



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 03:54 PM GMT

PDB ID : 4DRB
Title : The crystal structure of FANCM bound MHF complex
Authors : Tao, Y.; Niu, L.; Teng, M.
Deposited on : 2012-02-17
Resolution : 2.63 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

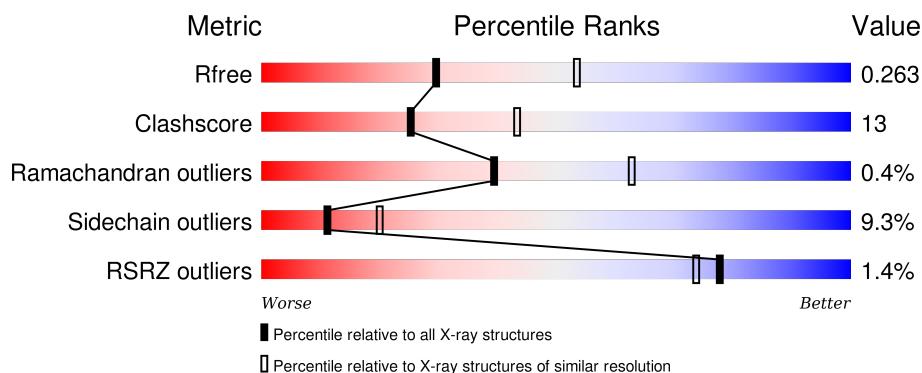
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.63 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.




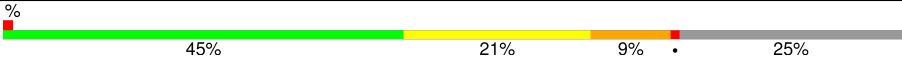








Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	91344	3377 (2.68-2.60)
Clashscore	102246	3781 (2.68-2.60)
Ramachandran outliers	100387	3722 (2.68-2.60)
Sidechain outliers	100360	3722 (2.68-2.60)
RSRZ outliers	91569	3388 (2.68-2.60)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions $\leq 5\%$. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	120	<div> <div>3%</div> <div>53% 23% 5% 18%</div> </div>
1	B	120	<div> <div>61% 24% • 14%</div> </div>
1	D	120	<div> <div>3%</div> <div>58% 22% • 19%</div> </div>
1	E	120	<div> <div>3%</div> <div>59% 21% • 19%</div> </div>
1	G	120	<div> <div>•</div> <div>73% 11% • 13%</div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	H	120	
2	C	141	% 
2	F	141	% 
2	I	141	
3	J	84	
3	K	84	% 
3	L	84	2% 
3	M	84	4% 
3	N	84	
3	O	84	

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 10508 atoms, of which 0 are hydrogens and 0 are deuteriums.

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

- Molecule 1 is a protein called Centromere protein S.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	98	Total	C	N	O	S	Se	0	0	0
			758	477	131	145	3	2			
1	B	103	Total	C	N	O	S	Se	0	0	0
			769	480	134	150	3	2			
1	D	97	Total	C	N	O	S	Se	0	1	0
			736	460	125	146	3	2			
1	E	97	Total	C	N	O	S	Se	0	0	0
			699	436	124	134	3	2			
1	G	105	Total	C	N	O	S	Se	0	0	0
			793	492	141	155	3	2			
1	H	94	Total	C	N	O	S	Se	0	0	0
			731	459	125	142	3	2			

There are 36 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-5	HIS	-	EXPRESSION TAG	UNP Q8N2Z9
A	-4	HIS	-	EXPRESSION TAG	UNP Q8N2Z9
A	-3	HIS	-	EXPRESSION TAG	UNP Q8N2Z9
A	-2	HIS	-	EXPRESSION TAG	UNP Q8N2Z9
A	-1	HIS	-	EXPRESSION TAG	UNP Q8N2Z9
A	0	HIS	-	EXPRESSION TAG	UNP Q8N2Z9
B	-5	HIS	-	EXPRESSION TAG	UNP Q8N2Z9
B	-4	HIS	-	EXPRESSION TAG	UNP Q8N2Z9
B	-3	HIS	-	EXPRESSION TAG	UNP Q8N2Z9
B	-2	HIS	-	EXPRESSION TAG	UNP Q8N2Z9
B	-1	HIS	-	EXPRESSION TAG	UNP Q8N2Z9
B	0	HIS	-	EXPRESSION TAG	UNP Q8N2Z9
D	-5	HIS	-	EXPRESSION TAG	UNP Q8N2Z9
D	-4	HIS	-	EXPRESSION TAG	UNP Q8N2Z9
D	-3	HIS	-	EXPRESSION TAG	UNP Q8N2Z9
D	-2	HIS	-	EXPRESSION TAG	UNP Q8N2Z9
D	-1	HIS	-	EXPRESSION TAG	UNP Q8N2Z9

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	0	HIS	-	EXPRESSION TAG	UNP Q8N2Z9
E	-5	HIS	-	EXPRESSION TAG	UNP Q8N2Z9
E	-4	HIS	-	EXPRESSION TAG	UNP Q8N2Z9
E	-3	HIS	-	EXPRESSION TAG	UNP Q8N2Z9
E	-2	HIS	-	EXPRESSION TAG	UNP Q8N2Z9
E	-1	HIS	-	EXPRESSION TAG	UNP Q8N2Z9
E	0	HIS	-	EXPRESSION TAG	UNP Q8N2Z9
G	-5	HIS	-	EXPRESSION TAG	UNP Q8N2Z9
G	-4	HIS	-	EXPRESSION TAG	UNP Q8N2Z9
G	-3	HIS	-	EXPRESSION TAG	UNP Q8N2Z9
G	-2	HIS	-	EXPRESSION TAG	UNP Q8N2Z9
G	-1	HIS	-	EXPRESSION TAG	UNP Q8N2Z9
G	0	HIS	-	EXPRESSION TAG	UNP Q8N2Z9
H	-5	HIS	-	EXPRESSION TAG	UNP Q8N2Z9
H	-4	HIS	-	EXPRESSION TAG	UNP Q8N2Z9
H	-3	HIS	-	EXPRESSION TAG	UNP Q8N2Z9
H	-2	HIS	-	EXPRESSION TAG	UNP Q8N2Z9
H	-1	HIS	-	EXPRESSION TAG	UNP Q8N2Z9
H	0	HIS	-	EXPRESSION TAG	UNP Q8N2Z9

- Molecule 2 is a protein called Fanconi anemia group M protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	106	Total	C	N	O	S	0	0	0
			860	542	151	161	6			
2	F	100	Total	C	N	O	S	0	1	0
			796	505	141	144	6			
2	I	103	Total	C	N	O	S	0	1	0
			857	542	151	158	6			

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	660	GLY	-	EXPRESSION TAG	UNP Q8IYD8
F	660	GLY	-	EXPRESSION TAG	UNP Q8IYD8
I	660	GLY	-	EXPRESSION TAG	UNP Q8IYD8

- Molecule 3 is a protein called Centromere protein X.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	J	74	Total	C	N	O	Se	0	0	0
			581	371	102	107	1			

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	K	74	Total	C	N	O	Se	0	0	0
			570	365	100	104	1			
3	L	74	Total	C	N	O	Se	0	0	0
			558	358	96	103	1			
3	M	74	Total	C	N	O	Se	0	0	0
			529	340	88	100	1			
3	N	74	Total	C	N	O	Se	0	0	0
			567	363	97	106	1			
3	O	74	Total	C	N	O	Se	0	0	0
			584	374	102	107	1			

There are 18 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
J	-2	GLY	-	EXPRESSION TAG	UNP A8MT69
J	-1	SER	-	EXPRESSION TAG	UNP A8MT69
J	0	HIS	-	EXPRESSION TAG	UNP A8MT69
K	-2	GLY	-	EXPRESSION TAG	UNP A8MT69
K	-1	SER	-	EXPRESSION TAG	UNP A8MT69
K	0	HIS	-	EXPRESSION TAG	UNP A8MT69
L	-2	GLY	-	EXPRESSION TAG	UNP A8MT69
L	-1	SER	-	EXPRESSION TAG	UNP A8MT69
L	0	HIS	-	EXPRESSION TAG	UNP A8MT69
M	-2	GLY	-	EXPRESSION TAG	UNP A8MT69
M	-1	SER	-	EXPRESSION TAG	UNP A8MT69
M	0	HIS	-	EXPRESSION TAG	UNP A8MT69
N	-2	GLY	-	EXPRESSION TAG	UNP A8MT69
N	-1	SER	-	EXPRESSION TAG	UNP A8MT69
N	0	HIS	-	EXPRESSION TAG	UNP A8MT69
O	-2	GLY	-	EXPRESSION TAG	UNP A8MT69
O	-1	SER	-	EXPRESSION TAG	UNP A8MT69
O	0	HIS	-	EXPRESSION TAG	UNP A8MT69

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	7	Total	O	0	0
			7	7		
4	B	10	Total	O	0	0
			10	10		
4	C	17	Total	O	0	0
			17	17		

Continued on next page...

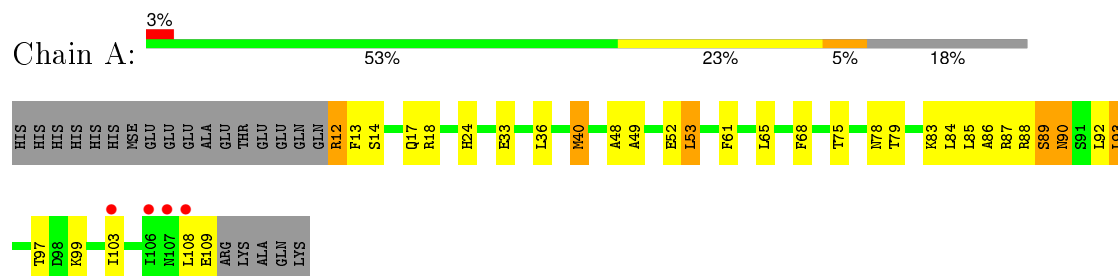
Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	D	3	Total 3	O 3	0	0
4	E	3	Total 3	O 3	0	0
4	F	3	Total 3	O 3	0	0
4	G	8	Total 8	O 8	0	0
4	H	12	Total 12	O 12	0	0
4	I	18	Total 18	O 18	0	0
4	J	3	Total 3	O 3	0	0
4	K	13	Total 13	O 13	0	0
4	L	8	Total 8	O 8	0	0
4	M	3	Total 3	O 3	0	0
4	N	4	Total 4	O 4	0	0
4	O	8	Total 8	O 8	0	0

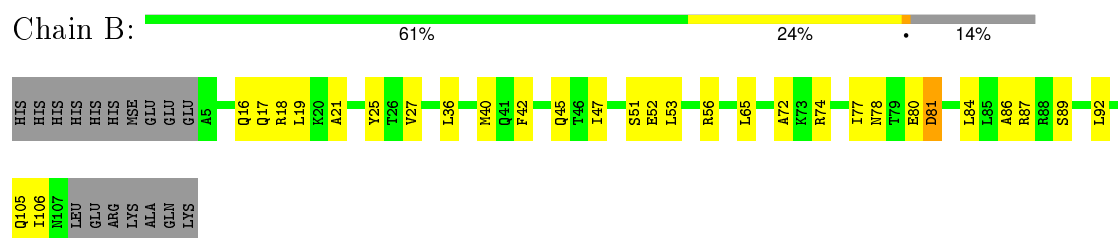
3 Residue-property plots [i](#)

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

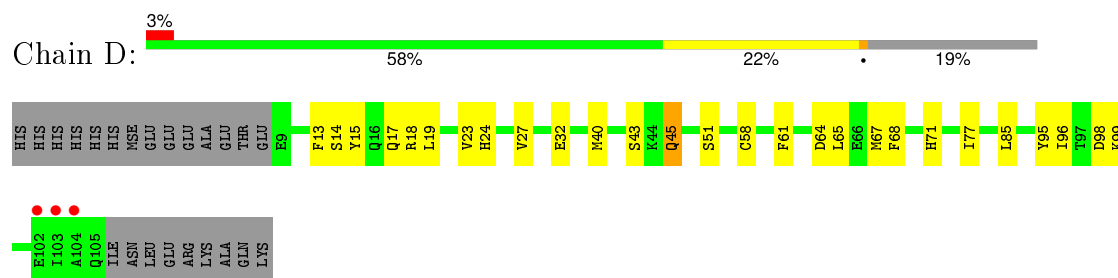
- Molecule 1: Centromere protein S



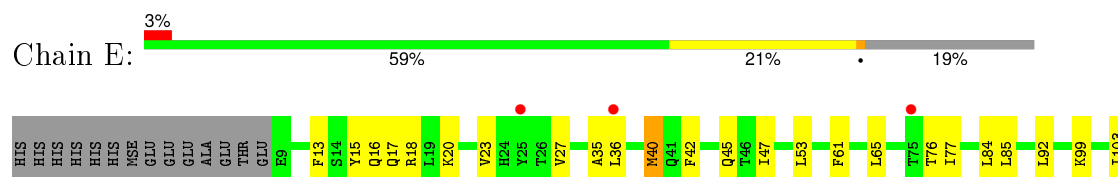
- Molecule 1: Centromere protein S



- Molecule 1: Centromere protein S



- Molecule 1: Centromere protein S



A104
Q105
ILE
ASN
LEU
GLU
ARG
LYS
ALA
GLN
LYS

• Molecule 1: Centromere protein S

Chain G:  73% 11% 13%

HIS HIS HIS HIS HIS HIS MSE GLU GLU E4 T7 Q11 R12 A21 A22 V23 H24 V27 L36 Q45 S51 S52 E52 L53 L55 L65 I77 A86 R87 R88 L92 I103 H107 L108 GLU ARG LYS LYS ALA GLN LYS


• Molecule 1: Centromere protein S

Chain H:  63% 16% 22%

HIS HIS HIS HIS HIS HIS MSE GLU GLU GLU GLU GLU THR GLU GLN GLN ARG F13 A22 V23 E33 M40 S43 K44 N689 E52 L53 C58 E59 K63 D64 L65 R70 T75 T76 I77 S89 L92 L93 I96 I106 ASN LEU GLU ARG LYS ALA

GLN
LYS

• Molecule 2: Fanconi anemia group M protein

Chain C:  45% 21% 9% 25%

GLY SER ILE PHE SER TYR ARG ASP GLY MET ARG GLN SER SER LEU K675 K676 D677 W678 F679 E682 E683 E684 F685 K686 W688 N689 R690 R693 L694 R695 E702 I703 T704 L705 P706 Q707 V708 Q714 ASN GLU GLU ASN LYS PRO ALA GLN GLU SER T725 T726 H729 H730

W736 R737 Q740 L744 H747 Q748 H751 R754 F758 L761 M764 I765 E766 R769 H770 E771 E772 G773 V781 Q786 V790 T791 T792 P793 P794 I795 P796 Q797 Q798 Q799 Q800 Q801 Q802 Q803 Q804 Q805 Q806 Q807 Q808 Q809 Q810 Q811 Q812 Q813 Q814 Q815 Q816 Q817 Q818 Q819 Q820 Q821 Q822 Q823 Q824 Q825 Q826 Q827 Q828 Q829 Q830 Q831 Q832 Q833 Q834 Q835 Q836 Q837 Q838 Q839 Q840 Q841 Q842 Q843 Q844 Q845 Q846 Q847 Q848 Q849 Q850 Q851 Q852 Q853 Q854 Q855 Q856 Q857 Q858 Q859 Q860 Q861 Q862 Q863 Q864 Q865 Q866 Q867 Q868 Q869 Q870 Q871 Q872 Q873 Q874 Q875 Q876 Q877 Q878 Q879 Q880 Q881 Q882 Q883 Q884 Q885 Q886 Q887 Q888 Q889 Q890 Q891 Q892 Q893 Q894 Q895 Q896 Q897 Q898 Q899 Q900 Q901 Q902 Q903 Q904 Q905 Q906 Q907 Q908 Q909 Q910 Q911 Q912 Q913 Q914 Q915 Q916 Q917 Q918 Q919 Q920 Q921 Q922 Q923 Q924 Q925 Q926 Q927 Q928 Q929 Q930 Q931 Q932 Q933 Q934 Q935 Q936 Q937 Q938 Q939 Q940 Q941 Q942 Q943 Q944 Q945 Q946 Q947 Q948 Q949 Q950 Q951 Q952 Q953 Q954 Q955 Q956 Q957 Q958 Q959 Q960 Q961 Q962 Q963 Q964 Q965 Q966 Q967 Q968 Q969 Q970 Q971 Q972 Q973 Q974 Q975 Q976 Q977 Q978 Q979 Q980 Q981 Q982 Q983 Q984 Q985 Q986 Q987 Q988 Q989 Q990 Q991 Q992 Q993 Q994 Q995 Q996 Q997 Q998 Q999

• Molecule 2: Fanconi anemia group M protein

Chain F:  51% 16% 29%

GLY SER ILE PHE SER TYR ARG ASP GLY MET ARG GLN SER SER LEU K675 W688 N689 R693 D696 L705 P706 Q707 S711 S712 L712 LEU GLN ASN GLU GLU GLU LYS LYS PRO P793 P794 I795 P796 Q797 Q798 Q799 Q800 Q801 Q802 Q803 Q804 Q805 Q806 Q807 Q808 Q809 Q810 Q811 Q812 Q813 Q814 Q815 Q816 Q817 Q818 Q819 Q820 Q821 Q822 Q823 Q824 Q825 Q826 Q827 Q828 Q829 Q830 Q831 Q832 Q833 Q834 Q835 Q836 Q837 Q838 Q839 Q840 Q841 Q842 Q843 Q844 Q845 Q846 Q847 Q848 Q849 Q850 Q851 Q852 Q853 Q854 Q855 Q856 Q857 Q858 Q859 Q860 Q861 Q862 Q863 Q864 Q865 Q866 Q867 Q868 Q869 Q870 Q871 Q872 Q873 Q874 Q875 Q876 Q877 Q878 Q879 Q880 Q881 Q882 Q883 Q884 Q885 Q886 Q887 Q888 Q889 Q890 Q891 Q892 Q893 Q894 Q895 Q896 Q897 Q898 Q899 Q900 Q901 Q902 Q903 Q904 Q905 Q906 Q907 Q908 Q909 Q910 Q911 Q912 Q913 Q914 Q915 Q916 Q917 Q918 Q919 Q920 Q921 Q922 Q923 Q924 Q925 Q926 Q927 Q928 Q929 Q930 Q931 Q932 Q933 Q934 Q935 Q936 Q937 Q938 Q939 Q940 Q941 Q942 Q943 Q944 Q945 Q946 Q947 Q948 Q949 Q950 Q951 Q952 Q953 Q954 Q955 Q956 Q957 Q958 Q959 Q960 Q961 Q962 Q963 Q964 Q965 Q966 Q967 Q968 Q969 Q970 Q971 Q972 Q973 Q974 Q975 Q976 Q977 Q978 Q979 Q980 Q981 Q982 Q983 Q984 Q985 Q986 Q987 Q988 Q989 Q990 Q991 Q992 Q993 Q994 Q995 Q996 Q997 Q998 Q999

• Molecule 2: Fanconi anemia group M protein

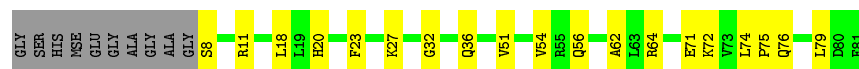
Chain I:  48% 20% 5% 27%

GLY SER ILE PHE SER TYR ARG ASP GLY MET ARG GLN SER SER LEU LEU K676 F679 L680 W688 N689 R693 E694 W698 S699 R699 L694 R695 D696 I700 L705 F710 S711 S712 L713 GLN ASN GLU GLU GLU ASN LYS PRO PRO ALA GLN GLU SER T725 T726 H729 H747 Q748

R754 H757 L761 M762 I765 M768 R769 H770 E771 GLU GLY E774 E780 V781 L785 Q786 V790 T791 T792 P793 P794 I795 P796 Q797 Q798 Q799 Q800 Q801 Q802 Q803 Q804 Q805 Q806 Q807 Q808 Q809 Q810 Q811 Q812 Q813 Q814 Q815 Q816 Q817 Q818 Q819 Q820 Q821 Q822 Q823 Q824 Q825 Q826 Q827 Q828 Q829 Q830 Q831 Q832 Q833 Q834 Q835 Q836 Q837 Q838 Q839 Q840 Q841 Q842 Q843 Q844 Q845 Q846 Q847 Q848 Q849 Q850 Q851 Q852 Q853 Q854 Q855 Q856 Q857 Q858 Q859 Q860 Q861 Q862 Q863 Q864 Q865 Q866 Q867 Q868 Q869 Q870 Q871 Q872 Q873 Q874 Q875 Q876 Q877 Q878 Q879 Q880 Q881 Q882 Q883 Q884 Q885 Q886 Q887 Q888 Q889 Q890 Q891 Q892 Q893 Q894 Q895 Q896 Q897 Q898 Q899 Q900 Q901 Q902 Q903 Q904 Q905 Q906 Q907 Q908 Q909 Q910 Q911 Q912 Q913 Q914 Q915 Q916 Q917 Q918 Q919 Q920 Q921 Q922 Q923 Q924 Q925 Q926 Q927 Q928 Q929 Q930 Q931 Q932 Q933 Q934 Q935 Q936 Q937 Q938 Q939 Q940 Q941 Q942 Q943 Q944 Q945 Q946 Q947 Q948 Q949 Q950 Q951 Q952 Q953 Q954 Q955 Q956 Q957 Q958 Q959 Q960 Q961 Q962 Q963 Q964 Q965 Q966 Q967 Q968 Q969 Q970 Q971 Q972 Q973 Q974 Q975 Q976 Q977 Q978 Q979 Q980 Q981 Q982 Q983 Q984 Q985 Q986 Q987 Q988 Q989 Q990 Q991 Q992 Q993 Q994 Q995 Q996 Q997 Q998 Q999

- Molecule 3: Centromere protein X

Chain J: 



- Molecule 3: Centromere protein X

Chain K: 



- Molecule 3: Centromere protein X

Chain L: 



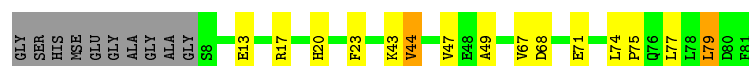
- Molecule 3: Centromere protein X

Chain M: 



- Molecule 3: Centromere protein X

Chain N: 



- Molecule 3: Centromere protein X

Chain O: 



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	111.32Å 70.03Å 115.75Å 90.00° 91.41° 90.00°	Depositor
Resolution (Å)	49.67 – 2.63 49.67 – 2.63	Depositor EDS
% Data completeness (in resolution range)	95.1 (49.67-2.63) 99.6 (49.67-2.63)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	4.87 (at 2.65Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.6 _289)	Depositor
R, R_{free}	0.221 , 0.256 0.224 , 0.263	Depositor DCC
R_{free} test set	2693 reflections (5.37%)	DCC
Wilson B-factor (Å ²)	42.7	Xtriage
Anisotropy	0.390	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.29 , 40.9	EDS
Estimated twinning fraction	0.000 for l,k,-h 0.020 for h,-k,-l 0.012 for l,-k,h	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtriage
Outliers	0 of 52875 reflections	Xtriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	10508	wwPDB-VP
Average B, all atoms (Å ²)	44.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.24% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 0.2 for perfectly twinned datasets.

5 Model quality [i](#)

5.1 Standard geometry [i](#)

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# Z >5	RMSZ	# Z >5
1	A	0.22	0/765	0.36	0/1027
1	B	0.26	0/775	0.38	0/1045
1	D	0.22	0/743	0.37	0/1001
1	E	0.25	0/704	0.40	0/953
1	G	0.21	0/799	0.36	0/1074
1	H	0.22	0/738	0.34	0/991
2	C	0.25	0/882	0.45	1/1195 (0.1%)
2	F	0.21	0/818	0.38	0/1111
2	I	0.21	0/882	0.36	0/1194
3	J	0.21	0/586	0.33	0/788
3	K	0.22	0/575	0.35	0/777
3	L	0.20	0/563	0.32	0/762
3	M	0.20	0/534	0.32	0/728
3	N	0.21	0/572	0.35	0/772
3	O	0.21	0/589	0.32	0/792
All	All	0.22	0/10525	0.37	1/14210 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	C	705	LEU	CA-CB-CG	5.50	127.96	115.30

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts [i](#)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	758	0	724	28	0
1	B	769	0	715	25	0
1	D	736	0	673	16	0
1	E	699	0	630	21	0
1	G	793	0	740	15	0
1	H	731	0	698	16	0
2	C	860	0	771	44	0
2	F	796	0	679	22	0
2	I	857	0	781	28	0
3	J	581	0	593	15	0
3	K	570	0	571	21	0
3	L	558	0	547	16	0
3	M	529	0	491	23	0
3	N	567	0	565	10	0
3	O	584	0	602	14	0
4	A	7	0	0	1	0
4	B	10	0	0	0	0
4	C	17	0	0	1	0
4	D	3	0	0	0	0
4	E	3	0	0	0	0
4	F	3	0	0	0	0
4	G	8	0	0	1	0
4	H	12	0	0	0	0
4	I	18	0	0	0	0
4	J	3	0	0	0	0
4	K	13	0	0	0	0
4	L	8	0	0	0	0
4	M	3	0	0	0	0
4	N	4	0	0	0	0
4	O	8	0	0	0	0
All	All	10508	0	9780	254	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 13.

The worst 5 of 254 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)
2:C:676:LYS:HG2	2:C:677:ASP:H	1.23	1.03
2:C:771:GLU:HA	2:C:772:GLU:CB	1.94	0.95
1:A:14:SER:H	1:A:17:GLN:HE21	1.19	0.90
1:E:99:LYS:HG2	3:M:40:GLU:HG2	1.54	0.86

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:E:61:PHE:HB3	3:M:38:MSE:HE1	1.58	0.85

There are no symmetry-related clashes.

5.3 Torsion angles [i](#)

5.3.1 Protein backbone [i](#)

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

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	96/120 (80%)	96 (100%)	0	0	100	100
1	B	101/120 (84%)	97 (96%)	3 (3%)	1 (1%)	19	37
1	D	96/120 (80%)	96 (100%)	0	0	100	100
1	E	95/120 (79%)	91 (96%)	3 (3%)	1 (1%)	17	34
1	G	103/120 (86%)	103 (100%)	0	0	100	100
1	H	92/120 (77%)	92 (100%)	0	0	100	100
2	C	102/141 (72%)	94 (92%)	5 (5%)	3 (3%)	6	9
2	F	95/141 (67%)	91 (96%)	4 (4%)	0	100	100
2	I	98/141 (70%)	92 (94%)	6 (6%)	0	100	100
3	J	72/84 (86%)	71 (99%)	1 (1%)	0	100	100
3	K	72/84 (86%)	72 (100%)	0	0	100	100
3	L	72/84 (86%)	72 (100%)	0	0	100	100
3	M	72/84 (86%)	69 (96%)	3 (4%)	0	100	100
3	N	72/84 (86%)	71 (99%)	1 (1%)	0	100	100
3	O	72/84 (86%)	72 (100%)	0	0	100	100
All	All	1310/1647 (80%)	1279 (98%)	26 (2%)	5 (0%)	39	63

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	105	GLN
2	C	772	GLU
2	C	706	PRO
2	C	707	GLN
1	E	103	ILE

5.3.2 Protein sidechains ⓘ

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

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	75/104 (72%)	65 (87%)	10 (13%)	5	8
1	B	74/104 (71%)	72 (97%)	2 (3%)	52	78
1	D	71/104 (68%)	65 (92%)	6 (8%)	13	24
1	E	63/104 (61%)	59 (94%)	4 (6%)	22	42
1	G	77/104 (74%)	71 (92%)	6 (8%)	16	29
1	H	74/104 (71%)	71 (96%)	3 (4%)	37	64
2	C	90/133 (68%)	74 (82%)	16 (18%)	2	3
2	F	77/133 (58%)	67 (87%)	10 (13%)	5	8
2	I	92/133 (69%)	81 (88%)	11 (12%)	6	10
3	J	62/67 (92%)	60 (97%)	2 (3%)	46	73
3	K	59/67 (88%)	53 (90%)	6 (10%)	9	16
3	L	56/67 (84%)	49 (88%)	7 (12%)	6	10
3	M	50/67 (75%)	44 (88%)	6 (12%)	6	10
3	N	59/67 (88%)	56 (95%)	3 (5%)	29	53
3	O	63/67 (94%)	58 (92%)	5 (8%)	15	28
All	All	1042/1425 (73%)	945 (91%)	97 (9%)	11	20

5 of 97 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	F	768	MET
1	H	58	CYS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
3	N	44	VAL
2	F	785	LEU
1	G	36	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 37 such sidechains are listed below:

Mol	Chain	Res	Type
2	F	740	GLN
2	I	689	ASN
3	M	69	GLN
2	F	757	HIS
1	G	45	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	A	96/120 (80%)	-0.27	4 (4%) 40 33	23, 40, 81, 95	0
1	B	101/120 (84%)	-0.26	0 100 100	27, 42, 75, 89	0
1	D	95/120 (79%)	-0.07	3 (3%) 51 45	29, 50, 71, 81	0
1	E	95/120 (79%)	0.10	3 (3%) 51 45	41, 63, 80, 92	0
1	G	103/120 (85%)	-0.28	1 (0%) 84 81	28, 41, 76, 89	0
1	H	92/120 (76%)	-0.39	0 100 100	25, 38, 66, 81	0
2	C	106/141 (75%)	-0.15	1 (0%) 85 83	26, 39, 68, 72	0
2	F	100/141 (70%)	0.00	1 (1%) 84 81	32, 54, 80, 86	0
2	I	103/141 (73%)	-0.34	0 100 100	25, 36, 58, 70	0
3	J	73/84 (86%)	-0.47	0 100 100	25, 34, 52, 59	0
3	K	73/84 (86%)	-0.31	1 (1%) 78 74	30, 38, 54, 58	0
3	L	73/84 (86%)	-0.23	2 (2%) 58 52	30, 43, 59, 72	0
3	M	73/84 (86%)	0.02	3 (4%) 41 34	35, 61, 75, 79	0
3	N	73/84 (86%)	-0.40	0 100 100	29, 38, 55, 64	0
3	O	73/84 (86%)	-0.38	0 100 100	23, 33, 50, 54	0
All	All	1329/1647 (80%)	-0.22	19 (1%) 78 74	23, 42, 74, 95	0

The worst 5 of 19 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	103	ILE	4.5
1	A	107	ASN	4.5
1	E	25	TYR	3.8
1	D	103	ILE	3.6
3	K	63	LEU	3.5

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.