



## wwPDB/EMDataBank EM Map/Model Validation Summary Report ⓘ

Sep 17, 2017 – 10:51 PM EDT

PDB ID : 5THR  
EMDB ID: : EMD-8407  
Title : Cryo-EM structure of a BG505 Env-sCD4-17b-8ANC195 complex  
Authors : Wang, H.; Bjorkman, P.J.  
Deposited on : unknown  
Resolution : 8.90 Å(reported)

This is a wwPDB/EMDataBank EM Map/Model Validation Summary Report  
for a publicly released PDB/EMDB entry.

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>  
with specific help available everywhere you see the ⓘ symbol.

---

MolProbity : 4.02b-467  
Mogul : 1.7.2 (RC1), CSD as538be (2017)  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)  
Validation Pipeline (wwPDB-VP) : rb-20029824

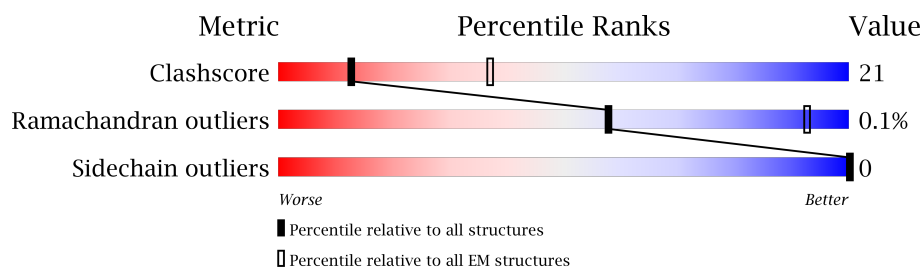
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*ELECTRON MICROSCOPY*

The reported resolution of this entry is 8.90 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	125131	1336
Ramachandran outliers	121729	1120
Sidechain outliers	121581	1026

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	153	
1	B	153	
1	C	153	
2	D	481	
2	E	481	
2	F	481	
3	G	192	
3	H	192	
3	I	192	

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
4	J	214	
4	K	214	
4	L	214	
5	M	229	
5	N	229	
5	O	229	
6	P	244	
6	R	244	
6	T	244	
7	Q	215	
7	S	215	
7	U	215	

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
10	MAN	D	2343	-	-	X	-
8	NAG	D	2340	-	-	X	-
8	NAG	D	2341	-	-	X	-
9	BMA	D	2342	-	-	X	-

## 2 Entry composition [i](#)

There are 10 unique types of molecules in this entry. The entry contains 25071 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 BG505 SOSIP gp41.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	124	Total	C	N	O	S	0	0
			983	619	170	188	6		
1	B	123	Total	C	N	O	S	0	0
			975	613	169	187	6		
1	C	123	Total	C	N	O	S	0	0
			975	613	169	187	6		

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	559	PRO	ILE	conflict	UNP Q2N0S6
A	605	CYS	THR	conflict	UNP Q2N0S6
B	559	PRO	ILE	conflict	UNP Q2N0S6
B	605	CYS	THR	conflict	UNP Q2N0S6
C	559	PRO	ILE	conflict	UNP Q2N0S6
C	605	CYS	THR	conflict	UNP Q2N0S6

- Molecule 2 is a protein called BG505 SOSIP gp120.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	D	364	Total	C	N	O	S	0	0
			2845	1792	496	533	24		
2	E	364	Total	C	N	O	S	0	0
			2845	1792	496	533	24		
2	F	364	Total	C	N	O	S	0	0
			2845	1792	496	533	24		

There are 21 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	332	ASN	THR	conflict	UNP Q2N0S6
D	501	CYS	ALA	conflict	UNP Q2N0S6
D	509	ARG	-	expression tag	UNP Q2N0S6

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
D	510	ARG	-	expression tag	UNP Q2N0S6
D	511	ARG	-	expression tag	UNP Q2N0S6
D	512	ARG	-	expression tag	UNP Q2N0S6
D	513	ARG	-	expression tag	UNP Q2N0S6
E	332	ASN	THR	conflict	UNP Q2N0S6
E	501	CYS	ALA	conflict	UNP Q2N0S6
E	509	ARG	-	expression tag	UNP Q2N0S6
E	510	ARG	-	expression tag	UNP Q2N0S6
E	511	ARG	-	expression tag	UNP Q2N0S6
E	512	ARG	-	expression tag	UNP Q2N0S6
E	513	ARG	-	expression tag	UNP Q2N0S6
F	332	ASN	THR	conflict	UNP Q2N0S6
F	501	CYS	ALA	conflict	UNP Q2N0S6
F	509	ARG	-	expression tag	UNP Q2N0S6
F	510	ARG	-	expression tag	UNP Q2N0S6
F	511	ARG	-	expression tag	UNP Q2N0S6
F	512	ARG	-	expression tag	UNP Q2N0S6
F	513	ARG	-	expression tag	UNP Q2N0S6

- Molecule 3 is a protein called T-cell surface glycoprotein CD4.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	G	98	Total	C	N	O	S	0	0
			786	496	137	151	2		
3	H	98	Total	C	N	O	S	0	0
			786	496	137	151	2		
3	I	97	Total	C	N	O	S	0	0
			775	487	136	150	2		

There are 30 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
G	183	ILE	-	expression tag	UNP P01730
G	184	ASP	-	expression tag	UNP P01730
G	185	GLY	-	expression tag	UNP P01730
G	186	ARG	-	expression tag	UNP P01730
G	187	HIS	-	expression tag	UNP P01730
G	188	HIS	-	expression tag	UNP P01730
G	189	HIS	-	expression tag	UNP P01730
G	190	HIS	-	expression tag	UNP P01730
G	191	HIS	-	expression tag	UNP P01730
G	192	HIS	-	expression tag	UNP P01730

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
H	183	ILE	-	expression tag	UNP P01730
H	184	ASP	-	expression tag	UNP P01730
H	185	GLY	-	expression tag	UNP P01730
H	186	ARG	-	expression tag	UNP P01730
H	187	HIS	-	expression tag	UNP P01730
H	188	HIS	-	expression tag	UNP P01730
H	189	HIS	-	expression tag	UNP P01730
H	190	HIS	-	expression tag	UNP P01730
H	191	HIS	-	expression tag	UNP P01730
H	192	HIS	-	expression tag	UNP P01730
I	183	ILE	-	expression tag	UNP P01730
I	184	ASP	-	expression tag	UNP P01730
I	185	GLY	-	expression tag	UNP P01730
I	186	ARG	-	expression tag	UNP P01730
I	187	HIS	-	expression tag	UNP P01730
I	188	HIS	-	expression tag	UNP P01730
I	189	HIS	-	expression tag	UNP P01730
I	190	HIS	-	expression tag	UNP P01730
I	191	HIS	-	expression tag	UNP P01730
I	192	HIS	-	expression tag	UNP P01730

- Molecule 4 is a protein called 17b Fab VL domain.

Mol	Chain	Residues	Atoms					AltConf	Trace
4	J	107	Total	C	N	O	S	0	0
			819	512	141	163	3		
4	K	107	Total	C	N	O	S	0	0
			819	512	141	163	3		
4	L	107	Total	C	N	O	S	0	0
			819	512	141	163	3		

- Molecule 5 is a protein called 17b Fab VH domain.

Mol	Chain	Residues	Atoms					AltConf	Trace
5	M	127	Total	C	N	O	S	0	0
			985	621	168	193	3		
5	N	127	Total	C	N	O	S	0	0
			985	621	168	193	3		
5	O	127	Total	C	N	O	S	0	0
			985	621	168	193	3		

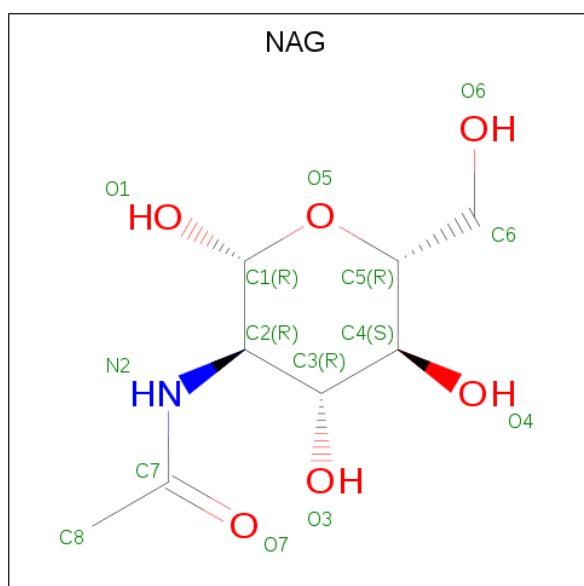
- Molecule 6 is a protein called 8ANC195 G52K5 VH domain.

Mol	Chain	Residues	Atoms					AltConf	Trace
6	P	131	Total	C	N	O	S	0	0
			980	622	167	188	3		
6	R	131	Total	C	N	O	S	0	0
			980	622	167	188	3		
6	T	131	Total	C	N	O	S	0	0
			980	622	167	188	3		

- Molecule 7 is a protein called 8ANC195 G52K5 VL domain.

Mol	Chain	Residues	Atoms					AltConf	Trace
7	Q	107	Total	C	N	O	S	0	0
			796	501	135	157	3		
7	S	107	Total	C	N	O	S	0	0
			796	501	135	157	3		
7	U	107	Total	C	N	O	S	0	0
			796	501	135	157	3		

- Molecule 8 is N-ACETYL-D-GLUCOSAMINE (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



Mol	Chain	Residues	Atoms				AltConf
8	D	1	Total	C	N	O	0
			84	48	6	30	
8	D	1	Total	C	N	O	0
			84	48	6	30	
8	D	1	Total	C	N	O	0
			84	48	6	30	
8	D	1	Total	C	N	O	0
			84	48	6	30	

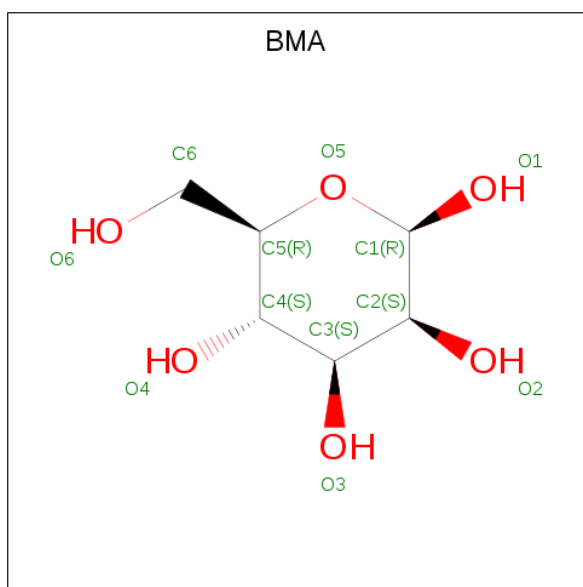
*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms				AltConf
8	D	1	Total	C	N	O	0
			84	48	6	30	
8	D	1	Total	C	N	O	0
			84	48	6	30	
8	E	1	Total	C	N	O	0
			84	48	6	30	
8	E	1	Total	C	N	O	0
			84	48	6	30	
8	E	1	Total	C	N	O	0
			84	48	6	30	
8	E	1	Total	C	N	O	0
			84	48	6	30	
8	E	1	Total	C	N	O	0
			84	48	6	30	
8	F	1	Total	C	N	O	0
			84	48	6	30	
8	F	1	Total	C	N	O	0
			84	48	6	30	
8	F	1	Total	C	N	O	0
			84	48	6	30	
8	F	1	Total	C	N	O	0
			84	48	6	30	
8	F	1	Total	C	N	O	0
			84	48	6	30	

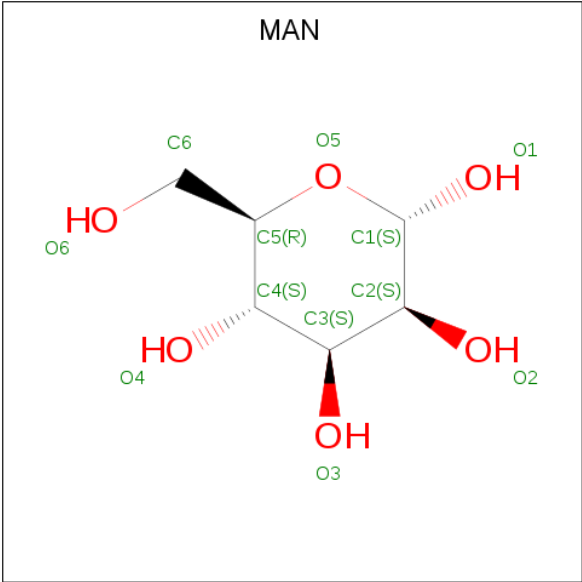
- Molecule 9 is BETA-D-MANNOSE (three-letter code: BMA) (formula: C<sub>6</sub>H<sub>12</sub>O<sub>6</sub>).





Mol	Chain	Residues	Atoms			AltConf
9	D	1	Total	C	O	0
			33	18	15	
9	D	1	Total	C	O	0
			33	18	15	
9	D	1	Total	C	O	0
			33	18	15	
9	E	1	Total	C	O	0
			33	18	15	
9	E	1	Total	C	O	0
			33	18	15	
9	E	1	Total	C	O	0
			33	18	15	
9	F	1	Total	C	O	0
			33	18	15	
9	F	1	Total	C	O	0
			33	18	15	
9	F	1	Total	C	O	0
			33	18	15	

- Molecule 10 is ALPHA-D-MANNOSE (three-letter code: MAN) (formula: C<sub>6</sub>H<sub>12</sub>O<sub>6</sub>).



Mol	Chain	Residues	Atoms			AltConf
10	D	1	Total	C	O	0
			55	30	25	
10	D	1	Total	C	O	0
			55	30	25	
10	D	1	Total	C	O	0
			55	30	25	
10	D	1	Total	C	O	0
			55	30	25	
10	D	1	Total	C	O	0
			55	30	25	
10	E	1	Total	C	O	0
			55	30	25	
10	E	1	Total	C	O	0
			55	30	25	
10	E	1	Total	C	O	0
			55	30	25	
10	E	1	Total	C	O	0
			55	30	25	
10	F	1	Total	C	O	0
			55	30	25	
10	F	1	Total	C	O	0
			55	30	25	
10	F	1	Total	C	O	0
			55	30	25	
10	F	1	Total	C	O	0
			55	30	25	

Continued on next page...

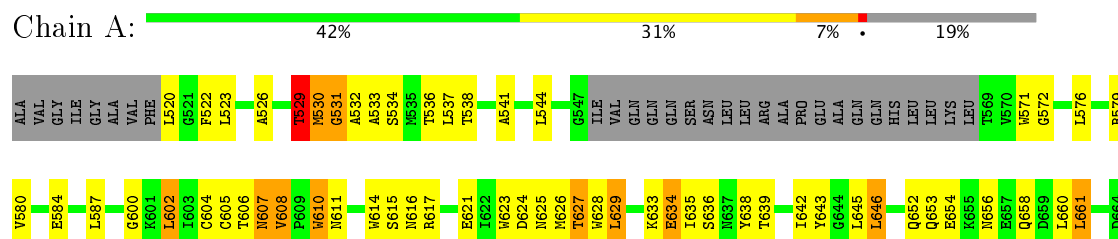
*Continued from previous page...*

Mol	Chain	Residues	Atoms			AltConf
			Total	C	O	
10	F	1	55	30	25	0

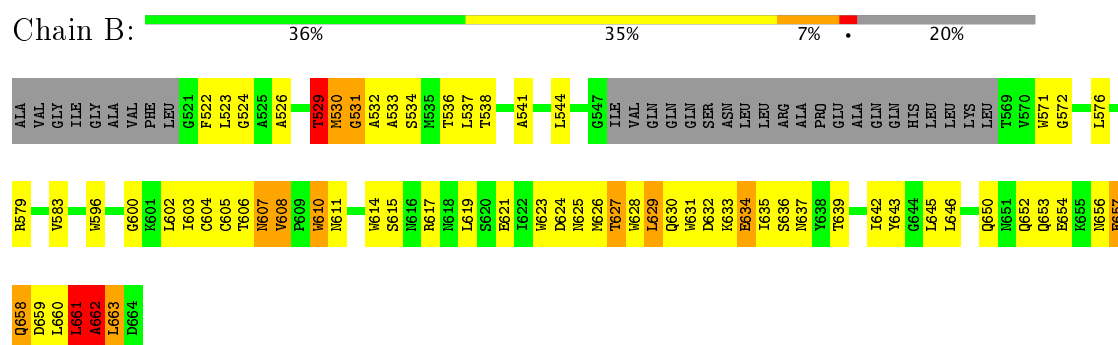
### 3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of 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.

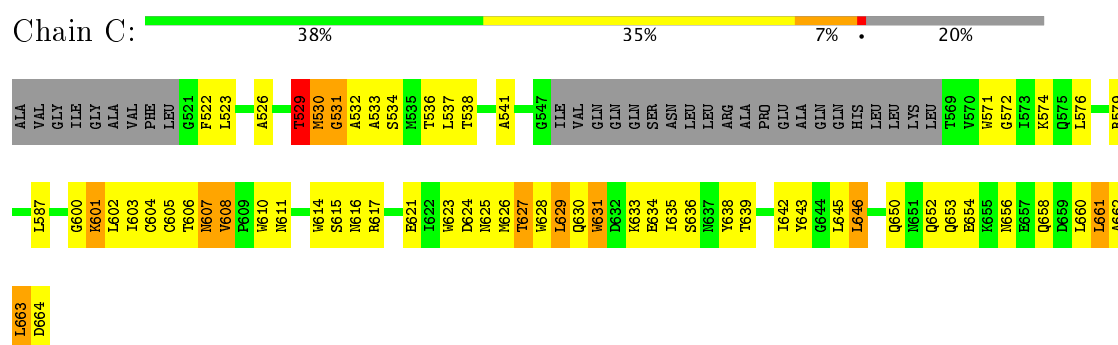
#### • Molecule 1: BG505 SOSIP gp41



#### • Molecule 1: BG505 SOSIP gp41

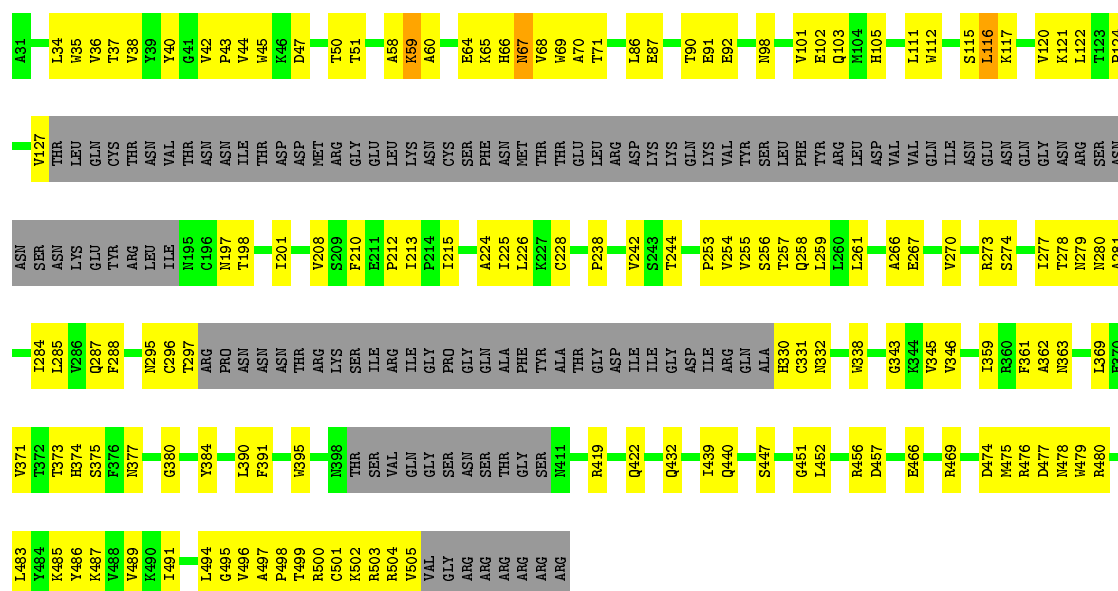


#### • Molecule 1: BG505 SOSIP gp41



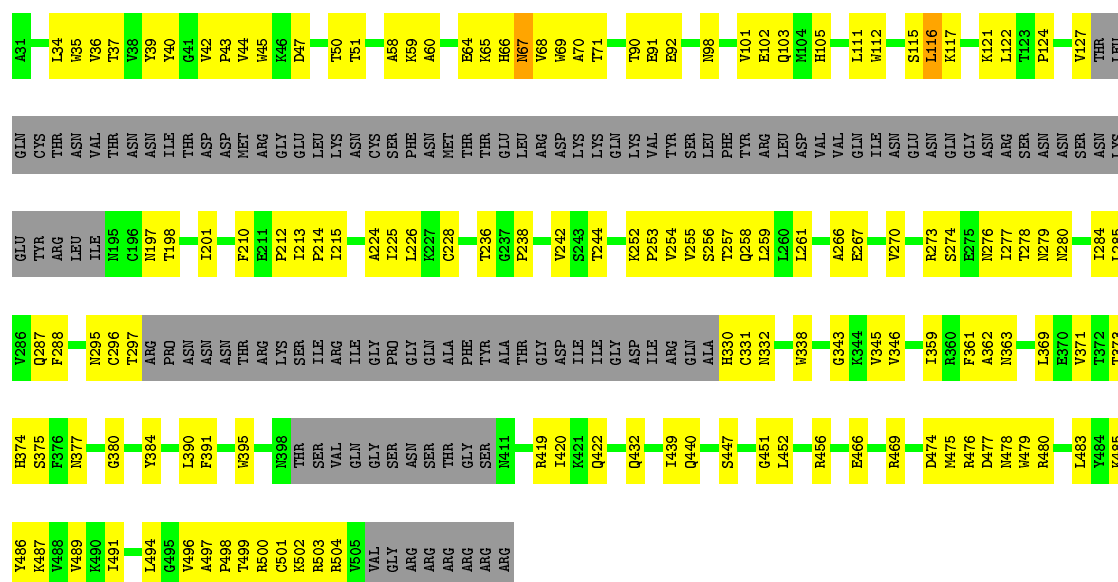
#### • Molecule 2: BG505 SOSIP gp120





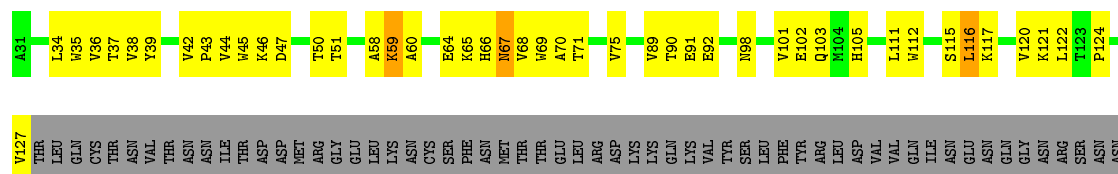
• Molecule 2: BG505 SOSIP gp120

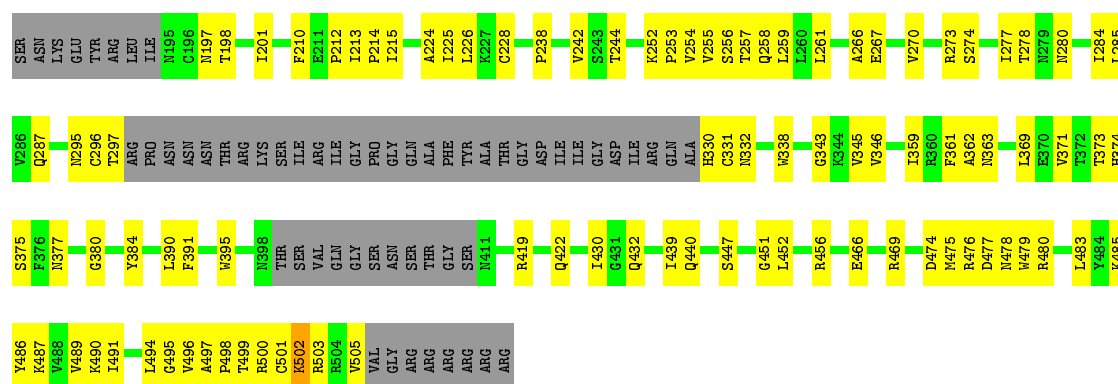
Chain E: 47% 29% 24%



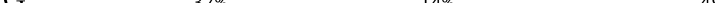
• Molecule 2: BG505 SOSIP gp120

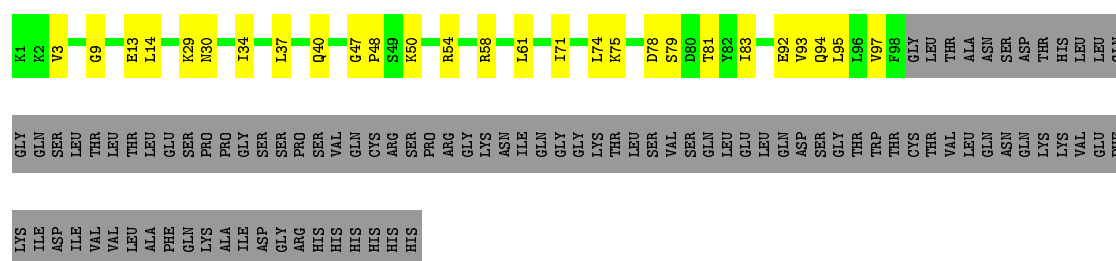
Chain F: 46% 29% 24%



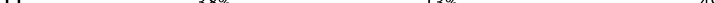


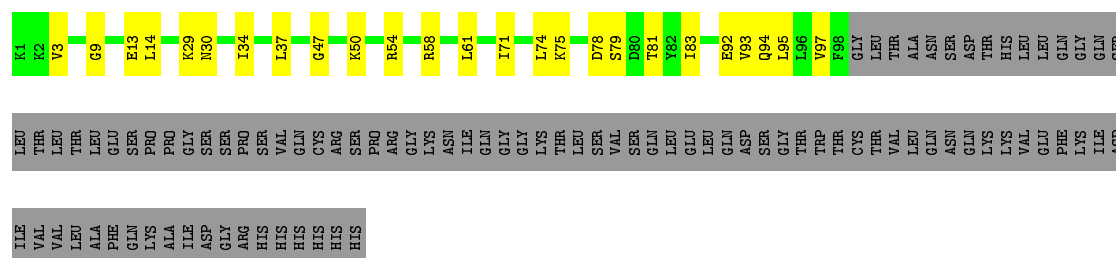
- Molecule 3: T-cell surface glycoprotein CD4

Chain G:  37% 14% 49%



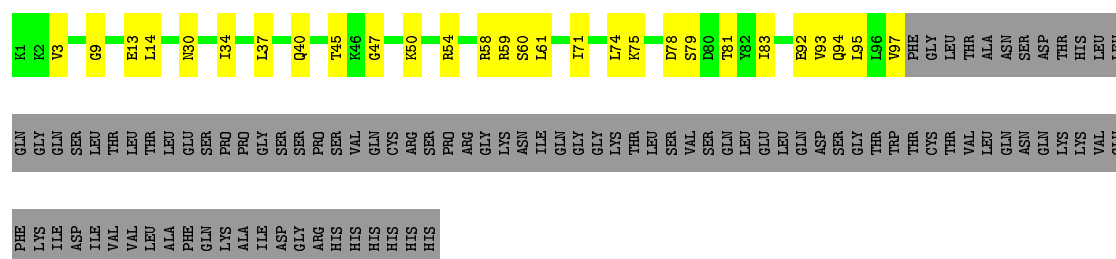
- Molecule 3: T-cell surface glycoprotein CD4

Chain H:  38% 13% 49%

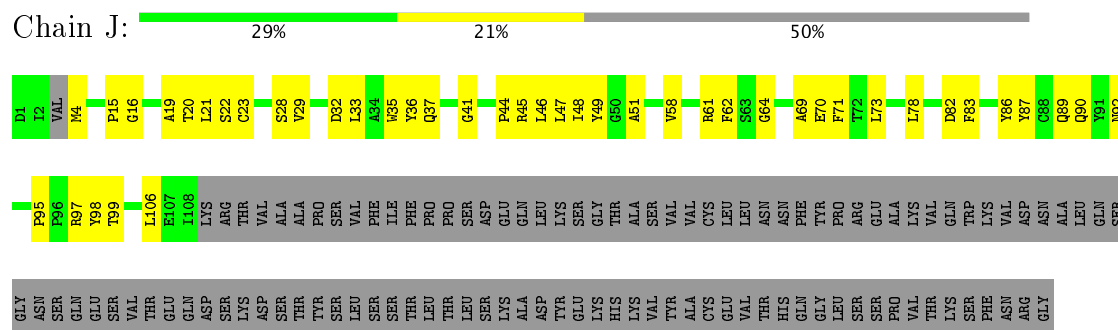


- Molecule 3: T-cell surface glycoprotein CD4

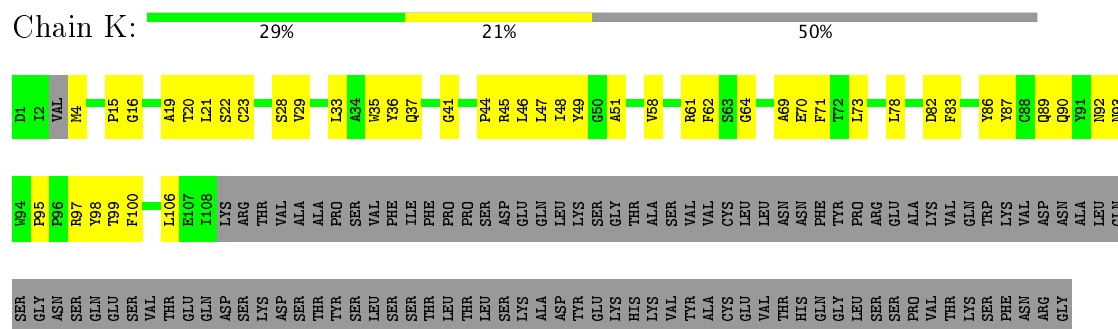
Chain I:  36% 15% 49%



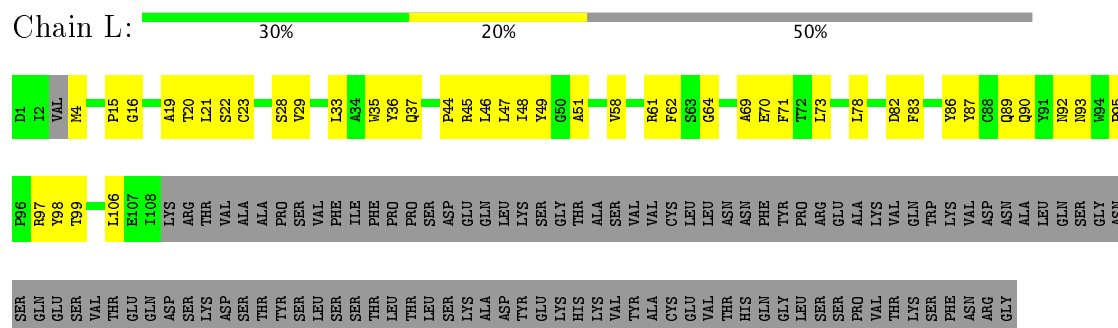
- Molecule 4: 17b Fab VL domain



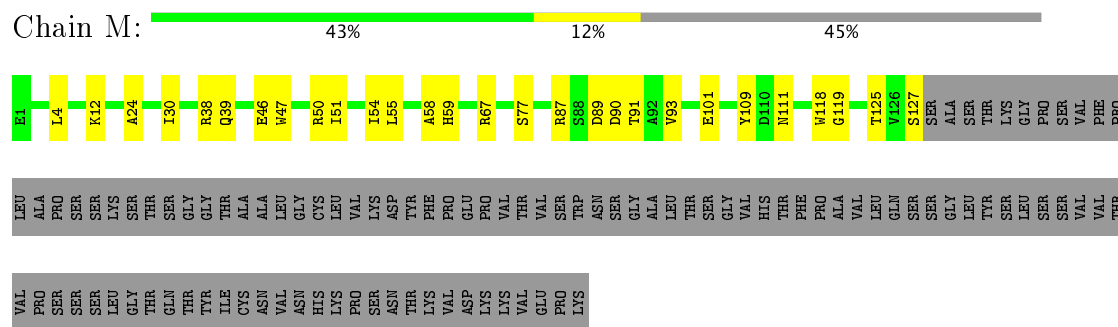
- Molecule 4: 17b Fab VL domain



- Molecule 4: 17b Fab VL domain

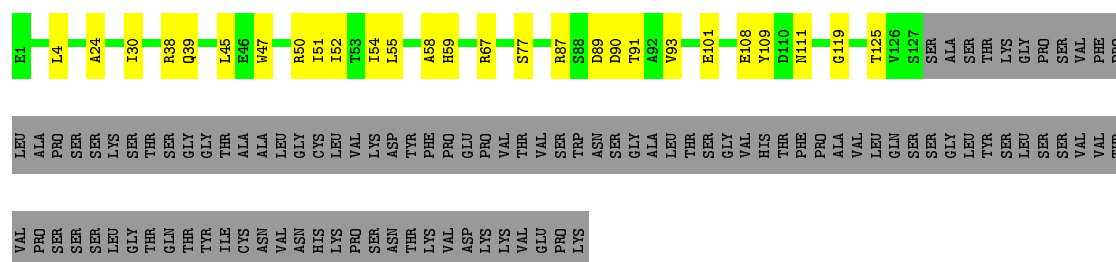


- Molecule 5: 17b Fab VH domain

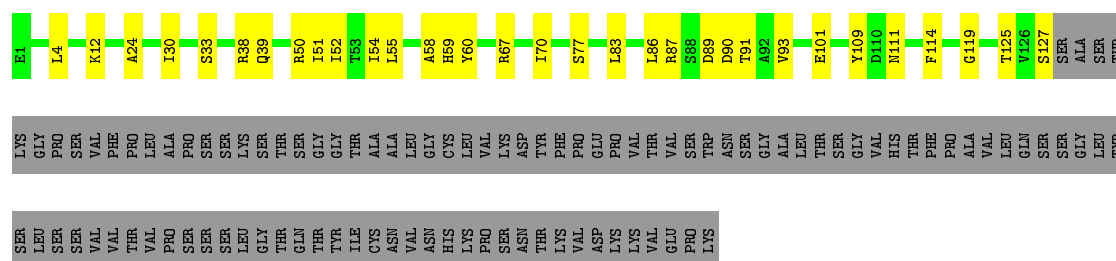


- Molecule 5: 17b Fab VH domain

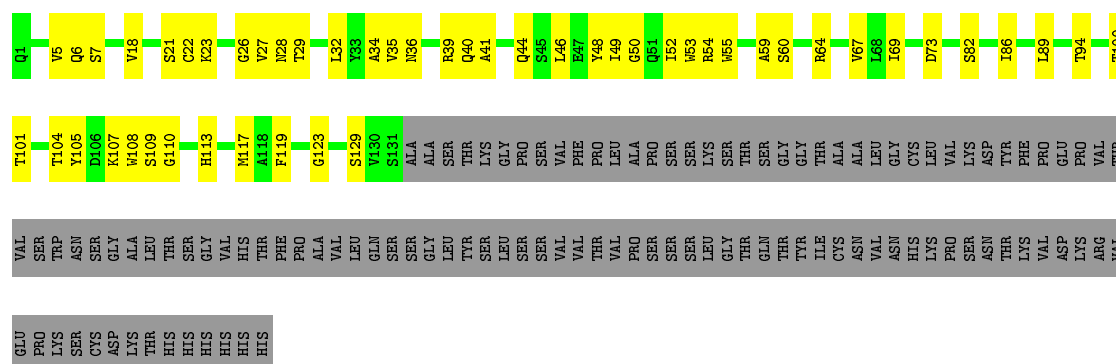
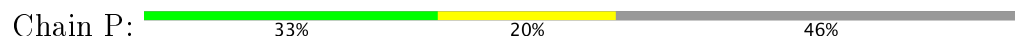




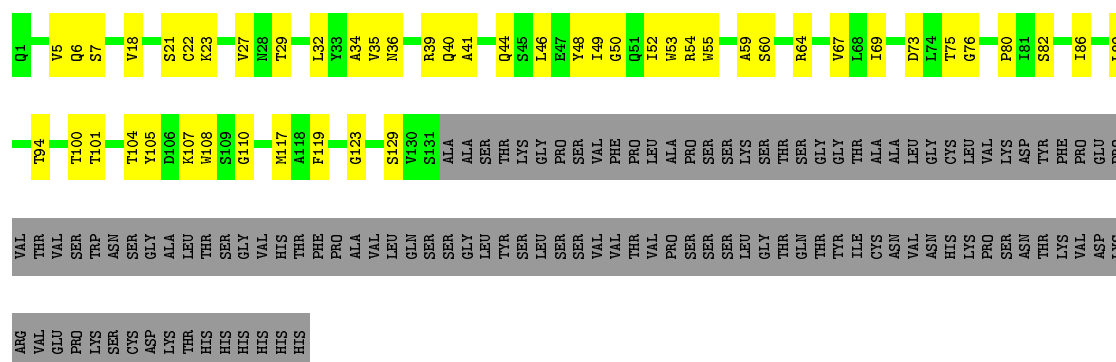
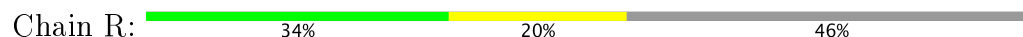
- Molecule 5: 17b Fab VH domain



- Molecule 6: 8ANC195 G52K5 VH domain

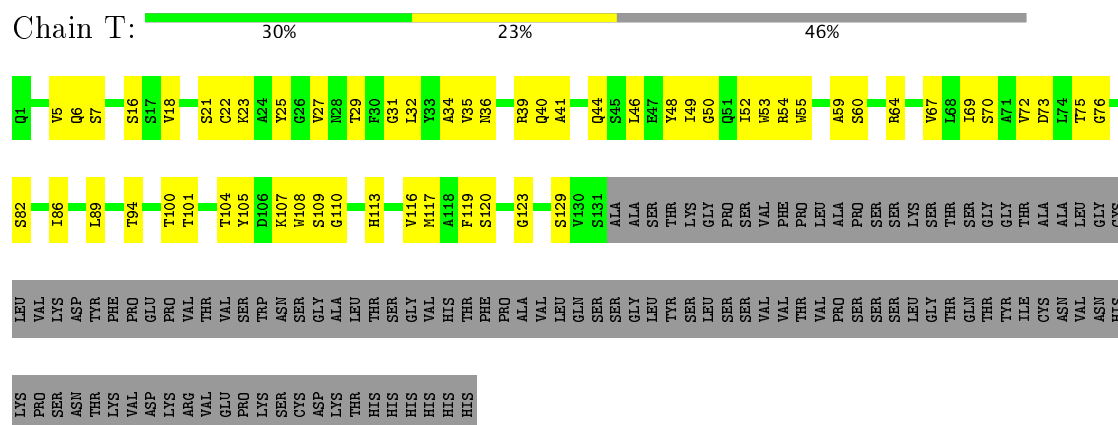


- Molecule 6: 8ANC195 G52K5 VH domain

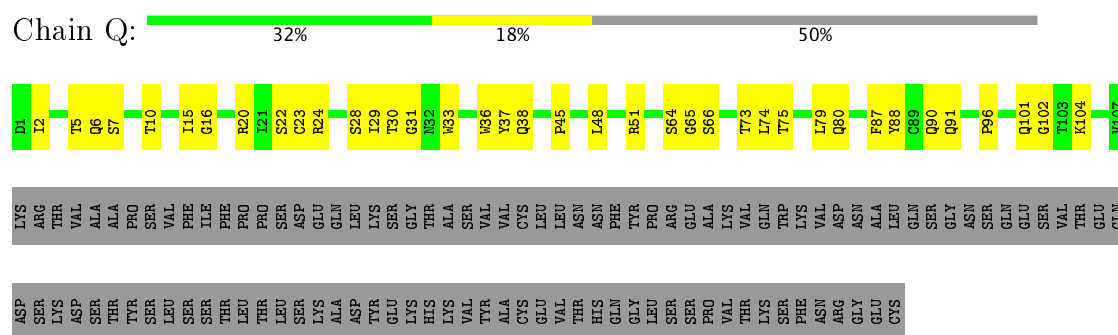




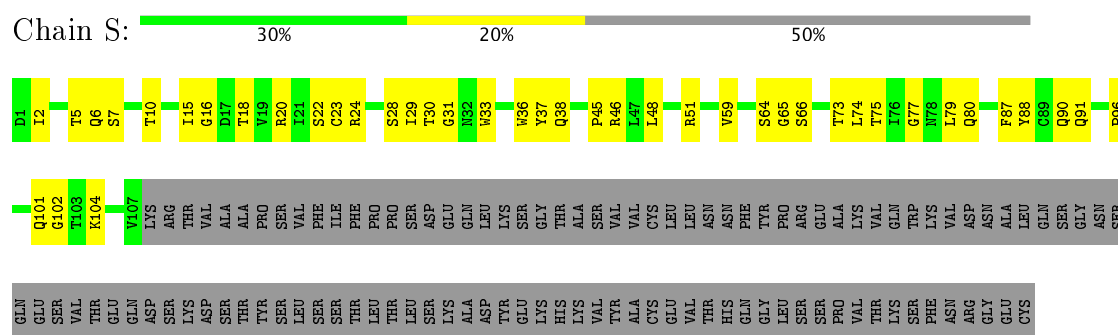
- Molecule 6: 8ANC195 G52K5 VH domain



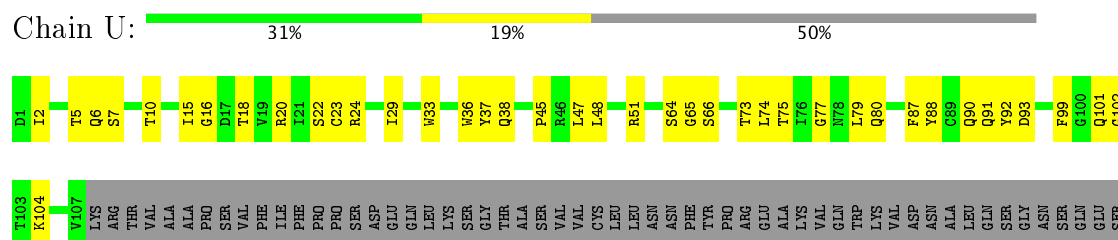
- Molecule 7: 8ANC195 G52K5 VL domain



- Molecule 7: 8ANC195 G52K5 VL domain



- Molecule 7: 8ANC195 G52K5 VL domain



VAL	THR	GLU	GLN	ASP	SER	LYS	ASP	SER	THR	TYR	SER	LEU	SER	SER	THR	THR	LEU	THR	LEU	SER	LYS	ALA	ASP	TYR	GLU	LYS	HIS	LYS	LYS	VAL	TYR	ALA	CYS	GLU	VAL	THR	HIS	GLN	GLY	LEU	SER	SER	PRO	VAL	THR	LYS	SER	PHE	ASN	ARG	GLY	GLU	CYS
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

## 4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C3	Depositor
Number of particles used	5175	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ( $e^-/\text{\AA}^2$ )	35	Depositor
Minimum defocus (nm)	Not provided	Depositor
Maximum defocus (nm)	Not provided	Depositor
Magnification	Not provided	Depositor
Image detector	GATAN K2 SUMMIT (4k x 4k)	Depositor

## 5 Model quality [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: BMA, NAG, MAN

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	0.84	3/1000 (0.3%)	1.30	19/1356 (1.4%)
1	B	1.03	8/992 (0.8%)	1.36	19/1345 (1.4%)
1	C	0.80	2/992 (0.2%)	1.31	21/1345 (1.6%)
2	D	0.45	1/2909 (0.0%)	0.76	5/3956 (0.1%)
2	E	0.45	1/2909 (0.0%)	0.76	5/3956 (0.1%)
2	F	0.50	3/2909 (0.1%)	0.78	6/3956 (0.2%)
3	G	0.29	0/797	0.61	0/1069
3	H	0.29	0/797	0.61	0/1069
3	I	0.29	0/785	0.62	0/1053
4	J	0.29	0/838	0.61	0/1139
4	K	0.29	0/838	0.62	0/1139
4	L	0.29	0/838	0.61	0/1139
5	M	0.27	0/1006	0.56	0/1365
5	N	0.27	0/1006	0.56	0/1365
5	O	0.27	0/1006	0.56	0/1365
6	P	0.38	0/1006	0.63	0/1372
6	R	0.38	0/1006	0.63	0/1372
6	T	0.38	0/1006	0.63	0/1372
7	Q	0.34	0/814	0.62	0/1109
7	S	0.34	0/814	0.62	0/1109
7	U	0.34	0/814	0.62	0/1109
All	All	0.47	18/25082 (0.1%)	0.78	75/34060 (0.2%)

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	4
1	B	0	5

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	#Chirality outliers	#Planarity outliers
1	C	0	4
All	All	0	13

The worst 5 of 18 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	608	VAL	CB-CG2	-15.19	1.21	1.52
1	A	608	VAL	CB-CG2	-15.03	1.21	1.52
1	C	608	VAL	CB-CG2	-14.94	1.21	1.52
1	B	657	GLU	CD-OE2	-8.64	1.16	1.25
1	B	658	GLN	CB-CG	-8.62	1.29	1.52

The worst 5 of 75 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	572	GLY	N-CA-C	12.55	144.47	113.10
1	B	572	GLY	N-CA-C	12.53	144.43	113.10
1	C	572	GLY	N-CA-C	12.53	144.43	113.10
1	C	628	TRP	CA-CB-CG	-9.97	94.75	113.70
1	A	628	TRP	CA-CB-CG	-9.90	94.89	113.70

There are no chirality outliers.

5 of 13 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	529	THR	Mainchain,Peptide
1	A	661	LEU	Mainchain,Peptide
1	B	529	THR	Mainchain,Peptide
1	B	661	LEU	Mainchain,Peptide
1	B	662	ALA	Peptide

## 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	983	0	961	66	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	B	975	0	950	91	0
1	C	975	0	950	75	0
2	D	2845	0	2786	163	0
2	E	2845	0	2786	122	0
2	F	2845	0	2785	124	0
3	G	786	0	804	23	0
3	H	786	0	804	16	0
3	I	775	0	795	18	0
4	J	819	0	782	33	0
4	K	819	0	782	34	0
4	L	819	0	782	33	0
5	M	985	0	952	19	0
5	N	985	0	952	19	0
5	O	985	0	952	20	0
6	P	980	0	935	56	0
6	R	980	0	935	55	0
6	T	980	0	935	173	0
7	Q	796	0	755	30	0
7	S	796	0	755	33	0
7	U	796	0	755	33	0
8	D	84	0	72	55	0
8	E	84	0	72	2	0
8	F	84	0	72	3	0
9	D	33	0	25	10	0
9	E	33	0	25	0	0
9	F	33	0	25	0	0
10	D	55	0	49	28	0
10	E	55	0	50	0	0
10	F	55	0	50	0	0
All	All	25071	0	24333	1031	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 21.

The worst 5 of 1031 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)
8:D:2341:NAG:H4	6:T:55:TRP:CG	1.27	1.68
10:D:2343:MAN:C4	6:T:70:SER:HB2	1.10	1.55
10:D:2343:MAN:C4	6:T:70:SER:CB	1.77	1.54
10:D:2343:MAN:C3	6:T:70:SER:HB3	1.38	1.51

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
9:D:2342:BMA:H2	6:T:55:TRP:CE3	1.54	1.41

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	120/153 (78%)	103 (86%)	17 (14%)	0	100	100
1	B	119/153 (78%)	102 (86%)	16 (13%)	1 (1%)	22	67
1	C	119/153 (78%)	99 (83%)	19 (16%)	1 (1%)	22	67
2	D	356/481 (74%)	322 (90%)	34 (10%)	0	100	100
2	E	356/481 (74%)	322 (90%)	34 (10%)	0	100	100
2	F	356/481 (74%)	323 (91%)	33 (9%)	0	100	100
3	G	96/192 (50%)	88 (92%)	8 (8%)	0	100	100
3	H	96/192 (50%)	88 (92%)	8 (8%)	0	100	100
3	I	95/192 (50%)	88 (93%)	7 (7%)	0	100	100
4	J	103/214 (48%)	97 (94%)	6 (6%)	0	100	100
4	K	103/214 (48%)	97 (94%)	6 (6%)	0	100	100
4	L	103/214 (48%)	97 (94%)	6 (6%)	0	100	100
5	M	125/229 (55%)	120 (96%)	5 (4%)	0	100	100
5	N	125/229 (55%)	120 (96%)	5 (4%)	0	100	100
5	O	125/229 (55%)	120 (96%)	5 (4%)	0	100	100
6	P	129/244 (53%)	120 (93%)	9 (7%)	0	100	100
6	R	129/244 (53%)	120 (93%)	9 (7%)	0	100	100
6	T	129/244 (53%)	120 (93%)	9 (7%)	0	100	100
7	Q	105/215 (49%)	99 (94%)	6 (6%)	0	100	100

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
7	S	105/215 (49%)	99 (94%)	6 (6%)	0	100	100
7	U	105/215 (49%)	99 (94%)	6 (6%)	0	100	100
All	All	3099/5184 (60%)	2843 (92%)	254 (8%)	2 (0%)	58	88

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	663	LEU
1	C	663	LEU

### 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	106/129 (82%)	106 (100%)	0	100	100
1	B	105/129 (81%)	105 (100%)	0	100	100
1	C	105/129 (81%)	105 (100%)	0	100	100
2	D	323/428 (76%)	323 (100%)	0	100	100
2	E	323/428 (76%)	323 (100%)	0	100	100
2	F	323/428 (76%)	323 (100%)	0	100	100
3	G	90/173 (52%)	90 (100%)	0	100	100
3	H	90/173 (52%)	90 (100%)	0	100	100
3	I	89/173 (51%)	89 (100%)	0	100	100
4	J	88/184 (48%)	88 (100%)	0	100	100
4	K	88/184 (48%)	88 (100%)	0	100	100
4	L	88/184 (48%)	88 (100%)	0	100	100
5	M	105/193 (54%)	105 (100%)	0	100	100
5	N	105/193 (54%)	105 (100%)	0	100	100
5	O	105/193 (54%)	105 (100%)	0	100	100
6	P	106/210 (50%)	106 (100%)	0	100	100

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
6	R	106/210 (50%)	106 (100%)	0	100	100
6	T	106/210 (50%)	106 (100%)	0	100	100
7	Q	82/182 (45%)	82 (100%)	0	100	100
7	S	82/182 (45%)	82 (100%)	0	100	100
7	U	82/182 (45%)	82 (100%)	0	100	100
All	All	2697/4497 (60%)	2697 (100%)	0	100	100

There are no protein residues with a non-rotameric sidechain to report.

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 45 such sidechains are listed below:

Mol	Chain	Res	Type
2	E	422	GLN
2	F	293	GLN
6	T	6	GLN
2	F	67	ASN
2	F	339	ASN

### 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 ⓘ

42 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. 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  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z  > 2$	Counts	RMSZ	$\# Z  > 2$
8	NAG	D	2340	8,2	14,14,15	0.40	0	15,19,21	0.76	1 (6%)
8	NAG	D	2341	9,8	14,14,15	0.46	0	15,19,21	0.54	0
9	BMA	D	2342	8,10	11,11,12	0.69	0	13,15,17	0.95	0
10	MAN	D	2343	9	11,11,12	1.24	2 (18%)	13,15,17	1.71	2 (15%)
8	NAG	D	2760	8,2	14,14,15	0.54	0	15,19,21	1.13	1 (6%)
8	NAG	D	2761	9,8	14,14,15	0.29	0	15,19,21	0.70	1 (6%)
9	BMA	D	2762	8,10	11,11,12	0.67	0	13,15,17	0.88	0
10	MAN	D	2763	9	11,11,12	0.79	0	13,15,17	0.98	1 (7%)
10	MAN	D	2764	9	11,11,12	0.75	0	13,15,17	1.08	2 (15%)
8	NAG	D	3860	8,2	14,14,15	0.58	0	15,19,21	1.10	1 (6%)
8	NAG	D	3861	9,8	14,14,15	0.95	1 (7%)	15,19,21	2.12	3 (20%)
9	BMA	D	3862	8,10	11,11,12	0.72	0	13,15,17	0.69	0
10	MAN	D	3863	9	11,11,12	0.77	0	13,15,17	1.17	2 (15%)
10	MAN	D	3864	9	11,11,12	0.75	0	13,15,17	1.01	2 (15%)
8	NAG	E	2340	8,2	14,14,15	0.40	0	15,19,21	0.76	1 (6%)
8	NAG	E	2341	9,8	14,14,15	0.48	0	15,19,21	0.54	0
9	BMA	E	2342	8,10	11,11,12	0.71	0	13,15,17	0.94	0
10	MAN	E	2343	9	11,11,12	1.24	2 (18%)	13,15,17	1.71	2 (15%)
8	NAG	E	2760	8,2	14,14,15	0.54	0	15,19,21	1.13	1 (6%)
8	NAG	E	2761	9,8	14,14,15	0.28	0	15,19,21	0.70	1 (6%)
9	BMA	E	2762	8,10	11,11,12	0.67	0	13,15,17	0.89	0
10	MAN	E	2763	9	11,11,12	0.78	0	13,15,17	0.97	1 (7%)
10	MAN	E	2764	9	11,11,12	0.74	0	13,15,17	1.08	2 (15%)
8	NAG	E	3860	8,2	14,14,15	0.58	0	15,19,21	1.10	1 (6%)
8	NAG	E	3861	9,8	14,14,15	0.96	1 (7%)	15,19,21	2.13	3 (20%)
9	BMA	E	3862	8,10	11,11,12	0.72	0	13,15,17	0.69	0
10	MAN	E	3863	9	11,11,12	0.77	0	13,15,17	1.17	2 (15%)
10	MAN	E	3864	9	11,11,12	0.74	0	13,15,17	1.01	2 (15%)
8	NAG	F	2340	8,2	14,14,15	0.45	0	15,19,21	0.84	1 (6%)
8	NAG	F	2341	9,8	14,14,15	0.42	0	15,19,21	0.52	0
9	BMA	F	2342	8,10	11,11,12	0.80	0	13,15,17	0.86	0
10	MAN	F	2343	9	11,11,12	1.41	2 (18%)	13,15,17	2.04	4 (30%)
8	NAG	F	2760	8,2	14,14,15	0.65	0	15,19,21	0.90	1 (6%)
8	NAG	F	2761	9,8	14,14,15	0.44	0	15,19,21	0.70	1 (6%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
9	BMA	F	2762	8,10	11,11,12	0.78	0	13,15,17	0.98	1 (7%)
10	MAN	F	2763	9	11,11,12	0.78	0	13,15,17	1.03	2 (15%)
10	MAN	F	2764	9	11,11,12	0.74	0	13,15,17	1.05	2 (15%)
8	NAG	F	3860	8,2	14,14,15	0.39	0	15,19,21	1.11	2 (13%)
8	NAG	F	3861	9,8	14,14,15	0.83	1 (7%)	15,19,21	2.09	3 (20%)
9	BMA	F	3862	8,10	11,11,12	0.73	0	13,15,17	0.79	0
10	MAN	F	3863	9	11,11,12	0.73	0	13,15,17	1.12	2 (15%)
10	MAN	F	3864	9	11,11,12	0.74	0	13,15,17	1.04	2 (15%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
8	NAG	D	2340	8,2	-	0/6/23/26	0/1/1/1
8	NAG	D	2341	9,8	-	0/6/23/26	0/1/1/1
9	BMA	D	2342	8,10	-	0/2/19/22	0/1/1/1
10	MAN	D	2343	9	-	0/2/19/22	0/1/1/1
8	NAG	D	2760	8,2	-	0/6/23/26	0/1/1/1
8	NAG	D	2761	9,8	-	0/6/23/26	0/1/1/1
9	BMA	D	2762	8,10	-	0/2/19/22	0/1/1/1
10	MAN	D	2763	9	-	0/2/19/22	0/1/1/1
10	MAN	D	2764	9	-	0/2/19/22	0/1/1/1
8	NAG	D	3860	8,2	-	0/6/23/26	0/1/1/1
8	NAG	D	3861	9,8	-	0/6/23/26	0/1/1/1
9	BMA	D	3862	8,10	-	0/2/19/22	0/1/1/1
10	MAN	D	3863	9	-	0/2/19/22	0/1/1/1
10	MAN	D	3864	9	-	0/2/19/22	0/1/1/1
8	NAG	E	2340	8,2	-	0/6/23/26	0/1/1/1
8	NAG	E	2341	9,8	-	0/6/23/26	0/1/1/1
9	BMA	E	2342	8,10	-	0/2/19/22	0/1/1/1
10	MAN	E	2343	9	-	0/2/19/22	0/1/1/1
8	NAG	E	2760	8,2	-	0/6/23/26	0/1/1/1
8	NAG	E	2761	9,8	-	0/6/23/26	0/1/1/1
9	BMA	E	2762	8,10	-	0/2/19/22	0/1/1/1
10	MAN	E	2763	9	-	0/2/19/22	0/1/1/1
10	MAN	E	2764	9	-	0/2/19/22	0/1/1/1
8	NAG	E	3860	8,2	-	0/6/23/26	0/1/1/1
8	NAG	E	3861	9,8	-	0/6/23/26	0/1/1/1
9	BMA	E	3862	8,10	-	0/2/19/22	0/1/1/1

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
10	MAN	E	3863	9	-	0/2/19/22	0/1/1/1
10	MAN	E	3864	9	-	0/2/19/22	0/1/1/1
8	NAG	F	2340	8,2	-	0/6/23/26	0/1/1/1
8	NAG	F	2341	9,8	-	0/6/23/26	0/1/1/1
9	BMA	F	2342	8,10	-	0/2/19/22	0/1/1/1
10	MAN	F	2343	9	-	0/2/19/22	0/1/1/1
8	NAG	F	2760	8,2	-	0/6/23/26	0/1/1/1
8	NAG	F	2761	9,8	-	0/6/23/26	0/1/1/1
9	BMA	F	2762	8,10	-	0/2/19/22	0/1/1/1
10	MAN	F	2763	9	-	0/2/19/22	0/1/1/1
10	MAN	F	2764	9	-	0/2/19/22	0/1/1/1
8	NAG	F	3860	8,2	-	0/6/23/26	0/1/1/1
8	NAG	F	3861	9,8	-	0/6/23/26	0/1/1/1
9	BMA	F	3862	8,10	-	0/2/19/22	0/1/1/1
10	MAN	F	3863	9	-	0/2/19/22	0/1/1/1
10	MAN	F	3864	9	-	0/2/19/22	0/1/1/1

The worst 5 of 9 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
10	F	2343	MAN	C2-C3	2.46	1.55	1.52
10	D	2343	MAN	C2-C3	2.47	1.55	1.52
10	E	2343	MAN	C2-C3	2.50	1.55	1.52
8	F	3861	NAG	C1-C2	2.62	1.56	1.52
10	E	2343	MAN	C1-C2	3.08	1.59	1.52

The worst 5 of 53 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
10	F	3864	MAN	O2-C2-C3	-2.32	105.63	110.17
10	D	2763	MAN	O2-C2-C3	-2.30	105.66	110.17
10	E	2763	MAN	O2-C2-C3	-2.29	105.68	110.17
10	F	2764	MAN	O2-C2-C3	-2.29	105.68	110.17
10	E	2764	MAN	O2-C2-C3	-2.28	105.69	110.17

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

9 monomers are involved in 98 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
8	D	2340	NAG	8	0
8	D	2341	NAG	42	0
9	D	2342	BMA	10	0
10	D	2343	MAN	28	0
8	D	2760	NAG	1	0
8	D	2761	NAG	4	0
8	E	2760	NAG	2	0
8	F	2760	NAG	2	0
8	F	2761	NAG	2	0

## 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.