wwPDB EM Map/Model Validation Report (i)

Apr 10, 2016 – 01:50 PM BST

PDB ID	:	3J31
EMDB ID:	:	EMD-5584
Title	:	Life in the extremes: atomic structure of Sulfolobus Turreted Icosahedral Virus
Authors	:	Veesler, D.; Ng, T.S.; Sendamarai, A.K.; Eilers, B.J.; Lawrence, C.M.; Lok, S.M.; Young, M.J.; Johnson, J.E.; Fu, C.-Y.
Deposited on	:	2013-02-18
Resolution	:	4.50 Å(reported)
Based on PDB ID	:	2BBD, 4IL7

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

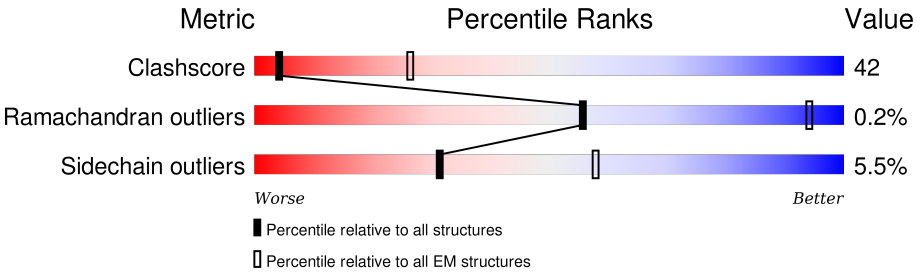
MolProbity	:	4.02b-467
Mogul	:	unknown
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk27241

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 4.50 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	Q	223	54% 36% 9%
2	R	15	80% 20%
3	A	345	39% 56%
3	B	345	38% 59%
3	C	345	39% 58%
3	D	345	45% 53%
3	E	345	41% 55%
3	F	345	39% 59%
3	G	345	41% 57%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	H	345	41% 55%
3	I	345	39% 58%
3	J	345	42% 54%
3	K	345	42% 52% 6%
3	L	345	36% 61%
3	M	345	42% 55%
3	N	345	43% 54%
3	O	345	38% 60%
4	P	381	69% 24% 5%

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 44549 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 A223 penton base.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	Q	220	Total	C	N	O	S	0	0
			1594	1031	261	301	1		

- Molecule 2 is a protein called A55 membrane protein.

Mol	Chain	Residues	Atoms				AltConf	Trace
2	R	15	Total	C	N	O	0	0
			75	45	15	15		

- Molecule 3 is a protein called Coat protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	A	344	Total	C	N	O	S	0	0
			2666	1712	439	511	4		
3	B	344	Total	C	N	O	S	0	0
			2666	1712	439	511	4		
3	C	344	Total	C	N	O	S	0	0
			2666	1712	439	511	4		
3	D	344	Total	C	N	O	S	0	0
			2666	1712	439	511	4		
3	E	344	Total	C	N	O	S	0	0
			2666	1712	439	511	4		
3	F	344	Total	C	N	O	S	0	0
			2666	1712	439	511	4		
3	G	344	Total	C	N	O	S	0	0
			2666	1712	439	511	4		
3	H	344	Total	C	N	O	S	0	0
			2666	1712	439	511	4		
3	I	344	Total	C	N	O	S	0	0
			2666	1712	439	511	4		
3	J	344	Total	C	N	O	S	0	0
			2666	1712	439	511	4		
3	K	344	Total	C	N	O	S	0	0
			2666	1712	439	511	4		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					AltConf	Trace
3	L	344	Total	C	N	O	S	0	0
			2666	1712	439	511	4		
3	M	344	Total	C	N	O	S	0	0
			2666	1712	439	511	4		
3	N	344	Total	C	N	O	S	0	0
			2666	1712	439	511	4		
3	O	344	Total	C	N	O	S	0	0
			2666	1712	439	511	4		

- Molecule 4 is a protein called C381 turret protein.

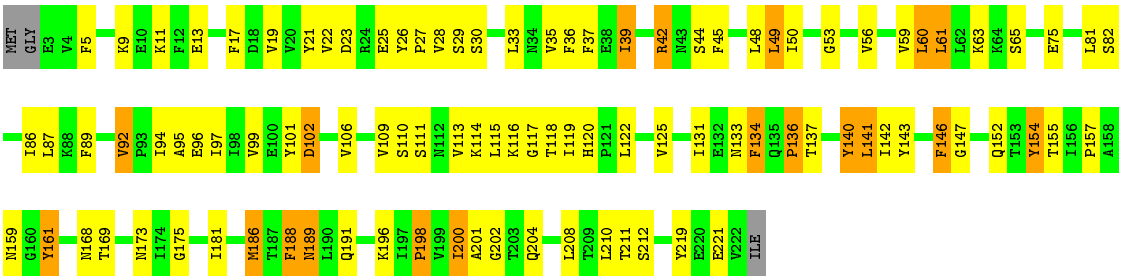
Mol	Chain	Residues	Atoms					AltConf	Trace
4	P	373	Total	C	N	O	S	0	0
			2890	1855	462	570	3		

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.

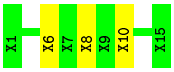
- Molecule 1: A223 penton base

Chain Q:



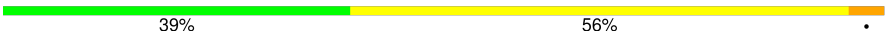
- Molecule 2: A55 membrane protein

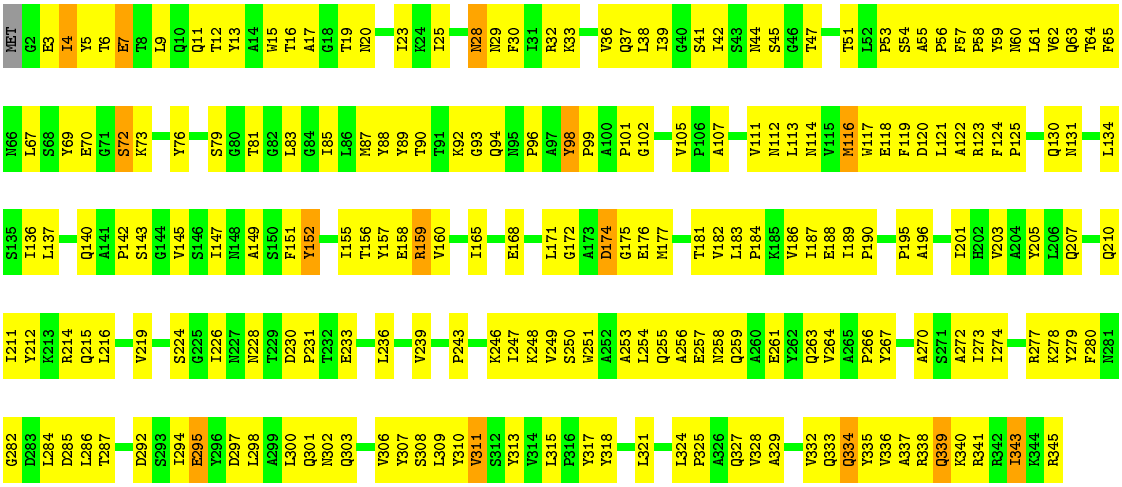
Chain R:



- Molecule 3: Coat protein

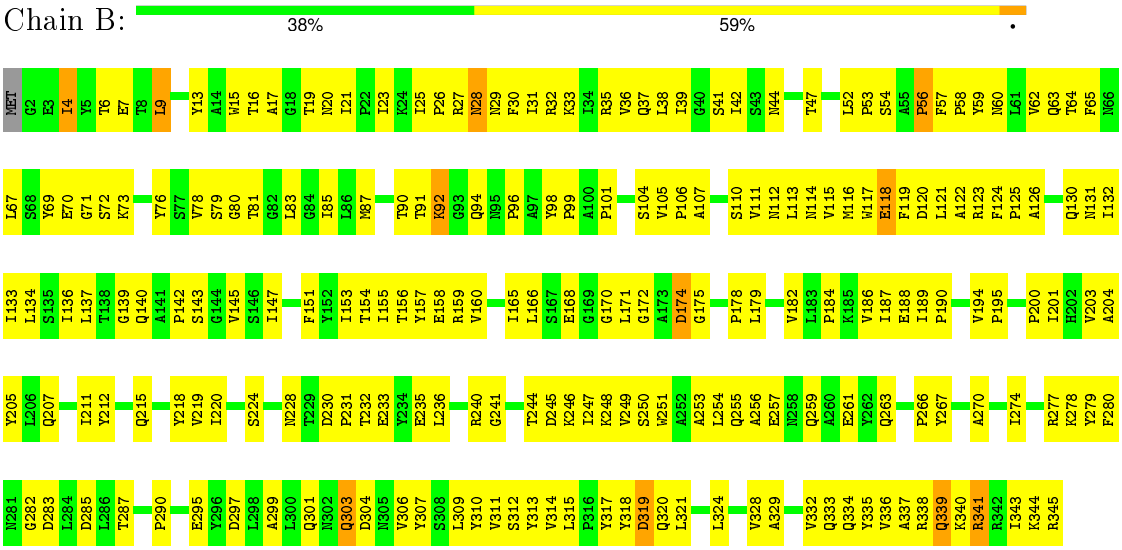
Chain A:





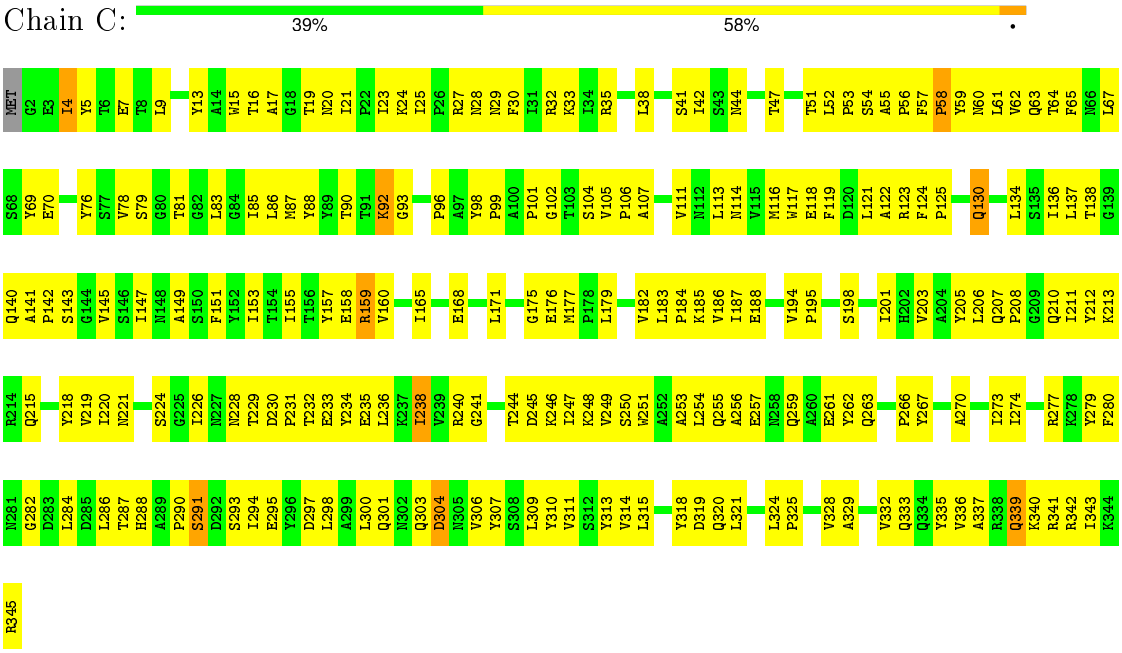
- Molecule 3: Coat protein

Chain B:



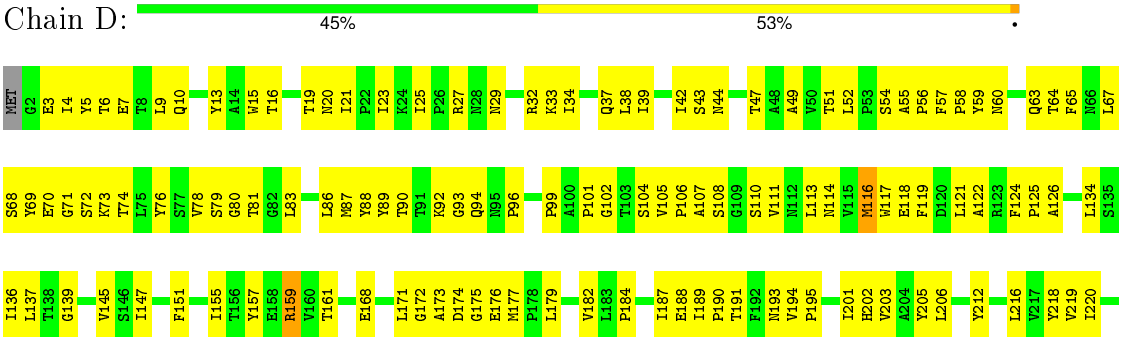
- Molecule 3: Coat protein

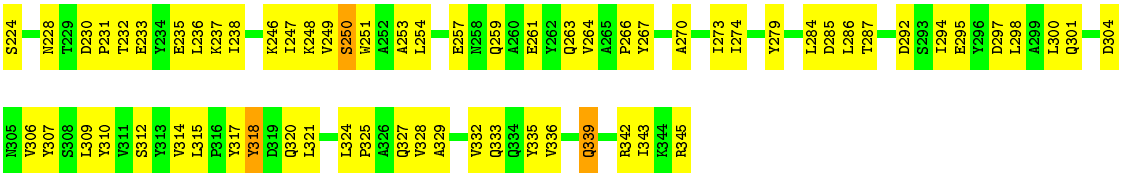
Chain C:



- Molecule 3: Coat protein

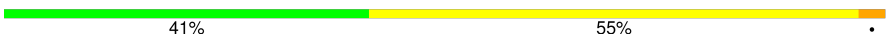
Chain D:

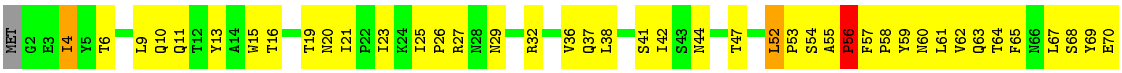


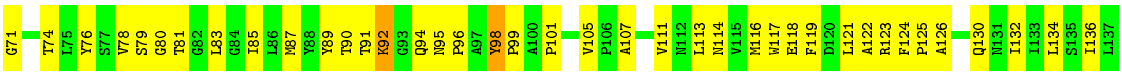


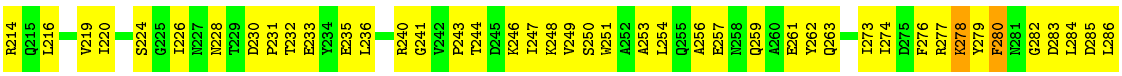
• Molecule 3: Coat protein

Chain E:



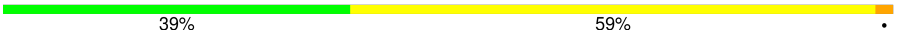


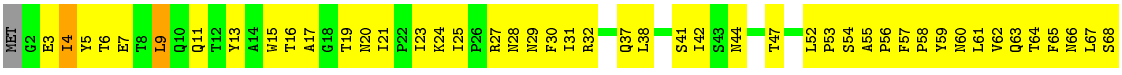


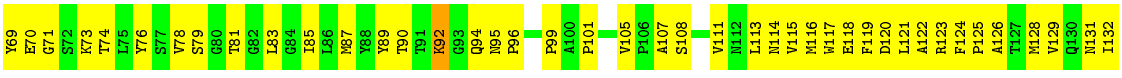


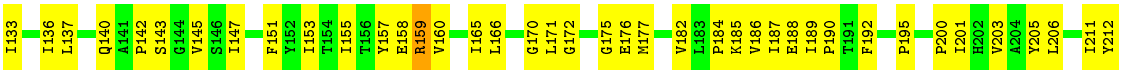
• Molecule 3: Coat protein

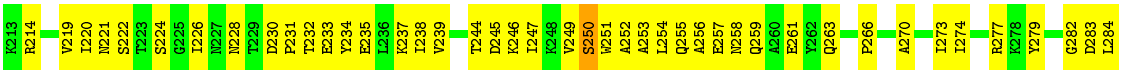
Chain F:





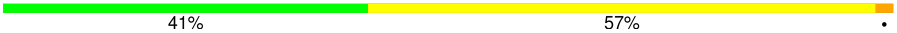


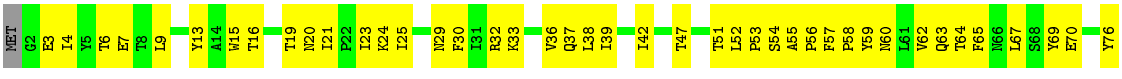


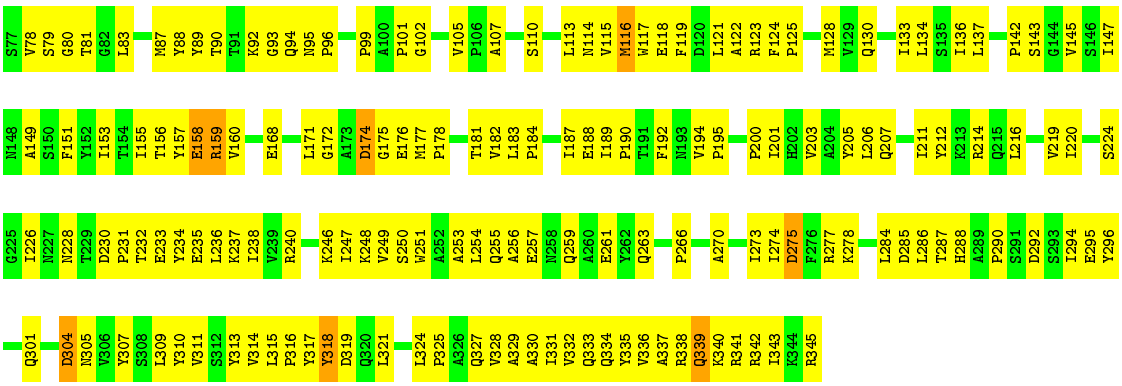


• Molecule 3: Coat protein

Chain G:

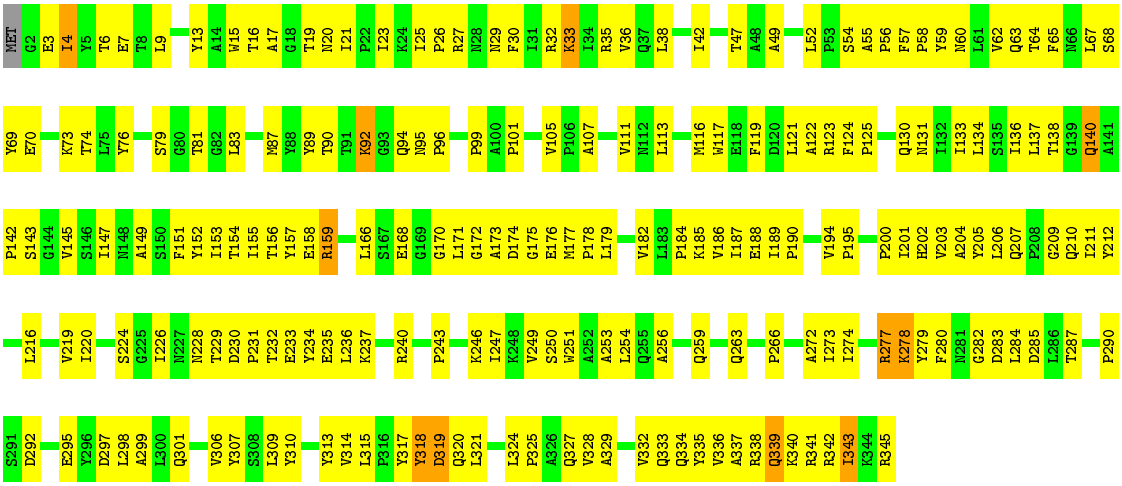






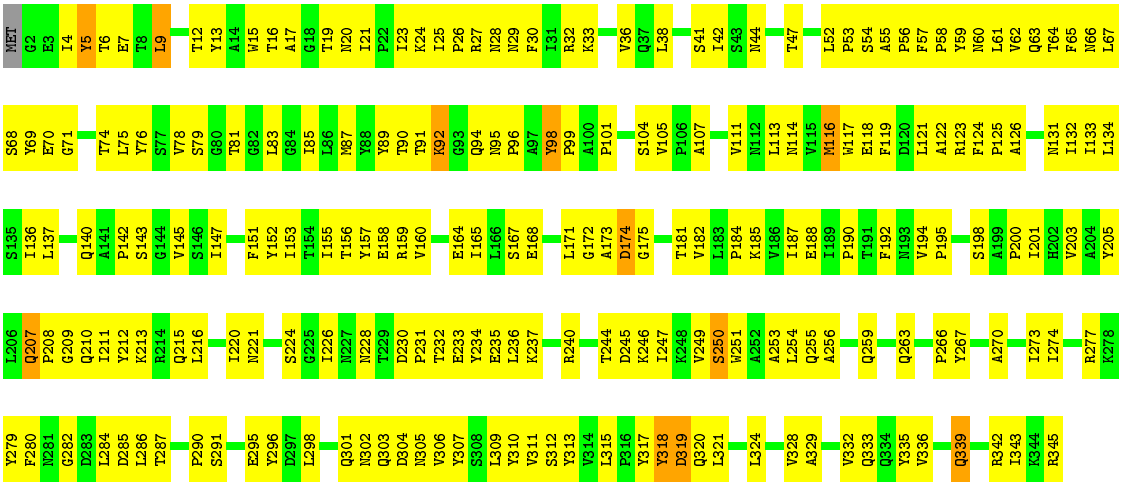
• Molecule 3: Coat protein

Chain H:



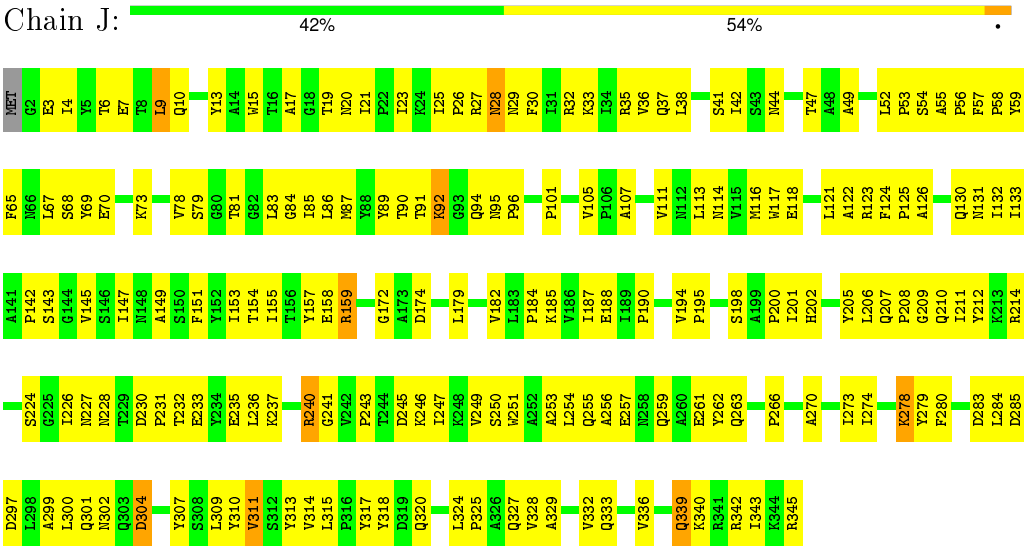
• Molecule 3: Coat protein

Chain I:



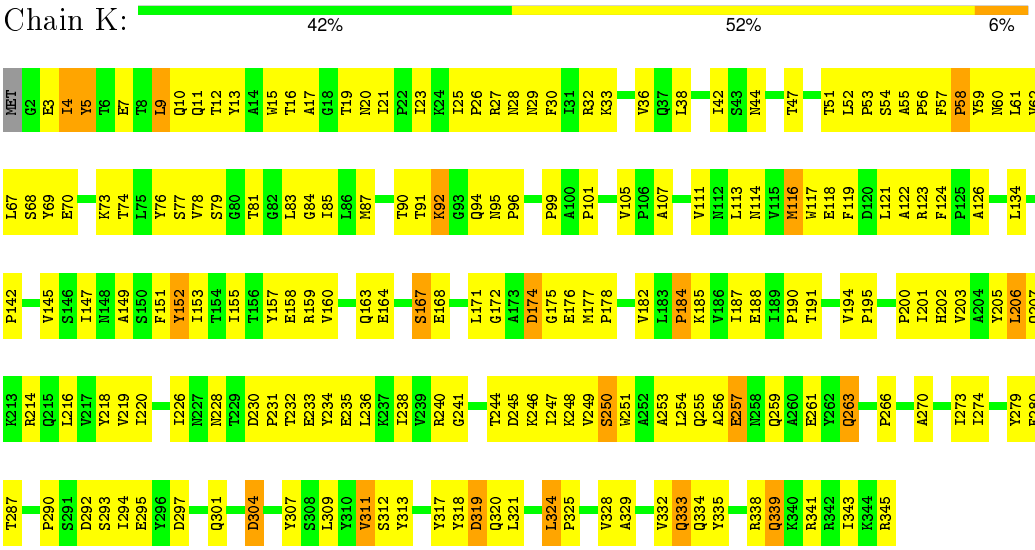
• Molecule 3: Coat protein

Chain J:



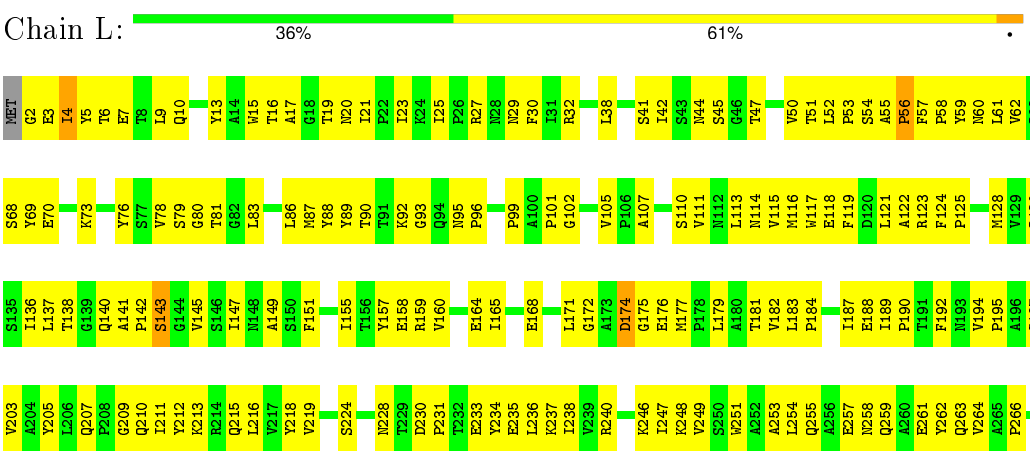
- Molecule 3: Coat protein

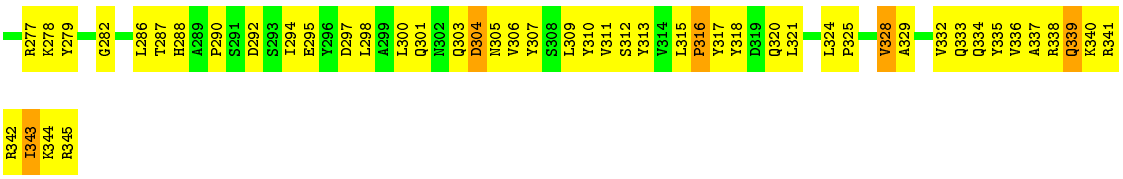
Chain K:



- Molecule 3: Coat protein

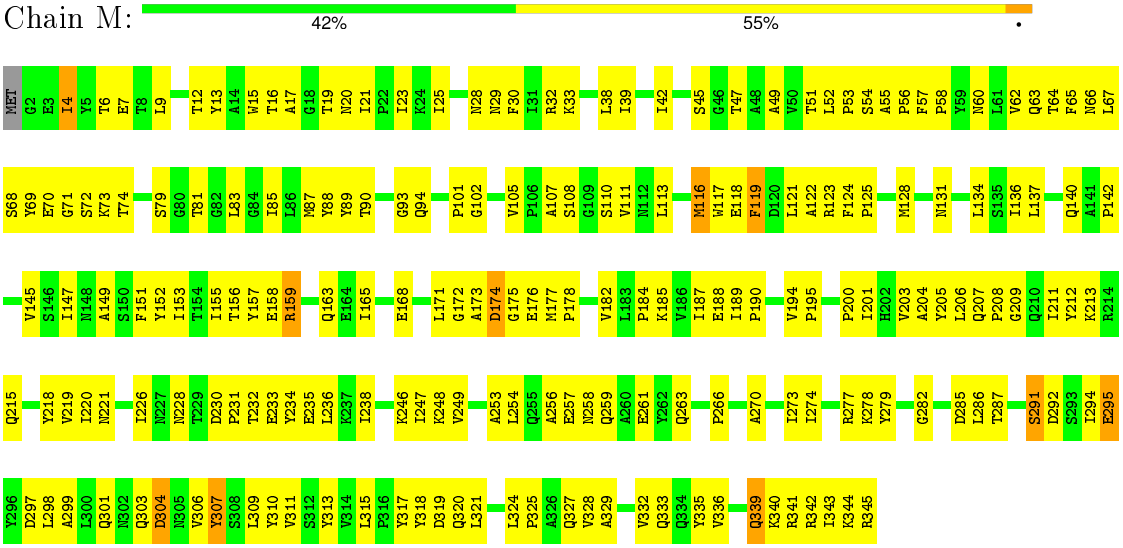
Chain L:





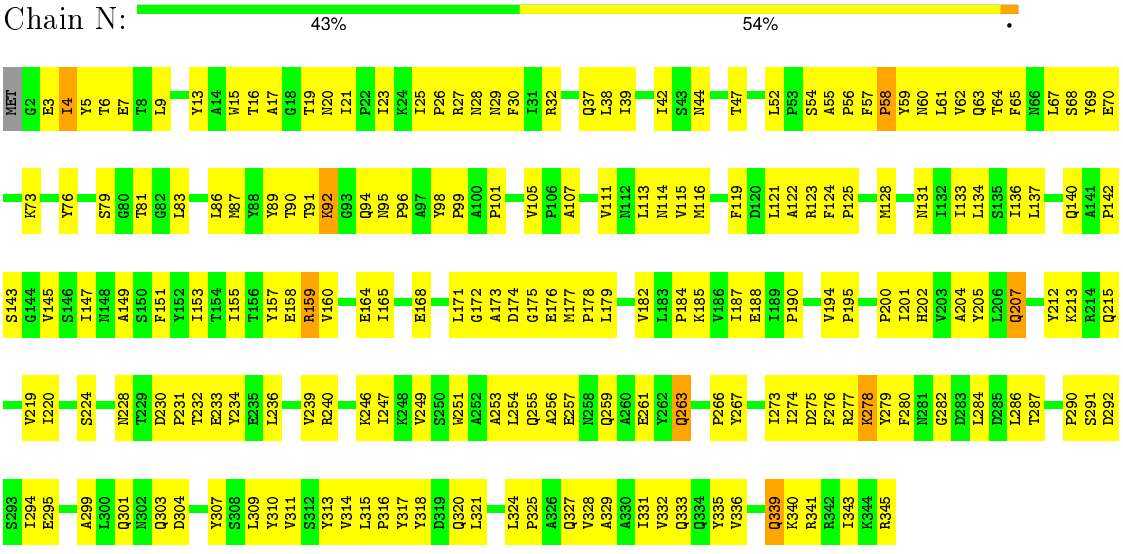
- Molecule 3: Coat protein

Chain M:



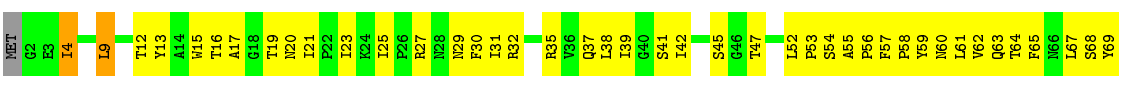
- Molecule 3: Coat protein

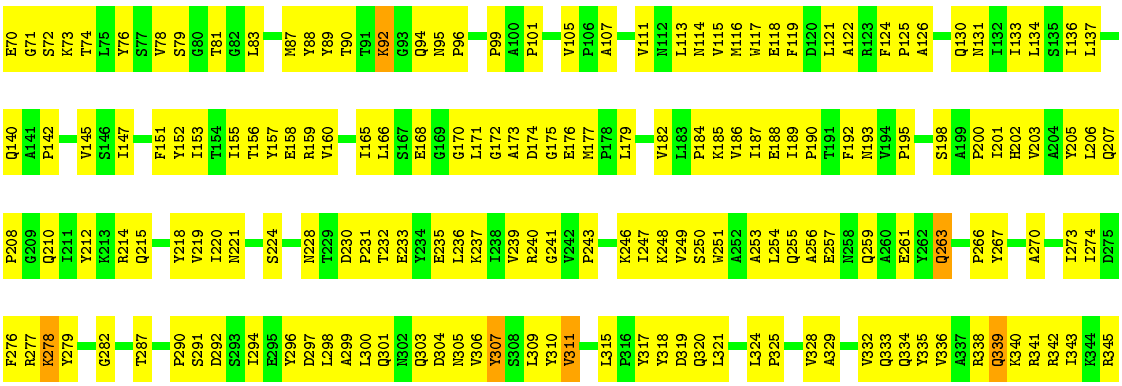
Chain N:



- Molecule 3: Coat protein

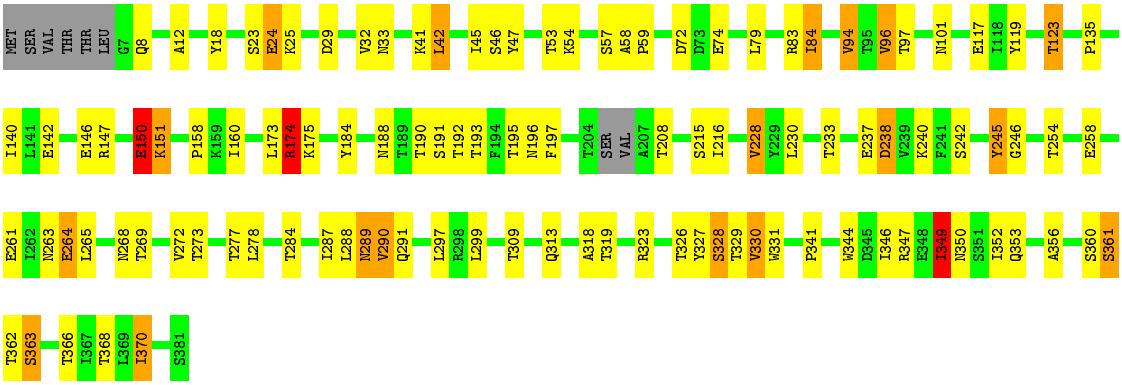
Chain O:





- Molecule 4: C381 turret protein

Chain P:



4 Experimental information (i)

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, I	Depositor
Number of images	8903	Depositor
Resolution determination method	FSC at 0.143 cut-off. The reported resolution is for the entire reconstruction. The resolution of the coat subunit region (B345) is estimated to 3.9 A using the same criterion.	Depositor
CTF correction method	Each particle	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/Å^2$)	18	Depositor
Minimum defocus (nm)	800	Depositor
Maximum defocus (nm)	2500	Depositor
Magnification	Not provided	Depositor
Image detector	FEI Falcon	Depositor

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	#$\|Z\|$ >2	RMSZ	#$\|Z\|$ >2
1	Q	1.43	20/1626 (1.2%)	1.27	14/2223 (0.6%)
3	A	0.60	0/2729	0.78	2/3732 (0.1%)
3	B	0.59	0/2729	0.76	1/3732 (0.0%)
3	C	0.61	0/2729	0.78	1/3732 (0.0%)
3	D	0.57	0/2729	0.77	0/3732
3	E	0.59	0/2729	0.79	3/3732 (0.1%)
3	F	0.58	0/2729	0.76	0/3732
3	G	0.60	0/2729	0.77	2/3732 (0.1%)
3	H	0.57	0/2729	0.78	0/3732
3	I	0.59	0/2729	0.78	0/3732
3	J	0.58	0/2729	0.77	1/3732 (0.0%)
3	K	0.62	2/2729 (0.1%)	0.82	6/3732 (0.2%)
3	L	0.65	0/2729	0.80	1/3732 (0.0%)
3	M	0.58	0/2729	0.75	0/3732
3	N	0.59	0/2729	0.76	0/3732
3	O	0.59	0/2729	0.77	1/3732 (0.0%)
4	P	1.49	23/2950 (0.8%)	1.40	28/4014 (0.7%)
All	All	0.73	45/45511 (0.1%)	0.85	60/62217 (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	Q	0	1
3	A	0	1
3	C	0	1
3	I	0	1
3	K	0	1
3	M	0	1
3	N	0	1
3	O	0	1

Continued on next page...

Continued from previous page...

Mol	Chain	#Chirality outliers	#Planarity outliers
All	All	0	8

The worst 5 of 45 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	Q	212	SER	CB-OG	-21.55	1.14	1.42
1	Q	198	PRO	N-CD	16.47	1.71	1.47
4	P	146	GLU	CG-CD	10.68	1.68	1.51
1	Q	152	GLN	CB-CG	-10.39	1.24	1.52
1	Q	204	GLN	CG-CD	-9.04	1.30	1.51

The worst 5 of 60 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	P	328	SER	N-CA-CB	-24.24	74.13	110.50
4	P	264	GLU	N-CA-CB	-17.25	79.56	110.60
4	P	8	GLN	N-CA-CB	-16.41	81.07	110.60
4	P	327	TYR	N-CA-CB	-15.55	82.61	110.60
4	P	238	ASP	N-CA-CB	-13.43	86.42	110.60

There are no chirality outliers.

5 of 8 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
3	A	152	TYR	Sidechain
3	C	307	TYR	Sidechain
3	I	98	TYR	Sidechain
3	K	152	TYR	Sidechain
1	Q	154	TYR	Sidechain

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	Q	1594	0	1530	147	0
2	R	75	0	18	6	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	A	2666	0	2661	236	0
3	B	2666	0	2661	255	0
3	C	2666	0	2661	266	0
3	D	2666	0	2661	204	0
3	E	2666	0	2661	258	0
3	F	2666	0	2661	258	0
3	G	2666	0	2661	237	0
3	H	2666	0	2661	240	0
3	I	2666	0	2661	246	0
3	J	2666	0	2661	226	0
3	K	2666	0	2661	229	0
3	L	2666	0	2661	246	0
3	M	2666	0	2661	252	0
3	N	2666	0	2661	236	0
3	O	2666	0	2661	255	0
4	P	2890	0	2865	82	0
All	All	44549	0	44328	3688	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 42.

The worst 5 of 3688 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:Q:140:TYR:CE1	1:Q:157:PRO:HG3	1.42	1.53
1:Q:198:PRO:N	1:Q:198:PRO:CD	1.70	1.35
1:Q:186:MET:CE	1:Q:186:MET:SD	2.15	1.33
1:Q:140:TYR:CD1	1:Q:157:PRO:HG3	1.65	1.30
1:Q:146:PHE:HD1	1:Q:147:GLY:N	1.34	1.24

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	Q	218/223 (98%)	209 (96%)	9 (4%)	0	100	100
3	A	342/345 (99%)	302 (88%)	40 (12%)	0	100	100
3	B	342/345 (99%)	305 (89%)	37 (11%)	0	100	100
3	C	342/345 (99%)	303 (89%)	39 (11%)	0	100	100
3	D	342/345 (99%)	304 (89%)	36 (10%)	2 (1%)	30	74
3	E	342/345 (99%)	304 (89%)	36 (10%)	2 (1%)	30	74
3	F	342/345 (99%)	297 (87%)	44 (13%)	1 (0%)	46	83
3	G	342/345 (99%)	300 (88%)	40 (12%)	2 (1%)	30	74
3	H	342/345 (99%)	302 (88%)	39 (11%)	1 (0%)	46	83
3	I	342/345 (99%)	305 (89%)	36 (10%)	1 (0%)	46	83
3	J	342/345 (99%)	299 (87%)	43 (13%)	0	100	100
3	K	342/345 (99%)	305 (89%)	36 (10%)	1 (0%)	46	83
3	L	342/345 (99%)	307 (90%)	34 (10%)	1 (0%)	46	83
3	M	342/345 (99%)	298 (87%)	43 (13%)	1 (0%)	46	83
3	N	342/345 (99%)	304 (89%)	37 (11%)	1 (0%)	46	83
3	O	342/345 (99%)	301 (88%)	40 (12%)	1 (0%)	46	83
4	P	369/381 (97%)	349 (95%)	20 (5%)	0	100	100
All	All	5717/5779 (99%)	5094 (89%)	609 (11%)	14 (0%)	56	86

5 of 14 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	F	71	GLY
3	G	318	TYR
3	H	318	TYR
3	E	318	TYR
3	I	318	TYR

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	Q	163/198 (82%)	150 (92%)	13 (8%)	15 53
3	A	289/290 (100%)	272 (94%)	17 (6%)	24 63
3	B	289/290 (100%)	275 (95%)	14 (5%)	31 69
3	C	289/290 (100%)	273 (94%)	16 (6%)	27 66
3	D	289/290 (100%)	281 (97%)	8 (3%)	51 79
3	E	289/290 (100%)	278 (96%)	11 (4%)	40 74
3	F	289/290 (100%)	276 (96%)	13 (4%)	34 70
3	G	289/290 (100%)	280 (97%)	9 (3%)	47 78
3	H	289/290 (100%)	276 (96%)	13 (4%)	34 70
3	I	289/290 (100%)	271 (94%)	18 (6%)	23 62
3	J	289/290 (100%)	276 (96%)	13 (4%)	34 70
3	K	289/290 (100%)	270 (93%)	19 (7%)	21 60
3	L	289/290 (100%)	276 (96%)	13 (4%)	34 70
3	M	289/290 (100%)	275 (95%)	14 (5%)	31 69
3	N	289/290 (100%)	275 (95%)	14 (5%)	31 69
3	O	289/290 (100%)	276 (96%)	13 (4%)	34 70
4	P	326/334 (98%)	281 (86%)	45 (14%)	4 29
All	All	4824/4882 (99%)	4561 (94%)	263 (6%)	31 66

5 of 263 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	I	66	ASN
3	K	9	LEU
4	P	272	VAL
3	I	143	SER
3	J	28	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 112 such sidechains are listed below:

Mol	Chain	Res	Type
3	F	255	GLN
3	I	131	ASN
3	O	37	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
3	G	114	ASN
3	H	140	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.