



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 07:09 PM GMT

PDB ID : 4NE1
Title : Human MHF1 MHF2 DNA complexes
Authors : Zhao, Q.; Saro, D.; Sachpatzidis, A.; Sung, P.; Xiong, Y.
Deposited on : 2013-10-28
Resolution : 6.50 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
<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 : 1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20026688
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac : 5.8.0135
CCP4 : 6.5.0
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : trunk26865

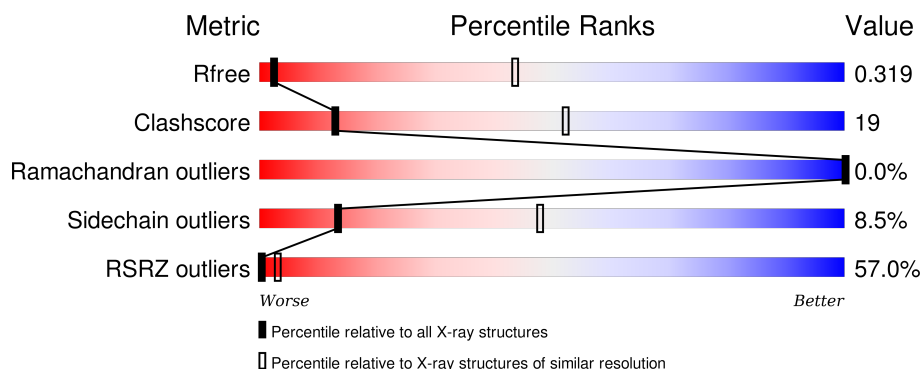
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 6.50 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	1012 (9.00-3.66)
Clashscore	102246	1060 (9.00-3.70)
Ramachandran outliers	100387	1033 (9.00-3.66)
Sidechain outliers	100360	1004 (9.00-3.66)
RSRZ outliers	91569	1011 (9.00-3.66)

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. 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	E	26	<div> <div>69%</div> <div>31%</div> </div>
1	O	26	<div> <div>46%</div> <div>50%</div> <div>.</div> </div>
1	s	26	<div> <div>4%</div> <div>100%</div> </div>
1	u	26	<div> <div>4%</div> <div>100%</div> </div>
2	F	26	<div> <div>8%</div> <div>46%</div> <div>54%</div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	P	26	
2	t	26	
2	v	26	
3	A	105	
3	C	105	
3	G	105	
3	I	105	
3	J	105	
3	K	105	
3	Q	105	
3	R	105	
3	S	105	
3	T	105	
3	Y	105	
3	a	105	
3	c	105	
3	e	105	
3	f	105	
3	g	105	
3	k	105	
3	l	105	
3	m	105	
3	n	105	
4	B	74	
4	D	74	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
4	H	74	
4	L	74	
4	M	74	
4	N	74	
4	U	74	
4	V	74	
4	W	74	
4	X	74	
4	Z	74	
4	b	74	
4	d	74	
4	h	74	
4	i	74	
4	j	74	
4	o	74	
4	p	74	
4	q	74	
4	r	74	

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 31827 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 DNA chain called DNA (26-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	E	26	Total	C	N	O	P	0	0	0
			546	260	130	130	26			
1	O	26	Total	C	N	O	P	0	0	0
			546	260	130	130	26			
1	s	26	Total	C	N	O	P	0	0	0
			546	260	130	130	26			
1	u	26	Total	C	N	O	P	0	0	0
			546	260	130	130	26			

- Molecule 2 is a DNA chain called DNA (26-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	F	26	Total	C	N	O	P	0	0	0
			520	260	52	182	26			
2	P	26	Total	C	N	O	P	0	0	0
			520	260	52	182	26			
2	t	26	Total	C	N	O	P	0	0	0
			520	260	52	182	26			
2	v	26	Total	C	N	O	P	0	0	0
			520	260	52	182	26			

- Molecule 3 is a protein called Centromere protein S.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	93	Total	C	N	O	S	0	0	0
			743	464	132	142	5			
3	G	93	Total	C	N	O	S	0	0	0
			743	464	132	142	5			
3	A	93	Total	C	N	O	S	0	0	0
			743	464	132	142	5			
3	I	105	Total	C	N	O	S	0	0	0
			844	526	154	159	5			

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	J	93	Total	C	N	O	S	0	0	0
			743	464	132	142	5			
3	K	105	Total	C	N	O	S	0	0	0
			844	526	154	159	5			
3	Q	105	Total	C	N	O	S	0	0	0
			844	526	154	159	5			
3	R	93	Total	C	N	O	S	0	0	0
			743	464	132	142	5			
3	S	93	Total	C	N	O	S	0	0	0
			743	464	132	142	5			
3	T	105	Total	C	N	O	S	0	0	0
			844	526	154	159	5			
3	Y	93	Total	C	N	O	S	0	0	0
			743	464	132	142	5			
3	a	93	Total	C	N	O	S	0	0	0
			743	464	132	142	5			
3	c	93	Total	C	N	O	S	0	0	0
			743	464	132	142	5			
3	e	105	Total	C	N	O	S	0	0	0
			844	526	154	159	5			
3	f	93	Total	C	N	O	S	0	0	0
			743	464	132	142	5			
3	g	105	Total	C	N	O	S	0	0	0
			844	526	154	159	5			
3	k	105	Total	C	N	O	S	0	0	0
			844	526	154	159	5			
3	l	93	Total	C	N	O	S	0	0	0
			743	464	132	142	5			
3	m	93	Total	C	N	O	S	0	0	0
			743	464	132	142	5			
3	n	105	Total	C	N	O	S	0	0	0
			844	526	154	159	5			

There are 40 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	39	ALA	GLU	CONFLICT	UNP Q8N2Z9
C	106	ALA	ILE	CONFLICT	UNP Q8N2Z9
G	39	ALA	GLU	CONFLICT	UNP Q8N2Z9
G	106	ALA	ILE	CONFLICT	UNP Q8N2Z9
A	39	ALA	GLU	CONFLICT	UNP Q8N2Z9
A	106	ALA	ILE	CONFLICT	UNP Q8N2Z9
I	39	ALA	GLU	CONFLICT	UNP Q8N2Z9

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
I	106	ALA	ILE	CONFLICT	UNP Q8N2Z9
J	39	ALA	GLU	CONFLICT	UNP Q8N2Z9
J	106	ALA	ILE	CONFLICT	UNP Q8N2Z9
K	39	ALA	GLU	CONFLICT	UNP Q8N2Z9
K	106	ALA	ILE	CONFLICT	UNP Q8N2Z9
Q	39	ALA	GLU	CONFLICT	UNP Q8N2Z9
Q	106	ALA	ILE	CONFLICT	UNP Q8N2Z9
R	39	ALA	GLU	CONFLICT	UNP Q8N2Z9
R	106	ALA	ILE	CONFLICT	UNP Q8N2Z9
S	39	ALA	GLU	CONFLICT	UNP Q8N2Z9
S	106	ALA	ILE	CONFLICT	UNP Q8N2Z9
T	39	ALA	GLU	CONFLICT	UNP Q8N2Z9
T	106	ALA	ILE	CONFLICT	UNP Q8N2Z9
Y	39	ALA	GLU	CONFLICT	UNP Q8N2Z9
Y	106	ALA	ILE	CONFLICT	UNP Q8N2Z9
a	39	ALA	GLU	CONFLICT	UNP Q8N2Z9
a	106	ALA	ILE	CONFLICT	UNP Q8N2Z9
c	39	ALA	GLU	CONFLICT	UNP Q8N2Z9
c	106	ALA	ILE	CONFLICT	UNP Q8N2Z9
e	39	ALA	GLU	CONFLICT	UNP Q8N2Z9
e	106	ALA	ILE	CONFLICT	UNP Q8N2Z9
f	39	ALA	GLU	CONFLICT	UNP Q8N2Z9
f	106	ALA	ILE	CONFLICT	UNP Q8N2Z9
g	39	ALA	GLU	CONFLICT	UNP Q8N2Z9
g	106	ALA	ILE	CONFLICT	UNP Q8N2Z9
k	39	ALA	GLU	CONFLICT	UNP Q8N2Z9
k	106	ALA	ILE	CONFLICT	UNP Q8N2Z9
l	39	ALA	GLU	CONFLICT	UNP Q8N2Z9
l	106	ALA	ILE	CONFLICT	UNP Q8N2Z9
m	39	ALA	GLU	CONFLICT	UNP Q8N2Z9
m	106	ALA	ILE	CONFLICT	UNP Q8N2Z9
n	39	ALA	GLU	CONFLICT	UNP Q8N2Z9
n	106	ALA	ILE	CONFLICT	UNP Q8N2Z9

- Molecule 4 is a protein called Centromere protein X.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	74	Total	C	N	O	S	0	0	0
			590	378	104	107	1			
4	H	74	Total	C	N	O	S	0	0	0
			590	378	104	107	1			
4	B	74	Total	C	N	O	S	0	1	0
			596	382	105	108	1			

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	L	74	Total 596	C 382	N 105	O 108	S 1	0	1	0
4	M	74	Total 596	C 382	N 105	O 108	S 1	0	1	0
4	N	74	Total 596	C 382	N 105	O 108	S 1	0	1	0
4	U	74	Total 596	C 382	N 105	O 108	S 1	0	1	0
4	V	74	Total 596	C 382	N 105	O 108	S 1	0	1	0
4	W	74	Total 596	C 382	N 105	O 108	S 1	0	1	0
4	X	74	Total 596	C 382	N 105	O 108	S 1	0	1	0
4	Z	74	Total 590	C 378	N 104	O 107	S 1	0	0	0
4	b	74	Total 590	C 378	N 104	O 107	S 1	0	0	0
4	d	74	Total 596	C 382	N 105	O 108	S 1	0	1	0
4	h	74	Total 596	C 382	N 105	O 108	S 1	0	1	0
4	i	74	Total 596	C 382	N 105	O 108	S 1	0	1	0
4	j	74	Total 596	C 382	N 105	O 108	S 1	0	1	0
4	o	74	Total 596	C 382	N 105	O 108	S 1	0	1	0
4	p	74	Total 596	C 382	N 105	O 108	S 1	0	1	0
4	q	74	Total 596	C 382	N 105	O 108	S 1	0	1	0
4	r	74	Total 595	C 382	N 105	O 107	S 1	0	1	0

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: DNA (26-MER)

Chain E:  69% 31%



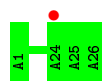
- Molecule 1: DNA (26-MER)

Chain O:  46% 50% .



- Molecule 1: DNA (26-MER)

Chain s:  4% 100%



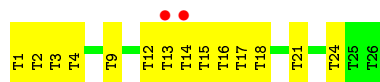
- Molecule 1: DNA (26-MER)

Chain u:  4% 100%



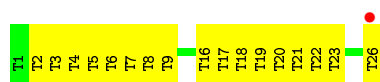
- Molecule 2: DNA (26-MER)

Chain F:  8% 46% 54%



- Molecule 2: DNA (26-MER)

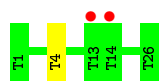
Chain P:  4% 35% 65%



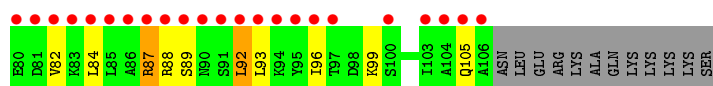
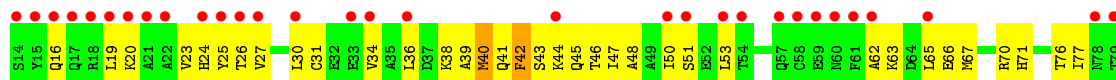
- Molecule 2: DNA (26-MER)



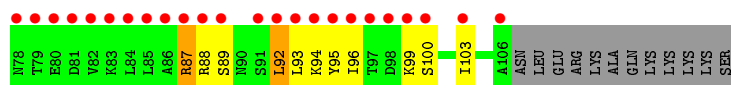
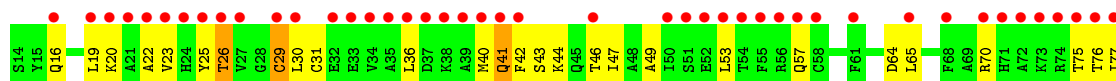
- Molecule 2: DNA (26-MER)



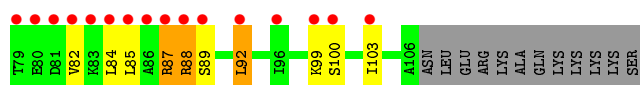
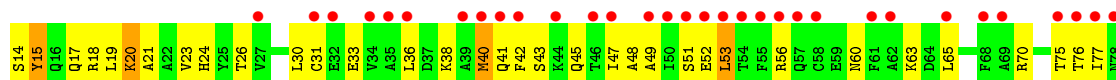
- Molecule 3: Centromere protein S



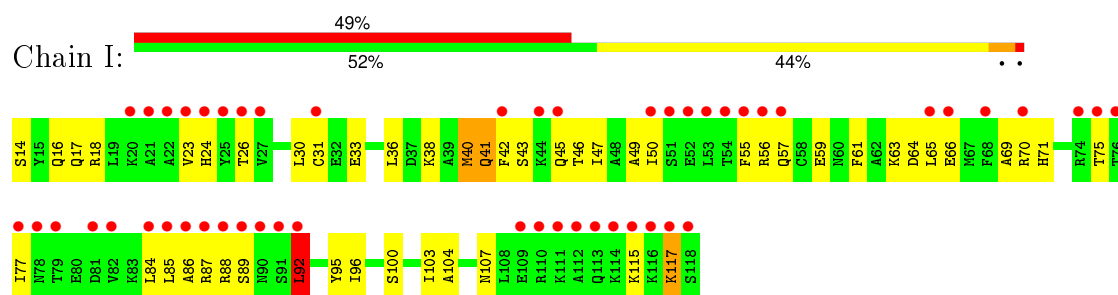
- Molecule 3: Centromere protein S



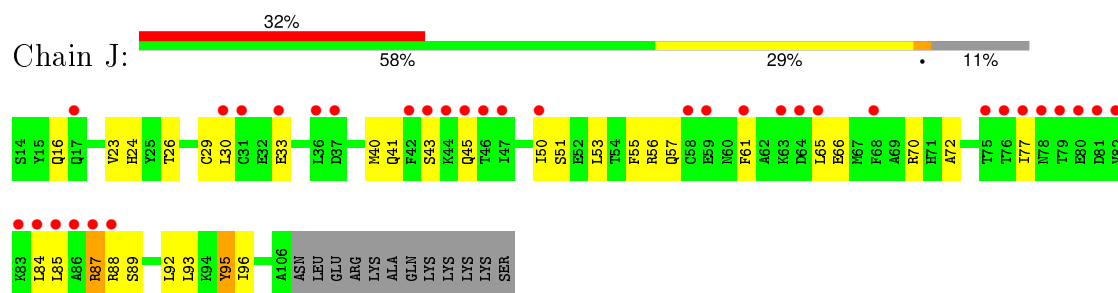
- Molecule 3: Centromere protein S



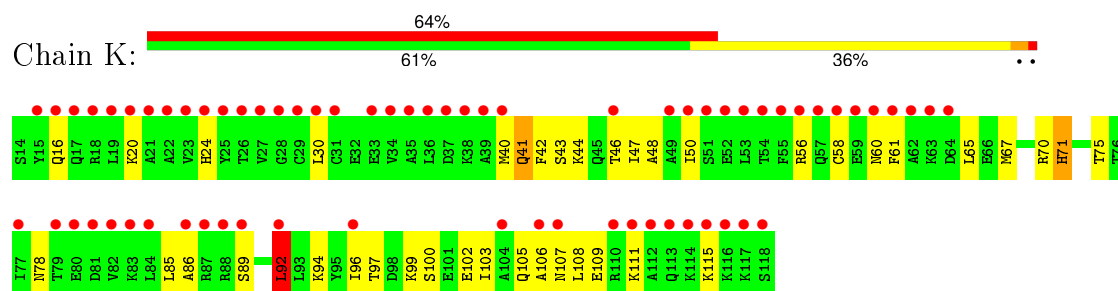
- Molecule 3: Centromere protein S



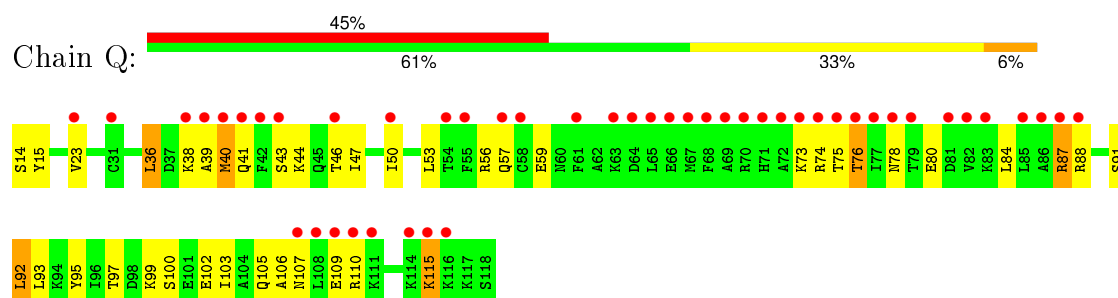
• Molecule 3: Centromere protein S



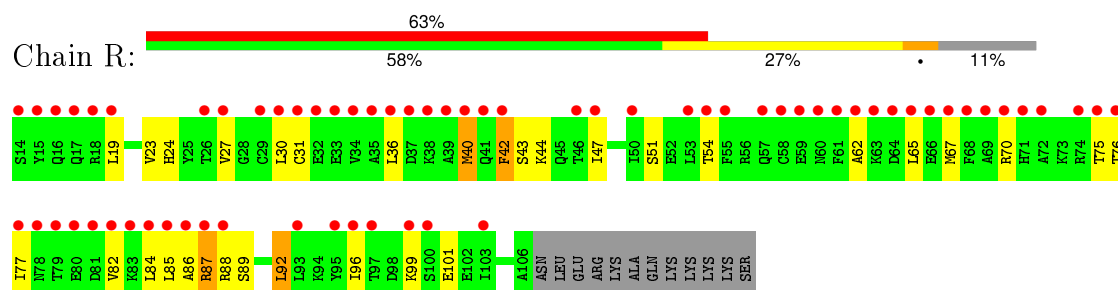
• Molecule 3: Centromere protein S



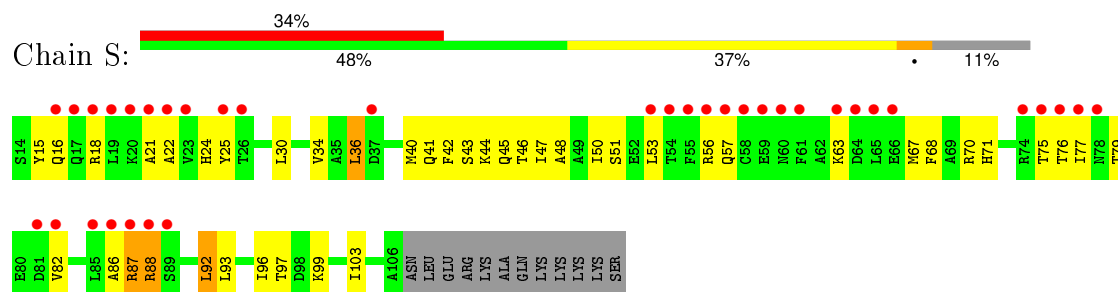
• Molecule 3: Centromere protein S



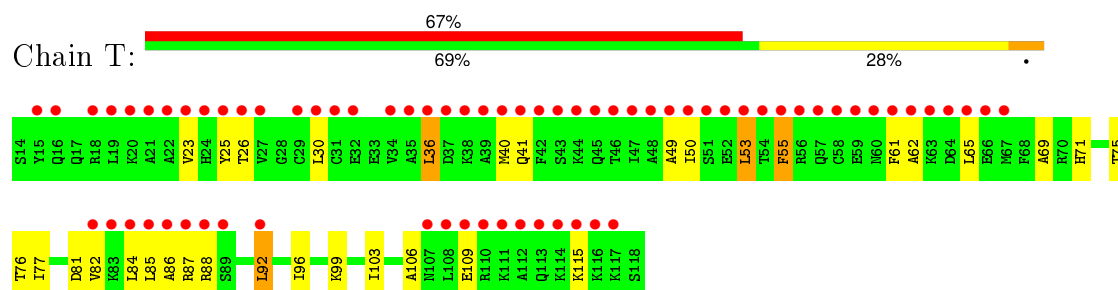
• Molecule 3: Centromere protein S



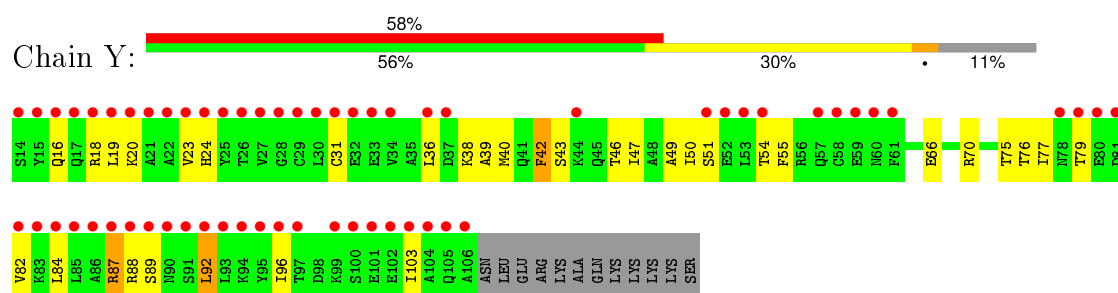
- Molecule 3: Centromere protein S



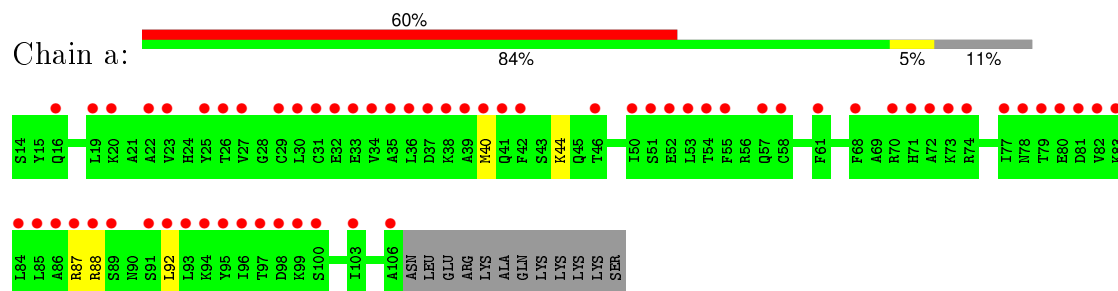
- Molecule 3: Centromere protein S



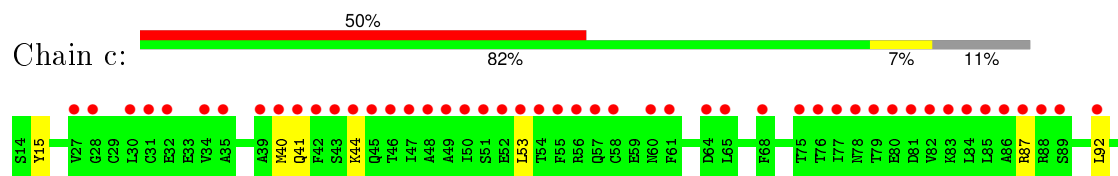
- Molecule 3: Centromere protein S

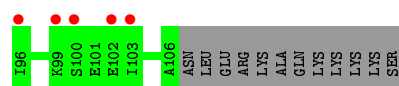


- Molecule 3: Centromere protein S

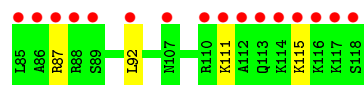
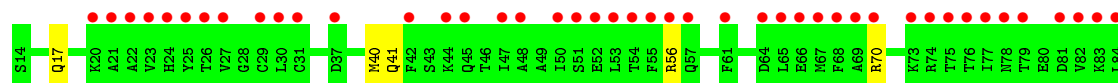
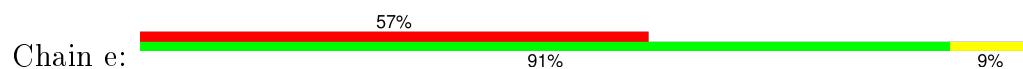


- Molecule 3: Centromere protein S

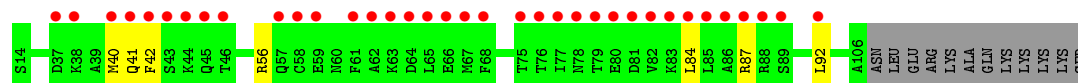
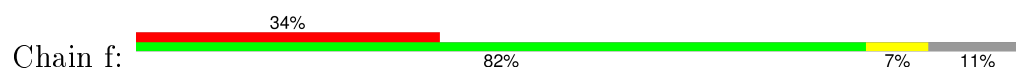




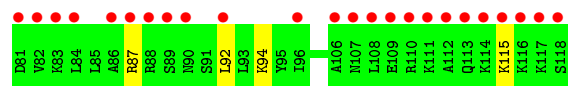
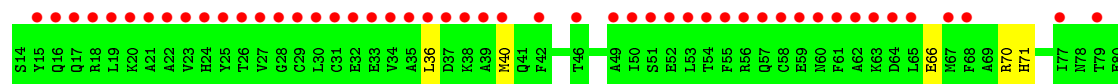
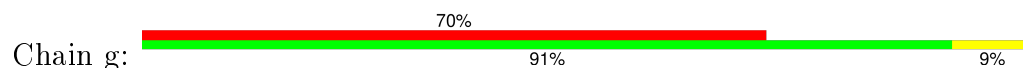
• Molecule 3: Centromere protein S



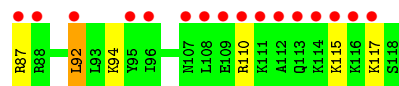
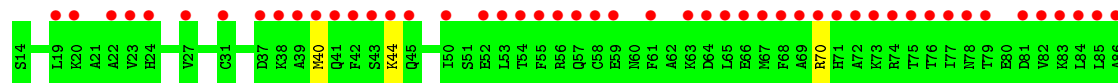
• Molecule 3: Centromere protein S



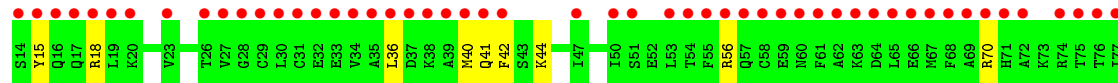
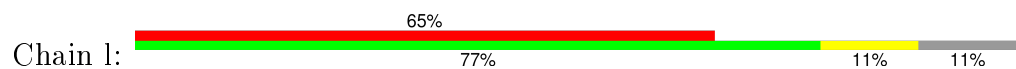
• Molecule 3: Centromere protein S

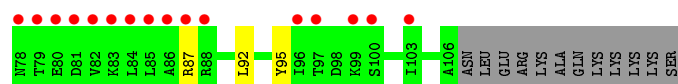


• Molecule 3: Centromere protein S

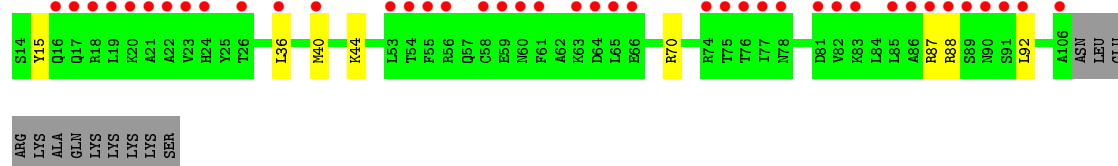
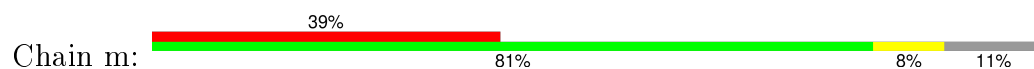


• Molecule 3: Centromere protein S

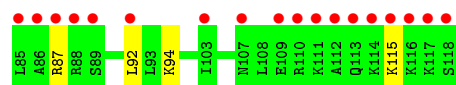
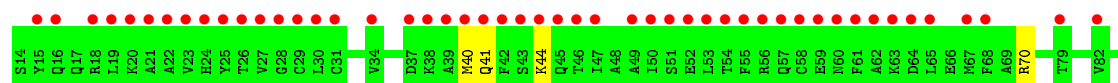
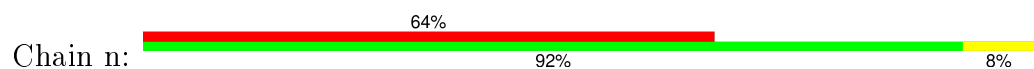




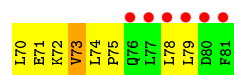
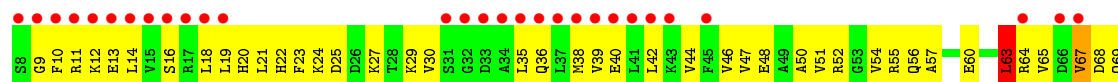
• Molecule 3: Centromere protein S



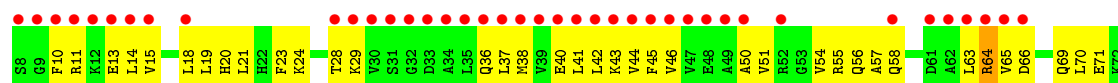
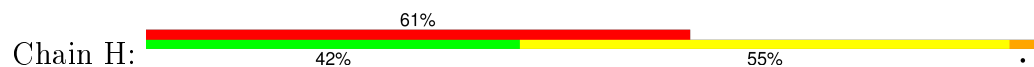
• Molecule 3: Centromere protein S



• Molecule 4: Centromere protein X

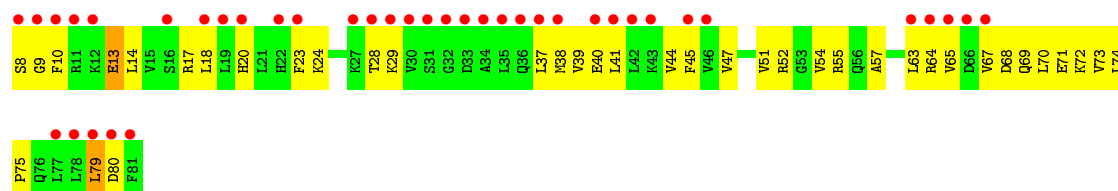


• Molecule 4: Centromere protein X

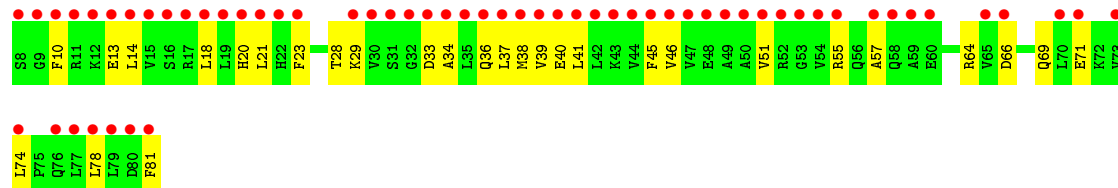
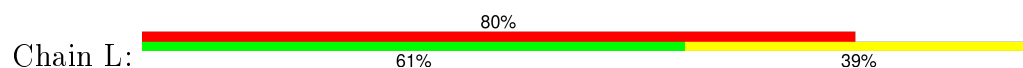


• Molecule 4: Centromere protein X

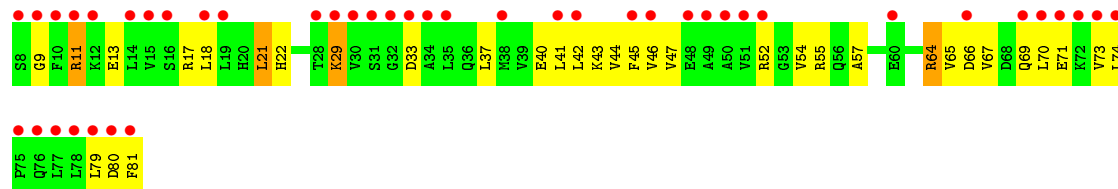




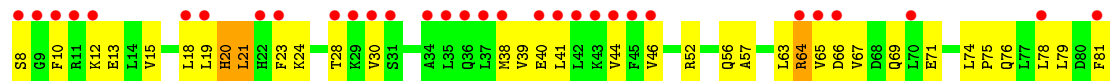
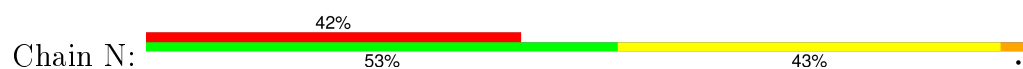
• Molecule 4: Centromere protein X



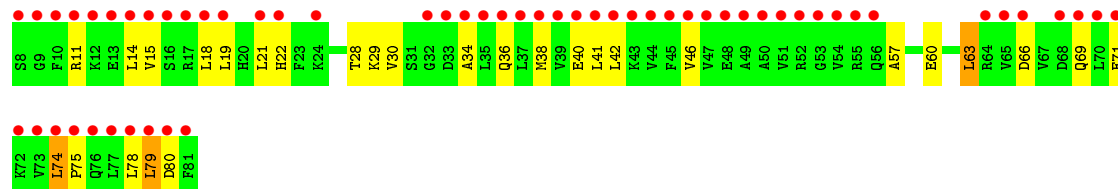
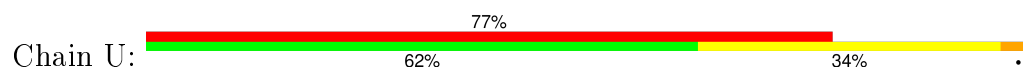
• Molecule 4: Centromere protein X



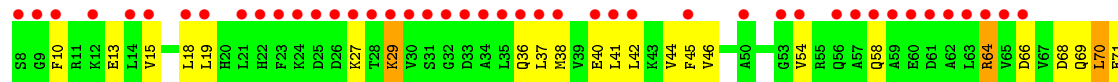
• Molecule 4: Centromere protein X

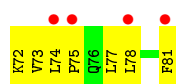


• Molecule 4: Centromere protein X

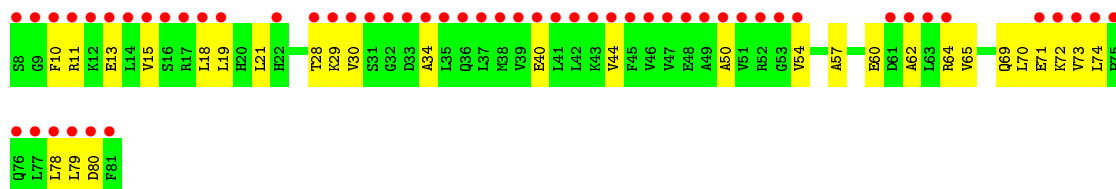
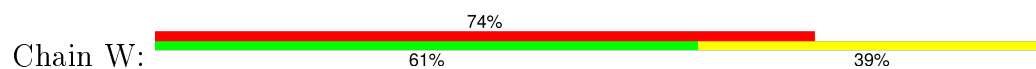


• Molecule 4: Centromere protein X

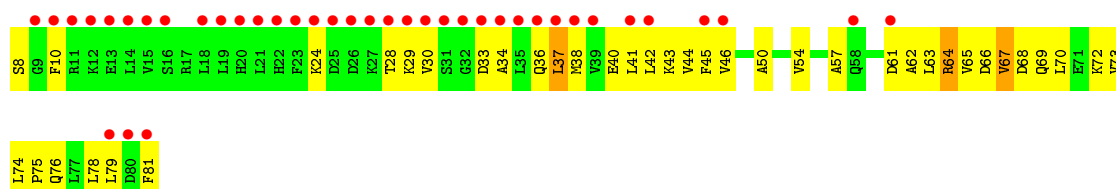




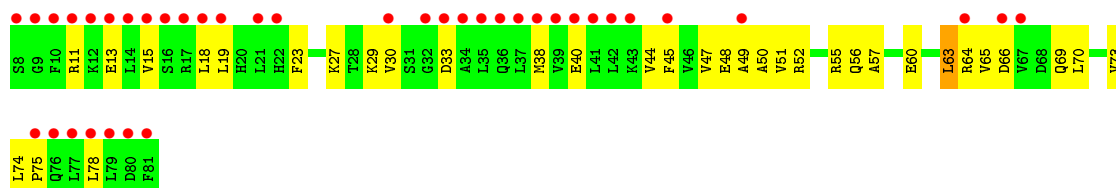
• Molecule 4: Centromere protein X



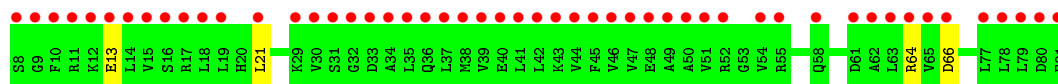
• Molecule 4: Centromere protein X



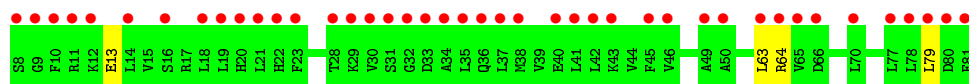
• Molecule 4: Centromere protein X



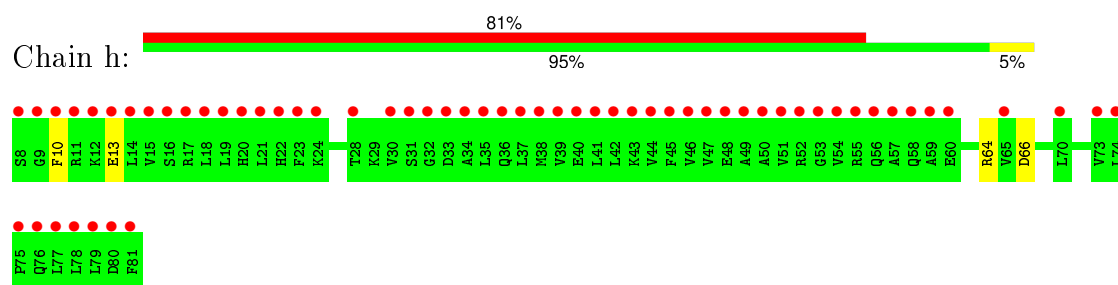
• Molecule 4: Centromere protein X



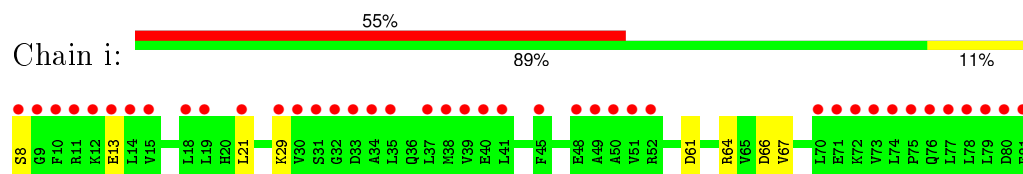
• Molecule 4: Centromere protein X



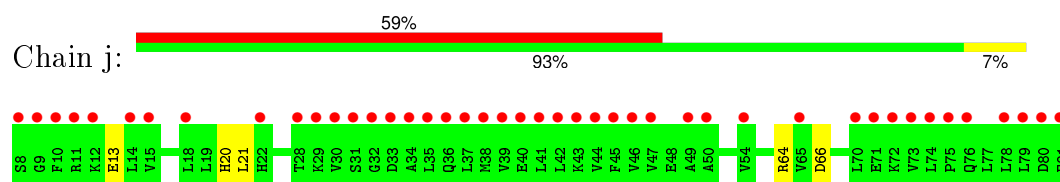
• Molecule 4: Centromere protein X



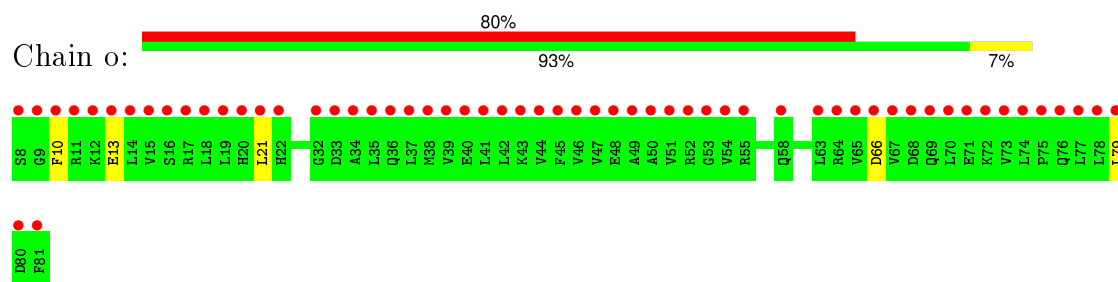
• Molecule 4: Centromere protein X



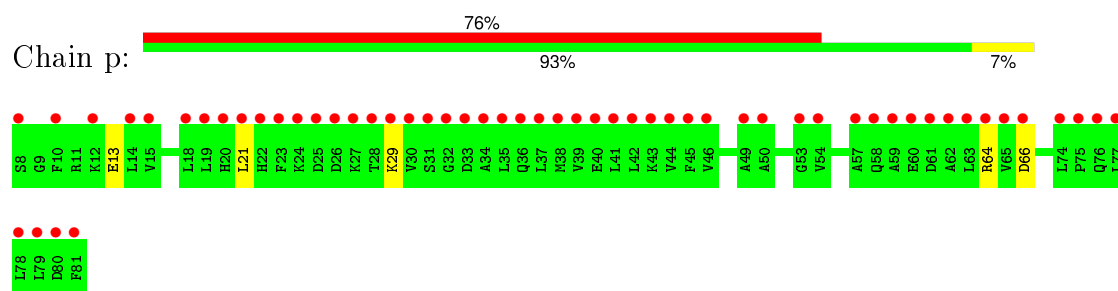
• Molecule 4: Centromere protein X



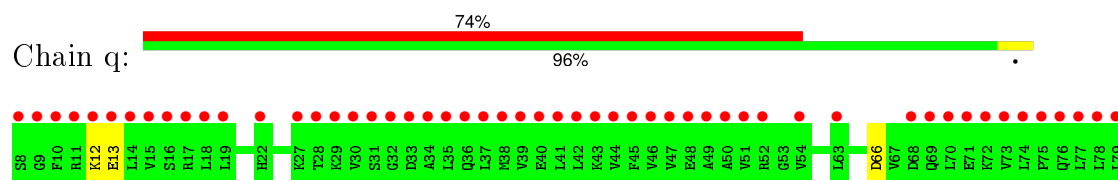
• Molecule 4: Centromere protein X



• Molecule 4: Centromere protein X

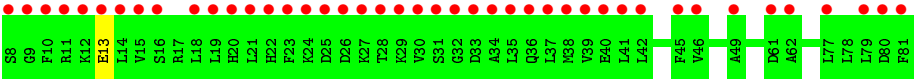


• Molecule 4: Centromere protein X





● Molecule 4: Centromere protein X



4 Data and refinement statistics

Property	Value	Source
Space group	P 3	Depositor
Cell constants a, b, c, α , β , γ	252.39Å 252.39Å 131.36Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	112.59 – 6.50 112.59 – 6.50	Depositor EDS
% Data completeness (in resolution range)	99.7 (112.59-6.50) 99.9 (112.59-6.50)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.02 (at 6.73Å)	Xtriage
Refinement program	REFMAC 5.7.0029	Depositor
R, R_{free}	0.266 , 0.295 0.284 , 0.319	Depositor DCC
R_{free} test set	816 reflections (4.47%)	DCC
Wilson B-factor (Å ²)	433.4	Xtriage
Anisotropy	0.287	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.37 , 446.5	EDS
Estimated twinning fraction	0.452 for H, K, L 0.032 for -K, -H, -L 0.022 for -H, -K, L 0.495 for K, H, -L 0.029 for -h,-k,l 0.067 for h,-h-k,-l 0.027 for -k,-h,-l	Xtriage
Reported twinning fraction	0.452 for H, K, L 0.032 for -K, -H, -L 0.022 for -H, -K, L 0.495 for K, H, -L	Depositor
L-test for twinning ²	$\langle L \rangle = 0.42$, $\langle L^2 \rangle = 0.24$	Xtriage
Outliers	0 of 18426 reflections	Xtriage
F_o, F_c correlation	0.83	EDS
Total number of atoms	31827	wwPDB-VP
Average B, all atoms (Å ²)	363.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 52.98 % of the origin peak, indicating pseudo translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo translational symmetry is equal to 4.5016e-05. The detected translational NCS is most likely*

¹ Intensities estimated from amplitudes.

² 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.

also responsible for the elevated intensity ratio.

5 Model quality ⓘ

5.1 Standard geometry ⓘ

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	E	0.38	0/623	0.88	0/958
1	O	0.46	1/623 (0.2%)	0.90	0/958
1	s	0.31	0/623	0.83	0/958
1	u	0.33	0/623	0.81	0/958
2	F	0.42	0/571	0.73	0/880
2	P	0.40	0/571	0.71	0/880
2	t	0.30	0/571	0.71	0/880
2	v	0.31	0/571	0.74	0/880
3	A	0.63	0/751	0.87	1/1007 (0.1%)
3	C	0.66	0/751	0.81	1/1007 (0.1%)
3	G	0.65	0/751	0.84	2/1007 (0.2%)
3	I	0.72	0/852	0.93	3/1137 (0.3%)
3	J	0.68	0/751	0.88	0/1007
3	K	0.63	0/852	0.82	1/1137 (0.1%)
3	Q	0.62	0/852	0.88	2/1137 (0.2%)
3	R	0.65	0/751	0.84	0/1007
3	S	0.65	0/751	0.85	1/1007 (0.1%)
3	T	0.56	0/852	0.77	2/1137 (0.2%)
3	Y	0.45	0/751	0.65	0/1007
3	a	0.42	0/751	0.64	0/1007
3	c	0.43	0/751	0.71	0/1007
3	e	0.42	0/852	0.70	0/1137
3	f	0.45	0/751	0.68	0/1007
3	g	0.44	0/852	0.63	0/1137
3	k	0.38	0/852	0.65	1/1137 (0.1%)
3	l	0.44	0/751	0.60	0/1007
3	m	0.43	0/751	0.63	0/1007
3	n	0.39	0/852	0.61	0/1137
4	B	0.62	0/605	0.85	1/813 (0.1%)
4	D	0.67	0/596	0.88	0/801
4	H	0.68	0/596	0.88	2/801 (0.2%)
4	L	0.62	0/605	0.79	0/813
4	M	0.68	0/605	0.81	0/813
4	N	0.71	0/605	0.95	0/813

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# Z >5	RMSZ	# Z >5
4	U	0.51	0/605	0.83	3/813 (0.4%)
4	V	0.63	0/605	0.93	1/813 (0.1%)
4	W	0.60	0/605	0.84	0/813
4	X	0.63	0/605	0.94	1/813 (0.1%)
4	Z	0.42	0/596	0.70	0/801
4	b	0.39	0/596	0.63	0/801
4	d	0.39	0/605	0.69	1/813 (0.1%)
4	h	0.39	0/605	0.65	0/813
4	i	0.42	0/605	0.71	0/813
4	j	0.39	0/605	0.62	0/813
4	o	0.37	0/605	0.68	1/813 (0.1%)
4	p	0.40	0/605	0.70	0/813
4	q	0.37	0/605	0.64	0/813
4	r	0.40	0/604	0.63	0/812
All	All	0.52	1/32667 (0.0%)	0.77	24/44743 (0.1%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
4	D	0	1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	O	24	DA	O3'-P	5.08	1.67	1.61

The worst 5 of 24 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	B	79	LEU	CA-CB-CG	7.28	132.03	115.30
3	T	53	LEU	CA-CB-CG	6.64	130.58	115.30
3	G	36	LEU	CA-CB-CG	6.51	130.28	115.30
4	H	79	LEU	CA-CB-CG	6.49	130.23	115.30
4	U	74	LEU	CA-CB-CG	6.37	129.96	115.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
4	D	63	LEU	Peptide

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	E	546	0	287	24	0
1	O	546	0	287	33	0
1	s	546	0	287	0	0
1	u	546	0	287	0	0
2	F	520	0	313	31	0
2	P	520	0	313	44	0
2	t	520	0	313	0	0
2	v	520	0	313	0	1
3	A	743	0	753	71	1
3	C	743	0	753	74	0
3	G	743	0	753	60	0
3	I	844	0	872	76	0
3	J	743	0	753	32	0
3	K	844	0	872	40	0
3	Q	844	0	872	50	0
3	R	743	0	753	33	0
3	S	743	0	753	68	0
3	T	844	0	872	37	0
3	Y	743	0	753	39	0
3	a	743	0	753	0	0
3	c	743	0	753	0	0
3	e	844	0	872	0	0
3	f	743	0	753	0	0
3	g	844	0	872	0	0
3	k	844	0	872	0	0
3	l	743	0	753	0	0
3	m	743	0	753	0	0
3	n	844	0	872	0	0
4	B	596	0	628	57	0
4	D	590	0	620	109	0
4	H	590	0	620	70	0
4	L	596	0	628	35	0
4	M	596	0	628	71	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	N	596	0	628	34	0
4	U	596	0	628	41	0
4	V	596	0	628	53	0
4	W	596	0	628	35	0
4	X	596	0	628	68	0
4	Z	590	0	620	45	0
4	b	590	0	620	0	0
4	d	596	0	628	0	0
4	h	596	0	628	0	0
4	i	596	0	628	0	0
4	j	596	0	628	0	0
4	o	596	0	628	0	0
4	p	596	0	628	0	0
4	q	596	0	628	0	0
4	r	595	0	628	0	0
All	All	31827	0	30940	831	1

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 19.

The worst 5 of 831 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:E:25:DA:C2	2:F:2:DT:O2	1.77	1.37
1:E:24:DA:C2	2:F:3:DT:O2	1.91	1.23
3:A:56:ARG:CD	4:X:79:LEU:HD13	1.70	1.19
3:K:30:LEU:HD11	4:L:10:PHE:HZ	1.05	1.17
3:A:56:ARG:HD2	4:X:79:LEU:HD13	1.24	1.15

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

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:v:4:DT:O2	3:A:15:TYR:OH[3_656]	2.19	0.01

5.3 Torsion angles

5.3.1 Protein backbone

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

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
3	A	91/105 (87%)	88 (97%)	3 (3%)	0	100	100
3	C	91/105 (87%)	88 (97%)	3 (3%)	0	100	100
3	G	91/105 (87%)	88 (97%)	3 (3%)	0	100	100
3	I	103/105 (98%)	99 (96%)	3 (3%)	1 (1%)	19	65
3	J	91/105 (87%)	89 (98%)	2 (2%)	0	100	100
3	K	103/105 (98%)	101 (98%)	2 (2%)	0	100	100
3	Q	103/105 (98%)	102 (99%)	1 (1%)	0	100	100
3	R	91/105 (87%)	90 (99%)	1 (1%)	0	100	100
3	S	91/105 (87%)	90 (99%)	1 (1%)	0	100	100
3	T	103/105 (98%)	102 (99%)	1 (1%)	0	100	100
3	Y	91/105 (87%)	89 (98%)	2 (2%)	0	100	100
3	a	91/105 (87%)	90 (99%)	1 (1%)	0	100	100
3	c	91/105 (87%)	91 (100%)	0	0	100	100
3	e	103/105 (98%)	101 (98%)	2 (2%)	0	100	100
3	f	91/105 (87%)	90 (99%)	1 (1%)	0	100	100
3	g	103/105 (98%)	103 (100%)	0	0	100	100
3	k	103/105 (98%)	102 (99%)	1 (1%)	0	100	100
3	l	91/105 (87%)	91 (100%)	0	0	100	100
3	m	91/105 (87%)	91 (100%)	0	0	100	100
3	n	103/105 (98%)	103 (100%)	0	0	100	100
4	B	73/74 (99%)	70 (96%)	3 (4%)	0	100	100
4	D	72/74 (97%)	65 (90%)	7 (10%)	0	100	100
4	H	72/74 (97%)	69 (96%)	3 (4%)	0	100	100
4	L	73/74 (99%)	72 (99%)	1 (1%)	0	100	100
4	M	73/74 (99%)	71 (97%)	2 (3%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
4	N	73/74 (99%)	70 (96%)	3 (4%)	0	100	100
4	U	73/74 (99%)	72 (99%)	1 (1%)	0	100	100
4	V	73/74 (99%)	72 (99%)	1 (1%)	0	100	100
4	W	73/74 (99%)	72 (99%)	1 (1%)	0	100	100
4	X	73/74 (99%)	73 (100%)	0	0	100	100
4	Z	72/74 (97%)	72 (100%)	0	0	100	100
4	b	72/74 (97%)	72 (100%)	0	0	100	100
4	d	73/74 (99%)	72 (99%)	1 (1%)	0	100	100
4	h	73/74 (99%)	73 (100%)	0	0	100	100
4	i	73/74 (99%)	73 (100%)	0	0	100	100
4	j	73/74 (99%)	73 (100%)	0	0	100	100
4	o	73/74 (99%)	73 (100%)	0	0	100	100
4	p	73/74 (99%)	73 (100%)	0	0	100	100
4	q	73/74 (99%)	73 (100%)	0	0	100	100
4	r	73/74 (99%)	73 (100%)	0	0	100	100
All	All	3372/3580 (94%)	3321 (98%)	50 (2%)	1 (0%)	100	100

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	I	117	LYS

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	
3	A	80/91 (88%)	71 (89%)	9 (11%)	7	33
3	C	80/91 (88%)	74 (92%)	6 (8%)	17	53
3	G	80/91 (88%)	71 (89%)	9 (11%)	7	33

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
3	I	91/91 (100%)	82 (90%)	9 (10%)	10	39
3	J	80/91 (88%)	70 (88%)	10 (12%)	6	30
3	K	91/91 (100%)	83 (91%)	8 (9%)	12	45
3	Q	91/91 (100%)	82 (90%)	9 (10%)	10	39
3	R	80/91 (88%)	71 (89%)	9 (11%)	7	33
3	S	80/91 (88%)	70 (88%)	10 (12%)	6	30
3	T	91/91 (100%)	82 (90%)	9 (10%)	10	39
3	Y	80/91 (88%)	73 (91%)	7 (9%)	12	45
3	a	80/91 (88%)	75 (94%)	5 (6%)	22	59
3	c	80/91 (88%)	73 (91%)	7 (9%)	12	45
3	e	91/91 (100%)	82 (90%)	9 (10%)	10	39
3	f	80/91 (88%)	73 (91%)	7 (9%)	12	45
3	g	91/91 (100%)	82 (90%)	9 (10%)	10	39
3	k	91/91 (100%)	82 (90%)	9 (10%)	10	39
3	l	80/91 (88%)	68 (85%)	12 (15%)	3	23
3	m	80/91 (88%)	72 (90%)	8 (10%)	9	38
3	n	91/91 (100%)	83 (91%)	8 (9%)	12	45
4	B	66/65 (102%)	63 (96%)	3 (4%)	34	69
4	D	65/65 (100%)	61 (94%)	4 (6%)	23	59
4	H	65/65 (100%)	62 (95%)	3 (5%)	33	68
4	L	66/65 (102%)	63 (96%)	3 (4%)	34	69
4	M	66/65 (102%)	59 (89%)	7 (11%)	8	36
4	N	66/65 (102%)	57 (86%)	9 (14%)	5	27
4	U	66/65 (102%)	64 (97%)	2 (3%)	48	77
4	V	66/65 (102%)	62 (94%)	4 (6%)	23	60
4	W	66/65 (102%)	63 (96%)	3 (4%)	34	69
4	X	66/65 (102%)	60 (91%)	6 (9%)	12	43
4	Z	65/65 (100%)	61 (94%)	4 (6%)	23	59
4	b	65/65 (100%)	61 (94%)	4 (6%)	23	59
4	d	66/65 (102%)	63 (96%)	3 (4%)	34	69
4	h	66/65 (102%)	62 (94%)	4 (6%)	23	60

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
4	i	66/65 (102%)	58 (88%)	8 (12%)	6	31
4	j	66/65 (102%)	61 (92%)	5 (8%)	16	53
4	o	66/65 (102%)	62 (94%)	4 (6%)	23	60
4	p	66/65 (102%)	61 (92%)	5 (8%)	16	53
4	q	66/65 (102%)	63 (96%)	3 (4%)	34	69
4	r	66/65 (102%)	65 (98%)	1 (2%)	72	88
All	All	3004/3120 (96%)	2750 (92%)	254 (8%)	13	48

5 of 254 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	T	87	ARG
4	Z	66	ASP
3	n	41	GLN
4	U	66	ASP
4	X	61	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 27 such sidechains are listed below:

Mol	Chain	Res	Type
3	S	71	HIS
4	W	58	GLN
3	m	60	ASN
4	V	20	HIS
3	I	17	GLN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

5.6 Ligand geometry [i](#)

There are no ligands in this entry.

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, 95th 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	E	26/26 (100%)	0.30	0 100 100	256, 341, 374, 396	0
1	O	26/26 (100%)	0.46	0 100 100	351, 422, 540, 558	0
1	s	26/26 (100%)	0.50	1 (3%) 44 42	369, 450, 514, 541	0
1	u	26/26 (100%)	0.25	1 (3%) 44 42	342, 398, 446, 459	0
2	F	26/26 (100%)	0.63	2 (7%) 16 19	259, 312, 432, 483	0
2	P	26/26 (100%)	0.33	1 (3%) 44 42	282, 344, 515, 538	0
2	t	26/26 (100%)	0.26	1 (3%) 44 42	330, 374, 525, 543	0
2	v	26/26 (100%)	0.59	2 (7%) 16 19	294, 363, 431, 536	0
3	A	93/105 (88%)	2.45	48 (51%) 0 3	266, 317, 378, 406	0
3	C	93/105 (88%)	3.74	54 (58%) 0 3	296, 396, 484, 503	0
3	G	93/105 (88%)	3.50	68 (73%) 0 2	247, 315, 394, 427	0
3	I	105/105 (100%)	2.54	51 (48%) 0 3	207, 283, 379, 427	0
3	J	93/105 (88%)	1.70	34 (36%) 0 4	199, 292, 365, 379	0
3	K	105/105 (100%)	3.29	67 (63%) 0 3	262, 371, 475, 499	0
3	Q	105/105 (100%)	2.12	47 (44%) 0 4	189, 279, 340, 390	0
3	R	93/105 (88%)	3.70	66 (70%) 0 2	226, 316, 399, 466	0
3	S	93/105 (88%)	1.66	36 (38%) 0 4	217, 290, 354, 409	0
3	T	105/105 (100%)	3.27	70 (66%) 0 2	302, 363, 420, 488	0
3	Y	93/105 (88%)	4.25	61 (65%) 0 3	369, 486, 567, 590	0
3	a	93/105 (88%)	3.69	63 (67%) 0 2	324, 412, 488, 512	0
3	c	93/105 (88%)	3.03	53 (56%) 0 3	276, 361, 423, 443	0
3	e	105/105 (100%)	2.93	60 (57%) 0 3	229, 352, 464, 508	0
3	f	93/105 (88%)	2.08	36 (38%) 0 4	291, 369, 455, 468	0
3	g	105/105 (100%)	3.59	73 (69%) 0 2	339, 447, 553, 590	0

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å ²)	Q<0.9
3	k	105/105 (100%)	2.84	65 (61%)	0	3	253, 358, 449, 496	0
3	l	93/105 (88%)	4.18	68 (73%)	0	2	296, 386, 471, 523	0
3	m	93/105 (88%)	2.10	41 (44%)	0	4	291, 395, 477, 507	0
3	n	105/105 (100%)	3.65	67 (63%)	0	3	341, 432, 499, 519	0
4	B	74/74 (100%)	2.31	39 (52%)	0	3	222, 294, 407, 468	0
4	D	74/74 (100%)	2.81	35 (47%)	0	4	278, 344, 472, 498	1 (1%)
4	H	74/74 (100%)	3.92	45 (60%)	0	3	224, 284, 388, 446	1 (1%)
4	L	74/74 (100%)	4.27	59 (79%)	0	2	248, 330, 418, 520	0
4	M	74/74 (100%)	2.85	43 (58%)	0	3	213, 286, 333, 352	0
4	N	74/74 (100%)	1.96	31 (41%)	0	4	207, 276, 395, 452	0
4	U	74/74 (100%)	4.23	57 (77%)	0	2	240, 340, 426, 473	0
4	V	74/74 (100%)	2.77	48 (64%)	0	3	211, 283, 347, 368	0
4	W	74/74 (100%)	3.35	55 (74%)	0	2	198, 290, 380, 413	0
4	X	74/74 (100%)	2.20	39 (52%)	0	3	221, 294, 375, 400	0
4	Z	74/74 (100%)	3.00	39 (52%)	0	3	318, 439, 504, 518	1 (1%)
4	b	74/74 (100%)	4.76	51 (68%)	0	2	306, 418, 506, 514	1 (1%)
4	d	74/74 (100%)	2.64	42 (56%)	0	3	276, 357, 411, 458	0
4	h	74/74 (100%)	5.66	60 (81%)	0	2	328, 442, 527, 591	0
4	i	74/74 (100%)	3.26	41 (55%)	0	3	247, 371, 423, 460	0
4	j	74/74 (100%)	2.73	44 (59%)	0	3	303, 400, 465, 497	0
4	o	74/74 (100%)	4.74	59 (79%)	0	2	320, 412, 507, 528	0
4	p	74/74 (100%)	3.26	56 (75%)	0	2	270, 346, 442, 475	0
4	q	74/74 (100%)	4.69	55 (74%)	0	2	308, 400, 458, 481	0
4	r	74/74 (100%)	3.32	43 (58%)	0	3	303, 360, 448, 467	0
All	All	3644/3788 (96%)	3.04	2077 (56%)	0	3	189, 357, 488, 591	4 (0%)

The worst 5 of 2077 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	Y	86	ALA	22.8
3	C	86	ALA	21.6
3	C	85	LEU	20.1
3	Y	88	ARG	18.6
4	h	9	GLY	17.9

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 carbohydrates in this entry.

6.4 Ligands [i](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.