



# wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 16, 2023 – 07:20 pm BST

PDB ID : 6YN1  
Title : Crystal structure of histone chaperone APLF acidic domain bound to the histone H2A-H2B-H3-H4 octamer  
Authors : Corbeski, I.; Guo, X.; Van Ingen, H.; Sixma, T.K.  
Deposited on : 2020-04-10  
Resolution : 2.35 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](mailto: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  
Mogul : 1.8.4, 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

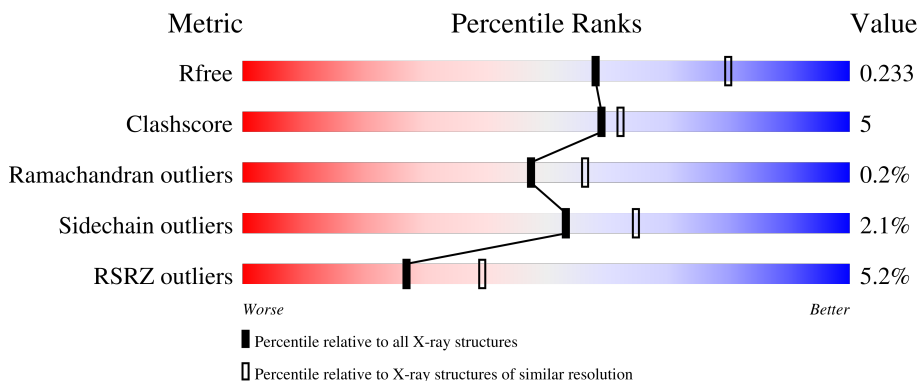
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.35 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	1164 (2.36-2.36)
Clashscore	141614	1232 (2.36-2.36)
Ramachandran outliers	138981	1211 (2.36-2.36)
Sidechain outliers	138945	1212 (2.36-2.36)
RSRZ outliers	127900	1150 (2.36-2.36)

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  $\geq 3$ , 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  $\leq 5\%$ . The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	107	 3% 91% 6% .
1	F	107	 7% 84% 12% .
1	K	107	 4% 90% 7% .
1	P	107	 6% 85% 12% .
1	U	107	 5% 79% 16% . .

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	Z	107	6% 87% 8% . .
1	e	107	9% 94% . 5%
1	j	107	4% 96% .
2	B	100	4% 77% 13% 10%
2	G	100	2% 71% 17% 12%
2	L	100	% 81% 7% 12%
2	Q	100	% 79% 9% 12%
2	V	100	% 75% 14% 11%
2	a	100	% 86% . 11%
2	f	100	2% 85% . 12%
2	k	100	3% 86% . 12%
3	C	99	3% 86% 9% 5%
3	H	99	3% 88% 6% . 5%
3	M	99	4% 89% 6% . .
3	R	99	2% 87% 7% . 5%
3	W	99	5% 88% 7% 5%
3	b	99	6% 94% . .
3	g	99	3% 94% . 5%
3	l	99	4% 94% . 5%
4	D	84	7% 80% 11% 10%
4	I	84	4% 80% 11% 10%
4	N	84	5% 83% 6% 11%
4	S	84	5% 74% 15% . 8%
4	X	84	5% 77% 12% . 8%
4	c	84	% 88% . 10%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
4	h	84	<p>5% 89% 11%</p>
4	m	84	<p>6% 89% 11%</p>
5	E	43	<p>12% 51% 49%</p>
5	J	43	<p>16% 53% 12% 33%</p>
5	O	43	<p>9% 44% 5% 51%</p>
5	T	43	<p>12% 47% 5% 49%</p>
5	Y	43	<p>14% 42% 9% 49%</p>
5	d	43	<p>12% 44% 53%</p>
5	i	43	<p>7% 42% 7% 51%</p>
5	n	43	<p>12% 47% 53%</p>

## 2 Entry composition

There are 8 unique types of molecules in this entry. The entry contains 24986 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 H2A.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
1	A	103	795	501	155	139	0	0	0
1	F	103	795	501	155	139	0	0	0
1	K	103	795	501	155	139	0	0	0
1	P	104	804	507	157	140	0	0	0
1	U	103	795	501	155	139	0	0	0
1	Z	103	795	501	155	139	0	0	0
1	e	102	786	495	153	138	0	0	0
1	j	103	795	501	155	139	0	0	0

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	12	MET	-	initiating methionine	UNP Q6AZJ8
F	12	MET	-	initiating methionine	UNP Q6AZJ8
K	12	MET	-	initiating methionine	UNP Q6AZJ8
P	12	MET	-	initiating methionine	UNP Q6AZJ8
U	12	MET	-	initiating methionine	UNP Q6AZJ8
Z	12	MET	-	initiating methionine	UNP Q6AZJ8
e	12	MET	-	initiating methionine	UNP Q6AZJ8
j	12	MET	-	initiating methionine	UNP Q6AZJ8

- Molecule 2 is a protein called Histone H2B.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	90	708	449	123	134	2	0	1	0
2	G	88	697	442	122	131	2	0	1	0
2	L	88	685	433	121	129	2	0	0	0
2	Q	88	685	433	121	129	2	0	0	0
2	V	89	694	438	122	132	2	0	0	0
2	a	89	703	446	122	133	2	0	1	0
2	f	88	685	433	121	129	2	0	0	0
2	k	88	685	433	121	129	2	0	0	0

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	26	MET	-	initiating methionine	UNP A0A1L8FQA5
G	26	MET	-	initiating methionine	UNP A0A1L8FQA5
L	26	MET	-	initiating methionine	UNP A0A1L8FQA5
Q	26	MET	-	initiating methionine	UNP A0A1L8FQA5
V	26	MET	-	initiating methionine	UNP A0A1L8FQA5
a	26	MET	-	initiating methionine	UNP A0A1L8FQA5
f	26	MET	-	initiating methionine	UNP A0A1L8FQA5
k	26	MET	-	initiating methionine	UNP A0A1L8FQA5

- Molecule 3 is a protein called Histone H3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	C	94	774	489	147	135	3	0	0	0
3	H	94	774	489	147	135	3	0	0	0
3	M	95	784	495	150	136	3	0	0	0
3	R	94	774	489	147	135	3	0	0	0
3	W	94	774	489	147	135	3	0	0	0
3	b	96	790	498	152	137	3	0	0	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	g	94	Total	C	N	O	S	0	0	0
			774	489	147	135	3			
3	l	94	Total	C	N	O	S	0	0	0
			774	489	147	135	3			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	37	MET	-	initiating methionine	UNP A0A310TTQ1
H	37	MET	-	initiating methionine	UNP A0A310TTQ1
M	37	MET	-	initiating methionine	UNP A0A310TTQ1
R	37	MET	-	initiating methionine	UNP A0A310TTQ1
W	37	MET	-	initiating methionine	UNP A0A310TTQ1
b	37	MET	-	initiating methionine	UNP A0A310TTQ1
g	37	MET	-	initiating methionine	UNP A0A310TTQ1
l	37	MET	-	initiating methionine	UNP A0A310TTQ1

- Molecule 4 is a protein called Histone H4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	76	Total	C	N	O	S	0	0	0
			606	385	117	103	1			
4	I	76	Total	C	N	O	S	0	0	0
			606	385	117	103	1			
4	N	75	Total	C	N	O	S	0	0	0
			602	383	116	102	1			
4	S	77	Total	C	N	O	S	0	0	0
			614	389	119	105	1			
4	X	77	Total	C	N	O	S	0	0	0
			618	391	119	107	1			
4	c	76	Total	C	N	O	S	0	0	0
			610	387	118	104	1			
4	h	75	Total	C	N	O	S	0	0	0
			602	383	116	102	1			
4	m	75	Total	C	N	O	S	0	0	0
			598	379	116	102	1			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	19	MET	-	initiating methionine	UNP P62799
I	19	MET	-	initiating methionine	UNP P62799

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
N	19	MET	-	initiating methionine	UNP P62799
S	19	MET	-	initiating methionine	UNP P62799
X	19	MET	-	initiating methionine	UNP P62799
c	19	MET	-	initiating methionine	UNP P62799
h	19	MET	-	initiating methionine	UNP P62799
m	19	MET	-	initiating methionine	UNP P62799

- Molecule 5 is a protein called Aprataxin and PNK-like factor.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
5	E	22	Total	C	N	O	0	0	0
			188	114	25	49			
5	J	29	Total	C	N	O	0	0	0
			248	150	32	66			
5	O	21	Total	C	N	O	0	0	0
			179	109	24	46			
5	T	22	Total	C	N	O	0	0	0
			188	114	25	49			
5	Y	22	Total	C	N	O	0	0	0
			188	114	25	49			
5	d	20	Total	C	N	O	0	0	0
			170	104	23	43			
5	i	21	Total	C	N	O	0	0	0
			179	109	24	46			
5	n	20	Total	C	N	O	0	0	0
			170	104	24	42			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
E	448	GLY	-	expression tag	UNP Q8IW19
J	448	GLY	-	expression tag	UNP Q8IW19
O	448	GLY	-	expression tag	UNP Q8IW19
T	448	GLY	-	expression tag	UNP Q8IW19
Y	448	GLY	-	expression tag	UNP Q8IW19
d	448	GLY	-	expression tag	UNP Q8IW19
i	448	GLY	-	expression tag	UNP Q8IW19
n	448	GLY	-	expression tag	UNP Q8IW19

- Molecule 6 is GLYCEROL (three-letter code: GOL) (formula: C<sub>3</sub>H<sub>8</sub>O<sub>3</sub>).





Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	1	Total C O 6 3 3	0	0
6	D	1	Total C O 6 3 3	0	0
6	H	1	Total C O 6 3 3	0	0
6	K	1	Total C O 6 3 3	0	0
6	N	1	Total C O 6 3 3	0	0
6	P	1	Total C O 6 3 3	0	0
6	S	1	Total C O 6 3 3	0	0
6	U	1	Total C O 6 3 3	0	0
6	b	1	Total C O 6 3 3	0	0
6	c	1	Total C O 6 3 3	0	0
6	e	1	Total C O 6 3 3	0	0
6	h	1	Total C O 6 3 3	0	0
6	j	1	Total C O 6 3 3	0	0
6	j	1	Total C O 6 3 3	0	0

- Molecule 7 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	J	1	Total Cl 1 1	0	0

- Molecule 8 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	A	16	Total O 16 16	0	0
8	B	10	Total O 10 10	0	0
8	C	14	Total O 14 14	0	0
8	D	13	Total O 13 13	0	0
8	E	1	Total O 1 1	0	0
8	F	17	Total O 17 17	0	0
8	G	7	Total O 7 7	0	0
8	H	19	Total O 19 19	0	0
8	I	13	Total O 13 13	0	0
8	J	4	Total O 4 4	0	0
8	K	10	Total O 10 10	0	0
8	L	10	Total O 10 10	0	0
8	M	23	Total O 23 23	0	0
8	N	6	Total O 6 6	0	0
8	O	1	Total O 1 1	0	0
8	P	10	Total O 10 10	0	0
8	Q	5	Total O 5 5	0	0
8	R	19	Total O 19 19	0	0

*Continued on next page...*

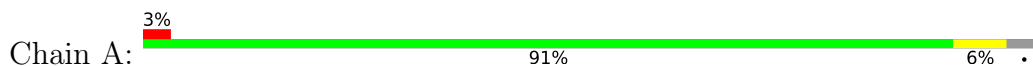
*Continued from previous page...*

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	S	7	Total O 7 7	0	0
8	U	16	Total O 16 16	0	0
8	V	6	Total O 6 6	0	0
8	W	19	Total O 19 19	0	0
8	X	13	Total O 13 13	0	0
8	Y	2	Total O 2 2	0	0
8	Z	9	Total O 9 9	0	0
8	a	6	Total O 6 6	0	0
8	b	25	Total O 25 25	0	0
8	c	7	Total O 7 7	0	0
8	d	2	Total O 2 2	0	0
8	e	8	Total O 8 8	0	0
8	f	3	Total O 3 3	0	0
8	g	25	Total O 25 25	0	0
8	h	9	Total O 9 9	0	0
8	i	1	Total O 1 1	0	0
8	j	15	Total O 15 15	0	0
8	k	8	Total O 8 8	0	0
8	l	18	Total O 18 18	0	0
8	m	18	Total O 18 18	0	0

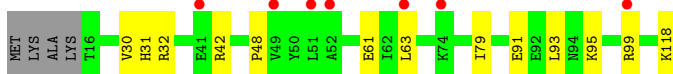
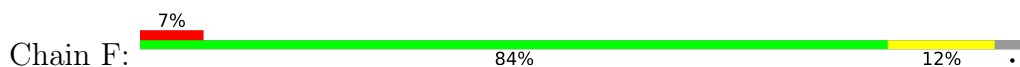
### 3 Residue-property plots [i](#)

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.

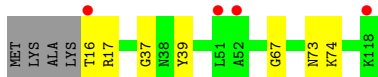
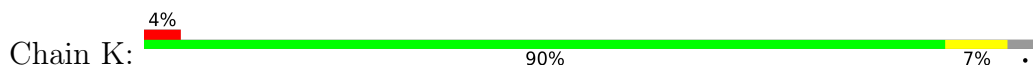
- Molecule 1: Histone H2A



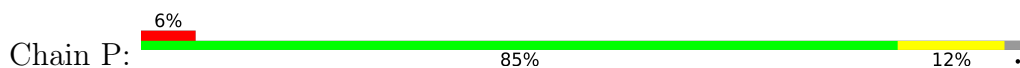
- Molecule 1: Histone H2A



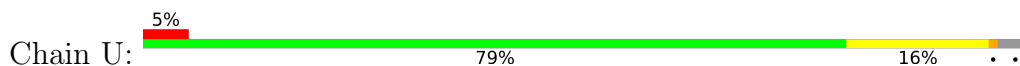
- Molecule 1: Histone H2A



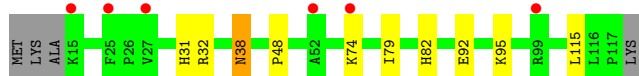
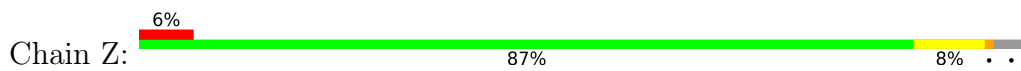
- Molecule 1: Histone H2A



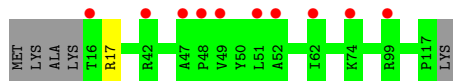
- Molecule 1: Histone H2A



- Molecule 1: Histone H2A



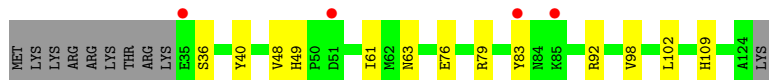
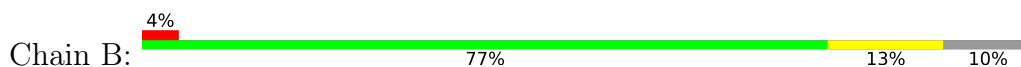
• Molecule 1: Histone H2A



• Molecule 1: Histone H2A



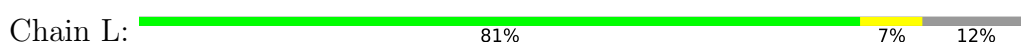
• Molecule 2: Histone H2B



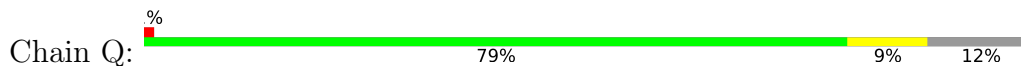
• Molecule 2: Histone H2B



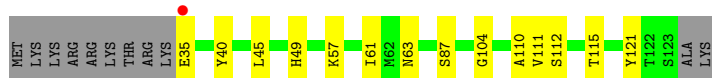
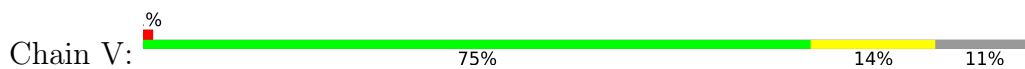
• Molecule 2: Histone H2B



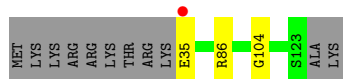
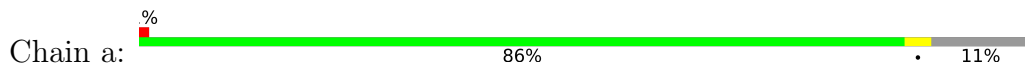
• Molecule 2: Histone H2B



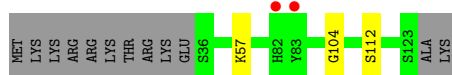
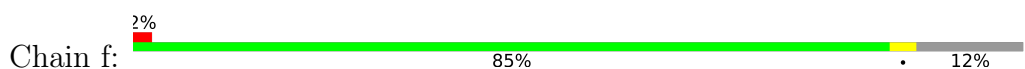
• Molecule 2: Histone H2B



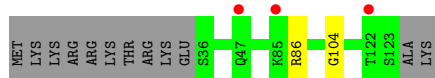
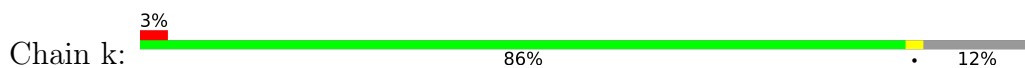
• Molecule 2: Histone H2B



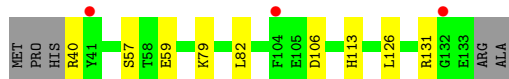
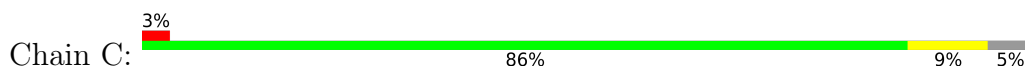
• Molecule 2: Histone H2B



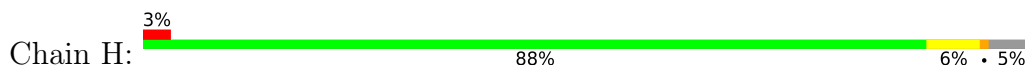
• Molecule 2: Histone H2B



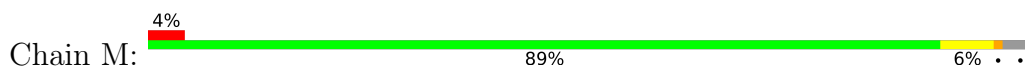
• Molecule 3: Histone H3



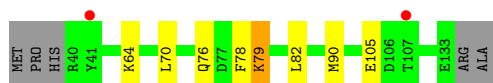
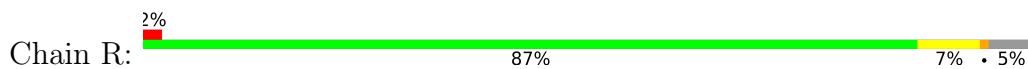
• Molecule 3: Histone H3



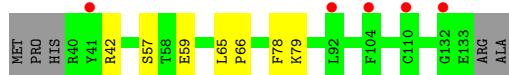
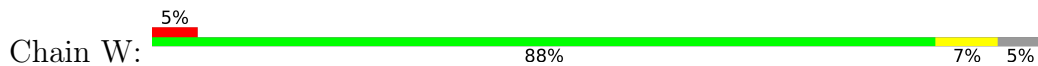
• Molecule 3: Histone H3



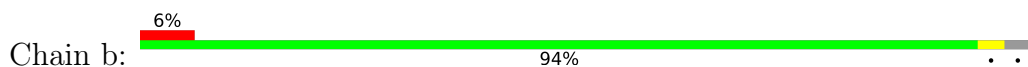
• Molecule 3: Histone H3



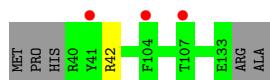
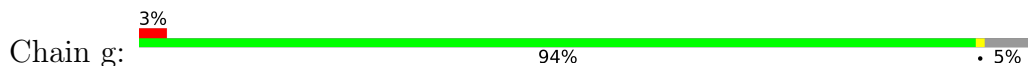
● Molecule 3: Histone H3



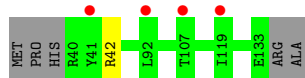
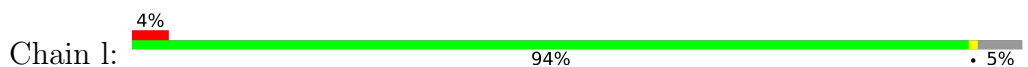
● Molecule 3: Histone H3



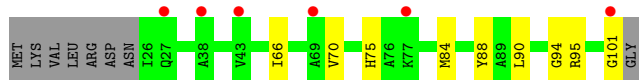
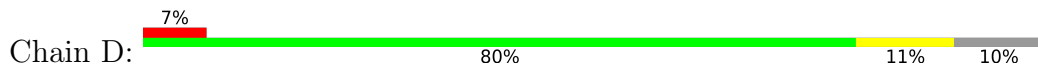
● Molecule 3: Histone H3



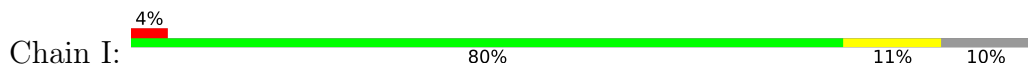
● Molecule 3: Histone H3



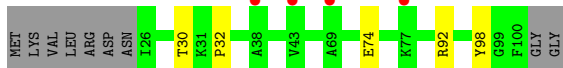
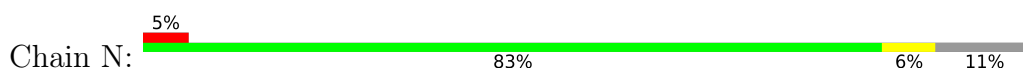
● Molecule 4: Histone H4



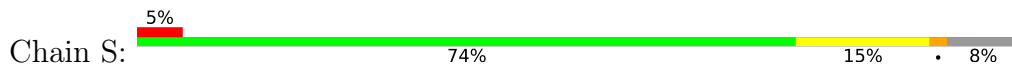
● Molecule 4: Histone H4



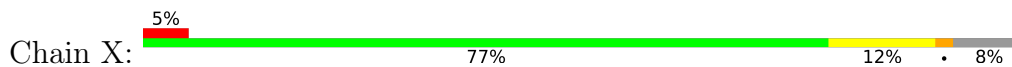
● Molecule 4: Histone H4



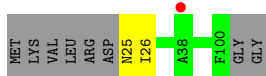
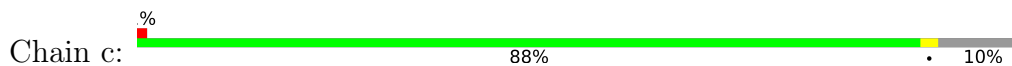
- Molecule 4: Histone H4



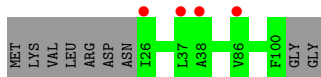
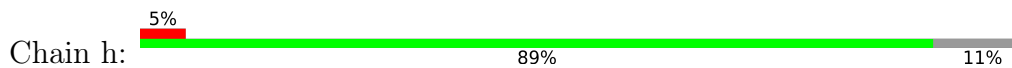
- Molecule 4: Histone H4



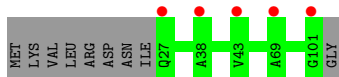
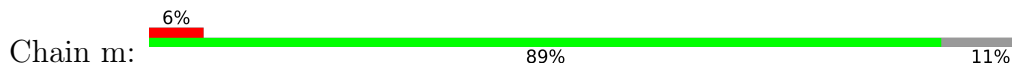
- Molecule 4: Histone H4



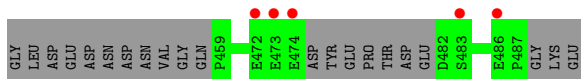
- Molecule 4: Histone H4



- Molecule 4: Histone H4



- Molecule 5: Aprataxin and PNK-like factor

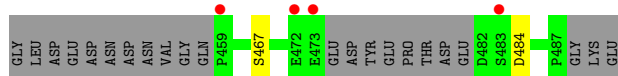
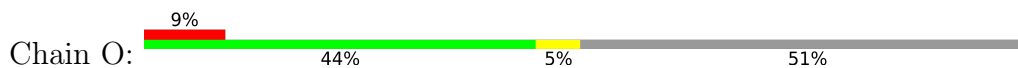


- Molecule 5: Aprataxin and PNK-like factor

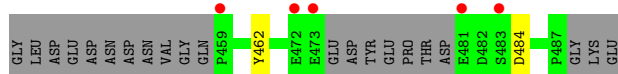




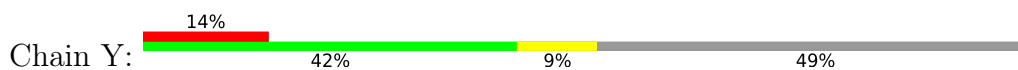
- Molecule 5: Aprataxin and PNK-like factor



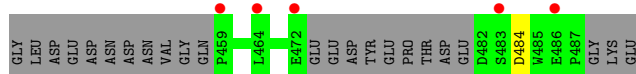
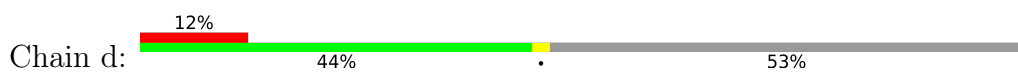
- Molecule 5: Aprataxin and PNK-like factor



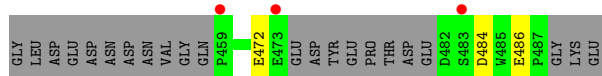
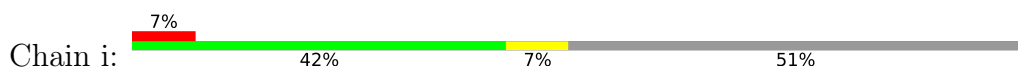
- Molecule 5: Aprataxin and PNK-like factor



- Molecule 5: Aprataxin and PNK-like factor



- Molecule 5: Aprataxin and PNK-like factor



- Molecule 5: Aprataxin and PNK-like factor



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	105.22Å 189.69Å 204.32Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	19.99 – 2.35 19.99 – 2.20	Depositor EDS
% Data completeness (in resolution range)	100.0 (19.99-2.35) 99.9 (19.99-2.20)	Depositor EDS
$R_{merge}$	0.14	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.21 (at 2.19Å)	Xtrriage
Refinement program	PHENIX 1.17.1_3660	Depositor
R, $R_{free}$	0.177 , 0.233 0.177 , 0.233	Depositor DCC
$R_{free}$ test set	10078 reflections (4.88%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	45.6	Xtrriage
Anisotropy	0.283	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 51.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	24986	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	62.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.20% of the height of the origin peak. No significant pseudotranslation is detected.*

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

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  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: CL, GOL

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	0.47	0/805	0.59	0/1088
1	F	0.38	0/805	0.56	0/1088
1	K	0.41	0/805	0.55	0/1088
1	P	0.41	0/814	0.58	0/1099
1	U	0.40	0/805	0.53	0/1088
1	Z	0.40	0/805	0.57	0/1088
1	e	0.39	0/796	0.53	0/1077
1	j	0.39	0/805	0.54	0/1088
2	B	0.54	0/723	0.55	0/976
2	G	0.53	0/709	0.53	0/957
2	L	0.56	0/696	0.53	0/939
2	Q	0.54	0/696	0.50	0/939
2	V	0.46	0/705	0.51	0/951
2	a	0.56	0/718	0.53	0/969
2	f	0.53	0/696	0.52	0/939
2	k	0.52	0/696	0.58	0/939
3	C	0.44	0/784	0.55	0/1052
3	H	0.45	0/784	0.61	0/1052
3	M	0.44	0/795	0.55	0/1067
3	R	0.42	0/784	0.58	0/1052
3	W	0.42	0/784	0.55	0/1052
3	b	0.46	0/800	0.56	0/1073
3	g	0.43	0/784	0.54	0/1052
3	l	0.43	0/784	0.55	0/1052
4	D	0.44	0/613	0.53	0/821
4	I	0.44	0/613	0.56	0/821
4	N	0.41	0/609	0.54	0/816
4	S	0.44	0/621	0.58	0/832
4	X	0.42	0/625	0.59	0/838
4	c	0.45	0/617	0.59	0/827
4	h	0.45	0/609	0.57	0/816
4	m	0.43	0/605	0.55	0/810

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# Z  >5	RMSZ	# Z  >5
5	E	0.35	0/192	0.50	0/260
5	J	0.63	0/255	0.59	0/349
5	O	0.35	0/183	0.51	0/248
5	T	0.50	0/192	0.52	0/260
5	Y	0.42	0/192	0.51	0/260
5	d	0.54	0/174	0.62	0/236
5	i	0.58	0/183	0.58	0/248
5	n	0.62	0/174	0.75	0/237
All	All	0.46	0/24835	0.56	0/33444

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

## 5.2 Too-close contacts [i](#)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	795	0	846	4	0
1	F	795	0	846	10	0
1	K	795	0	846	3	0
1	P	804	0	859	10	0
1	U	795	0	846	14	1
1	Z	795	0	846	6	0
1	e	786	0	833	0	0
1	j	795	0	846	0	0
2	B	708	0	723	12	0
2	G	697	0	711	13	0
2	L	685	0	703	3	0
2	Q	685	0	703	5	0
2	V	694	0	709	10	0
2	a	703	0	718	0	0
2	f	685	0	703	0	0
2	k	685	0	703	0	0
3	C	774	0	813	7	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	H	774	0	813	7	0
3	M	784	0	820	6	0
3	R	774	0	813	10	0
3	W	774	0	813	6	0
3	b	790	0	831	0	0
3	g	774	0	813	0	0
3	l	774	0	813	0	1
4	D	606	0	650	11	0
4	I	606	0	650	7	0
4	N	602	0	647	5	0
4	S	614	0	656	12	0
4	X	618	0	657	12	0
4	c	610	0	653	0	0
4	h	602	0	647	0	0
4	m	598	0	639	0	0
5	E	188	0	139	0	0
5	J	248	0	183	3	0
5	O	179	0	133	0	0
5	T	188	0	139	3	0
5	Y	188	0	139	2	0
5	d	170	0	127	0	0
5	i	179	0	133	0	0
5	n	170	0	128	0	0
6	A	6	0	8	0	0
6	D	6	0	8	0	0
6	H	6	0	8	0	0
6	K	6	0	8	1	0
6	N	6	0	8	2	0
6	P	6	0	8	0	0
6	S	6	0	8	1	0
6	U	6	0	8	2	0
6	b	6	0	8	0	0
6	c	6	0	8	0	0
6	e	6	0	8	0	0
6	h	6	0	8	0	0
6	j	12	0	16	0	0
7	J	1	0	0	0	0
8	A	16	0	0	0	0
8	B	10	0	0	0	0
8	C	14	0	0	0	0
8	D	13	0	0	0	0
8	E	1	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
8	F	17	0	0	0	0
8	G	7	0	0	0	0
8	H	19	0	0	0	0
8	I	13	0	0	0	0
8	J	4	0	0	0	0
8	K	10	0	0	0	0
8	L	10	0	0	0	0
8	M	23	0	0	0	0
8	N	6	0	0	0	0
8	O	1	0	0	0	0
8	P	10	0	0	0	0
8	Q	5	0	0	0	0
8	R	19	0	0	0	0
8	S	7	0	0	0	0
8	U	16	0	0	0	0
8	V	6	0	0	0	0
8	W	19	0	0	0	0
8	X	13	0	0	0	0
8	Y	2	0	0	0	0
8	Z	9	0	0	0	0
8	a	6	0	0	0	0
8	b	25	0	0	0	0
8	c	7	0	0	0	0
8	d	2	0	0	0	0
8	e	8	0	0	0	0
8	f	3	0	0	0	0
8	g	25	0	0	0	0
8	h	9	0	0	0	0
8	i	1	0	0	0	0
8	j	15	0	0	0	0
8	k	8	0	0	0	0
8	l	18	0	0	0	0
8	m	18	0	0	0	0
All	All	24986	0	25402	135	1

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 5.

The worst 5 of 135 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:D:84:MET:CE	4:D:101:GLY:HA3	1.88	1.02
4:X:26:ILE:CD1	4:X:59:LYS:HB2	2.06	0.85
1:F:118:LYS:N	1:F:118:LYS:HD3	1.97	0.79
4:D:84:MET:HE1	4:D:101:GLY:HA3	1.63	0.79
2:B:98:VAL:HG13	2:B:102:LEU:HD22	1.66	0.78

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:U:61:GLU:OE1	3:l:42:ARG:NH1[2_555]	1.80	0.40

## 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	101/107 (94%)	100 (99%)	1 (1%)	0	100	100
1	F	101/107 (94%)	99 (98%)	2 (2%)	0	100	100
1	K	101/107 (94%)	99 (98%)	2 (2%)	0	100	100
1	P	102/107 (95%)	100 (98%)	2 (2%)	0	100	100
1	U	101/107 (94%)	99 (98%)	2 (2%)	0	100	100
1	Z	101/107 (94%)	97 (96%)	4 (4%)	0	100	100
1	e	100/107 (94%)	96 (96%)	4 (4%)	0	100	100
1	j	101/107 (94%)	99 (98%)	2 (2%)	0	100	100
2	B	89/100 (89%)	88 (99%)	1 (1%)	0	100	100
2	G	87/100 (87%)	86 (99%)	0	1 (1%)	14	13
2	L	86/100 (86%)	84 (98%)	1 (1%)	1 (1%)	13	11
2	Q	86/100 (86%)	85 (99%)	0	1 (1%)	13	11
2	V	87/100 (87%)	85 (98%)	1 (1%)	1 (1%)	14	13

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	a	88/100 (88%)	87 (99%)	0	1 (1%)	14	13
2	f	86/100 (86%)	84 (98%)	1 (1%)	1 (1%)	13	11
2	k	86/100 (86%)	85 (99%)	0	1 (1%)	13	11
3	C	92/99 (93%)	92 (100%)	0	0	100	100
3	H	92/99 (93%)	89 (97%)	3 (3%)	0	100	100
3	M	93/99 (94%)	93 (100%)	0	0	100	100
3	R	92/99 (93%)	91 (99%)	1 (1%)	0	100	100
3	W	92/99 (93%)	91 (99%)	1 (1%)	0	100	100
3	b	94/99 (95%)	93 (99%)	1 (1%)	0	100	100
3	g	92/99 (93%)	92 (100%)	0	0	100	100
3	l	92/99 (93%)	92 (100%)	0	0	100	100
4	D	74/84 (88%)	73 (99%)	1 (1%)	0	100	100
4	I	74/84 (88%)	72 (97%)	2 (3%)	0	100	100
4	N	73/84 (87%)	70 (96%)	3 (4%)	0	100	100
4	S	75/84 (89%)	72 (96%)	3 (4%)	0	100	100
4	X	75/84 (89%)	74 (99%)	1 (1%)	0	100	100
4	c	74/84 (88%)	73 (99%)	1 (1%)	0	100	100
4	h	73/84 (87%)	72 (99%)	1 (1%)	0	100	100
4	m	73/84 (87%)	72 (99%)	1 (1%)	0	100	100
5	E	18/43 (42%)	18 (100%)	0	0	100	100
5	J	27/43 (63%)	24 (89%)	3 (11%)	0	100	100
5	O	17/43 (40%)	16 (94%)	1 (6%)	0	100	100
5	T	18/43 (42%)	18 (100%)	0	0	100	100
5	Y	18/43 (42%)	18 (100%)	0	0	100	100
5	d	16/43 (37%)	15 (94%)	1 (6%)	0	100	100
5	i	17/43 (40%)	17 (100%)	0	0	100	100
5	n	16/43 (37%)	15 (94%)	1 (6%)	0	100	100
All	All	2980/3464 (86%)	2925 (98%)	48 (2%)	7 (0%)	47	56

5 of 7 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	L	104	GLY

Continued on next page...



*Continued from previous page...*

Mol	Chain	Res	Type
2	G	104	GLY
2	f	104	GLY
2	a	104	GLY
2	V	104	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	82/85 (96%)	81 (99%)	1 (1%)	71	82
1	F	82/85 (96%)	81 (99%)	1 (1%)	71	82
1	K	82/85 (96%)	80 (98%)	2 (2%)	49	59
1	P	83/85 (98%)	81 (98%)	2 (2%)	49	59
1	U	82/85 (96%)	79 (96%)	3 (4%)	34	42
1	Z	82/85 (96%)	78 (95%)	4 (5%)	25	29
1	e	81/85 (95%)	80 (99%)	1 (1%)	71	82
1	j	82/85 (96%)	82 (100%)	0	100	100
2	B	77/86 (90%)	77 (100%)	0	100	100
2	G	76/86 (88%)	75 (99%)	1 (1%)	69	80
2	L	75/86 (87%)	73 (97%)	2 (3%)	44	55
2	Q	75/86 (87%)	75 (100%)	0	100	100
2	V	76/86 (88%)	74 (97%)	2 (3%)	46	56
2	a	77/86 (90%)	75 (97%)	2 (3%)	46	56
2	f	75/86 (87%)	73 (97%)	2 (3%)	44	55
2	k	75/86 (87%)	74 (99%)	1 (1%)	69	80
3	C	82/86 (95%)	81 (99%)	1 (1%)	71	82
3	H	82/86 (95%)	80 (98%)	2 (2%)	49	59
3	M	83/86 (96%)	82 (99%)	1 (1%)	71	82
3	R	82/86 (95%)	81 (99%)	1 (1%)	71	82

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
3	W	82/86 (95%)	82 (100%)	0	100	100
3	b	83/86 (96%)	80 (96%)	3 (4%)	35	43
3	g	82/86 (95%)	81 (99%)	1 (1%)	71	82
3	l	82/86 (95%)	82 (100%)	0	100	100
4	D	62/69 (90%)	62 (100%)	0	100	100
4	I	62/69 (90%)	61 (98%)	1 (2%)	62	75
4	N	62/69 (90%)	61 (98%)	1 (2%)	62	75
4	S	63/69 (91%)	60 (95%)	3 (5%)	25	30
4	X	64/69 (93%)	60 (94%)	4 (6%)	18	19
4	c	63/69 (91%)	61 (97%)	2 (3%)	39	47
4	h	62/69 (90%)	62 (100%)	0	100	100
4	m	61/69 (88%)	61 (100%)	0	100	100
5	E	22/40 (55%)	22 (100%)	0	100	100
5	J	29/40 (72%)	25 (86%)	4 (14%)	3	3
5	O	21/40 (52%)	19 (90%)	2 (10%)	8	7
5	T	22/40 (55%)	21 (96%)	1 (4%)	27	33
5	Y	22/40 (55%)	22 (100%)	0	100	100
5	d	20/40 (50%)	19 (95%)	1 (5%)	24	28
5	i	21/40 (52%)	18 (86%)	3 (14%)	3	3
5	n	20/40 (50%)	20 (100%)	0	100	100
All	All	2596/2928 (89%)	2541 (98%)	55 (2%)	53	65

5 of 55 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	U	71	ARG
1	Z	32	ARG
2	k	86	ARG
2	f	112	SER
2	V	87	SER

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

Mol	Chain	Res	Type
4	X	27	GLN
4	X	93	GLN
1	j	24	GLN
2	a	109	HIS
4	S	25	ASN

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

### 5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

### 5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

### 5.6 Ligand geometry [i](#)

Of 15 ligands modelled in this entry, 1 is monoatomic - leaving 14 for Mogul analysis.

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	$\# Z  > 2$
6	GOL	b	201	-	5,5,5	0.08	0	5,5,5	0.32	0
6	GOL	j	201	-	5,5,5	0.09	0	5,5,5	0.32	0
6	GOL	H	201	-	5,5,5	0.10	0	5,5,5	0.32	0
6	GOL	U	201	-	5,5,5	0.09	0	5,5,5	0.32	0
6	GOL	e	201	-	5,5,5	0.09	0	5,5,5	0.32	0
6	GOL	P	201	-	5,5,5	0.09	0	5,5,5	0.32	0
6	GOL	j	202	-	5,5,5	0.07	0	5,5,5	0.32	0
6	GOL	K	201	-	5,5,5	0.09	0	5,5,5	0.32	0
6	GOL	h	201	-	5,5,5	0.10	0	5,5,5	0.31	0
6	GOL	A	201	-	5,5,5	0.09	0	5,5,5	0.32	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
6	GOL	N	201	-	5,5,5	0.08	0	5,5,5	0.30	0
6	GOL	S	201	-	5,5,5	0.08	0	5,5,5	0.32	0
6	GOL	D	201	-	5,5,5	0.09	0	5,5,5	0.32	0
6	GOL	c	201	-	5,5,5	0.09	0	5,5,5	0.32	0

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
6	GOL	b	201	-	-	2/4/4/4	-
6	GOL	j	201	-	-	2/4/4/4	-
6	GOL	H	201	-	-	1/4/4/4	-
6	GOL	U	201	-	-	1/4/4/4	-
6	GOL	e	201	-	-	2/4/4/4	-
6	GOL	P	201	-	-	2/4/4/4	-
6	GOL	j	202	-	-	2/4/4/4	-
6	GOL	K	201	-	-	2/4/4/4	-
6	GOL	h	201	-	-	0/4/4/4	-
6	GOL	A	201	-	-	1/4/4/4	-
6	GOL	N	201	-	-	2/4/4/4	-
6	GOL	S	201	-	-	3/4/4/4	-
6	GOL	D	201	-	-	1/4/4/4	-
6	GOL	c	201	-	-	1/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 22 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	N	201	GOL	O1-C1-C2-C3
6	P	201	GOL	O1-C1-C2-O2
6	b	201	GOL	O1-C1-C2-O2
6	b	201	GOL	O1-C1-C2-C3
6	P	201	GOL	O1-C1-C2-C3

There are no ring outliers.

4 monomers are involved in 6 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	U	201	GOL	2	0
6	K	201	GOL	1	0
6	N	201	GOL	2	0
6	S	201	GOL	1	0

## 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 [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

In the following table, the column labelled ‘#RSRZ > 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95<sup>th</sup> percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q < 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
1	A	103/107 (96%)	-0.13	3 (2%) 51 62	39, 53, 80, 112	0
1	F	103/107 (96%)	0.04	7 (6%) 17 25	44, 64, 94, 105	0
1	K	103/107 (96%)	-0.05	4 (3%) 39 52	43, 57, 86, 110	0
1	P	104/107 (97%)	0.01	6 (5%) 23 33	44, 58, 88, 125	0
1	U	103/107 (96%)	0.01	5 (4%) 29 42	44, 59, 93, 94	0
1	Z	103/107 (96%)	0.02	6 (5%) 23 33	45, 59, 86, 96	0
1	e	102/107 (95%)	0.22	10 (9%) 7 12	48, 63, 97, 105	0
1	j	103/107 (96%)	0.02	4 (3%) 39 52	46, 60, 88, 114	0
2	B	90/100 (90%)	-0.32	4 (4%) 34 46	41, 55, 75, 113	0
2	G	88/100 (88%)	-0.09	2 (2%) 60 70	47, 61, 85, 99	0
2	L	88/100 (88%)	-0.42	0 100 100	45, 58, 84, 113	0
2	Q	88/100 (88%)	-0.22	1 (1%) 80 87	45, 58, 80, 95	0
2	V	89/100 (89%)	-0.20	1 (1%) 80 87	47, 62, 82, 98	0
2	a	89/100 (89%)	-0.27	1 (1%) 80 87	47, 58, 76, 96	0
2	f	88/100 (88%)	-0.11	2 (2%) 60 70	50, 65, 92, 103	0
2	k	88/100 (88%)	-0.08	3 (3%) 45 57	48, 61, 86, 104	0
3	C	94/99 (94%)	-0.04	3 (3%) 47 59	39, 52, 78, 118	0
3	H	94/99 (94%)	-0.04	3 (3%) 47 59	38, 49, 78, 128	0
3	M	95/99 (95%)	-0.04	4 (4%) 36 48	41, 55, 85, 129	0
3	R	94/99 (94%)	-0.06	2 (2%) 63 74	40, 52, 83, 118	0
3	W	94/99 (94%)	-0.14	5 (5%) 26 38	40, 55, 82, 104	0
3	b	96/99 (96%)	-0.00	6 (6%) 20 29	42, 53, 84, 146	0
3	g	94/99 (94%)	-0.08	3 (3%) 47 59	41, 53, 76, 126	0
3	l	94/99 (94%)	0.08	4 (4%) 35 47	43, 54, 78, 119	0

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
4	D	76/84 (90%)	0.07	6 (7%) 12 19	42, 54, 83, 95	0
4	I	76/84 (90%)	-0.10	3 (3%) 39 52	38, 50, 76, 102	0
4	N	75/84 (89%)	0.06	4 (5%) 26 38	43, 54, 83, 101	0
4	S	77/84 (91%)	0.05	4 (5%) 27 39	41, 51, 86, 123	0
4	X	77/84 (91%)	0.08	4 (5%) 27 39	44, 54, 92, 149	0
4	c	76/84 (90%)	-0.16	1 (1%) 77 84	41, 52, 78, 114	0
4	h	75/84 (89%)	0.03	4 (5%) 26 38	39, 53, 85, 108	0
4	m	75/84 (89%)	0.05	5 (6%) 17 26	42, 55, 79, 104	0
5	E	22/43 (51%)	0.88	5 (22%) 0 1	57, 81, 113, 133	0
5	J	29/43 (67%)	0.70	7 (24%) 0 1	56, 79, 127, 146	0
5	O	21/43 (48%)	0.76	4 (19%) 1 2	64, 80, 116, 121	0
5	T	22/43 (51%)	0.70	5 (22%) 0 1	53, 78, 109, 122	0
5	Y	22/43 (51%)	1.10	6 (27%) 0 0	54, 79, 119, 131	0
5	d	20/43 (46%)	0.75	5 (25%) 0 1	58, 79, 108, 113	0
5	i	21/43 (48%)	0.86	3 (14%) 2 4	50, 76, 105, 121	0
5	n	20/43 (46%)	1.13	5 (25%) 0 1	61, 84, 126, 138	0
All	All	3071/3464 (88%)	-0.00	160 (5%) 27 39	38, 57, 92, 149	0

The worst 5 of 160 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	R	41	TYR	6.8
3	M	41	TYR	6.8
3	H	41	TYR	6.5
5	J	472	GLU	6.1
3	l	41	TYR	6.1

## 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<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
6	GOL	D	201	6/6	0.76	0.19	75,84,86,87	0
6	GOL	h	201	6/6	0.78	0.15	88,90,92,94	0
6	GOL	N	201	6/6	0.80	0.23	61,85,91,98	0
6	GOL	e	201	6/6	0.82	0.24	67,72,77,81	0
6	GOL	U	201	6/6	0.82	0.17	68,72,76,78	0
6	GOL	H	201	6/6	0.87	0.18	93,93,96,100	0
6	GOL	b	201	6/6	0.87	0.24	84,84,87,88	0
6	GOL	j	202	6/6	0.89	0.15	79,82,89,90	0
6	GOL	j	201	6/6	0.91	0.15	55,65,72,73	0
6	GOL	S	201	6/6	0.91	0.21	92,93,94,96	0
6	GOL	P	201	6/6	0.92	0.12	59,71,73,83	0
6	GOL	K	201	6/6	0.92	0.14	61,69,73,74	0
7	CL	J	501	1/1	0.93	0.17	77,77,77,77	0
6	GOL	c	201	6/6	0.94	0.20	71,73,74,76	0
6	GOL	A	201	6/6	0.95	0.10	52,66,67,67	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.