



## wwPDB EM Map/Model Validation Report ⓘ

Apr 10, 2016 – 02:16 PM BST

PDB ID : 4V4U  
EMDB ID: : EMD-1111  
Title : The quasi-atomic model of Human Adenovirus type 5 capsid  
Authors : Fabry, C.M.S.; Rosa-Calatrava, M.; Conway, J.F.; Zubieta, C.; Cusack, S.;  
Ruigrok, R.W.H.; Schoehn, G.  
Deposited on : 2005-03-03  
Resolution : 10.00 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](mailto: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

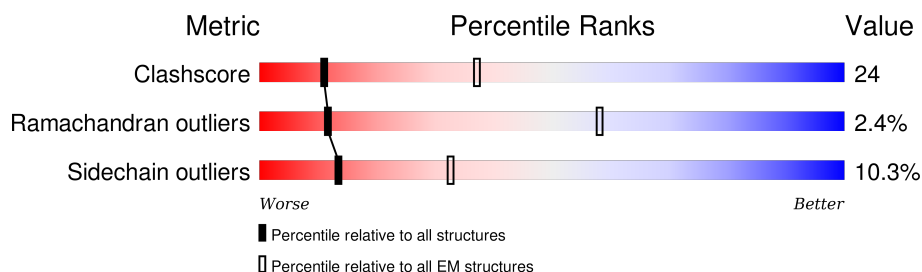
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*ELECTRON MICROSCOPY*

The reported resolution of this entry is 10.00 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)	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  $\geq 3$ , 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  $\leq 5\%$

Mol	Chain	Length	Quality of chain
1	A	523	30% 43% 11% • 16%
1	B	523	30% 42% 11% 16%
1	C	523	30% 43% 11% • 16%
1	D	523	31% 41% 11% 16%
1	E	523	30% 43% 11% 16%
2	S	10	40% 50% 10%
2	T	10	30% 60% 10%
2	U	10	50% 40% 10%
2	V	10	50% 40% 10%

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	W	10	
3	F	951	
3	G	951	
3	H	951	
3	I	951	
3	J	951	
3	K	951	
3	L	951	
3	M	951	
3	N	951	
3	O	951	
3	P	951	
3	Q	951	

## 2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 101329 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 PENTON PROTEIN.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	440	Total	C	N	O	S	0	0
			3519	2227	608	672	12		
1	B	440	Total	C	N	O	S	0	0
			3519	2227	608	672	12		
1	C	440	Total	C	N	O	S	0	0
			3519	2227	608	672	12		
1	D	440	Total	C	N	O	S	0	0
			3519	2227	608	672	12		
1	E	440	Total	C	N	O	S	0	0
			3519	2227	608	672	12		

- Molecule 2 is a protein called N-TERMINAL PEPTIDE OF FIBER PROTEIN.

Mol	Chain	Residues	Atoms				AltConf	Trace
2	S	10	Total	C	N	O	0	0
			86	58	11	17		
2	T	10	Total	C	N	O	0	0
			86	58	11	17		
2	U	10	Total	C	N	O	0	0
			86	58	11	17		
2	V	10	Total	C	N	O	0	0
			86	58	11	17		
2	W	10	Total	C	N	O	0	0
			86	58	11	17		

- Molecule 3 is a protein called HEXON PROTEIN.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	F	884	Total	C	N	O	S	0	0
			6942	4436	1172	1299	35		
3	G	884	Total	C	N	O	S	0	0
			6942	4436	1172	1299	35		
3	H	884	Total	C	N	O	S	0	0
			6942	4436	1172	1299	35		

*Continued on next page...*

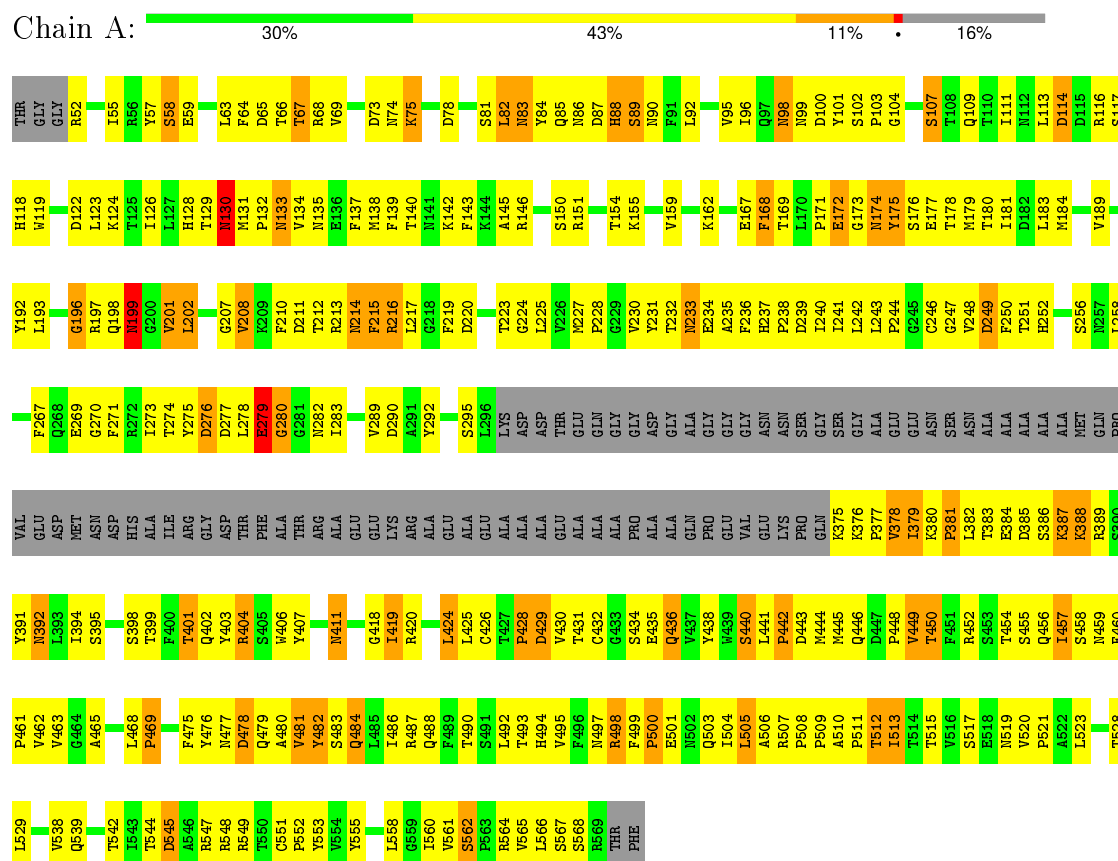
*Continued from previous page...*

Mol	Chain	Residues	Atoms					AltConf	Trace
3	I	884	Total	C	N	O	S	0	0
			6942	4436	1172	1299	35		
3	J	884	Total	C	N	O	S	0	0
			6942	4436	1172	1299	35		
3	K	884	Total	C	N	O	S	0	0
			6942	4436	1172	1299	35		
3	L	884	Total	C	N	O	S	0	0
			6942	4436	1172	1299	35		
3	M	884	Total	C	N	O	S	0	0
			6942	4436	1172	1299	35		
3	N	884	Total	C	N	O	S	0	0
			6942	4436	1172	1299	35		
3	O	884	Total	C	N	O	S	0	0
			6942	4436	1172	1299	35		
3	P	884	Total	C	N	O	S	0	0
			6942	4436	1172	1299	35		
3	Q	884	Total	C	N	O	S	0	0
			6942	4436	1172	1299	35		

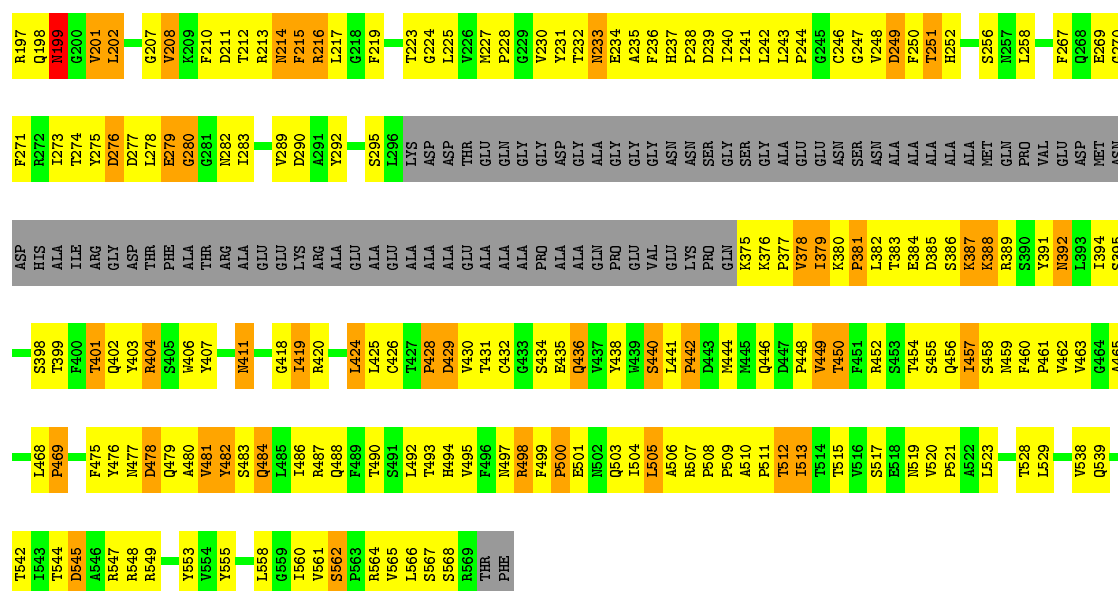
### 3 Residue-property plots

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.

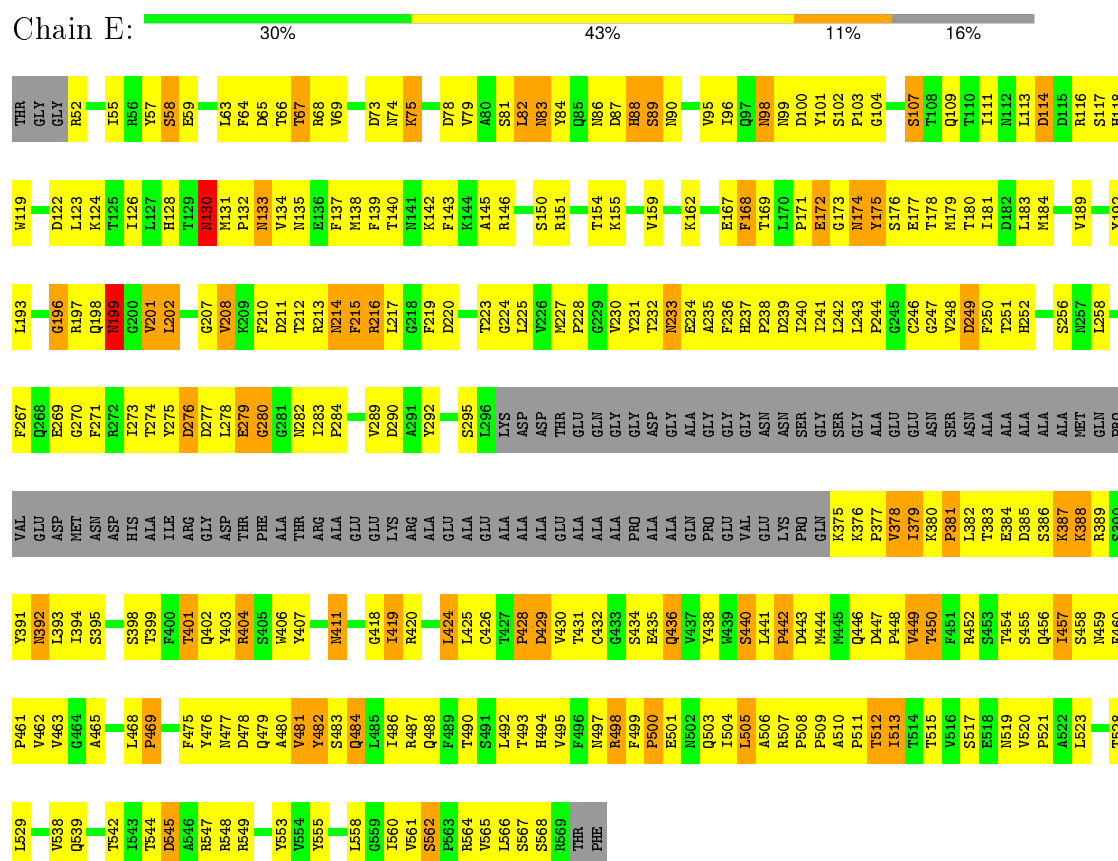
#### • Molecule 1: PENTON PROTEIN



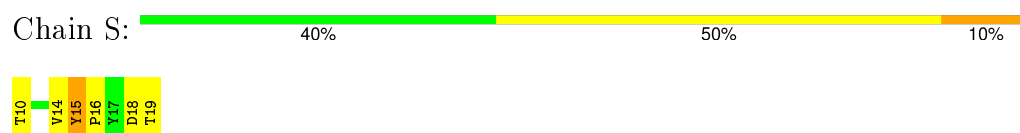




- Molecule 1: PENTON PROTEIN



- Molecule 2: N-TERMINAL PEPTIDE OF FIBER PROTEIN



- Molecule 2: N-TERMINAL PEPTIDE OF FIBER PROTEIN

Chain T: 



- Molecule 2: N-TERMINAL PEPTIDE OF FIBER PROTEIN

Chain U: 



- Molecule 2: N-TERMINAL PEPTIDE OF FIBER PROTEIN

Chain V: 



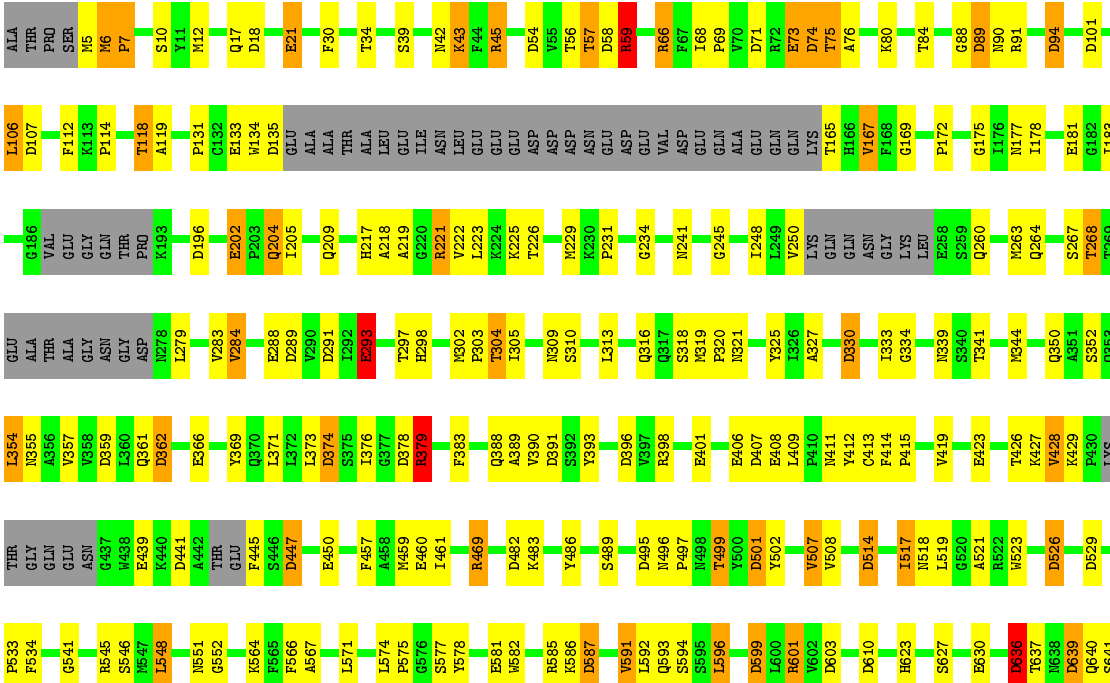
- Molecule 2: N-TERMINAL PEPTIDE OF FIBER PROTEIN

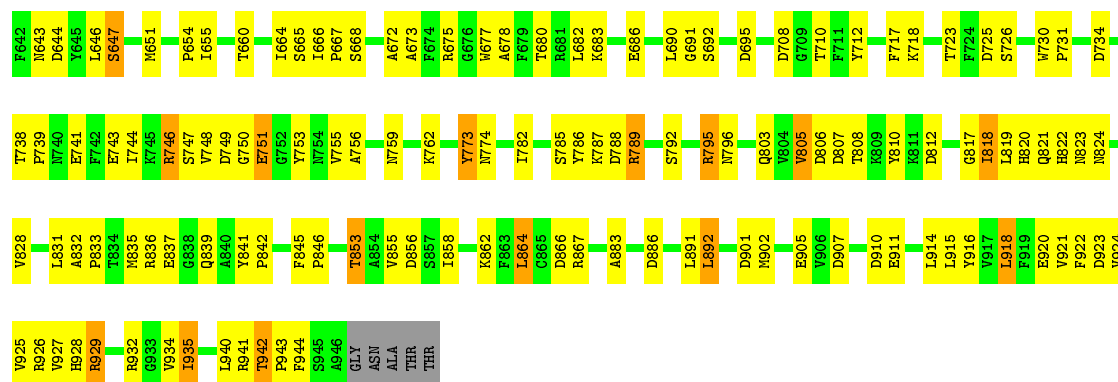
Chain W: 



- Molecule 3: HEXON PROTEIN

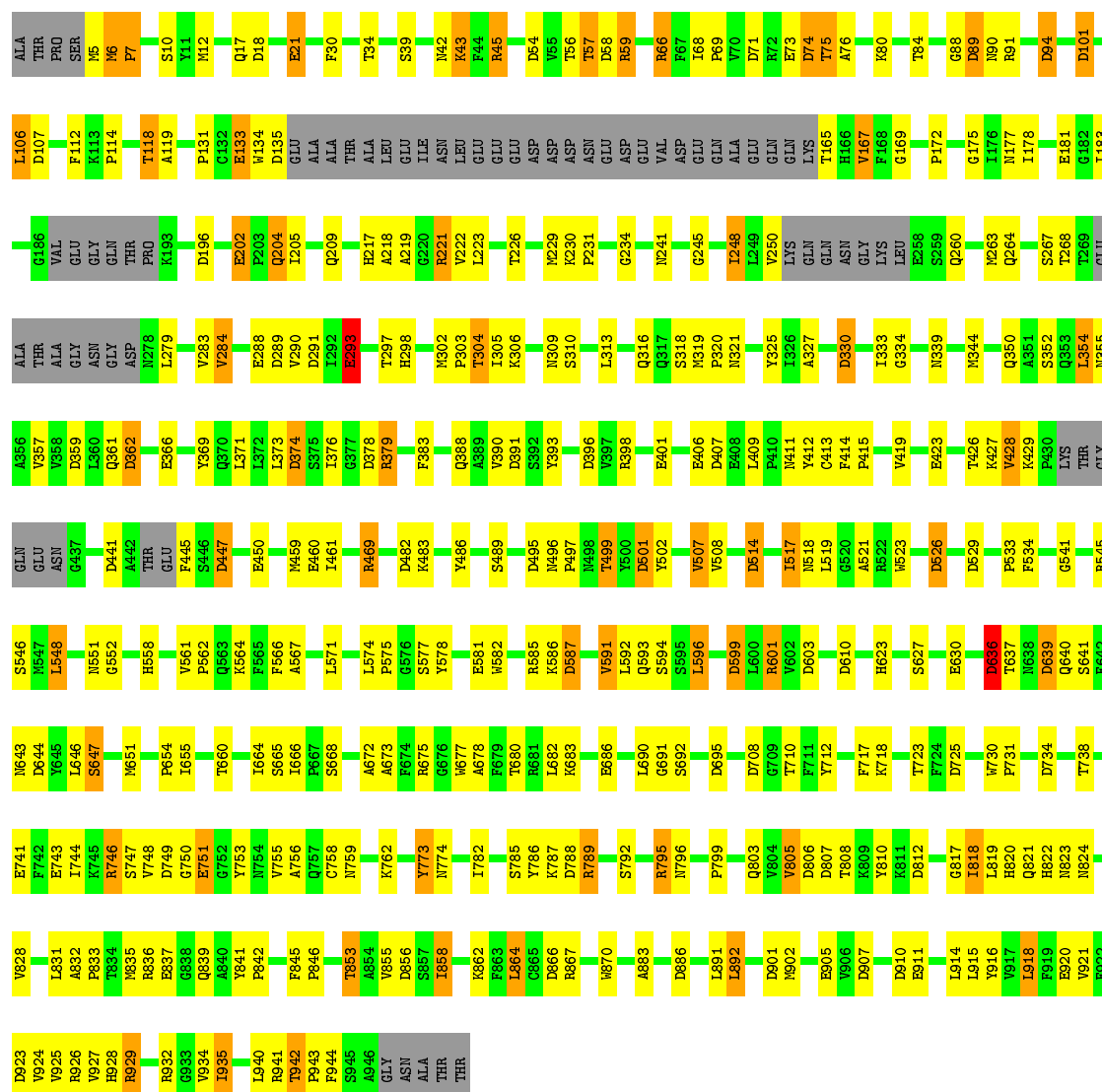
Chain F: 





### • Molecule 3: HEXON PROTEIN

Chain G: 57% 30% 6% 7%



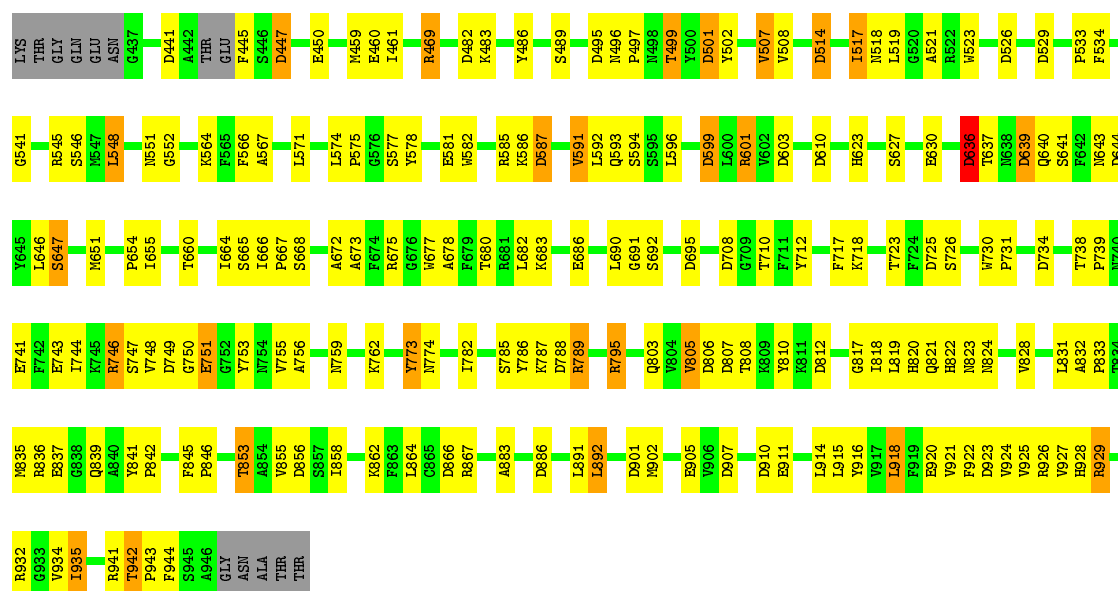
### • Molecule 3: HEXON PROTEIN

Chain H: 57% 30% 6% 7%

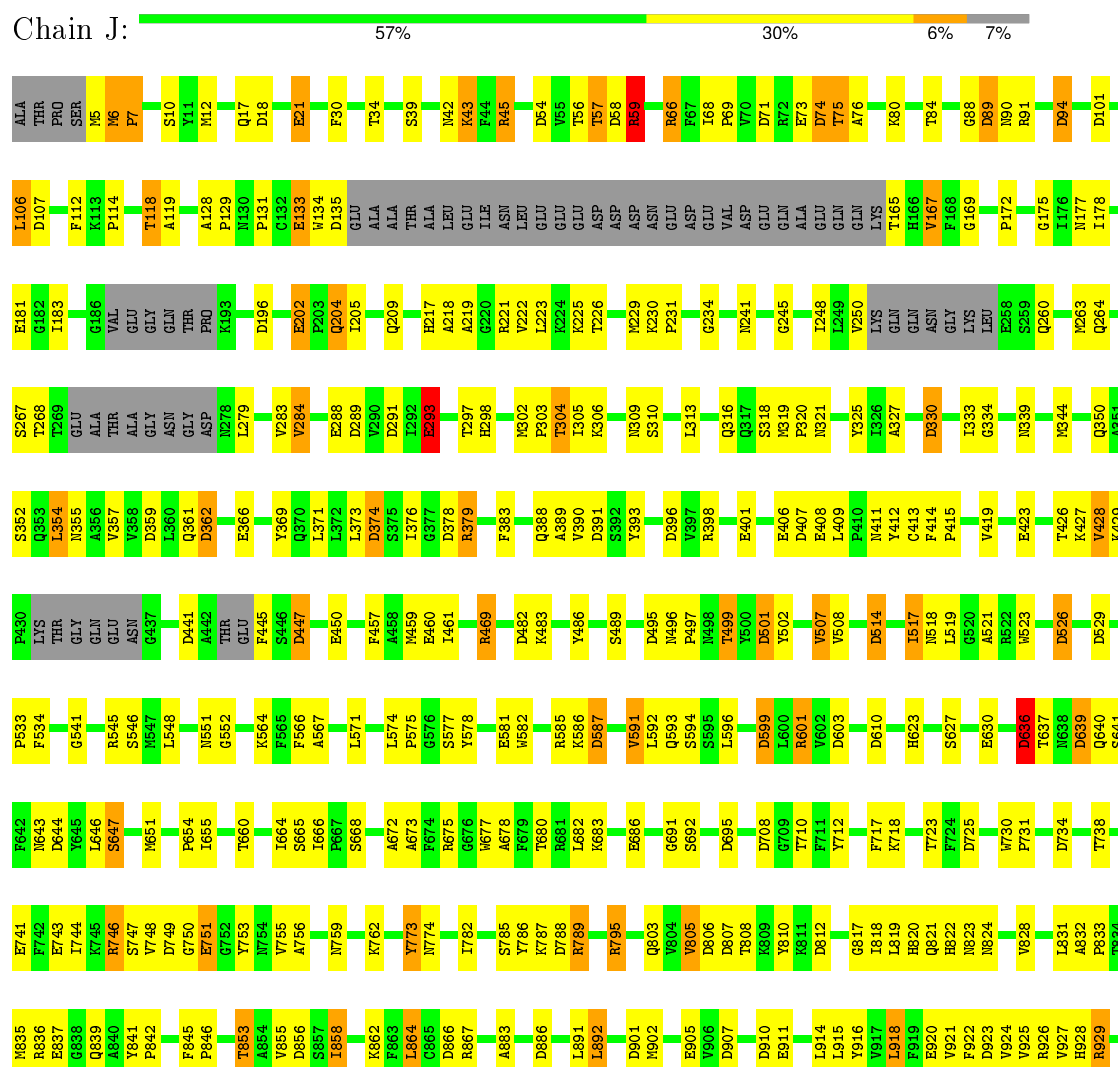


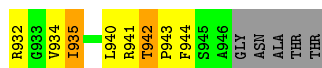
Response	Percentage
Yes, the U.S. should take action to address climate change	57%
No, the U.S. should focus on other issues	30%
Not sure	5%
Other	7%





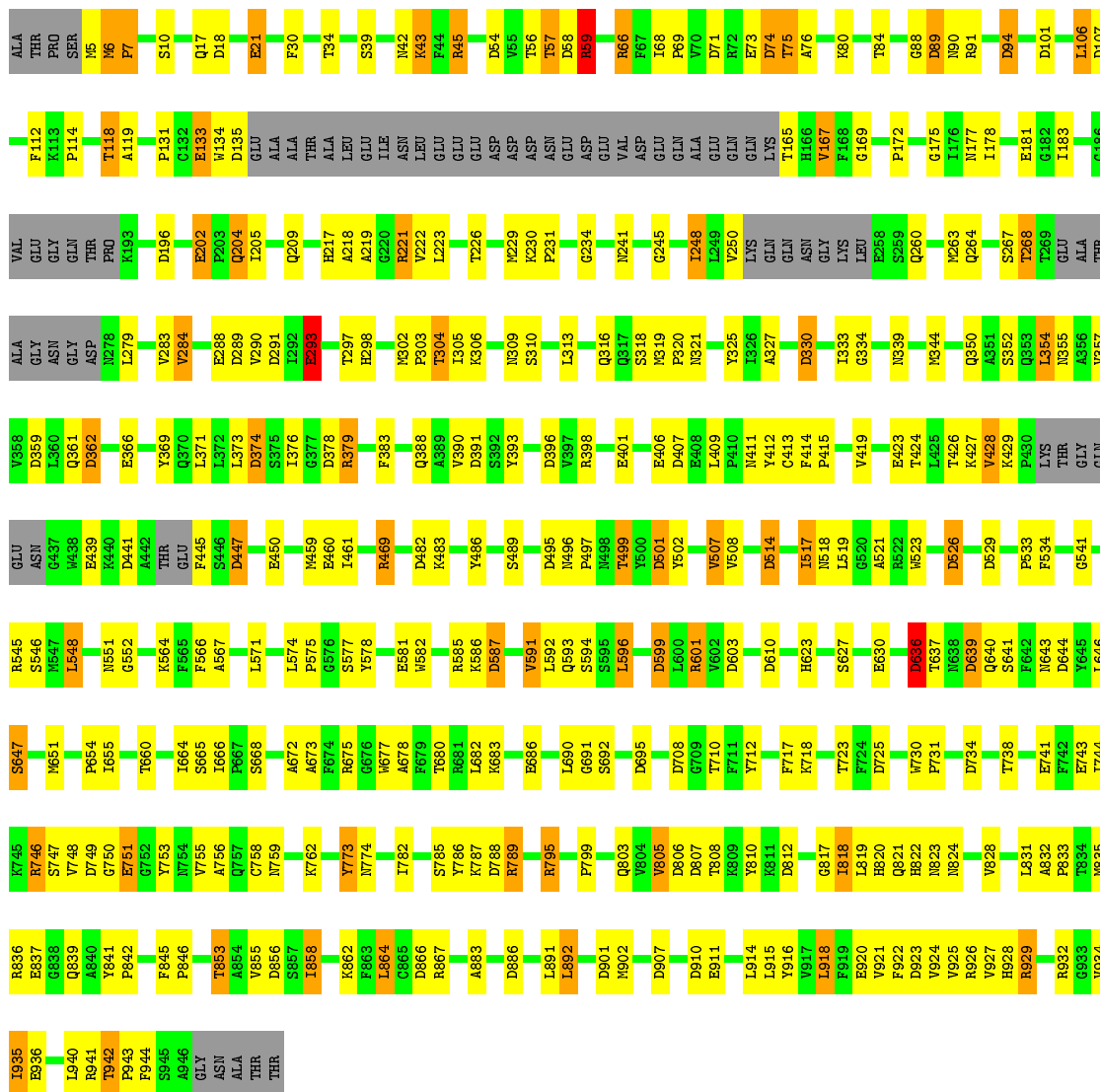
# Molecule 3: HEXON PROTEIN





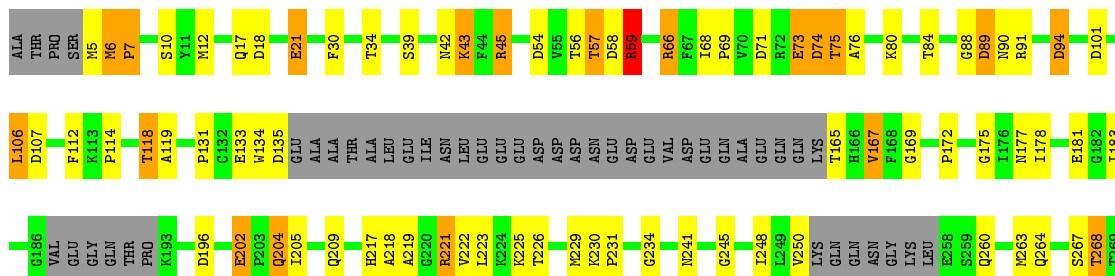
### • Molecule 3: HEXON PROTEIN

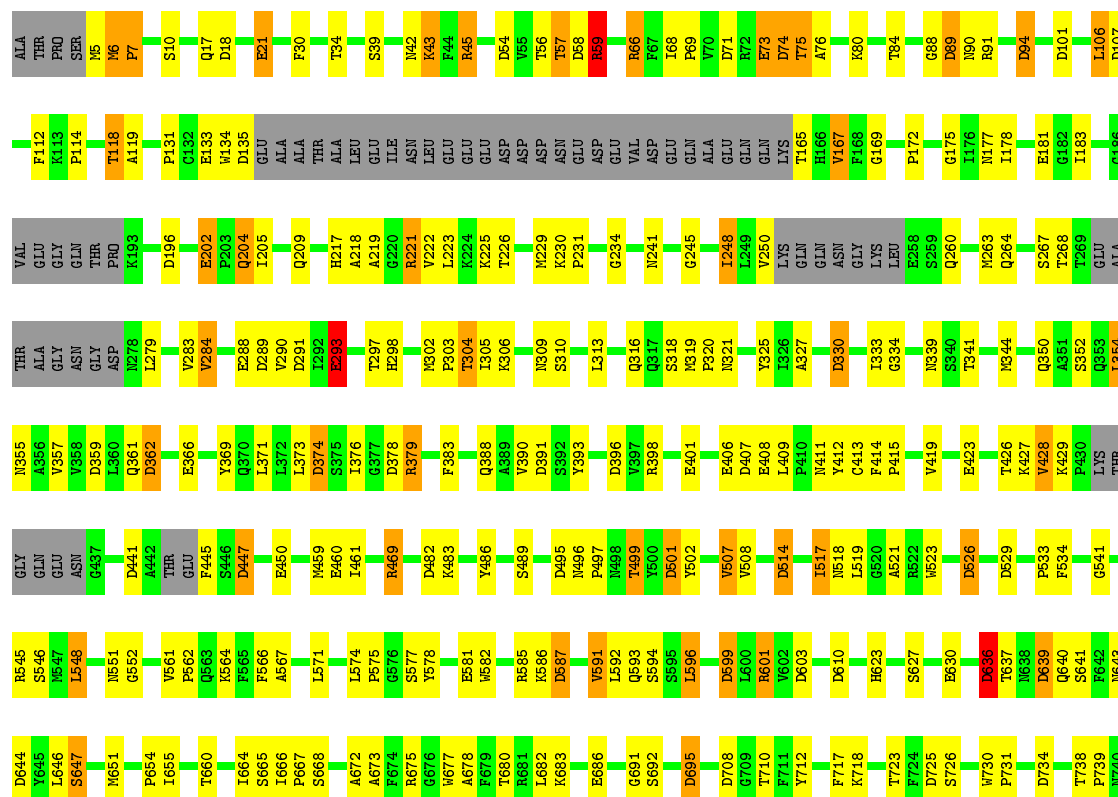
Chain K: 57% 29% 6% 7%

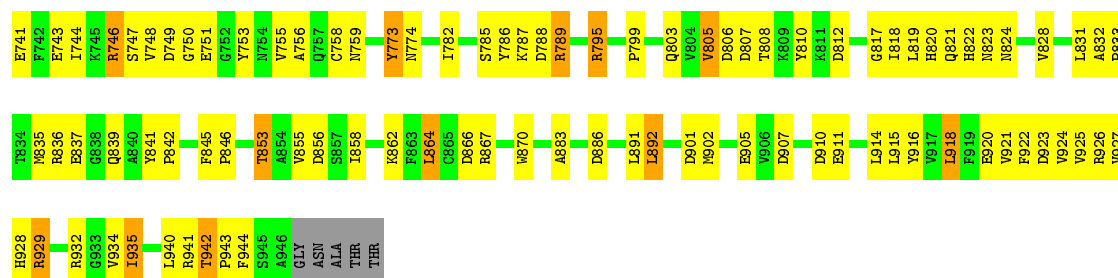


### • Molecule 3: HEXON PROTEIN

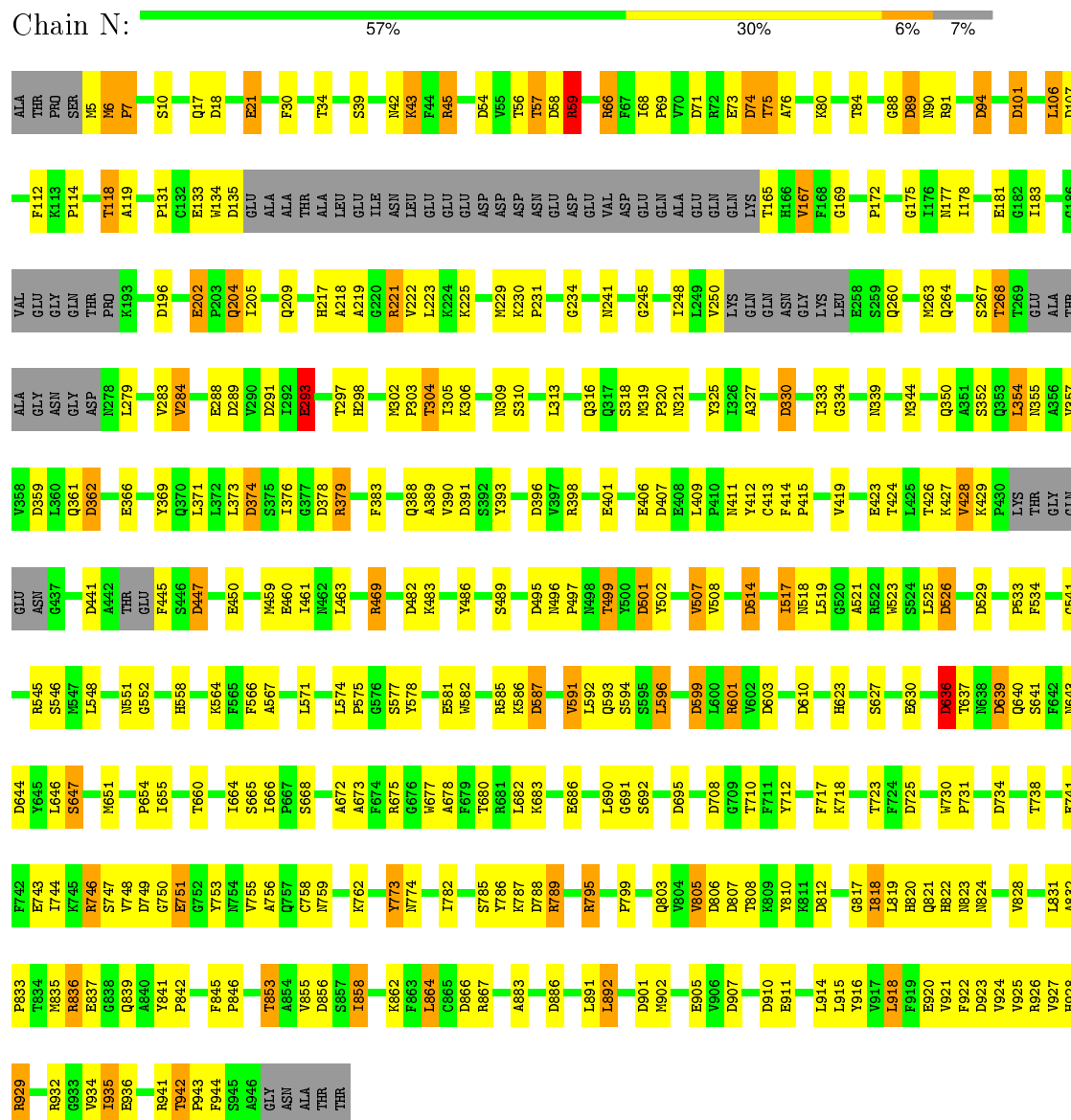
Chain L: 56% 30% 6% 7%



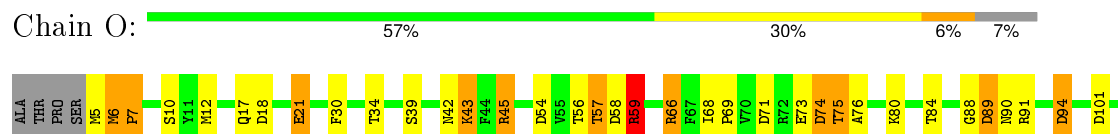




### • Molecule 3: HEXON PROTEIN



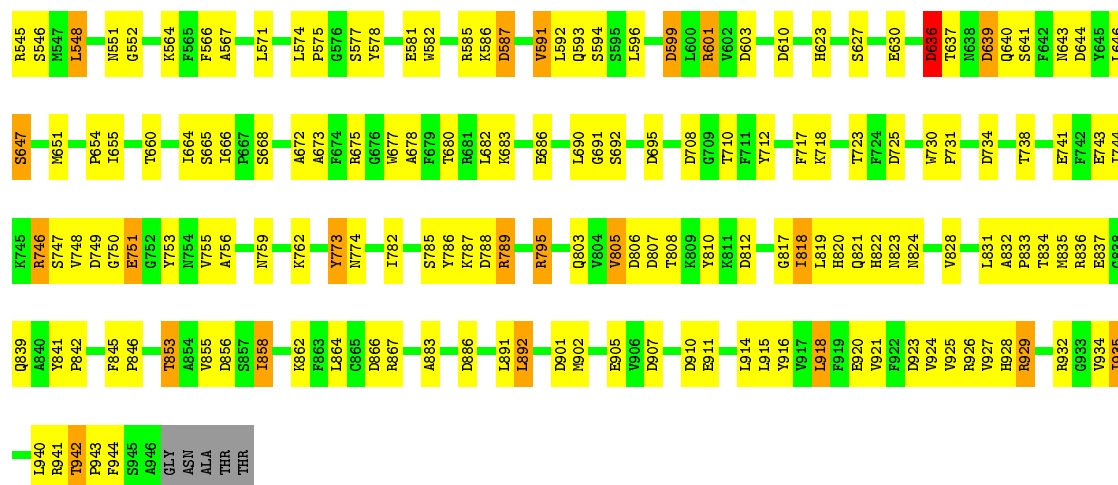
### • Molecule 3: HEXON PROTEIN





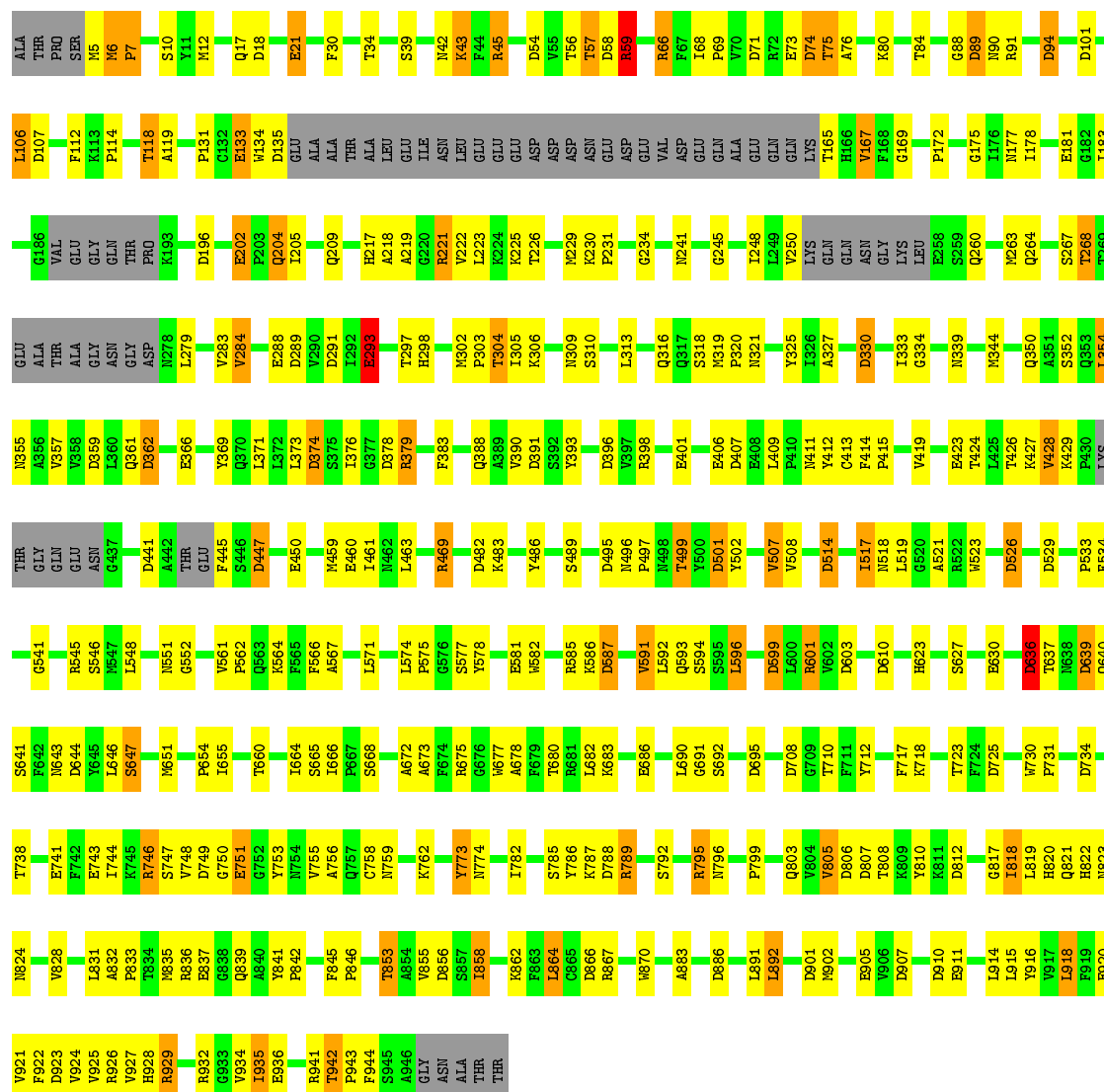
Situation	Percentage
Very good	57%
Good	30%
Not good	6%
Very bad	7%





### • Molecule 3: HEXON PROTEIN

Chain Q: 56% 30% 6% 7%



## 4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, I	Depositor
Number of images	Not provided	Depositor
Resolution determination method	Not provided	Depositor
CTF correction method	AMPLITUDE, PHASE TMV AND COMPARISON WITH X-RAY DATA	Depositor
Microscope	FEI/PHILIPS CM200	Depositor
Voltage (kV)	200	Depositor
Electron dose ( $e^-/\text{\AA}^2$ )	Not provided	Depositor
Minimum defocus (nm)	100	Depositor
Maximum defocus (nm)	250	Depositor
Magnification	27500	Depositor
Image detector	KODAK SO 163 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	0.61	0/3602	0.81	1/4904 (0.0%)
1	B	0.61	0/3602	0.81	1/4904 (0.0%)
1	C	0.61	0/3602	0.81	1/4904 (0.0%)
1	D	0.61	0/3602	0.81	1/4904 (0.0%)
1	E	0.61	0/3602	0.81	1/4904 (0.0%)
2	S	0.66	0/90	0.69	0/125
2	T	0.66	0/90	0.68	0/125
2	U	0.66	0/90	0.68	0/125
2	V	0.67	0/90	0.69	0/125
2	W	0.66	0/90	0.69	0/125
3	F	0.66	22/7133 (0.3%)	1.05	100/9725 (1.0%)
3	G	0.66	22/7133 (0.3%)	1.05	96/9725 (1.0%)
3	H	0.66	23/7133 (0.3%)	1.05	99/9725 (1.0%)
3	I	0.66	21/7133 (0.3%)	1.05	100/9725 (1.0%)
3	J	0.66	23/7133 (0.3%)	1.05	97/9725 (1.0%)
3	K	0.66	22/7133 (0.3%)	1.05	99/9725 (1.0%)
3	L	0.66	21/7133 (0.3%)	1.05	100/9725 (1.0%)
3	M	0.66	22/7133 (0.3%)	1.05	98/9725 (1.0%)
3	N	0.66	22/7133 (0.3%)	1.05	99/9725 (1.0%)
3	O	0.66	21/7133 (0.3%)	1.05	100/9725 (1.0%)
3	P	0.66	22/7133 (0.3%)	1.05	96/9725 (1.0%)
3	Q	0.66	23/7133 (0.3%)	1.05	99/9725 (1.0%)
All	All	0.65	264/104056 (0.3%)	1.01	1188/141845 (0.8%)

The worst 5 of 264 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	G	423	GLU	CD-OE1	5.79	1.32	1.25
3	H	911	GLU	CD-OE1	5.74	1.31	1.25
3	Q	911	GLU	CD-OE1	5.71	1.31	1.25
3	N	911	GLU	CD-OE1	5.71	1.31	1.25
3	P	423	GLU	CD-OE1	5.71	1.31	1.25

The worst 5 of 1188 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	482	TYR	N-CA-C	-7.27	91.36	111.00
1	A	482	TYR	N-CA-C	-7.27	91.37	111.00
1	B	482	TYR	N-CA-C	-7.26	91.40	111.00
1	D	482	TYR	N-CA-C	-7.26	91.40	111.00
1	E	482	TYR	N-CA-C	-7.25	91.42	111.00

There are no chirality outliers.

There are no planarity outliers.

## 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	A	3519	0	3451	436	0
1	B	3519	0	3452	415	0
1	C	3519	0	3452	416	0
1	D	3519	0	3452	419	0
1	E	3519	0	3452	422	0
2	S	86	0	73	19	0
2	T	86	0	73	17	0
2	U	86	0	73	17	0
2	V	86	0	73	16	0
2	W	86	0	73	19	0
3	F	6942	0	6506	267	0
3	G	6942	0	6506	248	0
3	H	6942	0	6506	250	0
3	I	6942	0	6506	260	0
3	J	6942	0	6506	238	0
3	K	6942	0	6506	251	0
3	L	6942	0	6506	277	0
3	M	6942	0	6506	263	0
3	N	6942	0	6506	247	0
3	O	6942	0	6506	260	0
3	P	6942	0	6504	249	0
3	Q	6942	0	6506	253	0
All	All	101329	0	95694	4683	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 24.

The worst 5 of 4683 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)
3:F:341:THR:CG2	3:M:739:PRO:HG2	1.55	1.36
3:L:341:THR:CG2	3:O:739:PRO:HG2	1.62	1.27
3:F:739:PRO:HG2	3:I:341:THR:CG2	1.63	1.26
3:L:73:GLU:CG	3:Q:68:ILE:HG21	1.67	1.23
3:L:341:THR:HG21	3:O:739:PRO:CG	1.70	1.19

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	436/523 (83%)	321 (74%)	69 (16%)	46 (11%)	0	11
1	B	436/523 (83%)	321 (74%)	69 (16%)	46 (11%)	0	11
1	C	436/523 (83%)	321 (74%)	69 (16%)	46 (11%)	0	11
1	D	436/523 (83%)	321 (74%)	69 (16%)	46 (11%)	0	11
1	E	436/523 (83%)	321 (74%)	69 (16%)	46 (11%)	0	11
2	S	8/10 (80%)	5 (62%)	3 (38%)	0	100	100
2	T	8/10 (80%)	5 (62%)	3 (38%)	0	100	100
2	U	8/10 (80%)	5 (62%)	3 (38%)	0	100	100
2	V	8/10 (80%)	5 (62%)	3 (38%)	0	100	100
2	W	8/10 (80%)	5 (62%)	3 (38%)	0	100	100
3	F	870/951 (92%)	808 (93%)	56 (6%)	6 (1%)	26	71
3	G	870/951 (92%)	807 (93%)	57 (7%)	6 (1%)	26	71

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
3	H	870/951 (92%)	808 (93%)	56 (6%)	6 (1%)	26	71
3	I	870/951 (92%)	808 (93%)	56 (6%)	6 (1%)	26	71
3	J	870/951 (92%)	806 (93%)	58 (7%)	6 (1%)	26	71
3	K	870/951 (92%)	808 (93%)	56 (6%)	6 (1%)	26	71
3	L	870/951 (92%)	808 (93%)	56 (6%)	6 (1%)	26	71
3	M	870/951 (92%)	807 (93%)	57 (7%)	6 (1%)	26	71
3	N	870/951 (92%)	808 (93%)	56 (6%)	6 (1%)	26	71
3	O	870/951 (92%)	808 (93%)	56 (6%)	6 (1%)	26	71
3	P	870/951 (92%)	807 (93%)	57 (7%)	6 (1%)	26	71
3	Q	870/951 (92%)	808 (93%)	56 (6%)	6 (1%)	26	71
All	All	12660/14077 (90%)	11321 (89%)	1037 (8%)	302 (2%)	12	47

5 of 302 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	82	LEU
1	A	130	ASN
1	A	172	GLU
1	A	174	ASN
1	A	201	VAL

### 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 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	399/451 (88%)	369 (92%)	30 (8%)	17	53
1	B	399/451 (88%)	370 (93%)	29 (7%)	17	54
1	C	399/451 (88%)	369 (92%)	30 (8%)	17	53
1	D	399/451 (88%)	370 (93%)	29 (7%)	17	54
1	E	399/451 (88%)	370 (93%)	29 (7%)	17	54
2	S	10/10 (100%)	9 (90%)	1 (10%)	9	38

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	T	10/10 (100%)	9 (90%)	1 (10%)	9	38
2	U	10/10 (100%)	9 (90%)	1 (10%)	9	38
2	V	10/10 (100%)	9 (90%)	1 (10%)	9	38
2	W	10/10 (100%)	9 (90%)	1 (10%)	9	38
3	F	727/828 (88%)	647 (89%)	80 (11%)	8	34
3	G	727/828 (88%)	649 (89%)	78 (11%)	8	36
3	H	727/828 (88%)	647 (89%)	80 (11%)	8	34
3	I	727/828 (88%)	647 (89%)	80 (11%)	8	34
3	J	727/828 (88%)	648 (89%)	79 (11%)	8	35
3	K	727/828 (88%)	647 (89%)	80 (11%)	8	34
3	L	727/828 (88%)	647 (89%)	80 (11%)	8	34
3	M	727/828 (88%)	648 (89%)	79 (11%)	8	35
3	N	727/828 (88%)	647 (89%)	80 (11%)	8	34
3	O	727/828 (88%)	647 (89%)	80 (11%)	8	34
3	P	727/828 (88%)	648 (89%)	79 (11%)	8	35
3	Q	727/828 (88%)	646 (89%)	81 (11%)	8	34
All	All	10769/12241 (88%)	9661 (90%)	1108 (10%)	13	37

5 of 1108 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	J	469	ARG
3	K	864	LEU
3	P	929	ARG
3	J	636	ASP
3	K	204	GLN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 177 such sidechains are listed below:

Mol	Chain	Res	Type
3	F	217	HIS
3	H	350	GLN
3	P	217	HIS
3	F	350	GLN
3	G	217	HIS

### 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 carbohydrates 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 ⓘ

There are no chain breaks in this entry.