WORLDWIDE wwPDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

Jun 18, 2024 – 09:47 PM EDT

PDB ID	:	4AJ5
Title	:	Crystal structure of the Ska core complex
Authors	:	Jeyaprakash, A.A.; Santamaria, A.; Jayachandran, U.; Chan, Y.W.; Benda, C.; Nigg, E.A.; Conti, E.
Deposited on	:	2012-02-15
Resolution	:	3.32 Å(reported)

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

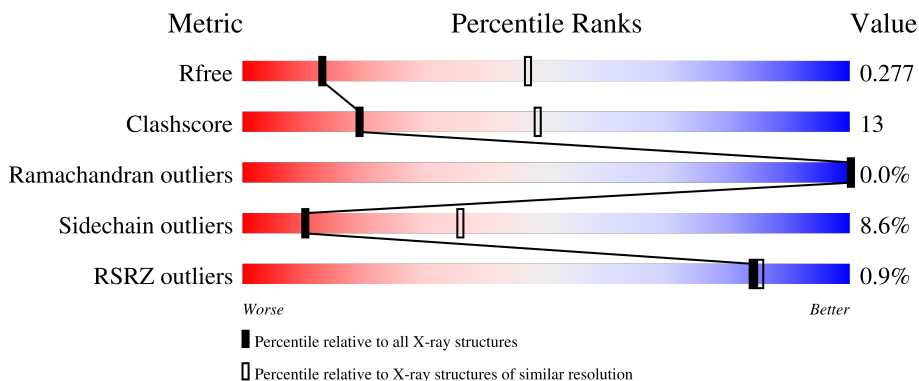|		
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.20.1
EDS	:	2.37.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

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 3.32 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.
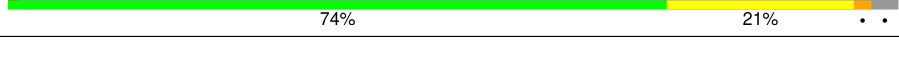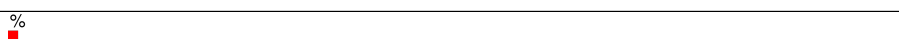
Metric	Percentile Ranks	Value
Rfree		0.277
Clashscore		13
Ramachandran outliers		0.0%
Sidechain outliers		8.6%
RSRZ outliers		0.9%

Worse — *Better*

▌ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution
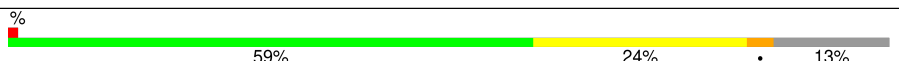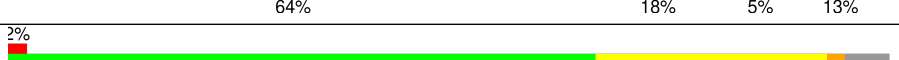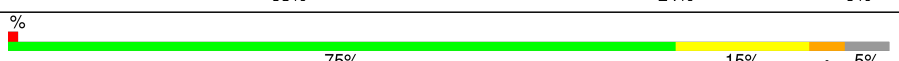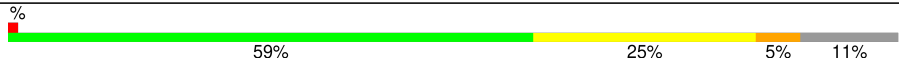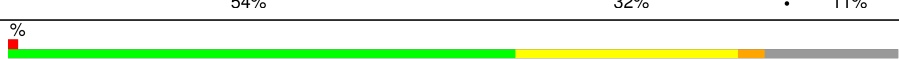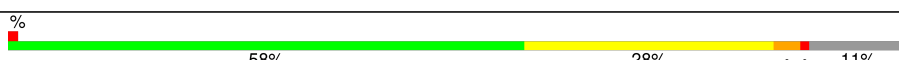
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1089 (3.36-3.28)
Clashscore	141614	1137 (3.36-3.28)
Ramachandran outliers	138981	1115 (3.36-3.28)
Sidechain outliers	138945	1114 (3.36-3.28)
RSRZ outliers	127900	1059 (3.36-3.28)

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

Mol	Chain	Length	Quality of chain
1	1	101	69% / 25% · ·
1	2	101	% / 64% / 30% / 5% ·
1	3	101	71% / 24% · ·
1	4	101	67% / 25% / 5% ·
1	U	101	63% / 33% · ·

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	V	101	68% ... 28% ...
1	W	101	67% ... 28% ...
1	X	101	74% ... 21% ...
1	Y	101	65% ... 26% ... 7%
1	Z	101	54% ... 40% ...
2	A	91	59% ... 24% ... 13%
2	B	91	71% ... 19% ... 5%
2	C	91	64% ... 18% ... 5% ... 13%
2	D	91	66% ... 26% ... 5%
2	E	91	68% ... 23% ... 5%
2	F	91	62% ... 27% ... 10%
2	G	91	76% ... 16% ... 5%
2	H	91	63% ... 24% ... 9%
2	I	91	75% ... 15% ... 5%
2	J	91	66% ... 25% ... 5%
3	K	123	60% ... 23% ... 5% ... 12%
3	L	123	63% ... 21% ... 14%
3	M	123	59% ... 25% ... 5% ... 11%
3	N	123	63% ... 18% ... 15%
3	O	123	54% ... 32% ... 11%
3	P	123	57% ... 25% ... 15%
3	Q	123	63% ... 19% ... 17%
3	R	123	54% ... 30% ... 12%
3	S	123	57% ... 27% ... 13%
3	T	123	58% ... 28% ... 11%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 21691 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 SPINDLE AND KINETOCHORE-ASSOCIATED PROTEIN 3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	1	99	Total	C	N	O	S	0	0	0
			740	458	128	151	3			
1	2	100	Total	C	N	O	S	0	0	0
			787	491	135	156	5			
1	3	98	Total	C	N	O	S	0	0	0
			738	459	126	149	4			
1	4	98	Total	C	N	O	S	0	0	0
			775	484	134	151	6			
1	U	100	Total	C	N	O	S	0	0	0
			773	484	131	152	6			
1	V	98	Total	C	N	O	S	0	0	0
			725	448	126	148	3			
1	W	98	Total	C	N	O	S	0	0	0
			727	449	128	146	4			
1	X	98	Total	C	N	O	S	0	0	0
			729	453	126	147	3			
1	Y	99	Total	C	N	O	S	0	0	0
			763	479	133	146	5			
1	Z	97	Total	C	N	O	S	0	0	0
			723	446	127	146	4			

There are 10 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
1	58	ILE	VAL	engineered mutation	UNP Q8IX90
2	58	ILE	VAL	engineered mutation	UNP Q8IX90
3	58	ILE	VAL	engineered mutation	UNP Q8IX90
4	58	ILE	VAL	engineered mutation	UNP Q8IX90
U	58	ILE	VAL	engineered mutation	UNP Q8IX90
V	58	ILE	VAL	engineered mutation	UNP Q8IX90
W	58	ILE	VAL	engineered mutation	UNP Q8IX90
X	58	ILE	VAL	engineered mutation	UNP Q8IX90

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
Y	58	ILE	VAL	engineered mutation	UNP Q8IX90
Z	58	ILE	VAL	engineered mutation	UNP Q8IX90

- Molecule 2 is a protein called SPINDLE AND KINETOCHORE-ASSOCIATED PROTEIN 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	A	79	Total 588	C 363	N 99	O 123	S 3	0	0	0
2	B	87	Total 656	C 406	N 109	O 138	S 3	0	0	0
2	C	79	Total 595	C 369	N 100	O 123	S 3	0	0	0
2	D	86	Total 646	C 399	N 108	O 136	S 3	0	0	0
2	E	86	Total 630	C 387	N 108	O 132	S 3	0	0	0
2	F	82	Total 585	C 362	N 102	O 118	S 3	0	0	0
2	G	86	Total 637	C 392	N 108	O 134	S 3	0	0	0
2	H	83	Total 633	C 391	N 104	O 135	S 3	0	0	0
2	I	86	Total 645	C 398	N 109	O 135	S 3	0	0	0
2	J	86	Total 639	C 396	N 105	O 135	S 3	0	0	0

- Molecule 3 is a protein called SPINDLE AND KINETOCHORE-ASSOCIATED PROTEIN 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	K	108	Total 813	C 511	N 132	O 166	S 4	0	0	0
3	L	106	Total 779	C 491	N 130	O 155	S 3	0	0	0
3	M	109	Total 800	C 505	N 133	O 158	S 4	0	0	0
3	N	104	Total 770	C 485	N 129	O 152	S 4	0	0	0
3	O	109	Total 840	C 528	N 140	O 168	S 4	0	0	0

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	P	105	Total	C	N	O	S	0	0	0
			776	491	125	155	5			
3	Q	102	Total	C	N	O	S	0	0	0
			729	457	123	145	4			
3	R	108	Total	C	N	O	S	0	0	0
			826	520	138	163	5			
3	S	107	Total	C	N	O	S	0	0	0
			785	495	129	157	4			
3	T	110	Total	C	N	O	S	0	0	0
			839	534	137	163	5			

There are 20 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
K	-1	GLY	-	expression tag	UNP Q8WVK7
K	0	HIS	-	expression tag	UNP Q8WVK7
L	-1	GLY	-	expression tag	UNP Q8WVK7
L	0	HIS	-	expression tag	UNP Q8WVK7
M	-1	GLY	-	expression tag	UNP Q8WVK7
M	0	HIS	-	expression tag	UNP Q8WVK7
N	-1	GLY	-	expression tag	UNP Q8WVK7
N	0	HIS	-	expression tag	UNP Q8WVK7
O	-1	GLY	-	expression tag	UNP Q8WVK7
O	0	HIS	-	expression tag	UNP Q8WVK7
P	-1	GLY	-	expression tag	UNP Q8WVK7
P	0	HIS	-	expression tag	UNP Q8WVK7
Q	-1	GLY	-	expression tag	UNP Q8WVK7
Q	0	HIS	-	expression tag	UNP Q8WVK7
R	-1	GLY	-	expression tag	UNP Q8WVK7
R	0	HIS	-	expression tag	UNP Q8WVK7
S	-1	GLY	-	expression tag	UNP Q8WVK7
S	0	HIS	-	expression tag	UNP Q8WVK7
T	-1	GLY	-	expression tag	UNP Q8WVK7
T	0	HIS	-	expression tag	UNP Q8WVK7

3 Residue-property plots ⓘ

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

- Molecule 1: SPINDLE AND KINETOCHORE-ASSOCIATED PROTEIN 3

Chain 1:
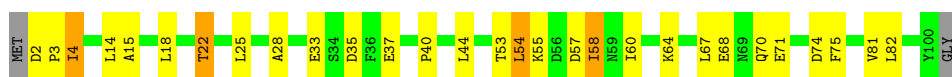
- Molecule 1: SPINDLE AND KINETOCHORE-ASSOCIATED PROTEIN 3

Chain 2:

- Molecule 1: SPINDLE AND KINETOCHORE-ASSOCIATED PROTEIN 3

Chain 3:

- Molecule 1: SPINDLE AND KINETOCHORE-ASSOCIATED PROTEIN 3

Chain 4:

- Molecule 1: SPINDLE AND KINETOCHORE-ASSOCIATED PROTEIN 3

Chain U:

- Molecule 1: SPINDLE AND KINETOCHORE-ASSOCIATED PROTEIN 3

Chain V:
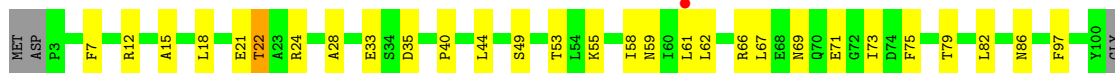
• Molecule 1: SPINDLE AND KINETOCHORE-ASSOCIATED PROTEIN 3

Chain W:

• Molecule 1: SPINDLE AND KINETOCHORE-ASSOCIATED PROTEIN 3

Chain X:

• Molecule 1: SPINDLE AND KINETOCHORE-ASSOCIATED PROTEIN 3

Chain Y:

• Molecule 1: SPINDLE AND KINETOCHORE-ASSOCIATED PROTEIN 3
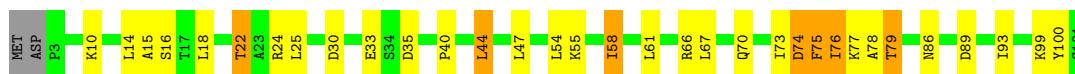
Chain Z:

• Molecule 2: SPINDLE AND KINETOCHORE-ASSOCIATED PROTEIN 1

Chain A:

• Molecule 2: SPINDLE AND KINETOCHORE-ASSOCIATED PROTEIN 1
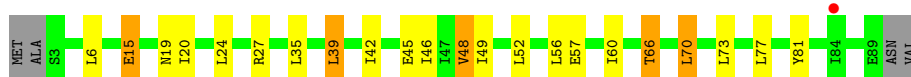
Chain B:

• Molecule 2: SPINDLE AND KINETOCHORE-ASSOCIATED PROTEIN 1

Chain C: 64%   18%   5%   13%

MET ALA S3 V13 N14 E15 K16 N19 I20 L24 L39 E45 I46 I47 V48 I49 L56 I60 Q63 E64 N67 L70 L73 C74 D80 Y81 LYS ASP ILE GLU HIS LEU LYS GLU ASN VAL

• Molecule 2: SPINDLE AND KINETOCHORE-ASSOCIATED PROTEIN 1

Chain D: 2%   66%   26%   5%

MET ALA S3 S4 D5 L6 V13 N14 E15 K16 N19 I20 R27 E32 P33 T34 L39 I42 E45 I46 I47 V48 I49 L53 L56 E59 I60 Q63 E64 N67 Y81 E85 H86 L87 K88 GLU ASN VAL

• Molecule 2: SPINDLE AND KINETOCHORE-ASSOCIATED PROTEIN 1

Chain E: 68%   23%   5%

MET ALA S3 L6 V13 N14 E15 N19 I20 L24 R27 E32 P33 T34 L35 L39 E45 I46 I47 V48 L53 L56 E57 I60 Q63 E64 Q65 T66 N67 L70 L77 I84 K88 GLU ASN VAL

• Molecule 2: SPINDLE AND KINETOCHORE-ASSOCIATED PROTEIN 1

Chain F: 62%   27%   10%

MET ALA S3 V13 N14 E15 K16 N19 I20 L24 S25 L26 R27 N28 E32 P33 T34 L35 L39 E45 I46 I47 V48 I49 L52 L53 L56 I60 Q63 E64 N67 N68 C74 I84 GLU HIS LEU LYS GLU ASN VAL
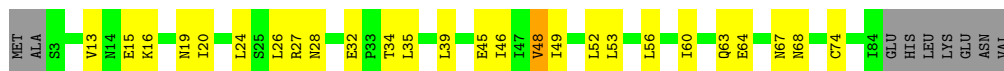
• Molecule 2: SPINDLE AND KINETOCHORE-ASSOCIATED PROTEIN 1

Chain G: 76%   16%   5%

MET ALA S3 V13 N14 E15 N19 I20 L24 R27 Q31 E32 P33 T34 L39 E45 I46 I47 V48 L53 I60 T66 N67 K88 GLU ASN VAL
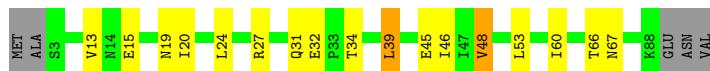
• Molecule 2: SPINDLE AND KINETOCHORE-ASSOCIATED PROTEIN 1

Chain H: %   63%   24%   9%

MET ALA S3 L6 E15 N19 I20 L24 C29 L35 L39 I42 E45 I46 I47 V48 I49 L56 E57 I60 Q61 Y62 Q63 E64 Q65 T66 N67 N68 S69 E72 L73 C74 L77 E85 HIS LEU LYS GLU ASN VAL
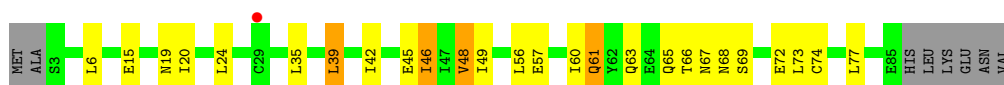
• Molecule 2: SPINDLE AND KINETOCHORE-ASSOCIATED PROTEIN 1

Chain I: %   75%   15%   5%

MET ALA S3 E15 N19 I20 L26 L39 E45 I46 I47 V48 I49 I60 Q63 E64 Q65 N68 S69 L70 C74 L77 D83 H86 L87 K88 GLU ASN VAL

• Molecule 2: SPINDLE AND KINETOCHORE-ASSOCIATED PROTEIN 1

Chain J:

MET ALA S3 L6 V13 N14 E15 N19 I20 L24 R27 E32 P33 T34 L35 L39 E45 I46 I47 V48 I49 L53 L56 E57 L58 E59 I60 Q63 E64 Q65 T66 N67 D83 I84 E85 K88 GLU ASN VAL

• Molecule 3: SPINDLE AND KINETOCHORE-ASSOCIATED PROTEIN 2

Chain K:

GLY HIS M1 L8 A15 E16 S17 D20 E29 I30 N33 H34 PRO ASP SER ALA SER E40 P43 K48 E49 L50 S51 V52 I53 K54 Y57 Q58 T59 L60 F64 V69 E70 Q71 I78 V82 K83 N87 M88 L92 Q93 L98 E99 L100

E106 T109 Q113 PHE LYS PHE HIS MET PRO ASP LEU

• Molecule 3: SPINDLE AND KINETOCHORE-ASSOCIATED PROTEIN 2

Chain L:

GLY H0 A3 M11 A15 D18 Q23 L26 E29 I30 N33 HIS PRO ASP SER ALA SER GLU LYS N42 P43 V44 L50 S51 S55 R56 Y57 L60 F64 V69 E70 Q71 I78 C79 A80 T81 V82 K83 K84 T85 M86 N87 T96 L103

T104 K105 T109 Q113 PHE LYS PHE HIS MET PRO ASP LEU

• Molecule 3: SPINDLE AND KINETOCHORE-ASSOCIATED PROTEIN 2

Chain M:

G-1 H0 M1 L8 A15 E16 S17 D18 L19 D20 Q23 Y24 R25 E29 H34 PRO ASP SER ALA SER GLU K41 N42 P43 V44 T45 L46 L50 S51 T59 L60 R63 F64 K65 P66 V69 E70 S74 I78 C79 V82 T85 D97 L98 E99 L100

S101 P102 L103 T104 K105 Q113 PHE LYS PHE HIS MET PRO ASP LEU

• Molecule 3: SPINDLE AND KINETOCHORE-ASSOCIATED PROTEIN 2

Chain N:

GLY HIS M1 L8 F12 A15 E16 S17 D20 Y21 H34 PRO ASP SER ALA SER GLU K41 L50 S51 K54 Y57 Q58 T59 K65 P66 V67 E70 Q71 S76 R77 I78 T81 V82 T85 I89 T96 ASP LEU GLU LEU S101 P102 L103 T104

T109 F114 LYS PHE HIS MET PRO ASP LEU

• Molecule 3: SPINDLE AND KINETOCHORE-ASSOCIATED PROTEIN 2

Chain O:

GLY H0 M1 E2 A3 E4 L8 M11 K14 A15 E16 S17 D18 L19 D20 L26 E29 H34 PRO ASP SER ALA SER GLU K41 N42 P43 V44 T45 L46 E49 L50 S51 V52 I53 R56 Y57 Q58 T59 F64 V67 A68 V69 E70 T81 T85 M86

I89 L92 E99 L100 S101 P102 L103 T104 K105 E106 E107 K108 T109 F114 LYS PHE HIS MET PRO ASP LEU

● Molecule 3: SPINDLE AND KINETOCHORE-ASSOCIATED PROTEIN 2

Chain P:

GLY H0 M1 K14 A15 E16 S17 D18 L19 D20 Y21 E29 I30 K31 T32 N33 H34 PRO ASP SER ALA SER GLU LYS N42 P43 V44 T45 L46 E49 L50 S51 V52 I53 K54 S55 Q58 T59 L60 F64 V69 E70 Q71 S74 I78 I89 L92 T96 D97

LEU GLU L100 L103 T104 K105 E106 T109 Q113 PHE LYS PHE HIS MET PRO ASP LEU

● Molecule 3: SPINDLE AND KINETOCHORE-ASSOCIATED PROTEIN 2

Chain Q:

GLY H0 M1 L8 K14 A15 E16 S17 D18 L19 D20 R25 L26 E29 N33 H34 PRO ASP SER ALA SER GLU LYS ASN PRO V44 L50 S51 V52 I53 R56 T59 L60 R63 F64 K65 P66 V69 E70 T81 I89 L100 SER PRO L103 E112

GLN PHE LYS PHE HIS MET PRO ASP LEU

● Molecule 3: SPINDLE AND KINETOCHORE-ASSOCIATED PROTEIN 2

Chain R:

GLY H0 E4 L8 M11 K14 A15 Q23 L26 E27 Y28 T32 N33 H34 PRO ASP SER ALA SER GLU LYS N42 P43 V44 T45 L46 L47 L50 S51 T59 L60 Y61 F64 K65 P66 V69 E70 Q71 S74 R77 I78 T81 V82 K83 N87

M88 I89 T96 D97 L98 E99 L100 L103 T104 K105 T109 Q113 F114 LYS PHE HIS MET PRO ASP LEU

● Molecule 3: SPINDLE AND KINETOCHORE-ASSOCIATED PROTEIN 2

Chain S:

G-1 E2 A3 L8 E9 L10 M11 A15 E16 S17 D18 L19 D20 Q23 Y28 E29 I30 K31 T32 N33 H34 PRO ASP SER ALA SER GLU LYS ASN P43 V44 L50 S51 V52 I53 K54 Q58 K65 V69 E70 S74 I78 C79 V82 T85 M86

I89 T96 L103 T104 K105 E106 E112 Q113 PHE LYS PHE HIS MET PRO ASP LEU

● Molecule 3: SPINDLE AND KINETOCHORE-ASSOCIATED PROTEIN 2

Chain T:

GLY H0 M1 L8 F12 A15 E16 S17 D20 Y21 I22 Q23 Y28 E29 T32 N33 H34 PRO ASP SER ALA SER GLU LYS N42 P43 V44 T45 L46 L50 S55 R56 Y57 Q58 T59 L60 Y61 A62 R63 F64 E70 K75 C79 M86 N87 M88 I89

T96 D97 I98 E99 L100 S101 P102 L103 T104 K105 K108 T109 E112 K115 F116 HIS MET PRO ASP LEU

4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	174.40Å 193.12Å 198.27Å 90.00° 95.21° 90.00°	Depositor
Resolution (Å)	69.03 – 3.32 81.12 – 3.32	Depositor EDS
% Data completeness (in resolution range)	94.2 (69.03-3.32) 94.4 (81.12-3.32)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.52 (at 3.33Å)	Xtriage
Refinement program	PHENIX (PHENIX.REFINE)	Depositor
R, R_{free}	0.234 , 0.280 0.233 , 0.277	Depositor DCC
R_{free} test set	4579 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å^2)	104.4	Xtriage
Anisotropy	0.261	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.27 , 96.3	EDS
L-test for twinning[2]	$< |L| > = 0.43$, $< L^2 > = 0.26$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	21691	wwPDB-VP
Average B, all atoms (Å^2)	140.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.87% of the height of the origin peak. No significant pseudotranslation is detected.*

---

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

5 Model quality ⓘ

5.1 Standard geometry ⓘ

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	1	0.39	0/747	0.64	0/1012
1	2	0.46	0/796	0.70	0/1071
1	3	0.38	0/745	0.65	0/1008
1	4	0.42	0/783	0.65	0/1050
1	U	0.41	0/782	0.69	0/1055
1	V	0.40	0/733	0.64	0/993
1	W	0.41	0/734	0.68	1/991 (0.1%)
1	X	0.42	0/738	0.64	0/1000
1	Y	0.48	0/772	0.75	1/1039 (0.1%)
1	Z	0.42	0/729	0.71	1/984 (0.1%)
2	A	0.37	0/590	0.52	0/800
2	B	0.37	0/659	0.53	0/894
2	C	0.38	0/598	0.51	0/811
2	D	0.45	0/649	0.54	0/879
2	E	0.38	0/632	0.53	0/858
2	F	0.40	0/587	0.53	0/799
2	G	0.40	0/639	0.51	0/867
2	H	0.47	0/636	0.61	0/862
2	I	0.40	0/648	0.47	0/879
2	J	0.39	0/642	0.50	0/873
3	K	0.40	0/822	0.63	1/1117 (0.1%)
3	L	0.42	0/788	0.58	0/1074
3	M	0.40	0/809	0.59	0/1101
3	N	0.40	0/778	0.57	1/1056 (0.1%)
3	O	0.50	0/850	0.70	0/1152
3	P	0.42	0/785	0.61	0/1070
3	Q	0.43	0/736	0.58	0/1001
3	R	0.46	0/836	0.69	0/1134
3	S	0.42	0/794	0.63	0/1081
3	T	0.42	0/851	0.67	2/1155 (0.2%)
All	All	0.42	0/21888	0.62	7/29666 (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	K	0	1
3	R	0	1
All	All	0	2

There are no bond length outliers.

The worst 5 of 7 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	Z	67	LEU	CA-CB-CG	5.96	129.02	115.30
3	T	98	LEU	CA-CB-CG	5.89	128.86	115.30
3	K	98	LEU	CA-CB-CG	5.69	128.39	115.30
3	T	103	LEU	CA-CB-CG	5.37	127.64	115.30
1	W	67	LEU	CA-CB-CG	5.07	126.96	115.30

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
3	K	99	GLU	Peptide
3	R	99	GLU	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	1	740	0	677	20	0
1	2	787	0	758	32	0
1	3	738	0	691	18	0
1	4	775	0	763	29	0
1	U	773	0	737	33	0
1	V	725	0	643	24	0
1	W	727	0	663	26	0
1	X	729	0	648	18	0
1	Y	763	0	737	35	1

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	Z	723	0	672	32	0
2	A	588	0	568	21	0
2	B	656	0	627	19	0
2	C	595	0	577	20	0
2	D	646	0	614	26	0
2	E	630	0	592	24	0
2	F	585	0	545	26	0
2	G	637	0	605	15	0
2	H	633	0	608	24	1
2	I	645	0	617	14	0
2	J	639	0	602	19	0
3	K	813	0	749	26	0
3	L	779	0	707	24	0
3	M	800	0	724	29	0
3	N	770	0	701	24	0
3	O	840	0	793	41	0
3	P	776	0	705	36	0
3	Q	729	0	621	23	0
3	R	826	0	779	43	0
3	S	785	0	715	31	0
3	T	839	0	784	34	0
All	All	21691	0	20222	540	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 13.

The worst 5 of 540 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:B:52:LEU:HB3	1:W:58:ILE:HD11	1.45	0.98
3:M:25:ARG:HH12	3:P:59:THR:HG21	1.36	0.91
1:Z:61:LEU:HD23	2:H:56:LEU:HD21	1.54	0.88
2:D:53:LEU:HD21	1:Y:58:ILE:HD11	1.57	0.87
2:F:49:ILE:HD11	1:Z:54:LEU:HD22	1.56	0.85

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:H:61:GLN:OE1	1:Y:70:GLN:NE2[3_445]	2.04	0.16

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	
1	1	97/101 (96%)	89 (92%)	8 (8%)	0	100	100
1	2	98/101 (97%)	88 (90%)	10 (10%)	0	100	100
1	3	96/101 (95%)	86 (90%)	10 (10%)	0	100	100
1	4	96/101 (95%)	85 (88%)	11 (12%)	0	100	100
1	U	98/101 (97%)	87 (89%)	11 (11%)	0	100	100
1	V	96/101 (95%)	85 (88%)	11 (12%)	0	100	100
1	W	96/101 (95%)	89 (93%)	7 (7%)	0	100	100
1	X	96/101 (95%)	87 (91%)	9 (9%)	0	100	100
1	Y	97/101 (96%)	85 (88%)	12 (12%)	0	100	100
1	Z	95/101 (94%)	86 (90%)	9 (10%)	0	100	100
2	A	77/91 (85%)	76 (99%)	1 (1%)	0	100	100
2	B	85/91 (93%)	83 (98%)	2 (2%)	0	100	100
2	C	77/91 (85%)	75 (97%)	2 (3%)	0	100	100
2	D	84/91 (92%)	82 (98%)	2 (2%)	0	100	100
2	E	84/91 (92%)	83 (99%)	1 (1%)	0	100	100
2	F	80/91 (88%)	78 (98%)	2 (2%)	0	100	100
2	G	84/91 (92%)	83 (99%)	1 (1%)	0	100	100
2	H	81/91 (89%)	77 (95%)	4 (5%)	0	100	100
2	I	84/91 (92%)	83 (99%)	1 (1%)	0	100	100
2	J	84/91 (92%)	82 (98%)	2 (2%)	0	100	100
3	K	104/123 (85%)	95 (91%)	8 (8%)	1 (1%)	15	47
3	L	102/123 (83%)	96 (94%)	6 (6%)	0	100	100
3	M	103/123 (84%)	96 (93%)	7 (7%)	0	100	100
3	N	98/123 (80%)	96 (98%)	2 (2%)	0	100	100
3	O	105/123 (85%)	90 (86%)	15 (14%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
3	P	99/123 (80%)	93 (94%)	6 (6%)	0	100	100
3	Q	96/123 (78%)	93 (97%)	3 (3%)	0	100	100
3	R	104/123 (85%)	94 (90%)	10 (10%)	0	100	100
3	S	103/123 (84%)	96 (93%)	7 (7%)	0	100	100
3	T	106/123 (86%)	96 (91%)	10 (9%)	0	100	100
All	All	2805/3150 (89%)	2614 (93%)	190 (7%)	1 (0%)	100	100

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	K	30	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	1	74/92 (80%)	67 (90%)	7 (10%)	8	30
1	2	84/92 (91%)	77 (92%)	7 (8%)	11	37
1	3	76/92 (83%)	69 (91%)	7 (9%)	9	32
1	4	84/92 (91%)	78 (93%)	6 (7%)	14	44
1	U	81/92 (88%)	77 (95%)	4 (5%)	25	57
1	V	70/92 (76%)	67 (96%)	3 (4%)	29	61
1	W	71/92 (77%)	65 (92%)	6 (8%)	10	36
1	X	70/92 (76%)	65 (93%)	5 (7%)	14	44
1	Y	79/92 (86%)	72 (91%)	7 (9%)	9	33
1	Z	73/92 (79%)	68 (93%)	5 (7%)	16	46
2	A	66/87 (76%)	60 (91%)	6 (9%)	9	32
2	B	73/87 (84%)	66 (90%)	7 (10%)	8	30
2	C	67/87 (77%)	61 (91%)	6 (9%)	9	33

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	D	71/87 (82%)	68 (96%)	3 (4%)	30	61
2	E	68/87 (78%)	62 (91%)	6 (9%)	10	34
2	F	60/87 (69%)	55 (92%)	5 (8%)	11	37
2	G	70/87 (80%)	66 (94%)	4 (6%)	20	52
2	H	72/87 (83%)	64 (89%)	8 (11%)	6	24
2	I	72/87 (83%)	65 (90%)	7 (10%)	8	30
2	J	70/87 (80%)	64 (91%)	6 (9%)	10	36
3	K	81/114 (71%)	70 (86%)	11 (14%)	3	16
3	L	74/114 (65%)	67 (90%)	7 (10%)	8	30
3	M	76/114 (67%)	66 (87%)	10 (13%)	4	17
3	N	73/114 (64%)	68 (93%)	5 (7%)	16	46
3	O	86/114 (75%)	78 (91%)	8 (9%)	9	31
3	P	76/114 (67%)	69 (91%)	7 (9%)	9	32
3	Q	63/114 (55%)	58 (92%)	5 (8%)	12	39
3	R	84/114 (74%)	77 (92%)	7 (8%)	11	37
3	S	75/114 (66%)	67 (89%)	8 (11%)	6	25
3	T	84/114 (74%)	76 (90%)	8 (10%)	8	30
All	All	2223/2930 (76%)	2032 (91%)	191 (9%)	10	36

5 of 191 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	N	67	VAL
3	R	100	LEU
3	O	46	LEU
3	P	69	VAL
3	S	82	VAL

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

Mol	Chain	Res	Type
3	Q	23	GLN
3	T	23	GLN
3	Q	33	ASN
3	R	93	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	G	40	ASN

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 ⓘ

The following chains have linkage breaks:

Mol	Chain	Number of breaks
3	M	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	M	96:THR	C	97:ASP	N	3.27

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	1	99/101 (98%)	-0.19	0	100	100	78, 143, 198, 211	0
1	2	100/101 (99%)	-0.20	1 (1%)	82	83	88, 122, 176, 211	0
1	3	98/101 (97%)	-0.27	0	100	100	79, 143, 196, 243	0
1	4	98/101 (97%)	-0.12	0	100	100	82, 125, 196, 242	0
1	U	100/101 (99%)	-0.37	0	100	100	77, 136, 173, 194	0
1	V	98/101 (97%)	-0.03	1 (1%)	82	83	79, 155, 211, 241	0
1	W	98/101 (97%)	-0.32	0	100	100	76, 147, 187, 203	0
1	X	98/101 (97%)	-0.24	1 (1%)	82	83	82, 148, 196, 239	0
1	Y	99/101 (98%)	-0.24	0	100	100	75, 115, 161, 195	0
1	Z	97/101 (96%)	-0.05	1 (1%)	82	83	81, 151, 207, 236	0
2	A	79/91 (86%)	-0.11	1 (1%)	77	77	99, 142, 224, 229	0
2	B	87/91 (95%)	-0.34	1 (1%)	80	81	90, 142, 207, 223	0
2	C	79/91 (86%)	-0.22	0	100	100	93, 137, 222, 250	0
2	D	86/91 (94%)	-0.09	2 (2%)	60	59	84, 129, 190, 263	0
2	E	86/91 (94%)	-0.41	0	100	100	95, 140, 191, 245	0
2	F	82/91 (90%)	-0.26	0	100	100	98, 142, 194, 224	0
2	G	86/91 (94%)	-0.02	0	100	100	91, 140, 248, 277	0
2	H	83/91 (91%)	-0.24	1 (1%)	79	79	73, 127, 165, 227	0
2	I	86/91 (94%)	-0.16	1 (1%)	79	79	90, 144, 210, 281	0
2	J	86/91 (94%)	-0.24	0	100	100	84, 136, 207, 272	0
3	K	108/123 (87%)	-0.46	1 (0%)	84	85	81, 144, 199, 224	0
3	L	106/123 (86%)	-0.02	4 (3%)	40	39	85, 163, 229, 248	0
3	M	109/123 (88%)	-0.31	1 (0%)	84	85	85, 139, 223, 292	0
3	N	104/123 (84%)	-0.27	2 (1%)	66	65	83, 139, 224, 245	0

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å²)	Q<0.9
3	O	109/123 (88%)	-0.38	0	100	100	83, 118, 159, 186	0
3	P	105/123 (85%)	-0.23	1 (0%)	82	83	79, 140, 217, 250	0
3	Q	102/123 (82%)	-0.33	0	100	100	81, 133, 214, 275	0
3	R	108/123 (87%)	-0.25	3 (2%)	53	51	75, 117, 176, 199	0
3	S	107/123 (86%)	-0.28	3 (2%)	53	51	89, 141, 211, 218	0
3	T	110/123 (89%)	-0.40	1 (0%)	84	85	81, 123, 187, 203	0
All	All	2893/3150 (91%)	-0.24	26 (0%)	84	85	73, 136, 206, 292	0

The worst 5 of 26 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	R	43	PRO	4.2
3	P	92	LEU	3.6
3	S	-1	GLY	3.5
2	A	73	LEU	2.9
3	S	53	ILE	2.8

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 monosaccharides 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.