



# wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 23, 2016 – 01:45 AM EDT

PDB ID : 4ZUX  
Title : SAGA DUB module Ubp8/Sgf11/Sus1/Sgf73 bound to ubiquitinated nucleosome  
Authors : Morgan, M.; Wolberger, C.  
Deposited on : 2015-05-17  
Resolution : 3.82 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  
<http://wwpdb.org/validation/2016/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  
Mogul : unknown  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20027107  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Refmac : 5.8.0122  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : rb-20027107

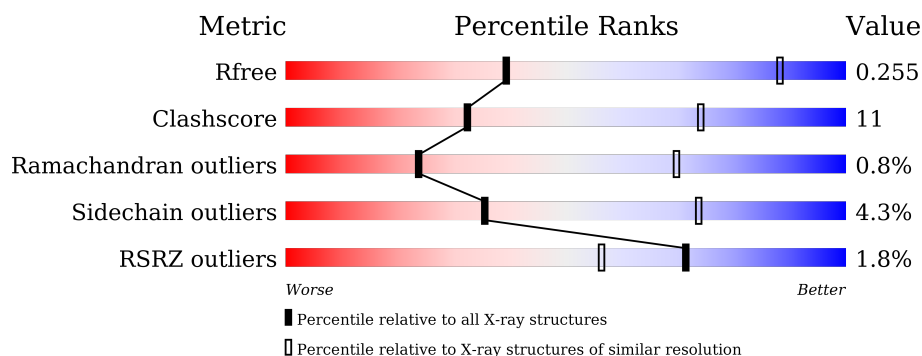
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.82 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))
$R_{free}$	91344	1324 (4.14-3.50)
Clashscore	102246	1028 (4.12-3.52)
Ramachandran outliers	100387	1404 (4.14-3.50)
Sidechain outliers	100360	1399 (4.14-3.50)
RSRZ outliers	91569	1332 (4.14-3.50)

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. 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	136	<div> <div>%</div> <div> <div></div> <div>46%</div> <div>24%</div> <div>•</div> <div>29%</div> </div> </div>
1	E	136	<div> <div>%</div> <div> <div></div> <div>46%</div> <div>26%</div> <div></div> <div>29%</div> </div> </div>
1	K	136	<div> <div></div> <div> <div></div> <div>50%</div> <div>21%</div> <div>•</div> <div>29%</div> </div> </div>
1	O	136	<div> <div>%</div> <div> <div></div> <div>46%</div> <div>25%</div> <div>•</div> <div>29%</div> </div> </div>
2	B	103	<div> <div>%</div> <div> <div></div> <div>51%</div> <div>24%</div> <div>5%</div> <div>19%</div> </div> </div>
2	F	103	<div> <div>2%</div> <div> <div></div> <div>57%</div> <div>18%</div> <div></div> <div>24%</div> </div> </div>

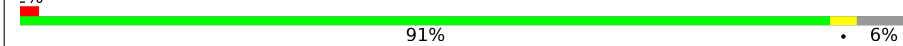



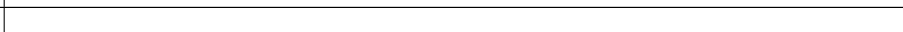
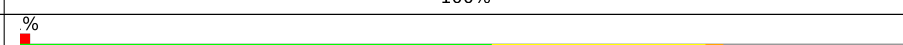
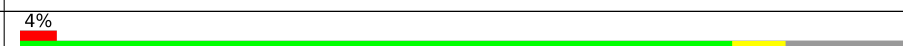
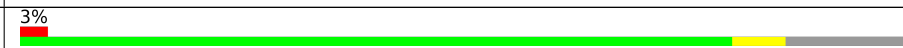
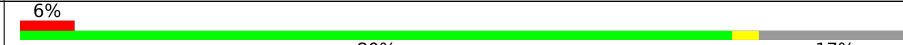
Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	L	103	
2	P	103	
3	C	130	
3	G	130	
3	M	130	
3	Q	130	
4	D	123	
4	H	123	
4	N	123	
4	R	123	
5	I	145	
5	S	145	
6	J	145	
6	T	145	
7	U	472	
7	Z	472	
7	e	472	
7	j	472	
8	V	96	
8	a	96	
8	f	96	
8	k	96	
9	W	99	
9	b	99	
9	g	99	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
9	l	99	 91% 2% 6%
10	X	76	 82% 17%
10	c	76	 99% 3%
10	h	76	 99% 7%
10	m	76	 100%
11	Y	104	 53% 24% 21% 1%
11	d	104	 80% 6% 14% 4%
11	i	104	 80% 6% 14% 3%
11	n	104	 80% 17% 6%

## 2 Entry composition [i](#)

There are 12 unique types of molecules in this entry. The entry contains 49079 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called Histone H3.2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	97	Total	C	N	O	S	0	0	0
			802	506	155	138	3			
1	E	97	Total	C	N	O	S	0	0	0
			801	504	155	139	3			
1	K	97	Total	C	N	O	S	0	0	0
			802	506	155	138	3			
1	O	97	Total	C	N	O	S	0	0	0
			801	504	155	139	3			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	102	ALA	GLY	conflict	UNP P84233
E	102	ALA	GLY	conflict	UNP P84233
K	102	ALA	GLY	conflict	UNP P84233
O	102	ALA	GLY	conflict	UNP P84233

- Molecule 2 is a protein called Histone H4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	83	Total	C	N	O	S	0	0	0
			662	418	129	114	1			
2	F	78	Total	C	N	O	S	0	0	0
			619	391	120	107	1			
2	L	83	Total	C	N	O	S	0	0	0
			662	418	129	114	1			
2	P	78	Total	C	N	O	S	0	0	0
			619	391	120	107	1			

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

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	C	103	Total	C	N	O	0	0	0
			795	501	155	139			
3	G	105	Total	C	N	O	0	0	0
			809	510	158	141			
3	M	103	Total	C	N	O	0	0	0
			795	501	155	139			
3	Q	105	Total	C	N	O	0	0	0
			809	510	158	141			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	99	ARG	GLY	conflict	UNP P06897
C	123	SER	ALA	conflict	UNP P06897
G	99	ARG	GLY	conflict	UNP P06897
G	123	SER	ALA	conflict	UNP P06897
M	99	ARG	GLY	conflict	UNP P06897
M	123	SER	ALA	conflict	UNP P06897
Q	99	ARG	GLY	conflict	UNP P06897
Q	123	SER	ALA	conflict	UNP P06897

- Molecule 4 is a protein called Histone H2B 1.1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	95	Total	C	N	O	S	0	0	0
			745	469	134	140	2			
4	H	93	Total	C	N	O	S	0	0	0
			726	457	130	137	2			
4	N	95	Total	C	N	O	S	0	0	0
			745	469	134	140	2			
4	R	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
D	0	MET	-	initiating methionine	UNP P02281
D	29	THR	SER	conflict	UNP P02281
H	0	MET	-	initiating methionine	UNP P02281
H	29	THR	SER	conflict	UNP P02281
N	0	MET	-	initiating methionine	UNP P02281
N	29	THR	SER	conflict	UNP P02281
R	0	MET	-	initiating methionine	UNP P02281

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
R	29	THR	SER	conflict	UNP P02281

- Molecule 5 is a DNA chain called DNA (145-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	I	145	Total	C	N	O	P	0	0	0
			2952	1404	537	867	144			
5	S	145	Total	C	N	O	P	0	0	0
			2952	1404	537	867	144			

- Molecule 6 is a DNA chain called DNA (145-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	J	145	Total	C	N	O	P	0	0	0
			2987	1416	558	869	144			
6	T	145	Total	C	N	O	P	0	0	0
			2987	1416	558	869	144			

- Molecule 7 is a protein called Ubiquitin carboxyl-terminal hydrolase 8.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	U	447	Total	C	N	O	S	0	0	0
			3569	2264	609	661	35			
7	Z	451	Total	C	N	O	S	0	0	0
			3600	2283	615	667	35			
7	e	453	Total	C	N	O	S	0	0	0
			3617	2292	618	672	35			
7	j	447	Total	C	N	O	S	0	0	0
			3566	2262	609	660	35			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
U	0	ALA	-	expression tag	UNP P50102
U	146	ALA	CYS	engineered mutation	UNP P50102
Z	0	ALA	-	expression tag	UNP P50102
Z	146	ALA	CYS	engineered mutation	UNP P50102
e	0	ALA	-	expression tag	UNP P50102
e	146	ALA	CYS	engineered mutation	UNP P50102
j	0	ALA	-	expression tag	UNP P50102
j	146	ALA	CYS	engineered mutation	UNP P50102

- Molecule 8 is a protein called Transcription and mRNA export factor SUS1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
8	V	89	Total	C	N	O	S	0	0	0
			719	451	118	148	2			
8	a	90	Total	C	N	O	S	0	0	0
			730	457	119	152	2			
8	f	85	Total	C	N	O	S	0	0	0
			690	432	112	144	2			
8	k	86	Total	C	N	O	S	0	0	0
			695	435	115	143	2			

- Molecule 9 is a protein called SAGA-associated factor 11.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
9	W	90	Total	C	N	O	S	0	0	0
			718	440	133	142	3			
9	b	89	Total	C	N	O	S	0	0	0
			710	434	132	141	3			
9	g	90	Total	C	N	O	S	0	0	0
			718	440	133	142	3			
9	l	93	Total	C	N	O	S	0	0	0
			739	452	136	148	3			

- Molecule 10 is a protein called Polyubiquitin-B.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
10	X	76	Total	C	N	O	S	0	0	0
			601	378	105	117	1			
10	c	76	Total	C	N	O	S	0	0	0
			601	378	105	117	1			
10	h	76	Total	C	N	O	S	0	0	0
			601	378	105	117	1			
10	m	76	Total	C	N	O	S	0	0	0
			601	378	105	117	1			

- Molecule 11 is a protein called SAGA-associated factor 73.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
11	Y	82	Total	C	N	O	S	0	0	0
			663	419	112	127	5			
11	d	89	Total	C	N	O	S	0	0	0
			710	447	120	138	5			
11	i	89	Total	C	N	O	S	0	0	0
			710	447	120	138	5			

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
11	n	86	Total	C	N	O	S	0	0	0
			693	437	117	134	5			

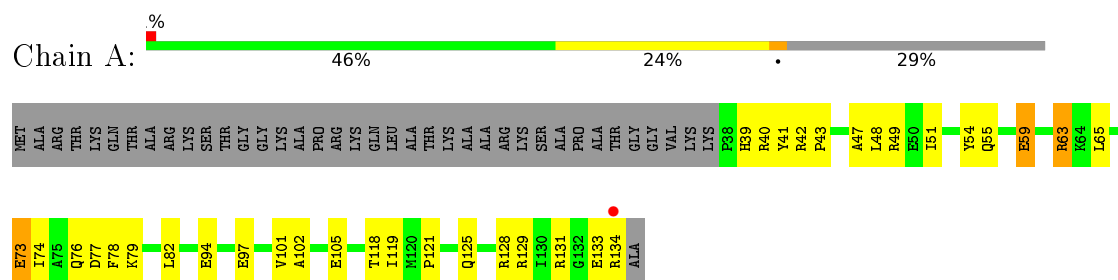
- Molecule 12 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
12	g	1	Total	Zn	0	0
			1	1		
12	j	6	Total	Zn	0	0
			6	6		
12	d	1	Total	Zn	0	0
			1	1		
12	e	6	Total	Zn	0	0
			6	6		
12	b	1	Total	Zn	0	0
			1	1		
12	i	1	Total	Zn	0	0
			1	1		
12	W	1	Total	Zn	0	0
			1	1		
12	Z	6	Total	Zn	0	0
			6	6		
12	n	1	Total	Zn	0	0
			1	1		
12	U	6	Total	Zn	0	0
			6	6		
12	Y	1	Total	Zn	0	0
			1	1		
12	l	1	Total	Zn	0	0
			1	1		

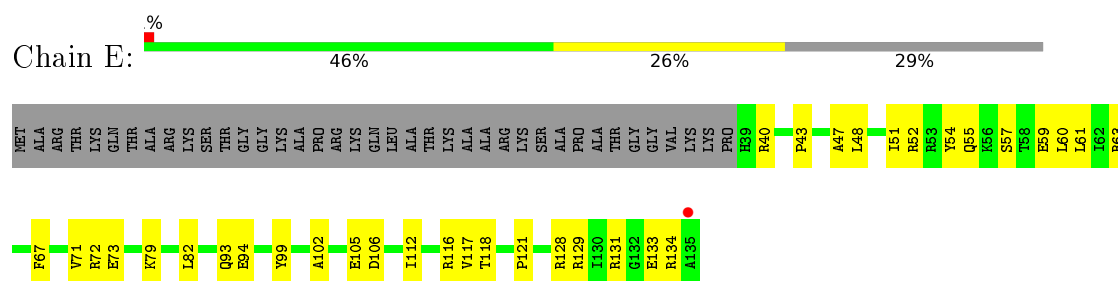
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density ( $RSRZ > 2$ ). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

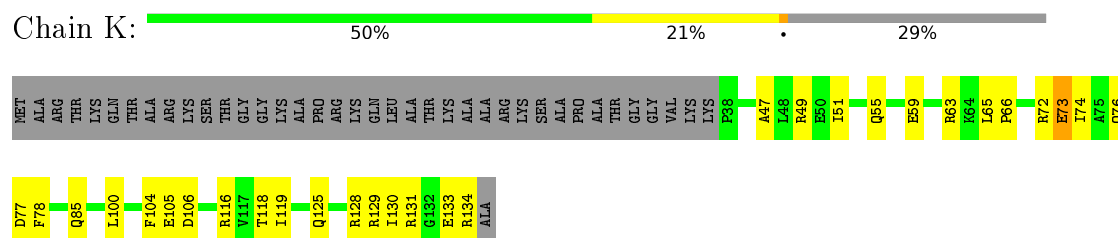
#### • Molecule 1: Histone H3.2



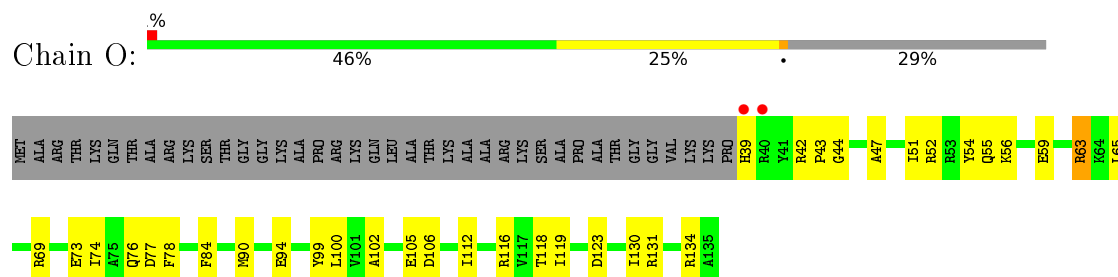
#### • Molecule 1: Histone H3.2



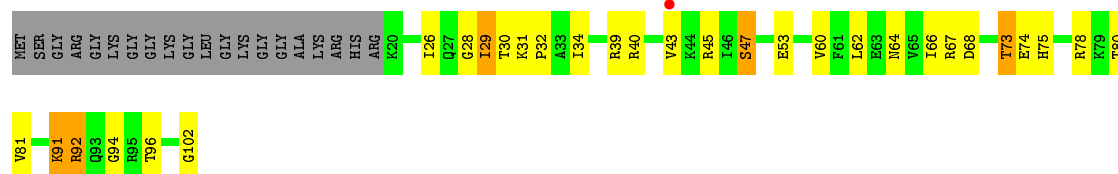
#### • Molecule 1: Histone H3.2



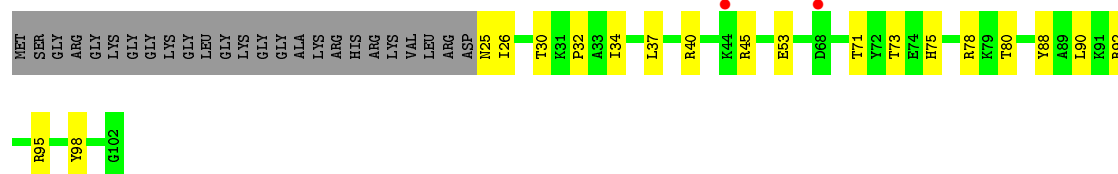
#### • Molecule 1: Histone H3.2



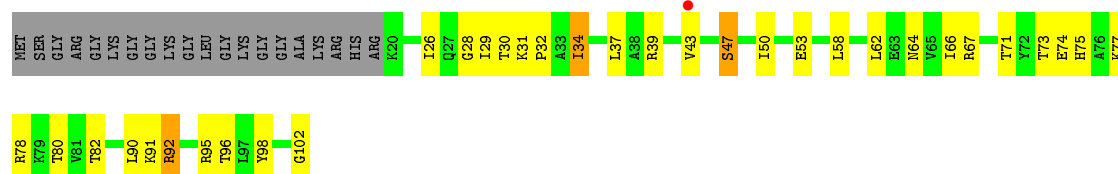
- Molecule 2: Histone H4



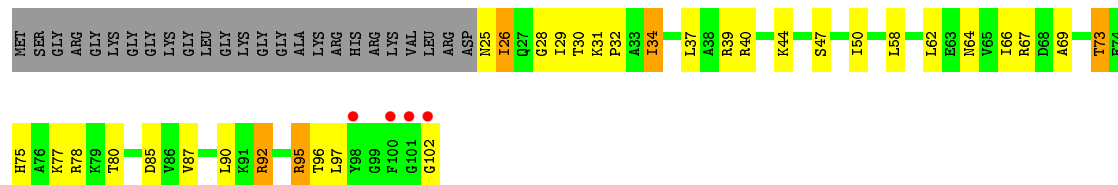
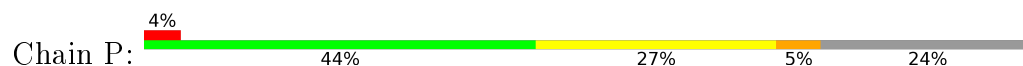
- Molecule 2: Histone H4



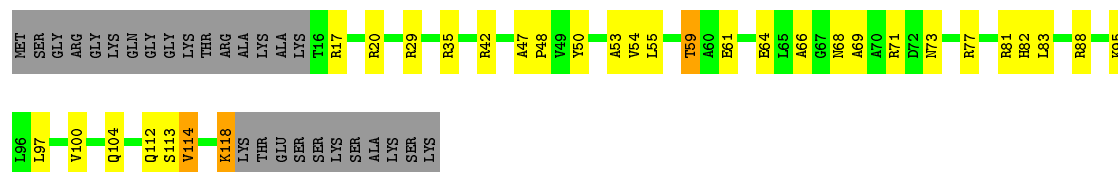
- Molecule 2: Histone H4



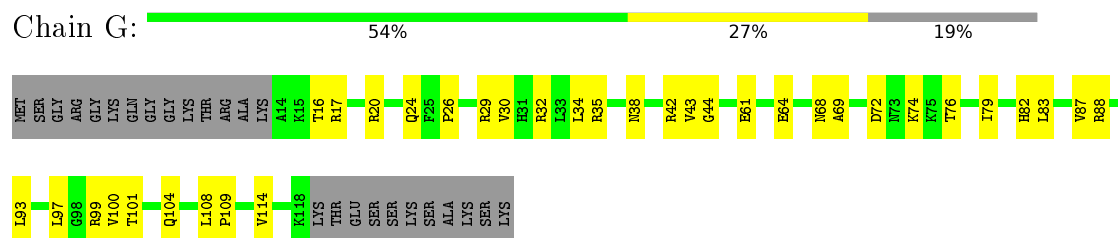
- Molecule 2: Histone H4



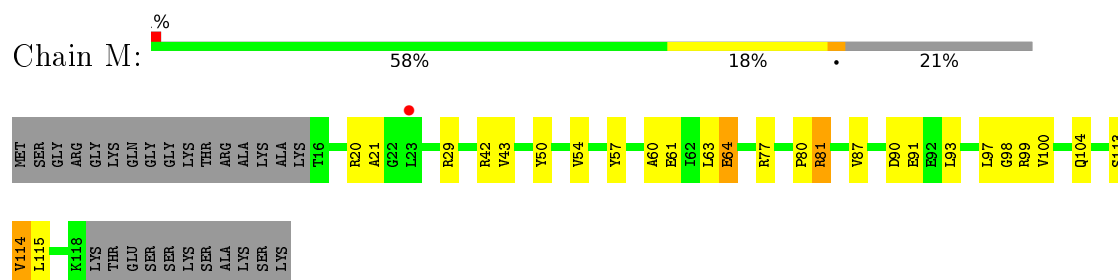
- Molecule 3: Histone H2A type 1



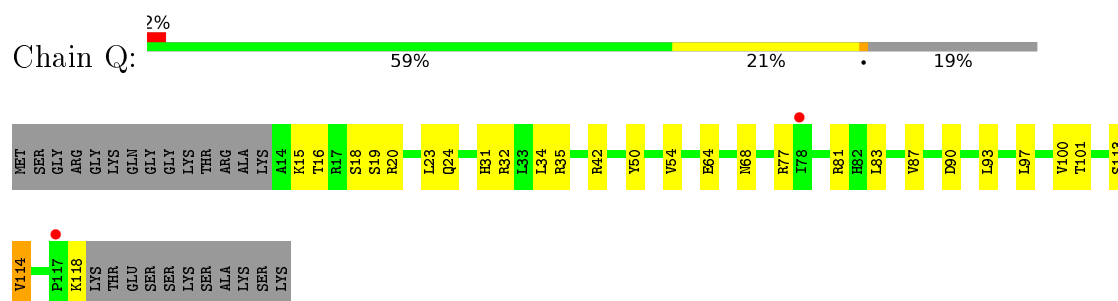
- Molecule 3: Histone H2A type 1



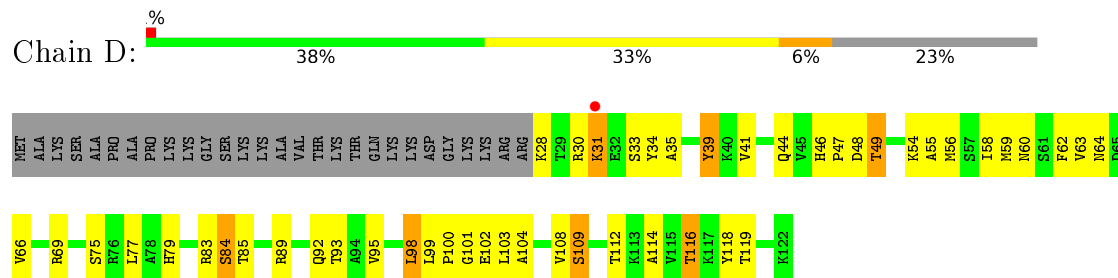
- Molecule 3: Histone H2A type 1



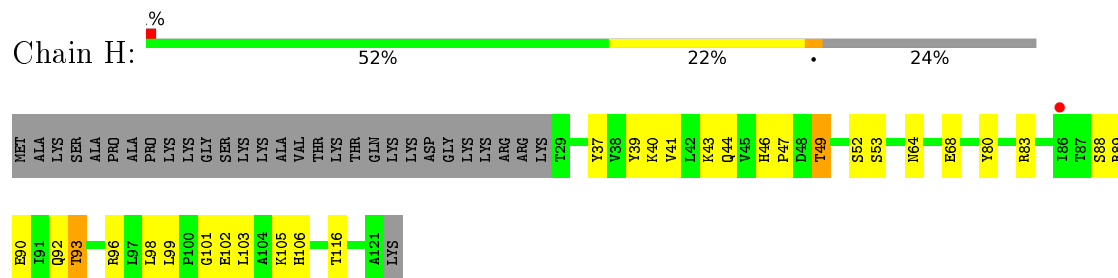
- Molecule 3: Histone H2A type 1



- Molecule 4: Histone H2B 1.1

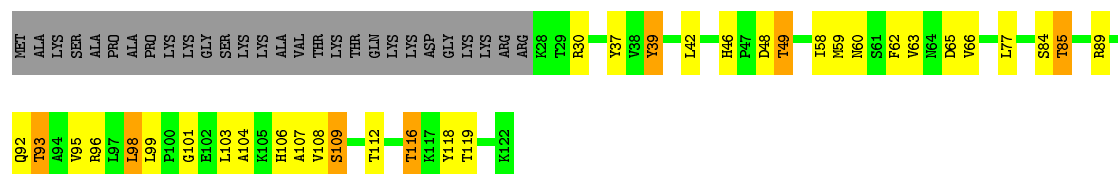


- Molecule 4: Histone H2B 1.1



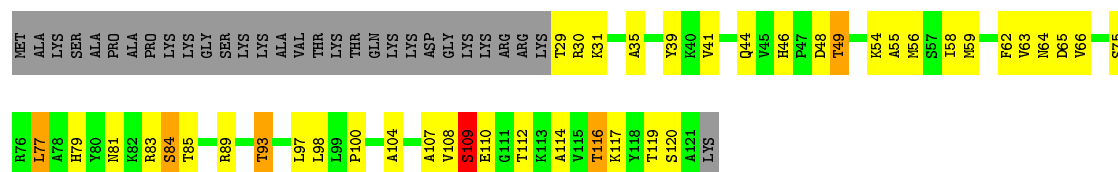
- Molecule 4: Histone H2B 1.1

Chain N: 



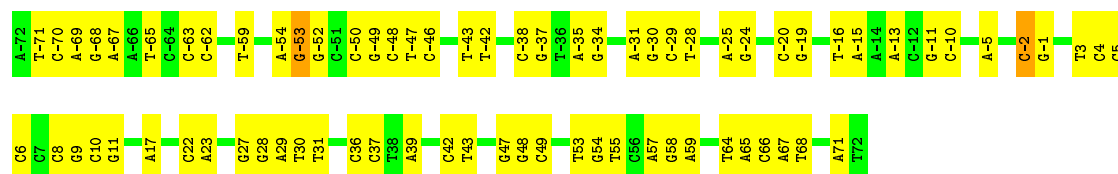
• Molecule 4: Histone H2B 1.1

Chain R: 



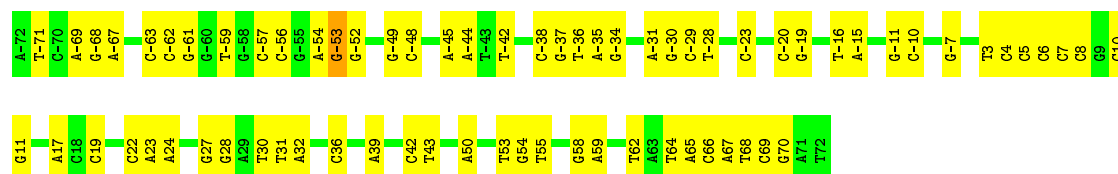
• Molecule 5: DNA (145-MER)

Chain I: 



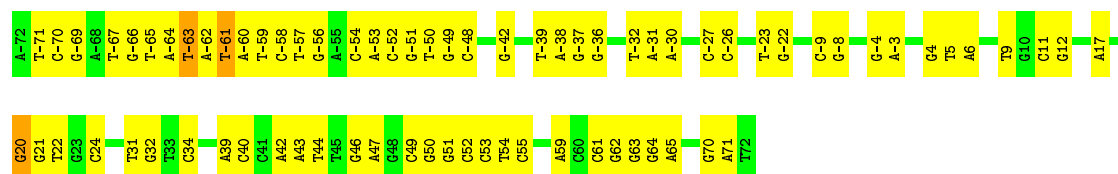
• Molecule 5: DNA (145-MER)

Chain S: 



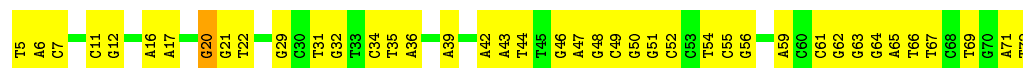
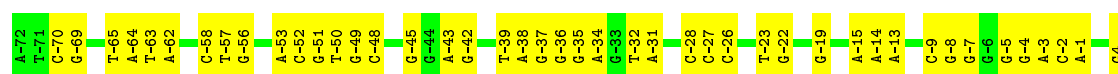
• Molecule 6: DNA (145-MER)

Chain J: 

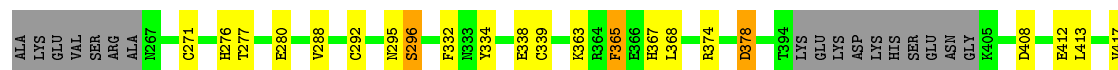
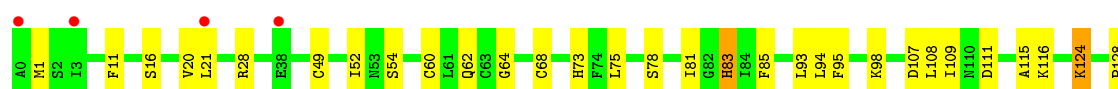
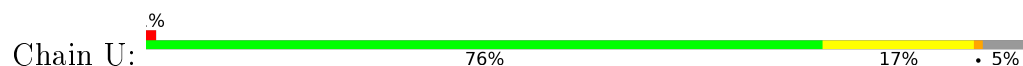


• Molecule 6: DNA (145-MER)

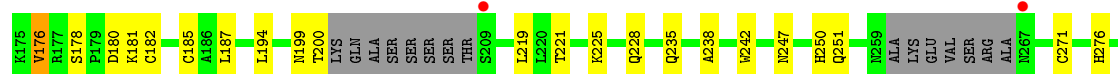
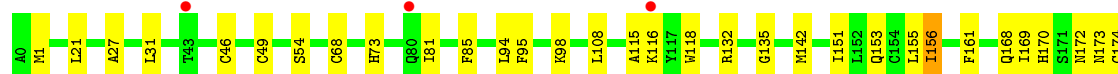
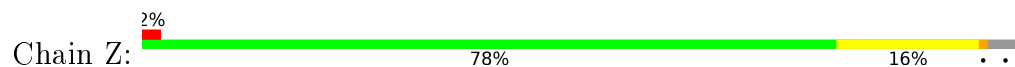
Chain T: 



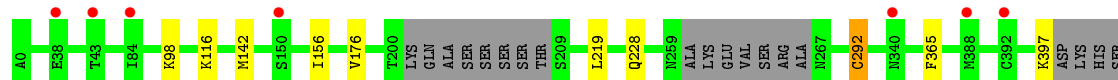
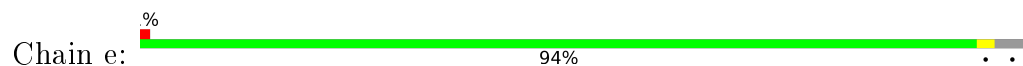
• Molecule 7: Ubiquitin carboxyl-terminal hydrolase 8



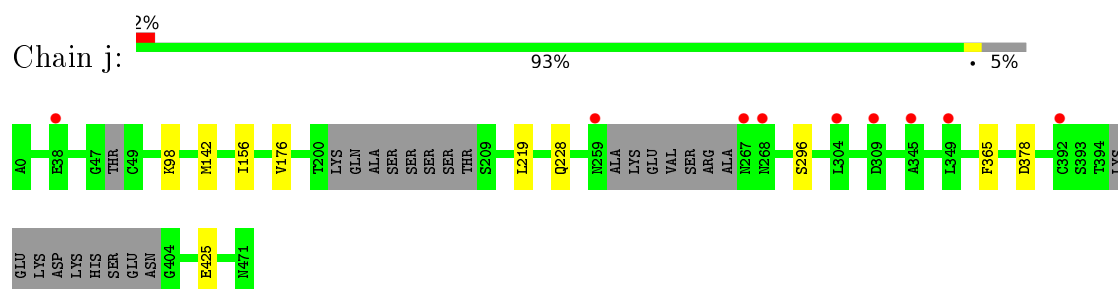
• Molecule 7: Ubiquitin carboxyl-terminal hydrolase 8



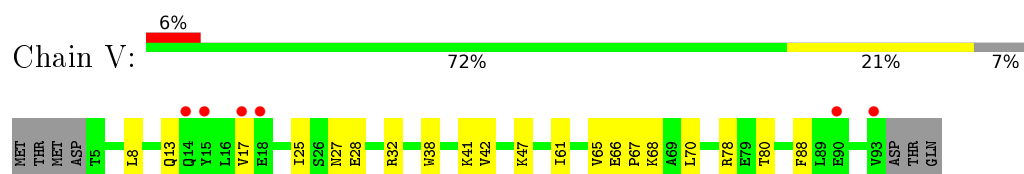
• Molecule 7: Ubiquitin carboxyl-terminal hydrolase 8



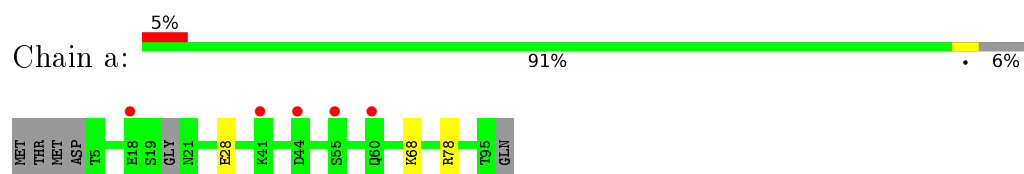
• Molecule 7: Ubiquitin carboxyl-terminal hydrolase 8



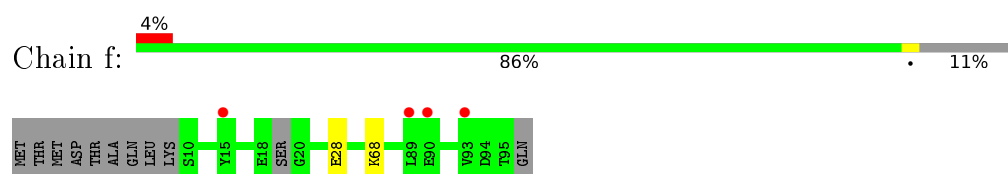
- Molecule 8: Transcription and mRNA export factor SUS1



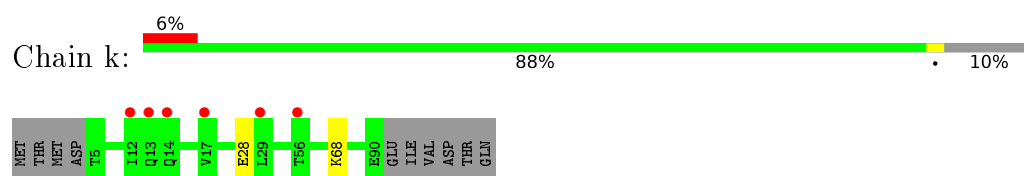
- Molecule 8: Transcription and mRNA export factor SUS1



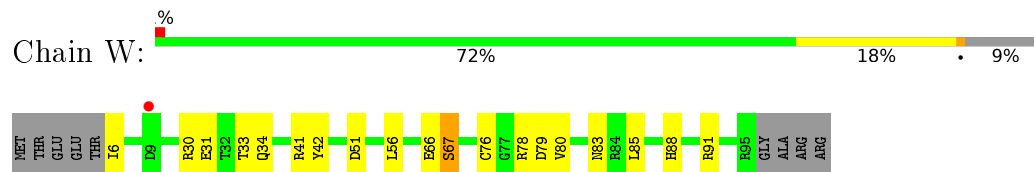
- Molecule 8: Transcription and mRNA export factor SUS1



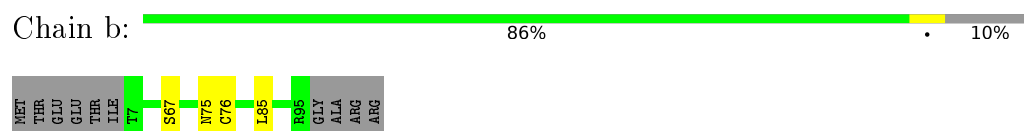
- Molecule 8: Transcription and mRNA export factor SUS1



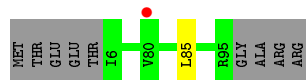
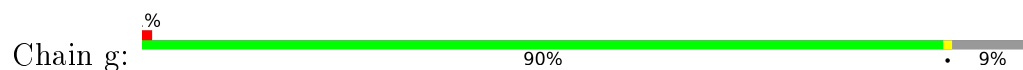
- Molecule 9: SAGA-associated factor 11



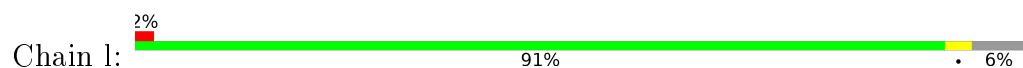
- Molecule 9: SAGA-associated factor 11



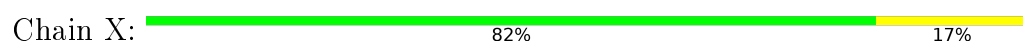
- Molecule 9: SAGA-associated factor 11



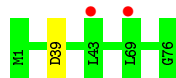
- Molecule 9: SAGA-associated factor 11



- Molecule 10: Polyubiquitin-B



- Molecule 10: Polyubiquitin-B



- Molecule 10: Polyubiquitin-B

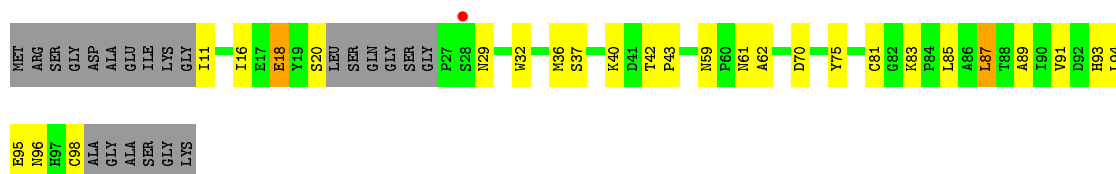


- Molecule 10: Polyubiquitin-B

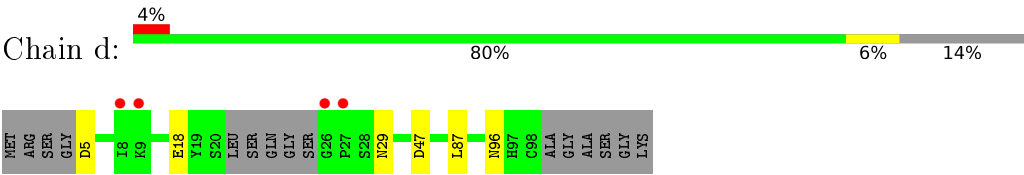


There are no outlier residues recorded for this chain.

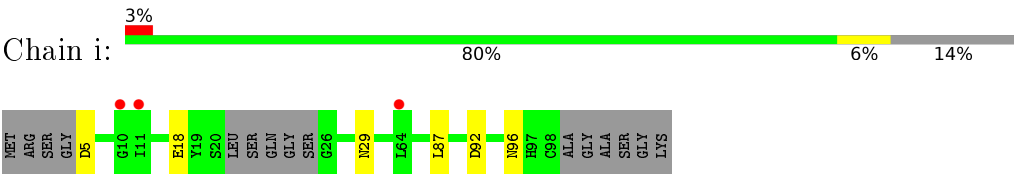
- Molecule 11: SAGA-associated factor 73



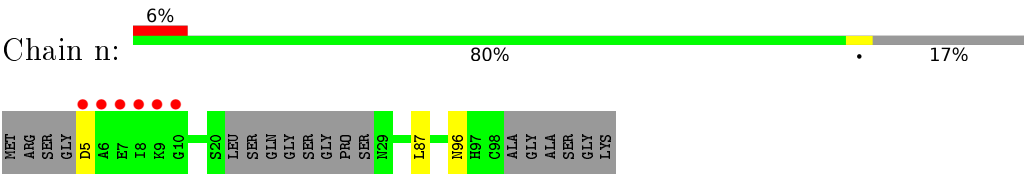
● Molecule 11: SAGA-associated factor 73



● Molecule 11: SAGA-associated factor 73



● Molecule 11: SAGA-associated factor 73



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	178.80Å 179.17Å 353.82Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.50 – 3.82 49.50 – 3.69	Depositor EDS
% Data completeness (in resolution range)	98.9 (49.50-3.82) 99.3 (49.50-3.69)	Depositor EDS
$R_{merge}$	0.11	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.47 (at 3.67Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.9_1692)	Depositor
R, $R_{free}$	0.234 , 0.256 0.230 , 0.255	Depositor DCC
$R_{free}$ test set	1790 reflections (1.62%)	DCC
Wilson B-factor (Å <sup>2</sup> )	168.1	Xtriage
Anisotropy	0.052	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.27 , 118.4	EDS
Estimated twinning fraction	0.228 for k,h,-l	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.37$ , $\langle L^2 \rangle = 0.20$	Xtriage
Outliers	0 of 121413 reflections	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	49079	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	197.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 10.69% 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, 0.375 respectively for untwinned datasets, and 0.333, 0.2 for perfectly twinned datasets.

## 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:  
ZN

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	0.27	0/814	0.50	0/1092
1	E	0.24	0/812	0.47	0/1088
1	K	0.26	0/814	0.50	0/1092
1	O	0.27	0/812	0.55	0/1088
2	B	0.28	0/669	0.59	0/894
2	F	0.32	0/626	0.55	0/837
2	L	0.30	0/669	0.57	0/894
2	P	0.28	0/626	0.59	0/837
3	C	0.25	0/805	0.53	0/1088
3	G	0.26	0/819	0.55	0/1106
3	M	0.25	0/805	0.57	1/1088 (0.1%)
3	Q	0.25	0/819	0.54	0/1106
4	D	0.28	0/756	0.56	0/1015
4	H	0.26	0/737	0.55	0/993
4	N	0.26	0/756	0.54	0/1015
4	R	0.29	0/737	0.63	1/993 (0.1%)
5	I	0.63	0/3308	1.05	2/5099 (0.0%)
5	S	0.64	0/3308	1.05	2/5099 (0.0%)
6	J	0.62	0/3354	1.05	5/5180 (0.1%)
6	T	0.62	0/3354	1.06	3/5180 (0.1%)
7	U	0.26	0/3647	0.49	0/4916
7	Z	0.25	0/3678	0.49	0/4955
7	e	0.25	0/3695	0.50	0/4978
7	j	0.25	0/3643	0.50	0/4908
8	V	0.23	0/725	0.46	0/977
8	a	0.24	0/735	0.48	0/990
8	f	0.29	0/695	0.57	0/936
8	k	0.25	0/701	0.44	0/944
9	W	0.27	0/726	0.52	0/982
9	b	0.26	0/718	0.52	0/971
9	g	0.30	0/726	0.55	0/982
9	l	0.27	0/747	0.54	0/1011

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# Z  >5	RMSZ	# Z  >5
10	X	0.22	0/607	0.47	0/816
10	c	0.22	0/607	0.44	0/816
10	h	0.23	0/607	0.48	0/816
10	m	0.22	0/607	0.45	0/816
11	Y	0.25	0/678	0.61	0/915
11	d	0.25	0/725	0.56	0/978
11	i	0.26	0/725	0.58	0/978
11	n	0.25	0/707	0.55	0/953
All	All	0.39	0/51099	0.72	14/71422 (0.0%)

There are no bond length outliers.

The worst 5 of 14 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	M	81	ARG	NE-CZ-NH1	-7.82	116.39	120.30
4	R	83	ARG	NE-CZ-NH1	-6.16	117.22	120.30
5	I	-53	DG	OP2-P-O3'	6.14	118.71	105.20
5	I	-2	DC	O4'-C1'-N1	5.95	112.17	108.00
5	S	-53	DG	OP2-P-O3'	5.87	118.11	105.20

There are no chirality outliers.

There are no planarity outliers.

## 5.2 Too-close contacts ⓘ

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	802	0	841	24	0
1	E	801	0	838	24	0
1	K	802	0	841	19	0
1	O	801	0	838	27	0
2	B	662	0	709	25	0
2	F	619	0	659	18	0
2	L	662	0	709	26	0
2	P	619	0	659	28	0
3	C	795	0	846	25	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	G	809	0	864	27	0
3	M	795	0	846	24	0
3	Q	809	0	864	25	0
4	D	745	0	773	35	0
4	H	726	0	747	26	0
4	N	745	0	773	29	0
4	R	726	0	747	29	0
5	I	2952	0	1629	81	0
5	S	2952	0	1629	88	0
6	J	2987	0	1630	84	0
6	T	2987	0	1630	97	0
7	U	3569	0	3476	63	0
7	Z	3600	0	3514	48	0
7	e	3617	0	3524	0	0
7	j	3566	0	3472	0	0
8	V	719	0	730	12	0
8	a	730	0	737	0	0
8	f	690	0	691	0	0
8	k	695	0	704	0	0
9	W	718	0	710	18	0
9	b	710	0	699	0	0
9	g	718	0	709	0	0
9	l	739	0	725	0	0
10	X	601	0	629	11	0
10	c	601	0	629	0	0
10	h	601	0	629	0	0
10	m	601	0	629	0	0
11	Y	663	0	653	21	0
11	d	710	0	697	0	0
11	i	710	0	697	0	0
11	n	693	0	682	0	0
12	U	6	0	0	0	0
12	W	1	0	0	0	0
12	Y	1	0	0	0	0
12	Z	6	0	0	0	0
12	b	1	0	0	0	0
12	d	1	0	0	0	0
12	e	6	0	0	0	0
12	g	1	0	0	0	0
12	i	1	0	0	0	0
12	j	6	0	0	0	0
12	l	1	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
12	n	1	0	0	0	0
All	All	49079	0	44008	690	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 11.

The worst 5 of 690 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)
7:Z:170:HIS:CE1	7:Z:182:CYS:SG	2.64	0.85
5:S:19:DC:H42	6:T:-19:DG:H1	1.21	0.84
5:S:70:DG:N2	6:T:-70:DC:O2	2.13	0.82
3:C:20:ARG:NH2	5:I:-42:DT:OP1	2.11	0.82
7:Z:46:CYS:SG	7:Z:73:HIS:ND1	2.54	0.78

There are no symmetry-related clashes.

## 5.3 Torsion angles [i](#)

### 5.3.1 Protein backbone [i](#)

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

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	95/136 (70%)	84 (88%)	8 (8%)	3 (3%)	5	44
1	E	95/136 (70%)	86 (90%)	9 (10%)	0	100	100
1	K	95/136 (70%)	84 (88%)	8 (8%)	3 (3%)	5	44
1	O	95/136 (70%)	84 (88%)	8 (8%)	3 (3%)	5	44
2	B	81/103 (79%)	71 (88%)	8 (10%)	2 (2%)	7	49
2	F	76/103 (74%)	69 (91%)	7 (9%)	0	100	100
2	L	81/103 (79%)	71 (88%)	7 (9%)	3 (4%)	4	40
2	P	76/103 (74%)	65 (86%)	8 (10%)	3 (4%)	4	38

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
3	C	101/130 (78%)	89 (88%)	9 (9%)	3 (3%)	5	45
3	G	103/130 (79%)	91 (88%)	12 (12%)	0	100	100
3	M	101/130 (78%)	90 (89%)	8 (8%)	3 (3%)	5	45
3	Q	103/130 (79%)	91 (88%)	9 (9%)	3 (3%)	6	46
4	D	93/123 (76%)	84 (90%)	7 (8%)	2 (2%)	8	52
4	H	91/123 (74%)	81 (89%)	9 (10%)	1 (1%)	17	65
4	N	93/123 (76%)	83 (89%)	8 (9%)	2 (2%)	8	52
4	R	91/123 (74%)	82 (90%)	7 (8%)	2 (2%)	8	52
7	U	439/472 (93%)	430 (98%)	9 (2%)	0	100	100
7	Z	443/472 (94%)	434 (98%)	9 (2%)	0	100	100
7	e	445/472 (94%)	435 (98%)	9 (2%)	1 (0%)	52	86
7	j	437/472 (93%)	426 (98%)	11 (2%)	0	100	100
8	V	87/96 (91%)	87 (100%)	0	0	100	100
8	a	86/96 (90%)	86 (100%)	0	0	100	100
8	f	81/96 (84%)	81 (100%)	0	0	100	100
8	k	84/96 (88%)	84 (100%)	0	0	100	100
9	W	88/99 (89%)	86 (98%)	2 (2%)	0	100	100
9	b	87/99 (88%)	84 (97%)	2 (2%)	1 (1%)	17	65
9	g	88/99 (89%)	86 (98%)	2 (2%)	0	100	100
9	l	91/99 (92%)	89 (98%)	2 (2%)	0	100	100
10	X	74/76 (97%)	73 (99%)	1 (1%)	0	100	100
10	c	74/76 (97%)	73 (99%)	1 (1%)	0	100	100
10	h	74/76 (97%)	72 (97%)	2 (3%)	0	100	100
10	m	74/76 (97%)	73 (99%)	1 (1%)	0	100	100
11	Y	78/104 (75%)	75 (96%)	3 (4%)	0	100	100
11	d	85/104 (82%)	80 (94%)	4 (5%)	1 (1%)	16	63
11	i	85/104 (82%)	82 (96%)	3 (4%)	0	100	100
11	n	82/104 (79%)	79 (96%)	3 (4%)	0	100	100
All	All	4552/5356 (85%)	4320 (95%)	196 (4%)	36 (1%)	24	70

5 of 36 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	73	GLU
2	B	29	ILE
4	D	101	GLY
1	K	73	GLU
2	L	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	A	85/111 (77%)	78 (92%)	7 (8%)	14	53
1	E	84/111 (76%)	82 (98%)	2 (2%)	57	84
1	K	85/111 (77%)	79 (93%)	6 (7%)	18	59
1	O	84/111 (76%)	77 (92%)	7 (8%)	14	52
2	B	68/79 (86%)	64 (94%)	4 (6%)	24	66
2	F	63/79 (80%)	60 (95%)	3 (5%)	31	71
2	L	68/79 (86%)	64 (94%)	4 (6%)	24	66
2	P	63/79 (80%)	57 (90%)	6 (10%)	11	46
3	C	82/102 (80%)	77 (94%)	5 (6%)	23	64
3	G	83/102 (81%)	79 (95%)	4 (5%)	31	71
3	M	82/102 (80%)	78 (95%)	4 (5%)	31	70
3	Q	83/102 (81%)	80 (96%)	3 (4%)	42	77
4	D	81/103 (79%)	70 (86%)	11 (14%)	5	30
4	H	79/103 (77%)	75 (95%)	4 (5%)	29	69
4	N	81/103 (79%)	70 (86%)	11 (14%)	5	30
4	R	79/103 (77%)	69 (87%)	10 (13%)	5	32
7	U	406/428 (95%)	391 (96%)	15 (4%)	41	76
7	Z	409/428 (96%)	398 (97%)	11 (3%)	52	82
7	e	411/428 (96%)	401 (98%)	10 (2%)	57	84
7	j	405/428 (95%)	395 (98%)	10 (2%)	55	83

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
8	V	84/91 (92%)	82 (98%)	2 (2%)	57	84
8	a	86/91 (94%)	83 (96%)	3 (4%)	43	78
8	f	81/91 (89%)	79 (98%)	2 (2%)	55	83
8	k	81/91 (89%)	79 (98%)	2 (2%)	55	83
9	W	82/89 (92%)	80 (98%)	2 (2%)	57	84
9	b	81/89 (91%)	78 (96%)	3 (4%)	41	76
9	g	82/89 (92%)	81 (99%)	1 (1%)	78	91
9	l	84/89 (94%)	81 (96%)	3 (4%)	42	77
10	X	68/68 (100%)	67 (98%)	1 (2%)	72	90
10	c	68/68 (100%)	67 (98%)	1 (2%)	72	90
10	h	68/68 (100%)	67 (98%)	1 (2%)	72	90
10	m	68/68 (100%)	68 (100%)	0	100	100
11	Y	77/90 (86%)	73 (95%)	4 (5%)	29	69
11	d	81/90 (90%)	76 (94%)	5 (6%)	23	64
11	i	81/90 (90%)	75 (93%)	6 (7%)	17	57
11	n	79/90 (88%)	76 (96%)	3 (4%)	40	76
All	All	4132/4644 (89%)	3956 (96%)	176 (4%)	35	74

5 of 176 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	Q	114	VAL
7	U	176	VAL
7	j	219	LEU
4	R	39	TYR
4	R	116	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 16 such sidechains are listed below:

Mol	Chain	Res	Type
7	U	83	HIS
7	U	90	ASN
9	b	75	ASN
7	U	53	ASN
7	e	348	GLN

### 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 carbohydrates in this entry.

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

Of 32 ligands modelled in this entry, 32 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](#)

There are no chain breaks in this entry.

## 6 Fit of model and data ⓘ

### 6.1 Protein, DNA and RNA chains ⓘ

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95<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	97/136 (71%)	-0.10	1 (1%) 84 72	134, 166, 198, 212	0
1	E	97/136 (71%)	-0.15	1 (1%) 84 72	139, 172, 213, 249	0
1	K	97/136 (71%)	-0.15	0 100 100	146, 175, 209, 216	0
1	O	97/136 (71%)	-0.15	2 (2%) 67 51	141, 177, 219, 249	0
2	B	83/103 (80%)	-0.18	1 (1%) 81 67	124, 155, 185, 198	0
2	F	78/103 (75%)	-0.05	2 (2%) 59 43	124, 166, 189, 201	0
2	L	83/103 (80%)	-0.11	1 (1%) 81 67	131, 152, 192, 204	0
2	P	78/103 (75%)	0.13	4 (5%) 32 21	153, 183, 219, 230	0
3	C	103/130 (79%)	-0.12	0 100 100	128, 162, 187, 197	0
3	G	105/130 (80%)	-0.01	0 100 100	148, 177, 198, 207	0
3	M	103/130 (79%)	-0.09	1 (0%) 84 72	143, 168, 195, 216	0
3	Q	105/130 (80%)	-0.06	2 (1%) 70 54	141, 179, 205, 222	0
4	D	95/123 (77%)	-0.19	1 (1%) 82 69	125, 155, 201, 243	0
4	H	93/123 (75%)	-0.22	1 (1%) 82 69	145, 172, 197, 209	0
4	N	95/123 (77%)	-0.18	0 100 100	137, 165, 190, 239	0
4	R	93/123 (75%)	-0.23	0 100 100	149, 172, 196, 207	0
5	I	145/145 (100%)	-0.85	0 100 100	176, 210, 241, 255	0
5	S	145/145 (100%)	-0.88	0 100 100	177, 208, 241, 257	0
6	J	145/145 (100%)	-0.87	0 100 100	169, 211, 243, 269	0
6	T	145/145 (100%)	-0.87	0 100 100	177, 211, 243, 264	0
7	U	447/472 (94%)	-0.12	6 (1%) 79 65	133, 183, 235, 285	0
7	Z	451/472 (95%)	-0.00	10 (2%) 65 50	161, 205, 261, 302	0
7	e	453/472 (95%)	-0.14	7 (1%) 76 62	148, 192, 244, 276	0
7	j	447/472 (94%)	-0.17	9 (2%) 68 53	146, 189, 232, 286	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
8	V	89/96 (92%)	-0.10	6 (6%) 21 12	160, 196, 260, 275	0
8	a	90/96 (93%)	0.09	5 (5%) 28 18	181, 238, 267, 283	0
8	f	85/96 (88%)	0.25	4 (4%) 35 23	196, 227, 267, 298	0
8	k	86/96 (89%)	-0.05	6 (6%) 19 11	188, 217, 249, 277	1 (1%)
9	W	90/99 (90%)	-0.31	1 (1%) 82 69	163, 196, 217, 231	0
9	b	89/99 (89%)	-0.24	0 100 100	173, 219, 252, 260	0
9	g	90/99 (90%)	-0.27	1 (1%) 82 69	170, 212, 266, 282	1 (1%)
9	l	93/99 (93%)	-0.19	2 (2%) 65 50	168, 201, 263, 275	0
10	X	76/76 (100%)	-0.07	0 100 100	157, 193, 221, 237	0
10	c	76/76 (100%)	-0.04	2 (2%) 59 43	159, 192, 218, 252	0
10	h	76/76 (100%)	0.07	5 (6%) 22 12	170, 195, 214, 239	0
10	m	76/76 (100%)	-0.15	0 100 100	159, 188, 218, 233	0
11	Y	82/104 (78%)	-0.01	1 (1%) 81 67	159, 193, 257, 301	0
11	d	89/104 (85%)	0.11	4 (4%) 37 24	174, 212, 279, 315	0
11	i	89/104 (85%)	0.03	3 (3%) 49 34	162, 200, 266, 297	0
11	n	86/104 (82%)	0.47	6 (6%) 19 11	167, 197, 281, 349	0
All	All	5242/5936 (88%)	-0.17	95 (1%) 71 56	124, 192, 248, 349	2 (0%)

The worst 5 of 95 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
11	n	6	ALA	22.3
11	n	5	ASP	10.3
11	d	26	GLY	10.3
11	n	7	GLU	9.7
7	Z	397	LYS	6.7

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

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

## 6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 6.4 Ligands ⓘ

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
12	ZN	Z	501	1/1	0.85	0.23	1.96	234,234,234,234	0
12	ZN	j	506	1/1	0.95	0.20	1.91	192,192,192,192	0
12	ZN	n	201	1/1	0.84	0.20	1.10	216,216,216,216	0
12	ZN	U	501	1/1	0.71	0.16	0.91	228,228,228,228	0
12	ZN	e	505	1/1	0.93	0.21	0.89	152,152,152,152	0
12	ZN	j	503	1/1	0.89	0.18	0.48	179,179,179,179	0
12	ZN	d	201	1/1	0.85	0.20	0.12	226,226,226,226	0
12	ZN	l	101	1/1	0.89	0.19	0.08	173,173,173,173	0
12	ZN	g	101	1/1	0.97	0.16	0.00	206,206,206,206	0
12	ZN	W	101	1/1	0.76	0.13	-0.01	199,199,199,199	0
12	ZN	j	502	1/1	0.99	0.15	-0.10	160,160,160,160	0
12	ZN	e	501	1/1	0.83	0.14	-0.34	251,251,251,251	0
12	ZN	i	201	1/1	0.98	0.16	-0.44	195,195,195,195	0
12	ZN	U	505	1/1	0.99	0.18	-0.45	160,160,160,160	0
12	ZN	U	502	1/1	0.98	0.17	-0.48	195,195,195,195	0
12	ZN	Z	502	1/1	0.58	0.20	-0.50	233,233,233,233	0
12	ZN	U	506	1/1	0.73	0.13	-0.57	252,252,252,252	0
12	ZN	Y	201	1/1	0.89	0.18	-0.59	184,184,184,184	0
12	ZN	e	504	1/1	0.99	0.14	-0.60	188,188,188,188	0
12	ZN	U	503	1/1	0.95	0.18	-0.61	209,209,209,209	0
12	ZN	j	504	1/1	0.84	0.16	-0.61	198,198,198,198	0
12	ZN	Z	505	1/1	0.83	0.17	-0.64	272,272,272,272	0
12	ZN	j	501	1/1	0.62	0.11	-0.65	286,286,286,286	0
12	ZN	j	505	1/1	0.96	0.12	-0.71	212,212,212,212	0
12	ZN	Z	504	1/1	0.93	0.12	-0.74	215,215,215,215	0
12	ZN	Z	506	1/1	0.91	0.14	-0.75	290,290,290,290	0
12	ZN	U	504	1/1	0.97	0.12	-0.89	201,201,201,201	0
12	ZN	e	503	1/1	0.70	0.13	-1.15	229,229,229,229	0
12	ZN	Z	503	1/1	0.94	0.10	-1.21	237,237,237,237	0
12	ZN	e	502	1/1	0.83	0.10	-1.21	279,279,279,279	0
12	ZN	b	101	1/1	0.86	0.12	-1.57	218,218,218,218	0
12	ZN	e	506	1/1	0.67	0.04	-1.90	251,251,251,251	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.