



wwPDB X-ray Structure Validation Summary Report ⓘ

Sep 17, 2023 – 10:50 AM EDT

PDB ID : 4XF2
Title : Tetragonal structure of Arp2/3 complex
Authors : Jurgenson, C.T.; Pollard, T.P.
Deposited on : 2014-12-25
Resolution : 5.00 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

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

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

MolProbity : 4.02b-467
Xtriage (Phenix) : 1.13
EDS : 2.35.1
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) : 2.35.1

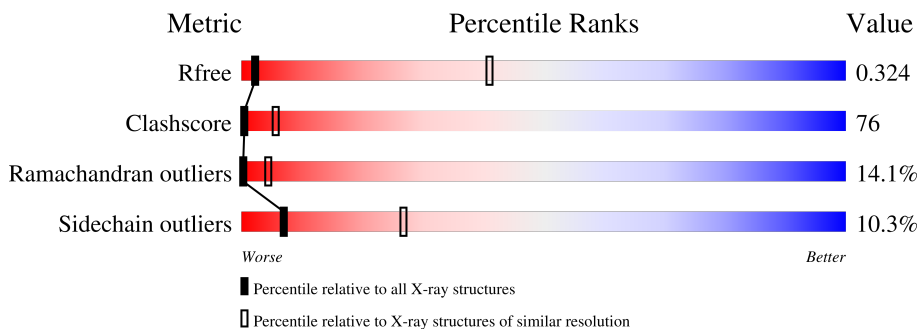
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 5.00 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}	130704	1140 (6.20-3.80)
Clashscore	141614	1000 (6.16-3.82)
Ramachandran outliers	138981	1146 (6.20-3.80)
Sidechain outliers	138945	1122 (6.20-3.80)

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 $\leq 5\%$

Mol	Chain	Length	Quality of chain
1	A	418	24% (green) 58% (yellow) 11% (orange) 6% (red) 1% (grey)
1	T	418	25% (green) 56% (yellow) 11% (orange) 6% (red) 1% (grey)
2	B	394	22% (green) 26% (yellow) 9% (orange) 42% (grey)
2	U	394	20% (green) 28% (yellow) 9% (orange) 42% (grey)
3	C	372	19% (green) 54% (yellow) 20% (orange) 5% (red) 1% (grey)
3	V	372	17% (green) 57% (yellow) 18% (orange) 5% (red) 1% (grey)
4	D	300	28% (green) 53% (yellow) 13% (orange) 5% (red) 1% (grey)

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
4	W	300	 29% 51% 14% 5%
5	E	178	 26% 62% 10%
5	X	178	 27% 57% 12%
6	F	168	 29% 58% 12%
6	Y	168	 24% 59% 15%
7	G	151	 27% 52% 13% 8%
7	Z	151	 25% 53% 14% 8%

2 Entry composition [i](#)

There are 7 unique types of molecules in this entry. The entry contains 27556 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 Actin-related protein 3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	392	3146	2022	522	587	15	0	0	0
1	T	392	3146	2022	522	587	15	0	0	0

- Molecule 2 is a protein called Actin-related protein 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	228	1734	1105	303	322	4	0	0	0
2	U	228	1738	1108	304	322	4	0	0	0

- Molecule 3 is a protein called Actin-related protein 2/3 complex subunit 1B.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	C	354	2758	1747	487	505	19	0	0	0
3	V	354	2758	1747	487	505	19	0	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	58	VAL	ILE	conflict	UNP Q58CQ2
V	58	VAL	ILE	conflict	UNP Q58CQ2

- Molecule 4 is a protein called Actin-related protein 2/3 complex subunit 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
4	D	284	2292	1456	397	431	8	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
4	W	284	2292	1456	397	431	8	0	0	0

- Molecule 5 is a protein called Actin-related protein 2/3 complex subunit 3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
5	E	174	1415	908	236	262	9	0	0	0
5	X	174	1415	908	236	262	9	0	0	0

- Molecule 6 is a protein called Actin-related protein 2/3 complex subunit 4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
6	F	167	1371	875	239	248	9	0	0	0
6	Y	167	1371	875	239	248	9	0	0	0

- Molecule 7 is a protein called Actin-related protein 2/3 complex subunit 5.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
7	G	139	1060	661	185	211	3	0	0	0
7	Z	139	1060	661	185	211	3	0	0	0

There are 4 discrepancies between the modelled and reference sequences:

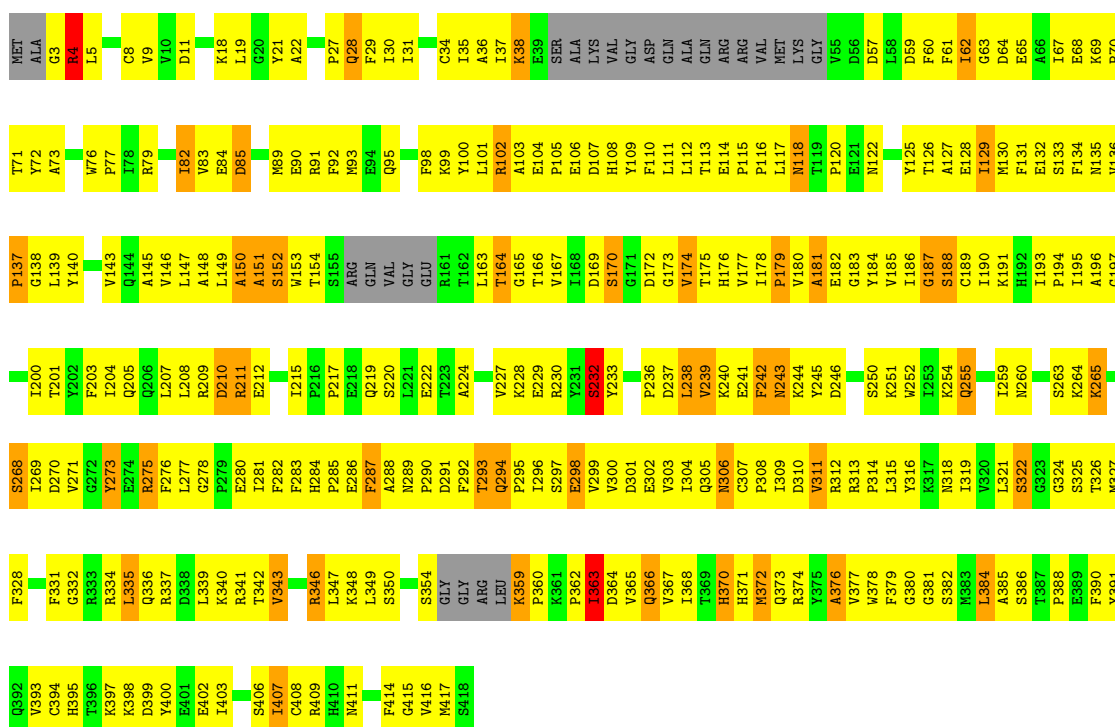
Chain	Residue	Modelled	Actual	Comment	Reference
G	17	ASP	GLY	conflict	UNP Q3SYX9
G	28	ASP	GLU	conflict	UNP Q3SYX9
Z	17	ASP	GLY	conflict	UNP Q3SYX9
Z	28	ASP	GLU	conflict	UNP Q3SYX9

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. 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.

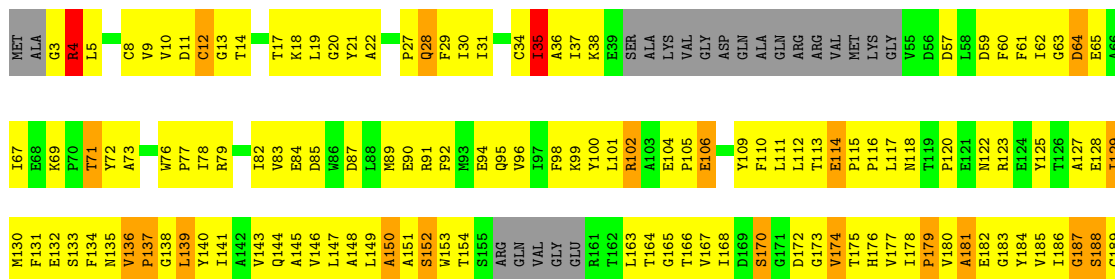
- Molecule 1: Actin-related protein 3

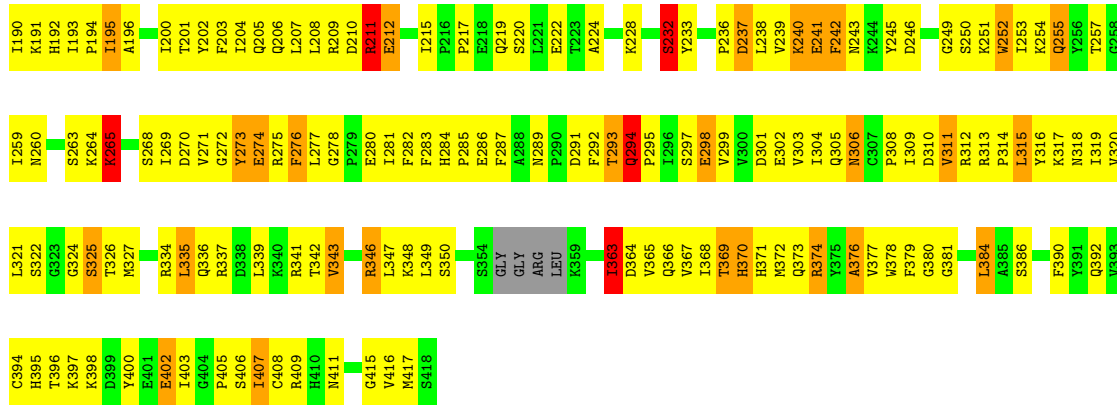
Chain A: 



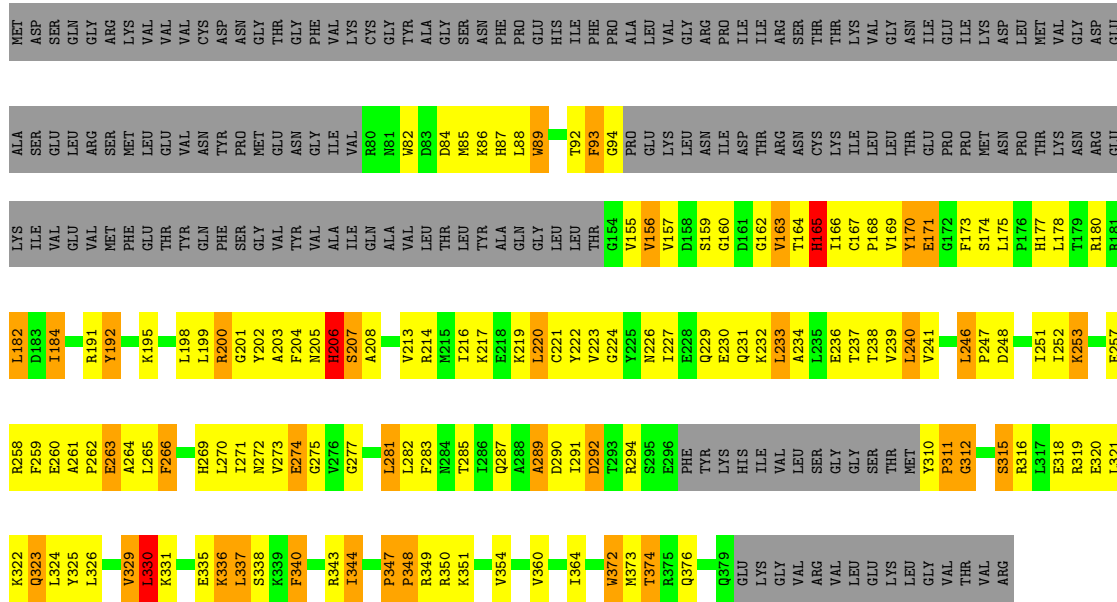
- Molecule 1: Actin-related protein 3

Chain T: 

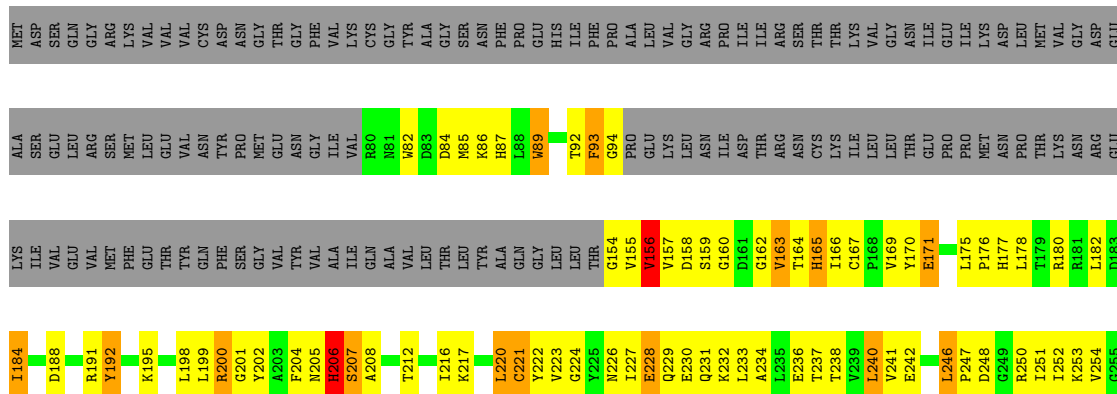


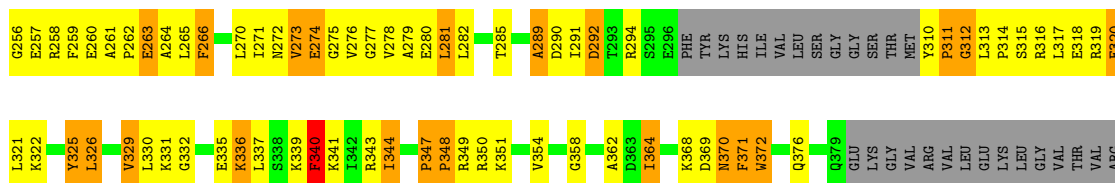


● Molecule 2: Actin-related protein 2

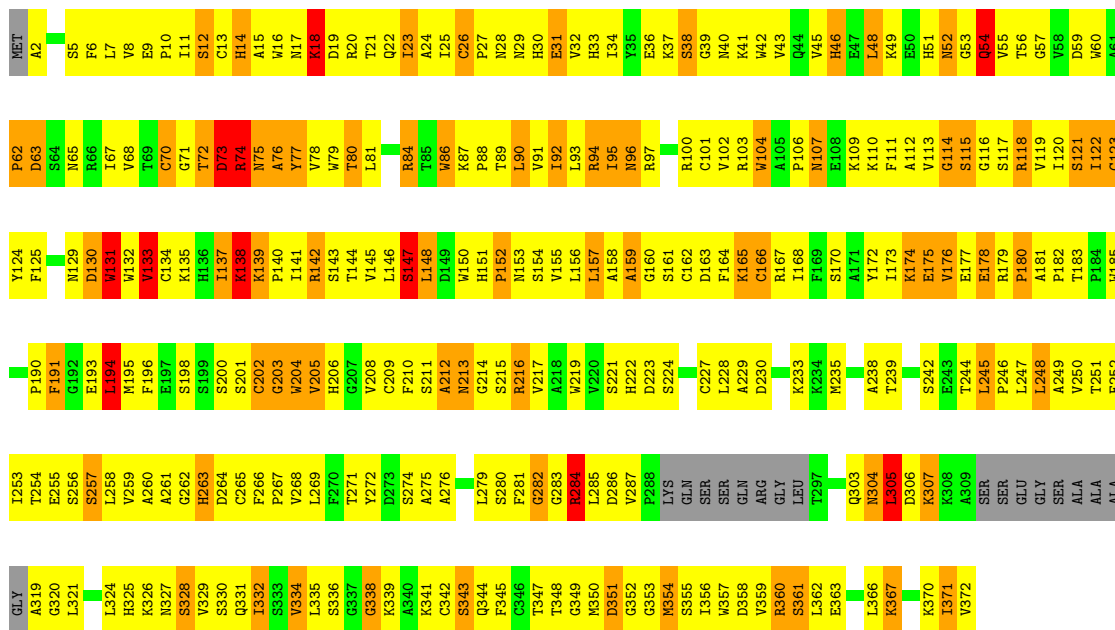
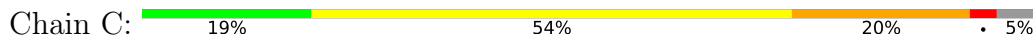


● Molecule 2: Actin-related protein 2

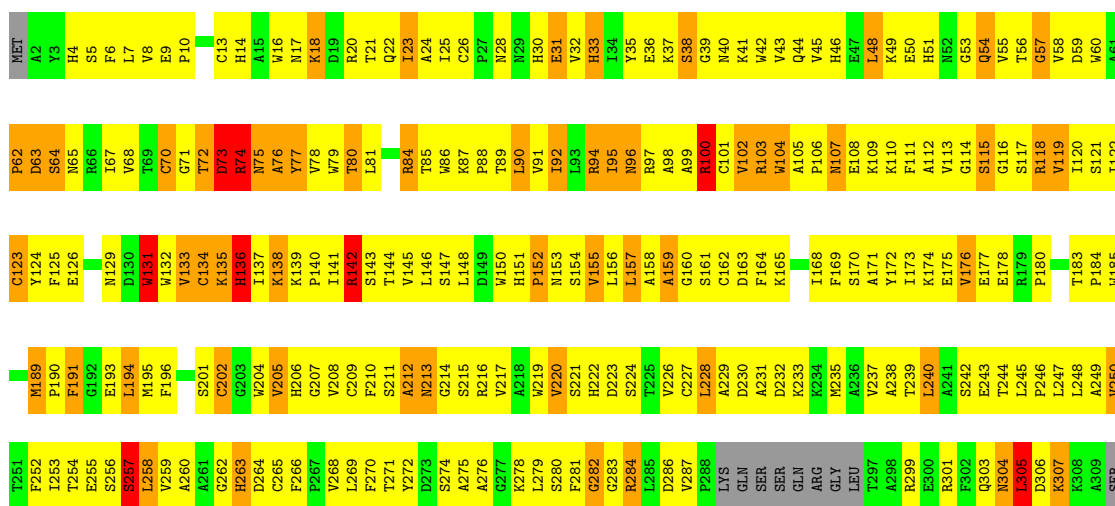
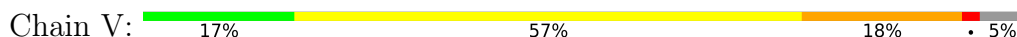




• Molecule 3: Actin-related protein 2/3 complex subunit 1B



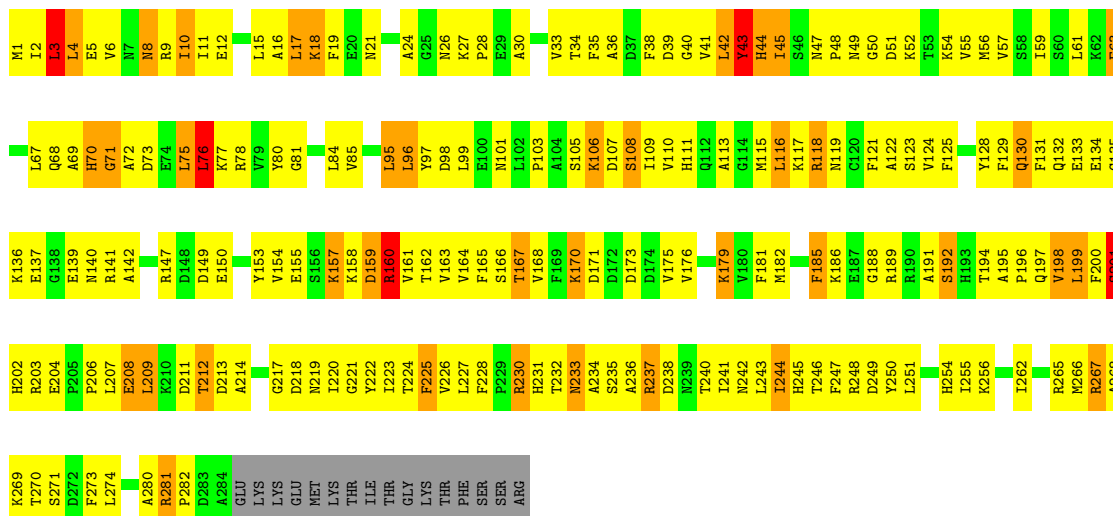
• Molecule 3: Actin-related protein 2/3 complex subunit 1B





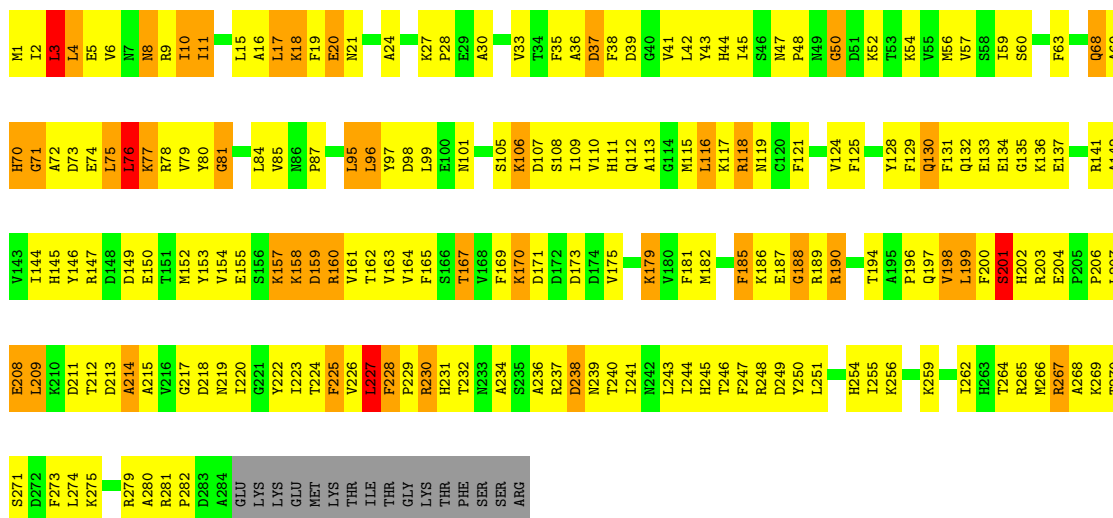
- Molecule 4: Actin-related protein 2/3 complex subunit 2

Chain D: 28% 53% 13% 5%



- Molecule 4: Actin-related protein 2/3 complex subunit 2

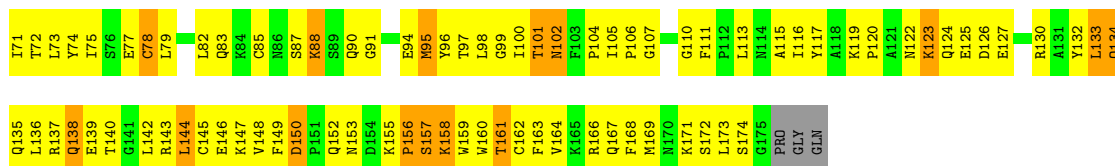
Chain W: 29% 51% 14% 5%



- Molecule 5: Actin-related protein 2/3 complex subunit 3

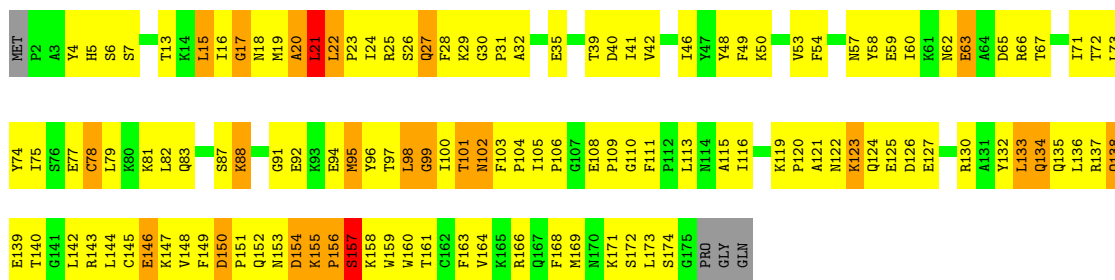
Chain E: 26% 62% 10% 5%





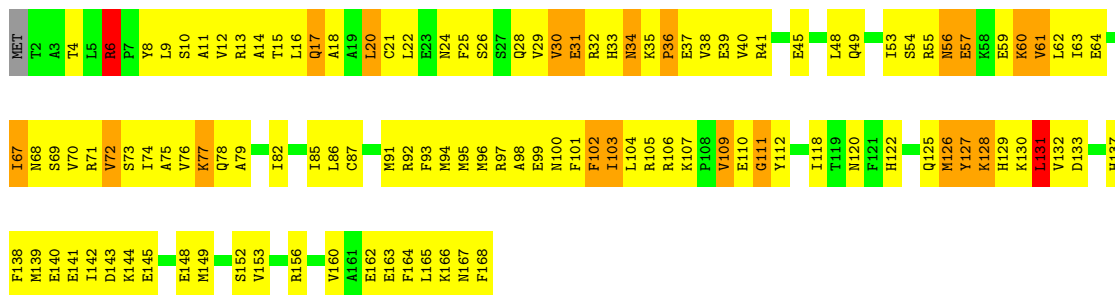
- Molecule 5: Actin-related protein 2/3 complex subunit 3

Chain X: 27% 57% 12% ..



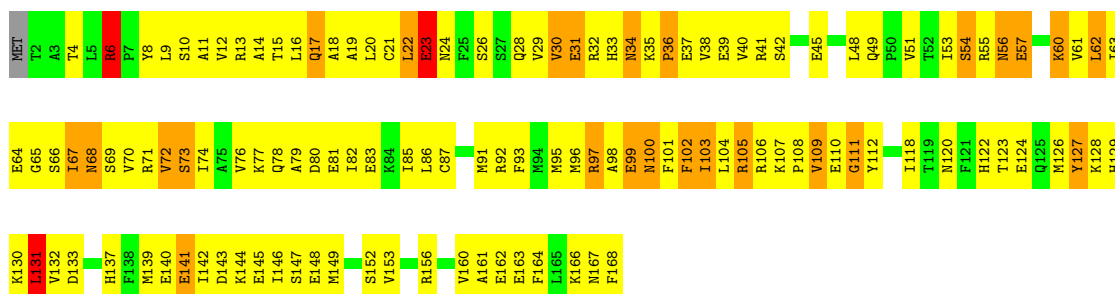
- Molecule 6: Actin-related protein 2/3 complex subunit 4

Chain F: 29% 58% 12% ..



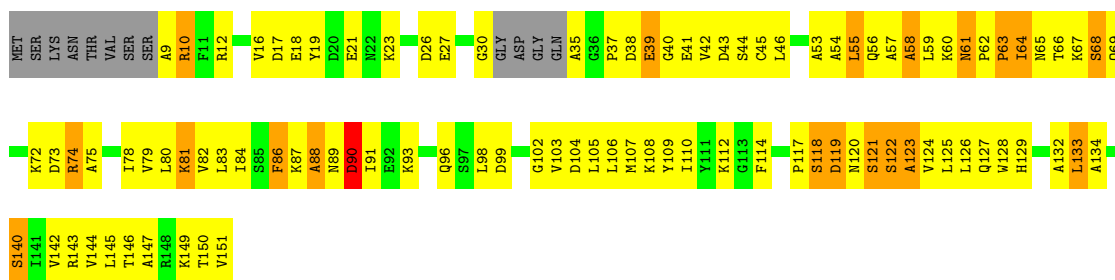
- Molecule 6: Actin-related protein 2/3 complex subunit 4

Chain Y: 24% 59% 15% ..



- Molecule 7: Actin-related protein 2/3 complex subunit 5

Chain G: 27% 52% 13% 8%



- Molecule 7: Actin-related protein 2/3 complex subunit 5

Chain Z: 25% 53% 14% 8%



4 Data and refinement statistics

Property	Value	Source
Space group	P 41	Depositor
Cell constants a, b, c, α , β , γ	149.93Å 149.93Å 265.91Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	47.54 – 5.00 47.54 – 4.30	Depositor EDS
% Data completeness (in resolution range)	96.5 (47.54-5.00) 81.1 (47.54-4.30)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.09	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.68 (at 4.29Å)	Xtrriage
Refinement program	PHENIX 1.9_1692	Depositor
R, R_{free}	0.275 , 0.303 0.311 , 0.324	Depositor DCC
R_{free} test set	1622 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å ²)	185.4	Xtrriage
Anisotropy	0.215	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.23 , 282.2	EDS
L-test for twinning ²	$\langle L \rangle = 0.47$, $\langle L^2 \rangle = 0.30$	Xtrriage
Estimated twinning fraction	0.368 for h,-k,-l	Xtrriage
F_o, F_c correlation	0.87	EDS
Total number of atoms	27556	wwPDB-VP
Average B, all atoms (Å ²)	256.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.55% 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, 0.333 respectively for untwinned datasets, and 0.375, 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	0.42	0/3226	0.83	4/4376 (0.1%)
1	T	0.33	0/3226	0.64	2/4376 (0.0%)
2	B	0.40	0/1764	0.84	5/2389 (0.2%)
2	U	0.33	0/1768	0.67	4/2393 (0.2%)
3	C	0.35	0/2827	0.74	2/3832 (0.1%)
3	V	0.32	0/2827	0.62	0/3832
4	D	0.37	0/2341	0.72	2/3161 (0.1%)
4	W	0.33	0/2341	0.63	2/3161 (0.1%)
5	E	0.37	0/1449	0.71	0/1954
5	X	0.36	0/1449	0.65	0/1954
6	F	0.40	0/1393	0.80	1/1868 (0.1%)
6	Y	0.36	0/1393	0.73	1/1868 (0.1%)
7	G	0.33	0/1072	0.63	0/1442
7	Z	0.38	0/1072	0.67	0/1442
All	All	0.36	0/28148	0.71	23/38048 (0.1%)

There are no bond length outliers.

The worst 5 of 23 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	359	LYS	C-N-CD	-9.01	100.78	120.60
4	D	267	ARG	NE-CZ-NH1	-6.56	117.02	120.30
4	D	96	LEU	CA-CB-CG	6.55	130.37	115.30
4	W	96	LEU	CA-CB-CG	6.54	130.35	115.30
6	Y	131	LEU	CA-CB-CG	6.50	130.24	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	3146	0	3089	442	0
1	T	3146	0	3087	561	0
2	B	1734	0	1641	156	0
2	U	1738	0	1653	224	0
3	C	2758	0	2713	539	0
3	V	2758	0	2711	609	0
4	D	2292	0	2257	444	2
4	W	2292	0	2257	391	1
5	E	1415	0	1416	181	1
5	X	1415	0	1416	230	2
6	F	1371	0	1410	202	0
6	Y	1371	0	1410	257	0
7	G	1060	0	1065	114	0
7	Z	1060	0	1065	119	0
All	All	27556	0	27190	4138	3

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 76.

The worst 5 of 4138 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)
1:T:193:ILE:HG23	1:T:292:PHE:CE1	1.28	1.67
4:D:165:PHE:CZ	4:D:247:PHE:CE2	1.77	1.67
1:T:194:PRO:HD2	1:T:292:PHE:CE1	1.29	1.64
3:V:99:ALA:HA	3:V:115:SER:CB	1.19	1.59
2:B:202:TYR:CE2	2:B:252:ILE:HB	1.11	1.57

All (3) 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)
4:D:200:PHE:O	5:X:88:LYS:NZ[3_555]	1.85	0.35
5:E:88:LYS:NZ	4:W:200:PHE:O[3_545]	1.90	0.30

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
4:D:199:LEU:CD2	5:X:146:GLU:OE1[3_555]	2.10	0.10

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	384/418 (92%)	270 (70%)	66 (17%)	48 (12%)	0	5
1	T	384/418 (92%)	264 (69%)	69 (18%)	51 (13%)	0	4
2	B	222/394 (56%)	139 (63%)	48 (22%)	35 (16%)	0	3
2	U	222/394 (56%)	132 (60%)	55 (25%)	35 (16%)	0	3
3	C	348/372 (94%)	207 (60%)	82 (24%)	59 (17%)	0	3
3	V	348/372 (94%)	208 (60%)	80 (23%)	60 (17%)	0	3
4	D	282/300 (94%)	189 (67%)	65 (23%)	28 (10%)	0	10
4	W	282/300 (94%)	186 (66%)	63 (22%)	33 (12%)	0	6
5	E	172/178 (97%)	124 (72%)	30 (17%)	18 (10%)	0	8
5	X	172/178 (97%)	116 (67%)	34 (20%)	22 (13%)	0	5
6	F	165/168 (98%)	120 (73%)	23 (14%)	22 (13%)	0	4
6	Y	165/168 (98%)	113 (68%)	29 (18%)	23 (14%)	0	4
7	G	135/151 (89%)	91 (67%)	21 (16%)	23 (17%)	0	3
7	Z	135/151 (89%)	84 (62%)	27 (20%)	24 (18%)	0	3
All	All	3416/3962 (86%)	2243 (66%)	692 (20%)	481 (14%)	0	4

5 of 481 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	4	ARG
1	A	28	GLN
1	A	150	ALA

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	181	ALA
1	A	232	SER

5.3.2 Protein sidechains [i](#)

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

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	345/363 (95%)	321 (93%)	24 (7%)	15	41
1	T	345/363 (95%)	319 (92%)	26 (8%)	13	39
2	B	168/345 (49%)	144 (86%)	24 (14%)	3	17
2	U	169/345 (49%)	147 (87%)	22 (13%)	4	20
3	C	301/313 (96%)	257 (85%)	44 (15%)	3	17
3	V	301/313 (96%)	264 (88%)	37 (12%)	4	21
4	D	249/264 (94%)	223 (90%)	26 (10%)	7	26
4	W	249/264 (94%)	225 (90%)	24 (10%)	8	29
5	E	156/159 (98%)	142 (91%)	14 (9%)	9	32
5	X	156/159 (98%)	136 (87%)	20 (13%)	4	20
6	F	154/155 (99%)	145 (94%)	9 (6%)	20	46
6	Y	154/155 (99%)	138 (90%)	16 (10%)	7	26
7	G	114/124 (92%)	103 (90%)	11 (10%)	8	29
7	Z	114/124 (92%)	105 (92%)	9 (8%)	12	38
All	All	2975/3446 (86%)	2669 (90%)	306 (10%)	7	27

5 of 306 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	V	284	ARG
6	Y	60	LYS
4	W	8	ASN
4	W	281	ARG
7	Z	68	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 41 such sidechains are listed below:

Mol	Chain	Res	Type
2	U	226	ASN
4	W	132	GLN
3	V	107	ASN
3	V	213	ASN
5	X	18	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 monosaccharides 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

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

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

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

6.3 Carbohydrates

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

6.4 Ligands

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

6.5 Other polymers

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