wwPDB EM Map/Model Validation Report (i)

Apr 10, 2016 – 02:22 PM BST

PDB ID	:	5A31
EMDB ID:	:	EMD-2925
Title	:	Structure of the human APC-Cdh1-Hsl1-UbcH10 complex.
Authors	:	Chang, L.; Zhang, Z.; Yang, J.; Mclaughlin, S.H.; Barford, D.
Deposited on	:	2015-05-26
Resolution	:	4.30 Å(reported)

This is a wwPDB EM Map/Model Validation Report for a publicly released PDB/EMDB entry.
For rigid body fitted models, validation errors reported here could
stem from errors in the original structure(s) used in the fitting.
We welcome your comments at validation@mail.wwpdb.org
A user guide is available at
http://wwpdb.org/validation/2016/EMValidationReportHelp

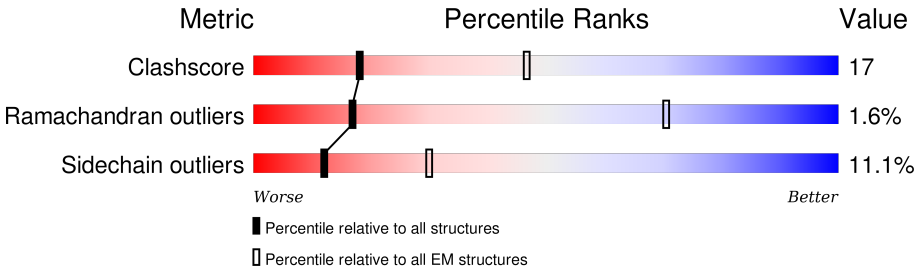
MolProbity	:	4.02b-467
Mogul	:	unknown
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk27241

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
ELECTRON MICROSCOPY

The reported resolution of this entry is 4.30 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)	EM structures (#Entries)
Clashscore	114402	924
Ramachandran outliers	111179	726
Sidechain outliers	111093	686

The table below summarises the geometric issues observed across the polymeric chains. The red, orange, yellow and green segments on the 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 <=5%

Mol	Chain	Length	Quality of chain
1	A	1441	61% 34% 5%
2	B	84	51% 36% 8% 5%
3	C	597	60% 23% • • 12%
3	P	597	57% 21% • 18%
4	D	121	32% 12% • 55%
5	E	110	35% 15% • 49%
6	F	824	44% 15% • 40%
6	H	824	42% 14% • 41%
7	G	85	25% • • 71%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
7	W	85	21% 7% 71%
8	I	808	67% 21% 10%
9	J	620	55% 21% 5% 19%
10	K	620	53% 23% 20%
11	L	185	61% 31% 5%
12	M	74	45% 28% 7% 20%
13	N	703	55% 36% 7%
14	O	755	60% 25% 6% 9%
15	Q	162	85% 15%
16	R	386	78% 17% 5%
17	T	21	90% 10%
18	U	24	88% 13%
19	V	13	69% 23% 8%
20	X	599	50% 28% 19%
20	Y	599	53% 26% 17%

2 Entry composition

There are 21 unique types of molecules in this entry. The entry contains 67685 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, 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 ANAPHASE-PROMOTING COMPLEX SUBUNIT 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	1441	Total	C	N	O	S	0	0
			10950	7046	1853	1977	74		

- Molecule 2 is a protein called ANAPHASE-PROMOTING COMPLEX SUBUNIT 11.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	B	84	Total	C	N	O	S	1	0
			650	418	117	98	17		

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	13	LEU	THR	CONFLICT	UNP Q9NYG5

- Molecule 3 is a protein called CELL DIVISION CYCLE PROTEIN 23 HOMOLOG.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	C	524	Total	C	N	O	S	0	0
			4305	2774	726	781	24		
3	P	491	Total	C	N	O	S	0	0
			4042	2611	678	729	24		

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	161	LEU	LYS	CONFLICT	UNP Q9UJX2
P	161	LEU	LYS	CONFLICT	UNP Q9UJX2

- Molecule 4 is a protein called ANAPHASE-PROMOTING COMPLEX SUBUNIT 15.

Mol	Chain	Residues	Atoms				AltConf	Trace
4	D	55	Total	C	N	O	0	0
			437	277	73	87		

There are 5 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	1	MET	-	EXPRESSION TAG	UNP P60006
D	2	SER	-	EXPRESSION TAG	UNP P60006
D	3	THR	-	EXPRESSION TAG	UNP P60006
D	4	LEU	-	EXPRESSION TAG	UNP P60006
D	5	TYR	-	EXPRESSION TAG	UNP P60006

- Molecule 5 is a protein called ANAPHASE-PROMOTING COMPLEX SUBUNIT 16.

Mol	Chain	Residues	Atoms					AltConf	Trace
5	E	56	Total	C	N	O	S	0	0
			450	290	74	85	1		

- Molecule 6 is a protein called ANAPHASE-PROMOTING COMPLEX SUBUNIT 3.

Mol	Chain	Residues	Atoms					AltConf	Trace
6	F	498	Total	C	N	O	S	0	0
			3923	2514	664	719	26		
6	H	483	Total	C	N	O	S	0	0
			3853	2473	650	704	26		

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
F	475	SER	ALA	CONFLICT	UNP P30260
F	484	SER	ALA	CONFLICT	UNP P30260
H	475	SER	ALA	CONFLICT	UNP P30260
H	484	SER	ALA	CONFLICT	UNP P30260

- Molecule 7 is a protein called ANAPHASE-PROMOTING COMPLEX SUBUNIT CDC26.

Mol	Chain	Residues	Atoms					AltConf	Trace
7	G	25	Total	C	N	O	S	0	0
			211	133	40	37	1		
7	W	25	Total	C	N	O	S	0	0
			213	133	40	39	1		

- Molecule 8 is a protein called ANAPHASE-PROMOTING COMPLEX SUBUNIT 4.

Mol	Chain	Residues	Atoms					AltConf	Trace
8	I	730	Total	C	N	O	S	0	0
			5709	3660	950	1066	33		

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
I	430	ASP	GLU	CONFLICT	UNP Q9UJX5

- Molecule 9 is a protein called ANAPHASE-PROMOTING COMPLEX SUBUNIT 6.

Mol	Chain	Residues	Atoms					AltConf	Trace
9	J	504	Total	C	N	O	S	0	0
			4047	2602	685	735	25		

There are 5 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
J	225	ASN	ASP	CONFLICT	UNP Q13042
J	228	GLU	GLN	CONFLICT	UNP Q13042
J	229	LYS	GLU	CONFLICT	UNP Q13042
J	347	ALA	GLU	CONFLICT	UNP Q13042
J	524	ALA	GLU	CONFLICT	UNP Q13042

- Molecule 10 is a protein called ANAPHASE-PROMOTING COMPLEX SUBUNIT 6.

Mol	Chain	Residues	Atoms					AltConf	Trace
10	K	493	Total	C	N	O	S	0	0
			3988	2565	673	726	24		

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
K	228	GLU	GLN	CONFLICT	UNP Q13042
K	229	LYS	GLU	CONFLICT	UNP Q13042
K	265	LYS	ALA	CONFLICT	UNP Q13042

- Molecule 11 is a protein called ANAPHASE-PROMOTING COMPLEX SUBUNIT 10.

Mol	Chain	Residues	Atoms					AltConf	Trace
11	L	182	Total	C	N	O	S	0	0
			1435	898	263	268	6		

- Molecule 12 is a protein called ANAPHASE-PROMOTING COMPLEX SUBUNIT 13.

Mol	Chain	Residues	Atoms					AltConf	Trace
12	M	59	Total	C	N	O	S	0	0
			493	310	79	102	2		

- Molecule 13 is a protein called ANAPHASE-PROMOTING COMPLEX SUBUNIT 2.

Mol	Chain	Residues	Atoms					AltConf	Trace
13	N	703	Total	C	N	O	S	0	0
			5400	3438	968	969	25		

- Molecule 14 is a protein called ANAPHASE-PROMOTING COMPLEX SUBUNIT 5.

Mol	Chain	Residues	Atoms					AltConf	Trace
14	O	685	Total	C	N	O	S	0	0
			5396	3439	939	991	27		

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
O	42	SER	ASN	CONFLICT	UNP Q9UJX4
O	55	VAL	MET	CONFLICT	UNP Q9UJX4
O	63	GLN	LEU	CONFLICT	UNP Q9UJX4
O	75	VAL	LEU	CONFLICT	UNP Q9UJX4
O	79	LEU	TYR	CONFLICT	UNP Q9UJX4
O	164	SER	ASN	CONFLICT	UNP Q9UJX4
O	165	ASP	GLY	CONFLICT	UNP Q9UJX4
O	167	ASN	LYS	CONFLICT	UNP Q9UJX4

- Molecule 15 is a protein called FUSION PROTEIN - UBIQUITIN-CONJUGATING ENZYME E2 C, UBIQUITIN-CONJUGATING ENZYME E2 S.

Mol	Chain	Residues	Atoms					AltConf	Trace
15	Q	162	Total	C	N	O	S	1	0
			1227	789	204	229	5		

- Molecule 16 is a protein called THE ANAPHASE-PROMOTING COMPLEX CHAIN R.

Mol	Chain	Residues	Atoms					AltConf	Trace
16	R	386	Total	C	N	O	S	0	0
			2990	1884	530	564	12		

- Molecule 17 is a protein called THE ANAPHASE-PROMOTING COMPLEX CHAIN T.

Mol	Chain	Residues	Atoms				AltConf	Trace
17	T	21	Total	C	N	O	0	0
			109	65	22	22		

- Molecule 18 is a protein called THE ANAPHASE-PROMOTING COMPLEX CHAIN U.

Mol	Chain	Residues	Atoms				AltConf	Trace
18	U	24	Total	C	N	O	0	0
			120	72	24	24		

- Molecule 19 is a protein called THE ANAPHASE-PROMOTING COMPLEX CHAIN V.

Mol	Chain	Residues	Atoms				AltConf	Trace
19	V	13	Total	C	N	O	0	0
			99	64	19	16		

- Molecule 20 is a protein called ANAPHASE-PROMOTING COMPLEX SUBUNIT 7.

Mol	Chain	Residues	Atoms					AltConf	Trace
20	X	484	Total	C	N	O	S	0	0
			3770	2394	650	705	21		
20	Y	496	Total	C	N	O	S	0	0
			3865	2450	667	725	23		

- Molecule 21 is ZINC ION (three-letter code: ZN) (formula: Zn).

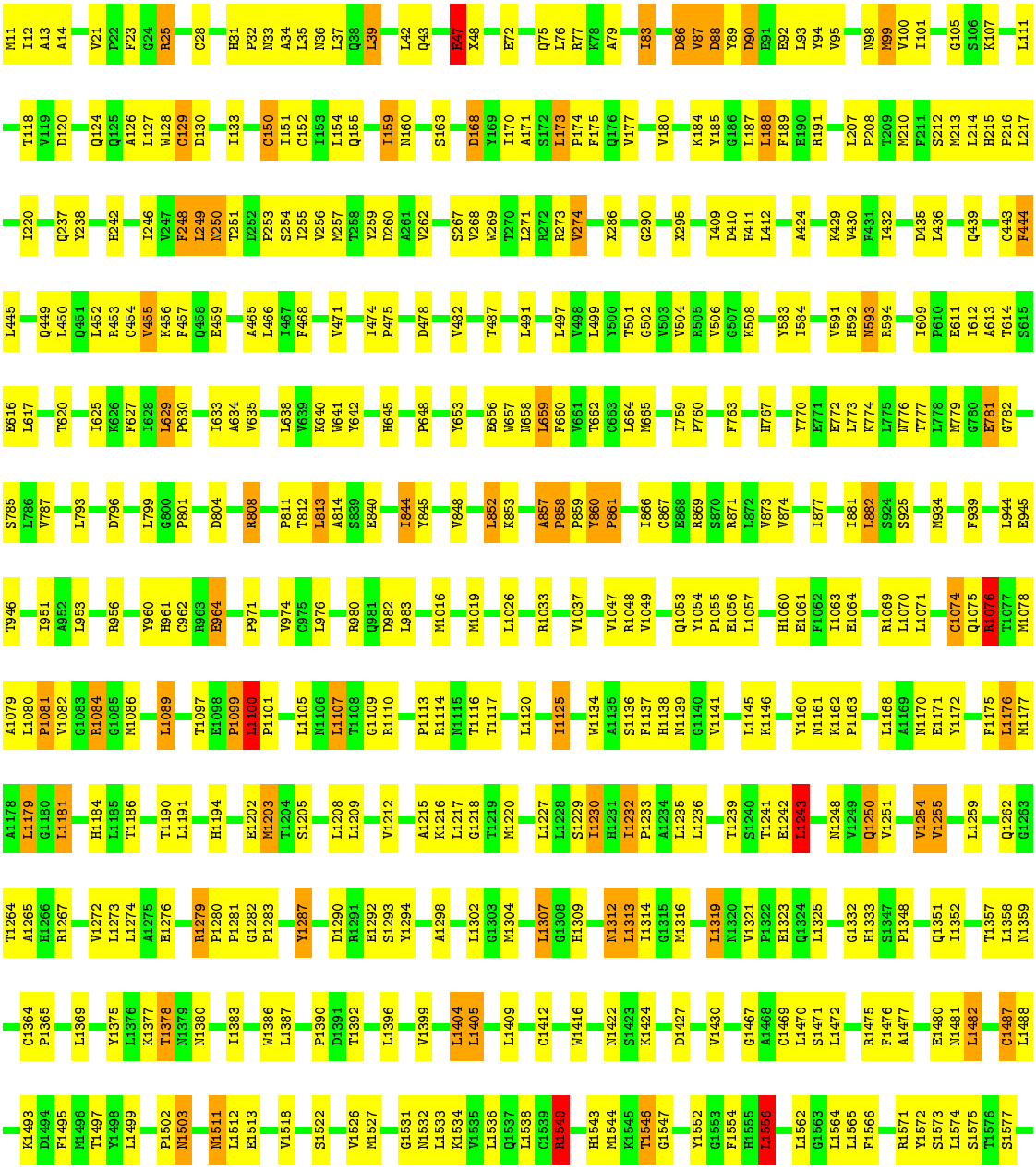
Mol	Chain	Residues	Atoms		AltConf
21	B	3	Total	Zn	0
			3	3	

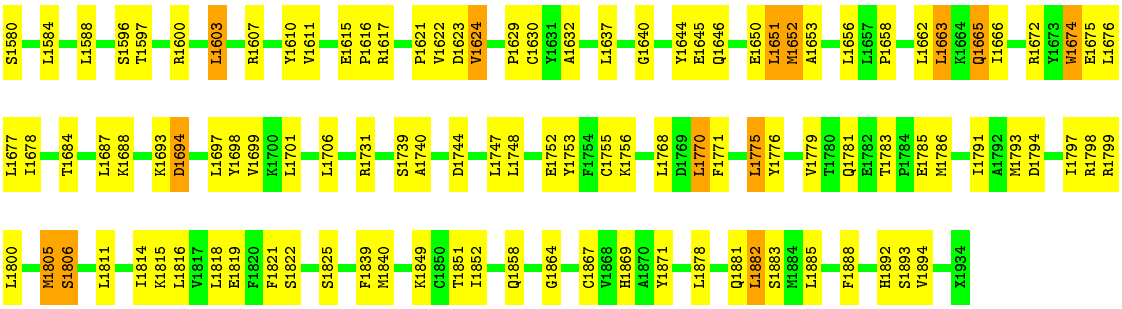
3 Residue-property plots (i)

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. 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: ANAPHASE-PROMOTING COMPLEX SUBUNIT 1

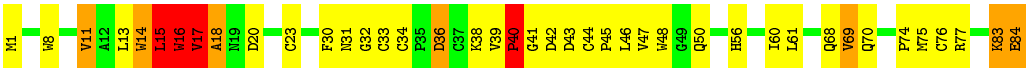
Chain A: 61% 34% 5%





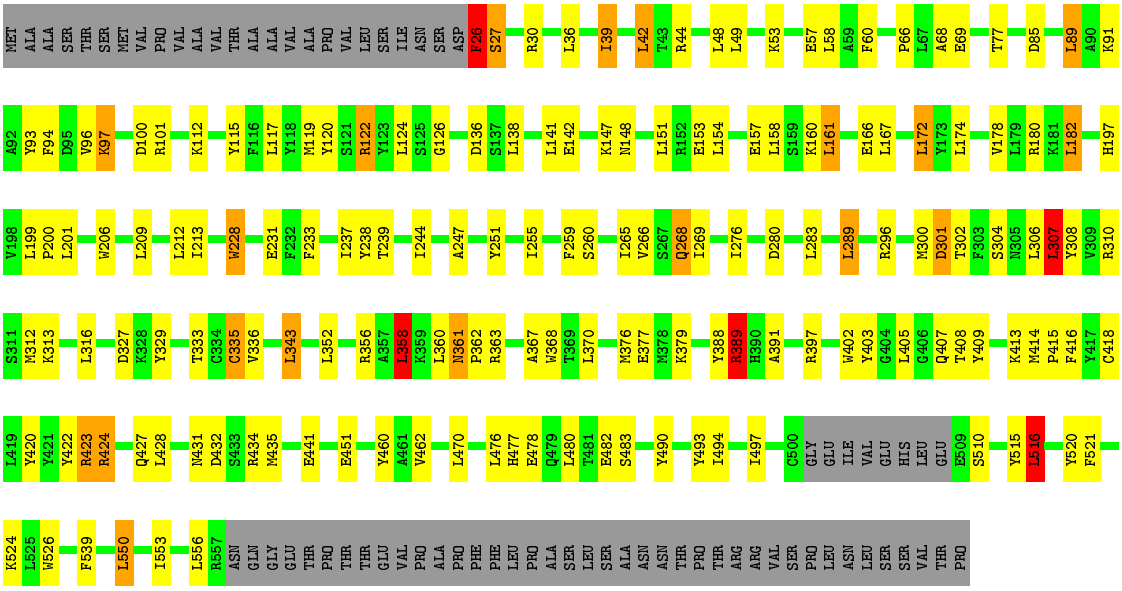
• Molecule 2: ANAPHASE-PROMOTING COMPLEX SUBUNIT 11

Chain B:



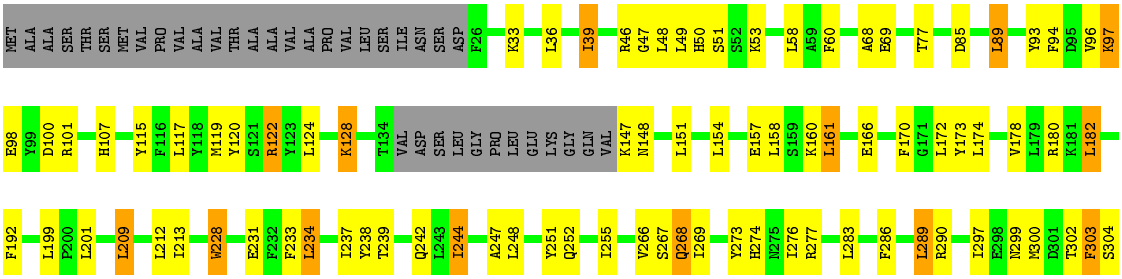
• Molecule 3: CELL DIVISION CYCLE PROTEIN 23 HOMOLOG

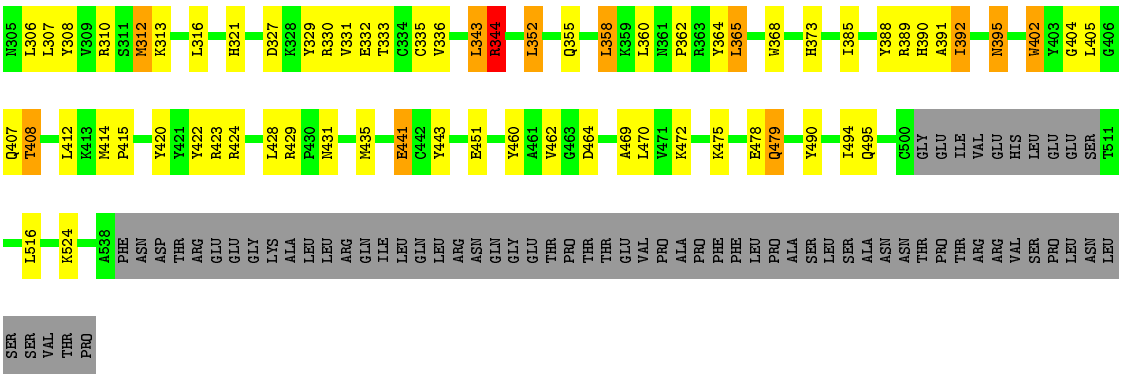
Chain C:



• Molecule 3: CELL DIVISION CYCLE PROTEIN 23 HOMOLOG

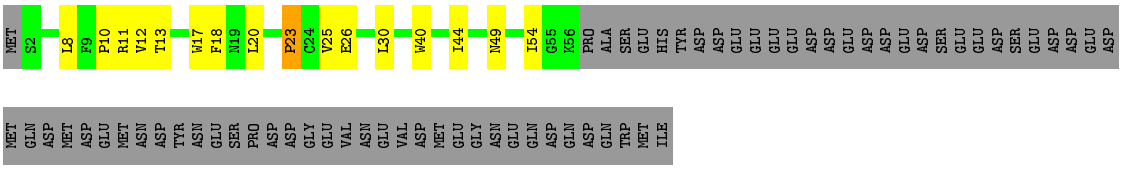
Chain P:





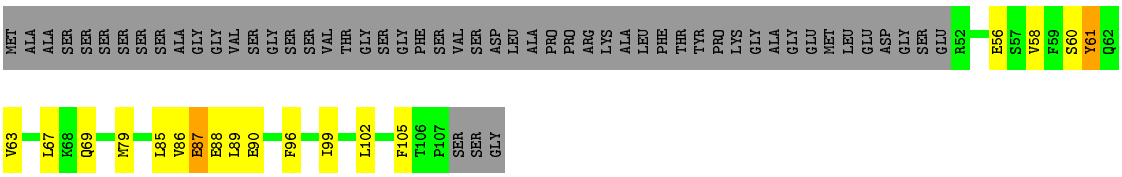
- Molecule 4: ANAPHASE-PROMOTING COMPLEX SUBUNIT 15

Chain D:



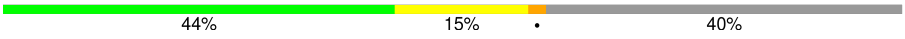
- Molecule 5: ANAPHASE-PROMOTING COMPLEX SUBUNIT 16

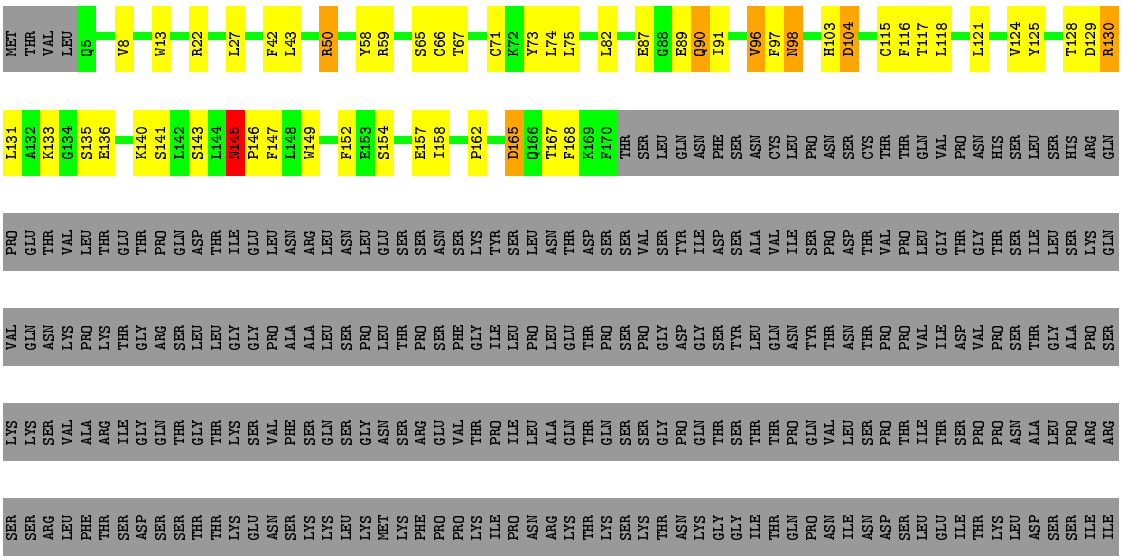
Chain E:

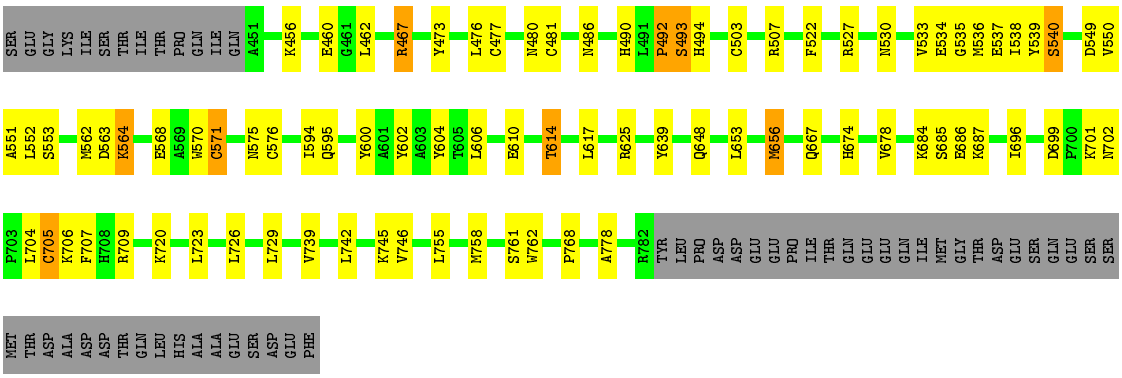


- Molecule 6: ANAPHASE-PROMOTING COMPLEX SUBUNIT 3

Chain F:

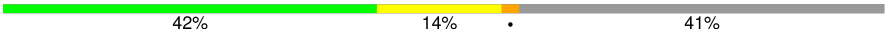


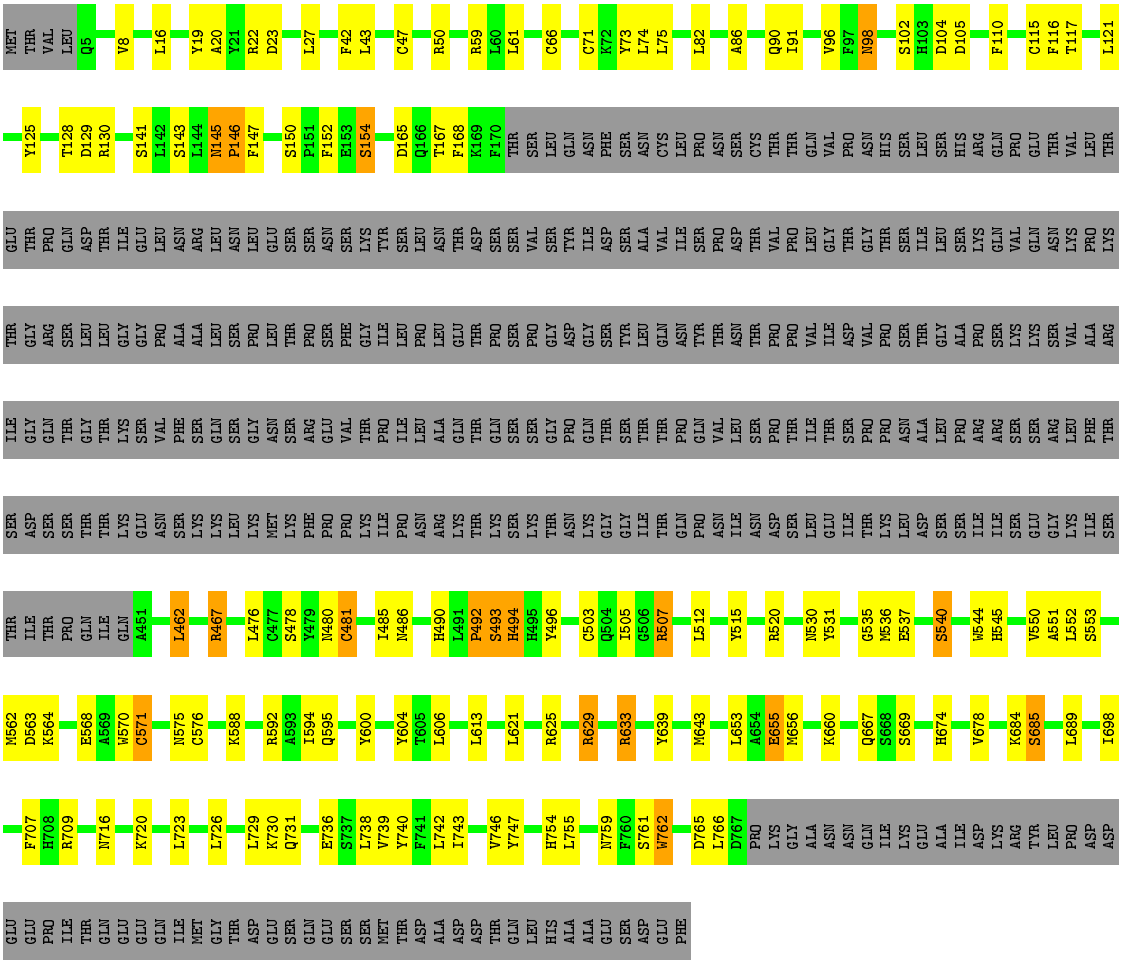




• Molecule 6: ANAPHASE-PROMOTING COMPLEX SUBUNIT 3

Chain H:



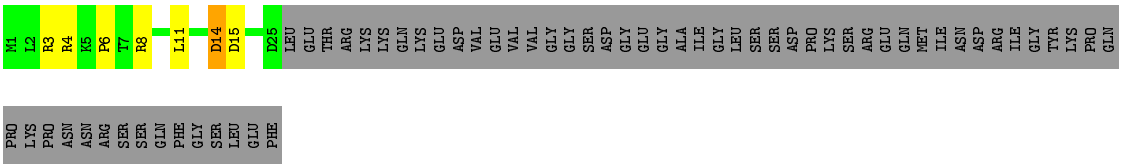


• Molecule 7: ANAPHASE-PROMOTING COMPLEX SUBUNIT CDC26

Chain G:

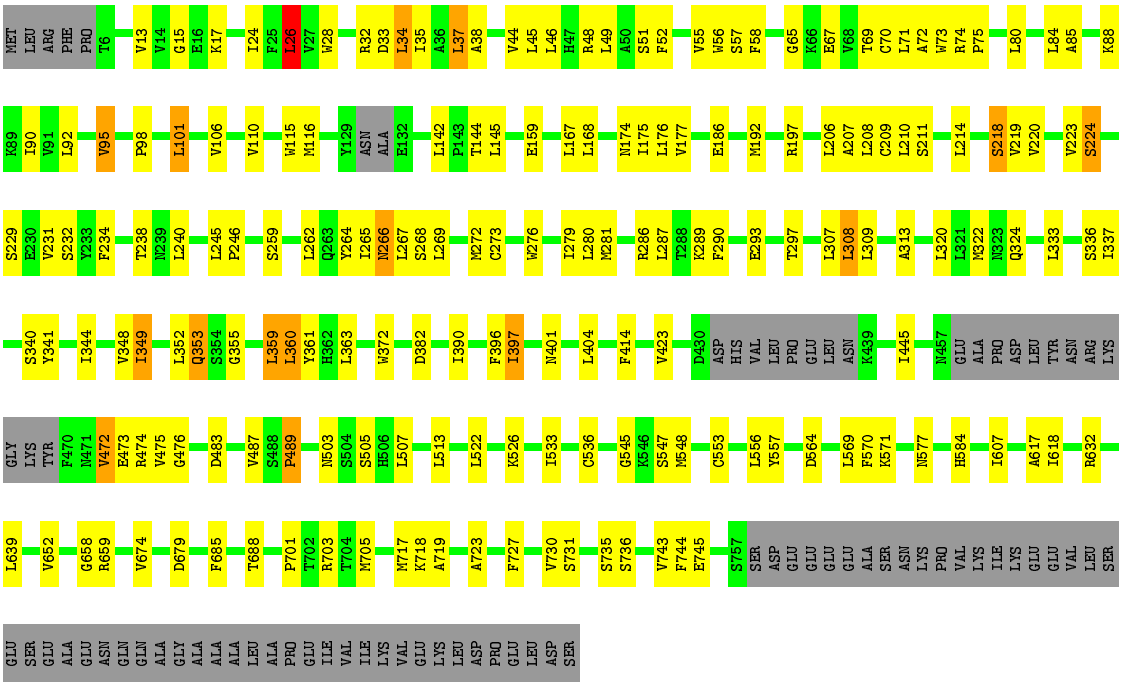
- Molecule 7: ANAPHASE-PROMOTING COMPLEX SUBUNIT CDC26

Chain W:



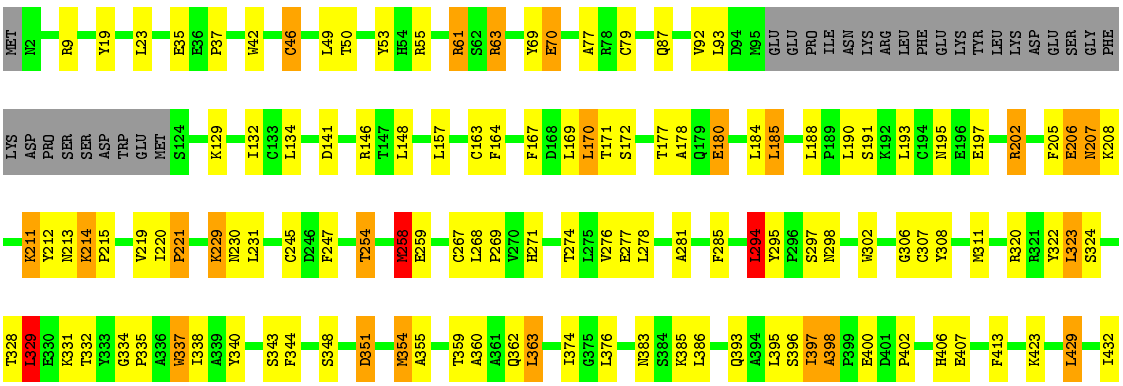
- Molecule 8: ANAPHASE-PROMOTING COMPLEX SUBUNIT 4

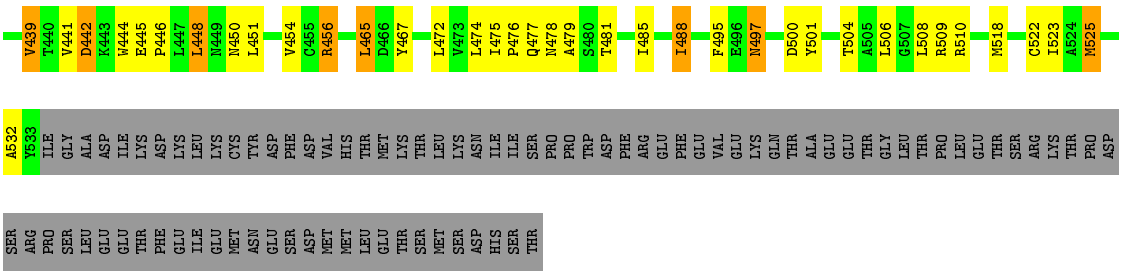
Chain I:



- Molecule 9: ANAPHASE-PROMOTING COMPLEX SUBUNIT 6

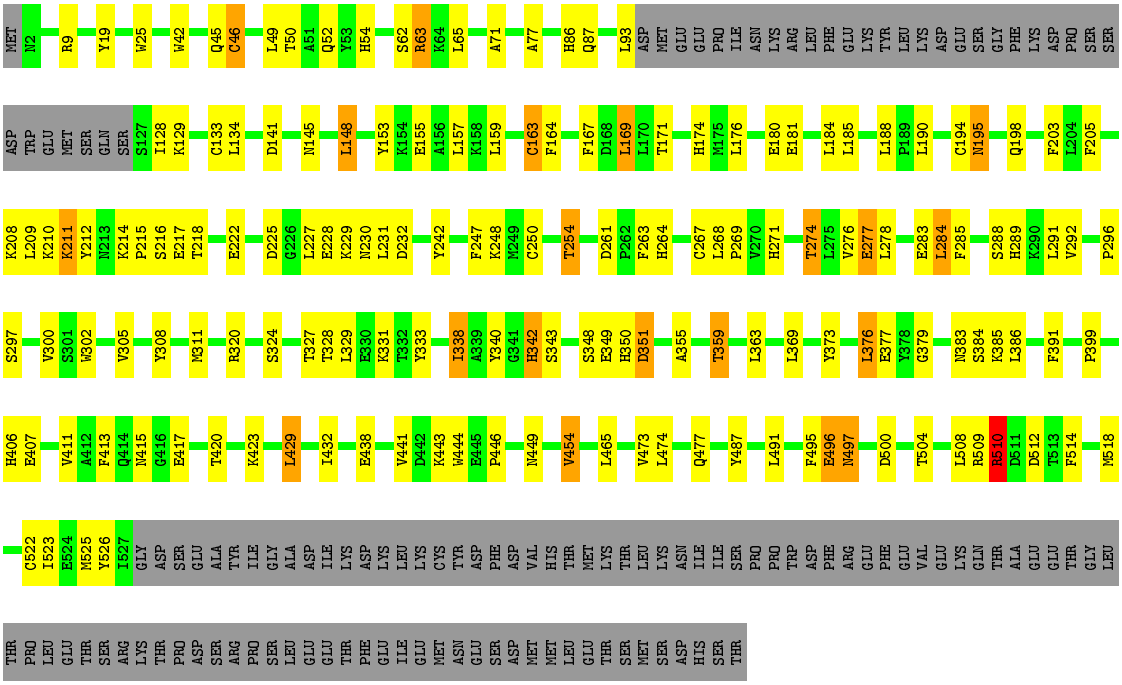
Chain J:





• Molecule 10: ANAPHASE-PROMOTING COMPLEX SUBUNIT 6

Chain K:

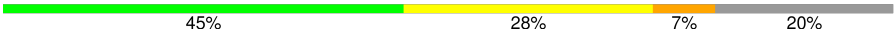


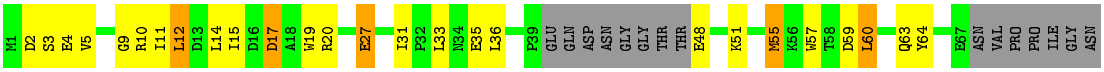
• Molecule 11: ANAPHASE-PROMOTING COMPLEX SUBUNIT 10

Chain L:

• Molecule 12: ANAPHASE-PROMOTING COMPLEX SUBUNIT 13

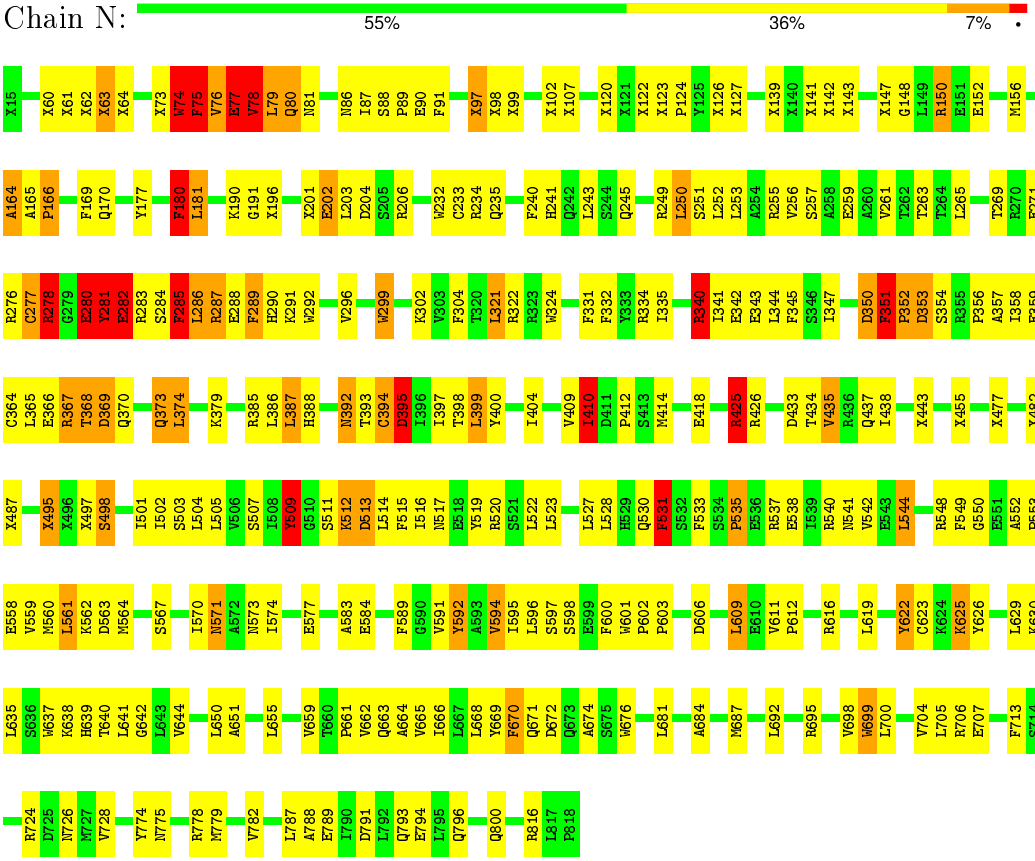
Chain M:





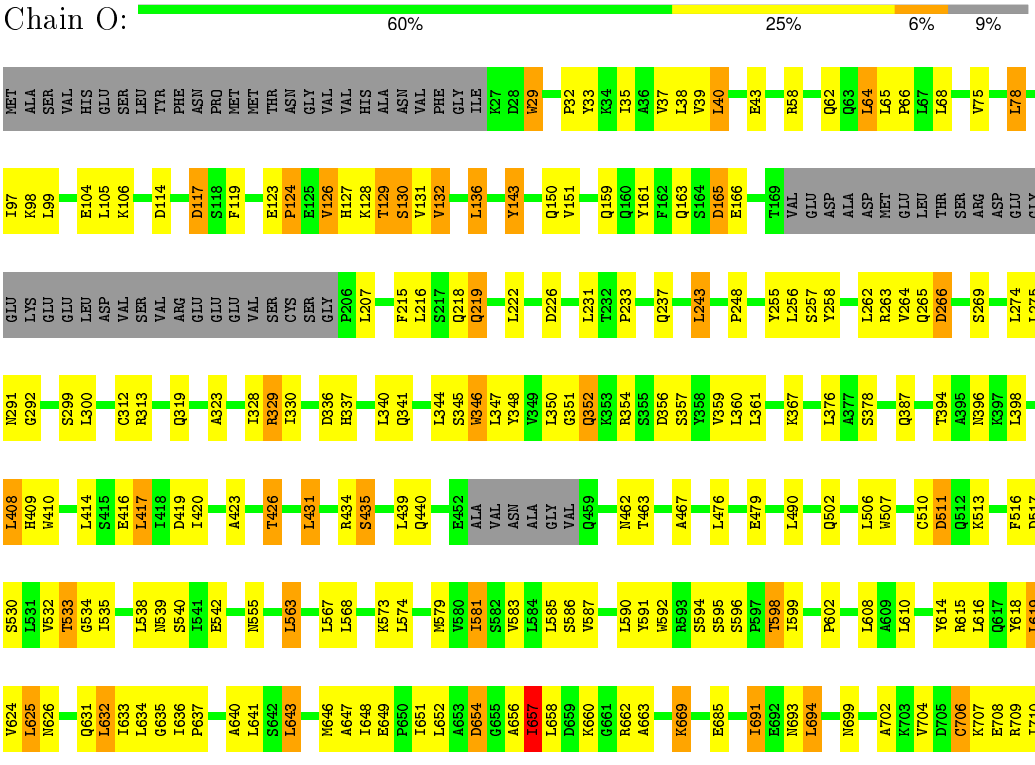
• Molecule 13: ANAPHASE-PROMOTING COMPLEX SUBUNIT 2

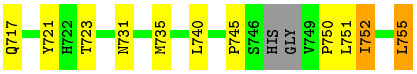
Chain N:



• Molecule 14: ANAPHASE-PROMOTING COMPLEX SUBUNIT 5

Chain O:



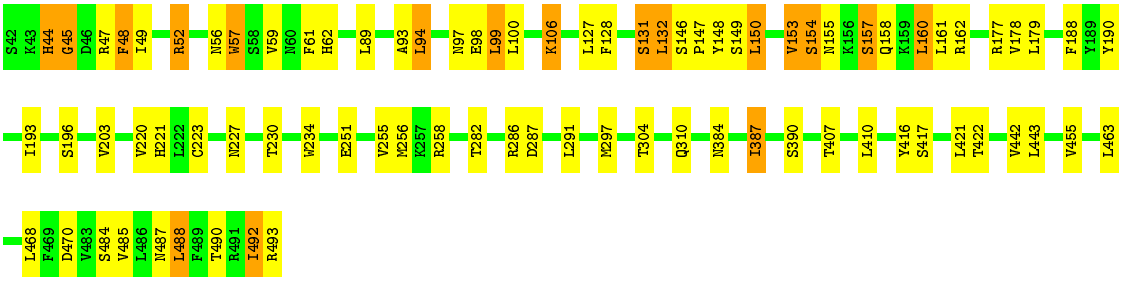


• Molecule 15: FUSION PROTEIN - UBIQUITIN-CONJUGATING ENZYME E2 C, UBIQUITIN-CONJUGATING ENZYME E2 S

Chain Q:

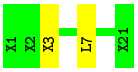
• Molecule 16: THE ANAPHASE-PROMOTING COMPLEX CHAIN R

Chain R:



• Molecule 17: THE ANAPHASE-PROMOTING COMPLEX CHAIN T

Chain T:

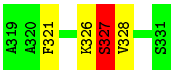


• Molecule 18: THE ANAPHASE-PROMOTING COMPLEX CHAIN U

Chain U:

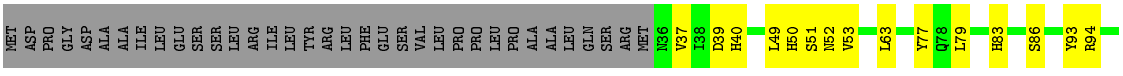
• Molecule 19: THE ANAPHASE-PROMOTING COMPLEX CHAIN V

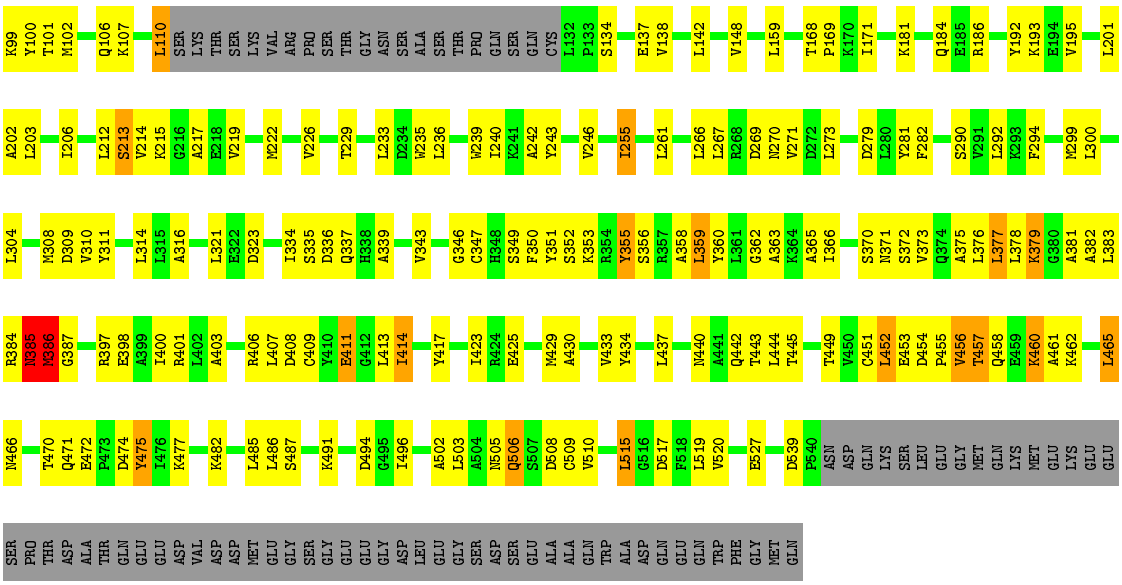
Chain V:



• Molecule 20: ANAPHASE-PROMOTING COMPLEX SUBUNIT 7

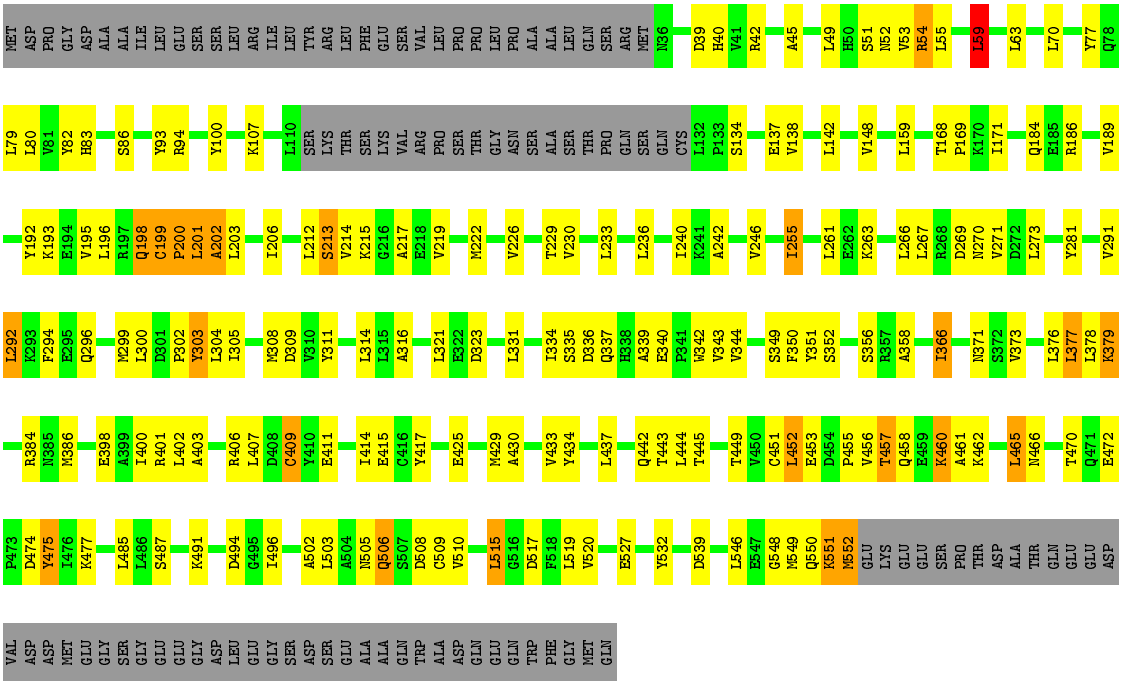
Chain X:





• Molecule 20: ANAPHASE-PROMOTING COMPLEX SUBUNIT 7

Chain Y:



4 Experimental information (i)

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	Depositor
Number of images	Not provided	Depositor
Resolution determination method	Not provided	Depositor
CTF correction method	Not provided	Depositor
Microscope	FEI TECNAI F30	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/Å^2$)	Not provided	Depositor
Minimum defocus (nm)	2000	Depositor
Maximum defocus (nm)	4000	Depositor
Magnification	78000	Depositor
Image detector	FALCON II	Depositor

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: ZN

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	#$\|Z\|>2$	RMSZ	#$\|Z\|>2$
1	A	0.75	0/10949	1.00	18/14903 (0.1%)
10	K	0.89	3/4086 (0.1%)	0.96	5/5532 (0.1%)
11	L	0.71	0/1468	0.96	5/1993 (0.3%)
12	M	0.73	0/502	1.05	1/680 (0.1%)
13	N	0.60	1/4885 (0.0%)	0.96	8/6596 (0.1%)
14	O	0.73	5/5494 (0.1%)	0.97	5/7425 (0.1%)
15	Q	0.59	0/1174	0.69	0/1601
16	R	0.61	0/3052	0.85	2/4139 (0.0%)
17	T	0.84	0/13	0.99	0/16
19	V	0.65	0/99	0.91	0/130
2	B	0.52	0/675	0.85	1/914 (0.1%)
20	X	0.60	4/3830 (0.1%)	0.84	6/5187 (0.1%)
20	Y	0.54	0/3925	0.85	4/5311 (0.1%)
3	C	0.75	1/4403 (0.0%)	0.96	9/5945 (0.2%)
3	P	0.70	2/4137 (0.0%)	0.92	4/5587 (0.1%)
4	D	0.70	0/447	0.98	1/612 (0.2%)
5	E	0.66	0/459	0.86	0/619
6	F	0.70	3/4013 (0.1%)	0.90	7/5428 (0.1%)
6	H	0.70	2/3943 (0.1%)	0.90	4/5329 (0.1%)
7	G	0.63	0/212	1.03	1/281 (0.4%)
7	W	0.65	0/214	1.01	0/284
8	I	0.58	0/5827	0.85	3/7899 (0.0%)
9	J	0.75	3/4146 (0.1%)	0.97	8/5615 (0.1%)
All	All	0.70	24/67953 (0.0%)	0.93	92/92026 (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
1	A	0	9
13	N	0	28
16	R	0	2
17	T	0	1
20	X	0	1
8	I	0	2
9	J	0	1
All	All	0	44

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
10	K	229	LYS	CB-CG	33.03	2.41	1.52
20	X	385	ASN	N-CA	9.74	1.65	1.46
9	J	302	TRP	CB-CG	-9.04	1.33	1.50
20	X	355	TYR	CE1-CZ	8.78	1.50	1.38
9	J	337	TRP	CB-CG	-8.44	1.35	1.50

The worst 5 of 92 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	1540	ARG	NE-CZ-NH1	16.30	128.45	120.30
10	K	229	LYS	CA-CB-CG	-13.30	84.15	113.40
20	X	379	LYS	CD-CE-NZ	10.58	136.04	111.70
13	N	425	ARG	NE-CZ-NH1	9.72	125.16	120.30
20	Y	379	LYS	CD-CE-NZ	9.03	132.47	111.70

There are no chirality outliers.

5 of 44 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	11	MET	Peptide
1	A	124	GLN	Peptide
1	A	14	ALA	Peptide
1	A	83	ILE	Peptide
1	A	86	ASP	Peptide

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	10950	0	10677	405	0
2	B	650	0	598	120	0
3	C	4305	0	4273	122	0
3	P	4042	0	3998	138	0
4	D	437	0	396	14	0
5	E	450	0	435	12	0
6	F	3923	0	3813	90	0
6	H	3853	0	3788	99	0
7	G	211	0	220	2	0
7	W	213	0	220	8	0
8	I	5709	0	5597	123	0
9	J	4047	0	3962	126	0
10	K	3988	0	3917	124	0
11	L	1435	0	1382	61	0
12	M	493	0	469	28	0
13	N	5400	0	4967	426	0
14	O	5396	0	5425	162	0
15	Q	1227	0	1128	27	0
16	R	2990	0	2913	62	0
17	T	109	0	32	3	0
18	U	120	0	27	2	0
19	V	99	0	111	1	0
20	X	3770	0	3829	254	0
20	Y	3865	0	3925	159	0
21	B	3	0	0	0	0
All	All	67685	0	66102	2317	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 17.

The worst 5 of 2317 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)
13:N:435:VAL:CG1	13:N:515:PHE:CE2	1.81	1.63
1:A:1114:ARG:HE	13:N:779:MET:CG	1.13	1.57
2:B:14:TRP:CH2	2:B:41:GLY:N	1.73	1.54
13:N:502:ILE:CG1	13:N:548:ARG:HH12	1.21	1.50
1:A:1116:THR:CG2	13:N:779:MET:CE	1.87	1.49

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 PDB entries followed by that with respect to all EM entries.

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	1354/1441 (94%)	1217 (90%)	106 (8%)	31 (2%)	8 50
2	B	83/84 (99%)	69 (83%)	8 (10%)	6 (7%)	1 23
3	C	520/597 (87%)	495 (95%)	23 (4%)	2 (0%)	39 80
3	P	485/597 (81%)	459 (95%)	25 (5%)	1 (0%)	52 86
4	D	53/121 (44%)	46 (87%)	6 (11%)	1 (2%)	10 54
5	E	54/110 (49%)	54 (100%)	0	0	100 100
6	F	494/824 (60%)	477 (97%)	11 (2%)	6 (1%)	16 62
6	H	479/824 (58%)	462 (96%)	12 (2%)	5 (1%)	19 65
7	G	23/85 (27%)	23 (100%)	0	0	100 100
7	W	23/85 (27%)	23 (100%)	0	0	100 100
8	I	722/808 (89%)	690 (96%)	28 (4%)	4 (1%)	30 74
9	J	500/620 (81%)	467 (93%)	28 (6%)	5 (1%)	19 65
10	K	489/620 (79%)	458 (94%)	26 (5%)	5 (1%)	19 65
11	L	180/185 (97%)	169 (94%)	9 (5%)	2 (1%)	17 64
12	M	55/74 (74%)	46 (84%)	9 (16%)	0	100 100
13	N	572/703 (81%)	501 (88%)	33 (6%)	38 (7%)	1 25
14	O	677/755 (90%)	645 (95%)	24 (4%)	8 (1%)	16 62
15	Q	144/162 (89%)	139 (96%)	5 (4%)	0	100 100
16	R	374/386 (97%)	334 (89%)	35 (9%)	5 (1%)	15 60
17	T	2/21 (10%)	2 (100%)	0	0	100 100
19	V	11/13 (85%)	8 (73%)	1 (9%)	2 (18%)	0 3
20	X	480/599 (80%)	462 (96%)	15 (3%)	3 (1%)	30 74
20	Y	492/599 (82%)	471 (96%)	16 (3%)	5 (1%)	19 65
All	All	8266/10313 (80%)	7717 (93%)	420 (5%)	129 (2%)	17 57

5 of 129 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	274	VAL
1	A	630	PRO
1	A	857	ALA
1	A	860	TYR
1	A	1125	ILE

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 PDB entries followed by that with respect to all EM entries.

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	1150/1206 (95%)	988 (86%)	162 (14%)	4 28
2	B	65/75 (87%)	54 (83%)	11 (17%)	2 20
3	C	452/520 (87%)	399 (88%)	53 (12%)	7 35
3	P	422/520 (81%)	373 (88%)	49 (12%)	7 35
4	D	46/115 (40%)	41 (89%)	5 (11%)	8 38
5	E	47/89 (53%)	37 (79%)	10 (21%)	1 11
6	F	407/729 (56%)	367 (90%)	40 (10%)	10 43
6	H	408/729 (56%)	372 (91%)	36 (9%)	12 48
7	G	22/77 (29%)	20 (91%)	2 (9%)	12 46
7	W	23/77 (30%)	22 (96%)	1 (4%)	35 72
8	I	620/730 (85%)	572 (92%)	48 (8%)	16 55
9	J	424/546 (78%)	368 (87%)	56 (13%)	5 30
10	K	423/549 (77%)	380 (90%)	43 (10%)	9 40
11	L	155/170 (91%)	140 (90%)	15 (10%)	10 43
12	M	55/67 (82%)	44 (80%)	11 (20%)	1 13
13	N	516/526 (98%)	459 (89%)	57 (11%)	8 38
14	O	578/651 (89%)	491 (85%)	87 (15%)	3 25
15	Q	122/127 (96%)	122 (100%)	0	100 100
16	R	326/334 (98%)	295 (90%)	31 (10%)	11 44

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
17	T	1/2 (50%)	1 (100%)	0	100 100
19	V	10/10 (100%)	7 (70%)	3 (30%)	0 3
20	X	407/513 (79%)	378 (93%)	29 (7%)	18 58
20	Y	418/513 (82%)	382 (91%)	36 (9%)	13 50
All	All	7097/8875 (80%)	6312 (89%)	785 (11%)	12 37

5 of 785 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
8	I	489	PRO
10	K	232	ASP
20	X	110	LEU
8	I	736	SER
9	J	329	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 139 such sidechains are listed below:

Mol	Chain	Res	Type
8	I	374	GLN
10	K	18	GLN
20	X	151	GLN
8	I	535	GLN
9	J	316	ASN

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

Of 3 ligands modelled in this entry, 3 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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.