

# wwPDB X-ray Structure Validation Summary Report (i)

#### Jan 8, 2024 – 02:17 pm GMT

PDB ID	:	50Y7
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 A user guide is available at https://www.wwpdb.org/validation/2017/XrayValidationReportHelp with specific help available everywhere you see the (i) symbol.

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

The following versions of software and data (see references (1)) were used in the production of this report:

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

# 1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:  $X\text{-}RAY \, DIFFRACTION$ 

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.



Motria	Whole archive	Similar resolution
wietric	$(\# {\rm Entries})$	$(\# { m Entries},  { m resolution}  { m range}({ m \AA}))$
R <sub>free</sub>	130704	1008 (7.70-3.86)
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 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%

Mol	Chain	Length	Quality of chain						
1	D	126	64% 10	)%	·	25%			
1	Н	126	62% 12	%		26%			
1	L	126	65%	9%	•	25%			
1	Р	126	63% 11	%		26%			
1	Т	126	65%	9%	•	25%			
1	Х	126	63% 11	%		26%			
1	b	126	65%	9%	•	25%			
1	f	126	63% 11	%		26%			



Mol	Chain	Length	Quality of chain		
2	А	135	64% 8%	-	28%
2	Е	135	· 68%		28%
2	Ι	135	64% 8%		28%
2	М	135	· 68% ·	-	28%
2	Q	135	64% 8%		28%
2	U	135	68% •		28%
2	Y	135	64% 8%		28%
2	с	135	68% •		28%
3	В	102	73%	9%	19%
3	F	102	73%	•	24%
3	J	102	73%	9%	19%
3	Ν	102	73%	•	24%
3	R	102	73%	8%	20%
3	V	102	73%	•	24%
3	Z	102	74%	8%	19%
3	d	102	73%	•	24%
4	С	130	72%	7% •	21%
4	G	130	74%	7%	19%
4	K	130	72%	7% •	21%
4	0	130	74%	7%	19%
4	S	130	72%	7% •	21%
4	W	130	74%	7%	19%
4	a	130	72%	7% •	21%
4	e	130	74%	7%	19%
5	g	634	73%		25% •
	Ь	0.01			



Mol	Chain	Length	Quality of chain		
6	h	628	74%	24%	



# 2 Entry composition (i)

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.

Mol	Chain	Residues		At	oms			ZeroOcc	AltConf	Trace		
1	т	05	Total	С	Ν	0	S	0	0	0		
		90	745	469	134	140	2	0	0	0		
1	D	03	Total	С	Ν	0	S	0	0	0		
1	1	90	726	457	130	137	2	0	0	0		
1	П	05	Total	С	Ν	0	S	0	0	0		
1	D	90	745	469	134	140	2	0	0	0		
1	п (	н	н	03	Total	С	Ν	0	S	0	0	0
1	11	95	726	457	130	137	2	0	0	0		
1	Т	T 95	Total	С	Ν	0	S	0	0	0		
1	L		745	469	134	140	2	0	0	0		
1	v	03	Total	С	Ν	0	S	0	0	0		
	Λ	90	726	457	130	137	2	0	0	0		
1	h	05	Total	С	Ν	0	S	0	0	0		
	D	90	745	469	134	140	2	0	0	0		
1	f	03	Total	С	Ν	Ο	S	0	0	0		
	1   f	93	726	457	130	137	2		0	U		

• Molecule 1 is a protein called Histone H2B 1.1.

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
L	29	THR	SER	conflict	UNP P02281
Р	29	THR	SER	conflict	UNP P02281
D	29	THR	SER	conflict	UNP P02281
Н	29	THR	SER	conflict	UNP P02281
Т	29	THR	SER	conflict	UNP P02281
Х	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		At	oms			ZeroOcc	AltConf	Trace
0	М	07	Total	С	Ν	0	S	0	0	0
	111	91	801	504	155	139	3	0	0	0
0	0	07	Total	С	Ν	0	S	0	0	0
	Q	91	802	506	155	138	3	0	0	0
2	Т	07	Total	С	Ν	0	S	0	0	0
	1	91	802	506	155	138	3	0	0	0
2	2 1	A 97	Total	С	Ν	0	S	0	0	0
	Л		802	506	155	138	3	0	0	0
2	F	E 97	Total	С	Ν	Ο	S	0	0	0
	Ľ		801	504	155	139	3		0	0
0	II	07	Total	С	Ν	0	S	0	0	0
	U	91	801	504	155	139	3	0	0	0
2	v	07	Total	С	Ν	Ο	$\mathbf{S}$	0	0	0
	1	91	802	506	155	138	3	0	0	0
2	0	07	Total	С	Ν	0	S	0	0	0
	C	97	801	504	155	139	3	0		0

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
М	102	ALA	GLY	conflict	UNP Q92133
М	111	ALA	GLY	conflict	UNP Q92133
Q	102	ALA	GLY	conflict	UNP Q92133
Q	111	ALA	GLY	conflict	UNP Q92133
Ι	102	ALA	GLY	conflict	UNP Q92133
Ι	111	ALA	GLY	conflict	UNP Q92133
А	102	ALA	GLY	conflict	UNP Q92133
А	111	ALA	GLY	conflict	UNP Q92133
Е	102	ALA	GLY	conflict	UNP Q92133
Е	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
с	102	ALA	GLY	conflict	UNP Q92133
с	111	ALA	GLY	conflict	UNP Q92133

• Molecule 3 is a protein called Histone H4.

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



Mol	Chain	Residues		At	oms			ZeroOcc	AltConf	Trace
2	D	80	Total	С	Ν	Ο	S	0	0	0
5	π	02	654	412	128	113	1	0	0	0
3	Т	83	Total	С	Ν	Ο	$\mathbf{S}$	0	0	0
5	J	00	662	418	129	114	1	0	0	0
3	В	83	Total	С	Ν	0	S	0	0 0	0
5	D	00	662	418	129	114	1	0		0
2	Б	78	Total	С	Ν	0	S	0	0	0
5	Г	10	619	391	120	107	1	0	0	0
2	V	78	Total	С	Ν	0	S	0	0	0
5	v	10	619	391	120	107	1	0	0	0
2	7	02	Total	С	Ν	0	S	0	0	0
5		00	662	418	129	114	1	0	0	0
3	d	78	Total	С	Ν	0	S	0	0	0
5	u	10	619	391	120	107	1			U

• Molecule 4 is a protein called Histone H2A.

Mol	Chain	Residues		Ato	ms		ZeroOcc	AltConf	Trace
4	0	105	Total	С	Ν	Ο	0	0	0
4	0	105	809	510	158	141	0	0	0
4	C	102	Total	С	Ν	Ο	0	0	0
4	U	105	795	501	155	139	0		
4	С	105	Total	С	Ν	Ο	0	0	0
4	G	105	809	510	158	141	0	0	0
4	K	102	Total	С	Ν	Ο	0	0 0	0
4	Γ	105	795	501	155	139	0		0
4	q	103	Total	С	Ν	Ο	0	0	0
4	U U	105	795	501	155	139	0	0	0
4	W	105	Total	С	Ν	Ο	0	0	0
4	vv	105	809	510	158	141	0	0	0
4	0	103	Total	С	Ν	Ο	0	0	0
4	a	105	795	501	155	139	0	0	0
4	0	105	Total	С	Ν	Ο	0	0	0
4	е	105	809	510	158	141		U	0

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

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace	
5	g	619	Total 12605	$\begin{array}{c} \mathrm{C} \\ 5987 \end{array}$	N 2278	O 3721	Р 619	0	0	0

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



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

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

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	О	1	Total Cl 1 1	1	0
7	С	1	Total Cl 1 1	1	0
7	G	1	Total Cl 1 1	1	0
7	К	1	Total Cl 1 1	1	0
7	S	1	Total Cl 1 1	1	0
7	W	1	Total Cl 1 1	1	0
7	a	1	Total Cl 1 1	1	0
7	е	1	Total Cl 1 1	1	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. 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.

Chain L:	65%	9% •	25%
MET PRO GLU PRO ALA PRO ALA ALA ALA ALA CVS CVS CVS CVS CVS CVS CVS CVS CVS CVS	LYS THR GLN LYS LYS ARG ARG K28 ARG K28 ARG CY3 CY3 CY3 CY3 CY3 CY3 CY3 CY3 CY3 CY3	149 L77 S84 T85 T93	L98 G101 T116 T119 K122
• Molecule 1: Histone H2B	1.1		
Chain P:	63%	11%	26%
MET PRO GLU GLU GLU PRO ALA ALA ALA PRO PRO CLYS SER LYS SER LYS SER LYS STA ALA ALA THR	LYS THR GLY GLY GLY GLY CLY CLY CLY CLY CLY CLY CLY CLY CLY C	100 000 000 000 000 000 000 000 000 000	L98 6101 8102 8109 8109 81106 81106 81112
A121 LYS			
• Molecule 1: Histone H2B	1.1		
Chain D:	64%	10% •	25%
MET PRO CLU PRO PRO ALA ALA PRO PRO PRO PRO CLYS CLYS CLYS CLYS CLYS CLYS CLYS THR	LYS GLN LLYS LLYS LLYS ARG ARG ARG ARG ARG ARG ARG ARG ARG ARG	149 L77 S84 T85 T93	L98 G101 S106 S106 T116 T116 T119 K122
• Molecule 1: Histone H2B	1.1		
Chain H:	62%	12%	26%
MET PRO CIUU CIUU PRO PRO PRO PRO PRO CIV SER CIV SER CIV SER CIV VAL ALA THR	LYS THR GLN CLVS CLVS CLVS ASP ASC CLVS CLVS CLVS CLVS CLVS CLVS CLVS CL	E68 R83 S88 T93 A94	V95 198 6101 6101 8100 8109 71112
K117 121 1718 1718			
• Molecule 1: Histone H2B	1.1		
Chain T:	65%	9% •	25%
MET PRO GLUU GLUU PRO PRO ALA ALA ALA ALA ALA CVS CLVS CLVS CLVS CLVS CLVS CLVS CLVS	LYS THR GLN CLYS CLYS CLYS ASP CLYS CLYS CLYS CLYS CLYS CLYS CLYS CLYS	149 L77 S84 T85 T93	198 6101 1116 1116 1119 1119 1119
		D E ANK	

• Molecule 1: Histone H2B 1.1

• Molecule 1: Histone H2B	1.1			
Chain X:	63%	11%	26%	
MET MET PRO GUU PRO PRO ALA ALA PRO PRO PRO LYS STR LYS STR LYS STR LYS STR VALA	THR THR CLN CLN CLN CLN CLN CLNS CLN CLNS CLN CLNS CLN CLNS CLN CLNS CLN CLNS CLN CLNS CLN CLNS CLN CLN CLN CLN CLN CLN CLN CLN CLN CLN	E68 R83 S88 S88 <b>A94</b> V95	L98 G101 E102 H106	S109 T112 K117
A121 LYS				
• Molecule 1: Histone H2B	1.1			
Chain b:	65%	9% •	25%	
MET PRO GLU GLU ALA PRO PRO PRO PRO PRO CLYS CLYS CLYS CLYS CLYS SER LYS SER LYS SER LYS SER THR	LITS THS CLN CLN CLN CLN CLN CLY CLY CLY CLY CLY CLY CLY CLY CLY CLN CLN CLN CLN CLN CLN CLN CLN CLN CLN	T49 L77 S84 T85 T93	L98 G101 S109 T116	T119 K122
• Molecule 1: Histone H2B	1.1			
Chain f:	63%	11%	26%	
MET PRIO GLU GLU ALLA ALLA ALLA ALLA ALLA CLYS CLYS CLYS CLYS CLYS CLYS CLYS CLYS	THR THR CLN CLN CLN CLN CLN CLN CLN CLN CLN CLN	E68 R83 R83 R83 R83 R83 R83 R83 R83 R94 R94 V95	L98 G101 E102 H106	S109 T112 K117
M21 LYS				
• Molecule 2: Histone H3				
Chain M:	68%		28%	1
ALA ARG THR LYS GLN THR ALA ARG CLY SER CLY GLY GLY GLY GLY GLY GLY GLY GLY CLYS CLYS CLYS CLYS CLYS CLN CLN CLN CLN CLN CLN CLN CLN CLN CLN	THR THR LYS ALA ALA ALA ARG SER SER SER ALA ALA ALA ALA ALA ALA CVY CVX CVY CVY	H39 E59 L65 V117 R129	A135	
• Molecule 2: Histone H3				
Chain Q:	64%	8%	28%	
ALA ARG THR LTR LTR LTR CLN ARG CLN SER ARG CLY GLY GLY GLY GLY GLY GLY GLY CLNS GLN	TALA TALA LYS ALA ALA ARG SER SER SER ALA ALA ALA CLY VAL VAL VAL VAL VAL VAL	L48 R49 E59 R63 K64 L65	E73 I74 D77 E105	K115 R129 R134
ALA				
• Molecule 2: Histone H3				
Chain I:	64%	8% 2	28%	
ALA ARG TTRR TTRR CLN CLN CLN CLN ARG CLN SER CLN CLN CLN CLN CLN CLN CLN CLN CLN CLN	TALR TALR TALR LYS ALA ALA ALA ALA ALA ALA ALA CLY GLY CLYS CLY CLYS CLY CLYS CLY CLYS CLY	L48 R49 E59 R63 K64 L65	E73 174 D77 E105	K115 R129 R134



# ALA

• Molecule 2: Histone H3

Chain A:	64%	8%	28%	
ALA ARG THR LYS GLN THR ALA ARG ARG	SER SER THR GLY CLYS CLYS CLYS ALA ALA ALA ALA ALA ALA ALA ALA ALA AL	L48 R49	E03 R63 K64 L65 E73 174 D77 E105	K115 R129 R134
ALA				
• Molecule	2: Histone H3			
Chain E:	68%	·	28%	
ALA ARG THR LYS GLN THR ALA ARG TXS	SER THR CLYS CLYS CLYS CLYS CLYS PRO CLYS CLN ALA ALA ALA ALA ALA ALA ALA ALA ALA A	H39 E59	117 117 117 1129	
• Molecule	2: Histone H3			
Chain U:	68%	•	28%	•
ALA ARG THR LYS GLN THR ALA ARG 1.YS	SER THR GLY CLYS GLY PRO CLYS CLYS CLYS CLY ALA ALA ALA ALA ALA ALA ALA ALA ALA CLY VAL CLYS CLY VAL CLYS CLY PRO CLY VAL CLYS CLY PRO CLYP PRO CLAP PRO CLAP PRO CLAP PRO CLAP PRO CLAP PRO CLAP PRO CLAP PRO CLAP PRO CLAP PRO CLAP PRO CLAP PRO CLAP PRO CLAP PRO CLAP PRO CLAPRO CLAP PRO CLAPR	H39 E59	V117 N129 A135	
• Molecule :	2: Histone H3			
Chain Y:	64%	8%	28%	•
Chain Y:	SER THR GLY GLY CLYS GLY PR.O A.L.A	148 R49 %8	28% <b>K83</b> <b>1</b> <b>1</b> <b>1</b> <b>1</b> <b>1</b> <b>1</b> <b>1</b> <b>1</b>	K115 R129 R134
Chain Y: SHELL SALE SHELL SA	SER SER GLY GLY GLY GLY ARG GLN ARG ARA ARA ARA ARA ARA ARA ARA ARA ARA	8%	28%	K115 R129 R134
Chain Y:	64% 5 5 4 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5	148 149 149 149 149 149 149 149 149 149 149	28%	K1 15 R1 29 R1 29
Chain Y:	64% 51 55 H 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	148 149 149 149 149 149 149 149 149 149 149	28%	K115 R129 R134
Chain Y:	64% 54% 54% 25 Histone H3 68% 68%	H139 148 148 148 148 148 148 148 148	28%	K115 R129 R134
Chain Y:	64% 56% 2: Histone H3 68% 3: Histone H4	H139 - 148 - 1	28%	K1 15 R1 29 R1 34
Chain Y:	64% 2: Histone H3 3: Histone H4 73%		28%	11 2 12 13 14 13 14 15 15 14 15 15 15 15 15 15 15 15 15 15 15 15 15
Chain Y:	64% 2: Histone H3 68% 3: Histone H4 73%		28%	11 15 11 20 11 20 11 24 11 24



Chain R:	73%	8%	20%
SER ARG ARG CLY CLY CLY CLY CLY CLY CLY CLY CLY CLY	HIS HIS KZ0 KZ0 126 126 134 135 R35 R35 R35 R35 R35 R35 R35 R35 R35 R	K91 R92 G102	
• Molecule 3: Histone H	14		
Chain J:	73%	9%	19%
SER GLY GLY GLY CLY CLY CLY CLY CLY CLY GLY CLY CLY CLY CLY CLY CLY CLY CLY CLY C	AHIS AHIS K20 K20 126 134 847 847 847 173 847 173 847 173 847 173 847	61 02 1 02	
• Molecule 3: Histone H	14		
Chain B:	73%	9%	19%
SER GLY ARG CLY CLY CLY CLY CLY CLY CLY CLY CLY CLY	HIS HIS K20 K20 126 134 134 134 134 134 134 134 134 134 134	G102	
• Molecule 3: Histone I	14		
Chain F:	73%	·	24%
SER GLY ARG CLY CLY CLY CLY CLY CLY CLY CLY CLY CLY	HIS HARG LYRS VAL VAL VAL ARG ASP ASP ASP ASP ASP ASP ASP ASP ASP ASP		
• Molecule 3: Histone I	44		
Chain V:	73%	·	24%
SER GLY ARG CLY CLY CLY CLY CLY CLY CLY CLY CLY CLY	HIS LYR VAL VAL LEU ARC ASP ASP ASP T73 T73 T73 T73 T73 C102		
• Molecule 3: Histone H	14		
Chain Z:	74%	8%	19%
SER GLY CLY CLY CLY CLY CLY CLY CLY CLY CLY C	MIS MIS N20 N20 126 134 134 173 173 173 173 180 180 180 180 180	6102	
• Molecule 3: Histone I	14		
Chain d:	73%	·	24%
SER ARG ALY ARG CLY LYS LYS CLY CLY CLY CLY CLY CLY CLY ALA ARG	H1S LYRG VAL VAL VAL LUU ARG ASP 125 125 125 125 125 125 125 125 125 125		
• Molecule 4: Histone H			
Chain O:	74%	7%	19%



MET SER GLY GLY CLY CLY CLY ALA ALA ALA ALA ALA ALA ALA ALA ALA A	D72 N73 K74 R81 R81 Q84 R88 R88	V114 K118 K118 LV75 LV75 CLU SER SER LV5 SER ALL ALL SER LV5 SER	
• Molecule 4: Histone H2A			
Chain C:	72%	7% •	21%
MET SER GLY GLY GLY CLYS LYS CLY ALA ALA ALA ALA ALA ALA ALA ALA ALA A	T59 E64 R81 E91 T101	S113 V114 V114 LYS LYS SER CVU SER SER ALA LYS SER LYS SER LYS	
• Molecule 4: Histone H2A			
Chain G:	74%	7%	19%
MET SER SER ALY GLY CLY CLY ARG CLY ARG ALA ARG ALA ARG ALA ARG ALA ALA ALA ALA ALA ALA ALA ALA ALA AL	D72 N73 K74 R81 Q84 Q84 R88 R88	V114 K118 LY18 LY18 LY18 SER SER SER LY5 SER LY5	
• Molecule 4: Histone H2A			
Chain K:	72%	7% •	21%
MET SER GLY GLY GLY CLYS LYS CLY ALA ALA ALA ALA ALA ALA ALA ALA ALA A	T59 E64 R81 E91 T101	S113 V114 K118 LVS LVS CLV SER SER SER LVS SER LVS SER LVS SER LVS	
• Moleculo 4: Histore H2A			
• Molecule 4. Instolle 112A			
Chain S:	72%	7% •	21%
Molecule 4. Instone HZA     Chain S:	72%	S113 VI 4 LYS THR LYS SER SER SER ALA SER LYS SER LYS SER LYS	21%
<ul> <li>Molecule 4: Histone H2A</li> <li>Chain S:</li> <li>Molecule 4: Histone H2A</li> </ul>	72%	S113 VI14 LTS LTS LTS RT SER SER SER SER LTS SER LTS SER LTS	21%
<ul> <li>Molecule 4: Histone H2A</li> <li>Chain S:</li> <li>Molecule 4: Histone H2A</li> <li>Chain W:</li> </ul>	72%	24% • 24%	21%
<ul> <li>Molecule 4: Histone H2A</li> <li>Chain S:</li> <li>Molecule 4: Histone H2A</li> <li>Chain W:</li> </ul>	72%	VI14 KII8 LIV8 SER THR THR THR THR THR SER SER SER SER SER SER SER SER SER SE	21%
<ul> <li>Molecule 4: Histone H2A</li> <li>Chain S:</li> <li>Molecule 4: Histone H2A</li> <li>Chain W:</li> <li>E E E E E E E E E E E E E E E E E E E</li></ul>	72%	VI14 KHB IVI IVI IVI IVI IVI IVI SER IVI SE IVI SI SI SI SI SI SI SI SI SI SI SI SI SI	21%
<ul> <li>Molecule 4: Histone H2A</li> <li>Chain S:</li> <li>Molecule 4: Histone H2A</li> <li>Chain W:</li> <li>Molecule 4: Histone H2A</li> <li>Molecule 4: Histone H2A</li> <li>Chain a:</li> </ul>	72%	7% .	21%
<ul> <li>Molecule 4: Histone H2A</li> <li>Chain S:</li> <li>Molecule 4: Histone H2A</li> <li>Chain W:</li> <li>Molecule 4: Histone H2A</li> <li>Chain W:</li> <li>Molecule 4: Histone H2A</li> <li>Chain a:</li> </ul>	72%	2113       2113         1118       2113         1118       2114         1118	21%
<ul> <li>Molecule 4: Histone H2A</li> <li>Chain S:</li> <li>Molecule 4: Histone H2A</li> <li>Chain W:</li> <li>Molecule 4: Histone H2A</li> <li>Molecule 4: Histone H2A</li> <li>Chain a:</li> <li>Molecule 4: Histone H2A</li> <li>Molecule 4: Histone H2A</li> <li>Molecule 4: Histone H2A</li> </ul>	72% 691 692 74% 74% 74% 72%	7% • 113 114 114 114 114 114 114 114	21%
<ul> <li>Molecule 4: Histone H2A</li> <li>Chain S:</li> <li>Molecule 4: Histone H2A</li> <li>Chain W:</li> <li>Molecule 4: Histone H2A</li> <li>Chain a:</li> </ul>	72% 69 69 74% 74%	1000       10000       1000	21%



#### 

• Molecule 5: DNA (619-MER)



A 113 C 104 C 104 C 104 C 104 C 104 C 104 C 106 C

# 4 Data and refinement statistics (i)

Property	Value	Source	
Space group	P 1	Depositor	
Cell constants	66.62Å 161.47Å 227.91Å	Dopositor	
a, b, c, $\alpha$ , $\beta$ , $\gamma$	$78.94^{\circ}$ $83.86^{\circ}$ $83.74^{\circ}$	Depositor	
Bosolution(Å)	111.41 - 5.77	Depositor	
Resolution (A)	141.52 - 5.77	EDS	
% Data completeness	93.8(111.41-5.77)	Depositor	
(in resolution range)	93.8(141.52-5.77)	EDS	
$R_{merge}$	0.11	Depositor	
$R_{sym}$	(Not available)	Depositor	
$< I/\sigma(I) > 1$	$1.31 (at 5.77 \text{\AA})$	Xtriage	
Refinement program	PHENIX (1.11.1_2575: ???)	Depositor	
B B.	0.219 , $0.238$	Depositor	
$\Lambda, \Lambda_{free}$	0.289 , $0.306$	DCC	
$R_{free}$ test set	1226 reflections $(5.04%)$	wwPDB-VP	
Wilson B-factor $(Å^2)$	222.9	Xtriage	
Anisotropy	0.199	Xtriage	
Bulk solvent $k_{sol}(e/Å^3)$ , $B_{sol}(Å^2)$	0.23, $230.1$	EDS	
L-test for $twinning^2$	$ < L >=0.40, < L^2>=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 $(Å^2)$	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>&</sup>lt;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.



<sup>&</sup>lt;sup>1</sup>Intensities estimated from amplitudes.

# 5 Model quality (i)

### 5.1 Standard geometry (i)

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

Mal	Chain	Bond lengths		E	Bond angles		
	Unain	RMSZ	# Z  > 5	RMSZ	# Z  > 5		
1	D	0.49	0/756	0.64	0/1015		
1	Н	0.62	0/737	0.70	0/993		
1	L	0.49	0/756	0.64	0/1015		
1	Р	0.62	0/737	0.70	0/993		
1	Т	0.49	0/756	0.64	0/1015		
1	Х	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	А	0.66	0/814	0.73	0/1092		
2	Е	0.41	0/812	0.57	0/1088		
2	Ι	0.66	0/814	0.73	0/1092		
2	М	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	с	0.41	0/812	0.58	0/1088		
3	В	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	Ζ	0.67	0/669	0.83	0/894		
3	d	0.46	0/626	0.61	0/837		
4	С	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	0	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	е	0.63	0/819	0.76	0/1106		



Mal	Chain	Bo	ond lengths	Bond angles		
IVIOI	Unain	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)	$\operatorname{Ideal}(\operatorname{\AA})$
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	$Observed(^{o})$	$Ideal(^{o})$
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)

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

### 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	Perce	entiles
1	D	93/126~(74%)	82~(88%)	9~(10%)	2(2%)	6	35
1	Н	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	Р	91/126~(72%)	80 (88%)	10 (11%)	1 (1%)	14	51
1	Т	93/126~(74%)	81 (87%)	10 (11%)	2 (2%)	6	35
1	Х	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	А	95/135~(70%)	83 (87%)	9 (10%)	3 (3%)	4	26
2	Е	95/135~(70%)	86 (90%)	9 (10%)	0	100	100
2	Ι	95/135~(70%)	83 (87%)	9 (10%)	3 (3%)	4	26
2	М	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	с	95/135~(70%)	86 (90%)	9 (10%)	0	100	100
3	В	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	Ν	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	С	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	Ο	$\overline{103/130}~(79\%)$	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



Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
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	Ι	73	GLU
3	J	29	ILE

#### 5.3.2 Protein sidechains (i)

In the following table, the Percentiles column shows the percent side chain 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 side chain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Outliers Per		entiles
1	D	81/106~(76%)	69~(85%)	12 (15%)		3	15
1	Н	79/106~(74%)	67~(85%)	12~(15%)		3	14
1	L	81/106~(76%)	69~(85%)	12 (15%)		3	15
1	Р	79/106~(74%)	67~(85%)	12~(15%)		3	14
1	Т	81/106~(76%)	69 (85%)	12 (15%)		3	15
1	Х	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	А	85/110~(77%)	77 (91%)	8 (9%)		8	28
2	Е	84/110 (76%)	79 (94%)	5 (6%)		19	44
2	Ι	85/110 (77%)	77 (91%)	8 (9%)		8	28
2	М	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	с	84/110 (76%)	79 (94%)	5 (6%)		19	44



Mol	Chain	Analysed	Rotameric	Outliers	Outliers Percen	
3	В	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	Ν	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	Ζ	68/78~(87%)	63~(93%)	5 (7%)	13	38
3	d	63/78~(81%)	59~(94%)	4 (6%)	18	42
4	$\mathbf{C}$	82/102~(80%)	75~(92%)	7~(8%)	10	33
4	G	83/102~(81%)	74~(89%)	9 (11%)	6	23
4	Κ	82/102~(80%)	75~(92%)	7~(8%)	10	33
4	Ο	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	е	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	$\mathbf{Res}$	Type
4	G	84	GLN
2	с	117	VAL
4	S	118	LYS
2	с	59	GLU
1	f	31	LYS

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

Mol	Chain	Res	Type
4	S	112	GLN
4	W	84	GLN
4	W	31	HIS
3	Ζ	75	HIS
4	С	38	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 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:

$\mathbf{Mol}$	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	Р	1.92



## 6 Fit of model and data (i)

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

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

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

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

### 6.3 Carbohydrates (i)

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

### 6.4 Ligands (i)

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

#### 6.5 Other polymers (i)

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

