wwPDB X-ray Structure Validation Summary Report (i)

Aug 20, 2020 – 03:32 PM BST

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
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

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

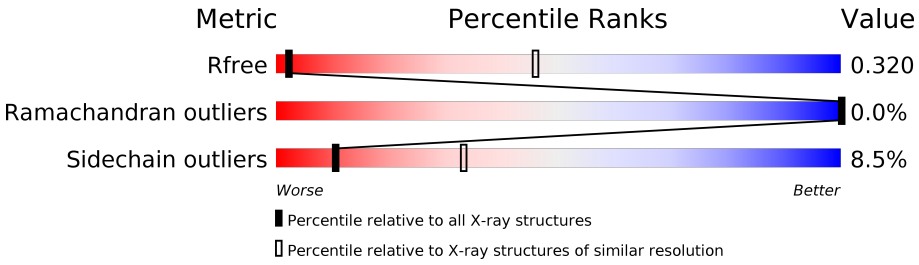
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.13
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation 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 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}	130704	1000 (9.00-3.90)
Ramachandran outliers	138981	1012 (9.00-3.88)
Sidechain outliers	138945	1010 (9.00-3.84)

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%

Mol	Chain	Length	Quality of chain
1	E	26	100%
1	O	26	96% .
1	s	26	100%
1	u	26	100%
2	F	26	100%
2	P	26	100%
2	t	26	100%
2	v	26	96% .
3	A	105	79% 9% . 11%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	C	105	82% 7% 11%
3	G	105	78% 10% 11%
3	I	105	90% 9%
3	J	105	79% 10% 11%
3	K	105	92% 7%
3	Q	105	91% 7%
3	R	105	80% 9% 11%
3	S	105	79% 9% 11%
3	T	105	90% 10%
3	Y	105	82% 7% 11%
3	a	105	84% 5% 11%
3	c	105	82% 7% 11%
3	e	105	91% 9%
3	f	105	82% 7% 11%
3	g	105	91% 9%
3	k	105	91% 8%
3	l	105	77% 11% 11%
3	m	105	81% 8% 11%
3	n	105	92% 8%
4	B	74	95% 5%
4	D	74	95%
4	H	74	93% 7%
4	L	74	96%
4	M	74	91% 9%
4	N	74	88% 12%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
4	U	74	95%
4	V	74	93% 7%
4	W	74	96%
4	X	74	91% 9%
4	Z	74	95% 5%
4	b	74	95% 5%
4	d	74	95% 5%
4	h	74	95% 5%
4	i	74	89% 11%
4	j	74	93% 7%
4	o	74	93% 7%
4	p	74	93% 7%
4	q	74	96%
4	r	74	99%

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...

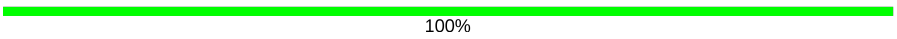
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	L	74	Total	C	N	O	S	0	1	0
			596	382	105	108	1			
4	M	74	Total	C	N	O	S	0	1	0
			596	382	105	108	1			
4	N	74	Total	C	N	O	S	0	1	0
			596	382	105	108	1			
4	U	74	Total	C	N	O	S	0	1	0
			596	382	105	108	1			
4	V	74	Total	C	N	O	S	0	1	0
			596	382	105	108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

3 Residue-property plots (i)

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

- Molecule 1: DNA (26-MER)

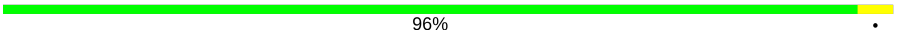
Chain E:



There are no outlier residues recorded for this chain.

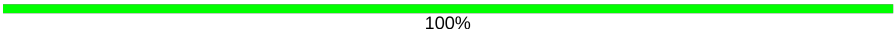
- Molecule 1: DNA (26-MER)

Chain O:



- Molecule 1: DNA (26-MER)

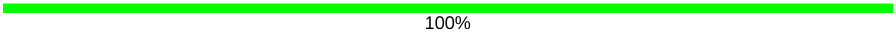
Chain s:



There are no outlier residues recorded for this chain.

- Molecule 1: DNA (26-MER)

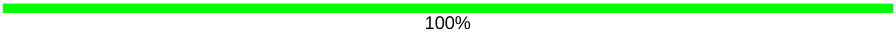
Chain u:



There are no outlier residues recorded for this chain.

- Molecule 2: DNA (26-MER)

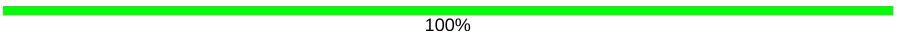
Chain F:



There are no outlier residues recorded for this chain.

- Molecule 2: DNA (26-MER)

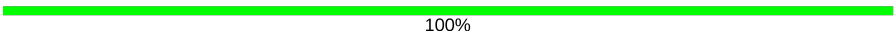
Chain P:



There are no outlier residues recorded for this chain.

- Molecule 2: DNA (26-MER)

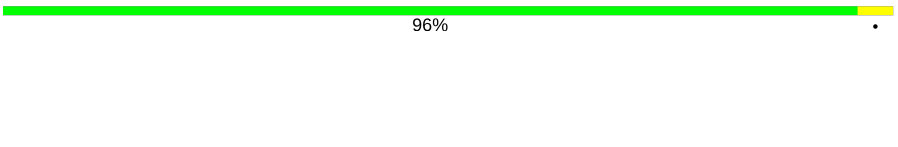
Chain t:



There are no outlier residues recorded for this chain.

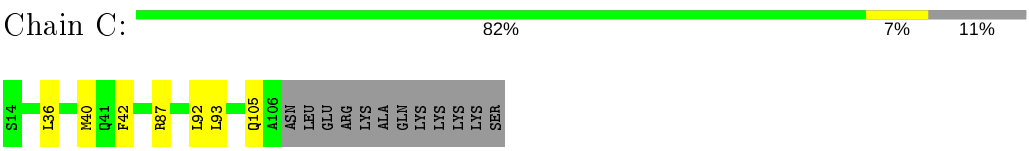
• Molecule 2: DNA (26-MER)

Chain v:



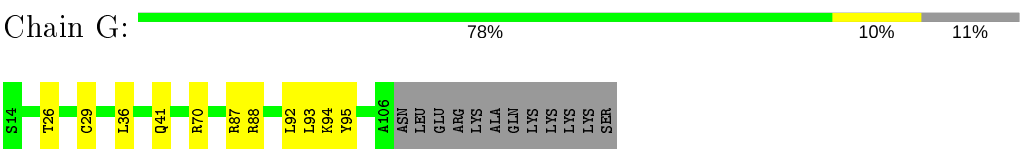
• Molecule 3: Centromere protein S

Chain C:



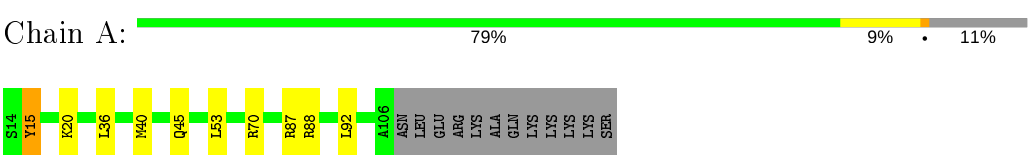
• Molecule 3: Centromere protein S

Chain G:



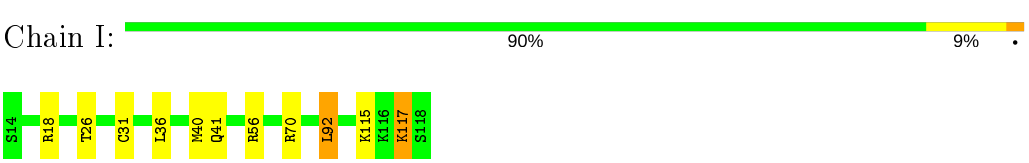
• Molecule 3: Centromere protein S

Chain A:



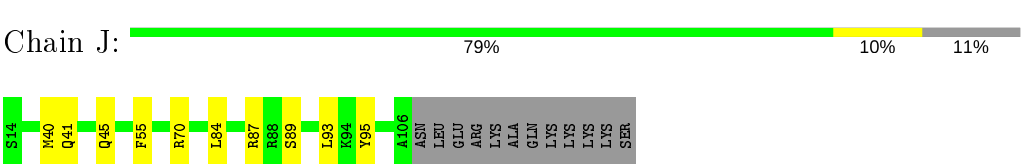
• Molecule 3: Centromere protein S

Chain I:



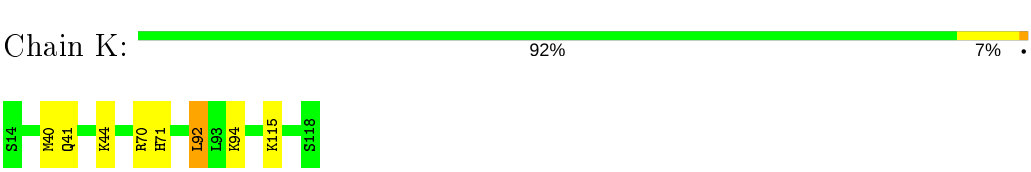
• Molecule 3: Centromere protein S

Chain J:



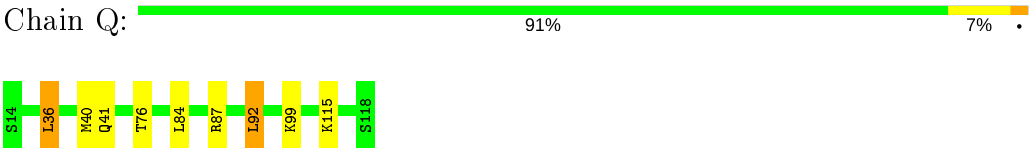
• Molecule 3: Centromere protein S

Chain K:



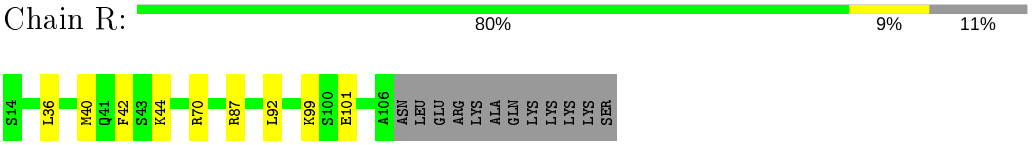
• Molecule 3: Centromere protein S

Chain Q:



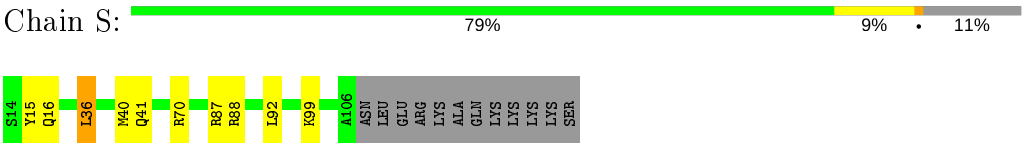
- Molecule 3: Centromere protein S

Chain R:



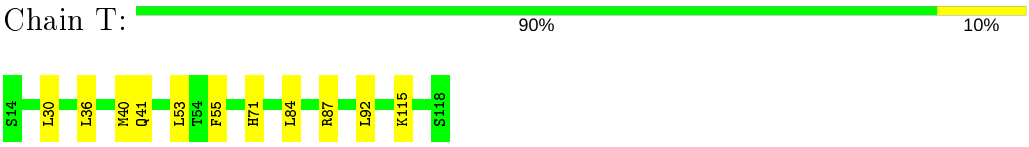
- Molecule 3: Centromere protein S

Chain S:



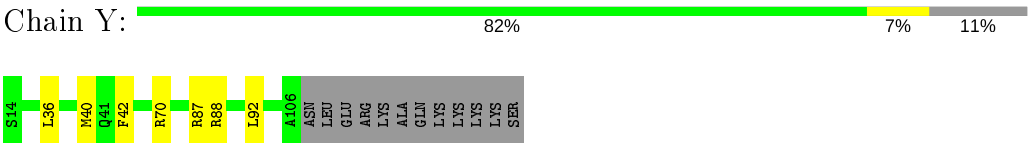
- Molecule 3: Centromere protein S

Chain T:



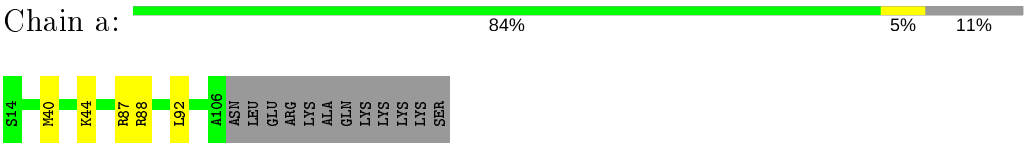
- Molecule 3: Centromere protein S

Chain Y:



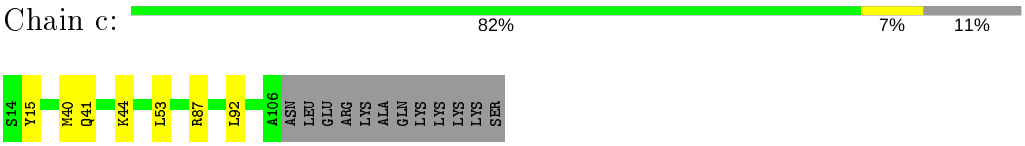
- Molecule 3: Centromere protein S

Chain a:



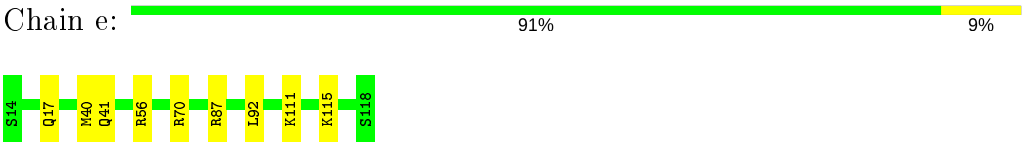
- Molecule 3: Centromere protein S

Chain c:



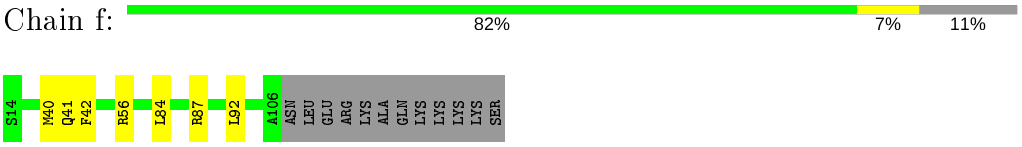
- Molecule 3: Centromere protein S

Chain e:



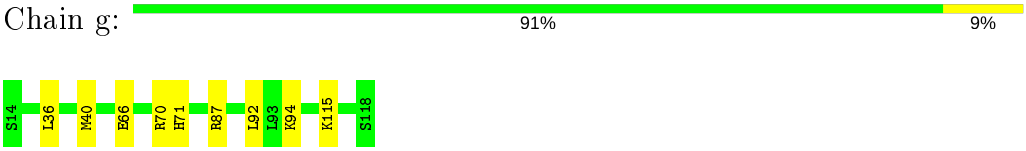
• Molecule 3: Centromere protein S

Chain f:



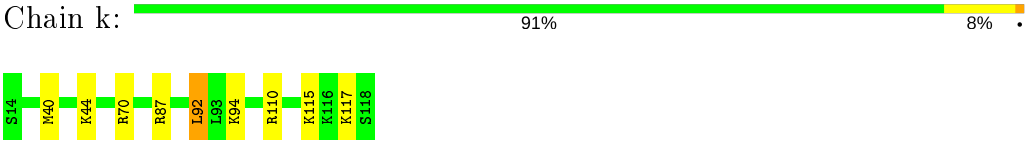
• Molecule 3: Centromere protein S

Chain g:



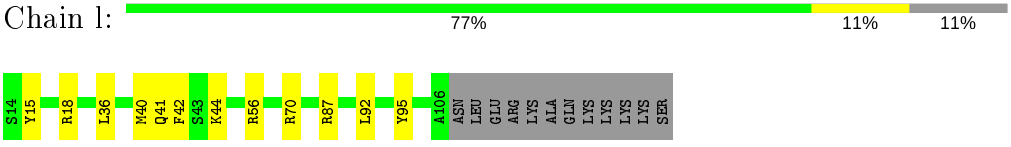
• Molecule 3: Centromere protein S

Chain k:



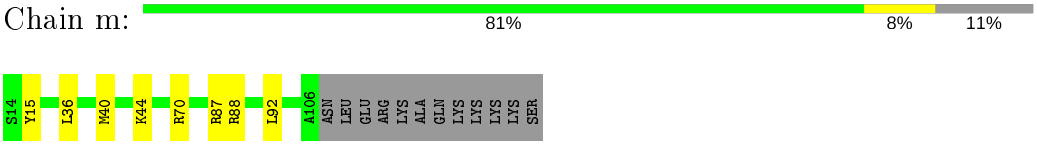
• Molecule 3: Centromere protein S

Chain l:



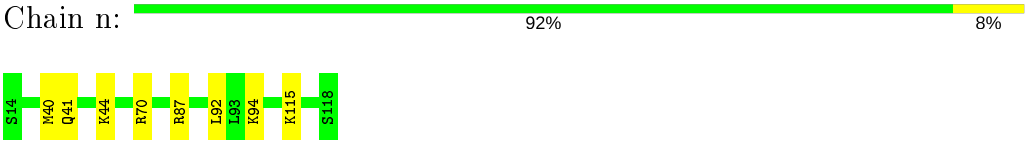
• Molecule 3: Centromere protein S

Chain m:



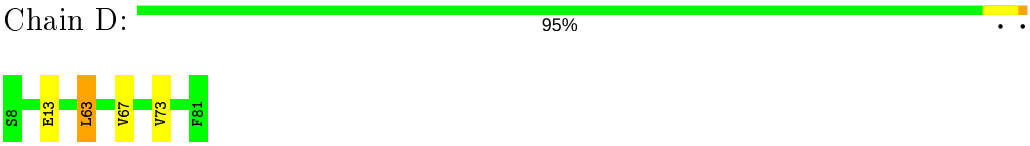
• Molecule 3: Centromere protein S

Chain n:



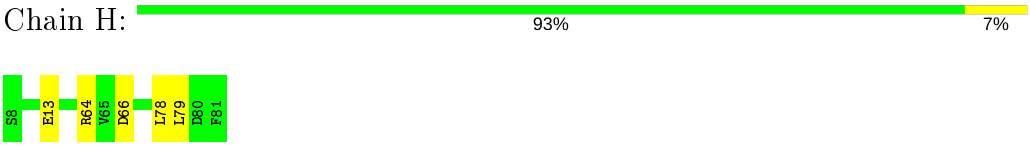
• Molecule 4: Centromere protein X

Chain D:



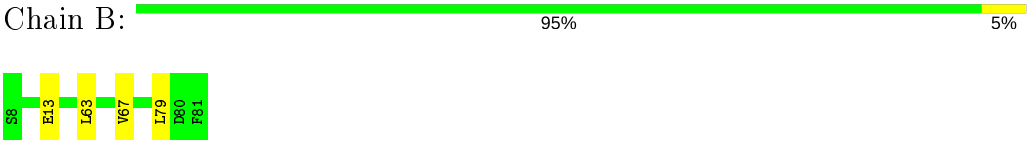
- Molecule 4: Centromere protein X

Chain H:



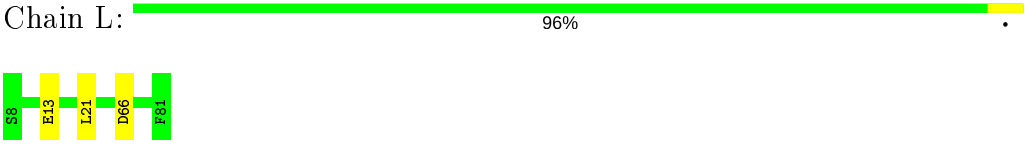
- Molecule 4: Centromere protein X

Chain B:



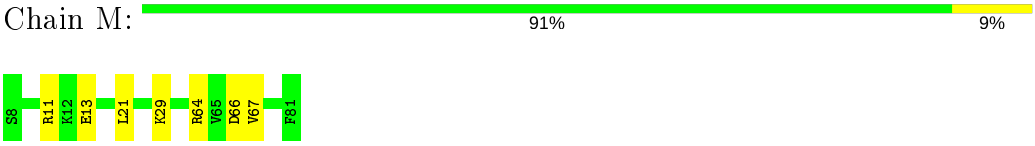
- Molecule 4: Centromere protein X

Chain L:



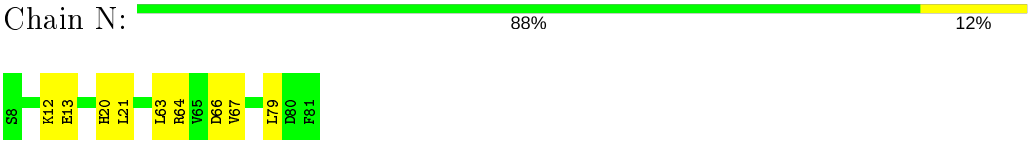
- Molecule 4: Centromere protein X

Chain M:



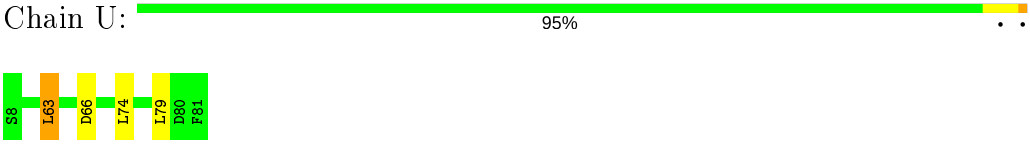
- Molecule 4: Centromere protein X

Chain N:



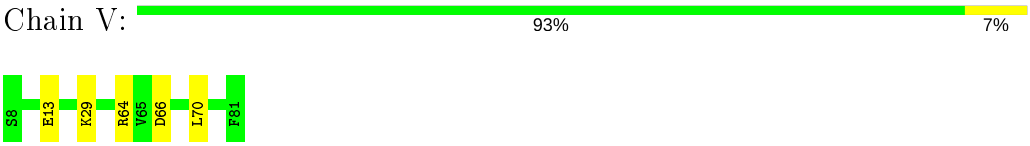
- Molecule 4: Centromere protein X

Chain U:



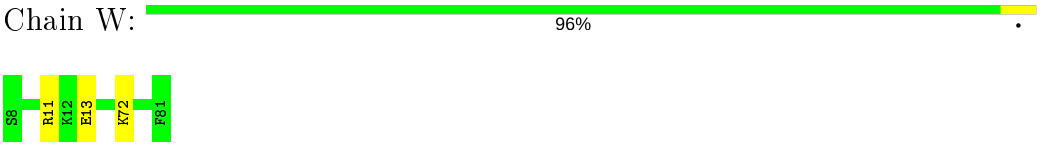
- Molecule 4: Centromere protein X

Chain V:



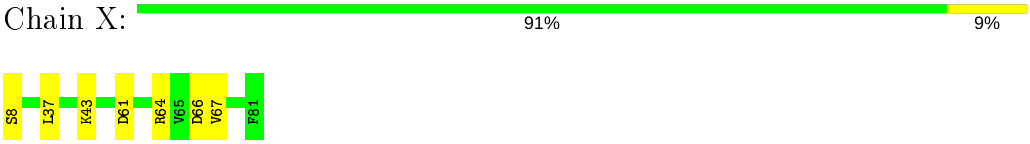
- Molecule 4: Centromere protein X

Chain W:



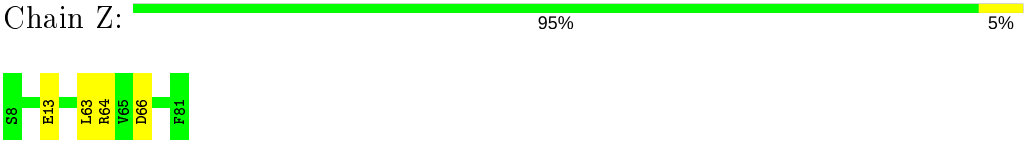
- Molecule 4: Centromere protein X

Chain X:



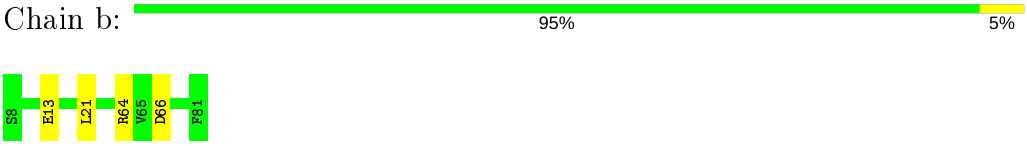
- Molecule 4: Centromere protein X

Chain Z:



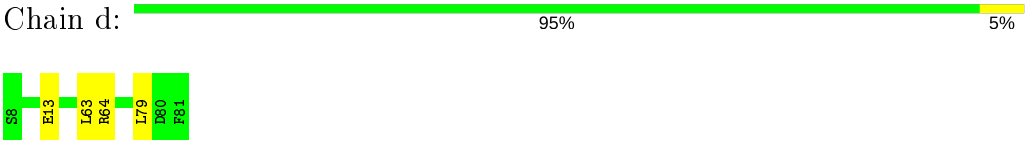
- Molecule 4: Centromere protein X

Chain b:



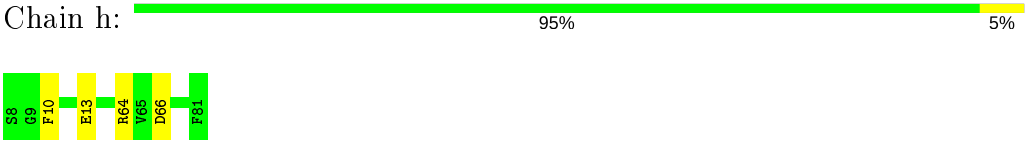
- Molecule 4: Centromere protein X

Chain d:



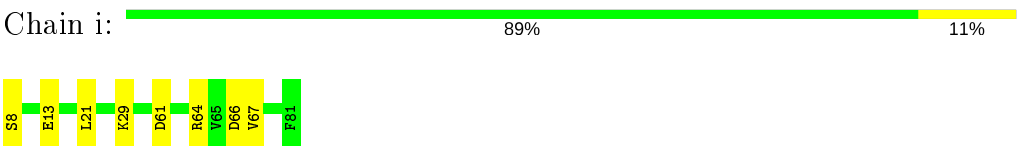
- Molecule 4: Centromere protein X

Chain h:



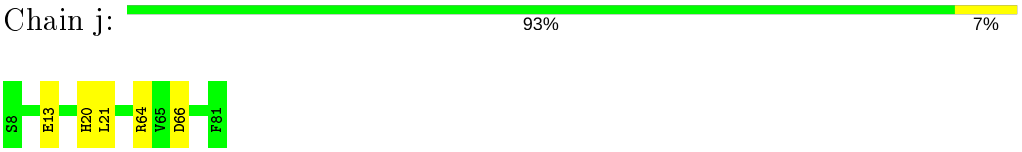
- Molecule 4: Centromere protein X

Chain i:



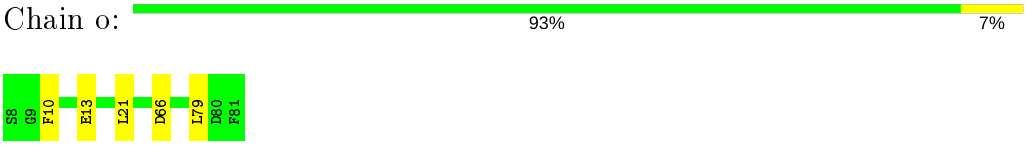
- Molecule 4: Centromere protein X

Chain j:



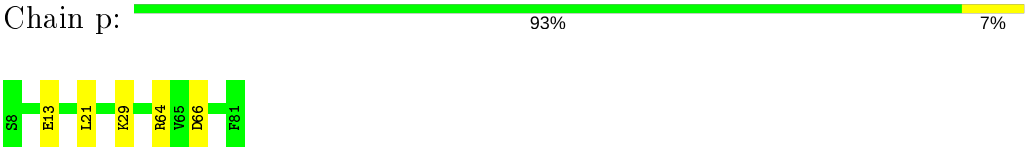
- Molecule 4: Centromere protein X

Chain o:



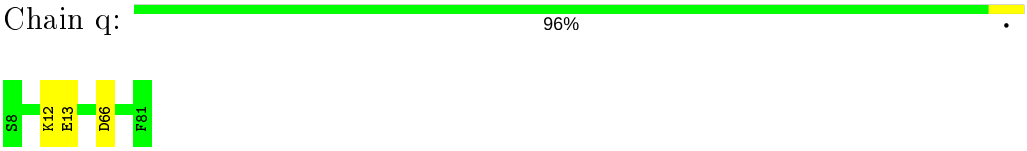
- Molecule 4: Centromere protein X

Chain p:



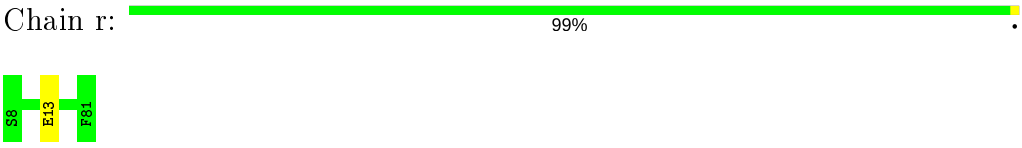
- Molecule 4: Centromere protein X

Chain q:



- Molecule 4: Centromere protein X

Chain r:



4 Data and refinement statistics (i)

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
$< I/\sigma(I) >$ [1]	2.02 (at 6.73Å)	Xtriage
Refinement program	REFMAC 5.7.0029	Depositor
R, R_{free}	0.266 , 0.295 0.286 , 0.320	Depositor DCC
R_{free} test set	826 reflections (4.48%)	wwPDB-VP
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
L-test for twinning[2]	$< \|L\| > = 0.42$, $< L^2 > = 0.24$	Xtriage
Estimated twinning fraction	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
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 also responsible for the elevated intensity ratio.*

[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

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(^{o})	Ideal(^{o})
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

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

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%)	15	54
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

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
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
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 (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
3	A	80/91 (88%)	71 (89%)	9 (11%)	6 21
3	C	80/91 (88%)	74 (92%)	6 (8%)	13 38
3	G	80/91 (88%)	71 (89%)	9 (11%)	6 21
3	I	91/91 (100%)	82 (90%)	9 (10%)	8 26
3	J	80/91 (88%)	70 (88%)	10 (12%)	4 19
3	K	91/91 (100%)	83 (91%)	8 (9%)	10 31
3	Q	91/91 (100%)	82 (90%)	9 (10%)	8 26
3	R	80/91 (88%)	71 (89%)	9 (11%)	6 21
3	S	80/91 (88%)	70 (88%)	10 (12%)	4 19
3	T	91/91 (100%)	82 (90%)	9 (10%)	8 26
3	Y	80/91 (88%)	73 (91%)	7 (9%)	10 31
3	a	80/91 (88%)	75 (94%)	5 (6%)	18 43
3	c	80/91 (88%)	73 (91%)	7 (9%)	10 31
3	e	91/91 (100%)	82 (90%)	9 (10%)	8 26
3	f	80/91 (88%)	73 (91%)	7 (9%)	10 31
3	g	91/91 (100%)	82 (90%)	9 (10%)	8 26
3	k	91/91 (100%)	82 (90%)	9 (10%)	8 26
3	l	80/91 (88%)	68 (85%)	12 (15%)	3 15
3	m	80/91 (88%)	72 (90%)	8 (10%)	7 26
3	n	91/91 (100%)	83 (91%)	8 (9%)	10 31
4	B	66/65 (102%)	63 (96%)	3 (4%)	27 52
4	D	65/65 (100%)	61 (94%)	4 (6%)	18 43
4	H	65/65 (100%)	62 (95%)	3 (5%)	27 52
4	L	66/65 (102%)	63 (96%)	3 (4%)	27 52
4	M	66/65 (102%)	59 (89%)	7 (11%)	6 24
4	N	66/65 (102%)	57 (86%)	9 (14%)	3 17
4	U	66/65 (102%)	64 (97%)	2 (3%)	41 63
4	V	66/65 (102%)	62 (94%)	4 (6%)	18 44
4	W	66/65 (102%)	63 (96%)	3 (4%)	27 52
4	X	66/65 (102%)	60 (91%)	6 (9%)	9 30

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
4	Z	65/65 (100%)	61 (94%)	4 (6%)	18 43
4	b	65/65 (100%)	61 (94%)	4 (6%)	18 43
4	d	66/65 (102%)	63 (96%)	3 (4%)	27 52
4	h	66/65 (102%)	62 (94%)	4 (6%)	18 44
4	i	66/65 (102%)	58 (88%)	8 (12%)	5 20
4	j	66/65 (102%)	61 (92%)	5 (8%)	13 37
4	o	66/65 (102%)	62 (94%)	4 (6%)	18 44
4	p	66/65 (102%)	61 (92%)	5 (8%)	13 37
4	q	66/65 (102%)	63 (96%)	3 (4%)	27 52
4	r	66/65 (102%)	65 (98%)	1 (2%)	65 80
All	All	3004/3120 (96%)	2750 (92%)	254 (8%)	10 33

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 monosaccharides in this entry.

5.6 Ligand geometry

There are no ligands in this entry.

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

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

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

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

6.3 Carbohydrates

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

6.4 Ligands

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

6.5 Other polymers

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