



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 29, 2020 – 08:30 am BST

PDB ID : 5OY7  
Title : Structure of the 4\_601\_157 tetranucleosome (P1 form)  
Authors : Ekundayo, B.; Richmond, T.J.; Schalch, T.  
Deposited on : 2017-09-07  
Resolution : 5.77 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 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	:	2.11
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)	:	2.11

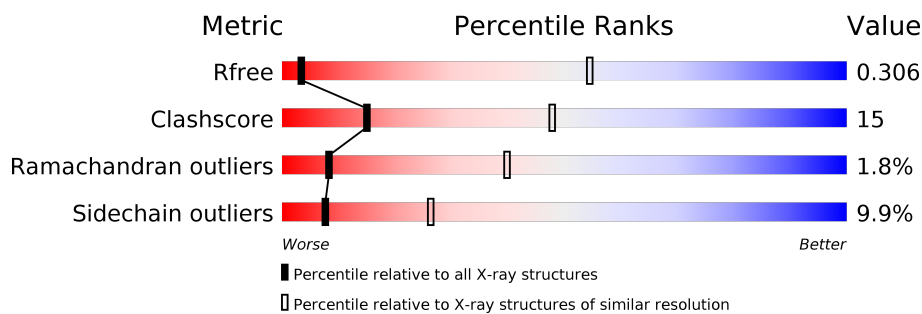
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 5.77 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	1008 (7.70-3.86)
Clashscore	141614	1034 (7.66-3.90)
Ramachandran outliers	138981	1003 (7.70-3.86)
Sidechain outliers	138945	1005 (7.70-3.82)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria 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  $\leq 5\%$ .

Mol	Chain	Length	Quality of chain
1	D	126	<div> <div>48%</div> <div>23%</div> <div>• •</div> <div>25%</div> </div>
1	H	126	<div> <div>45%</div> <div>26%</div> <div>•</div> <div>26%</div> </div>
1	L	126	<div> <div>44%</div> <div>26%</div> <div>5% •</div> <div>25%</div> </div>
1	P	126	<div> <div>46%</div> <div>25%</div> <div>•</div> <div>26%</div> </div>
1	T	126	<div> <div>45%</div> <div>25%</div> <div>5% •</div> <div>25%</div> </div>
1	X	126	<div> <div>45%</div> <div>26%</div> <div>•</div> <div>26%</div> </div>
1	b	126	<div> <div>65%</div> <div>9%</div> <div>•</div> <div>25%</div> </div>



*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	f	126	
2	A	135	
2	E	135	
2	I	135	
2	M	135	
2	Q	135	
2	U	135	
2	Y	135	
2	c	135	
3	B	102	
3	F	102	
3	J	102	
3	N	102	
3	R	102	
3	V	102	
3	Z	102	
3	d	102	
4	C	130	
4	G	130	
4	K	130	
4	O	130	
4	S	130	
4	W	130	
4	a	130	
4	e	130	

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
5	g	634	 73% 25% •
6	h	628	 74% 24% •

## 2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 49215 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 H2B 1.1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	L	95	Total	C	N	O	S	0	0	0
			745	469	134	140	2			
1	P	93	Total	C	N	O	S	0	0	0
			726	457	130	137	2			
1	D	95	Total	C	N	O	S	0	0	0
			745	469	134	140	2			
1	H	93	Total	C	N	O	S	0	0	0
			726	457	130	137	2			
1	T	95	Total	C	N	O	S	0	0	0
			745	469	134	140	2			
1	X	93	Total	C	N	O	S	0	0	0
			726	457	130	137	2			
1	b	95	Total	C	N	O	S	0	0	0
			745	469	134	140	2			
1	f	93	Total	C	N	O	S	0	0	0
			726	457	130	137	2			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
L	29	THR	SER	conflict	UNP P02281
P	29	THR	SER	conflict	UNP P02281
D	29	THR	SER	conflict	UNP P02281
H	29	THR	SER	conflict	UNP P02281
T	29	THR	SER	conflict	UNP P02281
X	29	THR	SER	conflict	UNP P02281
b	29	THR	SER	conflict	UNP P02281
f	29	THR	SER	conflict	UNP P02281

- Molecule 2 is a protein called Histone H3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	M	97	Total	C	N	O	S	0	0	0
			801	504	155	139	3			
2	Q	97	Total	C	N	O	S	0	0	0
			802	506	155	138	3			
2	I	97	Total	C	N	O	S	0	0	0
			802	506	155	138	3			
2	A	97	Total	C	N	O	S	0	0	0
			802	506	155	138	3			
2	E	97	Total	C	N	O	S	0	0	0
			801	504	155	139	3			
2	U	97	Total	C	N	O	S	0	0	0
			801	504	155	139	3			
2	Y	97	Total	C	N	O	S	0	0	0
			802	506	155	138	3			
2	c	97	Total	C	N	O	S	0	0	0
			801	504	155	139	3			

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
M	102	ALA	GLY	conflict	UNP Q92133
M	111	ALA	GLY	conflict	UNP Q92133
Q	102	ALA	GLY	conflict	UNP Q92133
Q	111	ALA	GLY	conflict	UNP Q92133
I	102	ALA	GLY	conflict	UNP Q92133
I	111	ALA	GLY	conflict	UNP Q92133
A	102	ALA	GLY	conflict	UNP Q92133
A	111	ALA	GLY	conflict	UNP Q92133
E	102	ALA	GLY	conflict	UNP Q92133
E	111	ALA	GLY	conflict	UNP Q92133
U	102	ALA	GLY	conflict	UNP Q92133
U	111	ALA	GLY	conflict	UNP Q92133
Y	102	ALA	GLY	conflict	UNP Q92133
Y	111	ALA	GLY	conflict	UNP Q92133
c	102	ALA	GLY	conflict	UNP Q92133
c	111	ALA	GLY	conflict	UNP Q92133

- Molecule 3 is a protein called Histone H4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	N	78	Total	C	N	O	S	0	0	0
			619	391	120	107	1			

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	R	82	Total	C	N	O	S	0	0	0
			654	412	128	113	1			
3	J	83	Total	C	N	O	S	0	0	0
			662	418	129	114	1			
3	B	83	Total	C	N	O	S	0	0	0
			662	418	129	114	1			
3	F	78	Total	C	N	O	S	0	0	0
			619	391	120	107	1			
3	V	78	Total	C	N	O	S	0	0	0
			619	391	120	107	1			
3	Z	83	Total	C	N	O	S	0	0	0
			662	418	129	114	1			
3	d	78	Total	C	N	O	S	0	0	0
			619	391	120	107	1			

- Molecule 4 is a protein called Histone H2A.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
4	O	105	Total	C	N	O	0	0	0
			809	510	158	141			
4	C	103	Total	C	N	O	0	0	0
			795	501	155	139			
4	G	105	Total	C	N	O	0	0	0
			809	510	158	141			
4	K	103	Total	C	N	O	0	0	0
			795	501	155	139			
4	S	103	Total	C	N	O	0	0	0
			795	501	155	139			
4	W	105	Total	C	N	O	0	0	0
			809	510	158	141			
4	a	103	Total	C	N	O	0	0	0
			795	501	155	139			
4	e	105	Total	C	N	O	0	0	0
			809	510	158	141			

- Molecule 5 is a DNA chain called DNA (619-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	g	619	Total	C	N	O	P	0	0	0
			12605	5987	2278	3721	619			

- Molecule 6 is a DNA chain called DNA (619-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	h	619	Total	C	N	O	P	0	0	0
			12774	6041	2407	3707	619			

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

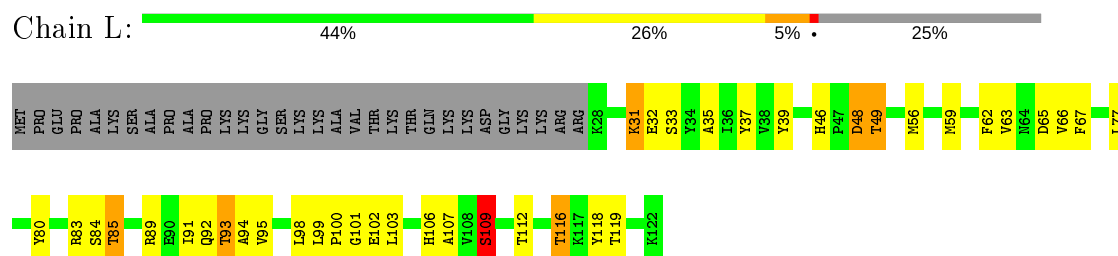
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	G	1	Total	Cl	1	0
			1	1		
7	K	1	Total	Cl	1	0
			1	1		
7	e	1	Total	Cl	1	0
			1	1		
7	C	1	Total	Cl	1	0
			1	1		
7	W	1	Total	Cl	1	0
			1	1		
7	a	1	Total	Cl	1	0
			1	1		
7	O	1	Total	Cl	1	0
			1	1		
7	S	1	Total	Cl	1	0
			1	1		



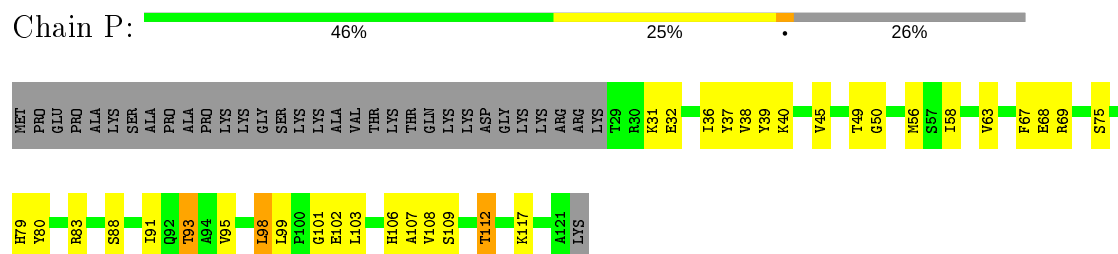
### 3 Residue-property plots [i](#)

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

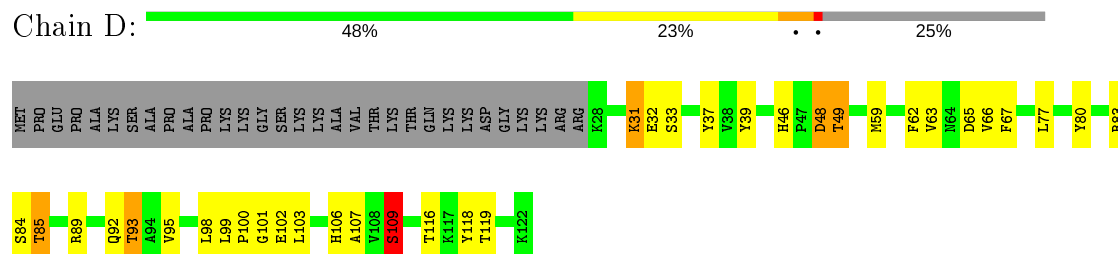
#### • Molecule 1: Histone H2B 1.1



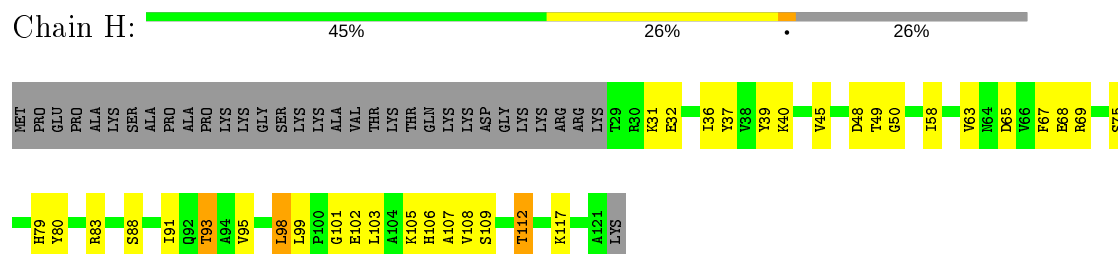
#### • Molecule 1: Histone H2B 1.1



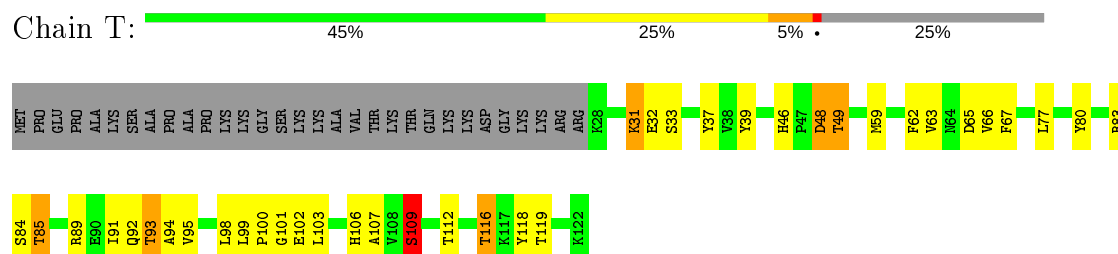
#### • Molecule 1: Histone H2B 1.1



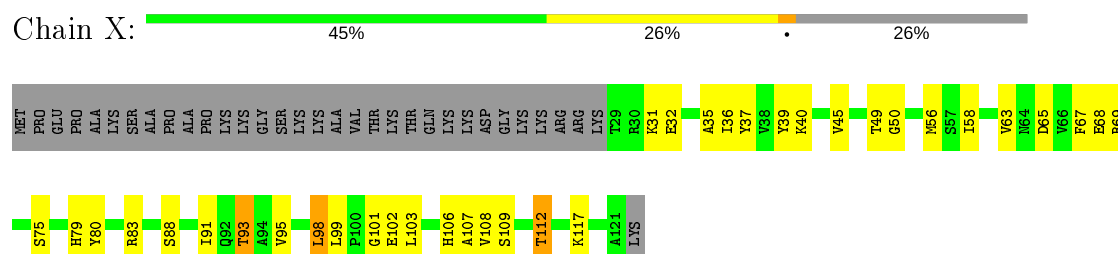
#### • Molecule 1: Histone H2B 1.1



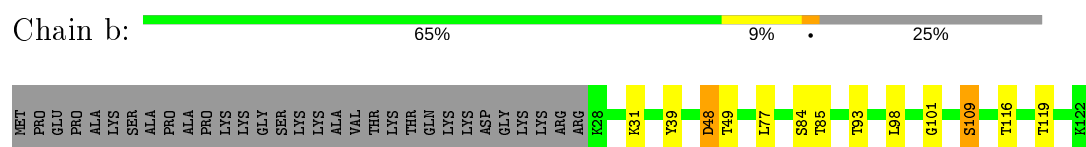
- Molecule 1: Histone H2B 1.1



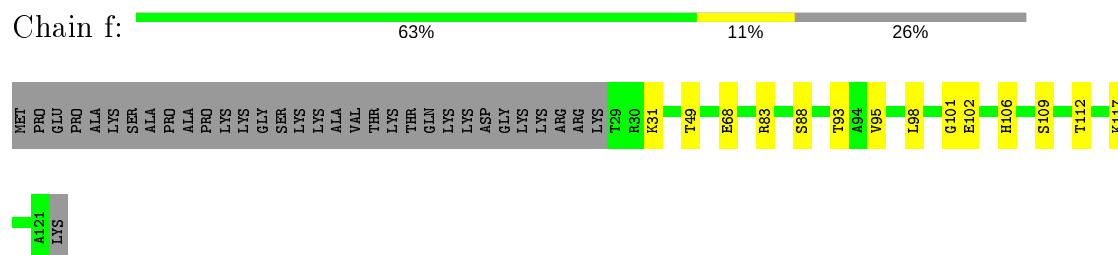
- Molecule 1: Histone H2B 1.1



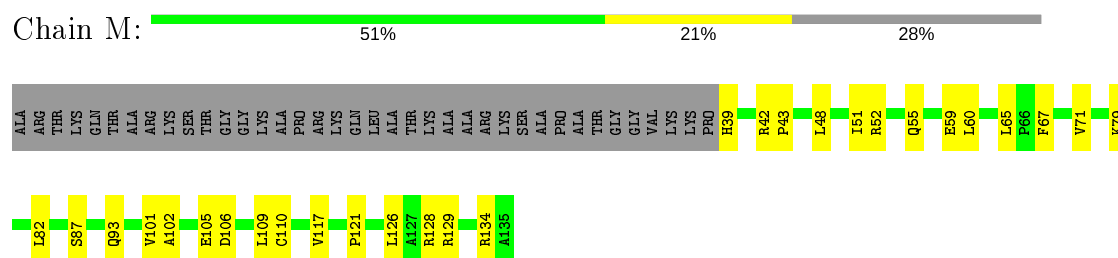
- Molecule 1: Histone H2B 1.1



- Molecule 1: Histone H2B 1.1

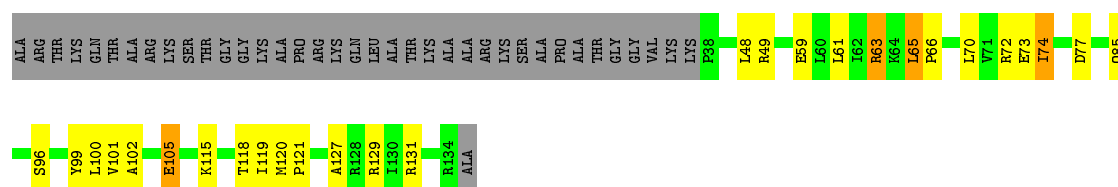


- Molecule 2: Histone H3

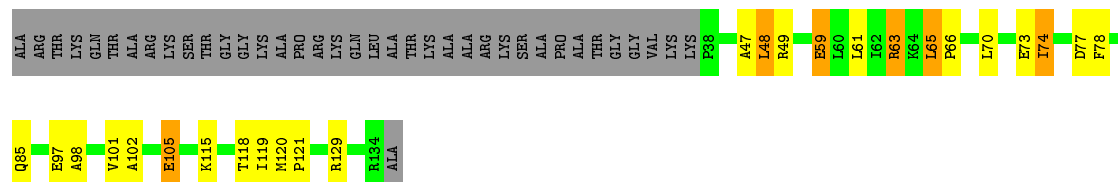


- Molecule 2: Histone H3

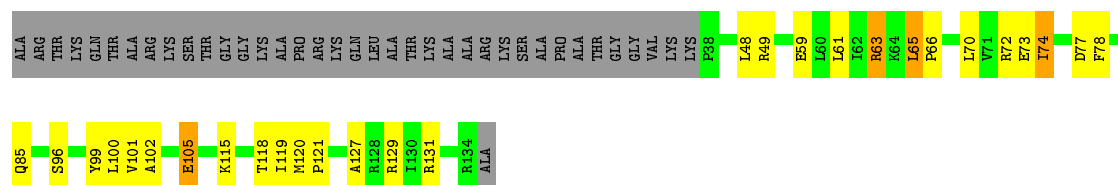




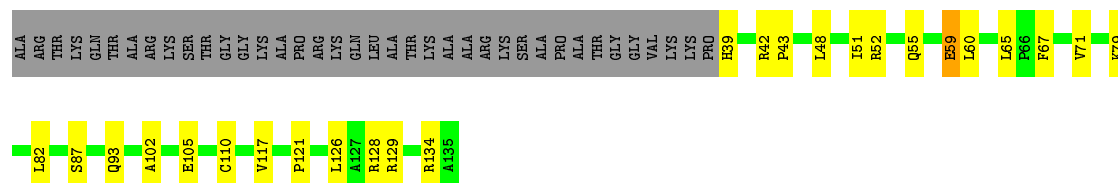
- Molecule 2: Histone H3



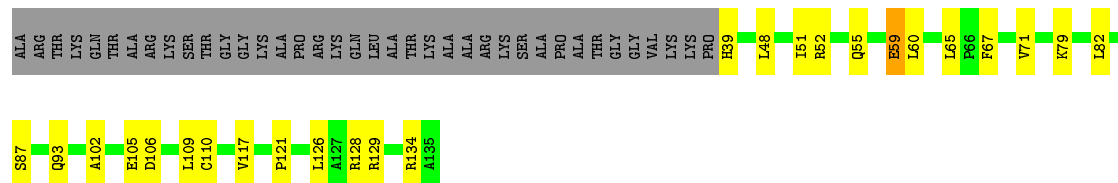
- Molecule 2: Histone H3



- Molecule 2: Histone H3

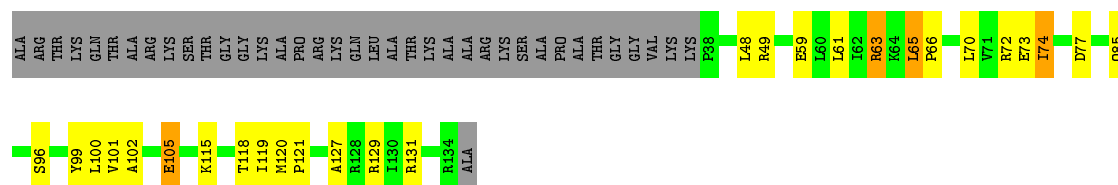


- Molecule 2: Histone H3



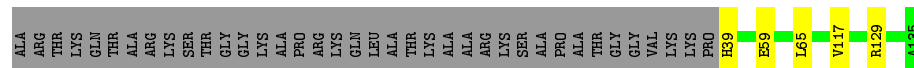
- Molecule 2: Histone H3





- Molecule 2: Histone H3

Chain c: 68% 28%



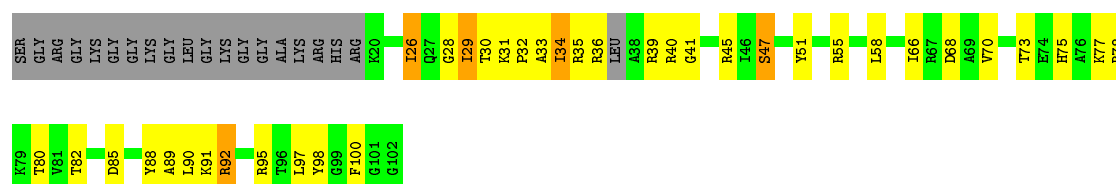
- Molecule 3: Histone H4

Chain N: 60% 15% 24%



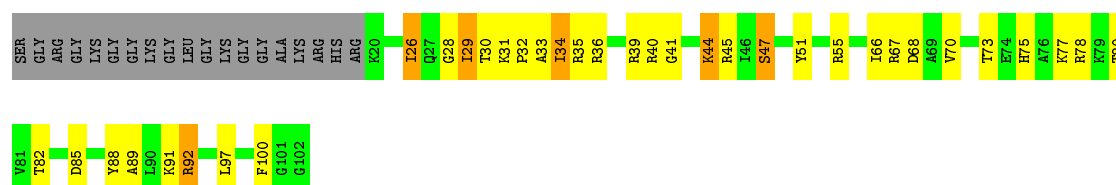
- Molecule 3: Histone H4

Chain R: 44% 31% 5% 20%



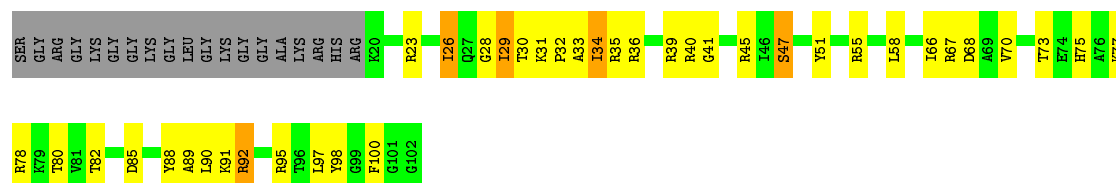
- Molecule 3: Histone H4

Chain J: 47% 28% 6% 19%



- Molecule 3: Histone H4

Chain B: 43% 33% 5% 19%





Chain e:  74% 7% 19%

MET	SER	GLY	ARG	GLY	LYS	GLN	GLY	THR	ARG	ALA	LYS	ALA	LYS	LYS
ALA	LYS	LYS	LYS	LYS	LYS	LYS	LYS	LYS	LYS	LYS	LYS	LYS	LYS	LYS

• Molecule 5: DNA (619-MER)

Chain g:  73% 25%

DA	DG	T120	C239	A358	G495	C526	C563	C569	T577	T585	T588	T618	G619	T620	DG	DT
DT	DA	C121	C239	C359	C496	A532	A533	A570	G574	T585	C589	G619	G621	G622	DG	DT
DA	DT	T122	T247	A360	T500	A534	C543	C574	T578	T585	C589	T618	G621	G622	DG	DT
DC	DA	G133	A251	C362	T504	C535	C545	G574	A579	C582	T590	T618	G622	DG	DT	DC
C4	G8	G151	A251	C370	T504	C535	C545	G574	A579	C582	T590	T618	G622	DG	DT	DC
G8	G26	T152	A258	A376	G507	A516	C545	G574	A579	C582	T590	T618	G622	DG	DT	DC
G26	G27	C154	G262	A378	C509	A517	C546	G575	A580	C583	T591	T619	G623	DG	DT	DC
G27	C28	G156	T265	C379	G513	A518	C547	G576	A581	C584	T592	T620	G624	DG	DT	DC
C28	C29	T157	T266	C387	A514	A519	C548	G577	A582	C585	T593	T621	G625	DG	DT	DC
C29	G30	A158	A267	G388	C515	A520	C549	G578	A583	C586	T594	T622	G626	DG	DT	DC
G30	C31	C159	T273	C389	C517	A521	C550	G579	A584	C587	T595	T623	G627	DG	DT	DC
C31	T32	G162	T276	C391	C518	A522	C551	G580	A585	C588	T596	T624	G628	DG	DT	DC
T32	T36	G165	T277	C392	C526	A523	C552	G581	A586	C589	T597	T625	G629	DG	DT	DC
T36	G39	A170	T278	T403	C527	A524	C553	G582	A587	C590	T598	T626	G630	DG	DT	DC
G39	T40	C174	G289	A407	T508	A525	C554	G583	A588	C591	T599	T627	G631	DG	DT	DC
C41	C45	DG	G293	C413	T509	A526	C555	G584	A589	C592	T600	T628	G632	DG	DT	DC
G45	A46	G175	G307	A414	C496	A527	C556	G585	A590	C593	T601	T629	G633	DG	DT	DC
A46	A48	G182	G307	A414	C496	A527	C556	G585	A590	C593	T601	T629	G633	DG	DT	DC
A48	G49	C183	G312	G418	T548	A528	C557	G586	A591	C594	T602	T630	G634	DG	DT	DC
G49	C50	C184	T313	T313	T549	A529	C558	G587	A592	C595	T603	T631	G635	DG	DT	DC
C50	C58	C185	A314	A423	T550	A530	C559	G588	A593	C596	T604	T632	G636	DG	DT	DC
C58	A64	G186	C315	C423	T551	A531	C560	G589	A594	C597	T605	T633	G637	DG	DT	DC
A64	A65	T188	C319	C426	A563	A532	C561	G590	A595	C598	T606	T634	G638	DG	DT	DC
A65	A66	T192	T320	T429	C563	A533	C562	G591	A596	C599	T607	T635	G639	DG	DT	DC
A66	C67	G195	G321	T429	C569	A534	C563	G592	A597	C600	T608	T636	G640	DG	DT	DC
C67	C75	T196	A326	T432	A570	A535	C564	G593	A598	C601	T609	T637	G641	DG	DT	DC
C75	G76	C197	A326	T434	G574	A536	C565	G594	A599	C602	T610	T638	G642	DG	DT	DC
G76	C77	G201	C330	T434	T577	A537	C566	G595	A600	C603	T611	T639	G643	DG	DT	DC
C77	G78	A202	G332	G445	T578	A538	C567	G596	A601	C604	T612	T640	G644	DG	DT	DC
G78	C79	C203	DT	G445	A579	A539	C568	G597	A602	C605	T613	T641	G645	DG	DT	DC
C79	T80	A204	G333	G449	C582	A540	C569	G598	A603	C606	T614	T642	G646	DG	DT	DC
T80	T91	G205	G333	G449	C588	A541	C570	G599	A604	C607	T615	T643	G647	DG	DT	DC
T91	A95	C206	G338	G465	T585	A542	C571	G600	A605	C608	T616	T644	G648	DG	DT	DC
A95	C101	C214	G339	C466	T588	A543	C572	G601	A606	C609	T617	T645	G649	DG	DT	DC
C101	A102	A220	C340	A467	C589	A544	C573	G602	A607	C610	T618	T646	G650	DG	DT	DC
A102	G106	A221	G342	T472	T590	A545	C574	G603	A608	C611	T619	T647	G651	DG	DT	DC
G106	T109	A222	C343	C473	T591	A546	C575	G604	A609	C612	T620	T648	G652	DG	DT	DC
T109	T110	C223	T344	T344	T618	A547	C576	G605	A610	C613	T621	T649	G653	DG	DT	DC
T110	A111	G231	T348	T348	G619	A548	C577	G606	A611	C614	T622	T650	G654	DG	DT	DC
A111	T117	C232	G351	G351	T620	A549	C578	G607	A612	C615	T623	T651	G655	DG	DT	DC
T117		C233	T352	T352	G621	A550	C579	G608	A613	C616	T624	T652	G656	DG	DT	DC
		G234	C353	C353	G622	A551	C580	G609	A614	C617	T625	T653	G657	DG	DT	DC
		C235	G357	G357	G623	A552	C581	G610	A615	C618	T626	T654	G658	DG	DT	DC
		T236	C357	C357	G624	A553	C582	G611	A616	C619	T627	T655	G659	DG	DT	DC

• Molecule 6: DNA (619-MER)

Chain h:  74% 24%

A-113	G-112	C-104	T-103	T-102	G-101	G-100	T-96	C-78	T-74	A-73	T-70	T-65	G-59	G-58	T-57	G-56	G-52	A-51	G-50	C-49	C-45	C-39	C-38	A-37	A-36	G-33	T-25	C-8	G-4	DG	DA	DT												
C-234	T-230	A-229	T-226	T-221	G-215	G-214	T-213	G-212	G-208	A-207	G-206	C-205	C-201	C-195	G-194	A-193	A-192	G-189	T-181	A-176	DC	C-175	G-174	G-173	C-168	C-165	A-160	T-155	G-154	C-151	A-150	C-133	A-132	G-130	C-127	A-117	G-116	G-115	G-114					
G-364	A-363	C-362	C-361	C-357	C-351	C-350	A-349	A-348	G-345	T-337	G-333	DA	C-332	T-326	T-325	C-324	G-317	C-312	T-311	C-307	A-306	A-299	C-287	G-286	T-285	G-284	C-283	A-273	G-272	G-271	G-270	A-269	G-268	C-260	T-259	T-258	G-257	G-256	C-255	G-254	T-252			
A-488	DC	C-487	G-486	G-485	T-482	G-475	A-472	G-471	T-467	G-466	C-465	A-462	T-448	C-445	A-444	G-443	G-442	T-441	G-440	C-439	A-429	G-428	G-427	G-426	A-425	G-424	C-416	T-415	T-414	G-413	G-412	C-411	T-408	C-390	T-386	A-385	T-382	T-377	G-371	G-370	T-369	G-368		
DA	DT	DC	G-622	G-617	A-600	C-599	G-598	T-597	G-596	C-595	A-585	G-584	G-583	G-582	A-581	G-580	T-579	C-572	T-571	T-570	G-569	G-568	C-567	T-564	C-546	T-542	A-541	T-538	T-533	G-527	G-526	T-525	G-524	G-520	A-519	G-518	C-517	C-513	C-507	C-506	A-505	A-504	G-501	T-493

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	66.62Å 161.47Å 227.91Å 78.94° 83.86° 83.74°	Depositor
Resolution (Å)	111.41 – 5.77 141.52 – 5.77	Depositor EDS
% Data completeness (in resolution range)	93.8 (111.41-5.77) 93.8 (141.52-5.77)	Depositor EDS
$R_{merge}$	0.11	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.31 (at 5.77Å)	Xtriage
Refinement program	PHENIX (1.11.1 _2575: ???)	Depositor
R, $R_{free}$	0.219 , 0.238 0.289 , 0.306	Depositor DCC
$R_{free}$ test set	1226 reflections (5.04%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	222.9	Xtriage
Anisotropy	0.199	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.23 , 230.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.40$ , $\langle L^2 \rangle = 0.22$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.87	EDS
Total number of atoms	49215	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	246.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.50% 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

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	D	0.49	0/756	0.64	0/1015
1	H	0.62	0/737	0.70	0/993
1	L	0.49	0/756	0.64	0/1015
1	P	0.62	0/737	0.70	0/993
1	T	0.49	0/756	0.64	0/1015
1	X	0.62	0/737	0.70	0/993
1	b	0.49	0/756	0.64	0/1015
1	f	0.62	0/737	0.70	0/993
2	A	0.66	0/814	0.73	0/1092
2	E	0.41	0/812	0.57	0/1088
2	I	0.66	0/814	0.73	0/1092
2	M	0.41	0/812	0.58	0/1088
2	Q	0.66	0/814	0.73	0/1092
2	U	0.41	0/812	0.57	0/1088
2	Y	0.66	0/814	0.73	0/1092
2	c	0.41	0/812	0.58	0/1088
3	B	0.67	0/669	0.84	0/894
3	F	0.45	0/626	0.61	0/837
3	J	0.67	0/669	0.89	1/894 (0.1%)
3	N	0.46	0/626	0.62	0/837
3	R	0.67	0/660	0.83	0/880
3	V	0.45	0/626	0.61	0/837
3	Z	0.67	0/669	0.83	0/894
3	d	0.46	0/626	0.61	0/837
4	C	0.42	0/805	0.61	0/1088
4	G	0.63	0/819	0.75	0/1106
4	K	0.43	0/805	0.61	0/1088
4	O	0.63	0/819	0.76	0/1106
4	S	0.42	0/805	0.61	0/1088
4	W	0.63	0/819	0.76	0/1106
4	a	0.43	0/805	0.61	0/1088
4	e	0.63	0/819	0.76	0/1106

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# Z  >5	RMSZ	# Z  >5
5	g	0.80	28/14121 (0.2%)	1.34	144/21771 (0.7%)
6	h	0.82	29/14351 (0.2%)	1.36	165/22170 (0.7%)
All	All	0.70	57/52615 (0.1%)	1.12	310/76379 (0.4%)

The worst 5 of 57 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	h	-448	DT	O3'-P	26.30	1.92	1.61
6	h	-475	DG	C1'-N9	-8.79	1.34	1.47
5	g	622	DC	C1'-N1	7.52	1.59	1.49
5	g	151	DG	C1'-N9	-7.18	1.37	1.47
5	g	157	DT	C3'-O3'	7.08	1.53	1.44

The worst 5 of 310 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	h	-448	DT	P-O3'-C3'	29.45	155.04	119.70
5	g	158	DA	OP1-P-OP2	17.27	145.50	119.60
5	g	157	DT	OP1-P-O3'	-14.16	74.04	105.20
5	g	157	DT	OP2-P-O3'	-13.26	76.02	105.20
6	h	-448	DT	O3'-P-O5'	12.06	126.92	104.00

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	D	745	0	773	29	5
1	H	726	0	747	33	5
1	L	745	0	773	33	5
1	P	726	0	747	32	5
1	T	745	0	773	30	5
1	X	726	0	747	33	5
1	b	745	0	773	0	4
1	f	726	0	747	0	4

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	A	802	0	841	17	0
2	E	801	0	838	13	0
2	I	802	0	841	23	0
2	M	801	0	838	14	0
2	Q	802	0	841	16	0
2	U	801	0	838	13	0
2	Y	802	0	841	17	0
2	c	801	0	838	0	0
3	B	662	0	709	37	1
3	F	619	0	659	14	0
3	J	662	0	709	44	0
3	N	619	0	659	16	0
3	R	654	0	697	34	0
3	V	619	0	659	15	0
3	Z	662	0	709	29	0
3	d	619	0	659	0	0
4	C	795	0	846	23	0
4	G	809	0	864	40	0
4	K	795	0	846	23	0
4	O	809	0	864	41	0
4	S	795	0	846	23	0
4	W	809	0	864	41	0
4	a	795	0	846	0	0
4	e	809	0	864	0	0
5	g	12605	0	6944	0	0
6	h	12774	0	6940	0	1
7	C	1	0	0	0	0
7	G	1	0	0	0	0
7	K	1	0	0	0	0
7	O	1	0	0	0	0
7	S	1	0	0	0	0
7	W	1	0	0	0	0
7	a	1	0	0	0	0
7	e	1	0	0	0	0
All	All	49215	0	38980	507	20

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 15.

The worst 5 of 507 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)
1:L:93:THR:HG21	3:J:75:HIS:CD2	1.58	1.37
3:B:75:HIS:CD2	1:D:93:THR:HG21	1.58	1.37
3:R:75:HIS:CD2	1:T:93:THR:HG21	1.58	1.36
3:B:75:HIS:HD2	1:D:93:THR:CG2	1.55	1.20
1:L:93:THR:CG2	3:J:75:HIS:HD2	1.55	1.19

The worst 5 of 20 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:D:48:ASP:OD2	1:H:106:HIS:ND1[1_455]	0.19	2.01
1:L:48:ASP:OD2	1:P:106:HIS:ND1[1_655]	0.47	1.73
1:T:48:ASP:OD2	1:X:106:HIS:ND1[1_655]	0.66	1.54
1:b:48:ASP:OD2	1:f:106:HIS:ND1[1_455]	0.75	1.45
1:D:48:ASP:OD2	1:H:106:HIS:CE1[1_455]	1.30	0.90

## 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	D	93/126 (74%)	82 (88%)	9 (10%)	2 (2%)	6	35
1	H	91/126 (72%)	80 (88%)	10 (11%)	1 (1%)	14	51
1	L	93/126 (74%)	82 (88%)	9 (10%)	2 (2%)	6	35
1	P	91/126 (72%)	80 (88%)	10 (11%)	1 (1%)	14	51
1	T	93/126 (74%)	81 (87%)	10 (11%)	2 (2%)	6	35
1	X	91/126 (72%)	80 (88%)	10 (11%)	1 (1%)	14	51
1	b	93/126 (74%)	82 (88%)	9 (10%)	2 (2%)	6	35
1	f	91/126 (72%)	80 (88%)	10 (11%)	1 (1%)	14	51
2	A	95/135 (70%)	83 (87%)	9 (10%)	3 (3%)	4	26
2	E	95/135 (70%)	86 (90%)	9 (10%)	0	100	100

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	I	95/135 (70%)	83 (87%)	9 (10%)	3 (3%)	4	26
2	M	95/135 (70%)	86 (90%)	9 (10%)	0	100	100
2	Q	95/135 (70%)	83 (87%)	9 (10%)	3 (3%)	4	26
2	U	95/135 (70%)	86 (90%)	9 (10%)	0	100	100
2	Y	95/135 (70%)	83 (87%)	9 (10%)	3 (3%)	4	26
2	c	95/135 (70%)	86 (90%)	9 (10%)	0	100	100
3	B	81/102 (79%)	70 (86%)	8 (10%)	3 (4%)	3	24
3	F	76/102 (74%)	67 (88%)	9 (12%)	0	100	100
3	J	81/102 (79%)	70 (86%)	8 (10%)	3 (4%)	3	24
3	N	76/102 (74%)	67 (88%)	9 (12%)	0	100	100
3	R	78/102 (76%)	68 (87%)	7 (9%)	3 (4%)	3	24
3	V	76/102 (74%)	67 (88%)	9 (12%)	0	100	100
3	Z	81/102 (79%)	70 (86%)	8 (10%)	3 (4%)	3	24
3	d	76/102 (74%)	67 (88%)	9 (12%)	0	100	100
4	C	101/130 (78%)	88 (87%)	9 (9%)	4 (4%)	3	23
4	G	103/130 (79%)	88 (85%)	15 (15%)	0	100	100
4	K	101/130 (78%)	88 (87%)	9 (9%)	4 (4%)	3	23
4	O	103/130 (79%)	87 (84%)	16 (16%)	0	100	100
4	S	101/130 (78%)	88 (87%)	9 (9%)	4 (4%)	3	23
4	W	103/130 (79%)	88 (85%)	15 (15%)	0	100	100
4	a	101/130 (78%)	88 (87%)	9 (9%)	4 (4%)	3	23
4	e	103/130 (79%)	88 (85%)	15 (15%)	0	100	100
All	All	2937/3944 (74%)	2572 (88%)	313 (11%)	52 (2%)	8	40

5 of 52 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	L	101	GLY
2	Q	73	GLU
3	R	29	ILE
2	I	73	GLU
3	J	29	ILE

### 5.3.2 Protein sidechains ⓘ

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	D	81/106 (76%)	69 (85%)	12 (15%)	3	15
1	H	79/106 (74%)	67 (85%)	12 (15%)	3	14
1	L	81/106 (76%)	69 (85%)	12 (15%)	3	15
1	P	79/106 (74%)	67 (85%)	12 (15%)	3	14
1	T	81/106 (76%)	69 (85%)	12 (15%)	3	15
1	X	79/106 (74%)	67 (85%)	12 (15%)	3	14
1	b	81/106 (76%)	69 (85%)	12 (15%)	3	15
1	f	79/106 (74%)	67 (85%)	12 (15%)	3	14
2	A	85/110 (77%)	77 (91%)	8 (9%)	8	28
2	E	84/110 (76%)	79 (94%)	5 (6%)	19	44
2	I	85/110 (77%)	77 (91%)	8 (9%)	8	28
2	M	84/110 (76%)	79 (94%)	5 (6%)	19	44
2	Q	85/110 (77%)	77 (91%)	8 (9%)	8	28
2	U	84/110 (76%)	79 (94%)	5 (6%)	19	44
2	Y	85/110 (77%)	77 (91%)	8 (9%)	8	28
2	c	84/110 (76%)	79 (94%)	5 (6%)	19	44
3	B	68/78 (87%)	63 (93%)	5 (7%)	13	38
3	F	63/78 (81%)	59 (94%)	4 (6%)	18	42
3	J	68/78 (87%)	63 (93%)	5 (7%)	13	38
3	N	63/78 (81%)	59 (94%)	4 (6%)	18	42
3	R	67/78 (86%)	62 (92%)	5 (8%)	13	37
3	V	63/78 (81%)	59 (94%)	4 (6%)	18	42
3	Z	68/78 (87%)	63 (93%)	5 (7%)	13	38
3	d	63/78 (81%)	59 (94%)	4 (6%)	18	42
4	C	82/102 (80%)	75 (92%)	7 (8%)	10	33
4	G	83/102 (81%)	74 (89%)	9 (11%)	6	23

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
4	K	82/102 (80%)	75 (92%)	7 (8%)	10	33
4	O	83/102 (81%)	74 (89%)	9 (11%)	6	23
4	S	82/102 (80%)	75 (92%)	7 (8%)	10	33
4	W	83/102 (81%)	74 (89%)	9 (11%)	6	23
4	a	82/102 (80%)	75 (92%)	7 (8%)	10	33
4	e	83/102 (81%)	74 (89%)	9 (11%)	6	23
All	All	2499/3168 (79%)	2251 (90%)	248 (10%)	8	26

5 of 248 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
4	G	72	ASP
4	S	59	THR
4	e	59	THR
4	G	84	GLN
1	H	98	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 24 such sidechains are listed below:

Mol	Chain	Res	Type
4	G	84	GLN
4	S	38	ASN
4	e	31	HIS
4	K	38	ASN
4	K	112	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 carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 8 ligands modelled in this entry, 8 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 [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues [i](#)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
6	h	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	h	-448:DT	O3'	-447:DG	P	1.92



## 6 Fit of model and data ⓘ

### 6.1 Protein, DNA and RNA chains ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.3 Carbohydrates ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.4 Ligands ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.5 Other polymers ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.