



# wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 8, 2026 – 06:31 AM UTC

PDB ID : 4LIN / pdb\_00004lin  
Title : Exploring the atomic structure and conformational flexibility of a 320 angstrom long engineered viral fiber using X-ray crystallography  
Authors : Bhardwaj, A.; Cingolani, G.  
Deposited on : 2013-07-02  
Resolution : 2.70 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)

A user guide is available at

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

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

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

---

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

MolProbity : 4-5-2 with Phenix2.0  
Xtrriage (Phenix) : 2.0  
EDS : 3.0  
Percentile statistics : 20250101.v01 (using entries in the PDB archive January 1st 2025)  
CCP4 : 9.0.010 (Gargrove)  
Density-Fitness : 1.0.12  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.49

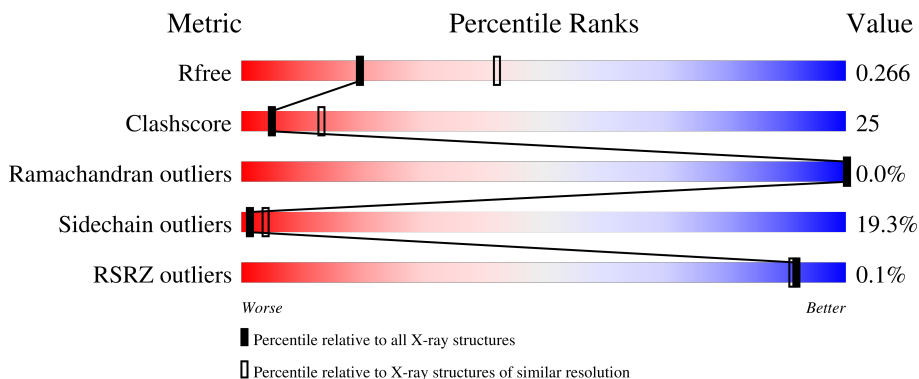
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.70 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	180053	3538 (2.70-2.70)
Clashscore	190562	3843 (2.70-2.70)
Ramachandran outliers	187476	3778 (2.70-2.70)
Sidechain outliers	187428	3778 (2.70-2.70)
RSRZ outliers	180081	3538 (2.70-2.70)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions  $\leq 5\%$ . The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	300	
1	B	300	
1	C	300	
1	D	300	
1	E	300	

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain		
1	F	300	47%	43%	6%
1	G	300	50%	39%	7%
1	H	300	45%	42%	9%
1	I	300	48%	40%	8%
1	J	300	49%	39%	8%
1	K	300	50%	38%	9%
1	L	300	49%	38%	9%

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
3	CL	A	1303	-	-	X	-
3	CL	H	1301	-	-	X	-
3	CL	I	1301	-	-	X	-
3	CL	L	1301	-	-	X	-

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 26552 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called Tail needle protein gp26.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	289	2148	1310	386	451	1	0	0	0
1	B	289	2148	1310	386	451	1	0	0	0
1	C	289	2148	1310	386	451	1	0	0	0
1	D	289	2148	1310	386	451	1	0	0	0
1	E	289	2148	1310	386	451	1	0	0	0
1	F	289	2148	1310	386	451	1	0	0	0
1	G	289	2148	1310	386	451	1	0	0	0
1	H	289	2148	1310	386	451	1	0	0	0
1	I	289	2148	1310	386	451	1	0	0	0
1	J	289	2148	1310	386	451	1	0	0	0
1	K	289	2148	1310	386	451	1	0	0	0
1	L	289	2148	1310	386	451	1	0	0	0

- Molecule 2 is CALCIUM ION (CCD ID: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	1	Total	Ca	0	0
			1	1		
2	E	1	Total	Ca	0	0
			1	1		

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	G	1	Total Ca 1 1	0	0
2	J	1	Total Ca 1 1	0	0

- Molecule 3 is CHLORIDE ION (CCD ID: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	2	Total Cl 2 2	0	0
3	D	1	Total Cl 1 1	0	0
3	E	1	Total Cl 1 1	0	0
3	H	1	Total Cl 1 1	0	0
3	I	1	Total Cl 1 1	0	0
3	K	1	Total Cl 1 1	0	0
3	L	1	Total Cl 1 1	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	70	Total O 70 70	0	0
4	B	81	Total O 81 81	0	0
4	C	54	Total O 54 54	0	0
4	D	68	Total O 68 68	0	0
4	E	61	Total O 61 61	0	0
4	F	70	Total O 70 70	0	0
4	G	58	Total O 58 58	0	0
4	H	54	Total O 54 54	0	0

*Continued on next page...*

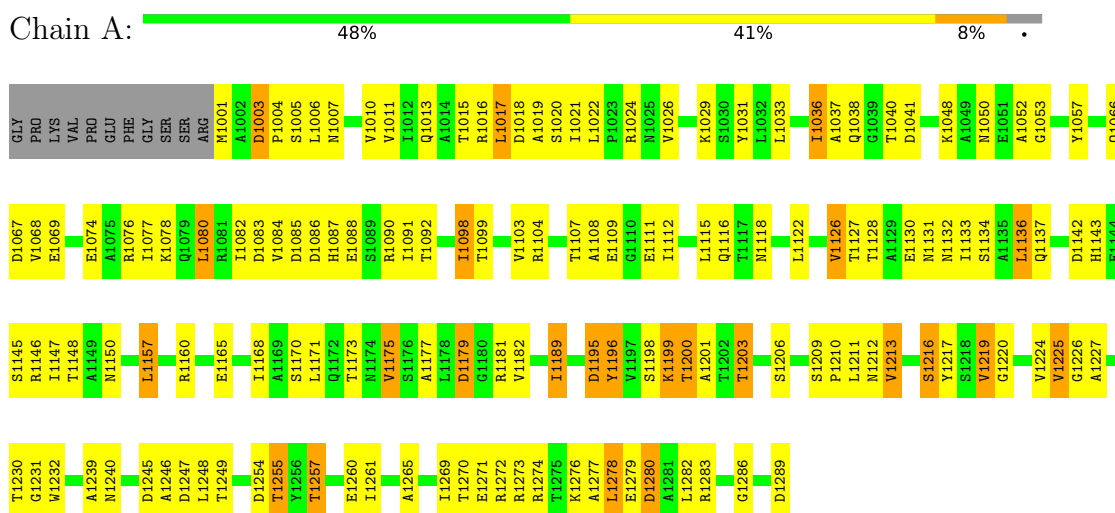
*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>	<b>ZeroOcc</b>	<b>AltConf</b>
4	I	61	Total O 61 61	0	0
4	J	67	Total O 67 67	0	0
4	K	53	Total O 53 53	0	0
4	L	67	Total O 67 67	0	0

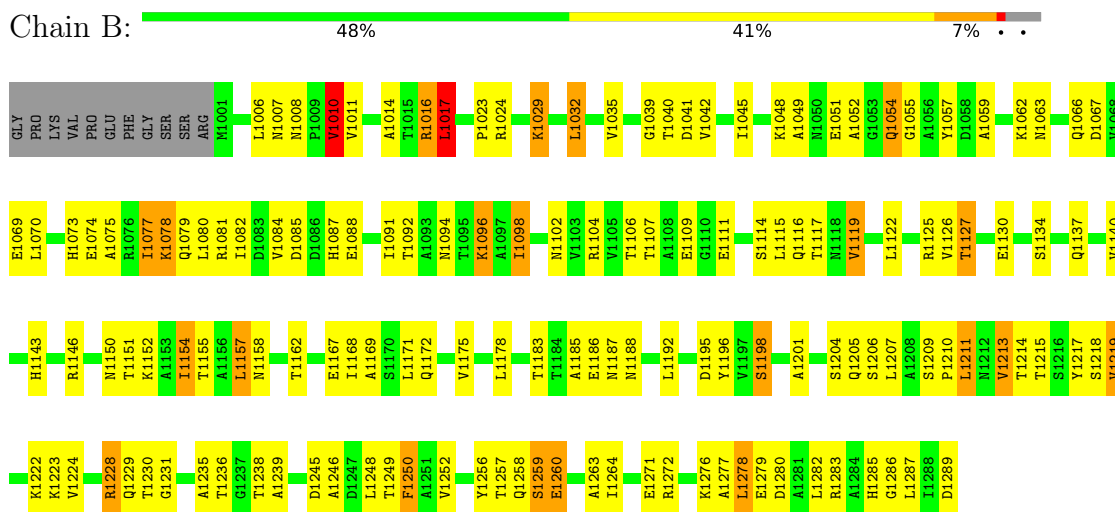
### 3 Residue-property plots

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. 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. A red dot above a residue indicates a poor fit to the electron density ( $RSRZ > 2$ ). 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: Tail needle protein gp26

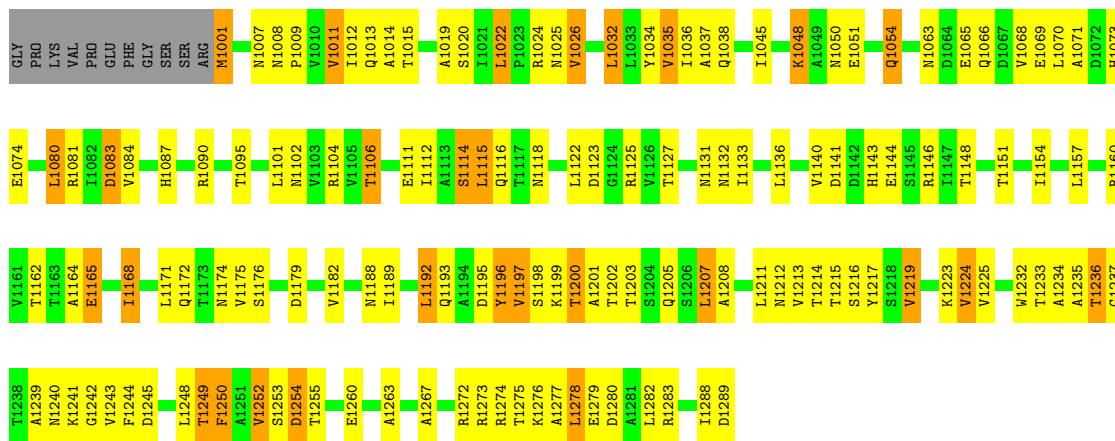


- Molecule 1: Tail needle protein gp26

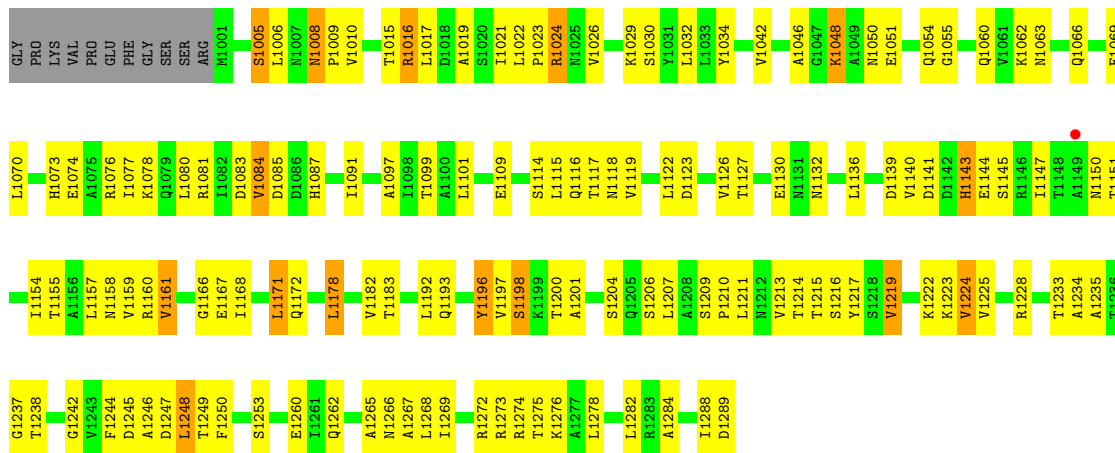


- Molecule 1: Tail needle protein gp26

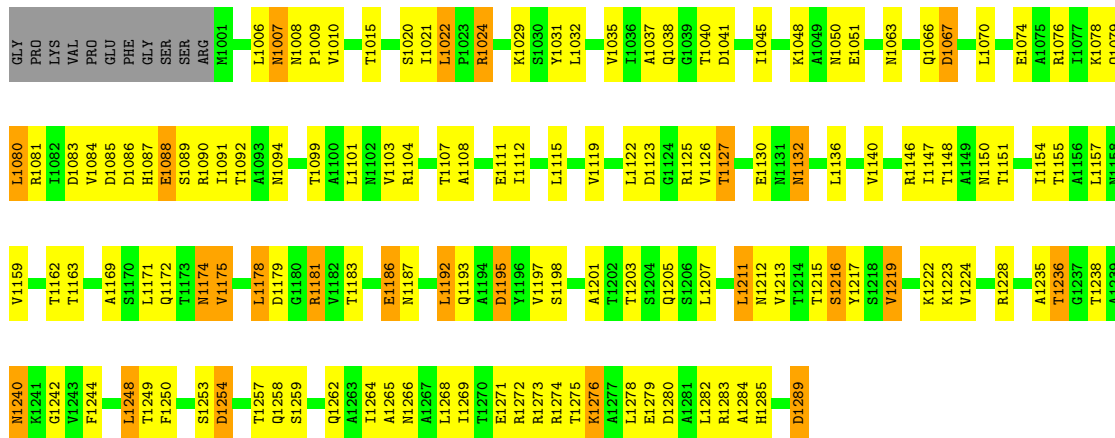




• Molecule 1: Tail needle protein gp26

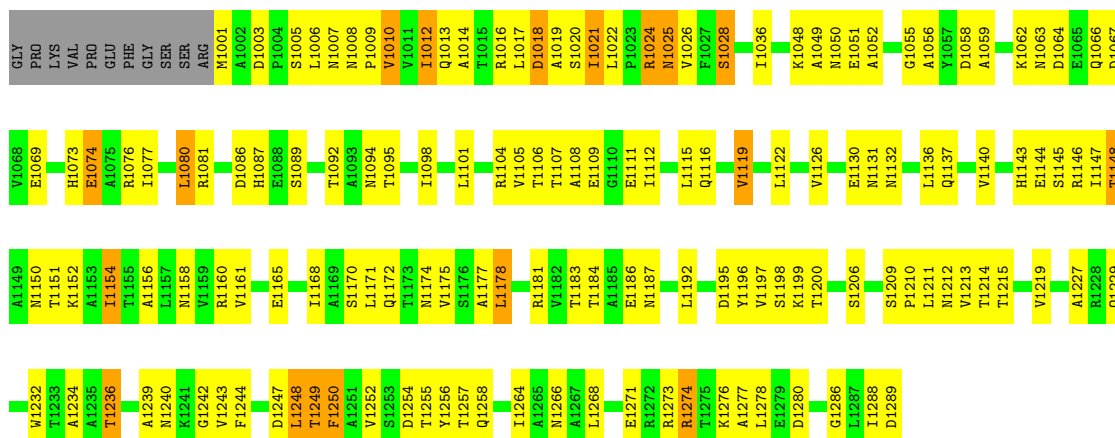


• Molecule 1: Tail needle protein gp26



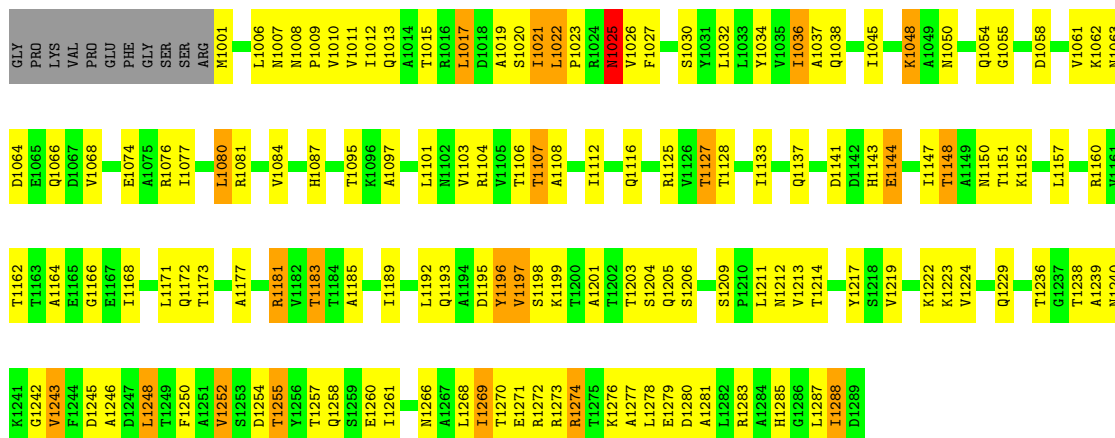
• Molecule 1: Tail needle protein gp26

Chain F:  47% 43% 6%



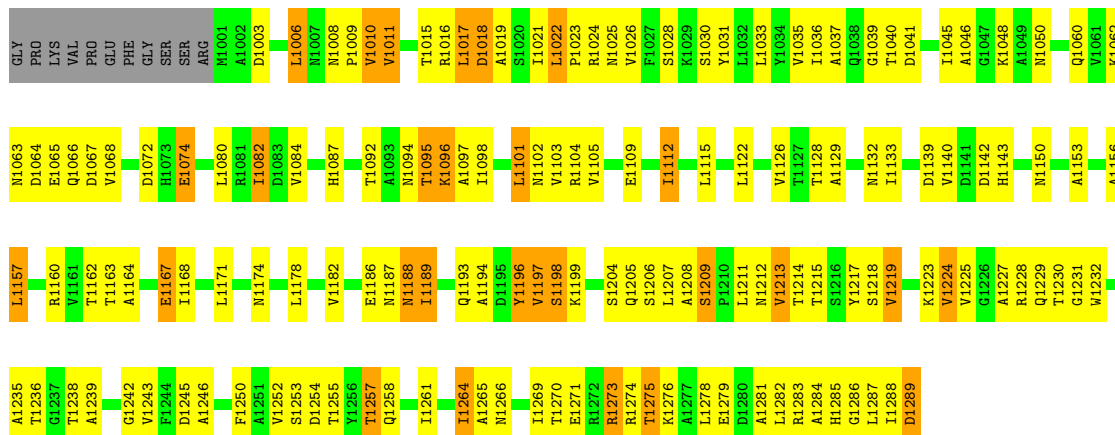
• Molecule 1: Tail needle protein gp26

Chain G:  50% 39% 7%



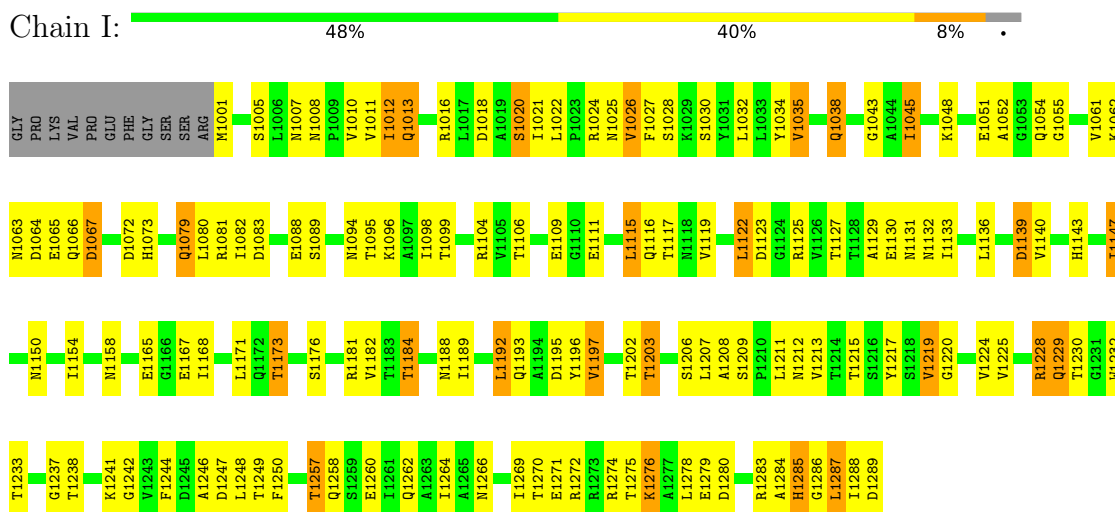
• Molecule 1: Tail needle protein gp26

Chain H:  45% 42% 9%



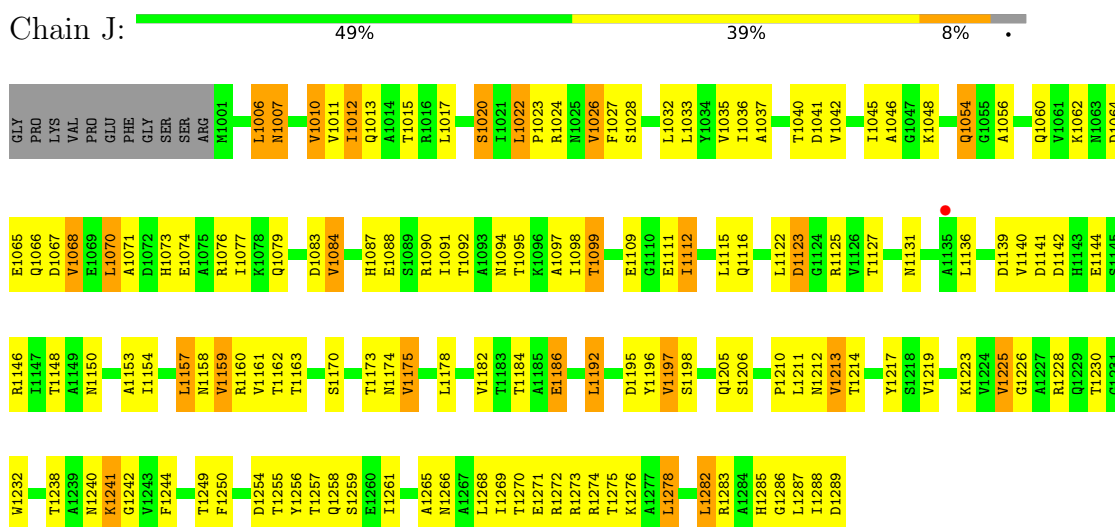
- Molecule 1: Tail needle protein gp26

Chain I:



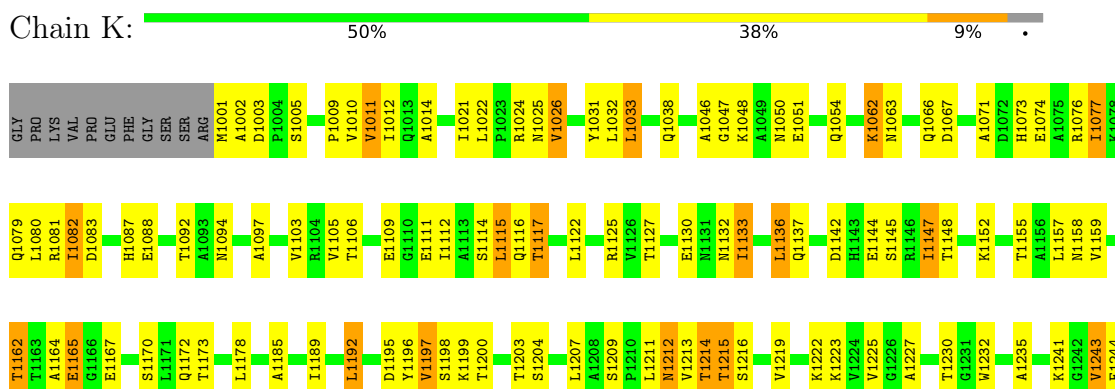
- Molecule 1: Tail needle protein gp26

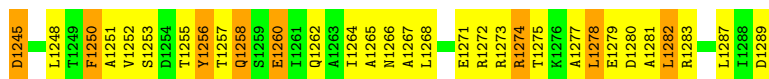
Chain J:



- Molecule 1: Tail needle protein gp26

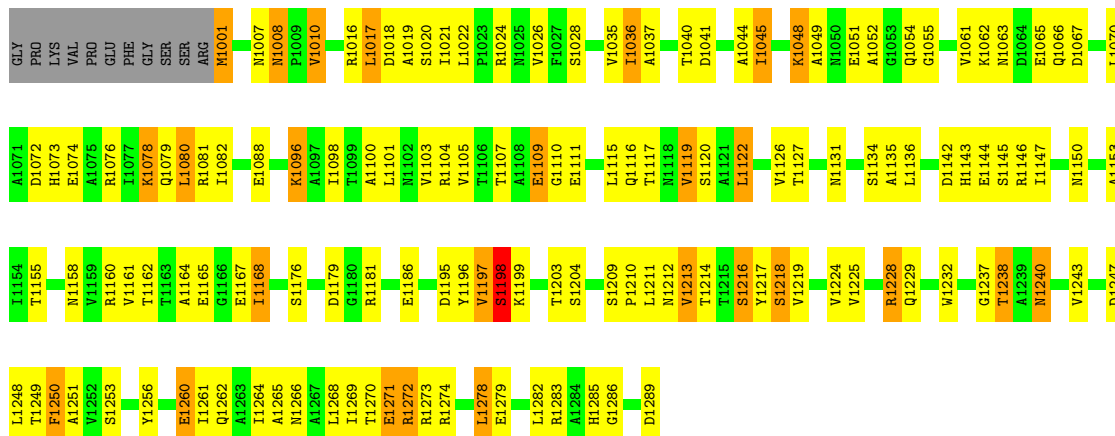
Chain K:





• Molecule 1: Tail needle protein gp26

Chain L: 49% 38% 9%



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	38.69Å 147.93Å 151.01Å 87.94° 90.05° 89.95°	Depositor
Resolution (Å)	14.99 – 2.70 14.99 – 2.70	Depositor EDS
% Data completeness (in resolution range)	91.5 (14.99-2.70) 73.2 (14.99-2.70)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.81 (at 2.68Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8.2_1309)	Depositor
R, $R_{free}$	0.229 , 0.268 0.228 , 0.266	Depositor DCC
$R_{free}$ test set	2018 reflections (2.40%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	32.9	Xtriage
Anisotropy	0.997	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.26 , 31.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	0.000 for h,-l,k 0.000 for h,l,-k 0.430 for h,-k,-l 0.000 for -h,k,-l 0.000 for -h,-k,l 0.000 for -h,-l,-k 0.000 for -h,l,k	Xtriage
Reported twinning fraction	0.480 for h,-k,-l	Depositor
Outliers	0 of 73933 reflections	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	26552	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	65.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 18.51% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

## 5 Model quality [i](#)

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

Bond lengths and bond angles in the following residue types are not validated in this section: CA, CL

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 5$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# Z  >5	RMSZ	# Z  >5
1	A	0.40	0/2166	0.83	3/2950 (0.1%)
1	B	0.39	0/2166	0.86	5/2950 (0.2%)
1	C	0.45	2/2166 (0.1%)	0.86	4/2950 (0.1%)
1	D	0.41	0/2166	0.82	2/2950 (0.1%)
1	E	0.40	0/2166	0.83	2/2950 (0.1%)
1	F	0.39	0/2166	0.87	3/2950 (0.1%)
1	G	0.41	0/2166	0.86	1/2950 (0.0%)
1	H	0.43	0/2166	0.86	0/2950
1	I	0.42	0/2166	0.86	3/2950 (0.1%)
1	J	0.40	0/2166	0.87	2/2950 (0.1%)
1	K	0.41	0/2166	0.89	4/2950 (0.1%)
1	L	0.43	0/2166	0.89	6/2950 (0.2%)
All	All	0.41	2/25992 (0.0%)	0.86	35/35400 (0.1%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	1143	HIS	CE1-NE2	-6.02	1.26	1.32
1	C	1143	HIS	CG-ND1	5.69	1.44	1.38

The worst 5 of 35 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	L	1261	ILE	N-CA-C	-10.52	102.42	112.83
1	K	1250	PHE	N-CA-C	7.82	119.44	111.07
1	L	1272	ARG	N-CA-C	-6.77	104.90	113.02
1	L	1008	ASN	CA-C-N	6.72	126.90	120.31
1	L	1008	ASN	C-N-CA	6.72	126.90	120.31

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	2148	0	2135	132	1
1	B	2148	0	2135	135	0
1	C	2148	0	2135	130	0
1	D	2148	0	2135	137	0
1	E	2148	0	2135	125	1
1	F	2148	0	2135	129	1
1	G	2148	0	2135	127	0
1	H	2148	0	2135	158	0
1	I	2148	0	2134	148	1
1	J	2148	0	2134	168	2
1	K	2148	0	2135	164	1
1	L	2148	0	2134	150	0
2	A	1	0	0	0	0
2	E	1	0	0	0	0
2	G	1	0	0	0	0
2	J	1	0	0	0	0
3	A	2	0	0	3	0
3	D	1	0	0	0	0
3	E	1	0	0	0	0
3	H	1	0	0	2	0
3	I	1	0	0	4	0
3	K	1	0	0	1	0
3	L	1	0	0	2	0
4	A	70	0	0	26	0
4	B	81	0	0	26	2
4	C	54	0	0	17	0
4	D	68	0	0	28	2
4	E	61	0	0	24	0
4	F	70	0	0	21	2
4	G	58	0	0	27	0
4	H	54	0	0	23	0
4	I	61	0	0	31	0
4	J	67	0	0	18	0
4	K	53	0	0	19	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	L	67	0	0	25	1
All	All	26552	0	25617	1280	7

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 25.

The worst 5 of 1280 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:1265:ALA:HB1	1:K:1250:PHE:CE2	1.41	1.49
1:H:1087:HIS:HD2	4:H:1454:HOH:O	1.08	1.35
1:K:1250:PHE:CE1	1:K:1264:ILE:HA	1.62	1.35
1:J:1217:TYR:CD2	1:J:1225:VAL:HB	1.72	1.23
1:K:1250:PHE:CD1	1:K:1264:ILE:HG12	1.73	1.22

The worst 5 of 7 symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:J:1244:PHE:O	4:F:1345:HOH:O[1_455]	1.98	0.22
1:J:1244:PHE:N	4:F:1345:HOH:O[1_455]	2.04	0.16
4:B:1346:HOH:O	4:D:1419:HOH:O[1_446]	2.06	0.14
1:F:1234:ALA:O	4:L:1408:HOH:O[1_655]	2.16	0.04
4:B:1338:HOH:O	4:D:1465:HOH:O[1_446]	2.18	0.02

## 5.3 Torsion angles [i](#)

### 5.3.1 Protein backbone [i](#)

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

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	287/300 (96%)	276 (96%)	11 (4%)	0	100	100
1	B	287/300 (96%)	279 (97%)	8 (3%)	0	100	100

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	C	287/300 (96%)	276 (96%)	11 (4%)	0	100	100
1	D	287/300 (96%)	281 (98%)	6 (2%)	0	100	100
1	E	287/300 (96%)	277 (96%)	10 (4%)	0	100	100
1	F	287/300 (96%)	274 (96%)	13 (4%)	0	100	100
1	G	287/300 (96%)	281 (98%)	6 (2%)	0	100	100
1	H	287/300 (96%)	278 (97%)	9 (3%)	0	100	100
1	I	287/300 (96%)	277 (96%)	10 (4%)	0	100	100
1	J	287/300 (96%)	275 (96%)	12 (4%)	0	100	100
1	K	287/300 (96%)	273 (95%)	14 (5%)	0	100	100
1	L	287/300 (96%)	270 (94%)	16 (6%)	1 (0%)	36	60
All	All	3444/3600 (96%)	3317 (96%)	126 (4%)	1 (0%)	100	100

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	L	1019	ALA

### 5.3.2 Protein sidechains [i](#)

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

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	229/238 (96%)	184 (80%)	45 (20%)	1	4
1	B	229/238 (96%)	183 (80%)	46 (20%)	1	4
1	C	229/238 (96%)	176 (77%)	53 (23%)	1	2
1	D	229/238 (96%)	199 (87%)	30 (13%)	4	10
1	E	229/238 (96%)	190 (83%)	39 (17%)	2	6
1	F	229/238 (96%)	194 (85%)	35 (15%)	3	7
1	G	229/238 (96%)	187 (82%)	42 (18%)	2	5
1	H	229/238 (96%)	178 (78%)	51 (22%)	1	3

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	I	229/238 (96%)	180 (79%)	49 (21%)	1	3
1	J	229/238 (96%)	186 (81%)	43 (19%)	1	4
1	K	229/238 (96%)	182 (80%)	47 (20%)	1	3
1	L	229/238 (96%)	180 (79%)	49 (21%)	1	3
All	All	2748/2856 (96%)	2219 (81%)	529 (19%)	1	4

5 of 529 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	K	1197	VAL
1	K	1260	GLU
1	K	1196	TYR
1	L	1240	ASN
1	E	1178	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 38 such sidechains are listed below:

Mol	Chain	Res	Type
1	K	1008	ASN
1	L	1063	ASN
1	K	1066	GLN
1	K	1212	ASN
1	L	1172	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 oligosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 12 ligands modelled in this entry, 12 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 5.7 Other polymers [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

## 6 Fit of model and data [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95<sup>th</sup> percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q< 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
1	A	289/300 (96%)	-0.96	0 100 100	29, 61, 89, 114	0
1	B	289/300 (96%)	-0.90	0 100 100	29, 59, 80, 96	0
1	C	289/300 (96%)	-0.82	0 100 100	26, 64, 92, 122	0
1	D	289/300 (96%)	-0.75	1 (0%) 90 89	26, 73, 101, 118	0
1	E	289/300 (96%)	-0.82	0 100 100	28, 64, 93, 115	0
1	F	289/300 (96%)	-0.85	0 100 100	28, 60, 98, 124	0
1	G	289/300 (96%)	-0.85	0 100 100	33, 67, 100, 141	0
1	H	289/300 (96%)	-0.77	0 100 100	32, 68, 99, 114	0
1	I	289/300 (96%)	-0.84	0 100 100	33, 65, 97, 125	0
1	J	289/300 (96%)	-0.83	1 (0%) 90 89	36, 67, 97, 116	0
1	K	289/300 (96%)	-0.85	0 100 100	37, 66, 102, 123	0
1	L	289/300 (96%)	-0.82	0 100 100	39, 65, 99, 119	0
All	All	3468/3600 (96%)	-0.84	2 (0%) 92 91	26, 65, 97, 141	0

All (2) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	1149	ALA	3.0
1	J	1135	ALA	2.1

### 6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

### 6.3 Carbohydrates [i](#)

There are no oligosaccharides in this entry.

### 6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
3	CL	I	1301	1/1	0.97	0.12	171,171,171,171	0
3	CL	A	1303	1/1	0.98	0.04	60,60,60,60	0
3	CL	D	1301	1/1	0.98	0.03	58,58,58,58	0
3	CL	H	1301	1/1	0.98	0.07	38,38,38,38	0
3	CL	A	1302	1/1	0.98	0.05	50,50,50,50	0
2	CA	G	1301	1/1	0.99	0.02	36,36,36,36	0
3	CL	E	1302	1/1	0.99	0.03	61,61,61,61	0
2	CA	A	1301	1/1	0.99	0.03	79,79,79,79	0
2	CA	E	1301	1/1	0.99	0.04	54,54,54,54	0
3	CL	K	1301	1/1	0.99	0.04	52,52,52,52	0
3	CL	L	1301	1/1	0.99	0.05	49,49,49,49	0
2	CA	J	1301	1/1	1.00	0.06	99,99,99,99	0

### 6.5 Other polymers [i](#)

There are no such residues in this entry.