Full wwPDB X-ray Structure Validation Report ⓘ

Aug 28, 2020 – 01:38 PM BST

PDB ID	:	4ZUX
Title	:	SAGA DUB module Ubp8/Sgf11/Sus1/Sgf73 bound to ubiqitinated nucleosome
Authors	:	Morgan, M.; Wolberger, C.
Deposited on	:	2015-05-17
Resolution	:	3.82 Å(reported)

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

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the ⓘ symbol.

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

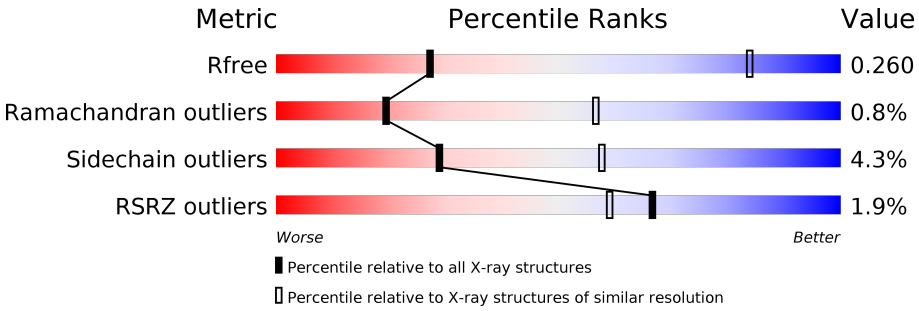
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.13
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.13

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 3.82 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}	130704	1231 (4.04-3.60)
Ramachandran outliers	138981	1261 (4.04-3.60)
Sidechain outliers	138945	1255 (4.04-3.60)
RSRZ outliers	127900	1139 (4.04-3.60)

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 >=3, 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5% The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	136	64% 7% 29%
1	E	136	70% 29%
1	K	136	65% 7% 29%
1	O	136	64% 7% 29%
2	B	103	75% 6% 19%
2	F	103	73% 24%
2	L	103	74% 7% 19%

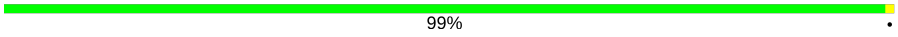
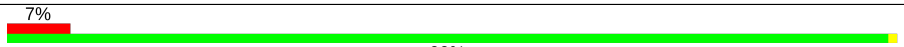
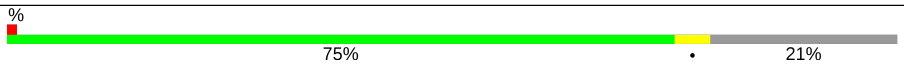
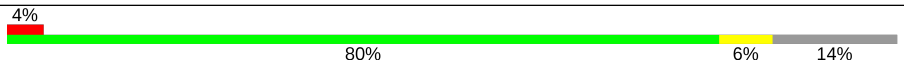
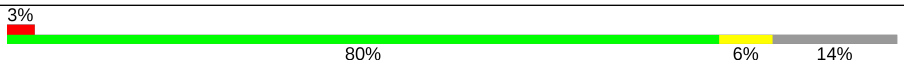
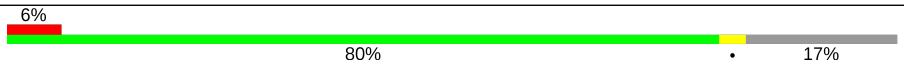
Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	P	103	4% 68% 7% 24%
3	C	130	74% 5% 21%
3	G	130	78% 19%
3	M	130	% 75% 21%
3	Q	130	2% 77% 19%
4	D	123	% 67% 9% 23%
4	H	123	% 72% 24%
4	N	123	67% 11% 23%
4	R	123	66% 9% 24%
5	I	145	99%
5	S	145	99%
6	J	145	98%
6	T	145	99%
7	U	472	% 92% 5%
7	Z	472	2% 93%
7	e	472	% 94%
7	j	472	2% 93% 5%
8	V	96	6% 91% 7%
8	a	96	6% 91% 6%
8	f	96	5% 86% 11%
8	k	96	6% 88% 10%
9	W	99	% 89% 9%
9	b	99	86% 10%
9	g	99	% 90% 9%
9	l	99	2% 91% 6%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
10	X	76	
10	c	76	
10	h	76	
10	m	76	
11	Y	104	
11	d	104	
11	i	104	
11	n	104	

2 Entry composition

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).

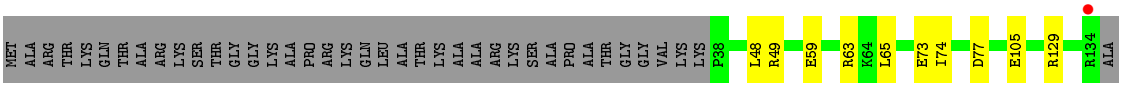
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
12	g	1	Total	Zn	0	0
			1	1		
12	j	6	Total	Zn	0	0
			6	6		
12	d	1	Total	Zn	0	0
			1	1		
12	e	6	Total	Zn	0	0
			6	6		
12	b	1	Total	Zn	0	0
			1	1		
12	i	1	Total	Zn	0	0
			1	1		
12	W	1	Total	Zn	0	0
			1	1		
12	Z	6	Total	Zn	0	0
			6	6		
12	n	1	Total	Zn	0	0
			1	1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

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

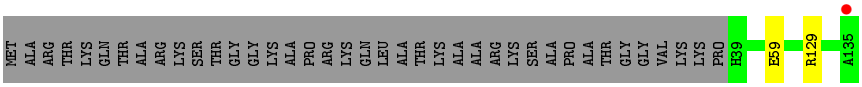
- Molecule 1: Histone H3.2

Chain A:



- Molecule 1: Histone H3.2

Chain E:

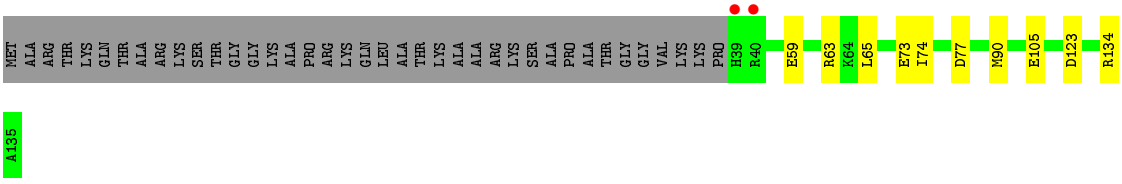


- Molecule 1: Histone H3.2

Chain K:

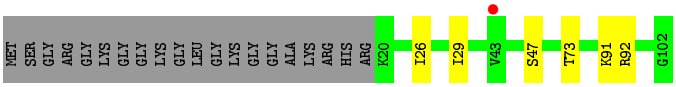
- Molecule 1: Histone H3.2

Chain O:



- Molecule 2: Histone H4

Chain B:

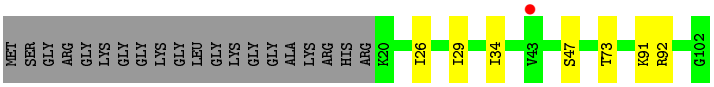


- Molecule 2: Histone H4

Chain F:

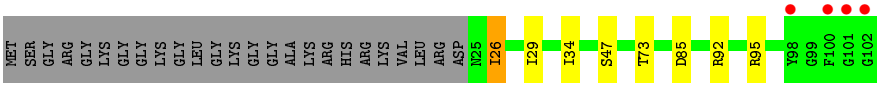
- Molecule 2: Histone H4

Chain L:



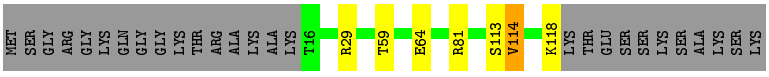
- Molecule 2: Histone H4

Chain P:



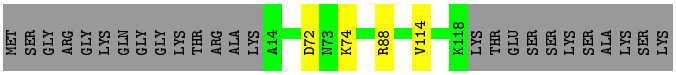
- Molecule 3: Histone H2A type 1

Chain C:



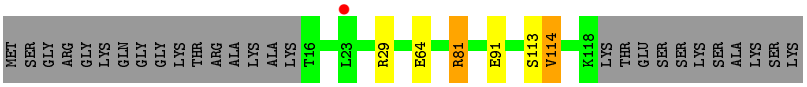
- Molecule 3: Histone H2A type 1

Chain G:



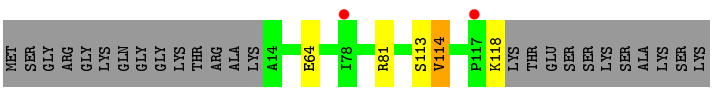
- Molecule 3: Histone H2A type 1

Chain M:



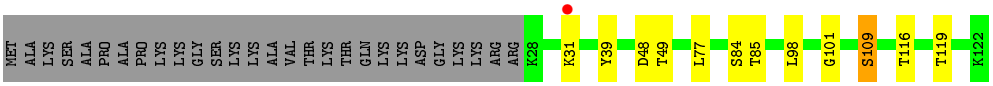
- Molecule 3: Histone H2A type 1

Chain Q:



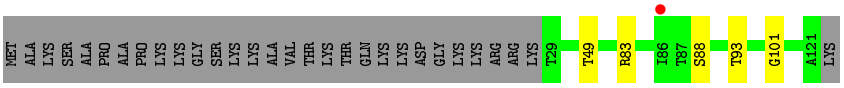
• Molecule 4: Histone H2B 1.1

Chain D:



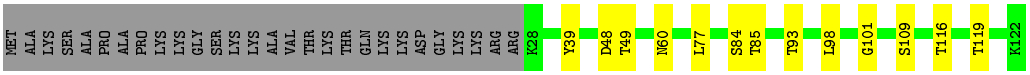
• Molecule 4: Histone H2B 1.1

Chain H:



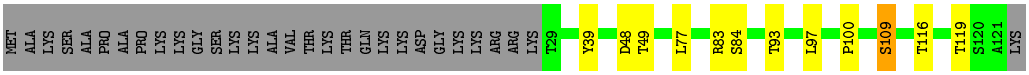
• Molecule 4: Histone H2B 1.1

Chain N:



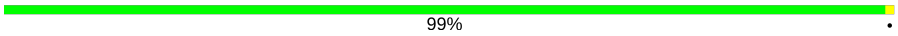
• Molecule 4: Histone H2B 1.1

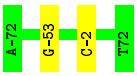
Chain R:



• Molecule 5: DNA (145-MER)

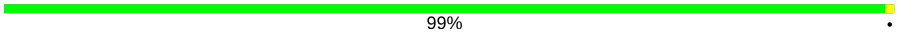
Chain I:





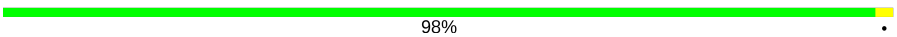
• Molecule 5: DNA (145-MER)

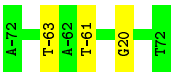
Chain S:



• Molecule 6: DNA (145-MER)

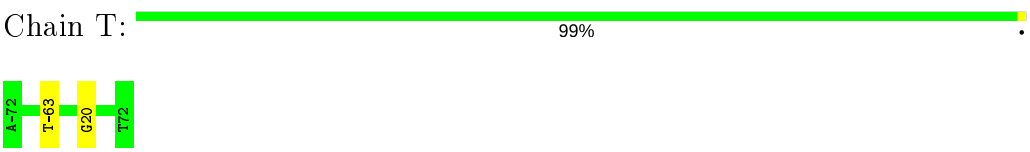
Chain J:





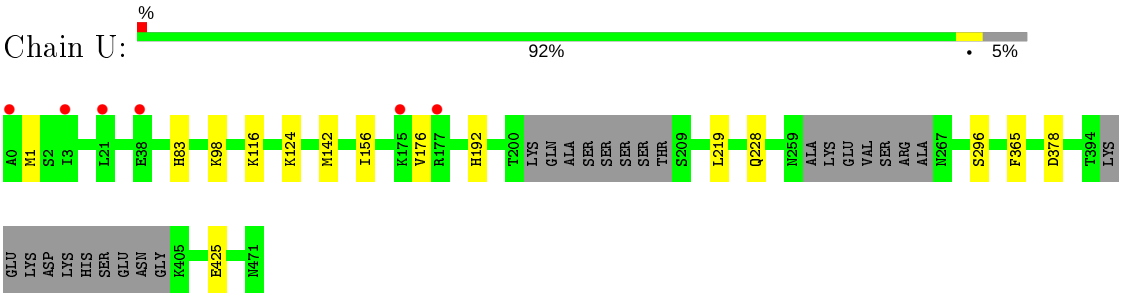
• Molecule 6: DNA (145-MER)

Chain T:



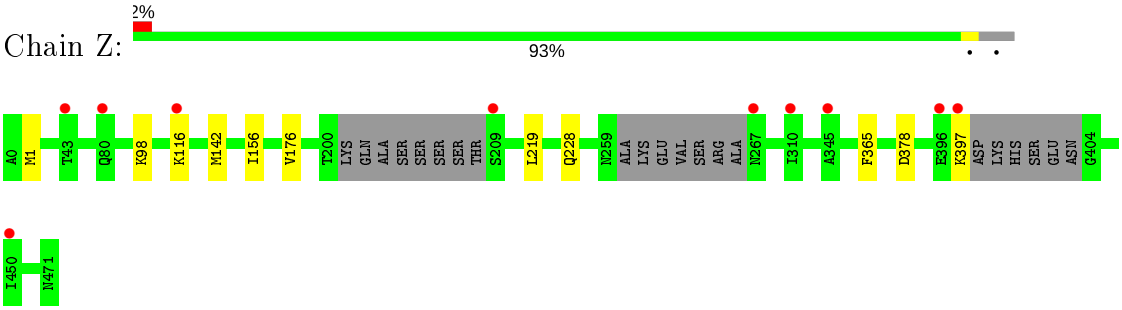
• Molecule 7: Ubiquitin carboxyl-terminal hydrolase 8

Chain U:



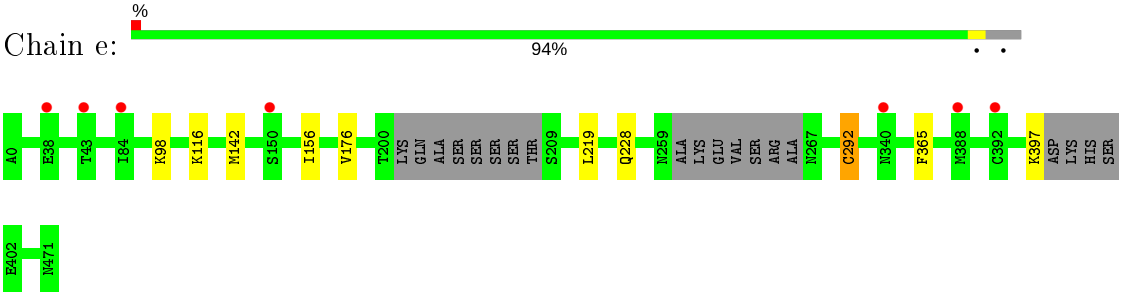
• Molecule 7: Ubiquitin carboxyl-terminal hydrolase 8

Chain Z:



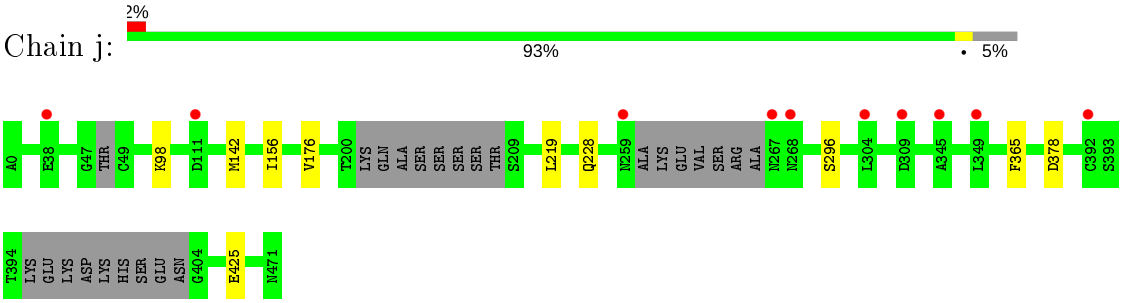
• Molecule 7: Ubiquitin carboxyl-terminal hydrolase 8

Chain e:



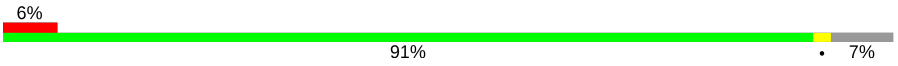
• Molecule 7: Ubiquitin carboxyl-terminal hydrolase 8

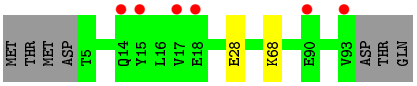
Chain j:



• Molecule 8: Transcription and mRNA export factor SUS1

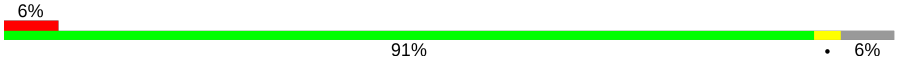
Chain V:

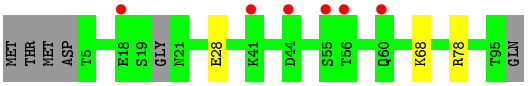




• Molecule 8: Transcription and mRNA export factor SUS1

Chain a:



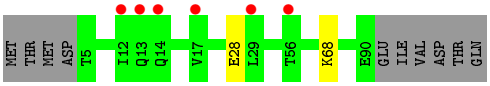


• Molecule 8: Transcription and mRNA export factor SUS1

Chain f:

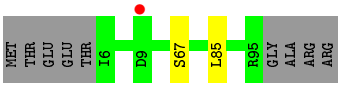
• Molecule 8: Transcription and mRNA export factor SUS1

Chain k:



• Molecule 9: SAGA-associated factor 11

Chain W:

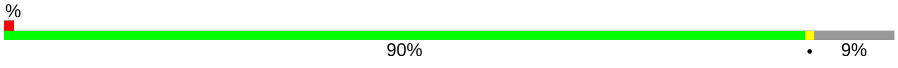


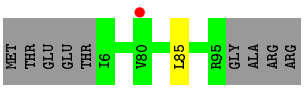
• Molecule 9: SAGA-associated factor 11

Chain b:

• Molecule 9: SAGA-associated factor 11

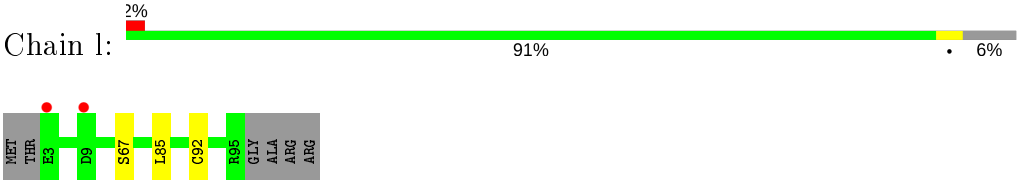
Chain g:





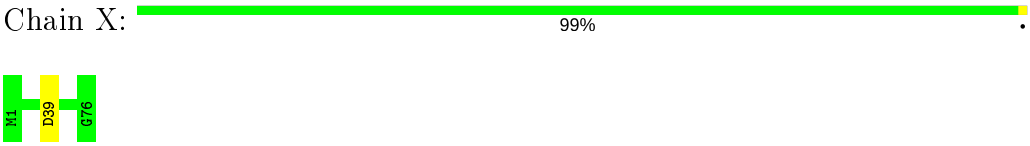
• Molecule 9: SAGA-associated factor 11

Chain l:



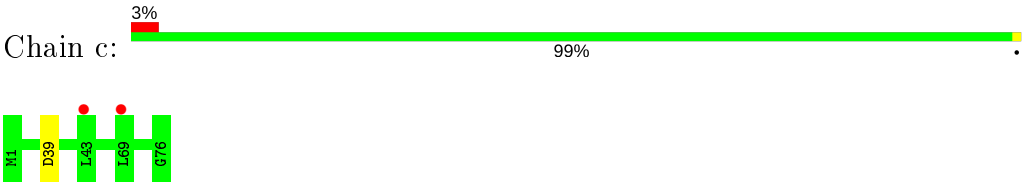
• Molecule 10: Polyubiquitin-B

Chain X:



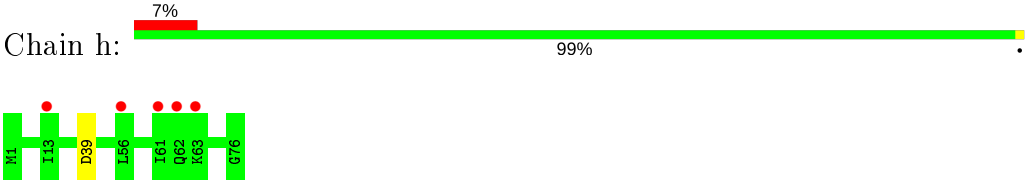
• Molecule 10: Polyubiquitin-B

Chain c:



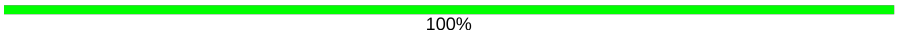
• Molecule 10: Polyubiquitin-B

Chain h:



• Molecule 10: Polyubiquitin-B

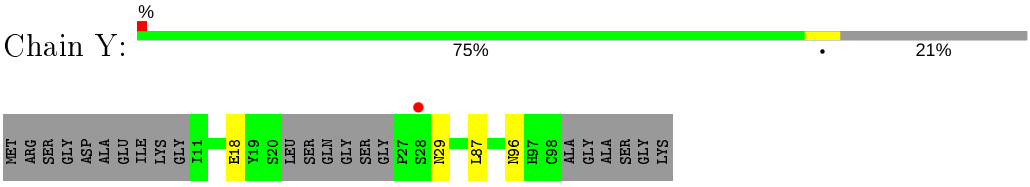
Chain m:



There are no outlier residues recorded for this chain.

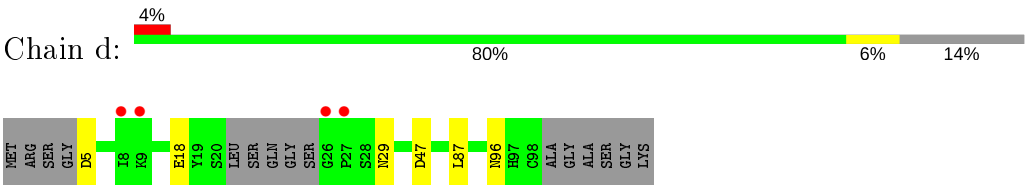
• Molecule 11: SAGA-associated factor 73

Chain Y:



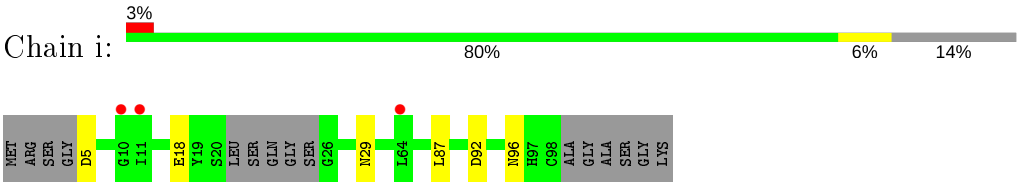
• Molecule 11: SAGA-associated factor 73

Chain d:



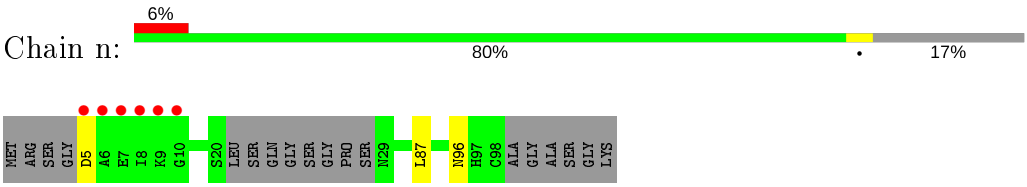
• Molecule 11: SAGA-associated factor 73

Chain i:



- Molecule 11: SAGA-associated factor 73

Chain n:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	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
$< I/\sigma(I) >$ [1]	1.47 (at 3.67Å)	Xtriage
Refinement program	PHENIX 1.9_1692	Depositor
R, R_{free}	0.234 , 0.256 0.232 , 0.260	Depositor DCC
R_{free} test set	1997 reflections (1.64%)	wwPDB-VP
Wilson B-factor (Å^2)	168.1	Xtriage
Anisotropy	0.052	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.27 , 121.5	EDS
L-test for twinning[2]	$< \|L\| > = 0.37$, $< L^2 > = 0.20$	Xtriage
Estimated twinning fraction	0.228 for k,h,-l	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	49079	wwPDB-VP
Average B, all atoms (Å^2)	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.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

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.

All (14) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
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
6	J	-61	DT	O4'-C1'-N1	5.34	111.73	108.00
6	T	20	DG	C4'-C3'-C2'	-5.33	98.30	103.10
6	T	20	DG	C3'-C2'-C1'	-5.27	96.17	102.50
5	S	-53	DG	P-O3'-C3'	5.25	126.00	119.70
6	T	-63	DT	O4'-C1'-N1	5.23	111.66	108.00
6	J	20	DG	C3'-C2'-C1'	-5.17	96.29	102.50
6	J	-63	DT	C1'-O4'-C4'	-5.17	104.93	110.10
6	J	-63	DT	O4'-C1'-N1	5.13	111.59	108.00
6	J	20	DG	C4'-C3'-C2'	-5.00	98.60	103.10

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	Percentiles
1	A	95/136 (70%)	84 (88%)	8 (8%)	3 (3%)	4 32
1	E	95/136 (70%)	86 (90%)	9 (10%)	0	100 100
1	K	95/136 (70%)	84 (88%)	8 (8%)	3 (3%)	4 32
1	O	95/136 (70%)	84 (88%)	8 (8%)	3 (3%)	4 32
2	B	81/103 (79%)	71 (88%)	8 (10%)	2 (2%)	5 36
2	F	76/103 (74%)	69 (91%)	7 (9%)	0	100 100
2	L	81/103 (79%)	71 (88%)	7 (9%)	3 (4%)	3 29
2	P	76/103 (74%)	65 (86%)	8 (10%)	3 (4%)	3 28
3	C	101/130 (78%)	89 (88%)	9 (9%)	3 (3%)	4 33
3	G	103/130 (79%)	91 (88%)	12 (12%)	0	100 100
3	M	101/130 (78%)	90 (89%)	8 (8%)	3 (3%)	4 33
3	Q	103/130 (79%)	91 (88%)	9 (9%)	3 (3%)	4 34
4	D	93/123 (76%)	84 (90%)	7 (8%)	2 (2%)	6 38
4	H	91/123 (74%)	81 (89%)	9 (10%)	1 (1%)	14 51
4	N	93/123 (76%)	83 (89%)	8 (9%)	2 (2%)	6 38
4	R	91/123 (74%)	82 (90%)	7 (8%)	2 (2%)	6 38
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%)	47 78
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

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
9	b	87/99 (88%)	84 (97%)	2 (2%)	1 (1%)	14 51
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%)	13 49
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%)	19 56

All (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
4	N	101	GLY
4	R	100	PRO
9	b	76	CYS
11	d	47	ASP
4	H	101	GLY
1	K	77	ASP
4	N	109	SER
1	O	73	GLU
1	O	77	ASP
2	P	29	ILE
1	A	77	ASP
4	D	109	SER
4	R	109	SER
7	e	292	CYS
2	B	26	ILE
3	C	64	GLU
3	M	64	GLU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
3	Q	64	GLU
1	A	74	ILE
1	K	74	ILE
2	L	26	ILE
1	O	74	ILE
3	C	113	SER
3	M	113	SER
3	Q	113	SER
2	P	26	ILE
3	Q	114	VAL
3	C	114	VAL
2	P	34	ILE
2	L	34	ILE
3	M	114	VAL

5.3.2 Protein sidechains (i)

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

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	A	85/111 (77%)	78 (92%)	7 (8%)	11 40
1	E	84/111 (76%)	82 (98%)	2 (2%)	49 71
1	K	85/111 (77%)	79 (93%)	6 (7%)	14 45
1	O	84/111 (76%)	77 (92%)	7 (8%)	11 40
2	B	68/79 (86%)	64 (94%)	4 (6%)	19 51
2	F	63/79 (80%)	60 (95%)	3 (5%)	25 55
2	L	68/79 (86%)	64 (94%)	4 (6%)	19 51
2	P	63/79 (80%)	57 (90%)	6 (10%)	8 34
3	C	82/102 (80%)	77 (94%)	5 (6%)	18 50
3	G	83/102 (81%)	79 (95%)	4 (5%)	25 55
3	M	82/102 (80%)	78 (95%)	4 (5%)	25 55
3	Q	83/102 (81%)	80 (96%)	3 (4%)	35 63
4	D	81/103 (79%)	70 (86%)	11 (14%)	3 22

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
4	H	79/103 (77%)	75 (95%)	4 (5%)	24 54
4	N	81/103 (79%)	70 (86%)	11 (14%)	3 22
4	R	79/103 (77%)	69 (87%)	10 (13%)	4 23
7	U	406/428 (95%)	391 (96%)	15 (4%)	34 61
7	Z	409/428 (96%)	398 (97%)	11 (3%)	44 68
7	e	411/428 (96%)	401 (98%)	10 (2%)	49 71
7	j	405/428 (95%)	395 (98%)	10 (2%)	47 70
8	V	84/91 (92%)	82 (98%)	2 (2%)	49 71
8	a	86/91 (94%)	83 (96%)	3 (4%)	36 63
8	f	81/91 (89%)	79 (98%)	2 (2%)	47 70
8	k	81/91 (89%)	79 (98%)	2 (2%)	47 70
9	W	82/89 (92%)	80 (98%)	2 (2%)	49 71
9	b	81/89 (91%)	78 (96%)	3 (4%)	34 61
9	g	82/89 (92%)	81 (99%)	1 (1%)	71 84
9	l	84/89 (94%)	81 (96%)	3 (4%)	35 63
10	X	68/68 (100%)	67 (98%)	1 (2%)	65 80
10	c	68/68 (100%)	67 (98%)	1 (2%)	65 80
10	h	68/68 (100%)	67 (98%)	1 (2%)	65 80
10	m	68/68 (100%)	68 (100%)	0	100 100
11	Y	77/90 (86%)	73 (95%)	4 (5%)	23 54
11	d	81/90 (90%)	76 (94%)	5 (6%)	18 49
11	i	81/90 (90%)	75 (93%)	6 (7%)	13 44
11	n	79/90 (88%)	76 (96%)	3 (4%)	33 60
All	All	4132/4644 (89%)	3956 (96%)	176 (4%)	29 58

All (176) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	48	LEU
1	A	49	ARG
1	A	59	GLU
1	A	63	ARG
1	A	65	LEU
1	A	105	GLU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	129	ARG
2	B	47	SER
2	B	73	THR
2	B	91	LYS
2	B	92	ARG
3	C	29	ARG
3	C	59	THR
3	C	81	ARG
3	C	114	VAL
3	C	118	LYS
4	D	31	LYS
4	D	39	TYR
4	D	48	ASP
4	D	49	THR
4	D	77	LEU
4	D	84	SER
4	D	85	THR
4	D	98	LEU
4	D	109	SER
4	D	116	THR
4	D	119	THR
1	E	59	GLU
1	E	129	ARG
2	F	26	ILE
2	F	73	THR
2	F	92	ARG
3	G	72	ASP
3	G	74	LYS
3	G	88	ARG
3	G	114	VAL
4	H	49	THR
4	H	83	ARG
4	H	88	SER
4	H	93	THR
1	K	49	ARG
1	K	59	GLU
1	K	63	ARG
1	K	65	LEU
1	K	105	GLU
1	K	129	ARG
2	L	47	SER
2	L	73	THR

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	L	91	LYS
2	L	92	ARG
3	M	29	ARG
3	M	81	ARG
3	M	91	GLU
3	M	114	VAL
4	N	39	TYR
4	N	48	ASP
4	N	49	THR
4	N	60	ASN
4	N	77	LEU
4	N	84	SER
4	N	85	THR
4	N	93	THR
4	N	98	LEU
4	N	116	THR
4	N	119	THR
1	O	59	GLU
1	O	63	ARG
1	O	65	LEU
1	O	90	MET
1	O	105	GLU
1	O	123	ASP
1	O	134	ARG
2	P	26	ILE
2	P	47	SER
2	P	73	THR
2	P	85	ASP
2	P	92	ARG
2	P	95	ARG
3	Q	81	ARG
3	Q	114	VAL
3	Q	118	LYS
4	R	39	TYR
4	R	48	ASP
4	R	49	THR
4	R	77	LEU
4	R	84	SER
4	R	93	THR
4	R	97	LEU
4	R	109	SER
4	R	116	THR

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
4	R	119	THR
7	U	1	MET
7	U	83	HIS
7	U	98	LYS
7	U	116	LYS
7	U	124	LYS
7	U	142	MET
7	U	156	ILE
7	U	176	VAL
7	U	192	HIS
7	U	219	LEU
7	U	228	GLN
7	U	296	SER
7	U	365	PHE
7	U	378	ASP
7	U	425	GLU
8	V	28	GLU
8	V	68	LYS
9	W	67	SER
9	W	85	LEU
10	X	39	ASP
11	Y	18	GLU
11	Y	29	ASN
11	Y	87	LEU
11	Y	96	ASN
7	Z	1	MET
7	Z	98	LYS
7	Z	116	LYS
7	Z	142	MET
7	Z	156	ILE
7	Z	176	VAL
7	Z	219	LEU
7	Z	228	GLN
7	Z	365	PHE
7	Z	378	ASP
7	Z	397	LYS
8	a	28	GLU
8	a	68	LYS
8	a	78	ARG
9	b	67	SER
9	b	75	ASN
9	b	85	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
10	c	39	ASP
11	d	5	ASP
11	d	18	GLU
11	d	29	ASN
11	d	87	LEU
11	d	96	ASN
7	e	98	LYS
7	e	116	LYS
7	e	142	MET
7	e	156	ILE
7	e	176	VAL
7	e	219	LEU
7	e	228	GLN
7	e	292	CYS
7	e	365	PHE
7	e	397	LYS
8	f	28	GLU
8	f	68	LYS
9	g	85	LEU
10	h	39	ASP
11	i	5	ASP
11	i	18	GLU
11	i	29	ASN
11	i	87	LEU
11	i	92	ASP
11	i	96	ASN
7	j	98	LYS
7	j	142	MET
7	j	156	ILE
7	j	176	VAL
7	j	219	LEU
7	j	228	GLN
7	j	296	SER
7	j	365	PHE
7	j	378	ASP
7	j	425	GLU
8	k	28	GLU
8	k	68	LYS
9	l	67	SER
9	l	85	LEU
9	l	92	CYS
11	n	5	ASP

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
11	n	87	LEU
11	n	96	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (16) such sidechains are listed below:

Mol	Chain	Res	Type
2	B	75	HIS
2	L	75	HIS
4	N	60	ASN
4	N	64	ASN
1	O	39	HIS
3	Q	112	GLN
7	U	53	ASN
7	U	83	HIS
7	U	90	ASN
7	Z	88	ASN
7	Z	90	ASN
7	Z	328	GLN
9	b	75	ASN
7	e	348	GLN
11	i	61	ASN
9	l	88	HIS

5.3.3 RNA

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains

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

5.5 Carbohydrates

There are no monosaccharides in this entry.

5.6 Ligand geometry

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

There are no such residues in this entry.

5.8 Polymer linkage issues

There are no chain breaks in this entry.

6 Fit of model and data

6.1 Protein, DNA and RNA chains

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

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å^2)	Q<0.9
1	A	97/136 (71%)	-0.09	1 (1%) 82 76	134, 166, 198, 212	0
1	E	97/136 (71%)	-0.14	1 (1%) 82 76	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%)	-0.14	2 (2%) 63 55	141, 177, 219, 249	0
2	B	83/103 (80%)	-0.18	1 (1%) 79 72	124, 155, 185, 198	0
2	F	78/103 (75%)	-0.04	2 (2%) 56 46	124, 166, 189, 201	0
2	L	83/103 (80%)	-0.10	1 (1%) 79 72	131, 152, 192, 204	0
2	P	78/103 (75%)	0.13	4 (5%) 28 24	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%) 82 76	143, 168, 195, 216	0
3	Q	105/130 (80%)	-0.05	2 (1%) 66 59	141, 179, 205, 222	0
4	D	95/123 (77%)	-0.18	1 (1%) 80 74	125, 155, 201, 243	0
4	H	93/123 (75%)	-0.21	1 (1%) 80 74	145, 172, 197, 209	0
4	N	95/123 (77%)	-0.17	0 100 100	137, 165, 190, 239	0
4	R	93/123 (75%)	-0.22	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%) 77 70	133, 183, 235, 285	0
7	Z	451/472 (95%)	0.00	10 (2%) 62 53	161, 205, 261, 302	0
7	e	453/472 (95%)	-0.14	7 (1%) 73 66	148, 192, 244, 276	0
7	j	447/472 (94%)	-0.16	10 (2%) 62 53	146, 189, 232, 286	0

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å^2)	Q<0.9
8	V	89/96 (92%)	-0.09	6 (6%)	17	13	160, 196, 260, 275	0
8	a	90/96 (93%)	0.10	6 (6%)	17	13	181, 238, 267, 283	0
8	f	85/96 (88%)	0.26	5 (5%)	22	17	196, 227, 267, 298	0
8	k	86/96 (89%)	-0.04	6 (6%)	16	12	188, 217, 249, 277	1 (1%)
9	W	90/99 (90%)	-0.31	1 (1%)	80	74	163, 196, 217, 231	0
9	b	89/99 (89%)	-0.23	0	100	100	173, 219, 252, 260	0
9	g	90/99 (90%)	-0.26	1 (1%)	80	74	170, 212, 266, 282	1 (1%)
9	l	93/99 (93%)	-0.19	2 (2%)	62	53	168, 201, 263, 275	0
10	X	76/76 (100%)	-0.06	0	100	100	157, 193, 221, 237	0
10	c	76/76 (100%)	-0.04	2 (2%)	56	46	159, 192, 218, 252	0
10	h	76/76 (100%)	0.07	5 (6%)	18	14	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%)	-0.00	1 (1%)	79	72	159, 193, 257, 301	0
11	d	89/104 (85%)	0.12	4 (4%)	33	28	174, 212, 279, 315	0
11	i	89/104 (85%)	0.02	3 (3%)	45	37	162, 200, 266, 297	0
11	n	86/104 (82%)	0.49	6 (6%)	16	12	167, 197, 281, 349	0
All	All	5242/5936 (88%)	-0.17	98 (1%)	66	59	124, 192, 248, 349	2 (0%)

All (98) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
11	n	6	ALA	22.8
11	n	5	ASP	10.5
11	d	26	GLY	10.3
11	n	7	GLU	9.9
7	Z	397	LYS	6.7
11	d	27	PRO	6.3
7	j	267	ASN	6.2
11	i	11	ILE	6.1
7	Z	267	ASN	5.7
8	f	93	VAL	4.9
2	P	98	TYR	4.7
9	l	3	GLU	4.4
2	P	102	GLY	4.4
11	n	9	LYS	4.3
2	P	100	PHE	4.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
8	f	15	TYR	4.0
11	Y	28	SER	4.0
2	P	101	GLY	3.8
7	Z	209	SER	3.8
8	k	56	THR	3.7
11	i	10	GLY	3.7
1	A	134	ARG	3.7
7	U	3	ILE	3.6
7	j	392	CYS	3.6
8	f	89	LEU	3.5
11	n	8	ILE	3.3
7	j	349	LEU	3.2
8	V	93	VAL	3.2
8	a	60	GLN	3.2
7	e	392	CYS	3.1
10	h	61	ILE	3.0
7	e	340	ASN	3.0
10	h	13	ILE	3.0
1	E	135	ALA	3.0
3	Q	78	ILE	3.0
8	V	17	VAL	3.0
7	Z	80	GLN	3.0
2	F	44	LYS	3.0
8	k	17	VAL	2.9
2	L	43	VAL	2.9
3	Q	117	PRO	2.9
7	j	304	LEU	2.8
9	l	9	ASP	2.8
7	j	268	ASN	2.8
7	e	38	GLU	2.8
7	j	309	ASP	2.8
8	k	14	GLN	2.8
8	V	90	GLU	2.7
10	c	43	LEU	2.7
7	Z	310	ILE	2.7
7	U	21	LEU	2.7
7	j	38	GLU	2.7
8	f	90	GLU	2.7
10	h	62	GLN	2.6
7	e	388	MET	2.6
7	j	259	ASN	2.6
8	V	14	GLN	2.6

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
11	n	10	GLY	2.6
8	a	55	SER	2.6
8	a	41	LYS	2.5
8	V	15	TYR	2.5
7	Z	450	ILE	2.5
7	Z	345	ALA	2.5
11	i	64	LEU	2.4
7	U	177	ARG	2.4
7	U	38	GLU	2.4
2	F	68	ASP	2.4
7	j	345	ALA	2.4
8	k	13	GLN	2.4
11	d	9	LYS	2.4
8	a	44	ASP	2.4
1	O	39	HIS	2.4
7	Z	396	GLU	2.4
7	e	150	SER	2.4
7	U	175	LYS	2.3
8	k	29	LEU	2.3
10	h	56	LEU	2.3
10	h	63	LYS	2.3
11	d	8	ILE	2.3
7	e	43	THR	2.2
9	W	9	ASP	2.2
8	V	18	GLU	2.2
7	e	84	ILE	2.2
4	D	31	LYS	2.2
9	g	80	VAL	2.2
1	O	40	ARG	2.2
2	B	43	VAL	2.2
8	k	12	ILE	2.2
10	c	69	LEU	2.1
7	Z	116	LYS	2.1
8	a	56	THR	2.1
3	M	23	LEU	2.1
7	U	0	ALA	2.1
7	Z	43	THR	2.1
4	H	86	ILE	2.0
7	j	111	ASP	2.0
8	f	92	ILE	2.0
8	a	18	GLU	2.0

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

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

6.3 Carbohydrates (i)

There are no monosaccharides in this entry.

6.4 Ligands (i)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
12	ZN	Z	502	1/1	0.58	0.20	233,233,233,233	0
12	ZN	j	501	1/1	0.62	0.11	286,286,286,286	0
12	ZN	e	506	1/1	0.68	0.04	251,251,251,251	0
12	ZN	e	503	1/1	0.70	0.13	229,229,229,229	0
12	ZN	U	501	1/1	0.71	0.16	228,228,228,228	0
12	ZN	U	506	1/1	0.74	0.13	252,252,252,252	0
12	ZN	W	101	1/1	0.76	0.14	199,199,199,199	0
12	ZN	Z	505	1/1	0.83	0.17	272,272,272,272	0
12	ZN	e	502	1/1	0.83	0.10	279,279,279,279	0
12	ZN	e	501	1/1	0.84	0.14	251,251,251,251	0
12	ZN	n	201	1/1	0.84	0.20	216,216,216,216	0
12	ZN	d	201	1/1	0.85	0.20	226,226,226,226	0
12	ZN	Z	501	1/1	0.85	0.23	234,234,234,234	0
12	ZN	j	504	1/1	0.85	0.16	198,198,198,198	0
12	ZN	b	101	1/1	0.86	0.13	218,218,218,218	0
12	ZN	l	101	1/1	0.88	0.19	173,173,173,173	0
12	ZN	j	503	1/1	0.89	0.18	179,179,179,179	0
12	ZN	Y	201	1/1	0.89	0.18	184,184,184,184	0
12	ZN	Z	506	1/1	0.91	0.14	290,290,290,290	0
12	ZN	Z	504	1/1	0.92	0.12	215,215,215,215	0
12	ZN	Z	503	1/1	0.94	0.10	237,237,237,237	0
12	ZN	e	505	1/1	0.94	0.21	152,152,152,152	0
12	ZN	j	506	1/1	0.95	0.20	192,192,192,192	0
12	ZN	U	503	1/1	0.95	0.18	209,209,209,209	0
12	ZN	U	504	1/1	0.97	0.12	201,201,201,201	0
12	ZN	g	101	1/1	0.97	0.17	206,206,206,206	0
12	ZN	j	505	1/1	0.97	0.12	212,212,212,212	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
12	ZN	U	502	1/1	0.98	0.17	195,195,195,195	0
12	ZN	i	201	1/1	0.98	0.17	195,195,195,195	0
12	ZN	U	505	1/1	0.99	0.18	160,160,160,160	0
12	ZN	e	504	1/1	0.99	0.14	188,188,188,188	0
12	ZN	j	502	1/1	0.99	0.15	160,160,160,160	0

6.5 Other polymers

There are no such residues in this entry.