



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 07:01 PM GMT

PDB ID : 4NDY
Title : Human MHF1-MHF2 DNA complex
Authors : Zhao, Q.; Saro, D.; Sachpatzidis, A.; Sung, P.; Xiong, Y.
Deposited on : 2013-10-28
Resolution : 7.00 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

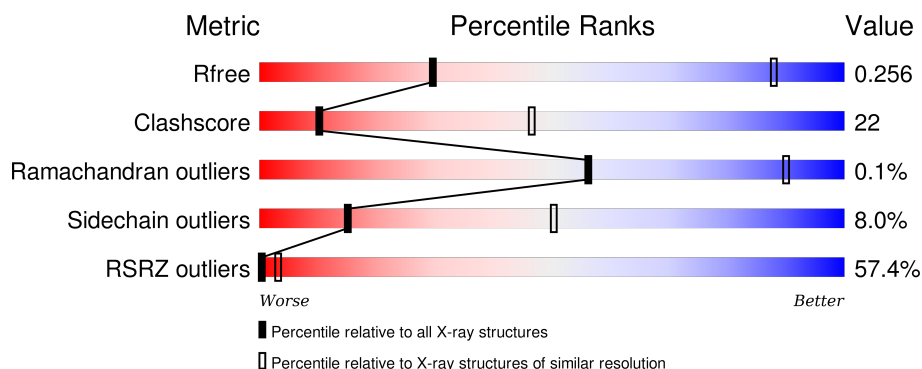
1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 7.00 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	1014 (9.50-3.66)
Clashscore	102246	1063 (10.00-3.70)
Ramachandran outliers	100387	1035 (9.50-3.66)
Sidechain outliers	100360	1005 (9.50-3.66)
RSRZ outliers	91569	1013 (9.50-3.66)

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>8%</div> <div>31%</div> <div>69%</div> </div>
1	O	26	<div> <div>12%</div> <div>27%</div> <div>73%</div> </div>
2	F	26	<div> <div>54%</div> <div>27%</div> <div>73%</div> </div>
2	P	26	<div> <div>42%</div> <div>31%</div> <div>69%</div> </div>
3	A	105	<div> <div>30%</div> <div>62%</div> <div>21%</div> <div>5%</div> <div>11%</div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
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	
4	B	74	
4	D	74	
4	H	74	
4	L	74	
4	M	74	
4	N	74	
4	U	74	
4	V	74	
4	W	74	
4	X	74	

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 15914 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			

- 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			

- 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			
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			

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
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			

There are 140 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	39	ALA	GLU	CONFLICT	UNP Q8N2Z9
C	106	ALA	-	EXPRESSION TAG	UNP Q8N2Z9
C	107	ASN	-	EXPRESSION TAG	UNP Q8N2Z9
C	108	LEU	-	EXPRESSION TAG	UNP Q8N2Z9
C	109	GLU	-	EXPRESSION TAG	UNP Q8N2Z9
C	110	ARG	-	EXPRESSION TAG	UNP Q8N2Z9
C	111	LYS	-	EXPRESSION TAG	UNP Q8N2Z9
C	112	ALA	-	EXPRESSION TAG	UNP Q8N2Z9
C	113	GLN	-	EXPRESSION TAG	UNP Q8N2Z9
C	114	LYS	-	EXPRESSION TAG	UNP Q8N2Z9
C	115	LYS	-	EXPRESSION TAG	UNP Q8N2Z9
C	116	LYS	-	EXPRESSION TAG	UNP Q8N2Z9
C	117	LYS	-	EXPRESSION TAG	UNP Q8N2Z9
C	118	SER	-	EXPRESSION TAG	UNP Q8N2Z9
G	39	ALA	GLU	CONFLICT	UNP Q8N2Z9
G	106	ALA	-	EXPRESSION TAG	UNP Q8N2Z9
G	107	ASN	-	EXPRESSION TAG	UNP Q8N2Z9
G	108	LEU	-	EXPRESSION TAG	UNP Q8N2Z9
G	109	GLU	-	EXPRESSION TAG	UNP Q8N2Z9
G	110	ARG	-	EXPRESSION TAG	UNP Q8N2Z9
G	111	LYS	-	EXPRESSION TAG	UNP Q8N2Z9
G	112	ALA	-	EXPRESSION TAG	UNP Q8N2Z9
G	113	GLN	-	EXPRESSION TAG	UNP Q8N2Z9
G	114	LYS	-	EXPRESSION TAG	UNP Q8N2Z9
G	115	LYS	-	EXPRESSION TAG	UNP Q8N2Z9
G	116	LYS	-	EXPRESSION TAG	UNP Q8N2Z9
G	117	LYS	-	EXPRESSION TAG	UNP Q8N2Z9
G	118	SER	-	EXPRESSION TAG	UNP Q8N2Z9
A	39	ALA	GLU	CONFLICT	UNP Q8N2Z9
A	106	ALA	-	EXPRESSION TAG	UNP Q8N2Z9
A	107	ASN	-	EXPRESSION TAG	UNP Q8N2Z9
A	108	LEU	-	EXPRESSION TAG	UNP Q8N2Z9
A	109	GLU	-	EXPRESSION TAG	UNP Q8N2Z9
A	110	ARG	-	EXPRESSION TAG	UNP Q8N2Z9

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	111	LYS	-	EXPRESSION TAG	UNP Q8N2Z9
A	112	ALA	-	EXPRESSION TAG	UNP Q8N2Z9
A	113	GLN	-	EXPRESSION TAG	UNP Q8N2Z9
A	114	LYS	-	EXPRESSION TAG	UNP Q8N2Z9
A	115	LYS	-	EXPRESSION TAG	UNP Q8N2Z9
A	116	LYS	-	EXPRESSION TAG	UNP Q8N2Z9
A	117	LYS	-	EXPRESSION TAG	UNP Q8N2Z9
A	118	SER	-	EXPRESSION TAG	UNP Q8N2Z9
I	39	ALA	GLU	CONFLICT	UNP Q8N2Z9
I	106	ALA	-	EXPRESSION TAG	UNP Q8N2Z9
I	107	ASN	-	EXPRESSION TAG	UNP Q8N2Z9
I	108	LEU	-	EXPRESSION TAG	UNP Q8N2Z9
I	109	GLU	-	EXPRESSION TAG	UNP Q8N2Z9
I	110	ARG	-	EXPRESSION TAG	UNP Q8N2Z9
I	111	LYS	-	EXPRESSION TAG	UNP Q8N2Z9
I	112	ALA	-	EXPRESSION TAG	UNP Q8N2Z9
I	113	GLN	-	EXPRESSION TAG	UNP Q8N2Z9
I	114	LYS	-	EXPRESSION TAG	UNP Q8N2Z9
I	115	LYS	-	EXPRESSION TAG	UNP Q8N2Z9
I	116	LYS	-	EXPRESSION TAG	UNP Q8N2Z9
I	117	LYS	-	EXPRESSION TAG	UNP Q8N2Z9
I	118	SER	-	EXPRESSION TAG	UNP Q8N2Z9
J	39	ALA	GLU	CONFLICT	UNP Q8N2Z9
J	106	ALA	-	EXPRESSION TAG	UNP Q8N2Z9
J	107	ASN	-	EXPRESSION TAG	UNP Q8N2Z9
J	108	LEU	-	EXPRESSION TAG	UNP Q8N2Z9
J	109	GLU	-	EXPRESSION TAG	UNP Q8N2Z9
J	110	ARG	-	EXPRESSION TAG	UNP Q8N2Z9
J	111	LYS	-	EXPRESSION TAG	UNP Q8N2Z9
J	112	ALA	-	EXPRESSION TAG	UNP Q8N2Z9
J	113	GLN	-	EXPRESSION TAG	UNP Q8N2Z9
J	114	LYS	-	EXPRESSION TAG	UNP Q8N2Z9
J	115	LYS	-	EXPRESSION TAG	UNP Q8N2Z9
J	116	LYS	-	EXPRESSION TAG	UNP Q8N2Z9
J	117	LYS	-	EXPRESSION TAG	UNP Q8N2Z9
J	118	SER	-	EXPRESSION TAG	UNP Q8N2Z9
K	39	ALA	GLU	CONFLICT	UNP Q8N2Z9
K	106	ALA	-	EXPRESSION TAG	UNP Q8N2Z9
K	107	ASN	-	EXPRESSION TAG	UNP Q8N2Z9
K	108	LEU	-	EXPRESSION TAG	UNP Q8N2Z9
K	109	GLU	-	EXPRESSION TAG	UNP Q8N2Z9
K	110	ARG	-	EXPRESSION TAG	UNP Q8N2Z9

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
K	111	LYS	-	EXPRESSION TAG	UNP Q8N2Z9
K	112	ALA	-	EXPRESSION TAG	UNP Q8N2Z9
K	113	GLN	-	EXPRESSION TAG	UNP Q8N2Z9
K	114	LYS	-	EXPRESSION TAG	UNP Q8N2Z9
K	115	LYS	-	EXPRESSION TAG	UNP Q8N2Z9
K	116	LYS	-	EXPRESSION TAG	UNP Q8N2Z9
K	117	LYS	-	EXPRESSION TAG	UNP Q8N2Z9
K	118	SER	-	EXPRESSION TAG	UNP Q8N2Z9
Q	39	ALA	GLU	CONFLICT	UNP Q8N2Z9
Q	106	ALA	-	EXPRESSION TAG	UNP Q8N2Z9
Q	107	ASN	-	EXPRESSION TAG	UNP Q8N2Z9
Q	108	LEU	-	EXPRESSION TAG	UNP Q8N2Z9
Q	109	GLU	-	EXPRESSION TAG	UNP Q8N2Z9
Q	110	ARG	-	EXPRESSION TAG	UNP Q8N2Z9
Q	111	LYS	-	EXPRESSION TAG	UNP Q8N2Z9
Q	112	ALA	-	EXPRESSION TAG	UNP Q8N2Z9
Q	113	GLN	-	EXPRESSION TAG	UNP Q8N2Z9
Q	114	LYS	-	EXPRESSION TAG	UNP Q8N2Z9
Q	115	LYS	-	EXPRESSION TAG	UNP Q8N2Z9
Q	116	LYS	-	EXPRESSION TAG	UNP Q8N2Z9
Q	117	LYS	-	EXPRESSION TAG	UNP Q8N2Z9
Q	118	SER	-	EXPRESSION TAG	UNP Q8N2Z9
R	39	ALA	GLU	CONFLICT	UNP Q8N2Z9
R	106	ALA	-	EXPRESSION TAG	UNP Q8N2Z9
R	107	ASN	-	EXPRESSION TAG	UNP Q8N2Z9
R	108	LEU	-	EXPRESSION TAG	UNP Q8N2Z9
R	109	GLU	-	EXPRESSION TAG	UNP Q8N2Z9
R	110	ARG	-	EXPRESSION TAG	UNP Q8N2Z9
R	111	LYS	-	EXPRESSION TAG	UNP Q8N2Z9
R	112	ALA	-	EXPRESSION TAG	UNP Q8N2Z9
R	113	GLN	-	EXPRESSION TAG	UNP Q8N2Z9
R	114	LYS	-	EXPRESSION TAG	UNP Q8N2Z9
R	115	LYS	-	EXPRESSION TAG	UNP Q8N2Z9
R	116	LYS	-	EXPRESSION TAG	UNP Q8N2Z9
R	117	LYS	-	EXPRESSION TAG	UNP Q8N2Z9
R	118	SER	-	EXPRESSION TAG	UNP Q8N2Z9
S	39	ALA	GLU	CONFLICT	UNP Q8N2Z9
S	106	ALA	-	EXPRESSION TAG	UNP Q8N2Z9
S	107	ASN	-	EXPRESSION TAG	UNP Q8N2Z9
S	108	LEU	-	EXPRESSION TAG	UNP Q8N2Z9
S	109	GLU	-	EXPRESSION TAG	UNP Q8N2Z9
S	110	ARG	-	EXPRESSION TAG	UNP Q8N2Z9

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
S	111	LYS	-	EXPRESSION TAG	UNP Q8N2Z9
S	112	ALA	-	EXPRESSION TAG	UNP Q8N2Z9
S	113	GLN	-	EXPRESSION TAG	UNP Q8N2Z9
S	114	LYS	-	EXPRESSION TAG	UNP Q8N2Z9
S	115	LYS	-	EXPRESSION TAG	UNP Q8N2Z9
S	116	LYS	-	EXPRESSION TAG	UNP Q8N2Z9
S	117	LYS	-	EXPRESSION TAG	UNP Q8N2Z9
S	118	SER	-	EXPRESSION TAG	UNP Q8N2Z9
T	39	ALA	GLU	CONFLICT	UNP Q8N2Z9
T	106	ALA	-	EXPRESSION TAG	UNP Q8N2Z9
T	107	ASN	-	EXPRESSION TAG	UNP Q8N2Z9
T	108	LEU	-	EXPRESSION TAG	UNP Q8N2Z9
T	109	GLU	-	EXPRESSION TAG	UNP Q8N2Z9
T	110	ARG	-	EXPRESSION TAG	UNP Q8N2Z9
T	111	LYS	-	EXPRESSION TAG	UNP Q8N2Z9
T	112	ALA	-	EXPRESSION TAG	UNP Q8N2Z9
T	113	GLN	-	EXPRESSION TAG	UNP Q8N2Z9
T	114	LYS	-	EXPRESSION TAG	UNP Q8N2Z9
T	115	LYS	-	EXPRESSION TAG	UNP Q8N2Z9
T	116	LYS	-	EXPRESSION TAG	UNP Q8N2Z9
T	117	LYS	-	EXPRESSION TAG	UNP Q8N2Z9
T	118	SER	-	EXPRESSION TAG	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			
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			

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
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			

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 ($\text{RSRZ} > 2$). 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.

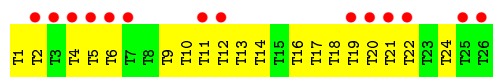
- Molecule 1: DNA (26-MER)



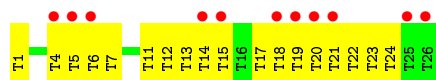
- Molecule 1: DNA (26-MER)



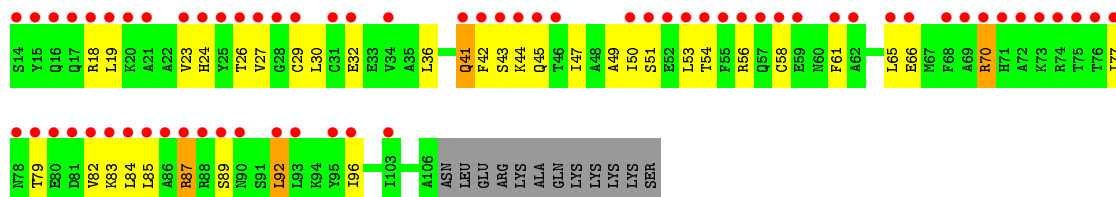
- Molecule 2: DNA (26-MER)



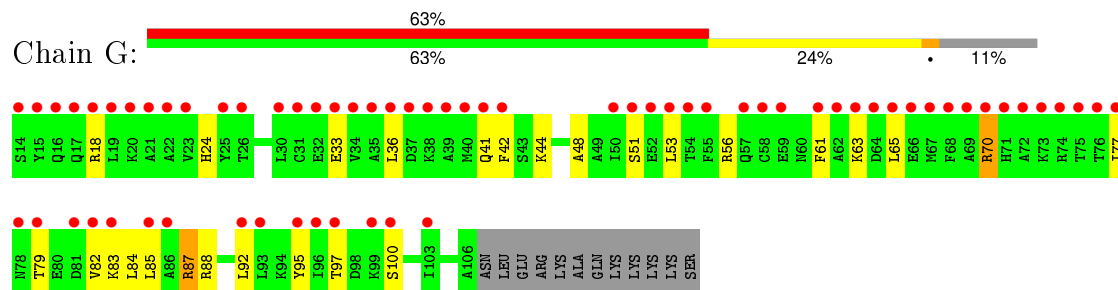
- Molecule 2: DNA (26-MER)



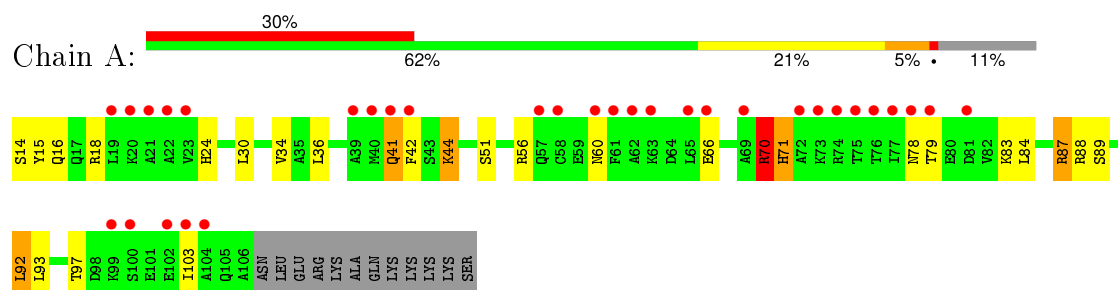
- Molecule 3: Centromere protein S



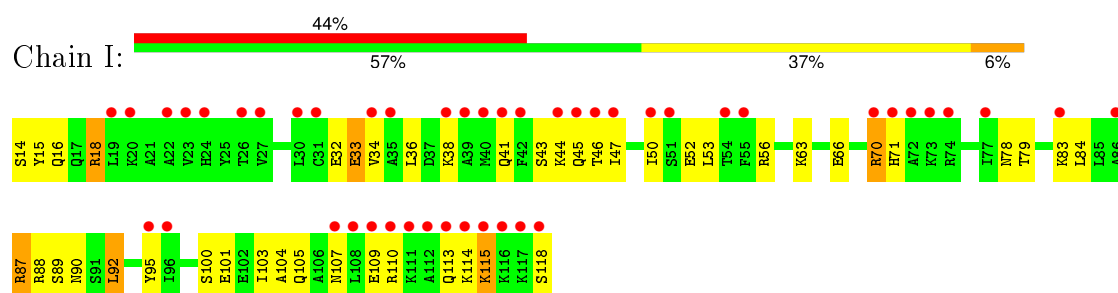
- Molecule 3: Centromere protein S



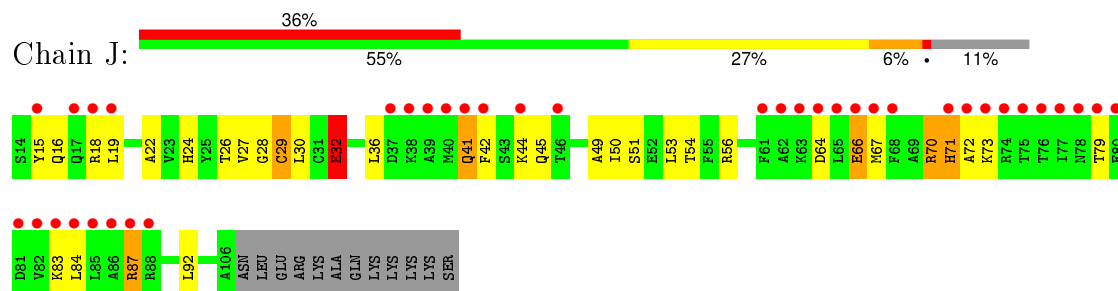
- Molecule 3: Centromere protein S



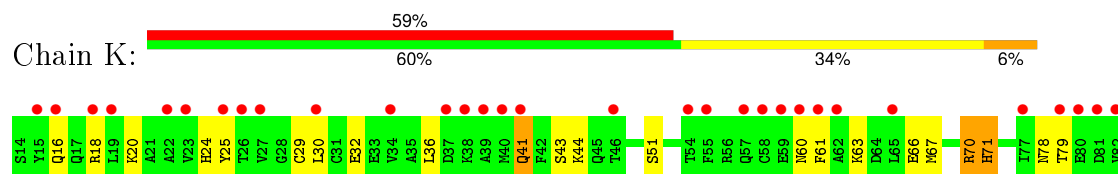
- Molecule 3: Centromere protein S

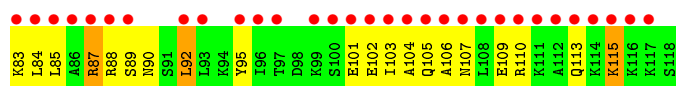


- Molecule 3: Centromere protein S

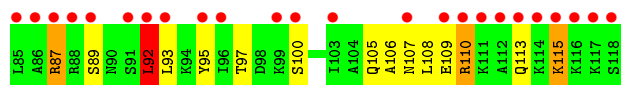
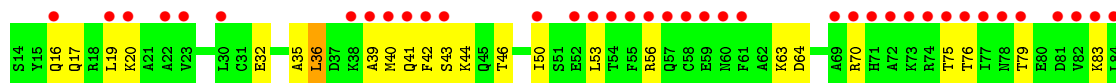


- Molecule 3: Centromere protein S

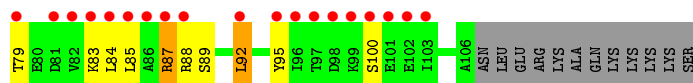
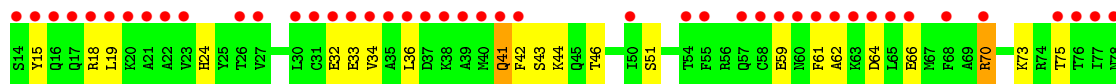




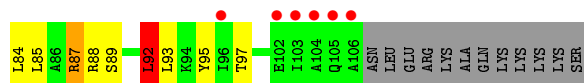
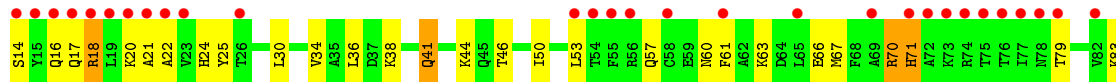
• Molecule 3: Centromere protein S



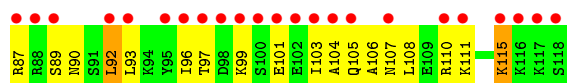
• Molecule 3: Centromere protein S



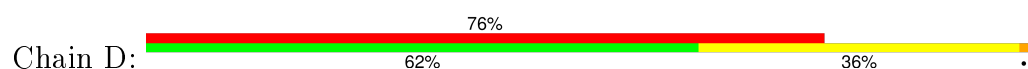
• Molecule 3: Centromere protein S

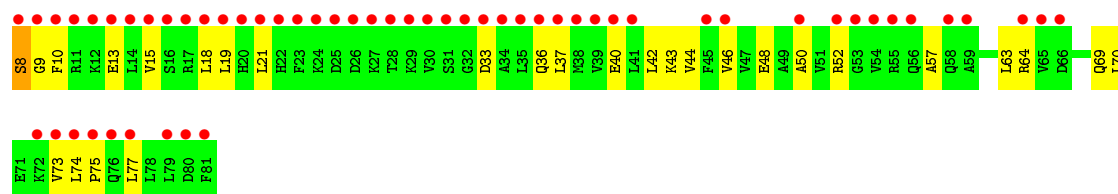


• Molecule 3: Centromere protein S

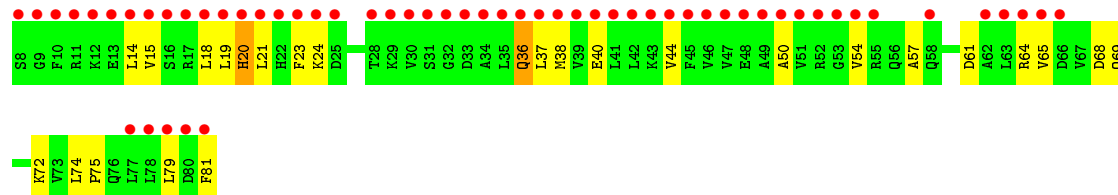
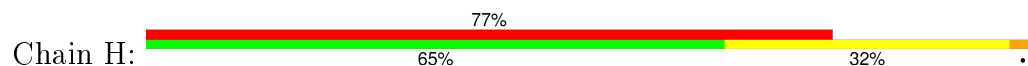


• Molecule 4: Centromere protein X

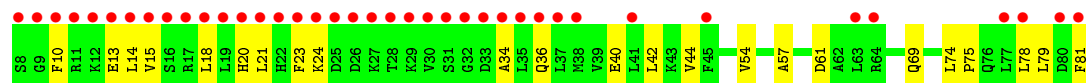




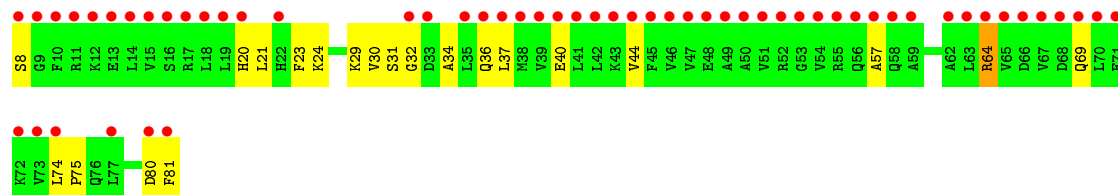
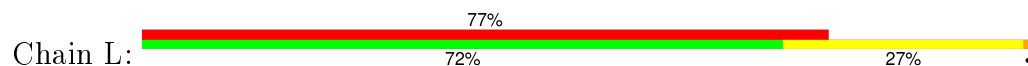
• Molecule 4: Centromere protein X



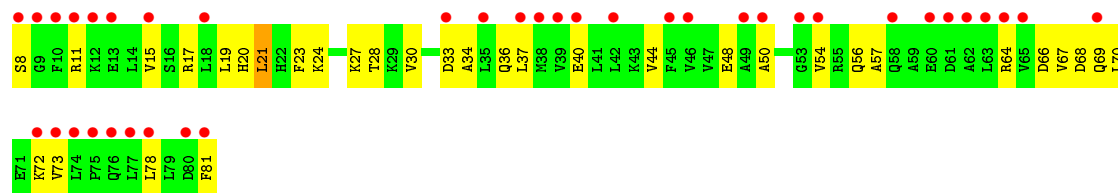
• Molecule 4: Centromere protein X



• Molecule 4: Centromere protein X

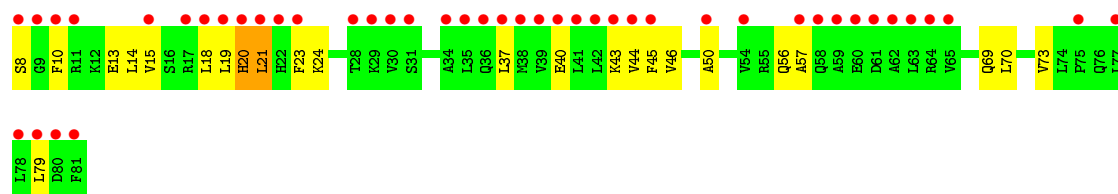


• Molecule 4: Centromere protein X

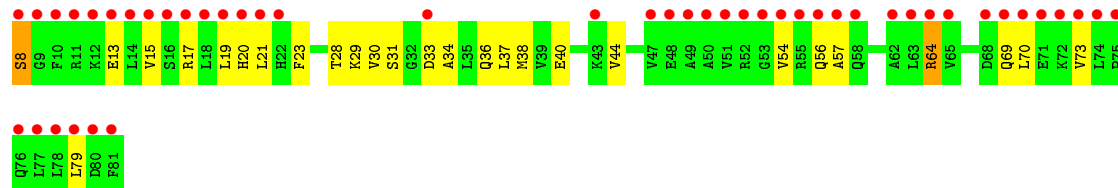


• Molecule 4: Centromere protein X

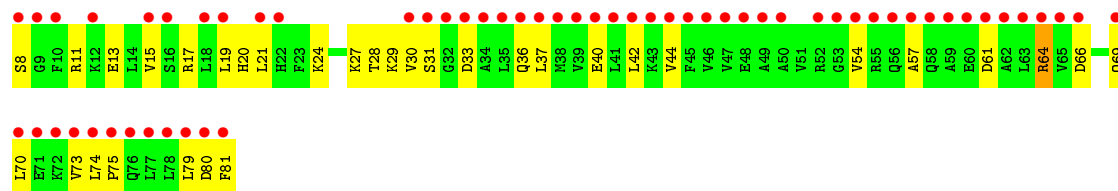
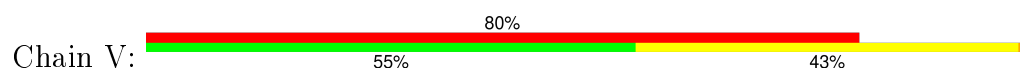




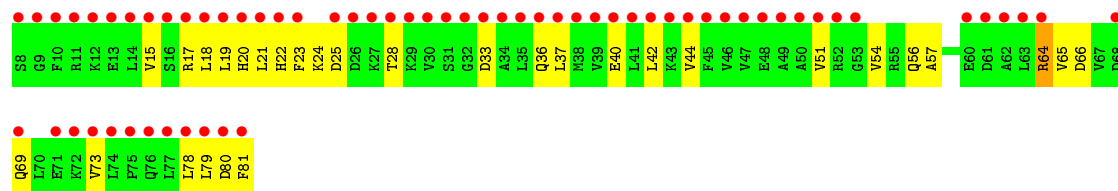
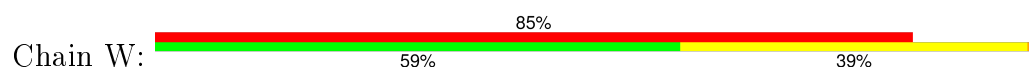
● Molecule 4: Centromere protein X



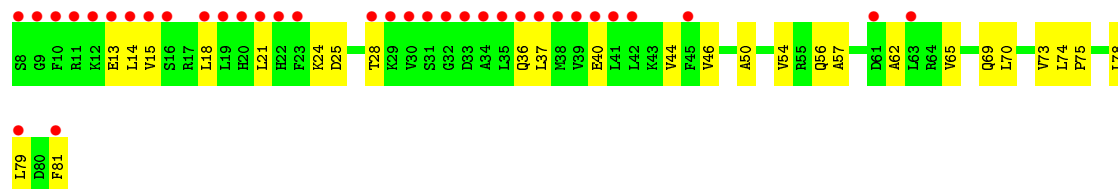
● Molecule 4: Centromere protein X



● 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	250.10 Å 250.10 Å 65.56 Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	108.30 – 7.00 108.30 – 7.00	Depositor EDS
% Data completeness (in resolution range)	99.4 (108.30-7.00) 99.8 (108.30-7.00)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.82 (at 6.73 Å)	Xtriage
Refinement program	REFMAC 5.7.0029	Depositor
R, R_{free}	0.242 , 0.279 0.248 , 0.256	Depositor DCC
R_{free} test set	449 reflections (6.27%)	DCC
Wilson B-factor (Å ²)	426.8	Xtriage
Anisotropy	0.139	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.37 , 435.5	EDS
Estimated twinning fraction	0.461 for H, K, L 0.046 for -K, -H, -L 0.045 for -H, -K, L 0.449 for K, H, -L 0.000 for -h,-k,l 0.107 for h,-h-k,-l 0.000 for -k,-h,-l	Xtriage
Reported twinning fraction	0.461 for H, K, L 0.046 for -K, -H, -L 0.045 for -H, -K, L 0.449 for K, H, -L	Depositor
L-test for twinning ²	$\langle L \rangle = 0.47$, $\langle L^2 \rangle = 0.30$	Xtriage
Outliers	0 of 7223 reflections	Xtriage
F_o, F_c correlation	0.78	EDS
Total number of atoms	15914	wwPDB-VP
Average B, all atoms (Å ²)	310.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.49% of the height of the origin peak. No significant pseudotranslation is detected.*

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.

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	0.55	1/623 (0.2%)	0.80	0/958
1	O	0.52	0/623	0.87	0/958
2	F	0.58	0/571	0.82	0/880
2	P	0.62	1/571 (0.2%)	0.86	1/880 (0.1%)
3	A	0.62	0/751	0.78	0/1007
3	C	0.51	0/751	0.76	0/1007
3	G	0.60	1/751 (0.1%)	0.74	0/1007
3	I	1.04	6/852 (0.7%)	0.94	3/1137 (0.3%)
3	J	0.73	3/751 (0.4%)	0.82	2/1007 (0.2%)
3	K	0.51	0/852	0.73	0/1137
3	Q	0.79	2/852 (0.2%)	0.80	2/1137 (0.2%)
3	R	0.86	3/751 (0.4%)	0.77	0/1007
3	S	0.67	0/751	0.83	2/1007 (0.2%)
3	T	0.50	0/852	0.69	0/1137
4	B	0.53	0/605	0.72	0/813
4	D	0.45	0/596	0.69	0/801
4	H	0.52	0/596	0.73	0/801
4	L	0.48	0/605	0.66	0/813
4	M	0.59	0/605	0.73	0/813
4	N	0.49	0/605	0.71	0/813
4	U	0.42	0/605	0.65	0/813
4	V	0.46	0/605	0.72	1/813 (0.1%)
4	W	0.51	0/605	0.73	0/813
4	X	0.48	0/605	0.73	0/813
All	All	0.61	17/16334 (0.1%)	0.77	11/22372 (0.0%)

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
3	A	0	1

The worst 5 of 17 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	I	33	GLU	CG-CD	18.16	1.79	1.51
3	R	32	GLU	CD-OE1	12.68	1.39	1.25
3	Q	32	GLU	CD-OE2	11.12	1.37	1.25
3	I	32	GLU	CD-OE2	10.73	1.37	1.25
3	J	66	GLU	CD-OE1	8.02	1.34	1.25

The worst 5 of 11 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	I	33	GLU	OE1-CD-OE2	-8.38	113.25	123.30
3	I	18	ARG	CG-CD-NE	-8.19	94.59	111.80
3	Q	110	ARG	NE-CZ-NH1	6.94	123.77	120.30
3	S	71	HIS	N-CA-CB	-6.29	99.28	110.60
3	J	29	CYS	CA-CB-SG	5.66	124.18	114.00

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
3	A	70	ARG	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	30	1
1	O	546	0	287	46	0
2	F	520	0	313	51	0
2	P	520	0	313	63	0
3	A	743	0	753	43	0
3	C	743	0	753	56	0
3	G	743	0	753	16	1
3	I	844	0	872	56	0
3	J	743	0	753	50	2
3	K	844	0	872	48	0
3	Q	844	0	872	62	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	R	743	0	753	39	0
3	S	743	0	753	66	0
3	T	844	0	872	58	0
4	B	596	0	628	26	1
4	D	590	0	620	69	0
4	H	590	0	620	30	1
4	L	596	0	628	33	0
4	M	596	0	628	57	1
4	N	596	0	628	43	1
4	U	596	0	628	49	0
4	V	596	0	628	58	0
4	W	596	0	628	37	0
4	X	596	0	628	43	0
All	All	15914	0	15470	673	4

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 22.

The worst 5 of 673 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)
3:Q:41:GLN:OE1	3:S:24:HIS:CD2	1.65	1.50
3:I:33:GLU:CD	3:I:33:GLU:CG	1.79	1.49
3:A:71:HIS:CE1	3:S:71:HIS:CE1	2.13	1.37
4:H:61:ASP:O	3:A:44:LYS:CE	1.73	1.34
1:E:11:DA:C2	2:F:17:DT:O2	1.86	1.29

All (4) 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)
3:G:48:ALA:CB	3:J:32:GLU:OE2[3_555]	2.00	0.20
4:H:61:ASP:OD2	4:M:64:ARG:NH1[3_554]	2.04	0.16
4:B:61:ASP:OD1	3:J:45:GLN:NE2[3_555]	2.10	0.10
1:E:1:DA:OP2	4:N:13:GLU:OE1[1_556]	2.11	0.09

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%)	1 (1%)	2 (2%)	8	49
3	C	91/105 (87%)	91 (100%)	0	0	100	100
3	G	91/105 (87%)	91 (100%)	0	0	100	100
3	I	103/105 (98%)	103 (100%)	0	0	100	100
3	J	91/105 (87%)	91 (100%)	0	0	100	100
3	K	103/105 (98%)	103 (100%)	0	0	100	100
3	Q	103/105 (98%)	103 (100%)	0	0	100	100
3	R	91/105 (87%)	91 (100%)	0	0	100	100
3	S	91/105 (87%)	91 (100%)	0	0	100	100
3	T	103/105 (98%)	103 (100%)	0	0	100	100
4	B	73/74 (99%)	73 (100%)	0	0	100	100
4	D	72/74 (97%)	72 (100%)	0	0	100	100
4	H	72/74 (97%)	72 (100%)	0	0	100	100
4	L	73/74 (99%)	73 (100%)	0	0	100	100
4	M	73/74 (99%)	73 (100%)	0	0	100	100
4	N	73/74 (99%)	73 (100%)	0	0	100	100
4	U	73/74 (99%)	73 (100%)	0	0	100	100
4	V	73/74 (99%)	73 (100%)	0	0	100	100
4	W	73/74 (99%)	73 (100%)	0	0	100	100
4	X	73/74 (99%)	73 (100%)	0	0	100	100
All	All	1686/1790 (94%)	1683 (100%)	1 (0%)	2 (0%)	56	90

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	A	15	TYR

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
3	A	71	HIS

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%)	73 (91%)	7 (9%)	12	45
3	C	80/91 (88%)	72 (90%)	8 (10%)	9	38
3	G	80/91 (88%)	73 (91%)	7 (9%)	12	45
3	I	91/91 (100%)	83 (91%)	8 (9%)	12	45
3	J	80/91 (88%)	70 (88%)	10 (12%)	6	30
3	K	91/91 (100%)	80 (88%)	11 (12%)	6	31
3	Q	91/91 (100%)	84 (92%)	7 (8%)	16	52
3	R	80/91 (88%)	72 (90%)	8 (10%)	9	38
3	S	80/91 (88%)	73 (91%)	7 (9%)	12	45
3	T	91/91 (100%)	80 (88%)	11 (12%)	6	31
4	B	66/65 (102%)	64 (97%)	2 (3%)	48	77
4	D	65/65 (100%)	61 (94%)	4 (6%)	23	60
4	H	65/65 (100%)	61 (94%)	4 (6%)	23	60
4	L	66/65 (102%)	64 (97%)	2 (3%)	48	77
4	M	66/65 (102%)	63 (96%)	3 (4%)	34	69
4	N	66/65 (102%)	63 (96%)	3 (4%)	34	69
4	U	66/65 (102%)	61 (92%)	5 (8%)	16	53
4	V	66/65 (102%)	60 (91%)	6 (9%)	12	43
4	W	66/65 (102%)	62 (94%)	4 (6%)	23	60
4	X	66/65 (102%)	64 (97%)	2 (3%)	48	77
All	All	1502/1560 (96%)	1383 (92%)	119 (8%)	15	51

5 of 119 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	K	41	GLN
4	N	8	SER
4	V	21	LEU
3	K	70	ARG
3	K	115	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 10 such sidechains are listed below:

Mol	Chain	Res	Type
3	J	71	HIS
3	K	60	ASN
3	S	24	HIS
3	I	71	HIS
3	Q	41	GLN

5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

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

5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

5.6 Ligand geometry [i](#)

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%)	0.90	2 (7%) 16 19	247, 283, 314, 333	0
1	O	26/26 (100%)	0.95	3 (11%) 6 12	251, 313, 356, 401	0
2	F	26/26 (100%)	2.05	14 (53%) 0 3	262, 291, 387, 437	0
2	P	26/26 (100%)	2.12	11 (42%) 0 4	289, 343, 397, 450	0
3	A	93/105 (88%)	1.93	32 (34%) 0 4	179, 253, 307, 350	0
3	C	93/105 (88%)	3.71	66 (70%) 0 3	325, 373, 421, 453	0
3	G	93/105 (88%)	3.34	66 (70%) 0 3	278, 329, 390, 427	0
3	I	105/105 (100%)	2.06	46 (43%) 0 4	188, 244, 388, 435	0
3	J	93/105 (88%)	1.91	38 (40%) 0 4	211, 274, 348, 404	0
3	K	105/105 (100%)	2.58	62 (59%) 0 3	231, 325, 405, 441	0
3	Q	105/105 (100%)	2.59	62 (59%) 0 3	173, 254, 335, 366	0
3	R	93/105 (88%)	3.29	63 (67%) 0 3	219, 284, 365, 410	0
3	S	93/105 (88%)	1.90	35 (37%) 0 4	203, 259, 317, 392	0
3	T	105/105 (100%)	2.52	50 (47%) 0 4	295, 395, 495, 571	0
4	B	74/74 (100%)	2.77	39 (52%) 0 4	219, 309, 379, 408	0
4	D	74/74 (100%)	4.52	56 (75%) 0 2	352, 405, 470, 498	1 (1%)
4	H	74/74 (100%)	4.07	57 (77%) 0 2	231, 292, 368, 415	1 (1%)
4	L	74/74 (100%)	4.32	57 (77%) 0 2	270, 341, 419, 445	0
4	M	74/74 (100%)	2.29	38 (51%) 0 4	163, 240, 312, 344	0
4	N	74/74 (100%)	2.88	45 (60%) 0 3	216, 269, 337, 368	0
4	U	74/74 (100%)	3.19	47 (63%) 0 3	386, 412, 465, 474	0
4	V	74/74 (100%)	3.85	59 (79%) 0 2	244, 281, 349, 386	0
4	W	74/74 (100%)	3.87	63 (85%) 0 2	202, 284, 364, 385	0
4	X	74/74 (100%)	3.17	35 (47%) 0 4	211, 272, 337, 364	0

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å ²)	Q<0.9
All	All	1822/1894 (96%)	2.89	1046 (57%) 0 3	163, 299, 435, 571	2 (0%)

The worst 5 of 1046 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
4	X	9	GLY	18.5
4	L	9	GLY	15.1
4	D	15	VAL	14.0
3	T	116	LYS	13.8
4	X	31	SER	13.3

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

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

6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.

6.4 Ligands ⓘ

There are no ligands in this entry.

6.5 Other polymers ⓘ

There are no such residues in this entry.