wwPDB EM Map/Model Validation Report (i)

Apr 10, 2016 – 02:03 PM BST

PDB ID	:	3IYN
EMDB ID:	:	EMD-5172
Title	:	3.6-Angstrom cryoEM structure of human adenovirus type 5
Authors	:	Liu, H.; Jin, L.; Koh, S.B.S.; Atanasov, I.; Schein, S.; Wu, L.; Zhou, Z.H.
Deposited on	:	2010-03-09
Resolution	:	3.60 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

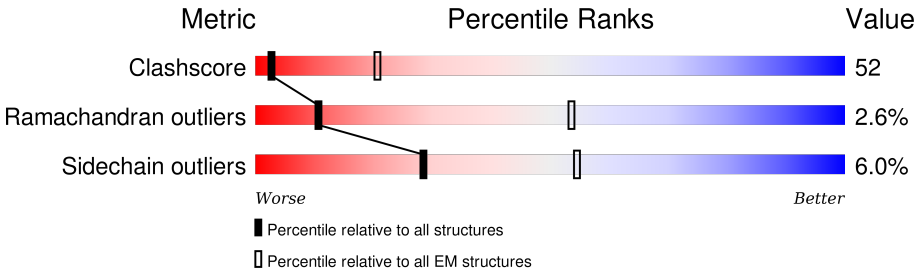
MolProbity	:	4.02b-467
Mogul	:	unknown
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk27241

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 3.60 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.



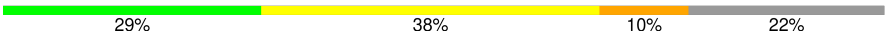
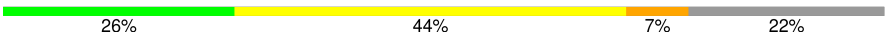
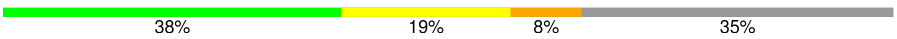
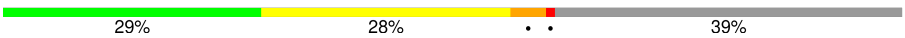
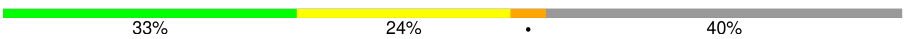
Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	114402	924
Ramachandran outliers	111179	726
Sidechain outliers	111093	686

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	952	56% 35% 5% •
1	B	952	54% 36% 5% •
1	C	952	52% 38% 6% • •
1	D	952	55% 36% 5% •
1	E	952	54% 37% 5% •
1	F	952	52% 38% 6% •
1	G	952	50% 40% 6% •
1	H	952	52% 38% 6% •
1	I	952	55% 36% 5% •

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	J	952	
1	K	952	
1	L	952	
2	M	571	
3	N	585	
4	O	227	
4	P	227	
5	Q	140	
5	R	140	
5	S	140	
5	T	140	

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 99595 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 Hexon protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	915	Total	C	N	O	S	0	0
			7327	4659	1241	1392	35		
1	B	915	Total	C	N	O	S	0	0
			7327	4659	1241	1392	35		
1	C	920	Total	C	N	O	S	0	0
			7362	4680	1246	1400	36		
1	D	916	Total	C	N	O	S	0	0
			7335	4664	1242	1393	36		
1	E	917	Total	C	N	O	S	0	0
			7342	4668	1243	1395	36		
1	F	917	Total	C	N	O	S	0	0
			7342	4668	1243	1395	36		
1	G	917	Total	C	N	O	S	0	0
			7342	4668	1243	1395	36		
1	H	919	Total	C	N	O	S	0	0
			7355	4676	1245	1398	36		
1	I	917	Total	C	N	O	S	0	0
			7342	4668	1243	1395	36		
1	J	918	Total	C	N	O	S	0	0
			7349	4672	1244	1397	36		
1	K	919	Total	C	N	O	S	0	0
			7355	4676	1245	1398	36		
1	L	919	Total	C	N	O	S	0	0
			7355	4676	1245	1398	36		

- Molecule 2 is a protein called Penton base protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	M	455	Total	C	N	O	S	0	0
			3642	2308	632	690	12		

- Molecule 3 is a protein called Peripentonal hexon-associated protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	N	284	Total	C	N	O	S	0	0
			2208	1368	406	430	4		

- Molecule 4 is a protein called Hexon-associated protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
4	O	176	Total	C	N	O	S	0	0
			1361	859	240	257	5		
4	P	176	Total	C	N	O	S	0	0
			1361	859	240	257	5		

- Molecule 5 is a protein called Hexon-associated protein.

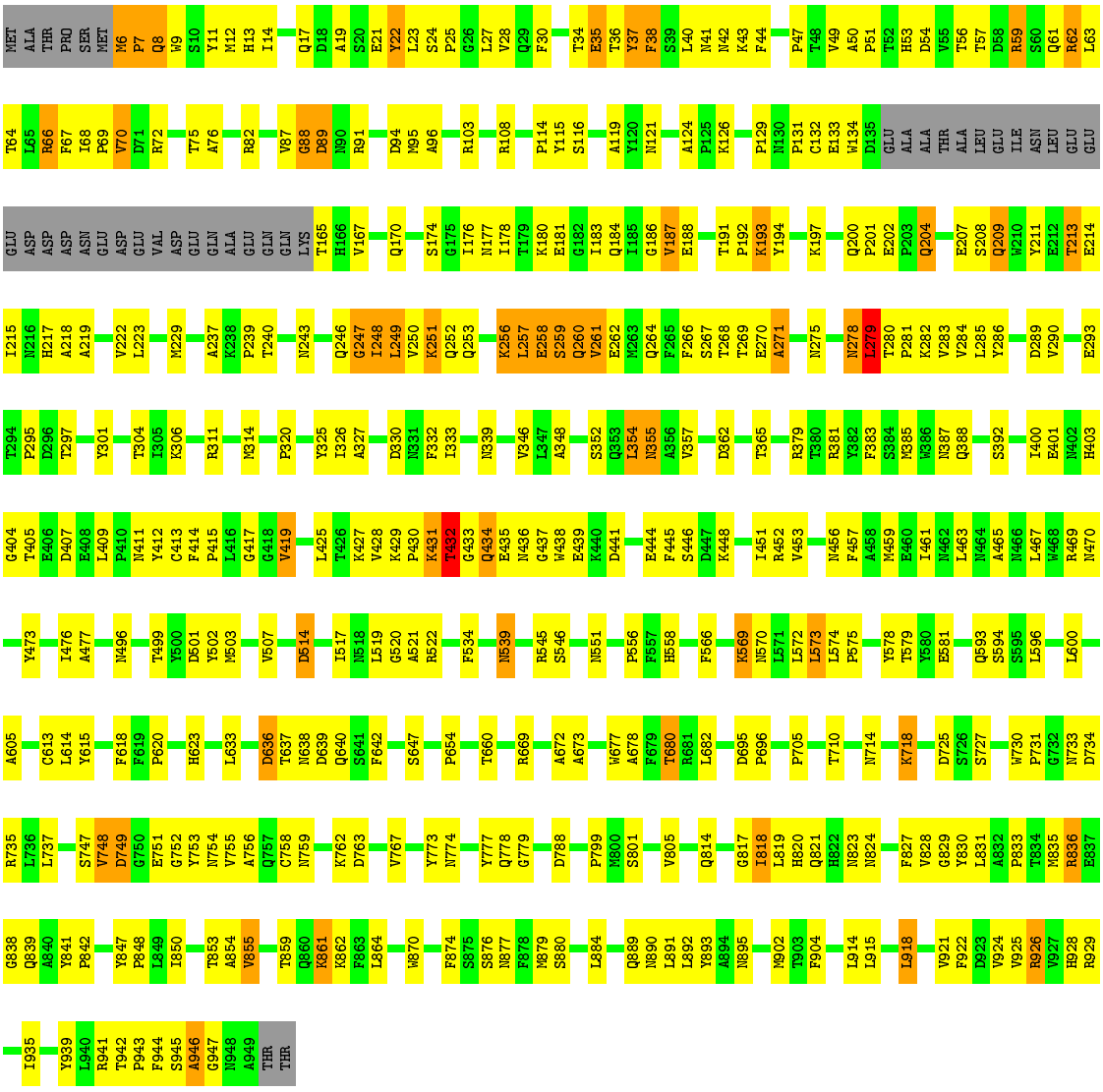
Mol	Chain	Residues	Atoms					AltConf	Trace
5	Q	128	Total	C	N	O	S	0	0
			927	571	164	190	2		
5	R	91	Total	C	N	O	S	0	0
			683	422	121	138	2		
5	S	86	Total	C	N	O	S	0	0
			647	397	115	133	2		
5	T	84	Total	C	N	O	S	0	0
			633	393	111	127	2		

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.

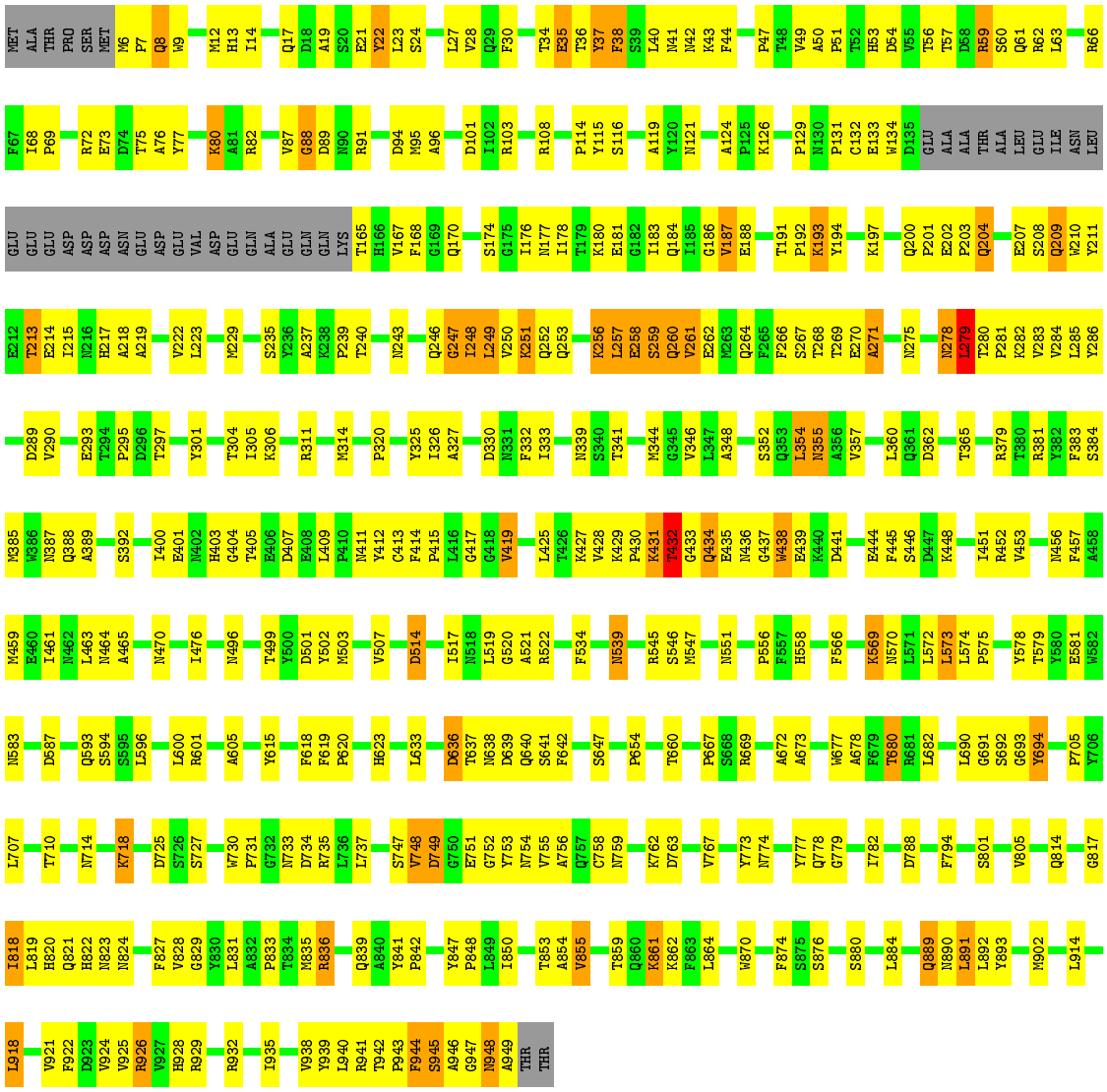
- Molecule 1: Hexon protein

Chain A:



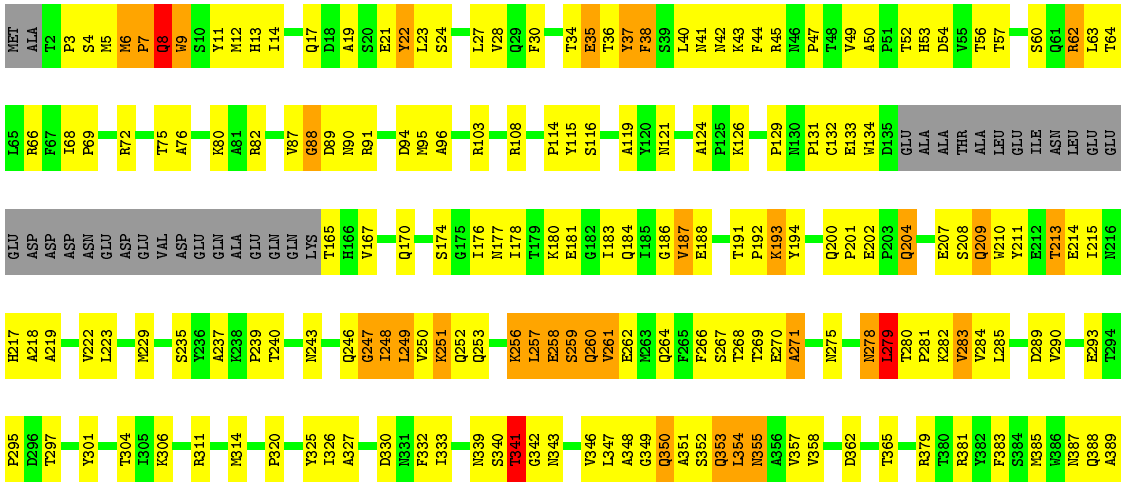
- Molecule 1: Hexon protein

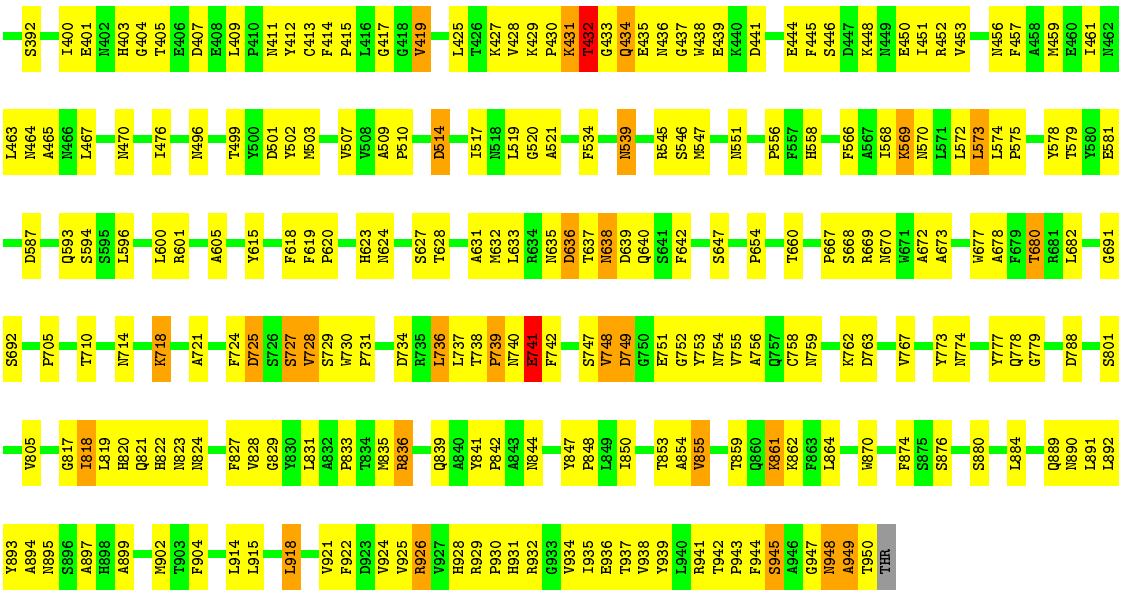
Chain B:



• Molecule 1: Hexon protein

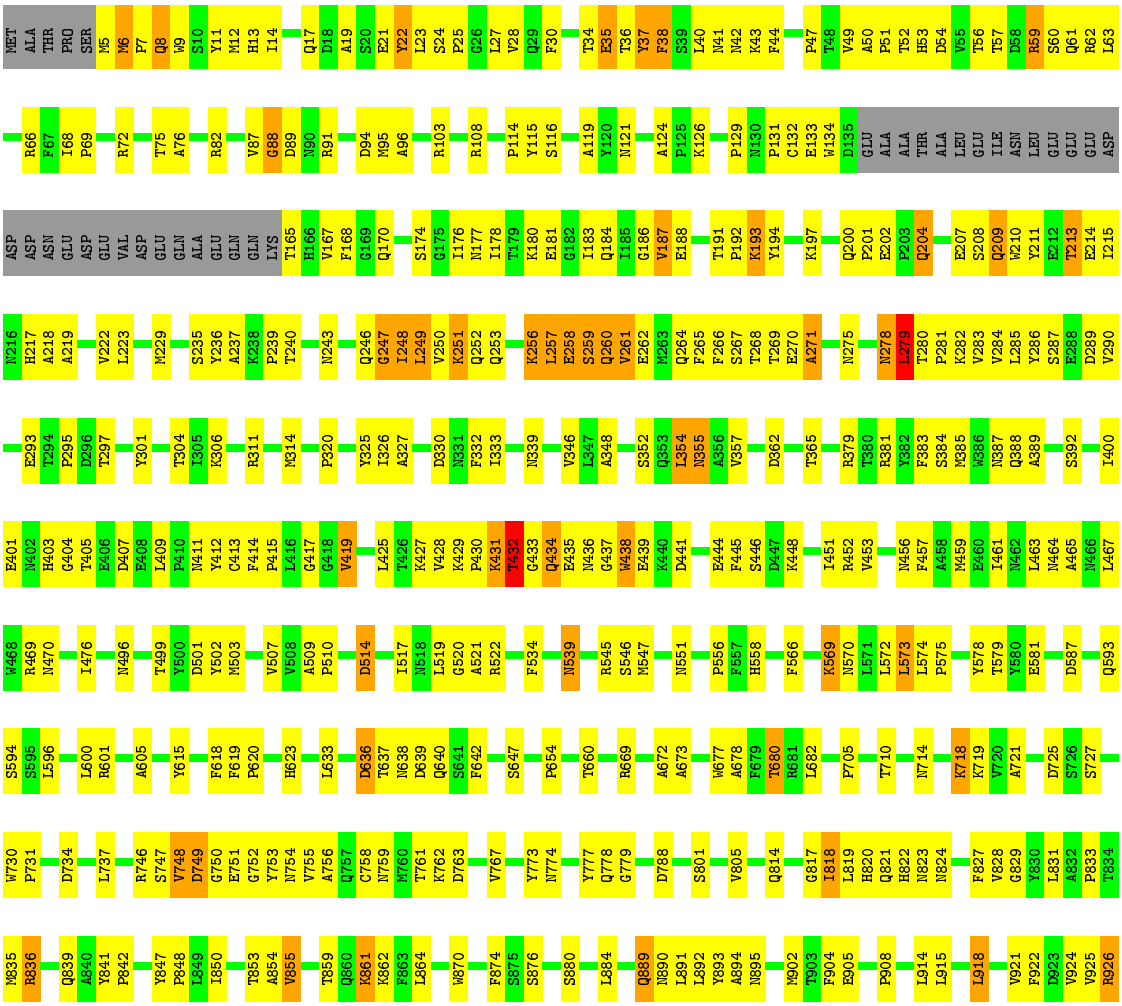
Chain C:





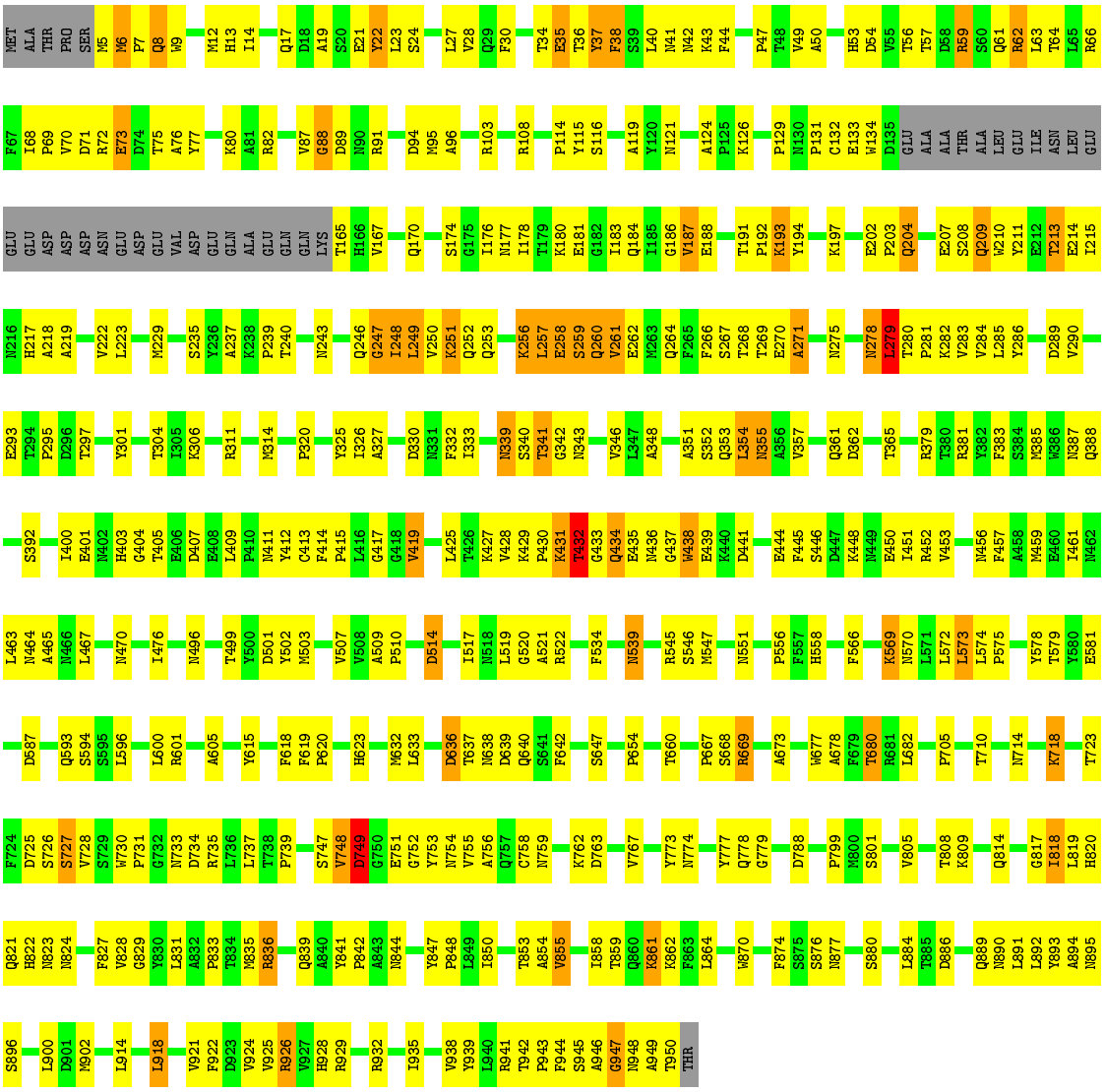
• Molecule 1: Hexon protein

Chain D:



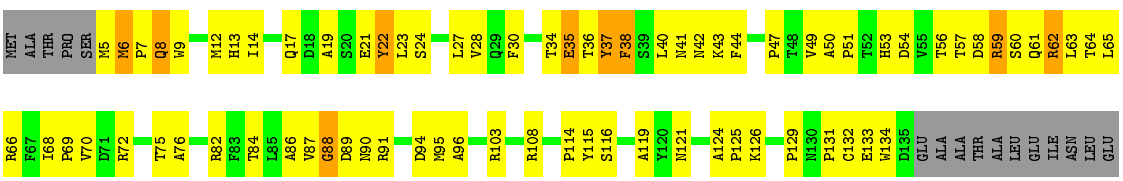
- Molecule 1: Hexon protein

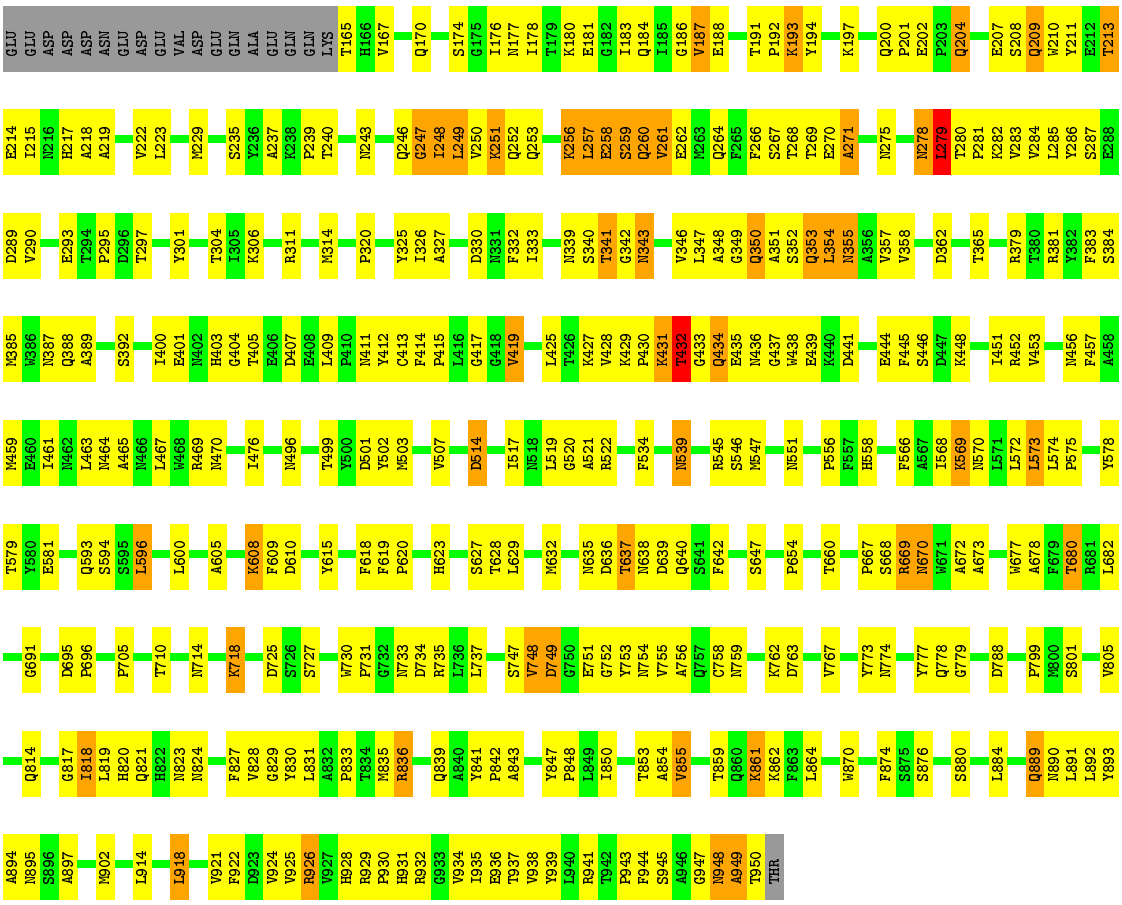
Chain E:



- Molecule 1: Hexon protein

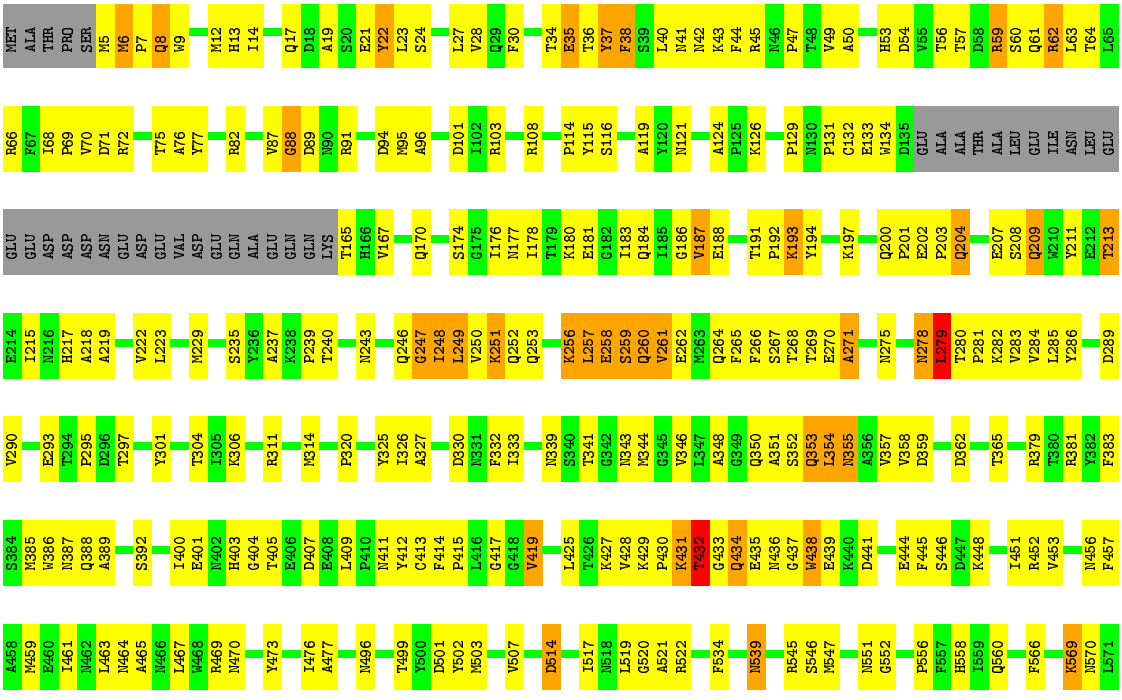
Chain F:

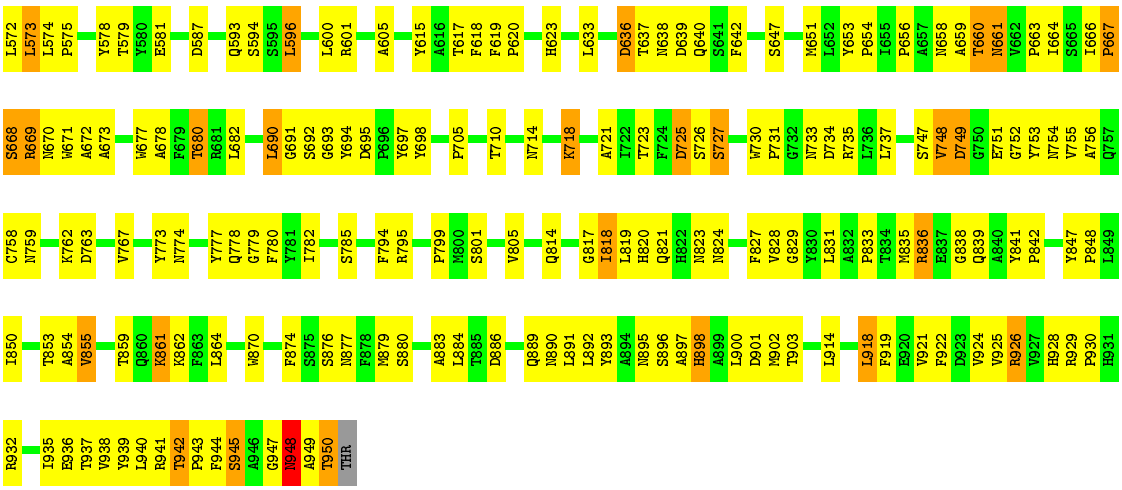




- Molecule 1: Hexon protein

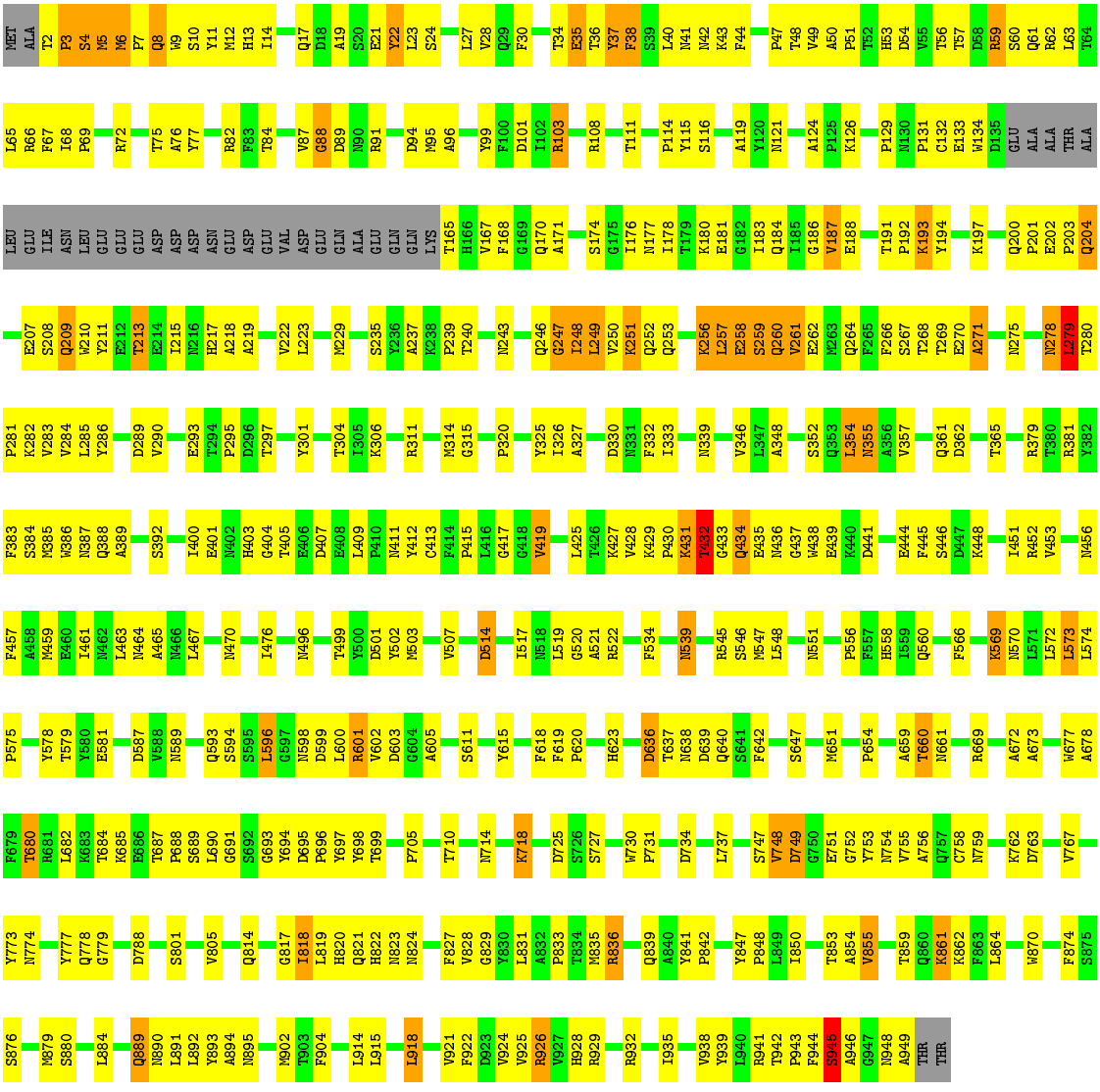
Chain G:





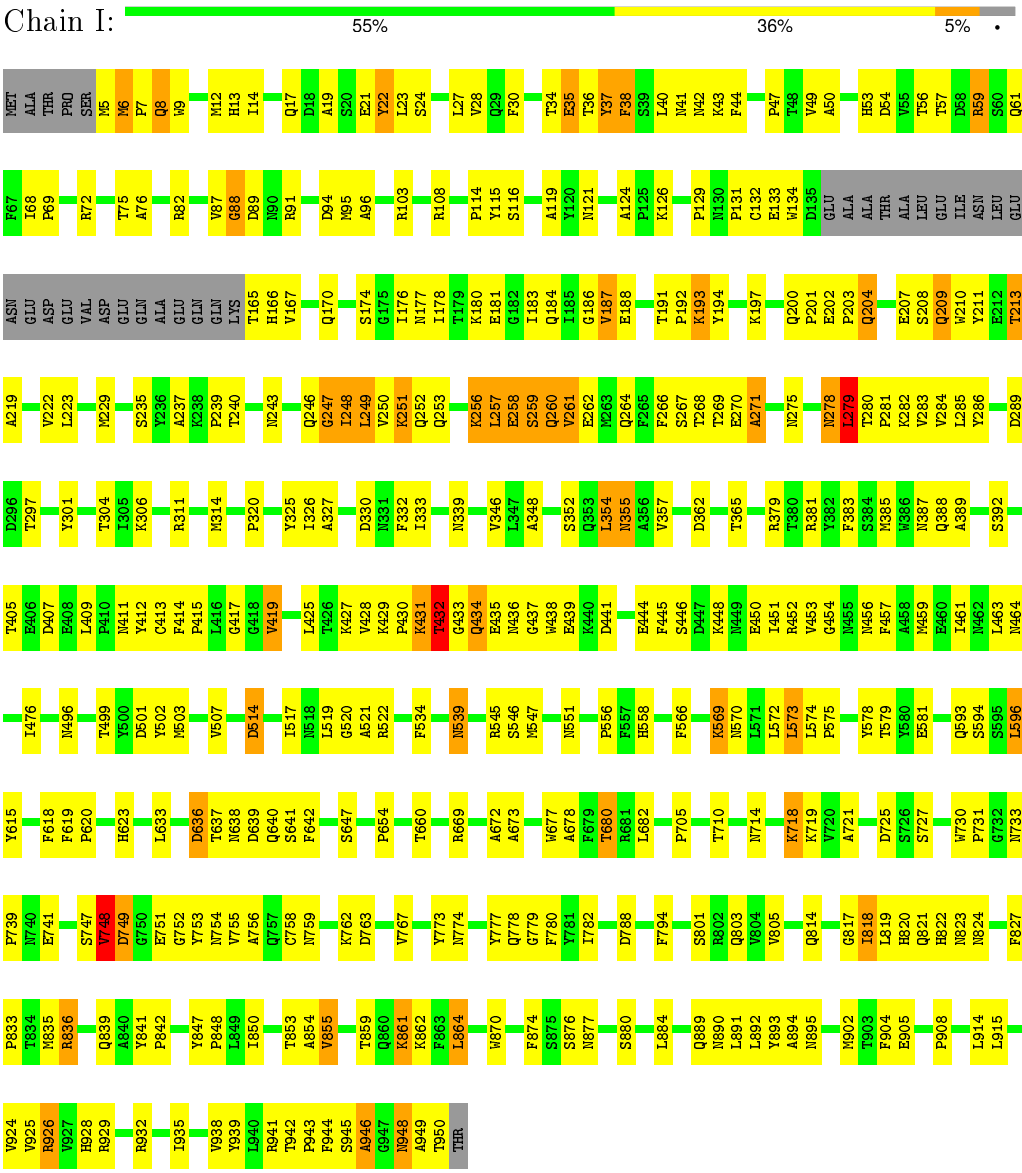
• Molecule 1: Hexon protein

Chain H:



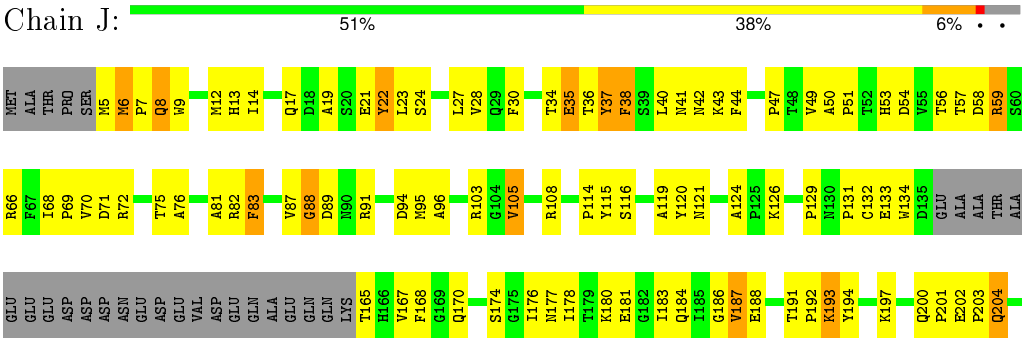
- Molecule 1: Hexon protein

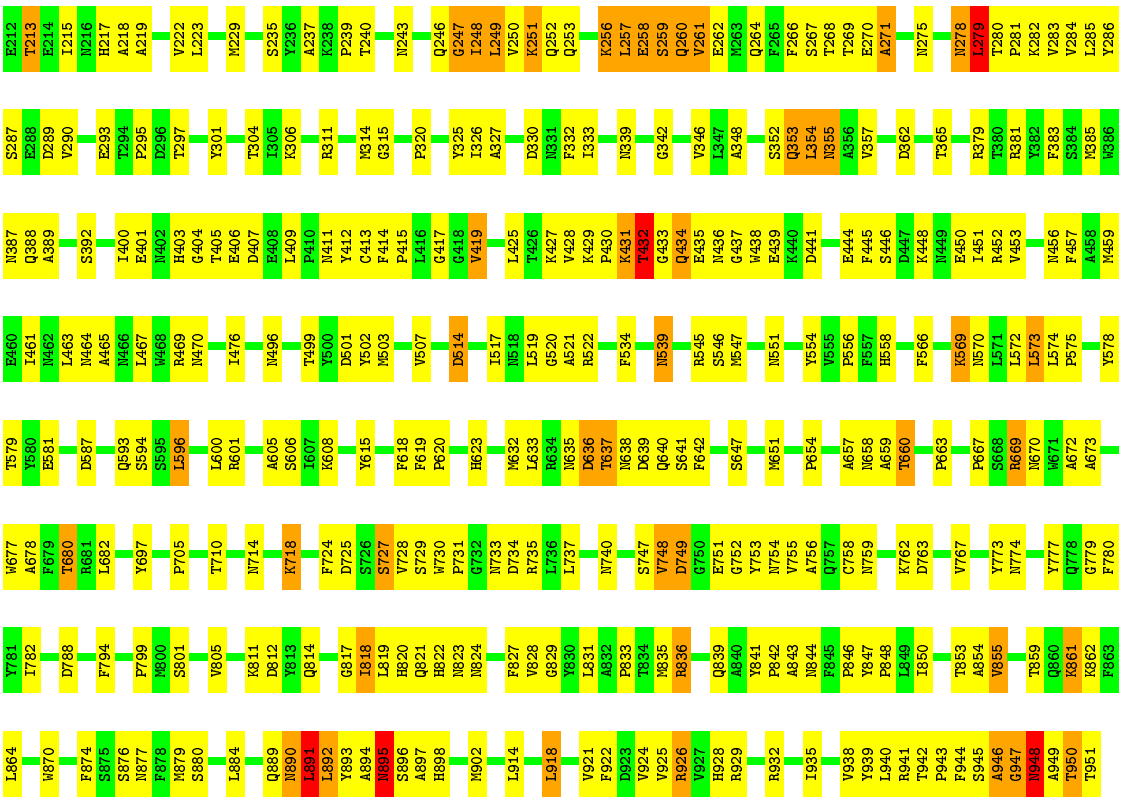
Chain I:



- Molecule 1: Hexon protein

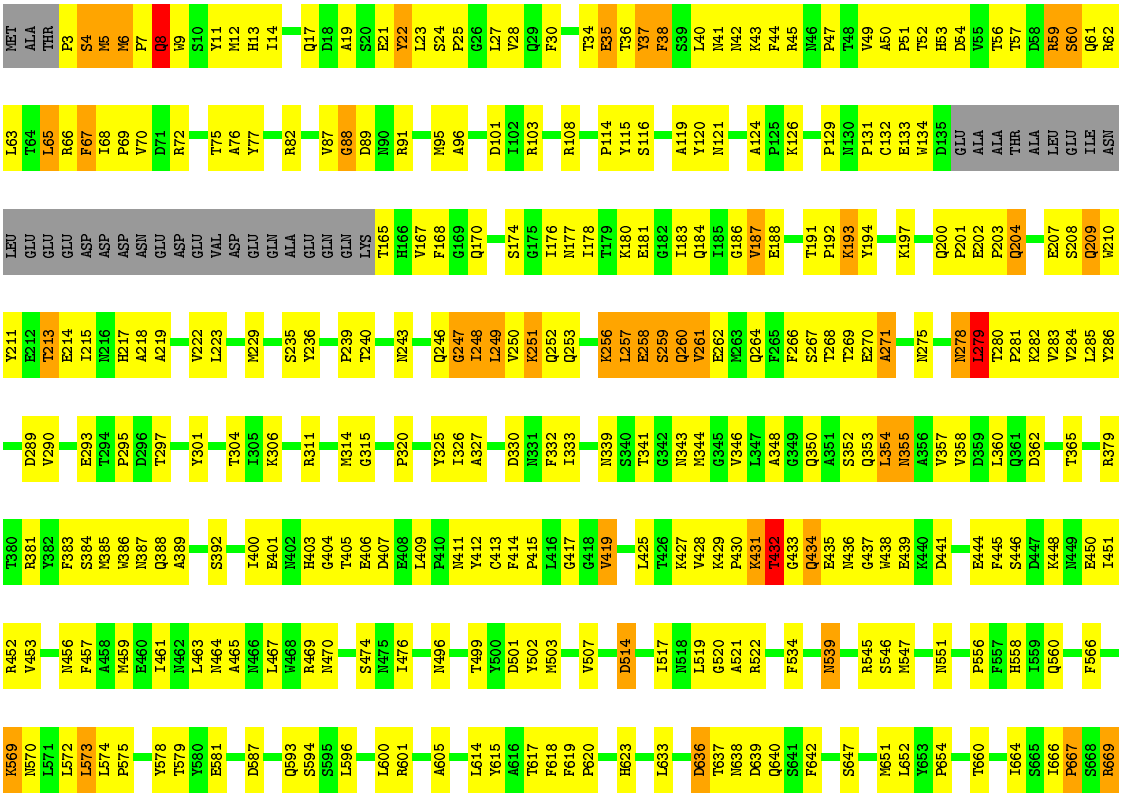
Chain J:

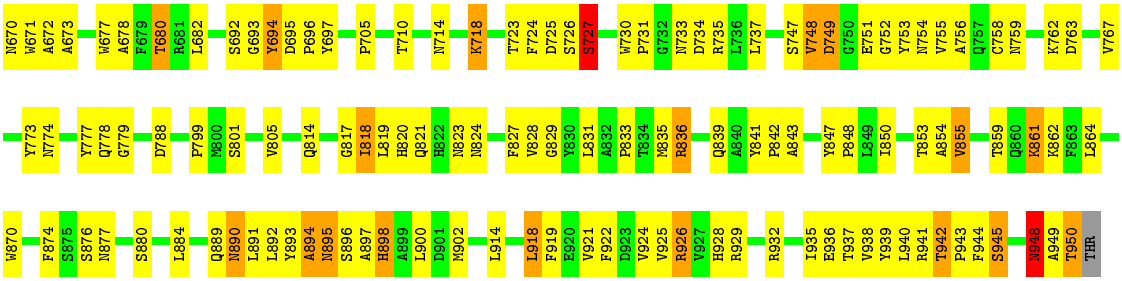


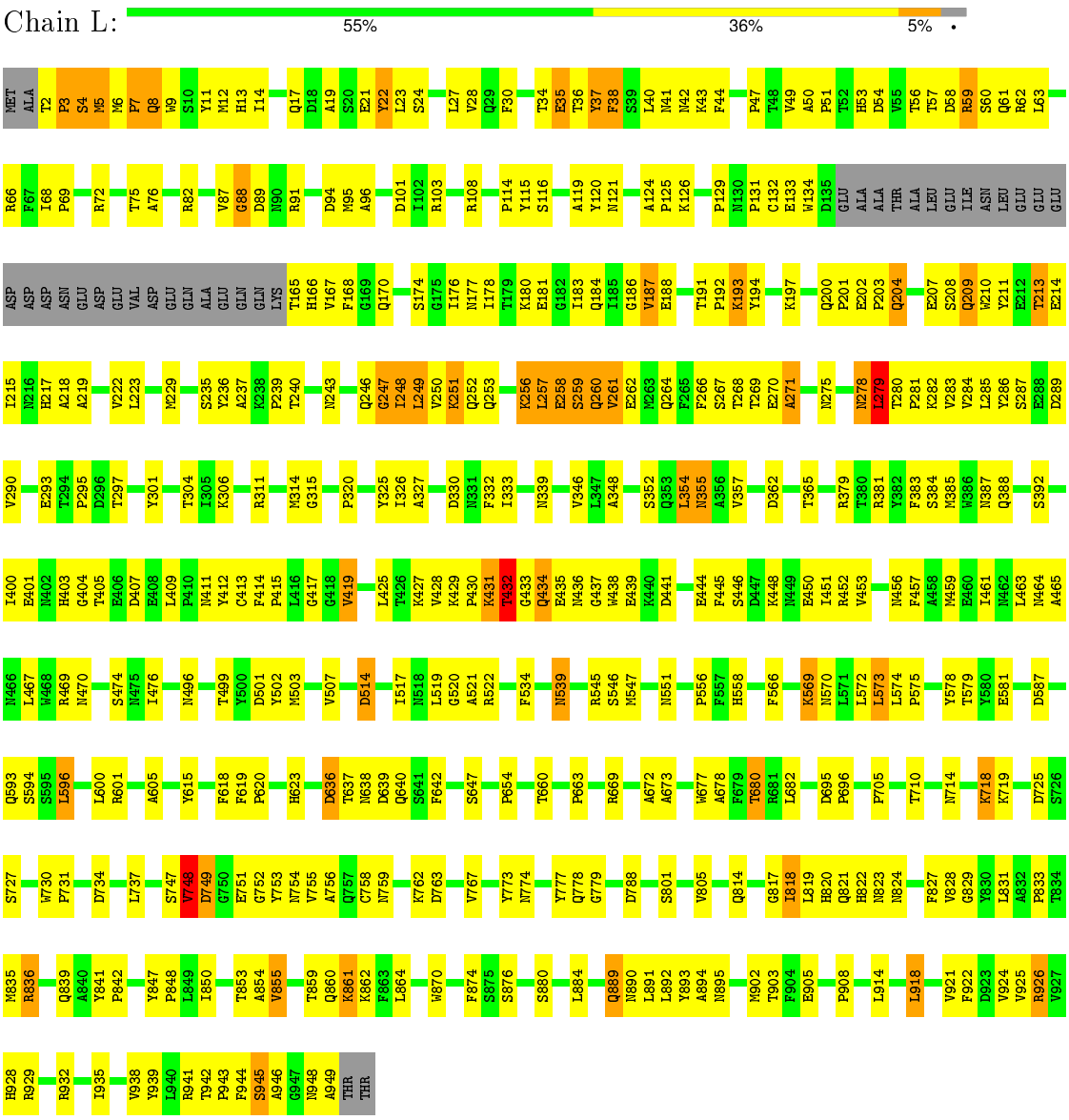


• Molecule 1: Hexon protein

Chain K:





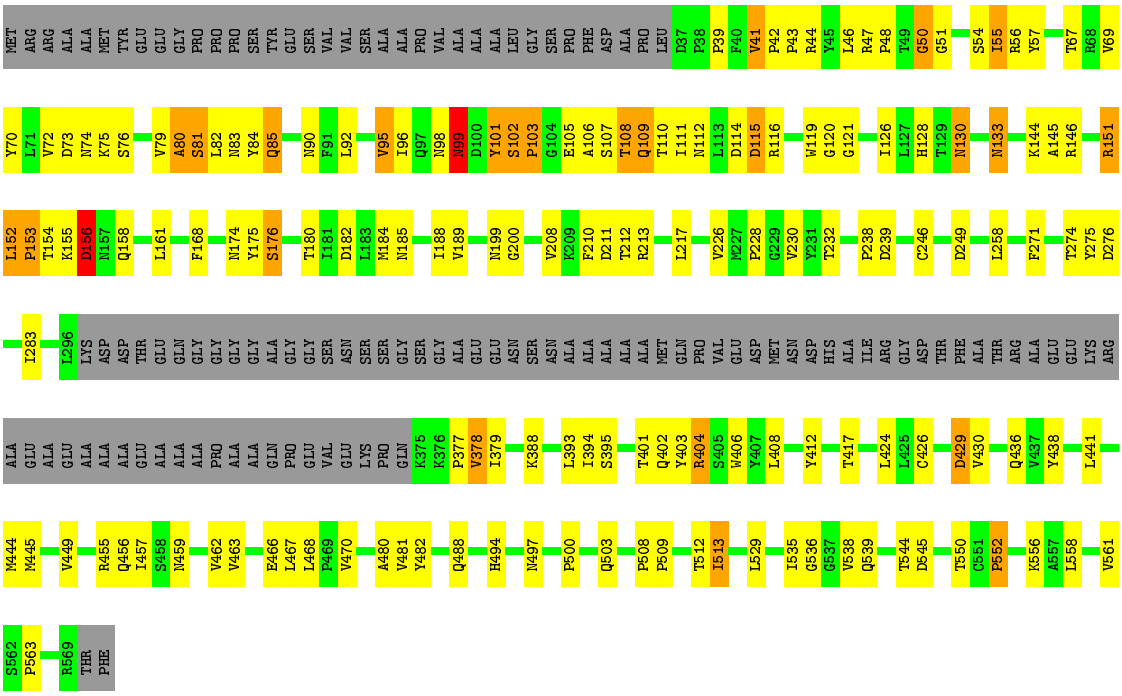


• Molecule 1: Hexon protein

Chain L:

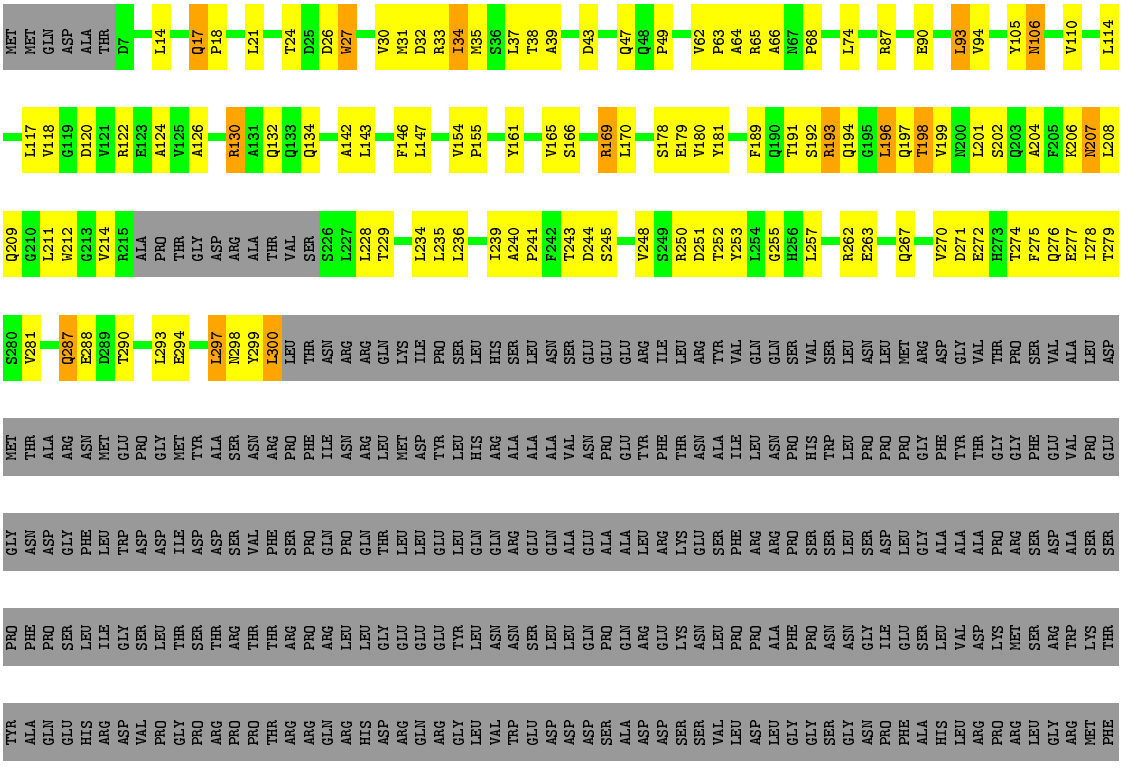
• Molecule 2: Penton base protein

Chain M:



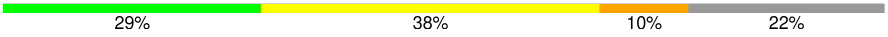
• Molecule 3: Peripentonal hexon-associated protein

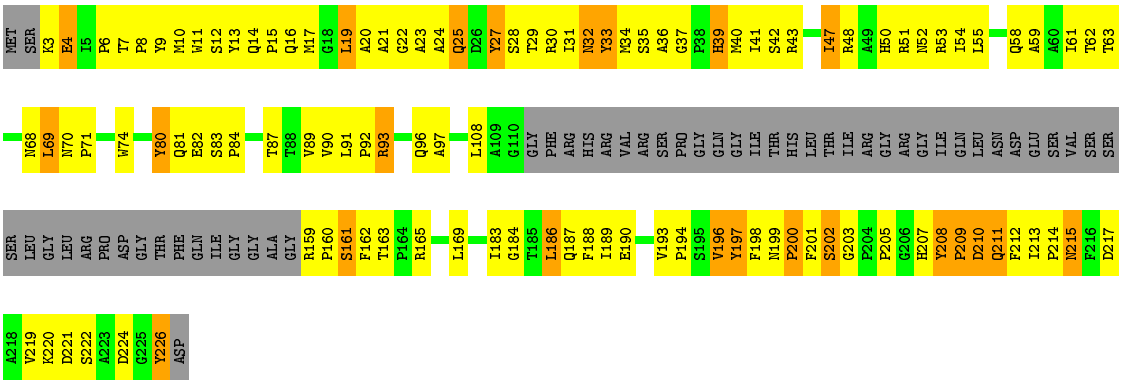
Chain N:



• Molecule 4: Hexon-associated protein

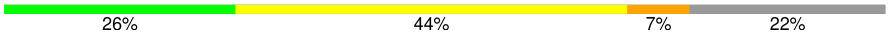
Chain O:

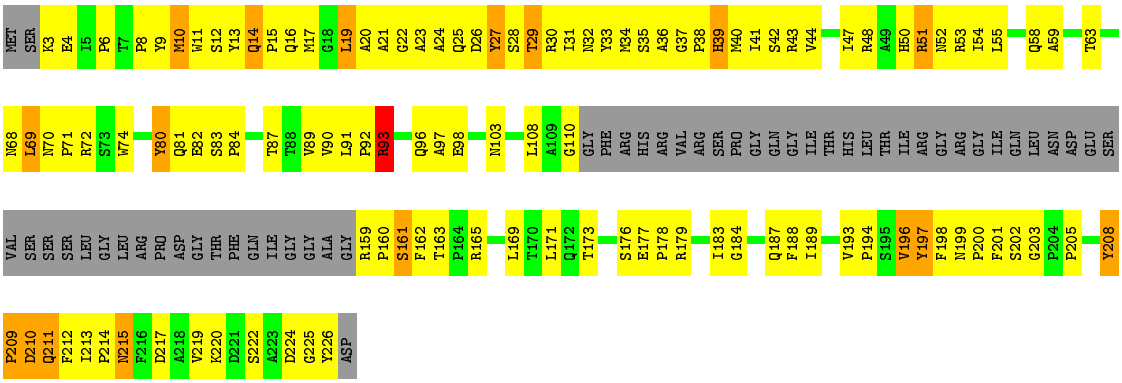




• Molecule 4: Hexon-associated protein

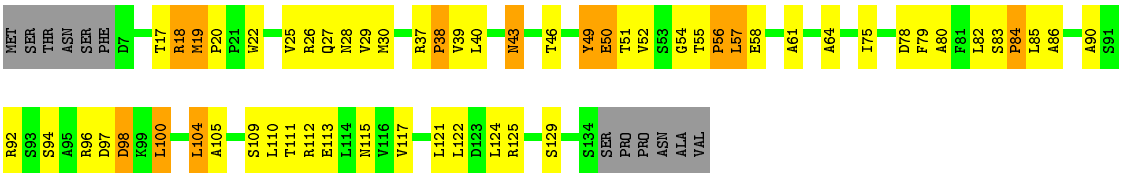
Chain P:





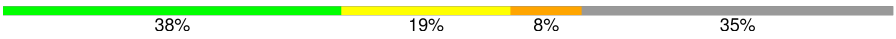
• Molecule 5: Hexon-associated protein

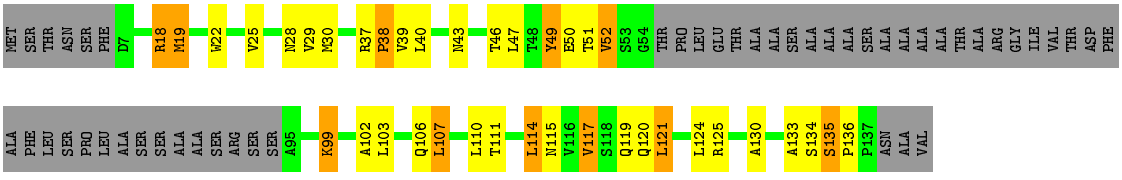
Chain Q:



• Molecule 5: Hexon-associated protein

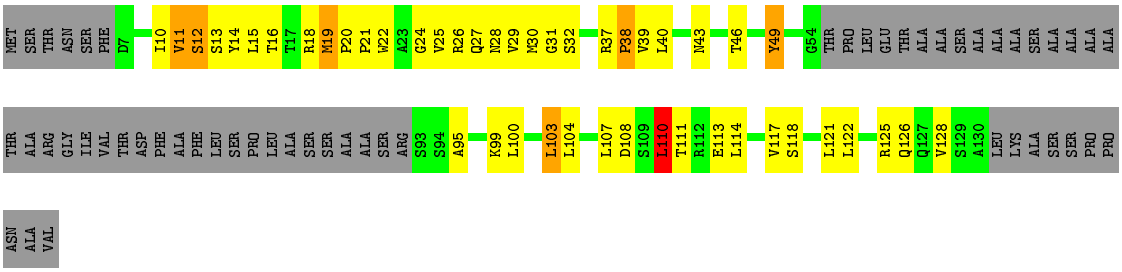
Chain R:





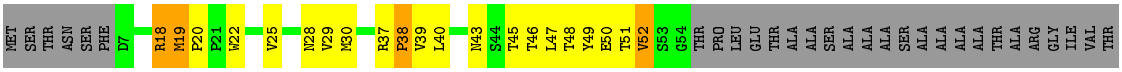
• Molecule 5: Hexon-associated protein

Chain S:



• Molecule 5: Hexon-associated protein

Chain T:



4 Experimental information (i)

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, I	Depositor
Number of images	31815	Depositor
Resolution determination method	FSC at 0.5 cut-off	Depositor
CTF correction method	Each particle	Depositor
Microscope	FEI Titan Krios	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/Å^2$)	20	Depositor
Minimum defocus (nm)	1000	Depositor
Maximum defocus (nm)	2500	Depositor
Magnification	59000	Depositor
Image detector	Kodak SO163 film	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	A	0.34	0/7525	0.61	2/10234 (0.0%)
1	B	0.34	0/7525	0.61	2/10234 (0.0%)
1	C	0.34	0/7561	0.60	2/10284 (0.0%)
1	D	0.34	0/7533	0.61	1/10244 (0.0%)
1	E	0.34	0/7540	0.61	2/10254 (0.0%)
1	F	0.34	0/7540	0.61	2/10254 (0.0%)
1	G	0.34	0/7540	0.61	2/10254 (0.0%)
1	H	0.34	0/7554	0.61	2/10274 (0.0%)
1	I	0.34	0/7540	0.61	2/10254 (0.0%)
1	J	0.34	0/7547	0.61	2/10264 (0.0%)
1	K	0.34	0/7554	0.60	2/10273 (0.0%)
1	L	0.34	0/7554	0.61	2/10274 (0.0%)
2	M	0.33	0/3733	0.54	0/5088
3	N	0.35	0/2243	0.68	0/3051
4	O	0.20	0/1400	0.36	0/1913
4	P	0.20	0/1400	0.36	0/1913
5	Q	0.30	0/938	0.49	0/1280
5	R	0.33	0/691	0.53	0/941
5	S	0.32	0/653	0.55	1/888 (0.1%)
5	T	0.33	0/639	0.53	0/869
All	All	0.33	0/102210	0.60	24/139040 (0.0%)

There are no bond length outliers.

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

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	H	247	GLY	N-CA-C	5.95	127.96	113.10
1	E	247	GLY	N-CA-C	5.93	127.93	113.10
1	K	247	GLY	N-CA-C	5.92	127.90	113.10
1	D	247	GLY	N-CA-C	5.92	127.89	113.10
1	B	247	GLY	N-CA-C	5.92	127.89	113.10

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	7327	0	7028	845	0
1	B	7327	0	7028	842	0
1	C	7362	0	7063	1044	0
1	D	7335	0	7037	846	0
1	E	7342	0	7044	1000	0
1	F	7342	0	7044	1083	0
1	G	7342	0	7044	1164	0
1	H	7355	0	7056	1002	0
1	I	7342	0	7044	851	0
1	J	7349	0	7051	1093	0
1	K	7355	0	7057	1139	0
1	L	7355	0	7056	884	0
2	M	3642	0	3571	227	0
3	N	2208	0	2180	288	0
4	O	1361	0	1311	565	0
4	P	1361	0	1311	572	0
5	Q	927	0	939	218	0
5	R	683	0	695	98	0
5	S	647	0	652	118	0
5	T	633	0	649	139	0
All	All	99595	0	95860	10167	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 52.

The worst 5 of 10167 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:J:654:PRO:CG	5:Q:84:PRO:HG2	1.33	1.58
1:J:728:VAL:HG11	4:P:226:TYR:CD1	1.37	1.57
1:C:632:MET:HG2	4:O:201:PHE:CD2	1.38	1.55
1:J:660:THR:HB	5:R:130:ALA:CB	1.37	1.54
1:A:88:GLY:HA3	1:E:353:GLN:CG	1.37	1.52

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	911/952 (96%)	821 (90%)	71 (8%)	19 (2%)	9	52
1	B	911/952 (96%)	824 (90%)	67 (7%)	20 (2%)	8	51
1	C	916/952 (96%)	817 (89%)	73 (8%)	26 (3%)	6	45
1	D	912/952 (96%)	821 (90%)	72 (8%)	19 (2%)	9	52
1	E	913/952 (96%)	821 (90%)	72 (8%)	20 (2%)	8	51
1	F	913/952 (96%)	816 (89%)	75 (8%)	22 (2%)	7	49
1	G	913/952 (96%)	819 (90%)	71 (8%)	23 (2%)	7	48
1	H	915/952 (96%)	824 (90%)	71 (8%)	20 (2%)	8	51
1	I	913/952 (96%)	827 (91%)	67 (7%)	19 (2%)	9	52
1	J	914/952 (96%)	821 (90%)	69 (8%)	24 (3%)	7	46
1	K	915/952 (96%)	817 (89%)	73 (8%)	25 (3%)	6	46
1	L	915/952 (96%)	824 (90%)	71 (8%)	20 (2%)	8	51
2	M	451/571 (79%)	341 (76%)	85 (19%)	25 (6%)	2	26
3	N	280/585 (48%)	248 (89%)	27 (10%)	5 (2%)	11	54
4	O	172/227 (76%)	140 (81%)	24 (14%)	8 (5%)	3	30
4	P	172/227 (76%)	139 (81%)	26 (15%)	7 (4%)	3	34
5	Q	126/140 (90%)	102 (81%)	18 (14%)	6 (5%)	3	30
5	R	87/140 (62%)	71 (82%)	12 (14%)	4 (5%)	3	31
5	S	82/140 (59%)	66 (80%)	13 (16%)	3 (4%)	4	38
5	T	80/140 (57%)	69 (86%)	8 (10%)	3 (4%)	4	37
All	All	12411/13594 (91%)	11028 (89%)	1065 (9%)	318 (3%)	11	46

5 of 318 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	35	GLU
1	A	248	ILE
1	A	249	LEU
1	A	257	LEU
1	A	261	VAL

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	796/829 (96%)	752 (94%)	44 (6%)	27 68
1	B	796/829 (96%)	751 (94%)	45 (6%)	25 67
1	C	801/829 (97%)	753 (94%)	48 (6%)	24 65
1	D	797/829 (96%)	754 (95%)	43 (5%)	27 68
1	E	798/829 (96%)	753 (94%)	45 (6%)	26 68
1	F	798/829 (96%)	754 (94%)	44 (6%)	27 68
1	G	798/829 (96%)	749 (94%)	49 (6%)	23 65
1	H	800/829 (96%)	754 (94%)	46 (6%)	25 66
1	I	798/829 (96%)	756 (95%)	42 (5%)	28 69
1	J	799/829 (96%)	751 (94%)	48 (6%)	24 65
1	K	800/829 (96%)	750 (94%)	50 (6%)	22 64
1	L	800/829 (96%)	756 (94%)	44 (6%)	27 68
2	M	412/489 (84%)	391 (95%)	21 (5%)	29 70
3	N	238/500 (48%)	220 (92%)	18 (8%)	16 56
4	O	145/186 (78%)	125 (86%)	20 (14%)	4 28
4	P	145/186 (78%)	128 (88%)	17 (12%)	7 35
5	Q	101/112 (90%)	91 (90%)	10 (10%)	10 44
5	R	79/112 (70%)	70 (89%)	9 (11%)	7 36
5	S	75/112 (67%)	67 (89%)	8 (11%)	8 40
5	T	73/112 (65%)	68 (93%)	5 (7%)	20 61

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	10849/11757 (92%)	10193 (94%)	656 (6%)	28 65

5 of 656 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	G	573	LEU
1	I	38	PHE
4	O	226	TYR
1	G	718	LYS
1	H	259	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 358 such sidechains are listed below:

Mol	Chain	Res	Type
1	G	184	GLN
1	H	370	GLN
2	M	503	GLN
1	G	339	ASN
1	G	740	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 (i)

There are no ligands in this entry.

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.