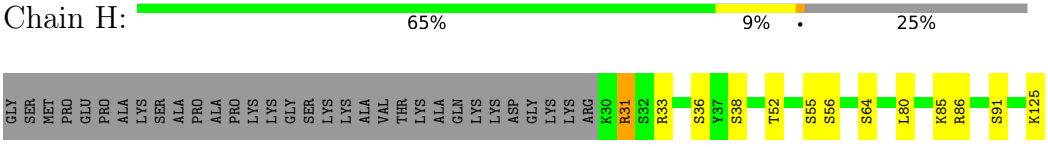
• Molecule 4: Histone H2B type 1-J

Chain D:



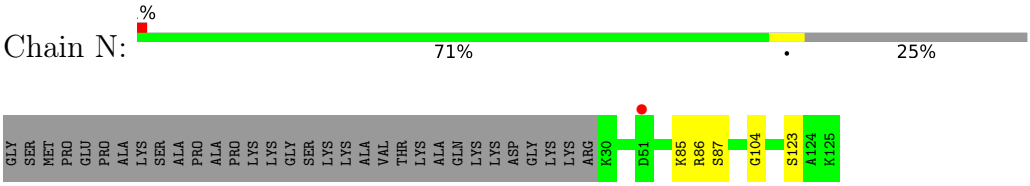
• Molecule 4: Histone H2B type 1-J

Chain H:



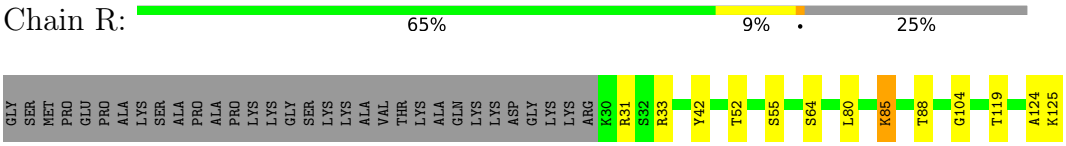
• Molecule 4: Histone H2B type 1-J

Chain N:



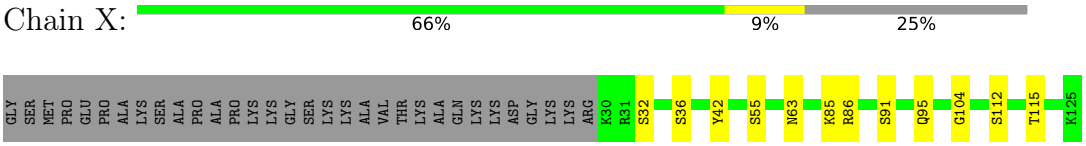
• Molecule 4: Histone H2B type 1-J

Chain R:



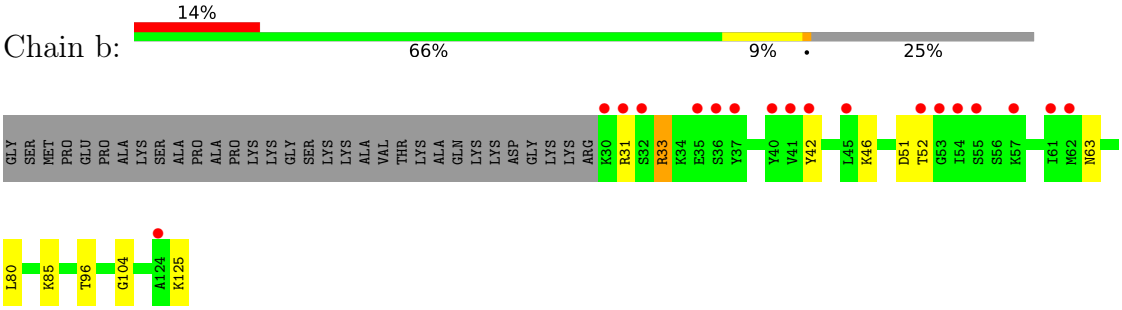
• Molecule 4: Histone H2B type 1-J

Chain X:



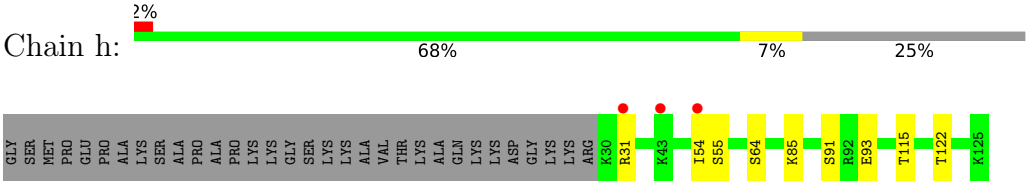
• Molecule 4: Histone H2B type 1-J

Chain b:



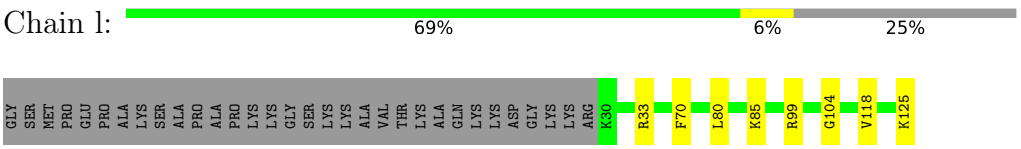
• Molecule 4: Histone H2B type 1-J

Chain h:



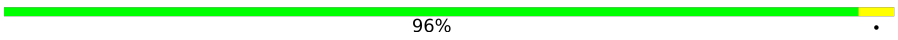
• Molecule 4: Histone H2B type 1-J

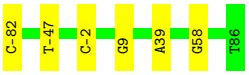
Chain l:



• Molecule 5: DNA (169-MER)

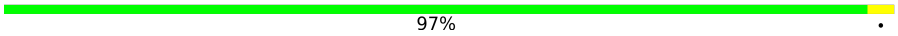
Chain I:





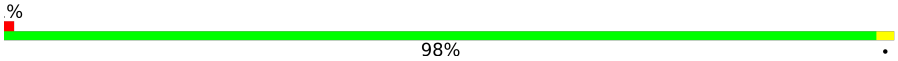
• Molecule 5: DNA (169-MER)

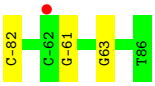
Chain S:



• Molecule 5: DNA (169-MER)

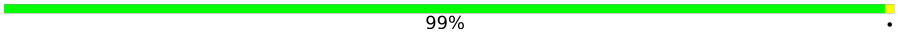
Chain c:





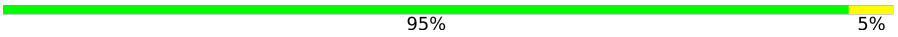
• Molecule 5: DNA (169-MER)

Chain m:



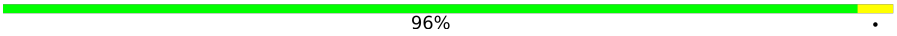
• Molecule 6: DNA (169-MER)

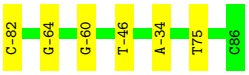
Chain J:



• Molecule 6: DNA (169-MER)

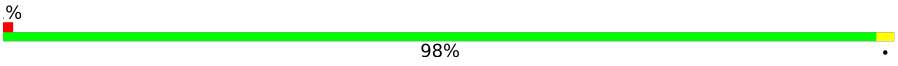
Chain T:

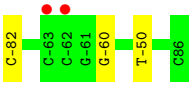




• Molecule 6: DNA (169-MER)

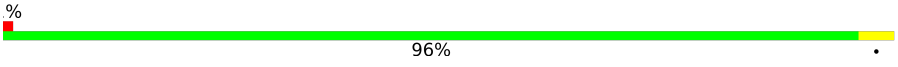
Chain d:

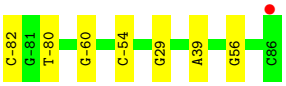




• Molecule 6: DNA (169-MER)

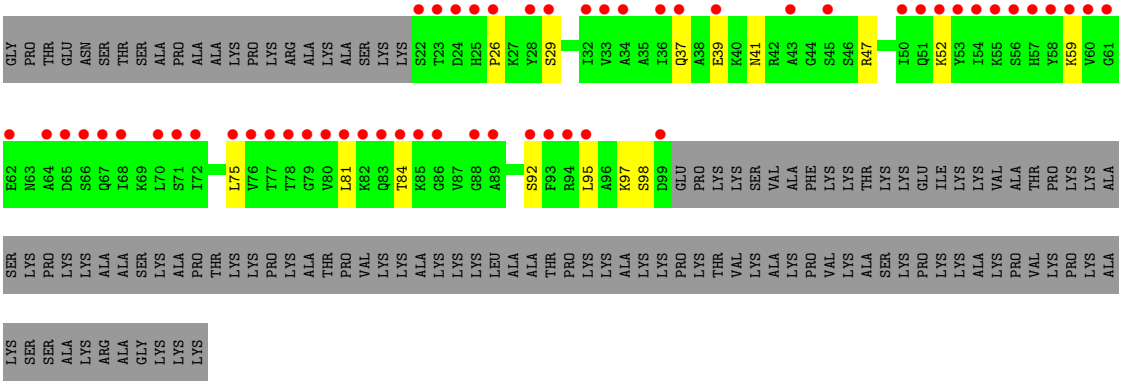
Chain n:





• Molecule 7: Histone H1.0

Chain o:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	213.85Å 102.46Å 218.29Å 90.00° 100.49° 90.00°	Depositor
Resolution (Å)	88.44 – 3.51 88.28 – 3.51	Depositor EDS
% Data completeness (in resolution range)	95.9 (88.44-3.51) 95.9 (88.28-3.51)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.11 (at 3.49Å)	Xtriage
Refinement program	REFMAC 5.8.0258	Depositor
R, R_{free}	0.206 , 0.280 0.209 , 0.276	Depositor DCC
R_{free} test set	2266 reflections (2.02%)	wwPDB-VP
Wilson B-factor (Å^2)	116.3	Xtriage
Anisotropy	0.111	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.29 , 100.2	EDS
L-test for twinning[2]	$< \|L\| > = 0.47$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	0.019 for l,-k,h	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	52574	wwPDB-VP
Average B, all atoms (Å^2)	159.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 8.55% 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)

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	0.68	0/819	0.90	0/1097
1	E	0.68	0/819	0.92	0/1097
1	K	0.65	0/819	0.85	0/1097
1	O	0.66	0/819	0.84	0/1097
1	U	0.67	0/819	0.81	0/1097
1	Y	0.66	0/819	0.82	0/1097
1	e	0.67	0/819	0.88	0/1097
1	i	0.66	0/819	0.84	0/1097
2	B	0.71	0/634	0.98	0/848
2	F	0.69	0/711	0.95	0/948
2	L	0.69	0/634	0.86	0/848
2	P	0.68	0/669	0.91	1/894 (0.1%)
2	V	0.68	0/634	0.85	0/848
2	Z	0.73	0/711	0.94	0/948
2	f	0.70	0/634	0.95	0/848
2	j	0.66	0/645	0.91	0/862
3	C	0.67	0/845	0.92	0/1139
3	G	0.71	0/820	0.90	0/1107
3	M	0.68	0/845	0.85	0/1139
3	Q	0.68	0/820	0.84	0/1107
3	W	0.68	0/845	0.83	0/1139
3	a	0.69	0/815	0.84	0/1100
3	g	0.68	0/820	0.86	0/1107
3	k	0.66	0/815	0.84	0/1100
4	D	0.69	0/766	0.89	0/1026
4	H	0.70	0/766	0.89	0/1026
4	N	0.69	0/766	0.85	0/1026
4	R	0.69	0/766	0.85	0/1026
4	X	0.70	0/766	0.82	0/1026
4	b	0.70	0/766	0.81	0/1026
4	h	0.69	0/766	0.86	0/1026
4	l	0.68	0/766	0.84	0/1026
5	I	0.55	3/3883 (0.1%)	0.87	3/5983 (0.1%)
5	S	0.48	2/3883 (0.1%)	0.85	3/5983 (0.1%)

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#|Z| >5	RMSZ	#|Z| >5
5	c	0.41	1/3883 (0.0%)	0.83	1/5983 (0.0%)
5	m	0.46	2/3883 (0.1%)	0.86	0/5983
6	J	0.55	5/3891 (0.1%)	0.88	4/6008 (0.1%)
6	T	0.52	4/3891 (0.1%)	0.84	2/6008 (0.0%)
6	d	0.40	1/3891 (0.0%)	0.84	1/6008 (0.0%)
6	n	0.48	2/3891 (0.1%)	0.86	6/6008 (0.1%)
7	o	0.74	0/601	0.88	0/802
All	All	0.58	20/56274 (0.0%)	0.86	21/81732 (0.0%)

All (20) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	I	-82	DC	OP3-P	-10.36	1.48	1.61
6	d	-82	DC	OP3-P	-10.30	1.48	1.61
6	J	-82	DC	OP3-P	-10.17	1.49	1.61
6	T	-82	DC	OP3-P	-10.05	1.49	1.61
5	S	-82	DC	OP3-P	-10.03	1.49	1.61
5	c	-82	DC	OP3-P	-9.95	1.49	1.61
5	m	-82	DC	OP3-P	-9.91	1.49	1.61
6	n	-82	DC	OP3-P	-9.82	1.49	1.61
6	T	-46	DT	O3'-P	-6.20	1.53	1.61
5	I	-47	DT	O3'-P	-6.06	1.53	1.61
6	J	-38	DA	O3'-P	-5.89	1.54	1.61
6	J	-47	DC	O3'-P	-5.78	1.54	1.61
6	J	-57	DT	O3'-P	-5.70	1.54	1.61
6	T	-34	DA	O3'-P	-5.61	1.54	1.61
6	n	56	DG	O3'-P	-5.56	1.54	1.61
5	m	45	DC	O3'-P	5.51	1.67	1.61
6	T	-60	DG	O3'-P	-5.50	1.54	1.61
5	I	58	DG	O3'-P	-5.36	1.54	1.61
6	J	-16	DT	O3'-P	-5.31	1.54	1.61
5	S	-7	DG	O3'-P	-5.17	1.54	1.61

All (21) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
6	J	-68	DA	O5'-P-OP1	-8.23	98.29	105.70
6	n	29	DG	O5'-P-OP1	-7.27	99.16	105.70
6	J	-64	DG	C1'-O4'-C4'	-6.66	103.44	110.10
6	n	-60	DG	O4'-C4'-C3'	-6.22	102.01	104.50
6	J	-5	DA	O5'-P-OP2	-5.95	100.34	105.70
6	J	29	DG	O5'-P-OP2	-5.84	100.44	105.70

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	S	67	DA	C1'-O4'-C4'	-5.73	104.37	110.10
5	I	9	DG	O5'-P-OP2	-5.58	100.67	105.70
6	n	39	DA	O5'-P-OP2	5.53	117.34	110.70
5	c	63	DG	C1'-O4'-C4'	-5.47	104.63	110.10
6	n	-54	DC	O5'-P-OP2	-5.44	100.80	105.70
6	T	-64	DG	C1'-O4'-C4'	-5.42	104.68	110.10
6	T	75	DT	C1'-O4'-C4'	-5.41	104.69	110.10
2	P	25	ASN	CB-CA-C	-5.32	99.76	110.40
5	S	-76	DA	C1'-O4'-C4'	-5.25	104.86	110.10
6	n	-60	DG	C1'-O4'-C4'	-5.25	104.86	110.10
6	d	-60	DG	C1'-O4'-C4'	-5.16	104.94	110.10
5	I	-2	DC	O5'-P-OP2	-5.12	101.09	105.70
5	I	39	DA	O5'-P-OP2	5.12	116.84	110.70
5	S	63	DG	C1'-O4'-C4'	-5.07	105.03	110.10
6	n	-80	DT	O5'-P-OP1	-5.03	101.17	105.70

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts

Due to software issues we are unable to calculate clashes - this section is therefore empty.

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	96/138 (70%)	88 (92%)	8 (8%)	0	100	100
1	E	96/138 (70%)	91 (95%)	5 (5%)	0	100	100
1	K	96/138 (70%)	86 (90%)	10 (10%)	0	100	100
1	O	96/138 (70%)	92 (96%)	4 (4%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	U	96/138 (70%)	87 (91%)	7 (7%)	2 (2%)	7 38
1	Y	96/138 (70%)	88 (92%)	8 (8%)	0	100 100
1	e	96/138 (70%)	89 (93%)	6 (6%)	1 (1%)	15 54
1	i	96/138 (70%)	88 (92%)	8 (8%)	0	100 100
2	B	77/105 (73%)	66 (86%)	11 (14%)	0	100 100
2	F	85/105 (81%)	71 (84%)	13 (15%)	1 (1%)	13 50
2	L	77/105 (73%)	66 (86%)	11 (14%)	0	100 100
2	P	81/105 (77%)	72 (89%)	8 (10%)	1 (1%)	13 50
2	V	77/105 (73%)	68 (88%)	7 (9%)	2 (3%)	5 33
2	Z	85/105 (81%)	72 (85%)	10 (12%)	3 (4%)	3 27
2	f	77/105 (73%)	67 (87%)	10 (13%)	0	100 100
2	j	78/105 (74%)	69 (88%)	9 (12%)	0	100 100
3	C	106/132 (80%)	94 (89%)	12 (11%)	0	100 100
3	G	103/132 (78%)	92 (89%)	11 (11%)	0	100 100
3	M	106/132 (80%)	94 (89%)	11 (10%)	1 (1%)	17 56
3	Q	103/132 (78%)	95 (92%)	7 (7%)	1 (1%)	15 54
3	W	106/132 (80%)	92 (87%)	13 (12%)	1 (1%)	17 56
3	a	102/132 (77%)	88 (86%)	12 (12%)	2 (2%)	7 39
3	g	103/132 (78%)	87 (84%)	16 (16%)	0	100 100
3	k	102/132 (77%)	89 (87%)	11 (11%)	2 (2%)	7 39
4	D	94/128 (73%)	84 (89%)	8 (8%)	2 (2%)	7 38
4	H	94/128 (73%)	86 (92%)	6 (6%)	2 (2%)	7 38
4	N	94/128 (73%)	86 (92%)	7 (7%)	1 (1%)	14 52
4	R	94/128 (73%)	78 (83%)	12 (13%)	4 (4%)	2 22
4	X	94/128 (73%)	79 (84%)	13 (14%)	2 (2%)	7 38
4	b	94/128 (73%)	81 (86%)	11 (12%)	2 (2%)	7 38
4	h	94/128 (73%)	83 (88%)	10 (11%)	1 (1%)	14 52
4	l	94/128 (73%)	84 (89%)	8 (8%)	2 (2%)	7 38
7	o	76/195 (39%)	64 (84%)	10 (13%)	2 (3%)	5 33
All	All	3064/4219 (73%)	2716 (89%)	313 (10%)	35 (1%)	14 52

All (35) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
4	N	104	GLY
2	P	21	VAL
4	R	33	ARG
2	V	28	GLY
2	Z	22	LEU
3	a	16	THR
4	b	33	ARG
4	D	31	ARG
4	D	104	GLY
4	H	33	ARG
4	R	104	GLY
4	R	124	ALA
1	U	43	PRO
2	V	96	THR
2	Z	20	LYS
4	b	104	GLY
4	l	104	GLY
7	o	26	PRO
4	H	31	ARG
2	Z	23	ARG
3	a	45	ALA
3	k	16	THR
4	l	70	PHE
4	R	85	LYS
2	F	23	ARG
7	o	98	SER
1	U	51	ILE
4	X	112	SER
1	e	51	ILE
3	M	109	PRO
4	X	104	GLY
3	Q	47	ALA
3	W	109	PRO
4	h	54	ILE
3	k	47	ALA

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	85/112 (76%)	82 (96%)	3 (4%)	36	67
1	E	85/112 (76%)	80 (94%)	5 (6%)	19	53
1	K	85/112 (76%)	82 (96%)	3 (4%)	36	67
1	O	85/112 (76%)	80 (94%)	5 (6%)	19	53
1	U	85/112 (76%)	83 (98%)	2 (2%)	49	76
1	Y	85/112 (76%)	84 (99%)	1 (1%)	71	87
1	e	85/112 (76%)	79 (93%)	6 (7%)	14	46
1	i	85/112 (76%)	80 (94%)	5 (6%)	19	53
2	B	64/80 (80%)	63 (98%)	1 (2%)	62	83
2	F	72/80 (90%)	71 (99%)	1 (1%)	67	85
2	L	64/80 (80%)	60 (94%)	4 (6%)	18	51
2	P	68/80 (85%)	63 (93%)	5 (7%)	13	44
2	V	64/80 (80%)	61 (95%)	3 (5%)	26	60
2	Z	72/80 (90%)	66 (92%)	6 (8%)	11	40
2	f	64/80 (80%)	60 (94%)	4 (6%)	18	51
2	j	65/80 (81%)	62 (95%)	3 (5%)	27	61
3	C	85/101 (84%)	83 (98%)	2 (2%)	49	76
3	G	83/101 (82%)	78 (94%)	5 (6%)	19	52
3	M	85/101 (84%)	80 (94%)	5 (6%)	19	53
3	Q	83/101 (82%)	76 (92%)	7 (8%)	11	40
3	W	85/101 (84%)	80 (94%)	5 (6%)	19	53
3	a	83/101 (82%)	79 (95%)	4 (5%)	25	60
3	g	83/101 (82%)	79 (95%)	4 (5%)	25	60
3	k	83/101 (82%)	74 (89%)	9 (11%)	6	30
4	D	82/106 (77%)	74 (90%)	8 (10%)	8	33
4	H	82/106 (77%)	70 (85%)	12 (15%)	3	18
4	N	82/106 (77%)	78 (95%)	4 (5%)	25	59
4	R	82/106 (77%)	72 (88%)	10 (12%)	5	23
4	X	82/106 (77%)	72 (88%)	10 (12%)	5	23
4	b	82/106 (77%)	72 (88%)	10 (12%)	5	23
4	h	82/106 (77%)	74 (90%)	8 (10%)	8	33
4	l	82/106 (77%)	76 (93%)	6 (7%)	14	45

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
7	o	65/158 (41%)	52 (80%)	13 (20%)	1 7
All	All	2604/3350 (78%)	2425 (93%)	179 (7%)	15 47

All (179) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	58	THR
1	A	63	ARG
1	A	129	ARG
2	B	95	ARG
3	C	74	LYS
3	C	101	THR
4	D	33	ARG
4	D	36	SER
4	D	48	VAL
4	D	56	SER
4	D	63	ASN
4	D	64	SER
4	D	85	LYS
4	D	86	ARG
1	E	45	THR
1	E	65	LEU
1	E	68	GLN
1	E	129	ARG
1	E	134	ARG
2	F	16	LYS
3	G	15	LYS
3	G	29	ARG
3	G	74	LYS
3	G	79	ILE
3	G	81	ARG
4	H	31	ARG
4	H	36	SER
4	H	38	SER
4	H	52	THR
4	H	55	SER
4	H	56	SER
4	H	64	SER
4	H	80	LEU
4	H	85	LYS
4	H	86	ARG
4	H	91	SER

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
4	H	125	LYS
1	K	63	ARG
1	K	86	SER
1	K	129	ARG
2	L	24	ASP
2	L	50	ILE
2	L	92	ARG
2	L	95	ARG
3	M	36	LYS
3	M	42	ARG
3	M	74	LYS
3	M	81	ARG
3	M	118	LYS
4	N	85	LYS
4	N	86	ARG
4	N	87	SER
4	N	123	SER
1	O	48	LEU
1	O	59	GLU
1	O	77	ASP
1	O	87	SER
1	O	129	ARG
2	P	21	VAL
2	P	23	ARG
2	P	24	ASP
2	P	25	ASN
2	P	27	GLN
3	Q	15	LYS
3	Q	19	SER
3	Q	24	GLN
3	Q	29	ARG
3	Q	81	ARG
3	Q	101	THR
3	Q	107	VAL
4	R	31	ARG
4	R	42	TYR
4	R	52	THR
4	R	55	SER
4	R	64	SER
4	R	80	LEU
4	R	85	LYS
4	R	88	THR

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
4	R	119	THR
4	R	125	LYS
1	U	63	ARG
1	U	134	ARG
2	V	26	ILE
2	V	92	ARG
2	V	95	ARG
3	W	24	GLN
3	W	74	LYS
3	W	81	ARG
3	W	100	VAL
3	W	101	THR
4	X	32	SER
4	X	36	SER
4	X	42	TYR
4	X	55	SER
4	X	63	ASN
4	X	85	LYS
4	X	86	ARG
4	X	91	SER
4	X	95	GLN
4	X	115	THR
1	Y	58	THR
2	Z	16	LYS
2	Z	17	ARG
2	Z	18	HIS
2	Z	19	ARG
2	Z	20	LYS
2	Z	92	ARG
3	a	15	LYS
3	a	16	THR
3	a	74	LYS
3	a	81	ARG
4	b	31	ARG
4	b	33	ARG
4	b	42	TYR
4	b	46	LYS
4	b	52	THR
4	b	63	ASN
4	b	80	LEU
4	b	85	LYS
4	b	96	THR

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
4	b	125	LYS
1	e	56	LYS
1	e	58	THR
1	e	63	ARG
1	e	65	LEU
1	e	86	SER
1	e	129	ARG
2	f	24	ASP
2	f	50	ILE
2	f	95	ARG
2	f	96	THR
3	g	24	GLN
3	g	73	ASN
3	g	74	LYS
3	g	114	VAL
4	h	31	ARG
4	h	55	SER
4	h	64	SER
4	h	85	LYS
4	h	91	SER
4	h	93	GLU
4	h	115	THR
4	h	122	THR
1	i	48	LEU
1	i	83	ARG
1	i	87	SER
1	i	129	ARG
1	i	134	ARG
2	j	23	ARG
2	j	26	ILE
2	j	27	GLN
3	k	15	LYS
3	k	16	THR
3	k	24	GLN
3	k	29	ARG
3	k	42	ARG
3	k	65	LEU
3	k	72	ASP
3	k	74	LYS
3	k	81	ARG
4	l	33	ARG
4	l	80	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
4	l	85	LYS
4	l	99	ARG
4	l	118	VAL
4	l	125	LYS
7	o	29	SER
7	o	37	GLN
7	o	39	GLU
7	o	41	ASN
7	o	47	ARG
7	o	52	LYS
7	o	59	LYS
7	o	75	LEU
7	o	81	LEU
7	o	84	THR
7	o	92	SER
7	o	95	LEU
7	o	97	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (39) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	76	GLN
1	A	93	GLN
3	C	31	HIS
3	C	73	ASN
4	D	95	GLN
2	F	18	HIS
2	F	27	GLN
2	F	64	ASN
3	G	31	HIS
4	H	47	GLN
4	H	95	GLN
1	K	68	GLN
1	K	108	ASN
4	N	84	ASN
4	N	95	GLN
2	P	27	GLN
3	Q	24	GLN
3	Q	110	ASN
4	R	95	GLN
4	X	67	ASN
4	X	95	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	Y	76	GLN
1	Y	108	ASN
2	Z	27	GLN
1	e	108	ASN
2	f	25	ASN
3	g	24	GLN
4	h	95	GLN
4	h	109	HIS
1	i	76	GLN
1	i	108	ASN
2	j	27	GLN
2	j	64	ASN
3	k	24	GLN
3	k	73	ASN
4	l	95	GLN
7	o	41	ASN
7	o	67	GLN
7	o	83	GLN

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

There are no ligands in this entry.

5.7 Other polymers

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	98/138 (71%)	-0.17	0 100 100	63, 88, 132, 174	0
1	E	98/138 (71%)	-0.26	1 (1%) 82 77	59, 84, 123, 147	0
1	K	98/138 (71%)	-0.38	0 100 100	90, 115, 155, 190	0
1	O	98/138 (71%)	-0.07	0 100 100	84, 116, 149, 189	0
1	U	98/138 (71%)	0.05	1 (1%) 82 77	101, 149, 191, 202	0
1	Y	98/138 (71%)	-0.25	0 100 100	118, 162, 210, 227	0
1	e	98/138 (71%)	-0.41	0 100 100	78, 108, 152, 185	0
1	i	98/138 (71%)	-0.04	1 (1%) 82 77	84, 109, 155, 185	0
2	B	79/105 (75%)	-0.22	0 100 100	60, 82, 114, 156	0
2	F	87/105 (82%)	-0.20	0 100 100	64, 82, 187, 217	0
2	L	79/105 (75%)	-0.24	0 100 100	84, 111, 137, 147	0
2	P	83/105 (79%)	-0.09	0 100 100	88, 115, 156, 224	0
2	V	79/105 (75%)	0.14	4 (5%) 28 25	104, 141, 172, 185	0
2	Z	87/105 (82%)	-0.18	1 (1%) 80 75	106, 159, 221, 253	0
2	f	79/105 (75%)	-0.36	0 100 100	84, 106, 138, 145	0
2	j	80/105 (76%)	-0.01	1 (1%) 77 71	80, 105, 150, 179	0
3	C	108/132 (81%)	-0.29	1 (0%) 84 79	63, 89, 127, 190	0
3	G	105/132 (79%)	-0.16	1 (0%) 82 77	59, 85, 136, 184	0
3	M	108/132 (81%)	0.00	3 (2%) 53 47	90, 124, 178, 226	0
3	Q	105/132 (79%)	-0.27	1 (0%) 82 77	91, 116, 159, 195	0
3	W	108/132 (81%)	-0.20	0 100 100	108, 152, 197, 250	0
3	a	104/132 (78%)	0.73	15 (14%) 2 3	119, 156, 208, 245	0
3	g	105/132 (79%)	0.23	0 100 100	86, 109, 159, 186	0
3	k	104/132 (78%)	-0.32	0 100 100	84, 112, 150, 185	0

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å^2)	Q<0.9
4	D	96/128 (75%)	-0.29	0 100 100	64, 90, 145, 216	0
4	H	96/128 (75%)	-0.26	0 100 100	54, 88, 160, 183	0
4	N	96/128 (75%)	-0.15	1 (1%) 82 77	92, 124, 179, 228	0
4	R	96/128 (75%)	-0.21	0 100 100	100, 127, 177, 200	0
4	X	96/128 (75%)	-0.22	0 100 100	121, 157, 206, 212	0
4	b	96/128 (75%)	0.50	18 (18%) 1 1	121, 166, 207, 222	0
4	h	96/128 (75%)	0.18	3 (3%) 49 43	77, 115, 174, 212	0
4	l	96/128 (75%)	-0.26	0 100 100	80, 113, 168, 231	0
5	I	169/169 (100%)	-0.67	0 100 100	89, 145, 240, 312	0
5	S	169/169 (100%)	-0.62	0 100 100	121, 186, 251, 316	0
5	c	169/169 (100%)	-0.45	1 (0%) 89 86	158, 221, 301, 363	0
5	m	169/169 (100%)	-0.60	0 100 100	118, 178, 244, 290	0
6	J	169/169 (100%)	-0.73	0 100 100	99, 142, 254, 339	0
6	T	169/169 (100%)	-0.67	0 100 100	127, 186, 258, 295	0
6	d	169/169 (100%)	-0.33	2 (1%) 79 73	162, 228, 285, 342	0
6	n	169/169 (100%)	-0.41	1 (0%) 89 86	130, 183, 265, 338	0
7	o	78/195 (40%)	3.39	55 (70%) 0 0	152, 196, 239, 272	1 (1%)
All	All	4482/5571 (80%)	-0.19	111 (2%) 57 51	54, 135, 236, 363	1 (0%)

All (111) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
7	o	22	SER	10.3
7	o	60	VAL	8.8
7	o	25	HIS	8.3
7	o	83	GLN	7.7
7	o	80	VAL	7.3
7	o	79	GLY	6.8
7	o	94	ARG	6.6
7	o	93	PHE	6.6
4	b	52	THR	6.5
7	o	23	THR	6.4
7	o	24	ASP	6.3
7	o	65	ASP	5.8
3	a	26	PRO	5.6
7	o	61	GLY	5.5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
7	o	84	THR	5.5
2	V	102	GLY	5.3
7	o	82	LYS	5.3
7	o	26	PRO	5.3
7	o	81	LEU	5.2
3	a	22	GLY	5.1
7	o	57	HIS	5.0
7	o	28	TYR	5.0
3	a	76	THR	4.6
7	o	54	ILE	4.5
7	o	33	VAL	4.5
7	o	55	LYS	4.4
7	o	43	ALA	4.3
4	b	41	VAL	4.3
7	o	45	SER	4.3
7	o	76	VAL	4.3
7	o	78	THR	4.3
2	V	101	GLY	4.1
7	o	99	ASP	4.0
7	o	77	THR	4.0
7	o	66	SER	3.9
4	b	40	TYR	3.9
4	b	31	ARG	3.8
7	o	64	ALA	3.6
3	a	78	ILE	3.5
7	o	62	GLU	3.5
4	b	53	GLY	3.5
4	b	54	ILE	3.4
3	M	13	LYS	3.4
7	o	32	ILE	3.3
7	o	37	GLN	3.3
2	V	100	PHE	3.3
7	o	68	ILE	3.3
7	o	53	TYR	3.3
7	o	89	ALA	3.2
7	o	92	SER	3.2
3	Q	14	ALA	3.2
7	o	59	LYS	3.2
7	o	29	SER	3.1
6	d	-63	DC	3.1
4	b	62	MET	3.0
4	b	35	GLU	3.0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
7	o	67	GLN	3.0
7	o	34	ALA	3.0
7	o	50	ILE	2.9
4	b	32	SER	2.9
7	o	70	LEU	2.9
3	a	20	ARG	2.9
7	o	72	ILE	2.8
6	d	-62	DC	2.8
7	o	95	LEU	2.8
2	V	99	GLY	2.8
3	G	14	ALA	2.7
3	a	82	HIS	2.7
3	a	21	ALA	2.7
4	b	42	TYR	2.7
7	o	51	GLN	2.6
7	o	75	LEU	2.6
4	b	61	ILE	2.6
3	a	118	LYS	2.6
4	b	30	LYS	2.6
3	a	25	PHE	2.5
4	h	31	ARG	2.5
3	a	18	SER	2.5
4	b	45	LEU	2.5
2	Z	44	LYS	2.5
7	o	85	LYS	2.4
4	b	55	SER	2.4
1	U	84	PHE	2.4
7	o	52	LYS	2.4
4	b	36	SER	2.4
3	a	34	LEU	2.3
4	b	57	LYS	2.3
7	o	86	GLY	2.3
3	a	30	VAL	2.3
1	i	39	HIS	2.2
3	C	12	ALA	2.2
3	M	104	GLN	2.2
5	c	-62	DC	2.2
7	o	71	SER	2.2
3	a	23	LEU	2.2
7	o	88	GLY	2.2
2	j	102	GLY	2.1
3	M	115	LEU	2.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
4	b	37	TYR	2.1
7	o	58	TYR	2.1
7	o	39	GLU	2.1
1	E	39	HIS	2.1
3	a	104	GLN	2.1
4	h	54	ILE	2.1
7	o	56	SER	2.1
4	N	51	ASP	2.1
3	a	77	ARG	2.1
4	h	43	LYS	2.0
4	b	124	ALA	2.0
7	o	36	ILE	2.0
6	n	86	DC	2.0

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)

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.